$d\sin\theta = m\lambda$ $m = 0,\ \pm1,\ \pm2,\ \pm3, \dots \dots$

$(b+c)\sin\theta = m\lambda$

which is called the grating equation and $(b+c)$ is grating element.

The zero order principal maximum occurs at $\theta = 0$, irrespective of the wavelength. Thus, if we are using white light source, then the central maxima will be of the same colour as the source itself. However, for orders other than zero, the angles of diffraction are different for different wavelength and therefore, various spectral components appear at different positions. Thus, by measuring the angles of diffraction for various colours, one can determine the values of the wavelengths. Note that, the intensity will be maximum for the zero order spectrum and falls of as the value of m increases.

4.3 FRESNEL DIFFRACTION

4.3.1 FRESNEL'S HALF PERIOD ZONES

Let BCDE represent a spherical wavefront of monochromatic light travelling towards the right (Figure 9). Every point on this sphere may be thought of as the origin of secondary wavelets. To find the resultant effect of these at a point P, divide the wavefront into zones by describing a series of circles, around the point O which is the foot of perpendicular from P and are such that each circle is a half wavelength farther from P, i.e. if OP = b, then the circles will be at distances $b + \lambda/2, b + 2\lambda/2, b + 3\lambda/2, \dots, b + m\lambda/2$ from P. The radius of the m^{th} circle will be given by

$$s_m = \left[\left(b + \lambda/2\right)^2 - b^2\right]^{\frac{1}{2}} = (mb\lambda)^{\frac{1}{2}}\left[1 + \frac{m\lambda}{4b}\right]^{\frac{1}{2}}$$

$$s_m = \sqrt{mb\lambda} \qquad \text{as } b \gg \lambda$$

The areas S_m of the zones, i.e. area of the rings between successive circles are practically equal.

$$S_m = \pi(s_m^2 - s_{m-1}^2) \cong \pi(mb\lambda - (m-1)b\lambda)$$

$$S_m \cong \pi b\lambda$$

To the approximation considered, it is therefore constant and independent of m.

These secondary wavelets will reach P with different phases, since each travels a different distance. Since each zone is $\lambda/2$ farther from P, the successive zones will produce resultants at P which differ in phase by π. The difference of half a period in the vibration from successive zones is the origin of the name half period zones.

If the resultant amplitude of the light from the m^{th} zone, the successive values of A_m will have alternating signs because changing the phase by means reversing the direction of the amplitude vector. When the resultant amplitude due to the whole wave is called A, it may be written as the sum of the series.

$$A = A_1 - A_2 + A_3 - A_4 + \dots + (-1)^{m-1}A_m$$

The amplitude produced by a particular zone is proportional to the area of the zone i.e., S_m and inversely proportional to the distance of the zones from the point P i.e., d_m. Further, it also depends on an obliquity factor which is proportional to $\frac{1}{2}(1+\cos\theta)$, where θ is the angle that the normal to the zone makes with the line connecting zone to P (Figure 10). Thus, we may express the amplitude due to m^{th} zone is,

$$A_m = constant\,\frac{S_m}{d_m}(1+\cos\theta)$$

where d_m is the average distance to P and θ is the angle at which the light leaves the zone.

The effect of obliquity factor $(1+\cos\theta)$ is causing the successive terms in the expression for A to decrease very slowly. The decrease is least slow at first, because of the rapid change of θ with m, but the amplitudes soon become nearly equal. With the knowledge of the variation in magnitude of the terms, we may evaluate the sum of the series by grouping its terms in the following two ways. Supposing m to be odd, we may write

$$A = \frac{A_1}{2} + \left(\frac{A_1}{2} - A_2 + \frac{A_3}{2}\right) + \left(\frac{A_3}{2} - A_4 + \frac{A_5}{2}\right) + \cdots + \frac{A_m}{2}$$

$$A = A_1 - \frac{A_2}{2} - \left(\frac{A_2}{2} - A_3 + \frac{A_4}{2}\right) - \left(\frac{A_4}{2} - A_5 + \frac{A_6}{2}\right) - \cdots - \frac{A_{m-1}}{2} + A_m$$

