9 WAVE PHENOMENA (AHL)

Introduction

In this topic we develop many of the concepts introduced in Topic 4. In general, a more mathematical approach is taken and we consider the usefulness of modelling using a spreadsheet – both for graphing and developing relationships through iteration.

9.1 Simple harmonic motion

Understanding

- The defining equation of SHM
- Energy changes

Nature of science

The importance of SHM

The equation for simple harmonic motion (SHM) can be solved analytically and numerically. Physicists use such solutions to help them to visualize the behaviour of the oscillator. The use of the equations is very powerful as any oscillation can be described in terms of a combination of harmonic oscillators using Fourier synthesis. The modelling of oscillators has applications in virtually all areas of physics including mechanics, electricity, waves and quantum physics. In this sub-topic we will model SHM using a simple spreadsheet and see how powerful this interpretation can be.

Applications and skills

- Solving problems involving acceleration, velocity and displacement during simple harmonic motion, both graphically and algebraically
- Describing the interchange of kinetic and potential energy during simple harmonic motion
- Solving problems involving energy transfer during simple harmonic motion, both graphically and algebraically

Equations

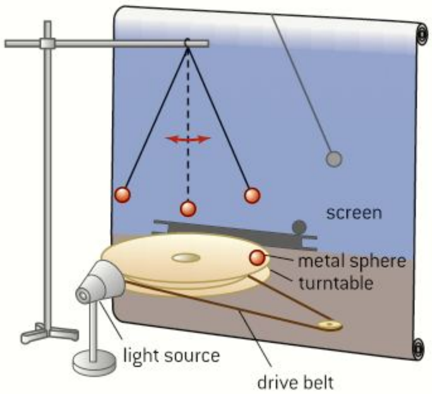
- angular velocity–period equation: $\omega = \frac{2\pi}{T}$
- defining equation for shm: $a = -\omega^2 x$
- displacement–time equations: $x = x_0 \sin \omega t$; $x = x_0 \cos \omega t$
- velocity–time equations: $v = \omega x_0 \cos \omega t$; $v = -\omega x_0 \sin \omega t$
- velocity-displacement equation: $v = \pm\omega\sqrt{(x_0^2 - x^2)}$
- kinetic energy equation: $E_K = \frac{1}{2}m\omega^2(x_0^2 - x^2)$
- total energy equation: $E_T = \frac{1}{2}m\omega^2 x_0^2$
- period of simple pendulum: $T = 2\pi\sqrt{\frac{l}{g}}$
- period of mass–spring: $T = 2\pi\sqrt{\frac{m}{k}}$

Introduction

In this sub-topic we treat SHM more mathematically but restrict ourselves to two systems – the simple pendulum and of a mass oscillating on a spring. Each of these systems is isochronous and is usually lightly damped; this means that a large number of oscillations occur before the energy in the system is transferred to the internal energy of the system and the surrounding air.

Angular speed or frequency (ω)

In Sub-topic 6.1 we considered the angular speed ω in relation to circular motion; it is the rate of change of angle with time and is also called **angular frequency**. It is measured in radians per second (rad s^{-1}). This quantity is important when we deal with simple harmonic motion because there is a very close relationship between circular motion and SHM. This relationship can be demonstrated using the apparatus shown in figure 1.

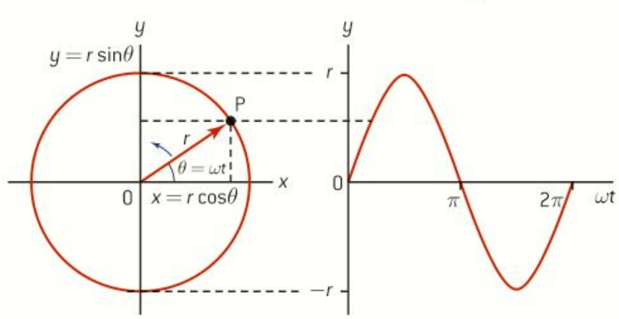


▲ Figure 1 Comparison of SHM and circular motion.

A metal sphere is mounted on a turntable that rotates at a constant angular speed. A simple pendulum is arranged so that it is in line with the sphere and oscillates with the same periodic time as that of the turntable – a little trial and error should give a good result here. The pendulum and the turntable are illuminated by a light that is projected onto a screen. The shadows projected, onto the screen, of the circular motion of the sphere and oscillatory motion of the pendulum show these motions to be identical.

Circular motion and SHM

In mathematical terms the demonstration in figure 1 is equivalent to projecting the two-dimensional motion of a point onto the single dimension of a line. Imagine a point P rotating around the perimeter of a circle with a constant angular speed ω. The radius of the circle r joins P with the centre of the circle O. At time $t = 0$ the radius is horizontal and at time t it has moved through an angle θ radians. For constant angular speed $\omega = \frac{\theta}{t}$, rearranging this gives $\theta = \omega t$. Projecting P onto the y-axis gives the vertical component of r as $r \sin\theta$. Projecting P onto the x-axis gives the horizontal component of r as $r \cos\theta$. The variation of the vertical component with time or angle is shown on the right of figure 2 and takes the form of a sine curve. Because the rate of rotation is constant, the angle θ or ωt is proportional to time. If we drew a graph of y against t the quantities 2π and π would be replaced by T and $\frac{T}{2}$ respectively.



▲ Figure 2 Projection of circular motion on to a vertical line.

The equations of the projections are $y = y_0 \sin\omega t$ and $x = x_0 \cos\omega t$.

Here y_0 and x_0 are the maximum values of y and x, which in this case are identical to r. These are the amplitudes of the motion.

You may remember from Sub-topic 4.1 that in SHM the displacement, velocity and acceleration all vary sinusoidally with time. Thus, the projection of circular motion on the vertical or horizontal takes the same shape as SHM. This is very useful in analysing SHM.

The relationship between displacement, velocity, and acceleration

In Sub-topic 4.1 we saw that, starting from the displacement–time curve, we could derive the velocity–time and acceleration–time curves from the gradients. This can be done graphically, but another technique is to differentiate the equations with respect to time (differentiation is equivalent to finding the gradient).

We have seen from the comparison with circular motion that the displacement takes the form: $x = x_0 \sin \omega t$ or $x = x_0 \cos \omega t$ depending on when we start timing. It really doesn't matter whether the projection is onto the x-axis or the y-axis – therefore we can use x and y interchangeably.

Let's start with $x = x_0 \sin \omega t$

This means that the velocity is given by $v\left(=\frac{dx}{dt}\right) = x_0\,\omega \cos \omega t$, where x_0 is the amplitude and ω the angular frequency.

The maximum value that cosine can take is 1 so the maximum velocity is $v_0 = x_0\,\omega$ thus making the equation for the velocity at time t become $v = v_0 \cos \omega t$

We know that the acceleration will be the gradient of a velocity–time graph so we have

$$a = \frac{dv}{dt} = -v_0\omega \sin \omega t = -x_0\,\omega^2 \sin \omega t$$

As for cosine, the maximum value that sine can take is 1 so $a_0 = v_0\,\omega = x_0\,\omega^2$ giving $a = -a_0 \sin \omega t$

Comparing the equations $x = x_0 \sin \omega t$ and $a = -x_0\,\omega^2 \sin \omega t$ we can see a common factor of $x_0 \sin \omega t$

meaning that

$$a = -\omega^2 x$$

This may remind you, in Sub-topic 4.1, we saw that $a = -kx$ for SHM.

So the constant k must actually be ω^2 (the angular frequency squared). We will look at the significance of ω^2 very soon.

In examinations, you can be asked to find maximum values for velocity and acceleration – this makes the calculations easier because the maximum of sine and cosine are each 1 and, therefore, we don't need to include the sine or cosine term in our calculations. Thus the maximum velocity will be $v_0 = x_0\,\omega$ and the magnitude of the maximum acceleration will be $a_0 = x_0\,\omega^2$ (the direction of a_0 will be opposite that of x_0).

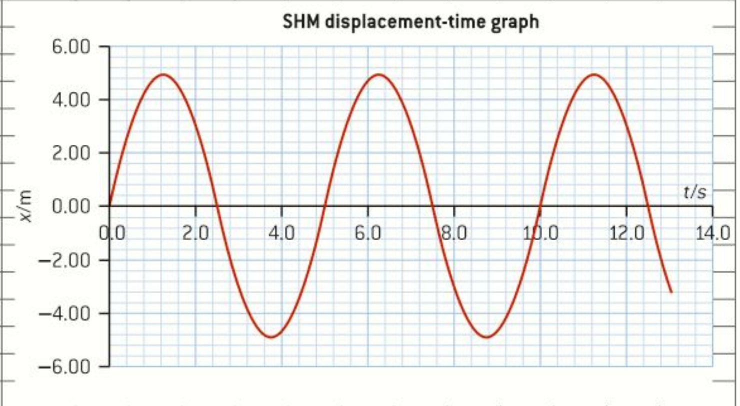
Modelling SHM with a spreadsheet

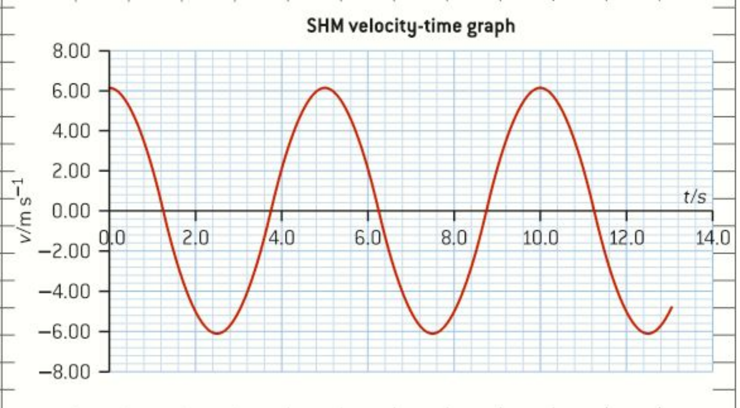
Much can be learned about SHM by using a spreadsheet to graph it. Figure 3 is a screen shot of part of a spreadsheet set up for this purpose

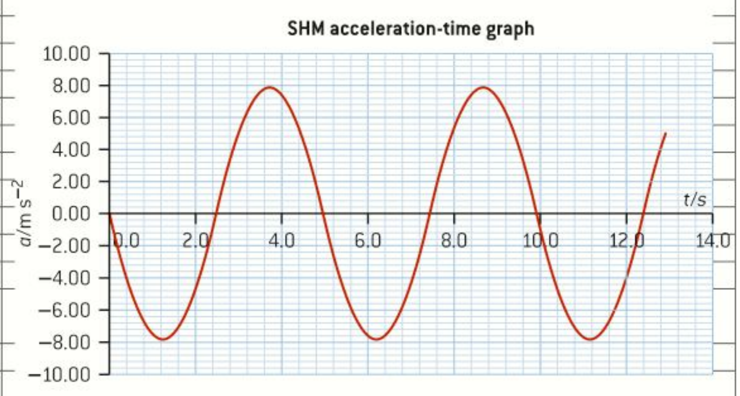
Note

- You will meet differential calculus in your Mathematics or Mathematics Studies course. This is not the place to teach you to differentiate – you will not be expected use calculus on your IB Physics course. However, for students studying physics, engineering and allied subjects at a higher level, calculus will form a major aspect of your course. In this case, we will differentiate the equations, but it is the results that are important not the method of obtaining them.

	A	B	C	D	E	F	G	H	I	J
1	time/s	displacement/m	velocity/m s^{-1}	acceleration/m s^{-2}			$t_1 =$	0	$x_0 =$	5
2	0.0	0.00	6.30	0.00			$\Delta t =$	0.2	$\omega =$	1.26
3	0.2	1.25	6.10	−1.98						
4	0.4	2.41	5.52	−3.83						
5	0.6	3.43	4.58	−5.45						
6	0.8	4.23	3.36	−6.71						
7	1.0	4.76	1.93	−7.56						
8	1.2	4.99	0.37	−7.92						
9	1.4	4.91	−1.21	−7.79						
10	1.6	4.51	−2.71	−7.16						
11	1.8	3.83	−4.05	−6.09						
12	20.	2.91	−5.12	−4.62						
13	2.2	1.81	−5.87	−2.87						
14	2.4	0.59	−6.26	−0.93						
15	2.6	−0.67	−6.24	1.06						
16	2.8	−1.88	−5.84	2.99						
17	3.0	−2.98	−5.06	4.73						
18	3.2	−3.89	−3.96	6.17						
19	3.4	−4.55	−2.62	7.22						
20	3.6	−4.92	−1.11	7.81						
21	3.8	−4.99	0.48	7.92						
22	4.0	−4.73	2.03	7.52						
23	4.2	−4.18	3.45	6.64						
24	4.4	−3.37	4.66	5.35						
25	4.6	−2.34	5.57	3.72						
26	4.8	−1.17	6.13	1.85						
27	5.0	0.08	6.30	−0.13						
28	5.2	1.33	6.07	−2.11						
29	5.4	2.49	5.46	−3.95						
30	5.6	3.49	4.51	−5.54						
31	5.8	4.27	3.27	−6.78						
32	6.0	4.79	1.83	−7.60						
33	6.2	5.0	0.26	−7.93						
34	6.4	4.89	−1.31	−7.76						
35	6.6	4.48	−2.81	−7.11						
36	6.8	3.78	−4.13	−6.00						
37	7.0	2.24	−5.18	−4.51						
38	7.2	1.73	−5.91	−2.74						
39	7.4	0.50	−6.27	−0.84						
40	7.6	−0.75	−6.23	1.20						
41	7.8	−1.96	−5.79	3.11						
42	8.0	−3.05	−5.00	4.84						
43	8.2	−3.94	−3.88	6.25						
44	8.4	−4.58	−2.52	7.28						
45	8.6	−4.94	−1.00	7.84						
46	8.8	−4.98	0.58	7.90						
47	9.0	−4.71	2.13	7.47						
48	9.2	−4.14	3.54	6.57						
49	9.4	−3.31	4.73	5.25						
50	9.6	−2.27	5.62	3.60						







▲ Figure 3 Spreadsheet for SHM.

(as before this uses Microsoft Excel but other spreadsheets will have similar functions). Column A contains incremental times starting from zero (cell H1 is copied into A2 – this means that the starting time can be changed). Cell H2 determines the time increments (in this case 0.2 s) by adding the contents of H2 to each previous cell; the times can be increased down the column. The formula in cell A3, in this case, is =A2+\$H\$2 and the formula in cell A4 is =A3+\$H\$2, etc.

