Additional Notes

Topic 12: Superposition

Suggested Lecture Plan

Lect	Concepts	Qns	Est. Video Time
1	A, B.1, B.2	CYU 1	40 min
2	B.3, B.4, C.1	Eg 1, Eg 2, Eg 3	42 min
3	C.2, C.3, C.4	Eg 4, Eg 5	42 min
4	C.5, D.1, D.2, E.1, E.2	Eg 6, Eg 7, Eg 8	36 min
5	E.3, E.4, E.5, E.6.1, E.6.2	CYU 2, Eg 9, Eg 10	38 min
6	E.6.3, E.6.4, E.6.5	Eg 11, Eg 12	46 min

Learning Outcomes:

Candidates should be able to:

A. Principle of superposition

(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

B. Two-source interference

(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 $\lambda = ax/D$ for double-slit interference

C. Single slit diffraction

(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

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

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

D. Multiple slit diffraction – diffraction grating

(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)

E. Stationary waves

(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

Wave

- **Nature of Wave**
  - Periodic Motion
  - Disturbance through Space and Time
  - Transfer Energy Not Matter
  - Transverse and Longitudinal Waves
  - Most Wave Types require Material Medium for Propagation, except EM Waves which can propagate in a vacuum as well as in a material medium.
- **Properties of Waves**
  - Wave speed
  - Wavelength
  - Frequency
  - Amplitude
  - Transverse Wave Crest & Trough
  - Longitudinal Wave Rarefaction & Compression
- **Behaviour of Wave**
  - Superposition
    - Stationary Waves
    - Single Slit Diffraction
    - Two-Source Interference
    - Multiple Slit Diffraction

Additional Notes

A Principle of Superposition

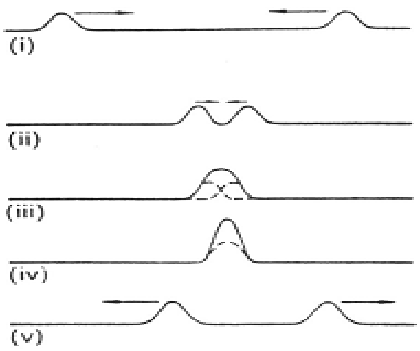
In this topic we are concerned with the interaction between two or more waves. What happens when two or more waves meet?

The Principle of Superposition states that when two waves <u>meet</u> at a point, the resultant displacement is equal to the <u>vector</u> sum of the individual displacements.

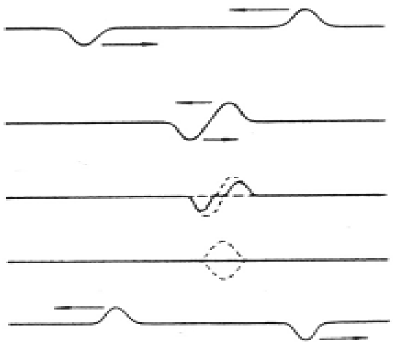
Memorize

- The waves must be of the <u>same type</u> (ie. light and sound wave cannot superpose but light and x-rays can) and if they are transverse waves, the direction of their polarisation must be the same.

Crest Meets Crest



Crest Meets Trough



Scan for animation of two pulses approaching each other in-phase / anti-phase

Guiding questions to help you navigate and understand the topic better:

1. What is the principle of superposition?
2. What are the general contexts under which it can be applied?
   A. Two-source interference – when waves from two sources overlap
   B. Single slit diffraction – when a wave passes through a single slit
   C. Multiple slit diffraction – when a wave passes through many slits
   D. Stationary waves – when two waves moving in opposite directions overlap
3. How is the principle of superposition being applied in these contexts?
4. What are the *patterns* that arise out of the interaction of waves in each of these contexts?

<u>Check Your Understanding 1 (solution and explanation in SLS).</u>

Select the scenario(s) where you think it makes sense to invoke the Principle of Superposition. You may select more than one option.

☐ A sound wave travels from A to B and a light wave travels from B to A.
☐ A sound wave from a tuning fork travels from A to B and a sound wave from another tuning fork of lower pitch travels from B to A.
☐ A beam of red laser light travels from A to B and a beam of green laser light travels from B to A.
☐ You are dancing as music plays from both sides of your headphones / earphones.
☐ Your teacher asks the class a question. You raise your hand and offer a possible answer.

B Two-Source Interference

Key Term: Interference
When two or more waves overlap and interact, we say they interfere with each other.

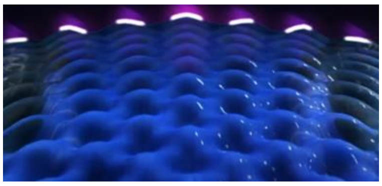
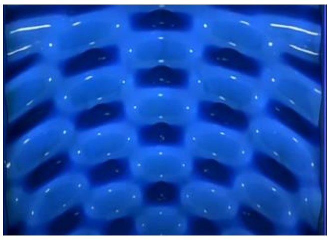
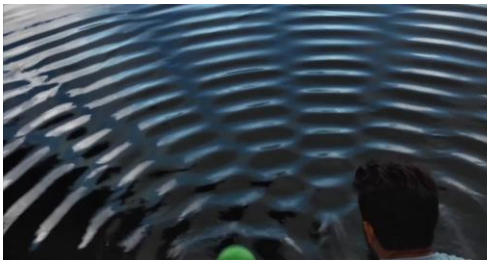
Interference is an effect that occurs when two or more waves **overlap** to produce a new wave pattern, i.e. a change in amplitude.

B.1 General patterns using water waves

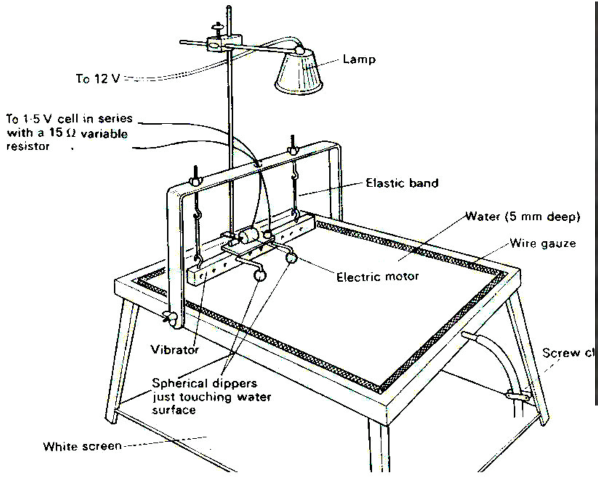
Source for left and centre picture: Fabric of the Cosmos by Brian Greene

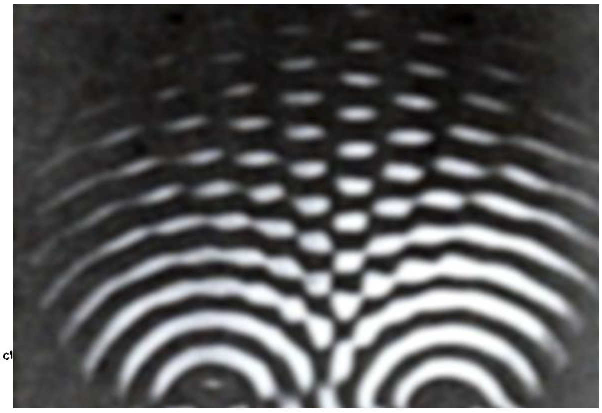
Source for picture on the right: The Original Double Slit Experiment by Derek Muller

Some visualisations:

Ripple Tank Experimental setup and observations





- Two dippers, attached to the vibrator, are set into vertical vibrations with the same amplitude, frequency and phase (in phase).

- By projecting the light from the lamp such that it passes through the water onto the white screen below, an interference pattern will be easily detected.

- A snapshot of the interference pattern observed is as shown. Note the lines of constructive interference (alternating bright and dark regions) and destructive interference (dark lines).

Additional Notes

B.1.1 General patterns using water waves – Mapping path differences

Fig. 1 below shows a time-freeze diagram of waves from two sources, S_1 and S_2. The two sources are *coherent*, and are in phase (phase difference = 0).

Key term: Coherence
Coherent sources are sources which produce waves of the **same frequency** with a **constant phase difference**. Waves produced by coherent sources are therefore coherent.

Two sources or waves **do not** have to be in phase to have a constant phase difference.

Coherence is used to indicate two waves that have a **constant phase difference**.

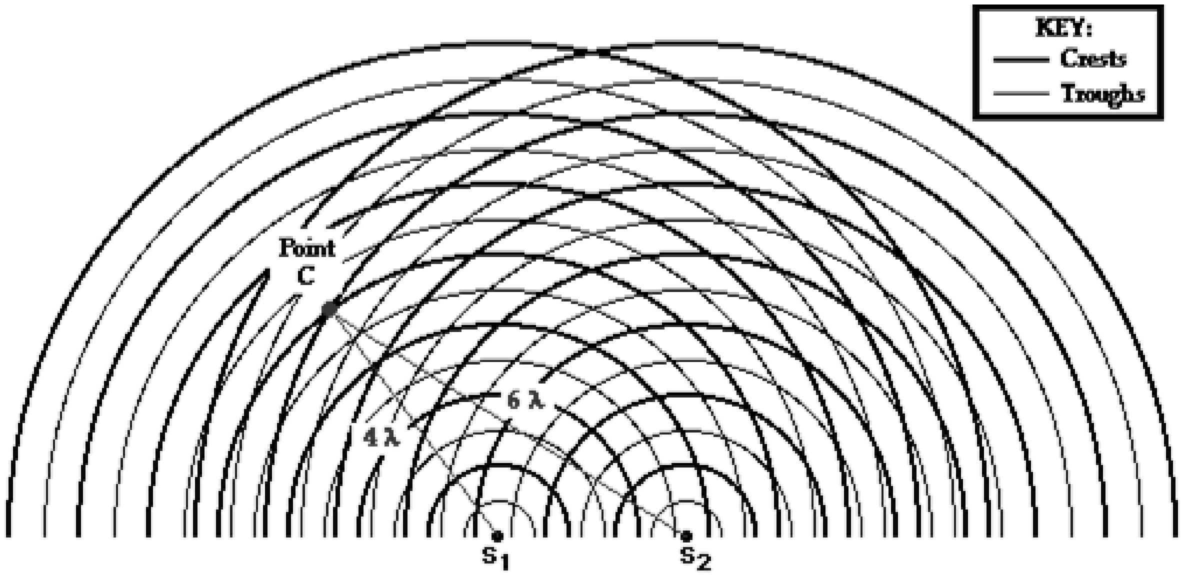


Fig. 1

Note that it is **incorrect** to use "**same phase difference**" to describe coherent sources which has "**constant** phase difference".

Wavefronts connecting crests and troughs of waves from each source are drawn. When two waves travel along paths of unequal length to meet at a point (eg. Point C), there is a *path difference* between the waves from their respective sources to that point.

Key term: Path difference
Path length is the distance travelled by a wave from its source to a point.
When waves from two sources meet at a point, the *difference* in their respective path lengths is called **path difference**.

In the figure above, the path difference for the waves to meet at C is $6\lambda - 4\lambda = 2\lambda$.

At which points in Fig. 1 are the path differences 0, ½ λ, λ, 3/2 λ or 2λ?

This path difference can be used to tell the *phase difference* of the waves when they arrive at the point.

Key term: Phase difference

- A measure of how much one wave is out of step with another. A phase difference of one cycle corresponds to 360° or 2π radians
- Two waves are said to meet in phase when they have a phase difference of $2\pi n$, where n = 0, 1, 2, …

- Two waves are said to meet in antiphase when they have a phase difference of $(2n+1)\pi$, where n = 0, 1, 2, …

- Phase difference is related to path difference:

$$\text{phase difference} = \frac{\text{path difference}}{\text{wavelength}} \times 2\pi \left(+\text{ phase difference between sources}\right)$$

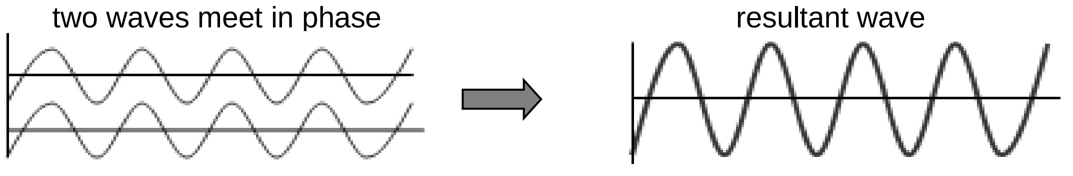
  The phase difference when waves meet at point C is hence 4π, which means they meet in phase.

  At which points in Fig. 1 are the phase differences 0, π, 2π, 3π, or 4π? At which of these points do the waves meet in phase, or antiphase?

The phase difference between waves when they meet at a point will determine *how they interfere* with each other.

Key term: Constructive interference

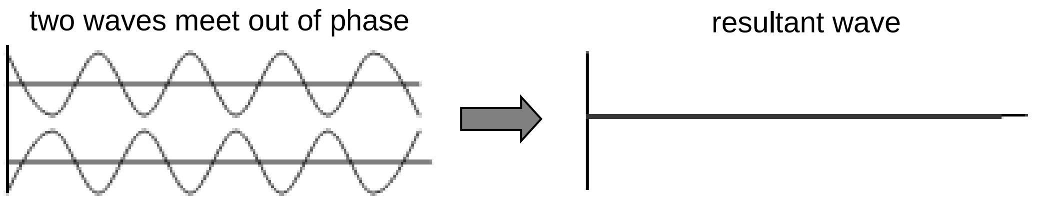
When waves from two or more sources arrive at a point in phase, they reinforce each other: The amplitude of the resultant wave is the *sum* of the amplitudes of the individual waves, giving rise to a maximum resultant amplitude. This is called **constructive interference**.



Therefore, at point C where the waves meet in phase, constructive interference occurs, resulting in a maximum resultant amplitude. At which other points in Fig.1 do constructive interference occur?

Key term: Destructive interference

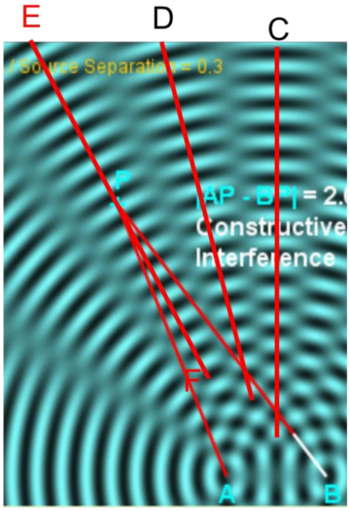
When waves from two or more sources arrive at a point exactly a half-cycle out of phase: The amplitude of the resultant wave is the *difference* of the amplitudes of the individual waves, giving rise to a minimum resultant amplitude. This is called **destructive interference**.

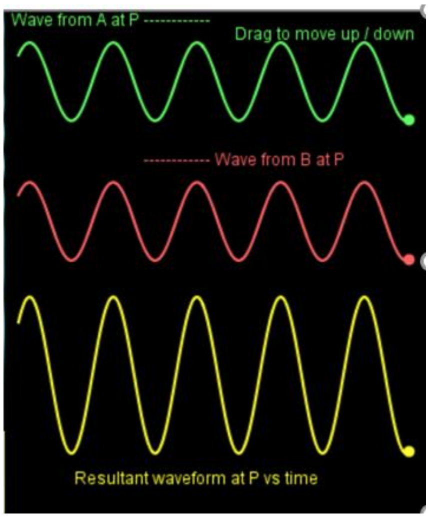


At which points in Fig.1 do destructive interference occur?

