DIFFRACTION OF LIGHT

Concepts of Fresnel and Fraunhoffer diffractions. Rectilinear propagation of light. Theory of Zone plate, comparison between zone plate and convergent lens. Fresnel's diffraction at a straight edge and wire. Fraunhoffer diffraction at a single slit [derivation of intensity expression], double slit with theory. Transmission grating – theory and experiment [determination of wavelength of light]. Dispersion and resolution of grating.

Formation of shadows, eclipses, and image formation in pinhole camera can be understood only on the basis of rectilinear propagation of light. Wave theory fails to account for rectilinear propagation of light and failed to explain the formation of shadow. A wave can bend and should reach the region of geometrical shadow. Thus according to wave theory, the shadow region should not be sharp at the edges.

Careful experiments conducted by Grimaldi [of Italy] in 1665 have revealed that the geometrical shadow of an obstacle is not very sharp, but, the edge of the shadow is not very sharp and contains bright and dark interference fringes. This was earlier thought to be due to interference of direct light and the light reflected from the edge of the obstacle. If this were true, then the shadow region must have larger area than expected according to rectilinear propagation of light. On the contrary, the shadow region has smaller area indicating that, the light bends into the region of geometric shadow. A satisfactory explanation was given by Fresnel in 1818.

The bending of light at the edge of an obstacle and hence its encroachment into the region of geometrical shadow is known as "diffraction of light".

Classification of diffraction phenomena:

Diffraction phenomenon is broadly divided into two groups, based on the method adopted to study diffraction effects. They are (1) Fresnel class of diffraction and (2) Fraunhofer class of diffraction.

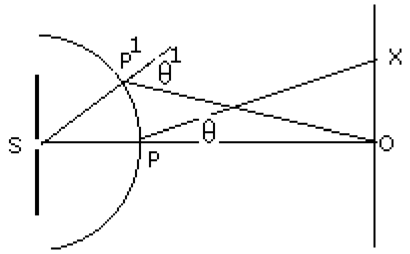
Sl. No.	Fresnel diffraction	Fraunhofer diffraction
1	Source and Screen are at finite distance from the obstacle/aperture.	Both the source and the Screen are at infinite distance from the obstacle/aperture.
2	Spherical/cylindrical wave front falls on the obstacle/aperture.	Plane wave front falls on the obstacle/ aperture.
3	No lenses are used in Fresnel diffraction	Converging lenses are used to study Fraunhofer diffraction.
4	Waves falling on the obstacle/aperture will not be in the same phase	Waves falling on the obstacle/aperture have the same phase.
5	Fresnel diffraction is general case of diffraction, which reduces to Fraunhofer case when the source and screen are at infinite distance from the obstacle/aperture.	Fraunhofer diffraction is a particular case of diffraction with source and screen at infinity from the obstacle/aperture.

Fresnel assumptions:

In order to explain diffraction phenomenon, Fresnel made the following assumptions.

(1) The entire wavefront can be divided into a large number of elements or zones of small area such that each of these elements acts as a source of secondary waves emitting waves in all directions.

(2) The effect at any point “O” will be the resultant of the secondary wavelets reaching “O” from various elements of the wavefront AB.
(3) The amplitude of disturbance at any point “X” due to the secondary waves from an element of the wavefront situated at “P” is a function of both the distance PX and the angle “θ” between the line PX and the normal to the wavefront PO at P. The variation of amplitude with the angle “θ“ is represented by an obliquity factor given by [1 + Cos θ]/2.



When θ = 0, the wave move along the normal to the wavefront along PO and hence the zone around P has maximum effect at O

When $\theta = 90^0$, cos θ = 0 and the effect in a direction perpendicular to the normal to the wavefront is one half of the effect along the direction of wave propagation.

When $\theta = 180^0$, Cos θ = - 1 and [1 + Cos θ]/2 = 0 and hence the wavefront has no effect along a direction opposite to the direction of wave normal.

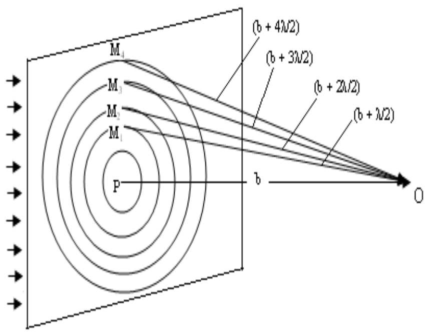
The failure of wave theory in explaining the absence of backward wave was accounted for by this explanation.

Rectilinear propagation of light OR diffraction:

Let us consider a plane wave ABCD propagating along X-axis towards a point “O” as shown in the following figure. Even though the waves emitted from different points of the wavefront with the same phase, while they reach “O”, they have phase difference due the following two reasons.

(1) The waves from different points travel different distances and
(2) Obliqueness of the wave reaching “O” from different points of the wavefront has different values.

To determine the resultant effect at “O” due to the entire wave front, the wavefront is imagined to be divided into a number of elementary zones as explained below.



Drop a perpendicular from O on the wavefront to meet it at P. Then P is known as the pole of the wavefront with respect to the point O. Let OP = b. If λ is the wavelength of the monochromatic light, then imaginary spheres are drawn with point O as centre and radii having values, OM_1= [b+λ/2], OM_2= [b+λ], OM_3 = [b+3λ/2] and so on, so that they cut the wavefront in to concentric circles of radii PM_1, PM_2, PM_3 and so on, as shown in figure.

The areas enclosed between these circles are known as **half period elements or half period zones**. The Fresnel zones are the annular rings in the wave front with the waves from the corresponding points of the adjacent areas differs in path by half wavelength [or reaches the points with a time delay of half time period]. The area of the first circle of radius PM_1 is called the first half period zone. The areas of the annular rings M_1M_2, M_2M_3, etc., are known as second, third, etc., half period zones.

The amplitude of the wavelet at the point O depends on the following factors:

(1) The amplitude at the point O is directly proportional to the area of the zone.
(2) The amplitude falls off inversely as the square of the distance of the zone from the point O.
(3) The amplitude varies with the obliquity of each zone [angle between the normal to the wavefront and the line joining the point O to the zone]. It is measured by the equation [1+Cosθ]/2. The amplitude decreases with the increase in obliquity. This decrease is extremely slow, but steady.

Area of n^th half period zone:

The radius of the first half-period zone, $PM_1 = \left[OM_1^2 - OP^2\right]^{\frac{1}{2}} = \left[\left(b+\frac{\lambda}{2}\right)^2 - b^2\right]^{1/2}$

$$PM_1 = \left[\left(b^2 + \frac{\lambda^2}{4} + b\lambda - b^2\right)\right] = \sqrt{b\lambda}$$ [Neglecting [$\lambda^2/4$] compared to other terms]

Similarly, the radius of the second half period zone is given by,

$$PM_2 = \sqrt{(b+\lambda)^2 - b^2} = \sqrt{b^2 + \lambda^2 + 2b\lambda - b^2} = \sqrt{2b\lambda}$$

More generally, the radius of (n – 1)^th zone can be shown to be $PM_{n-1} = \sqrt{(n-1)b\lambda}$ - - - - (1)

The radius of the n^th zone is given by, $PM_n = \sqrt{nb\lambda}$ - - - - (2)

The area of n^th zone = area of first n zones – area of first (n-1) zones

$= \pi[PM_n]^2 - \pi[PM_{n-1}]^2 = \pi\,[n\,b\lambda - (n-1)\,b\lambda] = \pi b\lambda$ - - - - - - (3)

Equation (3) shows that the area of n^th zone is independent of the zone number and hence, the area of each and every zone is a constant independent of the zone number.

The radii of the zones are proportional to $n^{1/2}$ even though the area of each zone is the same. Hence the width of the zone decreases from first to the outer zone. The contribution towards the amplitude of the wave at O has the same value for each and every zone.

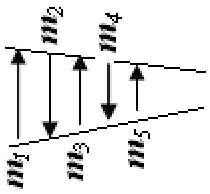
Distance and Obliquity factors:

Due to gradual increase in the distance of the zone and obliquity of the zones, the amplitude ot the wave received at O is of continuously decreasing values. If m_1, m_2, m_3, etc., are the amplitudes at O due to the first, second, third, etc., half-period zones respectively, then $m_1 > m_2 > m_3$, $m_{n-1} > m_n$, such that the amplitude of any zone may be taken as the arithmetic mean of the preceding and the succeeding zones. Thus to a close approximation we can write,

$$m_2 = \frac{m_1 + m_2}{2}, \quad m_4 = \frac{m_3 + m_5}{2}$$ And so on.

