

6

Holography

6.1 INTRODUCTION

Holography is the synthesis of interference and diffraction. In recording a hologram, two waves interfere to form an interference pattern on the recording medium. When reconstructing the hologram, the reconstructing wave is diffracted by the hologram. When looking at the reconstruction of a 3-D object, it is like looking at the real object. It is therefore said that: 'A photograph tells more than a thousand words and a hologram tells more than a thousand photographs'.

Although holography requires coherent light, it was invented by Gabor back in 1948, more than a decade before the invention of the laser. By means of holography an original wave field can be reconstructed at a later time at a different location. This technique therefore has many potential applications. In this book we concentrate on the technique of holographic interferometry. Because of the above-mentioned properties, we shall see that holographic interferometry has many advantages compared to standard interferometry.

6.2 THE HOLOGRAPHIC PROCESS

Figure 6.1(a) shows a typical holography set-up. Here the light beam from a laser is split in two by means of a beamsplitter. One of the partial waves is directed onto the object by a mirror and spread to illuminate the whole object by means of a microscope objective. The object scatters the light in all directions, and some of it impinges onto the hologram plate. This wave is called the object wave. The other partial wave is reflected directly onto the hologram plate. This wave is called the reference wave. In the figure this wave is collimated by means of a microscope objective and a lens. This is not essential, but it is important that the reference wave constitutes a uniform illumination of the hologram plate. The hologram plate must be a light-sensitive medium, e.g. a silver halide film plate with high resolution. We now consider the mathematical description of this process in more detail. For more comprehensive treatments, see Collier *et al.* (1971), Smith (1969), Caulfield (1979) and Hariharan (1984).

Let the object and reference waves in the plane of the hologram be described by the field amplitudes u_o and u respectively. These two waves will interfere, resulting in an intensity distribution in the hologram plane given by

$$I = |u + u_o|^2 = |u|^2 + |u_o|^2 + u_o^*u + u_o u^* \quad (6.1)$$

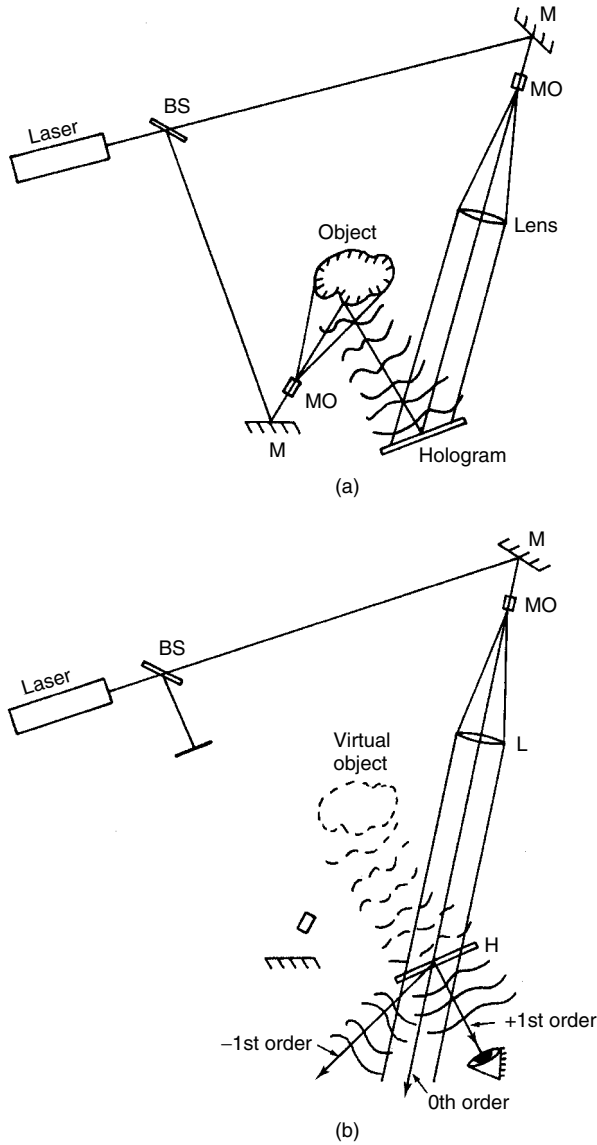


Figure 6.1 (a) Example of a holography set-up. BS = beamsplitter, M = mirrors, MO = microscope objectives and (b) Reconstruction of the hologram

We now expose the film plate to this intensity distribution until it gets a suitable blackening. Then it is removed from the plate holder and developed. We now have a hologram. The process so far is called a hologram recording.

This hologram has an amplitude transmittance t which is proportional to the intensity distribution given by Equation (6.1). This means that

$$t = \alpha I = \alpha |u|^2 + \alpha |u_0|^2 + \alpha u_0^* + \alpha u_0 u^* = t_1 + t_2 + t_3 + t_4 \tag{6.2}$$

We then replace the hologram back in the holder in the same position as in the recording. We block the object wave and illuminate the hologram with the reference wave which is now termed the reconstruction wave (see Figure 6.1(b)). The amplitude distribution u_a just behind the hologram then becomes equal to the field amplitude of the reconstruction wave multiplied by the amplitude transmittance of the hologram, i.e.

$$u_a = t \cdot u = \alpha[|u|^2 + |u_0|^2]u + \alpha(uu)u_0^* + \alpha|u|^2u_0 \quad (6.3)$$

As mentioned above, the reference (reconstruction) wave is a wave of uniform intensity. The quantity $|u|^2$ is therefore a constant and the last term of Equation (6.3) thus becomes (apart from a constant) identical to the original object wave u_0 . We therefore have been able to reconstruct the object wave, maintaining its original phase and relative amplitude distribution. The consequence is that, by looking through the hologram in the direction of the object, we will observe the object in its three-dimensional nature even though the physical object has been removed. Therefore this reconstructed wave is also called the virtual wave.

The other two terms of Equation (6.3) represent waves propagating in the directions indicated in Figure 6.1(b). In fact, a hologram can be regarded as a very complicated grating where the first term of Equation (6.3) represents the zeroth order and the second and third terms represent the \pm first side orders diffracted from the hologram. If we could use u^* , the conjugate of u , as the reconstruction wave, we see that the second term of Equation (6.3) would have become proportional to $|u|^2u_0^*$, i.e. the conjugate of the object wave would have been reconstructed. The physical meaning of this deserves some explanation. Complex conjugation of a field amplitude means changing the sign of its phase. It thus gives a wave field returning back on its own path. u_0^* therefore represents a wave propagating from the hologram back to the object forming an image of the object. It is therefore termed the real wave. To reconstruct the hologram with u^* in the case of a pure plane wave, the reconstruction wave can be reflected back through the hologram by means of a plane mirror. An easier way, which also applies for a general reference (reconstruction) wave, is to turn the hologram 180° around the vertical axis. By placing a screen in the real wave, we can observe the image of the object on the screen.

In Figure 6.2 another possible realization of a holography set-up is sketched. Here the expanded laser beam is wavefront-divided by means of a mirror which reflects the

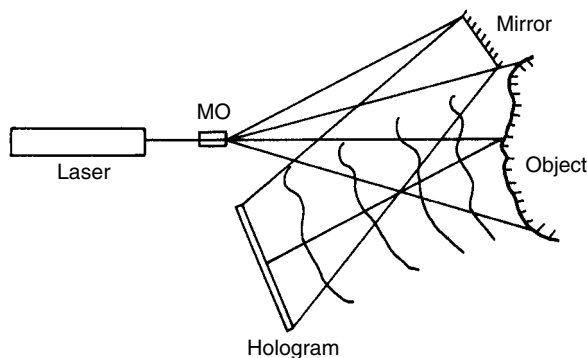


Figure 6.2

reference wave onto the hologram. This set-up is normally more stable than in Figure 6.1 since fewer components are involved.

6.3 AN ALTERNATIVE DESCRIPTION

An alternative and more physical description of the holographic process has already been touched on in Section 4.3.1. Let the point source P in Figure 4.7(a) represent the light from a point on the object, and let the plane wave represent the reference wave. The resulting zone plate pattern is recorded on a film. In Figure 4.7(b) this developed film (the hologram) is illuminated by a plane wave (the reconstruction wave). When viewed through the film, the diffracted, diverging spherical wave looks as if it is coming from P. This argument can be repeated for all points on the object and give us the virtual reconstructed object wave. The spherical wave converging to point P' represents the real wave.