Additional Notes

B.1.2 General patterns using water waves – Constructive Interference

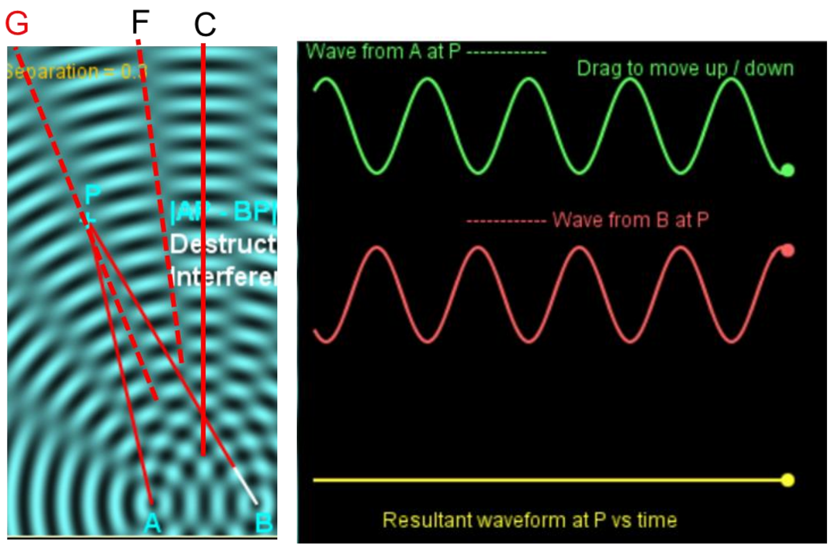




Sources A and B are **coherent** and **in phase**		
For point P anywhere along		
line C	line D	line E
Path difference = BP – AP		
= 0	= λ	= 2λ
⇨ **Phase difference** between the 2 waves at P		
= 0	= 2π	= 4π
⇨ Waves arrive in <u>phase / antiphase</u> at P		
⇨ **<u>Constructive / Destructive</u> interference** occurs at P		
⇒Resultant amplitude at P = <u>*sum* / difference</u> of individual amplitudes of the 2 waves at P, ⇒Resultant amplitude is a <u>maximum / minimum</u>		
Line C	Line D	Line E
is a(n) **<u>antinodal / nodal</u> line** connecting all points where		
path difference = 0, phase difference = 0	path difference = λ, phase difference = 2π	path difference = 2λ, phase difference = 4π
⇨ <u>Maximum/Minumum</u> resultant amplitude at all points along		
line C	line D	line E
Since intensity ∝ amplitude^2, ⇨ <u>Maximum/Minumum</u> intensity at all points along		
line C	line D	line E

See waves arriving at a point and undergoing Constructive interference:

B.1.3 General patterns using water waves – Destructive Interference

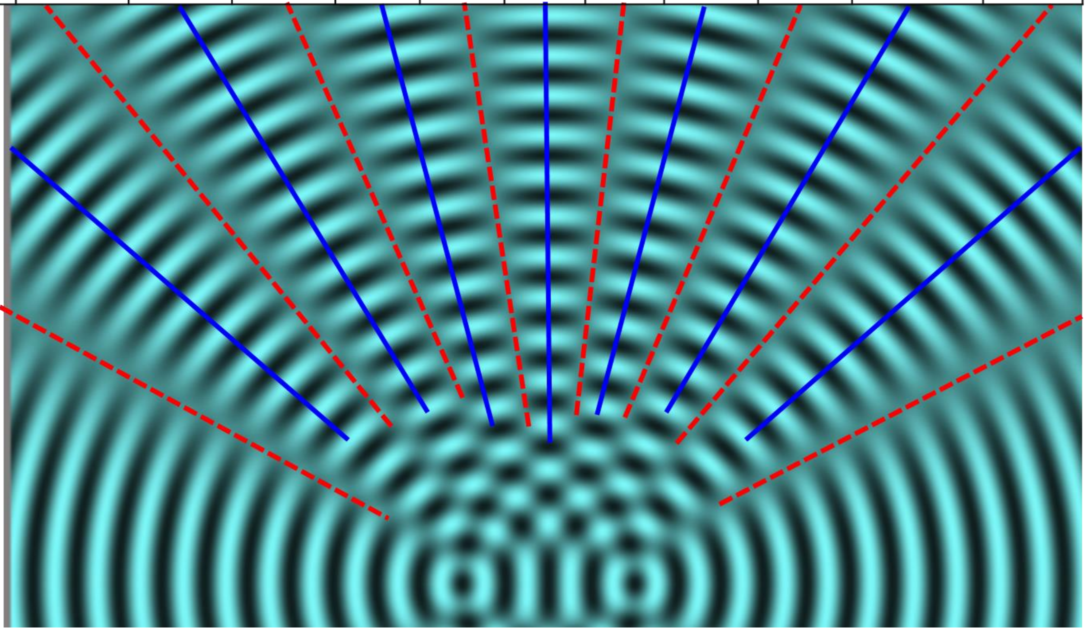


Sources A and B are **coherent** and **in phase**	
For point P anywhere along	
line F	line G
Path difference = BP – AP	
= ½ λ	= $^{3}/_{2}$ λ
⇨ **Phase difference** between the 2 waves at P	
= π	= 3π
⇨ Waves arrive in phase / antiphase at P	
⇨ **Constructive / Destructive interference** occurs at P	
⇒Resultant amplitude at P = *sum* / difference of individual amplitudes of the 2 waves at P, ⇒Resultant amplitude is a maximum / minimum	
Line F	Line G
is a(n) **antinodal / nodal line** connecting all points where	
path difference = ½ λ, phase difference = π	path difference = $^{3}/_{2}$ λ, phase difference = 3π
⇨ Maximum/Minimum resultant amplitude at all points along	
line F	line G
Since intensity ∝ amplitude^2, ⇨ Maximum/Minimum intensity at all points along	
line F	line G

B.1.4 General patterns using water waves– Consolidation

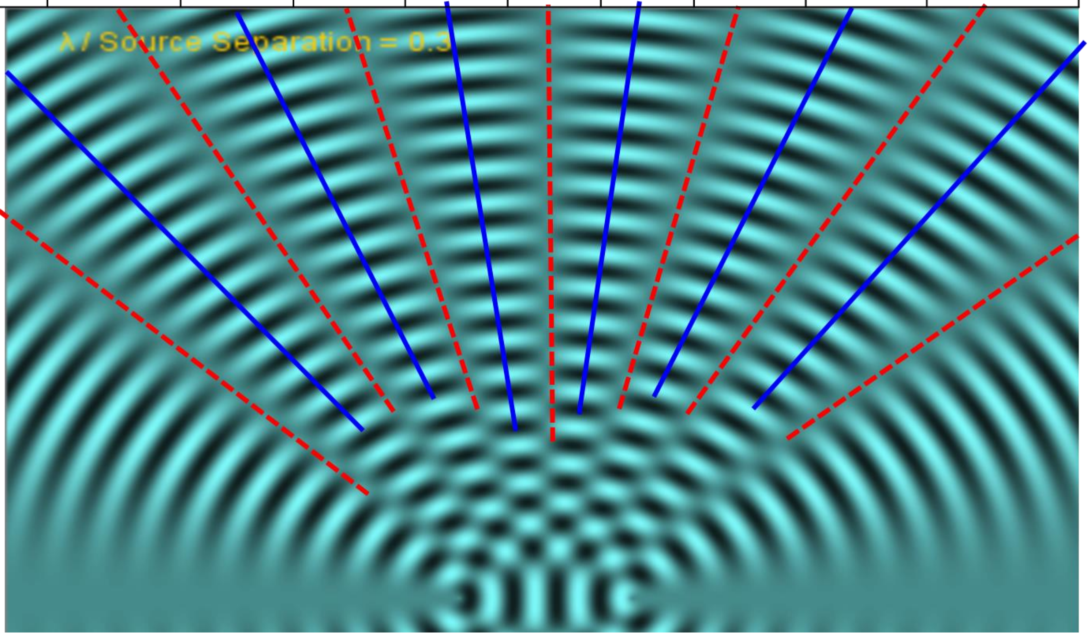
Sources in phase

Constructive or destructive	D	C	D	C	D	C	D	C	D	C	D
Phase diff	5π	4π	3π	2π	π	0	π	2π	3π	4π	5π
Path diff	$^5/_2\,\lambda$	2λ	$^3/_2\,\lambda$	λ	$^1/_2\,\lambda$	0	$^1/_2\,\lambda$	λ	$^3/_2\,\lambda$	2λ	$^5/_2\,\lambda$



Sources in half a cycle out of phase

Constructive or destructive	D	C	D	C	D	C	D	C	D
Phase diff	5π	4π	3π	2π	π	2π	3π	4π	5π
Path diff	2λ	$^3/_2\,\lambda$	λ	$^1/_2\,\lambda$	0	$^1/_2\,\lambda$	λ	$^3/_2\,\lambda$	2λ



Additional Notes

Scan to play with a virtual ripple tank!

Vary the wavelength and the distance between the sources. What do you observe?

You can pause / step through the simulation to better understand how the pattern evolves with time.

Additional Notes

Note: Two waves must always **meet in phase** for **constructive interference** to occur, irrespective of the phase difference between the two sources.

Two waves must always **meet in antiphase** for **destructive interference** to occur, irrespective of the phase difference between the two sources.

B.2 General conditions for two-source interference

For constructive/destructive interference

		Path Difference	
		$n\lambda$	$\left(n+\frac{1}{2}\right)\lambda$
Sources	**in phase**	Constructive	Destructive
	in antiphase	Destructive	Constructive

For steady observable interference

- When two waves superpose, interference occurs. But this does not mean that the interference pattern can be observed.

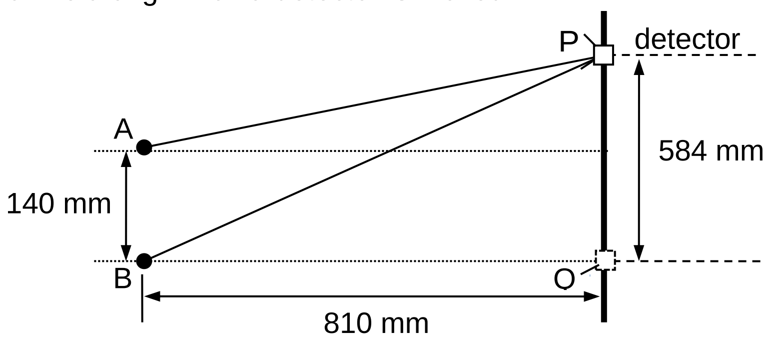
- For steady and observable interference patterns, the following conditions must be met:
  - The two waves must overlap.
  - The sources must be coherent. This means that they have a constant phase difference at all times (and therefore the waves need to have the same frequency and wavelength.)
  - The waves must have approximately the same amplitude (to achieve complete wave cancellation at the destructive region, and achieve good contrast between maxima and minima of intensity);
  - For transverse waves, they must either be unpolarized, or polarized in the same plane.

- Examples of coherent sources include: a single laser beam passing through 2 slits; 2 speakers fed by the same source; and 2 dippers attached to one oscillator.

B.3 Application to Microwaves (wavelength ~1 mm – 1 m)

Example 1

Two coherent microwave sources A and B are in phase with one another. They emit waves of equal amplitude and of wavelength 30 mm. They are placed 140 mm apart and a distance 810 mm from a line along which a detector is moved.



Additional Notes

(a) Determine if the intensity of microwaves received by the detector at P is high or low.

(b) Determine the number of locations where high intensities will be detected as the detector moves from P to Q, a point directly opposite source B.

(a) **Step 1: Calculate distance AP, BP and the path difference**

Step 2: Express the path difference in terms of multiples of wavelength

Step 3: Use the table on page 10 to conclude whether constructive or destructive interference has taken place and hence whether intensity will be a maximum or minimum.

Additional Notes

B.4 Application to visible light waves (wavelength ~ 400 nm – 700 nm)

B.4.1 Monochromatic light sources

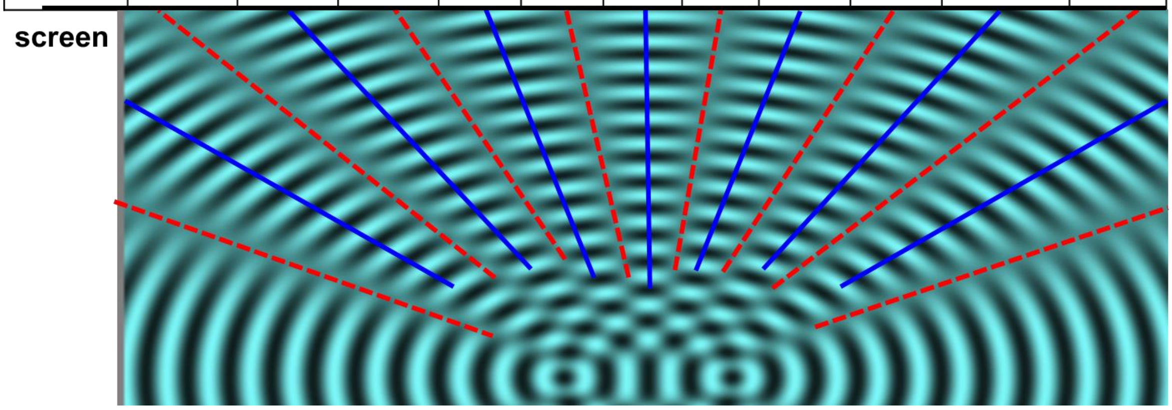
A monochromatic light is one that is of a single frequency or wavelength.

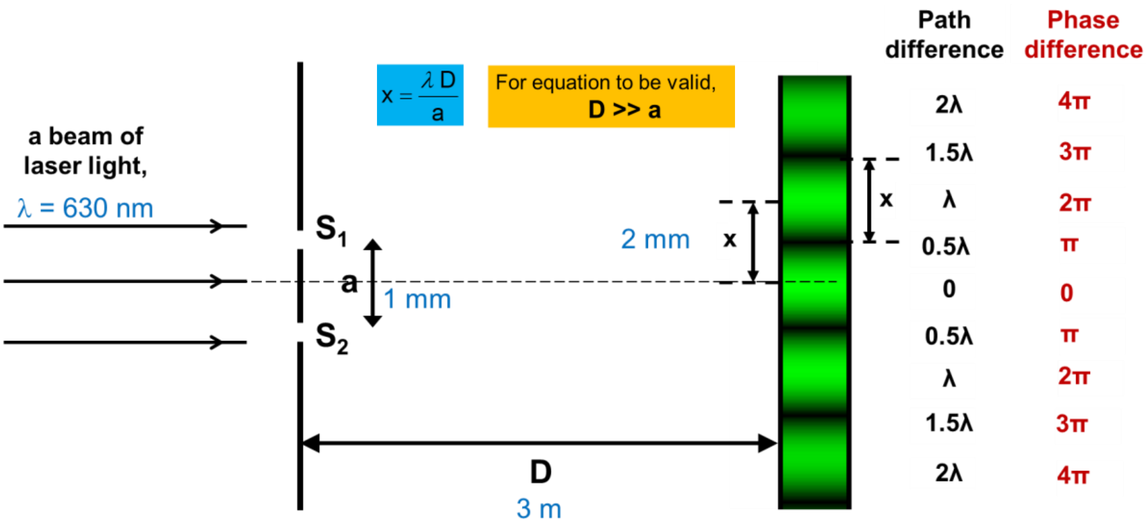
For water waves, we are able to observe the interference pattern as being made up of antinodal lines that trace points of maximum amplitude, and nodal lines that trace points of minimum amplitude. However, we are unable to observe this pattern for light waves, but we can place a screen and observe the interference pattern forming on the screen.

