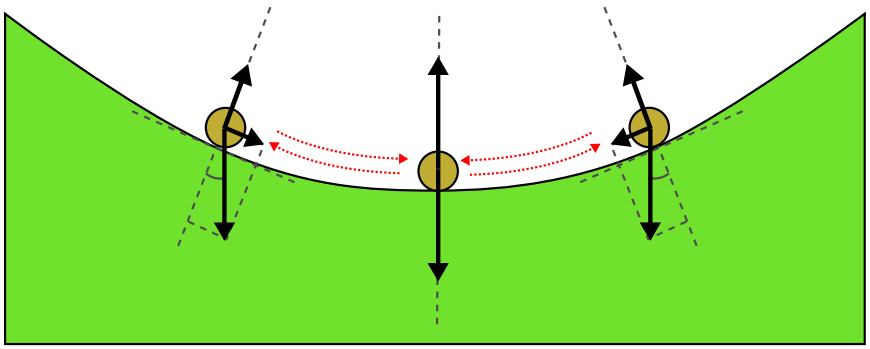
Slightly Disturbed

A Mathematical Approach to Oscillations and Waves

Brooks Thomas

Lafayette College

Second Edition



2017

Contents

Chapter 1

Simple Harmonic Motion

- **The Physics**: Stable and unstable equilibrium, springs and simple harmonic motion, state space, LC circuits, the energetics of oscillation
- **The Math**: Differential equations, initial conditions

.

1.1 Equilibrium, Restoring Forces, and Periodic Motion

Periodic motion — motion that repeats itself in finite time — is ubiquitous in nature. Objects shake back and forth when we bump into them; automobile engines, cell phones, drum heads, hummingbirds' wings, and audio speakers vibrate; rocking chairs rock; pendulums swing back and forth; and coffee sloshes back and forth in a cup after that cup is set down on a table. The electrical current used to power anything you might plug into a wall oscillates back and forth in a sinusoidal pattern 60 times each second. Nearly every medium of communication over distance relies on some sort of wave to transmit information, from direct speech (sound waves) to radio, television, and cell phones (electromagnetic waves).

Why is periodic motion such a common phenomenon? The reason is that oscillations are what generically happens when a system in a stable equilibrium state gets disturbed a little bit. In order to clarify precisely what this statement means, however, we're going to have to go into a little bit more detail about what we mean by "stable," "equilibrium," and "disturbed."

First, let's review what we mean by equilibrium. A rigid body is said to be in **mechanical equilibrium** if it is not accelerating. The acceleration of such a body is related to the force acting on it by Newton's Second Law. For example, in a one-dimensional system, this relationship takes the form

$$F \;=\; ma \;=\; m\frac{d^2x}{dt^2}\;, \tag{1.1}$$

where x is the position of the body, m is its mass, and t is time. A body in mechanical equilibrium is therefore one on which the net force is zero.[1]

Two examples of systems in mechanical equilibrium are illustrated in Fig. 1.1. The diagram on the left shows a ball at rest at the bottom of a valley, while the diagram on the right shows a similar ball at rest at the top of a hill. In both of these situations, the gravitational force $F_g = -mg$ and the normal force $F_N = mg$ are equal and opposite at the point where the ball is located, so the net force acting on the ball is zero.

Now let's review what we mean by "disturbance" and "stable." Nature is full of effects that disturb physical systems away from equilibrium. For example, in each of the situations illustrated in Fig. 1.1, a gust of wind might blow, an insect might land on the ball, a leaf might fall on it, raindrops might strike it, the ground might shift a little bit, someone or something might jostle it accidentally, and so on. However, the

[1] In the generalization of this principle to motion in more than one dimension, in which case rotation is possible, the net torque on the body must also vanish.

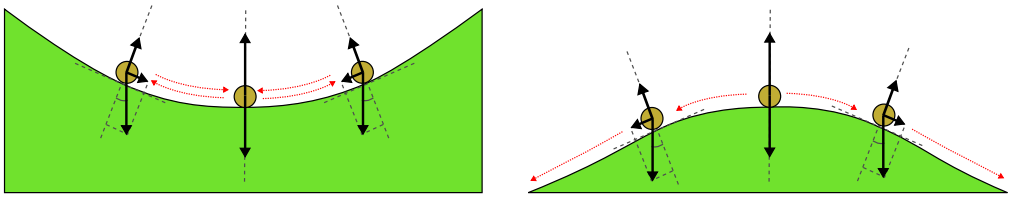


Figure 1.1: Examples of stable and unstable equilibrium. The diagram on the left shows a stable equilibrium state. In this case, the net force which acts on the ball when it moves away from its equilibrium position acts to drive it back toward that equilibrium position. The diagram on the right shows an unstable equilibrium state. In this case, the net force which acts on the ball when it moves away from its equilibrium position serves to drive it farther away from that equilibrium position.

two systems in this figure respond to these disturbances in very different ways. In the diagram on the left, the net force which acts on the ball when it is displaced slightly from its equilibrium position always serves to accelerate it back toward that equilibrium position. A force that serves to "correct" for any departure from equilibrium in this way is called a **restoring force**, an equilibrium state which is robust against small disturbances because of such "corrections" is called **stable**.[2] By contrast, in the diagram on the right, the force which acts on the ball when it is displaced from its equilibrium position accelerates it in a direction further away from that position. This precarious situation is an example of an system with an **unstable** equilibrium state.

While a restoring force in a system with a stable equilibrium state acts drive the system back toward that equilibrium state, that doesn't mean that it causes the system to return promptly to that equilibrium state and stop. For example, when the ball in the left diagram of Fig. 1.1 rolls back toward its equilibrium position at the bottom of the valley under the influence of the restoring force, it acquires momentum in the process. As a result, the ball will "overshoot" and roll right past the equilibrium point a the bottom of the valley because of **inertia** — *i.e.*, the tendency of an object to oppose change in its motion. It will continue rolling up the other side of the valley until the restoring force overcomes that inertial tendency and drives it back down toward the equilibrium position again, and the process repeats itself. This interplay between the action of the restoring force to drive the system toward equilibrium and and the inertial tendency of the system to remain in motion is what gives rise to oscillations. Indeed, oscillations can arise in just about any physical system in which a stable equilibrium state exists and in which that equilibrium is disturbed, provided that the disturbance is sufficiently small.[3]

That said, oscillations usually don't last forever. Nature is full of dissipative forces (*e.g.*, friction and air resistance) by which oscillating systems lose energy to their surroundings and eventually settles back into its equilibrium state. Over the course of the semester, we'll examine these effects too. For the moment, however, we're going to focus on what is probably the simplest and arguably the most important physical system which exhibits periodic motion: the simple harmonic oscillator.

1.2 Simple Harmonic Oscillator

One particularly simple and probably familiar example of a restoring force is the force provided by a spring on a mass attached to that spring. For an idealized spring, this force is given by Hooke's Law:[4]

$$F = -kx \,, \tag{1.2}$$

[2] As we shall make clear later, this heuristic definition indeed accords with the mathematical definition of a stable equilibrium point as an point at which the first derivative of the potential energy function $U(x)$ with respect to x vanishes and the second derivative is positive.

[3] What "small" means depends on the situation. However, for a system is in stable equilibrium, there is always some regime in which perturbations away from equilibrium are small.

[4] The restoring force provided by a real spring deviates from this ideal, but Hooke's law is a very good approximation in a wide variety of cases.

where k is a constant parameter (commonly called the "spring constant") with units[5]

$$[k] \;=\; \frac{\text{Newton}}{\text{Meter}} \;. \tag{1.3}$$

The negative sign in Eq. (1.2) is really what makes this a restoring force. I the mass is displaced in the positive direction then the force acts in the *negative* direction, accelerating it back towards its equilibrium position. Conversely, if the mass is displaced in the negative direction, the force acts in the *positive* direction, one again accelerating it back towards its equilibrium position.

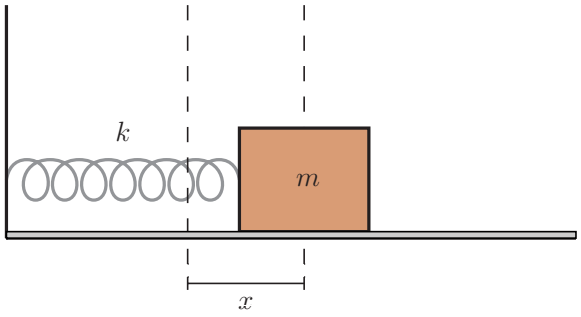


Figure 1.2: A mass m attached to a spring with spring constant k.

When we use Newton's Second Law to relate the force given by Hooke's law to the acceleration of the mass m attached to the spring, we get a differential equation which relates the second derivative of x (with respect to time t) to x itself:

$$m\frac{d^2x}{dt^2} \;=\; -\,kx \;. \tag{1.4}$$

This is the **equation of motion** for the mass m attached to the spring.

$$\frac{d^2x}{dt^2} \;=\; -\,\frac{k}{m}x \;, \tag{1.5}$$

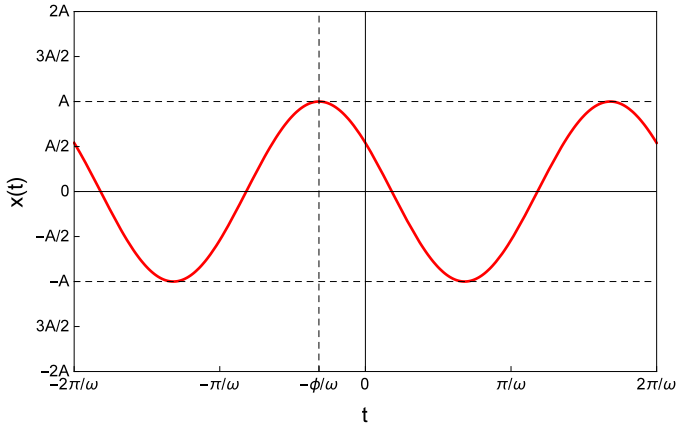


Figure 1.3: Illustration of the solution $x(t)$ to the equation of motion for a simple harmonic oscillator. The results shown here correspond to one particular choice of the parameters A, ϕ, and ω.

What do the solutions to this equation of motion — *i.e.*, the set of possible functions $x(t)$ which describe how the position of the mass attached to the spring evolves in time — look like? Well, Eq. (1.5) stipulates

[5]We will frequently use the notation $[X]$ throughout these lecture notes to indicate the units of the quantity X.

that each solution must be a function whose second derivative is equal to the function itself times a negative constant. You can easily verify for yourself by plugging that any function of the form

$$x(t) \;=\; A\cos(\omega t + \phi) \tag{1.6}$$

satisfies Eq. (1.5), where we have defined

$$\omega \;\equiv\; \sqrt{\frac{k}{m}}\;. \tag{1.7}$$

The parameters A and ϕ in this equation are arbitrary constants in the sense that $x(t)$ will still satisfy Eq. (1.5) no matter what values we happen to assign these parameters. In fact, this solution turns out to be the most general solution to the equation of motion for a simple harmonic oscillator. As we shall see later in this course, the fact that there are two free parameters A and ϕ is intimately related to the fact that the equation of motion involves a second derivative. An solution of this form is illustrated in Fig. 1.3.

In addition to having a function $x(t)$ which describes the position of the mass in our simple harmonic oscillator as a function of time, it is also useful to have a function $v(t)$ which describes its velocity. This is easily done: since the velocity is simply the instantaneous derivative of the position function $x(t)$ with respect to t, we find that

$$v(t) \;=\; \frac{dx}{dt} \;=\; -\,A\omega\sin(\omega t + \phi)\;. \tag{1.8}$$

Perhaps the most important aspect of the solution for $x(t)$ in Eq. (1.6), however, is that this solution is periodic. Indeed, since the cosine function has the property that $\cos(\theta + 2\pi) = \cos(\theta)$ for any real angle θ, we see that the mass will return to exactly the same position in which it started in finite time. Since this is the case, it might be a good idea to step back and review some of the quantities which are useful in describing periodic motion.

- The constant A represents the **amplitude** of the oscillation: the greatest displacement away from equilibrium that the mass m ever experiences during is oscillation.

- The **period** of oscillation T is defined as the time it takes for the system to go through one complete cycle of motion and return to the same state — *i.e.*, to the same position x and the same velocity v it started with. Since the cosine function has the property that $\cos(\theta + 2\pi) = \cos(\theta)$ for any real angle θ, we find that that

$$A\cos[\omega(t+T)+\phi] \;=\; A\cos(\omega t + \phi + 2\pi) \rightarrow \omega T = 2\pi\;. \tag{1.9}$$

  Thus, the period is just

$$T \;=\; \frac{2\pi}{\omega}\;. \tag{1.10}$$

  It's important to remember that the system being in the same *state* means more than merely returning to the same *position*. For example, consider that the mass on the spring in the simple harmonic oscillator shown in Fig. 1.2 passes through the point $x = 0$ twice during each cycle of oscillation. However, in one of these cases the mass is moving to the left, whereas in the other it's moving to the right. This idea of a "state" of the system characterized by the values of $x(t)$ and $v(t)$ together is actually an important one for other reasons as well, and we'll have more to say about it in Sect. 1.6.

- The **frequency** of oscillation f is number of complete cycles the system makes per unit time. In other words, it's just the inverse of the period:

$$f \;=\; \frac{1}{T} \;=\; \frac{\omega}{2\pi}\;. \tag{1.11}$$

  The frequency has units of inverse seconds or cycles per second. The SI unit of frequency is the hertz, which is equal to one cycle per second.

- The **angular frequency** is the number of radians that the system passes through per unit time. Since a full cycle is just 2π radians, the angular frequency is just $2\pi f$, or ω itself. Take care not to confuse f and ω. Even though it is common practice to refer to ω as the "oscillator frequency" or even just the "frequency" (and I confess that I will do this a great deal in these notes), these quantities represent different things.

- Finally, the constant ϕ in Eq. (1.5) is known as the **phase** of the oscillation. The phase encodes information about the times at which the system reaches its peak displacement $x(t) = A$. In particular since the cosine function reaches a maximum at $\cos(0) = 1$, the peak displacement occurs when $\omega t_{\text{peak}} + \phi = 0$, or

$$t_{\text{peak}} = -\phi/\omega. \tag{1.12}$$

  Note that t_{peak} doesn't depend on what the peak displacement A actually is, but only on ϕ and ω. Note also that since $\cos(\omega t + \phi + 2\pi) = \cos(\omega t + \phi)$, an oscillation with a phase ϕ and an oscillation with a phase $\phi' = \phi + 2\pi$ represent exactly the physics. In other words, ϕ can always be restricted to the range $0 \leq \phi < 2\pi$ without any loss of generality.

Finally, it's important to emphasize that while Eq. (1.6) is indeed the most general solution to the equation of motion for a simple harmonic oscillator, there are a lot of other ways of parameterizing this solution. Many of these look like completely independent solutions, but in fact are just clever ways of rewriting the same thing. For example, any function of the form

$$x(t) = B\sin(\omega t + \delta)\ , \tag{1.13}$$

where B and δ are arbitrary constants, also satisfies Eq. (1.5). However, this doesn't actually turn out to be a distinct solution from the solution in Eq. (1.6). The sine and cosine functions are related by $\sin(\theta) = \cos(\theta + \pi/2)$ for any angle θ, so any sine function of the form in Eq. (1.13) is actually completely equivalent to a cosine function of the form in Eq. (1.6) with $A = B$ and $\phi = \delta + \pi/2$.

There are plenty of other ways of parameterizing the general solution to Eq. (1.5) as well (see, for example, Problem 2), but in fact these solutions are all equivalent.

1.3 Initial Conditions

Up to this point, we still have not said anything about how to determine the values for the amplitude A and phase ϕ appearing in Eq. (1.6). Indeed, the equation of motion gives us no guidance here, since our general solution for $x(t)$ satisfies that equation no matter what values of A and ϕ we choose. In order to determine what the values of A and phase ϕ are, we need some additional information about the physical system we're studying. This information comes from the **initial conditions** which characterize the system at the moment when the mass m attached to the spring begins to oscillate. It is convenient to call that time $t = 0$, so that we "start the clock" when the mass starts oscillating.

Let's say that we release the mass from the position x_0 at $t = 0$. Perhaps we also give the block a push in one direction or the other as we release it so that it starts with an initial velocity v_0. Substituting our initial conditions $x(0) = x_0$ and $\dot{x}(0) = v_0$ into Eqs. (1.6) and (1.8), we find that

$$x_0 = A\cos(\phi) \tag{1.14}$$
$$v_0 = -A\omega\sin(\phi) \tag{1.15}$$

If we solve this system of equation equations for A and ϕ, we find that the amplitude of oscillation for an initial displacement x_0 and an initial velocity v_0 is

$$A = \sqrt{x_0^2 + \frac{v_0^2}{\omega^2}}\ . \tag{1.16}$$

Similarly, we also find that the phase of the oscillation is

$$\phi = -\arctan\left(\frac{v_0}{\omega x_0}\right)\ . \tag{1.17}$$

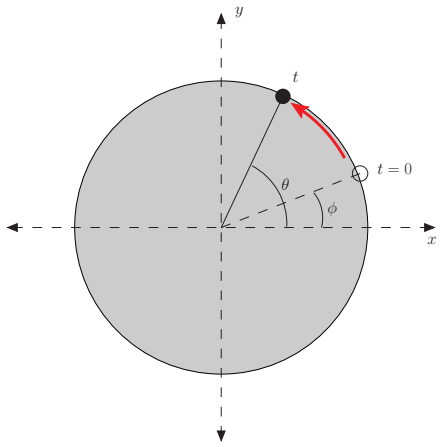


Figure 1.4: The motion of a dot marked on the rim of a flat disc with radius A which rotates counterclockwise with constant angular velocity ω. At time $t = 0$, the dot lies at an angle ϕ from the x axis.

1.4 Relation to Cirular Motion

There is a mathematical parallel between uniform circular motion and simple harmonic motion that is worth mentioning at this point. To illustrate this, let's consider the motion of a small dot drawn on the rim of a flat disc with radius A which rotates counterclockwise with constant angular velocity ω, as shown in Fig. 1.4. Moreover, let's say the dot is at an at the angle ϕ at the time $t = 0$ when the disc begins rotating, so that the angle that it makes with the x-axis as a function of time is

$$\theta \;=\; \omega t + \phi \;. \tag{1.18}$$

Let us ignore the motion in the y direction and focus only on the x coordinate. This coordinate evolves with time according to the relation

$$x(t) \;=\; A\cos(\omega t + \phi) \;. \tag{1.19}$$

This is completely identical to the formula in Eq. (1.6) for the position of the mass attached to the spring in a simple harmonic oscillator. We therefore see that if we set an object in uniform circular motion and then look at it edge-on so that we only observe its motion along one axis, this motion will be identical to a mass attached to a spring.

This correspondence provides a nice way to think about many of the features of our solution for the simple harmonic oscillator. While it's important to keep in mind that it is only a mathematical analogy, it it often easier to think of the angular frequency ω of a simple harmonic oscillator as the angular speed (*i.e.*, the magnitude of the angular velocity) with which the corresponding circle rotates. Likewise, it's also often easier to think of the phase ϕ as nothing more than the angle at which the object begins its motion on that circle at $t = 0$. For these two insights alone, the correspondence is with mentioning. The correspondence will also become even more useful later on, when we begin using complex numbers to describe oscillations.

1.5 Simple Harmonic Oscillators in Disguise

A mass oscillating on a spring in the absence of friction is probably the most familiar example of simple harmonic motion. However, it turns out that a lot of other physical systems are governed by an equation of motion that has exactly the same mathematical form as Eq. (1.5).

One example of a system which exhibits simple harmonic motion is the charge on a parallel-plate capacitor in an LC circuit — *i.e.*, a simple circuit which consists of an inductor and a parallel-plate capacitor, as shown in Fig. 1.5. The capacitor is a device which stores equal and opposite charges on its two plates. The capacitor equation

$$Q \;=\; VC \tag{1.20}$$

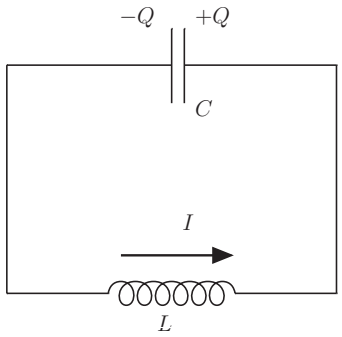


Figure 1.5: Diagram of a simple LC circuit. The direction of the current through the charges on the capacitor plates are indicated.

relates the potential difference V between the plates to the charge Q stored on the positive plate. The constant of proportionality C is called the **capacitance**, which depends on the size and geometry of the plates, as well as the properties of whatever material lies between them. The inductor consists of a coil of wire. When current flows through the wire, it produces a **magnetic flux** through the wire. Whenever the current I flowing through the wire changes, the magnetic flux through the coil changes as well. This generates an EMF[6] in the wire

$$\mathcal{E} \;=\; -\,L\frac{dI}{dt} \tag{1.21}$$

where L is the **inductance** L of the wire — a parameter that depends on the geometry of the coil, the number of turns of wire in it, *etc.*. The minus sign in Eq. (1.21) indicates that the EMF acts in the direction which opposes the change in the current.

How does charge flow in an LC circuit? Kirchhoff's second law (or loop rule) tells us that the sum of the potential drop across the inductor and potential drop across the capacitor must sum to zero:

$$L\frac{dI}{dt} + \frac{Q}{C} \;=\; 0\;. \tag{1.22}$$

Conservation of electric charge in the circuit tells us that if a current is flowing through the circuit, there must be a corresponding change in the charge Q on the capacitor:

$$I \;=\; \frac{dQ}{dt}\;. \tag{1.23}$$

Substituting this into Eq. (1.22) and rearranging things a bit, we arrive at the following differential equation for the charge Q stored on the capacitor plates:

$$\frac{d^2Q}{dt^2} \;=\; -\,\frac{1}{LC}Q\;. \tag{1.24}$$

This equation for $Q(t)$ has precisely the same mathematical form as the differential equation Eq. (1.5) for the position $x(t)$ of a mass attached to a spring! It therefore follows that $Q(t)$ likewise oscillates in time and is given by

$$Q(t) \;=\; A\cos(\omega t + \phi)\;. \tag{1.25}$$

with an oscillator frequency

$$\omega \;=\; \sqrt{\frac{1}{LC}}\;. \tag{1.26}$$

[6]This footnote contains the obligatory disclaimer that the term EMF or "electromotive force," is a horrible misnomer. As you have probably already heard countless times in other physics classes you've taken, an EMF is not a force at all and has the units of electric potential.

We see that in this circuit, the inductance L plays the role of the mass m in Eq. (1.5) and $1/C$ plays the role of the spring constant k. These analogies actually make good physical sense. The inductance L quantifies the "inertial" tendency of the inducting coil to resist a change in the current, just as the mass m of an object quantifies its tendency the to resist a change in its state of motion. Likewise, the electrostatic force exerted by the equal and opposite charges $+Q$ and $-Q$ on the capacitor plates provides the restoring force which always serves to drive the system toward the equilibrium state where $Q = 0$. This is completely analogous to the function that the spring performs in Fig. 1.2. However, Eq. (1.20) tell us us that the greater C is, the smaller the potential difference is between the plates for a given value of Q, and hence the smaller the restoring force. This is why it's the *inverse* of C in an LC circuit that plays an analogous role to the spring constant k.

There are other physical systems which are also really simple harmonic oscillators in disguise — though admittedly examples are scarce and usually fairly contrived. One example (see Problem 6) involves a buoy bobbing up and down in down in a body of water. Nevertheless, the simple harmonic oscillator turns out to be one of the most important physical systems in all of physics.

The reason is that while systems which are mathematically precisely identical to a simple harmonic oscillator are few and far between, systems which are *approximately* equivalent to a simple harmonic oscillator arise almost everywhere in nature. In fact, as we shall see in the next lecture, a physical system with a stable equilibrium solution almost always behaves like a simple harmonic oscillator when the departure from equilibrium is sufficiently small.

1.6 State Space

In Sect. 1.3, we saw that the behavior of a simple harmonic oscillator depends not only on the its initial position x_0, but also on its initial velocity v_0. In fact, if you know the position x and velocity v at any time t, you have sufficient information to be able to trace the future trajectory of the oscillator.

The concept of **state space** provides a useful way of visualizing how a physical system evolves from an arbitrary initial state. A state-space diagram is a plot in which the axes represent not x and t, but rather x and $\dot{x} = v$. Each point in state space represents a distinct state of the system — *i.e.*, a particular combination of x and v. As the system evolves in time and the values of x and v change, the system traces out a continuous curve in state space.

Some examples of such curves are shown in Fig. 1.6. The curves in the left panel of the figure correspond to an oscillator with a frequency $\omega = 0.5$; the ones in the right panel correspond to an oscillator with a frequency $\omega = 2$. The different curves in each panel differ only in terms of the initial conditions — *i.e.*, the values of x_0 and v_0. These values, which indicate the "starting point" for the system at $t = 0$, are indicated with a dot on each curve. The fact that the state-space trajectories are closed indicates that the motion is periodic. In particular, it reflects the fact that the system returns to the same state — the same combination of x and v — in finite time. It is not difficult to show that these trajectories are, in fact, ellipses.

$$\left(\frac{x}{A}\right)^2 + \left(\frac{v}{A\omega}\right)^2 \;=\; \cos^2(\omega t + \phi) + \sin^2(\omega t + \phi) \;=\; 1 \; . \tag{1.27}$$

This is the defining equation for an ellipse whose semimajor and semiminor axes have lengths A and $A\omega$ (not necessarily respectively: which axis is which depends on the value of ω).

While the direction in which the system evolves along that curve as t increases is not indicated in the figure, we can infer that direction from the fact that $v(t)$ is just the time derivative of $x(t)$. If the velocity $v(t)$ is positive, that means $x(t)$ is increasing, and the the location of the system in state space is evolving toward the right. Likewise, if v is negative, x is decreasing, and the location of the system is evolving toward the left. It therefore follows that the system always evolves clockwise in state space along its state-space trajectory. This isn't unique to the simple harmonic oscillator either, but is rather a general property of the way systems evolve in state space.

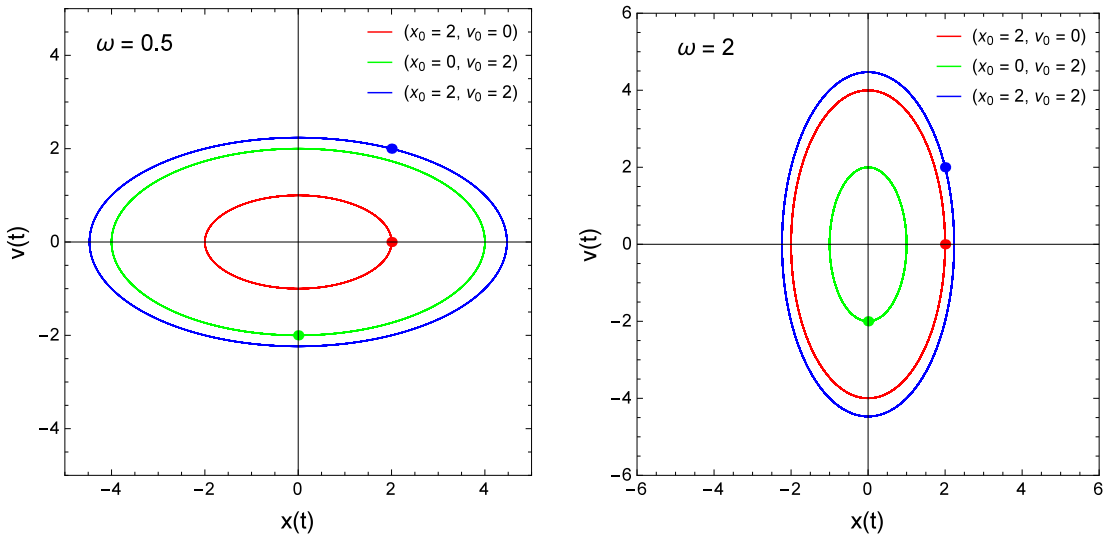


Figure 1.6: Some examples of trajectories in the state space for a simple harmonic oscillator that result from different choices of initial conditions. The dot on each curve indicates the initial state (x_0, v_0) in which the system starts at $t = 0$. From that initial point, the system proceeds in a clockwise direction along the corresponding curve, as discussed in the text. The diagram on the left illustrates the state-space trajectories for an oscillator with $\omega = 0.5$; the one on the right illustrates the trajectories for $\omega = 2$.

1.7 Energy in the Harmonic Oscillator

The fact that the motion of our simple harmonic oscillator is periodic and fact that its trajectories in state space are closed is intimately related to the conservation of mechanical energy in the system.

The kinetic energy of the mass attached to the spring is defined in terms of its mass m and its velocity v in the usual way:

$$K = \frac{1}{2}mv^2 \ . \tag{1.28}$$

For a simple harmonic oscillator, the velocity is given by Eq. (1.8), and so the kinetic energy is

$$\begin{aligned} K &= \frac{1}{2}mA^2\omega^2 \sin^2(\omega t + \phi) \\ &= \frac{1}{2}kA^2 \sin^2(\omega t + \phi) \ . \end{aligned} \tag{1.29}$$

In addition to kinetic energy, we can also define a potential energy for the system. This is possible because the restoring force provided by the spring is a **conservative force**. Recall that a conservative force is one for which the work done by the force on a particle moving from one point x_a to another x_b doesn't depend on the path the particle takes. A potential energy function U can always be associated with such a force. It turns out that for a particle moving in only one dimension, any force which depends only on the particle's position is conservative. The potential energy function for such a force is

$$U(x) = -\int_x^0 F(x')dx' \ . \tag{1.30}$$

For a simple harmonic oscillator, $F = -kx$, and so we have

$$U(x) = \int_x^0 kx'dx' = \frac{1}{2}kx^2 \ . \tag{1.31}$$

The total energy in the system is just the sum of the kinetic and potential contributions:

$$\begin{aligned} E_{\rm tot} &= \frac{1}{2}kA^2\Big[\sin^2(\omega t+\phi)+\cos^2(\omega t+\phi)\Big] \\ &= \frac{1}{2}kA^2\ . \end{aligned} \tag{1.32}$$

This is independent of t, which tells us that the total energy of the system is conserved. Again, this energy conservation is not an accident, but rather a consequence of the fact that the restoring force supplied by the spring is a conservative force for which a potential function is well-defined. To see this, we begin by noting that by definition, energy conservation means that

$$\frac{dE_{\rm tot}}{dt} \;=\; \frac{dK}{dt}+\frac{dU}{dt} \;=\; 0\ . \tag{1.33}$$

Plugging in the general forms for K and U in Eq. (1.28) and Eq. (1.30) and using the chain rule on dU/dt yields

$$0 \;=\; \frac{1}{2}m\left(2v\frac{dv}{dt}\right)-\frac{dU}{dx}\frac{dx}{dt} \;=\; mva-Fv \;=\; v(ma-F)\ . \tag{1.34}$$

Newton's Second Law tells us that $F=ma$, so $dE_{\rm tot}/dt$ indeed vanishes, implying that the total energy of the system is conserved.

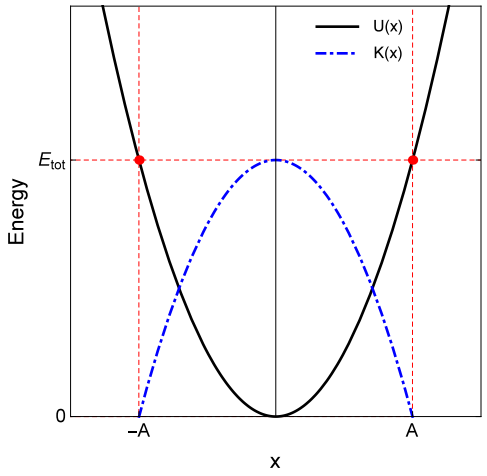


Figure 1.7: The kinetic and potential energies $K(x)$ (blue curve) and $U(x)$ (black curve) shown as functions of x for simple harmonic motion with an amplitude A. The red dots indicate the turning points at $\pm A$ where the direction of motion reverses. The horizontal line corresponds to the value of the total energy $E_{\rm tot}=K(x)+U(x)=kA^2/2$ for the system.

In Fig. 1.7, we show the kinetic and potential energies $K(x)$ and $U(x)$ as functions of x for simple harmonic motion with amplitude A. The red dots on the potential-energy curve represent the turning points at which the motion reverses direction. As can be seen in the plot, $K(x)$ vanishes at these points because $v=0$ there, while $U(x)=E_{\rm tot}$ is at its maximum. The motion is confined to the region between the vertical dashed lines.

Problems

1. Show explicitly that the solution for $x(t)$ in Eq. (1.6) satisfies the equation of motion in Eq. (1.5).

2. Another way of writing the general solution to the equation of motion Eq. (1.5) for a simple harmonic oscillator is

$$x(t) = C_1 \cos(\omega t) + C_2 \sin(\omega t) \ , \tag{1.35}$$

where C_1 and C_2 are constants.

   (a) Show explicitly that this expression for $x(t)$ is in fact a solution to Eq. (1.5).

   (b) Show that this solution is in fact equivalent to the more familiar form for $x(t)$ given in Eq. (1.6). What are the constants C_1 and C_2 in terms of A and ϕ?

3. A 5 kg block of metal is attached to a spring with a spring constant $k = 8.5$ N/m as in Fig. 1.2. You pull the block 80 cm away from its equilibrium position, and as you release it, you give it a push back toward its equilibrium position so that it is initially moving at 2 m/s.

   (a) Find the amplitude A, the phase ϕ, and the period T of oscillation.

   (b) Draw a graph of the position of the block as a function of time.

4. A mass m is suspended from a vertical spring with spring constant k, as illustrated in Fig. 1.8.

   (a) In terms of m, k, and the acceleration due to gravity g, find the distance d by which the spring will be stretched from its equilibrium length when the mass is hanging at rest.

   (b) Consider what happens if the mass is now pulled downwards an extra distance ℓ and released. Write down the equation of motion for the block the follows from Newton's Second Law, including both the gravitational force and the force provided by the spring. Make sure you are clear about which direction you are choosing as positive and what this implies for the signs of the gravitational and spring forces. Your differential equation will not look exactly like Eq. (1.5).

   (c) Show that the solution to this equation of motion has the form $x(t) = A\cos(\omega t + \phi) + C$, where C is a constant.

   (d) Express the period T of the resulting motion in terms of the displacement d that you found in part a and the acceleration due to gravity g.

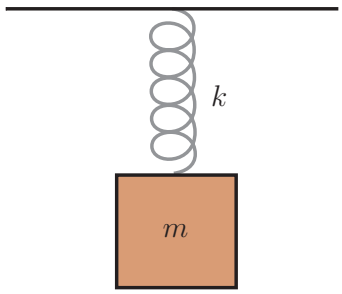


Figure 1.8: A mass m suspended vertically from a spring with spring constant k.

5. Find the frequency of oscillation for a block of mass m attached to two springs with equal spring constant k attached "in parallel" and "in series" as shown in Fig. 1.9, assuming that both springs have the same equilibrium length.

6. A cylindrical buoy with uniform density and radius r floats with ℓ of its total length submerged in water. If you push down on the buoy and release it, it will bob up and down vertically. Find the frequency of the oscillatory motion in terms of ℓ and the acceleration due to gravity g, assuming that the motion is purely in the vertical direction. Recall that an object submerged in a fluid feels a buoyant force given by $\rho_f V g$ where ρ_f is the density of the fluid and V is the volume of the fluid displaced.

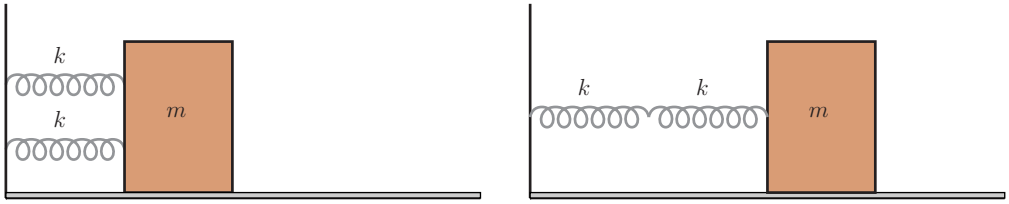


Figure 1.9: Simple harmonic oscillators with springs attached "in parallel" (left panel) and "in series" (right panel).

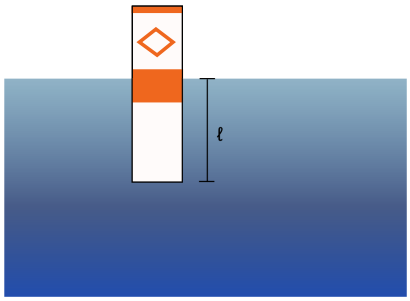


Figure 1.10: A cylindrical buoy bobbing up and down near the surface of a calm lake.

7. Show that the total energy stored in the electromagnetic fields in an LC circuit like that depicted in Fig. 1.5 is constant in time. Recall that the energy stored in the electric field beween the plates of a parallel-plate capacitor is $\frac{1}{2}Q^2/C$ and the energy stored in the magnetic field of an inductor is $\frac{1}{2}LI^2$.

8. A dielectric with a dielectric constant $\epsilon = 5$ is inserted between the plates of a parallel-plate capacitor in an LC circuit containing a capacitor with capacitance C and an inductor with inductance L. What effect does this have on the frequency of oscillation?

Chapter 2

Simple Harmonic Motion

- **The physics:** Simple pendulum, the harmonic approximation
- **The math:** Taylor and Maclaurin series, tests of convergence and remainders

2.1 Motivational Example: The Motion of a Simple Pendulum

At the end of the last section, I mentioned that a wide variety of physical systems behave approximately like a simple harmonic oscillator when the departure from equilibrium is small. In this section, we'll examine why that's the case. We'll begin with a simple example first, and then examine the mathematical machinery for characterizing small deviations in more generality.

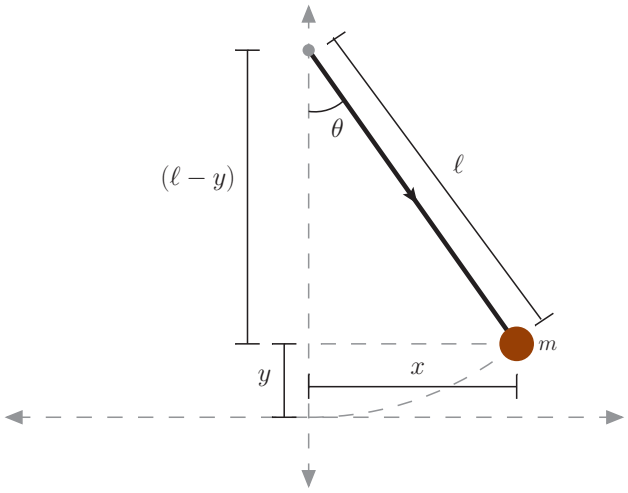


Figure 2.1: A simple pendulum consisting of a mass m attached to a rigid rod of length ℓ which swings back and forth on a pivot.

Let's begin by considering the motion of a simple pendulum consisting of a mass m attached to a rigid rod of length ℓ and negligible mass which swings freely back and forth on a pivot, as shown in Fig. 2.1. Once again, our goal will be to derive the equation of motion for the system and try to find a solution to that equation. We also want to compare that equation to the equation of motion for a simple harmonic oscillator that we saw in the previous section.

Once again, our derivation of this equation of motion begins with Newton's Second Law. For angular motion, Newton's Second Law law takes the form

$$\tau = I\alpha \tag{2.1}$$

where τ is the torque around the axis of rotation, I is the moment of inertia with respect to that axis, and α is the angular acceleration. In this case, the force acting on the mass at the end of pendulum is gravity, so $\tau = -mg\ell\sin\theta$. The moment of inertia for the mass is just $I = m\ell^2$. Since the angular acceleration, by definition, is just $\alpha = d^2\theta/dt^2$, the equation of motion is

$$m\ell^2 \frac{d^2\theta}{dt^2} \;=\; -\,mg\ell\sin\theta \;. \tag{2.2}$$

We can put this equation of motion for θ in a slightly more revealing form by moving all the constants to the right-hand side:

$$\frac{d^2\theta}{dt^2} \;=\; -\,\frac{g}{\ell}\sin\theta \;, \tag{2.3}$$

This is clearly not the equation of motion for a simple harmonic oscillator. However, as we shall soon demonstrate, it turns out that $\sin\theta \approx \theta$ for very small θ. This means that as long as θ remains very small as the system evolves over time, the approximate equation of motion that we obtain in this regime, which is

$$\frac{d^2\theta}{dt^2} \;\approx\; -\,\frac{g}{\ell}\theta \;+ [\text{small corrections}] \;, \tag{2.4}$$

does indeed have the simple-harmonic-oscillator form, with an oscillator frequency $\omega = \sqrt{g/\ell}$. However, it's only truly mathematically *equivalent* to a simple harmonic oscillator in the limit where $\theta \to 0$ — meaning either that the angle is zero (no motion)! The questions we should be asking ourselves, then, are what the "small correction" terms in Eq. (2.4) look like for small but non-zero θ.

Before we move on to discussing these correction terms, it's useful to note that one can also express the approximate equation of motion for the mass at the end of the pendulum in terms of the rectilinear coordinates x and y indicated in Fig. 2.1. Indeed, these coordinates are related to the angular coordinate θ by the relations

$$x \;=\; \ell\sin\theta \;, \qquad (\ell - y) \;=\; \ell\cos\theta \;. \tag{2.5}$$

Moreover, while both the x and y coordinates of the mass at the end of the pendulum change as it moves, we still need only one variable to completely characterize its motion. This is because the length of the rod is fixed, so x and y are related to each other by the constraint

$$\ell^2 \;=\; x^2 + (\ell - y)^2 \;. \tag{2.6}$$

Solving this equation for y gives us the equation

$$y \;=\; \ell - \sqrt{\ell^2 - x^2} \;. \tag{2.7}$$

This means that we only need to focus on the motion of the mass in the x direction, since the value of y is completely specified by the value of x.

The next step is to express the angular acceleration $d^2\theta/dt^2$ in terms of x and its derivatives. However, this is just a matter of inverting the relation in Eq. (2.5) and evaluating the time derivatives:

$$\frac{d^2\theta}{dt^2} \;=\; \frac{d}{dt}\left[\frac{d}{dt}\arcsin\left(\frac{x}{\ell}\right)\right] \;=\; \frac{d}{dt}\left[\frac{1}{\sqrt{\ell^2 - x^2}}\frac{dx}{dt}\right] \;=\; \frac{1}{\sqrt{\ell^2 - x^2}}\,\frac{d^2x}{dt^2} + \frac{x}{(\ell^2 - x^2)^{3/2}}\left(\frac{dx}{dt}\right)^2$$

Substituting this expression into Eq. (2.3), we obtain

$$\frac{1}{\sqrt{\ell^2 - x^2}}\,\frac{d^2x}{dt^2} + \frac{x}{(\ell^2 - x^2)^{3/2}}\left(\frac{dx}{dt}\right)^2 \;=\; -\,\frac{gx}{\ell^2} \;, \tag{2.8}$$

Finally, if we rearrange this expression a bit in order to get d^2x/dt^2 alone on the left-hand side, we find that

$$\frac{d^2x}{dt^2} \;=\; -\,\frac{g}{\ell}\,x\sqrt{1 - \frac{x^2}{\ell^2}} \;-\; \frac{x}{\ell^2 - x^2}\left(\frac{dx}{dt}\right)^2 \;. \tag{2.9}$$

As we might expect, the differential equation in Eq. (2.9) doesn't resemble the equation of motion for a simple-harmonic-oscillator any more than Eq. (2.3) did. However, let's consider what happens in the small-angle regime, which corresponds here to the regime in which x is very small compared to ℓ. First of all, in this regime, $\sqrt{1-x^2/\ell^2} \approx 1$. Second, we note that the velocity $v_x = dx/dt$ must also always be very small in order for x to remain very small as the system evolves; if v_x were very large, that would drive x towards very large values. Thus, in this regime, we are justified in dropping the second term in Eq. (2.9).[1] Thus, in the $x \ll \ell$ regime, the equation of motion takes the approximate form

$$\frac{d^2x}{dt^2} \;\approx\; -\frac{g}{\ell}\,x \;+\; [\text{small corrections}]\;. \tag{2.10}$$

Once again, we recover the equation for a simple harmonic oscillator with $\omega = \sqrt{g/\ell}$. Once again, however, it's only truly mathematically *equivalent* to a simple harmonic oscillator in the limit where $x/\ell \to 0$ — meaning either that the displacement is zero (no motion) or else that the length of the pendulum is infinite!

It is perhaps also worth noting that the second term in Eq. (2.9) that we are neglecting in the $x \ll \ell$ approximation is the term that corresponds to the additional centripetal force. (You may have already guessed this based on the fact that the term is proportional to the square of a velocity.) Why is this term negligible? The answer is essentially that for very small x, the mass at the end of the pendulum moves almost only in the horizontal direction. Indeed, according to Eq. (2.7), the y coordinate remains fixed at ℓ to a very good approximation.

Since this is the case, if we knew we were going to be working in the $x \ll \ell$ regime, we could have simply ignored the centripetal acceleration and set the net force in the radial direction — *i.e.*, the direction parallel to the rod – equal to zero. There are two forces which contribute to this net force: gravity and the tension force in the rod. The x and y components of the gravitational force are

$$\begin{aligned} F_{\text{grav},x} &= 0 \\ F_{\text{grav},y} &= -mg \end{aligned} \tag{2.11}$$

The tension force always acts along the radial direction. If the mass doesn't accelerate in the radial direction — which it doesn't to a very good approximation in the $x \ll \ell$ regime — the tension force must exactly cancel the component of the gravitational force acting in this direction. The magnitude of the tension force is therefore $F_{\text{tens}} = mg\cos\theta$, and the x and y components of this force are

$$\begin{aligned} F_{\text{tens},x} &= -mg\cos\theta\sin\theta \;=\; -\frac{mg}{\ell}\,x\sqrt{1-\frac{x^2}{\ell^2}} \\ F_{\text{tens},y} &= mg\cos^2\theta \;=\; mg\left(1-\frac{x^2}{\ell^2}\right)\;. \end{aligned} \tag{2.12}$$

Indeed, in the $x \ll \ell$ regime, $F_{\text{tens},y} \approx mg$, which cancels the contribution from $F_{\text{grav},y}$, so in this approximation, motion occurs only in the x direction to a very good approximation.

Since $F_{\text{grav},x} = 0$, the tension force is the only force acting in the x direction, so Newton's second law gives us

$$\frac{d^2x}{dt^2} \;\approx\; -\frac{g}{\ell}\,x\sqrt{1-\frac{x^2}{\ell^2}} \;+\; (\text{small corrections})\;. \tag{2.13}$$

Finally, setting $\sqrt{1-x^2/\ell^2} \approx 1$ as is appropriate in the $x \ll \ell$ approximation, we recover the same equation of motion we had in Eq. (2.10).

We have now seen, in a variety of different ways, how the equation of motion for a simple pendulum reduces to the equation of motion for a simple harmonic oscillator when θ (or x) is small. However, we have yet to address what the correction terms in Eq. (2.4) look like for small but non-zero θ and how important they are in terms of their effect on the motion of the pendulum. In order to address these questions, we now turn to examine a general method for making approximations of this sort: the method of Taylor series.

[1] If this makes you uneasy, think about it this way. In the regime in which the system behaves approximately like a harmonic oscillator with $\omega = \sqrt{g/\ell}$, we have $x \approx A\cos(\omega t+\phi)$, so the velocity is approximately $v_x \approx -A\omega\sin(\omega t+\phi)$. We therefore have $v_x = A^2\omega^2\sin^2(\omega t+\phi) = A^2\omega^2[1-\cos^2(\omega t+\phi)] = g(A^2-x^2)/\ell$, which means that the second term on the right side of Eq. (2.9) is suppressed relative to the first term by a factor of $(A^2-x^2)/(\ell^2-x^2)$. By assumption, the amplitude of the oscillation is small, so $\ell \gg A \geq x$, so we are justified in discarding this term.

2.2 Approximating Functions: Taylor Series

We are looking for a way of approximating a function $f(x)$ when x is very small. One fruitful first step would be to figure out a way of expressing $f(x)$ as a power series in x.

$$f(x) \;=\; \sum_{n=0}^{\infty} a_n x^n \;=\; a_0 + a_1 x + a_2 x^2 + \ldots \;, \tag{2.14}$$

where the a_n are some set of coefficients chosen such that the sum of terms on the right-hand side is equivalent to the function $f(x)$. Why is it helpful to be able to write $f(x)$ this way? The reason is basically that when x is small, the *partial* sum consisting of the first one or two terms in the whole infinite series may be an excellent approximation to $f(x)$.[2]

So how do we go about finding the coefficients a_n in Eq. (2.14)? Well, as a first step, we notice that at the point $x = 0$, all of the terms in the sum vanish except for the first one, so it must be true that

$$a_0 \;=\; f(0) \;. \tag{2.15}$$

So how do we pin down the other terms? The trick is take derivatives of both sides of Eq. (2.14). For example, taking the first derivative of both sides yields the relation

$$f'(x) \;=\; a_1 + 2a_2 x + 3a_3 x^2 + \ldots \;, \tag{2.16}$$

where the prime on $f'(x)$ indicates a derivative with respect to x. Once again, at the point $x = 0$, all terms on the right side of this equation vanish except for the first one, so we find that

$$a_1 \;=\; f'(0) \;. \tag{2.17}$$

Similarly, taking the second derivative of both sides of Eq. (2.14) yields

$$f''(x) \;=\; 2a_2 + 6a_3 x + 12a_4 x^2 + \ldots \;, \tag{2.18}$$

and by again setting $x = 0$, we find that

$$a_2 \;=\; \frac{1}{2} f''(0) \;. \tag{2.19}$$

We can continue to perform this procedure again and again, applying a different number of derivatives to each side of Eq. (2.14), in order to obtain the rest of the a_n. As is probably evident by now, the general formula for a_n is

$$a_n \;=\; \frac{1}{n!} f^{(n)}(0) \;, \tag{2.20}$$

where the notation $f^{(n)}(0)$ means that we take the nth derivative of $f(x)$ and then set $x = 0$. Thus, the series expansion of $f(x)$ is given by the general formula

$$f(x) \;=\; \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0) x^n \;. \tag{2.21}$$

Such a series expansion of $f(x)$ around the point $x = 0$ is known as the **Maclaurin series** for $f(x)$. A Maclaurin series is actually just a specific example of the more general construction known as a **Taylor series**, which is an expansion of $f(x)$ around any arbitrary value $x = a$ of its argument. The general formula for the Taylor-series expansion of $f(x)$ around $x = x_0$ is

$$f(x) \;=\; \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0)(x - x_0)^n \;. \tag{2.22}$$

[2]Later in this course, we will also use series of this sort to find solutions to differential equations, both numerically and analytically.

While you can always derive the Taylor-series expansion for any function using Eq. (2.22), there are certain Taylor series which crop up so frequently in physics applications that they're probably worth committing to memory:

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \ldots \tag{2.23}$$

$$\cos(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \ldots \tag{2.24}$$

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \ldots \tag{2.25}$$

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \ldots \tag{2.26}$$

$$(1+x)^p = 1 + px + \frac{1}{2!}p(p-1)x^2 + \ldots \ . \tag{2.27}$$

2.3 Taylor Series: Applications

Let's return to the example we were considering in Sect. 2.1 and apply what we now know about Taylor series to find an approximate form for the function $f(x) = \sqrt{1 - x^2/\ell^2}$ for small x. The first coefficient a_0 in the expansion of this function around $x = 0$ is simply $a_0 = \sqrt{1-0} = 1$. The second coefficient is

$$\begin{aligned} a_1 &= \left[\frac{d}{dx}\sqrt{1 - \frac{x^2}{\ell^2}}\right]\Bigg|_{x=0} \\ &= \left[\frac{1}{\sqrt{1 - x^2/\ell^2}}\frac{x}{\ell^2}\right]\Bigg|_{x=0} \\ &= 0 \ . \end{aligned} \tag{2.28}$$

The third is

$$\begin{aligned} a_2 &= \frac{1}{2!}\left[\frac{d^2}{dx^2}\sqrt{1 + \frac{x^2}{\ell^2}}\right]\Bigg|_{x=0} \\ &= \frac{1}{2}\left[\frac{1}{\sqrt{1 - x^2/\ell^2}}\frac{1}{\ell^2} - \frac{1}{(1 - x^2/\ell^2)^{3/2}}\frac{x^2}{\ell^4}\right]\Bigg|_{x=0} \\ &= \frac{1}{2\ell^2} \ , \end{aligned} \tag{2.29}$$

and so forth. If we were to continue this procedure explicitly up to a_4, we would find that

$$\sqrt{1 - \frac{x^2}{\ell^2}} = 1 - \frac{x^2}{2\ell^2} - \frac{x^4}{8\ell^4} + \ldots \tag{2.30}$$

We are now able to be more explicit about the size of the "small corrections" in Eq. (2.10). While both terms on the right-hand side of Eq. (2.9) contribute to these corrections, for purposes of illustration we'll concentrate on the corrections from the factor of $\sqrt{1 + x^2/\ell^2}$ in the first term. When $x \ll \ell$, the magnitude of each successive term in the Taylor series for $\sqrt{1 - x^2/\ell^2}$ is absolutely minuscule compared to the term that came before it. Therefore, in situations in which the displacement of the pendulum from equilibrium always remains extremely small, we would obtain a good approximation for this restoring-force term by keeping only the first two terms in Eq. (2.30) and approximating this term as

$$-\frac{g}{\ell}\,x\left(1 - \frac{x^2}{2\ell^2}\right) \approx -\frac{g}{\ell}\,x + \frac{g}{2\ell^3}\,x^3 \ . \tag{2.31}$$

Of course we'd get a slightly more accurate approximation by retaining more terms in this series, but the operative word is "slightly."

2.4 Tests of Convergence

Taylor series, as we shall soon see, are an extremely useful mathematical tool for solving physics problems. However, there are some limits to their applicability. For one thing, not every function $f(x)$ can be represented in this manner. First of all, in order for a function to be represented by a Taylor series, $f(x)$ must satisfy it must be infinitely differentiable in order that all of the coefficients a_n given by Eq. (2.20) to be well defined. Moreover, since Eq. (2.22) is an infinite series, there is no guarantee that it **converges** — *i.e.*, that the sum of its terms approaches a finite limiting value. Fortunately, there exist a number of tests to determine whether or not a series converges, some of the most useful of which I will now describe. In the descriptions of these tests, we'll use the symbol b_n to denote the nth term in the series we're considering. Of course $b_n = a_n x^n$ for a Taylor series, but we want to be as general as possible here because these convergence tests apply not only to Taylor series, but to other infinite series as well.

The Preliminary Test:

If $\lim_{n\to\infty} b_n \neq 0$, meaning that the terms in the series themselves do not tend toward zero, the series diverges. If $\lim_{n\to\infty} b_n = 0$, further tests are needed to determine whether the series converges.

The Integral Test:

In order for this test to be used on a given series, there must exist some finite value of n, which we'll call N, above which all of the b_n are positive and decreasing — *i.e.*, we must have $0 < b_n \leq b_N$ for $n > N$. If this criterion is satisfied, the series converges if the integral

$$\int_N^\infty b_n dn \tag{2.32}$$

is finite and diverges if it is infinite.[3]

The Ratio Test:

Define r_n to be the absolute value of the ratio of the two successive terms b_{n+1} and b_n in the series:

$$r_n \equiv \left| \frac{b_{n+1}}{b_n} \right| \tag{2.33}$$

Evaluate r_n and take the limit $r \equiv \lim_{n\to\infty} r_n$. If $r < 1$, the series converges. If $r > 1$, the series does not converge. If $r = 1$, the test is inconclusive and you need another test to determine whether or not the series converges.

Other tests for convergence exist as well, many of which involve performing term-by-term comparisons with the terms of some other series whose convergence properties you already know. Several of these tests are discussed, *e.g.*, in Sect. 1.5 - 1.8 of Boas, Mary L., *Mathematical Methods in the Physical Sciences*, Wiley 2005.

As an example of how these tests work in practice, let's try using them to determine the convergence properties of the function $\ln(1+x)$. First, we need to extract a general expression for the terms appearing in Eq. (2.26). Thus, we note that

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \ldots = \sum_{n=1}^{\infty} \frac{1}{n}(-1)^{n+1}x^n \ , \tag{2.34}$$

so we have

$$b_n = \frac{1}{n}(-1)^{n+1}x^n \ . \tag{2.35}$$

We are now ready to apply the preliminary test. Our results, as you might have expected, depend on the value of x. For $x > 1$, we find that the value of b_n diverges in the $n \to \infty$ limit, so the series itself must

[3] The lower limit of integration is essentially irrelevant here, since it's the behavior of the series as $n \to \infty$ that we're interested in. Some texts omit the lower limit completely.

also diverge. By contrast, $\lim_{n\to\infty} b_n = 0$ for $-1 \le x \le 1$, so we need to apply more tests. Since the sign of b_n alternates back and forth, there exists no N for which all b_n for $n > N$ are positive, to the integral test can't be applied here. Applying the ratio test, we find that

$$r_n \;=\; \left|\frac{n(-1)^{n+2}x^{n+1}}{(n+1)(-1)^{n+1}x^n}\right| \;=\; \frac{n}{n+1}|x| \;, \tag{2.36}$$

and the $n \to \infty$ limit of this is

$$r \;=\; \lim_{n\to\infty}\frac{n}{n+1}|x| \;=\; |x| \;, \tag{2.37}$$

so once again, the results of the test depend on the value of x. For $|x| > 1$, we have $r > 1$, and so the series diverges, but we already knew that from the preliminary test. The new piece of information we learn is that the series must converge for $|x| < 1$ because $r < 1$. For $|x| = 1$, the ratio test is inconclusive, and we'd still need to apply other tests to see whether the series converges for $x = \pm 1$.

A we have seen, the convergence properties of a Taylor series for a given function $f(x)$ frequently depend on the value of x. The **interval of convergence** for such a function is the range of values for x for which the series converges.[4] For example, we've just proved that the interval of convergence for $\ln(1+x)$ includes all points between -1 and $+1$.[5]

There is one final caveat I should mention about Taylor series. This is that sometimes the Taylor series for a function $f(x)$ will converge, but it won't actually converge to the value of the function! A classic example of a function which exhibits this "sick" behavior is

$$f(x) \;=\; e^{-1/x^2} \;. \tag{2.38}$$

This function is zero at $x = 0$, and so are all of its derivatives $f^{(n)}(0)$, so all of the Taylor-series coefficients a_n for this function vanish. However, for $x^2 > 0$, the value of $f(x)$ is clearly not zero, so the Taylor series fails to represent the function there.

2.5 Remainders

Whenever one makes any sort of approximation, it's always important to understand the error associated with that approximation. In the case, we would like to be able to compute the error associated with truncating the Taylor series expansion for $f(x)$ around $x = x_0$ after some finite number of terms — say, terms up to and including x^n. In other words, if we break the full Taylor series around $x = x_0$ into two parts, like so

$$f(x) \;=\; \sum_{p=0}^{n}\frac{1}{p!}f^{(p)}(x_0)(x-x_0)^p + \sum_{p=n+1}^{\infty}\frac{1}{p!}f^{(p)}(x_0)(x-x_0)^p \tag{2.39}$$

and keep only the first sum on the right-hand side, we want to be able to compute how big the second sum — *i.e.*, the **remainder**, often denoted R_n — actually is.

There are a number of different methods for determining R_n for different types of series. One useful, and fairly universally applicable method involves the use of a theorem[6] which states that the remainder can be written in the form

$$R_n \;=\; \frac{1}{(n+1)!}(x-x_0)^{n+1}f^{(n+1)}(\xi) \;, \tag{2.40}$$

where ξ is some as-yet unknown number that lies between x_0 and x. This theorem doesn't tell us what ξ actually is. However, it *does* tell us that if we can put a bound on ξ from some other consideration, we can use this to put a bound on R_n.

[4] Despite its name, the "interval" of convergence for certain series may consist of a single point — ie, one particular value of x — a set of such points, a set of multiple distinct intervals, *etc.*.

[5] Determining whether -1 and $+1$ fall in this window is the subject of Problem 5.

[6] A straightforward proof of this theorem and a list of other forms in which R_n can be written can be found in §5.6 of Arfken & Weber, *Mathematical Methods for Physicists*.

Example: Estimating a Remainder

As an example, let's consider the function $f(x) = e^x$. If we approximate this function by performing a Taylor-series expansion around $x = 0$, but keep only the first two terms $e^x \approx 1 + x$ in Eq. (2.25), how big can x get before the error could exceed 1% of the true result?

We can't derive the exact value of R_n from Eq. (2.40), because it doesn't tell us the value of ξ. However, *can* use it to place a conservative bound on R_n. In particular, this equation tells us that

$$R_1 \;=\; \frac{1}{2!}x^2 f^{(2)}(\xi) \;=\; \frac{1}{2}x^2 e^{\xi} \;, \tag{2.41}$$

so now we just need to figure out a way of constraining this expression. One helpful fact we can use is that $1 \leq e^{\xi} \leq e^x$ for any ξ in the range $0 \leq \xi \leq x$. Feeding this constraint into our expression for R_1 tell us us that the remainder must lie within the range

$$\frac{x^2}{2} \leq R_1 \leq \frac{1}{2}x^2 e^x \;. \tag{2.42}$$

These upper and lower bounds on R_1 are shown as functions of x in Fig. 2.2. The corresponding upper bound on the percent error associated with this approximation for any particular value of x is just the the upper limit on R_1 divided by the exact value of e^x:

$$[\text{Percent error}] \;\leq\; \frac{1}{2}x^2 \;. \tag{2.43}$$

Thus, we can guarantee that the error associated with this approximation doesn't exceed 1% as long as $x \leq \sqrt{2 \times 0.01} \approx 0.14$.

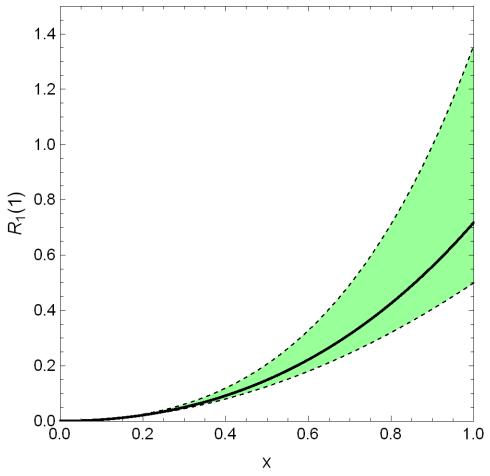


Figure 2.2: A plot showing the upper and lower limits (dashed lines) in Eq. (2.42) on the remainder R_1 in the Taylor expansion $e^x \approx 1 + x$ as a function of x. The solid curve indicates the exact value of R_1.

2.6 The Harmonic Approximation

Taylor series provide an important perspective on why harmonic motion is such a ubiquitous phenomenon in nature. The fact that the equation of motion for our pendulum in Sect. 2.1 looked an awful lot like the simple-harmonic-oscillator equation for small x was not special or unusual. In fact, if you take almost any physical system that has a stable equilibrium point x_{eq} and expand the restoring force $F(x)$ in a Taylor series around this equilibrium solution, you'll find that the equation of motion looks just like the equation for a simple harmonic oscillator for small $\Delta x = x - x_{\text{eq}}$.

In order to illustrate this, it's helpful to revisit in a bit more mathematical detail what we mean when we say that a system has a stable equilibrium point. We recall from Ch. 1 of these lecture notes that an equilibrium point — be it stable or unstable — is a point a which the net force vanishes — *i.e.*, where

$$F(x_{\rm eq}) \;=\; 0 \;. \tag{2.44}$$

As we also discussed briefly in Ch. 1, a *stable* equilibrium point is one for which the restoring force acts to push the system back towards equilibrium if it strays to either side. In other words, $F(x)$ must be negative for $x > x_{\rm eq}$ and positive for $x < x_{\rm eq}$. One way of testing whether a force meets this condition is to examine the derivative $F(x)$ with respect to x at the equilibrium point. If the derivative is positive, then the equilibrium point is clearly unstable, meaning that small deviations from $x_{\rm eq}$ result in the the system being pushed even further away from equilibrium. On the other hand, if the derivative is negative, the system is stable, meaning that small deviations by from $x_{\rm eq}$ result in the system being driven back toward $x_{\rm eq}$.

For a conservative force, we can also express these conditions on $F(x)$ as conditions on the potential-energy function $U(x)$. In particular, We saw in Ch. 1 that

$$F \;=\; -\frac{dU}{dx} \;, \tag{2.45}$$

so the stability criterion for dF/dx corresponds to a criterion for d^2U/dx^2 — *i.e.*, a condition on the *concavity* of the potential-energy function. Specifically, we have

$$\begin{aligned}
&\left.\frac{dF}{dx}\right|_{x=x_{\rm eq}} < 0 \quad \text{or} \quad \left.\frac{d^2U}{dx}\right|_{x=x_{\rm eq}} > 0 \quad \longrightarrow \quad \text{stable equilibrium point} \\
&\left.\frac{dF}{dx}\right|_{x=x_{\rm eq}} > 0 \quad \text{or} \quad \left.\frac{d^2U}{dx}\right|_{x=x_{\rm eq}} < 0 \quad \longrightarrow \quad \text{unstable equilibrium point} \\
&\left.\frac{dF}{dx}\right|_{x=x_{\rm eq}} = 0 \quad \text{or} \quad \left.\frac{d^2U}{dx}\right|_{x=x_{\rm eq}} = 0 \quad \longrightarrow \quad \text{test inconclusive}
\end{aligned} \tag{2.46}$$

If $d^2U/dx^2 > 0$ and the equilibrium solution at $x = x_{\rm eq}$ passes this **basic stability test**, it's not hard to show that the system will function like a simple harmonic oscillator if the deviation from equilibrium is sufficiently small. It's easy to demonstrate this by expanding $F(x)$ as a Taylor series around the equilibrium point $x_{\rm eq}$:

$$F(x) \;=\; a_1(x - x_{\rm eq}) + a_2(x - x_{\rm eq})^2 + a_3(x - x_{\rm eq})^3 + \ldots \tag{2.47}$$

I have omitted a_0 here because Eq. (2.44) tells me that $F(x_{\rm eq}) = 0$ at any equilibrium point, so it must be true that $a_0 = 0$. The derivative dF/dx at precisely $x = x_{\rm eq}$ is just the coefficient a_1

$$\left.\frac{dF}{dx}\right|_{x=x_{\rm eq}} \;=\; \Big[a_1 x + 2a_2(x - x_{\rm eq})^2 + 3a_3(x - x_{\rm eq})^2 + \ldots\Big]|_{x=x_{\rm eq}} \;=\; a_1 \;. \tag{2.48}$$

We therefore see that the basic stability test in Eq. (2.46) boils down to a statement about the sign of the Taylor-series coefficient a_1. If $a_1 < 0$, the stability test tells you that the equilibrium at the point $x_{\rm eq}$ is stable. If $a_1 > 0$, it tells you that the equilibrium is unstable. If $a_1 = 0$, the test is inconclusive and we need more information to determine the stability of the system for x near $x_{\rm eq}$.

