Fields and Waves

Tony Paxton

This is a first year physics course that I taught at King's College London between 2014 and 2017. There are both problems classes and tutorial problems with solutions.

I have taken a lot of material from David Griffiths's brilliant textbook which I am very glad to acknowledge; and also some figures from Halliday and Resnick (with permission).

4CCP1501: Fields and Waves, 2017-8

Lecturer: Prof Tony Paxton (office S3.10)

Lecture 1

1.1 Gravity and mass

We begin with gravity. Two objects of masses m_1 and m_2, a distance r apart, attract each other with a force whose magnitude is

$$F = G\frac{m_1 m_2}{r^2}$$

The force is in newtons, [N],

$$1 \text{ N} = 1 \text{ kg m s}^{-2} \qquad (\text{force} = \text{mass} \times \text{acceleration})$$

and the *universal gravitational constant* is

$$G = 6.674 \times 10^{-11} \text{ N m}^2 \text{ kg}^{-2}$$

Actually force is a vector having *magnitude* and *direction*. We agree on a cartesian coordinate system and then we can place mass number one at $\mathbf{r}_1$ and mass two at $\mathbf{r}_2$. The vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ then have *components*:

$$\mathbf{r}_1 = x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}} + z_1\hat{\mathbf{k}} = (x_1, y_1, z_1)$$
$$\mathbf{r}_2 = x_2\hat{\mathbf{i}} + y_2\hat{\mathbf{j}} + z_2\hat{\mathbf{k}} = (x_2, y_2, z_2)$$

$\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are *unit vectors* pointing in the agreed x, y and z directions. They have no units (dimensions)

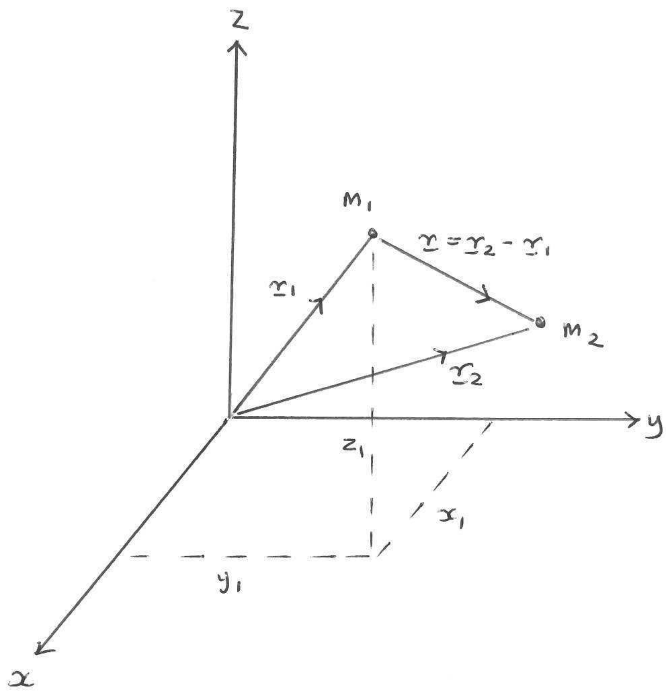


FIGURE 1–1

The force acting upon mass one, m_1, is directed along the vector $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. Vector addition (or subtraction) means adding (or subtracting) the vector components. Thus

$$\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$$

implies that if $\mathbf{r} = x\hat{\mathbf{\imath}} + y\hat{\mathbf{\imath}} + z\hat{\mathbf{k}}$ then

$$x = x_2 - x_1$$
$$y = y_2 - y_1$$
$$z = z_2 - z_1$$

So the force *on* mass one *due to* mass two is

$$\mathbf{F}_{12} = G\frac{m_1 m_2}{r^2}\,\hat{\mathbf{r}}$$

r not in bold face or underlined, is the *magnitude* of $\mathbf{r}$ (units of length)
$= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$
$\hat{\mathbf{r}}$ is the unit vector (dimensionless)
$= \frac{1}{r}\,\mathbf{r}$

The fact that $\mathbf{F}_{12} = -\mathbf{F}_{21}$ expresses *Newton's third law*—loosely: "every action has an equal and opposite reaction".

The gravitational force depends on the distance r like $1/r^2$. This is called an *inverse square law*, and such forces have very special properties as we'll see later.

If I introduce a third mass, m_3, this does *not* modify the force already acting between masses one and two. Therefore the force acting on mass one is now the force due to mass two plus the force due to mass three:

$$\mathbf{F}_{\text{total}} = \mathbf{F}_{12} + \mathbf{F}_{13}$$

and this is a *vector addition.* This is very important and is called the *principle of superposition.*

Because of the inverse square law we can show that the force acting on a mass m due to a spherically symmetric object of finite size and total mass M is the same as if all the mass of this object were concentrated at its centre. In fact we *will* show this when we come to do electrostatics. For example the force due to the Earth's gravity on an object of mass m a small distance d above the Earth's surface is, assuming the Earth to be a perfectly spherically symmetric body,

$$F_g = G\frac{M_E m}{(r_E + d)^2} = \frac{GM_E m}{r_E^2}$$

if $r_E \gg d$, and r_E is the radius of the Earth.

The magnitude of the force F_g is called the *weight* of the object having mass m. Since the weight is proportional to the mass at a given height h above the Earth's surface, we can write

$$F_g = mg$$

and g is the *acceleration due to gravity* at the height h. At the Earth's surface we use $g = 9.80\ \mathrm{ms}^{-2}$.

Do not confuse mass and weight. Mass is measured in kilograms [kg] and weight is in newtons [N]. The kg is *not* a unit of weight despite what your greengrocer might tell you.†

The attraction between masses one and two is an "action at a distance". We say that mass two sets up a *gravitational field* that is felt by mass one, and *vice versa*. The attraction is exactly the same if the two masses are immersed in some medium.

The mass we have talked of up to now should be called *gravitational mass*. Newton's first law states that an object that is stationary in some inertial frame of reference remains stationary in that frame unless acted upon by a force. The action of that force causes it to accelerate, that is, to change its velocity. It is found that the amount of acceleration, a, is proportional to the force applied; the proportionality constant is called the *inertial mass* of that object:

$$F = ma$$

We don't need to distinguish notationally between gravitational and inertial mass because for any given object they turn out to have the same numerical value in kg (or in any other unit of mass, for example, the slug‡.) This is a deep and non trivial result.

---

† Actually if you buy, say, "3 kg" of apples you do actually get 3 kg of apples. This is because you are actually getting the weight of your apples measured in a non SI unit, the kilogram-force [kg-f]. This is defined as the force exerted on a one kg mass by the Earth's gravity, whereas the SI unit of force, the newton, is the force exterted on a mass of one kg by an acceleration of $1\ \mathrm{m\ s}^{-2}$. Since a 3 kg mass experiences a force of $3g$ in the Earth's field you are actually buying 3 kg mass of apples. The tension in the greengrocer's spring balance is $3g$ newtons. It is best to use *only SI units—always, without exception.*

‡ In the United States of America they say that a force of one *pound* [lb] produces an acceleration of $1\ \mathrm{ft\ s}^{-2}$ on an object of mass one slug. So the pound is a unit of weight, *not* mass. So when people say one pound equals 0.45 kg they really mean 0.45 kg-f. Neither is an SI unit, so don't use them.

1.2 The mass on the spring

I don't want to get bogged down with gravity and the weight of the mass, so my mass is lying horizontally on a frictionless surface and connected to a rigid wall by an ideal linear spring having no inertia and a stiffness k. That is, k is the the tension per unit displacement. If the spring is stretched by an amount x, the restoring force is

$$F = -kx$$

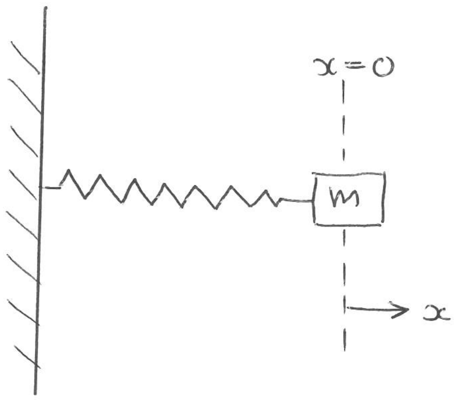


FIGURE 1–2

At $x = 0$ the spring is relaxed—neither extended nor compressed. If I pull it out to an extent $x = A$ and let it go the force in the spring will cause it to accelerate. In the differential calculus, we write

$$a = \frac{\mathrm{d}^2x}{\mathrm{d}t^2}$$

and since $F = ma$, we have

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\frac{1}{m}kx = -\omega_0^2 x$$

having defined

$$\omega_0^2 = \frac{k}{m}$$

Is this your first differential equation? The solution to

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\omega_0^2 x$$

is

$$x = A\cos(\omega_0 t + \phi) \tag{1.1}$$

in which A, ω_0 and ϕ are constant. If you don't believe me go on and try it! In fact try it anyway. You'll need to be able to differentiate sine and cosine. Whereas ω_0 is a physical property of the system, independent of whether or how it is moving, A and ϕ are *arbitrary constants* in the sense that *any* choice of these furnishes us with a solution of the differential equation. The theory of second order differential equations states that a solution must contain two arbitrarily variable constants. These are then later determined once we know a bit more about the *physics* of the problem, namely the *boundary conditions*. This becomes a lot clearer now.

A is called the *amplitude* of the oscillation; it is the greatest distance the mass travels away from $x = 0$.

ϕ is called the *phase* or *phase angle*. If I look at my watch when I release the mass and record the time as t_1 then I have:

$$(\text{at } t = t_1 \;, \quad x = A) \;\longleftarrow \text{a "boundary condition"}$$

This implies that

$$\cos(\omega_0 t_1 + \phi) = 1$$

and a solution of that is

$$\omega_0 t_1 + \phi = 0$$

or

$$t_1 = -\frac{\phi}{\omega_0}$$

Of course to make life a lot simpler I can always reset my watch to $t = 0$, when I release the mass; that is to say I use a stop watch. In that case $t_1 = 0$ and $\phi = 0$ and $x = A\cos\omega_0 t$.

The phase in that sense is *arbitrary*; but we certainly need it if we compare two or more oscillators, for example, whose masses have been released at different times.

Any system that is oscillating according to equation (1.1) is said to be executing *simple harmonic motion*.

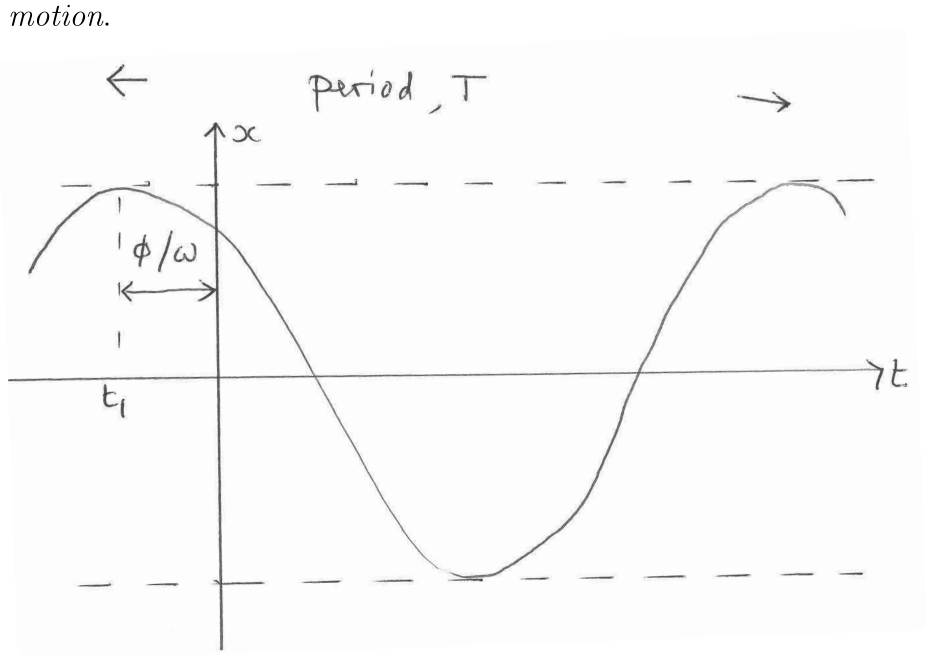


FIGURE 1–3

Already in this lecture we have encountered a *field* and a *wave*!

If the position of the mass is x, then its speed or the *magnitude of its velocity* (we often use "velocity" when we really mean "speed", but it will not cause any confusion) is

$$\begin{aligned} v &= \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left[A\cos\left(\omega_0 t + \phi\right)\right] \\ &= -A\omega_0 \sin\left(\omega_0 t + \phi\right) \end{aligned}$$

so its *kinetic energy* is

$$E_{\mathrm{kin}} = \frac{1}{2}mv^2 = \frac{1}{2}mA^2\omega_0^2 \sin^2\left(\omega_0 t + \phi\right)$$

Potential energy is stored in the spring, and this is

$$\begin{aligned} E_{\mathrm{pot}} &= \frac{1}{2}kx^2 = \frac{1}{2}m\omega_0^2 x^2 \\ &= \frac{1}{2}mA^2\omega_0^2 \cos^2\left(\omega_0 t + \phi\right) \end{aligned}$$

Since $\cos^2\theta + \sin^2\theta = 1$, the *total energy* is

$$\begin{aligned} E_{\mathrm{tot}} &= E_{\mathrm{kin}} + E_{\mathrm{pot}} \\ &= \frac{1}{2}mA^2\omega_0^2 = \frac{1}{2}kA^2 \end{aligned}$$

This is *constant*, independent of time; which is just as well. If it weren't then you'd be very worried about what has happened to the Law of Conservation of Energy.

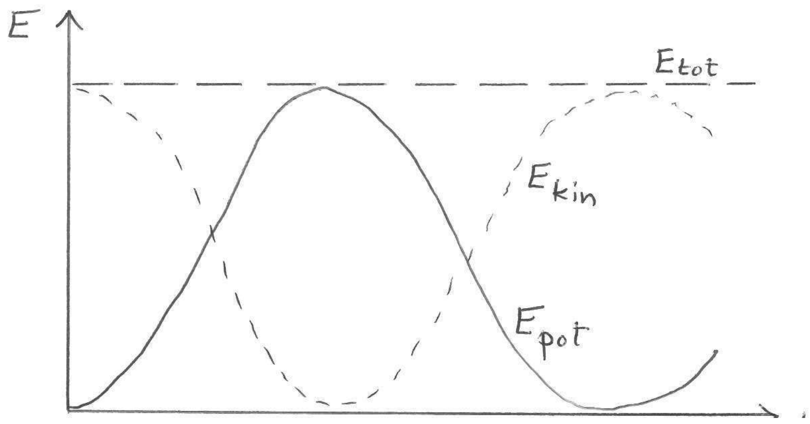


FIGURE 1–4

We say that E_{pot} and E_{kin} are oscillating 180° out of phase. At any time they add to give a constant E_{tot}.

Notice that x and v are also always 180° out of phase.

Cosine and sine are *periodic* or *circular* functions. As you can see, x and v return to their values at time t at a later time $t + T$ and again at $t + 2T$ and so on. T is called the *period* of the oscillation. We must have

$$\cos\left(\omega_0 t + n\omega_0 T\right) = \cos\omega_0 t$$

and this can only be true if

$$\omega_0 T = 2\pi$$

so

$$T = \frac{2\pi}{\omega_0}$$

If T is the period (in seconds) then there are

$$f_0 = \frac{1}{T}$$

oscillations, or *cycles*, per second. f_0 is called the frequency (often given the Greek symbol ν) and has units of Hertz [Hz] or *cycles per second*. ω_0 is hence equal to $2\pi f_0$ and is called the *angular frequency*. Its units are *radians per second*. Remember there are 2π radians of angle in a complete circle: $360° = 2\pi$ radians.

In terms of the constants of the system, k and m, remembering $\omega_0^2 = k/m$, we find

$$T = 2\pi\sqrt{\frac{m}{k}}$$

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$$

Lecture 2

Waves are everywhere in physics. We start with a disturbance called a *pulse* produced, say, by jerking one end of a rope.

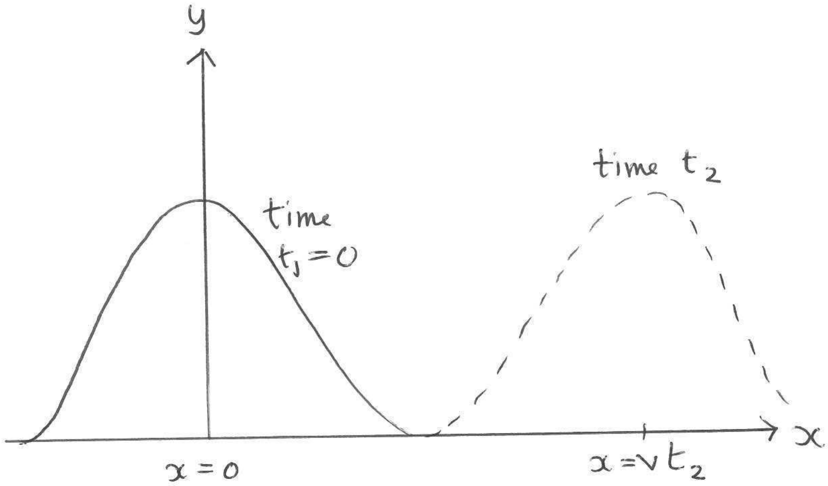


FIGURE 2–1

Figure 2–1 shows the height, y, of the pulse as a function of distance along the rope, x, at two times, $t_1 = 0$ and $t_2 \neq 0$. At time t_2 the crest of the pulse has travelled to position $x = vt_2$ along the rope if v is the speed of the pulse.

For any time t we can write

$$y = y(x, t)$$

meaning "y is a function of both time and position along the rope".† The graph shows that for any time t, including t_2,

$$y(x, t) = y(x - vt, 0)$$

This means that the variables x and t must appear in a special combination; if the pulse is travelling to the right $y(x, t)$ is a function of $x - vt$, that is,

$$y(x, t) = y(x - vt) \quad \longleftarrow \quad \text{wavefunction of a right travelling pulse}$$

and if it's travelling to the left

$$y(x, t) = y(x + vt) \quad \longleftarrow \quad \text{wavefunction of a left travelling pulse}$$

This is called the *wavefunction*.

† Please don't get confused! It is usually obvious if I write, say, $f(x)$, that I don't mean $f \times x$. However if I write $y(a + b)$ I probably mean $ya + yb$, but I could in principle mean "y is a function of $a + b$", so if in doubt, *ask!* $y(x, t)$ is unambiguous because of the comma.

Now we turn from a pulse to a wave.

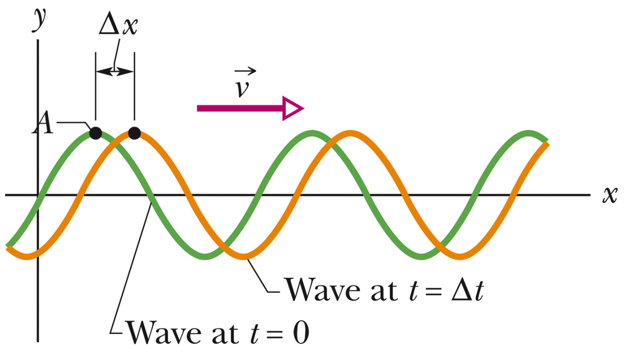


FIGURE 2–2

If we set our clock such that at $t = 0$, $x = 0$ and $y = 0$ then this is a sine wave,

$$y(x, 0) = A \sin ax$$

where a is to be determined[†] as follows; y is again zero at $x = \frac{1}{2}\lambda$, where λ is the distance between successive events, or *wavelength*. Therefore

$$y\left(\frac{1}{2}\lambda, 0\right) = A \sin \frac{1}{2} a\lambda = 0$$

which means that $\frac{1}{2}a\lambda = \pi$, that is

$$a = \frac{2\pi}{\lambda}$$

We can find the time dependence of the wavefunction in the same way as for the pulse. All this leads to

$$y(x, t) = A \sin\left(\frac{2\pi}{\lambda}(x - vt)\right)$$

v is the speed at which a crest of the wave travels to the right. It is called the *phase velocity*. If the period is T then the wave travels a distance λ in time T so

$$v = \frac{\lambda}{T} = f\lambda$$

where f is the *frequency*, having units of s^{-1} or Hz. To clear away the factor of 2π, I can work with the *angular frequency* (see Lecture 1)

$$\omega = 2\pi f \quad [\text{radians s}^{-1}]$$

† Here is *yet another* possible interpretation of $y(x, t)$. Here it means, “knowing the function y, evaluate it for particular values that I give you of x and t”. Don’t blame me: this is maths! What you can blame me for is writing $\sin ax$ when I should write $\sin(ax)$; but I don’t like needless parentheses if it seems pretty obvious what I mean. Again, if you’re puzzled, *ask*—the rest of the class will thank you for it.

and a quantity called the *wavenumber*

$$k = \frac{2\pi}{\lambda} \quad [\mathrm{m}^{-1}]$$

(this is not a good name for it as it's not a number, it has units of $[\mathrm{L}]^{-1}$. A better name is "the magnitude of the wavevector", or wavevector for short) and then the wavefunction is

$$y = A \sin(kx - \omega t)$$

Or, if the wavefunction is not zero at $x = 0$ we can account for this by introducing a *phase angle*

$$y = A \sin(kx - \omega t + \phi)$$

Usually for a single wave, the phase is arbitrary. If I set $\phi = \frac{1}{2}\pi$ then the wavefunction is

$$y = A \cos(kx - \omega t)$$

because $\sin(\theta + \frac{1}{2}\pi) = \cos\theta$.

Normally some device or experiment in physics produces more than one wave at a time. There is a *superposition principle*: if two or more waves exist in the same medium the resultant wavefunction is the sum of the wavefunctions of the individual waves.

I want to examine two cases.

1. Two waves having the same frequency but different phases—*interference*
2. Two waves having the same phase but different frequency—*beats*

Case 1—interference

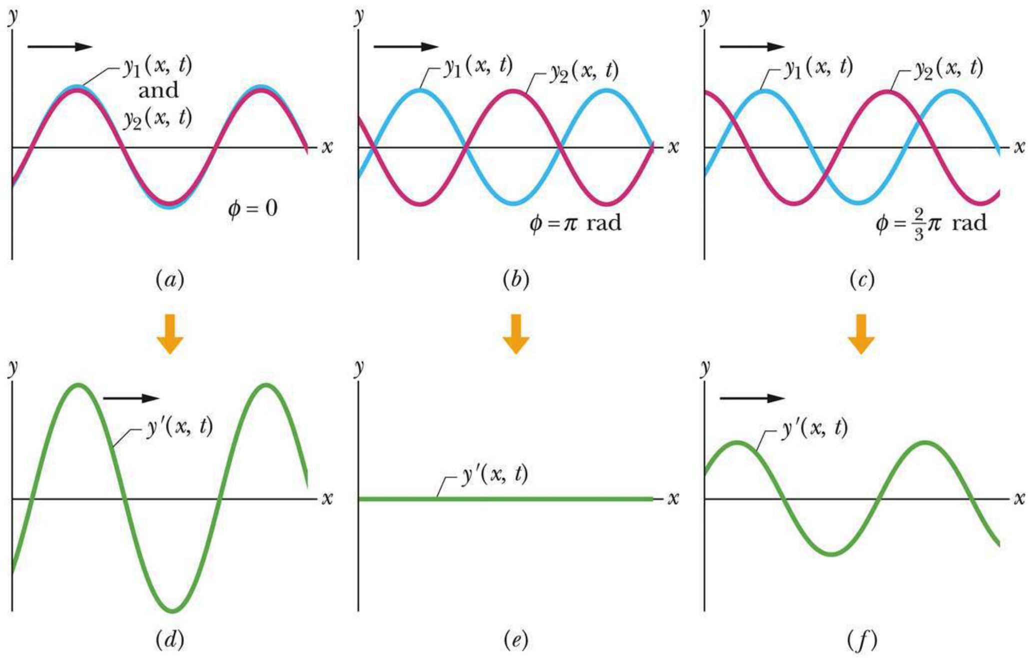


FIGURE 2–3

We describe this mathematically as follows. The two wavefunctions are

$$\begin{aligned} y_1 &= A\sin(kx - \omega t)\ , \quad \text{phase arbitrarily set to zero} \\ y_2 &= A\sin(kx - \omega t + \phi) \end{aligned}$$

By the principle of superposition, we add the wavefunctions,

$$y = y_1 + y_2 = A\left[\sin(kx - \omega t) + \sin(kx - \omega t + \phi)\right]$$

Now we use the identity

$$\sin a + \sin b = 2\cos\frac{1}{2}(a-b)\sin\frac{1}{2}(a+b)$$

and this leads to

$$y = \underbrace{2A\cos\frac{1}{2}\phi}_{\text{amplitude}}\ \sin\left(kx - \omega t + \frac{1}{2}\phi\right)$$

The wave now has an amplitude anything between zero (when, say, $\phi = \pi$ radians $= 180°$) and $2A$ (when, say, $\phi = 0$ which results in *twice* the original amplitude); that is, *destructive* or *constructive* interference. The new wave has the same wavelength and frequency as the two combining waves and its phase is now the *mean* of the original phases. (We could have done this with phases ϕ_1 and ϕ_2 and found the new phase to be $\phi = \frac{1}{2}(\phi_1 + \phi_2)$. Try it if you like.) What if the two amplitudes were different?

Case 2—beats

Be careful: sometimes I plot the wavefunction y against x. That means I am sketching the *waveform* at some fixed time. On other occasions, such as in what comes next, I plot y against t. That means I imagine standing in a fixed spot, x, and watching how the wavefunction (displacement, disturbance or whatever constitutes the wave) is varying in time.

Now we study the situation where two waves combine having the same phase and amplitude but different frequencies. The wavelengths are also different, but for now we are not interested in the x-dependence so we leave it out. Also for this case it's easier to deal with cosines than sines, so I fix the phase of each wave to $\phi = \frac{1}{2}\pi$. Then the two waves are

$$\begin{aligned} y_1 &= A\cos(2\pi f_1 t) \\ y_2 &= A\cos(2\pi f_2 t) \end{aligned}$$

and the superposition principle requires the combined wavefunction to be

$$y = A\left(\cos 2\pi f_1 t + \cos 2\pi f_2 t\right)$$

We will use the identity

$$\cos a + \cos b = 2\cos\frac{1}{2}(a-b)\cos\frac{1}{2}(a+b)$$

which leads to

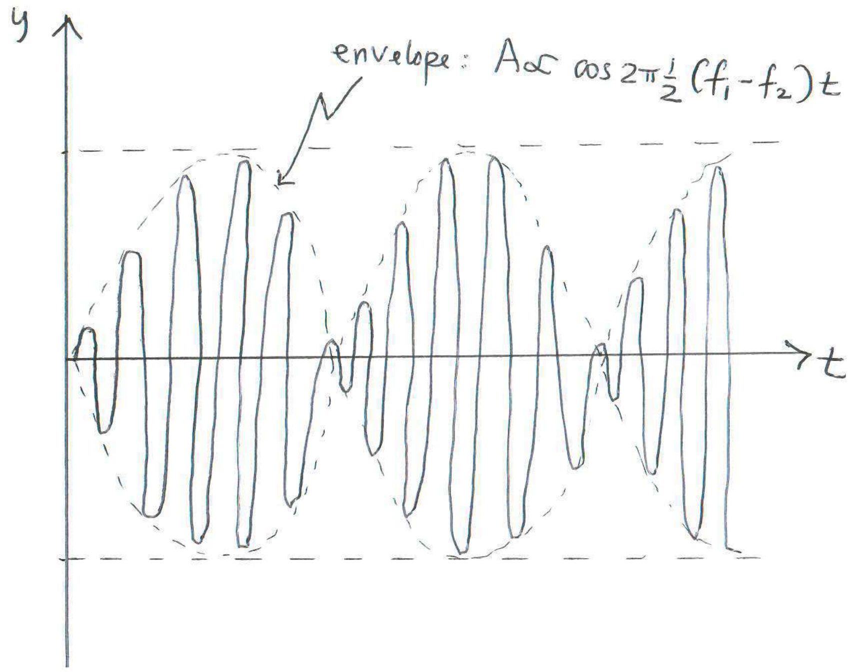
$$y = \underbrace{2A\cos\left(2\pi\frac{1}{2}\left(f_1 - f_2\right)t\right)}_{\text{amplitude}}\cos\left(2\pi\frac{1}{2}\left(f_1 + f_2\right)t\right)$$

We have a new wave whose whose wavefunction is proportional to

$$\cos\left(2\pi\frac{1}{2}\left(f_1 + f_2\right)t\right)$$

so the new frequency is the *mean* of the two combining frequencies.

But the amplitude *also* depends on time. It oscillates with a frequency that is half the difference of the two original frequencies. So at times the amplitude is twice the original and at other times it is zero and the waves are extinguished.



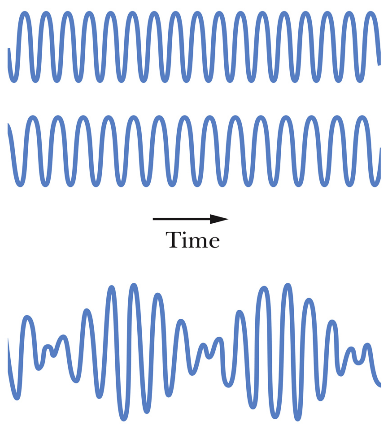


FIGURE 2–4

The amplitude will go through a maximum whenever the cosine function is either $+1$ or -1, that is *twice* in every period. Therefore we will observe oscillations in the amplitude at a *beat frequency* of $|f_1 - f_2|$.

This is the phenomenon of *beats*. A good example is the sound made by two guitar strings slighly out of tune; or two tuning forks of similar but not the same frequency.

Lecture 3

3.1 Wavepackets

Let us return to combinations of waves having different frequencies and wavelengths. The difference in wavelength, or *spread*, if many waves are combined, will be expressed using the measure

$$\Delta k = k_2 - k_1$$

where k_1 and k_2 are *wavenumbers* (see Lecture 2, page 6)

$$k_1 = \frac{2\pi}{\lambda_1}\ , \qquad k_2 = \frac{2\pi}{\lambda_2}$$

The difference in *angular frequencies* is

$$\Delta\omega = \omega_2 - \omega_1$$

and a combination of two waves, having the same amplitude for simplicity, is

$$\begin{aligned} y &= A\cos\left(k_1 x - \omega_1 t\right) + A\cos\left(k_2 x - \omega_2 t\right) \\ &= \underbrace{2A\cos\left(\frac{1}{2}\Delta k\, x - \frac{1}{2}\Delta\omega\, t\right)}_{\text{amplitude}}\ \cos\left[\frac{1}{2}\left(k_1 + k_2\right)x - \frac{1}{2}\left(\omega_1 + \omega_2\right)t\right] \end{aligned} \tag{3.1}$$

Now, again, we have an *amplitude* that varies in both *time* and *space*, and a new wavenumber and angular frequency each of which are the *mean* of those of the combining waves. It is not difficult, as you can possibly see to extend this argument to the case of a combination of many waves—but that case we will only treat qualitatively in these notes. Indeed we can combine a large number of waves in an experiment taking only wavenumbers between k_1 and k_2, hence having a spread of Δk, and similarly angular frequencies from a distribution of width $\Delta\omega$. The result is a train of localised *pulses*.

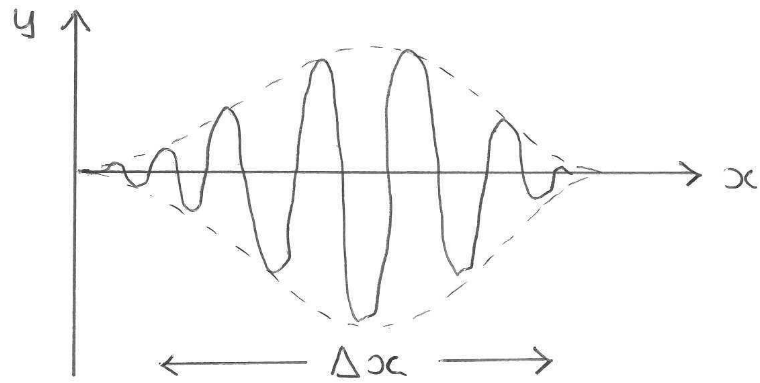


FIGURE 3–1

This pulse is called a *wavepacket*. If I stand at a fixed point and observe the pulse going past me, I can estimate the time it takes to do that and call this Δt, the *width in time* of the wavepacket. Now the wavepacket becomes localised in time because of destructive interference at the edges of the packet and constructive interference at its centre. To achieve this condition in the most economical way, we want the number of periods of the waves of maximum frequency, $\omega_2/2\pi$, in the time interval Δt to be just *one greater* than the number of periods of the waves of the minimum frequency, $\omega_1/2\pi$. This ensures constructive interference in the centre of the wavepacket and destructive interference at the edges, because if the two waves of minimum and maximum frequency are out of phase at $t = 0$ they will be out of phase again at $t = \Delta t$. You will need to draw a few sketches to convince yourself of this.

Now the number of periods in the interval Δt is $\Delta t/T = f\Delta t$, if T is the period and f is the frequency. So if the greatest frequency is f_{max} and the smallest is f_{min} then we want the difference in the number of periods contained in the interval Δt to be <u>one</u> and so

$$f_{\text{max}}\Delta t - f_{\text{min}}\Delta t = 1 \ \ \text{(or a larger odd integer)}$$

which is the same as

$$\Delta t\,\Delta\omega \geq 2\pi$$

So to achieve a *short* pulse we require a combination of waves with a *broad* spread of frequencies.

On the other hand, to achieve a narrow wavepacket *in space* the same argument leads to a spread of wavelengths $\lambda_{\text{max}} - \lambda_{\text{min}}$ such that

$$\Delta x\left(\frac{1}{\lambda_{\text{min}}} - \frac{1}{\lambda_{\text{max}}}\right) = 1 \ \ \text{(or a larger odd integer)}$$

In terms of wavenumbers, this is

$$\Delta x\,\Delta k \geq 2\pi$$

which is called the *uncertainty relation*. In Heisenberg's quantum mechanics we use de Broglie's relation $p = h/\lambda$ between the wavelength of matter waves and their momentum p; h is the Planck constant. Our formula then becomes

$$\Delta x\,\Delta p \geq h$$

which is the celebrated Heisenberg uncertainty principle.† It states that a simultaneous measurement of a quantum particle's position x and momentum p will lead to an uncertainty in outcomes so that the spread in measurements of momentum Δp times the spread in measurements of position Δx is such that their product *cannot* be smaller than half the reduced Planck constant.† Of course this is a statistical argument, so it can only make sense if we imagine possessing a very large number of copies of the particle and after making the simultaneous measurements of x and p we make a statistical analysis of the results. From that point of view Δx and Δp are exactly the *standard deviations*,

† Actually the uncertainty principle is $\Delta x\,\Delta p \geq h/4\pi$ so our crude analysis based on just the largest and smallest wavelength taken from all the waves that make up the wavepacket *overestimates* the actual lower bound on $\Delta x\,\Delta p$ by a factor of order ten.

σ_x and σ_p, arising from the statistical analysis. You will learn all this when you come to do quantum mechanics; and at that point it may be profitable to you to come back to these notes and appreciate that the uncertainty principle is really a consequence of the notion of a particle as a *wavepacket* of matter waves. In fact having made that leap forward, the uncertainty principle itself is just a simple result belonging to all kinds of wave motion—quantum or classical.

3.2 Group velocity

The wavepacket does not necessarily travel at the same speed as the waves that combine to produce it. The speed of an individual component of angular frequency ω and wavenumber k is

$$v = \lambda f = \frac{\omega}{k}$$

and this is usually called the *phase velocity.* It is the speed at which a crest is travelling. For a single wave, $y = A\cos(kx - \omega t)$, we see that the phase velocity is

$$v = -\frac{\text{coefficient of the time variable}}{\text{coefficient of the space variable}}$$

In the case of the wavepacket, we look at the *amplitude* term in equation (3.1) rather than the *wave* term because we are interested in the speed of the amplitude peak. So we write

$$\text{amplitude} = 2A\cos\left(\frac{1}{2}\Delta k\, x - \frac{1}{2}\Delta\omega\, t\right)$$

and apply the same argument to find the speed of the centre of the wavepacket. This is

$$v_g = -\frac{\text{coefficient of the time variable}}{\text{coefficient of the space variable}} = \frac{\Delta\omega}{\Delta k}$$

In the case of a very large number of waves this becomes the derivative

$$v_g = \frac{\mathrm{d}\omega}{\mathrm{d}k}$$

and because in order to interfere the waves must be all of the same kind, that is the same disturbance in the same medium, they will all have the same dependence of frequency on wavelength, which is expressed as a *dispersion relation*:

$$\omega = \omega(k) \quad \longleftarrow \quad \text{“}\omega \text{ is a function of } k\text{”}$$

v_g is called the *group velocity.* It may be the same as the phase velocity particularly if the dispersion relation is *linear* as for example in the case of monochromatic light in a vacuum for which $\omega = ck$ so that

$$\frac{\mathrm{d}\omega}{\mathrm{d}k} = v_g = \frac{\omega}{k} = v = c\ , \qquad \text{the speed of light}$$

In quantum mechanics, a free particle of mass m has the following dispersion relation,

$$\omega = \frac{1}{2}\frac{\hbar k^2}{m}$$

where $\hbar = h/2\pi$ is the *reduced Planck constant.* So the phase velocity is

$$v = \frac{1}{2}\frac{\hbar k}{m}$$

while the group velocity is

$$v_g = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{\hbar k}{m}$$

exactly *twice* the phase velocity.

Conversely for *deep water waves*, the group velocity is exactly half the phase velocity; you can observe this if you look very carefully at the ripples emerging when you drop a stone into a still pond. You will see a wave rising in the tail of the pulse, moving through the pulse at exactly *twice* the speed of the pulse and increasing in amplitude until it reaches the centre. Its amplitude then decays as it approaches the head of the pulse where it dies away.

3.3 Damping

We turn to a new topic now, but return to an old friend, the mass on the spring. We add some new physics, namely *damping.* No oscillator vibrates forever: it will always be damped by interaction with its environment, usually by some frictional process which generates heat and hence dissipates energy. We model this by causing the mass to vibrate inside a viscous fluid.

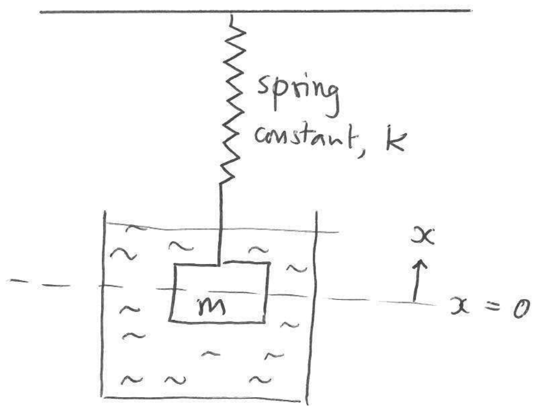


FIGURE 3–2

The mass is at rest at position $x = 0$. We ignore gravity and there are now two forces acting on the mass when it is displaced from equilibrium, $x = 0$.

1. The restoring force due to the spring. This depends linearly on the *displacement* x because we have used a linear spring

$$F_{\text{spring}} = -kx$$

The force is always in the opposite direction to the displacement because it's a *restoring force.*

2. The fluid exerts a force but only when the mass is moving. We suppose that the force is linearly proportional to the velocity,

$$F_{\text{damping}} = -bv = -b\frac{\mathrm{d}x}{\mathrm{d}t} = -b\dot{x}$$

in the opposite direction to the velocity. b is called the *damping coefficient* and has units of [N s m^{-1}] or [kg s^{-1}].

As in Lecture 1, page 4, we try to find the displacement as a function of time by solving Newton's second law, force = mass × acceleration, as a differential equation.

$$\text{acceleration is } \frac{\mathrm{d}^2x}{\mathrm{d}t^2} = \ddot{x}$$
$$\text{velocity is } \frac{\mathrm{d}x}{\mathrm{d}t} = \dot{x}$$
$$\text{force is } F_{\text{spring}} + F_{\text{damping}}$$

So we have, putting all this together,

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx - b\frac{\mathrm{d}x}{\mathrm{d}t}$$

You remember that the natural frequency of the oscillator, ω_0, without the damping is a function of the mass, m, and the spring constant, or *stiffness*, k.

$$\omega_0 = \sqrt{\frac{k}{m}} \qquad \text{[radian s}^{-1}\text{]}$$

Further, to simplify the differential equation, we define a new property of the damped oscillator called the *damping ratio*. Often it is given the symbol ζ which is hard to write, so I will use a Z. By definition

$$Z = \frac{1}{2}\frac{b}{m\omega_0} \qquad \text{[dimensionless]}$$

Our equation of motion becomes

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2Z\omega_0\frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2 x = 0 \qquad (3.2)$$

and please note that this is now independent of the nature of the mechanical device—it is true for any system possessing a natural frequency and a damping ratio, for example an electronic oscillator circuit, which is the subject of Lecture 20.

I will only tell you here about the solution of (3.2) when $Z < 1$: so called *underdamping*. Actually it's best first to sketch how we'd expect x to depend on t which I do in figure 3–3. The displacement will oscillate at some frequency we will call ω_D which will not be the natural frequency of the undamped counterpart; and we'd expect the amplitude to decay with time as energy is dissipated.

Only trigonometric and exponential functions solve this kind of differential equation so the amplitude must decay exponentially. In fact

$$x = x_m\, e^{-Z\omega_0 t}\, \cos \omega_D t$$

is a solution to (3.2) and the *damped frequency*, ω_D, is always *smaller* than the natural undamped frequency,

$$\begin{aligned} \omega_D &= \omega_0 \sqrt{1 - Z^2} \\ &= \omega_0 \sqrt{1 - \frac{1}{4}\frac{b^2}{mk}} \end{aligned}$$

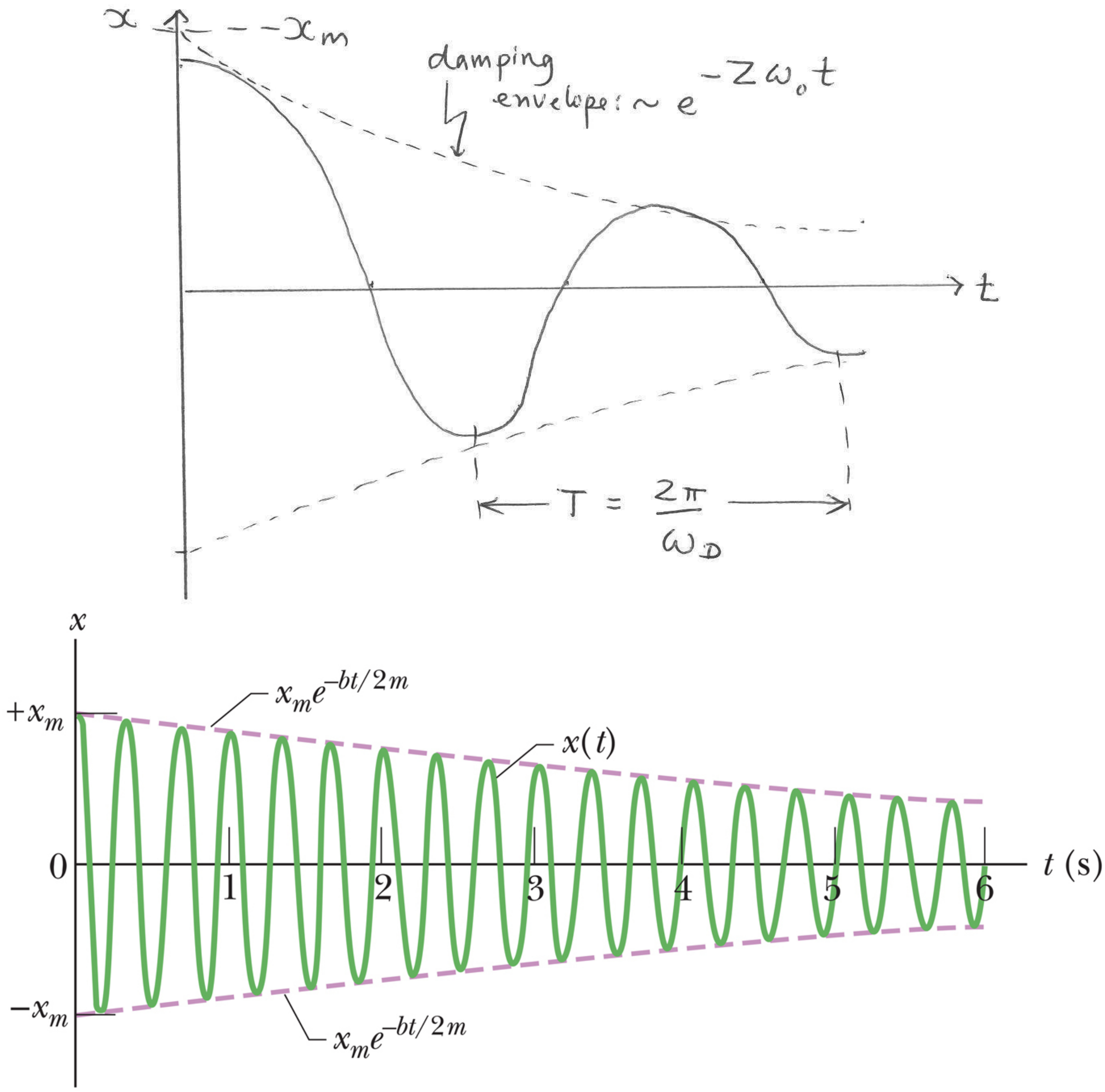


FIGURE 3–3

The total energy of the damped oscillator is

$$\begin{aligned} E &= \text{kinetic energy} + \text{potential energy} \\ &= \frac{1}{2}m\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \frac{1}{2}kx^2 = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega_0^2x^2 \end{aligned} \tag{3.3}$$

You can see how neat it is to use the notation that a "dot" over a quantity indicates a derivative with respect to time. Now, if I take equation (3.2) and write it using the overdot notation for first and second time derivatives it looks like this.

$$\ddot{x} + 2Z\omega_0\dot{x} + \omega_0^2x = 0$$

I multiply through by $m\dot{x}$ and get

$$m\dot{x}\ddot{x} + m\omega_0^2\dot{x}x = -2Z\omega_0m\dot{x}^2 \tag{3.4}$$

If I differentiate equation (3.3) and compare the result with equation (3.4) I see that

$$\begin{aligned} \frac{\mathrm{d}E}{\mathrm{d}t} &= m\ddot{x}\dot{x} + m\omega_0^2\dot{x}x = -2\omega_0Zm\dot{x}^2 \\ &= -2Z\omega_0 \times 2 \times (\text{kinetic energy}) \end{aligned} \tag{3.5}$$

The rate at which the damped oscillator loses energy is proportional to its instantaneous kinetic energy.

3.4 The energy of a weakly damped oscillator

I now have to restrict myself to the case of very weak damping, that is, weak enough that it makes sense to average the total energy over a large number of cycles during which time the amplitude decays only by an insignificant amount. In that case there is always a well-defined *average total energy*,

$$\bar{E} = \text{average kinetic energy} + \text{average potential energy}$$

I now appeal to the *virial theorem*† which states in the case of the harmonic oscillator for which $n = 2$

$$\text{average kinetic energy} = \text{average potential energy}$$

and I can use equation (3.5) to write

$$\frac{\mathrm{d}\bar{E}}{\mathrm{d}t} = -2Z\omega_0\bar{E}$$

† Roughly stated, if the potential energy of a system depends on its coordinate like x^n then the average kinetic energy is equal to $n/2$ times the average potential energy. In this case we are abusing the virial theorem because we neglect the damping force in comparison with the spring force. The damping force does not vary like x^n; indeed it is not even a *conservative force* in that it cannot be obtained as the gradient of any potential energy—it is in fact dependent on the mass's speed, $\dot{x}$.

which leads to

$$\bar{E} = E_0 e^{-2Z\omega_0 t}$$

This means that the energy is dissipated (into heating up the viscous medium) at a rate such that the energy is reduced by an amount e in a time $\tau = 1/2Z\omega_0$ We would normally call τ the *time constant* of this process.

Over one period of oscillation the amount of energy dissipated is

$$\Delta\bar{E} = -\frac{\mathrm{d}\bar{E}}{\mathrm{d}t} \times \text{period}$$

The period is $2\pi/\omega_D$ and so

$$\Delta\bar{E} = 2Z\omega_0 \frac{2\pi}{\omega_D} \bar{E}$$

We call the quantity

$$\frac{\Delta\bar{E}}{\bar{E}} = S$$

the *specific damping capacity*. Clearly

$$S = 4\pi Z \frac{\omega_0}{\omega_D} \approx 4\pi Z$$

The *quality factor*, Q, is defined to be

$$\begin{aligned} Q &= 2\pi \, \frac{\text{the total energy}}{\text{energy dissipated over the next cycle}} \\ &= \frac{2\pi}{S} = \frac{1}{2}\frac{1}{Z}\frac{\omega_D}{\omega_0} \approx \frac{1}{2Z} \end{aligned}$$

The *logarithmic decrement*, δ, is defined to be

$$\delta = \ln(\text{ratio of amplitudes of successive cycles})$$

If a peak occurs at t_1 with a displacement x_1 and the next peak occurs at t_2 with displacement x_2 then

$$\frac{x_1}{x_2} = \frac{x_m\, e^{-Z\omega_0 t_1} \cos\omega_D t_1}{x_m\, e^{-Z\omega_0 t_2} \cos\omega_D t_2}$$

but

$$t_2 = t_1 + \frac{2\pi}{\omega_D} \text{ , the period}$$

so

$$\cos\omega_D t_2 = \cos(\omega_D t_1 + 2\pi) = \cos\omega_D t_1$$

and therefore

$$\frac{x_1}{x_2} = e^{2\pi Z\omega_0/\omega_D}$$

and finally

$$\begin{aligned} \delta &= \ln\frac{x_1}{x_2} = 2\pi Z \frac{\omega_0}{\omega_D} \\ &= \frac{1}{2} S \approx 2\pi Z \end{aligned}$$

I have defined for you a whole bunch of parameters belonging to a damped oscillator. These are

— the natural frequency, ω_0

— the damped frequency, ω_D

— the damping coefficient, b

— the damping ratio, Z

— the specific damping capacity, S

— the quality factor, Q

— the logarithmic decrement, δ

It will be useful now, for revision purposes, but more importantly for general reference since the damped oscillator is ubiquitous in physics and engineering, to write a table of these for yourself showing their units and the relations between them. Use the approximate relations I have given as these are correct in as much as for most usual systems $\omega_D \approx \omega_0$.

Be warned, some authors, especially treating mechanical systems such as viscoelasticity and internal friction use a captial Δ for the logarithmic decrement. It gets worse because they then use a lower case δ for a different but related quantity. If the damping is small they write

$$\text{“}\Delta = \pi \tan \delta\text{”}$$

and $\tan \delta$ is called the loss tangent, the word "loss" referring to energy loss through dissipation. Loss tangent is also a widely used material property of semiconductors and dielectrics.

As a first year student and going on into higher level study you have to come to terms with the fact that there is *no agreed choice of symbols* among physicists for physical quantities (except possibly x for position and t for time—but many use q for position!) Even worse, there is no agreement over choice of units and this makes electrodynamics particularly troublesome; but on that subject my advice is *use only SI units, always.*

Lecture 4

4.1 Resonance

The damped oscillator will eventually come to rest and is of little interest except in the study of transients and in the establishment of system parameters, as we did in Lecture 3. Of much greater interest is the forced, or driven, oscillator. Here you take a damped or free oscillator with a natural frequency ω_0 and you force it to vibrate with a frequency ω. As ω approaches ω_0 *resonance* is reached; and this is a ubiquitous phenomenon in both physics and engineering. Examples are nuclear magnetic resonance and magnetic resonance imaging; Raman spectroscopy; infra-red spectroscopy; a.c. dielectric loss in capacitors; tuning of electronic circuits; microwave absorption; Mössbauer spectroscopy.

We now add a third force to F_{spring} and F_{damping}, namely an oscillating driving force,

$$F_{\text{driven}} = F_0 \sin \omega t$$

After transient effects have died away the oscillator must be vibrating at the driven frequency ω and may or may not be in phase with the driving force. Our task is to determine the amplitude of the final steady state and its phase. We might guess that the closer is the driven frequency to the natural frequency the greater will be the amplitude—think of the opera singer and the wine glass. It may be easier to see by examining a counter example: try pushing someone on a swing at a different frequency to the natural one; you'll be pushing when she's swinging back and you'll prevent her from swinging—her amplitude will suffer. But push in time to her swinging and she'll swing ever higher!

The differential equation that we'll need to solve is again Newton's second law,

$$\text{mass} \times \text{acceleration} = F_{\text{spring}} + F_{\text{damping}} + F_{\text{driven}}$$

That is,

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = F_0 \sin \omega t - b\frac{\mathrm{d}x}{\mathrm{d}t} - kx$$

or, by comparison with equation (3.2), Lecture 3,

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2Z\omega_0\frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2 x = \frac{F_0}{m}\sin \omega t$$

in which the left hand side is expressed only in terms of the natural frequency of the oscillator ω_0 and the damping ratio,

$$Z = \frac{1}{2}\frac{b}{m\omega_0}$$

A solution is†

$$x = A \sin(\omega t - \phi)$$

† Here and in Lecture 3, I am stating the solutions without proof. To some of you this will be unsatisfactory and for your benefit I have prepared a detailed set of notes giving a full mathematical treatment of the general forced damped oscillator and the problem of resonance. These notes are available to you online at `KEATS`.

A is called the *dynamic amplitude* and is given by

$$A = A_s \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2 + \left(2Z\frac{\omega}{\omega_0}\right)^2}} \tag{4.1}$$

in which A_s is the so called *static amplitude*,

$$A_s = \frac{F_0}{k}$$

The static amplitude is the amount of extension of the spring produced by the amplitude of the driving force, F_0. The phase difference between the oscillator and the driving force is

$$\phi = \arctan \frac{2Z\frac{\omega}{\omega_0}}{1 - \frac{\omega^2}{\omega_0^2}}$$

That all gets put together into a resonance curve.

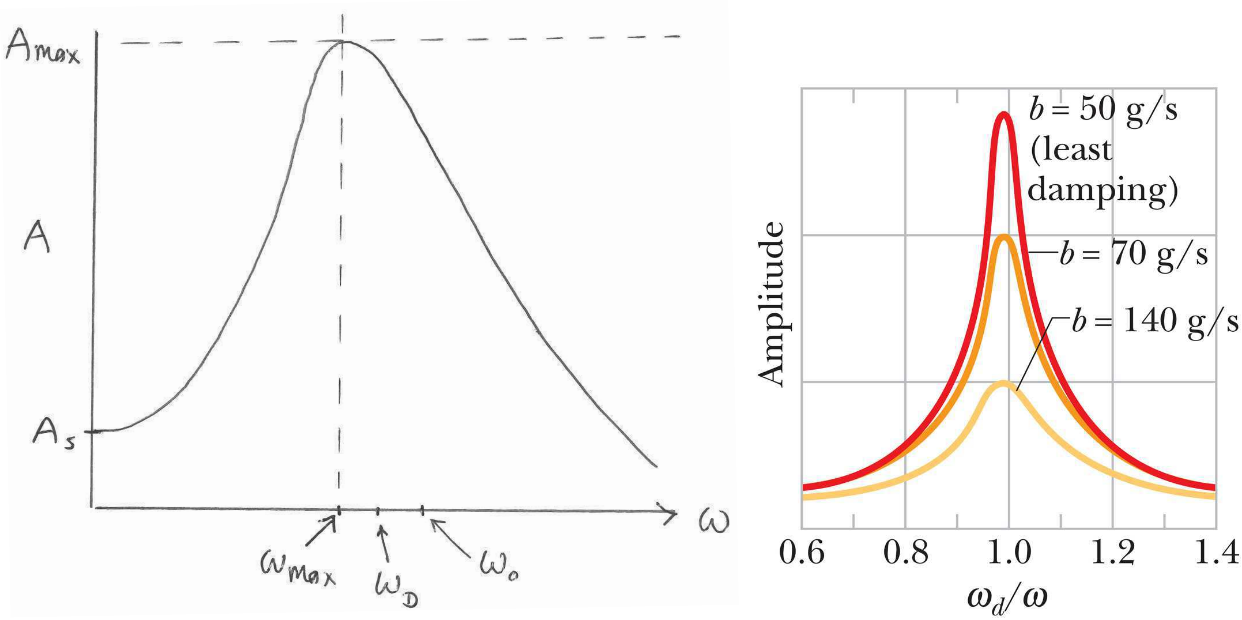


FIGURE 4–1

The maximum amplitude, $A_{\max}$, is found by finding the frequency that makes the denominator in (4.1) the smallest. By the usual means (that is, differentiate with respect to ω and set that to zero) we find

$$\omega_{\max} = \omega_0\sqrt{1 - 2Z^2}$$

and putting this back into (4.1) we get

$$A_{\max} = A_s\frac{m\omega_0}{b} = \frac{F_0}{b\omega_0}$$

It is important to understand that the resonant frequency of the driven, damped oscillator is neither the natural frequency ω_0 nor the damped frequency ω_D, which we defined in Lecture 3, page 6. In fact there are *three* frequencies belonging to the underdamped resonance problem, these are,

$$\omega_0 > \omega_D = \omega_0\sqrt{1-Z^2} > \omega_{\max} = \omega_0\sqrt{1-2Z^2}$$

Now

$$\frac{A_{\max}}{A_s} = \frac{1}{2Z} = \frac{m\omega_0}{b} = Q$$

is the *quality factor* of the oscillator (see Lecture 3, page 7). It is a property of the damped oscillator and does not depend on the driving force, F_{driven}, or the driving frequency, ω.

4.2 The Hertzian dipole

We now study what is perhaps the most important and commonly occuring oscillator in physics. I think that I can safely say that the origin of all electromagnetic radiation is the wiggling of little charges, usually electrons. To look at this in detail requires advanced physics, so I will just show you the qualitative features and I'll call upon a few concepts that you don't come to until later in the course. I will also have to give you some mathematical formulas without proof; but everything can be found in the textbook by Griffiths, which is on your reading list.

Consider, first, the object sketched in figure 4–2; it is an *electric dipole* (see Lecture 8)

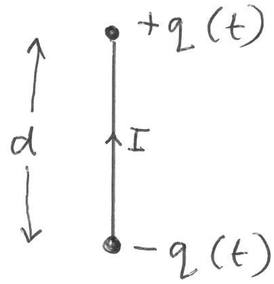


FIGURE 4–2

Two equal and opposite charges are separated by a conducting wire of length d. The magnitudes of the charges are varying sinusoidally in time in such a way that the total charge on the object is always zero.

$$q(t) = q_0 \sin \omega t$$

and as the charges change from positive to negative a current flows up and down the wire as sketched in figure 4–3.

The dipole emits electromagnetic radiation of frequency $f = \omega/2\pi$ and wavelength $\lambda = c/f$ where c is the speed of light. To illustrate in more detail, the radiation is emitted in the geometry shown in figure 4–4.

