Topic 14 – **Superposition**

Content

- Principle of superposition
- Stationary waves
- Diffraction
- Two-source interference
- Single slit and multiple slit diffraction

Learning Outcomes

Candidates should be able to:

(a) explain and use the principle of superposition in simple applications

(b) show an understanding of the terms interference, coherence, phase difference and path difference

(c) show an understanding of experiments which demonstrate stationary waves using microwaves, stretched strings and air columns

(d) explain the formation of a stationary wave using a graphical method, and identify nodes and antinodes

(e) explain the meaning of the term diffraction

(f) show an understanding of experiments which demonstrate diffraction including the diffraction of water waves in a ripple tank with both a wide gap and a narrow gap

(g) show an understanding of experiments which demonstrate two-source interference using water waves, sound waves, light waves and microwaves

(h) show an understanding of the conditions required for two-source interference fringes to be observed

(i) recall and solve problems using the equation $\lambda = \frac{ax}{D}$ for double-slit interference

(j) recall and use the equation $\sin\theta = \frac{\lambda}{b}$ to locate the position of the first minima for single slit diffraction

(k) recall and use the Rayleigh criterion $\theta \approx \frac{\lambda}{b}$ for the resolving power of a single aperture

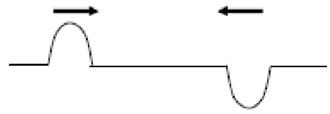
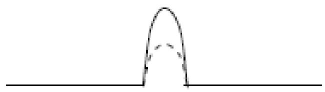
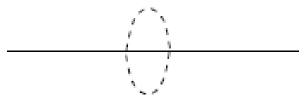
(l) recall and use the equation $d \sin\theta = n\lambda$ to locate the positions of the principal maxima produced by a diffraction grating

(m) describe the use of a diffraction grating to determine the wavelength of light (the structure and use of a spectrometer are not required).

(a)	explain and use the principle of superposition in simple applications.

1 PRINCIPLE OF SUPERPOSITION

In the previous topic, we only considered a single wave. Many interesting wave phenomena in nature are impossible to describe with a single progressive wave. In this topic, we will examine what happens when two (or more) waves pass through the same region of space at the same time.

	Equal waves	Opposite waves
(a) The two waves are some distance apart and approaching each other		
(b) They are about to meet		
(c) They are partially overlapping, the resultant is the vector sum of the two		
(d) They are exactly overlapping each other and maximum adding / cancellation is achieved		
(e) The waves are receding from each other.		

We see that after crossing each other, the waves travel as if nothing had happened and retain their original shape. Hence, pulses and waves pass through each other unaffected.

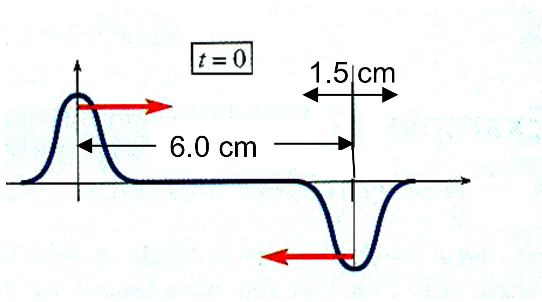
When the waves meet each other, they add up <u>vectorially</u> to give the resultant wave. This combination of two (or more) waves to form a resultant wave follows the Principle of Superposition.

The **Principle of Superposition** states that when two or more waves of <u>the same type</u> meet at a point at the same time, the displacement of the resultant wave is the <u>vector</u> sum of the displacements of the individual waves at that point at that time.

Note:
The two waves must be of the same type (sound, electromagnetic, etc).

Example 1:

Two identical waves travel at a speed of 2.0 m s^{-1} towards each other on a long cord. If the waves are 6.0 cm apart at t = 0 s, sketch the shape of the cord at t = 10, 15 and 20 ms.



In this topic, we are going to examine the application of this Principle of Superposition of waves to the following:

(i) formation of stationary waves
(ii) diffraction of waves
(iii) interference of waves

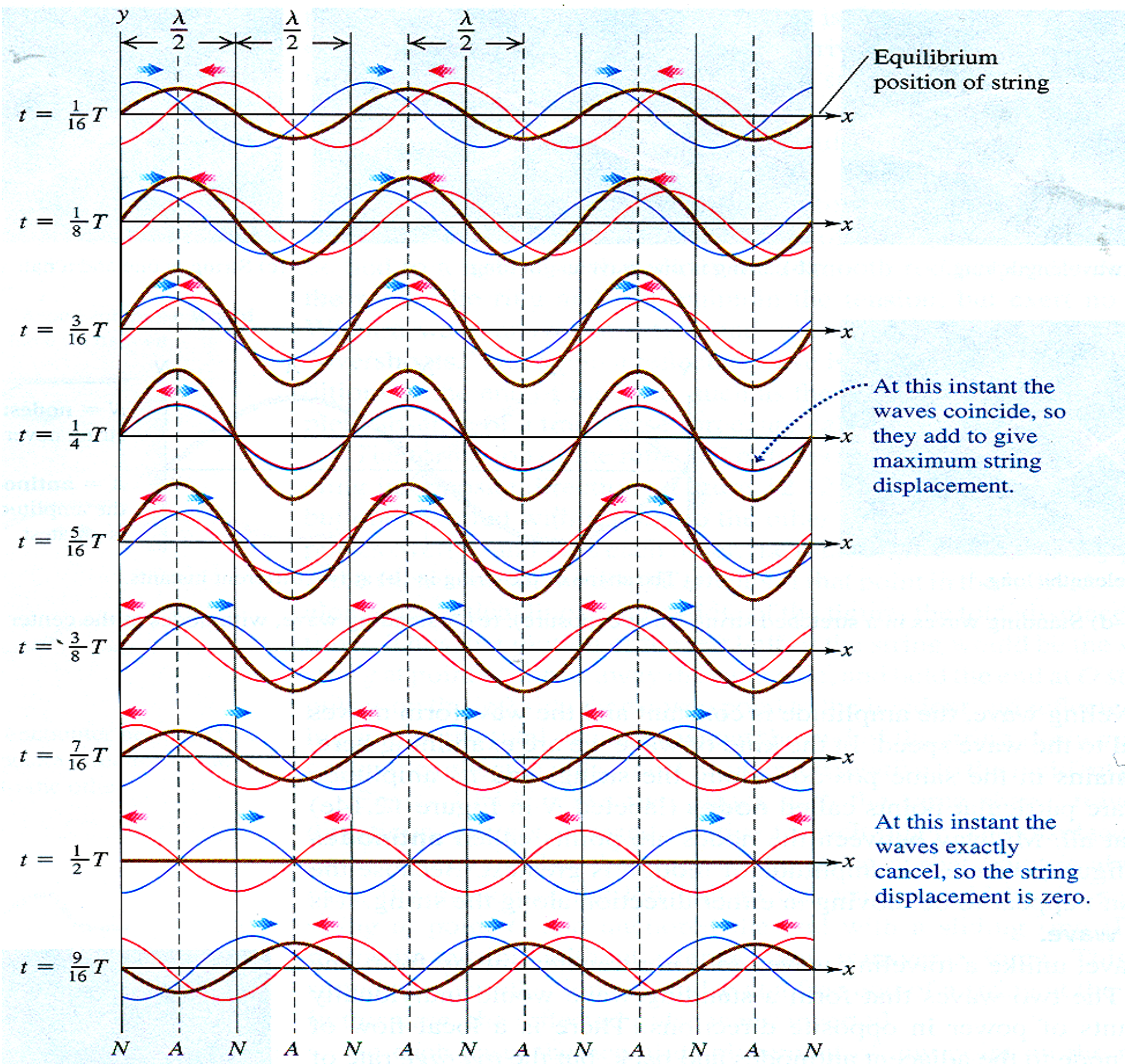
(d)	explain the formation of a stationary wave using a graphical method, and identify nodes and antinodes.

2 STATIONARY WAVES (ALSO CALLED STANDING WAVES)

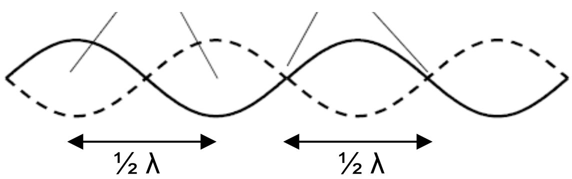
When two progressive waves of the same type of equal amplitude, equal frequency, equal speed travelling in opposite directions meet and undergo superposition with each other, a stationary wave is formed.

Vibrational energy of the wave is **not** transmitted from one point to another. Stationary wave has amplitudes of oscillation varying from maximum at antinodes to zero amplitude at nodes.

The red curve shows a wave traveling to the left while the blue curve shows a wave with the same speed, frequency and amplitude traveling to the right. The brown curve shows the resultant wave (stationary wave) obtained by applying the principle of superposition at instants $\frac{T}{16}$ apart.

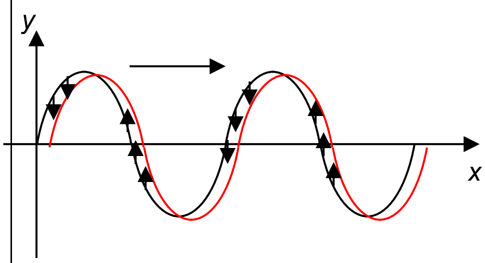
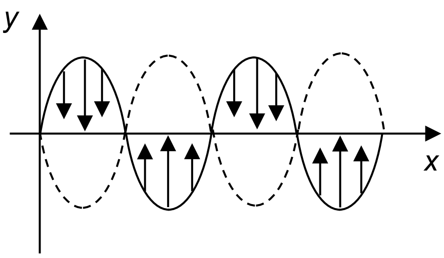


There are positions on the stationary wave where the two waves meet antiphase, resulting in the amplitude always zero. These are called **nodes (N).** Halfway between the nodes are the **antinodes (A)** where the waves meet in phase, resulting in maximum amplitudes. The distance between <u>two adjacent</u> nodes or <u>two adjacent</u> antinodes is ½ λ.



Note: Always draw stationary waves showing the *two* extremes of amplitude (one in solid, one in broken line.)

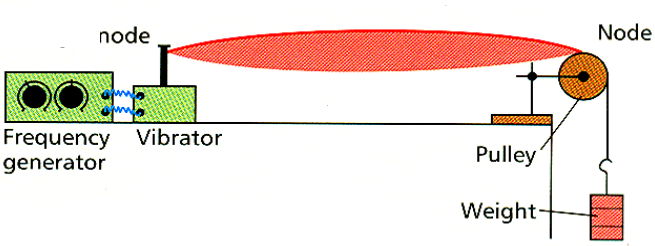
2.1 Comparing Progressive and Stationary Waves

	Progressive Wave	Stationary Wave
Amplitude	Same amplitude for all particles in the wave motion (if no loss of energy)	The amplitude of the oscillating particles varies from zero (at node) to a maximum (at anti-node)
Frequency	Particles of wave vibrate in s.h.m. with frequency of progressive wave.	Particles of wave vibrate in s.h.m. with frequency of stationary wave.
Wavelength	The wavelength of the wave is the shortest distance between any two successive points on a progressive wave which are vibrating in phase.	The wavelength is twice the distance between a pair of adjacent nodes (or a pair of adjacent anti-nodes).
Phase	All particles within one wavelength have different phases.	All particles within a loop vibrate in phase but have a phase difference of π rad with particles in the adjacent loop.
Wave Profile	Wave profile advances with the speed of the wave. 	Wave profile does not advance. 
Energy	Energy is transported in the direction of travel of the wave.	Energy is stored within the vibratory motion of the stationary wave.

(c) show an understanding of experiments which demonstrate stationary waves using microwaves, stretched strings and air columns.

2.2 Stationary Waves in Stretched Strings

The set up below shows how stationary waves can be produced in strings.

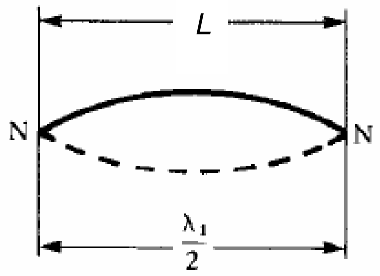
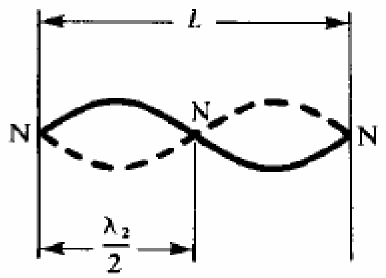
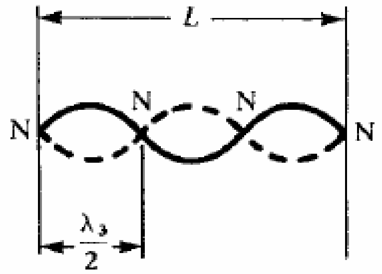


Fun fact: speed v of the waves on the string is given by $v = \sqrt{\frac{T}{\mu}}$ where T is the tension in the string and μ is the mass per unit length of the string.

