Imperial College London

Oscillations and Waves

First Year Physics, 2022-2023

Mike Tarbutt

Last Updated: December 14, 2022

Contents

This course

Aims

By the end of this course, I hope you will be able to:

- Explain the key phenomena associated with oscillations and waves;
- Use appropriate mathematical techniques to describe these phenomena;
- Appreciate that oscillations and waves are ubiquitous in nature and give diverse examples of systems that can be understood as oscillators or as waves.

Resources

These lecture notes are intended to be self-contained and sufficient to cover all the material of the course. The lectures will follow the lecture notes, but the notes will often contain more detail – not everything that's in the notes will be covered in lectures.

There are three problem sets and one APS. You will not learn anything unless you consolidate by doing the problems. Please do the problems. I will make the solutions available a little while after releasing the problems, but you should not just look at the solutions – you must do the problems yourselves. Struggling with them is the biggest part of learning.

There are many books on this topic and you can learn a lot from the books. If the problem sets are not challenging enough, the books provide many more problems. Some recommendations:

- *Feynman Lectures Vol. I*, especially chapters 21, 23, 24, 29, 47, 48, 49, 50.
- *Introduction to Vibrations and Waves* , Pain and Rankin.
- *Vibrations and Waves*, French.
- *A Student's Guide to Waves*, Fleisch and Kinnaman

I will have both in-person and remote office hours – see Blackboard for schedule. My office is Blackett 207. You are welcome to cone along to discuss the topics covered in the course. You can also email me – m.tarbutt@imperial.ac.uk – it may take a few days but I will always try to answer.

Chapter 1

Harmonic oscillators I

Harmonic oscillators are ubiquitous in physics. In this course, we will mainly look at mechanical examples such as pendulums and springs, as well as some examples in electrical circuits, but they are also found in acoustics, chemical processes, atoms interacting with light, particles in traps, most quantum systems and a huge variety of other phenomena.

1.1 Mass on a spring

We first consider a simple example of a mass m on a spring, as illustrated in figure 1.1(a). It exhibits the defining features of a harmonic oscillator:

- There is an equilibrium position where the force is zero.
- When displaced from equilibrium, there is a restoring force which is linear in the displacement.

Let x be the displacement from equilibrium. The force must be $F = -\kappa x$ – linear in x as required, opposite to x so that it brings the mass towards the equilibrium position, and zero at $x = 0$. The constant of proportionality, κ, is known as the spring constant. Using Newton's second law, we get the equation of motion:

$$\frac{d^2x}{dt^2} = -\frac{\kappa}{m}x. \tag{1.1}$$

This is a second-order differential equation describing the motion.

1.2 Pendulum

Our next example is the pendulum illustrated in figure 1.1(b). The pendulum is a mass m connected to a fixed point by a string of length l. The angle to the vertical is θ, and we will assume that the amplitude of the swing is small so that we can make a small angle approximation. The mass traces out an arc of length $s = l\theta$. The tangential component of the force (in the direction of s) is $F_s = -mg\sin\theta \approx -mg\theta$. The tension in the string is perpendicular to F_s so does not contribute to this component. Applying Newton's second law to the motion in the direction of s we have

$$m\frac{d^2s}{dt^2} = -mg\theta.$$

Using $s = l\theta$ and remembering that l is a constant, gives

$$\frac{d^2\theta}{dt^2} = -\frac{g}{l}\theta. \tag{1.2}$$

We see that the differential equation describing the pendulum [equation (1.2)] has the same form as the one describing the spring [equation (1.1)]. The variable that describes the motion has changed (θ instead of x) and the coefficient is different (g/l instead of κ/m), but the form is the same.

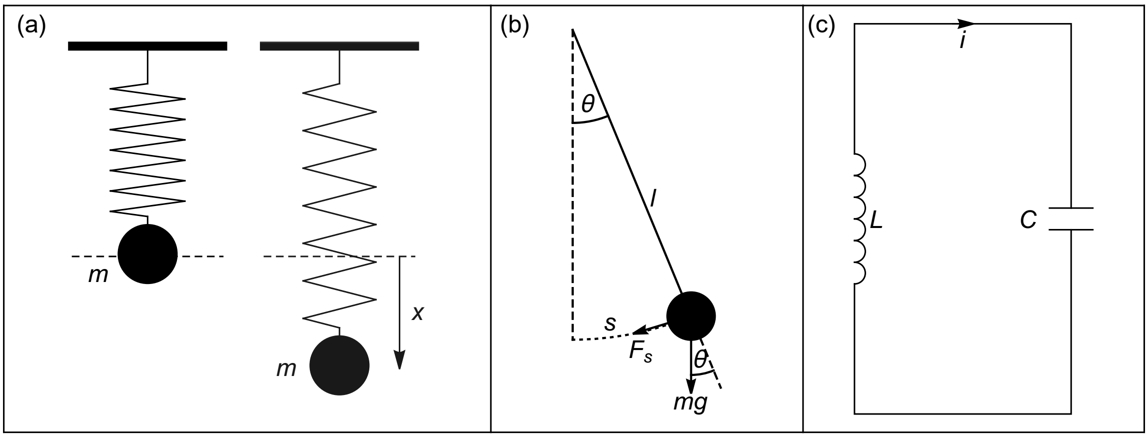

Figure 1.1: Examples of harmonic oscillators. (a) Mass on a spring, at equilibrium (left) and displaced by x (right). (b) Pendulum of length l and mass m. (c) Electrical circuit containing an inductor L and a capacitor C (LC circuit).

1.3 LC circuit

Our third example is an electrical circuit consisting of an inductor L and a capacitor C connected together as shown in figure 1.1(c). We suppose that the capacitor was initially charged, and then the inductor was connected across it so that the charge can flow from one capacitor plate to the other, producing a current i. Let the charge on the capacitor be q. Then the voltages across the capacitor and inductor are

$$V_C = \frac{q}{C},$$
$$V_L = L\frac{di}{dt}.$$

Since there is no voltage source in the circuit, the total voltage around the circuit must be zero, $V_L = -V_C$, so we have

$$L\frac{di}{dt} = -\frac{1}{C}q. \tag{1.3}$$

The current is the rate of change of charge, $i = dq/dt$. Using this we get,

$$\frac{d^2q}{dt^2} = -\frac{1}{LC}q. \tag{1.4}$$

Alternatively, we could differentiate both sides of equation (1.3), and again use $i = dq/dt$ to reach

$$\frac{d^2i}{dt^2} = -\frac{1}{LC}i. \tag{1.5}$$

We see that even though the system seems very different from a spring or a pendulum, it is described by a differential equation of exactly the same form. This time, the variable is the current i (or charge q) and the coefficient is $1/(LC)$.

1.4 Harmonic oscillators everywhere

The property of a harmonic oscillator is that the restoring force is linear in the displacement. Although oscillators often become non-linear when the displacement is large, they are usually harmonic for small displacements. Why is that?

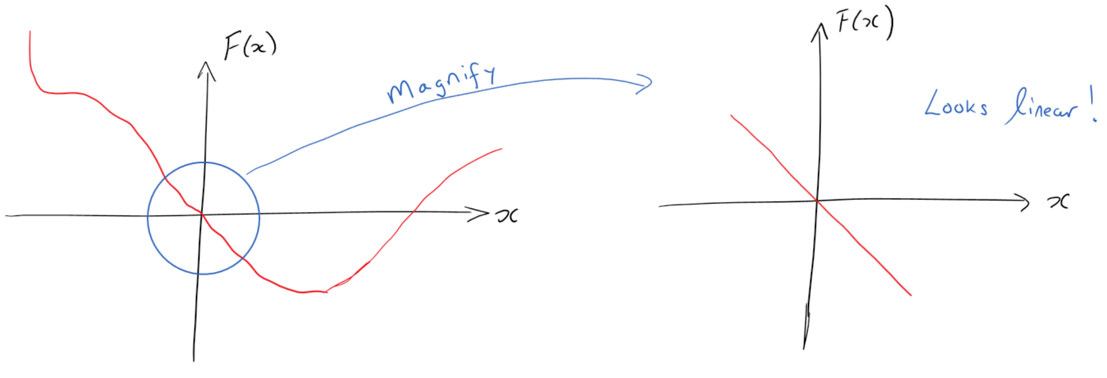

Figure 1.2: Force functions that are zero at the origin (equilibrium) and antisymmetric about the origin (restoring) tend to look linear near the origin.

To be an oscillator at all, there must be an equilibrium point (we'll put the origin here) where the force is zero, and there must be a restoring force towards that equilibrium - around the origin the force should be positive on the negative side and negative on the positive side. Try drawing a smooth, continuous force curve with these properties. You'll find yourself drawing something that is linear close to the origin. There's really no other choice! Figure 1.2 is my attempt.

More formally, we can make a Taylor expansion of the force around the equilibrium point:

$$F(x) = F(x=0) + \left(\frac{dF}{dx}\right)_{x=0} x + \frac{1}{2}\left(\frac{d^2F}{dx^2}\right)_{x=0} x^2 + ... \tag{1.6}$$

The first term is zero because the force must vanish at the equilibrium point. The second term is the linear term that gives rise to harmonic oscillation. For small enough values of x, the third and successive terms must be smaller than the first term - so the linear term dominates as long as x is small enough - small oscillations are harmonic.

1.5 Summary of key points

- The general differential equation describing harmonic motion is $\frac{d^2\psi}{dt^2} = -\omega_0^2\psi$.
- This equation describes all kinds of oscillators. ψ is the variable describing the thing that oscillates. The expression for ω_0 depends on the system. For a mass on a spring ψ is the displacement x and $\omega_0 = \sqrt{\kappa/m}$; for a pendulum ψ is the angle θ and $\omega_0 = \sqrt{g/l}$; for an LC circuit ψ is the current i and $\omega_0 = 1/\sqrt{LC}$.
- Any system where there is a restoring force towards an equilibrium point tends to be a harmonic oscillator for small enough disturbances from equilibrium.

Chapter 2

Harmonic oscillators II

In the last lecture, we looked at some examples of harmonic oscillators and saw that they all obey the same kind of equation – equations (1.1), (1.2) and (1.5) all have the same form. Once we know the solution to one of them, we know it for all of them, and indeed for any other system that is described by a differential equation of this form.

2.1 Solution of harmonic oscillator equation

Let's take the example of the mass on the spring, equation (1.1). We guess a trial solution

$$x = a\cos(\omega_0 t). \tag{2.1}$$

The derivatives with respect to time are $\dot{x} = -a\omega_0 \sin(\omega_0 t)$ and $\ddot{x} = -a\omega_0^2 \cos(\omega_0 t) = -\omega_0^2 x$. Substituting this into equation (1.1) we have $\ddot{x} = -\omega_0^2 x = -(\kappa/m)x$. So we see that equation (2.1) is indeed a solution of (1.1) provided

$$\omega_0 = \sqrt{\frac{\kappa}{m}}. \tag{2.2}$$

Note that we could equally well have used $x = b\sin(\omega_0 t)$ as our trial solution and found the same conclusion. Thus, the more general solution is

$$x = a\cos(\omega_0 t) + b\sin(\omega_0 t), \tag{2.3}$$

where a and b are constants that will be determined by the initial conditions (the values of x and $\dot{x}$ at $t = 0$). Alternatively, the solution can be written in the form

$$x = A\cos(\omega_0 t + \phi), \tag{2.4}$$

where A and ϕ are determined by the initial conditions. Using standard trigonometric identities, we see that forms (2.3) and (2.4) are related by $a = A\cos\phi$ and $b = -A\sin\phi$.

To see how the initial conditions fix A and ϕ, let's take an example where at $t = 0$ the mass is at rest and is displaced by x_0. Differentiating equation (2.4) gives $\dot{x} = -A\omega_0 \sin(\omega_0 t + \phi)$ and thus $\dot{x}(t = 0) = -A\omega_0 \sin(\phi) = 0$, giving $\phi = 0$. Also, $x(t = 0) = x_0 = A\cos\phi = A$. So, for this example, $A = x_0$ and $\phi = 0$.

Figure 2.1 is a plot of equation (2.4). We note the following important points:

1. ω_0 is called the *angular frequency*, or sometimes the *natural frequency* because it's the angular frequency that the system naturally oscillates at. The unit of ω_0 is rad/s. It is related to the oscillation period T as $T = 2\pi/\omega_0$.

2. A is called the *amplitude* of oscillation. It is determined by the initial conditions.

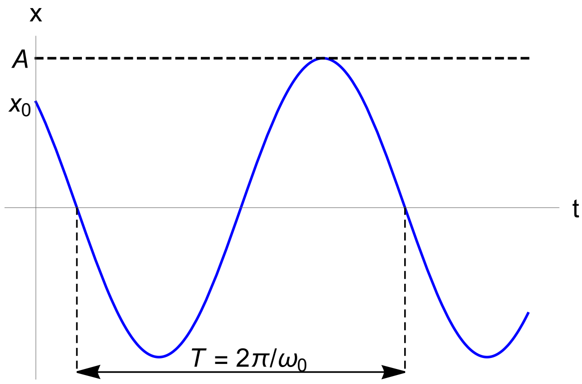

Figure 2.1: A plot of equation (2.4).

3. ϕ is called the *phase offset*. It is determined by the initial conditions. For example, $A\cos\phi$ is the displacement at $t = 0$, shown as x_0 in figure 2.1.

4. The angular frequency does not depend on the oscillation amplitude. This is a defining feature of harmonic oscillators – large oscillations take the same amount of time as small oscillations.[1]

5. ω_0 depends only on the coefficient in the differential equation, so can immediately be read off from the differential equation. For the mass on a spring, $\omega_0 = \sqrt{\kappa/m}$. For the pendulum, $\omega_0 = \sqrt{g/l}$. For the LC circuit, $\omega_0 = 1/\sqrt{LC}$.

Now, instead of taking a specific example, we can take the general case: a harmonic oscillator whose variable is ψ is described by the differential equation

$$\frac{d^2\psi}{dt^2} = -\omega_0^2\psi, \tag{2.5}$$

whose general solution is

$$\psi = A\cos(\omega_0 t + \phi). \tag{2.6}$$

2.2 Complex notation

Consider the complex number $\tilde{x} = x + iy$ and suppose that it satisfies the equation

$$\ddot{\tilde{x}} = -\omega_0^2\tilde{x}. \tag{2.7}$$

The solution can be written as

$$\tilde{x} = \tilde{A}e^{i\omega_0 t}, \tag{2.8}$$

where $\tilde{A} = Ae^{i\phi}$ is a complex number. This is easy to see: $\dot{\tilde{x}} = i\omega_0\tilde{A}e^{i\omega_0 t}$ and $\ddot{\tilde{x}} = -\omega_0^2\tilde{A}e^{i\omega_0 t} = -\omega_0^2\tilde{x}$. Since the real part of equation (2.7) is the harmonic oscillator equation ($\ddot{x} = -\omega_0^2 x$), the real part of equation (2.8) must be its solution, i.e.

$$x = \mathrm{Re}\{\tilde{x}\} = \mathrm{Re}\{\tilde{A}e^{i\omega_0 t}\} = \mathrm{Re}\{Ae^{i\phi}e^{i\omega_0 t}\} = \mathrm{Re}\{Ae^{i(\omega_0 t+\phi)}\} = A\cos(\omega_0 t + \phi). \tag{2.9}$$

We've arrived at the same solution we already knew, but in a slightly different way - pretending that x is allowed to be complex, solving the differential equation in that case, and then taking the real part of the solution at the end. We will see later, when we come to look at driven, damped

[1] Of course, this can never be true indefinitely. If you stretch the spring too much it will break or be permanently deformed. If you swing the pendulum too much the small angle approximation no longer holds and the oscillator becomes non-linear. If there's too much current in the LC circuit, the capacitor will pop or the inductor will fry. Harmonic oscillation is limited to "small" oscillation where the meaning of small depends on the context.

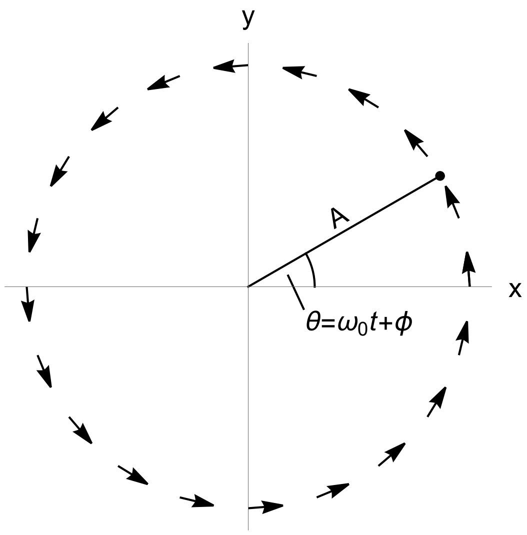

Figure 2.2: Complex number representation of the harmonic oscillator solution. The point in the complex plane goes around in a circle of radius A and angular frequency ω_0. The actual motion of the oscillator is the projection onto the x-axis.

oscillators, that this method make the mathematics much simpler to handle. The main reason is that in the form $x = Ae^{i\omega_0 t}$, all derivatives remain proportional to $e^{i\omega_0 t}$ and therefore proportional to x. That is not true for the cosine form of the solution where only the even-order derivatives[2] are proportional to x.

2.3 Connection with circular motion

The complex number representation introduced above is not just helpful mathematically; it also gives us a nice visual representation and a beautiful analogy between harmonic motion and circular motion. Our complex number $\tilde{x} = x + iy$ can be represented as a point in a diagram that has x and y as the horizontal and vertical axes (an Argand diagram). This is shown in figure 2.2. Written in the form $\tilde{x} = re^{i\theta}$, the vector from the origin to the point has length r and makes an angle θ with the x-axis. Our harmonic oscillator solution in complex form, equation (2.8), is $\tilde{x} = Ae^{i(\omega_0 t+\phi)}$, so we can immediately identify $r = A$ and $\theta = \omega_0 t + \phi$. We see that the point in the diagram goes round in a circle of radius A at a uniform rate given by the angular frequency ω_0. The phase angle at $t = 0$ is ϕ, and the point completes one rotation every time $\omega_0 t$ increases by 2π, i.e. every time t increases by $2\pi/\omega_0$. The harmonic motion is the projection of the point onto the x-axis.

