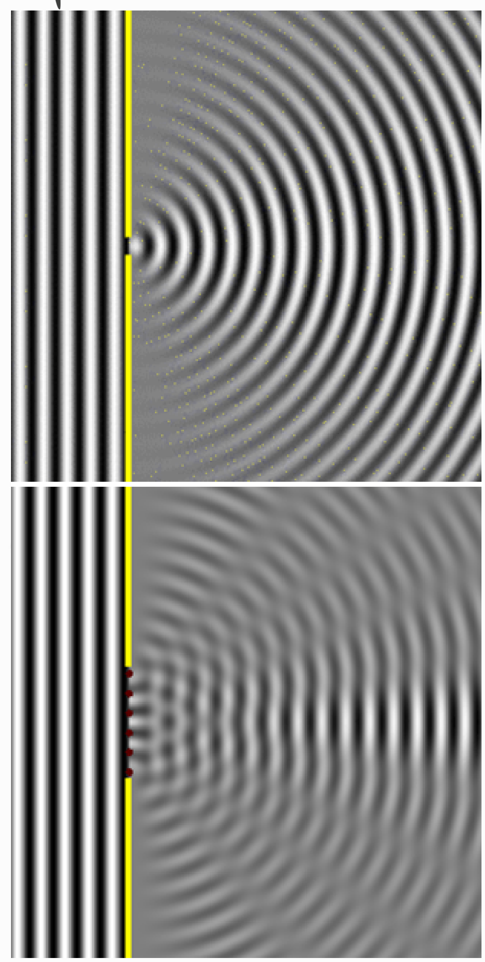
Lecture 26 : Diffraction

So far, we've discussed interference as an effect arising from point sources. However, real sources such as the two slits in Young's experiment, have finite width.



According to the Huygens Principle, each point on a wavefront acts as a source of new waves (wavelets).

The diffraction pattern is given by the interference of all the wavelets considering their amplitude and relative phases. Thus diffraction is an extension of our discussion on interference, and there is no significant distinction between the two effects:

Interference addresses the superposition of a few waves.

Diffraction is used when treating a large number of waves.

Diffraction and interference effects are technologically important in many regimes including x-rays, optical imaging, acoustic design, radio and cell phones.

Diffraction problems are solved in one of two limits, named after Fraunhofer (far-field) and Fresnel (near field).

A relative length scale is set by the wavelength, the size of the aperture, and the distance to the observe

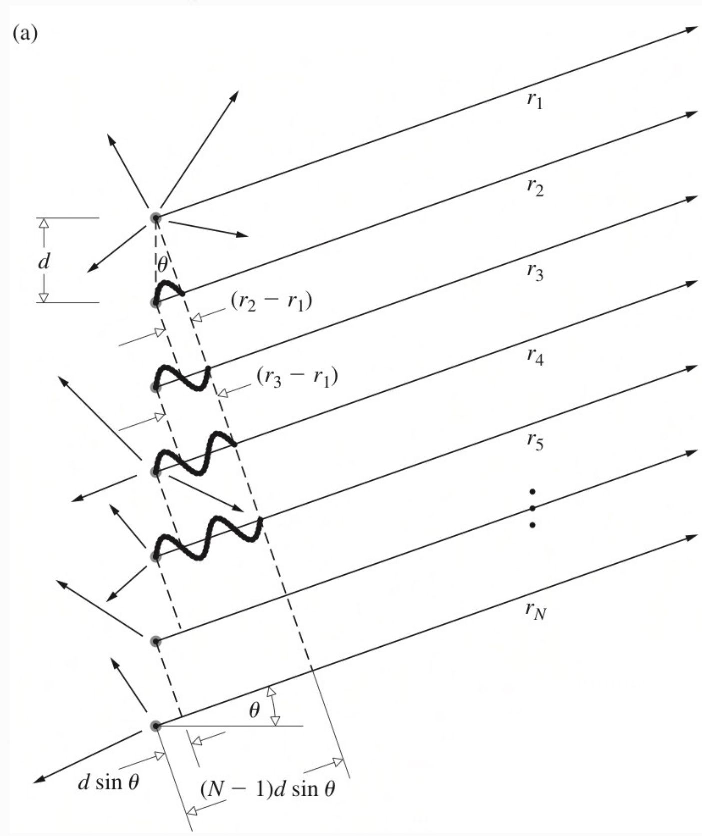
For a distance $\underline{d}$ along the axis of an aperture, the phase difference between waves at the center vs. the edge of the aperture of width $\underline{a}$ is:

$$\Delta\phi = \frac{2\pi}{\lambda}\left[\sqrt{d^2+\frac{a^2}{4}} - d\right] = \frac{2\pi}{\lambda} d\left[\sqrt{1-\frac{a^2}{4d^2}} - 1\right] \approx \frac{2\pi}{\lambda}\left(\frac{a^2}{8d}\right)$$

- If $d \gg \frac{a^2}{\lambda}$, then it is reasonable to consider rays from source to screen as parallel and the phase differences to be small (Fraunhofer)

- If $d < \frac{a^2}{\lambda}$, the phases differences become large and the rays cannot be approximated as parallel (Fresnel).

