Waves for the DSE (and more)

Alfred

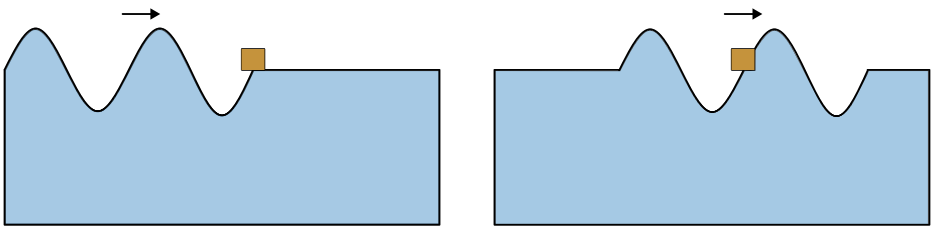
February 16, 2025

Contents

1 Introduction to waves

Ordinary objects, a ball, a bell, a stick, are regarded as matter, as in you can touch them. However, in real life there are also things that we can touch, think of heat, light, radio rays. With physics, these phenomena are described as waves.

Waves are oscillatory in nature, meaning that a point in the wave moves in a cycle. After a cycle, the point returns to its original position and repeats the same motion again.



(a) Original waveform　　(b) Waveform after 1 cycle

Figure 1: An object under the effect of water waves. The object returns to the equilibrium position after 1 cycle

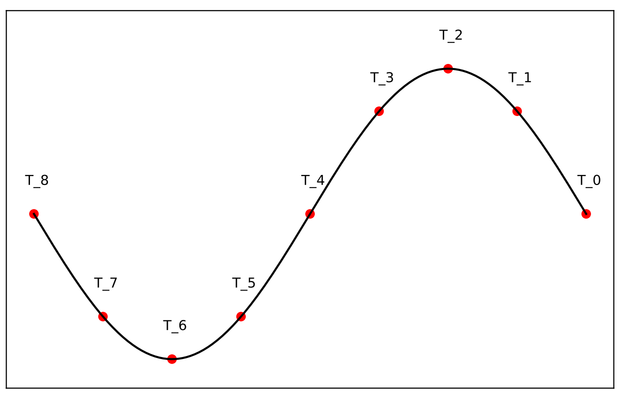


Figure 2: Vertical position of the object in figure 1 against time. Note that the wave moves towards the right, so the object first moves down and then back up

1.1 Terminology

To have a better understanding on waves, we first need to define what some words are and what property of a wave they represent.

Period T

The period is the time taken for a point in the wave to fully complete 1 cycle. For instance, for the object in figure 1, it is the time taken for the object to go from the state in figure 1(a) to figure 1(b). Alternatively, in figure 2 the period is the time $T = T_8 - T_0$. Period is measured in seconds (SI Unit), or any other measurement of time.

Note that after a period, the motion/shape of the wave should be repeating itself. Therefore, one period is not the time it takes for the same point on the wave to return to the equilibrium position. This interpretation is wrong as the wave can reach equilibrium position (more than once!) before it repeats itself.

Frequency f

Frequency is basically a measure of: *how many periods am I going through per second*. The SI unit for frequency is Hertz (Hz). If your period is not given in seconds, make sure to do the suitable conversion before expressing the frequency in Hz. Mathematically, the frequency is defined as the reciprocal of the period.

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

Another useful identity is the relationship between frequency and angular velocity ω.

$$\omega = 2\pi f$$

Wavelength λ

The wavelength refers to the length that one complete waveform takes up in space. Below is an example, where the wavelength of the water waves in figure 1 are indicated.

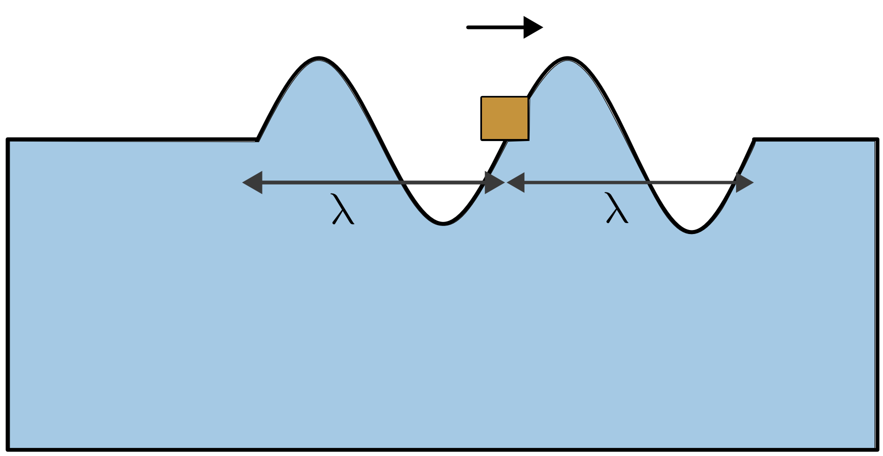


Figure 3: Wavelength indicated, using the situation in figure 1

1.2 Maths

Here I try to propose a more mathematical view of waves instead of just saying that they are "things that repeat itself". In mathematics, waves can be described as periodic functions that depend on space and time, and a periodic function is just a function that repeats itself over different values of x, t. The simplest functions we know of are the sin and cos function. Therefore, we can relate the displacement from equilibrium y of a wave, depending on its distance from the source x as:

$$y(x) = A\sin(kx + \phi)$$
$$k = \frac{2\pi}{\lambda}$$

Due to the way the sine and cosine functions work with radians, the wavenumber k is assigned as such to make sure one complete waveform of the wave takes up λ in space. ϕ is a phase angle, also in radians, which just states the offset of the sin wave. Finally, A is the amplitude of the wave. Figure 4 shows this function.

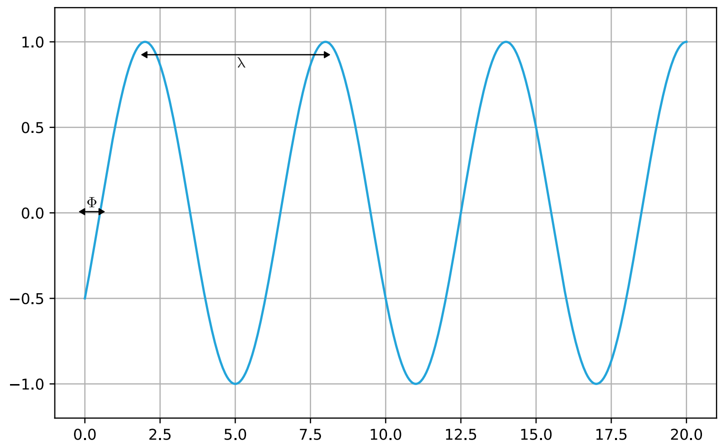


Figure 4: Plot of $y = \sin(\frac{2\pi x}{6} - \frac{\pi}{6})$, $\lambda = 6$, $\phi = \frac{\pi}{6}$

Finding wavelengths
Some questions may ask you to find the wavelength (or period) based on a graph. It is the easiest to find the wavelength if you consider the distance between 2 successive peaks.

Identifying phase
As for the phase, it is most easily visualized if you look at the offset of the wave from its supposed value at $x = 0$. The expected values for sin and cos are 0 and 1 respectively. Here the wave is shifted right by 1/12 of a period since the value at $x = 0$ is $y(x) = -0.5 = \sin(-\frac{\pi}{6})$.

To represent the change with time, we can simply notice that the phase angle ϕ changes with time. Hence, we arrive at the following equation. ω is used to make sure the period matches up with the calculation of the trigonometric functions.

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

Extension: Radians

Normally in compulsory mathematics we express angles in degrees (^{o}). As it turns out, it is not the best unit for advanced mathematics, including our usage here. Instead, we use the radian measure. Radians and degrees are related by the following relationship.

$$180^o = \pi \text{ rad}$$

Most importantly, just remember that 2π refers to one whole period of the trigonometric functions. Of course, if you take M2, this is probably already familiar to you.

1.3 Wave speed

The wave speed describes the speed at which a waveform propagates through space. If we instead freeze our view to a single point, we can rephrase the question to: How much of a wavelength passes through the point per unit time. Hence, we arrive at the defining equation for wave speed c.

$$c = f\lambda$$

In the mathematical description, wave speed is obtained with the following formula:

$$c = \frac{\omega}{k}$$

Reminder: Wave speed

Wave speed will not be changed by f or λ. Instead, those 2 quantities change accordingly to keep c constant. In DSE, wave speed is only dependent on the **medium** the wave travels in.
We can use $c = f\lambda$ to determine frequency or wavelength when we know the wave speed and one of the two quantities. On the contrary, we can also find c when given both f and λ.

Below is a table of how some waves' speed vary.

Wave Type	Wave Speed Dependency
Light	$v = \frac{c_0}{n}$
Sound	$v = \sqrt{\frac{K_s}{\rho}}$
Wave on string	$v = \sqrt{\frac{T}{\mu}}$
Water waves	$v \approx \sqrt{gh}$

n : Optical refractive index of medium
K_s : Bulk Modulus (Out of syllabus)
T : Tension applied to string
μ : Linear mass density of string
h : Depth of water

1.4 So what do waves do?

To examine this question let's think of where waves appear in our daily lives.

Sound Sound waves are the reason why we can hear]

Radio Radios capture the AM/FM waves in the atmosphere to play your favourite Keung To song on 903.

Light Light is a wave, that's why you can see

Heat The sun not only emits light, but also heat with infrared rays, that's why the Earth is not freezing.

All these things have something in common – They do not transport physical matter. Even in figure 1, the object doesn't move after the wave crests have completely passed. So what happened? During the process, notice how our object in figure 1 moves up and down, meaning that it has kinetic energy. **Wave, in fact, transport energy and information, rather than matter.** (Technically information is just the result of transforming energy)

1.5 Different types of waves

Waves are mainly split into 2 types: Transverse and Longitudinal. The difference is best illustrated visually in figure 5. The upper waveforms are the initial waves and the lower ones are at a later time. In a transverse wave, a point will move **perpendicularly** to the wave propagation direction, while a point moves **parallel** to direction of wave propagation in a longitudinal wave.

