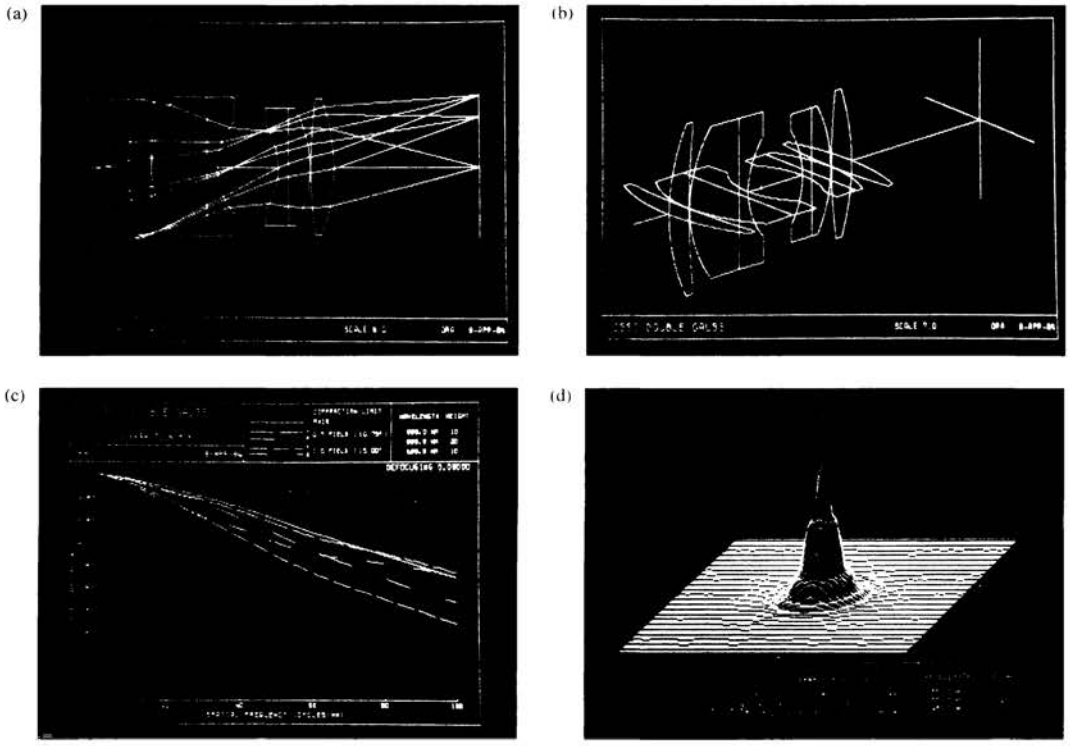


Figure 11.50 An example of the kind of lens design information available via computer techniques. (Photos courtesy Optical Research Associates.)

This discussion is only an introductory one designed more as a strong foundation than a complete structure. There are many other insights to be explored, such as the relationship between the autocorrelation of the pupil function and the OTF, and from there, the means of computing and measuring transfer functions (Fig. 11.50)—but for this the reader is directed to the literature.*

*See the series of articles "The Evolution of the Transfer Function," by F. Abbott, beginning in March 1970 in *Optical Spectra*; the articles "Physical Optics Notebook," by G. B. Parrent, Jr., and B. J. Thompson, beginning in December 1964, in the *S.P.I.E. Journal*, Vol. 3; or "Image Structure and Transfer," by K. Sayanagi, 1967, available from the Institute of Optics, University of Rochester. A number of books are worth consulting for practical emphasis, e.g., *Modern Optics*, by E. Brown; *Modern Optical Engineering*, by W. Smith; and *Applied Optics*, by L. Levi. In all of these, be careful of the sign convention in the transforms.

PROBLEMS

Complete solutions to all problems—except those with an asterisk—can be found in the back of the book.

11.1 Determine the Fourier transform of the function

$$E(x) = \begin{cases} E_0 \sin k_p x & |x| < L \\ 0 & |x| > L \end{cases}$$

Make a sketch of $\mathcal{F}\{E(x)\}$. Discuss its relationship to Fig. 11.11.

11.2* Determine the Fourier transform of

$$f(x) = \begin{cases} \sin^2 k_p x & |x| < L \\ 0 & |x| > L \end{cases}$$

Make a sketch of it.

11.3 Determine the Fourier transform of

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

$$f(t) = \begin{cases} \cos^2 \omega_p t & |t| < T \\ 0 & |t| > T \end{cases}$$

Make a sketch of $F(\omega)$, then sketch its limiting form as $T \to \pm\infty$.

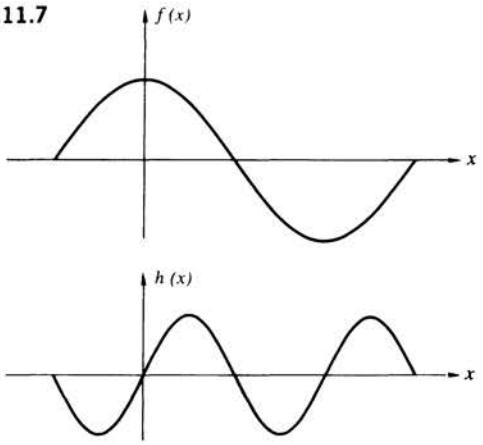
11.4* Show that $\mathcal{F}\{1\} = 2\pi\delta(k)$.

11.5* Determine the Fourier transform of the function $f(x) = A\cos k_0 x$.

11.6 Given that $\mathcal{F}\{f(x)\} = F(k)$ and $\mathcal{F}\{h(x)\} = H(k)$, if a and b are constants, determine $\mathcal{F}\{af(x) + bh(x)\}$.

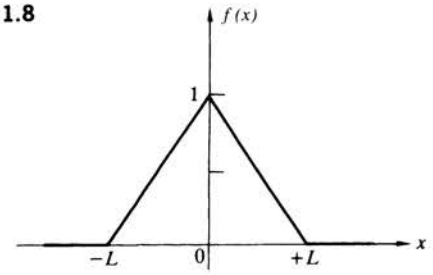
11.7* Figure P.11.7 shows two periodic functions, $f(x)$ and $h(x)$, which are to be added to produce $g(x)$. Sketch $g(x)$; then draw diagrams of the real and imaginary frequency spectra, as well as the amplitude spectra for each of the three functions.

Figure P.11.7



11.8 Compute the Fourier transform of the triangular pulse shown in Fig. P.11.8. Make a sketch of your answer, labeling all the pertinent values on the curve.

Figure P.11.8



11.9* Given that $\mathcal{F}\{f(x)\} = F(k)$, introduce a constant scaling factor $1/a$ and determine the Fourier transform of $f(x/a)$. Show that the transform of $f(-x)$ is $F(-k)$.

11.10* Show that the Fourier transform of the transform, $\mathcal{F}\{f(x)\}$, equals $2\pi f(-x)$, and that this is not the inverse transform of the transform, which equals $f(x)$. This problem was suggested by Mr. D. Chapman while a student at the University of Ottawa.

11.11* The rectangular function is often defined as

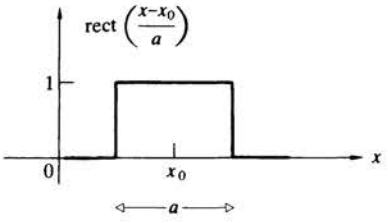
$$\text{rect}\left|\frac{x - x_0}{a}\right| = \begin{cases} 0, & |(x - x_0)/a| > \frac{1}{2} \\ \frac{1}{2}, & |(x - x_0)/a| = \frac{1}{2} \\ 1, & |(x - x_0)/a| < \frac{1}{2} \end{cases}$$

where it is set equal to $\frac{1}{2}$ at the discontinuities (Fig. P.11.11). Determine the Fourier transform of

$$f(x) = \text{rect}\left|\frac{x - x_0}{a}\right|$$

Notice that this is just a rectangular pulse, like that in Fig. 11.1*b*, shifted a distance x_0 from the origin.

Figure P.11.11



11.12* With the last two problems in mind, show that $\mathcal{F}\{(1/2\pi)\,\text{sinc}\,(\frac{1}{2}x)\} = \text{rect}(k)$, starting with the knowledge that $\mathcal{F}\{\text{rect}(x)\} = \text{sinc}\,(\frac{1}{2}k)$, in other words, Eq. (7.58) with $L = a$, where $a = 1$.

11.13* Utilizing Eq. (11.38), show that $\mathcal{F}^{-1}\{\mathcal{F}\{f(x)\}\} = f(x)$.

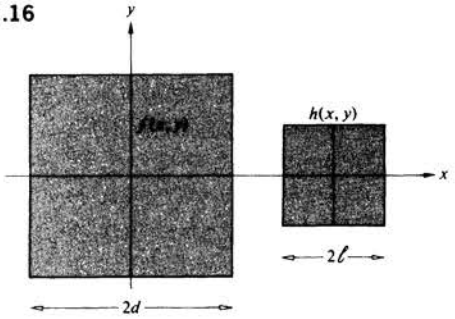
11.14* Given $\mathcal{F}\{f(x)\}$, show that $\mathcal{F}\{f(x - x_0)\}$ differs from it only by a linear phase factor.

11.15 Prove that $f \circledast h = h \circledast f$ directly. Now do it using the convolution theorem.

11.16* Suppose we have two functions, $f(x, y)$ and $h(x, y)$, where both have a value of 1 over a square region in the xy-plane and are zero everywhere else (Fig. P.11.16). If $g(X, Y)$ is their convolution, make a plot of $g(X, 0)$.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

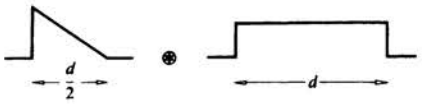
Figure P.11.16



11.17 Referring to the previous problem, justify the fact that the convolution is zero for $|X| \geq d + \ell$ when h is viewed as a spread functio00n.

11.18* Use the method illustrated in Fig. 11.23 to convolve the two functions depicted in Fig. P.11.18.

Figure P.11.18



11.19 Given that $f(x) \circledast h(x) = g(X)$, show that after shifting one of the functions an amount x_0, we get $f(x - x_0) \circledast h(x) = g(X - x_0)$.

11.20* Prove analytically that the convolution of any function $f(x)$ with a delta function, $\delta(x)$, generates the original function $f(X)$. You might make use of the fact that $\delta(x)$ is even.

11.21 Prove that $\delta(x - x_0) \circledast f(x) = f(X - x_0)$ and discuss the meaning of this result. Make a sketch of two appropriate functions and convolve them. Be sure to use an asymmetrical $f(x)$.

11.22* Show that $\mathcal{F}\{f(x) \cos k_0 x\} = [F(k - k_0) + F(k + k_0)]/2$ and that $\mathcal{F}\{f(x) \sin k_0 x\} = [F(k - k_0) - F(k + k_0)]/2i$.

11.23* Figure P.11.23 shows two functions. Convolve them graphically and draw a plot of the result.

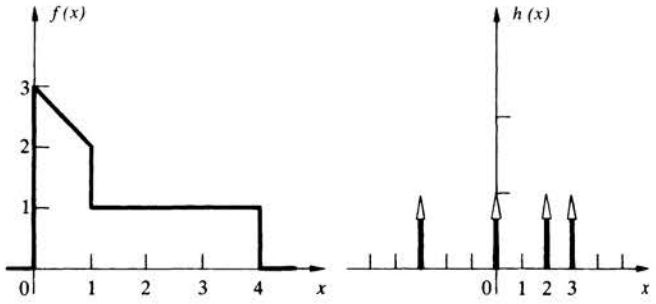
11.24 Given the function

$$f(x) = \text{rect}\left|\frac{x - a}{a}\right| + \text{rect}\left|\frac{x + a}{a}\right|$$

determine its Fourier transform. (See Problem 11.11.)

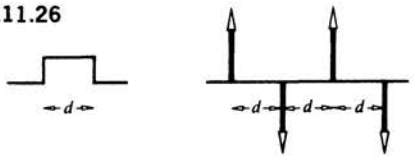
11.25 Given the function $f(x) = \delta(x + 3) + \delta(x - 2) + \delta(x - 5)$, convolve it with the arbitrary function $h(x)$.

Figure P.11.23



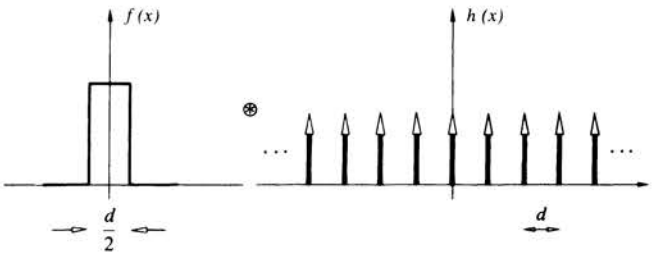
11.26* Make a sketch of the function arising from the convolution of the two functions depicted in Fig. P.11.26.

Figure P.11.26



11.27* Figure P.11.27 depicts a *rect* function (as defined above) and a periodic *comb* function. Convolve the two to get $g(x)$. Now sketch the transform of each of these functions against spatial frequency $k/2\pi = 1/\lambda$. Check your results with the convolution theorem. Label all the relevant points on the horizontal axes in terms of d—like the zeros of the transform of $f(x)$.

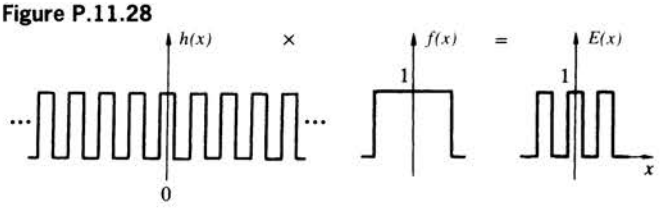
Figure P.11.27



11.28 Figure P.11.28 shows, in one dimension, the electric field across an illuminated aperture consisting of several opaque bars forming a grating. Considering it to be created by taking the product of a periodic rectangular wave $h(x)$ and a unit rectangular function $f(x)$, sketch the resulting electric field in the Fraunhofer region.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Figure P.11.28



11.29 Show (for normally incident plane waves) that if an aperture has a center of symmetry (i.e., if the aperture function is even), then the diffracted field in the Fraunhofer case also possesses a center of symmetry.

11.30 Suppose a given aperture produces a Fraunhofer field pattern $E(Y, Z)$. Show that if the aperture's dimensions are altered such that the aperture function goes from $\mathcal{A}(y, z)$ to $\mathcal{A}(\alpha y, \beta z)$, the newly diffracted field will be given by

$$E'(Y, Z) = \frac{1}{\alpha\beta} E\left(\frac{Y}{\alpha}, \frac{Z}{\beta}\right)$$

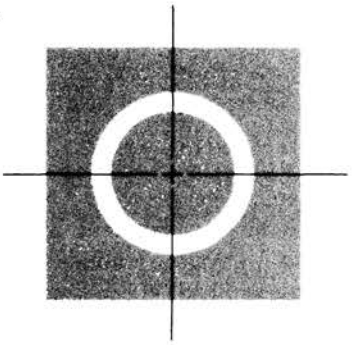
11.31 Show that when $f(t) = A \sin(\omega t + \varepsilon)$, $C_{ff}(\tau) = (A^2/2)\cos\omega t$, which confirms the loss of phase information in the autocorrelation.

11.32 Suppose we have a single slit along the y-direction of width b where the aperture function is constant across it at a value of $\mathcal{A}_0$. What is the diffracted field if we now apodize the slit with a cosine function amplitude mask? In other words, we cause the aperture function to go from $\mathcal{A}_0$ at the center to 0 at $\pm b/2$ via a cosinusoidal dropoff.

11.33* Show, from the integral definitions, that $f(x) \odot g(x) = f(x) \circledast g(-x)$.

11.34* Figure P.11.34 shows a transparent ring on an otherwise opaque mask. Make a rough sketch of its autocorrelation function, taking l to be the center-to-center separation against which you plot that function.

Figure P.11.34



11.35* Consider the function in Fig. 11.35 as a cosine carrier multiplied by an exponential envelope. Use the frequency convolution theorem to evaluate its Fourier transform.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

12 Basics of Coherence Theory

Thus far in our discussion of phenomena involving the superposition of waves, we've restricted the treatment to that of either completely coherent or completely incoherent disturbances. This was done primarily as a mathematical convenience, since, as is quite often the case, the extremes in a physical situation are the easiest to deal with analytically. In fact, both of these limiting conditions are more conceptual idealizations than actual physical realities. There is a middle ground between these antithetic poles, which is of considerable contemporary concern—the domain of ***partial coherence***. Even so, the need for extending the theoretical structure is not new; it dates back at least to the mid-1860s, when Emile Verdet demonstrated that a primary source commonly considered to be incoherent, such as the Sun, could produce observable fringes when it illuminated the closely spaced pinholes ($\lesssim$0.05 mm) of Young's Experiment (Section 9.3). Theoretical interest in the study of partial coherence lay dormant until it was revived in the 1930s by P. H. van Cittert and later by Fritz Zernike. And as the technology flourished, advancing from traditional light sources, which were essentially optical frequency noise generators, to the laser, a new practical impetus was given the subject. Moreover, the recent advent of individual-photon detectors has made it possible to examine related processes associated with the corpuscular aspects of the optical field.

Optical coherence theory is currently an area of active research. Thus, even though much of the excitement in the field is associated with material beyond the level of this book, we shall nonetheless introduce some of the basic ideas.

12.1 Introduction

Earlier (Section 7.10) we evolved the highly useful picture of quasimonochromatic light as resembling a series of randomly phased finite wavetrains (Fig. 7.38). Such a disturbance is nearly sinusoidal, although the frequency does vary slowly (in comparison to the rate of oscillation, 10^{15} Hz) about some mean value. Moreover, the amplitude fluctuates as well, but this too is a comparatively slow variation. The average constituent wavetrain exists roughly for a time Δt_c, which is the *coherence time* given by the inverse of the frequency bandwidth $\Delta\nu$.

It is often convenient, even if artificial, to divide coherence effects into two classifications, **temporal** and **spatial** (p. 390). *The former relates directly to the finite bandwidth of the source, the latter to its finite extent in space.*

To be sure, if the light were monochromatic, $\Delta\nu$ would be zero, and Δt_c infinite, but this is, of course, unattainable. However, over an interval much shorter than Δt_c an actual wave behaves essentially as if it were monochromatic. In effect, the coherence time is *the temporal interval over which we can reasonably predict the phase of the lightwave at a given point in space.* This then is what is meant by **temporal coherence**; namely, if Δt_c is large, the wave has a high degree of temporal coherence and vice versa.

The same characteristic can be viewed somewhat differently. To that end, imagine that we have two separate points P_1' and P_2' lying on the same radius drawn from a quasimonochromatic point source (see Fig. 9.4). If the coherence length, $c\Delta t_c$, is much larger than the distance (r_{12}) between P_1' and P_2', then a single wavetrain can easily extend over the whole separation. The disturbance at P_1' would then be highly correlated with the disturbance occurring at P_2'. On the other hand, if this longitudinal separation were much greater than the coherence length, many wavetrains, each with an unrelated phase, would span the gap r_{12}. In that case, the disturbances at the two points in space would be independent at any given time. The degree to which a correlation exists is sometimes spoken of alternatively as the amount of *longitudinal coher-*

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

ence. Whether we think in terms of coherence time (Δt_c) or coherence length ($c\Delta t_c$), the effect still arises from the finite bandwidth of the source.

The idea of **spatial coherence** is most often used to describe effects arising from the finite spatial extent of ordinary light sources. Suppose then that we have a classical broad monochromatic source. Two point radiators on it, separated by a lateral distance that is large compared with λ, will presumably behave quite independently. That is to say, there will be a lack of correlation existing between the phases of the two emitted disturbances. Extended sources of this sort are generally referred to as incoherent, but this description is somewhat misleading, as we shall see in a moment. Usually one is interested not so much in what is happening on the source itself but rather in what is occurring within some distant region of the radiation field. The question to be answered is really: How do the nature of the source and the geometrical configuration of the situation relate to the resulting phase correlation between two laterally spaced points in the light field?

This brings to mind Young's Experiment, in which a primary monochromatic source S illuminates two pinholes in an opaque screen. These in turn serve as secondary sources, S_1 and S_2, to generate a fringe pattern on a distant plane of observation, Σ_o (Fig. 9.8). We already know that if S is an idealized point source, the wavelets issuing from any set of apertures S_1 and S_2 on Σ_a will maintain a constant relative phase; they will be precisely correlated and therefore coherent. A well-defined array of stable fringes results, and the field is spatially coherent. At the other extreme, if the pinholes are illuminated by separate thermal sources (even with narrow bandwidths), no correlation exists; no fringes will be observable with existing detectors, and the fields at S_1 and S_2 are said to be incoherent. The generation of interference fringes is then seemingly a very convenient measure of the coherence.

We can gain some important insights into the process by returning to the general considerations of Section 9.1 and Eq. (9.7). Imagine two scalar waves $E_1(t)$ and $E_2(t)$ traveling toward, and overlapping at, point P, as in Fig. 9.2. If the light is monochromatic and both beams have the same frequency, the resulting interference pattern will depend on their relative phase at P. If the waves are in-phase, $E_1(t)E_2(t)$ will be positive for all t as the fields rise and fall in together. Hence, $I_{12} = 2\langle E_1(t)E_2(t)\rangle_T$ will be a nonzero positive number, and the net irradiance I will exceed $I_1 + I_2$. Similarly, if the lightwaves are completely out-of-phase, one will be positive when the other is negative, with the result that the product $E_1(t)E_2(t)$ will always be negative, yielding a negative interference term I_{12}, and the result that I will be less than $I_1 + I_2$. In both cases, the product of the two fields moment by moment is oscillatory, but it is nonetheless either totally positive or negative and so averages in time to a nonzero value.

Now consider the more realistic case in which the two lightwaves are quasimonochromatic, resembling the disturbance in Fig. 7.38, which has a finite coherence length. If we again form the product $E_1(t)E_2(t)$, we see in Fig. 12.1 that it varies in time, drifting from negative to positive values. Accordingly, the interference term $\langle E_1(t)E_2(t)\rangle_T$, which is averaged over a relatively long interval compared with the periods of the waves, will be quite small, if not zero: $I \approx I_1 + I_2$. In other words, insofar as the two lightwaves are uncorrelated in their risings and fallings, they will not preserve a constant phase relationship, they will not be completely coherent, and they will not produce the ideal high-contrast interference pattern considered in Chapter 9. We should be reminded here

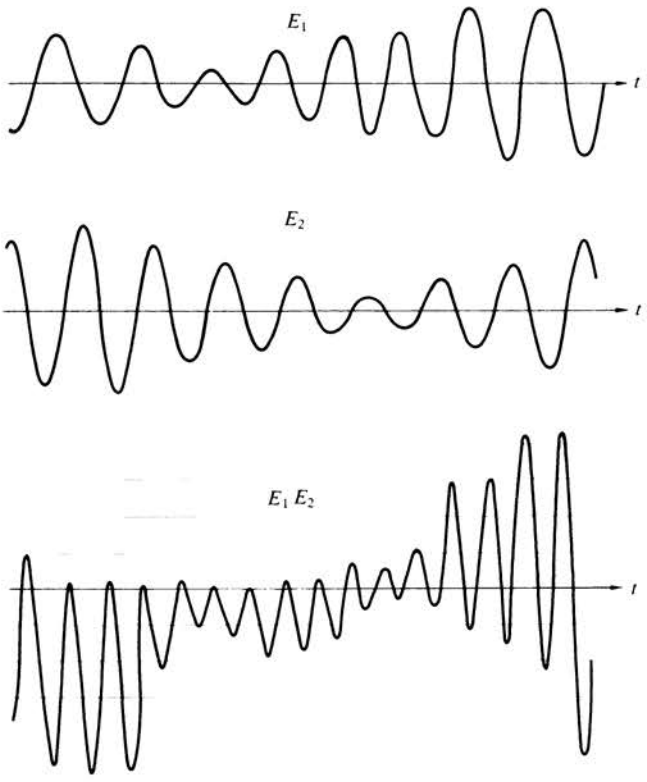


Figure 12.1 Two overlapping E-fields and their product as functions of time. The more uncorrelated the fields, the more nearly the product will average to zero.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

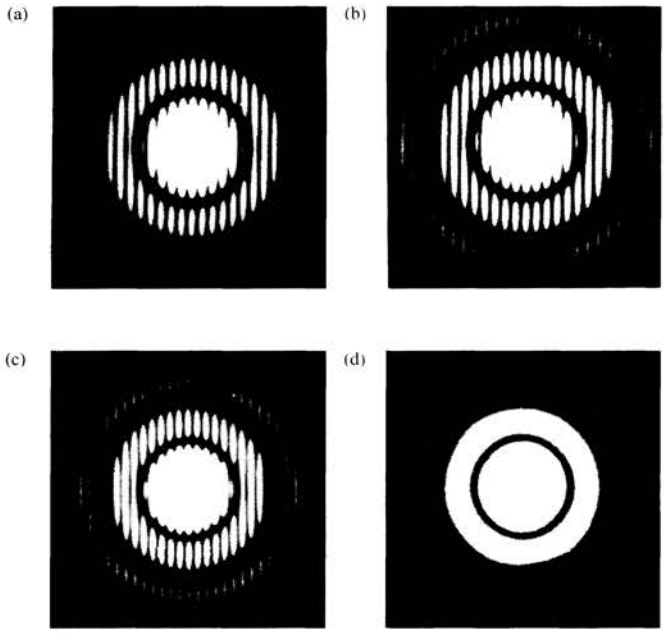


Figure 12.2 Double-beam interference from a pair of circular apertures. (*a*) He–Ne laserlight illuminating the holes. (*b*) Laserlight once again but now a glass plate, 0.5 mm thick, is covering one of the holes. (*c*) Fringes with collimated mercury-arc illumination but no glass plate. (*d*) This time the fringes disappear when the plate is inserted using mercury light. [From B. J. Thompson, *J. Soc. Photo. Inst. Engr.* **4**, 7 (1965).]

of Eq. (11.87), which expresses the cross-correlation of two functions—with $\tau = 0$. Indeed, if P is shifted in space (e.g., along the plane of observation in Young's Experiment), thereby introducing a relative time delay of τ between the two lightwaves, then the interference term becomes $\langle E_1(t)E_2(t + \tau)\rangle_T$, which is the cross-correlation. Coherence is correlation, a point that will be made formally in Section 12.3.

Young's Experiment can also be used to demonstrate temporal coherence effects with a finite bandwidth source. Figure 12.2*a* shows the fringe patterns obtained with two small circular apertures illuminated by a He–Ne laser. Before the photograph in Fig. 12.2*b* was taken, an optically flat piece of glass, 0.5 mm thick, was positioned over one of the pinholes (say S_1). No change in the form of the pattern (other than a shift in its location) is evident, because the coherence length of the laser light far exceeds the optical path length difference introduced by the glass. On the other hand, when the same experiment is repeated using the light from a collimated mercury arc [(*c*) and (*d*) in Fig. 12.2], the fringes disappear. Here the coherence length is short enough and the additional optical path length difference of the glass is long enough for uncorrelated wavetrains from the two apertures to arrive at the plane of observation. In other words, of any two coherent wavetrains that leave S_1 and S_2, the one from S_1 is now delayed so long in the glass that it falls completely behind the other and arrives at Σ_o to meet a totally different wavetrain from S_2.

In both cases of temporal and spatial coherence we are really concerned with one phenomenon, namely, the correlation between optical disturbances. That is, we are generally interested in determining the effects arising from relative fluctuations in the fields at two points in space–time. Admittedly, the term *temporal coherence* seems to imply an effect that is exclusively temporal. However, it relates back to the finite extent of the wavetrain in either space or time, and some people even prefer to refer to it as *longitudinal spatial* rather than temporal coherence. Even so, it does depend intrinsically on the stability of phase in time, and accordingly we will continue to use the term *temporal coherence*. Spatial coherence, or if you will, *lateral spatial coherence*, is perhaps easier to appreciate because it's so closely related to the concept of the wavefront. Thus if two laterally displaced points reside on the same wavefront at a given time, the fields at those points are spatially coherent (see Section 12.3.1).

12.2 Visibility

The quality of the fringes produced by an interferometric system can be described quantitatively using the **visibility** $\mathcal{V}$, which, as first formulated by Michelson, is given by

$$\mathcal{V}(\vec{r}) \equiv \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{12.1}$$

This is identical to the *modulation* of Eq. (11.89). Here $I_{\max}$ and $I_{\min}$ are the irradiances corresponding to the maximum and adjacent minimum in the fringe system. If we set up Young's Experiment, we could vary the separation of the apertures or the size of the primary incoherent quasimonochromatic source, measure $\mathcal{V}$ as it changes in turn, and then relate all this to the idea of coherence. An analytic expression can be derived for the flux-density distribution with the aid of Fig. 12.3.* Here we use a lens L to localize the fringe pattern

*This treatment in part follows that given by Towne in Chapter 11 of *Wave Phenomena*. See Klein, *Optics*, Section 6.3, or Problem 12.6 for different versions

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

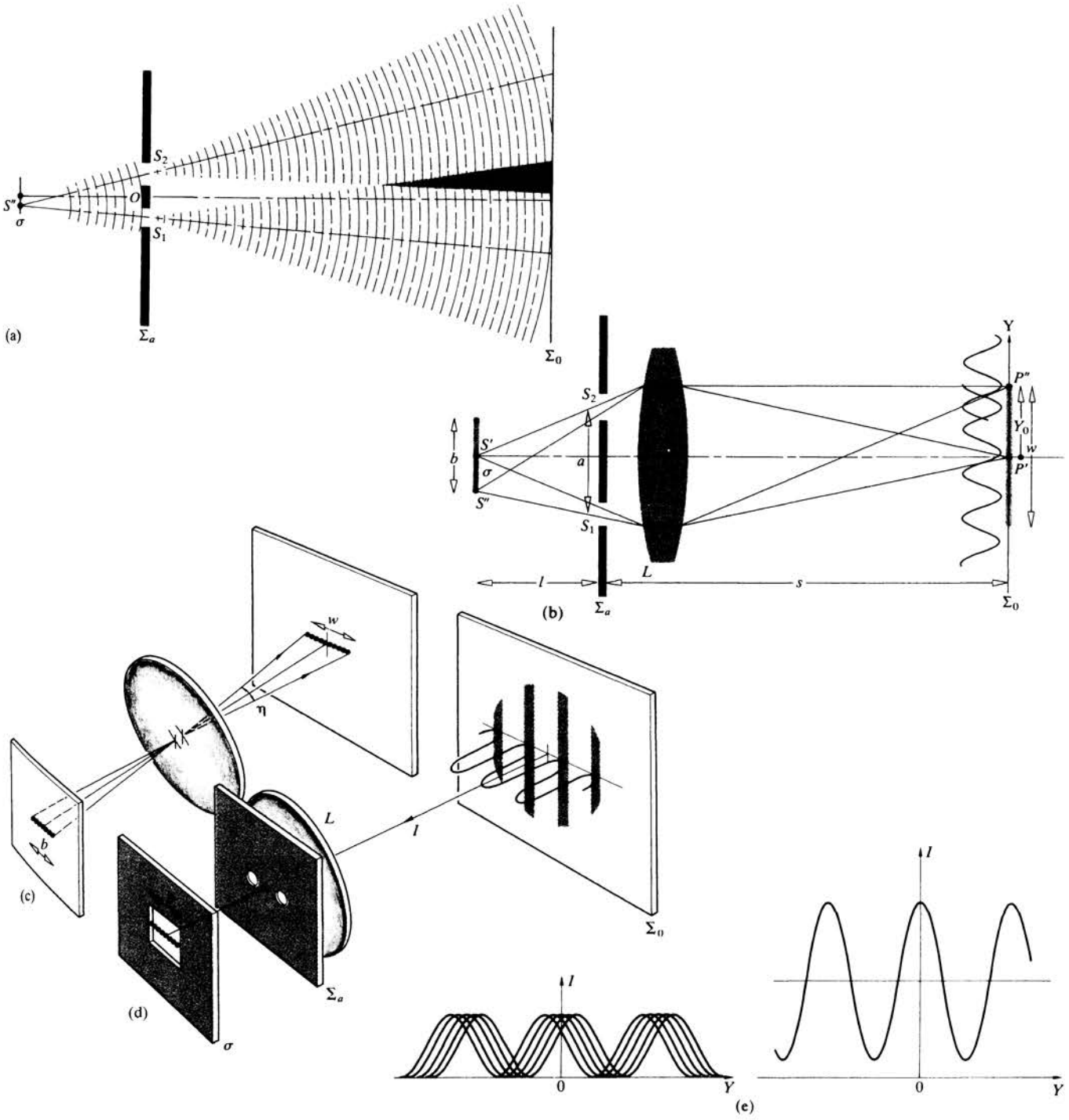


Figure 12.3 Young's Experiment with an extended slit source. (e) A simple representation of how shifted fringes with the same spatial frequency overlap and combine to form a net disturbance of that same spatial frequency with a reduced visibility (see Fig. 7.9).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

more effectively, that is, to make the cones of light diffracted by the finite pinholes more completely overlap on the plane Σ_o. A point source S' located on the central axis would generate the usual pattern given by

$$I = 4I_0 \cos^2\left(\frac{Ya\pi}{s\lambda}\right) \tag{12.2}$$

from Section 9.3. Similarly, a point source above or below S' and lying on a line normal to the line $\overline{S_1S_2}$ would generate the same straight band fringe system slightly displaced in a direction parallel to the fringes. Thus replacing S' by an incoherent line source (normal to the plane of the drawing) effectively just increases the amount of light available. This is something we presumably already knew. In contrast, an off-axis point source, at say S'', will generate a pattern centered about P'', its image point on Σ_o in the absence of the aperture screen. A "spherical" wavelet leaving S'' is focused at P''; thus all rays from S'' to P'' traverse equal optic paths, and the interference must be constructive; in other words, the central maximum appears at P''. The path difference $\overline{S_1P''} - \overline{S_2P''}$ accounts for the displacement $\overline{P'P''}$. Consequently, S'' produces a fringe system identical to that of S' but shifted by an amount $\overline{P'P''}$ with respect to it. Since these source points are incoherent, their irradiances add on Σ_o rather than their field amplitudes (Fig. 12.3*e*).