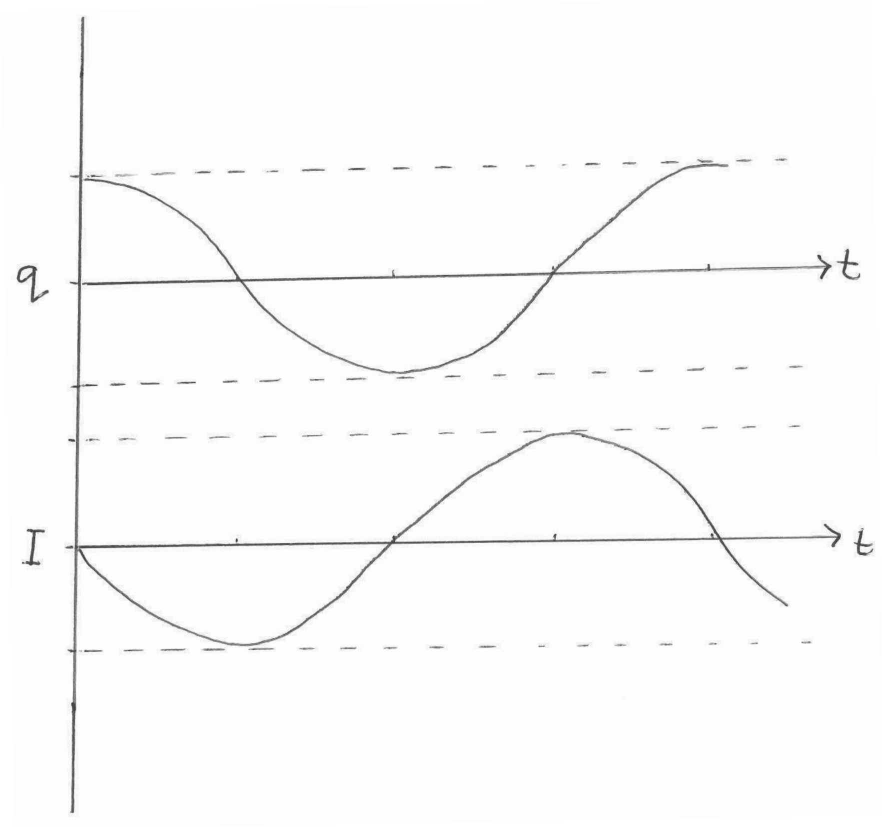


FIGURE 4–3

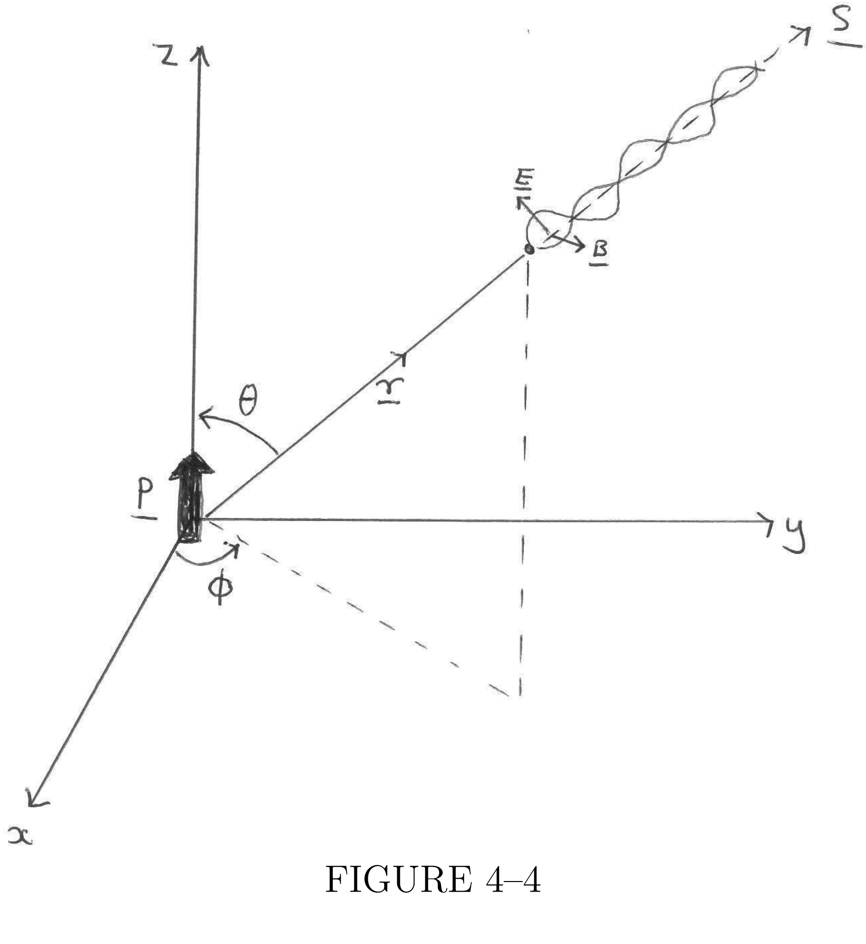


FIGURE 4–4

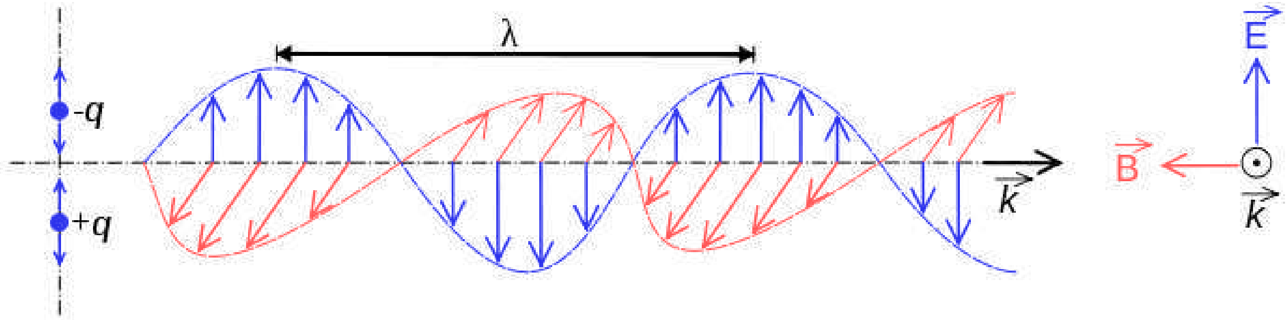


FIGURE 4–5

The dipole has electric dipole moment vector $\mathbf{p} = q\mathbf{d}$, and along a vector direction $\hat{\mathbf{r}}$ having polar and azimuthal angles θ and φ radiation is emitted having a *Poynting vector* $\mathbf{S}$. This vector is defined as

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E} \times \mathbf{B}$$

Here, $\mu_0 = 4\pi \times 10^7$ N amp^{-1} is a fundamental constant (see Lecture 15); $\mathbf{E}$ is the electric field vector (see Lecture 8) and $\mathbf{B}$ is the magnetic field vector (see Lecture 14). These fields are at right angles to each other and in phase. *This is electromagnetic radiation.*

The total power emitted by the dipole is

$$W = \frac{1}{4\pi\epsilon_0}\,\frac{1}{3}\,\frac{p_0^2\omega^4}{c^3} \qquad [\mathrm{J\ s^{-1}}]$$

where $p_0 = q_0 d$ and

$$\mu_0\epsilon_0 = \frac{1}{c^2}$$

Note how the power varies as the *fourth* power of the frequency, that is, the inverse fourth power of the wavelength of the radiation This "fourth power law" also crops up in Rutherford and Rayleigh scattering and lies at the heart of the explanation of why the sky is blue.

When constructed on a large scale this is an *antenna*. For example a half-wave antenna looks like figure 4–6.

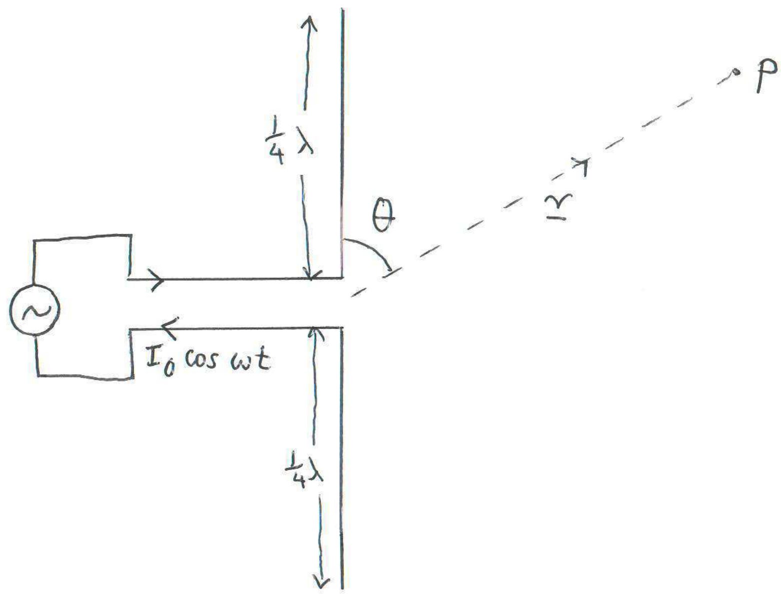


FIGURE 4–6

The flux of radiant energy at point P at which $r \gg \lambda$ is

$$S = \frac{1}{4\pi\epsilon_0} \frac{1}{2\pi c} \frac{I_0^2}{r^2} \frac{\cos^2\left(\frac{1}{2}\cos\theta\right)}{\sin^2\theta}$$

The power output is

$$W = 73.1\, I_{\rm rms}^2 \quad \text{[watt]}$$

where $I_{\rm rms}$ is the root mean square of the a.c. current.

An object which is similar to the "antenna" dipole that I have described is the *Hertzian dipole*, named after Heinrich Hertz (who gave his name to the unit of frequency) and studied by Sir Joseph Larmor in 1897. In this dipole the equal and opposite charges are fixed, but the upper charge is wiggling up and down as if the two charges were connected by a linear spring.

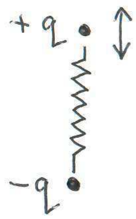


FIGURE 4–7

The power radiated by the Hertzian dipole is

$$W = \frac{1}{4\pi\epsilon_0} \frac{2}{3} \frac{q^2 a^2}{c^3} = \frac{\mu_0}{4\pi} \frac{2}{3} \frac{q^2 a^2}{c} \quad \text{in S.I. units [watt]}$$

$$\left(\text{Larmor actually wrote } W = \frac{2}{3} \frac{q^2 a^2}{c} \text{ in magnetic c.g.s. units}\right)$$

Here, a is the acceleration of the moving charge. After averaging over one cycle, it turns out that the power radiated is the *same* for the "antenna" and the Hertzian dipole. Therefore the Larmor radiation formula is really very general.

4CCP1501 Simple Harmonic Motion

1. Free oscillation

We are to solve Newton's second law, force = mass × acceleration, as a differential equation,

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx$$

which we write

$$\ddot{x} = -\omega_0^2 x \tag{1}$$

by using two dots to indicate a *second derivative with respect to time.* We will use one dot to indicate the first derivative. We have also combined the two constants, m, the mass and k, the spring constant, to define an *angular frequency*,

$$\omega_0^2 = \frac{k}{m}$$

We're not mathematicians, we just want a solution of this thing; so try $x = Ae^{st}$. Then by simple differentiating, we have

$$x = Ae^{st}\ ;\quad \dot{x} = sAe^{st}\ ;\quad \ddot{x} = s^2Ae^{st}$$

We only have to put this back into (1) to see that

$$s^2Ae^{st} + \omega_0^2Ae^{st} = 0 \longrightarrow s^2 + \omega_0^2 = 0 \longrightarrow s = \pm\mathrm{i}\omega_0$$

So we have two solutions:

$$x = Ae^{\mathrm{i}\omega_0\mathrm{t}} \text{ and } x = Ae^{-\mathrm{i}\omega_0\mathrm{t}}$$

The theory of second order, linear differential equations tells us that the most general solution is a linear combination of the two solutions with two arbitrary coefficients, that we will call A_1 and A_2:

$$\begin{aligned}
x &= A_1e^{\mathrm{i}\omega_0\mathrm{t}} + A_2e^{-\mathrm{i}\omega_0\mathrm{t}} \\
&= (A_1 + A_2)\cos\omega_0 t + \mathrm{i}(\mathrm{A}_1 - \mathrm{A}_2)\sin\omega_0\mathrm{t} \\
&= A\cos\omega_0 t + B\sin\omega_0 t && \text{(a)} \\
&= C\cos\phi\sin\omega_0 t + C\sin\phi\cos\omega_0 t && \text{(b)} \\
&= C\sin(\omega_0 t + \phi)
\end{aligned}$$

In going from line (a) to line (b) I have changed from the variables A and B to variables C and ϕ by making these two definitions,

$$A = C\sin\phi \text{ and } B = C\cos\phi$$

because then I can use the usual formula for $\sin(a + b)$ to arrive at the last line.

Now what we have is

$$x = C\sin(\omega_0 t + \phi)$$
$$\dot{x} = v = C\omega_0\cos(\omega_0 t + \phi)$$

To fix the, up to now arbitrary, constants requires us to know "boundary conditions." Let's suppose that at $t = 0$, $x = x_0$, say, and $v = v_0$, the initial velocity. These conditions give,

$$x_0 = C\sin\phi \ , \quad \sin\phi = \frac{x_0}{C} \tag{c}$$
$$v_0 = C\omega_0\cos\phi \ , \quad \cos\phi = \frac{v_0}{C\omega_0} \tag{d}$$

Now, square and add (c) and (d),

$$C = \sqrt{x_0^2 + \frac{v_0^2}{\omega_0^2}}$$

and divide (c) by (d)

$$\phi = \arctan\frac{x_0\omega_0}{v_0}$$

Finally, if we start off the oscillator at $t = 0$ with $v_0 = 0$ and $x = x_m$, for example we pull out the spring to maximum deflection, x_m, hold it still ($v_0 = 0$) and let it go; then the solution is

$$x = x_m\sin(\omega_0 t + \frac{1}{2}\pi) = x_m\cos(\omega_0 t)$$

2. Damping

To the differential equation (1), which is after all Newton's second law—force equals mass times acceleration—we add an additional force, $-b\dot{x}$. This force is *proportional to the velocity*, which is what you'd expect. Try swimming in syrup: the faster you swim the bigger is the drag, or viscous, force. So now we need to solve the differential equation

$$m\ddot{x} + b\dot{x} + kx = 0$$

which we re-write as

$$\ddot{x} + \frac{b}{m}\dot{x} + \omega_0^2 x = 0$$

We define a new constant, Z, such that

$$\frac{b}{m} = 2Z\omega_0$$

is the *frictional force per unit mass and unit speed.* Now our differential equation is

$$\ddot{x} + 2Z\omega_0\dot{x} + \omega_0^2 x = 0$$

As before we try

$$x = Ae^{st} \ ; \quad \dot{x} = sAe^{st} \ ; \quad \ddot{x} = s^2Ae^{st}$$

and so

$$s^2 + 2Z\omega_0 s + \omega_0^2 = 0$$

leads to

$$s = \omega_0 \left(-Z \pm \sqrt{Z^2 - 1}\right) \tag{2}$$

and the general solution must be

$$x = A_1 e^{st} + A_2 e^{-st} \tag{3}$$

Critial damping is defined as the condition $Z = 1$. For that case we define

$$b_{\text{crit}} = 2m\omega_0 = 2\sqrt{mk}$$

and we give a name to Z by

$$\frac{b}{b_{\text{crit}}} = Z$$

being called the *damping factor*, or *damping ratio.*

Underdamping is the condition $Z < 1$ or $b < b_{\text{crit}}$. This is usually the most interesting case, and for which

$$Z^2 - 1 < 0$$

meaning that there are two roots to (2), namely,

$$s_1 = \omega_0 \left(-Z + \mathrm{i}\sqrt{1 - \mathrm{Z}^2}\right)$$
$$s_1 = \omega_0 \left(-Z - \mathrm{i}\sqrt{1 - \mathrm{Z}^2}\right)$$

and then (3) is

$$x = e^{-Z\omega_0 t}\left(A_1 e^{\mathrm{i}\sqrt{1-\mathrm{Z}^2}\omega_0 \mathrm{t}} + A_2 e^{-\mathrm{i}\sqrt{1-\mathrm{Z}^2}\omega_0 \mathrm{t}}\right)$$

We then simplify this in the same manner as for equations (a) and (b):

$$\begin{aligned} x &= Ce^{-Z\omega_0 t}\sin\left(\sqrt{1-Z^2}\omega_0 t + \phi\right) \\ &= Ce^{-\alpha t}\sin(\omega_D t + \phi) \end{aligned}$$

where

$$\alpha = \frac{1}{2}\frac{b}{m} = Z\omega_0$$

is called the *damping constant*, and

$$\omega_D = \omega_0\sqrt{1 - Z^2} = \omega_0\sqrt{1 - \frac{1}{4}\frac{b^2}{mk}} < \omega_0$$

is the *damped frequency.*

Again, if at $t = 0$, $x = x_m$ and $v = 0$, the solution associated with these boundary conditions is

$$\begin{aligned} x &= x_m\, e^{-\alpha t}\, \sin\left(\omega_D t + \frac{1}{2}\pi\right) \\ &= x_m\, e^{-\alpha t}\, \cos\omega_D t \end{aligned}$$

which is the result I give you on page 6 of Lecture 3.

3. Driven oscillators

In real life we are less interested in an oscillator that is oscillating at its natural frequency, ω_0, or its natural damped frequency, ω_D, than in the behaviour of an undamped or damped oscillator when we choose to drive it at some frequency, ω, that *we* choose. Situations of this phenomenon are ubiquitous in physics and engineering. Try and write down some half a dozen examples of your own.

3.1 Undamped driven oscillator

The oscillator is driven by a periodic force of angular frequency ω and amplitude F_0. That means we have one more force to add in to Newton's second law, namely

$$F = F_0 \sin\omega t$$

and force = mass × acceleration now reads

$$m\ddot{x} = F_0 \sin\omega t - kx \tag{4}$$

Eventually the oscillator has no choice but to vibrate at the frequency of the driving force, whether it likes it or not, so we must have,

$$\begin{aligned} x &= A\sin\omega t \\ \dot{x} &= A\omega\cos\omega t \\ \ddot{x} &= -A\omega^2 \sin\omega t \end{aligned}$$

Equation (4) now reads

$$-mA\omega^2 \sin\omega t + kA\sin\omega t = F_0 \sin\omega t$$

That is,

$$\begin{aligned} A &= \frac{F_0}{k - m\omega^2} = \frac{F_0/k}{1 - \frac{\omega^2}{\omega_0^2}} \\ &= \frac{A_s}{1 - \frac{\omega^2}{\omega_0^2}} \end{aligned}$$

using

$$\omega_0 = \sqrt{\frac{k}{m}}$$

the natural frequency of the undamped oscillator. We call A_s the *static amplitude* and we call A the *dynamic amplitude*; their ratio is called the *magnification factor*,

$$D_s = \frac{A}{A_s} = \left(1 - \frac{\omega^2}{\omega_0^2}\right)^{-1}$$

If the driving frequency is less that the natural frequency the magnification factor is positive and the displacement is in phase with the driving force. Conversely if $\omega > \omega_0$, $D_s < 0$. An amplitude cannot be negative, so we'll have instead, for this case, to use the solution

$$x = -A \sin \omega t$$

which implies a phase difference of π (180°) between the displacement and the driving force. Thirdly, if $\omega = \omega_0$, $D_s \to \infty$ and we have *resonance*. In real life this never happens as there is always damping. But interesting things *do* happen when we drive an oscillator at a frequency close to its natural one.

3.1 Damped driven oscillator

Now we include the velocity dependent damping force into equation (4):

$$m\ddot{x} = F_0 \sin \omega t - b\dot{x} - kx$$

or

$$m\ddot{x} + b\dot{x} + kx = F_0 \sin \omega t \tag{4a}$$

Eventually after transients have died away, the oscillator must vibrate at the frequency of the driving force. It may not like it and it will protest unless the driving frequency is close to the natural frequency of the undriven oscillator. Its reluctance to cooperate is reflected in a reduction in amplitude. Nearer to *resonance* the amplitude is large. The so called *resonance curve* or relation between amplitude and driving frequency is what we will be seeking in the mathematical development that follows. The oscillator will necessarily vibrate at the frequency of the driving force, but it will not necessarily be in phase with it. Hence the solution for the amplitude must look like

$$\begin{aligned} x &= A \sin(\omega t - \phi) \\ \dot{x} &= A\omega \cos(\omega t - \phi) \\ \ddot{x} &= -A\omega^2 \sin(\omega t - \phi) \end{aligned}$$

when I plug these into (4a) I get

$$\begin{aligned} m\left[-A\omega^2 \sin(\omega t - \phi)\right] + b\left[A\omega \cos(\omega t - \phi)\right] + kA \sin(\omega t - \phi) &= F_0 \sin \omega t \\ &= F_0 \sin(\omega t - \phi + \phi) \end{aligned}$$

Rearranging this I have

$$\begin{aligned} A\left(k - m\omega^2\right) \sin(\omega t - \phi) &+ Ab\omega \cos(\omega t - \phi) \\ &= F_0\left[\sin(\omega t - \phi)\cos\phi + \cos(\omega t - \phi)\sin\phi\right] \end{aligned}$$

Now, equate the coefficients of $\sin(\omega t - \phi)$ and $\cos(\omega t - \phi)$ and obtain

$$\begin{aligned} Ab\omega &= F_0 \sin\phi \\ A\left(k - m\omega^2\right) &= F_0 \cos\phi \end{aligned}$$

We square and add these two, recalling that $\sin^2\phi + \cos^2\phi = 1$,

$$F_0^2 = A^2\left[\left(k - \omega^2\right) + b^2\omega^2\right]$$

which means that we have, for the *dynamic amplitude*,

$$\begin{aligned} A &= \frac{F_0}{\sqrt{\left(k - m\omega^2\right)^2 + b^2\omega^2}} \\ &= \frac{F_0/k}{\sqrt{\left(1 - \frac{m\omega^2}{k}\right)^2 + \frac{b^2\omega^2}{k^2}}} \end{aligned}$$

We also divide our two equations to find the *phase difference*, or phase angle, ϕ, between the oscillator and its driving force,

$$\tan\phi = \frac{b\omega}{k - m\omega^2}$$

We can simplify the formulas for A and ϕ using these definitions that we have encountered already in these notes,

$$\omega_0 = \sqrt{\frac{k}{m}}\ ,\quad b = 2mZ\omega_0\ ,\quad A_s = \frac{F_0}{k}$$

We also define the *frequency ratio*,

$$r = \frac{\omega}{\omega_0}$$

Then the magnification factor is

$$D_s = \frac{A}{A_s} = \frac{1}{\sqrt{\left(1 - r^2\right)^2 + \left(2rZ\right)^2}} \tag{5}$$

and the phase angle is

$$\phi = \arctan\frac{2rZ}{1 - r^2} \tag{6}$$

What is the frequency, ω_{max}, say, that gives us the greatest amplitude? Or to put the question another way, what is the *resonant frequency*? We need to minimise the denominator in (5); we do this in the usual way by setting its first derivative with respect to r equal to zero and solving for r which will then give us $\omega_{\text{max}}/\omega_0$.

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[\left(1 - r^2\right)^2 + \left(2rZ\right)^2\right] = 0$$

leads to

$$\omega_{\max} = \omega_0\sqrt{1-2Z^2} \tag{7}$$

which is neither ω_0, nor $\omega_D = \omega_0\sqrt{1-Z^2}$.

What is the *maximum ampltitude*; $A_{\max}$, say? Put (7) into (5) and neglect Z^4 when compared to Z^2. We find

$$\frac{A_{\max}}{A_s} = \frac{1}{2Z} = \frac{m\omega_0}{b} \approx Q$$

which is the "quality factor", and using $A_s = F_0/k$ and $\omega_0^2 = k/m$ we get

$$A_{\max} = \frac{F_0}{b\omega_0}$$

On page 9 (below) are two graphs I've taken from `wikipedia` showing a set of resonance curves and phase angles for a driven damped oscillator. On the abscissa is plotted the frequency ratio, r. They use the phrase "amplification ratio" for the magnification factor and have used the symbol ζ for the damping factor, Z. The first is essentially a plot of equation (5). In the second, note how in the case of the undamped forced oscillator there is an abrupt change from in phase to 180° out of phase as r goes through one, as we discuss on page 5 of these notes. Note how the frequency $\omega_{\max}$ is always smaller than the natural frequency ω_0 but appears to approach it as the peak becomes narrower, that is, the damping becomes less.

There are three interesting cases.

(i) If $r \ll 1$ the driving frequency is much smaller than the natural frequency of the oscillator,

$$\omega \ll \omega_0$$

Then the dynamic amplitude is close to the static amplitude,

$$A \approx A_s$$

and the phase difference is

$$\phi \approx \arctan 0 = 0$$

so the displacement and force are in phase.

(ii) If $r \approx 1$ then

$$\omega \approx \omega_0$$

and

$$\frac{A}{A_s} \approx \frac{1}{2Z} \approx Q \text{ , the quality factor}$$

Also,

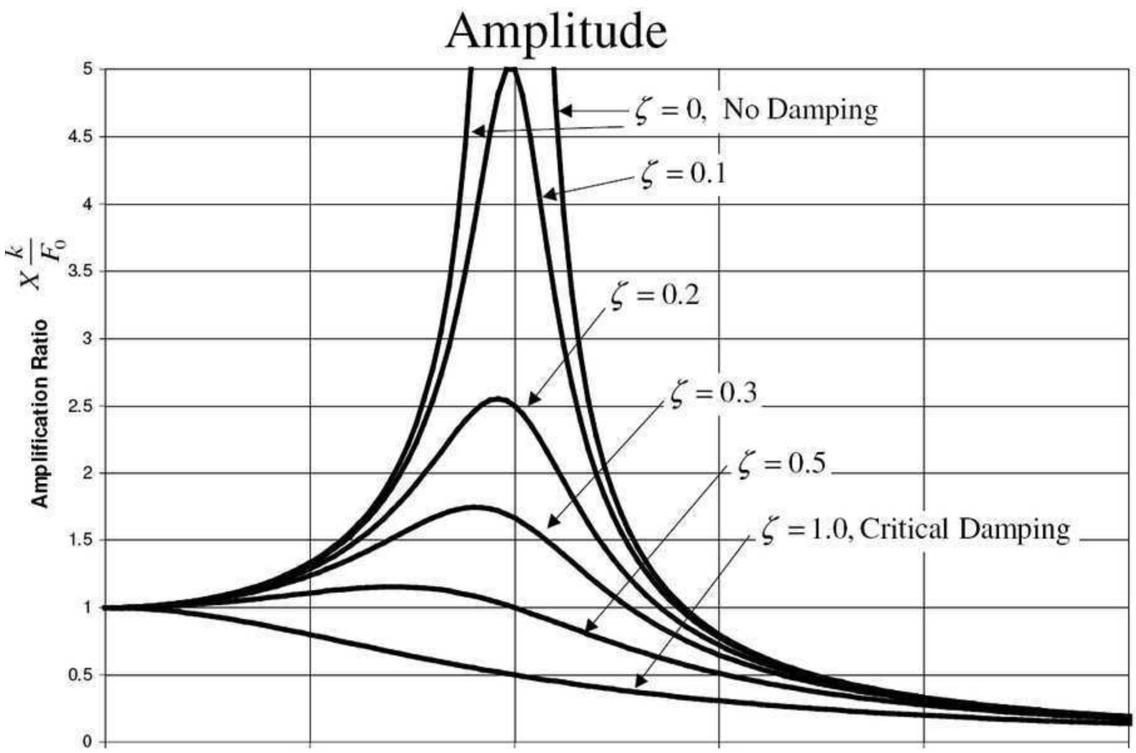
$$\phi \approx \arctan \infty = \frac{1}{2}\pi$$

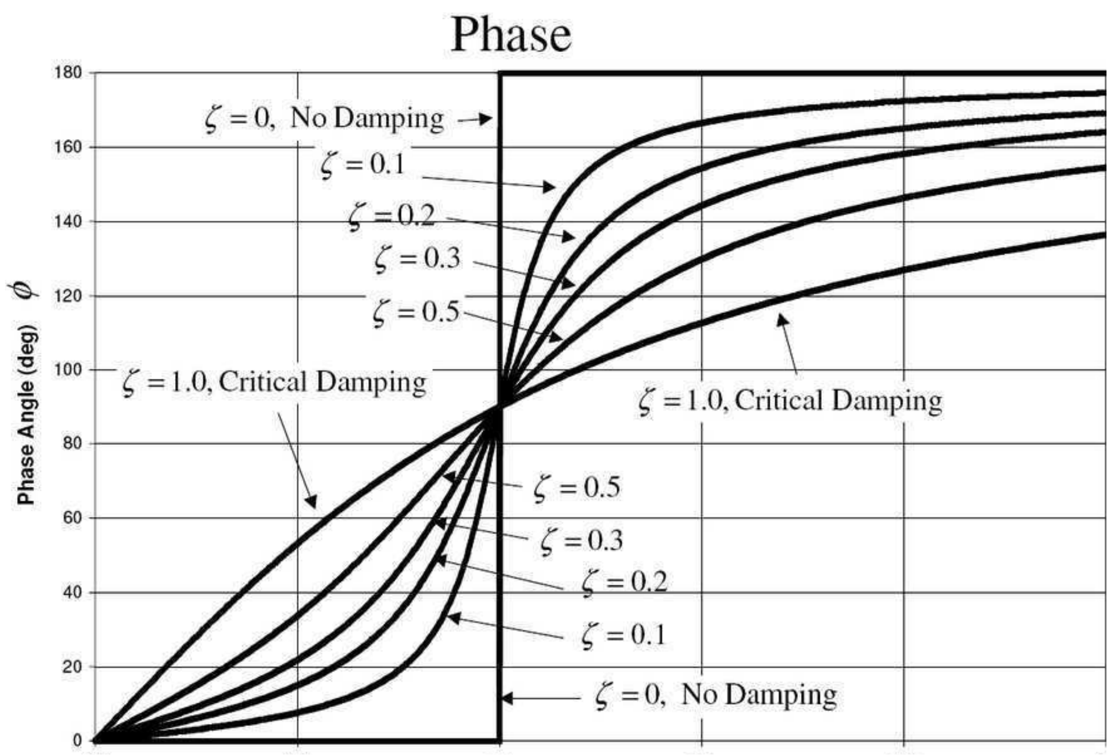
so the displacement and force are out of phase by 90°.

(iii) If $r \gg 1$, then $\omega \gg \omega_0$ and therefore

$$\frac{A}{A_s} \propto \frac{\omega_0^2}{\omega^2} = \frac{1}{r^2}$$

which is the shape of the high frequency tail of the resonance curve. The displacement and force are out of phase by 180°, for the same reason as given on page 6 for the driven undamped oscillator.


Amplitude
Amplification Ratio $X\frac{k}{F_0}$
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
$\zeta = 0$, No Damping
$\zeta = 0.1$
$\zeta = 0.2$
$\zeta = 0.3$
$\zeta = 0.5$
$\zeta = 1.0$, Critical Damping


Phase
Phase Angle (deg) ϕ
180
160
140
120
100
80
60
40
20
0
$\zeta = 0$, No Damping
$\zeta = 0.1$
$\zeta = 0.2$
$\zeta = 0.3$
$\zeta = 0.5$
$\zeta = 1.0$, Critical Damping
$\zeta = 1.0$, Critical Damping
$\zeta = 0.5$
$\zeta = 0.3$
$\zeta = 0.2$
$\zeta = 0.1$
$\zeta = 0$, No Damping

Lecture 5

5.1 Huygens's principle

You've seen in outline what an electromagnetic wave really is: electric and magnetic fields oscillating in the vacuum in phase and propagating with a phase velocity c. Optics is all about how light behaves in various media. The geometric optics is fine if the objects we deal with are much larger than the wavelength of the radiation (about 5/1000 mm in the case of yellow light). But ray optics cannot describe *diffraction* (light going around corners) and *interference*. For that, we use wave optics. This is also an idealisation—to get some proper insight, read *QED—The Strange Theory of Light and Matter*, by Richard Feynman.

We'll start with *Huygens's principle*. Light does not travel in rays. Light propagates as a *wavefront*. Every point acts as a source of spherical waves. For a plane wavefront this looks like this.

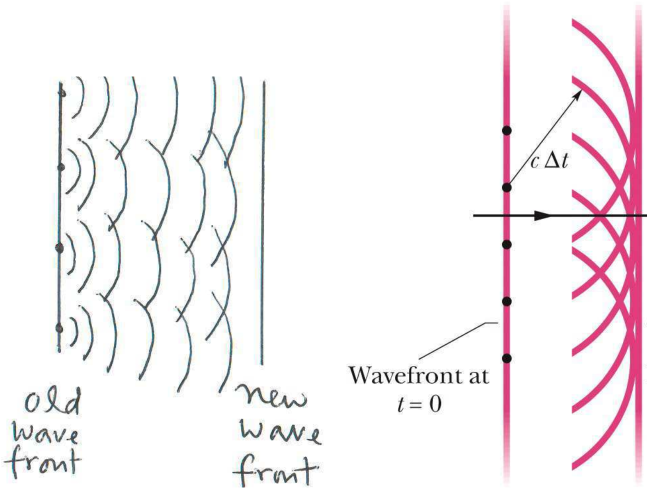


FIGURE 5–1

I've only drawn three point sources; there is actually an infinity of them. In the case of a spherical wavefront, I can draw this.

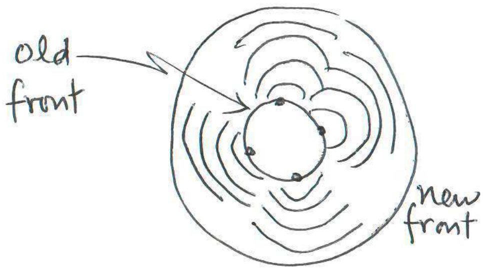


FIGURE 5–2

Again I only draw a few out of the infinity of new fronts. You can then see how light can *diffract* at a sharp edge.

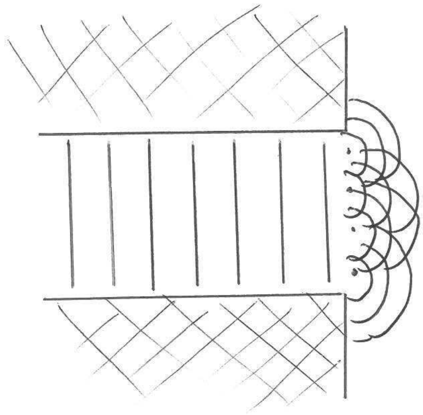


FIGURE 5–3

You can use Huygens's principle to derive the laws of reflection and refraction, but I won't do that here—you can find it in your textbook.

5.2 Young's slits

We consider a plane wave front of monochromatic light falling onto two closely spaces narrow slits. Because of Huygens's principle the light diffracts around the slits. If the apparatus is on a large scale, we get a situation described in ray optics:

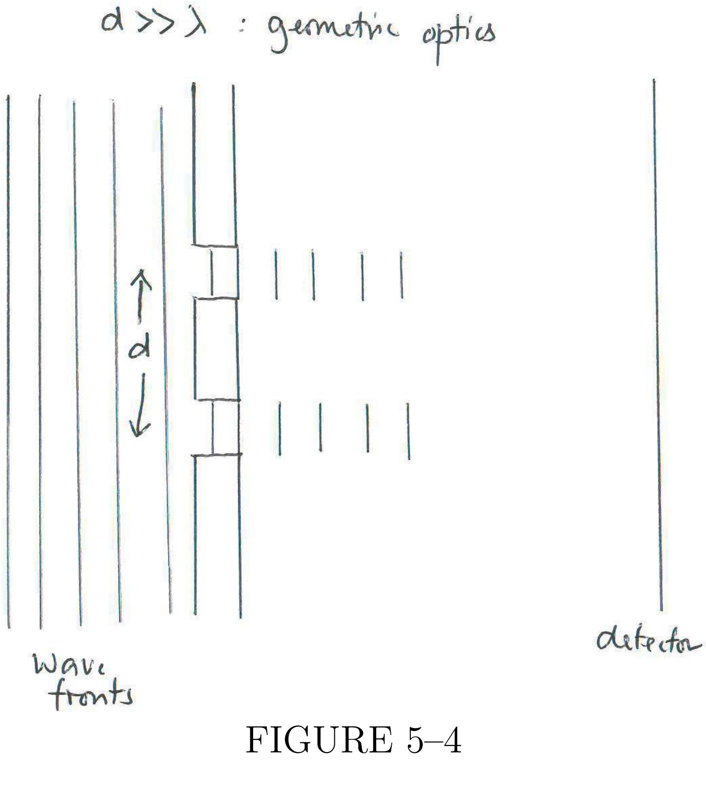
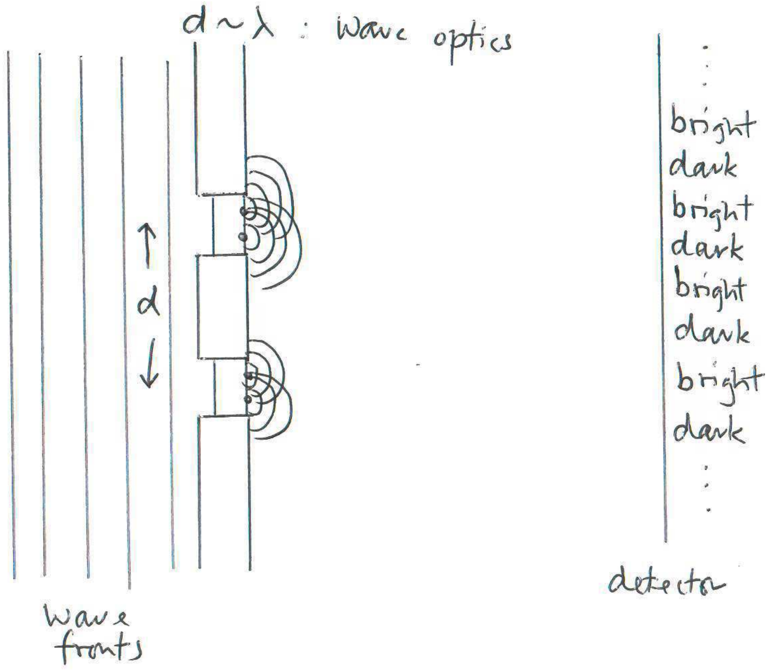


FIGURE 5–4

But in wave optics, which applies if the width and separation of the slits is on the same length scale as the wavelength of the light, there is diffraction, like this,



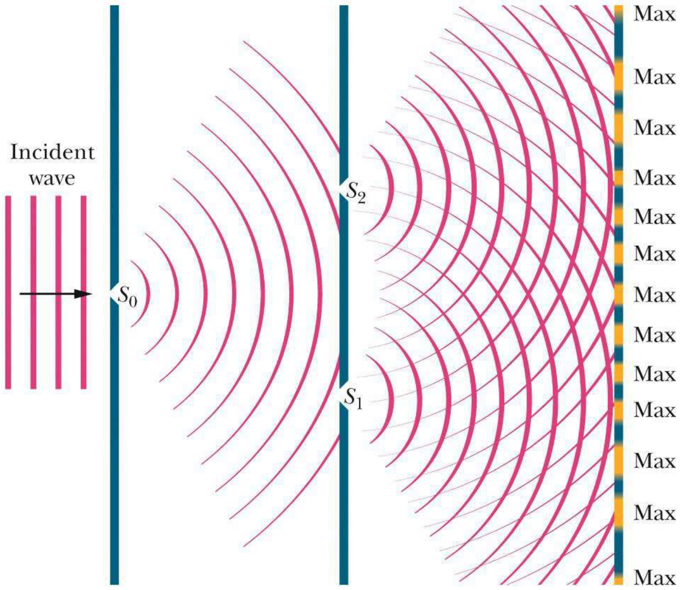


FIGURE 5–5

and there will appear a pattern on the detector screen of not just two spots a distance d apart as in figure 5–4, but an *interference pattern*. We choose a point P at the detector and ask, what will be the intensity of the light? Or, equivalently, do the light waves arriving at P interfere constructively or destructively? It is a matter in all of these problems of interference and diffraction to find the *path difference* between two waves. If this path difference is an *odd multiple* of $\frac{1}{2}\lambda$ then the waves are $\frac{1}{2}\pi$ (180°) out of phase and there is destructive interference. But if the path difference is any *integer multiple* of λ, say $m\lambda$. then there is maximum constructive interference and the image is brightest. We calculate the path difference as in figures 5–6 and 5–7.

You can see from the figure that the path difference is

$$\Delta L = r_1 - r_2 = d \sin\theta$$

so we have maxima in the light intensity on the detector screen when

$$d \sin\theta = m\lambda\ , \quad m = 0, 1, 2\ldots \qquad \longleftarrow \text{bright}$$

and darkness when

$$d \sin\theta = \left(m + \frac{1}{2}\right)\lambda\ , \quad m = 0, 1, 2\ldots \qquad \longleftarrow \text{dark}$$

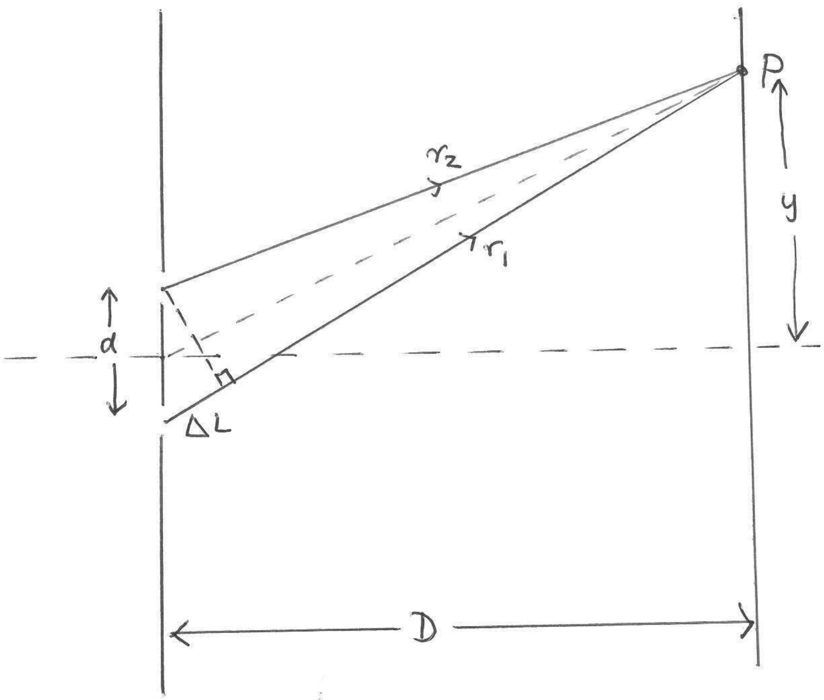


FIGURE 5–6

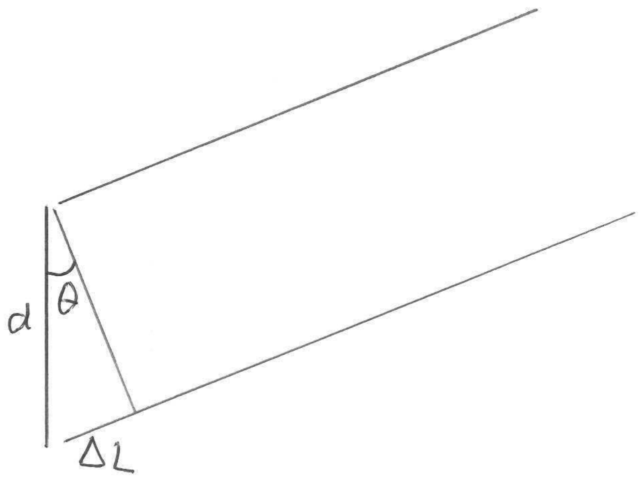


FIGURE 5–7

We can calculate the *intensity* of the image on the screen because it is proportional to the amplitude squared. Which amplitude, the electric or the magnetic field strength? It doesn't matter. Let's write the wavefunction as

$$E = E_0 \sin \omega t$$

in which E is the magnitude of the electric field. Now at any point on the screen two waves will combine to form an image and their phases will differ by an amount depending on their path difference. Now a path difference of zero corresponds to a phase difference of zero, and a path difference of λ results in a phase difference of 2π. Both are conditions for maximum constructive interference and we can interpolate between these two cases and expect to find this very useful relation between path difference and phase difference

$$\frac{\Delta L}{\lambda} = \frac{\phi}{2\pi} \qquad \text{in the range } [0, 1]$$

If the two combining waves are

$$E_1 = E_0 \sin \omega t \quad \text{and} \quad E_2 = E_0 \sin(\omega t + \phi)$$

we use the superposition principle to add the wavefunctions to find the total electric field strength when our two waves combine at point P on the screen. We get

$$\begin{aligned} E\,(\text{at point } P) &= E_1 + E_2 \\ &= E_0\left(\sin \omega t + \sin(\omega t + \phi)\right) \end{aligned}$$

and use the identity

$$\sin a + \sin b = 2 \sin \frac{1}{2}(a+b)\, \cos \frac{1}{2}(a-b)$$

and so the combined wavefunction is

$$E = 2E_0\, \cos \frac{1}{2}\phi\, \sin\left(\omega t + \frac{1}{2}\phi\right)$$

The *intensity* is proportional to the square of this.

$$I \propto E^2 = 4E_0^2\, \cos^2 \frac{1}{2}\phi\, \sin^2\left(\omega t + \frac{1}{2}\phi\right)$$

We want to average this over one cycle which is effectively what the detector does, and the average of $\sin^2$ (anything) is one half. So if I_0 is the intensity from a single wave, for example when one of the slits is covered, then we end up with

$$I = 4I_0\, \cos^2 \frac{1}{2}\phi$$

But we have from above that

$$\begin{aligned}\phi &= \frac{2\pi}{\lambda}\Delta L \\ &= \frac{2\pi}{\lambda}\, d\sin\theta \\ &\approx \frac{2\pi}{\lambda}\, d\theta \\ &= \frac{2\pi}{\lambda}\, d\,\frac{y}{D}\end{aligned}$$

if θ is small. That leads to

$$I = I_{\max}\cos^2\left(\frac{1}{2}\frac{2\pi}{\lambda}\frac{yd}{D}\right) \tag{5.1}$$

where $I_{\max}$ is the maximum intensity. The most important lesson from this is that to find the intensity, *first* add the *wavefunctions* contributing via the principle of superposition and *then* square them. This rule carries over into quantum mechanics: the wavefunction is called the *probability amplitude* and so long as there is no way of knowing which slit a particular particle goes through, then matter waves will interfere and by adding the probability amplitudes, which are generally complex numbers, and then taking the absolute value of the square, the *probability* is obtained for particles to be detected at point P. If you really want to know how this all works, even at this early stage of your education, you could do a lot worse than to read *QED—The Strange Theory of Light and Matter*, by Richard Feynman. But prepare to be shocked.

Lecture 6

6.1 Lloyd's mirror

Another interference experiment is *Lloyd's mirror.*

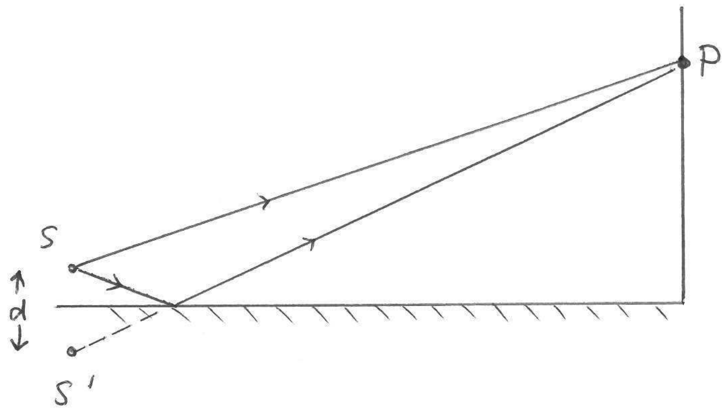


FIGURE 6–1

Light from the monochromatic source S reaches the point P on the screen or detector via a direct path and a reflected path. You can see that the reflected path is equivalent to the light having come from a virtual source S' so you'd expect to see the same pattern as from Young's slits. You do; but the contrast is reversed: so at the point directly ahead of the two "slits" one sees a dark, not a bright, fringe. So whereas the two paths have the same length the beams are nevertheless 180° out of phase. The reason for this is that when a light wave is reflected from the surface of a material with higher refractive index than the medium in which the light is propagating its phase is changed by exactly 180°. I cannot prove that to you here; you will have to wait until you have studied electrodynamics.

But you have learned an important principle: light in a medium, when reflected by a medium having higher refractive index is shifted in phase by an angle of π (180°). That is, cosine becomes minus cosine, sine becomes minus sine. Refraction never produces a change of phase.

6.2 Thin film interference

We consider light impinging on a film at right angles so it is reflected back the way it came. To illustrate this we draw slightly oblique rays in figure 6–2.

The wavelength in air is, say, λ and this is shortened in the medium of the film to

$$\lambda_n = \lambda/n_2$$

The path difference between rays 1 and 2 is $2L$ so these will be out of phase by 180° due to this path difference if

$$2L = \left(m + \frac{1}{2}\right)\lambda_n \ , \quad m = 1, 2 \ldots \qquad \longleftarrow \begin{array}{c}\text{constructive interference}\\ \text{(see below)}\end{array}$$

but ray 1 is shifted by 180° with respect to ray 2 because of its reflection by the the top surface of the film, which has a higher refractive index than air and this is a condition for *constructive interference* since the total phase shift is 2π. The condition for *destructive interference* is

$$\begin{aligned} 2L &= m\lambda_n \ , \quad m = 1, 2 \ldots \qquad \longleftarrow \text{destructive interference} \\ &= \frac{m\lambda}{n_2} \end{aligned}$$

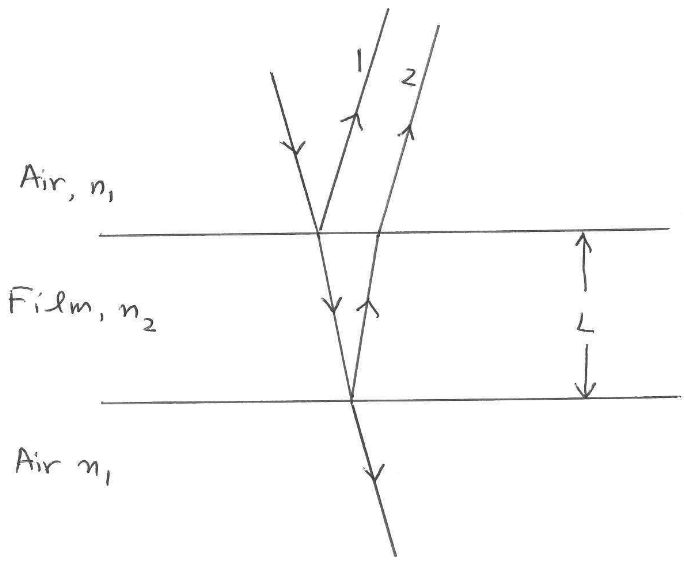


FIGURE 6–2

Of course you are very familiar with thin film interference from oil films, soap bubbles, hummingbird feathers and so on. Two special cases are,

1. Anti-reflection coatings of a certain thickness can be applied to glass so that at least for one range of wavelengths the path difference is such that rays 1 and 2 interfere destructively and very little light is reflected. This is why some lenses have a faint colour in their reflected light; that is the wavelength at which the destructive interference is least effective.

2. If films are very much thinner than the wavelength they appear black under reflected light. This is because the path difference is effectively zero and so because of the phase change of ray 1, rays 1 and 2 are out of phase by π and so interfere destructively.

Newton's rings are formed when a half-convex lens is balanced on a flat sheet of glass and viewed from above.

The air between the lens and the glass sheet is effectively a thin film with smaller refractive index than the glass and of continuously varying thickness. The effect can be used for the testing of lenses, probably exactly as Newton did using the lenses that he ground himself.

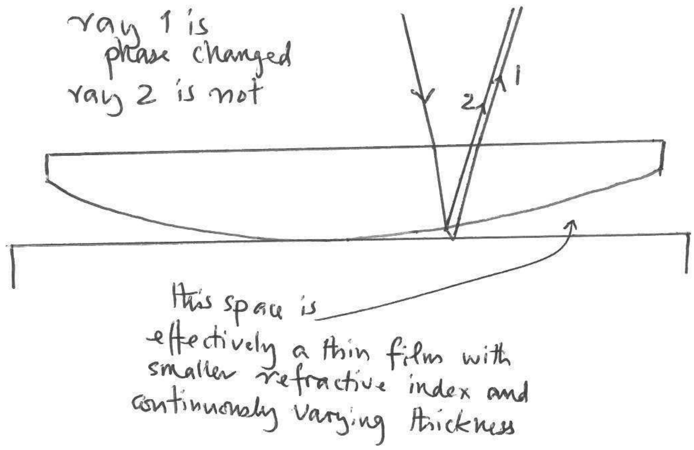


FIGURE 6–3

6.3 Diffraction

Diffraction is light going around corners. As we saw earlier, when a wavefront reaches a narrow slit, at each point of the wavefront a Huygens wavelet is emitted that expands outwards as a spherical wave. As a result, a single slit will produce a pattern on a distant screen which is qualitatively what is seen in a Young's slit experiment. This is called a *diffraction pattern* but this is really a misnomer—it is actually an *interference pattern*. How does this diffraction pattern get formed? We consider a single slit of width a as in figure 6–4.

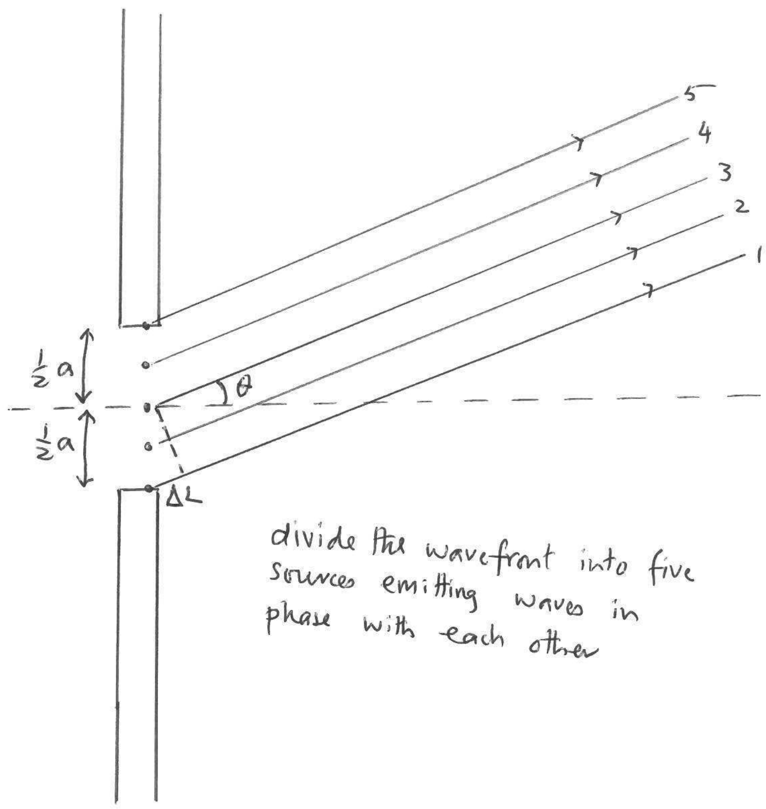


FIGURE 6–4

We notionally divide the wavefront into five sources emitting waves in phase with each other. The path difference between rays from sources 1 and 3 is

$$\Delta L = \frac{1}{2} a \sin\theta$$

It is the same for rays 2 and 4, and for rays 3 and 5. So if the path difference ΔL is

$$\Delta L = \frac{1}{2}\lambda$$

then rays from the upper half of the slit interfere *destructively* with rays from the lower half, so we get a dark fringe on the screen at an angle θ for which

$$\sin\theta = \pm\frac{\lambda}{a} \quad \longleftarrow \quad \text{dark}$$

(the $\pm$ sign is there because the same argument applies for negative θ).

But the division in two halves is really arbitrary; if we divide into quarters and proceed as above we find a dark fringe at angles θ given by

$$\sin\theta = \pm 2\frac{\lambda}{a} \quad \longleftarrow \quad \text{dark}$$

and dividing into six parts,

$$\sin\theta = \pm 3\frac{\lambda}{a} \quad \longleftarrow \quad \text{dark}$$

So in general we get a dark fringe at angles θ when

$$\sin\theta = m\frac{\lambda}{a}\ , \quad m = \pm 1, \pm 2, \pm 3 \ldots \quad \longleftarrow \quad \text{dark}$$

The intensity is not as easy to calculate as in the case of the double slit experiment, equation (5.1),

$$I = I_{\max}\cos^2\left(\frac{1}{2}\frac{2\pi}{\lambda}\frac{yd}{D}\right) \quad \longleftarrow \text{double slit} \tag{5.1}$$

so I'll just give you the result (see Halliday, ch 36 for a proof),

$$\begin{aligned} I &= I_{\max}\left[\frac{\sin\left(\frac{\pi a}{\lambda}\sin\theta\right)}{\frac{\pi a}{\lambda}\sin\theta}\right]^2 \quad \longleftarrow \text{single slit} \qquad (6.1) \\ &\equiv I_{\max}\left[\frac{\sin\alpha}{\alpha}\right]^2 \end{aligned}$$

So the pattern we obtain is a broad central bright fringe at $\theta = 0$ flanked by successively weaker bright fringes interspersing the dark fringes, as sketched in figure 6–5.

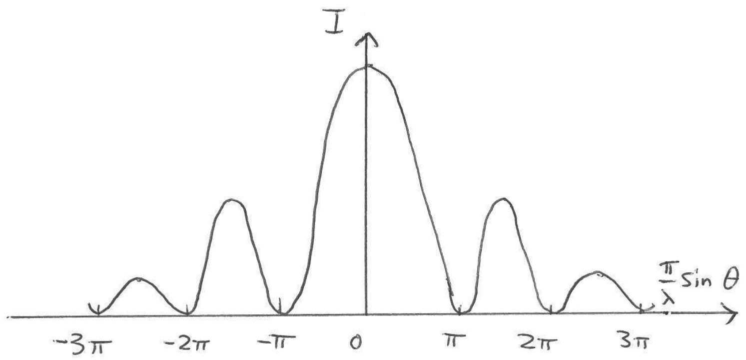


FIGURE 6–5

In figure 6–5 is plotted the quantity $\alpha = (\pi a/\lambda)\sin\theta$ which is in the round parentheses in equation (6.1). (There is a misprint in the figure—the a is missing.) Whenever α is an integer multiple of π its sine goes to zero as seen in figure 6–5.† In addition, because of the denominator α in equation (6-1) the peak intensity falls off with the angle like $1/\alpha$.

You may ask, why does the Young's slit experiment not show diffraction patterns? Well it does. If we plot the contrast expected in the *absence* of diffraction in the double slit experiment it looks like figure 6–6, which is a plot of equation (5.1). Note that there is no reducing of the intensity with angle as it is a simple cosine squared function with no denominator. (But bear in mind from lecture 5, page 6, that this result only holds in the small angle approximation.) Plotting the intensity in this case as a function of $d\sin\theta$ shows clearly the condition for maxima as $d\sin\theta = m\lambda$ (see page 4, lecture 5).

