18
FRESNEL DIFFRACTION

The diffraction effects obtained when either the source of light or the observing screen, or both, are at a finite distance from the diffracting aperture or obstacle come under the classification of *Fresnel diffraction*. These effects are the simplest to observe experimentally, the only apparatus required being a small source of light, the diffracting obstacle, and a screen for observation. In the Fraunhofer effects discussed in the preceding chapters, lenses were required to render the light parallel, and to focus it on the screen. Now, however, we are dealing with the more general case of divergent light which is not altered by any lenses. Since Fresnel diffraction is the easiest to observe, it was historically the first type to be investigated, although its explanation requires much more difficult mathematical theory than that necessary in treating the plane waves of Fraunhofer diffraction. In this chapter we consider some of the simpler cases of Fresnel diffraction, which are amenable to explanation by fairly direct mathematical and graphical methods.

18.1 SHADOWS

One of the greatest difficulties in the early development of the wave theory of light lay in the explanation of the observed fact that light appears to travel in straight lines. Thus if we place an opaque object in the path of the light from a point source, it casts

Scanned with CamScanner

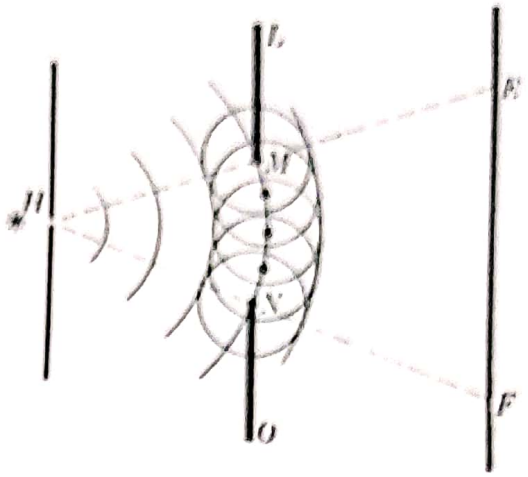

FIGURE 18A
Huygens' principle applied to secondary wavelets from a narrow opening.

a shadow having a fairly sharp outline of the same shape as the object. It is true, however, that the edge of this shadow is not absolutely sharp and that when examined closely it shows a system of dark and light bands in the immediate neighborhood of the edge. In the days of the corpuscular theory of light, attempts were made by Grimaldi and Newton to account for such small effects as due to the deflection of the light corpuscles in passing close to the edge of the obstacle. The correct explanation in terms of the wave theory we owe to the brilliant work of Fresnel. In 1815 he showed not only that the approximately rectilinear propagation of light could be interpreted on the assumption that light is a wave motion but also that in this way the diffraction fringes could in many cases be accounted for in detail.

To bring out the difficulty of explaining shadows by the wave picture, let us consider first the passage of divergent light through an opening in a screen. In Fig. 18A the light originates from a small pinhole H, and a certain portion MN of the divergent wave front is allowed to pass the opening. According to Huygens' principle, we may regard each point on the wave front as a source of secondary wavelets. The envelope of these at a later instant gives a divergent wave with H as its center and included between the lines HE and HF. This wave as it advances will produce strong illumination in the region EF of the screen. But also part of each wavelet will travel into the space behind LM and NO, and hence might be expected to produce some light in the regions of the geometrical shadow outside of E and F. Common experience shows that there is actually no illumination on these parts of the screen, except in the immediate vicinity of E and F. According to Fresnel, this is to be explained by the fact that in the regions well beyond the limits of the geometrical shadow the secondary wavelets arrive with phase relations such that they interfere destructively and produce practically complete darkness.

The secondary wavelets cannot have uniform amplitude in all directions, since if this were so, they would produce an equally strong wave in the backward direction. In Fig. 18A the envelope on the left side of the screen would represent a reverse wave converging toward H. Obviously such a wave does not exist physically, and hence we must assume that the amplitude at the back of a secondary wave is zero. The more exact formulation of Huygens' principle justifies this assumption and also gives

