Oscillating Systems

Natural Sciences

Physics Part 1A

J. S. Biggins

Cavendish Laboratory
University of Cambridge

Michaelmas 2014

Contents

1 Course Details

1.1 Synopsis

Simple harmonic motion (SHM): equation of undamped oscillation for a mass on a spring, its solution, relative phases of displacement, velocity and force. Approximations of oscillating systems to SHM: simple pendulum, torsional pendulum. Energy in SHM: vibration of two masses joined by a spring, quantum well.

Phasor diagrams: superposition of oscillations, beats, amplitude modulation.

SHM using complex numbers: Curves of time-dependence for an oscillator, amplitude, frequency, angular frequency and phase.

Damped oscillations: amplitude, energy decay. quality factor.

Forced oscillations: qualitative frequency response and resonance.

Revision of electrical circuits: voltage, current and charge in circuits, electrical resistance, Kirchhoff's Laws, resistors in series and parallel. Inductors and capacitors. Circuits with exponential decays: charge and discharge of a capacitor through a resistor, decay of current through an inductor.

Oscillations in electrical circuits and complex impedance: Oscillation in an LC circuit, relative phases of voltages, charge and currents, energy in an LC circuit. Complex current and voltage in resistors, capacitors and inductors. Complex impedance. Electrical resonance in an LCR circuit, simple filters, bandwidth, Q factor. Relationship of behaviors seen in electrical systems to those of mechanical systems. Concept of mechanical impedance.

1.2 Resources

"Understanding Physics", Mansfield M & O'Sullivan, *(2nd edition), (Wiley)*.
"Physics for Scientists and Engineers", Tipler P A & Mosca G, *(6th Edition, Extended version), (Freeman 2008)*.
"Fundamentals of Physics", Halliday D, Resnik R & Walker J, *(Extended (8th) Edition), (Wiley)*.

"A Cavendish Quantum Mechanics Primer", Warner M, Cheung ACH, & *(Periphyseos Press)*. Useful discussion of simple harmonic motion and relevant mathematics, including complex numbers, and accessible quantitative treatment of quantum mechanics using very similar equations.

Lots of extra questions and resources at `https://isaacphysics.org`, a Cambridge designed website with maths content on vectors, trigonometry, complex numbers and differential equations, and physics content on simple harmonic motion, motion in a circle and (soon!) circuits. Designed to help bridge the gap between school and university level physics and mathematics.

1.3 Practicalities

There are twelve lectures in the course on Oscillating Systems which will take place during the second half of the Michaelmas term. They will take place during weeks 5-8 inclusive, on Fridays, Mondays and Wednesdays, between 09:00 and 09:50.

There will an Example Sheet containing 23 problems issued to cover this course.

The examples sheet and notes can be downloaded from the NSTIA Physics Teaching information Service website at:

`http://www-teach.phy.cam.ac.uk/`

Hardcopies of the examples sheets and notes are also available from the filing cabinets outside the Pippard Lecture Theatre in the Cavendish.

For those who are interested, there are some **strictly optional** additional problems – less structured and usually of a slightly more mathematical nature – which can also be downloaded from the NSTIA Physics TiS website.

If you find any errors please contact the lecturer
Dr John Biggins (`jsb56@cam.ac.uk`).

2 Oscillations

2.1 Periodicity, Frequency and Angular Frequency

A system is said to be oscillating if it varies its state in a repetitive way. Examples abound, not just in science, but in every walk of life: the UK could be said to oscillate between day and night, winter and summer or between cloud and sunshine. Many oscillations are not just repetitive but periodic, meaning that the repetitions happen with a regular time interval. The oscillations between day and night, or between winter and summer are periodic, repeating every 24 h and every 1 yr respectively. The weather is rather less predictable. More formally, imagine a system characterized by $x(t)$ which varies in time as plotted in fig. 1. The system is periodic if, for some time T, $x(t) = x(t+T)$ for all times t. Of course, if this is true for T it is also true for $2T$, $3T$, $4T$ and so on, so we use T to denote the shortest time-period for for which $x(t) = x(t+T)$ and call it the fundamental period.

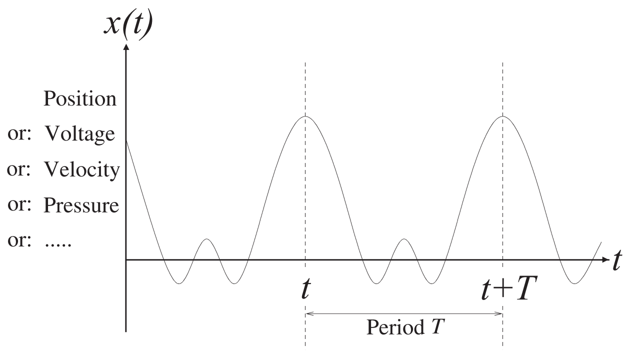

Figure 1: Periodic motion with fundamental time period T.

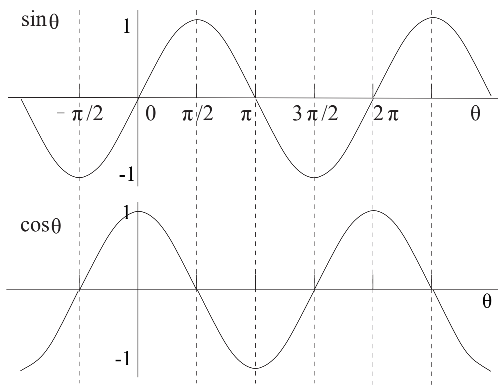

Figure 2: The sin and $\cos$ functions.

Instead of period, we often think of an oscillation's frequency, $\nu = 1/T$. If T is in seconds then ν is in s^{-1}, known as Hertz (Hz) but best though of as "per-second". If an oscillation has a time period T =0.1 s, it has frequency $\nu = 1/0.1 = 10$ Hz because it repeats ten times per-second.

To think about oscillations we need periodic functions. Two such functions, $\sin(\theta)$ and $\cos(\theta)$, plotted in fig. 2, are already familiar. Both repeat when θ increases by 2π because the argument, θ, is an angle and adding a whole turn (2π) to an angle brings you back to where you started. A particle whose position is given by $x = \cos(\omega t)$ is oscillating in time between $x = \pm 1$. The oscillation will repeat whenever ωt increases by 2π, so the time period is $T = 2\pi/\omega$ and the frequency is $\nu = \omega/(2\pi)$. The quantity ω is called the angular frequency and has units of radians/second. It is closely linked to angular velocity since, if we interpret ωt as an angle, ω is its velocity.

2.2 Why study oscillations?

We study oscillations in such detail because they are ubiquitous. Below is a list of oscillators from many sectors of physics and engineering and with frequencies spanning 37 orders of magnitude.

Oscillator	**Frequency /Hz**
Deformed nuclei	10^{21}
Light emitting atom	10^{15}
Molecular vibrations	10^{14}
Mobile Phone transmissions	10^{9}
TV transmitter	10^{8}
Radio (medium wave)	10^{6}
Musical instruments	$10^{2\to 4}$
Heart Beat	10^{0}
Old Faithful Geyser	10^{-4}
Tides	10^{-5}
Solar Activity Cycle	10^{-9}
Axial precession of the Earth	10^{-12}
Solar galactic rotation	10^{-16}
Cyclic Universe(!????)	?

Amazingly, much of this oscillatory behavior can be described by a handful of ideas encoded in simple equations. Obviously such ideas are worth studying in detail.

3 Simple Harmonic Motion

3.1 Mass on a spring

We start with a very simple oscillating system, a mass on a horizontal spring, sketched in fig. 3. We assume that the surface is frictionless and the motion horizontal, so the only force on the mass we need consider is that from the spring. If we displace the mass by x from its equilibrium position the spring pulls it back and, the further we pull it, the harder the spring pulls back. This spring behavior is encoded in Hooke's law, which states that a stretched spring exerts a restoring force proportional to its extension:

$$F = -kx. \tag{1}$$

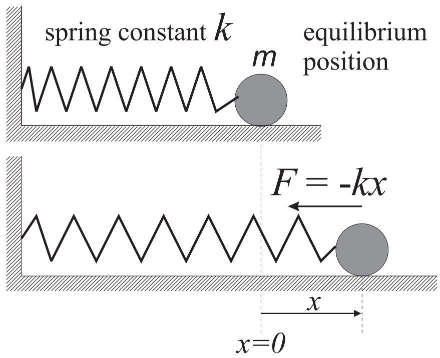

Figure 3: A mass on a horizontal spring. Top: The mass in it's equilibrium position with the spring unstretched. Bottom: The mass is displaced from equilibrium by an amount x, leading the spring to exert a force $-kx$ pulling it back towards the equilibrium point.

The constant of proportionality, k, depends on the spring. The minus sign encodes that the force acts to reduce x, that is, the spring pulls the mass back towards the equilibrium point.

We now ask what happens if we draw the mass out to a position $x = a_0$ and release it. What happens next is determined by Newton's second law,

$$F = ma, \tag{2}$$

where F is the force on an object, m is its mass and a is its acceleration. Applying this law qualitatively to the mass on a spring gives the sequence of events shown in fig. 4. Initially the mass is at rest but the spring is pulling it back towards the equilibrium point, causing it to accelerate towards the equilibrium point. As the mass moves towards the equilibrium point the spring

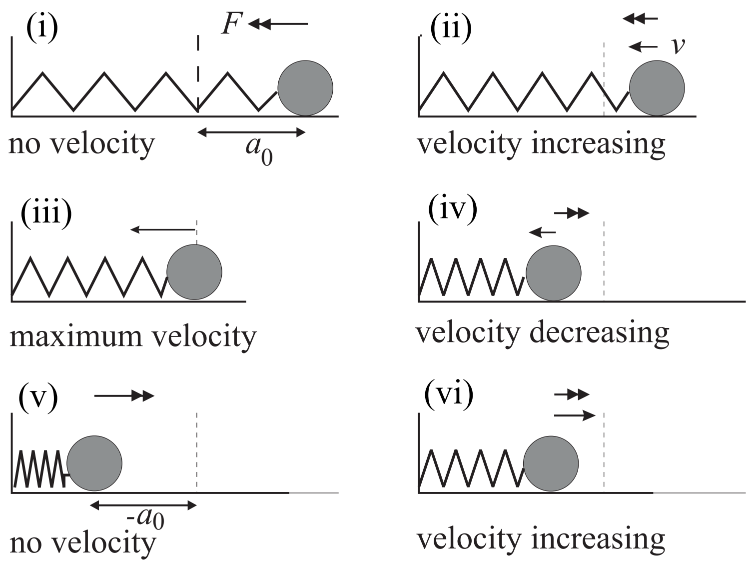

Figure 4: Snapshots of a mass oscillating on a spring. The mass is drawn back to $x = a_0$ then released, leading to oscillations. Double headed arrow indicates force, single headed arrow indicates velocity. Further description in the main text.

remains stretched, so it continues to pull towards the equilibrium point, and the mass continues to accelerate. When the spring reaches the equilibrium point the spring is no longer stretched so there is no force and the mass is moving at its maximum speed. It passes straight through the equilibrium point putting the spring into compression. This causes the spring to again push the mass back towards the equilibrium point, but now, this force is opposite to velocity, so it slows the mass down. Eventually the mass becomes stationary, but the mass is now far from the equilibrium point and the spring is deep in compression, pushing the mass back towards the equilibrium point, and the whole process starts again. The mass oscillates on the spring.

The above discussion reveals the two key ingredients for oscillatory behavior: we need a restoring force that pulls our system back towards its equilibrium point and inertia (i.e. resistance to change in velocity) so that when the system reaches the equilibrium point, its motion ensures that it passes through to the other side. In the above case, the spring provides the restoring force and the mass provides the inertia.

Of course $F = ma$ allows us to go beyond this qualitative analysis and predict the precise motion of the mass. The force acting on the mass is just the spring force, $F = -kx$, and its acceleration is simply $a = \ddot{x}$ (we use dots to denote derivatives with respect to time), so we have

$$-kx = m\ddot{x}. \tag{3}$$

This second order differential equations can be rearranged to

$$\ddot{x} = -\frac{k}{m}x, \tag{4}$$

which encodes that acceleration is proportional but opposite to displacement, the essential ingredient for simple-harmonic-motion.

3.2 Solving the SHM equation

Often the best way to solve differential equations is simply to guess the answer. Here we want an oscillating function that takes a maximum value of a_0 at $t = 0$, so we try $x = a_0 \cos(\omega t)$.

Substituting this into the equation of motion (eqn 4) we get

$$-\omega^2 a_0 = -\frac{k}{m} a_0 \tag{5}$$

which is true, provided

$$\omega = \sqrt{\frac{k}{m}}, \tag{6}$$

so the motion of our spring is given by

$$x = a_0 \cos\left(\sqrt{\frac{k}{m}}\, t\right). \tag{7}$$

We have just shown that our mass on a spring does indeed oscillate, and that it does so with angular frequency $\omega = \sqrt{\frac{k}{m}}$. The full motion is plotted in fig. 5

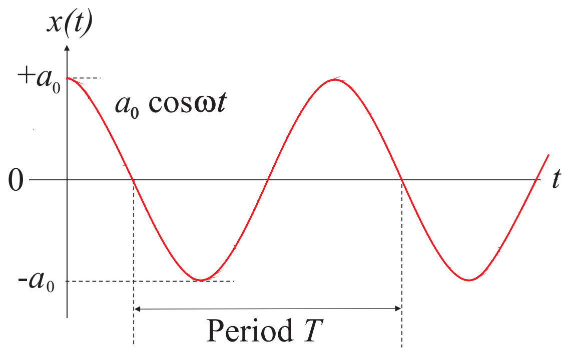

Figure 5: A mass on a spring is drawn back to $x = a_0$ and released from rest. This plot shows the resultant position of the mass, $x(t)$, oscillating as a function of time. The angular frequency of the oscillation is $\omega = \sqrt{k/m}$, so the time period is $T = 2\pi\sqrt{m/k}$.

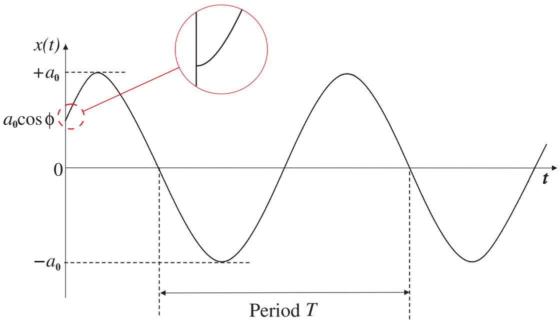

Figure 6: Plot of the full solution to the SHM equation, allowing that the oscillation need not start with zero velocity and maximum displacement. The inset shows a more realistic trajectory for at $t = 0$ since we cannot actually accelerate a mass to finite velocity instantaneously, but we will ignore this complication.

Many other systems produce equations of motion like eqn 4, but with different constants in place of k/m. We can write any such equations as

$$\ddot{x} + \omega_0^2 x = 0 \tag{8}$$

where the value of ω_0 depends on the specifics of the system, for the mass on a spring $\omega_0^2 = \frac{k}{m}$. This is the fundamental equation of SHM. We know it is solved by $x(t) = a_0 \cos(\omega_0 t)$, however this cannot be the complete solution as it has maximum displacement at $t = 0$, but if we had started the oscillations by giving the mass a kick rather than displacing it, we would instead need a solution with zero displacement at $t = 0$. With the right choice of displacement and velocity we could start the oscillations at any point in the cycle, so we consider the same solution but offset in time:

$$x(t) = a_0 \cos(\omega_0 t + \phi). \tag{9}$$

Substituting this into the fundamental equation of SHM, eqn 8, gives

$$-\omega_0^2 a_0 \cos(\omega_0 t + \phi) + \omega_0^2 a_0 \cos(\omega_0 t + \phi) = 0, \tag{10}$$

which is indeed true. Eqn. 9 is the most general solution to the fundamental equation of SHM: any system obeying eqn 8 will undergo a motion of this form, with a_0 and ϕ determined by how the oscillation is started. The displacement x varies between $\pm a_0$, so a_0 is the amplitude of the oscillation. The constant ϕ is called the phase-constant, and determines which point in the cycle the oscillation starts at. Setting $\phi = -\pi/2$ gives $x = a_0 \sin(\omega_0 t)$, which is appropriate when the motion is started at $t = 0$ by giving the mass a kick. A generic oscillation is plotted in fig. 6.

Although eqn. 9 is the full solution to the SHM equation, it is sometimes convenient to write it in a different form,

$$x(t) = A\cos(\omega_0 t) + B\sin(\omega_0 t). \tag{11}$$

This form also has two constants, A and B, which we can use to specify the motion of the system at $t = 0$. We can relate the two forms of the SHM solution by expanding the first form using the trig angle-addition formula to get

$$a_0\cos(\omega_0 t + \phi) = a_0\cos(\phi)\cos(\omega t) - a_0\sin(\phi)\sin(\omega t), \tag{12}$$

which matches the second form of the solution provided

$$A = a_0\cos(\phi) \quad B = -a_0\sin(\phi). \tag{13}$$

We may also wish to find the first form from the second, which we do by noting that

$$A^2 + B^2 = a_0^2\cos^2(\phi) + a_0^2\sin^2(\phi) = a_0^2(\cos^2(\phi) + \sin^2(\phi)) = a_0^2, \tag{14}$$

and

$$\frac{B}{A} = -\frac{a_0\sin(\phi)}{a_0\cos(\phi)} = -\tan(\phi). \tag{15}$$

Therefore, if we know A and B, we can calculate a_0 and ϕ as

$$a_0 = \sqrt{A^2 + B^2}, \qquad \phi = \arctan(-B/A). \tag{16}$$

3.3 Why *harmonic* motion?

The frequency of simple harmonic oscillations is independent of their amplitude. For a mass on a spring the angular frequency is $\omega = \sqrt{k/m}$, irrespective of a_0; however far we draw the mass back, it oscillates at the same rate when we release it. As the oscillations die down, their amplitude will diminish but their frequency will not. This is why these oscillations are called harmonic: they have a characteristic frequency. If that frequency lies in the audible range we will be able to hear the oscillation, and it will always have the same pitch, whether it is loud or quiet.

Not all oscillations are harmonic. Consider a bouncing ball. If I drop a ball from a height $z = h$ above the floor at $z = 0$, it will then fall under gravity, feeling a downwards weight force $-mg$. Applying Newton's second law gives

$$m\ddot{z} = -mg. \tag{17}$$

This is a motion with constant acceleration, which you have studied previously. Solving gives $z = h - \frac{1}{2}gt^2$, so the ball reaches the floor at $t = \sqrt{2h/g}$ and, if the ball bounces elastically, it will get back to its original height at $t = 2\sqrt{2h/g}$. It is an oscillator with time period $T = 2\sqrt{2h/g}$, which is dependent on h: if I drop the ball from higher, the time between bounces is longer. As a bouncing ball dissipates energy the frequency of its bounces increases. In general we have no right to expect frequency and amplitude of oscillations to be independent. It is a remarkable feature of SHM that in this case they are.

3.4 Position velocity and acceleration

We recall the displacement in SHM is given by

$$x(t) = a_0 \cos(\omega_0 t + \phi). \tag{18}$$

We can take a time derivative of this to find the velocity:

$$\dot{x}(t) = -\omega_0 a_0 \sin(\omega_0 t + \phi) \tag{19}$$
$$= \omega_0 a_0 \cos\left(\omega_0 t + \phi + \frac{\pi}{2}\right). \tag{20}$$

The velocity also oscillates with angular frequency ω_0, but it does so between $\pm a_0\omega$ and is $\pi/2$ (quarter of a cycle) ahead of the displacement. This makes sense if we think about the mass on a spring: the velocity is maximum when the displacement is zero. Quarter of a cycle later, the displacement is maximum, so it is quarter of a cycle behind velocity.

A time derivative of the velocity gives the mass's acceleration

$$\ddot{x}(t) = -\omega_0^2 a_0 \cos(\omega_0 t + \phi) \tag{21}$$
$$= \omega_0^2 a_0 \cos(\omega_0 t + \phi + \pi). \tag{22}$$

Acceleration also oscillates with angular frequency ω_0, but between $\pm a_0\omega^2$ and is π (half a cycle) ahead (or behind, for half a cycle these are the same thing) of the displacement. The acceleration behaves as the negative of displacement, exactly as the SHM equation requires. In terms of the spring, acceleration is always opposite to displacement because the spring always pulls the mass back back towards the equilibrium point. An example of the displacement, velocity and acceleration of a particle undergoing SHM is shown in fig. 7.

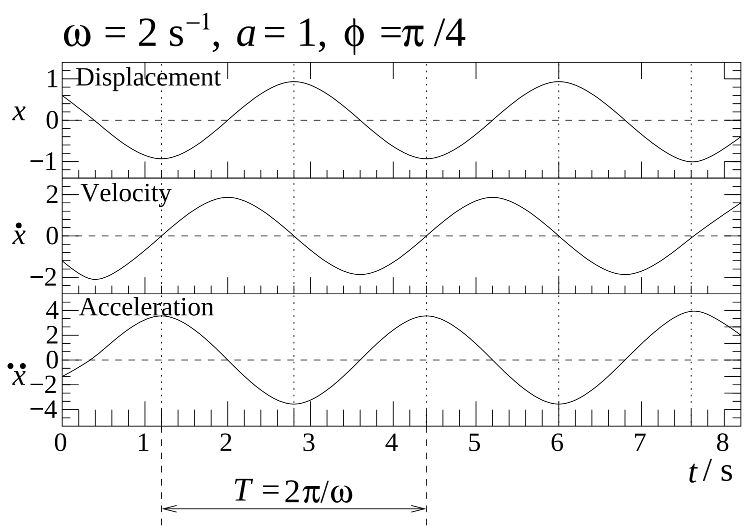

Figure 7: Displacement, velocity and acceleration of a particle undergoing SHM.

3.5 SHM Examples

We now look at some more examples of systems displaying simple harmonic motion.

3.5.1 Mass on a spring with gravity

A massless[1] spring stretches by 18 mm when a 2.8 kg mass is suspended vertically from one end. How much mass should be attached to the spring to make the frequency of oscillation $\nu = 3\,\text{Hz}$?

We first draw a diagram, seen in fig. 8. In the first stage we have force equilibrium between the gravity and the spring force

$$mg = kx. \tag{23}$$

Substituting in $m = 2.8$kg, $x = 18$mm and $g = 9.81$, we get $k = 1526\text{Nm}^{-1}$. In the second stage we have a different mass but the same spring. In equilibrium the spring extends by x_0 such that the spring and weight forces again balance:

$$-kx_0 + mg = 0. \tag{24}$$

If the mass is displaced from equilibrium by an amount x_1, the total displacement is $x_0 + x_1$ so Newton's second gives

$$-k(x_0 + x_1) + mg = m\ddot{x}_1. \tag{25}$$

However, since $-kx_0 = mg$, this reduces to

$$\ddot{x}_1 + \frac{k}{m}x_1 = 0, \tag{26}$$

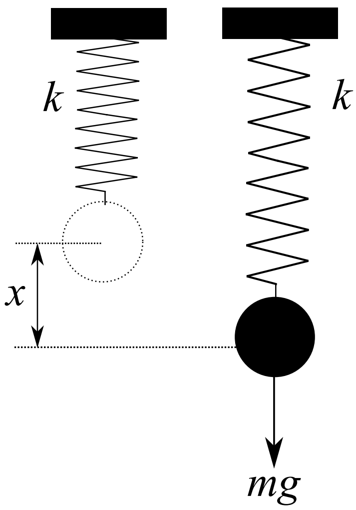

Figure 8: Left: A mass is held at the end of a spring. Right: When it is released, the spring stretches by x.

which is the fundamental equation of SHM, with angular frequency $\omega_0^2 = \frac{k}{m}$, and frequency $\nu = \omega_0/(2\pi)$. Gravity has not changed the frequency of oscillations. Putting in our previously calculated value for k, and $\nu = 3$Hz, we need $m = 4.3$kg.

3.5.2 Mass on two springs

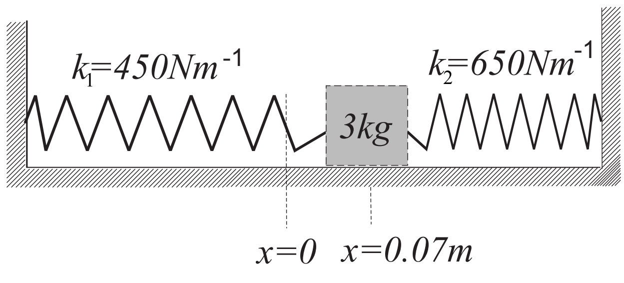

Figure 9: A mass oscillating under the influence of two springs.