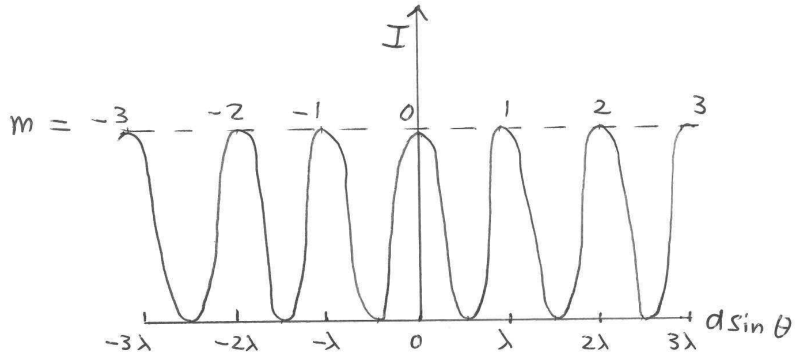


FIGURE 6–6

But actually the intensity in the peaks gets reduced as $\sin\theta$ increases because the diffraction effect multiplies the intensity with its "envelope" of increasingly weaker bright fringes, figure 6–5.

† You may wonder why the central fringe at $\alpha = 0$ is bright. But if you take the limit properly, using the rule of L'Hopital, you find

$$\lim_{\alpha\to 0}\frac{\sin\alpha}{\alpha} = \lim_{\alpha\to 0}\frac{\cos\alpha}{\text{one}} = 1$$

These are combined in an actual experiment. For example here is a pattern from 650 nm coherent light passing through a double slit of separation $d = 18 \times 10^{-6}$ m and width $a = 4 \times 10^{-6}$ m. (Again the a is missing in the x-axis label, sorry.)

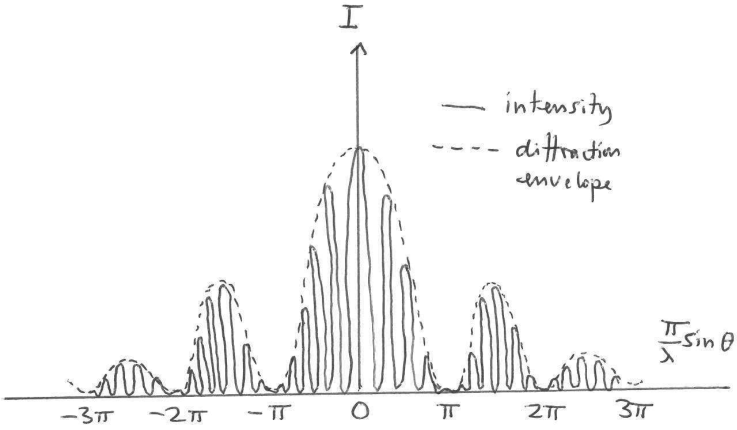


FIGURE 6–7a

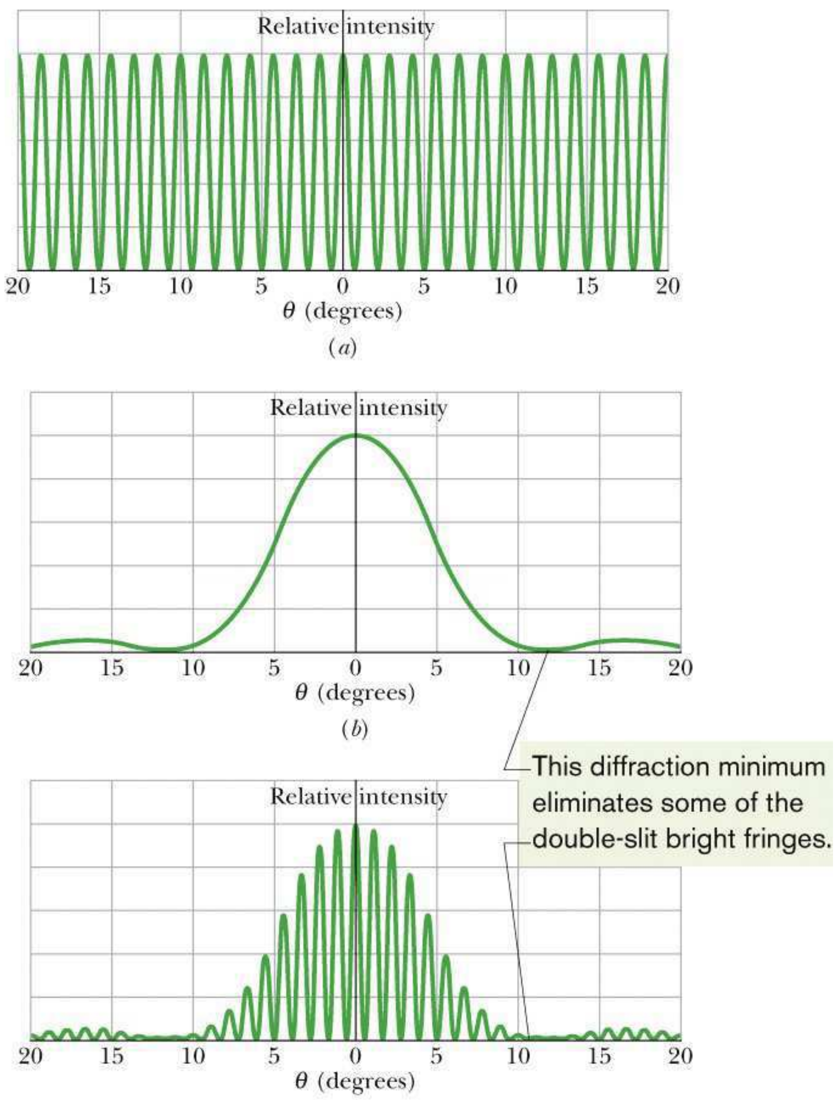


FIGURE 6–7b

Figure 6–7b (Halliday, ch 36, fig 15) shows the same thing. The point is that there are *two* spacings in the double slit experiment: the distance, d, between the slits; and the width, a, of the slits themselves. In the limit of a very small a you can see from equation (6-1) that $I \to I_{\max}$ (because $\sin\alpha \to \alpha$) so that for very narrow slits the central diffraction fringe is very wide. So effectively in Lecture 5 when I described the double-slit experiment I was tacitly assuming a limit of very narrow slits, or $a \ll d \sim \lambda$. But you need to know that in the real double-slit experiment the fringes become weaker as θ increases due to diffraction effects arising from the finite slit-width.

There are many practical applications of diffraction, which you will study later. They include,

1. Diffraction gratings
2. Holography
3. X-ray crystallography

6.4 Resolution and Rayleigh's criterion

Consider two distant, not necessarily coherent, sources, S_1 and S_2. They are a distance L apart such that rays make an angle θ at a slit aperture, of width a.

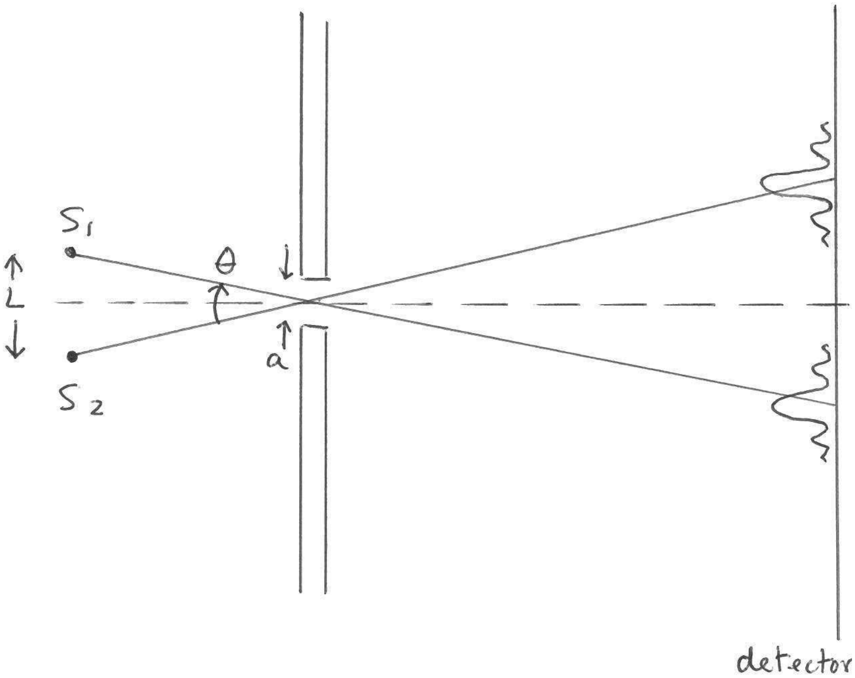


FIGURE 6–8

Each object produces a diffraction pattern at the detector. As the angle θ decreases, either by the objects moving away from the slit or getting closer together the patterns will move closer. If the central maximum of one conicides with the first minimum of the other we will have,

$$\sin\theta = \frac{\lambda}{a}$$

If the angle gets any smaller the detector is no longer able to resolve the two objects and they will appear merged as one. This is the *limit of resolution*, or *Rayleigh criterion*. Actually θ is so small generally that it's enough to assert that,

$$\theta = \frac{\lambda}{a}$$

Most apparatus, including your eye, use a circular, not a slit, aperture and the analysis is a lot harder; the result is that

$$\theta = 1.22\frac{\lambda}{d}$$

where d is the diameter of the aperture. Clearly the best resolution is achieved using a wide aperture, but this will reduce the depth of field. So the quantity to vary is the wavelength. Currently the greatest resolution is with electron microscopes that can resolve down to the distances between atoms in a crystal. Practically, though, the limit of resolution is governed by optical aberations.

6.5 Polarisation

We saw in Lecture 4 how an oscillating electric dipole gives rise to electromagnetic radiation in the sense of oscillating electric and magnetic fields.

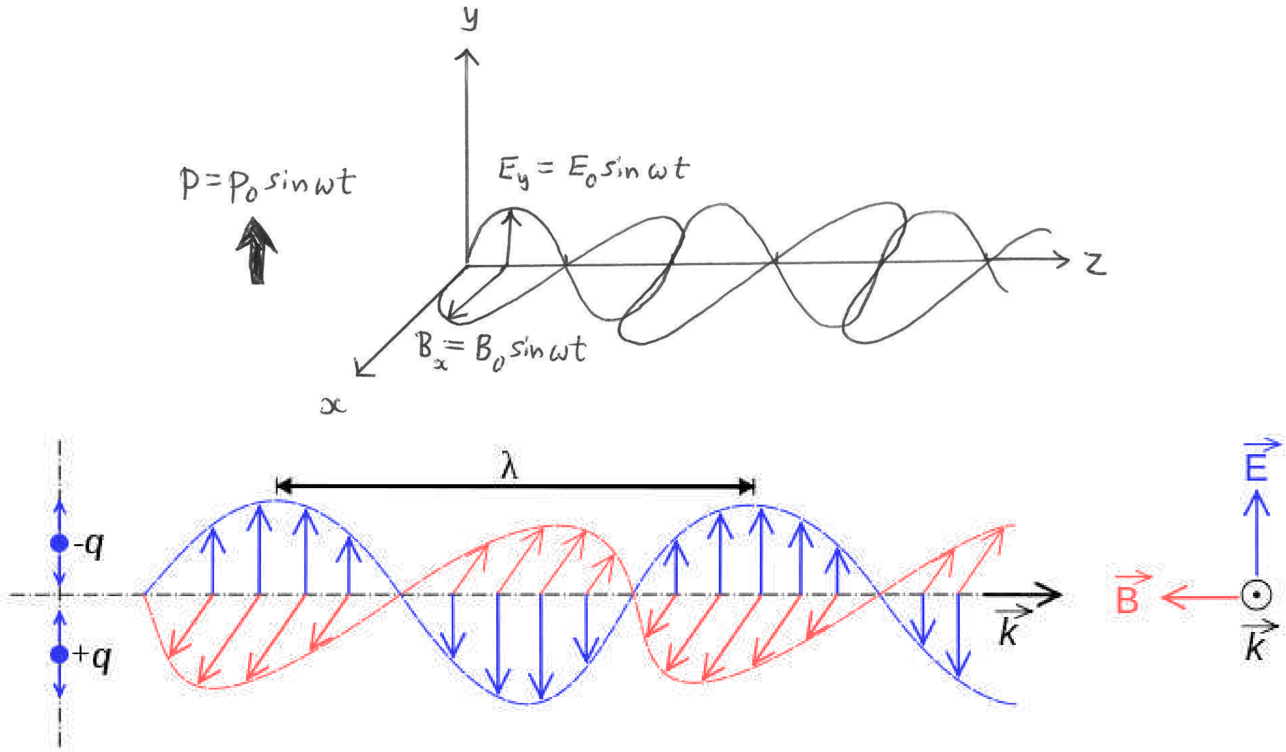


FIGURE 6–9

Of course in a usual source there are dipoles oscillating at all phase differences and pointing in random directions, and so the light is composed of waves having a random distribution of E-field and B-field orientations. Nevertheless any electromagnetic wave can be described as a linear combination of waves *polarised* in the x- and y-directions.

Devices can be made that allow light through in only one direction of polarisation and hence *polarised light* can be produced. You are probably familiar with polarised sunglasses. Look up *calcite* and *birefringence* on `wikipedia`.

Finally I need to tell you that light is *not* a wave. It is made up of particles called *photons* and their quantum mechanical nature causes them to act like waves do because of quantum mechanical interference. The particle nature of light cannot be doubted in view of observations of the photelectric effect and Compton scattering.

Wave optics is not really a physical explanation any more than geometric optics is, but it is closer to the truth. But, for example, in thin film interference don't imagine that rays only emanate from the front and back surfaces; why should they? Light falling on a glass film causes photon–electron interactions to set all the outer electrons into oscillations. These then act as Hertzian dipoles and emit radiation in all directions. All the complicated quantum mechanical interference gives rise to a final outcome that is precisely that produced in the wave optics.

Lecture 7

7.1 Coulomb's law

We have already talked about gravity; although it's very weak it *is* the force that we experience in everyday life. The force between two melons of mass 1 kg a metre apart is tiny: about 10^{-12} N and so in fact gravity is often neglected, as for example in the mass-on-a-spring example that we studied. The reason we rather rarely encounter the *electrostatic force* in everyday life is, paradoxically, because the force is so strong objects are are almost invariably *not* charged; however you have experienced it with balloons, and carpets, and combs and so on.

Whereas the gravitational force between two masses is

$$F = 6.67 \times 10^{-11} \frac{m_1 m_2}{r^2} \qquad [\text{N}]$$

the electrostatic force also follows an inverse square law called *Coulomb's law*:

$$F = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r^2} \qquad [\text{N}]$$

Here, the charges, which may be negative or positive take the role of mass, the distance between them is r and

$$\frac{1}{4\pi\epsilon_0}$$

takes the role of the gravitational constant. They both simply serve to get the units right: SI units in our case. If r is in metres and the charge is in coulombs the force is in newtons if

$$\frac{1}{4\pi\epsilon_0} = 9 \times 10^9 \text{ actually } 8.988 \times 10^9 \; [\text{N m}^2 \text{ C}^{-2}]$$

ϵ_0 is called the "permittivity of free space" and the 4π is there for sensible reasons.

As an example, if two protons in a nucleus are separated by 10^{-15} m (1 fm) the repulsive electrostatic force between them is

$$9 \times 10^9 \times \frac{\left(1.6 \times 10^{-19}\right)^2}{\left(10^{-15}\right)^2} = 230 \text{ N}$$

This is about 10^{36} times larger than their attraction due to gravity. How can nature produce such a huge ratio between two forces? I don't think we'll know until we understand gravity better. Why don't nucleii blast apart? Because there is a *third* even stronger, attractive force acting upon nucleons called the "strong force". It does not concern us in everyday life, because it has a range of only a few fm.

Force is a vector quantity: it has magnitude and direction. Consider two charges q_1 and q_2 at positions $\mathbf{r}_1$ and $\mathbf{r}_2$.

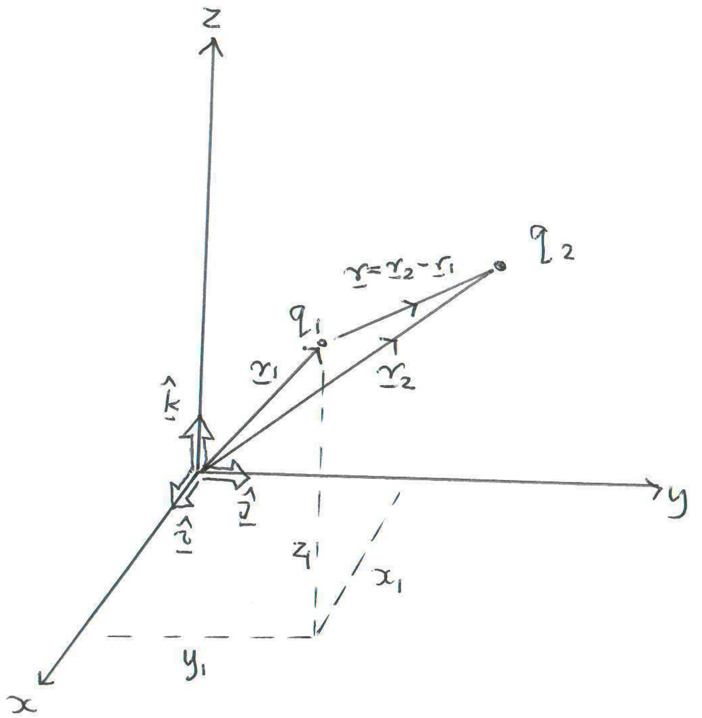


FIGURE 7–1

We write

$$\mathbf{r}_1 = x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}} + z_1\hat{\mathbf{k}} = (x_1, y_1, z_1)$$
$$\mathbf{r}_2 = x_2\hat{\mathbf{i}} + y_2\hat{\mathbf{j}} + z_2\hat{\mathbf{k}} = (x_2, y_2, z_2)$$

$\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are *unit vectors* in the x, y and z directions. They are *dimensionless*. The force *on* the charge q_2 *due to* the charge q_1 acts along the the vector joining the two charges, $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ shown in figure 7–1. Coulomb's law states that the *vector* force is

$$\begin{aligned}\mathbf{F} &= F_x\hat{\mathbf{i}} + F_y\hat{\mathbf{j}} + F_z\hat{\mathbf{k}}\\ &= 9\times 10^9\,\frac{q_1 q_2}{r^2}\,\hat{\mathbf{r}} \qquad [\mathrm{N}]\end{aligned}$$

in which r is the magnitude of the vector $\mathbf{r}$:

$$\begin{aligned}\mathbf{r} &= \mathbf{r}_2 - \mathbf{r}_1\\ &= (x_2 - x_1)\,\hat{\mathbf{i}} + (y_2 - y_1)\,\hat{\mathbf{j}} + (z_2 - z_1)\,\hat{\mathbf{k}}\\ &= x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}\end{aligned}$$

and

$$r = \sqrt{(x^2 + y^2 + z^2)}$$

$\hat{\mathbf{r}}$ is a unit vector pointing in the $\mathbf{r}$-direction:

$$\hat{\mathbf{r}} = \frac{1}{r}\,\mathbf{r}$$

and is dimensionless. It does *not* have dimensions of length.

What if there are more than two charged objects? The *principle of superposition* states that the Coulomb force between two charges is unaffected by the presence of other charges. This is the same principle we earlier learned in connection with the gravitational force. Hence the force on each charge is the *sum* of the forces due to all the others each calculated as if they were acting alone.

All of electrostatics can be deduced from just two postulates,

1. Coulomb's law
2. Principle of superposition

The study of electrostatics may be stated thus: there is a bunch of charges over there; what is the force on a charge that I hold up here?

If any of the charges are *moving* with respect to any others, then the problem is one in *electrodynamics*. However only one postulate needs to be added,

3. Eintein's postulate of special relativity
   (a) The laws of physics (not just mechanics) apply in all inertial frames.
   (b) The speed of light in a vacuum is the same for all inertial observers, whatever the motion of the source.

(Actually (a) and (b) can be derived one from the other so there really is only *one* Einstein postulate.)

7.2 Electric field

If you are holding up a bunch of point charges of strengths, $q_1, q_2, q_3, \ldots$ [Coulomb], all stationary with respect to each other then if I have a point charge of q_0 C, I can place it somewhere and *measure* the magnitude and direction of the force it experiences due to your charges. If I know the values of your charges, $q_1, q_2, q_3, \ldots$, and their positions, $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots$, I can also *calculate* the force on my "test" charge using Coulomb's law and the principle of superposition. I will find a force that is the sum of individual forces,

$$\begin{aligned}
\mathbf{F}_{01} &= \frac{1}{4\pi\epsilon_0}\frac{q_0 q_1}{r_{01}^2}\,\hat{\mathbf{r}}_{01} \ , \quad \hat{\mathbf{r}}_{01} = \frac{(\mathbf{r}_0 - \mathbf{r}_1)}{|\mathbf{r}_0 - \mathbf{r}_1|} \\
\mathbf{F}_{02} &= \frac{1}{4\pi\epsilon_0}\frac{q_0 q_2}{r_{02}^2}\,\hat{\mathbf{r}}_{02} \ , \quad \hat{\mathbf{r}}_{02} = \frac{(\mathbf{r}_0 - \mathbf{r}_2)}{|\mathbf{r}_0 - \mathbf{r}_2|} \\
\mathbf{F}_{03} &= \frac{1}{4\pi\epsilon_0}\frac{q_0 q_3}{r_{03}^2}\,\hat{\mathbf{r}}_{03} \ , \quad \hat{\mathbf{r}}_{03} = \frac{(\mathbf{r}_0 - \mathbf{r}_3)}{|\mathbf{r}_0 - \mathbf{r}_3|} \\
&\ \vdots \\
&\text{etc.}
\end{aligned}$$

For example in the case of, say, q_2 I can draw a relation like this.

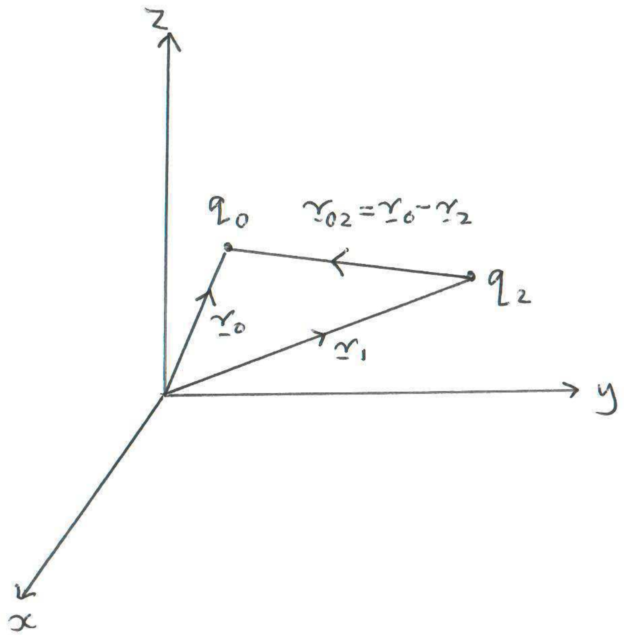


FIGURE 7–2

$\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots$ are called *source points* and $\mathbf{r}_0$ is called a *field point.*

At the point $\mathbf{r}_0$ I will find that the total *force per unit test charge* is

$$\mathbf{E}(\mathbf{r}_0) = \frac{1}{q_0}\left(\mathbf{F}_{01} + \mathbf{F}_{02} + \mathbf{F}_{03} + \ldots\right)$$

I can continue to do this by moving my test charge to any point, $\mathbf{r}$, while you keep your charges still and thereby I can map out the *electric field* $\mathbf{E}(\mathbf{r})$. This is a *vector field* because it defines a vector at each point in space—it is a vector that depends on another vector. Always remember, electric field is force per unit test charge.

You could imagine a little arrow at every point you wish to indicate the field strength, like this.

FIGURE 7–3

Examples of a *scalar field* are,

- $T(\mathbf{r})$, the temperature at any point in a metal bar
- $p(\mathbf{r})$, the pressure in the atmosphere
- $h(x, y)$, height above sea-level, as in a contour map
- $V(\mathbf{r})$, electric potential—we'll come to that

Examples of a *vector field* are,

- $\mathbf{E}(\mathbf{r})$, the electric field
- $\mathbf{B}(\mathbf{r})$, the magnetic field, or "magnetic induction"
- $\mathbf{H}(\mathbf{r})$, the "H-field", or magnetic excitation (equals $\mathbf{B}/\mu_0$ in a vacuum)†
- $\mathbf{v}(\mathbf{r})$, velocity of flow at point $\mathbf{r}$ in a river, or a wind

The are also *tensor fields*, for example

- $\mathsf{R}(\mathbf{r}, t)$, Riemann–Christoffel tensor, or *curvature*, of space time at the point $(\mathbf{r}, t)$.
- $\tau(\mathbf{r})$, the stress at the point $\mathbf{r}$ in a solid

but we shan't use those here.

7.3 Electric field due to a point charge

I will often come back to one of the simplest problems in electrostatics, namely the electric field due to a point charge. (The notion of *point charge* is not a complete idealisation: as far as we know an electron is a point charge). Now, we know that the electric field must be radially symmetric and that it becomes smaller, in proportion to $1/r^2$ as the field point is taken further away from the source point (the point charge). If the point charge is placed at the origin of a cartesian coordinate system and has charge q [C], the electric field at a field point $\mathbf{r}$ depends only on the magnitude, r, and is

$$E(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} \qquad [\mathrm{N\ C^{-1}}]$$

and *by convention* points *away* from the source charge if q is positive and *towards* the source charge if q is negative.

Take a positive point charge and try to illustrate the electric field by drawing a few representative arrows whose length depends on the magnitude of the field and which point in the direction of the field.

---

† The names given to the $\mathbf{B}$ and $\mathbf{H}$ fields are problematic (see Griffiths, "Introduction to Electrodynamics," p. 271). If you are not dealing with magnetism in matter, then there is no need to use $\mathbf{H}$ since $\mu_0\mathbf{H} = \mathbf{B}$. In the laboratory one deals with $\mathbf{H}$ because this is what can be controlled by varying the current in a solenoid; in fact $H = nI$ if a current I is passed through a solenoid of n turns per meter. If the solenoid is empty of matter then the magnetic field is just $B = \mu_0 nI$. But if you are dealing with magnetism in matter you may need to use both $\mathbf{B}$ and $\mathbf{H}$ fields. The problem is what to call them. Lorrain and Corson, your other recommended reading, call $\mathbf{H}$ the "magnetic field intensity," and $\mathbf{B}$ the "magnetic induction." Let us rather follow Griffiths and the sublime Arnold Sommerfeld, and call $\mathbf{B}$ the magnetic field or magnetic field strength and if we need use $\mathbf{H}$ we will call it the *magnetic excitation*, following Sommerfeld and by analogy with the *electric excitation*, $\mathbf{D}$; or we will call it the "H-field". But for our purposes neither $\mathbf{H}$ nor $\mathbf{D}$ are required since we do not deal with electricity and magnetism in matter.

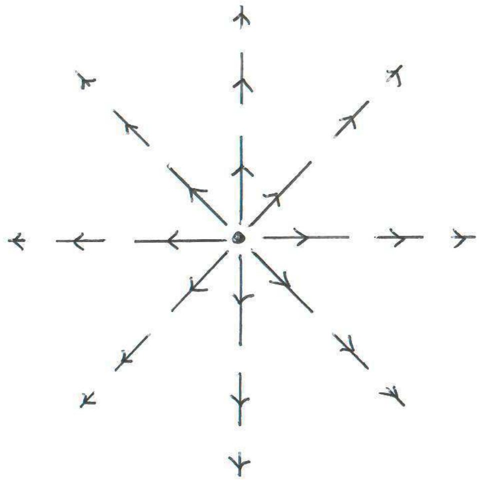


FIGURE 7–4

If you join the lines up like this,

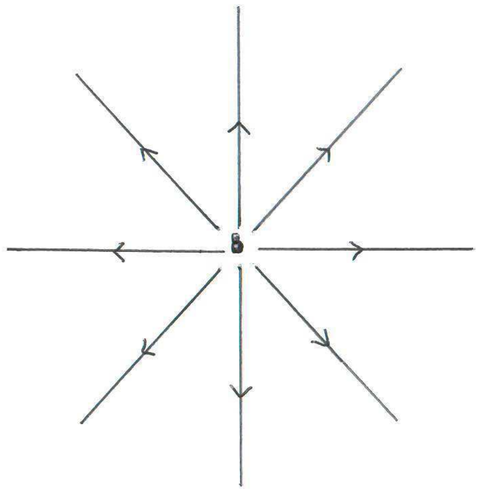


FIGURE 7–5

you may think that you have lost some information about the strength of the field which was illustrated by the length of the arrow, but you have not because this is still evident in the *density* of the lines. In this two dimensional drawing, the density of lines, or number of lines crossing a unit length of a circle of radius r, centred at the source charge, is proportional to $1/r$. But in real life, in three dimensions, the number of lines crossing unit area of a sphere of radius r, centred at the source charge is proportional to $1/r^2$, exactly reflecting the inverse square law. Therefore we can say that the *flux*—to be defined later, but roughly the number of lines penetrating total area—due to a point charge is constant, independent of r. Only an inverse square law, as in electrostatics or gravitation, allows us to take this liberty.

Drawing field lines is an illustration only as you choose how many lines to draw, based upon the amount of patience that you have and how sharp is your pencil. But having made this choice then the strength of the field is proportional to the density of the field lines that cut a given area whose normal is parallel to the field.

Rules for drawing field lines

1. Field lines run from positive to negative point charges. You may choose the number of lines to draw, but the number entering or leaving a point charge must be proportional to its charge.
2. Field lines can only end at a point charge or at infinity.
3. Field lines may not cross, because in all places the direction of $\mathbf{E}$ is single-valued.

Be cautious with two dimensional representations of the three dimensional field. Later, we will want to ask how many field lines intersect a unit area and this is not easily related to the number that cross a line in a two dimensional drawing of the thee dimensional situation.

Lecture 8

8.1 Electric field due to two point charges

Let us draw the field lines from two point charges of equal magnitude. If they are both positive we see this:

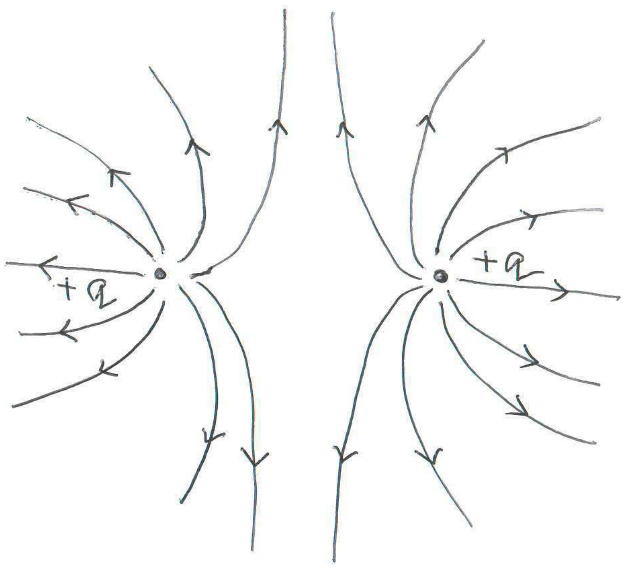


FIGURE 8–1

If they are both negative then the sense of all the arrows is reversed. If one is positive and the other negative then we call this an *electric dipole* and the field lines look like this:

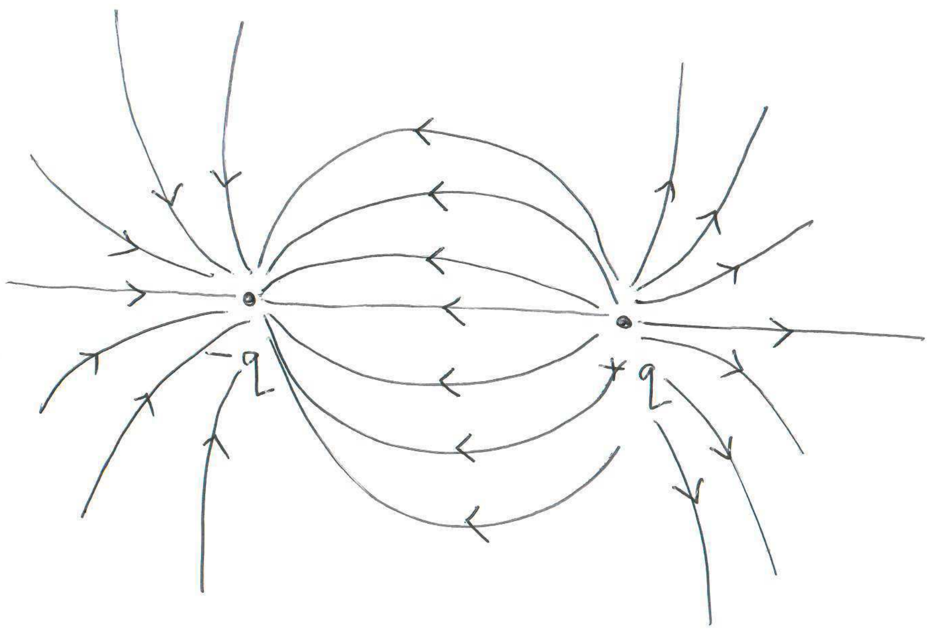


FIGURE 8–2

We now try and calculate the electric field due to a dipole of charges $+q$ and $-q$ separated along the x-axis by a distance d. We ask, what is the electric field at a field point $\mathbf{r}$?

By the principle of superposition this is the vector sum of the fields due to each point charge as if the other were not acting. We will assume the case $r \gg d$ so that we can regard d^2/r^2 as negligibly small. If you like, we are working *to first order* in d/r.

We choose an origin half-way between the two charges and we consider this diagram.

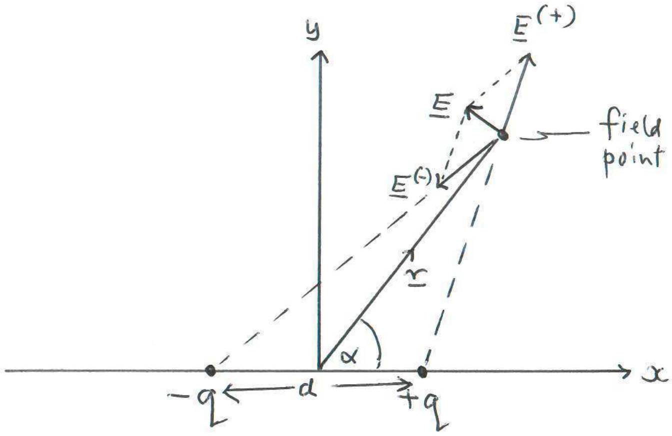


FIGURE 8–3

Note, the problem has cylindrical symmetry—nothing depends on the angle about the x-axis. $\mathbf{E}$ is the vector sum of $\mathbf{E}^{(+)}$ the electric field due to the positive charge, and $\mathbf{E}^{(-)}$, that due to the negative charge. Remember that by convention the field points away from positive charges and towards negative charges. To find $\mathbf{E}$ is pretty hard to work out. It is best done by calculating first the *electric potential*, but we haven't got to that yet. It will serve our purpose well enough to study two limiting cases, namely $\alpha = 90^\circ$ and $\alpha = 0$.

If $\alpha = 90^\circ$ the field point is on the y-axis as drawn below.

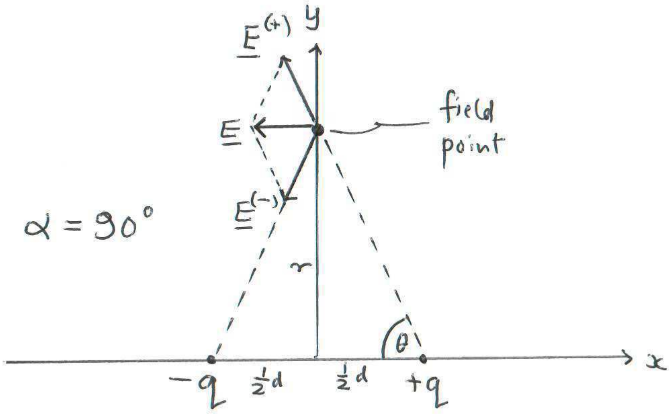


FIGURE 8–4

By inspection $\mathbf{E}$ only has an x-component, which is *twice* the x-component of either $\mathbf{E}^{(+)}$ or $\mathbf{E}^{(-)}$, the fields due to the individual charges. Let's find $\mathbf{E}^{(+)}$.

$$\mathbf{E}^{(+)} = \frac{1}{4\pi\epsilon_0}\,\frac{q}{r^2 + \frac{1}{4}d^2}\,(-\cos\theta\,\hat{\mathbf{i}} + \sin\theta\,\hat{\mathbf{j}})$$

In the denominator,

$$r^2 + \frac{1}{4}d^2 = r^2\left(1 + \frac{1}{4}\frac{d^2}{r^2}\right) \approx r^2$$

to first order, and so

$$\mathbf{E} = 2E_x^{(+)}\,\hat{\mathbf{i}} = -\frac{1}{4\pi\epsilon_0}\,\frac{2q}{r^2}\,\cos\theta\,\hat{\mathbf{i}}$$

but by inspecting figure 8–4, and again to first order

$$\cos\theta = \frac{\frac{1}{2}d}{\sqrt{r^2 + \frac{1}{4}d^2}} \approx \frac{1}{2}\frac{d}{r}$$

so that finally,

$$\mathbf{E} = -\frac{1}{4\pi\epsilon_0}\,\frac{qd}{r^3}\,\hat{\mathbf{i}}$$

The vector $qd\hat{\mathbf{i}}$ is called the *dipole moment vector*, $\mathbf{p}$ [C m], and so we rewrite the last equation as

$$\mathbf{E} = -\frac{1}{4\pi\epsilon_0}\,\frac{1}{r^3}\,\mathbf{p} \qquad \text{in the case } \alpha = 90^\circ \tag{8.1}$$

Please remember that by convention the dipole moment vector points *from* the negative charge *to* the positive charge.

Now for the case $\alpha = 0$ the field point is on the x-axis as here:

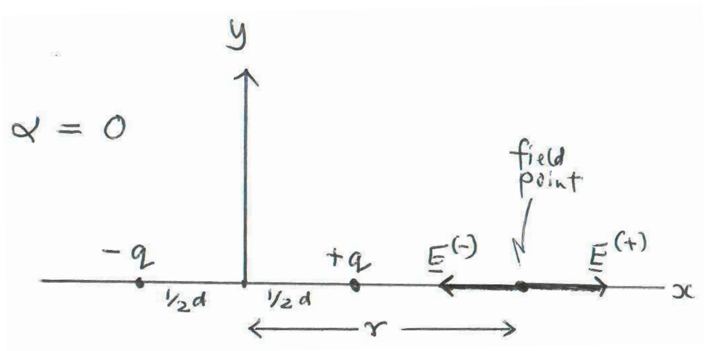


FIGURE 8–5

The summed contribution from the two charges to the electric field at the field point is

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\,\frac{q}{\left(r - \frac{1}{2}d\right)^2}\,\hat{\mathbf{i}} - \frac{1}{4\pi\epsilon_0}\,\frac{q}{\left(r + \frac{1}{2}d\right)^2}\,\hat{\mathbf{i}}$$

To first order,

$$\left(r + \frac{1}{2}d\right)^2 \approx r^2 + rd$$

$$\left(r - \frac{1}{2}d\right)^2 \approx r^2 - rd$$

and this leads to

$$\begin{aligned}\mathbf{E} &= \frac{1}{4\pi\epsilon_0}\, q \left(\frac{1}{r^2 - rd} - \frac{1}{r^2 + rd}\right)\hat{\mathbf{i}} \\ &= \frac{1}{4\pi\epsilon_0}\, q \left(\frac{(r^2 + rd) - (r^2 - rd)}{(r^2 - rd)\,(r^2 + rd)}\right)\hat{\mathbf{i}} \\ &= \frac{1}{4\pi\epsilon_0}\,\frac{2qd}{r^3}\,\hat{\mathbf{i}}\end{aligned}$$

the last line being correct to first order. In terms of the dipole moment vector we get

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\,\frac{1}{r^3}\,2\mathbf{p} \qquad \text{in the case } \alpha = 0 \tag{8.2}$$

This is the same as for the case $\alpha = 90°$ except for the factor of two and the minus sign. In fact for any value of α the electric field is proportional to the dipole moment and inversely proportional to r^3 as long as $r \gg d$,

$$E \propto \frac{p}{r^3}$$

Compare this with the field due to a point charge which is proportional to $1/r^2$. This should not surprise you: since the dipole is electrically neutral it should be harder to detect from a long distance than the point charge.

8.2 Electric field due to a uniform line of charge

What is the electric field at a perpendicular distance, s, from the centre of a charged wire of length $2L$? We do this using the integral calculus by summing (integrating) the fields due to infinitesimal segments of the wire and appealing to the principle of superposition. Consider those two segments of the wire of length dx at a distance x from the centre of the wire.

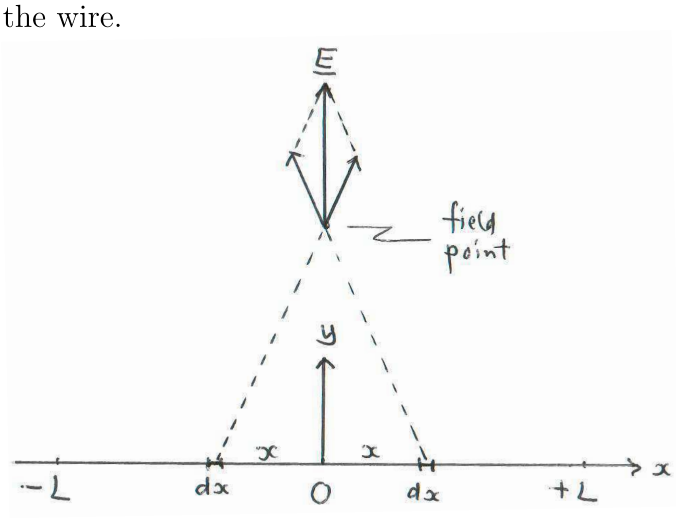


FIGURE 8–6

We suppose that the wire is charged uniformly to an amount λ [C m^{-1}]; so the elements at $\pm x$ of length dx each carry a charge

$$\mathrm{d}q = \lambda \mathrm{d}x$$

By combining the segments at $\pm x$ in pairs we exploit the symmetry of the problem and indeed we observe that for each pair the problem is that of the dipole that we solved in section 8.1 except that the elements of charge have the same sign. In contrast to the dipole case of $\alpha = 90°$ we seek twice the y-component of the field due to any one of the infinitesimal segments. By examining figure 8–6 we see that the infinitesimal electric field due to the two segments of charge is

$$\mathrm{d}\mathbf{E} = \frac{1}{4\pi\epsilon_0}\; 2\lambda \mathrm{d}x\, \frac{1}{x^2+s^2}\, \sin\theta\, \hat{\mathbf{j}}$$

and to get the total field at the field point we sum all these by an integration from zero to L, and using

$$\sin\theta = \frac{s}{\sqrt{x^2+s^2}}$$

as we see from figure 8–6 using the Pythagoras rule. This leads to

$$\begin{aligned}
\mathbf{E} &= \frac{1}{4\pi\epsilon_0}\; 2\lambda\, \hat{\mathbf{j}} \int_0^L \frac{s\,\mathrm{d}x}{(x^2+s^2)^{3/2}} \\
&= \frac{1}{4\pi\epsilon_0}\; 2\lambda s\, \hat{\mathbf{j}} \left[\frac{x}{s^2\sqrt{s^2+x^2}}\right]_0^L \\
&= \frac{1}{4\pi\epsilon_0}\; \frac{2\lambda L}{s\sqrt{s^2+L^2}}\, \hat{\mathbf{j}}
\end{aligned}$$

Let us consider two limiting cases.

1. If the length $2L$ is very small compared to s, then we are looking at a short wire from a very long way off and
$$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\, \frac{2\lambda L}{s^2}\, \hat{\mathbf{j}}$$
which is the field due to a point charge of charge $2\lambda L$ as you'd expect.
2. On the other hand if $L \to \infty$ we have
$$\frac{L}{s\sqrt{s^2+L^2}} = \frac{1}{s\sqrt{1+s^2/L^2}} \longrightarrow \frac{1}{s}$$
and
$$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\, \frac{2\lambda}{s}\, \hat{\mathbf{j}}$$
which is the electric field at a distance s from a long straight wire of charge density λ [C m^{-1}]. We will obtain this result again once we have learned Gauss's law. Actually this problem has cylindrical symmetry so we can say that the wire radiates field lines outwards and the magnitude of the electric field is
$$E = \frac{1}{4\pi\epsilon_0}\, \frac{2\lambda}{s} \tag{8.3}$$
noting that the field is inversely proportional to the distance from the wire. It should not surprise you that the field due to the wire decays away more slowly that the field due to a point charge.

It is a useful lesson in physics to check any complicated formula you have obtained to see that in certain limits it gives sensible results or results that you can obtain by another means.

Lecture 9

9.1 Electric field due to a uniform sheet of charge

We imagine a circular sheet of charge and ask, what is the electric field at a point a distance s above its centre?

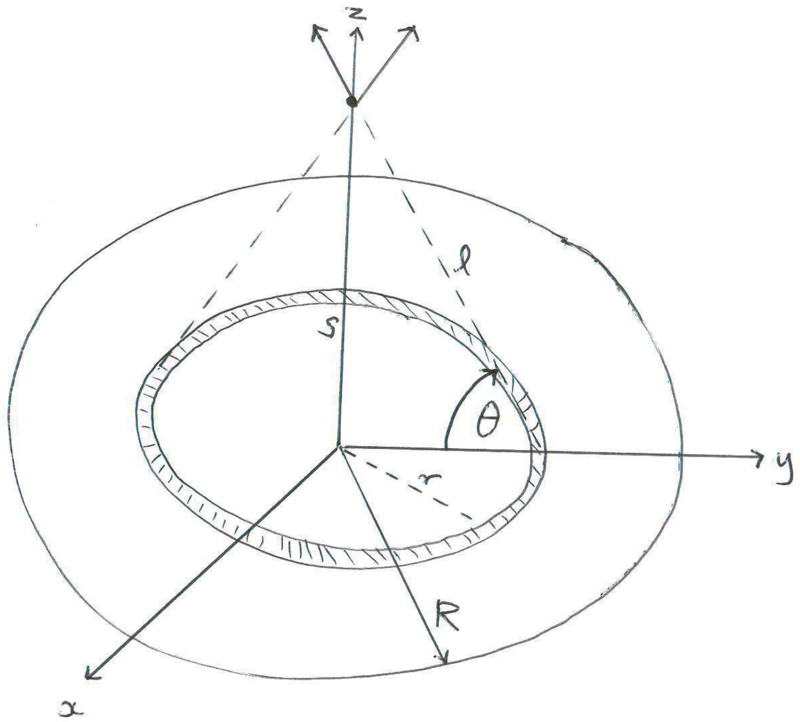


FIGURE 9–1

Again we *sum by integration* but now to exploit the symmetry of the problem the infinitesimal segments that we sum are narrow rings of charge. We suppose that the sheet is uniformly charged with a density of σ [C m^{-2}]. The element of charge is a ring of radius r and carries a charge of

$$\mathrm{d}q = \sigma\, 2\pi r\, \mathrm{d}r \qquad [\mathrm{C}]$$

At our field point the *magnitude* of the electric field due to any point within the ring is the same and the total field obviously points along the z-axis in figure 9–1 by symmetry. So the element of the electric field is

$$\mathrm{d}\mathbf{E} = \frac{1}{4\pi\epsilon_0}\,\frac{\mathrm{d}q}{l^2}\,\sin\theta\,\hat{\mathbf{k}}$$

in which l is defined in figure 9–1,

$$l = \sqrt{r^2 + s^2}$$

by Pythagoras rule; from the trigonometry of figure 9–1

$$\sin\theta = \frac{s}{\sqrt{r^2 + s^2}}$$

and so

$$\begin{aligned} d\mathbf{E} &= \frac{1}{4\pi\epsilon_0}\;\frac{s\,dq}{(r^2+s^2)^{3/2}}\,\hat{\mathbf{k}} \\ &= \frac{1}{4\pi\epsilon_0}\,2\pi\sigma s\,\frac{r\,dr}{(r^2+s^2)^{3/2}}\,\hat{\mathbf{k}} \end{aligned}$$

We now sum the field from all the infinitesimal rings from the centre at $r = 0$ to the edge of the circular sheet at $r = R$ by an integration,†

$$\begin{aligned} \mathbf{E} &= \frac{1}{4\pi\epsilon_0}\;2\pi\sigma s\int_0^R \frac{r\,dr}{(r^2+s^2)^{3/2}}\,\hat{\mathbf{k}} \\ &= \frac{1}{4\pi\epsilon_0}\;2\pi\sigma s\left(\frac{1}{s}-\frac{1}{\sqrt{R^2+s^2}}\right)\hat{\mathbf{k}} \end{aligned}$$

If we now let $R \to \infty$ we obtain the electric field due an infinite, uniformly charged sheet, whose normal is the z-axis, carrying a charge density σ C m^{-2}

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\;2\pi\sigma\,\hat{\mathbf{k}}$$

We can say that the electric field is *constant* and of magnitude

$$\frac{\sigma}{2\epsilon_0}$$

and pointing perpendicular to the sheet: out of the sheet if the charge is positive and into the sheet if it's negative. The field strength is *independent* of the distance from the sheet. Can you see why this is so?

You can now easily find the electric field inside and outside a parallel plate capacitor, neglecting end effects.

---

† To do this integral, write

$$I = \int_0^R \frac{r\,dr}{(r^2+s^2)^{3/2}}$$

Substitute $u = r^2$, $du = 2r dr$ giving

$$2\frac{d}{du}\left(u+s^2\right)^{-1/2} = -\left(u+s^2\right)^{-3/2}$$

Then

$$\begin{aligned} I &= \frac{1}{2}\int_0^{R^2} \frac{du}{(u+s^2)^{3/2}} \\ &= -\left[\left(u+s^2\right)^{-1/2}\right]_0^{R^2} \\ &= \left(\frac{1}{s}-\frac{1}{\sqrt{R^2+s^2}}\right) \end{aligned}$$

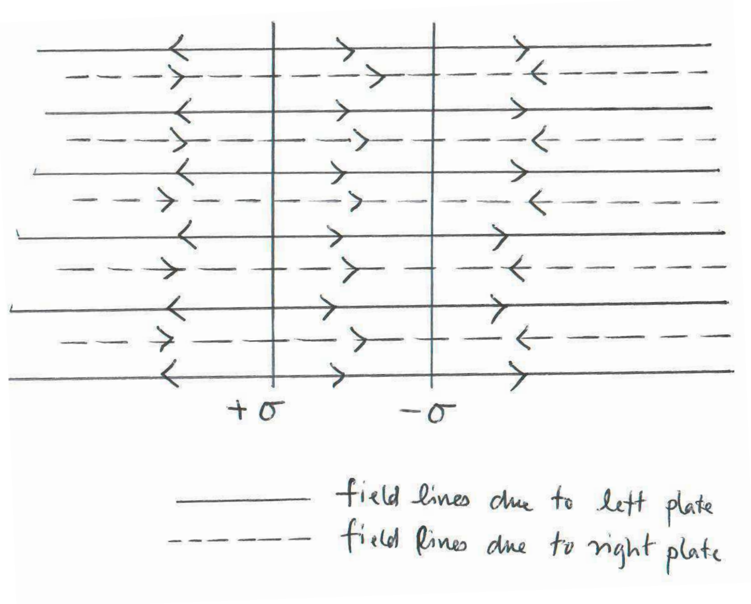


FIGURE 9–2

Outside the fields cancel and there is no field. Inside the field lines combine and so the total field is

$$E = \frac{\sigma}{\epsilon_0}$$

pointing from the postive to the negative plate.

Let us finish with the following useful table. You must become familiar with the electric fields (and later magnetic fields) associated with the most common and simple objects. In particular remember how the fields depend on the distance of the field point from the source charge density.

charge configuration	electric field
point	$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} \propto \frac{1}{r^2}$
dipole	$E \propto \frac{1}{r^3}$
line	$E = \frac{1}{4\pi\epsilon_0}\frac{2\lambda}{r} \propto \frac{1}{r}$
sheet	$E = \frac{\sigma}{2\epsilon_0} \propto$ constant

9.2 Electric flux

The number of electric field lines that pass through a surface orientated at right angles to the field is called the *flux*. The bigger the surface, the bigger the flux. Of course the number of field lines depends on how sharp is your pencil and how much patience you have and so is rather subjective; all the same I can define the flux Φ as

$$\Phi = \text{field} \times \text{area} = EA$$

noting that E is not a vector, it's its magnitude.

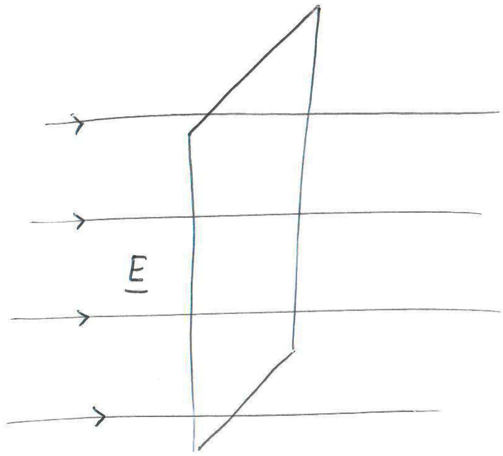


FIGURE 9–3

If the field is not uniform I must take a surface that is infinitesimally small. If I'm interested in a surface that is not perpendicular to the field, I must multiply E by the area projected perpendicularly to it. In that situation I draw this.

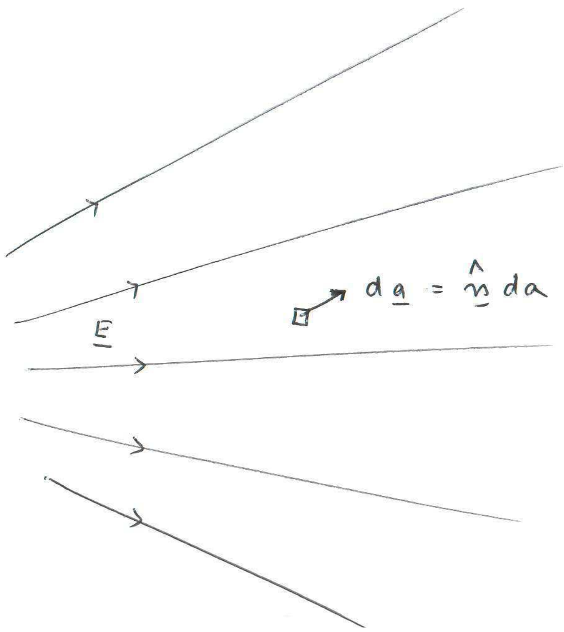


FIGURE 9–4

d**a** is a shorthand for $\hat{\mathbf{n}}\,\mathrm{d}a$, where $\mathrm{d}a$ is the infinitesimal area (in the sense of the differential calculus) and $\hat{\mathbf{n}}$ is the unit vector normal (perpendicular) to my chosen surface. $\mathbf{E}$ is the electric field at the centre of the infinitesimal area (actually we choose the infinitesimal area small enough so that the field is constant over its whole area, so it doesn't have to be specified as the centre). To repeat: d$\mathbf{a}$ is a vector whose direction is normal to the surface and whose length is equal to its area. So d$\mathbf{a}$ has units [Length]^2. The tiny bit of flux penetrating the infinitesimal area is

$$\mathrm{d}\Phi = \mathbf{E} \cdot \mathrm{d}\mathbf{a}$$

The scalar product is zero if $\mathbf{E}$ and d$\mathbf{a}$ are perpendicular; this is what you'd expect for in that case no flux lines cross the surface as they all lie *in* the surface. If $\mathbf{E}$ and d$\mathbf{a}$ are parallel then the flux is $E\mathrm{d}a$ which is the product of the magnitudes of $\mathbf{E}$ and d$\mathbf{a}$—again this is what you expect as you've orientated your surface to be perpendicular to the field lines.

If I now have a large surface and a non uniform electric field, I can divide the surface up into infinitesimal pieces and do an integral (a sum) over the surface to get the flux,

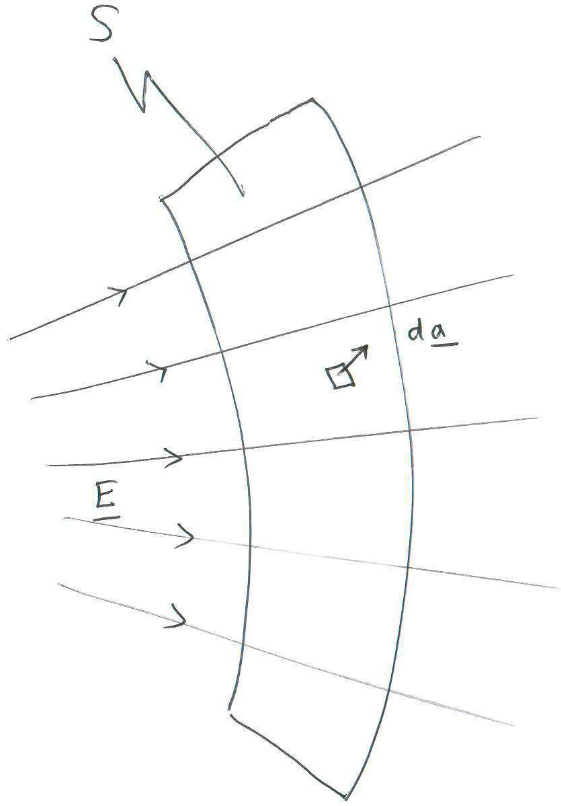


FIGURE 9–5

$$\Phi = \int_S \mathbf{E} \cdot \mathrm{d}\mathbf{a} \tag{9.1}$$

which means the sum of the increments of flux dΦ over all elements d$\mathbf{a}$ of the surface, S. You can probably see that the total area of the surface is

$$A = \int_S \mathrm{d}a$$

9.3 Gauss's law

Let's now think of a point charge, q, and its associated electric field lines. What is the flux? Imagine a sphere of radius r with the charge of amount q at its centre. The field is everywhere parallel to the normal to the surface, that is the radius vectors pointing from the the centre to the surface of your imaginary sphere, so the flux through the surface is

$$\Phi = E \times A = \frac{1}{4\pi\epsilon_0}\,\frac{q}{r^2} \times 4\pi r^2$$

Therefore

$$\Phi = \frac{1}{\epsilon_0}\,q$$

The flux is *independent of the radius of the imaginary sphere.* This is because the area is proportional to r^2 and because of Coulomb's law the field is proportional to $1/r^2$ and the two cancel in the product $E \times A$. Another way to see it is that the number of field lines intersecting the surface of the sphere is the same for any sphere. In fact it's the same for any *closed surface.* And the charge need not be at the centre.

I might have a number of point charges (some positive, some negative) but by the principle of superposition the flux from each will add together. In consequence of these facts I can assert that the flux penetrating any closed surface is $1/\epsilon_0$ times the charge inside,

$$\Phi = \frac{1}{\epsilon_0}\,Q_{\text{enclosed}}$$

If there is only charge *outside* then any field lines going in will have to come out so there can be no net flux entering or leaving due to external charges. Because of equation (9.1) we can write

$$\oint_S \mathbf{E}\cdot \mathrm{d}\mathbf{a} = \frac{1}{\epsilon_0}\,Q_{\text{enclosed}} \tag{9.2}$$

where the circle on the integral sign indicates that the surface, S, is closed, and Q_{enclosed} is all the charge inside S.