Indeed, the relation between harmonic oscillation and circular motion goes deeper. An object going around in a circle of radius R with angular velocity ω_0 has a linear velocity $v = R\omega_0$. It is kept in circular motion by an acceleration $a = -R\omega_0^2$, where the minus sign tells us that the acceleration is towards the centre. If θ is the angle made with the x-axis, the horizontal component of the acceleration is $a_x = -\omega_0^2 R\cos\theta$. Similarly, the horizontal coordinate of the object is $x = R\cos\theta$. Thus we have $a_x = \ddot{x} = -\omega_0^2 x$, which is the equation of motion for a harmonic oscillator. Circular motion projected onto the x-axis gives you harmonic oscillation.

2.4 Energy

Consider the potential energy U and kinetic energy K of our mechanical oscillator. We can set the zero of potential energy to be at the equilibrium point. Then, when the displacement is x

[2] $d^n x/dt^n$ with n even

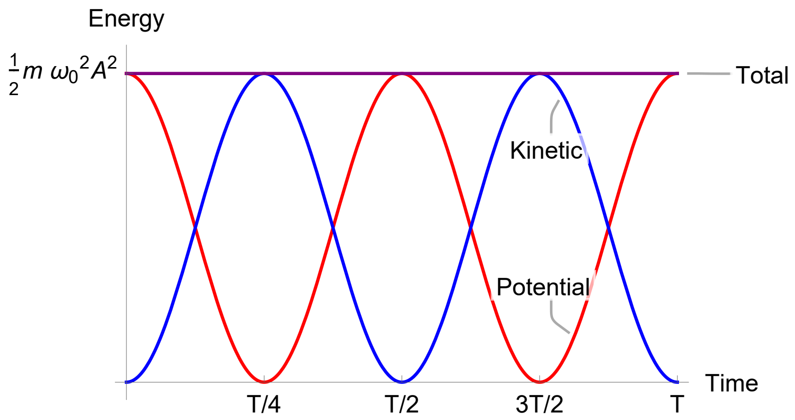

Figure 2.3: Potential, kinetic and total energy in a harmonic oscillator.

the potential energy is the negative of the work done in getting to x, which is the force times the displacement:

$$U = -\int_0^x F(x')dx' = \int_0^x \kappa x' dx' = \frac{1}{2}\kappa x^2 = \frac{1}{2}m\omega_0^2 x^2 = \frac{1}{2}m\omega_0^2 A^2 \cos^2(\omega_0 t + \phi), \tag{2.10}$$

where we have used equations (2.2) and (2.4). Noting that $\dot{x} = -\omega_0 A \sin(\omega_0 t + \phi)$, the kinetic energy is

$$K = \frac{1}{2}m\dot{x}^2 = \frac{1}{2}m\omega_0^2 A^2 \sin^2(\omega_0 t + \phi). \tag{2.11}$$

The total energy is

$$E = K + U = \frac{1}{2}m\omega_0^2 A^2 \left(\cos^2(\omega_0 t + \phi) + \sin^2(\omega_0 t + \phi)\right) = \frac{1}{2}m\omega_0^2 A^2. \tag{2.12}$$

The energy oscillates back and forth between potential and kinetic energy, with a period of $T/2$, but the total energy is constant as we would expect for an undamped oscillator. The energy is proportional to the square of the oscillation amplitude.

2.5 Summary of key points

- The harmonic oscillator equation $\frac{d^2\psi}{dt^2} = -\omega_0^2\psi$ has the general solution $\psi = A\cos(\omega_0 t + \phi)$.
- ω_0 is the angular frequency, A is the amplitude and ϕ is the phase offset. A and ϕ are determined by initial conditions.
- The angular frequency is independent of the oscillation amplitude.
- We very often use an equivalent complex solution $\tilde{\psi} = \tilde{A}e^{i\omega_0 t}$ where $\tilde{A} = Ae^{i\phi}$. The true solution is the real part of $\tilde{\psi}$.
- Circular motion projected onto an axis gives you harmonic oscillation.
- The energy of an oscillator is proportional to the square of the oscillation amplitude.
- The energy oscillates back and forth between kinetic energy and potential energy, but the total energy is a constant.

Chapter 3

Damped harmonic oscillator

3.1 Equation of motion

Real oscillators have some dissipation - the pendulum or spring will have some air resistance, electrical circuits contain resistors, the oscillating atom radiates its energy as light, and so on. For a mechanical system, the damping is often proportional to the velocity, and of course opposes the velocity, so we can write $F_{\mathrm{damp}} = -b\dot{x}$. Adding this to the restoring force, we get a total force $F = -\kappa x - b\dot{x}$ and a corresponding equation of motion

$$\ddot{x} = -\omega_0^2 x - \gamma\dot{x}, \tag{3.1}$$

where we have used the relation $\omega_0 = \sqrt{\kappa/m}$ and introduced the damping constant $\gamma = b/m$. The unit of γ is rad/s. Generalizing to any type of damped harmonic oscillator, we have the differential equation

$$\ddot{\psi} + \gamma\dot{\psi} + \omega_0^2\psi = 0. \tag{3.2}$$

3.2 Solution

We'll use our complex solution method by trying a complex solution of the form $\tilde{\psi} = \tilde{A}e^{i\omega t}$, and then taking the real part at the end. As previously, $\tilde{A} = Ae^{i\phi}$ is a complex number. Substituting into equation (3.2) we get

$$-\omega^2\tilde{\psi} + i\omega\gamma\tilde{\psi} + \omega_0^2\tilde{\psi} = 0.$$

So our trial solution works provided ω satisfies

$$\omega^2 - i\omega\gamma - \omega_0^2 = 0. \tag{3.3}$$

Solving the quadratic equation gives

$$\omega = i\frac{\gamma}{2} \pm \sqrt{\omega_0^2 - \left(\frac{\gamma}{2}\right)^2}. \tag{3.4}$$

We have a complex solution for ω. To study the solution, it is helpful to distinguish different regimes of damping.

3.3 Regimes of damping

3.3.1 Light damping

When $\gamma/2 < \omega_0$ we are in the light damping regime. In this case, we can define an angular frequency

$$\omega_{\mathrm{d}} = \sqrt{\omega_0^2 - (\gamma/2)^2}, \tag{3.5}$$

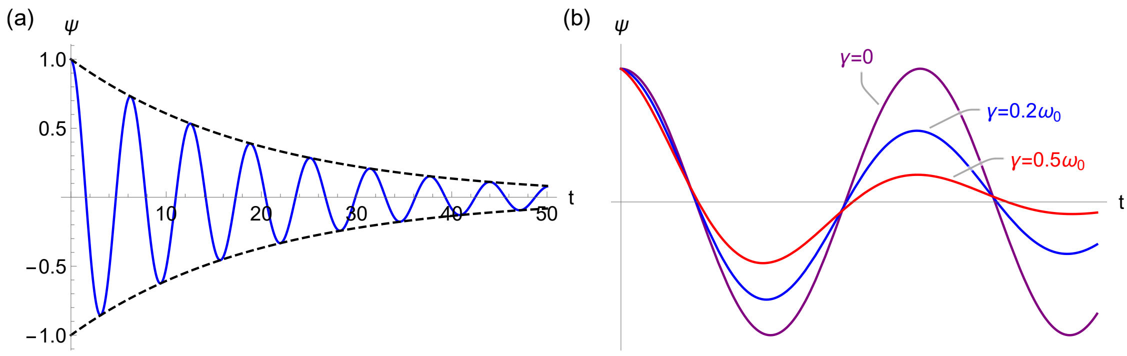

Figure 3.1: Harmonic oscillation with light damping. (a) $\omega_0 = 1$ rad/s and $\gamma = 0.1$ rad/s. The dashed lines show the envelope defined by $e^{-\gamma t/2}$. (b) Comparing oscillations with $\gamma = 0, 0.2, 0.5\omega_0$.

which is a real quantity, smaller than ω_0. Using this quantity, our solution is now

$$\tilde{\psi} = \tilde{A}e^{i(i\gamma/2 \pm \omega_\mathrm{d})t} = Ae^{i\phi}e^{-\gamma t/2}e^{\pm i\omega_\mathrm{d}t} = Ae^{-\gamma t/2}e^{i(\pm\omega_\mathrm{d}t+\phi)}.$$

Then, taking the real part, we obtain[1]

$$\psi = Ae^{-\gamma t/2}\cos(\omega_\mathrm{d}t + \phi). \tag{3.6}$$

The solution is plotted in figure 3.1. It is an oscillatory solution within an exponentially decaying envelope. The angular frequency is ω_d which is smaller than ω_0, though only slightly smaller when $\gamma \ll \omega_0$. The rate of decay of the oscillations is determined by γ - larger γ means a faster decay.

3.3.2 Heavy damping

When $\gamma/2 > \omega_0$, the square root in equation (3.4) is imaginary and so ω is pure imaginary. We can define

$$\gamma'/2 = \sqrt{(\gamma/2)^2 - \omega_0^2}$$

which is a real quantity, and

$$\gamma_\pm = \gamma \pm \gamma',$$

so that $\omega = i\gamma_\pm/2$. We have two possible values of ω, giving us two possible solutions. The general solution is the sum of the two solutions with coefficients that depend on the initial conditions:

$$\psi = Be^{-\gamma_+ t/2} + Ce^{-\gamma_- t/2}. \tag{3.7}$$

Both terms decay exponentially, one faster than $\gamma/2$ and one slower than $\gamma/2$. Figure 3.2 shows the values of $\gamma_\pm$ as a function of the ratio γ/ω_0. The two rates are equal when $\gamma/\omega_0 = 2$. As γ/ω_0 increase, γ_+ tends towards 2γ, and γ_- tends towards zero. At long times, the $e^{-\gamma_+ t/2}$ term will be nearly zero and the behaviour will be dominated by the slowly-decaying term, $e^{-\gamma_- t/2}$. In this limit, increasing γ makes the system decay more slowly towards equilibrium.

3.3.3 Critical damping

The particular case where $\gamma/2 = \omega_0$ is known as critical damping. In this case, $\gamma_+ = \gamma_- = \gamma$, and equation (3.7) reduces to $\psi = (B + C)e^{-\gamma t/2}$. Notice here that B and C are both multiplying the same thing, so are not independent constants at all (we could simply replace $B + C$ by a single

[1] Note that in the last step we have chosen the solution with $+\omega_\mathrm{d}$. The solution with $-\omega_\mathrm{d}$ is just a phase-shifted version which makes no difference to anything since the phase offset ϕ is still to be determined by the initial conditions.

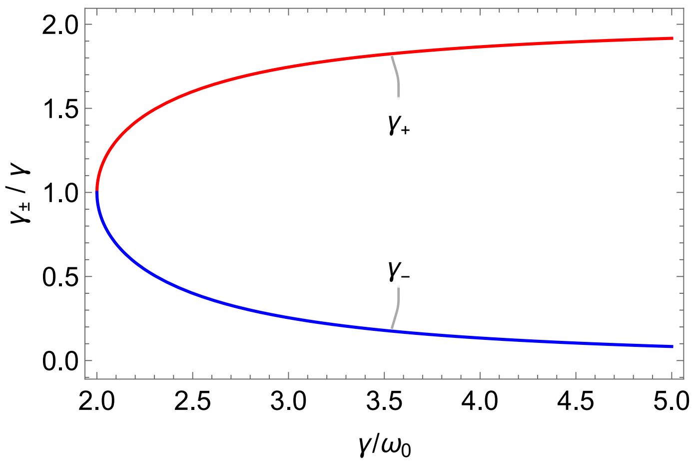

Figure 3.2: Values of the two damping rates, γ_+ and γ_- as a function of γ/ω_0. These are relevant only in the heavy damping limit where $\gamma/\omega_0 > 2$.

constant). That means that the solution is not complete in this special case. You will see in the problem set that the complete solution is

$$\psi = (A + Bt)e^{-\gamma t/2}. \tag{3.8}$$

Critical damping returns the system to equilibrium in the shortest possible time, so is quite a common choice for systems where the desired state is the equilibrium state.

3.4 Summary of key points

- For light (or weak) damping, the system oscillates at a frequency slightly below ω_0 and the oscillations damp away with a characteristic timescale of $\gamma/2$. Increasing the damping (γ) *decreases* the time taken to reach equilibrium.
- For heavy (or strong) damping, the system doesn't oscillate. Increasing the damping (γ) *increases* the time taken to reach equilibrium.
- At critical damping, the system doesn't oscillate but reaches equilibrium in the shortest possible time.

Chapter 4

Driven harmonic oscillator I

So far, we have looked at the free oscillations of various systems. We now turn to the case where the oscillator is driven periodically by some external force. This is a situation that occurs throughout physics and displays one of the most important phenomena in all of physics - *resonance*.

4.1 Equation of motion

We will study the mass on the spring again, but the phenomena we will find are common to all driven harmonic oscillators. The system is illustrated in figure 4.1. In addition to the force from the spring ($-\kappa x$) and the damping force ($-b\dot{x}$) that we had before, the mass is now being driven by a time-varying external force $F_0\cos(\omega t)$. Here, note carefully that the angular frequency of this external force (ω) is not necessarily the same as the natural frequency of the oscillator (ω_0). The total force acting on the mass is now $F = -\kappa x - b\dot{x} + F_0\cos(\omega t)$. The resulting equation of motion is

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = \frac{F_0}{m}\cos(\omega t). \tag{4.1}$$

Here, we have used the relations established previously: $\omega_0^2 = \kappa/m$ and $\gamma = b/m$.

4.2 Solution

To find the solution, we follow the same approach as before. First, we introduce $\tilde{x}$ whose real part is x, and write a complex differential equation whose real part is equation (4.1):

$$\ddot{\tilde{x}} + \gamma\dot{\tilde{x}} + \omega_0^2\tilde{x} = \frac{F_0}{m}e^{i\omega t}. \tag{4.2}$$

Then, we propose a trial solution $\tilde{x} = \tilde{A}e^{i\omega_1 t}$. For now, we've chosen some general frequency ω_1 because we don't yet know whether the oscillations will be at the natural frequency ω_0 or the drive frequency ω, or something else entirely. Substituting our trial solution into equation (4.2) we obtain

$$-\omega_1^2\tilde{A}e^{i\omega_1 t} + i\omega_1\gamma\tilde{A}e^{i\omega_1 t} + \omega_0^2\tilde{A}e^{i\omega_1 t} = \frac{F_0}{m}e^{i\omega t}.$$

Multiplying both sides by $e^{-i\omega t}$, we have

$$\tilde{A}e^{i(\omega_1-\omega)t}\left(-\omega_1^2 + i\omega_1\gamma + \omega_0^2\right) = \frac{F_0}{m}. \tag{4.3}$$

The right hand side of the equation is not a function of time, so the equation can only be satisfied if $\omega_1 = \omega$. This makes sense physically - the system oscillates at the frequency at which it's being driven. Then, equation (4.3) reduces to

$$\tilde{A} = \left(\frac{F_0}{m}\right)\left(\frac{1}{\omega_0^2 - \omega^2 + i\omega\gamma}\right) = \left(\frac{F_0}{m}\right)\left(\frac{\omega_0^2 - \omega^2 - i\omega\gamma}{(\omega_0^2-\omega^2)^2 + \omega^2\gamma^2}\right). \tag{4.4}$$

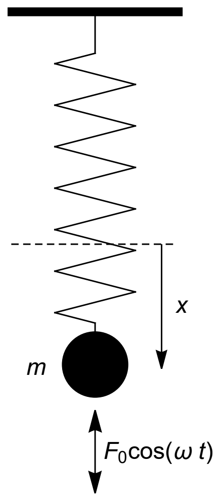

Figure 4.1: Harmonic oscillator driven by a periodic force.

In the last step, we multiplied the numerator and denominator by the complex conjugate of the denominator. We can write $\tilde{A} = Ae^{i\phi}$ and then use equation (4.4) to determine A and ϕ:

$$A = \left(\frac{F_0}{m}\right) \frac{1}{\sqrt{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2}}, \tag{4.5}$$

$$\phi = \tan^{-1}\left(\frac{-\omega\gamma}{\omega_0^2 - \omega^2}\right). \tag{4.6}$$

Then, taking the real part of the trial solution, we obtain the final result

$$x = A\cos(\omega t + \phi), \tag{4.7}$$

where A is given by equation (4.5) and ϕ is given by equation (4.6).

The system oscillates at the angular frequency of the driving force, ω. The oscillation has amplitude A and phase offset ϕ – this is the phase difference between the oscillation of the system and the oscillation of the driving force. Both A and ϕ are functions of ω and we have determined them by solving the differential equation – they are not determined by the initial conditions. Indeed, our solution is completely independent of any initial conditions! That's because we've only found part of the solution - it's the *steady state* solution that the system will reach if left for long enough. This is the most important part. The other part, called the *transient*, does depend on the initial conditions, and we will look at it soon.

4.3 Resonance

Let's look more closely at A and ϕ. They are plotted in figure 4.2. The amplitude peaks very strongly around $\omega = \omega_0$. This is resonance – the oscillation amplitude becomes very large when the system is driven at its natural frequency. The width of the peak is determined by the damping constant γ – the weaker the damping, the narrower the resonance peak becomes. In figure 4.2(b), we see how the phase offset ϕ changes with frequency. Well below the resonance frequency ϕ is close to zero meaning that the oscillations are in phase with the driving force. The driving force is "slow" compared to the system's natural frequency, so the oscillator stays in phase. Well above the resonance frequency ϕ approaches $-\pi$ meaning that the oscillations lag half a cycle behind the driving force. The force oscillates too quickly for the system to keep up. On resonance, the oscillator is $-\pi/2$ out of phase with the drive.