A string of length L is held taut by connecting a weight on one end through a pulley and the other end to a vibrator. A frequency generator is used to vary the frequency of vibration.

The progressive wave produced by the vibrator moves towards the pulley with a constant velocity v. The wave is reflected at the pulley and travels backwards. The incident and reflected waves have the same speed, frequency and amplitude, moving in opposite direction and superpose, thus forming a stationary wave.

Since the same string and same weight are used throughout the experiment, the speed of the waves along the string remains the same in all cases. A change in frequency causes the wavelength to change and different modes of vibration arise.

	Fundamental mode/ 1st harmonic (lowest frequency obtainable from the oscillation)
	2nd harmonic (a harmonic is a note whose frequency is a whole number multiple of the fundamental frequency)
	3rd harmonic

Note: For *stationary wave* on string, integer number of loops must fit *exactly* into the length of the string, i.e. $L = n\dfrac{\lambda}{2}$ where n = 1, 2, 3, 4 …………

Example 2:

A wire 1.20 m long is fixed at each end under tension. A transverse wave of speed 300 m s^{-1} is propagated along the wire and forms a standing wave pattern. In a certain mode of vibration, it is found that the nodes are 0.40 m apart. What is the frequency of this standing wave? What lower frequencies are possible? [375 Hz, 250 Hz and 125 Hz]

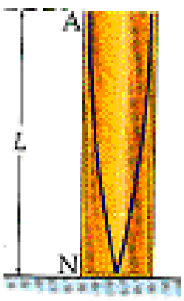
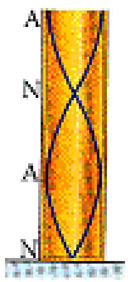
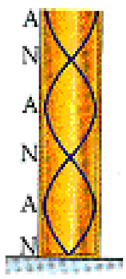
2.3 Stationary Waves in Air Columns

Stationary waves can be formed in both closed pipe (open at one end) and open pipe (open at both ends).

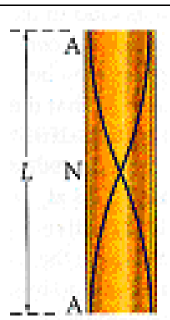
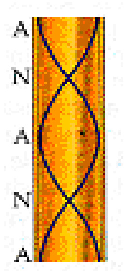
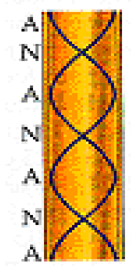
When air is blown into the open end of the pipe, a wave travels from the open end towards the closed end. The wave is reflected when it hits the wall of the closed end of the pipe. A stationary wave is formed when the incident and reflected waves superposed. Closed ends are nodes while open ends are antinodes.

Since the speed of sound remains the same in air (typically about 320 m s^{-1} to 340 m s^{-1}), different frequencies will produce different modes of vibration.

Closed Pipe

	Fundamental mode/ 1st harmonic
	3rd harmonic
	5th harmonic

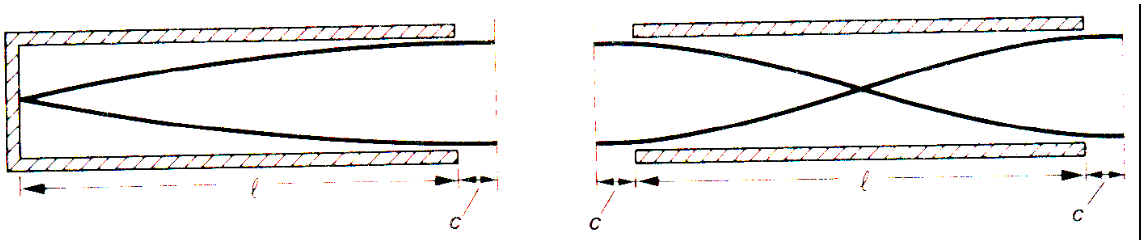
Open Pipe

	Fundamental mode/ 1st harmonic
	2nd harmonic
	3rd harmonic

Note:
1) In the above diagrams, the stationary wave is *represented* by drawing displacements of the particles at right angles to the axis of the pipe although the displacements of the particles are, in fact, along its axis since sound wave is a longitudinal wave.

2) Closed pipes only have odd harmonics.

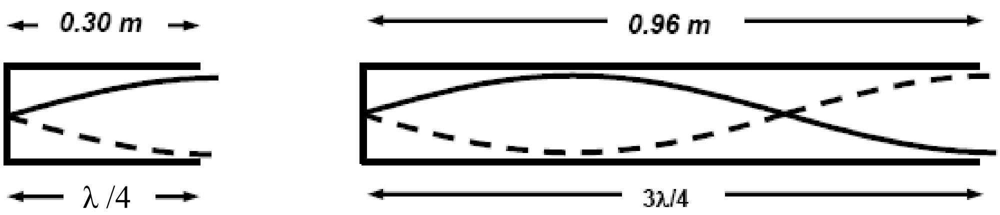
End Corrections



In practice, the antinode at the open end occurs slightly outside the pipe, hence, there is an end correction of length *c* that needs to be considered in the calculation of wavelength or frequency.

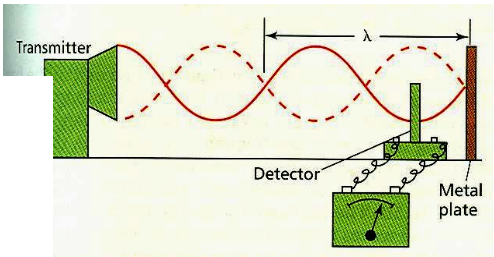
Example 3:

A source of sound of frequency 250 Hz is used with a resonant tube, closed at one end, to measure the speed of sound in air. Strong resonance is obtained at tube lengths of 0.30 m and 0.96 m. Find (a) the speed of sound, (b) the end correction of the tube



2.4 Stationary Waves in Microwaves

The transmitter emits an incident microwave which moves to the right. When it hits the metal plate, a reflected wave of the same speed, frequency, wavelength and amplitude is produced. The incident wave and the reflected wave move in opposite direction, undergo superposition and form the stationary wave.



Moving the microwave detector along the line between the microwave transmitter (wave source) and the metal plate (reflector) will enable alternate positions of maximum (antinodes) and minimum readings (nodes) to be found.

The distance between two adjacent nodes OR two adjacent antinodes corresponds to ½ λ.

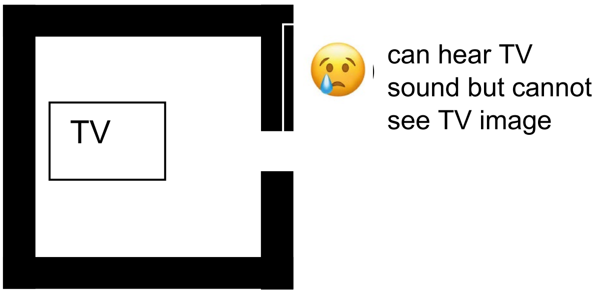
Example 4:

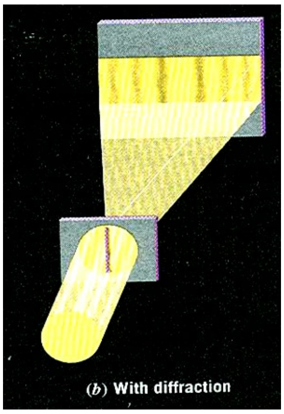
A sound source of frequency 2500 Hz is placed several meters from a plane reflecting wall in a large chamber containing a gas. A microphone, connected to a cathode ray oscilloscope, is used to detect nodes and antinodes formed along the normal from the source to the wall. The microphone is moved from one node through 20 antinodes to another node, a distance of 1.90 m. What is the speed of sound in the gas?

(e)	explain the meaning of the term diffraction.
(f)	show an understanding of experiments which demonstrate diffraction including the diffraction of water waves in a ripple tank with both a wide gap and a narrow gap.

3 DIFFRACTION

Everyone is used to the idea that sound spreads through an opening (e.g. door), which is why you can hear the sound of a TV which is inside the room (although you can't see the light from the TV). Light can also spread out or diffract as light is a wave. When waves encounter an obstacle or an opening, they will spread out after passing through an opening or around the edge of the obstacle. This effect which is peculiar to waves is known as diffraction.

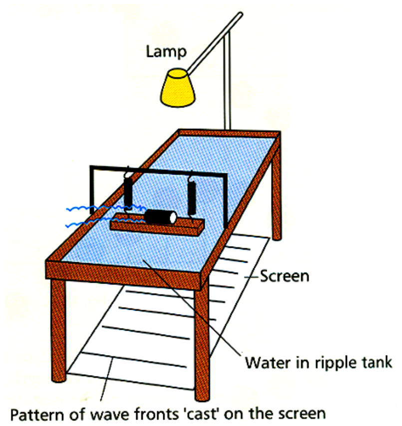


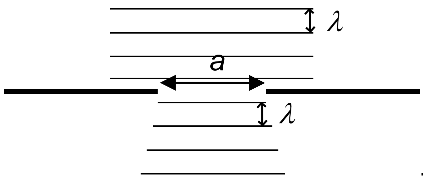


Diffraction is the <u>spreading</u> of waves at an edge of an obstacle or through a slit so that the waves do not travel in straight lines.

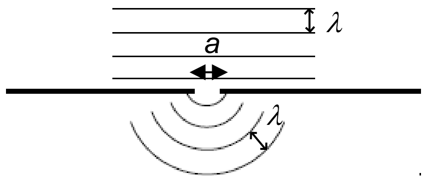
The spreading of the wavefronts becomes more pronounced when the size of the slit or aperture, a is of the same order as the wavelength of the wave λ, that is, $a \approx \lambda$, but is negligible when the width is large in comparison to the wavelength.

This can be demonstrated using a ripple tank with an oscillating horizontal narrow bar (set up as shown on the right). The patterns observed for different aperture sizes are as follows:





poor spreading

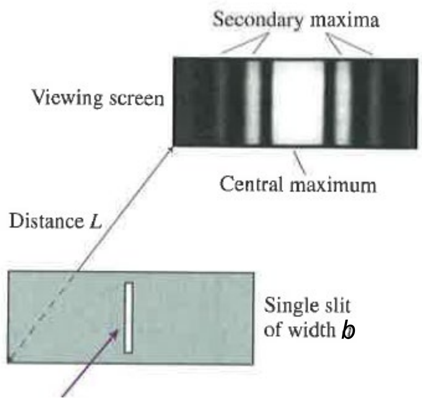


significant spreading

Note: When drawing diagram for <u>diffraction</u>, make sure <u>wavelength is the same</u> before and after passing through slit.

(j) recall and use the equation $\sin\theta = \dfrac{\lambda}{b}$ to locate the position of the first minima for single slit diffraction

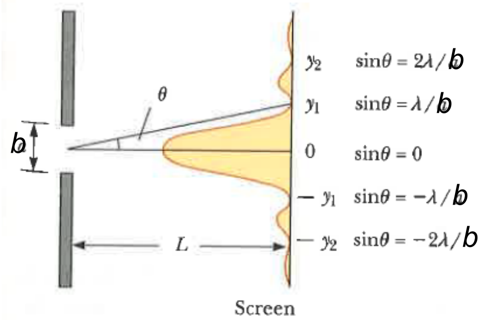
The figure on the right shows the experimental arrangement for observing the diffraction of light through a narrow slit of width b. A viewing screen is placed distance L (>> b) behind the slit



The light pattern on the screen consists of a central maximum flanked by a series of weaker secondary maxima and dark fringes.

It is also observed that the central maximum is significantly broader and brighter than the secondary maxima.

The figure on the right shows the intensity distribution.

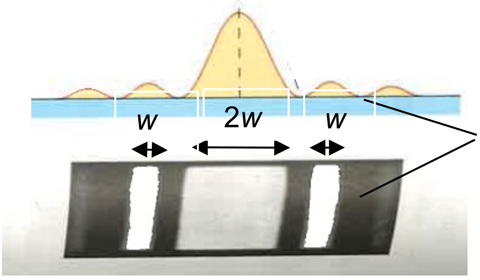


It can be shown in Annex A (p. 30) that points of <u>zero intensity (i.e. dark fringe)</u> occurs at angles given by the following equation

$$\boxed{\sin\theta = m\frac{\lambda}{b}}$$

where m = ±1, ±2, ±3, . . .

Notice that the central bright fringe is twice as wide as the secondary maxima (m > 1).



(k)	recall and use the Rayleigh criterion $\theta \approx \dfrac{\lambda}{b}$ for the resolving power of a single aperture

Resolution of two objects means the ability to see as distinct two objects that are distinct.

Light from each of the objects will diffract as it passes through the openings that admit light in cameras, telescopes, microscopes and the human eye. The light from each source will create its own diffraction pattern. The resulting diffraction pattern places a natural limit on the resolving power of these instruments and the human eye.

If these patterns are far apart the observer has no problem distinguishing the two sources as distinct, as shown in Figure A. We can say that the patterns are well resolved.

If on the other hand, the two sources are too close to each other, their diffraction patterns will overlap on the retina of the eye of the observer, as shown in Figure C. The observer cannot distinguish the two sources. We say that the patterns are not well resolved. The combined pattern looks very much like that from a single source.