Relative phases of the

Since the distance of the corresponding points of successive zones differ by λ/2, the waves from these zones arrive out of phase. In other words, the waves from odd zones arrive in phase with one another and the waves from even zones arrive in phase with one another, but the waves from odd zones have a phase difference of π with those arriving from even zones.



Due to this, the resultant amplitude at O is written as

$A = m_1 - m_2 + m_3 - m_4 + m_5 - .... \pm m_n$ According as m_n is odd or even.

$$\text{Or, } A = \frac{m_1}{2} + \left[\frac{m_1}{2} - m_2 + \frac{m_3}{2}\right] + \left[\frac{m_3}{2} - m_4 + \frac{m_5}{2}\right] + \frac{m_5}{2} - - -$$

The expressions in the brackets are each zero and resultant amplitude at O thus reduces to the following value.

$$A = \frac{m_1}{2} + \frac{m_n}{2}, \text{ if } m_n \text{ is odd - - - - -(4)}$$

$$\text{and } \quad A = \frac{m_1}{2} + \frac{m_{n-1}}{2} - m_n, \text{ if } m_n \text{ is even - - - - (5)}$$

If n is quite large then the effect due to the $(n-1)^{th}$ or n^{th} zone is almost negligible on account of the distance and obliquity of the zone and hence the resultant amplitude due to the whole wave-front reduces to only $[m_1/2]$, that is , one-half of the effect produced by the first half period zone alone.

As the intensity I is proportional to the square of the amplitude, so I = $\frac{m_1^2}{4}$, that is the intensity at O due to the whole wave front is one-fourth of the intensity due to the first half-period zone only.

This shows that a small obstacle of the size of half the area of the first half period zone is enough to stop the effect of the entire wavefront. Since bλ is very small owing to small wavelength of the light waves, a small obstacle of the size of bλ/2 is sufficient to produce zero illumination at O and hence an obstacle of that size produces shadow of the obstacle at O and hence rectilinear propagation of light is explained using diffraction effect.

On the other hand, if the size of the obstacle is less than half the area of the first half-period zone, the intensity of light at O is not zero and hence the light bends and enters into the region of geometrical shadow of the obstacle and hence the point O is illuminated. This explains the diffraction of light observed at the edges of obstacles.

ZONE PLATE

A zone plate is an optical device, which works on the principle of Fresnel's zone. In Fresnel's zones, the effect at a point due to alternate zones cancels each other. If the effect of either even or odd zones is blocked at that point, then the net effect due to alternate zones having a path difference of λ has a maximum value. A plate, which allows only one set of alternate zones of the wavefront to pass through it, is called a zone plate.

The radius of the zones is proportional to $\sqrt{n}$, where n = 1, 2, 3, etc. This property is used in the construction of a zone plate. On a drawing sheet, concentric circles are drawn with the radii proportional to square root of natural numbers as shown in the following figure (1). The alternate zones are painted black. A photograph of the pattern is taken. We get on developed negative a reduced pattern. This negative forms the zone plate. In this plate alternate zones are transparent and allow light and the remaining alternate zones act as opaque region. There are two types of zone plates (1) positive zone plate and (2) negative zone plate as shown in figure (1a) & (1b).

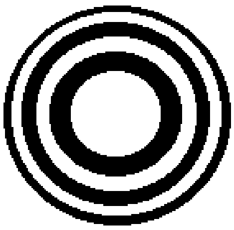


Figure (1a) Positive zone plate

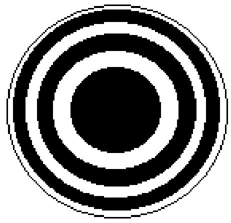


Figure (1b) Negative zone plate

(1) Positive zone plate: A zone plate in which odd zones are transparent and even zones are opaque is known as a positive zone plate.
(2) Negative zone plate: A zone plate in which even zones are transparent and odd zones are opaque is known as a negative zone plate.

THEORY OF ZONE PLATE:

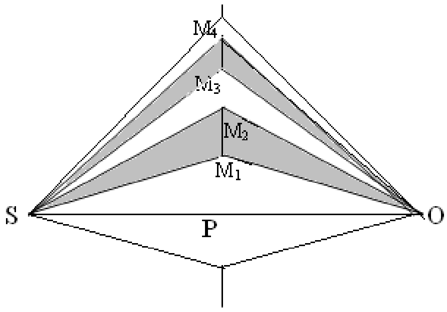


Figure (2)

Let AB be an imaginary transparent plate perpendicular to the plane of the paper as shown in figure (2). Let 'S' be a point source of monochromatic light giving out spherical wave of wavelength 'λ'. Draw SP perpendicular to the plane AB and produce it up to 'O'. Let us assume that we want to find out the resultant effect at the point "O" on the screen placed perpendicular to the plane of the paper. With 'P' as centre, imagine the plate AB to be divided into circular half period zones with radii equal to PM_1, PM_2, PM_3, etc., such that:

$$SM_1 + M_1O = SP + PO + \lambda/2$$
$$SM_2 + M_2O = SP + PO + 2[\lambda/2]$$
$$- \quad - - - - - - - -$$
$$SM_n + M_nO = SP + PO + n\,[\lambda/2]$$

To determine the radius r_n of the n^{th} circle, consider the relation,

$$SM_n + M_nO = SP + PO + n\,[\lambda/2]$$

Let SP = a, PO = b, then we can write,

$$SM_n = \sqrt{SP^2 + PM^2} = \sqrt{a^2 + r_n^2} = a\left[1 + \frac{r_n^2}{a^2}\right]^{1/2} = a + \frac{r_n^2}{2a} \text{ [approximately] - - - (1)}$$

['r_n' is very small compared to 'a' and hence by applying binomial theorem and neglecting the square and higher order term, we get the above equation.]

Similarly, $M_nO = \sqrt{PO^2 + PM_n^2} = \sqrt{b^2 + r_n^2} = b + \frac{r_n^2}{2b}$ - - - -(2)

Substituting for SM_n and M_nO, we get,

$$a + \frac{r_n^2}{2a} + b + \frac{r_n^2}{2b} = a + b + \frac{n\lambda}{2}$$

$$\therefore r_n^2 \left[\frac{1}{a} + \frac{1}{b}\right] = n\lambda$$

$$r_n^2 = n\lambda\left[\frac{ab}{a+b}\right] \text{ - - - -(3)}$$

Since a, b and λ are constants, $r_n \propto \sqrt{n}$ - - - - -(4)

It follows that the external radii of the various zones for given values of a and b are proportional to the square root of natural numbers.

Area of n[th] zone = $\pi\left[r_n^2 - r_{n-1}^2\right] = \pi\left[\frac{n\lambda ab}{a+b} - \frac{(n-1)\lambda ab}{a+b}\right] = \frac{\pi\lambda ab}{a+b}\left[n - n + 1\right] = \frac{\pi\lambda ab}{a+b}$ - - - - -(5)

Area of n[th] zone is independent of the order 'n'. This means that for the given positions of the point source 'S' and its image 'O', the area of all the zones remains practically constant. But, as the amplitude at 'O' m_1, m_2, m_3, etc., due to various zones decrease very slightly with the order of the zones and are alternately in opposite phase, the resultant amplitude is given by,

$$A = m_1 - m_2 + m_3 - m_4 + \text{ - - - - } = \frac{m_1}{2} \text{ [when n is very large] - - - - -(6)}$$

If monochromatic light from 'S' is allowed to fall on this zone plate and the emergent light is received on a screen, then for a certain value of 'b', the hypothetical half-period elements of the incident wavefront may coincide with the actual elemental areas of the plate. In such a case if even numbered elements are made opaque, then the resultant amplitude is given by,

$$A = m_1 + m_3 + m_5 + \text{ - - - - (7)}$$

This value is many times greater than that due to all the zones. This means that a zone plate has the focusing action. Under these conditions, 'O' can be said to be the image of 'S'.

Equation (4) can be arranged in the following form.

$$\frac{1}{a} + \frac{1}{b} = \frac{n\lambda}{r_n^2} \text{ - - - - (8)}$$

Comparing equation (8) with the lens formula, $\frac{1}{v} + \frac{1}{u} = \frac{1}{f}$, - - - -(9)

$$\text{We have, } \frac{1}{f} = \frac{n\lambda}{r_n^2} \text{ or } f = \frac{r_n^2}{n\lambda} \text{ - - -- - (10)}$$

Equation (10) gives the value of primary or first order focal length of the zone plate.