Now since the amplitude A_1, A_2, ... do not decrease at a uniform rate, each one is smaller than the arithmetic mean of the preceding and following ones. Therefore, the quantities in brackets in the above equations are all positive and the following inequalities must hold

$$\frac{A_1}{2} + \frac{A_m}{2} < A < A_1 - \frac{A_2}{2} - \frac{A_{m-1}}{2} + A_m$$

Because the amplitudes for any two adjacent zones are very nearly equal, it is possible to equate A_1 to A_2 and A_{m-1} to A_m. The result is

$$A = \frac{A_1}{2} + \frac{A_m}{2}$$

If m is taken to be even, we find by the same method that

$$A = \frac{A_1}{2} - \frac{A_m}{2}$$

Hence, the conclusion is that the resultant amplitude at P due to m zones is either the sum or difference of half of the amplitudes contributed by the first and last zones. If we allow m to become large enough for the entire spherical wave to be divided into zones, θ approaches 180 degree for the last zone. Therefore, the obliquity factor causes A_m to become negligible and the amplitude due to the whole wave is just half that due to the first zone acting alone.

The vector addition of the amplitudes A_1, A_2, ..., which are alternatively positive and negative, would be performed by drawing them along the same line, but here they are separated in a horizontal direction. The tail of each vector is put at the same height as the head of the previous one. Then the resultant amplitude A due to any given number of zones will be the height of the final arrowhead above the horizontal base line, as shown in Figure 11 for 12 zones.

SATISH CHANDRA

4.3.2 ZONE PLATE

Principle: A beautiful application of the concept of Fresnel's half period zones lies in the construction of the zone plate. This is a special screen designed to block off the light from every other half period zones. The result is to remove either all the positive terms or all the negative terms in the following equation,

$$A = A_1 - A_2 + A_3 - A_4 + \cdots + (-1)^{m-1} A_m$$

which is the resultant amplitude due to whole wave front divided into m number of zones.

In either case the amplitude at P will be increased to many times its value, because if the light is obstructed from either even half period zone or odd half period zone then the resultant amplitude at image point P becomes either

$$A = A_1 + A_3 + A_5 + \cdots + A_m$$

or, $A = -(A_2 + A_4 + A_6 + \cdots + A_m)$, respectively.

Construction: Its construction is based on the following there facts.

i. The areas of half period zones are equal.
ii. The radii of half period zones are proportional to the square root of natural number.
iii. If the light is obstructed from alternate zones, the intensity of the image is increases.

A zone plate can easily be made in practice by drawing concentric circles on white paper with radii proportional to the square roots of whole numbers. Every other zone is then blackened and the result is photographed on a reduced scale. The negative, when held in the light from a distant point source, produces a large intensity at a point on its axis at a distance corresponding to the size of the zones and the wavelength of the light used.

Theory: The zone plate (as shown in Figure 12) can also be used for imaging points on the axis, *e.g.* if we have a point source at S then a bright image will be formed at P, where the point P should be such that

$$SL + LP - SP = m\frac{\lambda}{2}$$

the point L being on the periphery of the first circle of the zone plate (shown in Figure 13). If the radius of the m^{th} circle is s_m, then

$$SL + LP - SP = (a^2 + s_m^2)^{1/2} + (b^2 + s_m^2)^{1/2} - (a + b)$$

$$SL + LP - SP = a\left(1 + \frac{s_m^2}{a^2}\right)^{1/2} + b\left(1 + \frac{s_m^2}{b^2}\right)^{1/2} - (a + b)$$

Using Binomial expansion,

$$SL + LP - SP \approx a\left(1 + \frac{s_m^2}{2a^2}\right) + b\left(1 + \frac{s_m^2}{2b^2}\right) - (a + b)$$

SATISH CHANDRA

Hence,

$$m\frac{\lambda}{2} \approx \frac{s_m^2}{2}\left(\frac{1}{a}+\frac{1}{b}\right)$$

