OPTICAL HOLOGRAPHY

Student Edition

Robert J. Collier
Christoph B. Burckhardt
Lawrence H. Lin
Bell Telephone Laboratories
Murray Hill, New Jersey
Chapter 8

ANALYSIS OF PLANE HOLOGRAMS

The fringes recorded by small, in-line holograms with nondiffuse light are coarsely spaced relative to the thickness of photographic emulsion. A ray in the wavefront illuminating such a hologram interacts with only one recorded fringe before passing through the hologram. Consequently the hologram response is closely approximated by that of a plane diffraction grating—with focusing properties. Gabor analyzed these properties by considering the hologram to be strictly two dimensional. Predictions of the analysis using the two-dimensional model were in good agreement with the experimental observations.

The off-axis technique, introduced by Leith and Upatnieks, led to holograms whose fringe frequency exceeded that of the in-line hologram by a term proportional to the angle between subject and reference beams [see Eq. (3.15)]. A typical value of the fringe spacing is obtained by considering the interference of two plane waves. Equation (1.10) \(2d \sin \theta = \lambda\), relates the half-angle \(\theta\) between beam directions and the wavelength \(\lambda\) to the fringe spacing \(d\). For \(\theta = 15^\circ\) and \(\lambda = 0.5 \mu m\) (green light), \(d = 1 \mu m\). Photographic emulsions used to record off-axis holograms are often 15 \(\mu m\) in thickness, and the hologram formed therein can no longer be realistically regarded as two dimensional. Nevertheless, a two-dimensional analysis, restated in terms of communication theory, was extended to the off-axis case by Leith and Upatnieks [8.1, 8.2]. Despite the discrepancy between the facts of photographic emulsion and the assumptions of the planar model, it provided a useful framework for the further development of holography. However its application to holograms which might be better described as volume diffraction gratings leads to partially fulfilled predictions and leaves many of the observed properties of holograms unexplained.
It is therefore important to remember that the results of plane hologram analysis are applicable with accuracy only to holograms formed on suitably thin media. A good example is thermoplastic film material whose thickness may be made comparable to the wavelength of light. The observed properties of holograms formed in thermoplastic are correctly predicted by planar analysis.

With analytic tools developed from the diffraction theory of Chapters 5 and 6, we can now treat those properties of plane holograms which are not revealed by the geometric analysis of Chapter 3. Diffraction theory allows us to discuss the Fourier transform hologram. We shall derive the condition for separation of the image-forming diffracted waves of off-axis holograms, consider factors affecting the quality of the image, and compute maximum values of diffraction efficiency for absorption and phase holograms.

8.1 Off-Axis Holography with Nondiffuse Subject Light

Hologram formation with a reference wave directed to interfere with a subject wave at some angle was described in Section 2.6 as the most successful method for separating the twin images. Analysis of the method in terms of spatial frequencies leads to the concept of a carrier wave (the reference) whose spatial frequency is modulated by subject information. Hence the name carrier-frequency hologram is an appropriate and descriptive alternate to off-axis hologram. The carrier-frequency method eliminates the need to derive a strong reference wave from the subject transmission. As a consequence Gabor’s restriction to transmission subjects with large transparent areas is no longer necessary.

Figure 8.1 represents a simple wavefront division method of illuminating a transmission subject with a coherent plane wave and deriving an off-axis plane reference wave from the same source. As subject we might choose a continuous-tone transparency. Suppose that the complex amplitude of the subject wave at the plane of the hologram is \( a(x, y) \) while that of the reference plane wave is \( r = r \exp(2\pi i \xi_t x) \). The reference wave spatial frequency \( \xi_{\text{Ref}} = -\xi_t = -(\sin \theta)/\lambda \) corresponds to the downward direction of the reference wave propagation vector and to the angle \( \theta \) which it makes in the \( xz \) plane with the \( z \) axis. We assume, as in Section 1.8, that we form an absorption hologram and that the amplitude transmittance of the fully processed hologram exposed to the interference pattern produced by \( a(x, y) \) and \( r \) is

\[
I = t_0 - kI, \tag{8.1}
\]

where \( t_0 \) is the transmittance of the unexposed (but developed) plate, \( k \) is a constant, and \( I \) is the intensity of the interference pattern. The intensity \( I \), as in Eq. (1.15), is

\[
I = aa^* + rr^* + ar^* + a^*r
- \exp(2\pi i \xi_t x) + \exp(2\pi i \xi_t x). \tag{8.2}
\]

When the hologram is illuminated by the original reference wave, the complex amplitude just behind the hologram takes the form

\[
w(x, y) = rt \exp(2\pi i \xi_t x)
- k[aa^* r \exp(2\pi i \xi_t x)
+ \exp(2\pi i \xi_t x) + ar^* + a^*r \exp(4\pi i \xi_t x)]. \tag{8.3}
\]

8.1.1 Separation of the Diffracted Waves

In Section 1.8 we state without proof that a properly directed reference wave would separate the desired reconstructed wave from the remaining waves diffracted by the hologram. Figure 2.10 of Section 2.6 indicates geometrically that the mean angle between subject and reference beam must be sufficiently large if this is to be satisfactorily accomplished. We now employ a spatial frequency analysis of Eq. (8.3) in order to relate the condition for angular separation of the diffracted waves to the maximum spatial frequency of the subject transmittance [8.1]. The subject transparency
is considered to have a transmittance \( s(x, y) \) and a spectrum \( S(\xi, \eta) \) where \( s(x, y) \rightarrow S(\xi, \eta) \). Its spectrum \( S(\xi, \eta) \) extends from \(-\xi_{\text{max}}\) to \(+\xi_{\text{max}}\) and from \(-\eta_{\text{max}}\) to \(+\eta_{\text{max}}\). A possible spectral distribution over the \( \xi \eta \) plane is plotted in Fig. 8.2 as \( |S(\xi, \eta)| \). When the transparency is illuminated by a \( z \)-directed plane wave, the complex amplitude of the subject wave received at the hologram plane is \( a(x, y) \), and the associated spectrum is given by Eq. (5.26) as

\[
A(\xi, \eta) = a_S(\xi, \eta) \exp\left[-i(2\pi d/\lambda)(1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2}\right] \tag{8.4}
\]

where \( a_s \) is the constant amplitude of the plane wave incident on the transparency and \( d \) is the distance separating the transparency from the hologram. Note that the maximum extent or width of \( A(\xi, \eta) \) in the spatial frequency plane is determined by the extent over which the spectrum \( S(\xi, \eta) \) has nonzero values. Equation (8.3) contains not only \( a(x, y) \) but also its conjugate \( a^*(x, y) \) whose spectrum is

\[
A'(\xi, \eta) = A^*(-\xi, -\eta) = a_s S^*(-\xi, -\eta) \exp[+i(2\pi d/\lambda)(1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2}] \tag{8.5}
\]

where we have used the correspondence (4.26). We now employ Eqs. (8.4) and (8.5) to plot the absolute value of the spectrum of \( w(x, y) \) in Eq. (8.3).

The first term on the right in Eq. (8.3), \( t_{\text{fr}} \exp(2\pi i \xi_0 x) \), represents undiffracted light proceeding in the direction of the illuminating wave. For normal absorption (low efficiency) holograms this usually will be a very strong component. According to correspondence (4.30), its spectrum is the delta function centered at \((-\xi_0, 0)\). We may write the transform relation as

\[
t_{\text{fr}} \exp(2\pi i \xi_0 x) \Rightarrow t_{\text{fr}} \delta(\xi + \xi_0) \tag{8.6}
\]

and plot the spectrum in Fig. 8.3 as a large vertical arrow.

The second term in Eq. (8.3) (the first in the brackets), \(-kaar \times \exp(2\pi i \xi_0 x)\), transforms into frequency space, according to correspondences (4.18) and (4.21), as a shifted autocorrelation function of the spectrum \( A(\xi, \eta) \). We are primarily interested in the maximum extent of the autocorrelation function over the spatial frequency plane; this will be a major determinant of the spatial frequency assigned to the reference wave. For this purpose we recall that the correlation integral like the convolution integral represents a scanning of one function by another (see Section 4.3, Fig. 4.4). The range of variables for which the integral is nonzero is given by the sum of the widths of the scanning and scanned functions; in the case of autocorrelation the maximum extent is twice the width of the function being autocorrelated. If we neglect constant multipliers, we have for the transform of the second term in Eq. (8.3)

\[
a^* a \exp(2\pi i \xi_0 x) \Rightarrow A^*(\xi, \eta) \ast A(\xi + \xi_0, \eta) = \left[S^*(\xi, \eta) \exp\left[i \frac{2\pi d}{\lambda} (1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2}\right]\right] \ast \left[S(\xi + \xi_0, \eta) \exp\left[i \frac{2\pi d}{\lambda} (1 - \lambda^2 (\xi + \xi_0)^2 - \lambda^2 \eta^2)^{1/2}\right]\right] \tag{8.7}
\]

As noted earlier, the width of \( A(\xi, \eta) \) is determined by its amplitude distribution and therefore given by the width of \( S(\xi, \eta) \). Hence the width of the autocorrelation function of Eq. (8.7) is twice that of \( S(\xi, \eta) \) and extends over a range \( 4[\xi_{\text{max}} - (-\xi_{\text{max}})] = 4\xi_{\text{max}} \) in the \( \xi \) direction and similarly \( 4\eta_{\text{max}} \) in the \( \eta \) direction. Its center \((-\xi_0, 0)\) corresponds to that of \( S(\xi + \xi_0, \eta) \) [see Eq. (4.12)] and coincides with the spatial frequency of the illuminating beam. The symmetric spread of diffracted light about the illumination beam direction is sometimes called intermodulation, referring to the modulation of light from one part of the subject by light from another.

