Matt Young, *Optics and Lasers: Including Fibers and Optical Waveguides*, 5th ed. (Springer, New York, 2000).

5.3 Interference by Division of Wavefront

The *wavefront* refers to the maxima (or other planes of constant phase) as they propagate. The wavefront is normal to the *direction of propagation*. One way of bringing about interference is by dividing the wavefronts into two or more segments and recombining the segments elsewhere.

5.3.1 Double-Slit Interference

Suppose a monochromatic *plane wave* (a collimated beam, or a beam with plane wavefronts) is incident on the opaque screen shown in Fig. 5.3. Two infinitesimal slits a distance d apart have been cut into the screen. Each slit behaves as a point source, radiating in all directions. We set up an observing screen a great distance L away from the slits. Light from both slits falls on this screen. The electric field at a point P is the sum of the fields originating from each slit

$$E = A(\mathrm{e}^{-ikr_1} + \mathrm{e}^{-ikr_2})\mathrm{e}^{\mathrm{i}\omega t} , \tag{5.34}$$

where A is the amplitude of the waves at the viewing screen and r_1, r_2 are the respective distances of the slits from P. Because the factor $\mathrm{e}^{\mathrm{i}\omega t}$ is common to all terms and will vanish from the intensity, we shall hereafter drop it.

If L is sufficiently large, r_1 and r_2 are effectively parallel and differ only by $d \sin\theta$. Thus

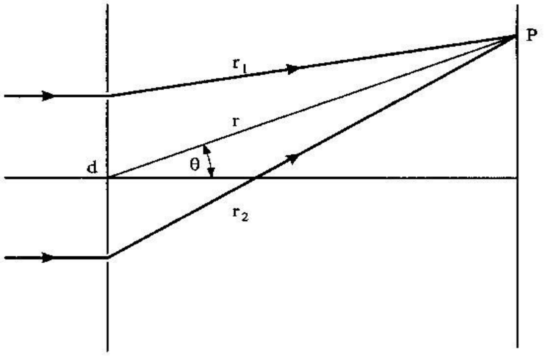

Fig. 5.3. Double-slit interference

$$E = Ae^{-ikr_1}(1 + e^{-ikd\sin\theta}) \ . \tag{5.35}$$

The phase difference between the two waves is

$$\phi = kd\sin\theta \ , \tag{5.36}$$

and we can immediately write

$$\boxed{I \propto 4\cancel{A^2}\cos^2(\frac{\pi}{\lambda}d\sin\theta)} \tag{5.37}$$

from the earlier treatment of superposition. $d\sin\theta$ is called the *optical path difference* (OPD) between the two waves. For small angles, $\sin\theta = x/L$, and the *interference pattern* has a $\cos^2$ variation with x. Maxima occur whenever the argument of the cosine is an integral multiple of π, or where

$$OPD = m\lambda \ \text{(constructive interference)} \ . \tag{5.38a}$$

This result is generally true and comes about because the waves have a relative phase equal to a multiple of 2π whenever the optical path difference between them is an integral multiple of the wavelength.

Similarly, minima (in this case, zeros) occur whenever

$$OPD = (m + 1/2)\lambda \ \text{(destructive interference)}. \tag{5.38b}$$

When this relation holds, the waves arrive at the observing screen exactly 180° out of phase. If the waves have equal amplitudes they cancel each other precisely.

In fact, the $\cos^2$ fringes do not extend infinitely far from the axis. This is so for at least two reasons: (a) the light is not purely monochromatic, and (b) the slits are not infinitesimal in width.

The first relates to the *coherence* of the light, which we shall discuss later. This effect brings about a superposition of many double-slit patterns, one for each wavelength, so to speak. Each wavelength brings about slightly

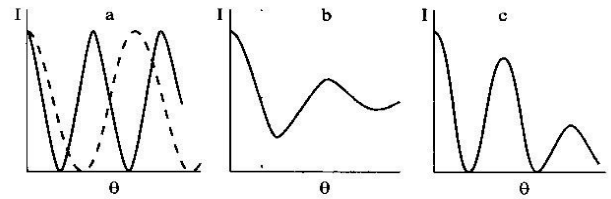

Fig. 5.4. **(a)** Superposition of two incoherent, double-slit interference patterns. **(b)** Double-slit pattern with light that is not monochromatic. **(c)** Double-slit pattern with finite slits

different fringe pattern from the rest, and at large angles θ the patterns do not coincide exactly (Fig. 5.4a). This results in the washing out and eventual disappearance of the fringes, as shown in Fig. 5.4b.