In the case in which the equilibrium point does pass the basic stability test, we can write our Taylor-series expression for $F(x)$ in a slightly more revealing form by defining a *positive* constant $k \equiv -a_1 > 0$:

$$F(x) \;=\; -k(x - x_{\rm eq}) + a_2(x - x_{\rm eq})^2 + a_3(x - x_{\rm eq})^3 + \ldots \tag{2.49}$$

When x is very close to $x_{\rm eq}$, the terms involving higher powers of the displacement $\Delta x = (x - x_{\rm eq})$ will be very small compared to the linear term, and the restoring force is very well approximated by $F(x) \;\approx\; -k(x - x_{\rm eq})$. In other words the system functions like a simple harmonic oscillator to a very good approximation when the displacement Δx is sufficiently small. This is true of *any* equilibrium point in any physical system

which passes this stability test. This is why nature is full of physical systems which exhibit simple harmonic motion: although the number of systems which function exactly like a simple harmonic oscillator is small, the number of systems which function approximately like a simple harmonic oscillator is huge!

We can also view the harmonic approximation s a Taylor-series expansion of the potential-energy function $U(x)$. The change in the potential energy when we displace the system from its equilibrium position $x_{\rm eq}$ to some other nearby position x is is just minus the integral of the restoring force from $x_{\rm eq}$ to x. Since we're assuming that x is very close to $x_{\rm eq}$, we can use the harmonic approximation for the restoring force, which gives us

$$U(x) \;=\; -\int_{x_{\rm eq}}^{x} F(x')dx' \;\approx\; \int_{x_{\rm eq}}^{x} k(x'-x_{\rm eq})dx' \;\approx\; \frac{1}{2}k(x-x_{\rm eq})^2 + C \;, \tag{2.50}$$

where C is an integration constant which represents the value of the potential-energy function at $x = x_{\rm eq}$. Explicitly setting $C = U(x_{\rm eq})$, we have:

$$U(x) \;=\; U(x_{\rm eq}) + \frac{1}{2}k(x-x_{\rm eq})^2 \tag{2.51}$$

which is in fact the Taylor series for $U(x)$ up to and including terms quadratic in x. Indeed, the Taylor-series coefficient a_2 is just $U^{(2)}(x_{\rm eq})/2! = k/2$, and the coefficient a_1 vanishes because, by definition, $U^{(1)}(x_{\rm eq}) = 0$ at an equilibrium point. We can therefore view the harmonic approximation as an approximation of the potential-energy function by a parabola in the vicinity of $x_{\rm eq}$, as illustrated in Fig. 2.3. Note that this approximation is reasonably good for x near $x_{\rm eq}$, but likely to fail miserably for large displacements.

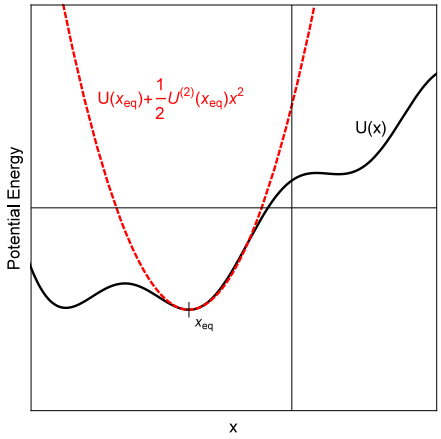


Figure 2.3: An illustration of the harmonic approximation of an arbitrary potential $U(x)$ for values of x near a stable equilibrium point at $x_{\rm eq}$.

There do exist physical systems which have stable equilibrium points for which the stability test in Eq. (2.46 is inconclusive. The equation of motion for such a system does not reduce to the simple-harmonic-oscillator form $\ddot{x} \approx -k(x - x_{\rm eq})$ for small Δx. Nevertheless, the approximate equation of motion that does emerge for small Δx can still exhibit periodic solutions $x(t)$. One example of this would be a system with a restoring force $F(x) \;=\; -kx^3$. The point $x = 0$ is clearly an equilibrium point, since $F(0) = 0$. However, because

$$\left.\frac{dF}{dx}\right|_{x=0} \;=\; 0 \;, \tag{2.52}$$

the Taylor-series coefficient a_1 vanishes.[7] Another example of a situation in which the the basic stability test would be inconclusive is one in which the derivative of $F(x)$ is discontinuous at $x = x_{\rm eq}$. In this case, the Taylor series for $F(x)$ isn't even well defined at this point because the function isn't infinitely differentiable.

[7]When $f(x)$ is a polynomial, they Taylor series for $f(x)$ is just $f(x)$ itself.

2.7 Applications of the Harmonic Approximation

One place where the harmonic approximation is frequently used in physics and chemistry is in calculating the interaction between pairs of atoms in molecules or crystals. To see how this works in practice, let's consider a pair of atoms with masses m_1 and m_2 which are exerting forces on each other, as shown in Fig. 2.4. Let's assume that Newton's Third Law holds, so that the force F_{12} exerted on atom 1 by atom 2 is equal and opposite to the force F_{21} exerted on atom 2 by atom 1. Furthermore, let's assume that the forces the atoms exert on each other depends only on the distance $r \equiv x_2 - x_1$ between them, where x_1 and x_2 are the respective coordinates of atoms 1 and 2.

Figure 2.4: A pair of atoms with masses m_1 and m_2 exerting forces on each other.

Newton's Second Law gives us a pair of differential equations for x_1 and x_2:

$$m_1 \frac{d^2 x_1}{dt^2} \;=\; F_{12} \;, \qquad m_2 \frac{d^2 x_2}{dt^2} \;=\; F_{21} \;=\; -\,F_{12} \tag{2.53}$$

If we divide each of these equations of motion by the respective mass of the atom and then subtract them, we get

$$\frac{d^2 x_2}{dt^2} - \frac{d^2 x_1}{dt^2} \;=\; \left(\frac{1}{m_2} + \frac{1}{m_1} \right) F_{21} \;=\; \frac{m_1 + m_2}{m_1 m_2} F_{21} \;. \tag{2.54}$$

You may recall from General Physics that the collection of masses appearing on the right-hand side of this equation is the inverse of what's called the **reduced mass**

$$\mu \;\equiv\; \frac{m_1 m_2}{m_1 + m_2} \;. \tag{2.55}$$

of the system. Since one of our basic assumptions was that F_{21} depends only on r, we'll drop the subscripts and just call it $F(r)$. Writing the force this way makes it clear that Eq. (2.54) is just an differential equation of motion for the relative coordinate r:

$$\mu \frac{d^2 r}{dt^2} \;=\; F(r) \;. \tag{2.56}$$

This equation has exactly the same form that one would obtain from Newton's Second Law for a particle with mass μ.

Example: Diatomic Molecule

As an example of how the harmonic approximation is applied, let's consider the interaction between the atoms in a diatomic molecule. This interaction is frequently modeled using the **Morse Potential**, which is given by

$$U(r) \;=\; D \left[1 - e^{-\beta(r-a)} \right]^2 \;, \tag{2.57}$$

where r, once again, represents the separation between the two atoms, and where D, β, and a are positive constants. For example, for an N_2 molecule, $D = 9.9$ eV, $a = 1.1$ Å, and $\beta = 2.85$ Å^{-1}.

Our first task is to find the equilibrium point, which we'll call $x_{\rm eq}$. We know that the derivative of the potential with respect to r must vanish at this point, so we have

$$\left. \frac{dU}{dr} \right|_{r = x_{\rm eq}} \;=\; 2D \left[1 - e^{-\beta(x_{\rm eq} - a)} \right] \beta e^{-\beta(x_{\rm eq} - a)} \;=\; 0 \;. \tag{2.58}$$

We can see that this condition is satisfied at $x_{\rm eq} = a$ because the quantity in brackets vanishes at that point. Now we want to test whether this equilibrium point is stable by taking the second derivative of $U(r)$ and applying the basic stability test in Eq. (2.46). Doing do, we find that

$$\begin{aligned} \frac{d^2U}{dr^2} &= -2D\beta^2\left[1-e^{-\beta(r-a)}\right]e^{-\beta(r-a)} + 2D\beta^2 e^{-2\beta(r-a)} \\ &= 2D\beta^2\left[2e^{-\beta(r-a)}-1\right]e^{-\beta(r-a)} \ , \end{aligned} \tag{2.59}$$

and at the equilibrium point, where $r = x_{\rm eq} = a$, this becomes

$$\left.\frac{d^2U}{dr^2}\right|_{r=x_{\rm eq}} = 2D\beta^2 \ . \tag{2.60}$$

This quantity is positive, so we know that $x_{\rm eq}$ is a stable equilibrium point and that the restoring force $F(r) = -dU/dr$ reduces to the approximate, harmonic-oscillator form $F(r) \approx -k(r - x_{\rm eq})$ when r is very near $x_{\rm eq}$. Moreover, since

$$\left.\frac{d^2U}{dr^2}\right|_{r=x_{\rm eq}} = -\left.\frac{dF}{dr}\right|_{r=x_{\rm eq}} = k \ , \tag{2.61}$$

the effective "spring constant" is just

$$k = 2D\beta^2 \ . \tag{2.62}$$

Now that we have a general formula, it's helpful to plug in some realistic numbers in order to get a sense of what the frequencies associated with molecular vibrations actually are. Plugging in the parameters given above for N_2, we find that $k = 2573$ N/m. The reduced mass for two nitrogen atoms is $\mu = m_{\rm N}/2 = 7$ amu $= 1.162 \times 10^{-26}$ kg, and so the vibrational frequency is

$$f_{\rm N_2} = \frac{\omega}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} = 75 \text{ THz}. \tag{2.63}$$

This is very close to the result $f_{\rm N_2} = 71$ THz obtained from direct measurement.[8]

Problems

1. Derive each of the Taylor-series expansions in Eqs. (2.23) through (2.27) up to the same order quoted here in x using the general formula for the coefficients in Eq. (2.20).

2. Find the first two non-zero terms in the Maclaurin series for the error function

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt \ . \tag{2.64}$$

   Use these to obtain numerical estimates for erf(0.1). Check your answer using the built-in function "Erf" in Mathematica.

3. Taylor series can be added, multiplied, or divided. Mulitply the series for $\cos x$ and $\sin x$ together, keeping the first three non-zero terms in the product. Verify that the trigonometric relation

$$\sin(2x) = 2\sin x \cos x \ , \tag{2.65}$$

   holds to this order in the Taylor-series expansion of each side of the equation.

[8] While this is a nice cross-check, the reasoning is actually somewhat circular because the parameters D, β, and a for the Morse potential are derived from spectroscopic measurements.

4. In special relativity, the energy of a particle of mass m and velocity v is given by

$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \, . \tag{2.66}$$

This includes both the rest energy of the particle and its kinetic energy. Find the first three non-zero terms in the Maclaurin-series expansion of E around $v = 0$. The first two should be familiar. Note that it may be easier to find the expansion for $E(u)$, where $u = v^2/c^2$, and then substitute back in for u at the end.

5. Determine whether $\ln(1+x)$ converges for the value $x = 1$. In order to do this, you'll need to apply one or more of the additional tests of convergence described in Sect. 1.5 - 1.8 of Boas, Mary L., *Mathematical Methods in the Physical Sciences*, Wiley 2005.

6. The **Lennard-Jones potential** is an empirically-derived formula for the potential-energy function which describes the interaction between a pair of neutral atoms separated by a distance r. It is given by

$$U(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] \, , \tag{2.67}$$

where the constants ϵ and σ depend on the particular atomic species in question. The second (attractive) term in this potential arises from a fluctuating electric-dipole interaction; the first term is purely empirical. For neon (Ne), the constants are given by $\epsilon = 3.1$ meV and $\sigma = 2.74$ Å. Find, the effective "spring constant" k (in N/m) for the force between two Ne atoms. How does this compare to the k value for a typical mechanical spring that you might have encountered in Introductory Physics Lab?

Figure 2.5: A wineglass with a spherical shape.

7. The bowl of a wine glass has a spherical shape with a radius of $r = 4$ cm, as shown in Fig. 2.5. It is half full of wine so that the wine constitutes exactly one half of a hemisphere. Determine the "sloshing period" for small oscillations. Assume that the surface remains flat and find the moment of inertia and center of mass of the hemisphere of wine; the problem is then the same as a physical pendulum. In practice, the easiest way to measure this sloshing period would be to drive the oscillation with your hand and measure the (angular) frequency at which the wine in the glass responds most violently. As we'll see later on, this "resonant frequency" is approximately equal to the natural frequency ω of the system.

Chapter 3

Complex Variables

- **The physics:** Complex impedances
- **The math:** Complex numbers, the complex plane, Euler's formula

3.1 Complex Numbers

Chances are, you already have some experience dealing with complex numbers from solving for the roots of certain polynomial equations. For example, consider the quadratic equation $2x^2 - 4x + 6 = 0$. The solutions to this equation, as given by the quadratic formula, are given by the quadratic formula:

$$x \;=\; \frac{4 \pm \sqrt{4^2 - 4 \times 2 \times 6}}{2 \times 2} \;=\; 1 \pm \sqrt{-2}\;. \tag{3.1}$$

There is no real number whose square is -2. Let us therefore define a non-real number, which we'll call the **imaginary number**, so that[1]

$$i \;\equiv\; \sqrt{=1}\;. \tag{3.2}$$

Thus, it is understood that $i^2 = -1$. We can express any quantity which involves the square root of a negative real number by pulling out a factor of $\sqrt{-1} = i$. For example, we can write the solutions for x in Eq. 3.1 in the form

$$x \;=\; 1 \pm i\sqrt{2}\;. \tag{3.3}$$

These two solutions are examples of **complex numbers** — numbers which have both a purely real piece and a piece proportional to the imaginary number. In general, a complex number z can be written in the form

$$z \;=\; x + iy\;, \tag{3.4}$$

where x and y are real numbers. The number x is called the **real part** of z, often written as $x = \mathrm{Re}[z]$. The coefficient y of the imaginary number i is called the **imaginary part** of z, often written as $y = \mathrm{Im}[z]$. It's important to remember that although this coefficient is called the "imaginary part" of z, the quantity y is a purely real number!

Adding or subtracting complex numbers is as simple as adding or subtracting their real and imaginary parts. For example, if we have two complex numbers $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, their sum is just

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)\;. \tag{3.5}$$

Multiplying complex numbers is likewise straightforward. Indeed, it's simply a matter of applying the distributive property:

$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) \;=\; x_1x_2 + ix_1y_2 + iy_1x_2 + i^2y_1y_2 \;=\; (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2)\;. \tag{3.6}$$

[1] Physicists and most mathematicians use the symbol i to refer to the imaginary number. Engineers typically use the symbol j. I will use the symbol i to refer to the imaginary number throughout these notes.

The division of one complex number by another is also a straightforward extension of the way division works for real numbers:

$$\frac{z_1}{z_2} \;=\; \frac{x_1+iy_1}{x_2+iy_2}\;. \tag{3.7}$$

The trick to expressing this ratio of complex numbers in the form given in Eq. (3.4) is to multiply both the numerator and the denominator by $x_2 - iy_2$:

$$\frac{z_1}{z_2} \;=\; \frac{(x_1+iy_1)(x_2-iy_2)}{(x_2+iy_2)(x_2-iy_2)} \;=\; \left(\frac{x_1x_2+y_1y_2}{x_2^2+y_2^2}\right) + i\left(\frac{y_1x_2-x_1y_2}{x_2^2+y_2^2}\right)\;. \tag{3.8}$$

The reason why multiplying both the numerator and denominator by $x_2 - iy_2$ was helpful in simplifying this expression is that the product $(x_2+iy_2)(x_2-iy_2) = x^2+y^2$ is purely real. For this reason, the complex number $x_2 - iy_2$ has a very special relationship to $x_2 + iy_2$. In general, for any complex number $z = x + iy$, the number[2]

$$z^* \;\equiv\; x - iy \tag{3.9}$$

is called the **complex conjugate** of z. The product z^*z is always a purely real number. For this reason, the absolute value of a complex number is defined to be

$$|z| = \sqrt{z^*z}\;. \tag{3.10}$$

There is one other important thing to keep in mind about working with complex quantities. There are some operations you can perform on complex numbers for which it doesn't matter whether you take the real parts of those numbers before you perform the operation or whether you perform the operation first and then take the real part of what you get. These include:

- Taking the derivative of a complex-valued function $z(t) = x(t) + iy(t)$ with respect to t:

$$\mathrm{Re}\left[\frac{d}{dt}z(t)\right] \;=\; \mathrm{Re}\left[\frac{d}{dt}x(t) + i\frac{d}{dt}y(t)\right] \;=\; \frac{d}{dt}x(t) \;=\; \frac{d}{dt}\mathrm{Re}\big[z(t)\big]\;. \tag{3.11}$$

- Adding (or subtracting) a pair of complex numbers $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$:

$$\mathrm{Re}\big[z_1+z_2\big] \;=\; \mathrm{Re}\big[x_1+iy_1+x_2+iy_2\big] \;=\; x_1+x_2 \;=\; \mathrm{Re}[z_1]+\mathrm{Re}[z_2]\;. \tag{3.12}$$

However, there are also operations for which the result you get from taking the real parts first and then perform the operations are *not* the same as the result you get by performing the operation first and then taking the real part of what you get. These include:

- The multiplication of two complex numbers z_1 and z_2. If we multiply first and then take the real part of the product, we get

$$\mathrm{Re}\big[z_1z_2\big] \;=\; \mathrm{Re}\big[(x_1+iy_1)(x_2+iy_2)\big] \;=\; \mathrm{Re}\big[x_1x_2-y_1y_2+i(x_1y_2+y_1x_2)\big] \;=\; x_1x_2-y_1y_2\;. \tag{3.13}$$

  By contrast, if we take the real parts and then multiply them, we get a different result:

$$\mathrm{Re}[z_1]\mathrm{Re}[z_2] \;=\; x_1x_2\;. \tag{3.14}$$

- The division of one complex number z_1 by another z_2. If we divide and then take the real part, we get

$$\mathrm{Re}\left[\frac{z_1}{z_2}\right] \;=\; \mathrm{Re}\left[\frac{x_1+iy_1}{x_2+iy_2}\right]\;, \tag{3.15}$$

  which we can simplify by multiplying both the numerator and the denominator by the complex conjugate z_2^* of z_2:

$$\mathrm{Re}\left[\frac{z_1}{z_2}\right] \;=\; \mathrm{Re}\left[\frac{(x_1+iy_1)(x_2-iy_2)}{(x_2+iy_2)(x_2-iy_2)}\right] \;=\; \mathrm{Re}\left[\frac{x_1x_2+y_1y_2+i(y_1x_2-y_2x_1)}{x_2^2+y_2^2}\right] \;=\; \frac{x_1x_2+y_1y_2}{x_2^2+y_2^2}\;. \tag{3.16}$$

[2]The complex conjugate of a complex variable z is typically denoted either z^* or $\bar{z}$.

By contrast, if we take the real parts of z_1 and z_2 and then divide, we get a different result:

$$\frac{\mathrm{Re}[z_1]}{\mathrm{Re}[z_2]} \;=\; \frac{x_1}{x_2}\;. \tag{3.17}$$

It's important to keep this in mind when working with complex quantities. If you're not careful in your calculations about when it's appropriate to take the real (or imaginary) part of an expression, you're likely to run into problems.

3.2 The Complex Plane

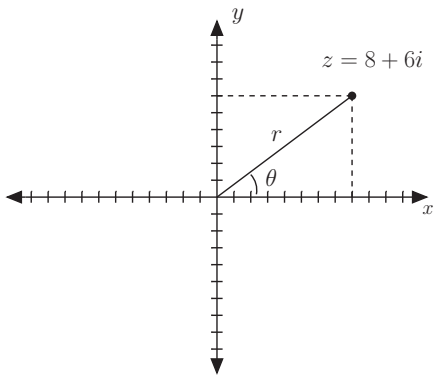


Figure 3.1: The complex number $z = 8+6i$ plotted as a point on the complex plane. The x and y coordinates are just $\mathrm{Re}[z]$ and $\mathrm{Im}[z]$, respectively. The geometric interpretations of the magnitude r and the complex phase θ are also shown.

The **magnitude** of z, usually denoted r, of a complex number is just the distance from the origin to the corresponding point on the complex plane. In other words, it's the length of the corresponding "vector" in the complex plane. The **complex phase** of z is the angle that "vector" makes with the real axis. Since r, x, and y form a right triangle, the relationship between the rectilinear coordinate x and y and the polar coordinates r and θ is

$$x \;=\; r\cos\theta\;, \qquad\qquad y \;=\; r\sin\theta\;. \tag{3.18}$$

By solving this system of equations for r and θ, we obtain the inverse relations

$$r \;=\; \sqrt{x^2+y^2}\;, \qquad\qquad \theta \;=\; \arctan\left(\frac{y}{x}\right)\;. \tag{3.19}$$

Note that $r = |z|$, meaning that the modulus of a complex number z is simply its absolute value, and can be found using Eq. (3.10)

The relations in Eq. (3.18) tell us that any complex number can be written in the form

$$z \;=\; r(\cos\theta + i\sin\theta)\;. \tag{3.20}$$

We can actually use some of what we've learned about Taylor series to write z in an even more compact and useful form. We begin by recalling that the Maclaurin series for $\sin x$ and $\cos x$ are

$$\sin x \;=\; x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots \tag{3.21}$$

$$\cos x \;=\; 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots\;, \tag{3.22}$$

while the Maclaurin series for e^x is

$$e^x \;=\; 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots \;. \tag{3.23}$$

Now let's examine what this Maclaurin series looks like when the argument is imaginary:

$$\begin{aligned} e^{i\theta} &= 1 + (i\theta) + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \ldots \\ &= 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \ldots \end{aligned} \tag{3.24}$$

If we look back at Eqs. (3.21) - (3.22), we see that this is just the sum of the Maclaurin series for $\cos\theta$ and i times the Maclaurin series for $\sin\theta$:

$$e^{i\theta} \;=\; \cos\theta + i\sin\theta \;. \tag{3.25}$$

This relationship between the exponential function and the trigonometric functions sine and cosine is known as **Euler's Formula**, which is one of the most important identities in all of mathematics. For example, in conjunction with Eq. (3.20), Euler's formula tells us that any complex number can be written in the form

$$z \;=\; re^{i\theta} \;, \tag{3.26}$$

which is often referred to as the **polar form** for parametrizing complex numbers. For reasons that a glance at Fig. 3.1 makes clear, the form $z = x + iy$ is often referred to as the **rectangular form**.

3.3 Complex Variables and the Simple Harmonic Oscillator

Consider the equation of motion for a simple harmonic oscillator:

$$\frac{d^2x}{dt^2} \;=\; -\,\omega^2 x \;. \tag{3.27}$$

We've already seen that the general solution to this equation takes the form $x(t) = A\cos(\omega t + \phi)$. However, we now note that the function

$$x(t) \;=\; Ae^{i(\omega t + \phi)} \tag{3.28}$$

also satisfies Eq. (3.27). Indeed, you can verify for yourself that this is the case. So why haven't we discussed this solution before? The reason is that this $x(t)$ is a complex quantity, whereas a physical quantity like the position of a mass on a spring is always a real. Thus, while this complex solution satisfied the equation of motion, it made no sense physically.

However, we have already seen that it's sometimes useful to think about physical systems in ways that involve additional, unphysical coordinates. In particular, we have seen that there is a very close correspondence between harmonic motion and circular motion. For example, we can think of the position of a mass on a spring undergoing simple harmonic motion as being equivalent to the projection in the x direction of the position of a point on the circumference of an imaginary circle that rotates clockwise with an angular frequency equal to the oscillator frequency ω. This circle is a pure fiction, but it's a useful construct for describing the motion of the oscillating mass — provided that we recognize that the y coordinate on the imaginary circle has no direct physical meaning. In the same way, it's sometimes useful to artificially promote real, physical quantities to complex quantities. There are certain equations which, when written in terms of these "complexified" quantities, often turn out to be much easier to solve. Of course the physical quantities we're interested in are the real ones rather than the complexified ones, so we must remember to go back and take the real part of the resulting expression at the end of the day.

For the case of simple harmonic motion, the analogy between this complexification procedure and the circular-motion construction runs even deeper. Indeed, if we promote the physical position coordinate $x(t) = A\cos(\omega t + \phi)$ to the complex quantity $z(t) = Ae^{i(\omega t + \phi)}$, Euler's formula tells us that

$$\mathrm{Re}[z(t)] \;=\; A\cos(\omega t + \phi) \;, \qquad \mathrm{Im}[z(t)] \;=\; A\sin(\omega t_\phi) \;. \tag{3.29}$$

We recognize that the imaginary part of $z(t)$ is the equation for the y coordinate of an object undergoing circular motion as a function of time. In other words, when we promote $x(t)$ to a complex coordinate $z(t)$, we are once again exploiting the circular-motion analogy, with the circular motion occurring in the complex plane.

3.4 Where Making Things Complex Makes Them Simple: AC Circuits

One situation in which complex quantities are very useful is in analyzing alternating-current (AC) circuits. To see how this works, consider the simple LRC circuit shown in Fig. 3.2 containing a potential source that supplies voltage $V_{\rm PS}$, a resistor with resistance R, a capacitor with capacitance C, and an inductor with inductance L. The voltage $V_{\rm PS} = V_0 \cos(\omega t)$ supplied by the potential source oscillates sinusoidally with an amplitude V_0 and angular frequency ω.

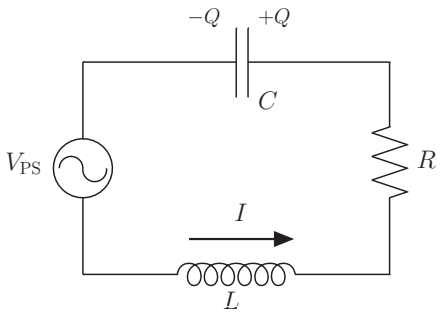


Figure 3.2: Diagram of an LRC circuit with a power source supplying a potential $V_{\rm PS}(t) = V_0 \cos(\omega t)$.

If we apply Kirchhoff's Second Law (the "loop rule") to this circuit, we have

$$V_0 \cos(\omega t) - L\frac{dI}{dt} - IR - \frac{Q}{C} \;=\; 0 \;, \tag{3.30}$$

where Q is the chage on the positive capacitor plate. In this simple circuit, $I = dQ/dt$, so this relation can be interpreted as a differential equation for Q:

$$V_0 \cos(\omega t) - L\frac{d^2Q}{dt^2} - R\frac{dQ}{dt} - \frac{Q}{C} \;=\; 0 \;. \tag{3.31}$$

This differential equation is far more complicated than the familiar simple-harmonic-oscillator equation. We won't attempt to derive a general solution to this equation yet (although we will do so later in this course). For the moment, we will merely demonstrate that there exists one particular solution to this equation of the form

$$Q(t) \;=\; \frac{I_0}{\omega}\sin(\omega t + \phi) \;, \tag{3.32}$$

where I_0 is a constant. We have labeled this constant I_0 because it represents the amplitude of the current — *i.e.*, the first derivative of Q:

$$\frac{dQ}{dt} \;=\; I \;=\; I_0 \cos(\omega t + \phi) \;. \tag{3.33}$$

The second derivative of Q is

$$\frac{d^2Q}{dt^2} \;=\; -\,I_0\omega \sin(\omega t + \phi) \;. \tag{3.34}$$

Plugging these expressions into Eq. (3.31), we find that

$$V_0 \cos(\omega t) + LI_0\omega \sin(\omega t + \phi) - RI_0 \cos(\omega t + \phi) - \frac{I_0}{\omega C}\sin(\omega t + \phi) \;=\; 0 \;. \tag{3.35}$$

The next step in showing that our proposed solution for $Q(t)$ in Eq. (3.32) indeed satisfies Eq. (3.31) would be to find values for I_0 and ϕ which make the relation in Eq. (3.35) true for all times t. This does not seem like particularly easy task, given that this relation includes both sine and cosine terms.

It turns out to be much easier to tackle this problem when we express things in terms of complex quantities. Euler's formula in Eq. (3.25) tells us that

$$\mathrm{Re}\big[e^{i\theta}\big] \;=\; \cos\theta \tag{3.36}$$

$$\mathrm{Re}\big[-ie^{i\theta}\big] \;=\; \mathrm{Re}\big[-i\cos\theta+\sin\theta\big] \;=\; \sin\theta\ . \tag{3.37}$$

for any real variable θ. We can therefore promote Eq. (3.35) to a complex equation by writing by substituting the corresponding exponential expressions in for $\sin(\omega t+\phi)$ and $\cos(\omega t+\phi)$:

$$V_0 e^{i\omega t} - iLI_0\omega e^{i(\omega t+\phi)} - RI_0 e^{i(\omega t+\phi)} + i\frac{I_0}{\omega C}e^{i(\omega t+\phi)} \;=\; 0\ . \tag{3.38}$$

We know from Sect. 3.1 that the real part of a sum of complex numbers is the same as the sum of the real parts. Therefore, when we take the real part of this equation, we'll get the same result as we would have obtained if we had tried to solve Eq. (3.35) for for I_0 and ϕ by brute force.

The first thing we notice about Eq. (3.38) is that the time dependence drops out, so we can write

$$V_0 - iL\omega I_0 e^{i\phi} - RI_0 e^{i\phi} + \frac{iI_0}{\omega C}e^{i\phi} \;=\; 0\ . \tag{3.39}$$

We also notice that I_0 and ϕ only appear in the particular combination

$$\tilde{I}_0 \;\equiv\; I_0 e^{i\phi}\ . \tag{3.40}$$

Note that I_0 and ϕ are real numbers (they're the amplitude and phase of the physical current, respectively), so we can interpret them as the modulus and complex phase of the complex parameter $\tilde{I}_0$. Moreover, when we write Eq. (3.39) in terms of $\tilde{I}_0$, it becomes even simpler:

$$V_0 - iL\omega\tilde{I}_0 - R\tilde{I}_0 + \frac{i}{\omega C}\tilde{I}_0 \;=\; 0\ . \tag{3.41}$$

Solving for $\tilde{I}_0$ in this equation is now straightforward. We find that

$$\tilde{I}_0 \;=\; \frac{V_0}{R+iL\omega-\frac{i}{\omega C}}\ . \tag{3.42}$$

We can put this complex expression in the standard $\tilde{I}_0 = \mathrm{Re}[\tilde{I}_0]+i\mathrm{Im}[\tilde{I}_0]$ form by multiplying both numerator and denominator by the complex conjugate of the denominator:

$$\tilde{I}_0 \;=\; \frac{V_0 R}{R^2+\big(L\omega-\frac{1}{\omega C}\big)^2} \;-\; i\,\frac{V_0\big(L\omega-\frac{1}{\omega C}\big)}{R^2+\big(L\omega-\frac{1}{\omega C}\big)^2}\ . \tag{3.43}$$

Finally, since I_0 and ϕ are just the modulus and the phase of $\tilde{I}_0$, we can use the relations in Eq. (3.19) to obtain

$$\begin{aligned} I_0 \;&=\; \sqrt{\mathrm{Re}[\tilde{I}_0]^2+\mathrm{Im}[\tilde{I}_0]^2} \;=\; \sqrt{\frac{V_0^2R^2+V_0^2\big(L\omega-\frac{1}{\omega C}\big)^2}{\Big[R^2+\big(L\omega-\frac{1}{\omega C}\big)^2\Big]^2}} \;=\; \frac{V_0}{\sqrt{R^2+\big(L\omega-\frac{1}{\omega C}\big)^2}} \\ \phi \;&=\; \arctan\left(\frac{\mathrm{Im}[\tilde{I}_0]}{\mathrm{Re}[\tilde{I}_0]}\right) \;=\; \arctan\left(-\frac{L\omega}{R}+\frac{1}{\omega RC}\right)\ . \end{aligned} \tag{3.44}$$

Let's step back a moment and reflect on what we have just done. We have shown that there indeed do exist values of the constants I_0 and ϕ for which our purported solution for $Q(t)$ in Eq. (3.32) satisfies the equation of motion for our LRC circuit. Thus, we have verified that indeed

$$Q(t) \;=\; \frac{V_0}{\sqrt{\omega^2R^2+\big(L\omega^2-\frac{1}{C}\big)^2}}\sin\left(\omega t+\arctan\left[-\frac{L\omega}{R}+\frac{1}{\omega RC}\right]\right) \tag{3.45}$$

is a solution to this Eq. (3.31), and the corresponding current in the circuit is

$$I(t) \;=\; \frac{V_0}{\sqrt{R^2 + \left(L\omega - \frac{1}{\omega C}\right)^2}} \cos\left(\omega t + \arctan\left[-\frac{L\omega}{R} + \frac{1}{\omega RC}\right]\right) \;, \tag{3.46}$$

3.5 Complex Impedances

This strategy for studying the dynamics of charges and currents in AC circuits by using complex quantities is an example of a more general technique called the method of **complex impedances**. This technique is applicable to a wide variety of circuits, including circuits which don't contain capacitors. For this reason, it is more convenient to express things in terms of the current $I(t)$ (which is relevant in any circuit) rather than in terms of the charge $Q(t)$ stored on the capacitor plates (which is clearly only relevant for circuits with capacitors).

First, we define the **complex current** in the following way:

$$\tilde{I} \;\equiv\; \tilde{I}_0 e^{i\omega t} \;=\; I_0 e^{i(\omega t + \phi)} \;. \tag{3.47}$$

The reason we define $\tilde{I}$ in this way is so that the real current I is just the real part of this expression:

$$I \;=\; \mathrm{Re}[\tilde{I}] \;=\; I_0 \cos(\omega t + \phi) \;. \tag{3.48}$$

Next, we define the **complex impedances**[3] for the resistors, capacitors, and inductors in the circuit in the following manner:

$$\begin{aligned} \text{Resistor:} \quad & Z_R \;=\; R \\ \text{Capacitor:} \quad & Z_C \;=\; \frac{-i}{\omega C} \\ \text{Inductor:} \quad & Z_I \;=\; i\omega L \;. \end{aligned} \tag{3.49}$$

Finally, we apply a generalized version of Kirchhoff's loop rule to the circuits in which the potential drop across a given circuit element E, where $E = \{R, C, I\}$, is given by

$$V_E \;=\; \tilde{I} Z_E \;. \tag{3.50}$$

Why does this work? The answer is that the product of $\tilde{I}_0$ with each of the complex impedances listed in Eq. (3.49) is really just the "complexified" version of the corresponding voltage drop IR for a resistors, Q/C for capacitors, and $L\,dI/dt$ for inductors. The additional factors of i and ω appearing in these expressions are proportionality constants that relate the complex impedance $\tilde{I}(t)$ to its time derivatives and integrals. Indeed, if the physical current $I(t)$ in a circuit as a function of time is described by Eq. (3.48), then we have

$$\begin{aligned} \frac{dI}{dt} \;&=\; -I_0\omega \sin(\omega t + \phi) \;=\; \mathrm{Re}\big[i\omega\tilde{I}\big] \;=\; \mathrm{Re}\left[\frac{d\tilde{I}}{dt}\right] \\ \int_0^t I dt' \;&=\; Q(t) \;=\; \frac{I_0}{\omega}\sin(\omega t + \phi) \;=\; \mathrm{Re}\left[-\frac{i}{\omega}\tilde{I}\right] \;=\; \mathrm{Re}\left[\int_0^t \tilde{I} dt'\right] \;. \end{aligned} \tag{3.51}$$

In other words, if we can assume that the real current $I(t)$ is described by Eq. (3.48), then both the derivative and the integral of the complex current $\tilde{I}(t)$ turn out to be proportional to $\tilde{I}(t)$ itself! Thus, while the real expression for the voltage drop associated with an inductor involves the derivative of I and the corresponding expression for a capacitor involves Q (the integral of I), the factors of i and ω in Eq. (3.49) take into account the effect of these operations.

[3] In general, the impedance of a circuit element is just the relationship between the current and the voltage across that circuit element.

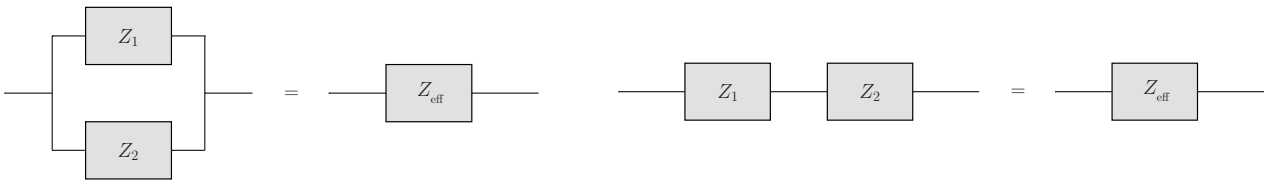


Figure 3.3: The effective complex impedance of a pair of complex impedances connected in series (left panel) and in parallel (right panel).

Since the expression in Eq. (3.50) for the complex impedance of any of these circuit elements has the same mathematical form as Ohm's Law, the rules for combining complex impedances are analogous to the rules for combining resistances. For example, the effective complex impedance for two complex impedances connected in series is

$$Z_{\rm eff} \;=\; Z_1 + Z_2 \;, \tag{3.52}$$

whereas the corresponding result for complex impedances connected in parallel is

$$\frac{1}{Z_{\rm eff}} \;=\; \frac{1}{Z_1} + \frac{1}{Z_2} \;. \tag{3.53}$$

Problems

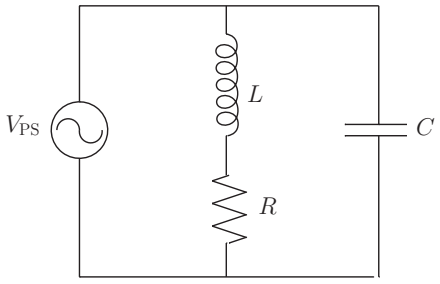


Figure 3.4: A circuit containing an inductor, a resistor, a capacitor, and an AC power source.

1. Find the effective complex impedance for the entire circuit shown in Fig. 3.4? If the voltage supplied by the power source is $V_{\rm PS} = V_0 \cos(\omega t)$, what is the (real) current through the resistor?

Chapter 4

Introduction to Differential Equations

- **The physics:** Projectile motion with damping
- **The math:** Differential equations, separation of variables, first-order linear differential equations, linear differential equations with constant coefficients, linear independence

4.1 Differential Equations

We're now going to begin discussing differential equations and techniques for solving them in more generality. We'll begin with some terminology that's useful for characterizing different kinds of differential equations.

- A differential equation is said to be **ordinary** if it has only one independent variable. For example, the familiar equation of motion for a simple harmonic oscillator

$$\frac{d^2x}{dt^2} \;=\; -\,\omega^2 x \;, \tag{4.1}$$

where ω is a constant, is an ordinary differential equation because t is the only independent variable on which the dependent variable $x(t)$ depends as it involves in time. Note that although $x(t)$ also depends on ω, this quantity is a **parameter** rather than an independent variable because it is constant in time. By contrast, an equation which contains more than one independent variable is called **partial**. An example of a partial differential equation is

$$\frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial f^2}{\partial z^2} \;=\; 0 \;. \tag{4.2}$$

This equation is called Poisson's equation and it appears frequently in many branches of physics, including fluid dynamics and electromagnetic theory. In this equation, the dependent variable $f(x, y, z)$ is a function of three independent variables: x, y, and z. In this class, we will focus primarily on ordinary differential equations.

- A differential equation is **linear** if if the only terms that appear in the equation are either independent or proportional to the dependent variable or one of its derivatives. If the equation contains powers or functions of the dependent variable or its derivatives, it is said to be **nonlinear**. For example, the equation

$$\frac{d^3x}{dt^3} + a\frac{dx}{dt} + bx + c \;=0 \;, \tag{4.3}$$

is a linear differential equation because each term is either independent of $x(t)$ or proportional to one of its derivatives. The simple-harmonic-oscillator equation in Eq. (4.1) is also linear. However, the equations

$$x\frac{d^2x}{dt^2} + a\frac{dx}{dt} \;=\; e^{-x^2/b^2} \tag{4.4}$$

and

$$\frac{dx}{dt} \;=\; -ax^2 \tag{4.5}$$

are both nonlinear: the first contains both x multiplied by its second derivative with respect to t and en exponential function involving x in its argument, while the second contains the square of x.

- The **order** of a differential equation is the order of the highest derivative in the equation. For example, the differential equation

$$\frac{d^3x}{dt^3} + x^2\frac{dx}{dt} + a\cos^2(bt) \;=\; 0\ , \tag{4.6}$$

is third order, while the equation

$$(1+x^2)\frac{dx}{dt} + x^3(a-x) \;=\; b^2x \tag{4.7}$$

is first order.

- A linear differential equation is **homogeneous**[1] if it contains no terms which do not involve the dependent variable. For example, the linear differential equation

$$\frac{d^2x}{dt^2} + ax^2 \;=\; 0 \tag{4.8}$$

is homogeneous because all of the terms involve x or its derivatives. By contrast, the related equation

$$\frac{d^2x}{dt^2} + ax^2 + bt^2 \;=\; 0 \tag{4.9}$$

is **inhomogeneous** because the term bt^2 does not involve the dependent variable $x(t)$.

A **solution** to a differential equation is a functional form for $x(t)$ for which the two sides of the equation are always equal, regardless of the value of the dependent variable or variables. For example, we have already seen that

$$x(t) \;=\; A\cos(\omega t + \phi) \tag{4.10}$$

is a solution to Eq. (4.1) regardless of the values of A and ϕ. We have also seen that

$$Q(t) \;=\; \frac{I_0}{\omega}\cos(\omega t + \phi) \tag{4.11}$$

is a solution to the differential equation

$$V_0\cos(\omega t) - L\frac{d^2Q}{dt^2} - R\frac{dQ}{dt} - \frac{Q}{C} \;=\; 0 \tag{4.12}$$

that one obtains from Kirchhoff's loop rule for an LRC circuit with an AC voltage source.

There is a useful theorem that exists regarding the solutions to linear differential equations. In particular, it can be proven that for any linear differential equation of order n, a solution can be constructed which contains n arbitrary parameters and from which any solution to the differential equation can be obtained by assigning particular values to those n parameters. Such a solution is called the **general solution** to the differential equation. For example, the solution $x(t)$ in Eq. (4.10) represents the general solution to the simple-harmonic-oscillator equation. This solution contains two arbitrary parameters A and ϕ, which is expected because the simple-harmonic-oscillator equation is second-order. This theorem is a powerful one, but one should keep in mind that it pertains specifically to *linear* differential equations — for nonlinear differential equations, it may not apply.

[1] While the definition presented here is typically what is meant when one uses the word "homogeneous" in relation to a differential equation, you should know that there is also a special class of first-order linear equations called "homogeneous differential equations." In this class, however, you may assume that the word is being used according to the definition presented here unless explicitly stated otherwise.

By contrast, the solution given in Eq. (4.11) is *not* the general solution to the LRC-circuit equation in Eq. (4.12). Although this equation is second order and although our solution for $Q(t)$ involves two parameters, we've already seen that these two parameters are not *arbitrary*. Indeed, we've already shown that their values are completely determined by the values of the physical parameters L, R, C, V_0, and ω using the method of complex impedances. Thus, we say that the solution for $Q(t)$ in Eq. (4.11) is a **particular solution** to Eq. (4.12). It certainly solves the differential equation, but there may be other solutions with different functional forms as well.

Solving differential equations is a much an art as it is a science. There is no universally applicable method for solving differential equations. However, there are a lot of tricks and techniques for solving particular kinds of differential equations. Let's now turn to examine some of these techniques and how they can be applied to solving problems in physics.

4.2 Separation of Variables

First-order differential equations are among the easiest differential equations to solve. We'll therefore begin by discussing some of the techniques for solving these equations, beginning with a particular class of differential equations called separable equations. A **separable** differential equation is one of the form

$$\frac{dy}{dx} = f(x)g(y) \ , \tag{4.13}$$

where x is the independent variable, $y(x)$ is the dependent variable, and $f(x)$ and $g(y)$ are arbitrary functions of x and y, respectively. Separable differential equations may be either linear or nonlinear and may be either homogeneous or inhomogeneous. The trick to solving a separable differential equation is to rearrange the equation so that the dependent variable y appears only on one side, while the independent variable x appears only on the other:

$$\frac{dy}{g(y)} = f(x)dx \ . \tag{4.14}$$

One can then integrate both sides of the equation:

$$\int \frac{dy}{g(y)} = \int f(x)dx \ . \tag{4.15}$$

The result is an equation which is no longer a differential equation, but rather just a regular equation without any derivatives. We can then simply solve this equation for $y(x)$ using the methods we would normally use to solve any ordinary, non-differential equation. This technique for solving first-order, separable differential equations is called **separation of variables**.

Example: Projectile Motion

To how separation of variables can be used to solve linear first-order differential equations, let's begin by considering a familiar example from mechanics: projectile motion. In particular, consider the motion of a ball that's tossed up into the air at time $t = 0$. The force of gravity acting on the ball is constant: $F = -mg$. Newton's Second Law tells us that the height of the ball y evolves in time according to by the equation

$$m\frac{d^2y}{dt^2} = -mg \ . \tag{4.16}$$

This differential equation for y is second order, so we can't use separation of variables to solve it. However, the equation for the velocity $v_y = dy/dt$ of the ball in the y direction *is* a first-order equation:

$$m\frac{dv_y}{dt} = -mg \ . \tag{4.17}$$

Moreover, this equations is indeed of the separable form given in Eq. (4.13), with $f(x)g(y)$ being a constant in this case, so we can use separation of variables to solve it. We first write the equation in the form

$$dv_y = -gdt \ , \tag{4.18}$$

and then integrate both sides

$$\int dv_y \;=\; -\,g\int dt\;, \tag{4.19}$$

to obtain

$$v_y(t) \;=\; -\,gt + C_1\;, \tag{4.20}$$

where C_1 is a constant of integration. This is our solution for $v_y(t)$. The constant C_1 is as yet undetermined, but we can understand what it means physically by noting that at $t = 0$ the velocity of the ball is $v_y(0) = C_1$. The constant C_1 therefore represents the initial velocity of the ball at $t = 0$.

Of course one can also take this calculation a step further. In particular, since $v_y = dy/dt$, we can substitute this solution back into the differential equation which relates v_y and y in order to obtain a differential equation for the height of the ball:

$$\frac{dy}{dt} \;=\; v_y(t) \;=\; -\,gt + v_0\;. \tag{4.21}$$

This differential equation is also separable. Indeed, in the language of Eq. (4.13), we have $f(t) = -gt + C_1$ and $g(y) = 1$. Thus, we can once again apply separation of variables to solve for $y(t)$. Rearranging the equation to get y on one side and t on the other gives

$$dy \;=\; v_y(t) \;=\; \big(-\,gt + v_0\big)dt\;. \tag{4.22}$$

Integrating both sides, we obtain

$$\int dy \;=\; y(t) \;=\; \int \big(-\,gt + v_0\big)dt \;=\; -\,\frac{1}{2}gt^2 + v_0t + C_2\;, \tag{4.23}$$

where C_2 is yet another integration constant. Once again, we can assess the physical significance of C_2 by observing that at $t = 0$, the height of the ball is $y(0) = C_2$. Thus, C_2 represents the initial position y_0 of the ball at $t = 0$. In other words, we have used separation of variables (twice) to show that

$$y(t) \;=\; -\,\frac{1}{2}gt^2 + v_0t + y_0\;. \tag{4.24}$$

This is the familiar ballistics equation — the equation which describes the trajectory of a projectile moving under the influence of gravity.

Example: Falling Motion with Quadratic Damping

We can actually use the method of separation of variables to extend our study of projectile motion to slightly more complicated but more realistic situations. For example, we can examine the effect of adding a drag force on the trajectory of the ball. In particular, let's consider adding a drag force which is proportional to the square of the velocity. Physically, a drag force always acts in the direction opposite the velocity, so the corresponding term in Newton's law would have to flip sign whenever v_y switched from negative to positive or vice versa. Let's avoid these issues for the moment by focusing on the portion of the ball's trajectory when its velocity is negative — *i.e.*, when it's falling. In this case, we can write $F_{\rm drag} = bv_y^2$. Thus, the differential equation for the vertical velocity v_y of the ball becomes

$$m\frac{dv_y}{dt} \;=\; -\,mg + bv_y^2\;. \tag{4.25}$$

This equation is also separable. Indeed, in the language of Eq. (4.13), we have $f(t) = 1$ and $g(v_y) = -mg + bv_y^2$, so rearranging gives us

$$\frac{mdv_y}{mg - bv_y^2} \;=\; -\,dt. \tag{4.26}$$

Integrating both sides gives us

$$m\int \frac{dv_y}{mg - bv_y^2} \;=\; -\int dt. \tag{4.27}$$

To perform the integral on the left-hand side of this equation, we begin by first rewriting the equation in a slightly more tractable form by multiplying both sides of the equation by mg to obtain

$$\int \frac{dv_y}{1-(v_y/v_T)^2} \;=\; -\,g\int dt\;, \tag{4.28}$$

where, for convenience, we have defined the constant parameter

$$v_T \;\equiv\; \sqrt{\frac{mg}{b}}\;. \tag{4.29}$$

As we shall see, this parameter represents the **terminal velocity** which v_y approaches as the ball falls. In addition, since we will want to choose our constants of integration such that the initial velocity of the ball is specified by a parameter v_0, we'll write the integral on each side of the equation as a definite integral:

$$\int_{v_0}^{v_y} \frac{dv'_y}{1-(v'_y/v_T)^2} \;=\; -\,g\int_0^t dt'\;. \tag{4.30}$$

This way, $v_y(0)=v_0$, and v_0 has the meaning that we want.

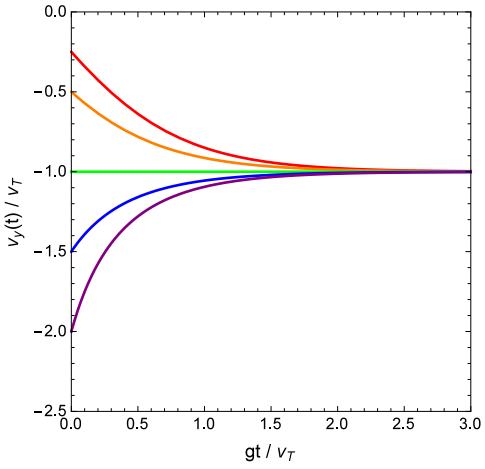


Figure 4.1: This plot shows how the vertical velocity v_y of a falling object with quadratic damping evolves with time. The curves shown each represent a different initial condition – *i.e.*, a different value of the initial velocity v_0, which corresponds to the value of the curve at $gt/v_T = 0$. Note that because the object is assumed to be falling (*i.e.*, the position y is assumed to be decreasing with time), we have $v_y(t) < 0$. Note alse that at $t\to\infty$, $v_y(t)\to v_T$ all of the curves shown.

We are now ready to perform the integral on the left-hand side of Eq. (4.30). We can do this by making the hyperbolic substitution

$$v'_y \;=\; v_T\tanh\eta\;, \qquad\qquad dv'_y \;=\; v_T\,\mathrm{sech}^2\eta\,d\eta\;. \tag{4.31}$$

This yields

$$\int_{\mathrm{arctanh}(v_0/v_T)}^{\mathrm{arctanh}(v_y/v_T)} \frac{v_T\,\mathrm{sech}^2\eta}{1-\tanh^2\eta} \;=\; \int_{\mathrm{arctanh}(v_0/v_T)}^{\mathrm{arctanh}(v_y/v_T)} d\eta \;=\; v_T\,\mathrm{arctanh}\left(\frac{v_y}{v_T}\right) - v_T\,\mathrm{arctanh}\left(\frac{v_0}{v_T}\right)\;, \tag{4.32}$$

so we have

$$v_T\,\mathrm{arctanh}\left(\frac{v_y}{v_T}\right) - v_T\,\mathrm{arctanh}\left(\frac{v_0}{v_T}\right) \;=\; -\,gt\;. \tag{4.33}$$

Finally, we solve this equation to obtain our solution for the velocity $v_y(t)$ of the falling ball:

$$v_y(t) \;=\; v_T \tanh\left[\operatorname{arctanh}\left(\frac{v_0}{v_T}\right) - \frac{gt}{v_T}\right] \;. \tag{4.34}$$

We can use the identity

$$\tanh(\alpha+\beta) \;=\; \frac{\tanh\alpha + \tanh\beta}{1+\tanh\alpha\tanh\beta} \tag{4.35}$$

and the fact that $\tanh(-\alpha) = -\tanh\alpha$ to rewrite this solution in the more revealing form

$$v_y(t) \;=\; v_T\,\frac{v_0 - v_T\tanh(gt/v_T)}{v_T - v_0\tanh(gt/v_T)} \;. \tag{4.36}$$

In Fig. 4.36, the way in which the velocity $v_y(t)$ evolves with time is shown for a variety of different values of the initial velocity v_0. Indeed, each different curve in the figure corresponds to a different value for v_0. Note that because the object is assumed to be falling (*i.e.*, the position y is assumed to be decreasing with time), we have $v_y(t) < 0$. Note also that at $t \to \infty$, $v_y(t) \to v_T$ all of the curves shown. This is what we mean when we say that v_T is a "terminal velocity": at very late times, $v_y(t)$ tends toward v_T regardless of the initial conditions.

Example: Radioactive Decay

In general, the rate at which a radioactive isotope decays is proportional to the number of atoms N of the isotope remaining in the sample. In other words, N is governed by a differential equation of the form

$$\frac{dN}{dt} \;=\; -\lambda N \;, \tag{4.37}$$

where λ is a proportionality constant with units $[\lambda] = 1/\text{s}$. This constant is called **decay constant** for the isotope. This equation is separable and may be rearranged to give

$$\frac{dN}{N} \;=\; -\lambda dt \;. \tag{4.38}$$

The next step is to integrate both sides of this equation. However, as with the last example (projectile motion with quadratic damping), it is useful to stop and consider what our boundary conditions are so that we can write the constants of integration in a physically meaningful way right from the start, before we integrate. In this case, our boundary condition is that at $t = 0$, we begin with a certain number of radioactive atoms $N(0) = N_0$ in our sample. Thus, we incorporate our boundary condition by integrating the left-hand side of this equation from N_0 to N and the right-hand side over from 0 to t:

$$\int_{N_0}^{N} \frac{dN'}{N'} \;=\; -\int_0^t \lambda dt' \;. \tag{4.39}$$

This gives us the regular, non-differential equation

$$\ln(N) - \ln(N_0) \;=\; \ln\left(\frac{N}{N_0}\right) \;=\; -\lambda t \;, \tag{4.40}$$

which we can then solve for $N(t)$. Doing so, we find that

$$N(t) \;=\; N_0 e^{-\lambda t} \;. \tag{4.41}$$

This is the expected result: the number of radioactive atoms $N(t)$ in the sample decays exponentially with t.

The constant λ sets the time scale for the decay. Indeed, we see the see that inverse quantity 1λ plays a role in this solution which is very similar to the role that the time constant $\tau = RC$ plays in setting the time

scale to which an RC circuit charges or discharges. In fact, the **half-life** $t_{1/2}$ of the isotope — *i.e.*, the value of t for which $N(t) = N_0/2$ — is proportional to $1/\lambda$. To see this, we simply set $N(t) = N_0/2$ in Eq. (4.41):

$$\frac{N}{N_0} \;=\; \frac{1}{2} \;=\; e^{-\lambda t_{1/2}} \;, \tag{4.42}$$

which yields

$$\lambda \;=\; -\frac{1}{t_{1/2}} \ln\left(\frac{1}{2}\right) \;=\; \frac{\ln(2)}{t_{1/2}} \;. \tag{4.43}$$

4.3 First-Order Linear Differential Equations

While separation of variables is a useful technique when we can apply it, not all first-order differential equations are separable. However, there are a number of other useful techniques for solving different types of first-order differential equations. For example, there is a completely general method for solving *linear* first-order differential equations.

A first-order linear differential equation can always be written in the form

$$\frac{dy}{dx} + P(x)y \;=\; Q(x) \;, \tag{4.44}$$

where $P(x)$ and $Q(x)$ are arbitrary functions of the independent variable x. Note that according to our definition from Sect. 4.1, an equation of this form is homogeneous if $Q(x) = 0$ and inhomogeneous otherwise. As we shall see, this distinction will turns out to be quite important in

Let's begin with the homogeneous case where $Q(x) = 0$ and Eq. (4.44) reduces to

$$\frac{dy}{dx} + P(x)y \;=\; 0. \tag{4.45}$$

We actually already know how to solve this equation. Regardless of what form the function $P(x)$ takes, we observe that the equation is separable. Indeed, it has precisely the form indicated in Eq. (4.13), with $f(x) = -P(x)$ and $g(y) = y$. Thus, applying separation of variables on this equation and integrating, we get

$$\int \frac{dy}{y} \;=\; -\int P(x)dx \;. \tag{4.46}$$

When we perform the integral on each side of this equation, we pick up an overall integration constant, which we'll call C:

$$\ln y \;=\; -\int P(x)dx + C \;. \tag{4.47}$$

The solution to Eq. (4.45) for any function $P(x)$ is therefore

$$y(x) \;=\; e^{-\int P(x)dx + C} \;. \tag{4.48}$$

This is in fact the *general* solution for $y(x)$.

It's important to emphasize that at this stage, the value of the integration constant C is undetermined. You can verify for yourself (using the fundamental theorem of calculus) that the solution for $y(x)$ in Eq. (4.48) satisfies Eq. (4.45) regardless of the value of C. Indeed, just like the amplitude A and the phase ϕ that appear in general solution to the simple-harmonic-oscillator equation, its value cannot be specified without additional information — information that comes in the form of the boundary conditions which characterize the particular problem you happen to be dealing with. The solution in Eq. (4.48) is in fact the general solution to Eq. (4.45). As appropriate for a first-order differential equation, this general solution contains one undetermined constant. However, there are frequently many ways of writing the undetermined constants in a differential equation. Some of these turn out to be more useful or revealing than others. For example, in this case we shall soon see that it is useful to write Eq. (4.48) in the form

$$y(x) \;=\; Ae^{-\int P(x)dx} \;, \tag{4.49}$$

where $A = e^C$ plays the role of our undetermined constant.

We are now ready to generalize this technique for solving homogeneous first-order linear differential equations to solve inhomogeneous differential equations as well. In order to do so, we begin by introducing a new piece of terminology. For any inhomogeneous linear differential equation, let us define the **complementary equation** to be the homogeneous differential equation that you would obtain by removing all of the terms that do not involve the dependent variable or its derivatives. For example, the complementary equation to any inhomogeneous differential equation of the form appearing in Eq. (4.44) would be the corresponding equation of the form appearing in Eq. (4.45) — *i.e.*, the equation without the $Q(x)$ term.

Why is the complementary equation relevant for solving the full, inhomogeneous differential equation? The answer is that there's a trick we can use in order to construct a solution $y(x)$ to the full equation by exploiting the properties of the solution $y_c(x)$ to the complementary equation. Since the complementary equation is a homogeneous first-order linear differential, equation, we already know that $y_c(x)$ has the form given in Eq. (4.48):

$$y_c(x) \;=\; Ae^{-I(x)} \;, \tag{4.50}$$

where for convenience we have used the shorthand

$$I(x) \;=\; \int P(x)dx \tag{4.51}$$

for the integral over the function $P(x)$. Multiplying both sides of this equation by $e^I(x)$ yields

$$y_c(x)e^{I(x)} \;=\; A \;. \tag{4.52}$$

Now let's examine what happens when we take the derivative of both sides of this equation. Since A is a constant, we have

$$\frac{d}{dx}\Big[y_c(x)e^{I(x)}\Big] \;=\; e^{I(x)}\frac{d}{dx}y_c(x) + y_c(x)e^{I(x)}\frac{d}{dx}I(x) \;=\; 0. \tag{4.53}$$

Moreover, the fundamental theorem of calculus tells us that

$$\frac{d}{dx}I(x) \;=\; \frac{d}{dx}\int P(x)dx \;=\; P(x) \;, \tag{4.54}$$

so we learn from Eq. (4.53) that

$$\frac{d}{dx}y_c(x) + P(x)y_c(x) \;=\; 0 \;. \tag{4.55}$$

It turns out that we can use the result in Eq. (4.55) to generate a solution to the full, inhomogeneous differential equation. The trick is to look for solutions for $y(x)$ of the form

$$y(x) \;=\; u(x)y_c(x) \tag{4.56}$$

where $u(x)$ is some unknown function of x.[2] Why would we want to write $y(x)$ this way? Well, if we plug this candidate solution into Eq. (4.44), we obtain

$$\begin{aligned} Q(x) \;&=\; \frac{d}{dx}y(x) + P(x)y(x) \\ &=\; y_c(x)\frac{d}{dx}u(x) + u(x)\frac{d}{dx}y_c(x) + P(x)u(x)y_c(x) \\ &=\; y_c(x)\frac{d}{dx}u(x) + u(x)\left[\frac{d}{dx}y_c(x) + P(x)y_c(x)\right] \;. \end{aligned} \tag{4.57}$$

However, Eq. (4.55) tells us that the quantity in brackets mush vanish. Thus, we find that

$$y_c(x)\frac{d}{dx}u(x) \;=\; Q(x) \;, \tag{4.58}$$

[2]It should be emphasized that we're not pulling a solution to Eq. (4.44) out of thin air here. We haven't said anything about the functional form of $u(x)$, so this is merely a *rewriting* of $y(x)$, not a guess or ansatz as to what the solution might look like look like.

which is a differential equation for $u(x)$. This equation is separable and can we can solve it in the standard manner. The solution is:

$$u(x) \;=\; \int \frac{Q(x)}{y_c(x)}dx + C \;=\; \int \frac{1}{A}Q(x)e^{I(x)}dx + C, \tag{4.59}$$

where once again C is an integration constant. Plugging this result for $u(x)$ back into Eq. (4.56), we arrive at the final form of our solution to Eq. (4.44):

$$\begin{aligned} y(x) \;=\; u(x)y_c(x) \;&=\; Ae^{-I(x)}\int \frac{1}{A}Q(x)e^{I(x)}dx + CAe^{-I(x)} \\ &=\; e^{-I(x)}\int Q(x)e^{I(x)}dx + Be^{-I(x)}\ , \end{aligned} \tag{4.60}$$

where we have written our undetermined constant in the form $B \equiv CA$. This represents the general solution for any inhomogeneous first-order linear differential equation, and it involves a single undetermined constant, as we would expect.

Example: RC Circuit

As an example of how this general technique for solving first-order linear differential equations can be applied, consider the RC circuit pictured in Fig 4.2. This circuit includes a switch and an AC power source which supplies a voltage $V_{\rm PS} = V_0\cos(\omega t)$. At some time $t = t_0$, the switch is closed and current begins to flow within the circuit. We are principally interested in how the charge $Q(t)$ on the capacitor evolves in time. We recall that the differential equation we get for $Q(t)$ from Kirchhoff's loop rule is

$$V_0\cos(\omega t) - \frac{Q}{C} - R\frac{dQ}{dt} \;=\; 0\ , \tag{4.61}$$

where we have used the fact that the current through the resistor is given by $I = dQ/dt$. While we have analyzed circuits of this form using other methods (for example, the method of complex impedances) these methods furnished us with one particular solution to the above equation. We will now attempt to derive the *general* solution to this equation.

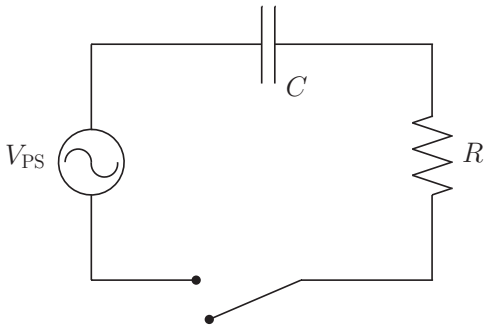


Figure 4.2: An RC circuit with an AC power source and a switch in the open position.

The first step in deriving a general solution for $Q(t)$ is to recognize that the equation in Eq. (4.61) is an inhomogeneous first-order linear differential equation for Q(t). Indeed, if we rewrite it in the form

$$\frac{dQ}{dt} + \frac{Q}{RC} \;=\; \frac{V_0}{R}\cos(\omega t)\ , \tag{4.62}$$

we can readily equate the functions $P(x)$ and $Q(x)$ in Eq. (4.44) with $1/RC$ and $(V_0/R)\cos(\omega t)$, respectively. We therefore know that the general solution is of the form given in Eq. (4.60).

The next step in deriving this general solution out is to evaluate $I(t)$, by which we mean the integral that corresponds to Eq. (4.51). Since the quantity $1/RC$ which plays the role of $P(x)$ is a constant in this case, this is easily done:

$$I(t) \;=\; \int \frac{dt}{RC} \;=\; \frac{t}{RC}\;. \tag{4.63}$$

Plugging this into the general solution, we have

$$Q(t) \;=\; e^{-t/RC}\int \frac{V_0}{R}\cos(\omega t)e^{t/RC}dx + Be^{-t/RC}\;. \tag{4.64}$$

We can evaluate this integral by using Euler's formula to express the cosine function in terms of exponentials:

$$\begin{aligned}
\int \cos(\omega t)e^{t/RC}dx &= \frac{1}{2}\int \left(e^{t/RC+i\omega t} + e^{t/RC-i\omega t}\right)dx \\
&= \frac{1}{2}\left[\frac{1}{(RC)^{-1}+i\omega}e^{t/RC+i\omega t} + \frac{1}{(RC)^{-1}-i\omega}e^{t/RC-i\omega t}\right] \\
&= \frac{1}{2}e^{t/RC}\left[\frac{RC - i\omega(RC)^2}{1+(\omega RC)^2}e^{i\omega t} + \frac{RC + i\omega(RC)^2}{1+(\omega RC)^2}e^{-i\omega t}\right] \\
&= \frac{RC}{1+(\omega RC)^2}e^{t/RC}\Big[\cos(\omega t) + \omega RC\sin(\omega t)\Big]\;.
\end{aligned} \tag{4.65}$$

Substituting this result into Eq. (4.64) gives us

$$Q(t) \;=\; \frac{V_0 C}{1+(\omega RC)^2}\Big[\cos(\omega t) + \omega RC\sin(\omega t)\Big] + Be^{-t/RC}\;. \tag{4.66}$$

This is our general solution for $Q(t)$, so in principle we have what we've been looking for. However, it's also nice to rearrange this expression slightly so the the physical interpretation of this solution is a bit more clear. In order to do this, we begin by defining the quantity $\phi \equiv \arctan(\omega RC)$ and using it to rewrite the quantity in square brackets in Eq. (4.66) as follows:

$$\begin{aligned}
\cos(\omega t) + \omega RC\sin(\omega t) &= \cos(\omega t) + \tan\phi\sin(\omega t) \\
&= \frac{1}{\cos\phi}\big[\cos\phi\cos(\omega t) + \sin\phi\sin(\omega t)\big] \\
&= \frac{\cos(\omega t - \phi)}{\cos\phi}\;.
\end{aligned} \tag{4.67}$$

We can then use the trigonometric identity

$$\cos(\arctan x) \;=\; \frac{1}{\sqrt{1+x^2}} \tag{4.68}$$

to write this expression in the form

$$\cos(\omega t) + \omega RC\sin(\omega t) \;=\; \sqrt{1+(\omega RC)^2}\,\cos\big[\omega t - \arctan(\omega RC)\big]\;. \tag{4.69}$$

This may not seem like much of an improvement, but when we substitute this result beck into Eq. (4.64), the resulting expression has some pieces that we recognize from earlier in the course:

$$Q(t) \;=\; \frac{V_0 C}{\sqrt{1+(\omega RC)^2}}\cos\big[\omega t - \arctan(\omega RC)\big] + Be^{-t/RC}\;. \tag{4.70}$$

Indeed, the first term on the left-hand side is exactly the form of the solution for $Q(t)$ that we obtained for the same RC circuit using the method of complex impedances! What's new is the second piece, which has the form of a decaying exponential. We also see that the quantity ϕ that we defined above is related to the phase shift of the current in the circuit.