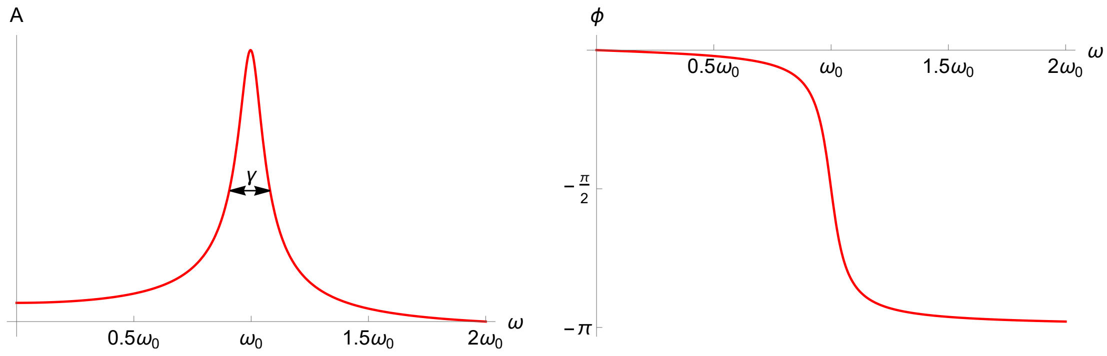

Figure 4.2: (a) Amplitude and (b) phase offset as functions of ω/ω_0. In this example, $\gamma = 0.1\omega_0$.

It is very common for γ to be much smaller than ω_0. In this case, the resonance becomes very narrow and very tall, and the only interesting regime is close to resonance where $\omega \approx \omega_0$. Since it's such an important case, let's look at this situation a bit further. First, note that when $\omega \approx \omega_0$, $(\omega_0^2 - \omega^2) = (\omega_0 + \omega)(\omega_0 - \omega) \approx 2\omega_0(\omega_0 - \omega)$. We can also take $\omega\gamma \approx \omega_0\gamma$. Applying these approximations to equations (4.5) and (4.6) we obtain the simpler results

$$A \approx \left(\frac{F_0}{2m\omega_0}\right)\frac{1}{\sqrt{(\omega-\omega_0)^2 + (\gamma/2)^2}}, \tag{4.8}$$

and

$$\phi \approx \tan^{-1}\left(\frac{\gamma}{\omega - \omega_0}\right). \tag{4.9}$$

These are usually excellent approximations, so are very commonly used to describe resonance. Equation (4.8) peaks at $\omega = \omega_0$ and has a width proportional to γ.

4.4 Quality factor, Q

Driven oscillators are often characterized by their quality factors, or Q-factor. It is defined as[1]

$$Q = \omega_0/\gamma. \tag{4.10}$$

In the context of a free (not driven) damped harmonic oscillator, like the one shown in figure 3.1(a), we see that Q tells us roughly how many oscillations the system makes before the oscillations damp away. In the context of resonance, it tells us how narrow the resonant peak will be relative to the natural frequency – larger Q means a narrower and taller peak.

We see from equation (4.5) that far below resonance ($\omega \ll \omega_0$) the amplitude has a constant (ω-independent) value of $A_0 = F_0/(m\omega_0^2) = F_0/\kappa$. On resonance ($\omega = \omega_0$), the amplitude is $F_0\omega_0/(m\omega_0^2\gamma) = QA_0$. So the amplitude on resonance is Q times higher than the amplitude below resonance. Many systems exhibit extremely large values of Q so that the resonances become enormous.

To take some examples, a tuning fork has a Q of about 10^3, the quartz crystal oscillators used in most clocks have Q factors of about 10^5, many atomic transitions have Q factors of 10^8 or higher, and the world's most stable lasers have Q factors of 10^{15}. To illustrate, consider an atom interacting with light. The light is an oscillating electric field which drives the electrons in the atom, and this system behaves just like a driven harmonic oscillator. To take a typical example, consider sodium atoms that have a natural frequency ω_0 (their transition frequency) close to 3×10^{15} rad/s, being

[1] There are several different definitions of Q, but they all become equal in the limit when Q is large, which is the only regime where it's a useful number.

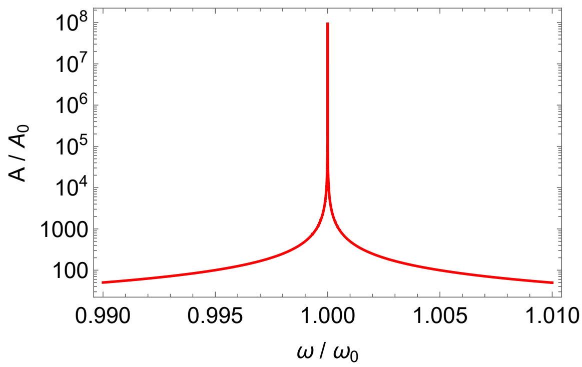

Figure 4.3: Amplitude, A/A_0, as a function of ω/ω_0 for a driven oscillator with a Q-factor of 10^8.

driven by near-resonant light (the yellow light that you may have seen from a sodium street lamp). The Q-factor for this oscillator is 10^8. Figure 4.3 shows the response of the atom to the light as a function of the light's frequency. To show what happens, the plot covers a tiny range of frequencies and the amplitude is plotted on a logarithmic scale. The atom does hardly anything if the light is even slightly off resonance, but goes crazy when the light is exactly resonant!

4.5 Summary of key points

- In the steady state, a driven harmonic oscillator oscillates at the drive frequency, ω.
- The amplitude and phase depend on ω.
- The amplitude is far larger when the drive frequency matches the natural frequency of the oscillator, $\omega = \omega_0$. This is resonance.
- The width of the resonance peak is proportional to the damping constant γ. The height is inversely proportional to γ. Lighter damping gives narrower and taller resonance peaks.
- Below resonance, the oscillations are in phase with the drive ($\phi \approx 0$). Above resonance, the oscillations lag half a cycle behind the drive, $\phi \approx -\pi$. On resonance, there is a 90 degree phase lag, $\phi = -\pi/2$.
- Oscillators are often characterized by their Q-factor, $Q = \omega_0/\gamma$.
- Resonance is a universal phenomenon found everywhere in nature and in all branches of physics.

Chapter 5

Driven harmonic oscillator II

5.1 Power

We have found the steady-state solution of the driven oscillator. In the steady-state, the energy stored in the oscillator is not changing, even though there is damping (dissipation) in the system. The power lost due to the damping is provided by the driving force. Let's work out the power being supplied by the drive.

To help with this, recall that our solution is

$$x = A\cos(\omega t + \phi) = A\left(\cos(\omega t)\cos\phi - \sin(\omega t)\sin\phi\right).$$

We'll find it helpful to use the second form here. Differentiating gives the velocity

$$\dot{x} = -A\omega\left(\sin(\omega t)\cos\phi + \cos(\omega t)\sin\phi\right).$$

Power is force times velocity, $P = F\dot{x}$, and the driving force is $F = F_0\cos(\omega t)$, so the power supplied by the drive is

$$P = -F_0 A\omega\left(\sin(\omega t)\cos(\omega t)\cos\phi + \cos^2(\omega t)\sin\phi\right). \tag{5.1}$$

The power changes as a function of time within each cycle. The more interesting quantity is the power averaged over one cycle,

$$\langle P\rangle = \frac{1}{T}\int_0^T P dt, \tag{5.2}$$

where T is the oscillation period. It is easy to show that $\frac{1}{T}\int_0^T \sin(\omega t)\cos(\omega t)dt = 0$, and that $\frac{1}{T}\int_0^T \cos^2(\omega t)dt = 1/2$. Substituting (5.1) into (5.2) we obtain

$$\langle P\rangle = -\frac{1}{2}F_0 A\omega\sin\phi. \tag{5.3}$$

Recall from figure 4.2(b) that $\sin\phi$ is negative for all drive frequencies, so our expression for the power is positive. Using equation (4.5) for A and (4.6) for ϕ, we find after a little algebra that the above result can also be written as

$$\langle P\rangle = \frac{1}{2}m\omega^2\gamma A^2. \tag{5.4}$$

This is the power supplied by the driving force. We can also calculate the power dissipated due to the damping force, which is $P = -F_{\rm damp}\dot{x}$, where the minus sign is because the power is lost. Using $F_{\rm damp} = -b\dot{x} = -m\gamma\dot{x}$ we obtain

$$P = -F_{\rm damp}\dot{x} = m\gamma\dot{x}^2 = m\omega^2\gamma\ A^2\sin^2(\omega t + \phi). \tag{5.5}$$

Taking the cycle average, we see that the power lost due to dissipation is

$$\langle P\rangle = \frac{1}{2}m\omega^2\gamma A^2. \tag{5.6}$$

This is the same as equation (5.4), showing that the power lost to damping is equal to the power supplied by the driving force, as we would expect in the steady state.

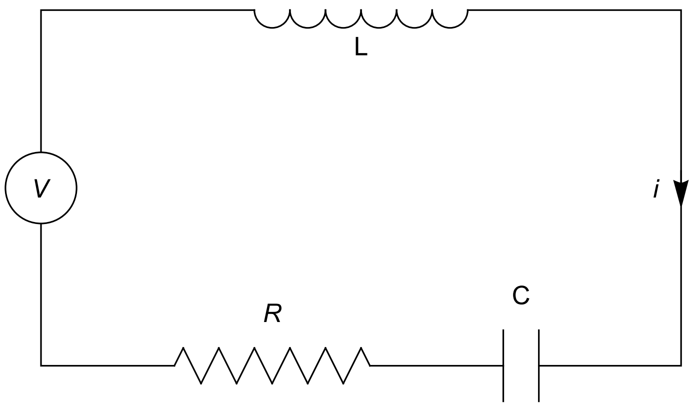

Figure 5.1: A series LCR circuit.

5.2 Energy

Next, we calculate the time-average energy stored in the oscillator. The kinetic energy is $K = 1/2m\dot{x}^2$. The potential energy is $U = 1/2m\omega_0^2x^2$ as shown in equation (2.10). So the total energy is

$$E = U + K = \frac{1}{2}m\omega_0^2x^2 + \frac{1}{2}m\dot{x}^2 = \frac{1}{2}m\omega_0^2A^2\cos^2(\omega t + \phi) + \frac{1}{2}m\omega^2A^2\sin^2(\omega t + \phi). \tag{5.7}$$

Taking the cycle average we obtain

$$\langle E\rangle = \frac{1}{2}m\omega_0^2\frac{A^2}{2} + \frac{1}{2}m\omega^2\frac{A^2}{2} = \frac{1}{4}mA^2(\omega_0^2 + \omega^2). \tag{5.8}$$

Note that the stored energy is proportional to the square of the amplitude, which becomes very large when there's not much damping and the oscillator is driven close to resonance. Moreover, when there's not much damping, it doesn't take much energy to maintain the oscillations. So there can be a lot of stored energy, without doing much work. In the last lecture, we introduced the quality factor, Q, defined by equation (4.10). There is another definition of Q - 2π times the stored energy divided by the work done in each cycle of the oscillator. The work done in each cycle is the mean power provided by the drive (equation (5.4)) multiplied by the time for one cycle ($T = 2\pi/\omega$). So we have

$$Q = 2\pi \times \frac{\text{Stored energy}}{\text{Work done per cycle}} = 2\pi \times \frac{1/4mA^2(\omega_0^2 + \omega^2)}{1/2m\omega^2\gamma A^2 \times 2\pi/\omega} = \frac{\omega_0^2 + \omega^2}{2\omega\gamma}. \tag{5.9}$$

Near a narrow resonance, where $\omega \approx \omega_0$, this reduces to

$$Q = \frac{\omega_0}{\gamma}, \tag{5.10}$$

which is the same result as in equation (4.10).

5.3 Resonant electrical circuits

In lecture 1, we looked at an LC circuit and noticed that the current in the circuit obeys the equation of a harmonic oscillator with natural frequency $\omega_0 = 1/\sqrt{LC}$. We can introduce damping into this circuit by adding a resistor, and we can provide a drive by adding an ac voltage source, $V = V_0\sin(\omega t)$. This gives us the driven LCR circuit illustrated in figure 5.1.

The voltage across the inductor is $V_L = L\frac{di}{dt}$, where i is the current in the circuit. The voltage across the capacitor is $V_C = q/C$, where q is the charge on the capacitor. The voltage across the

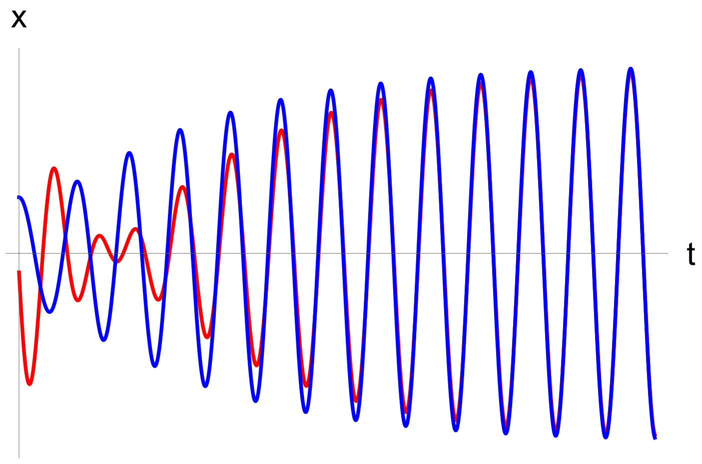

Figure 5.2: Driven harmonic oscillator for two different initial conditions. The transient depends on the initial conditions, but the steady state does not.

resistor is $V_R = iR$. The sum of these three voltages must equal the voltage supplied by the source:

$$\begin{aligned} &V_L + V_R + V_C = V_0 \sin(\omega t), \\ &L\frac{di}{dt} + Ri + \frac{q}{C} = V_0 \sin(\omega t), \\ &\frac{d^2 i}{dt^2} + \frac{R}{L}\frac{di}{dt} + \frac{1}{LC} i = \frac{V_0 \omega}{L}\cos(\omega t). \end{aligned} \tag{5.11}$$

In the last step, we've taken the derivative with respect to time, divided through by L, and used $i = \frac{dq}{dt}$. This equation can be written as

$$\frac{d^2 i}{dt^2} + \gamma\frac{di}{dt} + \omega_0^2 i = \frac{V_0\omega}{L}\cos(\omega t) \tag{5.12}$$

where $\omega_0 = 1/\sqrt{LC}$ and $\gamma = R/L$. This equation is that of a driven harmonic oscillator – it's the same as equation (4.1). The solution for the current must be $i = i_0 \cos(\omega t + \phi)$ where i_0 is given by equation (4.5) with F_0/m replaced by $V_0\omega/L$, and ϕ given by equation (4.6). Most notably, the circuit exhibits resonance when $\omega = \omega_0$, and the current can be very large in this case.

5.4 Transients

So far, we have found the "steady-state" solution to the driven harmonic oscillator – this is the solution that the system will tend to once the drive has been acting for long enough. We have also noted that our solution has no parameters that depend on initial conditions, so cannot be the full solution.

Let's write down equation (4.1) for the driven harmonic oscillator again:

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = \frac{F_0}{m}\cos(\omega t). \tag{5.13}$$

When you study differential equations, you learn that the general solution to equations like (5.13) can be found in two steps. First, we find any particular solution of (5.13) which is not the general solution. We've already found one – it's the steady-state solution we've been working with all this time. Let's call that solution x_1. Then, we find the general solution to the equivalent equation with zero on the right hand side (known as the *homogeneous* equation):

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = 0. \tag{5.14}$$

This is just the equation of a free harmonic oscillator with damping, whose solution we already found in lecture 3. This part of the solution is called the transient because it decays away over time. Let's write it as x_2. The general solution to (5.13) is $x_1 + x_2$. It's as simple as that! We can easily prove this by substituting $x_1 + x_2$ into the left hand side of (5.13):

$$\text{LHS} = \left[\ddot{x_1} + \gamma\dot{x_1} + \omega_0^2 x_1\right] + \left[\ddot{x_2} + \gamma\dot{x_2} + \omega_0^2 x_2\right] = [\frac{F_0}{m}\cos(\omega t)] + [0] = \text{RHS}. \tag{5.15}$$

The first square bracket is equal to $\frac{F_0}{m}\cos(\omega t)$ because we know x_1 satisfies (5.13), and the second square bracket is equal to zero because we know x_2 satisfies (5.14).

In the case of light damping, which is usually the case of interest, the complete solution is

$$x = A\cos(\omega t + \phi) + Be^{-\gamma t/2}\cos(\omega_\mathrm{d} t + \theta), \tag{5.16}$$

where A is given by equation (4.5), ϕ is given by (4.6), ω_d is given by (3.5), and B and θ are determined by the initial conditions.

5.5 Summary of key points

- When a driven harmonic oscillator reaches steady state, the power absorbed from the drive balances the power lost due to damping. This power is proportional to the damping constant γ and to the square of the oscillation amplitude.
- The stored energy is proportional to the square of the oscillation amplitude.
- A driven LCR circuit is a resonant electrical circuit. The resonant frequency is $\omega_0 = 1/\sqrt{LC}$.
- When the drive is first turned on, there is a transient oscillation at angular frequency ω_d (very close to ω_0 for light damping), which damps away to leave the driven, steady state, oscillation at angular frequency ω.

Chapter 6

Coupled oscillators

We now turn to another remarkable phenomenon found often in nature - what happens when two harmonic oscillators are coupled together so that the motion of one depends on the motion of the other? We will find that although the motion is quite complicated in general, it can always be written as the sum of two simple motions that we call the *normal modes*.