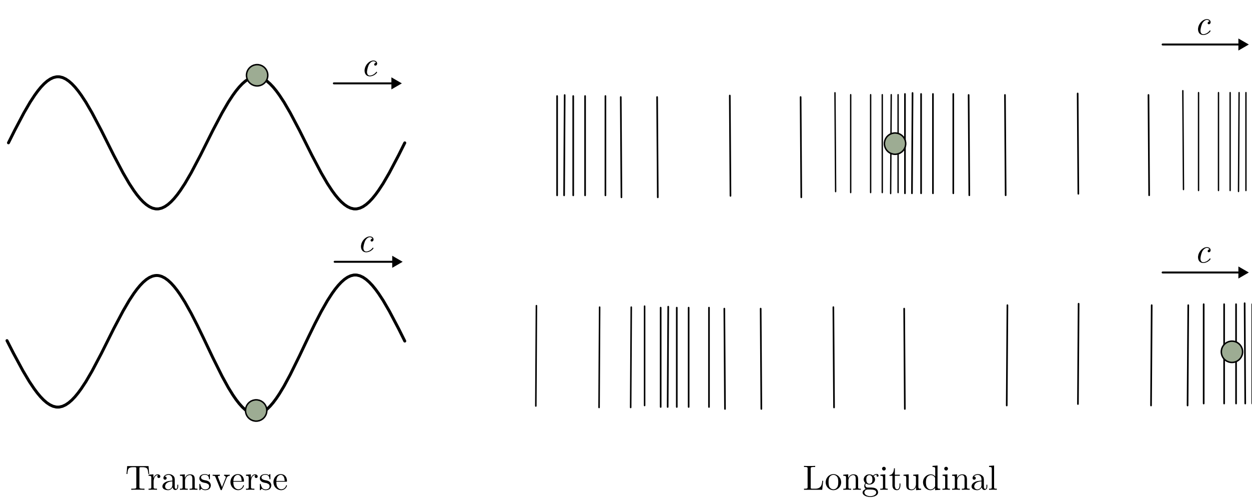


Figure 5: Transverse waves (left) and longitudinal waves (right)

Classification of common waves:

Transverse waves

- Light
- Water waves
- Waves in a string

Longitudinal waves

- Sound
- Waves in a spring

2 Wave motion

2.1 Particle on a transverse wave

In this section we answer the following few questions about a particle at a point on a transverse wave:

- How will it move?
- Where will it be a certain point of time later?

Let's answer the first question first. We can tackle this problem by just inspecting the waveform in figure 6. The wave propagates to the right, we draw a "new" version of the wave a slight bit later, and mark out the displacement from equilibrium of the "new" wave at our old location. Then, compare the old and new locations to find the direction of movement. In our case, the particle is travelling downwards. Note: At a trough or peak, the particle is temporarily at rest.

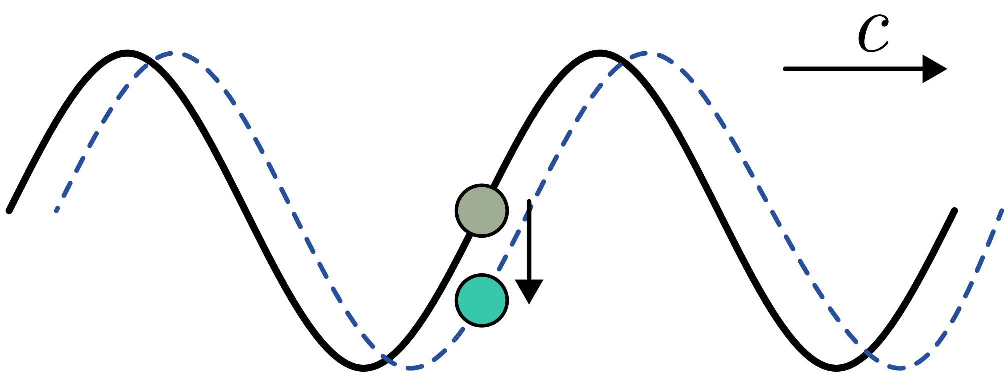


Figure 6: Motion of particle on wave.
Black line: Waveform at $t = 0$s; Blue dashed line: t at some small step of time later.

Extension: Calculus and waves

That method described above is enough to tackle 99% of questions you see in the DSE. However, we do realise that we have only determined the direction of movement, but missing the velocity. This is where math can come in. Once again we model the curve with our trusty $y = A\sin(kx - \omega t + \phi)$. Our first goal is to figure out what the phase angle ϕ is.

We notice that to the right of the particle in figure 6, a normal sine wave appears. Hence, assuming $x = 0$ at the particle, and $t = 0$, we can deduce that $\phi = 0$. From then on, we do a simple differentiation to find the velocity of the particle.

$$\vec{v}(x,t) = \frac{\partial y}{\partial t} = -\omega A\cos(kx - \omega t)$$
$$\vec{v}(x=0,t=0) = -\omega A\cos(k(0) - \omega(0))$$
$$\vec{v}(x=0,t=0) = -\omega A = -2\pi f A$$

Hence, the velocity direction can be determined (negative means going down). On top of that, we see the velocity is proportional to the frequency and amplitude of the wave, on top of the position.

Now for the second question. The easiest way is to see where it is in its current period. In the DSE, questions will only ask you about the displacement of the point at a certain quarter after the initial time. E.g. they can ask you about $\frac{T}{4}$, $\frac{T}{2}$, $\frac{3T}{4}$, etc. but not something like $\frac{5T}{12}$. With this, we just need to remember the sequence. The most important thing is to realise whether you are heading to the peak or the trough first, but you can do that with the method in the previous section.

$$\text{eqm position} \xrightarrow{\frac{T}{4}} \text{Peak/trough} \xrightarrow{\frac{T}{4}} \text{eqm position} \xrightarrow{\frac{T}{4}} \text{Peak/trough} \xrightarrow{\frac{T}{4}} \text{eqm position} \xrightarrow{\text{Start over}}$$

Extension: Mathematical method

With our trusty $y = A\sin(kx - \omega t + \phi)$. There's not so much guessing to be done here.

Figure out ϕ based on your initial condition and then substitute in t to directly arrive at the answer. For most cases, it is the most convenient to take the point of interest to be $x = 0$. Most of the time, this method is slower for finding waveforms at $\frac{nT}{4}, n \in \mathbb{Z}$, but it is always helpful to have another method at your disposal.

Quick question

1. Answer the following questions given the displacement-distance graph (Figure 7). Point A is as indicated in figure 7.
   (a) What is the amplitude of the wave?
   (b) Identify the wavelength of the wave.
   (c) If the wave propagates to the right. Sketch the wave at 3/4 period later.
   (d) To which direction is particle A currently moving towards?

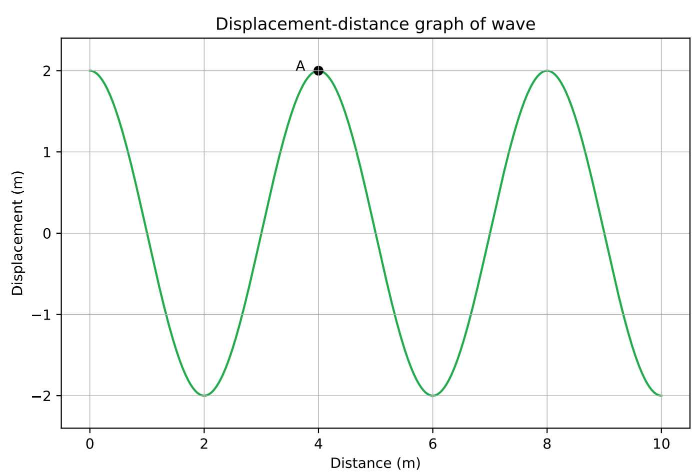


Figure 7: Displacement-distance graph

The displacement-time graph of A is now given as follows. The moment in figure 7 is taken to be $t = 0$s.

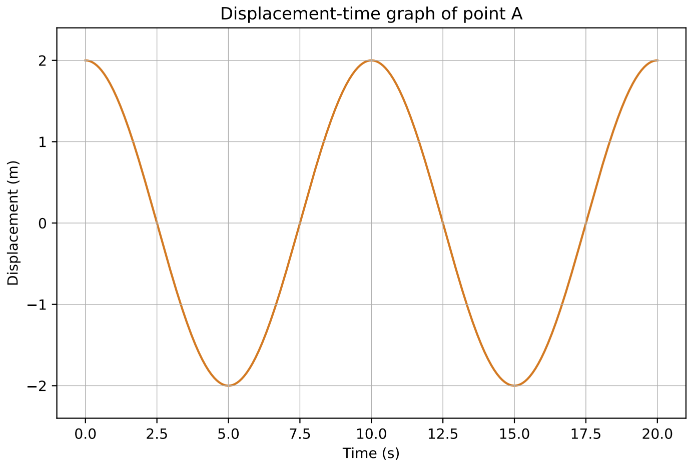


Figure 8: Displacement-time graph of point A

(e) What is the period of the wave?

(f) Suppose the frequency is doubled, sketch the new displacement-time graph on figure 8.

(g) What is the speed of the wave? How will the change in part (f) affect the wave speed?

(h) Starting from $t = 0$s, when will point A reach the equilibrium position for the 5^{th} time?

(i) Are the two graphs sufficient to determine whether the wave propagates to the left or to the right?

2.2 Wave motion on a longitudinal wave

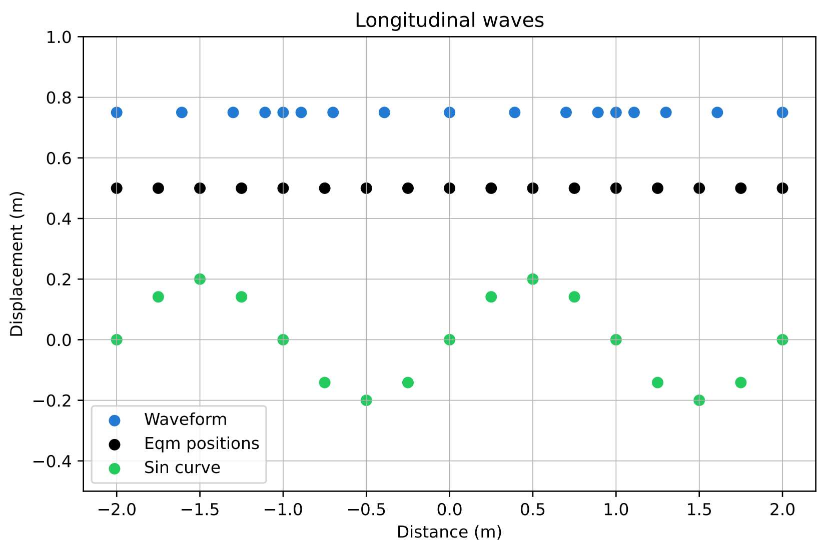


Figure 9: Forming a displacement-distance graph for longitudinal waves

The hardest part of these questions for longitudinal wave is that we don't instantly have a clear indication of the displacement-distance relationship. Hence, the heavy lifting consists of transforming a waveform into a sin graph such as figure 7. The rest is the same as before. You can do the process by finding the difference in position of each particle to its equilibrium position, and then plotting. This process is shown in figure 9.