The times generated, together with chosen values of x_0 and ω, are used to generate the displacement, velocity and time curves. The values for x_0 and ω are inserted into cells J1 and J2 respectively – these can then be changed at will. The equation for the displacement, $x = x_0 \sin \omega t$ is converted into the formula =$J$1*SIN(A2*$J$2), which is copied into column B. The equation for the velocity, $v = x_0 \omega \cos \omega t$ is converted into the formula =$J$1*$J$2*COS(A2*$J$2) which is copied into column C. Finally, the equation for the acceleration, $a = -x_0 \omega^2 \sin \omega t$, is converted into the formula =(−1)*$J$1*($J$2)^2*SIN(A2*$J$2) and is copied into column D.

Note

We have used a spreadsheet to show the solutions to the SHM equation. Later we will discuss how we can actually solve the SHM equation using iteration.

To produce the curves shown in figure 3, you now need to:

- insert a scatter chart
- choose the data series
- do a little formatting to improve the size and position of the chart and label the axes, etc.

The SHM equation and ω^2

Re-visiting the equation $a = -\omega^2 x$ discussed earlier in this sub-topic, we see that it fits the definition of SHM (**motion in which the acceleration is proportional to the displacement from a fixed point and is always directed towards that fixed point**). The equation is the defining equation for SHM.

The sinusoidal graphs provide the solutions to this equation with respect to time. This may seem strange because there is no apparent time factor in $a = -\omega^2 x$. This is where ω^2 comes in. We saw with circular motion that ω represents the angular speed. Therefore ω^2 is simply the square of this measured in rad² s⁻². ω^2 has dimensions equivalent to $\left(\frac{1}{\text{time}}\right)^2$.

In circular motion we defined ω as being $\frac{\Delta\theta}{\Delta t}$ or simply $\frac{\theta}{t}$ when it is constant. For a complete revolution the angle will be 2π radians and the time will be the periodic time T. This means that $\omega = \frac{2\pi}{T}$ or, alternatively, $\omega = 2\pi f$; comparing this equation with $f = \frac{1}{T}$ explains why we can call ω the angular frequency.

Worked example

An object performs SHM with a period of 0.40 s and has amplitude of 0.20 m. The displacement is zero at time zero. Calculate:

a) the maximum velocity

b) the magnitude of the velocity after 0.10 s

c) the maximum acceleration of the object.

Solution

a) $v_0 = x_0 \omega = x_0 \times \dfrac{2\pi}{T}$

$v_0 = 0.20 \times \dfrac{2\pi}{0.40} = 3.1 \text{ m s}^{-1}$

b) As the displacement is zero at time zero this must be a sine or negative sine wave. Thus the velocity will be a cosine or negative cosine.

We are only asked to find the magnitude of the velocity so it doesn't matter which of the cosine curves the motion really takes.

Using $v = x_0 \omega \cos \omega t$

$\omega t = \dfrac{2\pi t}{T} = \dfrac{(2\pi \times 0.10)}{0.40} \approx \dfrac{\pi}{2}$ rad

So $v = 0.20 \times \dfrac{2\pi}{0.40} \times \cos \dfrac{\pi}{2} = 0$

This could have been done without the full calculation once we had decided that it was a cosine. 0.10 s represents the time for a quarter of a period (0.40 s); the value of any cosine at a quarter of a period (or $\frac{\pi}{2}$ radian) is zero.

c) $a_0 = x_0 \omega^2 = 0.20 \times \left(\dfrac{2\pi}{0.40}\right)^2 \approx 49 \text{ m s}^{-1}$

The velocity equation

The derivation of the SHM equation is for reference and you don't need to reproduce it – following it through, however, will help you to understand what is going on.

We have already seen that $v = x_0\,\omega \cos \omega t$ and $x = x_0 \sin \omega t$ and you may know that $\sin^2\theta + \cos^2\theta = 1$.

This means that $\cos\theta = \pm\sqrt{1-\sin^2\theta}$ (don't forget that squaring either the positive or negative cosine will give $(+)\cos^2\theta$. This means that $\pm$ must be included in the square root of this.

So $v = \pm x_0\omega\sqrt{1-\sin^2\omega t}$

But $\sin\omega t = \frac{x}{x_0}$ so $\sin^2\omega t = \left(\frac{x}{x_0}\right)^2$

$$v = \pm x_0\,\omega\sqrt{1-\left(\frac{x}{x_0}\right)^2} = \pm\omega\sqrt{{x_0}^2 - {x_0}^2\left(\frac{x}{x_0}\right)^2} = \pm\omega\sqrt{{x_0}^2 - x^2}$$

The equation $v = \pm\omega\sqrt{{x_0}^2 - x^2}$ is useful for finding the velocity at a particular position when you know the amplitude and period (or frequency or angular frequency) – you don't need to know the time being considered.

Worked example

An object oscillates simple harmonically with frequency 60 Hz and amplitude 25 mm. Calculate the velocity at a displacement of 8 mm.

Solution

$\omega = 2\pi f = 120\pi$ (there is no need to calculate the value here but do not leave π in the answer in an examination.

$$v = \pm\omega\sqrt{{x_0}^2 - x^2}$$
$$= \pm 120\pi\sqrt{(25\times10^{-3})^2 - (8\times10^{-3})^2}$$
$$= \pm 8.9\ \text{m s}^{-1}$$

Note

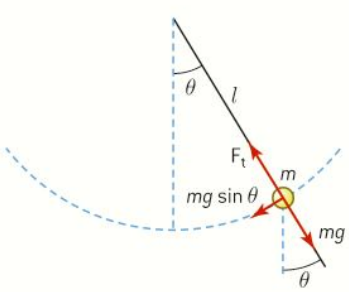
- $\pm$ tells us that the object could be going in either of the two opposite directions.
- Don't forget ${x_0}^2 - x^2 \neq (x_0 - x)^2$; it is a common mistake for students to equate these two expressions.

Simple harmonic systems

1. The simple pendulum

The simple pendulum represents a straightforward system that oscillates with SHM when its amplitude is small. When the pendulum bob (the mass suspended on the string) is displaced from the rest position there is a component of the bob's weight that tends to restore the bob to its normal rest or equilibrium position. A condition of a system oscillating simple harmonically is that there is a restoring force that is proportional to the displacement from the equilibrium position – this is, in effect, tying in with the equation defining SHM because $F = ma$ and $a = -\omega^2 x$ this means that $F = -m\omega^2 x$.

Figure 4 shows the forces acting on a pendulum bob. The bob is in equilibrium along the radius when the tension in the string F_t equals the component of the weight in line with the string ($=mg\cos\theta$). The component of the weight perpendicular to this is not in equilibrium and provides the restoring force.



▲ Figure 4 Restoring force for simple pendulum.

So the restoring force must be equal to the mass multiplied by the acceleration according to Newton's second law of motion:

$$mg \sin \theta = ma$$

For a small angle $\quad \sin \theta \approx \theta \approx \frac{x}{l}$

$$\text{rearranging gives } -m\left(\frac{g}{l}\right)x = ma$$

(The minus sign is because the displacement (to the right) is in the opposite direction to the acceleration (to the left) in figure 4.

Cancelling m

$$a = -\left(\frac{g}{l}\right)x$$

this compares with the defining equation for SHM $a = -\omega^2 x$ leading to

$$\omega^2 = \frac{g}{l}$$

As the periodic time for SHM is given by $T = \frac{2\pi}{\omega}$, this shows that $T = 2\pi\sqrt{\frac{l}{g}}$.

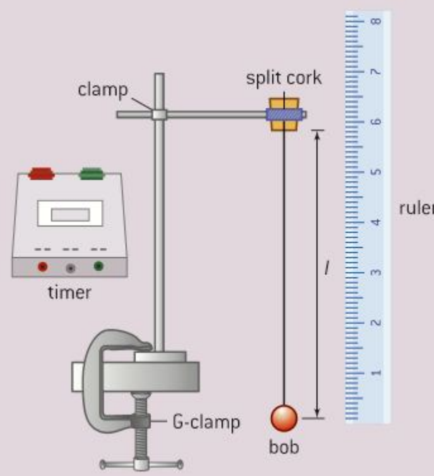
Note

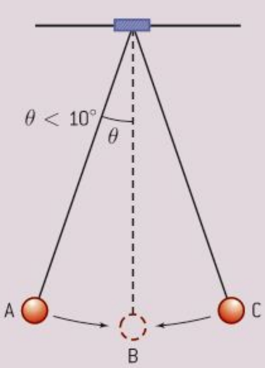
- The period of a simple pendulum is independent of the mass of the pendulum.
- The period of a simple pendulum is independent of the amplitude of the pendulum.
- l is the length of the pendulum from the point of suspension to the centre of mass of the bob.
- The equation only applies to small oscillations (swings making angle of less than 10° with the rest position).

Investigate!

1 Experimenting with the simple pendulum:

- attach a piece of thread of length about 2 metres to a pendulum bob (a lump of modelling clay would do for this)
- suspend the thread through a split cork to provide a stable point of suspension
- align a pin mounted in another piece of modelling clay with the rest position of the bob – to act as a reference point (sometimes called a fiducial marker)
- set the bob oscillating through a small angle and start timing as the bob passes the fiducial marker (you should consider why your measurements are likely to be more reliable when the bob is moving at its fastest)
- time thirty oscillations (remember that each oscillation is from **A** to **C** and back to **A**)
- measure the length of the thread from the point of suspension to the centre of mass of the bob (take this to be the centre of the bob)
- repeat the procedure and find an average period for the pendulum





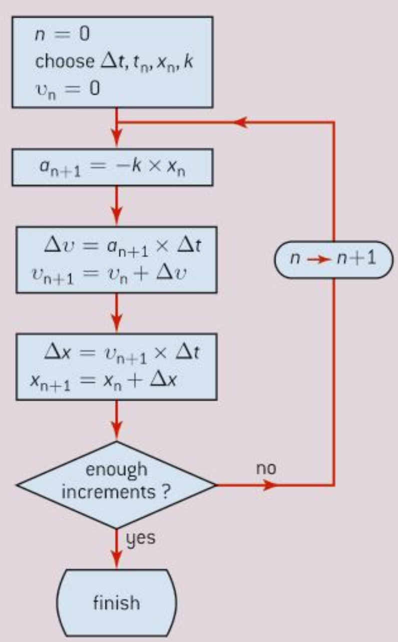
- repeat your measurements for a range of lengths
- first plot a graph of periodic time T against length l
- referring to the simple pendulum equation you may see that this graph should not be linear – what should you plot in order to linearize your results?
- how can you use the graph to calculate a value for the acceleration of freefall g?
- don't forget to do an error analysis and compare your result with the accepted value of g.

2 Use of iteration

Making use of the powerful iterative functionality of a spreadsheet can be a valuable learning aid. With iteration a graphical investigation is possible from basic principles and basic equations without knowing the solution and without advanced mathematics.

- you will need to set up a worksheet with the headings time, displacement, acceleration, change in velocity, velocity and change in displacement
- you now need to set up the initial conditions:
  1. choose a time increment for the iteration (make it fairly large until your spreadsheet is working well)
  2. start your time column at zero and decide how long you want to run the iteration for (make this, say, 10 time increments initially)
  3. choose an amplitude value for your oscillation
  4. set the initial displacement value equal to the amplitude
  5. choose a constant (k or ω^2)
  6. set the initial velocity to be zero
- the acceleration will always be $-k$ multiplied by the displacement
- because acceleration is the change in velocity divided by your time increment the change in velocity will be the acceleration multiplied by the time increment – when you add this to your previous velocity you will get the new velocity (actually it's the average velocity through the time increment)
- because velocity is the change in displacement divided by the time increment, you can find the change in displacement by multiplying your current velocity value by your chosen time increment
- the new value of the displacement will be the previous value added to the change in displacement
- This now feeds back into finding the next value of the acceleration and the cycle repeats and you generate your data for which you can plot displacement-time, velocity-time and acceleration-time graphs (a example of the spreadsheet is included on the website).

The following flow chart illustrates the iteration:



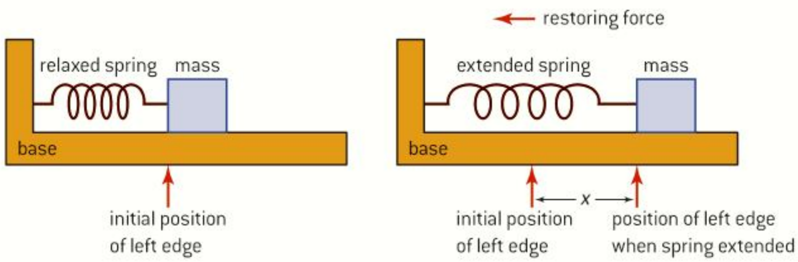
Subscripts n represent the current value of a variable and $n + 1$ the next value; at each loop the next value always replaces the previous one.

2. Mass–spring system

We have focused on a simple pendulum as being a very good approximation to SHM. A second system which also behaves well and gives largely undamped oscillations is a mass–spring system. We will consider a mass being oscillated horizontally by a spring (see figure 5); this is more straightforward than taking account of including the effects of gravity experienced in vertical motion. We will assume that the friction between the mass and the base is negligible. The mass, therefore, exchanges elastic potential energy (when it is fully extended and compressed) with kinetic energy (as it passes through the equilibrium position).

Note

- The period of the mass-spring system is independent of amplitude (for small oscillations).
- The period of the mass-spring system is independent of the acceleration of gravity.



▲ Figure 5 Restoring force for mass-spring system.

When a spring (having spring constant k) is extended by x from its equilibrium position there will be a restoring force acting on the mass given by $F = -kx$ (the force is in the opposite direction to the extension).

Using Newton's second law

$$ma = -kx$$

which can be written as

$$a = -\left(\frac{k}{m}\right)x$$

this compares with the defining equation for SHM

$$a = -\omega^2 x$$

$$\omega^2 = \left(\frac{k}{m}\right)$$

As the periodic time for SHM is given by

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

this shows that

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

When the spring is compressed the quantity x represents the compression of the spring. When the mass is to the left of the equilibrium position the compression is positive but the restoring force will be negative. This, therefore, leads to the same outcomes as for extensions.

Worked example

The mass–spring system is used in many common accelerometer designs. A mass is suspended by a pair of springs which displaces when acceleration occurs. An accelerometer contains a mass of 0.080 kg coupled to a spring with spring constant of 4.0 kN m^{-1}. The amplitude of the mass is 20 mm. Calculate:

a) the maximum acceleration

b) the natural frequency of the mass.

Solution

a) $a = -\left(\frac{k}{m}\right)x$

$$a = -\left(\frac{4.0 \times 10^3}{0.08}\right) 20 \times 10^{-3}$$

$$= 1000 \text{ m s}^{-2}$$

b) $T = 2\pi\sqrt{\frac{m}{k}}$ so $f = \frac{1}{T}$

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

$$f = \frac{1}{2\pi}\sqrt{\frac{4.0 \times 10^3}{0.08}} = 36 \text{ Hz}$$

Energy in SHM systems

We have seen that, in a simple pendulum, there is energy interchange between gravitational potential and kinetic; in a horizontal mass–spring system the interchange is between elastic potential and kinetic energy. Because each system involves kinetic energy we will focus on this form of energy and bear in mind that the potential energy will always be the difference between the total energy and the kinetic energy at a particular time. The total energy will be equal to the maximum kinetic energy.