6.1 Equation of motion

To analyze what happens, we will study the system illustrated in figure 6.1. Two identical pendulums, each of length l, are coupled together by a spring whose spring constant is κ. The pendulums are arranged so that when they are hanging vertically downwards (i.e. not displaced) the spring is at its equilibrium length. The horizontal displacement of the left pendulum is x and that of the right pendulum is y. We will make the small angle approximation at all times.

If there were no spring, the equation of motion for the left pendulum would be $m\ddot{x} = -m\omega_0^2 x$ where $\omega_0 = \sqrt{g/l}$ [see equation (1.2)]. An exactly analogous equation of motion would describe the right pendulum, with the same value of ω_0 because the lengths are the same. The extension of the spring is $y - x$, so the force exerted by the spring on the left pendulum is $\kappa(y - x)$ (towards the right). The spring exerts an equal and opposite force on the right pendulum. So the equations of motion are

$$m\ddot{x} = -m\omega_0^2 x + \kappa(y - x), \tag{6.1a}$$

$$m\ddot{y} = -m\omega_0^2 y - \kappa(y - x). \tag{6.1b}$$

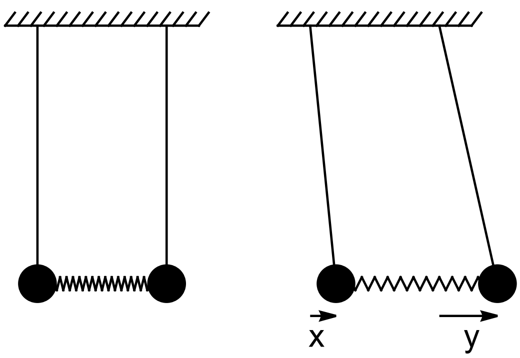

Figure 6.1: Two pendulums coupled by a spring. Left: Equilibrium. Right: Displaced.

6.2 Normal modes

We will look for the special solutions where both pendulums oscillate periodically with the same frequency ω. Using the complex exponential method, we propose the trial solutions $\tilde{x} = \tilde{A}e^{i\omega t}$, $\tilde{y} = \tilde{B}e^{i\omega t}$. Substituting into equations (6.1), cancelling the common factor of $e^{i\omega t}$ and collecting together the terms, we get

$$\left(\omega_0^2 - \omega^2 + \kappa/m\right)\tilde{A} = (\kappa/m)\tilde{B}, \tag{6.2a}$$

$$\left(\omega_0^2 - \omega^2 + \kappa/m\right)\tilde{B} = (\kappa/m)\tilde{A}. \tag{6.2b}$$

Multiplying these two equations together gives

$$\left(\omega_0^2 - \omega^2 + \kappa/m\right)^2 \tilde{A}\tilde{B} = (\kappa/m)^2\tilde{A}\tilde{B}.$$

We don't want $\tilde{A}$ or $\tilde{B}$ to be zero, because this corresponds to stationary pendulums. So, after taking a square root, we obtain

$$\omega_0^2 - \omega^2 + \kappa/m = \pm\kappa/m.$$

This gives us two possible solutions for ω, which we will call ω_1 and ω_2. They are $\omega_1^2 = \omega_0^2$ and $\omega_2^2 = \omega_0^2 + 2\kappa/m$. Introducing the natural frequency of the spring $\omega_s = \sqrt{\kappa/m}$, we have

$$\omega_1 = \omega_0, \tag{6.3a}$$

$$\omega_2 = \sqrt{\omega_0^2 + 2\omega_s^2}. \tag{6.3b}$$

These are the *normal mode frequencies.* Substituting these frequencies back into equations (6.2) we see that for ω_1 we must have $\tilde{A} = \tilde{B}$, and for ω_2 we have $\tilde{A} = -\tilde{B}$.

The special solutions we have found are called the *normal modes.* The frequencies ω_1 and ω_2 are the *normal mode frequencies.* Writing $\tilde{A} = Ae^{i\phi}$ and taking the real part, we see that these normal modes are

$$x_1 = A_1\cos(\omega_1 t + \phi_1), \quad y_1 = A_1\cos(\omega_1 t + \phi_1). \tag{6.4a}$$

$$x_2 = A_2\cos(\omega_2 t + \phi_2), \quad y_2 = -A_2\cos(\omega_2 t + \phi_2). \tag{6.4b}$$

We have used A_1, ϕ_1 for mode 1 and A_2, ϕ_2 for mode 2, recognizing that the overall amplitudes and phases of the modes are set by initial conditions and are allowed to be different. In the first mode, both pendulums move together, with exactly the same amplitude, frequency and phase. The spring never stretches, so never exerts any force, and the motion is exactly as it would be without any spring. That is why the frequency is the same as the natural frequency of the uncoupled pendulums. In the second mode, the two pendulums again oscillate with the same frequency and amplitude, but now they are exactly out of phase. In this case, the spring is stretched and compressed as the pendulums oscillate, so the spring contributes to the restoring force and raises the oscillation frequency.

Note that there are 2 normal modes, because there are 2 degrees of freedom (2 coordinates needed to describe the motion). If we have 3 coupled pendulums, there will be 3 normal modes. More generally, a system of coupled oscillators with N degrees of freedom will have N normal modes.

6.3 General solution

Equations (6.1) are linear in the displacements x and y. An important property of linear systems is that their solutions can be added together – this is the superposition principle. Thus, the general solution for the coupled harmonic oscillators is

$$x = A_1\cos(\omega_1 t + \phi_1) + A_2\cos(\omega_2 t + \phi_2), \tag{6.5a}$$

$$y = A_1\cos(\omega_1 t + \phi_1) - A_2\cos(\omega_2 t + \phi_2). \tag{6.5b}$$

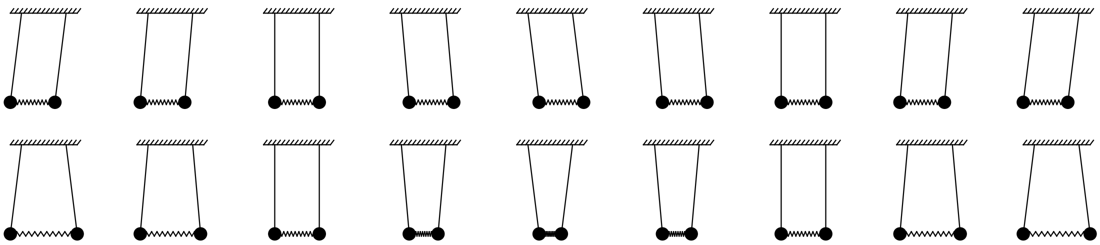

Figure 6.2: Normal modes of two coupled oscillators. Top: Symmetric mode with $\omega = \omega_0$. Bottom: Anti-symmetric mode with $\omega = \sqrt{\omega_0^2 + 2\omega_s^2}$.

Note that we have 4 undetermined coefficients (A_1, A_2, ϕ_1, ϕ_2) and 4 initial conditions (the initial values of $x, y, \dot{x}, \dot{y}$). It follows that these equations are sufficient to describe all general motions of the two coupled oscillators.

Although the general motion can be quite complicated and is not sinusoidal, it can always be expressed as the sum of sinusoidal motions of different frequencies.

6.4 Interesting case: one oscillator initially at rest

Consider the specific case where, at $t = 0$, one oscillator is at equilibrium and at rest ($y = 0, \dot{y} = 0$ at $t = 0$), while the other is released from rest with amplitude x_0 ($x = x_0, \dot{x} = 0$ at $t = 0$). Using these initial conditions in equations (6.5), we get $\phi_1 = \phi_2 = 0$, $A_1 = A_2 = x_0/2$. So the solution for this case is

$$x = \frac{x_0}{2}\left(\cos(\omega_1 t) + \cos(\omega_2 t)\right) = x_0 \cos\left(\frac{\omega_1 + \omega_2}{2} t\right) \cos\left(\frac{\omega_1 - \omega_2}{2} t\right), \tag{6.6a}$$

$$y = \frac{x_0}{2}\left(\cos(\omega_1 t) - \cos(\omega_2 t)\right) = -x_0 \sin\left(\frac{\omega_1 + \omega_2}{2} t\right) \sin\left(\frac{\omega_1 - \omega_2}{2} t\right). \tag{6.6b}$$

We have used trigonometric identities to get the forms on the right hand sides. These equations show that the motion is an oscillation at the mean frequency of the two normal modes, modulated by an envelope at the difference frequency. Figure 6.3 illustrates this motion. Initially the left pendulum has all the energy, but over time this is all transferred to the right pendulum. The left oscillator is driving the right oscillator, and because the total energy has to be conserved the energy must drain out of one and into the other. There is a moment in time when the left pendulum becomes stationary and the right one has all the energy. Then the energy starts transferring back again. Note that this exchange of energy is effective because the two oscillators are resonant with one another (the pendulums have the same lengths).

We will see the same phenomenon again when we consider the superposition of waves that have two different frequencies. Once again, we will see that the result is an oscillation at the mean frequency, modulated by an oscillation at the difference frequency. In the context of waves, we refer to this modulation as *beats*. Similarly, we can think of the motion of the coupled oscillators as a beat between the normal modes.

6.5 Summary of key points

- To analyze the motion of coupled oscillators, we first find the normal mode solutions.
- A normal mode is a special solution where the oscillators all move sinusoidally with the same frequency and with a fixed phase relation. Each normal mode has its own characteristic frequency.

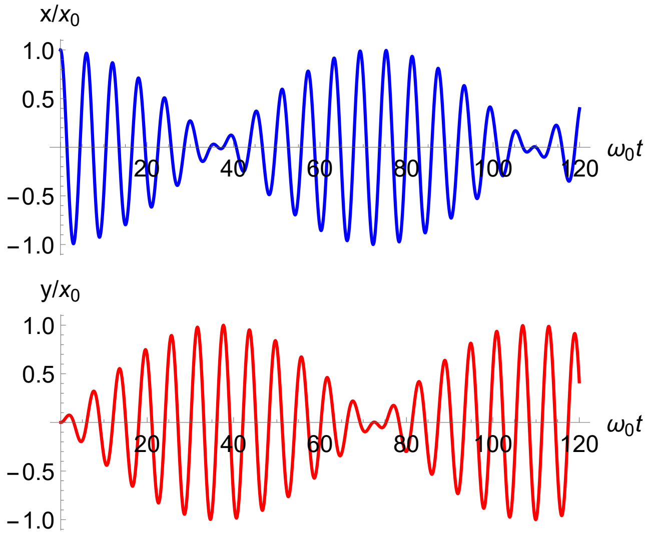

Figure 6.3: Motion of a pair of coupled harmonic oscillators where one oscillator starts at equilibrium and at rest, while the other starts at rest but displaced. The motion is illustrated for the case where $\omega_s = 0.3\omega_0$.

- For two coupled oscillators, there are two normal modes.
- One is a symmetric mode where the two oscillators have equal amplitudes and are in phase. Its frequency is the same as the frequency of the uncoupled oscillators.
- The second is an antisymmetric mode where the two have the same amplitude but are exactly out of phase. Its frequency is higher because of the extra restoring force provided by the coupling.
- The general solution is a superposition of the normal modes.
- When the oscillators are resonant (have the same ω_0) energy can be exchanged back and forth between them.

Chapter 7

Wave equation

A wave is the propagation of any kind of disturbance away from some equilibrium. Waves appear everywhere in nature. Examples include water waves, sound waves, waves on strings, seismic waves, electromagnetic waves, gravitational waves and the wavefunctions of quantum mechanics. Propagating waves typically transport energy, but there is no net transport of the medium itself (the water, air, string, earth etc). The waves obey a wave equation. The quantities involved may differ from one field to another, but the form of the equation is often the same. To find this form, we will look first at waves on a string.

7.1 Transverse waves

We consider transverse waves on a long string that lies parallel to the x axis. The string has a mass per unit length of ρ. There is a uniform tension T along the string. We would like to know how the transverse displacement, y, of each piece of the string varies with position x and with time t. A very small element of the string of length Δx is depicted in figure 7.1. The string makes a very small arc whose angle to the horizontal is θ at x and $\theta + \Delta\theta$ at $x + \Delta x$. The vertical and horizontal forces on this segment of string are

$$F_y = T\sin(\theta + \Delta\theta) - T\sin(\theta), \tag{7.1a}$$

$$F_x = T\cos(\theta + \Delta\theta) - T\cos(\theta). \tag{7.1b}$$

Making the small angle approximation, these forces are

$$F_y \approx T\Delta\theta, \quad F_x \approx 0.$$

The gradient of the arc at position x is $\tan\theta$ and is also $\frac{\partial y}{\partial x}$. Here, we use partial derivatives because the displacement y is a function of both x and t. Since our angles are small we have

$$\theta \approx \tan\theta = \left(\frac{\partial y}{\partial x}\right)_x, \quad \theta + \Delta\theta \approx \tan(\theta + \Delta\theta) = \left(\frac{\partial y}{\partial x}\right)_{x+\Delta x}.$$

The subscripts indicate the positions at which the derivatives are evaluated. It follows that[1]

$$\Delta\theta = \left(\frac{\partial y}{\partial x}\right)_{x+\Delta x} - \left(\frac{\partial y}{\partial x}\right)_x = \frac{\partial^2 y}{\partial x^2}\Delta x.$$

[1] In the last step, we have used the elementary definition of the derivative of a function f with respect to x, namely that

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

An alternative way to reach the same result is

$$\tan\theta = \frac{\partial y}{\partial x} \quad \therefore \sec^2\theta\Delta\theta = \frac{\partial^2 y}{\partial x^2}\Delta x$$

and $\sec^2\theta \approx 1$ for small θ.

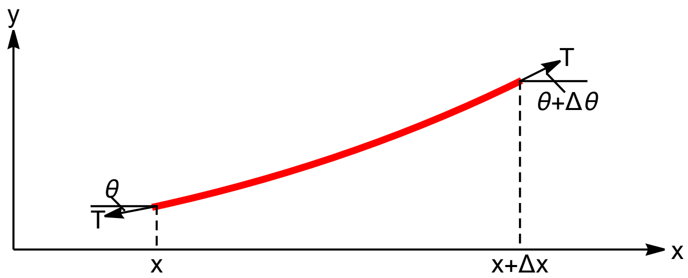

Figure 7.1: Forces on a tiny element of string that is displaced in the vertical direction.

Thus we have

$$F_y = T\frac{\partial^2 y}{\partial x^2}\Delta x. \tag{7.2}$$

The mass of the element of string is $\rho\Delta x$ and its acceleration in the vertical direction is $\partial^2 y/\partial t^2$. So we obtain the equation of motion

$$\frac{\partial^2 y}{\partial x^2} = \frac{\rho}{T}\frac{\partial^2 y}{\partial t^2}. \tag{7.3}$$

This is the equation that describes transverse waves on a string. We will write it in the standard form for a wave equation

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}. \tag{7.4}$$

where

$$v = \sqrt{T/\rho}. \tag{7.5}$$

We will understand soon that v can be interpreted as the speed of wave propagation.

7.2 Longitudinal waves

Now we consider longitudinal waves. We take the example of a long rod lying along the x-axis, as illustrated in figure 7.2. The rod has mass density[2] ρ and cross-sectional area A. Imagine banging on one end of the rod. This exerts a force on the particles at that end which causes them to move slightly. They cannot move far because of the very strong restoring forces present in the material, but they move a bit. The particles a little further down haven't moved yet, so the small slice of material at the end of the rod is compressed. The resulting stress exerts a force further on, and in this way a wave of compression is set up that propagates down the rod.

To analyze this situation, we have to remember about stress and strain. If you try to stretch a rod by pulling in opposite directions on the two ends with force F, the stress in the rod is simply $\sigma = F/A$, where A is the cross-sectional area. If the equilibrium length of the rod is L and you stretch it by ΔL, the strain is $s = \Delta L/L$. The Young's modulus, Y, is the ratio of stress to strain, $Y = \sigma/s$.

Consider a very small section of the rod, such as the blue shaded section in figure 7.2. At equilibrium, this section is at position x and has a tiny length Δx. Now the section is forced out of equilibrium - it moves a little bit and changes its length. Let the displacement of the left hand side of the section be ϵ, and the displacement of the right hand side be $\epsilon + \Delta\epsilon$. The section has changed length by the amount $\Delta\epsilon$, and since its equilibrium length is Δx the strain is $s = \Delta\epsilon/\Delta x$. The corresponding stress at position x is

$$s_x = Y\left(\frac{\partial \epsilon}{\partial x}\right)_x, \tag{7.6}$$

[2]Beware - here ρ is mass per unit volume, whereas in the case of the string it is mass per unit length.