The pattern arising from a broad source having a rectangular aperture of width b can be determined by finding the irradiance due to an incoherent continuous line source parallel to $\overline{S_1S_2}$. Notice, in Fig. 12.3*b*, that the variable Y_0 describes the location of any point on the image of the source when the aperture screen is absent. With Σ_a in place, each differential element of the line source will contribute a fringe system centered about its own image point, a distance Y_0 from the origin on Σ_o. Moreover, its contribution to the flux-density pattern dI is proportional to the differential line element or, more conveniently, to its image, dY_0, on Σ_o. Thus, using Eq. (9.31), the contribution to the total irradiance arising from dY_0 becomes

$$dI = A\, dY_0 \cos^2\left[\frac{a\pi}{s\lambda}(Y - Y_0)\right] \tag{12.3}$$

where A is an appropriate constant. This, in analogy to Eq. (12.2), is the expression for an entire fringe system of minute irradiance centered at Y_0 contributed by the tiny piece of the source whose image corresponds to dY_0 at Y_0. By integrating over the extent w of the image of the line source, we effectively integrate over the source and get the entire pattern:

$$I(Y) = A\int_{-w/2}^{+w/2} \cos^2\left[\frac{a\pi}{s\lambda}(Y - Y_0)\right] dY_0 \tag{12.4}$$

After a good bit of straightforward trigonometric manipulation, this becomes

$$I(Y) = \frac{Aw}{2} + \frac{A}{2}\frac{s\lambda}{a\pi}\sin\left(\frac{a\pi}{s\lambda}w\right)\cos\left(2\frac{a\pi}{s\lambda}Y\right) \tag{12.5}$$

The irradiance oscillates about an average value of $\bar{I} = Aw/2$, which increases with w, which in turn increases with the width of the source slit. Accordingly,

$$\frac{I(Y)}{\bar{I}} = 1 + \left(\frac{\sin a\pi w/s\lambda}{a\pi w/s\lambda}\right)\cos\left(2\frac{a\pi}{s\lambda}Y\right) \tag{12.6}$$

or

$$\frac{I(Y)}{\bar{I}} = 1 + \operatorname{sinc}\left(\frac{a\pi w}{s\lambda}\right)\cos\left(2\frac{a\pi}{s\lambda}Y\right) \tag{12.7}$$

It follows that the extreme values of the relative irradiance are given by

$$\frac{I_{\max}}{\bar{I}} = 1 + \left|\operatorname{sinc}\left(\frac{a\pi w}{s\lambda}\right)\right| \tag{12.8}$$

and

$$\frac{I_{\min}}{\bar{I}} = 1 - \left|\operatorname{sinc}\left(\frac{a\pi w}{s\lambda}\right)\right| \tag{12.9}$$

When w is very small in comparison to the fringe width $(s\lambda/a)$, the sinc function (p. 49) approaches 1 and $I_{\max}/\bar{I} = 2$, while $I_{\min}/\bar{I} = 0$ (see Fig. 12.4). As w increases, $I_{\min}$ begins to differ from zero, and the fringes lose contrast until they finally vanish entirely at $w = s\lambda/a$. Between the arguments of π and 2π (i.e., $w = s\lambda/a$ and $w = 2s\lambda/a$), the sinc is negative. As the primary slit source widens beyond $w = s\lambda/a$, the fringes reappear but are shifted in phase; in other words, whereas previously there was a maximum at $Y = 0$, now there will be a minimum.

In actuality, the light diffracted by the apertures is localized (Section 10.2) so that the fringe system does not continue out uniformly indefinitely as Y increases. Instead, the pattern of Fig. 12.4*a* will look more like Fig. 12.5.

As a rule, the extent of the source (b) and the separation of the slits (a) are very small compared with the distances between the screens (l) and (s), and consequently we can make some simplifying approximations. While the above consider-

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

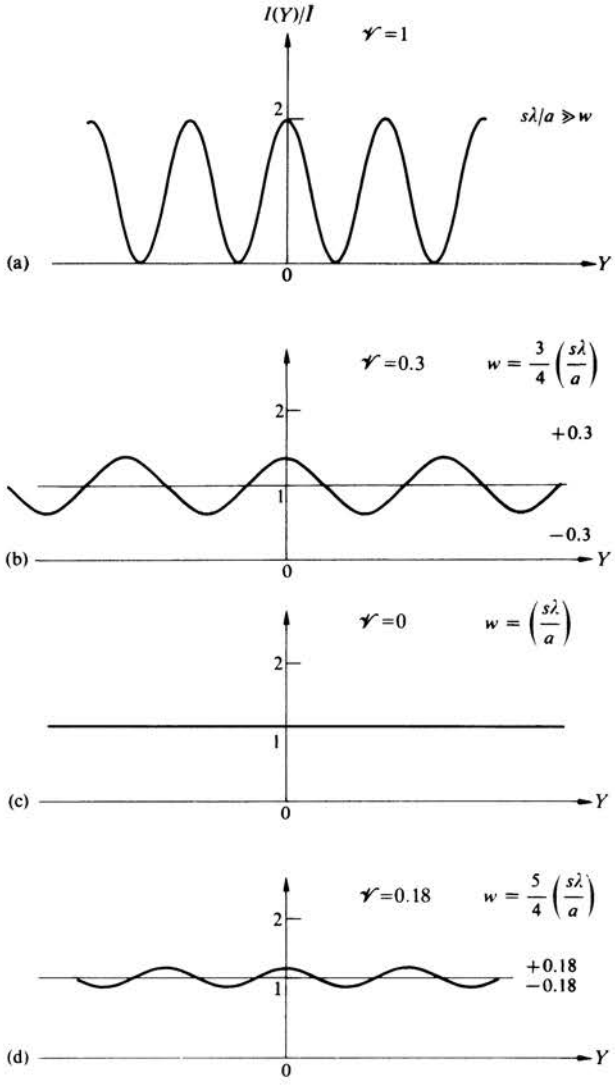


Figure 12.4 Fringes with varying source slit size. Here w is the width of the image of the slit, and $s\lambda/a$ is the peak-to-peak width of the fringes.average to zero.

ations were expressed in terms of w and s, it follows from Fig. 12.3*c*, using the central angle η, that $b \approx l\eta$ and $w \approx s\eta$; hence $w/s \approx b/l$. Accordingly, $(a\pi w/s\lambda) \approx (a\pi\eta/\lambda) \approx (a\pi b/l\lambda)$. The visibility of the fringes follows from Eq. (12.1):

$$\mathcal{V} = \left|\operatorname{sinc}\left(\frac{a\pi w}{s\lambda}\right)\right| = \left|\operatorname{sinc}\left(\frac{a\pi b}{l\lambda}\right)\right| \qquad (12.10)$$

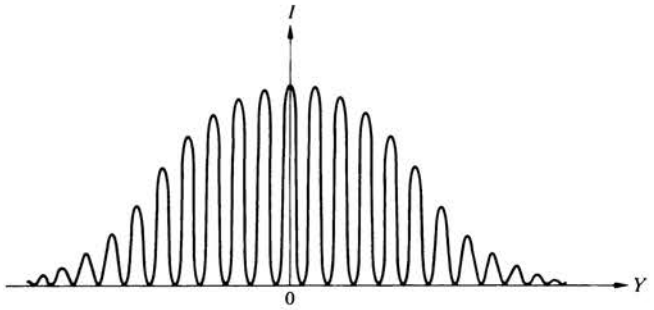


Figure 12.5 Double-beam interference fringes showing the effect of diffraction.

which is plotted in Fig. 12.6. Observe that $\mathcal{V}$ is a function of both the source breadth and the aperture separation a. Holding either one of these parameters constant and varying the other will cause $\mathcal{V}$ to change in precisely the same way. Note that the visibilities in both Figs. 12.4*a* and 12.5 are equal to one because $I_{\min} = 0$. Clearly then, the visibility of the fringe system on the plane of observation is linked to the way the light is distributed over the aperture screen. If the primary source were in fact a point, b would equal zero, and the visibility would be a perfect 1. Shy of that, the smaller $(a\pi b/l\lambda)$ is, the better; that is, the bigger $\mathcal{V}$ is and the clearer the fringes are. We can think of $\mathcal{V}$ as a measure of the degree of coherence of the light from the primary source as spread over the aperture screen. Keep in mind that we have encountered the sinc function before, in connection with the diffraction pattern resulting from a rectangular aperture.

When the primary source is circular, the visibility is a good deal more complicated to calculate. It turns out to be propor-

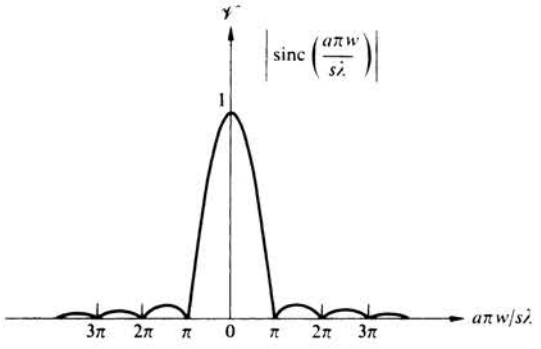


Figure 12.6 The visibility as given by Eq. (12.10).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

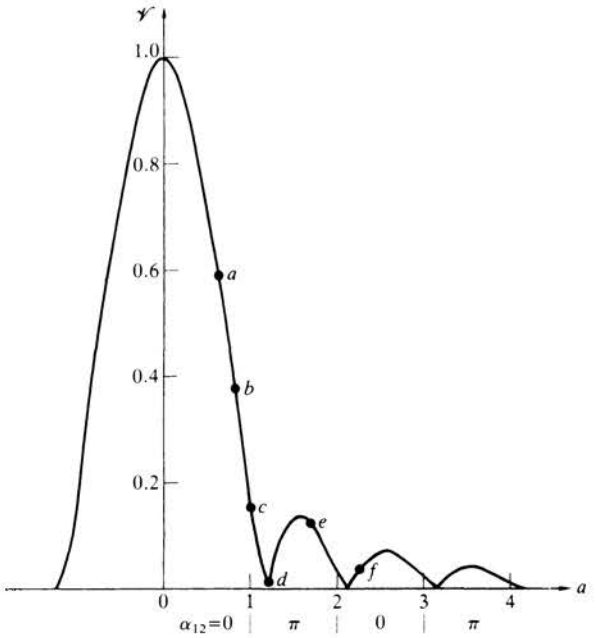


Figure 12.7 The visibility for a circular source.

tional to a first-order Bessel function (Fig. 12.7). This too is quite reminiscent of diffraction, this time at a circular aperture [Eq. (10.56)]. These similarities between expressions for $\mathcal{V}$ and the corresponding diffraction patterns for an aperture of the same shape are not merely fortuitous but rather are a manifestation of something called the van Cittert–Zernike theorem, as we will see presently.

Figure 12.8 shows a sequence of fringe systems in which the circular incoherent primary source is increased in size while the separation a between S_1 and S_2 is held constant.

Figure 12.9 on page 568 results when the separation a is increased while the primary incoherent source diameter is held constant. Since the width of the fringes is inversely proportional to a, the spatial frequency of the bright and dark bands increases as a increases from its value in (a) to that in (f) in Fig. 12.9. The visibility decreases from (a) to (d) in the figure, then increases for (e) and decreases again at (f). All the associated $\mathcal{V}$-values are plotted in Fig. 12.7. Note the shift in the peaks, that is, the change in phase at the center of the pattern for each point on the second lobe of Fig. 12.7 (the Bessel function is negative over that range). In other words, (a), (b), and (c) have a central maximum, while (d) and (e) have a central minimum, and (f) on the third lobe is back to a maximum. In the same way, for a slit source the domain where sinc $(a\pi w/s\lambda)$ in Eq. (12.7) is positive or negative will yield a maximum or minimum, respectively, in $I(0)/\bar{I}$. These in turn correspond to the odd or even lobes of the visibility curve of Fig. 12.6. Bear in mind that we could define a complex visibility of magnitude $\mathcal{V}$, having an argument corresponding to the phase shift—we'll come back to this idea later.

We should also mention that the effects of the finite bandwidth will show up in a given fringe pattern as a gradually decreasing value of $\mathcal{V}$ with Y, as in Fig. 12.10 (see Problem 12.3). When the visibility is determined in these cases, using the central region of each of a series of patterns, the dependence of $\mathcal{V}$ on aperture separation will again match Fig. 12.7.

12.3 The Mutual Coherence Function and the Degree of Coherence

Let's now carry the discussion a bit further in a more formal fashion. Again suppose we have a broad, narrow bandwidth source, which generates a light field whose complex representation is $\tilde{E}(\mathbf{r}, t)$. We'll overlook polarization effects, and therefore a scalar treatment will do. The disturbances at two points in space S_1 and S_2 are then $\tilde{E}(S_1, t)$ and $\tilde{E}(S_2, t)$ or, more succinctly, $\tilde{E}_1(t)$ and $\tilde{E}_2(t)$. If these two points are then isolated using an opaque screen with two circular apertures (Fig. 12.11, p. 570), we're back to Young's Experiment. The two apertures serve as sources of secondary wavelets, which propagate out to some point P on Σ_o. There the resultant field is

$$\tilde{E}_P(t) = \tilde{K}_1\tilde{E}_1(t - t_1) + \tilde{K}_2\tilde{E}_2(t - t_2) \qquad (12.11)$$

where $t_1 = r_1/c$ and $t_2 = r_2/c$. This says that the field at the space–time point (P, t) can be determined from the fields that existed at S_1 and S_2 at t_1 and t_2, respectively, these being the instants when the light, which is now overlapping, first emerged from the apertures. The quantities $\tilde{K}_1$ and $\tilde{K}_2$, which are known as *propagators*, depend on the size of the apertures and their relative locations with respect to P. They mathematically affect the alterations in the field resulting from its having traversed either of the apertures. For example, the secondary wavelets issuing from the pinholes in this setup are out-of-phase by $\pi/2$ rad, with the primary wave incident on

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

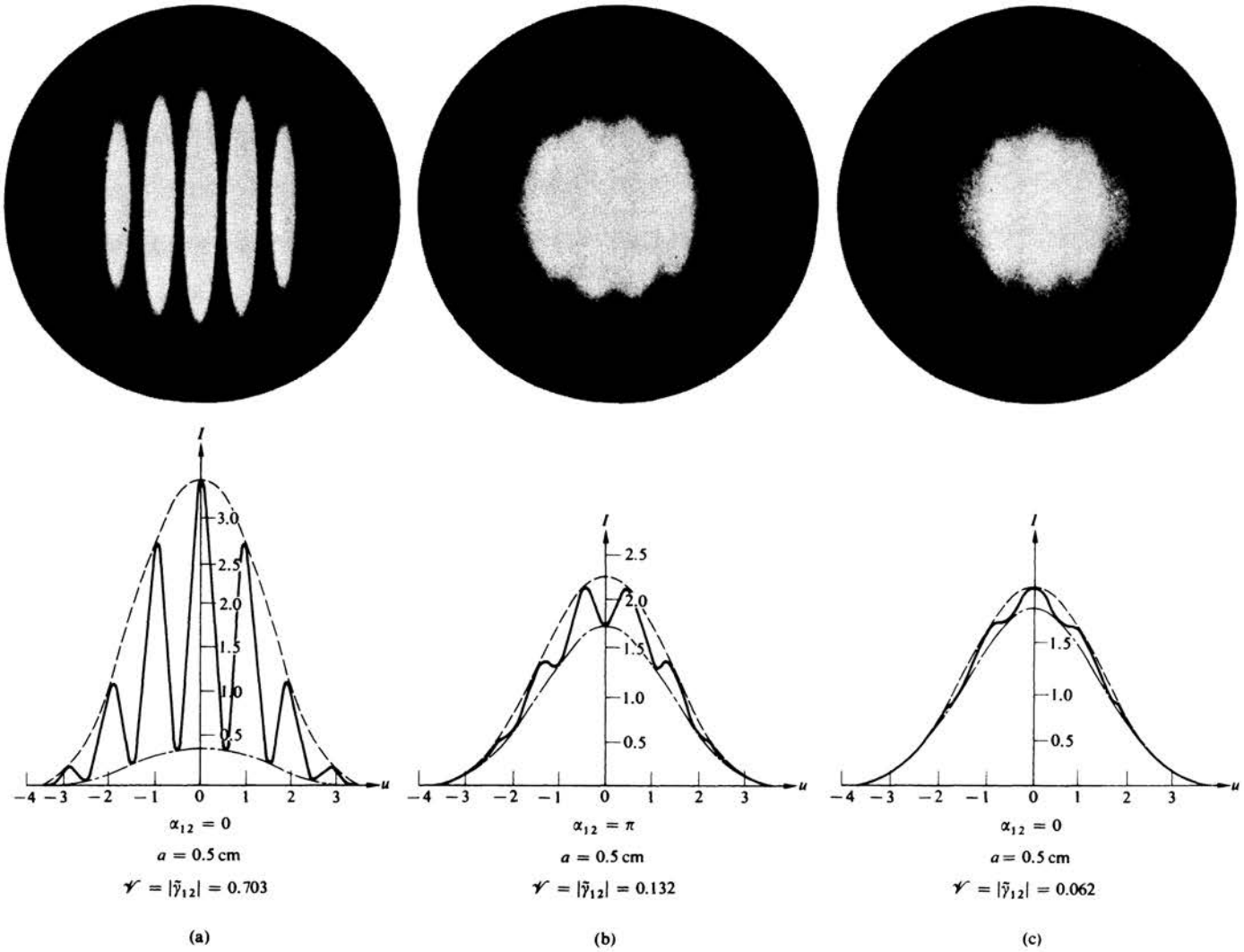


Figure 12.8 Double-beam interference patterns. Here the aperture separation was held constant, thereby yielding a constant number of fringes per unit displacement in each photo. The visibility was altered by varying the size of the primary incoherent source. [From B. J. Thompson, *J. Soc. Photo. Inst. Engr.* **4**, 7 (1965).]

the aperture screen, Σ_a (Section 10.3.1). Clearly, someone is going to have to tell $\tilde{E}(\vec{r}, t)$ to shift phase beyond Σ_a—that's just what the $\tilde{K}$ factors are for. Moreover, they reflect a reduction in the field that might arise from a number of physical causes: absorption, diffraction, and so forth. Here, since there is a $\pi/2$ phase shift in the field, which can be introduced by multiplying by $\exp i\pi/2$, $\tilde{K}_1$ and $\tilde{K}_2$ are purely imaginary numbers.

The resultant irradiance at P measured over some finite time interval, which is long compared with the coherence time, is $\vec{r}$

$$I = \langle \tilde{E}_P(t)\, \tilde{E}_P^*(t) \rangle_T \tag{12.12}$$

It should be remembered that Eq. (12.12) is written sans several multiplicative constants. Hence using Eq. (12.11),

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

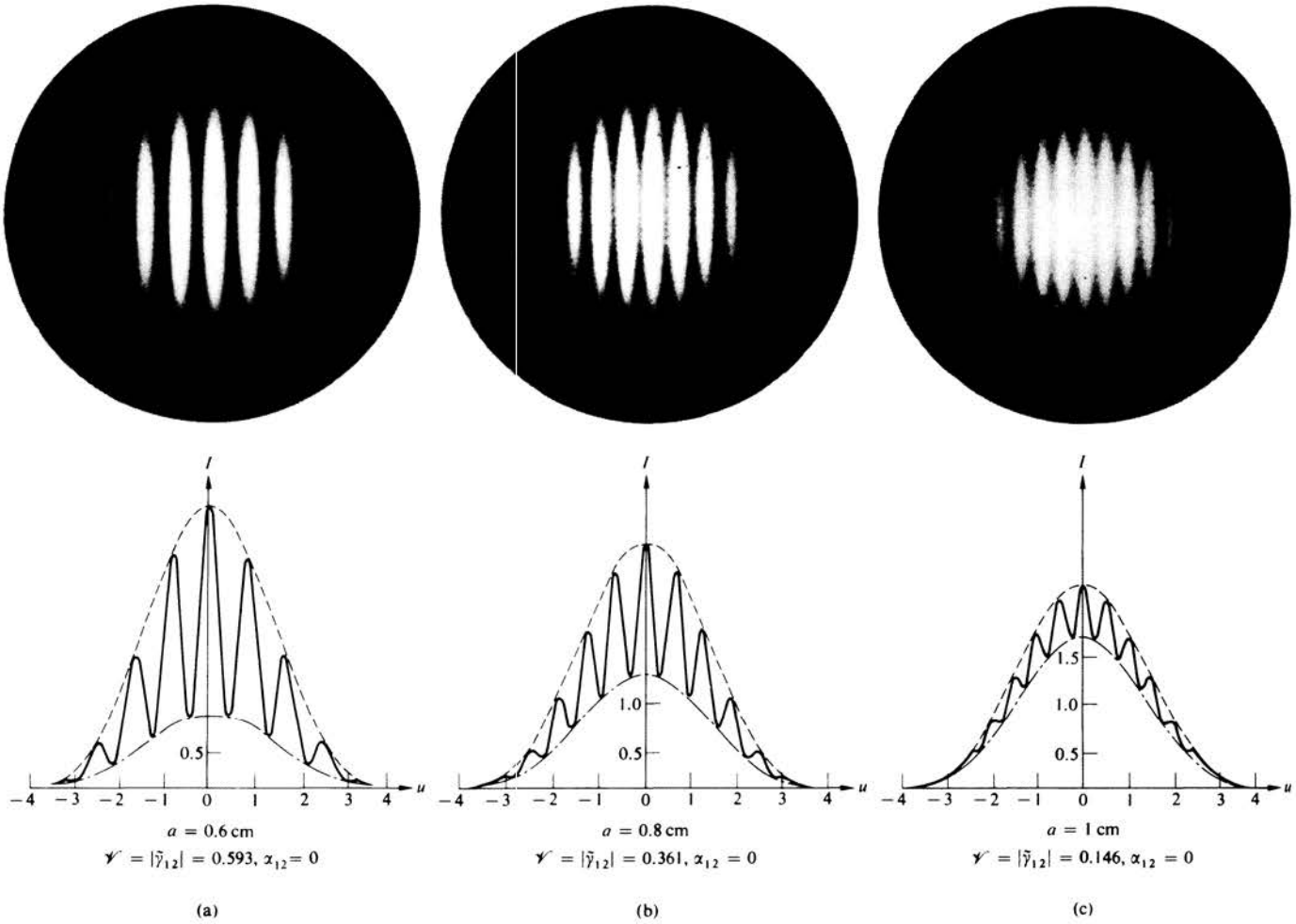


Figure 12.9 Double-beam interference patterns using partially coherent light. The photographs correspond to a variation in visibility associated with changes in a, the separation between the apertures. In the theoretical curves $I_{max} \propto 1 + |2J_1(u)/u|$ and $I_{min} \propto 1 - |2J_1(u)/u|$. Several of the symbols will be discussed later. [From B. J. Thompson and E. Wolf, *J. Opt. Soc. Am.* **47**, 895 (1957).]

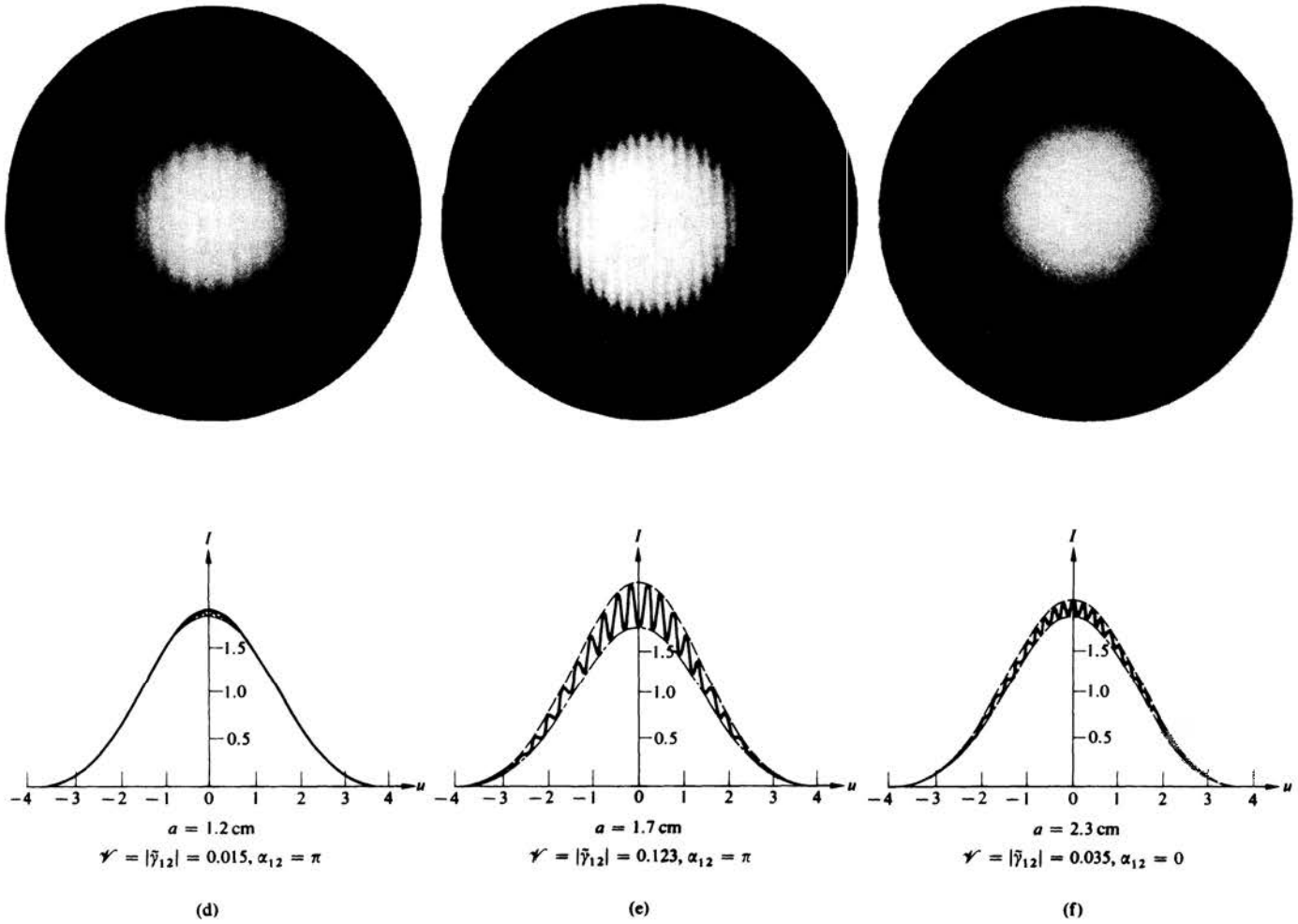
$$\begin{aligned} I = {} & \tilde{K}_1\tilde{K}_1^*\langle \tilde{E}_1(t-t_1)\tilde{E}_1^*(t-t_1)\rangle_T \\ & + \tilde{K}_2\tilde{K}_2^*\langle \tilde{E}_2(t-t_2)\tilde{E}_2^*(t-t_2)\rangle_T \\ & + \tilde{K}_1\tilde{K}_2^*\langle \tilde{E}_1(t-t_1)\tilde{E}_2^*(t-t_2)\rangle_T \\ & + \tilde{K}_1^*\tilde{K}_2\langle \tilde{E}_1^*(t-t_1)\tilde{E}_2(t-t_2)\rangle_T \end{aligned} \qquad (12.13)$$

It is now assumed that the wave field is *stationary*, as is almost universally the case in classical Optics; in other words, it does not alter its statistical nature with time, so that the time average is independent of whatever origin we select. Thus, even though there are fluctuations in the field variables, the time origin can be shifted, and the averages in Eq. (12.13) will be unaffected. The particular moment over which we decide to measure I shouldn't matter. Accordingly, the first two time averages can be rewritten as

$$I_{S_1} = \langle \tilde{E}_1(t)\tilde{E}_1^*(t)\rangle_T \quad \text{and} \quad I_{S_2} = \langle \tilde{E}_2(t)\tilde{E}_2^*(t)\rangle_T$$

where the origin was displaced by amounts t_1 and t_2, respec-

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



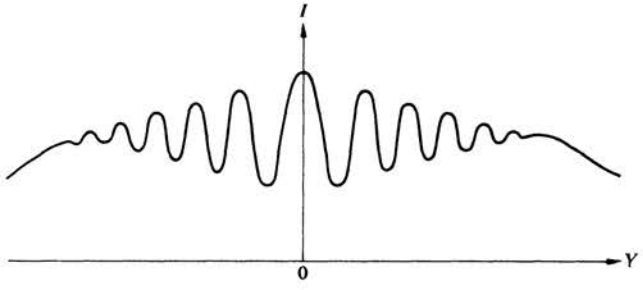


Figure 12.10 A finite bandwidth results in a decreasing value of $\mathscr{V}$ with increasing Y.

tively. Here the subscripts underscore the fact that these are the irradiances at points S_1 and S_2. Furthermore, if we let $\tau = t_2 - t_1$, we can shift the time origin by an amount t_2 in the last two terms of Eq. (12.13) and write them as

$$\tilde{K}_1\tilde{K}_2^*\langle\tilde{E}_1(t+\tau)\tilde{E}_2^*(t)\rangle_T + \tilde{K}_1^*\tilde{K}_2\langle\tilde{E}_1^*(t+\tau)\tilde{E}_2(t)\rangle_T$$

But this is a quantity plus its own complex conjugate and is therefore just twice its real part; that is, it equals

$$2\,\mathrm{Re}\left[\tilde{K}_1\tilde{K}_2^*\langle\tilde{E}_1(t+\tau)\tilde{E}_2^*(t)\rangle_T\right]$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

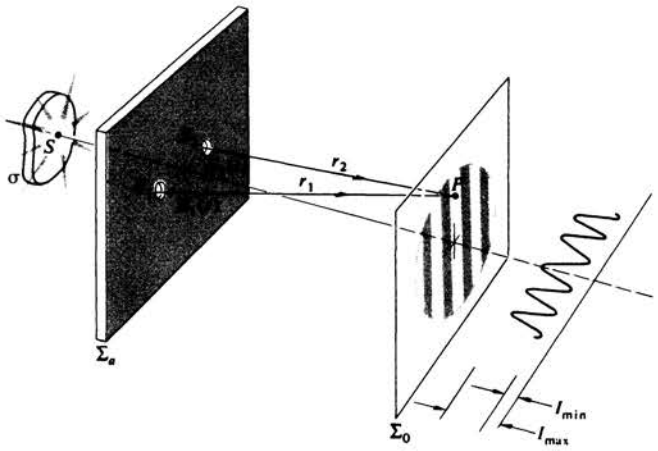


Figure 12.11 Young's Experiment.

The $\tilde{K}$-factors are purely imaginary, and so $\tilde{K}_1\tilde{K}_2^* = \tilde{K}_1^*\tilde{K}_2 = |\tilde{K}_1||\tilde{K}_2|$. The time-average portion of this term is a cross-correlation function [Section 11.3.4(iii)], which we denote by

$$\tilde{\Gamma}_{12}(\tau) \equiv \langle \tilde{E}_1(t+\tau)\tilde{E}_2^*(t)\rangle_T \tag{12.14}$$

and refer to as the **mutual coherence function** of the light field at S_1 and S_2. If we make use of all this, Eq. (12.13) takes the form

$$I = |\tilde{K}_1|^2 I_{S_1} + |\tilde{K}_2|^2 I_{S_2} + 2|\tilde{K}_1||\tilde{K}_2|\,\mathrm{Re}\,\tilde{\Gamma}_{12}(\tau) \tag{12.15}$$

The terms $|\tilde{K}_1|^2 I_{S_1}$ and $|\tilde{K}_2|^2 I_{S_2}$, if we again overlook multiplicative constants, are the irradiance at P arising when one or the other of the apertures is open alone, in other words, $\tilde{K}_2 = 0$ or $\tilde{K}_1 = 0$, respectively. Denoting these as I_1 and I_2, Eq. (12.15) becomes

$$I = I_1 + I_2 + 2|\tilde{K}_1||\tilde{K}_2|\,\mathrm{Re}\,\tilde{\Gamma}_{12}(\tau) \tag{12.16}$$

Note that when S_1 and S_2 are made to coincide, the mutual coherence function becomes

$$\tilde{\Gamma}_{11}(\tau) = \langle \tilde{E}_1(t+\tau)\tilde{E}_1^*(t)\rangle_T$$

or

$$\tilde{\Gamma}_{22}(\tau) = \langle \tilde{E}_2(t+\tau)\tilde{E}_2^*(t)\rangle_T$$

We can imagine that two wavetrains emerge from this coalesced source point and somehow pick up a relative phase delay proportional to τ. In the present situation τ becomes zero (since the optical path difference goes to zero), and these functions are reduced to the corresponding irradiances $I_{S_1} = \langle \tilde{E}_1(t)\tilde{E}_1^*(t)\rangle_T$ and $I_{S_2} = \langle \tilde{E}_2(t)\tilde{E}_2^*(t)\rangle_T$ on Σ_a. Hence

$$\Gamma_{11}(0) = I_{S_1} \quad \text{and} \quad \Gamma_{22}(0) = I_{S_2}$$

and these are called *self-coherence functions*. Thus

$$I_1 = |\tilde{K}_1|^2\Gamma_{11}(0) \quad \text{and} \quad I_2 = |\tilde{K}_2|^2\Gamma_{22}(0)$$

Keeping Eq. (12.16) in mind, observe that

$$|\tilde{K}_1||\tilde{K}_2| = \sqrt{I_1}\sqrt{I_2}\,/\,\sqrt{\Gamma_{11}(0)}\sqrt{\Gamma_{22}(0)}$$

Hence the normalized form of the mutual coherence function is defined as

$$\tilde{\gamma}_{12}(\tau) \equiv \frac{\tilde{\Gamma}_{12}(\tau)}{\sqrt{\Gamma_{11}(0)\Gamma_{22}(0)}} = \frac{\langle \tilde{E}_1(t+\tau)\ \tilde{E}_2^*(t)\rangle_T}{\sqrt{\langle|\tilde{E}_1|^2\rangle\langle|\tilde{E}_2|^2\rangle}} \tag{12.17}$$

and it's spoken of as the **complex degree of coherence**, for reasons that will be clear imminently. Equation (12.16) can then be recast as

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\,\mathrm{Re}\,\tilde{\gamma}_{12}(\tau) \tag{12.18}$$

which is the *general interference law for partially coherent light.*

For quasimonochromatic light, the phase-angle difference concomitant with the optical path difference is given by

$$\varphi = \frac{2\pi}{\bar{\lambda}}(r_2 - r_1) = 2\pi\bar{\nu}\tau \tag{12.19}$$

where $\bar{\lambda}$ and $\bar{\nu}$ are the mean wavelength and frequency. Now $\tilde{\gamma}_{12}(\tau)$ is a complex quantity expressible as

$$\tilde{\gamma}_{12}(\tau) = |\tilde{\gamma}_{12}(\tau)|e^{i\Phi_{12}(\tau)} \tag{12.20}$$

The phase angle of $\tilde{\gamma}_{12}(\tau)$ relates back to Eq. (12.14) and the phase angle between the fields. If we set $\Phi_{12}(\tau) = \alpha_{12}(\tau) - \varphi$, then

$$\mathrm{Re}\ \tilde{\gamma}_{12}(\tau) = |\tilde{\gamma}_{12}(\tau)|\cos[\alpha_{12}(\tau) - \varphi]$$

Equation (12.18) is then expressible as

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}|\tilde{\gamma}_{12}(\tau)|\cos[\alpha_{12}(\tau) - \varphi] \tag{12.21}$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

It can be shown from Eq. (12.17) and the Schwarz inequality that $0 \leq |\tilde{\gamma}_{12}(\tau)| \leq 1$. In fact, a comparison of Eqs. (12.21) and (9.14), the latter having been derived for the case of complete coherence, makes it evident that if $|\tilde{\gamma}_{12}(\tau)| = 1$, I is the same as that generated by two *coherent* waves out-of-phase at S_1 and S_2 by an amount $\alpha_{12}(\tau)$. If at the other extreme $|\tilde{\gamma}_{12}(\tau)| = 0$, $I = I_1 + I_2$, there is no interference, and the two disturbances are said to be *incoherent*. When $0 < |\tilde{\gamma}_{12}(\tau)| < 1$ we have *partial coherence*, the measure of which is $|\tilde{\gamma}_{12}(\tau)|$ itself; this is known as the **degree of coherence**. In summary then,

$$|\tilde{\gamma}_{12}| = 1 \quad \text{coherent limit}$$

$$|\tilde{\gamma}_{12}| = 0 \quad \text{incoherent limit}$$

$$0 < |\tilde{\gamma}_{12}| < 1 \quad \text{partial coherence}$$

The basic statistical nature of the entire process must be underscored. Clearly $\tilde{\Gamma}_{12}(\tau)$ and, therefore, $\tilde{\gamma}_{12}(\tau)$ are the key quantities in the various expressions for the irradiance distribution; they are the essence of what we previously called the interference term. It should be pointed out that $\tilde{E}_1(t + \tau)$ and $\tilde{E}_2(t)$ are in fact two disturbances occurring at different points in both space and time. We anticipate, as well, that the amplitudes and phases of these disturbances will somehow fluctuate in time. If these fluctuations at S_1 and S_2 are completely independent, then $\tilde{\Gamma}_{12}(\tau) = \langle \tilde{E}_1(t + \tau)\, \tilde{E}_2^*(t) \rangle_T$ will go to zero, since $\tilde{E}_1$ and $\tilde{E}_2$ can be either positive or negative with equal likelihood, and their product averages to zero. In that case no correlation exists, and $\tilde{\Gamma}_{12}(\tau) = \tilde{\gamma}_{12}(\tau) = 0$. If the field at S_1 at a time $(t + \tau)$ were perfectly correlated with the field at S_2 at a time t, their relative phase would remain unaltered despite individual fluctuations. The time average of the product of the fields would certainly not be zero, just as it would not be zero even if the two were only slightly correlated.