The circular zone plate is therefore also termed a unit hologram. In the general case when the object- and reference waves are not normally incident on the hologram, the pattern changes from circular to elliptical zone plate patterns, and the diffracted virtual and real waves propagate in different directions in the reconstruction process.

6.4 UNCOLLIMATED REFERENCE AND RECONSTRUCTION WAVES

We now consider in more detail the locations of the virtual and real images for the most general recording and reconstructing geometries. To do this, it suffices to consider a single object point source with coordinates (x_o, y_o, z_o) : see Figure 6.3. Here the hologram film is placed in the xy -plane and the reference wave is coming from a point source with coordinates (x_r, y_r, z_r) . Using quadratic (Fresnel) approximations to the spherical waves, the object and reference fields of wavelength λ_1 incident on the xy -plane may be written

$$u_o = U_o \exp \left\{ i \frac{\pi}{\lambda_1 z_o} [(x - x_o)^2 + (y - y_o)^2] \right\} \quad (6.4)$$

$$u = U \exp \left\{ i \frac{\pi}{\lambda_1 z_r} [(x - x_r)^2 + (y - y_r)^2] \right\} \quad (6.5)$$

The transmittance of the resulting hologram we write as

$$t \propto |u_o + u|^2 = t_1 + t_2 + t_3 + t_4 \quad (6.6)$$

where the interesting terms (cf. Equation (6.2)) are

$$t_3 = \alpha U U_o \exp \left\{ i \frac{\pi}{\lambda_1 z_r} [(x - x_r)^2 + (y - y_r)^2] - i \frac{\pi}{\lambda_1 z_o} [(x - x_o)^2 + (y - y_o)^2] \right\} \quad (6.7)$$

$$t_4 = \alpha U U_o \exp \left\{ -i \frac{\pi}{\lambda_1 z_r} [(x - x_r)^2 + (y - y_r)^2] + i \frac{\pi}{\lambda_1 z_o} [(x - x_o)^2 + (y - y_o)^2] \right\} \quad (6.8)$$

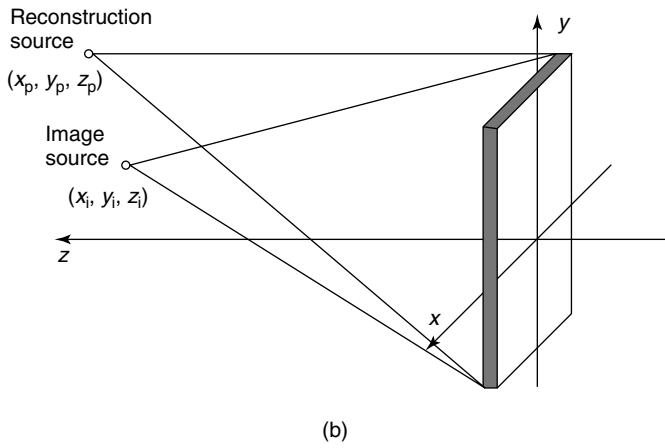
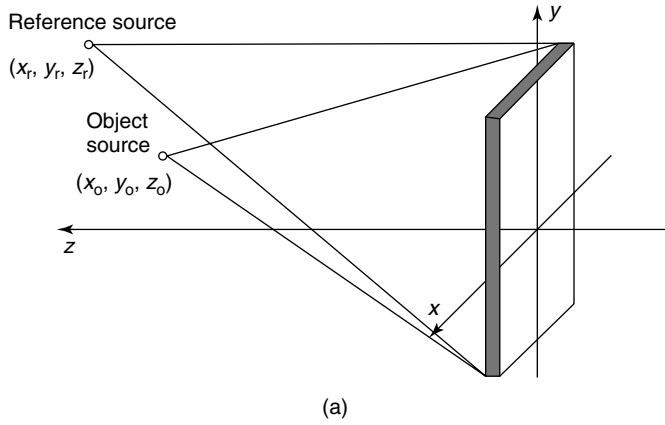


Figure 6.3 (a) Recording and (b) reconstruction geometries of point sources

In reconstruction, the hologram is illuminated by the spherical wave

$$u_p = U_p \exp \left\{ i \frac{\pi}{\lambda_2 z_p} [(x - x_p)^2 + (y - y_p)^2] \right\} \tag{6.9}$$

where we have allowed for both a displaced (relative to the reference wave) point source and a different wavelength λ_2 . The two reconstructed waves of interest are $u_3 = t_3 u_p$ and $u_4 = t_4 u_p$ which gives (writing out the x -dependence only)

$$u_3 = t_3 u_p \propto \exp \left\{ i \frac{\pi}{\lambda_1 z_r} (x^2 + x_r^2 - 2x_r x) - i \frac{\pi}{\lambda_1 z_o} (x^2 + x_o^2 - 2x_o x) + i \frac{\pi}{\lambda_2 z_p} \right. \\ \left. \times (x^2 + x_p^2 - 2x_p x) \right\}$$

$$\begin{aligned}
 &= \exp \left\{ i\pi \left[\frac{x_r^2}{\lambda_1 z_r} - \frac{x_o^2}{\lambda_1 z_o} + \frac{x_p^2}{\lambda_2 z_p} \right] \right\} \exp \left\{ i\pi \left(\frac{1}{\lambda_1 z_r} - \frac{1}{\lambda_1 z_o} + \frac{1}{\lambda_2 z_p} \right) x^2 \right\} \\
 &\quad \times \exp \left\{ -2i\pi \left(\frac{x_r}{\lambda_1 z_r} - \frac{x_o}{\lambda_1 z_o} + \frac{x_p}{\lambda_2 z_p} \right) x \right\}
 \end{aligned} \tag{6.10}$$

By performing the same calculations for the wave u_4 , we get for the phase terms depending on x^2 and x

$$u_4 \propto \exp \left\{ i\pi \left(-\frac{1}{\lambda_1 z_r} + \frac{1}{\lambda_1 z_o} + \frac{1}{\lambda_2 z_p} \right) x^2 \right\} \exp \left\{ -2i\pi \left(-\frac{x_r}{\lambda_1 z_r} + \frac{x_o}{\lambda_1 z_o} + \frac{x_p}{\lambda_2 z_p} \right) x \right\} \tag{6.11}$$

A spherical wave diverging from a point (x_i, y_i, z_i) (writing out only the x -dependence) is given as:

$$\begin{aligned}
 u_i &= U_i \exp \left\{ i \frac{\pi}{\lambda_2 z_i} (x - x_i)^2 \right\} = U_i \exp \left\{ i \frac{\pi}{\lambda_2 z_i} (x^2 + x_i^2 - 2x_i x) \right\} \\
 &= U_i \exp \left\{ i \frac{\pi}{\lambda_2 z_i} x_i^2 \right\} \exp \left\{ i \frac{\pi}{\lambda_2 z_i} x^2 \right\} \exp \left\{ -2i\pi \frac{x_i}{\lambda_2 z_i} x \right\}
 \end{aligned} \tag{6.12}$$

By comparing this with the above expressions for u_3 and u_4 , we get

$$\frac{1}{\lambda_2 z_i} = \pm \frac{1}{\lambda_1 z_r} \mp \frac{1}{\lambda_1 z_o} + \frac{1}{\lambda_2 z_p}, \quad \text{i.e.} \quad z_i = \left(\frac{1}{z_p} \pm \frac{\lambda_2}{\lambda_1 z_r} \mp \frac{\lambda_2}{\lambda_1 z_o} \right)^{-1} \tag{6.13}$$

and

$$\frac{x_i}{\lambda_2 z_i} = \pm \frac{x_r}{\lambda_1 z_r} \mp \frac{x_o}{\lambda_1 z_o} + \frac{x_p}{\lambda_2 z_p}, \quad \text{i.e.} \quad x_i = \mp \frac{\lambda_2 z_i}{\lambda_1 z_o} x_o \pm \frac{\lambda_2 z_i}{\lambda_1 z_r} x_r + \frac{z_i}{z_p} x_p \tag{6.14}$$

and with a completely analogous expression for y_i :

$$y_i = \mp \frac{\lambda_2 z_i}{\lambda_1 z_o} y_o \pm \frac{\lambda_2 z_i}{\lambda_1 z_r} y_r + \frac{z_i}{z_p} y_p \tag{6.15}$$

Here the upper set of signs applies for u_3 , the real reconstructed wave, and the lower set for u_4 , the virtual wave. What we have done is to find the coordinates (x_i, y_i, z_i) of the image point expressed by the coordinates of the object point, the source point of the reference and the reconstruction waves. We see that when $\lambda_2 = \lambda_1$ and $z_p = z_r$, we get for the virtual wave $z_i = z_o$. When, in addition, $z_r = \infty$ (collimated reference and reconstruction waves), $z_i = -z_o$ for the real wave.

