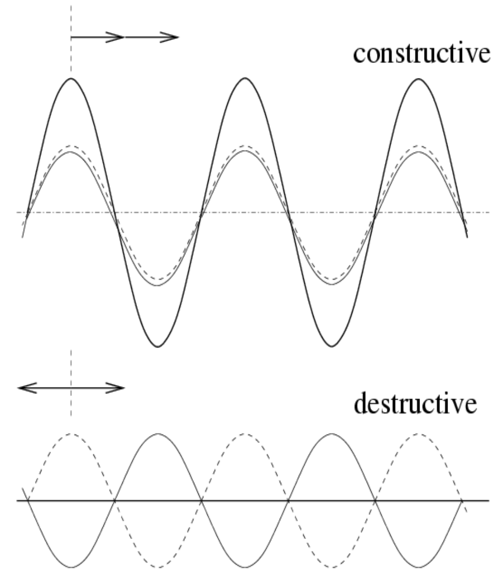
3 Interference and Diffraction

3.1 Introduction

Interference is the superposition of two or more harmonic waves meeting in space. If there is a phase difference of 0 (or $2n\pi$ where n is an integer) constructive interference occurs. On the other hand a phase difference of $(2n+1)\pi$ leads to destructive interference. We will see later how we can use phasors to help see what is happening.

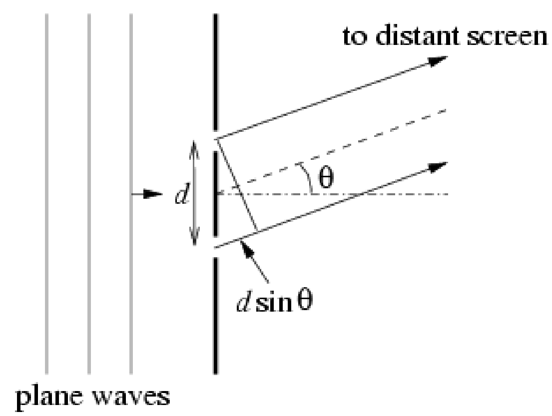


Diffraction occurs when a wave is partially obstructed and it 'bends round' such obstructions. The effect is only noticeable when the size of the diffracting obstacle is comparable to the wavelength λ.

It can be understood using the same Huygens' Principle as before: imagine the unobstructed part of the wave acts as a set of sources of secondary wavelets and then superpose all the resulting waves, taking account both of their amplitude and phase.

3.2 Young's Double Slits

So far in 1A you showed that light is also a transverse wave obeying a wave equation for the electric field. The Young's slits configuration shows the wave properties and consists of two identical and very narrow slits separated by a distance d. Each slit acts as a single source of secondary wavelets. In order to work out the resultant distribution of light, we need to sum the two wavelets, one originating from each slit and which have equal amplitudes. We will assume the incident wave is a plane wave normal to the slits so that the two secondary sources are in phase at the aperture. We want to work out the diffraction pattern observed on some very distant screen, distance R away.

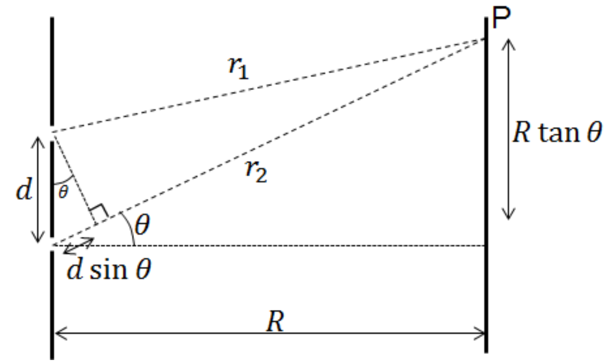


- When the path difference $d\sin\theta$ is an integral number of wavelengths, corresponding to a phase difference of $2n\pi$ there is constructive interference leading to maximum amplitude.