Both $|\tilde{\gamma}_{12}(\tau)|$ and $\alpha_{12}(\tau)$ are slowly varying functions of τ in comparison to $\cos 2\pi\bar{\nu}\tau$ and $\sin 2\pi\bar{\nu}\tau$. In other words, as P is moved across the resultant fringe system, the point-by-point spatial variations in I are predominantly due to the changes in φ as $(r_2 - r_1)$ changes.

The maximum and minimum values of I occur when the cosine term in Eq. (12.21) is $+1$ and -1, respectively. The visibility at P (Problem 12.7) is then

$$\mathcal{V} = \frac{2\sqrt{I_1}\sqrt{I_2}}{I_1 + I_2} |\tilde{\gamma}_{12}(\tau)| \tag{12.22}$$

Perhaps the most common arrangement occurs when things are adjusted so that $I_1 = I_2$, whereupon

$$\mathcal{V} = |\tilde{\gamma}_{12}(\tau)| \tag{12.23}$$

That is, *the modulus of the complex degree of coherence is identical to the visibility of the fringes* (take another look at Fig. 12.8).

It is essential to realize that Eqs. (12.17) and (12.18) clearly suggest the way in which the real parts of $\tilde{\Gamma}_{12}(\tau)$ and $\tilde{\gamma}_{12}(\tau)$ can be determined from direct measurement. When the flux densities of two disturbances are adjusted to be equal, Eq. (12.23) provides an experimental means of obtaining $|\tilde{\gamma}_{12}(\tau)|$ from the resultant fringe pattern. Furthermore, the off-axis shift in the location of the central fringe (from $\varphi = 0$) is a measure of $\alpha_{12}(\tau)$, the apparent relative retardation of the phase of the disturbances at S_1 and S_2. Thus, measurements of the visibility and fringe position yield both the amplitude and phase of the complex degree of coherence.

By the way, it can be shown* that $|\tilde{\gamma}_{12}(\tau)|$ will equal 1 for all values of τ and any pair of spatial points, if and only if the optical field is strictly monochromatic, and therefore such a situation is unattainable. Moreover, a nonzero radiation field for which $|\tilde{\gamma}_{12}(\tau)| = 0$ for all values of τ and any pair of spatial points cannot exist in free space either.

12.3.1 Temporal and Spatial Coherence

Let's now relate the ideas of temporal and spatial coherence to the above formalism.

If the primary source S in Fig. 12.11 shrinks down to a point source on the central axis having a finite frequency bandwidth, temporal coherence effects will predominate. The optical disturbances at S_1 and S_2 will then be identical. In effect, the mutual coherence [Eq. (12.14)] between the two points will be the self-coherence of the field. Hence $\tilde{\Gamma}(S_1, S_2, \tau) = \tilde{\Gamma}_{12}(\tau) = \tilde{\Gamma}_{11}(\tau)$ or $\tilde{\gamma}_{12}(\tau) = \tilde{\gamma}_{11}(\tau)$. The same thing obtains when S_1 and S_2 coalesce, and $\tilde{\gamma}_{11}(\tau)$ is sometimes referred to as the **complex degree of temporal coherence** at that point for two instances of time separated by an interval τ. This would be the case in an amplitude-splitting interferometer, such as Michelson's, in which τ equals the

*The proofs are given in Beran and Parrent, *Theory of Partial Coherence*, Section 4.2.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

path length difference divided by c. The expression for I, that is, Eq. (12.18), would then contain $\tilde{\gamma}_{11}(\tau)$ rather than $\tilde{\gamma}_{12}(\tau)$.

Suppose a lightwave is divided into two identical disturbances of the form

$$\tilde{E}(t) = E_0 e^{i\phi(t)} \tag{12.24}$$

by an amplitude-splitting interferometer, which later recombines them to generate a fringe pattern. Then

$$\tilde{\gamma}_{11}(\tau) = \frac{\langle \tilde{E}(t+\tau)\tilde{E}^*(t)\rangle_T}{|\tilde{E}|^2} \tag{12.25}$$

or $$\tilde{\gamma}_{11}(\tau) = \langle e^{i\phi(t+\tau)} e^{-i\phi(t)}\rangle_T$$

Hence

$$\tilde{\gamma}_{11}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T e^{i[\phi(t+\tau)-\phi(t)]}\,dt \tag{12.26}$$

and $$\tilde{\gamma}_{11}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T (\cos\Delta\phi + i\sin\Delta\phi)\,dt$$

where $\Delta\phi = \phi(t+\tau) - \phi(t)$. For a strictly monochromatic plane wave of infinite coherence length, $\phi(t) = \vec{\mathbf{k}}\cdot\vec{\mathbf{r}} - \omega t$, $\Delta\phi = -\omega\tau$, and

$$\tilde{\gamma}_{11}(\tau) = \cos\omega\tau - i\sin\omega\tau = e^{-i\omega\tau}$$

Hence $|\tilde{\gamma}_{11}| = 1$; the argument of $\tilde{\gamma}_{11}$ is just $-2\pi\nu\tau$, and we have complete coherence. In contradistinction, for a quasi-monochromatic wave where τ is greater than the coherence time, $\Delta\phi$ will be random, varying between 0 and 2π such that the integral averages to zero, $|\tilde{\gamma}_{11}(\tau)| = 0$, corresponding to complete incoherence. A path difference of 60 cm, produced when the two arms of a Michelson Interferometer differ in length by 30 cm, corresponds to a time delay between the recombining beams of $\tau \approx 2$ ns. This is roughly the coherence time of a good isotope discharge lamp, and the visibility of the pattern under this sort of illumination will be quite poor. If white light is used instead, $\Delta\nu$ is large, Δt_c is very small, and the coherence length is less than one wavelength. In order for τ to be less than Δt_c (i.e., in order that the visibility be good), the optical path difference will have to be a small fraction of a wavelength. The other extreme is laserlight, in which Δt_c can be so long that a value of $c\tau$ that will cause an appreciable decrease in visibility would require an impractically large interferometer.

We see that $\tilde{\Gamma}_{11}(\tau)$, being a measure of temporal coherence, must be intimately related to the coherence time and therefore the bandwidth of the source. Indeed, ***the Fourier transform of the self-coherence function, $\tilde{\Gamma}_{11}(\tau)$, is the power spectrum, which describes the spectral energy distribution of the light*** (Section 11.3.4).

If we go back to Young's Experiment (Fig. 12.11) with a very narrow-bandwidth extended source, spatial coherence effects will predominate. The optical disturbances at S_1 and S_2 will differ, and the fringe pattern will depend on $\tilde{\Gamma}(S_1, S_2, \tau) = \tilde{\Gamma}_{12}(\tau)$. By examining the region about the central fringe where $(r_2 - r_1) = 0$, $\tau = 0$ and $\tilde{\Gamma}_{12}(0)$ and $\tilde{\gamma}_{12}(0)$ can be determined. This latter quantity is the **complex degree of spatial coherence** of the two points at the same instant in time. $\tilde{\Gamma}_{12}(0)$ plays a central role in the description of the Michelson stellar interferometer to be discussed forthwith.

There is a very convenient relationship between the complex degree of coherence in a region of space and the corresponding irradiance distribution across the extended source giving rise to the light fields. We shall make use of that relationship, the **van Cittert–Zernike theorem**, as a calculational aid without going through its formal derivation. Indeed, the analysis of Section 12.2 already suggests some of the essentials. Figure 12.12 represents an extended quasimonochromatic incoherent source, S, located on the plane σ and having an irradiance given by $I(y, z)$. Also shown is an observation screen on which are two points, P_1 and P_2. These are at distances R_1 and R_2, respectively, from a tiny element of S. It is on this plane that we wish to determine $\tilde{\gamma}_{12}(0)$, which describes the correlation of the field vibrations at the two points. Note that although the source is incoherent, the light reaching P_1 and P_2 will generally be correlated to some degree, since each source element contributes to the field at each such point.

Calculation of $\tilde{\gamma}_{12}(0)$ from the fields at P_1 and P_2 results in an integral that has a familiar structure. The integral has the same form and will yield the same results as a well-known diffraction integral, provided we reinterpret each term appropriately. For instance, $I(y, z)$ appears in that coherence integral where an aperture function would be if it were, in fact, a diffraction integral. Thus, suppose that S is not a source but an aperture of identical size and shape, and suppose that $I(y, z)$ is not a description of irradiance, but instead its functional form corresponds to the field distribution across that aperture. In other words, imagine that there is a transparency at the aperture with amplitude transmission characteristics that correspond functionally to $I(y, z)$. Furthermore, imagine that the

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

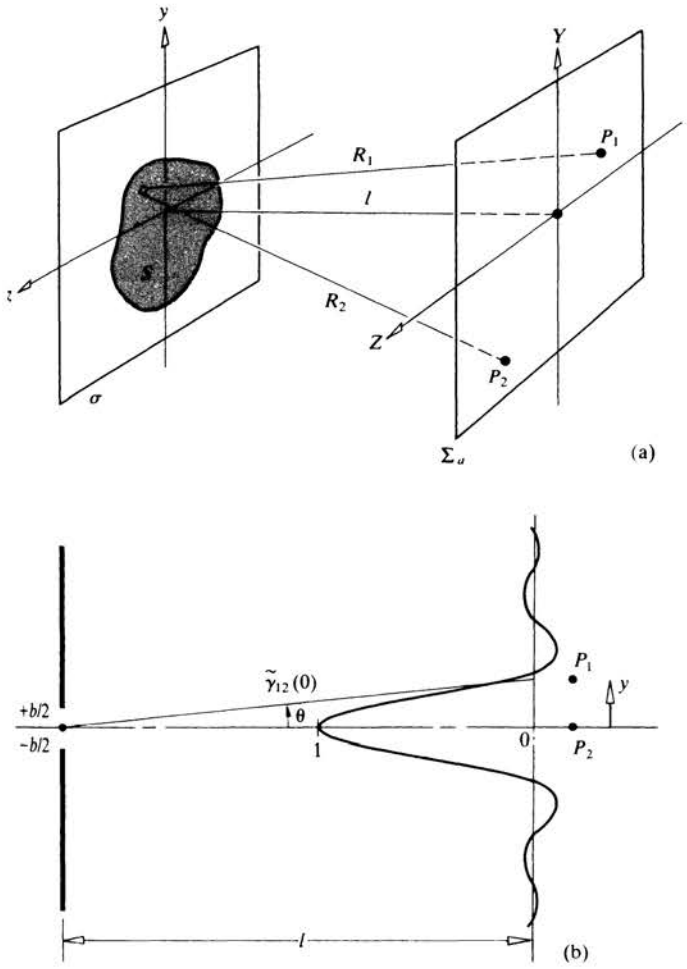


Figure 12.12 (*a*) The geometry of the van Cittert–Zernike theorem. (*b*) The normalized diffraction pattern corresponds to the degree of coherence. Here for a rectangular source slit the diffraction pattern is sinc ($\pi by/l\lambda$)

aperture is illuminated by a spherical wave converging toward the fixed point P_2 (see Fig. 12.12*b*), so that *there will be a diffraction pattern centered on* P_2. This diffracted field distribution, normalized to unity at P_2, is everywhere (i.e., at P_1) equal to the value of $\tilde{\gamma}_{12}(0)$ at that point. This is the van Cittert–Zernike theorem.

When P_1 and P_2 are close together and S is small compared with l, the complex degree of coherence equals the normalized Fourier transform of the irradiance distribution across the source. Furthermore, if the source has a uniform irradiance, then $\tilde{\gamma}_{12}(0)$ is simply a sinc function when the source is a slit and a Bessel function when it's circular. Observe that in Fig. 12.12*b* the sinc function corresponds to that used in Fig. 10.10, where $\beta = (kb/2)\sin\theta$ and $\theta \approx \sin\theta$. Thus if P_1 is a distance y from P_2, $\beta = kb\theta/2$ and $\theta = y/l$, hence $|\tilde{\gamma}_{12}(0)| = |\text{sinc}\,(\pi by/l\lambda)|$. This result is explored further in the problem set.

12.4 Coherence and Stellar Interferometry

12.4.1 The Michelson Stellar Interferometer

In 1890 A. A. Michelson, following an earlier suggestion by Fizeau, proposed an interferometric device (Fig. 12.13) that is of interest here both because it was the precursor of some important modern techniques, and because it lends itself to an interpretation in terms of coherence theory. The function of the *stellar interferometer*, as it is called, is to measure the small angular dimensions of remote astronomical bodies.

Two widely spaced movable mirrors, M_1 and M_2, collect rays, assumed to be parallel, from a very distant star. The light is then channeled via mirrors M_3 and M_4 through apertures S_1 and S_2 of a mask and thence into the objective of a telescope. The optical paths $M_1M_3S_1$ and $M_2M_4S_2$ are made equal, so that the relative phase-angle difference between a disturbance at M_1 and M_2 is the same as that between S_1 and S_2. The two apertures generate the usual Young's Experiment fringe system in the focal plane of the objective. Actually, the mask and openings are not really necessary; the mirrors alone could serve as apertures.

Suppose we now point the device so that its central axis is directed toward one of the stars in a closely spaced double-star configuration. Because of the tremendous distances involved, the rays reaching the interferometer from either star are well collimated. Furthermore, we assume, at least for the moment, that the light has a narrow linewidth centered about a mean wavelength of $\bar{\lambda}_0$. The disturbances arising at S_1 and S_2 from the axial star are in-phase, and a pattern of bright and dark bands forms, centered on P_0.

Similarly, rays from the other star arrive at some angle θ, but this time the disturbances at M_1 and M_2 (and therefore at S_1 and S_2) are out-of-phase by approximately $\bar{k}_0 h\theta$ or, if you will,

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

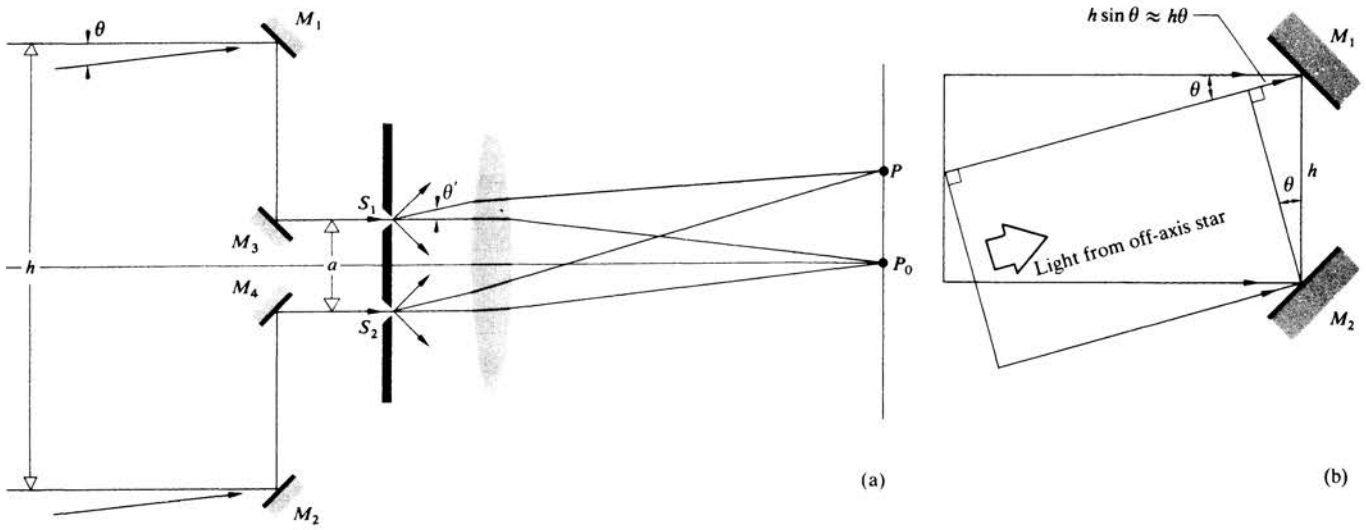


Figure 12.13 Michelson stellar interferometer.

retarded by a time $h\theta/c$, as indicated in Fig. 12.13*b*. The resulting fringe system is centered about a point P shifted by an angle θ' from P_0 such that $h\theta/c = a\theta'/c$. Since these stars behave as though they were incoherent point sources, the individual irradiance distributions simply overlap. The separation between the fringes set up by either star is equal and dependent solely on a. Yet the visibility varies with h. Thus, if h is increased from nearly zero until $\bar{k}_0 h\theta = \pi$, that is, until

$$h = \frac{\bar{\lambda}_0}{2\theta} \tag{12.27}$$

the two fringe systems take on an increasing relative displacement, until finally the maxima from one star overlap the minima from the other, at which point, if their irradiances are equal, $\mathcal{V} = 0$. Hence, when the fringes vanish, one need only measure h to determine the angular separation between the stars, θ. Notice that the appropriate value of h varies inversely with θ.

Note that even though the source points, the two stars, are assumed to be completely uncorrelated, the resulting optical fields at any two points (M_1 and M_2) are not necessarily incoherent. For that matter, as h becomes very small, the light from each point source arrives with essentially zero relative phase at M_1 and M_2; $\mathcal{V}$ approaches 1, and the fields at those locations are highly coherent.

In much the same way as with a double star system, the angular diameter (θ) of certain single stars can be measured. Once again the fringe visibility corresponds to the degree of coherence of the optical field at M_1 and M_2. If the star is assumed to be a circular distribution of incoherent point sources such that it has a uniform brilliance, its visibility is equivalent to that already plotted in Fig. 12.7. Earlier, we alluded to the fact that $\mathcal{V}$ for this sort of source was given by a first-order Bessel function, and in fact it is expressible as

$$\mathcal{V} = |\tilde{\gamma}_{12}(0)| = 2\left|\frac{J_1(\pi h\theta/\bar{\lambda}_0)}{\pi h\theta/\bar{\lambda}_0}\right| \tag{12.28}$$

Recall that $J_1(u)/u = \frac{1}{2}$ at $u = 0$, and the maximum value of $\mathcal{V}$ is 1. The first zero of $\mathcal{V}$ occurs when $\pi h\theta/\bar{\lambda}_0 = 3.83$, as in Fig. 10.23. Equivalently, the fringes disappear when

$$h = 1.22\frac{\bar{\lambda}_0}{\theta} \tag{12.29}$$

and as before, one simply measures h to find θ.

In Michelson's arrangement, the two outrigged mirrors were movable on a long girder, which was mounted on the 100-inch reflector of the Mount Wilson Observatory. Betelgeuse (α Orionis) was the first star whose angular diameter was measured with the device. It's the orange-looking star in the upper left of the constellation Orion. In fact, its name is a contraction for the Arabic phrase meaning *the armpit of the central one* (i.e., Orion). The fringes formed by the interfer-

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

ometer, one cold December night in 1920, were made to vanish at $h = 121$ inches, and with $\bar{\lambda}_0 = 570$ nm, $\theta = 1.22(570 \times 10^{-9})/121(2.54 \times 10^{-2}) = 22.6 \times 10^{-8}$ rad, or 0.047 second of arc. Using its known distance, determined from parallax measurements, the star's diameter turned out to be about 240 million miles, or roughly 280 times that of the Sun. Actually, Betelgeuse is an irregular variable star whose maximum diameter is so tremendous that it's larger than the orbit of Mars about the Sun. The main limitation on the use of the stellar interferometer is due to the inconveniently long mirror separations required for all but the largest stars. This is true as well in radio astronomy, where an analogous setup has been widely used to measure the extent of celestial sources of radiofrequency emissions.

Incidentally, we assume, as is often done, that "good" coherence means a visibility of 0.88 or better. For a disk source this occurs when $\pi h\theta/\bar{\lambda}_0$ in Eq. (12.28) equals one, that is, when

$$h = 0.32\,\frac{\bar{\lambda}_0}{\theta} \tag{12.30}$$

For a narrow-bandwidth source of diameter D a distance R away, there is an **area of coherence** equal to $\pi(h/2)^2$ over which $|\tilde{\gamma}_{12}| \geq 0.88$. Since $D/R = \theta$,

$$h = 0.32\,\frac{R\bar{\lambda}_0}{D} \tag{12.31}$$

These expressions are very handy for estimating the required physical parameters in an interference or diffraction experiment. For example, if we put a red filter over a 1-mm-diameter disk-shaped flashlight source and stand back 20 m from it, then

$$h = 0.32(20)(600 \times 10^{-9})/10^{-3} = 3.8 \text{ mm}$$

where the mean wavelength is taken as 600 nm. This means that a set of apertures spaced at about h or less should produce nice fringes. Evidently, the area of coherence increases with R, and this is why you can always find a distant bright streetlight to use as a convenient source.

12.4.2 Correlation Interferometry

Let's return for a moment to the representation of a disturbance emanating from a thermal source, as discussed in Section 7.4.3. Here the word *thermal* connotes a light field arising predominantly from the superposition of spontaneously emitted waves issuing from a great many independent atomic sources.* A quasimonochromatic optical field can be represented by

$$E(t) = E_0(t)\cos[\varepsilon(t) - 2\pi\bar{\nu}t] \tag{7.65}$$

The amplitude is a relatively slowly varying function of time, as is the phase. For that matter, the wave might undergo tens of thousands of oscillations before either the amplitude (i.e., the envelope of the field vibrations) or the phase would change appreciably. Thus, just as the coherence time is a measure of the fluctuation interval of the phase, it is also a measure of the interval over which $E_0(t)$ is fairly predictable. Large fluctuations in ε are generally accompanied by correspondingly large fluctuations of E_0. Presumably, knowledge of these amplitude fluctuations of the field could be related to the phase fluctuations and therefore to the correlation (i.e., coherence) functions. Accordingly, at two points in space–time where the phases of the field are correlated, we could expect the amplitudes to be related as well.

When a fringe pattern exists for the Michelson stellar interferometer, it is because the fields at M_1 and M_2, the apertures, are somehow correlated, that is, $\tilde{\Gamma}_{12}(0) = \langle \tilde{E}_1(t)\tilde{E}_2^*(t)\rangle_T \neq 0$. If we could measure the field amplitudes at these points, their fluctuations would likewise show an interrelationship. Since this isn't practicable because of the high frequencies involved, we might instead measure and compare the fluctuations in irradiance at the locations of M_1 and M_2 and from this, in some as yet unknown way, infer $|\tilde{\gamma}_{12}(0)|$. In other words, if there are values of τ for which $\tilde{\gamma}_{12}(\tau)$ is nonzero, the field at the two points is partially coherent, and a correlation between the irradiance fluctuations at these locations is implied. This is the essential idea behind a series of remarkable experiments conducted in the years 1952 to 1956 by R. Hanbury–Brown in collaboration with R. Q. Twiss and others. The culmination of their work was the so-called *correlation interferometer.*

Thus far we have evolved only an intuitive justification for the phenomenon rather than a firm theoretical treatment. Such an analysis, however, is beyond the scope of this discussion, and we shall have to content ourselves with merely outlining its salient features.* Just as in Eq. (12.14), we are interested in

*Thermal light is sometimes spoken of as *Gaussian light* because the amplitude of the field follows a Gaussian probability distribution (p.12).

*For a complete discussion, see, for example, L. Mandel, "Fluctuations of Light Beams," *Progress in Optics*, Vol. II, p. 193, or Françon, *Optical Interferometry*, p. 182.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

determining the cross-correlation function, this time, of the irradiances at two points in a partially coherent field, $\langle I_1(t+\tau)I_2(t)\rangle_T$. The contributing wavetrains, which are again represented by complex fields, are assumed to have been randomly emitted in accord with Gaussian statistics, with the final result that

$$\langle I_1(t+\tau)I_2(t)\rangle_T = \langle I_1\rangle_T\langle I_2\rangle_T + |\tilde{\Gamma}_{12}(\tau)|^2 \qquad (12.32)$$

or

$$\langle I_1(t+\tau)I_2(t)\rangle_T = \langle I_1\rangle_T\langle I_2\rangle_T[1 + |\tilde{\gamma}_{12}(\tau)|^2] \qquad (12.33)$$

The instantaneous irradiance fluctuations $\Delta I_1(t)$ and $\Delta I_2(t)$ are given by the variations of the instantaneous irradiances $I_1(t)$ and $I_2(t)$ about their mean values $\langle I_1(t)\rangle_T$ and $\langle I_2(t)\rangle_T$, as in Fig. 12.14. Consequently, if we use

$$\Delta I_1(t) = I_1(t) - \langle I_1\rangle_T \quad \text{and} \quad \Delta I_2(t) = I_2(t) - \langle I_2\rangle_T$$

and the fact that

$$\langle \Delta I_1(t)\rangle_T = 0 \quad \text{and} \quad \langle \Delta I_2(t)\rangle_T = 0$$

Eqs. (12.32) and (12.33) become

$$\langle \Delta I_1(t+\tau)\Delta I_2(t)\rangle_T = |\tilde{\Gamma}_{12}(\tau)|^2 \qquad (12.34)$$

or

$$\langle \Delta I_1(t+\tau)\Delta I_2(t)\rangle_T = \langle I_1\rangle_T\langle I_2\rangle_T|\tilde{\gamma}_{12}(\tau)|^2 \qquad (12.35)$$

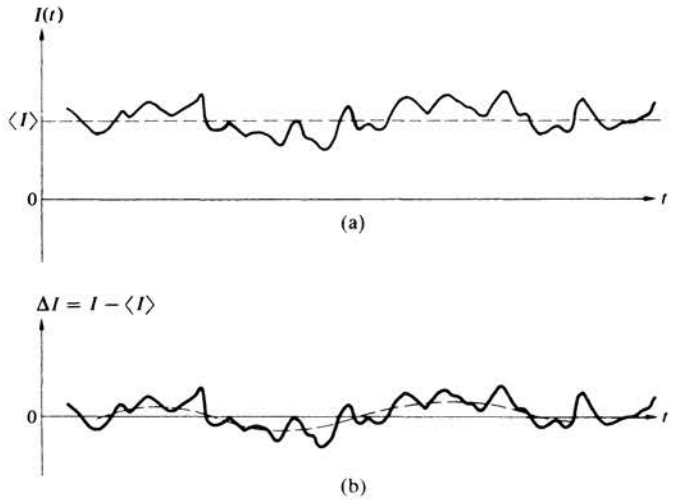


Figure 12.14 Irradiance variations.

(Problem 12.11). These are the desired cross-correlations of the irradiance fluctuations. They exist as long as the field is partially coherent at the two points in question. Incidentally, these expressions correspond to linearly polarized light. When the wave is unpolarized. a multiplicative factor of $\frac{1}{2}$ must be introduced on the right-hand side.

The validity of the principle of correlation interferometry was first established in the radiofrequency region of the spectrum, where signal detection was a fairly straightforward matter. Soon afterward, in 1956, Hanbury–Brown and Twiss proposed the optical stellar interferometer illustrated in Fig. 12.15. But the only suitable detectors that could be used at optical frequencies were photoelectric devices whose very operation is keyed to the quantized nature of the light field. Thus

> ...it was by no means certain that the correlation would be fully preserved in the process of photo-electric emission. For these reasons a laboratory experiment was carried out as described below.*

That experiment is shown in Fig. 12.16. Filtered light from a Hg arc was passed through a rectangular aperture, and different portions of the emerging wavefront were sampled by two photomultipliers, PM_1 and PM_2. The degree of coherence was altered by moving PM_1, that is, by varying h. The signals from the two photomultipliers were presumably proportional to the incident irradiances $I_1(t)$ and $I_2(t)$. These were then filtered and amplified, such that the steady, or dc, component of each of the signals (being proportional to $\langle I_1\rangle_T$ and $\langle I_2\rangle_T$) was removed, leaving only the fluctuations, in other words, $\Delta I_1(t) = I_1(t) - \langle I_1\rangle_T$ and $\Delta I_2(t) = I_2(t) - \langle I_2\rangle_T$. The two signals were then multiplied together in the correlator, and the time average of the product, which was proportional to $\langle \Delta I_1(t)\Delta I_2(t)\rangle_T$, was finally recorded. The values of $|\tilde{\gamma}_{12}(0)|^2$ for various separations, h, as deduced experimentally via Eq. (12.35), were in fine agreement with those calculated from theory. For the given geometry, the correlation definitely existed; moreover, it was preserved through photoelectric detection.

The irradiance fluctuations have a frequency bandwidth roughly equivalent to the bandwidth ($\Delta\nu$) of the incident light, in other words, $(\Delta t_c)^{-1}$, which is about 100 MHz or more. This is much better than trying to follow the field alternations at

*Taken from R. Hanbury–Brown and R. Q. Twiss, "Correlation Between Photons in Two Coherent Beams of Light," *Nature* **127**, 27 (1956).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

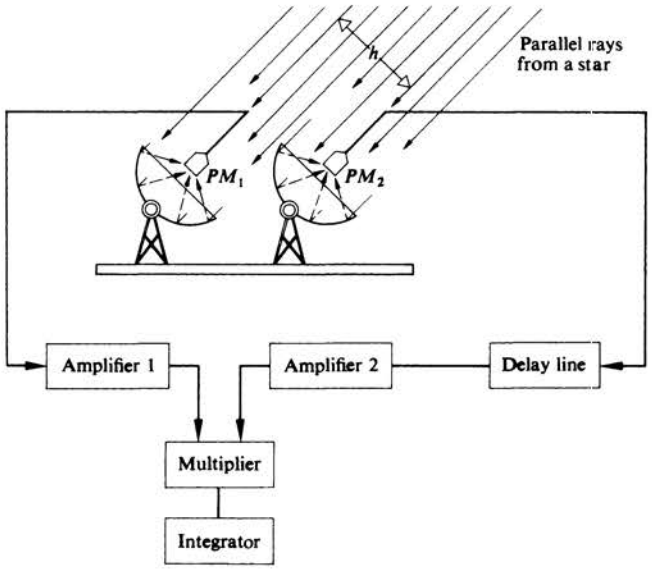


Figure 12.15 Stellar correlation interferometer.

10^{15} Hz. Even so, fast circuitry with roughly a 100-MHz pass bandwidth is required. In actuality the detectors have a finite resolving time T, so that the signal currents $\mathcal{I}_1$ and $\mathcal{I}_2$ are actually proportional to averages of $I_1(t)$ and $I_2(t)$ over T and not their instantaneous values. In effect, the measured fluctuations

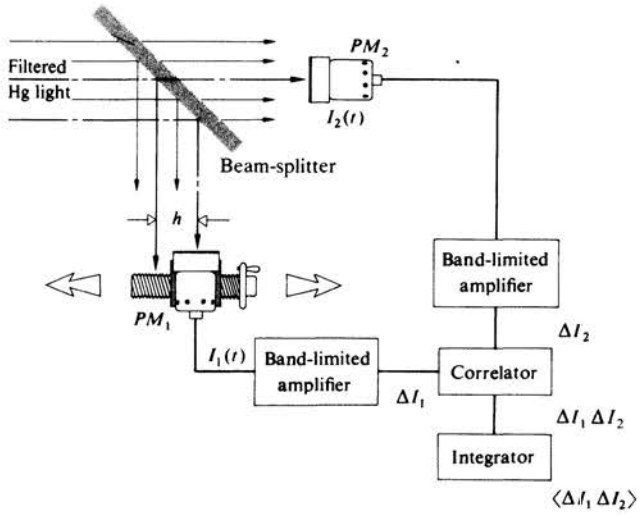


Figure 12.16 Hanbury–Brown and Twiss experiment.

are smoothed out, as illustrated by the dashed curve of Fig. 12.14*b*. For $T > \Delta t_c$, which is normally the case, this just leads to a reduction, by a factor of $\Delta t_c/T$, in the correlation actually observed:

$$\langle \Delta\mathcal{I}_1(t)\Delta\mathcal{I}_2(t)\rangle = \langle \mathcal{I}_1\rangle\langle \mathcal{I}_2\rangle \frac{\Delta t_c}{T} |\tilde{\gamma}_{12}(0)|^2 \qquad (12.36)$$

For example, in the preceding laboratory arrangement, the filtered mercury light had a coherence time of about 1 ns, while the electronics had a reciprocal pass bandwidth or effective integration time of ≈40 ns. Note that Eq. (12.36) isn't any different conceptually from Eq. (12.35)—it's just been made a bit more realistic.

Shortly after their successful laboratory experiment, Hanbury–Brown and Twiss constructed the stellar interferometer shown in Fig. 12.15. Searchlight mirrors were used to collect starlight and focus it onto two photomultipliers. One arm contained a delay line, so that the mirrors could physically be located at the same height, with compensation for any differences in the arrival times of the light. The measurement of $\langle \Delta\mathcal{I}_1(t)\Delta\mathcal{I}_2(t)\rangle_T$ at various separations of the detectors allowed the square of the modulus of the degree of coherence, $|\tilde{\gamma}_{12}(0)|^2$, to be deduced, and this in turn yielded the angular diameter of the source, just as it did with the Michelson stellar interferometer. This time, however, the separation h could be very large, because one no longer had to worry about messing up the phase of the waves, as was the case in the Michelson device. There, a slight shift in a mirror of a fraction of a wavelength was fatal. Here, in contrast, the phase was discarded, so that the mirrors didn't even have to be of high optical quality. The star Sirius was the first to be examined, and it was found to have an angular diameter of 0.006 9 second of arc. More recently, a correlation interferometer with a baseline of 618 feet has been constructed at Narrabri, Australia. For certain stars, angular diameters of as little as 0.000 5 second of arc can be measured with this instrument—that's a long way from the angular diameter of Betelgeuse (0.047 second of arc).*

The electronics involved in irradiance correlation could be greatly simplified if the incident light were very nearly monochromatic and of considerably higher flux density. Laserlight isn't thermal and doesn't display the same statistical fluctuations, but it can nonetheless be used to generate *pseudother-*

*For a discussion of the photon aspects of irradiance correlation, see Garbuny, *Optical Physics*, Section 6.2.5.2, or Klein, *Optics*, Section 6.4.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

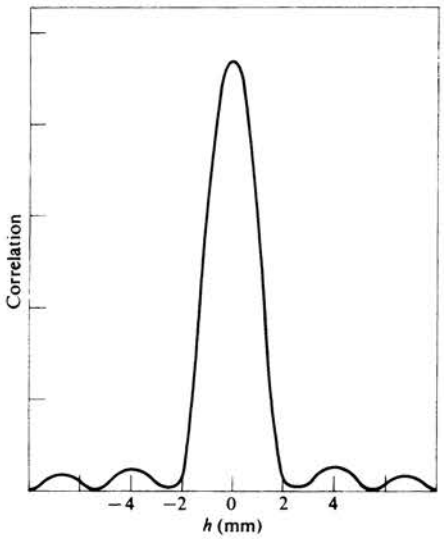


Figure 12.17 A correlation function for a pseudothermal source. [From A. B. Haner and N. R. Isenor, *Am J. Phys.* **38**, 748 (1970).]