From Topic 2 we know that the kinetic energy of an object of mass m, moving at velocity v, is given by

$$E_K = \frac{1}{2}mv^2$$

We also know that the equation for the velocity at a particular position is

$$v = \pm\omega\sqrt{x_0^2 - x^2}$$

Combining these equations gives the kinetic energy at displacement x:

$$E_K = \frac{1}{2}m\,\omega^2\left(x_0^2 - x^2\right)$$

This tells us that the maximum kinetic energy will be given by

$$E_{K\max} = \frac{1}{2}m\,\omega^2 x_0^2$$

and this must be the total energy (when the potential energy is zero) so we can say

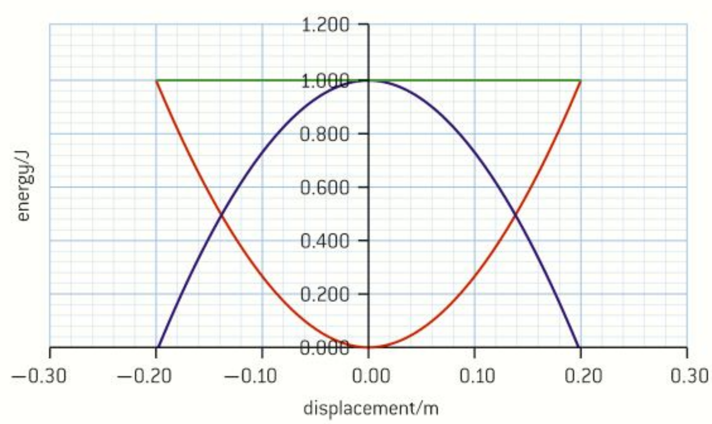
$$E_T = \frac{1}{2}m\,\omega^2 x_0^2$$

The potential energy at any position will be the difference between the total energy and the kinetic energy

so

$$E_P = E_T - E_K = \frac{1}{2}m\,\omega^2 x_0^2 - \frac{1}{2}m\,\omega^2\left(x_0^2 - x^2\right) = \frac{1}{2}m\,\omega^2 x^2$$

Figure 6 illustrates the variation with displacement of energy: the green line shows the total energy, the red the potential energy and the blue the kinetic energy. At any position the total energy is the sum of the kinetic energy and the potential energy – as we would expect.



▲ Figure 6 Variation of energy with displacement.

The variation of the potential energy and kinetic energy with displacement are both parabolas. With all the quantities in the total energy equation being constant, the total energy is, of course, constant for an undamped system.

In addition to looking at the variation of energy with displacement we should consider the variation of energy with time. Again, let us start with the kinetic energy.

The velocity varies with time according to the equation

$$v = x_0\,\omega \cos \omega t$$

so the kinetic energy will be

$$E_K = \frac{1}{2}mv^2 = \frac{1}{2}m(x_0\,\omega \cos \omega t)^2$$

or

$$E_K = \frac{1}{2}m{x_0}^2\,\omega^2 \cos^2 \omega t$$

When the cosine term equals 1, this gives the maximum kinetic energy. The maximum kinetic energy occurs when the potential energy is zero and so is numerically equal to the total energy at that instant.

$$E_T = \frac{1}{2}m{x_0}^2\,\omega^2$$

The potential energy will be the difference between the total energy and the kinetic energy so

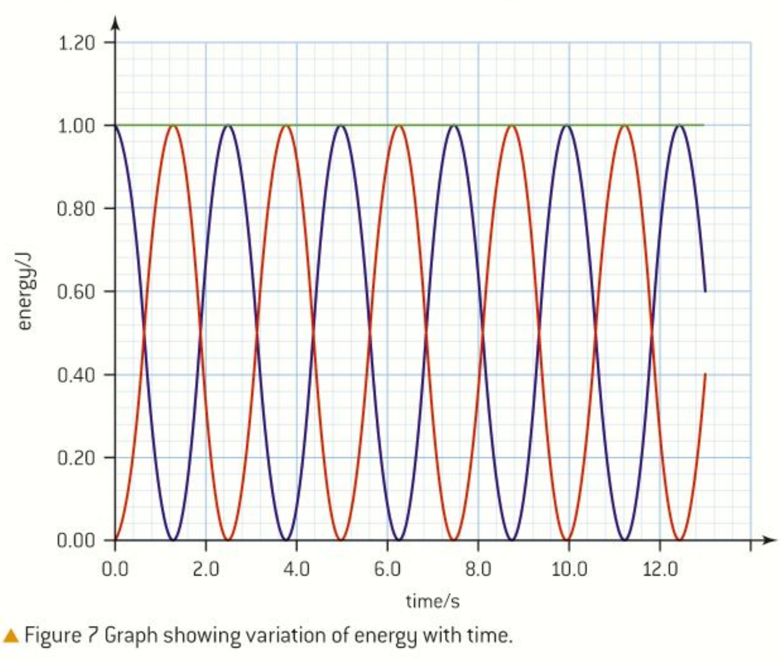
$$E_P = E_T - E_K$$

$$= \frac{1}{2}m{x_0}^2\,\omega^2 - \frac{1}{2}m{x_0}^2\,\omega^2 \cos^2 \omega t$$

$$= \frac{1}{2}m{x_0}^2\,\omega^2 \sin^2 \omega t$$

This relationships are shown on figure 7.

The green line represents the total energy, the red curve the potential energy and the blue curve the kinetic energy.



▲ Figure 7 Graph showing variation of energy with time.

Note

- The total energy is always the sum of the kinetic and potential energies.
- The graphs (unlike sine and cosine) never become negative.
- The period of the energy change is half that of the variation with time of displacement, velocity, or acceleration.
- The frequency of the energy change is twice that of the variation with time of displacement, velocity, or acceleration.

9.2 Single-slit diffraction

Understanding

→ The nature of single-slit diffraction

Nature of science

Development of theories

That rays "travel in straight lines" is one of the first theories of optics that students encounter. It comes as a surprise when this theory cannot explain the diffraction seen at the edges of shadows cast by small objects illuminated using point sources. Although partial shadows can be explained by considering light sources to be extended, this cannot account for the diffraction from a point source. The wave theory of diffraction and how diffraction can be explained in terms of wavefront propagation from secondary sources is a good example of how theories have been developed in order to explain a wider variety of phenomena.

Applications and skills

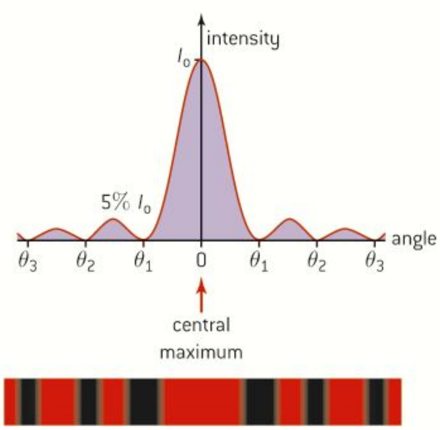
→ Describing the effect of slit width on the diffraction pattern

→ Determining the position of first interference minimum

→ Qualitatively describing single-slit diffraction patterns produced from white light and from a range of monochromatic light frequencies

Equations

→ *angle between first minimum and central maximum* $\theta = \frac{\lambda}{a}$



▲ Figure 1 Variation of intensity with angle for a diffraction pattern.

Introduction

In Sub-topic 4.4 we introduced diffraction and saw that when a wave passes through an aperture it spreads into the geometric shadow region. We also saw that the diffraction pattern consists of a series of bright and dark fringes. We will now consider the diffraction pattern in more detail.

Graph of intensity against angle

Figure 1 shows a single-slit diffraction pattern together with a graph of the variation of the intensity of the diffraction pattern with the angle measured from the straight-through position.

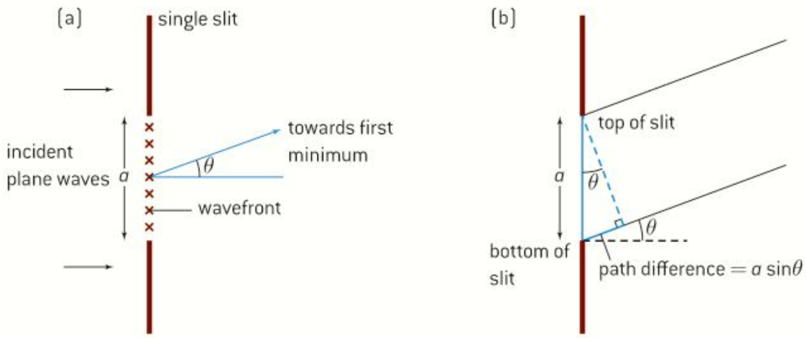
θ_1, θ_2, and θ_3 are the angles with the straight-through position made by the minima.

Note

- The central maximum has **twice the angular width** of the secondary maxima (each of these have the same angular width).
- The intensity falls off quite significantly from the principal maximum to the secondary maxima – the intensity of the first secondary maximum is approximately 5% of that of the principal maximum, the second is about 2% and the third is about 1% (figure 1 is not drawn to scale – the secondary maxima are all larger than a scale-diagram would show).
- No light reaches the centre of the minima but, in going towards the maxima, the intensity gradually increases – it is very difficult to decide the exact positions of the maxima and minima.
- The intensity is proportional to the square of the amplitude.

The single-slit equation

Returning to the work of Sub-topic 4.4 we saw how waves can superpose to give constructive interference (when they meet in phase) or destructive interference (when they meet anti-phase).



▲ Figure 2 Deriving the single slit equation.

Figure 2(a) shows a single-slit of width a. Each point on the wavefront within the slit behaves as a source of waves. These waves interfere when they meet beyond the slit. For the two waves coming from the edges of the slit making an angle θ with the straight through, there is a path difference of $a \sin \theta$. Waves from a point halfway along the slit will have a path difference of $\frac{a}{2} \sin \theta$ from the waves coming from each of the edges. When this path difference equals half a wavelength, the waves from halfway along the slit will interfere destructively with the waves coming from the bottom edge of the slit. It follows that for each point in the bottom half of the slit there will be a point in the top half of the slit at a distance $\frac{a}{2} \sin \theta$ from it. This means that when $\frac{a}{2} \sin \theta = \frac{\lambda}{2}$, there is destructive interference between a wave coming from a point in the upper half of the slit and an equivalent wave from the lower half of the slit.

Because we are dealing with small angles we can approximate $\sin \theta$ to θ and, by cancelling the factor of two, we arrive at the equation

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

The position of other minima (shown in figure 1) will be given by $\theta = \frac{n\lambda}{a}$ (where $n = 2, 3, 4, ...$) but you will not be examined on this relationship.

Note

- This is the equation for the angle of the first **minimum**.
- It follows from the equation $\theta = \frac{n\lambda}{a}$ that the minima are not actually separated by equal distances. However, for values of θ that are less than about 10°, it is a good approximation to consider the minima to be equally spaced.
- The principal or central maximum occurs because the pairs of waves from the top and bottom halves of the slit will travel the same distance and have no path difference.
- The angular width of the principal maximum (from first minimum on one side to first minimum on the other) is 2θ.

TOK

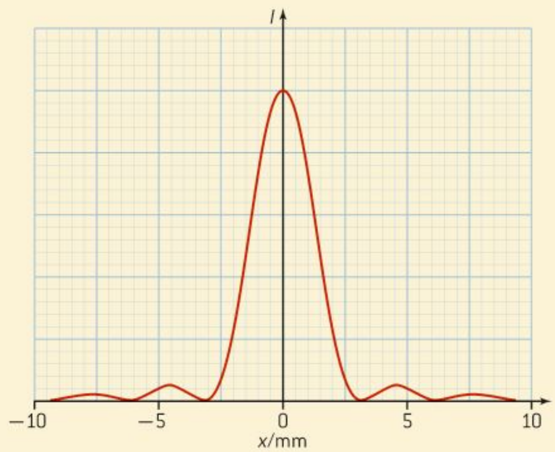
Small angle approximation

We say that diffraction is most effective when the aperature is of the same order of magnitude as the width of the slit; we can demonstrate this effectively using a ripple tank. In the equation $\theta = \frac{\lambda}{a}$ if we make $\lambda \approx a$, then $\theta = 1$ radian (or 57.3°) or, if you avoid the small angle approximation, $\sin \theta = 1$ and $\theta = 90°$. Examine how closely diffraction in a ripple tanks agrees with small angle approximation prediction. With poor agreement (as this should show) why do we continue to use the equation $\theta = \frac{\lambda}{a}$?

Worked example

a) Explain, by reference to waves, the diffraction of light at a single slit.

b) Light from a helium–neon laser passes through a narrow slit and is incident on a screen 3.5 m from the slit. The graph below shows the variation with distance *x* along the screen of intensity *I* of the light on the screen.



(i) The wavelength of the light emitted by the laser is 630 nm. Determine the width of the slit.

(ii) State **two** changes to the intensity distribution of the central maximum when the single slit is replaced by one of greater width.

Solution

a) The wavefront within the slit behaves as a series of point sources which spread circular wave fronts from them. These secondary waves superpose in front of the slit and the superposition of the waves produces the diffraction pattern.

b) (i) In this case the graph is drawn for distance not angle, so we need to calculate the angle θ in order to be able to use $\theta = \frac{\lambda}{a}$.

$\theta = \frac{s}{D} = \frac{3.0 \times 10^{-3}}{3.5}$

$= 0.86 \times 10^{-3}$ rad with s being the distance from the centre of the principal maximum to the first minimum. It is more reliable to measure 2θ from the first minimum on one side of the principal maximum to the first minimum on the other side of the principal maximum. D is the distance from the slit to the screen.

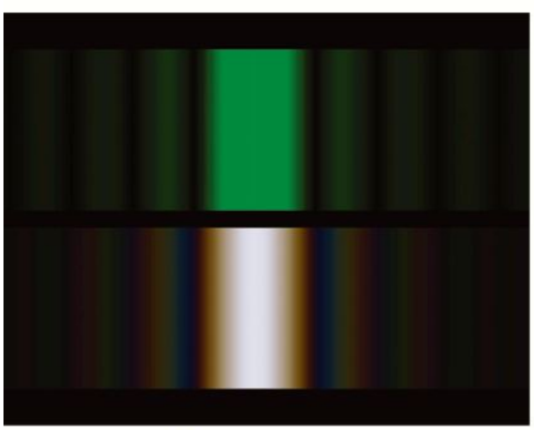
$a = \frac{\lambda}{\theta} = \frac{630 \times 10^{-9}}{0.86 \times 10^{-3}}$

$= 0.73 \times 10^{-3}$ m or 0.73 mm

(ii) With a wider slit more light is able to pass through. This will result in an increase in the intensity of the beam and so the peaks will all be higher.

a in the equation $\theta = \frac{\lambda}{a}$ increases but λ remains constant; the angle θ will decrease which means the principal maximum will become narrower and the minima will move closer together.