- When the path difference is $(n+\frac{1}{2})\lambda$, corresponding to a phase difference of $(2n+1)\pi$, destructive interference occurs and there is zero (minimum) amplitude.

We will consider the case where the distance to the screen R is much greater than the slit spacing d. We want to evaluate the intensity at some point P on the screen.

$$r_2 \approx r_1 + d\sin\theta$$



By superposing the waves arriving at P from the two slits, we find the resultant ψ_p is given by

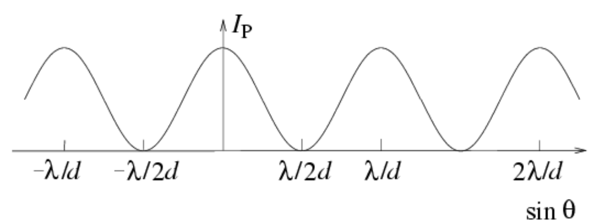
$$\psi_p = A\cos[\omega t - kr_1] + A\cos[\omega t - kr_2]$$
$$= A\cos[\omega t - kr_1] + A\cos[\omega t - k(r_1 + d\sin\theta)]$$

$$\boxed{\psi_p = 2A\cos\left[\omega t - kr_1 + k\left(\frac{d}{2}\right)\sin\theta\right]\cos\left[k\left(\frac{d}{2}\right)\sin\theta\right]}$$

This means that the intensity at P is $\propto \cos^2[k(d/2)\sin\theta]$ (the first $\cos^2$ term averages out over time) and this is seen to oscillate in amplitude with angle. The intensity at P (see 1A Gravitational and Electromagnetic Fields)

$$\boxed{I_P \propto \psi_p^2 \propto \cos^2[k(d/2)\sin\theta]}$$

Changing θ as we move across the screen leads to alternating regions of light and dark, the intensities varying as cos-squared. Thus *fringes* are seen and the successive maxima occur when



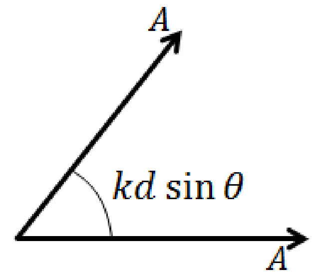
$$\cos[k(d/2)\sin\theta] = \pm 1$$

$$\Rightarrow \frac{kd\sin\theta}{2} = n\pi$$

$$\Rightarrow \sin\theta = \frac{2n\pi}{kd} = \frac{2n\pi}{d}\frac{\lambda}{2\pi} = \frac{n\lambda}{d}$$

For the case of small θ (so $\sin\theta \approx \tan\theta \approx \theta$), the angular separation of successive fringes is given by λ/d, corresponding to a linear separation of $R\lambda/d$ on the screen. The 'order' of the maxima is given by n.

3.3 Phasor Diagrams



Phasor diagrams provide a graphical way to sum amplitudes. We can use them explicitly to work out the resultant in the above Young's slit experiment. In this case we need to evaluate the sum of two oscillations with equal amplitudes and the same frequency. The phases of these two waves are given by

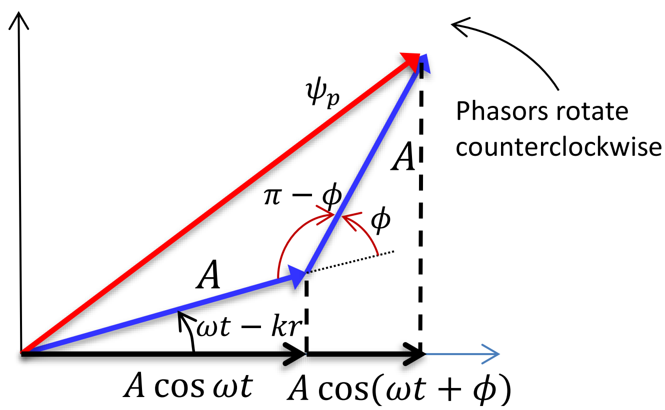
$\omega t - kr, \;\; \omega t - kr + kd\sin\theta.$

We are only interested in the time averages, so we set $\omega t - kr = 0$.