MULTIPLE FOCI OF A ZONE PLATE:

When $a = \infty$, the focal length of the zone plate is equal to 'b', and hence $f = \frac{r_n^2}{n\lambda}$ or $f_1 = \frac{r_1^2}{\lambda}$ - - (11)

For this focal length each actual zone of the plate contains one half-period element and the focal point O_1 is called the first order focal point, the intensity of the image is maximum being proportional to $\left(A_1^2\right)$. [Evident from equation (7)]

If each zone contains three half-period elements with respect to some other point 'O_2', nearer to the plate, then the resultant intensity at O_2 is given by,

$A_2 = (m_1 - m_2 + m_3) + (m_7 - m_8 + m_9) + (m_{13} - m_{14} + m_{15}) + - - - -$

$$= \left(\frac{m_1 + m_3}{2}\right) + \left(\frac{m_7 + m_9}{2}\right) + \left(\frac{m_{13} + m_{15}}{2}\right) + - - - - = \frac{1}{2}(m_1 + m_3 + m_7 + m_9) < A_1.$$

The rays transmitted through each successive transparent zone have a mean path difference of 3λ. Hence, O_2 is known as the third order focal point.

As the first transparent zone contains three half-period elements all of equal area, we can write,

$$3\pi x^2 = \pi r_1^2 \qquad \text{or} \qquad x^2 = \frac{r_1^2}{3}$$

Hence, the second focal length of the zone plate is given by,

$$f_2 = \frac{x^2}{\lambda} = \frac{r_1^2}{3\lambda} = \frac{f_1}{3} \quad - - - - - (12)$$

Similarly, we can show that the third focal length of the zone plate is given by,

$$f_3 = \frac{f_1}{5} \quad - - - - (13)$$

In general, the focal length of various foci of the zone plate can be given by the formula,

$$f = \frac{r_1^2}{(2p+1)\lambda} \quad - -- \; -- -(14)$$

Where, p = 0, 1, 2, 3, etc.

Comparison between a zone plate and a convex lens:

Zone plate	Convex lens
(1) Zone plate works on the phenomenon of diffraction	(1) Convex lens works on the phenomenon of refraction of light.
(2) Odd zone transparent plate produces real image on the opposite side of the plate.	(2) Convex lens produces real image of the object on the other side of the lens
(3) It shows chromatic aberration as the focal length depends on the wavelength λ	(3) It shows chromatic aberration as the focal length depends on the wavelength λ
(4) For a given wavelength λ, zone plate has different focal lengths given by equation $f = \frac{r_1^2}{(2p+1)\lambda}$	(4) A convex lens has a single focal length for a given wavelength given by the formula, $\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$
(5) Intensity of the image decreases with the order of focal length	(5) Intensity does not depend on the focal length.
(6) In a zone plate, the red focus is nearer the plate than the violet focus, hence, $f_r < f_v$	(6) In a lens the violet focus is nearer than the red focus, hence, $f_v < f_r$

FRESNEL DIFFRACTION OF A CYLINDRICAL WAVE-FRONT AT A STRAIGHT EDGE:

Let a cylindrical wavefront from a rectangular slit 'S', illuminated by a monochromatic source of wavelength 'λ', falls on a sharp and straight edge of an opaque obstacle AB [such as blade] such that the edge of the blade is parallel to the silt. Join SA and produce it to meet the screen CD placed perpendicular to the plane of the paper at 'O'.

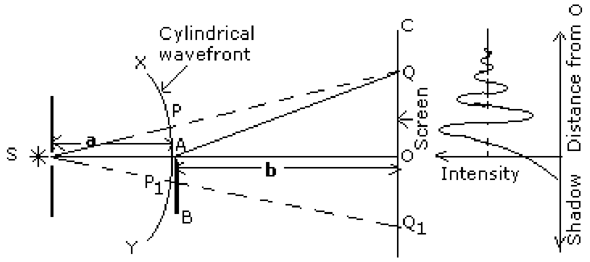


Figure (1): Diffraction at a straight edge

According to geometric optics, the region below 'O' should have shadow of the obstacle and the region above 'O' should be illuminated.

But, experiments show that the intensity falls off continuously and rapidly as we move deep into the region below 'O' and becomes totally dark after a short distance.

Just above 'O', alternate dark and bright bands of unequal width and varying intensity parallel to the straight edge are observed. The bands become closer and less distinct as we move away from 'O' until the screen is uniformly illuminated.

Explanation based on Fresnel theory:

(a) Intensity variation inside the geometrical shadow: With reference to the point 'O' on the screen the point A acts as the pole of the wavefront. The light reaches only from the wavefront AX above the pole and the part AY of the wavefront is blocked by the obstacle AB. If the wavefront above A is divided in to several half period strips of rectangular area, then the net amplitude of the entire wavefront at O is given by the equation,

$$A = m_1 - m_2 + m_3 - m_4 + \ldots = \frac{m_1}{2} \quad \text{- - - (1)}$$

Hence the intensity at "O" is proportional to $\frac{m_1^2}{4}$.

As we move to a point below O, into the geometrical shadow, say to a point Q_1, the pole also moves from A to P_1 and gradually the first, the first two, the first three, etc., half period elements of the upper half of the wave are also successively cut off by the obstacle. If the first half period element is only obstructed, then the resultant amplitude becomes,

$$A = -m_2 + m_3 - m_4 + \ldots = -\frac{m_2}{2} \quad \text{- -- - (2)}$$

Similarly, if two or three half period elements are obstructed, the amplitude at the corresponding point inside the shadow will become $\frac{m_3}{2}, \frac{m_4}{2}$ and so on. As the intensity is proportional to the square of the maximum displacement, the magnitude of intensity of the points below 'O' decreases rapidly and become zero finally.

(b) Formation of bands outside the geometrical shadow:

Consider a point Q in the illuminated region on the screen. With respect to the point Q, P is the pole of the wave front. For any point Q in the illuminated region, the effect due to the wave front above the pole P has the same value and it results in the general illumination. Depending on the number of half period elements between the pole P and the edge A of the obstacle, the intensity at Q will be either maximum or minimum. If, with respect to Q, only one half period strip is exposed between A and P, then the resultant amplitude at Q will have a maximum. In general if the region PA contains odd number of zones, then the net effect at Q will be maximum and if PA contains even number of zones then the net effect at Q will be minimum.

To find the positions of maximum and minimum and their widths let us consider the effect at a point Q outside the geometrical shadow. Join SQ cutting the wave front at P. Also join AQ. The number of half period elements contained in PA is equal to the number of half wavelengths in the path difference AQ – PQ.

Taking SA = a and AO = b, we have,

$$AQ = \sqrt{AO^2 + OQ^2} = \sqrt{b^2 + y^2} = b\left(1 + \frac{y^2}{b^2}\right)^{1/2} = b + \frac{y^2}{2b^2} \text{ - - -(3)}$$

[Using binomial expansion and neglecting square and higher order terms.
Also,

$$PQ = SQ - SP = \sqrt{(a+b)^2 + y^2} - a = (a+b)\left(1 + \frac{y^2}{(a+b)^2}\right)^{1/2} - a$$

$$PQ = (a+b)\left(1 + \frac{y^2}{2(a+b)^2}\right) - a = a + b + \frac{y^2}{2(a+b)} - a = b + \frac{y^2}{2(a+b)} \text{ -- - - (4)}$$

Hence the path difference is given by,

$$AQ - PQ = \left(b + \frac{y^2}{2b}\right) - \left(b + \frac{y^2}{2(a+b)}\right) = \frac{y^2}{2}\left(\frac{1}{b} - \frac{1}{a+b}\right) = \frac{ay^2}{2b(a+b)} \text{ -- - - (5)}$$

For maximum intensity at Q the path difference given by equation (5) must satisfy the following condition.

$$\frac{ay^2}{2b(a+b)} = (2n+1)\frac{\lambda}{2}$$

$$\text{Or, } y = \sqrt{\frac{b(a+b)(2n+1)\lambda}{a}} \text{ - - - - - (6)}$$

For minimum brightness at Q the path difference given by equation (5) must satisfy the following condition.

$$\frac{ay^2}{2b(a+b)} = 2n\frac{\lambda}{2},$$

$$\text{Or, } y = \sqrt{\frac{b(a+b)2n\lambda}{a}} \text{ - - - - (7)}$$

Where, n is whole number, 0, 1, 2, 3, etc.