You may notice that there is a central region consisting of alternating bright and dark fringes followed by regions of alternating bright and dark fringes at the 2 sides of the central region. This will be further explained in Concept C.4.

When monochromatic light (eg. laser light) passes through 2 identical slits that acts as coherent sources, it forms an interference pattern on the screen, that consists of alternating bright and dark fringes. We can use the general pattern established in section B.1.1 to understand this phenomenon.

Constructive or destructive	D	C	D	C	D	C	D	C	D	C	D
Phase diff	5π	4π	3π	2π	π	0	π	2π	3π	4π	5π
Path diff	$^5/_2\,\lambda$	2λ	$^3/_2\,\lambda$	λ	$^1/_2\,\lambda$	0	$^1/_2\,\lambda$	λ	$^3/_2\,\lambda$	2λ	$^5/_2\,\lambda$
Bright or dark	D	B	D	B	D	B	D	B	D	B	D





Points on the screen that coincide with antinodal lines are points where constructive interference took place, resulting in maximum intensity or bright fringes. Points on the screen that coincide with nodal lines are points where destructive interference took place, resulting in minimum intensity or dark fringes. The explanation for constructive and destructive interference in pages 7 and 8 are applicable, as they are for all cases of two-source interference.

The bright and dark fringes alternate with one another, much like how antinodal and nodal lines alternate.

- When the distance from the slits to the screen (D) is much larger than the slit separation (a), i.e. $a << D$, the fringes are <u>evenly spaced</u> and fringe separation x is approximately

$$\boxed{x = \frac{\lambda D}{a}}$$

where x = fringe separation i.e. distance between 2 adjacent bright fringes (or 2 dark fringes)
a = slit separation
D = the distance from the double slits to the screen
λ = wavelength of the light (of the order of 10^{-7} m for visible light)

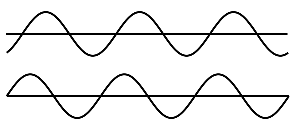
The derivation of this equation is not required by the syllabus, and can be found in the appendix.

- An example of typical dimensions used for such an experiment would be $a \approx 1$ mm, $D \approx 3$ m. When light of $\lambda \approx 630$ nm is used, it results in $x \approx 2$ mm.
- This equation is valid for any type of two-source interference, provided $a << D$.

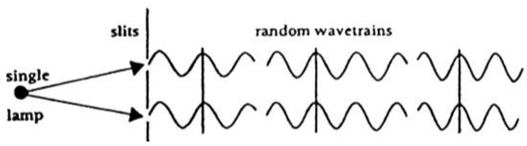
<u>Purpose of double slit</u>

The double slit plays an important role – the interference pattern would not be observed if two identical laser sources were to be used in place of the double slits to provide for two-source interference.

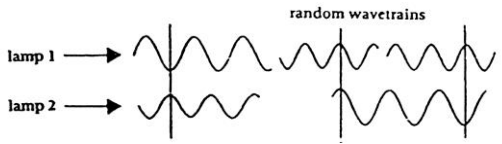
- Light is emitted from sources as a series of pulses or packets of energy, each lasting for a very short time, about a nanosecond (10^{-9} s). Between each pulse there is an abrupt change in the phase of the waves in a random manner. Waves from two separate light sources may be in phase in one instant, but out of phase in the next. Hence, separate light sources, even of the same frequency, produce incoherent waves.
- When a single laser beam enters a double slit, waves emitted from the two slits will always be in phase, so the two slits act as coherent sources.
- As waves spread after passing through each slits (more on this in the next Section), this spreading allows waves from the two slits to overlap and undergo interference.



<u>Coherent</u>:
Constant phase difference between 2 waves

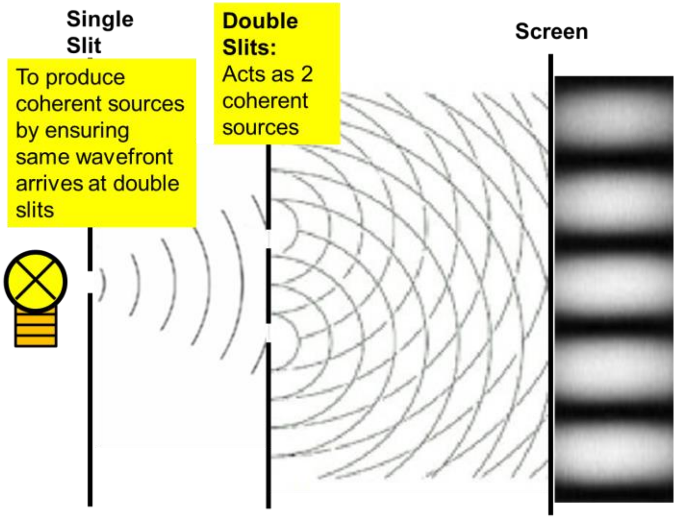


<u>Coherent</u>:
Phase difference between random wave is constant.



<u>Incoherent</u>:
The phase difference is different as the sources are independent of each other.

B.4.2 Young's double slit experiment using filament lamp



When a white light source such as a filament lamp is used, a single slit is required to be placed between the light source and the double slit.

- Light from this source is diffracted (spread) at the single slit, producing wavefronts that reach both slits of the double slit.
- This ensures the double slit act as coherent sources, thus creating a sustained and observable interference pattern.

<u>Nature of Science</u>

Physicists have long debated and proposed different models of light, such as the wave model proposed by Christian Huygens in 1690 and the particle model proposed by Isaac Newton in 1704. The first experimental evidence came from Thomas Young who in 1803 (a hundred years later!) conducted an experiment that is now named after him, showing that light behaved as waves by undergoing interference. The wave theory thus became the dominant theory, although later in 1905 (another hundred years later!), Albert Einstein showed that new evidence demonstrated light could behave as a particle too! More of that in the topic Quantum Physics to be taught next year.

This is an example how ***science is an evidence-based, model-building enterprise concerned with the natural world. Scientists require the use of both logic and creativity in proposing new models and testing them against their predictions.*** *Also,* ***scientific knowledge is reliable – able to provide accurate predictions that can be tested, yet subject to revision in the light of new evidence.***

Example 2

In a Young's double slit experiment, a coherent monochromatic blue light of wavelength 4.0×10^{-7} m illuminates two narrow parallel slits separated by 0.50 mm. A white screen is placed 2.0 m away from the slits, as shown below.

Additional Notes

(a) Describe what can be seen on the screen.

(b) Calculate the separation between the fringes on the screen.

(c) Determine how the spacing between the fringes on the screen would be affected if the following changes were made separately:

- **(i)** the screen is brought nearer to the slits;
- **(ii)** the slits are separated further apart; and
- **(iii)** a red light source is used.

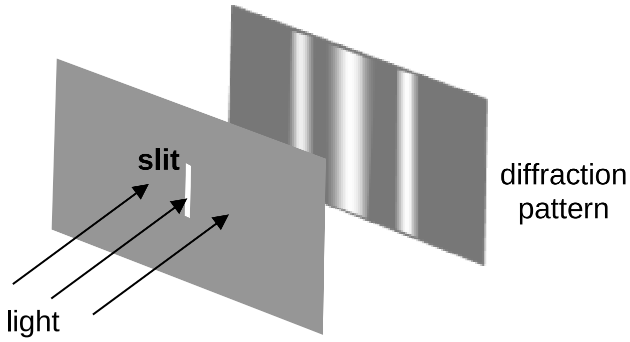
C Single Slit Diffraction

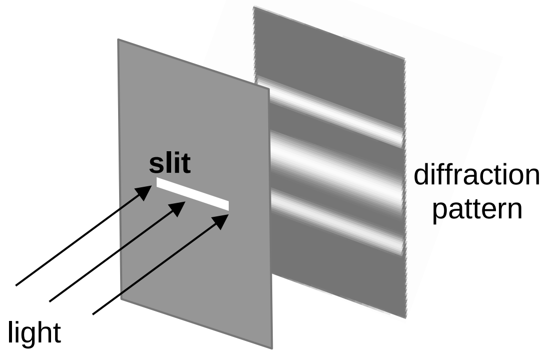
C.1.1 Single slit diffraction of light

From the previous section, we saw that when a coherent beam of light passes through a double slit, it forms an interference pattern – alternating bright and dark fringes, on the screen.

In this section, we look at what happens when a wave passes through a single slit. The diagram below shows a single slit diffraction pattern formed on the screen, when a beam of laser passes through a single vertical slit.

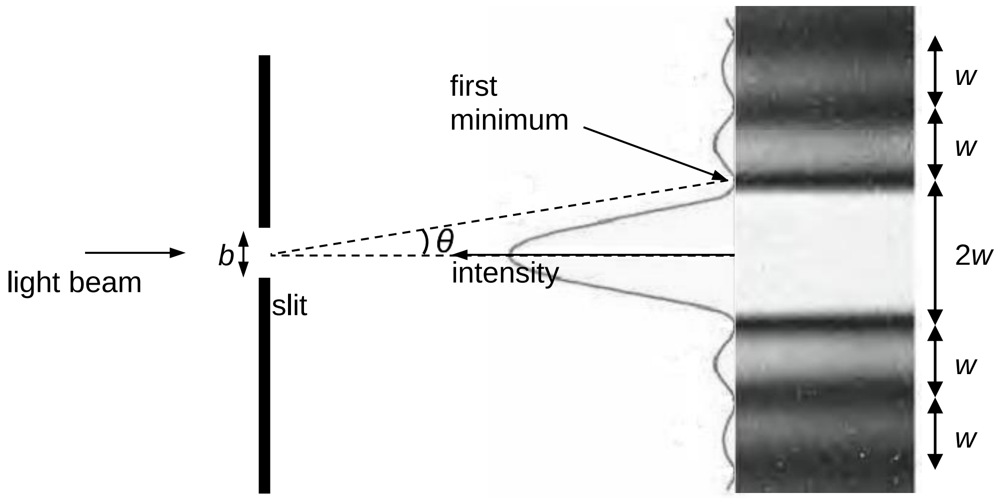
- The energy is redistributed in space, such that the diffraction pattern shows a bright central maximum wider than the slit, with narrower and less intense secondary maxima on both sides. The direction of spreading is perpendicular to the orientation of the slit.





- The **intensity distribution** of the diffracted light beam is shown below:

Note that the width w of the secondary maxima are the same, and are half the width of the central maximum.



Additional Notes

- When light beam of wavelength λ passes through a single slit of width b, the angle θ between the perpendicular to the slit and a line from the center of the slit to the first minimum is given by

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

- The general condition for destructive interference for single slit diffraction is given by

$$b \sin\theta = m\lambda \text{ where } m = \pm 1, \pm 2, \pm 3 \ldots \text{ (order of minima)}$$

Like in two-source interference, we can use superposition to explain the positions of the minima in the single slit diffraction pattern. But, in two-source interference, there are two sources producing waves that can interfere and undergo superposition. How can a wave produced by a single slit undergo superposition?

To appreciate how this can work out, you need to be introduced to Huygens' Principle (not in syllabus). According to Huygen's Principle, every point along a wavefront acts as a point source that produces waves coherently with other point sources along the same wavefront. For a single slit of a finite width, we can see it as having many point sources along the width of the slit.

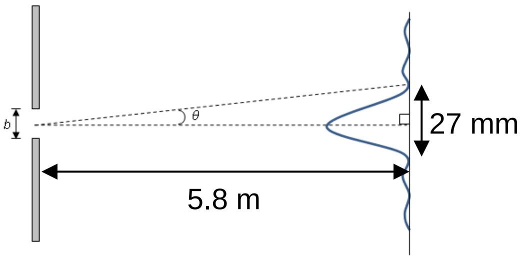
The waves from these point sources overlap and, when they meet on the screen with phase differences that cancels out each other's individual displacements, destructive interference occurs, resulting in minimum intensity.

You may refer to the appendix for a detailed derivation of the equation above. This derivation is not required by the syllabus.

Example 3

A laser light of wavelength 650 nm is passed normally through a narrow slit. A screen placed parallel to the slit 5.8 m away from the slit. An interference pattern is formed on the screen. The distance between the centers of the first minima outside the central bright fringe is 27 mm.

Determine the width of the slit.



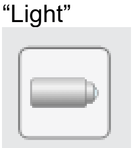
Hints
Method 1:
Find θ using trigo then use single slit formula to find b

Method 2:
Use the concept that for small θ, sin θ = tan θ

Additional Notes

Investigate a single-slit diffraction pattern.

Steps:
1. Select "Slits"
2. Select "Light"



3. Check the boxes for "Screen" and "Intensity".
4. Click the green button on the source.

Investigation 1: Keeping other factors constant, change the width. Observe how spreading and intensity change.

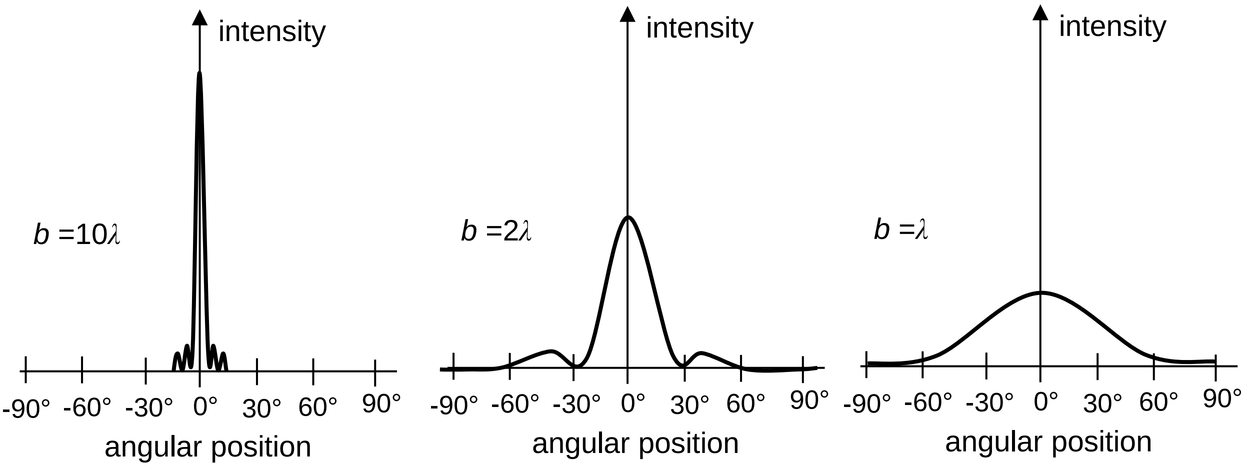
Investigation 2: Keeping other factors constant, change the wavelength. Observe how spreading and intensity change.

Note that the amplitude A of the resultant wave arriving at θ = 0 is proportional to the slit width b.

C.1.2 Variation of the single-slit diffraction pattern with slit width

- From the equation $\sin\theta = \dfrac{\lambda}{b}$, it is clear that as slit width decreases, the angle of the 1st order minimum increases.

- The diagrams below show the intensity of the wave detected for varying slit sizes. (Intensities are not drawn to scale.)