Here we also observe that longitudinal waves have *centers of compression* and *centers of rarefaction*. Centers of compression refers to locations where the particles are more concentrated. In figure 9, the centers of compression are at $x = -1$ and $x = 1$. Centers of rarefaction are the locations where the concentration of particles are at a minimum. In figure 9, this would be $x = -2, 0, 2$.

Wavelength of a longitudinal wave

A keen eye should notice that the centres of compressions and centers of rarefaction in figure 9 all correspond to points on the equilibrium position. We can use this to our advantage.
The wavelength of a longitudinal wave is the distance between two successive centers of rarefaction or centers of compression.

3 Basic wave phenomena

3.1 Reflection

When a wave collides with a boundary, it changes direction of propagation. This is known as reflection, perhaps best characterized by the law of reflection.

Law of reflection

The angle of reflection is equal to the angle of incidence.

When drawing out the wavefronts, it is best to first sort out the direction of propgation, then add perpendicular lines to the propagation arrow to show the wavefronts. An example is shown in figure 10

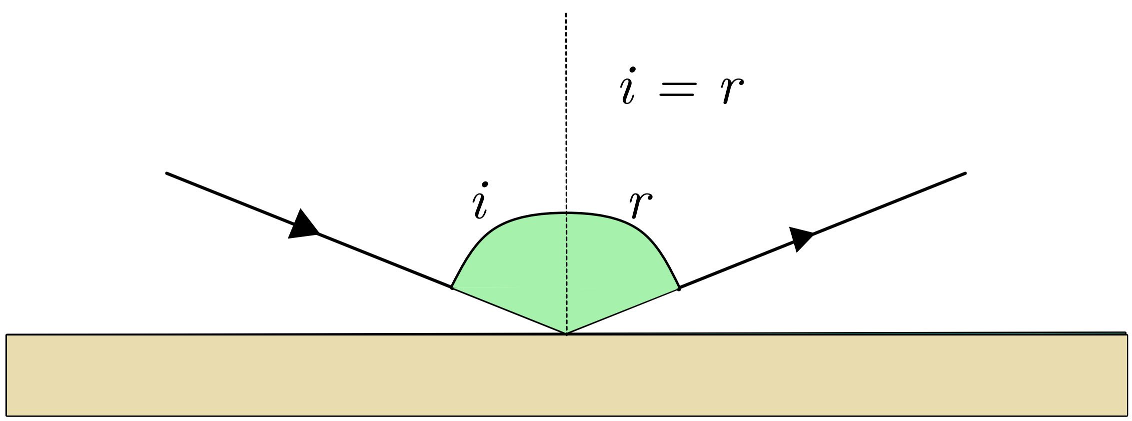


Figure 10: Forming a displacement-distance graph for longitudinal waves

Extension: Matter can reflect as well

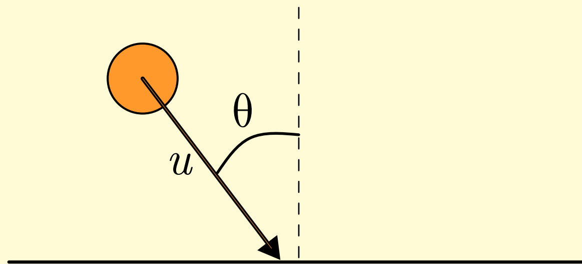


Figure 11: Ball colliding with a wall at an angle

Figure 11 shows a ball, at an incident angle θ and initial velocity u colliding with a wall. As the collision is only in the vertical direction, with conservation of momentum we can write the following. Along the y-direction:

$$\begin{aligned} \vec{p}_{yi} &= \vec{p}_{yf} \\ u\cos\theta &= v_y \end{aligned}$$

Along the x-direction:

$$\begin{aligned} \vec{p}_{xi} &= \vec{p}_{xf} \\ u\sin\theta &= v_x \end{aligned}$$

Now we calculate the angle of reflection θ_r.

$$\begin{aligned} \tan\theta_r &= \frac{v_x}{v_y} \\ &= \tan\theta \\ \Rightarrow \theta_r &= \theta \end{aligned}$$

We hence come to the conclusion that angle of reflection is the same as angle of incidence. This shows that matter has the same reflection behaviour as waves. Hence, being able to reflect is NOT a defining characteristic of a wave.

3.2 Refraction

When waves pass through a boundary between two mediums at some incidence angle, you may notice that the wave direction bends by a small angle. This phenomenon is called refraction.

Cause of refraction

Refraction occurs when waves change their propagation speed, usually due to a change in medium.

Change in propagation direction
The change in direction during refraction is described by Snell's law.

$$\frac{\sin i}{\sin r} = \frac{v_1}{v_2}$$

As we can see, the ratio of incident to refracted angle is equal to the ratio between the speed of the incident wave and the speed of the emerging wave. For light, the refractive index is a conventional way of expressing the speed of light in different media. We can here express Snell's law in the more familiar form.

Snell's law

$$n_k = \frac{c_0}{v_k}$$

$$n_i \sin\theta_i = n_r \sin\theta_r$$

Change in wavelength
From the discussion on wave speed we know wave speed is related to frequency and wavelength by the following relationship.

$$c = f\lambda$$

The frequency of a wave only depends on the source. If you wiggle a string once per second, any point down the string is only going to wiggle once per second, the same frequency. Hence, when c changes, and f is required to remain constant, the only quantity accounting for the change is λ.

Wavelength change when wave speed changes

For waves coming out of the same source at the same frequency, the wavelength is proportional to the wave speed by $\lambda = \frac{c}{f}$.

Summary When waves go into another medium with a different propagation speed the following occurs.

- Propagation direction changes: if wave speed decreases, the refraction angle is smaller to the incident angle
- Wavelength changes: If wave speed decreases, wavelength decreases; successive wave crests become closer to one another
- Continuity of wavefronts: The wave crests are continuous across the boundary

Reminder: Incident and refraction angles

The incident and refraction angles are taken to be the angle between the wave propagation direction and the *normal* of the boundary. For an example of this, see figure 12.

Putting refraction in action

When asked to draw one of these wave diagrams, you can:

1. Identify whether the wave speed increases or decreases to determine if the refracted ray bends towards or away from the normal.
2. Draw the wave direction. (Dark blue lines in figure 12).
3. Draw the crests, which are perpendicular to the wave direction.
4. Make sure your crests are spaced out correctly and connected at the boundary.

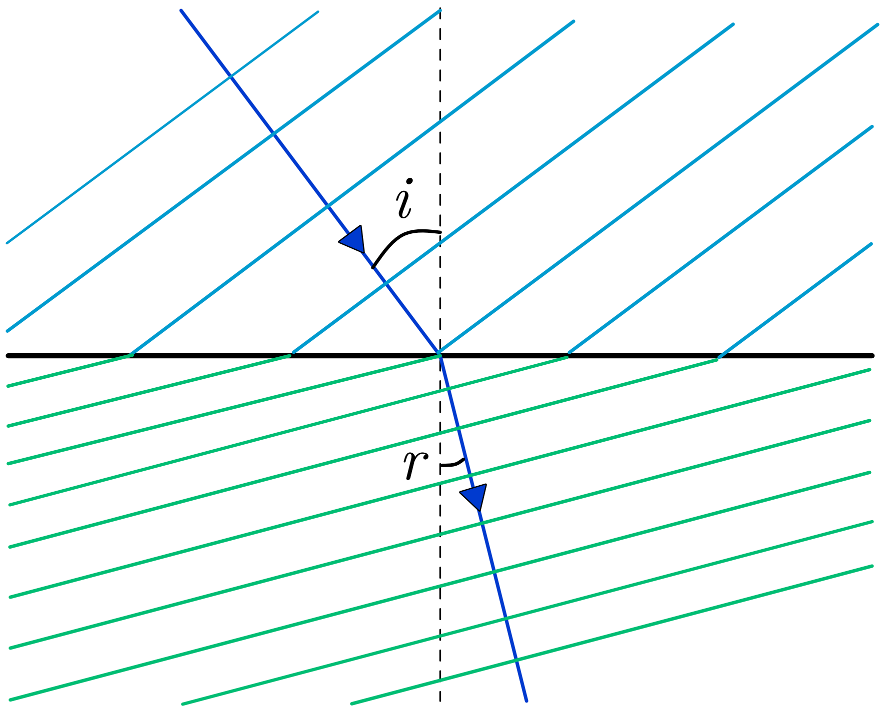


Figure 12: Water waves traveling from a deep region into a shallow region.

3.3 Total internal reflection

When increasing in wave speed, the refraction angle increases. So, what if the refraction angle becomes too big? Let's say our wave is incident at 45 deg to the normal, and the wave speed is doubled in the new medium. We try to find the angle of refraction as follows:

$$\sin r = \frac{v_2}{v_1} \sin i$$

$$\sin r = \sqrt{2}$$

The co-domain of $y = \sin x$ is $y \in [-1, 1]$. Since $\sqrt{2} > 1$, there is no possible solution for r! Therefore, no refraction can occur. Instead the wave reflects on the boundary as if it was a mirror.

We can find the maximum angle of incidence, defined as the critical angle (θ_c) which can produce a refracted ray as follows by setting $\sin\theta_r$ to its max value of 1:

$$\frac{\sin\theta_c}{\sin\theta_r} = \frac{v_1}{v_2}$$

$$\theta_c = \arcsin\left(\frac{v_1}{v_2}\theta_r\right)$$

Reminder: Total internal reflection

Total internal reflection only occurs with **both** of the follow conditions satisfied:

- Incident angle of the wave is larger than critical angle. i.e. $\theta_i > \theta_c$
- The wave is **increasing** in speed when going into the new medium.