Returning to Eq. (8.3) once again we find that the second term in the brackets \(-kr^2 \exp(2\pi i \xi_0 x)\) is similar to the initial term in the equation, an exponential which transforms to a delta function at \((\xi = -\xi_0, 0)\). It generally has a lesser strength than that corresponding to the initial term. All three terms which we have thus far discussed are called zero-order
terms because they represent light waves emerging from the hologram with their mean directions undeviated from the propagation axis of the illuminating wave.

The third term in the bracketed portion of Eq. (8.3) is proportional to the original wave $a(x, y)$ which arrived at the hologram from the subject. The absolute value of its spectrum $|A(\xi, \eta)|$ is proportional to $|S(\xi, \eta)|$ as is evident from Eq. (8.4). We assume that $|S(\xi, \eta)|$ is a symmetric distribution about a center spatial frequency $\xi_0 = 0$, $\eta = 0$ with limits $\pm \xi_{\text{max}}$ and $\pm \eta_{\text{max}}$ (see Fig. 8.2), and therefore the same distribution represents $a(x, y)$ in the spectral plot of Fig. 8.3.

The final term in Eq. (8.3), $-k \alpha r^2 \exp(4\pi i \xi r)$ represents the conjugate to the complex amplitude of the subject wavefront at the hologram, modulated onto a high carrier frequency. (Note that the corresponding diffracted wave is not antiparallel to the original subject wave and not its conjugate as defined in Section 1.8.) It has a spectrum given by the correspondence

$-k \alpha r^2 \exp(4\pi i \xi r) \Rightarrow -kr^2 A^*[-(\xi + 2\xi_r), -\eta].$

Its absolute value according to Eq. (8.5) is proportional to $|S^*[-(\xi + 2\xi_r), -\eta]|$, a distribution similar to that shown in Fig. 8.2 but inverted and shifted in the $-\xi$ direction by $2\xi_r$, and its limits are $\xi = -2\xi_r \pm \xi_{\text{max}}$ and $\eta = \pm \eta_{\text{max}}$.

In Fig. 8.3 are plotted the absolute values of the spectra of the waves emerging from the back side of the hologram. It is clear that the employment of a reference wave with a suitably high spatial frequency or large angle $\theta$ insures angular separation of the image-forming waves. To avoid angular overlap of the image-forming waves by the zero-order waves, Fig. 8.3 indicates that the reference spatial frequency $\xi_\text{Ref}$ must satisfy

$|\xi_\text{Ref} - \xi_0| = | -\xi_r - \xi_0 | = \xi_r + \xi_0 \geq 3\xi_{\text{max}}$ (8.8)

where $\xi_0$ is the center spatial frequency of the subject spectrum (assumed to be 0). To obtain even the minimum condition for angular separation $|\xi_\text{Ref} - \xi_0| = 3\xi_{\text{max}}$, however, requires a recording medium with high resolution in the $x$ direction as may be seen by reconsidering the effective hologram transmittance in Eq. (8.2). Let us substitute for the complex amplitude of the subject light at the hologram, $a(x, y)$, its highest spatial frequency component in the $+x$ direction

$A(\xi_0 + \xi_{\text{max}}) \exp[-2\pi i (\xi_0 + \xi_{\text{max}}) x].$

Equation (8.2) can then be written

$t(x) \propto I = A^2 + r^2 + A(\xi_0 + \xi_{\text{max}}) \exp[-2\pi i (\xi_0 + \xi_{\text{max}} + \xi_r) x]$

$+ A(\xi_0 + \xi_{\text{max}}) \exp[2\pi i (\xi_0 + \xi_{\text{max}} + \xi_r) x]$

$= I_0 + \text{(constant)} \cos[2\pi i (\xi_0 + \xi_{\text{max}} + \xi_r) x].$

In the argument of the cosine is the fringe frequency $\xi_r + \xi_0 + \xi_{\text{max}}$, which the medium must record. Taken with the condition $\xi_r + \xi_0 + 3\xi_{\text{max}}$ from Eq. (8.8), this amounts to a spatial frequency of $4\xi_{\text{max}}$, four times the highest spatial frequency in the subject. The high resolution needed in the recording medium is the price of twin image separation by means of the off-axis method. In terms of image quality and of flexibility in the choice of subject, the benefits outweigh this disadvantage, especially when high resolution photographic emulsion is available. Furthermore, some overlap of intermodulation diffraction can be tolerated since the amplitude of this term falls off rapidly away from its center frequency (see Fig. 8.3). It is possible to further reduce the effect of the overlap by making the amplitude of the reference wave large in comparison to that of the subject wave. The first term in the brackets of Eq. (8.3) is then small compared to the third and fourth terms.

Although Eq. (8.8) is the condition for no overlap in the spatial frequency domain (no angular overlap), it does not guarantee that in the spatial plane of the image formed by one of the diffracted waves there will be no unwanted light from the other waves. This is evident from Fig. 8.4 which shows...
the formation of a real image in a plane relatively close to the hologram plane (Position 1). Figure 8.4 indicates the cones of light emitted by two points of the illuminated hologram. Although there is angular separation of the diffracted waves, the site of the real image in Position 1 is illuminated by undesirable light from the zero-order waves as well as by the desired real image wave. One solution, the usual one, is to choose the reference wave spatial frequency sufficiently high and the distance of the subject from the hologram sufficiently large that there is no overlap in the image plane (Position 2).

A second, more complicated solution is to filter out the undesired spatial frequency components. This may be done by using a lens to display in its back focal plane the frequency spectrum of the complex amplitude just behind the hologram and then blocking out all but the useful portion of the spectrum. The complexity of the method relegates its use to situations where the resolution capacity of the recording medium is very low. As an example, let us suppose we wish to form a hologram on Kodak Tri-X pan-chromatic film of a subject transparency whose center spatial frequency is $\xi_0 = 0$ and whose maximum spatial frequency is $\xi_{\text{max}} = \eta_{\text{max}} = 20$ cycles/mm (corresponding to subject detail 0.025 mm in dimension). If we wish to record an off-axis hologram as described in this section, the absolute value of minimum spatial frequency of the reference wave required for complete angular separation of the diffracted waves is $|\xi_{\text{Ref}}| = 3\xi_{\text{max}} = 60$ cycles/mm, and the highest fringe frequency to be recorded is $4\xi_{\text{max}} = 80$ cycles/mm. This is just within the resolving power of the emulsion. Although further increase in the angular separation of the waves can be achieved by increasing the reference-to-subject beam angle, i.e., increasing $|\xi_{\text{Ref}}|$, the resolution limit of the emulsion will then be exceeded. The angle in the arrangement of Fig. 8.1 corresponding to 60 cycles/mm is given by

$$\sin \theta = \theta = \lambda \xi_{\text{Ref}} = (0.633 \times 10^{-3})(6 \times 10^2) = 3.79 \times 10^{-2} \text{ rad} = 2.16^\circ,$$

where $\lambda = 0.633 \times 10^{-3}$ mm is the wavelength of radiation from the He–Ne laser. Because the angle is so small, the reference beam must be brought in by means of a beam splitter placed between subject transparency and hologram. A possible arrangement for forming the hologram is shown in Fig. 8.5. Since $\theta$ is only 2.16° the diffracted waves will certainly overlap in the plane of the real image when the lateral dimensions of the subject transparency are comparable to its distance from the hologram. Spatial frequency filtering is required in this case. Methods involving total reflection of unwanted waves have been suggested [8.3, 8.4].

It should be noted in passing that when a number of holograms are superimposed on the same photographic plate, the proper choice of the reference waves employed in their formation can produce an angular separation of the reconstructed image-forming waves. Suppose we consider two superimposed holograms. Each is formed with a subject located on the z-axis and illuminated by an axial plane wave. Each subject has a mean
spatial frequency $\xi_n = 0$, $\eta_n = 0$ and a spatial frequency bandwidth $\pm \xi_{\text{max}}$, $\pm \eta_{\text{max}}$. One hologram is formed with a plane reference wave of spatial frequency $\xi_{\text{Ref}} = -\xi_x - 3\xi_{\text{max}}$ and $\eta_{\text{Ref}} = 0$; the other is formed with a plane reference wave of spatial frequency $\xi_{\text{Ref}} = -\xi_x - 2\xi_{\text{max}}$ and $\eta_{\text{Ref}} = 0$. Based on analysis similar to that leading to Fig. 8.3, we find that the spectral response of each hologram taken individually and illuminated by its respective reference wave is as shown in Fig. 8.6a,b. When the two superimposed holograms are illuminated by a plane wave of spatial frequency $\xi = -\xi_x$, the spectrum is as given by Fig. 8.6c. The image-forming waves are angularly separated, and the condition for achieving this separation is

$$\Delta \xi_{\text{Ref}} \geq 2\xi_{\text{max}} \quad (8.9)$$

where $\Delta \xi_{\text{Ref}}$ is the difference in the reference-wave spatial frequencies and $2\xi_{\text{max}}$ is the spatial frequency bandwidth of the subject in the $x$ direction. Choice of the illuminating wave spatial frequency is arbitrary.