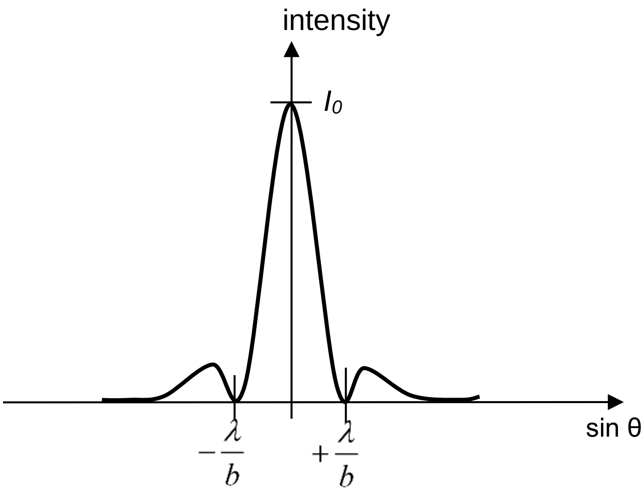
- Note that the degree of spreading of the wave as it passes through the slit is larger when the wavelength gets closer to the slit width.

- Also, the maximum intensity decreases with decreasing slit width.

Example 4

Light of wavelength λ incident on an adjustable single slit produces on a screen the diffraction pattern shown below. This pattern is obtained when the slit width is b.

Sketch with label on the axis the intensity pattern when the slit width is reduced to $\dfrac{b}{2}$.



Additional Notes

C.2 Diffraction is a general wave phenomenon

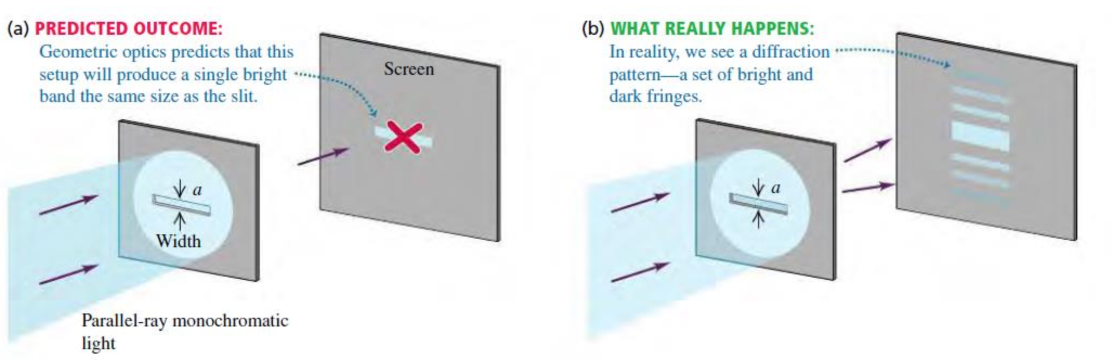


Diagram source: Hugh D. Young and Roger A. Freedman

In the previous section you learnt that when a wave passes through a single slit, a diffraction pattern is formed. This is really surprising and unintuitive – if we hadn't learnt this topic, we would expect the wave to pass through without spreading, and form a bright image on the screen that is exactly the same size as the slit. Except it does not happen.

This behaviour is called *diffraction*, and is a general wave phenomenon.

Key Term: Diffraction

Diffraction is the **spreading of waves** when they pass through an opening or round an obstacle. Diffraction effects are the greatest when the **width of the opening** is **comparable with the wavelength of the waves**.

Memorize

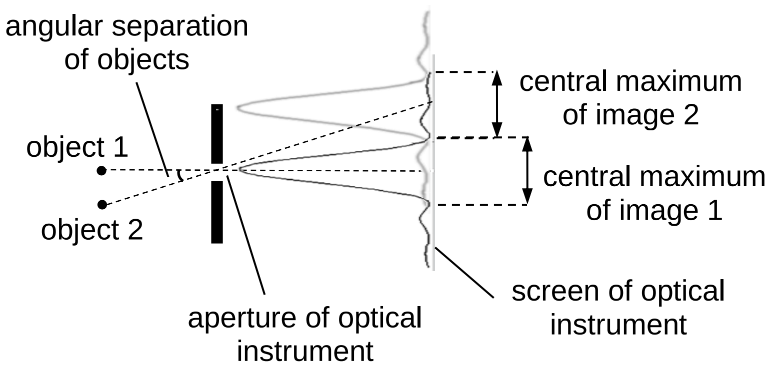
- The **speed, frequency and wavelength** of the waves **remain unchanged** after diffraction.
- Diffraction is an important characteristic of waves. It is **one of the defining characteristics of a wave**.

C.3 Resolving Power and Rayleigh Criterion

Diffraction has far-reaching implications for image formation by optical instruments, since light waves will diffract when they pass through a slit (eg. camera aperture or pupil of the eye) and form a diffraction pattern on the screen (eg. film/sensor of camera or retina of the eye). When we have two point objects, their images are not two points, but two diffraction patterns. The size of the central maximum of the diffraction pattern thus imposes a limit on resolution.

- When the objects are sufficiently spaced apart, the central maximum of their diffraction patterns do not overlap, so their images can be distinguished from each other, as illustrated below.

Angular separation of objects = angular separation of images



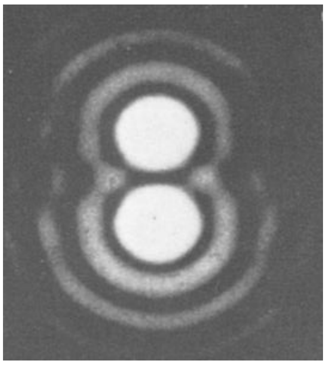


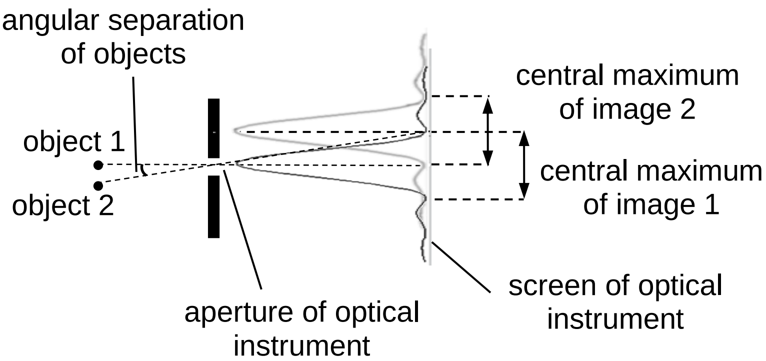
Diagram not drawn to scale. Typically, distance between objects and the aperture is greater than distance between aperture and screen.

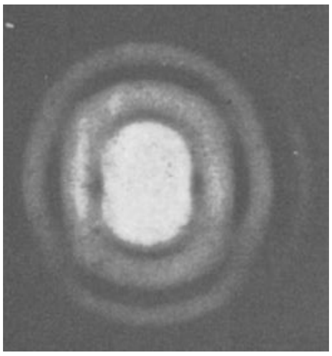
- As the angular separation between the objects decrease, either by moving the optical instrument further away, or moving the objects closer together, the central maximum of the two diffraction patterns starts to overlap, and their images are less distinguishable from each other.

Key Term: Rayleigh criterion

- The images of two point objects is considered just resolved (that is, barely distinguishable) when the **central maximum in the diffraction pattern of one image coincides with the first minimum of the diffraction pattern of the other image**.
- This is known as the Rayleigh criterion, and is the generally accepted criterion for the minimum resolvable detail.

Illustration of Rayleigh criterion





- Far away from the aperture, the angle at which the first minimum occurs, measured from the direction of incoming light, is given by the approximate formula (where θ is the angular separation of the objects, or that of the images) :

$$\sin\theta = \lambda / b$$

or, for small angle, simply

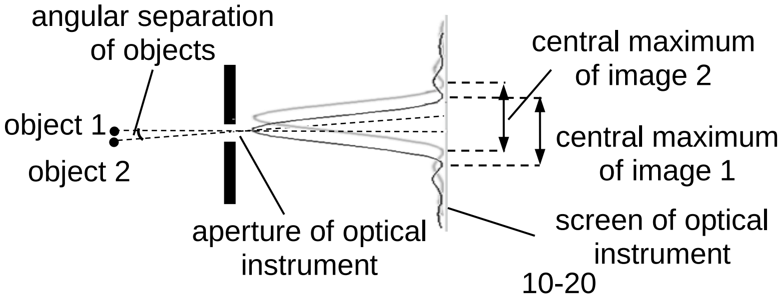
$$\theta \approx \lambda / b$$

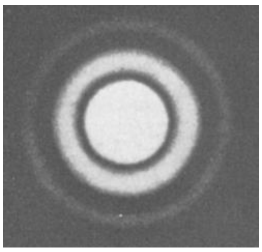
This θ is the limit of angular separation of the two objects; in other words, it is the smallest angular separation of the two objects such that their images are just resolved.

Key Term: Resolving power

- The minimum separation of two objects that can just be resolved by an optical instrument is called the limit of resolution of the instrument. The smaller the limit of resolution, the greater the resolution, or **resolving power**, of the instrument.
- Hence, in the above diagram, since the images are just resolved, the distance between the two objects is the limit of resolution of the instrument.
- Diffraction sets the ultimate limits on resolving power of lenses.

- When the angular separation of the objects is smaller than that given by the Rayleigh criterion (implying that the distance between the objects is smaller than the limit of resolution), the two images are not resolved.





Additional Notes

Improving Resolution

- Resolution is said to improve when θ decreases (less diffraction, narrower central maximum) as that means that the angular separation between the 2 objects can be smaller, but yet the images can still be resolved.

- Since $\theta \approx \lambda / b$, resolution improves:

  (i) With shorter wavelengths

  E.g. Ultraviolet microscopes have higher resolution than visible-light microscopes;

  (ii) With larger aperture sizes

  E.g. In a human's eye, the maximum pupil diameter is about 5 mm; in an eagle's eye it is about 9 mm.

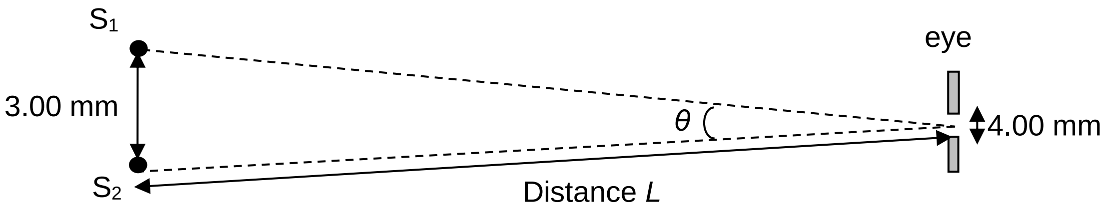
For distant point sources, $\because \dfrac{b_{eagle}}{b_{human}} = \dfrac{9}{5}$

$$\dfrac{\theta_{eagle}}{\theta_{human}} = \dfrac{5}{9}, \text{ i.e. better resolution of images on the eagle's eye}$$

Example 5

Under certain lighting conditions, the diameter of the pupil of a student's eye is 4.0 mm. With one eye, she is able to just resolve the images of two small light sources of wavelength 640 nm which are placed 3.0 mm apart at one end of a large assembly hall.

(a) Determine the maximum distance from which the student can just resolve the images of the two sources.



Step 1: For small θ, the distance S_1S_2 can be taken to be the same as the arc length S_1S_2 of part of a circle of radius L with centre of the circle at the eye. Using "s = rθ", find an expression for θ

Step 2: Use Rayleigh's criterion to find another expression for θ and solve simultaneously

(b) The sources are then replaced by another pair of sources, separated by the same distance but emitting light of wavelength appreciably less than 640 nm. Determine the direction the student should move so that she can again just resolve the images of the two sources.

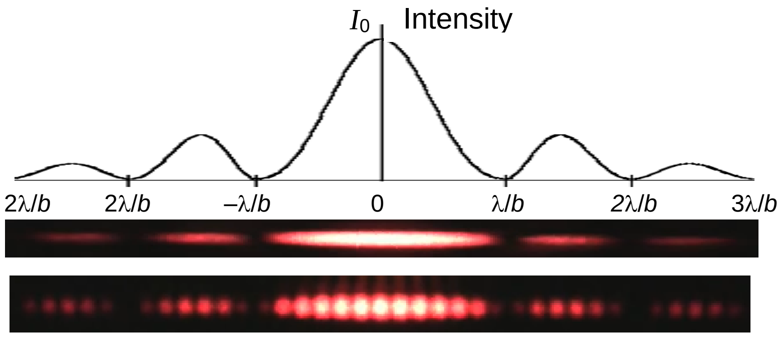
Additional Notes

C.4 Interference of light waves (Revisited!) – intensity of the double slit interference pattern

- In section C.1.1, we learnt that when a light wave passes through a single slit whose width is of the order of the wavelength of the light wave, a single slit diffraction pattern is observed, as shown below. We let the peak intensity of the central bright fringe be I_0.

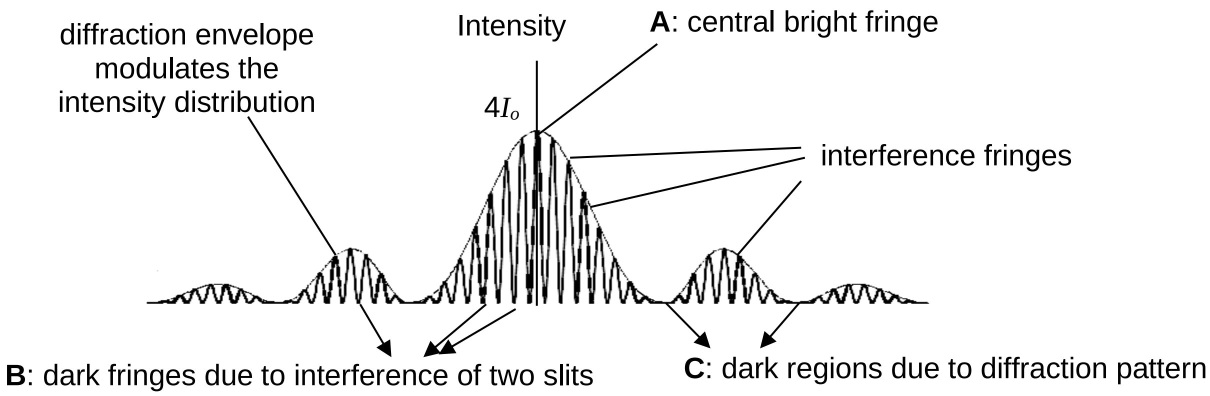
- Now we pass a light wave through two of such slits (a double slit), where the slit separation is small compared to the distance from the slits to the screen. The double slit diffraction pattern is shown right below the single slit diffraction pattern for ease of comparison. What do you observe in terms of the positions of the dark fringes of the single slit diffraction pattern?

Intensity of single slit diffraction pattern



single slit diffraction pattern

double slit diffraction pattern



Intensity of double slit interference pattern

- As the intensity of the single slit diffraction pattern decreases from the central maximum towards the first minimum, the intensity of maxima of the double slit interference pattern decreases accordingly.
- When the single slit diffraction pattern reaches zero intensity, the maxima of the double slit interference pattern also becomes zero intensity – there is an absence of bright fringes at those positions. We can use the principle of superposition to explain the features of the intensity graph, as given below.

A. At the **central bright fringe**, waves from the two slits arrive in phase and interfere constructively. If both waves have amplitude A, the resultant wave will have amplitude A + A = 2 A.

Intensity is proportional to the square of amplitude of the wave. Hence, at the central bright fringe, the intensity of the resultant fringe is four times that of the individual waves, ie 4 I_0.