Equation (9.2) is *Gauss's law.* It follows from Coulomb's law and the principle of superposition. It holds for any inverse square law, so there is a Gauss's law also in gravitation.

Gauss's law is hugely powerful as a way to find the field due to a distribution of charge (or mass in the case of gravitation) particularly under either of two circumstances:

1. Gauss's law can give you an immediate answer to a general question.
2. Gauss's law furnishes us with a ready solution to the electric field in the case of charge distributions having certain symmetries. The trick is to find an imaginary surface of suitable shape so that the flux through its faces is either constant or zero.

Lecture 10

10.1 Applications of Gauss's Law

To demonstrate point 2 from the end of the previous lecture, I now want to do the point, line and sheet of charge again to show you how much easier the problem is if we use Gauss's law. *Always remember* that the surface that we call the "gaussian surface" is an imaginary construction that you set up as a device to aid your calculation.

10.1.1 Point charge

Here the argument is the reverse of the one we used in Lecture 9, section 9.3, to obtain Gauss's law. The electric field due to a point charge obviously has spherical symmetry. So the clear choice of gaussian surface is a sphere centred on the point charge. This is because the flux is constant over the surface of the sphere. By choosing a surface of radius r we can find the electric field as a function of the distance r from the charge. Of course we do this just once for an arbitrary choice of radius. We also know that the electric field is *radial*, that is it points directly away or towards the point charge in straight field lines, so we only need to concern ourselves with the magnitude, E. So here is our point charge and our gaussian surface, a sphere of radius r,

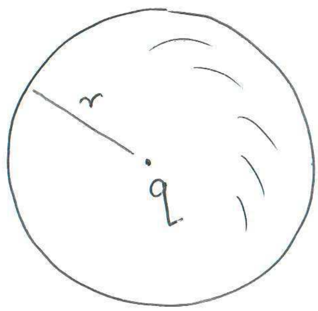


FIGURE 10–1

The flux crossing the surface of area

$$A = 4\pi r^2$$

is

$$\Phi = EA$$

and by Gauss's law, this is

$$\Phi = EA = \frac{1}{\epsilon_0} Q_{\text{enclosed}} = \frac{q}{\epsilon_0}$$

Therefore

$$4\pi r^2 E = \frac{q}{\epsilon_0}$$

or

$$E = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2}$$

which of course we know already.

10.1.2 An infinitely long, uniform line of charge

The symmetry here is obviously cylindrical: the field lines leave (or enter) the wire radially in straight lines and so the field strength E is constant at a distance s from the wire, independent of the direction. The obvious gaussian surface to choose is a cylinder since for any cylinder of radius s the field is uniform on the curved surface and of magnitude $E(s)$. As the wire has no ends there is no flux crossing the ends of the cylinder.

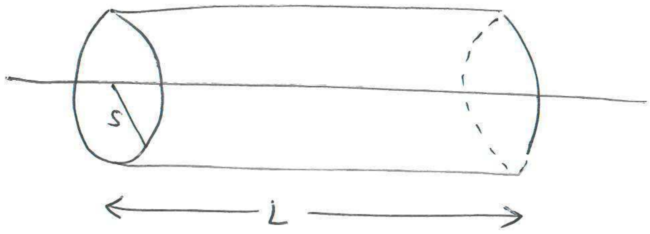


FIGURE 10–2

So we ignore the ends as there's no flux crossing them and we have the flux equal to

$$\begin{aligned}\Phi &= EA = \text{field} \times \text{area}\\ &= E \times 2\pi s L\end{aligned}$$

If the wire carries a uniform charge λ [C m^{-1}] then the charge inside the cylinder is

$$Q_{\text{enclosed}} = \lambda L$$

and Gauss's law tells us that

$$\Phi = EA = \frac{1}{\epsilon_0} Q_{\text{enclosed}}$$

or

$$E \times 2\pi s L = \frac{1}{\epsilon_0} L\lambda$$

naturally the length, L, of the cylinder cancels as you'd hope it would as this construction must not depend on how long we choose our cylindrical imaginary surface, and we get

$$E = \frac{1}{4\pi\epsilon_0} \frac{2\lambda}{s}$$

pointing radially with respect to the wire. This is of course the result we obtained in Lecture 8, equation (8.3), when we obtained it with a much more difficult piece of mathematics. Mind you, then we got the more general result for a wire of finite length; you cannot use Gauss's law in that problem as there is insufficient symmetry.

10.1.3 An infinite uniformly charged sheet

The infinite sheet carries a uniform surface charge density of σ [C m^{-2}]. Field lines obviously emanate perpendicular to the surface. I need a gaussian surface that has a flat parallel face above and below the sheet a distance d away and having sides(s) perpendicular to the surface whose shape is irrelevant as no flux penetrates them anyway. So I could use a rectangular box or a cylindrical, so called, "pill box".

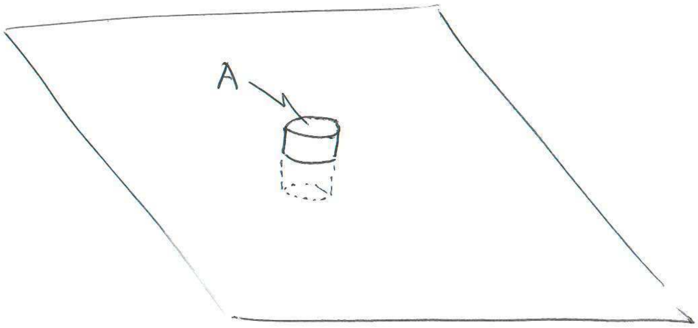


FIGURE 10–3

The area will anyway cancel as you'd anticipate as the construction must not depend on this. The flux is the same penetrating the upper and lower surfaces, so I have

$$\Phi = E \times 2A$$

and no flux penetrates the sides. The charge inside the imaginary box is

$$Q_{\text{enclosed}} = \sigma A$$

and so by Gauss's law

$$\Phi = \frac{1}{\epsilon_0} Q_{\text{enclosed}}$$

or

$$E \times 2A = \frac{1}{\epsilon_0} \sigma A$$

which leads to

$$E = \frac{\sigma}{2\epsilon_0}$$

which naturally is the same as we got earlier in Lecture 9, page 2, by direct integration. There, we also got the field due do a finite disc.

Why do you think the field due to the *infinite* sheet does not depend on the distance away?

10.2 Charge density

So far we have dealt mostly with point charges or uniform one and two dimensional distributions of charge. In real life we want to work out the electric field due to non uniform distributions of charge having a density denoted by

$$\rho(\mathbf{r}) \quad [\mathrm{C\ m^{-3}}]$$

which varies throughout some volume, V, as a function of position vector $\mathbf{r}$. We recognise $\rho(\mathbf{r})$ as a *scalar field*. Ultimately any charge distribution must be a collection of point charges with charge $\pm e^\dagger$ and by the principle of superposition the electric field at a field point $\mathbf{r}_0$ is the sum over all the fields due to the elemental charges that make up the distribution. Here, I want to equip you with the formal mathematics that you will need in later years of your studies in physics.

Let us re-draw figure 7–1 (Lecture 7) for the case when the source charges are distributed into a charge density function $\rho(\mathbf{r})$.

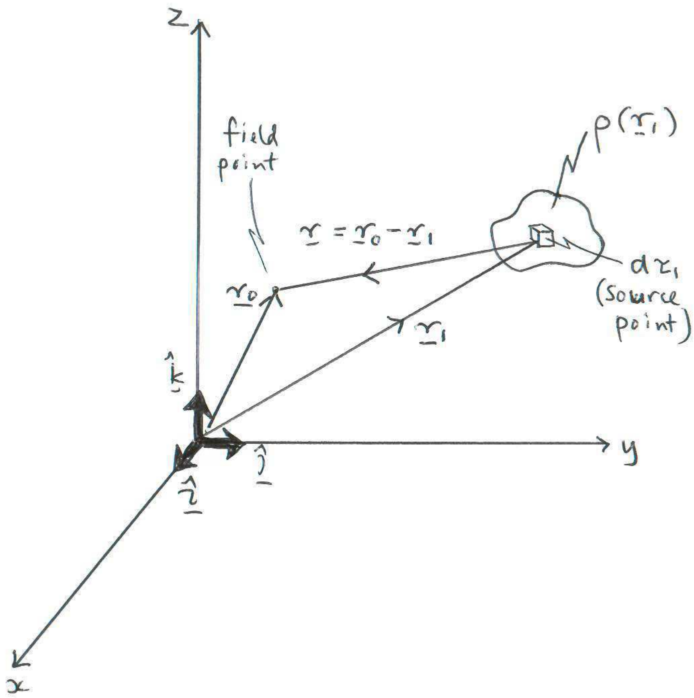


FIGURE 10–4

In the little cube, there is an infinitesimal amount of charge dq so the field at the field point $\mathbf{r}_0$ is

$$\mathrm{d}\mathbf{E}(\mathbf{r}_0) = \frac{1}{4\pi\epsilon_0}\,\frac{\mathrm{d}q}{r^2}\,\hat{\mathbf{r}}$$

† This is the smallest known charge—the charge of one proton. It is true that quarks have a fractional charge but they only appear in combinations of two or three quarks such that the total charge is e or zero; the inability to separate nucleons into individual quarks is called "quark confinement".

The total field at $\mathbf{r}_0$ is the sum of all these increments of field as I reposition the box at all values of $\mathbf{r}_1$ that range over the volume, V, containing the charge distribution.

Summing the field over contributions from each infinitesimal box amounts to an integration, in the sense of the integral calculus. Each box has a volume

$$\mathrm{d}x_1\,\mathrm{d}y_1\,\mathrm{d}z_1 = \mathrm{d}\tau_1$$

and the charge in the box is

$$\mathrm{d}q = \rho(\mathbf{r}_1)\,\mathrm{d}\tau_1$$

so the field at $\mathbf{r}_0$ is

$$\mathbf{E}(\mathbf{r}_0) = \frac{1}{4\pi\epsilon_0}\int_V \frac{\rho(\mathbf{r}_1)}{r^2}\,\hat{\mathbf{r}}\,\mathrm{d}\tau_1 \tag{10.1}$$

Note that this is a tricky integral because the vector $\mathbf{r}$ is not a constant—it depends upon $\mathbf{r}_1$, the position vector of the box $\mathrm{d}\tau_1$. Please note that in electrostatics we always define $\mathbf{r}$ as the vector that points *from* a source point *to* the field point.

In the case of a two dimensional distribution of charge it is conventional to denote the charge density by

$$\sigma(\mathbf{r}) \quad [\mathrm{C\ m^{-2}}]$$

The infinitesimal box becomes an infinitesimal area

$$\mathrm{d}a_1 = \mathrm{d}x_1\,\mathrm{d}y_1$$

and now the electric field at $\mathbf{r}_0$ is

$$\mathbf{E}(\mathbf{r}_0) = \frac{1}{4\pi\epsilon_0}\int_S \frac{\sigma(\mathbf{r}_1)}{r^2}\,\hat{\mathbf{r}}\,\mathrm{d}a_1$$

You may compare this with the integral on page 2, Lecture 9, which is the special case for which the charge density is uniform, namely,

$$\sigma(\mathbf{r}) = \sigma$$

and can be taken as a constant to the front of the integral sign.

If we are interested in a one dimensional distribution of charge such as along a straight wire, we usually denote the density by

$$\lambda(\mathbf{r}) \quad [\mathrm{C\ m^{-1}}]$$

and the box becomes an increment of line length

$$\mathrm{d}l_1 = \mathrm{d}x_1$$

and the electric field at the field point $\mathbf{r}_0$ is

$$\mathbf{E}(\mathbf{r}_0) = \frac{1}{4\pi\epsilon_0}\int_L \frac{\lambda(\mathbf{r}_1)}{r^2}\,\hat{\mathbf{r}}\,\mathrm{d}l_1$$

and this reduces to the integral on page 5, Lecture 8, in the case that the distribution

$$\lambda(\mathbf{r}) = \lambda$$

is uniform.

Lecture 11

11.1 Electric field is irrotational

If all charges in a problem are stationary with respect to each other, we can prove that the electric field is "irrotational".

Consider, again, the electric field due to a point charge,

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\,\frac{q}{r^2}\,\hat{\mathbf{r}}$$

which points radially away from the charge. Let us now examine the *line integral*

$$\int_a^b \mathbf{E}\cdot \mathrm{d}\boldsymbol{\ell}$$

from some point a to some other point b along some path. In figure 11–1 there is a point charge at the origin of the coordinate system.

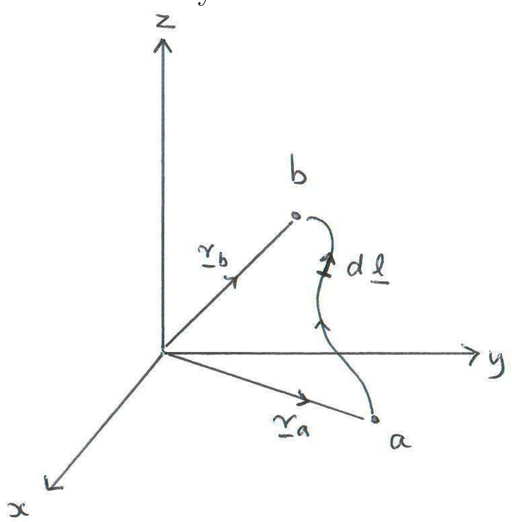


FIGURE 11–1

Whatever the orientation of $\mathrm{d}\boldsymbol{\ell}$, $\mathbf{E}\cdot\mathrm{d}\boldsymbol{\ell}$ is the magnitude of the electric field times the component of $\mathrm{d}\boldsymbol{\ell}$ in the radial direction $\mathrm{d}\mathbf{r}$ pointing to or from the point charge, so

$$\begin{aligned}\mathbf{E}\cdot\mathrm{d}\boldsymbol{\ell} &= \frac{1}{4\pi\epsilon_0}\,\frac{q}{r^2}\,\hat{\mathbf{r}}\cdot\mathrm{d}\boldsymbol{\ell}\\ &= \frac{1}{4\pi\epsilon_0}\,\frac{q}{r^2}\,\mathrm{d}r\end{aligned}$$

Therefore

$$\begin{aligned}\int_a^b \mathbf{E}\cdot\mathrm{d}\boldsymbol{\ell} &= \frac{1}{4\pi\epsilon_0}\int_a^b \frac{q}{r^2}\,\mathrm{d}r\\ &= -\frac{1}{4\pi\epsilon_0}\left[\frac{q}{r}\right]_{r_a}^{r_b}\end{aligned}$$

In the same way that $\mathrm{d}r$ is the radial component of $\mathrm{d}\boldsymbol{\ell}$, so r_a and r_b are the radial components of the vectors defining the points a and b. That is

$$r_a = \hat{\mathbf{r}}_a \cdot \mathbf{r}_a \ ; \ \ r_b = \hat{\mathbf{r}}_b \cdot \mathbf{r}_b$$

By *radial component* of a vector $\mathbf{r}$ I mean the distance from the origin to the point $\mathbf{r}$. Therefore we have,

$$\int_a^b \mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell} = \frac{1}{4\pi\epsilon_0} \left(\frac{q}{r_a} - \frac{q}{r_b} \right)$$

Now suppose that the path were closed, that is we deal with any path having $a = b$. Then,

$$\oint \mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell} = 0$$

I have only proved this for taking a path in the electric field due to a point charge at the origin. But the final result makes no reference to where that origin is. Furthermore this must be true for the electric field due to any distribution of stationary charge, according to the principle of superposition. So this is generally true for any electric field arising from static charges.

When you learn Stokes's theorem you will see that the above result is equivalent to writing

$$\boldsymbol{\nabla} \times \mathbf{E} = 0$$

that is, the curl of $\mathbf{E}$ is zero.

When you come to dealing with moving charges you will find that this formula becomes modified to

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t}$$

where $\mathbf{B}$ is the magnetic field. This is one of "Maxwell's equations" also known as Faraday's law of induction. It says that if you grow or shrink a magnetic field inside a wire loop you will cause a current to flow in the loop. This is how a generator works.

11.2 The electric potential

If I write out

$$\boldsymbol{\nabla} \times \mathbf{E} = 0$$

in full, it looks like this

$$\left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \right) \hat{\mathbf{\imath}} + \left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \right) \hat{\mathbf{\jmath}} + \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \right) \hat{\mathbf{k}} = \mathbf{0}$$

These are partial derivatives of the components of

$$\mathbf{E} = (E_x, E_y, E_z)$$

with respect to components of the position vector $\mathbf{r} = (x, y, z)$ and since the resulting vector is zero, each of its components must be zero and I am left with three equations relating the three components of $\mathbf{E}$ to one another. So in *electrostatics* the $\mathbf{E}$-field is a rather special field because of the relations between E_x, E_y and E_z. Only certain vector fields will be consistent with these interrelations. And what is more, we shouldn't need all three numbers E_x, E_y and E_z to specify an electric field in electrostatics, one will do. Therefore there must be a *scalar* field from which I can uniquely deduce the electric field. How can I find this scalar field?

First we ask, what is the work done in moving a charge through an electric field? Work is force times distance. I do work if I move a positive test charge *towards* a positive charge, that is against the direction of the field lines. Otherwise, if it is negative, the charge does work on my positive test charge. We won't labour this point, in the latter case the work I do is negative. Hence if I move a charge q' a little distance $\mathrm{d}\boldsymbol{\ell}$ in an electric field the work done on the test charge is

$$-\mathbf{F} \cdot \mathrm{d}\boldsymbol{\ell}$$

and if the charge is a unit test charge, one Coulomb in SI units, the work done is

$$-\mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell}$$

since $\mathbf{E}$ is the force per unit charge. We require here a scalar (dot) product because work is only done over that component of the path that is parallel to the field—no work is done by moving a test charge perpendicular to the field lines.

Go back and look at figure 11-1; the work done on a unit test charge in moving from a to b is

$$\begin{aligned} -\int_a^b \mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell} &= -\frac{1}{4\pi\epsilon_0} \int_a^b \frac{q}{r^2}\,\mathrm{d}r \\ &= -\frac{1}{4\pi\epsilon_0}\, q \left(\frac{1}{r_a} - \frac{1}{r_b}\right) \end{aligned} \tag{11.1}$$

You can see why this only depends on the difference in (inverse) radial distance from the point charge at the origin by looking at the two vectors $\mathbf{r}_a$ and $\mathbf{r}_b$,

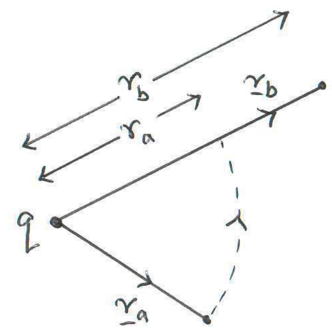


FIGURE 11–2

(In this figure the work done is negative because the test charge ends up further away from q.) No work is done in moving along the dotted arc since the test charge remains at a fixed distance from q.

The work done is *independent of the path taken.* It depends only on the inverse difference between the initial and final radial distances to the point charge. You can appreciate this by using either of the following arguments.

1. Any path consisting of radial segments and arcs will give the same answer. Try drawing a few.
2. If you move the test charge from a to b along one path and then back along a different path, you would not expect to have done any net work. Otherwise you could create a perpetual motion machine.

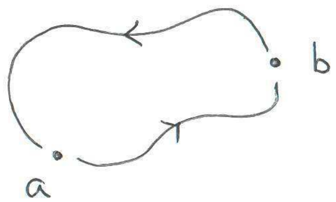


FIGURE 11–3

According to point 2., and figure 11-3, it must be true that

$$\oint \mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell} = 0$$

which in fact we have already proved.

As you see from equation (11.1) the work done is just the difference between two numbers. Let's fix what these numbers are by agreeing that in taking our test charge from a to b we will always stop off at an agreed point P. Then we can define a number $V(\mathbf{r})$ as the work done per unit charge in moving the test charge *from* P *to* $\mathbf{r}$

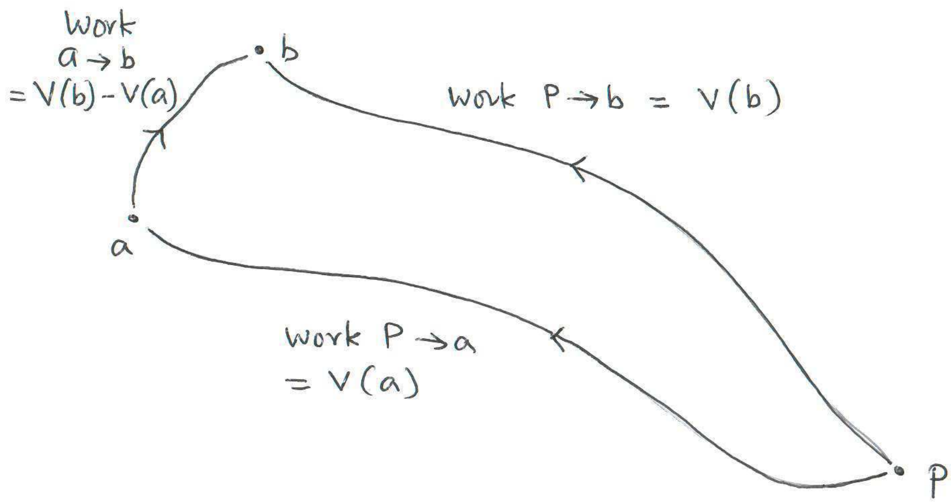


FIGURE 11–4

So we have,

$$V(\mathbf{r}) = -\int_P^{\mathbf{r}} \mathbf{E}\cdot \mathrm{d}\boldsymbol{\ell}$$

which is independent of the path taken. $V(\mathbf{r})$ is called the *electric potential* at point $\mathbf{r}$ and you must remember that it is only uniquely defined once we all agree where P is. Almost always we agree to put P at *infinity* which means away from the influence of any of the charges in the world.

In that case we can say that the electric potential due to a point charge, q, at the origin is

$$\begin{aligned} V(\mathbf{r}) &= -\frac{1}{4\pi\epsilon_0}\int_\infty^{\mathbf{r}} \frac{q}{r^2}\mathrm{d}r \\ &= -\frac{1}{4\pi\epsilon_0}\frac{q}{r} \end{aligned}$$

There's no problem if there are other charges present: by the principle of superposition the electric potential is just the sum of the contributions from each charge. In the case of a charge distribution, $\rho(\mathbf{r})$ [C m^{-3}] existing in a volume Ω,

$$V(\mathbf{r}_0) = \frac{1}{4\pi\epsilon_0}\int_\Omega \frac{\rho(\mathbf{r}_1)}{r}\mathrm{d}\tau_1 \tag{11.2}$$

if $\mathbf{r}_0$ is the field point and $\mathbf{r}_1$ a source point. To get the geometry look again at figure 10–4 and compare this with equation (10.1) in lecture 10, page 5. They're both horrible looking but believe me (11.2) is a lot easier to deal with as V is a scalar, not a vector, and there is no unit vector in the integrand. After all, any equation for a vector quantity like $\mathbf{E}$ is really three equations—one for each component.

The units of electric potential are [N m C^{-1}] (electric field times distance) or Joules per Coulomb. One J C^{-1} is called a *volt* [V].

Be very very careful of confusing *electric potential* with *potential energy*. Electric potential is work (energy) per unit charge: a point charge q' has potential energy relative to infinity of $q'\,V(\mathbf{r})$ [J] when it's placed at position $\mathbf{r}$ where the electric potential is $V(\mathbf{r})$.

Because of the inverse square law there is also of course a *gravitational potential* which is work done against a gravitational field per unit mass.

Now I told you that the basic problem in electrostatics is, there's a bunch of charges (or charge distribution) *over there* what is the electric field *over here*? The answer, if you can work it out, is contained in equation (10.1). *Now* I've told you, don't worrry about equation (10.1) because equation (11.2) is easier. But now you say, how to I get $\mathbf{E}(\mathbf{r})$ if I know $V(\mathbf{r})$?

We obtain the relation between them using the vector form of the *fundamental theorem of calculus*,

$$\int_a^b F\mathrm{d}x = \int_a^b \frac{\mathrm{d}f}{\mathrm{d}x}\mathrm{d}x = f(b) - f(a)$$

if

$$F = \frac{\mathrm{d}f}{\mathrm{d}x}$$

The vector form of this in our instance is

$$V(b) - V(a) = \int_a^b \boldsymbol{\nabla} V \cdot \mathrm{d}\boldsymbol{\ell}$$

where $\boldsymbol{\nabla}$ is the vector with the property that

$$\boldsymbol{\nabla} V(x, y, z) = \frac{\partial V}{\partial x}\hat{\mathbf{i}} + \frac{\partial V}{\partial y}\hat{\mathbf{j}} + \frac{\partial V}{\partial z}\hat{\mathbf{k}}$$

and these are partial derivates. But we saw earlier that

$$V(b) - V(a) = -\int_a^b \mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell}$$

and since the two equations are true for any choice of points a and b and any path between them it must be true that

$$\mathbf{E} = -\boldsymbol{\nabla} V \tag{11.3}$$

That means

$$E_x = -\frac{\partial V}{\partial x}$$
$$E_y = -\frac{\partial V}{\partial y}$$
$$E_z = -\frac{\partial V}{\partial z}$$

so if we know $V(\mathbf{r}) = V(x, y, z)$ then we can get $\mathbf{E}(\mathbf{r})$ by differentiating. You may wonder how can we obtain a *vector*, which is three numbers, at a point $\mathbf{r}$ from a *scalar*, which is one number, at the point $\mathbf{r}$. The point is that it's *not enough* just to know the value of V at $\mathbf{r}$. You need also to know its *derivatives* which means you need to know its values in the *neighbourhood* of $\mathbf{r}$ also.

11.3 The electric potential due to a spherical shell of charge

A spherical shell of radius R carries a uniform surface charge of σ [C m^{-2}].† What is the electric potential relative to infinity at points inside and outside the sphere?

† A solid conductor has all its charge on its surface so this problem also solves the problem of a charged metal sphere.

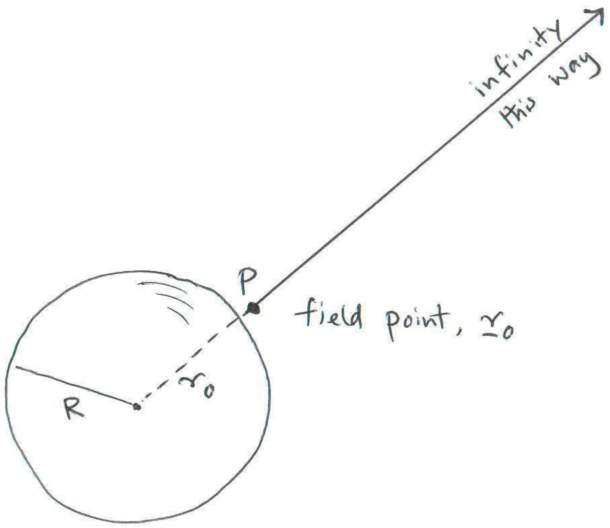


FIGURE 11–5

Step 1. Find the electric field. This must point radially outwards or inwards and can be found using Gauss's law. Imagine a concentric sphere of radius $r > R$. The flux is

$$\Phi = \frac{1}{\epsilon_0} Q_{\text{enclosed}}$$

Now,

$$Q_{\text{enclosed}} = 4\pi R^2 \, \sigma \equiv Q$$

so

$$\Phi = EA$$

becomes

$$\frac{1}{\epsilon_0} \, Q = E \, 4\pi r^2$$

and since the field is radial

$$E(r) = \frac{1}{4\pi\epsilon_0} \frac{Q}{r^2} \quad \longleftarrow \quad r > R$$

outside the sphere. *Inside* the sphere the electric field is zero because any gaussian sphere having radius less than R encloses no charge.

Step 2. So, the electric potential *outside* the sphere is

$$\begin{aligned} V(r) &= -\int_\infty^r \mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell} \\ &= -\frac{1}{4\pi\epsilon_0} \int_\infty^r \frac{q}{r'^2} \, \mathrm{d}r' \\ &= \frac{1}{4\pi\epsilon_0} \frac{Q}{r} \end{aligned}$$

(Please don't be confused but I have put a "prime" on the r in the integrand since that is a dummy variable, and r itself appears in the limit to the integral—it is the place where we want to find the potential.)

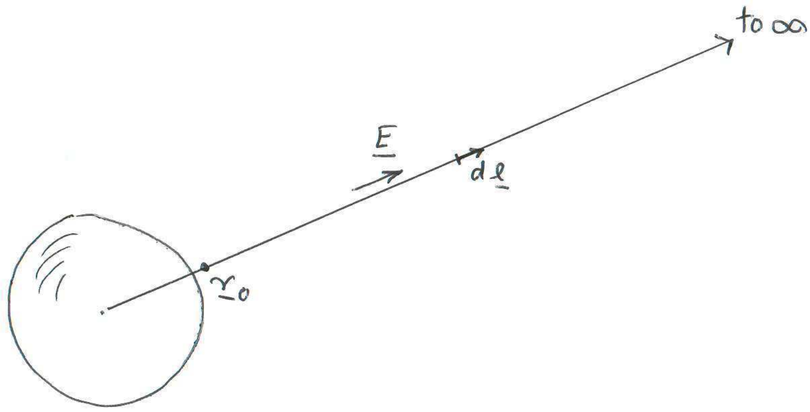


FIGURE 11–6

Although the electric *field* inside the shell is zero, the electric *potential* is not. The question is, how much work do I have to do on a unit test charge to move it from infinity to inside the shell and to get it there I have to move it through the electric field that is outside the shell. I need to do the integral in two pieces,

$$\begin{aligned} V(r) &= -\frac{1}{4\pi\epsilon_0}\int_\infty^R \frac{Q}{r'^2}\,\mathrm{d}r' - \frac{1}{4\pi\epsilon_0}\int_R^r (\text{zero})\,\mathrm{d}r' \\ &= \frac{1}{4\pi\epsilon_0}\frac{Q}{R} \qquad \longleftarrow \quad r > R \end{aligned}$$

The electric potential inside the shell is constant: once I've got my test charge inside I can move it around without having to do any work because the field inside is zero.

It is very instructive, indeed obligatory, in problems of this type to *plot* E and V as functions of distance, in this case from the centre of the shell.

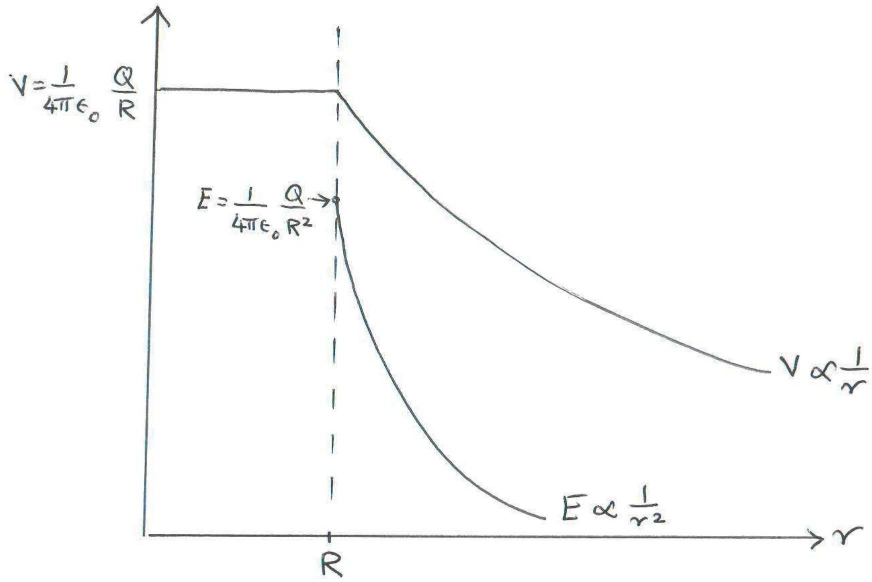


FIGURE 11–7

Lecture 12

12.1 Electric dipole revisited

In Lecture 8 we struggled to find the electric field due to a dipole and ended up settling for solutions only for field points either in line with or perpendicular to the dipole. Now I'll show you how to get a general solution using the electric potential. Here is the dipole again, illustrated in figure 12–1.

The vector $\mathbf{p} = q\mathbf{d}$ is again the dipole moment vector, and remember it points *from* the negative *to* the positive charge by convention; and we'll work to *first order* in d^2/r^2, that is we only treat the case $r \gg d$ meaning that we cannot allow the field point P to get too close. This problem is sometimes called an "ideal" as opposed to a "physical" dipole.

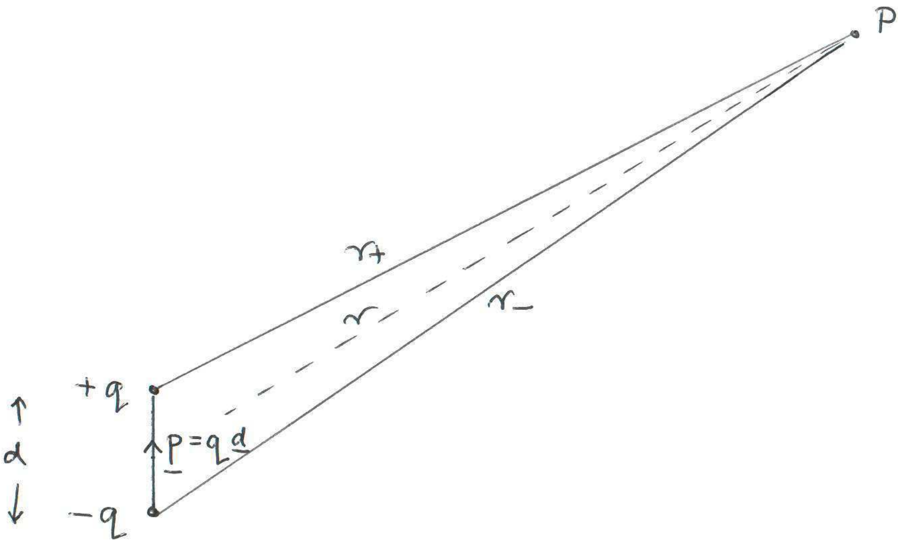


FIGURE 12–1

The electric potential at P, using the superposition principle is,

$$V = \frac{1}{4\pi\epsilon_0}\, q \left(\frac{1}{r_+} - \frac{1}{r_-}\right)$$

We use the cosine rule to write

$$\begin{aligned}
r_-^2 &= r^2 + \left(\frac{1}{2}d\right)^2 + rd\cos\theta \\
\frac{r}{r_-} &= \left(1 + \frac{d^2}{4r^2} + \frac{d}{r}\cos\theta\right)^{-\frac{1}{2}} \\
&= 1 - \frac{1}{2}\left(\frac{d^2}{4r^2} + \frac{d}{r}\cos\theta\right) + \frac{3}{8}\left(\frac{d^2}{4r^2} + \frac{d}{r}\cos\theta\right)^2 + \ldots
\end{aligned}$$

In the third line the square root is expanded in a Taylor series. Neglecting terms smaller than d^2/r^2 this comes out to

$$\frac{r}{r_-} = 1 - \frac{d}{2r}\cos\theta + \frac{d^2}{8r^2}\left(3\cos^2\theta - 1\right)$$
$$\frac{r}{r_+} = 1 + \frac{d}{2r}\cos\theta + \frac{d^2}{8r^2}\left(3\cos^2\theta - 1\right)$$

and subtracting these and subsituting into the above equation for V,

$$V = \frac{1}{4\pi\epsilon_0}\, qd\, \frac{1}{r^2}\, \cos\theta$$

You can see from figure 12–1 that

$$qd\cos\theta = p\cos\theta = \mathbf{p}\cdot\hat{\mathbf{r}}$$

so

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\,\frac{1}{r^2}\mathbf{p}\cdot\hat{\mathbf{r}} \tag{12.1}$$

Note that the potential falls away like $1/r^2$ compared to that due to a single point charge which falls off like $1/r$. The electric *fields* remember decay like $1/r^3$ and $1/r^2$ respectively.

So how do we get the field from the potential? We need to calculate

$$\mathbf{E} = -\boldsymbol{\nabla} V$$

by doing partial differentiation (see Lecture 11, page 6 and your Mathematical Notes on `KEATS`). It's not completely straightforward so I'll put the maths into the additional material on `KEATS`. The result is

$$\mathrm{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\,\frac{1}{r^3}\left[3\left(\mathbf{p}\cdot\hat{\mathbf{r}}\right)\hat{\mathbf{r}} - \mathbf{p}\right] \tag{12.2}$$

As always, $\mathbf{r}$ points *towards* the field point and $\hat{\mathbf{r}}$ is its dimensionless unit vector. I hope you can see that the results we got on pages 3 and 4 in Lecture 8, are consistent with equation (12.2). In the first case, equation (8.1), $\mathbf{p}\cdot\hat{\mathbf{r}}$ is zero, and in the second case, equation (8.2), it is p so that $(\mathbf{p}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}} = \mathbf{p}$.

12.2 Electrostatic energy

Associated with a collection of point charges is an *electrostatic energy* which is the work done in assembling this collection of point charges, each being brought in from infinity where the electric potential is zero. Suppose I put a point charge of amount q_1 at $\mathbf{r}_1$, q_2 at $\mathbf{r}_2$ and so on. To add q_2 at $\mathbf{r}_2$ when q_1 is already at $\mathbf{r}_1$ requires an amount of work,

$$\frac{1}{4\pi\epsilon_0}\, q_2\, \frac{q_1}{r_{12}}$$

which is the charge q_2 times the electric potential at $\mathbf{r}_2$ due to the charge q_1 at $\mathbf{r}_1$.

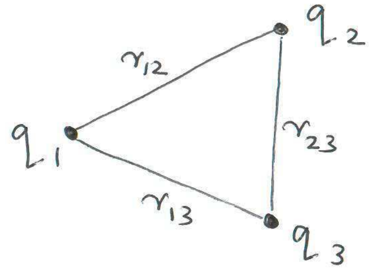


FIGURE 12–2

To add the charge q_3 I need to do this much more work:

$$\frac{1}{4\pi\epsilon_0}\, q_3 \left(\frac{q_1}{r_{13}} + \frac{q_2}{r_{23}}\right)$$

To add the next charge, q_4 at $\mathbf{r}_4$, the work I do is

$$\frac{1}{4\pi\epsilon_0}\, q_4 \left(\frac{q_1}{r_{14}} + \frac{q_2}{r_{24}} + \frac{q_3}{r_{34}}\right)$$

The total work done is the sum of all these terms,

$$\frac{1}{4\pi\epsilon_0}\left(\frac{q_1q_2}{r_{12}} + \frac{q_1q_3}{r_{13}} + \frac{q_2q_3}{r_{23}} + \frac{q_1q_4}{r_{14}} + \frac{q_2q_4}{r_{24}} + \frac{q_3q_4}{r_{34}}\right)$$

Note that by working with the electric *potential*, not the electric *field*, we don't need to deal with vectors.

If I want to assemble N charges then I hope that you can see that the work I have to do can be written using "summation signs" as

$$W = \frac{1}{4\pi\epsilon_0}\,\frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{q_iq_j}{r_{ij}}$$

and the factor of a half appears because I've counted every pair twice in the double sum. Now, I can write this in a different way,

$$W = \frac{1}{2}\sum_{i=1}^{N} q_i \left(\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{1}{4\pi\epsilon_0}\,\frac{q_j}{r_{ij}}\right)$$

but the sum in parentheses is the potential at the point $\mathbf{r}_i$ due to all the other charges. We say that this is the electric potential, $V(\mathbf{r}_i)$, "seen by" q_i.

Therefore

$$W = \frac{1}{2}\sum_{i=1}^{N} q_i\, V(\mathbf{r}_i) \tag{12.3}$$

Now if I had a continuous distribution of charge ρ [C m^{-3}] then for q_i I would write an infinitesimal bit of charge density, $\rho(\mathbf{r}_i)\mathrm{d}\tau_i$ (see Lecture 10, section 10.2, p. 5).† I will now have an energy associated with this distribution of charge that is

$$W = \frac{1}{2}\int \rho V \mathrm{d}\tau \tag{12.4}$$

where you remember that an integral is just a sum over infinitesimal elements.

Now there's some difficult mathematics, but a very significant result emerges. Using the differential form of Gauss's law (see additional material at KEATS) $\boldsymbol{\nabla}\cdot\mathbf{E} = \rho/\epsilon_0$,

$$\begin{aligned} W &= \frac{1}{2}\epsilon_0 \int_\Omega (\boldsymbol{\nabla}\cdot\mathbf{E})\ V\,\mathrm{d}\tau \\ &= \frac{1}{2}\epsilon_0 \left\{ -\int_\Omega \mathbf{E}\cdot(\boldsymbol{\nabla}V)\,\mathrm{d}\tau + \oint_S V\,\mathbf{E}\cdot\mathrm{d}\mathbf{a} \right\} \\ &= \frac{1}{2}\epsilon_0 \left\{ \int_\Omega E^2\,\mathrm{d}\tau + \oint_S V\,\mathbf{E}\cdot\mathrm{d}\mathbf{a} \right\} \end{aligned}$$

I am using Ω for volume rather than V in this lecture so as not to confuse it with the electric potential. Going from the first to the second line I have used integration by parts, and to get to the third line from the second I have used equation (11.3) from Lecture 11, namely,

$$\boldsymbol{\nabla}V = -\mathbf{E}$$

Now, S is the surface bounding the volume Ω. I can assume for the present purposes that my charge distribution is isolated in a volume infinitely far from any other charges so I can make that volume big enough so that the potential vanishes at its surface. In that case the second term is zero and taking what remains we arrive at this insight provoking result,

$$W = \frac{1}{2}\epsilon_0 \int_{\substack{\text{all}\\ \text{space}}} E^2\,\mathrm{d}\tau \tag{12.5}$$

The electrostatic energy is evidently contained *within* the field itself, or if you will, it is stored in the empty space surrounding the charge distribution.

12.3 Electrostatic energy of a charged capacitor

Two plates separated by vacuum have area A, separation d, and are charged to $+Q$ and $-Q$ respectively. The surface charge density is

$$\sigma = \frac{Q}{A} \qquad [\mathrm{C\ m^{-2}}]$$

† At this point I am cheating you and I'll come back later in the lecture to show you how you've been misled. I'll be surprised if you can spot right now where the inconsistency lies in this paragraph.

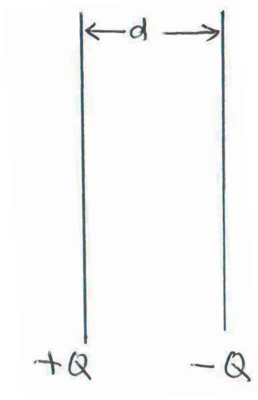


FIGURE 12–3

The electric field is constant between the plates and zero outside the plates. Inside the field is

$$E = \frac{\sigma}{\epsilon_0} = \frac{1}{\epsilon_0}\frac{Q}{A}$$

(See Lecture 10, p. 3.) The potential difference is

$$\begin{aligned} V &= -\int_{\text{right plate}}^{\text{left plate}} \mathbf{E}\cdot \mathrm{d}\boldsymbol{\ell} = -E\int_{\text{right plate}}^{\text{left plate}} \mathrm{d}\ell = Ed \\ &= \frac{1}{\epsilon_0}\frac{Q}{A}\,d = \frac{Q}{C} \end{aligned}$$

where $C = Q/V$ is the *capacitance*. In this case

$$C = \frac{A\epsilon_0}{d} \qquad [\mathrm{C\ V^{-1}}]$$

One Coulomb per volt is called a *farad*. The electrostatic energy is

$$\begin{aligned} W &= \frac{1}{2}\epsilon_0 \int_{\substack{\text{inside}\\ \text{the plates}}} E^2\,\mathrm{d}\tau \\ &= \frac{1}{2}\epsilon_0\,E^2 \int_{\substack{\text{inside}\\ \text{the plates}}} \mathrm{d}\tau \\ &= \frac{1}{2}\epsilon_0 E^2 Ad \end{aligned}$$

because the last integral is just the volume Ad inside the capacitor. Hence

$$W = \frac{1}{2}\frac{1}{\epsilon_0}Q^2\frac{d}{A} = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}QV$$

The purpose of the capacitor is to store charge, not energy; so a high capacitance, Q/V, means a large charge is stored using a low voltage. We can increase the capacitance by replacing vacuum with a material of large relative permittivity ϵ_r to

$$C = \frac{A\epsilon_0\epsilon_r}{d}$$

12.4 Electrostatic energy of a charged spherical shell

Let us go back to the problem of Lecture 11, section 11.3, p. 6. What is the electrostatic energy of the spherical shell of radius R carrying a charge Q? We need to square the electric field and integrate over all space. The field inside the shell is zero so we just need to consider the space outside. We must be careful to do it properly. Divide the space into concentric infinitesimal shells of radius r and thickness $\mathrm{d}r$; these have a volume of $4\pi r^2 \mathrm{d}r$ (area $\times$ thickness). In each of these shells the field is uniform (because they're infinitesimally thin) and has the value

$$E(r) = \frac{1}{4\pi\epsilon_0} \frac{Q}{r^2}$$

We want E^2 times the elemental volume of the shell integrated from R to ∞,

$$\begin{aligned} W &= \frac{1}{2}\epsilon_0 \int_R^\infty \frac{Q^2}{16\pi^2\epsilon_0^2} \frac{1}{r^4} 4\pi r^2 \,\mathrm{d}r \\ &= \int_R^\infty \frac{Q^2}{8\pi\epsilon_0} \frac{1}{r^2} \,\mathrm{d}r \\ &= \frac{Q^2}{8\pi\epsilon_0} \left[-\frac{1}{r} \right]_R^\infty \\ &= \frac{1}{4\pi\epsilon_0} \frac{1}{2} \frac{Q^2}{R} \end{aligned} \tag{12.6}$$

Now suppose I shrink the shell by reducing the radius R while at the same time increasing the surface charge density σ [C m^{-2}] in such a way that the total charge remains equal to Q. In the limit of $R \to 0$ I obtain a *point charge*—the subject of much of these lectures. So I evaluate equation (12.4) using $R = 0$ and I get the energy contained within the field of a point charge,

$$W = \infty$$

What the hell is happening here? Another thing may have struck you as odd. According to equation (12.5) the electrostatic energy is always positive; but if I bring a negative point charge up to a positive point charge I will gain energy from the work that the charge does for me. W should be able to have either sign!

The difficulty can be *partially* resolved. In going from equation (12.3) to (12.4) I did something sneaky. I wrote

$$q_i \longrightarrow \rho(\mathbf{r}_i)\mathrm{d}\tau_i$$

I smeared all the point charges into a distribution so that equation (12.4) contains all the energy to assemble the charge distribution whereas equation (12.3) accepts the point charges as given and expresses just the energy to move them about. This is evidently just as well if you are dealing with point charges since the energy to make one is apparently infinite. In that sense the difference in numerical value between equations (12.4) and (12.3) is an infinite number! Which equation is actually the "right" one?

This point is very hard to resolve. Take the electron. It's probably a point charge—at least no-one has yet measured its radius, and if that were not zero then surely we could break the electron into pieces; but no-one has managed to do that either. We explore this matter a bit more in the tutorials.

Lecture 14

14.1 The Lorentz force

Suppose I position two point charges, q_a and q_b; q_a at the origin O_2 of a cartesian coordinate system, and q_b at some point (x_2, y_2) in the x_2–y_2-plane of the same coordinate system.

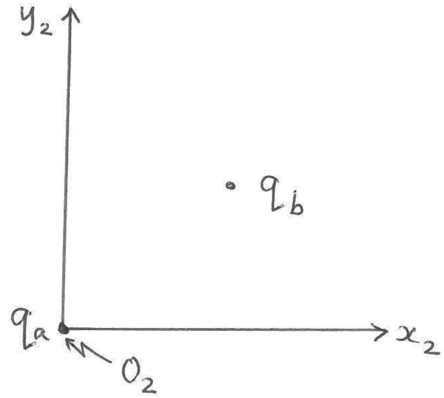


FIGURE 14–1

What's the force acting between them? Easy. Let's treat q_b as the test charge and q_a as the source charge. Then the force acting on q_b is

$$q_b\,\mathbf{E}_a$$

where $\mathbf{E}_a$ is the electric field at (x_2, y_2) due to the source charge q_a.

Now suppose that both charges are moving with a constant velocity $\mathbf{u}$ in the x_2-direction. What is now the force on q_b? Measured by an observer also moving at velocity $\mathbf{u}$ the force is the same, $q_b\mathbf{E}_a$, because according to the principle of special relativity it is not possible to devise an experiment to detect the uniform motion of a frame of reference by making measurements within that frame.

But suppose I make the observation of the force on q_b from within an inertial frame with origin O_1 with respect to which O_2 is travelling at constant velocity $\mathbf{u}$ in the x_2-direction. Let us call O_1 the *laboratory* frame. Then I could be observing two electrons leaving a field emitter or electron gun and travelling parallel to each other at a velocity $\mathbf{u}$.

I will not go through the mathematics. Consult chapters 11–13 of "Electromagnetic Phenomena" or chapters 13–17 of "Electromagnetic Fields and Waves" (chapters 5–6 in the second edition) by Lorrain, Corson and Lorrain. The point is that you have to make a Lorentz transformation of the force

$$\mathbf{F} = \frac{\mathrm{d}}{\mathrm{d}t}(m\mathbf{u})$$

the rate of change of relativistic momentum. This involves transforming from frame O_2 into frame O_1. The result is <u>astonishing</u>. In addition to the electrostatic force $q_b\mathbf{E}_a$, there arises a force on the charge q_b that acts in a direction *perpendicular* to the trajectory of q_a.

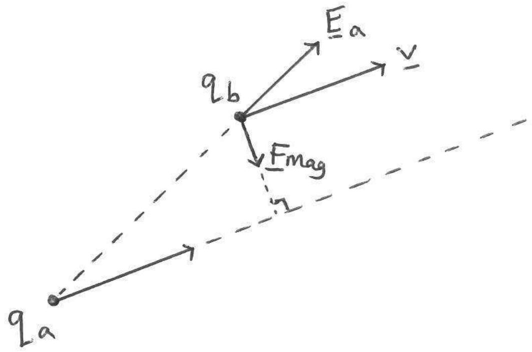


FIGURE 14–2

We may regard this additional *force* as the effect of an additional *field* whose field lines encircle the trajectory of the charge q_a.

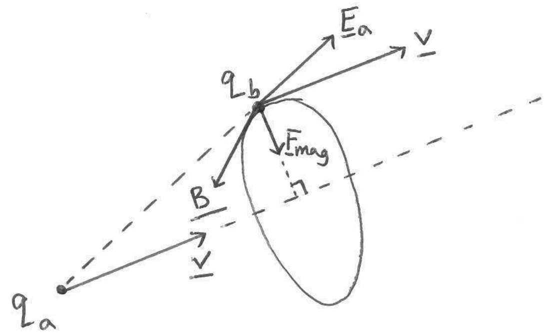


FIGURE 14–3

Viewed from the laboratory the force acting on the charge q_b is

$$\begin{aligned}\mathbf{F}_b &= q_b\,\mathbf{E}_a + q_b\,(\mathbf{u}\times\mathbf{B})\\ &= \mathbf{F}_{\text{electric}} + \mathbf{F}_{\text{magnetic}}\end{aligned}$$

and this introduces the so called *magnetic field*, $\mathbf{B}$, sometimes also called the *magnetic induction.*

This case can be generalised to one in which the charge q_a is fixed in frame O_2 and the charge q_b is moving relative to q_a. The moving charge, in this case q_a, generates a magnetic field. This field will act upon another moving, test, charge, in this case q_b, to produce a "magnetic" force, $\mathbf{F}_{\text{mag}}$ that is perpendicular to both $\mathbf{B}$ and the velocity $\mathbf{v}$, which is the velocity of q_b as observed in the laboratory frame.

The total force acting upon a point charge, q, moving in electric and magnetic fields is called the *Lorentz force*,

$$\mathbf{F} = q\,(\mathbf{E} + \mathbf{v}\times\mathbf{B})$$

Study this formula carefully. The *electric force* $q\mathbf{E}$ on a test charge in electric and magnetic fields acts in a direction *parallel* to the electric field, but the *magnetic force*

$q\mathbf{v} \times \mathbf{B}$ acts in a direction perpendicular to the magnetic field *and* perpendicular to the velocity of the test charge. Moreover the force due to the magnetic field depends on the *speed* of the test charge. Technically we say that the magnetic force is *non conservative* that means it cannot be written as the gradient of a potental energy; also because the magnetic force is perpendicular to a particle's velocity *magnetic forces cannot do work.*

A simple manifestation of the magnetic force is the fact that two current carrying wires will repel each other,

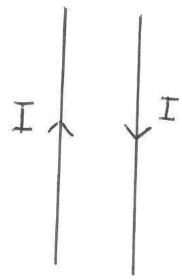


FIGURE 14–4

or attract each other,

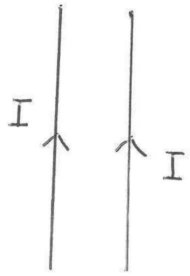


FIGURE 14–5

which is easily demonstrated in a laboratory. Of course the force is on the *electrons* not on the wire itself. We have to imagine that the electrons try to move in the direction of the force and by the electrostatic force between them and the nuclei of the atoms in the metal these nuclei have to come along for the ride and the whole wire then moves.

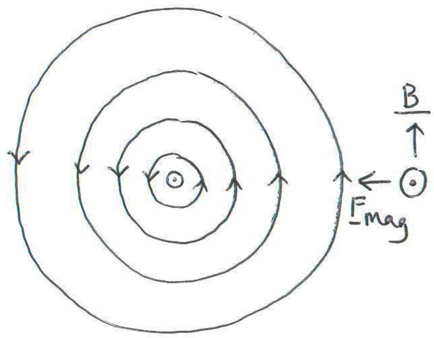


FIGURE 14–6

Around each wire is a circulating magnetic field. Moving charges in the other wire experience a magnetic force

$$\mathbf{F}_{\rm mag} = q\mathbf{v} \times \mathbf{B} \tag{14.1}$$

Just for clarity, in figure 14–6 I've not drawn the magnetic field due to the right hand wire. Note the convention that a tiny circle with a "dot" in it refers to a current flowing towards you (as if you were seeing the point of an arrow approaching) and a tiny circle with a cross in it denotes the opposite (as if you were observing the feathers of a receding arrow). In figure 14–6 I show a current flowing towards you—that is a current of positive charge ($q > 0$); you always have to be careful because thanks to an unfortunate misunderstanding over a hundred years ago, electrons are assigned a negative charge. So regrettably the usual carriers of current in a metal wire move in the opposite direction to the current as defined conventionally.

Please note also that the "right hand rule" that you have no doubt learned is a picture of the vector cross product in a right handed cartesian system of coordinates.

14.2 Current and current density

In a sense, *current* in *magnetostatics* playes the role of *charge* in *electrostatics*. Electrostatics is the study of relatively stationary charges; "magnetostatics" by analogy is the study of fixed steady currents. When we become interested in time-varying charges and currents we appeal to the subject called "electrodynamics". By *current* I mean the amount of positive charge that passes a given point in unit time. A flow of negative charge is equivalent to a flow of positive charge in the reverse direction. Current is measured in the unit of Coulomb per second, or ampere ("A" or "amp"). One amp is one C s^{-1}.

If a line of charge of density λ C m^{-1} travels down a wire at a speed v then the current is

$$I = \lambda v \qquad [\mathrm{A}]$$

Even at constant speed the velocity $\mathbf{v}$ may vary along the wire, unless of course the wire is straight.

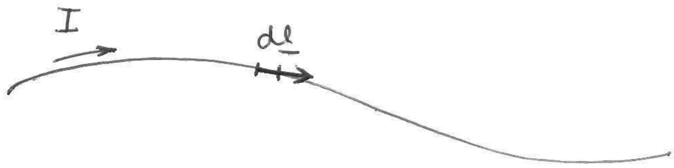


FIGURE 14–7

Figure 14–7 shows an "element of line length" d$\boldsymbol{\ell}$ having dimensions of length which we will be using a lot in the integral calculus that follows. You may think of it as you do the increment dx in the differential and integral calculus, but here we need to account for the possible changes in direction over the interval of integration. This leads us to the concept of a *line integral*, which you may wish to study separately, but will be illustrated in these lectures by way of examples.

So, the current in a wire is really a vector,

$$\mathbf{I} = \lambda \mathbf{v}$$

If a current carrying wire is placed in a magnetic field then each segment of wire behaves as does a moving charge and experiences a Lorentz force.

The segment of vector length of wire $\mathrm{d}\boldsymbol{\ell}$ carries a charge

$$\mathrm{d}q = \lambda \, \mathrm{d}\ell$$

and this charge is moving with a velocity $\mathbf{v}$ in the direction of $\mathrm{d}\boldsymbol{\ell}$. So the magnetic force acting on the segment of wire is

$$\mathrm{d}\mathbf{F}_{\mathrm{mag}} = \mathrm{d}q \, (\mathbf{v} \times \mathbf{B})$$

The total magnetic force on the wire is this, integrated along the length of the wire,

$$\begin{aligned} \mathbf{F}_{\mathrm{mag}} &= \int (\mathbf{v} \times \mathbf{B}) \, \mathrm{d}q \\ &= \int (\mathbf{v} \times \mathbf{B}) \, \lambda \, \mathrm{d}\ell \\ &= \int (\mathbf{I} \times \mathbf{B}) \, \mathrm{d}\ell \end{aligned} \tag{14.2}$$

The current is usually constant along the wire and so

$$\mathbf{F}_{\mathrm{mag}} = I \int (\mathrm{d}\boldsymbol{\ell} \times \mathbf{B})$$

Normally a wire has some thickness. *Current density* is the charge crossing a unit area in unit time. Consider an element of length $\mathrm{d}\boldsymbol{\ell}$ of a wire of cross sectional area a carrying a uniform mobile charge density ρ moving at a constant speed v in the direction of $\mathrm{d}\boldsymbol{\ell}$.

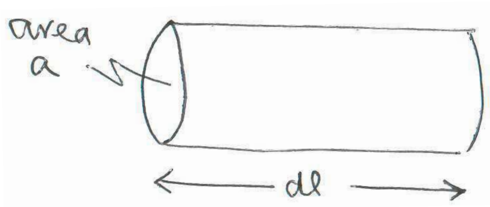


FIGURE 14–8

The object in figure 14–8 is a volume element and for these we always use the notation from the integral calculus $\mathrm{d}\tau = \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$ so the infinitesimal volume

$$\mathrm{d}\tau = a \, \mathrm{d}\ell$$

contains a charge

$$\mathrm{d}q = \rho \, \mathrm{d}\tau = \rho \, a \, \mathrm{d}\ell \qquad \text{(density} \times \text{volume)}$$

This charge passes through the end-surface of area a in a time

$$\mathrm{d}t = \frac{1}{v}\,\mathrm{d}\ell$$

so the current is

$$I = \frac{\mathrm{d}q}{\mathrm{d}t} = \frac{v}{\mathrm{d}\ell}\,\rho\, a\, \mathrm{d}\ell = \rho v a$$

Hence we can deduce a current per unit area vector, having magnitude I/a, that we call $\mathbf{J}$, the *current density*,

$$\mathbf{J} = \rho\,\mathbf{v}$$

If the element of wire is placed in a magnetic field $\mathbf{B}$ the magnetic force on the whole wire, in comparison with equation (14.2), is the integral

$$\begin{aligned}\mathbf{F}_{\mathrm{mag}} &= \int (\mathbf{v}\times\mathbf{B})\,\mathrm{d}q \\ &= \int (\mathbf{v}\times\mathbf{B})\,\rho\,\mathrm{d}\tau \\ &= \int (\mathbf{J}\times\mathbf{B})\,\mathrm{d}\tau\end{aligned}$$

over all the volume elements $\mathrm{d}\tau$ in the wire.

