

III. Light is a Wave (Physical Optics)

III.B. 2-Wave Interference

1. Superposition of Harmonic Waves

When two or more waves are present, the total wave amplitude is simply the sum of all of the individual wave amplitudes, or

$$\psi(x,t) = \psi_1(x,t) + \psi_2(x,t) + \psi_3(x,t) + \dots .$$

To understand the simplicity of this superposition principle, we look at some examples below. For all of these examples we consider the case of fixed time (say $t = 0$).

a. *Superposition of waves with the same wavelength:*

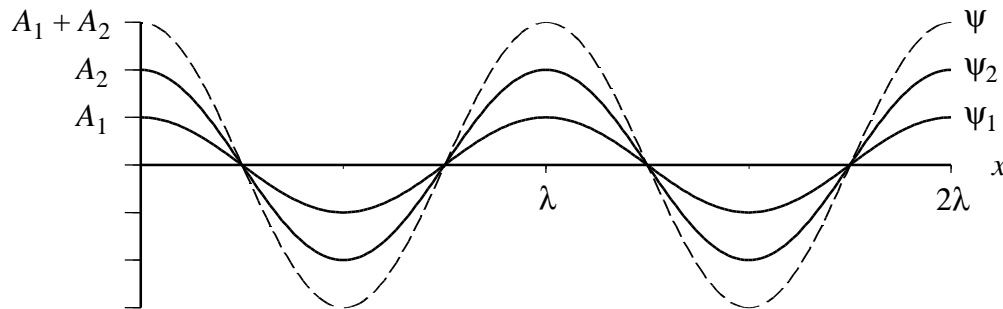
The simplest example of superimposing 2 waves that have exactly the same wavelength is when the 2 waves also have the same constant relative phase — this is called the in-phase case. Suppose

$$\psi_1(x) = A_1 \cos(kx) \quad \text{and} \quad \psi_2(x) = A_2 \cos(kx).$$

Then

$$\psi(x) = \psi_1(x) + \psi_2(x) = (A_1 + A_2) \cos(kx).$$

Graphically, the 2 waves and their sum might look like this:



More generally, the 2 waves might have a different constant phase at a particular position and time. This is called the out-of-phase case. Suppose

$$\psi_1(x) = A_1 \cos(kx + \phi_1) \quad \text{and} \quad \psi_2(x) = A_2 \cos(kx + \phi_2),$$

where ϕ_1 and ϕ_2 are 2 different constant phase values. Then

$$\psi(x) = \psi_1(x) + \psi_2(x) = A_1 \cos(kx + \phi_1) + A_2 \cos(kx + \phi_2)$$

We can simplify this expression for the total amplitude by using the trigonometric identity

$$\cos(\theta_1 + \theta_2) = \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2,$$

so that

$$\psi(x) = A_1 \cos(kx) \cos\phi_1 - A_1 \sin(kx) \sin\phi_1 + A_2 \cos(kx) \cos\phi_2 - A_2 \sin(kx) \sin\phi_2,$$

or

$$\psi(x) = (A_1 \cos \phi_1 + A_2 \cos \phi_2) \cos(kx) - (A_1 \sin \phi_1 + A_2 \sin \phi_2) \sin(kx).$$

If we define a new total amplitude A and a new total phase ϕ according to the following relations

$$A \cos \phi = A_1 \cos \phi_1 + A_2 \cos \phi_2 \quad \text{and} \quad A \sin \phi = A_1 \sin \phi_1 + A_2 \sin \phi_2 ,$$

then we see

$$\psi(x) = A(\cos \phi \cos(kx) - \sin \phi \sin(kx)),$$

or, using the trigonometric identity above,

$$\boxed{\psi(x) = A \cos(kx + \phi)} .$$

This appears to be a very simple result, only we have not yet determined what A and ϕ are. If we square and add the two relations above that were used to define A and ϕ , we see that

$$\begin{aligned} A^2 \cos^2 \phi + A^2 \sin^2 \phi &= A^2 (\cos^2 \phi + \sin^2 \phi = 1) = A^2 \\ &= (A_1 \cos \phi_1 + A_2 \cos \phi_2)^2 + (A_1 \sin \phi_1 + A_2 \sin \phi_2)^2 \\ &= A_1^2 \cos^2 \phi_1 + 2A_1 A_2 \cos \phi_1 \cos \phi_2 + A_2^2 \cos^2 \phi_2 + A_1^2 \sin^2 \phi_1 + 2A_1 A_2 \sin \phi_1 \sin \phi_2 + A_2^2 \sin^2 \phi_2 . \\ &= A_1^2 (\cos^2 \phi_1 + \sin^2 \phi_1) + 2A_1 A_2 (\cos \phi_1 \cos \phi_2 + \sin \phi_1 \sin \phi_2) + A_2^2 (\cos^2 \phi_2 + \sin^2 \phi_2) \\ &= A_1^2 + 2A_1 A_2 \cos(\phi_1 - \phi_2) + A_2^2 \end{aligned}$$