From our calculations, we can associate a transversal magnification

$$m = \left| \frac{x_i}{x_o} \right| = \left| \frac{y_i}{y_o} \right| = \left| \frac{\lambda_2 z_i}{\lambda_1 z_o} \right| = \left| 1 - \frac{z_o}{z_r} \mp \frac{\lambda_1 z_o}{\lambda_2 z_p} \right|^{-1} \tag{6.16}$$

6.5 DIFFRACTION EFFICIENCY. THE PHASE HOLOGRAM

Assume the object- and reference waves to be described by

$$u_o = U_o e^{i\phi_o} \quad (6.17a)$$

and

$$u = U e^{i\phi} \quad (6.17b)$$

respectively. The resulting amplitude transmittance then becomes

$$\begin{aligned} t &= \alpha[U^2 + U_o^2 + UU_o e^{i(\phi-\phi_o)} + UU_o e^{-i(\phi-\phi_o)}] \\ &= \alpha(I + I_o)[1 + V \cos(\phi - \phi_o)] \end{aligned} \quad (6.18)$$

which can be written as

$$t = t_b \left[1 + \frac{V}{2} e^{i(\phi-\phi_o)} + \frac{V}{2} e^{-i(\phi-\phi_o)} \right] \quad (6.19)$$

where $I = U^2$, $I_o = U_o^2$ and where we have introduced the visibility V (see eq. (3.29)) and the bias transmittance $t_b = \alpha(I + I_o)$. Since the transmittance t never can exceed unity and $0 \leq V \leq 1$, we see from Equation (6.18) that $t_b \leq 1/2$.

The reconstructed object wave u_r is found by multiplying the last term of Equation (6.19) by the reconstruction wave u :

$$u_r = t_b \frac{V}{2} U e^{i\phi_o} \quad (6.20)$$

and the intensity

$$I_r = |u_r|^2 = \frac{1}{4} U^2 t_b^2 V^2 \quad (6.21)$$

The diffraction efficiency η of such a hologram we define as the ratio of the intensities of the reconstructed wave and the reconstruction wave, i.e.

$$\eta = I_r/I = \frac{1}{4} t_b^2 V^2 \quad (6.22)$$

From this expression we see that the diffraction efficiency is proportional to the square of the visibility. η therefore reaches its maximum when $V = 1$, i.e. when $I_o = I$, which means that the diffraction efficiency is highest when the object and reference waves are of equal intensity.

Maximum possible diffraction efficiency is obtained for $V = 1$ and $t_b = \frac{1}{2}$, which gives

$$\eta_{\max} = \frac{1}{16} = 6.25\%$$

This type of hologram is called an amplitude hologram because its transmittance is a pure amplitude variation. A hologram with a pure phase transmittance is called a phase

hologram. Such holograms can be produced in different ways. A commonly applied method consists of bleaching the exposed silver grains in the film emulsion of a standard amplitude hologram. The recorded amplitude variation then changes to a corresponding variation in emulsion thickness. The transmittance t_p of a phase hologram formed by bleaching of an amplitude hologram can be written as

$$t_p = e^{iM \cos(\phi_0 - \phi)} = \sum_{n=-\infty}^{\infty} i^n J_n(M) e^{in(\phi_0 - \phi)} \quad (6.23)$$

where J_n is the n th-order Bessel function. Here M is the amplitude of the phase delay. From this expression we see that a sinusoidal phase grating will diffract light into n orders in contrast to a sinusoidal amplitude grating which has only ± 1 st orders. The amplitude of the first-order reconstructed object wave is found by multiplying Equation (6.23) by the reconstruction wave u for $n = 1$, i.e.

$$u_r = J_1(M) U e^{i\phi_0} \quad (6.24)$$

and the intensity

$$I_r = U^2 J_1^2(M) \quad (6.25)$$

The diffraction efficiency becomes

$$\eta_p = I_r/I = J_1^2(M) \quad (6.26)$$

Since $J_{1\max}(M) = 0.582$ for $M = 1.8$, the maximum possible diffraction efficiency of a phase hologram is

$$\eta_{p,\max} = 0.339 = 34\%$$

6.6 VOLUME HOLOGRAMS

Up to now we have regarded the hologram film emulsion as having negligible thickness.

For emulsions of non-negligible thickness, however, volume effects, hitherto not considered, must be taken into account. For example, a thick phase hologram can reach a theoretical diffraction efficiency of 100 per cent.

Consider Figure 6.4(a) where two plane waves are symmetrically incident at the angles $\theta/2$ to the normal on a thick emulsion. These waves will form interference planes parallel to the yz -plane with spacings (cf. eq. (3.21)).

$$d = \frac{\lambda}{2 \sin(\theta/2)} \quad (6.27)$$

After development of this hologram, the exposed silver grains along these interference planes will form silver layers that can be regarded as partially reflecting plane mirrors. In Figure 6.4(b) this hologram is reconstructed with a plane wave incident at an angle ψ . This wave will be reflected on each 'mirror' at an angle ψ .

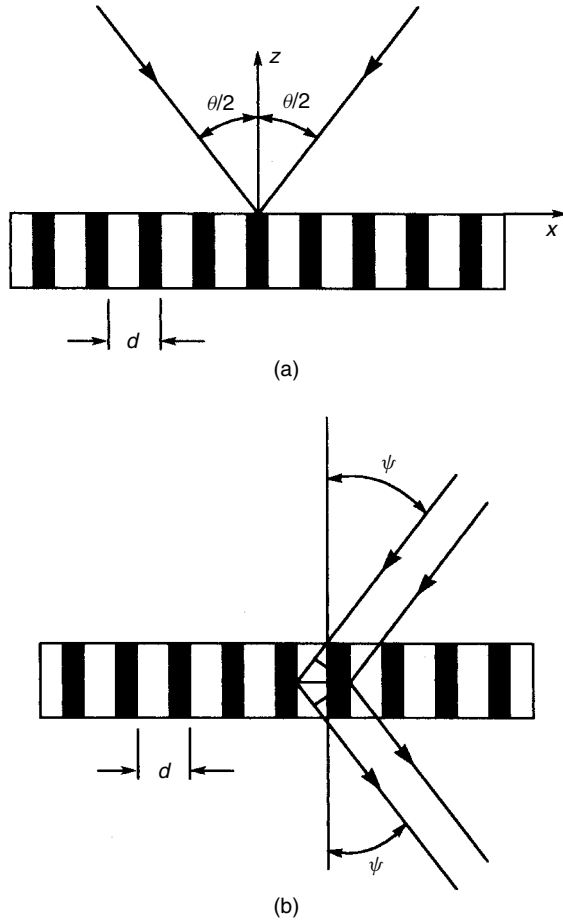


Figure 6.4

To obtain maximum intensity of the reflected, reconstructed wave, the path length difference between light reflected from successive planes must be equal to λ . From the triangles in Figure 6.4(b) this gives

$$2d \sin \psi = \lambda \tag{6.28}$$

which, by substitution of Equation (6.27), gives

$$\sin \psi = \sin \theta/2 \tag{6.29}$$

i.e. the angles of incidence of the reconstruction and reference waves must be equal. It can be shown that for a thick hologram, the intensity of the reconstructed wave will decrease rapidly as ψ deviates from $\theta/2$; see Section 13.6. This is referred to as the Bragg effect and Equation (6.29) is termed the Bragg law.