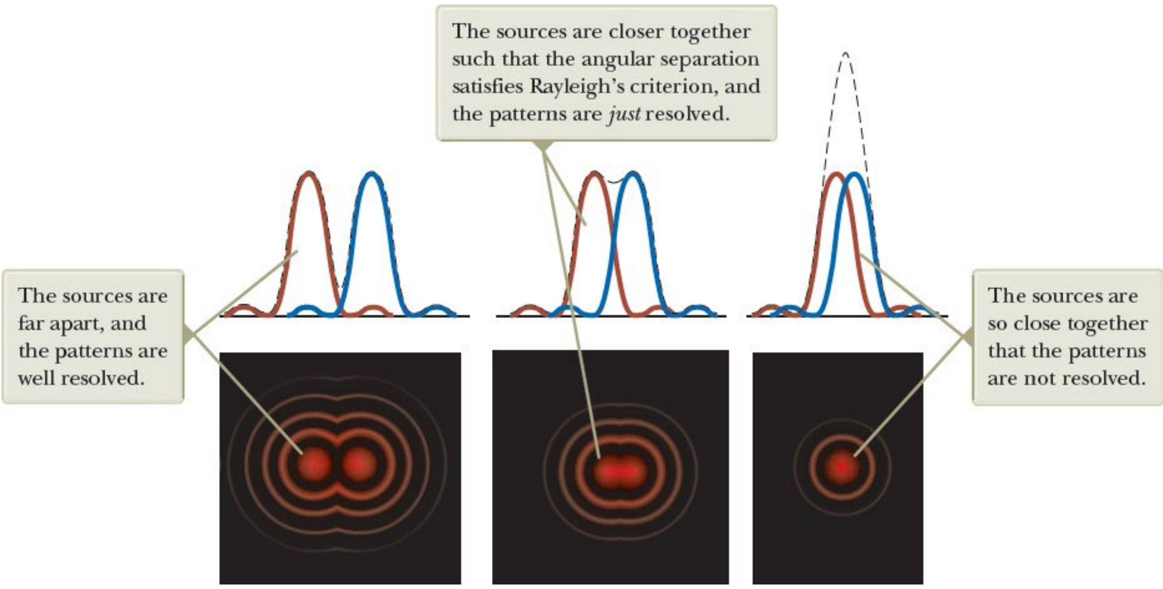
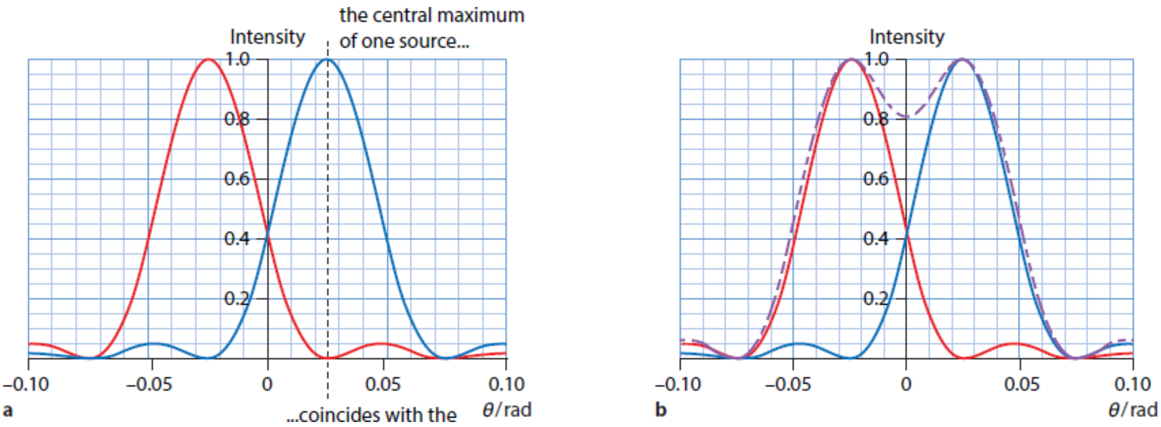


Figure A Figure B Figure C

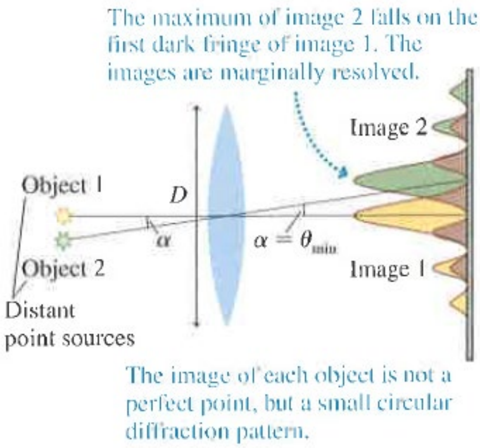
The limiting case, i.e. when the two sources can just be resolved, is when the first minimum of the diffraction pattern of one source coincides with the central maximum of the diffraction pattern of the other source, as shown in Figure B. This is known as the Rayleigh criterion.



The limiting case where resolution is thought to be just barely possible. The first minimum of one source coincides with the central maximum of the other. The combined pattern for the two sources shows a small dip in the middle.

Rayleigh Criterion:

For the two patterns to be <u>just distinguishable (or resolved)</u>, the <u>central</u> maximum of one must lie on the first minimum of the other.



Recall that the first minimum in a single-slit diffraction pattern occurs at an angle θ where

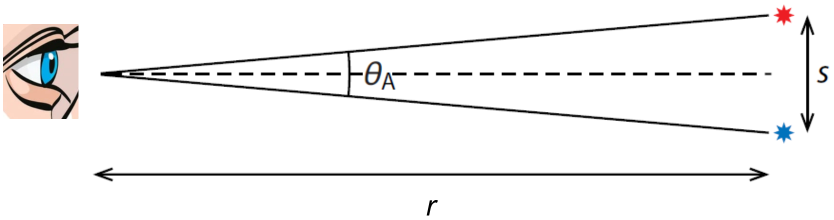
$$\sin\theta = \frac{\lambda}{b}$$

Because $\lambda \ll b$, sin θ is small, so that we can use the approximation that sin $\theta \approx \theta$, so

$$\boxed{\theta \approx \frac{\lambda}{b}}$$

where b represents the width of the slit and θ is the <u>minimum</u> angle for the two sources to be resolved

Suppose the two sources are a distance r from the observer. The angle that the separation of the sources subtends at the eye is called the angular separation θ_A of the two sources and $\theta_A = s / r$ since θ_A is small and is measured in radians.

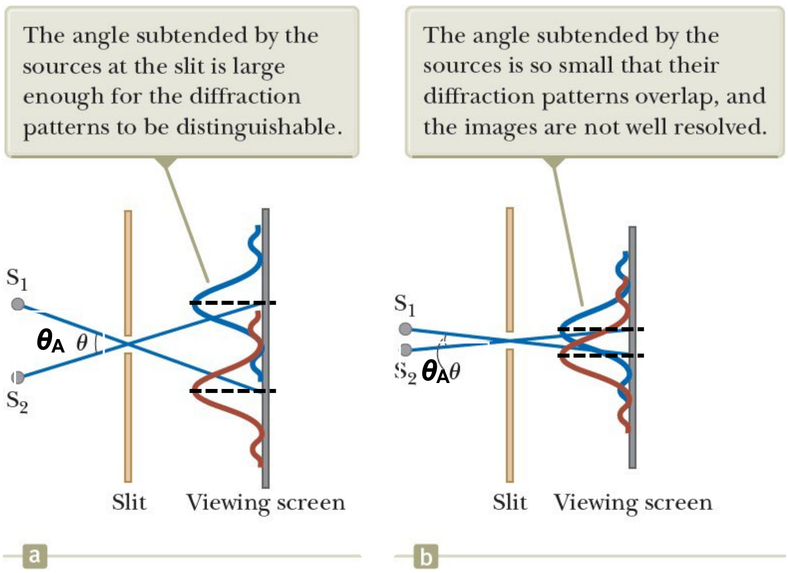


According to the Rayleigh criterion, resolution is just possible when this angular separation θ_A is equal to the angle of the first diffraction minimum θ.

Hence, we can resolve the two sources if the angular separation is greater than the minimum angle

$$\theta_A > \theta \qquad \text{or} \qquad s / r > \lambda / b$$

Thus, solving resolution problems involves the comparison of two angles: the angular separation of the two sources and the minimum angle due to the Rayleigh criterion.



$\theta_A > \theta$ $\qquad$ $\theta_A < \theta$

Example 5:

Parallel light of wavelength 590 nm is incident normally on a rectangular slit of width 0.60 mm. Light passing through the slit is incident on a screen that is 2.4 m from the plane of the slit.

(b)	show an understanding of the terms interference, coherence, phase difference, path difference

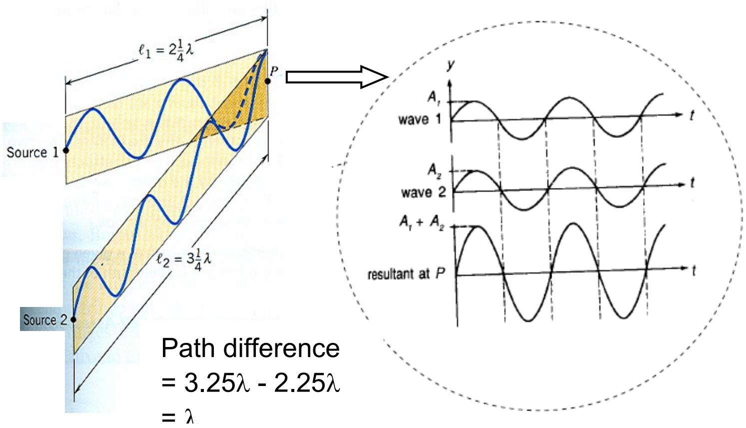
4 INTERFERENCE

In the region where coherent waves meet, superposition produces regular patterns of maxima (constructive interference) at some points and minima (destructive interference) at other points. This series of maxima and minima in the overlap region is known as the interference pattern.

- The two sources need to be **coherent** for interference to take place. This means that the waves produced by the sources have a **constant phase difference**.
- Interference patterns provide evidence for the wave nature of electromagnetic radiation.

4.1 Constructive Interference

When two coherent waves of same amplitude start in phase, and arrive at point P *in phase*, the superposition of waves will lead to a wave at P with displacement that is twice the displacements of either wave. We say constructive interference occurs at P.



Path difference
= 3.25λ - 2.25λ
= λ

Hence, when the waves start in phase, constructive interference occurs if

path difference = nλ

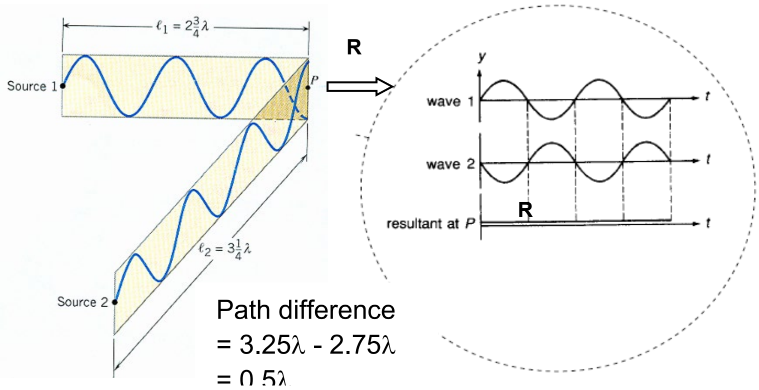
phase difference, ϕ = 2nπ

where *path difference* = $|l_2 - l_1|$, and *n* is an integer

Constructive interference	When 2 waves meet *in phase* at a point, the resultant displacement is the *sum* of the magnitude of individual displacements of the two waves.

4.2 Destructive interference

When two coherent waves of same amplitude start in phase, but arrive at point R *antiphase* (or 180° out of phase), the superposition of waves will lead to a wave at R with zero displacement. We say destructive interference occurs at R.



Hence, when the **waves start in phase**, destructive interference occurs if

path difference = (n + ½) λ

phase difference, ϕ = (2n+1)π

where *path difference* = | l_2 - l_1 |, and *n* is an integer

Destructive interference	When two waves meet *antiphase* (or out of phase by 180° or π rad) at a point, the resultant displacement is the *difference* of the magnitude of the individual displacements of the two waves. If the amplitudes of the two waves are the same, we obtain zero resultant displacement.