Thus we find that the total amplitude A is

$$\boxed{A = \sqrt{A_1^2 + A_2^2 + 2A_1 A_2 \cos(\phi_1 - \phi_2)}} .$$

Taking the ratio of the two relations we used to define A and ϕ , we see that

$$\frac{A \sin \phi}{A \cos \phi} = \tan \phi = \frac{A_1 \sin \phi_1 + A_2 \sin \phi_2}{A_1 \cos \phi_1 + A_2 \cos \phi_2} .$$

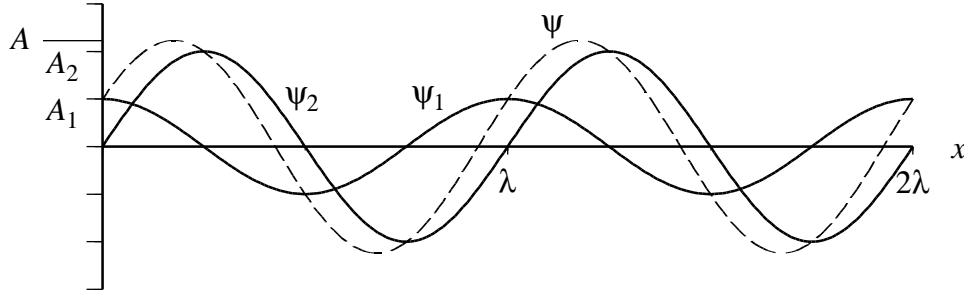
Thus

$$\boxed{\phi = \arctan \left(\frac{A_1 \sin \phi_1 + A_2 \sin \phi_2}{A_1 \cos \phi_1 + A_2 \cos \phi_2} \right)} .$$

The above three boxed results are important. They tell us that

If we add any 2 waves that have the same wavelength λ (and thus the same wavenumber k), the resulting wave has exactly the same wavelength, but a new amplitude and a new constant phase.

An example of this result is illustrated in the plot below.



b. *Superposition of waves with different wavelengths:*

This case is just a little more complicated than the previous one. Suppose we have 2 waves with the same amplitude and the same constant phase (at $x = t = 0$), but different wavelengths λ_1 and λ_2 (and thus different wavenumbers k_1 and k_2 , respectively):

$$\psi_1(x) = A \cos(k_1 x) \quad \text{and} \quad \psi_2(x) = A \cos(k_2 x).$$

Then

$$\psi(x) = \psi_1(x) + \psi_2(x) = A(\cos(k_1 x) + \cos(k_2 x)).$$

If we use the trigonometric identity

$$\cos \theta_1 + \cos \theta_2 = 2 \cos\left(\frac{\theta_1 - \theta_2}{2}\right) \cos\left(\frac{\theta_1 + \theta_2}{2}\right),$$

then we find

$$\boxed{\psi(x) = 2A \cos\left(\frac{k_1 - k_2}{2} x\right) \cos\left(\frac{k_1 + k_2}{2} x\right)},$$

which shows us that *the superposition of waves with different wavelengths is no longer a simple harmonic wave!*

We can interpret this result in a simple way if the 2 waves have approximately equal wavelengths, such that

$$\lambda_1 \cong \lambda_2 \quad \Rightarrow \quad \Delta\lambda = \lambda_2 - \lambda_1 \ll \lambda_1, \lambda_2.$$

In this case

$$\frac{k_1 + k_2}{2} = \frac{2\pi}{2} \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} \right) = \frac{2\pi}{2} \left(\frac{\lambda_2 + \lambda_1}{\lambda_1 \lambda_2} \right) \cong \frac{2\pi}{2} \frac{2\lambda_1}{\lambda_1^2} = \frac{2\pi}{\lambda_1}$$

and

$$\frac{k_1 - k_2}{2} = \frac{2\pi}{2} \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} \right) = \frac{2\pi}{2} \left(\frac{\lambda_2 - \lambda_1}{\lambda_1 \lambda_2} \right) \cong \frac{2\pi}{2} \frac{\Delta\lambda}{\lambda_1^2} = \frac{2\pi}{\lambda_1} \frac{\Delta\lambda}{2\lambda_1} = \frac{2\pi}{\Lambda}$$