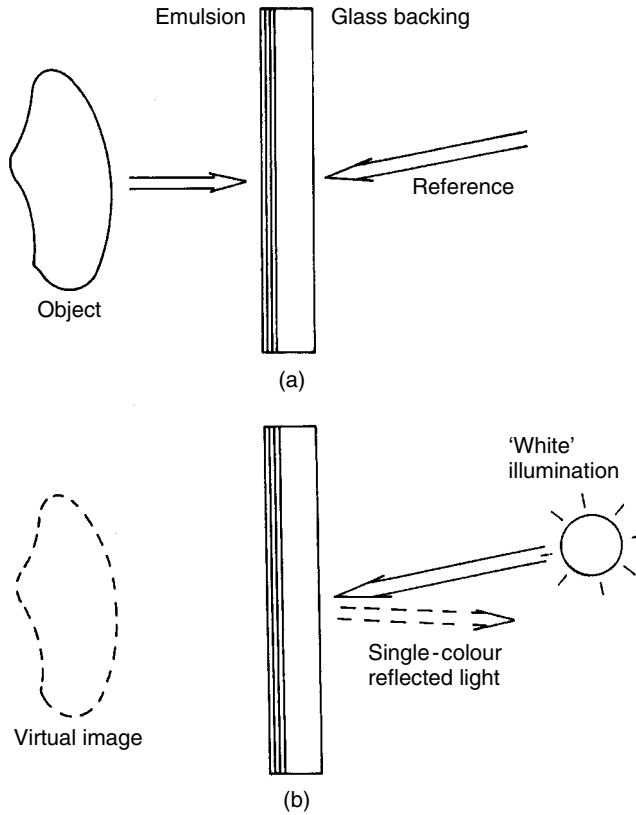


Figure 6.5

A special type of volume hologram, called a reflection hologram, is obtained by sending the object and reference waves from opposite sides of the emulsion, as shown in Figure 6.5(a). Then $\theta = 180^\circ$ and the stratified layers of metallic silver of the developed hologram run nearly parallel to the surface of the emulsion with a spacing equal to $\lambda/2$ (see Equation (6.27)). Owing to the Bragg condition, the reconstruction wave must be a duplication of the reference wave with the same wavelength, i.e. the hologram acts as a colour filter in reflection. Therefore a reflection hologram can be reconstructed in white light giving a reconstructed wave of the same wavelength as in the recording (see Figure 6.5(b)). In practice the wavelength of the reflected light is shorter than that of the exposing light, the reason being that the emulsion shrinks during the development process and the silver layers become more closely spaced.

6.7 STABILITY REQUIREMENTS

In the description of the holographic recording process we assumed the spatial phases of both the object- and reference waves to be time independent during exposure. It is clear, however, that relative movements between the different optical components (like mirrors, beamsplitters, the hologram, etc.) in the hologram set-up will introduce such phase

changes. If, for instance, a mirror makes vibrations of amplitude greater than $\lambda/4$ during the exposure time, adjacent dark and bright interference fringes interchange their positions randomly, which can lead to a uniform blackening of the hologram film and therefore ruin the experiment. The exposure time using a 5 mW H-Ne laser is typically of the order of seconds. This poses stringent requirements on the stability of the set-up. Therefore the standard methods of holography are normally performed on vibration-isolated heavy tables with the optical components mounted in massive holders. There are, however, special techniques by which unwanted movements can to a certain extent be compensated for or subtracted from. Thus, successful holographic experiments executed on the factory floor using continuous wave lasers have been reported.

By using pulse lasers, exposure times down to the order of nanoseconds can be achieved. In such cases, unwanted movements become less important. The application of pulse lasers therefore substantially reduces the stringent stability requirements.

6.8 HOLOGRAPHIC INTERFEROMETRY

In Section 3.5 we mentioned an imaginary experiment where two waves reflected from two identical objects could interfere. With the method of holography now at hand, we are able to realize this type of experiment by storing the wavefront scattered from an object in a hologram. We then can recreate this wavefront by hologram reconstruction, where and when we choose. For instance, we can let it interfere with the wave scattered from the object in a deformed state. This technique belongs to the field of holographic interferometry (Vest 1979; Erf 1974; Jones and Wykes 1989). In the case of static deformations, the methods can be grouped into two procedures, double-exposure and real-time interferometry.

6.8.1 Double-Exposure Interferometry

In this method, two exposures of the object are made on the same hologram. This might be recordings before and after the object has been subject to load by, for instance, external forces or two other object states that are to be compared. By reconstructing the hologram, the two waves scattered from the object in its two states will be reconstructed simultaneously and interfere. This double-exposed hologram can be stored and later reconstructed for analysis of the registered deformation at the time appropriate for the investigator. If a lot of different states of the object (e.g. different load levels) are to be investigated, many holograms have to be recorded, which makes the method time-consuming and elaborate.

6.8.2 Real-Time Interferometry

In this method, a single recording of the object in its reference state is made. Then the hologram is processed and replaced in the same position as in the recording. By looking through the hologram we are now able to observe the interference between the reconstructed object wave and the wave from the real object in its original position. Thus we are able to follow the deformation as it develops in real time by observing the changes in the interference pattern. These changes might be recorded on film for

later playback and analysis. A disadvantage of the method is that the hologram must be replaced in its original position with very high accuracy. This can be overcome by developing the hologram *in situ* in a transparent cuvette or using a thermoplastic film, see Section 5.5.2. Also the contrast of the interference fringes is not as good as in the double-exposure method.

6.8.3 Analysis of Interferograms

As we have seen, holographic interferometry enables the wave scattered from the object in its reference state described by the field amplitude $u_1 = U_1 e^{i\phi_1}$ and the wave scattered from the object in a deformed state described by the field amplitude $u_2 = U_2 e^{i\phi_2}$ to occur simultaneously. The actual deformations will be so small that we can put $U_1 = U_2 = U$. These two waves will form an interference pattern in the usual way given by

$$I = 2U^2[1 + \cos \Delta\phi] \tag{6.30}$$

where

$$\Delta\phi = \phi_1 - \phi_2 \tag{6.31}$$

The problem is then to find the relation between $\Delta\phi$ and the deformation.

Consider Figure 6.6 where a point O on the object is moved along the displacement vector \mathbf{d} to the point O' due to a deformation of the object. The object is illuminated by a plane wave (point source placed at infinity) which propagation direction \mathbf{n}_1 makes an angle θ_1 with the displacement vector \mathbf{d} . Assume that we are looking through the hologram from infinity along the direction \mathbf{n}_2 making an angle θ_2 with \mathbf{d} . We realize that the geometrical path length from the light source via the object point to the point of observation will be different before and after the deformation has taken place. In our case this difference is equal to the path length AO + OB which by applying simple trigonometry becomes equal to

$$d(\cos \theta_1 + \cos \theta_2) \tag{6.32}$$

From Section 1.4 we know that the phase difference $\Delta\phi$ is equal to the path length multiplied by the wave number k :

$$\Delta\phi = kd(\cos \theta_1 + \cos \theta_2) \tag{6.33}$$

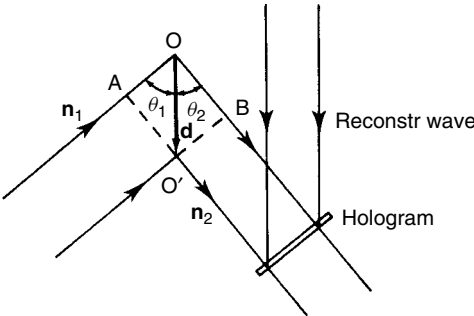


Figure 6.6

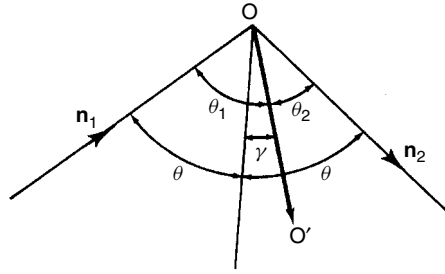


Figure 6.7

In Figure 6.7 a portion of Figure 6.6 is redrawn and the line bisecting the angle 2θ between \mathbf{n}_1 and \mathbf{n}_2 is introduced. This bisector is inclined at an angle γ to d . By using trigonometric formulas we find the geometry factor g to be

$$g = \cos \theta_1 + \cos \theta_2 = 2 \cos \gamma \cos \theta \quad (6.34)$$

which yields

$$\Delta\phi = (2\pi/\lambda)2d(\cos \gamma) \cos \theta \quad (6.35)$$

By inserting Equation (6.35) into Equation (6.30) we find that the interference pattern has a maximum (bright fringe) whenever

$$\Delta\phi = (2\pi/\lambda)2d(\cos \gamma) \cos \theta = n2\pi \quad \text{for } n = 0, 1, 2, \dots$$

i.e. when

$$d \cos \gamma = \frac{n\lambda}{2 \cos \theta} \quad (6.36)$$

and a minimum (dark fringe) whenever

$$\Delta\phi = (2\pi/\lambda)2d(\cos \gamma) \cos \theta = (2n + 1)\pi \quad \text{for } n = 0, 1, 2, \dots$$

i.e. when

$$d \cos \gamma = \frac{2n + 1}{2} \frac{\lambda}{2 \cos \theta} \quad (6.37)$$

Here $d \cos \gamma$ is the component of the displacement vector onto the line bisecting the angle between the illumination and observation directions. This applies also when \mathbf{d} does not lie in the plane defined by \mathbf{n}_1 and \mathbf{n}_2 as in Figure 6.7.