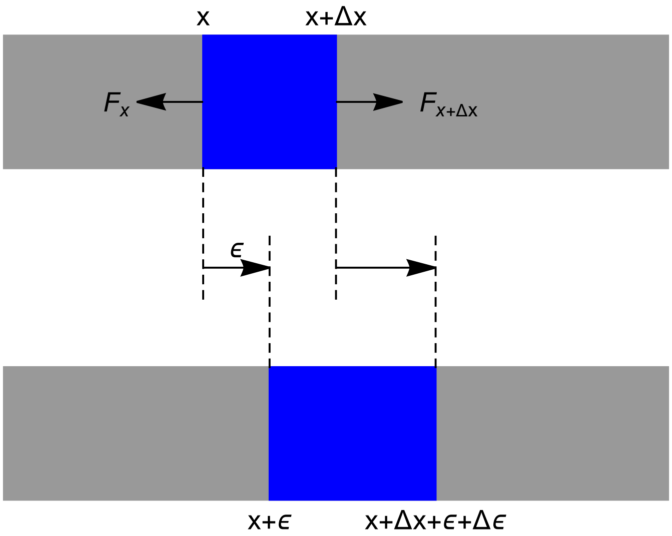

Figure 7.2: Wave propagation in a rod, showing how a small section of the rod is displaced and stretched by the forces acting on the section. The section gets displaced by ϵ and stretched by $\Delta\epsilon$. The aim is to find an equation showing how ϵ changes with x and t.

where we have taken the limit of $\Delta x \to 0$. We need a partial derivative, because we expect ϵ to be a function of x and t, and we use the subscript x to indicate the position where the derivative is taken. From the definition of the stress, we get the force at x to be

$$F_x = AY \left(\frac{\partial \epsilon}{\partial x}\right)_x . \tag{7.7}$$

Similarly, the force at $x + \Delta x$ is

$$F_{x+\Delta x} = AY \left(\frac{\partial \epsilon}{\partial x}\right)_{x+\Delta x} = AY \left[\left(\frac{\partial \epsilon}{\partial x}\right)_x + \left(\frac{\partial^2 \epsilon}{\partial x^2}\right)_x \Delta x\right]. \tag{7.8}$$

In the last step we have used a Taylor expansion of $\partial\epsilon/\partial x$. The net force on the section is the difference between the two forces,

$$F = F_{x+\Delta x} - F_x = AY \frac{\partial^2 \epsilon}{\partial x^2} \Delta x. \tag{7.9}$$

This net force is responsible for the acceleration of the section, which is $\partial^2\epsilon/\partial t^2$. The mass of the section is $\rho A \Delta x$, so from Newton's second law we have

$$AY \frac{\partial^2 \epsilon}{\partial x^2} \Delta x = \rho A \Delta x \frac{\partial^2 \epsilon}{\partial t^2}. \tag{7.10}$$

Cancelling the common factors and rearranging we reach

$$\frac{\partial^2 \epsilon}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 \epsilon}{\partial t^2}, \tag{7.11}$$

where $v = \sqrt{Y/\rho}$.

This is identical to equation (7.4), except that the transverse displacement y has been replaced with the longitudinal displacement ϵ, and the expression for v is slightly different in the two cases. Thus we see that transverse waves on a string and longitudinal waves in a rod are both described by an equation of the same form.

The longitudinal wave in the rod is a type of sound wave. We usually think of sound travelling through a gas, but it can travel through any medium and when it travels through a solid it is described by the wave equation we have found. When it travels through a gas we have the same equation, except we don't use a Young's modulus when thinking about a gas. Instead, we use the pressure of the gas, P. Everything is exactly the same, but now the expression for the speed is $v = \sqrt{\gamma P/\rho}$. The constant γ is a number close to 1 and is called the ratio of specific heats - you will meet it when you study thermodynamics. For air at room temperature, $\gamma \approx 1.4$.

7.3 Wave equation in three dimensions

Often, the quantity of interest – let's call it ψ – can vary in all three dimensions. Good examples are the electric and magnetic fields of an electromagnetic wave, i.e. light. In that case, the wave equation generalizes to

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2}. \tag{7.12}$$

This can be written in the more compact form (see vector calculus course)

$$\nabla^2\psi = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2}. \tag{7.13}$$

7.4 Summary of key points

- The wave equation in 1D is $\frac{\partial^2\psi}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2}$. In 3D, it is $\nabla^2\psi = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2}$
- $\psi(x,t)$ is the quantity disturbed from equilibrium – the transverse displacement for a wave on a string, the longitudinal displacement for a sound wave, the electric or magnetic field for an electromagnetic wave etc.
- v is the speed of wave propagation. For transverse waves on a string, $v = \sqrt{T/\rho}$ where T is the tension in the string and ρ is its mass per unit length.

Chapter 8

Waves I

In the last lecture, we saw that the wave equation has the form

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}. \tag{8.1}$$

We found this in the context of waves on a string, where y is the transverse displacement of the string, but we also emphasized that the same equation appears in many other contexts.

8.1 Solution

The solution to our wave equation needs to be a function of x and t. We can make an educated guess about the form of the solution using the following argument. The solution is meant to represent a wave. If we give one end of the string a little shake, making a pulse such as the one illustrated in figure 8.1, we might expect that pulse to propagate along the string at some speed v. Suppose we represent the initial pulse by a function $f(x)$. After a time t, the pulse will have moved a distance vt, so is represented by the displaced function $f(x - vt)$. Thus, we propose a solution to the wave equation of the form $y = f(x - vt)$. We won't specify the function f at this point - we'll just substitute into the wave equation and see what happens.

To keep the notation clear, we introduce the variable $q = x - vt$ so that our trial solution is $y = f(q)$. Take derivatives with respect to x:

$$\frac{\partial f}{\partial x} = \frac{df}{dq}\frac{\partial q}{\partial x} = \frac{df}{dq}, \tag{8.2}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{d}{dq}\left(\frac{df}{dq}\right)\frac{\partial q}{\partial x} = \frac{d^2 f}{dq^2}, \tag{8.3}$$

Similarly, the derivatives with respect to t are

$$\frac{\partial f}{\partial t} = \frac{df}{dq}\frac{\partial q}{\partial t} = -v\frac{df}{dq}, \tag{8.4}$$

$$\frac{\partial^2 f}{\partial t^2} = \frac{\partial}{\partial t}\left(\frac{\partial f}{\partial t}\right) = \frac{d}{dq}\left(-v\frac{df}{dq}\right)\frac{\partial q}{\partial t} = v^2\frac{d^2 f}{dq^2}. \tag{8.5}$$

Comparing equations (8.2) and (8.4) we see that

$$\frac{\partial y}{\partial x} = -\frac{1}{v}\frac{\partial y}{\partial t}. \tag{8.6}$$

Similarly, comparing equations (8.3) and (8.5) we see that

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}. \tag{8.7}$$

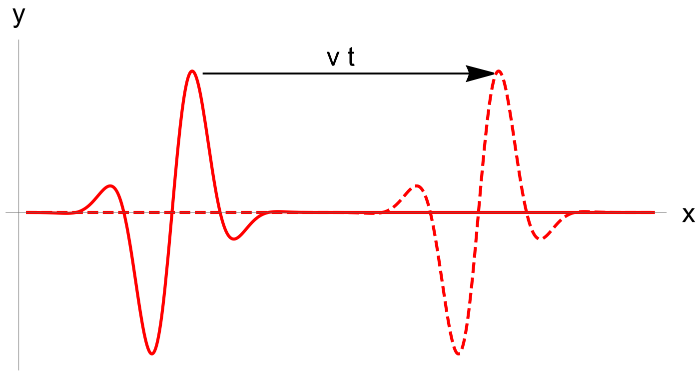

Figure 8.1: A pulse represented as $f(x)$ moves at speed v, so that after time t it is represented by $f(x - vt)$.

This is the same as equation (8.1). Thus $y = f(x-vt)$ satisfies the wave equation and we understand v to be the speed at which the wave propagates.

It is common to solve differential equations by proposing some trial solution, substituting it in, and then finding that the proposed solution has to satisfy some constraints. But here we didn't find any constraints on the form of f. It can be any function of $x - vt$. This makes sense. We could choose to shake one end of the string in all manner of ways, setting up all kinds of different initial pulse shapes on the string. And it seems reasonable to expect that pulse to propagate down the string at some speed that depends on the properties of the string but *not* on the shape of the pulse. That is indeed what the wave equation tells us.

One more thing. Instead of $f(x-vt)$ we could have chosen $g(x+vt)$ where g is also an arbitrary function. This is a wave travelling in the opposite direction. It clearly also satisfy the wave equation, because the quantity that appears in equation (8.5) is v^2 – the sign of v doesn't matter. This also makes sense – we are free to shake either end of the string, and the string must be able to support waves travelling in either direction.

8.2 Sinusoidal waves

So far, we have seen that $f(x - vt)$ or $g(x + vt)$, are solutions of the wave equation, where f and g can be any function. That is very general! We also know from experience that waves are often periodic. So let's choose our function f to be sinusoidal. The argument of the sinusoidal function has to be an angle, but $x - vt$ is a length , so we'll need to multiply it by a constant (we'll call it k) that has the units of 1/length, in order to make an angle. Thus, as a simple periodic solution of the wave equation, we propose

$$y = A\cos[k(x - vt)].$$

It is often preferable to write this as

$$y = A\cos(kx - \omega t) \tag{8.8}$$

where $\omega = kv$.

The wave is an oscillation in both space and time. Figure 8.2(a) shows how y changes with t at a fixed point in space. The period (T) of the wave is the time taken to complete one cycle, e.g. from t_1 to t_2 in the figure. The argument of the cosine must increase by 2π in going from t_1 to t_2. Since the position is fixed, we must have $\omega(t_2 - t_1) = \omega T = 2\pi$. So we see that

$$\omega = \frac{2\pi}{T}, \tag{8.9}$$

and thus interpret ω as the angular frequency (units: rad/s). The number of oscillations per second, known as the frequency (unit: Hz), is $f = 1/T = \omega/(2\pi)$.

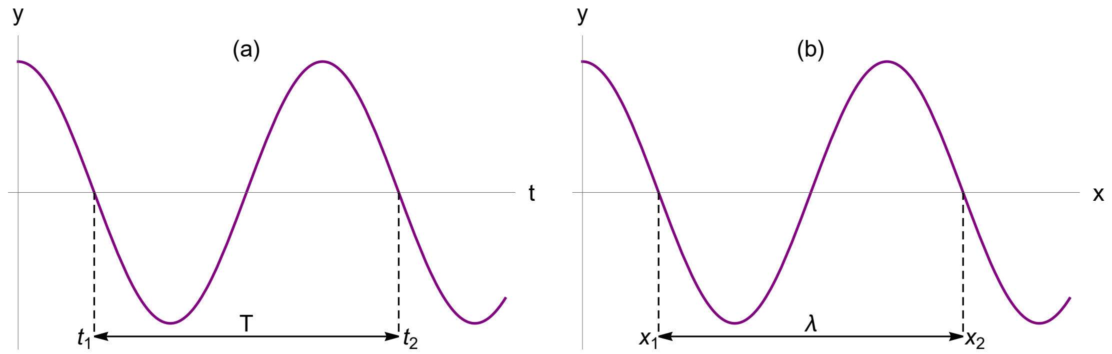

Figure 8.2: (a) Wave as a function of t at a fixed position. (b) Wave as a function of x at a fixed time.

Figure 8.2(b) shows how y changes with x at a fixed point in time. The wavelength (λ) of the wave is the distance taken to complete one cycle, e.g. from x_1 to x_2 in the figure. The argument of the cosine must increase by 2π in going from x_1 to x_2. Since the time is fixed, we must have $k(x_2 - x_1) = k\lambda = 2\pi$. So we see that

$$k = \frac{2\pi}{\lambda}. \tag{8.10}$$

k is called the wavenumber (units: rad/m).

Suppose we fix our attention on a particular point on the wave, say a particular crest, and we follow it as it moves. It is at position x at time t and has moved to $x+\Delta x$ at time $t+\Delta t$. Saying that it is the same point on the wave is the same as saying that the argument of the cosine is unchanged, thus $kx - \omega t = k(x + \Delta x) + \omega(t + \Delta t)$, i.e. $k\Delta x = \omega\Delta t$. So we see that

$$\frac{\Delta x}{\Delta t} = \frac{\omega}{k} = v. \tag{8.11}$$

But $\Delta x/\Delta t$ is speed – the speed at which the crest (or any other point on the wave) is moving – so we confirm our interpretation of v as the speed of the wave.

8.3 Plane waves in three dimensions

We have seen that $\psi = A\cos(kx - \omega t)$ is a solution of the wave equation in 1D. It is also a solution to the 3D wave equation as is easily seen by substitution into equation (7.12). It is called a *plane wave* meaning that there is no variation in the directions perpendicular to the propagation direction. The surface where ψ is a constant (a *wavefront*) is defined by x = constant – meaning any plane parallel to the yz-plane. Figure 8.3(a) illustrates this plane wave solution.

A more general plane wave is

$$\psi(\vec{r}, t) = A\cos\Big(\vec{k}.\vec{r} - \omega t\Big), \tag{8.12}$$

where $\vec{r}$ is the displacement vector and $\vec{k} = k_x\hat{x} + k_y\hat{y} + k_z\hat{z}$ is called the wavevector. The direction of $\vec{k}$ is the direction of propagation of the wave, and its magnitude is $|\vec{k}| = 2\pi/\lambda$.

8.4 Spherical waves in two or three dimensions

Often, a wave originates from a localized disturbance, e.g. dropping a pebble in a lake, or a small source of light such as a bulb in a room or a star in a galaxy. In these cases, the wave propagates equally in all directions. The wavefronts are circles in 2D (e.g. waves on a lake) or spheres in 3D

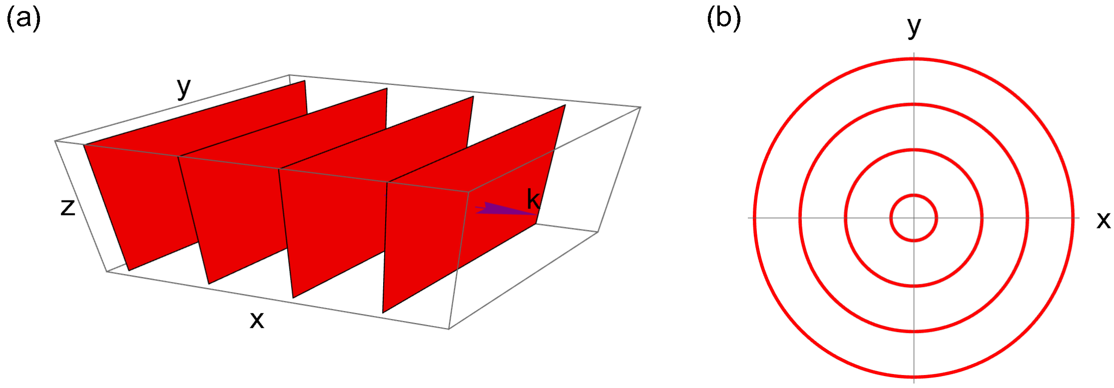

Figure 8.3: (a) Wavefronts of a plane wave propagating in the x direction. (b) Wavefronts of a 2D spherical wave.

(e.g. light waves). These waves are called spherical waves. They are illustrated in figure 8.3(b) and are described by

$$\psi(\vec{r}, t) = \frac{A}{r}\cos(kr - \omega t). \tag{8.13}$$

Here $\vec{r}$ is the displacement from the origin of the wave. The wave depends only on the magnitude of $\vec{r}$ ($r = \sqrt{x^2 + y^2}$ in 2D and $r = \sqrt{x^2 + y^2 + z^2}$ in 3D). It has no dependence on the angles as we would expect for this spherically symmetric situation. We will see in the next lecture that the intensity of the wave is proportional to the square of the amplitude. The amplitude of the spherical wave is A/r so the intensity drops off with distance as $1/r^2$. This is the inverse-square law.

8.5 Summary of key points

- Any function of the form $f(x \pm vt)$ is a solution of the wave equation.
- Waves are often sinusoidal, $y = A\cos(kx - \omega t)$. A is the amplitude of the wave, k is the wavenumber, and ω is the angular frequency.
- A plane wave is a wave that does not vary in the directions perpendicular to the propagation direction.
- A spherical wave propagates equally in all directions and has an intensity that drops as the inverse-square of the distance.

Chapter 9

Waves II

9.1 Energy

Let's work out the energy carried by the wave. To do this, we look again at our short length of string, shown in figure 9.1. The equilibrium length of the string is Δx and the stretched length due to the transverse wave is Δs.

The mass of the small section is $\rho\Delta x$, so the kinetic energy of the small section is

$$\Delta K = \frac{1}{2}\rho\Delta x\left(\frac{\partial y}{\partial t}\right)^2. \tag{9.1}$$

The potential energy is the tension times the extension, which is $U = T(\Delta s - \Delta x)$. If the length is very tiny, the segment is very nearly straight, so

$$\Delta s = \sqrt{\Delta x^2 + \Delta y^2} = \Delta x\sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2} \approx \Delta x\left[1 + \frac{1}{2}\left(\frac{\partial y}{\partial x}\right)^2\right]. \tag{9.2}$$

In the last step, we have expanded the square root using the fact that we are dealing always with small amplitude waves (small angles) and so $\Delta y/\Delta x \ll 1$. We have also converted the ratio to a partial derivative which is valid in the limit of vanishing Δx. Thus for the potential energy we have

$$\Delta U = T(\Delta s - \Delta x) = \frac{1}{2}T\left(\frac{\partial y}{\partial x}\right)^2\Delta x. \tag{9.3}$$

Finally, using equation (8.6) to relate derivatives with respect to x and t, and equation (7.5) for v, we obtain

$$\Delta U = \frac{1}{2}\rho\left(\frac{\partial y}{\partial t}\right)^2\Delta x. \tag{9.4}$$

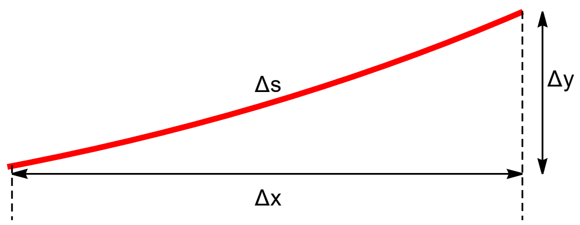

Figure 9.1: Coordinates used to work out the energy carried by a wave. The short length of string has equilibrium length Δx. When displaced, it has the new length Δs.