4 Interference

Interference is one of the core defining features of waves. Waves can undergo interference, while matter cannot. Interference describes how two waves interact with each other. For normal matter, like a cup, you cannot put two cups in the same position. You can stack them, but you can never "overlap" them in the exact same place.

4.1 Superposition

Principle of superposition

The net response from multiple waves is the sum of the individual disturbances of the waves.

We can express this mathematically:

$$y(x) = f(x) + g(x) + ... = \sum_i^N f_i(x)$$

In other words, if a particle is under the effect of many waves, we can always find its displacement (y) by adding up its displacement for all the individual waves.

Reminder: Superposition

Superposition of waves is a general phenomenon that can occur for any two waves of the same type. (e.g. sound, light, etc.) The two waves need not have any other qualities in common.

Question on superposition

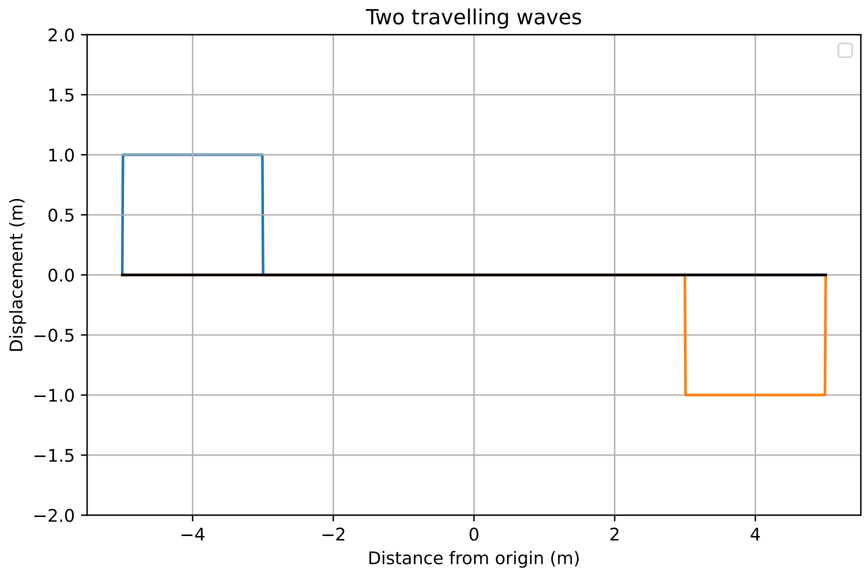


Figure 13: Two traveling waves

In figure 13, there are two square waves on the left and right side of the origin. The left wave is moving to the right at a speed of 1 ms^{-1}, and the right wave is moving to the left at a speed of 1 ms^{-1}. Sketch the total waveform at the following times, supposing that figure 13 is showing the waveform at $t = 0$s.

1. $t = 3$s
2. $t = 3.5$s
3. $t = 4$s
4. $t = 4.5$s
5. $t = 5$s
6. $t = 6$s

Hint: Draw the individual waveforms first, then add them up.

We have investigated superposition for simple-looking waves, but in fact, most waves are not square. Below is an example of superposition of cosine waves.

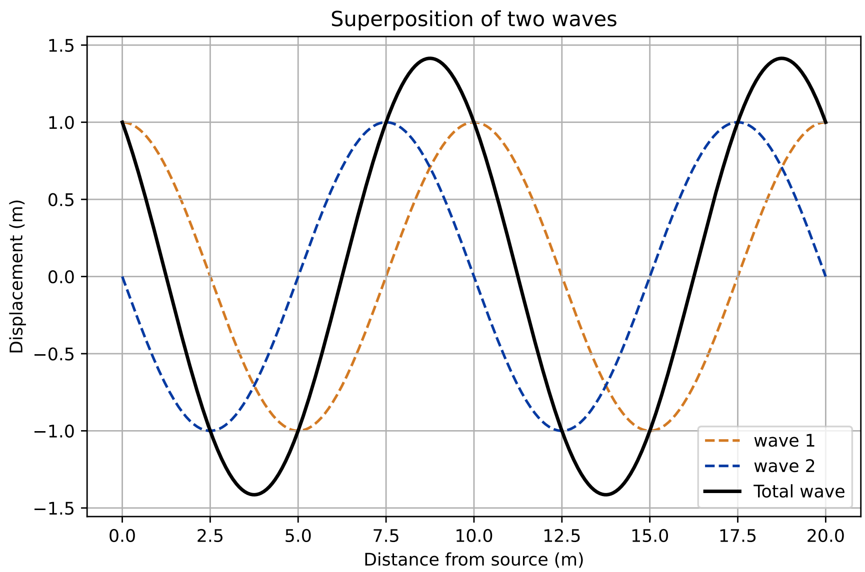


Figure 14: Two traveling waves

4.2 Interference as specific cases of superposition

In the previous section, we saw how waves essentially add to one another in superposition. Here we will see how in two very special cases, waves can reinforce each other or cancel each other. We first consider two waves. For simplicity we will look at the sine wave acting on the same position.

$$y_1 = A \sin \omega t$$
$$y_2 = A \sin(\omega t - \phi)$$
$$y = A(\sin \omega t + \sin(\omega t - \phi))$$

We recall our basic trigonometric identities:

$$\sin(\pi - \theta) = \sin \theta$$
$$\sin(2\pi + \theta) = \sin \theta$$
$$\sin(-\theta) = -\sin \theta$$

Hence we obtain (with $n \in \mathbb{Z}$):

$$\sin(\theta - (2n+1)\pi) = -\sin \theta$$
$$\sin(\theta + 2n\pi) = \sin \theta$$

Going back to y, we see if $\phi = (2n+1)\pi$, then $y = 0$. Conversely, if $\phi = 2n\pi$, $y = 2A \sin \omega t$.

Interference types

We see that if the *phase difference* is $2n\pi$, then the resultant wave doubles. On the other hand, resultant wave cancels out if the *phase difference* is $(2n+1)\pi$. Note that n is an integer. We hence define the two types of interference as follows:

Constructive interference Phase difference of $2n\pi$, the waves reinforce each other, and amplitude is doubled.

Destructive interference Phase difference of $(2n+1)\pi$, the waves cancel each other. The amplitude is zero.

Reminder: Interference

We have assumed phase difference to be a constant. In practice, this means that the frequency (and wavelength) of our waves from the two sources are the same. We call these sources **coherent sources**. **For interference to occur, we need coherent sources.**

Translating to path difference

Now we impose the path factor. With the basic wave equation:

$$\begin{aligned} y_1 &= A\sin\left(kx_1 - \omega t + \phi_1\right) \\ y_2 &= A\sin\left(kx_2 - \omega t + \phi_2\right) \\ \phi &= k(x_2 - x_1) + \phi_2 - \phi_1 \end{aligned}$$

Assuming that the sources are in phase (i.e. $\phi_2 - \phi_1 = 0$), we see:

$$\phi = k(x_2 - x_1) = \frac{2\pi}{\lambda}(\Delta x)$$

Δx is the difference of the distances of the paths to both sources. For constructive interference to occur:

$$\phi = \frac{2\pi}{\lambda}\Delta x = 2n\pi \quad (n \in \mathbb{Z})$$
$$\boldsymbol{\Delta x = n\lambda}$$

Similarly, for destructive interference to occur:

$$\phi = \frac{2\pi}{\lambda}\Delta x = (2n+1)\pi \quad (n \in \mathbb{Z})$$
$$\boldsymbol{\Delta x = \left(n + \frac{1}{2}\right)\lambda}$$

Reminder: Phase difference of sources and interference

If the sources are out of phase – Their waveforms at the individual sources cancel out each other, then the conditions the constructive and destructive interference flip.

Extension: A simpler mathematical treatment

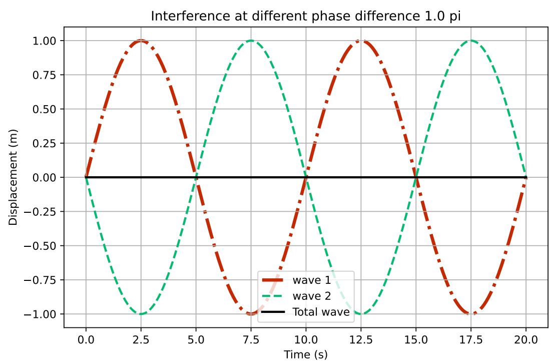
We can use the sum to product formulas to get:

$$y = A\left(\sin\omega t + \sin\left(\omega t - \phi\right)\right) = 2A\sin\left(\omega t - \frac{\phi}{2}\right)\cos\frac{\phi}{2}$$

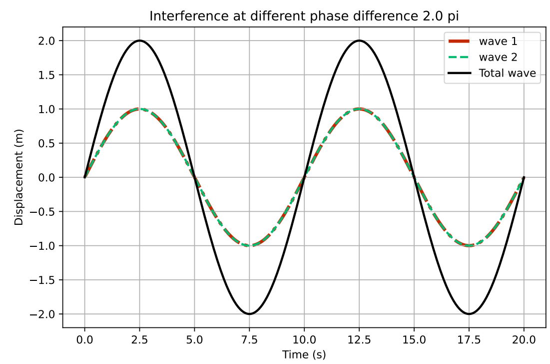
The same results come out when considering the extreme values of $\cos\frac{\phi}{2}$, which is a constant in our equation.

Putting it in pictures

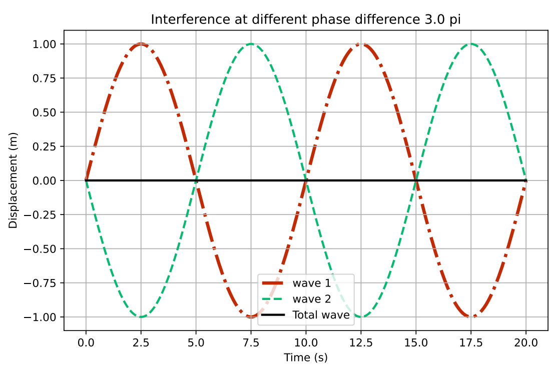
Here we see what coherent waves with different phase differences look like.