When interpreting interference patterns (also called interferograms) due to deformations of extended objects, we therefore can imagine the space to be filled with equispaced parallel planes which are normal to the bisector of \mathbf{n}_1 and \mathbf{n}_2 with a spacing equal to $\lambda/(2 \cos \theta)$. Each time the surface of the deformed object intersects one of these planes we get a bright (or dark) fringe. To measure the deformation at a given point, one therefore simply has to count the number of fringes and multiply it by $\lambda/(2 \cos \theta)$. This is illustrated

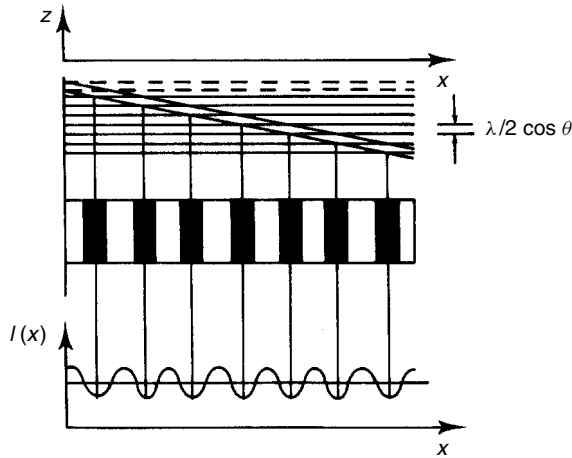


Figure 6.8

in Figure 6.8 which shows the interference pattern seen on a beam that is rotated as a rigid body about an axis (the y-axis).

Highest sensitivity is obtained when both the illumination and observation directions are parallel to the displacement vector, i.e. when $\gamma = \theta_1 = \theta_2 = \theta = 0$. The displacement corresponding to one fringe spacing is then

$$d = \lambda/2$$

from Equation (6.37) we see that the first-order dark fringe ($n = 0$) then occurs at

$$d_{\min} = \lambda/4$$

For the He-Ne laser wavelength $\lambda = 0.6328 \mu\text{m}$, this corresponds to a displacement equal to $0.15 \mu\text{m}$ which therefore gives a representative figure for the sensitivity of standard holographic interferometry.

To obtain the best measurement accuracy when analysing holographic interferograms, the fringe positions should be determined with the highest possible accuracy (or fractional fringe width). This can be achieved in many different ways as will be discussed in chapter 11. Here we mention one method based on the heterodyne principle (see Section 3.6.4) called heterodyne holographic interferometry (HHI) (Dandliker and Thalmann 1985). In this method the two wavefronts corresponding to the object in its two states have different frequencies ν_1 and ν_2 . Equation (6.30) then becomes

$$I = 2U^2\{1 + \cos[2\pi(\nu_1 - \nu_2)t + \Delta\phi]\} \tag{6.38}$$

where $\nu_1 - \nu_2$ corresponds to the intermediate frequency of Equation (3.41) and the analysis to recover $\Delta\phi$ follows the principle described in Section 3.6.4. Such a measurement can be made to an accuracy of typically 10^{-1} radians or $3 \cdot 10^{-3}$ of a cycle and hence enables very small deformations to be resolved. Moreover the measurement is independent of U which affects only the amplitude of the signal. Heterodyning may be used in

both real-time and double exposure holography. In real-time HHI, the 'live' object beam and the reconstructing beam have different frequencies, whilst in double exposure HHI, a single frequency is used to make both exposures but two separate reference beams are employed. At the reconstruction stage these two reference beams are set at different frequencies. In both methods two detectors are required which are used in one of two ways. For the first of these, one of the detectors is tracked across the fringe pattern whilst the other is held static and hence generates the fixed reference frequency. Alternatively the two detectors may be maintained at a fixed distance with respect to one another and both tracked across the fringe field. This enables a differential measurement to be made and hence the variations in the gradient of $\Delta\phi$ is found. In practice the two frequencies are obtained by splitting a single continuous laser beam and passing the two beams through separate optoacoustical modulators.

When the point source and the point of observation are placed at finite distances from the object, the illumination and observation directions (\mathbf{n}_1 and \mathbf{n}_2) will vary across the object surface. This will turn the above-mentioned plane equispaced surfaces into equidistant ellipsoids with the point source and the point of observation as foci. The sensitivity (displacement per fringe) then varies across the object surface but for small objects and reasonably long distances to the source point and the observation point this variation will be quite small. As can be seen from Equation (6.35) and Figure 6.8 the method is insensitive to displacements parallel to the equispaced planes or along the ellipsoids in the general case.

This should be kept in mind when making a set-up for holographic interferometry. The concept called the holo-diagram (which is essentially the ellipsoids mentioned above) developed by Abramson (1970, 1972) could be helpful in this respect.

Figure 6.9 shows some typical examples of interferograms obtained by means of holographic interferometry on different objects. It should be noted that holographic interferometry is incapable of measuring surface contour deviations between two different objects as pointed out in Section 3.5. It is therefore also impossible to measure deformations if the microstructure of the object changes drastically, as for example in plastic deformations.

6.8.4 Localization of Interference Fringes

Another difficulty in evaluating holographic interferograms comes from the phenomenon of fringe localization. It is an annoying fact that the fringes in holographic interferometry only in special cases are localized on the object surface.

This means that when imaging an interferogram by focusing on the object, the fringes may be completely lost because they focus (localize) in a plane (which might be curved) which lies far away from the object surface. That the fringes are localized in a certain plane means that they have maximum contrast or visibility in that plane. Loss of fringes in the plane of the object or other planes therefore means that their contrast is too low to be detected in that plane.

To see this, we must remember that the interferogram is formed by interference between light scattered from pairs of corresponding points on the object surface in the first and second exposure. Interference between non-identical points on the two displaced surfaces will contribute a random noisy background to the interferogram, thereby reducing its contrast.