Scanned with CamScanner

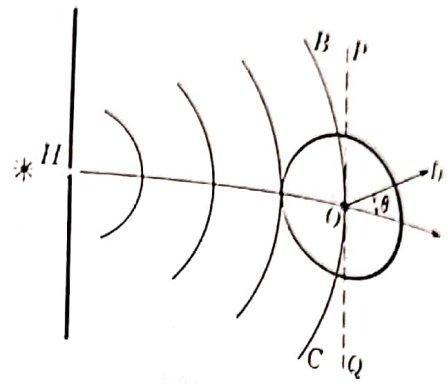

FIGURE 18B
The obliquity factor for Huygens' secondary wavelets.

quantitatively the variation of the amplitude with direction. The so-called *obliquity factor*, as is illustrated in Fig. 18B, requires an amplitude varying as $1 + \cos\theta$, where θ is the angle with the forward direction. At right angles, in the directions P and Q of the figure, the amplitude falls to one-half and the intensity to one-quarter of its maximum value. Another property that the wavelets must be assumed to have, in order to give the correct results, is an advance of phase of one-quarter period ahead of the wave that produces them. The consequences of these two rather unexpected properties and the manner in which they are derived will be discussed later.

18.2 FRESNEL'S HALF-PERIOD ZONES

As an example of Fresnel's approach to diffraction problems, we first consider his method of finding the effect that a slightly divergent spherical wave will produce at a point ahead of the wave. In Fig. 18C let $BCDE$ represent a spherical wave front of monochromatic light traveling toward the right. Every point on this sphere may be thought of as the origin of secondary wavelets, and we wish to find the resultant effect of these at a point P. To do this, we divide the wave front into *zones* by the following construction. Around the point O, which is the foot of the perpendicular from P, we describe a series of circles whose distances from O, measured along the arc, are $s_1, s_2, s_3, \ldots, s_m$ and are such that each circle is a half wavelength farther from P. If the distance $OP = b$, the circles will be at distances $b + \lambda/2$, $b + 2\lambda/2$, $b + 3\lambda/2, \ldots, b + m\lambda/2$ from P.

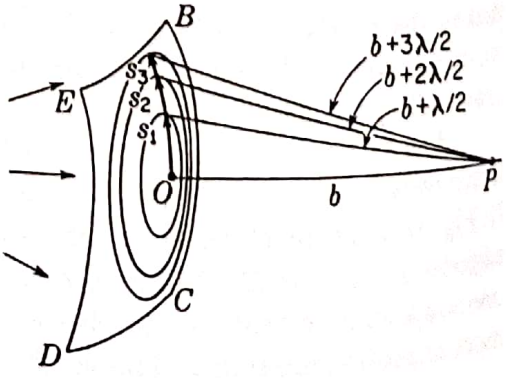

FIGURE 18C
Construction of half-period zones on a spherical wave front.

Scanned with CamScanner

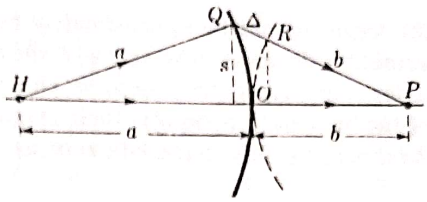

FIGURE 18D
Path difference Δ at a distance s from the pole of a spherical wave.

The areas S_m of the zones, i.e., of the rings between successive circles, are practically equal. In proving this, we refer to Fig. 18D, where a section of the wave spreading out from H is shown with radius a. If a circle of radius b is now drawn (broken circle) with its center at P and tangent to the wave front at its "pole" O, the path HQP is longer than HOP by the segment indicated by Δ. For the borders of the zones, this path difference must be a whole multiple of $\lambda/2$. To evaluate it, we note first that in all optical problems the distance s is small compared with a and b. Then s may be considered as the vertical distance of Q above the axis, and Δ may be equated to the sum of the sagittas of the two arcs OQ and OR. By the sagitta formula we then have

$$\Delta = \frac{s^2}{2a} + \frac{s^2}{2b} = s^2 \frac{a+b}{2ab} \tag{18a}$$

The radii s_m of the Fresnel zones are such that

$$m\frac{\lambda}{2} = {s_m}^2 \frac{a+b}{2ab} \tag{18b}$$

and the area of any one zone becomes

$$S_m = \pi({s_m}^2 - {s_{m-1}}^2) = \pi \frac{\lambda}{2} \frac{2ab}{a+b} = \frac{a}{a+b}\pi b\lambda \tag{18c}$$

To the approximation considered, it is therefore constant and independent of m. A more exact evaluation would show that the area increases very slowly with m.