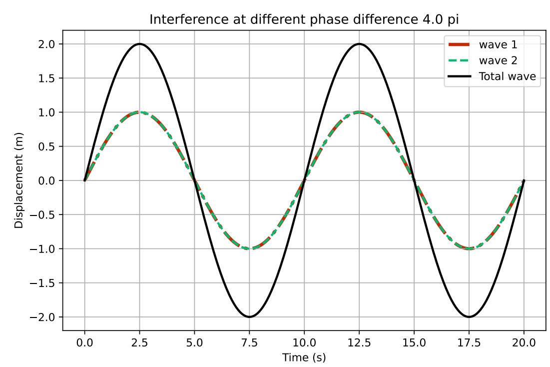
(a) $\phi = \pi$



(b) $\phi = 2\pi$



(c) $\phi = 3\pi$



(d) $\phi = 4\pi$

Reminder: Phase difference of sources and interference

The phase difference is the more fundamental relation to describe interference. The path difference is suitable when the two sources are completely in phase or out of phase.

Fortunately, most of your questions will be with such sources.

4.3 Interference pattern of two sources

Any two sources form an interference pattern as in figure 15. The dashed lines in figure 15 represent troughs, and the solid lines represent peaks. The two dippers are placed at (0, 3.5) and (0, -3.5) respectively.

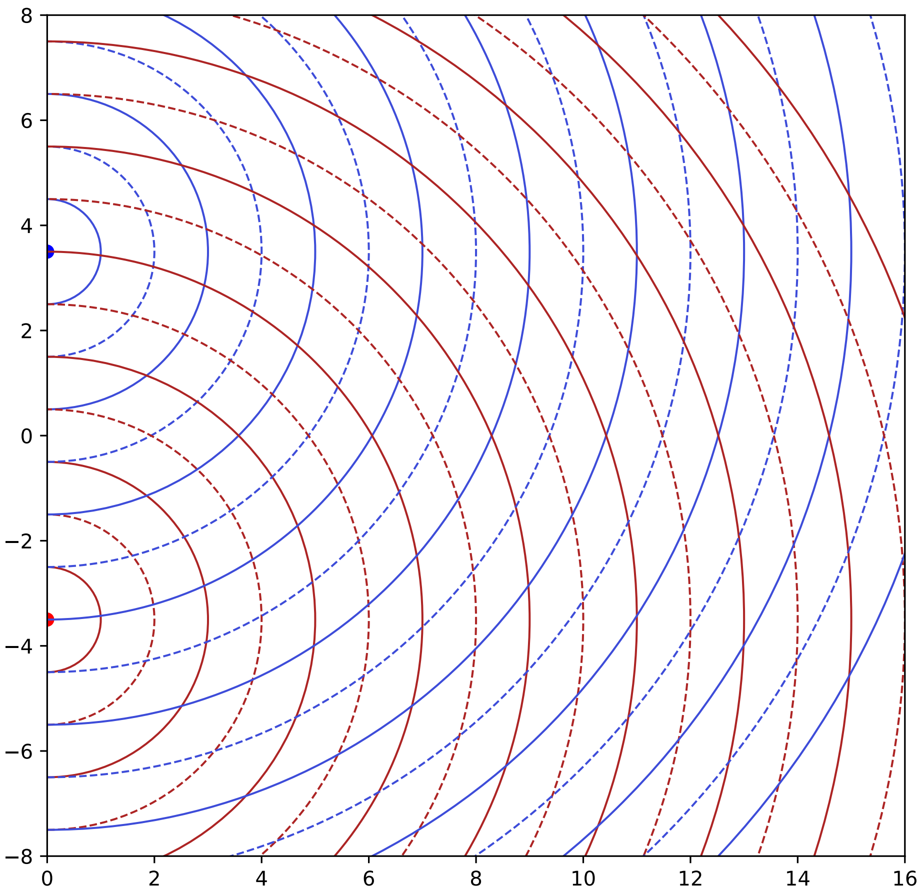


Figure 15: Interference pattern of two dipper sources, placed 7m away from each other.

Identifying different quantities from interference patterns

Wavelength
The wavelength is usually given in the question. Here with the length scales given, one can identify the wavelength by measuring the distance between successive peaks (or troughs). In figure 15, $\lambda = 2m$. If the question further provides the frequency, then the wave speed can be deduced by $c = f\lambda$.

Phase relationship
The phase relationship can be seen from the nature of the first extreme value from each source. In figure 15, the first line from the sources are both peaks. Therefore, both sources are in phase. Generalising, if the first extreme values are the same, then both sources are in phase, and vice versa.

Reminder: Frequency in interference patterns

You cannot identify frequency of a wave based on a displacement-distance relationship alone.

4.4 Lines of minimum and maximum amplitude

As you may have expected, minimum amplitude occurs at locations of destructive interference; Maximum amplitude occurs at locations of constructive interference. On the diagram they can be identified as follows:

Identifying locations of constructive and destructive interference

Constructive interference Locations where "trough meets trough", or "peak meets peak".

Destructive interference Locations where "trough meets peak".

After that, you have to decide the path difference. This can be done by "counting" the circles from each source and subtracting to find the difference. Assuming you keep the path difference constant, you should be able to find a group of points. Connect them up with a line and this is called a *line of constructive/destructive interference.*

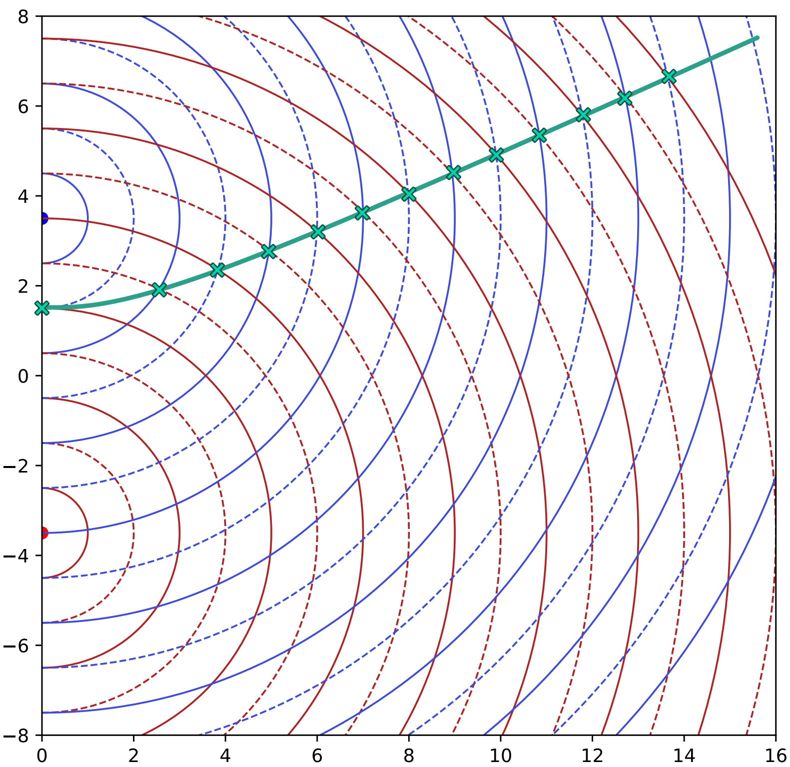


Figure 16: Interference pattern as figure 15, with line of destructive interference $\Delta x = 1.5\lambda$

Extension: Shapes of interference lines

We can use coordinate geometry to find out the shapes of the lines of destructive or constructive interference. We start off with denoting the sources to be vertically spaced from each other, at (0, a) and (0, -a). Then, we calculate the path difference (Δq) at a point (x, y).

$$\begin{aligned}
\Delta q &= \sqrt{x^2+(y-a)^2} - \sqrt{x^2+(y+a)^2} \\
\Delta q + \sqrt{x^2+(y+a)^2} &= \sqrt{x^2+(y-a)^2} \\
\Delta q^2 + x^2 + (y+a)^2 + 2\Delta q\sqrt{x^2+(y+a)^2} &= x^2+(y-a)^2 \quad \text{Squaring both sides} \\
\Delta q^2 + y^2 + 2ay + a^2 + 2\Delta q\sqrt{x^2+(y+a)^2} &= y^2 - 2ay + a^2 \\
\Delta q^2 + 4ay + 2\Delta q\sqrt{x^2+(y+a)^2} &= 0 \\
\Delta q^2 + 4ay &= -2\Delta q\sqrt{x^2+(y+a)^2} \\
\Delta q^4 + 16(ay)^2 + 8ay\Delta q^2 &= 4\Delta q^2(x^2+(y+a)^2) \\
\Delta q^4 + 16(ay)^2 + 8ay\Delta q^2 &= 4\Delta q^2x^2 + 4\Delta q^2y^2 + 8ay\Delta q^2 + 4\Delta q^2a^2 \\
&\text{Rearranging then gives} \\
4\Delta q^2x^2 - (16a^2 - 4\Delta q^2)y^2 &= \Delta q^4 - 4\Delta q^2a^2
\end{aligned}$$

The final equation is in the form of $\frac{x^2}{c_1^2} - \frac{y^2}{c_2^2} = c_3^2$, which is the standard equation for a hyperbola. The shape of a hyperbola is shown in figure 17. The center of the graph looks like a quadratic, while the edges of the graph tend to a linear relationship.