A 3.0 kg block is attached between two horizontal springs and placed on a frictionless surface as shown in fig. 9. Neither spring is strained when the block is positioned at the equilibrium position $x = 0$. The block is now displaced a distance of 0.07 m in a direction along the longitudinal axes of the springs and released from rest.

- *Determine the angular frequency of the system.*
- *At what time does the mass first cross the point $x = 0$?*

[1]All springs in the course are massless.

- *What is the speed of the block as it passes through $x = 0$?*

If the mass is displaced to a position x both springs are stretched by x so the total restoring force is $k_1 x + k_2 x = (k_1 + k_2)x$. Newton's second law gives

$$-(k_1 + k_2)x = m\ddot{x} \tag{27}$$

$$\implies \ddot{x} + \frac{k_1 + k_2}{m}x = 0, \tag{28}$$

the fundamental equation of SHM, with angular frequency $\omega_0^2 = \frac{k_1+k_2}{m}$. Putting in our values for m, k_1 and k_2 gives $\omega_0 = 19.15\text{rad s}^{-1}$. The general solution is

$$x = a_0 \cos(\omega_0 t + \phi), \quad \dot{x} = -a_0\omega_0 \sin(\omega_0 t + \phi). \tag{29}$$

Since the mass is released from rest with displacement 0.07m, we need $x(0) = 0.07$m and $\dot{x}(0) = 0$, which requires $\phi = 0$ and $a_0 = 0.07$m, so the displacement and velocity are

$$x = a_0 \cos(\omega_0 t) \quad \dot{x} = -a_0\omega_0 \sin(\omega_0 t). \tag{30}$$

The displacement is first zero when $\omega_0 t = \pi/2$, requiring $t = \pi/(2\omega_0) = 0.082$s. Since $\sin(\pi/2) = 1$, the velocity at this point is just $-a_0\omega_0 = -1.34\text{ms}^{-1}$.

3.5.3 The Hydrometer

A prismatic hydrometer of mass m and cross-sectional area A floats in a fluid of density ρ with its axis in the vertical direction, as depicted in fig. 10. It is then displaced vertically and released. Find the frequency of the resulting oscillations.

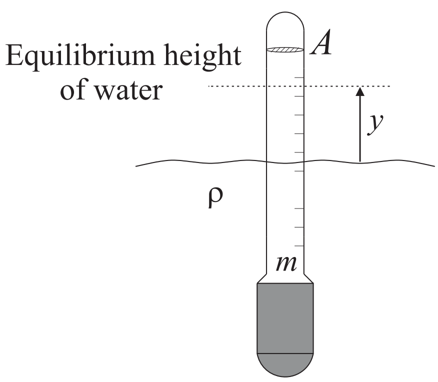

Figure 10: A floating hydrometer.

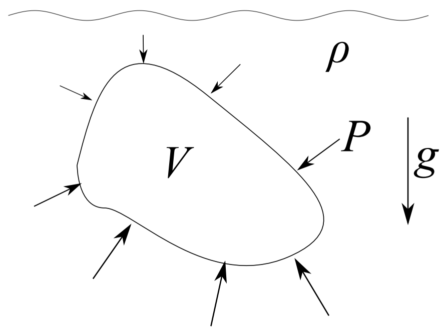

Figure 11: Pressure forces on an object with volume V submerged in a fluid of density ρ in a gravitational field g. The pressures are higher at the bottom of the object, leading to a net upwards buoyancy force.

Before tackling the hydrometer we need to understand buoyancy forces. Imagine a light hollow object of volume V submerged in a fluid of density ρ and in a gravitational field g, sketched in fig. 11. The fluid's pressure P pushes in on the object and, since the pressure is greater deeper in the fluid, this gives an upwards buoyancy force. To calculate its size we imagine the object is filled with fluid. We then have a homogeneous fluid, which is certainly in equilibrium, so the buoyancy force must equal the object's weight, $\rho g V$. Any object with the same shape will experience the same pressure field, and the same buoyancy force $\rho g V$. If this is greater than the objects weight it will float, otherwise it will sink. We have just deduced Archimedes' principle: *The buoyancy force on a partially or fully submerged body is equal to the weight of the volume of fluid it displaces.*

Returning to our hydrometer, it first floats in equilibrium between buoyancy and gravity,

$$mg = \rho g V, \tag{31}$$

where m is hydrometer's mass and V is its submerged volume. If we displace the hydrometer by y from this equilibrium we change the submerged volume by Ay, so Newton's second law gives

$$\rho g(V - Ay) - mg = m\ddot{y}. \tag{32}$$

However, using eqn. 31, this reduces to

$$\ddot{y} + \frac{\rho g A}{m} y = 0, \tag{33}$$

which is the fundamental equation of SHM, with angular frequency $\omega_0^2 = \frac{\rho g A}{m}$.

3.5.4 Simple Pendulum

A simple pendulum consists of a mass m hung from a fixed pivot by a light string of length l. The mass is drawn back to an angle θ and released. What is the frequency of oscillation?

We first draw a diagram, fig. 12, and consider the mass's motion. It follows a circle so, when it is at θ, it has traversed an arc-length $s = l\theta$. Its velocity is $\dot{s} = l\dot{\theta}$ and its acceleration is $\ddot{s} = l\ddot{\theta}$. Two forces act on the mass, gravity and tension T. Velocity and acceleration are both perpendicular to the string, so resolving forces perpendicular to the string and applying Newton's second law, we have

$$-mg\sin(\theta) = ml\ddot{\theta}. \tag{34}$$

For small angles $\sin(\theta) \approx \theta$, so this is approximately

$$-mg\theta = ml\ddot{\theta} \tag{35}$$

$$\implies \ddot{\theta} + \frac{g}{l}\theta = 0, \tag{36}$$

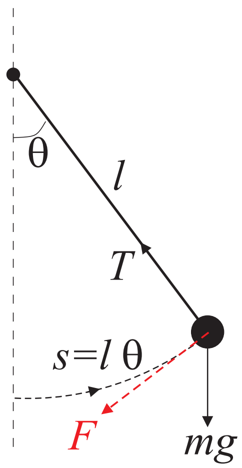

Figure 12: A simple pendulum.

which is the fundamental equation of SHM, with angular frequency $\omega_0^2 = g/l$. This is why it is important for pendulum clocks to work at low amplitude: otherwise the frequency depends on amplitude and the clock will not keep good time.

3.6 Energy in Simple Harmonic Motion

We now return to the oscillations of a horizontal mass on a spring. This motion involves two types of energy, the mass's kinetic energy and the spring's the potential energy. We already know the kinetic energy is given by

$$KE = \frac{1}{2}mv^2 = \frac{1}{2}m\dot{x}^2. \tag{37}$$

To calculate the potential energy we imagine stretching the spring very slowly. To stretch the spring from X to $X + \mathrm{d}X$, we must apply a force kX, so we do work $\mathrm{d}W = kX\mathrm{d}X$. This energy is stored in the spring, and can be recovered if we unstretch it. To calculated the total energy stored in a spring stretched to x we must sum these contributions using an integral

$$PE = \int_0^x kX\mathrm{d}X = \frac{1}{2}kx^2. \tag{38}$$

We recall that a mass on a spring executes simple harmonic motion with

$$x(t) = a_0 \cos(\omega_0 t + \phi) \qquad \dot{x}(t) = -\omega_0 a_0 \sin(\omega_0 t + \phi), \tag{39}$$

where $\omega_0^2 = k/m$. In such a motion the kinetic and potential energy both oscillate:

$$\begin{aligned} KE &= \tfrac{1}{2}ma_0^2\omega_0^2 \sin^2(\omega_0 t + \phi) \\ &= \tfrac{1}{2}ka_0^2 \sin^2(\omega_0 t + \phi) \qquad \text{using } \omega_0^2 = k/m \\ PE &= \tfrac{1}{2}ka_0^2 \cos^2(\omega_0 t + \phi). \end{aligned} \tag{40}$$

The total energy of the system is the sum of these two contributions,

$$\begin{aligned} E = \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}kx^2 &= \tfrac{1}{2}ka_0^2 \sin^2(\omega_0 t + \phi) + \tfrac{1}{2}ka_0^2 \cos^2(\omega_0 t + \phi) \\ &= \tfrac{1}{2}ka_0^2 \left(\sin^2(\omega_0 t + \phi) + \cos^2(\omega_0 t + \phi)\right) \\ &= \tfrac{1}{2}ka_0^2, \end{aligned} \tag{41}$$

which is independent of time. Although the kinetic energy and the potential energy both oscillate, the total energy is constant. We can see how this works by plotting the kinetic energy and potential energy as functions of displacement, as seen in fig. 13. At maximum displacement the mass is not moving but the spring is very stretched and all the energy is potential. As the mass moves inwards the spring's potential energy is released and the mass gains a corresponding amount of kinetic energy. At zero displacement the spring is not stretched but the mass is moving at maximum velocity and all the energy is kinetic. Energy oscillates between kinetic and potential, but the sum is conserved.

It is also informative to ask how the potential and kinetic energies vary in time, which we plot in fig. 14. The kinetic energy is maximal when the displacement is zero, and the potential energy when displacement is $\pm a_0$. Both happen twice a cycle, so the energy oscillates between potential and kinetic at twice the oscillation's frequency. Mathematically this is because, though the trigonometric functions in (eqn 40) still oscillate at ω_0, they are squared. The function $\sin^2 \omega_0 t$ is maximum whenever $\sin \omega_0 t = \pm 1$, so it oscillates twice as fast. We can verify this by using the identities $\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta)$ and $\sin^2\theta = \frac{1}{2}(1 - \cos 2\theta)$. Applying these to eqn. 40 gives

$$\begin{aligned} KE = \tfrac{1}{2}ka_0^2 \sin^2(\omega_0 t + \phi) &= \tfrac{1}{2}ka_0^2\tfrac{1}{2}\left(1 - \cos(2\omega_0 t + 2\phi)\right) \\ &= \tfrac{1}{4}ka_0^2 - \tfrac{1}{4}ka_0^2 \cos(2\omega_0 t + 2\phi) \end{aligned} \tag{42}$$

$$\begin{aligned} PE = \tfrac{1}{2}ka_0^2 \cos^2(\omega_0 t + \phi) &= \tfrac{1}{2}ka_0^2\tfrac{1}{2}\left(1 + \cos(2\omega_0 t + 2\phi)\right) \\ &= \tfrac{1}{4}ka_0^2 + \tfrac{1}{4}ka_0^2 \cos(2\omega_0 t + 2\phi), \end{aligned} \tag{43}$$

which clearly oscillate with angular frequency $2\omega_0$. We can also use these forms to deduce the time average of potential and kinetic energy. The first term in each is a constant, $\frac{1}{4}ka_0^2$, while the second oscillates at $2\omega_0$ and time averages to zero, so the time averaged energies are

$$\langle KE\rangle = \langle PE\rangle = \tfrac{1}{4}ka_0^2, \tag{44}$$

both half the total energy, $E = \frac{1}{2}ka_0^2$. *All energies are quadratic in amplitude* a_0.

The above was a rather *ad-hoc* way of deducing a time average. A more systematic way to do it is by integration. For example, we can find the average of the potential energy by integrating (summing) it over one time period then dividing by the time period $T = 2\pi/\omega_0$.

$$\langle PE\rangle = \frac{1}{T}\int_0^T \frac{1}{2}ka_0^2\cos^2(\omega_0 t+\phi)\mathrm{d}t = \frac{1}{4}ka_0^2. \tag{45}$$

However it is typically easiest just to recall that, averaged over one cycle, $\langle\sin^2\theta\rangle = \frac{1}{2}$ and $\langle\cos^2\theta\rangle = \frac{1}{2}$, so we can find the average energies just by replacing the squared trigonometric functions in eqn. 40, by their time averages of 1/2.

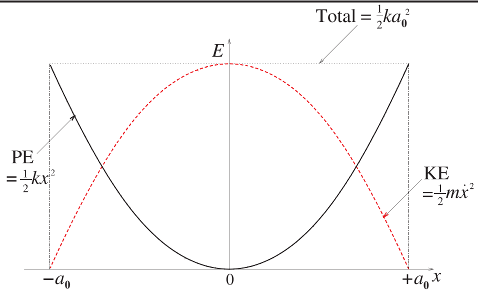

Figure 13: Kinetic and potential energy as a function of displacement x for SHM.

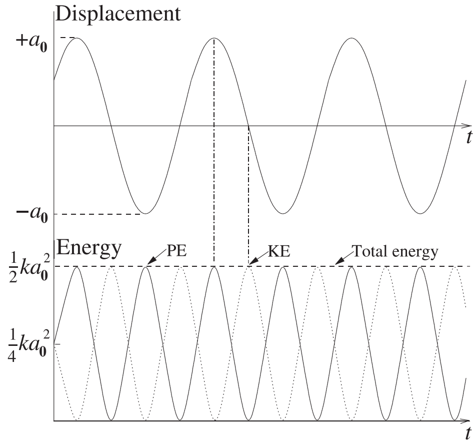

Figure 14: Displacement (top) and kinetic and potential energy (bottom) in an oscillation as a function of time. The energy oscillates between PE and KE at twice the underlying frequency.

3.6.1 The ubiquity of SHM: Harmonic motion in a general potential well

We can also work the other way. If total energy is conserved, its time derivative is zero. For our mass on a spring this means

$$\dot{E} = \frac{1}{2}m\frac{\mathrm{d}}{\mathrm{d}t}\dot{x}^2 + \frac{1}{2}k\frac{\mathrm{d}}{\mathrm{d}t}x^2 = 0. \tag{46}$$

We can conduct both derivatives using the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}t}x^2 = \left(\frac{\mathrm{d}}{\mathrm{d}x}x^2\right)\frac{\mathrm{d}x}{\mathrm{d}t} = 2x\dot{x}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\dot{x}^2 = \left(\frac{\mathrm{d}}{\mathrm{d}\dot{x}}\dot{x}^2\right)\frac{\mathrm{d}\dot{x}}{\mathrm{d}t} = 2\dot{x}\ddot{x}. \tag{47}$$

Putting these into the energy derivative gives

$$\dot{E} = m\dot{x}\ddot{x} + kx\dot{x} = 0 \quad\Longrightarrow\quad \ddot{x} + \frac{k}{m}x = 0, \tag{48}$$

which is the SHM equation for a mass on a spring, but this time derived from conservation of energy. We have just learnt something remarkable. Kinetic energy is always $\frac{1}{2}m\dot{x}^2$ so, whenever a particle moves in a quadratic ($\propto x^2$) potential energy, the total energy will be like that in eqn. 46 and the particle will undergo SHM.

A more complicated potential a particle might move in is the Lennard-Jones potential,

$$PE = V_0\left(\left(\frac{a}{x}\right)^{12} - 2\left(\frac{a}{x}\right)^6\right), \tag{49}$$

which is used to describe the potential energy of a pair of neutral atoms a distance x apart. It is plotted in fig. 15, which shows the potential has a minimum at $x = a$. This arrises as a trade off between the first term, which is strongly repulsive for small x, reflecting the fact that if the atoms are too close together their electron orbitals overlap and nuclear-nuclear electrostatic repulsion takes place, and the second term, which is attractive and dominates at large x, reflecting the long range van-der-Waals attraction between neutral atoms.

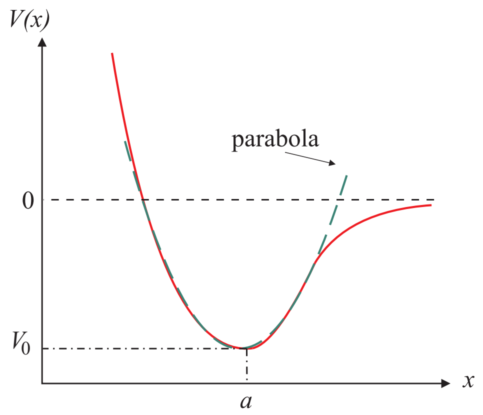

Figure 15: The Lennard-Jones potential for the interaction of neutral atoms.

The Lennard-Jones potential is not quadratic, so we do not expect a particle moving in it to undergo SHM. However, in the region of the minimum at $x = a$, we expect the potential to be well approximated by its Taylor series about $x = a$. Conducting this Taylor series gives

$$V(a+u) = V(a) + u\left(\frac{\mathrm{d}V}{\mathrm{d}x}\right)_{x=a} u + \frac{1}{2}\left(\frac{\mathrm{d}^2V}{\mathrm{d}x^2}\right)_{x=a} u^2 + \frac{1}{6}\left(\frac{\mathrm{d}^3V}{\mathrm{d}x^3}\right)_{x=a} u^3 + \ldots \tag{50}$$

However, since V has a minimum at a, we know $\left(\frac{\mathrm{d}V}{\mathrm{d}x}\right)_{x=a} = 0$. For sufficiently small u the cubic term is negligible compared to the quadratic term, so we can approximate the potential by

$$V(a+u) = V(a) + \frac{1}{2}\left(\frac{\mathrm{d}^2V}{\mathrm{d}x^2}\right)_{x=a} u^2, \tag{51}$$

which is a quadratic potential! Indeed, almost all potentials can be approximated by quadratic potentials for sufficiently small perturbations around their minima. This is why SHM is ubiquitous. If a particle moves in a general potential, it will eventually settle down in a minima. If we then perturb it a little, it will be moving in what is effectively a quadratic potential and undergo SHM around the minima. We expect that any system in a stable equilibrium (i.e. at a potential energy minima) will undergo SHM if we give it a small perturbation.

3.6.2 Energy method for finding ω_0

For any system in equilibrium in a potential energy minima, the total energy, at least for small disturbances from the minima, is of the form

$$E = KE + PE = \frac{1}{2}\alpha\dot{x}^2 + \frac{1}{2}\beta x^2. \tag{52}$$

Conservation of energy then gives

$$\dot{E} = \alpha\dot{x}\ddot{x} + \beta x\dot{x} = 0 \qquad \Longrightarrow \qquad \ddot{x} + \frac{\beta}{\alpha}x = 0, \tag{53}$$

so the system performs SHM with $\omega_0^2 = \frac{\beta}{\alpha}$. It is often much simpler to find the frequency of SHM this way than by direct consideration of forces.

3.6.3 Simple pendulum via conservation of energy

The potential energy in a pendulum is gravitational potential energy, $PE = mgh$. Inspecting fig. 16, we see that when the mass is at θ it has risen a height $l - lg\cos(\theta)$, so its potential energy is

$$PE = mgl(1 - \cos(\theta)). \tag{54}$$

This is not a quadratic potential, but it does have a minima at $\theta = 0$. For small θ we can approximate $\cos(\theta)$ by its Taylor series about zero, $\cos(\theta) = 1 - \frac{1}{2}\theta^2$, so our potential is approximately

$$PE = \frac{1}{2}mgl\theta^2, \tag{55}$$

which is quadratic. Adding in the kinetic energy, $KE = \frac{1}{2}mv^2$, the total energy is

$$E = \frac{1}{2}ml^2\dot{\theta}^2 + \frac{1}{2}mgl\theta^2. \tag{56}$$

Comparing these with eqn 52, we have $\alpha = ml^2$ and $\beta = mgl$ so, as expected, the pendulum will undergo SHM with $\omega_0^2 = \beta/\alpha = g/l$. We now understand that the behavior of the pendulum — SHM for small amplitude, more complicated for large amplitude — is not a quirk of pendulums but the generic behavior for energy conserving oscillating systems. The mass on a spring and the hydrometer are unusual cases where the potential energy is quadratic not just near the minima but everywhere, so the systems exhibit large amplitude SHM.

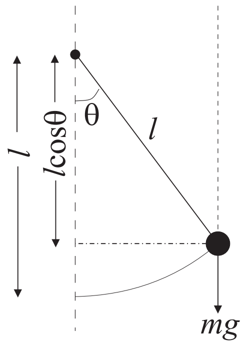

Figure 16: Potential energy for a simple pendulum.

3.6.4 Water in a U-tube via conservation of energy

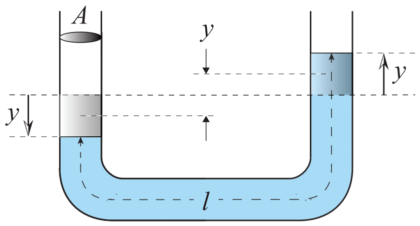

Figure 17: Water in a U shaped tube.

A U shaped tube with cross-sectional area A, is filled with water as sketched in fig. 17. If the water is pushed down on one side then released, what is the frequency of the resulting oscillations?

Again here the potential energy is gravitational. Since the volume of water is constant, if, as sketched, the level in the LHS is a distance y below the equilibrium level, the RHS must rise by y above. In potential energy terms, this is like moving volume of fluid Ay from the left to the right, requiring us to elevate its center of mass by a height y, so the potential energy is

$$PE = \rho Aygy = \rho Agy^2, \tag{57}$$

which is quadratic. If the level is changing at a rate $\dot{y}$ then all the fluid is moving at velocity $\dot{y}$. The total volume of fluid in the tube is Al, so the kinetic energy is

$$KE = \frac{1}{2}\rho Al\dot{y}^2. \tag{58}$$

Comparing these with eqns 52, we have $\alpha = \rho Al$ and $\beta = 2\rho Ag$ so the level water will undergo SHM with $\omega_0^2 = \frac{\beta}{\alpha} = \frac{2g}{l}$. This is much simpler than finding the equations of motion directly.

3.6.5 Torsional Pendulum via conservation of energy

A uniform disk of mass m and radius R is suspended from a wire (a torsional fibre or spring), as sketched in fig. 18. When the wire is twisted through an angle θ, the fibre stores a potential energy

$$PE = \frac{1}{2}\tau\theta^2 \tag{59}$$

where τ is called the torsional stiffness and has units N m rad^{-1}. The disk is twisted and released. What is the frequency of the resulting oscillations?

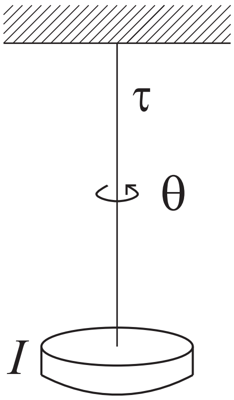

Figure 18: A torsional pendulum.

We already know the potential energy, but we need to find the kinetic energy. This is slightly tricky as the mass at the center of the disk is not moving at all, whereas the mass at the edge moves quickly. However, if the disk is rotating at $\dot{\theta}$ then the thin ring of mass of width $\mathrm{d}r$ and at a distance r from the center of the disk, shown in fig. 19, is all moving at the same speed, $v = r\dot{\theta}$. The mass-per-unit-area of the disk is $m/(\pi R^2)$, so the ring has mass $\mathrm{d}m = (m/(\pi R^2)) \times 2\pi r \mathrm{d}r$ and kinetic energy

$$\mathrm{d}(KE) = \frac{1}{2}(\mathrm{d}m)v^2 = \frac{m\dot{\theta}^2}{R^2} r^3 \mathrm{d}r. \tag{60}$$

To find the total kinetic energy of the disk we must add up all of these contributions using an integral

$$KE = \sum d(KE) \to \int d(KE) = \int_{r=0}^{r=R} \frac{m\dot{\theta}^2}{R^2} r^3 \mathrm{d}r = \frac{1}{4} m R^2 \dot{\theta}^2. \tag{61}$$

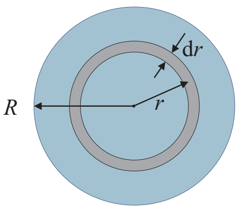

Figure 19: A disk can be broken into a series of small rings.

Comparing with eqn 52, we have $\alpha = \frac{1}{2}mR^2$ and $\beta = \tau$, so the disk oscillates with angular frequency $\omega_0^2 = \frac{\beta}{\alpha} = \frac{2\tau}{mR^2}$. In Lent term you will study rotational motion, and learn to do this calculation in a force-like way. You will also learn that the kinetic energy of any spinning body can be written as $\frac{1}{2}I\dot{\theta}^2$, where I is called the body's moment of inertia. For the disk we effectively just calculated $I = \frac{1}{2}mR^2$. The general result for a torsional pendulum is $\omega_0^2 = \frac{\tau}{I}$.

3.6.6 Diatomic molecule via conservation of energy

We model a diatomic molecule, such as HCl, as two different masses, m_1 and m_2, connected by a spring of constant k, as sketched in fig. 20. What is the vibrational frequency of the molecule.

