Superposition

A-Level Physics

Principle of superposition

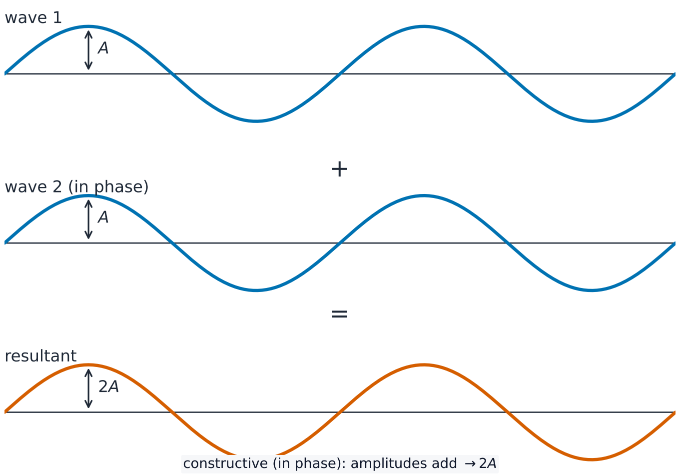
When two or more **waves** 波 overlap at a point, the **displacement** 位移 there is the **vector sum** 矢量和 of the displacements each wave would make on its own. This is the **principle of superposition** 叠加.

The waves pass through each other and come out unchanged. Superposition is the base of everything in this topic.

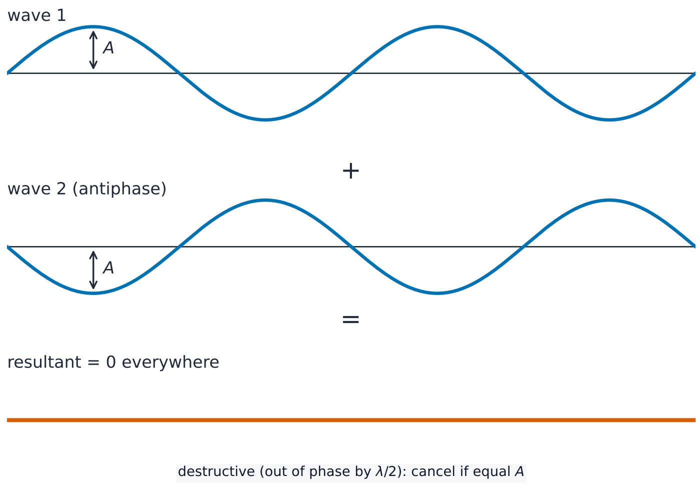
For the two-mark statement: **when two or more waves meet at a point, the resultant displacement is the sum of the displacements of the individual waves.** It is the displacements that add, not the amplitudes or the intensities, and the principle applies whenever waves of the same type overlap; the waves do not have to be coherent or of equal amplitude.

If two waves of **amplitude** 振幅 A_1 and A_2 meet:

- **in phase** 同相 (**crest** 波峰 meets crest): the amplitude is $A_1 + A_2$ (**constructive interference** 相长干涉).
- exactly out of phase (crest meets **trough** 波谷, **phase difference** 相位差 π): the amplitude is $|A_1 - A_2|$ (**destructive interference** 相消干涉).
- any other phase difference ϕ: the amplitude is somewhere between these two.



Two waves arriving in phase add to give double the amplitude (constructive)



Two waves arriving exactly out of phase cancel to zero (destructive)

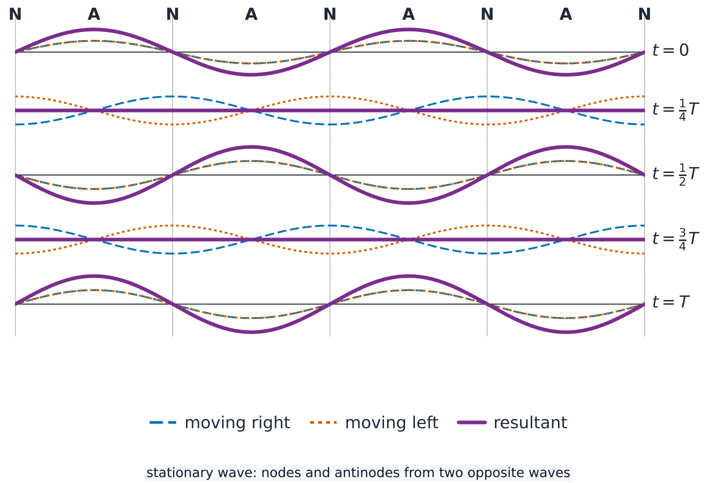
For **intensity** 强度, $I \propto A^2$. Two equal waves meeting in phase give intensity $(2A)^2 = 4A^2$ — **four times** the intensity of one wave alone.