*mal** light. A pseudothermal source is composed of an ordinary bright source (a laser is most convenient) and a moving medium of *nonuniform* optical thickness, such as a rotating ground glass disk. If the scattered beam emerging from a stationary piece of ground glass is examined with a *sufficiently slow detector*, the inherent irradiance fluctuations will be smoothed out completely. By setting the ground glass in motion, irradiance fluctuations appear with a simulated coherence time commensurate with the disk's speed. In effect, one has an extremely brilliant thermal source of variable Δt_c (from, say, 1 s to 10^{-5} s), which can be used to examine a whole range of coherence effects. For example, Fig. 12.17 shows the correlation function, which is proportional to $[2J_1(u)/(u)]^2$, for a pseudothermal circular aperture source determined from irradiance fluctuations. The experiment setup resembles that of Fig. 12.16, although the electronics is considerably simpler.*

*See W. Martienssen and E. Spiller, "Coherence and Fluctuations in Light Beams," *Am J. Phys.* **32**, 919 (1964), and A. B. Haner and N. R. Isenor, "Intensity Correlations from Pseudothermal Light Sources," *Am. J. Phys.* **38**, 748 (1970). Both of these articles are well worth studying.

*A good overall reference for this chapter is the review article by L. Mandel and E. Wolf, "Coherence Properties of Optical Fields," *Revs. Modern Phys.* **37**, 231 (1965); this is rather heavy reading. Take a look at K. I. Kellermann, "Intercontinental Radio Astronomy," *Sci. Am.* **226**, 72 (February 1972).

PROBLEMS

Complete solutions to all problems—except those with an asterisk—can be found in the back of the book.

12.1 Suppose we set up a fringe pattern using a Michelson Interferometer with a mercury vapor lamp as the source. Switch on the lamp in your mind's eye and discuss what will happen to the fringes as the mercury vapor pressure builds to its steady-state value.

12.2* We wish to examine the irradiance produced on the plane of observation in Young's Experiment when the slits are illuminated simultaneously by two monochromatic plane waves of somewhat different frequency, E_1 and E_2. Sketch these against time, taking $\lambda_1 = 0.8\lambda_2$. Now draw the product $E_1 E_2$ (at a point P) against time. What can you say about its average over a relatively long interval? What does $(E_1 + E_2)^2$ look like? Compare it with $E_1^2 + E_2^2$. Over a time that is long compared with the periods of the waves, approximate $\langle (E_1 + E_2)^2 \rangle_T$.

12.3* With the previous problem in mind, now consider things spread across space at a given moment in time. Each wave separately would result in an irradiance distribution I_1 and I_2. Plot both on the same space axis and then draw their sum $I_1 + I_2$. Discuss the meaning of your results. Compare your work with Fig. 7.16. What happens to the net irradiance as more waves of different frequency are added in? Explain in terms of the coherence length. Hypothetically, what

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

would happen to the pattern as the frequency bandwidth approached infinity?

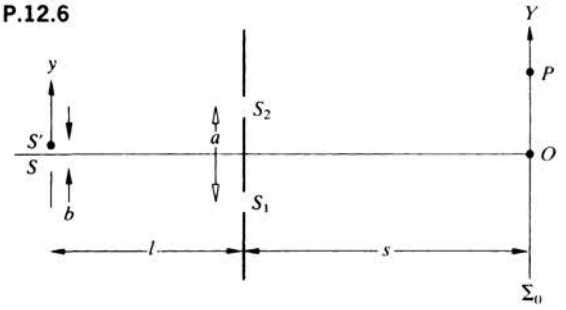
12.4 With the previous problem in mind, return to the autocorrelation of a sine function, shown in Fig. 11.37. Now suppose we have a signal composed of a great many sinusoidal components. Imagine that you take the autocorrelation of this complicated signal and plot the result (use three or four components to start with), as in part (*e*) of Fig. 11.37. What will the autocorrelation function look like when the number of waves is very large and the signal resembles random noise? What is the significance of the $\tau = 0$ value? How does this compare with the previous problem?

12.5* Imagine that we have the arrangement depicted in Fig. 12.3. If the separation between fringes (max. to max.) is 1 mm and if the projected width of the source slit on the screen is 0.5 mm, compute the visibility.

12.6 Referring to the slit source and pinhole screen arrangement of Fig. P.12.6, show by integration over the source that

$$I(Y) \propto b + \frac{\sin(\pi a/\lambda l)b}{\pi a/\lambda l} \cos(2\pi aY/\lambda s)$$

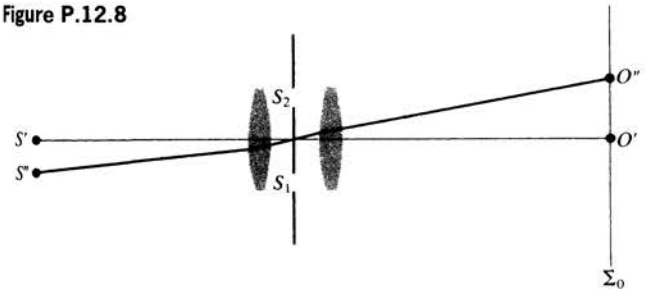
Figure P.12.6



12.7 Carry out the details leading to the expression for the visibility given by Eq. (12.22).

12.8 Under what circumstances will the irradiance on Σ_o in Fig. P.12.8 be equal to $4I_0$, where I_0 is the irradiance due to either incoherent point source alone?

Figure P.12.8



12.9* Suppose we set up Young's Experiment with a small circular hole of diameter 0.1 mm in front of a sodium lamp ($\bar{\lambda}_0 = 589.3$ nm) as the source. If the distance from the source to the slits is 1 m, how far apart will the slits be when the fringe pattern disappears?

12.10 Taking the angular diameter of the Sun viewed from the Earth to be about $1/2°$, determine the diameter of the corresponding area of coherence, neglecting any variations in brightness across the surface.

12.11 Show that Eqs. (12.34) and (12.35) follow from Eqs. (12.32) and (12.33).

12.12* Return to Eq. (12.21) and separate it into two terms representing a coherent and an incoherent contribution, the first arising from the superposition of two coherent waves with irradiances of $|\tilde{\gamma}_{12}(\tau)|I_1$ and $|\tilde{\gamma}_{12}(\tau)|\ I_2$ having relative phase of $\alpha_{12}(\tau) - \varphi$, and the second from the superposition of incoherent waves of irradiance $[1 - |\tilde{\gamma}_{12}(\tau)|]I_1$ and $[1 - |\tilde{\gamma}_{12}(\tau)|]I_2$. Now derive expressions for $I_{\text{coh}}/I_{\text{incoh}}$ and for $I_{\text{incoh}}/I_{\text{total}}$. Discuss the physical significance of this alternative formulation and how we might view the visibility of fringes in terms of it.

12.13 Imagine that we have Young's Experiment, where one of the two pinholes is now covered by a neutral-density filter that cuts the irradiance by a factor of 10, and the other hole is covered by a transparent sheet of glass, so there is no relative phase shift introduced. Compute the visibility in the hypothetical case of completely coherent illumination.

12.14* Suppose that Young's double-slit apparatus is illuminated by sunlight with a mean wavelength of 550 nm. Determine the separation of the slits that would cause the fringes to vanish.

12.15 We wish to construct a double-pinhole setup illuminated by a uniform, quasimonochromatic, incoherent slit source of mean wavelength 500 nm and width b, a distance of 1.5 m from the aperture screen. If the pinholes are 0.50 mm apart, how wide can the source be if the visibility of the fringes on the plane of observation is not to be less than 85%?

12.16* Suppose that we have an incoherent, quasimonochromatic, uniform slit source, such as a discharge lamp with a mask and filter in front of it. We wish to illuminate a region on an aperture screen 10.0 m away, such that the modulus of the complex degree of coherence everywhere within a region 1.0 mm wide is equal to or greater than 90% when the wavelength is 500 nm. How wide can the slit be?

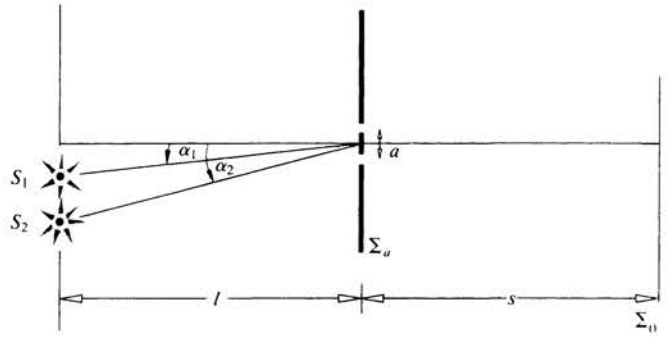
Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

12.17* Figure P.12.17 shows two incoherent quasimonochromatic point sources illuminating two pinholes in a mask. Show that the fringes formed on the plane of observation have minimum visibility when

$$a(\alpha_2 - \alpha_1) = \tfrac{1}{2}m$$

where $m = \pm 1, \pm 3, \pm 5 \ldots$.

Figure P.12.17



12.18 Imagine that we have a wide quasimonochromatic source ($\lambda = 500$ nm) consisting of a series of vertical, incoherent, infinitesimally narrow line sources, each separated by 500 μm. This is used to illuminate a pair of exceedingly narrow vertical slits in an aperture screen 2.0 m away. How far apart should the apertures be to create a fringe system of maximum visibility?

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

13 Modern Optics: Lasers and Other Topics

13.1 Lasers and Laserlight

During the early 1950s a remarkable device known as the *maser* came into being through the efforts of a number of scientists. Principal among these people were Charles Hard Townes of the United States and Alexandr Mikhailovich Prokhorov and Nikolai Gennadievich Basov of the USSR, all of whom shared the 1964 Nobel Prize in Physics for their work. The maser, which is an acronym for Microwave Amplification by Stimulated Emission of Radiation, is, as the name implies, an extremely low-noise, microwave amplifier.* It functioned in what was then a rather unconventional way, making direct use of the quantum-mechanical interaction of matter and radiant energy. Almost immediately after its inception, speculation arose as to whether or not the same technique could be extended into the optical region of the spectrum. In 1958 Townes and Arthur L. Schawlow prophetically set forth the general physical conditions that would have to be met in order to achieve Light Amplification by Stimulated Emission of Radiation. And then in July of 1960 Theodore H. Maiman announced the first successful operation of an optical maser or **laser**—certainly one of the great milestones in the history of Optics, and indeed in the history of science, had been achieved.

The laser is a quantum-mechanical device that manages to produce its "marvelous light" by taking advantage of the subtle ways in which atoms interact with electromagnetic radiation. To gain a solid, if only introductory, understanding of how the laser works and what makes its emissions so special, we'll first lay out some basic theory about ordinary thermal sources, such as lightbulbs and stars. That will require an introduction to blackbody radiation, but those insights are also basic to any treatment of the interaction of EM-radiation and matter. To that will be added a discussion of the Boltzmann distribution (p. 586) as applied to atomic energy levels. With this to stand on, we can appreciate the central notion of stimulated emission via the Einstein A and B coefficients (p. 586); the rest, more or less, follows.

*See James P. Gorden, "The Maser," *Sci. Am.* 199, **42** (December 1958).

13.1.1 Radiant Energy and Matter in Equilibrium

It shouldn't surprise anyone that if physics was to be turned on its head, it would be done while trying to figure out what light (i.e., radiant energy) was all about. Quantum theory had its earliest beginnings back in 1859 with the study of a seemingly obscure phenomenon known as **blackbody radiation**. That year, Charles Darwin published *The Origin of Species*, and Gustav Robert Kirchhoff proffered an intellectual challenge that would lead to a revolution in physics.

Kirchhoff was involved in analyzing the way bodies in thermal equilibrium behave in the process of exchanging radiant energy. This *thermal radiation* is electromagnetic energy emitted by all objects, the source of which is the random motion of their constituent atoms. He characterized the abilities of a body to emit and absorb electromagnetic energy by an *emission coefficient* ε_λ and an *absorption coefficient* α_λ. Epsilon is the energy per unit area per unit time emitted in a tiny wavelength range around λ (in units of $W/m^2/m$): thermal radiation comprises a wide range of frequencies, and an energy-measuring device by necessity admits a band of wavelengths. *Alpha is the fraction of the incident radiant energy absorbed per unit area per unit time in that wavelength range*; it's unitless. The emission and absorption coefficients depend on both the nature of the surface of the body (color, texture,

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

etc.) and the wavelength—a body that emits or absorbs well at one wavelength may emit or absorb poorly at another.

Consider an isolated chamber of some sort in thermal equilibrium at a fixed temperature T. Presumably, it would be filled with radiant energy at a myriad of different wavelengths—think of a glowing furnace. Kirchhoff assumed there was some formula, or *distribution function* $I_\lambda(\lambda)$, which depends on T and which provides values of the ***energy per unit area per unit time at each wavelength***; call it the ***spectral flux density*** within the cavity (or ***spectral exitance*** when it leaves it). He concluded that the *total* amount of energy at all wavelengths being absorbed by the walls versus the amount emitted by them must be the same, or else T would change, and it doesn't. Furthermore, Kirchhoff argued that if the walls were made of different materials (which behave differently with T), that same balance would have to apply for *each* wavelength range individually. The energy absorbed at λ, namely, $\alpha_\lambda I_\lambda$, must equal the energy radiated, ε_λ, *and this is true for all materials no matter how different.* **Kirchhoff's Radiation Law** is therefore

$$\frac{\varepsilon_\lambda}{\alpha_\lambda} = I_\lambda \tag{13.1}$$

wherein the distribution I_λ, in units of $J/m^3 \cdot s$ or W/m^3, is a universal function the same for every type of cavity wall regardless of material, color, size, and shape and is only dependent on T and λ. That's quite extraordinary! Still, the British ceramist Thomas Wedgwood had commented long before (1792) that objects in a fired kiln all turned glowing red together along with the furnace walls, regardless of their size, shape, or material constitution.

Although Kirchhoff could not provide the energy distribution function in general, he did observe that a perfectly absorbing body, one for which $\alpha_\lambda = 1$, will appear black and, in that special case, $I_\lambda = \varepsilon_\lambda$. Moreover, the distribution function for a perfectly black object is the same as for an isolated chamber at that same temperature (visualize such a blackbody at equilibrium inside a hot oven). The radiant energy distribution at equilibrium within an isolated cavity is in every regard the same, "as if it came from a completely blackbody of the same temperature." Therefore ***the energy that would emerge from a small hole in the chamber should be identical to the radiation coming from a perfectly black object at the same temperature.***

The scientific community accepted the challenge of experimentally determining I_λ, but the technical difficulties were great and progress came slowly. The basic setup (Fig. 13.1*a*) is simple enough, although coming up with a reliable source was a daunting problem for a long time. Data must be extracted that is independent of the construction of the specific detector, and so the best thing to plot is the radiant energy per unit time, which enters the detector per unit area (of the entrance

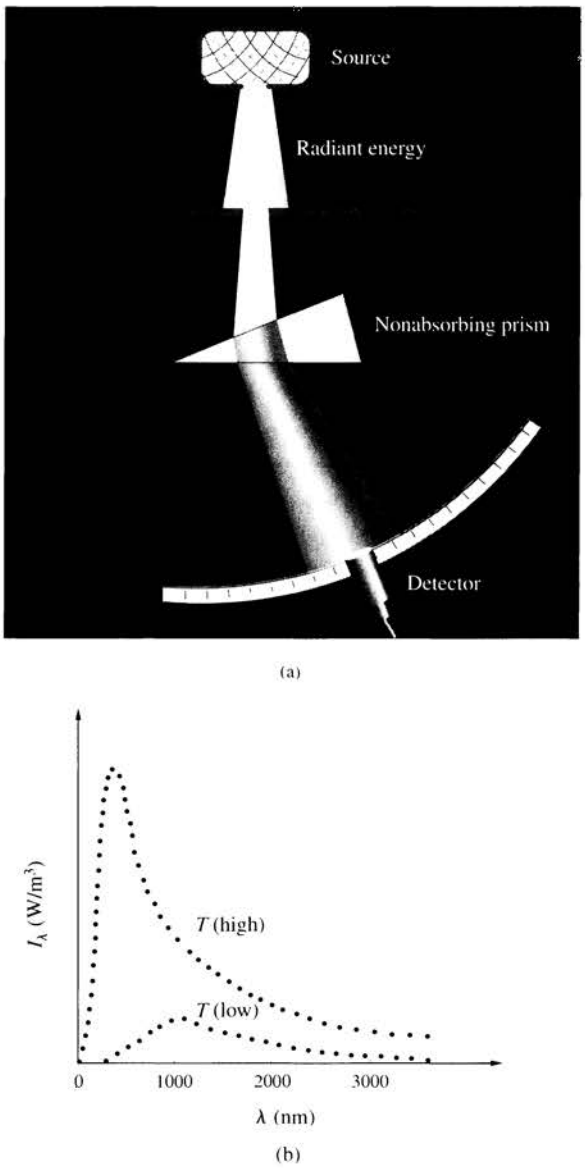


Figure 13.1 (*a*) A basic experimental setup for measuring blackbody radiation. (*b*) Values of I_λ at successive wavelengths as measured by a detector. Each curve corresponds to a specific source temperature.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

window) per unit wavelength range (admitted by the detector). The kind of curves that were ultimately recorded are shown in Fig. 13.1*b*, and each is a plot of I_λ at a specific temperature.

Stefan-Boltzmann Law

In 1865 John Tyndall published some experimental results, including the determination that the total energy emitted by a heated platinum wire was 11.7 times greater when operating at 1200°C (1473 K) than it was at 525°C (798 K). Rather amazingly, Josef Stefan (1879) noticed that the ratio of $(1473\ \text{K})^4$ to $(798\ \text{K})^4$ was 11.6, nearly 11.7, and he surmised that the rate at which energy is radiated is proportional to T^4. In this observation Stefan was quite right (and quite lucky); Tyndall's results were actually far from those of a blackbody. Still, the conclusion was subsequently given a theoretical foundation by L. Boltzmann (1884). His was a traditional treatment of the radiation pressure exerted on a piston in a cylinder using the laws of thermodynamics and Kirchhoff's Law. The analysis progressed in much the same way one would treat a gas in a cylinder, but instead of atoms, the active agency was electromagnetic waves. The resulting Stefan–Boltzmann Law for blackbodies (which is correct, though nowadays we would derive it differently) is

$$P = \sigma A T^4 \tag{13.2}$$

where P is the total radiant power at all wavelengths, A is the area of the radiating surface, T is the absolute temperature in kelvins, and σ is a universal constant now given as

$$\sigma = 5.670\,33 \times 10^{-8}\ \text{W/m}^2\cdot\text{K}^4$$

The total area under any one of the blackbody-radiation curves of Fig. 13.1*b* for a specific T is the power per unit area, and from Eq. (13.2) that's just $P/A = \sigma T^4$.

Real objects are not perfect blackbodies; carbon black has an absorptivity of nearly one, but only at certain frequencies (obviously including the visible). Its absorptivity is much lower in the far infrared. Nonetheless, most objects resemble a blackbody (at least at certain temperatures and wavelengths)—you, for instance, are nearly a blackbody for infrared. Because of that, it's useful to write a similar expression for ordinary objects. This can be done by introducing a multiplicative factor called the total emissivity (ε), which relates the radiated power to that of a blackbody for which $\varepsilon = 1$, at the same temperature, thus

$$P = \varepsilon \sigma A T^4$$

Table 13.1 provides a few values of ε (at room temperature), where $0 < \varepsilon < 1$. Note that emissivity is unitless.

If an object with a *total absorptivity* of α is placed in an enclosure such as a cavity or a room having an emissivity ε_e and a temperature T_e, the body will radiate at a rate $\varepsilon\sigma A T^4$ and absorb energy inside the enclosure at a rate $\alpha(\varepsilon_e \sigma A T_e^4)$. Yet at any temperature at which the body and enclosure are in equilibrium (i.e., $T = T_e$), these rates must be equal; hence, $\alpha\varepsilon_e = \varepsilon$ and that has to be true for all temperatures. The net power radiated (when $T > T_e$) or absorbed (when $T < T_e$) by the body is then

$$P = \varepsilon \sigma A (T^4 - T_e^4)$$

All bodies not at zero kelvin radiate, and the fact that T is raised to the fourth power makes the radiation highly sensitive to temperature changes. When a body at 0°C (273 K) is brought up to 100°C (373 K), it radiates about 3.5 times the previous power. Increasing the temperature increases the net power radiated; that's why it gets more and more difficult to increase the temperature of an object. (Try heating a steel spoon to 1300°C.) Increasing the temperature of an object also shifts the emitted distribution of energy among the various wavelengths present. At the moment when the filament of a lightbulb "blows," the resistance, current, and temperature rise; it goes from its normal operating reddish-white color to a bright flash of blue-white.

TABLE 13.1 Some Representative Values of Total Emissivity*

Material	ε
Aluminum foil	0.02
Copper, polished	0.03
Copper, oxidized	0.5
Carbon	0.8
White paint, flat	0.87
Red brick	0.9
Concrete	0.94
Black paint, flat	0.94
Soot	0.95

*T = 300 K, room temperature.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Wien Displacement Law

Perhaps the last notable success in applying classical theory to the problem of blackbody radiation came in 1893 at the hands of the German physicist and Nobel laureate Wilhelm Otto Fritz Franz Wien (1864–1928), known to his friends as Willy. He derived what is today called the **Displacement Law**. Each blackbody curve reaches a maximum height at a value of wavelength (λ_{max}) that is particular to it and therefore to the absolute temperature T. At that wavelength, the blackbody radiates the most energy. Wien was able to show that

$$\lambda_{max} T = \text{constant} \qquad (13.3)$$

where the constant was found experimentally to be 0.002 898 m·K. The peak wavelength is inversely proportional to the temperature. *Raise the temperature, and the bulk of the radiation shifts to shorter wavelengths and higher frequencies* (see the dashed curve in Fig. 13.2). As a glowing coal or a blazing star gets hotter, it goes from IR warm to red-hot to blue-white. A person or a piece of wood, both only roughly blackbodies, radiates for the most part in the infrared and would only begin to glow faintly in the visible at around 600°C or 700°C, long after either had decomposed. The bright cherry red of a chunk of "red" hot iron sets in at around 1300°C.

In 1899 researchers greatly advanced the state of experimentation by using, as a source of blackbody radiation, a small hole in a heated cavity (Fig. 13.3). Energy entering such an aperture reflects around inside until it's absorbed. (The pupil of the eye appears black for precisely the same reason.) A near-perfect absorber is a near-perfect emitter, and the region of a small hole in the face of an oven is a wonderful source of *blackbody radiation*.

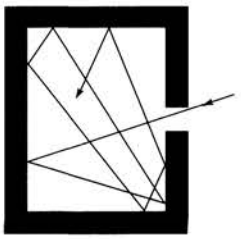


Figure 13.3 Radiant energy entering a tiny hole in a chamber will rattle around with little chance of ever emerging through the aperture, and so the hole looks black. In reverse, the aperture of a heated chamber appears as a blackbody source.

It was at this point in time that classical theory began to falter. All attempts to fit the entire radiation curve (Fig. 13.2) with some theoretical expression based on electromagnetism led only to the most limited successes. Wien produced a formula that agreed with the observed data fairly well in the short wavelength region but deviated from it substantially at large λ. Lord Rayleigh and later Sir James Jeans (1877–1946) developed a description in terms of the standing-wave modes of the field within the enclosure. But the resulting *Rayleigh–Jeans formula* matched the experimental curves only in the very long wavelength region. The failure of classical theory was totally inexplicable; a turning point in the history of physics had arrived.

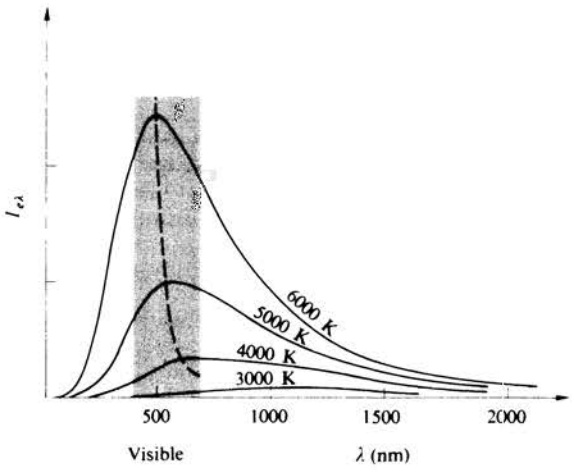


Figure 13.2 Blackbody radiation curves. The hyperbola passing through peak points corresponds to Wien's Law.

Planck Radiation Law

Max Karl Ernst Ludwig Planck at 42 was the somewhat reluctant father of quantum theory. Like so many other theoreticians at the turn of the century, he, too, was working on blackbody radiation. But Planck would succeed not only in producing Kirchhoff's distribution function, but also in turning physics upside-down in the process. We cannot follow the details of his derivation here; besides, the original version was wrong. (Bose and Einstein corrected it years later.) Still, it had such a powerful impact that it's worth looking at some of the features that are right.

Planck knew that if an arbitrary distribution of energetic molecules was injected into a constant-temperature chamber, it would ultimately rearrange itself into the Maxwell–Boltzmann distribution of speeds as it inevitably reached equilibrium. Presumably, if an arbitrary distribution of radiant energy

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

is injected into a constant-temperature cavity, it, too, will ultimately rearrange itself into the Kirchhoff distribution of energies as it inevitably reaches equilibrium.

In October 1900, Planck produced a distribution formula that was based on the latest experimental results. This mathematical contrivance, concocted "by happy guesswork," fit all the data available. It contained two fundamental constants, one of which (h) would come to be known as **Planck's Constant**. That much by itself was quite a success, even if it didn't explain anything. Although Planck had no idea of it at the time, he was about to take a step that would inadvertently revolutionize our perception of the physical Universe.

Naturally enough, Planck set out to construct a theoretical scheme that would logically lead to the equation he had already devised. He assumed that the radiation in a chamber interacted with simple microscopic oscillators of some unspecified type. These vibrated on the surfaces of the cavity walls, absorbing and reemitting radiant energy independent of the material. (In fact, the atoms of the walls do exactly that. Because of their tightly packed configuration in the solid walls, the atoms interact with a huge number of their neighbors. That completely blurs their usual characteristic sharp resonance vibrations, allowing them to oscillate over a broad range of frequencies and emit a continuous spectrum.) Try as he might, Planck was unsuccessful. At that time, he was a devotee of E. Mach, who had little regard for the reality of atoms, and yet the obstinate insolubility of the problem ultimately led Planck to "an act of desperation." He hesitantly turned to Boltzmann's "distasteful" statistical method, which had been designed to deal with the clouds of atoms that constitute a gas.

Boltzmann, the great proponent of the atom, and Planck were intellectual adversaries for a while. And now Planck was forced to use his rival's statistical analysis, which—ironically—he would misapply. If Boltzmann's scheme for counting atoms was to be applied to something continuous, such as energy, some adjustments would have to be made in the procedure. Thus, according to Planck, the total energy of the oscillators had to be thought of, at least temporarily, as apportioned into "energy elements" so that they could be counted. These energy elements were given a value proportional to the frequency ν of the resonators. Remember that he already had the formula he was after, and in it there appeared the term $h\nu$. Planck's Constant,

$$6.626\,075\,5 \times 10^{-34}\,\text{J}\cdot\text{s} \quad \text{or} \quad 4.135\,669\,2 \times 10^{-15}\,\text{eV}\cdot\text{s}$$

is a very small number and so $h\nu$, which has the units of energy, is itself a very small quantity. Accordingly, he set the value of the energy element equal to it: $\mathscr{E} = h\nu$.

This was a statistical analysis, and counting was central. Still, when the method was applied as Boltzmann intended, it naturally smoothed out energy, making it continuous as usual. Again, we needn't worry about the details. The amazing thing was that Planck had stumbled on a hidden mystery of nature: **energy is quantized**—it comes in tiny bursts, but he didn't realize it then.

Planck derived the following formula for the spectral exitance (or spectral irradiance)—which he had already arrived at by fitting curves to the data—and it's the answer to Kirchhoff's challenge:

$$I_\lambda = \frac{2\pi hc^2}{\lambda^5}\left[\frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1}\right] \tag{13.4}$$

where k_B is Boltzmann's Constant. This is **Planck's Radiation Law**, and, of course, it fit blackbody data splendidly (Fig. 13.4). Notice how the expression contains the speed of light, Boltzmann's Constant, and Planck's Constant (h). It bridges Electromagnetic Theory to the domain of the atom.

Although Eq. (13.4) represents a great departure from previous ideas, Planck did not mean to break with classical theory. It would have been unthinkable for him even to suggest that radiant energy was anything but continuous. "That energy is forced, at the outset, to remain together in certain quanta ...," Planck later remarked, "was purely a formal assumption and I really did not give it much thought." It was only around 1905, at the hands of a much bolder thinker, Albert Einstein, that we learned that the atomic oscillators

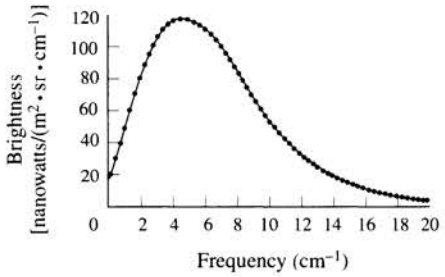


Figure 13.4 The cosmic background radiation of the Universe. Since the creation of the Universe with the Big Bang, it has expanded and cooled. The data points (measured in the microwave spectrum) were detected by the Cosmic Background Explorer (COBE) satellite. The solid line is the Planck blackbody curve for a temperature of 2.735 ± 0.06 K.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

were real and that their energies were quantized. Each oscillator could only exist with an energy that was a whole-number multiple of $h\nu$ (a little like the *gravitational*-PE of someone walking up a flight of stairs). Moreover, **radiant energy itself is quantized**, existing in localized blasts of an amount $\mathscr{E} = h\nu$.

13.1.2 Stimulated Emission

The LAser accomplishes "light amplification" by making use of energetic atoms in a medium to reinforce the light field. Let's therefore examine the manner in which the energy states of a system of atoms at some arbitrary temperature is normally distributed. The problem is part of the broader discipline of Statistical Mechanics and is addressed specifically in terms of the Maxwell–Boltzmann distribution.

Population of Energy Levels

Imagine a chamber filled with a gas in equilibrium at some temperature T. If T is relatively low, as it is in a typical room, most of the atoms will be in their ground states, but a few will momentarily pick up enough energy to "rise" into an excited state. The classical Maxwell–Boltzmann distribution maintains that, on average, a number of atoms per unit volume, N_i, will be in any excited state of energy $\mathscr{E}_i$ such that

$$N_i = N_0 e^{-\mathscr{E}_i/k_B T}$$

where N_0 is a constant for a given temperature. *The higher the energy state, that is, the greater the value of* $\mathscr{E}$ *(the smaller is the exponential) and the fewer atoms there will be in that state.*

Since we will be interested in atomic transition between arbitrary states, consider the jth energy level where $\mathscr{E}_j > \mathscr{E}_i$. Then for it $N_j = N_0 e^{-\mathscr{E}_j/k_B T}$, and the ratio of the populations occupying these two states is

$$\frac{N_j}{N_i} = \frac{e^{-\mathscr{E}_j/k_B T}}{e^{-\mathscr{E}_i/k_B T}} \tag{13.5}$$

This is the *relative population*, and it follows that

$$N_j = N_i e^{-(\mathscr{E}_j - \mathscr{E}_i)/k_B T} = N_i e^{-h\nu_{ji}/k_B T} \tag{13.6}$$

where use was made of the fact that a transition for the jth-state to the ith-state corresponds to an energy change of $(\mathscr{E}_j - \mathscr{E}_i)$ and since such transitions are accompanied by the emission of a photon of frequency ν_{ji}, we can substitute $(\mathscr{E}_j - \mathscr{E}_i) = h\nu_{ji}$.

The Einstein *A* and *B* Coefficients

In 1916 Einstein devised an elegant and rather simple theoretical treatment of the dynamic equilibrium existing for a material medium bathed in electromagnetic radiation, absorbing and reemitting. The analysis was used to affirm Planck's Radiation Law, but more importantly it also created the theoretical foundation for the laser. The reader should already be familiar with the basic mechanism of *absorption* (see Fig. 3.35, p. 64). Suppose the atom is in its lowest energy or ground-state configuration. A photon having an adequate amount of energy interacts with the atom, imparting that energy to the atom, thereby causing the electron cloud to take on a new configuration. The atom jumps into a higher-energy excited state (Fig. 13.5). In a dense medium, the atom is likely to interact with its jiggling neighbors and pass off its bounty of energy via collisions.

Such an excess-energy configuration is usually (though not always) exceedingly short-lived, and in 10 ns or so, without the intercession of any external influence, the atom will emit its overload of energy as a photon. As it does, it reverts to a stable state in a process called *spontaneous emission* (Fig. 13.5*b*).

The remarkable thing is that there is a third alternative process, one first appreciated by Einstein and crucial to the operation of the laser—which wasn't invented until almost a half century later. For a medium inundated with EM-radiation, it's possible for a photon to interact with an excited atom while that atom is still in its higher-energy configuration. The atom can then dump its excess energy in-step with the incoming photon, in a process now called **stimulated emission** (Fig. 13.5).