Note:

Path difference is the difference between the distance travelled by two waves meeting at a point. Once the path difference *x* is known, the phase difference ϕ can be calculated using the formula $\phi = \dfrac{2\pi x}{\lambda}$

Note: If the two **waves start antiphase** (180° out of phase), the above **formula swap** between constructive interference and destructive interference, that is,

constructive interference occurs if

path difference = (n + ½) λ

phase difference, ϕ = (2n+1)π

where *path difference* = | l_2 - l_1 |, and *n* is an integer

destructive interference if

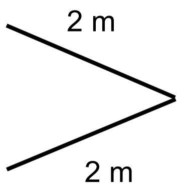
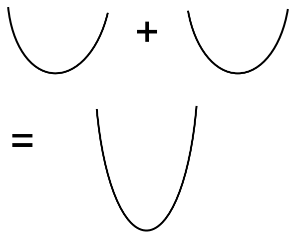
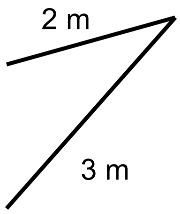
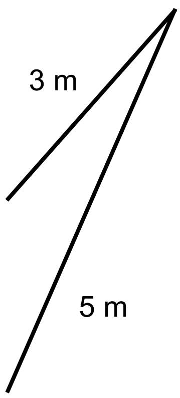
path difference = nλ

phase difference, ϕ = 2nπ

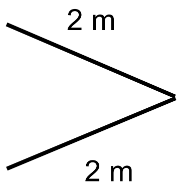
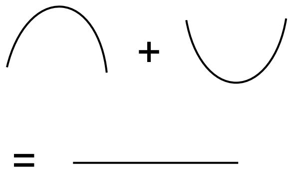
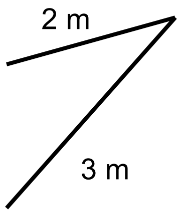
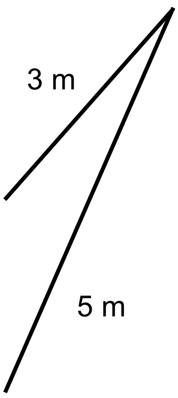
where *path difference* = | l_2 - l_1 |, and *n* is an integer

Exercise:

Let wavelength = 2 m. Assume waves <u>start in phase</u>.

Paths	Path difference (m)	Path difference (λ)	Constructive/Destructive Interference
	0	0	 constructive
			
			

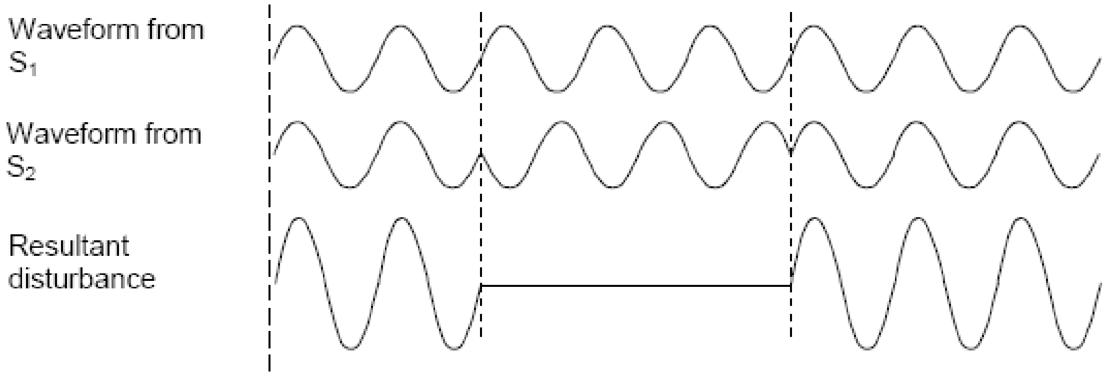
Let wavelength = 2 m. Assume waves start <u>anti-phase</u>.

Paths	Path difference (m)	Path difference (λ)	Constructive/Destructive Interference
	0	0	 destructive
			
			

Common Misconceptions:

No.	Wrong	Correct
1	• Constructive Interference occurs only when crests meet crests or troughs meet troughs. • Destructive interference occurs only when crests meet troughs.	• Constructive interference occurs whenever two waves meet in phase. • Destructive interference occurs whenever two waves meet antiphase or 180° out of phase.
2	Coherence means the 2 waves must be in phase	Coherence means the 2 waves have a constant phase difference
3	Same frequency implies coherence	Coherence implies same frequency

The third point is especially true for light waves because light sources do not emit continuous streams of light waves. Instead, they emit light in short pulses (photons). Between two consecutive pulses, there will be a random jump of phase, making the phase difference between two sources of exactly the same frequency inconsistent. The following diagram shows an example of how the phase of light waves can change between pulses.



(g) show an understanding of experiments which demonstrate two-source interference using water waves, sound waves, light and microwaves (h) show an understanding of the conditions required if two-source interference fringes are to be observed.

Two-Source Interference

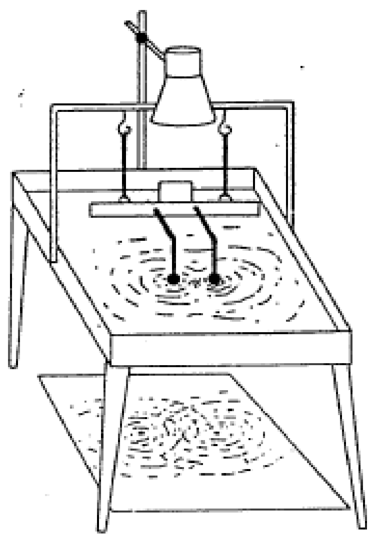
We can demonstrate two-source interference pattern using water waves, sound waves, microwaves and light.

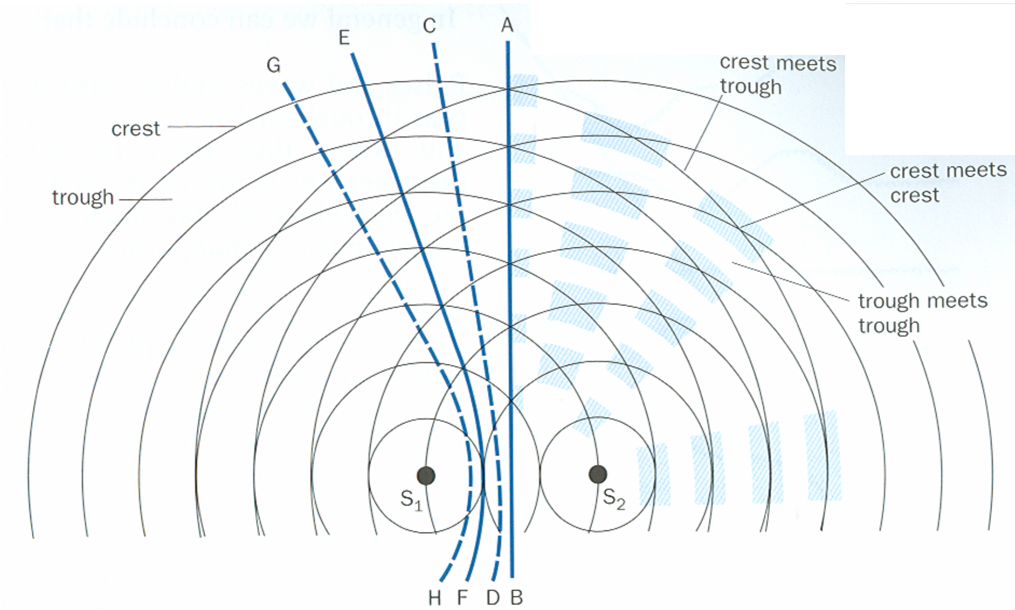
Conditions for 2-source interference fringes to be observable	(i) Waves must meet / overlap. (ii)Waves must be coherent i.e. **constant** phase difference. (iii)Waves have the equal or approximately equal amplitudes. (iv)Transverse waves must be either unpolarised or polarised in the same plane. (v) The slit separation *d* is of the same order as wavelength, λ.

4.3 Interference of water waves using ripple tank

- When two ball-ended dippers are attached to the vibrator of the ripple tank, two sets of circular waves (ripples) are sent out. These pass through one another as shown.

- The two dippers are coherent sources as they are attached to the same vibrator.



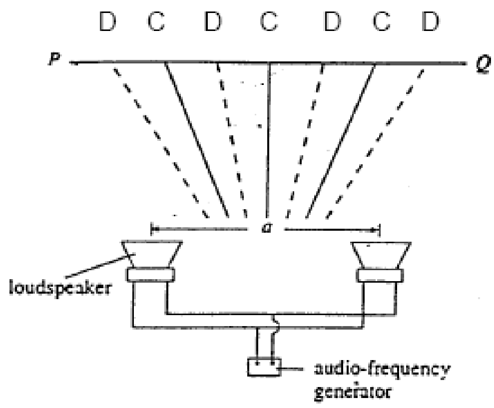


- Using the principle of superposition, where the two waves meet in phase, we obtain constructive interference as shown by lines of increased disturbance, e.g. lines AB and EF. These are called anti-nodal lines.

- In between these anti-nodal lines are the nodal lines along which the waves are anti-phase, e.g. lines CD and GH. Destructive interference takes place. If the amplitudes of these two waves are the same, we obtain zero resultant disturbance of the water surface.

4.4 <u>Interference of sound waves</u>

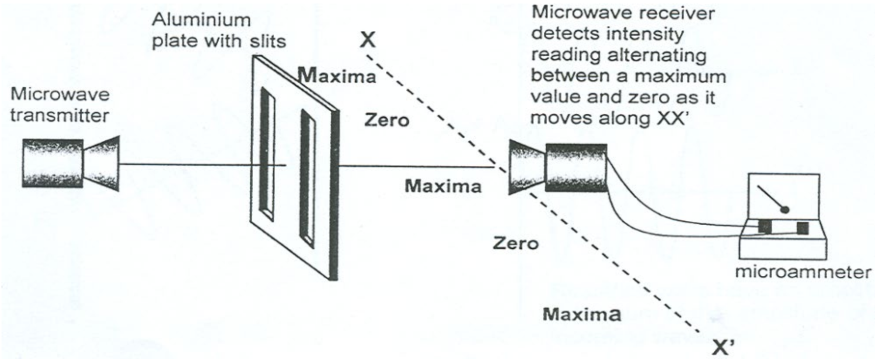
Two identical loudspeakers are connected to a audio-frequency generator. Sounds produced from the loudspeakers have the same amplitude, frequency and phase.



In the region where the sound waves overlap, superposition takes place and there are places of constructive interference (loud) and of destructive interference (soft). If one moves along the line PQ indicated in the above diagram, one would come across alternating loud and soft regions.

Constructive interference occurs at Points C when there is maximum loudness. Destructive interference occurs at Points D when there is minimum loudness.

4.5 Interference of microwaves



- The transmitter sent microwaves onto two vertical slits in an aluminium plate. The two slits act as coherent sources of microwaves which produce two sets of diffracted microwaves.

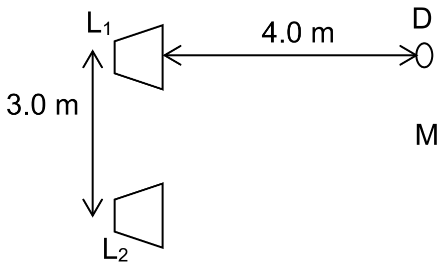
- When these two sets of diffracted microwaves interfere, constructive interference or destructive interference takes place, depending on the path difference between the waves.

- By moving the microwave receiver along the axis XX', the microammeter registers a current that varies from maximum to minimum alternately, showing regions of constructive interference and destructive interference respectively.

Note: For microwave, we do NOT use "bright / dark" or "loud / soft" to describe the interference pattern, we should use "high intensity" and "low intensity" instead.

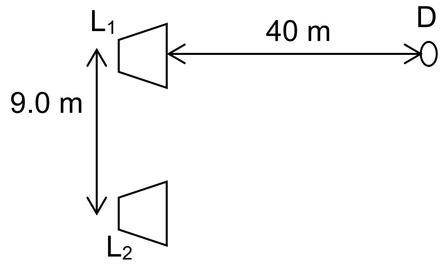
Example 6:

Two loudspeakers L_1 and L_2, connected from a common sound source and amplifier is set up 3.0 m apart. A detector D is placed at 4.0 m away directly in front of loudspeaker L_1. Given that the speed of sound = 340 m s^{-1} and the frequency of the sound source is 170 Hz, determine if the signal detected at D is a maxima or minima.



Example 7:

Two loudspeakers L_1 and L_2, connected from a common sound source and amplifier is set up 9.0 m apart. A detector D is placed at 40 m away directly in front of loudspeaker L_1. It is found that as the frequency of the sound source is gradually increased from zero, the signal detected passes through a series of maxima and minima. Explain this phenomenon and find the frequency at which the first maximum is heard.



[Speed of sound = 343 m s^{-1}]

Solution:

Path length L_2D = $\sqrt{(9.0)^2 + (40)^2}$ = 41 m

Path difference between two waves from L_1 and L_2 arriving at D = 41 – 40 = 1 m

The path difference is fixed at 1 m since the detector does not move.