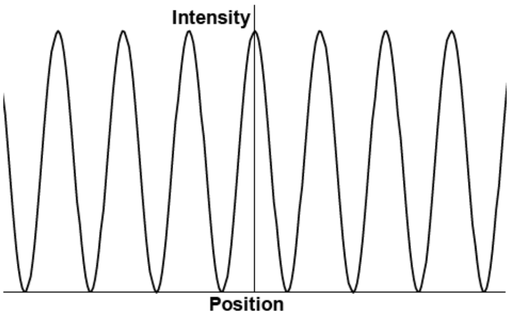
B. When wavefronts arrive in antiphase and interfere destructively, the resultant amplitude is zero. Hence, this gives rise to zero intensity at the dark fringes.

C. Due to single slit diffraction pattern, these are positions where the wave from each slit will arrive and interfere destructively to result in zero amplitude (ie. minima in single slit diffraction pattern). Hence, regardless of the path difference of the two waves from the double slits, the resultant amplitude at these positions is always zero. This gives rise to dark regions due to diffraction pattern.

Note: *Interference* and *diffraction* are both *superposition* effects, in that they result from the combining of waves with different phases at a given point. If the combining waves originate from a finite number of elementary sources, the process is called interference. If the combining waves originate in a single wavefront, the process is called diffraction. This distinction between interference and diffraction is a convenient one; both are superposition effects and usually both are present simultaneously.

Hence, when there are two or more slits, the intensity pattern on the screen is due to the combination of both the diffraction and interference effects.

Recall from earlier section ($\sin\theta = \lambda / b$), when the width of slits is decreased, the central bright fringe of the single-slit diffraction pattern broadens. When this happens for both slits, the interference fringes will be of approximately equal intensity.



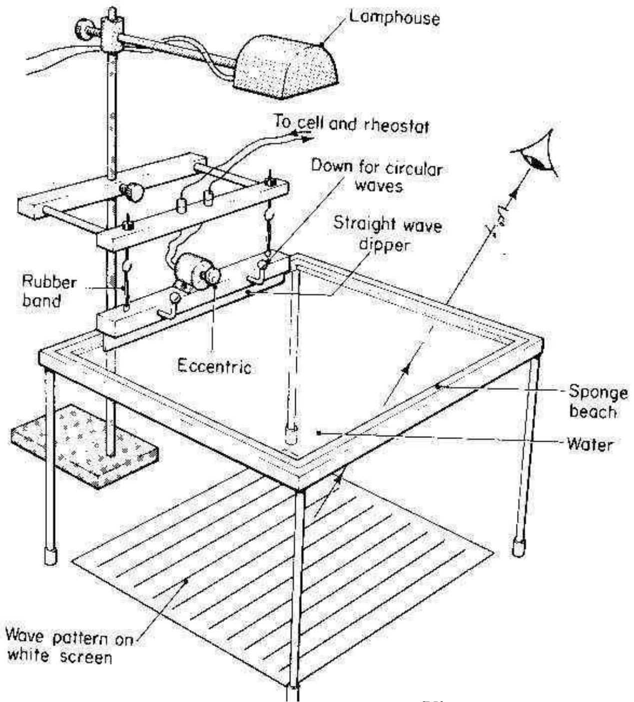
Intensity and interference pattern for double slits, with width of slits decreased

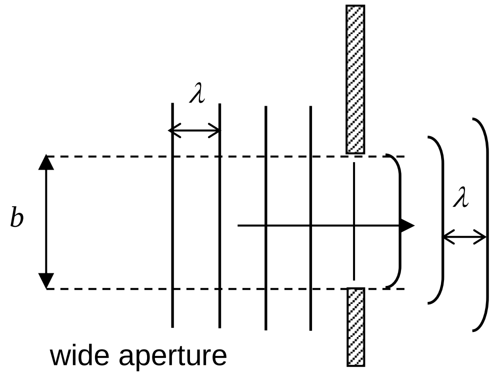
Additional Notes

C.5 Diffraction of water waves

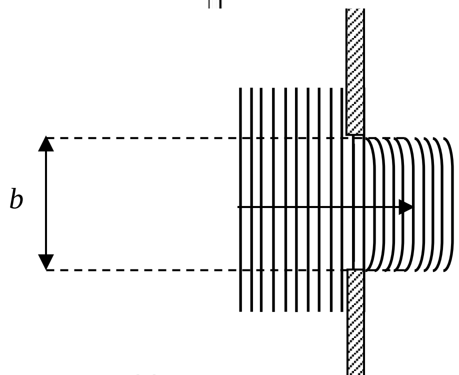
You learnt in previous sections that a wave undergoes diffraction when it passes through a slit. This effect can be demonstrated using by observing water waves in a ripple tank.

- A straight dipper produces plane waves, which can be observed on the screen below the ripple tank.

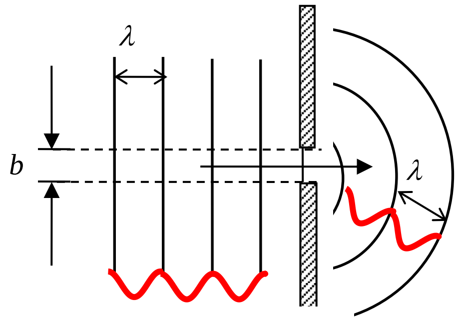




wide aperture
(slit width $b > \lambda$ but of same order of magnitude)
Diffraction can be observed



very wide aperture
(slit width $b >> \lambda$)
Diffraction not apparent



narrow aperture
(slit width $b \approx \lambda$)
Diffraction observed

- When plane water waves pass through a gap of width *a* much larger than their wavelength λ, there is no observable diffraction effect. As the width *b* starts to decrease, the ends of the emerging waves become rounded out at the ends. When *b* is made about the same size as the λ, the emerging waves are much more spread out and rounded. Hence, for observable diffraction effort, $b \approx \lambda$.
- For a more detail explanation of the spreading of waves as it passes through an aperture, refer to Appendix.

As the wave passes through the aperture, its wavelength remains unchanged. Hence, the distance between successive wavefronts should be the same after the wave passes through the aperture.

Scan to self-assess how well you understand single-slit diffraction 😊 SLS log-in required.

D Multiple Slit Diffraction – Diffraction Grating

D.1 Equation for diffraction grating

- The diffraction grating contains many closely spaced slits through which a beam is split.

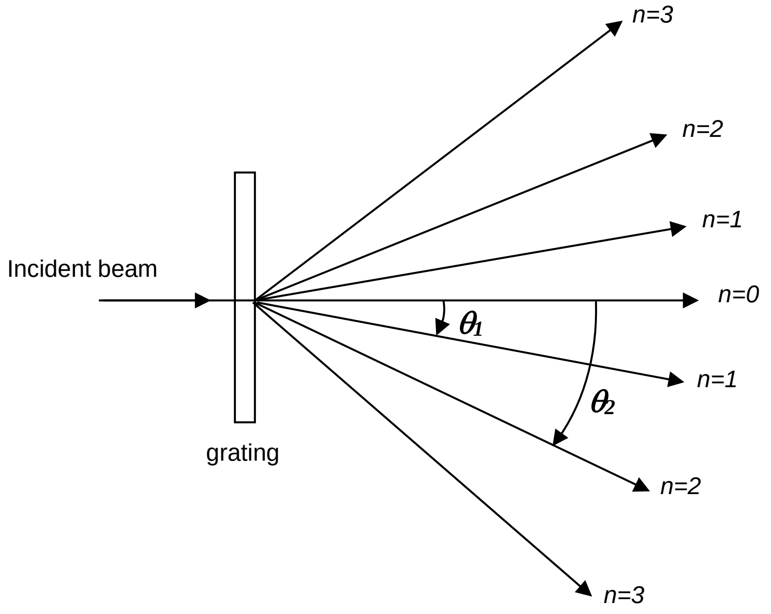
- The spacing between the slits is often indirectly indicated by indicating the number of slits in a certain length of grating, e.g. "600 lines per mm". There is often a need to determine the slit spacing otherwise known as slit separation.
  To find the slit separation, *d* for "600 lines per mm"
  $$d = \frac{1 \times 10^{-3}\ \text{m}}{600\ \text{lines}} = 1.67 \times 10^{-6}\ \text{m}$$

- The beams passing through the slits then interfere and produce a many-slit interference pattern. With a large number of slits, the intensity maxima become very well-defined (sharp and bright). However they are no longer evenly spaced.

- The position (or angle) of the n^{th} order intensity maximum may be determined using:

$$\boxed{d \sin\theta = n\lambda}$$

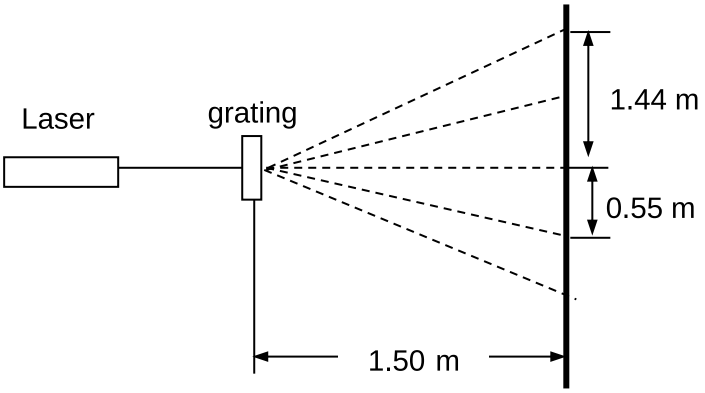
where d = separation between adjacent slits
n = order of diffraction, an integer
θ = angle between the n^{th} order beam and the normal to the grating
λ = wavelength of the incident beam



Additional Notes

Example 6

Monochromatic red light emitted by a laser passes normally through a diffraction grating as shown. The diffraction grating has 550 lines per mm. A large screen is set up 1.50 m from the grating. Bright spots are observed at 0.55 m and 1.44 m from the central bright spot.



Using measurements of the first-order diffraction, calculate the wavelength of the light.

Example 7

Monochromatic light of wavelength 600 nm is used in a spectrometer* to illuminate a diffraction grating set normally to the collimator. The grating has 3.0 x 10^5 lines per metre. The 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?

** refer to Appendix D on how a spectrometer works*

<u>Hint</u>
Make sin θ *the subject and the expression for* sin θ *must be* ≤ 1.

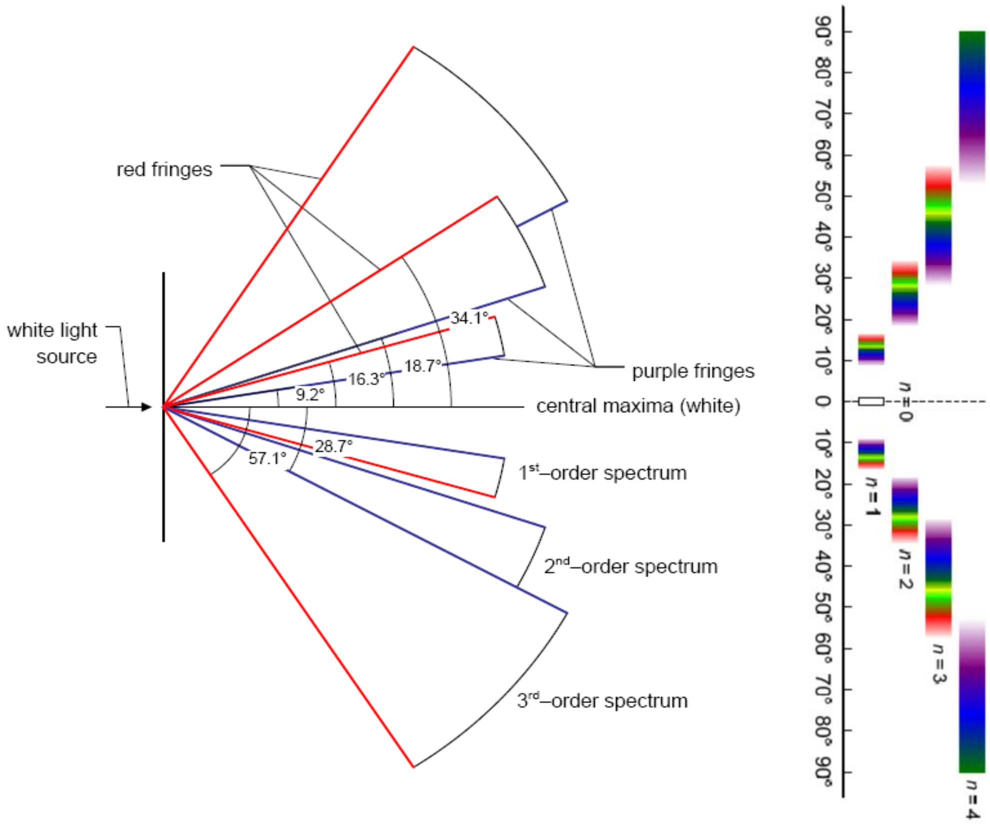
D.2 Diffraction of White Light

- Since $d \sin \theta_n = n\lambda$, for the same order n, waves with longer wavelength λ will be diffracted more than waves with a shorter wavelength. Thus red light (with $\lambda \approx 700$nm) will have a larger diffraction angle compared to blue light (with $\lambda \approx 400$nm).

- White light is a combination of light of different colours. When white light passes through a diffraction grating, a <u>spectrum</u> of all seven colours is obtained for each order, except at the zeroth order.

- Consecutive higher-order spectra can also overlap.

To gain better understanding, view the diagram in colours here.



Example 8

In the spectrum of white light obtained using a certain diffraction grating, the second order and third orders partially overlap. What wavelength in the third-order spectrum will appear at the angle corresponding to a wavelength of 650 nm in the second-order spectrum?

<u>Hint</u>
When two images overlap, which quantity is the same for both images?

Additional Notes

<u>Relating Science and Society</u>

In this Section, we learnt that if we pass light of a known wavelength through a grating of unknown grating spacing, we can gain knowledge of the unknown grating's structure (grating spacing) by analysing the diffraction pattern (angles of diffracted rays) and applying the known relationship between wavelength of light, diffraction pattern and structure of the grating.

Similarly, one could pass X-ray of known wavelength through molecules of unknown structure, and gain knowledge of its structure by studying the diffraction pattern of the X-ray. This was done by Rosalind Franklin in 1953 on DNA molecules, thus making a ground-breaking first evidence of the helical structure of the DNA molecule. X-ray diffraction went on to play an important role in the subsequent advances in molecular genetics.

E Stationary Waves

- In the previous topic of Waves, we have seen how energy can be transmitted by a progressive wave. Now we shall see how superposition of two waves can cause interference such that energy is localised despite wave motion – these waves formed are called stationary waves or standing waves.
- The wave is termed "stationary" because although the particles still oscillate, the wave profile does not seem to propagate.
- Energy is stored rather than transmitted along the wave.

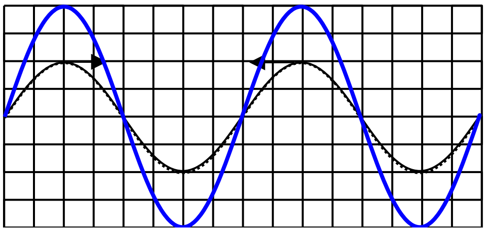
E.1 Formation of Stationary Waves

- A stationary wave is formed when 2 identical progressive waves (same amplitude, frequency and speed) travelling in opposite directions towards each another superpose, resulting in the formation of maxima (antinodes) and minima (nodes).