Since ‘y’ is proportional to $\sqrt{2n+1}$ for maximum and $\sqrt{2n}$ for minimum, the width of the bands goes on decreasing with the increase in the order of the number of the band.

The diffraction pattern is as represented in the graph on the right side of figure (1). In the geometrical shadow the intensity falls off rapidly until with in a short distance the region is totally dark. Outside the geometrical shadow there is a system of dark and bright bands till uniform illumination results.

Note: In the case of white light, a few coloured bands running parallel to the edge of the shadow will be observed in the illuminated region above O. This is due to the fact that white light contains several wavelengths and the condition for maximum and minimum is satisfied at different positions above "O" for different wavelengths.

Fresnel diffraction at a narrow obstacle (cylindrical wire):

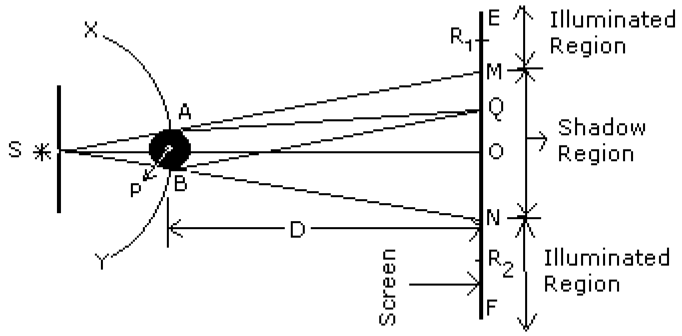


Figure (1) Fresnel diffraction at a wire

Let AB be a narrow cylindrical obstacle such as a wire of thickness 'd' held parallel to a narrow rectangular slit 'S', such that they are parallel to one another and perpendicular to the plane of the paper. When the slit is illuminated by a monochromatic source of light [of wavelength-λ] it gives rise to a cylindrical wavefront XY falling on the wire as shown in figure (1). Due to this a shadow of the wire is formed between MN on the screen and the region above M or the region below N is illuminated.

On the screen, in the region above M and below N diffraction bands of decreasing intensity and decreasing thickness are formed as in the case of a straight edge. Where as in the shadow region MN, the intensity of illumination depends on the thickness of the wire as discussed in the following two cases.

(1) Thin wire:

In the case of a thin wire, a few half period elements of the upper half and lower half of the wave front are obstructed. At any point R_1 above M [or a point R_2 below N], the obstacle produces the same effect as a straight edge. Hence, the illuminated region has diffraction bands formed. When path difference between the rays from the pole of the wave front with respect to the point R_1 and the end A of the obstacle is an odd multiple of λ, the intensity of the band has a maximum value. On the other hand, when the path difference is an integer value, then the intensity of the band has a minimum value. Similarly, the region below N also has diffraction bands of varying intensity and thickness due to the reasons mentioned above.

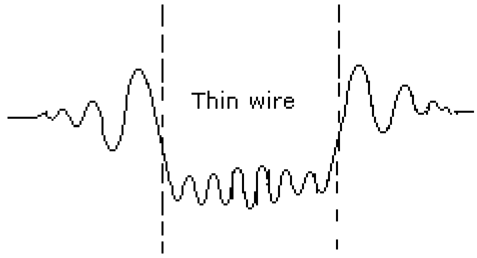


Figure (2): Intensity distribution in the diffraction pattern on a thin wire.

In the geometrical shadow region MN, the intensity at the centre of the shadow has a maximum value. This is due to the reason that waves from the first, second, third etc. exposed half period zones reach O with the same path traveled and hence they produce maximum at the point O. At a point Q in the shadow region, the number of half period elements obstructed in the upper and lower half of the wave front has different number. As a result, depending on the path difference [BQ –AQ] the intensity at Q will have a maximum or minimum.

(i) A bright band is formed at Q if [BQ –AQ] = nλ and

(ii) A dark band is formed at Q if [BQ –AQ] = (2n+1)λ/2

If 'd' is the thickness of the wire and 'D' is the distance of the point O from P [midpoint of wire], then OQ = y for the n[th] bright maximum at Q can be shown to be $\boxed{y = \frac{D}{d} n\lambda}$

Also the fringe width β of the interference fringes can be shown to be $\beta = \frac{D}{d}\lambda$

Theory of interference fringes is same as the interference fringes in the case of two coherent sources. The upper half of the wave front above the wire acting as one source and the lower half of the wave front below the wire acting as another source and these two sources behave as coherent sources with respect to any point on the screen in the geometrical shadow.

(2) Thick wire:

When the diameter of the wire is increased, the fringe width inside the geometrical shadow decreases and when the wire is sufficiently thick, the fringes inside the shadow vanishes and the intensity decreases from the edge of the shadow as we move deep into the shadow region as shown in figure (3). The intensity variation of diffraction bands in the illuminated region is same as in the case of thin wire as shown in figures (2) and (3).

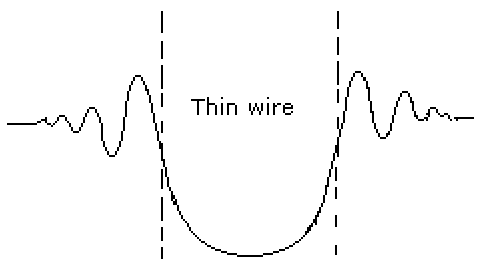
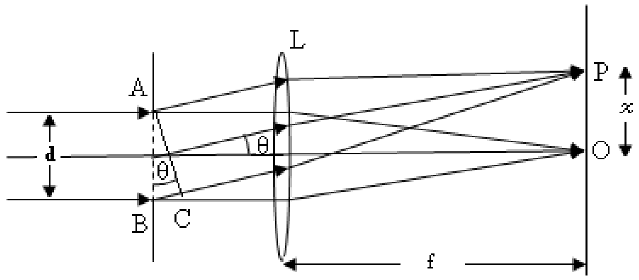


Figure (3) case of thick wire

FRAUNHOFER DIFFRACTION AT A SINGLE SLIT



A plane wave front incident normally on a narrow slit AB of width ‘a’ is diffracted as shown in figure (1). A converging lens L placed in the path of the diffracted beam forms a real image of the diffraction pattern on the screen in the focal plane of the lens. All the waves moving in the incident direction have the same initial phase and are brought to focus at a point O on the screen and as all the waves arrive in phase, the point O on the screen has maximum brightness and is known as central bright band.

The rays diffracted at an angle ‘θ’ have been brought to focus at P by the lens L. At this point the waves from different parts of the slit AB do not arrive in phase. From A drop a perpendicular to the diffracted ray from A to meet it at E. Then AC represents the diffracted wave front and BC = d sin θ represents the path difference between the rays from two edges A and B of the slit. If this path difference is λ, then the aperture AB can be divided in to two equal parts, so that, the wavelets from the corresponding points in each half will differ in path by λ/2 when they reach P and hence they would mutually interfere and cancel out each other’s effect thus producing the first minimum at P.

Thus, the condition for first minimum is d Sin θ = λ

Or $Sin\theta = \frac{\lambda}{d} \Rightarrow \theta = \frac{\lambda}{d}$ - - - - - (1) [since θ is very small]

If the path difference, d Sin θ = 2λ between the rays reaching P from the points A and B, then imagining the slit to be divided into four equal parts, the rays from corresponding points separated by a distance [d/4] in the adjacent part of the slit have a path difference of [λ/2] and will mutually cancel each other’s effect in pair and hence a second minimum is formed at the point P. Thus the condition for second minimum is BC = d Sin θ = 2λ - - - - (2)

In general, n^{th} minimum is formed at a point on the screen when the waves reaching the point on the screen from the extreme end of the slit have a path difference BC = d Sin θ = nλ - - - - (3)

The minima are symmetrically formed on the other side of central maximum as shown in figure (2).

Also, secondary maxima are formed in between the minima on the two sides of the central maximum at positions for which the path difference BC = d Sin θ = [2n+1] (λ/2) - - - - - (4)

The intensity of these secondary maxima is much less and falls off rapidly as we move outwards. The integer n = 1 corresponds to first maximum after central maximum and n = 2 corresponds to second maximum after the central maximum and so on.