Lecture 15

15.1 The law of Biot and Savart

The Lorentz force formula tells us the force on a moving charge or current carrying wire in a magnetic field. But the magnetic field is itself the product of moving charges or currents in a wire.† So if I know the current, what is the associated magnetic field? I will answer this in this lecture only for the case of *steady* currents, that is **I** and **J** do not vary in time; this is the study of magnetostatics.

The magnetic field due to a steady current **I** is (see figure 15–1)

$$\mathbf{B}(\mathbf{r}_0) = \frac{\mu_0}{4\pi} \int \frac{1}{r^2} \left(\mathbf{I} \times \hat{\mathbf{r}}\right) \mathrm{d}\ell_1$$
$$= \frac{\mu_0}{4\pi} I \int \frac{1}{r^2} \mathrm{d}\boldsymbol{\ell}_1 \times \hat{\mathbf{r}} \tag{15.1}$$

and this is called the law of *Biot and Savart.* You may think of this as a phenomenological statement as you do Coulomb's law which is the result of many observations. This is of course how many great edifices in physics are built: gravitation, thermodynamics, electrodynamics; each a set of theories and conclusions that allow us to *predict* events in the physical world constructed from a very small set of *postulates.* For example thermodynamics is a very powerful discipline based on its famous laws which themselves are regarded as either postulates or generalisations of our experience. On the other hand the Biot–Savart law follows directly from a Lorentz transformation of the force as does the Lorentz force itself. I have said that *all* electrodynamics follows from Coulomb's law and the principle of superposition and I stand by this as long as you are happy to invoke also the Lorentz transformation of the force. It is as Feynman emphasises in his Lectures, "magnetism is a relativistic consequence of electricity." Of course you've not studied relativity yet and so I am happy that you follow many textbooks and believe that the magnetic force is another, almost unrelated, force additional to the Coulomb force which is dictated by an experimental law, namely the Biot–Savart law, equation (15.1). On the other hand I want you to be aware that electricity and magnetism are manifestations of the same phenonenon, and to me at least, the fundamental underlying laws are as I keep stating, Coulomb's law and the principle of superposition. Unfortunately I cannot take you through all the steps to arrive at the Lorentz force and the Biot–Savart law just because you have yet to discover special relativity and also because of lack of time. My favorite place for learning this is one of the books by Lorrain and Corson, partly because they put their chapters on relativity *in between* those of electricity and magnetism rather than at the end of the book as does Griffiths. On the other hand, as always, Griffiths's account of special relativity is excellent. Maybe as an interim measure if you are feeling a little faint hearted, study the chapter on special relativity in your first year textbook "Principles of Physics" at least up to the transformation of momentum.

† This reciprocity, you will see later, is captured in the symmetry between Faraday's law and Ampère's law.

Now, to return to the Biot–Savart law, equation (15.1), the constant μ_0 is

$$\mu_0 = \frac{1}{\epsilon_0 c^2} = 4\pi \times 10^{-7} \qquad [\mathrm{N\ A^{-2}}]$$

The units of $\mathbf{B}$ come out as [N $\mathrm{A^{-1}}$ $\mathrm{m^{-1}}$]. One N $\mathrm{A^{-1}}$ $\mathrm{m^{-1}}$ is called a *Tesla* [T]. You may wonder why we need the extra constant μ_0 as we already have ϵ_0 from getting the SI units into Coulomb's law and we already have the speed of light, c. Well, I'm sorry you'll have to keep on wondering. It gets a little clearer when you study electric and magnetic fields in matter. In these lectures we treat these fields only in the vacuum. Of course the appearance of the constant c reminds you that we are dealing with relativity, but we conceal this in the SI system of units by using mostly ϵ_0 and μ_0 in our formulas. ϵ_0 is often called the "permittivity of free space", and as we say briefly in Lecture 12, p. 5, this is multiplied in matter by the material's "relative permittivity" ϵ_r to arrive at the permittivity ϵ of the material. In similar vein μ_0 is called the permeability of free space and in matter we deal with the permeability μ of a material.

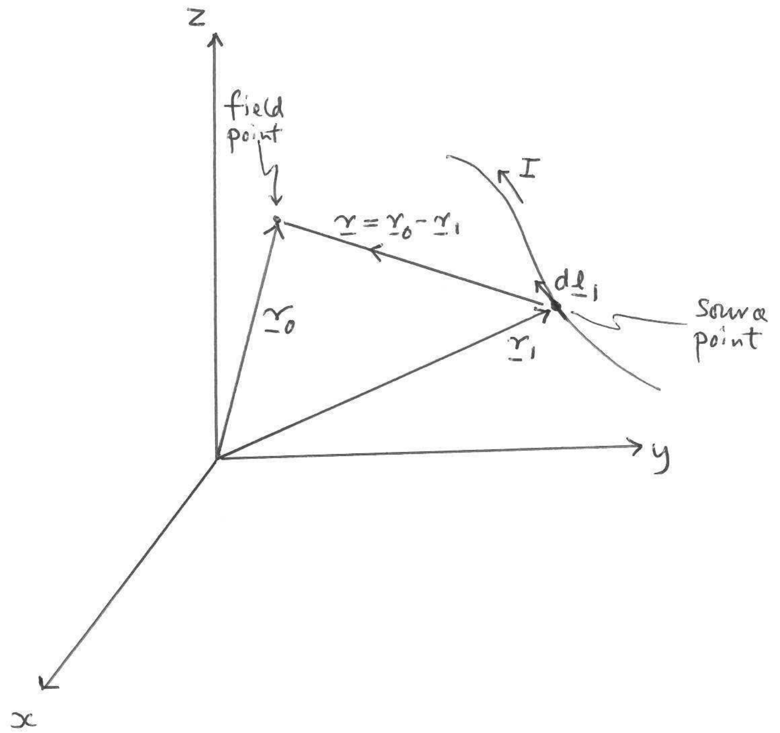


FIGURE 15–1

Next, let us be *absolutely clear* about equation (15.1). We have a current flowing through a line element of wire (or space) $d\boldsymbol{\ell}_1$ and we put on a subscript "1" as we do in figure 10–4 in on p. 4 of Lecture 10 as this labels a *source point*, $\mathbf{r}_1$. We then ask for the magnetic field at a *field point* $\mathbf{r}_0$ which as we did in figure 10–4 under our study of Coulomb's law we indicate with a subscript "0". Again as in Lectures 7 and 10 we define the vector $\mathbf{r}$ as the vector that points *from* the source point $\mathbf{r}_1$ *to* the field point, $\mathbf{r}_0$. To find the magnetic field at the field point I then have to integrate equation (15.1) along the whole length of the wire. As in electrostatics this integral is in general difficult to do because the vector $\mathbf{r}$ is not a constant—it depends on the position vector of the varying source point $\mathbf{r}_1$.

Note the similarity between the Biot–Savart and Coulomb laws, in particular the inverse square dependence on the distance between field and source points. In SI units magnetic field is "force per unit charge per unit velocity," whereas electric field remember is force per unit charge.

15.2 Magnetic field due to a long straight current carrying wire

We now use the Biot–Savart law to deduce the magnetic field due to a long straight current carrying wire.

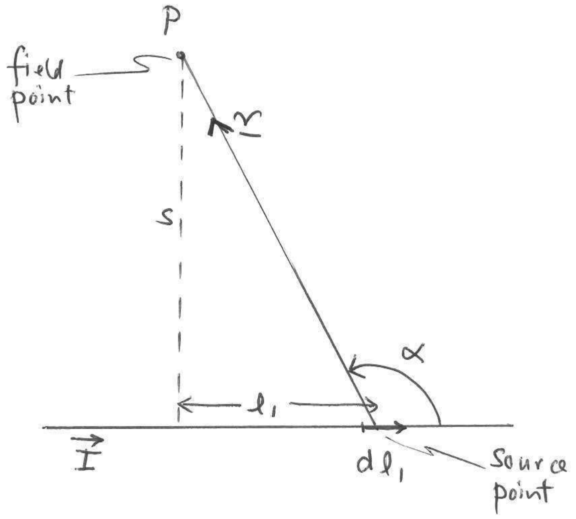


FIGURE 15–2

In figure 15–2, P is a field point, a perpendicular distance s from the wire. The magnetic field at P is in the direction of $\mathrm{d}\boldsymbol{\ell}_1 \times \hat{\mathbf{r}}$ and so points *out* of the page above the wire (and into the page below the wire); you can use the right hand rule also to get the directions. As we have done in electrostatic problems we try and avoid the full vector algebra by deducing the *direction* of the field from the geometry of the problem leaving us only to find the *magnitude* of the field. Now the magnitude of $\mathrm{d}\boldsymbol{\ell}_1 \times \hat{\mathbf{r}}$ is

$$\mathrm{d}\ell_1 \sin\alpha = \mathrm{d}\ell_1 \cos\theta$$

since the magnitude of $\hat{\mathbf{r}}$ is one, and using some trigonometry we can also see that

$$\ell_1 = s\tan\theta$$

and that by differentiating this we can arrive at

$$\mathrm{d}\ell_1 = \frac{s}{\cos^2\theta}\,\mathrm{d}\theta$$

Also from figure 15–2 we can see that

$$s = r\cos\theta$$

which means, after squaring this, that

$$\frac{1}{r^2} = \frac{1}{s^2}\cos^2\theta$$

Putting all this into the Biot–Savart law, equation (15.1) we find the magnitude of the magnetic field,

$$B = \frac{\mu_0}{4\pi}\ I \int_{\theta_1}^{\theta_2} \left(\frac{1}{s^2}\cos^2\theta\right)\left(\frac{s}{\cos^2\theta}\right)\cos\theta\,\mathrm{d}\theta$$

The idea here is to integrate between the end points of the wire after changing the variable to θ as shown in figure 15–3.

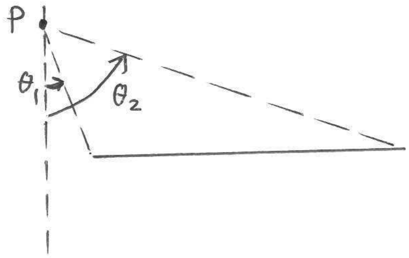


FIGURE 15–3

We don't ask where the current goes at the ends—this is the contribution to the magnetic field *due* to this finite segment of current; the integration results in

$$B = \frac{\mu_0}{4\pi}\ \frac{I}{s}\,(\sin\theta_2 - \sin\theta_1)$$

Let's suppose we want the magnetic field strength at a perpendicular distance from the *centre* of the wire. Then the angle θ belongs to a right angled triangle whose lengths s and $\frac{1}{2}L$ are at right angles and you can convince yourself using a diagram that in this case

$$\sin\theta = \frac{\frac{1}{2}L}{\sqrt{\frac{1}{4}L^2 + s^2}} = \frac{1}{\sqrt{1 + \frac{4s^2}{L^2}}}$$

which leads to

$$\begin{aligned} B &= \frac{\mu_0}{4\pi}\ \frac{I}{s}\,(\sin\theta - \sin(-\theta)) \\ &= \frac{\mu_0}{4\pi}\ \frac{I}{s}\,\frac{2}{\sqrt{1 + 4s^2/L^2}} \end{aligned}$$

If we are interested in an "infinitely long" wire then we take the limit as $L \to \infty$ and so we arrive at this very simple and important result

$$B = \frac{\mu_0 I}{2\pi s} \tag{15.2}$$

Compare this with the *electric* field due to a line of charge which we obtained in Lecture 8 by brute force as we have done here, and very elegantly using Gauss's law in Lecture 10. Especially note the dependence of the field strength on the inverse distance from the wire. Instead of radiating in or out of the wire as do the electric field lines, the magnetic field lines encircle the wire. And of course in the electric case the wire is charged, while in the magnetic case it may be neutral but must be carrying a current.

You may ask if we have a simple way of doing high symmetry problems in *electrostatics* using Gauss's law, is there a similar scheme in *magnetostatics*? The answer is yes, we can use Ampère's law, which we will come to in Lecture 17.

15.3 Magnetic field due to a current carrying loop

You may have wondered whether there is a potential associated with a magnetic field in analogy to the electric potential. Well as I've said the magnetic force is non conservative so it cannot be obtained as the gradient of a scalar potential as can the electric force. Nonetheless there *is* a "magnetic potential" but it's not a scalar field it's a vector field, given the symbol $\mathbf{A}$ and called the *vector potential*. It is related to the magnetic field by the formula

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$$

So it's not obviously easier to work with and we won't use it this year. You just need to know it exists and to accept some of the results in the following lectures when I say that they may only be obtained using the vector potential. The vector potential really comes into its own in quantum mechanics and both $V(\mathbf{r})$ and $\mathbf{A}(\mathbf{r})$ enter the Schrödinger equation for charged particles in electromagnetic fields. Indeed in quantum mechanics V and $\mathbf{A}$ appear to be more *fundamental* than $\mathbf{E}$ and $\mathbf{B}$.

Without using the vector potential it's hard to study the current carrying loop. This is similar to what we found in the case of the electric dipole—we could only obtain a complete solution using the electric potential; when dealing with the electric field in lecture 8 we were only able to solve the problem in two special highly symmetric cases, namely when the field point was perpendicular to or along the axis of the electric dipole. The same situation arises here and using $\mathbf{B}$ we can only find the magnetic field at a field point perpendicularly above the centre of the current carrying loop. You will see in a moment why I make an analogy between an electric dipole and a current carrying loop—the latter turns out to behave as a "magnetic dipole".

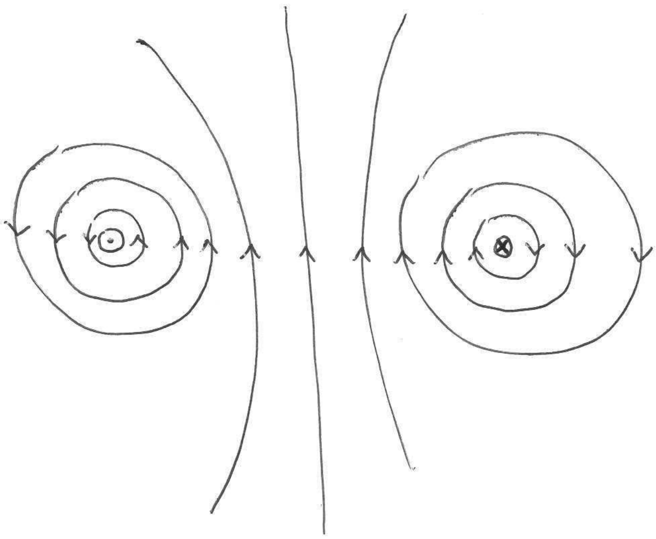


FIGURE 15–4

We can easily guess what the magnetic field lines look like if we sketch the loop end-on so there's a current coming towards us and a current flowing into the page (see the "dot-cross" convention I described in Lecture 14, p. 4).

So right in the centre there is a field line running straight through the loop. Let's find the strength of the magnetic field along this field line.

Please compare figure 15–5 with figure 15–2. The little element $\mathrm{d}\boldsymbol{\ell}_1$ induces a magnetic field $\mathrm{d}\mathbf{B}$ at the field point. The vector $\mathrm{d}\mathbf{B}$ is perpendicular in a right handed sense to both $\mathrm{d}\boldsymbol{\ell}_1$, which is in the velocity direction of the current, and $\mathbf{r}$ the vector pointing from source to field point. R is the radius of the wire loop which carries a current I and we are interested in the *magnitude* of the magnetic field at the field point. We only need the magnitude as we know its direction. Now the field $\mathrm{d}\mathbf{B}$ is, according to the Biot–Savart law,

$$\mathrm{d}\mathbf{B} = \frac{\mu_0}{4\pi}\, I\, \frac{1}{r^2}\, \mathrm{d}\boldsymbol{\ell}_1 \times \hat{\mathbf{r}}$$

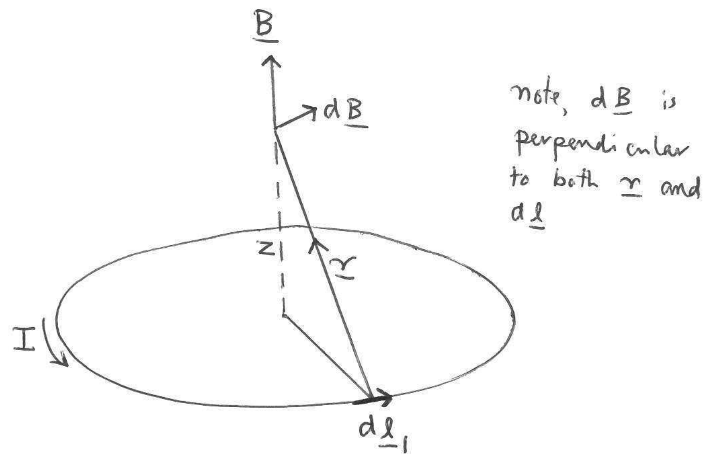


FIGURE 15–5

We only want the z-component of this; other components will cancel by symmetry as you can see from figure 15–5, when all the elements from around the loop are added up. So we only want the magnitude of $\mathrm{d}\boldsymbol{\ell}_1 \times \hat{\mathbf{r}}$ which is simply $\mathrm{d}\ell_1$ because $\mathrm{d}\boldsymbol{\ell}_1$ is perpendicular to $\hat{\mathbf{r}}$ (whose magnitude is one as it's a unit vector). We will multiply this by $\cos\theta$ to get the z-component. This leads us to

$$\begin{aligned} B &= \frac{\mu_0}{4\pi}\, I \oint \frac{1}{r^2} \cos\theta \,\mathrm{d}\ell_1 \\ &= \frac{\mu_0}{4\pi}\, I\, \frac{\cos\theta}{r^2} \oint \mathrm{d}\ell_1 \end{aligned}$$

since in this particular case both θ and r are constant and may move outside the integral sign. The integral of $\mathrm{d}\ell_1$ around the loop is its circumference $2\pi R$, so now

$$\begin{aligned} B &= \frac{\mu_0}{4\pi}\, I\, \frac{\cos\theta}{r^2}\, 2\pi R \\ &= \frac{1}{2}\mu_0 I \frac{R^2}{(R^2+z^2)^{3/2}} \end{aligned} \qquad (15.3)$$

because $\cos\theta = R/r$ and $r^2 = z^2 + R^2$. If we stay a long way away from the loop along the straight field line then $z \gg R$ and the magnetic field is

$$\mathbf{B}(z) = \frac{1}{2}\mu_0\, I\, R^2\, \frac{1}{z^3}\, \hat{\mathbf{k}} \tag{15.4}$$

and I have put back the vector direction since I know what it is. Note the *inverse cube* dependence on the distance z from the loop to the field point. You can compare this with the electric field along the axis of an electric dipole, equation (8.2), in Lecture 8, p. 4. In fact we *define* the magnetic dipole moment of the loop as

$$\begin{aligned}\mathbf{m} &= \text{current} \times \text{area} \times \text{vector perpendicular to the loop}\\ &= \pi R^2 I\, \hat{\mathbf{k}}\end{aligned}$$

so that

$$\mathbf{B} = \frac{1}{2\pi}\mu_0 \frac{1}{z^3}\, \mathbf{m}$$

You may remember that for the electric dipole we obtained this general formula for the electric field, which is independent of the coordinate system,

$$\mathbf{E}_{\text{dip}}(\mathbf{r}) \frac{1}{4\pi\epsilon_0}\, \frac{1}{r^3}\, (3\, (\mathbf{p}\cdot\hat{\mathbf{r}})\, \hat{\mathbf{r}} - \mathbf{p})$$

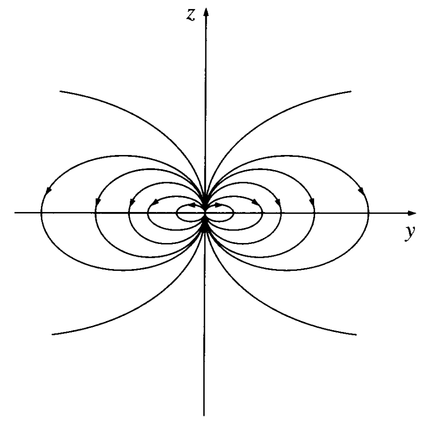
Now, without proof, I can tell you that the general formula for the magnetic field due to a dipole of moment $\mathbf{m}$ is

$$\mathbf{B}_{\text{dip}}(\mathbf{r}) = \frac{\mu_0}{4\pi}\, \frac{1}{r^3}\, (3\, (\mathbf{m}\cdot\hat{\mathbf{r}})\, \hat{\mathbf{r}} - \mathbf{m})$$

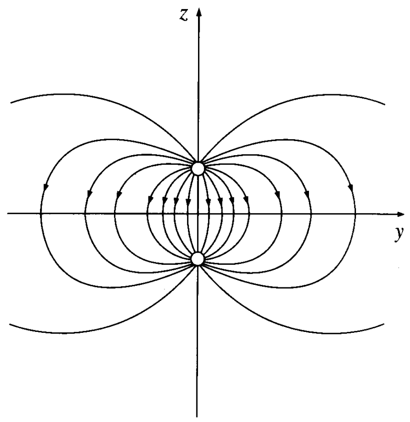
Figure 15–6 illustrates both types of dipole and note that we can talk of “ideal” and “physical dipoles”.

Electric Dipole

Pure Dipole



Physical Dipole

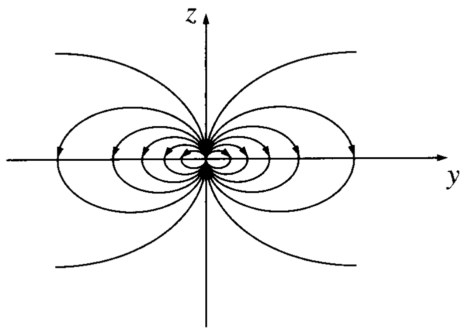


Dipole moment: Dipole electric field: Torque in applied electric field $\mathbf{E}$:

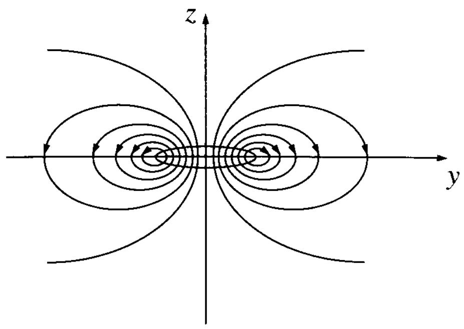
$$\mathbf{p} = qd\,\hat{\mathbf{k}}\,, \qquad \mathbf{E}_{\text{dip}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{1}{r^3}\left(3\,(\mathbf{p}\cdot\hat{\mathbf{r}})\,\hat{\mathbf{r}} - \mathbf{p}\right), \qquad \mathbf{T} = \mathbf{p}\times\mathbf{E}$$

Magnetic Dipole

Pure Dipole



Physical Dipole



Dipole moment: Dipole magnetic field: Torque in applied magnetic field $\mathbf{B}$:

$$\mathbf{m} = IA\,\hat{\mathbf{k}}\,, \qquad \mathbf{B}_{\text{dip}}(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{1}{r^3}\left(3\,(\mathbf{m}\cdot\hat{\mathbf{r}})\,\hat{\mathbf{r}} - \mathbf{m}\right), \qquad \mathbf{T} = \mathbf{m}\times\mathbf{B}$$

FIGURE 15–6

Lecture 16

16.1 Magnetic field at the centre of a solenoid

A solenoid is a tightly wound coil carrying a current I. The radius is a and there are n turns per meter. By "tightly bound" I mean we can neglect the pitch of the turns and treat them as loops perpendicular to the axis. In this approximation there is no current parallel to the axis of the solenoid.

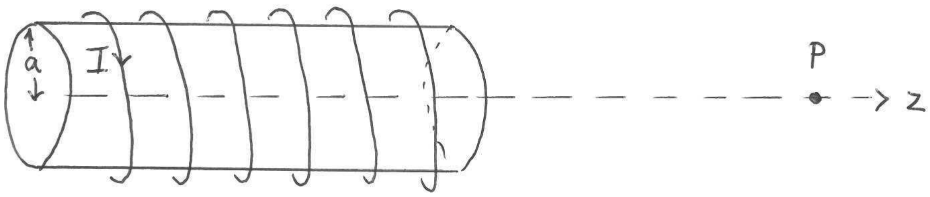


FIGURE 16–1

What is the magnetic field at the point P? We consider an elemental length of the coil $\mathrm{d}l$ and we treat it as a loop of current,

$$\mathrm{d}I = In\,\mathrm{d}l$$

The loop subtends an angle θ at the point P

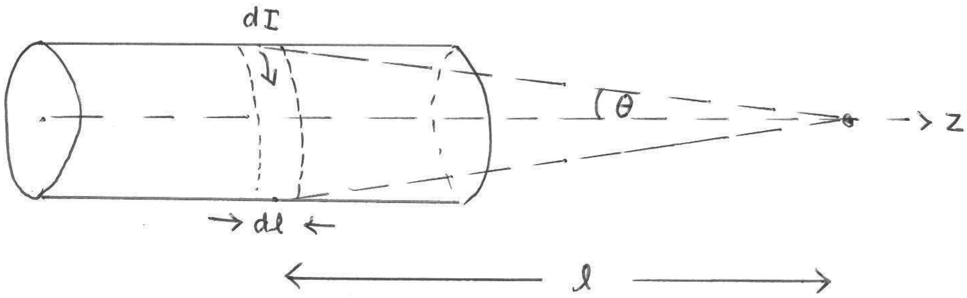


FIGURE 16–2

We now use our formula for the magnitude of the magnetic field due to a loop of current, equation (15.3). The contribution to the z-component of the field, B_z, at P due to the loop of infinitesimal thickness $\mathrm{d}l$ is

$$\mathrm{d}B_z = \frac{1}{2}\mu_0\,\mathrm{d}I\,\frac{a^2}{(a^2+l^2)^{3/2}}$$

with

$$\mathrm{d}I = In\,\mathrm{d}l$$

the right hand side being the current times the number of turns per unit length times the length of the infinitesimal element. Consider this triangle

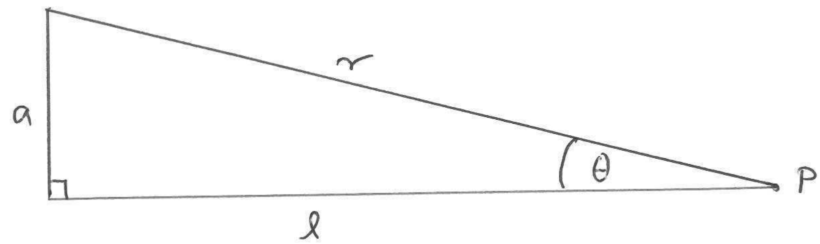


FIGURE 16–3

By trigonometry we see that

$$
\begin{aligned}
a^2 + l^2 &= r^2 \\
r &= \frac{a}{\sin\theta} \quad , \quad r^3 = \frac{a^3}{\sin^3\theta} \\
l &= \frac{a}{\tan\theta} \quad , \quad \mathrm{d}l = -\frac{a}{\sin^2\theta}\,\mathrm{d}\theta
\end{aligned}
$$

and using all of these constructions, we get

$$
\begin{aligned}
\mathrm{d}B_z = &\ \frac{1}{2}\mu_0\, n\, I\, \frac{a^2}{(a^2+l^2)^{3/2}}\,\mathrm{d}l \\
= &\ \frac{1}{2}\mu_0\, n\, I\, \frac{a^2}{r^3}\,\mathrm{d}l \\
= &\ \frac{1}{2}\mu_0\, n\, I\, \frac{\sin^3\theta}{a}\,\mathrm{d}l \\
= &-\frac{1}{2}\mu_0\, n\, I\, \sin\theta\,\mathrm{d}\theta
\end{aligned}
$$

Now we have $\mathrm{d}B_z$ in terms of the subtended angle θ and to get the total magnetic field at P we need to add all the incremental elements by an integration between θ_1 and θ_2 which are the angles subtended at P by the right and left hand ends of the solenoid in figure 16–2. So we have

$$
\begin{aligned}
B_z &= -\frac{1}{2}\mu_0\, n\, I \int_{\theta_2}^{\theta_1} \sin\theta\,\mathrm{d}\theta \\
&= \frac{1}{2}\mu_0\, n\, I\, (\cos\theta_1 - \cos\theta_2)
\end{aligned}
$$

If the solenoid is "infinitely long" then P is obviously inside the coil and you should be able to see that

$$\theta_2 = 0 \quad \text{and} \quad \theta_1 = \pi$$

so that

$$B_z = \mu_0\, n\, I \tag{16.1}$$

This is very important and well known result. The field along the axis of a long solenoid is

$$\mu_0 \times \text{the current} \times \text{the number of turns per unit length}$$

and it points along the axis in a direction such that if you look down the coil and the current is turning anticlockwise then the magnetic field is pointing towards you.

Later we will find the field everywhere near a solenoid, but we'll need help. Just as in electrostatics when the problem got hard we could appeal to Gauss's law, in magnetostatics we'll be able to use Ampère's law. And in place of *gaussian surfaces* we will use *amperian loops*. But first we need to define the *magnetic flux* and obtain Ampère's law.

16.2 Magnetic flux

By analogy with electric flux we may define magnetic flux which is again a measure of the spacing of the field lines. The flux penetrating a surface S is

$$\Phi = \int_S \mathbf{B} \cdot \mathrm{d}\mathbf{a}$$

which you can compare with equation (9.1) in Lecture 9, p. 5. I will not invent another symbol for magnetic flux as we will never deal with the two simultaneously and there will be no confusion as to which of the two I mean—we could of course give them subscripts like $\Phi_{\mathbf{E}}$ and $\Phi_{\mathbf{B}}$ (or indeed Φ_m for gravitational flux) but we won't. We illustrate the magnetic flux like this.

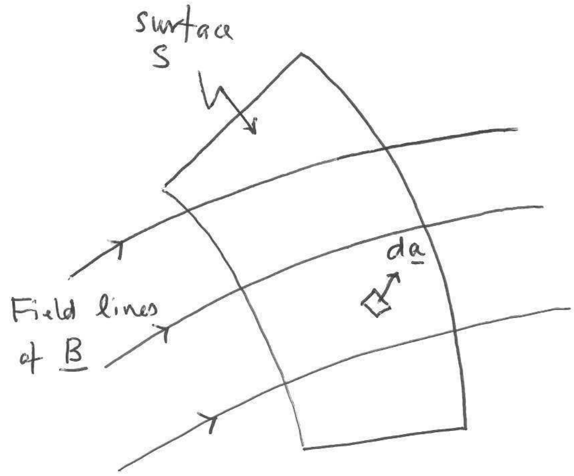


FIGURE 16–4

Magnetic flux has SI units of [T m^{-2}] called *webers*. One weber is one Tesla square meter or one volt-second [V s].

16.3 Divergence of the magnetic field

If our surface is *closed* then we know in the case of *electric flux* from Gauss's law that the flux integrated over the surface is equal to $Q_{\text{enclosed}}/\epsilon_0$, that is, proportional to the enclosed *source* of electric field (namely charge). There are no sources of magnetic field, that is isolated north or south poles. No-one has ever found a magnetic monopole. So until they do, we can assert that

$$\oint_S \mathbf{B} \cdot \mathrm{d}\mathbf{a} = 0$$

which is the same thing as writing (see the additional material at `KEATS`)

$$\nabla \cdot \mathbf{B} = 0 \tag{16.2}$$

In fact this result can be deduced from the Biot–Savart law and hence from the Lorentz transformation of the force (see the text books by Lorrain and Corson). So if a monopole is ever found it will throw a big spanner into a lot of well established physics.

16.3 Curl of the magnetic field

The magnetic field lines circulate around a long straight current carrying wire. The magnetic field has *no* divergence as we just saw in section 16.2, and its *curl* is parallel to the wire.

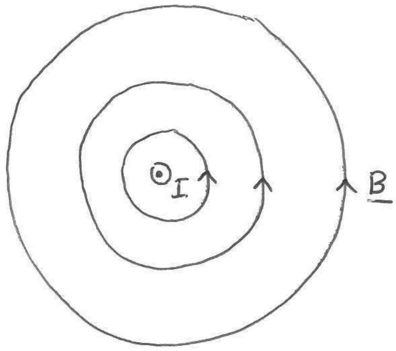


FIGURE 16–5

The magnitude of $\mathbf{B}$ a distance s from the wire is

$$B = \frac{\mu_0 I}{2\pi s}$$

as we saw in Lecture 15, equation (15.2). So the integral around a closed loop of radius s concentric with the wire, is,

$$\begin{aligned}\oint \mathbf{B} \cdot \mathrm{d}\boldsymbol{\ell} &= \oint \frac{\mu_0 I}{2\pi s}\, \mathrm{d}\ell \\ &= \frac{\mu_0 I}{2\pi s} \oint \mathrm{d}\ell \\ &= \frac{\mu_0 I}{2\pi s}\, 2\pi s \\ &= \mu_0 I\end{aligned}$$

where I have exploited the fact that $\mathbf{B}$ and $\mathrm{d}\boldsymbol{\ell}$ are always parallel so the dot product turns into just the product of the magnitudes B and $\mathrm{d}\ell$. Although it's not immediately obvious it is true for *any* loop that encircles the wire. On the other hand for any loop that does *not* enclose the wire, the integral is zero. See figure 16.6.

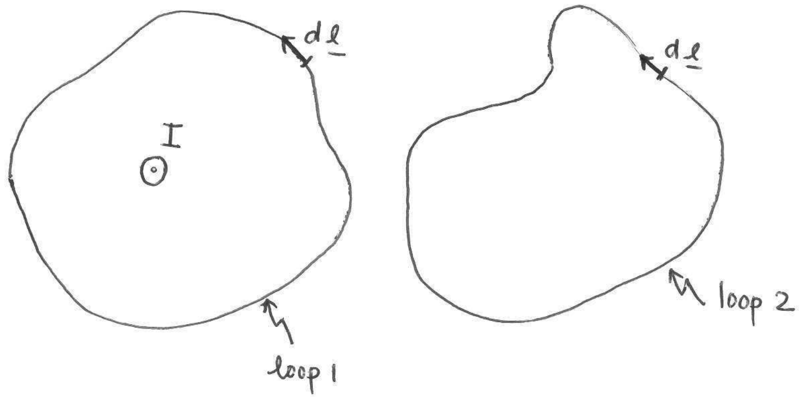


FIGURE 16–6

If I have a bundle of long straight current carrying wires, with currents $I_1, I_2, I_3 \ldots$

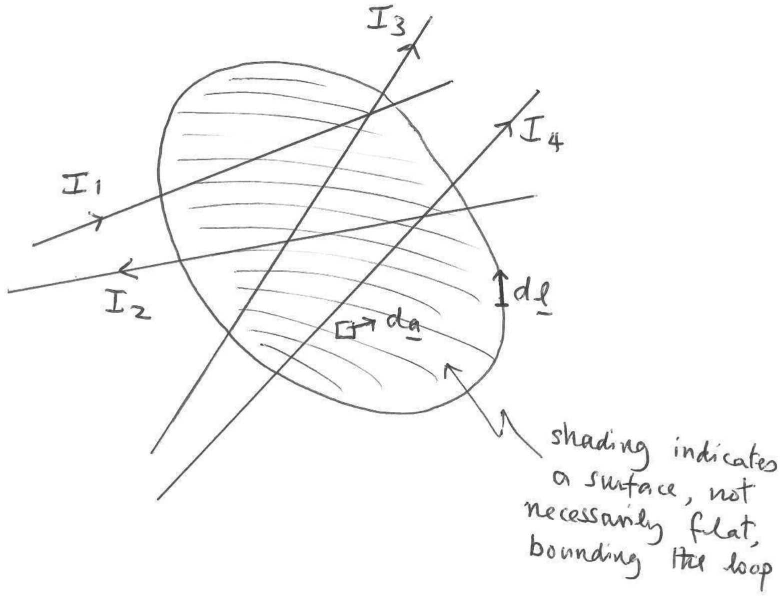


FIGURE 16–7

I will have

$$\oint \mathbf{B} \cdot \mathrm{d}\boldsymbol{\ell} = \mu_0 I_{\text{enclosed}}$$

The magnetic field integrated around any closed loop is equal to μ_0 times the current that penetrates the loop. Some people say the current "linked by" the loop.

Note the similarity to Gauss's law which relates electric field to charge; here we are relating magnetic field to current. If we focus as in figure 16–7 on any surface that is bounded by the loop then we can see that

$$I_{\text{enclosed}} = \int_S \mathbf{J} \cdot \mathrm{d}\mathbf{a}$$

I hope that you can see this; you will have to study figure 16–7 carefully. Now when you have learned Stokes's theorem you will come back to these notes and agree with me that

$$\begin{aligned}\oint \mathbf{B}\cdot \mathrm{d}\boldsymbol{\ell} &= \int_S (\boldsymbol{\nabla}\times\mathbf{B})\cdot \mathrm{d}\mathbf{a}\\ &= \mu_0 \int_S \mathbf{J}\cdot \mathrm{d}\mathbf{a}\end{aligned}$$

and for this to be true over any of the infinite number of surfaces bounded by our loop, the integrands must be equal, leading to

$$\boldsymbol{\nabla}\times\mathbf{B} = \mu_0\,\mathbf{J} \tag{16.3}$$

which is Ampère's law and states that the curl or *circulation* of the magnetic field is equal to μ_0 times the current density.

I have only demonstrated Ampère's law for the case of currents in straight wires. It is in fact generally true as can be shown from the Biot–Savart law. Indeed Ampère's law and the Biot–Savart law are equally fundamental and one can derive each from the Lorentz transformation and each from the other.

Lecture 17

In applications of Ampère's law we always employ the integral form, as we do with Gauss's law, that is

$$\oint \mathbf{B} \cdot \mathrm{d}\boldsymbol{\ell} = \mu_0 \int_S \mathbf{J} \cdot \mathrm{d}\mathbf{a}$$

That is to say, if we can identify a suitable loop through which we know a certain current passes, then if we know from symmetry that the magnetic field has constant magnitude along that loop then we can easily work out the magnetic field once we know the current. We might also be helped by knowing the surface integral on the right hand side may be made over *any* convenient surface as long as it is bounded by our chosen loop.

17.1 The long straight current carrying wire revisited

Once we'd learned Gauss's law, we went back and solved some problems that display a lot of symmetry and found them to be much easier. In Lecture 15, section 15.2 we used the Biot–Savart law to find the magnetic field due to a long straight current carrying wire. We obtained equation (15.2)

$$B = \frac{\mu_0 I}{2\pi s}$$

for the magnitude of the magnetic field a distance s from the wire. Now let's do it using Ampère's law.

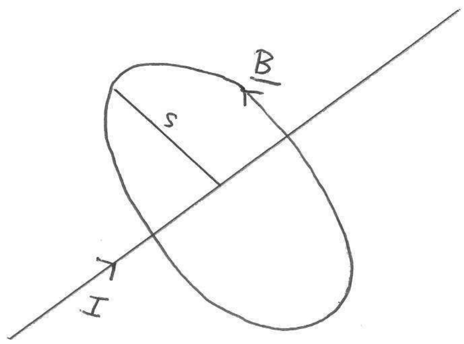


FIGURE 17–1

By the "right hand rule" and by symmetry we know that if we draw an imaginary amperian loop as a circle of radius s concentric with the wire then the magnetic field must be constant around the loop and point along the loop, anticlockwise when looking down the wire with the current flowing towards you. Now the problem becomes a simple one. Ampère's law is

$$\oint \mathbf{B} \cdot \mathrm{d}\boldsymbol{\ell} = \mu_0 \, I_{\text{enclosed}} = \mu_0 \, I$$

and the left hand side simplifies because B is constant and the integral of $\mathrm{d}\boldsymbol{\ell}$ around the loop becomes the circumference of the amperian loop, namely $2\pi s$. Therefore

$$\oint \mathbf{B} \cdot \mathrm{d}\boldsymbol{\ell} = B \oint \mathrm{d}\ell = B \, 2\pi s$$

because $\mathbf{B}$ is parallel to $d\boldsymbol{\ell}$ everwhere around the loop so the dot product becomes the product of the magnitudes of $\mathbf{B}$ and $d\boldsymbol{\ell}$. This leads us to

$$2\pi s B = \mu_0 I$$

which is

$$B = \frac{\mu_0 I}{2\pi s}$$

the same result that we obtained using the law of Biot and Savart.

17.2 The solenoid revisited

To illustrate the application of Ampère's law we return to the solenoid.

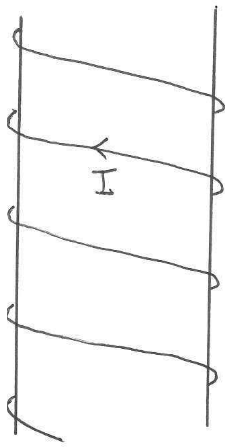


FIGURE 17–2

The solenoid is long enough so that we can neglect end effects. We're going to break down the problem into regions and finally get the field everywhere. First we determine the *direction* of $\mathbf{B}$ and then we find its magnitude. Firstly could there be a radial component, inside or outside the coil?

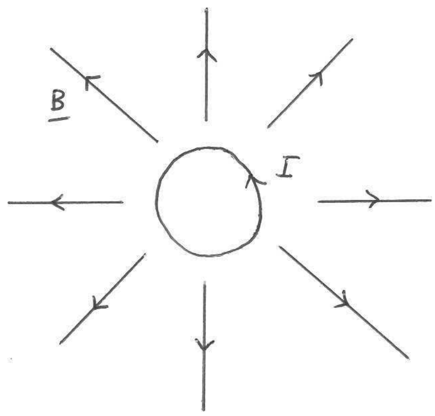


FIGURE 17–3

No, because if I reversed the current I would reverse the direction (the arrows would then point inwards). But I could also reverse the current by turning the solenoid upside down—and that surely could not reverse the field.

Secondly, is there a magnetic field circulating around the outside of the coil? To find out I place an amperian loop outside the solenoid, concentric with its axis (figure 17–4)

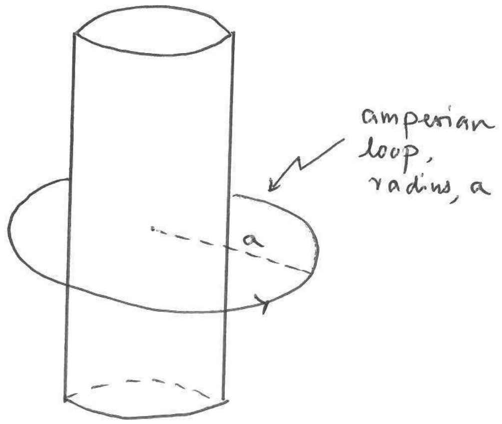


FIGURE 17–4

Remember, as with a gaussian surface, this is an imaginary construction. In diagrams do not confuse the amperian loop with an actual current carrying wire.

The field circulating in the direction of the arrow on the loop is expressed as a line integral

$$\begin{aligned}\oint \mathbf{B} \cdot \mathrm{d}\boldsymbol{\ell} &= \mathbf{B} \cdot \oint \mathrm{d}\boldsymbol{\ell} \\ &= B\, 2\pi a\end{aligned}$$

because B is constant and the line integral just becomes the circumference of the circle, radius a. Now by Ampère's law this is equal to the current poking through the loop. In the idealisation we make earlier of a "tightly wound coil" this current is zero. But in a real solenoid, independently of the number of turns per unit length, the current along the axis is I. Think about it—it has to be or there will be a build up of charge at the ends of the solenoid. So by Ampère's law

$$2\pi a B = \mu_0 I$$

and so

$$B = \frac{\mu_0 I}{2\pi a}$$

so there is a circulating magnetic field, decaying inversely with the distance from the coil, exactly as if the solenoid were a long straight wire carrying a current I. On the other hand the circulating magnetic field *inside* the coil is zero because the amperian loop "links" no current.

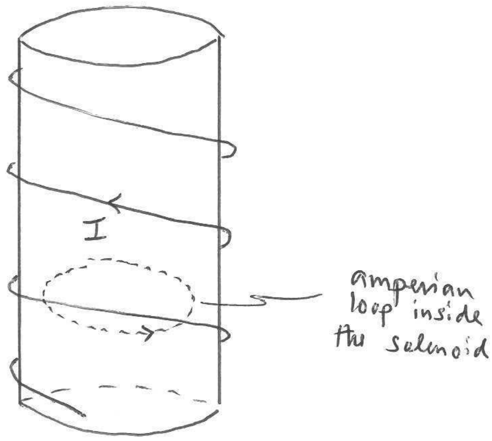


FIGURE 17–5

Thirdly and finally, the only remaining direction is along the axis of the solenoid, our z-direction in figure 17–2. To find this component of the field *outside* consider this amperian loop.

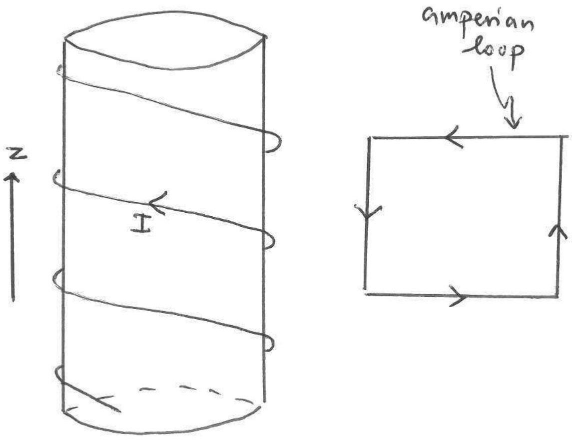


FIGURE 17–6

The loop links no current and we are interested in the z-component of $\mathbf{B}$. The field pointing along the upward leg must cancel that along the downward leg because the integral of $\mathbf{B}$ around the loop is zero (it links no current). This requires that if there *is* a non zero z-component of the magnetic field it is *independent* of the distance from the axis of the solenoid. The field cannot just continue out forever so the z-component must be zero. Now draw the loop with its downward leg *inside* the coil.

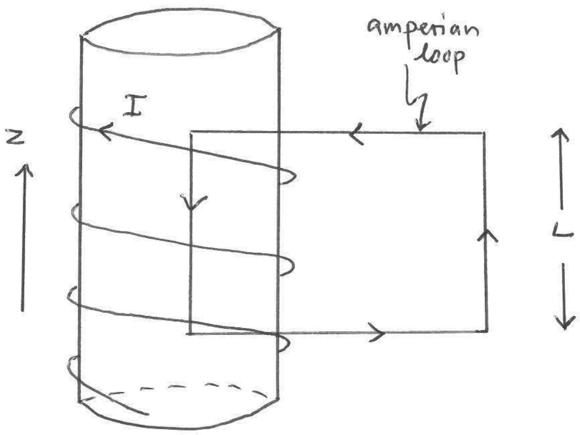


FIGURE 17–7

The magnetic field is zero along three of the legs so the field pointing along the left leg is given by

$$\oint \mathbf{B} \cdot \mathrm{d}\boldsymbol{\ell} = B\,L = \mu_0 I_{\mathrm{enclosed}}$$
$$= \mu_0\, n\, I\, L$$

so, cancelling the L as you'd expect, for the z-component of the field we arrive at

$$B_z = \begin{cases} \mu_0 n I & \text{inside} \\ 0 & \text{outside} \end{cases}$$

There is, in addition, the circulating field outside the coil that we found first which has magnitude $\mu_0 I$; usually this is negligible compared to B_z because $nL \gg 1$.

Note the way I have drawn figure 17–7 (and the other solenoids in this lecture) is such that the current is flowing *clockwise* as I look down from the top. The magnetic field is then pointing in the downward, or negative z-direction, and so carries the unit vector $-\hat{\mathbf{k}}$. This is consistent with figures 16–5 and 16–7 which show that the sense of the magnetic field is *anticlockwise* looking down onto a loop with the "linked" current flowing towards you; this is also the case in figure 17–7 in which the current linking the rectangular amperian loop points towards you and hence the line integral is taken anticlockwise, leading to the magnetic field pointing downwards along the left hand leg of the loop.

Lecture 18

In the next two lectures we come to some electrodynamics. This is the subject of time-varying electric and magnetic fields, and so extends beyond electrostatics and magnetostatics. Here I present only an introduction; next year you will learn how to use the famous "Maxwell equations".

18.1 Ohm's law

Currents flow in wires because something is pushing them. This could be a charged capacitor, an electrochemical battery, a piezoelectric crystal or even a Van der Graaf generator. The current density produced is very often proportional to the *force per unit charge*, $\mathbf{f}$. We write

$$\mathbf{J} = \sigma \mathbf{f}$$

and σ is called the *conductivity* of the substance carrying the current.

$$\rho = \frac{1}{\sigma}$$

is called the *resistivity.* (Do not confuse these with surface and volume charge densites; most people use the same symbols that I am using for all these.) Most often the force is produced by an electric field in which case the force per unit charge is $\mathbf{E}$ (see Lecture 8) and

$$\mathbf{J} = \sigma \mathbf{E}$$

is called *Ohm's law.* It's not really a law because it's essentially never obeyed exactly but it's a very good approximation true for many materials at reasonably small fields. You can also think of Ohm's law as the first term in a Taylor expansion of the current density as a function of electric field, so in that sense it is correct to first order in E.

If I have a conductor of cross sectional area A and length L and I apply an electric potential difference V across its ends, what current will flow?

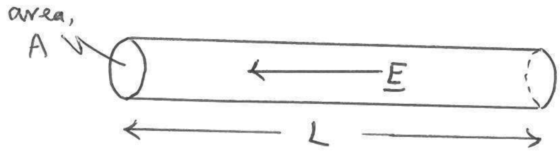


FIGURE 18–1

If the current is uniformly distributed in the conductor then the current is

$$\begin{aligned} I &= JA \\ &= \sigma E A \\ &= \frac{\sigma A}{L} V \end{aligned}$$

if σ is the conductivity. I can write this as

$$V = IR$$

which is the way that Ohm's law is usually written. *In this case only*

$$R = \frac{L}{\sigma A} = \frac{L}{A}\rho$$

is the *resistance* of the wire. In general R depends on the geometry, the distribution of current over the cross section and resistivity. But the resistivity is a *material property.*

18.2 Electromotance

Imagine a wire having some resistance connected to a battery and completing a circuit. Inside the battery, or *source*, S, a force per unit charge $\mathbf{f}_s$ is applied to the electrons.

What keeps them going around the circuit at constant current are electric fields that are generated by the electrons, and these electric fields prevent the electrons from piling up anywhere. You have learned that there can be no electric fields in a metal, but that is only true in electrostatic equilibrium.

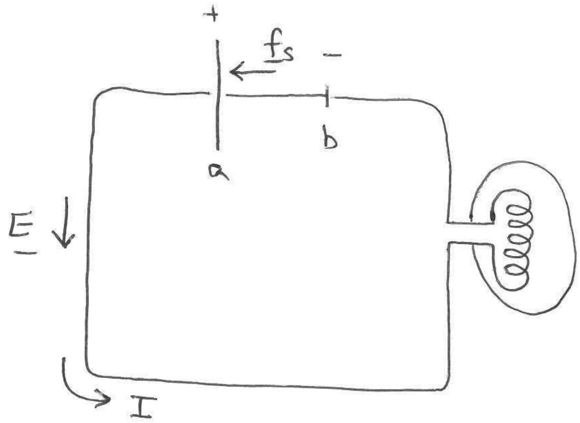


FIGURE 18–2

In the circuit in figure 18–2 the force per unit charge at some point in the circuit is

$$\mathbf{f} = \mathbf{f}_s + \mathbf{E}$$

If there is no "internal resisitance" in the battery then $\mathbf{f}_s$ is exactly balanced by $\mathbf{E}$ *inside* the battery,

$$\mathbf{E} = -\mathbf{f}_s \qquad \longleftarrow \text{inside the battery} \tag{18.1}$$

If we sum the total force around the circuit, including inside the battery, we may define a quantity

$$\begin{aligned}\mathcal{E} &= \oint \mathbf{f}\cdot \mathrm{d}\boldsymbol{\ell} \\ &= \oint \mathbf{f}_s \cdot \mathrm{d}\boldsymbol{\ell}\end{aligned}$$

as the *electromotance.* The second equation follows because $\mathbf{f} = \mathbf{f}_s + \mathbf{E}$ and

$$\oint \mathbf{E}\cdot \mathrm{d}\boldsymbol{\ell} = 0$$

which is true in electrodynamics as long as there are no time varying magnetic fields.

The electromotance, sometimes called the *electromotive force* or e.m.f., here is the work done per unit charge by the source (battery). The potential difference between the terminals a and b is the voltage of the battery,

$$
\begin{aligned}
V &= -\int_a^b \mathbf{E}\cdot \mathrm{d}\boldsymbol{\ell} \ \text{ by the definition of electric potential} \\
&= \int_a^b \mathbf{f}_s \cdot \mathrm{d}\boldsymbol{\ell} \ \text{ by equation (18.1)} \\
&= \oint \mathbf{f}_s \cdot \mathrm{d}\boldsymbol{\ell} \qquad (18.2) \\
&= \mathcal{E} \ \text{ by equation (18.1)}
\end{aligned}
$$

To get to the third line I can extend the line integral to around the whole circuit because $\mathbf{f}_s$ anyway only acts inside the battery. You can see from this why the e.m.f. of a battery is the same as its voltage. In the case that the battery *does* have some internal resistance, call it $\mathcal{R}$, the voltage is

$$V = \mathcal{E} - I\mathcal{R} = IR$$

if R is the resistance of the circuit. So the voltage of a *real life* battery is usually smaller than its e.m.f.

18.3 Motional electromotance

We all know how a generator works. I move a wire in a magnetic field and this induces an e.m.f. that produces a current. Imagine the following simple example.

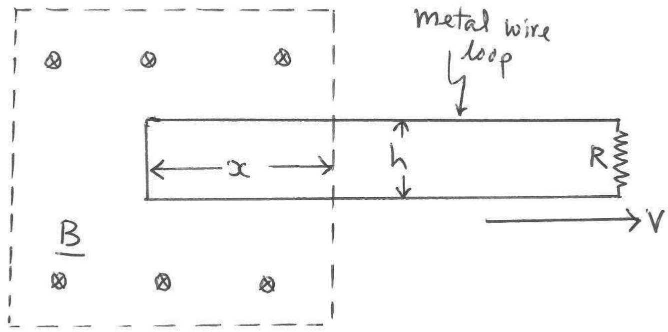


FIGURE 18–3

A uniform magnetic field pointing away from you exists in the the square region outlined by a broken line. I pull the loop with a steady velocity $\mathbf{v}$. Charges in the wire experience a Lorentz force per unit charge

$$\mathbf{F}_{\text{mag}} = \frac{1}{q}\, q\mathbf{v} \times \mathbf{B}$$

and the vertical segment of length h acts like a source of electromotive force whose electromotance according to equation (18.2) is

$$\begin{aligned}\mathcal{E} &= \oint \mathbf{F}_{\mathrm{mag}} \cdot \mathrm{d}\boldsymbol{\ell} \\ &= vBh\end{aligned}$$

since $\mathbf{F}_{\mathrm{mag}}$ is parallel to $\mathrm{d}\boldsymbol{\ell}$ and the integral only has a contribution along the segment of length h.

As I pull out the loop the total magnetic flux that penetrates the area enclosed by the loop is changing because the area of the loop that intersects the region in which the magnetic field exists is getting smaller. The magnetic flux in this case is

$$\Phi = \text{field} \times \text{area} = Bhx$$

and the rate of change of flux is

$$\begin{aligned}\frac{\mathrm{d}\Phi}{\mathrm{d}t} &= Bh\frac{\mathrm{d}x}{\mathrm{d}t} = -Bhv \\ &= -\mathcal{E}\end{aligned}$$

One can show that this is a general result called the *flux rule*,

$$\mathcal{E} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t} \tag{18.3}$$

Whenever, and for whatever reason, the magnetic flux penetrating any shaped loop changes, an e.m.f. is generated around the loop. The minus sign tells us that if the e.m.f. were to cause a current to grow in the loop this current would produce a magnetic field in the *opposite* sense of $\mathbf{B}$. That is to say, induced current will flow in such a way as to oppose the change in flux. This is *Lenz's law.* I said "for whatever reason" and indeed if instead of moving the wire I move the magnetic field the same thing happens—I get the same e.m.f. and the same induced current. Is this obvious to you? Why? In the second instance there are no moving charges and hence no Lorentz force. In fact there are three classes of experiment all conducted by Michael Faraday in the Royal Institution in London in 1831.

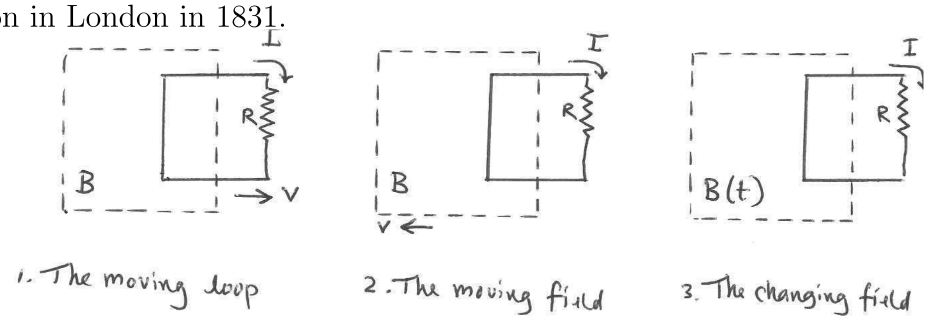


FIGURE 18–4

In each experiment the induced electromotance is given by equation (18.3). It was the coincidence of experiments 1. and 2. that led Albert Einstein first to think about relativity. Faraday's supposition was, "a changing magnetic field induces an electric field."

$$\mathcal{E} = \oint \mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}$$

in which $\mathbf{E}$ is the *induced electric field.*

From the definition of magnetic flux,

$$\mathrm{d}\Phi = \mathbf{B} \cdot \mathrm{d}\mathbf{a}$$

we can see that

$$\oint \mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell} = -\int \frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t} \cdot \mathrm{d}\mathbf{a}$$

This is the partial derivative with respect to time. The magnetic field also depends on the coordinate $\mathbf{r}$. If we apply Stokes's theorem to the left hand side,

$$\int_S (\boldsymbol{\nabla} \times \mathbf{E}) \cdot \mathrm{d}\mathbf{a} = -\int_S \frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t} \cdot \mathrm{d}\mathbf{a}$$

and the surface S is any surface bounded by the wire loop, or notional circuit around which we require the induced electric field, then as we argued in Lecture 16 just before equation (16.3), since this is true for an infinity of such surfaces, the integrands must be equal leading us to *Faraday's law of induction*,

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\mathrm{d}\mathbf{B}}{\mathrm{d}t}$$

which we have already encountered in Lecture 11.