In order to understand the physical meaning behind this solution for $Q(t)$ at a deeper level, it's useful to examine how his solution behaves in certain limiting cases. First, let's consider what happens at very late times, when $t \gg RC$. In the regime, the term with the decaying exponential rapidly dies away and the first term — the familiar one from our study of complex impedances — is the only term that matters. Because $Q(t)$ always reduces to this solution when t becomes sufficiently large, it's typically referred to as the **steady-state solution** for the system. It's also worth noting that this steady-state solution is independent of the initial conditions which characterize the system, which enter into the general solution for $Q(t)$ through the undetermined constant B. By contrast, the decaying-exponential term in Eq, (4.66) is an example of a **transient solution** — a term which depends sensitively on the initial conditions and only contributes significantly at early times.

In addition, let's also consider what happens to our solution for $Q(t)$ for very low frequencies $\omega \ll RC$. Physically, this means that the potential supplied by the power source is very nearly constant on short timescales — *i.e.*, for t much smaller than the period T. In this limit, we recover another familiar result. In particular, we see that Eq. (4.66) reduces to

$$Q(t) \stackrel{\omega \to 0}{\longrightarrow} V_0 C \cos(0) + B e^{-t/RC} = V_0 C + B e^{-t/RC} . \tag{4.71}$$

To see the connection between this expression and a more familiar result we've already seen in lab, let's specify a value for B by explicitly applying some initial conditions. In particular, if the capacitor is initially uncharged — *i.e.*, if we specify $Q(0) = 0$ as our boundary condition — we find that

$$B = -V_0 C , \tag{4.72}$$

and thus we have

$$Q(t) \approx V_0 C\big(1 - e^{-t/RC}\big) . \tag{4.73}$$

We recognize this as the formula for the charge on a charging capacitor in a DC circuit. This is indeed what we might have expected: for extremely low frequencies the voltage supplied by the power source is effectively constant, so it functions like a DC circuit for $t \ll T$.

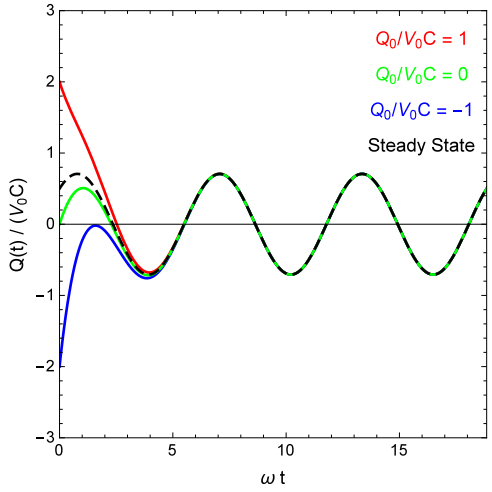


Figure 4.3: The charge $Q(t)$ on the capacitor (normalized to the value of $V_0 C$) in the RC circuit pictured in Fig. 4.2 as a function of ωt for several different choices of the initial charge Q_0. The steady-state solution (the black dashed curve) is also shown. For all of the curves shown, we have taken $\omega RC = 1$. We observe that all of the solutions indeed settle into the steady-state behavior at sufficiently late times.

Finally, let's examine how we can apply the general solution in Eq. (4.70) to find particular solutions to specific problems with specific boundary conditions. This amounts to choosing the appropriate value of B

which corresponds to whatever those boundary conditions happen to be. For example, we might start with a particular initial charge Q_0 on the capacitor at $t = 0$. Setting $Q(0) = Q_0$ in Eq. (4.70) yields

$$Q_0 \;=\; \frac{V_0 C}{\sqrt{1+(\omega RC)^2}}\,\cos\big[\arctan(\omega RC)\big] + B\;. \tag{4.74}$$

We can then solve this equation for B to determine how this undetermined constant is related to Q_0. When we do so — and use the trigonometric identity in Eq. (4.68) to write the result in a simpler, more compact form — we find that

$$B \;=\; Q_0 - \frac{V_0 C}{1+(\omega RC)^2}\;. \tag{4.75}$$

Substituting this expression for B back into Eq. (4.70) gives us the corresponding expression for $Q(t)$:

$$Q(t) \;=\; \frac{V_0 C}{\sqrt{1+(\omega RC)^2}}\,\cos\big[\omega t - \arctan(\omega RC)\big] + \left[Q_0 - \frac{V_0 C}{1+(\omega RC)^2}\right] e^{-t/RC}\;. \tag{4.76}$$

We emphasize that this is just another way of writing the general solution for $Q(t)$. Indeed, we still have one undetermined constant — we've just traded B for Q_0. However, the nice thing about writing the general solution in this form is that Q_0 has a more direct physical interpretation than B does.

In Fig. 4.3, we plot the ratio of $Q(t)$ to the parameter combination $V_0 C$ at a function of ωt for a variety of different choices of the initial charge Q_0 on the capacitor. For each curve shown, we have taken $\omega RC = 1$. Also shown is a black dashed curve which corresponds to the steady-state solution. At early times, the transient term — which depends sensitively on the choice of boundary conditions — is important, and the curves for different Q_0 look very different. By contrast, at late times, the curves all settle into the steady-state solution, regardless on Q_0, as we'd anticipated.

4.4 General Solutions from Solutions to the Complementary Equation

The method we used to obtain the solution in Eq. (4.60) is applicable to any first-order linear differential equation. We emphasize again that this solution represents the general solution to an equation of this sort: any function which solves such a first-order linear differential equation can be written in this form, and different choices of initial conditions simply correspond to different assignments of the constant B.

It is worth mentioning another technique which can be used for obtaining the general solution to an inhomogeneous linear differential equation. While this technique is not as universally applicable as the method for solving first-order linear differential equations outlined in Sect. 4.3, it has the advantage of being applicable not only just to first-order equations, but to inhomogeneous linear differential equations of *any* order. In particular, it can be shown[3] that every solution to a linear differential equation of arbitrary order may be written in the form

$$y(x) \;=\; y_c(x) + y_p(x)\;, \tag{4.77}$$

where $y_c(x)$ is the general solution to the complementary equation and $y_p(x)$ is *any* particular solution to the full equation. Of course, since we already have an explicit method for calculating the general solution to any first-order linear differential equation, this method is far more useful for dealing with higher-order differential equations. It's important to emphasize, however, that this method is applicable only for solving *linear* differential equations. For or nonlinear equations, Eq. (4.77) does not necessarily hold true.

Example: Mass Suspended from a Spring

As an example of wow Eq. (4.77) can be used, let's apply it to the familiar problem of a mass m suspended vertically from a spring with spring constant k. Both gravity and the spring force act on the mass, so the

[3]For a proof of the corresponding theorem, see, *e.g.*, Section 3.6 of Boyce, William E. and DiPrima, Richard C., *Elementary Differential Equations*, Wiley 1992.

equation of motion we get from from Newton's Second Law for the extension y of the spring is

$$m\frac{d^2y}{dt^2} \;=\; -ky - mg \;. \tag{4.78}$$

This is a second-order linear differential equation, and it's inhomogeneous because the gravitational term doesn't involve the dependent variable y. The complementary equation, which we obtain by dropping this inhomogeneous term, is just the simple-harmonic-oscillator equation:

$$m\frac{d^2y_c}{dt^2} \;=\; -ky_c \;. \tag{4.79}$$

We already know that the general solution to this equation is

$$y_c(t) \;=\; A\cos(\omega t + \phi) \;, \qquad \omega \;=\; \sqrt{\frac{k}{m}} \;. \tag{4.80}$$

Therefore, all that we need to do in order to obtain a solution to the full equation in Eq. (4.78) to find a particular solution — *any* particular solution — that satisfies that equation.

Since any particular solution will do, we'll work with the simplest possible solution we can find. (This is typically a good strategy whenever you're applying this technique.) One trivial solution that satisfies Eq. (4.78) is the solution for which the right-hand side vanishes and we just have $dy/dt = 0$ for all time. Physically, this corresponds to the mass just hanging there undisturbed at the equilibrium point and not oscillating at all. Our particular solution is therefore

$$y_p(t) \;=\; -\frac{mg}{k} \;. \tag{4.81}$$

Like any particular solution to a differential equation, this solution involves a particular choice of boundary conditions. In particular it corresponds to an initial position $y(0) = y_{\rm eq} = -mg/k$ and an initial velocity $v_y(0) = 0$. It's admittedly a thoroughly uninteresting solution in a lot of ways, but it's sufficient to allow us to construct a general solution to the full inhomogeneous differential equation in Eq. (4.78):

$$y(t) \;=\; y_c(t) + y_p(t) \;=\; A\cos(\omega t + \phi) - \frac{mg}{k} \;. \tag{4.82}$$

This was exactly the expression we obtained in Chapter 1 of these Lecture Notes using other methods.

Problems

1. For each of the following ordinary differential equations, identify the order of the equation and state whether it is linear or nonlinear. If it is linear, identify whether it homogeneous or inhomogeneous. Note that in each equation, a, b, and c are assumed to be constant parameters.

   (a) $\displaystyle\frac{d^2x}{dt^2} - ax\cos(bt) \;=\; c$

   (b) $\displaystyle y^2\frac{dy}{dx} \;=\; -2cy(y-b)$

   (c) $\displaystyle e^{-t/a}\frac{dy}{dt} - 4b\frac{dy^2}{dt^2} \;=\; cy$

   (d) $\displaystyle\frac{d^3y}{dx^3} \;=\; 2cy$

   (e) $\displaystyle\frac{d^2f}{dx^2} \;=\; af + bf^2 + cf^3$

(f) $$\left[1 - a\frac{dx}{dt}\right]^2 = bx - ct^2$$

2. Use the identities in Sect. 2.12 of Boas, Mary L., *Mathematical Methods in the Physical Sciences*, Wiley 2005 to derive the identity for $\tanh(\alpha + \beta)$ in Eq. (4.35) under the assumption that α and β are real numbers.

3. Consider a spherical droplet which is evaporating at a rate proportional to its surface area.

$$\frac{dV}{dt} = -kA \ , \tag{4.83}$$

where V is the volume of the droplet, A is it's surface area, and k is a constant with units $[k] = \text{m/s}$.

   (a) Solve for $V(t)$ for a droplet which begins at $t = 0$ with an initial volume $V(0) = V_0$.

   (b) If half of the droplet has disappeared after 1 day, how many additional days must you wait before the entire droplet will have disappeared?

4. In one of the examples in Sect. 4.2, we examined the motion of a falling object acted upon by a quadratic drag force. Let us now examine the motion of an object acted upon by a *linear* drag force $F_{\text{drag}} = cv_y$, where c is a constant with units $[c] = \text{kg/s}$. Such a form for the drag force is appropriate for slow, small objects, like the oil droplets in Millikan's classic experiment for measuring the charge of the electron. Find the solution for the velocity $v_y(t)$ in the y direction for such an object, assuming that it is released from rest — *i.e.*, that $v_y(0) = 0$ — and that the only forces acting on the object are gravity and the linear drag force described above.

5. Consider the RL circuit shown in Fig. 4.4. The switch is closed at time t = 0. Write down a differential equation for the current $I(x)$ using Kirchhoff's loop rule Solve this equation for $I(t)$, and plot the voltage drop across the resistor as a function of time.

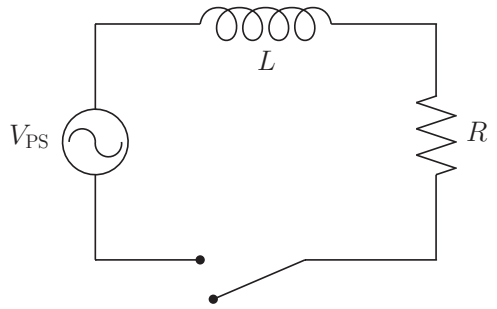


Figure 4.4: An RL circuit with an AC power source and a switch in the open position.

6. Find the general solution to the differential equation

$$x\frac{dy}{dx} + 2y = \sin x \ . \tag{4.84}$$

Find the value of the constant in your solution for the boundary condition $y(\pi/2) = 0$.

Chapter 5

Second-Order Differential Equations and Damped Oscillations

- **The physics:** Decaying oscillations, quality factors, frictional damping, underdamped and overdamped oscillators
- **The math:** Second-order homogeneous linear differential equations, linear independence, reduction of order, piecewise solutions to differential equations

5.1 Second-Order Homogeneous Linear Differential Equations

Now that we have explored a number of useful techniques for solving first-order differential equations, let us now turn to discuss second-order differential equations. Such equations crop up almost everywhere in physics, in part due to the fact that Newton's Second Law inherently involves acceleration — the second derivative of position with respect to time. As one might expect, dealing with second-order differential equations — even linear ones — can be a much more involved process than dealing with first-order equations, and the range of phenomena they display can be far richer, In this section, we will develop a number of methods for solving particular types of *linear* second-order differential equations. Nonlinear second-order differential equations are often far more complicated.

The general form for a second-order linear differential equation is

$$P(x)\frac{dy^2}{dx^2} + Q(x)\frac{dy}{dx} + R(x)y \;=\; G(x)\;, \tag{5.1}$$

where $P(x)$, $Q(x)$, $R(x)$, and $G(x)$ are functions of the independent variable x. We will begin by restricting our attention to homogeneous equations — *i.e.*, equations for which $G(x) = 0$. In other words, we will focus on equations of form

$$P(x)\frac{dy^2}{dx^2} + Q(x)\frac{dy}{dx} + R(x)y \;=\; 0\;. \tag{5.2}$$

First, before we begin discussing some of the techniques that apply to specific kinds of homogeneous second-order linear differential equations, it is worth calling attention to a generic property which the general solution to such equations share. In particular, there exists a theorem [1] which states that the the general solution to such a differential equation can always be written in the form

$$y(x) \;=\; C_1 y_1(x) + C_2 y_2(x)\;, \tag{5.3}$$

where C_1 and C_2 are constants and $y_1(x)$ and $y_2(x)$ are a pair of **linearly-independent** solutions to the equation. Formally speaking, two functions $y_1(x)$ and $y_2(x)$ are linearly independent if the only solution to the equation

$$a y_1(x) + b y_2(x) \;=\; 0 \qquad \text{for all } x \tag{5.4}$$

[1] For a proof of this see, *e.g.*, Sect. 3.2 of Boyce, William E. and DiPrima, Richard C., *Elementary Differential Equations*, Wiley 1992.

for constant parameters a and b is $a = b = 0$. Practically speaking, this basically means that the functions $y_1(x)$ and $y_2(x)$ are not simply multiples of each other or merely different ways or writing the same function. One can verify that two functions $y_1(x)$ and $y_2(x)$ are linearly independent by calculating a quantity called the **Wronksian**, given by

$$W(x) \;=\; y_1(x)\frac{d}{dx}y_2(x) - y_2(x)\frac{d}{dx}y_1(x) \;. \tag{5.5}$$

If the Wronskian for is not identically zero for all x, then these functions are linearly independent.

The fact that the general solution to any homogeneous second-order linear differential equation can be written in the form given in Eq. (5.3) is actually quite powerful. Essentially, it means that all you need to do in order to obtain the general solution to such an equation is find any two distinct (*i.e.*, linearly-independent) functions $y_1(x)$ and $y_2(x)$ that satisfy the equation. The general solution $y(x)$ is simply a linear combination of these functions with arbitrary coefficients. We are now going to show that once you know one solution $y_1(x)$, you can always find the second solution $y_2(x)$ using techniques that we learned for solving first-order differential equations.

In passing, however, it's worth remarking that we have already seen one example in which the general solution to a homogeneous second-order linear differential equation can be expressed in the form given in Eq. (5.3). The simple-harmonic-oscillator equation is a homogeneous second-order linear differential equation, and we have shown that the general solution can be written as

$$x(t) \;=\; C_1\cos(\omega t) + C_2\sin(\omega t) \;. \tag{5.6}$$

Indeed, as you will demonstrate in Problem 1, $\cos(\omega t)$ and $\sin(\omega t)$ are linearly independent.

5.2 Reduction of Order

At the end of the previous section we alluded to a method which would allow you to use one solution $y_1(x)$ to a second-order differential equation to derive the other, linearly-independent solution. The method is called **reduction of order**, and it can always be used to find a second solution $y_2(x)$ to a homogeneous second-order differential equation if you know the first. The trick is to write $y_2(x)$ in the form

$$y_2(x) \;=\; u(x)y_1(x) \;, \tag{5.7}$$

where $u(x)$ is some arbitrary function of x. Since we haven't specified anything about the form of $u(x)$, this is just a rewriting of $y_2(x)$.

When we plug this expression for $y_2(x)$ into Eq. (5.2), we find that

$$\begin{aligned}
0 \;&=\; P(x)\frac{d^2}{dx^2}(uy_1) + Q(x)\frac{d}{dx}(uy_1) + R(x)(uy_1) \\
&=\; P(x)\left[y_1\frac{d^2u}{dx^2} + 2\frac{du}{dx}\frac{dy_1}{dx} + u\frac{d^2y_1}{dx^2}\right] + Q(x)\left[y_1\frac{du}{dx} + u\frac{dy_1}{dx}\right] + R(x)uy_1 \\
&=\; u\left[P(x)\frac{d^2y_1}{dx^2} + Q(x)\frac{dy_1}{dx} + R(x)y_1\right] + P(x)y_1\frac{d^2u}{dx^2} + \left[2P(x)\frac{dy_1}{dx} + Q(x)y_1\right]\frac{du}{dx} \;.
\end{aligned} \tag{5.8}$$

The first term on the right-hand side of this equation must be equal to zero because we already know that $y_1(x)$ is a solution to Eq. (5.2), from which is follows that

$$P(x)\frac{d^2y_1}{dx^2} + Q(x)\frac{dy_1}{dx} + R(x)y_1 \;=\; 0 \;. \tag{5.9}$$

We therefore have

$$P(x)y_1\frac{d^2u}{dx^2} + \left[2P(x)\frac{dy_1}{dx} + Q(x)y_1\right]\frac{du}{dx} \;=\; 0 \;. \tag{5.10}$$

Since we're assuming that we already know the functional form of $y_1(x)$ (and therefore its derivatives with respect to x as well), all that remains is for us to solve this differential equation for $u(x)$ and we'll have our solution for $y_2(x)$.

This equation in Eq. (5.10) is yet another second-order homogeneous differential equation. However, it does not contain any terms proportional to $u(x)$ — only terms proportional to it derivatives. This means that we define

$$v(x) \equiv \frac{d}{dx}u(x) , \tag{5.11}$$

we can cast Eq. (5.10) as a first-order differential equation for $v(x)$:

$$P(x)y_1\frac{dv}{dx} + \left[2P(x)\frac{dy_1}{dx} + Q(x)y_1\right]v = 0 . \tag{5.12}$$

This equation is separable, so we can solve it for $v(x)$ using separation of variables. We can then plug the result into Eq. (5.11) and solve for $u(x)$. Finally, we can substitute this result into Eq. (5.7) to obtain the second solution $y_2(x)$ to our original differential equation.

5.3 Finding Roots: Equations with Constant Coefficients

One class of homogeneous second-order linear differential equations which can be solved a straightforward method is the class of equations for which the functions $P(x)$, $Q(x)$, and $R(x)$ appearing in Eq. (5.2) are constants — *i.e.*, equations of the form

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0 , \tag{5.13}$$

where a, b, and c are constant parameters. Equations of this form arise quite commonly in physics. One canonical example, which we shall study in more detail below, is the case of a harmonic oscillator subject a velocity-dependent damping force.

Our first step in solving Eq. (5.13) will simply be to use our intuition about functions and their derivatives to guide us as to what functional form the solution might be likely to take. For example, we know that the derivative of the exponential function e^x is just the function itself. Therefore, for any constant parameter r, the chain rule tells us that

$$\frac{d}{dx}e^{rx} = re^{rx} , \tag{5.14}$$

and taking further derivatives of this expression will just result in our original function being multiplied by additional factors of r. This means that if we substitute a function of the form $y(x) = e^{rx}$ into Eq. (5.13), we get

$$ar^2e^{rx} + bre^{rx} + ce^{rx} = \big(ar^2 + br + c\big)e^{rx} = 0 . \tag{5.15}$$

This equation will be satisfied identically for all x if the quantity in parentheses vanishes — *i.e.*, when r is a solution the quadratic equation

$$ar^2 + br + c = 0 . \tag{5.16}$$

This quadratic equation is known as the **characteristic equation** for the differential equation in Eq. (5.13). In other words, we have shown that the solutions to Eq. (5.13) are just exponential functions of the form

$$y_\pm(x) = e^{r_\pm x} , \tag{5.17}$$

where $r_\pm$ are the two solutions to Eq. (5.16) given by the quadratic formula

$$r_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{5.18}$$

As long as $b^2 \neq 4ac$, these two solutions will be linearly independent.[2] Therefore, according to Eq. (5.3), we can construct the general solution to Eq. (5.13) by adding them together with arbitrary coefficients:

$$y(x) = C_1e^{r_+x} + C_2e^{r_-x} . \tag{5.19}$$

[2]The special case in which $b = 4ac$ is considered in Problem 4.

5.4 The Damped Harmonic Oscillator

One of the most important applications of the results we derived in the previous section arises in the context of physical systems with dissipative forces — *i.e.*, forces which cause physical systems to lose energy to their surroundings. Nature is full of dissipative forces. Air resistance and friction are two familiar examples. Joule heating (the dissipation of energy as heat by resistors) plays a analogous role in electric circuits to the role these forces play in mechanics. Our first-hand experience tells us that a real mass suspended from from a spring does not continue to oscillate forever. Rather, the amplitude of its oscillations decreases over time until the mass effectively comes to a standstill.

One example of a dissipative force is the **drag force** which acts to slow the motion of an object moving through a fluid such as air or water. In general, the drag force on a moving body is a complicated non-linear function of the velocity v. However, just as we saw that the restoring force in many physical systems is approximately proportional to the displacement $\Delta x = x - x_0$ from equilibrium when that displacement is sufficiently small (the Harmonic approximation), it likewise turns out that for small velocities, the drag force on an object moving through a fluid can likewise be approximated as linear in the velocity v:

$$F_{\rm drag} \;\approx\; -bv \;. \tag{5.20}$$

This is just the first term in a Maclaurin-series expansion for $F_{\rm drag}$. For more rapidly moving objects, the drag force will contain terms involving higher powers of v, and such terms often actually dominate over the linear term shown in Eq. (5.20). For example, the drag force on a soccer ball moving through the air is dominated by the quadratic term $F_{\rm drag} = -cv^2$.

Let us now consider the case of a mass on a spring which is subject to a linear damping force of the form given in Eq. (5.20). Since $v = dx/dt$, the equation of motion for the mass is

$$m\frac{d^2x}{dt^2} \;=\; -kx - b\frac{dx}{dt} \;. \tag{5.21}$$

By defining the parameter combinations $\omega_0 \equiv \sqrt{k/m}$ and $\beta \equiv b/2m$, we can recast this expression in the form

$$\frac{d^2x}{dt^2} + 2\beta\frac{dx}{dt} + \omega_0^2 x \;=\; 0 \;. \tag{5.22}$$

Note that angular frequency ω_0 correspond to the familiar angular frequency of oscillation ω for the simple harmonic oscillator. Since β and ω are constants, this equation has precisely the form given in Eq. (5.13), with[3] $a = 1$, $b = 2\beta$, and $c = \omega_0^2$. Thus, we know that the solutions are

$$x(t) \;=\; C_1 e^{r_+ t} + C_2 e^{r_- t} \;, \tag{5.23}$$

where the roots $r_\pm$ of the characteristic equation are

$$r_\pm \;=\; -\beta \pm \sqrt{\beta^2 - \omega_0^2} \;. \tag{5.24}$$

In principle, that's it. We're done. The solution in Eq. (5.23) represents the most general solution to the equation of motion for a linearly-damped harmonic oscillator. Applying this result to any particular physical system is simply a matter of specify our boundary conditions and using those boundary conditions to solve for the undetermined constants C_1 and C_2. However, as simple as the solution in Eq. (5.23) may seem at first glance, it can give rise to an astonishing variety of possible behaviors, depending on the relationship between the parameters ω and β. It is therefore worth exploring the properties of the solutions to this equation in more detail.

5.5 Underdamping, Overdamping, and Critical Damping

It turns out that there are three different kinds of behavior which can arise from our general solution for $x(t)$ in Eq. (5.23). Which kind of behavior occurs in a particular physical system depends on the relationship between the parameters β and ω_0. We now examine each of these three cases in turn.

[3] Make sure not to confuse the quantity b appearing in Eq. (5.13) with the drag coefficient in Eq (5.20).

5.5.1 Underdamped Motion

The first of these three cases we'll examine is the one which bears the most resemblance to the simple harmonic oscillator. This is the case in which $\omega_0 > \beta$. In this case, the expressions in Eq. (5.24) for the roots $r_\pm$ of the characteristic equation are complex. In order to see what this means physically, let's separate out the real and imaginary parts of these roots and write them in rectangular form:

$$r_\pm \;=\; -\beta \pm i\omega_1 \;, \tag{5.25}$$

where we have defined the real parameter

$$\omega_1 \;\equiv\; = \sqrt{\omega_0^2 - \beta^2} \;. \tag{5.26}$$

This case is commonly referred to as the **underdamped** case. We shall soon see that ω_1 represents the oscillation frequency of the solution.

When we plug the expressions for $r_\pm$ in Eq. (5.25) back into the general solution for $x(t)$ in Eq. (5.23), we get

$$\begin{aligned} x(t) &= e^{-\beta t}\left[C_1 e^{i\omega_1 t} + C_2 e^{-i\omega_1 t}\right] \\ &= e^{-\beta t}\left[(C_1 + C_2)\cos(\omega_1 t) + i\,(C_1 - C_2)\sin(\omega_1 t)\right] \;, \end{aligned} \tag{5.27}$$

where we have used Euler's theorem to write the complex exponentials in the first line in terms of sines and cosines. You might be concerned that the imaginary number appears in this expression, since the position $x(t)$ is a real physical quantity and must therefore be a purely real-valued function. However, this requirement just represents a constraint on the values of the undetermined constants C_1 and C_2. Indeed, for our solution to make any physical sense, we must require that that $C_1 - C_2$ be purely imaginary — or, in other words, we must have that $C_2 = C_1^*$. This condition also implies that $C_1 + C_2$ is be purely real. Thus, we can rewrite Eq. (5.27) in terms of a pair of purely real constants

$$B_1 \;\equiv\; C_1 + C_2 \;, \qquad B_2 \;\equiv\; i\,(C_1 - C_2) \;, \tag{5.28}$$

which gives us

$$x(t) \;=\; e^{-\beta t}\left[B_1 \cos(\omega_1 t) + B_2 \sin(\omega_1 t)\right] \;. \tag{5.29}$$

We can write this expression for $x(t)$ in an even more compact form using the same trigonometric identities we used to simplify our general solution for $Q(t)$ in the RC-circuit example from Sect. 4.3. In particular, we find that we can write our general solution for $x(t)$ as

$$x(t) \;=\; Ae^{-\beta t}\cos(\omega_1 t + \phi) \;, \tag{5.30}$$

where the two undetermined constants A and ϕ are related to B_1 and B_2 in the following manner:

$$A \;=\; \sqrt{B_1^2 + B_2^2} \;, \qquad \phi \;=\; -\arctan\left(\frac{B_2}{B_1}\right) \;. \tag{5.31}$$

When we write it in this form, the physical interpretation of our general solution for an underdamped oscillator becomes much more transparent. In particular, the solution looks quite a bit like the general solution for the simple harmonic oscillator. However, there is very one important difference. We have seen that the amplitude of oscillation for the general solution to the simple-harmonic-oscillator equation is constant in time — a result that ultimately stems from the fact that energy is conserved within the system. By contrast, the amplitude

$$A(t) \;=\; Ae^{-\beta t} \tag{5.32}$$

for the underdamped-oscillator solution in Eq. (5.30) is *not* constant; rather, it includes a falling exponential that suppresses the amplitude for late times. This reflects the fact that energy is not conserved within this system. The drag force removes mechanical energy from the system over time by transferring it to the surroundings (*e.g.*, as heat). As a result, the amplitude of oscillation diminishes over time. The quantity β, which has units $[t] = 1/\text{s}$, represents the rate at which the amplitude function $A(t)$ decreases. In particular,

we see that $1/\beta$ is the time it takes for $A(t)$ to decrease to $1/e$ of its original value. For this reason, β is often referred to as the **decay parameter** for underdamped harmonic motion:

$$\Gamma_{\rm dec} \;=\; \beta \qquad \text{(underdamped case)} \; . \tag{5.33}$$

Since the the undetermined coefficients in the general solution for $x(t)$ are at this point simply arbitrary numbers, it would be good to examine how these coefficients are related to quantities that have a direct physical interpretation — quantities of the sort that we might use in setting up boundary conditions for a given oscillator problem. We'll focus on the parametrization in Eq. (5.29), in which the the undetermined coefficients are B_1 and B_2. One potentially relevant set of physical quantities which we could use in establishing boundary conditions are the initial position x_0 and the initial velocity v_0 of the oscillator at $t = 0$. Indeed, we often used these same quantities that when specifying our boundary condition for the simple harmonic oscillator.

Imposing the boundary condition $x(0) = x_0$ on our general solution in Eq. (5.29) for the damped harmonic oscillator is relatively straightforward. We find that

$$x_0 \;=\; x(0) \;=\; e^0 \left[B_1 \cos(0) + B_2 \sin(0) \right] \;=\; B_1 \; . \tag{5.34}$$

In other words, the value of the coefficient B_1 is apparently just the initial position x_0 of the oscillator. Imposing the initial condition $v(0) = v_0$ is simply a matter of taking the derivative of $x(t)$ with respect to t and setting the result equal to v_0 at $t = 0$. The derivative is

$$\begin{aligned} \frac{dx}{dt} \;&=\; -\beta e^{-\beta t} \left[B_1 \cos(\omega_1 t) + B_2 \sin(\omega_1 t) \right] + e^{-\beta t} \left[-B_1 \omega_1 \sin(\omega_1 t) + B_2 \omega_1 \cos(\omega_1 t) \right] \\ &=\; e^{-\beta t} \Big[(-B_1 \beta + B_2 \omega_1) \cos(\omega_1 t) - (B_1 \omega_1 + B_2 \beta) \sin(\omega_1 t) \Big] \; , \end{aligned} \tag{5.35}$$

and so we have

$$v_0 \;=\; \left. \frac{dx}{dt} \right|_{t=0} \;=\; e^0 \Big[(-B_1 \beta + B_2 \omega_1) \cos(0) - (B_1 \omega_1 + B_2 \beta) \sin(0) \Big] \;=\; -B_1 \beta + B_2 \omega_1 \; . \tag{5.36}$$

Since we already know that $B_1 = x_0$, we can solve this equation for B_2 in terms of x_0 and v_0. Doing so, we obtain our set of relations between the coefficients B_1 and B_2 and the initial values x_0 and v_0:

$$B_1 = x_0 \; , \qquad B_2 \;=\; \frac{v_0 + \beta x_0}{\omega_1} \; . \tag{5.37}$$

The relations in Eq. (5.37) allow us to specify $x(t)$ as a function of time for any combination of x_0 and v_0. Some example solutions of the form given in Eq. (5.29) for the case of underdamped motion are illustrated in Fig. 5.2. Once again, the four different panels shown in the figure correspond to different choices of the initial position x_0 and velocity v_0 of the oscillator. Indeed, these solutions are what we might have expected: in each case, $x(t)$ oscillates around its equilibrium point at $x_{\rm eq} = 0$ with an amplitude that diminishes over time.

Because there are two timescales involved in underdamped harmonic motion — the period of oscillation and the timescale on which the amplitude of oscillation decays away — it is often of interest to see how these two timescales compare to one another. Information about the ratio of these timescales is typically expressed in terms of a quantity known as the **quality factor**:

$$Q \;\equiv\; \frac{\omega_0}{2\beta} \; . \tag{5.38}$$

The reason for the factor of 2 in this expression — as well as the reason for choosing ω_0 rather than the physical frequency of oscillation ω_1 — may not be at all obvious at this point. However, as we shall soon see, these factors are included to give Q a direct interpretation in terms of how the energy of the system decreases over time.

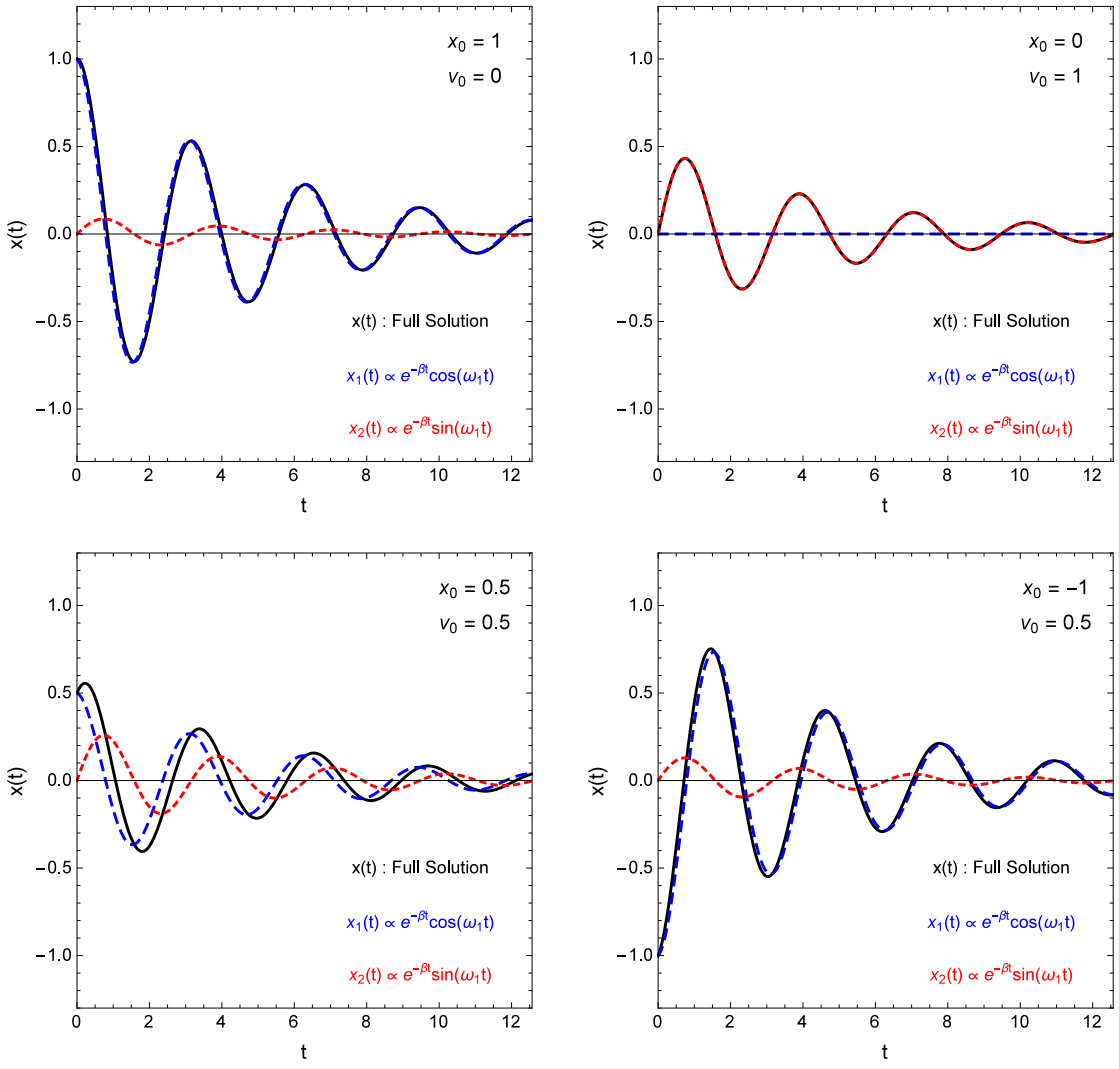


Figure 5.1: Each panel in this figure shows the full solution $x(t)$ for underdamped motion (solid black curve) for a different combination of the initial position x_0 and initial velocity v_0 (in arbitrary units). The results in each panel correspond to the parameter choices $\omega_0 = 2$ and $\beta = 0.2$ (in the same, arbitrary units). The contributions from the two individual solutions $x_1(t)$ (blue long-dashed curve) and $x_2(t)$ (red short-dashed curve) are also indicated in each figure. The full solution $x(t)$ is the sum of these two individual contributions.

5.5.2 Overdamped Motion

Let us now turn to the second kind of behavior that can be realized from the general solution to the damped-harmonic-oscillator equation. This is the behavior that arises in the case in which $\omega_0 < \beta$. In this case, the expressions in Eq. (5.24) for $r_\pm$ are purely real. Thus, the general solution takes the form

$$\begin{aligned} x(t) &= C_1 e^{-\beta t + \sqrt{\beta^2 - \omega_0^2}\,t} + C_2 e^{-\beta t - \sqrt{\beta^2 - \omega_0^2}\,t} \\ &= e^{-\beta t}\left[C_1 e^{\sqrt{\beta^2 - \omega_0^2}\,t} + C_2 e^{-\sqrt{\beta^2 - \omega_0^2}\,t}\right] , \end{aligned} \tag{5.39}$$

where the undetermined coefficients C_1 and C_2 in this case are purely real. It's important to note that we always have $\beta > \sqrt{\beta^2 - \omega_0^2}$ in this case, so both of the terms in Eq. (5.39) are actually decaying exponentials, and both asymptotically tend toward zero as $t \to \infty$. This solution is called the **overdamped** solution by contrast with the underdamped solution in Eq. (5.30). It is also frequently expressed in terms of the hyperbolic sine and cosine functions

$$x(t) = e^{-\beta t}\left[B_1 \cosh\left(\sqrt{\beta^2 - \omega_0^2}\,t\right) + B_2 \sinh\left(\sqrt{\beta^2 - \omega_0^2}\,t\right)\right] , \tag{5.40}$$

where the coefficients B_1 and B_2 are related to the coefficients C_1 and C_2 as follows:

$$B_1 = C_1 + C_2 , \qquad B_2 = C_2 - C_1 . \tag{5.41}$$

We can make contact between the arbitrary coefficients B_1 and B_2 in Eq. (5.40) and the respective initial values x_0 and v_0 for the position and velocity for the case of overdamped motion in much the same way as we did for the underdamped case. Proceeding in essentially the same way that we did in Sect. 5.5.1, we find that

$$B_1 = x_0 , \qquad B_2 = \frac{v_0 + \beta x_0}{\sqrt{\beta^2 - \omega_0^2}} . \tag{5.42}$$

Some example solutions of the form given in Eq. (5.39) for the case of overdamped motion are illustrated in Fig. 5.2. Once again, the four different panels shown in the figure correspond to different choices of the initial position x_0 and velocity v_0 of the oscillator. However, you can see that calling the system an "oscillator" at all is a bit of a misnomer. Indeed, for $\beta > \omega_0$, our solution for $x(t)$ never even manages to complete a single oscillation — or even half an oscillation — but rather tends asymptotically toward the equilibrium point.

Although the two terms in Eq. (5.39) both decay away to zero as $t \to \infty$, we observe that the second term decays away more quickly than the first; thus, the late-time behavior of $x(t)$ is dominated by the first term. If we wanted to define a decay parameter for this overdamped solution, we would therefore focus on the decay behavior of this first term and write

$$\Gamma_{\rm dec} = \beta - \sqrt{\beta^2 - \omega_0^2} \qquad \text{(overdamped case)} . \tag{5.43}$$

We note that when $\beta \gg \omega$, we can use a Taylor series to approximate the second term. Doing so, we find that

$$\Gamma_{\rm dec} = \beta - \beta\sqrt{1 - \frac{\omega_0^2}{\beta^2}} = \beta - \beta\left[1 + \frac{1}{2}\frac{\omega^2}{\beta^2} - \frac{1}{8}\frac{\omega^2}{\beta^2} - \frac{1}{16}\frac{\omega^2}{\beta^2} + \ldots\right] \approx \frac{\omega_0^2}{2\beta} ,$$

where we have assumed that ω_0 is sufficiently small compared to β that we are justified in keeping only the leading term in the expansion. This is actually highlights an interesting difference between the underdamped and overdamped cases. For underdamped motion, the rate at which the $x(t)$ decays with time is governed solely by the damping parameter β. By contrast, for overdamped motion, both β and ω_0 play a critical role in determining this rate. Moreover, for a fixed value of ω_0, we see that increasing β actually makes this rate *smaller*, which is the opposite of what happens in the underdamped case.

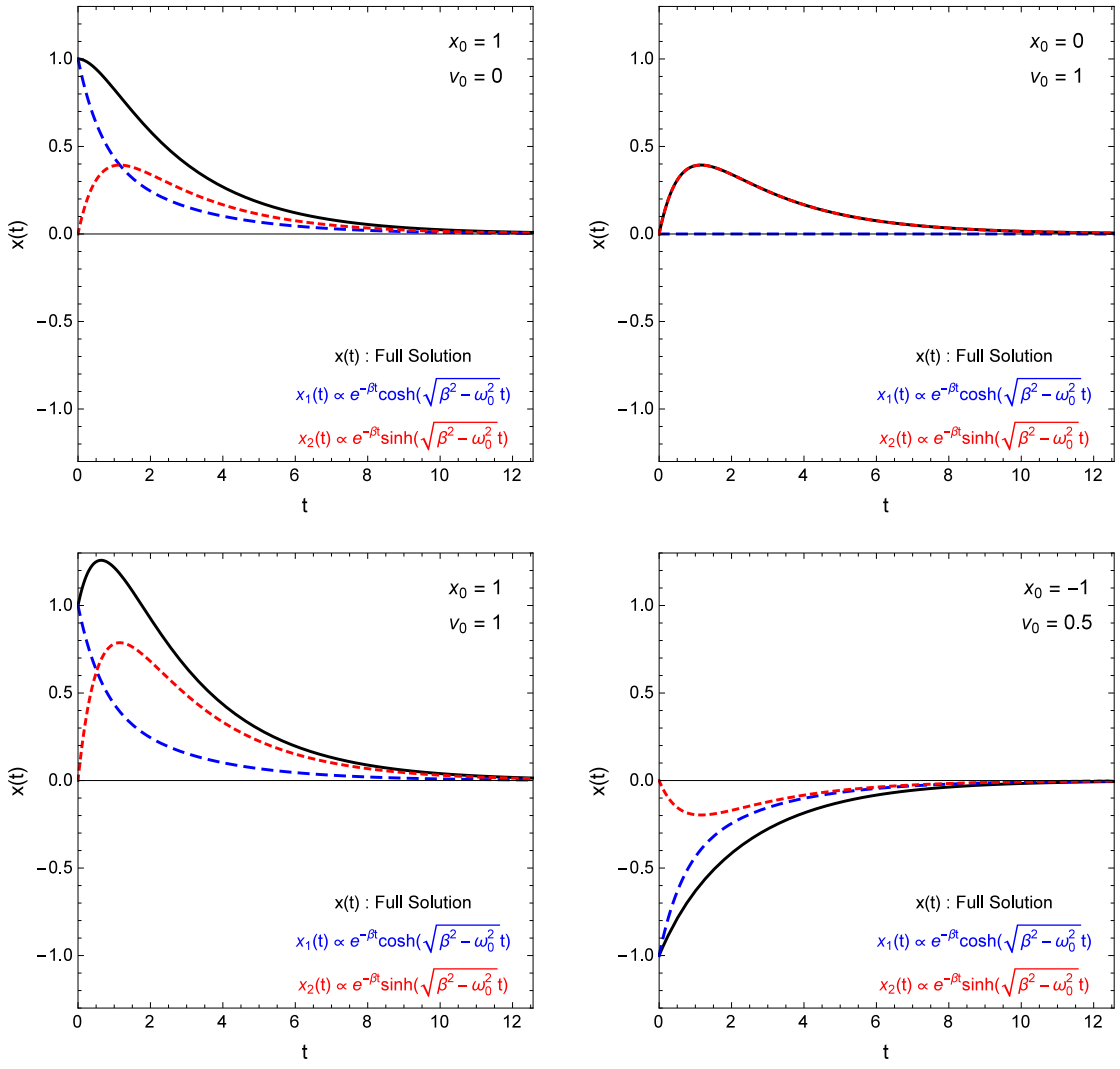


Figure 5.2: Each panel in this figure shows the full solution $x(t)$ for overdamped motion (solid black curve) for a different combination of the initial position x_0 and initial velocity v_0 (in arbitrary units). The results in each panel correspond to the parameter choices $\omega_0 = 0.8$ and $\beta = 1$ (in the same, arbitrary units). The contributions from the two individual solutions $x_1(t)$ (blue long-dashed curve) and $x_2(t)$ (red short-dashed curve) are also indicated in each figure. The full solution $x(t)$ is the sum of these two individual contributions.

5.5.3 Critically-Damped Motion

So far, we have considered the behavior of the general solution $x(t)$ to the-damped-oscillator equation for the case in which $\beta < \omega_0$ (which corresponds to underdamped motion) and the case in which $\beta > \omega_0$ (which corresponds to overdamped motion). There is, however, a third case which we have not yet considered. This is the special case in which $\beta = \omega_0$. In this case, the roots $r_\pm$ of the characteristic equation are degenerate and both given by

$$r_+ \;=\; r_- \;=\; -\,\beta\;. \tag{5.44}$$

We therefore know that one of the solutions is given by

$$x_1(t) \;=\; e^{-\beta t}\;, \tag{5.45}$$

but we still need a second linearly-independent solution in order to construct the general solution for $x(t)$. However, as you have already demonstrated in Problem 4, the second solution has the form

$$x_2(t) \;=\; te^{-\beta t}\;. \tag{5.46}$$

Thus, the full solution is

$$x(t) \;=\; \big(B_1 + B_2 t\big)e^{-\beta t}\;, \tag{5.47}$$

where the undetermined constants B_1 and B_2 are once again purely real. This special case is referred to as **critically-damped** motion.

Once again, we can relate the coefficients B_1 and B_2 in Eq. (5.47) to x_0 and v_0 for critically-damped motion in the same way as we did for the underdamped and overdamped cases. Doing so, we find that

$$B_1 \;=\; x_0\;, \qquad B_2 \;=\; v_0 + \beta x_0\;. \tag{5.48}$$

Some example solutions of the form given in Eq. (5.47) for the case of critically-damped motion are illustrated in Fig. 5.3. Once again, the four different panels shown in the figure correspond to different choices of the initial position x_0 and velocity v_0 of the oscillator. These figures demonstrate that indeed, just as in the overdamped case, the solutions for the critically-damped case do not oscillate, but rather asymptotically approach zero as $t \to \infty$.

5.6 The Energetics of Damped Harmonic Motion

The total energy $E_{\rm tot}$ of the system is once again the sum of the kinetic energy of the oscillator and the potential energy stored in the spring:

$$E_{\rm tot} \;=\; \frac{1}{2}mv^2 + \frac{1}{2}kx^2\;. \tag{5.49}$$

Since the drag force is not a conservative force, there is no potential-energy term associated with this force. Rather, this force serves to dissipate the energy stored in the system to its surroundings. Thus, for the damped harmonic oscillator, $E_{\rm tot}$ is a (decreasing) function of time.

Once again, let's begin with the underdamped case, which bears the most resemblance to the simple harmonic oscillator. Computing $E_{\rm tot}$ is simply a matter of plugging our results for $x(t)$ and $v(t)$ for this case into Eq. (5.49). It's easiest to work with the parametrization for $x(t)$ in Eq. (5.30). In this parametrization, the velocity takes the form

$$\begin{aligned} v(t) \;&=\; -A\beta e^{-\beta t}\cos(\omega_1 t + \phi) - A\omega_1 e^{-\beta t}\sin(\omega_1 t + \phi) \\ &=\; -A\omega_0 e^{-\beta t}\cos\left[\omega_1 t + \phi - \arctan\left(\frac{\omega_1}{\beta}\right)\right]\;. \end{aligned} \tag{5.50}$$

where we have used a few trigonometric identities from our "standard toolbox" to simplify this result (see Problem 8). The total energy of an underdamped oscillator system is therefore

$$\begin{aligned} E_{\rm tot} \;&=\; \frac{1}{2}kA^2e^{-2\beta t}\left\{\cos^2(\omega_1 t + \phi) + \cos^2\left[\omega_1 t + \phi - \arctan\left(\frac{\omega_1}{\beta}\right)\right]\right\} \\ &=\; \frac{1}{2}kA^2e^{-2\beta t}\left\{1 + \frac{\beta}{\omega_0}\cos\left[2\omega_1 t + 2\phi - \arctan\left(\frac{\omega_1}{\beta}\right)\}\right]\right)\;, \end{aligned} \tag{5.51}$$

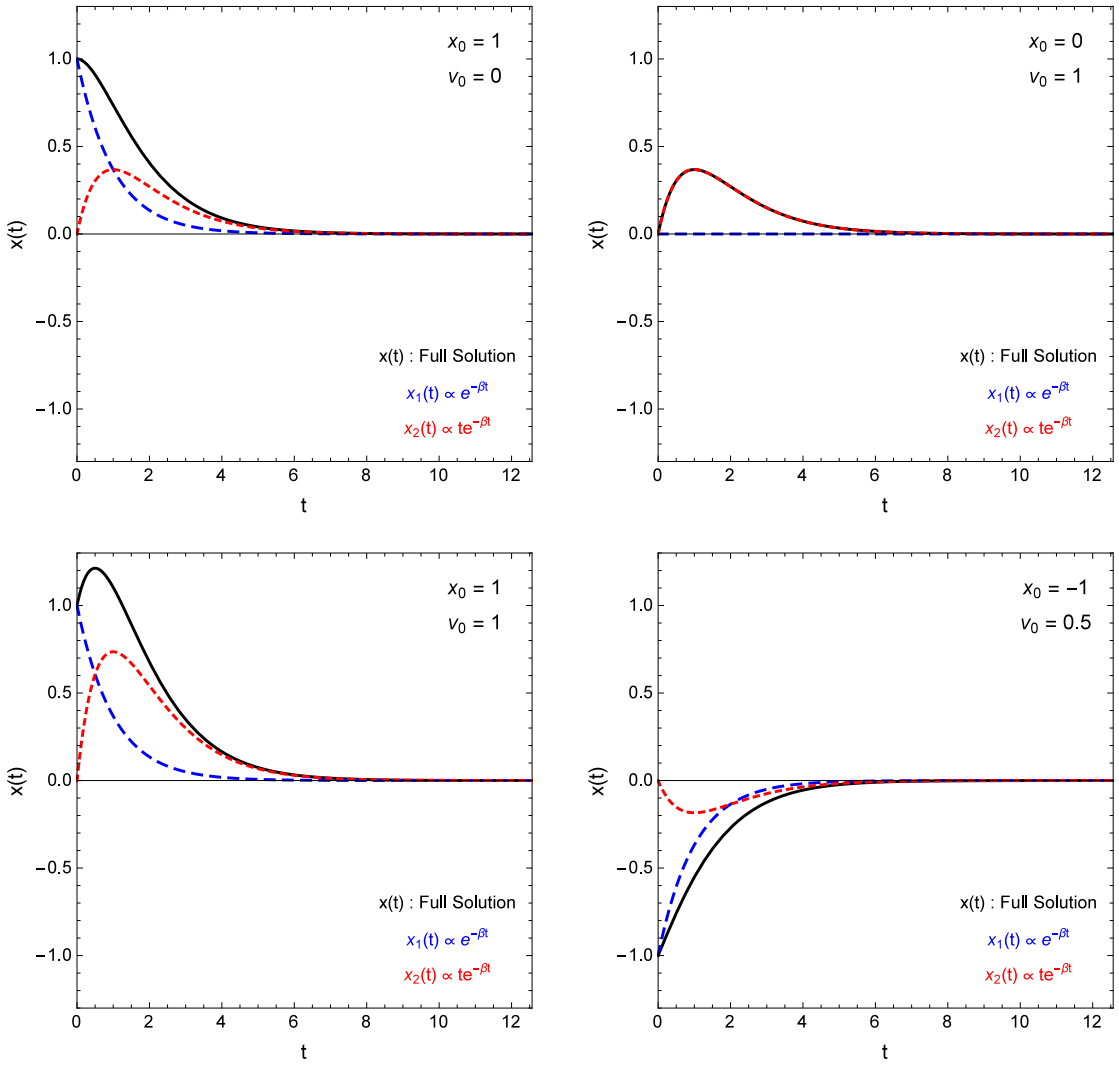


Figure 5.3: Each panel in this figure shows the full solution $x(t)$ for critically-damped motion (solid black curve) for a different combination of the initial position x_0 and initial velocity v_0 (in arbitrary units). The results in each panel correspond to the parameter choice $\omega_0 = \beta 1$ (in the same, arbitrary units). The contributions from the two individual solutions $x_1(t)$ (blue long-dashed curve) and $x_2(t)$ (red short-dashed curve) are also indicated in each figure. The full solution $x(t)$ is the sum of these two individual contributions.

where we have once again used trigonometric identities to simplify the result (see Problem 9).

This result may seem a bit bizarre at first glance. We might have anticipated that the expression for $E_{\rm tot}$ would contain an exponential suppression factor which causes the energy of the system to decrease over time as the system gives up energy to its surroundings. Indeed, for the simple harmonic oscillator, we found that

$$E_{\rm tot}^{\rm (SHO)} \;=\; \frac{1}{2}kA^2 \;, \tag{5.52}$$

and so it makes sense that since the amplitude $A(t)$ for an underdamped oscillator falls off exponentially with time, it makes sense that the energy would be suppressed by a factor of $A^2(t)$. However, the fact that there is an *additional* time-dependence to the expression in Eq. (5.51) — a time-dependence which comes in the form of a cosine term no less — is certainly less obvious *a priori*.

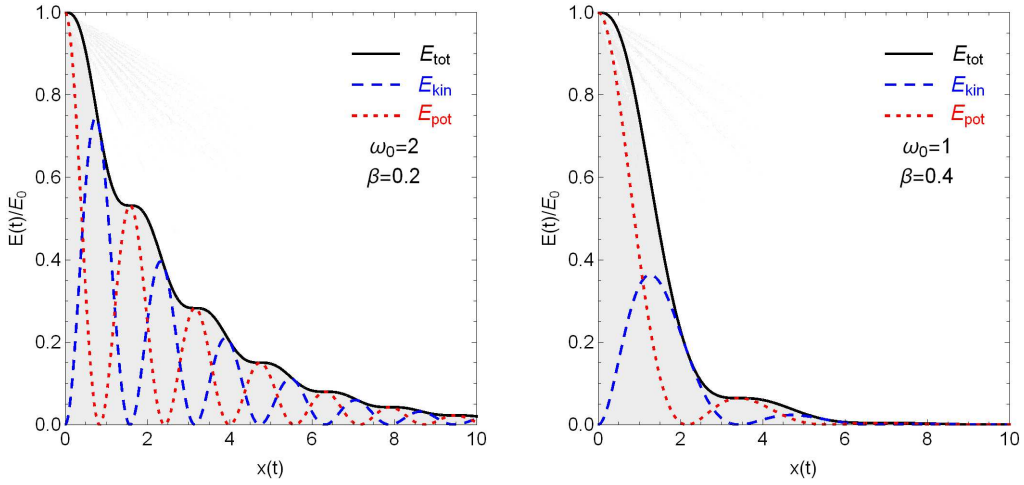


Figure 5.4: The total energy, kinetic energy, and potential energy (normalized to the initial energy E_0) of underdamped harmonic motion as a function of time, assuming that the oscillator starts from rest — *i.e.*, we take $v_0 = 0$, which means that $\phi = 0$. These results in the left panel correspond to an oscillator with $\omega = 2$ and $\beta = 0.2$ (in arbitrary units), while the results in the left panel correspond to an oscillator with $\omega = 1$ and $\beta = 0.4$.

In order to try to understand the physical meaning of the result in Eq. (5.51), let's begin by examining it graphically. In Fig. 5.4, we show some examples of how the total energy $E_{\rm tot}$ evolves with time for different choices of the parameters β and ω_0. The left panel shows the results for $\omega = 2$ and $\beta = 0.2$ (in arbitrary units); the right panel shows the results for $\omega = 1$ and $\beta = 0.4$. In each case, to simplify things, we have chosen the initial conditions so that the oscillator starts from rest — *i.e.*, *i.e.*, that $v_0 = 0$ — which means that $\phi = 0$ in Eq. (5.51). We also the individual contributions from the kinetic and potential energy terms in each case.

The first thing we notice in Fig. 5.4 is that the $E_{\rm tot}$ curves have "ripples." These ripples are the consequence of the cosine term in Eq. (5.51). It's important to emphasize that despite the presence of these ripples, there is never any point in time at which $E_{\rm tot}$ actually increases. If there were, that would mean that the system was gaining additional energy from somewhere! The ripples simply mean that $E_{\rm tot}$ is decreasing more rapidly at certain times than at others. We can appreciate why this is the case when we compare each $E_{\rm tot}$ curve in Fig. 5.4 to the corresponding curves for the kinetic- and potential-energy contributions individually. You'll notice that he rate of energy loss is highest when the kinetic energy is large. This is because the drag force $F_{\rm drag} = -bv$ responsible for dissipating the energy of the system to its surroundings is proportional to v. The rate of energy loss is therefore greater when v (and hence also the kinetic energy) is large. When the kinetic energy is small, as it is around the turning points of the motion where $x(t) \approx A(t)$, we see that $E_{\rm tot}$ decreases much more slowly.

The quality factor we defined in Eq. (5.38) actually turns out to be a very useful quantity for characterizing how the total energy in a damped oscillator diminishes over time. To see this, let's consider the case in which

the damping is very light, in the sense that $\beta \ll \omega_0$. In this case, the cosine term in Eq. (5.51 can be neglected, and the total energy of the system is approximately

$$E_{\rm tot} \approx \frac{1}{2}kA^2e^{-2\beta t} . \tag{5.53}$$

We can use the relationship between Q and β Eq. (5.38) to rewrite this equation in terms of Q:

$$E_{\rm tot} \approx E_0e^{-\omega_0 t/Q} , \tag{5.54}$$

where $E_0 \equiv kA^2/2$ is the initial energy of the oscillator at $t = 0$. Since $\omega_1 = \sqrt{\omega_0^2 - \beta^2} \approx \omega_0$ for $\beta \ll \omega_0$, we see that the period of oscillation for a lightly damped oscillator is approximately

$$T_1 = \frac{2\pi}{\omega_1} \approx \frac{2\pi}{\omega_0} . \tag{5.55}$$

Substituting this result into Eq. (5.54), we have

$$E_{\rm tot} \approx E_0e^{-2\pi t/QT_1} . \tag{5.56}$$

This expression gives us a direct physical interpretation of the quality factor Q. Specifically, it tells us that $Q/2\pi$ is the number of periods of oscillation the oscillator experiences before the energy $E_{\rm tot}$ of the oscillator drops to a fraction $1/e$ of its initial value. A large value of Q means that the system will pass through a substantial number of oscillations before its energy significantly diminishes.

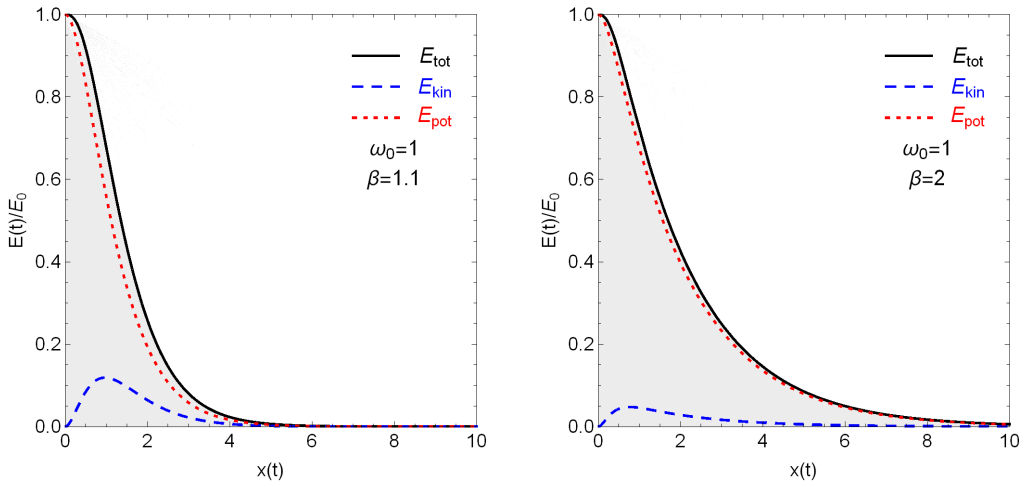


Figure 5.5: The total energy, kinetic energy, and potential energy (normalized to the initial energy E_0) of overdamped harmonic motion as a function of time. Is in Fig. 5.4, we assume that the oscillator starts from rest. These results in the left panel correspond to an oscillator with $\omega = 1$ and $\beta = 1.1$ (in arbitrary units), while the results in the left panel correspond to an oscillator with $\omega = 1$ and $\beta = 2$.

The energy associated with both overdamped and critically-damped oscillator systems can be computed in much the same way as the energy associated with underdamped systems. Indeed, one simply substitutes the appropriate expressions for $x(t)$ and $v(t)$ into Eq. (5.49). In Fig. 5.5, we show the $E_{\rm tot}$, kinetic-energy, and potential-energy curves for an overdamped oscillator (again normalized to the initial value of the energy E_0) with an initial velocity $v_0 = 0$. We observe that in the case of overdamped motion, $E_{\rm tot}$ decreases smoothly without the "ripple" effect which arises in the underdamped cases. We also observe that the potential-energy contribution to $E_{\rm tot}$ dominates over the kinetic-energy contribution at all times. This is because the damping is far more significant for overdamped motion than it is for underdamped motion. This means that the velocity never becomes terribly large in this case.

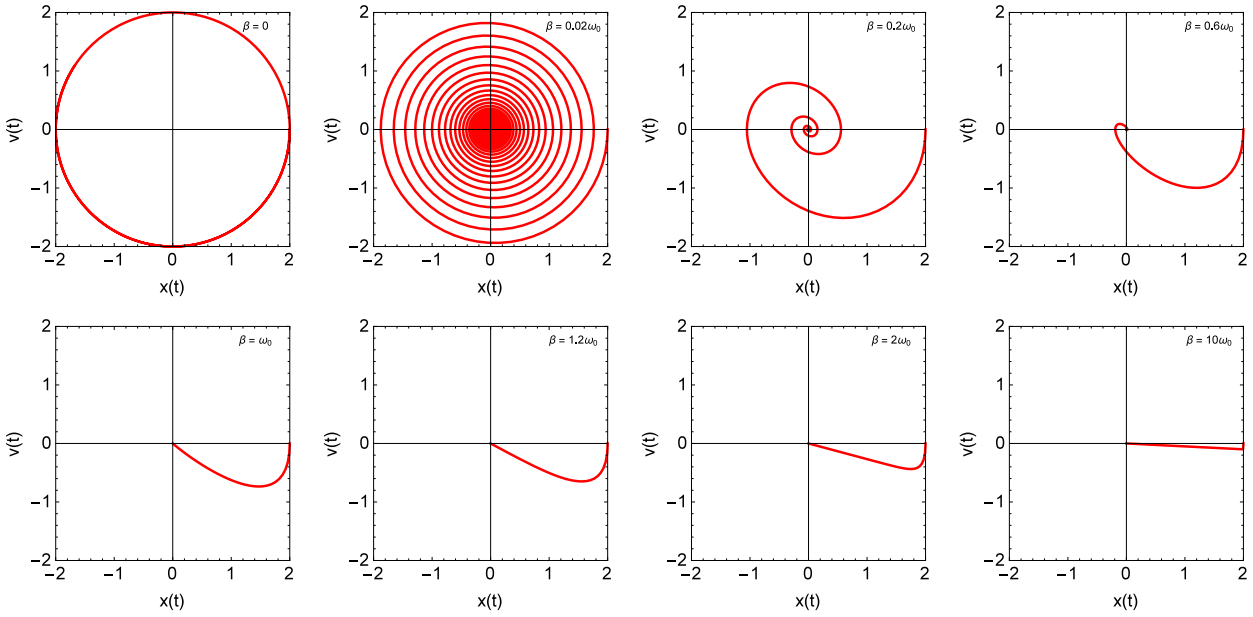


Figure 5.6: Curves showing the evolution of a damped-harmonic-oscillator system in state space. The eight panels of this figure, from top left to bottom right, illustrate how the state-space trajectory for such a system changes as the damping parameter β is increased. In each panel, we have taken $\omega_0 = 1$ (in arbitrary units) and chosen $x_0 = 2$ and $v_0 = 0$ as our boundary conditions conditions.