The second effect has to do with diffraction, which we also discuss later. In the derivation, we assumed each slit to radiate uniformly in all directions. This assumption is valid only for zero slit width. A finite slit radiates primarily into a cone whose axis is the direction of the incident light. For this reason, the intensity of the pattern falls nearly to 0 for large θ. With a good, monochromatic source, this is usually the important effect and is shown in Fig. 5.4c.

5.3.2 Multiple-Slit Interference

If we generalize from two slits to many (Fig. 5.5), we find that the OPD between rays coming from adjacent slits is $d\sin\theta$. Thus, the OPD between the first and the jth slit is $(j-1)d\sin\theta$. The total electric field at a point on the distant observation screen is a sum not of two terms, but of many,

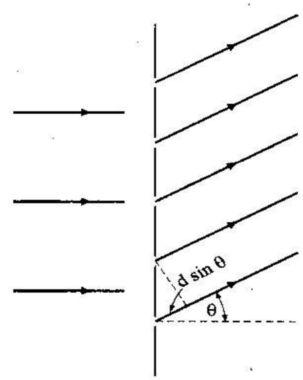

Fig. 5.5. Multiple-slit interferometer

$$E = Ae^{-ikr_1}[1 + e^{-i\phi} + e^{-2i\phi} + e^{-3i\phi} + \cdots + e^{-(N-1)i\phi}] \ , \tag{5.39}$$

where N is the number of slits and $\phi = kd \sin\theta$ as before.

The term in brackets is a geometric series whose common ratio is $e^{-i\phi}$. The sum of the terms in the series may be found by a well-known formula to be

$$\text{series sum} = \frac{1 - e^{-iN\phi}}{1 - e^{-i\phi}} \ . \tag{5.40}$$

We use the same technique as before: factor $e^{-iN\phi/2}$ from the numerator and $e^{-i\phi/2}$ from the denominator, and rewrite the sum as

$$\text{series sum} = e^{-i(N-1)\phi/2}\,\frac{\sin N\phi/2}{\sin \phi/2} \ . \tag{5.41}$$

Thus, the intensity of the interference pattern is

$$\boxed{I(\theta) = A^2 \frac{\sin^2 N\phi/2}{\sin^2 \phi/2} = A^2 \frac{\sin^2(\frac{\pi}{\lambda} N d \sin\theta)}{\sin^2(\frac{\pi}{\lambda} d \sin\theta)}} \tag{5.42}$$

At certain values of θ, the denominator vanishes. Fortunately, the numerator vanishes at (among others) the same values of θ. The indeterminate form 0/0 must be evaluated by studying the limit of $I(\theta)$ as θ approaches one of these values. The evaluation is particularly simple as θ approaches 0, where the sine is replaced by its argument. Thus,

$$\lim_{\theta \to 0} I(\theta) = A^2 N^2 \ . \tag{5.43}$$

The denominator is 0 at other values of θ, and intuition shows that $I(\theta)$ approaches N^2A^2 in those cases as well.

If N is a fairly large number, $I(\theta)$ is large at these angles. Conservation of energy requires that $I(\theta)$ be relatively small at all other angles, and direct calculation will bear this out.

A typical interference pattern is sketched in Fig. 5.6. The sharp peaks are known as *principal maxima* and appear only when

$$\frac{\pi}{\lambda} d \sin\theta = m\pi; \ m = 0, \pm 1, \pm 2, \ldots \ , \tag{5.44}$$

or when

$$\boxed{m\lambda = d \sin\theta \ . \simeq d\theta} \tag{5.45}$$

This is known as the *grating equation*, and m is known as the *order number* or *order*.