As θ varies, so does the resultant amplitude. Maxima occur when $kd\sin\theta = 2n\pi$, and the contributions are in phase. Conversely minima occur when $kd\sin\theta = (2n+1)\pi$. Then the contributions are out of phase and the net sum is zero.

3.3.1 Young's Slit Experiment using Phasors

We can use phasors to look again at the solution to the Young's slit problem.



In the diagram, we can use the cosine rule to work out the resultant wave amplitude ψ_p. However what we measure is the intensity of light, given by

$$\begin{aligned}\psi_p^2 &= A^2 + A^2 - 2A^2\cos(\pi - \phi)\\ &= A^2 + A^2 + 2A^2\cos\phi\\ &= 2A^2(1+\cos\phi)\end{aligned}$$

In this diagram, ϕ is the phase difference between the two slits, so $\phi = kd\sin\theta$ and it is the total length of the vector which is important,

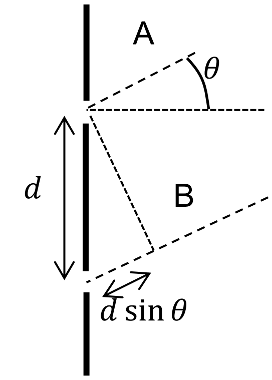
$$\psi_p^2 = 4A^2\cos^2\left(\frac{\phi}{2}\right) = 4A^2\cos^2\left(\frac{kd\sin\theta}{2}\right)$$

As before, when we worked it out explicitly, we see the intensity ψ_p^2 exhibits $\cos^2$ fringes.

Example

Two narrow slits separated by 1 mm, are illuminated by light of wavelength 600 nm and the interference pattern is viewed on a screen 2 m away. Use the phasor method to indicate how you can deduce the positions of successive minima and find the linear separation of these minima on the screen.

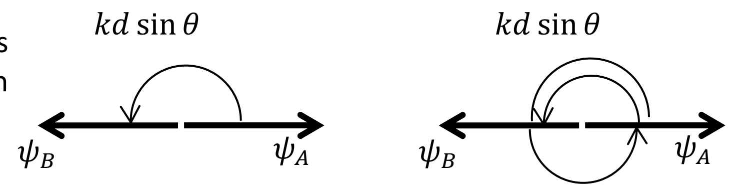
The relative phase between light from the two slits is given by $kd\sin\theta$. Looking at the phasor diagrams we can see that the first minimum occurs when



$$kd\sin\theta = \pi \quad \Rightarrow \quad \frac{2\pi}{\lambda} d\sin\theta = \pi$$
$$\Rightarrow \quad \sin\theta = \frac{\lambda}{2d}$$

The second minimum will be when

$$kd\sin\theta = 3\pi \qquad \Rightarrow \quad \sin\theta = \frac{3\lambda}{2d}$$



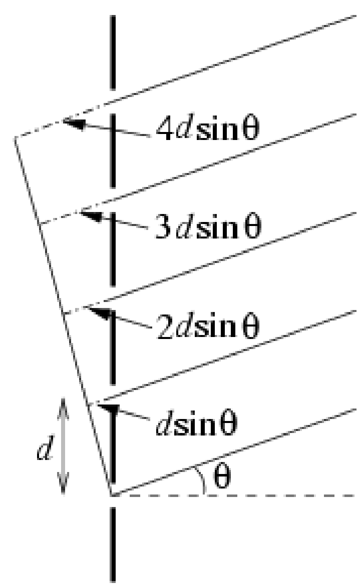
If $\theta \ll 1$ then the angular separation of minima is given by λ/d , leading to a separation of minima on screen given by

$$\frac{R\lambda}{d} = \frac{2 \times 600 \times 10^{-9}}{1 \times 10^{-3}} = 1.2 \times 10^{-3}\mathrm{m}$$