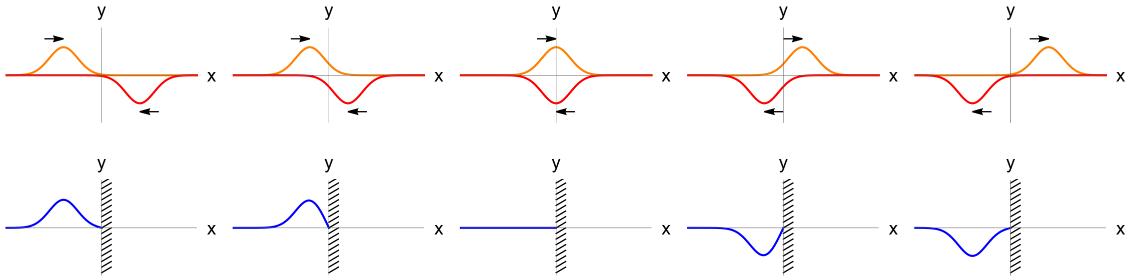

Figure 9.2: Reflection of a wave on a string as it approaches a fixed end point at $x = 0$. Time is increasing from the left plot to the right plot. Upper row shows $f(x - vt)$ (orange) and $-f(-x - vt)$ (red). The bottom row shows their sum in the region $x < 0$. and the solid line is their sum. The plots show the wave at different moments in time.

This is exactly the same as equation (9.1) for the kinetic energy. So we see that the potential energy and kinetic energy are equal.

Finally, let us consider the power delivered by the wave. We can imagine a perfect absorber at one end of the string that completely absorbs the wave. The power is the total energy per unit time, $P = (\Delta K + \Delta U)/\Delta t$. Since $\Delta x/\Delta t = v$, we obtain

$$P = \rho v \left(\frac{\partial y}{\partial t}\right)^2 . \tag{9.5}$$

Now we can evaluate this for our sinusoidal wave, equation (8.8):

$$P = \rho v \omega^2 A^2 \sin^2(kx - \omega t). \tag{9.6}$$

The first important thing to take away from this result is that the power in the wave is proportional to the square of the amplitude of the wave (A^2). The second important thing to notice is that the expression for the power is of the form $f(x - vt)$, which we already saw is a quantity propagating at speed v – waves transport energy.

Equation (9.6) gives us the instantaneous power. Often, the detector used to measure the power cannot respond on the timescale of a single cycle of the wave, so we are very often interested in the power averaged over many cycles. The time average of $\sin^2(kx - \omega t)$ is $1/2$. So we get a time-average power of

$$\langle P \rangle = \frac{1}{2} \rho v \omega^2 A^2 \tag{9.7}$$

9.2 Reflection at one end

Consider again the wave propagating along a string. What happens when it gets to the end? Let's suppose that the end is connected to a rigid wall so that it's impossible for the very end of the string to move. We'll put the wall at $x = 0$ so that the string stretches from $x = -\infty$ to $x = 0$ and we have the boundary condition $y(x = 0) = 0$.

We already know that the general solution for the wave on the string is

$$y = f(x - vt) + g(x + vt) \tag{9.8}$$

where f and g are any functions. The first term is a wave travelling one way, and the second is a wave travelling the other way. Imposing the boundary condition at $x = 0$ gives us

$$f(-vt) + g(vt) = 0.$$

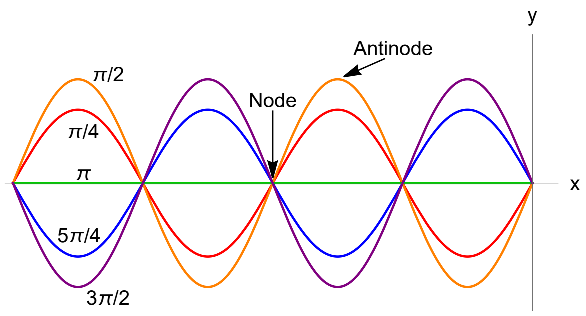

Figure 9.3: Snapshots of a standing wave at various moments in time. The labels give the values of ωt.

This has to hold at all times, so it should be read as $g(\text{anything}) = -f(-\text{anything})$. Using this in equation (9.8) we obtain

$$y = f(x - vt) - f(-x - vt). \tag{9.9}$$

Figure 9.2 illustrates the reflection of a pulse propagating on the string. The upper row shows $f(x - vt)$ and $-f(-x - vt)$ separately. Here, we pretend there's no wall and plot the functions for all x. The bottom row shows the sum of these two, just in the region $x < 0$, showing that the pulse is reflected and inverted. Notice that, at all times, $y = 0$ at $x = 0$. We see that the reflection at the wall is equivalent to having an infinitely long string that has waves travelling in both directions related by a reflection and inversion through $x = 0$.

9.3 Standing wave

Now let's suppose our wave is a sinusoidal wave like the ones studied in the last lecture i.e. $f(x-vt) = A\cos(kx - \omega t)$. Then the total wave is

$$y = A\cos(kx - \omega t) - A\cos(-kx - \omega t). \tag{9.10}$$

With the help of a trigonometric identity, this can be written as

$$y = 2A\sin(kx)\sin(\omega t). \tag{9.11}$$

This is a very different type of wave. Suppose we take a snapshot of the string at some fixed point in time, so that $\sin(\omega t)$ is a constant. We will see a sine wave on the string, but the amplitude of this wave depends on when we take the snapshot. There are some moments of time, e.g. $t = 0$, where the entire string is flat. Similarly, if we concentrate on some particular point, so that $\sin(kx)$ is constant, we will see that point oscillating in time with angular frequency ω. The point is a harmonic oscillator. This is true regardless of which point we choose - all points are harmonic oscillators all oscillating with the same frequency. We could regard all these points as coupled harmonic oscillators, in which case we would call this motion a normal mode, since 'everything oscillating sinusoidally at the same frequency' is the definition of a normal mode. We also see that there are some particular points, where the amplitude of oscillation is zero – there are points that do not move at all. These points are at $kx = n\pi$ where n is any integer, and thus are separated by $\Delta x = \lambda/2$. These stationary points are called the *nodes*. Similarly, there are points half way between where the oscillation has its maximum extent – these are called the *antinodes*. Note that the nodes and antinodes (or any other chosen point on the wave) do not move along the string – the wave is not propagating so we call it a *standing wave*.

9.4 Impedance

For an electrical circuit, impedance is the voltage divided by the current. For a mechanical wave like the transverse wave on a string, the impedance, Z, is force divided by velocity – it tells you how much force is needed to impart a certain velocity. For the wave on the string, the force is the tension in the string, so $Z = T/v$. Since $v = \sqrt{T/\rho}$, we can also write the impedance as $Z = \sqrt{T\rho}$. Note that impedance is a property of the medium (here the string), not a property of the wave (i.e. does not depend on the amplitude or frequency of the wave).

9.5 Reflection, transmission and impedance matching

Suppose we have a wave incident on some interface between two different media, for example a sound wave propagating from air into water, or a light wave propagating from air to glass, or a wave passing from one string to a different one. We can expect some part of the wave to be reflected at the interface and some part to be transmitted. We would like to know how much is reflected and how much transmitted.

To analyze this, we once again take the example of transverse waves on a string. We join two different strings at $x = 0$ so that string 1 goes from $x = -\infty$ to zero and string 2 from zero to ∞. A wave initially propagates on string 1, heading towards $x = 0$. Previously, we wrote a general solution of the wave equation as $f(x - vt)$. We could equally well write this is as $f(t - x/v)$ – it still satisfies the wave equation – and because we are interested in what happens at $x = 0$ it will be convenient to write it that way. We expect part of the wave to be reflected from the interface at $x = 0$, so in general the wave on string 1 is

$$y_1(x,t) = f_1(t - x/v_1) + g_1(t + x/v_1). \tag{9.12}$$

f_1 is the incident wave and g_1 is the reflected wave. We expect another part of the wave to be transmitted so there will be a wave on string 2 travelling in the positive x direction:

$$y_2(x,t) = f_2(t - x/v_2). \tag{9.13}$$

Note that we have used a different speed for string 2, because it's a different string (e.g. different mass per unit length). Now we can apply some continuity conditions. First, since the two strings are connected at $x = 0$ their transverse displacements must be equal at $x = 0$, i.e $y_1(0,t) = y_2(0,t)$. This gives us

$$f_1(0) + g_1(0) = f_2(0), \tag{9.14}$$

for all t. Next, we realize that the transverse force must also be continuous at $x = 0$. For it to be discontinuous, the force would need to change by a finite amount across an infinitesimally small section of string, which would generate an infinite acceleration which is not physical. We saw in chapter 7 that the transverse force is $F_y = T\frac{\partial y}{\partial x}$, where T is the tension. So, for the force to be continuous at $x = 0$ we must have

$$T_1 \left(\frac{\partial y_1}{\partial x}\right)_{x=0} = T_2 \left(\frac{\partial y_2}{\partial x}\right)_{x=0}. \tag{9.15}$$

This means

$$T_1 \frac{\partial f_1}{\partial x} + T_1 \frac{\partial g_1}{\partial x} = T_2 \frac{\partial f_2}{\partial x}, \tag{9.16}$$

where the derivatives are evaluated at $x = 0$. Note from equation (8.6) that the derivatives with respect to x and t are related to each other: $\frac{\partial f_1}{\partial x} = -\frac{1}{v_1}\frac{\partial f_1}{\partial t}$ etc. So we can re-write the above equation as

$$-\frac{T_1}{v_1}\frac{\partial f_1}{\partial t} + \frac{T_1}{v_1}\frac{\partial g_1}{\partial t} = -\frac{T_2}{v_2}\frac{\partial f_2}{\partial t}. \tag{9.17}$$

But T/v is the impedance Z. Furthermore, since the derivatives are evaluated at $x = 0$ they are only functions of t so we can convert to a complete derivative, giving us

$$\frac{d}{dt}\left(-Z_1 f_1 + Z_1 g_1 + Z_2 f_2\right) = 0. \tag{9.18}$$

For the derivative to be zero, the quantity in brackets must be a constant: $-Z_1 f_1 + Z_1 g_1 + Z_2 f_2 =$ const. In fact, the constant has to be zero to make sure that there is no displacement of string 2 when there is no incident wave. So, remembering that everything is evaluated at $x = 0$ we have

$$-Z_1 f_1(0) + Z_1 g_1(0) + Z_2 f_2(0) = 0. \tag{9.19}$$

Equations (9.14) and (9.19) are a pair of simultaneous equations that can be solved to find the ratio of the reflected wave to the incident wave, known as the reflection coefficient f_R:

$$f_R = \frac{g_1(0)}{f_1(0)} = \frac{Z_1 - Z_2}{Z_1 + Z_2}. \tag{9.20}$$

Alternatively, we can solve for the ratio of the transmitted wave to the incident wave, known as the transmission coefficient f_T:

$$f_T = \frac{f_2(0)}{f_1(0)} = \frac{2Z_1}{Z_1 + Z_2}. \tag{9.21}$$

We see that the reflection and transmission coefficients depend only on the ratio of the two impedances. These depend only on the medium, not on the amplitude or frequency of the wave. We also see that if $Z_1 = Z_2$ the wave will pass from one medium to the other without any reflection. This is known as impedance matching.

9.6 Summary of key points

- A travelling wave transports energy.
- The energy and power are proportional to the square of the amplitude of the wave.
- When there is a boundary condition – often that the wave must vanish at a particular place – there is a reflection.
- When there is an incident wave and a reflected wave, their sum produces a standing wave. The standing wave oscillates but does not travel.
- The standing wave has nodes – points where there is no motion. The nodes are separated by half a wavelength.

Chapter 10

Modes

In the last lecture, we looked at the case where a string was fixed at one end. We saw how this sets up a reflection of the wave, and how this can lead to a standing wave on the string. Now we consider what happens when the string is fixed at both ends. We will find that the string can only oscillate sinusoidally at certain special frequencies and that there is a certain pattern of displacement associated with each – these are the normal modes. This is a general phenomenon – wherever there are confined waves there will be modes. We will look at some examples.

10.1 Modes of a vibrating string fixed at both ends

The string is fixed at both ends, say at $x = 0$ and $x = L$. We require $y = 0$ at both of these points. Imposing the boundary condition at $x = 0$ gives us equation (9.11), as already seen. Then, imposing the condition at $x = L$ and requiring this to hold for all t we see that we must have $\sin(kL) = 0$. The only way to satisfy this is by choosing particular values of k, namely

$$k_n = n\frac{\pi}{L}, \tag{10.1}$$

where n is an integer. We have a set of *allowed values* for k, and we distinguish between them using the subscript n. We can equivalently talk about the allowed wavelengths which are

$$\lambda_n = \frac{2L}{n}. \tag{10.2}$$

This has a simple interpretation. We know there has to be fixed points (nodes) at $x = 0$ and $x = L$, and we know that the nodes of a standing wave are spaced by $\lambda/2$. So the longest wavelength that

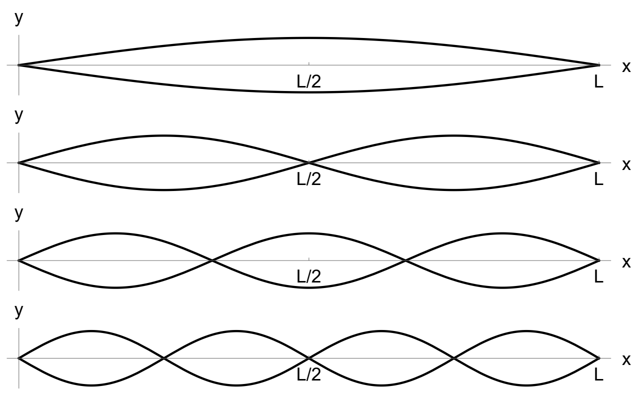

Figure 10.1: The first four modes of a vibrating string of length L fixed at both ends.

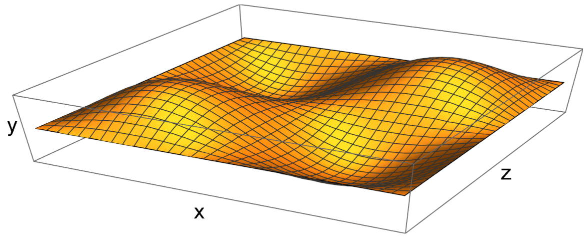

Figure 10.2: Waves on a rectangular plate or membrane whose edges are fixed. At equilibrium, the plate lies in the xz-plane. Waves on the plate are described by the displacement $y(x, z, t)$.

will 'fit' on the string is $\lambda_1 = 2L$. But we could also put an extra node half way along and still satisfy the boundary conditions, and this wave would have wavelength $\lambda_2 = L$. Or you could put 2 nodes in between and get a standing wave of wavelength $\lambda_3 = 2L/3$. And so on – there are a discrete set of standing waves that will fit. These are illustrated in figure 10.1.

Finally, since $\omega = kv$ and $v = \sqrt{T/\rho}$, we see that there are an equivalent set of discrete angular frequencies,

$$\omega_n = n\frac{\pi}{L}\sqrt{\frac{T}{\rho}}. \tag{10.3}$$

The string will only oscillate at these particular frequencies. The simplest mode (with $n = 1$) is the *fundamental mode*. It is also called the *first harmonic*. The next mode has twice the frequency and we call it the second harmonic. If you want to change the fundamental frequency you have to change the length of the string or the tension in the string or the density of the string. You get a higher frequency by shortening the length or increasing the tension or lowering the density. These are principles of stringed musical instruments.

The allowed frequencies define a set of normal modes. Associated with each normal mode there is a pattern – in the case of the string it is a pattern of displacement and these are drawn in figure 10.1. The two always come together - each special frequency has an associated special pattern. They are sometimes called the eigenvalues and eigenfunctions.

10.2 Modes of a rectangular plate

Next we consider a two dimensional example – a rectangular plate or membrane as illustrated in figure 10.2. At equilibrium, the plate lies in the xz-plane. It has length a along x and length b along z. The displacement of any point in the plate is $y(x, z, t)$. The wave equation for this system is

$$\frac{\partial^2 y}{\partial x^2} + \frac{\partial^2 y}{\partial z^2} = \frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}. \tag{10.4}$$

All four edges of the plate will be fixed, so we have the boundary conditions $y(x = 0, z, t) = y(x = a, z, t) = y(x, z = 0, t) = y(x, z = b, t) = 0$.