The smaller peaks are called *secondary maxima* and appear because of the oscillatory nature of the numerator of $I(\theta)$. When $N \gg 1$, the secondary maxima are relatively insignificant, and the intensity appears to be 0 at all angles where the grating equation is not satisfied. At all angles that satisfy the grating equation, the intensity is N^2A^2; it falls rapidly to 0 at other angles.

5.5.1 Single-Slit Diffraction

This is shown in one dimension in Fig. 5.12. We appeal to Huygens's construction and assume that each element ds of the slit radiates a spherical wavelet. The observing screen is located a distance L away from the aperture, and we seek the intensity of the light diffracted at angle θ to the axis.

The center O of the aperture is located a distance r from the observation point P. The optical path difference between the paths from θ and from the element ds (at s) is $s \sin\theta$, in the Fraunhofer approximation.

The electric field at P arising from the element is

$$dE = A\frac{e^{-ik(r + s\sin\theta)}}{r} ds \ . \tag{5.55}$$

Here, A is the amplitude of the incident wave, assumed constant across the aperture. We obtain the r in the denominator by realizing that the element is essentially a point source. The intensity from the point source obeys the inverse-square law, so the amplitude falls off as $1/r$. We drop $s \sin\theta$ from the denominator because it is small compared with r. We cannot, however, drop it from the phase term $k(r + s\sin\theta)$ because very small changes of $s \sin\theta$ cause pronounced changes of the phase of the wavelet relative to that of another wavelet.

The total field at P is the sum of the fields due to individual elements. If the width of the slit is b and its center, $s = 0$, this sum is just the integral

$$E(\theta) = A\frac{e^{-ikr}}{r} \int_{-b/2}^{b/2} e^{-(ik\sin\theta)s} ds \ , \tag{5.56}$$

where constant terms have been removed from the integral. The integrand is of the form $\exp(as)$, so the integral is easily evaluated:

$$E(\theta) = A\frac{e^{-ikr}}{r} \frac{2\sin[(kb\sin\theta)/2]}{ik\sin\theta} \ . \tag{5.57}$$

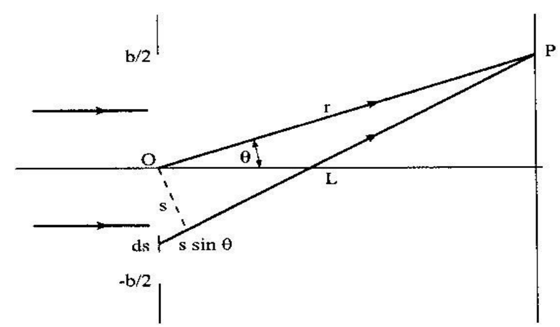

Fig. 5.12. Fraunhofer diffraction by a single opening

If we multiply both numerator and denominator by b, and define

$$\beta = \frac{1}{2} kb \sin\theta, \tag{5.58}$$

we may write

$$E(\theta) = \frac{Ab}{r} e^{-ikr} \left(\frac{\sin\beta}{\beta}\right) \tag{5.59a}$$

or

$$\boxed{I(\theta) \propto \cancel{\frac{I_0 b^2}{r^2}} \left(\frac{\sin\beta}{\beta}\right)^2} . \tag{5.59b}$$

More proper analysis, based on electromagnetic theory and a two-dimensional integration would include an additional factor of i/λ in the expression (5.59a) for $E(\theta)$.

In addition, rigorous theory predicts another factor called the *obliquity factor* whose functional form has been the subject of debate. The correct form is most likely $\cos\theta$, though the function $(1 + \cos\theta)/2$ has also been proposed. Both forms depend on the assumption that the plane wave remains planar in spite of possible distortion of the electric field inside the aperture. The obliquity factor is 0 in the reverse direction and explains why Huygens wavelets do not propagate backward. Otherwise, it is generally unimportant, but may have to be considered in studies of optical fibers that have small cores and therefore high numerical apertures (see Sect. 11.7). For our purposes, only the variable $(\sin\beta)/\beta$ is important, since we will generally be interested in relative intensities only.