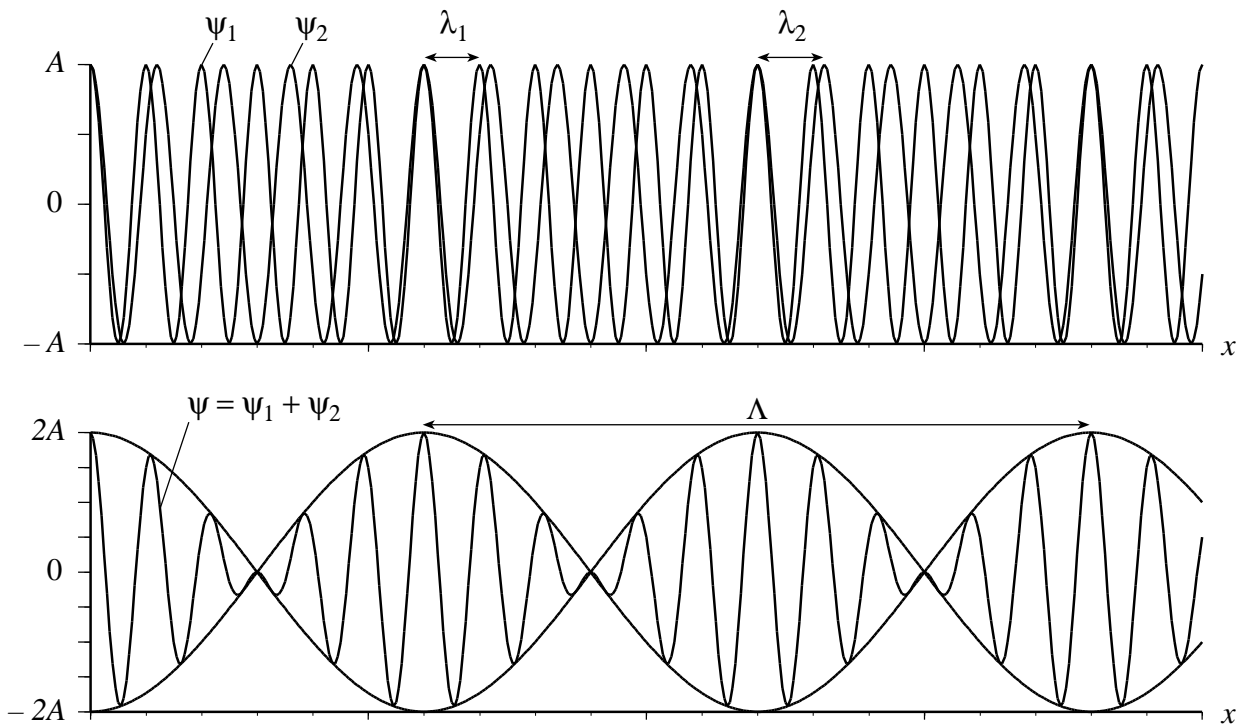
where we define the beat period Λ to be

$$\Lambda \equiv \left(\frac{2\lambda_1}{\Delta\lambda} \right) \cdot \lambda_1.$$

Note that the beat period is much larger than either of the wavelengths, or $\Lambda \gg \lambda_1, \lambda_2$. With these approximations that result from assuming similar wavelengths, we can thus write

$$\psi(x) \equiv 2A \cos\left(\frac{2\pi}{\Lambda} x\right) \cos\left(\frac{2\pi}{\lambda_1} x\right).$$

In this expression the first cosine factor can be thought of as a *slowly varying wave amplitude*, while the second cosine factor is like a *rapidly oscillating harmonic wave* that is similar to each of the 2 waves we started with. The plots below illustrate 2 harmonic waves with approximately equal wavelengths and their resulting superposition.



The lines in the lower plot that form a *slowly varying envelope* around the actual waveform ψ do not actually exist, and are merely drawn to illustrate the effect of the first cosine factor.