In the case of absorption, the rate-of-change of the number of atoms in some initial state, as they leave to some higher state, must depend on the strength of the photon field inundating those atoms. In other words, it must depend on the energy density u, given by Eq. (3.34), but more specifically it must depend on the energy density in the frequency range driving the transition, that is, the spectral energy density u_ν, which has units of joules per meter-squared per inverse second ($J \cdot s/m^2$). (Note that if we consider the radiation field as a photon gas, the spectral energy density can be thought of as the photon

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

density per unit frequency range.) The rate-of-change of the number of atoms, the **transition rate**, will also be proportional to the population, that is, the number density of atoms in that state (N_i); the more there are, the more can leave (via absorption) per second. Because the process is driven by the photon field, let's call it **stimulated absorption**, whereupon the transition rate is

[stimulated absorption]
$$\left(\frac{dN_i}{dt}\right)_{\text{ab}} = -B_{ij}N_i u_\nu \tag{13.7}$$

Here B_{ij} is a constant of proportionality, and the minus arises because N_i is decreasing. Similarly, for stimulated emission

[stimulated emission]
$$\left(\frac{dN_j}{dt}\right)_{\text{st}} = -B_{ji}N_j u_\nu \tag{13.8}$$

In the case of spontaneous emission, the process is independent of the field environment and

[spontaneous emission]
$$\left(\frac{dN_j}{dt}\right)_{\text{sp}} = -A_{ji}N_j \tag{13.9}$$

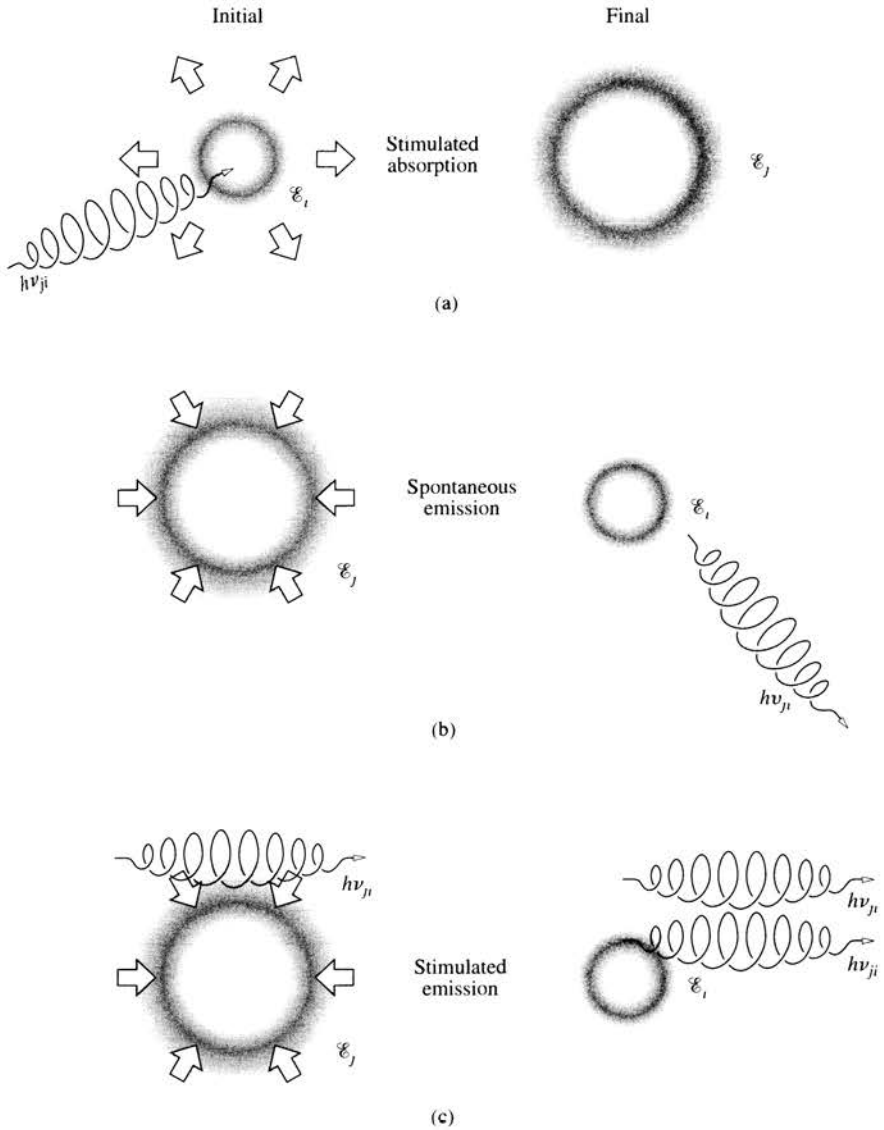


Figure 13.5 A schematic representation of (*a*) stimulated absorption, (*b*) spontaneous emission, and (c) stimulated emission.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Keep in mind that the transition rate, the number of atoms making transitions per second, divided by the number of atoms, is the probability of a transition occurring per second, $\mathscr{P}$. Consequently, the probability per second of spontaneous emission is $\mathscr{P}_{sp} = A_{ji}$.

For a single excited atom making a spontaneous transition to a lower state, the inverse of the transition probability per second is the **mean life** or **lifetime** of the excited state τ. Thus (operating under conditions that exclude any other mechanism but spontaneous emission), if N atoms are in that excited state, the total rate of transitions, that is, the number of emitted photons per second, is $N\mathscr{P}_{sp} = NA_{ji} = N/\tau$. A low-transition probability means a long lifetime.

The three constants A_{ji}, B_{ji}, and B_{ij} are Einstein's coefficients. Following his lead, we assume (1) that thermodynamic equilibrium exists between the radiation field and the atoms in it at any T; (2) that the energy density has the characteristics of a blackbody at T; and (3) that the number densities of the two states are in accord with the Maxwell–Boltzmann distribution.

Given that the system is in equilibrium, the rate of upward ($i \rightarrow j$) transitions must equal the rate of downward transitions ($j \rightarrow i$):

$$B_{ij}N_i u_\nu = B_{ji}N_j u_\nu + A_{ji}N_j$$

Dividing both sides by N_i and rearranging terms yields

$$\frac{N_j}{N_i} = \frac{B_{ij}u_\nu}{A_{ji} + B_{ji}u_\nu}$$

Making use of Eq. (13.6), that is, what we found from the application of the Maxwell–Boltzmann distribution, this becomes

$$e^{-h\nu_{ji}/k_B T} = \frac{B_{ij}u_\nu}{A_{ji} + B_{ji}u_\nu}$$

and solving for u_ν leads to

$$u_\nu = \frac{A_{ji}/B_{ji}}{(B_{ij}/B_{ji})e^{h\nu_{ji}/k_B T} - 1} \tag{13.10}$$

Here Einstein pointed out that as $T \rightarrow \infty$, the spectral energy density, that is, the spectral photon density, approaches infinity. Figure 13.3 shows that I_λ increases with T, and that implies that u_ν will behave in a like fashion. In fact, $I_\nu = \frac{c}{4}u_\nu$, a point we will address presently. In any event, since $e^0 = 1$, the only way u_ν will be large is if

$$B_{ij} = B_{ji} = B$$

for large T, but since these constants are temperature independent, they must be equal at all T. The probabilities of stimulated emission and absorption are $\mathscr{P}_{st} = B_{ji}u_\nu$ and $\mathscr{P}_{ab} = B_{ij}u_\nu$, respectively. Hence, **the probability of stimulated emission is identical to the probability of stimulated absorption**; an atom in the lower state is just as likely to make a stimulated transition up, as an excited atom is to make a stimulated transition down.

Simplifying the notation (let $A = A_{ji}$), Eq. (13.10) becomes

$$u_\nu = \frac{A}{B}\left[\frac{1}{e^{h\nu_{ji}/k_B T} - 1}\right] \tag{13.11}$$

The ratio A/B can be expressed via basic quantities by comparing this equation with

$$I_\lambda = \frac{2\pi hc^2}{\lambda^5}\left[\frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1}\right] \tag{13.4}$$

But first transform I_λ into I_ν where these are expressions for exitance (which is irradiance going outward) per interval $d\lambda$ and $d\nu$, respectively. Using the fact that $\lambda = c/\nu$, differentiating yields $d\lambda = -cd\nu/\nu^2$. Because $I_\lambda d\lambda = I_\nu d\nu$, and dropping the sign (since it just says that one differential increases while the other decreases), we get $I_\lambda c/\nu^2 = I_\nu$; and so

$$I_\nu = \frac{2\pi h\nu^3}{c^2}\left[\frac{1}{e^{\frac{h\nu}{k_B T}} - 1}\right] \tag{13.12}$$

Now as a last step we need only to compare the spectral energy density u_ν in the chamber with the spectral exitance,

$$I_\nu = \frac{c}{4}u_\nu \tag{13.13}$$

emerging from it. Rather than burden the reader with a complete derivation of this relationship, let it suffice merely to justify it. Keep in mind that I_ν corresponds to a flow of energy across a unit normal area, in one side and out the other—a beam leaving the chamber. In Section 3.3.1 we saw that the instantaneous flow of power per unit normal area, the Poynting vector, was given by $S = cu$ and so on average $I = cu$ for a beam. Inside a chamber, however, with light traveling in every direction, not all the photons that contribute to u will contribute to the exitance in a particular direction. Presumably, inside the chamber a unit area held horizontally would have as much energy flowing up through it as down. Moreover, only the components perpendicular to the area contribute to S, so a factor of $1/4$ is not unreasonable.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

From Eqs. (13.11), (13.12), and (13.14) it follows that

$$\frac{A}{B} = \frac{8\pi h\nu^3}{c^3} \tag{13.14}$$

The probability of spontaneous emission is proportional to the probability of stimulated emission; an atom susceptible to one mechanism is proportionately susceptible to the other. Lasers work by stimulated emission, and anything that enhances spontaneous emission (i.e., A) at the price of stimulated emission (i.e., B) can be expected to work to the detriment of the process. Because the ratio of A/B varies as ν^3, it would seem that X-ray lasers ought to be very difficult to build—they are!

Imagine a system of atoms in thermal equilibrium having only two possible states. Furthermore, require that the atoms have a long mean life so that we can ignore spontaneous emission. When the system is inundated by photons of the proper energy, stimulated absorption depopulates the lower i-level, while stimulated emission depopulates the upper j-level. The number of photons vanishing from the system per second via stimulated absorption is proportional to $\mathscr{P}_{ab}N_i$, and the number entering it via stimulated emission is proportional to $\mathscr{P}_{st}N_j$, but from the equality of the B-coefficients it follows that $\mathscr{P}_{st} = \mathscr{P}_{ab}$. Therefore $\mathscr{P}_{ab}N_j = \mathscr{P}_{st}N_j$. However, if the system is in thermal equilibrium, $N_i > N_j$, which means that the number of photons vanishing per second exceeds the number entering per second; there's a net absorption of photons by the lower state because there are more atoms in the lower state at any given temperature. The reverse would be true if we could create a situation—a *population inversion*—in which $N_i < N_j$; then stimulated emission would dominate over stimulated absorption.

13.1.3 The Laser

Consider an ordinary medium in which a few atoms are in some excited state; call it $|j\rangle$ to conform with quantum-mechanical notation. If a photon in an incident beam is to trigger one of these excited atoms into stimulated emission, it must have the frequency ν_{ji}, as per Fig. 13.5*c*. A remarkable feature of this process is that ***the emitted photon is in-phase with, has the polarization of, and propagates in the same direction as, the stimulating radiation***. The emitted photon is said to be in the same radiation mode as the incident wave and tends to add to it, increasing its flux density. However, since most atoms are ordinarily in the ground state, absorption is usually far more likely than stimulated emission.

This raises an intriguing point: What would happen if a substantial percentage of the atoms could somehow be excited into an upper state, leaving the lower state all but empty? For obvious reasons this is known as **population inversion**. An incident photon of the proper frequency could then trigger an avalanche of stimulated photons—*all in-phase*. The initial wave would continue to build, so long as there were no dominant competitive processes (such as scattering) and provided the population inversion could be maintained. In effect, energy (electrical, chemical, optical, etc.) would be pumped in to sustain the inversion, and a beam of light would be extracted after sweeping across the *active medium*.

The First (Pulsed Ruby) Laser

To see how all of this is accomplished in practice, let's take a look at Maiman's original device (Fig. 13.6). The first operative laser had as its active medium a small, cylindrical, synthetic, pale pink ruby, that is, an Al_2O_3 crystal containing about 0.05 percent (by weight) of Cr_2O_3. Ruby, which is still one of the most common of the crystalline laser media, had been used earlier in maser applications and was suggested for use in the laser by Schawlow. The rod's end faces were polished flat, parallel and normal to the axis. Then both were silvered (one only partially) to form a **resonant cavity**.

It was surrounded by a helical gaseous discharge flashtube, which provided broadband **optical pumping**. Ruby appears red because the chromium atoms have absorption bands in the blue and green regions of the spectrum (Fig. 13.7*a*). Firing the flashtube generates an intense burst of light lasting a few milliseconds. Much of this energy is lost in heat, but many of the Cr^{3+} ions are excited into the absorption bands. A simplified energy-level diagram appears in Fig. 13.7*b*. The excited ions rapidly relax (in about 100 ns), giving up energy to the crystal lattice and making nonradiative transitions. They preferentially drop "down" to a pair of closely spaced, especially long-lived, interim states. They remain in these so-called **metastable states** for up to several milliseconds ($\approx$3 ms at room temperature) before randomly, and in most cases spontaneously, dropping down to the ground state. This is accompanied by the emission of the characteristic red fluorescent radiation of ruby. The lower-level transition dominates, and the resulting emission occurs in a relatively broad spectral range centered about 694.3 nm; it emerges in all directions and is incoherent.

When the pumping rate is increased somewhat, a population inversion occurs, and the first few spontaneously emitted

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

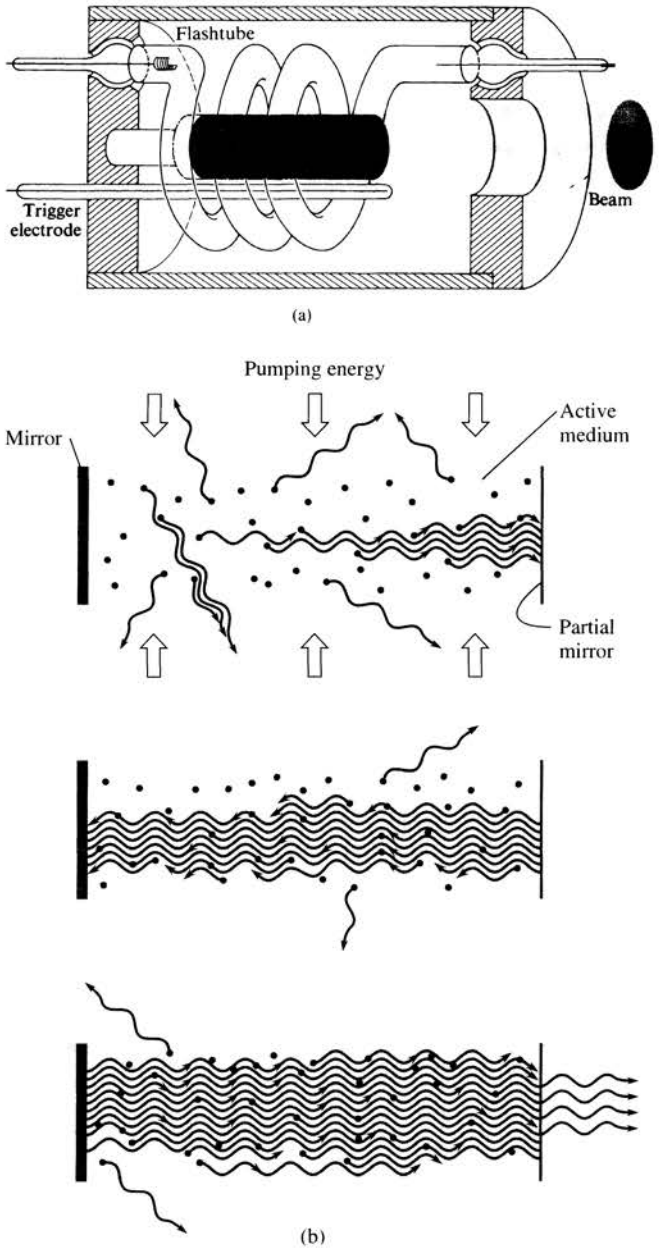


Figure 13.6 The first ruby-laser configuration, just about life-sized.

photons stimulate a chain reaction. One quantum triggers the rapid, in-phase emission of another, dumping energy from the metastable atoms into the evolving lightwave (Fig. 13.6*b*). The wave continues to grow as it sweeps back and forth across the active medium (provided enough energy is available to overcome losses at the mirrored ends). Since one of those reflecting surfaces was partially silvered, an intense pulse of red laser light (lasting about 0.5 ms and having a line-width of about 0.01 nm) emerges from that end of the ruby rod.

Notice how neatly everything works out. The broad absorption bands make the initial excitation rather easy, while the long lifetime of the metastable state facilitates the population inversion. The atomic system in effect consists of (1) the absorption bands, (2) the metastable state, and (3) the ground state. Accordingly, it is spoken of as a *three-level* laser.

Today's ruby laser is generally a high-power source of pulsed coherent radiation used extensively in work on interferometry, plasma diagnostics, holography, and so forth. Such devices operate with coherence lengths of from 0.1 m to 10 m.

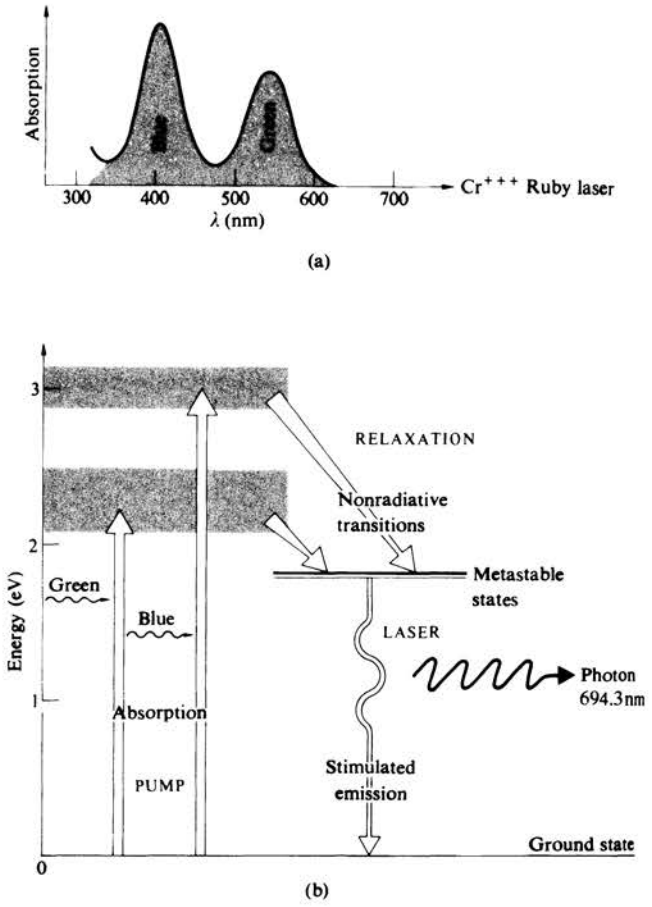


Figure 13.7 Ruby-laser energy levels.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Modern configurations usually have flat external mirrors, one totally and the other partially reflecting. As an oscillator, the ruby laser generates millisecond pulses in the energy range from around 50 J to 100 J, but by using a tandem oscillator-amplifier setup, energies well in excess of 100 J can be produced. The commercial ruby laser typically operates at a modest overall efficiency of less than 1%, producing a beam that has a diameter ranging from 1 mm to about 25 mm, with a divergence of from 0.25 mrad to about 7 mrad.

Optical Resonant Cavities

The resonant cavity, which in this case is of course a Fabry–Perot etalon, plays a significant role in the operation of the laser. In the early stages of the laser process, spontaneous photons are emitted in every direction, as are the stimulated photons. But all of these, with the singular exception of those propagating very nearly along the cavity axis, quickly pass out of the sides of the ruby. In contrast, the axial beam continues to build as it bounces back and forth across the active medium. This accounts for the amazing degree of collimation of the issuing laserbeam, which is then effectively a coherent plane wave. Although the medium acts to amplify the wave, the *optical feedback* provided by the cavity converts the system into an oscillator and hence into a light generator—the acronym is thus somewhat of a misnomer.

In addition, the disturbance propagating within the cavity takes on a standing-wave configuration determined by the separation (L) of the mirrors. The cavity resonates (i.e., standing waves exist within it) when there is an integer number (m) of half wavelengths spanning the region between the mirrors. The idea is simply that there must be a node at each mirror, and this can only happen when L equals a whole number multiple of $\lambda/2$ (where $\lambda = \lambda_0/n$). Thus

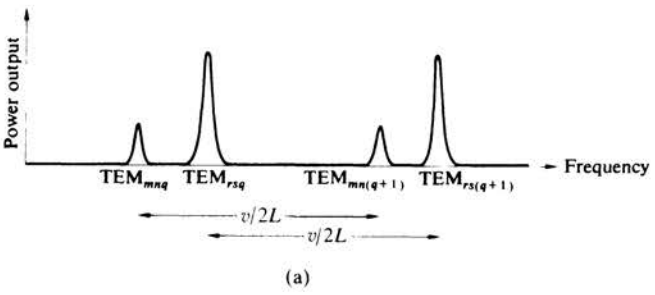
$$m = \frac{L}{\lambda/2}$$

and

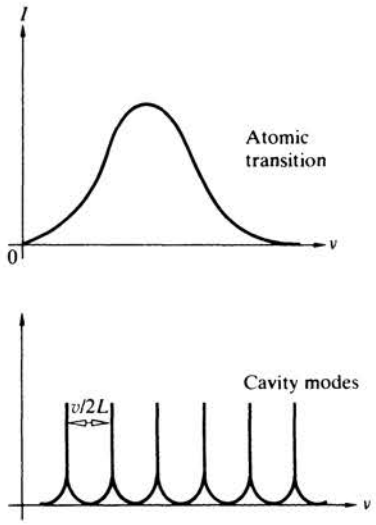
$$\nu_m = \frac{mv}{2L} \tag{13.15}$$

There are therefore an infinite number of possible oscillatory **longitudinal cavity modes**, each with a distinctive frequency ν_m. Consecutive modes are separated by a constant difference,

$$\nu_{m+1} - \nu_m = \Delta\nu = \frac{v}{2L} \tag{13.16}$$



(a)



(b)



(c)

Figure 13.8 Laser modes: (*a*) illustrates the nomenclature; (*b*) compares the broad atomic emission with the narrow cavity modes; (c) depicts three operation configurations for a c-w gas laser, first showing several longitudinal modes under a roughly Gaussian envelope, then several longitudinal and transverse modes, and finally a single longitudinal mode.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

which is the free spectral range of the etalon [Eq. (9.79)] and, incidentally, the inverse of the round-trip time. For a gas laser 1 m long, $\Delta\nu \approx 150$ MHz.

The resonant modes of the cavity are considerably narrower in frequency than the bandwidth of the normal spontaneous atomic transition. These modes, whether the device is constructed so that there is one or more, will be the ones that are sustained in the cavity, and hence the emerging beam is restricted to a region close to those frequencies (Fig. 13.8). In other words, the radiative transition makes available a relatively broad range of frequencies out of which the cavity will select and amplify only certain narrow bands and, if desired, even only one such band. This is the origin of the laser's extreme quasimonochromaticity. Thus, while the bandwidth of the ruby transition to the ground state is roughly a rather broad 0.53 nm (330 GHz)—because of interactions of the chromium ions with the lattice—the corresponding laser cavity bandwidth, the frequency spread of the radiation of a single resonant mode, is a much narrower 0.000 05 nm (30 MHz). This situation is depicted in Fig. 13.8*b*, which shows a typical transition lineshape and a series of corresponding cavity spikes—in this case each is separated by $v/2L$, and each is 30 MHz wide.

A possible way to generate only a single mode in the cavity would be to have the mode separation, as given by Eq. (13.16), exceed the transition bandwidth. Then only one mode would fit within the range of available frequencies provided by the transition. For a ruby laser (with an index of refraction of 1.76) a cavity length of a few centimeters will easily ensure single longitudinal mode operation. The drawback of this particular approach is that it limits the length of the active region contributing energy to the beam and so limits the output power of the laser.

In addition to the longitudinal or axial modes of oscillation, which correspond to standing waves set up along the cavity or z-axis, **transverse modes** can be sustained as well. Since the fields are very nearly normal to z, these are known as TEM_{mn} modes (transverse electric and magnetic). The m and n subscripts are the integer number of transverse nodal lines in the x- and y-directions across the emerging beam. That is to say, the beam is segmented in its cross section into one or more regions. Each such array is associated with a given TEM mode, as shown in Figs. 13.9 and 13.10. The lowest order or TEM_{00} transverse mode is perhaps the most widely used, and this for several compelling reasons: the flux density is ideally

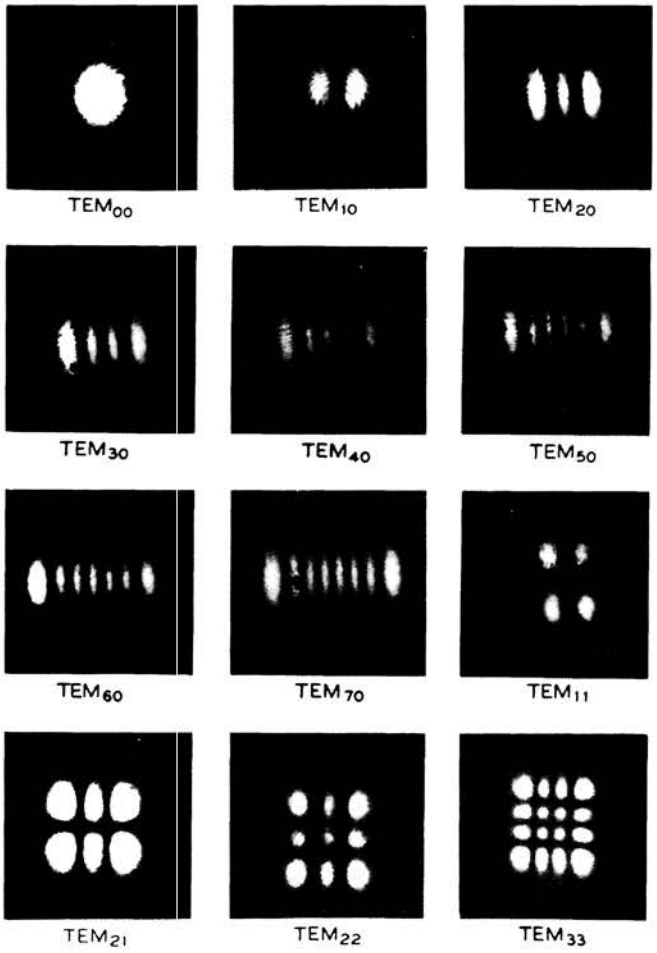


Figure 13.9 Mode patterns (without the faint interference fringes this is what the beam looks like in cross section). (Photos courtesy Bell Telephone Laboratories.)

Gaussian over the beam's cross section (Fig. 13.11); there are no phase shifts in the electric field across the beam, as there are in other modes, and so it is completely spatially coherent; the beam's angular divergence is the smallest; and it can be focused down to the smallest-sized spot. Note that the amplitude in this mode is actually not constant over the wavefront, and it is consequently an inhomogeneous wave.

A complete specification of each mode has the form TEM_{mnq}, where q is the longitudinal mode number. For each transverse mode (m, n) there can be many longitudinal modes

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

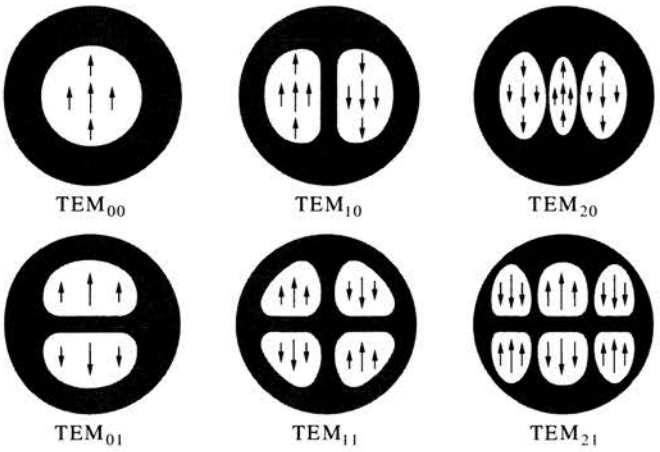


Figure 13.10 Mode configurations (rectangular symmetry). Circularly symmetric modes are also observable, but any slight asymmetry (such as Brewster windows) destroys them.

(i.e., values of q). Often, however, it's unnecessary to work with a particular longitudinal mode, and the q subscript is usually simply dropped.*

Several additional cavity arrangements are of considerably more practical significance than is the original plane-parallel setup (Fig. 13.12). For example, if the planar mirrors are replaced by identical concave spherical mirrors separated by a distance very nearly equal to their radius of curvature, we have the *confocal* resonator. Thus the focal points are almost coincident on the axis midway between mirrors—ergo the name confocal.

If one of the spherical mirrors is made planar, the cavity is termed a *hemispherical* or *hemiconcentric* resonator. Both of these configurations are considerably easier to align than is the plane-parallel form. Laser cavities are either *stable* or *unstable* to the degree that the beam tends to retrace itself and so remain relatively close to the optical axis (Fig. 13.13). A beam in an unstable cavity will "walk out," going farther from the axis on each reflection until it quickly leaves the cavity altogether. By contrast, in a stable configuration (with mirrors that are, say, 100% and 98% reflective) the beam might traverse the resonator 50 times or more. Unstable resonators are commonly used in high-power lasers, where the fact that the beam traces across a wide region of the active medium enhances the amplification and allows for more energy to be extracted. This approach will be especially useful for media (like carbon dioxide or argon) wherein the beam gains a good deal of energy on each sweep of the cavity. The needed number of sweeps is determined by the so-called *small-signal gain* of the active medium. The actual selection of a resonator configuration is governed by the specific requirements of the system—there is no universally best arrangement.

The decay of energy in a cavity is expressed in terms of the Q or **quality factor** of the resonator. The origin of the expression dates back to the early days of radio engineering, when it was used to describe the performance of an oscillating (tuning) circuit. A high-Q, low-loss circuit meant a narrow bandpass and a sharply tuned radio. If an optical cavity is somehow disrupted, as for example by the displacement or removal of one of the mirrors, the laser action generally ceases. When this is done deliberately in order to delay the onset of oscillation in the laser cavity, it's known as *Q-spoiling* or **Q-switching**. The power output of a laser is self-limited in the sense that the population inversion is continuously depleted through stimulated emission by the radiation field within the cavity. However, if oscillation is prevented, the number of atoms pumped into the (long-lived) metastable state can be considerably increased, thereby creating a very extensive population inversion. When the cavity is switched on at the proper moment, a tremendous-

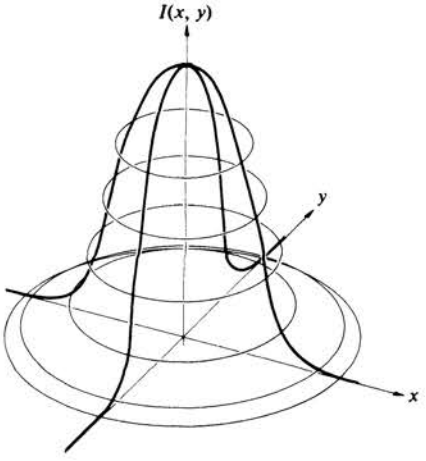


Figure 13.11 Gaussian irradiance distribution.

*Take a look at R. A. Phillips and R. D. Gehrz, "Laser Mode Structure Experiments for Undergraduate Laboratories," *Am. J. Phys.* **38**, 429 (1970).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

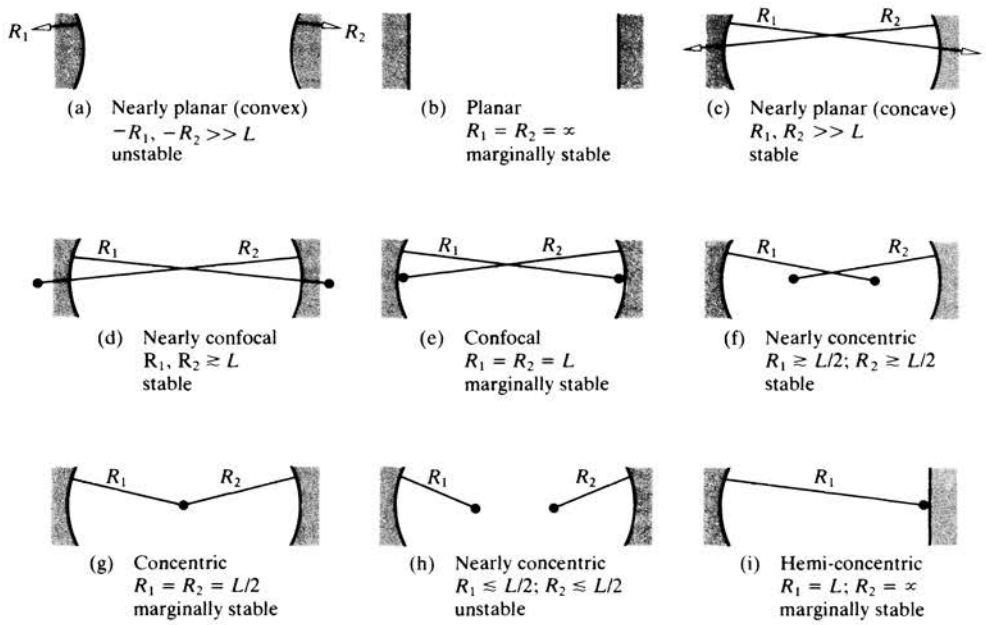


Figure 13.12 Laser cavity configurations. (Adapted from O'Shea, Callen, and Rhodes, *An Introduction to Lasers and Their Applications.*)

ly powerful *giant pulse* (perhaps up to several hundred megawatts) will emerge as the atoms drop down to the lower state almost in unison. A great many *Q-switching* arrangements utilizing various control schemes, for example, bleachable absorbers that become transparent under illumination, rotating prisms and mirrors, mechanical choppers, ultrasonic cells, or electro-optic shutters such as Kerr or Pockels cells, have all been used.

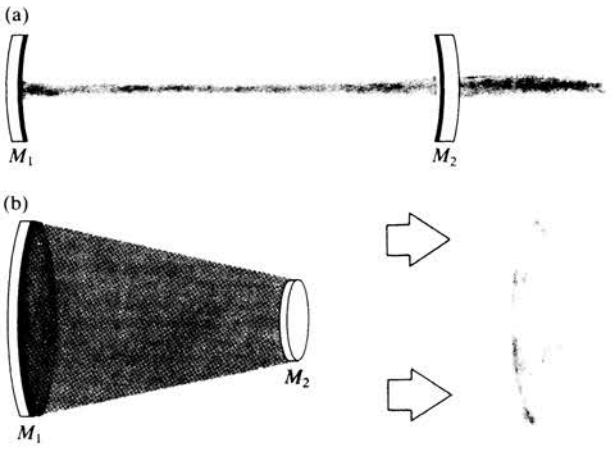


Figure 13.13 Stable and unstable laser resonators. (Adapted from O'Shea, Callen, and Rhodes, *An Introduction to Lasers and Their Applications.*)

Gaussian Laserbeams

The TEM_{00} mode that develops within a resonator has a Gaussian profile (Fig. 13.11); that is, the strength of the beam-like wave falls off transversely following a bell-shaped curve that's symmetrical around the central axis (Fig. 13.14*a*). Recall (p. 12) that a Gaussian is a negative exponential that's a function of the square of the variable, in this case, the distance (r) measured, in a transverse plane, from the central axis of propagation (z). Because the beam trails off radially it's useful to put an arbitrary boundary to its width. Accordingly, let $r = w$ be the *beam half-width*, the distance at which the electric field of the beam drops from its maximum axial value of E_0 to E_0/e or $37\%E_0$. At $r = w$ the beam's irradiance, which depends on the square of the amplitude, is then I_0/e^2 which is only $14\%I_0$. Most of the energy of the beam resides within this imaginary cylinder of radius w, where (Fig. 13.14*b*)

$$I = I_0 e^{-2r^2/w^2}$$

and $I = I_0 e^{-2}$, as it's supposed to, at $r = w$.