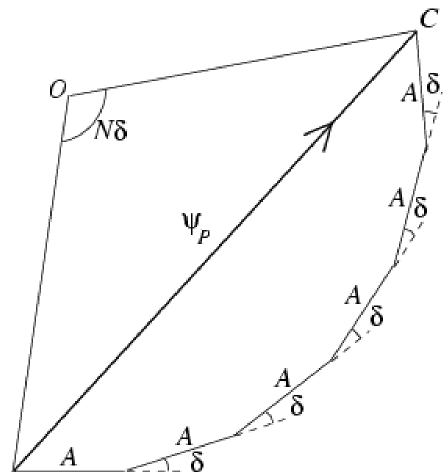
3.4 Diffraction Gratings

The same arguments can now be used to discuss another common situation, in which there are many slits. This is known as a diffraction grating.

- Suppose there are N slits. There will be constructive interference if the path difference between successive slits is an integral number of wavelengths.
- This leads to regions of maximum amplitude whenever $d\sin\theta = n\lambda$.
- The order of the maximum is given by n, so when n=0 we have the zeroth order maximum, while n=1 corresponds to the first order maximum etc.



As before, we can use a phasor diagram to evaluate the intensity at points on the screen. In this case we need to superpose the oscillations at a point P on the screen, each with equal amplitude, A, and the same frequency, and with phases given by



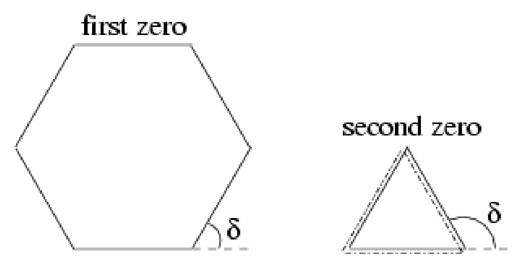
$$\omega t - k(r - md\sin\theta) = \omega t - kr + m\delta$$

In this expression, m is an integer with values from 0 to N–1 and $\delta = kd\sin\theta$. In this case the phasor diagram is a section of a polygon as shown.

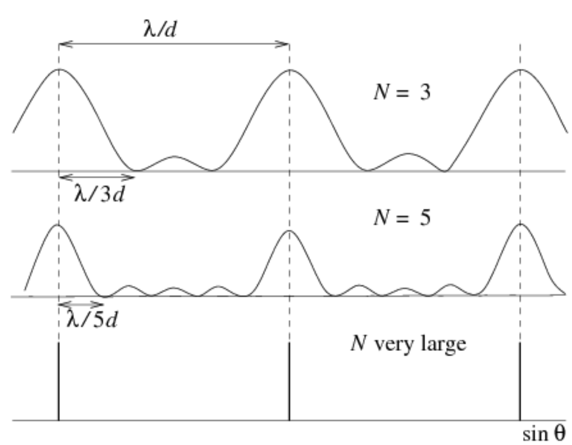
Maxima will occur whenever all the vectors add in phase. This requires that

$$\boxed{\delta = \pm 2n\pi \quad \text{ie. } kd\sin\theta = 2n\pi \quad \text{or } d\sin\theta = n\lambda}$$

- Positions of maxima will be wavelength dependent.
- Positions of the maxima do not depend upon the number of slits, although their sharpness (or narrowness) and intensity do.
- From the phasor diagram it should be clear that the intensity varies as N^2.
- The sharpness is determined by the value of δ for which the resultant goes to zero for the first time ie when $\delta = kd\sin\theta = \frac{2\pi}{N}$ which implies that $\sin\theta = \frac{\lambda}{Nd}$
- Hence the width of each maximum is proportional to $1/N$.



Zeros (minima) occur whenever $\delta = 2p\pi/N$ (p an integer which is not a multiple of N). This implies there are N–1 minima and N–2 subsidiary maxima between principal maxima; the subsidiary maxima are very weak compared with the principle maxima. As the number of slits N is increased, the principal maxima become sharper and more intense, and the intensities of the subsidiary maxima become negligible by comparison.