If m_1 is displaced outwards by x_1 and m_2 outwards by x_2 the the spring is extended by $x_1 + x_2$. There is no external force on the molecule so its center of mass cannot move, meaning $m_1 x_1 = m_2 x_2$. We can therefore write the potential and kinetic energy entirely in terms of x_1 as

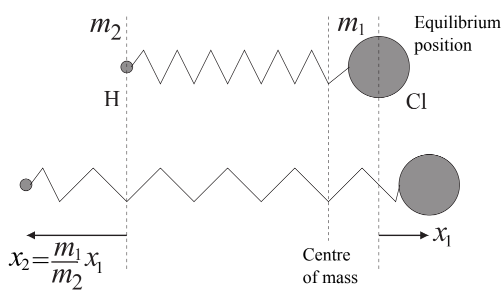

Figure 20: Model of a diatomic molecule.

$$PE = \frac{1}{2}k(x_1 + x_2)^2 = \frac{1}{2}k\left(1 + \frac{m_1}{m_2}\right)^2 x_1^2, \tag{62}$$

and

$$KE = \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2\dot{x}_2^2 = \frac{1}{2}m_1\left(1 + \frac{m_1}{m_2}\right)x_1^2. \tag{63}$$

Comparing with eqn 52, $\alpha = m_1\left(1 + \frac{m_1}{m_2}\right)$ and $\beta = k\left(1 + \frac{m_1}{m_2}\right)^2$ so we have SHM with

$$\omega_0^2 = \frac{\beta}{\alpha} = \frac{\left(1 + \frac{m_1}{m_2}\right)k}{m_1} = k\left(\frac{1}{m_1} + \frac{1}{m_2}\right). \tag{64}$$

The system behaves as a single mass on a spring with constant k, but with an effective mass m known as the reduced mass given by

$$\frac{1}{m} = \frac{1}{m_1} + \frac{1}{m_2}. \tag{65}$$

This simple calculation allows us to estimate the "spring-constant" of an atomic bond. HCl strongly absorbs radiation of around $\nu \sim 10^{14}$Hz because such radiation excites the vibrations. The masses of chlorine and hydrogen are 1 a.m.u and 35 a.m.u. respectively (where 1 a.m.u.$=$ 1.67×10^{-27}kg). Putting in these numbers gives

$$k = m\omega_0^2 = \left(\frac{1}{m_1} + \frac{1}{m_2}\right)^{-1} 2\pi\nu \sim 640\text{Nm}^{-1}, \tag{66}$$

which is a typical every-day value: such a spring would extend 1.5cm under a 1kg load.

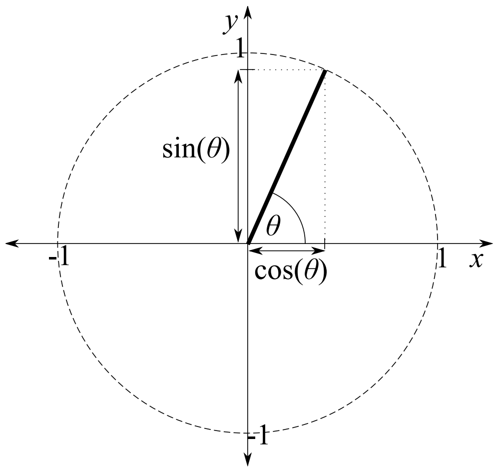

Figure 21: A unit length rod making an angle θ with the x axis.

3.7 SHM and Circles

Imagine a rod of unit length which protrudes from the origin making an angle θ with the x axis, as sketched in fig. 21. The projection of the rod onto the x axis — that is, its shadow on the x axis if it is illuminated with rays parallel to the y axis — has length $\cos(\theta)$. As θ increases, the end of the rod moves around the unit circle and it's projection onto the x axis changes between ± 1. In this geometry, it is obvious that the cosine of a quarter turn is zero, the cosine of a half turn is -1, the cosine of a three-quarter turn is again zero, and the cosine of a full turn is 1. The rod's projection onto the y axis has length $\sin(\theta)$, which also varies between ± 1 and, similarly, it is obvious that the sine of zero is zero, the sine of a quarter turn is 1, and so on.

If we now imagine the rod is spinning at constant angular velocity ω_0 then, at time t, it makes an angle $\omega_0 t$ with the x axis and its projection onto the x axis is $\cos(\omega_0 t)$. The rod moves in circles, but its projection onto the x axis (or indeed any other diameter of the circle) does simple harmonic motion.

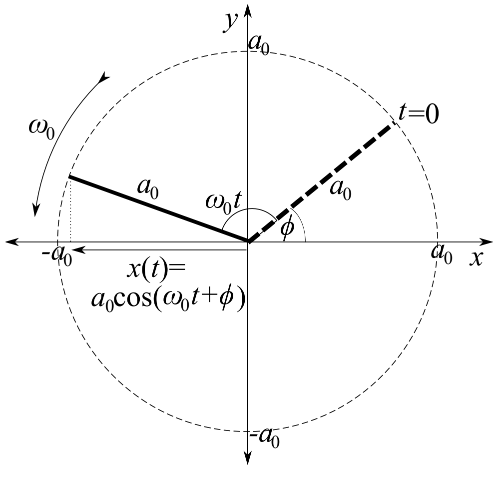

Figure 22: A vector of length a_0 rotates around the origin. Its projection onto the x axis undergoes SHM.

We now replace the rod by a vector of length a_0 that at $t = 0$ makes an angle ϕ with the x axis, but which still spins at angular velocity ω_0. At a time t it makes an angle $\phi + \omega_0 t$ (see fig. 22) so its projection onto the x axis has length

$$x = a_0 \cos(\omega_0 t + \phi). \tag{67}$$

The projection undergoes the general form of SHM. This representation of SHM gives a geometric meaning to the previously abstract angle $\omega t + \phi$, called the oscillation's phase, and clarifies the link between angular frequency and angular velocity.

3.7.1 Phasor diagrams

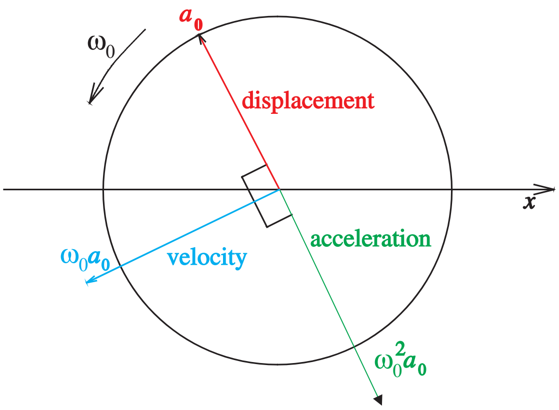

Figure 23: Displacement velocity and acceleration of SHM represented by three co-rotating vectors with different lengths. Their projections onto the x axis give the instantaneous value of their respective variables.

The velocity and acceleration can be represented the same way. The velocity is given by

$$\dot{x}(t) = -\omega_0 a_0 \sin(\omega_0 t + \phi) = \omega_0 a_0 \cos\left(\omega_0 t + \phi + \frac{\pi}{2}\right), \tag{68}$$

so it can also be represented by a vector spinning at ω_0 but with a length of $\omega_0 a_0$ and that is $\pi/2$ (quarter of a turn) ahead of the displacement's vector. Similarly, the acceleration is

$$\ddot{x}(t) = -\omega_0^2 a_0 \cos(\omega_0 t + \phi) = \omega_0^2 a_0 \cos(\omega_0 t + \phi + \pi), \tag{69}$$

so it can also be represented by a vector spinning at ω_0 but with length of $\omega_0^2 a_0$ and that is π (half a turn) ahead of the displacement's vector. This leads us to fig. 23, a diagram with all three vectors. They all rotate at the same speed, maintaining their relative angles, and the values of the displacement, velocity and acceleration at a given moment in time are given by their projections onto the x axis. Such diagrams are called phasor diagrams. They allow us to visualize the phase difference between different quantities in SHM.

3.8 Superposition of simple harmonic motions

Suppose we have two solutions to the same fundamental SHM equation, $\ddot{x}_1 = -\omega_0^2 x_1$ and $\ddot{x}_2 = -\omega_0^2 x_2$. The SHM equation (eqn 8) is linear, so the sum of these, $x_3 = x_1 + x_2$, is also a solution:

$$\ddot{x}_3 = \ddot{x}_1 + \ddot{x}_2 = -\omega_0^2 x_1 - \omega_0^2 x_2 = -\omega_0^2 (x_1 + x_2) = -\omega_0^2 x_3. \tag{70}$$

N.B. *This only works because the restoring force is linear, it would not work if, for example, we had $F \propto x^n$ and $n \neq 1$.* Since x, x_1 and x_2 are all SHM solutions, we must be able to write them all in the form of the general solution:

$$x_1 = a_1 \cos(\omega_0 t + \phi_1) \qquad x_2 = a_2 \cos(\omega_0 t + \phi_2)$$
$$x_3 = a_1 \cos(\omega_0 t + \phi_1) + a_2 \cos(\omega_0 t + \phi_2) = a \cos(\omega_0 t + \phi). \tag{71}$$

We can understand how this works by representing x_1 and x_2 as vectors of different lengths and phases (but both spinning at ω_0) on the same phasor diagram, as shown in fig. 24. The vector sum of the x_1 and x_2 phasors gives a third one representing x_3, also spinning at ω_0. From the diagram we can visualize how the phase and amplitude of x depend on the phase and amplitude of x_1 and x_2. With some geometry we can actually find a and ϕ. First, applying the cosine rule to the vector-sum triangle we have

$$a^2 = a_1^2 + a_2^2 - 2a_1 a_2 \cos(\alpha), \tag{72}$$

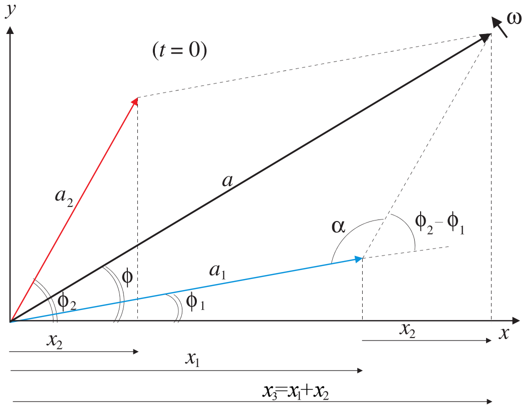

Figure 24: Two phasors representing SHM with the same frequency but different amplitude and phase. The sum of the two motions is also SHM, represented by a phasor that is the vector sum of the first two.

but from the diagram we know $\alpha = \pi - (\phi_2 - \phi_1)$, so $\cos(\alpha) = -\cos(\phi_2 - \phi_1)$, giving the resultant amplitude a as

$$a^2 = a_1^2 + a_2^2 + 2a_1a_2\cos(\phi_2 - \phi_1). \qquad (73)$$

Secondly, looking at the right-angle triangle formed by the x_3 phasor and the x axis, we have

$$\phi = \tan^{-1}\left(\frac{a_1\sin\phi_1 + a_2\sin\phi_2}{a_1\cos\phi_1 + a_2\cos\phi_2}\right). \qquad (74)$$

We could prove these results by manipulating eqn 71, but using phasors builds geometric intuition.

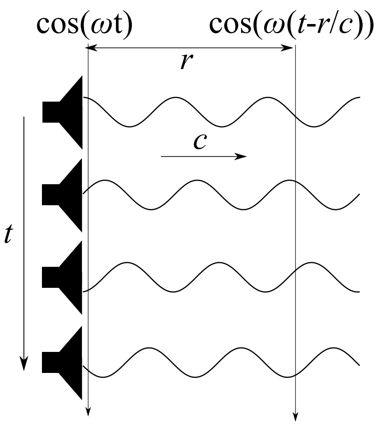

Figure 25: A sequence of snapshots of the sound wave propagating out of a speaker. An observer next the the speaker observes $\cos(\omega t)$, whereas an observer a distance r away sees $\cos(\omega(t - r/c))$.

3.8.1 Interference via superposition of matching frequencies

Let us imagine an emitter of waves that is oscillating as $a_0\cos(\omega t)$. These oscillations travel away from the emitter at the wave speed c, as sketched in fig. 25, so if we measure a distance r from the emitter we see what the emitter was doing a time r/c ago, that is, we measure $a_0\cos(\omega(t - r/c)) = a_0\cos(\omega t - \omega r/c)$. This is exactly the same form as the SHM solution (with $\phi = -\omega r/c$) so we can analyze it using phasors.

We first imagine that we have two identical sources a distance R apart, and we stand between them, a distance r from the first thus a distance $R - r$ from the second, as shown in fig. 26. We experience the sum of two signals, $a_0\cos(\omega t - \omega r/c)$ from the first emitter and $a_0\cos(\omega t - \omega(R - r)/c)$ from the second. We can represent this sum as the sum of two phasors of equal amplitude, but with phase constants $\phi_1 = -\omega r/c$ and $\phi_2 = \omega(r - R)/c$ respectively. If we stand half way between the emitters, these are

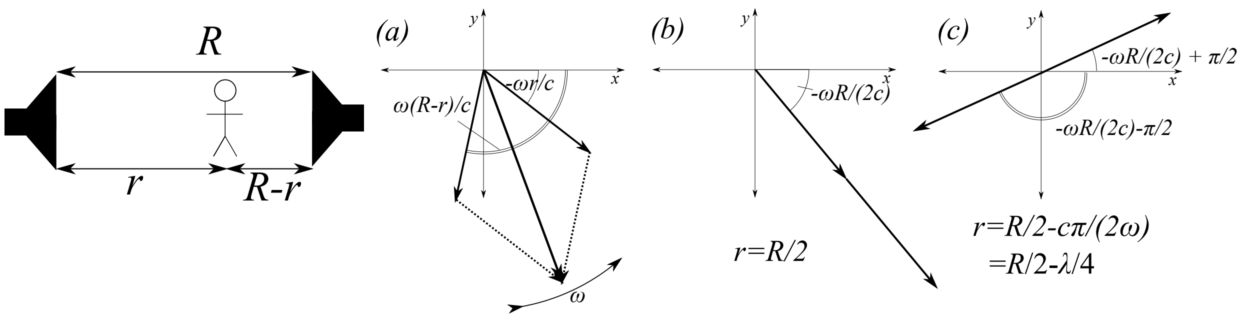

Figure 26: Left: Two emitter, a distance R apart, oscillate together as $\cos(\omega t)$, and an observer stands between them. (a) The wave observed from each source can be represented by a phasor, and the total wave observed as their sum. The two phasors have the same amplitude but different phases. (b) If the observer is half way between the sources the phasors have equal phase and add constructively. (c) But as the observer moves away from $R/2$ the two phases change in opposite directions and eventually differ by π, leading to destructive interference and no resultant wave.

both $\phi = -R/(2c)$ and the phasors are aligned and add constructively, giving a result with amplitude $2a_0$. However, as we move towards the first emitter, ϕ_1 increases and ϕ_2 decreases so the amplitude of the result falls. When we reach $r = R/2 - c\pi/(2\omega)$ the phasors are opposite to each other and the resultant amplitude is zero. From $c = \nu\lambda$, we know that $c\pi/(2\omega) = \lambda/4$, is quarter of the waves's wavelength. As we move between the two sources, we move between regions of high amplitude ($2a_0$) oscillations (anti-nodes) and regions of zero amplitude (nodes), separated in space by $\lambda/4$. This is called a standing wave.

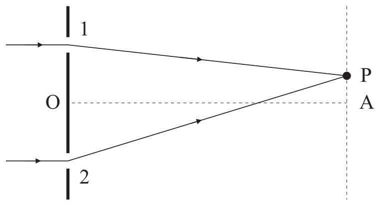

Figure 27: Two slits are illuminated from behind, and the resultant intensity pattern is collected on a screen on the right.

The same thing happens if we imagine illuminating two slits and collecting the resultant intensity pattern on a screen, as sketched in fig. 27. The two slits act as synchronous sources of waves and at each point on the screen we collect the sum of the two waves, which we can represent as the sum of two phasors. Exactly between the two slits (at A) both waves have traveled the same distance, so the phasors have the same phase and add constructively. Moving to the side (towards P) the distances to the two slits start to differ and the phasors move out of phase, eventually canceling. This leads to a pattern of "fringes" (stripes of high and low intensity) on the screen and, historically, was considered the definitive proof that first light then, later, electrons, have a wave character.

3.8.2 Beats via superposition of different frequencies

We can also imagine two wave sources with different frequencies so that, at a given point in space, the two waves produce oscillations

$$x_1 = a_0 \cos(\omega_1 t) \qquad x_2 = a_0 \cos(\omega_2 t), \quad (75)$$

and our observer measures the sum:

$$x = x_1 + x_2 = a_0 \left(\cos(\omega_1 t) + \cos(\omega_2 t)\right). \quad (76)$$

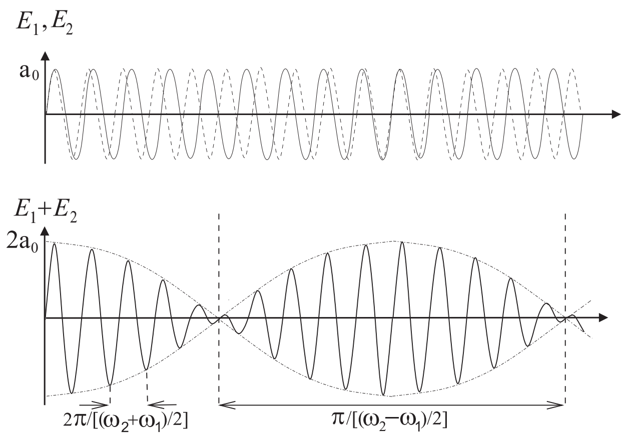

Figure 28: Top: Two oscillations with slightly different frequencies plotted as a function of time. They start in phase and add constructively but, as time goes on, they become out of phase and add destructively. Bottom: Sum of the two oscillations, showing a low frequency oscillation between constructive and destructive interference, known as beats.

We can easily see what happens by drawing a graph of the two oscillations assuming their frequencies are close, as shown in fig. 28. The oscillations start in phase but and add constructively. However, since they have slightly different frequencies, over many cycles they drift out of phase. Eventually the two are completely out of phase and cancel out. They then drift back into phase, and add constructively. This leads to a low frequency alternation between constructive and destructive interference known as beats. We can analyze this more formally using the trig angle addition formula:

$$
\begin{aligned}
x &= a_0\left(\cos(\omega_1 t) + \cos(\omega_2 t)\right) \qquad (77)\\
&\equiv a_0\left(\cos\left(\frac{\omega_2+\omega_1}{2}t - \frac{\omega_2-\omega_1}{2}t\right)\right.\\
&\left.+\cos\left(\frac{\omega_2+\omega_1}{2}t + \frac{\omega_2-\omega_1}{2}t\right)\right)\\
&= 2a_0\cos\left(\frac{\omega_2+\omega_1}{2}t\right)\cos\left(\frac{\omega_2-\omega_1}{2}t\right).
\end{aligned}
$$

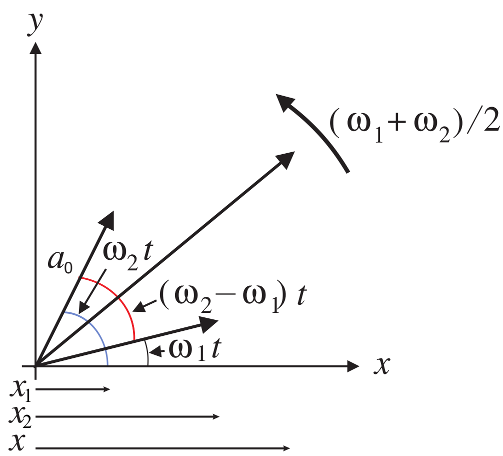

Figure 29: Analyses of beats via phasors.

We see that x varies between $\pm 2a_0$ and is the product of a rapid oscillation, which oscillates at the average of ω_1 and ω_2, and a slow oscillation with frequency $(\omega_2-\omega_1)/2$, i.e. proportional to the difference in the two frequencies. This is exactly what fig. 28 looks like. The beats are caused by the slow "envelope" oscillation, which gets slower as the frequencies get closer. You hear a beat (maximum) whenever $\cos((\omega_2-\omega_1)t/2) = \pm 1$, so the actual angular frequency of the beats is simply $(\omega_2-\omega_1)$.

Finally, we can also analyze beats with phasors. Our sum is now the sum of two phasors that spin at slightly different frequencies, as shown in fig. 29. They start aligned but one slowly gets ahead of the other reducing the amplitude of their sum. However, the resultant always bisects the original two phasors, so it makes an angle of $(\omega_1 t + \omega_2 t)/2$, i.e. it spins at the average frequency. The phase difference between the two is $(\omega_2-\omega_1)t$, and when this is half a turn (π, 3π, 5π...) the two are out of phase and the resultant amplitude is zero, when it is a full turn (0, 2π, 4π...) they are in phase and the amplitude is maximum. As promised, the beats thus come with angular frequency $\omega_2-\omega_1$, or real frequency $\nu = (\omega_2-\omega_1)/(2\pi)$.

3.9 Complex representation of SHM

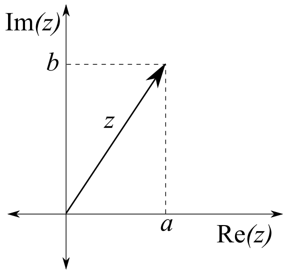

Figure 30: Argand diagram showing the complex number $z = a + \mathrm{i}b$ in the complex plane.

Phasor diagrams have a natural representation in terms of complex numbers. A complex number is the sum of a real and an imaginary number such as

$$z = a + \mathrm{i}b, \qquad a,b \in \mathbb{R}. \tag{78}$$

We can plot z on an Argand diagram, a representation of the complex plane in which the x axis represents the real part of z and the y axis the imaginary part, seen in fig. 30. By Pythagoras the "vector" z has length $A = \sqrt{a^2+b^2}$, and, by trigonometry, it makes an angle $\phi = \arctan(b/a)$ with the real axis. We can thus write our complex number as

$$z = A(\cos(\phi) + \mathrm{i}\sin(\phi)). \tag{79}$$

The power of complex numbers comes from the fact we can also write z as a complex exponential:

$$\cos(\phi) + \mathrm{i}\sin(\phi) = \mathrm{e}^{\mathrm{i}\phi} \qquad \Longrightarrow \; z = A\mathrm{e}^{\mathrm{i}\phi}. \tag{80}$$

The complex number $z = a_0\mathrm{e}^{\mathrm{i}(\omega_0 t+\phi)}$ thus has constant length a_0 but spins around the Argand diagram with initial phase ϕ, as sketched in fig. 31. This is just like a phasor, and indeed the projection of z onto the real-axis (i.e. its real part) undergoes general SHM:

$$
\begin{aligned}
x &= \mathrm{Re}\{a_0 \mathrm{e}^{\mathrm{i}(\omega_0 t+\phi)}\} \\
&= \mathrm{Re}\{a_0(\cos(\omega_0 t+\phi)+\mathrm{i}\sin(\omega_0 t+\phi))\} \\
&= a_0\cos(\omega_0 t+\phi). \qquad (81)
\end{aligned}
$$

Complex numbers allow us to turn rotation into multiplication. If I have a complex number that makes an angle ϕ_1 with the real axis, $z = a_0\mathrm{e}^{\mathrm{i}\phi_1}$, and I then rotate it a further angle ϕ_2, my new complex number, $\tilde{z}$, is

$$\tilde{z} = a_0\mathrm{e}^{\mathrm{i}(\phi_1+\phi_2)} = a_0\mathrm{e}^{\mathrm{i}\phi_1}\mathrm{e}^{\mathrm{i}\phi_2} = z\mathrm{e}^{\mathrm{i}\phi_2}. \qquad (82)$$

This makes complex numbers a powerful way of dealing with phases. For example, we can write our SHM solution as

$$z = a_0\mathrm{e}^{\mathrm{i}(\omega_0 t+\phi)} = a_0\mathrm{e}^{\mathrm{i}\phi}\mathrm{e}^{\mathrm{i}\omega_0 t} = A\mathrm{e}^{\mathrm{i}\omega_0 t}, \qquad (83)$$

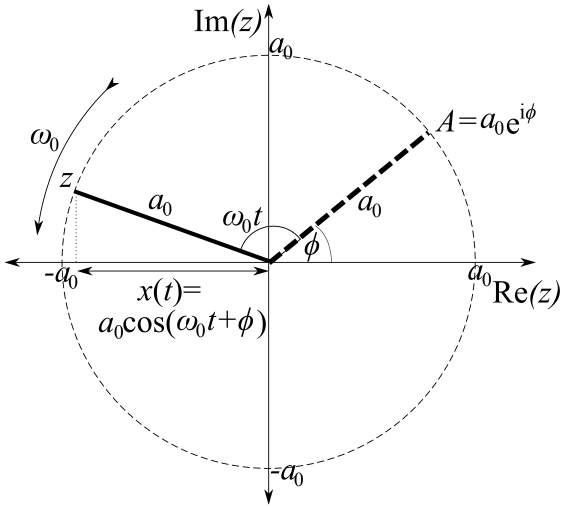

Figure 31: Argand diagram showing the complex number $z = a_0\mathrm{e}^{\mathrm{i}(\omega_0 t+\phi)}$ spinning in the complex plane like a phasor.

where the complex number $A = a_0\mathrm{e}^{\mathrm{i}\phi}$ encodes the amplitude and the phase constant (fig. 31).