The bright spot produced by a zone plate is so intense that the plate acts much like a lens. Thus suppose that the first 10 odd zones are exposed. This leaves the amplitudes A_1, A_3, . . . , A_{19}; the sum of which is nearly 10 times of A_1. The whole wavefront gives $A_1/2$, so that using only 10 exposed zones, we obtain amplitude at P which is 20 times as great as when the plate is removed. The intensity is therefore 400 times as great (because intensity is proportional to the square of amplitude). If the odd zones are covered, the amplitude A_2, A_4, . . . , A_{20} will give the same effect. The object and image distances obey the ordinary lens formula since,

$$\left(\frac{1}{a}+\frac{1}{b}\right) = m\frac{\lambda}{s_m^2} = \frac{1}{f}$$

where

$$f = \frac{s_m^2}{m\lambda}$$

the focal length f being the value of b for $a = \infty$, namely

$$b = f = \frac{s_m^2}{m\lambda}$$

Here, f is called the principal focal length of the zone plate. Thus, the zone plate acts as a convergent (convex) lens with multiple foci for a particular wavelength depending on the values of m and s_m.

Multiple foci of zone plate: From above equation, we can obtain more correct expression for the area of any zone. It becomes

$$S_m = \pi(s_m^2 - s_{m-1}^2) = \left(\frac{a}{a+b}\right)\pi b\lambda$$

Now consider a point Q along the axis of the zone plate at a distance $b/3$ from the zone plate then the area of each half period zone with respect to Q will be $\frac{1}{3}\pi b\lambda$ i.e., one third of the area of each zone on the zone plate. Hence, each zone on the zone plate will contain three half period elements contain three half period elements corresponding to Q, thus the resultant amplitude would be,

$$A = (A_1 - A_2 + A_3) + (A_7 - A_8 + A_9) + \cdots$$

The effect of two of them cancel almost, but that of the third is left over in each bracket which would again correspond to a maximum, but it would not be as intense as the point P. Point Q represents the position of the second focal point. It is given by

$$f_3 = \frac{f}{3} = \frac{s_m^2}{3m\lambda}$$

Similarly, there are more fainter images corresponding to focal length $f/3, f/5, f/7, \ldots$ because at distances $b/3, b/5, b/7, \ldots$ each zones of the plate includes 3, 5, 7, . . . Fresnel's zones.

SATISH CHANDRA

In the similar way we can understand for points at distances $b/2$, $b/4$, $b/6$, . . . each zone of the plate includes 2, 4, 6, . . . Fresnel's zones. Thus, the effect of each pair in any case would almost cancel out and therefore, the points at these distances correspond to minima, which decrease gradually. Thus, we can say that zone plate has multiple foci.

4.3.3 STRAIGHT EDGE

Introduction: The diffraction by a single screen with a straight edge is the simplest application of half period zones. Figure 14, represents a section of such a screen having its edge parallel to the slit S, so that the straight diffraction fringes produced by each element of its length are all lined up on the observing screen. Therefore, it is possible to regard the wavefront as cylindrical as shown in Figure 14.

To construct half period elements on a cylindrical wave front, divide it into strips, the edge of which is successively half wavelength farther from point P. In Figure 14, the half period strips corresponding to the point P being situated on the edge of the geometrical shadow are marked off on the wavefront. The points M_0, M_1, M_2, . . . on the circular section of the cylindrical wave are at distances $b, b+{}^{\lambda}/_{2}, b+{}^{2\lambda}/_{2}, \ldots$ from P. M_0 is on the straight line SP.

According to the law of geometrical optics, we should get uniform brightness above point P and complete darkness below this. But actually it is observed that there are a few unequally spaced diffraction fringes in the illuminated region close to P, and in the region of darkness the intensity does not become zero at P.

However, unlike the half period zones, the areas of the half period strips will not be equal and thus the analysis become quite difficult. Even then one can draw the following conclusions.

(i) Corresponding to the edge of the geometrical shadow, shown as P, half of the wavefront is obstructed by the edge. Hence, the amplitude will be given by

$$A=\frac{1}{2}A_0=\frac{1}{4}A_1$$

where A_0 represents the amplitude that would be produced by the unobstructed wavefront (i.e. in the absence of the edge). Thus, the intensity at the edge of shadow is just ¼ of that found above for unobstructed wave.