The following figure (2) shows the variation of intensity with the distance from the central maximum. The intensity of illumination of the first and the second secondary maxima on either side of the central maximum O are [1/22] and [1/61] of the intensity of the central maximum respectively. Only a few secondary minima are visible as they decrease in intensity rapidly with the increase in angle of diffraction. If the slit is wide, "d" is large and hence θ is small and hence the central maximum is very narrow and the image is sharp. On the other hand, if the slit is narrow, θ is quite large and the central maximum is broad and flat. The image is no longer sharp and the first minima lie far apart.

Intensity distribution in the diffraction pattern of a slit:

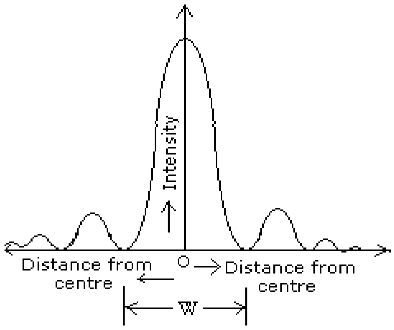


Figure (2)

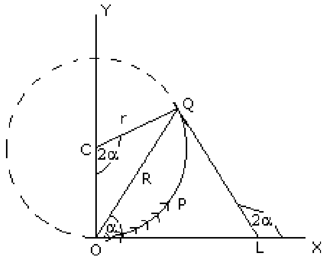


Figure (3)

Intensity variation with the distance from the central maximum in Fraunhofer diffraction due to a single slit is as shown in figure (2) and the resultant amplitude and hence the intensity at any point on the screen can be calculated as follows.

The path difference between the waves reaching the point P on the screen from the extreme ends A and B of the slit is given by,

BC = d Sin θ

The corresponding phase difference "2α" between the waves while they reach P_1 is given by,

$$2\alpha = \left[\frac{2\pi}{\lambda}\right] d\, Sin\theta \quad - - - - (5)$$

If the entire aperture of the slit is divided into a large number of strips [n] of equal width [d/n] parallel to its length, then each strip acts as an elementary source. The waves from different strips reaching P have the same amplitude, but have their phases gradually increasing for waves from B to A. This is due to increase in phase of the wave that reach the point P for the waves from the strips B to A. Thus, we have to find the resultant amplitude of a number of waves of equal amplitude of regularly increasing phase. In this case, the phases can be added according to vector addition to get the resultant amplitude as in figure (3).

In figure (3), the length of the chord R = OQ gives the resultant amplitude & hence

R = 2 r Sin α

Where "r" is the radius of the circumscribing circle and "2α" is the phase difference between the initial phase of the first and the last elementary waves and it is equal to $\angle QLX$

If "S" is the length of the arc OPQ, then S = n a = A_0 [say]. But, S = 2 α r = A_0 [from figure]

Where, 'a' is the amplitude of waves from individual strips.

$$\therefore r = \frac{A_0}{2\alpha} \quad - - - (6)$$

Hence, the chord length $R = 2r Sin\alpha = 2\left(\frac{A_0}{2\alpha}\right) Sin\alpha = A_0 \frac{Sin\alpha}{\alpha} \quad - - - - (7)$

Since OQ represent the resultant amplitude, the intensity of the wave is given by,

$$I = A_0^2 \frac{Sin^2\alpha}{\alpha^2} \quad - - - - - (8)$$

Where 'α' is given by the equation, $\alpha = \frac{\pi}{\lambda} d\, Sin\theta$ - - - - (9) [follows from equation (5)]

Intensity of central maximum:

In this case θ = 0 [waves move in incident direction], and hence α = 0 and hence from equation (9) we can write, $I_0 = \lim_{\alpha \to 0} A_0^2 \frac{Sin^2\alpha}{\alpha^2} = A_0^2 \left[\lim_{\alpha \to 0} \frac{Sin\alpha}{\alpha} \right]^2 = A_0^2$ - - - - (10)

Thus intensity at "O" has a maximum value as all the waves travel and reach O with the same amplitude and same phase.

Directions of minimum intensity:

According to equation (8), the intensity I has minimum value when α = ± n π

$$\Rightarrow \frac{\pi}{\lambda} d\, Sin\theta = \pm n\pi$$

Or, d Sin θ = ± n λ

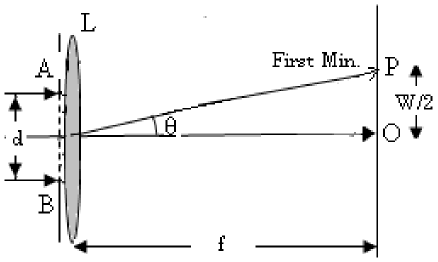
For the first minimum adjacent to central maximum, n = ± 1, therefore, $Sin\theta = \pm \frac{\lambda}{d}$

Other minima occur along directions obtained by giving n the integer values ± 2, ±3, etc. In other words, the direction "θ" of n^{th} minimum is given by the equation,

$$Sin\theta = \pm n \frac{\lambda}{d} \text{ - - - - (11)}$$

Note: When "d = λ", Sin θ = 1 and hence θ = 90 ^{0}, The light spreads out in all directions and the intensity decreases from the centre outwards giving us a broad central maximum, but no other maximum is obtained.

Width of Central maximum:



If the lens is placed close to the slit, then the distance between the slit and the screen is 'f'.

$$\sin\theta = \frac{W/2}{f} = \pm \frac{\lambda}{d} \Rightarrow W = \frac{2f\lambda}{d} \text{ - - - (12)}$$

Directions of secondary maxima:

In equation (8), the position of secondary minima after central maximum is given by the condition $\alpha = \frac{\pi}{\lambda} d\, Sin\theta = (2n+1)\frac{\pi}{2} \Rightarrow d\, Sin\theta = (2n+1)\frac{\lambda}{2}$; where n = 1, 2, 3, etc.

The value n = 1 gives the position of 1^{st} secondary maximum after the central maximum and for this maximum the value of α = [3π/2] and the corresponding intensity, according to equation (8) is given by,

$$I_1 = A_0^2 \frac{Sin^2\left(3\pi/2\right)}{\left(3\pi/2\right)^2} = \left(\frac{2}{3\pi}\right)^2 A_0^2 = \left[\frac{1}{22}\right] A_0^2 = \frac{I_0}{22} \text{ - - - - (13)}$$

The value n = 2 gives the position of 2^{nd} secondary maximum after the central maximum and for this maximum the value of α = [5π/2] and the corresponding intensity, according to equation (8) is given by

$$I_2 = \left[\frac{2}{5\pi}\right]^2 A_0^2 = \left[\frac{1}{66}\right] A_0^2 = \frac{I_0}{66} \text{ - - - - - (14)}$$

Similarly, n = 3 gives the position of III[rd] secondary maximum after the central maximum and for this maximum the value of $\alpha = [7\pi/2]$ and the corresponding intensity, according to equation (8) is given by,

$$I_3 = \left[\frac{2}{7\pi}\right]^2 A_0^2 = \left[\frac{1}{121}\right] A_0^2 = \frac{I_0}{121} \quad ----(15)$$

The intensity variation on either side of the central maximum is as shown in figure (2).

FRAUNHOFER DIFFRACTION AT A SINGLE SLIT [CACLULUS METHOD]

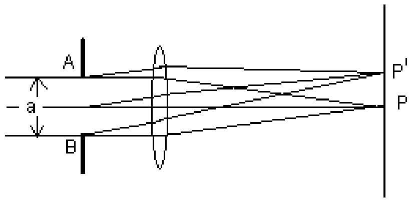
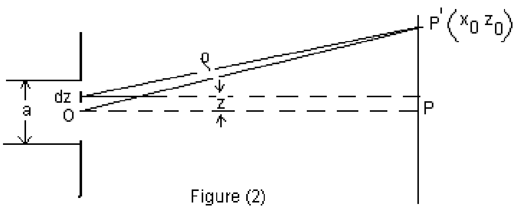


Figure (1) Diffraction at a single slit.

Let a monochromatic parallel beam of light be incident on the slit AB of width ‘a’. The secondary waves traveling in the direction of incident light converge at the point P. the secondary waves traveling along direction inclined to incident direction at an angle “θ” converge at the point P^1.



Let the centre of the slit be taken as the origin of the co-ordinate system and let x_0 be the distance of the screen from the slit of width ‘a’. If we assume that the entire slit is divided in to several small element of width dz. Let us consider an element dz with co-ordinates (0, z). Let the co-ordinates of a point P^1 be (x_0 z_0). Let ‘ρ’ is the distance of the point ‘P^1’ from the element.