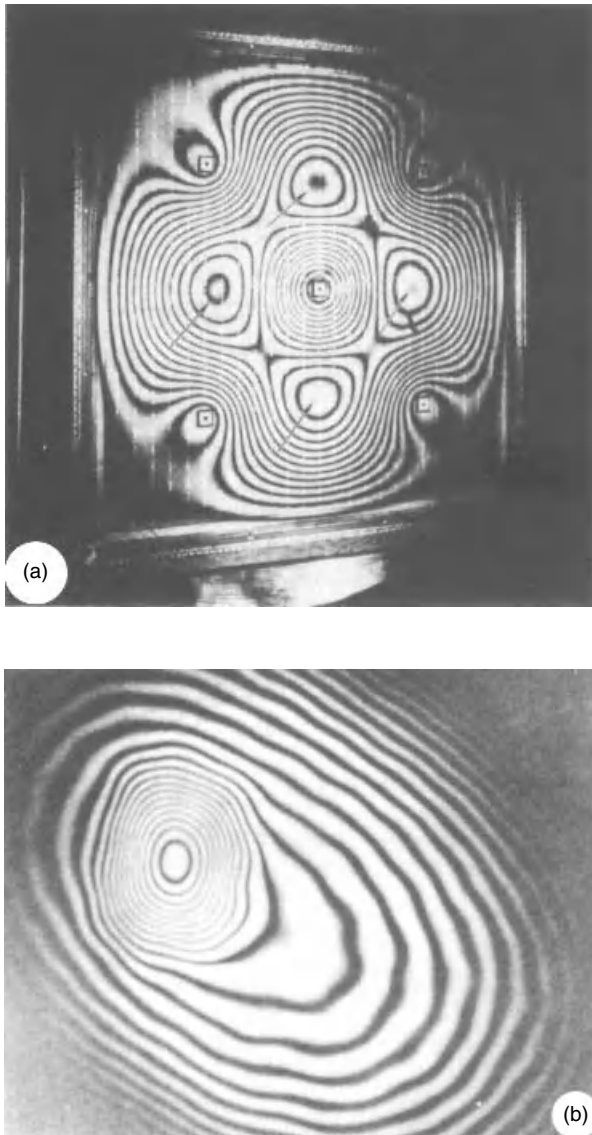
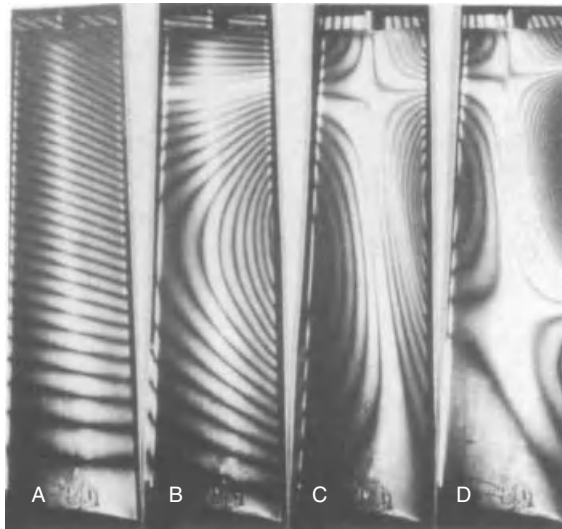
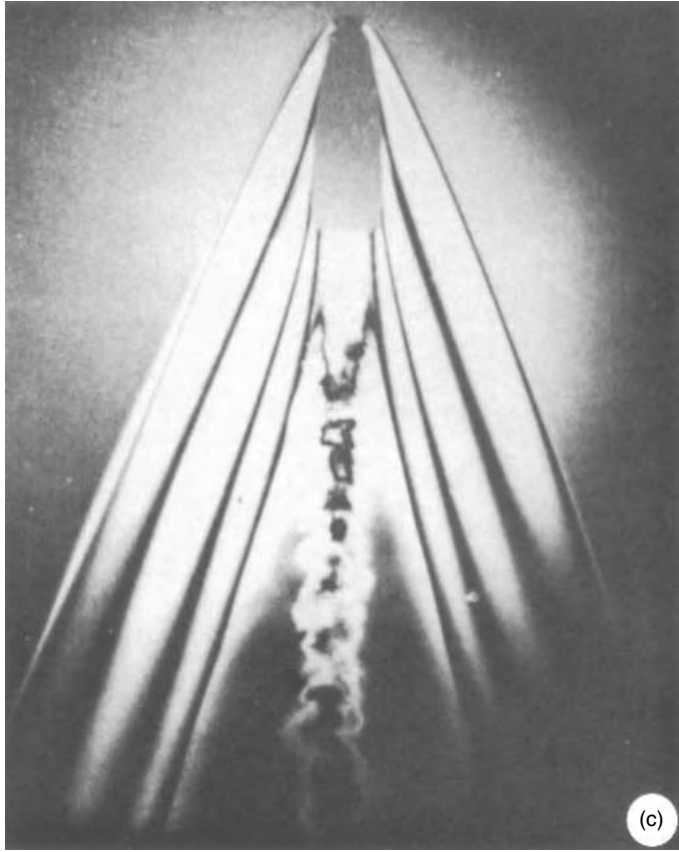


Figure 6.9 Examples of holographic interferograms. (a) Deflection of a rectangular plate fastened with five struts and subjected to a uniform pressure. From Wilson *et al.* 1971. (Photograph courtesy of Dr A. D. Wilson, Thomas J. Watson Research Center, Yorktown Heights, New York. Reproduced by permission of SEM.); (b) Detection of debonded region of a honeycomb construction panel. From Vest 1979. (Reproduced by permission of John Wiley & Sons Inc.); (c) A bullet in flight observed through a doubly-exposed hologram. From Collier *et al.* 1971. (Reproduced by permission of Dr R. F. Wuerker, TRW Inc.); and (d) Holographic reconstruction of a solid turbine blade illustrating (A) the first flexural resonance at 981 Hz, (B) a second-order flexural resonance at 4624 Hz, and the 2nd and 3rd torsional resonances at (C) 6406 Hz. From Erf 1974. (Photograph courtesy of Dr R. K. Erf, United Technologies Research Center, East Hartford, Connecticut. Reproduced by permission of Academic Press, New York)



(d)

Figure 6.9 (continued)

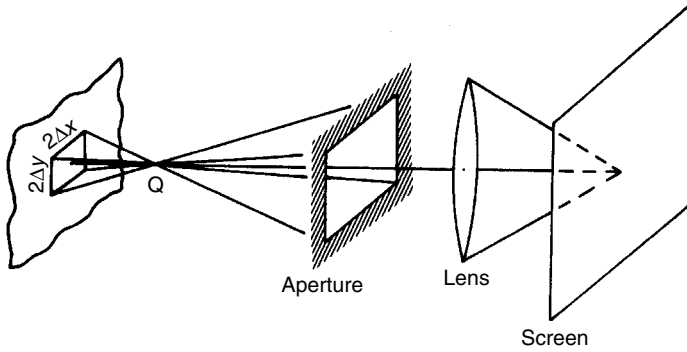


Figure 6.10

In the analysis below, we shall show that this problem might be overcome by stopping down the aperture of the imaging system, thereby increasing the depth of focus to obtain simultaneously an image of sufficient quality of both the fringe pattern and the object surface. For more details, see Vest (1979), Walles (1969) and Molin and Stetson (1971).

Consider Figure 6.10 where the holographic interferogram due to displacement of the object surface (drawn as a single surface in the figure) is imaged onto a viewing screen by a lens through a rectangular aperture. Assume that the lens images a plane containing the point Q onto the viewing screen. The central ray passing through Q emanates from a point P(x₀, y₀) on the object surface. The intensity at Q is the integral of the intensity of all the ray pairs within the ray cone defined by the aperture. This cone subtends a rectangular area of dimension 2Δx, 2Δy on the object surface, hence the intensity at Q is

$$I(Q) = \int_{x_0-\Delta x}^{x_0+\Delta x} \int_{y_0-\Delta y}^{y_0+\Delta y} [1 + \cos \delta(x, y)] dx dy \tag{6.39}$$

where

$$\delta(x, y) = kd(\cos \theta_1 + \cos \theta_2) \tag{6.40}$$

is the phase difference due to the displacement of each point on the object between the two exposures (see Equation (6.33)). For small Δx, Δy, δ(x, y) can be approximated by the initial terms of a Taylor series expansion about x₀, y₀:

$$\begin{aligned} \delta(x, y) &= \delta(x_0, y_0) + (x - x_0) \left. \frac{\partial \delta}{\partial x} \right|_{x_0, y_0} + (y - y_0) \left. \frac{\partial \delta}{\partial y} \right|_{x_0, y_0} \\ &= \delta_0 + (x - x_0) \delta_0^x + (y - y_0) \delta_0^y \end{aligned} \tag{6.41}$$

where δ₀^x and δ₀^y denote partial derivatives of δ at x₀, y₀ with respect to x and y respectively.

Substituting Equation (6.41) into Equation (6.39) and evaluating the integral yields

$$I(Q) = 4\Delta x \Delta y \left[1 + \frac{\sin(\delta_0^x \Delta x)}{\delta_0^x \Delta x} \frac{\sin(\delta_0^y \Delta y)}{\delta_0^y \Delta y} \cos \delta_0 \right] \tag{6.42}$$

The contrast of this intensity distribution is found from the definition of contrast or visibility (cf. Equation (3.8)):

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (6.43)$$

which by putting $\cos \delta_0 = 1$ for I_{\max} and $\cos \delta_0 = -1$ for I_{\min} in Equation (6.42) gives

$$V = \left| \frac{\sin(\delta_0^x \Delta x)}{\delta_0^x \Delta x} \frac{\sin(\delta_0^y \Delta y)}{\delta_0^y \Delta y} \right| \quad (6.44)$$

From this expression we see that the visibility is equal to 1 for δ_0^x or $\delta_0^y = 0$, i.e. when the variations in phase over the viewing cone are minimized. This defines the area of fringe localization. If the value of $\delta_0^x \Delta x$ or $\delta_0^y \Delta y$ increases rapidly as the distance z is changed from the localization distance, the region of fringe localization is sharply defined. If $\delta_0^x \Delta x$ or $\delta_0^y \Delta y$ increases slowly as z is varied, the region of localization is broad. Since δ_0^x and δ_0^y are determined by the system geometry and object displacement field, we see from Equation (6.44) that the sharpness of localization can be controlled by the viewing aperture $\Delta x \Delta y$. Therefore, by decreasing the aperture, relatively distinct fringes can be observed over an extended region and fringes can be seen in the plane of the object even though it is at some distance from the region of localization.