(ii) Let the next assume that the point P' satisfies the following relation,

$$SN+NP'-SP'=\tfrac{1}{2}\lambda$$

i.e., P' lies in the direction SM_1, where M_1 is the upper edge of the first half period strip. For this point, the centre M_0 of the half period strips lies on the straight line joining S with P' and the Figure 14 must be reconstructed as in Figure 15. The edge now lies at the point M'_1 so that not only all the half period strips above M_0 are exposed but also the first one below M_0. Thus the resultant amplitude would approximately be

$$\frac{A_1}{2}+\frac{A_1}{4}=\frac{3}{4}A_1=\frac{3}{2}A_0$$

© Satish Chandra

SATISH CHANDRA

where $A_1/2$ is the amplitude produced by the first half period strip in the lower portion, i.e. by $M_0M'_1$ and $A_1/4$ is the resultant amplitude produced by the upper half of the wavefront i.e. by M_0G region. The intensity would be 9/4 of that found unobstructed.

(iii) Consider next the intensity at P" such that

$$SN + NP'' - SP'' = \lambda$$

we will have a minimum and the resultant amplitude will be

$$\left(\frac{A_1}{2} - \frac{A_2}{2}\right) + \frac{A_1}{4} = \frac{3A_1}{4} - \frac{A_2}{2} = \frac{3}{2}A_0 - \frac{1}{2}A_2$$

Intensity Distribution: In general, an arbitrary point P, will correspond to maximum intensity if,

$$SN + NP - SP = \left(m + \tfrac{1}{2}\right)\lambda \qquad \text{where } m = 0, 1, 2, \ldots\ldots$$

and minimum intensity if,

$$SN + NP - SP = m\lambda \qquad \text{where } m = 0, 1, 2, \ldots\ldots$$

Now, from Figure 14 we have that

$$NP = (b^2 + x^2)^{1/2} \approx b\left(1 + \frac{x^2}{2b^2}\right) = b + \frac{x^2}{2b}$$

$$SP = [(a+b)^2 + x^2]^{1/2} \approx (a+b)\left(1 + \frac{x^2}{2(a+b)^2}\right) = (a+b) + \frac{x^2}{2(a+b)}$$

here x is the distance of point P from the edge of the geometrical shadow. Hence,

$$SN + NP - SP = \left[a + b + \frac{x^2}{2b} - (a+b) - \frac{x^2}{2(a+b)}\right] = \frac{a}{2b(a+b)}x^2$$

Thus, when

$$x = \left[\frac{2b(a+b)}{a}\left(m + \tfrac{1}{2}\right)\lambda\right]^{1/2}$$

we will have a maximum. Thus, the first maximum will occur near to the edge of the shadow and the second, third and other maxima will occur at some increasing distances from the edge. The distance between two consecutive maxima will decrease as we go away from the edge of the geometrical shadow.

Similarly the position of the minima are given by,

$$x = \left[\frac{2b(a+b)}{a}m\lambda\right]^{1/2}$$

SATISH CHANDRA

The distribution of minima along screen will be in similar way as the maxima. By determining the positions of these maxima and minima one can calculate the wavelength.

The intensity distribution on the screen due to a straight edge is shown in Figure 16. Here, S is the source, AD is the straight edge and MN is the screen. In illuminating part PM alternative bright and dark bands of gradually decreasing intensity will be observed and the intensity falls off gradually in the region of geometrical shadow. Thus, according to the wave theory, the shadow cast by obstacles is not sharp and hence rectilinear propagation of light is only approximately true.

4.3.4 CIRCULAR APERTURE

Let AB be a small circular aperture (a pin hole, say) having radius r and centre at O and S is a point source of monochromatic light wavelength, λ. Let XY is a screen and P is an axial point on the screen lying on the line passing through O and normal to the plane of aperture. GOG' is the trace of incident spherical wavefront (Figure 17).

Intensity at axial points: Let us examine the effect upon the intensity at P. If the radius of the hole r is made equal to the distance S_1 to the outer edge of the first half period zone, the amplitude will be A_1 and this is twice the amplitude due to the unscreened wave. Thus, the intensity at P is four times as great as if the screen were absent. When the radius of the hole is increased until it includes the first two zones, the amplitude is $A_1 - A_2$, or almost zero. A further increase of r will cause the intensity to pass through maxima and minima each time the number of zones included become odd or even.