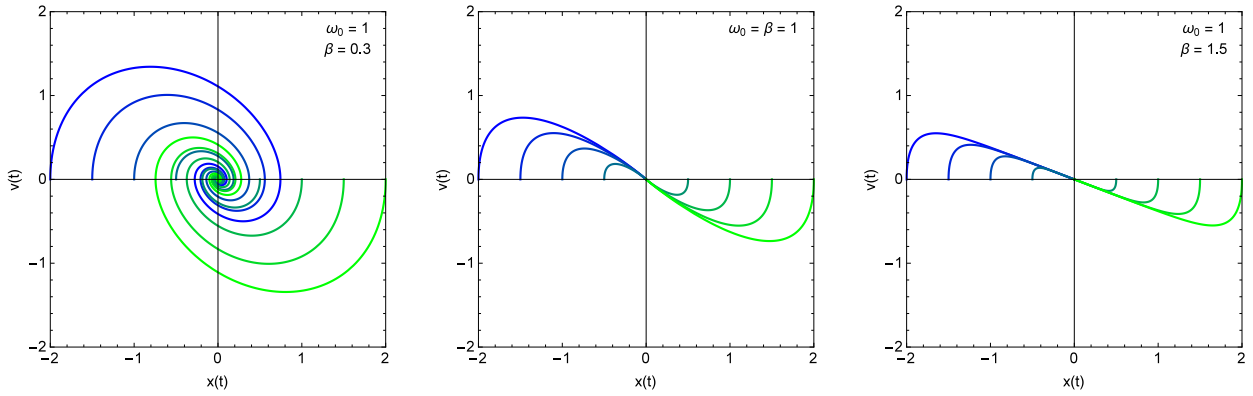


Figure 5.7: Examples of state-space trajectories for a damped harmonic oscillator in the underdamped regime (left panel), critically-damped regime (center panel), and overdamped regime (right panel). The curves in each panel correspond to different choices of the initial position x_0 with an initial velocity $v_0 = 0$.

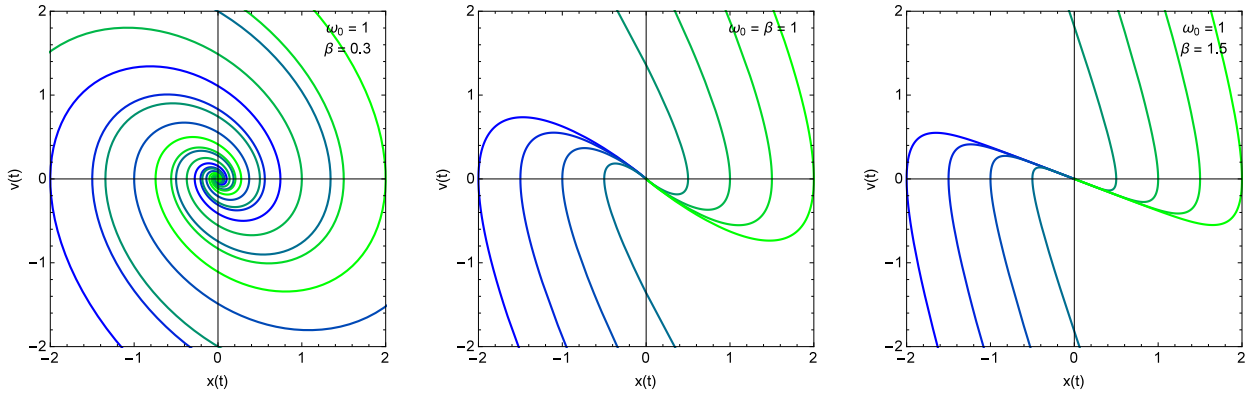


Figure 5.8: The *full* curves in state space along which the trajectories shown in Fig. 5.7 lie.

5.7 The State-Space Picture of Damped Harmonic Motion

We can also gain some further intuition about the physics of damped-oscillator systems by looking at how these systems evolve in state space. First, let's consider how the state-space trajectories depend on the fundamental parameters β and ω_0 which characterize the system. In Fig. 5.6, we show how the state-space trajectories for the damped harmonic oscillator depend on the relationship between β and ω_0. The eight panels of this figure, from top left to bottom right, illustrate how the state-space trajectory for a damped harmonic oscillator changes as the damping parameter β is increased. In each panel, we have set $\omega_0 = 1$ (in arbitrary units) and chosen $x_0 = 2$ and $v_0 = 0$ as our boundary conditions.

The panel at the top corresponds to the $\beta = 0$ limit, in which the system reduces to the simple harmonic oscillator and the state-space trajectories are circles.[4] The fact that the state-space trajectories form closed curves in this limit is a consequence of energy conservation. As β increases from zero, the system begins to lose energy to its surroundings and the phase space trajectories change from circles to spirals. These spirals tend toward the origin more and more rapidly as β increases until we reach the value $\beta = \omega_0$ which corresponds to the critically damped case. At this critical value of β, the state-space trajectories change from spirals to curves which flow directly toward the origin without "overshooting."

It's also worth taking a look at how the state-space trajectories depend on the choice of boundary conditions. In Fig. 5.7, we show a variety of state-space curves for an overdamped oscillator (left panel), a critically-damped oscillator (center panel), and an overdamped oscillator (right panel). Each curve corresponds to a different choice of the initial position x_0 for an initial velocity $v_0 = 0$. Just as we saw for the simple harmonic oscillator, the state-space trajectories for a particular oscillator system (*i.e.*, a particular choice of β and ω_0) never cross. However, they do converge to the origin as $t \to \infty$ because of the energy loss due to damping.

Each of the state-space trajectories shown in Fig. 5.7 starts at a particular point in state space which is specified by the boundary conditions x_0 and v_0. However, one can also view these trajectories as truncated "pieces" of the full state-space curves shown in Fig. 5.8. For every point (x, v) in state space, there is a unique curve which passes through that point. An oscillator will therefore follow the curve that passes through the point (x_0, v_0), always proceeding in a clockwise direction as time progresses.

5.8 Frictional Damping

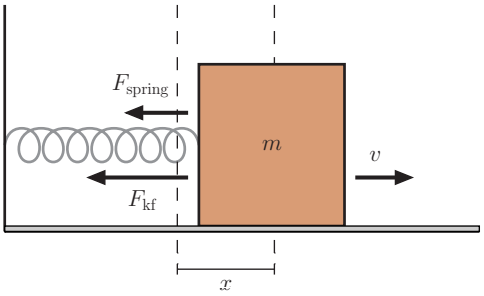


Figure 5.9: A mass oscillating at the end of a spring with a frictional damping force. The diagram shows the mass during the part of its motion when $x > x_{\text{eq}}$ and $v > 0$.

Up to this point, our discussion of dissipative forces have been focused almost exclusively on linear drag forces. We have seen that oscillator systems damped by such a force can display a broad range of behavior. However, we have not yet said anything about another dissipative force which plays a very important role in real physical systems: friction. You may recall from mechanics that the force of kinetic friction on sliding object takes the form

$$F_{\text{kf}} = -\mu_k N \text{sign}(v) \, , \tag{5.57}$$

[4] In the more general case in which $\omega_0 \neq 0$, the state-space trajectories would be ellipses.

where N is the magnitude of the normal force acting on the object, μ_k is the coefficient of kinetic friction, and $\text{sign}(v)$ simply indicates the sign of the velocity:

$$\text{sign}(v) \;=\; \begin{cases} +1 & v > 0 \\ -1 & v < 0 \; . \end{cases} \tag{5.58}$$

Let's consider what happens if we add such a frictional damping force to the equation of motion for a simple harmonic oscillator, as illustrated in Fig. 5.9. Since μ_k and N are both constants, it might seem like incorporating a frictional damping force this equation would be even easier than incorporating the linear drag force we dealt with in Sect. 5.4. However, the quantity $\text{sign}(v)$ in Eq. (5.57) actually makes dealing with frictional damping a non-trivial matter. The reason is that when we add a frictional damping force to the equation of motion for a simple harmonic oscillator, the resulting differential equation

$$m\frac{d^2x}{dt^2} \;=\; -kx - \mu_k mg\,\text{sign}\left(\frac{dx}{dt}\right) \tag{5.59}$$

is nonlinear because of the way in which it depends on dx/dt.

In this case, however, the non-linearity is merely a consequence of the friction term in Eq. (5.59) flipping its sign at the turning points of the motion when the block in Fig. 5.9 reverses its direction. During the time interval between any two of those turning points, the equation actually looks quite linear indeed. For example, while the block is traveling to the right, its velocity v is positive. Thus, the differential equation for $x(t)$ is

$$\frac{d^2x}{dt^2} \;=\; -\frac{k}{m}x - \mu_k g \; . \tag{5.60}$$

We have seen equations of this form before and know what the solutions look like. Moreover, in Sect. 4.4, we discussed the fact that the general solution to an inhomogeneous linear differential equation of arbitrary order may be written as a sum of the general solution $y_c(x)$ to the complementary equation and any particular solution $y_p(x)$ to the full equation. The complementary equation is just the simple-harmonic-oscillator equation. Furthermore, we know that there exists one particular solution $x_p(t)$ to the full equation consists of the block sitting inert at the equilibrium point $x_{\text{eq}} = -\mu_k mg/k$. Thus, the general solution to Eq. (5.60) can be written in the form

$$x(t) \;=\; B\cos(\omega t) + C\sin(\omega t) - \frac{\mu_k mg}{k} \; . \tag{5.61}$$

We can solve for $x(t)$ at times when the velocity v of the block is negative in a similar fashion. For $v < 0$, the the equation of motion is

$$\frac{d^2x}{dt^2} \;=\; -\frac{k}{m}x + \mu g \; , \tag{5.62}$$

and the general solution takes the form

$$x(t) \;=\; B\cos(\omega t) + C\sin(\omega t) + \frac{\mu_k mg}{k} \; . \tag{5.63}$$

We emphasize that the undetermined coefficients B and C appearing in Eq. (5.61) and the coefficients B and C appearing in Eq. (5.63) are *not* necessarily equal. On the contrary, like any undetermined coefficients which appear in the general solution to a differential equation, they are determined by the appropriate boundary conditions.

But what are those boundary conditions and how to we deal with them? For every other differential equation we've encountered thus far, we simply had to apply our boundary conditions once in order to obtain a solution that was valid for all values of the independent variable. Here, however, the solution that we get by plugging in our initial values for $x(t)$ and $v(t)$ at $t = 0$ is only valid until the block reaches the first turning point at some time t_1 and reverses its direction. After that, the form of the solution changes. However, since the solution that applies for this first time interval is valid all the way up to the turning point, we can use the values of $x(t_1)$ and $v(t_1)$ obtained from this solution as our initial conditions for $x(t)$ and $v(t)$ for the subsequent time interval from t_1 until the second turning point at time t_2. We can then use the values of $x(t_2)$ and $v(t_2)$ obtained from this second solution as the initial conditions for the next time interval, and so on.

To understand how this procedure works in practice, it's best to begin with a concrete example. In particular, let's consider the case in which the block begins at time $t = 0$ with an initial position $x_0 > 0$ and an initial velocity $v_0 = 0$. In this case, the mass will begin moving to the left, so the appropriate solution for $x(t)$ is the solution in Eq. (5.63). The boundary condition on $x(t)$ gives us

$$x_1(0) \;=\; x_0 \;=\; B_1\cos(0) + C_1\sin(0) + \frac{\mu_k mg}{k} \;=\; B_1 + \ell \;, \tag{5.64}$$

where, for convenience, we have defined $\ell \equiv \mu_k mg/k$. Note that ℓ has dimensions of length. The coefficient C_1 can be determined from the boundary condition

$$\left.\frac{dx}{dt}\right|_{t=0} \;=\; 0 \;=\; -B_1\omega\sin(0) + C_1\omega\cos(0) \;=\; C_1 \;. \tag{5.65}$$

Thus, our initial solution is

$$x_1(t) \;=\; (x_0 - \ell)\cos(\omega t) + \ell \;. \tag{5.66}$$

We're not done yet, however. The solution in Eq. (5.66) is only valid so long as $v < 0$ and the block is moving to the left. At $t = \pi/\omega = T/2$, the velocity — *i.e.*, the derivative of $x_1(t)$ — changes from negative to positive. At this point, Eq. (5.60) becomes the equation of motion for $x(t)$, and Eq. (5.61) becomes the solution. Let's call this second piece of the solution $x_2(t)$. The values of the coefficients B and C for $x_2(t)$ can be determined by applying the appropriate boundary conditions. In this case, those conditions are that the initial values for $x_2(t)$ and $v_2(t)$ match up with the final values for $x_1(t)$ and $v_1(t)$ at the turning point $t = T/2$. Thus, we have

$$\begin{aligned} x_2(T/2) \;&=\; B_2\cos(\pi) + C\sin(\pi) - \ell \;=\; -B_2 - \ell \;=\; (x_0 - \ell)\cos(\pi) + \ell \;=\; 2\ell - x_0 \\ v_2(T/2) \;&=\; -B_2\omega\sin(\pi) + C_2\omega\cos(\pi) \;=\; -C_2\omega - (x_0 - \ell)\sin(\pi) \;=\; 0 \;. \end{aligned} \tag{5.67}$$

Solving for B_2 and C_2 and plugging the result back into Eq. (5.60), we get

$$x_2(t) \;=\; (x_0 - 3\ell)\cos(\omega t) - \ell \;. \tag{5.68}$$

Once again, however, this solution is only valid until $t = 2\pi = T$ and the derivative of $x_2(t)$ changes from positive to negative. After that, the solution once again takes the form in Eq. (5.63), with B and C determined by the boundary conditions $x_3(T) = x_2(T)$ and $v_3(T) = v_2(T)$. The solution turns out to be

$$x_3(t) \;=\; (x_0 - 5\ell)\cos(\omega t) + \ell \;. \tag{5.69}$$

By repeating this procedure each time the velocity switches sign, we obtain a set of individual solutions for the position of the block, each of which is valid only within a particular range of the independent variable t:

$$x(t) \;=\; \begin{cases} (x_0 - \ell)\cos(\omega t) + \ell & 0 \le t < T/2 \\ (x_0 - 3\ell)\cos(\omega t) - \ell & T/2 \le t < T \\ (x_0 - 5\ell)\cos(\omega t) + \ell & T \le t < 3T/2 \\ (x_0 - 7\ell)\cos(\omega t) - \ell & 3T/2 \le t < 2T \\ \dots & \dots \end{cases} \tag{5.70}$$

This solution for $x(t)$ is an example of a **piecewise solution** to a differential equation: set of individual solutions valid for particular ranges of the independent variable which are ultimately "sewn together" by boundary conditions at the endpoints of each range. During the first half-period of oscillation, $x(t)$ looks like the solution for a simple harmonic oscillator oscillating around the point $x_{\rm eq} = \ell$ with amplitude $x_0 - \ell$; during the second half-period, the $x(t)$ looks like a simple harmonic oscillator oscillating around the point $x_{\rm eq} = -\ell$ with amplitude $x_0 - 3\ell$; and so forth. This behavior is illustrated in Fig. 5.10.

The reason is that we have only dealt with *kinetic* friction thus far. At the points $x_0 - (2n+1)\ell$ where $x(t)$ reaches a minimum or a maximum, the velocity is zero and *static* friction becomes relevant. You may recall from your mechanics course that the force of static friction (which we'll call $F_{\rm sf}$) acts on an object at rest on a surface, and acts in the direction opposite to the direction of the component of the applied force

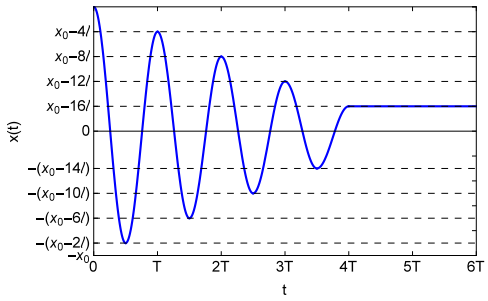


Figure 5.10: The position $x(t)$ of a mass oscillation on a spring with frictional damping, as given by the piecewise solution in Eq. (5.70). The mass oscillates with an amplitude that decreases with each half-cycle until the spring force can no longer overcome the force of static friction at one of the turning points at which $v = 0$. After this occurs, the mass just remains "stuck" at the turning point.

$F_{\rm app}$ parallel to that surface. The magnitude of $F_{\rm sf}$ is equal to the magnitude of $F_{\rm app}$ up to a maximum value $|F_{\rm sf}| = \mu_s N$, where N is the normal force and μ_s is the coefficient of static friction. In other words,

$$F_{\rm sf} \;=\; \begin{cases} -F_{\rm app} & \text{for } F_{\rm app} < \mu_s N \\ -\mu_s N \,\text{sign}(F_{\rm app}) & \text{for } F_{\rm app} \geq \mu_s N \;. \end{cases} \tag{5.71}$$

This static-friction force,is generically stronger than the force of kinetic friction. Consequently, the mass has a tendency to get "stuck" at the first extremum it reaches for which the restoring force from the spring is weaker than the force of static friction — *i.e.*, for which

$$\mu_s mg \;>\; k\big[x_0 - (2n+1)\ell\big] \;=\; kx_0 - (2n+1)\mu_k mg \;. \tag{5.72}$$

The number of oscillations which the mass undergoes before this condition is met depends on the particular values which these parameters happen to take. For example, in the case illustrated in Fig. (5.10), the we have chosen our parameters such that the mass becomes stuck at $t = 4T$, after four complete oscillations.

Problems

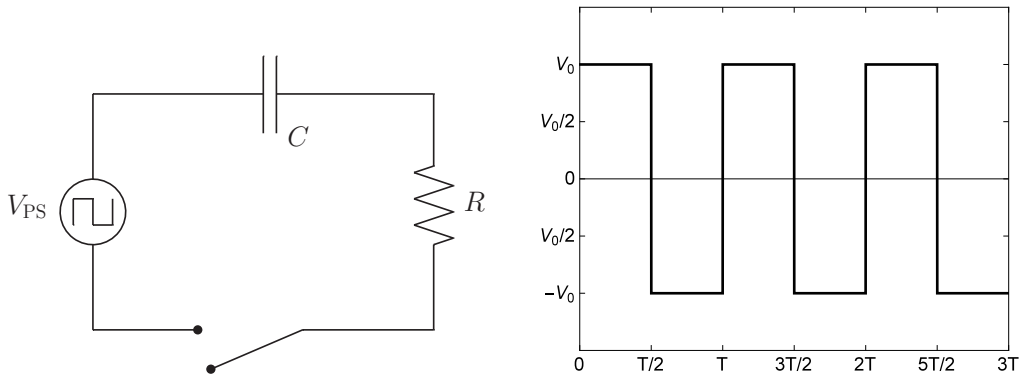


Figure 5.11: An RC circuit hooked up to power supply which delivers a square-wave voltage $V_{\rm PS}$ which alternates between $V_{\rm PS} = V_0$ and $V_{\rm PS} = -V_0$.

1. Verify that the functions $y_1(x) = \cos x$ and $y_2(x) = \sin x$ are linearly independent by computing the Wronskian and showing that it does not vanish identically.

2. Consider the differential equation

$$x^2 \frac{d^2y}{dx^2} + 2x\frac{dy}{dx} + x^2 y \;=\; 0 \; . \tag{5.73}$$

   (a) Verify that the function $y_1(x) = \sin(x)/x$ is a solution to this differential equation.

   (b) Use reduction of order to find a second linearly-independent solution $y_2(x)$ and write down the general solution $y(x)$.

3. In Sect. 5.2, we saw how reduction of order could be used to derive a second, linearly-independent solution $y_2(x)$ to a homogeneous linear differential equation, provided that you already know one solution $y_1(x)$. This method can actually be extended to inhomogeneous equations of the form

$$P(x)\frac{d^2y}{dx^2} + Q(x)\frac{dy}{dx} + R(x)y \;=\; G(x) \tag{5.74}$$

   as well. In this problem, you'll examine how this method works in the inhomogeneous case.

   (a) Let's say you already know one solution $y_{1c}(x)$ to the *complementary* equation to Eq. (5.74) — *i.e.*, the corresponding differential equation with the inhomogeneous term removed:

$$P(x)\frac{dy_{1c}^2}{dx^2} + Q(x)\frac{dy_{1c}}{dx} + R(x)y_{1c} \;=\; 0 \; . \tag{5.75}$$

   Begin by setting $y_2 = u(x)y_{1c}(x)$ and derive a second-order differential equation for $u(x)$ by following an analogous set of steps to those we followed in Sect. 5.2.

   (b) Define $v(x) \equiv du/dx$ and use the methods you know for solving first-order differential equations to solve the resulting first-order linear differential equation for $v(x)$ in terms of $y_{1c}(x)$ (and its derivatives) and the functions $P(x)$, $Q(x)$, $R(x)$, and $G(x)$. Make sure to retain all relevant constants of integration.

   (c) Use this expression for $v(x)$ to derive a formula for $y_2(x)$ in terms of these quantities. Again, make sure to retain all relevant constants of integration.

4. In the case that the two roots to the characteristic equation in Eq. (5.16) are degenerate — *i.e.*, when $r_+ = r_-$ — our solution in Eq. (5.19) does not hold because the functions $y_+(x)$ and $y_-(x)$ are not linearly independent in this case. Use the solution $y_1(x) = e^{rx}$, where $r_+ = r_- = r$ is the degenerate solution to the characteristic equation, and derive the second solution $y_2(x)$ using reduction of order, as outlined in Sect. 5.2.

5. Show that the distance between adjacent maxima in $x(t)$ for underdamped harmonic motion is given by the period $T_1 = 2\pi/\omega_1$. Show that this time is also twice the distance between adjacent zeroes of $x(t)$.

6. Show that one can determine the decay parameter β for underdamped harmonic motion from a plot of $x(t)$. In particular, show that

$$\beta \;=\; \frac{\ln R}{T_1} \; , \tag{5.76}$$

   where R is the ratio of the amplitudes at any two adjacent maxima of $x(t)$. Note that in deriving this result, you will need the fact about T_1 that you proved in Problem 5.

7. Consider a mass m hanging vertically from a spring with a spring constant k. In addition to gravity, assume that the mass experiences a linear drag force $F_{\rm drag} = -bv$, where b is the drag coefficient. At time $t = 0$, the mass is released from rest with the spring unstretched. Assuming that the motion is critically damped, find $x(t)$ and make a plot of your solution. Note that depending on how you choose your origin in this problem, you may need to add a constant to Eq. (5.47) in order to enforce that this solution satisfies the equation of motion.

8. Show that the expression for $v(t)$ on the second line of Eq. (5.50) is equivalent to the expression on the first line. (Hint: the identity in Eq. (68) in Chapter 4 of these lecture notes might be helpful.)

9. Show the expression for the energy of an underdamped oscillator on the second line of Eq. (5.51) is equivalent to the expression on the first line.

10. Take the derivative of Eq. (5.49) and use the equation of motion for the damped harmonic oscillator in Eq. (5.22) to show that

$$\frac{dE_{\rm tot}}{dt} \;=\; F_{\rm drag}v \;. \tag{5.77}$$

11. In this problem, we're going to examine the behavior of the total energy $E_{\rm tot}$ as a function of time for overdamped and critically-damped motion in greater detail. We're also going to see why shock absorbers are designed to be critically damped.

    (a) First, derive a general expression for $E_{\rm tot}$ as a function of time for the case of an overdamped oscillator. Express your result in terms of the initial position x_0 and the initial velocity v_0 of the oscillator using Eq. (5.41). (Note: you are not *not* required to express your result in any special form through the use of hyperbolic-function identities.)

    (b) Now derive the corresponding expression for $E_{\rm tot}$ for a critically-damped oscillator. Express your result in terms of x_0 and v_0 using the relations in Eq. (5.48).

    (c) You are now going to examine the effect that varying the ratio β/ω_0 has on the time-evolution of $E_{\rm tot}$. Use Mathematica to plot $E_{\rm tot}$ vs. time for the overdamped case in part (a) for several different values of β but with all other parameters fixed. In particular, plot four such curves: one for $\beta = 1.1$ s^{-1}, one for $\beta = 1.2$ s^{-1}, one for $\beta = 1.3$ s^{-1}, and one for $\beta = 1.4$ s^{-1}. For all of these curves, set $\omega_0 = 1$ s^{-1} and $k = 10$ N/m and take $x_0 = 20$ cm and $v_0 = 0$ as your boundary conditions. Finally, add a curve to the same plot which corresponds to the critically-damped case in part (b) for this same choice of parameters.

    (d) Plot the value of the decay parameter $\Gamma_{\rm dec}$ as a function of β for this system for $0.1\omega_0 \leq \beta \leq 5\omega_0$. For what value of β is $\Gamma_{\rm dec}$ maximized?

    (e) Use the graphical results from parts (c) and (d) to justify why you would typically want to engineer a shock absorber to be critically damped rather than overdamped or underdamped.

12. In the case of very light damping, as we discussed in Sect. 5.6, the period of oscillation is given by $T_1 \approx T_0$. Derive the first non-vanishing correction to T_1 in terms of T_0 and Q.

13. The piecewise solution for $x(t)$ in Eq. (5.70) for an oscillator with frictional damping corresponds to the initial conditions $x_0 > 0$ and $v_0 = 0$. Derive the piecewise solution for the same system which would result from taking the initial conditions to be $x_0 = 0$, $v_0 < 0$.

14. A block of mass $m = 0.3$ kg is attached to a long spring with spring constant $k = 10$ N/m is sliding along a concrete floor. Use Mathematica to create a state-space plot describing the motion of the mass if it starts oscillating from rest (*i.e.*, with initial velocity $v_0 = 0$) with the spring at a distance $x_0 = 3$ m away from its equilibrium position. The coefficients of kinetic and static friction between the block and the table are $\mu_k = 0.5$ and $\mu_s = 1.0$, respectively.

15. An RC circuit is hooked up to a power source which delivers a square-wave voltage $V_{\rm PS}$ that alternates back and forth between $V_{\rm PS} = V_0$ and $V_{\rm PS} = -V_0$ with period T, as shown in Fig. 5.11. At time $t = 0$, the switch is closed.

    (a) Find the piecewise solution for the charge $Q(t)$ on the capacitor as a function of t within the range $0 \leq t \leq 2T$, assuming that the capacitor is initially uncharged.

    (b) Use Mathematica to plot the solution for $Q(t)$ over this range of t for $V_0 = 1$ V, a capacitance $C = 0.002$ F, a resistance $R = 250\ \Omega$, and a period $T = 5$ s.

Chapter 6

Driven Oscillations and Resonance

- **The physics:** Driven oscillations, resonance
- **The math:** Second-order inhomogeneous linear differential equations, the method of undetermined coefficients, complexification

6.1 Second-Order Inomogeneous Linear Differential Equations

In the last chapter of these notes, we discussed a number of techniques for solving homogeneous second-order linear differential equations. We then put these techniques to use to study the behavior of physical systems like damped harmonic oscillators and LRC circuits which include both a restoring force and a dissipative force which causes the system to lose energy to its surroundings. In this chapter, we will extend this discussion to examine the effect of incorporating an external "driving force" into the system which injects energy into to the system over time.

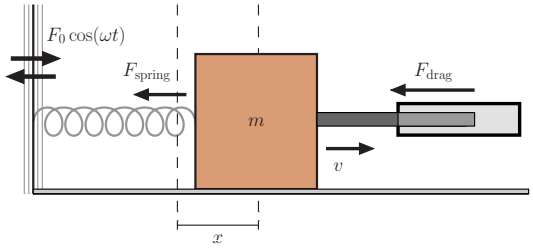


Figure 6.1: Illustration of a damped, driven harmonic oscillator system involving a block of mass m attached to a spring. The damping in this system is the result of a piston which provides a linear drag force $F_{\text{drag}} = -bv$. A driving force $F_{\text{drive}} = F_0\cos(\omega t)$ results from shaking the wall to which the other end of the spring is attached.

To explain what we mean by a "driving force," it's probably best to start with an example. In Fig. 6.1, we provide an illustration of a damped, driven harmonic-oscillator system. In this example, the oscillator consists of a block with mass m attached to a spring which supplies a restoring force $F_{\text{spring}} = -kx$). In addition, the block is also attached to the end of a piston which provides a linear drag force $F_{\text{drag}} = -bv$. A driving force $F_{\text{drive}} = F_0\cos(\omega)$ is applied to the block by shaking the wall to which the other end of the spring is attached so that it vibrates back and forth with a period $T = 2\pi/\omega$. Putting together all of these contributions to the net force

$$F \;=\; F_{\text{spring}} + F_{\text{drag}} + F_{\text{drive}}\ , \tag{6.1}$$

we find that the equation of motion for the position $x(t)$ of the block that we get from Newton's second law

for this example looks like

$$m\frac{d^2x}{dt^2} \;=\; -kx - b\frac{dx}{dt} + F_0\cos(\omega t)\ . \tag{6.2}$$

Physically, the interpretation of the driving term in Eq. (6.2) is reasonably straightforward. The motion of the wall compresses or extends the spring and thereby changes the equilibrium position $x_{\rm eq}$ of the block. This is essentially the same effect we've seen many times in the presence of additional constant forces that act on the system and change $x_{\rm eq}$ — for example, the force of gravity acting on a block suspended from a vertical spring. However, there is one important difference here: in this case, the additional force being applied is time-dependent.

Mathematically, the interpretation of the driving term is straightforward as well. The driving term does not involve $x(t)$ or any of its derivatives, so it represents an inhomogeneous term in our differential equation. In this example, the inhomogeneous term happened to take the form of a cosine function, but of course we could imagine a lot of other functional forms which this term might take in different physical scenarios. Clearly, if we want to understand driven harmonic oscillations, we're first going to need to understand how to deal with inhomogeneous linear differential equations.

Fortunately, there's actually a lot that we already *do* understand about how to deal with equations of this sort. In particular, we have already seen that the general solution $y(x)$ to an inhomogeneous linear differential equation of any order can be determined from the relation

$$y(x) \;=\; y_c(x) + y_p(x)\ , \tag{6.3}$$

where $y_c(x)$ is the general solution to the complementary equation and $y_p(x)$ is any particular solution — no matter how trivial — to the full, inhomogeneous equation. In the last chapter of these lecture notes, we examined a few useful techniques for obtaining general solutions to certain kinds of homogeneous linear differential equations. Thus, we already have an intuition about how to obtain $y_c(x)$ for a variety of systems of physical interest, including the damped harmonic oscillator. However, we haven't yet discussed how to obtain a particular solution $y_p(x)$ to the full equation except for in the most trivial of cases — the case in which the function $f(x)$ in Eq. (6.4) is a constant. We therefore begin our discussion of inhomogeneous second-order linear differential equations with a discussion of how to go about finding particular solutions to these equations.

6.2 The Method of Undetermined Coefficients

The most general form for an inhomogeneous linear second-order differential equation is

$$P(x)\frac{d^2y}{dx^2} + Q(x)\frac{dy}{dx} + R(x)y \;=\; f(x)\ , \tag{6.4}$$

where $P(x)$, $Q(x)$, $R(x)$, and the inhomogeneous term $f(x)$ are all arbitrary functions of the independent variable x. However, our primary motivation is to study the behavior of the damped, driven harmonic oscillator and other mathematically analogous systems, we're primarily going to focus on a particular subset of equations with this general form. To wit, we're going to focus on equations of the form

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy \;=\; f(x)\ , \tag{6.5}$$

where a, b, and c are constants.

A variety of useful techniques exist for solving differential equations of this form. We'll begin by discussing one of the simplest of these techniques — one which goes by the name of the **Method of Undetermined Coefficients**. Really, this "method" is essentially a glorified name for what boils down to making an educated guess about the form of $y_p(x)$ for the differential equation you're trying to solve. This "method" is actually very similar to the method that we used to obtain our general solution for a homogeneous linear differential equation with constant coefficients. We begin by using use our intuition about the relationship between functions and their derivatives to guess the functional form of $y_p(x)$. However, there will be a few constant parameters — let's call them A_i — in our parametrization for $y_p(x)$ whose values we still need to

determine. In order to determine the values for these parameters (our "undetermined coefficients"), we need to plug our guess for $y_p(x)$ back into the differential equation and solve for them.

For example, let's say that we were trying to find a particular solution to the differential equation

$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y \;=\; e^x \;. \tag{6.6}$$

First, we need The function e^x appears on the right-hand side of this equation, while the left-hand side involves both y and its derivatives. We know that the derivative of e^x with respect to x is just e^x itself, so one reasonable guess we could make for $y_p(x)$ is

$$y_p(x) \;\stackrel{?}{=}\; Ae^x \;, \tag{6.7}$$

where A is our undetermined coefficient. Plugging this guess back into Eq. (6.6) gives us

$$\frac{d^2}{dx^2}(Ae^x) + 2\frac{d}{dx}(Ae^x) + Ae^x \;=\; 4Ae^x \;=\; e^x \;. \tag{6.8}$$

Thus, our educated guess for $y_p(x)$ indeed is a solution to this equation — provided that our coefficient takes the value $A = 1/4$. Thus, one particular solution to Eq. (6.6) is

$$y_p(x) \;=\; \frac{1}{4}e^x \;. \tag{6.9}$$

The process of making successful guesses about the form of $y_p(x)$ for different kinds of inhomogeneous functions $f(x)$ is admittedly as much of an art as it is a science. Thus, it makes sense to draw from the well of accumulated wisdom about such guesswork rather than stumbling afresh through the process of trial and error each time we're faced with a new differential equation that we're trying to solve. Indeed, there are certain functional forms for $f(x)$ — including polynomial functions, exponentials, and certain trigonometric functions — for which the correct "guess" for the form of $y_p(x)$ is well known. These functional forms for $f(x)$, along with the corresponding forms for $y_p(x)$, are listed in Table 6.1.

Functional form of $f(x)$		Form of $y_p(x)$
$f(x) \;=\; C_0 + C_1x + C_2x^2 + \ldots C_nx^n$	$\rightarrow$	$y_p(x) \;=\; A_0 + A_1x + A_2x^2 + \ldots A_nx^n$
$f(x) \;=\; Ce^{ax}$	$\rightarrow$	$y_p(x) \;=\; Ae^{ax}$
$f(x) \;=\; C_1\cos(ax) + C_2\sin(ax)$	$\rightarrow$	$y_p(x) \;=\; A_1\cos(ax) + A_2\sin(ax)$
$f(x) \;=\; C_1e^{ax}\cos(bx) + C_2e^{ax}\sin(bx)$	$\rightarrow$	$y_p(x) \;=\; A_1e^{ax}\cos(bx) + A_2e^{ax}\sin(bx)$

Table 6.1: Functional forms for the inhomogeneous term $f(x)$ and the corresponding "guesses" for the functional form of the particular solution $y_p(x)$ to be used in the Method of Undetermined Coefficients. The coefficients C_n and the parameters a and b appearing in the expressions for $f(x)$ are constants.

It should be clear that not every differential equation of the form Eq. (6.5) can be solved using the Method of Undetermined Coefficients. This method is only useful for solving equations where the inhomogeneous term $f(x)$ has a particular functional form for which it's easy to guess the form of $y_p(x)$. If $f(x)$ doesn't have such a form, we'll need to resort to other methods for finding our particular solution.

One additional comment is in order. This is that you should take care not to confuse the undetermined coefficients A_i in Table 6.1 with the n undetermined constants B_i which characterize the general solution $y(x)$ an nth-order linear differential equation! The A_i are only "undetermined" in the sense that we not know what their values are going to be at the moment when we write down our guess for $y_p(x)$. We have no freedom to choose those values: they're completely determined by the underlying physical parameters which characterize the system. Instead, we solve for them by plugging $y_p(x)$ into the differential equation and find

the specific values for the A_i for which that equation is valid for all x. At that point, these coefficients aren't undetermined anymore. By contrast, the undetermined constants B_i appearing in the general solution $y(x)$ truly are free parameters in the sense that $y(x)$ will solve the corresponding differential equation no matter matter what values I choose for the B_i. Additional information — such as information about the boundary conditions which characterize the problem — is required in order to specify their values.

Example: Applying the Method of Unetermined Coefficients

As a further example of how the Method of Undetermined Coefficients is applied in practice, let's use it to find a particular solution to the differential equation

$$\frac{d^2y}{dx^2} + 3\frac{dy}{dx} + y \;=\; 5\cos(2x)\;. \tag{6.10}$$

The inhomogeneous term in this equation has the form $f(x) = C_1\cos(ax) + C_2\sin(ax)$ with $C_2 = 0$ and $a = 2$, so Table 6.1 tells us that our "guess" for $y_p(x)$ should be $y_p(x) = A_1\cos(2x) + A_2\sin(2x)$. Plugging this guess into Eq. (6.10) gives us

$$-4A_1\cos(2x) - 4A_2\sin(2x) - 6A_1\sin(2x) + 6A_2\cos(2x) + A_1\cos(2x) + A_2\sin(2x) \;=\; 5\cos(2x)\;. \tag{6.11}$$

Collecting the sine and cosine terms together, we have

$$(6A_2 - 3A_1 - 5)\cos(2x) - (6A_1 + 3A_2)\sin(2x) \;=\; 0\;. \tag{6.12}$$

This equation can only be satisfied for all x if the coefficients of the sine and cosine terms vanish independently. Thus, it must be true that

$$\begin{aligned} 6A_2 - 3A_1 - 5 \;&=\; 0 \\ 6A_1 + 3A_2 \;&=\; 0 \end{aligned} \tag{6.13}$$

We can easily solve this system of equations for our two unknowns A_1 and A_2. The second equation tells us that $A_2 = -2A_1$, and when we substitute this into the first equation, we arrive at the solution

$$A_1 \;=\; -\frac{1}{3}\;, \qquad A_2 \;=\; \frac{2}{3}\;. \tag{6.14}$$

Thus, our particular solution to Eq. (6.10) is

$$y_p(x) \;=\; -\frac{1}{3}\cos(2x) + \frac{2}{3}\sin(2x)\;. \tag{6.15}$$

Now that we have a particular solution for Eq. (6.10), we can use it to construct the general solution as well. Indeed, Eq. (6.3) tells us that all we need to do is find the general solution to the complementary equation

$$\frac{d^2y}{dx^2} + 3\frac{dy}{dx} + y \;=\; 0\;. \tag{6.16}$$

Since this is just a homogeneous linear second-order equation with constant coefficients, the solutions are

$$y_c(x) \;=\; B_1e^{r_+x} + B_2e^{r_-x}\;, \tag{6.17}$$

where the roots $r_\pm$ of the characteristic equation in this case are

$$r_\pm \;=\; -\frac{3}{2} \pm \frac{\sqrt{5}}{2}\;. \tag{6.18}$$

Thus, our general solution to the full equation is

$$y(x) \;=\; B_1e^{-(3-\sqrt{5})x/2} + B_2e^{-(3+\sqrt{5})x/2} - \frac{1}{3}\cos(2x) + \frac{2}{3}\sin(2x)\;. \tag{6.19}$$

6.3 Solving the Driven Harmonic Oscillator Equation

Now that we have some sense of how to solve inhomogeneous linear differential equations using the Method of Undetermined coefficients, let's apply this knowledge to the physical system which motivated us to study this technique in the first place: the damped, driven harmonic-oscillator. We'll begin by considering a driving force of the same form $F_{\rm drive} = F_0\cos(\omega t)$ that we had in the example in Sect. 6.1. In other words, we will assume that the driving force oscillates sinusoidally in time with an amplitude F_0 and a frequency ω.

We'll begin by using our usual definitions $\omega_0 = \sqrt{k/m}$ and $\beta = b/2m$ in order to rewrite Eq. (6.2) in the form

$$\frac{d^2x}{dt} + 2\beta\frac{dx}{dt} + \omega_0^2 x \;=\; \omega_0^2\frac{F_0}{k}\cos(\omega t)\;. \tag{6.20}$$

It's important to remember that although ω_0 and ω are both angular frequencies, they are different quantities and don't necessarily have anything to do with each other. The first, ω_0, is determined by the physical characteristics of the oscillator system itself — in particular, the mass m and the spring constant k. By contrast, the second, ω, is the angular frequency the external driving force.

Let's apply the method of undetermined coefficients to obtain the general solution for the driven-harmonic-oscillator equation in Eq. (6.20). Since the inhomogeneous term in this equation — *i.e.*, the term associated with the driving force $F_{\rm drive}$ — is a cosine function, we know that our educated guess for the form of $x_p(t)$ should be

$$x_p(t) \;=\; A_1\cos(\omega t) + A_2\sin(\omega t)\;. \tag{6.21}$$

Substituting this into Eq. (6.20), we get

$$-\omega^2\big[A_1\cos(\omega t)+A_2\sin(\omega t)\big] - 2\omega\beta\big[A_1\sin(\omega t)-A_2\cos(\omega t)\big] + \omega_0^2\big[A_1\cos(\omega t)+A_2\sin(\omega t)\big] \;=\; \omega_0^2\frac{F_0}{k}\cos(\omega t)\;. \tag{6.22}$$

Collecting the $\cos(\omega t)$ terms and the $\sin(\omega t)$ in this equation together, we have

$$\left[A_1(\omega_0^2-\omega^2) + 2A_2\omega\beta - \omega_0^2\frac{F_0}{k}\right]\cos(\omega t) + \Big[A_2(\omega_0^2-\omega^2) - 2A_1\omega\beta\Big]\sin(\omega t) \;=\; 0\;. \tag{6.23}$$

We are looking for a solution that satisfies the driven-harmonic-oscillator equation for all times t. In order for this to be true of our proposed solution in Eq. (6.21), the coefficients of the $\cos(\omega t)$ and $\sin(\omega t)$ terms in Eq. (6.23) must vanish independently. Imposing this condition gives us a pair of equations:

$$\begin{aligned} A_1(\omega_0^2-\omega^2) + 2A_2\omega\beta &= \omega_0^2\frac{F_0}{k} \\ A_2(\omega_0^2-\omega^2) - 2A_1\omega\beta &= 0 \end{aligned} \tag{6.24}$$

We can solve this system of two equations for the two unknowns A_1 and A_2. The result is

$$\begin{aligned} A_1 &= \frac{\omega_0^2-\omega^2}{(\omega_0^2-\omega^2)^2+4\beta^2\omega^2}\,\omega_0^2\frac{F_0}{k} \\ A_2 &= \frac{2\beta\omega}{(\omega_0^2-\omega^2)^2+4\beta^2\omega^2}\,\omega_0^2\frac{F_0}{k}\;. \end{aligned} \tag{6.25}$$

Plugging these results back into Eq. (6.21), we find that our particular solution to the driven-harmonic-oscillator equation is

$$x_p(t) \;=\; \frac{(\omega_0^2-\omega^2)\cos(\omega t) + 2\beta\omega\sin(\omega t)}{(\omega_0^2-\omega^2)^2+4\beta^2\omega^2}\,\omega_0^2\frac{F_0}{k}\;. \tag{6.26}$$

There are several other, equivalent ways of parametrizing this particular solution. For example, we can make use of the trigonometric identities we have invoked many times before for similar purposes in order to combine the sine and cosine terms together into a single cosine term with a phase. In other words, we can use these identities to recast $x_p(t)$ in the form

$$x_p(t) \;=\; A(\omega)\cos\big[\omega t - \delta(\omega)\big]\;, \tag{6.27}$$

where the amplitude $A(\omega)$ and phase $\delta(\omega)$ for a particular driving frequency ω are given by

$$\begin{aligned} A(\omega) &= \frac{\omega_0^2}{\sqrt{(\omega_0^2-\omega^2)^2+4\beta^2\omega^2}}\,\frac{F_0}{k} \\ \delta(\omega) &= \arctan\left(\frac{2\beta\omega}{\omega_0^2-\omega^2}\right) . \end{aligned} \tag{6.28}$$

Now that we have a particular solution to Eq. (6.20), all we need in order to construct the general solution to this equation is the general solution $y_c(x)$ to the complementary equation. In this case, the complementary equation is just the damped-harmonic-oscillator equation. For the case of underdamped motion — *i.e.*, the case in which $\beta < \omega_0$ — we already know that the general solution to this equation can be written in the form

$$x_c(t) = A_c e^{-\beta t}\cos(\omega t+\phi) , \tag{6.29}$$

where we have called the initial amplitude A_c in order to distinguish it from the amplitude $A(\omega)$ in Eq. (6.27). Obtaining the general solution $x(t)$ for the driven harmonic oscillator is simply a matter of adding $x_c(t)$ and $x_p(t)$ together:

$$x(t) = A_c e^{-\beta t}\cos(\omega_1 t+\phi) + A(\omega)\cos\left[\omega t-\delta(\omega)\right] . \tag{6.30}$$

The first piece of this solution — the one that came from $y_c(x)$ — contains a decaying exponential which suppresses its contribution to $x(t)$ at times $t \gg 1/\beta$. Thus, this first piece represents a **transient** contribution to $x(t)$. By contrast, the second piece of the solution — the one that came from $x_p(t)$ has no such exponential suppression and therefore dominates at late times. This piece therefore represents the steady-state solution into which $x(t)$ eventually settles after that transient contributions have died away.

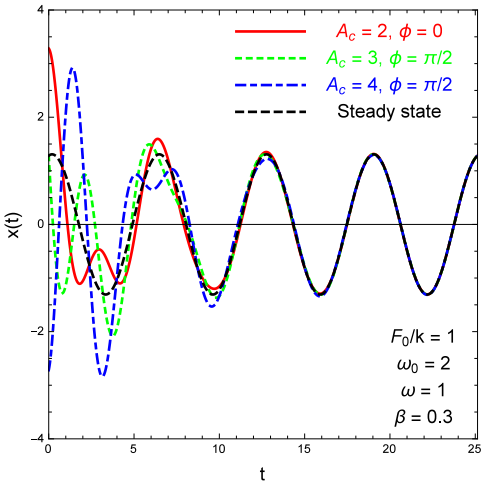


Figure 6.2: Example curves showing the behavior of the general solution $x(t)$ for a damped, driven harmonic oscillator with the parameters $\omega_0 = 2$, $\beta = 0.3$, $F_0/k = 1$, and $\omega = 1$. The solid red, green dotted, and blue dash-dotted curves correspond to different choices of initial conditions, while the black dashed curve corresponds to the steady-state solution. We see that all of the solutions shown settle into the steady state at late times.

In Fig. 6.2, we provide a few example $x(t)$ curves corresponding to different choices of initial conditions at $t = 0$ (as parametrized by the choice of A_c and ϕ). We see that indeed, regardless of their behavior at early times, all of these solutions settle into the steady state at late times.

6.4 Resonance

We can see from Eq. (6.28) that the amplitude $A(\omega)$ of the particular solution $x_p(t)$ to the driven-harmonic-oscillator equation clearly depends on the driving frequency ω in a non-trivial way. Let's explore this

dependence a little further.

$$\frac{d}{d\omega}A(\omega) \;=\; \frac{2\omega(\omega_0^2-\omega^2)-4\beta^2\omega}{\left[(\omega_0^2-\omega^2)^2+4\beta^2\omega^2\right]^{3/2}}\,\omega_0^2\frac{F_0}{k} \;=\; 0\;, \tag{6.31}$$

which reduces to

$$\omega(\omega_0^2-\omega^2+2\beta^2) \;=\; 0\;. \tag{6.32}$$

There are two solutions to this equation. One is $\omega = 0$; the other is

$$\omega_{\rm res} \;\equiv\; \sqrt{\omega_0^2-2\beta^2}\;. \tag{6.33}$$

This latter solution corresponds to a maximum of $A(\omega)$, and it occurs when the driving frequency ω coincides with a particular "natural frequency" $\omega_{\rm res}$ associated with the oscillator. This maximum is illustrated in Fig. 6.3, which shows a plot of $A(\omega)$ (normalized to the value of F_0/k) as a function of ω (in arbitrary units) for an oscillator with $\omega_0 = 1$ and $\beta = 0.2$. For values of , In this plot, we see a pronounced peak centered around $\omega_{\rm res}$ which represents a a significant enhancement to the amplitude of oscillation for frequencies $\omega \approx \omega_{\rm res}$. This is our first example of **resonance** — a phenomenon in which two frequencies coincide, resulting in a dramatic enhancement to the amplitude of oscillation. The frequency $\omega_{\rm res}$ at which this enhancement is the greatest is known as the **resonant frequency** of the system.

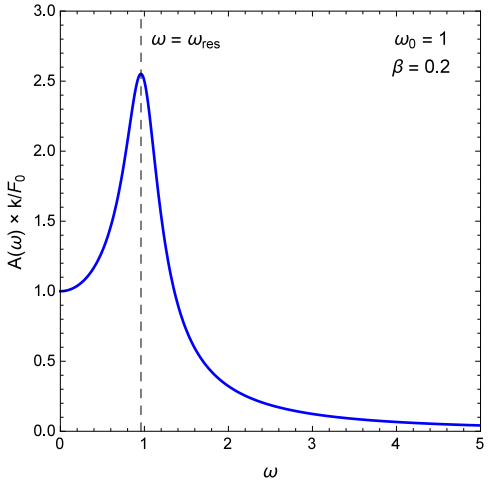


Figure 6.3: The amplitude $A(\omega)$ of the particular solution $x_p(t)$ for a driven harmonic oscillator with $\omega_0 = 1$ and $\beta = 0.2$, shown as a function of ω. The amplitude function peaks at the resonant frequency $\omega_{\rm res}$ in Eq. (6.33).

The size of the resonant enhancement to the amplitude — *i.e.*, the value of $A(\omega)$ at $\omega \approx \omega_{\rm res}$ — is quite sensitive to the value of β. Indeed, if we set $\omega = \omega_{\rm res}$ in the formula for $A(\omega)$ in Eq. (6.28), we find that

$$A(\omega_{\rm res}) \;=\; \frac{\omega_0^2}{2\beta\omega_1}\,\frac{F_0}{k}\;. \tag{6.34}$$

This means that the smaller β is, the larger the resonant enhancement to the amplitude will be. This behavior is illustrated in Fig. 6.4, which shown the $A(\omega)$ curves for a variety of different values of β with ω_0 held fixed. It's also worth remarking that when β is very small (in the sense that $\beta \ll \omega_0$) and we have $\omega_1 \approx \omega_{\rm res} \approx \omega_0$, the resonant amplitude $A(\omega)$ is approximately proportional to the quality factor Q:

$$A(\omega_{\rm res}) \;\approx\; Q\frac{F_0}{k}\;, \qquad \text{for } \beta \ll \omega_0\;. \tag{6.35}$$

One further thing to note about the expression for $A(\omega_{\rm res})$ in Eq. (6.34) is that it blows up in the $\beta \to 0$ limit. In other words, in the absence of damping, the resonant enhancement to the amplitude becomes infinite. This result can be understood as follows.

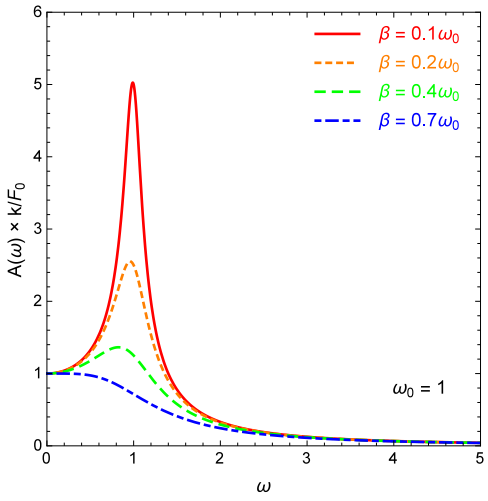


Figure 6.4: The amplitude $A(\omega)$ of the particular solution $x_p(t)$ for a driven harmonic oscillator, shown as a function of ω for a variety of different values of β. For each curve, we have once again taken $\omega_0 = 1$, as in Fig. 6.3. We see that as β becomes smaller relative to ω_0, the resonant enhancement to the amplitude at $\omega \approx \omega_{\rm res}$ becomes more pronounced.

6.5 Energy in a Driven-Oscillator System

Let's now turn to consider the energy stored in the damped, driven harmonic oscillator. We have seen that for a simple harmonic oscillator, the total energy $E_{\rm tot}$ of the system is conserved. We have seen that for a damped oscillator, $E_{\rm tot}$ decreases monotonically as a function of time due to the effect of the damping force. However, if we include a driving force, the situation changes because the driving force serves to inject energy into the system over time. Thus, in the presence of both a damping force and a driving force, there is a flow of energy both into and out of the system.

The rate at which energy flows into this system at a particular time t is described by the instantaneous power $P(t)$ delivered by the external driving force. The instantaneous power $P(t)$ delivered by a force to an object is the rate at which work is done to that object:

$$P(t) \;=\; \frac{dW}{dt} \;. \tag{6.36}$$

We can also use the definition of the work

$$W \;=\; \int F(x)dx \tag{6.37}$$

in order to write $P(t)$ as a product of the instantaneous force $F(t)$ acting on the object and its velocity:

$$P(t) \;=\; \frac{d}{dt}\int F dx \;=\; \frac{d}{dt}\int F\frac{dx}{dt}dt \;=\; F(t)v(t) \;. \tag{6.38}$$

The power delivered to the system at early times, while the transient in Eq. (6.30) dominates, is highly dependent on the initial conditions. Thus, we'll focus on the situation at late times, after the system has settled into the steady state. For the steady-state solution, the velocity $v(t)$ is just the time derivative of the expression for $x_p(t)$ in Eq. (6.27). Thus, we have

$$\begin{aligned} P(t) &= -\omega A(\omega)\sin[\omega t - \delta(\omega)]F_0\cos(\omega t) \\ &= \frac{1}{2}F_0\omega A(\omega)\Big[\sin\delta - \sin(2\omega t - \delta)\Big] \;, \end{aligned} \tag{6.39}$$

where in going from the first to the second line, we have used the identity

$$\sin\alpha\cos\beta \;=\; \frac{1}{2}\Big[\sin(\alpha-\beta)+\sin(\alpha+\beta)\Big]\;. \tag{6.40}$$

We see from Eq. (6.39) that even in the steady state, the instantaneous power $P(t)$ is time-dependent. However, this is to be expected, since the driving force itself oscillates with an angular frequency ω. For this reason, it's often more revealing to look at the average power $\langle P\rangle_T$ delivered to the system over a period of oscillation. The average power delivered over a period is just

$$\langle P\rangle_T \;=\; \frac{1}{T}\int_0^T P(t)dt\;. \tag{6.41}$$

Plugging our expression for $P(t)$ in Eq. (6.39) into this equation yields

$$\langle P\rangle_T \;=\; \frac{1}{2}F_0\omega A(\omega)\frac{1}{T}\int_0^T\Big[\sin\delta-\sin(2\omega t-\delta)\Big]dt \;=\; \frac{1}{2}F_0\omega A(\omega)\sin\delta\;. \tag{6.42}$$

This expression for $\langle P\rangle_T$ is certainly very compact, but its not terribly revealing. Therefore, in order to get a better sense of how this quantity depends on the fundamental, underlying parameters ω, ω_0, and β, let's plug our expressions for $A(\omega)$ and $\delta(\omega)$ in Eq. (6.28) into this formula. Using the identity

$$\sin(\arctan x) \;=\; \frac{x}{\sqrt{1+x^2}} \tag{6.43}$$

to simplify the result, we find that

$$\begin{aligned}\langle P\rangle_T &= \frac{F_0\omega}{2}\frac{F_0}{k}\frac{\omega_0^2}{\sqrt{(\omega^2-\omega_0^2)^2+4\beta^2\omega^2}}\frac{\frac{2\beta\omega}{\omega_0^2-\omega^2}}{\sqrt{1+\left(\frac{2\beta\omega}{\omega_0^2-\omega^2}\right)^2}}\\ &= \frac{F_0^2}{2k}\frac{2\beta\omega^2\omega_0^2}{(\omega^2-\omega_0^2)^2+4\beta^2\omega^2}\end{aligned} \tag{6.44}$$

This result indicates that the average power delivered by the driving force to the oscillator is constant in the steady state. We also note that this result can be expressed in terms of the the quality factor $Q=\omega_0/2\beta$ rather than the damping parameter β. In fact the resulting expression even turns out to be a little cleaner:

$$\langle P\rangle_T \;=\; \frac{F_0^2\omega_0}{2k}\frac{Q\omega^2\omega_0^2}{\omega^2\omega_0^2+Q^2(\omega^2-\omega_0^2)^2}\;. \tag{6.45}$$

While Eq. (6.44) indicates that the driving force is constantly transferring energy to the oscillator while the system is in the steady state, this does not mean that the total energy $E_{\rm tot}$ of the system is increasing. Indeed, as we saw in Eq. (6.28) the oscillation amplitude $A(\omega)$ is constant in the steady state — and thus so is $E_{\rm tot}$. It's important to keep in mind that a driven, damped harmonic oscillator is also constantly losing energy to its surroundings as a result of the damping force. At early times, the rate at which energy is injected into the system by the driving force (*i.e.*, the power delivered) and the rate at which energy is dissipated by the damping force can differ, and as a result there can be a net gain or loss in $E_{\rm tot}$. However, as the system settles into the steady state, these two rates become equal, and as a result, $E_{\rm tot}$ becomes effectively constant.

Our formula for $\langle P\rangle_T$ in Eq. (6.44) makes it clear that the average power delivered to the a driven oscillator in the steady state is constant in time. However, it also tells us that this average power delivered to the system depends sensitively of ω. This dependence is illustrated in Fig. 6.5, where we show $\langle P\rangle_T$ as a function of ω for $\omega_0=1$ and $\beta=0.2$. We see from the figure that like the steady-state oscillation amplitude $A(\omega)$, the average power $\langle P\rangle_T$ delivered to the system also receives a resonant enhancement for values of ω near some particular, resonant frequency. This frequency is not hard to determine. Simply by inspecting

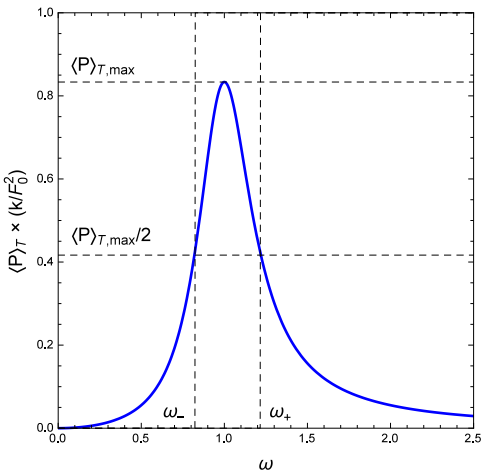


Figure 6.5: The average power $\langle P\rangle_T$ over a cycle for a damped, driven harmonic oscillator as a function of the driving frequency ω. In this plot, we have taken $\omega_0 = 1$ and $\beta = 0.2$. The two frequencies ω_+ and ω_- at which $\langle P\rangle_T = \langle P\rangle_{T,\max}/2$ are also indicated.

the denominator of Eq. (6.45), we can see that $\langle P\rangle_T$ is maximized for $\omega = \omega_0$. The corresponding value of $\langle P\rangle_T$ at this resonant frequency is

$$\langle P\rangle_{T,\max} \;=\; \frac{F_0^2\omega_0 Q}{2k} \;=\; \frac{F_0^2\omega_0^2}{4k\beta}\;. \tag{6.46}$$

It's also important to keep in mind that these resonance effects are felt not merely when the driving frequency ω is precisely equal to ω_0, but also over a range of nearby values where $\omega \approx \omega_0$ — the resonant frequency just represents the frequency at which this effect is maximized. It is therefore pertinent to ask exactly how far away ω can be from ω_0 before the resonance effect ceases to be important. In other words, we don't just want to know about the *height* of the resonance peak in Fig. 6.5, but about its *width* as well.

There are a variety of ways of characterizing how broad a resonance peak is. However, one of the most commonly used conventions is to quote the **full width at half maximum** (FWHW). The full width at half maximum is the is the difference

$$(\Delta\omega)_{\rm FWHM} \;\equiv\; \omega_+ - \omega_- \tag{6.47}$$

between the two frequencies $\omega_\pm$ on either side of the peak at which $\langle P\rangle_T = \langle P\rangle_{T,\max}/2$.[1] These frequencies are indicated in Fig. 6.5. It can be shown (see Problem 5) that the values of ω_+ and ω_- in this case are

$$\omega_\pm \;=\; \sqrt{\beta^2 + \omega_0^2} \pm \beta \tag{6.48}$$

so the width of the peak is

$$(\Delta\omega)_{\rm FWHM} \;=\; 2\beta\;. \tag{6.49}$$

Thus, we see that as the damping parameter β increases (or, equivalently, as the quality factor $Q = \omega_0/2\beta$ decreases for a fixed value of ω_0), the width of the resonance peak increases. However, at the same time, increasing β also decreases the height of the resonance peak, as indicated in Eq. (6.46). The net result of these two effects on $\langle P\rangle_T$ is illustrated in Fig. 6.6

6.6 Complexification

There is an alternative technique for obtaining our particular solution to the damped-harmonic-oscillator equation. This technique can be used in conjunction with the method of undetermined coefficients to find

[1] It should be emphasized here that the use of FWHM as a standard for characterizing the width of a resonance peak is a common convention in a wide variety (no pun intended) of fields from particle physics to astronomy to electrical engineering.

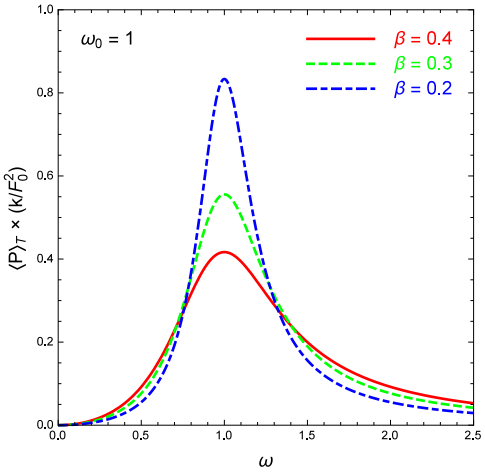


Figure 6.6: The average power $\langle P \rangle_T$ over a cycle for a damped, driven harmonic oscillator as a function of the driving frequency ω for several different choices of the damping parameter β. We have taken $\omega_0 = 1$ for all of the curves shown.

particular solutions to differential equations of the form given in Eq. (6.5) in cases where $f(x)$ is a sine or cosine function. This technique is called **complexification**, and it's typically more efficient than the method we used for obtaining $x_p(t)$ in Sect. 6.3.

As we mentioned above, complexification is a procedure which is applicable in cases where the inhomogeneous term $f(x)$ in the differential equation we're trying to solve takes the form of a since or cosine function. The first step in this procedure is to replace the function $f(x)$ in our differential equation with a complex function $\widetilde{f}(x)$. In particular, we define $\widetilde{f}(x)$ such that $f(x)$ is the real part of this complex function:

$$f(x) \;=\; \mathrm{Re}[\widetilde{f}(x)] \;. \tag{6.50}$$

For example, let's say that we were dealing with a differential equation where the inhomogeneous term $f(x)$ was just a single cosine function of the form $f(x) = f_0 \cos(\lambda t)$, where λ and f_0 are both constants. In this case, we would make the replacement

$$f(x) \;=\; f_0 \cos(\lambda t) \quad \longrightarrow \quad \widetilde{f}(x) \;=\; f_0 e^{i\lambda x} \;. \tag{6.51}$$

Likewise, if we had $f(x) = f_0 \sin(\lambda t)$, we wold make the replacement

$$f(x) \;=\; f_0 \sin(\lambda x) \quad \longrightarrow \quad \widetilde{f}(x) \;=\; -\,i f_0 e^{i\lambda x} \;. \tag{6.52}$$

Replacing $f(x)$ by the complex function $\widetilde{f}(x)$ turns our original, real differential equation into a complex equation. It therefore makes sense that the solutions to this differential equation will also be complex functions. We'll use the symbol $\widetilde{y}(x)$ for these functions to emphasize the fact that they're complex. However, these $\widetilde{y}(x)$ aren't just any old complex functions. We know that the real part of the complex equation

$$a\frac{d^2\widetilde{y}}{dx^2} + b\frac{d\widetilde{y}}{dx} + c\widetilde{y} \;=\; \widetilde{f}(x) \tag{6.53}$$

for $\widetilde{y}(x)$ is just the original, real differential equation we were trying to solve.[2] Thus, the function $y(x)$ which solves our original differential equation must be the real part of $\widetilde{y}(x)$:

$$y(x) \;=\; \mathrm{Re}[\widetilde{y}(x)] \;. \tag{6.54}$$

[2] As we showed in Chapter 3 of these lecture notes, taking the real part of a complex function and taking the derivative of that function *commute*, meaning that you'll get the same result regardless of the order in which you perform these operations.

This means that if we can find a particular solution $\widetilde{y}_p(x)$ to the complex differential equation in Eq. (6.53), all we need to do in order to obtain the corresponding particular solution $y_p(x)$ for our original equation is to take the real part of $\widetilde{y}_p(x)$.

So how do we find our particular solution to the complex equation in Eq. (6.53)? The easiest way is to use the method of undetermined coefficients. The only difference is that now our undetermined coefficients will in general be complex. We are assuming here that $f(x)$ is a sine or cosine function, and thus that $\widetilde{f}(x)$ is the complex exponential, as in the examples in Eqs. (6.51) and (6.51). For a function $\widetilde{f}(x)$ of this form, our intuition — or a glance at our table of educated guesses in Sect. 6.2 — tells us that $\widetilde{y}_p(x)$ will also have the form of a complex exponential:

$$\widetilde{y}_p(x) \;=\; \widetilde{A}e^{i\lambda x} \;, \tag{6.55}$$

where we have written the coefficient $\widetilde{A}$ with a tilde to emphasize that it is in general complex. Solving our complexified differential equation is now just a matter of plugging this expression into Eq. (6.53) and solving for the undetermined coefficient $\widetilde{A}$. Once we've done this, we simply take the real part of $\widetilde{y}_p(x)$ to get $y_p(x)$.

You may wonder why anyone would want to bother going through this complexification procedure in order to obtain $y_p(x)$ when we know perfectly well how to obtain $y_p(t)$ directly from the real differential equation. The reason is that exponential functions are easier to deal with than trigonometric functions when one is taking derivatives. Indeed, the derivative of $e^{i\lambda x}$ with respect to x is proportional to $e^{i\lambda x}$ itself, while the derivative of $\cos(\lambda x)$ is proportional to $\sin(\lambda x)$ and vice versa. This means that solving for $\widetilde{y}(x)$ is often much easier and faster than solving for $y(x)$ directly because there's only one functional form you need to keep track of.

Example: Driven LRC Circuit

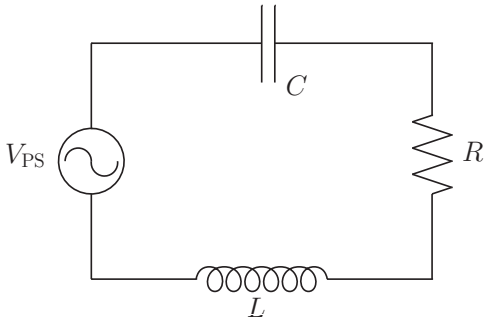


Figure 6.7: An LRC circuit hooked up to an AC power supply which delivers a driving voltage $V_{\rm PS} = V_0\cos(\omega t)$.