Single slit with monochromatic and white light



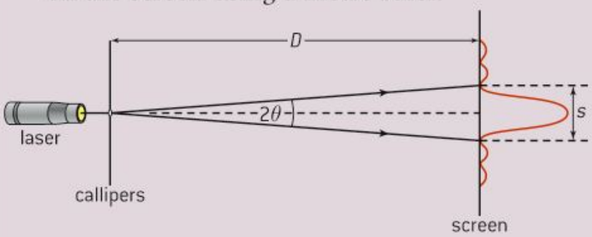
▲ Figure 3 Single slit with monochromatic and white light.

Figure 3 shows two images of the light emerging from a single-slit. The upper image is obtained using monochromatic (one frequency) green light and the lower image is obtained using white light. Both the angular width of the central maximum **and** the angular separation of successive secondary maxima depend on the wavelength of the light – this is the reason why the secondary maxima produced by the diffraction of white light are coloured. You will see that, for the secondary maxima, the blue light is less deviated than the other colours as it has the shortest wavelength. The edges of the principal maximum are coloured rather than pure white. This is because the principal maxima for the colours at the blue end of the visible spectrum are less spread than the colours at the red end; the edges are therefore a combination of red, orange, and yellow and have an orange hue.

Investigate!

Diffraction at a single slit

- Using a laser pointer, shine the light on the gap between the jaws of a pair of digital or vernier callipers (forming a slit) and project it onto a white screen. The screen needs to be at least 3 m from the callipers.
- Adjust the callipers so that the image is clear.
- Measure the separation, s, of the first minima on the screen using a metre ruler.



▲ Figure 4 Investigating laser light passing through slit.

- Measure the distance between the callipers and the screen, D, using a tape-measure.
- Calculate the wavelength of the light using $\lambda \approx a\theta$ as we know that $2\theta \approx \frac{s}{D}$
- Both the dependence of the width of the central maximum and the separation of the maxima on the wavelength of light can be investigated by using laser pointers with different colours.
- Research the wavelength ranges of the laser light and then suggest how you could modify the experiment to check the calibration of the callipers at small jaw separation.

We will return to diffraction when we look at resolution in Sub-topic 9.4.

9.3 Interference

Understanding

- → Young's double-slit experiment
- → Modulation of two-slit interference pattern by one-slit diffraction effect
- → Multiple slit and diffraction grating interference patterns
- → Thin film interference

Nature of science

Thin film interference

The observation of colour is not simply a question of colour pigmentation. Certain mollusc and beetle "shells", butterfly wings, and the feathers of hummingbirds and kingfishers can produce beautiful colours as a result of thin film interference known as "iridescence". The observed colour changes depend on the angle of illumination or viewing, and are caused by multiple reflections from the surfaces of films with thicknesses similar to the wavelength of light. These natural aesthetics require the analysis of the physics of interference.

Applications and skills

- → Qualitatively describing two-slit interference patterns, including modulation by one-slit diffraction effect
- → Investigating Young's double-slit experimentally
- → Sketching and interpreting intensity graphs of double-slit interference patterns
- → Solving problems involving the diffraction grating equation
- → Describing conditions necessary for constructive and destructive interference from thin films, including phase change at interface and effect of refractive index
- → Solving problems involving interference from thin films

Equations

fringe separation for double slit: $s = \frac{\lambda D}{d}$

diffraction grating equation: $n\lambda = d \sin\theta$

light reflection by parallel-sided thin film:

constructive interference $2dn = (m + \frac{1}{2})\lambda$

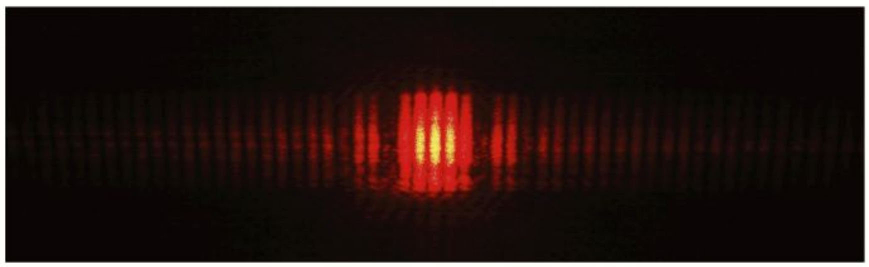
destructive interference $2dn = m\lambda$

Introduction

In Sub-topic 4.4 we considered interference of the waves emitted by two coherent sources. We saw that interference is a property of all waves, and we considered the equal fringe spacing in the Young double-slit experiment. In this sub-topic we focus on the intensity variation in a double-slit experiment and see how interference is achieved using multiple slits. We then look at a second way of achieving interference, using division of amplitude instead of division of wavefront as with double-slit interference.

Intensity variation with the double-slit

Figure 1 shows the image of the light from a helium–neon laser that has passed through a double slit. The alternate red and dark fringes are equally spaced as we saw in Sub-topic 4.4. However, looking at the image closely we see that there are extra dark regions.



▲ Figure 1 Double-slit diffraction pattern for light from He–Ne laser.

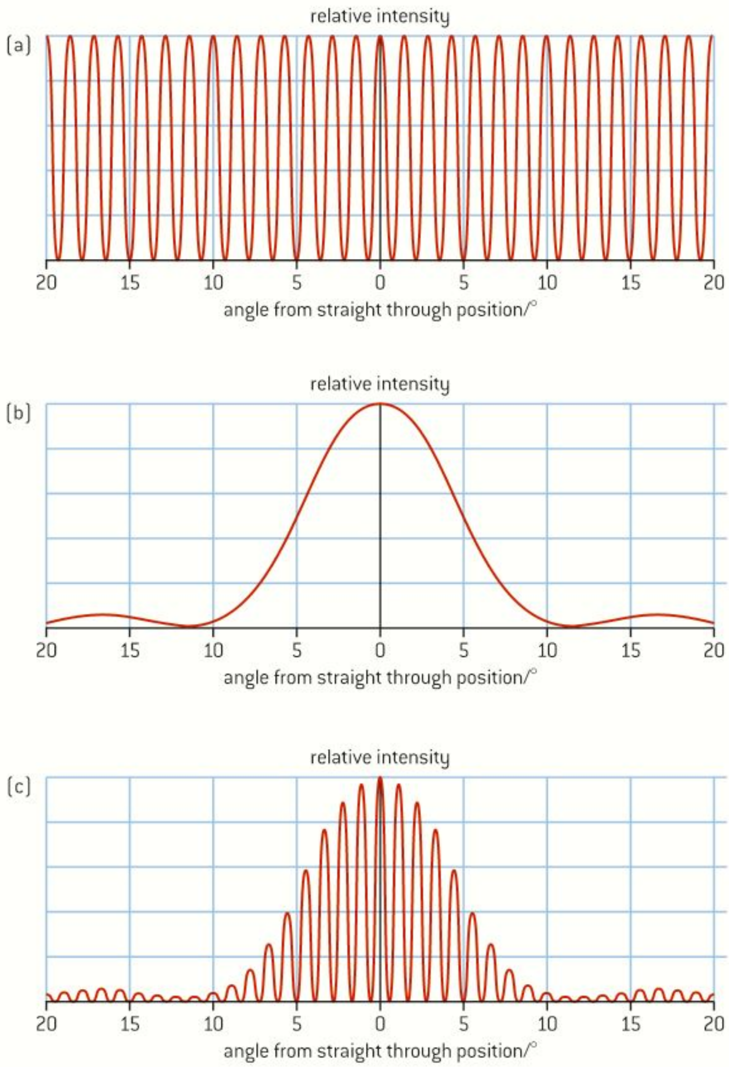
We know from Sub-topic 9.2 that a single slit produces a diffraction pattern with a very intense principal maximum and much less intense secondary maxima. A double slit is, of course, two single slits so each of the slits produces a diffraction pattern and the waves from the two slits interfere. The two effects mean that the intensity of the interference pattern is not constant, but is modified by the diffraction pattern to produce the intensity.

Figure 2(a) below shows how the relative intensity would vary for a double-slit interference pattern without any modification due to diffraction. By using relative intensity we avoid the need to think about the actual intensity values and units. Figure 2(b) shows the variation of relative intensity with angle for a single slit. Figure 2(c) shows the superposition of the two effects so that the single-slit diffraction pattern behaves as the envelope of the interference pattern. Shaping a pattern in this way is called *modulation* and is important in the theory of AM (amplitude modulation) radio. You will note that the fringe spacing does not change between figures 2(a) and (c) but that the bright fringes occurring between 11° and 14° in figure 2(a) are reduced to a much lower intensity. An interference maximum coinciding with a diffraction minimum is suppressed and does not appear in the overall pattern.

We saw in Sub-topic 4.4 that the fringe separation s for light of wavelength λ is given by

$$s = \frac{\lambda D}{d}$$

where D is the distance from the double slit to the screen and d is the separation of the slits. The value of s used in this equation is for the interference pattern not the modulated pattern caused by diffraction.



▲ Figure 2 Combination of diffraction and interference.

Multiple-slit interference

We have now seen the interference patterns produced both by a single slit and a double slit. What happens if there are more than two slits? The answer is that the bright fringes, which come from constructive interference of the light waves from different slits, remain in the same positions as for a double slit but the pattern becomes sharper. The bright fringes are narrower and their intensity is proportional to the square of the number of slits. Why is this?

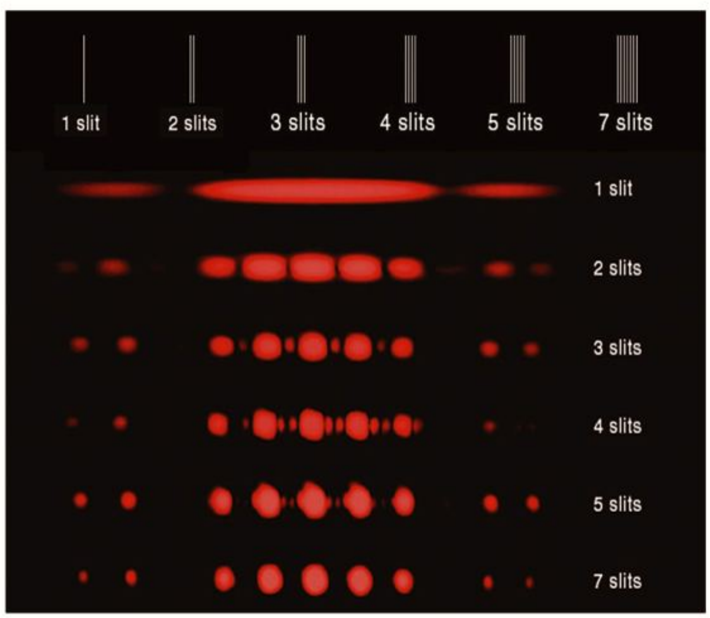
At the centre of the principal maximum the waves reaching the screen from all of the slits are in phase and so it is very bright here. By moving to a position close to the centre of this maximum the path difference between the light from two adjacent slits has changed by such a small amount that it hardly affects the interference and so it is still bright. Slits further away from each other will have more likelihood of a greater path difference and therefore meeting out of phase.

Imagine having three slits: at the principal maximum the path difference between the waves will be small as their paths are nearly parallel. Each of the three wave trains will meet nearly in phase at positions around the principal maximum peak, which makes it fairly wide.

Now imagine having 100 slits: for every slit there will a second slit somewhere that transmits a wave with a half wavelength path difference from the first. When the waves interfere at a position close to the centre of the maximum, destructive interference occurs and reduces the overall intensity. This will be true for the waves coming from the two slits adjacent to this pair, the two next to those and so on. Increasing the number of slits will give destructive interference close to the centre of the maxima – this reduces the width of the central maximum (and the other maxima too). The extra energy that is needed to increase the intensity of the maxima must come from the regions that are now darker, so the maxima are more intense.

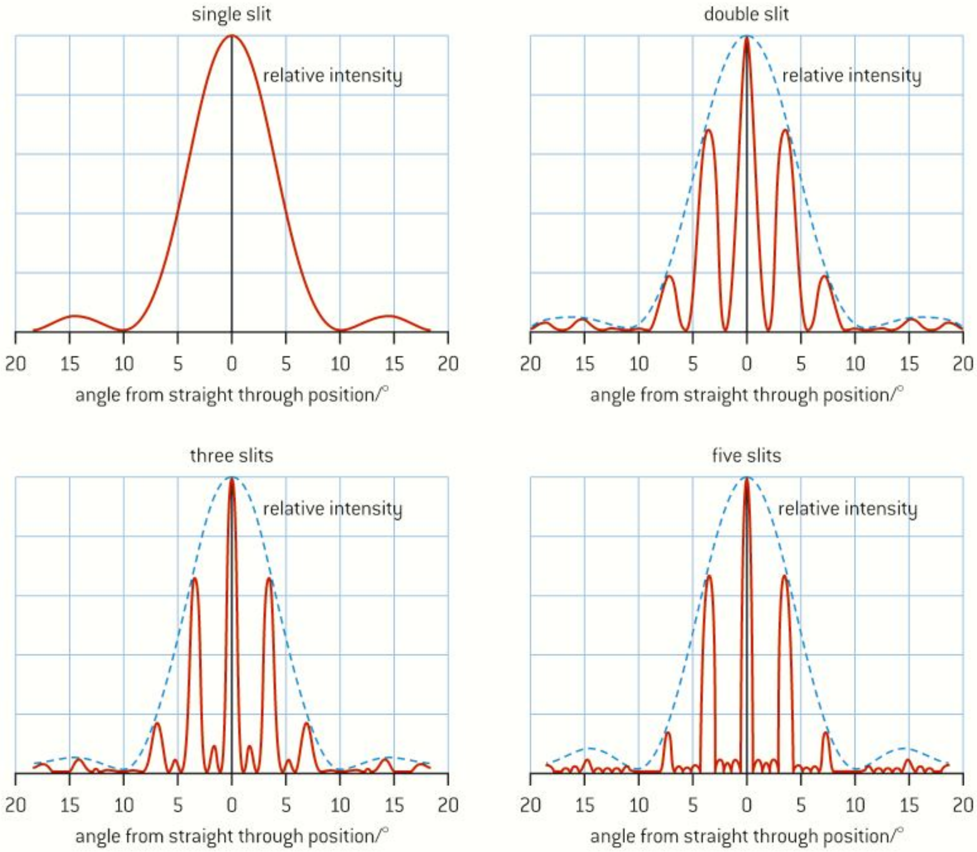
The mathematics of modulation of waves is quite complex and for the purpose of the IB Physics course you will simply need to recognize that the modulation is happening and remember the effect of increasing the number of slits.