We also see from Equation (6.44) that the contrast is a periodic function of $\delta_0^x \Delta x$, $\delta_0^y \Delta y$. As we move away from the region of localization, the contrast therefore can assume periodic maxima but with a much lower value due to the sinc-function dependence (see Figure B.1.c).

From Equation (6.40) we see that δ is dependent on the illumination and observation directions as well as the displacement vector d . The determination of the regions of fringe localization must therefore be calculated in each separate case by maximizing Equation (6.44). Here we shall not go into such detailed calculations but merely quote two general results which are:

- (1) For a rigid body translation the fringes localize at infinity.
- (2) For a rigid body rotation about an axis inside the object's surface and normal to the illumination and observation directions the fringes localize on the object.

In conclusion, fringe localization rarely poses any problem to the experimentalist since it can be solved by decreasing the viewing aperture. On the other hand, the phenomenon can be taken advantage of since it conveys additional information about the deformation under investigation.

6.9 HOLOGRAPHIC VIBRATION ANALYSIS

Up to now we have been dealing with holographic interferometry applied on static deformations. In this section we shall show that this method can also be applied to vibrating objects.

Assume that the object point in Figure 6.6 executes harmonic vibrations given by

$$d(x, t) = D(x) \cos \omega t \quad (6.45)$$

where $D(x)$ is the amplitude, x represents the spatial coordinates of the object point and ω is the vibration frequency. Light scattered from this point can be described by a field amplitude in the hologram plane given by

$$u_o(x, t) = U_o(x)e^{i\phi} \quad (6.46)$$

where

$$\phi = kgd(x, t). \quad (6.47)$$

Here $g = \cos \theta_1 + \cos \theta_2$ is the geometry factor determined by the illumination and observation directions in the same way as in the static case.

Let u_o of Equation (6.46) be the object wave in a hologram recording and u the reference wave. Just as in the static case, the reconstructed wave will be given by the last term of Equation (6.3). However, since u_o is time-varying during the exposure, the recorded intensity distribution (Equation (6.1)) and hence the hologram amplitude transmittance (Equation (6.2)) will be averaged over the exposure time. The reconstructed object wave therefore becomes equal to

$$u_a = \alpha |u|^2 \bar{u}_o \quad (6.48)$$

where the bar denotes time averaging. For exposure times much longer than the vibrating period of the object (Løkgberg 1979) this time averaging is equivalent to averaging over one vibration period T . Therefore

$$\begin{aligned} u_a &= \alpha |u|^2 \frac{1}{T} \int_0^T u_o(x, t) dt \\ &= \alpha |u|^2 U_o \frac{1}{T} \int_0^T e^{ikgD(x)\cos\omega t} dt = \alpha |u|^2 U_o J_0[kgD(x)] \end{aligned} \quad (6.49)$$

where J_0 is the zeroth-order Bessel function. Here we have applied the relation

$$\frac{1}{2\pi} \int_0^{2\pi} e^{i\eta \cos \xi} d\xi = J_0(\eta) \quad (6.50)$$

The observable intensity distribution of the reconstructed wave becomes

$$I_a = |u_a|^2 = K J_0^2[kgD(x)] \quad (6.51)$$

where all constants are gathered into a common constant K .

As an illustrative example, consider Figure 6.11(a) where a bar is vibrating as a rigid body about an axis. In Figure 6.11(b) we have drawn the Bessel function squared which represents the intensity distribution that will be observed in the reconstruction of a hologram recording of the vibrating bar. We see that the region around the axis which is at rest (the nodal point) will show up as a bright zero-order fringe of much higher intensity than the higher-order bright fringes farther along the bar. This is in contrast to the cos-fringes obtained in the case of a static deformation where all bright fringes have equal intensity.

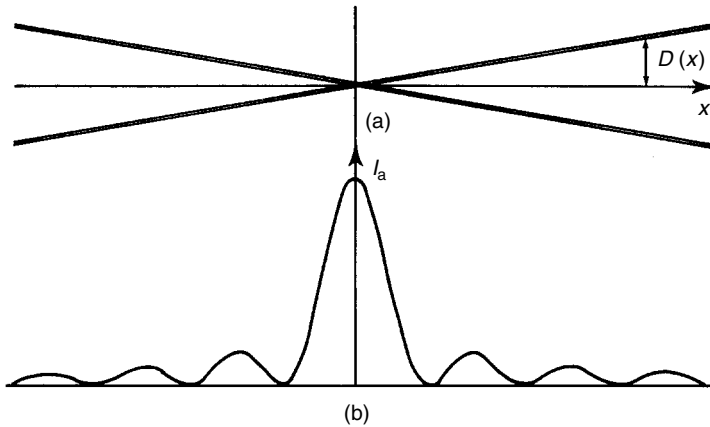


Figure 6.11 (a) Vibration of a bar about an axis and (b) Intensity distribution of the resulting time-average holographic recording

To find the vibration amplitude at the higher-order fringes we consult a table of Bessel function values from which we find that

$$J_0^2(\eta) = \max \text{ for } \eta = 0, 3.83, 7.02, 10.17, 13.32, 16.47, \dots$$

$$J_0^2(\eta) = 0 \text{ for } \eta = 2.40, 5.52, 8.65, 11.79, 14.93, \dots$$

For $g = 2$ (illumination and observation directions parallel to the displacement) and $\lambda = 632.8 \text{ nm}$ (wavelength of the He-Ne laser) this gives:

$$\text{Bright fringes when } D(x) = 0, 0.19, 0.35, 0.51, 0.67, 0.83, \dots [\mu\text{m}]$$

$$\text{Dark fringes when } D(x) = 0.12, 0.28, 0.44, 0.59, 0.75, \dots [\mu\text{m}]$$

By means of standard holographic techniques one has been able to observe fringes up to the 50th order by this method. A very detailed map of the amplitude distribution is therefore obtained. The frequency range of holographic vibration measurements is only limited by the method of object excitation. Using piezoelectric transducers for excitation, values of hundreds of kilohertz are easily obtained.

The method described above is called the time-average method. Another method similar to the real-time method for static displacements can also be applied. It consists of first recording a hologram while the object is at rest, then replacing the hologram in its original position and observing the resulting fringe pattern when the object vibrates. The contrast of this pattern is very low due to the resulting $(1 - J_0)$ -dependence (Vest 1979). While observing this real-time pattern the laser beam can be chopped with the same frequency. This is equivalent to the real-time method for static displacements. Here one observes fringes due to the displacement of the object, when at rest and when illuminated by the light pulse. In that way, by slowly varying the phase between the light pulse and the object vibration, one can observe the vibration of the object in slow motion. Thus it is also possible to observe the phase of the displacements on the different parts of the object surface.

Another method is to phase modulate the reference wave (Alekssoff 1971, 1974). This can be done by placing a vibration mirror in the light path of the reference wave. The argument of the Bessel function of Equation (6.51) then changes from $(4\pi/\lambda)D(x)$ (for $g = 2$) to

$$\frac{4\pi}{\lambda} [D_o^2(x) + R^2 - 2D_o(x)R \cos(\psi_o(x) - \psi_r)]^{1/2} \quad (6.52)$$

where

$D_o(x)$ = the vibrating amplitude of the object,

R = the vibrating amplitude of the reference mirror,

$\psi_o(x) - \psi_r$ = the phase difference between the vibration of the object and reference mirror.

The result of this reference wave modulation is that the centre of the Bessel function is moved from the nodal points of the object to points that vibrate at the same amplitude and phase as the modulating mirror. These points then show up as zero-order bright fringe areas. In that way it is possible to extend the measurable amplitude range considerably, in practice up to about 10 μm . By varying the phase of the reference mirror it is also possible to trace out the areas of the object vibrating in the same phase as the reference mirror, thereby mapping the phase distribution of the object.

Reference mirror modulation can also be used to measure very small vibration amplitudes by moving the steepest part of the central maximum of the Bessel function to coincide with zero object vibration amplitude (Metherell *et al.* 1969; Hogmoen and Løkberg 1976).

Application of a TV camera as the recording medium gives a very versatile instrument for vibration analysis using the methods described above. This will be treated in more detail in Section 12.2.