### 8.1.2 Generation of the Real Image

If one illuminates a transparency with a plane wave, the diffraction pattern in the near field is a projection of the subject transparency. Each small area of the hologram which records such a pattern contains information about only a small part of the subject. Suppose the hologram formed as in Fig. 8.1 is now illuminated by the original reference plane wave. Figure 8.7 shows the illuminating and reconstructed waves as well as the on-axis virtual image. (A simple subject transparency consisting of three small transparent holes in an opaque background has been assumed.) Because the subject is neither illuminated by diffuse light nor does itself diffuse the illuminating plane wave, the observer receives light from only one of the virtual image spots at a time. He can learn about all three spots by moving his head so as to scan the image, a procedure tiresome and impractical. It is therefore preferable in the case of nondiffuse subject light to view the projected real image. This may be done by allowing the image to form on a ground-glass diffusing screen placed in the real-image plane. Diffusion by the screen permits viewing the entire image from a single position. (A less convenient viewing method is to convert the virtual image to a real image by means of a lens.)

When the hologram is illuminated by the original reference wave, as in Fig. 8.7, the virtual image appears on-axis in the position of the original subject, and it is free of spherical and off-axis aberrations (see Section 3.4, [3.2]). The real image, formed off axis, is not. Let us now analyze two methods of illuminating the hologram to obtain aberration-free, on-axis real images. We assume the hologram has been formed with the arrangement of Fig. 8.1, where the subject is centered on the $z$ axis and where the reference plane wave propagates at an angle $-\theta$ to the $z$ axis. In the first method the hologram is illuminated by a plane wave proceeding in a direction $+\theta$ to the $z$ axis, as shown in Fig. 8.8. Illumination of the hologram in this manner corresponds to multiplying the hologram transmittance, as expressed in Eq. (8.2), by $r \exp(-2\pi n\xi x)$ where $\xi = (\sin \theta)/\lambda$. Of all the
waves diffracted, we shall be concerned only with the real-image forming wave at the hologram plane, $a^*r'$. This wave has a complex amplitude conjugate to that of the original subject wave at the same hologram plane. Its properties are best analyzed in the spatial frequency plane. The frequency spectrum of $a^*(x,y)$ is given in Eq. (8.5) as

$$A'(\xi, \eta) = A^*(-\xi, -\eta) = a_sS^*(-\xi, -\eta) \exp \left[ + \frac{2\pi d}{\lambda} (1 - \lambda^2\xi^2 - \lambda^2\eta^2)^{1/2} \right].$$

If we take the inverse Fourier transform of both sides of Eq. (8.10) and use correspondence (4.26) we have that the complex amplitude in a plane a distance $d$ from the hologram is proportional to the complex conjugate of the transmittance of the original subject transparency $s^*(x,y)$. The intensity at $d$ is then proportional to that in the original subject plane. Thus the hologram generates an image of the original transparency which is real (the image-forming wave propagates to the image) and which is located on the $z$ axis. The image plane lies to the right of the hologram the same distance $d$ as the original subject had been to the left. A photographic plate placed in the image plane at $d$ will record the image without the need for a lens.

A second method for generating a real image is indicated in Fig. 8.9. Note that the illuminating beam is incident on the right side of the hologram (whereas the original reference illuminated the left surface) and is antiparallel to the original reference. It is therefore the conjugate to the original reference beam. However, in the plane of the hologram the complex amplitude

---

**Fig. 8.7.** Observation of the virtual image of a subject illuminated with nondiffuse light.

**Fig. 8.8.** Generation of an on-axis real image.

**Fig. 8.9.** Alternate method for generating an on-axis real image.
of the illuminating wavefront is \( r \exp(-i2\pi \xi_x) \), identical to that of the illuminating beam shown in Fig. 8.8. In contrast to the latter beam, the illuminating beam in Fig. 8.9 travels from right to left. Since the complex amplitude in the plane of the hologram is the same for both arrangements, our previous discussion applies to Fig. 8.9. The only difference is that the real image now forms on axis a distance \( d \) to the left of the hologram, consistent with the direction of propagation of the illuminating beam, and the image coincides with the original subject location. To accomplish this, the image-forming wave must be antiparallel to the original subject and therefore must be its conjugate.

The arrangements of Figs. 8.8 and 8.9 for generating the real image are equally good providing the hologram is thin. For thick holograms the arrangement of Fig. 8.9 is superior, since Bragg’s law, the criterion for significant diffraction from a volume grating, is satisfied by the conjugate to the original reference wave as well as the original (see Section 1.6). When the recording medium cannot be characterized as strictly plane, Figs. 8.7 and 8.9 indicate the preferred arrangement for generating the virtual and real images respectively.

8.1.3 COHERENCE REQUIREMENTS FOR OFF-AXIS HOLOGRAMS

The coherence length required of a laser which is used, as in Fig. 8.1, to form an off-axis hologram is greater than required of a laser used to form an in-line hologram. We can relate the necessary coherence length to the hologram-formation geometry by first considering the interference of two plane waves indicated in Fig. 8.10. The (unmodulated) signal beam is axial and the reference beam makes an angle \( \theta \) with it. Hence the paths from source to photographic plate taken by all rays in the signal are equal while those for the reference beam rays differ by as much as \( l_1 \). We may set reference and signal paths equal for the central ray of the reference beam in which case the maximum signal-to-reference beam path difference is given by

\[
\frac{l_1}{2} = \frac{a}{2} \sin |\theta| = \frac{a}{2} \lambda \xi, \tag{8.11}
\]

where \( a \) is the length of the plate, \( |\theta| \) is the magnitude of the angle the reference beam makes with the \( z \) axis, and \( \xi \) is the corresponding reference wave spatial frequency. Thus, even with no information in the signal beam, the coherence length must at least equal \( l_1/2 \).

Next consider the added path length in the signal beam due to diffraction by an information-bearing transparency as in Fig. 8.11. The longest subject-

![Fig. 8.10. Maximum path difference \( l_1 \) at the hologram plane for rays of an off-axis plane reference wave.](image)

![Fig. 8.11. Maximum path difference \( l_s \) for a signal beam with subject information.](image)

ray path differs from the unmodulated signal-beam path by

\[
l_s = \frac{d}{\cos \theta_{\text{max}}} - d = \frac{d}{2} \theta_{\text{max}} \approx \frac{d}{2} \lambda \xi_{\text{max}} \tag{8.12}
\]

for small values of \( \theta_{\text{max}} \). Here \( d \) is the distance separating transparency and photographic plate, \( \theta_{\text{max}} \) is the maximum angle through which the plane wave incident on the transparency is diffracted, and \( \xi_{\text{max}} \) is the maximum spatial frequency of the transparency. Thus, if the path length of the central ray of the reference beam is set equal to that of the central ray of the signal
beam, the total requirement on the coherence length of the laser light is

$$\Delta L_H > \frac{L_1}{2} + L_2 = \frac{a}{2} \lambda_{\xi_r} + \frac{d}{2} \lambda_{\xi_{\text{max}}}.$$ (8.13)

In the Gabor in-line arrangement $\xi_r = 0$ and the required coherence length $L_H$ need only equal or exceed $L_2$. Several techniques exist for forming off-axis holograms in a manner requiring this lesser coherence length. The geometries are specialized, however, and intended for use only where laser light is not available [8.5, 8.6] or where holograms are formed with multimode pulsed lasers (see Section 11.6.1).

8.2 Off-Axis Holography with a Diffuse Signal

Difficulty in observing the virtual image has already been cited as one disadvantage of illuminating subject transparencies with a uniform wave. Another is of concern even when viewing the real image. A speck of dust or imperfection on the surfaces of optical components used to expand the subject-illuminating beam (Section 7.3) can give rise on reconstruction to a pattern of circular rings located on the image. Figure 8.12 illustrates this ring structure. It is a zone plate pattern caused by interference of spherical waves (scattered by the speck) with the unperturbed subject illumination and projected onto the subject.

A third disadvantage is that the intensity of the light transmitted to the hologram by the plane-wave-illuminated subject transparency varies over a large range determined by the transmittance variation of the subject. Again this is a consequence of the projection character of the signal wave. If linear holographic recording is desirable, then the intensity of the reference beam must exceed that of the signal over the entire hologram (see Section 7.2.2). Where the signal is weak, the beam ratio is too high, and the diffraction efficiency is consequently poor.

These disadvantages can be avoided by using diffuse light to illuminate the subject transparency [8.7]. This is usually accomplished by inserting a diffusing screen such as a sheet of ground glass between the illuminating laser source and the transparency. Since the diffusing screen scatters light over a wide angular range, the transparency can now be viewed comfortably from a single position of the viewer's head. The same is true for viewing the virtual image generated by the reconstructed subject wave, at least over the angular range of subject rays received and recorded by the hologram.

![Figure 8.12](image-url)  
**Fig. 8.12.** Ringlike noise in the real image of a nondiffusely illuminated transparency. (Courtesy Leith and Upatnieks [8.2].)

Although the phase of the diffuse subject light is a rapidly varying spatial function of the hologram coordinates, the light at the hologram plane can still be coherent. For this to be so, (1) the initial wave illuminating the diffusing screen must be spatially coherent over the full area of the screen, (2) the maximum path length from source to hologram via the diffusing screen must not differ from that via the reference beam by more than the coherence length, and (3) the screen must be motionless. An arrangement for forming a hologram of a diffusely illuminated transparency is shown schematically in Fig. 8.13a and the viewing of the virtual image is indicated in Fig. 8.13b.
8. Analysis of Plane Holograms

8.2 Off-Axis Holography with a Diffuse Signal

The hologram of a diffusely illuminated subject has a number of unusual properties. Since the diffusing screen has a much broader spectrum than the transparency, it scatters light over a wide angular range, and every point on the hologram plate receives light from every point of the transparency. In the reconstruction step the entire virtual image of the subject can be observed by looking through any part of the hologram. As the viewing position is changed the image is seen from a different aspect. If the subject is a two-dimensional transparency and one wishes to display the real image on some observation plane, a complete image can be obtained even if the hologram is broken or partially destroyed, with but a fragment remaining. Of course, the resolution in the image decreases (as with any lens of finite size) as the fragment becomes smaller. Figure 8.14 displays photographs of three real images obtained from the same hologram illuminated by a laser beam of decreasing diameter. This property of the diffuse light hologram to record information in a nonlocalized way makes it attractive for information storage. In contrast to storing information by means of microimages where a scratch or blemish entirely removes or obliterates information, information stored in a hologram made with diffuse subject light is relatively immune to such imperfections of the recording medium.