Because $I \propto A^2$, work in amplitudes first. Two waves of intensities I and $4I$ have amplitudes A and $2A$: superposing in phase gives amplitude $3A$ and intensity $9I$, in antiphase amplitude A and intensity I. So two coherent waves of different intensities never cancel completely: the minima have intensity $(A_2 - A_1)^2$, not zero. A wave of amplitude $2A$ meeting one of amplitude $A/2$ travelling the opposite way gives a resultant that varies between $2.5A$ and $1.5A$, a stationary pattern with no true nodes. Doubling the amplitude of one of two equal waves that meet in phase takes the resultant from $2A$ to $3A$, so the intensity rises from $4A^2$ to $9A^2$: 2.25 times.

Stationary (standing) waves

When two identical **progressive waves** 行波 travel in **opposite directions** and overlap, they make a **stationary wave** 驻波. Examples: a wave on a string **reflected** 反射 from a fixed end overlapping the incoming wave; sound in an air column reflected from a closed end; microwaves between an emitter and a metal sheet.

Explain how the stationary wave is formed (three or four marks, asked for a string, an air column and microwaves alike): the wave from the source travels to the far end (the wall, the closed end, the metal plate) and is **reflected**; the incident and reflected waves have the same frequency, wavelength and speed and travel in **opposite directions**; they **superpose** where they overlap; at the points where they always meet in phase the displacements add to give the maximum amplitude, an antinode, and at the points where they always meet in antiphase they cancel, a node. For “state the conditions”: two waves of the same type, with the same frequency (and wavelength) and speed, travelling in opposite directions along the same line; equal amplitudes are needed only for the nodes to have zero amplitude. The metal plate in the microwave experiment is there to reflect the waves back along their own path.



A stationary wave forms where two opposite waves overlap (N marks a node, A an antinode)

Nodes and antinodes

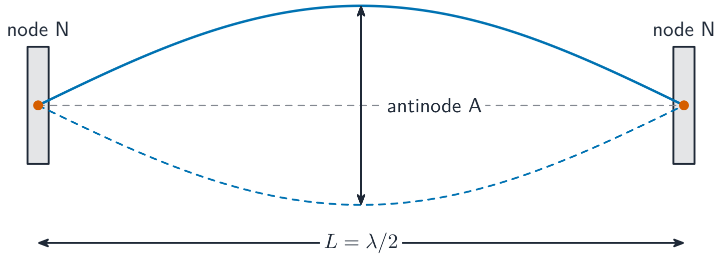
In a stationary wave:

- **node** 波节 — a point that is always at zero displacement (the two waves always cancel). The distance between next-door nodes is $\lambda/2$.
- **antinode** 波腹 — a point of largest amplitude (the two waves always add). The distance between next-door antinodes is $\lambda/2$.
- a node and the next antinode are $\lambda/4$ apart.

Particles between two nodes oscillate **in phase** with each other, but with different amplitudes (largest at the antinode, zero at the nodes). Particles on opposite sides of a node oscillate in **antiphase** 反相 (phase difference π).

So the phase difference between two points on the same loop (between adjacent nodes) is 0°; between points on neighbouring loops it is 180°; and between points on loops separated by one whole loop it is 0° again. Every point reaches its maximum displacement at the same instant and passes through zero at the same instant. A sketch of the

string a quarter of a cycle after the instant of maximum displacement is a straight line along the rest position; half a cycle later it is the mirror image of the first sketch. In one period a particle at an antinode travels four amplitudes, so a particle that moves 72 mm in one and a half periods has an amplitude of 12 mm.



Fundamental mode on a stretched string — one loop, with L equal to half a wavelength

Worked example. A string of length 0.80 m is fixed at both ends and vibrates in its fundamental mode, where the wave speed is 240 m s^{-1}. Find the fundamental frequency.