To understand diffraction, let us start by considering interference arising from N sources, which we assume for simplicity are identical.



Assuming:

- No initial phase difference;
- Sources are small compared to λ;
- Observation point P is far away, at distance $r_1, r_2, \dots, r_N$ from sources.

The amplitudes of the waves arriving at P are equal: $E_0(r_1) = E_0(r_2) = \dots E_0(r_N) = E_0(r)$

The sum of the interfering wavelets at P is the real part of:

$$E = E_0(r)\, e^{i(kr_1 - \omega t)} + e^{i(kr_2 - \omega t)} + \dots + e^{i(kr_N - \omega t)}$$

$$= E_0(r)\, e^{-i\omega t}\, e^{ikr_1}\left[1 + e^{ik(r_2 - r_1)} + e^{ik(r_3 - r_1)} + \dots + e^{ik(r_N - r_1)}\right.$$

From the figure above, we can see that $r_2 - r_1 = d\sin\theta$

and in general, $r_N - r_1 = (N-1)d\sin\theta$

The phase difference between adjacent sources is

$$\delta = \frac{2\pi}{\lambda}(r_N - r_{N-1}) = \frac{2\pi}{\lambda} d\sin\theta$$

And maxima occur when $\delta = 2\pi m = \frac{2\pi}{\lambda} d\sin\theta_m$

$$\Rightarrow \boxed{d\sin\theta_m = m\lambda}$$

Diffraction gratings make use of this effect:

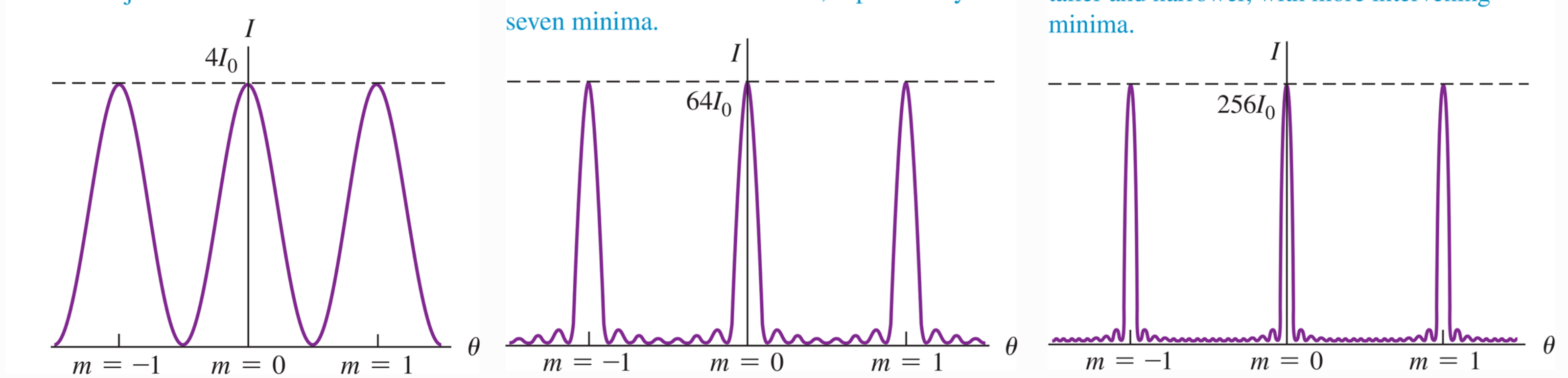
$$m\Delta\lambda = d\cos\theta\,\Delta\theta \quad \Rightarrow \quad \Delta\theta = \frac{m\Delta\lambda}{d\cos\theta}$$

Two different wavelengths show principle maxima at different values of θ_m. Thus, diffraction gratings separate wavelengths and are often used as "spectral" devices. Moreover, the width of a strong peak is approximately $\delta\theta = \lambda/(Nd\cos\theta)$ where N = number of slits.

(a) $N = 2$: two slits produce one minimum between adjacent maxima.

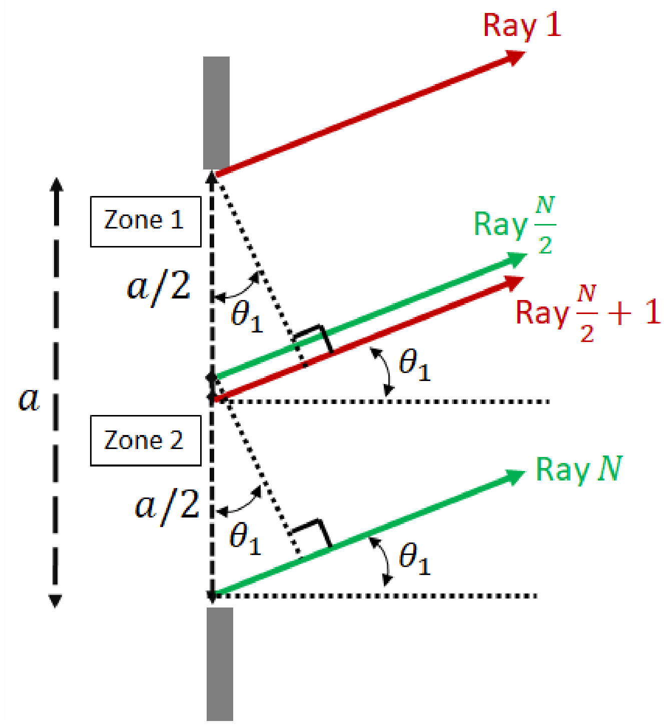
(b) $N = 8$: eight slits produce taller, narrower maxima in the same locations, separated by seven minima.