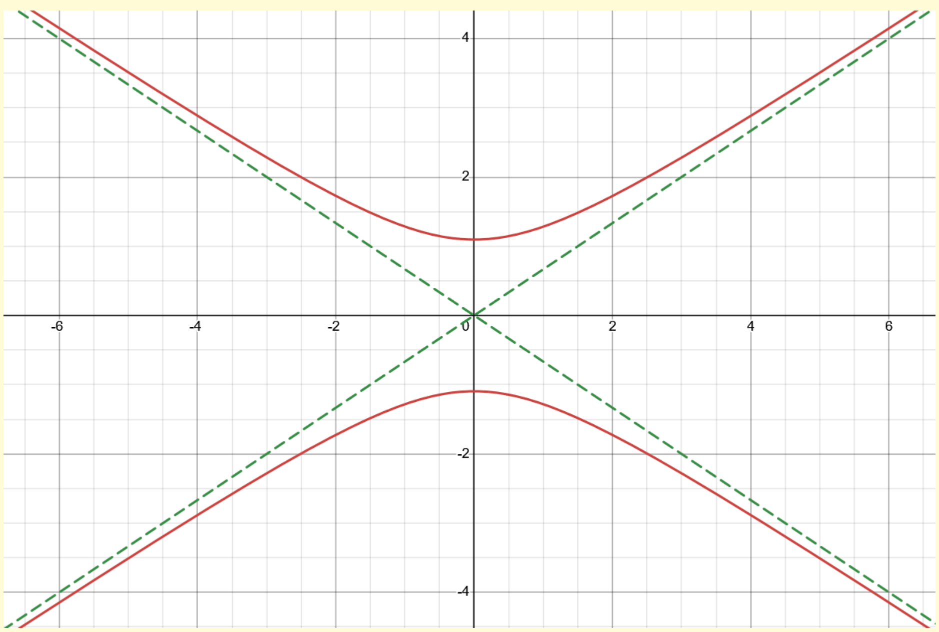


Figure 17: Shape of hyperbola

Finally, if we set $\Delta q \to 0$, then we retrieve $y = 0$, which is just the perpendicular bisector of the two sources, as expected.

4.5 Maximum interference order

We have previously seen in figure 16, that there are numerous lines of constructive and destructive interference from the two-source setup. However, exactly how many of these lines are there? We know that for each interference pattern, we have the following path differences:

$$\Delta x = m\lambda \rightarrow \text{Constructive Interference}$$
$$\Delta x = (m + \frac{1}{2})\lambda \rightarrow \text{Destructive Interference}$$
$$m \in \mathbb{Z}$$

Therefore, to say there is a limit to the amount of interference lines we see is the same as saying that there is a limit to the path difference. To figure the max path difference out, let's look at the sources again.

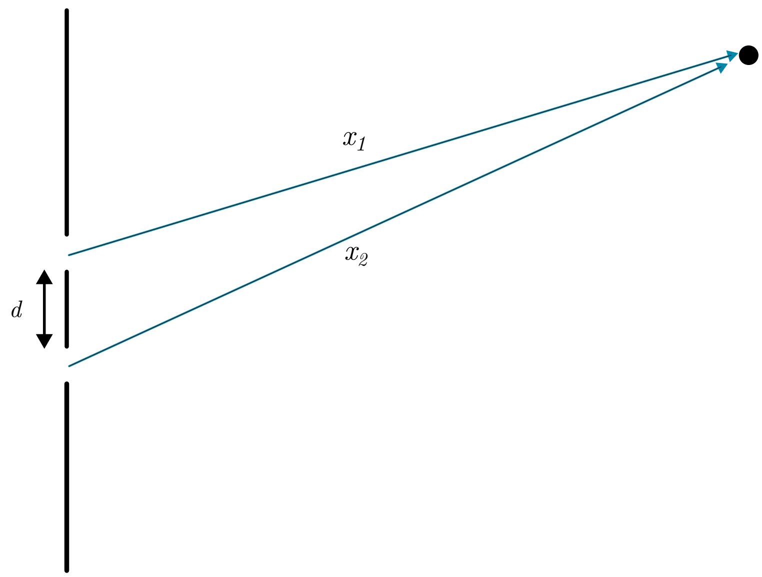


Figure 18: Maximum path difference from two sources

In figure 18, the sources and our point form a triangle, with that we can use the triangle inequality.

$$x_1 + d > x_2$$
$$x_2 - x_1 < d$$
$$\Delta x < d$$

Therefore, the largest interference order is limited by the distance between the two sources. We further substitute m to find the max interference order.

$$m < \frac{d}{\lambda} \text{ (Constructive Interference)}$$
$$m < \frac{d}{\lambda} - \frac{1}{2} \text{ (Destructive Interference)}$$

Reminder: Number of lines of constructive and destructive interference

When you are asked about how many lines there will be, remember to count **both** sides of the pattern. For a given m_{max} you calculated using the method in section 4.5, the maximum lines is given as:

$$l = 2m_{max} + 1 \text{ (Constructive Interference)}$$
$$l = 2m_{max} + 2 \text{ (Destructive Interference)}$$

The difference comes from the number of lines of the two types of interference at the zeroth order $m = 0$. Make sure that you that m_{max} as an integer.

5 Diffraction

The final wave phenomenon we will cover is more like an extension of interference. Diffraction is when a wave travels through a small gap in a barrier, the resultant wave "diverges". An example of this is shown in figure 19.

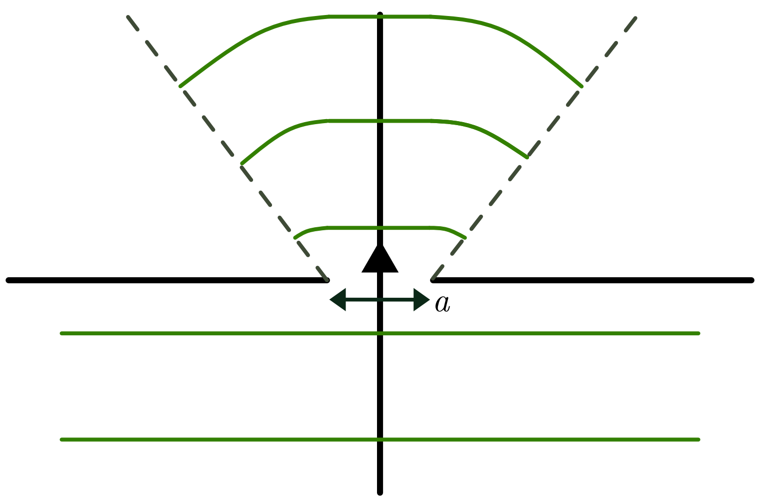


Figure 19: Diffraction through a small slit

Drawing diffraction patterns

- Start by drawing two lines diverging from the mouth of the slit. In figure 19, this is the 2 dotted gray lines.
- Then, draw a straight plane waves for the portion of the wave that directly travels through the slit.
- Connect the edges by drawing circular arcs from the partially compeleted wavefronts of the previous step.

Factors affecting degree of diffraction

There are only two factors affecting how much a wave can "spread" by diffraction when traveling through a small slit.

Wavelength A wave with longer wavelength diffracts more

Slit width A smaller slit will result in larger degree of diffraction

Reminder: Quantities in diffraction

Only the amplitude and wave propagation direction changes during diffraction. Everything else, including frequency, wavelength and wave speed remain constant.

Extension: Intensity of diffraction pattern

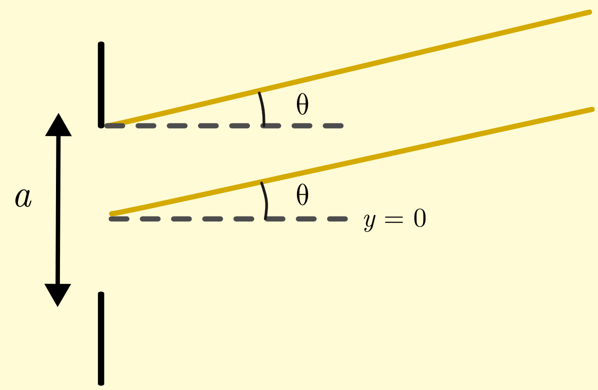


Figure 20: Using integration to solve for diffraction patterns

In this extension section we will find the intensity pattern of diffraction from a single-slit. We first assume the *far field condition*, which states that the light rays coming out of the slit are parallel. This then allows us to find the diffraction intensity pattern by taking the *fourier transform of a top-hat function with width a*. Don't worry if you know nothing about that, just know that the diffraction wave function can be found with the following integral.

$$
\begin{aligned}
\psi(p) &= \int_{-a/2}^{a/2} \frac{1}{a} e^{-ipy} dy \\
&= \frac{i}{ap} e^{-ipy} \Big|_{-a/2}^{a/2} \\
&= \frac{2}{ap} \frac{e^{ipa/2} - e^{-ipa/2}}{2i} \\
&= \frac{2}{ap} \sin \frac{pa}{2} \\
\psi(p) &= \operatorname{sinc}\left(\frac{pa}{2}\right)
\end{aligned}
$$

Extension: Intensity of diffraction pattern

In the integral we have used quite a few tricks in math, I will list them here. (Note: Calculating this integral is unimportant, obviously. We are only interested in the result)

- $e^{ix} = \cos x + i \sin x$ – Euler's identity, this makes it so we can express our wavefunctions with that.
- $\frac{e^{i\theta} - e^{-i\theta}}{2i} = \sin\theta$, the complex definition of sin function. You can actually derive this from the above identity.
- $\frac{\sin x}{x} = \mathrm{sinc}(x)$ – Just a mathematical shorthand.

We will define p as the spatial frequency: $p = k \sin\theta = 2\pi/\lambda \sin\theta$. We finally arrive at the intensity of the wave:

$$\begin{aligned} I &= \psi^2 \\ &= \mathrm{sinc}^2\left(\frac{pa}{2}\right) \\ &= \mathrm{sinc}^2\left(\frac{\pi a \sin\theta}{\lambda}\right) \end{aligned}$$

When plotted, the functions looks as follows:

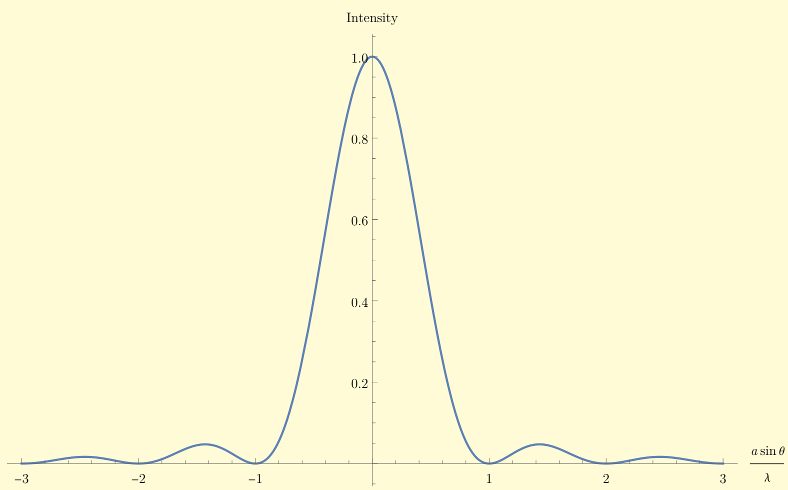


Figure 21: $\mathrm{sinc}^2\left(\frac{\pi a \sin\theta}{\lambda}\right)$

Crucially, the function is zero when $a \sin\theta = n\lambda, n \in \mathbb{Z}$. Hence the first order diffraction angular width is given as $\sin\theta = \lambda/a$, proving the relations we discussed in this section.