One loop fits the string, so $\lambda = 2L = 1.60$ m. Then

$$f = \frac{v}{\lambda} = \frac{240}{1.60} = 150 \text{ Hz}.$$

How a stationary wave differs from a progressive wave: it does not carry **energy** 能量 along its length, the pattern does not move along, and the nodes stay fixed; a progressive wave has the same amplitude everywhere and carries energy. The three differences the scheme lists: a progressive wave transfers energy and a stationary wave does not; in a progressive wave every point has the same amplitude, in a stationary wave the amplitude varies from zero at a node to a maximum at an antinode; in a progressive wave neighbouring points differ in phase, in a stationary wave all the points between two nodes are in phase.

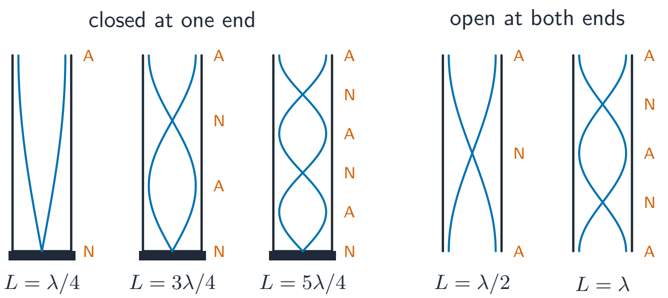
Measuring wavelength from node spacing

Drive a string with a vibrator at frequency f until a stationary pattern appears. Measure the distance between two well-separated nodes and divide by the number of half-**wavelengths** 波长 between them. Then λ is known, and $v = f\lambda$ gives the wave speed.

For a tube closed at one end and open at the other (a **resonance tube** 共鸣管), the closed end is a displacement node and the open end is a displacement antinode. The

fundamental 基频 has $L = \lambda/4$; the next resonance is at $L = 3\lambda/4$; and so on. For a tube open at both ends, both ends are antinodes; the fundamental is at $L = \lambda/2$.

The general rule: a closed end is a node and an open end an antinode, and a string fixed at both ends has a node at each end. So a closed pipe fits an odd number of quarter-wavelengths ($L = \lambda/4, 3\lambda/4, 5\lambda/4, \ldots$), while an open pipe and a fixed string fit whole half-wavelengths ($L = \lambda/2, \lambda, 3\lambda/2, \ldots$); a pipe open at both ends with n nodes has $n+1$ antinodes. Frequencies follow from $f = v/\lambda$: a closed pipe whose lowest note is 820 Hz has $\lambda = 4L$, and a corridor 13.2 m long with a reflecting door at each end resonates lowest at $f = v/2L = 330/26.4 = 12.5$ Hz.

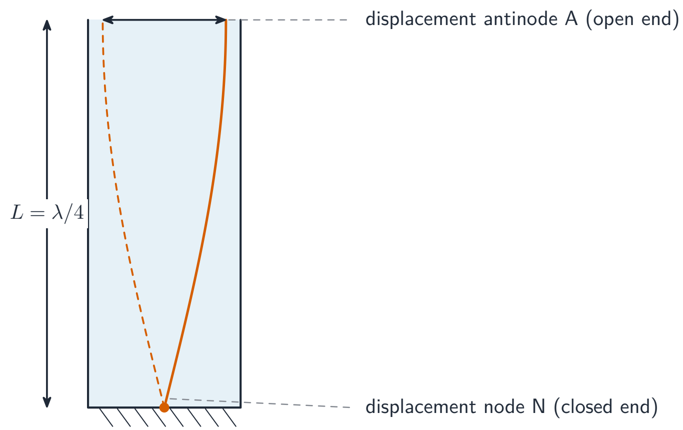


Modes of an air column: a closed end is a node, an open end an antinode

Worked example. A loudspeaker at the open end of a tube 0.51 m long, closed at the other end, sets up a stationary wave with two nodes and two antinodes. Find the wavelength and, with $v = 340$ m s^{-1}, the frequency.

Two nodes and two antinodes is the second mode of a closed pipe, $L = 3\lambda/4$, so $\lambda = 4 \times 0.51/3 = 0.68$ m and $f = 340/0.68 = 500$ Hz. The longest wavelength that can resonate in this tube is $4L = 2.0$ m.