(c) $N = 16$: with 16 slits, the maxima are even taller and narrower, with more intervening minima.



Now let us consider a single slit of width $\underline{a}$, as a source of N wavelets.

The N sources emit N parallel rays at angle θ. If we divide the sources into two zones of width $a/2$, then angle θ_1 of the first intensity min occurs when all zone 1 rays destructively interfere with all zone 2 rays.

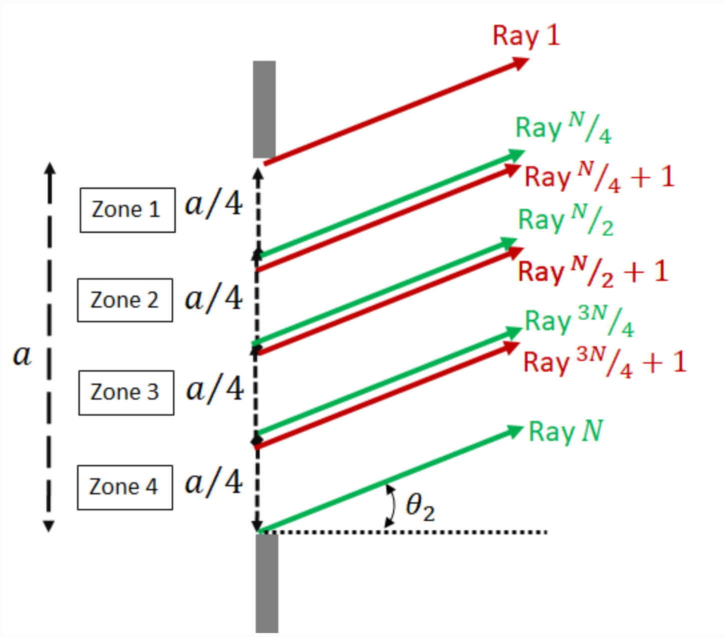


Pairwise destructive interference of the rays (e.g. Ray 1 and $\frac{N}{2}+1$ in zone 2, or Ray $\frac{N}{2}$ in zone 1 and ray N in 2 occurs when the path difference between a ray in zone 1 and its counterpart in zone 2 satisfy:

$(a/2 \sin\theta_1 = (\overset{\nearrow 0}{m} + 1/2))\lambda$

$$\frac{a}{2}\sin\theta_1 = \frac{\lambda}{2} \qquad \text{or} \qquad a\sin\theta_1 = \lambda$$

Angle of second diffraction minimum can be found

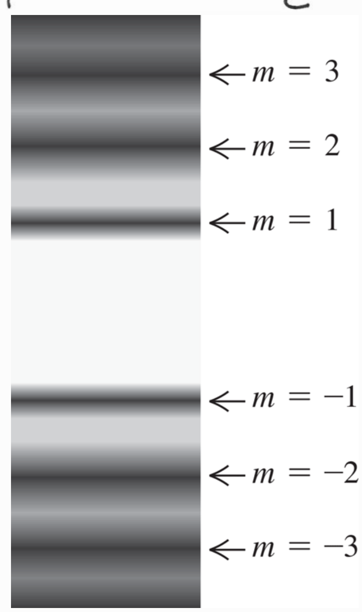


in a completely analogous fashion:

Break the slit up into four equal zones of width $a/4$, and consider pairwise destructive interference of rays in zone 1 & 2, and those in zones 3 & 4.

Destructive interference occurs when $\frac{a}{4} \sin\theta_2 = \frac{\lambda}{2}$

or, equivalently $a \sin\theta_2 = 2\lambda$



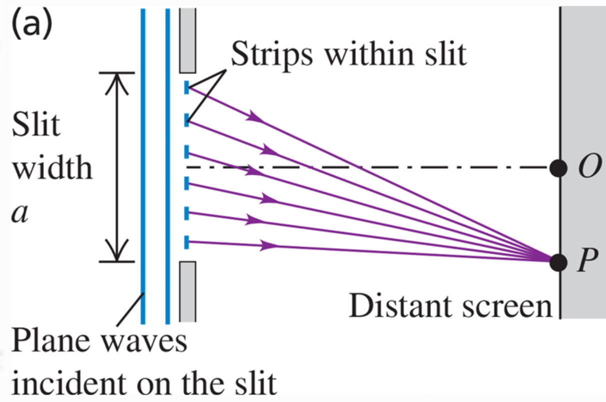
Other minima can be found in the same way by dividing the slit into 6, 8, etc.