In order to gain a better sense of how to implement this technique in practice, let's walk through an explicit example. Consider the LRC circuit shown in Fig. 6.7. The circuit includes an AC power supply which delivers a voltage $V_{\rm PS} = V_0\cos(\omega t)$. We have studied circuits like this before. Indeed, we have used the method of complex impedances to find the steady-state solution for the charge $Q(t)$ on the capacitor, the current $I(t)$ in the circuit, *etc.*However, we have not yet derived a general solution for these quantities which includes transient solutions as well as the steady-state piece. Let's now use complexification in conjunction with the method of undetermined coefficients to obtain such a general solution.

Applying Kirchhoff's loop rule for this circuit gives

$$V_0\cos(\omega t) - R\frac{dQ}{dt} - L\frac{d^2Q}{dt^2} - \frac{Q}{C} \;=\; 0 \;, \tag{6.56}$$

which we can rearrange into

$$\frac{d^2Q}{dt^2} + \frac{R}{L}\frac{dQ}{dt} + \frac{Q}{LC} \;=\; \frac{V_0}{L}\cos(\omega t) \;. \tag{6.57}$$

This is a second-order linear differential equation with constant coefficients and an inhomogeneous term of the form $f(x) = f_0 \cos(\lambda t)$, so we can solve for $Q(t)$ by complexifying the equation and using the method of undetermined coefficients to find a solution for the complexified charge $\widetilde{Q}(t)$. Our complexified equation is

$$\frac{d^2\widetilde{Q}}{dt^2} + \frac{R}{L}\frac{d\widetilde{Q}}{dt} + \frac{\widetilde{Q}}{LC} \;=\; \frac{V_0}{L}e^{i\omega t} \;, \tag{6.58}$$

and our intuition tells us that that our solution is going to have the form

$$\widetilde{Q}(t) \;=\; \widetilde{A}e^{i\omega t} \;. \tag{6.59}$$

Plugging this ansatz for $\widetilde{Q}(t)$ into Eq. (6.58) gives us

$$-\widetilde{A}\omega^2 e^{i\omega t} + i\widetilde{A}\frac{\omega R}{L}e^{i\omega t} + \widetilde{A}\frac{1}{LC}e^{i\omega t} \;=\; \frac{V_0}{L}e^{i\omega t} \;. \tag{6.60}$$

Since $e^{i\omega}$ appears in every term, we can cancel this factor and then multiply both sides of the equation by LC to get

$$-\widetilde{A}LC\omega^2 + i\widetilde{A}RC\omega + \widetilde{A} \;=\; V_0 C \;. \tag{6.61}$$

The equation in Eq. (6.61) is a complex equation. It is equivalent to two real equations — one from the real part of the equation and one from the imaginary part. Likewise, the complex parameter $\widetilde{A} = \mathrm{Re}[\widetilde{A}] + i\mathrm{Im}[\widetilde{A}]$ embodies two unknowns: the real part $\mathrm{Re}[\widetilde{A}]$ and the imaginary part $\mathrm{Im}[\widetilde{A}]$. Specifically, the equations that we get from Eq. (6.61) for these two unknowns are

$$\begin{aligned} \mathrm{Re}[\widetilde{A}](1 - LC\omega^2) - \mathrm{Im}[\widetilde{A}]RC\omega &= V_0 C \\ \mathrm{Im}[\widetilde{A}](1 - LC\omega^2) + \mathrm{Re}[\widetilde{A}]RC\omega &= 0 \;. \end{aligned} \tag{6.62}$$

Solving this system of equations for $\mathrm{Re}[\widetilde{A}]$ and $\mathrm{Im}[\widetilde{A}]$, we find that

$$\begin{aligned} \mathrm{Re}[\widetilde{A}] &= V_0 C\,\frac{1 - (LC\omega^2)}{(RC\omega)^2 + (1 - LC\omega^2)^2} \\ \mathrm{Im}[\widetilde{A}] &= -V_0 C\,\frac{RC\omega}{(RC\omega)^2 + (1 - LC\omega^2)^2} \;. \end{aligned} \tag{6.63}$$

These results indicate that that the value of the complex coefficient $\widetilde{A}$ depends on the driving frequency ω (but *not* on t). To emphasize this, we'll make the ω dependence of $\widetilde{A}$ explicit from this point forward by writing it as $\widetilde{A}(\omega)$. Putting the real and imaginary pieces from Eq. (6.63), we have

$$\widetilde{A}(\omega) \;=\; \frac{V_0 C\Big[(1 - LC\omega^2) - iRC\omega\Big]}{(RC\omega)^2 + (1 - LC\omega^2)^2} \;. \tag{6.64}$$

Plugging this result into Eq. (6.59) and using Euler's theorem to expand the complex exponential gives us

$$\begin{aligned} \widetilde{Q}_p(t) &= -\frac{V_0 C\Big[(LC\omega^2 - 1) + iRC\omega\Big]}{(RC\omega)^2 + (1 - LC\omega^2)^2}\Big[\cos(\omega t) + i\sin(\omega t)\Big] \\ &= \frac{V_0 C\Big[(1 - LC\omega^2)\cos(\omega t) + RC\omega\sin(\omega t)\Big]}{(RC\omega)^2 + (1 - LC\omega^2)^2} + i\,\frac{V_0 C\Big[(1 - LC\omega^2)\sin(\omega t) - RC\omega\cos(\omega t)\Big]}{(RC\omega)^2 + (1 - LC\omega^2)^2} \end{aligned} \tag{6.65}$$

The corresponding solution $Q_p(t)$ to our original differential equation in Eq. (6.57) is just the real part of this expression:

$$Q(t) \;=\; \mathrm{Re}[\widetilde{Q}(t)] \;=\; \frac{V_0 C}{(RC\omega)^2 + (LC\omega^2 - 1)^2}\Big[(1 - LC\omega^2)\cos(\omega t) + RC\omega\sin(\omega t)\Big] \;. \tag{6.66}$$

There is actually an interesting connection between this particular solution to Eq. (6.57) and the solution we would have obtained if we had used the method of complex impedances. This connection is worth exploring in a little bit more detail. If we had used the method of complex impedances, we would have begun by writing down the complex version of our Kirchoff's-loop-rule equation

$$\widetilde{V}_{\rm PS} - \widetilde{I}Z_R - \widetilde{I}Z_L - \widetilde{I}Z_C \;=\; 0 \;, \tag{6.67}$$

where $\widetilde{V}_{\rm PS} = V_0 e^{i\omega t}$ and $\widetilde{I}$ is the complex current. Solving this equation for $\widetilde{I}$, we obtain

$$\widetilde{I} \;=\; \frac{\widetilde{V}_{\rm PS}}{Z_R + Z_L + Z_C} \;=\; \frac{V_0 e^{i\omega t}}{R + iL\omega - \frac{i}{\omega C}} \;. \tag{6.68}$$

The real current $I(t)$ is just the real part of $\widetilde{I}$. It's easiest to identify the real part of a complex quantity when that quantity is written in rectangular form. Do rewrite $\widetilde{I}$ in this form, we begin by multiplying and dividing the expression in Eq. (6.68) by the complex conjugate of the denominator and expanding the complex exponential using Euler's theroem:

$$\widetilde{I} \;=\; V_0 \frac{R - i\left(L\omega - \frac{1}{\omega C}\right)}{R^2 + \left(L\omega - \frac{1}{\omega C}\right)^2} \big[\cos(\omega t) + i\sin(\omega t)\big] \;. \tag{6.69}$$

Thus, we find that

$$\begin{aligned} \widetilde{I} \;&=\; V_0 C\omega \, \frac{RC\omega - i(LC\omega^2 - 1)}{(RC\omega)^2 + (1 - LC\omega^2)^2} \big[\cos(\omega t) + i\sin(\omega t)\big] \\ &=\; V_0 C\omega \, \frac{RC\omega\cos(\omega t) - (1 - LC\omega^2)\sin(\omega t)}{(RC\omega)^2 + (1 - LC\omega^2)^2} + iV_0 C\omega \, \frac{RC\omega\sin(\omega t) + (1 - LC\omega^2)\cos(\omega t)}{(RC\omega)^2 + (1 - LC\omega^2)^2} \;. \end{aligned} \tag{6.70}$$

Since the real current $I(t)$ is just the real part of this expression, we have

$$I(t) \;=\; V_0 C\omega \, \frac{RC\omega\cos(\omega t) - (1 - LC\omega^2)\sin(\omega t)}{(RC\omega)^2 + (1 - LC\omega^2)^2} \;. \tag{6.71}$$

Finally, since $Q(t)$ is just the integral of $I(t)$ with respect to t, we integrate this result to obtain our final expression for the charge on the capacitor as a function of time:[3]

$$Q(t) \;=\; \int_0^t I(t')dt' \;=\; V_0 C\omega \, \frac{RC\omega\sin(\omega t) + (LC\omega^2 - 1)\cos(\omega t)}{(RC\omega)^2 + (1 - LC\omega^2)^2} \;. \tag{6.72}$$

This expression is identical to the solution $Q_p(t)$ in Eq. (6.66) that we obtained from our complexification procedure in conjuction with the method of undetermined coefficients!

6.7 The Principle of Superposition

The educated guesswork that underlies the method of undetermined coefficients is certainly useful for solving differential equations where the inhomogeneous term $f(x)$ has one of the particular functional forms appearing in Table 6.1. However, we don't yet have a good method for finding $y_p(x)$ in cases where $f(x)$ doesn't happen to have one of these functional forms. For example, let's say that $f(x)$ includes both a cosine piece and a polynomial piece. What does the solution for $y(x)$ look like in that case?

It turns out that for any linear differential equation in which the inhomogeneous term $f(x)$ is a direct sum of several individual functions $f_i(x)$ — *i.e.*, where

$$f(x) \;=\; \sum_i^N f_i(x) \;, \tag{6.73}$$

[3] If you're worried about the fact that there's no constant of integration in Eq. (6.72), don't be. We're not interested in applying boundary conditions yet. At this point, we're just looking for a *particular* solution which we can use in order to construct the *general* solution for $Q(t)$. Any particular solution will do, regardless of what the value of $Q_p(t)$ happens to be at $t = 0$. After we have our general solution, *then* it's time to apply boundary conditions.

where N is the number of functions in the sum — there is actually a general principle that we can use in order to obtain a particular solution $y_p(x)$ to that equation. This principle is called the **principle of superposition**. It states that whenever $f(x)$ takes this form, there exists a particular solution of the form

$$y_p(x) \;=\; \sum_{i=1}^{N} y_{pi}(x) \;, \tag{6.74}$$

where each of the N functions $y_{pi}(x)$, *etc.*, is the particular solution to the differential equation with $f(x)$ replaced by $f_i(x)$. For example, if we had a second-order differential equation with the general form given in Eq. (6.4), the $y_p(x)$ would be the solutions to the equations

$$P(x)\frac{d^2y_{pi}}{dx^2} + Q(x)\frac{dy_{pi}}{dx} + R(x)y_{pi} \;=\; f_i(x) \;. \tag{6.75}$$

The principle of superposition is actually not terribly difficult to prove. We'll begin by demonstrating that it works for a second-order equation in the $N = 2$ case in which $f(x) = f_1(x) + f_2(x)$. In this case, Eq. (6.74) says that $y_p(x)$ is the sum of two functions $y_{p1}(x)$ and $y_{p2}(x)$. Plugging this into the general expression in Eq. (6.4), we find that

$$\begin{aligned} f_1(x) + f_2(x) \;&=\; P(x)\frac{d^2}{dx^2}(y_{p1}+y_{p2}) + Q(x)\frac{d}{dx}(y_{p1}+y_{p2}) + R(x)(y_{p1}+y_{p2}) \\ &=\; \left[P(x)\frac{d^2y_{p1}}{dx^2} + Q(x)\frac{dy_{p1}}{dx} + R(x)y_{p1}\right] + \left[P(x)\frac{d^2y_{p2}}{dx^2} + Q(x)\frac{dy_{p2}}{dx} + R(x)y_{p2}\right] \end{aligned} \tag{6.76}$$

This equality will certainly be satisfied if the first term in brackets on the right-hand side is equal to $f_1(x)$ and the second term in brackets is equal to $f_2(x)$. (There may be other solutions too, but we're just looking for one particular solution that works.) Thus, in this case, the principle of superposition holds. Moreover, it's pretty clear that adding more terms to $f(x)$ won't qualitatively change anything: for each additional $f_i(x)$ that we add to the left-hand side of Eq. (6.76), a corresponding term in $y_{pi}(x)$ will appear on the right-hand side. Furthermore, generalizing this result to linear differential equations of any arbitrary order is straightforward. This is because the derivative of a sum of functions is equal to the sum of the derivatives of those functions. Thus, for a derivative of any order, we have

$$\frac{d^n}{dx^n}\sum_{i=1}^{N} y_{pi} \;=\; \sum_{i=1}^{N}\frac{d^n y_{pi}}{dx^n} \;. \tag{6.77}$$

This means that as long as the differential equation we're dealing with is *linear* — *i.e.*, that all of the terms in that equation are proportion to $y(x)$ or to one of its derivatives — we can group terms in exactly the same way that we grouped them in Eq. (6.76).

The principle of superposition is actually an incredibly powerful result. Indeed, as we shall see before too long, it's the underlying principle behind two methods which allow us to solve a inhomogeneous linear differential equations for any arbitrary functional form that $f(x)$ might happen to take. These methods are Fourier decomposition and the method of Green's functions, and they are the subject of the next two chapters of these lecture notes.

Problems

1. Consider the differential equation

$$3\frac{d^2y}{dx^2} + 2\frac{dy}{dx} \;=\; -\,e^{3x} \;. \tag{6.78}$$

   (a) Find a particular solution to this equation.

   (b) Find the general solution to the equation.

   (c) Find the solution to the equation for the boundary conditions $y(0) = 0$ and $dy/dx = 0$ at $x = 0$.

2. Find a particular solution to the differential equation

$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + 17y \;=\; 60e^{-4x}\sin(5x)\;. \tag{6.79}$$

3. For a damped, driven harmonic oscillator in the steady state, the acceleration $a(t)$ of the oscillating object as a function of time is given by

$$a(t) \;=\; \frac{d^2x}{dt^2} \;=\; -\,\omega^2 A(\omega)\cos[\omega t - \delta(\omega)]\;. \tag{6.80}$$

   Let us define the "acceleration amplitude" $a(\omega) = \omega^2 A(\omega)$, which represents the maximum acceleration the object attains as a function of the driving frequency ω. What is the value of ω for which this acceleration amplitude is maximized?

4. In Sect. 6.5, we showed that the average power $\langle P\rangle_T$ delivered to a damped, driven oscillator by the driving force in the steady state was given by Eq. (6.45). Show that the average power $\langle P_{\rm drag}\rangle_T$ dissipated by the drag force in the steady state is equal in magnitude to $\langle P\rangle_T$.

5. Derive the expression for $\omega_\pm$ in Eq. (6.48).

6. Find a particular solution to the differential equation

$$\frac{d^2y}{dx^2} + y \;=\; \sinh x\;. \tag{6.81}$$

7. In this problem, you are going to investigate the case of an *undamped*, driven oscillator — *i.e.*, a simple harmonic oscillator oscillator with a driving force, but with $\beta = 0$. The equation of motion for such an oscillator is

$$\frac{d^2x}{dt^2} + \omega_0^2 x \;=\; \omega_0^2\frac{F}{k}\cos(\omega t)\;. \tag{6.82}$$

   (a) Find a particular solution $x_p(t)$ to this equation for general ω. What happens to $x_p(t)$ when $\omega = \omega_0$?

   (b) The case in which $\omega_0 = \omega$ is a special case in which the solution in part (a) does not apply. Show that the function

$$x_p(t) \;=\; \frac{F_0}{4k}\big[\cos(\omega_0 t) + 2t\omega_0\sin(\omega_0 t)\big] \tag{6.83}$$

   is a particular solution to the undamped, driven oscillator equation for $\omega = \omega_0$.

   (c) Use the particular solutions you found in parts (a) and (b) to construct the general solution $x(t)$ both for the general case in which $\omega \neq \omega_0$ and for the special case in which $\omega = \omega_0$. Express your answer in each case in terms of the initial position x_0 and initial velocity v_0 of the oscillating object.

   (d) Now let's see what these solutions actually look like for a give choice of parameters. In particular, as an example, let's take $\omega_0 = 1$ s^{-1}, $F_0 = 1$ N, and $k = 1$ N/m. Use Mathematica to plot the $x(t)$ curves for $\omega = 0.7$ s^{-1}, for $\omega = 0.9$ s^{-1}, for $\omega = 0.95$ s^{-1}, and for $\omega = 1$ s^{-1} together on the same plot for the boundary conditions $x_0 = 0$ and $v_0 = 0$. Make a similar plot for the boundary conditions $x_0 = 40$ m, $v_0 = 0$.

   (e) Qualitatively speaking, what happens to $x(t)$ for $\omega \neq \omega_0$ in the limit as $t \to \infty$? What happens to the solution for $\omega = \omega_0$ in this same limit?

8. In this problem, we're going to compare the solution to the damped, driven oscillator equation with a driving force $F(t) = F_0\cos(\omega t)$ to the solution we get from the corresponding equation with $F(t)$ approximated by a Maclaurin series.

   (a) Find the Maclaurin-series approximation for $F(t)$ as a function of t, retaining terms up to and including the t^2 term.

(b) Find the particular solution $x_{p,\text{approx}}(t)$ to the driven, damped oscillator equation with the exact functional form for $F(t)$ replaced by your Taylor-series approximation from part (a).

(c) Use Mathematica to plot your solution $x_{p,\text{approx}}(t)$ from part (b) together with the particular solution $x_p(t)$ to the equation with the exact $F(t) = F_0 \cos(\omega t)$. For your plot, take $F_0 = 1$ N, $k = 1$ N/m, and $\omega_0 = 1$ s^{-1}, with $\beta = 0.2\,\omega_0$ and $\omega = 0.5\,\omega_0$.

(d) For this choice of parameters, how long does take (in units of the period $T = 2\pi/\omega$) before $x_{p,\text{approx}}(t)$ differs from $x_p(t)$ by 10%? (Note: you may use Mathematica to answer this if you wish.) The lesson here is that Taylor series are not necessarily the best way of approximating periodic functions. You typically need to keep a significant number of terms in order for the solution to remain accurate over even a few periods of oscillation. As we shall see in the next chapter, there are far better ways of approximating periodic functions!

Chapter 7

Fourier Analysis

- **The physics:** Driven oscillations with periodic driving functions
- **The math:** Fourier series, Fourier decomposition, orthogonal functions

7.1 Superposition and the Decomposition of Functions

In Chapter 6 of these lecture notes, we discussed some of the techniques which can be used to solve inhomogeneous linear differential equations. One of these techniques was based upon the principle of superposition — a principle which becomes relevant when the inhomogeneous term $f(x)$ takes the form of a direct sum. This principle states that if $f(x) = f_1(x) + f_2(x) + \ldots$, then $y_p(x) = y_{p1} + y_{p2} + \ldots$ is a particular solution to the differential equation, where $y_{p1}(x)$ is the solution to the corresponding equation with $f(x)$ replaced by $f_1(x)$, where $y_{p2}(x)$ is the solution to the corresponding equation with $f(x)$ replaced by $f_2(x)$, and so on.

This principle of superposition might initially seem of only limited applicability. Indeed, in order for us to put it to use, not only must $f(x)$ take the form of a direct sum, but we must also be able to find the solution $y_{pi}(x)$ for each of the individual functions $f_i(x)$ appearing in that sum. However, as we shall see both in this chapter and in the next, there are general methods for writing any arbitrary function $f(x)$ as a sum of functions $f_i(x)$ for which obtaining the corresponding $y_{pi}(x)$ is straightforward.

7.2 Fourier Series

Let's begin by examining the case in which our function $f(t)$ is periodic in the independent variable t. In other words, after some fixed interval T, the function $f(t)$ repeats itself, and we have

$$f(t+T) \;=\; f(t)\;. \tag{7.1}$$

It can be shown that nearly any periodic function of this sort can be expressed as an infinite sum of sine and cosine functions angular whose angular frequencies which are integer multiples of the fundamental angular frequency $\omega = 2\pi/T$ associated with the period T. In particular, it can be shown that

$$f(t) \;=\; \sum_{n=0}^{\infty} \Big[a_n \cos(n\omega t) + b_n \sin(n\omega t) \Big]\;, \tag{7.2}$$

where a_n and b_n are constant coefficients for each term. These coefficients represent the amplitudes of the individual cosine and sine terms in the series, respectively. This way of representing a function is called a **Fourier series**, and the a_n and b_n are called **Fourier coefficients**. A given function $f(x)$ is specified by a particular set of values for these coefficients. The process of decomposing a function into a set of sine and cosine terms in this way is often referred to as **Fourier decomposition**.

The assertion that practically any periodic function can be represented as a sum of sine and cosine waves of different frequencies might at first seem somewhat surprising. Therefore, in order to illustrate how these

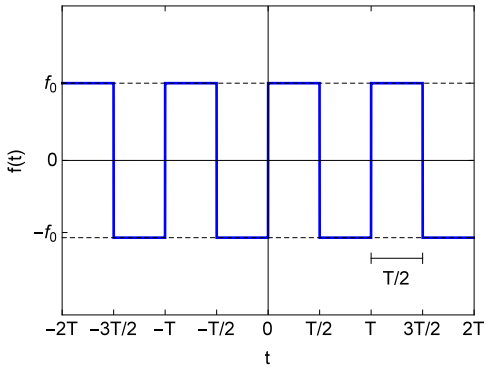


Figure 7.1: A square wave with amplitude f_0 and period T.

sine and cosine waves conspire to reproduce a given function, we'll begin by looking a a concrete example. In particular, let's examine how this works for the square wave shown in Fig. 7.1. For the moment, we're just interested in getting a sense of how Fourier decomposition works qualitatively. We'll deal with the quantitative aspects of Fourier analysis — including how to explicitly calculate the coefficients a_n and b_n for any arbitrary function.

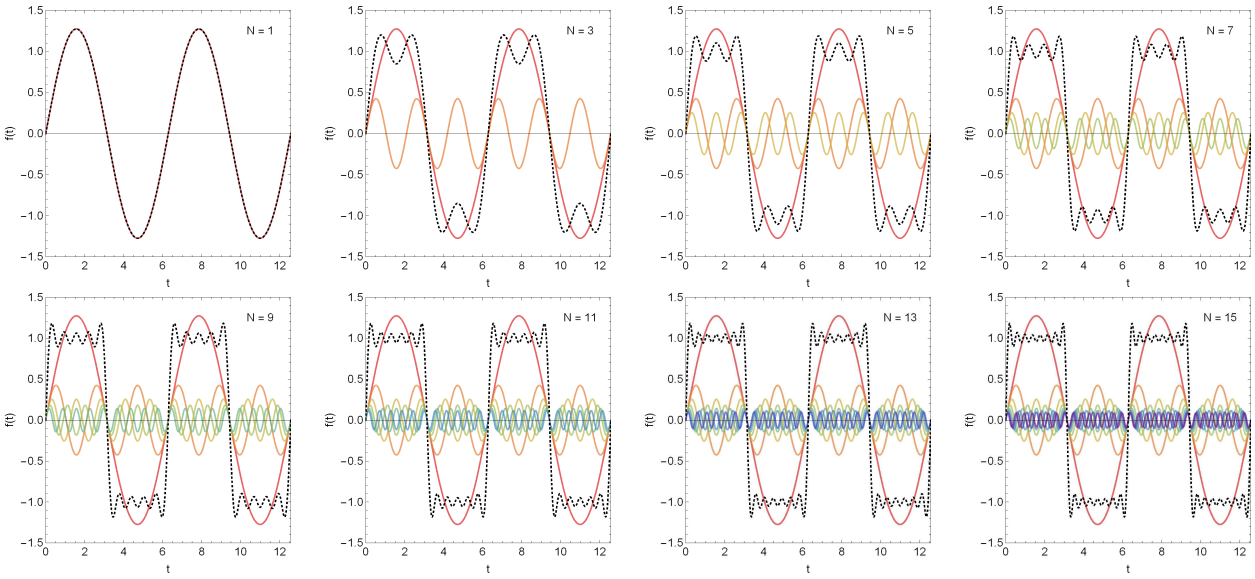


Figure 7.2: In each panel of this figure, we show the first N individual terms (solid curves) in the Fourier-series representation for a square wave with period $T = 2\pi$. In addition, we also show the sum (black dashed curve) of these N contributions. As more and more terms are included, this partial sum looks more and more like a square wave. In the limit as $N \to \infty$, the sum converges to the square wave.

In Fig. 7.2, we illustrate how the individual terms in the Fourier series conspire to reproduce our original square wave. Each panel of this figure shows the individual terms in the Fourier series for which $n \leq N$ for different values of N. In addition, in each panel, we also display the curve curve which represents the sum of these individual contributions. In general, the sum

$$f(t) \;=\; \sum_{n=0}^{N} \Big[a_n \cos(n\omega t) + b_n \sin(n\omega t) \Big] \tag{7.3}$$

of terms in a Fourier series up to some finite value of N is called a **partial sum**. As we shall verify explicitly in Sect. 7.4, all of the coefficients a_n vanish for this particular function, so the Fourier series for our square wave consists only of sine terms.[1] Moreover, it turns out that the b_n for even values of n also vanish, which is why we have included plots only for odd values of N.

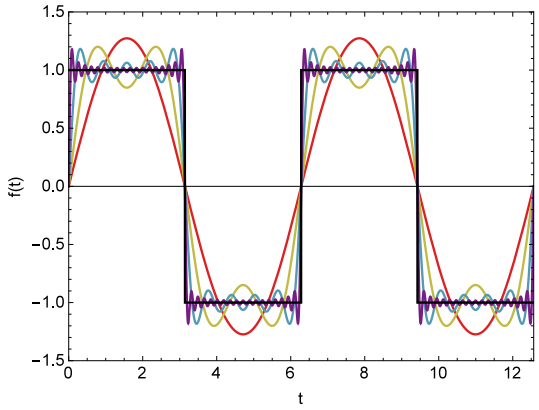


Figure 7.3: Curves corresponding to partial sums in the Fourier expansion of a square wave for $N = 1$ (red curve), $N = 3$ (yellow curve), $N = 9$ (light blue curve), and $N = 35$ (purple curve). As more and more terms in the Fourier series are included, the partial sum becomes a better and better approximation to the square wave. In the $n \to \infty$ limit, the series converges to the value of the function.

We see from Fig. 7.2 that when N is small and only a few terms are included, the resulting curve doesn't look much like a square wave. However, when more and more terms are included in the sum, the resemblance becomes more and more apparent. In the $N \to \infty$ limit, the sum converges to the square wave. In Fig. 7.3, we compare the curve for the partial sums with $N = 1$, $N = 3$, $N = 9$, and $N = 35$ to the square-wave function itself on the interval $0 \le t \le 2T$. We see from the figure that the partial sums which include only the first few terms bear only a vague resemblance to the original function, but the partial sum for $N = 35$ already provides a very good approximation to a square wave.

When we said above that "nearly any" periodic function $f(t)$ can be expressed as a Fourier series, we should probably be a bit more explicit about what we meant. Not every periodic function can be written in this way. However, the criteria that $f(t)$ must satisfy are not terribly restrictive. These criteria are called the **Dirichlet conditions**, and they are

1. The number of discontinuities for the function $f(t)$ on the interval $0 \le t \le T$ is finite.
2. The number of extrema — *i.e.*, maxima or minima — of $f(t)$ on this same interval is finite.
3. The integral $\int_0^T |f(t)|dt$ is finite.

If a function $f(t)$ satisfies these criteria, the series in Eq. (7.2) it is guaranteed to converge to the value of $f(t)$ at every point where the function is defined. Moreover, at any point $t = t_d$ where $f(t)$ is discontinuous, the series is guaranteed to converge to midpoint

$$f_{\rm mid}(t_d) \;=\; |f_+(t_d) - f_-(t_d)| \;\equiv\; \Big| \lim_{t\to t_d^+} f(t) - \lim_{t\to t_d^-} f(t) \Big| \tag{7.4}$$

between the two limiting values $f_+(x)$ and $f_-(x)$ on either side of the discontinuity. The Dirichlet conditions are satisfied by a wide variety of functions, including square waves (which have discontinuities), triangle waves (which have discontinuous derivatives), and nearly any other periodic function[2] you'd ever be interested in analyzing from a physics perspective.

[1] One might have guessed that this might be the case based on the fact that this step function is an odd function, while the cosine function is even.

[2] In fact, a lot of periodic distributions which technically aren't even functions can be represented this way.

7.3 Orthogonal Functions

Now that we have a qualitative sense of what Fourier decomposition is all about, we're ready to start looking at how this procedure works in more mathematical detail. However, in order to understand how Fourier decomposition works at a deeper level, we first need to discuss a particular property which the functions $\sin(n\omega t)$ and $\cos(n\omega t)$ possess. This property is called **orthogonality**, and it is the foundation on which the entire mechanism of Fourier decomposition rests.

Rather than beginning with an abstract definition of what it means for two functions to be orthogonal right off the bat, we're instead going to build toward our definition of orthogonality from the ground up — *i.e.*, by generalizing from properties of the $\sin(n\omega t)$ and $\cos(n\omega t)$ that we already understand. In particular, our starting point will be to consider what the average of the *product* of two such functions is over the interval $0 \leq t \leq T$, where $\omega = 2\pi/T$. In other words, we're looking to evaluate the integrals

$$\begin{aligned} \mathcal{I}_{mn}^{(cc)} &= \frac{1}{T}\int_0^T \cos(m\omega t)\cos(n\omega t)dt \\ \mathcal{I}_{mn}^{(ss)} &= \frac{1}{T}\int_0^T \cos(m\omega t)\cos(n\omega t)dt \\ \mathcal{I}_{mn}^{(cs)} &= \frac{1}{T}\int_0^T \cos(m\omega t)\sin(n\omega t)dt \ , \end{aligned} \tag{7.5}$$

where m and n are two integers and where $\omega = 2\pi/T$.

We can evaluate these integrals fairly easily by using Euler's Theorem to express the sine and cosine functions in terms of complex exponentials. For example, for $\mathcal{I}_{mn}^{(cc)}$ we have

$$\begin{aligned} \mathcal{I}_{mn}^{(cc)} &= \int_0^T \cos(m\omega t)\cos(n\omega t)dt \\ &= \frac{1}{4T}\int_0^T \Big[e^{im\omega t} + e^{-im\omega t}\Big]\Big[e^{im\omega t} + e^{-im\omega t}\Big]dt \\ &= \frac{1}{4T}\int_0^T \Big[e^{i(m+n)\omega t} + e^{-i(m+n)\omega t} + e^{i(m-n)\omega t} + e^{-i(m-n)\omega t}\Big]dt \\ &= \frac{1}{2T}\int_0^T \Big\{\cos[(m-n)\omega t] + \cos[(m+n)\omega t]\Big\}dt \ . \end{aligned} \tag{7.6}$$

For cases in which $m \neq n$, we

$$\begin{aligned} \mathcal{I}_{mn}^{(cc)} &= \frac{\sin[(m-n)\omega t]}{2(m-n)\omega T}\bigg|_0^T + \frac{\sin[(m+n)\omega t]}{2(m+n)\omega T}\bigg|_0^T \\ &= \frac{\sin[2\pi(m-n)] - \sin(0)}{4\pi(m-n)} + \frac{\sin[2\pi(m+n)] - \sin(0)}{4\pi(m+n)\omega} \\ &= 0 \end{aligned} \tag{7.7}$$

where we have used the fact that $\omega T = 2\pi$ in going from the first to the second line and the fact that $\sin(2\pi p) = 0$ for any integer p. Thus, we see that $\mathcal{I}_{mn}^{(cc)}$ vanishes when the integers m and n are different, regardless of their values. By contrast, when m and n are equal, the corresponding integral does *not* vanish. For example, for the special case in which $m = n = 0$, we have

$$\mathcal{I}_{mm}^{(cc)} \ = \ \frac{1}{2T}\int_0^T \Big[\cos(0) + \cos(0)\Big]dt \ = \ \frac{1}{T}\int_0^T dt \ = \ 1 \ .$$

Likewise, in the special case in which $m = n \neq 0$, we have

$$\mathcal{I}_{mm}^{(cc)} \ = \ \frac{1}{2T}\int_0^T \Big[\cos(0) + \cos(2m\omega t)\Big]dt \ = \ \left[\frac{t}{2T} - \frac{\sin(2m\omega t)}{2m\omega T}\right]\bigg|_0^T \ = \ \frac{1}{2} + \frac{\sin(4\pi m) - \sin(0)}{4\pi m} \ = \ \frac{1}{2} \ .$$

Collecting all of these different cases together, we can write

$$\frac{1}{T}\int_0^T \cos(m\omega t)\cos(n\omega t)dt \;=\; \begin{cases} 0 & m \neq n \\ \frac{1}{2} & m = n \neq 0 \\ 1 & m = n = 0\ . \end{cases} \tag{7.8}$$

Performing the $\mathcal{I}_{mn}^{(ss)}$ integral yields a similar (but not quite identical) result:

$$\frac{1}{T}\int_0^T \sin(m\omega t)\sin(n\omega t)dt \;=\; \begin{cases} 0 & m \neq n \\ \frac{1}{2} & m = n \neq 0 \\ 0 & m = n = 0\ . \end{cases} \tag{7.9}$$

Finally, it can be shown that the $\mathcal{I}_{mn}^{(cs)}$ integral vanishes regardless of the values of m and n:

$$\frac{1}{T}\int_0^T \cos(m\omega t)\sin(n\omega t)dt \;=\; 0\ . \tag{7.10}$$

This last result is not really a surprise, since $\cos(m\omega t)$ is an even function of t and $\sin(m\omega t)$ is an odd function of t. Thus, if we had been integrating product of these two functions from $-T/2$ to $T/2$ (or some other interval symmetric around $t = 0$) rather than from 0 to T, we'd have known that the value of the integral would have had to be zero without even having to calculate anything. However, we also know that $\cos(n\omega t + T) = \cos(n\omega t)$ and $\sin(n\omega t + T) = \sin(n\omega t)$ for any integer n. This means that

$$\int_{-T/2}^{0} \cos(m\omega t)\sin(n\omega t)dt \;=\; \int_{T/2}^{T} \cos(m\omega t)\sin(n\omega t)dt\ . \tag{7.11}$$

Now we are ready to discuss what we mean by orthogonality. The relations in Eq. (7.8), Eq. (7.9), and Eq. (7.10) tell us that unless $m = n$, the integral of the product of any two sine or cosine functions vanishes on the interval $0 \leq t \leq T$. The functions $\cos(n\omega t)$ and $\sin(m\omega t)$ are therefore examples of what are called **orthogonal functions**. Two functions $f(x)$ and $g(x)$ of the independent variable x are said to be orthogonal over the interval $a \leq x \leq b$ if

$$\int_a^b f^*(x)g(x)dt \;= 0\ , \tag{7.12}$$

where $f^*(x)$ denotes the complex conjugate of $f(x)$. We see from Eq. (7.10) that $\cos(m\omega t)$ and $\sin(n\omega t)$ are orthogonal to each other according to this definition for any pair or integers m and n. (Since $\cos(n\omega t)$ and $\sin(n\omega t)$ are purely real functions of t, the complex conjugate of each function is equal to the function itself.) Moreover, Eqs. (7.8) and (7.9) tell us that $\cos(m\omega t)$ is orthogonal to $\cos(n\omega t)$ and that $\sin(m\omega t)$ is orthogonal to $\sin(n\omega t)$ for all $m \neq n$.

The sine and cosine functions are certainly not the only familiar functions which exhibit orthogonality properties of this sort. For example, the set of complex exponential functions of the form $e^{in\omega t}$, where n is an integer and $\omega = 2\pi/T$, are also mutually orthogonal. Indeed, if we take $f^*(x) = e^{-im\omega t}$ and $g(x) = e^{in\omega t}$ in Eq. (7.12) with $m \neq n$, we find that

$$\begin{aligned} \frac{1}{T}\int_0^T e^{-im\omega t}e^{in\omega t}dt &= \frac{1}{T}\int_0^T e^{-i(m-n)\omega t}dt \\ &= \frac{ie^{-i(m-n)\omega t}}{(m-n)\omega T}\bigg|_0^T \\ &= \frac{i}{2\pi}\Big[e^{-2\pi i(m-n)} - e^0\Big] \\ &= \frac{i}{2\pi}\Big\{\cos[2\pi(m-n)] - i\sin[2\pi(m-n)] - 1\Big\} \\ &= 0\ , \end{aligned} \tag{7.13}$$

where in the last step, we have once again used the fact that $\cos(2\pi p) = 1$ and $\sin(2\pi p) = 0$ for any integer p. By contrast, when $m = n$, we find that

$$\frac{1}{T}\int_0^T e^{-im\omega t}e^{im\omega t}dt \;=\; \frac{1}{T}\int_0^T e^0 dt \;=\; 1\;. \tag{7.14}$$

Collecting these together, we can write

$$\frac{1}{T}\int_0^T e^{-im\omega t}e^{in\omega t}dt \;=\; \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases} \tag{7.15}$$

As we shall soon see, it is also possible to use complex exponentials of this sort to construct a Fourier expansion for a complex function $f(t)$.

7.4 Determining the Fourier Coefficients

As for the other coefficients in the Fourier expansion, there is a trick we can use to determine them for any periodic function $f(x)$. We'll begin by determining the coefficients a_n of the cosine terms in Eq. (7.2). Our first step will be to multiply each side of Eq.‘(7.2) by $\cos(m\omega t)/T$, where m is a positive integer:

$$\frac{1}{T}f(t)\cos(m\omega t) \;=\; \frac{1}{T}\sum_{n=0}^{\infty}\Big[a_n\cos(n\omega t)\cos(m\omega t) + b_n\sin(n\omega t)\cos(\omega t)\Big]\;. \tag{7.16}$$

Our next step will be to integrate both sides of this equation over one period of oscillation from $t = 0$ to $t = T$:

$$\frac{1}{T}\int_0^T f(t)\cos(m\omega t)dt \;=\; \frac{1}{T}\sum_{n=0}^{\infty}\left[a_n\int_0^{\infty}\cos(n\omega t)\cos(m\omega t)dt + b_n\int_0^{\infty}\sin(n\omega t)\cos(\omega t)dt\right]\;. \tag{7.17}$$

The right-hand side of this equation involves an infinite sum of integrals. However, nearly all of the integrals in this sum are zero because $\cos(m\omega t)$ and $\sin(n\omega t)$ are orthogonal functions on the interval $0 \le t \le T$. Indeed, Eq. (7.8) tells us that the second term on the right side of this equation vanishes for all values of n in the sum. Moreover, Eq. (7.8) tells us that the first term vanishes for every term in the sum except one: the term where $n = m$. Thus, for $m \neq 0$ we have

$$\frac{1}{T}\int_0^T f(t)\cos(m\omega t)dt \;=\; \frac{a_m}{2}\;. \tag{7.18}$$

Inverting this relation gives us an expression for the value of the coefficient a_m in the Fourier series for the function $f(x)$

$$a_m \;=\; \frac{2}{T}\int_0^T f(t)\cos(m\omega t)dt\;, \qquad \text{for } m \neq 0\;. \tag{7.19}$$

Likewise, for the special case in in which $m = 0$, we get a similar expression from the orthogonality relation in Eq. (7.8), but without the factor of two:

$$\frac{1}{T}\int_0^T f(t)\cos(0)dt \;=\; a_0\;. \tag{7.20}$$

This tells us that the Fourier coefficient a_0 actually has an important physical interpretation. Since $\cos(0) = 1$, the quantity on the left-hand side of this equation is just the average value $\langle f\rangle$ of the function $f(t)$ over one cycle of oscillation. Thus, we write

$$a_0 \;=\; \langle f\rangle \;=\; \frac{1}{T}\int_0^T f(t)dt\;. \tag{7.21}$$

In hindsight, this result actually makes a great deal of sense. The a_0 term in a Fourier series is a special term because the the "cosine" term with which it's associated is really just a constant. Indeed, since $\cos(0) = 1$ and $\sin(0) = 0$, we could just as well have written our Fourier series for $f(x)$ in Eq. 7.2 in the alternative form

$$f(t) \;=\; a_0 + \sum_{n=1}^{\infty} \Big[a_n \cos(n\omega t) + b_n \sin(n\omega t) \Big] \; . \tag{7.22}$$

The average value of a constant over any interval is, of course, just the value of the constant. By contrast, the average value of $\cos(n\omega t)$ or $\sin(m\omega t)$ over one cycle of oscillation is zero for any integer $n \neq 0$, so none of the other terms in the Fourier series contribute to $\langle f \rangle$. In other words, the Fourier coefficient a_0 can be interpreted as an overall shift up or down of the waveform made up by the sum of all of the other, oscillating terms in the Fourier series.[3]

We can play a similar trick to obtain the coefficients b_n of the sine terms in Eq. (7.2). We simply multiply both sides of this equation by $\sin(m\omega t)/T$ and integrate. The result (see Problem 2) for any value of m (including $m = 0$) is

$$b_m \;=\; \frac{2}{T} \int_0^T f(t) \sin(m\omega t) dt \; . \tag{7.23}$$

Example: Square Wave

To get an idea of how this technique for finding the Fourier coefficients of a function works in practice, let's use it to determine the general form of a_n and b_n for the square wave in our example from Sect. 7.2.

The first step is to write down the functional form for $f(x)$ within the interval from $0 \leq t < T$.

$$f(t) \;=\; \begin{cases} f_0 & 0 \leq t \leq T/2 \\ -f_0 & T/2 < t < T \end{cases} \tag{7.24}$$

Moreover, it is clear from Fig. 7.1 that this square wave is an odd function of t. We can therefore anticipate that our Fourier series for $f(t)$ will consist exclusively of odd functions of t. Since $\cos(m\omega t)$ is an even function, this means that that all of the coefficients a_m must vanish. Indeed, this turns out to be the case, but let's prove it. For the special case of the a_0 coefficient, Eq. (7.21) tells us that

$$a_0 \;=\; \frac{f_0}{T} \int_0^{T/2} dt - \frac{f_0}{T} \int_{T/2}^{T} dt \;=\; 0 \; . \tag{7.25}$$

This makes intuitive sense, since a_0 is just the average of $f(t)$ over a cycle, and we have $f(t) = f_0$ for half a cycle and $f(t) = -f_0$ for the other half. Likewise, for the rest of the a_m, Eq. (7.19) gives us

$$\begin{aligned} a_m \;&=\; \frac{2f_0}{T} \int_0^{T/2} \cos(m\omega t) dt - \frac{2f_0}{T} \int_{T/2}^{T} \cos(m\omega t) dt \\ &=\; \frac{2f_0}{T} \left[\frac{\sin(\pi m)}{m\omega} - \frac{\sin(0)}{m\omega} \right] - \frac{2f_0}{T} \left[\frac{\sin(2\pi m)}{m\omega} - \frac{\sin(\pi m)}{m\omega} \right] \\ &=\; 0 \; , \end{aligned} \tag{7.26}$$

since $\sin(\pi m) = 0$ for any integer m. So indeed, as we anticipated, the Fourier series for our square wave consists entirely of sine rather than cosine terms.

[3] Another useful analogy is to think of the a_0 coefficient as performing the same role that the "DC offset" function plays on our waveform generators in lab.

Now, let's determine the coefficients b_m of those sine terms. Plugging our expression for $f(t)$ in Eq. (7.24) into Eq. (7.23) gives us a general formula for b_m:

$$\begin{aligned} b_m &= \frac{2f_0}{T}\int_0^{T/2}\sin(m\omega t)dt - \frac{2f_0}{T}\int_{T/2}^{T}\sin(m\omega t)dt \\ &= -\frac{2f_0}{T}\left[\frac{\cos(\pi m)}{m\omega} - \frac{\cos(0)}{m\omega}\right] + \frac{2f_0}{T}\left[\frac{\cos(2\pi m)}{m\omega} - \frac{\cos(\pi m)}{m\omega}\right] \\ &= \frac{2f_0}{T\omega m}\big[2 - 2\cos(\pi m)\big] \\ &= \frac{2f_0}{\pi m}\big[1 - (-1)^m\big] \ . \end{aligned} \tag{7.27}$$

We see that $b_m = 0$ for even m, but that the b_m for odd m are non-vanishing. Moreover, we also see that the value of b_m for the non-zero coefficients decreases with increasing m. This implies that the higher-frequency terms in the Fourier series have a smaller impact on the overall waveform than the lower frequency terms do. Indeed, we have already seen in Fig. (7.2) that the amplitudes of the individual sine terms in the Fourier series — which is, after all, what the coefficients b_m represent — are smaller for the higher-frequency modes. In Fig. 7.4, we present a plot of b_m vs. m for our square wave.

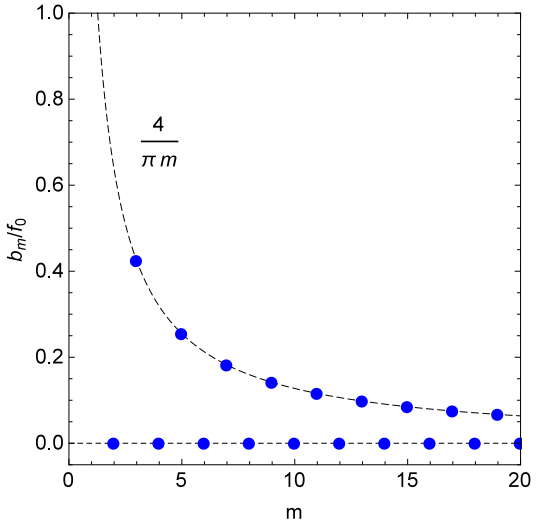


Figure 7.4: The values of the first few Fourier coefficients b_m (blue dots) for a square wave. The coefficients for even m vanish, while the coefficients for odd m lie along the curve $b(m) = 4/(\pi m)$.

Finally, putting the results from Eq. (7.25), Eq. (7.26), and Eq. (7.27) together, we arrive at our final, closed-form expression for the Fourier-series expansion of our square wave:

$$f(t) = \frac{2f_0}{\pi}\sum_{n=0}^{\infty}\frac{1-(-1)^n}{n}\sin(n\omega t) \ . \tag{7.28}$$

Example: $f(t) \propto \cos^2(\omega t)$

As another example, let's consider the Fourier expansion of the function

$$f(t) = f_0\cos^2(\omega t) \ . \tag{7.29}$$

This function is even, so we know that the coefficients b_m of the sine terms in the Fourier series must vanish. Thus, we'll focus on the coefficients a_0 of the cosine terms.

We being with the special case of a_0. Plugging Eq. (7.29) into Eq. (7.21) gives us

$$a_0 \;=\; \frac{f_0}{T}\int_0^T \cos^2(\omega t)dt \;. \tag{7.30}$$

We can evaluate the integral over $\cos^2(\omega t)$ using integration by parts. Doing so, we find that

$$a_0 \;=\; \frac{f_0}{2T\omega}\Big[1+\cos(\omega t)\sin(\omega t)\Big]\Big|_0^T \;=\; \frac{f_0}{2T\omega}\;. \tag{7.31}$$

The rest of the a_m can be evaluated using Eq. (7.19):

$$\begin{aligned} a_m &= \frac{2f_0}{T}\int_0^T \cos^2(\omega t)\cos(m\omega t)dt \\ &= \frac{2f_0}{T}\int_0^T \frac{1}{2}\big[\cos(0)+\cos(2\omega t)\big]\cos(m\omega t)dt \\ &= f_0\left[\frac{1}{T}\int_0^T \cos(m\omega t)dt + \frac{1}{T}\int_0^T \cos(2\omega t)\cos(m\omega t)dt\right]\;, \end{aligned} \tag{7.32}$$

where in going from the first to the second line, we have used the identity

$$\cos\alpha\cos\beta \;=\; \frac{1}{2}\big[\cos(\alpha-\beta)+\cos(\alpha+\beta)\big]\;. \tag{7.33}$$

The first term in the square brackets on the right side of Eq. (7.32) is easy to evaluate. It's just the average of $\cos(m\omega t)$ over the interval $0 \le t \le T$, which is zero for any integer $m \neq 0$. This leaves us with the second term, which is a little less trivial. However, we can take advantage of the orthogonality relation in Eq. (7.8), which tells us that this term also vanishes except in the special case where $m = 2$. In this case, we have

$$\frac{1}{T}\int_0^T \cos(2\omega t)\cos(2\omega t)dt \;=\; \frac{1}{2}\;, \tag{7.34}$$

so the coefficient a_2 is

$$a_2 \;=\; \frac{f_0}{2}\;. \tag{7.35}$$

For all other values of m, both terms on the right-hand side of Eq. (7.32) vanish, and thus a_m vanishes as well.

The punch line is that the Fourier series for the cosine-squared function in Eq. (7.29) turns out to be remarkably simple. Indeed, only two of the coefficients in the entire infinite sum — the coefficients a_0 and a_2 — are non-zero. Thus, the Fourier series for $f(x)$ is just

$$f(x) \;=\; a_0 + a_2\cos(2\omega t) \;=\; \frac{f_0}{2} + \frac{f_0}{2}\cos(2\omega t)\;. \tag{7.36}$$

The individual contributions from these two non-vanishing terms, as well as their sum — *i.e.*, the function $f(x)$ itself — are shown in Fig. fig:CosSqdPlotWithComponents.

As it turns out, "punch line" is actually an apt turn of phrase. In a sense, the joke is on us, because we could actually have obtained this result in one line by applying the identity in Eq. (7.33) directly to our original function:

$$f(x) \;=\; f_0\cos^2(\omega t) \;=\; \frac{f_0}{2}\big[cos(0)+\cos(2\omega t)\big] \;=\; \frac{f_0}{2}+\frac{f_0}{2}\cos(2\omega t) \tag{7.37}$$

However, going through the exercise of determining the Fourier coefficients in this case was valuable for two reasons. First, it is one of the few cases in which we have an independent way of checking the values of those coefficients that we obtained using the method outlined in Sect. 7.4. Second, this case also provided an example of how the orthogonality relations discussed in Sect. 7.3 can be used in order to evaluate integrals.

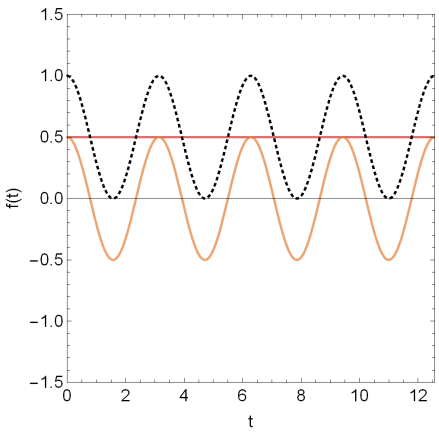


Figure 7.5: The individual contributions (solid curves) from the two non-zero terms in the Fourier series for the cosine-squared function in Eq. (7.29). The sum of these two terms (dashed curve), which corresponds to the function $f(x)$ itself, is also shown.

7.5 Fourier Series in Terms of Complex Exponentials

In Sect. 7.3, we saw that the set of complex exponentials $e^{in\omega t}$, where n is an integer and $\omega = 2\pi/T$, are orthogonal functions on the interval $0 \le t \le T$. In this section, we will see how to construct Fourier series using these complex exponentials instead of sines and cosines. The analogue of Eq. (7.2) is

$$f(t) \;=\; \sum_{n=-\infty}^{\infty} c_n e^{in\omega t} \;, \tag{7.38}$$

where the c_n are constant *complex* coefficients. Note that in contrast to the sum over sines and cosines in Eq. (7.2), the sum here extends from $-\infty$ to ∞ rather than from 0 to ∞.

In order to determine the Fourier coefficients c_n for a given function $f(x)$, we can employ essentially the same trick we used in Sect. 7.4. We begin by multiplying both sides by $e^{-im\omega t}/T$ and then integrating:

$$\frac{1}{T}\int_0^T f(t)e^{-im\omega t}dt \;=\; \frac{1}{T}\sum_{n=-\infty}^{\infty} c_n \int_0^T e^{-im\omega t}e^{in\omega t}dt \;. \tag{7.39}$$

Our orthogonality relation in Eq. (7.15) tells us that all of the terms in the sum on the right-hand side vanish except for the term where $m = n$. For this lone surviving term, the value of the integral is 1, so we have

$$c_m \;=\; \frac{1}{T}\int_0^T f(t)e^{-im\omega t}dt \;. \tag{7.40}$$

It's also worth mentioning that since the integrand in Eq. (7.40) is periodic with period T, it must be true that

$$\int_{T/2}^{T} f(t)e^{-im\omega t}dt \;=\; \int_{-T/2}^{0} f(t)e^{-im\omega t}dt \;. \tag{7.41}$$

Therefore, we could just as well have written Eq. (7.40) in terms of an integral from $-T/2$ to $T/2$, rather than as an integral from 0 to T:

$$c_m \;=\; \frac{1}{T}\int_{-T/2}^{T/2} f(t)e^{-im\omega t}dt \;. \tag{7.42}$$

Writing the limits of integration in this more symmetric way is useful for many purposes. For example, as we shall see in Sect. 7.7, it makes the behavior of the c_n in the $T \to \infty$ limit more transparent.

The advantage of working with Fourier series in this form is that it's a lot easier and faster once you get the hang of it. When we expand a function $f(x)$ in terms of sine and cosine functions as we did in 7.4, we have two different sets of coefficients a_n and b_n to keep track of. Moreover, the coefficient a_0 is a special case which must be computed separately. By contrast, when we expand a function in terms of complex exponentials, we only have one set of coefficients c_n to keep track of, and the formula for c_0 takes the same form as the formula for every other coefficient. As you gain familiarity with complex numbers, you will find yourself increasingly preferring to work with Fourier series in this latter form.

7.6 Solving Differential Equations with Fourier Series

Now that we've seen how Fourier decomposition works and how to find the Fourier coefficients for an arbitrary function $f(x)$, we're ready to put these techniques to work in solving differential equations.

As discussed in Sect. 7.1, the principle of superposition tells us that if the inhomogeneous term $f(x)$ in a linear differential equation can be written as a sum of the form

$$f(t) \;=\; f_1(x) + f_2(x) + f_3(x) + \ldots \;, \tag{7.43}$$

then there always exists a particular solution $y_p(x)$ of the form

$$y_p(x) \;=\; y_{p1}(x) + y_{p2}(x) + y_{p3}(x) + \ldots \;, \tag{7.44}$$

where $y_{pn}(x)$ is the solution to the corresponding differential equation with the full function $f(x)$ replaced by $f_n(x)$. We have now shown that essentially *any* arbitrary periodic function can be expressed as an Fourier series — *i.e.*, an infinite sum of sine and cosine terms. The principle of superposition therefore implies that an inhomogeneous linear differential equation where $f(x)$ is a periodic function has a particular solution of the form

$$y_p(x) \;=\; \sum_{n=0}^{\infty} y_{pn}(x) \;. \tag{7.45}$$

Moreover, since the individual terms in the Fourier series have the form

$$f_n(x) \;=\; C_n \cos(\lambda x) + D_n \sin(\lambda x) \;, \tag{7.46}$$

where C_n and D_n are constants, the method of undetermined coefficients tells us that each of the $y_{pn}(x)$ has the form

$$y_{pn}(x) \;=\; \sum_{n=0}^{\infty} \Big[A_n \cos(n\omega t) + B_n \sin(n\omega t) \Big] \;. \tag{7.47}$$

As an example of how this works, let's consider the case of a damped, driven harmonic oscillator with a driving force $F(x)$ which takes the form of our square wave from Fig 7.1 with amplitude $f_0 = F_0$. We have already computed the Fourier coefficients for this function in Sect. 7.4. Plugging our result from Eq. (7.28) into the damped, driven harmonic-oscillator equation gives us

$$\frac{d^2x}{dt^2} + 2\beta\frac{dx}{dt} + \omega_0^2 x \;=\; \frac{2F_0\omega_0^2}{\pi k} \sum_{n=0}^{\infty} \frac{[1-(-1)^n]}{n} \sin(n\omega t) \;. \tag{7.48}$$

The inhomogeneous term on right-hand side of this equation is an infinite sum, each of the terms in which has the form

$$f_n(x) \;=\; \frac{2F_0\omega_0^2[1-(-1)^n]}{\pi n k} \sin(n\omega t) \;. \tag{7.49}$$

Thus, the principle of superposition tells us that there exists a particular solution to this equation of the form

$$x(t) \;=\; \sum_{n=0}^{\infty} x_{pn}(t) \;, \tag{7.50}$$

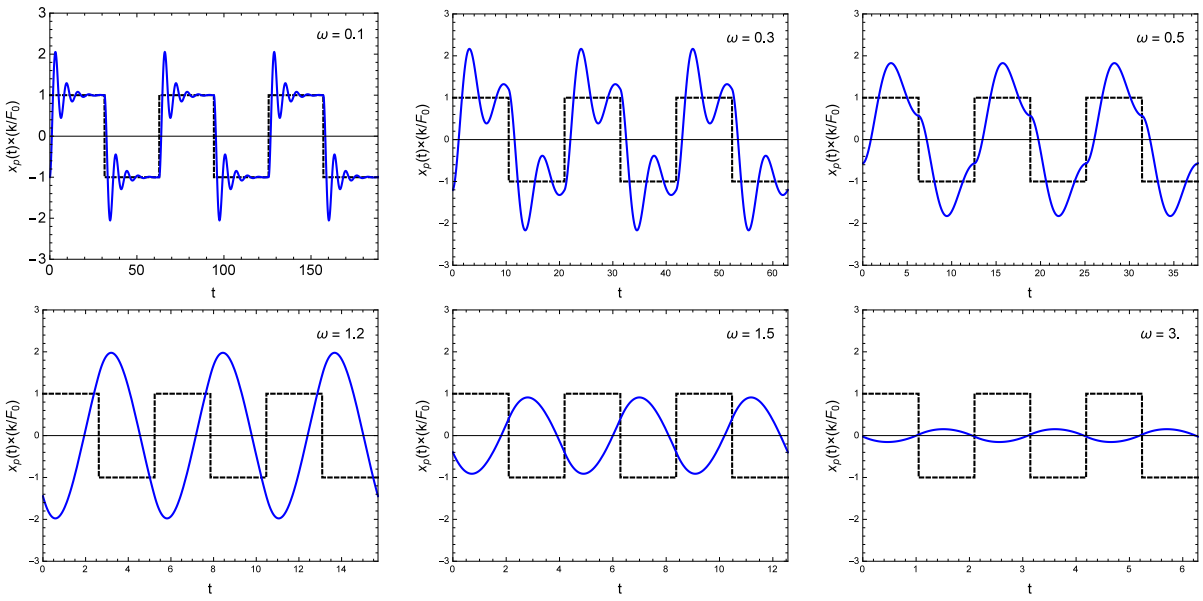


Figure 7.6: Solutions $x_p(t)$ (normalized to the value of F_0/k) to the damped, driven harmonic oscillator equation with a square-wave driving force with amplitude F_0. Each panel shows the solution for a different value of ω In each case, we have taken $\omega_0 = 1$ and $\beta = 0.2$ (in arbitrary units) and plotted $x_p(t)$ on the range $0 \leq t \leq 3T$, where T is the period of oscillation. (Note that since the value of $\omega = 2\pi/T$ is different in each panel, the range of t which corresponds to $0 \leq t \leq 3T$ is also different.) For comparison purposes, we also include a curve in each panel corresponding to driving force itself (normalized to the value of F_0).

where $x_{pn}(t)$ is the solution to the differential equation

$$\frac{d^2x_{pn}}{dt^2} + 2\beta\frac{dx_{pn}}{dt} + \omega_0^2 x_{pn} \;=\; \frac{2F_0\omega_0^2[1-(-1)^n]}{\pi nk}\sin(n\omega t)\;. \tag{7.51}$$

The inhomogeneous term in Eq. (7.51) is just a sine wave, so the method of undetermined coefficients tells us that our "guess" for each of the $x_{pn}(x)$ should be of the form

$$x_{pn}(t) \;=\; A_n\cos(n\omega t) + B_n\sin(n\omega t)\;, \tag{7.52}$$

where A_n and B_n are our undetermined coefficients. Determining the values of these coefficients is simply a matter of plugging this guess back into Eq. (7.51) and solving the resulting system of equations. However, we can take a shortcut. In Sect. 6.3, we already derived a particular solution in this way for a differential equation in which the inhomogeneous term took the form of a cosine rather than a sine function. We know that $\sin(n\omega t) = \cos(n\omega t - \pi/2)$, so we can rewrite Eq. (7.51) in the form

$$\frac{d^2x_{pn}}{dt^2} + 2\beta\frac{dx_{pn}}{dt} + \omega_0^2 x_{pn} \;=\; \frac{2F_0\omega_0^2[1-(-1)^n]}{\pi nk}\cos(n\omega t - \pi/2)\;. \tag{7.53}$$

This isn't exactly the same equation we solved in Chapter 6 because the cosine term contains an additional phase shift of $-\pi/2$. However, it's not too hard to convince ourselves that the "guess" we should use when applying the method of undetermined coefficients in this case is

$$x_{pn}(t) \;=\; A_n\cos(n\omega t - \pi/2) + B_n\sin(n\omega t - \pi/2)\;. \tag{7.54}$$

The phase shift doesn't affect the way that the derivatives act on $x_{pn}(t)$, so this phase just "comes along for the ride," so to speak. As a result, we find that our solution for $x_{pn}(t)$ is identical to the result we derived in Sect. 6.3, except that there's an additional phase shift of $-\pi/2$ which turns the cosine function into a sine:

$$\begin{aligned} x_{pn}(t) &= A_n(\omega)\cos\left[n\omega t - \delta_n(\omega) - \pi/2\right] \\ &= A_n(\omega)\sin\left[n\omega t - \delta_n(\omega)\right]\;, \end{aligned} \tag{7.55}$$

where the amplitude $A_n(\omega)$ and phase $\delta_n(\omega)$ are given by the familiar expressions from Chapter 6 with the replacements

$$\omega \to n\omega \ , \qquad F_0 \to \frac{2F_0[1-(-1)^n]}{\pi n k} \ . \tag{7.56}$$

In other words, we have

$$\begin{aligned} A_n(\omega) &= \frac{\omega_0^2}{\sqrt{(\omega_0^2 - n^2\omega^2)^2 + 4\beta^2 n^2 \omega^2}} \, \frac{2F_0[1-(-1)^n]}{\pi n k} \\ \delta_n(\omega) &= \arctan\left(\frac{2\beta n\omega}{\omega_0^2 - n^2\omega^2}\right) \ . \end{aligned} \tag{7.57}$$

The solution $x_p(t)$ to the full differential equation in Eq. (7.48) is just the sum of these individual contributions:

$$x_p(t) = \sum_{n=0}^{\infty} A_n(\omega) \sin\left[n\omega t - \delta_n(\omega)\right] \ . \tag{7.58}$$

In Fig. 7.6, we show what the solution in Eq. (7.58) looks like for a variety of different values of the driving frequency ω ranging from $\omega \ll \omega_0$ (upper left) to $\omega \gg \omega_0$ (lower right). We see that these curves exhibit a wide variety of possible behaviors, depending on the relationship between ω and ω_0.

7.7 Fourier Transforms

We have already seen that Fourier decomposition provides us with an incredibly powerful tool for solving linear differential equations. Indeed, this technique — in conjunction with the principle of superposition — gives us the tools to derive a particular solution for essentially any such equation in which the inhomogeneous function $f(x)$ is periodic in t. However, it turns out there is also a procedure for decomposing a function $f(x)$ which is *not* periodic. The resulting expansion is called the **Fourier transform** of $f(x)$.

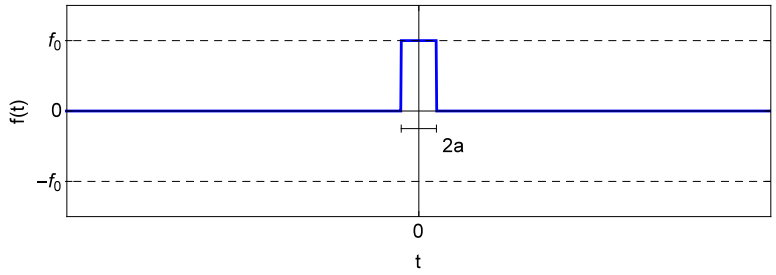


Figure 7.7: A single, non-repeating square pulse with height f_0 and width $2a$.

In order to understand how Fourier transforms work, it's again best to start with a concrete example. In particular, let us consider the function $f(x)$ illustrated in Fig. 7.7. This function consists of a single, non-repeating square pulse of height f_0 and width $2a$. In other words, it's a piecewise function of the form

$$f(x) = \begin{cases} 0 & t < -a \\ f_0 & -a \le t \le a \\ 0 & t > a \ . \end{cases} \tag{7.59}$$

This function is clearly not periodic in t, so it cannot be represented as a Fourier series.

Now let us consider the function shown in the top panel of Fig. 7.10, which consists of a train of such pulses with period T. This function *is* periodic, so we can represent it as a Fourier series. Indeed, we already

know from Problem 4 that the Fourier series (in complex-exponential form) for this function is

$$f(t) \;=\; \sum_{n=0}^{\infty} c_n e^{i\omega_n t} \;, \qquad c_n \;=\; \frac{2a}{T} f_0 \frac{\sin(\omega_n a)}{\omega_n a} \;, \tag{7.60}$$

where ω_n is just the frequency of the corresponding term:

$$\omega_n \;\equiv\; n\omega \;=\; \frac{2\pi n}{T} \;. \tag{7.61}$$

This function is obviously not the same as the square-wave pulse from Eq. (7.59). Nevertheless, it is true that these two functions *do* coincide at times $|t| < T - a$. This means that if we want a pulse train that coincides with this square-wave pulse over a larger range of t, all we need to do increase the period T while keeping the width a of the pulses fixed, as illustrated in the bottom panel of Fig. 7.8.