The experiments the syllabus names all measure λ from node or antinode spacing. In the **resonance tube**, a tube is raised out of water while a loudspeaker or tuning fork sounds at its open end; the sound is suddenly loud at the first resonance, when the air column is $\lambda/4$ long, and again at $3\lambda/4$, so the tube moves $\lambda/2$ between the two (a student needs no equipment to detect the resonance: the note becomes loud). In a **dust tube**, fine powder settles into heaps at the displacement nodes, $\lambda/2$ apart. With **microwaves**, a receiver moved along the line between the transmitter and a metal reflecting plate reads a minimum every $\lambda/2$: a receiver that starts at a minimum and passes six more minima in 1.05 m has crossed 3λ, so $\lambda = 0.35$ m and, with $v = 340$ m s^{-1} for the equivalent sound experiment, $f = 970$ Hz. A stationary wave in a microwave oven melts chocolate at the antinodes, $\lambda/2$ apart, so the spot spacing and the oven's frequency (2.45 GHz) give the speed of light from $c = f\lambda$.



Fundamental mode in a closed pipe — a node at the closed end, an antinode at the open end

Diffraction

Diffraction 衍射 is the spreading of a wave after it passes through a gap or around an **obstacle** 障碍物. All waves diffract — water, sound, light, microwaves.

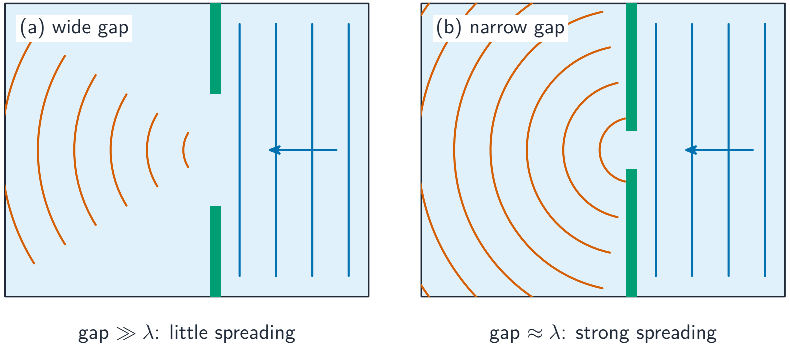
For the two-mark "state what is meant by diffraction": **the spreading of a wave as it passes through a gap (an aperture) or around the edge of an obstacle**, into the region behind it. Diffraction changes the direction the wave travels in, but not its speed, frequency or wavelength.

The amount of spreading depends on the ratio of **wavelength to gap width**:

- gap **much wider** than λ: very little spreading; the wave goes nearly straight through.
- gap **about the size of** λ: a lot of spreading; the wave fans out.
- gap **smaller than** λ: very strong spreading; the gap acts almost like a point source.

Show this with water waves in a **ripple tank** 水波槽: straight waves meet a barrier with a gap, and the waves curve more as the gap is made narrower. The same idea is why you can hear someone around a corner (speech has λ near 1 m, close to the gap size) but cannot see them (visible light has $\lambda \sim 500$ nm, far smaller than the gap).

For the strongest spreading the gap should be about one wavelength wide. So to increase the diffraction of a given wave, make the gap narrower; for a given gap, use a longer wavelength, which means a lower frequency. Radio waves of wavelength 1.5 km diffract around a mountain and reach an aerial behind it; microwaves of 1.5 cm do not. Sound of 0.44 kHz in air has $\lambda = 330/440 = 0.75$ m, so features of about 0.75 m diffract it most, and of the sounds passing through a doorway 0.80 m wide the low frequencies spread out most. Making the gap many wavelengths wide, or raising the frequency, reduces the spreading.