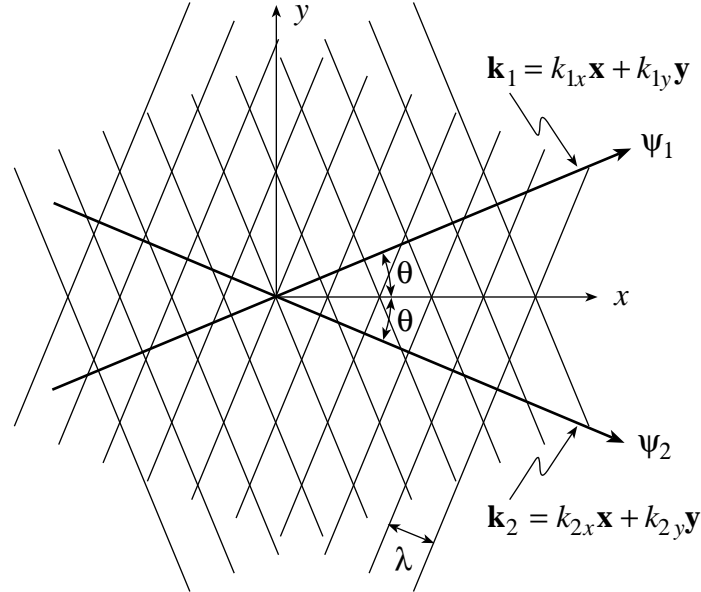
2. Interference of 2 Plane Waves of Light

In this section we will see how interference fringes can form when 2 plane waves of light pass through one another in space. Suppose the 2 plane waves are described by the amplitudes below

$$\psi_1(x, y, t) = A_1 \cos(k_{1x}x + k_{1y}y - \omega t) = A_1 \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega t)$$

$$\psi_2(x, y, t) = A_2 \cos(k_{2x}x + k_{2y}y - \omega t) = A_2 \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega t).$$

A view of these waves in the x - y plane might look like



For the 2 plane waves shown in the drawing, the wavenumber components can be determined with the help of the triangle diagrams below on the right, and they are given by:

$$\begin{aligned} k_{1x} &= k_1 \cos \theta & k_{1y} &= k_1 \sin \theta \\ k_{2x} &= k_2 \cos \theta & k_{2y} &= -k_2 \sin \theta \end{aligned}$$

If the 2 waves have the same wavelength, then

$$k_1 = k_2 = k = \frac{2\pi}{\lambda}$$

The total amplitude ψ is thus given by the superposition

$$\psi(x, y, t) = \psi_1 + \psi_2 = A_1 \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega t) + A_2 \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega t)$$

and the time-averaged intensity $\langle I \rangle$ is

$$\begin{aligned} \langle I \rangle &= \langle \psi^2(x, y, t) \rangle = \langle (A_1 \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega t) + A_2 \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega t))^2 \rangle \\ &= \langle A_1^2 \cos^2(\mathbf{k}_1 \cdot \mathbf{r} - \omega t) \rangle + \langle A_2^2 \cos^2(\mathbf{k}_2 \cdot \mathbf{r} - \omega t) \rangle + \langle 2A_1A_2 \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega t) \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega t) \rangle \end{aligned}$$

Let us look more carefully at each term in the expression for the time-averaged intensity. The first two terms are simply the time-averaged intensities of each of the plane waves in the absence of the other:

$$\langle A_1^2 \cos^2(\mathbf{k}_1 \cdot \mathbf{r} - \omega t) \rangle = \frac{A_1^2}{2} = \langle I_1 \rangle \quad \text{and} \quad \langle A_2^2 \cos^2(\mathbf{k}_2 \cdot \mathbf{r} - \omega t) \rangle = \frac{A_2^2}{2} = \langle I_2 \rangle.$$

We can simplify the third term using the trigonometric identity

$$\cos \theta_1 \cos \theta_2 = \frac{1}{2} \cos(\theta_1 - \theta_2) + \frac{1}{2} \cos(\theta_1 + \theta_2)$$

such that

$$\begin{aligned} \langle 2A_1A_2 \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega t) \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega t) \rangle &= \langle A_1A_2 \cos((\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r}) \rangle \\ &+ \langle A_1A_2 \cos((\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r} - 2\omega t) \rangle \end{aligned}$$