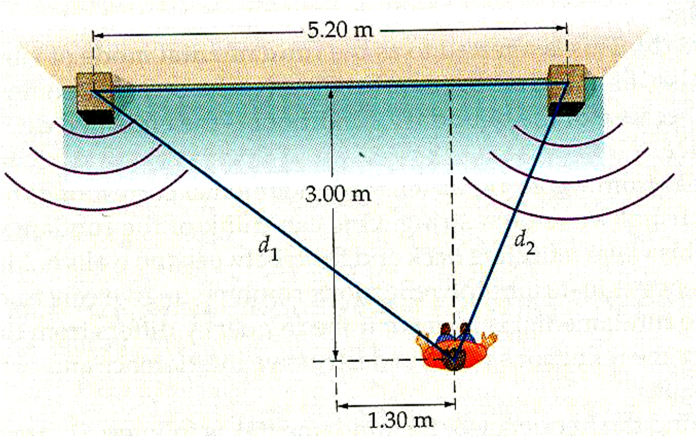
The wavelength of sound $\lambda = \dfrac{v}{f}$, so as the frequency f increases, the wavelength λ decreases.

Waves from L_1 and L_2 start in phase since it is from a common source.

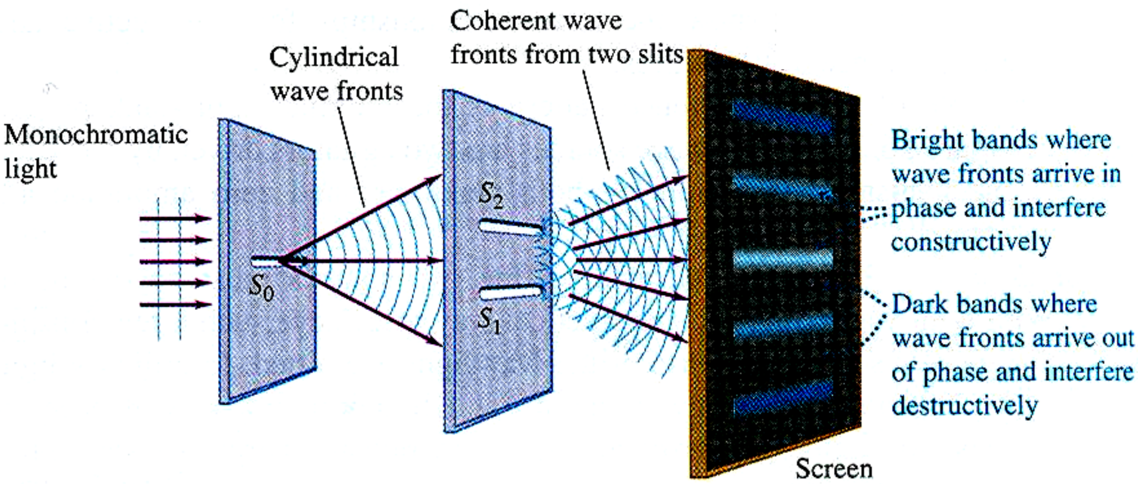
Example 8:

Two speakers, separated by a distance 5.20 m, emit sound of frequency 104 Hz in antiphase. A person listens from a location 3.00 m in front of the speakers and 1.30 m to one side of the centre line between the speakers.

(a) What does the person hear?
(b) What is next higher frequency that would produce constructive interference at the person's location?
[Speed of sound = 343 m s^{-1}]

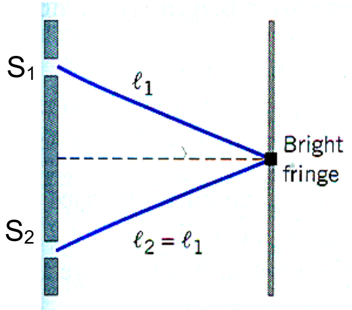


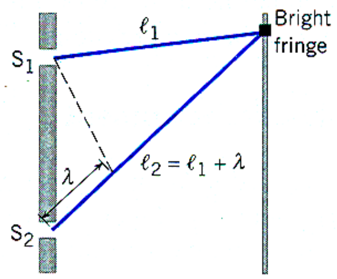
4.6 Interference of light waves (Young double-slit experiment)



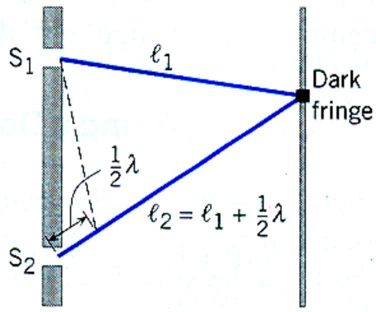
Monochromatic light sources provides light of a single wavelength. Slit S_0 acts as a single source of wave. Two closely spaced slits S_1 and S_2 act as coherent sources. On the screen, alternate bright and dark fringes are formed due to the superposition of coherent waves.

A bright fringe is due to constructive interference as waves superposed are in phase. The path difference of the waves arriving is = $n\,\lambda$





A dark fringe is due to destructive interference as waves superposed are antiphase. The path difference of the waves arriving is = (n + ½) λ



For interference of light to be observable, the following conditions need to be satisfied:

- Monochromatic light
- The combination of the single light source through single slit and then double slits is used instead two separate monochromatic light sources. The double slits act as coherent sources of waves.

Note: Two separate light sources are not coherent. This is because two monochromatic light sources do NOT emit coherent waves because there are abrupt phase shifts within each wave train due to the random nature of the atoms emitting the light. Light consists of short pulse of waves (photons), and although the photons have the same frequency, they are emitted at random from the sources and consequently have no constant phase difference between any two of them.

- The width of the slits must of the same order as the wavelength of light. This is to allow sufficient diffraction to occur so that the waves overlap to form the interference pattern.
- The distance between the two slits *a* must be small compared with the distance between the source and the screen *D*. **Typically, *a* ≈ 1 mm and *D* ≈ 3 m**.

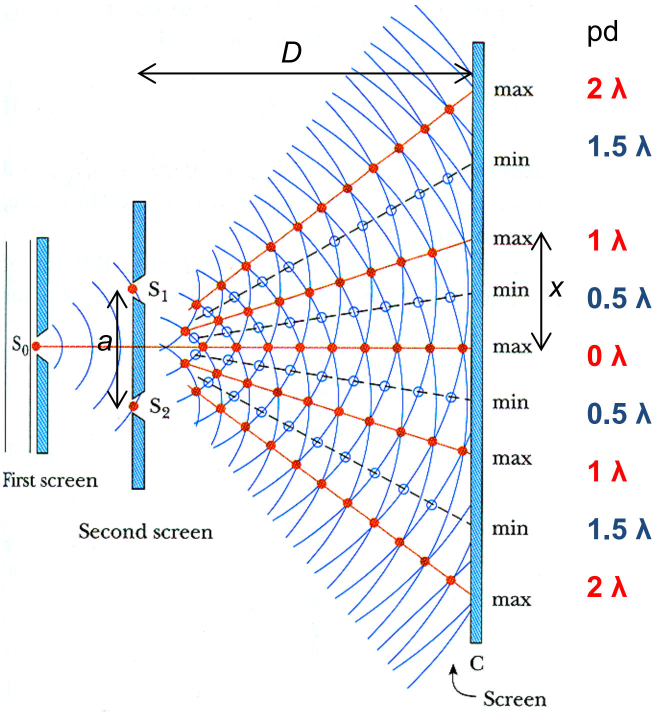
(i) recall and solve problems using the equation $\lambda = \frac{ax}{D}$ for double-slit interference using light.

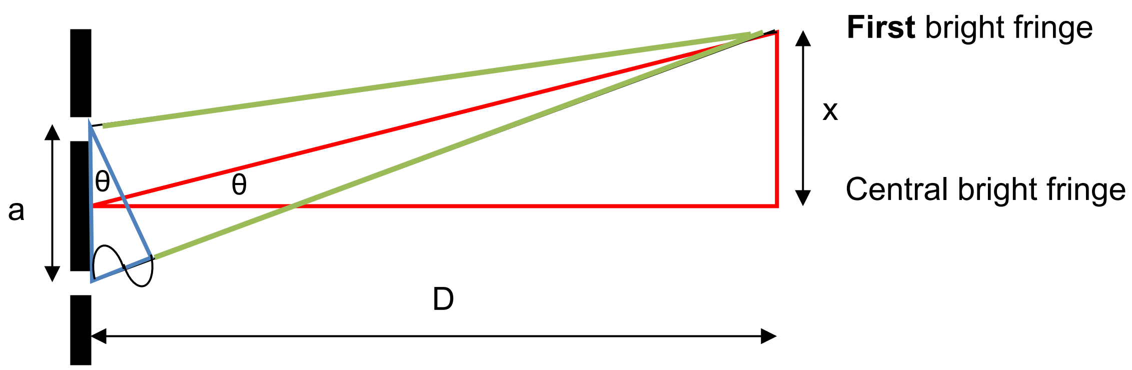
A relationship can be derived between the wavelength of the light λ, the distance of separation between the two slits *a*, the distance from the slit to the screen *D* and the fringe separation *x*. See Annex B (p.31) for derivation – not in syllabus.

$$\boxed{\lambda = \frac{xa}{D}}$$

This formula allows us to determine the wavelength of the waves by measuring the separation of the fringes in a double-slit experiment. However, it should be noted that this formula is only **valid for D >> a.**

tan θ ≈ sin θ $\rightarrow \frac{x}{D} = \frac{\lambda}{a}$



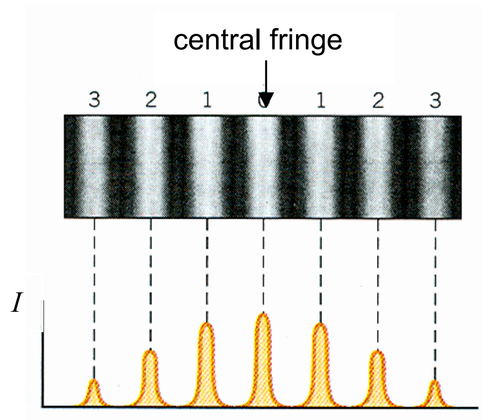


Intensity distribution of the interference fringes

If the waves from S_1 and S_2 are each of amplitude *A*, the intensity of each wave $I \propto A^2$.

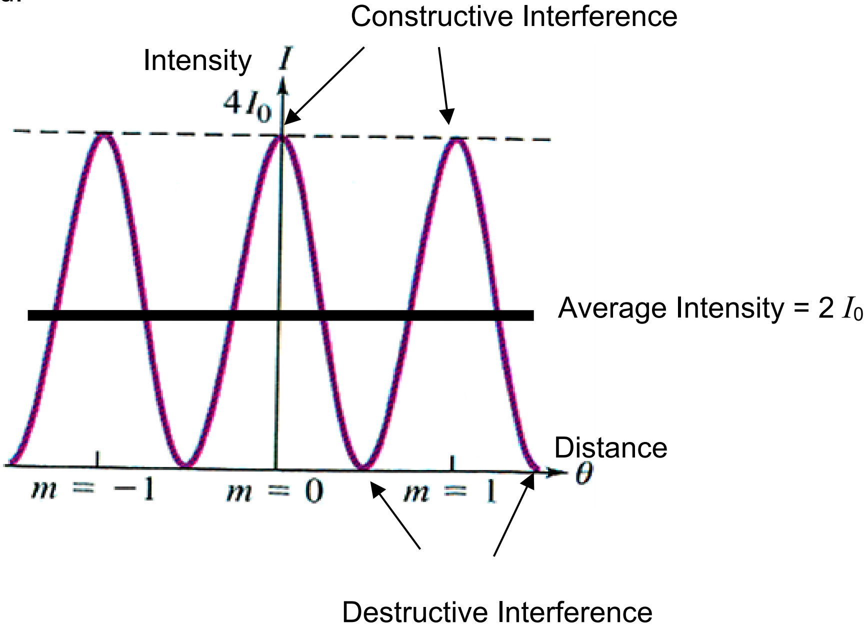
For a bright fringe on the screen, the total amplitude is *A* + *A* = 2*A*. The intensity $I \propto (2A)^2 \rightarrow I \propto 4A^2$

For a dark fringe on the screen, the total amplitude is zero. So intensity is zero.



As the width of the slits are comparable to the wavelength of light, diffraction occurs at the two slits and the intensity pattern obtained is one that is due to both the effect of interference and diffraction.

The effects of interference *redistribute* the energy into bright and dark fringes. Total energy is still conserved.



Example 9:

In a Young double-slit experiment, the separation between the second and sixth bright fringe is 2.0 mm and the wavelength of light used is 250 nm. If the distance from the slit to the screen is 1.2 m, find the separation between the slits.

Example 10:

In a Young double-slit experiment, a small detector measures intensity, *I* at the centre of the fringe pattern. If one of the two slits is now covered, what would be the measured intensity?