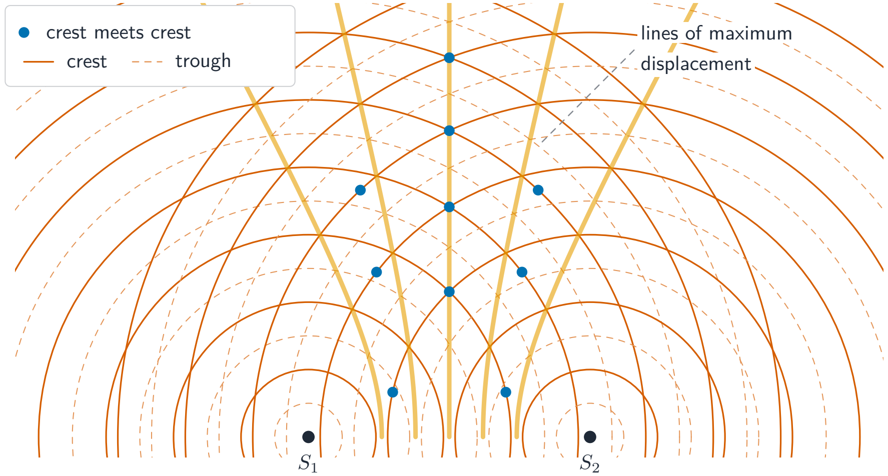


Diffraction in a ripple tank — a wide gap (a) spreads the waves little, a narrow gap (b) much more

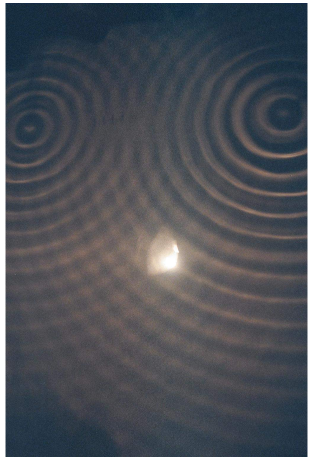
Interference

The shifting colours on a soap bubble come from the interference of light.

Interference 干涉 is the superposition of two **coherent** 相干 waves to give a steady pattern of high-amplitude regions (constructive) and low-amplitude regions (destructive).



Two coherent sources give lines of maximum displacement where crests meet crests



The same effect in a real ripple tank — two coherent sources give a steady interference pattern

Coherence

Two sources are **coherent** when they emit waves with a **constant phase difference** (which also needs the same frequency). Two separate lamps are not coherent — their phase changes randomly, so any pattern flickers too fast to see and you get only an average.

The one-mark definition: coherent waves have a **constant phase difference**, which requires the same frequency. They need not be in phase with each other: two coherent sources emitting 180° apart give a pattern whose central line is a minimum. Two separate lasers, or a lamp and a laser, are not coherent even when their frequencies happen to match, so no steady pattern forms.

To make coherent light from one source, pass it through two **slits** 狭缝 in a **double-slit** 双缝 setup. Both slits are lit by the same wavefront, so the two beams keep a fixed phase relationship.

Conditions for a clear pattern

To see two-source fringes you need:

1. two **coherent** sources (constant phase difference).
2. roughly equal amplitudes (or the dark regions are not very dark).
3. the waves overlap where you look.
4. for light (a transverse wave), the same plane of **polarisation** 偏振.

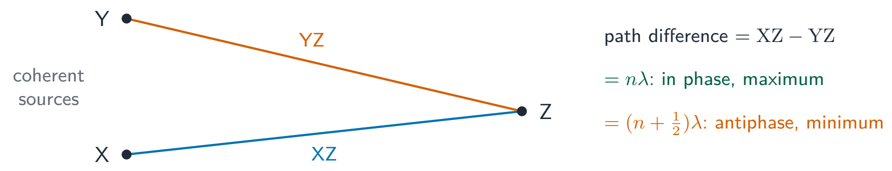
In practice, for light: a single slit (or a laser) makes the two slits coherent; the slits are narrow, so each diffracts the light into the region where the two beams overlap; and the slits are close together with the screen far away, so that the fringes are wide enough to see.

Path difference

For two coherent sources, what happens at a point depends on the **path difference** 路程差 Δx between the two waves arriving there:

- constructive: $\Delta x = n\lambda$ (for whole numbers $n = 0, 1, 2, \ldots$).
- destructive: $\Delta x = (n + \frac{1}{2})\lambda$.