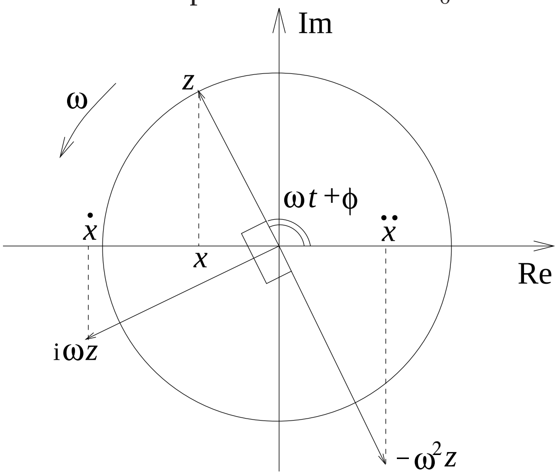

Figure 32: Argand diagram showing $z = a_0\mathrm{e}^{\mathrm{i}(\omega_0 t+\phi)}$ and its first and second time derivatives, $\dot{z} = \mathrm{i}\omega_0 z$ and $\ddot{z} = -\omega_0^2 z$.

We can also differentiate our complex solution to find the velocity:

$$\dot{z} = \frac{\mathrm{d}}{\mathrm{d}t}\left(A\mathrm{e}^{\mathrm{i}\omega_0 t}\right) = \mathrm{i}\omega_0 A\mathrm{e}^{\mathrm{i}\omega_0 t} = \mathrm{i}\omega_0 z. \qquad (84)$$

The multiplication by $\mathrm{i} = \mathrm{e}^{\mathrm{i}\pi/2}$ rotates z by a quarter-turn ($\pi/2$) on the Argand diagram, ensuring that $\dot{z}$ spins a quarter turn ahead of z, just as with the phasors. As with displacement, the physical velocity is given by the real part:

$$
\begin{aligned}
\mathrm{Re}\{\dot{z}\} &= \mathrm{Re}\{\mathrm{i}\omega_0 A\mathrm{e}^{\mathrm{i}\omega_0 t}\} = \mathrm{Re}\{\mathrm{i}\omega_0 a_0\mathrm{e}^{\mathrm{i}\phi}\mathrm{e}^{\mathrm{i}\omega_0 t}\} \\
&= \mathrm{Re}\{\mathrm{i}\omega_0 a_0\left(\cos(\omega_0 t+\mathrm{i}\phi)+\mathrm{i}\sin(\omega_0 t+\mathrm{i}\phi)\right)\} \\
&= -\omega_0 a_0\sin\left(\mathrm{i}\omega_0 t+\mathrm{i}\phi\right) \\
&= \dot{x}. \qquad (85)
\end{aligned}
$$

Similarly the complex acceleration is

$$\ddot{z} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{z}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{i}\omega_0 A\mathrm{e}^{\mathrm{i}\omega_0 t}\right) = \mathrm{i}^2\omega_0^2 A\mathrm{e}^{\mathrm{i}\omega_0 t} = -\omega_0^2 z. \qquad (86)$$

The multiplication by $-1 = \mathrm{e}^{\mathrm{i}\pi}$ rotates z by half a turn (π), so, just like phasors, the complex acceleration is π ahead of the complex displacement. Fig. 32 is an Argand diagram of z, $\dot{z}$ and $\ddot{z}$, which reproduces the phasor diagram (fig. 23). Again, $\mathrm{Re}\{\ddot{z}\}$ gives the physical acceleration:

$$\mathrm{Re}\{\ddot{z}\} = \mathrm{Re}\{-\omega_0^2 z\} = \mathrm{Re}\{-\omega_0^2 a_0\mathrm{e}^{\mathrm{i}\phi}\mathrm{e}^{\omega_0 t}\} = -\omega_0^2 a_0\cos\left(\omega_0 t+\phi\right) = \ddot{x}.$$

Although the physical solution is the real part of the complex one, we note from eqn 86 that, $\ddot{z} + \omega_0^2 z = 0$, that is, the whole complex solution satisfies the fundamental equation of SHM. This is because, in addition to the real part of z undergoing SHM, the imaginary part $\mathrm{Im}\{z\} = a_0\sin(\omega_0 t+\phi)$ also undergoes SHM with the same amplitude but $\pi/2$ behind the real part.

3.9.1 Energy in the complex representation

When calculating energies in the complex representation, we must be careful to take real parts before squaring. For example, with a mass on a spring, the potential energy is

$$PE = \frac{1}{2}kx^2 = \frac{1}{2}k\left(\mathrm{Re}\{z\}\right)^2 = \frac{1}{2}a_0^2\cos^2\left(\omega_0 t + \phi\right). \quad (87)$$

We must *not* work out $\mathrm{Re}\{\frac{1}{2}kz^2\}$, the real part of the "complex-potential energy": in general $\mathrm{Re}\{z\}^2 \neq \mathrm{Re}\{z^2\}$, easily verified by trying $z = \mathrm{i}$. We must be similarly careful with the KE,

$$KE = \frac{1}{2}m\dot{x}^2 = \frac{1}{2}m\mathrm{Re}\{\dot{z}\}^2 = \frac{1}{2}m\mathrm{Re}\{\mathrm{i}\omega_0 z\}^2. \quad (88)$$

Here it is useful to note that $\mathrm{Re}\{iz\} = -\mathrm{Im}\{z\}$ (easy to check if you set $z = a + \mathrm{i}b$), so

$$KE = \frac{1}{2}m\omega_0^2\mathrm{Im}\{z\}^2 = \frac{1}{2}m\omega_0^2 a_0^2\sin^2(\omega_0 t + \phi). \quad (89)$$

Recalling $\omega_0^2 = k/m$, we see the total energy does have a nice complex representation:

$$E = KE + PE = \frac{1}{2}k\mathrm{Im}\{z\}^2 + \frac{1}{2}k\mathrm{Re}\{z\}^2 = \frac{1}{2}k(\mathrm{Im}\{z\}^2 + \mathrm{Re}\{z\}^2) = \frac{1}{2}k|z|^2, \quad (90)$$

where the modulus of a complex number $z = a + \mathrm{i}b$ is $|z| = \sqrt{a^2 + b^2}$, which, geometrically, is its distance from the origin on the Argand diagram. The process is conserving energy because the length of z is constant as it spins. We can verify that this form for E agrees with eqn 41:

$$E = \frac{1}{2}k|z|^2 = \frac{1}{2}k\left|A\mathrm{e}^{\mathrm{i}\omega_0 t}\right|^2 = \frac{1}{2}k|A|^2 = \frac{1}{2}k|a_0\mathrm{e}^{\mathrm{i}\phi}|^2 = \frac{1}{2}ka_0^2. \quad (91)$$

3.9.2 Comparison of the complex and standard methods

A mass on a spring oscillates with $\omega_0 = 3\mathrm{rad\ s}^{-1}$. *It is released with displacement -3 cm and velocity 6 cms*$^{-1}$*, as shown in fig. 33. Find the amplitude of the oscillations and the times when the mass is at rest.*

Solution 1: Using $\sin$ *and* $\cos$. The initial conditions are easiest to impose if we use the general form of the solution

$$x = A\cos(3t) + B\sin(3t) \quad \Longrightarrow \quad \dot{x} = -3A\sin(3t) + 3B\cos(3t). \quad (92)$$

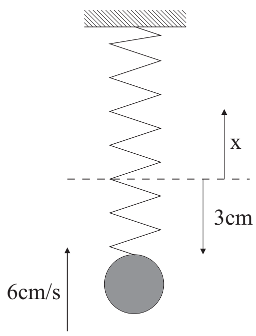

Figure 33: Initial conditions for a mass oscillating on a spring.

Working in cm, our initial conditions are $x(0) = -3$, requiring $A = -3$ and $\dot{x}(0) = 6$ requiring $B = 2$. The amplitude is $a_0 = \sqrt{A^2 + B^2}$ (eqn 16), so $a_0 = \sqrt{13}\mathrm{cm}$. The mass is at rest whenever $\dot{x} = 0$, that is when,

$$\begin{aligned} 3A\sin(3t) = 3B\cos(3t) \quad &\Longrightarrow \tan 3t = \frac{B}{A} \\ &\Longrightarrow t = \frac{1}{3}\left(\arctan\left(\frac{B}{A}\right) + n\pi\right) \quad (93) \\ &= \frac{1}{3}\left(\arctan\left(\frac{-2}{3}\right) + n\pi\right) \quad = 0.85, 1.90, 2.95, \ldots\mathrm{s}. \end{aligned}$$

Solution 2: Using complex numbers. The complex solution is

$$z = A\mathrm{e}^{\mathrm{i}3t} \quad \Longrightarrow \quad \dot{z} = 3\mathrm{i}A\mathrm{e}^{\mathrm{i}3t}. \tag{94}$$

Our initial conditions are $\mathrm{Re}\{z\} = -3$, giving $\mathrm{Re}\{A\} = -3$, and $\mathrm{Re}\{\mathrm{i}3A\} = -3\mathrm{Im}\{A\} = 6$, giving $\mathrm{Im}\{A\} = -2$. The full solution is thus

$$z = (3 - 2\mathrm{i})\mathrm{e}^{\mathrm{i}3t}. \tag{95}$$

The amplitude is $a_0 = |A| = \sqrt{13}$cm and the complex velocity is $\dot{z} = (-9\mathrm{i} + 6)\mathrm{e}^{\mathrm{i}3t} = (-9\mathrm{i} + 6)(\cos(3t) + \mathrm{i}\sin(3t))$. The physical velocity is zero whenever $\mathrm{Re}\{\dot{z}\} = 0$, i.e. whenever

$$\mathrm{Re}\{\dot{z}\} = 6\cos(3t) + 9\sin(3t) = 0 \quad \Longrightarrow \tan 3t = -\frac{2}{3} \tag{96}$$

$$\Longrightarrow t = \frac{1}{3}\left(\arctan\left(\frac{-2}{3}\right) + n\pi\right) = 0.85, 1.90, 2.95, \ldots\mathrm{s}.$$

3.9.3 Why use the complex representation

For pure SHM the advantages of the complex representation are marginal. However, the complex representation has several advantages for more complicated problems. Firstly, complex numbers turn rotations into multiplications. This makes life much easier if you have several phase shifts to keep track of. Secondly, the exponential function is the easiest possible function to differentiate. Indeed, since $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{\mathrm{i}\omega t} = \mathrm{i}\omega e^{\mathrm{i}\omega t}$, exponential functions also turn differentiation into multiplication. Finally, the complex exponential makes a link between exponential decay and the oscillating functions $\sin$ and $\cos$, which is very useful when we analyze damped and driven oscillators.

3.10 Summary of SHM

- The fundamental equation of SHM, $\ddot{x} + \omega_0^2 x = 0$, arrises whenever we have an equilibrium with a restoring force proportional to displacement, such as $F = -kx$ for the spring.
- The solution to the SHM equation can be written in two different ways

$$x = a_0\cos(\omega_0 t + \phi) \quad \text{or} \quad x = A\cos(\omega_0 t) + B\sin(\omega_0 t), \tag{97}$$

  where the constants (a_0 and ϕ or A and B) are fixed by the initial conditions of the oscillation but ω_0 only depends on the oscillator.
- The amplitude of the oscillation, a_0, does not depend on the frequency $\nu = \omega_0/(2\pi)$.
- The SHM equation is linear, so we can add (superpose) solutions to form new solutions.
- SHM conserves total energy. On average it is shared equally between two forms (KE and PE), but it oscillates between the two at $2\omega_0$.
- We can derive SHM from conservation of energy, provided our potential energy is quadratic. All potentials are approximately quadratic near their minima so SHM is ubiquitous.
- SHM is the projection of motion in a circle onto a diameter of the circle. This gives a geometric meaning to the phase of an oscillating variable.
- Circular motion can be neatly described by a rotating complex number $z = A\mathrm{e}^{\mathrm{i}\omega_0 t}$. The real part of this is a third general way of writing the SHM solution $x = \mathrm{Re}\{z\} = \mathrm{Re}\{A\mathrm{e}^{\mathrm{i}\omega_0 t}\}$. Both amplitude and initial phase are encoded by the complex number $A = a_0\mathrm{e}^{\mathrm{i}\phi}$. However, z also satisfies the SHM equation in its own right, $\ddot{z} + \omega_0^2 z = 0$.

4 Damped Harmonic Motion

4.1 Equation of damped harmonic motion

SHM conserves energy perfectly; after we set our system oscillating the amplitude of the oscillations remains constant. No oscillators actually behave like this, instead they slowly dissipate energy through friction-like processes, and the amplitude of the oscillations dies down over time. Consider a horizontal mass on a spring, as sketched in fig. 42, that, in addition to the spring force $F_s = -kx$ is also subject to a friction[2] force $F_f = -b\dot{x}$. This is a force that always points opposite to the velocity of the mass, acting to slow it down. Applying Newton's second law, the equation of motion of the mass is now

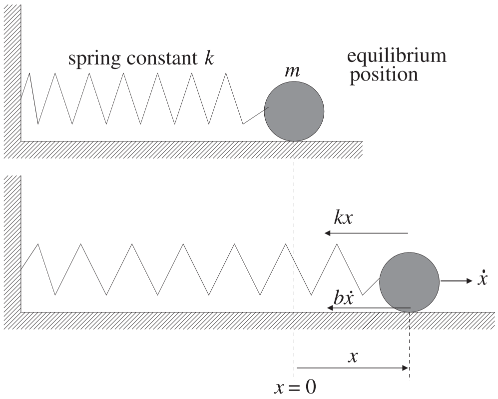

Figure 34: Mass on a spring with damping. Top: Mass at rest in its equilibrium position. Bottom: When the mass is at x and has velocity $\dot{x}$ it feels both a spring force $-kx$ and a friction force $-b\dot{x}$.

$$\begin{aligned} F = F_s + F_f &= m\ddot{x} \\ \implies -kx - b\dot{x} &= m\ddot{x} \\ \implies \ddot{x} + \frac{b}{m}\dot{x} + \frac{k}{m}x &= 0. \end{aligned} \tag{98}$$

This is an example of the general form of the equation of damped SHM,

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0, \tag{99}$$

with $\gamma = b/(2m)$ and $\omega_0^2 = k/m$. Many other systems produce this equation of motion. Before solving it properly, it is worth asking what we expect to happen. If $\gamma = 0$ there is no friction and we have SHM: the system will oscillate at ω_0 in perpetuity. If $\omega_0 = 0$ there is no spring force and we just have a mass moving against friction. We now don't expect oscillations, rather we expect that, if we give the mass a velocity, it will just be slowed down by friction. The equation of motion in this case, $\ddot{x} = -2\gamma\dot{x}$, is an exponential decay equation for $\dot{x}$ solved by

$$\dot{x} = \dot{x}(0)\mathrm{e}^{-2\gamma t}, \tag{100}$$

i.e. the mass slows down, with its velocity decaying by a factor of e in the time $\tau = 1/(2\gamma)$.

The constants γ and ω_0 both have units of $1/\mathrm{time}$ but with very different interpretations: $T = 2\pi/\omega_0$ is the period of an oscillation while $\tau = 1/(2\gamma)$ is a decay time. This leads us to expect two different regimes. If $T \ll \tau$ then the system will oscillate many times before it decays, leading to many oscillations with slowly falling amplitude. We call this light damping. If $T \gg \tau$ then any velocity we give the mass will decay in much less than one oscillation, and the mass will then just move slowly back to the equilibrium point. We call this heavy damping.

4.2 Solving the equation of damped harmonic motion

If there is heavy damping we expect our system to simply decay without oscillation, so we try the $x = A\mathrm{e}^{-pt}$. Substituting this into eqn 98 turns the differential equation into a quadratic one:

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0 \implies p^2(A\mathrm{e}^{-pt}) - 2\gamma p(A\mathrm{e}^{-pt}) + \omega_0^2(A\mathrm{e}^{-pt}) = 0. \tag{101}$$

[2] Full disclosure: this isn't actually a good model for friction. Sliding friction always points against velocity, as this force does, but it is independent of speed. This is really a viscous drag force.

Assuming $x \neq 0$ (i.e. we have some displacement) we can cancel $A\mathrm{e}^{-pt}$ from this to get

$$p^2 - 2\gamma p + \omega_0^2 = 0 \qquad \Longrightarrow \; p = \gamma \pm \sqrt{\gamma^2 - \omega_0^2}. \tag{102}$$

We have found not one but two solutions! The most general solution to eqn 98 is the their sum:

$$x = A\mathrm{e}^{-\left(\gamma + \sqrt{\gamma^2 - \omega_0^2}\right)t} + B\mathrm{e}^{-\left(\gamma - \sqrt{\gamma^2 - \omega_0^2}\right)t}. \tag{103}$$

We should not be surprised to have found two solutions, as this gives us a general solution with two undetermined constants, A and B, which we can use to fix the initial displacement and velocity of the oscillation. The SHM solution also has two undetermined constants, a and ϕ, for this reason as, in general, will the solution to any second order ordinary differential equation — i.e. one containing second derivatives.

4.2.1 Heavy damping, $\gamma > \omega_0$

If $\gamma > \omega_0$ then the square root in eqn 102 is of a positive number and p is real. Our solution, eqn 103, is thus the sum of two exponentially decaying terms with different decay rates. The decay rates are set by the system, while the unknown constants, A and B, allow us to chose the initial displacement and velocity. The system is non-oscillatory. Some typical solutions for an oscillator with $\omega_0 = 1$ but with different levels of heavy damping are shown in fig. 35. All the solutions start with a pure displacement, $x(0) = 1$, $\dot{x}(0) = 0$. The higher the damping, the longer the system takes to get back to equilibrium.

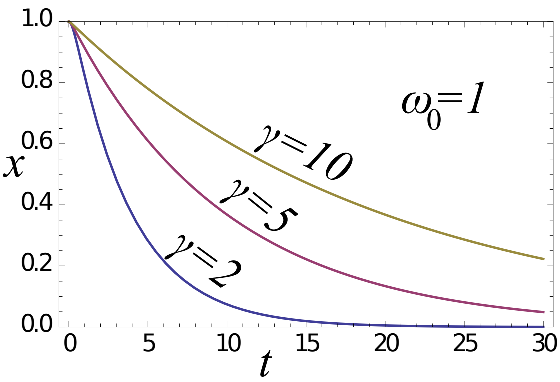

Figure 35: Heavily damped oscillators with different damping coefficients. Higher damping leads to slower relaxation to equilibrium.

Example: Heavily damped pendulum

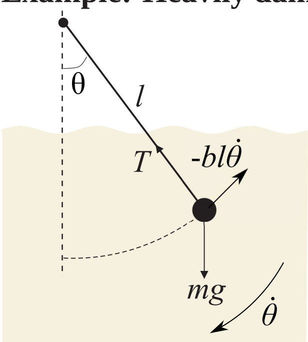

Figure 36: A simple pendulum moves in treacle.

A simple pendulum, sketched in fig. 36, has $l = 1\mathrm{m}$, $m = 1\mathrm{kg}$ and swings in treacle which exerts a viscous drag force on the mass $F = -bv$, with $b = 100\mathrm{Nsm}^{-1}$. If I start the pendulum at $\theta = 0.2\mathrm{rad}$ and with an inward velocity of 3m/s, what is the mass's subsequent motion?

As previously, the mass's velocity is $l\dot{\theta}$ and its acceleration is $l\ddot{\theta}$. Applying Newton's second law perpendicular to the pendulum gives

$$ml\ddot{\theta} = -mg\sin(\theta) - bl\dot{\theta}. \tag{104}$$

For small angles $\sin(\theta) \approx \theta$, so we can rearrange this to get

$$\ddot{\theta} + \frac{b}{m}\dot{\theta} + \frac{g}{l}\theta = 0, \tag{105}$$

which is eqn 99 with $\gamma = b/(2m) = 50\mathrm{s}^{-1}$ and $\omega_0^2 = g/l \approx 10\mathrm{s}^{-2}$. Therefore the solution is:

$$\theta = A\mathrm{e}^{-\left(\gamma + \sqrt{\gamma^2 - \omega_0^2}\right)t} + B\mathrm{e}^{-\left(\gamma - \sqrt{\gamma^2 - \omega_0^2}\right)t} \approx A\mathrm{e}^{-99t} + B\mathrm{e}^{-1t} \;\Longrightarrow\; \dot{\theta} \approx -99A\mathrm{e}^{-99t} - B\mathrm{e}^{-t}. \tag{106}$$

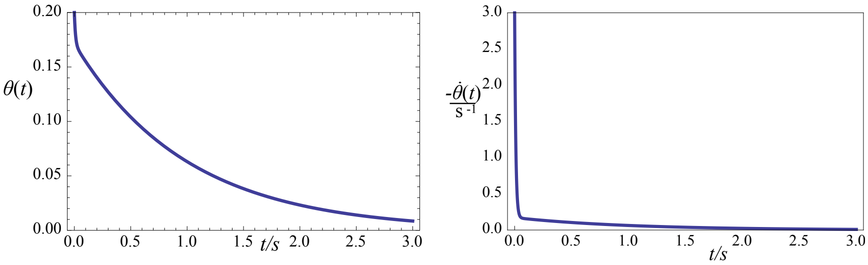

Figure 37: Decay of angle (left) and velocity (right) for the heavily damped pendulum. The initial velocity decays very rapidly whereas the angle decays very slowly.

At $t = 0$ we need $\theta = A + B = 0.2$ and $l\dot{\theta} = l(-99A - B) = -3$m/s, requiring $B \approx 0.17$ and $A \approx 0.03$, so the motion is

$$\theta = 0.03\mathrm{e}^{-99t} + 0.17\mathrm{e}^{-t}. \tag{107}$$

It is notable that the two terms in this solution decay at very different rates. Looking at the general solution, eqn 103, we see this is generic for extremely high damping. If $\gamma \gg \omega_0$ then $\sqrt{\gamma^2 - \omega_0^2} \approx \gamma$ so one of the decay rates tends to 2γ while the other tends to zero, exactly as we see here. The meaning of these two different decay rates becomes clear if we plot the angle and velocity of our damped pendulum, shown in fig. 37. The process starts with a very high velocity, generating lots of drag, which slows the pendulum down very rapidly. This decay gets faster if there is more drag, and accounts for our rapid decay rate. However, although the pendulum is brought close to rest very rapidly, it is still far from its equilibrium point. It then falls back to $\theta = 0$ very slowly as gravity works against drag. This process gets slower as drag gets higher, and accounts for our slow decay rate. We have two very different decay rates because velocities decay very rapidly but displacements decay very slowly.