Now recall our knowledge of time averages. The first term on the right-hand-side of the above equation is a constant in time (it does not depend on t), and thus averaged over a long time its value is simply equal to the value shown in brackets. The second term is just a simple harmonic function of time, and thus its average over a long time is simply zero. Therefore

$$\langle 2A_1A_2 \cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega t) \cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega t) \rangle = A_1A_2 \cos((\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r}).$$

Combining all of these results, *the time-averaged intensity associated with the interference of 2 plane waves of light is*

$$\boxed{\langle I \rangle = \langle I_1 \rangle + \langle I_2 \rangle + 2\sqrt{\langle I_1 \rangle \langle I_2 \rangle} \cos((\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r})} .$$

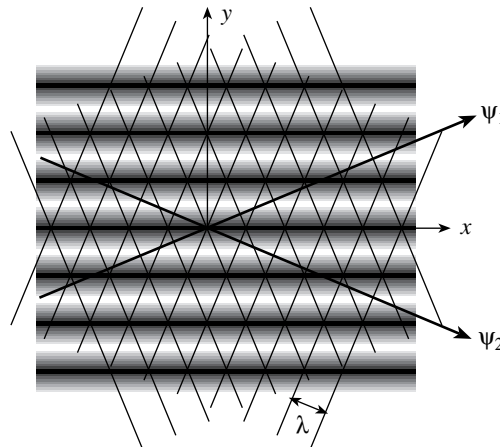
For our example in the drawing above

$$\begin{aligned} (\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r} &= \mathbf{k}_1 \cdot \mathbf{r} - \mathbf{k}_2 \cdot \mathbf{r} = k_{1x}x + k_{1y}y - k_{2x}x - k_{2y}y \\ &= k(x \cos \theta + y \sin \theta - x \cos \theta - (-y \sin \theta)) \end{aligned}$$

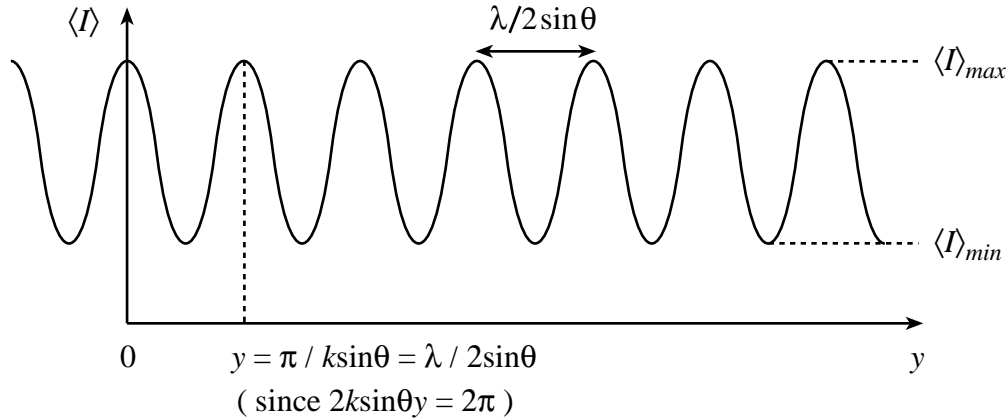
or

$$\boxed{(\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{r} = 2k \sin \theta y} .$$

Notice that the resulting intensity pattern does not depend on the position x . The pattern consists of interference fringes that are parallel to the x axis, but vary spatially along the y direction. The fringes might look like the following.



If we pick a fixed value of x and look in more detail at the variation of the intensity along the y direction, then we see a sinusoidally varying intensity pattern:



As shown in the diagram above, the fringes are spaced by a distance

$$\Delta y = \frac{\lambda}{2 \sin \theta} .$$

Often we wish to quantify the fringe contrast, where the *fringe contrast is the ratio of the peak intensity to the valley intensity*. Poor contrast means the fringes are barely visible, while perfect contrast means they are very clear. The contrast is usually measured by a parameter called the fringe visibility, V , which varies between 0 and 1 ($0 \leq V \leq 1$), and is defined as follows:

Fringe Visibility, V : A measure of the fringe contrast in a spatially varying intensity pattern, where a $V = 0$ implies the poorest contrast (no fringes!) and $V = 1$ implies perfect contrast (peaks are infinitely larger than valleys). Mathematically,

$$V = \frac{\langle I \rangle_{max} - \langle I \rangle_{min}}{\langle I \rangle_{max} + \langle I \rangle_{min}} .$$