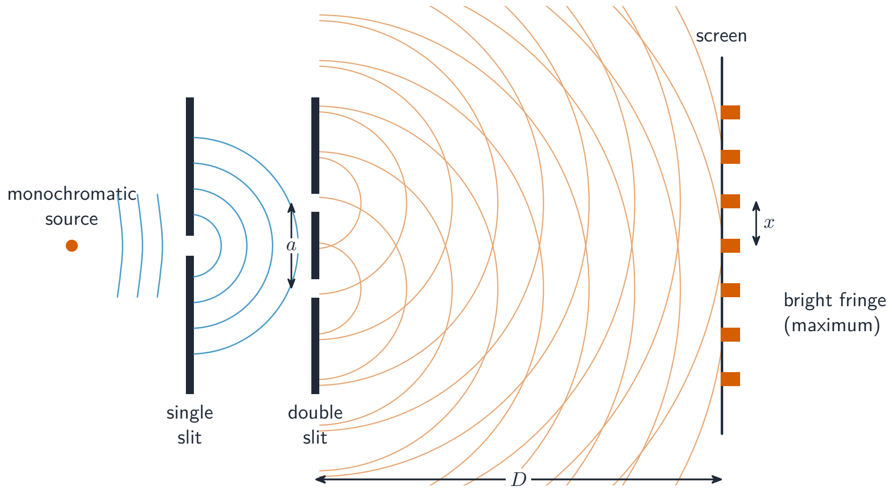
The path difference fixes the **phase difference**: one wavelength of path is 360°. Waves that have travelled 100 cm and 80 cm from two in-phase sources of wavelength 8.0 cm arrive with a path difference of 2.5λ, a phase difference of 180°, and cancel; microwaves of wavelength 4 cm whose paths differ by 6 cm (1.5λ) give a minimum, of zero intensity only if the two amplitudes are equal. Lowering both source frequencies equally lengthens the wavelength, so the same path difference is a smaller number of wavelengths and the point is no longer a minimum.



What happens at Z depends on the path difference XZ − YZ, measured in wavelengths

Double-slit (Young's) experiment

For two slits a distance a apart, with a screen a distance D away (assume $D \gg a$), light of wavelength λ makes fringes on the screen.



Young's double-slit experiment — the single slit makes the two slits coherent sources

The **fringe spacing** 条纹间距 x (one **fringe** 条纹 to the next) is

$$\lambda = \frac{ax}{D}, \qquad x = \frac{\lambda D}{a}.$$

Bright fringes (**maximum** 极大) are where the path difference is a whole number of λ; dark fringes (**minimum** 极小) where it is $(n+\frac{1}{2})\lambda$. The fringes are equally spaced.

To make the fringe spacing smaller: increase a (slits further apart), reduce D (screen closer), or use a shorter λ (bluer light).

Explain how the pattern of bright and dark fringes is formed (three marks): the light **diffracts** at each slit; the two diffracted beams **overlap** and superpose; where the path difference from the two slits is a whole number of wavelengths the waves arrive in phase and interfere constructively, giving a bright fringe, and where it is an odd number of half-wavelengths they arrive in antiphase and interfere destructively, giving a dark fringe. A brighter source makes the bright fringes brighter but does not change their spacing; changing to blue light makes the spacing smaller, so to keep the same spacing the slits must be moved closer together; making each slit narrower increases the diffraction, so fringes appear across a wider region, with the spacing unchanged.

Worked example. In a double-slit experiment the slits are 0.50 mm apart and lit by light of wavelength 600 nm. The screen is 2.0 m away. Find the fringe spacing.

$$x = \frac{\lambda D}{a} = \frac{(600 \times 10^{-9})(2.0)}{0.50 \times 10^{-3}} = 2.4 \times 10^{-3} \text{ m} = 2.4 \text{ mm}.$$

Worked example. Red light of wavelength 680 nm falls on slits 0.16 mm apart. The distance between the centres of the first and ninth dark fringes is 3.2 cm. Find the distance D to the screen.

Eight fringe spacings make 3.2 cm, so $x = 4.0$ mm and $D = ax/\lambda = (0.16 \times 10^{-3})(4.0 \times 10^{-3})/(680 \times 10^{-9}) = 0.94$ m. On a screen 5.0 cm wide centred on the pattern, with $x = 2.4$ mm, ten bright fringes fit on each side of the central one: 21 in all. A graph of x against a is a curve falling as $1/a$; a graph of x against λ (or against D) is a straight line through the origin with gradient D/a (or λ/a), from which a can be found.

Diffraction grating

A **diffraction grating** 衍射光栅 has many equally spaced slits — often hundreds or thousands per millimetre. Each slit is a coherent source. A maximum is seen at angle θ from the **normal** 法线 to the grating when

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

where d is the **slit spacing** 缝间距 (centre to centre), $n = 0, \pm 1, \pm 2, \ldots$ is the **order** 级次, and λ is the wavelength.