Figure 3 shows the interference patterns obtained when red laser light passes through various numbers of slits of identical width. Figures 3 and 4 show that the single-slit diffraction pattern always acts as an envelope for the multiple-slit interference patterns.



▲ Figure 3 The effect of increasing the number of slits on an interference pattern.

In figure 4 the relative intensities have been drawn the same size in order to show the increased sharpness; however, for two, three, and five slits the actual intensities are in the ratio 1 : 9 : 25.

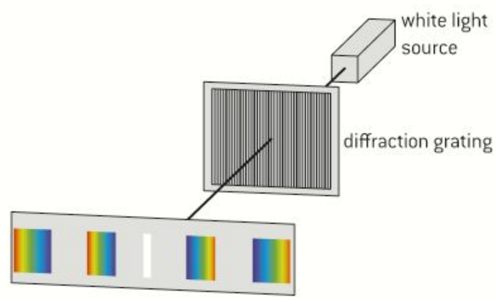


▲ Figure 4 The effect of increasing the number of slits on variation of intensity.

The diffraction grating

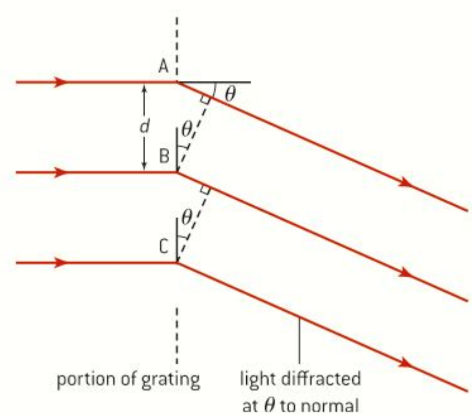
A diffraction grating is a natural consequence of the effect on the interference pattern when the number of slits is increased. Diffraction gratings are used to produce optical spectra. A grating contains a large number of parallel, equally spaced slits or "lines" (normally etched in glass or plastic) – typically there are 600 lines per millimetre (see figure 5). When light is incident on a grating it produces interference maxima at angles θ given by

$$n\lambda = d\sin\theta$$

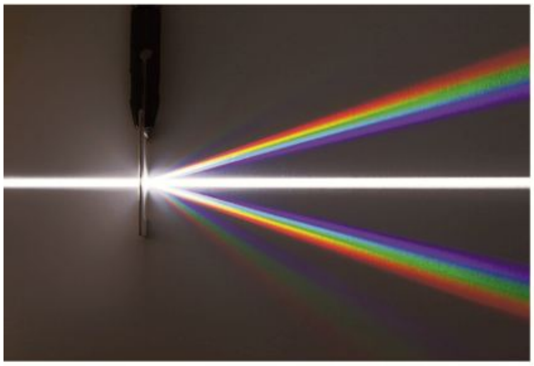


▲ Figure 5 Diffraction grating.

The spacing between the slits is small, which makes the angle θ large for a fixed wavelength of light and n. This means that we cannot use the small angle approximation for relating the wavelength to the position of the maxima as we did for a double-slit. Figure 6 shows a section of a diffraction grating in which three consecutive slits deviate the incident waves towards a maximum. The slits are so narrow and the screen is so far away from the grating that the angles made by the diffracted waves are virtually identical. Providing the path difference between waves coming from the same part of successive slits is an integral number of wavelengths, the waves will reach the screen in phase and give a maximum. In the case of a diffraction grating the distance d is taken as the length of both the transparent and opaque sections of the slit (they are taken to be of equal width). From figure 6 we see that the path difference between those waves coming from A and B and the path between waves coming B and C will each be $d\sin\theta$. This means



▲ Figure 6 Deriving the diffraction grating equation.



▲ Figure 7 Dispersing white light with a diffraction grating.

that $n\lambda = d\sin\theta$ (where n is the "order" of the maximum and is zero for the central maximum, 1 for the first maximum on each side of the centre, etc.).

We can see that the angular positions of the interference maxima depend on the grating spacing, d. The shape of the diffraction envelope, however, is determined by the width of the clear spaces (as for a single slit). The maxima must lie within the envelope of the single slit diffraction pattern if they are to be relatively intense; this sets an upper limit on how wide the transparent portions of the grating can be.

One of the uses of a diffraction grating is to disperse white light into its component colours: this is because different wavelengths produce maxima at different angles. Figure 7 shows that light of greater wavelength (for any given order) is deviated by a larger angle. This is in line with what we would predict from using $n\lambda = d\sin\theta$. Each successive visible spectrum repeats the order of the colours of the previous one but becomes less intense and more spread out.

Grating spacing and number of lines per mm

It is usual for the "number of lines per mm" or (N) to be quoted for a diffraction grating, rather than the spacing. N must be converted into d in order to use the diffraction grating equation $n\lambda = d\sin\theta$. This is straightforward because $d = \frac{1}{N}$ but N must first be converted to the number of lines per metre by multiplying by 1000 before taking the reciprocal.

Nature of science

Appropriate wavelengths for effective dispersion

For a diffraction grating to produce an observable pattern, the grating spacing must be comparable to the wavelength of the waves. The wavelength of visible light is between approximately 400–700 nm. A grating with 600 lines per mm has a spacing of approximately 2000 nm or four times the wavelength of green light. This spacing produces very clear images.

The wavelengths of low-energy X-rays is around 10^{-10} m or 0.1 nm. The 600 lines per millimetre optical grating will not produce any observable maximum ... therefore, what will diffract X-rays? The spacing of ions in crystals is of the same order of magnitude as X-rays and the regular lattice shape of a crystal can perform the same task with X-rays as a diffraction grating does with visible light. The diffracted patterns are now commonly detected with charge-coupled device (CCD) detectors.

Worked example

A diffraction grating having 600 lines per millimetre is illuminated with a parallel beam of monochromatic light, which is normal to the grating. This produces a second-order maximum which is observed at 42.5° to the straight-through direction. Calculate the wavelength of the light.

Solution

$N = 6.00 \times 10^5$ lines per metre.

$d = \frac{1}{N} = \frac{1}{6.00 \times 10^5} = 1.67 \times 10^{-6}$ m.

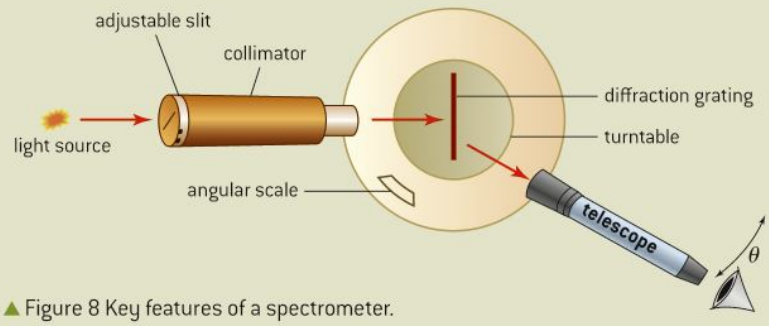
$n\lambda = d\sin\theta$ so $\lambda = \frac{d\sin\theta}{n} = \frac{1.67 \times 10^{-6} \times \sin 42.5}{2}$ (don't forget to have your calculator in degrees)

$\lambda = 5.64 \times 10^{-7} = 564$ nm (all data is given to three significant figures so the answer should also be to this precision).

Nature of science

Using a spectrometer

A spectrometer is a useful, if expensive, piece of laboratory equipment that allows the wavelengths of light emitted by sources to be analysed. The instrument consists of a collimator, turntable and telescope. A diffraction grating or other light disperser is placed on the turntable, which is carefully levelled. The collimator uses lenses to produce a parallel beam of light from a source – this light beam is then incident on the grating that disperses it. The end of the collimator furthest from the grating has a vertical adjustable slit that serves as the source – when we talk about spectral lines we infer that a spectrometer is being used because the "lines" are the dispersed images of the collimator slit. Before measurements are made the telescope is focused on the collimator slit – this means that the dispersed light will also be in focus. It is usual for sources that are said to be "monochromatic" to actually emit a variety of different wavelengths. By using the spectrometer, the angle θ in the diffraction grating equation can be measured. This is usually done by measuring the angle between the two first-order images on either side of the straight-through position i.e. 2θ. With a knowledge of the grating spacing d, the wavelength λ can be determined.



▲ Figure 8 Key features of a spectrometer.

Interference by division of amplitude

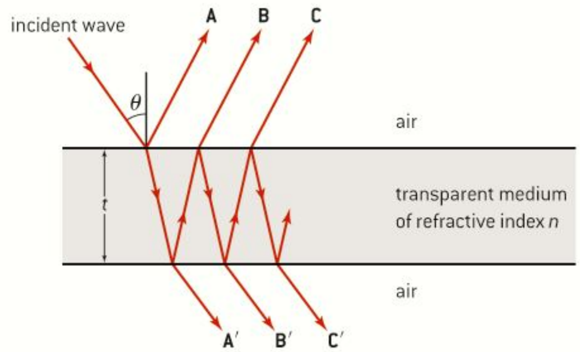
It was mentioned in the introduction to this sub-topic that there are two ways of providing coherent sources that are able to interfere. Young's double slit and multiple slits all derive their interfering waves by taking waves from different parts of the same wavefront. Because the interfering waves have all come from the same wavefront they will be in phase with each other. Wherever the waves meet they will interfere and a fringe pattern can be obtained anywhere in front of the sources (the slits). Since this interference can be found anywhere the fringes are said to be "non-localized".

Division of amplitude is a method of achieving interference using two waves that have come from the same point on a wavefront. Each wave has a portion of the amplitude of the original wave. In order to achieve interference by division of amplitude, the source of light must come from a much bigger source than the slit used for division of wavefront interference. The image produced will, however, be "localized" to one place instead of being found anywhere in front of the sources.

Thin film interference

Figure 9 shows a wave incident at an angle θ to the normal to the surface of a film of transparent material (such as low-density oil or detergent) having refractive index n. This diagram is not drawn to scale

– θ and t are both very small so that incident wave is effectively normal to the surface. The incident wave partially reflects at the top surface of the film and partially refracts into the film. This refracted wave, on reaching the lower surface of the film, again partially reflects (remaining in the film) and partially refracts into the air below the film. This process can occur several times for the same incident wave.



▲ Figure 9 Interference at parallel-sided thin film.

Waves reflected by the film *(you may be tested on this in IB Physics examinations)*

In this case **A** has been reflected from the top surface of the film and, because the reflection is at an optically denser medium, there is a phase change of π radian (equivalent to half a wavelength). The wave **B** travels an optical distance of $2tn$ before it refracts back into the air. Thus the optical path difference between **A** and **B** will be $2tn$. If there had been no phase change then this optical distance would equal $m\lambda$ for constructive interference. However, because of **A**'s phase change at the top surface the overall effect will be destructive interference.

Thus for the light reflecting from the film when

$$2tn = m\lambda$$

there will be destructive interference and when

$$2tn = (m + \tfrac{1}{2})\lambda$$

there will be constructive interference.

Waves transmitted through the film (you will **not** be tested on this in IB Physics examinations)

If the film is thin and θ is small the (geometrical) path difference between waves that have passed through the film (i.e. between **A'** and **B'** or between **B'** and **C'**) will be very nearly $2t$, in other words **B'** travels an extra $2t$ compared with **A'**. Because the waves are travelling in a material of refractive index n, the waves will slow down and the wavelength becomes shorter. This means that, compared with travelling in air, they will take longer to pass through the film. This is equivalent to them travelling through a thicker film at their normal speed. The optical path difference will, therefore, be $2tn$. If this distance is equal to $m\lambda$ where $m = 0, 1, 2$ etc. (we are using m to avoid confusion with the refractive index n) then there will be constructive interference. If the optical path difference is equal to an odd number of half wavelengths, $(m + \frac{1}{2})\lambda$, then there will be destructive interference.

Nature of science

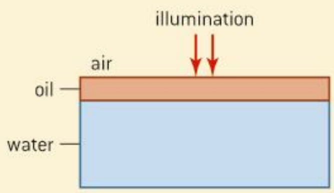
Coating of lenses

When light is incident on a lens some of it will be reflected and some transmitted. The reflected light is effectively wasted, reducing the intensity of any image formed (by the eye using corrective lenses, in a camera or in a telescope). By coating the lens with a transparent material of quarter of a wavelength thickness, the light reflected by the coating and the light reflected by the lens can be made to interfere destructively and thus eliminate the reflection altogether. Magnesium fluoride is often used for the coating and is optically denser than air but less dense than glass (having a refractive index of 1.38). When it is used to coat a lens the waves reflecting from both the magnesium fluoride and the glass will undergo a phase change of π radian ... so the phase changes effectively cancel each other out. The optical path difference between the waves will be equal to the refractive index of magnesium fluoride multiplied by twice the thickness of the coating. For destructive interference the optical path difference is $2nt = \frac{\lambda}{2}$ so the thickness of the coating is given by $t = \frac{\lambda}{4n}$. With white light there will be a range of wavelengths and so it is impossible to match the thickness of the coating to all of these wavelengths. If the thickness is matched to green light, then red and blue will still be reflected giving the lens the appearance of being magenta (= red + blue). By using multiple layers of a material of low refractive index and one of higher refractive index, it is possible to reduce the amount of light reflected to as little as 0.1% for a chosen wavelength.

Worked example

a) Name the wave phenomenon that is responsible for the formation of regions of different colour when white light is reflected from a thin film of oil floating on water.

b) A film of oil of refractive index 1.45 floats on a layer of water of refractive index 1.33 and is illuminated by white light at normal incidence.



When viewed at near normal incidence a particular region of the film looks red, with an average wavelength of about 650 nm.

(i) Explain the significance of the refractive indices of oil and water with regard to observing the red colour.

(ii) Calculate the minimum film thickness.