For the case of 2-wave interference described above, we can easily determine $\langle I \rangle_{max}$ and $\langle I \rangle_{min}$. In particular, looking at the expression for $\langle I \rangle$, the maximum value occurs when the cosine = 1, and the minimum occurs when the cosine = -1. Therefore

$$\langle I \rangle_{max} = \langle I_1 \rangle + \langle I_2 \rangle + 2\sqrt{\langle I_1 \rangle \langle I_2 \rangle}$$

and

$$\langle I \rangle_{min} = \langle I_1 \rangle + \langle I_2 \rangle - 2\sqrt{\langle I_1 \rangle \langle I_2 \rangle} .$$

so we can rewrite the visibility in terms of $\langle I_1 \rangle$ and $\langle I_2 \rangle$ as follows:

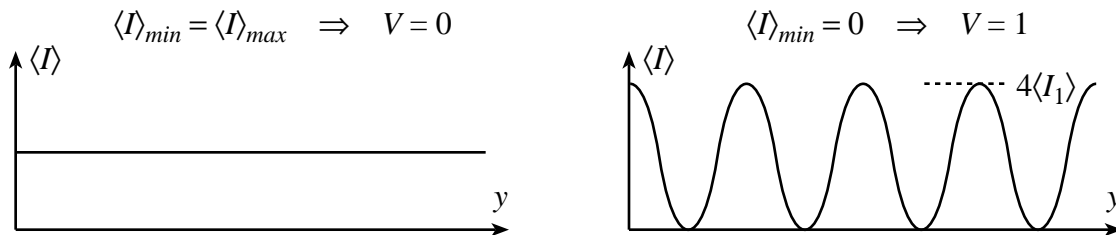
$$V = \frac{2\sqrt{\langle I_1 \rangle \langle I_2 \rangle}}{\langle I_1 \rangle + \langle I_2 \rangle} .$$

Thus we see

$$\text{when } \langle I_1 \rangle = 0 \text{ or } \langle I_2 \rangle = 0 \Rightarrow \langle I \rangle_{max} = \langle I \rangle_{min} \Rightarrow V = 0 ,$$

$$\text{and when } \langle I_1 \rangle = \langle I_2 \rangle \Rightarrow \langle I \rangle_{min} = 0 \Rightarrow V = 1 .$$

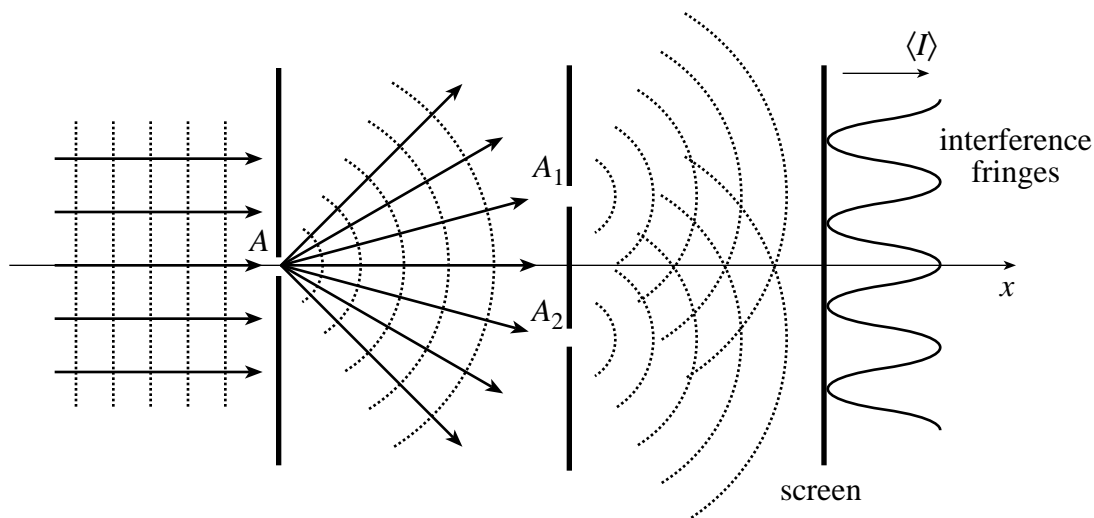
These two extreme cases are shown in the drawings below.