4.2.2 Light damping, $\gamma < \omega_0$

In the case of light damping we expect our system to oscillate many times with slowly decreasing amplitude. At first sight this doesn't seem connected to our non-oscillatory solution in the heavily damped case, but actually complex numbers allow us to use exactly the same solution. If $\gamma < \omega_0$ the square root in eqn 102 is of a negative number (i.e. imaginary) so p is complex:

$$p = \gamma \pm \mathrm{i}\sqrt{\omega_0^2 - \gamma^2} = \gamma \pm \mathrm{i}\omega_d, \tag{108}$$

where $\omega_d = \sqrt{\omega_0^2 - \gamma^2}$. Our solution is now a complex number $A\mathrm{e}^{-pt} = A\mathrm{e}^{-\gamma t}\mathrm{e}^{\mathrm{i}\omega_d t}$. The first exponential in this expression is a regular real exponential decay, while the second is a complex exponential representing oscillation at ω_d as we saw in SHM. Allowing $A = a_0\mathrm{e}^{\mathrm{i}\phi}$ also to be

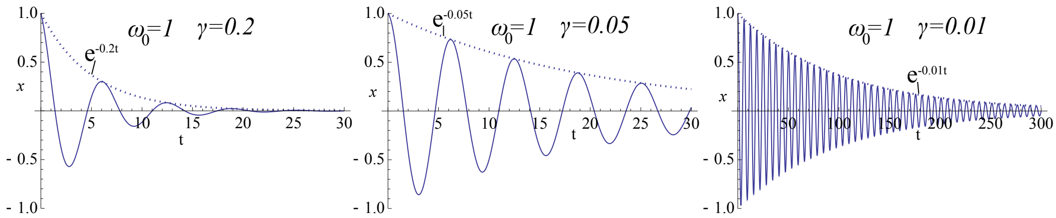

Figure 38: Motion of a lightly damped oscillator with different damping coefficients. Dotted lines show the decay of the oscillations as $\mathrm{e}^{-\gamma t}$. *N.B. the right-hand figure decays slowly and has a different time-axis.*

complex, the real (physical) part of the solution is

$$x = \mathrm{Re}\{Ae^{-\gamma t}\mathrm{e}^{\mathrm{i}\omega_d t}\} = a_0 e^{-\gamma t}\cos(\omega_d t + \phi), \tag{109}$$

which is oscillations at the damped angular frequency $\omega_d = \sqrt{\omega_0^2 - \gamma^2}$, (which for light damping is less than but close to ω_0) but with decaying amplitude $a_0 e^{-\gamma t}$. Interestingly the oscillations are still harmonic: the frequency of the oscillations does not change as their amplitude diminishes. The decay rate γ and the angular frequency ω are fixed for a given system, while a_0 and ϕ are set by the initial conditions. We might worry what happened to the second solution, $A\mathrm{e}^{-pt} = A\mathrm{e}^{-\gamma t}\mathrm{e}^{-\mathrm{i}\omega t}$, but in fact it has the same real part, and generates the same physical solution.

Some examples of lightly damped solutions (eqn 109) are plotted in fig. 38. As in the heavy damping case, these are all for systems with $\omega_0 = 1$ and started with pure displacement, $x(0) = 1$, $\dot{x}(0) = 0$, but with a range of levels of damping. We see that even for moderately heavy damping, the period of the oscillations $2\pi/\omega_d$ scarcely differs from $2\pi/\omega_0$, and that the lighter the damping, the more oscillations take place before the amplitude decays.

Example: Lightly damped pendulum

The pendulum of fig. 36 is now in a much less viscous fluid (air?) and hence is lightly damped with $b = 0.01\mathrm{Nsm}^{-1}$. *At* $t = 0$ *I do not displace the pendulum, but I give it an angular velocity of* $0.2\mathrm{rads}^{-1}$. *How large is the initial amplitude. How long does the amplitude take to fall by a factor of* e*?*

The equation for the pendulum is again

$$\ddot{\theta} + \frac{b}{m}\dot{\theta} + \frac{g}{l}\theta = 0, \tag{110}$$

which is eqn 99 with $\gamma = b/(2m) = 0.005\mathrm{s}^{-1}$ and $\omega_0^2 = g/l \approx 10\mathrm{s}^{-2} \gg \gamma^2$, so the pendulum is indeed lightly damped. It therefore follows a motion of the form of eqn 109 with $\omega_d = \sqrt{\omega_0^2 - \gamma^2} \approx \omega_0^2 \approx 10$. Since the displacement is zero at $t = 0$ we know $\phi = \pi/2$, so the solution is actually of the form

$$\theta = a_0\mathrm{e}^{-0.005t}\sin(10t) \implies \dot{\theta} = a_0\mathrm{e}^{-0.005t}(-0.005\sin(10t) + 10\cos(10t)). \tag{111}$$

The initial angular velocity is then $\dot{\theta}(0) = 10a_0 = 0.2$, requiring the initial amplitude to be $a_0 = 0.02$. The full motion is

$$\theta = 0.02\mathrm{e}^{-0.005t}\sin(10t). \tag{112}$$

The amplitude decays to a_0/e (i.e. to about 37% of a_0) when $0.005t = 1$, requiring $t = 200s$. The time period is $T = 2\pi/\omega \approx 0.63\mathrm{s}$, so this decay takes $200/0.63 \approx 320$ cycles.

4.2.3 Critical damping

If $\gamma = \omega_0$ then the square-root in eqn 102 is zero, and we only have one solution, $p = \gamma$. As in the heavy damping case, this is real so the solution is pure exponential decay $x = A\mathrm{e}^{-\gamma t}$. This cannot be the whole story as there is only one constant, but we need two since we can specify the mass's initial velocity and displacement. In fact, just in this case, there is an entirely different second solution,

$$x = A\mathrm{e}^{-\gamma t} + Bt\mathrm{e}^{-\gamma t}. \tag{113}$$

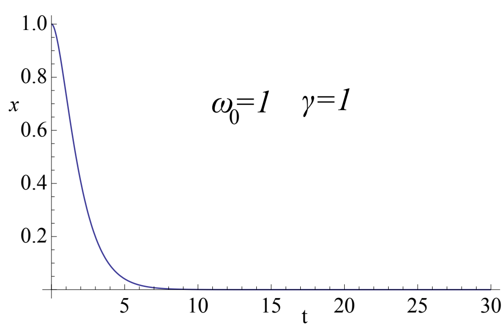

Figure 39: Critically damped oscillator.

You can verify this by substituting it into eqn 99.

We again plot an example of this solution for a system with $\omega_0 = 1$ and started with pure displacement ($x(0) = 1$, $\dot{x}(0) = 0$) in fig. 39. We see critically damped systems are the fastest decaying systems: they strike a balance between being overdamped, where friction prevents the mass from getting back to equilibrium, and underdamped where the mass oscillates for many cycles before the amplitude decays. Many damped systems are engineered to be critically damped so that they settle to their equilibrium point as quickly as possible: examples include car suspension, measuring instruments, weighing scales, closing doors and, latterly, the Millennium Bridge.

Example: Spring balance.

I am designing a spring balance to weigh babies, seen in fig. 40. When I put a typical 4kg baby on the spring it extends by 0.25m. What damping coefficient, b, do I need to critically damp my system? How long will it take the reading on the scale to settle down in this case?

This is just a mass on a spring, with $m = 4\text{kg}$. When the baby hangs in equilibrium there is a balance between gravity and the spring force $mg = kx = 0.25k$, so $k = 160\text{Nm}^{-1}$. The equation of motion is then just eqn 98, so we have $\gamma = \frac{b}{2m}$ and $\omega_0^2 = k/m = 40\text{s}^{-2}$. For critical damping we need $\gamma = \omega_0 = \sqrt{40}$, requiring $b \approx 50\text{Nsm}^{-1}$. The system will then decay as $\mathrm{e}^{-\gamma t} = \mathrm{e}^{-\sqrt{40}t}$, so the reading on the balance will settle down when $\sqrt{40}t \gtrsim 1$, or in time $t \gtrsim 0.16\text{s}$. If we damp less than this the baby will bounce for an extended period. If we damp more than this the spring will extend slowly and we will be waiting a long time for the reading.

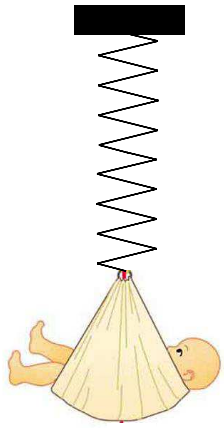

Figure 40: Baby being weighed on a spring balance.

4.3 Energy and amplitude decay

4.3.1 Energy dissipation

For a mass on a spring, the total energy, as before, is

$$E = KE + PE = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2. \tag{114}$$

However, unlike in SHM, with damping the total energy is not conserved. Its rate of change is:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = m\dot{x}\ddot{x} + kx\dot{x} = \dot{x}(m\ddot{x} + kx). \tag{115}$$

Recalling that the equation of motion for a damped mass on a spring is $m\ddot{x} + b\dot{x} + kx = 0$, we see that the rate of loss of energy is

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \dot{x}(-b\dot{x}) = -b\dot{x}^2. \tag{116}$$

This is exactly the work done by the frictional force: the friction force has magnitude $-b\dot{x}$ and the power is force times velocity. $\dot{E}$ is always negative as friction only removes energy.

4.3.2 Amplitude and energy dissipation for light damping

For light damping our mass and spring still performs many oscillations with displacement

$$x = a_0 e^{-\gamma t}\cos(\omega_d t + \phi). \tag{117}$$

An example of this motion is plotted in fig. 41. The mass's maximum displacement in each cycle occurs when $\omega t + \phi = 0$, which occurs periodically with time period $T = 2\pi/\omega$. If the nth maxima occurs at time t_n, its displacement, which we call the amplitude of the oscillator at t_n, is

$$a_n = a_0 e^{-\gamma t_n}. \tag{118}$$

At this point the mass is stationary, so all the energy is potential energy, and the total energy is

$$E_n = \frac{1}{2}kx^2 = \frac{1}{2}ka_0^2 e^{-2\gamma t_n}. \tag{119}$$

The amplitude decays in time as $\mathrm{e}^{-\gamma t}$, while the energy decays at twice the rate, as $\mathrm{e}^{-2\gamma t}$.

N.B. We have only calculated the energy at the maximum of each cycle. We could however work out the the energy at any time by substituting eqn 117 into eqn 114.

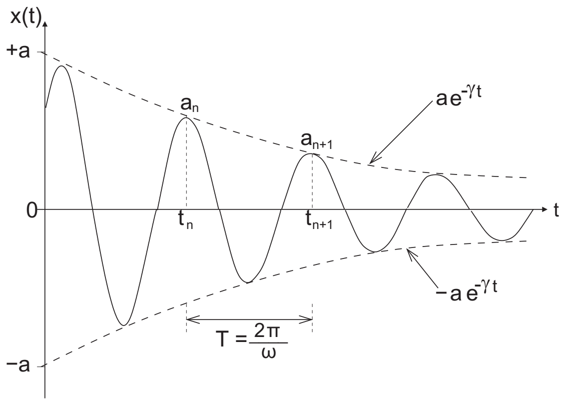

Figure 41: Successive maximum displacements in a lightly damped oscillator.

4.4 Comparing oscillators

We are often interested to know how good an oscillator is, by which we mean how many oscillations it performs before its amplitude decays. We saw earlier that the general damped harmonic equation (eqn 99) contains two characteristic times, a decay time $\tau = 1/(2\gamma)$ and an oscillation period $T = 2\pi/\omega_0$. The ratio of these two, τ/T, is an estimate of the number of oscillations performed in one decay time. If this ratio is large damping is very light and the oscillator is very good, performing many oscillations before the amplitude decays. However, convention dictates that we in fact measure the quality of oscillators using two slightly different measures, the logarithmic decrement and the quality factor.

4.4.1 The logarithmic decrement

The logarithmic decrement measures how much the amplitude of a lightly damped oscillator falls by in one cycle. From eqn 118, the ratio of the amplitude of successive oscillations is

$$\frac{a_{n+1}}{a_n} = \frac{\mathrm{e}^{-\gamma t_{n+1}}}{\mathrm{e}^{-\gamma t_n}}. \tag{120}$$

However, the time period for oscillations is $T = 2\pi/\omega$, so we know $t_{n+1} = t_n + 2\pi/\omega$, giving

$$\frac{a_{n+1}}{a_n} = \frac{\mathrm{e}^{-\gamma(t_n+T)}}{\mathrm{e}^{-\gamma t_n}} = \mathrm{e}^{-\gamma T} = \mathrm{e}^{-\frac{2\pi\gamma}{\omega_d}}. \tag{121}$$

The logarithm of this is called the logarithmic decrement

$$\Delta = \log\left(\mathrm{e}^{-\frac{2\pi\gamma}{\omega_d}}\right) = \frac{2\pi\gamma}{\omega_d}. \tag{122}$$

The amplitude of successive oscillations decays by a factor of $e^{-\Delta}$, and the energy of successive oscillations decays by a factor of $e^{-2\Delta}$. A good oscillator has a small Δ.

The logarithmic decrement is often relatively easy to measure as it just requires us to measure the ratio of successive amplitudes.

4.4.2 The Quality factor

The quality factor of an oscillator is defined as

$$Q = \frac{\omega_0}{2\gamma}. \tag{123}$$

For a good oscillator $\omega_0 > \gamma$ so Q is very large. We initially estimated that the number of oscillations an oscillator would perform before significantly decaying is $\tau/T = \omega_0/(4\pi\gamma)$. The quality factor is simply larger than this by a factor of 2π, so it estimates the numbers of radians of oscillation rather than the number of cycles.

For high quality oscillators we have $\omega_d = \sqrt{\omega_0^2 - \gamma^2} \approx \omega_0$, i.e. they oscillate at close to their undamped frequency. This allows us to relate the quality factor and the logarithmic decrement,

$$\Delta = \frac{2\pi\gamma}{\omega_d} \approx \frac{2\pi\gamma}{\omega_0} = \frac{\pi}{Q}. \tag{124}$$

This is useful because Δ is easy to measure, but most people prefer to think in terms of Q.

The energy decays by $\mathrm{e}^{-2\Delta}$ each cycle, so, if we start with E_0, after n cycles we have left

$$E = E_0 \left(\mathrm{e}^{-2\Delta}\right)^n = E_0 \mathrm{e}^{-2n\Delta}. \tag{125}$$

The number of cycles required for the energy to fall by a factor of e (i.e. to $1/\mathrm{e}$ or 37%) is

$$N = \frac{1}{2\Delta} = \frac{Q}{2\pi} \implies Q = 2\pi N. \tag{126}$$

We see that Q is the number of radians of oscillation required for the energy to fall by a factor of e. It takes half as many radians for the amplitude to fall by the same factor.

4.4.3 Quality factor of Big-Ben

Big Ben is a bell that, when struck, rings at around 100Hz for around 3s. Estimate Big Ben's quality factor.

The bell rings at 100Hz, so in 3s it performs $3 \times 100 = 300$ cycles, i.e. the oscillation takes 300 cycles to die down. However, Q is the number of radians not the number of cycles, so $Q = 2\pi \times 300 \approx 1800$.

4.4.4 Quality factor of a radiating atom

Atoms emit light via quantum transitions. We can model this as the atom oscillating at the frequency of the emitted light, just as Big-Ben oscillates at the frequency of the emitted sound. The oscillation is damped because the light carries away energy. If, during a transition, an atom emits a 3m long wave-train of visible light with wavelength $\lambda = 500$nm, what is its quality factor.

The 3m long wave train contain $3/(5 \times 10^{-7}) = 6 \times 10^6$ wavelengths, so the atom undergoes 6×10^6 cycles before its amplitude significantly decays. Again the quality factor is the number of radians, not the number of cycles, so $Q = 2\pi \times 6 \times 10^6 \approx 4 \times 10^7$. The atom is a much better oscillator than Big Ben.

5 Forced Oscillations

Many interesting oscillating systems are driven by external forces. The classic example is a child on a swing. The swing is effectively a pendulum, and will, if the parent gives it a single push, perform damped oscillations. However, the child has more fun if the parent pushes the swing periodically, so that it keeps swinging with high amplitude for an extended period. This is an example of a forced or driven oscillator: if we apply a periodic force to a damped oscillator we can keep it oscillating indefinitely. Forced oscillations are everywhere, for example the pendula in clocks are driven by wound springs, the water molecules in food undergo forced oscillations when microwaved, and the air in a trumpet undergoes forced oscillations when it is blown.

However, as always, we start our analysis with a horizontal mass on a spring (fig. 42). We add an oscillatory driving force applied to the mass $F_d = F_0 \cos(\omega t)$, so Newton's second law becomes

$$F_s + F_f + F_d = m\ddot{x} \implies -kx - b\dot{x} + F_0 \cos(\omega t) = m\ddot{x}. \tag{127}$$

We can rearrange this to get

$$\ddot{x} + \frac{b}{m}\dot{x} + \frac{k}{m}x = \frac{F_0}{m}\cos(\omega t), \tag{128}$$

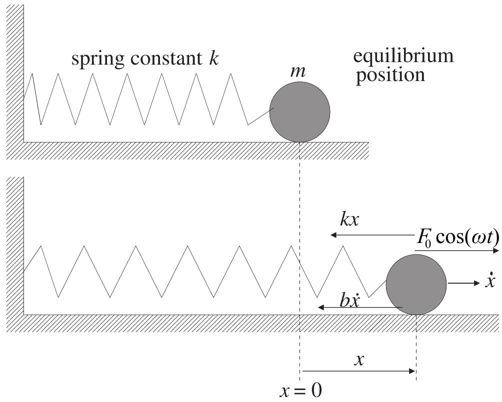

Figure 42: Damped driven mass on a spring. Top: Mass at rest in equilibrium. Bottom: Mass at x and with velocity $\dot{x}$. It feels a spring force $-kx$ a friction force $-b\dot{x}$ and a driving force $F_0 \cos(\omega t)$.

which is in the form of the general equation for the damped driven harmonic oscillator

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = f\cos(\omega t), \tag{129}$$

with $\gamma = b/(2m)$, $\omega_0^2 = k/m$ and $f = F_0/m$. Again, before solving formally, it is worth asking what we expect to happen. If we start from rest, at first the the amplitude of the oscillations will build up, but this growth should not continue indefinitely: eventually, the mass will settle down into a steady oscillation at the driving frequency. We also know from our experience with swings that the amplitude of the swing's oscillations are greatest if we match the frequency of our pushes with the natural frequency of the swing, so in general we expect the amplitude of the steady oscillations to depend not only on the strength of our driving, but also on its frequency. Returning to eqn 129, in the long run we are expecting the mass to oscillate steadily at the driving frequency, so we try a solution of the form

$$x = a_0 \cos(\omega t + \phi). \tag{130}$$

Substituting this proposed solution into eqn 129 gives

$$a_0 \left((\omega_0^2 - \omega^2)\cos(\omega t + \phi) - 2\gamma\omega \sin(\omega t + \phi) \right) = f\cos(\omega t). \tag{131}$$

5.1 Low frequency response

If we apply a constant force F to a mass on a spring, we know that, eventually, the spring will settle down with constant extension $x = F/k$ as given by Hooke's law. If we then change the force a little and wait, the spring will settle down with a slightly different extension. Extending this idea to an oscillating force $F = F_0 \cos(\omega t)$, if the force varies sufficiently slowly the mass will always be essentially stationary, in equilibrium between the applied force and the spring force, $-kx = F_0\cos(\omega t)$, so the extension will be

$$x = \frac{F_0\cos(\omega t)}{k}, \tag{132}$$

i.e. the displacement of the mass has amplitude F_0/k and is in-phase with the driving force. Mathematically, we can see this regime emerge from eqn 131 by realizing that at, very low frequencies, we can neglect the terms proportional to ω and ω^2, leaving

$$a_0\omega_0^2\cos(\omega t + \phi) = f\cos(\omega t), \tag{133}$$

which is clearly solved by $\phi = 0$ and $a_0 = f/\omega_0^2$, giving the solution

$$x = \frac{f}{\omega_0^2}\cos(\omega t) = \frac{F_0\cos(\omega t)}{k}, \tag{134}$$

where the latter equality is for the mass-on-a-spring case. The key idea is that, if the driving force is very slow, the mass's velocity and acceleration are negligible, so Newton's second law turns into a force balance between the applied force and the driving force. *At low frequencies, the resistance to the driving force is entirely provided by the spring.*

5.2 High frequency response

We next imagine driving the mass at very high frequency. If, as sketched in fig. 43, we apply a rapidly oscillating force $F_0\cos(\omega t)$ to an isolated mass Newton's second law gives

$$m\ddot{x} = F_0\cos(\omega t), \tag{135}$$

which, integrating twice with respect to time, is solved by

$$x = -\frac{F_0\cos(\omega t)}{m\omega^2}. \tag{136}$$

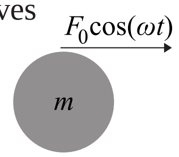

Figure 43: Isolated mass driven by a force $F_0\cos(\omega t)$.

We see the mass's displacement is always opposite to the driving force and, as the frequency gets high, the amplitude of the oscillations vanishes. Taking a derivative reveals the mass's velocity also vanishes. If the mass is on a spring, driving the mass at high frequency results in negligible displacements and velocities, so neither the spring force nor the damping force are relevant, and the displacement is just given by eqn 136. We can also see this emerge from eqn 131 since, if ω is very large, the left side is dominated by the term proportional to ω^2, so it reduces to

$$-a_0\omega^2\cos(\omega t + \phi) = f\cos(\omega t). \tag{137}$$

This is solved by $\phi = -\pi$ and $a_0 = f/\omega^2$, (we could also choose $\phi = 0$ and $a_0 = -f/\omega$, but we prefer to keep a_0 positive), so the displacement is

$$x = \frac{f\cos(\omega t - \pi)}{\omega^2} = -\frac{F_0\cos(\omega t)}{m\omega^2}, \tag{138}$$

where the latter equality is again for the mass-on-a-spring case. The key is that at high frequency the mass's velocity and displacement are negligible, so drag and spring force are negligible and Newton's second law reduces to a balance between the applied force and the acceleration. *At high frequencies the resistance to the driving force is provided by the mass's inertia.*

5.3 Resonant response

If we imagine driving a mass on a spring at exactly $\omega = \omega_0 = \sqrt{k/m}$ something interesting happens. The system oscillates at ω_0, but at this frequency of oscillation, independent of amplitude, there is always a perfect balance between the spring force and the acceleration, $\ddot{x} + \omega_0^2 x = 0$, so neither the spring force nor the acceleration can counter-balance the driving force in Newton's second law. If the system were undamped, the amplitude of the oscillations would simply diverge. However, real systems are damped, albeit often lightly, so when the oscillation gets big enough the damping force can counterbalance the driving force, and the system reaches a steady high amplitude state. This is called resonance. We can see it clearly in eqn 131: if we drive the system at exactly $\omega = \omega_0$ then the acceleration and spring terms exactly cancel and we are left with

$$-2a_0\gamma\omega_0 \sin(\omega_0 t + \phi) = f\cos(\omega_0 t), \tag{139}$$

solved by $a_0 = f/(2\gamma\omega_0)$ and $\phi = -\pi/2$ (again we want a_0 to be positive), so the displacement is

$$x = \frac{f\cos(\omega_0 t - \pi/2)}{2\gamma\omega_0}. \tag{140}$$

The displacement is quarter of a cycle behind the driving force, and the amplitude of the response is only limited by the damping, so a good oscillator will produce a very high amplitude response. This effect is called resonance. *At resonance the resistance to the driving force is provided entirely by the damping*.

The $\pi/2$ phase shift between displacement and driving force means that the driving force is in-phase with velocity, and thus does work each cycle, adding energy to the system. In the steady state, this addition is exactly balanced by the energy dissipated by the damping.