Diffraction gratings are used to investigate wavelengths and shapes of spectral lines – the greater N the greater the detail observed.

Example

A diffraction grating with slit spacing d is used to analyse the spectrum emitted by a gas.

(i) What is the angular separation in the nth order spectrum of two lines of wavelength λ_1 and λ_2 ($\lambda_2 > \lambda_1$)?

(ii) What is the minimum width of the grating required to resolve these two lines (i.e. to see them clearly as two distinct lines) in the first order spectrum?

(iii) The sodium D-lines have wavelengths of 589.00 and 589.59 nm. A diffraction grating with 2000 lines per centimetre is to be used to resolve these lines in the first-order spectrum. What width must the beam of sodium light have when it falls on the grating? (You may assume all angles are very small so $\sin\theta \approx \theta$).

(i) In general for the nth order maximum

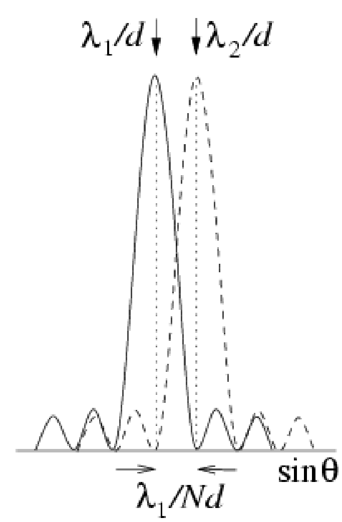
$$\sin\theta = \frac{n\lambda}{d}$$

$$\Rightarrow \quad \sin\theta_1 = (n\lambda_1)/d \qquad \text{and} \qquad \sin\theta_2 = (n\lambda_2)/d$$

$$\Delta\theta = \frac{n(\lambda_2 - \lambda_1)}{d}$$

(ii) The criterion that two lines can be resolved from each other is that: at that their minimum separation the maximum of one line lies over the first zero of the other – this is known as the *Rayleigh Criterion*. Thus

$$\frac{\lambda_2}{d} - \frac{\lambda_1}{d} = \frac{\lambda_1}{Nd} \qquad \Rightarrow \quad N = \frac{\lambda_1}{\lambda_2 - \lambda_1}$$



Thus the required width w is given by

$$w > Nd = \frac{\lambda_1}{\lambda_2 - \lambda_1} d$$

iii) First order spectrum, so n=1. Using the Rayleigh criterion, the number of slits the beam must illuminate is $N = \frac{\lambda_1}{\lambda_2 - \lambda_1}$

The slits are a distance d apart where

$$d = \frac{10^{-2}}{2000} \text{ m}$$

Thus the width of the beam of sodium light must be

$$Nd = \frac{\lambda_1}{\lambda_2 - \lambda_1} d = \frac{589.00 \times 10^{-9}}{0.59 \times 10^{-9}} \times \frac{1}{2 \times 10^5} \text{ m}$$

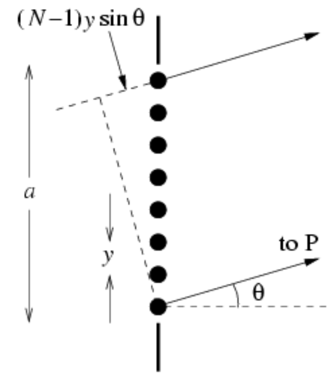
$$= 4.9 \times 10^{-3} \text{m}$$

Thus the beam must be about 5 mm wide.