Fig. 8.14. Photographs of real images of a diffusely illuminated subject obtained by illuminating the hologram with a beam of decreasing diameter. The beam diameter from left to right was 0.81 cm, 0.26 cm, and 0.08 cm.

Diffuse illumination places an increased demand on the resolution capability of the recording medium. The maximum fringe frequency which must be recorded by a hologram is determined by the maximum angle which light rays from the subject make with the reference beam direction. Figure 8.15a illustrates an angle $\psi$ formed with a plane reference wave and nondiffused light passing through a transparency. When a diffusing screen is interposed between illuminating source and transparency, a substantial increase in $\psi$ results (Fig. 8.15b).

8.2.1 Three-Dimensional Images

Most three-dimensional objects reflect light more or less diffusely so that holograms made with such subjects exhibit the properties we have just discussed. Apart from these are properties related to the three-dimensional
8. Analysis of Plane Holograms

FIG. 8.15. Maximum angles between subject and reference rays reaching the hologram in the case (a) where the subject is illuminated nondiffusely and (b) where the subject is diffusely illuminated.

nature of the subject. Since the hologram can reconstruct a wave which is an exact replica of the original subject wave, the virtual image from which the reconstructed wave appears to diverge contains all the depth and parallax properties of the subject. Each eye of the observer views the virtual image through a different area of the hologram and thus sees the image from a different aspect (Fig. 8.16). These aspects, moreover, are identical to those seen by an observer of the original subject looking through an aperture defined by the extent of the hologram. Both the observer of the image and the observer of the subject itself have the same perception of depth. Should the observer move his head the aspects change, and, as with the original subject, he perceives parallax. Figure 8.17 shows two photographs of the same virtual image taken from different angles to exhibit the parallax. Of course, the 3D nature of the image can be fully appreciated only by actually viewing through the hologram.

The real image of a three-dimensional object generated by a hologram usually has the curious property that its depth is inverted. We say that the image is pseudoscopic. Consider the formation of a hologram with a simple three-dimensional subject consisting of two separated point sources (Fig. 8.18a) and the subsequent generation of a real image by illuminating the hologram with the conjugate to the reference wave (Fig. 8.18b). In the formation step the point \( P_1 \) is closer to the hologram plane than the point \( P_2 \).
The reconstruction step causes real images of $P_1$ and $P_2$ to form at their original locations (as discussed in Section 8.1.2). However, the observer of the real image must be positioned as in Fig. 8.18b if he is to receive the light diffracted by the hologram from right to left. To him $P_1$ is the closer point. Thus he observes an inverted depth relation relative to his normal view of the original sources, where light propagates from left to right through the photographic plate. This depth inversion leads to a number of confusing visual clues. Suppose that $P_1$ and $P_2$ are points on the surface of some three-dimensional subject and that there is a range of viewing angles over which the surface in the vicinity of $P_1$ obstructs the view of $P_2$. Over this angular range the hologram records information only about the surface in the vicinity of $P_1$ and records no information about $P_2$. Upon reconstruction,
202

8. Analysis of Plane Holograms

the observer of the real image who views from the corresponding angles, but from the other side of the hologram, sees only the area about \( P_1 \). It appears to him that a point \( P_2 \) in the front of the image is obscured by an area in the rear of the image, a sensation contrary to normal experience. Competition between what he sees in the image and what he remembers to be true of the object produces conflict in the mind of the observer and makes viewing of pseudoscopic images unsatisfactory.

A real image generated by a hologram need not be pseudoscopic. Suppose a lens is employed to form a real image of a laser-illuminated subject and a photographic plate is inserted into the converging wave as in Fig. 8.19.

![Fig. 8.19. Formation of a hologram which gives an orthoscopic real image.](image)

A suitable reference is added and a hologram of the converging wave is formed. When the hologram is illuminated with the original reference wave, the original subject wave is reconstructed. The reconstructed wave converges to a real image to the right of the hologram just as the original wave, and the image is therefore normal in its depth properties or orthoscopic. On the other hand, if the conjugate to the reference wave traveling from right to left illuminates the hologram, the conjugate to the original subject wave will appear to diverge from the original subject location. That is, a virtual image of the image produced by the lens is generated at the original image plane of the lens. Looking through the hologram from the left, points formerly in the rear of the image produced by the lens now appear closer to the observer, so that it is the virtual image which is pseudoscopic [8.8].

If, with holography, one forms a pseudoscopic image of an image which is already pseudoscopic relative to some original subject, then the resulting holographic image is orthoscopic with respect to that subject. Hence if the subject wave for a second hologram is the pseudoscopic image-forming wave emerging from a first hologram, then illumination of the second hologram with the conjugate to its original reference will yield an orthoscopic image [8.9]. The same results can be obtained with autocollimating devices which do not require a second recording to convert pseudoscopic images into orthoscopic ones and vice versa [8.10].

8.2 Off-Axis Holography with a Diffuse Signal

8.2.2 Speckle Pattern

Since the images generated by holograms made with diffuse-light signals are replicas of the original laser-illuminated subjects, they suffer from a problem common to laser-illuminated, diffusely scattering objects. An observer of either image or object sees an annoying, granular speckle pattern which gives the observed surface a spangled or scintillating appearance. The grain size depends on the smallest aperture of the recording or viewing instrument (e.g., the iris of the eye) and on its distance from the observed surface [8.11]. In any area of the surface smaller than can be resolved by the observer there are many scattering centers which impose random phases on the complex amplitude of the light they scatter. Because the light is coherent and the relative phases constant in time, these amplitude components add. The observer, unable to resolve the individual scattering centers, sees the light emitted from the small area as a small spot of uniform intensity whose value depends on the net phase resulting from the amplitude addition. Even when the surface is uniformly illuminated, the observed reflected intensity will seem to vary randomly from spot to spot, since at each spot the intensity depends on the net amplitude resulting from the coherent addition of the randomly phased, scattered amplitudes. Motion of the observer so as to view the surface from a new aspect requires that the light rays traveling from surface to viewer proceed along a new path. The phases of the net amplitude of light from the minimum resolvable spots under observation change accordingly and, along with these, the intensity also changes. A continuous motion of the head tends to average out the intensity variation contributing to the speckle pattern and thereby tends to improve perception. However this is not very convenient when examining detail, and the presence of speckle pattern must be regarded as detrimental both to perception and viewing comfort. (For a more detailed discussion of speckle see Chapter 12.)
8. Analysis of Plane Holograms

8.3 Hologram-Forming Geometries

8.3.1 Fresnel Holograms

When the photosensitive medium or plate for recording a hologram is placed in the near-field or Fresnel diffraction region of the subject and at an arbitrary distance from the reference source, the record formed is called a Fresnel hologram. It is a natural way to form a hologram since no lenses or other imaging devices are required either in formation or in the reconstruction. Except for aspects of the lenless Fourier transform hologram discussed in Chapter 3, all of the hologram properties discussed thus far are those exhibited by the Fresnel hologram. It is the common hologram formed with the arrangements indicated in Figs. 7.1, 7.19, and 7.20. The hologram itself appears as in Fig. 7.21; it can be illuminated with the simple apparatus of Fig. 7.22 and is generally used to produce images of 3D objects such as those in Fig. 7.23. Figure 2.4 of Section 2.2 shows an early Fresnel hologram-forming arrangement with an on-axis reference wave.

8.3.2 Image Holograms

When special geometries and sometimes lenses or other imaging devices are used in the formation of the hologram, a number of useful properties are realized. Suppose that the photographic plate in Fig. 8.19 is moved until it lies in the central plane of the image formed by the lens, as in Fig. 8.20. Addition of the reference beam then allows formation of an Image hologram \[8.12, 8.13\]. On reconstruction with the original reference wave, part of the image generated by the hologram is virtual and part real. To the observer there is no marked difference between this image and those generated from lensless Fresnel holograms. However the angle from which the image can be viewed is limited by the aperture of the lens and the 3D image appears centered about the hologram plate. The technique has the merit of minimizing the requirements on the hologram-illuminating source. This advantage was pointed out in Section 7.6 where the distance \(z_i\) of the image from the hologram was related to the requirements for spatial and temporal coherence of the source. Equation (7.52),

\[
\Delta s = \left(\frac{z_i}{z_e}\right) \Delta r,
\]

describes how the degree of spatial coherence, represented by the extent of the source \(\Delta r\) and the distance of the source from the hologram \(z_e\), determine the resolution of the image \(\Delta s\). As \(z_i \to 0\) (the case of the image hologram) even large sources with little spatial coherence can be used to generate hologram images with sufficient resolution. The consequence is...
that image holograms can be brightly illuminated by extended sources. Of course, \( z_i = 0 \) only for a single plane of the subject, and image resolution on either side of the plane will degrade when a broad source is used (Fig. 8.21).