5.4 Steady state solution at any frequency

We have just seen that the phase of the displacement shifts with the driving frequency. This is exactly the sort of problem where complex numbers are particularly useful. The physical solution is the real part of $z = A\mathrm{e}^{\mathrm{i}\omega t}$ (where $A = a_0\mathrm{e}^{\mathrm{i}\phi}$ encodes both the amplitude and the phase of the displacement), and the force is the real part of $f\mathrm{e}^{\mathrm{i}\omega t}$. This leads us to write eqn 129 as

$$\ddot{z} + 2\gamma\dot{z} + \omega_0^2 z = fe^{\mathrm{i}\omega t}. \tag{141}$$

If we can solve this equation, the real part of the solution will be the physical solution to eqn 129. Substituting in our proposed solution, $z = \mathrm{e}^{\mathrm{i}\omega t}$, reduces this equation to an algebraic one

$$-\omega^2 A\mathrm{e}^{\mathrm{i}\omega t} + 2\mathrm{i}\gamma\omega A\mathrm{e}^{\mathrm{i}\omega t} + \omega_0^2 A\mathrm{e}^{\mathrm{i}\omega t} = fe^{\mathrm{i}\omega t}, \tag{142}$$

from which we can cancel a factor of $\mathrm{e}^{\mathrm{i}\omega t}$ to get

$$A(-\omega^2 + 2\mathrm{i}\gamma\omega + \omega_0^2) = f \quad \Longrightarrow \quad A = \frac{f}{\omega_0^2 - \omega^2 + 2\mathrm{i}\gamma\omega}. \tag{143}$$

To recover the physical solution, we need to cast this is the form $a_0\mathrm{e}^{\mathrm{i}\phi}$. We first check we can recover our three previous special cases. If ω is very small we have $A \approx f/\omega_0^2$. This is real and positive, so to write it in the form $a_0\mathrm{e}^{\mathrm{i}\phi}$ we simply need $a_0 = f/\omega_0^2$ and $\phi = 0$. At very high frequencies we have $A \approx -f/\omega^2$. This is real and negative, so we need $a_0 = f/\omega^2$ and $\phi = -\pi$. When $\omega = \omega_0$ we have $A = f/(2\mathrm{i}\gamma\omega_0) = -\mathrm{i}f/(2\gamma\omega_0)$. This is negative imaginary, so it needs $\phi = -\pi/2$ (since $\mathrm{e}^{-\mathrm{i}\pi/2} = -\mathrm{i}$)and $a_0 = f/(2\gamma\omega_0)$. We thus recover all three of

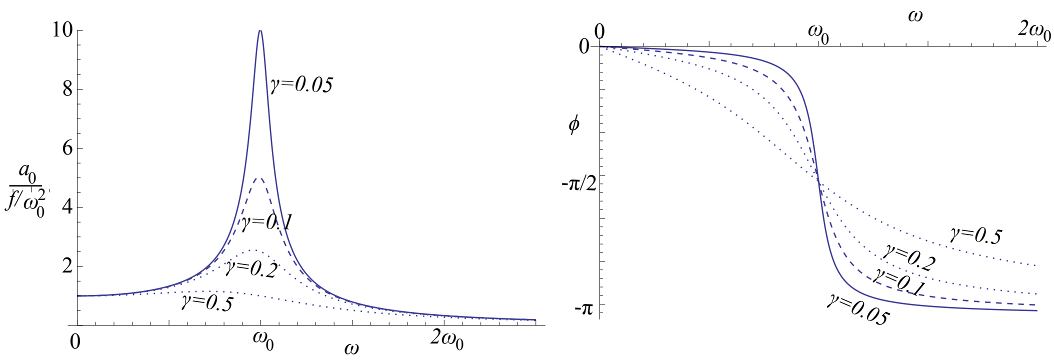

Figure 44: Amplitude (left) and phase (right) of a damped harmonic oscillator as a function of driving frequency. At low driving frequency the amplitude is low but finite, and the system oscillates in phase with the driving force. When driven at ω_0 the system resonates, responding with high amplitude and quarter of a cycle behind the force. At high driving frequency the amplitude decays to nothing and the displacement is half a cycle out of phase with the driving force. Systems with lower damping have higher resonances and the phase shifts from 0 to $-\pi$ more sharply.

our special cases correctly. More generally, the complex number in A's denominator has length $|\omega_0^2 - \omega^2 + 2i\gamma\omega| = \sqrt{(\omega_0^2 - \omega^2)^2 + 4\gamma^2\omega^2}$, and makes an angle $\tan(\beta) = 2\gamma\omega/(\omega_0^2 - \omega^2)$ with the real axis. Therefore A has length and phase given by:

$$a_0 = |A| = \frac{f}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\gamma^2\omega^2}}, \qquad \tan(\phi) = \frac{2\gamma\omega}{\omega^2 - \omega_0^2}. \tag{144}$$

N.B. We must be careful inverting this for ϕ as it varies between 0 and $-\pi$, but most arctan functions only return values between $-\pi/2$ and $\pi/2$.

In fig. 44 we plot the amplitude and phase for driven oscillators with a range of damping coefficients, as a function of driving frequency. We see the high-amplitude resonance emerge in each case as a peak around $\omega = \omega_0$, with impressively high amplitude when the damping is low. During resonance the phase shifts from 0 to $-\pi$. The shift is sharper with lighter damping.

5.4.1 Power and Resonance

In a driven oscillating system the driving force puts energy into the oscillation while the damping removes it. In the steady state the amount of energy in the system is constant, so these two processes must balance: all the energy added by the driving force is dissipated by the damping. To calculate how much energy is being dissipated, we first work out the velocity:

$$\dot{x} = -a_0\omega\sin(\omega t + \phi) = \frac{-f\omega\sin(\omega t + \phi)}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\gamma^2\omega^2}} = \frac{-f\sin(\omega t + \phi)}{\sqrt{((\omega_0^2 - \omega^2)/\omega)^2 + 4\gamma^2}}. \tag{145}$$

In the mass-and-spring case, the friction force is $b\dot{x}$, so the instantaneous power dissipation is $b\dot{x}^2$. Since the average of $\sin^2(\omega t + \phi)$ over one cycle is one half, the average rate power dissipation is

$$\langle b\dot{x}^2 \rangle = \frac{1}{2} b \frac{f^2}{((\omega_0^2 - \omega^2)/\omega)^2 + 4\gamma^2}. \tag{146}$$

In fig. 45a we plot the power absorbed by three oscillators with different damping as a function of driving frequency. We see a resonance peak in the power absorbed when the oscillators are driven at ω_0, and that the reducing the damping dramatically increases this peak power.

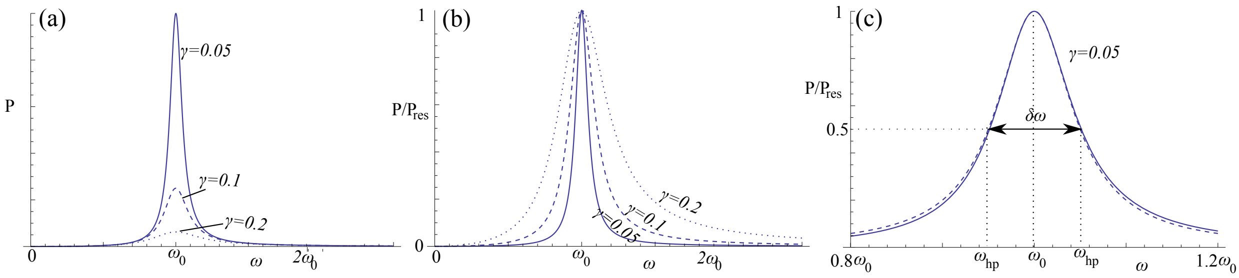

Figure 45: (a) Power absorbed by damped driven oscillators as a function of driving frequency. The maximum power is absorbed at the resonant frequency, ω_0 and the magnitude of the maximum power increases if the damping coefficient γ is decreased. (b) Same curves as (a) but with each curve normalized to the same height. We see that the resonance peaks are thinner for the higher quality oscillators. (c) Zoomed in graph of the same power curve for the least damped system, $\gamma = 0.05$, showing the definition of $\delta\omega$, the width of the peak at half maximum power. The dashed line is the Lorentzian power spectrum given by eqn 147 instead of eqn 146, which agrees very well.

5.4.2 Example: IR Absorption Spectra

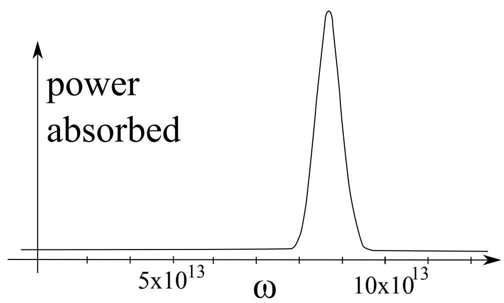

Figure 46: HCl's infrared absorption spectrum.

Molecules can be thought of as assemblies of masses (atoms) connected by springs (bonds), so they exhibit many vibrational modes of oscillation, each with a characteristic frequency. If we fire a beam of light into a beaker containing many identical molecules it drives these oscillations at its frequency. If we then measure how much of the light passes through the beaker, we know how much power the oscillations have absorbed. The absorption at each frequency is described by eqn 146 so, if we scan through frequencies, we can plot a graph like fig. 45, showing high absorption peaks at each natural frequency. Such a graph is called an infra-red absorption spectrum because the vibrational frequencies lie in the infra-red. From the resonant frequencies we can work out which molecule we have in our beaker. Fig. 46 shows the absorbance spectrum of HCl, which, as we saw earlier, we can model as two masses connected by a spring, so the spectrum shows a single big peak at the resulting resonant frequency.

5.4.3 Lorentzian Peaks in Spectra (Non-examinable)

Other types of spectrography depend on resonances of different oscillations, for example nuclear-magnetic resonance depends on atomic nuclei oscillating their magnetic moment in a magnetic field. In all these spectra, we see peaks described by eqn 146 poking out of background noise. Since we can't see the peak's tails, we describe them with a simplified version of eqn 146 that only applies around the peak. The ω dependent term in the denominator is $((\omega_0^2 - \omega^2)/\omega)^2 = (\omega - \omega_0)^2(\omega + \omega_0)^2/\omega^2$. When $\omega \sim \omega_0$ (i.e. near resonance) this is well approximated by

$$\frac{(\omega-\omega_0)^2(\omega+\omega_0)^2}{\omega^2} \approx \frac{(\omega-\omega_0)^2(2\omega_0)^2}{\omega_0^2} = 4\omega_0^2(\omega-\omega_0)^2 \implies \langle b\dot{x}^2\rangle \approx \frac{1}{8}b\frac{f^2}{(\omega-\omega_0)^2+\gamma^2}. \tag{147}$$

This symmetric function, which describes the peak, is called a Lorentzian or a Cauchy distribution.

5.4.4 Resonance width and the quality factor

In fig. 45b we plot the same absorption curves as in the left-most one, but normalized so they all have the same height. This reveals a second feature of resonance: the width of the resonant peak becomes narrower as the oscillator gets better. This means that, if we wish to cause a low damping oscillator to resonate, we need to drive it very close to its resonant frequency.

We quantify this by asking how far from the resonant frequency our driving frequency can be before the total power absorbed falls by a half. We call the width of the resonant peak at half power $\delta\omega$, as sketched in fig. 45c. From eqn 146, we see that if we drive the oscillator at exactly ω_0 the power absorbed is $f^2 b/(8\gamma^2)$. To reduce this by half we need $(\omega_0^2 - \omega^2)/\omega = \pm 2\gamma$, a quadratic equation for ω solved by

$$\omega_{hp} = \mp\gamma + \sqrt{\omega_0^2 + \gamma^2} \quad \text{(we take the positive discriminant so } \omega_{hp} > 0\text{)}. \tag{148}$$

The resonant peak thus has width $\delta\omega = 2\gamma$. Recalling the definition of the quality factor, we have

$$\frac{\delta\omega}{\omega_0} = \frac{2\gamma}{\omega_0} = \frac{1}{Q}. \tag{149}$$

Although high quality oscillators have dramatic resonances, they only occur if you drive them very close to their resonant frequency.

5.4.5 Example: Breaking a wine glass

*When I strike a wine glass, it rings at 700*Hz*, and the sound decays*[3] *in half a second. If I seek to shatter the glass by singing at its resonant frequency, how accurate a singer do I need to be?*

In half a second, the glass performs $750 \times 0.5 = 350$ cycles. Quality factor is the number of radians for the oscillation to decay, so it is $Q \approx 350 \times 2\pi \approx 2100$. The glass is thus a very good oscillator, so when struck it rings very close to its natural frequency ω_0. To break the glass I must sing at this resonant frequency. My margin for error is given by

$$\frac{\delta\omega}{\omega_0} = \frac{\delta\nu}{\nu_0} = \frac{1}{Q} \approx 0.0005 \implies \delta\nu = 0.0005\nu_0 = 0.3\text{Hz}. \tag{150}$$

At this frequency a semitone is around 20Hz, so this is difficult. It's easier if you simultaneously sing at many glasses, each with a slightly different resonant frequency.

5.5 Transients and the quality factor (Non examinable)

Thus far we have only dealt with steady state driven oscillations. However, we are also interested in how long it takes our system to reach the steady state: if I drive a system at the resonant frequency how long do I have to wait for the amplitude reach its final value? Our general damped driven oscillator is governed by by the equation

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = f\cos(\omega t). \tag{151}$$

We can write the general solution to this equation as $x = x_{ss} + x_d$, the sum of the steady state solution, x_{ss}, which we have already found, and a part decaying part, x_d. However eqn 152 is linear in x so substituting this form in gives

$$(\ddot{x}_{ss} + 2\gamma\dot{x}_{ss} + \omega_0^2 x_{ss}) + (\ddot{x}_d + 2\gamma\dot{x}_d + \omega_0^2 x_d) = f\cos(\omega t). \tag{152}$$

[3]We actually hear the glass ring for several decay times, but this is because our ears are amazing and can perceive sound over many magnitudes of amplitude.

We know that $\ddot{x}_{ss} + 2\gamma\dot{x}_{ss} + \omega_0^2 x_{ss} = f\cos(\omega t)$, so this reduces to $\ddot{x}_d + 2\gamma\dot{x}_d + \omega_0^2 x_d = 0$. This is the equation for a damped undriven oscillator, which we already know is solved by $x_d = a_d e^{-\gamma t}\cos(\omega_d t + \phi_d)$, where $\omega_d^2 = \omega_0^2 - \gamma^2$ and a_d and ϕ_d are constants. The full solution is thus

$$x = x_{ss} + x_d = x = a_0\cos(\omega t + \phi) + a_d e^{-\gamma t}\cos(\omega_d t + \phi_d), \tag{153}$$

where the steady state amplitude and phase, a_0 and ϕ, are fixed by eqn 144, while a_d and ϕ_d are fixed by the initial displacement and velocity. The system reaches its steady state when x_d has decayed which, since it is the same solution, takes the same length of time as undriven oscillations in the system take to decay. In this time the system conducts about Q radians of oscillation. Thus *Q is also the number of radians of oscillation a driven system needs to reach its steady state.*

5.5.1 Example: Breaking a wine glass II

If, to break the wine glass from the previous example, I need the full amplitude of the resonance, how long must I sustain the accurate note for?

When struck the glass's oscillations decay in 0.5s. The decaying part of eqn 153 thus also decays in 0.5s so, to reach the steady state (and break the glass) I need to sing for at least 0.5s.

5.6 More examples of resonance

- **Mechanical resonance in bridges:** Engineers must be careful to design bridges that don't have resonances that are driven by the bridge's environment. The Millennium Bridge in London has a resonant frequency close to walking frequency, and people walking on it initially caused it to sway disconcertingly. The bridge was closed while damping was added to reduce the resonant amplitude. More dramatically, the Tacoma Narrows Bridge in Washington State USA had a resonant frequency excited by the wind, causing it to collapse. You can watch this at `www.youtube.com/watch?v=j-zczJXSxnw`.

Figure 47: Tacoma Narrows Bridge. Still from film taken by Barney Elliott

Figure 48: The Bay of Fundy in Canada. Map by Decumanus at en.wikipedia.

- **Tidal resonances:** The Bay of Fundy in Canada (fig. 48) is a shallow rectangular bay facing the ocean. It takes about 12 h for a wave to travel from the bay's mouth to its back, reflect, and return to the mouth. High tide is every 12 h 25mins, providing a periodic forcing at close to the bay's natural frequency, so the the wave resonates and builds up a large amplitude. In the open ocean tides are less than a meter tall; in the Bay of Fundy tides reach 16m.

- **Acoustic resonances in musical instruments:** A trumpet is a column of air with certain resonant frequencies. When the player drives the trumpet, by buzzing with his/her lips into the mouthpiece, the trumpet produces a loud sound, but only if the driving (buzzing) frequency matches one of the trumpet's resonant frequencies.

- **Electrical resonances in Radio and TV Tuners:** These have electrical resonances which are used to selectively amplify the frequency of the broadcast signal, but not unwanted channels or noise. These resonances need a high Q so the resonant peak is thin and we only amplify a very small frequency range. Traditional radios are tuned by changing the resonant frequency until it matches the channel. Modern TVs and radios are more complicated.

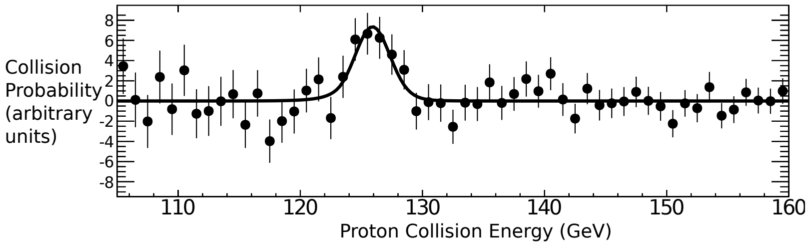

Figure 49: Higgs Boson Resonance from CERN

- **Higgs Boson Resonance:** According to the standard model of particle physics, a Higgs-Field permeates all of space, and it has a resonant frequency of oscillation. If we fire protons together with a range of energies (which in quantum mechanics are equivalent to frequencies, related via $\hbar$) then, when we approach this resonant energy/frequency, the protons drive an oscillation of the Higgs Field, which we see as them colliding to form a Higgs Boson. This shows up in our experiment as an increase in the probability that the protons collide that looks just like a resonance peak. CERN recently found this peak, shown in fig. 49, verifying the existence of the Higgs Boson.

6 Electrical Circuits

6.1 Charge and Coulomb's Law

Some fundamental particles carry a property called charge, which comes in two types, positive and negative. Most matter we encounter contains an almost exactly equal number of protons, which are positivity charged, and electrons, which are equally but negatively charged, so the matter is electrically neutral. However, electrons are sufficiently mobile that we can easily move some from one object to another, leaving both objects carrying a net electrical charge. When we do this, we discover that objects carrying the same sign of charge repeal each other, while those carrying opposite sign charges attract each other. If, as sketched in fig. 50, one object carries a charge q_1 and the other q_2, and they are a distance r apart, the magnitude of this force is given by Coulomb's law,

$$F \propto \frac{q_1 q_2}{r^2}, \qquad (154)$$

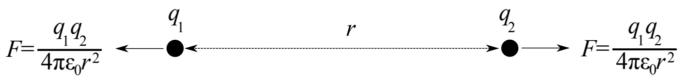

Figure 50: Coulomb force between two charged particles.

where a negative force denotes attraction and a positive force repulsion. If we work in SI units (i.e. we measure length in meters and charge in Coulombs) then the constant of proportionality is $\frac{1}{4\pi\epsilon_0}$, where $\epsilon_0 = 8.854187817... \times 10^{-12}\mathrm{F/m}$. In this unit system, the charge on a proton is $e = 1.60217657^{-19}\mathrm{C}$, while the charge on an electron is $-e$.

We interpret this force as follows. There exists at every point in space a vector $\mathbf{E}$, which we call the electric field, and if we put a particle with charge q into an electric field, it feels a force

$$\mathbf{F} = q\mathbf{E}. \qquad (155)$$

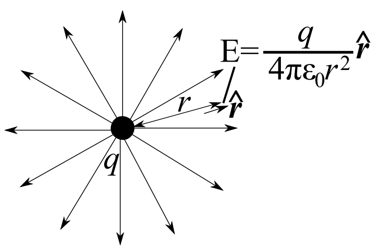

Figure 51: Electric field from caused by a positive charge.

However, charges also *produce* an electric field, which points radially away from them if they are a positive and radially towards them if they are negative. More specifically, a distance r from a charge, it produces an electric field

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0 r^2}\hat{\mathbf{r}}, \qquad (156)$$

where $\hat{\mathbf{r}}$ is a unit-length vector pointing away from the charge, as sketched in fig. 51. Putting these two results together to find the force on one charge because of the electric field of a second charge gives us back Coulomb's law. This should remind you of Newton's law of gravitation, although with the important difference that we can have positive and negative charge, whereas mass is always positive.

You will study static fields and charges extensively in Easter Term. However, in this course we are going to jump ahead of ourselves and think about flows of charge around circuits. These underpin all electronic devices, so their importance is hard to overstate.

6.2 Current

An electrical conductor is a material which contains mobile charge carriers. Typically this is a metal containing free electrons, though it could also be a fluid containing mobile ions. When we apply an electric field to a conductor, the charges feel a force $q\mathbf{E}$ and the mobile ones consequently accelerate along $\mathbf{E}$. If the mobile charges were in a vacuum they would accelerate indefinitely.

However, they are actually in amongst densely packed atoms (for metallic conductors, a lattice of atoms) with which they undergo energy-dissipating inelastic collisions. The mobile charge carriers thus reach a finite average terminal velocity, known as the drift velocity.

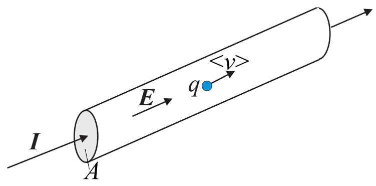

Figure 52: Electric field in a wire exerts a force on mobile charges, causing them to move with average velocity $\langle v \rangle$.

Consider a wire with cross-sectional area A, shown in fig. 52, containing n mobile charges per unit volume, each carrying a charge q. If we apply an electric field along the of the wire, the charge carriers drift along the wire with average velocity $\langle v \rangle$. In a time t, the charge in a volume $A \langle v \rangle t$ passes an observer standing alongside the wire, which is a total charge $Q = nqA \langle v \rangle t$. The rate at which charge flows past the observer is thus

$$I = \frac{\mathrm{d}Q}{\mathrm{d}t} = nqA \langle v \rangle \,, \tag{157}$$

which we call the electric current, I. If $\langle v \rangle$ changes as a function of time so will the current. The total charge that flows past is then given by integrating eqn 157:

$$Q = \int I \mathrm{d}t. \tag{158}$$

In the special case where I is constant, we can trivially integrate this to get

$$Q = It. \tag{159}$$

It is important to understand that this does not mean the wire is becoming charged: Q indicates how much charge has flowed through the wire in a time t, but in every section of wire Q flows in and Q flows out so the wire remains neutral.

The SI unit of current is the Ampere, which is one Coulomb flowing past each second.[4] Since the charge on an electron is $-1.6 \times 10^{-19} C$, if we have a current on 1A, then we have $1/(1.6 \times 10^{-19}) \sim 6 \times 10^{18}$ electrons flowing past each second.

6.2.1 Example: speed of electrons in a current

A copper wire with cross-sectional area $A = 1\mathrm{mm}^2$ *carries 1 A of current. How fast do the electrons drift? Copper has one free electron per atom, density* $\rho = 9\mathrm{g/cm}^3$ *and atomic weight 63.5*$\mathrm{g/mol}$.

A cubic centimeter of copper weighs 9 g, so it contains $9/63.5 \sim 0.14$ moles of copper. Recalling that Avogadro's number is $N_A \approx 6 \times 10^{23}$, this is 8.5×10^{22} copper atoms. Each copper atom contributes one free electron, so the density of free electrons is $n = 8.5 \times 10^{22}\mathrm{cm}^{-3}$. The current flowing is $I = -neA \langle v \rangle = 1\mathrm{A}$, and we know $e = 1.6 \times 10^{-19} C$ and $A = 1\mathrm{mm}^2 = 0.01\mathrm{cm}^2$, so the drift velocity is $\langle v \rangle = 1/(nqA) \approx -0.01\mathrm{cms}^{-1}$, or a tenth of a millimeter per second. This is remarkably slow. Electric signals travel so much faster than this because, when we fire up a circuit, the electrons all along the wires start moving almost simultaneously.

The electron velocity above is negative because electrons carry a negative charge, so they move in the opposite direction to the current. This is a simple matter of convention: we are stuck with one introduced by Benjamin Franklin in the 1700s.

[4] Actually, the Ampere has a more fundamental definition and defines the Coulomb, but that's beyond this course.

6.3 Voltage

Consider a ball of mass m held a distance h above the ground. If it is released, it will fall under gravity gaining kinetic energy. When it reaches the ground, gravity has exerted a force mg though a distance h, so it has done work $W = mgh$ on the ball, which has been converted to kinetic energy. If the ball were not freely falling but sinking slowly through a viscous fluid, gravity would still have done the same amount of work, but the energy would be dissipated as heat in the fluid. We therefore say that when the ball is at a height h it has gravitational potential energy mgh, which is released when the ball moves down to the ground. More generally, we might say there is a gravitational potential difference of $\phi_{h\to 0} = gh$ between the height h and the ground, which is an energy per unit-mass, meaning that if a mass M moves from h to ground, it releases gravitational potential energy $M\phi_{h\to 0}$.

Exactly the same considerations apply to a charge moving in an electric field. If we have an electric field E pointing along a wire, and a charge in the wire moves a distance $\mathrm{d}l$ along the field direction, as shown in fig. 53, the electric field does work on the charge $\mathrm{d}W = qE\mathrm{d}l$. If the charge moves a long distance down the wire, from a to b, the total work done by the electric field is

$$W = q\int_a^b E\mathrm{d}l. \tag{160}$$

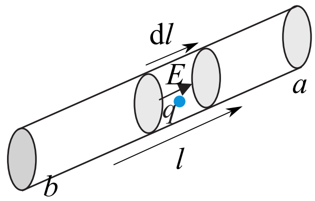

Figure 53: If I move a charge q a distance $\mathrm{d}l$ against the electric field $E(l)$ in a wire, I must do work $qE(l)\mathrm{d}l$.