6.10 HOLOGRAPHIC INTERFEROMETRY OF TRANSPARENT OBJECTS

Up to now we have mainly considered the application of holographic interferometry to opaque objects. The method can, however, equally well be used for the analysis of transparent objects (Vest 1979). In fact, the set-up becomes slightly simpler, for example that shown in Figure 6.12.

The quantity actually measured by this method is the change in refractive index due to some change in the object volume. For three-dimensional objects, the corresponding phase difference is an integrated value through the object volume along the ray path. In the same way as for opaque objects, the resulting interferogram is given as (cf. Equation (6.30))

$$I = 2U^2(1 + \cos \Delta\phi) \quad (6.53)$$

where $\Delta\phi$ is the phase difference in the two recordings. In most applications the refractive index during one exposure, say the first, is uniform and can be denoted by n_0 . Then the difference in the general case is given as

$$\Delta\phi = k \int [n(x, y, z) - n_0] dz \quad (6.54)$$

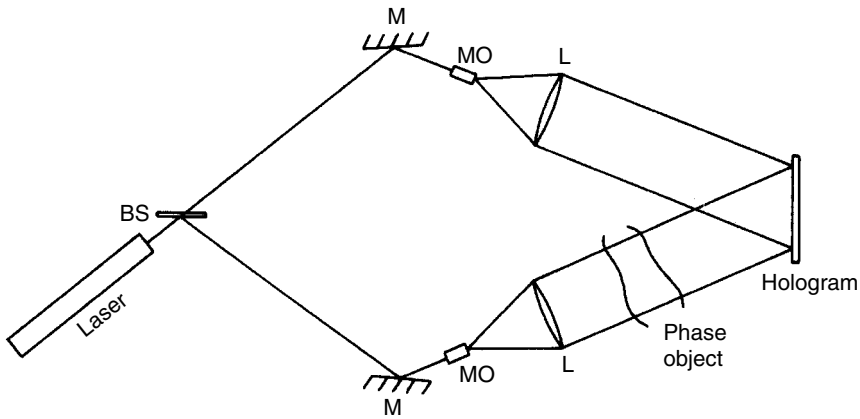


Figure 6.12 Holography set-up for transparent phase objects

where $n(x, y, z)$ is the refractive index distribution during the second exposure and where we have assumed plane wave object illumination along the z -axis. From this equation we see that it is impossible to determine the phase contribution from a single specific point inside the object volume. There are, however, special cases, for example refractive index variations only in the x - and y -directions, radially symmetric distributions, etc. for which Equation (6.54) can be explicitly solved.

Holographic interferometry of transparent phase objects has nevertheless become a versatile tool in research areas such as aerodynamics, heat transfer, plasma diagnostics and stress analysis of transparent models. The latter will be treated in more detail in Section 9.7.

Figure 6.9(c) shows a typical example from the field of aerodynamics. It is a two-exposure holographic interferogram of high speed air flow past a cone. The change in refractive index due to air compression can be found by counting the number of fringes starting from the tip of the cone.

We shall not here go into details about the various techniques used in this field. In fact, a lot of different methods have been applied. For reference, we merely quote some relations between the refractive index and some physical quantities appropriate to the different measurement problems.

In aerodynamics and flow visualization, interferometry is used to determine the distribution of density (the mass per unit volume) in a gas. Density, denoted by ρ , is related to the refractive index of the gas by the Gladstone-Dale equation:

$$n - 1 = K\rho \quad (6.55)$$

where K , the Gladstone-Dale constant, is a property of the gas. K is nearly independent of the wavelength of light and of temperature and pressure under moderate physical conditions.

In heat and mass transfer, interferometry is used to determine the spatial distribution of temperature or concentration of chemical species. For gases, the relation between

temperature and refractive index is found by combining the ideal gas equation of state

$$\rho = \frac{MP}{RT} \quad (6.56)$$

and the Gladstone-Dale equation (Equation (6.55)) to yield

$$n - 1 = \frac{KMP}{RT} \quad (6.57)$$

Where P is the pressure, M is the molecular weight of the gas, T is the absolute temperature and R is the universal gas constant. The slope of a curve of refractive index versus temperature is therefore

$$\frac{dn}{dT} = -\frac{KMP}{RT^2} \quad (6.58)$$

For small temperature changes, the right-hand side of Equation (6.58) is approximately constant, giving a linear relation for the rate of change of refractive index with temperature. For liquids, the density is not a simple function of temperature so empirical relations must be found between n and T .

In plasma diagnostics, interferometry is used to determine the spatial distribution of densities of the plasma which is a collection of atoms, ions and electrons created by very high temperatures. The refractive index of a plasma is the sum of the refractive indices of these particles weighted by their number densities. For atoms and ions, the refractive index is given by the Gladstone-Dale relation, Equation (6.55). For electrons, the refractive index n_e is given as

$$n_e = \left(1 - \frac{N_e e^2 \lambda^2}{2\pi m_e c^2}\right)^{1/2} \quad (6.59)$$

where e is the charge and m_e the mass of an electron, c is the speed of light and N_e is the number density of electrons. We see that this relation is strongly dependent on the wavelength λ .

PROBLEMS

6.1 In Section 3.4 we found an expression for the interfringe distance d_x measured along the x -axis when two plane waves are incident on the xy -plane, see Figure 3.2 and Equation (3.23).

Calculate the interfringe distance along the x -axis when the two plane waves are refracted into a medium (a hologram plate) of refractive index n .

6.2 Imagine that you cut a small piece out of a recorded hologram plate (you might have dropped the hologram glass plate on the floor!) and use this piece in the reconstruction. What do you see? Explain.

6.3 In Section 6.5 we stated that maximum diffraction efficiency for a thin amplitude hologram is obtained when the visibility (or contrast) of the intensity distribution is unity, which means that the ratio $R = I_o/I_r$ between the object and reference intensities is equal to 1.

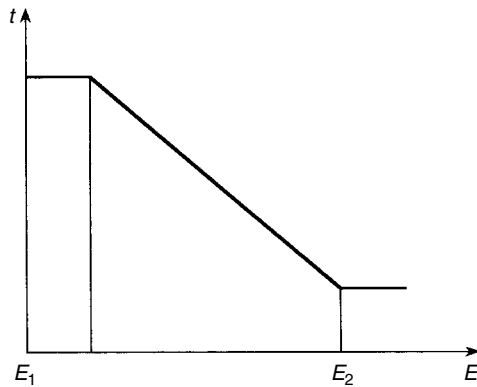


Figure P6.1

The transmittance versus exposure ($t - E$) curve for a holographic film typically looks like that sketched in Figure P6.1 with a linear portion between the exposures E_1 and E_2 . To get a linear response, it is therefore advantageous to have the exposure lying between these values. Calculate R in terms of E_1 and E_2 when the whole linear portion of the $t - E$ curve is utilized.

- 6.4 A thin hologram is recorded with the object- and reference waves being unit amplitude, plane waves with angles of incidence θ_o and θ_r respectively, and with wavelength λ_1 .
 - (a) Calculate the intensity distribution in the hologram plane.
The hologram is reconstructed with a plane wave with angle of incidence θ_i and wavelength λ_2 .
 - (b) Use the grating equation (Equation (4.21)) to find an expression for the angle θ_s of the reconstructed wave and the ray angle magnification $M_\alpha \simeq \sin \theta_s / \sin \theta_o$.
- 6.5 As mentioned in Section 5.5.2, a thermoplastic film has a bandlimited spatial frequency response centred at about 1500 lines/mm. What should preferably be the angle between the object- and reference waves when using this film?
- 6.6 The transmittance of a phase hologram is given by Equation (6.23). Do you get a single reconstructed wave? What happens to the zero-order wave as M approaches 2.4?
- 6.7 (a) Consider the interferogram in Figure 6.9(a). If we have $\theta_1 = 10^\circ$, $\theta_2 = 20^\circ$ and an He-Ne laser is used, what is the maximum deflection of the plate?
 - (b) Assuming $g = 2$ and $\lambda = 632.8$ nm, what is the maximum vibration amplitude on the left and right side of the turbine blade in B of Figure 6.9(d) (Consult a table of Bessel functions).
- 6.8 Consider the vibrating bar in Figure 6.11(a) and let the angular vibrating amplitude be α_m . Assume that we record a time-average hologram with a modulated reference wave with a vibrating amplitude of the reference mirror equal to R and the phase $\psi_r = \pi$. Sketch (qualitatively) the intensity distribution of the resulting interferogram.