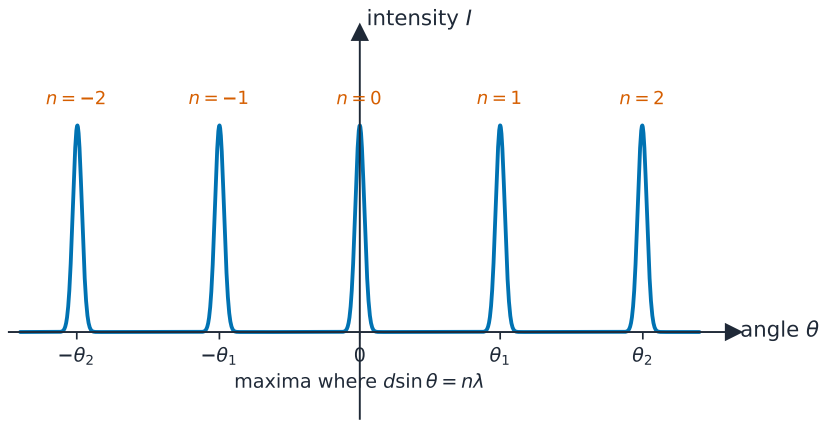
Worked example. A diffraction grating has 500 lines per mm. Light of wavelength 600 nm is shone normally on it. Find the angle of the first-order ($n = 1$) maximum.

The slit spacing is $d = \dfrac{1}{500}$ mm $= 2.0 \times 10^{-6}$ m, so

$$\sin\theta = \frac{n\lambda}{d} = \frac{600 \times 10^{-9}}{2.0 \times 10^{-6}} = 0.30 \quad \Rightarrow \quad \theta \approx 17°.$$

Compared with the double slit, a grating gives much **sharper** maxima, because more slits add together — every other direction is cancelled by many slits.

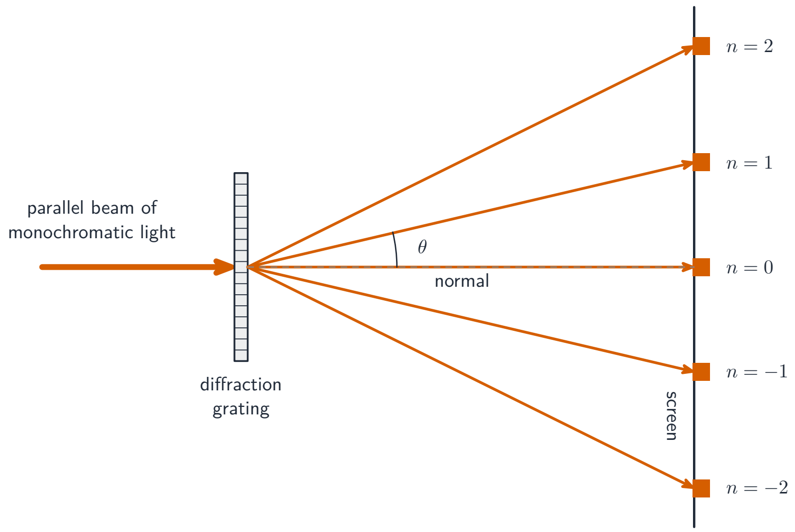
To **describe the diffraction at the grating**: the light spreads out (diffracts) at every slit, and the waves from all the slits superpose; in the directions where the path difference between neighbouring slits is a whole number of wavelengths they are all in phase, so sharp maxima form there and almost nothing in between. A graph of intensity against angle is a set of narrow peaks at $\theta = 0$ and at $\pm\theta_1, \pm\theta_2, \ldots$, where $\sin\theta_n = n\lambda/d$: equally spaced in $\sin\theta$, so slightly further apart in θ at the higher orders.



Intensity against angle for a grating: sharp maxima where $d \sin\theta = n\lambda$

Worked example. Light of wavelength 680 nm is incident normally on a grating with 450 lines per mm. Find the angle between the two second-order maxima.

$d = 1/450$ mm $= 2.22 \times 10^{-6}$ m, so $\sin\theta_2 = 2 \times 680 \times 10^{-9}/(2.22 \times 10^{-6}) = 0.612$ and $\theta_2 = 37.7°$; the two second-order beams are $2\theta_2 = 75°$ apart.