The general relation for the angles at which there is an intensity min is:

$$\boxed{a \sin\theta_{m'} = m'\lambda}$$

where $m' = \pm 1, \pm 2, \pm 3, \ldots$

The intensity for single-slit diffraction can be obtained by adding all the wavelets to get the resultant wave and amplitude.



We can do this using phasors.

Assume there are N sources in a slit, creating N waves with equal spacing $\Delta x = \frac{a}{N}$ between sources.

The phase difference between wave 1 and 2 is:

$$\Delta\phi = k(\Delta x \sin\theta) = \frac{2\pi}{\lambda}(\Delta x \sin\theta)$$

$$\Rightarrow \quad \Delta\phi = \frac{2\pi a}{\lambda N}\sin\theta$$

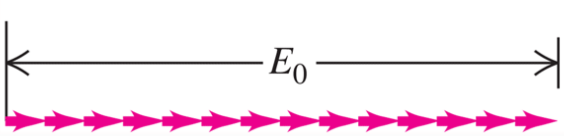
In fact, this is the phase difference between all adjacent waves.

Assuming that the amplitude of each wavelet is E_0/N:

For $\theta = 0$, $\Delta\phi = \frac{2\pi a}{\lambda N}\sin(0) = 0$, i.e.- all phasors line up and the net amplitude is

$$N\left(\frac{E_0}{N}\right) = E_0$$

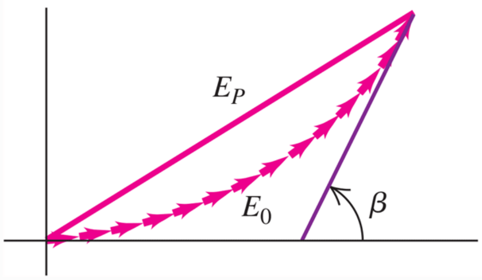
(b) At the center of the diffraction pattern (point O), the phasors from all strips within the slit are in phase.



For angles $\theta \neq 0$, the sum of the phasors along the arc is still

$$N\left(\frac{E_0}{N}\right) = E_0$$

(c) Phasor diagram at a point slightly off the center of the pattern; β = total phase difference between the first and last phasors.

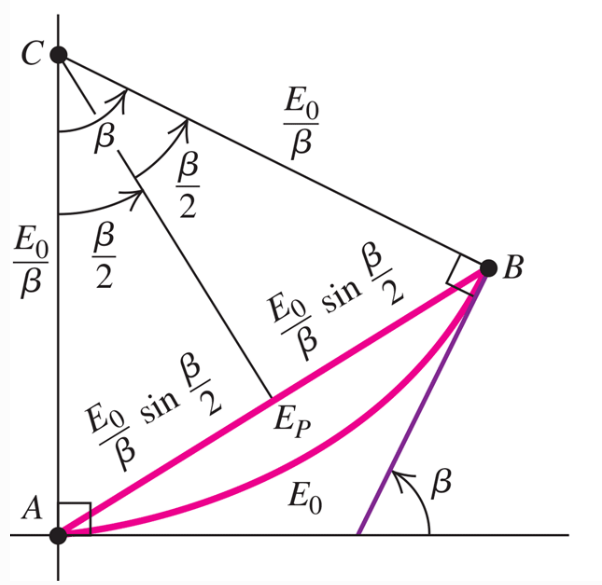


The angle the last phasor (for wave N) makes with the horizontal axis (direction of wave 1) is

$$\beta = N\Delta\phi = N\left(\frac{2\pi a}{\lambda N}\sin\theta\right) = \frac{2\pi a}{\lambda}\sin\theta$$

The amplitude of the electric field at point P on the screen is the resultant phasor vector E_p.

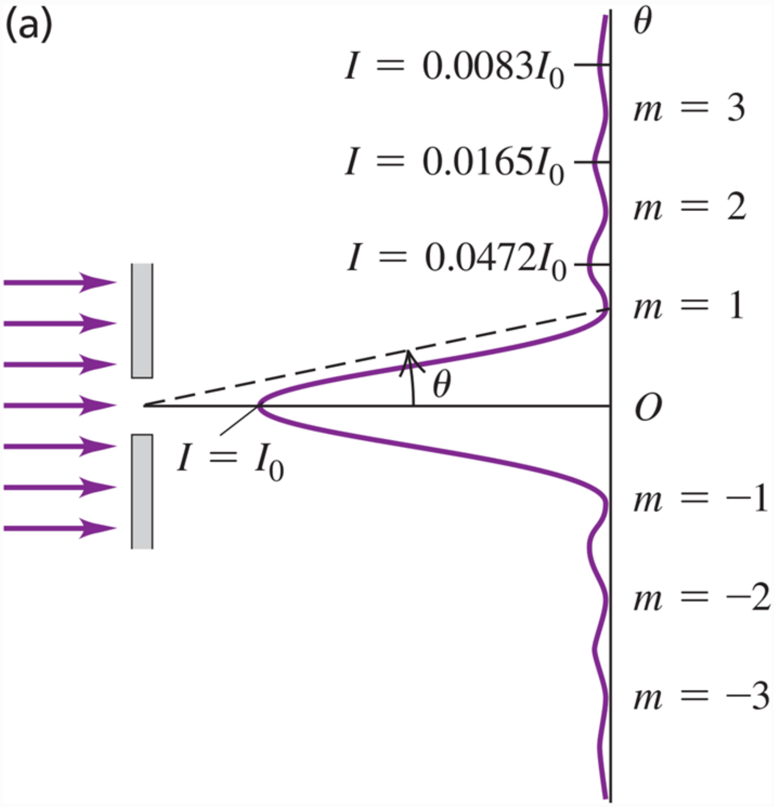
(d) As in (c), but in the limit that the slit is subdivided into infinitely many strips