(l) recall and use the equation $d \sin\theta = n\lambda$ to locate the positions of the principal maxima produced by a diffraction grating

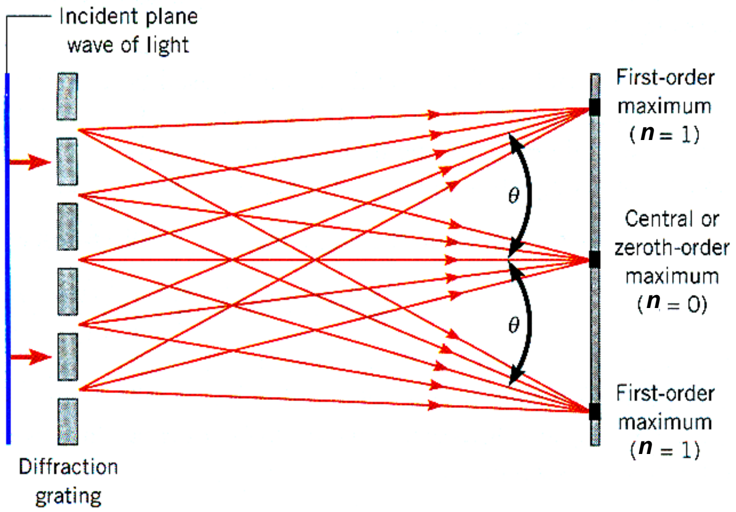
(m) describe the use of a diffraction grating to determine the wavelength of light (the structure and use of the spectrometer is not required).

5 DIFFRACTION GRATING

A diffraction grating consists of a large number of fine, equidistant and closely spaced parallel lines of equal width, ruled on glass. A typical grating has 300 lines per mm, but gratings with as many as 4,000 lines per mm have been made. The spaces between the lines serve as slits making the grating a multiple slit device. We can calculate slit separation d using $d = \dfrac{1}{\text{number of lines per unit length}}$.

When monochromatic light passes through a grating, diffraction occurs at each of the slits. The waves emerging from the slits are coherent (since they are from the same incident wavefront) and interference occurs in the region where the waves overlap.

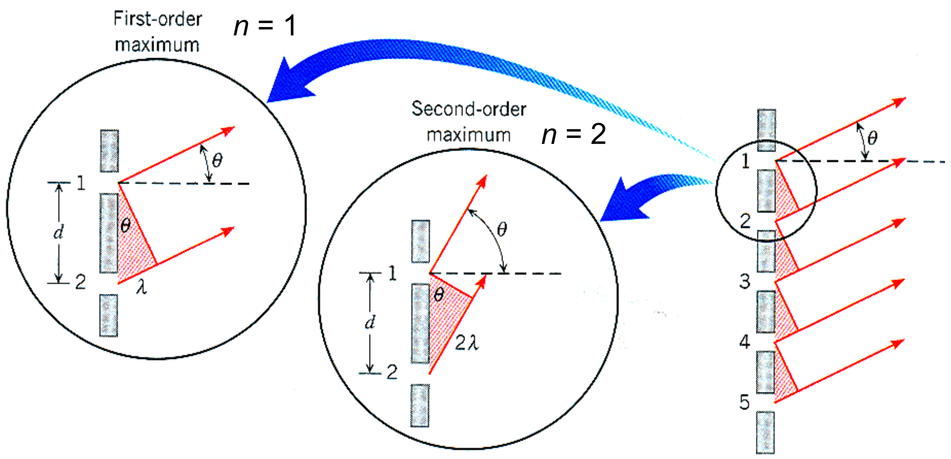
The diagram below illustrates how light travels to a screen from five slits in a grating to form the central bright fringe (n = 0) and the first-order bright fringe (n = 1) on either side. (Higher order bright fringes are not shown in the diagram.) Each bright fringe is located at an **angle θ relative to the central fringe.**



The bright fringes are due to constructive interference.

You may refer to Annex C (p. 32) to see how the fringes would be affected when the number of slits increase.

To reach the screen where the first-order maximum (n = 1) is located, light from slit 2 travels a distance of one wavelength further than light from slit 1. Similarly, light from slit 3 travels one wavelength further than light from slit 2 and so forth, as shown by the four coloured triangles on the right drawing below. The enlarged view of one of these triangles, as shown on the left drawing, shows that constructive interference occurs when $d \sin\theta = \lambda$.
Thus, the first-order (n = 1) is the result of constructive interference between light from adjacent slits where the path difference is one wavelength λ.



n = 0 (or θ = 0) is due to constructive interference of light with zero path difference between light from adjacent slits.

The second-order maximum forms when the extra distance travelled by light from adjacent slits is two wavelengths, as shown by the middle drawing so that $d \sin \theta = 2\lambda$.

Generalising, for the n^{th} order maximum, path difference $\boxed{d \sin \theta = n\lambda}$

Example 11:

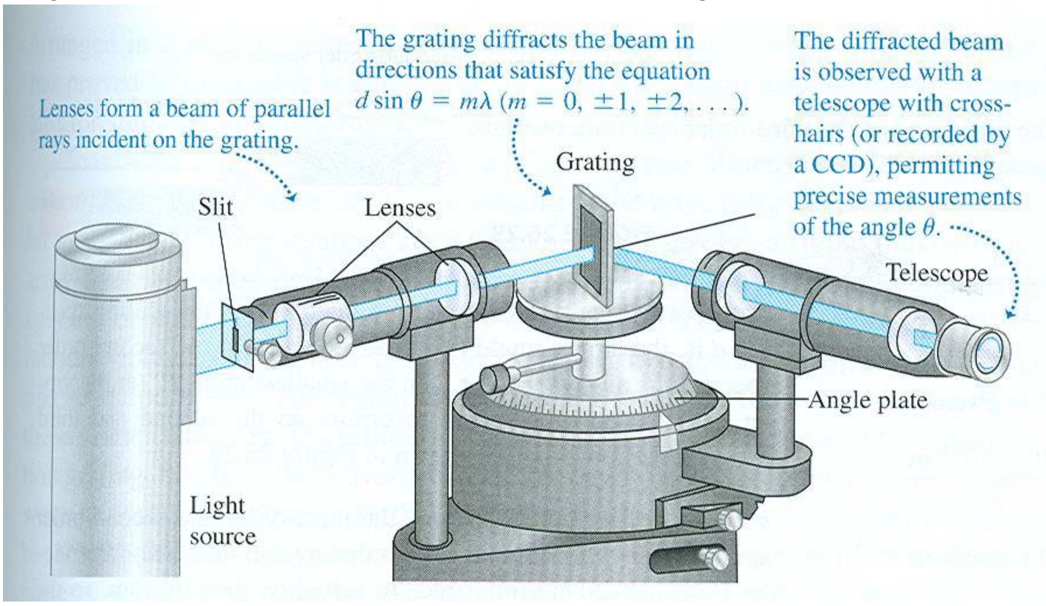
A diffraction grating is ruled with 500 lines per mm. When monochromatic light falls normally on the grating, the first order spectral line is observed at an angle of 20° with respect to the normal to the grating.
(a) What is the slit separation?
(b) What is the wavelength of the monochromatic light? State the colour of the light.
(c) Find the angle the second order spectrum can be seen.
(d) What is the maximum order of spectrum that can be observed?

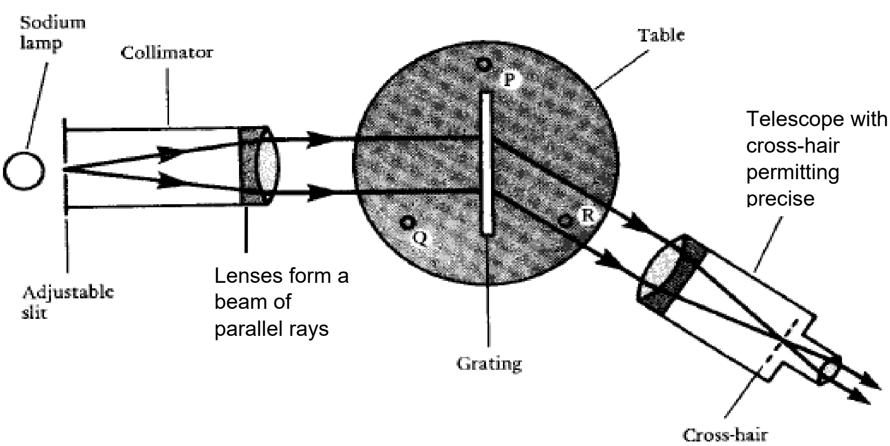
Example 12:

Light from a white source passes through a filter that transmits only the band of wavelengths from 400 nm to 600 nm. When this filtered light is incident normally on a diffraction grating, the 400 nm light in one order of the spectrum is diffracted at the same angle 30° as the 600 nm light in the adjacent order. Find the spacing between the lines in the grating.

Example 13:

Monochromatic light of wavelength 600 nm is used in a spectrometer to illuminate a diffraction grating set normally to the collimator. The grating has 3000 lines per cm. A telescope is used to scan the field to one side of the straight through position. Not counting the straight-through image, what is the maximum number of diffracted images of the slit visible to the observer?

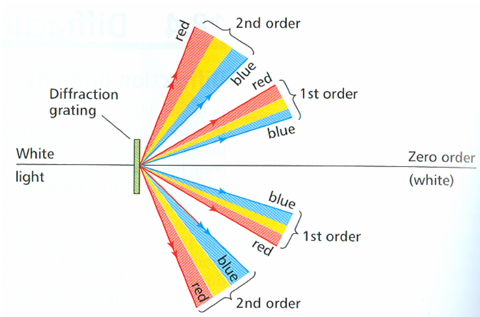




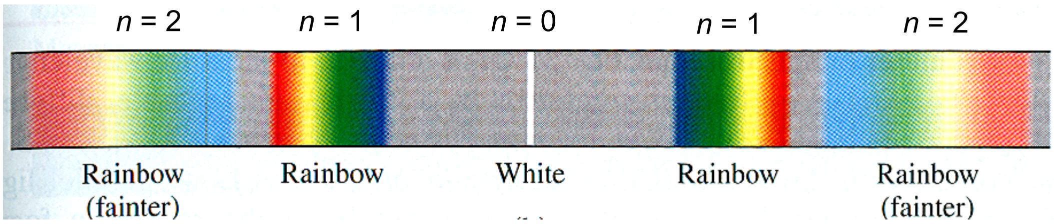
Solution:

Example 14:

Visible light ranging in wavelength from 450 nm to 660 nm falls on a grating with 7500 lines per cm. Describe the spectrum that is observed.



The central order (straight through) position is white because all the colours overlap there. At other orders, the rainbow-like dispersion of the colours occurs on both sides of the central white fringe. The spectrum of colours associated with n = 2 order is completely separate from that of n = 1 order.



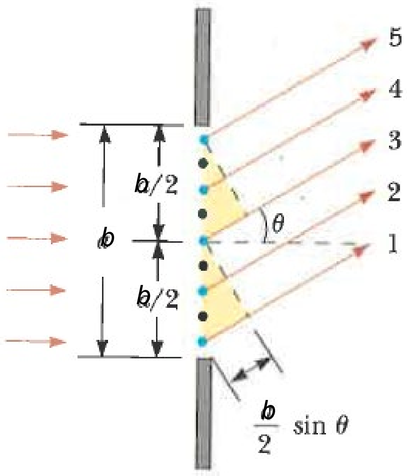
It is possible that the spectra from adjacent orders may overlap. See tutorial question 17.

Annex A

To derive the equation $\sin\theta = \frac{\lambda}{b}$ for the position of first minima in single slit diffraction.

According to Huygens principle, each portion of the slit acts as a source of waves. Hence light from one portion of the slit can interfere with light from another portion and the resultant intensity on the screen depends on the direction θ.

The figure shows light rays directed from each source toward the first dark fringe. The angle θ gives the position of this dark fringe relative to the line between the midpoint of the slit and the midpoint of the central bright fringe. Since the screen is very far from the slit, the rays from each Huygens source are nearly parallel and oriented at nearly the same angle θ.