Diffraction pattern width

The width of diffraction pattern is given by $a \sin\theta = \lambda$

6 Light as a wave

This section examines the two experiments that demonstrates the wave nature of light. In particular, we show that light can undergo both interference and diffraction – The characteristic phenomenon of waves.

6.1 Double slit interference

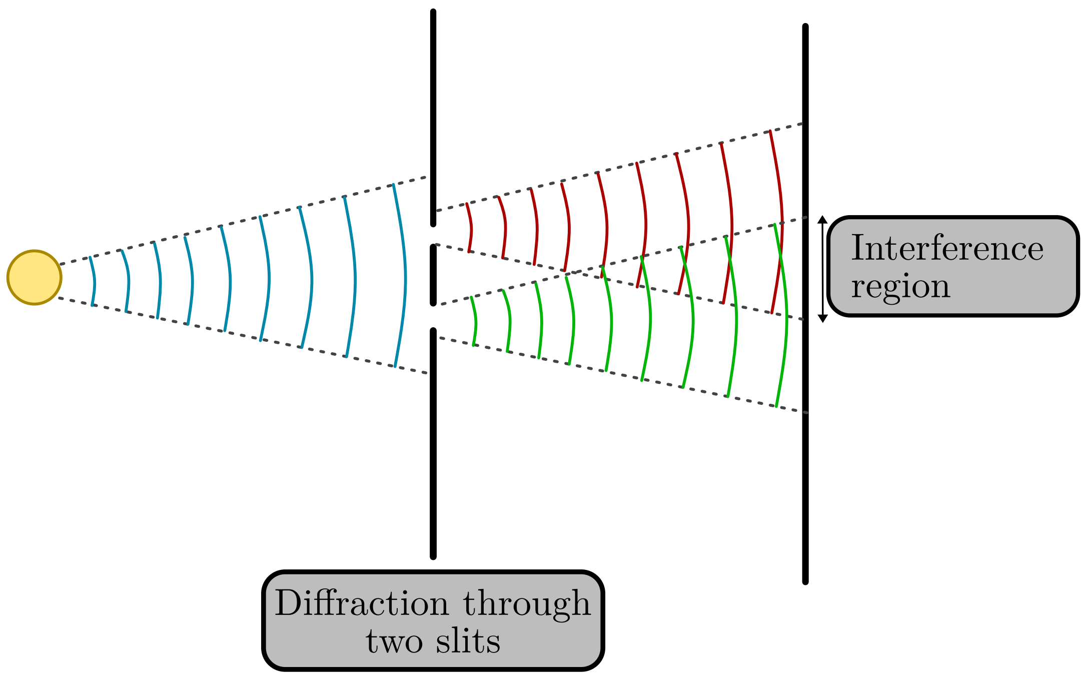


Figure 22: Young's double slit setup

In the experiment, light comes from a monochromatic source. A monochromatic source refers to a light source producing light of only one frequency (and hence one wavelength). Most light sources, like LEDs and Light bulbs actually emit light of several wavelengths.

Light reaches the two slits, and then diffracts as shown in figure 22. The diffracted wavefronts interfere near the centre of the screen, as marked by the interference region. If light is a wave, we should be able to find spots of constructive and destructive interference, leading to bright and dark fringes.

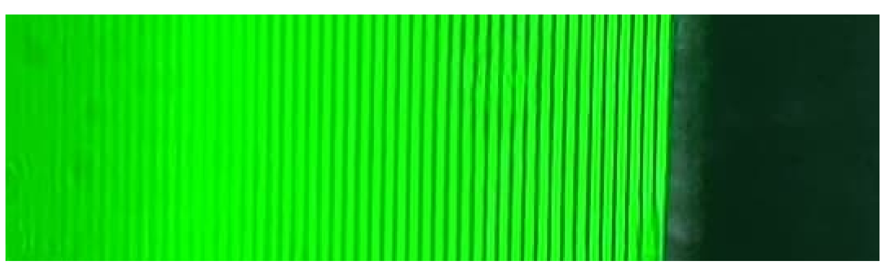


Figure 23: Double slit interference pattern, produced on a Lloyd's mirror setup.

6.2 Deeper look into Young's double slit setup

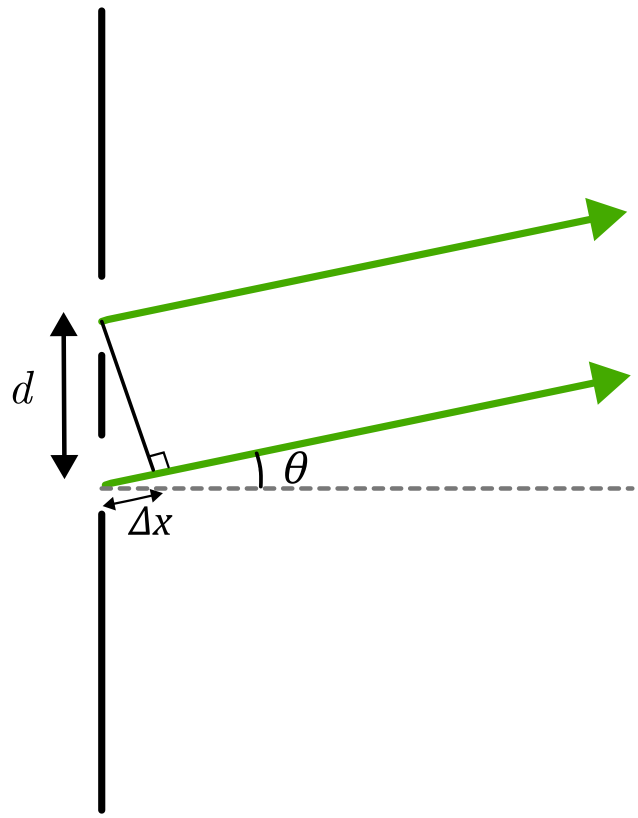


Figure 24: A closer look on the double slit setup

We first make the following assumptions:

Far field conditions The screen distance is much larger than the slit separation, i.e. $D >> d$. This allows us the portray the light waves from the two slits to a certain point on the far-away screen as parallel.

Slit width The slits are considered to be very narrow, so we can treat them as point sources. Later on, we will quickly brief over what happens if the slits had non-zero width.

Location on screen The angle θ is small, so that means we are only looking at the center of the screen. Normally, this is achieved by having a large screen distance D. This allows us to use the small angle approximation later.

By considering some simple geometry, we can find the path difference Δx. Δx is only that extra bit labeled in figure 24, since the two waves reach the final point of the screen with the same path length after that point.

$$\Delta x = d \sin\theta$$

We then consider the conditions for constructive interference.

$$\Delta x = d\sin\theta = m\lambda \quad \text{You will see this again...}$$

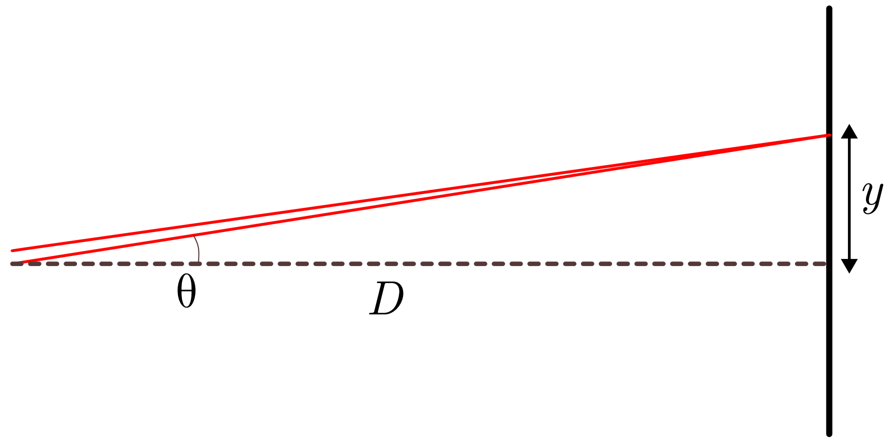


Figure 25: Screen setup of Young's double slit experiment

We use the small angle approximation.

$$\sin\theta \approx \tan\theta = \frac{y}{D}$$
$$\Rightarrow d\frac{y}{D} = m\lambda$$
$$y = m\frac{D\lambda}{d}$$

The distance between successive bright fringes is:

$$\Delta y = (m+1)\frac{D\lambda}{d} - m\frac{D\lambda}{d} = \frac{D\lambda}{d}$$

So we find the following relationship between the fringe separation Δy and the three quantities:

- Screen distance D – A screen placed further away yield larger fringe separation Δy
- Wavelength λ – Waves with longer wavelengths give larger fringe separation Δy
- Slit separation d – Small slit separation gives larger fringe separation Δy.