Solution

a) Although there is reflection involved, the colours come about because of interference.

b) (i) n is the refractive index of the oil. Because the waves are travelling in oil, they move more slowly than they would do in air and so the effective path difference between the waves reflected at the air–oil interface and the waves at the oil–water interface is longer by a factor of 1.45. The wave reflected at the air–oil interface undergoes a phase change equivalent to $\frac{\lambda}{2}$ (oil is denser than air). The waves reflected at the oil–water interface undergo no phase change on reflection at the less dense medium. So for the bright constructive red interference $2tn = (m + \frac{1}{2})\lambda$

if $m = 0$,

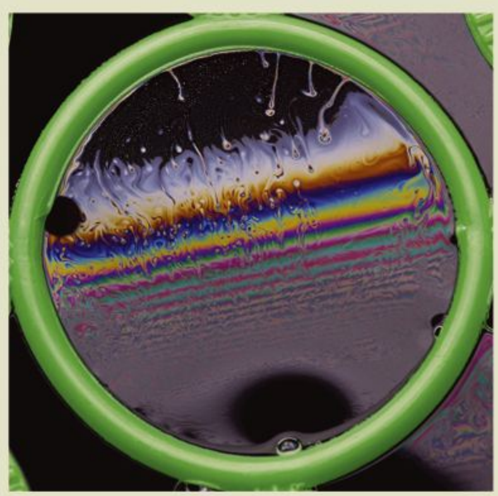
$2tn = \frac{\lambda}{2}$ or $\lambda = 4tn$

(ii) $t = \frac{650}{(4 \times 1.45)} = 110$ nm

Nature of science

Vertical soap films

Figure 10 shows the colours of light transmitted by a vertical soap film. Over a few seconds the film drains and becomes thinner at the top and thicker at the bottom. When the film is illuminated with white light, the reflected light appears as a series of horizontal coloured bands. The bands move downwards as the film drains and the top becomes thinner. The top of the film appears black just before the film breaks. It has now become too thin for there to be a path difference between the waves coming from the two surfaces of the film. The phase change that occurs for the light reflected by the surface of the film closest to the source means that, for all colours, there is cancellation and so no light can be seen.



▲ Figure 10 Thin film interference in a vertical soap film.

TOK

The aesthetics of physics

The colour in the soap film makes a beautiful image. Does physics need to rely upon the arts to be aesthetic? Herman Bondi the Anglo-Austrian mathematician and cosmologist implied that Einstein believed otherwise when he wrote:

> *"What I remember most clearly was that when I put down a suggestion that seemed to me cogent and reasonable, Einstein did not in the least contest this, but he only said, 'Oh, how ugly.' As soon as an equation seemed to him to be ugly, he really rather lost interest in it and could not understand why somebody else was willing to spend much time on it. He was quite convinced that beauty was a guiding principle in the search for important results in theoretical physics."*
>
> *— H. Bondi*

How does a mathematical equation convey beauty? Is the beauty in the mathematics itself or what the mathematics represents?

9.4 Resolution

Understanding

- The size of a diffracting aperture
- The resolution of simple monochromatic two-source systems

Applications and skills

- Solving problems involving the Rayleigh criterion for light emitted by two sources diffracted at a single slit
- Describing diffraction grating resolution

Equations

- Rayleigh's criterion: $\theta = 1.22\frac{\lambda}{b}$
- resolvance of a diffraction grating: $R = \frac{\lambda}{\Delta\lambda} = mN$

Nature of science

How far apart are atoms?

When we view objects, we are limited by the wavelength of the light used to make an observation – shorter wavelengths such as X-rays, gamma rays and fast-moving electrons will improve the resolution that is achievable. However, even using the shortest wavelengths obtainable does not allow us to locate the exact position of objects on the atomic and sub-atomic scale. The process of making an observation using these waves disturbs the system and increases the uncertainty with which we can locate an object. Heisenberg's uncertainty principle places a limit on how close we can be to finding the exact position of objects on the quantum scale.

Introduction

Resolution is the ability of an imaging system to be able to produce two separate distinguishable images of two separate objects. The imaging system could be an observer's eye, a camera, a radio telescope, etc. Whether objects can be resolved will depend on the wavelength coming from the objects, how close they are to each other and how far away they are from the observer.

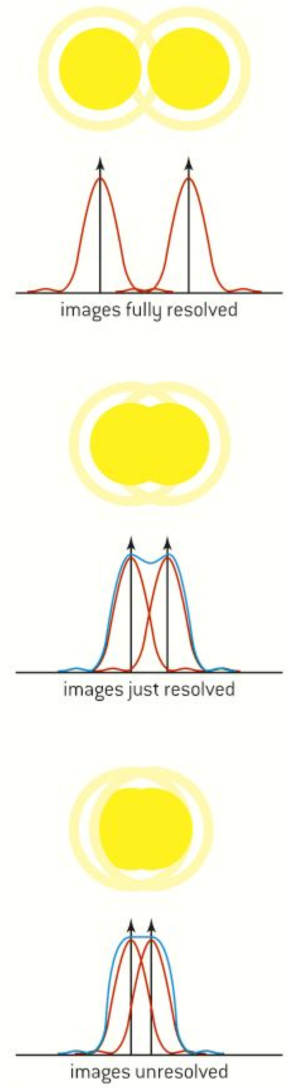
Diffraction and resolution

We have seen that when light passes through an aperture a diffraction pattern is formed. For an optical system the aperture could be the pupil of the observer's eye or the objective lens of a telescope. When there are two sources of light two diffraction patterns will be formed by the system.

How close can these patterns be for us to still recognize that there are two sources?

Two objects observed through an aperture will produce two diffracted images which may or may not overlap. In the late nineteenth century the English physicist, John William Strutt, 3rd Baron Rayleigh, proposed what is now known as the ***Rayleigh criterion*** for resolution of images. This states that **two sources are resolved if the principal maximum from one diffraction pattern is no closer than the first minimum of the other pattern.**

The limit to resolution is when the principal maximum of the diffraction pattern from one source lies on the first minimum diffraction pattern from the second source (and vice versa). Diffracted images further apart than this limit will be resolved and those closer will be unresolved.



▲ Figure 1 Diffraction intensity patterns of two objects viewed through a circular aperture.

Figure 1 shows the diffraction intensity patterns produced by two objects the same distance apart but viewed through a circular aperture from different distances. The variation of intensity with angle for each of the diffraction patterns is shown below the image. According to the Rayleigh criterion, the uppermost pair of images are fully resolved because the principal maximum of each diffraction pattern lies further from the other than the first minimum. The central images are just resolved since the principal maximum of one diffraction pattern is at the same position as the first minimum of the second diffraction pattern. The bottom images are unresolved as the principal maximum of one diffraction pattern lies closer to the second pattern than its first minimum.

Nature of science

Other criteria for resolution?

Rayleigh's criterion is not the only one used in optics. Many astronomers believe that they can resolve better than Rayleigh predicts. C M Sparrow developed another criterion for telescopes that leads to an angular separation at resolution about half that of Rayleigh. His criterion is that the two diffraction patterns when added together give a constant amplitude in the regions of the two central maxima.

Resolution equation

For single slits we saw in Sub-topic 9.2 that the first minimum occurs when the angle with the straight-through position is given by $\theta = \frac{\lambda}{a}$ where λ is the wavelength of the waves and b is the slit width. With a circular aperture the equation is modified by a factor of 1.22, but derivation of this factor is beyond the scope of the IB Physics course. So we have

$$\theta = 1.22\,\frac{\lambda}{a}$$

in this case a is the diameter of the circular aperture (or very commonly the diameter of the lens or mirror forming the image).

The pupil of the eye has a diameter of about 3 mm and, taking visible light to have a wavelength in the order of 6×10^{-7} m, the minimum angle of resolution for the eye is $\theta = \frac{1.22 \times 6 \times 10^{-7}}{3 \times 10^{-3}} \approx 2 \times 10^{-4}$ rad. Optical defects in the eye mean that this limit is probably a little small.

The primary mirror of the Hubble Space Telescope has a diameter of 2.4 m. For this telescope, the minimum angle of resolution is $\theta = \frac{1.22 \times 6 \times 10^{-7}}{2.4} \approx 3 \times 10^{-7}$ rad. This is a factor of about a thousand smaller than that achieved by the unaided eye – which means that the Hubble Space Telescope is much better at resolving images. As this telescope is in orbit above the atmosphere, it avoids the atmospheric distortion which degrades images achieved by Earth-based telescopes. These concepts are covered in more detail in Option C (Imaging).

Worked example

A student observes two distant point sources of light. The wavelength of each source is 550 nm. The angular separation between these two sources is 2.5×10^{-4} radians subtended at the pupil of a student's eye.

a) State the Rayleigh criterion for the two images on the retina to be just resolved.

b) Estimate the diameter of the circular aperture of the eye if the two images are just to be resolved.

Solution

a) The images will be just resolved when the diffraction pattern from one of the point sources has its central maximum at the same position as the first minimum of the diffraction pattern of the other point source.

b) $\theta = 1.22\,\frac{\lambda}{b} => b = 1.22\,\frac{\lambda}{\theta} = \frac{1.22 \times 550 \times 10^{-9}}{2.5 \times 10^{-4}}$

$= 2.7 \times 10^{-3}\text{ m} \approx 3\text{ mm}$

Note

- You may wish to support your answer by drawing the two intensity–angle curves if you think your answer may not be clear.
- It is common for students to miss out the word "diffraction" in their answers – this is crucial to score all marks.

Nature of science

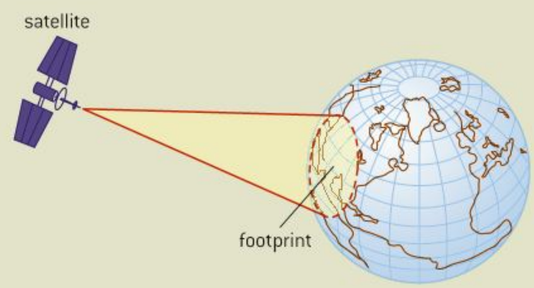
Diffraction and the satellite dish

When radiation is emitted from a transmitting satellite dish the waves diffract from the dish. The diameter of the dish behaves as an aperture. The angle θ made by the first minimum with the straight-through is related to diameter b by the equation

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

When the wavelength increases so does the angle through which the waves become diffracted. This means that the diffracted beam now covers a larger area. However, increasing the diameter of a dish narrows the beam and it covers a smaller area. In satellite communications the **footprint** is the portion of the Earth's surface over which the satellite dish delivers a specified amount of signal power. The footprint will be less than the region covered by the principal maximum because the signal will be too weak to be useful at the edges of the principal maximum. Figure 2 shows the footprint of a communications satellite.

- Small values of the dish diameter will give a large footprint but the intensity may be quite low since the energy is spread over a large area.
- The footprint of a satellite has social and political implications ranging from unwarranted observation to sharing television programmes.



▲ Figure 2 Satellite footprint.

Resolvance of diffraction gratings

We have seen that diffraction gratings are used to disperse light of different colours. Such gratings are usually used with spectrometers to allow the angular dispersion for each of the colours to be measured. When using a spectrometer, knowing the number of lines per millimetre on the diffraction grating and measuring the angles that the wavelengths of light are deviated through allows the wavelengths of the colours to be determined. We have seen that there is a limit to how close two objects can be before their diffraction images are indistinguishable – the same is true of the different wavelengths that can be resolved using a particular diffraction grating.

When we previously looked at interference patterns for multiple slits in Sub-topic 9.3, we saw that increasing the number of slits improves the sharpness of the maxima formed. Figure 3 shows that, when light of the same wavelength is viewed from the same distance x, the angular dispersion θ_D for the principal maximum with the double slit is larger than angular dispersion θ_G for the diffraction grating. A sharper principal maximum is one with less angular dispersion. With wider maxima there is more overlap of images from different sources and lower resolution.

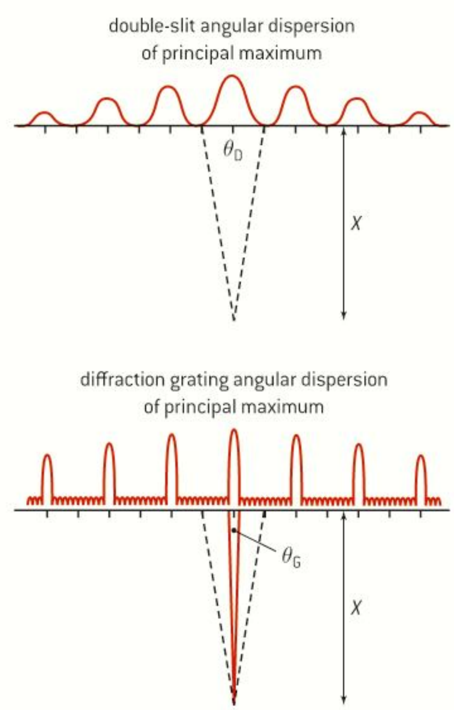
Using this argument we see that, when beams of light are incident on a diffraction grating, a wider beam covers more lines (a greater number of slits) and will produce sharper images and better resolution.

The resolvance R for a diffraction grating (or other device used to separate the wavelengths of light such a multiple slits) is defined as the ratio of the wavelength λ of the light to the smallest difference in wavelength that can be resolved by the grating $\Delta\lambda$.

The resolvance is also equal to Nm where N is the total number of slits illuminated by the incident beam and m is the order of the diffraction.

$$R = \frac{\lambda}{\Delta\lambda} = Nm$$

The larger the resolvance, the better a device can resolve.



▲ Figure 3 Angular dispersion of principal maximum for diffraction grating compared with a double-slit.

Nature of science

Resolution in a CCD

Charge-coupled devices (CCDs) were originally developed for use in computer memory devices but, today, appear in all digital cameras and smartphones. When you buy a camera or phone you will no doubt be interested in the number of pixels that it has. The pixel is a picture element and, for example, a 20 megapixel camera will have 2×10^7 pixels on its CCD. The resolution of a CCD depends on both the number of pixels and their size when compared to the projected image. The smaller the camera, the more convenient it is to carry, and (at present) cameras with pixels of dimensions as small as 2.7 μm × 2.7 μm are mass-produced. In general CCD images are resolved better with larger numbers of smaller pixels. Figure 4 shows two images of the words "IB Physics" – the top uses a small number of large pixels, while the lower one is far better resolved by using a large number of small pixels. The upper image is said to be "pixellated".

▲ Figure 4 Images with a small number of large pixels and a large number of small pixels.

Worked example

Two lines in the emission spectrum of sodium have wavelengths of 589.0 nm and 589.6 nm respectively. Calculate the number of lines per millimetre needed in a diffraction grating if the lines are to be resolved in the second-order spectrum with a beam of width 0.10 mm.

Solution

$R = \frac{\lambda}{\Delta\lambda} = \frac{589.0}{0.6}$ (both values are in nanometres so this factor cancels)

You would be equally justified in using 589.6 (or 589.3 – the mean value) in the numerator here.

$R = 981.7$ (no units)

Thus $981.7 = Nm = N \times 2$

$\therefore N = 498.8$ lines

This is in a beam of width 0.10×10^{-3} m so, in 1 mm, there needs to be $4988 \approx 5000$ lines.

Since diffraction gratings are not normally made with 4988 lines mm^{-1}, the sensible choice is to use one with 5000 lines mm^{-1}!