The displacement at the point P^1 due to the element ‘dz’ at any instant is given by,

$$dy = K\ dz\ Sin\ 2\pi\left(\frac{t}{T} - \frac{\rho}{\lambda}\right) \quad ----(1)$$

The resultant displacement at P^1 due to the whole wave front is given by,

$$y = K \int_{-a/2}^{a/2} Sin\ 2\pi\left(\frac{t}{T} - \frac{\rho}{\lambda}\right) dz \quad ----(2)$$

Now, $\rho^2 = x_0^2 + (z_0 - z)^2 \quad \& \quad r^2 = x_0^2 + z_0^2 \Rightarrow x_0^2 = r^2 - z_0^2$

Therefore, $\rho^2 = r^2 - z_0^2 + (z_0 - z)^2 = r^2 - z_0^2 + z_0^2 - 2z z_0 + z^2$

$$\therefore \quad \rho^2 = r^2\left[1 - \frac{2z z_0}{r^2} - \frac{z^2}{r^2}\right]$$

In Fraunhofer diffraction, the screen is at a very large distance from the slit, and r > > z & hence $\frac{z^2}{r^2}$ is negligible.

$$\rho^2 = r^2\left[1 - \frac{2z z_0}{r^2}\right] \quad \therefore \quad \rho = r\left[1 - \frac{2z z_0}{r^2}\right]^{\frac{1}{2}}$$

$$\therefore \quad \rho = r\left[1 - \frac{2z z_0}{2r^2}\right] = r\left[1 - \frac{z z_0}{r^2}\right] = r - \frac{z z_0}{r} = r - z\ Sin\ \theta \quad ---(3)$$

Substituting the value of ‘ρ’ in equation (2), we get,

$$y = K\int_{-a/2}^{a/2} Sin\ 2\pi\left(\frac{t}{T}-\frac{r-z\ Sin\ \theta}{\lambda}\right)dz \quad y = -K\left[\frac{Cos\ 2\pi\left(\frac{t}{T}-\frac{r-z\ Sin\ \theta}{\lambda}\right)}{\frac{2\pi\ Sin\ \theta}{\lambda}}\right]_{-a/2}^{a/2} \text{ - - - - (4)}$$

Substituting the limits of integration, we get,

$$y = \frac{-K\lambda}{2\pi\ Sin\ \theta}\left[Cos\ 2\pi\left(\frac{t}{T}-\frac{r}{\lambda}+\frac{a\ Sin\ \theta}{\lambda}\right)-Cos\ 2\pi\left(\frac{t}{T}-\frac{r}{\lambda}-\frac{a\ Sin\ \theta}{\lambda}\right)\right]$$

$$y = \frac{-K\lambda}{2\pi\ Sin\ \theta}\left[-2\ Sin\ 2\pi\left(\frac{t}{T}-\frac{r}{\lambda}\right)\times Sin\left(\frac{\pi\ a\ Sin\ \theta}{\lambda}\right)\right]$$

$$y = \frac{K\lambda}{\pi\ Sin\ \theta}\left[Sin\ 2\pi\left(\frac{t}{T}-\frac{r}{\lambda}\right)\times Sin\ \alpha\right] \text{ - - - - (5)}$$

Where, $\alpha = \left(\frac{\pi}{\lambda}\right) a\ Sin\ \theta$ - - - - (6)

Equation (5) may be written as,

$$y = \frac{Ka\lambda}{\pi\ aSin\ \theta}\left[Sin\ 2\pi\left(\frac{t}{T}-\frac{r}{\lambda}\right)\times Sin\ \alpha\right] = Ka\frac{Sin\ \alpha}{\alpha}\times Sin\ 2\pi\left(\frac{t}{T}-\frac{r}{\lambda}\right)$$

$$\Rightarrow \quad y = y_0\ Sin\ 2\pi\left(\frac{t}{T}-\frac{r}{\lambda}\right) \text{ - - - (7)}$$

In equation (7), $y_0 = Ka\frac{Sin\ \alpha}{\alpha}$ is the amplitude of the resultant vibration at the point P[1]. The corresponding intensity is given by,

$$I = y_0^2 = K^2a^2\left(\frac{Sin^2\ \alpha}{\alpha^2}\right) \text{ - - - - (8)}$$

Intensity at the point P[1] is determined b the term $\left(\frac{Sin^2\ \alpha}{\alpha^2}\right)$

Intensity at P: [Central maximum]

Along the direction of P, the value of θ = 0 and hence α = 0, But in the limit α → 0 $\frac{Sin\ \alpha}{\alpha}$ → 1 & hence, $I_0 = K^2a^2$ - - - (9)

Intensity of secondary maxima:

At P secondary maxima are observed when $\alpha = (2n+1)\frac{\pi}{2}$

For first secondary minima, n = 1& hence $\alpha = \left(\frac{3\pi}{2}\right)$ or the corresponding path difference *a Sin θ = (2n+1) λ/2.*

The intensity of first secondary maximum is given by,

$$I_1 = y_0^2 = K^2a^2\left(\frac{Sin^2\ 3\pi/2}{(3\pi/2)^2}\right) = I_0\frac{4}{9\pi^2} = \frac{I_0}{22} \text{ - - -}$$

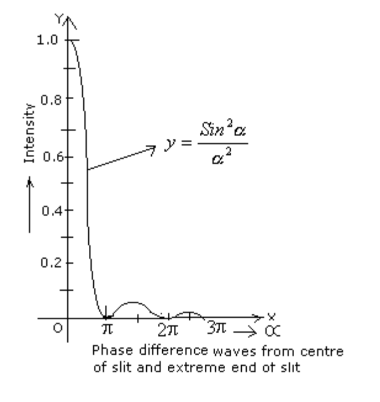


Figure (3)

Similarly, for second secondary maximum, *n=2* & $\alpha = \left(\frac{5\pi}{2}\right)$

Therefore, the intensity is given by,

$$I_2 = y_0^2 = K^2 a^2 \left(\frac{Sin^2 \, 5\pi/2}{\left(5\pi/2\right)^2} \right) = I_0 \frac{4}{25\pi^2} = \frac{I_0}{66} \text{ - - - - (11)}$$

Similarly, the intensity of third secondary maximum is given by,

$$I_3 = y_0^2 = K^2 a^2 \left(\frac{Sin^2 \, 7\pi/2}{\left(7\pi/2\right)^2} \right) = I_0 \frac{4}{49\pi^2} = \frac{I_0}{121} \text{ - - - - (12)}$$

FRAUNHOFFER DIFFRACTION AT DOUBLE SLIT

Fraunhofer diffraction at a slit is as shown in figure (1). AB and CD are two rectangular parallel slits arranged perpendicular to the plane of the paper. 'a' is the width of the slit and 'b' is the width of the opaque portion between slits. The light passing through the double slit is focused on the screen MN using a convex lens. Let P is a point on the screen such that OP is perpendicular to the screen. All the secondary waves moving in the direction of OP come to focus at P and P corresponds to the position of central maximum.

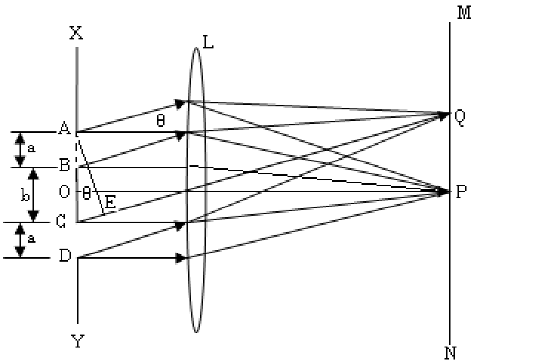


Figure (1)

In this case the diffraction pattern consists of two parts.

(i) Interference pattern due to secondary waves from the corresponding points of the two slits and

(ii) The diffraction pattern due to the secondary waves from the two slits individually.

For calculating the positions of interference maxima and minima the angle of diffraction is denoted as 'θ' and for the diffraction maxima and minima it is denoted as φ. Both θ & φ denote the angle between the direction of secondary waves and the initial direction of incident light

(i) Interference maxima and minima

Let us consider the secondar angle ϑ with the initial direction.

In the triangle CAE, $\sin\theta = \frac{CE}{AC} = \frac{CE}{a+b}$

Therefore, the path difference between the waves $CE = (a+b)\sin\theta$ - - - (1)

If this path difference is equal to an odd multiple of λ/2, θ gives the direction of minima due to interference of secondary waves from the corresponding points of the two slits.