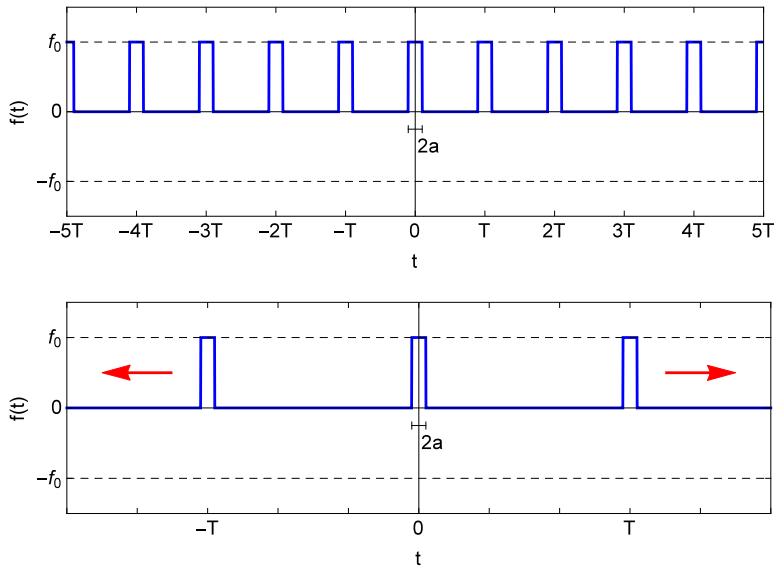


Figure 7.8: A pulse train with period T (top panel) consisting of pulses with height f_0 and width $2a$. By taking the limit as $T \to \infty$ while holding a fixed (as illustrated schematically in the bottom panel), we recover the single pulse shown in Fig. 7.7.

Let's now consider what happens to the Fourier series in Eq. (7.60) as we increase T while holding a fixed. First of all, Eq. (7.61) tells us that each ω_n decreases with increasing T. However, we should keep in mind that there are an infinite number of modes in the Fourier series, so what this really means is that the interval

$$\Delta\omega \;\equiv\; \omega_{n+1} - \omega_n \;=\; \frac{2\pi}{T} \tag{7.62}$$

between the frequencies of any two successive modes in the Fourier series decreases. Another way of saying this is that within a fixed frequency range — say, for example, from $1\ \mathrm{s}^{-1}$ to $2\ \mathrm{s}^{-1}$ — the number of frequencies ω_n within that frequency range which appear in our Fourier series for $f(x)$ increases.

At the same time, however, increasing T also affect the Fourier coefficients c_n. We can get a better sense of exactly how these coefficients scale with T by using Eq. (7.61) in order to rewrite ω_n in terms of T and thereby make the T-dependence of the c_n more explicit:

$$c_n \;=\; \frac{2af_0}{T} \frac{\sin(2\pi na/T)}{(2\pi na/T)} \;. \tag{7.63}$$

Now let's consider what happens to this expression when T becomes large. For fixed values of a and n, the quantity $2\pi na/T$ becomes small. We have already proved in Chapter 3 of these notes that

$$\lim_{x\to 0} \frac{\sin(x)}{x} \;=\; 1 \;, \tag{7.64}$$

so for very large T, the expression in Eq. (7.63) becomes approximately $c_n \approx 2af_0/T$. This means that if we multiply the coefficient c_n by $T/2\pi$,[4] the resulting quantity

$$\frac{Tc_n}{2\pi} \;=\; \frac{af_0}{\pi}\frac{\sin(\omega_n a)}{\omega_n a} \tag{7.65}$$

does not tend toward zero when T is large. In fact, if we view Eq. (7.65) as a function of the variable ω_n, we see that the *shape* of this function is independent of T. What changes as T is increased, as discussed above, is the number of values for ω_n within a given frequency range at which this function is "sampled" — *i.e.*, the number of values for ω_n which actually correspond to modes in the Fourier series. This behavior is illustrated in the different panels of Fig. 7.9. In each panel, the dashed curve corresponds to the quantitiy in Eq. 7.65, plotted as a function of ω_n. The blue dots in each panel correspond to the the actual values of ω_n for the Fourier modes. As T increases, the dashed curve remains the same, but the density of points increases.

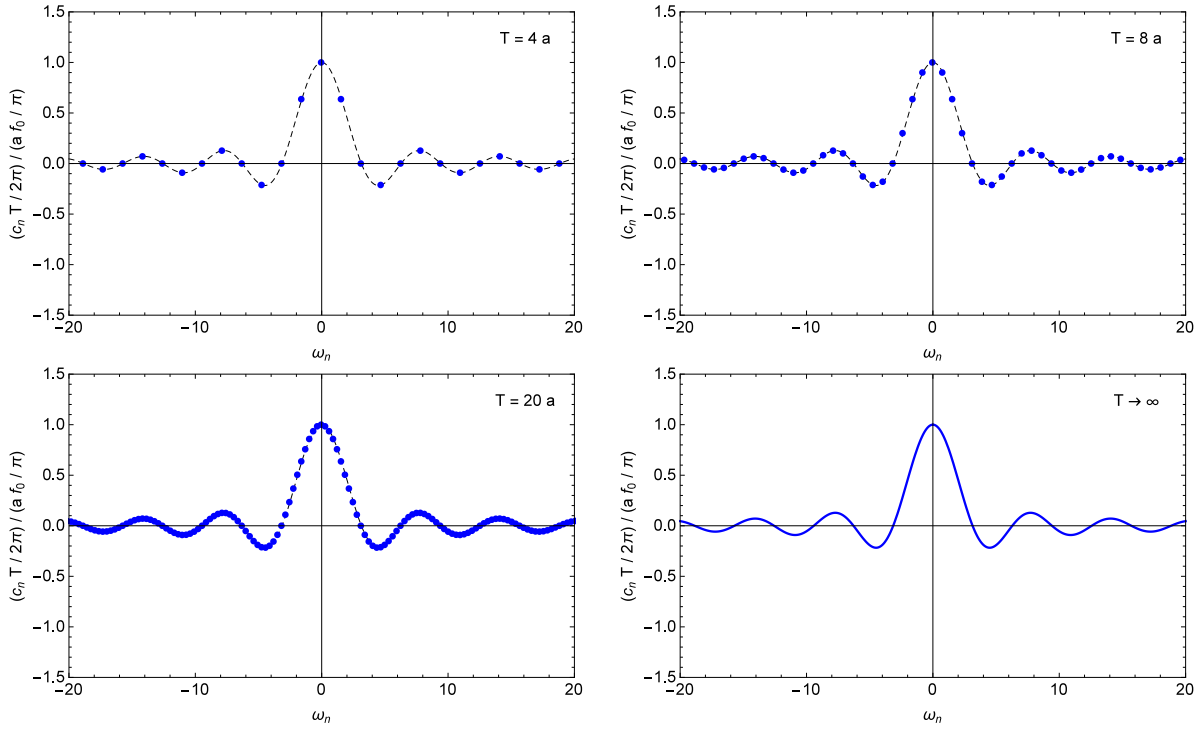


Figure 7.9: The values (blue dots) of the rescaled Fourier coefficients $c_nT/2\pi$ for the pulse train shown in Fig. 7.8, normalized to the value of af_0/π. The four different panels shown here correspond to different values for the of T. For $T = 4a$, as shown in the top left panel, the interval $\Delta\omega$ between the frequencies of any two successive modes is quite large. As T increases, this interval decreases, which means that more and more values of ω_n are being "sampled" in the Fourier sum. In the $T \to \infty$ limit (illustrated in the bottom right panel), the $\Delta\omega \to 0$ and the spectrum of ω_n becomes a continuum.

[4]The factor of 2π that we have included here is merely a particular convention that we've adopted in order to make our subsequent calculations cleaner. The important thing is that we're multiplying by T in order to cancel out the factor of $1/T$ in c_n.

Finally, let us consider what happens in the limit in which T is not merely large, but infinite. In this limit, the distance between the pulses in our pulse train becomes infinite, and we recover the single square pulse from Fig. 7.7. Let's consider what happens to the the frequencies ω_n and amplitudes c_n of the individual Fourier modes in this limit. First of all, the interval $\Delta\omega$ in Eq. (7.62) goes to zero as $T \to \infty$. This means that the spectrum of frequencies included in the Fourier series becomes continuous. The Fourier coefficients c_n also go to zero in this limit, but they do so in such a way that the curve of $Tc_n/2\pi$ vs. ω_n in Eq. (7.65) remains unchanged. The result is that $Tc_n/2\pi$ really becomes a continuous function of the frequency in this limit, as shown in the bottom right panel of Fig. 7.9. This continuous function, which we will call $F(\omega)$ is given by

$$F(\omega) \equiv \lim_{T\to\infty} \frac{Tc_n}{2\pi} = \frac{af_0}{\pi}\frac{\sin(\omega a)}{\omega a} . \tag{7.66}$$

where we have replaced the symbol ω_n with the symbol ω because the spectrum of frequencies becomes continuous in the $T \to \infty$ limit. The quantity in Eq. (7.66) is called the **Fourier transform** of our square wave pulse Eq. (7.59). Physically, the Fourier transform represents the spectrum of amplitudes for the continuous spectrum of Fourier modes that make up our square-wave pulse.

The notion of a Fourier transform $F(\omega)$ for a non-periodic function $f(t)$ is a general concept that extends far beyond our square-wave pulse example. Indeed, we can likewise define the Fourier transform for any well-behaved[5] function $f(t)$ in a similar manner. Since Eq. 7.42 gives us a general expression for the Fourier coefficients c_n for any such function, we have

$$\begin{aligned} F(\omega) &= \lim_{T\to\infty} \frac{T}{2\pi} c_n \\ &= \lim_{T\to\infty} \frac{T}{2\pi}\frac{1}{T}\int_{-T/2}^{T/2} f(t)e^{-i\omega_n t}dt \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt , \end{aligned} \tag{7.67}$$

where once again we have replaced the symbol ω_n with the symbol ω because the spectrum of frequencies becomes continuous in the $T \to \infty$ limit.

Finally, it is also possible to invert the relation in Eq. (7.67) and obtain a formula for $f(x)$ in terms of $F(\omega)$. As with the Fourier transform itself, the simplest way to derive this formula is to begin with the corresponding relation for a periodic function and take the $T \to \infty$ limit. In this case, the relation is the expression for the Fourier series itself in Eq. (7.38), which we can write as

$$\begin{aligned} f(t) &= \sum_{n=-\infty}^{\infty} c_n e^{i\omega_n t} \\ &= \sum_{n=-\infty}^{\infty} \frac{Tc_n}{2\pi} e^{i\omega_n t}\frac{2\pi}{T} \\ &= \sum_{n=-\infty}^{\infty} \frac{Tc_n}{2\pi} e^{i\omega_n t}\Delta\omega \end{aligned} \tag{7.68}$$

Now it's a little more clear what happens to this expression when we take the $T \to \infty$ limit. The discrete set of frequencies ω_n becomes a continuous variable ω, the quantity $Tc_n/2\pi$ becomes the Fourier transform $F(\omega)$, and the sum weighted by the spacing $\Delta\Omega$ between frequencies becomes an integral over $d\omega$:

$$\begin{aligned} f(t) &= \lim_{T\to\infty} \sum_{n=-\infty}^{\infty} \frac{Tc_n}{2\pi} e^{i\omega_n t}\Delta\omega \\ &= \int_{-\infty}^{\infty} F(\omega)e^{i\omega t}d\omega . \end{aligned} \tag{7.69}$$

[5]In this context, "well-behaved" means that the function satisfies a set of criteria which look like the Dirichlet conditions in Sect. 7.2, but on the interval $-\infty$ to ∞ rather than 0 to T.

The "Fourier integral" in Eq. (7.69) is called the *inverse Fourier transform*. It is the analogue of the Fourier series Eq. (7.38) for a non-periodic function. In other words, Eq. (7.69) tells us that it's actually still possible for us to write a "Fourier series" for such a function — provided that we are prepared to include a continuum of frequencies. Moreover, this equation makes the physical meaning of the Fourier transform $F(\omega)$ very clear. Indeed, it represents the spectrum of amplitudes for the continuous spectrum of Fourier modes that make up our square-wave pulse. Likewise, the Fourier transform in Eq. (7.67) can be seen as the analog[6] of the formula for the Fourier coefficients c_n in Eq. (7.42).

Example: Inverse Fourier Transform of a Square Pulse

As an example of how Eq. (7.69) can be used to reconstruct a function $f(t)$ from its Fourier transform, let's use it to obtain $f(t)$ from the Fourier transform of our square-wave pulse from Fig. 7.7. Plugging the expression for $F(\omega)$ in Eq. (7.66) into Eq. (7.69) gives us

$$f(t) \;=\; \int_{-\infty}^{\infty} \frac{af_0}{\pi}\frac{\sin(\omega a)}{\omega a}e^{i\omega t}d\omega \;. \tag{7.70}$$

We can simplify things a bit using Euler's formula:

$$\begin{aligned} f(t) &= \frac{af_0}{\pi}\int_{-\infty}^{\infty}\frac{\sin(\omega a)}{\omega a}e^{i\omega t}d\omega \\ &= \frac{f_0}{\pi}\int_{-\infty}^{\infty}\frac{1}{2i\omega}\Big[e^{i\omega(t+a)}-e^{i\omega(t-a)}\Big]d\omega \\ &= \frac{f_0}{\pi}\int_{-\infty}^{\infty}\left\{\frac{\cos[\omega(t+a)]+i\sin[\omega(t+a)]}{2i\omega}-\frac{\cos[\omega(t-a)]+i\sin[\omega(t-a)]}{2i\omega}\right\}d\omega \;. \end{aligned} \tag{7.71}$$

This integral may not look much simpler, but its now in a form that's more straightforward to evaluate. First of all, we know that $\cos[\omega(t\pm a)]$ is an even function of ω, so the ratio $\cos[\omega(t\pm a)]/\omega$ is an odd function. This implies that the integral of each of the cosine terms in Eq. (7.71) from $-\infty$ to ∞ is zero. This just leaves the sine terms:

$$f(t) \;=\; \frac{f_0}{2\pi}\int_{-\infty}^{\infty}\left\{\frac{\sin[\omega(t+a)]}{\omega}-\frac{\sin[\omega(t-a)]}{\omega}\right\}d\omega \;. \tag{7.72}$$

We can evaluate the integral of each of the terms in the integrand separately by performing an appropriate change of variables. Using $u=\omega(t+a)$ for the first term term and $u=\omega(t-a)$ for the second term, we find that

$$\begin{aligned} f(t) &= \frac{f_0}{2\pi}\left\{\int_{-\infty}^{\infty}\frac{\sin[\omega(t+a)]}{\omega}d\omega-\int_{-\infty}^{\infty}\frac{\sin[\omega(t-a)]}{\omega}d\omega\right\} \\ &= \frac{f_0}{2\pi}\left\{\mathrm{sign}(t+a)\int_{-\infty}^{\infty}\frac{\sin(u)}{u}du-\mathrm{sign}(t-a)\int_{-\infty}^{\infty}\frac{\sin[u)]}{u}du\right\} \;. \end{aligned} \tag{7.73}$$

The reason that the $\mathrm{sign}(t\pm a)$ coefficient arises in each term is that the u-substitution flips the limits of integration when $t\pm a$ is negative, but not when it's positive. When it's negative, you pick up a minus sign when you put the limits of integration back in the correct order. Each of the integrals in Eq. (7.73) evaluates to π, so we find that

$$f(t) \;=\; \frac{f_0}{2}\big[\mathrm{sign}(t+a)-\mathrm{sign}(t-a)\big] \;. \tag{7.74}$$

The value of the term in square brackets in Eq. (7.74) depends on the relationship between t and a. For $t<-a$, we have $\mathrm{sign}(t+a)=\mathrm{sign}(t-a)=-1$, so $f(t)=0$. Likewise, when $t>a$, we have $\mathrm{sign}(t+a)=\mathrm{sign}(t-a)=+1$, so again $f(t)=0$. However, for $-a<t<a$, we have $\mathrm{sign}(t+a)=+1$ and $\mathrm{sign}(t-a)=-1$, which means that $f(t)=f_0$. Collecting these results together, we obtain our final piecewise expression for $f(t)$:

$$f(x) \;=\; \begin{cases} 0 & t<-a \\ f_0 & -a<t<a \\ 0 & t>a \;. \end{cases} \tag{7.75}$$

[6] No pun intended.

This is nothing but our original square-wave-pulse function from Eq. (7.59)![7] Indeed, when we performing the inverse Fourier transform of $F(\omega)$, we recover $f(t)$, as expected.

Problems

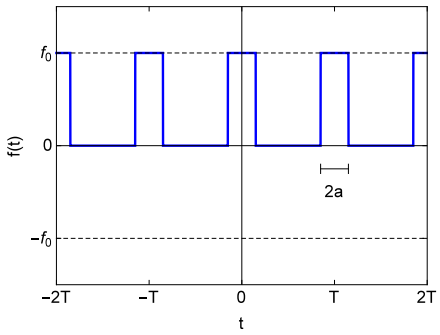


Figure 7.10: A pulse train with an overall period T, which consists of a series of rectangular pulses of width $2a$.

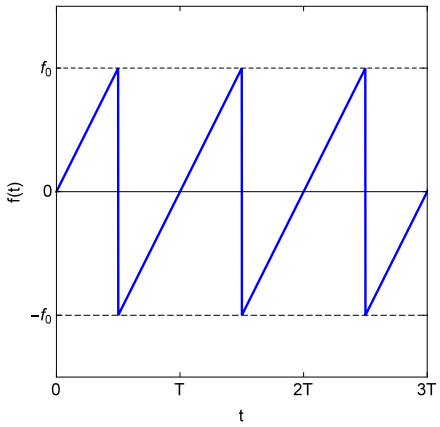


Figure 7.11: A sawtooth wave with amplitude f_0 and period T.

1. Explicitly derive the orthogonality relations for in Eqs. (7.9) and (7.10).
2. Verify that the the expression in Eq. (7.23) gives the correct result for the Fourier coefficients b_n.
3. Show the following about the complex Fourier coefficients c_n appearing in Eq. (7.38).
   (a) For an even function $f(t)$, show that the coefficients satisfy $c_{-n} = c_n$.
   (b) For an odd function $f(t)$, show that the coefficients satisfy $c_{-n} = -c_n$.
   (c) For a real function $f(t)$, show that the coefficients satisfy $c_{-n} = c_n^*$.

[7] There is one minor technical discrepancy here, but it's an expected one. It turns out that at a value of t where $f(t)$ is discontinuous, the Fourier transform converges to the midpoint between the values of $f(t)$ on either side of the discontinuity just like a regular Fourier series does [see Eq. (7.4)]. You can verify for yourself that $f(t) = f_0/2$ for $t = \pm a$ for our square-wave pulse.

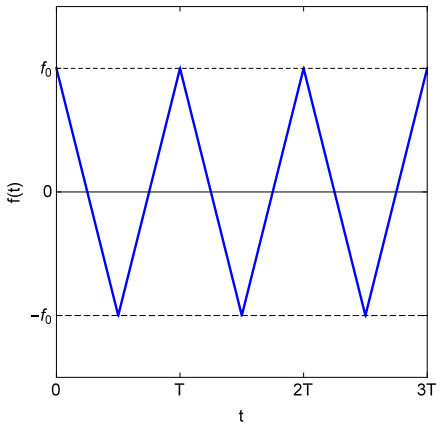


Figure 7.12: A triangle wave with amplitude f_0 and period T.

(d) For an imaginary function $f(t)$, show that the coefficients satisfy $c_{-n} = -c_n^*$.

4. Find the Fourier series, in complex-exponential form, for the pulse train pictured in Fig. 7.10. Check that any applicable relations among the c_n that you derived in Problem 3 are satisfied in this case.

5. Consider a sawtooth wave with amplitude f_0 and period T, as shown in Fig. 7.11.

   (a) Determine the Fourier series for this function in closed form.

   (b) Use Mathematica to make a single plot that shows both the original sawtooth wave and the partial sum including the first three non-vanishing terms in the Fourier Series.

6. Consider a triangle wave with amplitude f_0 and period T, as shown in Fig. 7.12,

   (a) Determine the Fourier series for this function in closed form.

   (b) Use Mathematica to make a single plot that shows both the original triangle wave and the partial sum including the first three non-vanishing terms in the Fourier Series.

Chapter 8

Impulses and Green's Functions

- **The physics:** Impulsive forces, driven oscillators with generic driving terms
- **The math:** The Dirac delta function, the Heaviside theta function, the method of Green's functions.

8.1 Introduction and Motivation

Over the course of the last two chapters of these notes, we have developed a variety of increasingly sophisticated techniques for solving inhomogeneous linear differential equations. In Chapter 6, we were introduced to the principle of superposition. This principle states that whenever the inhomogeneous term takes the form of a direct sum — *i.e.*, whenever $f(x) = f_1(x) + f_2(x) + f_3(x) + \ldots$ — there exists a particular solution of the form $y_p(x) = y_{p1}(x) + y_{p2}(x) + y_{p3}(x) + \ldots$, where $y_{pi}(x)$ is the solution to the corresponding differential equation with $f_i(x)$ in place of $f(x)$. In Chapter 7, we saw that it was possible to express nearly any periodic function as a direct sum. Specifically, we saw that it was possible to express such a function as an infinite sum of sine and cosine functions (or, alternatively, complex exponentials) with different frequencies — a construction known as a Fourier series. Furthermore, we even saw that non-periodic functions could be represented as a "Fourier series" of sorts — provided we were willing to include a continuous spectrum of frequencies in the sum. Thus, Fourier decomposition provided us with a method of solving essentially any inhomogeneous linear differential equation of up to second order.

In this chapter, we will examine another, equally powerful technique for decomposing an arbitrary function into a set of individual functions for which the corresponding differential equation is more tractable. This technique is called the method of Green's functions. Instead of decomposing functions into sine and cosine functions, we decompose them into a continuous spectrum of infinitely thin, infinitely high spikes called Dirac delta functions. As with Fourier analysis, the method of Green's functions provides us with a tool which can be used to evaluate the particular solution to nearly any inhomogeneous linear differential equation. Green's functions are likewise ubiquitous in physics: they are used extensively in electromagnetic theory, fluid dynamics, astrophysics, particle physics, and condensed-matter theory.

8.2 The Dirac Delta Function

The **Dirac delta function** (sometimes just referred to as the delta function), usually written $\delta(x)$, is a generalized mathematical function[1] whose value is defined to be infinite at precisely $x = 0$ but zero everywhere else:

$$\delta(x) \equiv \begin{cases} 0 & x \neq 0 \\ \infty & x = 0 \end{cases} . \tag{8.1}$$

[1]The Dirac delta function is not truly a function in the precise, technical sense, but rather what mathematicians call a generalized function or distribution. However, this distinction is not particularly important for any of the physics applications we'll be exploring in this chapter.

You can think of the Dirac delta function as an infinitely high, infinitely narrow spike located at the origin, as illustrated in Fig. 8.1. There is one more crucially important property that defines the Dirac delta function, which is that the area under the "spike" is defined to be 1. In other words, the integral of $\delta(x)$ over even an infinitesimally small range $-\epsilon \leq x \leq \epsilon$ centered around the origin is 1:

$$\int_{-\epsilon}^{\epsilon} \delta(x)dx \;=\; 1 \;. \tag{8.2}$$

However, since $\delta(x)$ evaluates to zero everywhere else except at the origin, the contribution to the integral from all other values of x is zero. It therefore follows that if the interval over which we're integrating $\delta(x)$ includes the origin, the value of the integral is exactly 1; if not, the value of the integral is zero:

$$\int_{a}^{b} \delta(x)dx \;=\; \begin{cases} 1 & a \leq 0 \leq b \\ 0 & \text{otherwise} \;. \end{cases} \tag{8.3}$$

For example, consider the three points x_1, x_2, and x_3 indicated on the x-axis in Fig. 8.1. If we were to integrate from x_1 to x_3 or from x_2 to x_3, we would get

$$\int_{x_1}^{x_3} \delta(x)dx \;=\; \int_{x_2}^{x_3} \delta(x)dx \;=\; 1 \tag{8.4}$$

whereas if we were to integrate from x_1 to x_2, we would get

$$\int_{x_1}^{x_2} \delta(x)dx \;=\; 0 \;, \tag{8.5}$$

because this interval doesn't include the origin.

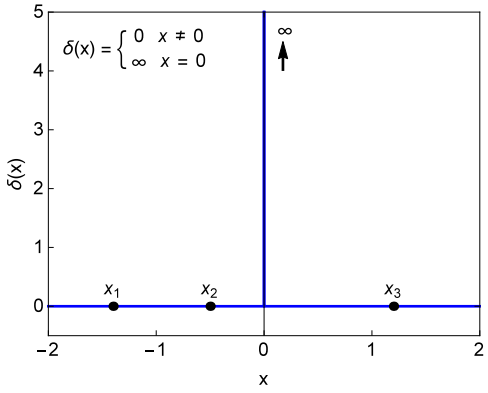


Figure 8.1: Illustration of the Dirac Delta function $\delta(x)$. The "function" consists of an infinitely high spike at $x = 0$, but evaluates to zero for all other values of x.

Now that we've been introduced to the Dirac Delta function, let's see why it's useful. In particular, let's consider what happens when we integrate the product of $\delta(x)$ with some other function $f(x)$ over some interval:

$$\int_{a}^{b} f(x)\delta(x)dx \;. \tag{8.6}$$

The delta function is zero everywhere except at $x = 0$, so the only non-vanishing contribution to this integral comes from the single point $x = 0$. At this point $f(x)$ evaluates to some particular value $f(0)$. Since the only contribution to the integral comes from this point, we can pull $f(0)$ outside the integral. We therefore find that

$$\int_{a}^{b} \delta(x)dx \;=\; f(0)\int_{a}^{b} \delta(x)dx \;=\; \begin{cases} f(0) & a \leq 0 \leq b \\ 0 & \text{otherwise} \;. \end{cases} \tag{8.7}$$

In other words, we see that the delta function serves to "pick out" the value of the function $f(x)$ at the point $x = 0$ when we integrate over an interval which includes this point. It's also straightforward to generalize Eq. (8.7) and "pick out" the value of $f(x)$ at any other value of x. Since $\delta(x)$ is only non-zero at the point $x = 0$, the function $\delta(x - c)$ will only be non-zero at the point $x = c$. Thus, we have

$$\int_a^b f(x)\delta(x-c)dx \;=\; \begin{cases} f(c) & a \le c \le b \\ 0 & \text{otherwise} \end{cases} . \tag{8.8}$$

We can also use what we know about the fundamental properties of Delta functions in order to evaluate integrals in which the argument of the delta function is a more complicated expression than just x or $x - c$. For example, let's say we wanted to evaluate the integral

$$\int_{-2\pi}^{2\pi} \cos(4x)\delta(-2x)dx \; . \tag{8.9}$$

How would we do this? The easiest way is to perform a change of variables to rewrite our expression in the form appearing in Eq. (8.7). In particular, if we define the variable $u = -2x$, we have

$$du \;=\; \frac{du}{dx}dx \;=\; -\,2dx \; , \tag{8.10}$$

and our integral becomes

$$\begin{aligned} \int_{-2\pi}^{2\pi} \cos(4x)\delta(-2x)dx &= \int_{4\pi}^{-4\pi} \cos(-2u)\delta(u)\left(-\frac{1}{2}du\right) \\ &= \frac{1}{2}\int_{-4\pi}^{4\pi} \cos(2u)\delta(u)du \\ &= \frac{1}{2}\cos(0) \\ &= \frac{1}{2} \; . \end{aligned} \tag{8.11}$$

It's important to note that the minus sign that we picked up from changing the measure of integration from dx to du was canceled by our switching the limits of integration in the second line of Eq. (8.11). However, if the argument of the delta function in Eq. (8.9) had been $2x$ instead of $-2x$, we wouldn't have had to switch the limits of integration and would have ended up with the same result.

It's not too difficult to generalize from this example and write down a formula for the integral of the product of $f(x)$ with the delta function $\delta(cx)$, where c is a constant. In particular, we can perform a similar change of variables $u = cx$ to get

$$\int_a^b f(x)\delta(cx)dx \;=\; \int_a^b f\left(\frac{u}{c}\right)\frac{1}{|c|}\delta(u)du \; . \tag{8.12}$$

The absolute value of c appears in this expression rather than c because switching the limits of integration cancels the minus sign we pick up from changing the measure of integration from dx to du, as discussed above. Finally, since the variable u is just a dummy variable, we can rename it x and write

$$\begin{aligned} \int_a^b f(x)\delta(cx)dx &= \int_a^b f\left(\frac{x}{c}\right)\frac{1}{|c|}\delta(x)du \\ &= \int_a^b f(x)\frac{1}{|c|}\delta(x)dx \; , \end{aligned} \tag{8.13}$$

where in going from the first to the second line, we have used the fact that $x = 0$ wherever $x/c = 0$. Comparing this expression with our original expression on the left-hand side of Eq. (8.12), we see that

$$\delta(cx) \;=\; \frac{1}{|c|}\delta(x) \; . \tag{8.14}$$

8.3 Impulses

There are a lot of situations in physics in which a force acts for only a very short duration but causes a dramatic change in the velocity of some object. Some examples include the force that the floor exerts on a basketball when it bounces, the force between two colliding objects, and the force you exert when you push someone on a swing set. It's often useful to model a force like this as an **impulse** — *i.e.*, a force which acts for an infinitesimal amount of time but which is infinite so that it can nevertheless produce a change in the motion of the object on which is acts. Such impulsive forces can be modeled by a Dirac delta function

$$F(t) \;=\; J\delta(t-t_*)\;, \tag{8.15}$$

where t_* is the time at which the force acts, and where J is a constant with dimensions $[J] = \mathrm{kg\,m/s}$. The integral of this impulse over any interval $t_i \le t \le t_f$ which includes t_* is

$$\int_{t_i}^{t_f} F(t)dt \;=\; J\int_{t_i}^{t_f}\delta(t-t_*)dt \;=\; J\;. \tag{8.16}$$

Now let's see what happens when an impulse acts within a physical system and alters the trajectories of the objects within that system. In particular, let's once again consider a damped, driven harmonic oscillator which consists of a block attached to a spring and ask what the solutions look like when the driving force which acts on the block takes the form of an impulse. Our equation of motion in this case is

$$\frac{dx^2}{dt^2} + 2\beta\frac{dx}{dt} + \omega_0^2 x \;=\; \frac{F(t)}{m} \;=\; \frac{J}{m}\delta(t-t_*)\;. \tag{8.17}$$

For simplicity, we'll focus on the case in which $\beta < \omega_0$ and the oscillator is underdamped. However, the overdamped and critically-damped cases can be analyzed in a completely analogous way. At times $t < t_*$, the delta function on the right-hand side of Eq. (8.17) evaluates to zero, so Eq. (8.17) is just

$$\frac{dx^2}{dt^2} + 2\beta\frac{dx}{dt} + \omega_0^2 x \;=\; 0\;. \tag{8.18}$$

This is just the equation of motion for a damped harmonic oscillator, so we already know that the solutions are

$$x_-(t) \;=\; e^{-\beta t}\big[A_-\cos(\omega_1 t) + B_-\sin(\omega_1 t)\big]\;, \tag{8.19}$$

where the subscripts on the coefficients A_1 and B_- indicate that these are the solutions for $t < t_*$ — *i.e.*, at times before the impulse acts on the system. Likewise, at times $t > t_*$, the delta function also evaluates to zero, so the solutions have exactly the same form

$$x_+(t) \;=\; e^{-\beta t}\big[A_+\cos(\omega_1 t) + B_+\sin(\omega_1 t)\big]\;. \tag{8.20}$$

Now comes the point where we need to take into account the effect of the impulse. Since the time interval during which this impulse is "turned on" is infinitesimally short, it doesn't correspond to an additional piece of our piecewise soluion for $x(t)$; rather, it corresponds to a modification of the boundary conditions which join $x_-(t)$ to $x_+(t)$ at t_*. The boundary condition for the position of the block is straightforward to evaluate. Clearly, the block cannot instantaneously change its position, no matter how strong the impulse might be, so we must have

$$x_+(t_*) \;=\; x_-(t_*) \tag{8.21}$$

The boundary condition for the velocity $v = dx/dt$ is a bit trickier to evaluate because the the velocity is not continuous at t_*. Rather, the impulsive force acts to produce an instantaneous change in the momentum p — and hence the velocity — of the block. Thus, the impulse leads to a discontinuity

$$\Delta v \;\equiv\; v_+(t_*) - v_-(t_*) \tag{8.22}$$

in the velocity of the block at $t = t_*$. In order to derive this discontinuity, we integrate both sides Eq. (8.17) over an infinitesimal region centered around $t_* = 0$:

$$\int_{-\epsilon}^{\epsilon}\left[\frac{dx^2}{dt^2} + 2\beta\frac{dx}{dt} + \omega_0^2 x\right]dt \;=\; \int_{-\epsilon}^{\epsilon}\frac{J}{m}\delta(t-t_*)dt\;. \tag{8.23}$$

Using the fundamental theorem of calculus to evaluate the first two terms on the left-hand side of this equation and using Eq (8.8) to evaluate the term on the right-hand side, we find that

$$\left.\frac{dx}{dt}\right|_{-\epsilon}^{\epsilon} + 2\beta x\Big|_{-\epsilon}^{\epsilon} + \omega_0^2 \int_{-\epsilon}^{\epsilon} x dt \;=\; \frac{J}{m}\,. \tag{8.24}$$

The first term on the left-hand side of this equation represents the difference in velocity across the discontinuity at $t = t_*$. By contrast, the second two terms, on this side of the equation vanish in the $\epsilon \to 0$ limit because, as we've already discussed, the position $x(t)$ itself is continuous at t_*. Thus, if we take the $\epsilon \to 0$ limit of Eq. (8.24), we find that the discontinuity in the velocity of the block at t_* is

$$\Delta v \;=\; \left.\frac{dx}{dt}\right|_{-\epsilon}^{\epsilon} \;=\; \frac{J}{m}\,. \tag{8.25}$$

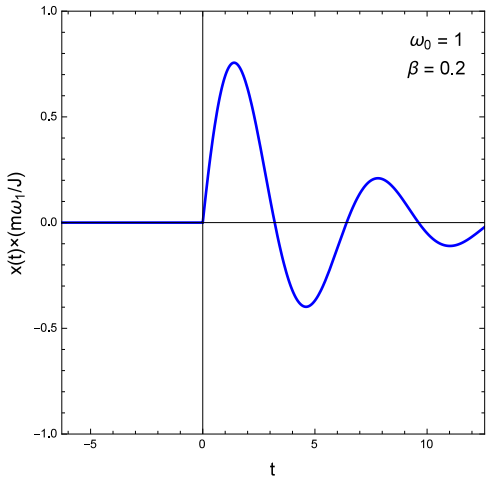


Figure 8.2: The pieceiwse solution $x(t)$ for an underdamped harmonic oscillator which is disturbed from its equilibrium state by an impulse $F(t) = J\delta(t)$ at $t = 0$, normalized to $J/m\omega_1$. The trajectory shown here corresponds to the parameter choices $\omega_0 = 1$ and $\beta = 0.2$.

Now that we know the discontinuity Δv in the velocity at t_*, we're ready to apply the boundary conditions in Eqs. (8.21) and (8.22) and derive our piecewise solution for $x(t)$. To make things simpler, we'll choose our time coordinate so that $t_* = 0$. Moreover, let's focus on the case in which the block is initially at rest at at its equilibrium point at some initial $t_0 < 0$ and is only later set into motion by the action of the impulse. In this case, the boundary conditions for $x_-(t)$ at $t = t_0$ are $x_-(t_0) = 0$ and $v_-(t_0) = 0$. This gives us the trivial solution

$$x_-(t) \;=\; 0 \tag{8.26}$$

for times $t < 0$. However, at $t = 0$, the impulse acts and causes an instantaneous jump in velocity given by Eq. (8.25). As a result, we have

$$x_+(0) \;=\; 0\,, \qquad v_+(0) \;=\; \Delta v \;=\; \frac{J}{m}\,. \tag{8.27}$$

We know the relationship between the initial position and velocity values $x_+(0)$ and $v_+(0)$ and the coefficients A_+ and B_+ in Eq. (8.20) — indeed, we already derived this relationship in Chapter 5 of these notes. Plugging in the values from Eq. (8.27) gives us

$$A_+ \;=\; x_+(0) \;=\; 0\,, \qquad B_+ \;=\; \frac{v_+(0) + \beta x_+(0)}{\omega_1} \;=\; \frac{J}{m\omega_1}\,. \tag{8.28}$$

Thus, we find that the solution for $t > 0$ takes the form

$$x_+(t) \;=\; e^{-\beta t}\frac{J}{m\omega_1}\sin(\omega_1 t)\;. \tag{8.29}$$

By stitching the individual pieces in Eqs. (8.26) and (8.29) together, we finally arrive at our piecewise solution for $x(t)$:

$$x(t) \;=\; \begin{cases} 0 & t < 0 \\ e^{-\beta t}\frac{J}{m\omega_1}\sin(\omega_1 t) & t \geq 0 \end{cases}\;. \tag{8.30}$$

We can write this expression in a slightly more compact form by defining the function

$$\Theta(t) \;\equiv\; \begin{cases} 0 & t < 0 \\ \frac{1}{2} & t = 0 \\ 1 & t > 0 \end{cases} \tag{8.31}$$

This function is called the **Heaviside theta function**, and it allows us to write Eq. (8.30) in the form

$$x(t) \;=\; e^{-\beta t}\frac{J}{m\omega_1}\sin(\omega_1 t)\Theta(t)\;. \tag{8.32}$$

In Fig. 8.2, we show what this solution (normalized to the value of the quantity $J/m\omega_1$) looks like as a function of time for the choice of parameters $\omega_0 = 1$ and $\beta = 2$. For $t_0 < t < 0$, the block simply sits at equilibrium. At $t = 0$, the impulse endows the block with a momentum $\Delta p = J$ and it starts undergoing damped oscillations around the equilibrium point.

It's not difficult to generalize our result in Eq. (8.32) to the case of an impulse which acts at a time $t_* \neq 0$. You can easily verify (see Problem 3) for yourself that the corresponding solution for general t_* is

$$x(t) \;=\; e^{-\beta(t-t_*)}\frac{J}{m\omega_1}\sin[\omega_1(t-t_*)]\Theta(t-t_*)\;. \tag{8.33}$$

8.4 Green's Functions

In Sect. 8.3, we saw how to solve linear differential equations in which the inhomogeneous term takes the form of a delta function. As you might imagine, it's not difficult to generalize this to the case in which multiple different impulses act on the system. Let's say, for example, that in the example from the previous section, we had given our block a "push" not merely once, but several times in succession. In other words, let's say that the force $F(t)$ applied to the block had consisted not of a single impulse, but rather of N independent impulses, each of which acts at a different time t_n and has its own particular strength J_n, where the index $n = 1, 2, \ldots, N$ labels the impulse in chronological order. In this case, $F(t)$ takes the form of a sum of Dirac delta functions representing these individual impulses

$$F(t) \;=\; \sum_{n=1}^{N} J_n\delta(t-t_n)\;. \tag{8.34}$$

What would the particular solution to such a differential equation look like? In order to answer this question, we once again turn to the principle of superposition for guidance. This principle tells us that there exists particular solution to the corresponding differential equation of the form

$$x_p(t) \;=\; \sum_{n=1}^{N} x_{pn}(t)\;, \tag{8.35}$$

where $x_{pn}(t)$ is the solution to the corresponding differential equation with the full inhomogeneous term $F(t)$ replaced by the single delta function $J_n\delta(t-t_n)$. Each of these individual equations has exactly the same form as Eq. (8.17), so we know that the corresponding solution $x_{pn}(t)$ to each equation must likewise

have the same form as Eq. (8.33). Thus, the solution to the equation with the full inhomogeneous term in Eq. 8.34 is

$$x_p(t) \;=\; \sum_{n=0}^{N} e^{-\beta(t-t_n)} \frac{J_n}{m\omega_1} \sin[\omega_1(t-t_n)]\Theta(t-t_n) \;. \tag{8.36}$$

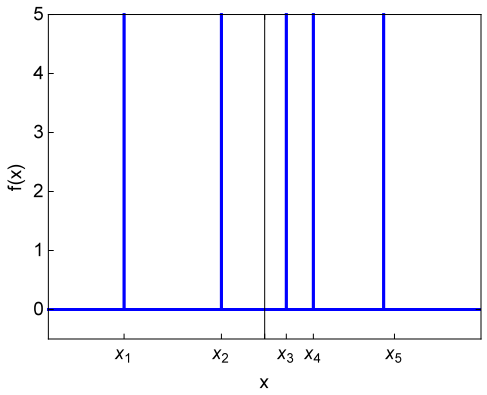


Figure 8.3: Illustration of a function $f(x)$ which consists of a sum of multiple different delta functions, as in Eq. (8.37)

We have thus far been speaking about summing over delta functions in the context of one particular physics example involving impulsive forces. However, these same principles apply more broadly to other mathematically analogous contexts as well. In particular, whenever we have an inhomogeneous linear differential equation in which the inhomogeneous term can be written as a sum of delta functions — *i.e.*, in which

$$f(x) \;=\; \sum_{n=1}^{N} f_n \delta(x-x_n) \;, \tag{8.37}$$

we can use the same general procedure that we used in Sect. 8.3 in order to determine the solution in a fairly straightforward manner using the principle of superposition.

All well and good, but now we're going to take this one step further. In particular, let us consider the case in which our function $f(x)$ is not merely a sum of some finite number of impulses weighted by some set of coefficients f_n, but rather a "sum" — *i.e.*, an integral — over a continuous spectrum of delta functions $f(x')\delta(x-x')$ weighted by a set of coefficients $f(x')$, where the location of the delta-function spike is labeled by the continuous variable x'. In other words, let us consider functions of the form

$$f(x) \;=\; \int_{-\infty}^{\infty} f(x')\delta(x-x')dx' \;, \tag{8.38}$$

I want to stress that mathematically we've done nothing new or daring here. In fact, this equation is really just a restatement of Eq. (8.8) with x' as the integration variable and $x = c$.[2] Thus, *any* function $f(x)$ can be written this way.

Conceptually, however, this equation is actually quite profound. It says that *any* function $f(x)$ can be decomposed into continuum of Dirac delta functions $\delta(x-x')$ labeled by the continuous variable x', each with a coefficient $f(x')$. This decomposition is illustrated in Fig. 8.4. It's important to remember that each delta function in this decomposition is still infinitely high — it's the act of *integrating* over these delta functions that gives us a finite value for $f(x)$ at each value of x. You can think of this decomposition as being similar to the Fourier transform $F(\omega)$ we saw in Chapter 7 in that it's a way of representing a function in terms of a continuum of constituent functions labeled by some continuous variable. For the Fourier transform, that

[2] You've already shown in your homework that $\delta(x-c) = \delta(c-x)$, so the minus sign in the argument of the delta function relative to Eq. (8.8) doesn't matter.

variable is the frequency ω of the Fourier mode; for this delta-function decomposition, it's the location x' of the delta-function spike. The difference is that we're decomposing our function into a set of such spikes instead of sine waves or complex exponentials.

What is the advantage of expressing a function in terms of a continuum of Dirac delta functions? Once again, just as with Fourier transforms, the answer has everything to do with the principle of superposition. Let's say, for example, that we we trying to solve an inhomogeneous linear equation of the form

$$\frac{d^2y}{dx^2} + P(x)\frac{dy}{dx} + Q(x)y \;=\; f(x) \;, \tag{8.39}$$

where the inhomogeneous term $f(x)$ doesn't have one of the forms for which we can easily guess the solution and use the method of undetermined coefficients. If we use Eq. (8.38) to re-express $f(x)$ in the form

$$\frac{d^2y}{dx^2} + P(x)\frac{dy}{dx} + Q(x)y \;=\; \int_{-\infty}^{\infty} f(x')\delta(x-x')dx' \;, \tag{8.40}$$

we see that the right-hand side takes the form of a direct sum — a sum over a continuous variable x', mind you (*i.e.*, an integral), but a sum nonetheless! This means that the principle of superposition applies, and it tells us that there exists a particular solution $y_p(x)$ to this equation which takes the form of a sum — or in this case an integral, since x' is continuous — over the particular solutions to the corresponding differential equations with $f(x)$ replaced a single delta-function term. We have already seen that differential equations of this sort are quite tractable — indeed, we just solved one in Sect. 8.15. Once we've obtained the solution to this delta-function equation for an arbitrary value of x' — we'll call this solution $Y_p(x;x')$, since it depends on both x and x' — obtaining the solution to the full differential equation in Eq. (8.39) is simply a matter of integrating over x':

$$y_p(x) \;=\; \int_{-\infty}^{\infty} Y_p(x;x')dx' \;. \tag{8.41}$$

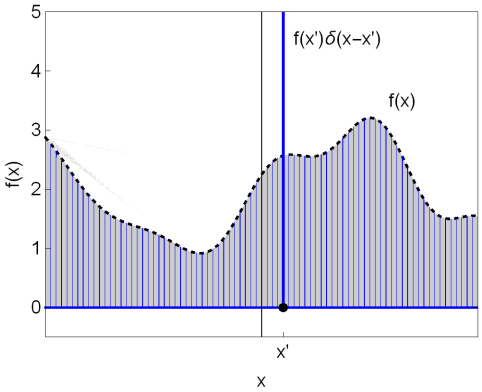


Figure 8.4: Illustration of how a function $f(x)$ is decomposed into a continuum of infinitely high delta-function spikes $f(x')\delta(x-x')$, weighted by the value of the function at each point. The integral over this continuum of delta functions returns a finite value for $f(x)$ at every value of x.

This strategy for solving inhomogeneous linear differential equations by decomposing the inhomogeneous term $f(x)$ into a continuum of delta functions is the basis for the method of Green's functions. The **Green's function** $G(x;x')$ for a a differential equation is just the solution to the corresponding equation with $f(x)$ replaced by a single, unweighted delta function. For example, for a generic second-order linear equation of the form given in Eq. (8.39), $G(x;x')$ is a solution to the equation

$$\frac{d^2G(x;x')}{dx^2} + P(x)\frac{dG(x;x')}{dx} + Q(x)G(x;x') \;=\; \delta(x-x') \tag{8.42}$$

The Green's function is useful because no matter what form $f(x)$ happens to take, the solution $Y_p(x;x')$ for each value of x' is just the Green's function, weighted by the value of the function $f(x)$ at the point $x = x'$ where the corresponding delta function $\delta(x - x')$ is non-zero:

$$Y_p(x;x') \;=\; f(x')G(x;x') \;. \tag{8.43}$$

The particular solution $y_p(x)$ to the full differential equation can therefore be written in the form

$$y_p(x) \;=\; \int_{-\infty}^{\infty} Y_p(x;x')dx' \;=\; \int_{-\infty}^{\infty} f(x')G(x;x')dx' \;. \tag{8.44}$$

Thus, if we can determine the form of the Green's function which corresponds to a particular differential equation, all we need to do in order to find a solve that equation is to evaluate the integral in Eq. (8.44).

If you're skeptical about this procedure or about the underlying idea of taking the principle of superposition to the continuum limit, there are a number of reassuring cross-checks that we can perform in order to show that it works. For example, it's not difficult to verify that the expression for $y_p(x)$ in Eq. (8.44) is indeed a particular solution to our original differential equation. If we simply plug this alleged solution back into Eq. (8.39), we find

$$\begin{aligned} f(x) \;&\stackrel{?}{=}\; \frac{d^2}{dx^2}\int_{-\infty}^{\infty} f(x')G(x;x')dx' + P(x)\frac{d}{dx}\int_{-\infty}^{\infty} f(x')G(x;x')dx' + Q(x)\int_{-\infty}^{\infty} f(x')G(x;x')dx' \\ &\stackrel{?}{=}\; \int_{-\infty}^{\infty}\left[\frac{d^2G(x;x')}{dx^2} + P(x)\frac{dG(x;x')}{dx} + Q(x)G(x;x')\right] f(x')dx' \\ &\stackrel{?}{=}\; \int_{-\infty}^{\infty}\delta(x-x')f(x')dx' \\ &=\; f(x) \;, \end{aligned} \tag{8.45}$$

where in going from the second to the third line, we have used Eq. (8.42). Thus, we see that $y_p(x)$ indeed satisfies our original differential equation in Eq. (8.39) for all x. This is yet another testament to how useful the principle of superposition can be — even in the continuum limit! Thus we see that, Green's functions, like Fourier series, provide us with another powerful tool for solving otherwise seemingly intractable differential equations.

Example: Green's Function for a Damped, Driven Harmonic Oscillator

As an example of how to derive the Green's function for a particular differential equation, let's determine $G(x;x')$ for the damped, driven harmonic oscillator with a driving force $F(x)$. Once again, for purposes of illustration, we'll focus on the underdamped case. The equation of motion for this system is

$$\frac{d^2x}{dt^2} + 2\beta\frac{dx}{dt} + \omega_0^2 x \;=\; \frac{\omega_0^2}{k}F(t) \;, \tag{8.46}$$

The Green's function $G(t;t')$ is the solution to the corresponding equation with the inhomogeneous term replaced by a delta function:

$$\frac{d^2G(t;t')}{dt^2} + 2\beta\frac{dG(t;t')}{dt} + \omega_0^2 G(t;t') \;=\; \delta(t-t') \;. \tag{8.47}$$

This equation has exactly the same form as Eq. (8.17), with $J/m \to 1$ and $t_*- \to t'$, so the solution for $G(t;t')$ is the solution we derived for $x(t)$ in Eq. (8.33) with the corresponding replacements for J/m and t_*:

$$G(t;t') \;=\; e^{-\beta(t-t')}\frac{1}{\omega_1}\sin[\omega_1(t-t')]\Theta(t-t') \;. \tag{8.48}$$

8.5 Determining the Green's Functions

Admittedly, we didn't have to do much work in order to find the Green's function for the damped, driven harmonic-oscillator equation in the example given above, since we had basically already derived it in Sect. 8.3 up to an overall constant. However, the procedure for finding the Green's function for an arbitrary second-order linear differential equation of the form given in Eq. (8.39) is pretty much analogous. The corresponding equation which the Green's function must satisfy is

$$\frac{d^2G(x;x')}{dx^2} + P(x)\frac{dG(x;x')}{dx} + Q(x)G(x;x') \;=\; \delta(x-x') \;. \tag{8.49}$$

For every value of x other than $x = x'$, the delta function evaluates to zero, so for $x < x'$ or $x > x'$, this equation reduces to the *complementary* equation

$$\frac{d^2G(x;x')}{dx^2} + P(x)\frac{dG(x;x')}{dx} + Q(x)G(x;x') \;=\; 0 \;. \tag{8.50}$$

Provided that we can find the general solution to this homogeneous equation, we can use this general solution to construct a piecewise solution for $G(x;x')$ in the same way we constructed our solution $x(t)$ for the damped, driven harmonic-oscillator equation Sect. 8.3.

We can make some more concrete statements about the form of this piecewise solution based on some things we already know about the properties of general solutions to equations like Eq. (8.50) In particular we saw in Chapter 5 of these notes that we can always write the general solution $y(x)$ to a second-order homogeneous linear differential equation as the sum of two linearly independent solutions with arbitrary coefficients

$$y(x) \;=\; C_1y_1(x) + C_2y_2(x) \;, \tag{8.51}$$

where values of C_1 and C_2 are determined by the boundary conditions. Therefore, since $G(x;x')$ satisfies Eq. (8.50) on either side of the delta-function "spike," we know that our piecewise solution for $G(x;x')$ can be written in the form

$$G(x;x') \;=\; \begin{cases} A_-y_1(x) + B_-y_2(x) & x < x' \\ A_+y_1(x) + B_+y_2(x) & x > x' \;. \end{cases} \tag{8.52}$$

The coefficients A_- and B_- which characterize the solution for $x < x'$ and the coefficients A_+ and B_+ which characterize the solution for $x > x'$ will in general be different.

The next step is to impose the boundary conditions at $x = x'$ and thereby join the two pieces of our piecewise solution in Eq. (8.52) together in a manner that properly accounts for the effect of the delta function. These boundary conditions provide us with two constraints on the four unknown coefficients which appear in this equation. The remaining two constraints that we would need in order to completely determine all of these coefficients would have to come from some other input such as, for example, a set of boundary conditions for $G(x;x')$ at $x \to -\infty$. However, we are once again only interested in obtaining a particular solution to Eq. (8.49), so as long as our solution for $G(x;x')$ respects the boundary conditions at $x = x'$, we can choose any convenient values for the remaining two coefficients. One particularly convenient choice is to set $A_- = B_- = 0$ so that Eq. (8.52) reduces to

$$G(x;x') \;=\; \begin{cases} 0 & x < x' \\ A_+y_1(x) + B_+y_2(x) & x > x' \;. \end{cases} \tag{8.53}$$

Now let's apply the boundary conditions at $x = x'$. First of all, we know that $G(x;x')$ itself must be continuous across the boundary, so the two pieces of our piecewise solution must be equal at $x = x'$. This give us

$$0 \;=\; A_+y_1(x') + B_+y_2(x') \;. \tag{8.54}$$

The boundary condition relating the derivatives of the two pieces of $G(x;x')$ across the boundary is a little bit more complicated because the delta function introduces a discontinuity in the x-derivative of $G(x;x')$ at this point. We can determining this discontinuity in essentially the same way that we determined the discontinuity in $v(t)$ for the impulse-driven oscillator in Sect. 8.3. We simply integrate both sides of Eq. (8.49)

over an infinitesimal range of x values centered around $x = x'$ and then take the $\epsilon \to 0$ limit of the resulting expression. Doing so, we find that

$$\begin{aligned}
\int_{x'-\epsilon}^{x'+\epsilon} \left[\frac{d^2G(x;x')}{dx^2} + P(x)\frac{dG(x;x')}{dt} + Q(x)G(x;x') \right] dx &= \int_{x'-\epsilon}^{x'+\epsilon} \delta(x-x')dx \\
\frac{dG(x;x')}{dx}\bigg|_{x'-\epsilon}^{x'+\epsilon} + \int_{x'-\epsilon}^{x'+\epsilon} \left[P(x)\frac{dG(x;x')}{dx} + Q(x)G(x;x') \right] dx &= 1 \\
\frac{dG(x;x')}{dx}\bigg|_{x'-\epsilon}^{x'+\epsilon} + P(x)G(x;x')\bigg|_{x'-\epsilon}^{x'+\epsilon} + \int_{x'-\epsilon}^{x'+\epsilon} \left[Q(x)G(x;x') - \frac{dP(x)}{dx}G(x;x') \right] dx &= 1 \ , \qquad (8.55)
\end{aligned}$$

where in going from the second to the second line, we have used integration by parts. Now we know that $G(x;x')$ is continuous at $x = x'$, so as long as the function $P(x)$ and its first derivative are continuous at $x = x'$, the second and third terms on the left-hand side of this equation vanish in the $\epsilon \to 0$ limit. Thus, we find that the discontinuity in the x-derivative of $G(x;x')$ is

$$\Delta \left[\frac{dG(x;x')}{dx} \right] \equiv \lim_{\epsilon \to 0} \ , \frac{dG(x;x')}{dt}\bigg|_{x'-\epsilon}^{x'+\epsilon} = 1 \ . \qquad (8.56)$$

Applying this boundary condition to the function form of $G(x;x')$ in Eq. (8.53) gives us our second equation for the coefficients A_+ and B_+:

$$A_+ \frac{dy_1(x)}{dx} + B_+ \frac{dy_2(x)}{dx} = \Delta \left[\frac{dG(x;x')}{dx} \right] = 1 \ . \qquad (8.57)$$

We now have two equations — Eqs. (8.54) and (8.57) — and two unknowns, so we can solve the system and determine the values of A_+ and B_+. The result can be most cleanly expressed in the form

$$A_+ = -\frac{y_2(x')}{W(x')} \ , \qquad B_+ = \frac{y_1(x')}{W(x')} \ , \qquad (8.58)$$

where the quantity $W(x')$ appearing in the denominator in each of these expressions is the Wronksian

$$W(x) = y_1(x)\frac{dy_2(x)}{dx} - y_2(x)\frac{dy_1(x)}{dx} \qquad (8.59)$$

that we met in Chapter 5, evaluated at $x = x'$. Plugging this result into Eq. (8.53), we obtain our final formula for the Green's function for

$$G(x;x') = \frac{1}{W(x')}\Big[y_1(x')y_2(x) - y_2(x')y_1(x) \Big] \Theta(x-x') \ , \qquad (8.60)$$

where $\Theta(x-x')$ once again denotes the Heaviside theta function from Eq. (8.31).

Problems

1. Evaluate the following integrals:

   (a) $\displaystyle\int_1^2 \ln x \, \delta(4x+3)dx$

   (b) $\displaystyle\int_{-\infty}^{\infty} e^{3x}\delta(5x)dx$

   (c) $\displaystyle\int_{-10}^{4} (5x+3)\delta(e^{3x}-1)dx$.

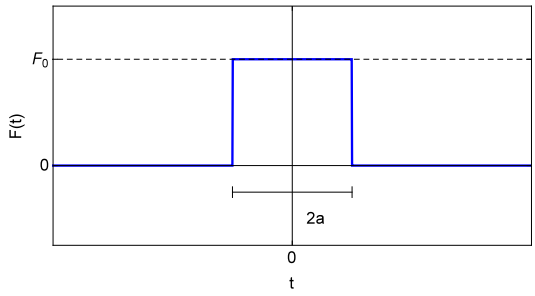


Figure 8.5: A driving force $F(t)$ which consists of of a square pulse with height F_0 and width $2a$.

2. Show that for $a \neq b$, we have

$$\delta\big[(x-a)(x-b)\big] \;=\; \left|\frac{1}{a-b}\right| \big[\delta(x-a)+\delta(x-b)\big] \;. \tag{8.61}$$

3. Verify that the solution in Eq. (8.33) satisfies the equation of motion in Eq. (8.17) for general t_* when the block begins from rest at its equilibrium position at some time $t_0 < t_*$.

4. In Sect. 8.3, we calculated the solution $x(t)$ for an underdamped, driven harmonic oscillator with an impulsive driving force $F(t) = J\delta(t-t_*)$.

   (a) Determine the corresponding solutions for an overdamped oscillator with the same impulsive driving force.

   (b) Determine the corresponding solutions for an critically-damped oscillator with the same impulsive driving force.

5. Show that the Dirac delta function is the derivative of the Heaviside theta function defined in Eq. (8.31). In other words, show that

$$\frac{d\Theta(t)}{dt} \;=\; \delta(t) \;. \tag{8.62}$$

6. Consider the differential equation

$$x^2\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + 2y \;=\; f(x) \;, \tag{8.63}$$

   where $f(x)$ is some function of the independent variable x.

   (a) Verify that $y_1(x) = x$ and $y_2(x) = x^2$ are solutions to the complementary equation.

   (b) Find the Green's function for this equation. Be careful about how you apply Eq. (8.60). This formula assumes that the original differential equation has the form Eq. (8.49), which is not the case here.

   (c) Use your result from part (b) to find a particular solution $y_p(t)$ to the original differential equation in the case where the inhomogeneous term has the form

$$f(x) \;=\; \begin{cases} 0 & x < a \\ x-a & x > a \;, \end{cases} \tag{8.64}$$

   where a is a positive constant.

7. Consider a underdamped harmonic oscillator which consists of a block attached to a spring, subject to a linear drag force. The oscillator is is driven by a force $F(t)$ which consists of a square pulse with width $2a$ and height F_0, as shown in Fig. 8.5.

(a) Use the method of Green's functions to find a piecewise solution for the position $x(t)$ of the block as a function of time, assuming the block starts from rest at its equilibrium position. This piecewise solution will consist of three pieces: one for $t < -a$, one for $-a < t < a$, and one for $t > a$.

(b) Use Mathematica to create a plot of this piecewise solution for $\omega_0 = 1\ \mathrm{s}^{-1}$, $\beta = 0.2\ \mathrm{s}^{-1}$, $F_0 = 1\ \mathrm{N/m}$, $m = 1\ \mathrm{kg}$, and $a = 1\ \mathrm{m}$. Make sure you adjust the range of your plot so that all three pieces of the solution are shown.

Chapter 9

Coupled Oscillations and Linear Systems of Equations

- **The physics:** Coupled oscillators, normal modes, beats, longitudinal and transverse oscillations, dispersion relations, phonons and vibrations.
- **The math:** Linear algebra, inner and outer products of vectors, matrix inversion, determinants and traces, eigenvalues and eigenvectors, diagonalization.

9.1 Systems of Differential Equations

Over the last several chapters of these lecture notes, we have developed a series of increasingly sophisticated techniques for solving differential equations. However, up to this point, we have focused on systems that can be described by a single differential equation involving a single dependent variable $y(x)$ and a single independent variable x. In this chapter, we will take a look at some examples of physical systems which are described by multiple differential equations involving some number $N > 1$ of dependent variables $y_i(x)$, where $i = \{1, \ldots, N\}$, all of which are functions of a single independent variable x.[1] We will examine a few techniques for solving coupled *systems* of differential equations — including systems of coupled harmonic oscillators. Finally, we shall see how wave phenomena can emerge out of the collective oscillations of large numbers of coupled oscillators.

9.2 Our First Coupled System: Two Oscillators

We'll begin our discussion of coupled oscillations by first examining a simple system consisting of two blocks, each of mass m, attached to each other and to a pair of fixed walls by a set of three springs, as shown in Fig. 9.1. Each of the springs has the same spring constant k. For now, let's assume that the effects of friction and other damping forces are negligible. In this case, the total force acting on each block at any given time t is simply the sum of the forces exerted by the two springs to which the block is directly attached.

Let's consider what happens when the blocks are set in motion. We'll begin by defining a pair of coordinates — which we'll call x_1 and x_2 — which represent the displacement of the blocks from their respective equilibrium positions $x_{1,\mathrm{eq}}$ and $x_{2,\mathrm{eq}}$, as shown in the bottom panel of Fig. 9.1. Provided that Hooke's Law holds, the force exerted by each spring is proportional to the total distance that the spring is compressed relative to its equilibrium length. At any given time t, the springs which connect the left and right blocks to the walls are compressed a distance $-x_1(t)$ and $x_2(t)$, respectively, whereas the spring between the two blocks is compressed a distance $x_1 - x_2$. Thus, for the block on the left, Newton's Second

[1] Increasing the number of *independent* variables takes us from the real of ordinary differential equations to the realm of partial differential equations, which is another topic entirely.

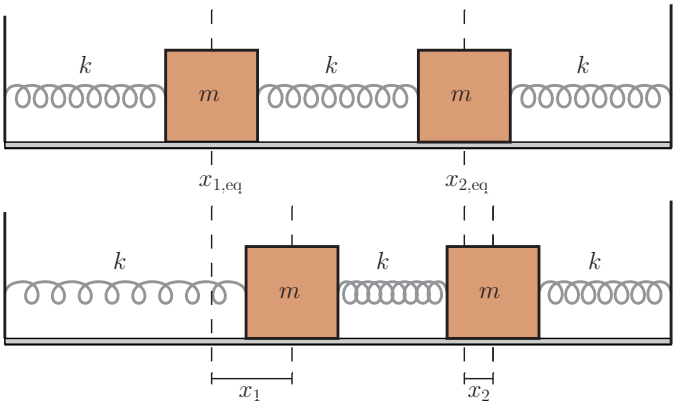


Figure 9.1: A pair of blocks, each with mass m, attached to each other and to the walls by springs with spring constant k. The top panel shows the blocks at equilibrium. The bottom panel shows the blocks oscillating. The coordinates x_1 and x_2 are defined relative to the initial positions $x_{1,\mathrm{eq}}$ and $x_{2,\mathrm{eq}}$ of the respective blocks.

Law gives

$$m\frac{d^2x_1}{dt^2} \;=\; -kx_1 - k(x_1 - x_2) \;=\; -2kx_1 + kx_2. \tag{9.1}$$

Similarly, for the block on the right, Newton's Law gives

$$m\frac{d^2x_2}{dt^2} \;=\; -kx_2 + k(x_1 - x_2) \;=\; -2kx_2 + kx_1. \tag{9.2}$$

The important thing to notice about about Eqs. (9.1) and (9.2) is that the spring force provided on each block by the central spring depends on both x_1 and x_2. The acceleration of the the left block therefore depends on the position of the right block and vice verse. In other words, these two equations of motion are **coupled** in the sense that we can't solve for $x_1(t)$ and $x_2(t)$ separately. Instead, we must simultaneously solve these two equations as a system in order to find the positions of either block as a function of time.

The easiest way to solve this system of equations is to see if we can decouple the system by combining Eq. (9.1) and Eq. (9.2) in such a way that the resulting equations can be solved separately. For example, if we add these two equations together, we get

$$m\frac{d^2}{dt^2}(x_1 + x_2) \;=\; -k(x_1 + x_2)\ . \tag{9.3}$$

We observe that $x_1(t)$ and $x_2(t)$ appear in this equation only in the combination $x_+(t) \equiv x_1(t) + x_2(t)$. We can therefore view it as an equation of motion for the quantity $x_+(t)$:

$$\frac{d^2x_+}{dt^2} \;=\; -\frac{k}{m}x_+\ . \tag{9.4}$$

Likewise, if we subtract Eq. (9.2) from Eq. (9.1), we get

$$m\frac{d}{dt}(x_1 - x_2) \;=\; -3k(x_1 - x_2)\ . \tag{9.5}$$

We observe that $x_1(t)$ and $x_2(t)$ appear in this equation only in the combination $x_-(t) = x_1(t) - x_2(t)$. Thus, we may view it as an equation of motion for the quantity x_-:

$$\frac{d^2x_-}{dt^2} \;=\; -\frac{3k}{m}x_-\ . \tag{9.6}$$

Now that we've decoupled the equations of motion for our coupled-oscillator system, solving for $x_1(t)$ and $x_2(t)$ is straightforward. Indeed, Eq. (9.4) and Eq. (9.6) both look like the equation of motion for a simple harmonic oscillator! Thus, we know that the general solutions for $x_\pm(t)$ can be written in the form

$$\begin{aligned} x_+(t) &= A_+ \cos(\omega_+ t + \phi_+) \\ x_-(t) &= A_- \cos(\omega_- t + \phi_-) \ , \end{aligned} \tag{9.7}$$

where we have defined the oscillator frequencies

$$\omega_+ \ \equiv \ \sqrt{\frac{k}{m}} \ , \qquad \omega_- \ \equiv \ \sqrt{\frac{3k}{m}} \ . \tag{9.8}$$

The amplitudes $A_\pm$ and phases $\phi_\pm$ of these solutions are determined by the boundary conditions, as usual. Now that we have the solutions for $x_+(t)$ and $x_-(t)$, it's easy to determine the solutions for our original coordinates $x_1(t)$ and $x_2(t)$ as well. Indeed, since the relationship between our two sets of coordinates is such that

$$x_1(t) \ = \ \frac{x_+(t) + x_-(t)}{2} \ , \qquad x_2(t) = \frac{x_+(t) - x_-(t)}{2} \ , \tag{9.9}$$

we find that the solutions for $x_1(t)$ and $x_2(t)$ are

$$x_1(t) = \frac{A_+}{2} \cos(\omega_+ t + \phi_+) + \frac{A_-}{2} \cos(\omega_- t + \phi_-) \tag{9.10}$$

$$x_2(t) = \frac{A_+}{2} \cos(\omega_+ t + \phi_+) - \frac{A_-}{2} \cos(\omega_- t + \phi_-) \ . \tag{9.11}$$

In order to visualize better what these solutions for $x_1(t)$ and $x_2(t)$ actually look like, let's choose a particular set of boundary conditions for our system so that we can plot the results. In particular, let's consider a situation in which both of the blocks are initially at rest at time $t = 0$. Moreover, let's say that one of them — let's make it the block on the left — is initially displaced from equilibrium by an amount a, while the other is initially at its equilibrium position. The boundary conditions that correspond to this initial physical state of the system are

$$v_1(0) \ = \ v_2(0) \ = \ 0 \ , \qquad x_1(0) \ = \ a \ , \qquad x_2(0) \ = \ 0 \ . \tag{9.12}$$

We'll start by applying the boundary condition on the initial velocities. Differentiating the expressions in Eqs. (9.10) and (9.11) gives the general expressions for $v_1(t)$ and $v_2(t)$:

$$v_1(t) = -\frac{A_+\omega_+}{2} \sin(\omega_+ t + \phi_+) - \frac{A_-\omega_-}{2} \sin(\omega_- t + \phi_-) \tag{9.13}$$

$$v_2(t) = -\frac{A_+\omega_+}{2} \sin(\omega_+ t + \phi_+) + \frac{A_-\omega_-}{2} \sin(\omega_- t + \phi_-) \ . \tag{9.14}$$

The boundary conditions at $t = 0$ require that

$$\begin{aligned} -\frac{A_+\omega_+}{2} \sin(\phi_+) - \frac{A_-\omega_-}{2} \sin(\phi_-) &= v_2(0) \ = \ 0 \\ -\frac{A_+\omega_+}{2} \sin(\phi_+) + \frac{A_-\omega_-}{2} \sin(\phi_-) &= v_2(0) \ = \ 0 \ . \end{aligned} \tag{9.15}$$

These equations can both be satisfied by taking $\phi_+ = \phi_- = 0$. Plugging this result back into Eqs. (9.10) and (9.11) and imposing the boundary conditions on the $x_1(0)$ and $x_2(0)$ gives us a pair of equations for A_+ and A_-:

$$\frac{A_+ + A_-}{2} \ = \ x_1(0) \ = \ a \ , \qquad \frac{A_+ - A_-}{2} \ = \ x_2(0) \ = \ 0 \ . \tag{9.16}$$

The solution is

$$A_+ \ = \ A_- \ = \ a \ . \tag{9.17}$$

Substituting this result back into Eqs. (9.10) and. (9.11) gives us our final result expressions for $x_1(t)$ and $x_2(t)$ for this choice of boundary conditions:

$$x_1(t) = a\Big[\cos(\omega_+ t) + \cos(\omega_- t)\Big] \tag{9.18}$$

$$x_2(t) = a\Big[\cos(\omega_+ t) - \cos(\omega_- t)\Big] \ . \tag{9.19}$$

These solutions are plotted in Fig. 9.2 for the parameter choices choice $a = 1$ and $k/m = 1$.