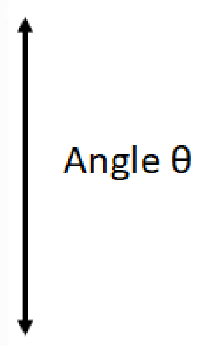


Approximating this phasor as resulting from a smooth, a continuous arc (i.e., taking $N \to \infty$) trig gives us:

$$E_p = \frac{2E_0}{\beta}\sin\left(\frac{\beta}{2}\right)$$

Thus,
$$I = I_0\left[\frac{\sin(\beta/2)}{\beta/2}\right]^2$$





Intensity pattern:

Note that the intensi(ty) of the first side max is about 1/20 of the central max.

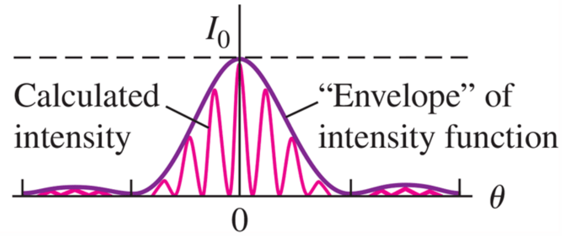
For two-slit interference, including the effect of diffraction for each slit, the resultant intensity is the product of the diffraction pattern of a single slit multiplied by the pattern for two-slit interference:

$$I(\theta) = I_0 \underbrace{\left[\frac{\sin(\beta/2)}{(\;\;\;}\right]^2}_{\text{single slit diffraction}} \underbrace{\cos^2\left(\frac{\alpha}{2}\right)}_{\text{two slit interference}}$$

where $\beta = \frac{2\pi a}{\lambda}\sin\theta$ and $\alpha = \frac{2\pi d}{\lambda}\sin\theta$

Two-slit interference peaks are in the same position, but their intensities are now modulated by single-slit diffraction which acts as an "envelope" for the intensity function. That is, the overall pattern is a convolution of the diffraction and two-slit interference effects.

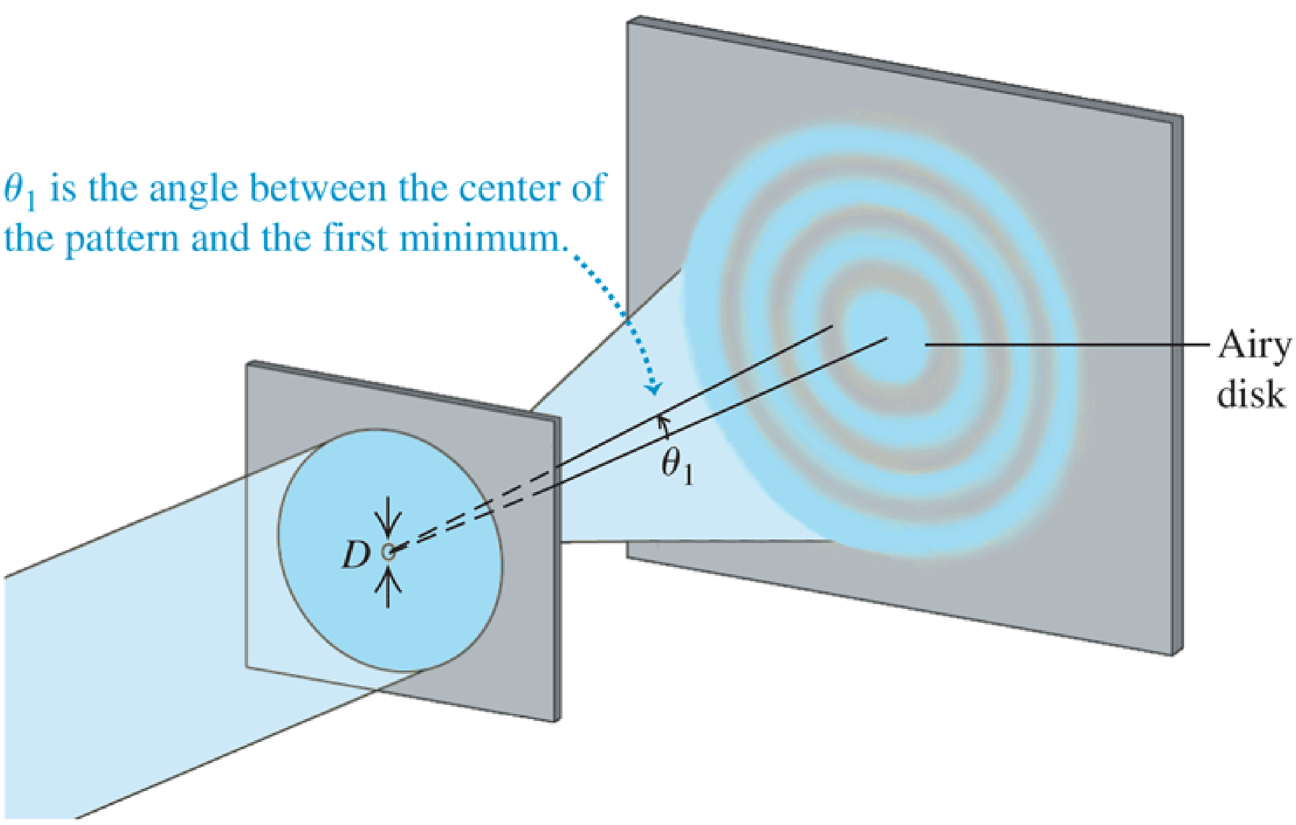
(c) Calculated intensity pattern for two slits of width a and separation $d = 4a$, including both interference and diffraction effects