Thus for minimum intensity, $(a+b)\sin\theta = (2n+1)\frac{\lambda}{2}$ - - - (2)

Taking n = 1,2,3, etc., the values of θ_1, θ_2, θ_3 etc., corresponding to the direction of minima are obtained. In general the direction of n^{th} minimum is given by, $\sin\theta_n = \frac{(2n+1)\lambda}{2(a+b)}$ - - - (3)

On the other hand, if this path difference is equal to an even multiple of λ/2, θ gives the direction of maxima due to interference of secondary waves from the corresponding points of the two slits.

Thus, for maximum intensity, $(a+b)\sin\theta = 2n\left(\frac{\lambda}{2}\right) = n\lambda$ - - - (4)

Taking n = 1,2,3, etc., the values of θ_1, θ_2, θ_3 etc., corresponding to the direction of maxima are obtained. In general the direction of nth minimum is given by, $\sin\theta_n = \frac{n\lambda}{(a+b)}$ - - - (5)

The angular separation between nth & (n-1)th maxima is given by,

$$\sin\theta_n - \sin\theta_{n-1} = \frac{n\lambda}{(a+b)} - \frac{(n-1)\lambda}{(a+b)} = \frac{\lambda}{(a+b)} \text{ - - (6)}$$

This gives the value of fringe width of minimum and is independent of order of maxima. Similarly, we can show that the width of maximum intensity is also given by equation (6) and is independent of fringe number.

(ii) Diffraction maxima and minima:

Let us consider the secondary waves traveling in a direction inclined at an angle φ to the initial direction of incident light. If the path difference between the waves from the extreme ends of the slits A and B in reaching P is λ, the φ gives the direction of minimum. In this case the wave front AB passing through the slit is divided in to two parts. The path differences between the corresponding points of the two parts cancel each other's effect resulting in minimum intensity at that point. If 'a' is the width of the slit, the path difference between the waves from the extreme ends of the slit AB in reaching P is "a Sin θ". The condition of nth minimum is given by $a\sin\phi_n = n\lambda$ - - - (7)

Taking n = 1, 2, 3, etc., we get the direction θ_1, θ_2, θ_3, of different minima corresponding to each slit.

Similarly, the condition for nth diffraction maximum due to each slit is given by, $a\sin\phi_n = (2n+1)\frac{\lambda}{2}$ - - - (8)

In the combined effect of interference due two slits acting as two coherent sources and individual slits showing diffraction phenomenon individually, the net effect is minimum when the condition for minimum is satisfied for both phenomenon, where as maxima occurs when the condition for maximum is satisfied for interference phenomenon only. In other words, interference maxima and minima are seen with in the diffraction pattern due to the two slits as shown in the following figure (2).

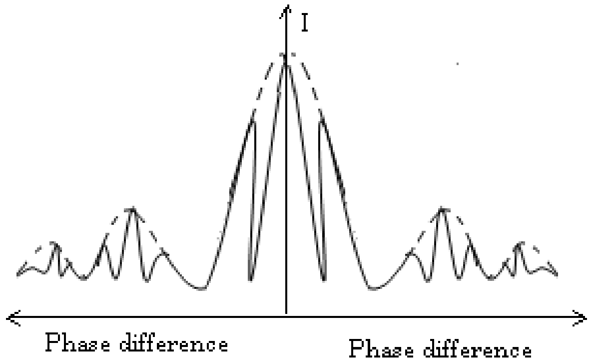


Figure (2)

Case 1: If a = b, then (a + b) = 2a and the condition for nth interference maxima is

$2a\sin\theta = n\lambda$ − − − (1)

The condition for nth diffraction minimum is given by,

$$a\sin\varphi_n = n\lambda \text{ −−−−(2)}$$

From equations (9) and (10) it follows that the condition for second interference maximum coincides with the condition for first diffraction minimum and hence the second interference maximum is absent in the pattern. In between central maximum and the first diffraction minimum, an interference maximum occurs. Therefore, in this case, the cental maximum contains three maxima as shown in figure (2). In this same way each of the successive diffraction maxima contains three interference maxima.

DIFFRACTION GRAGING

An arrangement of a large number of equidistant narrow rectangular slits of equal width placed side by side parallel to one another is known as a ***grating***. Fraunhofer produced first grating by arranging about 120 narrow slits per cm using fine wires parallel to one another. Later he developed ruled grating over glass by using fine diamond point.

There are two types of grating. They are (1) Transmission grating and (2) Reflection grating.

Transmission grating: A transmission grating consists of a well polished glass surface upon which a large number of fine equidistant parallel lines varying from 5000 to 12000 lines per cm are ruled with the help of a dividing engine. Such ruled surface has a length varying from 5 to 15 cm. Transmission gratings are ruled on plane glass surfaces. In this case, the ruled portions of the grating act as opaque regions and do not allow the light to pass through them. The region between the ruled portions acts as slits and allows the light to pass through them. The ruled regions are also known as opacities and the slit portions are known as transparencies. If the widths of opaque and transparent portions are respectively 'a' and 'b', then, the distance [a+b] is known as ***"grating constant or grating element".***

Ruled gratings are costly and they are duplicated with reduced cost using celluloid or plastic films by the following technique. In this method, a thin layer of a solution of celluloid dissolved in a volatile solvent is poured over the surface of an original grating and it is allowed to dry. The plastic film is stripped from the grating and is firmly mounted on a glass plate to prepare transmission grating.

Theory of plane transmission grating:

Let a plane wavefront falls normally on a plane transmission grating as shown in the following figure (1). The figure shows a cross section of the grating, in which, AB, CD, etc., represents the slit portions and the regions adjacent to slits, such as, BC, DE, etc., represents the opaque regions. L is a convex lens, which converges the diffracted wavefront propagating in a particular direction on a screen S arranged at the focal plane of the lens.

All the rays starting from AB on the surface of grating traveling in the direction of the incident beam reach O in phase with one another and so reinforce each other producing an image of the slits at the centre of the screen and it is known as central maximum. The rays diffracted at an angle θ with the incident direction reach a point P_1 on the screen after passing through the lens in different phases. Draw AK perpendicular to the direction of the diffracted light, then CN is the path difference between the rays diffracted from the two corresponding points A and C of the adjacent slit at an angleθ.

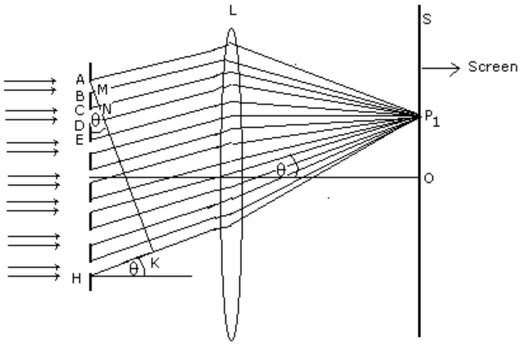


Figure (1) Fraunhofer diffraction at a plane transmission graging

The path difference, CN = AC Sin θ = (a + b) Sin θ - - - - - (1)

If, this path difference is an even multiple of [λ/2], then the point P_1 will be bright and the path difference is an odd multiple of [λ/2], then the point P_1 will be dark.

Thus, (a +b) Sin θ = ± nλ - - - - (2) [for maximum intensity at P_1]

Also, (a +b) Sin θ = ±(2n +1)λ/2 - - - - (3) [for minimum intensity at P_1]

Where n is an integer that takes values 0, 1, 2, 3, etc.

This condition holds true for all the rays from the corresponding points of any pair of adjoining slits in the entire grating surface. Due to this the bright and dark bands are alternately formed.

Equation (2) is known as grating law and the direction of n^{th} maximum with respect to incident direction is given by the formula,

$$Sin\theta = \frac{\pm n\lambda}{(a+b)} \text{ - - - - (4)}$$

In equation n = 0 corresponds to the position of central maximum and this occurs along the incident direction. The value n =1 corresponds to the position of first principal maximum and in such case, when grating element (a+b) = λ, sin θ = 1 ⇒ θ = 90^0. This means that the central maximum spreads on the entire screen and the first principal maximum on either side of central maximum is not seen. The value n = 2 gives the position of second order principal maximum

Secondary maxima and minima:

In between any two principal maxima, several secondary maxima and secondary minima are formed when the following condition is satisfied. If, for a small change in the angle of diffraction [say dθ] the path difference between rays from the adjacent slits of the grating reaching the point on the screen changes by [λ/N], where N is the total number of slits in the grating, then the path difference between the rays from the extreme slits is equal to 'λ' and as a result the effect due to one half of the slit is annulled by the effect due to the remaining half of the slit and hence the resultant intensity is minimum. Due to the same reason, if the path difference between the rays from adjacent slit changes, in general, by an integral multiple of [λ/N], the corresponding direction gives the position of secondary minima. Thus there are [N-1] secondary minima between any two principal maxima.