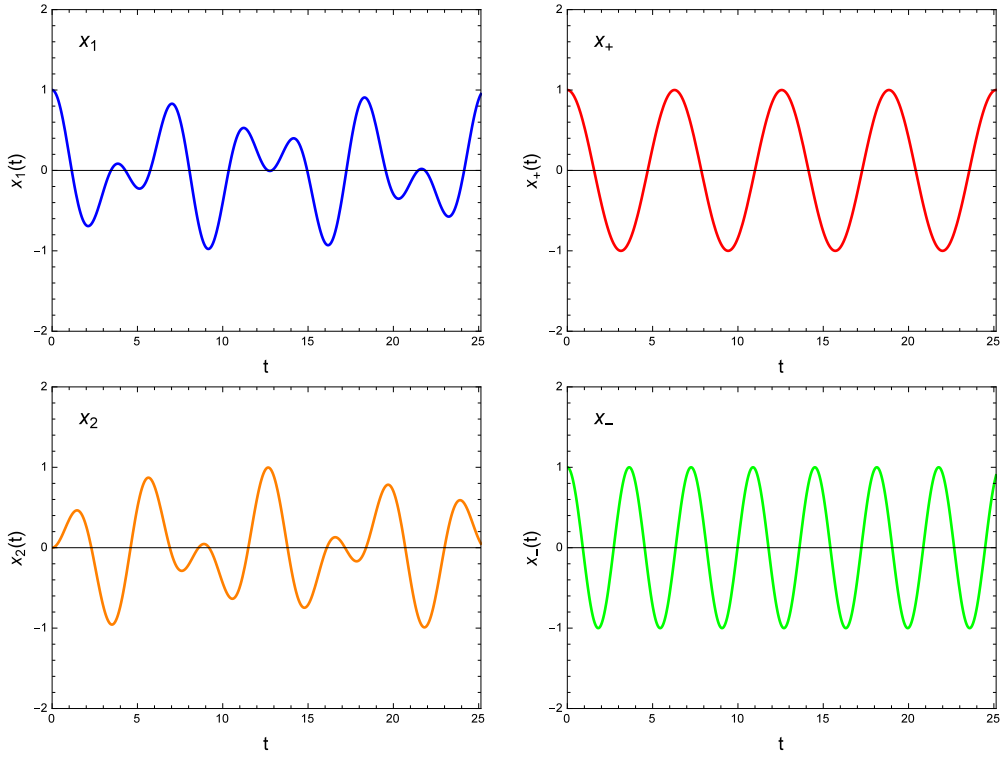


Figure 9.2: The top left and bottom left panels show the positions $x_1(t)$ and $x_2(t)$ of the blocks as functions of time, respectively. The top right and bottom right panels show the value of the variables $x_+(t) = x_1(t)+x_2(t)$ and $x_-(t) = x_1(t) - x_2(t)$ as functions of time, respectively. For all of of the results shown in all of the panels, we have taken $a = 1$ and $k/m = 1$.

Our end goal in analyzing this system of oscillators was, of course, to obtain general expressions for the coordinates $x_1(t)$ and $x_2(t)$ which describe the physical positions of the blocks. However, the coordinates $x_+(t)$ and $x_-(t)$ for which the corresponding differential equations decoupled are also are also clearly very important in terms of characterizing the dynamics of those blocks. It's therefore useful to attempt to understand their physical significance a bit better. The easiest way to do this is to consider what happens when either A_+ or A_- is set to zero and so that only one of these coordinates is non-zero.

For example, let's consider what happens if we were to set $A_- = 0$ with $A_+ \neq 0$. This implies that $x_-(t) = x_1(t) - x_2(t) = 0$ for all time, which means that the distance between the positions of the two blocks remains constant as $x_+(t)$ evolves. In other words, the "mode" of oscillation with $A_- = 0$ corresponds to the two blocks sliding back and forth in unison, as shown in the left panels of Fig. 9.3. By contrast, the if we set $A_+ = 0$ with $A_- \neq 0$, we have $x_+(t) = x_1(t) + x_2(t) = 0$ for all time, which means that the displacements of the two blocks from their equilibrium positions must always have the same magnitude but opposite sign. This "mode" of oscillation is illustrated in the right panels of Fig. 9.3. These two modes of oscillation are our first example of what are called **normal modes** for a system of coupled oscillators. Since $x_1(t)$ and $x_2(t)$ are just linear combinations of $x_+(t)$ and $x_-(t)$, we can think of any motion of the positions of the

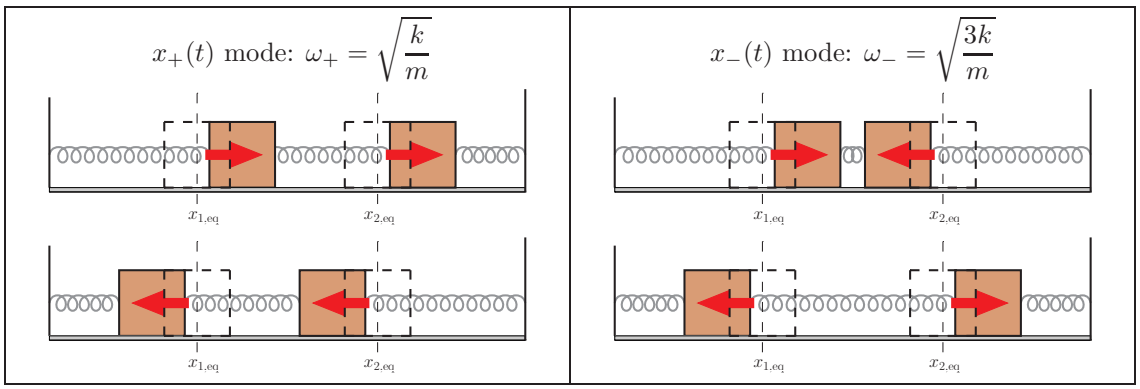


Figure 9.3: Illustration of the two normal modes of oscillation for the coupled system of two harmonic oscillators shown in Fig. 9.1. The panels on the left correspond to oscillations of the coordinate $x_+(t)$ alone; the panels on the right correspond to oscillations of the coordinate $x_-(t)$ alone. Any patter of oscillations of the physical positions $x_1(t)$ and $x_2(t)$ of the two blocks can be thought of a a superposition of contributions from these two normal modes.

blocks as being a linear combination of these normal modes. As we shall soon see, the motion of even more complicated systems of oscillators can be represented in terms of a set of normal modes which characterize the system.

9.3 Our Second Coupled System: Charged Particle in a Magnetic Field

Another example of a commonplace physical system which is described by a set of coupled differential equations is the motion of a charged particle moving in a magnetic field. This is a useful example to consider alongside the coupled-harmonic-oscillator example from Sect. 9.2 because it highlights a different technique which can be useful for solving coupled differential equations which can't be decoupled by adding and subtracting the individual equations of motion.

Let's begin by reminding ourselves how the motion of a charged particle is affected by the presence of a magnetic fields. You may recall from introductory physics that that the Lorentz force on a particle with an electric charge q and mass m moving in a magnetic field $\vec{\mathbf{B}}$ is given by

$$\vec{\mathbf{F}} \;=\; q\vec{\mathbf{v}} \times \vec{\mathbf{B}} \;, \tag{9.20}$$

where $\vec{\mathbf{v}}$ denotes the velocity vector for the particle. We'll focus here on the particular case in which the magnetic field $\vec{\mathbf{B}} = B_0\hat{\mathbf{z}}$ is constant and points in the z-direction. In this case, the force vector is

$$\vec{\mathbf{F}} \;=\; qB_0 v_y \hat{\mathbf{x}} - qB_0 v_x \hat{\mathbf{y}} \;. \tag{9.21}$$

The trajectory of the particle is governed, as usual, by Newton's Second Law. The vector form of this law is

$$\vec{\mathbf{F}} \;=\; m\vec{\mathbf{a}} \;. \tag{9.22}$$

When we write this vector equation out in components for the force in Eq. (9.21), we obtain a system of three coupled second-order differential equations describing the motion of the particle:

$$\begin{aligned} \frac{d^2x}{dt^2} \;&=\; F_x \;=\; \frac{qB_0}{m}\frac{dy}{dt} \\ \frac{d^2y}{dt^2} \;&=\; F_y \;=\; -\frac{qB_0}{m}\frac{dx}{dt} \\ \frac{d^2z}{dt^2} \;&=\; F_z \;=\; 0 \;. \end{aligned} \tag{9.23}$$

The differential equation for the z coordinate of the particle is trivial: the acceleration in the z-direction is zero, so the particle just continues traveling along in this direction with whatever initial velocity $v_{z,0}$ it started with. In other words, the z component of the velocity is a constant:

$$v_z(t) \;=\; v_{z,0} \;. \tag{9.24}$$

By contrast, the differential equations for x and y are non-trivial and coupled. However, we can at least simplify things a little bit by observing that neither of these equations involves the coordinates x or y directly, but only the time-derivatives of these coordinates. This means that we can rewrite these second-order differential equations for the coordinates x and y as first-order differential equations for the velocity components v_x and v_y:

$$\frac{dv_x}{dt} \;=\; \frac{qB_0}{m}v_y \tag{9.25}$$

$$\frac{dv_y}{dt} \;=\; -\frac{qB_0}{m}v_x \tag{9.26}$$

Can we simplify things further? Well, the trick that we used above to solve the system of equations in Eqs. (9.1) and (9.2) doesn't quite work in this case. Indeed, we find that adding and subtracting Eqs. (9.25) and (9.26) gives us

$$\frac{d}{dt}(v_x + v_y) \;=\; \frac{qB_0}{m}(v_y - v_x) \;, \qquad \frac{d}{dt}(v_x - v_y) \;=\; \frac{qB_0}{m}(v_x + v_y) \;, \tag{9.27}$$

which doesn't help in decoupling the equations because the linear combination of x_1 and x_2 which appears on the left side of each equation isn't the same one which appears on the right side. Apparently we need an alternative approach.

Fortunately, there is indeed another trick that we can use in order to solve this system of equations for $v_x(t)$ and $v_y(t)$. The idea is to combine the two first-order differential equations in Eq. (9.26) and Eq. (9.26) together into a single second-order differential equation which involves only one of the two variables. After we've isolated that variable, we can use our original first-order equations to obtain the expression for the other variable. We begin by taking the time derivative of both sides of Eq. (9.25):

$$\frac{d^2v_x}{dt^2} \;=\; \frac{qB_0}{m}\frac{dv_y}{dx} \;. \tag{9.28}$$

Now we can substitute in for dv_y/dx using Eq. (9.26). The resulting equation

$$\frac{d^2v_x}{dt^2} \;=\; -\frac{q^2B_0^2}{m^2}v_x \tag{9.29}$$

involves v_x alone, so we can solve it directly. We don't have to work terribly hard to find the solution either. This equation has the same form as the the simple-harmonic-oscillator equation with an oscillator frequency

$$\omega_c \;\equiv\; \frac{qB_0}{m} \;. \tag{9.30}$$

We therefore know that the general solution for $v_x(t)$ can be written in the form

$$v_x(t) \;=\; V\cos(\omega_c t + \phi) \;, \tag{9.31}$$

where V and phase ϕ are constant parameters to be determined by the boundary conditions applicable to the problem. It's also worth that the angular frequency ω_c is just the angular version of the **cyclotron frequency** f_c:

$$f_c \;\equiv\; \frac{\omega_c}{2\pi} \;=\; \frac{qB_0}{2\pi m} \;. \tag{9.32}$$

Now that we have our solution for $v_x(t)$, we can obtain a solution for $v_y(t)$ by substituting Eq. (9.31) back into Eq. (9.25):

$$\frac{dv_x}{dx} \;=\; -V\omega_c\sin(\omega_c t + \phi) \;=\; \frac{qB_0}{m}v_y(t) \;=\; \omega_c v_y(t) \;. \tag{9.33}$$

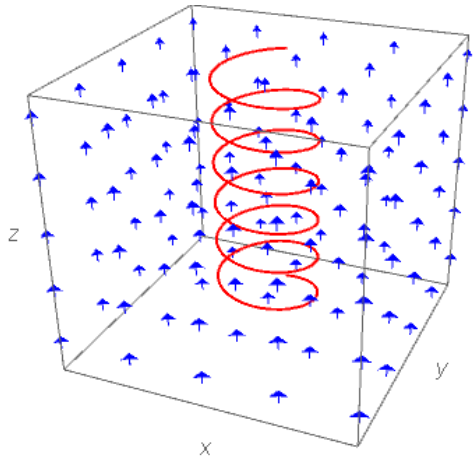


Figure 9.4: The trajectory (red curve) of a charged particle moving in a constant magnetic field pointing in the z direction. The magnetic field itself is shown in blue. The particle begins its motion at at bottom of the figure at $t = 0$ and travels upward, coiling around the magnetic field lines as it goes.

Collecting this result for $v_y(t)$ together with the results for $v_z(t)$ and $v_x(t)$ in Eqs. (9.24) and (9.24), we find that the components of the particle's velocity vector $\vec{\mathbf{v}}$ are

$$\begin{aligned} v_x(t) &= V\cos(\omega_c t + \phi) \\ v_y(t) &= -V\sin(\omega_c t + \phi) \\ v_z(t) &= v_{z,0} \end{aligned} \tag{9.34}$$

To obtain the corresponding components of the particle's position vector $\vec{\mathbf{x}}$ are simply the integrals of these expressions with respect to time. Evaluating these integrals and choosing our time coordinate so that $t_0 = 0$, we find that the trajectory of a charged particle as it moves through the constant magnetic field in the z direction is

$$\begin{aligned} x(t) &= \frac{V}{\omega_c}\sin(\omega_c t + \phi) + x_0 \\ y(t) &= \frac{V}{\omega_c}\cos(\omega_c t + \phi) + y_0 \\ z(t) &= v_{z,0}t + z_0 \ , \end{aligned} \tag{9.35}$$

where x_0, y_0, and z_0 are the initial values of $x(t)$, $y(t)$, and $z(t)$ at $t = 0$. Note that these three components of the initial position vector of the particle, along with V, $v_{z,0}$, and ϕ, amount to six free parameters. Indeed, this is the appropriate number of free parameters we need in order to characterize the general solution to the three second-order differential equations with which we began in Eq. (9.23).

The trajectory described by the three expressions in Eq. (9.35) is illustrated in Fig. 9.4. The charged particle begins its motion at at bottom of the figure at $t = 0$ and travels upward, coiling around the magnetic field lines as it goes.

9.4 Vector Spaces

In each of the two previous sections, we focused on a different physical system whose dynamics were described by a set of coupled differential equations. We were able to decouple each of these sets of equations by using a different trick. We solved the equations governing the coupled-oscillator system in Sect. 9.2 by adding and subtracting the equations of motion for the individual oscillators. We solved the coupled equations of motion for a charged particle moving in a magnetic field in Sect. 9.3 by substituting one of these first-order

equations into the other to obtain a second-order equation. However, each of these tricks was useful only because the corresponding equations of motion happened to take a particular form.

Fortunately, there is a more general approach for decoupling and solving systems of coupled linear differential equations. This approach involves recasting these equations in terms of vectors and matrices and using the machinery of linear algebra in order to decouple the system and obtain a solution. Therefore, we begin our discussion of this more general approach to coupled systems of linear differential equations with a brief review of some of the fundamental concepts from linear algebra which we will need, beginning with the concept of a vector space.

A **vector space** comprises of a set of objects called vectors, which we'll denote $\{|\alpha\rangle, |\beta\rangle, |\gamma\rangle, \ldots\}$, along with another set of objects[2] called scalars or numbers, which we'll denote $\{a, b, c, \ldots\}$. In any vector space, two fundamental mathematical operations are always defined for how these objects can be combined together. First, there must be some sense in which we can add a pair of vectors $|\alpha\rangle$ and $|\beta\rangle$ together to get another vector $|\gamma\rangle = |\alpha\rangle + |\beta\rangle$. Second, there must be some sense in which we can multiply a scalar a by a vector $|\alpha\rangle$ and get another vector. In order for the set of objects on which these two operations act to be considered a vector space, these operation must satisfy the following criteria.

- The vectors in the vector space must satisfy a property called **closure**.[3] What this means is that whenever we add any two vectors $|\alpha\rangle$ and $|\beta\rangle$ in the space together, their sum $|\gamma\rangle = \alpha + \beta$ must also be a vector in the space. Likewise, whenever we multiply a vector by a scalar, the product $|\beta\rangle = a|\alpha\rangle$ must also be a vector in the space.

- Vector addition must be commutative in the sense that

$$|\alpha\rangle + |\beta\rangle \;=\; |\beta\rangle + |\alpha\rangle \;. \tag{9.36}$$

- Vector addition must be associative in the sense that

$$|\alpha\rangle + \big(|\beta\rangle + |\gamma\rangle\big) \;=\; \big(|\alpha\rangle + |\beta\rangle\big) + |\gamma\rangle \;. \tag{9.37}$$

- Scalar multiplication must be associative in the sense that

$$a\big(b|\alpha\rangle\big) \;=\; (ab)|\alpha\rangle \;. \tag{9.38}$$

- The multiplication of vectors by scalars must be distributive in two different senses:

$$(a+b)|\alpha\rangle \;=\; a|\alpha\rangle + b|\alpha\rangle \tag{9.39}$$

$$a\big(|\alpha\rangle + \beta\rangle\big) \;=\; a|\alpha\rangle + a\beta\rangle \;. \tag{9.40}$$

- A vector space must include a special vector which is called the **additive identity element**. This vector, which we will denote $|0\rangle$, has the property that the sum of this vector with any vector $|\alpha\rangle$ in the space is just $|\alpha\rangle$ itself:

$$|0\rangle + |\alpha\rangle \;=\; \alpha \qquad \text{for all } |\alpha\rangle \;. \tag{9.41}$$

- Likewise, a vector space must include a special scalar which is called the **multiplicative identity element**. This scalar, which we will denote 1, has the property that the product of this scalar and any vector $|\alpha\rangle$ in the space is just $|\alpha\rangle$ itself:

$$1|\alpha\rangle \;=\; |\alpha\rangle \qquad \text{for all } |\alpha\rangle \;. \tag{9.42}$$

- Finally, every vector in a vector space must have an **additive inverse**. The additive inverse $|-\alpha\rangle$ of a vector is another vector in the space[4] for which

$$|\alpha\rangle + |-\alpha\rangle \;=\; |0\rangle \;. \tag{9.43}$$

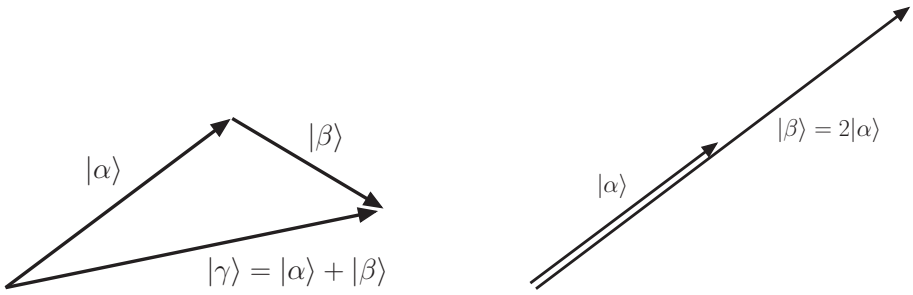


Figure 9.5: An illustration of how the two basic operations on the elements of a vector space are carried out within the familiar vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$. The left panel shows two vectors $|\alpha\rangle$ and $|\beta\rangle$ being added together tip to tail; the right panel shows the multiplication of a vector $|\alpha\rangle$ by the scalar $a = 2$.

The properties in this list are admittedly defined in a very abstract way. However, these properties also conform quite nicely to our intuitive notion of what a vector is. Indeed, if think about it for a moment, you'll realize that these requirements are satisfied by the familiar kinds of vectors we commonly use to describe quantities like positions, velocities, or forces in two or three spatial dimensions. Indeed, vectors of this sort are members of the vector spaces which mathematicians call $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. In these cases, you know that one adds two vectors by placing them head to tail and that the multiplication of a vector by a number simply scales the vector by the corresponding amount. These two operations are illustrated in Fig. 9.5. Each one clearly produces another vector in the space.

However, there are many other collections of objects that qualify as vector spaces. Many of these objects do not conform to our intuitive idea of what a vector should look like, but they nevertheless satisfy the criteria for a vector space. A few examples are given below.

- The null vector $|0\rangle$ itself, along with any set of scalars. This is a strange and somewhat trivial example, but it indeed satisfies all of the above criteria for a vector space. The closure requirement is satisfied because $|0\rangle + |0\rangle = |0\rangle$ and $a|0\rangle = |0\rangle$ for any scalar a. Moreover, $|0\rangle$ is its own additive inverse and functions as the additive identity element.

- The set of functions $f(x)$ for which $f(0) = 0$ (our vectors) along with the set of real numbers (our scalars). A space like this, in which the vectors are functions of some independent variable x, is called a **function space**. In such a space, the vector-addition operation simply corresponds to adding two functions $f(x)$ and $g(x)$ together in the standard manner to produce a new function $h(x) = f(x)+g(x)$. This particular function space is closed under vector addition because $h(0) = f(0) + g(0) = 0 + 0 = 0$ for any vectors $f(x)$ and $g(x)$ in the space. It is also closed under the multiplication of vectors by scalars. Indeed, $h(x) = af(x)$ is also in the space because $h(0) = af(0) = 0$ for any scalar a.

- The set of functions of the form $f(x) = f_0 + f_1 x$, where f_0 and f_1 are arbitrary constants. The sum of two such functions is
$$h(x) \;=\; f(x) + g(x) \;=\; (f_0 + g_0) + (f_1 + g_1)x \;, \tag{9.44}$$
so $h(x)$ is also a vector in the space, with $h_0 = f_0 + g_0$ and $h_1 = f_1 + g_1$. Closure under scalar multiplication is also easy to prove.

Linear Combinations of Vectors:

[2]Technically, the set of scalars associated with a vector space cannot be completely arbitrary. In particular, they must collectively constitute what mathematicians call a field. This technicality won't concern us here, since the collections of scalars we'll be using for all of the vector fields we'll be dealing with in this chapter, including both the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$, are fields in this sense.

[3]The corresponding adjective is "closed."

[4]This "other vector" need not necessarily be different from $|\alpha\rangle$ itself. In other words, it is perfectly legitimate for a vector to be its own additive inverse. For example, the additive identity element $|0\rangle$ in any vector space is always its own inverse.

A **linear combination** of vectors is simply the sum of N vectors $|\alpha_n\rangle$, where $n = 1, 2, \ldots, N$, each of which is multiplied by a scalar a_n. Since vector addition and the multiplication of vectors by scalars are operations which are guaranteed to return another vector in the space, we know that any linear combination of vectors

$$|\beta\rangle \;=\; \sum_{n=1}^{N} a_n |\alpha_n\rangle \tag{9.45}$$

must be another vector in the space.

Linear Independence:

In Chapter 5, we introduced the concept of linear independence for functions. In particular, we defined a pair of functions $y_1(x)$ and $y_2(x)$ to be linearly independent if the only solution to the equation

$$a y_1(x) + b y_2(x) \;=\; 0 \tag{9.46}$$

valid for all x, where a and b are arbitrary coefficients, is $a = b = 0$. The definition of linear independence for vectors is quite similar. In particular, a set of N vectors $|\alpha_n\rangle$, where $n = 1, 2, \ldots, N$, are said to be linearly independent if the only solution to the equation

$$\sum_{n=1}^{N} a_n |\alpha_n\rangle \;=\; |0\rangle \;, \tag{9.47}$$

where the a_n are arbitrary scalars, is $a_n = 0$ for all n.

If our vector space happens to be a function space, this condition corresponds to a generalization of the linear-independence criterion for two functions $y_1(x)$ and $y_2(x)$ stated above. In particular, a set of N functions $y_n(x)$, where $n = 1, 2, \ldots, N$, are said to be linearly independent if the only solution to the equation

$$a_1 y_1(x) + a_2 y_2(x) + \ldots + a_N y_N(x) \;=\; 0 \tag{9.48}$$

valid for all x is $a_n = 0$ for all n.[5]

Spanning the Space:

A set of N vectors $|\alpha_n\rangle$, where $n = \{1, 2, \ldots, N\}$ in a vector space is said to **span the space** if every vector $|\beta\rangle$ in the space can be written as a linear combination of the these vectors for some set of scalar coefficients a_n:

$$|\beta\rangle \;=\; \sum_{n=1}^{N} a_n |\alpha\rangle \;\;=\; a_1|\alpha_1\rangle + a_2|\alpha_2\rangle + \ldots a_N|\alpha_N\rangle \;. \tag{9.49}$$

Basis:

A set of vectors in a vector space which are both linearly independent and span the space are called a **basis**. The concept of a basis is sufficiently important in the analysis of vector spaces that we will use a special notation $|e_n\rangle$ to denote a set of vectors which form a basis. Since the $|e_n\rangle$ span the space, we know that any vector $|\alpha\rangle$ can be written as a linear combination of these basis vectors:

$$|\alpha\rangle \;=\; \sum_{n=1}^{N} a_n |e_n\rangle \;=\; a_1|e_1\rangle + a_2|e_2\rangle + \ldots a_N|e_N\rangle \;. \tag{9.50}$$

However, because the $|e_n\rangle$ are also linearly independent, we know that there is no redundancy among them. This means that the set of a_n which characterize a given vector $|\alpha\rangle$ in a particular basis is unique. It also means that a basis always contains precisely the minimal number of vectors necessary to express any vector

[5] If you're wondering whether there is a generalization of the Wronksian for a set of N functions $y_n(x)$, there is indeed. In particular, the Wronksian $W(x)$ is the determinant of the $N \times N$ matrix whose elements are given by $W_{ij} = y_j^{(i-1)}(x)$, where $y_j^{(i-1)}(x)$ denotes the $(i-1)$th derivative of the function $y_j(x)$. You can read more about determinants in Sect. 9.6. You can also verify that $W(x)$ reduces to the familiar form from Chapter 5 in the case where $N = 2$.

in the space through linear combinations of the form given in Eq. (9.50). This minimal number of vectors is referred to as the **dimension** of the vector space.

Since a given vector $|\alpha\rangle$ is characterized by a unique set of N coefficients a_n in any particular basis, we can unambiguously describe $|\alpha\rangle$ by providing a list of these N coefficients — provided, of course, that we are clear about what basis we're working in. The coefficients in this list are called the **components** of $|\alpha\rangle$ in that basis. These components are sometimes expresses as an N-tuple of the form $(a_1, a_2, \ldots, a_N)$. Alternatively, it is often convenient to express them as an array of the form

$$|\alpha\rangle \;\leftrightarrow\; \mathbf{a} \;=\; \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} . \tag{9.51}$$

This notation for vectors is likely to be more familiar. The double arrow is used here to indicate that the component-wise expression $\mathbf{a}$ for the vector $|\alpha\rangle$ is only defined with respect to a particular basis. In what follows, we will use $|\alpha\rangle$ to refer to a vector in the most general, abstract sense and $\mathbf{a}$ to refer to a vector in a particular basis.[6]

One of the advantages of working in a particular basis is that it makes performing the fundamental operations on vectors and scalars within a vector spaces operationally much simpler. Vector addition in a particular basis just corresponds to adding the individual components of the two vectors together:

$$|\alpha\rangle + \beta\rangle \;=\; \mathbf{a} + \mathbf{b} \;=\; \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_N + b_N \end{pmatrix} . \tag{9.52}$$

Likewise, multiplying a vector by a scalar corresponds to multiplying each component by that same scalar:

$$a|\beta\rangle \;\leftrightarrow\; a\mathbf{b} \;=\; \begin{pmatrix} ab_1 \\ ab_2 \\ \vdots \\ ab_N \end{pmatrix} . \tag{9.53}$$

It's important to keep in mind that the description of a vector $|\alpha\rangle$ in terms of components — either as an N-tuple or as a column of numbers — only makes sense if we know the basis in which we're working. For example, when we write down vectors in $\mathbb{R}^3$ in terms of their components without explicitly declaring what the basis is, the tacit assumption is that we are working in a basis where the three basis vectors $|e_1\rangle$, $|e_2\rangle$, and $|e_3\rangle$ correspond to the unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$, respectively.

Example: Basis for a Function Space

As an example of how one goes about finding a basis for an abstract vector space, let's find a basis for the function space which consists of all functions of the form $f(x) = a_1 x + a_2 x^2$. It's not difficult to verify that this set of functions is closed under both vector addition and scalar multiplication and therefore indeed constitutes a vector space.

In order to construct a basis for this vector space, we need to find a set of linearly-independent functions which span the space. One good guess might be the functions $y_1(x) = x$ and $y_2(x) = x^2$. Indeed, these functions are linearly independent because the Wronksian for $y_1(x)$ and $y_2(x)$ is

$$W(x) \;=\; y_1 \frac{dy_2}{dx} - y_2 \frac{dy_1}{dx} \;=\; x \cdot 2x - x^2 \cdot 1 \;=\; x^2 \,, \tag{9.54}$$

[6] In lecture, I will use the notation $\vec{a}$ rather than $\mathbf{a}$ to refer to the description of a vector in a particular basis.

which is not identically zero for all x. Moreover, they span the space because any function $f(x)$ in our function space can be written as a linear combination of these two functions:

$$f(x) \;=\; a_1 x + a_2 x^2 \;=\; a_1 y_1(x) + a_2 y_2(x) \;. \tag{9.55}$$

Thus, the functions x and x^2 constitute a basis for this function space. Since this basis contains two basis vectors the dimension of our function space is 2.

9.5 The Inner Product and Inner Product Spaces

The fundamental operations which define a vector space — vector addition and the multiplication of a vector by a scalar — are recipes for combining two vectors to get a vector and combining two scalars to get a vector. These two operations are the only ones that need to be defined in order for a collection of scalars and vectors to constitutes vector space. However, in many vector spaces — including the familiar examples $\mathbb{R}^2$ and $\mathbb{R}^3$ — it is possible to define additional operations, such as a procedure for combining any two vectors to get a scalar. Such an operation is called the **inner product** of those vectors, and a vector space for which an inner product is defined is called an **inner-product space**.

The inner product of two vectors $|\alpha\rangle$ and $|\beta\rangle$ is usually written $\langle\alpha|\beta\rangle$, and it must satisfy the following criteria.

- The inner product of any vector $|\alpha\rangle$ with any other vector $|\beta\rangle$ in the space must be equal to the complex conjugate of the inner product of $|\beta\rangle$ with $|\alpha\rangle$:

$$\langle\alpha|\beta\rangle \;=\; \langle\beta|\alpha\rangle^* \;. \tag{9.56}$$

- The inner product must be distributive in the sense that

$$\langle\gamma|\big(|\alpha\rangle + |\beta\rangle\big) \;=\; \langle\gamma|\alpha\rangle + \langle\gamma|\beta\rangle \;. \tag{9.57}$$

- The inner product must have the property that

$$\langle\beta|\big(c|\alpha\rangle\big) \;=\; c\langle\beta|\alpha\rangle. \tag{9.58}$$

- The inner product of a vector with itself must satisfy the condition

$$\langle\alpha|\alpha\rangle \;\geq\; 0 \;, \tag{9.59}$$

  with equality only in the case in which $|\alpha\rangle = |0\rangle$ is the null vector.

Norm:

The norm of a vector $|\alpha\rangle$ vector, which is typically denoted $||\alpha||$, is the square root of the inner product of the vector with itself:

$$||\alpha|| \;=\; \sqrt{\langle\alpha|\alpha\rangle} \;. \tag{9.60}$$

The norm of a vector has a direct, physical interpretation for vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$: it corresponds to the length or magnitude of the vector. A vector which has a norm $||\alpha|| = 1$ is said to be **normalized**.

Orthogonality:

Two vectors are said to be **orthogonal** if their inner product vanishes — *i.e.*, if

$$\langle\alpha|\beta\rangle \;=\; 0 \;. \tag{9.61}$$

For a pair of vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, this condition corresponds to those vectors being perpendicular to one another.

An **orthonormal basis** is a set of basis vectors $|e_n\rangle$ which are all normalized and mutually orthogonal. In other words, we must have $||e_n|| = 1$ for all n and $\langle e_m|e_n\rangle = 0$ for $m \neq n$. These conditions can be written in the compact form

$$\langle e_m|e_n\rangle \;=\; \delta_{mn} \tag{9.62}$$

where we have introduced the symbol known as the **Kronecker delta**,[7] which is defined such that

$$\delta_{mn} \;=\; \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} . \tag{9.63}$$

In $\mathbb{R}^2$ or $\mathbb{R}^3$, this condition corresponds to the basis vectors being not only perpendicular to one another, but also each having length 1. For example, the unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ constitute an orthonormal basis for $\mathbb{R}^3$.

The useful thing about working in an orthonormal basis is that it makes computing the inner product of two vectors much simpler. First of all, we note that in *any* basis, I can always expand a vector $|\alpha\rangle$ is terms of the basis vectors, as in Eq. (9.50). Moreover, since the inner product of any two vectors must satisfy Eq. (9.56), the corresponding expansion for $\langle\alpha|$ in terms of the basis vectors must be

$$\langle\alpha| \;=\; \sum_{n=1}^{N} \langle e_n|a_n^* \; . \tag{9.64}$$

Thus, when we are working in an orthonormal basis, the inner product of two vectors reduces to

$$\begin{aligned} \langle\alpha|\beta\rangle \;&=\; \left(\sum_{m=1}^{N} \langle e_n|a_m^*\right)\left(\sum_{n=1}^{N} |e_n|a_n^*\rangle\right) \\ &=\; \sum_{m=1}^{N}\sum_{n=1}^{N} a_m^* b_n \langle e_m|e_n\rangle \\ &=\; \sum_{m=1}^{N}\sum_{n=1}^{N} a_m^* b_n \delta_{mn} \\ &=\; \sum_{n=1}^{N} a_n^* b_n \;=\; a_1^* b_1 + a_2^* b_2 + \ldots a_N^* b_N \; . \end{aligned} \tag{9.65}$$

Working in an orthonormal basis makes finding the components of a vector much simpler as well. In particular, we can obtain the component a_n of a vector $|\alpha\rangle$ by taking the inner product of $|\alpha\rangle$ with the corresponding basis vector $|e_n\rangle$:

$$\langle e_n|\alpha\rangle \;=\; \langle e_n|\left(\sum_{m=1}^{N} a_m|e_m\rangle\right) \;=\; \sum_{m=1}^{N} a_m\langle e_n|e_m\rangle \;=\; \sum_{m=1}^{N} a_m\delta_{mn} \;=\; a_n \; .$$

These considerations in an of themselves make it clear how advantageous it is to work in an orthonormal basis when dealing with vectors. Therefore, in the rest of this chapter, we will assume that we are working in an orthonormal basis, unless we state otherwise. Thus, if a vector is expressed as an n-tuple or array of numbers from this point forward, you should assume that the components in that array have been evaluated with respect to an orthonormal basis unless you are explicitly told otherwise.

So how to we go about finding an orthonormal basis for the vectors in our vector space if we aren't explicitly given one? The answer is that there is a procedure which can be used to obtain an orthonormal basis from any arbitrary set of vectors which are linearly independent and span the space. This procedure

[7] Take care not to confuse the Kronecker delta with the Dirac Delta function, as they represent very different things. The Kronecker delta δ_{mn} represents a condition on a pair of integers m and n. By contrast, the Dirac delta function $\delta(x)$ is a function of the continuous variable x.

is called the **Gram-Schmidt procedure**. To see how it works, it's best to begin with an example from $\mathbb{R}^3$. Consider the set of vectors

$$\mathbf{a}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{a}_2 = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}, \qquad \mathbf{a}_3 = \begin{pmatrix} 1 \\ 1 \\ 7 \end{pmatrix} \tag{9.66}$$

whose components are given with respect to the usual basis in which the basis vectors are $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$. In order to create an orthonormal basis out of these vectors, we begin by simply choosing one of them — let's say $\mathbf{a}_1$ — and normalizing it to obtain the first basis vector e_1 in our new basis:

$$\mathrm{e}_1 = \frac{\mathbf{a}_1}{||\mathbf{a}_1||} = \frac{1}{\sqrt{2^2+1^2+0^2}} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}. \tag{9.67}$$

Now we need to find a set of additional vectors which are orthogonal to $\mathbf{e}_1$. The easiest way of doing this is to choose another vector — say, $\mathbf{a}_2$ — from our original set and subtract off the piece of this vector which is in the direction along $\mathbf{e}_1$. This piece is just $(\mathbf{e}_1 \cdot \mathbf{a}_2)\mathbf{e}_1$, so we have

$$\mathbf{a}_2 - (\mathbf{e}_1 \cdot \mathbf{a}_2)\mathbf{e}_1 = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{5}}\big(2 \cdot 2 + 1 \cdot 3 + 0 \cdot 0\big) = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \frac{4}{5} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}. \tag{9.68}$$

All we need to do in order to obtain our second basis vector $\mathbf{a}_2$ is normalize:

$$\mathbf{e}_2 = \frac{\mathbf{a}_2 - (\mathbf{e}_1 \cdot \mathbf{a}_2)\mathbf{e}_1}{||\mathbf{a}_2 - (\mathbf{e}_1 \cdot \mathbf{a}_2)\mathbf{e}_1||} = \frac{1}{\sqrt{(-1)^2+2^2+0^2}} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}. \tag{9.69}$$

Our third vector must be orthogonal to both $\mathbf{e}_1$ and $\mathbf{e}_2$. Once again, the easiest way to do this is to take our remaining vector $\mathbf{a}_3$ and subtract off the pieces which are in the directions along both $\mathbf{e}_1$ and $\mathbf{e}_2$:

$$\mathbf{a}_3 - (\mathbf{e}_1 \cdot \mathbf{a}_3)\mathbf{e}_1 - (\mathbf{e}_1 \cdot \mathbf{a}_3)\mathbf{e}_1 = \begin{pmatrix} 1 \\ 1 \\ 7 \end{pmatrix} - \frac{3}{5} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{5} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix}. \tag{9.70}$$

Normalizing the result gives us the third and final basis vector $\mathbf{e}_3$ in our orthonormal basis:

$$\mathbf{e}_3 = \frac{\mathbf{a}_3 - (\mathbf{e}_1 \cdot \mathbf{a}_3)\mathbf{e}_1 - (\mathbf{e}_2 \cdot \mathbf{a}_3)\mathbf{e}_2}{||\mathbf{a}_3 - (\mathbf{e}_1 \cdot \mathbf{a}_3)\mathbf{e}_1 - (\mathbf{e}_2 \cdot \mathbf{a}_3)\mathbf{e}_2||} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{9.71}$$

In summary, we started with the three vectors in Eq. (9.66) and derived an orthonormal basis which comprises the vectors

$$\mathbf{a}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{a}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \qquad \mathbf{a}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{9.72}$$

You can verify for yourself that indeed these vectors indeed constitute an orthonormal basis.

The above calculation is an example of the Gram-Schmidt procedure in action. Basically, this recursive procedure consists of picking a vector from our original set, subtracting off the pieces of this vector which are parallel to any of the new basis vectors we've already derived, and normalizing the result. In general, if we begin with a set of N linearly-independent vectors $|\alpha_n\rangle$ with $n = 1, 2, \ldots, N$ which span the space, the

Gram-Schmidt procedure yields the following set of basis vectors:

$$\begin{aligned}
|e_1\rangle &= \frac{|a_1\rangle}{||\alpha_1||} \\
|e_2\rangle &= \frac{|a_2\rangle - |e_1\rangle\langle e_1|a_2\rangle}{\sqrt{||\alpha_2||^2 - |\langle e_1|a_2\rangle|^2}} \\
|e_3\rangle &= \frac{|a_3\rangle - |e_1\rangle\langle e_1|a_3\rangle - |\mathbf{e}_2\rangle\langle e_2|a_3\rangle}{\sqrt{||\alpha_3||^2 - |\langle e_1|a_3\rangle|^2 - |\langle e_1|a_3\rangle|^2}} \\
&\vdots \\
|e_3\rangle &= \frac{|a_N\rangle - \sum_{n=1}^{N-1} |e_n\rangle\langle e_n|a_N\rangle}{\sqrt{||\alpha_3||^2 - \sum_{n=1}^{N-1} |\langle e_n|a_N\rangle|^2}}
\end{aligned} \tag{9.73}$$

Computing the orthonormal basis of vectors in Eq. (9.72) in the above example might seem somewhat silly or unnecessary. After all, the vectors Eq. (9.66) were already expressed in terms of an orthonormal basis — namely, the standard basis for vectors in $\mathbb{R}^3$ in which the basis vectors are $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$. However, this fact actually highlights an important aspect of the Gram-Schmidt procedure. This procedure is guaranteed to give you *an* orthonormal basis, but not necessarily any *particular* orthonormal basis.

On a final note, while the definition of the inner product presented in this section may once again seem a trifle abstract, it also conforms to our notion of what a vector product should be. In particular, for vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, the inner product corresponds to our familiar notion of the dot product.[8] For example, the dot product

$$\mathbf{a} \cdot \mathbf{b} \;=\; a_x b_x + a_y b_y + a_z b_z \;. \tag{9.74}$$

of two vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^3$ satisfies all of the criteria for an inner product because the components of a vector in $\mathbb{R}^3$ are real by definition. However, the definition of an inner product can be extended to a far broader class of operations in different kinds of vector spaces. For example, in vector spaces where the vectors are functions $f(x)$ defined on some interval from $a < x < b$, the definite integral

$$\langle f|g\rangle \;\equiv\; \int_a^b f^*(x) g(x) dx \tag{9.75}$$

satisfies all requirements for an inner product as well.

9.6 Matrices

Matrices provide a compact way of representing transformations on the vectors in vector spaces — including transformations from one basis to another. Structurally, a matrix is simply a rectangular array of numbers, called the **entries** of the matrix. These entries may in general be either real or complex. A matrix consisting of m rows and n columns is referred to as an $m \times n$ matrix. For example, the matrix

$$\mathbf{A} \;=\; \begin{pmatrix} 2 & -1 & 3i \\ 2+2i & 4 & -3 \end{pmatrix} \tag{9.76}$$

is a 2×3 matrix. A matrix which is has the same number of rows and columns is sometimes referred to as as **square** matrix.

Complex Conjugate:

The complex conjugate $\mathbf{A}^*$ of a matrix $\mathbf{A}$ is the matrix whose individual elements are the complex conjugates of the corresponding elements of $\mathbf{A}$. In other words,

$$[\mathbf{A}^*]_{ij} \;=\; A^*_{ij} \;. \tag{9.77}$$

[8] If you're wondering about the cross product — the other familiar vector product from $\mathbb{R}^3$, which combines two vectors together to yield something like a vector — this is an example of a different sort of product called an outer product with very different properties. We won't be dealing with outer products in this chapter, but they are an interesting subject in their own right.

For example, the complex conjugate of the 2 matrix

$$\mathbf{A} \;=\; \begin{pmatrix} 1 & 2i & 2 \\ 3-2i & -2 & -3i \end{pmatrix} \tag{9.78}$$

is the matrix

$$\mathbf{A}^* \;=\; \begin{pmatrix} 1 & -2i & 2 \\ 3+2i & -2 & 3i \end{pmatrix} \;. \tag{9.79}$$

The **transpose** of a matrix, usually denoted $\mathbf{A}^T$, is the matrix with the rows and columns of the original matrix A exchanged. In other words, the transpose of $\mathbf{A}$ is the matrix whose elements are given by

$$\left[\mathbf{A}^T\right]_{ij} \;=\; \mathbf{A}_{ji} \;. \tag{9.80}$$

This means that the transpose of an $m \times n$ matrix is an $n \times m$ matrix.

Hermitian Conjugate:

The **Hermitian conjugate** (sometimes also called the "conjugate transpose" or "Hermitian adjoint") of a matrix, usually denoted $A^\dagger$, is the transpose of the complex conjugate of that matrix. The elements of the Hermitian conjugate are given by

$$\left[\mathbf{A}^\dagger\right]_{ij} \;=\; A^*_{ji} \;. \tag{9.81}$$

Once again, as with the transpose, the Hermetian conjugate of an $m \times n$ matrix is an $m \times n$ matrix. This means that it is possible for a square matrix to be equal to its own Hermitian conjugate — *i.e.*, for the matrix to satisfy the criterion

$$A^\dagger \;=\; \mathbf{A} \quad \text{(Hermitian)} \;. \tag{9.82}$$

. A matrix for which satisfies this criterion is called a **Hermitian** matrix. Likewise, a matrix which satisfies the criterion

$$\mathbf{A}^\dagger \;=\; -\,\mathbf{A} \quad \text{(anti-Hermitian)} \tag{9.83}$$

is called an **anti-Hermitian**
indexanti-Hermitian matrix (or sometimes "skew Hermitian") matrix.

Trace:

The **trace** of a matrix is only defined for square matrices — *i.e.*, matrices which have the same number of rows and columns. The trace of a matrix is the sum of its diagonal elements:

$$\mathrm{Tr}(\mathbf{A}) \;=\; \sum_i A_{ii} \;. \tag{9.84}$$

For example, the trace of the matrix

$$\mathbf{A} \;=\; \begin{pmatrix} 2 & -1 & 4 \\ 3 & 0 & -6 \\ -1 & 5 & -3 \end{pmatrix} \tag{9.85}$$

is simply

$$\mathrm{Tr}(\mathbf{A}) \;=\; 2+0-3 \;=\; -\,1 \;. \tag{9.86}$$

Determinant:

The **determinant** of a matrix, like the trace, is only defined for matrices which have the same number of rows and columns. The determinant is a scalar (*i.e.*, a number), usually denoted $\det(\mathbf{A})$. The rules for computing determinants are as follows. First of all, the determinant of a 1×1 matrix (*i.e.*, a number) is equal to the sole element A_{11} of that matrix (*i.e.*, the number itself). For any $n \times n$ matrix with $n > 1$, there is a general procedure for computing the determinant. This procedure is called the Laplace expansion in terms of minors, and it proceeds via the following steps:

1. Choose one row of the matrix **A**.
2. For each element A_{ij} in that row, write down the sum-matrix obtained by crossing out the ith row and the jth column of **A**. The determinant M_{ij} of this sub-matrix is called the **minor** of A_{ij}.
3. The determinant of the matrix **A** is given by

$$\det(\mathbf{A}) \;=\; \sum_j (-1)^{i+j} A_{ij} M_{ij} \;. \tag{9.87}$$

Often, for sake of convenience, the factor $(-1)^{i+j}$ and the minor M_{ij} are often combined together into what is called the **cofactor** C_{ij} of the element A_{ij}:

$$C_{ij} \;\equiv\; (-1)^{i+j} M_{ij} \;. \tag{9.88}$$

Thus, the determinant of **A** can be written in the form

$$\det(\mathbf{A}) \;=\; \sum_j A_{ij} C_{ij} \;. \tag{9.89}$$

As an example, let's begin by computing the determinant of a general 2×2 matrix of the form

$$\mathbf{A} \;=\; \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \;. \tag{9.90}$$

When a matrix expressed as array of elements, the determinant is usually denoted by replacing the parentheses that typically enclose the array with a pair of vertical lines:[9]

$$\det(\mathbf{A}) \;=\; \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \;. \tag{9.91}$$

The first step in the Laplace expansion is to choose a row of the matrix, so we'll choose the first row. Next, we need to compute the respective minors M_{11} and M_{12} of the elements A_{11} and A_{12} in this row. In this case, however, this because the "sub-matrices" obtained by crossing out the ith row and $j$the column of the matrix for our chosen A_{ij} are just the individual elements A_{22} and A_{21}:

$$\begin{aligned} i=1,\; j=1 \;:\; & \begin{pmatrix} \cancel{A_{11}} & \cancel{A_{12}} \\ A_{12} & A_{22} \end{pmatrix} \quad\longrightarrow\quad A_{22} \\ i=1,\; j=2 \;:\; & \begin{pmatrix} \cancel{A_{11}} & \cancel{A_{12}} \\ A_{12} & A_{22} \end{pmatrix} \quad\longrightarrow\quad A_{21} \;, \end{aligned} \tag{9.92}$$

The determinant of a 1×1 matrix is just the the sole element of that matrix, as discussed above, so the minors are just $M_{11} = A_{22}$ and $M_{21} = A_{21}$ and athe corresponding cofactors are

$$\begin{aligned} C_{11} \;&=\; (-1)^{1+1} A_{22} \;=\; A_{22} \\ C_{12} \;&=\; (-1)^{1+2} A_{21} \;=\; -A_{21} \;. \end{aligned} \tag{9.93}$$

The determinant of the 2×2 matrix **A** is therefore

$$\det(\mathbf{A}) \;=\; \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \;=\; A_{11}A_{22} - A_{12}A_{21} \;. \tag{9.94}$$

Now let's consider a slightly more complicated example, namely the determinant

$$\det(\mathbf{A}) \;=\; \begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} \tag{9.95}$$

[9] This notation might superficially resemble an absolute-value sign around the matrix, but don't read too much into that. The determinant of a matrix is not an "absolute value" of the matrix in any sense. For example, the determinant of a matrix can be negative.

of a 3×3 matrix. As always, we must begin by choosing a row, and we'll once again choose the first row. The sub-matrices obtained by crossing out the ith row and $j$the column of the matrix are

$$
\begin{aligned}
i=1,\ j=1\ :\ & \begin{pmatrix} \cancel{A_{11}} & \cancel{A_{12}} & \cancel{A_{13}} \\ A_{12} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \longrightarrow \begin{pmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{pmatrix} \\
i=1,\ j=2\ :\ & \begin{pmatrix} \cancel{A_{11}} & \cancel{A_{12}} & \cancel{A_{13}} \\ A_{12} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \longrightarrow \begin{pmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{pmatrix} \\
i=1,\ j=3\ :\ & \begin{pmatrix} \cancel{A_{11}} & \cancel{A_{12}} & \cancel{A_{13}} \\ A_{12} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \longrightarrow \begin{pmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix} ,
\end{aligned}
\tag{9.96}
$$

The minors M_{11}, M_{12}, and M_{13} are the respective determinants of these matrices 2×2 matrices. However, we already know how to evaluate the determinant of a 2×2 matrix. Therefore, determining the corresponding cofactors is simply a matter of plugging each of the sub-matrices above into Eq. (9.94):

$$
\begin{aligned}
C_{11} &= (-1)^{1+1} \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} = A_{22}A_{33} - A_{23}A_{32} \\
C_{12} &= (-1)^{1+2} \begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{13} \end{vmatrix} = -\big(A_{21}A_{33} - A_{23}A_{31}\big) \\
C_{13} &= (-1)^{1+3} \begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix} = A_{21}A_{32} - A_{22}A_{31} \ .
\end{aligned}
\tag{9.97}
$$

Substituting these results into Eq. (9.89), we find that the determinant of any 3×3 matrix is given by the general formula

$$
\det(\mathbf{A}) = A_{11}\big(A_{22}A_{33} - A_{23}A_{32}\big) - A_{12}\big(A_{21}A_{33} - A_{23}A_{31}\big) + A_{13}\big(A_{21}A_{32} - A_{22}A_{31}\big) \ . \tag{9.98}
$$

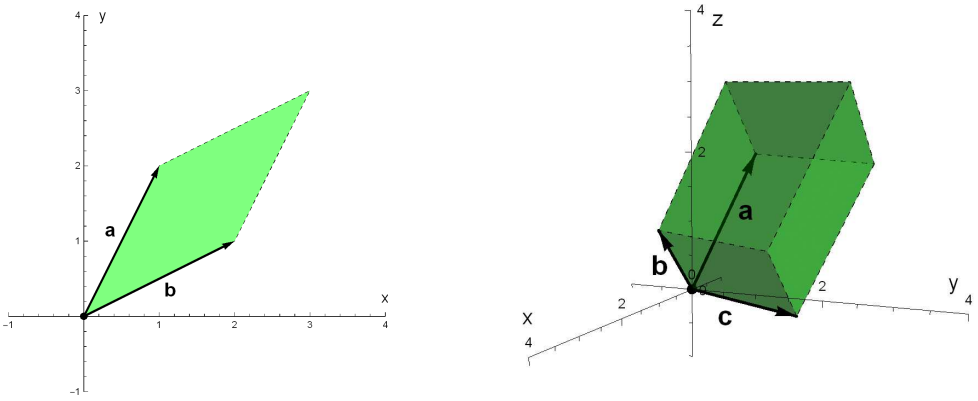


Figure 9.6: The absolute value $|\det(\mathbf{A})|$ of the determinant of a 2×2 matrix $\mathbf{A}$ represents the area of the parallelogram constructed from the vectors $\mathbf{a} = (A_{11}, A_{12})$ and $\mathbf{b} = (A_{21}, A_{22})$, which is shown in the left panel. The Likewise, the absolute value $|\det(\mathbf{A})|$ of the determinant of a 3×3 matrix $\mathbf{A}$ represents the area of the parallelepiped constructed from the vectors $\mathbf{a} = (A_{11}, A_{12}, A_{13})$, $\mathbf{b} = (A_{21}, A_{22}, A_{23})$, and $\mathbf{c} = (A_{31}, A_{32}, A_{33})$, which is shown in the right panel.

The intuitive meaning of the determinant may not be immediately obvious from this definition. However, the determinant does have a direct geometrical interpretation. The absolute value $|det(\mathbf{A})|$ of the determinant

of a 2×2 matrix with elements A_{ij} is equal to the area of the parallelogram constructed from the two vectors

$$\mathbf{a} \;=\; \begin{pmatrix} A_{11} \\ A_{12} \end{pmatrix} , \qquad \mathbf{b} \;=\; \begin{pmatrix} A_{21} \\ A_{22} \end{pmatrix} \tag{9.99}$$

as shown in the left panel of Fig. 9.6. Likewise, The absolute value $|det(\mathbf{A})|$ of the determinant of a 3×3 matrix with elements A_{ij} is equal to the area of the parallelepiped constructed from the three vectors

$$\mathbf{a} \;=\; \begin{pmatrix} A_{11} \\ A_{12} \\ A_{13} \end{pmatrix} , \qquad \mathbf{b} \;=\; \begin{pmatrix} A_{21} \\ A_{22} \\ A_{23} \end{pmatrix} , \qquad \mathbf{c} \;=\; \begin{pmatrix} A_{31} \\ A_{32} \\ A_{33} \end{pmatrix} \tag{9.100}$$

as shown in the right panel of Fig. 9.6. In general, the absolute value of the determinant of an $n \times n$ matrix $\mathbf{A}$ represents the volume of the n-dimensional paralellepiped constructed from a set of n vectors, where the components of each vector are given by the elements in a particular row of $\mathbf{A}$.

9.7 Operations on Matrices

In the Sect. 9.6, we discussed some of the general properties of matrices, including the determinant and the trace. In this section, we turn to consider what kinds of mathematical operations we can perform on matrices.

The first operation that we will consider is matrix addition. Two matrices can be added together if and only if they have the same number of rows m and the same number of columns n. The sum of two $m \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ is another $m \times n$ matrix $\mathbf{C}$ whose elements are given by $C_{ij} = A_{ij} + B_{ij}$. For example, the sum of the 3×2 matrices is

$$\mathbf{A} \;=\; \begin{pmatrix} 1 & -i \\ 2 & 4 \\ 1+i & 2 \end{pmatrix} , \qquad \mathbf{B} \;=\; \begin{pmatrix} 0 & 1 \\ -2 & 2i \\ -3 & -2 \end{pmatrix} \tag{9.101}$$

is the matrix

$$\begin{pmatrix} 1 & -i \\ 2 & 4 \\ 1+i & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ -2 & 2i \\ -3 & -2 \end{pmatrix} \;=\; \begin{pmatrix} 1 & 1-i \\ 0 & 4+2i \\ -2+i & 0 \end{pmatrix} . \tag{9.102}$$

The rules for subtracting matrices are completely analogous. For example,

$$\begin{pmatrix} 1 & -i \\ 2 & 4 \\ 1+i & 2 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ -2 & 2i \\ -3 & -2 \end{pmatrix} \;=\; \begin{pmatrix} 1 & -1-i \\ 4 & 4-2i \\ 4+i & 4 \end{pmatrix} . \tag{9.103}$$

In addition to adding and subtracting matrices, we can also multiply matrix by a scalar or by another matrix. The product $c\mathbf{A}$ of a scalar c with a matrix $\mathbf{A}$ and is another matrix $\mathbf{B}$ with elements $B_{ij} = cA_{ij}$. For example,

$$2 \begin{pmatrix} 2 & -1 \\ 5 & 3i \end{pmatrix} \;=\; \begin{pmatrix} 4 & -2 \\ 10 & 6i \end{pmatrix} . \tag{9.104}$$

The product $\mathbf{AB}$ of two matrices $\mathbf{A}$ and $\mathbf{B}$ is only defined when the number of rows in the matrix $\mathbf{A}$ is equal to the number of columns in the matrix $\mathbf{B}$. When $\mathbf{A}$ is an $m \times p$ matrix and $\mathbf{B}$ is a $p \times n$ $\mathbf{B}$ matrix, their product

$$\mathbf{C} \;=\; \mathbf{AB} \tag{9.105}$$

is an $m \times n$ matrix whose elements are given by

$$C_{ij} \;=\; \sum_{k=1}^{p} A_{ik} B_{kj} \; . \tag{9.106}$$

Note that order matters when multiplying matrices. We are used to the multiplication of two numbers together being commutative in the sense that $ab = ba$. By contrast, for matrices, $\mathbf{AB}$ is not necessarily equal to $\mathbf{BA}$. Moreover, it is possible for $\mathbf{AB}$ to be well-defined and for $\mathbf{BA}$ not to be defined at all! However, matrix multiplication *is* associative in the sense that

$$\mathbf{A}(\mathbf{BC}) \;=\; (\mathbf{AB})\mathbf{C}\;. \tag{9.107}$$

Furthermore, it's easy to verify that the transpose of the product of two matrices is the product of the transposes in the reverse order:

$$(\mathbf{AB})^T \;=\; \mathbf{B}^T\mathbf{A}^T\;. \tag{9.108}$$

An analogous property also holds for the Hermitian conjugate of a product of two matrices, since the Hermitian conjugate is just the complex conjugate of the transpose:

$$(\mathbf{AB})^\dagger \;=\; \mathbf{B}^\dagger\mathbf{A}^\dagger\;. \tag{9.109}$$

As an example of how this matrix-multiplication procedure works, let's calculate the product $\mathbf{AB}$ of the two matrices

$$\mathbf{A} \;=\; \begin{pmatrix} 3 & 0 & -1 \\ 2 & -2 & 0 \end{pmatrix}\;, \qquad \mathbf{B} \;=\; \begin{pmatrix} 1 & 0 \\ -1 & 4 \\ 0 & 2 \end{pmatrix}\;. \tag{9.110}$$

Since $\mathbf{A}$ is a 2×3 matrix and $\mathbf{B}$ is a 3×2 matrix, this product is well-defined and given by the 2×2 matrix

$$\begin{aligned} \begin{pmatrix} 3 & 0 & -1 \\ 2 & -2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 4 \\ 0 & 2 \end{pmatrix} \;&=\; \begin{pmatrix} 3\cdot 1 + 0\cdot(-1) + (-1)\cdot 0 & 3\cdot 0 + 0\cdot 4 + (-1)\cdot 2 \\ 2\cdot 1 + (-2)\cdot(-1) + 0\cdot 0 & 2\cdot 0 + (-2)\cdot 4 + 0\cdot 2 \end{pmatrix} \\ &=\; \begin{pmatrix} 3 & -2 \\ 4 & -8 \end{pmatrix}\;. \end{aligned} \tag{9.111}$$

In this example, the product of these matrices $\mathbf{BA}$ in the reverse order is also well-defined (see Problem 6), but $\mathbf{AB} \neq \mathbf{BA}$.

The same rules that govern the multiplication of two matrices together can also be used to multiply matrices and vectors together. In particular, let's say that I have chosen a basis and can therefore express any vector $|\alpha\rangle$ in my vector space as a column of numbers, as in Eq. (9.51). Such an array can be viewed as an $n\times 1$ matrix $\mathbf{a}$ whose elements a_i are the components of the vector in that basis. The product of an $m\times n$ matrix $\mathbf{M}$ with such a vector is therefore given by

$$\mathbf{b} \;=\; \mathbf{M}\mathbf{a}\;, \tag{9.112}$$

where $\mathbf{b}$ is an $m\times 1$ matrix — *i.e.*, another column of numbers. The m components b_i of this array are given by Eq. (9.106):

$$b_i \;=\; \sum_{j=1}^{m} M_{ij}a_j\;. \tag{9.113}$$

This formula is quite general and applies to the multiplication of any $n\times 1$ column vector by any $m\times n$ matrix. However, in order for the resulting array $\mathbf{b}$ to correspond to another vector in the same vector space as $\mathbf{a}$, it must have the same number of components as $\mathbf{a}$. Thus, in order for $\mathbf{b}$ to have any possibility of being a vector in the same vector space, $\mathbf{M}$ must be a square matrix with $m = n$.

Not only can we use the rules that govern matrix multiplication to multiply a matrix by a vector, but we can also use them to recast the expression in Eq. (9.65) for the inner product of two vectors as a matrix product. In particular, let's say that we have two vectors a and b expressed as $n\times 1$ arrays of the "column-vector" form given in Eq. (9.51) with respect to some orthonormal basis. The Hermitian conjugate $\mathbf{a}^\dagger$ of the vector $\mathbf{a}$ is a $1\times n$ matrix — a "row vector," if you will — components

$$\mathbf{a}^\dagger \;=\; \left(\mathbf{a}^*\right)^T \;=\; \begin{pmatrix} a_1^* & a_2^* & \ldots & a_n^* \end{pmatrix}\;. \tag{9.114}$$

Since the array $\mathbf{a}^\dagger$ is a $1 \times n$ matrix and $\mathbf{b}$ is an $n \times 1$ matrix, the product $\mathbf{a}^\dagger\mathbf{b}$ is well-defined. In particular, it's a 1×1 "matrix" — *i.e.*, a single number or scalar — whose sole component is given by

$$\mathbf{a}^\dagger\mathbf{b} \;=\; \begin{pmatrix} a_1^* & a_2^* & \ldots & a_n^* \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_n \end{pmatrix} \;=\; a_1^* b_1 + a_2^* b_2 + \ldots a_n^* b_n \;. \tag{9.115}$$

The expression on the right-hand side matches the result for for the inner product of two vectors in an orthonormal basis in Eq. (9.65). This means that if we're working in such a basis, we can always express the inner product of two vectors in the language of matrix multiplication. As we shall see in Sect. 9.8, the ability to do this comes in handy whenever one is using matrices to perform transformations on vectors.

The division of one number by another is not an operation for which for there is a straightforward parallel for matrices. However, there *is* an operation for $n \times n$ matrices which functions as an inverse operation to matrix multiplication in a sense that performing this operation will "undo" the action of multiplying one such matrix $\mathbf{B}$ by another such matrix $\mathbf{A}$. This operation consists of multiplication by a different matrix $\mathbf{A}^{-1}$, called the **inverse** of $\mathbf{A}$, defined such that

$$\mathbf{A}^{-1}(\mathbf{AB}) \;=\; \mathbf{B} \;. \tag{9.116}$$

So how do we find the inverse matrix $\mathbf{A}^{-1}$ for a given matrix $\mathbf{A}$? In order to answer this question, it is helpful to begin by introducing one additional concept: **identity matrix**. The identity matrix $\mathbf{I}_{n\times n}$ is an $n \times n$ matrix whose entries are given by $I_{ij} = \delta_{ij}$. In other words, all of the diagonal elements in the identity matrix are equal to 1 and all the off-diagonal elements are equal to zero. For example, the 3×3 identity matrix is

$$\mathbf{I}_{3\times 3} \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \;. \tag{9.117}$$

This matrix has the special property that for any $n \times n$ matrix $\mathbf{A}$,

$$\mathbf{I}_{n\times n}\mathbf{A} \;=\; \mathbf{A}\mathbf{I}_{n\times n} \;=\; \mathbf{A} \;, \tag{9.118}$$

which can be verified directly using Eq. (9.106). In other words, $\mathbf{I}_{n\times n}$ has the same fundamental properties under matrix multiplication that the number 1 has under scalar multiplication.