The same effect is produced by moving the point of observation P towards or away from the aperture along the perpendicular. This varies the size of zones, so that if P is originally at a position such that PR – PO is $\lambda/2$ (one zone included), moving P towards the screen will increase this path difference to $2\lambda/2$ (two zone included), $3\lambda/2$ (three zone included) etc. We thus have maxima along the axis of the aperture.

If the screen is placed at such a distance or the size of aperture is such that the aperture allows m full half period zones, then the path difference between the waves reaching at P along paths SAP and SOP is

$$SA + AP - SP = m\frac{\lambda}{2}$$

where $m = 0, 1, 2, \ldots ....$ Using trigonometry,

$$SA = (r^2 + a^2)^{1/2} \approx a\left(1 + \frac{r^2}{2a^2}\right) = a + \frac{r^2}{2a}$$

$$AP = (r^2 + b^2)^{1/2} \approx b\left(1 + \frac{r^2}{2b^2}\right) = b + \frac{r^2}{2b}$$

$$SP = a + b$$

Thus,

$$\frac{r^2}{2}\left(\frac{1}{a} + \frac{1}{b}\right) = m\frac{\lambda}{2}$$

SATISH CHANDRA

$$\frac{1}{a}+\frac{1}{b}=m\frac{\lambda}{r^2}$$

The point P will be bright or dark corresponding as m is odd or even. If the same is situated at infinite, $a=\infty$. Then,

$$r=\sqrt{mb\lambda}$$

Intensity at non-axial points: In order to calculate the intensity at a non-axial point consider a point P' off the axis slightly above P. The pole of the wavefront with respect to P' is O' slightly above O. Suppose that the wave surface within the aperture contains m zones with respect to P. If m is even the point P will be dark. The amplitude at P will be

$$A_p=A_1-A_2+A_3-A_4+\cdots-A_m \qquad \text{at P}$$

Now consider that the point P' is such that with respect to it, the m^{th} half period zone (constructed with O' as the pole of the wavefront as shown in Figure 18) is cut from the upper half $(m+1)^{th}$ half period zone is exposed from the lower half. As m is even A_m is negative and A_{m+1} is positive. Thus, the amplitude at P' is

$$A_{p'}=A_1-A_2+A_3-A_4+\cdots-\frac{A_m}{2}+\frac{A_{m+1}}{2}$$

$$A_{p'}=(A_1-A_2+A_3-A_4+\cdots-A_m)+\frac{1}{2}(A_m+A_{m+1})$$

$$A_{p'}=\text{amplitude at point P}+\text{amplitude due to one zone} \qquad \text{at P'}$$

So under these conditions P' will be bright.

Now consider another point P'' slightly above P'. The pole of the wavefront with respect to P'' is O'' slightly above O'. Let point P'' be such that with respect to it $(m-1)^{th}$ zone is cut from the upper half and $(m+2)^{th}$ zone is exposed from the lower half (Figure 19). As m is even, A_{m-1} is positive and A_{m+2} is negative. Therefore, amplitude at P'' is

$$A_{p''}=A_1-A_2+A_3-A_4+\cdots+\frac{A_{m-1}}{2}-\frac{A_m}{2}+\frac{A_{m+1}}{2}-\frac{A_{m+2}}{2}$$

$$A_{p''}=(A_1-A_2+A_3-A_4+\cdots+A_{m-1}-A_m)+\frac{1}{2}(A_m+A_{m+1})-\frac{1}{2}(A_{m-1}+A_{m+2})$$

$$A_{p''}=\text{amplitude at point P}'-\text{amplitude due to one zone} \qquad \text{at P''}$$

So under these conditions P'' will almost be dark.

Thus, if the point P is dark and we move sideways (below or above P), we get alternatively bright and dark points. As the circular aperture is symmetrical about the axis, we get alternate bright and dark rings about the centre at P. If m is odd, the axial point P will be bright and we get alternate dark and bright rings about point P.

SATISH CHANDRA