Similarly, for a small change in the angle 'θ', when the path difference between the rays from the corresponding points of adjacent slits changes by $\left[\frac{(2n_s+1)\lambda}{2N}\right]$, where n_s is giving the direction of n^{th} secondary maxima after any principal maximum. The value n_s = 1 gives the position of first secondary maxima after any principal maximum. Similarly n_s = 2, 3, etc., gives the positions of second, third etc., maxima between any two principal maxima. Thus there are [N-2] secondary maxima between any two principal maxima.

Width of a principal maximum:

The first secondary minimum is formed when the following condition is satisfied.

$$(a+b)Sin(\theta+d\theta)=n\lambda+\frac{\lambda}{N}$$

Expanding and simplifying, we get,

$$(a+b)Sin\theta+(a+b)Cos\theta d\theta=n\lambda+\frac{\lambda}{N};\ [\Theta\ for\ small\ d\theta, Sin\,d\theta=d\theta\ \&\ Cosd\theta=1]$$

Using grating law, we get, $(a+b)Cos\theta\ d\theta=\frac{\lambda}{N}$; or $d\theta=\frac{\lambda}{N(a+b)Cos\theta}=\frac{\lambda}{a^1}$; where a^1 is the projection of the grating width on a plane perpendicular to the direction of the given principal maximum.

DISPERSION IN A GRATING

When monochromatic light of wavelength λ is passed through a grating, several maxima are observed at positions satisfying the following condition.

$$(a+b)\ Sin\ \theta = n\lambda \quad \text{- - - (1)}$$

In equation (1) "θ" is the angle at which n^{th} maximum is observed for the wavelengthλ.

If white light is incident on the grating surface, the above equation is satisfied for all the wavelengths present in the incident light at suitable angle of diffraction. This results in a spectrum of light in the diffracted light. The white light is dispersed into its constituent colours on passing through the grating. In other words, the grating produces dispersion of white light into its constituent colours.

The dispersive power of a grating is defined as the change in the angle of diffraction corresponding to a unit change in the wavelength of light used.

If the angle of diffraction changes from θ to (θ +dθ) when the wavelength changes from λ to (λ+dλ), then the ratio [dθ/dλ] is known as the dispersive power of the grating.
Differentiating equation (1) with respect to λ, we get,

$$(a+b)Cos\theta\frac{d\theta}{d\lambda}=n$$

$$\text{Or}\ \frac{d\theta}{d\lambda}=\frac{n}{(a+b)}\frac{1}{Cos\theta}=\frac{n\times m}{Cos\theta} \quad \text{- - - - (2)}$$

In equation (2), $m=\frac{1}{(a+b)}$ denote the number of lines per cm length of the grating. Since for small values of θ, Cos θ is constant, we can write that, in a particular order, dθ α dλ. Such a spectrum is known as "normal spectrum". Also, equation (2) shows that dispersive power of a grating depends on (1) m, the number of lines per unit length of the grating and (2) n, the order of the spectrum

If 'f' is the focal length of the lens used to form a spectrum on the screen, and if 'dl ' is the separation between the spectral lines on the screen between the wavelengths λ and (λ+dλ), then
dθ = dl/f, and hence the linear dispersion per unit wavelength is given by,

$$\frac{dl}{d\lambda}=f\frac{d\theta}{d\lambda}=\frac{nf}{(a+b)Cos\theta} \quad \text{- - - - (3)}$$

Thus the linear separation between the wavelengths λ and (λ+dλ) in the n^{th} order is given by,

$$dl=\frac{nf}{(a+b)Cos\theta}\times d\lambda \quad \text{- - - - (4)}$$

Grating spectrum v/s prism spectrum:

Spectrum formed by grating is pure in the sense that there is no overlapping. Dispersion depends on the wavelength and the grating constant. Thus, the spectrum formed by different grating of same grating constant is similar and the spectrum is invariable. This is not the case in a prism spectrum, because, even though the angle of the prism is the same, change of the nature of material changes the position of the spectral lines itself causing a change in the dispersion and hence the spectrum obtained using one prism cannot be compared with that obtained with another prism.

In a prism, the angle of deviation of a spectral line has less value for light of longer wavelength, while, in the case of a grating, the angle of deviation of a spectral line has less value for light of shorter wavelength.

Resolution of optical image of two closely lying point objects:

According to Lord Rayleigh, the image of two point objects are said to be resolved if they are seen as separate points.

(a) The image of two closely lying point objects is said to be well resolved if the distance between the images of two point objects is greater than the distance between the central maximum of the first and its first minimum.

(b) The image of two closely lying point objects is said to be just resolved if the distance between the images of two point objects is equal to the distance between the central maximum of the first and its first minimum

© The image of two closely lying point objects is said to be un-resolved if the distance between the images of two point objects is less than the distance between the central maximum of the first and its first minimum

Resolving power of a grating:

Resolving power of a grating is defined as its ability to show two neighbouring lines in a spectrum as separate. It is measured by the ratio $\frac{\lambda}{d\lambda}$, where λ is the wavelength of a spectral line and dλ is the least difference in the wavelengths of the two neighbouring spectral lines which can be just resolved.

Consider two spectral lines closely lying in the spectrum, so that they are seen as just separate. According to Lord Rayleigh's criterion the two lines are seen as just separate in any order of the spectrum if the principal maximum of one line coincides with the first minimum of the other. If the angular deviation for the two just resolved spectral lines are respectively θ and θ +dθ, then following two conditions must be simultaneously satisfied for the just resolved spectral lines.

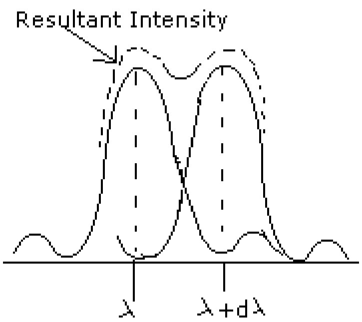


Figure (1): Resultant intensity pattern at the limit of resolution

For the spectral line of wavelength λ, n^{th} maximum is formed at an angle θ when the following condition is satisfied.

$(a+b)\ \mathrm{Sin}\ \theta = n\lambda$ - - - - (1)

In the above equation (1), (a+b) is the grating element.

For the same wavelength, the first minimum is formed at an angle (θ+dθ) if the following condition is satisfied

$(a+b)\ \mathrm{Sin}\ (\theta + d\theta) = n\lambda + \lambda/N$ - - - - (2)

Also, (θ+dθ) corresponds to the direction of n^{th} maximum for the wavelength (λ+dλ) and hence,

$(a+b)\ \mathrm{Sin}\ (\theta + d\theta) = n\ (\lambda+d\lambda)$ - - - - (3)

When the two equations (2) and (3) are simultaneously satisfied, the two closely lying spectral lines are just resolved.

Comparing equations (2) and (3) we can write,

$n\lambda + \lambda/N = n(\lambda + d\lambda) = n\lambda + n\,d\lambda \quad \Rightarrow \lambda/N = n\,d\lambda$ Or, $\frac{\lambda}{d\lambda} = n \times N$ - - - - (4)

Equation (4) shows that the resolving power is independent of the grating constant, but it increases with (i) the order of the spectrum and (ii) the total number of lines N in the aperture of the grating.

Note:

(1) Dispersive power is independent of the width (size) of the grating, while, the resolving power is inversely proportional to the total number of lines N on the grating and hence on the width (size) of the grating.

(2) Angular separation between the spectral line is unaffected by the value of N, but the sharpness of the spectral and hence the resolving power of the grating is increased by the increase in the value of N.

(3) Two different gratings having the same grating constant but different number of lines on it has the same dispersive power, but the resolving power has different value for the two gratings.

(4) Dispersive power and the resolving power are related by the following equation.

$$\frac{\lambda}{d\lambda} = N \times n = N \times [(a+b)Cos\theta]\frac{d\theta}{d\lambda} \quad ---- (5)$$

In equation (5), N (a+b) Cos θ is known as the total effective aperture.

$\therefore$ **Resolving power = dispersive power X Total effective aperture - - -(6)**