We could solve equation (10.4) formally, subject to these boundary conditions. Instead, taking our inspiration from the 1D case studied already, we propose the solution

$$y = A\sin(k_x x)\sin(k_z z)\sin(\omega t). \tag{10.5}$$

Substituting into (10.4) and cancelling common factors we see that this is a good solution provided

$$k_x^2 + k_z^2 = \omega^2/v^2. \tag{10.6}$$

Equation (10.5) already satisfies the boundary conditions at $x = 0$ and $z = 0$. To satisfy the conditions at $x = a$ and $z = b$, we require

$$k_x = n_1\frac{\pi}{a}, \quad k_z = n_2\frac{\pi}{b}, \tag{10.7}$$

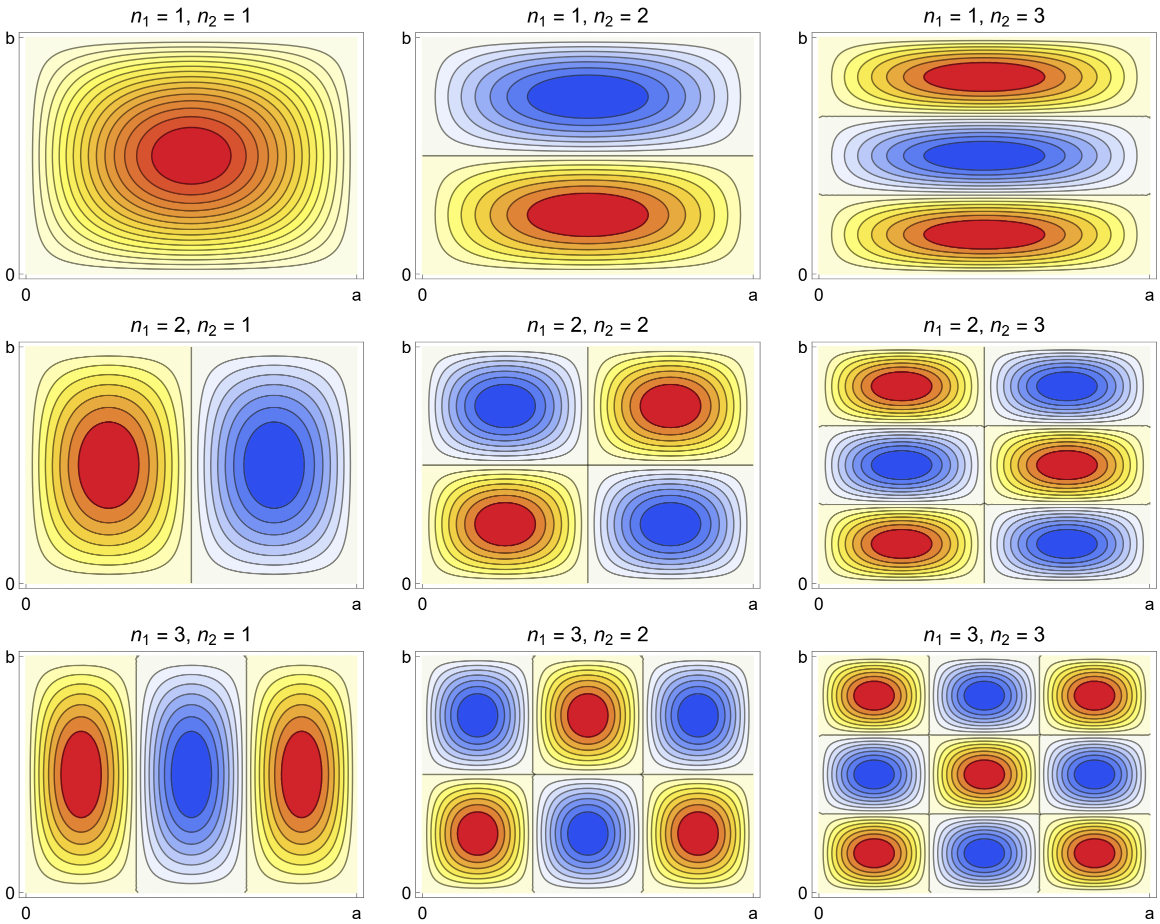

Figure 10.3: Contour plots showing various modes of a vibrating plate held fixed along its four edges. Here, we have chosen the side lengths of the plate to be in the ratio $a/b = 1.4$. The colour map shows the displacement of the plate - red for large positive displacements, blue for large negative displacements, and white for zero. There are nodal lines where the displacement is zero at all times.

where n_1 and n_2 are integers. Thus, we have the normal modes

$$y = A \sin\left(\frac{n_1 \pi x}{a}\right) \sin\left(\frac{n_2 \pi z}{b}\right) \sin(\omega t), \tag{10.8}$$

where

$$\omega = v\sqrt{\left(\frac{n_1\pi}{a}\right)^2 + \left(\frac{n_2\pi}{b}\right)^2}. \tag{10.9}$$

Figure 10.3 shows the first few modes of the plate, shown by making contour plots of equation (10.8). These are the 'patterns' associated with each of the mode frequencies. Bright red and blue regions are antinodes, while the white regions are the nodes. On the 1D string we had nodal points, but on the 2D plate we have nodal lines. Looking at equation (10.9) we see that, unlike the string, the normal mode frequencies of the vibrating plate are not harmonics of one another in general.

10.3 Summary of key points

- When a wave is confined to some region of space, the system can only oscillate sinusoidally at certain special frequencies. These are the normal modes of the system.

- There is a certain pattern of displacement associated with each normal mode.
- For a wave on a string fixed at both ends, the normal mode frequencies and patterns are easily found by requiring that a standing wave fits on the string (length L) with nodes at each end. The modes are described as $y = A\sin(k_n x)\sin(\omega_n t)$ where $k_n = n\pi/L$.
- This generalizes to waves on a rectangular plate (side lengths a, b) fixed along all four edges. The modes are $y = A\sin(k_x x)\sin(k_z z)\sin(\omega t)$ where $k_x = n_1\pi/a$, $k_z = n_2\pi/b$.

Chapter 11

Interference

In this lecture we look at one of the most important topics in physics and one that you will encounter over and over – interference.

11.1 Superposition

Suppose we have found two solutions to the wave equation, $f_1(x,t)$ and $f_2(x,t)$. Is $f(x,t) = f_1(x,t) + f_2(x,t)$ also a solution?

The derivative with respect to x is

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f_1}{\partial x^2} + \frac{\partial^2 f_2}{\partial x^2}.$$

The derivative with respect to t is

$$\frac{\partial^2 f}{\partial t^2} = \frac{\partial^2 f_1}{\partial t^2} + \frac{\partial^2 f_2}{\partial t^2}.$$

Thus

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2} = \left(\frac{\partial^2 f_1}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 f_1}{\partial t^2}\right) + \left(\frac{\partial^2 f_2}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 f_2}{\partial t^2}\right).$$

Each of the brackets on the right hand side is equal to zero, because f_1 and f_2 are solutions of the wave equation. Thus, $f(x,t)$ is also a solution of the wave equation.

Don't be fooled by the apparent triviality of the above maths. It tells us something really important - you can add waves together, and the thing you have is still a wave.

11.2 Intensity

We do not normally measure the amplitude of the wave directly. Instead, we typically have a detector that measures the power. Moreover, it's common for the detector to measure the power averaged over many periods of the wave. Consider visible light waves for example. The oscillation frequency is of order 10^{15} Hz. There are no detectors that can respond anywhere near that quickly, so their output is the average over a very large number of cycles. In section 9.1, we discussed the cycle-averaged power of the wave, and saw that this was proportional to the square of the amplitude – see equation (9.7). We typically talk about the intensity of the wave, especially when we are dealing with light or sound. The intensity, I is defined as the cycle-averaged power per unit area, so it is also proportional to the square of the amplitude. We often use the complex notation to describe waves, e.g. $\psi = Ae^{i(kx-\omega t)}$. The intensity is usually very easy to calculate when written in this form – it is simply

$$I = \alpha|\psi|^2, \tag{11.1}$$

where α is a proportionality constant that depends on the kind of wave we are considering.

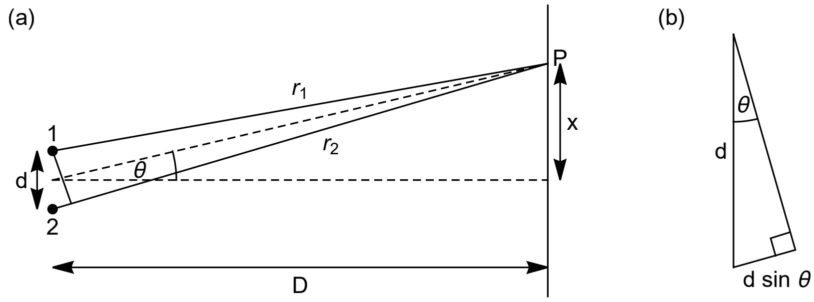

Figure 11.1: Interference of two sources emitting waves of the same amplitude, frequency and phase. (a) Coordinates used to analyze the interference. (b) Expanded region in the vicinity of the two sources showing that the path difference between the two waves is $d \sin\theta$.

11.3 Interference – the main idea

Let's consider two waves

$$\psi_1 = A_1 e^{i(\omega t+\phi_1)}, \tag{11.2a}$$

$$\psi_2 = A_2 e^{i(\omega t+\phi_2)}. \tag{11.2b}$$

Here, the phases ϕ_1 and ϕ_2 will contain the spatial part of the phase as well as any additional phase offsets, for example $\phi_1 = kx_1 + \Phi$. The intensities of each of the waves are $I_1 = \alpha A_1^2$ and $I_2 = \alpha A_2^2$. We add the two waves together to make a total wave

$$\psi = \psi_1 + \psi_2 = \left(A_1 e^{i\phi_1} + A_2 e^{i\phi_2}\right) e^{i\omega t}. \tag{11.3}$$

The corresponding intensity is

$$\begin{aligned} I &= \alpha \left(A_1 e^{i\phi_1} + A_2 e^{i\phi_2}\right)\left(A_1 e^{-i\phi_1} + A_2 e^{-i\phi_2}\right) \\ &= \alpha \left(A_1^2 + A_2^2 + A_1 A_2 (e^{i(\phi_1-\phi_2)} + e^{-i(\phi_1-\phi_2)})\right) \\ &= \alpha \left(A_1^2 + A_2^2 + 2A_1A_2\cos(\phi_1-\phi_2)\right) \end{aligned}$$

Since the intensities of the individual waves are $I_1 = \alpha A_1^2$ and $I_2 = \alpha A_2^2$, we can write this result as

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\phi_1 - \phi_2). \tag{11.4}$$

This equation expresses the key point about interference - the total intensity is not just the sum of the individual intensities. Instead, there is an extra term which is proportional to the cosine of the phase difference between the two waves. When the two waves have the same intensity and the cosine has its maximum value, the total intensity is *4 times* the intensity of a single wave – this is constructive interference. Conversely, when the cosine has its minimum value, the intensity is zero – this is destructive interference. The phase difference $\phi_1 - \phi_2$ can arise in many ways – differences in path lengths, differences in frequencies, sources emitting with different phases – we will see some examples in this lecture and you will meet many other examples throughout your course. No matter how it arises the basic idea is always the same and is expressed by equation (11.4).

11.4 Interference between two point sources

Figure 11.1(a) illustrates the situation we want to analyze. A pair of point sources separated by a small distance d emit waves with the same amplitude, frequency and phase.[1] A screen is placed at

[1] A common way to make this happen is to have a plane wave incident from the left on a pair of small apertures or slits. The wave can only pass through the open area of the slits, and if they are small enough each slit looks like a source of waves.

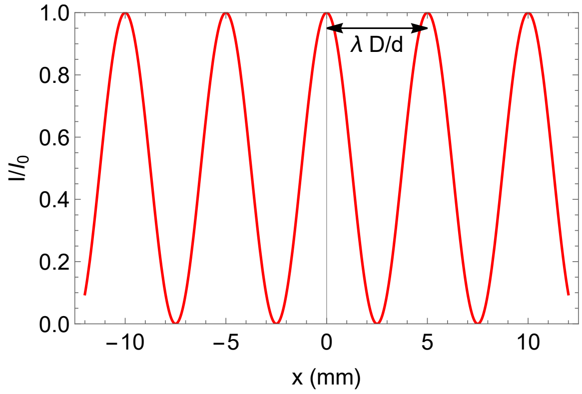

Figure 11.2: Interference fringes due to the pair of sources illustrated in figure 11.1. We have used $\lambda = 500$ nm, $d = 100$ μm, $D = 1$ m, which are typical parameters when looking at a two-slit interference pattern of visible light.

a distance D from the sources, with $D >> d$, and we are interested in the intensity of the waves arriving at each point on the screen. We focus on a particular point P which is a distance r_1 from source 1 and r_2 from source 2. The line from the sources to P makes an angle θ to the horizontal.[2]

Let's think about what happens. When $\theta = 0$ the distance from each of the sources to the observation point is the same. Since the waves leave in phase and the travel distance is the same they must arrive in phase. The waves add up at this point – we have constructive interference. As we move away from this point on the screen, say by increasing θ, the wave from source 2 has to travel a little bit further than the wave from source 1. When that extra distance is half a wavelength, the waves will arrive exactly out of phase and will cancel each other perfectly, giving zero amplitude – we have destructive interference. If we keep going so that the extra distance becomes a complete wavelength, the waves arriving at the screen will be back in phase again and we get the next maximum. We see from figure 11.1(b) that the extra distance for source 2 is $d \sin\theta$. So we expect to get a maximum at angles where $d \sin\theta = n\lambda$, where n is an integer.

Now we'll do the analysis mathematically. The wave arriving at P from sources 1 and 2 are

$$\psi_1 = A\cos(kr_1 - \omega t), \tag{11.5a}$$
$$\psi_2 = A\cos(kr_2 - \omega t). \tag{11.5b}$$

The total wave arriving at P is the sum of these two which is

$$\psi = \psi_1 + \psi_2 = 2A\cos\left(\frac{k(r_2 - r_1)}{2}\right)\cos(kr - \omega t), \tag{11.6}$$

where we have used a trigonometric identity and introduced the mean distance $r = (r_1 + r_2)/2$. Referring to figure 11.1(b), we see that $k(r_2 - r_1) = kd\sin\theta = 2\pi d\sin\theta/\lambda$. So we obtain

$$\psi = 2A\cos\left(\frac{\pi d\sin\theta}{\lambda}\right)\cos(kr - \omega t). \tag{11.7}$$

Now we calculate the intensity, which is proportional to the square of the wave. After averaging over the part oscillating at ω, which is too fast to detect, we find that the intensity can be written as

$$I(\theta) = I_0\cos^2\left(\frac{\pi d\sin\theta}{\lambda}\right), \tag{11.8}$$

[2] d is so small that it won't matter whether we measure θ from the lower source, or the upper source, or half way in between.

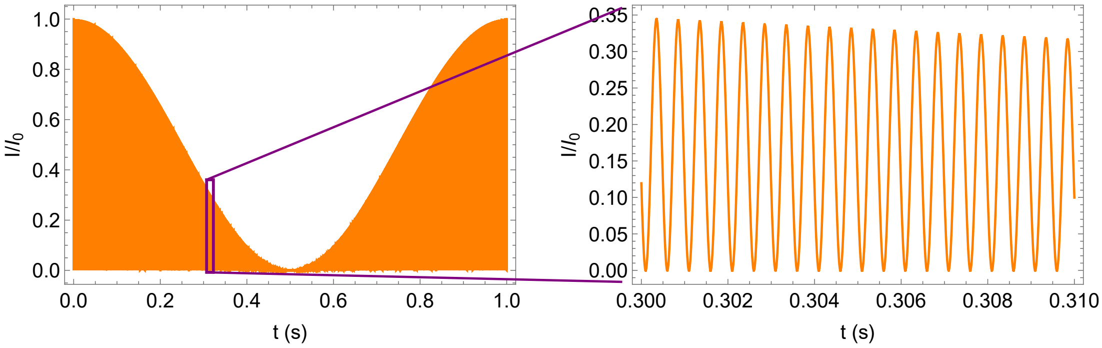

Figure 11.3: Interference of two waves with frequencies 1000 Hz and 1001 Hz. In the left plot we see the beat note between the two waves at their difference frequency. If we take the small portion indicated by the rectangle and magnify it, we get the right plot where we can see the fast oscillation at the average frequency.

where I_0 is the intensity at $\theta = 0$. For small angles, we have $\sin\theta \approx \tan\theta = x/D$. So we can write the above result in terms of the position on the screen, x:

$$I(x) = I_0 \cos^2\left(\frac{\pi d}{\lambda D}x\right). \tag{11.9}$$

We see that the intensity on the screen oscillates with x as a $\cos^2$ function, and we often refer to these as '$\cos^2$ interference fringes'. They are plotted in figure 11.2. The interference fringes have a regular spacing. There is an intensity maximum at the centre. The next maximum occurs when the argument of the $\cos^2$ function is π. Thus, using equation (11.8) and $\sin\theta \approx \theta$, we see that the angular spacing of the fringes is $\Delta\theta = \lambda/d$. Equivalently, from equation (11.9) we can say that the linear spacing of the fringes is $\Delta x = \lambda D/d$. As the two sources are brought closer together, the fringes get further apart.

11.5 Interference between waves with different frequencies

Above, we looked at interference of two waves that have the same frequency and saw that the intensity varies from place to place, but for a fixed point is constant in time. Now we look at interference between waves of different frequencies. As an example, suppose the frequency of one source is 1000 Hz and the frequency of the other is 1001 Hz. We stay at a fixed point in space and ask what happens as a function of time. At some moment in time the waves arrive in phase so we have constructive interference and a big wave. But 0.5 s later one wave will have completed 500 cycles and the other will have completed 500.5. That means that the two waves are a half cycle out of phase, which is the condition for destructive interference. The wave will have no amplitude at this point. After a further 0.5 s the waves will be back in phase again and the wave will be big. If the waves are sound waves, we we will hear the sound oscillate from loud to quiet to loud, and this oscillation will be at the difference frequency (1 Hz in our example). If the waves are visible light, we will see the light blink on and off at the difference frequency. This is the phenomenon known as beats.

Now let's see this mathematically. The situation is the same as before - two sources emitting waves and an observer a large distance away, but now the sources have different frequencies. To keep things simple, let's put the observer at an equal distance (R) from the two sources (that means choosing $\theta = 0$ in figure 11.1). The waves arriving from each of the sources are

$$\psi_1 = A\cos(kR - \omega_1 t), \tag{11.10a}$$
$$\psi_2 = A\cos(kR - \omega_2 t). \tag{11.10b}$$

The total wave arriving at P is the sum of these two which is

$$\psi = \psi_1 + \psi_2 = 2A\cos\left(\frac{(\omega_2 - \omega_1)t}{2}\right)\cos(kR - \bar{\omega}t), \tag{11.11}$$

where $\bar{\omega} = (\omega_1 + \omega_2)/2$ is the average of the two frequencies. Note that kR is just a constant phase offset. So we see a wave oscillating at the average frequency modulated by an envelope that oscillates at the difference frequency. The intensity is proportional to ψ^2, so we have

$$I = I_0\cos^2\left(\frac{(\omega_2 - \omega_1)t}{2}\right)\cos^2(kR - \bar{\omega}t) \tag{11.12}$$

11.6 Summary of key points

- You can add waves together.
- You cannot add intensities together. You have to first add the waves (the quantity that is the solution to the wave equation), and then calculate the square of the resulting amplitude.
- After adding in this way, the total intensity depends on the phase difference between the waves. This is interference.
- When you have two point sources separated by d e.g. due to plane waves incident on a pair of tiny slits, the intensity on a distant screen has a $\cos^2$ distribution. The distance between the maxima is inversely proportional to d.
- When you add two waves that differ in frequency you get an interference in time which are known as beats. The beat frequency is the difference in frequency between the two waves.