Lecture 19

19.1 Inductance

We know from the Biot–Savart law that the magnetic field due to a current is *proportional* to the current. So if a current I flows around a loop of wire, the magnetic flux penetrating that loop is proportional to I; we write

$$\Phi = LI$$

and the proportionality constant L is called the *inductance* of the loop. It depends on the geometry of the loop and is measured in a unit called a *Henry*,

$$1 \text{ H} = 1 \text{ V s A}^{-1} \text{ (volt second per amp)}$$

If I change the current in the loop I change the flux and this change of flux induces an e.m.f. in the loop. By Lenz's law, which may be paraphrased like this,

"Nature acts to oppose changes in flux"

this e.m.f. acts against the change in current attempting to prevent it from changing. Mathematically, using the flux rule, this is stated as

$$\mathcal{E} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t} = -L\frac{\mathrm{d}I}{\mathrm{d}t} \tag{19.1}$$

and so here $\mathcal{E}$ is sometimes called "back e.m.f." It can cause a spark in a light switch when you turn it off because the back e.m.f. attempts to keep the current flowing even if it has to jump across the contacts in the switch. Every electric circuit has its own inductance. A good inductor is a coil, or *solenoid*: its purpose is to resist changes in current. In this sense it provides *inertia*; compare with the inertial mass in Newton's second law,

$$\text{force} = \text{rate of change of momentum}$$

$$\mathbf{F} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}$$

19.2 Mutual inductance

If you have two loops, loop 1 and loop 2, carrying currents I_1 and I_2

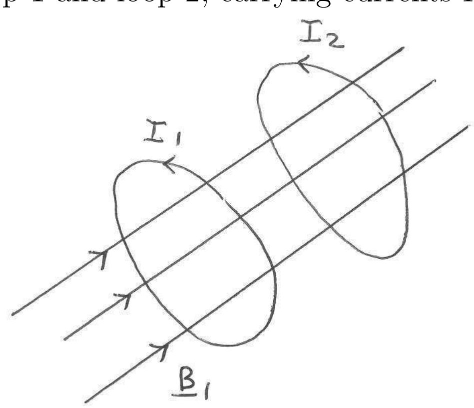


FIGURE 19–1

the magnetic field $\mathbf{B}_1$ due to loop 1 creates a flux Φ_2 that penetrates loop 2,

$$\Phi_2 = \int_{\substack{\text{a surface bounded}\\ \text{by loop 2}}} \mathbf{B}_1 \cdot \mathrm{d}\mathbf{a}_2$$

We don't need to work this out here; just note that Φ_2 is proportional to I_1 by the Biot–Savart law so we can invent a proportionality constant and write

$$\Phi_2 = M_{21}\, I_1$$

and we call M_{21} the *mutual inductance* of the two loops. Conversely of course the current I_2 produces a flux Φ_1 through loop 1 such that

$$\Phi_1 = M_{12}\, I_2$$

and the remarkable fact is that

$$M_{12} = M_{21}$$

To prove this we would need to use the vector potential (see Lorrain, Corson and Lorrain, section 19.1). But the conclusion is that whatever the shapes of the two loops, if I send a current I around loop 1 then the flux through loop 2 is the same as the flux through loop 1 if I drive the same current I through loop 2. This can be very useful in solving certain problems.

19.3 Magnetic energy and momentum

An amount of work has to be done to establish the magnetic field in a circuit. This is in addition to any work that is done in heating the wires by current. The energy I am referring to is fully recoverable when the circuit is switched off. So when I turn on a circuit I have to do work against the back e.m.f. and this amount of work is

$$\frac{W}{Q} = -\mathcal{E}$$

acting on a charge Q. Hence the rate of doing work is

$$\begin{aligned}\frac{\mathrm{d}W}{\mathrm{d}t} &= -\mathcal{E}\,\frac{\mathrm{d}Q}{\mathrm{d}t}\\ &= -\mathcal{E}\, I\\ &= L\, I\,\frac{\mathrm{d}I}{\mathrm{d}t}\end{aligned}$$

where I have used equation (19.1) after noting that I is the quantity of charge passing in unit time. Now if I integrate this from zero current to the final steady current, which I shall call I, I get

$$W = \frac{1}{2} L I^2$$

where L is the inductance of the circuit. Where is this energy stored? One answer to this is arrived at by showing that

$$W = \frac{1}{2\mu_0} \int_{\substack{\text{all}\\\text{space}}} B^2 \, d\tau$$

where B is the magnetic field due to the current I; but you need to use the vector potential to prove it.

Evidently energy is stored in the magnetic field itself. This is quite analogous to what we learned in Lecture 12, namely that the energy stored in the electric field is

$$W = \frac{1}{2}\epsilon_0 \int_{\substack{\text{all}\\\text{space}}} E^2 \, d\tau \tag{12.5}$$

The magnetic field also stores angular momentum as I can demonstrate to you now. Consider this experimental arrangement: a solenoid producing a uniform magnetic field inside and (approximately) zero field outside, encircled by a ring of line-charge, for example a charged plastic hula-hoop somehow suspended so that it is able to rotate freely.

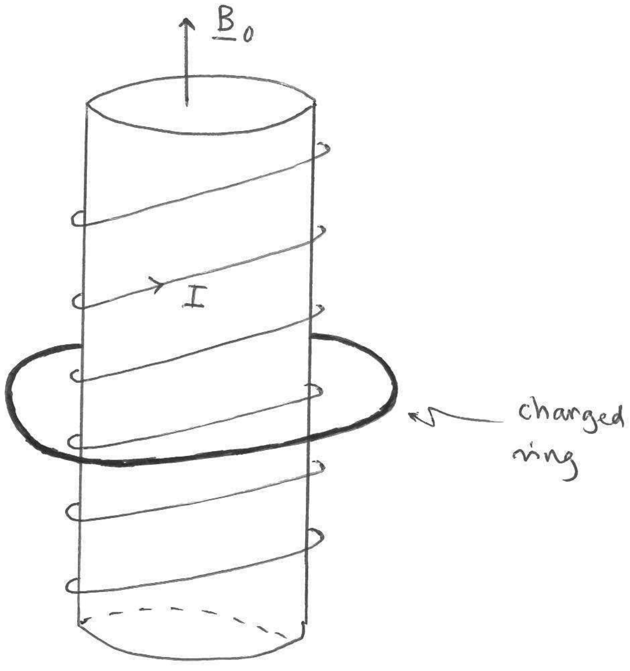


FIGURE 19–2

Now I cut off the current in the coil. What happens? Looking down the axis, the set-up looks like this,

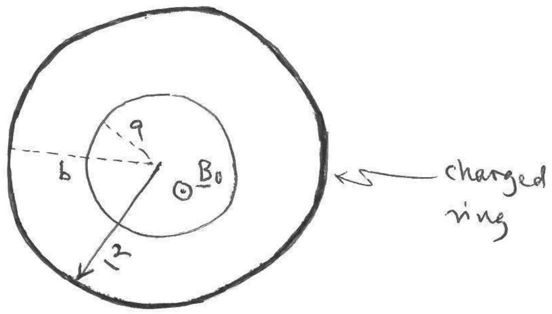


FIGURE 19–3

A magnetic field $\mathbf{B}_0$ exists in the shaded region and points towards you (upwards in figure 19–2 as the current is flowing anticlockwise when look down the coil). When the current is turned off, the changing magnetic field induces an electric field as in Faraday's experiment number 3 in figure 18–4, Lecture 18. This electric field $\mathbf{E}$ according to Faraday's law is given by

$$\oint_{\text{ring}} \mathbf{E} \cdot \mathrm{d}\boldsymbol{\ell} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t} = -\pi a^2 \frac{\mathrm{d}B}{\mathrm{d}t}$$

because flux is magnetic field times area and the field only exists in the region of radius a in figure 19–3. The element of length $\mathrm{d}\boldsymbol{\ell}$ carries a charge $\lambda \mathrm{d}\ell$ where λ is the charge per unit length. The induced electric field exerts a force acting upon $\mathrm{d}\boldsymbol{\ell}$ such that it experiences a torque

$$\mathrm{d}\mathbf{T} = \mathbf{r} \times \mathbf{E}\, \lambda \mathrm{d}\ell$$

whose magnitude is

$$\mathrm{d}T = b\lambda E \mathrm{d}\ell$$

since b is the length of the lever arm $\mathbf{r}$, that is, the radius of the ring. The total torque on the ring is

$$\begin{aligned} T &= \oint_{\text{ring}} \mathrm{d}T \\ &= b\lambda \oint_{\text{ring}} E \mathrm{d}\ell \\ &= -b\lambda\pi a^2 \frac{\mathrm{d}B}{\mathrm{d}t} \end{aligned}$$

Now torque is rate of change of angular momentum,† so

$$T = \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t}$$

where $\mathcal{L}$ is the angular momentum. We have by integration

$$\begin{aligned} \mathcal{L} &= \int T \mathrm{d}t \\ &= -\pi a^2 b\lambda \int_{B_0}^{0} \mathrm{d}B \\ &= \pi a^2 b\lambda B_0 \end{aligned}$$

† This is the *circular motion* equivalent of Newton's second law, namely force equals rate of change of *linear* momentum.

in which B_0 is the magnetic field inside the solenoid before the current is switched off.

So as I turn off the current to the coil the ring starts to rotate anticlockwise if λ is positive. There are two curious features to this experiment.

1. There is no magnetic field outside the solenoid. How does the ring detect the collapsing magnetic field inside the coil? Well, it's the induced electric field that produces the torque on the charged ring. Actually note for next year that whereas there is no magnetic field outside the coil the *vector potential* is *not* zero. You will encounter a simular situation when you study the Aharonov–Bohm effect in quantum mechanics.
2. What about the principle of conservation of angular momentum? Clearly after turning off the current the system possesses angular momentum as the ring is spinning; but before that the ring is not spinning so from where does that angular momentum arise? The conclusion has to be that there is angular momentum *stored* in the magnetic field which is transferred to the ring when the field collapses.

Lecture 20

In this lecture we study a simple electrical circuit. The principal reason is to learn about a non mechanical resonant device and the learn the lesson, "the same equations have the same solutions".

We will discuss an electrical circuit called LCR which is a resistor, an inductor and a capacitor connected in series to an AC power supply. This is a *resonant circuit* so we are to recall what we learned about resonance in Lecture 4. First we study the three components separately.

20.1 The R-circuit

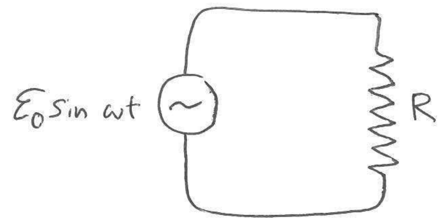


FIGURE 20–1

In the circuit in figure 20–1, $\mathcal{E}_0$ is the peak e.m.f. (voltage); ω is the driving angular frequency; t is time; R is the resistance in ohms. If I is the current, then

$$I = \frac{\mathcal{E}_0}{R} \sin \omega t \tag{20.1}$$

by Ohm's law. The current is *in phase with* the voltage

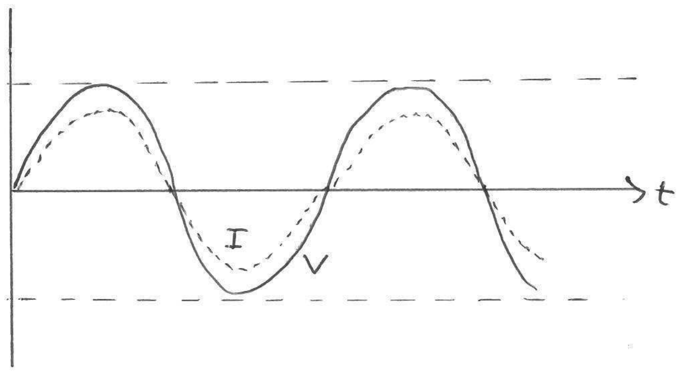


FIGURE 20–2

20.2 The C-circuit

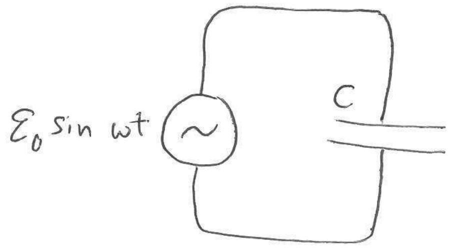


FIGURE 20–3

The capacitance is

$$C = \frac{Q}{V}$$

so the voltage is

$$\frac{Q}{C} = \mathcal{E}_0 \sin \omega t \tag{20.2}$$

Current is

$$I = \frac{\mathrm{d}Q}{\mathrm{d}t} = C\mathcal{E}_0\omega \cos \omega t$$

That is

$$I = \frac{\mathcal{E}_0}{1/\omega C} \sin(\omega t + \frac{1}{2}\pi) = \frac{\mathcal{E}_0}{X_C} \sin(\omega t + \frac{1}{2}\pi)$$

so the current *leads* the voltage by a phase angle of 90° as shown in figure 20–4.

$$X_C = \frac{1}{\omega C}$$

is called the *capacitive reactance*.

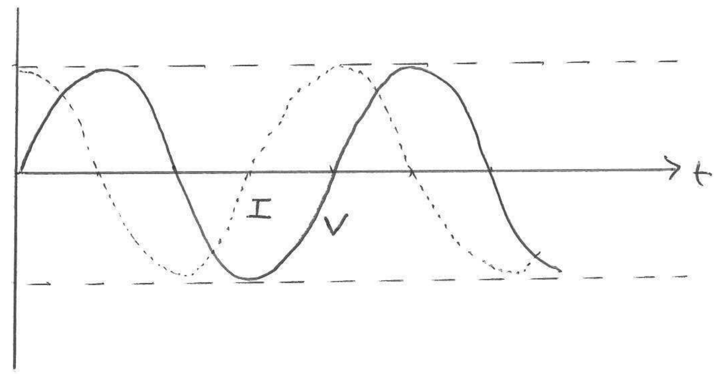


FIGURE 20–4

20.3 The L-circuit

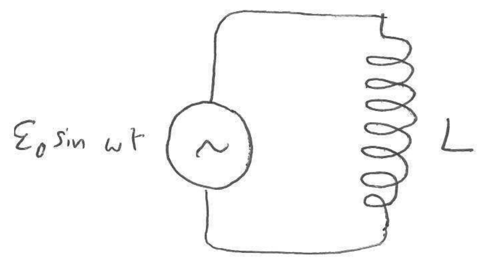


FIGURE 20–5

An inductor is a coil that produces a back e.m.f. as a magnetic field is grown inside the coil. The back e.m.f. is proportional to the rate of increase of current, the constant of proportionality is the *inductance* L. The back e.m.f. is $-L\dot{I}$ by Lenz's law and by Kirchhoff's loop law the back e.m.f. is $-\mathcal{E}$, then,

$$\mathcal{E}_0 \sin \omega t = L\frac{\mathrm{d}I}{\mathrm{d}t} \tag{20.3}$$

that is,

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\mathcal{E}_0}{L}\sin\omega t$$

and by integration,

$$\begin{aligned} I &= -\frac{\mathcal{E}_0}{\omega L}\cos\omega t \\ &= \frac{\mathcal{E}_0}{\omega L}\sin(\omega t - \frac{1}{2}\pi) \\ &= \frac{\mathcal{E}_0}{X_L}\sin(\omega t - \frac{1}{2}\pi) \end{aligned}$$

Now the voltage *leads* the current, or if you prefer the current *trails* the voltage.

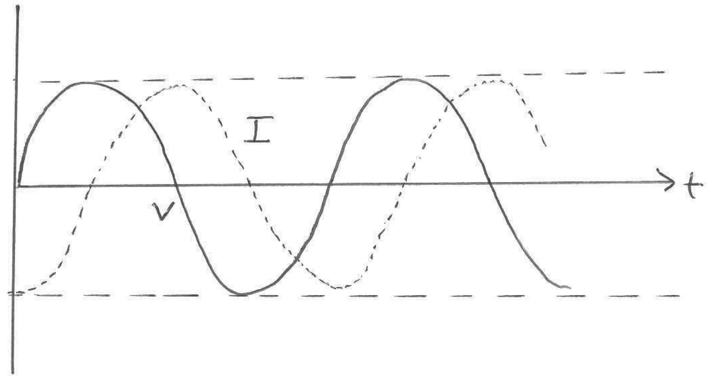


FIGURE 20–6

and

$$X_L = \omega L$$

is called the *inductive reactance.*

20.4 The LCR circuit

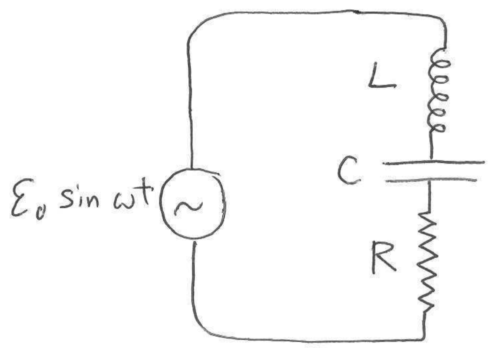


FIGURE 20–7

We now combine equations (20.1), (20.2) and (20.3)

$$L\frac{\mathrm{d}I}{\mathrm{d}t} + RI + \frac{Q}{C} = \mathcal{E}_0\sin\omega t$$

and since

$$I = \frac{\mathrm{d}Q}{\mathrm{d}t}$$

this is

$$L\frac{\mathrm{d}^2Q}{\mathrm{d}t^2} + R\frac{\mathrm{d}Q}{\mathrm{d}t} + \frac{Q}{C} = \mathcal{E}_0 \sin\omega t$$

or

$$\frac{\mathrm{d}^2Q}{\mathrm{d}t^2} + \frac{R}{L}\frac{\mathrm{d}Q}{\mathrm{d}t} + \frac{1}{CL}Q = \frac{\mathcal{E}_0}{L}\sin\omega t$$

Compare this with the equation in Lecture 4, p. 1, which is the equation of motion of a damped mechanical oscillator,

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2Z_m\omega_0\frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2 x = \frac{F_0}{m}\sin\omega t$$

In the case of the LCR circuit we will write

$$\frac{\mathrm{d}^2Q}{\mathrm{d}t^2} + 2\mathcal{Z}\omega_0\frac{\mathrm{d}Q}{\mathrm{d}t} + \omega_0^2 Q = \frac{\mathcal{E}_0}{L}\sin\omega t$$

and we use the symbol Z_m for the damping ratio of the mechanical device and $\mathcal{Z}$ for the damping ratio of the LCR circuit.

The following table shows the correspondence between the parameters of the two devices.

	damped mass and spring	LCR circuit
inertial element	m	L
stiffness	k	$1/C$
damping coefficient	b	R
damping ratio	$\frac{1}{2}b/\sqrt{mk}$	$\frac{1}{2}R\sqrt{C/L}$
static amplitude A_s	F_0/k	$\mathcal{E}_0 C$
quality factor $\mathcal{Q}$	$\sqrt{mk}/b$	$\frac{1}{R}\sqrt{L/C}$
natural frequency ω_0	$\sqrt{k/m}$	$1/\sqrt{LC}$
$A_{\max} = A_s\mathcal{Q}$	$F_0/b\omega_0$	$\mathcal{E}_0/R\omega_0$

We can use this table as a “dictionary” to translate the solution from Lecture 4 into the physics of the present situation. There, we had

$$x = A\sin(\omega t - \phi)$$

with

$$\frac{A}{A_s} = \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2 + \left(2Z_m\frac{\omega}{\omega_0}\right)^2}}$$

and

$$A_s = \frac{F_0}{k}$$

For convenience in what follows, we will shift the phase and write for the time dependence of the charge, Q,

$$Q = -A\cos(\omega t - \phi)$$

This is of course just as good a solution of the differential equation for Q. The shift of phase in our choice of solution of the differential equation does not affect the amplitude, so we still have A/A_s as above, but after substituting $\mathcal{Z}$ for Z_m and using $A_s = \mathcal{E}_0 C$. To obtain the current, we differentiate the charge with respect to time,

$$\begin{aligned} I = \frac{\mathrm{d}Q}{\mathrm{d}t} &= A\omega\sin(\omega t - \phi) \\ &= I_0 \sin(\omega t - \phi) \end{aligned} \tag{20.4}$$

with†

$$I_0 = \frac{\omega \mathcal{E}_0 C}{\sqrt{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2 + \left(2\mathcal{Z}\frac{\omega}{\omega_0}\right)^2}}$$

and after a lot of easy algebra, and using $\mathcal{Z} = \frac{1}{2}R\sqrt{C/L}$ and $\omega_0 = 1/\sqrt{LC}$, this turns into

$$I_0 = \frac{\mathcal{E}_0}{\sqrt{\left(\omega L - \frac{1}{\omega C}\right)^2 + R^2}} \tag{20.5}$$

The phase angle turns out to be‡

$$\begin{aligned} \phi &= \arctan \frac{\omega L - \frac{1}{\omega C}}{R} \\ &= \arctan \frac{\text{reactance}}{\text{resistance}} \end{aligned} \tag{20.6}$$

in which by "reactance" I mean the the inductive reactance take away the capacitive reactance. Now you are going to see why I have used new symbols Z_m and $\mathcal{Z}$ for the damping ratios (as well as $\mathcal{Q}$ for quality factor, so as not to confuse it with Q for charge)—everybody writes equation (20.5) as

$$I_0 = \frac{\mathcal{E}_0}{Z}$$

† I chose the $-\cos$ solution so that we now have $\mathcal{E} = \mathcal{E}_0 \sin\omega t$ and $I = I_0\sin(\omega t - \phi)$ and this instantly identifies ϕ as the phase difference between the current and the voltage, which we what we want to know.

‡ Compare with page 6 of your additional material on `KEATS` on simple harmonic motion after changing $x = A\sin(\omega t - \phi)$ into $x = -A\cos(\omega t - \phi)$. In that case $\tan\phi = (r^2 - 1)/2rZ_m$ with $r = \omega/\omega_0$; then use the dictionary.

That is, the peak current equals the peak voltage divided by, not *resistance* as in a direct current circuit but, *impedance*. Impedance is always given the symbol Z. The impedance of the LCR circuit, according to equation (20.5) is given by

$$Z^2 = \left(\omega L - \frac{1}{\omega C}\right)^2 + R^2$$

and we remember that ωL is the inductive reactance and $1/\omega C$ is the capacitive reactance. So

$$\begin{aligned} Z^2 &= (X_L - X_C)^2 + R^2 \\ &= (\text{reactance})^2 + (\text{resistance})^2 \end{aligned}$$

Depending on the relative sizes of the inductive and capacitive reactances, the total reactance may be positive or negative. We must therefore interpret Z^2 as the "square" of a complex number:†

$$Z^2 = Z^* Z = |Z|^2$$

and

$$Z = R + i\,(X_L - X_C)$$

is called the *complex impedance* of the LCR circuit. We can also write this as

$$Z = |Z|\, e^{i\phi}$$

and plot Z in the Argand diagram

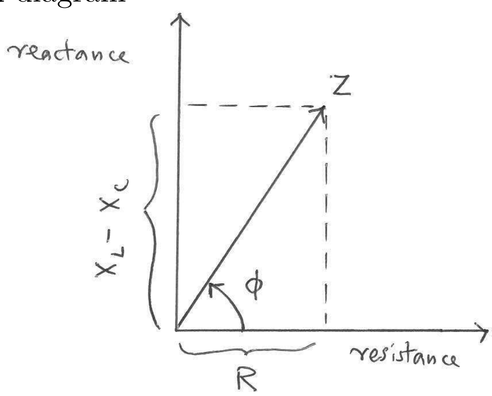


FIGURE 20–8

We then see that the phase is exactly as in equation (20.6), namely,

$$\tan\phi = \frac{\text{reactance}}{\text{resistance}}$$

This is very important because we now see that however difficult the maths has been in all this development, for whatever LCR circuit we construct once we know the resistance,

† We need to do this because we wish to write $Z = a + b$, say, that is the sum of a reactive part and a resistive part; while what we have is $Z^2 = A^2 + B^2$, say. If Z is real, then we have $Z^2 = (a+b)^2 = a^2 + b^2 + 2ab$ which is not in the form $a^2 + b^2$ so this doesn't work. But if Z is *complex* we can write $Z = a + ib$ and then $Z^2 = (a - ib)(a + ib) = a^2 + b^2$ as we require. So we need to insist that either the reactance or the resistance is "imaginary".

capacitance and inductance of the three elements and the driving angular frequency, ω, of the a.c. power supply, then we can easily calculate the reactance and then with a diagram like figure 20–8, we have a simple *graphical construction* to find the impedance and the phase difference—that is, whether and by how much the current leads or trails the voltage in our LCR circuit.

We now write down the oscillating e.m.f. as in figure 20–7,

$$\begin{aligned}\mathcal{E} &= \mathcal{E}_0 \sin \omega t \qquad (20.7)\\ &= \mathrm{Im}\, \mathcal{E}_0\, e^{i\omega t}\end{aligned}$$

in which Im means "imaginary part of", and so the oscillating current in the LCR circuit is

$$\begin{aligned}I = \frac{\mathcal{E}}{Z} &= \mathrm{Im}\, \frac{\mathcal{E}_0}{|Z|} e^{-i\phi}\, e^{i\omega t}\\ &= \frac{\mathcal{E}_0}{|Z|} \sin(\omega t - \phi)\\ &= I_0 \sin(\omega t - \phi) \qquad (20.4)\end{aligned}$$

which is the same as equation (20.4). This is why we wanted a "sine" solution for the current and hence needed a "minus cosine" solution for the charge at the top of page 5. By comparison with equation (20.7), the angle ϕ determines the angle by which the current trails (or leads if $\phi < 0$) the voltage. We can think of this as arising in a diagram like figure 20–9 which provides a graphical means to find the phase relation between voltage and current in an LCR circuit, given the values of L, R and C.

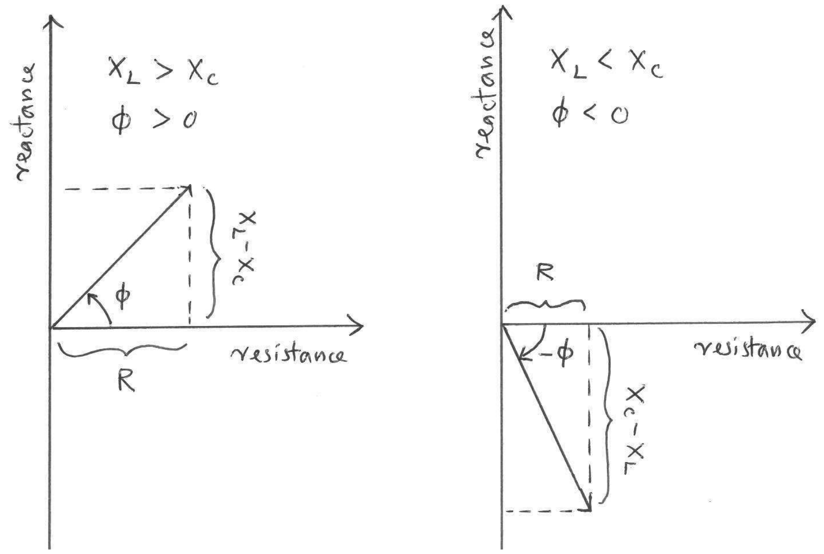


FIGURE 20–9

This illustrates how the phase is zero in a purely resistive circuit as in figure 20–2. The phase is positive or negative depending whether the reactance is positive or negative, or in other words, whether the inductive reactance is greater or less than the capacitive reactance. In a purely *capacitive* circuit we see that $\phi = -\frac{1}{2}\pi = -90°$; and in a purely *inductive* circuit we see that $\phi = +\frac{1}{2}\pi = +90°$ which is entirely consistent with figures 20–4 and 20–6.

The LCR circuit is a *resonant oscillator* just as is its mechanical counterpart. The angular frequency ω_0 is the natural frequency of the *undamped* circuit, that is a circuit with $R = 0$, so ω_0 is the solution of the equation†

$$\omega L - \frac{1}{\omega C} = 0$$

that is,

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

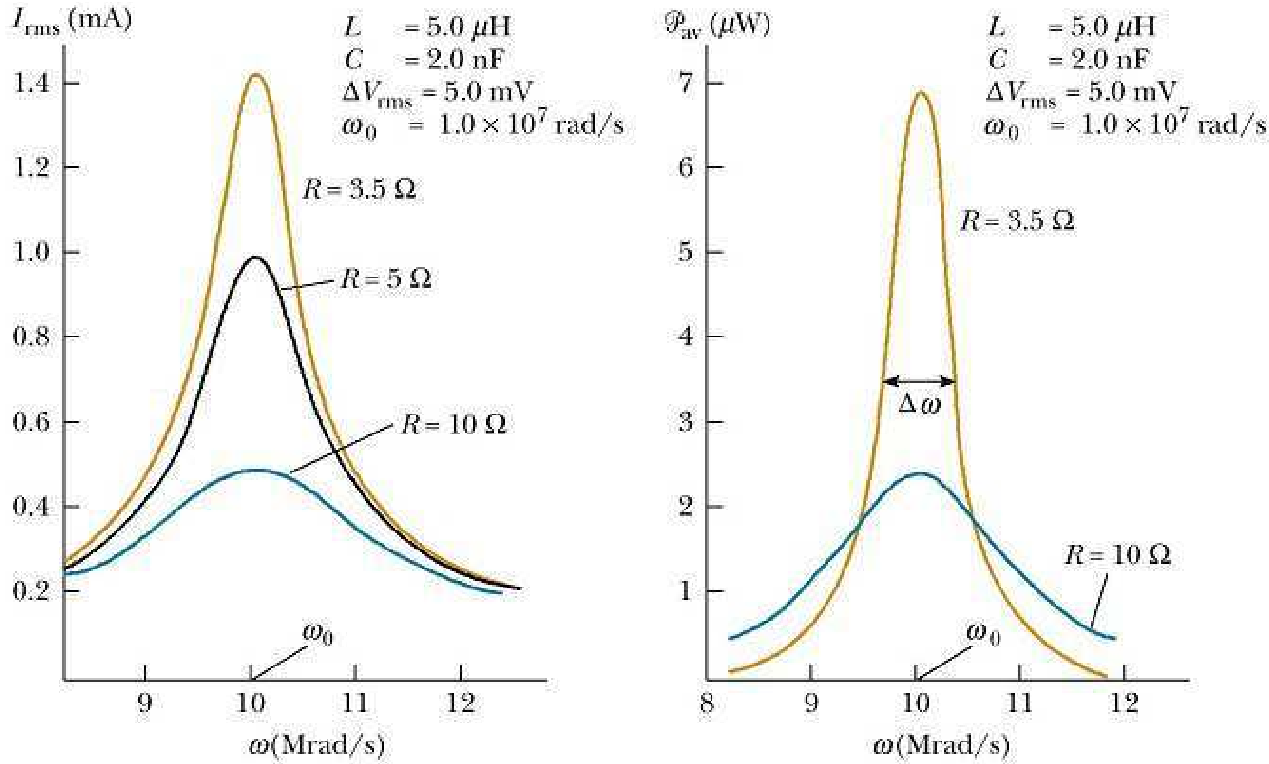
The quality factor is

$$\mathcal{Q} = \frac{\omega_0 L}{R} = \frac{1}{R}\sqrt{\frac{L}{C}}$$

and the static amplitude is $A_s = \mathcal{E}_0 C$ which is the charge stored in the capacitor if the frequency is zero so that the circuit becomes a DC circuit. A circuit with a high $\mathcal{Q}$ has a narrower resonance peak or "bandwidth". This is how a radio is tuned. The capacitance of a resonant circuit is varied until the resonant frequency matches the frequency of the signal being sought. By exploiting a narrow resonance peak, signals at nearby frequencies do not affect the current in the circuit.

Figure 20–10 shows a typical resonance and power output curve for an LCR circuit. You should be able to identify A_s and A_{max} in the left hand graph. In the right hand graph, $\Delta\omega$ is the bandwidth. There's a good `wikipedia` page on the LCR circuit (`en.wikipedia.org/wiki/RLC_circuit`). Note that they call it an RLC circuit. Also in your textbooks, so when you look it up in the index, try under "R" and "C" as well as "L".

† The reason for this is that I am looking for the value of ω that *maximises* the peak current in equation (20.5) in the absence of damping, that is, $R = 0$. So I need to *minimise* the denominator. Actually it is smallest when it is zero. And this means that at resonance the undamped oscillator has infinite amplitude. This is also the case for the mass on the spring as we see in the notes on `KEATS` on SHM. Of course in real life there is no such thing as a totally undamped oscillator—there is always some damping—but at resonance the amplitude can be very large in a high $\mathcal{Q}$ LCR circuit (or any other oscillator with a large quality factor).



I_{rms} (mA)
$L = 5.0\ \mu H$
$C = 2.0$ nF
$\Delta V_{rms} = 5.0$ mV
$\omega_0 = 1.0 \times 10^7$ rad/s
1.4
1.2
1.0
0.8
0.6
0.4
0.2
$R = 3.5\ \Omega$
$R = 5\ \Omega$
$R = 10\ \Omega$
ω_0
9
10
11
12
ω(Mrad/s)
$\mathcal{P}_{av}$ (μW)
$L = 5.0\ \mu H$
$C = 2.0$ nF
$\Delta V_{rms} = 5.0$ mV
$\omega_0 = 1.0 \times 10^7$ rad/s
7
6
5
4
3
2
1
$R = 3.5\ \Omega$
$\Delta\omega$
$R = 10\ \Omega$
ω_0
8
9
10
11
12
ω(Mrad/s)

FIGURE 20–10

4CCP1501 Problem class 1

Do these problems in class with the help of tutors

C1.1 Two waves are travelling along a stretched string in the same direction. They are out of phase by an angle $\frac{1}{2}\pi$ radians. Each wave has an amplitude of 0.04 m. Find the amplitude of the resultant wave.

C1.2 Somewhere out at sea the waves happen to be described as a superposition of these two wavefunctions,

$$y_1 = 3.0\cos(4.0x - 1.6t)$$
$$y_2 = 4.0\sin(5.0x - 2.0t)$$

where x is in meters, and t is in seconds. Find the height of the water above (or below) the mean level at these points and times:

a. $x = 1$, $t = 1$

b. $x = 1$, $t = 0.5$

c. $x = 0.5$, $t = 0$

*C1.3 Two waves are described by the wave functions

$$y_1 = 5\sin(2x - 10t)$$
$$y_2 = 10\cos(2x - 10t)$$

with x in meters and t in seconds. When these are superposed the resulting wave can be described as a single sine function. Find what this function is and calculate the amplitude and phase of the combined wave.

[HINT: use the identity $\sin(a + b) = \sin a \cos b + \cos a \sin b$]

4CCP1501 Problem class 1—solutions

C1.1 Use the formula on page 5 of Lecture 2:

$$\text{amplitude} = 2A\cos\frac{1}{2}\phi$$

This results in an amplitude of 0.0566 m

C1.2 No need to do any algebraic manipulation, just add y_1 to y_2 in the calculator. Negative numbers correspond to the level being below the mean.

a. -1.65 m

b. -6.02 m

c. 1.15 m

*C1.3 The new wave must have the same wavevector and angular frequency of the two combining waves as they are the same for each. So we write

$$5\sin(2x - 10t) + 10\cos(2x - 10t) = A\sin(2x - 10t + \phi)$$

and try and find A and ϕ. First use the identity $\sin(a + b) = \sin a\cos b + \cos a\sin b$ to write

$$A\sin(2x - 10t + \phi) = A\sin(2x - 10t)\cos\phi + A\cos(2x - 10t)\sin\phi$$

This will work as long as we can find a solution to the simultaneous equations

$$\begin{aligned} 5 &= A\cos\phi \\ 10 &= A\sin\phi \end{aligned}$$

Squaring and adding we get $5^2 + 10^2 = A^2$ which means that $A = 11.2$ m. Dividing, we get $\tan\phi = 10/5 = 2$ so that $\phi = 63.4°$. These are the amplitude and phase angle that we are looking for. And so the combined wavefunction is

$$y = 11.2\sin\left(2x - 10t + 63.4 \times \frac{2\pi}{360}\right)$$

4CCP1501 Problem class 2

Do these problems in class with the help of tutors

C2.1 A bunch of bananas is attached to the lower end of a vertical spring in a greengrocer's shop. The spring has a stiffness of 16 N m^{-1}. The bananas are pulled down to a distance 0.2 m from equilibrium, released and thereby set into oscillatory motion. The maximum speed of the bananas is 0.4 m s^{-1}. What is the weight of the bananas in newtons?

Take the acceleration due to gravity as $g = 9.8$ m s^{-1}.

C2.2 A 1 kg glider attached to a spring with force constant 25 N m^{-1} oscillates on a frictionless air track. At $t = 0$ the glider is released from rest at $x = -0.03$ m (that is, the spring is compressed by 0.03 m).

a. Find the period of the glider's motion

b. Find the maximum values of its speed and acceleration

c. Find expressions for the position, velocity and acceleration as functions of time.

*C2.4 Damping is negligible for a 0.15 kg mass hanging from a light spring of spring rate 6.3 N m^{-1}. A sinusoidal force with amplitude 1.7 N drives the mass into forced oscillations. At what driving frequency will the mass oscillate with an amplitude of 0.44 m? (You should actually be able to find *two* possible driving frequencies that give the same amplitude of vibration.)

4CCP1501 Problem class 2—solutions

C2.1 The maximum speed of a simple harmonic oscillator is

$$v_{\max} = A\omega_0 = A\sqrt{\frac{k}{m}}$$

We are told that $A = 0.2$ m, $k = 16$ N m^{-1} and $v_{\max} = 0.4$ m s^{-1}. Thus the mass of the bananas is

$$m = \frac{kA^2}{v_{\max}^2} = 4 \text{ kg}$$

and so their weight is

$$F_g = mg = 39.2 \text{ N}$$

C2.2 We are given that

$$m = 1 \text{ kg},\ k = 25 \text{ N m}^{-1},\ A = 0.03 \text{ m}$$

The boundary condition is

$$\text{At } t = 0,\ x = -0.03 \text{ m}$$

a.

$$\omega_0 = \sqrt{\frac{k}{m}} = 5 \text{ rad s}^{-1}$$

$$\text{period } T = \frac{2\pi}{\omega_0} = \mathbf{1.26} \text{ s}$$

b.

$$v_{\max} = A\omega_0 = \mathbf{0.15} \text{ m s}^{-1}$$

$$a_{\max} = A\omega_0^2 = \mathbf{0.75} \text{ m s}^{-2}$$

c. Because $x = -0.03$ m and $v = 0$ at $t = 0$ the solution we want is $x = -A\cos\omega_0 t$, which for the given values of the constants is

$$x = 0.03\cos(5t + \pi)$$

Therefore the formulas we are asked for are

$$v = \frac{\mathrm{d}x}{\mathrm{d}t} = -\mathbf{0.15}\sin(\mathbf{5t} + \pi)$$

$$a = \frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\mathbf{0.75}\cos(\mathbf{5t} + \pi)$$

with x in meters, v in m s^{-1}and a in m s^{-2}.

*C2.3 We use equation (4.1) from Lecture 4, but neglecting damping, $Z = 0$,

$$A = A_s \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2}} = A_s \frac{1}{\pm\left(1 - \frac{\omega^2}{\omega_0^2}\right)}$$

This leads to

$$\frac{A_s}{A} = \pm\left(1 - \frac{\omega^2}{\omega_0^2}\right)$$

which is rewritten as

$$\frac{\omega^2}{\omega_0^2} = 1 \pm \frac{A_s}{A}$$

where ω is the driving frequency, ω_0 is the natural frequency and A_s is the "static amplitude". We also have

$$\omega_0^2 = \frac{k}{m}$$

and

$$A_s = \frac{F_0}{k}$$

where F_0 is the amplitude of the driving force. Therefore

$$\begin{aligned}
\omega^2 &= \omega_0^2\left(1 \pm \frac{A_s}{A}\right) \\
&= \frac{k}{m}\left(1 \pm \frac{F_0}{kA}\right) \\
&= \frac{k}{m} \pm \frac{F_0}{mA} \\
&= \frac{6.3}{0.15} \pm \frac{1.7}{0.15 \times 0.44}
\end{aligned}$$

giving either $\omega = 4.030$ radians s^{-1} or $\omega = 8.232$ radians s^{-1}. There are thus two frequencies which will produce the required amplitude: $f = $ **0.641 Hz** and $f = $ **1.310 Hz**. Actually you can see why this is by looking at the resonance curve. If you draw a horizontal line at some amplitude, you see that it intersects the curve twice indicating that there are two frequencies that give rise to that particular amplitude.

4CCP1501 Problem class 3

Do these problems in class with the help of tutors

C3.1 Two slits are separated by 0.32 mm. A beam of coherent light of wavelength $\lambda =$ 500 nm impinges on the slits and produces an interference pattern. How many bright fringes are observed in a range of angles between $-30°$ and $+30°$?

C3.2 A Young's double slit experiment is made using a blue-green argon laser. The slits are separated by a distance of half a millimeter and the screen is placed 3.3 m away from the slits. The first bright fringe is located 3.4 mm from the centre of the interference pattern. Find the wavelength of the laser light.

*C3.3 A tanker dumps one cubic metre of oil into the ocean and this spreads evenly into a slick on a perfectly calm sea. Viewed from above using a light beam of variable wavelength, it is found that a first order maximum appears in the brightness of reflected light at a wavelength of 500 nm. You are given that the refractive index of the oil is 1.25 and that this is smaller than the refractive index of seawater. Calculate the area of the oil slick.

4CCP1501 Problem class 3—solutions

C3.1 The bright fringes are observed at angles θ given by the formula

$$d \sin\theta = m\lambda$$

and so the number of fringes in the range $0 < \theta < 30°$ is

$$m = \frac{d\sin\theta}{\lambda} = \frac{3.2 \times 10^{-4}\ \sin 30°}{500 \times 10^{-9}} = 320$$

There is an additional 320 at negative angles and one central bright fringe, so the total is **641** bright fringes.

C3.2 Bright fringes are found at distances

$$y = \frac{\lambda L}{d}\, m$$

from the centerline. We are told the position of the first bright fringe and so we have $m = 1$ and

$$\lambda = \frac{yd}{L} = \frac{3.4 \times 10^{-3} \cdot 0.5 \times 10^{-3}}{3.3} = \mathbf{515\ nm}$$

*C3.3 Because the three media, air, oil, seawater have increasingly larger refractive indices, the light reflected from both surfaces of the oil film will be phase changed by 180°. This gives a path difference of $2t$ if t is the thickness of the film. For a maximum in intensity of the reflected light we require

$$2t = \frac{m\lambda}{n}$$

where $n = 1.25$ is the refractive index of the oil. Since we are told this is a first order maximum, $m = 1$ and

$$t = \frac{\lambda}{2n} = 200\ \text{nm}$$

The volume of the oil is

$$1\ \text{m}^3 = 200\ \text{nm} \times\ \text{area}$$

which leads to an area of **5 square km**.

4CCP1501 Problem class 4

Do these problems in class with the help of tutors

C4.1 I have two cylindrical glass rods of radius 10 mm and length 0.1 m.

a. The density of this glass is 2200 kg m^{-3}. Calculate the mass of each rod.

b. The molar mass of SiO_2 (glass) is $28 + 2 \times 16 = 60$ and each SiO_2 molecule has $4 + 2 \times 6 = 16$ valence electrons. How many valence electrons are there in each rod?

Now I rub the rods with a silk cloth until each has acquired a charge of 100 nC (1 nC $= 10^{-9}$ C).

c. The charge on one electron is $-e = -1.602 \times 10^{-19}$ C. How many electrons have I "rubbed off" each rod? What is this as a fraction of the number of valence electrons you have just found?

d. If I hold these rods 1 m apart, what is the electrostatic force of repulsion between them? What is their gravitational force of attraction? What is the ratio of the two?

C4.2 Using the rules for the drawing of field lines, draw field lines and mark them with arrows for two point charges carrying charges of $+2q$ and $-q$ and separated by a distance d.

C4.3 Eight charges of $q = 1 \times 10^{-3}$ C each are placed at the corners of a cube having a side of $a = 1$ m. A test charge of $q_0 = 0.1 \times 10^{-3}$ C is placed in the centre of one of the faces. Find the magnitude and direction of the force it experiences due to the eight charges. [HINT: exploit the symmetry of the problem.]

4CCP1501 Problem class 4—solutions

C4.1

a. mass $= \pi \times (0.01)^2 \times 0.1 \times 2200 \times 10^3 = 69$ g

b. molar mass = 60 gram per gram-mole

$$\Rightarrow \text{number of moles} = 69/60 = 1.15$$
$$\Rightarrow \text{number of SiO}_2 \text{ molecules} = 1.15 \times 6.02 \times 10^{23} = 6.92 \times 10^{23}$$
$$\Rightarrow \text{number of valence electrons} = 16 \times 6.92 \times 10^{23} = 1.1 \times 10^{25}$$

c.

$$\frac{100 \text{ nC}}{e} = \frac{10^{-7}}{1.602 \times 10^{-19}} = 6.24 \times 10^{11} \text{ electrons}$$

Fraction is $6.24 \times 10^{11}/1.1 \times 10^{25} = 6 \times 10^{-14}$

d. Coulomb force:

$$F_e = 9 \times 10^9 \times (10^{-7})^2 = 9 \times 10^{-5} \text{ N}$$

Gravitational force:

$$F_g = 6.7 \times 10^{-11} \times (0.069)^2 = 3.2 \times 10^{-13} \text{ N}$$

Ratio is 0.28×10^9

C4.2 Consult the following figure taken from a textbook.

EXAMPLE 22–6 Draw the electric field lines for a system of two charges, $+2q$ and $-q$, separated by a fixed distance.

Setting It Up We sketch a 2-dimensional representation of the field due to the two charges $+2q$ and $-q$ in Fig. 22–13a. Twice as many field lines leave the charge $+2q$ as end at the charge $-q$.

Strategy Arbitrarily close to each charge the field lines will be radial and uniformly spread over the area surrounding the charge. With twice as many lines coming from $+2q$ as go into charge $-q$, we can take half of the lines from the charge $+2q$ and connect them to the lines going into charge $-q$. The remaining lines go off to infinity.

Working It Out We choose, arbitrarily, to show 24 lines coming from $+2q$ so that 12 lines will go into $-q$ and 12 will go off to infinity. The final sketches are presented in Figs. 22–13b and 22–13c.

What Do You Think? What happens to the remaining 12 lines that emerge from $+2q$ that don't terminate on $-q$?

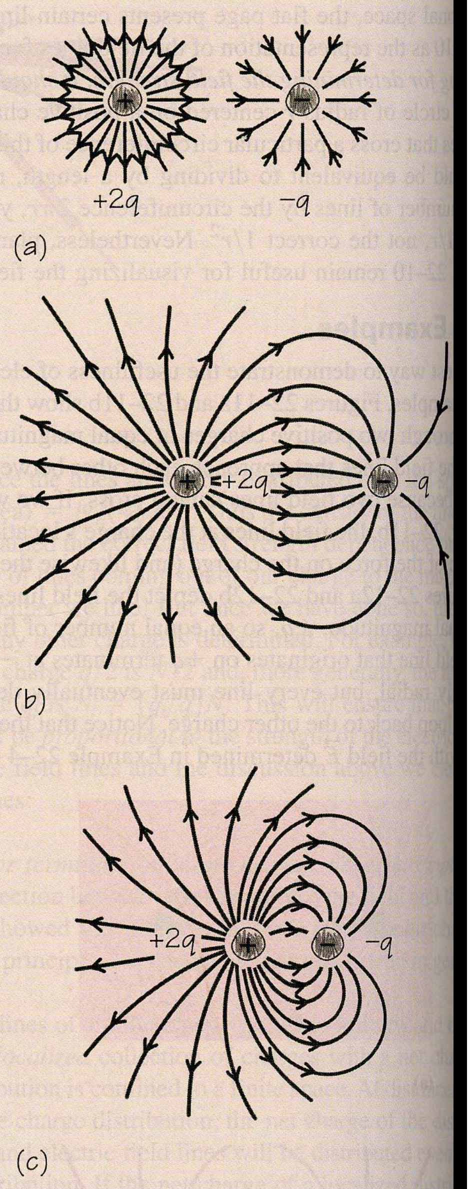
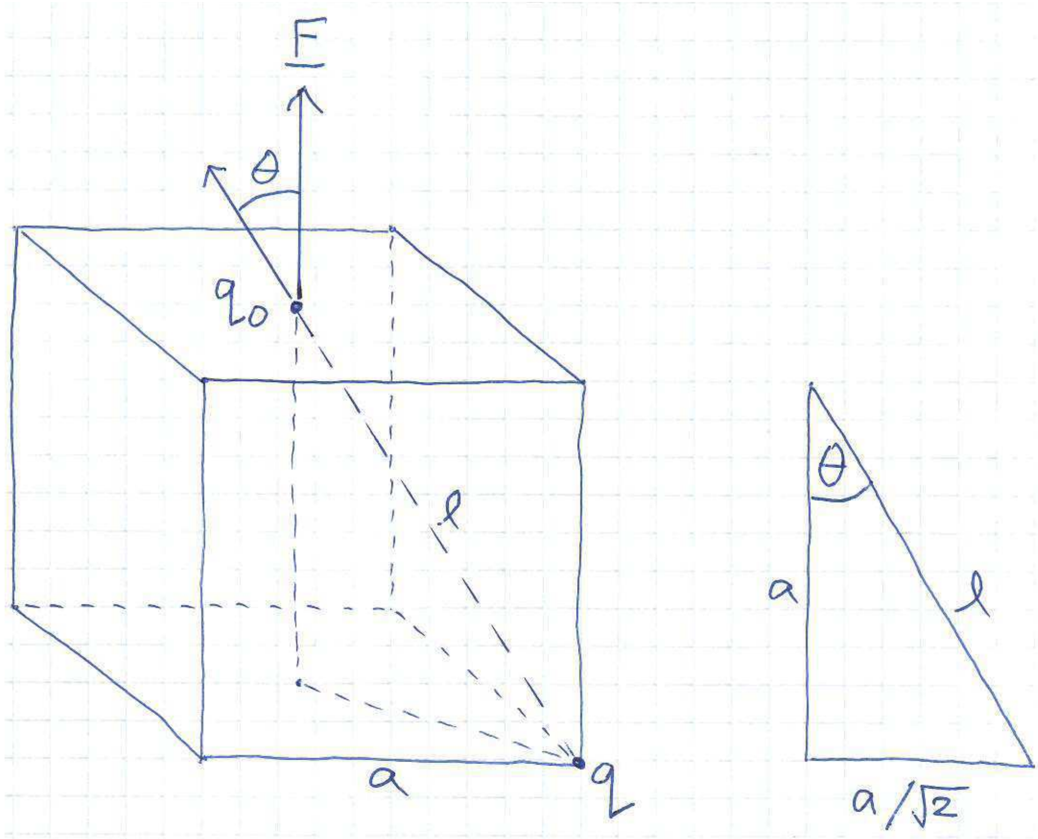


FIGURE 22–13 (a) The electric field lines close to the $+2q$ and $-q$ point charges are those of a point charge. (b) Half the electric field lines that emerge from $+2q$ end up on $-q$. (c) Far from the point charges, the electric field lines are those of a point charge $+q$. In this view you can begin to see the lines spread out to form a radial distribution.

C4.3 Refer to the figure below.



The forces due to the charges in the upper face cancel each other and the forces due to charges in the lower face are equal in magnitude and sum to give a resultant force pointing perpendicular to the upper face. Its magnitude is four times the force from one of the charges, say the front right, multiplied by $\cos\theta$.

$$F = 4\left(\underbrace{9\times 10^9}_{1/4\pi\epsilon_0}\times \frac{qq_0}{\ell^2}\cos\theta\right)$$

By examining the right hand figure we see that $\tan\theta = \frac{1}{\sqrt{2}}$ and $\ell^2 = \frac{3}{2}a^2$ and so

$$\begin{aligned} F &= 4\times 9\times 10^9\times \frac{qq_0}{\frac{3}{2}a^2}\cos\theta \\ &= 4\times 9\times 10^9\times 0.1\times 10^{-3}\times 10^{-3}\times \frac{2}{3}\cos 0.6155 \\ &= 1.9603\times 10^3 \qquad \text{[N]} \end{aligned}$$

It is useful to get an idea of the magnitudes of forces and charges. Here, we see that charges on the order of μC acting over distances of the order of metres give rise to forces of the order of kilonewtons (kN). These are large forces; remember one Newton is roughly the force due to gravity exerted on your hand when you hold an apple in it. If you want to buy a tensile testing machine for ripping apart specimens of steel with diameters of about 1 cm you can get one that can exert a pulling force of nearly a thousand kN. From that perspective question T6.1 in your tutorials, based on one first posed by Richard Feynman, is particularly astonishing.

4CCP1501 Problem class 5

Do these problems in class with the help of tutors

C5.1 The electric field inside a large parallel plate capacitor is 10^7 N C^{-1}. Calculate the total flux penetrating the following imaginary surfaces placed within the capacitor.

a. A square of side 1 mm aligned parallel to the plates

b. The same square having its normal tilted at 30° with respect to the normal to the capacitor plates

c. A hemispherical surface of radius 1 mm with its circumference lying parallel to the plates

d. The same hemispherical surface tilted at 90° so that its circumference lies in a plane perpendicular to the plates.

C5.2 In your lecture notes you have calculations of the electric field due to an *infinite line* of uniform charge density λ C m^{-1}, and an *infinite sheet* of uniform charge density σ C m^{-2}. Use Gauss's Law and suitable Gaussian surfaces to solve these problems again and confirm that you get the same results as in your notes.

4CCP1501 Problem class 5—solutions

C5.1

a. $E = 10^7$, $A = 10^{-6}$, $\Phi = EA = 10\ \mathrm{NC^{-1}m^2}$

b. $E = 10^7$, $A = 10^{-6}\cos 30° = 10^{-6}\sqrt{3}/2$, $\Phi = EA = 8.66\ \mathrm{NC^{-1}m^2}$

c. $E = 10^7$, $A = \pi \times 10^{-6}$, $\Phi = EA = 31.4\ \mathrm{NC^{-1}m^2}$

d. zero

C5.2 Please consult your lecture notes where you will find all these problems solved.

4CCP1501 Problem class 6

Do these problems in class with the help of tutors

C6.1 A solid spherical conductor has radius a and carries a charge Q. It is enclosed by a very thin concentric spherical shell of conductor of radius b carrying a charge $-Q$.

a. Use Gauss's Law to calculate the electric field everywhere in this example and plot the field as a function of distance from the centre of this object.

b. Calculate the electric potential difference between a and b.

c. Hence calculate the *capacitance* of this object.

[There can be no charge inside a conductor: all its charge must be found at its surface.]

*C6.2 We have learned that the curl of the electrostatic field is zero and hence the electric field is special in that there are three equations relating the three components of the vector field. For example consider the following two possible forms for an electrostatic field.

a. $\mathbf{E} = C\left(y^2\hat{\mathbf{i}} + (2xy + z^2)\,\hat{\mathbf{j}} + 2yz\hat{\mathbf{k}}\right)$

b. $\mathbf{E} = Cy\hat{\mathbf{i}}$ which is a field pointing everywhere in the x-direction whose magnitude is proportional to how far you are along the y-axis. Sketch this field with arrows.

Which, if any, of these is a legitimate electric field? Can you see from your sketch why b. is not?

4CCP1501 Problem class 6—solutions

C6.1 I'll leave you to make the sketch plot in part a.

a.

$$\begin{aligned} 0 < r < a\ , &\quad E = 0 \\ a < r < b\ , &\quad E = \frac{1}{4\pi\epsilon_0}\,\frac{Q}{r^2} \quad [\text{N C}^{-1}] \\ r > b\ , &\quad E = 0 \end{aligned}$$

b.

$$\begin{aligned} V(a) - V(b) &= -\int_b^a E\,\mathrm{d}r \\ &= -\int_b^a \frac{1}{4\pi\epsilon_0}\,\frac{Q}{r^2} = \left[\frac{1}{4\pi\epsilon_0}\,\frac{Q}{r}\right]_b^a \\ &= \frac{1}{4\pi\epsilon_0}\,Q\left(\frac{1}{a} - \frac{1}{b}\right) \\ &= \frac{1}{4\pi\epsilon_0}\,Q\,\frac{b-a}{ab} \quad \text{volt [N m C}^{-1}] \end{aligned}$$

c.

$$\begin{aligned} C &= \frac{Q}{\Delta V} \\ &= 4\pi\epsilon_0\frac{ab}{b-a} \quad \text{farad [C V}^{-1}] \end{aligned}$$

*C6.2 The answer is that a. is a legitimate electric field, while b. is not. You can probably see when you sketch the vector field of b. that it looks like it has a non zero "circulation".

a. We resolve the vector $\mathbf{E}$ into its components,

$$\mathbf{E} = E_x\hat{\mathbf{i}} + E_y\hat{\mathbf{j}} + E_z\hat{\mathbf{k}}$$

and so, by inspection,

$$\begin{aligned} \frac{1}{C}\,E_x &= y^2 \\ \frac{1}{C}\,E_y &= 2xy + z^2 \\ \frac{1}{C}\,E_z &= 2yz \end{aligned}$$

Now we check the derivatives,

$$\begin{aligned} \frac{1}{C}\left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right) &= 2z - 2z = 0 \\ \frac{1}{C}\left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\right) &= 0 - 0 = 0 \\ \frac{1}{C}\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right) &= 2y - 2y = 0 \end{aligned}$$

so indeed for this vector field $\boldsymbol{\nabla} \times \mathbf{E} = \mathbf{0}$.

b. Now we do the same and find

$$\begin{aligned}
\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} &= 0 \\
\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} &= 0 \\
\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} &= -1
\end{aligned}$$

so not all components of the curl are zero—the z-component is -1. So for this field $\boldsymbol{\nabla} \times \mathbf{E} \neq \mathbf{0}$.

4CCP1501 Problem class 7

Do these problems in class with the help of tutors

C7.1 A uniform magnetic field, $\mathbf{B}$, points in the z-direction. A particle of mass m and charge q is travelling at a constant velocity $\mathbf{v}$ in the x-direction when it enters the magnetic field.

a. Explain why the particle moves in a circle due to the magnetic field.

b. Looking so that the magnetic field is pointing away from you, if q is positive is the motion of the particle clockwise or anticlockwise?

c. Use the equation for the Lorentz force to show that the magnitude of the magnetic force is qvB.

d. Since the particle is travelling in a circle it is accelerating. Equate the force with the mass times the centripetal acceleration v^2/r, where r is the radius of the particle's orbit, and hence show that the radius is $r = mv/qB$.

e. Find the angular frequency ω of the particle's orbit. This is called the "cyclotron frequency." Find the period, T, of the orbit.

f. Suppose that the particle entered the magnetic field in a direction slightly inclined to the x-direction. Describe qualitatively the shape of the path it will take.

g. Suppose that in addition to the uniform magnetic field $\mathbf{B}$ pointing in the z-direction, there is also a uniform electric field $\mathbf{E}$ pointing in the y-direction. If the particle is released from rest, describe qualitatively the path that it takes. Sketch the path.

C7.2 (After Griffiths) In 1897 J. J. Thomson "discovered" the electron by measuring the charge-to-mass ratio of "cathode rays" (actually streams of electrons with charge q and mass m) as follows.

a. First he passed the beam through uniform crossed electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ (mutually perpendicular and both of them perpendicular to the beam), and adjusted the electric field until he got zero deflection. What then was the speed of the particles in terms of E and B?

b. Then he turned off the electric field, and measured the radius of curvature, r, of the beam, as deflected by the magnetic field alone. In terms of E and B, what is the charge-to-mass ratio, q/m, of the electrons?