3.5 Slit of Finite Width

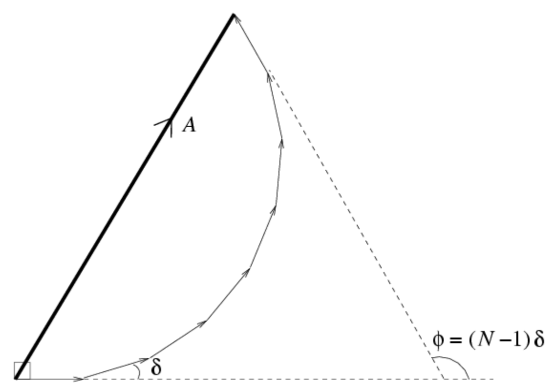
3.5.1 Phasor Approach

Consider a slit of width a, which we can imagine as divided into N equal intervals. Using Huygen's Principle we then assume a secondary wavelet starts at the midpoint of each of these, and we can sum the effect of all these at a point P on some distant screen using phasors. The total phase difference between the top and bottom of the slit is $ka \sin\theta$.



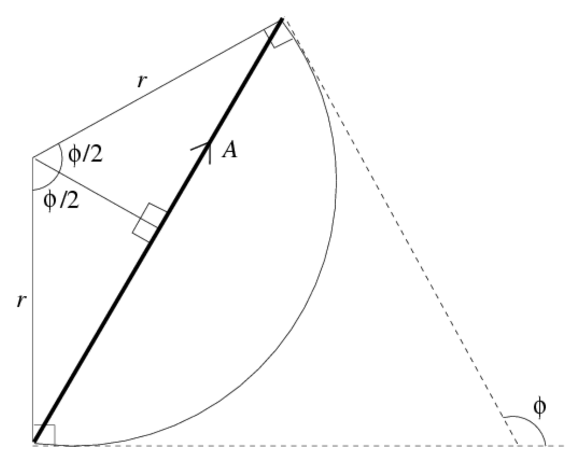
slit with N secondary sources

When N is very large the phasor diagram for calculating the resultant amplitude tends to an arc, and the chord gives us the resultant. Then



$$\sin\left(\frac{\phi}{2}\right) = \frac{A/2}{r} \quad \text{so} \quad A = 2r \sin\left(\frac{\phi}{2}\right)$$

where r is the radius of the arc.



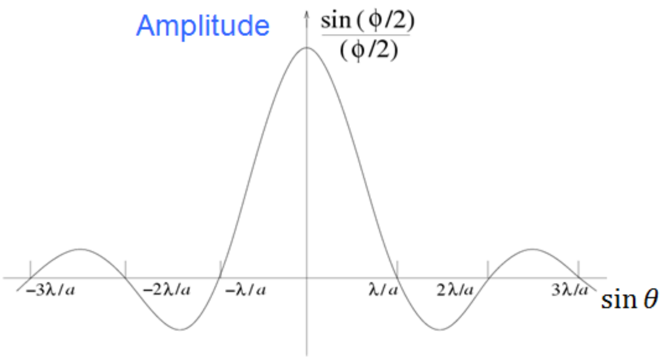
The length of the arc is NA_0, where A_0 is the amplitude due to a single source. So $\phi = NA_0/r = A_{max}/r$ and

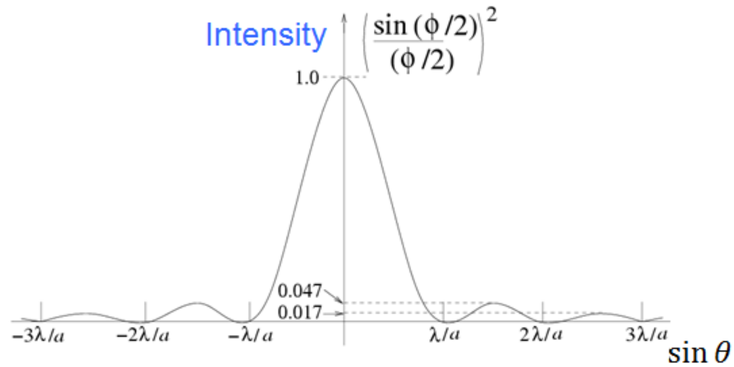
$$A = \frac{2A_{max}}{\phi} \sin\left(\frac{\phi}{2}\right) = A_{max} \frac{\sin(\phi/2)}{\phi/2}$$
$$= A_{max} \operatorname{sinc}(\phi/2)$$