Chapter 12

Wave packets

12.1 Adding more waves

In the last lecture we saw that when you add together two waves that have a frequency difference $\Delta\omega$, the interference between them results in a total wave that is modulated at the difference frequency. The resulting wave looks like a train of pulses with period $T = 2\pi/\Delta\omega$. This train of pulses goes on forever, and is like the one shown in the first graph (labelled "2") in figure 12.1. Suppose we wanted to make just a single pulse, which is usually given the name *wavepacket*. How could we do that? We need to arrange things so that we have constructive interference around a particular point in space or time, and destructive interference *everywhere else*. This clearly can't be done by adding two waves because if the two go out of phase after a certain time, they must come back into phase at twice that time (and so on).

Perhaps we need to add more waves. We'll keep the two we already have with frequencies ω and $\omega + \Delta\omega$ but add an extra one with frequency in the middle, $\omega + \Delta\omega/2$. Suppose everything is in phase at time $t = 0$. As time advances, they slip out of phase. At $t = T = 2\pi/\Delta\omega$ two of the waves will be back in phase, but the third one is completely out of phase, so the constructive interference here is incomplete. That means the pulse either side of the central one is suppressed. However, at $t = 2T$ everything is back in phase again. So now we get a pulse train that alternates between a big pulse and a little pulse, like the one labelled "3" in figure 12.1. Let's keep going, adding more and more waves with slightly different frequencies, always spanning the range $\Delta\omega$. So, when there are N waves the frequency of wave i is $\omega + \Delta\omega(i-1)/(N-1))$. Figure 12.1 shows the result of doing this. As more waves are added it takes longer for everything to come back into phase so the big pulses become further separated with more little pulses in between. When there are N waves, there are $N-2$ little pulses between the big ones. You can see that if we carry on with this process, making N very large we will end up with something approaching the single isolated pulse we wanted. Though not quite – in the addition to the big pulse we wanted, we still have the long train of little pulses. Is there a way to get rid of them? Yes – so far we've only used waves of the same amplitude, but of course we could choose different amplitudes for each. That gives us a lot of freedom, and it turns out that if we choose the amplitudes right we can indeed end up with the single pulse we want. Amazingly, you can actually make any function you like by adding sine waves of different frequencies and amplitudes. You will learn how to do that when you study Fourier Analysis later on.

We see from figure 12.1 that adding more waves gives us a better approximation to a single isolated pulse, but the duration of that pulse hardly changes. That's because we fixed the frequency span to always be the same value of $\Delta\omega$. We already know that in the case of just two waves, the duration of each pulse (from one zero to the next) is $2\pi/\Delta\omega$, and we know that fitting more waves into the same range of frequencies doesn't change the pulse duration very much. So we find that if we add lots of waves spanning a frequency range $\Delta\omega$ (often called the *bandwidth*) we get a pulse of duration

$$\Delta t \sim 1/\Delta\omega. \tag{12.1}$$

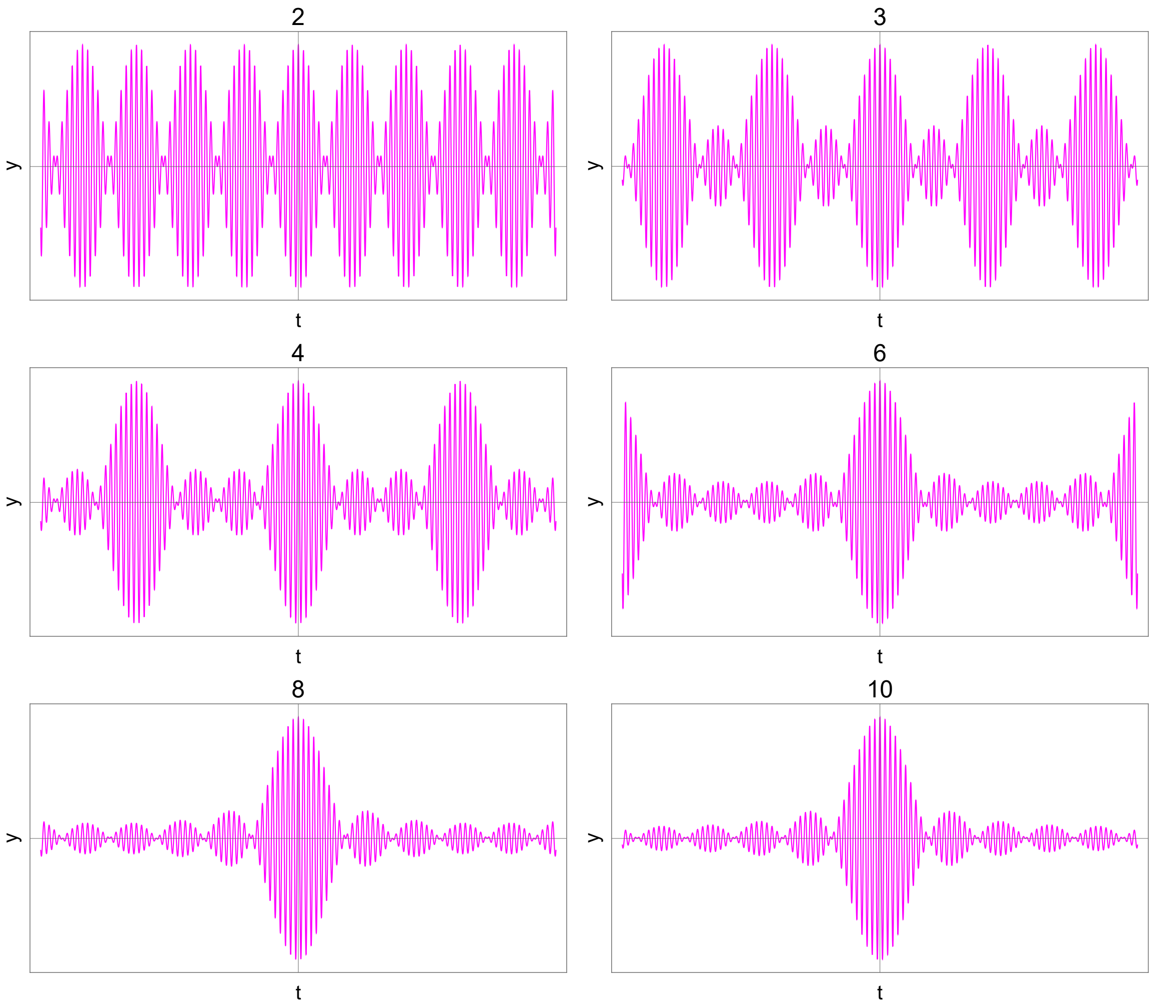

Figure 12.1: Attempting to make a single pulse by adding more and more waves. The number of waves is shown on the top of each plot. The waves have frequencies from ω to $\omega + \Delta\omega$ in equal intervals.

Here the $\sim$ sign means 'of order' – the exact value depends on the details of the shape of the pulse and how it's made and what exactly we mean by the duration. What's important here is not the exact relation but the concept – the pulse duration is inversely proportional to the range of frequencies. For example, if you wish to make a pulse that is 1 ms long it will need to contain a range of frequencies of at least 1 kHz (again, we are not worrying here about where the factors of 2π go).

We have looked at this as a pulse in time, but you could equally look at it as a pulse in space. From that perspective, we are adding waves of different k spanning a range Δk to make a pulse of length Δx. Since $\Delta\omega = v\Delta k$ and $\Delta x = v\Delta t$, we find the relation

$$\Delta x \sim 1/\Delta k. \tag{12.2}$$

You will see these relations again in Fourier Analysis, Optics and Quantum Mechanics. The quantum version is Heisenberg's uncertainty principle.

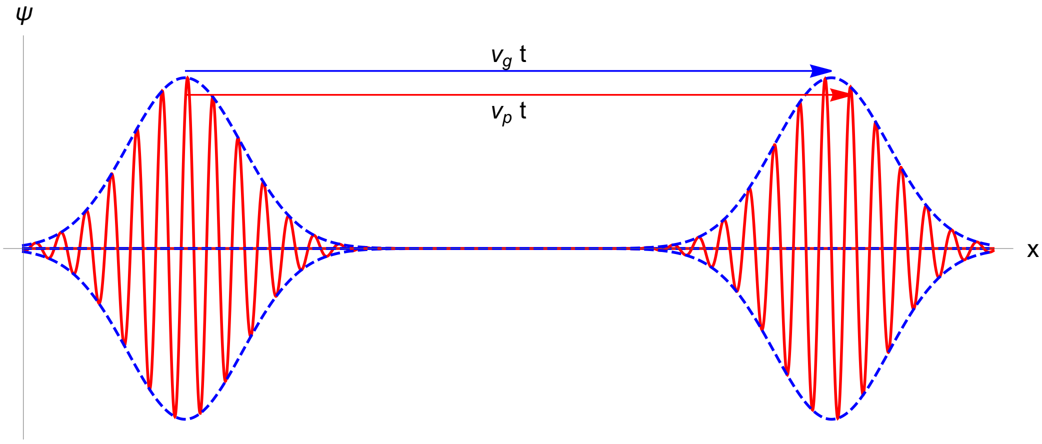

Figure 12.2: Motion of a wavepacket in a time t. The pulse envelope moves a distance $v_\mathrm{g}t$ whereas the underlying fast oscillation moves a distance $v_\mathrm{p}t$.

12.2 Phase and group velocities

Now that we know how to make an isolated pulse – a wavepacket – let's examine what happens to it over time. To keep the analysis simple, we can again start with our pulse train formed by adding just two waves that differ in frequency by $\Delta\omega$ and differ in wavenumber by Δk. We have seen this many times:

$$\begin{aligned}\psi &= A\cos(kx-\omega t) + A\cos((k+\Delta k)x - (\omega+\Delta\omega)t)) \\ &= 2A\cos\left[(k+\Delta k/2)x - (\omega+\Delta\omega/2)t\right]\cos\left[\frac{1}{2}(\Delta kx - \Delta\omega t)\right] \\ &\approx 2A\cos(kx-\omega t)\cos\left[\frac{1}{2}(\Delta kx - \Delta\omega t)\right]. \end{aligned} \tag{12.3}$$

The approximation made in the last step assumes $\Delta k \ll k$ and $\Delta\omega \ll \omega$, which is usually the case of interest.

Suppose we follow one of the peaks of the rapidly oscillating part i.e. the first factor in equation (12.3). It moves along with speed $v = \omega/k$. We have seen that many times already. This speed is called the *phase velocity*. It is the speed that appears in the wave equation and is the only wave speed encountered so far. Now consider one of the peaks of the slowly-varying envelope, i.e. the second factor in equation (12.3). It moves along with a speed $\Delta\omega/\Delta k$. This has to do with how a change in frequency relates to a change in wavenumber. This velocity is called the *group velocity* and is not necessarily the same as the phase velocity. We have looked at the case of just two frequencies, but the same idea holds for any number of frequencies added together to make a wavepacket. In this general case we define the phase velocity

$$v_\mathrm{p} = \frac{\omega}{k}, \tag{12.4}$$

and the group velocity

$$v_\mathrm{g} = \frac{d\omega}{dk}. \tag{12.5}$$

12.3 Dispersion

In the cases considered until now, we have always supposed that the phase velocity of the wave is independent of its frequency (or wavelength). We have $\omega = kv_\mathrm{p}$ with v_p independent of k and ω. In

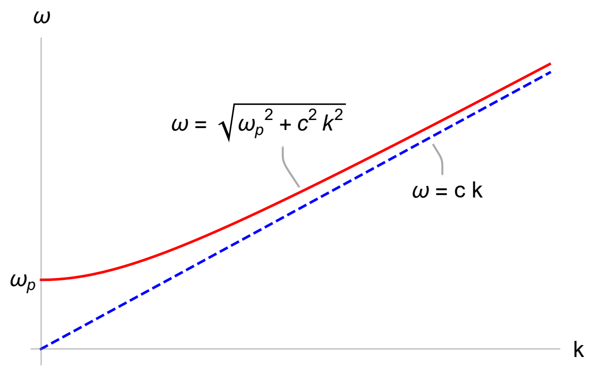

Figure 12.3: Solid red line shows the dispersion relation for light propagating in a plasma. Dashed blue line is the equivalent relation for light in vacuum.

that case it is obvious that $\frac{d\omega}{dk} = v_{\rm p}$, i.e. $v_{\rm g} = v_{\rm p}$ – the group velocity and phase velocity are equal and there is no need to distinguish between them.

However, for many waves travelling in many media the phase velocity has some dependence on the frequency. This is called *dispersion*. Light travelling through glass is a good example – red light travels faster than blue light in the glass. This is, of course, how prisms separate the colours of light. When we have dispersion, the group and phase velocities are not the same. That means that a wavepacket can move at a different speed to the fast oscillations that make up the wavepacket! Figure 12.2 illustrates a propagating wavepacket in the case where $v_{\rm g} \neq v_{\rm p}$. The fast oscillating peaks are not fixed relative to the slow envelope – they move relative to each other as the wavepacket travels. When there is dispersion, the different frequency components that make up a wavepacket all travel with different phase velocities. This causes the shape of the wavepacket to change as it propagates. Often (though not always) the wavepacket spreads out as it travels.

An interesting example of dispersion is an electromagnetic wave propagating through a plasma. A plasma is an ionized gas consisting of free electrons and ions. Because the electrons are so much lighter, they are easier to move than the ions. If the electrons are all displaced from equilibrium in one direction they feel a restoring force due to the Coulomb force of the ions and they execute simple harmonic motion with a natural frequency $\omega_{\rm p}$ which is known as the plasma frequency. For an electromagnetic wave travelling through the plasma, the relationship between wavenumber k and angular frequency ω is

$$\omega = \sqrt{\omega_{\rm p}^2 + k^2c^2}, \tag{12.6}$$

where c is the speed of light in vacuum. This is called the *dispersion relation* of the medium. Figure 12.3 plots this relation and compares it to light in vacuum. The plasma dispersion relation approaches that of vacuum in the limit of high k (short wavelengths), but at low k (long wavelengths) they are very different. The phase velocity is $v_{\rm p} = \omega/k$. Note that for all k the solid line is above the dashed line, which means $v_{\rm p} > c$ – the phase velocity is higher than the speed of light in vacuum! This is also obvious from equation (12.6) since $v_{\rm p} = \sqrt{c^2 + \omega_{\rm p}^2/k^2}$, which is always larger than c. The group velocity is the gradient of the curve, and we can see immediately from figure 12.3 that this is always smaller than the gradient of the dashed line and approaches it asymptotically as k becomes large. Specifically, we have

$$v_{\rm g} = \frac{d\omega}{dk} = \frac{c^2k}{\sqrt{\omega_{\rm p}^2 + k^2c^2}} = \frac{c^2k}{\omega} = \frac{c}{v_{\rm p}}c. \tag{12.7}$$

Since $v_{\rm p} > c$, we see that $v_{\rm g} < c$.

Special relativity tells us that information cannot be transmitted faster than the speed of light. So now we may wonder whether information travels at the phase velocity, which would mean our result violates relativity, or at the group velocity, meaning our result is consistent with relativity.

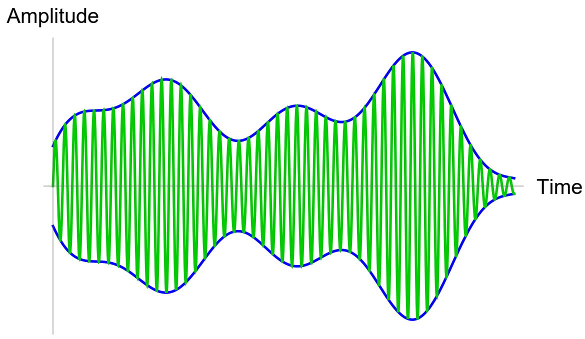

Figure 12.4: Transmitting information by amplitude modulation of a carrier signal. The carrier is the high frequency wave shown in green, and the modulation is the slowly varying envelope shown in blue.

12.4 Sending information

Waves are used all the time to transmit information. The world is full of sound waves and electromagnetic waves all sending information. How does this work? A simple sine wave is uniform for all time so does not contain any information. The information has to be encoded into the wave by modulating (changing) either its amplitude (AM) or its frequency (FM). There is typically a high frequency carrier wave whose amplitude or frequency is modulated at a much lower frequency. This modulation contains the information. Figure 12.4 shows an example of an amplitude modulated carrier wave. An AM radio station would send a signal of this kind. To receive it, the radio is tuned to the frequency of the carrier and then demodulates the signal in order to retrieve the slowly varying amplitude which contains the information being transmitted. Note that this slowly varying amplitude – the information – travels at the group velocity.

12.5 Summary of key points

- We can make a wavepacket of duration Δt by adding together many waves whose frequencies span an angular frequency range $\Delta\omega \approx 1/\Delta t$.
- A wavepacket consists of a rapidly oscillating part which travels at the phase velocity and a slowly-varying envelope that travels at the group velocity.
- The phase velocity is $v_{\mathrm{p}} = \omega/k$.
- The group velocity is $v_{\mathrm{g}} = d\omega/dk$.
- Different frequencies can have different phase velocities. This is called dispersion.
- An equation relating ω and k is know as a dispersion relation. When the dispersion relation is linear, the phase and group velocities are the same. Otherwise, they are different.
- Dispersion causes a wavepacket to change shape as it propagates. Often, the wavepacket spreads out.
- Information is typically transmitted by modulating the amplitude or frequency of a carrier wave.
- Information travels at the group velocity.