By Huygens' principle we now regard every point on the wave as sending out secondary wavelets in the same phase. These will reach P with different phases, since each travels a different distance. The phases of the wavelets from a given zone will not differ by more than π, and since each zone is on the average $\lambda/2$ farther from P, it is clear that the successive zones will produce resultants at P which differ by π. This statement will be examined in more detail in Sec. 18.6. The difference of half a period in the vibrations from successive zones is the origin of the name *half-period zones*. If we represent by A_m the resultant amplitude of the light from the mth zone, the successive values of A_m will have alternating signs because changing the phase by π 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

$$A = A_1 - A_2 + A_3 - A_4 + \cdots + (-1)^{m-1}A_m \tag{18d}$$

Three factors determine the magnitudes of the successive terms in this series: (1) because the area of each zone determines the number of wavelets it contributes,

REDMI NOTE 7 PRO
AI DUAL CAMERA
2020/4/2 18:17
Scanned with CamScanner

the terms should be approximately equal but should increase slowly; (2) since the amplitude decreases inversely with the average distance from P of the zone, the magnitudes of the terms are reduced by an amount which increases with m; and (3) because of the increasing obliquity, their magnitudes should decrease. Thus we may express the amplitude due to the mth zone as

$$A_m = \text{const}\,\frac{S_m}{d_m}(1 + \cos\theta) \tag{18e}$$

where d_m is the average distance to P and θ the angle at which the light leaves the zone. It appears in the form shown because of the obliquity factor assumed in the preceding section. Now an exact calculation of the S_m's shows that the factor b in Eq. (18c) must be replaced by $b + \Delta$, where Δ is the path difference for the middle of the zone. Since at the same time $d_m = b + \Delta$, we find that the ratio S_m/d_m is a constant, independent of m. Therefore we have left only the effect of the obliquity factor $1 + \cos\theta$, which causes the successive terms in Eq. (18d) 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 this 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,

$$\begin{aligned} A &= \frac{A_1}{2} + \left(\frac{A_1}{2} - A_2 + \frac{A_3}{2}\right) + \left(\frac{A_3}{2} - A_4 + \frac{A_5}{2}\right) + \cdots + \frac{A_m}{2} \\ &= A_1 - \frac{A_2}{2} - \left(\frac{A_2}{2} - A_3 + \frac{A_4}{2}\right) - \left(\frac{A_4}{2} - A_5 + \frac{A_6}{2}\right) - \cdots - \frac{A_{m-1}}{2} + A_m \end{aligned} \tag{18f}$$

Now since the amplitudes $A_1, A_2, \ldots$ 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 parentheses 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

$$A = \frac{A_1}{2} + \frac{A_m}{2} \tag{18g}$$

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

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

ence the conclusion is that the resultant amplitude at P due to m zones is either half e sum or half the difference of the amplitudes contributed by the first and last zones. we allow m to become large enough for the entire spherical wave to be divided into

Scanned with CamScanner

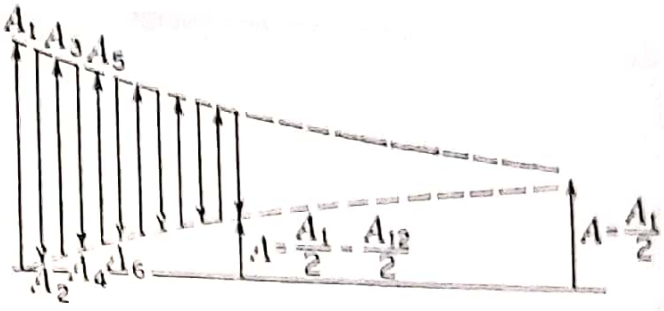

FIGURE 18E
Addition of the amplitudes from half-period zones.

zones, θ approaches 180° 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.

Figure 18E shows how these results can be understood from a graphical construction. The vector addition of the amplitudes $A_1, A_2, A_3, \ldots$, which are alternately positive and negative, would normally be performed by drawing them along the same line, but here for clarity 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. In the figure, it is shown for 12 zones and also for a very large number of zones.

Scanned with CamScanner