As can be seen in Fig. 13.13*a*, when curved mirrors form the laser cavity there is a tendency to "focus" the beam, giving it a minimum cross section or *waist* of radius w_0. Under such circumstances, the external divergence of the laserbeam is essentially a continuation of the divergence out from this waist (Fig. 13.15). In general, there will be a beam waist somewhere

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

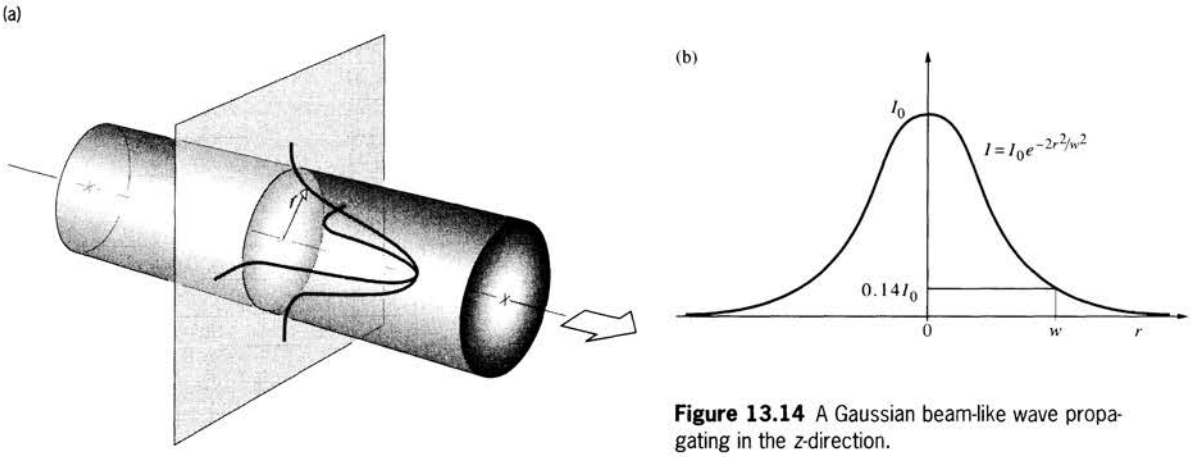


Figure 13.14 A Gaussian beam-like wave propagating in the z-direction.

between the mirrors of a laser resonator; its exact location depends on the specific design. For example, a confocal resonator (Fig. 13.12) has a waist halfway between the mirrors.

A more complete analysis of EM-waves in the cavity, setting $z = 0$ at the beam waist, yields the expression

$$w(z) = w_0\left[1 + \left(\frac{\lambda z}{\pi w_0^2}\right)^2\right]^{1/2} \tag{13.17}$$

for the half-width at any location z. The shape of the beam as specified by this expression for $w(z)$ is a hyperbola of revolution about the z-axis. A practical measure of the divergence of the beam is *the distance over which its cross-sectional area doubles*, or equivalently, the value of z for which $w(z) = \sqrt{2}w_0$. This special distance, z_R, is known as the **Rayleigh range**, and it follows from the above equation for $w(z)$ that

$$z_R = \frac{\pi {w_0}^2}{\lambda} \tag{13.18}$$

The smaller the waist (or equivalently, the smaller the minimum cross-sectional area), the smaller the Rayleigh range and the faster the beam diverges. At large distances from the waist ($z >> z_R$) the *full-angular width* of the beam (Θ, in radians) approaches $2w(z)/z$. In other words, as the line of length z, rotates through the angle Θ, its endpoint sweeps out a distance of $\approx 2w(z)$. Thus, when z is large and w_0 is small, the second term in the expression for $w(z)$ is much greater than 1 and

$$w(z) \approx w_0\left[\left(\frac{\lambda z}{\pi w_0^2}\right)^2\right]^{1/2} \approx \frac{\lambda z}{\pi w_0}$$

Since $\Theta \rightarrow 2w(z)/z$,

$$\Theta = \frac{2\lambda}{\pi w_0} = 0.637\,\frac{\lambda}{w_0}$$

Again, ***the smaller w_0 is, the larger will be Θ***, the beam divergence. In part, that's why people used to use megaphones—waves emerging from a larger aperture diverge less.

While two plane mirrors forming a laser cavity will produce a beam that is aperture limited via diffraction, this will

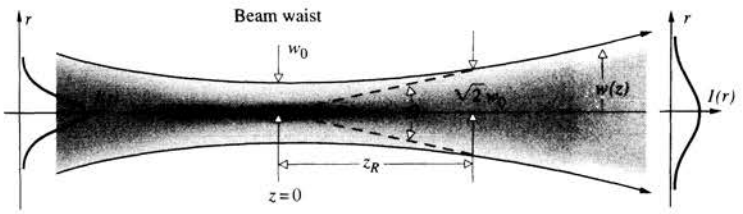


Figure 13.15 The spreading of a Gaussian beam.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

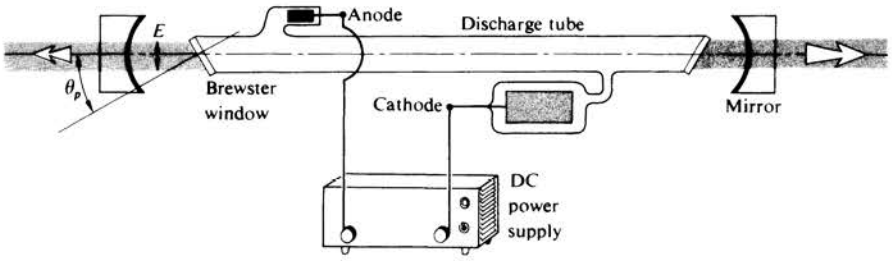


Figure 13.16 A simple, early He–Ne laser configuration.

not now be the case. Recall Eq. (10.58), which describes the radius of the Airy disk, and divide both sides by f to get the half-angular width of the diffracted circular beam of diameter D. Doubling this yields Θ, the *full-angular width* or **divergence of an aperture-limited laserbeam**:

$$\Theta \approx 2.44\lambda/D$$

By comparison, far from the region of minimum cross section, the full-angular width of a waisted laserbeam is

$$\Theta \approx 1.27\lambda/D_0 \qquad (13.19)$$

where $D_0 = 2w_0$, the beam-waist diameter, can be calculated from the particular cavity configuration.

The Helium–Neon Laser

Maiman's announcement of the first operative laser came at a New York news conference on July 7, 1960.* By February of 1961 Ali Javan and his associates W. R. Bennett, Jr., and D. R. Herriott had reported the successful operation of a *continuous-wave* (c-w) helium–neon, gas laser at 1152.3 nm. The He–Ne laser (Fig. 13.16) is still among the most popular devices of its kind, most often providing a few milliwatts of continuous power in the visible (632.8 nm). Its appeal arises primarily because it's easy to construct, relatively inexpensive, and fairly reliable and in most cases can be operated by a flick of a single switch. Pumping is usually accomplished by electrical discharge (via either dc, ac, or electrodeless rf excitation). Free electrons and ions are accelerated by an applied field and, as a result of collisions, cause further ionization and excitation of the gaseous medium (typically, a mixture of about 0.8 torr of He and about 0.1 torr of Ne). Many helium atoms, after dropping down from several upper levels, accumulate in the long-lived 2^1S- and 2^3S-states. These are metastable states (Fig. 13.17) from which there are no allowed radiative transitions. The excited He atoms inelastically collide with and transfer energy to ground-state Ne atoms, raising them in turn to the 5s- and 4s-states. These are the upper laser levels, and there then exists a population inversion with respect to the lower 4p- and 3p-states. Transitions between the 5s- and 4s-states are forbidden. Spontaneous photons initiate stimulated emission, and the chain reaction begins. The dominant laser transitions

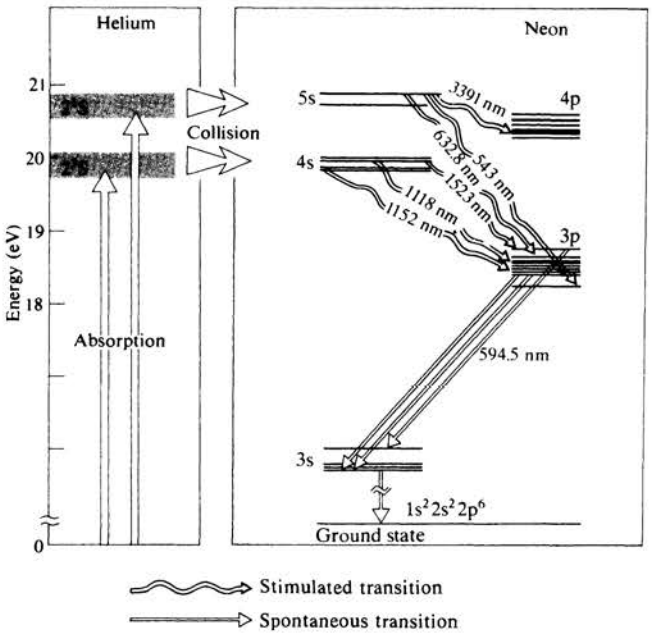


Figure 13.17 He–Ne laser energy levels.

*His initial paper, which would have made his findings known in a more traditional fashion, was rejected for publication by the editors of *Physical Review Letters*—this to their everlasting chagrin.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

correspond to 1152.3 nm and 3391.2 nm in the infrared and, of course, the ever-popular 632.8 nm in the visible (bright red). The *p*-states drain off into the 3s-state, thus themselves remaining uncrowded and thereby continuously sustaining the inversion. The 3s-level is metastable, so that 3s-atoms return to the ground state after losing energy to the walls of the enclosure. This is why the plasma tube's diameter inversely affects the gain and is, accordingly, a significant design parameter. In contrast to the ruby, where the laser transition is down to the ground state, stimulated emission in the He–Ne laser occurs between two upper levels. The significance of this, for example, is that since the 3p-state is ordinarily only sparsely occupied, a population inversion is very easily obtained, and this without having to half empty the ground state.

Return to Fig. 13.16, which pictures the relevant features of a basic early He–Ne laser. The mirrors are coated with a multilayered dielectric film having a reflectance of over 99%. The laser output is made linearly polarized by the inclusion of Brewster end windows (i.e., plates tilted at the polarization angle) terminating the discharge tube. If these end faces were instead normal to the axis, reflection losses (4% at each interface) would become unbearable. By tilting them at the polarization angle, the windows presumably have 100% transmission for light whose electric-field component is parallel to the plane-of-incidence (the plane of the drawing). This polarization state rapidly becomes dominant, since the normal component is partially reflected off-axis at each transit of the windows. Linearly polarized light in the plane-of-incidence soon becomes the preponderant stimulating mechanism in the cavity, to the ultimate exclusion of the orthogonal polarization.*

Epoxying the windows to the ends of the laser tube and mounting the mirrors externally was a typical though dreadful approach used commercially until the mid-1970s. Inevitably, the epoxy leaked, allowing water vapor in and helium out. Today, such lasers are *hard sealed*; the glass is bonded directly to metal (Kovar) mounts, which support the mirrors within the tube. The mirrors (one of which is generally $\approx 100\%$ reflective) have modern resistive coatings so they can tolerate the discharge environments within the tube. Operating lifetimes of 20 000 hours and more are now the rule (up from only a few hundred hours in the 1960s). Brewster windows are usually optional, and most commercial He–Ne lasers generate more or less "unpolarized" beams. The typical mass-produced He–Ne laser (with an output of from 0.5 mW to 5 mW) operates in the TEM_{00} mode, has a coherence length of around 25 cm, a beam diameter of approximately 1 mm, and a low overall efficiency of only 0.01% to about 0.1%. Although there are infrared He–Ne lasers, and even a new green (543.5 nm) He–Ne laser, the bright red 632.8-nm version remains the most popular.

A Survey of Laser Developments

Laser technology is so dynamic a field that what was a laboratory breakthrough a year or two ago may be a commonplace off-the-shelf item today. The whirlwind will certainly not pause to allow descriptive terms like "the smallest," "the largest," "the most powerful," and so on to be applicable for very long. With this in mind, we briefly survey the existing scene without trying to anticipate the wonders that will surely come after this type is set. Laserbeams have already been bounced off the Moon; they have spot welded detached retinas, generated fusion neutrons, stimulated seed growth, served as communications links, read CD discs, guided milling machines, missiles, ships, and grating engines, carried color television pictures, drilled holes in diamonds, levitated tiny objects,* and intrigued countless among the curious.

Along with ruby there are a great many other **solid-state lasers** whose outputs range in wavelength from roughly 170 nm to 3900 nm. For example, the trivalent rare earths Nd^{3+}, Ho^{3+}, Gd^{3+}, Tm^{3+}, Er^{3+}, Pr^{3+}, and Eu^{3+} undergo laser action in a host of hosts, such as $CaWO_4$, Y_2O_3, $SrMoO_4$, LaF_3, yttrium aluminum garnet (YAG for short), and glass, to name only a few. Of these, neodymium-doped glass and neodymium-doped YAG are of particular importance. Both constitute high-powered laser media operating at approximately 1060 nm. Nd: YAG lasers generating in excess of a kilowatt of continuous power have been constructed. Tremendous power outputs in pulsed systems have been obtained by operating several lasers in tandem. The first laser in the train serves as a

*Half of the output power of the laser is *not* lost in reflections at the Brewster windows when the transverse $\mathcal{P}$-state light is scattered. Energy simply isn't continuously channeled into that polarization component by the cavity. If it's reflected out of the plasma tube, it's not present to stimulate further emission.

*See M. Lubin and A. Fraas, "Fusion by Laser," *Sci. Am.* 224, **21** (June 1971); R. S. Craxton, R. L. McCrory, and J. M. Soures, "Progress in Laser Fusion," *Sci. Am.* 255, **69** (August 1986); and A. Ashkin, "The Pressure of Laser Light," *Sci. Am.* 226, **63** (February 1972).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

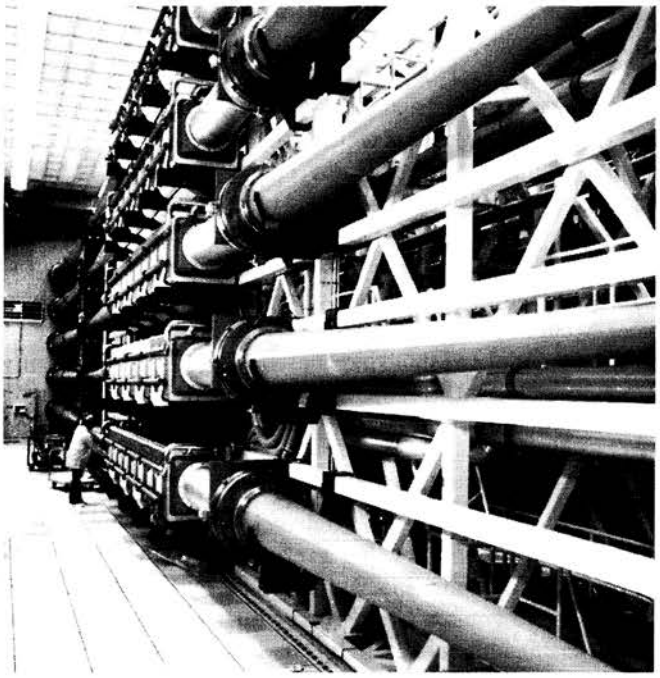
Q-switched oscillator that fires into the next stage, which functions as an amplifier; and there may be one or more such amplifiers in the system. By reducing the feedback of the cavity, a laser will no longer be self-oscillatory, but it will amplify an incident wave that has triggered stimulated emission. Thus the amplifier is, in effect, an active medium, which is pumped, but for which the end faces are only partially reflecting or even nonreflecting. Ruby systems of this kind, delivering a few GW (gigawatts, i.e., 10^9 W) in the form of pulses lasting several nanoseconds, are available commercially.

On December 19, 1984, the largest laser then in existence, the Nova at the Lawrence Livermore National Laboratory in California, fired all 10 of its beams at once for the first time, producing a warm-up shot of a mere 18 kJ of 350-nm radiation in a 1-ns pulse (see photo). This immense neodymium-doped glass laser was designed to focus up to 120 TW onto a fusion pellet—that's roughly 500 times more power than all the electrical generating stations in the United States—albeit only for about 10^{-9} s. In the late 1900s, the last years of its operation, using just one beamline of the Nova, LLNL researchers were able to produce 1.25-PW pulses, each lasting 490 fs and carrying 580 J.

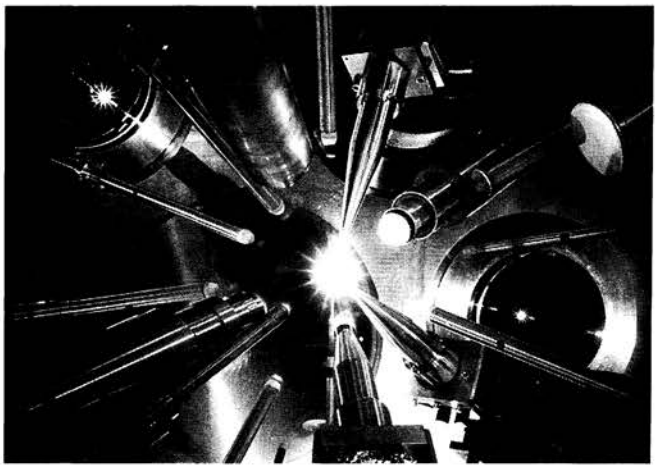
Nova's successor, which came on line in 1980 (with 24 solid state lasers), is housed in the Laboratory for Laser Energetics at the (LLE) University of Rochester. At present, LLE operates the 30 to 45 kJ Omega laser and is the world's premier laser fusion research facility. Upgraded in 1995, Omega is a 60-beam neodymium-doped phosphate glass laser that can concentrate 60×10^{12} W of radiant power onto a pinhead sized target. To accomplish that feat, the initial laser output is split repeatedly and each beam is subsequently amplified using Nd:glass discs and rods. Just before reaching the target the several beams are frequency-tripled to 351 nm using KDP crystals (p. 641). Supplying the needs of a variety of researchers, Omega is being operated at its maximum rate of one shoot per hour.

Still under construction, Omega's successor will be housed in the Department of Energy's immense National Ignition Facility. Its 192 beams will deliver 50 times more energy to its targets than does Omega.

A large group of **gas lasers** operate across the spectrum from the far IR to the UV (1 mm to 150 nm). Primary among these are helium–neon, argon, and krypton, as well as several molecular gas systems, such as carbon dioxide, hydrogen fluoride, and molecular nitrogen (N_2). Argon lases mainly in the green, blue-green, and violet (predominantly at 488.0 and 514.5 nm) in either pulsed or continuous operation. Although



The Nova laser. (Photo courtesy Lawrence Livermore National Laboratory.)



Inside the target chamber of the LLE laser-fusion device. Fusion reactions take place in a tiny target-sphere filled with deuterium/tritium and irradiated by the 30-kJ Omega laser. (Photo courtesy University of Rochester's Laboratory for Laser Energetics. Eugene Kowaluk, Image specialist.)

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

its output is usually several watts c-w, it has gone as high as 150 W c-w. The argon ion laser is similar in some respects to the He–Ne laser, although it evidently differs in its usually greater power, shorter wavelength, broader linewidth, and higher price. All of the noble gases (He, Ne, Ar, Kr, Xe) have been made to lase individually, as have the gaseous ions of many other elements, but the former grouping has been studied most extensively.

The CO_2 molecule, which lases between vibrational modes, emits in the IR at 10.6 μm, with typical c-w power levels of from a few watts to several kilowatts. Its efficiency can be an unusually high 15% when aided by additions of N_2 and He. While it once took a discharge tube nearly 200 m long to generate 10 kW c-w, considerably smaller "table models" are now available commercially. For a while in the 1970s, the record output belonged to an experimental gas-dynamic laser utilizing thermal pumping on a mixture of CO_2, N_2, and H_2O to generate 60 kW c-w at 10.6 μm in multimode operation.

The pulsed nitrogen laser operates at 337.1 nm in the UV, as does the c-w helium–cadmium laser. A number of metal vapors (e.g., Zn, Hg, Sn, Pb) have displayed laser transitions in the visible, but problems such as maintaining uniformity of the vapor in the discharge region have handicapped their exploitation. The He–Cd laser emits at 325.0 nm and 441.6 nm. These are transitions of the cadmium ion arising after excitation resulting from collisions with metastable helium atoms.

The **semiconductor laser**—alternatively known as the junction or diode laser—was invented in 1962, soon after the development of the light-emitting diode (LED). Today it serves a central role in electro-optics, primarily because of its spectral purity, high efficiency ($\approx$100%), ruggedness, ability to be modulated at extremely rapid rates, long lifetimes, and moderate power (as much as 200 mW) despite its pinhead size. Junction lasers have already been used in the millions in fiberoptic communications, laser disk audio systems, and so forth.

The first such lasers were made of one material, gallium arsenide, appropriately doped to form a p–n junction. The associated high lasing threshold of these so-called homostructures limited them to pulsed mode operation and cryogenic temperatures; otherwise the heat developed in their small structures would destroy them. The first tunable lead–salt diode laser was developed in 1964, but it was not until almost a dozen years later that it became commercially available. It operates at liquid nitrogen temperatures, which is certainly inconvenient, but it can scan from 2 μm to 30 μm.

Later advances have since allowed a reduction in the threshold and resulted in the advent of the continuous-wave (c-w), room temperature diode laser. Transitions occur between the conduction and valence bands, and stimulated emission results in the immediate vicinity of the p–n junction (Fig. 13.18). Quite generally, as a current flows in the forward direction through a semiconductor diode, electrons from the n-layer conduction band will recombine with p-layer holes, thereupon emitting energy in the form of photons. This radiative process, which competes for energy with the existing absorption mechanisms (such as phonon production), comes to predominate when the recombination layer is small and the current is large. To make the system lase, the light emitted

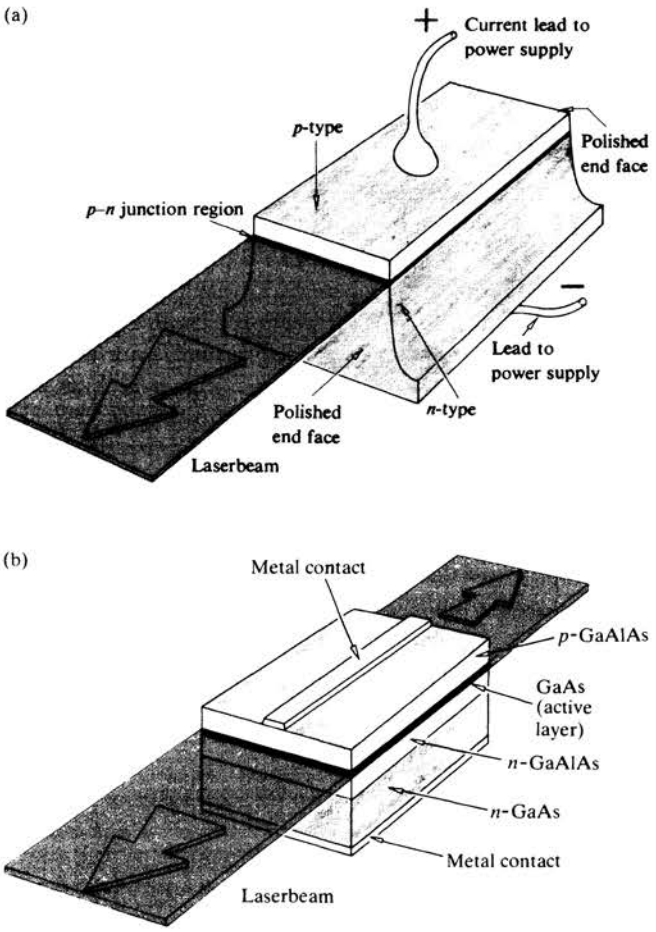


Figure 13.18 (*a*) An early GaAs p–n junction laser. (*b*) A more modern diode laser.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

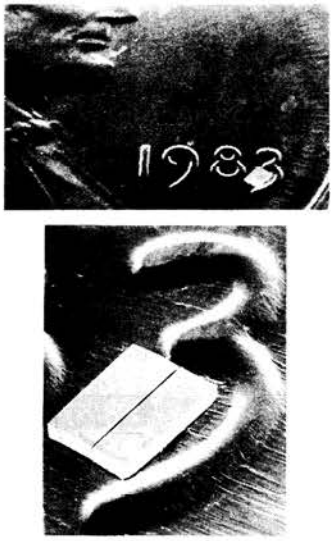
from the diode is retained within a resonant cavity, and that's usually accomplished by simply polishing the end faces perpendicular to the junction channel.

Nowadays semiconductor lasers are created to meet specific needs, and there are many designs producing wavelengths ranging from around 700 nm to about 30 μm. The early 1970s saw the introduction of the c-w GaAs/GaAlAs laser. Operating at room temperature in the 750-nm to 900-nm region (depending on the relative amounts of aluminum and gallium), the tiny diode chip is usually about a sixteenth of a cubic centimeter in volume. Figure 13.18*b* shows a typical heterostructure (a device formed of different materials) diode laser of this kind. Here the beam emerges in two directions from the 0.2-μm thick active layer of GaAs. These little lasers usually produce upward of 20 mW of continuous wave power. To take advantage of the low-loss region ($\lambda \approx 1.3$ μm) in fiberoptic glass the GaInAsP/InP laser was devised in the mid-1970s with an output of 1.2 μm to 1.6 μm.

The cleaved–coupled-cavity laser is shown in the accompanying photo. In it the number of axial modes is controlled in order to produce very-narrow-bandwidth tunable radiation. Two cavities coupled together across a small gap restrict the radiation to the extremely narrow bandwidth that can be sustained in both resonant chambers.*

The first **liquid laser** was operated in January of 1963.† All of the early devices of this sort were exclusively *chelates* (i.e., metallo-organic compounds formed of a metal ion with organic radicals). That original liquid laser contained an alcohol solution of europium benzoylacetonate emitting at 613.1 nm. The discovery of laser action in nonchelate organic liquids was made in 1966. It came with the fortuitous lasing (at 755.5 nm) of a chloroaluminum phthalocyanine solution during a search for stimulated Raman emission in that substance.‡

A great many fluorescent dye solutions of such families as the fluoresceins, coumarins, and rhodamines have since been made to lase at frequencies from the IR into the UV. These have usually been pulsed, although c-w operation has been obtained. There are so many organic dyes that it would seem possible to build such a laser at any frequency in the visible. Moreover, these devices are distinctive in that they inherently can be tuned continuously over a range of wavelengths (of perhaps 70 nm or so, although a pulsed system tunable over 170 nm exists). Indeed, there are other arrangements that will vary the frequency of a primary laserbeam (i.e., the beam enters with one color and emerges with another, Section 13.4), but in the case of the dye laser, the primary beam itself is tuned internally. This is accomplished, for example, by changing the concentration or the length of the dye cell or by adjusting a diffraction grating reflector at the end of the cavity. Multicolor dye laser systems, which can easily be switched from one dye to another and thereby operate over a very broad frequency range, are available commercially.



The cleaved–coupled-cavity laser. (Photos courtesy of Bell Laboratories.)

A **chemical laser** is one that is pumped with energy released via a chemical reaction. The first of this kind was operated in 1964, but it was not until 1969 that a continuous-wave chemical laser was developed. One of the most promising of these is the deuterium fluoride–carbon dioxide (DF–CO_2) laser. It is self-sustaining in that it requires no external power source. In brief, the reaction $F_2 + D_2 \rightarrow 2DF$, which occurs on the mixing of these two fairly common gases, generates enough energy to pump a CO_2 laser.

There are solid-state, gaseous, liquid, and vapor (e.g., H_2O) lasers; there are semiconductor lasers, free electron (600 nm to

*See Y. Suematsu, "Advances in Semiconductor Lasers," *Phys. Today*, **32** (May 1985). For a discussion of heterostructure diode lasers, refer to M. B. Panish and I. Hayashi, "A New Class of Diode Lasers," *Sci. Am.* 225, **32** (July 1971).

†See Adam Heller, "Laser Action in Liquids," *Phys. Today* (November 1967), p. 35, for a more detailed account.

‡P. Sorokin, "Organic Lasers," *Sci. Amer.* 220, **30** (February 1969).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

3 mm) lasers, X-ray lasers, and lasers with very special properties, such as those that generate extremely short pulses, or those that have extraordinary frequency stability. These latter devices are very useful in the field of high-resolution spectroscopy, but there is a growing need for them in other research areas as well (e.g., in the interferometers used to attempt to detect gravity waves). In any event, these lasers must have precisely controlled cavity configurations despite the disturbing influences of temperature variations, vibrations, and even sound waves. To date, the record is held by a laser at the Joint Institute for Laboratory Astrophysics in Boulder, Colorado, which maintains a frequency stability (p. 315) of nearly one part in 10^{14}.

13.1.4 The Light Fantastic

Laserbeams differ somewhat from one type of laser to another; yet there are several remarkable features that are displayed, to varying degrees, by all of them. Quite apparent is the fact that most laserbeams are exceedingly directional, or if you will, highly collimated. One need only blow some smoke into the otherwise invisible, visible laserbeam to see (via scattering) a fantastic thread of light stretched across a room. A He–Ne beam in the TEM_{00} mode generally has a divergence of only about one minute of arc or less. Recall that in that mode the emission closely approximates a Gaussian irradiance distribution; that is, the flux density drops off from a maximum at the central axis of the beam and has no side lobes. The typical laserbeam is quite narrow, usually issuing at no more than a few millimeters in diameter. Since the beam resembles a truncated plane wave, it is of course highly *spatially coherent*. In fact, its directionality may be thought of as a manifestation of that coherence. Laserlight is quasimonochromatic, generally having an exceedingly narrow frequency bandwidth (p. 313). In other words, it is highly *temporally coherent*.

Another attribute is the large flux or *radiant power* that can be delivered in that narrow frequency band. As we've seen, the laser is distinctive in that it emits all its energy in the form of a narrow beam. In contrast, a 100-W incandescent lightbulb may pour out considerably more radiant energy in toto than a lower-power c-w laser, but the emission is incoherent, spread over a large solid angle, and it has a broad bandwidth as well. A good lens* can totally intercept a laserbeam and focus essentially all of its energy into a minute spot (whose diameter varies directly with λ and the focal length and inversely with the beam diameter). Spot diameters of just a few thousandths of an inch can readily be attained with lenses that have a conveniently short focal length. And a spot diameter of a few hundred-millionths of an inch is possible in principle. Thus flux densities can readily be generated in a focused laserbeam of over 10^{17} W/cm^2, in contrast to, say, an oxyacetylene flame having roughly 10^3 W/cm^2. To get a better feel for these power levels, note that a focused CO_2 laserbeam of a few kilowatts c-w can burn a hole through a quarter-inch stainless steel plate in about 10 seconds. By comparison, a pinhole and filter positioned in front of an ordinary source will certainly produce spatially and temporally coherent light, but only at a minute fraction of the total power output.

Femtosecond Optical Pulses

The advent of the mode-locked dye laser in the early part of the 1970s gave a great boost to the efforts then being made at generating extremely short pulses of light.* Indeed, by 1974 subpicosecond (1 ps = 10^{-12} s) optical pulses were already being produced, although the remainder of the decade saw little significant progress. In 1981 two separate advances resulted in the creation of femtosecond laser pulses (i.e., <0.1 ps or <100 fs)—a group at Bell Labs developed a colliding-pulse ring dye laser, and a team at IBM devised a new pulse-compression scheme.

Above and beyond the implications in the practical domain of electro-optical communications, these accomplishments have firmly established a new field of research known as *ultrafast phenomena*. The most effective way to study the progression of a process that occurs exceedingly rapidly (e.g., carrier dynamics in serniconductors, fluorescence, photochemical biological processes, and molecular configuration changes) is to examine it on a time scale that is comparatively short with respect to what's happening. Pulses lasting $\approx$10 fs allow an entirely new access into previously obscure areas in the study of matter.

Pulses lasting a mere 8 fs (10^{-15} s), which corresponds to wavetrains only about 4 wavelengths of red light in length, can now be produced routinely. One of the techniques that makes these femtosecond wavegroups possible is based on an idea

*Spherical aberration is usually the main problem, since laserbeams are, as a rule, both quasimonochromatic and incident along the axis of the lens.

*See Chandrashekhar Joshi and Paul Corkum, "Interactions of Ultra-Intense Laser Light with Matter," *Phys. Today* **36** (January 1995).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

used in radar work in the 1950s called *pulse compression*. Here an initial laser pulse has its frequency spectrum broadened, thereby allowing the inverse or temporal pulse width to be shortened—remember that $\Delta\nu$ and Δt are conjugate Fourier quantities [Eq. (7.63)]. The input pulse (several picoseconds long) is passed into a nonlinear dispersive medium, namely, a single-mode optical fiber. When the light intensity is high enough, the index of refraction has an appreciable nonlinear term (Section 13.4), and the carrier frequency of the pulse experiences a time-dependent shift. On traversing perhaps 30 m of fiber, the frequency of the pulse is drawn out or "chirped." That is, a spread occurs in the spectrum of the pulse, with the low frequencies leading and the high frequencies trailing. Next the spectrally broadened pulse is passed through another dispersive system (a delay line), such as a pair of diffraction gratings. By traveling different paths, the blue-shifted trailing edge of the pulse is made to catch up to the red-shifted leading edge, creating a time-compressed output pulse.

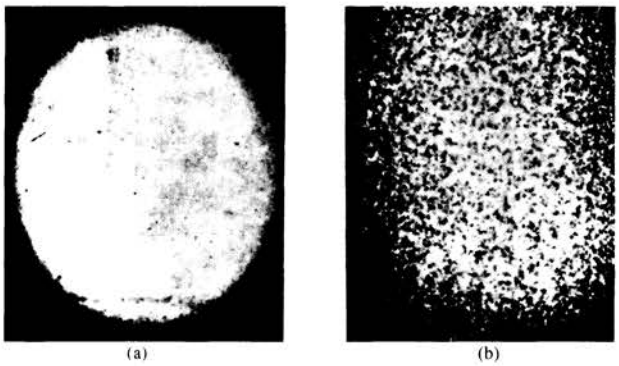
The Speckle Effect

A rather striking and easily observable manifestation of the spatial coherence of laserlight is its granular appearance on reflection from a diffuse surface. Using a He–Ne laser (632.8 nm), expand the beam a bit by passing it through a simple lens and project it onto a wall or a piece of paper. The illuminated disk appears speckled with bright and dark regions that sparkle and shimmer in a dazzling psychedelic dance. Squint and the grains grow in size; step toward the screen and they shrink; take off your eyeglasses and the pattern stays in perfect focus. In fact, if you are nearsighted, the diffraction fringes caused by dust on the lens blur out and disappear, but the speckles do not. Hold a pencil at varying distances from your eye so that the disk appears just above it. At each position, focus on the pencil; wherever you focus, the granular display is crystal clear. Indeed, look at the pattern through a telescope; as you adjust the scope from one extreme to the other, the ubiquitous granules remain perfectly distinct, even though the wall is completely blurred.