This work done by the electric field must be converted into some other form of energy. For current in a wire, it is released as heat but in other circumstances the energy may be turned into light or motion. It is useful to introduce a new concept, the electric potential difference between a and b, defined as

$$V_{a\to b} = \int_a^b E\mathrm{d}l, \tag{161}$$

so that, if a charge Q flows between from $a \to b$, the total electric energy released is

$$W = QV_{a\to b}. \tag{162}$$

This electric potential difference is measured in Volts, where one Volt is one Joule of energy per Coulomb of charge. When one Coulomb of charge moves through one a potential difference of one Volt, one Joule of energy is released. Note that, in both the electric and the gravitational case, we only speak of the potential difference between two points, not the absolute potential at a point.

6.4 EMF and Batteries

A battery is a device with two terminals, that generates a sustained potential difference, conventionally labeled ξ, between them. If we connect a wire between the two terminals, there will be a voltage drop along the wire (and hence an electric field within the wire) so a current, I, will flow through it. Consequently a current I is departing from the high voltage terminal and a current I is arriving at the low voltage terminal, so a current I also appears to be flowing through the battery. When a Coulomb of charge flows through the wire it releases ξ Joules of energy as heat, so when it then flows through the battery, from low voltage terminal to high voltage terminal, it must acquire ξ Joules of electrical energy to release on its next trip through the wire. We say that the battery has an electro-motive-force (or *emf*) ξ, which is also measured in Volts (i.e. Joules per

Coulomb) but in this case, meaning Joules of electrical energy gained by a Coulomb of charge when it flows through the battery.

In a typical battery, energy is stored as chemical energy, and only converted to electrical energy when a current flows. Other devices generate *emfs* by converting other forms of energy into electrical energy: generators convert mechanical energy to electrical energy, solar panels convert light energy and microphones convert sound energy.

While the picture of current flowing through a battery in this way is a good way to think about circuits, it isn't necessarily how sources of *emf* actually work. Some sources (notably generators) do work in exactly this way, but, in general, we have no guarantee that the charges that flow into the low voltage terminal are those that subsequently flow out of the high voltage terminal. Indeed in most batteries charges do not move from low to high voltage within the battery (except when you charge it) rather the high voltage terminal has a large stock of charge at high voltage, and once it has all flowed to the low voltage terminal the battery is used up.

6.5 Resistance

If we connect a wire between two terminals of a battery then current flows through it, but how much current? The answer depends both on the material and the geometry of the wire, but for almost all wires we find that the current is proportional to the applied potential difference

$$V = IR, \tag{163}$$

where the constant of proportionality is called the wire's resistance. This result is known as Ohm's law and the units of resistance are Ohms ($\Omega = \mathrm{VA}^{-1}$), The number of Ohms of a component tells you how many Volts you must apply to drive a current of one Amp though it.

We can understand Ohms law better by thinking microscopically. Imagine a cylindrical wire of length l connected between the terminals of a battery with an *emf* of V. The wire will then contain an electric field which will push the charge carriers through the wire with an average velocity drift proportional E, giving $\langle v \rangle = \beta E$. We recall there is an average drift velocity because, although the electric field accelerates the charges, they then collide with the atoms in the wire and slow down again, so β depends on the atoms in question, and hence the material of the wire. The drift velocity will be constant along the wire's length, so the electric field must also be constant. This means that, looking at eqn 161, we can see the electric field must have size $E = V/l$. Putting this into eqn 157, the total current in the wire is $I = nqA\beta V/l$, which is Ohm's law and, as a bonus, we have derived the wire's resistance[5], $R = l/(nqA\beta)$.

Distilling this argument to its bare essentials, if we double the voltage over the wire, we double the electric field in the wire, so we double the drift velocity of the charge carriers, so we double the current. Current is thus proportional to voltage. We also see that the resistance of a wire depends on its geometry: if the wire is twice as long its resistance doubles, if it doubles in cross-sectional area its resistance halves.

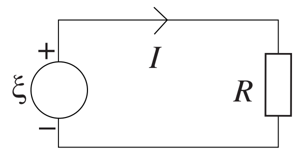

Figure 54: A battery with *emf* ξ drives a current I through a resistor.

6.5.1 Resistors

Although all wires have some resistance, in practice the resistance of wires is very low. When we build circuits we normally limit the current by using components known as resistors, which have resistances that are orders of magnitude higher than a typical wire. We can make these, for

[5]We typically merge the three material dependent quantities into a single material parameter, the conductivity of the material, $\sigma = nq\beta$ or, if you prefer, its resistivity $\rho = 1/\sigma$.

example, by using a material such as graphite that has a much higher resistance than a metal. We can now analyze the simplest possible circuit: a resistor of magnitude R connected across a perfect battery producing an *emf* of ξ. This is shown as a circuit diagram in fig. 54. The question we ask is: what is the current through the resistor? The answer is simple: the potential difference across the resistor is equal to the *emf* of the battery, and the current through the resistor is given by Ohm's law:

$$I = \frac{\xi}{R}. \tag{164}$$

6.5.2 Resistors in series

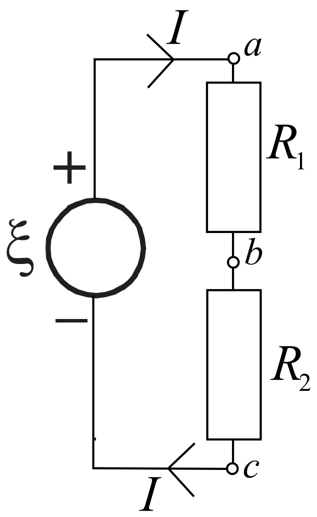

Figure 55: A battery drives a current I through two resistors in series.

Suppose we now have two resistors, of resistance R_1 and R_2, connected in series across the battery, as shown in fig. 55. The same current, I must flow through both resistors (otherwise we have more charge flowing out of the first resistor than into the second, which would result in it building up between them) and the potential difference across the each resistor is given by Ohm's law

$$V_1 = V_{a\to b} = IR_1 \qquad V_2 = V_{b\to c} = IR_2. \tag{165}$$

A Coulomb of charge liberates potential energy V_1 when it passes through resistor 1, then V_2 when it passes through resistor two. All this energy must have been given to the Coulomb of charge by the *emf* of the battery, so we have

$$\xi = V_1 + V_2 = IR_1 + IR_2 = I(R_1 + R_2). \tag{166}$$

We see that the two resistors offer a combined resistance to the current of $R = R_1 + R_2$. This easily generalizes to n resistors in series, giving

$$R = R_1 + R_2 + R_3 + ...R_n. \tag{167}$$

This is consistent with our observation that a wire's resistance is proportional to its length. This result also confirms our intuition that we can neglect the resistance of the wires in a circuit like fig. 54. The wire has a tiny resistance R_w and is in series with a resistor R, so the total resistance $R_w + R$ is scarcely bigger than R. Correspondingly the voltage drop across the wire $V_w = IR_w$ is also tiny, so the big resistor feels almost the the whole *emf* of the battery. In practice we completely neglect the wire's resistance in this type of circuit.

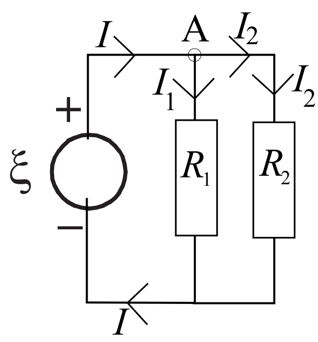

Figure 56: A battery drives a current I through two resistors in parallel.

6.5.3 Resistors in parallel

Next we consider two resistors, R_1 and R_2, connected in parallel across the battery, as shown in fig. 56. Now there is no need for the two to carry the same current, but in both cases any charge that flows through the resistor must lose the entire amount of energy it was given by the battery so the potential difference across each resistor is the whole *emf*, ξ. The current in both resistors is thus

$$I_1 = \frac{\xi}{R_1}, \quad I_2 = \frac{\xi}{R_2}, \tag{168}$$

and the total current provided by the battery is

$$I = I_1 + I_2 = V\left(\frac{1}{R_1} + \frac{1}{R_2}\right). \tag{169}$$

From the battery's perspective, the resistors are equivalent to a single resistor with resistance

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}. \tag{170}$$

Again we can easily generalize this to a n resistors in parallel,

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + ... \frac{1}{R_n}. \tag{171}$$

This is consistent with our observation that a wire's resistance is inversely proportional to its area.

6.5.4 Realistic Batteries

Ideal batteries produce the same *emf* regardless of how much current they are providing. Real batteries are not so impressive, since they do offer some resistance to current passing though. A more realistic model is that they behave like an ideal battery in series with a small internal resistance R_{int}, as shown in fig. 57. When the battery is called on to produce current I, the voltage drop across the internal resistance is $V = IR_{int}$, so the voltage seen by the rest of the circuit is $\xi - R_{int}$. For most purposes we neglect this complication.

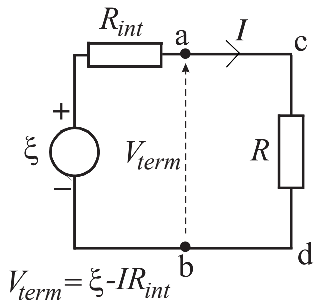

Figure 57: A realistic battery can be modeled as an ideal battery in series with an internal resistance R_{int}.

6.6 Power in electric circuits

If a total charge Q flows through a potential difference V then the total electrical energy released is $W = QV$. The rate at which electrical energy is being released is thus

$$P = \frac{\mathrm{d}W}{\mathrm{d}t} = V\frac{\mathrm{d}Q}{\mathrm{d}t} = IV, \tag{172}$$

This applies in any situation where a current flows through a voltage so, for example, if a battery produces an *emf* of ξ and a current I, it is providing electrical energy at a rate ξI. If the voltage drop is accross a resistor (and the current through the resistor), then the power we are calculating is the rate at which electrical energy is dissipated as heat in the resistor. However, in this case, voltage and current are related by Ohm's law, so we can write the general result as

$$P = IV = \frac{V^2}{R} = I^2 R. \tag{173}$$

6.7 Kirchhoff's Laws

Kirchhoff's laws encode two principles that we have already implicitly used in our treatment of resistors in series and parallel. However, writing them down formally allows us to tackle more complicated networks of resistors and batteries.

6.7.1 Kirchhoff's Current Law

Kirchhoff's current law states that current is conserved at a junction, that is, the total current flowing into a junction of a circuit is equal to the total current flowing out. If this were not true there would be more charge arriving than departing, so charge would be building up at the junction, which is not observed.

An example of a junction in a circuit is shown in fig. 58. Kirchhoff's law tells us that $I_1 = I_2 + I_3 + I_4$. In general we write this as

$$\sum_i I_i = 0, \tag{174}$$

where currents are assigned positive values if they are flowing into the junction and negative values if they are flowing out.

We already used this principle when we discussed resistors in parallel: when we stated that the total current provided by the battery is equal to the sum of the currents flowing through the two resistors, we were implicitly applying Kirchhoff's current law to the junction marked A in fig. 56.

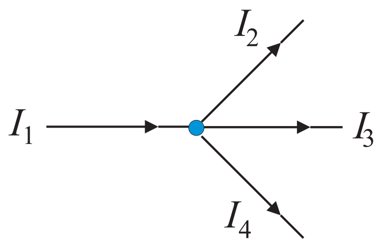

Figure 58: A battery drives a current I through two resistors in parallel.

6.7.2 Kirchhoff's Voltage Law

Kirchhoff's voltage law is a somewhat disguised version of the principal of conservation of energy. If we imagine a charge moving around any loop in our circuit, the amount of electrical potential energy it picks up going through any sources of *emf* should equal the amount it loses passing through any resistors. If this were not true, then we would have a situation where we were producing more heat energy in the resistors than we were extracting from our sources of *emf*, violating conservation of energy. When a charge q passes through a component over which there is a voltage change V_i, it gains or loses energy qV_i. The formal statement of Kirchhoff's voltage law is thus that around any loop

$$\sum_i V_i = 0, \tag{175}$$

so that the total energy picked up by a charge going around the loop, $q \sum_i V_i$, is zero. The V_i are positive if a charge moving through the component gains energy, and negative if the charge moving through the component loses energy, where the movement is in the direction of the loop we are imagining taking the charge in.

We already used this principle when we discussed resistors in parallel: when we stated that the voltage across each resistor was equal to the battery's *emf* we were implicitly applying Kirchhoff's voltage law to the loop in the circuit containing the battery and R_1, and the loop in the circuit containing the battery and R_2. Similarly with resistors in series, when we stated that the sum of the voltage drop across the two resistors should equal the battery's *emf*, we were implicitly applying Kirchhoff's voltage law to the single loop in the circuit.

6.7.3 Using Kirchhoff's laws: signs

When applying Kirchhoff's laws it is all too easy to get confused about the signs of the various currents and voltages. The current law is easy. Draw a diagram of the circuit and label each branch with the current flowing through it, I_i, and an arrow indicating the current's direction. It doesn't matter if the arrow points in the direction the current actually flows — if it doesn't you will just find your current is negative — but it is imperative to choose a direction for each current and stick to it. At each junction, the signs in Kirchhoff's current law are dictated by whether the currents as marked flow into or out of the junction. The voltage law is harder because there are two sets of signs to keep track of. The voltage across a component labels how much energy a charge gains or loses when it flows through the component in a particular direction. This can be positive if the component extracts electrical energy, like a resistor, of positive if it supplies energy like a battery. Secondly, when we apply the voltage law to a loop, we must also keep track of whether we pass each component with or against the assigned direction of the voltage. A simple way through is to mark next to each component a voltage V_i which is the potential lost by a charge flowing through the component in the direction of the indicated current. For a resistor, this is positive, related to the current by Ohm's law $V_i = I_i R$, while for a battery, if the current flows through from the negative terminal to the positive terminal it will be negative, $V_i = -\xi$. In the voltage law we then imagine taking a charge around a loop, so it loses energy qV_i if it passes the component in the direction of the indicated current and gains energy qV_i if it passes against the direction of the indicated current. The voltage law then says that the total energy change around the loop must be zero.

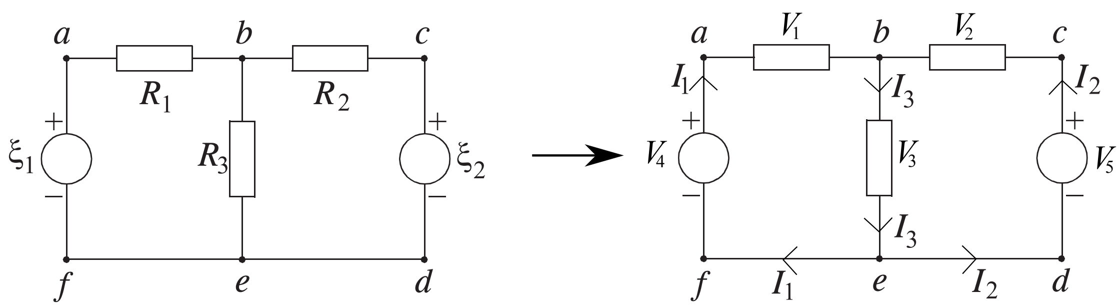

Figure 59: Left: Diagram of a circuit containing two batteries. Right: We label a current in each branch and a voltage for each component.

As an example of using Kirchhoff's laws, consider the circuit shown if fig. 59. On the left we see the circuit problem as posed: a network of three resistors and two batteries. We wish to find the voltage over and current through R_3. On the right, we have the same circuit, but now each of the three branches has a current assigned, with the directions indicated by an arrow, and each component has a voltage change assigned. Applying the current law to the junction at b (or equivalently e) gives

$$I_1 + I_2 - I_3 = 0 \tag{176}$$

Applying the voltage law to the left loop going clockwise ($a \to b \to e \to f \to a$) we pass three components, each in the direction of the indicated current, so we have

$$-V_1 - V_3 - V_4 = 0 \quad \Longrightarrow \quad \xi_1 - I_1 R_1 - I_3 R_3 = 0. \tag{177}$$

In the latter equality, we have applied Ohm's law to the two resistors, and used the fact that we know the battery's *emf*, which appears with a positive sign because the V_i indicate the energy lost by a Coulomb of charge flowing through a component, but a Coulomb of charge passing through the battery gains energy so $V_4 = -\xi_1$. Similarly, applying the voltage law to the right loop ($c \to b \to e \to d \to c$) we have

$$-V_2 - V_3 - V_5 = 0 \quad \Longrightarrow \quad \xi_2 - I_2 R_2 - I_3 R_3 = 0. \tag{178}$$

These three equations can be solved for the three unknown currents, I_1, I_2 and I_3. Using the latter two to eliminate I_1 and I_2 from the first gives a linear equation for I_3,

$$\frac{\xi_1 - I_3 R_3}{R_1} + \frac{\xi_2 - I_3 R_3}{R_2} - I_3 = 0 \quad \Longrightarrow \quad I_3 = \frac{\xi_1/R_1 + \xi_2/R_2}{1 + R_3/R_2 + R_3/R_1}. \tag{179}$$

The voltage drop over R_3 is correspondingly

$$V = I_3 R_3 = \frac{\xi_1/R_1 + \xi_2/R_2}{1/R_3 + 1/R_2 + 1/R_1}. \tag{180}$$

6.7.4 Using Kirchhoff's laws: current loops

The fact that current must be conserved, means that charges must continuously flow in loops. If we assign to each loop in a circuit a current flowing around it, we generate a set of currents that automatically obey Kirchhoff's current law, leaving us with less work to do. To treat the previous problem in this manner, we assign currents I_1 and I_2 to the two loops, as shown in fig. 60, so that the current flowing through R_3, which is in both loops, is $I_1 + I_2$. Notice that Kirchhoff's current law is automatically satisfied at b and e. All that remains is to apply the voltage law to the two loops, giving

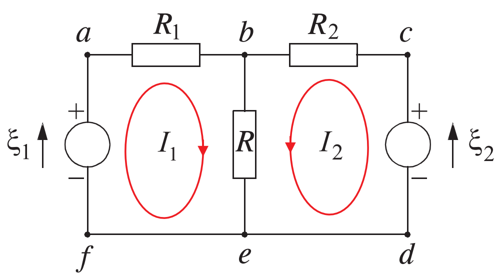

Figure 60: Analyzing the circuit in fig. 59 via current loops.

$$\xi_1 - I_1 R_1 - (I_1 + I_2) R_3 = 0 \quad \text{and} \quad \xi_2 - I_2 R_2 - (I_1 + I_2) R_3 = 0. \tag{181}$$

We wish to find the current through R_3, which is $I_1 + I_2$, so we simply add the first equation divided by R_1 to the second divided by R_2 to get

$$\xi_1/R_1 + \xi_2/R_2 - (I_1+I_2)(1+R_3/R_1+R_3/R_2) = 0 \Longrightarrow I_1+I_2 = \frac{\xi_1/R_1 + \xi_2/R_2}{1 + R_3/R_2 + R_3/R_1}, \tag{182}$$

as before.

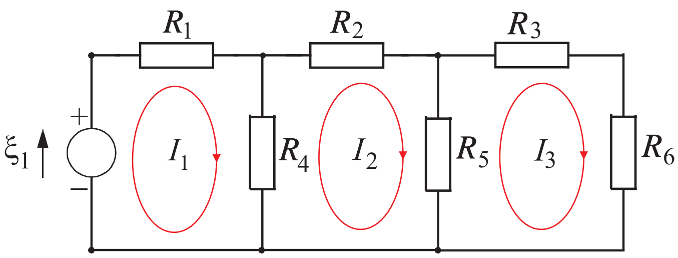

Figure 61: A more complicated circuit analyzed via current loops

This technique saves more effort in circuits with more loops such as that shown in fig. 61. Suppose here we wish to find the current though R_6. The circuit has six branches, so our original approach would require us to introduce six currents, but using loops we can use only three currents, and immediately satisfy Kirchhoff's current law. Applying the voltage law to the three loops then gives three linear equations in our three unknown currents,

$$\begin{aligned} \xi_1 - I_1 R_1 - (I_1 - I_2) R_4 &= 0 \\ -I_2 R_2 - (I_2 - I_3) R_5 - (I_2 - I_1) R_4 &= 0 \\ -I_3 R_3 - I_3 R_6 - (I_3 - I_2) R_5 &= 0 \end{aligned}$$

which, in principle, we can easily solve to find I_3.

6.7.5 Using Kirchhoff's laws: Assigning Voltages

We can do an analogous trick to automatically implement Kirchhoff's voltage law. We know from Ohm's law that there is negligible potential difference across each wire in the circuit, so we label each wire in the circuit with a voltage relative to some arbitrary reference point in the circuit, normally taken for convenience as the negative terminal of the battery. The potential difference across each component is then given by the difference in voltages of the wires on either side of it. If we do this then Kirchhoff's voltage law is automatically satisfied since, whatever loop we move a charge around, it gets back to the same voltage it started at. Consider, for example, our resistors in series diagram, fig. 55. If we assign a voltage V_a to the wire marked a, V_b to wire b and $V_c = 0$ to wire c (which connects to the negative terminal) then, working clockwise around the loop, the voltage drop over R_1 is $V_1 = V_a - V_b$, the voltage drop over R_2 is $V_2 = V_b - V_c$ and the voltage drop over the battery is $V_3 = V_c - V_a$. Kirchhoff's voltage law is thus automatically satisfied:

$$-V_1 - V_2 - V_3 = -(V_a - V_b) - (V_b - V_c) - (V_c - V_a) = 0. \tag{183}$$

Recalling $V_c = 0$, we know the battery's *emf* is ξ, so $V_a = \xi$. Applying Ohm's law to the two resistors then gives $\xi - V_b = IR_1$ and $V_b = IR_2$, which we can solve to find $IR_1 + IR_2 = \xi$, as before, and $V_b = \xi R_2/(R_1 + R_2)$.

6.8 Capacitors

Consider two parallel conducting plates, one holding a charge $+Q$ and the other a charge $-Q$, as sketched in fig. 62. If a small charge $\mathrm{d}Q$ is released from the $+Q$ plate, it will be repelled by the $+Q$ plate and attracted by the $-Q$ plate so, if there is a vacuum between them, it will fly towards the $-Q$ plate gaining kinetic energy. When it collides with the $-Q$ plate it will stick, turning all its kinetic energy into heat. Where did this heat energy come from? The separated $+Q$ and $-Q$ charges are storing electrical potential energy. When the charge $\mathrm{d}Q$ flows between them, both plates become slightly less charged, so they are storing less electrical potential energy, and the balance has been released as heat. We define the voltage between the two plates as the amount of electrical potential energy released when one Coulomb of charge moves between them, so, in the previous situation, the amount of electrical potential energy converted to heat was $V\mathrm{d}Q$.

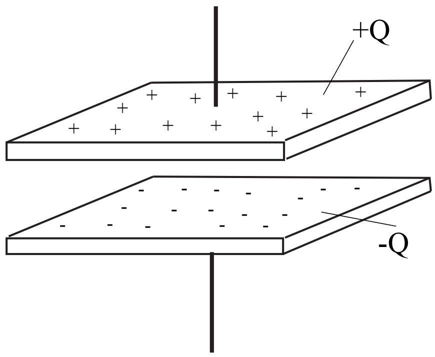

Figure 62: A capacitor formed from two parallel plates with charge $\pm Q$.

Two such charged plates are an example of a circuit component called a capacitor. Typically between the plates we put a good insulating material, so no charge can pass between them as discussed above. However, if we then connect a resistor R between the two plates, as shown in fig. 63 it provides path for charge to flow between them and a current $I = V/R$ will flow through the resistor until the plates are completely uncharged. The charged plates thus behaves like a battery with *emf* of V, except that, rather than storing energy chemically, they store energy directly as electrical potential energy. However, unlike a battery which has fixed voltage between its terminals, the voltage between the two plates depends on the charge they are

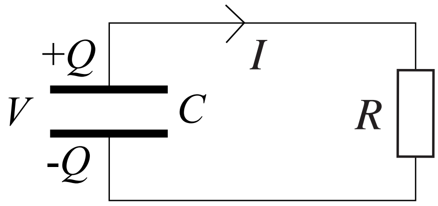

Figure 63: When a resistor connected across the plates of a charged capacitor, it drives a current through the resistor.

holding. Returning briefly to the case where a charge $\mathrm{d}Q$ is allowed to pass between the plates, if we double the charge on the plates, we double the electric field between them, so the charge is pushed twice as hard, and it arrives at the negative plate with twice as much energy, meaning the voltage between the plates has also doubled. In general the voltage between the plates is proportional to the charge on them, so we write

$$Q = CV, \tag{184}$$

where the constant of proportionality C, is called the plates' capacitance, and is measured in Farads, where one Farad is one Coulomb per Volt.