9.5 The Doppler effect

Understanding

- The Doppler effect for sound waves and light waves

Nature of science

From water waves to the expansion of the universe

In his 1842 paper, *Über das farbige Licht der Doppelsterne* (Concerning the coloured light of the double stars), Doppler used the analogy of the measurement of the frequency of water waves to reason that the effect that bears his name should apply to all waves. Three years later, Buys Ballot verified Doppler's hypothesis for sound using stationary and moving groups of trumpeters. During his lifetime Doppler's hypothesis had no practical application; one hundred years on it has far-reaching implications for cosmology, meteorology and medicine.

Applications and skills

- Sketching and interpreting the Doppler effect when there is relative motion between source and observer
- Describing situations where the Doppler effect can be utilized
- Solving problems involving the change in frequency or wavelength observed due to the Doppler effect to determine the velocity of the source/observer

Equations

Doppler equation

- for a moving source:

$$f' = f\left(\frac{v}{v \pm u_s}\right)$$

- for a moving observer:

$$f' = f\left(\frac{v \pm u_o}{v}\right)$$

- for electromagnetic radiation:

$$\frac{\Delta f}{f} = \frac{\Delta \lambda}{\lambda} \approx \frac{v}{c}$$

Introduction

When there is relative motion between a source of waves and an observer, the observed frequency of the waves is different to the frequency of the source of waves. The apparent change in pitch of an approaching vehicle engine and a sounding siren are common examples of this effect. The Doppler effect has wide-ranging implications in both atomic physics and astronomy.

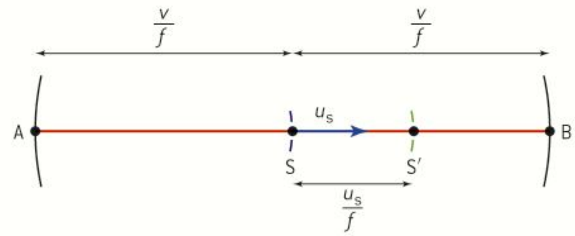
The Doppler effect with sound waves

Although we use general equations for this effect, we build up the equations under different conditions before combining them. In the following derivations (that need *not* be learned but will help you to understand the equations and how to answer questions on this topic) the letter s refers to the source of the waves (the object giving out the sound) and the letter o refers to the observer; f will always be the frequency of the source and f' the apparent frequency as measured by the observer.

Note

It is only the component of the wave in the source–observer direction that is used; perpendicular components of the motion do not alter the apparent frequency of the wave.

1. Moving source and stationary observer



▲ Figure 1 The Doppler effect for a moving source and stationary observer.

At time $t = 0$ the source is at position S and it emits a wave that travels outwards in all directions, with a velocity v as shown in figure 1. At time $t = T$ (i.e. one period later) the wave will have moved a distance equivalent to one wavelength or (using $v = f\lambda$) a distance $= \frac{v}{f}$ to reach positions A and B. When the source is moving to the right with a velocity u_s, in time T it will have travelled to S′, a distance $u_sT(= \frac{u_s}{f})$. To a stationary observer positioned at B it will appear that the previously emitted crest has reached B but the source that emitted it has moved forwards and now is at S′. Therefore, to the observer at B, the apparent wavelength (λ') is the distance S′B $= \frac{v}{f} - \frac{u_s}{f}$

$$\therefore \quad \lambda' = \frac{v - u_s}{f}$$

Thus, the wavelength appears to be squashed to a smaller value. This means that the observer at B will hear a sound of a higher frequency than would be heard from a stationary source – the sound waves travel at speed v (which is unchanged by the motion of the source) so

$$f' = \frac{v}{\lambda'} = f\left(\frac{v}{v - u_s}\right)$$

To an observer at A the wavelength would have appeared to be stretched to a longer value given by S′A meaning that

$$\lambda' = \frac{v + u_s}{f}$$

and the observed frequency would be lower than that of a stationary source, meaning that

$$f' = \frac{v}{\lambda'} = f\left(\frac{v}{v + u_s}\right)$$

The two equations for f' can be combined into a single equation

$$f' = f\left(\frac{v}{v \pm u_s}\right)$$

in which the $\pm$ sign is changed to $-$ for a source moving towards a stationary observer and to $+$ for a source moving away from a stationary observer.

2. Moving observer and stationary source

In this case the source remains stationary but the observer at B moves towards the source with a velocity u_o. The source emits crests at a frequency f but the observer, moving towards the source, encounters the crests more often, in other words at a higher frequency f'. Relative to the observer, the waves are travelling with a velocity $v + u_o$ and the frequency is f'. The wavelength of the crests does not appear to have changed and will be $\lambda = \frac{v}{f}$. The wave equation ($v = f\lambda$) applied by the observer becomes $v + u_o = f'\frac{v}{f}$ meaning that the frequency measured by the observer will be

$$f' = f\left(\frac{v + u_o}{v}\right)$$

When the observer moves away from the source the wave speed appears to be $v - u_o$ and so

$$f' = f\left(\frac{v - u_o}{v}\right)$$

Again the two equations for f' can be combined into a single equation

$$f' = f\left(\frac{v \pm u_o}{v}\right)$$

in which the $\pm$ sign is changed to $-$ for an observer moving away from a stationary source and to $+$ for an observer moving towards a stationary source.

Worked example

A stationary loudspeaker emits sound of frequency of 2.00 kHz. A student attaches the loudspeaker to a string and swings the loudspeaker in a horizontal circle at a speed of 15 m s^{-1}. The speed of sound in air is 330 m s^{-1}.

An observer listens to the sound at a close, but safe, distance from the student.

a) Explain why the sound heard by the observer changes regularly.

b) Determine the maximum frequency of the sound heard by the observer.

Solution

a) As the loudspeaker approaches the observer, the frequency appears higher than the stationary frequency because the apparent wavelength is shorter – meaning that the wavefronts are compressed. As it moves away from the observer, the frequency appears lower than the stationary frequency because the apparent wavelength is stretched. The overall effect is, therefore, a continuous rise and fall of pitch heard by the observer.

b) The maximum observed frequency occurs when the speaker is approaching the observer so:

$f' = f\left(\frac{v}{v - u_s}\right) = 2000\left(\frac{330}{330 - 15}\right) = 2100$ Hz (this is 2 s.f. precision in line with 15 m s^{-1})

The Doppler effect with light

The Doppler effect occurs not only with sound but also with light (and other electromagnetic waves); in which case the frequency and colour of the light differs from that emitted by the source. There is a significant difference in the application of Doppler effect for sound and light waves. Sound is a mechanical wave and requires a medium through which to travel; electromagnetic waves need no medium. Additionally, one of the assumptions or postulates of special relativity is that the velocity of light waves is constant in all inertial reference frames – this means that, when measured by an observer who is not accelerating, the observer will measure the speed of light to be 3.00×10^8 m s^{-1} irrespective of whether the observer moves towards the source or away from it and, therefore, it is impossible to distinguish between the motion of a source and an observer. This is not true for sound waves, as we have seen.

Although the Doppler effect equations for light and sound are derived on completely different principles, providing the speed of the source (or observer) is much less than the speed of light, the equations give approximately the correct results for light or sound. The quantities u_s and u_o in the Doppler equations for sound have no significance for light, and so we need to deal with the relative velocity v between the source and observer. Thus, in either of the equations

$$f' = f\left(\frac{v}{v \pm u_s}\right)$$

or

$$f' = f\left(\frac{v \pm u_o}{v}\right)$$

the wave speed is that of electromagnetic waves (c) and one of the velocities u_s or u_o is made zero while the other is replaced by the relative velocity v.

The first of these equations becomes

$$f' = f\left(\frac{c}{c+v}\right) = f\left(\frac{1}{1+\frac{v}{c}}\right) = f\left(1+\frac{v}{c}\right)^{-1}$$

This can be expanded using the binomial theorem to approximate to

$$f' = f\left(1-\frac{v}{c}\right) \approx f - f\frac{v}{c}$$

(ignoring all the terms after the second in the expansion).

This can be written as

$$f - f' \approx f\frac{v}{c}$$

or

$$\Delta f \approx f\frac{v}{c}$$

This equation is equivalent to

$$\Delta\lambda \approx \lambda\frac{v}{c}$$

Note

- Substituting values into the second equation will give the same resulting relationship – you may like to try this but remember you do not need to know any of these derivations.
- The equation is only valid when $c \gg v$ and so cannot usually be used with sound (where the wave speed is $\approx$ 300 m s^{-1} – unless the source or observer is moving much more slowly than this speed).

Worked example

As the Sun rotates, light waves received on Earth from opposite ends of a diameter show equal but opposite Doppler shifts. The speed of the edge of the Sun relative to the Earth is 1.90 km s^{-1}. What wavelength shift should be expected in the helium line having wavelength 587.5618 nm?

Solution

Using $\Delta f = f\frac{v}{c}$ this is equivalent to $\Delta\lambda \approx \lambda\frac{v}{c}$.

$1.90 \text{ km s}^{-1} = 1.90 \times 10^3 \text{ m s}^{-1}$

$$\Delta\lambda = 587.5618 \times \frac{1.90 \times 10^3}{3.00 \times 10^8} = 0.003\,7 \text{ nm}$$

Since one edge will approach the Earth – the shift from this edge will be a decrease in wavelength (blue shift) and that of the other edge (receding) will be an increase in wavelength (red shift).

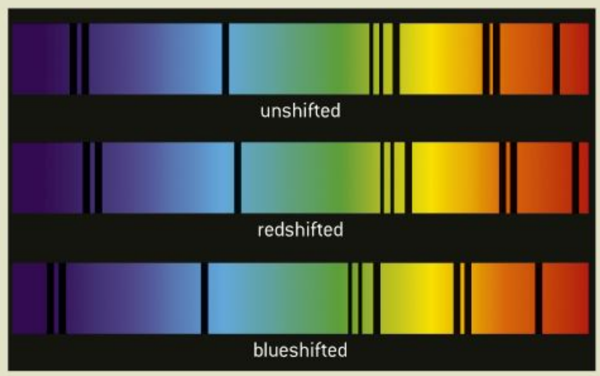
Nature of science

Applications of the Doppler effect

1. Astronomy

The Doppler effect is of particular interest in astronomy – it has been used to provide evidence about the motion of the objects throughout the universe.

The Doppler effect was originally studied in the visible part of the electromagnetic spectrum. Today, it is applied to the entire electromagnetic spectrum. Astronomers use Doppler shifts to calculate the speeds of stars and galaxies with respect to the Earth. When an astronomical body emits light there is a characteristic spectrum that corresponds to emissions from the elements in the body. By comparing the position of the spectral lines for these elements with those emitted by the same elements on the Earth, it can be seen that the lines remain in the same position relative to each other but shifted either to longer or shorter wavelengths (corresponding to lower or higher frequencies). Figure 2 shows an unshifted absorption spectrum imaged from an Earth-bound source together with the same spectral lines from distant astronomical objects. The middle image is red-shifted indicating that the source is moving away from the Earth. The lower image is blue-shifted showing the source to be local to the Earth and moving towards us.



▲ Figure 2 The Doppler shifted absorption spectra.

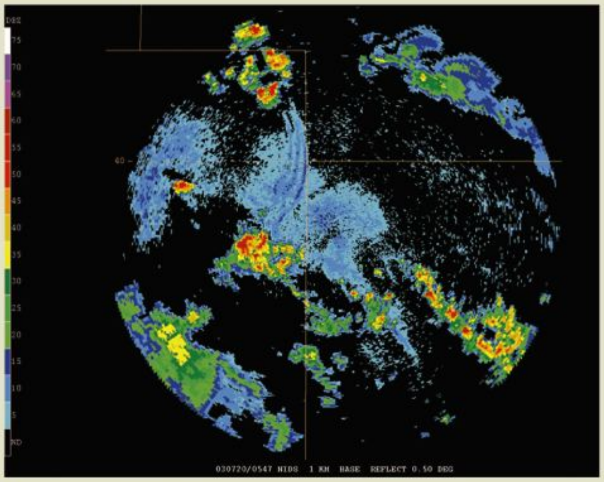
Frequency or wavelength shifts can occur for reasons other than relative motion. Electromagnetic waves moving close to an object with a very strong gravitational field can be red-shifted – this, unsurprisingly, is known as *gravitational red-shift* and is discussed further in Option A. The cosmological red-shift arises from the expansion of space following the Big Bang and is what we currently detect as the cosmic microwave background radiation (this is further discussed in Option D).

2. Radar

Radar is an acronym for "radio detection and ranging". Although it was developed for tracking aircraft during the Second World War, the technique has wide-ranging uses today, including:

- weather forecasting
- ground-penetrating radar for locating geological and archaeological artefacts
- providing bearings
- radar astronomy

- use in salvaging
- collision avoidance at sea and in the air.



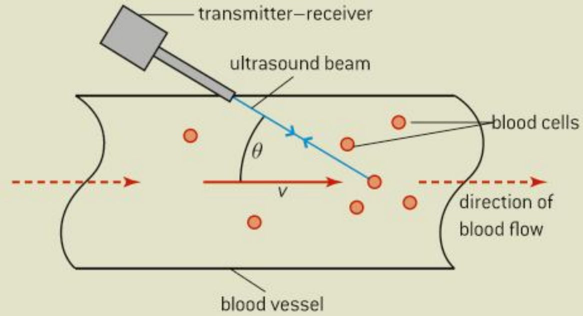
▲ Figure 3 Radar screen used in weather forecasting.

Radar astronomy differs from radio astronomy as it can only be used for Moon and planets close to the Earth. This depends on microwaves being transmitted to the object which then reflects them back to the Earth for detection. In order to use the Doppler equations, we must recognize that the moving object first of all behaves as a moving observer and then, when it reflects the microwaves, behaves as a moving source.

This means, for microwave sensing, the equation $\Delta f \approx f\frac{v}{c}$ is adapted to become $\Delta f \approx 2f\frac{v}{c}$.

3. Measuring the rate of blood flow

The Doppler effect can be used to measure the speed of blood flow in blood vessels in the body. In this case ultrasound (i.e. longitudinal mechanical waves of frequency above



▲ Figure 4 The Doppler effect used to measure the speed of blood cells.

approximately 20 kHz) is transmitted towards a blood vessel. The change in frequency of the beam reflected by a blood cell is detected by the receiver. The speed of sound and ultrasound is around 1500 m s^{-1} in body tissue so the equation $\Delta f \approx f\frac{v}{c}$ is appropriate for blood cells moving at speeds of little more than 1 m s^{-1}. As the blood does not flow in the direction of the transmitter–receiver, there needs to be a factor that will give the necessary component of the blood velocity in a direction parallel to the transmitter – receiver. As with radar, there will need to be a factor of two included in the equation – the shift is being caused by the echo from a moving reflector. As can be seen from figure 4, the equation for the rate of blood flow will be $\Delta f \approx 2f\frac{v\cos\theta}{c}$.