Equation (7.53) relates the resolution \( \Delta \sigma \) in the image generated by the illuminated hologram to the bandwidth \( \Delta \lambda \) of the illuminating source,

\[
\Delta \sigma = \theta_0 z_i (\Delta \lambda / \lambda)
\]

where \( z_i \) again is the distance between hologram and image, \( \lambda \) is the center wavelength, and \( \theta_0 \) is the angle the reference beam makes with the normal to the hologram plate (the subject and image are considered to be on-axis). We see that when both \( \theta_0 \) and \( z_i \) are small, the illuminating source bandwidth may be large and yet have small effect on the minimum resolvable image spot. One can even use a white-light source; the central plane of the image, located on the hologram, will then appear achromatic. However image points out of this plane will exhibit color dispersion and blurring which degrade the resolution.

8.3.3 Fourier Transform Holograms

In the next three sections we shall distinguish between several methods of producing holograms which generate at the hologram plane wave amplitudes which are either the exact Fourier transform of the subject or the Fourier transform multiplied by a slowly varying phase factor. Common to all of the methods is the restriction that the reference source must effectively lie in the same (input) plane as the subject. As a consequence, the analysis is intended to apply strictly to planar subjects (e.g., transparencies) and is less applicable as the subject extends out of the input plane. We generally require the subject to be illuminated with a plane wave. In some of the hologram-forming arrangements to be discussed a lens is used. When the lens precedes the subject, it is the lens which is illuminated with the plane wave. When the lens follows the input plane, it is assumed to operate on light from both the subject and the reference source.

We call a Fourier transform hologram one which records the interference of two waves whose complex amplitudes at the hologram are the Fourier transforms of both the subject and reference source. As we shall see in Chapter 14, such holograms are employed as spatial filters for pattern recognition, and the properties of the Fourier transform provide the basis for the recognition process. There the transmittance of the reference source is spatially modulated (an extended source). Here we confine our analysis to point reference sources.

The Fourier transform of a two-dimensional subject can be displayed in the back focal plane of a lens as discussed in Section 6.3.3 (see Fig. 6.6). An arrangement for forming Fourier transform holograms in the manner of Vander Lugt [8.14] is shown in Fig. 8.22. If \( s(x, y) \) is the transmittance of the transparency in the front focal plane of the lens, the subject amplitude at the hologram located in the back focal plane is \( S(\xi, \eta) \) where \( s(x, y) \rightarrow S(\xi, \eta) \). Also located in the front focal plane is a point source \( \delta(x+b, y) \) whose transform, a plane wave amplitude given by \( \exp(-2\pi i \xi b) \), acts as the reference wave and illuminates the back focal plane along with \( S(\xi, \eta) \). The intensity of the interference pattern formed by the two transforms is

\[
I = [\exp(-2\pi i \xi b) + S(\xi, \eta)][\exp(2\pi i \xi b) + S^*(\xi, \eta)]
\]

\[
= 1 + |S(\xi, \eta)|^2 + S(\xi, \eta) \exp(2\pi i \xi b) + S^*(\xi, \eta) \exp(-2\pi i \xi b). \quad (8.14)
\]

We assume the developed hologram has a transmittance \( t(x, y) \propto I \). If the hologram is illuminated with a plane wave propagating along the \( z \) axis with constant amplitude \( r_0 \), the product \( r_0 t(x, y) \) represents the complex
amplitude $W$ of the diffracted light just behind the hologram, where

$$W \propto r_0(x, y) \propto I = 1 + |S|^2 + S \exp(2ni\xi b) + S^* \exp(-2ni\xi b). \quad (8.15)$$

A lens placed immediately before or after the hologram (Fig. 8.23) will display in its back focal plane the product of the inverse Fourier transform of $W$ and a spherical phase factor (see Section 6.2). If we detect only the intensity in the back focal plane then we can neglect the spherical phase factor. As indicated in Fig. 8.23, the zero-order terms in Eq. (8.15) will be focused about the origin of that plane. The inverse Fourier transform of the third term on the right of Eq. (8.15), $S(x - b, y)$, is the original transmittance shifted $b$ units from the origin in the positive $x$ direction while the transform of the fourth term yields $S^*(x + b, -y)$, the conjugate of the original transmittance, inverted and shifted $b$ units from the origin in the negative $x$ direction. In each case the diffracted light converges to a real image formed in a common plane. A photograph of the output of a Fourier transform hologram displayed in the back focal plane of the reconstructing lens is shown in Fig. 8.24. Since only intensity is detected by the photographic film, the conjugate image differs from the other image only by its inversion.

![Fig. 8.23. Generation of two real images from a Fourier transform hologram.](image)

A useful property of the Fourier transform hologram formed with a plane reference wave is that the generated image remains stationary when the hologram is translated. This, for example, would allow holograms recorded on a reel of film to project stationary images while the film was moving. To demonstrate the invariance of image position to hologram translation, consider the displacement of the complex amplitude in the third term of Eq. (8.15) by $\xi_0$ units in the $+\xi$ direction so that it is now represented by

$$S(\xi - \xi_0, \eta) \exp[2ni(\xi - \xi_0)b].$$

When the hologram is illuminated as before and the corresponding diffracted wave is transformed, we then have for the complex amplitude at the back focal plane

$$\mathcal{F}^{-1}[S(\xi - \xi_0, \eta) \exp(2ni\xi_0b)] \exp(-2ni\xi_0b)$$

$$= [s(x, y) \exp(-2ni\xi_0x) * \delta(x - b)] \exp(-2ni\xi_0b)$$

$$= s(x - b, y) \exp(-2ni\xi_0x),$$

where $\mathcal{F}^{-1}$ signifies the inverse Fourier transformation and where we have used correspondences (4.11), (4.21), (4.29), and the consequences of convolving a function with a delta function. The phase factor $\exp(-2ni\xi_0x)$ drops out of the expression for intensity and the image intensity

$$s(x - b, y)s^*(x - b, y)$$

is identical to that observed with no hologram displacement.

Holograms of the Fourier transform of a transparency must record the large intensity variations present in the transform. Light passing through the transparency undeviated (zero-order) is focused by the lens to a high-intensity peak at the origin of the frequency or hologram plane. By comparison, the higher spatial frequencies diffracted by the transparency and focused to other regions of the frequency plane are much less intense. A reference beam intensity great enough to linearly record the low frequency
peak may be too great for the weak high-frequency intensities. Diffraction efficiency is then very poor for these high frequencies. When their complex amplitudes in the reconstruction are no greater than the amplitudes of light scattered by the hologram as noise, then subject information is lost.

8.3.4 QUASI-FOURIER TRANSFORM HOLOGRAMS

We designate a quasi-Fourier transform hologram as one formed under conditions where (1) the hologram plate is located in the back focal plane of a lens and (2) the subject transparency and reference point source are coplanar in a plane which is in front or in back of the lens but not in the front focal plane. Either the subject or the lens, whichever precedes, is illuminated with a plane wave.

A specific arrangement with the transparency and reference point source adjacent to the lens is shown in Fig. 8.25. We know (Section 6.2) that the subject wave amplitude at the hologram is the Fourier transform of the transparency times a spherical phase factor, and we inquire as to the effect of this factor. As in Eq. (6.23) the complex amplitude of the subject wave at the hologram plane can be represented by

\[
a(x_a, y_a) = \frac{ia}{\lambda f} \exp \left[-\frac{i\pi}{\lambda f} (x_a^2 + y_a^2)\right] S(\xi, \eta) \\
= e \exp \left[-\frac{i\pi}{\lambda f} (x_a^2 + y_a^2)\right] S(\xi, \eta)
\]  

(8.16)

Interference between \( r \) and \( a(x_a, y_a) \) from Eq. (8.16) produces a pattern whose intensity is

\[
l = r_0^2 + aa^* \\
+ crS(\xi, \eta) \exp \left[-\frac{i\pi}{\lambda f} (x_a^2 + y_a^2)\right] \exp \left[\frac{i\pi}{\lambda f} (x_a^2 + y_a^2 + 2x_b b)\right] \\
+ e^*r_0 S^*(\xi, \eta) \exp \left[\frac{i\pi}{\lambda f} (x_a^2 + y_a^2)\right] \exp \left[-\frac{i\pi}{\lambda f} (x_a^2 + y_a^2 + 2x_b b)\right] \\
= r_0^2 + aa^* + crS(\xi, \eta) \exp \left[i2\pi \frac{x_a}{f} \cdot b\right] + e^*r_0 S^*(\xi, \eta) \exp \left(-i2\pi \frac{x_a}{f} \cdot b\right) \\
= r_0^2 + |cS|^2 + crS(\xi, \eta) \exp \left(2\pi i \frac{b}{f}\right) + e^*r_0 S^*(\xi, \eta) \exp \left(-2\pi i \frac{b}{f}\right).
\]

(8.18)

where \( c = ia/\lambda f, \xi = x_0/\lambda f, \eta = y_0/\lambda f, S(\xi, \eta) = s(x_1, y_1) \), and \( \exp[-(i\pi/\lambda f) \cdot (x_a^2 + y_a^2)] \) is the spherical phase factor. Since \( \exp[-(i\pi/\lambda f) \cdot (x_a^2 + y_a^2)] \) may be recognized as the transmittance of a thin diverging lens of focal length \( -f \) [see Eq. (6.15)], \( a(x_a, y_a) \) may be considered to represent the combined transmittance formed by placing a transparency of transmittance \( S(\xi, \eta) \) adjacent to a diverging lens of focal length \( -f \). Similarly \( a^*(x_a, y_a) \) is equivalent to the combined transmittance formed by placing a transparency, whose transmittance is \( S^*(\xi, \eta) \), adjacent to a converging lens of focal length \( f \).