As we discussed above, the inverse matrix $\mathbf{A}^{-1}$ which satisfies Eq. (9.116) for any $n \times n$ matrix $\mathbf{B}$. Since matrix multiplication is associative, we have

$$\left(\mathbf{A}^{-1}\mathbf{A}\right)\mathbf{B} \;=\; \mathbf{B} \;, \tag{9.119}$$

which implies that

$$\mathbf{A}\mathbf{A}^{-1} \;=\; \mathbf{I}_{n\times n} \;. \tag{9.120}$$

The elements of $\mathbf{A}^{-1}$, if such a matrix exists, are given by

$$\left[\mathbf{A}^{-1}\right]_{ij} \;=\; \frac{1}{\det(\mathbf{A})} C_{ji} \;, \tag{9.121}$$

where C_{ji} is the cofactor of the element A_{ji} of $\mathbf{A}$, as defined in Eq. (9.88). Note that the indices i and j in C_{ji} are reversed on the right-hand side of this equation relative to the left-hand side.

However, not every square matrix has an inverse. In particular, we can see that the denominator on the right-hand side of Eq. (9.121) blows up whenever $\det(\mathbf{A}) = 0$. Indeed, it turns out that a matrix has a well-defined inverse if and only if $\det(\mathbf{A}) \neq 0$. Such a matrix is called an **invertable** matrix. A matrix whose determinant vanishes and therefore cannot be inverted is called a **singular** matrix.

Finally, there is even a sense in which matrices can be used as the arguments of functions. What exactly it means for a function to take a matrix as an argument may not be obvious, but there is a straightforward and unambiguous prescription for evaluating functions with such arguments. In particular, a function $f(\mathbf{A})$

whose argument is a matrix is to be evaluated according to its Taylor-series expansion. For example, the Taylor-series expansion of the exponential e^a of a scalar a is

$$e^a \;=\; 1 + a + \frac{1}{2!}a^2 + \frac{1}{3!}a^3 + \ldots \tag{9.122}$$

The Taylor-series expansion of the exponential of an $n \times n$ matrix $\mathbf{A}$ is defined in a completely analogous way:

$$e^{\mathbf{A}} \;=\; \mathbf{I}_{n\times n} + \mathbf{A} + \frac{1}{2!}\mathbf{A}\mathbf{A} + \frac{1}{3!}\mathbf{A}\mathbf{A}\mathbf{A} + \ldots \tag{9.123}$$

Note that the sum which appears on the right-hand side of this equation represents another $n \times n$ matrix. Other functions of a matrix, including trigonometric functions, logarithms, *etc.*, are defined in a similar manner.

Example:

Let's evaluate the matrix exponential $e^{i\theta\boldsymbol{\sigma}_1}$, where θ is a scalar and where

$$\boldsymbol{\sigma}_1 \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} . \tag{9.124}$$

The Taylor series for the exponential of a matrix was given in Eq. (9.123). Here, the matrix appearing in the exponential is the product $i\theta\boldsymbol{\sigma}_1$ of the scalar $i\theta$ and the matrix $\boldsymbol{\sigma}_1$, so we have

$$e^{i\theta\boldsymbol{\sigma}_1} \;=\; \mathbf{I}_{2\times 2} + i\theta\boldsymbol{\sigma}_1 + \frac{(i\theta)^2}{2!}\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1 + \frac{(i\theta)^3}{3!}\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1 + \ldots \tag{9.125}$$

The product of two $\boldsymbol{\sigma}_1$ matrices with each other turns out to be the identity matrix:

$$\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1 \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \;=\; \mathbf{I}_{2\times 2} . \tag{9.126}$$

This, in turn, means that the product of three $\boldsymbol{\sigma}_1$ matrices is just $\mathbf{I}_{2\times 2}\boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_1$, the product of four $\boldsymbol{\sigma}_1$ matrices is $\boldsymbol{\sigma}_1$, and so on. In other words, the product of any odd number of $\boldsymbol{\sigma}_1$ matrices is $\boldsymbol{\sigma}_1$, while the product of and even number of $\boldsymbol{\sigma}_1$ matrices is $\mathbf{I}_{2\times 2}$. Thus, we have

$$\begin{aligned} e^{i\theta\boldsymbol{\sigma}_1} &= \mathbf{I}_{2\times 2}\left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \ldots\right) + \boldsymbol{\sigma}_1\left(i\theta - i\frac{\theta^3}{3!} + i\frac{\theta^5}{5!} - \ldots\right) \\ &= \mathbf{I}_{2\times 2}\cos\theta + i\boldsymbol{\sigma}_1\sin\theta \\ &= \begin{pmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix} \end{aligned} \tag{9.127}$$

9.8 Matrices and Transformations of Vectors

In Sect. 9.7, we saw that the product of an $n \times n$ matrix $\mathbf{A}$ and an $n \times 1$ matrix a is another $n \times 1$ matrix b with components b_i given by Eq. (9.113). Let's consider the case in which the elements of a represent the components of a vector in some vector space with respect to some particular basis. Since the resulting matrix b is a linear combination of the same basis vectors, it follows that b must also be a vector in the space. This is a very important observation. It means that multiplying a vector by a matrix has the effect of transforming it into another vector in the same space.

One example of a transformation that can be represented by a matrix is the rotation of a vector in $\mathbb{R}^2$ by an angle θ. The matrix $\mathbf{R}$ which corresponds to such a (counterclockwise) rotation is

$$\mathbf{R} \;=\; \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} . \tag{9.128}$$

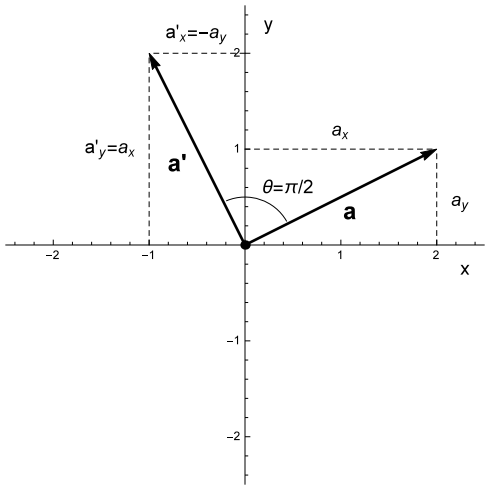


Figure 9.7: A vector $\mathbf{a}$ in $\mathbb{R}^2$ and the vector $\mathbf{a}' = \mathbf{R}\mathbf{a}$ which results from multiplying $\mathbf{a}$ by the matrix in Eq. (9.128) with $\theta = \pi/2$. We see that the effect of multiplying $\mathbf{a}$ by $\mathbf{R}$ is to produce a vector which corresponds to our original vector rotated counterclockwise in the plane by $\pi/2$ radians.

The multiplication of any vector $\mathbf{a}$ in the space by this vector yields another vector $\mathbf{a}'$ whose components are given in terms of the components of the original vector by Eq. (9.113):

$$\mathbf{a}' \;=\; \mathbf{R}\mathbf{a} \;=\; \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a_x \\ a_y \end{pmatrix} \;=\; \begin{pmatrix} a_x\cos\theta - a_y\sin\theta \\ a_x\sin\theta + a_y\cos\theta \end{pmatrix} . \tag{9.129}$$

We can check that this does in fact correspond to a counterclockwise rotation of our original vector $\mathbf{a}$ by plugging in some explicit values for θ. For example, plugging in $\theta = \pi/2$ gives us

$$\theta \;=\; \frac{\pi}{2} \quad\longrightarrow\quad \mathbf{a}' \;=\; \begin{pmatrix} -a_y \\ a_x \end{pmatrix} , \tag{9.130}$$

which is a vector with components $a'_x = -a_y$ and $a'_y = a_x$, as shown in Fig. 9.7. Indeed, this vector corresponds to a counterclockwise rotation of our original vector by $\pi/2$ radians. Likewise, setting $\theta = \pi$ yields a vector with components $a'_x = -a_x$ and $a'_y = -a_y$, which corresponds to a rotation of $\mathbf{a}$ by π radians.

The rotation of a vector in $\mathbb{R}^2$ is an example of a special kind of transformation, called a **unitary transformation**. A unitary transformation is a transformation that preserves length. In other words, it's a transformation that leaves the norm $||\alpha||$ of the original vector $|\alpha\rangle$ unchanged. A matrix $\mathbf{U}$ which corresponds to a unitary transformation is called a **unitary matrix**. [10]

Clearly, in order to preserve length, a unitary matrix must possess some special properties. Let's investigate a little further exactly what these properties are. If the norm $||\alpha|| = \sqrt{\langle\alpha|\alpha\rangle}$ of the original vector is equal to the norm $||\alpha'|| = \sqrt{\langle\alpha'|\alpha'\rangle}$ of the transformed vector, it must be true that

$$\langle\alpha'|\alpha'\rangle \;=\; \langle\alpha|\alpha\rangle \; . \tag{9.131}$$

If we are working in an orthonormal basis, we can use Eq (9.115) to express each side of this equation as a matrix product:

$$\mathbf{a'}^{\dagger}\mathbf{a}' \;=\; \mathbf{a}^{\dagger}\mathbf{a} \tag{9.132}$$

Moreover, since $\mathbf{a} \;=\; \mathbf{U}\mathbf{a}$, we have

$$\big(\mathbf{U}\mathbf{a}\big)^{\dagger}\big(\mathbf{U}\mathbf{a}\big) \;=\; \mathbf{a}^{\dagger}\mathbf{U}^{\dagger}\mathbf{U}\mathbf{a} \; , \tag{9.133}$$

[10] The corresponding noun is "unitarity."

where, in the second step, we have used Eq. (9.109) to rewrite the Hemitian conjugate of the matrix product **Ua** in terms of the Hermitian conjugates of **U** and **a** themselves in the reverse order. Thus, our unitarity condition in Eq. (9.132) becomes

$$\langle \alpha' | \alpha' \rangle \;=\; \langle \alpha | \alpha \rangle \quad \leftrightarrow \quad \mathbf{a'}^{\dagger}\mathbf{a'} \;=\; \mathbf{a}^{\dagger}\mathbf{U}^{\dagger}\mathbf{U}\mathbf{a} \;=\; \mathbf{a}^{\dagger}\mathbf{a}\;, \tag{9.134}$$

which implies that

$$\mathbf{U}^{\dagger}\mathbf{U} \;=\; \mathbf{I}_{n\times n}\;. \tag{9.135}$$

Comparing this result with Eq. (9.120), we see that a matrix is unitary if and only if its Hermitian conjugate is equal to its inverse:

$$\mathbf{U}^{\dagger} \;=\; \mathbf{U}^{-1} \quad \text{(Unitary)}\;. \tag{9.136}$$

.

Another kind of transformation which corresponds to multiplication by a unitary matrix is the conversion from one orthonormal basis to another. Indeed, changing the basis with respect to which we are writing the components of a vector shouldn't alter any of the the fundamental properties of that vector such as its length. Thus, it stands to reason that such a transformation should be unitary. Moreover, it's straightforward to determine what the corresponding matrix **U** looks like for any such transformation. Let's call the set of basis vectors in our original basis $|e_i\rangle$ and the basis vectors in our new basis $|e'_i\rangle$. The relationship between the components a'_i of the vector in the new basis and the components a_i in the old basis are given by Eq. (9.66):

$$a'_i \;=\; \langle e'_i|\alpha\rangle \;=\; \langle e'_i| \left(\sum_{j=1}^{n} a_j |e_j\rangle \right) \;=\; \sum_{j=1}^{N} a_j \langle e'_i|e_j\rangle\;. \tag{9.137}$$

Comparing this expression to Eq. (9.113), we see that this basis transformation is completely equivalent to multiplying our original vector a_j by a matrix whose elements are given by

$$U_{ij} \;=\; \langle e'_i|e_j\rangle\;. \tag{9.138}$$

Thus, if we have a vector **a** whose components are expressed with respect to one orthonormal basis, the components of the corresponding vector a′ in another basis are given by

$$\mathbf{a'} \;=\; \mathbf{U}\mathbf{a}\;, \tag{9.139}$$

where the elements of the matrix **U** are given by Eq. (9.138).

As an example of how this works, let's take the vector in $\mathbb{R}^2$ whose components are given by

$$\mathbf{a} \;=\; \begin{pmatrix} 4 \\ -2 \end{pmatrix} \tag{9.140}$$

in the standard orthonormal basis

$$\mathbf{e}_1 \;=\; \hat{\mathbf{x}}\;, \qquad \mathbf{e}_2 \;=\; \hat{\mathbf{y}} \tag{9.141}$$

and re-express this vector in the orthonormal basis

$$\mathbf{e}'_1 \;=\; \frac{1}{\sqrt{2}}(\hat{\mathbf{x}}+\hat{\mathbf{y}})\;, \qquad \mathbf{e}'_2 \;=\; \frac{1}{\sqrt{2}}(\hat{\mathbf{x}}-\hat{\mathbf{y}})\;. \tag{9.142}$$

The matrix which corresponds to this transformation is

$$\mathbf{U} \;=\; \begin{pmatrix} \langle e'_1|e_1\rangle & \langle e'_1|e_2\rangle \\ \langle e'_2|e_1\rangle & \langle e'_2|e_2\rangle \end{pmatrix} \;=\; \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\;. \tag{9.143}$$

First, let's verify that this matrix is unitary. The Hermetian conjugate of **U** turns out to be equal to **U** itself

$$\mathbf{U}^{\dagger} \;=\; [\mathbf{U}^{*}]^{T} \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \;=\; \mathbf{U}\;. \tag{9.144}$$

Thus, we find that $\mathbf{U}$ indeed satisfies Eq. (9.136):

$$\mathbf{U}^\dagger\mathbf{U} \;=\; \left(\frac{1}{\sqrt{2}}\right)^2 \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \;=\; \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \;=\; \mathbf{I}_{2\times 2}\;. \tag{9.145}$$

Now let's use this unitary matrix to determine the components of the vector from Eq. (9.140) in our new basis. Rewriting $\mathbf{a}$ to this new basis is simply a matter of multiplying $\mathbf{a}$ by $\mathbf{U}$ gives us, so the corresponding vector $\mathbf{a}'$ in the new basis is

$$\mathbf{a}' \;=\; \mathbf{U}\mathbf{a} \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 4 \\ -2 \end{pmatrix} \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 4-2 \\ 4+2 \end{pmatrix} \;=\; \begin{pmatrix} \sqrt{2} \\ 3\sqrt{2} \end{pmatrix}\;. \tag{9.146}$$

In other words, our vector is specified by the linear combination of basis vectors

$$\mathbf{a}' \;=\; \sqrt{2}\,\mathbf{e}_1' + 3\sqrt{2}\,\mathbf{e}_2'\;. \tag{9.147}$$

The operation in Eq. (9.139) gives us a prescription for transforming a vector whose components are expressed with respect to some original basis into a new basis. However, we can also just as easily use this prescription to transform the resulting vector from the new basis back into to the old one. In particular, to "undo" this operation and bring ourselves back to the original basis, all we need to do is multiply both sides of Eq. (9.139) by the inverse $\mathbf{U}^{-1} = \mathbf{U}^\dagger$ of our unitary matrix:

$$\mathbf{U}^\dagger\mathbf{a}' \;=\; \mathbf{U}^\dagger\mathbf{U}\mathbf{a} \;=\; \mathbf{I}_{2\times 2}\mathbf{a} \;=\; \mathbf{a}\;. \tag{9.148}$$

In other words, the transformation which brings us back to the old basis corresponds to multiplication by the Hermitian conjugate of $\mathbf{U}$:

$$\mathbf{a} \;=\; \mathbf{U}^\dagger\mathbf{a}' \tag{9.149}$$

We can also use Eq. (9.149) to figure out how a matrices transform when we convert from one basis to another. For example, let's say that the product of a matrix $\mathbf{A}$ and a vector $\mathbf{a}$ gave us the vector b in our original basis:

$$\mathbf{b} \;=\; \mathbf{A}\mathbf{a}\;. \tag{9.150}$$

To obtain the corresponding equation in a new basis in which $\mathbf{a}' = \mathbf{U}\mathbf{a}$ and $\mathbf{b}' = \mathbf{U}\mathbf{b}$, all we need to do is use Eq. (9.149) to express $\mathbf{a}$ and $\mathbf{b}$ in terms of $\mathbf{a}'$ and $\mathbf{b}'$

$$\mathbf{U}^\dagger\mathbf{b}' \;=\; \mathbf{A}\mathbf{U}^\dagger\mathbf{a}'\;. \tag{9.151}$$

If we multiply both sides of this equation by $\mathbf{U}$, we get

$$\mathbf{U}\mathbf{U}^\dagger\mathbf{b}' \;=\; \mathbf{b}' \;=\; \mathbf{U}\mathbf{A}\mathbf{U}^\dagger\mathbf{a}'\;. \tag{9.152}$$

Comparing the form of this equation to Eq. (9.150), we see that the matrix which corresponds to $\mathbf{A}$ in the new basis is

$$\mathbf{A}' \;=\; \mathbf{U}\mathbf{A}\mathbf{U}^\dagger\;. \tag{9.153}$$

This is in fact a general rule which specifies how a matrix transforms under any unitary transformation.

9.9 Eigenvalues and Eigenvectors

It turns out there are often certain choices of basis one can adopt within which equations involving matrices — for example, equations of the form given in Eq. (9.150) — become particularly simple. For example, a basis may exist wherein a particular matrix consists entirely of diagonal entries. Clearly, if we can identify such a basis for any particular matrix, it makes calculations a lot easier. The question, then, is how we go about finding such a basis — or, for that matter, determining whether one even exists. Fortunately, there is a method for doing this. This method involves calculating a set of numbers called the eigenvalues of the matrix, along with a corresponding set of vectors called its eigenvectors. We therefore begin this section with a discussion of exactly what these quantities are and how to determine them.

For a given matrix $\mathbf{A}$, there is often a set of vectors $\mathbf{b}_i$ for which the matrix product of $\mathbf{A}$ and $\mathbf{b}_i$ is just the vector $\mathbf{b}_i$ itself multiplied by some overall scalar coefficient λ_i:

$$\mathbf{A}\mathbf{b}_i \;=\; \lambda_i \mathbf{b}_i \;. \tag{9.154}$$

The set of vectors $\mathbf{b}_i$ which have this property for any particular matrix are called the **eigenvectors** of that matrix. The corresponding scalars λ_i are called the **eigenvalues**. Of course the null vector $\mathbf{0}$ will always satisfy this relation, but that's a bit too trivial — we're interested in non-trivial solutions to this equation, and thus $\mathbf{0}$ is generally excluded from the set of eignevectors for a given matrix.

So how can we find these non-trivial solutions? Moreover, how do we even know such solutions exist for a given matrix? The first step in answering these questions is to rewrite Eq. (9.154) in a slightly more revealing form. If we insert the identity matrix on the right-hand side of this equation

$$\mathbf{A}\mathbf{b} \;=\; \lambda \mathbf{I}_{n\times n}\mathbf{b} \;, \tag{9.155}$$

we can rewrite it in the form

$$\big(\mathbf{A} - \lambda_i \mathbf{I}_{n\times n}\big)\mathbf{b}_i \;=\; \mathbf{0} \;. \tag{9.156}$$

We can determine whether or not there exist non-trivial solutions to this vector equation using the following reasoning. Let's say that the matrix $\mathbf{A} - \lambda \mathbf{I}_{n\times n})$ in this equation had a well-defined inverse. If that were true, then we could multiply both sides of the equation by this inverse to obtain

$$\big(\mathbf{A} - \lambda_i \mathbf{I}_{n\times n}\big)^{-1}\big(\mathbf{A} - \lambda_i \mathbf{I}_{n\times n}\big)\mathbf{b}_i \;=\; \mathbf{I}_{n\times n}\mathbf{b}_i \;=\; \mathbf{b}_i \;=\; \big(\mathbf{A} - \lambda_i \mathbf{I}_{n\times n}\big)^{-1}\mathbf{0} \;=\; \mathbf{0} \;. \tag{9.157}$$

In other words, the only solution to Eq. (9.156) is the trivial one, with $\mathbf{b} = 0$. This means that if Eq. (9.156) is to have non-trivial solutions, the matrix $\mathbf{A} - \lambda_i \mathbf{I}_{n\times n}$ must not be invertible. As discussed in Sect. 9.7, this implies that its determinant must vanish:

$$\det\big(\mathbf{A} - \lambda_i \mathbf{I}_{n\times n}\big) \;=\; 0 \tag{9.158}$$

This equation is called the **characteristic equation** for the matrix $\mathbf{A}$. The eigenvalues λ_i are the set of solutions to this equation. To find the corresponding eigenvector $\mathbf{b}_i$ for each λ_i, all we need to do is plug that value of λ_i back into Eq. (9.156) and solve for the components of $\mathbf{b}_i$

Example: Eigenvalues of a Matrix

As an example of how this works, let's determine the eigenvalues and eigenvectors of the 2×2 matrix

$$\mathbf{A} \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} \;. \tag{9.159}$$

The characteristic equation is

$$\det\big(\mathbf{A} - \lambda_i \mathbf{I}_{2\times 2}\big) \;=\; \begin{vmatrix} -\lambda_i & 1 \\ 1 & 2-\lambda_i \end{vmatrix} \;=\; -\lambda_i(2-\lambda_i) - 1 \;=\; 0 \;. \tag{9.160}$$

This is just a quadratic equation for λ_i, so the eigenvalues of $\mathbf{A}$ are the two solutions given by

$$\lambda_i \;=\; \frac{-(-2) \pm \sqrt{2^2 - 4\cdot 1\cdot(-1)}}{2\cdot 1} \;=\; 1 \pm \sqrt{2} \;. \tag{9.161}$$

In other words, the two eigenvalues are

$$\lambda_1 \;=\; 1 + \sqrt{2} \;, \qquad \lambda_2 \;=\; 1 - \sqrt{2} \;. \tag{9.162}$$

The eigenvectors $\mathbf{b}_1$ and $\mathbf{b}_2$ are the vectors whose components satisfy Eq. (9.156) for λ_1 and λ_2, respectively. We'll begin by solving the equation for $\mathbf{b}_1$. Let's call the two components of this vector $b_{1,1}$ and $b_{1,2}$.

The

$$\begin{aligned}\begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \left[\begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} - \begin{pmatrix} 1+\sqrt{2} & 0 \\ 0 & 1+\sqrt{2} \end{pmatrix}\right]\begin{pmatrix} b_{1,1} \\ b_{1,2} \end{pmatrix} \\ &= \begin{pmatrix} -1-\sqrt{2} & 1 \\ 1 & 1-\sqrt{2} \end{pmatrix}\begin{pmatrix} b_{1,1} \\ b_{1,2} \end{pmatrix} \\ &= \begin{pmatrix} b_{1,2} - (1+\sqrt{2})b_{1,1} \\ b_{1,1} + (1-\sqrt{2})b_{1,2} \end{pmatrix} . \end{aligned} \tag{9.163}$$

This is equivalent to the system of equations

$$\begin{aligned} -(1+\sqrt{2})b_{1,1} + b_{1,2} &= 0 \\ b_{1,1} + (1-\sqrt{2})b_{1,2} &= 0 . \end{aligned} \tag{9.164}$$

The solution to this system of equations is $b_{1,1} = -1+\sqrt{2}$ and $b_{1,2} = 1$, so the eigenvector associated with the eigenvalue λ_1 is

$$\mathbf{b}_1 = \begin{pmatrix} \sqrt{2}-1 \\ 1 \end{pmatrix} . \tag{9.165}$$

The procedure for finding the eigenvector $\mathbf{b}_2$ associated with the eigenvalue λ_2 is completely analogous. The result is

$$\mathbf{b}_2 = \begin{pmatrix} -\sqrt{2}-1 \\ 1 \end{pmatrix} . \tag{9.166}$$

It's worth noting that the two eigenvectors $\mathbf{b}_1$ and $\mathbf{b}_2$ are not only linearly independent, but actually orthogonal:

$$\mathbf{b}_1 \cdot \mathbf{b}_2 = \big(\sqrt{2}-1\big)\big(-\sqrt{2}-1\big) + 1^2 = 0 . \tag{9.167}$$

Thus, we can form an orthonormal basis out of these eigenvectors simply by normalizing them. In this basis, the eigenvectors are

$$\begin{aligned} \mathbf{e}'_1 &= \frac{\mathbf{b}_1}{||\mathbf{b}_1||} = \frac{1}{\big[4-2\sqrt{2}\big]^{1/2}}\begin{pmatrix} \sqrt{2}-1 \\ 1 \end{pmatrix} \\ \mathbf{e}'_2 &= \frac{\mathbf{b}_2}{||\mathbf{b}_2||} = \frac{1}{\big[4+2\sqrt{2}\big]^{1/2}}\begin{pmatrix} -\sqrt{2}-1 \\ 1 \end{pmatrix} . \end{aligned} \tag{9.168}$$

In the above example, we found that our 2×2 matrix had two distinct eigenvalues λ_1 and λ_2. Moreover, the corresponding eigenvectors even turned out to be linearly independent, to be orthogonal, and to span the space. Indeed, all we had to do in order to construct an orthonormal basis out of them was to normalize them. It is not always the case that an $n\times n$ matrix has n distinct eigenvalues, however, much less that they turn out to be orthogonal. Indeed, the eigenvalues of a matrix can be degenerate, meaning that two or more distinct eigenvectors correspond to the same eigenvalue. Moreover, an $n\times n$ matrix need not necessarily have n distinct eigenvectors either; rather, it may have any number of eigenvectors from one — and there will always be at least one — up to n.

However, there is a broad class of matrices which are guaranteed to have n linearly-independent, orthogonal eigenvectors which span the vector space. Matrices in this class are called **normal matrices**. It turns out that any matrix which satisfies the condition

$$\mathbf{A}\mathbf{A}^\dagger - \mathbf{A}^\dagger\mathbf{A} = \mathbf{0} \tag{9.169}$$

is a member of this class. For example, Eq. (9.135) implies that every unitary matrix is a normal matrix because

$$\mathbf{U}\mathbf{U}^\dagger - \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}_{2\times 2} - \mathbf{I}_{2\times 2} = \mathbf{0} . \tag{9.170}$$

In addition, any Hermetian or anti-Hermetian matrix is normal, as is any real symmetric matrix. There are plenty of other examples as well. Since the eigenvectors for a normal matrix are linearly independent, spanning, and orthogonal, we can always use them to construct an orthonormal basis.

Why is it significant or useful that the set of eigenvectors of a matrix form an othonormal basis? In order to answer this, let's return to the above example in which we constructed an orthonormal basis out of the eigenvalues of the matrix $\mathbf{A}$ in Eq. (9.159). Let us now ask ourselves what the matrix $\mathbf{A}$ would look like if we transformed it into this new basis. Calculating the unitary matrix $\mathbf{U}$ which takes us from our original basis

$$\mathbf{e}_1 \;=\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \qquad \mathbf{e}_2 \;=\; \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{9.171}$$

to the basis defined in Eq. (9.168) is straightforward. Its elements U_{ij} are given by Eq. (9.138), so we have

$$\mathbf{U} \;=\; \begin{pmatrix} \frac{\sqrt{2}-1}{[4-2\sqrt{2}]^{1/2}} & \frac{1}{[4-2\sqrt{2}]^{1/2}} \\ \frac{-\sqrt{2}-1}{[4+2\sqrt{2}]^{1/2}} & \frac{1}{[4+2\sqrt{2}]^{1/2}} \end{pmatrix} . \tag{9.172}$$

The matrix which corresponds to $\mathbf{A}'$ in the new basis is given by substituting this result into Eq. (9.153). You can verify for yourself that the resulting expression for $\mathbf{A}'$ ultimately reduces to

$$\mathrm{A}' \;=\; \mathbf{U}\mathbf{A}\mathbf{U}^\dagger \;=\; \begin{pmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{pmatrix} . \tag{9.173}$$

Apparently, in this special basis, the matrix $\mathbf{A}'$ is a diagonal matrix whose diagonal entries are the eigenvalues of our original matrix! Clearly, working in this special basis — often called the **eigenbasis** of the matrix $\mathbf{A}$ — definitely has its advantages.

This example is in fact an illustration of a general property of normal matrices — and an exceedingly useful one. Whenever we make a change of basis to the eigenbasis of a normal matrix $\mathbf{A}$, the corresponding matrix $\mathbf{A}'$ will always be purely diagonal. In particular, the diagonal entries A'_{ii} of $\mathbf{A}'$ in this basis are given by the eigenvalues of the original matrix. In other words, the elements of $\mathbf{A}'$ are given by

$$A'_{ij} \;=\; \lambda_i \delta_{ij} \; . \tag{9.174}$$

For this reason, the procedure of converting our original matrix $\mathbf{A}$ into this special basis is called **diagonalization**.

9.10 Coupled Differential Equations as Matrix Equations

For example, let's return to the coupled-oscillator system from Sect. 9.2 and see how we can describe this system using the language of linear algebra. The state of the system at any given moment is completely described by a pair of coordinates x_1 and x_2 which describe positions of the two blocks. We can view these two position coordinates as the components of a vector

$$\mathbf{x} \;=\; \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{9.175}$$

in an abstract, two-dimensional vector space.[11] Specifically, x_1 and x_2 are the components of this vector with respect to an orthonormal basis in which the basis vectors are

$$\mathbf{e}_1 \;=\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \qquad \mathbf{e}_2 \;=\; \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{9.176}$$

Each vector $\mathbf{x}$ in this vector space specifies a unique configuration of the positions of the individual blocks. I will therefore refer to this vector as the "configuration vector" for the system.[12]

Two examples of vectors in this abstract vector space are shown in Fig. 9.8, along with the physical configurations of the blocks to which each vector corresponds.

[11] Since the coordinates x_1 and x_2 are real-valued, this vector space is equivalent to $\mathbb{R}^2$. Admittedly, not every choice of these coordinates is physically sensible For example, if the distance between the equilibrium positions of the two blocks is d, any vector with $x_1 > x_2 + d$ is physically nonsensical because it corresponds to the block with coordinate x_1 lying to the right of the block with coordinate x_2. This does not mean that the space fails to satisfy the criteria for a vector space (*e.g.*, closure). Mathematically speaking, the space is still $\mathbb{R}^2$. Rather, we just need to be careful that we are operating within a physically sensible region of this space.

[12] I use the word "configuration" rather than "state" to describe this set of positions because the word "state" is typically reserved for a set of both position and momentum coordinates together.

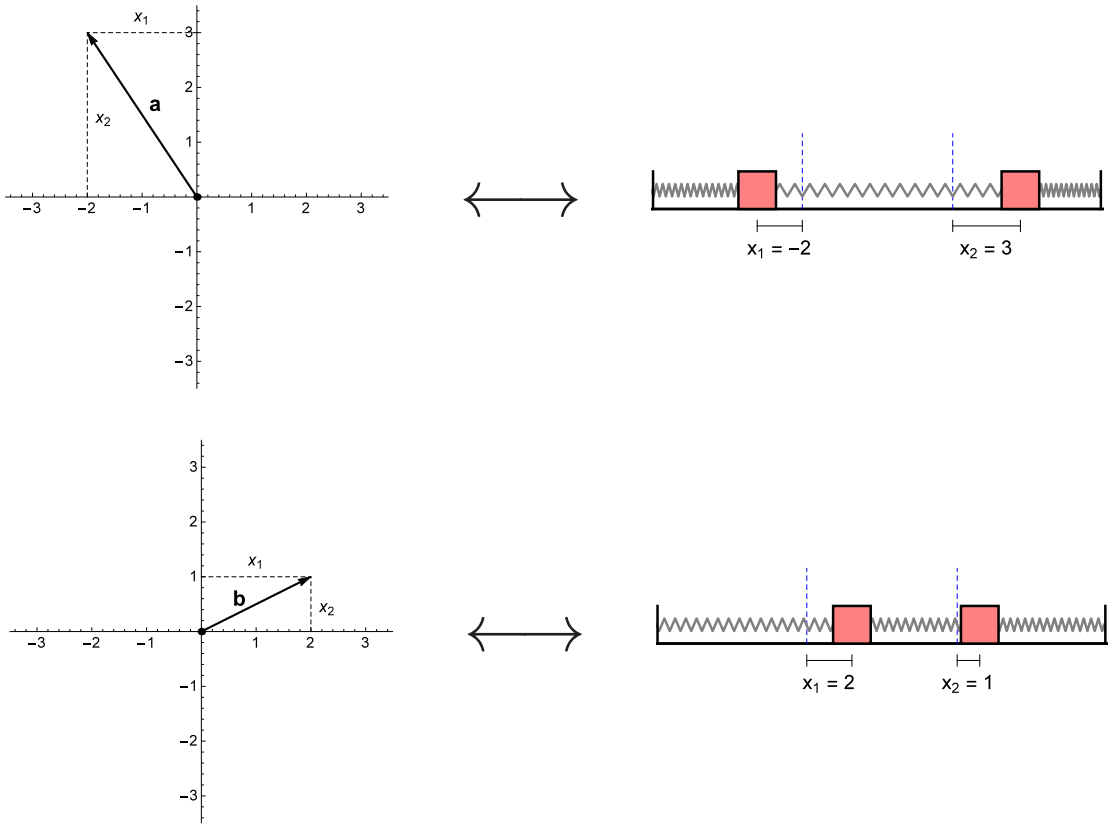


Figure 9.8: An illustration of how the overall position configuration of a two-oscillator system can be represented as a vector in a two-dimensional vector space. It is convenient to work in the (orthonormal) basis in which the components of the vector correspond to the two position coordinates x_1 and x_2 of the individual oscillators. The top panel shows the vector $\mathbf{a} = (-2, 3)$ in this abstract space, along with an illustration of the corresponding physical configuration of the blocks. Likewise, the bottom panel shows the vector $\mathbf{b} = (2, 1)$ in the same abstract space, along with an illustration of the corresponding physical configuration.

Another advantage of our coupled-oscillator system as a vector also allows us to write the equations of motion for x_1 and x_2 in a much more compact form. In particular, we can combine Eqs. (9.1) and (9.2) into a single matrix equation:

$$m\frac{d^2}{dt^2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2k & k \\ k & -2k \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} . \tag{9.177}$$

This equation can be viewed as merely a repackaging of the equations of motion for the coordinates x_1 and x_2 using matrix notation. However, it can also be viewed as an equation of motion for the overall configuration vector $\mathbf{x}$.

$$m\frac{d^2}{dt^2}\mathbf{x} = k\mathbf{M}\mathbf{x} , \tag{9.178}$$

where we have defined

$$\mathbf{M} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} . \tag{9.179}$$

The fact that our original differential equations for x_1 and x_2 were coupled is reflected in the fact that this matrix is not diagonal. It is the off-diagonal elements of $\mathbf{M}$ that "mix" the different components so that the derivative of on component of our configuration vector depends on the value of the other. If, on the other hand, the matrix had been diagonal, then the equations for the two components would have decoupled and we could have solved each of these equations separately.

However, one of the crucial lessons from Sect. 9.9 is that there exists a basis we can choose in which the matrix *is* diagonal and the corresponding equations for the components *do* decouple. That basis is the eigenbasis of the matrix $\mathbf{M}$. Indeed, $\mathbf{M}$ is a real symmetric matrix. All real symmetric matrices are normal matrices, as we saw in Sect. 9.9, so we know that there exists a basis in which $\mathbf{M}$ is purely diagonal. Once we identify that basis, we can decouple the equations of motion by performing a basis transformation. Then the system can be solved. Of course the result won't be a surprise, since we already know what the solutions look like from Sect. 9.2. However, in contrast with the trick we used there to obtain those solutions, the method we will walk through below can be applied to nearly any coupled system of physical interest. Moreover, it will shed a lot more light on many of the concepts to which we were introduced in Sect. 9.2 — including the concept of normal modes.

Finding the basis vectors for the basis in which $\mathbf{M}$ is diagonal is simply a matter of finding the eigenvalues and eigenvectors of this matrix. The first step is to determine the eigenvalues by solving the characteristic equation

$$\det\left(\mathbf{M} - \lambda \mathbf{I}_{2\times 2}\right) \;=\; 0 \;, \tag{9.180}$$

where λ represents either of the two eigenvalues of $\mathbf{M}$. Evaluating the determinant gives us

$$\left| \begin{array}{cc} -2-\lambda & 1 \\ 1 & -2-\lambda \end{array} \right| \;=\; (-2-\lambda)^2 - (-1)^2 \;=\; \lambda^2 + 4\lambda + 3^2 \;=\; 0 \;. \tag{9.181}$$

The solutions $\lambda_\pm$ to this quadratic equation are

$$\lambda_\pm \;=\; \frac{-4 \pm \sqrt{4^2 - 4\cdot 1\cdot 3}}{2\cdot 1} \;=\; \frac{-4\pm\sqrt{4}}{2} \;=\; -2 \pm 1 \;. \tag{9.182}$$

In other words, the eigenvalues of $\mathbf{M}$ are

$$\lambda_+ \;=\; -1 \;, \qquad \lambda_- \;=\; -3 \;. \tag{9.183}$$

The next step is to determine the eigenvectors $\mathbf{e}_+$ and $\mathbf{e}_-$, which represent the respective solutions to the equations

$$\left(\mathbf{M} - \lambda_\pm \mathbf{I}_{2\times 2}\right)\mathbf{e}_\pm \;=\; 0 \;. \tag{9.184}$$

Explicitly, for $\mathbf{e}_+$, we have

$$\left(\begin{array}{cc} -2+1 & 1 \\ 1 & -2+1 \end{array} \right) \left(\begin{array}{c} e_{+,1} \\ e_{+,2} \end{array} \right) \;=\; 0 \;, \tag{9.185}$$

where $e_{+,1}$ and $e_{+,2}$ are just the components of $\mathbf{e}_+$ with respect to our original basis. This matrix relation corresponds to the system of equations

$$\begin{aligned} -e_{+,1} + e_{+,2} &= 0 \\ e_{+,1} - e_{+,2} &= 0 \;. \end{aligned} \tag{9.186}$$

These two equations are redundant; any pair of numbers for which $e_{+,1} = e_{+,2}$ is a solution to the system. However, since we want to construct an orthonormal basis from $\mathbf{e}_\pm$, we'll choose $e_{+,1} = e_{+,2} = 1/\sqrt{2}$ so that $\mathbf{e}_+$ is normalized:

$$||\mathbf{e}_+|| \;=\; \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2} \;=\; 1 \;. \tag{9.187}$$

Similarly, for $\mathbf{x}_-$, we have

$$\left(\begin{array}{cc} -2+3 & 1 \\ 1 & -2+3 \end{array} \right) \left(\begin{array}{c} e_{-,1} \\ e_{-,2} \end{array} \right) \;=\; 0 \;, \tag{9.188}$$

which corresponds to the system of equations

$$\begin{aligned} e_{-,1} + e_{-,2} &= 0 \\ e_{-,1} + e_{-,2} &= 0 \;. \end{aligned} \tag{9.189}$$

These equations are once again redundant; any pair of numbers for which $e_{-,1} = -e_{-,2}$ is a solution. We'll choose $e_{-,1} = 1/\sqrt{2}$ and $e_{-,2} = -1/\sqrt{2}$ so that we have $||\mathbf{e}_-|| = 1$. In summary, then, the normalized eigenvectors which correspond to the eigenvalues in Eq. (9.190) are

$$\mathbf{e}_+ \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\;, \qquad \mathbf{e}_- \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{9.190}$$

with respect to the original basis in Eq. (9.176).

Now that we've found the eigenbasis for $\mathbf{M}$, let's see what the equation of motion in Eq. (9.178) looks like in this new basis. Ww know that the configuration vector $\mathbf{x}'$ in the new basis in Eq. (9.190) and the configuration vector $\mathbf{x}$ in the original basis in Eq. (9.176) are related by $\mathbf{x} = \mathbf{U}^\dagger\mathbf{x}'$, where $\mathbf{U}$ is a unitary matrix. The elements of $\mathbf{U}$ are given by Eq. (9.138), so we have

$$\mathbf{U} \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\;. \tag{9.191}$$

Replacing $\mathbf{x}$ with $\mathbf{U}\mathbf{x}'$ in Eq. (9.178) gives us

$$m\frac{d^2}{dt^2}\big(\mathbf{U}^\dagger\mathbf{x}'\big) \;=\; k\mathbf{M}\mathbf{U}^\dagger\mathbf{x}'\;. \tag{9.192}$$

The elements of $\mathbf{U}^\dagger$ are all just constants, so we can pull this matrix through the derivative operator:

$$m\mathbf{U}^\dagger\frac{d^2}{dt^2}\mathbf{x}' \;=\; k\mathbf{M}\mathbf{U}^\dagger\mathbf{x}'\;. \tag{9.193}$$

Multiplying both sides of this equation by $\mathbf{U}$, we have

$$m\mathbf{U}\mathbf{U}^\dagger\frac{d^2}{dt^2}\mathbf{x}' \;=\; m\frac{d^2}{dt^2}\mathbf{x}' \;=\; k\mathbf{U}\mathbf{M}\mathbf{U}^\dagger\mathbf{x}'\;. \tag{9.194}$$

The matrix $\mathbf{M}' \equiv \mathbf{U}\mathbf{M}\mathbf{U}^\dagger$ on the right side of Eq. (9.194) is just a diagonal matrix whose diagonal entries are λ_+ and λ_-:

$$\mathbf{M}' \;=\; \mathbf{U}\mathbf{M}\mathbf{U}^\dagger \;=\; \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \;=\; \begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}\;. \tag{9.195}$$

The expression $\mathbf{x}'$ for the configuration vector in this new basis can likewise be found by multiplying $\mathbf{x}$ by the transformation matrix $\mathbf{U}$:

$$\mathbf{x}' \;=\; \mathbf{U}\mathbf{x} \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix}\;. \tag{9.196}$$

Substituting the expressions in Eqs. (9.195) and (9.196) back into Eq. (9.194) and canceling the overall factor of $1/\sqrt{2}$ from both sides and using the notation $x_\pm = x_+ \pm x_-$ we defined back in Sect. 9.2, we find that the equation of motion for the configuration vector in our new basis is

$$\frac{d^2}{dt^2}\begin{pmatrix} x_+ \\ x_- \end{pmatrix} \;=\; \frac{k}{m}\begin{pmatrix} -1 & 0 \\ 0 & -3 \end{pmatrix}\begin{pmatrix} x_+ \\ x_- \end{pmatrix}\;. \;=\; \begin{pmatrix} -\omega_+^2 & 0 \\ 0 & -\omega_-^2 \end{pmatrix}\begin{pmatrix} x_+ \\ x_- \end{pmatrix}\;, \tag{9.197}$$

where ω_+ and ω_- are the frequencies we defined in Eq. (9.8) for the normal modes of the system.

Now let's pause a moment to reflect on what this equation means. First of all, the components of the configuration vector in this basis are nothing but the coordinates x_+ and x_- that we obtained in Sect. 9.2 by adding and subtracting x_1 and x_2! In other words, the two basis vectors $\mathbf{e}_+$ and $\mathbf{e}_-$ correspond to the two normal modes of oscillation for the system with frequencies ω_+ and ω_-, respectively.

Second, as anticipated, we see that Eq. (9.197) is diagonal in this basis. In other words, we have decoupled the differential equations for the different components of our configuration vector by diagonalizing M. The individual equations for x_+ and x_- can be read off from Eqn. 9.197:

$$\begin{aligned} \frac{d^2x_+}{dt^2} &= -\omega_+^2 x_+ \\ \frac{d^2x_-}{dt^2} &= -\omega_-^2 x_- \end{aligned} \tag{9.198}$$

The solutions are those given in Eq. (9.7) of course, but the point is that we could have obtained these solutions via this method had we not known them already.

Third and finally, we emphasize again that the diagonalization method that we used here to decouple the equations of motion for our coupled-oscillator system is applicable to a broad variety of coupled systems.

Example: More Modes

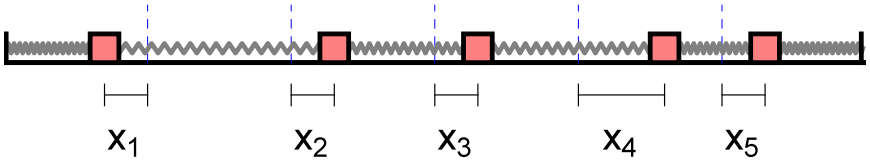


Figure 9.9: A coupled-oscillator system consisting of five blocks, each of mass m, connected to each other and to a pair of fixed walls by a set of springs, each with spring constant k.

We've now seen how to solve a system comprising two coupled oscillators — but why stop at two? For example, let's consider a coupled-oscillator system consisting of five blocks coupled together by a set of springs, as shown in Fig. 9.9. The system can be characterized by a set of position coordinates x_n, where $n = 1, 2, \ldots, 5$, each of which represents the displacement of a different block away from its equilibrium position. The force F_n on any particular block receives a contribution from both the spring on its left and the spring on its right. Thus, Newton's Second Law yields an equation of motion

$$\begin{aligned} m\frac{d^2x_i}{dt^2} \;=\; F_n \;&=\; -k(x_i - x_{i-i}) + k(x_{n+1}) - x_n) \\ &=\; -2kx_n + kx_{n+1} + kx_{n-1} \end{aligned} \tag{9.199}$$

for each block. These five coupled equations can be combined into the single matrix equation

$$\frac{d^2}{dt^2}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \;=\; \omega_0^2 \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} , \tag{9.200}$$

where we have defined $\omega_0 \equiv \sqrt{k/m}$. Solving the system is therefore a matter of diagonalizing the matrix

$$\mathbf{M} \;=\; \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix} . \tag{9.201}$$

Once again, just as we did for the two-oscillator systems, we begin by finding the eigenvalues of M. The characteristic equation is

$$\det\left(\mathbf{M} - \lambda \mathbf{I}_{5\times 5}\right) \;=\; \begin{vmatrix} -2-\lambda & 1 & 0 & 0 & 0 \\ 1 & -2-\lambda & 1 & 0 & 0 \\ 0 & 1 & -2-\lambda & 1 & 0 \\ 0 & 0 & 1 & -2-\lambda & 1 \\ 0 & 0 & 0 & 1 & -2-\lambda \end{vmatrix} \;=\; 0 \; . \tag{9.202}$$

This determinant is slightly more cumbersome to solve than the determinant of a 2×2 matrix, but it can be evaluated using the method of relative minors nonetheless. The resulting equation involves a quintic (*i.e.*, fifth-degree) polynomial:

$$\lambda^5 + 10\lambda^4 + 36\lambda^3 + 56\lambda^2 + 35\lambda + 6 \;=\; 0 \; . \tag{9.203}$$

The solutions to this equation are

$$\begin{aligned} \lambda_1 &= -2+\sqrt{3} \\ \lambda_2 &= -1 \\ \lambda_3 &= -2 \\ \lambda_4 &= -3 \\ \lambda_5 &= -2-\sqrt{3}\,. \end{aligned} \tag{9.204}$$

The corresponding normalized eigenvectors $\mathbf{e}_n$ are the solutions to the equation

$$\big(\mathbf{M}-\lambda_n\mathbf{I}_{5\times 5}\big)\mathbf{e}_n \;=\; 0 \tag{9.205}$$

with $||\mathbf{e}_n|| = 1$. Physically, the $\mathbf{e}_n$ represent the normal modes of the system — *i.e.*, the independent ways in which the oscillators in the system can be excited. Expressed in terms of their components in the original basis, these eigenvectors turn out to be

$$\mathbf{e}'_1 \;=\; \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 \\ \sqrt{3} \\ 2 \\ \sqrt{3} \\ 1 \end{pmatrix}\,, \qquad \mathbf{e}'_2 \;=\; \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}\,, \qquad \mathbf{e}'_3 \;=\; \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

$$\mathbf{e}'_4 \;=\; \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}\,, \qquad \mathbf{e}'_5 \;=\; \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 \\ -\sqrt{3} \\ 2 \\ -\sqrt{3} \\ 1 \end{pmatrix}\,. \tag{9.206}$$

It's important to keep in mind what the components of these basis vectors represent. The jth component $e'_{n,j}$ of the vector $\mathbf{e}'_n$ tells us proportionally how large the amplitude of oscillation for a particular block is when the corresponding mode of oscillation is excited. For example, if only the $n = 2$ mode (*i.e.*, the mode corresponding to the basis vector $\mathbf{e}'_2$) is excited, for example, the leftmost two blocks in Fig. 9.9 oscillate back and forth in unison with equal amplitude. The rightmost two blocks also oscillate in unison with the same amplitude, but they're a half-cycle out of phase with the blocks on the left. The block at the center remains fixed at its equilibrium position. The set of relative oscillation amplitudes associated with each normal mode of the system is illustrated in Fig. 9.10.

9.11 From Oscillations to Waves

Now let us take our analysis of coupled-oscillator systems one step further still. In particular, we'd like to be able to generalize the results we obtained in the previous section to the more general case of a system consisting of some arbitrary number N of coupled oscillators. If we know how to do this, we will be equipped to study the oscillations of vast numbers of coupled oscillators.

But how often do we really encounter a system consisting of a vast number of coupled oscillators in nature? Actually, we encounter them all the time! In Chapter 2, we saw that the interatomic forces in a molecule could in a wide variety of cases be modeled (for small deviations from equilibrium) as simple harmonic oscillators via the harmonic approximation. Thus, we can think of a long polyatomic molecule as a system of coupled oscillators, as shown in the left panel of Fig. 9.11. Likewise, a crystal whose atoms are arranged in a rectangular lattice, as shown in the right panel of Fig. 9.11, can also be modeled as a three-dimensional grid of coupled oscillators. For this reason, we will in this section refer to the oscillation objects as "particles" rather than "blocks." However, we emphasize that the results in this section apply to any system of oscillators which are coupled together in a mathematically analogous manner — be they block on springs or atoms in a lattice.

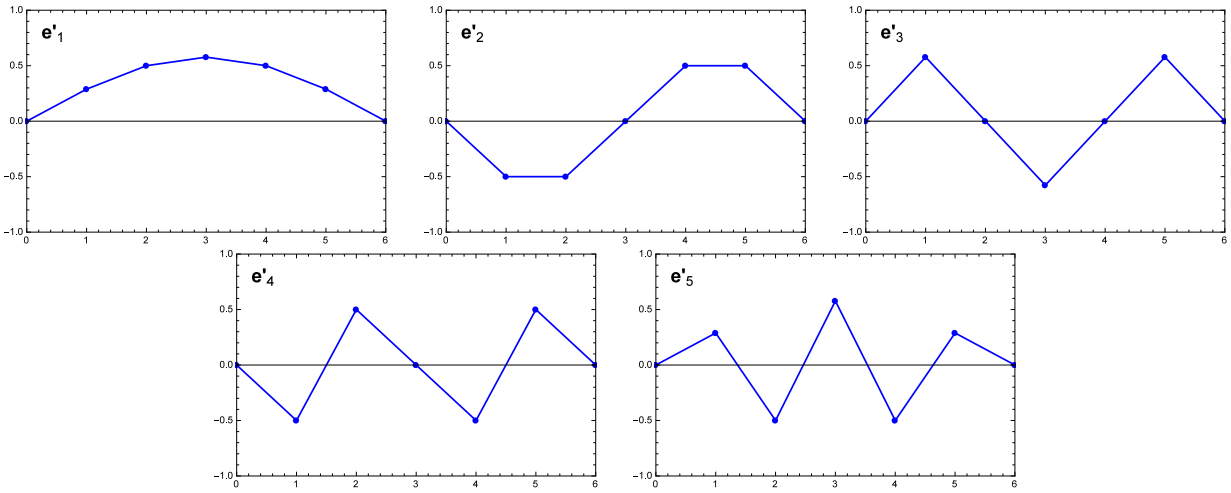


Figure 9.10: The five normal modes $\mathbf{e}'_n$ of a system consisting of a chain of five coupled oscillators. The dots in each panel indicate the components $e'_{n,j}$ of the corresponding $\mathbf{e}'_n$ with respect to the original basis of the individual oscillators. In particular, the horizontal position of each dot indicates the label j of the block and the vertical position indicates the value of $e_{n,j}$ for that block. The magnitude of $e'_{n,j}$ reflects the relative amplitude of oscillation of the jth block when the corresponding mode is excited.

The mathematical description of a system of N coupled oscillators is easy to generalize from the five-oscillator system we studied in Sect. 9.10. We begin by labeling each of the particles in our system by an index $j = 1, 2, \ldots, N$ which labels the particles in order in order of increasing distance away from the left "wall." The total force F_j on any particular particle once again receives a contribution from both the "spring" on its left and the "spring" on its right. Thus, Newton's Second Law yields an equation of motion

$$\begin{aligned} m\frac{d^2x_i}{dt^2} &= F_n &= -k(x_i - x_{i-i}) + k(x_{n+1}) - x_n) \\ &&= -2kx_n + kx_{n+1} + kx_{n-1} \end{aligned} \tag{9.207}$$

These N equations can be combined into a single matrix equation of the form

$$\frac{d^2}{dt^2}\mathbf{x} = \omega_0^2 \mathbf{M}\mathbf{x} \, , \tag{9.208}$$

where the $N \times N$ matrix $\mathbf{M}$ appearing in this equation is

$$\mathbf{M} = \begin{pmatrix} -2 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & -2 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & -2 & 1 & \ldots & 0 & 0 \\ 0 & 0 & 1 & -2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & -2 & 1 \\ 0 & 0 & 0 & 0 & \ldots & 1 & -2 \end{pmatrix} . \tag{9.209}$$

It turns out that for a set of N coupled oscillators, the components $e_{n,j}$ of the eigenvector $\mathbf{e}_n$ are given by

$$e'_{n,j} = C_n \sin\left(\frac{nj\pi}{N+1}\right) \, , \tag{9.210}$$

where C_n is an overall normalization constant. You can verify for yourself (see Problem 12) that the components of the eigenvectors given in Eq. (9.206) for a coupled system of five oscillators satisfy this

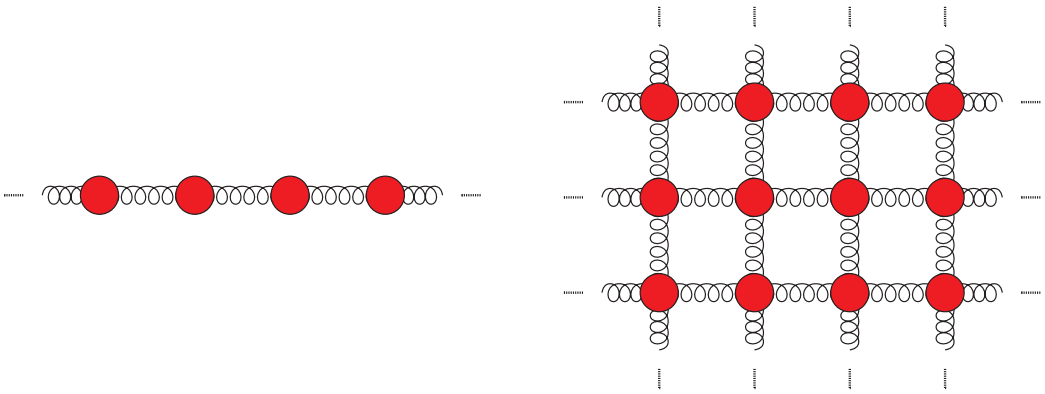


Figure 9.11: A linear molecule (left panel) can be modeled as a set of coupled oscillators when the interatomic forces can be modeled via the harmonic approximation. A crystal with a rectangular lattice structure (right panel) can likewise be modeled as a three-dimensional grid of coupled oscillators.

relation. As for the eigenvalues λ_n associated with these modes, we note that the eigenvector equation

$$\mathbf{M}\mathbf{e}_n \;=\; \lambda_n \mathbf{e}_n \tag{9.211}$$

takes the explicit form

$$\begin{pmatrix} -2 & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & -2 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 \\ 0 & 0 & 1 & -2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -2 & 1 \\ 0 & 0 & 0 & 0 & \dots & 1 & -2 \end{pmatrix} \begin{pmatrix} e_{n,1} \\ e_{n,2} \\ e_{n,3} \\ \vdots \\ e_{n,N-1} \\ e_{n,N} \end{pmatrix} \;=\; \lambda \begin{pmatrix} e_{n,1} \\ e_{n,2} \\ e_{n,3} \\ \vdots \\ e_{n,N-1} \\ e_{n,N} \end{pmatrix} . \tag{9.212}$$

This matrix relation implies that the various components of $\mathbf{e}_n$ are related to each other by a set of N individual equations. For $j = 0$ and $j = N$ the equations are respectively

$$\begin{aligned} e_{n,2} - 2e_{n,1} &= \lambda_n e_{n,1} \\ -2e_{n,N} + e_{n,N-1} &= \lambda_n e_{n,N} \; . \end{aligned} \tag{9.213}$$

For all other values of j, the corresponding equation has the general form

$$e_{n,j+1} - 2e_{n,j} + e_{n,j-1} \;=\; \lambda_n e_{n,j} \; . \tag{9.214}$$

Solving this equation for λ_n and plugging in the expression in Eq. (9.210) for $e_{n,j}$ into the resulting equation gives us

$$\begin{aligned} \lambda_n &= \frac{e_{n,j+1} + e_{n,j-1}}{e_{n,j}} - 2 \\ &= \frac{\sin\left[\frac{n(j+1)\pi}{N+1}\right] + \sin\left[\frac{n(j-1)\pi}{N+1}\right]}{\sin\left(\frac{nj\pi}{N+1}\right)} - 2 \\ &= \frac{\sin\left(\frac{nj\pi}{N+1} + \frac{n\pi}{N+1}\right) + \sin\left(\frac{nj\pi}{N+1} - \frac{n\pi}{N+1}\right)}{\sin\left(\frac{nj\pi}{N+1}\right)} - 2 \end{aligned} \tag{9.215}$$

We can simplify this relation further by using the trigonometric identity

$$\sin\alpha\cos\beta \;=\; \frac{1}{2}\Big[\sin(\alpha-\beta)+\sin(\alpha+\beta)\Big]\;. \tag{9.216}$$

The sine terms in the numerator and denominator of the resulting expression cancel, and we find that

$$\begin{aligned}\lambda_n &= 2\cos\left(\frac{n\pi}{N+1}\right)-2\\ &= 4\cos^2\left[\frac{n\pi}{2(N+1)}\right]-4\\ &= -4\sin^2\left[\frac{n\pi}{2(N+1)}\right]\;,\end{aligned} \tag{9.217}$$

where in going from the first to the second line, we have used the identity $\cos(2\theta) = 2\cos^2\theta - 1$, and in going from the second to the third, we have used $\cos^2\theta + \sin^2\theta = 1$. These are the eigenvalues associated with our set of N normal modes. Moreover, Eq. (9.208) tells us that the relation between each λ_n and the corresponding oscillator frequency ω_n is given by

$$\omega^2 = -\lambda\omega_0^2\;. \tag{9.218}$$

Thus, the normal-mode frequencies are

$$\omega_n \;=\; \omega_0\sqrt{-\lambda} \;=\; 2\omega_0\left|\sin\left[\frac{n\pi}{2(N+1)}\right]\right|\;. \tag{9.219}$$

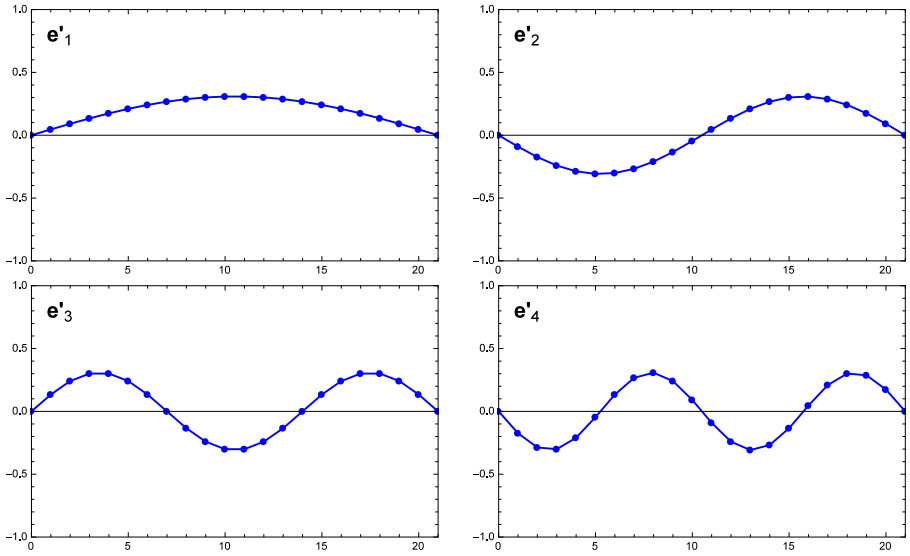


Figure 9.12: The first four normal modes $\mathbf{e}'_n$ of a system consisting of a chain of $N = 20$ coupled oscillators. Just as in Fig. 9.10, the dots in each panel indicate the components $e'_{n,j}$ of the corresponding $\mathbf{e}'_n$ with respect to the original basis. The horizontal position of each dot indicates the label j of the particle and the vertical position indicates the value of $e_{n,j}$.

In Fig. 9.12, we plot the set of $e_{n,j}$ as a function of the particle index j. As discussed in Sect. 9.10, $e_{n,j}$ indicates the relative amplitude with which block j oscillates when the normal mode e_n is excited. Another way of saying this is that each normal mode $\mathbf{e}_n$ of the system corresponds to a different collective excitation of its constituent particles. The patterns of $e_{n,j}$ values shown in each of the four panels of Fig. 9.12 clearly resemble sine waves — and indeed, waves they are! Specifically, these collective excitations correspond to **standing compressional waves** (*i.e.*, longitudinal waves).

Problems

1. Find the normal frequencies $\omega_\pm$ of the normal modes for a coupled-oscillator system which is identical to the one shown in Fig. 9.1 except that the spring constant k' of the central spring has a different value from the spring constant k of each of the other two springs.

2. Find the values of the constants A_+, A_-, ϕ_+, and ϕ_- in Eq. (9.7) for the initial conditions $x_1(0) = x_2(0) = 0$ with $v_1(0) = v_0 \neq 0$ and $v_2(0) = 0$. Use Mathematica to make a plot of the coordinates $x_1(t)$ and $x_2(t)$ of the two blocks as functions of time for $k/m = 1$ with $v_0 = 5$ (in arbitrary units).

3. Determine whether each of the following collections of objects constitutes a vector space. The collection of objects listed are the would-be vectors in the space. You may assume that the associated set of scalars is the set of real numbers in each case.

   (a) The set of all real functions $f(x)$ for which that are $f(x) = 1$ at $x = 0$.

   (b) The set of all periodic functions with period π.

   (c) The set of all vectors of the form $z\hat{\mathbf{x}} + z\hat{\mathbf{y}}$, where z is a complex number.

   (d) The set of all vectors of the form $z\hat{\mathbf{x}} + 2z\hat{\mathbf{y}}$, where z is a complex number.

   (e) The set of all polynomials of degree less than or equal to 5 — *i.e.*, the set of polynomials $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5$, where the a_n are constants.

   (f) The set of all polynomials of degree less than or equal to 4 with $a_1 = a_3$.

4. For each of the items (a) − (f) in Problem. 3 which you determined to be a vector space, give the dimension of the space.

5. In each of the following cases, determine whether the set of vectors given form a basis in the space.

   (a) The functions $2x^2 + 3$, $(x-1)^2$, and x^2 in the function space of quadratic polynomials.

   (b) The vectors in the vector space $\mathbb{R}^3$ whose component-wise descriptions in the standard $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ basis are

$$\mathbf{a}_1 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \qquad \mathbf{a}_2 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \qquad \mathbf{a}_3 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}. \tag{9.220}$$

   (c) The functions $x+1$, x^2+x, x^3+x^2, and x^4+x^3 in the function space of quartic (*i.e.*, 4th-order) polynomials.

6. Evaluate the product $\mathbf{BA}$ of the two matrices in Eq. (9.110) in reverse order.

7. For the pair of square matrices $\mathbf{A}$ and $\mathbf{B}$ given below

$$\mathbf{A} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} -1 & 4 \\ 0 & 2 \end{pmatrix}, \tag{9.221}$$

   evaluate the following:

   (a) The products $\mathbf{AB}$ and $\mathbf{BA}$

   (b) The products $\mathbf{A}^2 = \mathbf{AA}$ and $\mathbf{B}^2 = \mathbf{BB}$

   (c) Their sum $\mathbf{A} + \mathbf{B}$ and difference $\mathbf{A} - \mathbf{B}$

8. The **commutator** of two square matrices $\mathbf{A}$ and $\mathbf{B}$, usually denoted $[\mathbf{A}, \mathbf{B}]$, is defined to be the difference between the product of the matrices and their product in the reverse order:

$$[\mathbf{A}, \mathbf{B}] \equiv \mathbf{AB} - \mathbf{BA}\,. \tag{9.222}$$

Evaluate the commutators $[\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2]$, $[\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_3]$, and $[\boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3]$ for the matrices

$$\boldsymbol{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.223}$$

These three matrices are called the **Pauli matrices**, and their commutators play an important role in the quantum-mechanical description of angular momentum.

9. Given the square matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 5i \\ -2i & 2 & 0 \\ 1 & 1+i & 0 \end{pmatrix}, \tag{9.224}$$

find all of the following:

   (a) its transpose $\mathbf{A}^T$

   (b) its complex conjugate $\mathbf{A}^*$

   (c) its Hermitian conjugate $\mathbf{A}^\dagger$

   (d) its trace $\mathrm{Tr}(\mathbf{A})$

   (e) its inverse $\mathbf{A}^{-1}$

10. Evaluate the exponential $e^{\mathbf{A}}$ of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{9.225}$$

Express your answer in the simplest, most compact form possible, without any matrix products or infinite sums.

11. Consider the 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}. \tag{9.226}$$

   (a) Find the eigenvalues λ_1 and λ_2 of this matrix.

   (b) Find the corresponding eigenvectors $\mathbf{e}'_1$ and $\mathbf{e}'_2$.

   (c) Find the unitary matrix $\mathbf{U}$ which diagonalizes $\mathbf{A}$ and verify that $\mathbf{UAU}^\dagger$ is indeed diagonal, with diagonal entries equal to the eigenvalues of $\mathbf{A}$.

12. Verify that the general formula in Eq. (9.210) for the components $e'_{n,j}$ of the eigenvectors for a system of N coupled oscillators is satisfied for the components of the vectors $\mathbf{e}'_1$ and $\mathbf{e}'_2$ given in Eq. (9.206) for $N = 5$.

Index