Figure 5.13 shows $I(\theta)$ vs θ for a single slit, normalized to 1. The *principal maximum* occurs when θ approaches 0, and $(\sin\beta)/\beta$ becomes 1. The diffracted intensity is 0 at angles (except 0) for which $\sin\beta = 0$. The first such zero occurs at angle

$$\boxed{\theta_1 = \lambda/b ,} \tag{5.60}$$

where θ is assumed small.

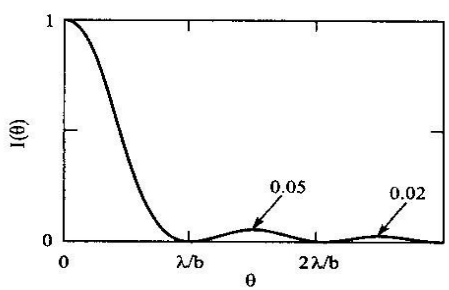

Fig. 5.13. Single-slit diffraction pattern

If the viewing screen is the focal plane of a lens, then the first minimum is located a distance

$$RL = \lambda f'/b \tag{5.61}$$

from the center of the pattern, which extends in the direction perpendicular to the edges of the aperture. Over 80 % of the diffracted light falls within $2\lambda f'/b$ of the center of the pattern, and the first *secondary maximum* is only about 5 % as intense as the principal maximum.

Similar analysis can be carried out with a circular aperture in two dimensions. The result is similar, except that the pattern is a disk, known as the Airy disk, with radius defined by the first zero as

$$RL = 1.22\lambda f'/D , \tag{5.62}$$

where D is the diameter of the aperture. It is the finite size of the Airy disk that limits the theoretical resolving power of any optical system.

Example 5.2. Calculate the Fraunhofer-diffraction pattern of a slit whose center is located a distance s_0 away from the axis of the system. Show that the result is identical with (5.59a) multiplied by a complex-exponential function, $\exp(-iks_0 \sin\theta)$. Show further that the intensity is identical with (5.59b) and is centered about the angle $\theta = 0$.

This result applies only to Fraunhofer diffraction and therefore presumes that $s_0 \ll L, r$. The argument $ks_0 \sin\theta$ of the complex exponential function is a *phase factor* that results from the shift of the aperture.

5.5.2 Interference by Finite Slits

Earlier, we noted that division-of-wavefront interference occurs because light is diffracted by the individual apertures. This implies, for example, that the interference pattern will vanish in those directions in which the diffracted intensity is 0. Similarly, the pattern will be strongest in those directions where the diffracted intensity is greatest. If the slits are identical, this implies that the diffraction pattern with finite slits should be given by

$$\boxed{\text{(interference pattern)} \times \text{(diffraction pattern of single slit)} ,} \tag{5.63}$$

where "interference pattern" refers to the pattern derived with infinitesimal slits. It is possible to verify this relation by direct integration over an aperture consisting of several finite slits.

The significance is mainly for multiple-slit interference. As we shall see in Chap. 6, a *diffraction grating* may well have slits whose widths are about equal to their spacing. Figure 5.14 shows the diffraction pattern in such a case. The dashed line is the diffraction pattern of a single slit, and the various orders of interference are indicated as peaks. Zero-order diffraction is of no interest, but the first and higher orders are weak because very little light is diffracted into their directions. Occasionally a principal maximum will fall so close to the diffraction minimum that it is barely detectable. In this case we speak of a *missing order*.

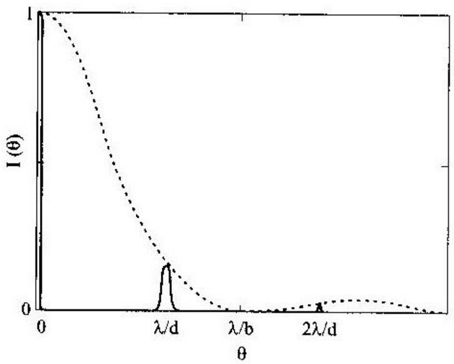

Fig. 5.14. Multiple-slit interference with finite slits