This gives us the definition of the *sinc function* as

$$\boxed{\operatorname{sinc} x = \sin x / x}$$

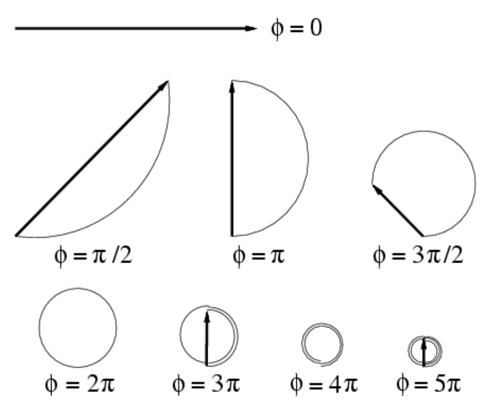
The behaviour of the sinc function is shown in the diagrams. Minima occur whenever $\theta = \frac{n\lambda}{a}$.





As θ is varied, corresponding to moving across the screen, the total amplitude goes through alternating maxima and minima. The maxima decrease in amplitude as ϕ increases. There is a minima every time $\phi = 2n\pi$. This gives us the condition $ka\sin\theta = 2n\pi$ or equivalently $\frac{2\pi}{\lambda}a\sin\theta = 2n\pi$. So

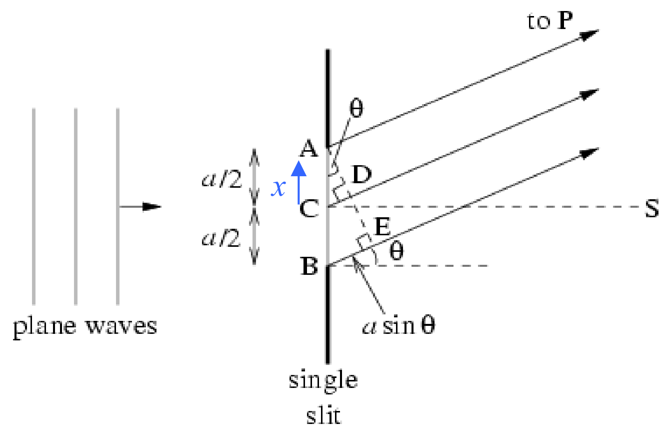
$$\boxed{\sin\theta \sim \theta = \frac{n\lambda}{a}}$$



On axis there will always be a maximum since all contributions are in phase.

3.5.2 Slit of Finite Width via Integration

Considering phase differences relative to the middle of the slit C, the phase difference for A is $-k(a/2)\sin\theta$, and for B is $+k(a/2)\sin\theta$. In general the phase difference is $kx\sin\theta$ and we use $\phi = ka\sin\theta$.



The maximum amplitude $A_{max} = A_0 a$, where A_0 is the amplitude per unit length of the slit. We add waves from each small section of the slit dx. We can then integrate all the contributory wavelets across the slit to give the total amplitude at P as:

$$A_P = A_0 \int_{-a/2}^{+a/2} \exp(ikx\sin\theta)\, dx$$

$$= \frac{A_0}{ik\sin\theta}[\exp(ikx\sin\theta)]_{-a/2}^{+a/2}$$

$$= \frac{A_0}{ik\sin\theta}[\exp(ik(a/2)\sin\theta) - \exp(-ik(a/2)\sin\theta)]$$

$$= \frac{2A_0}{k\sin\theta}\sin[k(a/2)\sin\theta]$$

Hence

$$A_P = \frac{2A_{max}}{ak\sin\theta}\sin[k(a/2)\sin\theta]$$
$$= A_{max}\,\mathrm{sinc}[\phi/2]$$