(d) Photograph of the pattern calculated in **(c)**

For $d = 4a$, every fourth interference maximum at the sides ($m_i = \pm 4, \pm 8, \ldots$) is missing.

Diffraction from circular apertures also occurs, for the same reasons.



This function has a sharp peak, where the central max corresponds to a circular spot called an Airy disk. Its radius is given by the first zero of $J_1(\alpha)$. (Bessel function of the first order)

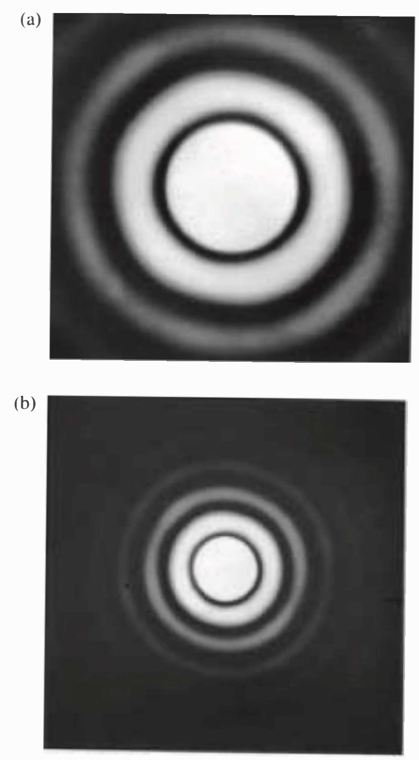
the first zero in J_1 occurs at:

$$ka\sin\theta = 3.832$$

$$\frac{2\pi}{\lambda} a\sin\theta = 3.832$$

$$\sin\theta = \frac{3.832}{2\pi}\frac{\lambda}{a} = 1.22\frac{\lambda}{2a} = 1.22\frac{\lambda}{D} \approx \theta \quad \text{(small angles)}$$

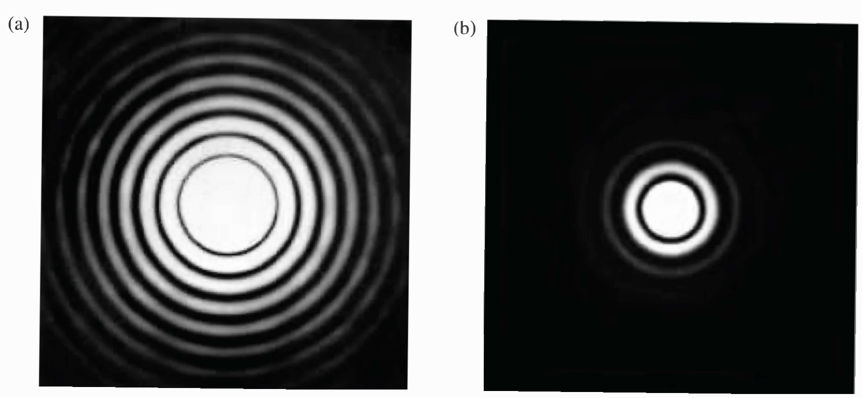
In an optical system, the image of a point source will be an Airy disk. This defines the fundamental limit to the resolution of the system.



Airy rings using (*a*) a 0.5-mm hole diameter and (*b*) a 1.0-mm hole diameter. (E.H.)