Extension: A mathematical method on the double slit setup

We treat the two sources as point sources emitting waves that hit the screen at a point y away from the center, just like in figure 25.

$$\psi_1 = A\cos(kx_1 - \omega t)$$
$$\psi_2 = A\cos(kx_2 - \omega t)$$

Extension: A mathematical method on the double slit setup

Our distances can be defined as follows

$$x_1 = \sqrt{D^2 + \left(y - \frac{d}{2}\right)^2}$$
$$x_2 = \sqrt{D^2 + \left(y + \frac{d}{2}\right)^2}$$

We approximate $x_2 - x_1$ as follows, with the assumption that $d \approxeq y << D$, we can use the approximation $(1+x)^n \approx 1 + nx$ for $x << 1$.

$$\begin{aligned} x_1 &= D\left(1 + (\frac{y}{D} - \frac{d}{2D})^2\right)^{1/2} \\ &\approx D\left(1 + \frac{1}{2}\left(\frac{y}{D} - \frac{d}{2D}\right)^2\right) \\ \text{and } x_2 &= D\left(1 + (\frac{y}{D} + \frac{d}{2D})^2\right)^{1/2} \\ &\approx D\left(1 + \frac{1}{2}\left(\frac{y}{D} + \frac{d}{2D}\right)^2\right) \end{aligned}$$

Giving us the path and hence phase difference of:

$$x_2 - x_1 = \frac{D}{2}\left(\frac{2yd}{D^2}\right)$$
$$\Delta x = \frac{yd}{D}$$
$$\phi = \frac{kyd}{D}$$

We add the two sources up by the superposition principle with the sum to product formula.

$$\psi = \psi_1 + \psi_2 = A\cos(kx_1 - \omega t) + A\cos(kx_1 - \omega t + \phi)$$
$$\psi = 2\cos(kx_1 - \omega t + \frac{\phi}{2})\cos(\frac{\phi}{2})$$

We see the amplitude is given by $2\cos(\phi/2)$, using $I \propto A^2$,

$$I(y) \propto \cos^2\frac{\phi}{2} = \cos^2\frac{\pi y d}{\lambda}$$

$I(y)$ peaks whenever $\pi yd/D\lambda = m\pi$, so we once again retrieve:

$$y = \frac{mD\lambda}{d}$$

Plots of double slit intensities

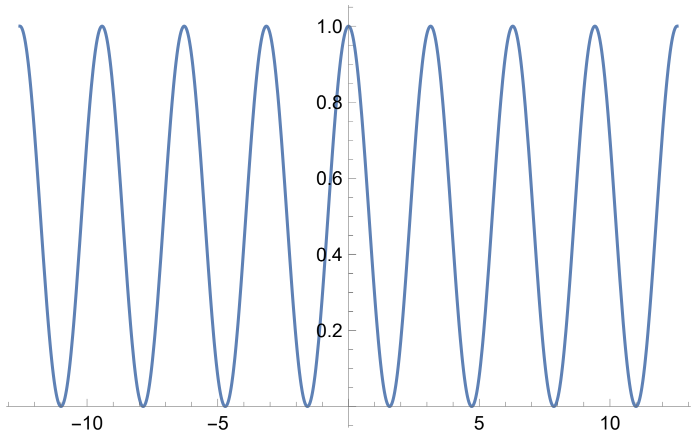


Figure 26: Intensity pattern of double slit interference

Extension: Double slit plots with the effects of diffraction

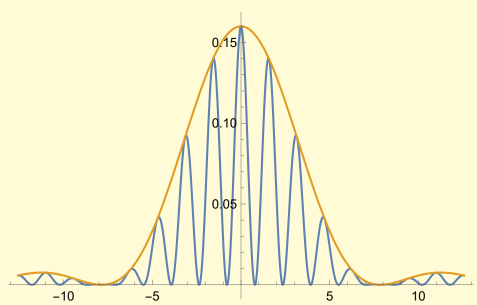


Figure 27: Double slit interference pattern with finite slit size; $\cos^2$ function with a sinc^2 envelope.

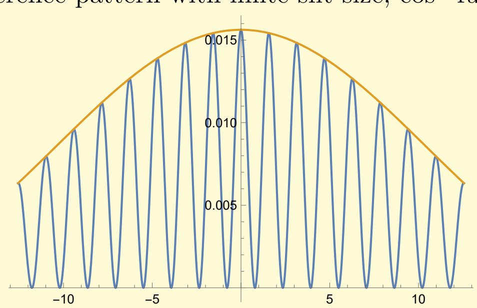


Figure 28: Double slit interference pattern with smaller finite slit size; Brightness is more evenly distributed

Reminder: Slit width in double slit setups

The slit width has no effect on the fringe separation Δy. However, a smaller slit width will lead to a more even distribution of intensity across the screen as the degree of diffraction when light passes through the two slits increases. With a bigger slit width, the central fringes are much brighter, but the fringes' brightness quickly decrease as you move away from the center.

6.3 The diffraction grating

The diffraction grating is essentially the same as the double slit setup, except this this, it is many, many slits interfering with each other.

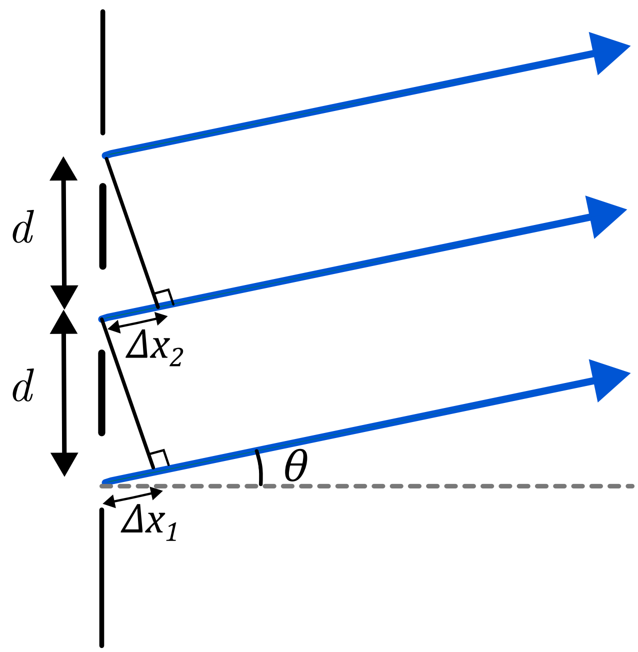


Figure 29: A close up look on the diffraction grating

For constructive interference to occur, all path differences **between any two slits** must be an integer multiple of the wavelength. The only way for this to happen is for the path difference of successive slits, indicated by Δx_1 and Δx_2 in figure 29 $= m\lambda$. We use Δx_i to represent the path difference from two adjacent slits.

$$\Delta x_i = m\lambda$$

$$d \sin\theta = m\lambda$$

Once again we have assumed a few items:

Screen distance $D >> d$, i.e. the far-field conditions, so the emerging light rays can be taken to be parallel.

Wave incidence The incident wave is completely horizontal, coherent and in phase.

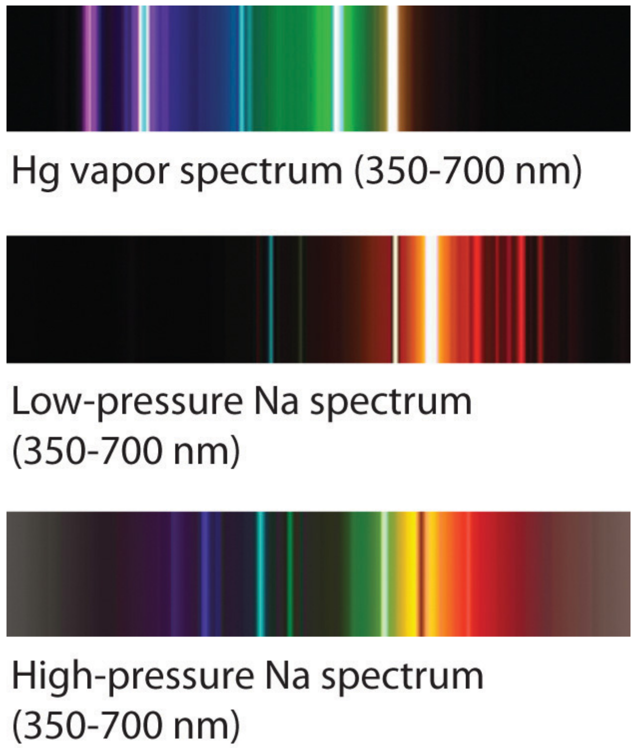


Figure 30: Light from difference sources passed through a diffraction grating

As you can see, while the double slit produces evenly (and closely) spaced fringes, a diffraction grating image creates narrow and bright lines.

6.4 Maximum diffraction order

Typically your grating is labeled in lines per unit length, (e.g. 500 lines/mm), you can thus find the slit to slit distance d as follows:

$$d = \frac{l}{n}$$

Where n is the number of lines in said length, and l is the length. Typically, $d \sim 1\mu$m. This gives rise to a maximum diffraction order observed. In $d\sin\theta = m\lambda$, m can only take up certain values, or else if m is too large we run into the issue of $\sin\theta > 1$, which has no solutions.

$$\sin\theta \leq 1$$

$$m\frac{\lambda}{d} \leq 1$$

$$m \leq \frac{d}{\lambda}$$

Taking the largest integral value of m gives you the maximum order of diffraction which can be observed.

Reminder: Slit width in double slit setups

The number of "bright" lines corresponding to maximas is given by $2m_{max} + 1$, accounting for both sides of the diffraction pattern as well as the central maximum.

6.5 Diffraction of white light

White light is basically light of every wavelength coming together. Hence, if we pass white light through a grating, we can observe a spectrum of colours (somewhat like a rainbow). This is because different colours of light have different wavelengths, which of course gives rise to different diffraction angles by $d \sin\theta = m\lambda$.

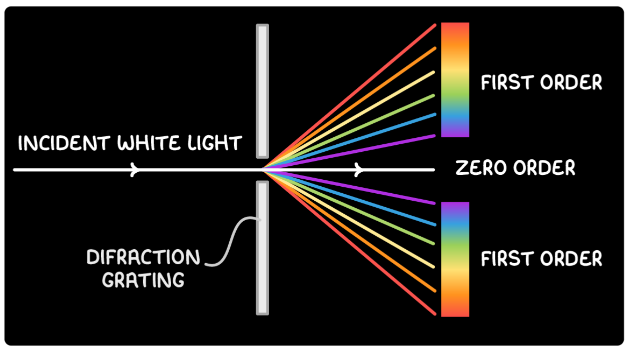


Figure 31: White light going through a diffraction grating

Note that the central maximum is still white as all colours have the central spot as their maxima. Thus, all colours superimpose on each other to create white light. Beyond that, the light with lower wavelength (purple) is diffracted less, and hence closer to the center.

The final item to note is the overlapping of the second and third order spectrums.

$$\sin\theta_{r2} = \frac{2\lambda}{d} = \frac{14 \cdot 10^-6}{d}$$

$$\sin\theta_{p3} = \frac{3\lambda}{d} = \frac{12 \cdot 10^-6}{d}$$

Substituting in the correct wavelengths (700nm for red, 400nm for purple), we find

$$\sin\theta_{r2} > \sin\theta_{p3}$$

, which means the red light in the second order is *further* away from the center than the third order purple light. In other words, the two spectrum will overlap.