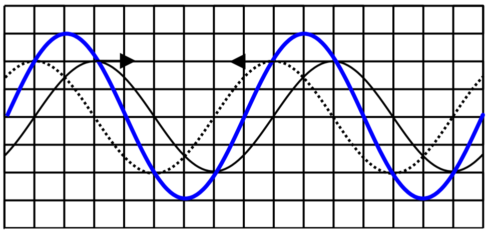
E.2 Graphical Representation of the formation of a Stationary Wave

Access to simulation of stationary waves

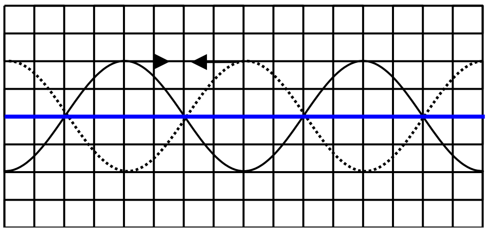
- The following diagrams show the resultant wave observed at intervals of $\frac{T}{8}$ when 2 similar waves travelling in opposite directions superpose.



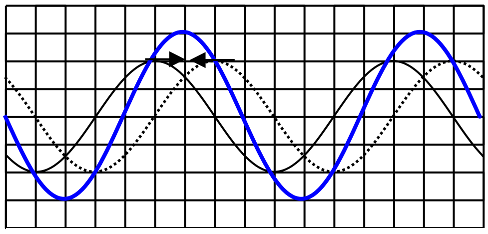
t = 0



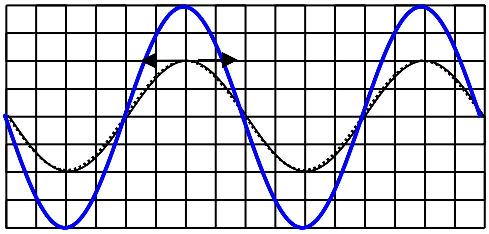
t = T/8



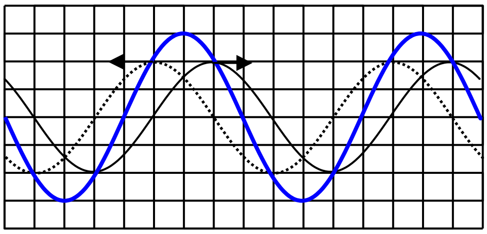
t = 2T/8



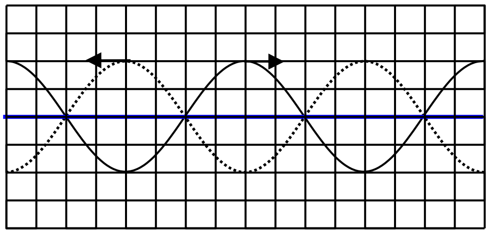
t = 3T/8



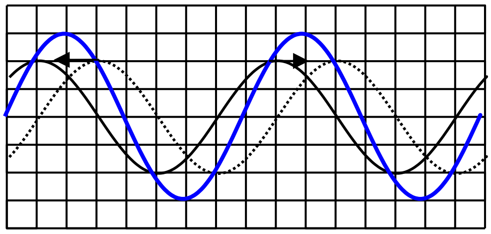
t = 4T/8



t = 5T/8



t = 6T/8



t = 7T/8

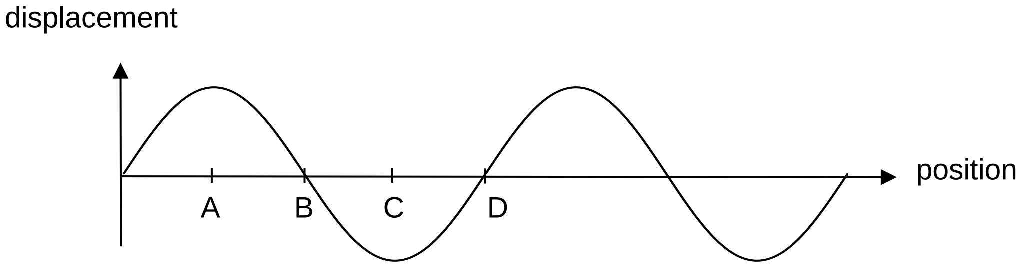
Key Terms

- A **node** is a point on a stationary wave where the amplitude of the vibration is zero or minimum
- An **antinode** is a point on a stationary wave where the amplitude of the vibration is maximum.

Additional Notes

Check your understanding 2 (solution and explanation in SLS).

The displacement-position graph of a stationary wave is shown below. The solid line represents the resultant wave at time t = 0, where displacement at each position is at its maximum.



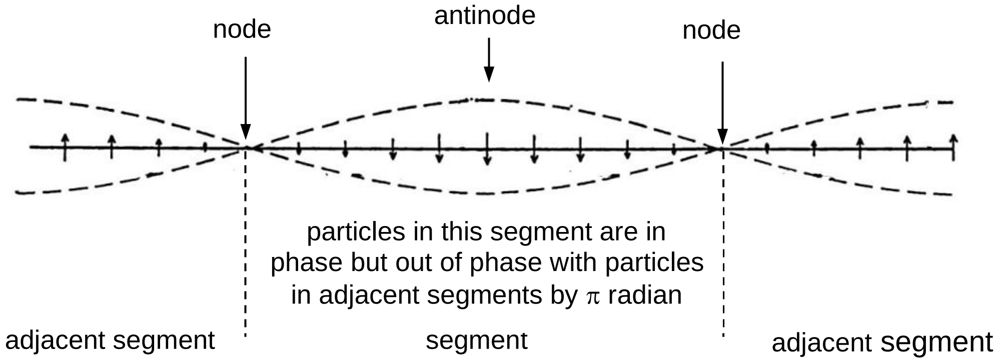
(a) Sketch the shape of the wave when
- **(i)** $t = T/4$,
- **(ii)** $t = T/2$,
- **(iii)** $t = (7/8)\ T$, where T is the period of the wave motion.

(b) Label these lines with their respective t value.

(c) Label the nodes with N, and antinodes with A.

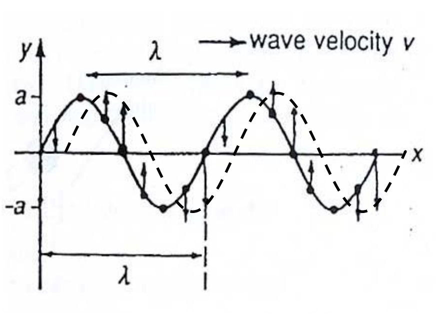
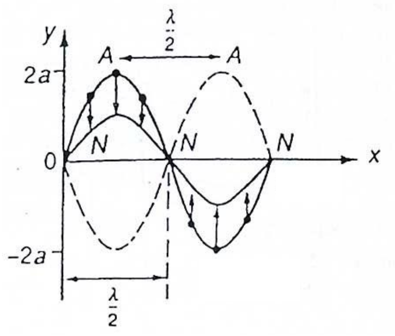
E.3 Properties of Stationary Waves

1. Stationary waves do not transfer any energy through the medium.
2. The nodes and antinodes are alternately situated. The distance between any two successive nodes (or antinodes) is constant and equal to half the wavelength. The distance between a node and the adjacent antinode is equal to a quarter wavelength.
3. All the particles in the medium, except those at the nodes, vibrate in simple harmonic motion with a time period equal to that of the component waves.
4. During each vibration, all the particles simultaneously pass through their mean positions twice, with maximum velocity which is different for different particles.
5. The amplitude of vibration increases gradually from zero to maximum from a node to an antinode.
6. The nodes divide the medium into loops or segments of equal lengths.
7. At any instant, all the particles in one segment are in phase. The particles in one segment are out of phase with the particles in the adjacent segment by π radian.



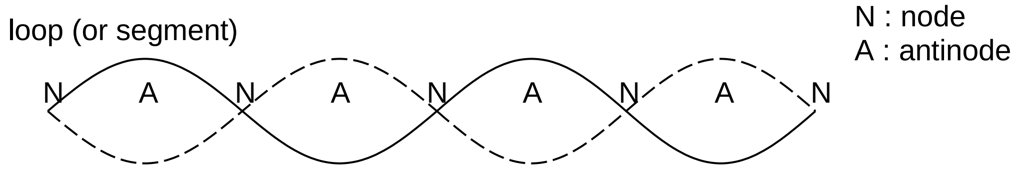
Additional Notes

E.4 Comparing Properties of Progressive and Stationary Waves

	Progressive Wave	Stationary Wave
Displacement-position graphs		
Waveform	• Advances in the direction of energy transfer of the wave at the velocity of the wave.	• Does not advance.
Energy of wave	• Energy is carried along in the direction of wave propagation.	• Although there **is** energy associated with the wave, **no** energy is carried **along** the wave. • The energy is stored as the kinetic and potential energy of the vibrating particles.
Amplitude of oscillation of individual particles	• Same for all particles in the wave regardless of position (assuming no energy loss), equal to the amplitude of the wave.	• Varies according to their positions • Particles at **A** (the antinodes) oscillate with maximum amplitude • Particles at **N** (the nodes) do not oscillate and have zero amplitude
Wavelength	• Distance between <u>any 2 adjacent points</u> on the wave with the <u>same phase</u>.	• Twice the distance between 2 <u>adjacent</u> **A** or **N**. • Equal to the wavelength of the component waves.
Phase of wave particles in a wavelength	• Wave particles have <u>different</u> phases (0 to 2π) within a wavelength.	• Within 2 adjacent **N**, all particles have the same phase. • Particles in adjacent segments are π radians out of phase.

Do note that the solid line represents the resultant waveform at a certain moment in time when the particles are at their maximum displacement and the dotted line represents the resultant waveform half a period later.

Stationary waves are represented by the diagram below. When sketching a diagram to represent one, both the solid and dotted lines should be drawn.



Additional Notes

E.5 Reflection of Waves at a Boundary

- We learnt that a stationary wave is formed when 2 identical progressive waves (same amplitude, frequency and speed) travelling in opposite directions towards each another superpose. How can we produce 2 such identical progressive waves?
- One possible way for this to occur is when a wave is reflected at a boundary.
- When a progressive wave encounters a boundary, reflection occurs.
- The reflected wave has the same frequency, speed and amplitude as the incident wave and travels in a direction opposite to that of the incident wave.
- Depending on the nature of the boundary, the reflected wave is either in phase or in antiphase when compared to the incident wave.
- Upon hitting a fixed and rigid boundary, the reflected wave is 180° out of phase with the incident wave.
- Upon hitting an open boundary, the reflected wave is in phase with the incident wave.

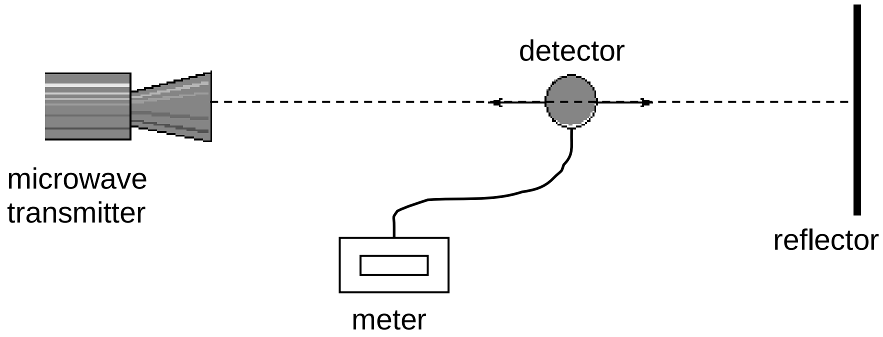
Scan for simulation of wave reflection at fixed / open boundary.

In the next section, we shall look at how stationary waves can be formed using microwaves, strings and air columns through reflection of a progressive wave at a boundary.

E.6.1 Demonstration of Stationary Waves using Microwaves, String Waves and Air Columns

Microwaves

The experimental set up is as shown below.

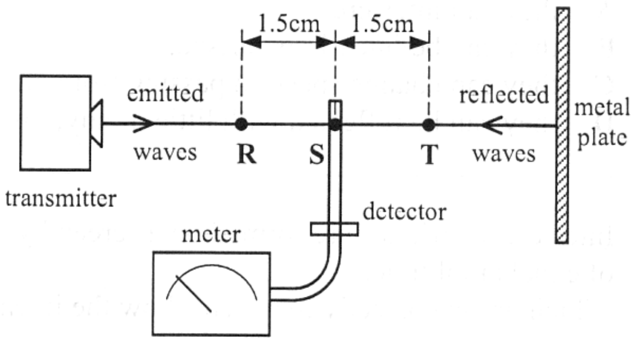


- As the detector moves along the path, the strength of the received signal will fluctuate between a series of maxima (antinodes) and minima (nodes).
- Measure the average distance between pairs of minima (two 'lows' of radiation or nodes).
- The wavelength of the microwave is twice the distance between adjacent minima.

Note: Microwaves are a form of electromagnetic radiation with wavelengths ranging from one metre to one millimetre.

Example 9

The given diagram shows a detector placed between a transmitter of microwaves and a metal plate. At three adjacent points, R, S and T, the meter shows zero intensity.



Calculate the frequency of the emitted waves.

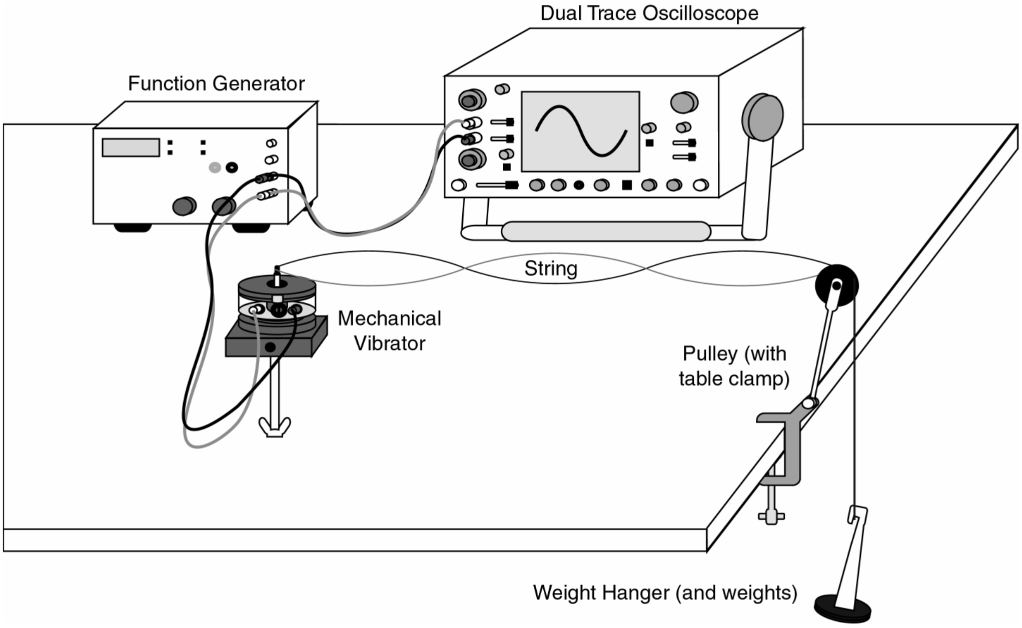
Hint
Find λ using concept taught.

Microwaves is part of the electromagnetic spectrum, hence you know the speed and can use $v = f\lambda$ to solve

Additional Notes

E.6.2 String Waves

The experimental set up is as shown below.