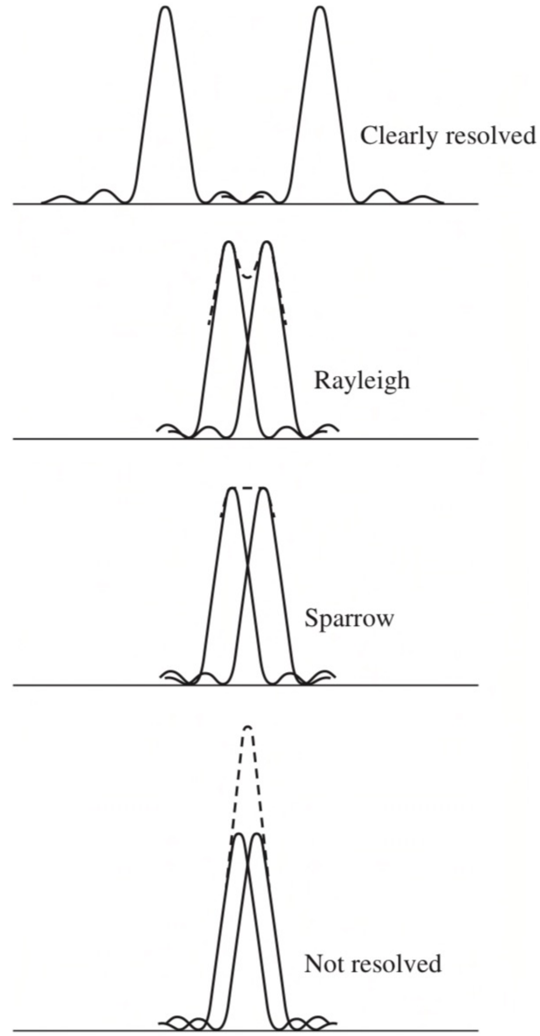
For example, if a lens has a focal length f, the radius of a point source on the screen will be:

$$\boxed{d \approx f\theta \approx f\left(\frac{1.22\lambda}{D}\right)}$$



(*a*) Airy rings—long exposure (1.5-mm hole diameter). (*b*) Central Airy disk—short exposure with the same aperture. (E.H.)

If we want to image two points, they need to be far enough apart to be distinguished.



The Rayleigh criterion for resolving two sources is when the center of one Airy disk falls on the first minimum of the other. In that case, the min resolvable angular separation is:

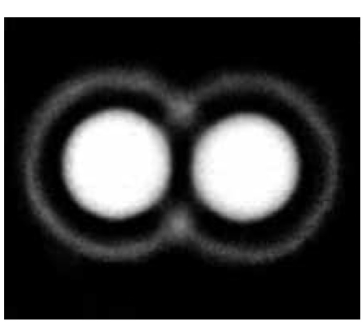
$$\boxed{(\Delta\varphi)_{min} = \theta = 1.22\,\frac{\lambda}{D}}$$

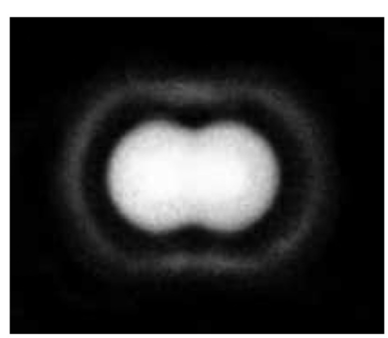
If $\Delta\ell$ is the center-to-center separation, the resolution limit is:

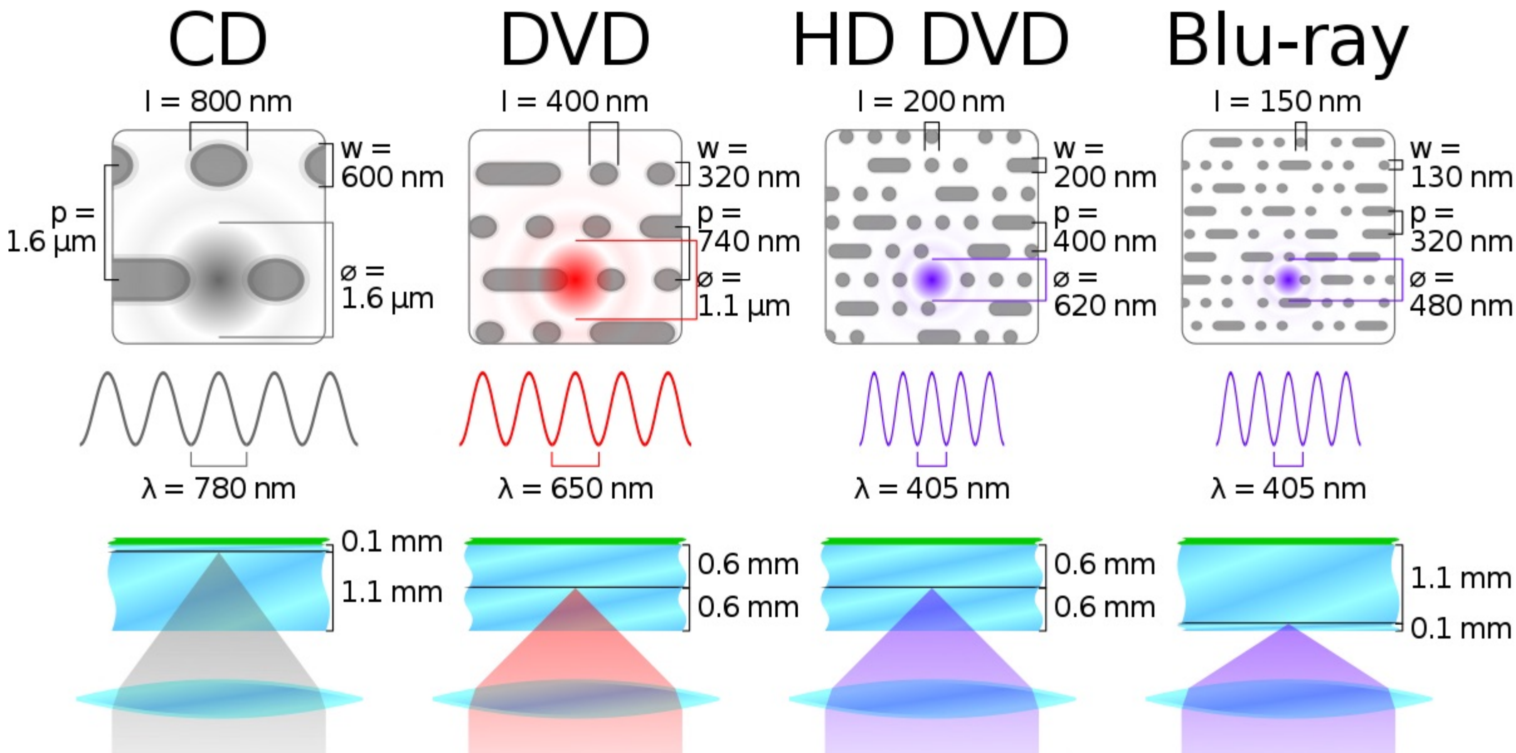
$$\boxed{(\Delta\ell)_{min} = 1.22\,\frac{f\lambda}{D}}$$