All the waves that originate at the slit are in phase. Consider wave 1 and wave 5 which originate at the bottom and top of the slit respectively. Wave 1 travels further than wave 3 by an amount equal to path difference $\frac{b}{2}\sin\theta$ where b is the width of the slit. Similarly the path difference between waves 3

If this path difference is exactly half a wavelength (corresponding to a phase difference of π rad), the two waves cancel each other and destructive interference occurs. This is true, in fact, for any two waves that originate at points separated by half the slit width, because the phase difference between two such points is π rad. Therefore, waves from the upper half of the slit interfere destructively with waves from the lower half of the slit when

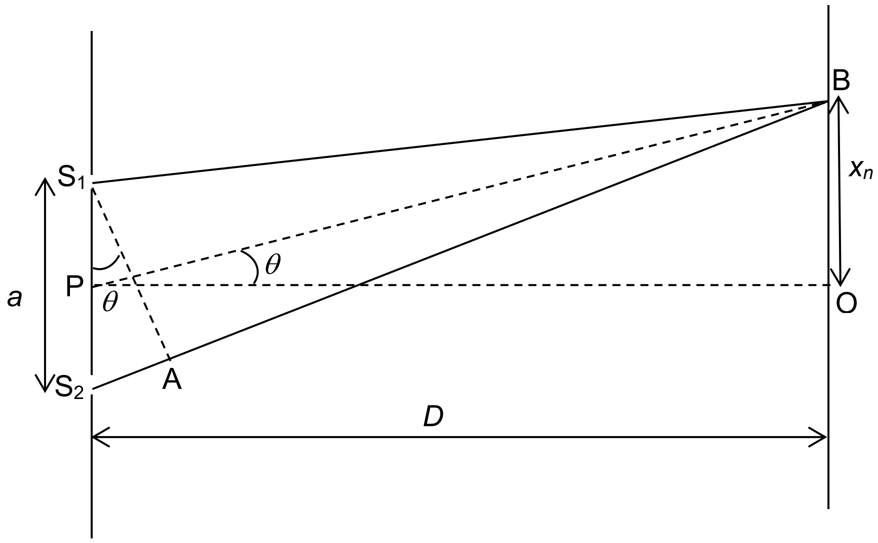
$$\frac{b}{2}\sin\theta = \frac{\lambda}{2}$$

or when

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

Annex B

To derive the equation $\lambda = \dfrac{ax}{D}$ in Young's double-slit experiment



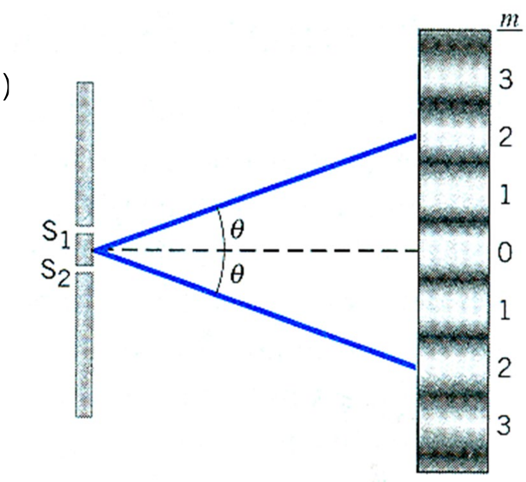
For the *n*th bright fringe formed at position x_n

Path difference between S_1B and S_2B = $S_2B - S_1B$
= S_2A
= $a \sin\theta$

Since constructive interference takes place, path difference = $a \sin\theta = n\lambda$ . . . . (1)

From ΔOPB, $\tan\theta = \dfrac{OB}{OP} = \dfrac{x_n}{D}$ . . . . (2)

n



Since θ is small (because $a \ll D$), $\tan\theta \approx \sin\theta$ . . . (3)

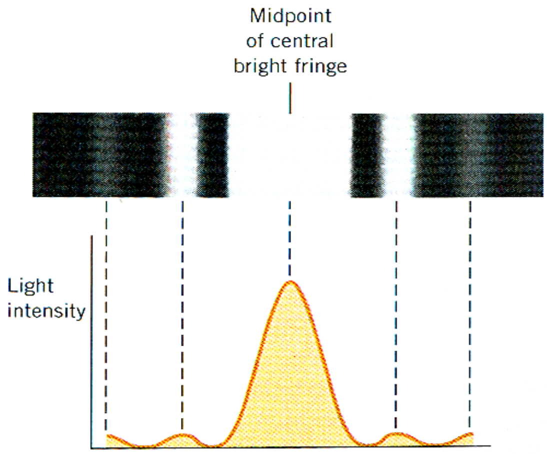
Using (3) and (2) into (1) gives $a\dfrac{x_n}{D} = n\lambda$

or $x_n = n\left(\dfrac{\lambda D}{a}\right)$

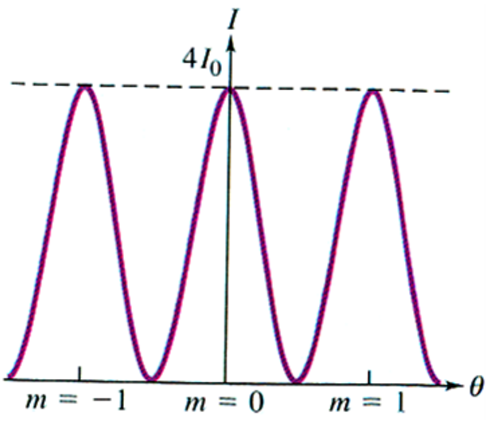
The fringe separation $x = x_{n+1} - x_n = (n+1)\left(\dfrac{\lambda D}{a}\right) - n\left(\dfrac{\lambda D}{a}\right)$

$= \dfrac{\lambda D}{a}$

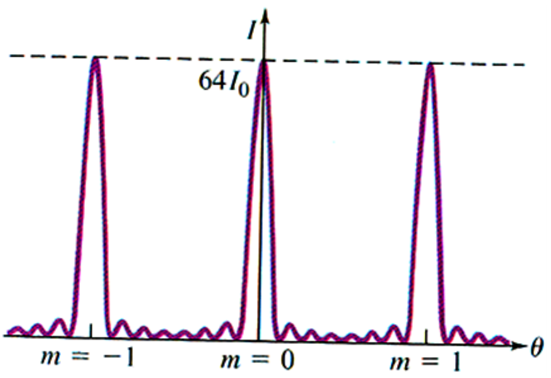
<u>Annex C</u> **Characteristics of diffraction grating**



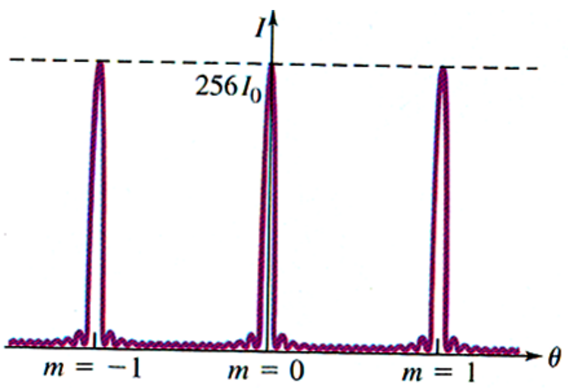
1 slit



2 slits



8 slits – larger, narrower maxima



16 slits – maxima are even taller and narrower

As we <u>increase</u> the number of slits,

1) more light passes through the grating and the absolute <u>intensities</u> of the maxima <u>increases</u>.

2) the <u>sharpness</u> of the maxima <u>increases</u>.

Definition List

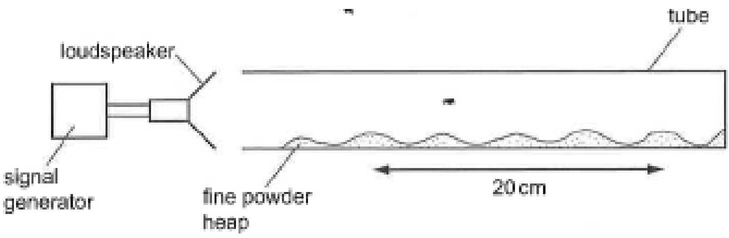
Principle of Superposition	It states that when two or more waves of the same kind meet at a point at the same time, the displacement of the resultant wave is the vector sum of the displacements of the individual waves at that point at that time.
Diffraction	It is the spreading of waves at an edge or a slit so that the waves do not travel in straight lines. *Note: When drawing diagram for diffraction, make sure wavelength is the same before and after passing through slit.*
Coherence	Coherent waves have **constant** phase difference.
Conditions for 2-source interference fringes to be observable	(i) Waves must meet / overlap. (ii)Waves must be coherent i.e. **constant** phase difference. (iii)Waves have the equal or approximately equal amplitudes. (iv)Transverse waves must be either unpolarised or polarised in the same plane. (v) The slit separation d is a few times of wavelength, λ.
Interference	It is an effect that occurs when two or more waves overlap to produce a new wave pattern, i.e. change in amplitude.
Constructive interference	When two waves meet in phase, the resultant amplitude (maximum) is the sum of individual amplitudes of the waves.
Destructive interference	When two waves meet out of phase by 180^o (or π rad), the resultant amplitude (minimum) is the difference of the amplitudes of the waves. If the amplitudes of the two waves are the same, we obtain zero resultant amplitude or *complete destructive interference*.
Stationary waves	When two progressive waves of the same type of equal amplitude, equal frequency, equal wavelength, equal speed travelling in opposite directions meet and undergo superposition with each other, a stationary wave is formed.
Antinode	A point on a *stationary wave* vibrating with maximum amplitude.
Node	A point on a *stationary wave* vibrating with zero amplitude.
Rayleigh criterion	For the diffraction patterns to be just resolved, i.e., able to be distinguished as two distinct objects, the central maximum of one pattern must lie on the first minimum of the other pattern.

Tutorial Questions

Standing Wave

1 Calculate the length of (i) a closed pipe and (ii) an open pipe for each to emit a fundamental note of frequency 256 Hz. Take the speed of sound in air to be 330 m s^{-1}. Neglect end corrections. [0.32 m, 0.64 m]

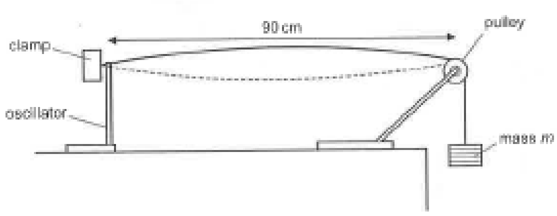
2 A long horizontal tube, containing fine powder, is closed at one end. A loudspeaker, connected to a signal generator, is positioned at the other end.



At a particular frequency, a stationary wave is set up inside the tube and the powder forms heaps as shown. The speed of sound is 330 m s^{-1}.
What type of wave is the stationary wave and what is its frequency?

	type of wave	frequency / kHz
A	longitudinal	3.3
B	transverse	3.3
C	longitudinal	6.6
D	transverse	6.6

3 The speed of a wave on a stretched string is proportional to the square root of the tension in the string.
The string is held under tension between a clamp and a pulley 90 cm apart by a mass *m* of 400 g hanging from the end of the string.

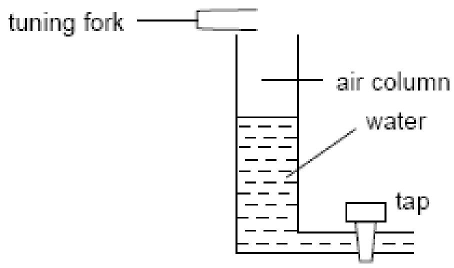


When the frequency *f* of the mechanical oscillator close to the clamped end is 15 Hz, the stationary wave shown is set up between the clamp and the pulley. The hanging mass *m* and frequency *f* are varied until another stationary wave is formed.

Which combination of *m* and *f* gives a stationary wave?

	m / g	*f* / Hz
A	500	30
B	600	40
C	800	60
D	900	90

4 The diagram shows an experiment to produce a stationary wave in an air column. A tuning fork, placed above the column, vibrates and produces a sound wave. The length of the air column can be varied by altering the volume of the water in the tube.



The tube is filled and then water is allowed to run out of it. The first two resonances occur when the air column lengths are 0.14 m and 0.46 m. What is the wavelength of the sound wave? *[0.64 m, LO (c)]*

5 a) If a stationary wave has frequency f and wavelength λ,
- i) State, in terms of λ, the minimum separation of two points that have zero amplitude of vibration.
- ii) State the phase difference between the vibrations of points on the string situated at adjacent antinodes.

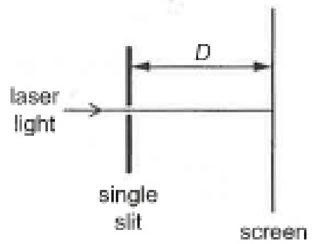
b) The speed v of a progressive wave is given by $v = f\lambda$. Since a stationary wave does not have a speed, explain the significance of the product $f\lambda$ for a stationary wave. *[LO (d)]*

Diffraction

6 A wave of single wavelength is passed through an gap in a barrier and undergoes diffraction. The angle at which the first minimum appears is less than 90°. Which combination of changes in gap size and wavelength must reduce the angle of diffraction?

	gap size	wavelength
A	decrease	decrease
B	decrease	increase
C	increase	decrease
D	increase	increase

7 Laser light of frequency f passes through a single slit. The resulting diffraction pattern is observed on a screen. The width of the central maximum is x. The distance D between the single slit and the screen is much greater than the slit width.



What is the correct expression for the slit width?

A $cx/2Df$ **B** cx/Df **C** cD/fx **D** $2cD/fx$

8. Laser light is incident normally on a single slit of width 12 μm. The light diffracts and a diffraction pattern is observed on a screen 2.7 m away. Part of the diffraction pattern showing the variation with displacement from the centre of the pattern of the intensity of the light is shown.

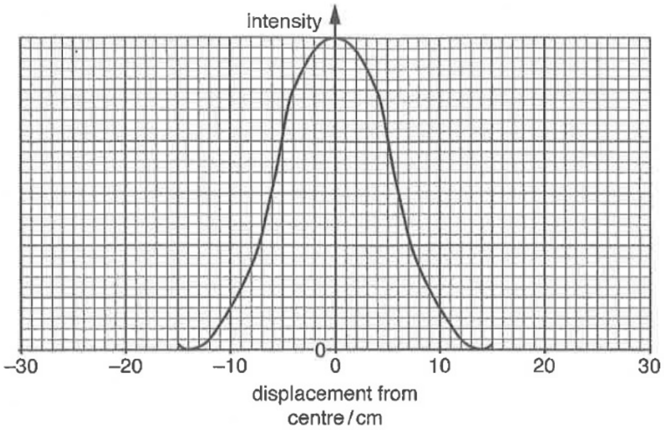


Fig 6

(a) Calculate the wavelength of the laser light. [621 nm, LO (j)]

(b) The slit is now replaced by a narrower slit. On Fig 6, sketch the new diffraction pattern.