The hologram recording process will remove the spherical phase factor in \( a(x_a, y_a) \) if the reference point source is coplanar with the subject. Figure 8.25 shows a point source located \( -b \) units below the origin of the \( x_2y_2 \) plane on the \( x \) axis. It produces a spherical wave whose phase variation over the \( x_2y_2 \) plane, relative to an arbitrarily assigned phase of zero at \( x_2 = 0, y_2 = 0 \), has the form of Eq. (3.3), viz.,

\[
\phi(x_2, y_2) = \frac{\pi}{kz_1} (x_2^2 + y_2^2 + 2x_2b).
\]

(The thin lens with small curvature to its surface will contribute only a nearly constant phase factor to the light emitted by a point adjacent to it.) Recall that in Eq. (3.3) \( z_1 \) is a negative quantity. In the arrangement of Fig. 8.25 \( z_1 = -f \). The complex amplitude of the spherical reference wave at the hologram is therefore

\[
r = r_0 \exp \left[-\frac{i\pi}{\lambda f} (x_a^2 + y_a^2 + 2x_2 b)\right].
\]

(8.17)
This is equivalent to the intensity in Eq. (8.14) characterizing the Fourier transform hologram and hence the hologram produced with the arrangement of Fig. 8.25 has the properties of the Fourier transform hologram.

We can verify that these same properties are obtained for quasi-Fourier transform holograms formed such that the plane of the subject and reference source is not adjacent to the lens. Let us return to Chapter 6 and to Eq. (6.59) as an aid. There we have arriving at the (hologram) plane in the back focal plane of a lens, the complex amplitude of light coming from a transmittance \( t(x_1, y_1) \). The latter is illuminated by a plane wave and is located some arbitrary distance \( d \) in front of the lens. We find the complex amplitude to be a product of the Fourier transform of \( t(x_1, y_1) \) times a spherical phase factor

\[
\exp \left[ -\frac{in}{\lambda} \left( \frac{1}{f} - \frac{d}{f^2} \right) (x_1^2 + y_1^2) \right]
\]

which is independent of the form of \( t(x_1, y_1) \). Equation (6.59) holds for both subject and reference waves. The phase factor is independent of whether we consider the subject transmittance or the reference source to be \( t(x_1, y_1) \) since each is the same distance \( d \) from the lens. Consequently in the products \( a^*r \) and \( a^*r \) arising from the recording process, the phase factors cancel out leaving only the Fourier transforms of subject and reference source as in Eq. (8.18).

Suppose now that subject and reference source are behind the lens in a plane a distance \( d \) from the back focal plane and suppose the lens to be illuminated with a plane wave. Equations (8.16)-(8.18) and the associated conclusions still apply if we substitute \( d \) for \( f \). In this case the converging light falling on the subject can be considered to derive from parallel light illuminating a lens of focal length \( d \) adjacent to the subject.

### 8.3.5 Lensless Fourier Transform Holograms

Suppose that the lens is removed from the hologram-forming arrangement of Fig. 8.25, but the reference point source is maintained in the plane of the subject transparency (Fig. 8.26). The subject wavefront to be recorded at the hologram plane is now the near-field or Fresnel diffraction pattern (Fresnel transform) of the subject transparency. Nevertheless we shall see that the hologram formed with the arrangement of Fig. 8.26 has a transmittance resembling that of the Fourier transform hologram. The image-forming terms are again products of a Fourier transform and a phase factor linear in the coordinate of the hologram plane. Hence the name lensless Fourier transform hologram is applied to holograms formed without lenses but with a point reference source in the plane of the subject. As before, the subject is illuminated with a plane wave.

According to Eq. (5.33) and the discussion following in Section 5.5, the complex amplitude of the subject light in the \( x_2\eta_2 \) hologram plane of Fig. 8.26 can be written as

\[
a(x_2, \eta_2) = \frac{i a_1}{\lambda d} \exp \left[ -\frac{in}{\lambda d} (x_2^2 + \eta_2^2) \right]
\]

\[
\times \iint \left[ s(x_1, \eta_1) \exp \left[ -\frac{in}{\lambda d} (x_1^2 + \eta_1^2) \right] \right] \exp [2\pi i (\xi x_1 + \eta \eta_1)] \, dx_1 \, d\eta_1
\]

\[
= c \exp \left[ -\frac{in}{\lambda d} (x_2^2 + \eta_2^2) \right] F(\xi', \eta')
\]

where

\[
F(\xi', \eta') = s(x_1, \eta_1) \exp \left[ -\frac{in}{\lambda d} (x_1^2 + \eta_1^2) \right], \quad c = \frac{i a_1}{\lambda d},
\]

\( d \) is the separation between the \( x_2\eta_2 \) and \( x_1\eta_1 \) planes, and \( \xi' = x_2/\lambda d, \eta' = \eta_2/\lambda d \). When a hologram is formed with a reference wave \( r = r_0 \exp \left[ -(in/\lambda d)(x_2^2 + \eta_2^2) \right] \exp (-2\pi i \xi b) \) [of the form given by Eq. (8.17) but with \( f = d \)], the image-forming terms of the hologram transmittance are represented as in Eq. (8.18) by

\[
a(x_2, \eta_2) r^* + a^*(x_2, \eta_2) r
\]

\[
= c r_0 F(\xi', \eta') \exp (2\pi i \xi b) + c^* r_0 F^*(\xi', \eta') \exp (-2\pi i \xi b).
\]
Once again the spherical phase factor dependent on the hologram coordinates $x_2$ and $y_2$ has been removed. The image generated by the hologram thus remains stationary with translation of the hologram. A spherical phase factor in $x_2$, $y_2$ multiplying the subject transmittance $s(x_1, y_1)$ in the transform does not affect the reconstruction properties. It can be regarded as merely part of the subject, i.e., a lens adjacent to the transparency.

When the hologram is illuminated by an axial plane wave, the complex amplitudes of the diffracted waves at the hologram which lead to image formation are proportional to the right-hand side of Eq. (8.20). Apart from the fact that $F'(\xi', \eta')$ and its conjugate are not Fourier transforms of $s(x_1, y_1)$ or of $s^*(-x_1, -y_1)$ but Fourier transforms of their products with spherical phase factors, the hologram transmittance in Eq. (8.20) is similar to that obtained with Fourier transform holograms. Plane wave illumination requires a lens, as in Fig. 8.23 to carry out a Fourier transformation of the hologram transmittance. [We note that $\xi$' and $\eta'$ in Eq. (8.20) differ by a constant scale factor $a = df/l$ from the spatial frequencies $\xi = x_2/\lambda f$ and $\eta = y_2/\lambda f$ defined in Eq. (8.16) for Fourier transformation by a lens of focal length $f$. Let us write $\xi'$ and $\eta'$ in the form $\xi' = \xi/a$, $\eta' = \eta/a$ and employ the similarity theorem, correspondence (4.22). We find that transformation of the terms of Eq. (8.20) by a lens of focal length $l$ divides the spatial dimensions of the image by the factor $a$ as compared to those obtained when $d = f$.] Carrying out the transformation, we obtain complex amplitudes in the image plane $x_3y_3$, located at the back focal plane of the lens which are proportional to

$$s \left[ (ax_3 - b), ay_3 \right] \exp \left\{ -\frac{i\pi}{\lambda d} \left[ (ax_3 - b)^2 + a^2y_3^2 \right] \right\}$$

and

$$s^* \left[ -(ax_3 + b), -ay_3 \right] \exp \left\{ \frac{i\pi}{\lambda d} \left[ (ax_3 + b)^2 + a^2y_3^2 \right] \right\}$$

times the spherical phase factors introduced by the reconstructing lens. All the phase factors drop out when the intensities of the images are observed. The image intensities $|s[(ax_3 - b), ay_3]|^2$ and $|s^* - (ax_3 + b), -ay_3)|^2$ represent one upright image centered at $(b/a, 0)$ and one inverted image centered at $(-b/a, 0)$ in the $x_3y_3$ plane.

While it is of interest to show that the hologram formed with the lensless Fourier transform method can be illuminated with a plane wave so as to produce an image display similar to that of the Fourier transform hologram, in practice it is often simpler to illuminate the hologram with the original reference point source. As with most holograms the virtual image then coincides with the original subject position.

Basic to the lensless Fourier transform method is the requirement that the curvature in the spherical phase factor associated with the reference wave be the same as that associated with the subject light. Translation invariance has been effectively used to produce composite holograms which reduce information content in pictorial holograms to the minimum necessary and allow alteration of the stereoscopic aspects of a 3D image (see Chapter 18).

### 8.3.6 Lensless Fraunhofer Holograms

In Section 2.5.4 it was noted that in-line holograms formed in the far field of a subject permitted the examination of one image without confusion from the other. Thompson et al. [8.15] employed the method to measure the size and shape of dynamic aerosol particles, perhaps the first practical application of holography. The Fraunhofer or far-field diffraction pattern of a subject can be recorded on a photographic plate a distance $d$ from the subject provided

$$(x_1^2 + y_1^2)/\lambda \ll d, \quad \text{as in Eq. (5.37).} \quad (8.21)$$

Here $x_1$ and $y_1$ are the coordinates of any point on the subject and $\lambda$ is the wavelength of the illuminating light.