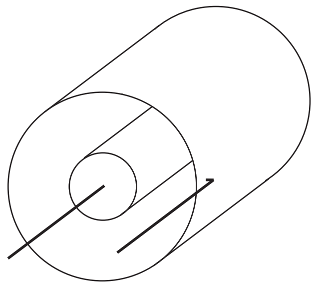

Figure 64: Concentric cylindrical conductors as a capacitor.

A capacitor can be formed by any pair of separated conductors that carry equal and opposite charges: for example a pair of separated conducting spheres, or concentric conducting cylinders (fig. 64) or concentric conducting spheres. However, all capacitors are described by eqn 184, with the value of the capacitance depending on the geometry of the capacitor. A capacitor is a circuit component, like a battery or a resistor, but with the voltage across it described by eqn 184 rather than Ohm's law (resistors) or simply being a constant (batteries). We say that a capacitor carrying charges $\pm Q$ on its two conductors is "charged" or "carries a charge Q", but it is important to realize that, although the two conductors are charged, overall, the capacitor is neutral.

6.8.1 Energy stored in a Capacitor

To charge a capacitor we must move charge from the negatively charged plate to the positively charged plate. If the capacitor is holding a charge q then to move a further charge $\mathrm{d}q$ we must do work $\mathrm{d}W = V\mathrm{d}q$ which is stored as electrical potential energy. However, since it is a capacitor, this voltage depends on the charge the capacitor is holding via eqn 184, giving $\mathrm{d}W = (q/C)\mathrm{d}q$. To find the total work needed to charge a capacitor for 0 to Q we must sum these contributions using an integral, giving

$$W = \int_0^Q (q/C)\mathrm{d}q = \frac{Q^2}{2C} = \frac{1}{2}QV, \tag{185}$$

where Q is the final charge on the capacitor and V its final voltage. We can verify that all this energy is stored in the capacitor by considering discharging it through a resistor. Each time a charge $\mathrm{d}q$ flows through from the positive plate to the negative one through the resistor, it converts $\mathrm{d}W = V\mathrm{d}q$ of electric energy into heat. However, again, this voltage depends on the charge the capacitor is holding via eqn 184, giving $\mathrm{d}W = (q/C)\mathrm{d}q$. If we sum these contributions to find the total energy dissipated in the resistor during the discharge, we have exactly the same integral as eqn 185. Thus all the energy we thought we stored in the capacitor is released as heat in the resistor during discharge, confirming it was stored in the capacitor.

6.8.2 Exponential decay in an RC circuit

We are also interested in how long it takes to charge and discharge a capacitor. If we have an RC circuit, as shown in fig. 63, in which the capacitor has a charge Q, then the voltage across the resistor is $V = Q/C$, and the current through the resistor discharging the capacitor is $I = V/R = Q/(CR)$. However, current is rate of flow of charge so, if a current I flows for a time $\mathrm{d}t$ then a charge $\mathrm{d}Q = I\mathrm{d}t$ has flowed from the positive plate to the negative plate, reducing the charge on the capacitor by $\mathrm{d}Q$. We thus have

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -I = -\frac{V}{R} = -\frac{Q}{CR}, \tag{186}$$

where the negative sign arises because the capacitor is discharging, so the current acts to reduce Q. Integrating this gives a simple exponential decay of the charge on the capacitor as a function of time:

$$Q(t) = Q(0)\mathrm{e}^{-\frac{t}{RC}}. \tag{187}$$

The charge on the capacitor, as sketched in fig. 65, decays with characteristic decay time $\tau = RC$. As we might expect, a big resistor means little current flows, so the capacitor discharges very slowly.

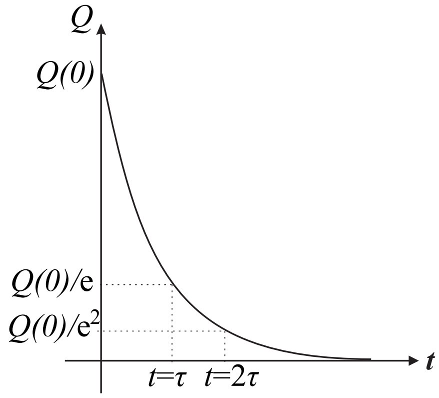

Figure 65: Decay of the charge on a capacitor as it discharges through a resistor.

The energy $W(t)$ stored in a discharging capacitor is thus

$$W(t) = \frac{1}{2}\frac{Q(t)^2}{C} = \frac{1}{2}\frac{Q(0)^2}{C}e^{-\frac{2t}{RC}} = W(0)e^{-\frac{2t}{RC}}, \tag{188}$$

so the energy decays twice as quickly as the charge, with decay time $RC/2 = \tau/2$.

6.8.3 Charging a capacitor with a battery

Finally, we are also interested in the mechanics of charging a capacitor with a battery. Consider a circuit such as that shown in fig. 66 containing a battery a capacitor and a resistor. If the capacitor starts uncharged, we expect that current will flow in the circuit, charging the capacitor.

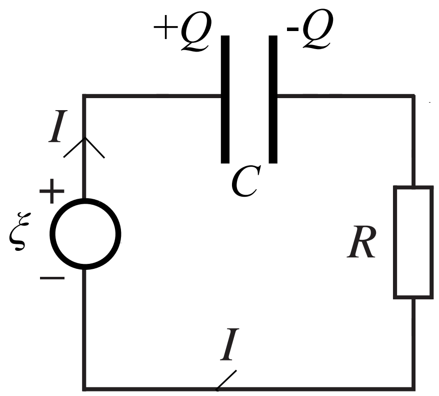

Figure 66: Charging a capacitor with a battery.

To analyze this circuit, we first apply Kirchhoff's voltage law, which tells us that the voltage across the resistor and capacitor must sum to the *emf* of the battery

$$\xi = IR + \frac{Q}{C}. \tag{189}$$

Secondly, we know that the capacitor remains overall uncharged, so the current flowing on to its positive plate must equal the current flowing off its negative plate, so we have $I = \frac{\mathrm{d}Q}{\mathrm{d}t}$. Taking a time derivative of the above equation thus gives a simple exponential decay law for the current in the circuit:

$$0 = R\frac{\mathrm{d}I}{\mathrm{d}t} + \frac{I}{C} \implies \quad I = I(0)\mathrm{e}^{-\frac{t}{RC}}. \tag{190}$$

We see that the current in the circuit is high at early times, and decays thereafter. This is because at first the capacitor is uncharged, so there is no voltage across it, and the whole *emf* of the battery is lost across the resistor, driving a large current through it. As charging proceeds, the voltage over the capacitor rises, so the voltage across the resistor falls, and hence the current also falls. At $t = 0$ the capacitor is uncharged ($Q = 0$) so substituting this current (eqn 190) into eqn 189 gives $I(0) = \xi/R$. The same substitution gives the charge at t as

$$Q(t) = C\xi\left(1 - \mathrm{e}^{-\frac{t}{RC}}\right). \tag{191}$$

We see that the capacitor charges with the same characteristic time $\tau = RC$ as it takes to decay.

Finally, it is interesting to think about energy in this charging process. There is something of a mystery here, since a charge $Q = C\xi$ has flowed off the positive terminal of the battery and onto the negative terminal, releasing energy $Q\xi$, but the energy stored on the capacitor is only $\frac{1}{2}Q\xi$. The other half of the energy is dissipated in the resistor. We can see this directly, since we know the power dissipated in a resistor is I^2R, so the total energy dissipated in the resistor in the charging process is

$$W = \int_0^\infty I(t)^2 R\mathrm{d}t = \frac{\xi^2}{R}\int_0^{\mathrm{inf}} \mathrm{e}^{-\frac{2t}{RC}}\mathrm{d}t = \frac{1}{2}C\xi^2 = \frac{1}{2}Q\xi. \tag{192}$$

If our resistor is very small, the capacitor will charge very quickly, but half the energy lost by the battery is always dissipated in the resistor.

6.9 Inductors

This is not a course on magnetic fields, which are covered in greater detail in Easter term. However, here we need to understand a final circuit component called an inductor, which works via magnetic fields, so we first recap their properties.

6.9.1 Magnetic Fields

The magnetic field, much like the electric field, is described by a vector at each point in space, which we label $\mathbf{B}$. However, the forces produced by magnetic fields are altogether more complicated than those produced by the electric field: unlike in the electric case, the universe doesn't contain any objects that carry a net "magnetic-charge" $q_\mathbf{B}$ which experience a force $\mathbf{F} = q_\mathbf{B}\mathbf{B}$. If we put a regular electrically charged particle into a magnetic field, it does experience a force, but only if it's moving. More precisely, if a charge q has velocity $\mathbf{v}$ in a magnetic field, it experiences a magnetic force which is orthogonal to both the magnetic field and the velocity of the particle given by:

$$\mathbf{F} = q(\mathbf{v} \times \mathbf{B}). \tag{193}$$

The unit of the magnetic field is the Tesla (T), where one 1T=1Ns/(Cm): a charge of one Coulomb moving at one meter per second perpendicular to a field of 1 Tesla feels a force of one Newton.

We can thus observe magnetic fields by looking at the forces they exert on moving charges, or currents in wires. However, nature also provides us with another tool. Some objects, such as bar magnets, carry magnetic dipoles, meaning that they appear to carry a magnetic charge $+q_\mathbf{B}$ or north-pole at one of $-q_\mathbf{B}$ (a south-pole) at the other end. These magnetic charges appear to produce magnetic fields that radiate out of the positive charge (north-pole) and into the negative charge (south pole), so the whole dipole produces the well-known field of a bar magnet, shown in fig. 67. Furthermore, just like electric charges, the north-poles of two magnets repel, as do the south-poles, while north and south poles attract each other. However, unlike in the electric case, we cannot isolate

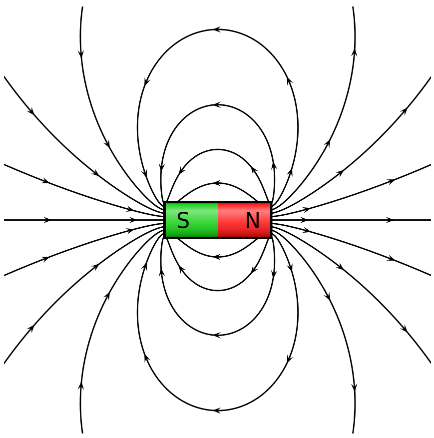

Figure 67: Magnetic Field of a bar magnet. The field lines radiate out of the north pole and into the south pole. Image by wikipedia user by Geek3, licensed as GFDL.

a net charge: if we try, for example by cutting a bar magnet in half, we find it divides into two shorter dipoles rather than separate north and south poles. Even if we get down to fundamental particles, we find that they still cary an intrinsic magnetic dipole. The magnetic charges are a useful way of thinking, but, as far as we know, they don't actually exist.

If we put a dipole in a homogeneous magnetic field, the field does not exert a net force on the dipole (since the forces on the two poles are opposite and cancel out) but it does produce a turning moment that tries to align the dipole with the magnetic field. This is how a magnetic compass works: the needle in it is a magnetic dipole, which aligns with the magnetic field of the earth. Iron filings behave similarly since they acquire a magnetic dipole when placed in a magnetic field. This allows us to visualize magnetic fields by sprinkling iron filings or little bar magnets around another source of magnetic field and observing which way they align.

6.9.2 Magnetic Field from a wire

If we have a long straight wire and we run a current I through it, it produces a magnetic field in its surroundings. We can observe the form of this field with iron filings, and we discover that, as sketched in fig. 68, the magnetic field forms as rings around the wire, with a direction given by a right-hand-rule (point your right thumb along the current, and your fingers wrap in the direction of the field). The strength of the field is proportional to the strength of the current.

Historical note: this was first reported by Oersted in 1820, who conducted a famous experiment placing a compass over a conducting wire, and showing that the needle of the compass aligned perpendicular with the wire. This was the first time a force had been seen to act between two objects (the needle and the wire) that wasn't along their separation. It then took Ampere less than a week to show that the magnetic field lines lie in concentric circles around the wire.

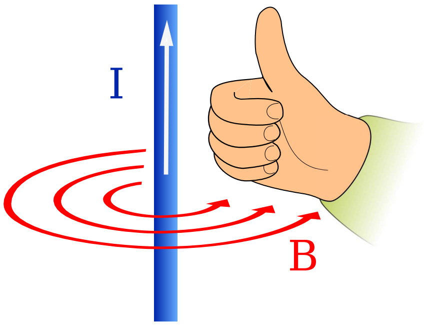

Figure 68: Magnetic Field of a current carrying wire. The field forms concentric circles around the wire, with the direction given by the right-hand-rule. Image by wikipedia user by Jfmelero, licensed as GFDL.

Figure 69: Magnetic Field of a current loop traced with iron filings.

The magnetic field of a wire bent into a loop is shown traced in fig. 69. We can qualitatively understand the shape of the field by breaking the wire into small arc lengths that are effectively straight, and imagining that each produce concentric rings of field. Each of these contributions produces a field through the loop in the same direction, so there is a large field through the loop. Outside the loop the contributions point in different directions, so the field is weaker. The strength of the field is proportional to the strength of the current.

6.9.3 Electromagnetic Induction

Thus far, electricity and magnetism have been separate phenomena, albeit both acting on the same charges. However, a simple experiment reveals that, at least when the fields vary in time, they are intimately linked. If , as sketched in fig. 70, we have a wire loop, and we move a bar magnet towards it, we discover it causes a current to flow in the loop. However, the current only flows while the magnet is being moved: if we hold the magnet still the current stops. More precisely, the current is proportional to the strength of the magnet's field and the velocity of the magnet through the loop. If we reverse the velocity of the magnet, we reverse the direction of the current. We can also establish that the current is inversely proportional to the wire's resistance, so, invoking Ohm's law, we can say that the movement of the magnet through the loop is inducing an *emf* in the loop proportional to the velocity and strength of the magnet.

More careful experiments reveal that the *emf* around any loop induced by changing magnetic fields is given by

$$\xi = -\frac{\mathrm{d}\phi_B}{\mathrm{d}t}, \tag{194}$$

where ϕ_B, called the magnetic flux, measures how much magnetic field passes through the loop. If we think of the magnetic field as a velocity field for a fluid, we can then calculate ϕ_B as the rate at which the fluid flows through the loop. For example, in the very simple case where we have a loop of area A and a constant magnetic field B flowing through it perpendicular, we simply have $\phi_B = BA$. The units of ϕ_B are thus Tm^2. In our initial experiment, when the magnet is far from the loop, there are very few magnetic field lines flowing through the loop, so ϕ_B is small. When the magnet is close to the loop, many field lines pass through the loop, so ϕ_B is big. The *emf* in the loop arrises when ϕ_B changes, which occurs when the magnet approaches the loop.

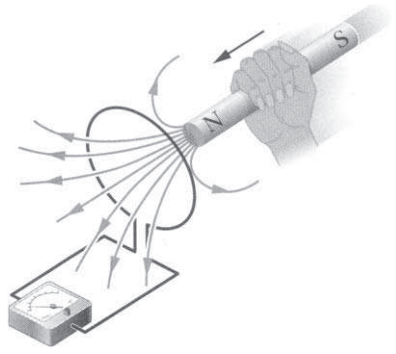

Figure 70: When a magnet moves towards a coil, it induces a current to flow in the coil

The minus sign in eqn 194 encodes an important idea. As we saw in fig. 69, when current flows in a loop it produces its own magnetic field, which itself flows through the loop. The sign in eqn 194 indicates that the direction of the current induced in the coil will be such that its magnetic field opposes the change in ϕ_B driving the current. The result is known as Lenz's law.

6.9.4 Self Inductance

When a current flows around a loop it produces a magnetic field that itself flows through the loop. Thus the magnetic field of the loop itself has a magnetic flux ϕ_B through the loop. Although calculating this flux is difficult, we know that if we double the current in the loop, we will double the strength of the field, and thus have twice the flux, so the flux through the loop is proportional to the current in the loop. We therefore write

$$\phi_B = LI, \tag{195}$$

where the constant L is known as the loop's self-inductance, and is measured in Henrys, where $1\mathrm{H} = 1\mathrm{Tm}^2/A$. This linear relationship holds for all geometries of loops, but the constant of proportionality, L, is different for loops of different shapes.

If the current through a loop changes, then the flux through the loop from its own magnetic field, $\phi_B = LI$, changes. By the law of electromagnetic induction, eqn 194, this means that an *emf* is generated in the loop

$$\xi = -\frac{\mathrm{d}\phi_B}{\mathrm{d}t} = -L\frac{\mathrm{d}I}{\mathrm{d}t}. \tag{196}$$

This *emf* is proportional to the rate of change of the current in the loop. However, the *emf* is also within the loop, so it drives current in the loop. Lenz's law allows us to understand how this influence works: the *emf* acts to oppose the change in flux through the loop that occurs when the current changes, so it acts against the change in current. If we try to increase the current in the loop, an *emf* appears around the loop trying to drive a current in the opposite direction to slow the build-up of current, while, if we decrease the current in the loop, an *emf* appears around the loop that acts to push more current around the loop, slowing the decrease in the current.

6.9.5 Inductors as circuit elements

Figure 71: A wire coil forms an air-core inductor.

We use the effect of self-inductance to make a circuit component called an inductor. As with our previous circuit elements, this is a two terminal device, but between the terminals is a wire wound into coil, as shown in fig. 71. The wire coil has negligible resistance, but if a current flows through it, each loop in the coil makes a magnetic field like that in fig. 69, which passes through both itself and many other loops of the coil, so the inductor has a very high self inductance L. When the current through the coil changes, an *emf* given by given by eqn 196 appears between the two ends of the coil, i.e. between the terminals of our inductor. An inductor is thus a two terminal circuit element, where the voltage across it is determined by the rate of change of current passing through it. This contrasts with a resistor, for which the voltage is determined by the current flowing through it, a capacitor, for which the voltage is determined by the charge, or a battery, for which it is constant.

6.9.6 Inductor and a battery

The simplest circuit we can make involving an inductor is a loop containing an inductor L and a battery with an *emf* of ξ, as shown in fig. 72. The battery tries to drive a current through the inductor, but this is resisted by the back *emf* of the inductor. The current at every point is the same. By Kirchoff's voltage law, the *emf* of the battery must equal the *emf* across the inductor, so we have

$$\xi - L\frac{\mathrm{d}I}{\mathrm{d}t} = 0 \qquad \Longrightarrow \qquad I(t) = \frac{\xi}{L}t, \tag{197}$$

i.e. the current through the inductor grows linearly in time. This makes sense: a linearly increasing current gives a linearly rising flux through the inductor, and hence a constant back *emf* across the inductor, which matches the *emf* of the battery.

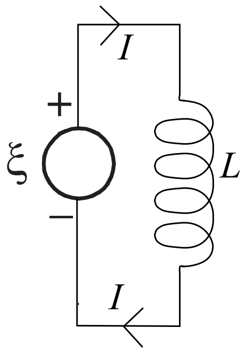

Figure 72: A battery and an inductor.

6.9.7 Energy in an inductor

In the above example, an ever increasing amount of current flows from the positive terminal of the battery to the negative one, releasing the battery's energy at a rate ξI, leading us to ask where this energy has gone. There is no conventional resistance in our circuit to dissipate the energy as heat. The answer is that the energy is stored in the ever increasing magnetic field inside the inductor. We can calculate how much energy the inductor stores by looking at how much energy the battery has released by time T

$$W = \int_0^T \xi I(t)\mathrm{d}t = \int_0^T \frac{\xi^2 t}{L}\mathrm{d}t = \frac{\xi^2 T^2}{2L} = \frac{1}{2}LI(T)^2. \tag{198}$$

More generally, the power absorbed by an inductor is, as always the product of the current through it and the voltage over it $P = VI = IL\frac{\mathrm{d}I}{\mathrm{d}t}$. The power absorbed in a time $\mathrm{d}t$ is thus $\mathrm{d}W = P\mathrm{d}t = VI = IL\frac{\mathrm{d}I}{\mathrm{d}t}\mathrm{d}t = IL\mathrm{d}I$, so the power absorbed building the current up from zero to I is

$$W = \int_0^I IL\mathrm{d}I = \frac{1}{2}LI^2, \tag{199}$$

which agrees with our previous calculation. This is the general result of the energy stored in an inductor when it is carrying a current I.

6.9.8 Exponential decay of an RL circuit

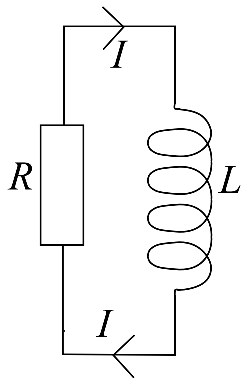

Figure 73: A inductor and a resistor.

Consider a circuit consisting of an inductor and a resistor in a loop, as shown in fig. 73. If, at $t = 0$, we prepare the circuit with a current I flowing in it, what happens next? The flow of the current is resisted by the resistor, but a reduction in the current is resisted by the inductor. Applying Kirchoff's voltage law to the loop, we the back *emf* from the inductor must match the potential difference across the resistor, giving

$$-IR - L\frac{\mathrm{d}I}{\mathrm{d}t} = 0 \qquad \Longrightarrow \qquad I(t) = I(0)\mathrm{e}^{-\frac{tR}{L}}, \tag{200}$$

i.e. the current in the circuit decays exponential with decay time $\tau = L/R$.

We can understand this from an energy perspective. At the outset the inductor is storing energy $\frac{1}{2}LI(0)^2$. However, the same current, I, is flowing through the resistor, dissipating energy as heat at a rate $P = I^2R$. This energy must ultimately come from the stock of energy in the inductor, so this stock falls in time, meaning the current in the loop falls in time. The total energy dissipated in the resistor is

$$W = \int_0^\infty I^2R\mathrm{d}t = \int_0^\infty I(0)^2R\mathrm{e}^{-\frac{2Rt}{L}}\mathrm{d}t = \frac{1}{2}LI(0)^2, \tag{201}$$

i.e. all the energy that starts in the inductor is ultimately dissipated in the resistor.

The energy stored in the inductor decays as $E_L(t) = \frac{1}{2}LI(t)^2 = E_L(0)\mathrm{e}^{-\frac{2Rt}{L}}$, so it also decays exponentially but at twice the rate as the current, with decay time $\tau = L/(2R)$.

6.10 Summary of Circuits

- The electric current through a surface is the rate at which charge passes through the surface:

$$I = \frac{\mathrm{d}q}{dt}. \tag{202}$$

- The potential difference between two points is the amount of electrical energy turned to other forms of energy when a unit of charge moves between them, thus when a charge q moves,

$$W = qV. \tag{203}$$

- A battery gives electrical energy to the charges that flow through it. In a circuit the potential difference across a perfect battery is always its *emf*, ξ,

$$V = -\xi. \tag{204}$$

- When a capacitor of capacitance C holds a charge Q, the potential difference across it is given by

$$V = Q/C. \tag{205}$$

- When a current I flows through a resistor R the potential difference across it is

$$V = IR. \tag{206}$$

- When the current through an inductor, L, is changing the potential difference across it is

$$V = L\frac{\mathrm{d}I}{\mathrm{d}t}. \tag{207}$$

- To analyze circuits, we need the above constitutive laws for each component, and Kirchoff's laws. Firstly, at a junction in a circuit current is conserved $\sum_i I_i = 0$. Secondly, the sum of the potential differences around every loop in a circuit add to zero $\sum_i V_i = 0$.

- When a current I flows through a potential difference V, the rate at which electrical energy is converted to other forms is

$$P = VI \tag{208}$$

- A battery thus puts electrical energy into a circuit at a rate

$$P = I\xi. \tag{209}$$

- The power that flows into a resistor is dissipated as heat at a rate

$$P = VI = V^2/R = I^2R. \tag{210}$$

- The power that flows into capacitors and inductors is stored and can be recovered later. The energies stored in a capacitor and an inductor are

$$W_C = \tfrac{1}{2}QV = \tfrac{1}{2}CV^2 = \tfrac{1}{2}Q^2/(2C) \qquad W_L = \tfrac{1}{2}LI^2. \tag{211}$$