The spatially coherent light scattered from a diffuse surface fills the surrounding region with a *stationary* interference pattern (just as in the case of the wavefront-splitting arrangements of Section 9.3). At the surface the granules are exceedingly small, and they increase in size with distance. At any location in space, the resultant field is the superposition of many contributing scattered wavelets. These must have a constant relative phase determined by the optical path length from each scatterer to the point in question, if the interference pattern is to be sustained. The accompanying photo illustrates this point rather nicely. It shows a cement block illuminated in one case by laserlight and in the other by collimated light from a Hg arc lamp, both of about the same spatial coherence. While the laser's coherence length is much greater than the height of the surface features, the coherence length of the Hg light is not. In the former case, the speckles in the photograph are large, and they obscure the surface structure; in the latter, despite its spatial coherence, the speckle pattern is not observable in the photograph, and the surface features predominate. Because of the rough texture, the optical path length difference between two wavelets arriving at a point in space, scattered from different surface bumps, is generally greater than the coherence length of the mercury light. This means that the relative phases of the overlapping wavetrains change rapidly and randomly in time, washing out the large-scale interference pattern.

A real system of fringes is formed of the scattered waves that converge in front of the screen. The fringes can be viewed by intersecting the interference pattern with a sheet of paper at a convenient location. After forming the real image in space, the rays proceed to diverge, and any region of the image can therefore be viewed directly with the eye appropriately focused. In contrast, rays that initially diverge appear to the eye as if they had originated behind the scattering screen and thus form a virtual image.

It seems that as a result of chromatic aberration, normal and farsighted eyes tend to focus red light behind the screen. Contrarily, a nearsighted person observes the real field in



Speckle patterns. (*a*) A cement block illuminated by a mercury arc and (*b*) a He–Ne laser. [From B. J. Thompson, *J. Soc. Phot. Inst. Engr.* **4**, 7 (1965).]

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

front of the screen (regardless of wavelength). Thus if the viewer moves her head to the right, the pattern will move to the right in the first instance (where the focus is beyond the screen) and to the left in the second (focus in front). The pattern will follow the motion of your head, if you're viewing it very close to the surface. The same apparent parallax motion can be seen by looking through a window; outside objects will seem to move with your head, inside ones opposite to it. The brilliant, narrow-bandwidth, spatially coherent laserbeam is ideally suited for observing the granular effect, although other means are certainly possible.* In unfiltered sunlight the grains are minute, on the surface, and multicolored. The effect is easy to observe on a smooth, flat-black material (e.g., poster-painted paper), but you can see it on a fingernail or a worn coin as well.

Although it provides a marvelous demonstration, both aesthetically and pedagogically, the granular effect can be a real practical nuisance in coherently illuminated systems. For example, in holographic imagery the speckle pattern corresponds to troublesome background noise. Incidentally, very much the same kind of thing is observable when listening to a mobile radio where the signal strength fluctuates from one location to the next, depending on the environment and the resulting interference pattern.

The Spontaneous Raman Effect

It is possible that an excited atom will not return to its initial state after the emission of a photon. This kind of behavior had been observed and studied extensively by George Stokes prior to the advent of quantum theory. Since the atom drops down to an interim state, it emits a photon of lower energy than the incident primary photon, in what is usually referred to as a ***Stokes transition***. If the process takes place rapidly (roughly 10^{-7} s), it is called **fluorescence**, whereas if there is an appreciable delay (in some cases seconds, minutes, or even many hours), it is known as **phosphorescence**. Using ultraviolet quanta to generate a fluorescent emission of visible light has become an accepted occurrence in our everyday lives. Any number of commonplace materials (e.g., detergents, organic dyes, and tooth enamel) will emit characteristic visible photons so that they appear to glow under ultraviolet illumination; ergo the widespread use of the phenomenon for commercial display purposes and for "whitening" cloths.

If quasimonochromatic light is scattered from a substance, it will thereafter consist mainly of light of the same frequency. Yet it is possible to observe very weak additional components having higher and lower frequencies (side bands). Moreover, the difference between the side bands and the incident frequency ν_i is found to be characteristic of the material and therefore suggests an application to spectroscopy. The **Spontaneous Raman Effect**, as it is now called, was predicted in 1923 by Adolf Smekal and observed experimentally in 1928 by Sir Chandrasekhara Vankata Raman (1888–1970), then professor of physics at the University of Calcutta. The effect was difficult to put to actual use, because one needed strong sources (usually Hg discharges were used) and large samples. Often the ultraviolet from the source would further complicate matters by decomposing the specimen. And so it is not surprising that little sustained interest was aroused by the promising practical aspects of the Raman Effect. The situation was changed dramatically when the laser became a reality. ***Raman spectroscopy*** is now a unique and powerful analytical tool.

To appreciate how the phenomenon operates, let's review the germane features of molecular spectra. A molecule can absorb radiant energy in the far-infrared and microwave regions, converting it to rotational kinetic energy. Furthermore, it can absorb infrared photons (i.e., ones within a wavelength range from roughly 10^{-2} mm down to about 700 nm), transforming that energy into vibrational motion of the molecule. Finally, a molecule can absorb energy in the visible and ultraviolet regions through the mechanism of electron transitions, much like those of an atom. Suppose then that we have a molecule in some vibrational state, which, using quantum-mechanical notation, we call $|b\rangle$, as indicated diagrammatically in Fig. 13.19*a*. This need not necessarily be an excited state. An incident photon of energy $h\nu_i$ is absorbed, raising the system to some intermediate or virtual state, whereupon it immediately makes a Stokes transition, emitting a (scattered) photon of energy $h\nu_s < h\nu_i$. In conserving energy, the difference $h\nu_i - h\nu_s = h\nu_{cb}$ goes into exciting the molecule to a higher vibrational energy level $|c\rangle$. It is possible that electronic or rotational excitation results as well.

Alternatively, if the initial state is an excited one (just heat the sample), the molecule, after absorbing and emitting a photon, may drop back to an even lower state (Fig. 13.19*b*), thereby making an ***anti-Stokes transition***. In this instance $h\nu_s > h\nu_i$, which means that some vibrational energy of the molecule

*For further reading on this effect, see L. I. Goldfischer, *J. Opt. Soc. Am.* **55**, 247 (1965); D. C. Sinclair, *J. Opt. Soc. Am.* **55**, 575 (1965); J. D. Rigden and E. I. Gordon, *Proc. IRE* **50**, 2367 (1962); B. M. Oliver, *Proc. IEEE* **51**, 220 (1963).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

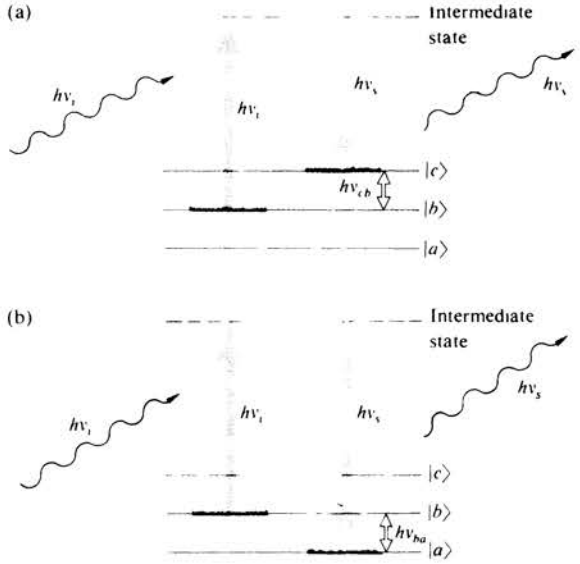


Figure 13.19 Spontaneous Raman Scattering.

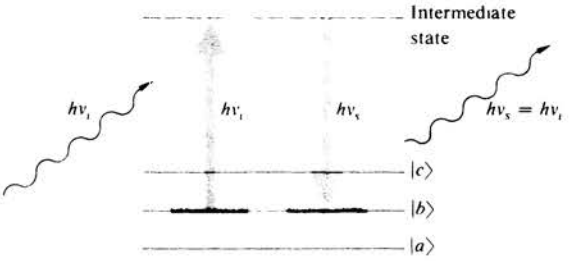


Figure 13.20 Rayleigh Scattering.

($h\nu_{ba} = h\nu_s - h\nu_i$) has been converted into radiant energy. In either case, the resulting differences between ν_s and ν_i correspond to specific energy-level differences for the substance under study and as such yield insights into its molecular structure. Figure 13.20, for comparison's sake, depicts Rayleigh Scattering where $\nu_s = \nu_i$.

The laser is an ideal source for spontaneous Raman Scattering. It is bright, quasimonochromatic, and available in a wide range of frequencies. Figure 13.21 illustrates a typical laser-Raman system. Complete research instruments of this sort are commercially available, including the laser (usually helium–neon, argon, or krypton), focusing lens systems, and photon-counting electronics. The double scanning monochromator provides the needed discrimination between ν_i and ν_s, since unshifted laserlight (ν_i) is scattered along with the

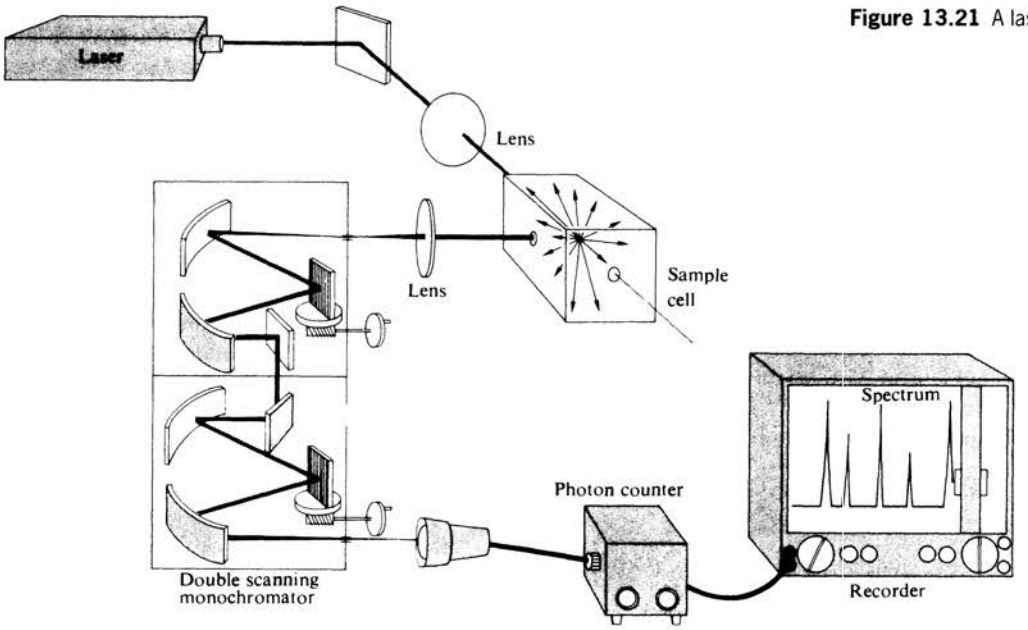


Figure 13.21 A laser–Raman system.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

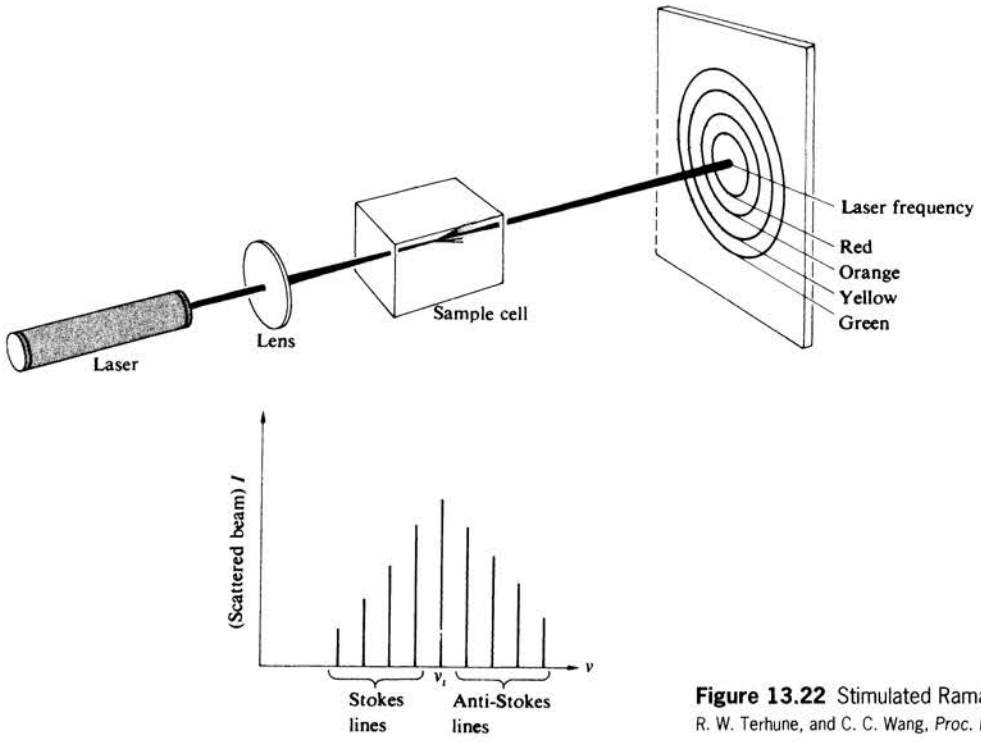


Figure 13.22 Stimulated Raman Scattering. [See R. W. Minck, R. W. Terhune, and C. C. Wang, *Proc. IEEE* **54**, 1357 (1966).]

Raman spectra (ν_s). Although Raman Scattering associated with molecular rotation was observed prior to the use of the laser, the increased sensitivity now available makes the process easier and allows even the effects of electron motion to be examined.

The Stimulated Raman Effect

In 1962 Eric J. Woodbury and Won K. Ng rather fortuitously discovered a remarkable related effect known as *Stimulated Raman Scattering*. They had been working with a million-watt pulsed ruby laser incorporating a nitrobenzene Kerr cell shutter (see Section 8.11.3). They found that about 10% of the incident energy at 694.3 nm was shifted in wavelength and appeared as a *coherent* scattered beam at 766.0 nm. It was subsequently determined that the corresponding frequency shift of about 40 THz was characteristic of one of the vibrational modes of the nitrobenzene molecule, as were other new frequencies also present in the scattered beam. Stimulated Raman Scattering can occur in solids, liquids, or dense gases under the influence of focused high-energy laser pulses (Fig. 13.22). The effect is schematically depicted in Fig. 13.23. Here two photon beams are simultaneously incident on a molecule, one corresponding to the laser frequency ν_i, the other having the scattered frequency ν_s. In the original setup, the scattered beam was reflected back and forth through the specimen, but the effect can occur without a resonator. The laserbeam loses a photon $h\nu_i$, while the scattered beam gains a photon $h\nu_s$ and

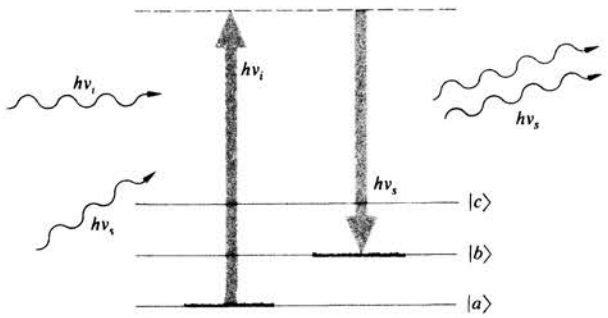


Figure 13.23 Energy-level diagram of Stimulated Raman Scattering.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

is subsequently *amplified*. The remaining energy ($h\nu_i - h\nu_s = h\nu_{bd}$) is transmitted to the sample. The chain reaction in which a large portion of the incident beam is converted into stimulated Raman light can only occur above a certain high-threshold flux density of the exciting laserbeam.

Stimulated Raman Scattering provides a whole new range of high flux-density coherent sources extending from the infrared to the ultraviolet. It should be mentioned that in principle each spontaneous scattering mechanism (e.g., Rayleigh and Brillouin Scattering) has its stimulated counterpart.*

13.2 Imagery — The Spatial Distribution of Optical Information

The manipulation of all sorts of data via optical techniques has already become a technological *fait accompli*. The literature since the 1960s reflects, in a diversity of areas, this far-reaching interest in the methodology of optical data processing. Practical applications have been made in the fields of television and photographic image enhancement, radar and sonar signal processing (phased and synthetic array antenna analysis), as well as in pattern recognition (e.g., aerial photointerpretation and fingerprint studies), to list only a very few.

Our concern here is to develop the nomenclature and some of the ideas necessary for an appreciation of this contemporary thrust in Optics.

13.2.1 Spatial Frequencies

In electrical processes one is most frequently concerned with signal variations in time, that is, the moment-by-moment alteration in voltage that might appear across a pair of terminals at some fixed location in space. By comparison, in Optics we are most often concerned with information spread across a region of space at a fixed location in time. For example, we

*For further reading on these subjects you might try the review-tutorial paper by Nicolaas Bloembergen, "The Stimulated Raman Effect," *Am. J. Phys.* **35**, 989 (1967). It contains a fairly good bibliography as well as a historical appendix. Many of the papers in *Lasers and Light* also deal with this material and are highly recommended reading.

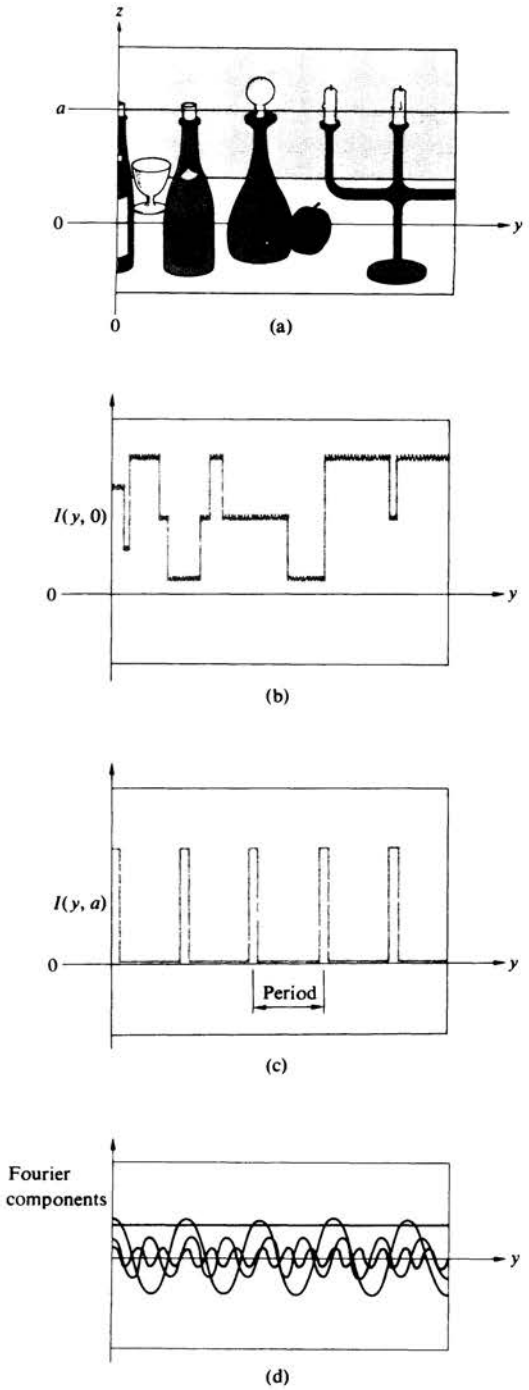


Figure 13.24 A two-dimensional irradiance distribution.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

can think of the scene depicted in Fig. 13.24*a* as a two-dimensional flux-density distribution. It might be an illuminated transparency, a television picture, or an image projected on a screen; in any event there is presumably some function $I(y, z)$, which assigns a value of I to each point in the picture. To simplify matters a bit, suppose we scan across the screen on a horizontal line ($z = 0$) and plot point-by-point variations in irradiance with distance, as in Fig. 13.24*b*. The function $I(y, 0)$ can be synthesized out of harmonic functions, using the techniques of Fourier analysis treated in Chapters 7 and 11. In this instance, the function is rather complicated, and it would take many terms to represent it adequately. Yet if the functional form of $I(y, 0)$ is known, the procedure is straightforward enough. Scanning across another line, for example, $z = a$, we get $I(y, a)$, which is drawn in Fig. 13.24*c* and which just happens to turn out to be a series of equally spaced square pulses. This function is one that was considered at length in Section 7.3, and a few of its constituent Fourier components are roughly sketched in Fig. 13.24*d*. If the peaks in (*c*) are separated, center to center, by say, 1-cm intervals, the spatial period equals 1 cm per cycle, and its reciprocal, which is the spatial frequency, equals 1 cycle per cm.

Quite generally, we can transform the information associated with any scan line into a series of sinusoidal functions of appropriate amplitude and spatial frequency. In the case of either of the simple sine- or square-wave targets of Fig. 13.25, each such horizontal scan line is identical, and the patterns are effectively one-dimensional. The spatial frequency spectrum of Fourier components needed to synthesize the square wave is shown in Fig. 7.32. On the other hand, $I(y, z)$ for the wine bottle candelabra scene is two-dimensional, and we have to think in terms of two-dimensional Fourier transforms (Sections 7.4.4 and 11.2.2). We might mention as well that, at least in principle, we could have recorded the amplitude of the electric field at each point of the scene and then performed a similar decomposition of that signal into its Fourier components.

Recall (Section 11.3.3) that the far-field or Fraunhofer diffraction pattern is essentially identical to the Fourier transform of the aperture function $\mathscr{A}(y, z)$. The aperture function is proportional to $\mathscr{E}_A(y, z)$, the source strength per unit area [Eq. (10.37)] over the input or object plane. In other words, if the field distribution on the object plane is given by $\mathscr{A}(y, z)$, its two-dimensional Fourier transform will appear as the field distribution $E(Y, Z)$ on a very distant screen. As in Figs. 7.43 and 10.17, we can introduce a lens (L_t) after the object in order to

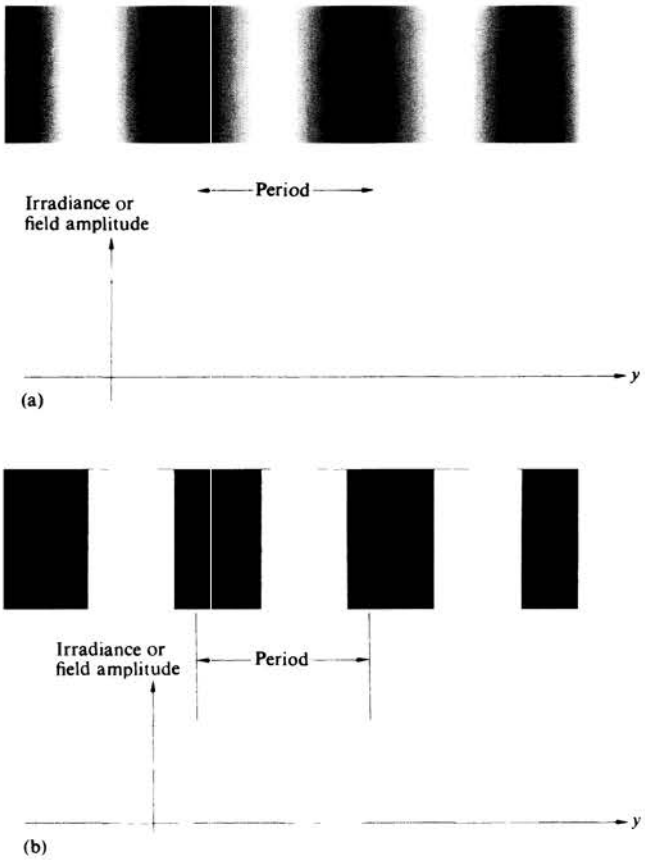


Figure 13.25 (*a*) Sine-wave target and (*b*) square-wave target.

shorten the distance to the image plane. That objective lens is commonly referred to as the ***transform lens***, since we can imagine it as if it were an *optical computer* capable of generating instant Fourier transforms. Now, suppose we illuminate a somewhat idealized transmission grating with a spatially coherent, quasimonochromatic wave, such as the plane wave emanating from a laser or a collimated, filtered Hg arc source (Fig. 13.26). In either case, the amplitude of the field is assumed to be fairly constant over the incident wavefront. The aperture function is then a periodic step function (Fig. 13.27); in other words, as we move from point to point on the object plane, the amplitude of the field is either zero or a constant. If a is the grating spacing, it is also the spatial period of the step function, and its reciprocal is the fundamental spatial frequency of the grating. The central spot ($m = 0$) in the diffraction

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

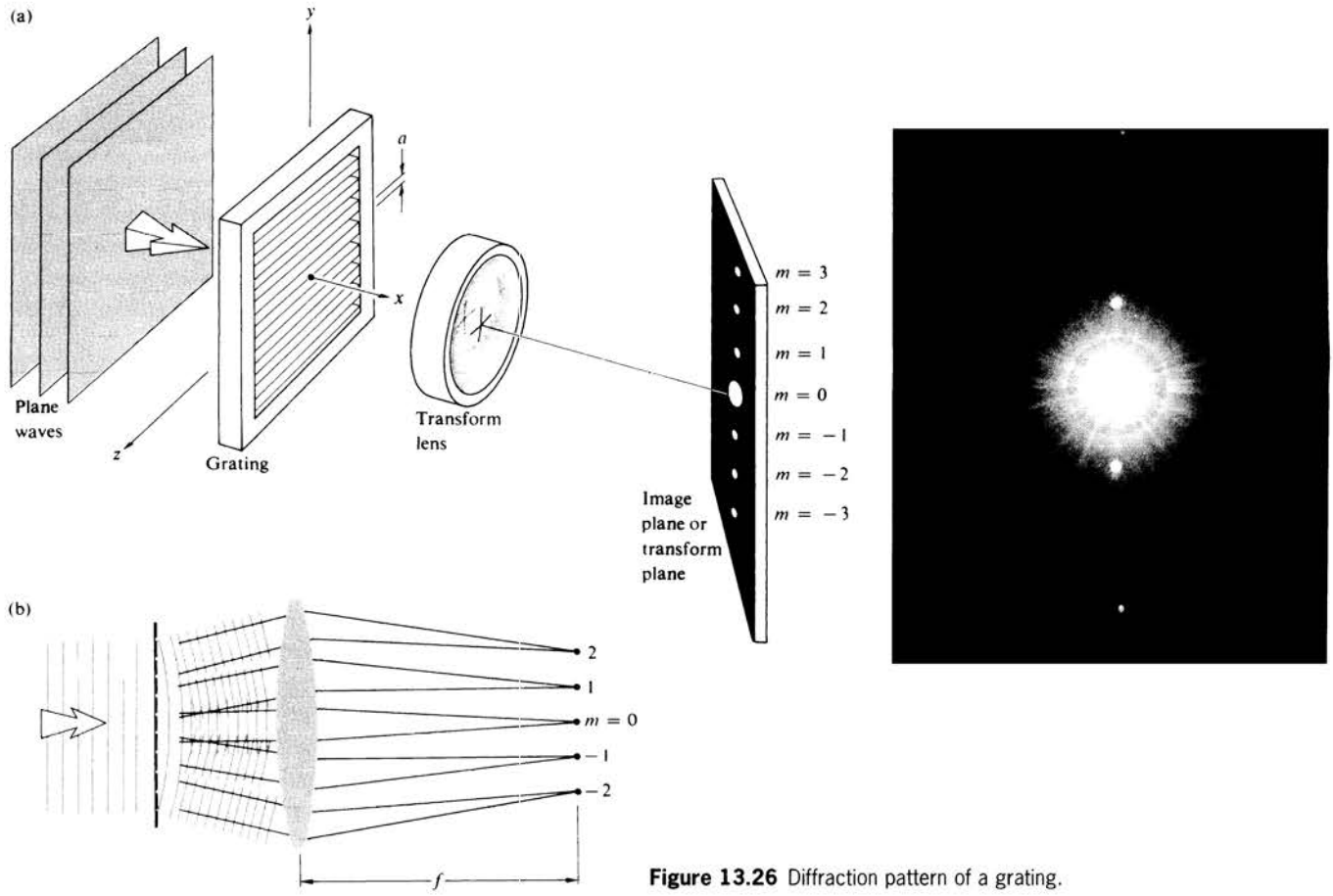


Figure 13.26 Diffraction pattern of a grating.

pattern is the dc term corresponding to a zero spatial frequency—it's the bias level that arises from the fact that the input $\mathcal{A}(y)$ is everywhere positive. This bias level can be shifted by constructing the step-function pattern on a uniform gray background. As the spots in the image (or in this case the transform) plane get farther from the central axis, their associated spatial frequencies (m/a) increase in accord with the grating equation $\sin \theta_m = \lambda(m/a)$. A coarser grating would have a larger value of a, so that a given order (m) would be concomitant with a lower frequency, (m/a), and the spots would all be closer to the central or optical axis.

Had we used as an object a transparency resembling the sine target (Fig. 13.25*a*), such that the aperture function varied sinusoidally, there would ideally have only been three spots on the transform plane, these being the zero-frequency central peak and the first order or fundamental $(m = \pm 1)$ on either side of the center. Extending things into two dimensions, a crossed grating (or mesh) yields the diffraction pattern shown in the accompanying photo. Note that in addition to the obvious periodicity horizontally and vertically across the mesh, it is also repetitive, for example, along diagonals. A more involved object, such as a transparency of the surface of the Moon, would generate an extremely complex diffraction pattern. Because of the simple periodic nature of the grating, we could think of its Fourier-series components, but now we will certainly have to think in terms of Fourier transforms. In any case, *each spot of light in the diffraction pattern denotes the presence of a specific spatial frequency, which is proportional to its distance from the optical axis (zero-frequency location)*. Frequency components of positive and negative sign appear diametrically opposite each other about the central axis. If we could measure the electric field at each point in the

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

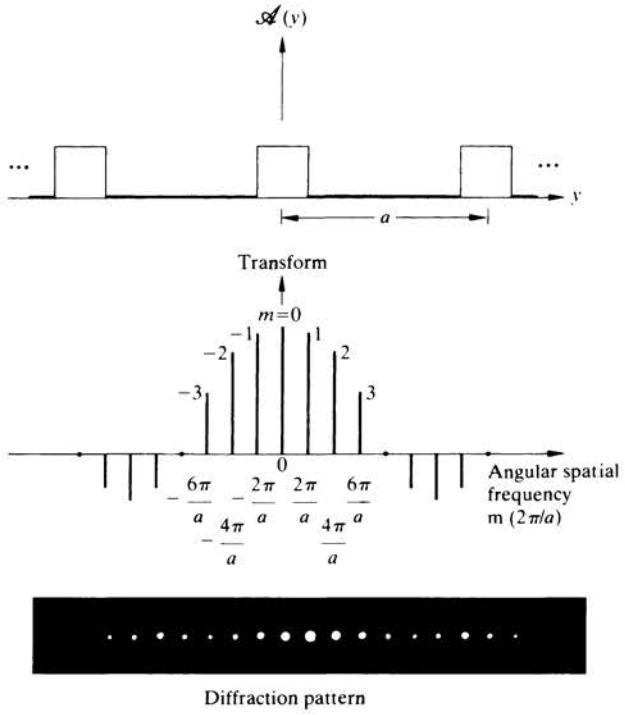
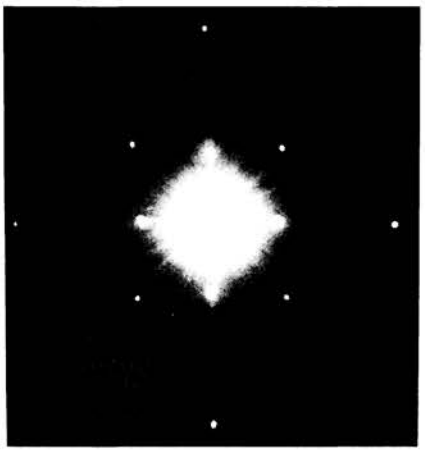


Figure 13.27 Square wave and its transform.

transform plane, we would indeed observe the transform of the aperture function, but this is not practicable. Instead, what will be detected is the flux-density distribution, where at each point the irradiance is proportional to the time average of the electric field squared or equivalently to the square of the amplitude of the particular spatial frequency contribution at that point.



Diffraction pattern of a crossed grating or mesh.

13.2.2 Abbe's Theory of Image Formation

Consider the system depicted in Fig. 13.28*a*, which is just an elaborated version of Fig. 13.28*b*. Plane monochromatic wavefronts emanating from the collimating lens (L_c) are diffracted by a grating. The result is a distorted wavefront, which we resolve into a new set of plane waves, each corresponding to a given order $m = 0, \pm 1, \pm 2, \ldots$ or spatial frequency and each traveling in a specific direction (Fig. 13.28*b*). The objective lens (L_t) serves as a *transform lens*, forming the Fraunhofer diffraction pattern of the grating on the transform plane Σ_t (which is also the back focal plane of L_t). The waves, of course, propagate beyond Σ_t and arrive at the image plane Σ_i. There they overlap and interfere to form an inverted image of the grating. Accordingly, points G_1 and G_2 are imaged at P_1 and P_2, respectively. The objective lens forms two distinct patterns of interest. One is the Fourier transform on the focal plane conjugate to the plane of the source, and the other is the image of the object, formed on the plane conjugate to the object plane. Figure 13.29 shows the same setup for a long, narrow, horizontal slit coherently illuminated.

We can envision the points S_0, S_1, S_2, and so forth in the transform plane of Fig. 13.28*a* as if they were point emitters of Huygens wavelets, and the resulting diffraction pattern on Σ_t is then the grating's image. In other words, *the image arises from a double diffraction process.* Alternatively, we can imagine that the incoming wave is diffracted by the object, and the resulting diffracted wave is then diffracted once again by the objective lens. If that lens were not there, a diffraction pattern of the object would appear on Σ_i in place of the image.

These ideas were first propounded by Professor Ernst Abbe (1840–1905) in 1873.* His interest at the time concerned the theory of microscopy, whose relationship to the above discussion is clear if we consider L_t as a microscope objective. Moreover, if the grating is replaced by a piece of some thin translucent material (i.e., the specimen being examined), which is illuminated by light from a small source and condenser, the system certainly resembles a microscope.