3. Young's Double-Slit Experiment — Locating Fringes by Tracking Phase

Young's Double-Slit experiment has tremendous historical significance, since it was instrumental in demonstrating that light indeed behaves like a wave. However, its importance in optics is more than historical — this experiment is an ideal system for illustrating, understanding, and analyzing numerous properties of the wave nature of light, including *interference*, *diffraction*, and *coherence*.

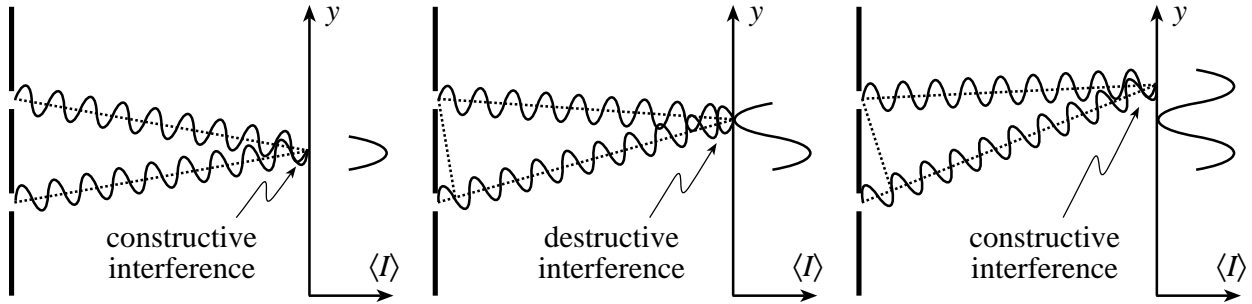
The figure below shows the experimental arrangement. A point source of light illuminates 2 tiny apertures in an opaque screen, and the resulting intensity pattern is observed on a viewing screen.



The first aperture A ensures that the light waves that arrive at apertures A_1 and A_2 originate from a single point source and thus have equal (or at least correlated) phase. The degree to which the phase at two points in a light wave is correlated is called *coherence*. The fact that light spreads out as if from a point source when it is passed through a small aperture is a fundamental consequence of the wave nature of light called *diffraction*.

The secondary apertures A_1 and A_2 act like point sources of 2 spherically (in 3 dimensions) or cylindrically (in 2 dimensions) expanding waves which interfere and thus produce fringes on the viewing screen.

What is the spacing of the fringes on the screen? To find out, we can look for constructive and destructive interference of the 2 source waves at the screen by *tracking the phase*, as illustrated in the drawing below.



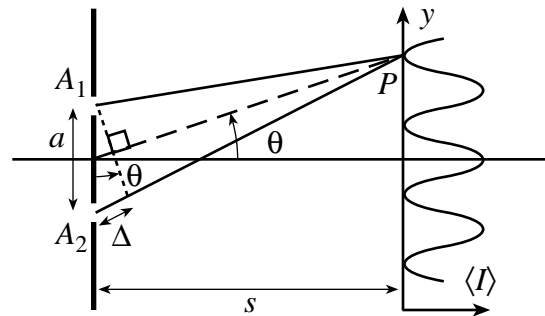
We can see that as we trace our way along the interference fringes, we pass through 1 full fringe period every time the path length difference changes by 1 wavelength. This result is an important principle of 2-wave interference.

When 2 waves interfere, each fringe in the resulting intensity pattern corresponds to a path length difference between the 2 waves of an integer number of wavelengths.

We can quantify the fringe spacing Δy on a screen placed a distance s behind the 2 apertures separated by a spacing a as follows.

Constructive interference occurs when
 $\overline{A_1P} - \overline{A_2P} \equiv \Delta = m\lambda$ ($m = \text{integer}$).

Destructive interference occurs when
 $\overline{A_1P} - \overline{A_2P} \equiv \Delta = m\lambda + \lambda/2 = (m + \frac{1}{2})\lambda$.



As long as we place the screen far enough away from the apertures, so that $s \gg a$, then

$$\Delta \cong a \sin \theta \quad (s \gg a).$$

Furthermore, for small angles θ , such that $y \ll s$, we can make the approximation

$$\sin \theta \cong \tan \theta \cong y/s \quad (y \ll s).$$

Putting these results together, we see that the locations y_m of the bright fringes (which correspond to constructive interference) and the spacing Δy between these fringes are given by

$$y_m \cong m \frac{\lambda s}{a} \Rightarrow \text{bright fringe};$$

$$\Delta y = y_{m+1} - y_m \cong \frac{\lambda s}{a}.$$