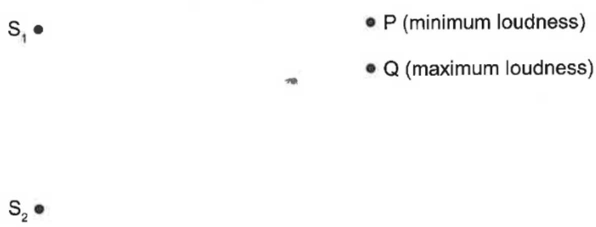
(c) (i) State and explain the changes to the observed diffraction pattern in Fig 6 when visible light of a longer wavelength is used.

(ii) When white light is incident on a single slit, the central fringe is coloured at the edges and has a white central region. Explain this observation.

9 A double star is at a distance 20 light years from the Earth. A telescope with a diameter of 3.0 m is used to view the star. (A light year is the distance light travels in a vacuum in one year, which is 9.5×10^{15} m.)
What is the approximate minimum separation between the two stars of the double star that can be detected by the telescope? *[3.0×10^{10} m, LO (k)]*

<u>Interference</u>

10 Two coherent sources of sound S_1 and S_2 give rise to a minimum loudness at point P and a maximum loudness at point Q.



The amplitude of the sound produced by S_2v is reduced. The frequency of the sound produced by both sources is unchanged. What happens to the loudness of the sound at P and at Q?

	loudness at P	loudness at Q
A	decreases	decreases
B	increases	decreases
C	increases	no change
D	no change	no change

11 Two coherent sources produce wavefronts as shown in Fig. 10.

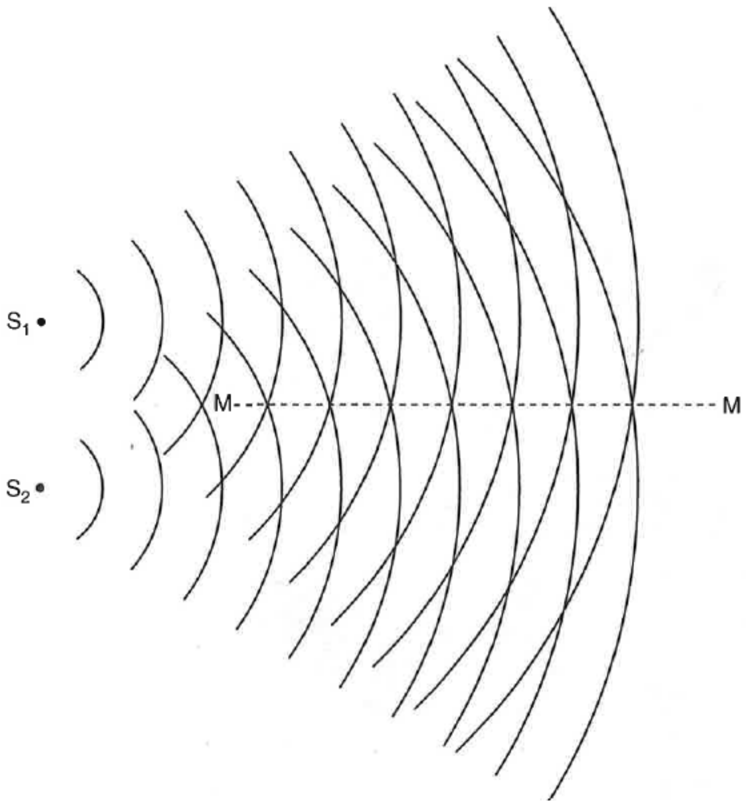
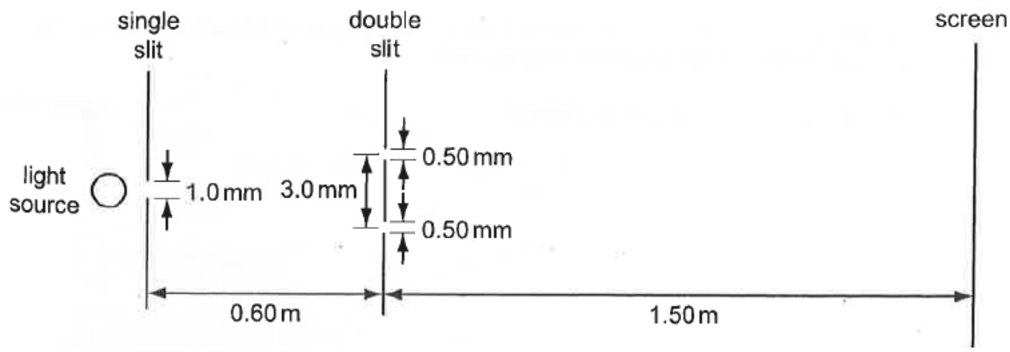


Fig. 10

The amplitude of the resultant of the waves is seen to be a maximum along the direction MM. On Fig. 10, draw a line to show

1. a second direction along which the amplitude is a maximum (label this line AA),
2. a direction in which the amplitude is a minimum (label this line NN). *[LO (g)]*

12 A student sets up an experiment to demonstrate double-slit interference, using light of wavelength 6.0×10^{-7} m. The main features of the apparatus, and some of the dimensions, are illustrated.



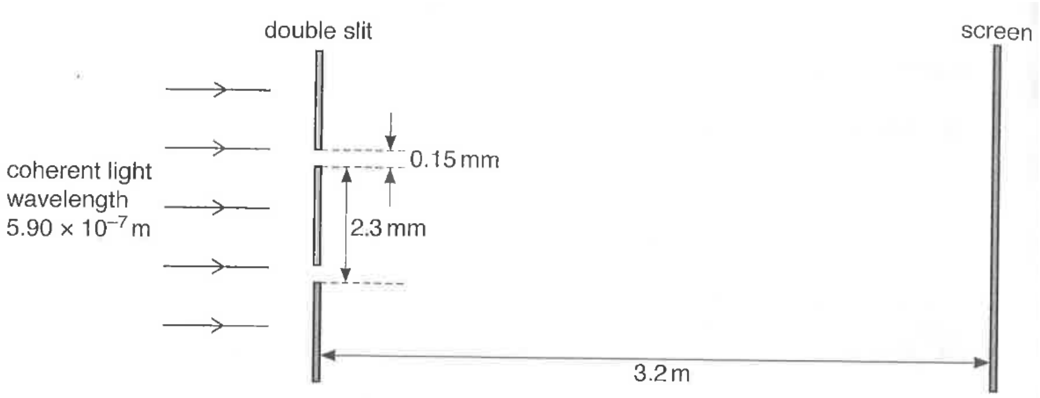
What is the separation of the bright fringes on the screen?

A 0.12 mm **B** 0.30 mm **C** 0.90 mm **D** 1.80 mm

13 In the Young double-slit experiment, state and explain the effect on the intensity and separation of the fringes if the following changes were carried out independent of one another:

(i) the distance between the slits is decreased
(ii) the distance from the slits to the screen is increased
(iii) the intensity of the light on one of the slits is reduced
(iv) one of the slit is covered with a transparent medium of refractive index higher than air
(v) the width of both slits is reduced without altering their separation
(vi) the screen is rotated so that it is no longer parallel to the plane of the two slits

14 Coherent light of wavelength 590 nm is incident normally on a double slit as shown. Both slits are rectangular with a width of 0.15 mm. The separation of the two slits is 2.3 mm. A screen is placed parallel to the plane of the double slit at a distance 3.2 m from the double slit.



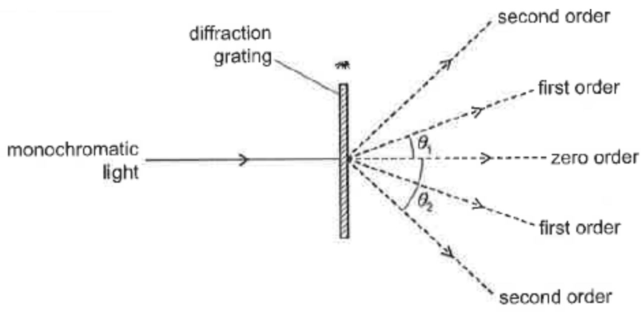
(i) Initially, one of the two slits is covered. Calculate the width of the central bright maximum formed on the screen by diffraction through the uncovered slit. *[2.52 cm]*
(ii) Both slits are now uncovered. Use Rayleigh criterion to explain whether the diffraction patterns produced by the two slits are seen on the screen as being separate.
(iii) Calculate the width of a fridge produced by diffraction through the uncovered double slit. *[8.2 x 10^{-4} m]*
(iv) Use your answers in (i) and (iii) to estimate, to one significant figure, the number of fringes observed in the central bright maximum produced by diffraction at one of the slits. *[30, LO (i), (j), (k)]*

Diffraction Grating

15 Light of frequency 6.0 x 10^{14} Hz passes through a diffraction grating wit 4.0 x 10^3 lines per centimetre. What is the angle between the two third-order diffraction maxima?

A 12° **B** 23° **C** 37° **D** 74°

16 Monochromatic light is incident normally on a diffraction grating. First and second order images are observed at angles of diffraction θ_1 and θ_2 respectively. The diagram is not to scale.



What is the relation between the angles of diffraction and between the intensities of light observed in the two orders?

	angles of diffraction	intensities
A	$2\theta_1 = \theta_2$	first order brighter
B	$2\theta_1 = \theta_2$	same
C	$2\theta_1 < \theta_2$	first order brighter
D	$2\theta_1 < \theta_2$	same

17 A diffraction grating has 6000 lines per cm.

(i) What is the angle of dispersion between the red light and violet light in the first order spectrum produced?

(ii) What is the angle of dispersion in the second order? Is there any overlap between the second and third order spectrum?

Let wavelength of red light be 700 nm and wavelength of violet light be 450 nm.

[1st order: violet – 15.7°. red – 24.8°; 2nd order: violet – 32.7°, red – 57.1°; yes, 3rd order violet – 54.1°]

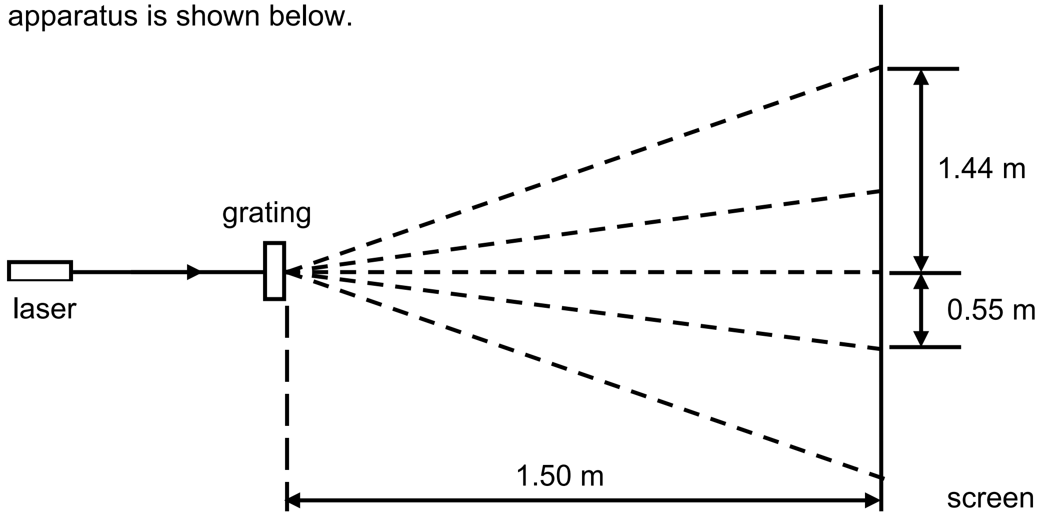
18 A narrow beam of coherent light of wavelength 589 nm is incident normally on a diffraction grating having 4.00×10^5 lines per metre.

(i) Determine the number of orders of diffracted light that are visible on each side of the zero order.

(ii) A student suspects that there are two wavelengths of light in the incident beam, one at 589.0 nm and the other at 589.6 nm.

1. State the order of the diffracted light at which the two wavelengths are most likely to be distinguished.
2. The minimum angular separation of the diffracted light for which two wavelengths may be distinguished is 0.10°. Make calculations to determine whether the student can observe the two wavelengths as separate images.

19 You are asked to determine the wavelength of monochromatic red light emitted by a laser. The apparatus is shown below.



The diffraction grating has 5.5 x 10^5 lines per metre and is set so that its plane is normal to the incident light and is situated 1.50 m from a large screen. Bright spots are observed at 0.55 m and 1.44 m from the central bright spot.

a) Calculate the wavelength of the laser light from observations of the first-order diffracted light. *[(626 nm, LO (l))]*

b) Suggest one advantage and one disadvantage of obtaining the wavelength by using observations of the second-order diffracted light rather than the first-order diffracted light.