4CCP1501 Problem class 7—solutions

C7.1

a. The magnetic force is $\mathbf{F}_{\text{mag}} = q\mathbf{v} \times \mathbf{B}$ and so is at all times perpendicular to both the velocity and the field. This implies that the magnetic force is equivalent to a centripetal force which drives the particle into a circular orbit. In a circular orbit the velocity is always a tangent vector to the circle and the centripetal force is directed along the radius vector and is hence perpendicular to the velocity. Both these vectors are perpendicular to the magnetic field if the field vector is perpendicular to the plane of the circular orbit. This is how a *cyclotron* works.

b. Anticlockwise.

c. $\mathbf{F}_{\text{mag}} = q\mathbf{v} \times \mathbf{B}$ and so the magnitude is $F_{\text{mag}} = qvB$.

d. $F_{\text{mag}} = qvB = mv^2/r$, hence $r = mv/qB$.

e. $\omega = v/r = qB/m$. $T = 2\pi/\omega = 2\pi m/qB$.

f. The particle will describe a helix whose axis is the z-direction.

g. When the particle is released it will feel no magnetic force as its velocity is zero, and it will be accelerated along the y-direction by the electric field. But it immediately acquires a velocity in the y-direction so that the magnetic field begins to deflect it into the x direction. As the speed increases, so does the magnetic force which increases while the electric force remains constant. As the particle describes a curved trajectory it reverses the direction of travel along the y-direction so that it is now travelling against the electric field and so it begins to slow down. The magnetic force then weakens so that the particle slows down and comes to rest back on the x-axis but displaced along it. It has thereby completed a semi-eliptic orbit and returned stationary at the x-axis, displaced a certain distance along x; after which it repeats the above process. This is called "cycloid motion." (See Griffiths, example 5.2) It is actually the path traced out by a point on the rim of a bicycle wheel.

C7.2

a. $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = 0 \Rightarrow E = vB \Rightarrow v = E/B$

b. Use the result of C7.1d, $r = mv/qB$ and $v = E/B$ from C7.2a. Hence $q/m = v/rB = E/B^2r$

4CCP1501 Problem class 8

Do these problems in class with the help of tutors

C8.1 The dipole fields due to electric and magnetic dipoles are

$$\mathbf{E}_{\rm dip}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{1}{r^3}\left(3\left(\mathbf{p}\cdot\hat{\mathbf{r}}\right)\hat{\mathbf{r}} - \mathbf{p}\right)$$

and

$$\mathbf{B}_{\rm dip}(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{1}{r^3}\left(3\left(\mathbf{m}\cdot\hat{\mathbf{r}}\right)\hat{\mathbf{r}} - \mathbf{m}\right).$$

These formulas are independent of the choice of axes. But taking the z direction to be along the axis of the dipole use these to obtain the dipole field in the special cases that the field point $\mathbf{r}$ is (i) along the axis, and (ii) in the plane of the dipole. Compare these answers with the results obtained in your lecture notes.

*C8.2 (After Griffiths) A square loop of wire carries a steady current I. The sides of the square have length $2R$. Calculate the magnetic field $\mathbf{B}$ at the mid point in the plane of the loop. [HINT: look in Lecture 15, p. 4.]

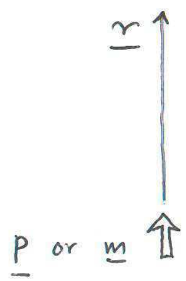
a. Now find $\mathbf{B}$ at the centre of a regular n-sided polygon, for which R is the perpendicular distance from the centre to one of the sides, and check that your answer is consistent with what you got in a. for the case $n = 4$.

b. Using your formula obtained in b., take the limit as $n \to \infty$. You will need to use the rule of L'Hopital. Confirm that the magnetic field you obtain is that of a circular loop in your lecture notes.

4CCP1501 Problem class 8—solutions

C8.1

(i) $\mathbf{p} = p\hat{\mathbf{k}}\,,\ \hat{\mathbf{r}} = \hat{\mathbf{k}}\,,\ \mathbf{p}\cdot\hat{\mathbf{r}} = p$



$$\mathbf{E}_{\text{dip}} = \frac{1}{4\pi\epsilon_0}\,\frac{1}{r^3}(3p\hat{\mathbf{k}} - p\hat{\mathbf{k}}) = \frac{1}{4\pi\epsilon_0}\,\frac{1}{r^3}2p\hat{\mathbf{k}}$$

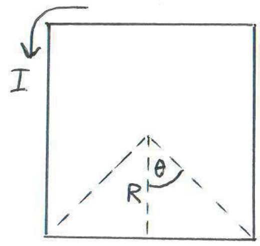
$$\mathbf{B}_{\text{dip}} = \frac{\mu_0}{4\pi}\frac{1}{r^3}2m\hat{\mathbf{k}}$$

b(ii) $\mathbf{p} = p\hat{\mathbf{k}}\,,\ \mathbf{p}\cdot\hat{\mathbf{r}} = 0$

$$\mathbf{E}_{\text{dip}} = -\frac{1}{4\pi\epsilon_0}\,\frac{1}{r^3}p\hat{\mathbf{k}}$$

$$\mathbf{B}_{\text{dip}} = -\frac{\mu_0}{4\pi}\frac{1}{r^3}m\hat{\mathbf{k}}$$

*C8.2 At the field point at the centre of the square, we calculate the magnetic field due to the lower, horizontal segment of current.



We use our equation from Lecture 15, just after figure 15–3, p. 4:

$$B = \frac{\mu_0}{4\pi}\frac{I}{R}(\sin\theta_2 - \sin\theta_1)$$

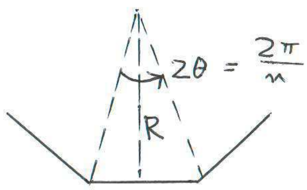
with $\theta_2 = -\theta_1 = \theta = 45° = \pi/4$. The field at the centre is *four times* this—an equal contribution coming from all four sides,

$$B = 4 \times \frac{\mu_0}{4\pi}\frac{I}{R} \times 2\sin\frac{\pi}{4} = \frac{\mu_0 I}{\pi}\frac{\sqrt{2}}{R}$$

a. The contribution from n sides of the polygon is

$$B = n \times \frac{\mu_0}{4\pi}\frac{I}{R} \times 2\sin\frac{\pi}{n} = \frac{\mu_0}{2\pi}\frac{I}{R} n\sin\frac{\pi}{n}$$

and you can see that for $n = 4$ the first expression is exactly that you obtained for the square.



b. As $n \to \infty$ we need

$$\lim_{n\to\infty} n\sin\frac{\pi}{n} = \infty \cdot 0$$

an “indeterminate form”. We write this as

$$\lim_{n\to\infty} \frac{\sin(\pi/n)}{1/n} = \frac{0}{0}$$

and differentiate top and bottom with respect to n,

$$\frac{(-\pi/n^2)\cos(\pi/n)}{-1/n^2} = \pi\cos\frac{\pi}{n}$$

$$\longrightarrow \pi \quad \text{as } n \to \infty$$

Hence

$$\lim_{n\to\infty} \frac{\mu_0 I}{2\pi}\frac{1}{R} n \sin\frac{\pi}{n} = B = \frac{1}{2}\mu_0 \frac{I}{R}$$

which is consistent with equation (15.3) in Lecture 15, p. 6, in the limit that $z = 0$, that is, in the centre of the loop.

4CCP1501 Problem class 9

Do these problems in class with the help of tutors

C9.1 A piece of wire having a resistance $R = 2\,\Omega$ (ohm) is bent into a ring enclosing an area of 8×10^{-4} m^2. It is oriented so that the plane of the ring is perpendicular to a magnetic field which increases at a fixed rate from 0.5 T to 2.5 T in a time interval of one second. Find the induced current in the wire.

C9.2 A particle with mass $m = 2 \times 10^{-16}$ kg and a charge $q = 30$ nC (nano-Coulomb) is accelerated from rest in a potential difference ΔV. It emerges from the accelerator into a region free of electric field but containing a uniform magnetic field of strength 0.6 T (tesla) pointing perpendicular to the particle's velocity. The particle moves in a circular orbit and returns to the point where it first emerged from the accelerator. In doing so its circular orbit encloses a magnetic flux of 15×10^{-6} Wb (weber).

a. Calculate the particle's orbital speed.

b. Find what was the accelerating voltage, ΔV.

C9.3 A long solenoid has radius a. An alternating current is passed through the coils so that the magnetic field inside the solenoid is

$$B(t) = B_0 \sin(2\pi\nu t)$$

as a function of time, t; ν is the frequency of the a.c. supply. A circular wire loop of radius r is placed concentric with the axis of the solenoid. In terms of the radius, r, the amplitude and frequency of the magnetic field, B_0 and ν, and the electrical *resistivity*, ρ, of the wire in the loop find an expression for the induced *current density*, J, in the wire as a function of time if the radius of the loop is

a. $r < a$, that is, *inside* the coil;

b. $r > a$, that is, *outside* the coil. Since there is no magnetic field outside the coil, comment on how an e.m.f. can be generated in this case.

Check that both your formulas give the same result for $r = a$ and plot the amplitude of the current density J as a function of the radius r of the loop.

[For the resistance of the wire, use $R = \rho L/A$ where L is the length of the wire, and A its cross sectional area.]

4CCP1501 Problem class 9—solutions

C9.1 The induced e.m.f. is given by the flux rule,

$$-\mathcal{E} = \frac{\mathrm{d}\Phi}{\mathrm{d}t}$$

and since the field increases at a fixed rate, we can write this as

$$\begin{aligned} -\mathcal{E} &= \frac{\Delta\Phi}{\Delta t} = \frac{(2.5-0.5)\times \text{area}}{\text{one second}} \\ &= \frac{2\times 8\times 10^{-4}}{\text{one}} = 1.6\times 10^{-3}\ \text{V} \end{aligned}$$

Therefore the current is

$$I = \frac{|\mathcal{E}|}{R} = \frac{1.6\ \text{mV}}{2\ \text{ohm}} = 0.8\ \text{mA}$$

C9.2 The flux enclosed by the particle's orbit is

$$\Phi = BA = B\,\pi r^2$$

where B is the magnetic field strength and r is the radius of the orbit. The particle experiences a Lorentz force which is equal to the centripetal force around the orbit. This leads to the force balance,

$$qvB = \frac{mv^2}{r}$$

which results in

$$r = \frac{mv}{qB}$$

and

$$\Phi = \frac{\pi m^2 v^2}{q^2 B}$$

a. From this we can calculate the speed,

$$v = \sqrt{\frac{\Phi q^2 B}{\pi m^2}} = 2.54\times 10^5 \quad \text{m s}^{-1}$$

b. The accelerating voltage gave the particle all its kinetic energy since it was accelerated from rest, so we must have

$$q\Delta V = \text{kinetic energy} = \frac{1}{2}mv^2$$

Hence

$$\Delta V = \frac{mv^2}{2q} = 215 \text{ V}$$

C9.3 We use the flux rule to find the induced e.m.f. in both cases. We start with

$$\begin{aligned} -\mathcal{E} &= \frac{\mathrm{d}\Phi}{\mathrm{d}t} = \text{area} \times \frac{\mathrm{d}B}{\mathrm{d}t} \\ &= \mathcal{A}_\Phi \, 2\pi\nu \, B_0 \cos 2\pi\nu t \end{aligned}$$

where $\mathcal{A}_\Phi$ is the area inside the loop which encloses magnetic field lines, and so the current is

$$\begin{aligned} I &= \frac{|\mathcal{E}|}{R} \\ &= \frac{\mathcal{A}_\Phi}{R} \, 2\pi\nu B_0 \cos 2\pi\nu t \end{aligned}$$

In each case the resistance of the wire is given by

$$R = 2\pi r \frac{\rho}{A} \longrightarrow \frac{1}{R} = \frac{A}{2\pi r\rho}$$

where A is the cross sectional area of the wire. The current density is

$$J = \frac{I}{A}$$

so A will not appear in our final formulas.

a. In this case the entire loop is within the magnetic field and so the area associated with the magnetic flux is the area of the loop,

$$\mathcal{A}_\Phi = \pi r^2$$

and we find for the current,

$$I = \frac{A}{2\pi r\rho} \, \pi r^2 \, 2\pi\nu B_0 \cos 2\pi\nu t$$

and so the current density is

$$J = \frac{I}{A} = \pi \frac{r}{\rho} \, \nu B_0 \cos 2\pi\nu t \qquad \longleftarrow r < a$$

b. In this case, whatever is the radius of the loop, when it's greater than the radius of the coil the area of magnetic flux is just the cross sectional area of the coil

$$\mathcal{A}_\Phi = \pi a^2$$

and we find

$$I = \frac{A}{2\pi r\rho} \, \pi a^2 \, 2\pi\nu B_0 \cos 2\pi\nu t$$

and so the current density is

$$J = \frac{I}{A} = \pi \frac{a^2}{r\rho}\, \nu B_0 \cos 2\pi\nu t \qquad \longleftarrow r > a$$

One explanation for this phenomenon is that while there is no *magnetic field* outside the coil, there is *vector potential* and it is this that induces the e.m.f. Interestingly the induced e.m.f. is *independent* of the radius of the wire loop, however the resistance of the loop increases with its radius so as $r \to \infty$ the induced current tends to zero.

Note that when $r = a$ the two expressions coincide as you'd hope they would. Note also that *inside* the solenoid the induced current density increases linearly with the radius of the loop and *outside* the coil it decays like $1/r$. You should plot this.

4CCP1501 Problem class 10

Do these problems in class with the help of tutors

C10.1 A source of alternating current produces an output electromotive force described by the formula $\mathcal{E} = \mathcal{E}_0 \sin(\omega t)$. This is connected in turn to a resistor of resistance R ohm; a capacitor of capacitance C farad; and an inductor of inductance L henry. For each of these three circuits, find an expression for the current I as a function of time. In the cases of the capacitor and the inductor identify what is the reactance of the circuit and by plotting voltage and current as functions of time, indicate in each case what is the phase difference and hence whether the current leads or trails the voltage. Think very carefully to find an explanation in your head as to how the phase difference between current and voltage is achieved in terms of the way charges are moving in each circuit.

C10.2 An electric circuit is constructed using a resistor of 2330 ohm, an inductor of 0.15 henry and a capacitor of 5×10^{-6} farad connected in series with an AC power supply.

(i) Find the natural frequency in Hz of the circuit.

(ii) Find the total reactance at frequencies of 50 Hz and 500 Hz. Be sure to include the units of reactance in your answer.

(iii) Use a graphical construction to estimate the phase difference between current and voltage at frequencies of 50 Hz and 500 Hz. In each case state whether the current leads or trails the voltage.

C10.3 **Why is there a 4π in the constant $1/4\pi\epsilon_0$?** Compare Coulomb's law with Newton's law of gravitation (which doesn't have a 4π). Then deduce Gauss's law for gravitation and use it to find the gravitational field due to an infinite sheet of mass of uniform density ρ_m. Now you find a π in the answer to a problem that contains no spherical shapes. Is that as it should be? Most professionals in physics *still* don't use SI units in electrodynamics. So πs appear in all the wrong places. Sommerfeld was urging against this as long ago as the 1920s.

4CCP1501 Problem class 10—solutions

C10.1 You will find the answer exactly reproduced in your Lecture 20 notes.

C10.2

(*i*) This is the frequency corresponding to the peak current in the undamped ($R = 0$) circuit. Hence it is the solution of

$$\omega L = \frac{1}{\omega C}$$

namely

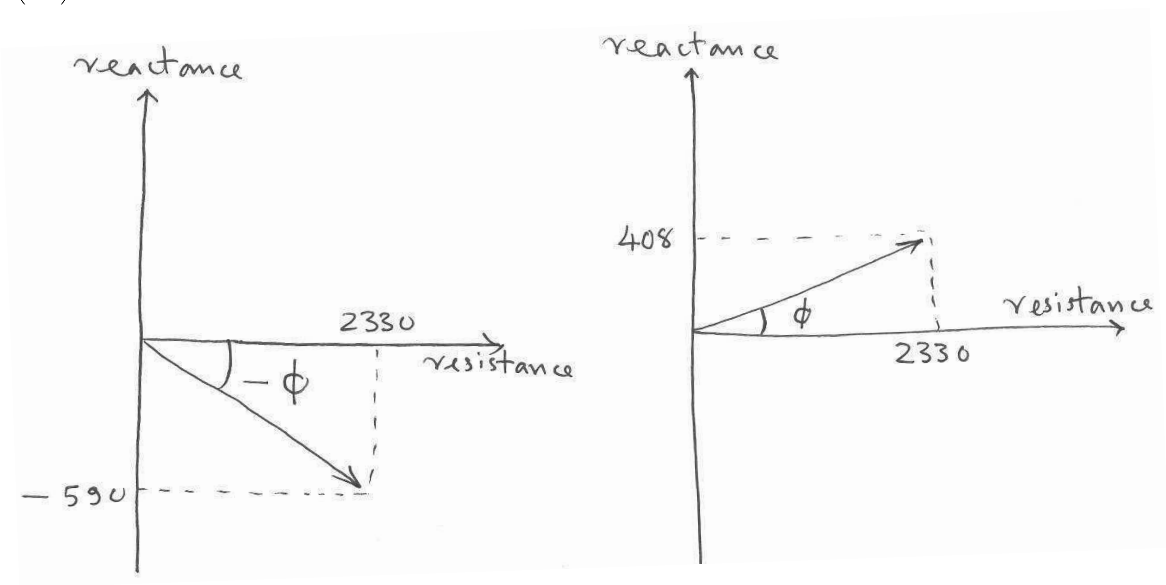
$$\omega_0 = \frac{1}{\sqrt{LC}} = 1154.7 \text{ rad s}^{-1}$$

Therefore the natural frequency is $\omega_0/2\pi = 183.3$ Hz

(*ii*)

50 Hz	:	reactance = −589.5 ohm
500 Hz	:	reactance = 407.6 ohm

(*iii*)



current leads voltage — current trails voltage

C10.3 I'll leave you to figure this out.

4CCP1501 Tutorial 1

T1.1 What is the *principle of superposition* of forces?

T1.2 An object of mass 1.5 kg experiences three forces. These are, in newton,

$$\mathbf{F}_1 = 1.2\hat{\mathbf{\imath}} + 3.3\hat{\mathbf{\jmath}} + 1.6\hat{\mathbf{k}}$$
$$\mathbf{F}_2 = 3.3\hat{\mathbf{\imath}} + 1.4\hat{\mathbf{\jmath}} + 2.9\hat{\mathbf{k}}$$
$$\mathbf{F}_3 = 1.5\hat{\mathbf{\imath}} + 2.7\hat{\mathbf{\jmath}} - 5.1\hat{\mathbf{k}}$$

Find the total force and the acceleration of the object.

T1.3 A greengrocer attaches a bunch of bananas to her hanging scales. The spring extends by 0.3 m. She then sets the bananas into oscillation. What is the frequency of this oscillation in Hz? What approximations have you made?

T1.4 Show that for a linear spring of spring constant k, the potential energy stored in the spring is

$$\text{potential energy} = \frac{1}{2}k(\Delta x)^2$$

if it is stretched by an amount Δx.

4CCP1501 Solutions to Tutorial 1

T1.1 See your lecture notes.

T1.2

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3 = 6.0\hat{\mathbf{\imath}} + 7.4\hat{\mathbf{\jmath}} - 0.6\hat{\mathbf{k}}$$

The magnitude of the total force is

$$F = \sqrt{6.0^2 + 7.4^2 + 0.6^2} = 9.6 \text{ N}$$

The mass, m, is given as 1.5 kg and using $F = ma$ we get

$$a = 9.6/1.5 = 6.4 \text{ m s}^{-2}$$

T1.3 The spring constant, k, is *force per unit displacement*:

$$k = \frac{mg}{\Delta x} \tag{1}$$

Here, m is the mass and g is the acceleration due to gravity, so mg is the force acting on the bananas due to gravity. Δx is the displacement. (We always use the Greek upper case Δ to indicate a "finite change in…". Conversely we use the Roman lower case d to indicate an infinitesimal change as in the differential calculus. Sometimes the Greek lower case δ is used to mean a "small, but yet finite change in…")

The number of oscillations per second is the *frequency*, f. The *angular frequency* is

$$\omega = 2\pi f$$

and is given in terms of the spring constant and the mass for the mass-on-a-spring as

$$\omega = \sqrt{\frac{k}{m}} \tag{2}$$

We may use $g = 9.81$ m s^{-2}, and we are told that the displacement of the spring due to the bananas is $\Delta x = 0.3$ m. This allows us to calculate ω and hence f. You may think that we need to know the mass and that I've not given you enough information to solve the problem. But if you substitute equation (1) into (2) then the mass cancels. You have,

$$\omega = \sqrt{\frac{g}{\Delta x}} = \sqrt{\frac{9.81}{0.3}} = 5.72 \text{ rad s}^{-1}$$

Hence the frequency is

$$\frac{\omega}{2\pi} = 0.91 \text{ Hz}.$$

The approximations you have made are (i) to neglect the mass of the spring itself; (ii) that the spring is *linear*, that means the amount of stretching (the "displacement") is directly proportional to the force applied, or in other words that the spring constant is indeed *constant* and independent of the extension of the spring; and (iii) that there is no damping. In real life it is impossible to remove all damping effects, such as those arising from heating of the spring due to *anelastic* processes among the atoms of the metal and friction from the air resistance, for example. It is also not easy to design a perfectly linear spring. In physics we are always making idealisations like this so that we can get down to the fundamental theory. Later when we need to interpret experiments we can make *corrections*.

Students often make the mistake of confusing ω with f, and getting the dimensions wrong. f is in units of s^{-1} (or Hz), and ω is in radians s^{-1}. So students often make errors of 2π in their answers. Also, if the question asks you for the frequency, then work out the angular frequency and convert that to Hz.

T1.4 As I stretch the spring from zero to a displacement of Δx the amount of force I have to apply increases as the spring is stretched. When the spring is stretched to an amount x, I have applied an amount of force $F = kx$. Then to stretch it by a further infinitesimal amount $\mathrm{d}x$ I need to do an infinitesimal amount of additional work,

$$\mathrm{d}W = \text{force} \times \text{distance} = F \times \mathrm{d}x = kx\,\mathrm{d}x$$

So the total amount of work I have to do to stretch the spring from zero to Δx can be found by an integration of the work,

$$W = \int_0^{\Delta x} \mathrm{d}W = k \int_0^{\Delta x} x\mathrm{d}x = \frac{1}{2}k(\Delta x)^2 \tag{3}$$

The work done is the potential energy stored in the spring. Note that k is a constant so it comes to the front of the integral sign.

You can see this easily if you just plot force against distance, which is a straight line with slope k if the spring is linear. Then you note that the work done is the area under the line between zero and Δx.

You may have argued that if energy is force times distance then if the force is $k\Delta x$ and the distance is Δx then the energy is $k(\Delta x)^2$, but then you're out by a factor of two. Make sure you understand why there has to be a half in equation (3): it's because, as I said, the amount of work you need to stretch the spring by a given amount increases as the spring becomes stretched. You can try it with an elastic band.

4CCP1501 Tutorial 2

T2.1 Zoe waits on a railway platform while two trains approach from the same direction and at the same speed of 8 m s^{-1}. Neither of them is going to stop, so both trains are sounding their whistles, which have the same frequency; and one train is some distance behind the other. After the first train passes Zoe but before the second train has passed her, she hears a beating in the volume of the sound of the whistles having a frequency of 4 Hz. What is the frequency of the trains' whistles?

You will need to know about the *Doppler effect*: if a wave has a phase velocity of v and is emitted at frequency f from a source travelling *towards* an observer with speed v_s then the frequency perceived by the observer travelling *towards* the source with speed v_o is

$$f' = \left(\frac{v + v_o}{v - v_s}\right) f$$

[The speed of sound in air is 343 m s^{-1}]

4CCP1501 Solutions to Tutorial 2

T2.1 Since Zoe is stationary, we use the formula

$$f' = \left(\frac{v}{v - v_s}\right) f$$

We need to find f. From the approaching train, Zoe hears a whistle of frequency

$$f_1 = \left(\frac{343}{343 - 8}\right) f$$

and from the receding train, she hears a frequency

$$f_2 = \left(\frac{343}{343 + 8}\right) f$$

From our notes we know that the beat frequency is $(f_1 - f_2) = 4$ Hz because the receding train produces a note of lower frequency. This leads to

$$4 \text{ Hz} = \left(\frac{343}{343 - 8}\right) f - \left(\frac{343}{343 + 8}\right) f$$

which we solve for f to get $f = 85.7$ Hz.

4CCP1501 Tutorial 3

T3.1 A mass of 0.5 kg is suspended from a rigid support by a spring which stretches by 0.08 m. Find the angular frequency.

Write down an equation for the displacement of the mass as a function of time in terms of the amplitude and angular frequency. Write down an equation for the speed of the mass. Using these two equations show that the total energy is constant and independent of the time.

The oscillator is now subjected to a damping force such that the damping ratio is 0.01. Calculate the frequency of the oscillator. As the oscillation is damped so that the mass comes to rest, estimate the fraction of energy that is lost over each period. Hence determine the quality factor of the damped oscillator.

The oscillator is now driven by a sinusoidal force of amplitude 1 N. Calculate the resonant angular frequency. Calculate the static amplitude and the amplitude at resonance. What is the amplification factor at resonance?

4CCP1501 Solutions to Tutorial 3

T3.1 Spring constant is

$$k = \frac{0.5 \times g}{0.08} = \frac{4.905}{0.08} = 61.313 \text{ M m}^{-1}$$

Hence

$$\omega_0 = \sqrt{\frac{k}{m}} = 11.074 \text{ rad s}^{-1}$$

$$\begin{aligned}
x &= A\cos\omega_0 t \\
\text{speed} &= -A\omega_0 \sin\omega_0 t \\
\text{Total energy} &= \text{kinetic energy} + \text{potential energy} \\
&= \frac{1}{2}mv^2 + \frac{1}{2}kx^2 \\
&= \frac{1}{2}mv^2 + \frac{1}{2}\omega_0^2 m x^2 \\
&= \frac{1}{2}mA^2\omega_0^2 \sin^2\omega_0 t + \frac{1}{2}mA^2\omega_0^2\cos^2\omega_0 t \\
&= \frac{1}{2}mA^2\omega_0^2 \\
&= \frac{1}{2}kA^2
\end{aligned}$$

This is independent of t.

The damping ratio, $Z = 0.01$; $\omega_D = \omega_0\sqrt{1-Z^2} = 11.073 \text{ rad s}^{-1}$. So the damped frequency is 1.762 Hz.

The rate of decay of the average energy is

$$\frac{\mathrm{d}\overline{E}}{\mathrm{d}t} = -2Z\omega_0\overline{E}$$

Over one period of oscillation the energy dissipated is

$$\begin{aligned}
\Delta\overline{E} &= -\frac{\mathrm{d}\overline{E}}{\mathrm{d}t} \times \text{period} \\
&= 2Z\omega_0\overline{E} \times \frac{2\pi}{\omega_D}
\end{aligned}$$

so the fraction of energy lost is

$$\frac{\Delta\overline{E}}{\overline{E}} = 4\pi\, Z\,\frac{\omega_0}{\omega_D} \approx 4\pi Z = 0.126$$

which is the *specific damping capacity*, S. The *quality factor*, Q, is

$$Q = \frac{2\pi}{S} \approx \frac{1}{2Z}$$

Hence $S = 0.126$ and $Q = 50$.

$$\omega_{\rm max} = \omega_0\sqrt{1-2Z^2} = 11.072 \text{ rad s}^{-1}$$

Hence the resonant freqency is 1.762 Hz.

The static amplitude is $A_{\rm s} = F_0/k = 1/61.3 = 0.016$ m. The amplitude at resonance is $A_{\rm s}Q = 0.016 \times 50 = 0.8$ m. The amplification factor at resonance is $A_{\rm max}/A_{\rm s} = 0.8/0.016 = 50$ (which is of course Q).

4CCP1501 Tutorial 4

T4.1 A mechanical simple harmonic oscillator is affected by viscous forces, having a damping ratio, $Z = 0.1$. It is driven by a sinusoidal force of angular frequency ω. Use a computer plotting program to make a graph of the *magnification factor*, $D_s = A/A_\mathrm{s}$, as a function of the frequency ratio, $r = \omega/\omega_0$. Here, A is the amplitude and A_s is the "static amplitude" (or amplitude in the limit of zero driving frequency). ω_0 is the natural angular frequency of the oscillator in the absence of damping. On your plot, indicate the values of ω_D/ω_0 and $\omega_\mathrm{max}/\omega_0$, where ω_D is the natural angular frequency of the damped oscillator and ω_max is the resonant frequency. Also mark on your plot the value of the quality factor, Q.

If you are good at plotting, then try and vary the damping ratio to produce a family of resonance curves. Please don't use excel: teach yourself a grown-up plotting program like gnuplot, python or MATLAB.

[HINT: You are being asked to plot the function

$$D_s = \frac{A}{A_s} = \frac{1}{\sqrt{\left(1 - r^2\right)^2 + \left(2rZ\right)^2}}$$

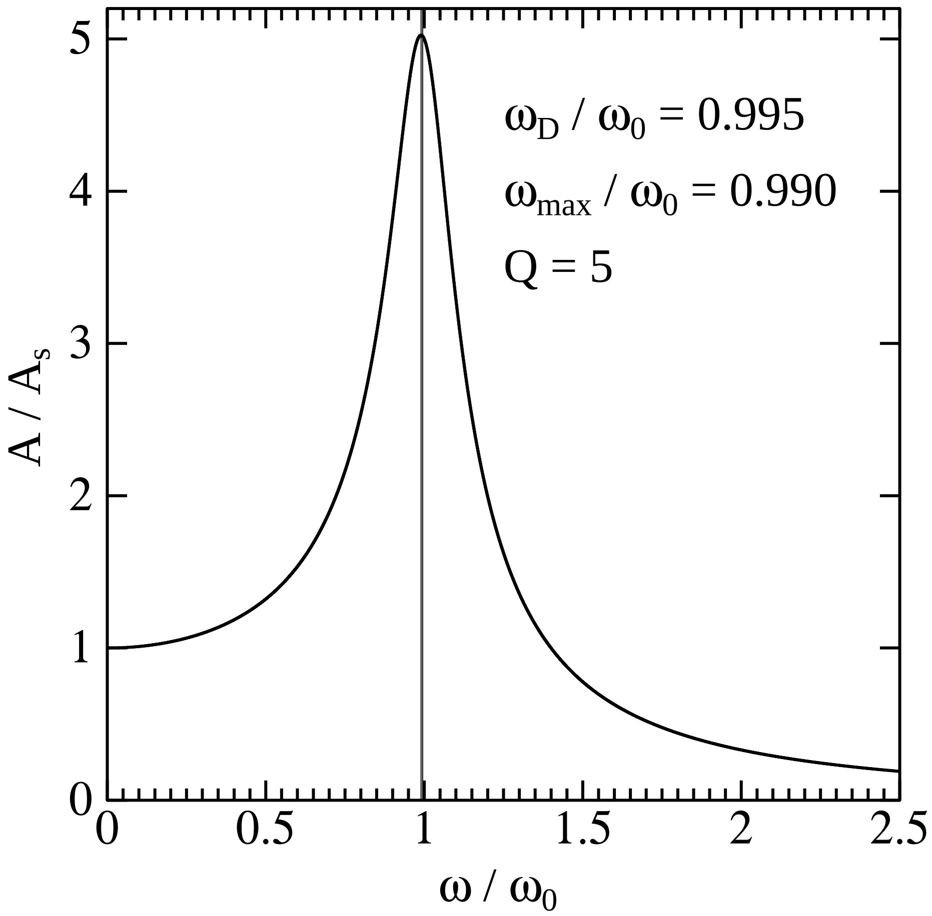
which is equation (5) in your *additional notes* on simple harmonic motion (at `KEATS`). You will find examples of the curves on page 9 of these notes. If you are bold and wish to go further you may plot and think about the *phase angle*

$$\phi = \arctan\frac{2rZ}{1 - r^2}$$

which is the phase difference between the driving force and the oscillator. It is not at all obvious how and why the response of the driven oscillator is out of phase with the driving force.]

4CCP1501 Solutions to Tutorial 4

T4.1 Here's my curve:



Note that even with fairly large damping, $\omega_{\max} = \omega_D = \omega_0$ to within 1%. So on a cartoon graph such as in the lecture notes I mark these off to indicate $\omega_{\max} > \omega_D > \omega_0$ but in most practical cases of under damping they can be regarded as equal.

Note that the high frequency tail goes like $1/\omega^2$.

4CCP1501 Tutorial 5

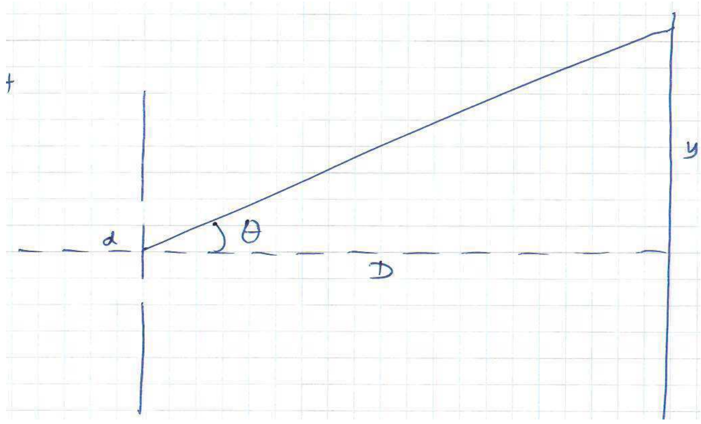
T5.1 Light with wavelength 442×10^{-9} m passes through a Young's double slit apparatus in which the slits are separated by $d = 0.4 \times 10^{-3}$ m. How far away must the screen be placed so that dark fringes appear directly opposite each slit with only one bright fringe in between them? [Use the small angle approximation, $\sin\theta \approx \tan\theta \approx \theta$.]

T5.2 Very low frequency (VLF) radio waves have wavelengths in the range 10 to 100 kilometers. A surfaced submarine sends a VLF signal which is received by a receiving station in Antarctica via two paths. One is the direct path of 1000 km between submarine and receiver; the other is by reflection of the signal from the ionosphere, which is a layer of ionised molecules in the upper atmosphere having a smaller refractive index than air. Reflection happens at a point midway between the station and the receiver, and this leads to a signal of wavelength 15 km being particularly weak at the receiver. Calculate the minimum height of the ionosphere.

T5.3 You are holding a laser that emits at a wavelength of 632.8 nm. You attach an opaque sheet containing two slits to the front of the laser and project the interference pattern on to a screen. The slits are separated by 0.3 mm. Now you walk towards the screen at a speed of 3 m s^{-1} so that the fringes move closer together. What is the speed of the 50$^{\text{th}}$-order bright fringe?

4CCP1501 Solutions to Tutorial 5

T5.1



For dark fringes, $d\sin\theta = \left(m+\frac{1}{2}\right)\lambda$. We want the zero-order fringe, $m=0$, so

$$d\sin\theta = \frac{1}{2}\lambda \approx \theta d$$

From the diagram,

$$\tan\theta = \frac{y}{D} \approx \theta$$

and we want $y = \frac{1}{2}d$ so the dark fringe is opposite the slit. We write

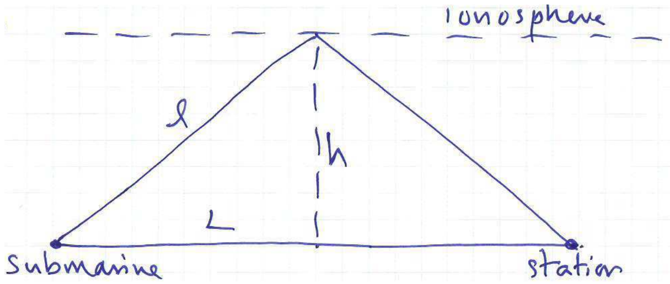
$$\theta d \approx \frac{1}{2}\lambda \Rightarrow \theta = \frac{1}{2}\frac{\lambda}{d}$$

$$\tan\theta = \frac{d/2}{D} \approx \theta \Rightarrow D = \frac{1}{2}\frac{d}{\theta}$$

Therefore

$$D = \frac{1}{2}d\,\frac{2d}{\lambda} = \frac{d^2}{\lambda} = \frac{(0.4\times10^{-3})^2}{442\times10^{-9}} = 0.362\ \mathrm{m}$$

T5.2



We want to find h in the diagram. The path difference for a first order *minimum* is $\frac{1}{2}\lambda$ and so we have

$$\text{path difference} = \frac{1}{2}\lambda = 2\ell - 2L = 2\left(\sqrt{L^2+h^2} - L\right)$$

$$\Rightarrow \frac{1}{2}\lambda + 2L = 2\sqrt{L^2+h^2}$$

$$\Rightarrow \frac{1}{4}\left(\frac{1}{2}\lambda + 2L\right)^2 - L^2 = h^2$$

$$\Rightarrow h^2 = \frac{1}{4}(7.5+1000)^2 - 1000^2 \text{ km}^2$$

$$\Rightarrow h = 61.352 \text{ km}$$

T5.3 The 50^{th}-order bright fringe is located at an angle θ given by

$$d\sin\theta = m\lambda \quad \Rightarrow \theta = \arcsin\frac{m\lambda}{d}$$

In our notes we use D for the distance between the slits and the screen, and y for the distance from the centreline to the fringe, that is the distance from the central maximum to the m^{th}-order fringe. As $\tan\theta = y/D$ and θ is unchanging, as you walk towards the screen the ratio y/D remains constant. Now,

$$\begin{aligned} y &= D\tan\theta \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= \frac{\mathrm{d}D}{\mathrm{d}t}\tan\theta \\ &= -v\tan\theta \end{aligned}$$

where v is the speed you are walking at. The minus sign appears because $\dot{D}$ is negative since D is getting smaller as you walk. So the speed of the fringe is

$$\begin{aligned} v_{\text{fringe}} &= -\frac{\mathrm{d}y}{\mathrm{d}t} = v\tan\theta \\ &= v\tan\left[\arcsin\left(\frac{m\lambda}{d}\right)\right] \\ &= 3\times\tan\left[\arcsin\left(\frac{50\times 632.8\times 10^{-9}}{0.3\times 10^{-3}}\right)\right] \\ &= \mathbf{0.318\ m\ s^{-1}} \end{aligned}$$

4CCP1501 Tutorial 6

T6.1 We found in Lecture 7 that two melons of mass 1 kg one metre apart experience an attractive force of about 10^{-12} N due to gravity. Suppose each melon has 0.01% more electrons than protons and make an order of magnitude estimate of the repulsive force between them.

T6.2 In a hydrogen atom, an electron of charge $-e = -1.602 \times 10^{-19}$ C is found near a proton of charge $+e$ at an average distance of $a_0 = 0.0529$ nm (1 nm = 10^{-9} m); this distance is called the "Bohr radius." Calculate the work I would have to do to remove this electron. [HINT: integrate the force, $\int F \mathrm{d}r$, from a_0 to infinity.] You can do this in SI units if you like, but it's especially easy to use "Rydberg atomic units," in which $a_0 = 1$, $e^2 = 2$ and $4\pi\epsilon_0 = 1$. The energy you get is then in Rydbergs. The known binding energy of the hydrogen atom is -1 Ry = –13.6 eV (eV is an "electron volt," 1 eV = 1.602×10^{-19} J). In your tutorial, discuss why you don't get this answer.

4CCP1501 Solutions to Tutorial 6

T6.1 Each melon has 0.01% more electrons than protons. So the number of electrons is

$$n_e = (1 + 0.0001)n_p$$

where n_p is the number of protons. The total charge Q on each melon is the difference in numbers of electrons and protons times the charge $-e$ on each electron,

$$\begin{aligned} Q &= -e(n_e - n_p) \\ &= -e(1.0001n_p - n_p) \\ &= -0.0001n_p e \end{aligned} \tag{1}$$

Now, how many protons are there in a melon? Most atoms have roughly equal numbers of protons and neutrons and the mass of the electrons is small enough to be neglected; so about half the mass of a melon is made up of protons. So if the melon has a mass m and the mass of a proton in m_p then

$$n_p = \frac{1}{2}\frac{m}{m_p}$$

With $m = 1$ kg and $m_p = 1.67 \times 10^{-27}$ kg this gives us $n_p = 3 \times 10^{26}$ and putting this in equation (1) with $e = 1.6 \times 10^{-19}$ C we get $Q = -4790$ C. Then finally the force acting between the two melons, one meter apart, is

$$\begin{aligned} F &= 9 \times 10^9 \frac{Q^2}{\text{one}^2} \\ &= 2 \times 10^{17} \quad [\text{N}] \end{aligned}$$

This is of course a truly enormous force. Even if the melon had only 0.000001% more electrons that protons the force would still be over 2000 kN—greater than the capability of a large tensile testing machine.

T6.1 Working in atomic units, the force between the proton and the electron when a distance r apart has magnitude $-e^2/r^2 = -2/r^2$. I want to integrate force × distance from a_0 to infinity to get the work done in removing the electron. The force that I apply is minus the force acting on the electron due to the proton, namely e^2/r^2, so

$$\begin{aligned} E &= e^2 \int_{a_0}^{\infty} \frac{\mathrm{d}r}{r^2} \\ &= e^2 \left[-\frac{1}{r}\right]_{a_0}^{\infty} \\ &= \frac{e^2}{a_0} = 2 \qquad [\text{Rydberg}] \end{aligned}$$

Why isn't this the known binding energy of minus one Rydberg? Because we have calculated only the potential energy of the hydrogen atom. We need to add the kinetic energy of the motion of the two particles. Surely this is a hard problem in quantum mechanics? Yes, it is; but we can appeal to the *virial theorem* which applies in quantum mechanics as well as in classical mechanics. For a system whose potential energy is proportional to r^{-1}, we know that

$$2 \times \text{kinetic energy} = -\text{potential energy}$$

and of course

$$\text{total energy} = \text{potential energy} + \text{kinetic energy}$$

So you see from these two expressions that the *total energy* of the hydrogen atom with respect to the proton and electron at infinite separation is indeed minus one Rydberg. The calculation of the potential energy by a simple integration is quite easy. When we come to study the *electric potential* you will see there is an even easier way to obtain this result.

Additional note: when I wrote this tutorial I was not sure whether we would have covered the electric potential yet. But we have and so you should check in your head that you can actually write down the work done in removing the electron without doing an integration.

4CCP1501 Tutorial 7

T7.1 We're pretty certain that the electron is an elementary particle. According to `wikipedia` its radius is certainly less than 10^{-22} m. But if it *is* a point charge then as we have seen in Lecture 12 its electrostatic self energy is infinite. So let's try the following line of argument. Suppose the electron is a sphere of radius r_e and its charge $-e$ is uniformly distributed throughout the sphere. Show that the electrostatic energy is

$$W = \frac{1}{4\pi\epsilon_0}\,\frac{3}{5}\frac{e^2}{r_e}$$

If you find this too hard, do the case where the charge $-e$ is distributed evenly over the surface of the sphere and obtain

$$W = \frac{1}{4\pi\epsilon_0}\,\frac{1}{2}\frac{e^2}{r_e}$$

If that's also too hard, don't worry just take the results as given and drop the half or the three-fifths. Then equate this energy with the rest energy of the electron by writing

$$\frac{1}{4\pi\epsilon_0}\,\frac{e^2}{r_e} = mc^2$$

Then visit `http://physics.nist.gov/cuu/Constants` and work out the numerical value of r_e which is called the "classical electron radius". Incidentally, what are the classical radii of the μ and τ leptons?

T7.1 Let's think a bit more about the "size" of an electron. Find out a bit about the Compton experiment. This was the experiment that drove the final nail into the coffin of the notion that light is a wave. Photons were scattered by the free electrons in a graphite sheet and the difference in wavelength between incident and scattered X-rays was found to depend on the scattering angle, θ, through the *Compton shift equation*

$$\Delta\lambda = \frac{h}{mc}\,(1-\cos\theta)$$

The largest observed Compton shift is called the "Compton wavelength",

$$\lambda_c = \frac{h}{mc}$$

We could use this, or maybe,

$$\bar{\lambda}_c = \frac{\hbar}{mc}$$

as a measure of the size of an electron. Here h is the Planck constant and $\hbar = h/2\pi$ is called the "reduced Planck constant". In your last tutorial, T6, you encountered the hydrogen atom and in its ground state the electron is in an orbit whose "radius" is a_0 (the *Bohr radius*). Actually this is the average, or "expectation" value of the measured distance from the proton to the electron. This provides a *third* estimate of the "size" of the electron. Using the definition of the fine structure constant,

$$\alpha = \frac{1}{4\pi\epsilon_0} \frac{e^2}{\hbar c} \approx \frac{1}{137}$$

show that

$$r_e = \alpha \lambdabar_c = \alpha^2 a_0$$

that is, each estimate, $r_e < \lambdabar_c < a_0$, becomes smaller in powers of the fine structure constant. You may want to read around a bit about the Planck constant or the fine structure constant, and discuss in your tutorial any consequences of this calculation.

4CCP1501 Solutions to Tutorial 7

T7.1 The problem of the uniformly charged sphere is an important one; the result is needed in pseudopotential theory and the atomic spheres approximation in solid state physics. As is often the case, there are two ways to solve it—using Gauss's law and the electric field, and using the electric potential.

If we use Gauss's law then we first need to find the field outside the sphere. That's easy, it's

$$E = \frac{1}{4\pi\epsilon_0}\frac{-e}{r^2} \quad , \qquad r > r_e$$

exactly as for a point charge. So the energy stored in that part of space is

$$\begin{aligned} W_{\text{out}} &= \frac{1}{2}\epsilon_0 \int_{r_e}^{\infty} E^2\, 4\pi r^2 \mathrm{d}r \\ &= \frac{1}{4\pi\epsilon_0}\,\frac{1}{2}e^2 \int_{r_e}^{\infty} \frac{1}{r^2}\mathrm{d}r \end{aligned}$$

It's as for the point charge in Lecture 12 but now the lower limit doesn't trouble us and the result is finite. So we have

$$W_{\text{out}} = \frac{1}{4\pi\epsilon_0}\,\frac{1}{2}\frac{e^2}{r_e}$$

This is the solution if the charge is confined to the surface of the sphere, as it would be if the we were dealing with a *metal* sphere or with the charged shell of Lecture 11, since inside such a sphere the electric field is zero as a gaussian surface contains no charge. But to do the case of a uniformly charged sphere we also need the field *inside*. A gaussian surface of radius $r < r_e$ encloses a volume $(4/3)\pi r^3$ and hence an amount of charge

$$Q_{\text{enclosed}} = \rho\,\frac{4}{3}\pi r^3 \qquad \text{density} \times \text{volume}$$

and the charge density within the sphere, being uniform, is

$$\rho = \frac{-e}{(4/3)\pi r_e^3} \qquad \text{charge} \div \text{volume} \tag{1}$$

Gauss's law states

$$\text{flux through gaussian surface} = \text{electric field} \times \text{area} = \frac{1}{\epsilon_0}Q_{\text{enclosed}}$$

That is,

$$E \times 4\pi r^2 = \frac{1}{\epsilon_0}\rho\,\frac{4}{3}\,\pi r^3$$

so the electric field at radius r inside the sphere is

$$E = \frac{1}{4\pi\epsilon_0}\,\frac{4}{3}\pi r \rho \quad , \qquad r < r_e$$

(note it increases with r as you'd expect; it's a good idea to sketch the field as a function of r at this point, both inside and outside the sphere) and so

$$\begin{aligned} W_{\text{in}} &= \frac{1}{2}\epsilon_0 \int_0^{r_e} E^2\, 4\pi r^2\, \mathrm{d}r \\ &= \frac{1}{4\pi\epsilon_0}\,\frac{1}{2}\left(\frac{4}{3}\pi\rho\right)^2 \underbrace{\int_0^{r_e} r^4\, \mathrm{d}r}_{=\frac{1}{5}r_e^5} \end{aligned}$$

So from the electric field inside the sphere the energy is, using (1),

$$W_{\text{in}} = \frac{1}{4\pi\epsilon_0}\,\frac{1}{10}\,\frac{e^2}{r_e}$$

and if we add this to the energy from the field outside the sphere we get the final result, which is

$$W = W_{\text{in}} + W_{\text{out}} = \frac{1}{4\pi\epsilon_0}\,\frac{3}{5}\,\frac{e^2}{r_e}$$

To do the same problem using electric potential proceed as follows. We will bring in from infinity little increments of charge and plaster them evenly over the surface of a growing sphere of uniform charge density

$$\rho = \frac{-e}{(4/3)\pi r_e^3}$$

At every stage of this process the sphere has a radius r and its total charge up to this stage is

$$q(r) = \frac{4}{3}\pi r^3\,\rho \qquad \text{volume of growing sphere} \times \text{density}$$

We bring in an increment of charge $\mathrm{d}q$ from infinity and add it to the sphere in the form of a spherical shell of charge of volume $4\pi r^2\mathrm{d}r$ so the increment of charge is

$$\mathrm{d}q = \rho\, 4\pi r^2 \mathrm{d}r \qquad \text{density} \times \text{volume of shell}$$

The electric potential at radius r due to the charge already accumulated is

$$V(r) = \frac{1}{4\pi\epsilon_0}\,\frac{q(r)}{r} = \frac{1}{4\pi\epsilon_0}\,\frac{4}{3}\pi r^2\,\rho$$

and so the work done in bringing in this increment of charge $\mathrm{d}q$ is

$$\mathrm{d}W = V(r)\,\mathrm{d}q = \frac{1}{4\pi\epsilon_0}\,\frac{16}{3}\pi^2\rho^2 r^4 \mathrm{d}r$$

So the *total* work done in assembling the charged sphere is

$$W = \int_0^{r_e} = \frac{1}{4\pi\epsilon_0}\frac{16}{3}\pi^2\rho^2 \underbrace{\int_0^{r_e} r^4\,\mathrm{d}r}_{=\frac{1}{5}r_e^5}$$

$$= \frac{1}{4\pi\epsilon_0}\frac{3}{5}\frac{e^2}{r_e}$$

using the formula (1) for ρ. This is of course the same result that we got using Gauss's law and as is generally the case fewer steps are needed using the electric potential rather than the field.

It is interesting to compare the two approaches to calculating the energy: from the electric *field* and from the electric *potential*. In the first instance you are required to do two integrals and when combined they reach from the centre of the electron to infinity. This underlines the point made in lecture 12, at equation (12.5), that the electrostatic energy is "stored" in the empty space surrounding the point charge. This is especially evident if you did the case of the metal sphere or hollow shell because then *none* of the electrostatic energy arises from within the sphere where the field is zero. On the other hand, in the case of using the electric potential to do the calculation you are only required to do an integral from the centre to the surface of the sphere, so where does the empty space come in? But remember that the space outside is highly important there too—you need to carry each little increment of charge through that space from infinity. Maybe this picture is physically more appealing, in the sense that you're not required to believe, as the Victorians did, that the electrostatic energy is stored in the elastic distortion of the ether, in the same way as energy in a metal spring is stored in the elastic strain of the metal crystal lattice. Thanks to Einstein we know that there is no ether.

So to finish the problem we now have

$$r_e = \frac{1}{4\pi\epsilon_0}\frac{e^2}{mc^2}$$

and if I look up the masses of the leptons and I use

$$\frac{1}{4\pi\epsilon_0} = 9\times 10^9 \qquad [\mathrm{N\ m^2\ C^{-2}}] \qquad \text{and} \qquad c = 3\times 10^8 \qquad [\mathrm{m\ s^{-2}}]$$

I get

$$\begin{aligned} r_e &= 2.8\times 10^{-15} \quad [\mathrm{m}] \\ r_\mu &= 1.4\times 10^{-17} \quad [\mathrm{m}] \\ r_\tau &= 8.1\times 10^{-19} \quad [\mathrm{m}] \end{aligned}$$

Note the more massive particles have a smaller classical radius.

T7.2

$$
\begin{aligned}
\alpha &= \frac{1}{4\pi\epsilon_0}\,\frac{e^2}{\hbar c} \\
a_0 &= 4\pi\epsilon_0 \frac{\hbar^2}{me^2} \\
\lambdabar_c &= \frac{\hbar}{mc} \\
r_e &= \frac{1}{4\pi\epsilon_0}\,\frac{e^2}{mc^2} \\
\alpha\lambdabar_c &= \frac{1}{4\pi\epsilon_0}\,\frac{e^2}{mc^2} \\
\alpha^2 a_0 &= \frac{1}{4\pi\epsilon_0}\,\frac{e^2}{mc^2}
\end{aligned}
$$

Note that all of these quantities *except* r_e involve the Planck constant. That's why it's called the *classical* electron radius. Compton scattering is the *inelastic* scattering of photons by free electrons. *Elastic* scattering of photons by free electrons is called Thomson scattering (named after J. J. Thomson) and in that experiment the differential cross section for scattering into an angle θ is $\frac{1}{2}\,r_e^2\,(1+\cos^2\theta)$. This is interesting because the Thomson scatter into zero, or glancing, angle gives a *measure* of the size of the electron which is consistent with the construction in question T7.1 using electrostatics and special relativity. It is curious, that when electrons scatter photons inelastically (that is they exchange energy with the photon) they appear to be larger, by a factor of α, than when they scatter elastically. You might ponder on why that is, or ask your tutor.

4CCP1501 Tutorial 8

T8.1 Use the Biot–Savart law to show that the magnetic field at a perpendicular distance s from the mid-point of a straight wire of length L carrying an electric current I has magnitude

$$B = \frac{\mu_0}{4\pi} \frac{I}{s} \frac{2}{\sqrt{1 + 4s^2/L^2}}$$

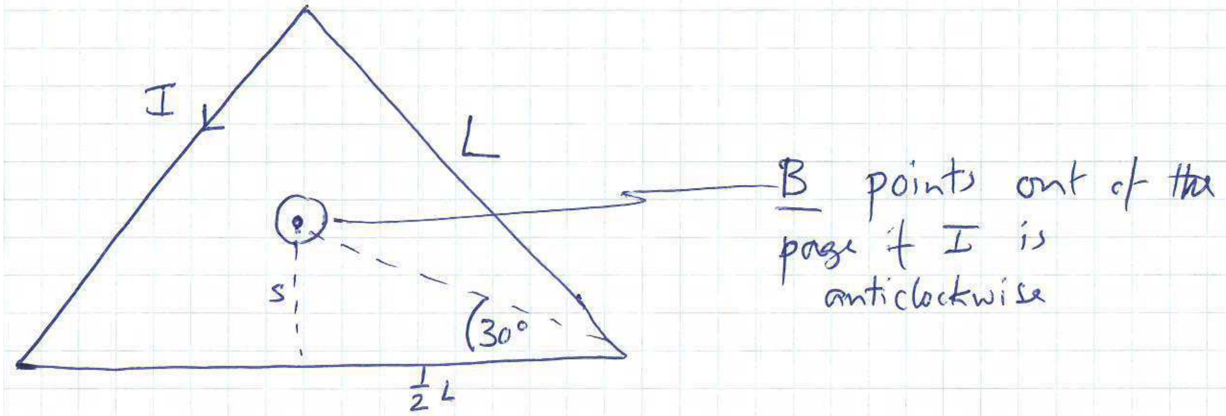
T8.2 Find an expression for the magnitude of the magnetic field at the centre of a loop of wire that is shaped into an equilateral triangle of side L and is carrying a current I. Draw a sketch to show the direction of the current and the direction of the magnetic field.

T8.3 What is the magnitude of the magnetic moment of this object? Draw a sketch to show the direction of the magnetic moment vector. A magnetic field of strength B is applied along a direction that is 30° from the direction of the magnetic moment. Find an expression for the torque acting on the triangular loop in terms of B, I and L. In what direction is the torque pointing?

4CCP1501 Solutions to Tutorial 8

T8.1 See your lecture notes.

T8.2



First find the length, s:

$$s = \frac{1}{2}L\,\tan 30^\circ = \frac{1}{2\sqrt{3}}\,L$$

Hence

$$4s^2 = 4 \times \frac{L^2}{12} = \frac{1}{3}L^2$$

So

$$\frac{4s^2}{L^2} = \frac{1}{3} \qquad ; \qquad \sqrt{1 + \frac{4s^2}{L^2}} = \sqrt{\frac{4}{3}} = \frac{2}{\sqrt{3}}$$

Now we use the formula from the question, recognising that the field in the centre, by the principle of superposition, is three times that from one side of length, L,

$$\begin{aligned} B &= 3\,\frac{\mu_0}{4\pi}\,\frac{I}{s}\,\frac{2}{\sqrt{1+4s^2/L^2}} \\ &= 3\,\frac{\mu_0}{4\pi}\,I\,\frac{1}{s}\,\sqrt{3} = \frac{\mu_0}{4\pi}\,I\,\frac{6\sqrt{3}}{L}\,\sqrt{3} \\ &= \frac{\mu_0}{4\pi}\,I\,\frac{18}{L} \end{aligned}$$

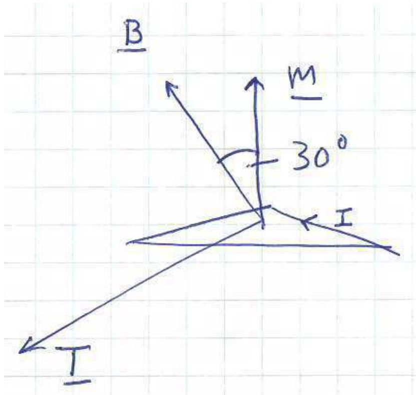
T8.3

$$\text{magnetic moment} = \text{current} \times \text{area}$$

$$\text{area of triangle } = \frac{1}{2} \times L \times \frac{\sqrt{3}}{2} L = \frac{\sqrt{3}}{4} L^2$$

Hence

$$m = \frac{\sqrt{3}}{4} I L^2$$



$$\text{torque, } \mathbf{T} = \mathbf{m} \times \mathbf{B}$$

Hence

$$T = \frac{\sqrt{3}}{4} I L^2 B \sin 30^\circ = \frac{\sqrt{3}}{8} I L^2 B$$

4CCP1501 Tutorial 9

T9.1 An infinitely long straight cylindrical wire of radius a carries a current I. Use Ampère's Law and a suitably chosen amperian loop to calculate the magnetic field $\mathbf{B}$ as a function of distance r from the axis of the wire, both inside and outside the wire, and draw a sketch graph of B as a function of r, for the following three distributions of the currrent over the cross section.

(a) The current is uniformly distributed over the outside surface of the wire.

(b) The current is uniformly distributed over the whole cross section of the wire.

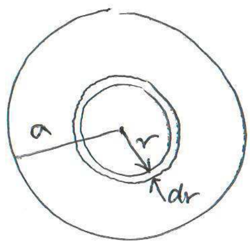
(c) The current is distributed such that the current density J is proportional to r, the distance from the axis of the wire.

In each case sketch the magnetic field from the centre to a point distance from the surface of the rod.