Suppose a subject transmittance $s(x_1, y_1)$, whose Fourier transform is $S(\xi, \eta)$, is illuminated with a plane wave traveling in a direction normal...
to the \( x_1 y_1 \) plane (see Fig. 8.27). Equation (5.39) expresses the complex amplitude \( a(x_2', y_2) \) of the light arriving at a plane \( x_2 y_2 \) separated from \( x_1 y_1 \) by a “far-field” distance \( d \):

\[
a(x_2', y_2) = \frac{i\lambda d}{\lambda d} \exp\left[-\frac{i\pi}{\lambda d} (x_2'^2 + y_2'^2)\right] S(\xi, \eta). \tag{8.22}
\]

As employed by Thompson et al. [8.15], a **Fraunhofer hologram** is formed by illuminating a subject with plane coherent light, placing a photographic plate in a plane a far-field distance from the subject and properly exposing it. Undiffracted light passing through the subject acts as the reference wave. As noted in Chapter 2 and in common with other in-line holograms, the image possesses the original contrast of the subject providing a positive hologram is properly made. In the reconstruction step, the hologram is illuminated by a plane wave similar to that which illuminated the subject.

Since the plane reference wave does not remove the spherical phase factor in Eq. (8.22) during formation of the hologram, the image generated by the Fraunhofer hologram does not have a position which is invariant to hologram translation. We may consider that the hologram has a built-in lens, represented by the spherical phase factor, which translates with the hologram and in turn translates the image.

When the subject becomes so small that it approaches an on-axis pointscatterer, as in the case of an aerosol particle, Fig. 8.27 becomes equivalent to Fig. 2.5 and a zone plate interference pattern is formed. If we can approximate the small subject by a delta function \( \delta(x) \rightarrow S(\xi, \eta) = 1 \), then \( a(x_2', y_2) \) in Eq. (8.22) approaches an unmodulated spherical wave. The reconstruction of \( a(x_2', y_2) \) by the illuminated Fraunhofer hologram is an identical wave diverging from the virtual image. This produces only a uniform background at the plane of the real image on the far side of the hologram and permits relatively undisturbed observation of that image.

We may define the **Fraunhofer hologram** in a general way as one which records interference of the far-field diffraction pattern of a subject with a spherical wave from a reference source not coplanar with the subject. Often the reference source is effectively at infinity and the reference wave is plane.

### 8.3.7 Fraunhofer Holograms Formed with a Lens

The far-field diffraction pattern \( a(x_2', y_2) \) in Eq. (8.22) is equivalent to the complex amplitude given in Eq. (8.16). Equation (8.16) applies to a coherently illuminated subject placed up against a lens of focal length \( f \); the light amplitude is observed in the back focal plane. It is apparent then that a **Fraunhofer hologram formed with a lens** will result when (1) a subject adjacent to a lens is illuminated with a plane wave, (2) a plane reference wave is introduced (e.g., off-axis), and (3) subject and reference waves are made to interfere over a hologram plate placed in the back focal plane of the lens. Properties similar to those of the lensless Fraunhofer hologram will be obtained. Figure 8.28 indicates the arrangement for forming such a Fraunhofer hologram.

---

**Fig. 8.28.** Formation of a Fraunhofer hologram with a lens.
8.4.1 LIMITED RESOLUTION IN THE RECORDING MEDIUM

Let us initially carry out our analysis in the $xz$ plane of Fig. 8.29 where a subject, a reference point source $R(x_0, 0, -z_s)$, and a recording plate are indicated. The plane normal to the axis and containing the reference point source is separated from the recording plate by the axial distance $z_r$. As indicated in Fig. 8.29 a ray from the reference source to an arbitrary point $Q(x, 0, 0)$ on the plate makes an angle $\theta_0$ with the normal to the plate (the $z$ axis) and is therefore characterized by the spatial frequency $\xi_0$. We suppose a small bundle of such rays from the reference interferes in the vicinity of $Q$ with a similar bundle of rays from the subject making an angle $\theta_0'$ to the $z$ axis and characterized by a spatial frequency $\xi_0'$. The interference pattern intensity which must be recorded is

$$I = [\exp(2\pi i \xi_0 x_0) + \exp(2\pi i \xi_0' x_0)] \exp(-2\pi i \xi_0 x_2) + \exp(-2\pi i \xi_0' x_2)$$

where we have assumed unit amplitude waves and where the frequency of the cosine term $\xi_0 - \xi_0'$ is the fringe frequency. For small angles, we approximate the fringe frequency of the pattern by

$$\xi_0 - \xi_0' \approx \frac{\sin \theta_0 - \sin \theta_0'}{\lambda} \approx \frac{\theta_0 - \theta_0'}{\lambda}. \quad (8.24)$$

Suppose now that the recording medium is perfectly capable of recording all fringe frequencies below a cutoff frequency $\xi_s$ but is totally incapable of recording fringe frequencies exceeding $\xi_s$. For a fixed value of $\xi_s$ and a subject which scatters light over a broad band of spatial frequencies $\xi_0$, this means that there will be limits to the subject spatial frequencies that can be stored in the hologram. That is, if rays from the subject are to be stored in the hologram, their angles with the $z$ axis must lie within certain limits. In Fig. 8.29 one such limiting ray passing from the subject to an arbitrary point $Q$ on the hologram plate intersects the plane of the reference source at the point $x_{\text{max}}$. We call this ray a marginal ray; its spatial frequency $\xi_{0,\text{max}}$ satisfies

$$\xi_{0,\text{max}} = \xi_s. \quad (8.25a)$$

If the angle the marginal ray makes with the $z$ axis is $\theta_{0,\text{max}}$, then we can write, with the help of Eq. (8.24), that

$$\frac{\theta_{0,\text{max}} - \theta_s}{\lambda} \approx \xi_s = \frac{\theta_s}{\lambda} \quad (8.25b)$$

or

$$\theta_{0,\text{max}} - \theta_s \approx \theta_s. \quad (8.25c)$$

The point $x_{\text{max}}$ can be obtained by examining Fig. 8.29 and is given by

$$\frac{x_{\text{max}} - x_0}{x_r} = \tan(\theta_s + \theta_0) \approx \theta_s + \theta_0 \approx \frac{x_r - x_2}{x_r} + \xi_s \lambda$$

or

$$x_{\text{max}} \approx x_r + x_r \lambda \xi_s. \quad (8.26)$$

In the same manner we obtain for $x_{\text{min}}$, the intersection coordinate of the other marginal ray,

$$x_{\text{min}} \approx x_r - x_r \lambda \xi_s. \quad (8.27)$$

Note that the marginal rays intersect the plane of the reference source at a coordinate independent of the point $Q$ in the recording plane.

We now define a hypothetical opaque mask located in a plane transverse to the $z$ axis and containing the reference source. It has an aperture which is centered at the source $R$ and extends between limits given by the points $(x_{\text{max}}, 0, -z_r)$ and $(x_{\text{min}}, 0, -z_r)$. Through the mask pass all subject rays whose interference with reference rays at some point $Q$ on the hologram plate will produce fringes capable of being resolved and recorded by the photosensitive medium. If a ray from the subject to an arbitrary point $Q$...
on the hologram plate, or an extension of that ray, is intercepted by the hypothetical mask, then that ray will not be recorded by the hologram. Several such rays are indicated in Fig. 8.30. The rule as stated holds for subjects in front of as well as behind the reference source plane. We can extend the analysis to the case of a two-dimensional mask by considering spatial frequencies \( \nu \) where \( \nu^2 = \xi^2 + \eta^2 \), and a cutoff frequency \( \nu_c \); the aperture in the mask then becomes a circle of radius \( z_c \nu_c \) centered at \( R \).

**Fig. 8.30.** Subject rays which are not recorded by the hologram.

Let us now employ the mask concept to illustrate the consequences of a cutoff fringe frequency \( \xi_0 \) for several of the hologram-forming geometries previously considered. We begin with the lensless Fourier transform geometry shown in Fig. 8.26. Since the subject transparency and the reference point source are in the same plane, the plane of the mask, only the size of the subject which can be recorded is restricted by the mask. To the extent that the small-angle approximations used in the analysis are valid, rays of all spatial frequencies coming from the subject are faithfully recorded by the hologram. The same is true for the arrangement in Fig. 8.25 using a lens. These two geometries therefore allow the holographic recording of high spatial frequency subjects on low resolution media.

Next consider the geometry for forming in-line Fresnel holograms with an axial plane wave illuminating the subject transparency. The reference source \( R \) lies on the \( z \) axis at infinity \((x_s = 0, z_s = \infty)\). The effect of the mask at infinity can best be described in terms of the limitation on angle which it imposes. If we define \( \theta_{\max} = x_{\max}/z_r \) and \( \theta_{\min} = x_{\min}/z_r \), Eqs. (8.26) and (8.27) become

\[
\theta_{\max} = \theta_c, \quad \theta_{\min} = \theta_c. \tag{8.28}
\]

These relations indicate that the extreme angles \( \theta_{\max} \) and \( \theta_{\min} \) which subject rays can make with the \( z \) axis and still be recorded are directly determined by the cutoff angle \( \theta_c \) (Fig. 8.31). The subject plane wave with ray direction \( | \xi_r | \) and spatial frequency \( | \xi_e | = | \xi_r | / \lambda \) corresponds to a spatial frequency component of the subject transmittance which also has the value \( | \xi_r | \) (see Section 5.3). Hence the hologram will not record spatial frequency components of the subject transmittance greater than \( | \xi_r | \); the resolution in the

**Fig. 8.31.** Effect of limited recording resolution with in-line hologram geometry.

**Fig. 8.32.** Effect of limited recording resolution with off-axis geometry and a plane reference wave.
image is limited accordingly. Image resolution obtainable with this forming method is equivalent to that which can be obtained from a contact print of the subject made on the same recording medium.