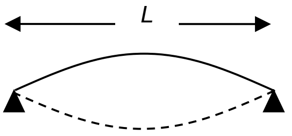
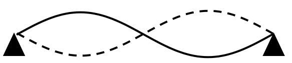
- A string is attached to a mechanical vibrator at one end and a weight hanger at the other, stretching over the pulley which is clamped to the bench.
- The stretched string can be made to vibrate using the mechanical vibrator connected to a function generator, from which the frequency and amplitude of vibration can be adjusted. The function generator output is fed into an oscilloscope so that the time varying voltage can be displayed.
- Not all frequencies will produce an observable stationary wave. This is because the formation of stationary waves is closely linked to the concept of resonance.
- To obtain observable stationary waves on a string, resonance must occur, where the driving frequency of the vibrator is equal to the resonant frequencies of the string. For a fixed tension (weight hanger), the resonant frequencies depend on the length of the string.
- When the frequency of vibration is slowly increased from a low value until it reaches the resonant frequency, a stationary wave can then be seen (see table on next page). This is called the fundamental mode of vibration. This lowest frequency at which a stationary wave is formed is known as the **fundamental frequency**.
- When the frequency of vibration is further increased, stationary waves can also be formed at higher frequencies – these modes of oscillation are known as harmonics. These different modes of oscillations may occur simultaneously (though not with the same intensity).

For cases like the stretched string, in order for observable stationary waves to be formed, not only must the conditions in E.1 be met, the driving frequency must be equal to the resonant frequencies (or we can say the boundary conditions must be met, where the two ends of the string form a node).

Additional Notes

Stationary Wave on a Stretched String

Do not memorise these formulae – you should learn to derive them on the spot when needed.

Harmonic series	Mode of vibration	Wavelength	Frequency
1		$\lambda_1 = 2L$	$f_1 = \dfrac{v}{2L}$ fundamental
2		$\lambda_2 = L$	$f_2 = \dfrac{v}{L} = 2f_1$
3		$\lambda_3 = \dfrac{2L}{3}$	$f_3 = \dfrac{3v}{2L} = 3f_1$
n	All harmonics possible	$\lambda_n = \dfrac{2L}{n}$	$f_n = \dfrac{nv}{2L}$

Example 10

A taut wire is clamped at two points 3.0 m apart. It is plucked near one end and the first 3 resonance frequencies exist on the string.

(a) Calculate the fundamental wavelength.

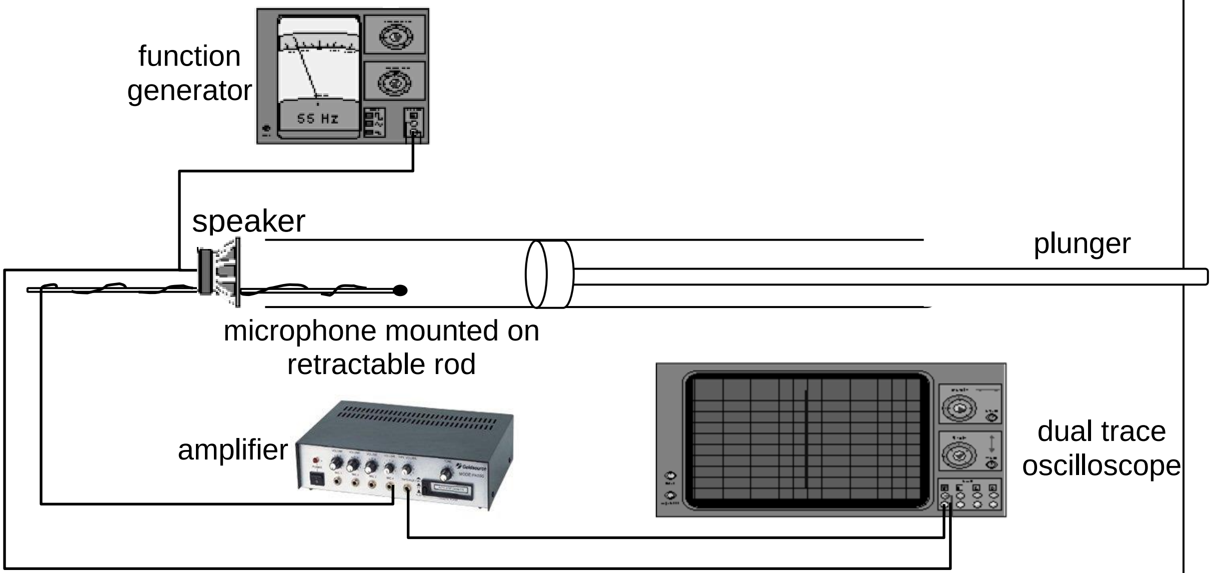
(b) Calculate the wavelengths for the next two resonant frequencies.

(c) Suggest a way to form a stationary wave that is predominantly the 2nd harmonic.

Additional Notes

E.6.3 Air Column

Sound waves travelling in an air column can be reflected off its ends. The incoming and reflected waves can form stationary waves within the air column. The experimental set up is as shown below.



- The speaker driven by a function generator produces a single frequency sound wave.
- The plunger that can be slid back and forth to vary the effective length of the column of air in the tube (to form a one-end open, one end closed pipe) or be completely removed (to form a both-ends open pipe).
- The frequency of the sound wave can be increased from a low value until it reaches the resonant frequency. A stationary wave will then form, and a loud sound will be heard. This is called the fundamental mode of vibration.
- Similarly, if the frequency is further increased, stationary waves can also be formed at higher frequencies.

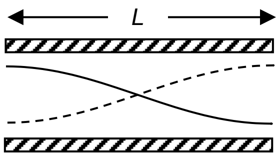
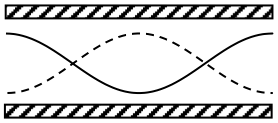
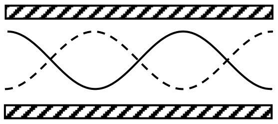
Note: A node forms at the closed end of the pipe, as there can be no vibration of air molecules. An antinode forms at the open end of the pipe, as the amplitude of vibration of air molecules is a maximum.

Stationary (Sound) Wave within a both-ends opened pipe (as <u>frequencies increases</u>)

- When the plunger is **removed**, the pipe is **open at both ends**. Sound can be reflected from the open ends and form a stationary wave within the pipe.
- Antinodes are formed at both ends of the open pipe.

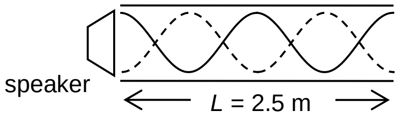
Do not memorise these formulae – you should learn to derive them on the spot when needed.

Actual antinodes occur at a short distance beyond the open ends of a pipe.This is known as an end correction. Refer to E.6.4.

Harmonic series	Mode of vibration	Wavelength	Frequency
1		$\lambda_1 = 2L$	$f_1 = \dfrac{v}{2L}$ fundamental
2		$\lambda_2 = L$	$f_2 = \dfrac{v}{L} = 2f_1$
3		$\lambda_3 = \dfrac{2L}{3}$	$f_3 = \dfrac{3v}{2L} = 3f_1$
n	All harmonics possible	$\lambda_n = \dfrac{2L}{n}$	$f_n = \dfrac{nv}{2L} = nf_1$

Example 11

An air column with both ends open is placed in a medium. A speaker producing sound of frequency 280 Hz is placed near the opening of the pipe, as shown in the figure below.

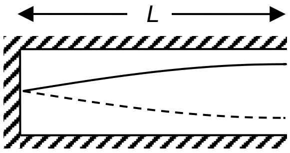
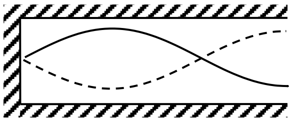
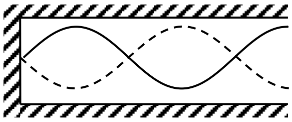


Calculate
(a) the speed of sound in the medium, and
(b) the fundamental frequency.

Additional Notes

Stationary (Sound) Wave within a one-end opened pipe (as <u>frequency increases</u>)

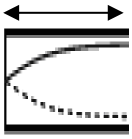
- When the plunger is **fixed at a particular position**, the pipe is **closed at one end** and **open at the other end**. Sound can be reflected from the two ends and form a stationary wave within the pipe.
- A **node** will be formed at the **closed end**, and an **antinode** at the **open end**.

Harmonic series	Mode of vibration	Wavelength	Frequency
1		$\lambda_1 = 4L$	$f_1 = \dfrac{v}{4L}$ fundamental
2	Not possible		
3		$\lambda_3 = \dfrac{4L}{3}$	$f_3 = \dfrac{3v}{4L} = 3f_1$
4	Not possible		
5		$\lambda_5 = \dfrac{4L}{5}$	$f_5 = \dfrac{5v}{4L} = 5f_1$
n (odd intergers)	Only odd-numbered harmonics are possible	$\lambda_n = \dfrac{4L}{n}$	$f_n = \dfrac{nv}{4L} = nf_1$

Stationary (Sound) Wave within a one-end opened pipe (as <u>effective length of air column increases</u>)

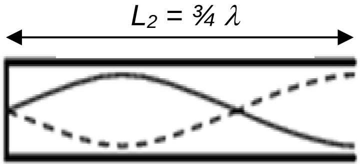
- The position of the plunger can be **shifted** such that the effective length of the air column L is varied until resonance is heard. The **<u>wavelength λ, and the frequency f</u>**, of the wave, **<u>are fixed</u>**.

$L_1 = ¼\ \lambda$



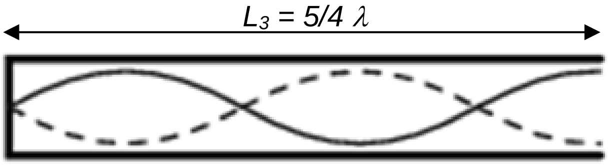
$L_1 = ¼\ \lambda \Rightarrow \lambda = 4L_1$

$L_2 = ¾\ \lambda$



$L_2 = ¾\ \lambda \Rightarrow \lambda = 4/3\ L_2$

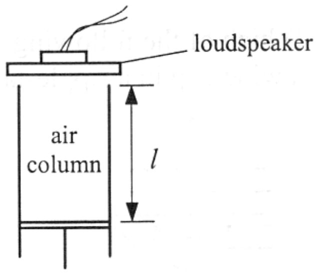
$L_3 = 5/4\ \lambda$



$L_3 = 5/4\ \lambda \Rightarrow \lambda = 4/5\ L_3$

Example 12

A note of constant frequency is produced by a loudspeaker placed above air column. The length *l* of the air column is slowly increased from zero. When *l* reaches 17 cm, the sound increases greatly in volume.

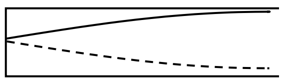
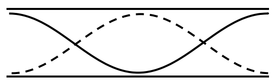
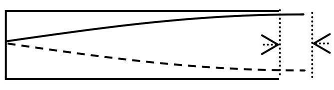
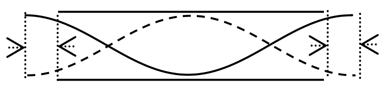


(a) Calculate the wavelength of the note produced by the loudspeaker.

(b) Calculate the next length *l* when the sound increases volume greatly.

E.6.4 End correction

- In the discussion on the modes of vibration of air column in a closed pipe and an open pipe, it was assumed that antinodes are formed exactly at the open end and the length of the pipe L was taken as the length of the air column.
- In practice, the air at the open end is free to move and hence the actual boundary (location of the antinode) extends a little beyond the open end.

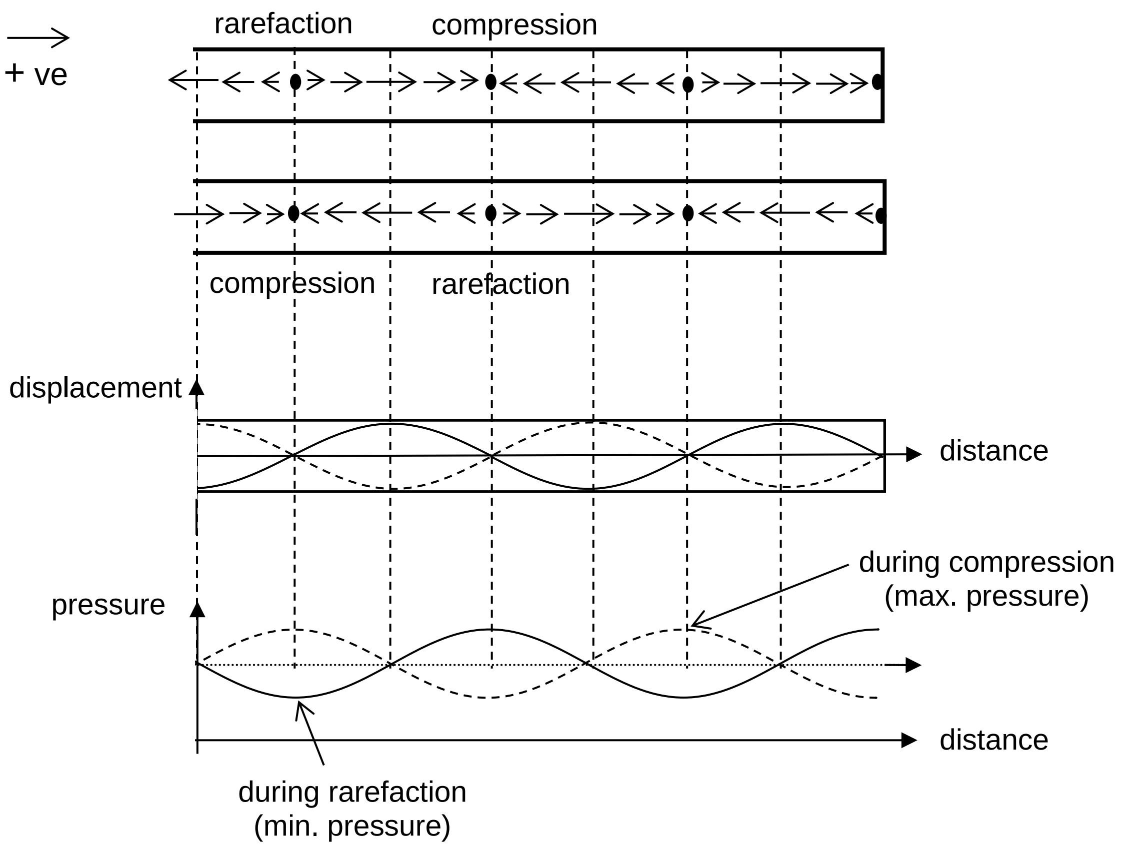
	Closed Pipe	**Open Pipe**
Theoretical	 $\lambda / 4 = L$	 $\lambda = L$
Experimental	 $\lambda /4 = L + \varepsilon$	 $\lambda = L + 2\varepsilon$

Additional Notes

E.6.5 Detection of Signals by microphone in a Stationary Longitudinal Wave (e.g. Sound wave)

- Consider a stationary sound wave (a longitudinal wave) being set up in a pipe with one end closed.
- When a stationary wave is formed in a pipe, we learnt earlier that a **loud sound would be heard/detected outside the pipe** as resonance has occurred.
- If a microphone is placed **within the pipe** and moved along the air column, the strength of the received signal will fluctuate between a series of maxima and minima.