[HINT: in the case (c) we are told that the current *density* J is proportional to r, but we want our answer in terms of the *total current*, I. So write down

$$J = \kappa r$$

say, and first we need to find the constant κ. To do this, ask what is the current carried by an infinitesimal tubular shell of radius r and thickness $\mathrm{d}r$? Looking down the wire we see this



The area of the cross section of the tube is $2\pi r\mathrm{d}r$ so the element of current is

$$\mathrm{d}I = J \times 2\pi r\mathrm{d}r = 2\pi\kappa r^2\mathrm{d}r$$

Then the total current is

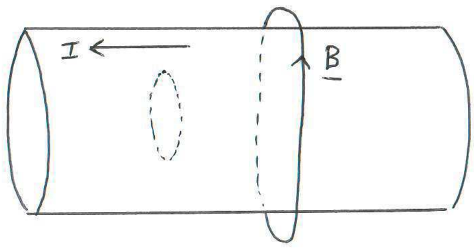
$$I = \int_0^a \mathrm{d}I = 2\pi\kappa \int_0^a r^2\mathrm{d}r$$

When you've done that integral then you can find κ and hence the current density at a distance r from the centre as a function of I and the radius a. (You should get $J = 3Ir/2\pi a^3$.) Then when you want the current flowing through an amperian loop of radius r, say, you need to do the integral again, but this time with the upper limit r instead of a.]

4CCP1501 Solutions to Tutorial 9

T9.1

(a) All the current is carried on the outside surface so there is no current "linking" the amperian loop having radius $r < a$. So there's no magnetic field inside.



The amperian loop of radius $r \geq a$ is linked by the total current I, so

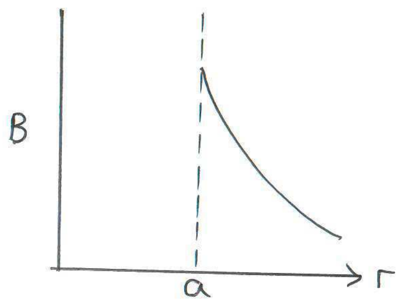
$$\oint \mathbf{B} \cdot \mathrm{d}\boldsymbol{\ell} = \mu_0 I$$

or

$$B \times 2\pi r = \mu_0 I$$

Therefore

$$B = \frac{\mu_0 I}{2\pi r} \quad r \geq a$$



(b) The current density is uniform across the section of the wire and equal to

$$J = \frac{\text{current}}{\text{area}} = \frac{I}{\pi a^2}$$

so the current linked by the amperian loop having radius $r \leq a$ is

$$\begin{aligned} I_{\text{enclosed}} &= J \times \text{area} \\ &= \frac{I}{\pi a^2} \times \pi r^2 = I \frac{r^2}{a^2} \end{aligned}$$

So

$$B \times 2\pi r = \mu_0 I \frac{r^2}{a^2}$$

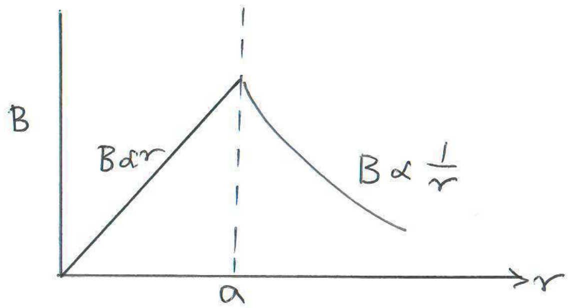
giving us

$$B = \frac{\mu_0 I}{2\pi} \frac{r}{a^2} \qquad r \leq a$$

If $r \geq a$ the current linked is I and the field outside is the same as in (a)

$$B = \frac{\mu_0 I}{2\pi r} \qquad r \geq a$$

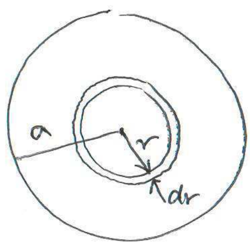
Note that, of course, for $r = a$ the two expressions coincide, as they should; see the sketch:



(c) Here we have to write that J is proportional to r, that is,

$$J = \kappa r$$

say, and first we need to find the constant κ. To do this, ask what is the current carried by an infinitesimal tubular shell of radius r and thickness dr? Looking down the wire we see this



The area of the cross section of the tube is $2\pi r \mathrm{d}r$ so the element of current is

$$\mathrm{d}I = J \times 2\pi r \mathrm{d}r = 2\pi\kappa r^2 \mathrm{d}r$$

Then the total current is

$$I = \int_0^a \mathrm{d}I = 2\pi\kappa \int_0^a r^2 \mathrm{d}r = 2\pi\kappa \frac{1}{3} a^3 \tag{1}$$

Therefore we've found κ:

$$\kappa = \frac{3I}{2\pi a^3}$$

and the current density is

$$J = \frac{3I}{2\pi}\frac{r}{a^3}$$

Now we want the current flowing through an amperian loop of radius $r < a$. This means we do the integral in (1) again but just between the limits 0 and r. So the current enclosed by the amperian loop is

$$\begin{aligned} I_{\text{enclosed}}(r) = 2\pi\kappa \int_0^r r'^2 \mathrm{d}r' &= 2\pi\kappa\frac{1}{3}r^3 \\ &= 2\pi\frac{3I}{2\pi a^3}\frac{1}{3}r^3 \\ &= I\frac{r^3}{a^3} \end{aligned}$$

So by Ampère's law

$$B \times 2\pi = \mu_0 I \frac{r^3}{a^3}$$

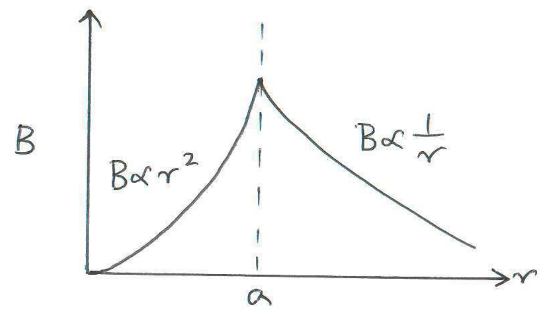
and

$$B = \frac{\mu_0 I}{2\pi}\frac{r^2}{a^3} \qquad r \le a$$

<u>Again</u> as in parts (a) and (b) if $r \ge a$ the current linked is I and

$$B = \frac{\mu_0 I}{2\pi r} \quad r \ge a$$

Note that, of course, for $r = a$ the two expressions coincide, as they should; see the sketch:

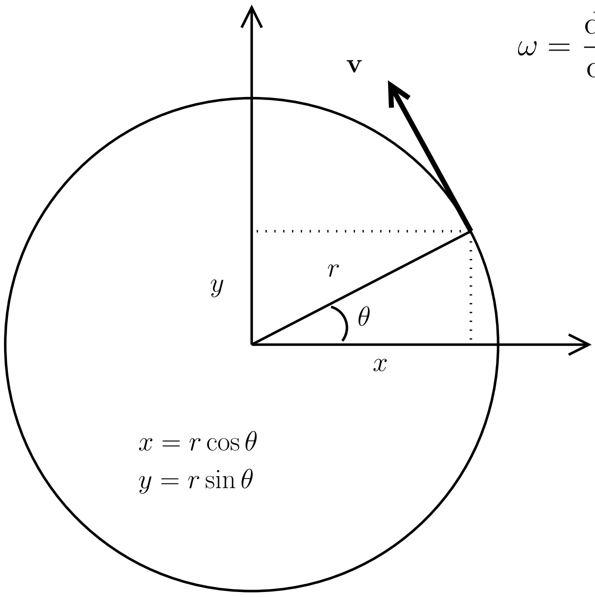


4CCP1501 Tutorial 10

T10.1 An insulating sphere, of radius a, carries a charge Q. The charge is distributed with spherical symmetry such that the charge density is zero at the centre and increases as a positive integer power with the distance from the centre to the surface of the sphere (that is $\rho \propto r^n,\ n \in \mathbb{Z}^+$). Find the electric field inside the sphere as a function the distance from the centre. What is the electric field outside the sphere? Show that the electric field is continuous across the surface of the sphere and make a sketch of the electric field from the centre to a point distant from the surface in the case that the charge density is quadratically increasing with radius.

T10.2 A vortex of water occurs in a whirlpool, or when you drain your bathwater. Suppose that water is flowing around the centre of a vortex with uniform angular velocity

$$\omega = \frac{\mathrm{d}\theta}{\mathrm{d}t}$$



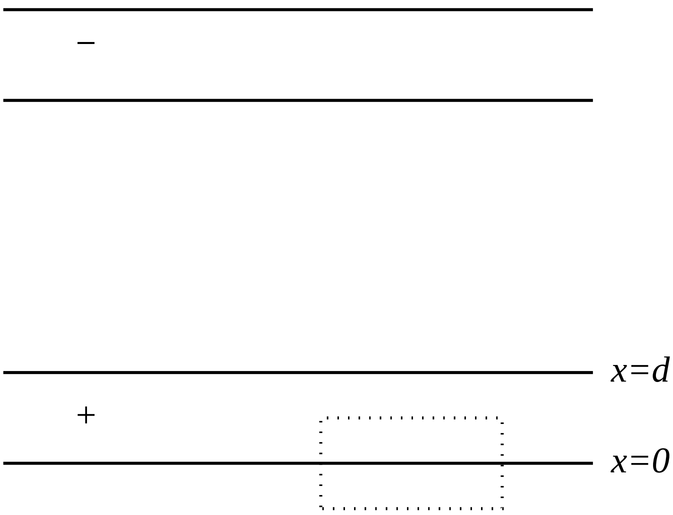
(a) Using the figure, demonstrate that the velocity of the water at the point $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$ is given by the formula

$$\mathbf{v}(\mathbf{r}) = -\omega y\hat{\mathbf{i}} + \omega x\hat{\mathbf{j}}$$

(b) Find the curl, $\boldsymbol{\nabla} \times \mathbf{v}$, of the vector velocity field. In which direction is the curl pointing? What is the divergence of the velocity field, $\boldsymbol{\nabla} \cdot \mathbf{v}$?

OVER...

T10.3 Two parallel sheets of insulator of thickness d are separated by vacuum and carry uniform charge densities $\pm\rho$ C m^{-3}.



Concentrate on the lower, positive, sheet. Using the Gaussian pillbox shown, find an expression for the electric field, $\mathbf{E}$, inside the sheet as a function of x. Note that the electric field is zero below the sheet (that's what the negatively charged sheet is for: to make the problem easier. There is only flux through the top surface of the pillbox, and the charge inside is the volume times ρ). Calculate the divergence, $\boldsymbol{\nabla} \cdot \mathbf{E}$, of the field and show that this agrees with Gauss's Law. Calculate the curl of the field, and confirm that $\boldsymbol{\nabla} \times \mathbf{E} = 0$.

4CCP1501 Solutions to Tutorial 10

T10.1 We start by writing $\rho = kr^n$, and then we need to find the constant k. It's a bit like the problem of last week. First, how are we going to set up the integration? Imagine the infinitesimally thin spherical shell of matter within the sphere, having radius r and thickness dr. Its volume is

$$\mathrm{d}V = \frac{4}{3}\pi\left[(r+\mathrm{d}r)^3 - r^3\right] = 4\pi r^2\mathrm{d}r + \mathcal{O}(\mathrm{d}r^2)$$

after throwing away terms smaller than dr. The amount of charge in this shell is its volume times the charge density, $\mathrm{d}Q = \rho\mathrm{d}V$. Now we can write the total charge in the sphere by summing all these infinitesimal shells from radius zero to radius a. We do this with an integration,

$$\begin{aligned} Q &= \int_0^a \mathrm{d}Q = \int_0^a \rho\mathrm{d}V = 4\pi k\int_0^a r^{n+2}\mathrm{d}r \\ &= 4\pi k\left[\frac{1}{n+3}r^{n+3}\right]_0^a = 4\pi k\frac{1}{n+3}a^{n+3} \end{aligned} \tag{1}$$

Good. We now have

$$k = \frac{1}{4\pi}\,Q\,\frac{n+3}{a^{n+3}} \tag{2}$$

To find the electric field at a distance r from the centre of the sphere we use Gauss's law: flux is charge enclosed divided by ϵ_0. The enclosed charge is found by doing the integral again, now between zero and r. So by comparison with (1) and substituting (2)

$$Q_{\text{enclosed}} = 4\pi k\frac{1}{n+3}r^{n+3} = \frac{Q}{a^{n+3}}\,r^{n+3}$$

(note the 4π and $n+3$ cancel out). The area of the the gaussian surface is $4\pi r^2$ and so Gauss's law reads

$$E\times 4\pi r^2 = \frac{1}{\epsilon_0}\,Q_{\text{enclosed}} = \frac{1}{\epsilon_0}\,\frac{Q}{a^{n+3}}\,r^{n+3}$$

The final result is

$$E = \frac{1}{4\pi\epsilon_0}\,\frac{Q}{a^{n+3}}r^{n+1}$$

Note a few things. (i) The field increases *inside* the sphere as r^{n+1}. To get things dimensionally right, there's a a^{n+3} in the denominator; this is a useful check that you're on the right track. (ii) If you can, always do the most general case. Then if you're asked about a problem in which the charge density goes like, say, the *cube* of the distance, then you've already done it. (iii) Now go back to the last problem, T8.1 from last week, and do the general case $J = \kappa r^n$.

If $r > a$ then the electric field is the same as that of a point charge of amount Q at the origin and you can check that at $r = a$ the field is

$$E(a) = \frac{1}{4\pi\epsilon_0}\,\frac{Q}{a^2}$$

as you'd hope it would be.

T10.2 Angular velocity is

$$\omega = \frac{\mathrm{d}\theta}{\mathrm{d}t} \quad [\text{radian s}^{-1}]$$

We have

$$\text{distance travelled in one period} = 2\pi r$$
$$\text{time taken over one period} = \frac{2\pi}{\omega}$$

and therefore

$$\text{speed} = \text{distance} \div \text{time} = \omega r$$

Then

$$\begin{aligned}\text{velocity, } \mathbf{v}(\mathbf{r}) &= \text{speed} \times \text{unit vector } \hat{\mathbf{v}} \\ &= \omega r \hat{\mathbf{v}} \\ &= \omega r(-\sin\theta \hat{\mathbf{i}} + \cos\theta \hat{\mathbf{j}}) \\ &= -\omega y \hat{\mathbf{i}} + \omega x \hat{\mathbf{j}}\end{aligned}$$

You can show it another way,

$$\begin{aligned}\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} x\hat{\mathbf{i}} + \frac{\mathrm{d}}{\mathrm{d}t} y\hat{\mathbf{j}} &= \frac{\mathrm{d}x}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}t}\hat{\mathbf{i}} + \frac{\mathrm{d}y}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}t}\hat{\mathbf{j}} \\ &= -r\sin\theta\frac{\mathrm{d}\theta}{\mathrm{d}t}\hat{\mathbf{i}} + r\cos\theta\frac{\mathrm{d}\theta}{\mathrm{d}t}\hat{\mathbf{j}} \\ &= -y\omega\hat{\mathbf{i}} + x\omega\hat{\mathbf{j}}\end{aligned}$$

Or yet another way by calculating $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$ ...

Now you have the velocity vector field you can find its curl and divergence. Instinctively I hope you can see that the field has circulation, that is, non zero curl; but zero divergence—there is no *source* of water anywhere. To prove it just do

$$\begin{aligned}\boldsymbol{\nabla} \times \mathbf{v}(\mathbf{r}) &= \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ -\omega y & \omega x & 0 \end{vmatrix} \\ &= 2\omega\hat{\mathbf{k}}\end{aligned}$$

so the curl vector points in the z-direction—perpendicular to the plane of the vortex—and has a magnitude of twice the angular velocity. To show that the divergence vanishes,

$$\begin{aligned}\boldsymbol{\nabla} \cdot \mathbf{v}(\mathbf{r}) &= \frac{\partial}{\partial x}(-\omega y) + \frac{\partial}{\partial y}(\omega x) \\ &= 0\end{aligned}$$

T10.3 Area of pillbox is A. The only flux is through the top surface. So,

$$E(x) \times A = \frac{1}{\epsilon_0} Q_{\text{enclosed}}$$

The volume of the pillbox lying inside the sheet is Ax, therefore

$$Q_{\text{enclosed}} = \rho A x$$

and so

$$E(x) = \frac{1}{\epsilon_0} \rho x$$

and if we want to include the direction as well as magnitude of the field, we write

$$\mathbf{E} = \frac{1}{\epsilon_0} \rho x \, \hat{\mathbf{i}}$$

I can take the divergence of both sides of this,

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{E} &= \left(\hat{\mathbf{i}} \frac{\partial}{\partial x} + \hat{\mathbf{j}} \frac{\partial}{\partial y} + \hat{\mathbf{k}} \frac{\partial}{\partial z} \right) \cdot \frac{1}{\epsilon_0} \rho x \, \hat{\mathbf{i}} \\ &= \frac{1}{\epsilon_0} \rho \, \hat{\mathbf{i}} \cdot \hat{\mathbf{i}} = \frac{1}{\epsilon_0} \rho \end{aligned}$$

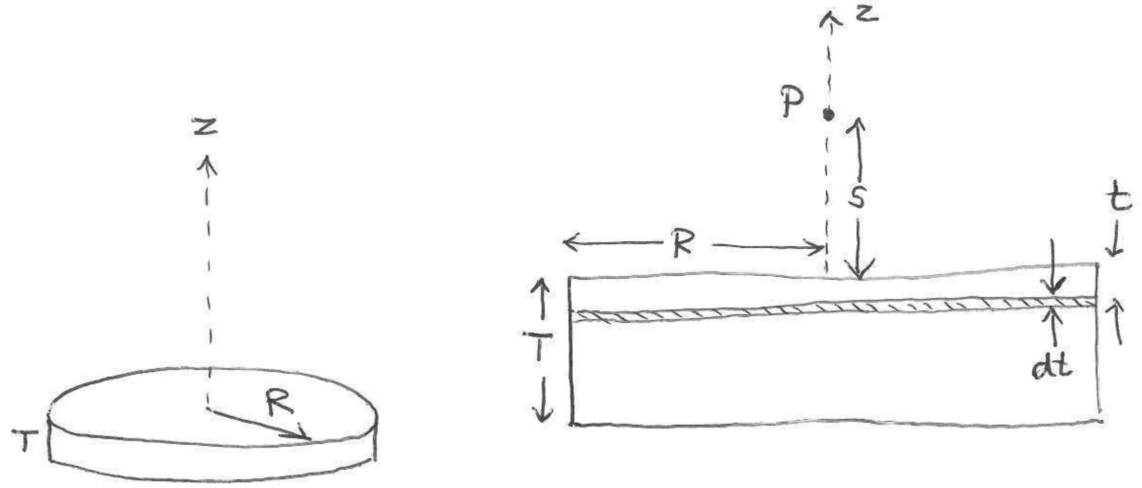
which is, indeed, Gauss's Law. To show that the circulation, or curl, of the electric field is zero, write

$$\boldsymbol{\nabla} \times \mathbf{E} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \frac{1}{\epsilon_0} \rho x & 0 & 0 \end{vmatrix} = 0$$

4CCP1501 Worked Problem

Find the electric field along the central axis outside a uniformly charged disc.

This has the shape of a pound coin, or hockey puck. The thickness is T; the radius is R and the total charge is Q. This is like the problem that we solved in Lecture 9, but now instead of a two dimensional circular sheet, which has no thickness, the disc has a thickness T. But the strategy is the same: we set up a point, P, which is a distance s above the surface of the disc, and along the central axis, and ask what is the electric field due to an infinitesimal sheet, of thickness $\mathrm{d}t$ within the disc that is a distance t below the top surface. We know the solution to that problem from page 2 of Lecture 9; and now we just have to integrate, that is, to sum up all the infinitesimal sheets of charge from $t = 0$ to $t = T$.



Shape of the puck and illustration of the integration.

The infinitesimal sheet is a distance t below the surface so it is a distance $s + t$ from point P. So, the increment of electric field at point P due to the sheet is, using the equation from Lecture 9 and substituting $s + t$ for s,

$$\mathrm{d}E = \frac{1}{4\pi\epsilon_0}\; 2\pi\, \mathrm{d}\sigma\, (s+t) \left(\frac{1}{s+t} - \frac{1}{\sqrt{R^2 + (s+t)^2}} \right) \tag{1}$$

and we only need the magnitude of the vector because we know it is pointing along the $\hat{\mathbf{k}}$-direction. I have written that the charge density on the infinitesimal sheet is $\mathrm{d}\sigma$ [C m^{-2}] and I can find what this is in terms of the total charge Q as follows.

The charge density is

$$\rho = \frac{\text{charge}}{\text{volume of puck}} = \frac{Q}{\pi R^2 T} \tag{2}$$

the volume of the infinitesimal sheet is

$$\mathrm{d}V = \pi R^2 \mathrm{d}t$$

and the charge in the infinitesimal sheet is

$$\mathrm{d}q = \rho\,\mathrm{d}V = \frac{Q}{\pi R^2 T}\,\pi R^2 \mathrm{d}t = \frac{Q\,\mathrm{d}t}{T}$$

So the infinitesimal charge density of the sheet is

$$\mathrm{d}\sigma = \frac{\text{charge}}{\text{area}} = \frac{\mathrm{d}q}{\pi R^2} = \frac{Q\,\mathrm{d}t}{T}\,\frac{1}{\pi R^2} = \rho\,\mathrm{d}t$$

where I have used equation (2). Now I can rewrite equation (1) in terms of the volume charge density, rather than the surface charge density,

$$\begin{aligned}\mathrm{d}E &= \frac{1}{4\pi\epsilon_0}\,2\pi\,\rho\,\mathrm{d}t\,(s+t)\left(\frac{1}{s+t} - \frac{1}{\sqrt{R^2+(s+t)^2}}\right)\\ &= \frac{1}{4\pi\epsilon_0}\,2\pi\,\rho\,\mathrm{d}t\,\left(1 - \frac{(s+t)}{\sqrt{R^2+(s+t)^2}}\right)\end{aligned}$$

Now I will integrate this from $t = 0$ to $t = T$ and note that there are two terms. The first I can do in my head since

$$\int_0^T \mathrm{d}t = T$$

and so the total electric field at point P is

$$E = \int_0^T \mathrm{d}E = \frac{1}{4\pi\epsilon_0}\,2\pi\rho T - \frac{1}{4\pi\epsilon_0}\,2\pi\rho\int_0^T \frac{(s+t)}{\sqrt{R^2+(s+t)^2}}\mathrm{d}t \tag{3}$$

You can at once inspect the limit as $R \to \infty$. What do you expect? Well in this limit you have an infinite disc of thickness T and volume charge density ρ so the charge density per unit area is $\sigma = \rho T$. And as $R \to \infty$ the integrand becomes zero and you are left with

$$E(R \to \infty) = \frac{1}{2}\frac{\rho T}{\epsilon_0} = \frac{\sigma}{2\epsilon_0} \tag{4}$$

You can easily confirm this using Gauss's law.

What we now have in equation (3) is an expression for E which is the value it would have in the limit of infinite radius plus a correction term that we will call E_{corr}, due the finite size of the disc. By comparison with equation (3)

$$E_{\mathrm{corr}} = -\frac{1}{4\pi\epsilon_0}\,2\pi\rho\int_0^T \frac{(s+t)}{\sqrt{R^2+(s+t)^2}}\mathrm{d}t$$

We can do the integral by substitution, $u = s + t$, $\mathrm{d}t = \mathrm{d}u$,

$$\int \frac{s+t}{\sqrt{R^2+(s+t)^2}}\mathrm{d}t = \sqrt{R^2+(s+t)^2} + \text{constant}$$

and this gives us the correction term due to the finite radius of the disc,

$$E_{\text{corr}} = \frac{1}{4\pi\epsilon_0}\, 2\pi\rho \left[\sqrt{R^2+s^2} - \sqrt{R^2+(s+T)^2}\right]$$

Adding this to the first term in equation (3) gives me my final answer,

$$E = \frac{1}{4\pi\epsilon_0}\, 2\pi\rho \left[T + \sqrt{R^2+s^2} - \sqrt{R^2+(s+T)^2}\right] \tag{5}$$

If s becomes very large compared to R and T, then from a long distance the field must look like that of a point charge $Q = \pi R^2 T\rho$. If I expand the term in brackets in equation (5) around $s = \infty$ I get

$$T + \sqrt{R^2+s^2} - \sqrt{R^2+(s+T)^2} = \frac{R^2 T}{2s^2} + \mathcal{O}(s^{-3})$$

and so

$$E(s\to\infty) = \frac{1}{4\pi\epsilon_0}\, 2\pi\rho \frac{R^2T}{2s^2} = \frac{1}{4\pi\epsilon_0}\,\frac{Q}{s^2} \tag{6}$$

as expected.

If the disc is *thin*, that is, $T \ll R$; and if also $s \ll R$ so that my point P is close to the surface of the disc,

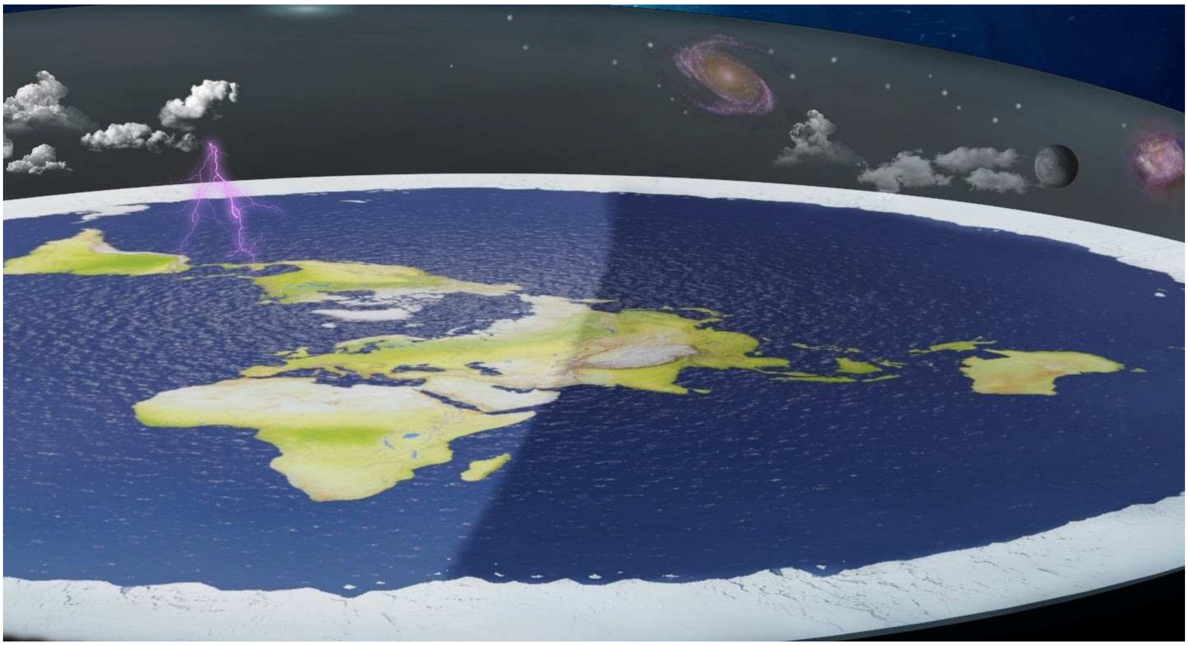
$$\begin{aligned}
\sqrt{R^2+s^2} - \sqrt{R^2+(s+T)^2} &= R\sqrt{1+\left(\frac{s}{R}\right)^2} - R\sqrt{1+\left(\frac{s+T}{R}\right)^2} \\
&\approx R\left[1+\frac{1}{2}\left(\frac{s}{R}\right)^2\right] - R\left[1+\frac{1}{2}\left(\frac{s+T}{R}\right)^2\right] \\
&= -\frac{T}{2R}\,(2s+T)
\end{aligned}$$

So putting this in place of the two square roots in (5), to first order in s the electric field is

$$\begin{aligned}
E &= \frac{1}{4\pi\epsilon_0}\, 2\pi\rho T \left(1 - \frac{2s+T}{2R}\right) && (7) \\
&= \frac{1}{4\pi\epsilon_0}\;\, 2\pi\sigma \left(1 - \frac{2s+T}{2R}\right) && (8)
\end{aligned}$$

This makes sense, because in the limit that the radius is infinite I recover from equation (8) the well known solution, equation (4), for the infinite sheet of charge carrying a charge density σ [C m^{-2}] having the well known fact that the field is *independent* of the height s. The second term which corrects for a finite radius is correct to first order in s, and has just a *linear* dependence on the height, s, above the surface. In fact it's not at all obvious from equation (5) that in the limit of very large s the electric field falls off like the *square* of the distance, as it must do in accordance with Coulomb's law. This is however demonstrated by equation (6).

Logo of the Flat Earth Society (`www.tfes.org`*)*



Artist's impression of the flat-earth.

Some people believe that the Earth is flat (`en.wikipedia.org/wiki/Modern_flat_Earth_societies`). That is, that it has exactly this shape of a hockey puck; that the north pole is at the centre and that the antarctic ice is distributed around the rim. I got to wondering whether we could test this hypothesis by measuring the acceleration due to gravity, g, at different distances, s, above the surface of the Earth. After all, if the Earth is spherical then $g \propto 1/s^2$ whereas if we could assume that the disc had a very large radius, then like the infinite sheet of charge g would be the same at any height above the earth and this would be an easy way to convince the flat-earthers that they are wrong. Actually a friend pointed out to me that this is a false logic because the same argument (assuming an infinite radius) applied to the Earth would also result in

g being independent of the height. Let's start by seeing how this arises from Coulomb's law. The analogy can be made with the electrostatic case, very simply by replacing charge in Coulomb with mass in kg and making the substitution

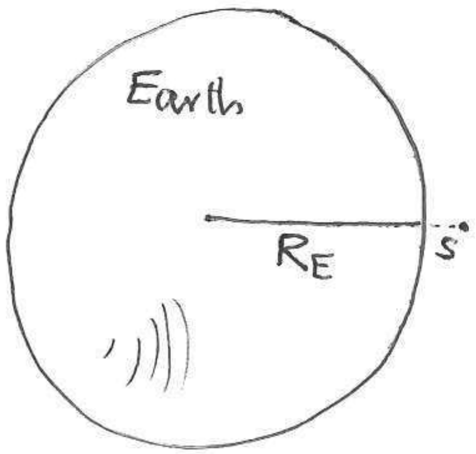
$$\frac{1}{4\pi\epsilon_0} \longrightarrow G$$

where $G = 6.674 \times 10^{-11}$ m^3 kg^{-1} s^{-2} is the universal gravitational constant. You all know from Gauss's law that the electric field at a point outside a uniformly charged sphere of radius R_E is

$$E = \frac{1}{4\pi\epsilon_0}\,\frac{Q}{(R_\mathrm{E}+s)^2} = \frac{1}{4\pi\epsilon_0}\,\frac{4}{3}\pi R_\mathrm{E}^3\,\rho\,\frac{1}{(R_\mathrm{E}+s)^2}$$

where Q is the total charge and ρ is the charge density. The distance from the surface to the field point is s. So I can use this result to write down the acceleration due to gravity, or *force per unit mass*, g_E, (*cf.* force per unit charge, E) at a height s above the surface of a perfectly spherical Earth with uniform mass density ρ. (From now on, ρ will always mean mass density, whereas up to now I have used ρ for charge density.) So,

$$g_\mathrm{E} = \frac{4}{3}\pi\,GR_\mathrm{E}^3\,\rho\,\frac{1}{(R_\mathrm{E}+s)^2} \tag{9}$$



Illustrating equation (9)

Just as I did in equation (7) I want to write this as a term that is independent of the height plus a correction term for the finite curvature (radius) of the Earth. So I expand

$$(R_\mathrm{E}+s)^{-2} = R_\mathrm{E}^{-2}\left(1+\frac{s}{R_\mathrm{E}}\right)^{-2} = R_\mathrm{E}^{-2}\left(1-2\frac{s}{R_\mathrm{E}}\cdots\right)$$

and then to first order in s/R_E

$$g_\mathrm{E}^{(1)} = \frac{4}{3}\pi\,G\,\rho\,R_\mathrm{E}\left(1-\frac{2s}{R_\mathrm{E}}\right) \tag{10}$$

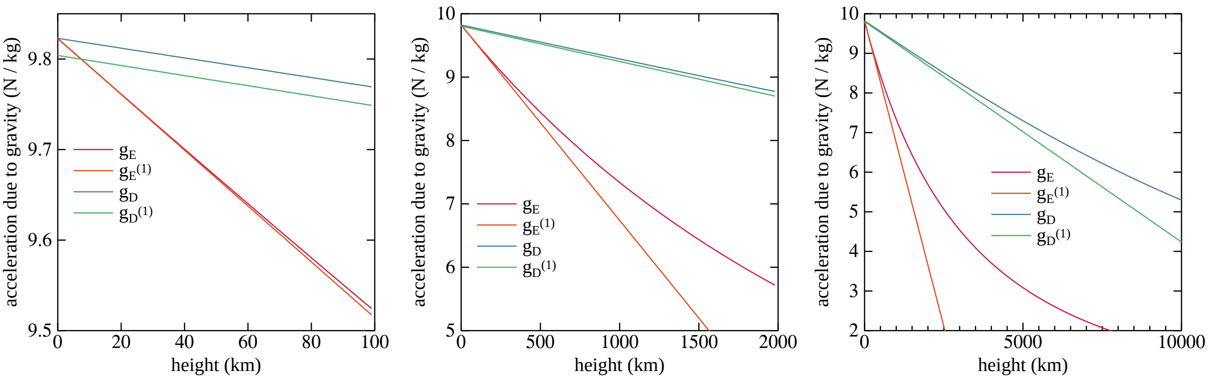
Of course equations (9) and (10) are the same at the surface of the Earth where $s = 0$. In the case of the disc, I can do the same and define by comparison with equation (5), the acceleration due to gravity a distance s above the "north pole" of the flat earth having a distance R between the "north pole" and the "antarctic rim" and the depth, or thickness, T,

$$g_{\mathrm{D}} = G\,2\pi\rho\left[T + \sqrt{R^2+s^2} - \sqrt{R^2+(s+T)^2}\right] \tag{11}$$

And to first order, from equation (7) I can write

$$g_{\mathrm{D}}^{(1)} = 2\pi G\,\rho\,T\left(1 - \frac{2s+T}{2R}\right) \tag{12}$$

If you compare equations (10) and (12) it becomes much less obvious that I can tell whether I'm on a large flat plate or the surface of a large sphere because to zero order in either case g is independent of the height above the surface and the first order correction is in both cases linear in the height.



Acceleration due to gravity at a height s km above the Earth and above the flat disc. Both have the same density $\rho = 5\,515$ kg m^{-3}; the Earth's radius is $R_{\mathrm{E}} = 6\,371$ km; the radius and thickness of the disc are $R = 20\,015$ km and $T = 4\,820$ km. Exact results are g_{E}, equation (9) and g_{D}, equation (11); linear approximations are $g_{\mathrm{E}}^{(1)}$, equation (10) and $g_{\mathrm{D}}^{(1)}$, equation (12). Because I was constrained to have the flat-earth's radius equal to half the circumference of the Earth and for us to experience the same force of gravity at the surface, the Earth and flat-earth masses are not the same: they are 6×10^{24} kg and 7×10^{23} kg respectively.

Now the astronomers and Earth scientists have measured the radius and mass density of the Earth and we believe we know that $R_{\mathrm{E}} = 6\,371$ km and $\rho = 5\,515$ kg m^{-3}. We also feel that the acceleration due to gravity is 9.8 m s^{-2} (N kg^{-1}). On the other hand the flat-earthers must hold that the radius of the disc is the distance between the north and south poles, which they will have measured by walking and boating till it became too cold to continue, and so $R = \pi R_{\mathrm{E}} = 20\,015$ km. They will also have to accept that the density is $\rho = 5\,515$ kg m^{-3} since they take no issue with the Earth scientists about the density of rock and so on. Indeed they feel the same gravitational pull as we do, and using equation (11) it turns out that the thickness of the disc or "depth" of the flat earth is $T = 4\,820$ km. As my friend pointed out it is not a good approximation to neglect the finite radius. If I use just the first term in equation (12) which is equivalent to using Gauss's law as in equation (4) then I find that the acceleration due to gravity is 11.15 N kg^{-1}. Because of this I may not assert that g_{D} is independent of the height above the surface of the flat earth and so I cannot so easily confound the flat-earthers.

All the same it is instructive to compare the two cases. Here I have plotted the acceleration due to gravity for the Earth and the flat-earth in both the exact formula and to first order.

In the left hand figure which looks at low heights from zero to 100 km it is clear that the height dependence is much stronger for the Earth than for the flat-earth and this may convince the flat-earthers. In both cases the slope is very close to linear and for the Earth the linear approximation is close to exact. The approximation $g_{\mathrm{D}}^{(1)}$ has the right slope, but there is a constant error of 0.02 N kg^{-1} which arises because T/R is not that small for our geometry. In the middle figure you can see more clearly that the gravitional field is falling off like s^{-2} as it should, equation (9); this is still not evident for the flat-earth: according to the right hand figure you have to get to nearly 10 000 km away (which is more than twice its thickness and half its radius) before it starts to look like a "point mass". This second difference between the two cases may convince the flat-earther but it would be hard to get that far away to make the measurements.

Mind you, in a post-truth argument worthy of the United States President, I read that some flat-earthers don't anyway accept the existence of gravity. Their claim is that the reason we feel it is that the flat-earth is accelerating in a direction normal to the disc at exactly 9.8 m s^{-2}. By now we must be going pretty fast; so don't tell them about special relativity...

4CCP1501 Simple Harmonic Motion

1. Free oscillation

We are to solve Newton's second law, force = mass × acceleration, as a differential equation,

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx$$

which we write

$$\ddot{x} = -\omega_0^2 x \tag{1}$$

by using two dots to indicate a *second derivative with respect to time.* We will use one dot to indicate the first derivative. We have also combined the two constants, m, the mass and k, the spring constant, to define an *angular frequency*,

$$\omega_0^2 = \frac{k}{m}$$

We're not mathematicians, we just want a solution of this thing; so try $x = Ae^{st}$. Then by simple differentiating, we have

$$x = Ae^{st}\ ;\quad \dot{x} = sAe^{st}\ ;\quad \ddot{x} = s^2Ae^{st}$$

We only have to put this back into (1) to see that

$$s^2Ae^{st} + \omega_0^2Ae^{st} = 0 \longrightarrow s^2 + \omega_0^2 = 0 \longrightarrow s = \pm\mathrm{i}\omega_0$$

So we have two solutions:

$$x = Ae^{\mathrm{i}\omega_0\mathrm{t}} \text{ and } x = Ae^{-\mathrm{i}\omega_0\mathrm{t}}$$

The theory of second order, linear differential equations tells us that the most general solution is a linear combination of the two solutions with two arbitrary coefficients, that we will call A_1 and A_2:

$$\begin{aligned}
x &= A_1e^{\mathrm{i}\omega_0\mathrm{t}} + A_2e^{-\mathrm{i}\omega_0\mathrm{t}} \\
&= (A_1 + A_2)\cos\omega_0 t + \mathrm{i}(\mathrm{A}_1 - \mathrm{A}_2)\sin\omega_0\mathrm{t} \\
&= A\cos\omega_0 t + B\sin\omega_0 t && \text{(a)} \\
&= C\cos\phi\sin\omega_0 t + C\sin\phi\cos\omega_0 t && \text{(b)} \\
&= C\sin(\omega_0 t + \phi)
\end{aligned}$$

In going from line (a) to line (b) I have changed from the variables A and B to variables C and ϕ by making these two definitions,

$$A = C\sin\phi \text{ and } B = C\cos\phi$$

because then I can use the usual formula for $\sin(a + b)$ to arrive at the last line.

Now what we have is

$$x = C\sin(\omega_0 t + \phi)$$
$$\dot{x} = v = C\omega_0 \cos(\omega_0 t + \phi)$$

To fix the, up to now arbitrary, constants requires us to know "boundary conditions." Let's suppose that at $t = 0$, $x = x_0$, say, and $v = v_0$, the initial velocity. These conditions give,

$$x_0 = C\sin\phi\ , \quad \sin\phi = \frac{x_0}{C} \tag{c}$$
$$v_0 = C\omega_0\cos\phi\ , \quad \cos\phi = \frac{v_0}{C\omega_0} \tag{d}$$

Now, square and add (c) and (d),

$$C = \sqrt{x_0^2 + \frac{v_0^2}{\omega_0^2}}$$

and divide (c) by (d)

$$\phi = \arctan\frac{x_0\omega_0}{v_0}$$

Finally, if we start off the oscillator at $t = 0$ with $v_0 = 0$ and $x = x_m$, for example we pull out the spring to maximum deflection, x_m, hold it still ($v_0 = 0$) and let it go; then the solution is

$$x = x_m\sin(\omega_0 t + \frac{1}{2}\pi) = x_m\cos(\omega_0 t)$$

2. Damping

To the differential equation (1), which is after all Newton's second law—force equals mass times acceleration—we add an additional force, $-b\dot{x}$. This force is *proportional to the velocity*, which is what you'd expect. Try swimming in syrup: the faster you swim the bigger is the drag, or viscous, force. So now we need to solve the differential equation

$$m\ddot{x} + b\dot{x} + kx = 0$$

which we re-write as

$$\ddot{x} + \frac{b}{m}\dot{x} + \omega_0^2 x = 0$$

We define a new constant, Z, such that

$$\frac{b}{m} = 2Z\omega_0$$

is the *frictional force per unit mass and unit speed.* Now our differential equation is

$$\ddot{x} + 2Z\omega_0\dot{x} + \omega_0^2 x = 0$$

As before we try

$$x = Ae^{st}\ ; \quad \dot{x} = sAe^{st}\ ; \quad \ddot{x} = s^2Ae^{st}$$

and so

$$s^2 + 2Z\omega_0 s + \omega_0^2 = 0$$

leads to

$$s = \omega_0 \left(-Z \pm \sqrt{Z^2 - 1}\right) \tag{2}$$

and the general solution must be

$$x = A_1 e^{st} + A_2 e^{-st} \tag{3}$$

Critial damping is defined as the condition $Z = 1$. For that case we define

$$b_{\text{crit}} = 2m\omega_0 = 2\sqrt{mk}$$

and we give a name to Z by

$$\frac{b}{b_{\text{crit}}} = Z$$

being called the *damping factor*, or *damping ratio.*

Underdamping is the condition $Z < 1$ or $b < b_{\text{crit}}$. This is usually the most interesting case, and for which

$$Z^2 - 1 < 0$$

meaning that there are two roots to (2), namely,

$$s_1 = \omega_0 \left(-Z + \mathrm{i}\sqrt{1 - \mathrm{Z}^2}\right)$$
$$s_1 = \omega_0 \left(-Z - \mathrm{i}\sqrt{1 - \mathrm{Z}^2}\right)$$

and then (3) is

$$x = e^{-Z\omega_0 t} \left(A_1 e^{\mathrm{i}\sqrt{1-\mathrm{Z}^2}\omega_0 \mathrm{t}} + A_2 e^{-\mathrm{i}\sqrt{1-\mathrm{Z}^2}\omega_0 \mathrm{t}}\right)$$

We then simplify this in the same manner as for equations (a) and (b):

$$\begin{aligned} x &= Ce^{-Z\omega_0 t} \sin\left(\sqrt{1 - Z^2}\omega_0 t + \phi\right) \\ &= Ce^{-\alpha t} \sin(\omega_D t + \phi) \end{aligned}$$

where

$$\alpha = \frac{1}{2}\frac{b}{m} = Z\omega_0$$

is called the *damping constant*, and

$$\omega_D = \omega_0\sqrt{1 - Z^2} = \omega_0\sqrt{1 - \frac{1}{4}\frac{b^2}{mk}} < \omega_0$$

is the *damped frequency.*

Again, if at $t = 0$, $x = x_m$ and $v = 0$, the solution associated with these boundary conditions is

$$\begin{aligned} x &= x_m\, e^{-\alpha t} \sin\left(\omega_D t + \frac{1}{2}\pi\right) \\ &= x_m\, e^{-\alpha t} \cos\omega_D t \end{aligned}$$

which is the result I give you on page 6 of Lecture 3.

3. Driven oscillators

In real life we are less interested in an oscillator that is oscillating at its natural frequency, ω_0, or its natural damped frequency, ω_D, than in the behaviour of an undamped or damped oscillator when we choose to drive it at some frequency, ω, that *we* choose. Situations of this phenomenon are ubiquitous in physics and engineering. Try and write down some half a dozen examples of your own.

3.1 Undamped driven oscillator

The oscillator is driven by a periodic force of angular frequency ω and amplitude F_0. That means we have one more force to add in to Newton's second law, namely

$$F = F_0 \sin\omega t$$

and force = mass × acceleration now reads

$$m\ddot{x} = F_0 \sin\omega t - kx \tag{4}$$

Eventually the oscillator has no choice but to vibrate at the frequency of the driving force, whether it likes it or not, so we must have,

$$\begin{aligned} x &= A\sin\omega t \\ \dot{x} &= A\omega\cos\omega t \\ \ddot{x} &= -A\omega^2 \sin\omega t \end{aligned}$$

Equation (4) now reads

$$-mA\omega^2 \sin\omega t + kA\sin\omega t = F_0 \sin\omega t$$

That is,

$$\begin{aligned} A &= \frac{F_0}{k - m\omega^2} = \frac{F_0/k}{1 - \frac{\omega^2}{\omega_0^2}} \\ &= \frac{A_s}{1 - \frac{\omega^2}{\omega_0^2}} \end{aligned}$$

using

$$\omega_0 = \sqrt{\frac{k}{m}}$$

the natural frequency of the undamped oscillator. We call A_s the *static amplitude* and we call A the *dynamic amplitude*; their ratio is called the *magnification factor*,

$$D_s = \frac{A}{A_s} = \left(1 - \frac{\omega^2}{\omega_0^2}\right)^{-1}$$

If the driving frequency is less that the natural frequency the magnification factor is positive and the displacement is in phase with the driving force. Conversely if $\omega > \omega_0$, $D_s < 0$. An amplitude cannot be negative, so we'll have instead, for this case, to use the solution

$$x = -A \sin \omega t$$

which implies a phase difference of π (180°) between the displacement and the driving force. Thirdly, if $\omega = \omega_0$, $D_s \to \infty$ and we have *resonance*. In real life this never happens as there is always damping. But interesting things *do* happen when we drive an oscillator at a frequency close to its natural one.

3.1 Damped driven oscillator

Now we include the velocity dependent damping force into equation (4):

$$m\ddot{x} = F_0 \sin \omega t - b\dot{x} - kx$$

or

$$m\ddot{x} + b\dot{x} + kx = F_0 \sin \omega t \tag{4a}$$

Eventually after transients have died away, the oscillator must vibrate at the frequency of the driving force. It may not like it and it will protest unless the driving frequency is close to the natural frequency of the undriven oscillator. Its reluctance to cooperate is reflected in a reduction in amplitude. Nearer to *resonance* the amplitude is large. The so called *resonance curve* or relation between amplitude and driving frequency is what we will be seeking in the mathematical development that follows. The oscillator will necessarily vibrate at the frequency of the driving force, but it will not necessarily be in phase with it. Hence the solution for the amplitude must look like

$$\begin{aligned} x &= A \sin(\omega t - \phi) \\ \dot{x} &= A\omega \cos(\omega t - \phi) \\ \ddot{x} &= -A\omega^2 \sin(\omega t - \phi) \end{aligned}$$

when I plug these into (4a) I get

$$\begin{aligned} m\left[-A\omega^2 \sin(\omega t - \phi)\right] + b\left[A\omega \cos(\omega t - \phi)\right] + kA \sin(\omega t - \phi) &= F_0 \sin \omega t \\ &= F_0 \sin(\omega t - \phi + \phi) \end{aligned}$$

Rearranging this I have

$$\begin{aligned} &A\left(k - m\omega^2\right) \sin(\omega t - \phi) + Ab\omega \cos(\omega t - \phi) \\ &\qquad = F_0 \left[\sin(\omega t - \phi) \cos \phi + \cos(\omega t - \phi) \sin \phi\right] \end{aligned}$$

Now, equate the coefficients of $\sin(\omega t - \phi)$ and $\cos(\omega t - \phi)$ and obtain

$$A b\omega = F_0 \sin\phi$$
$$A\left(k - m\omega^2\right) = F_0 \cos\phi$$

We square and add these two, recalling that $\sin^2\phi + \cos^2\phi = 1$,

$$F_0^2 = A^2\left[\left(k - \omega^2\right) + b^2\omega^2\right]$$

which means that we have, for the *dynamic amplitude*,

$$\begin{aligned} A &= \frac{F_0}{\sqrt{\left(k - m\omega^2\right)^2 + b^2\omega^2}} \\ &= \frac{F_0/k}{\sqrt{\left(1 - \frac{m\omega^2}{k}\right)^2 + \frac{b^2\omega^2}{k^2}}} \end{aligned}$$

We also divide our two equations to find the *phase difference*, or phase angle, ϕ, between the oscillator and its driving force,

$$\tan\phi = \frac{b\omega}{k - m\omega^2}$$

We can simplify the formulas for A and ϕ using these definitions that we have encountered already in these notes,

$$\omega_0 = \sqrt{\frac{k}{m}}\ ,\quad b = 2mZ\omega_0\ ,\quad A_s = \frac{F_0}{k}$$

We also define the *frequency ratio*,

$$r = \frac{\omega}{\omega_0}$$

Then the magnification factor is

$$D_s = \frac{A}{A_s} = \frac{1}{\sqrt{\left(1 - r^2\right)^2 + \left(2rZ\right)^2}} \tag{5}$$

and the phase angle is

$$\phi = \arctan\frac{2rZ}{1 - r^2} \tag{6}$$

What is the frequency, $\omega_{\max}$, say, that gives us the greatest amplitude? Or to put the question another way, what is the *resonant frequency*? We need to minimise the denominator in (5); we do this in the usual way by setting its first derivative with respect to r equal to zero and solving for r which will then give us $\omega_{\max}/\omega_0$.

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[\left(1 - r^2\right)^2 + \left(2rZ\right)^2\right] = 0$$

leads to

$$\omega_{\max} = \omega_0\sqrt{1-2Z^2} \tag{7}$$

which is neither ω_0, nor $\omega_D = \omega_0\sqrt{1-Z^2}$.

What is the *maximum ampltitude*; $A_{\max}$, say? Put (7) into (5) and neglect Z^4 when compared to Z^2. We find

$$\frac{A_{\max}}{A_s} = \frac{1}{2Z} = \frac{m\omega_0}{b} \approx Q$$

which is the "quality factor", and using $A_s = F_0/k$ and $\omega_0^2 = k/m$ we get

$$A_{\max} = \frac{F_0}{b\omega_0}$$

On page 9 (below) are two graphs I've taken from `wikipedia` showing a set of resonance curves and phase angles for a driven damped oscillator. On the abscissa is plotted the frequency ratio, r. They use the phrase "amplification ratio" for the magnification factor and have used the symbol ζ for the damping factor, Z. The first is essentially a plot of equation (5). In the second, note how in the case of the undamped forced oscillator there is an abrupt change from in phase to 180° out of phase as r goes through one, as we discuss on page 5 of these notes. Note how the frequency $\omega_{\max}$ is always smaller than the natural frequency ω_0 but appears to approach it as the peak becomes narrower, that is, the damping becomes less.

There are three interesting cases.

(i) If $r \ll 1$ the driving frequency is much smaller than the natural frequency of the oscillator,

$$\omega \ll \omega_0$$

Then the dynamic amplitude is close to the static amplitude,

$$A \approx A_s$$

and the phase difference is

$$\phi \approx \arctan 0 = 0$$

so the displacement and force are in phase.

(ii) If $r \approx 1$ then

$$\omega \approx \omega_0$$

and

$$\frac{A}{A_s} \approx \frac{1}{2Z} \approx Q \text{ , the quality factor}$$

Also,

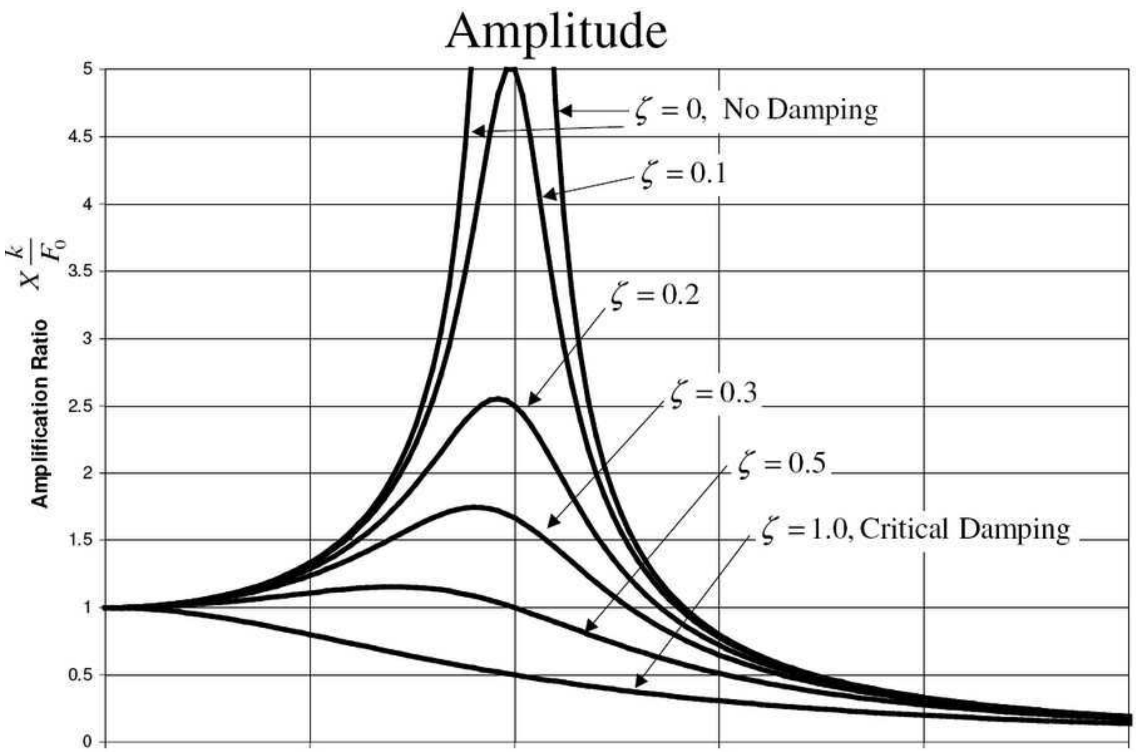
$$\phi \approx \arctan \infty = \frac{1}{2}\pi$$

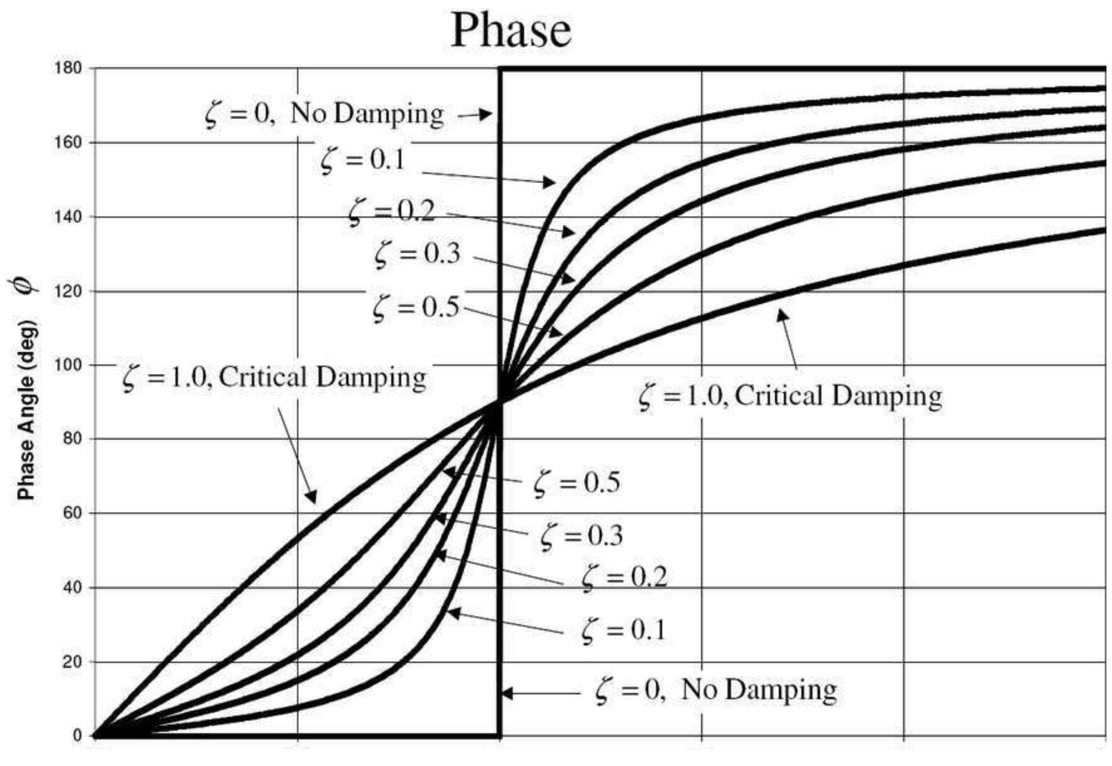
so the displacement and force are out of phase by 90°.

(iii) If $r \gg 1$, then $\omega \gg \omega_0$ and therefore

$$\frac{A}{A_s} \propto \frac{\omega_0^2}{\omega^2} = \frac{1}{r^2}$$

which is the shape of the high frequency tail of the resonance curve. The displacement and force are out of phase by 180°, for the same reason as given on page 6 for the driven undamped oscillator.


Amplitude
$\zeta = 0$, No Damping
$\zeta = 0.1$
$\zeta = 0.2$
$\zeta = 0.3$
$\zeta = 0.5$
$\zeta = 1.0$, Critical Damping
Amplification Ratio $X\frac{k}{F_0}$
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0


Phase
$\zeta = 0$, No Damping
$\zeta = 0.1$
$\zeta = 0.2$
$\zeta = 0.3$
$\zeta = 0.5$
$\zeta = 1.0$, Critical Damping
Phase Angle (deg) ϕ
180
160
140
120
100
80
60
40
20
0

4CCP1501 Mass on a vertical spring

In lecture 2 I used a horizontal mass on a spring so that I could introduce you to simple harmonic motion without worrying about the force of gravity acting on the mass in addition to the force due to the spring. But don't worry about gravity in the case of a mass on a spring. Here's why.

Imagine I hang a mass on a vertical spring having somehow turned off gravity. Then the spring will feel no force and will be relaxed; the tension in the spring is zero. Let us say that in this condition when there is neither force on the mass or the spring that the position of the mass is at $x = 0$. Now I contrive to turn on gravity and the mass descends and stretches the spring by an amount x_1, so it is now in the position $x = x_1$. The force on the mass is mg and this is therefore the tension in the spring. If the spring constant is k, the force (or tension) per unit extension then

$$x_1 = \frac{mg}{k} \tag{1}$$

Now I pull down the mass to the amplitude A before I release it and so the total extension of the spring is $x_1 + A$ and so the total tension in the spring is

$$\begin{aligned} F_{\text{spring}} &= -k(x_1 + A) \\ &= -kx_1 - kA \\ &= -mg - kA \end{aligned}$$

The total force on the spring is this tension plus the force due to gravity so

$$\begin{aligned} F_{\text{total}} &= F_{\text{spring}} + mg \\ &= -mg - kA + mg \\ &= -kA \end{aligned}$$

This is just the same as for the horizontal spring and now I don't need to invent a frictionless surface. So the vertical and horizontal springs obey the same equation of motion—the gravity cancels out in the former and is absent in the latter.