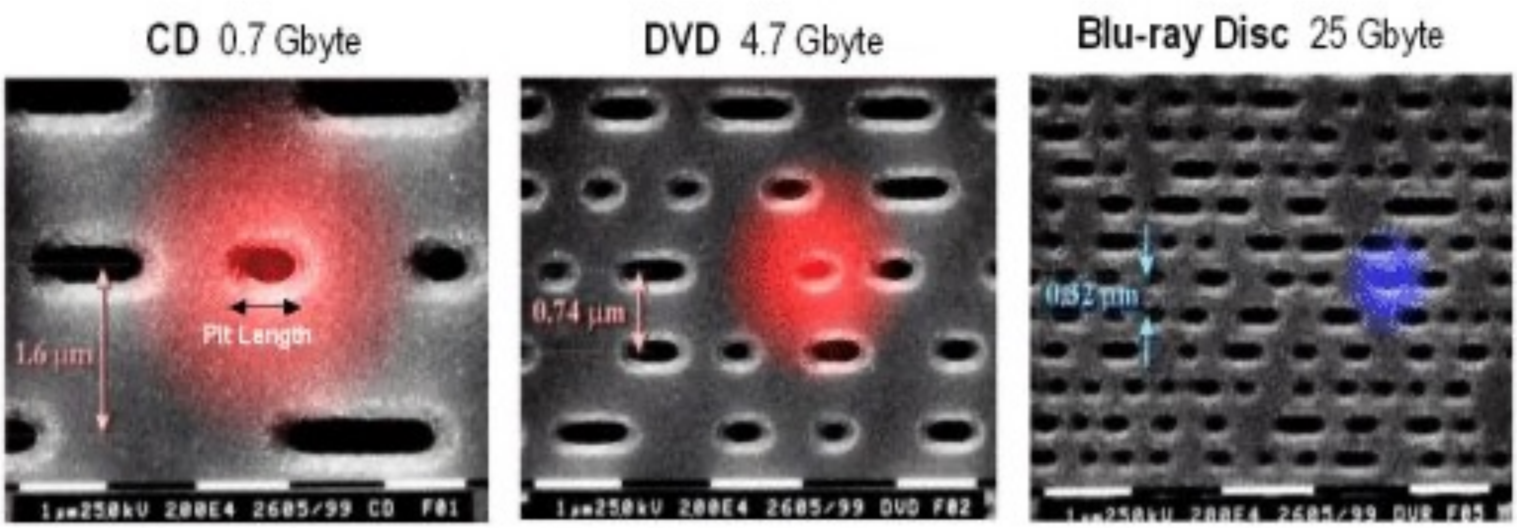
The resolving power for an image-forming system is generall defined as either $\dfrac{1}{(\Delta\varphi)_{min}}$ or $\dfrac{1}{(\Delta\ell)_{min}}$

A more sensitive criterion for resolving power is Sparrow's limit. In that case, two points can be resolved when their superposition results in a flat top.









Track Pitch: 1,6 micron	Track Pitch: 0,74 micron	Track Pitch: 0,32 micron
Minimum Pit Length: 0,8 µm	Minimum Pit Length: 0,4 µm	Minimum Pit Length: 0,15 µm
Storage Density: 0,41 Gb/inch²	Storage Density: 2,77 Gb/inch²	Storage Density: 14,73 Gb/inch²