A diffraction grating splits monochromatic light into sharp maxima on a screen

Slit spacing from "lines per mm"

If a grating has N lines per millimetre, then $d = 1/N$ millimetres $= 10^{-3}/N$ metres. For 450 lines per mm, $d = 1/450$ mm $\approx 2.22\ \mu$m.

Highest order

For a given grating and wavelength, $\sin\theta = n\lambda/d$ cannot be more than 1, so the highest order seen is

$$n_{\max} = \left\lfloor \frac{d}{\lambda} \right\rfloor .$$

If $d/\lambda = 3.27$, orders up to $n = 3$ exist; $n = 4$ would need $\sin\theta > 1$ and is not seen.

So the total number of maxima on a wide screen is $2n_{\max} + 1$: for 700 nm light and 400 lines per mm, $d/\lambda = 3.57$, so $n_{\max} = 3$ and seven beams are seen. A shorter wavelength gives more orders at smaller angles. With white light every order except the zero order is a spectrum, violet nearest the centre and red furthest out, because θ grows with λ; the zero order stays white. Two wavelengths give a maximum at the same angle when $n_1\lambda_1 = n_2\lambda_2$: the third order of 400 nm coincides with the second order of 600 nm.

Finding λ with a grating

Shine parallel light of unknown wavelength straight at the grating. Measure the angle θ_1 of the first-order maximum from the centre. Then $\lambda = d\sin\theta_1$. Repeating for higher orders and averaging reduces error.

Measure the angle between the first-order maxima on the two sides and halve it, which cancels any error in setting the zero; higher orders give larger angles and so a smaller percentage uncertainty; and plotting $\sin\theta$ against n for several orders gives a straight line through the origin of gradient λ/d, so $\lambda = Gd$ (or, for a known wavelength, $d = \lambda/G$). Two things must be right: θ is measured from the **normal** to the grating, not from its surface, and d is the distance between adjacent lines, so 400 lines per mm means $d = 2.5\ \mu$m, never 400.

Definitions the examiner accepts

A definition question is marked against fixed wording. Learn these exactly, and give one answer only.

Term	Definition
principle of superposition	when two or more waves meet at a point, the resultant displacement is the sum of the displacements of the individual waves
stationary wave	the pattern formed when two progressive waves of the same frequency and speed travel in opposite directions and superpose, with nodes and antinodes that do not move
node	a point on a stationary wave where the displacement is always zero
antinode	a point on a stationary wave where the amplitude is a maximum
diffraction	the spreading of a wave as it passes through a gap or around the edge of an obstacle
interference	the superposition of waves from coherent sources, giving a steady pattern of maxima and minima
coherence	waves that have a constant phase difference (and so the same frequency)
path difference	the difference between the distances travelled by two waves from their sources to a point
fringe spacing	the distance between the centres of two adjacent bright (or dark) fringes

Term	Definition
order of a maximum	the whole number n in $d \sin\theta = n\lambda$, the number of wavelengths of path difference between adjacent slits

Exam tips

- Two-source interference: **constructive** when path difference $= n\lambda$, **destructive** when $= (n + \frac{1}{2})\lambda$; the sources must be **coherent**.
- Double slit: $\lambda = ax/D$; diffraction grating: $d \sin\theta = n\lambda$ — know every symbol.
- On a **stationary wave** mark nodes and antinodes; adjacent nodes are $\lambda/2$ apart; it stores energy but does not transfer it.
- A stationary wave needs two waves of the **same frequency travelling in opposite directions**.

Common mistakes

- Adding amplitudes or intensities in the principle of superposition. Displacements add; the intensity then follows from the resultant amplitude squared.
- "Adjacent nodes are one wavelength apart." Half a wavelength; a node to the next antinode is a quarter.
- Using $d =$ lines per millimetre in $d \sin\theta = n\lambda$. Invert: $d = 10^{-3}/N$ metres.
- Measuring θ from the grating surface, or forgetting that the angle between the two first-order beams is $2\theta_1$.
- Writing "in phase" for coherent. Coherent means a constant phase difference; the sources may be permanently out of step.
- Saying the fringe spacing changes when the source is made brighter, or that narrower slits change the spacing. Brightness and slit width change the contrast and the number of visible fringes, not $x = \lambda D/a$.