Figure 8.32 illustrates the consequences of limited resolution in the recording medium when diffraction from a subject transparency illuminated by an axial plane wave interferes with an off-axis plane reference wave. The beam arrangement forms a carrier-frequency, Fresnel hologram. Equations (8.26) and (8.27) written in terms of the limiting mask angles \( \theta_{\text{max}} \) and \( \theta_{\text{min}} \) now take the form

\[
\theta_{\text{max}} = \theta_c + \theta_s, \quad \theta_{\text{min}} = \theta_c - \theta_s.
\]  

(8.29)

We see from the figure that unless \( \theta_c \) is large enough, only a limited range of positive spatial frequencies diffracted by the subject will pass to the plate without intercepting the angular mask. One such ray from the subject making an angle \( \theta_s \) to the \( z \) axis is shown. The negative subject spatial frequencies are far less limited (the case would be reversed if the reference wave had a positive spatial frequency). This uneven recording of positive and negative spatial frequencies can reduce resolution in the image. To avoid this, the cutoff frequency \( \xi_c \) must be high. Moreover, for complete separation of the zero order and the image-forming diffracted waves, \( \xi_c \) must exceed the highest subject spatial frequency by a factor of four (see Section 8.1.1).

### 8.4.2 Hologram Size

When arrays of holograms are considered for use as document storage files (Chapter 16), the individual holograms must necessarily be small in size if the number of stored documents is to be large. We consider the effect of limited hologram size on image quality to be similar to that for the finite-size lens in Section 6.4.2. The main result developed there holds for the hologram if in the discussion we replace the lens aperture by the hologram aperture.

Resolution in the image from either a hologram or a lens is determined by the Fourier transform of the coherent transfer function of the imaging device, i.e., the spread function. The larger the hologram the greater the maximum subject spatial frequency that can be accepted and the narrower the spread function (see Fig. 6.11). When the hologram size is limited by other considerations, such as maximizing information storage density, it is important to optimize the use of the limited area. If we assume that the recording medium has sufficiently high resolution to record all spatial frequencies coming from the subject, then the goal is to record uniformly all frequency components from all parts of the subject.

In Section 6.4.2 it is shown that a lens imaging system, where the subject is illuminated with an axial plane wave, yields maximum image resolution only for subject points near the optic axis. Resolution in the image falls off linearly with distance from the axis. A better way to illuminate the subject, shown in Fig. 6.9, is to use a converging spherical wave which focuses at the lens. With this method the highest possible image resolution consistent with the fixed size of the lens is obtained for all subject points. The imaging system is said to be space-invariant and characterized by the top-hat transfer function of Fig. 6.10. The analogous, space-invariant, hologram-forming arrangements are shown in Figs. 8.25 and 8.28. These methods make optimum use of the limited hologram area by causing the small hologram to receive the same range of spatial frequencies from each subject point and consequently produce an image which is uniformly well resolved over its entire extent. Since a cone of rays from each subject point fills the hologram area (see Fig. 6.9) information about every subject point is stored at every point of the hologram. Holograms formed as in Fig. 8.25 or 8.28 are thus resistant to dust and scratches, as are holograms formed with diffuse subject light, but do not have the problem of speckle. The main problem associated with the hologram-forming arrangements of Figs. 8.25 and 8.28 is the large range of subject light intensity that must be recorded. As previously noted, there is a large intensity peak at zero spatial frequency. Some useful compromises between the redundancy with which information is stored and this intensity range will be discussed in Chapter 16.

### 8.5 Maximum Efficiency of Plane Holograms

If bright reconstructions are desirable, it is useful to know the maximum diffraction efficiency obtainable from various hologram types. Here we compute the efficiencies [8.18] of plane absorption and phase holograms formed by interference of an off-axis reference plane wave \( r = r \exp(2\pi i \xi x) \) with an axial, unmodulated plane wave signal of amplitude \( a \).

We begin with the absorption hologram. If we assume that the exposure-dependent term \( t_r \) of the amplitude transmittance of the developed hologram is proportional to the intensity of the interference pattern to which the hologram was exposed, then

\[
t_r \propto I = [a + r \exp(2\pi i \xi x)][a + r \exp(-2\pi i \xi x)]
\]

\[
= a^2 + r^2 + 2ar \cos(2\pi \xi x)
\]
224 8. Analysis of Plane Holograms

or

\[ t_B = t_0 + t_1 \cos(2\pi \xi x). \] (8.30)

For the plane absorption hologram the total amplitude transmittance \( t \) is defined in Eq. (1.14) as

\[ t = t_0 - t_B \]

where \( t_0 \) is the transmittance of the unexposed plate. At most, \( t \) can vary between 0 and 1 (when \( t_0 = 1 \)). This maximum range is achieved when \( t_B = \frac{1}{2} \) and \( t_1 = \frac{1}{4} \). Under these conditions

\[ t = t_0 - t_B = 1 - \frac{1}{4} - \frac{1}{4} \cos(2\pi \xi x) \]

\[ = \frac{1}{4} - \frac{1}{4} \cos(2\pi \xi x) \cos(2\pi \xi x) \]

\[ = \frac{1}{4} - \frac{1}{4} \exp(2\pi i \xi x) - \frac{1}{4} \exp(-2\pi i \xi x). \] (8.31)

Suppose the hologram is illuminated by an axial plane wave of unit amplitude. The amplitude of the light emerging from the hologram is then \( t \) in Eq. (8.31). We see that light is diffracted only into the zero order and into the +1 and -1 orders. Since each first-order diffracted wave receives \( \frac{1}{4} \) of the incident light amplitude, each first-order wave has \( \frac{1}{4} t \) the intensity of the incident light. Diffraction efficiency is defined as the ratio of power diffracted into one first-order wave to the power illuminating the hologram. In the present case where the hologram is uniformly exposed, we may substitute intensity for power in the definition. Thus the efficiency is \( \frac{1}{4} t \) or 6.25%. Practical recording media are not usually linear over the full range of exposure required to make the transmittance \( t \) vary from 0 to 1. Therefore the maximum efficiency of 6.25% cannot be achieved if it is also necessary to reconstruct a wavefront proportional to the original signal wavefront.

A somewhat higher diffraction efficiency can be obtained if the transmittance \( t_B \) varies as a square-wave function of \( x \), e.g., a Ronchi grating. Computer-generated holograms may be of this nature. In this case \( t_B \) is 0 in \( \frac{1}{2} \) of the square-wave period and 1 in the remainder. The first two terms of the Fourier series representation of the square wave are [8.19]

\[ t_B = \frac{1}{4} + \frac{1}{4} \cos(2\pi \xi x) + \cdots + \frac{1}{4} \cos(2\pi \xi x) \]

\[ = \frac{1}{4} + \frac{1}{4} \cos(2\pi \xi x) \exp(2\pi i \xi x) + \exp(-2\pi i \xi x) \] (8.32)

When \( t = 1 - t_B \), the diffraction efficiency is \( (1/4)^2 = 10.1\% \). Square-wave gratings also diffract appreciably into higher orders.

In Section 7.2.2 the complex transmittance of a photosensitive medium is expressed as

\[ t = t \exp[i\varphi(x)]. \]

8.5 Maximum Efficiency of Plane Holograms

For a lossless, phase hologram we can consider \( t \) to be constant and equal to unity so that

\[ t = \exp[i\varphi(x)]. \] (8.33)

Although we note in Section 7.2.2 that phase holograms are capable of linear recording only when \( \varphi \) is very small, let us, for the purpose of computing maximum possible efficiency, remove this restriction on \( \varphi \). The photosensitive medium is exposed so that the phase shift \( \varphi(x) \), imposed on a plane wave traveling through the developed hologram is proportional to the intensity of the exposure. Thus

\[ \varphi(x) \propto a^2 + r^2 + 2ar \cos 2\pi \xi x = \varphi_0 + \varphi_1 \cos(2\pi \xi x). \] (8.34)

The transmittance of the hologram \( t \) becomes

\[ t = \exp(i\varphi_0) \exp[i\varphi_1 \cos(2\pi \xi x)] \]

\[ \propto \exp[i\varphi_1 \cos(2\pi \xi x)]. \] (8.35)

When we neglect the constant phase factor \( \exp(i\varphi_0) \), Eq. (8.35) may be represented by the Fourier series [8.19]

\[ t = \exp[i\varphi_1 \cos(2\pi \xi x)] = \sum_{n=-\infty}^{\infty} i^n J_1(\varphi_1) \exp(in2\pi \xi x), \]

\[ J_1(\varphi_1) \]

\[ \int 0 0.6 0.4 0.2 0 \]

\[ 1 2 \]

\[ \phi_1 \]

FIG. 8.33. The amplitude of a first-order diffracted wave \( J_1(\varphi_1) \) for a sinusoidal phase grating.
where $J_n$ is the Bessel function of the first kind and $n$th order. When $t$ is illuminated with an axial plane wave of unit amplitude, the diffracted amplitude in the $+1$ order is $J_1(\varphi)$, a function plotted in Fig. 8.33. Its maximum amplitude is 0.582 and the maximum efficiency is 33.9%.

Somewhat more light can be diffracted into the first order if the phase modulation varies as a square-wave function of $x$ and assumes the value $\varphi = 0$ during half the square-wave period and $\varphi = \pi$ during the remainder. The transmittance $t$ in Eq. (8.33) is then $+1$ when $\varphi = 0$ and $-1$ when $\varphi = \pi$. This is analogous to the square-wave, absorption hologram where now the amplitude of the incident light diffracted into the first order is twice that for amplitude modulation. Hence efficiency is four times as large or 40.4%.

REFERENCES