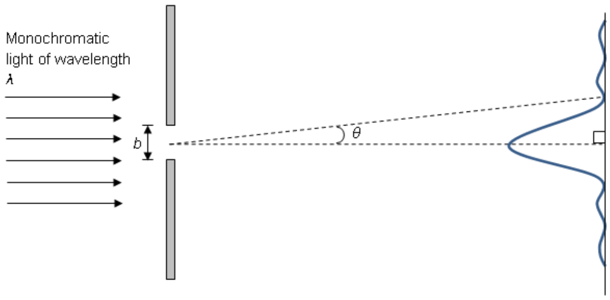
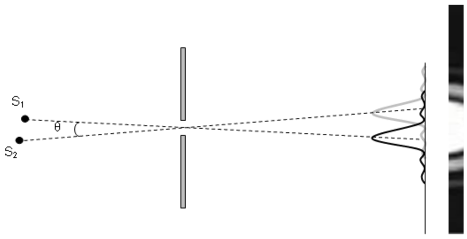
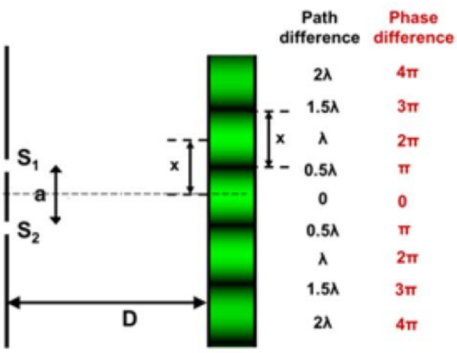
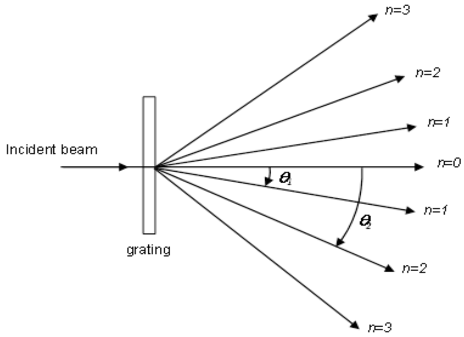
Scan for access to simulation. Observe the displacement of the air molecules as the stationary wave profile changes with time.



Do not memorise this information – learn to read the graph and understand how the graphs are related to each other.

- At the displacement node, the air molecules will be alternately squeezing that point during part of the cycle, and expanding away from it in another part of the cycle.
- The pressure variation at that point is the greatest, so a displacement node is also a pressure antinode.
- At the displacement antinode, the air molecules in the region will be moving with maximum displacement in the same direction with surrounding molecules, and are not changing their separation relative to the surrounding molecules.
- The pressure variation at that point is the smallest, so a displacement antinode is also a pressure node.
- Microphones work by detecting pressure variation in the air, instead of displacements of air molecules. The larger the pressure variation, the larger the signal detected. Thus, the maxima and minima detected by the microphone correspond to pressure antinodes and pressure nodes respectively.

Summary

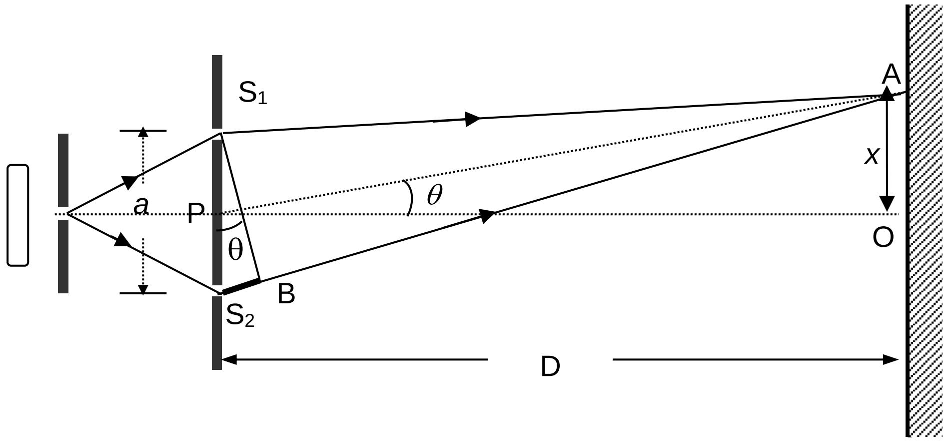
Phenomenon	Equation	Remarks
Single Slit Diffraction 	1st Order **Minimum** $\sin\theta = \dfrac{\lambda}{b}$	Note that unlike the diffraction grating, θ is the angle to the minimum. As the equation has sin θ, the angle may be in both degrees or radians.
Rayleigh's Criterion for Resolution 	Minimum angular separation for resolution $\theta \approx \dfrac{\lambda}{b}$	Note that the θ is angular distance between two objects. Note that θ is in radians.
Double Slit Interference 	Fringe separation when *a* << *D* $x = \dfrac{\lambda D}{a}$	
Diffraction Grating 	For **Maxima** $d\sin\theta = n\lambda$	Note that unlike the single slit, θ is is for the angle to the maximum. As the equation has sin θ, the angle may be in both degrees or radians.

Additional Notes

Principle of Superposition	The **Principle of Superposition** states that when two waves meet at a point, the resultant displacement is equal to the vector sum of the individual displacements.
Stationary waves	The wave formed when 2 identical waves (same amplitude, frequency and speed) travelling in opposite directions towards one another along a line superpose, resulting in regions of maxima (antinodes) and minima (nodes).
Diffraction	Spreading of waves when they pass through an opening or round an obstacle. Diffraction effects are the greatest when the width of the opening is comparable with the wavelength of the waves.
Interference	An effect that occurs when two or more waves overlap to produce a new wave pattern.
Coherent sources	Sources that produce waves with a constant phase difference at all times. The waves will have the same wavelength, frequency and speed.

Context	**How principle of superposition is applied**	**Patterns arising from interaction of waves**
Two-source interference	When waves arrive in phase – constructive interference When waves arrive anti-phase – destructive interference	• Interference pattern • Nodal/anti-nodal lines • High / low intensity regions • Fringe separation
Single slit diffraction	Slit of finite size made up of many pairs of point source, each producing coherent waves and undergoing interference.	• Single slit diffraction pattern • High / low intensity regions • Angular width of central fringe • Angular separation and resolving power • Single slit diffraction envelope for two source interference with sources of finite sizes (instead of point sources)
Multiple slits	Superposition of waves from many sources – at angle where waves from all sources undergo constructive interference, result in high intensity.	• Diffraction grating pattern – similar to antinodal lines • Diffraction grating equation
Stationary waves	When 2 waves travel in opposite direction, they interfere constructively and destructively at fixed points, giving rise to nodes and antinodes.	• Nodes/antinodes • Harmonics in strings/air columns

Appendix A: Derivation of Young's Double-Slit Formula



- Trigonometry enables us to derive a relationship between the distance from the slits to the screen (D), the slit separation (a), the distance between successive light or dark fringes (x) and the wavelength of light used (λ):

  Considering A to be a point of constructive interference,
  path difference between the two rays of light from S_1 and S_2 to A on the screen,
  $S_2A - S_1A = \lambda = S_2B$

  If angle $S_2S_1B = \theta$, then $\sin\theta = \dfrac{S_2B}{S_1S_2} = \dfrac{\lambda}{a}$

- But the diagram shows that $\sin\theta = \dfrac{AO}{AP}$.

  Since θ is very small (in practice AO is about 2 mm while OP is about 1 m) AP ≈ OP,

  $\sin\theta = \dfrac{AO}{OP}$.

  Since OP = D and AO = x, then $\dfrac{\lambda}{a} = \dfrac{x}{OP} = \dfrac{x}{D}$.

  Hence $\boxed{x = \dfrac{\lambda D}{a}}$

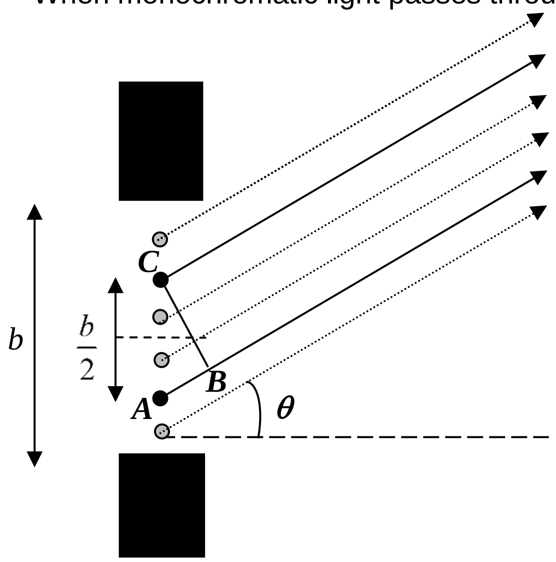
  Where x = the distance between 2 adjacent bright fringes (or 2 dark fringes) (m)
  a = slit separation (m)
  D = distance from the double slits to the screen (m)
  λ = wavelength of the light (m)

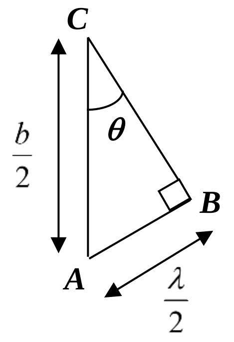
Additional Notes

Appendix B: Derivation of the 1st order minimum of a single slit

<u>Derivation of 1st order minimum of a single slit</u>

- When monochromatic light passes through a single slit of width b,





- At the slit, according to Huygens's principle, each portion of the slit will act as a source of light waves. These sources will cause an interference pattern.
- For simplicity, we are drawing 6 points, 3 in the upper half of the slit and 3 in the lower half of the slope.
- When the screen is very far away from the single slit, the light from the point sources may be considered to be parallel to each other.
- If we consider two points which are a distance of $\frac{b}{2}$ as indicated by the black dots,

  At the first minimum where destructive interference occurs, the waves from these two sources must have a path difference of $\frac{\lambda}{2}$.
- Path difference = AB = $\frac{\lambda}{2}$
- Since $\sin\theta = \frac{AB}{AC}$, we have $\sin\theta = \frac{\frac{\lambda}{2}}{\frac{b}{2}}$,

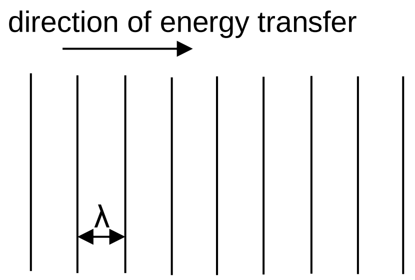
  Therefore, the expression given is

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

where b = slit width
λ = wavelength of monochromatic light
θ = angle between the 1st order minima and the normal to the slit.

Appendix C: Diffraction

- In representing progressive waves, a wavefront diagram is often used.



The figure above is an example of a plane progressive wave. The lines representing the wavefronts show locations where all particles along the wavefront are in phase. Usually for ease of thinking, the wavefronts show the crest (or trough) of the wave. The distance between each wavefront is the wavelength λ.

- According to Huygen's Principle, every point along the wavefront acts as a source of secondary wavelets (in other words every point along a wavefront acts as a point source), giving rise to their own outward-spreading circular wavelets. All these wavelets spread and *superpose* to produce the resulting wavefront. Figure (a) below is an example of imagining multiple point sources along the wavefront. (There should be an infinite number of point sources but 7 is sufficient to see the effect.)

Fig. (a)

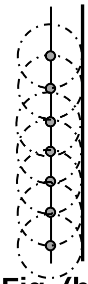
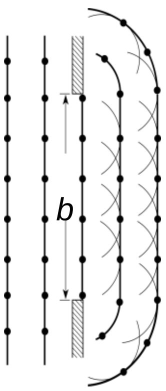


Fig. (b)

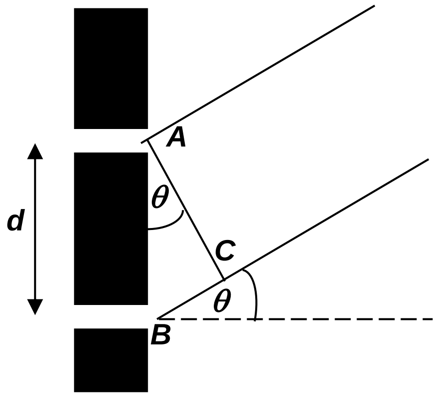
- Figure (b) above shows the effect of drawing how each point source sends the waves out radially (as seen by the circle), it is not difficult to see that the point sources still causes a tangential line whereby all points along that point are in phase. This means that the combined effect of the point sources will still be planar.

- However, when the plane waves path is blocked by a slit of width *b*, we will notice that in this case, the spreading of the wavefronts will occur. The curvature arises as a result of the missing contributions which are now blocked by the walls of the slit. The wave is now spread out by the restriction imposed by the slit. This spreading effect is known as diffraction.



Additional Notes

Appendix D: Diffraction Grating



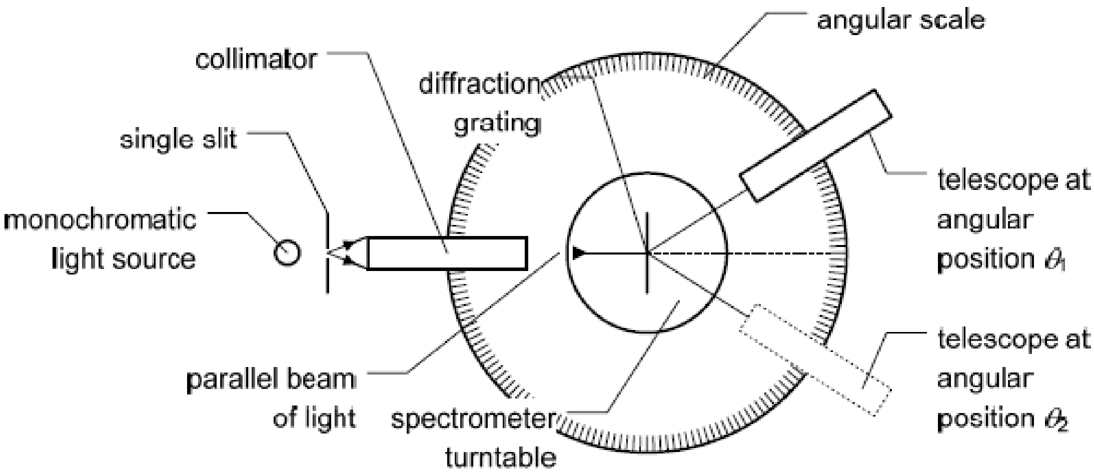
For maximum intensity, path difference BC must be an integral number of wavelengths.

Since $\sin\theta = \dfrac{BC}{AB}$, we have $\sin\theta = \dfrac{n\lambda}{d}$

$$\boxed{\text{path difference} = d \sin \theta_n = n\lambda}$$

where
- d = separation between adjacent slits
- n = order of diffraction, an integer
- θ_n = angle between the nth order beam and the normal to the grating.
- λ = wavelength

Determination of Wavelength of Light Using Diffraction Grating and Spectrometer



- A parallel beam of monochromatic light, after emerging from the collimator, is incident normally on the diffraction grating. As a result, bright fringes are produced at various angular positions.

- The telescope is first rotated such that the n^{th}–order bright fringe is at the centre of the crosswire and its angular position θ_1 is noted. It is then positioned on the opposite side of the normal to the grating and its angular position θ_2 is noted.

- The average diffraction angular position of the n^{th} –order bright fringe can then be obtained to a greater precision and accuracy.

Appendix E: Resolving power of your eyes

Source: https://stokes.byu.edu/teaching_resources/resolve.html

Refer to the section on **resolving power** in page 19 of the lecture notes.

When viewed at a distance, the two patterns in this page look identical, but as you approach them, there is a point at which you can barely resolve the lines and tell the difference between the two images. From this distance L, you can calculate the angular resolution of your eyes.

Or, just have fun, compare distance L with that of your peers or family members, and *see* the effects of diffraction come to life. ☺ Just remember, you learnt it in the topic of Superposition! :D