The Doppler effect flowmeter may be used in preference to "in-line" flowmeters because it is non-invasive and its presence does not affect the rate of flow of fluids. The fact that it will remain outside the vessel that is carrying the fluid means it will not suffer corrosion from contact with the fluid.

Worked example

Microwaves of wavelength 150 mm are transmitted from a source to an aircraft approaching the source. The shift in frequency of the reflected microwaves is 5.00 kHz. Calculate the speed of the aircraft relative to the source.

Solution

The data gives a wavelength and a change of frequency so we need to find the frequency of the microwaves.

Using

$c = f\lambda$ gives $f = \frac{c}{\lambda} = \frac{3.00 \times 10^8}{150 \times 10^{-3}} = 2.00 \times 10^9$ Hz.

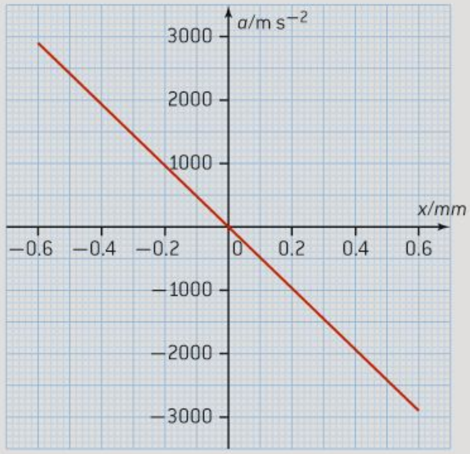
Using the Doppler shift equation for radar

$\Delta f \approx 2f\frac{v}{c} => v \approx \frac{\Delta f c}{2f} = \frac{5.00 \times 10^3 \times 3.00 \times 10^8}{2 \times 2.00 \times 10^9}$
$= 375$ m s^{-1}

Questions

1 *(IB)*

The variation with displacement x of the acceleration a of a vibrating object is shown below.



a) State and explain **two** reasons why the graph indicates that the object is executing simple harmonic motion.

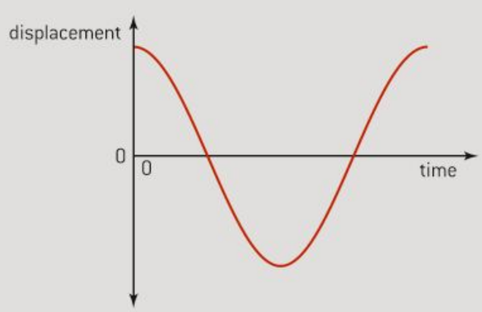
b) Use data from the graph to show that the frequency of oscillation is 350 Hz.

c) State the amplitude of the vibrations.

(9 marks)

2 *(IB)*

a) A pendulum consists of a bob that is suspended from a rigid support by a light inextensible string. The pendulum bob is moved to one side and then released. The sketch graph shows how the displacement of the pendulum bob undergoing simple harmonic motion varies with time over one time period.



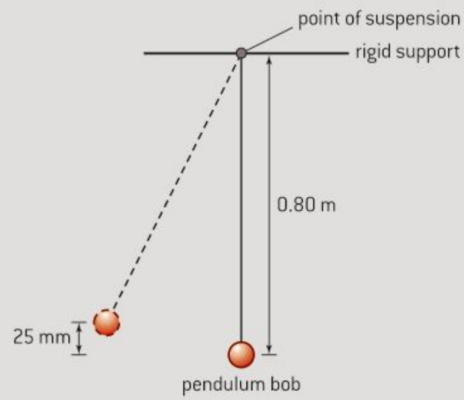
Copy the sketch graph and on it clearly label

(i) a point at which the acceleration of the pendulum bob is a maximum.

(ii) a point at which the speed of the pendulum bob is a maximum.

b) Explain why the magnitude of the tension in the string at the midpoint of the oscillation is greater than the weight of the pendulum bob.

c) The pendulum bob is moved to one side until its centre is 25 mm above its rest position and then released.



(i) Show that the speed of the pendulum bob at the midpoint of the oscillation is 0.70 m s^{-1}.

(ii) The mass of the pendulum bob is 0.057 kg. The centre of the pendulum bob is 0.80 m below the support. Calculate the magnitude of the tension in the string when the pendulum bob is vertically below the point of suspension.

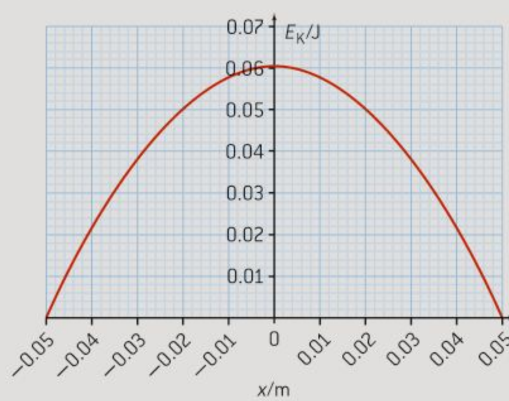
(10 marks)

3 *(IB)*

a) A particle of mass m attached to a light spring is executing simple harmonic motion in a **horizontal direction**.

State the condition (relating to the net force acting on the particle) that is necessary for it to execute simple harmonic motion.

b) The graph shows how the kinetic energy E_K of the particle in (a) varies with the displacement x of the particle from equilibrium.



(i) On a copy of the axes above, sketch a graph to show how the potential energy of the particle varies with the displacement *x*.

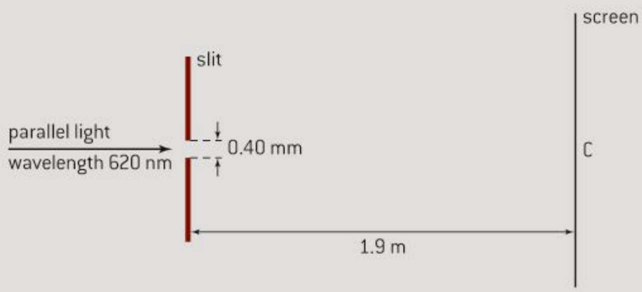
(ii) The mass of the particle is 0.30 kg. Use data from the graph to show that the frequency *f* of oscillation of the particle is 2.0 Hz.

(8 marks)

4 *(IB)*

a) Describe what is meant by the *diffraction of light*.

b) A parallel beam of monochromatic light from a laser is incident on a narrow slit. The diffracted light emerging from the slit is incident on a screen.



The centre of the diffraction pattern produced on the screen is at C. Sketch a graph to show how the intensity *I* of the light on the screen varies with the distance *d* from C.

c) The slit width is 0.40 mm and it is 1.9 m from the screen. The wavelength of the light is 620 nm. Determine the width of the central maximum on the screen.

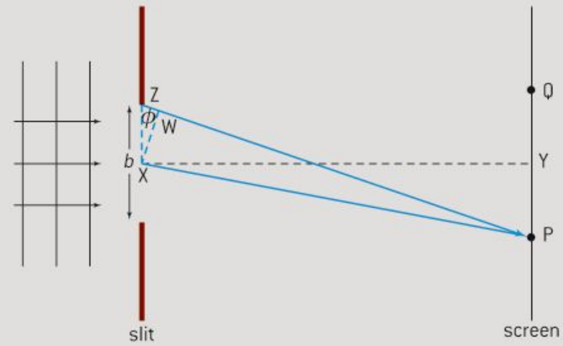
(8 marks)

5 *(IB)*

Plane wavefronts of monochromatic light are incident on a narrow, rectangular slit whose width b is comparable to the wavelength λ of the light. After passing through the slit, the light is brought to a focus on a screen.

The line XY, normal to the plane of the slit, is drawn from the centre of the slit to the screen. The points P and Q are the first points of minimum intensity as measured from point Y.

The diagram also shows two rays of light incident on the screen at point P. Ray ZP leaves one edge of the slit and ray XP leaves the centre of the slit.



The angle ϕ is small.

a) On a copy of the diagram, label the half angular width θ of the central maximum of the diffraction pattern.

b) State and explain an expression, in terms of λ, for the path difference ZW between the rays ZP and XP.

c) Deduce that the half angular width θ is given by the expression

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

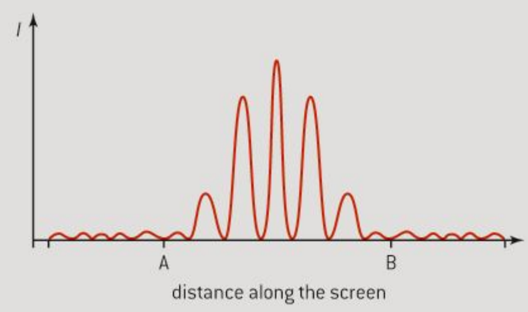
d) In a certain demonstration of single slit diffraction, $\lambda = 450$ nm, $b = 0.15$ mm and the screen is a long way from the slits.

Calculate the angular width of the central maximum of the diffraction pattern on the screen.

(8 marks)

6 *(IB)*

Monochromatic parallel light is incident on two slits of equal width and close together. After passing through the slits, the light is brought to a focus on a screen. The diagram below shows the intensity distribution of the light on the screen.



a) Light from the same source is incident on many slits of the same width as the widths of the slits above. On a copy of the diagram, draw a possible new intensity distribution of the light between the points A and B on the screen.

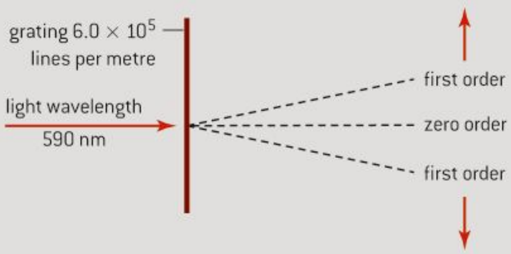
A parallel beam of light of wavelength 450 nm is incident at right angles on a diffraction grating. The slit spacing of the diffraction grating is 1.25×10^{-6} m.

b) Determine the angle between the central maximum and first order principal maximum formed by the grating.

(4 marks)

7 *(IB)*

Light of wavelength 590 nm is incident normally on a diffraction grating, as shown below.



The grating has 6.0×10^5 lines per metre.

a) Determine the **total** number of orders of diffracted light, including the zero order, that can be observed.

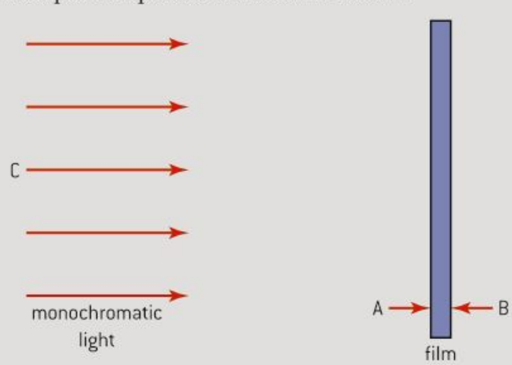
b) The incident light is replaced by a beam of light consisting of two wavelengths, 590 nm and 589 nm.

State **two** observable differences between a first-order spectrum and a second-order spectrum of the diffracted light.

(6 marks)

8 *(IB)*

Monochromatic light is incident on a thin film of transparent plastic as shown below.



The plastic film is in air.

Light is partially reflected at both surface A and surface B of the film.

a) State the phase change that occurs when light is reflected from

(i) surface A

(ii) surface B.

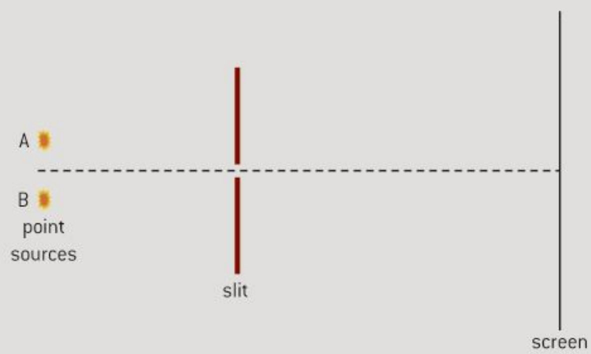
The light incident on the plastic has a wavelength of 620 nm. The refractive index of the plastic is 1.4.

b) Calculate the minimum thickness of the film needed for the light reflected from surface A and surface B to undergo destructive interference.

(5 marks)

9 *(IB)*

The two point sources A and B emit light of the same frequency. The light is incident on a rectangular narrow slit and, after passing through the slit, is brought to a focus on the screen.



a) B is covered. Sketch a graph to show how the intensity I of the light from A varies with distance along the screen. Label the curve you have drawn A.

b) B is now uncovered. The images of A and B on the screen are just resolved. Using your axes, sketch a graph to show how the intensity I of the light from B varies with distance along the screen. Label this curve B.

c) The bright star Sirius A is accompanied by a much fainter star, Sirius B. The mean distance of the stars from Earth is 8.1×10^{16} m. Under ideal atmospheric conditions, a telescope with an objective lens of diameter 25 cm can just resolve the stars as two separate images.

Assuming that the average wavelength emitted by the stars is 500 nm, estimate the apparent, linear separation of the two stars.

(6 marks)

10 *(IB)*

a) Explain what is meant by the term *resolvance* with regards to a diffraction grating.

b) A grating with a resolvance of 2000 is used in an attempt to separate the red lines in the spectra of hydrogen and deuterium.

(i) The incident beam has a width of 0.2 mm. For the first order spectrum, how many lines per mm must the grating have?

(ii) Explain whether or not the grating is capable of resolving the hydrogen lines which have wavelengths 656.3 nm and 656.1 nm.

(6 marks)

11 A source of sound approaches a stationary observer. The speed of the emitted sound and its wavelength, measured at the source, are v and λ respectively. Compare the wave speed and wavelength, as measured by the observer, with v and λ. Explain your answers. (4 marks)

12 The sound emitted by a car's horn has frequency f, as measured by the driver. An observer moves towards the stationary car at constant speed and measures the frequency of the sound to be f'.

a) Explain, using a diagram, any difference between f' and f.

b) The frequency f is 3.00×10^2 Hz. An observer moves towards the stationary car at a constant speed of 15.0 m s^{-1}. Calculate the observed frequency f' of the sound. The speed of sound in air is 3.30×10^2 m s^{-1}.

(5 marks)

13 *(IB)*

The wavelength diagram shown below represents three lines in the emission spectrum sample of calcium in a laboratory.

A distant star is known to be moving directly away from the Earth at a speed of 0.1c. The light emitted from the star contains the emission spectra of calcium. Copy the diagram and sketch the emission spectrum of the star as observed in the laboratory. Label the lines that correspond to A, B, and C with the letters A*, B*, and C*. Numerical values of the wavelengths are **not** required. (3 marks)