Carl Zeiss (1816–1888), who in the mid-1800s was running a small microscope factory in Jena, realized the shortcomings

*An alternative and yet ultimately equivalent approach was put forth in 1896 by Lord Rayleigh. He envisaged each point on the object as a coherent source whose emitted wave was diffracted by the lens into an Airy pattern. Each of these in turn was centered on the ideal image point (on Σ_i) of the corresponding point source. Thus Σ_i was covered with a distribution of somewhat overlapping and interfering Airy patterns.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

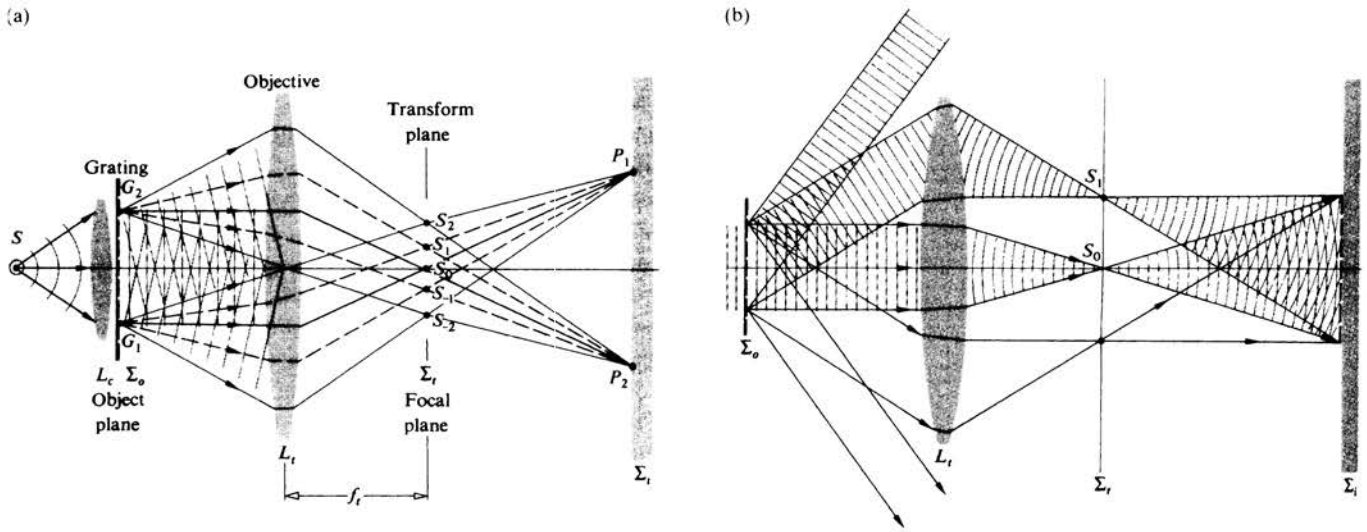


Figure 13.28 Image formation.

of the trial-and-error development techniques of that era. In 1866 he enlisted the services of Ernst Abbe, then lecturer at the University of Jena, to establish a more scientific approach to microscope design. Abbe soon found by experimentation that a larger aperture resulted in higher resolution, even though the apparent cone of incident light filled only a small portion of the objective. Somehow the surrounding "dark space" contributed to the image. Consequently, he took the approach that the then well-known diffraction process that occurs at the edge of a lens (leading to the Airy pattern for a point source) was not operative in the same sense as it was for an incoherently illuminated telescope objective. Specimens, whose size was of the order of λ, were apparently scattering light into the "dark space" of the microscope objective. Observe that if, as in Fig. 13.28*b*, the aperture of the objective is not large enough to collect all of the diffracted light, the image does not correspond exactly to that object. Rather, it relates to a fictitious object whose complete diffraction pattern matches the one collected by L_t. We know from the previous section that these lost portions of the outer region of the Fraunhofer pattern are associated with the higher spatial frequencies. And, as we shall see presently, their removal will result in a loss in image sharpness and resolution.

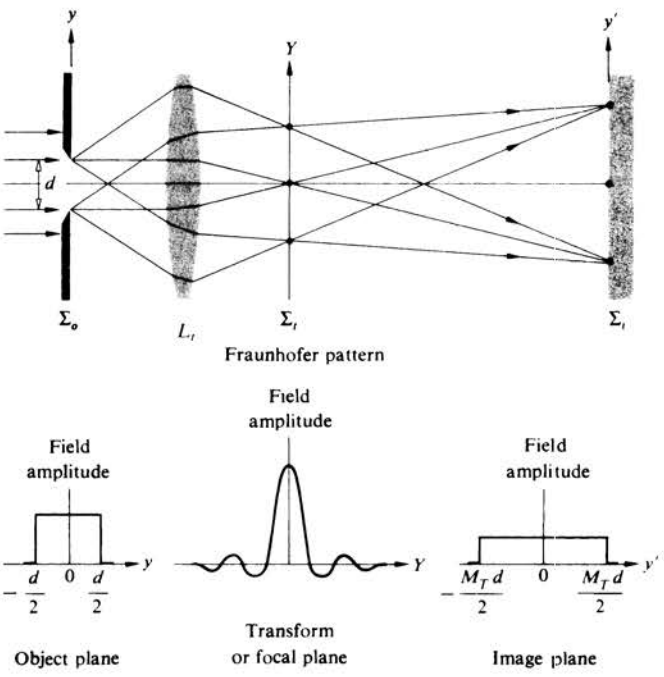


Figure 13.29 The image of a slit.

Practically speaking, unless the grating considered earlier has an infinite width, it cannot be strictly periodic. This means that it has a continuous Fourier spectrum dominated by the usual discrete Fourier-series terms, the other being much smaller in amplitude. Complicated, irregular objects clearly display the continuous nature of their Fourier transforms. In any event, it should be emphasized that *unless the objective*

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

lens has an infinite aperture, it functions as a low-pass filter rejecting spatial frequencies above a given value and passing all those below (the former being those that extend beyond the physical boundary of the lens). Consequently, all practical lens systems will be limited in their ability to reproduce the high spatial frequency content of an actual object under coherent illumination.* It might be mentioned as well that there is a basic nonlinearity associated with optical imaging systems operating at high spatial frequencies.†

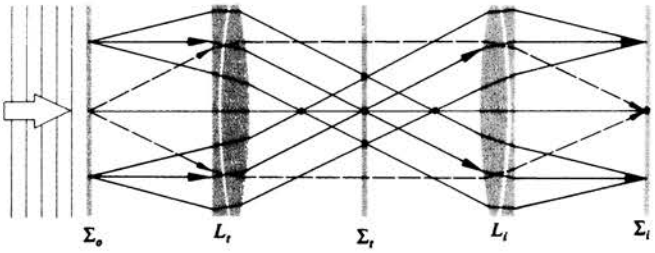


Figure 13.30 Object, transform, and image planes.

13.2.3 Spatial Filtering

Suppose we actually set up the system shown in Fig. 13.28*a*, using a laser as a plane-wave source. If the points S_0, S_1, S_2, and so on are to be the sources of a Fraunhofer pattern, the image screen must presumably be located at $x = \infty$ (although 30 or 40 ft will often do). At the risk of being repetitious, recall that the reason for using L_t originally was to bring the diffraction pattern of the object in from infinity. We now introduce an *imaging lens* L_i (Figs. 13.30 and 13.31) in order to bring in from infinity the diffraction pattern of the set of source points S_0, S_1, S_2,..., thereby relocating Σ_i at a convenient distance. The transform lens causes the light from the object to converge in the form of a diffraction pattern on the plane Σ_t; that is, it produces on Σ_t a two-dimensional Fourier transform of the object. To wit, the spatial frequency spectrum of the object is spread across the transform plane. Thereafter, L_i (the *"inverse" transform lens*) projects the diffraction pattern of the light distributed over Σ_t onto the image plane. In other words, it diffracts the diffracted beam, which effectively means that it generates an (inverted) inverse transform. Thus essentially an inverse transform of the data on Σ_t appears as the final image.

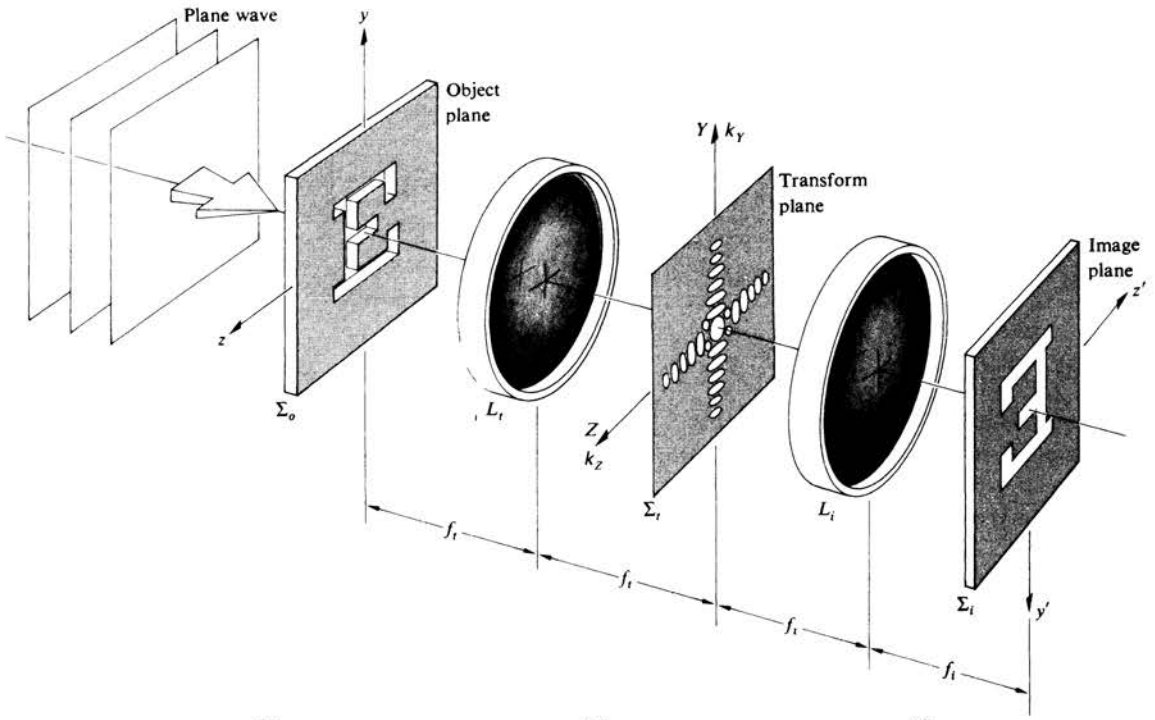
Quite frequently, in practice L_t and L_i are identical ($f_t = f_i$) well-corrected multi-element lenses [for quality work these might have resolutions of about 150 line pairs/mm—one line pair being a period in Fig. 13.25*b*]. For less demanding applications, two projector objectives of large aperture (about 100 mm) having convenient focal lengths of roughly 30 or 40 cm serve quite nicely. One of these lenses is then merely turned around so that both their back focal planes coincide with Σ_t. Incidentally, the input or object plane need not be located a focal length away from L_t; the transform still appears on Σ_t. Moving Σ_0 affects only the phase of the amplitude distribution, and that is generally of little interest. The device shown in Figs. 13.30 and 13.31 is referred to as a ***coherent optical computer***. It allows us to insert obstructions (i.e., masks or filters) into the transform plane and in so doing partially or completely block out certain spatial frequencies, stopping them from reaching the image plane. *This process of altering the frequency spectrum of the image is known as* **spatial filtering** (see Section 7.4.4).

From our earlier discussion of Fraunhofer diffraction we know that a long narrow slit at Σ_0, regardless of its orientation and location, generates a transform at Σ_t consisting of a series of dashes of light lying along a straight line perpendicular to the slit (see Fig. 10.6) and *passing through the origin*. Consequently, if the straight-line object is described by $y = mz + b$, the diffraction pattern lies along the line $Y = -Z/m$ or equivalently, from Eqs. (11.64) and (11.65), $k_Y = -k_Z/m$. With this and the Airy pattern in mind, we should be able to anticipate some of the gross structure of the transforms of various objects. Be aware as well that these transforms are centered about the zero-frequency optical axis of the system. For example, a transparent plus sign whose horizontal line is thicker than its vertical one has a two-dimensional transform again shaped more or less like a plus sign. The thick horizontal line generates a series of short vertical dashes, while the thin vertical element produces a line of long horizontal dashes. Remember that object elements with small dimensions diffract through relatively large angles. Along with Abbe, one could think of this entire subject in these terms rather than using the concepts of spatial frequency filtering and transforms, which represent the more modern influence of communication theory.

*Refer to H. Volkmann, "Ernst Abbe and His Work," *Appl. Opt.* **5**, 1720 (1966), for a more detailed account of Abbe's many accomplishments in Optics.

†R. J. Becherer and G. B. Parrent, Jr., "Nonlinearity in Optical Imaging Systems," *J. Opt. Soc. Am.* **57**, 1479 (1967).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



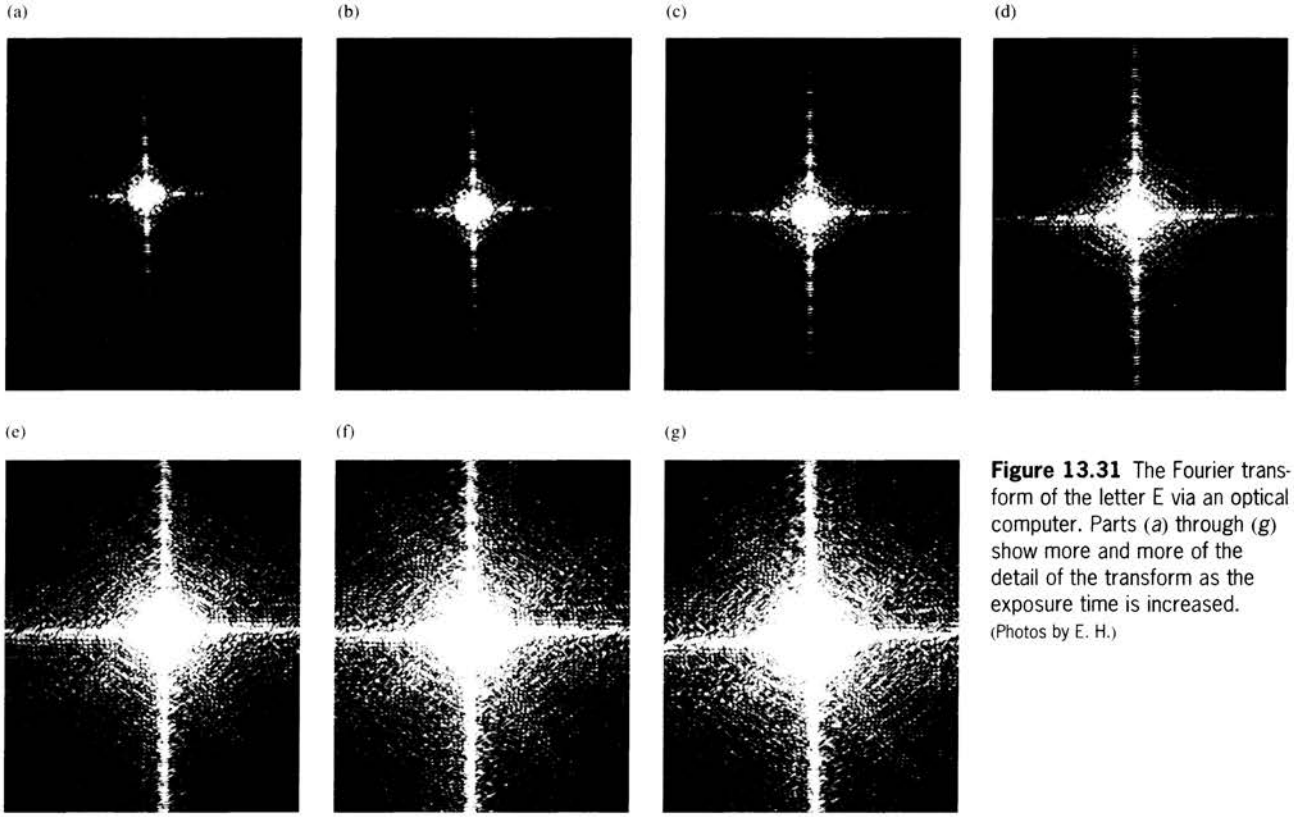


Figure 13.31 The Fourier transform of the letter E via an optical computer. Parts (*a*) through (*g*) show more and more of the detail of the transform as the exposure time is increased. (Photos by E. H.)

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

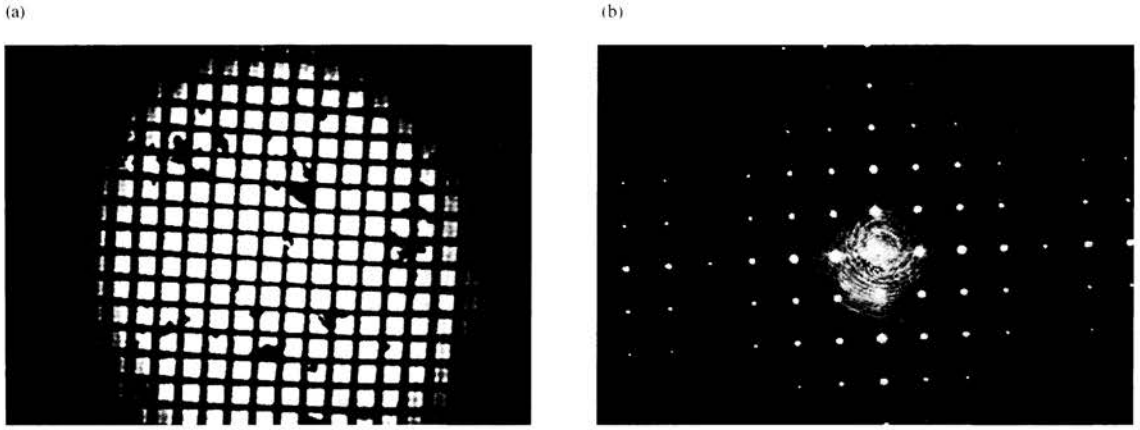


Figure 13.32 A fine, slightly dusty mesh and its transform. (Photos from D. Dutton, M. P. Givens, and R. E. Hopkins, *Spectra-Physics Laser Technical Bulletin Number 3.*)

The vertical portions of the symbol **E** in Fig. 13.31 generate the broad frequency spectrum appearing as the horizontal pattern. Note that all parallel line sources on a given object correspond to a single linear array on the transform plane. This, in turn, passes through the origin on Σ_f (the intercept is zero), just as in the case of the grating. A transparent figure 5 will generate a pattern consisting of both a horizontal and vertical distribution of spots extending over a relatively large frequency range. There will also be a comparatively low-frequency, concentric ring-like structure. The transforms of disks and rings and the like will obviously be circularly symmetric. Similarly, a horizontal elliptical aperture will generate vertically oriented concentric elliptical bands. Most often, far-field patterns possess a center of symmetry (see Problems 10.14 and 11.29).

We are now in a better position to appreciate the process of spatial filtering and to that end will consider an experiment similar to one published in 1906 by A. B. Porter. Figure 13.32*a* shows a fine wire mesh whose periodic pattern is disrupted by a few particles of dust. With the mesh at Σ_0, Fig. 13.32*b* shows the transform as it would appear on Σ_f. Now the fun starts—since the transform information relating to the dust is located in an irregular cloud-like distribution about the center point, we can easily eliminate it by inserting an opaque mask at Σ_f. If the mask has holes at each of the principal maxima, thus passing on only those frequencies, the image appears dustless (Fig. 13.33*a*). At the other extreme, if we just pass the cloud-like pattern near center, very little of the periodic structure appears, leaving an image consisting of essentially just the dust particles (Fig. 13.33*b*). Passing only the zero-order central spot generates a uniformly illuminated (dc) field, just as if the mesh were no longer in position. Observe that as more and more of the higher frequencies are eliminated, the detail of the image deteriorates markedly [(*d*), (*e*), and (*f*) in Fig. 13.33]. This can be understood quite simply by remembering how a function, with what we might call "sharp edges," was synthesized out of harmonic components. The square wave of Fig. 7.29 serves to illustrate the point. It is evident that the addition of higher harmonics serves predominantly to square up the corners and flatten out the peaks and troughs of the profile. In this way, ***the high spatial frequencies contribute to the sharp edge detail between light and dark regions of the image***. The removal of the high-frequency terms causes a rounding out of the step function and a consequent loss of resolution in the two-dimensional case.

What would happen if we took out the dc component (Fig. 13.33*c*) by passing everything but the central spot? A point on the original image that appears black in the photo denotes a near-zero irradiance and perforce a near-zero field amplitude. Presumably, all of the various optical field components completely cancel each other at that point—ergo, no light. Yet with the removal of the dc term, the point in question must certainly then have a nonzero field amplitude. When squared ($I \propto E_0^2/2$) this will generate a nonzero irradiance. It follows that regions that were originally black in the photo will now appear whitish, while regions that were white will become grayish, as is shown in the accompanying photo.

Let's now examine some of the possible applications of this technique. Figure 13.34*a* shows a composite photograph of the Moon consisting of film strips pieced together to form a

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

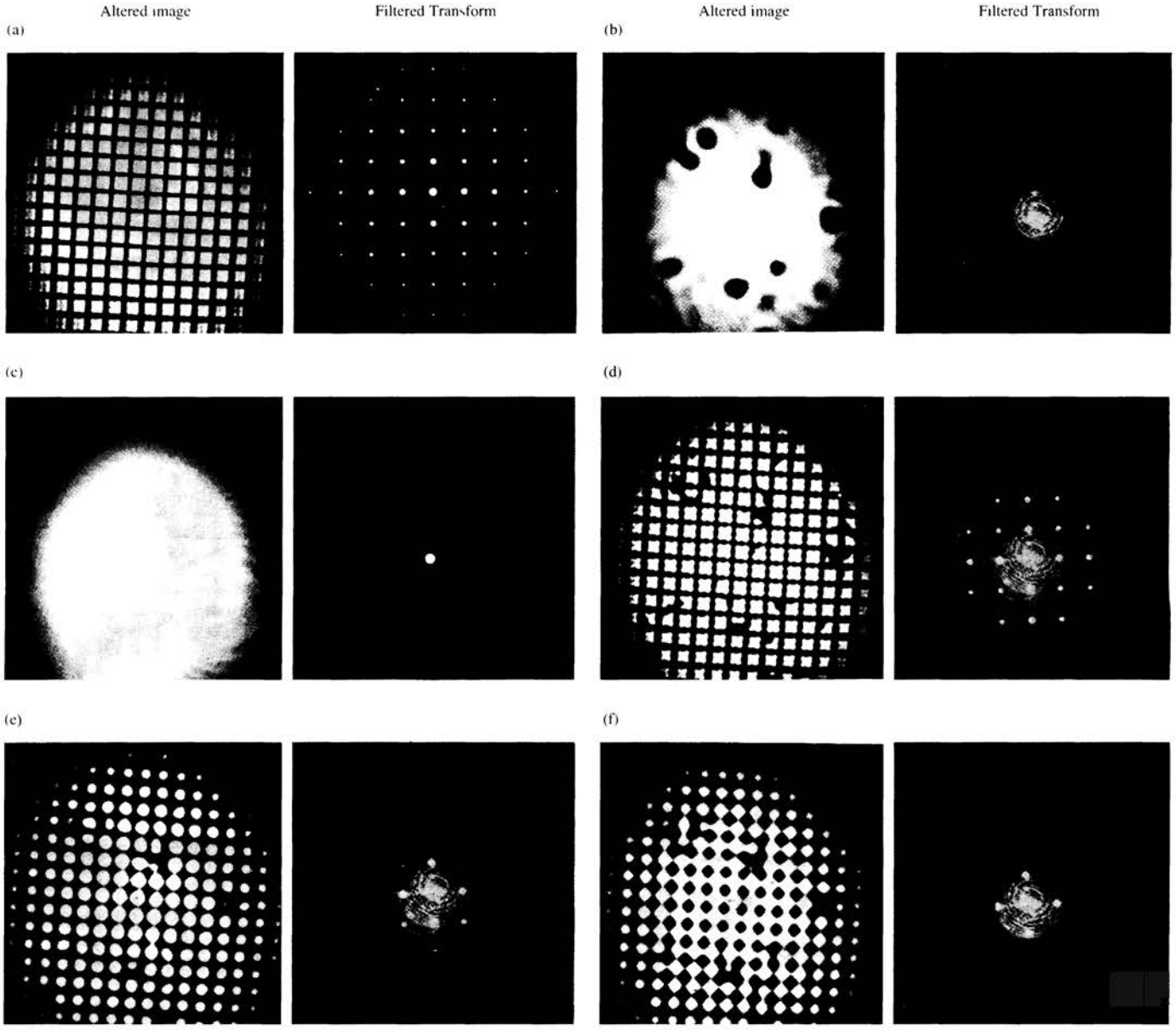
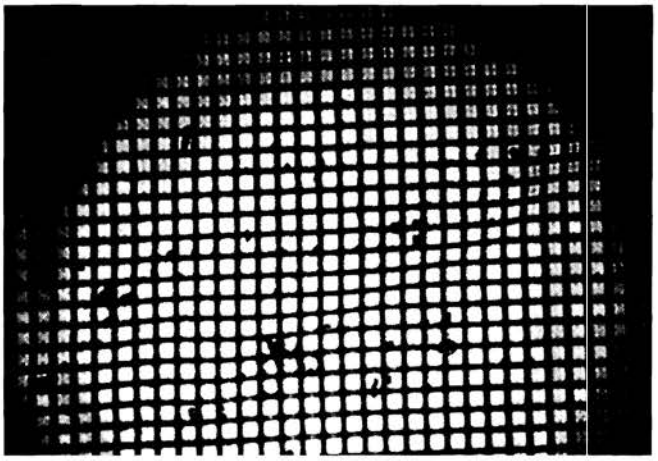


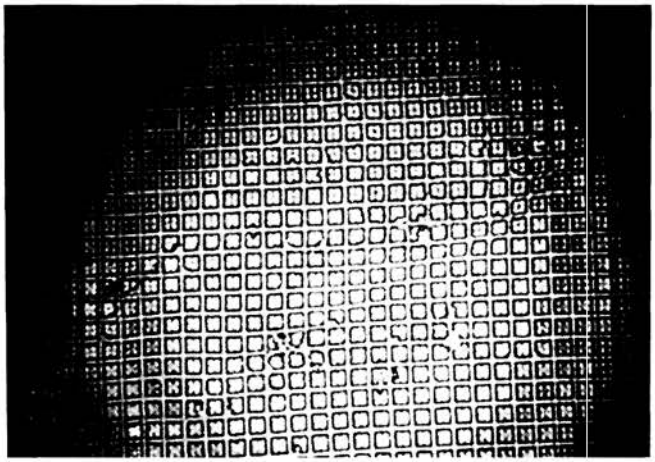
FIGURE 13.33 Images resulting when various portions of the diffraction pattern of Fig. 13.32*b* are obscured by the accompanying masks or spatial filters. (Photos from D. Dutton, M. P. Givens, and R. E. Hopkins, *Spectra-Physics Laser Technical Bulletin Number 3.*

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

(a)



(b)



Part (*b*) is a filtered version of (*a*) where the zeroth order was removed. (Photos from D. Dutton, M. P. Givens, and R. E. Hopkins, *Spectra-Physics Laser Technical Bulletin Number 3.*)

single mosaic. The video data were telemetered to Earth by *Lunar Orbiter 1*. Clearly, the grating-like regular discontinuities between adjacent strips in the object photograph generate the broad-bandwidth, vertical-frequency distribution evident in Fig. 13.34*c*. When these frequency components are blocked, the enhanced image shows no sign of having been a mosaic. In very much the same way, one can suppress extraneous data in bubble chamber photographs of subatomic particle tracks.* These photographs are made difficult to analyze because of the presence of the unscattered beam tracks (Fig. 13.35), which, since they are all parallel, are easily removed by spatial filtering.

Consider the familiar half-tone or facsimile process by which a printer can create the illusion of various tones of gray while using only black ink and white paper (take a close look at a newspaper photograph). If a transparency† of such a facsimile is inserted at Σ_0 in Fig. 13.30, its frequency spectrum will appear on Σ_f. Once again the relatively high-frequency components arising from the half-tone mesh can easily be eliminated. This yields an image in shades of gray (Fig. 13.36 on page 617) showing none of the discontinuous nature of the original. One could construct a precise filter to obstruct only the square mesh frequencies by actually using a negative transparency of the transform of the basic checkerboard array. Alternatively, it usually suffices to use a low-pass circular aperture filter, and in so doing inadvertently discard some of the high-frequency detail of the original scene, at least as long as the mesh frequency is comparatively high.

The same procedure can be used to remove the graininess of highly enlarged photographs, which is of value, for example, in aerial photo reconnaissance. In contrast, we could sharpen up the details in a slightly blurred photograph by emphasizing its high-frequency components. This could be done with a filter that preferentially absorbed the low-frequency portion of the spectrum. A great deal of effort, beginning in the 1950s, has gone into the study of photographic image enhancement, and the ensuing successes have been notable indeed. Prominent among these contributors was A. Maréchal of the Institut d'Optique, Université de Paris, who combined absorbing and phase-shifting filters to reconstitute the detail in badly blurred photographs. These filters are transparent coatings deposited on optical flats so as to retard the phase of various portions of the spectrum (Section 13.2.4).

As this work in optical data processing continues into the coming decades, we will surely see the replacement of the

*D. G. Falconer, "Optical Processing of Bubble Chamber Photographs," *Appl. Opt.* **5**, 1365 (1966), includes some additional uses for the coherent optical computer.

†Polaroid 55 P/N film is satisfactory for medium resolution work, while Kodak 649 plates are good where higher resolution is required of the transparency.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

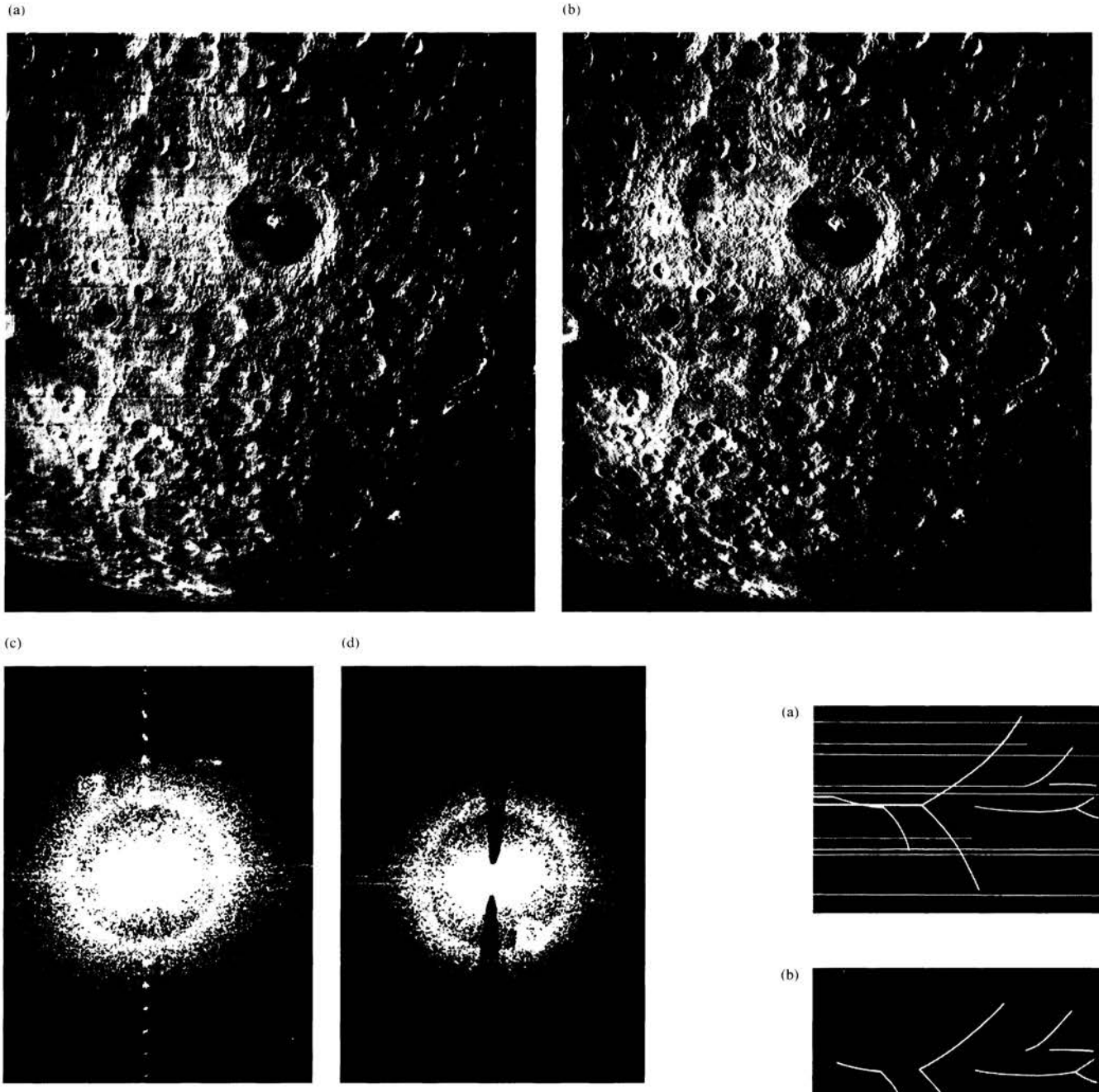


Figure 13.34 Spatial filtering. (*a*) A *Lunar Orbiter* composite photo of the Moon. (*b*) Filtered version of the photo sans horizontal lines. (*c*) A typical unfiltered transform (power spectrum) of a moonscape. (*d*) Diffraction pattern with the vertical dot pattern filtered out. (Photos courtesy D. A. Ansley, W. A. Blikken, The Conductron Corporation, and NASA.)

Figure 13.35 Unfiltered and filtered bubble-chamber tracks.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

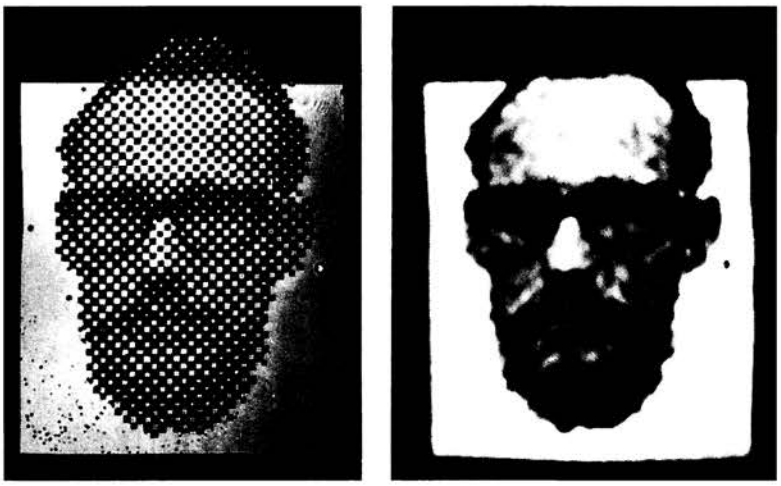


Figure 13.36 A self-portrait of K. E. Bethke consisting of only black and white regions as in a halftone. When the high frequencies are filtered out, shades of gray appear and the sharp boundaries vanish. [From R. A. Phillips, *Am. J. Phys.* **37**, 536 (1969).]

photographic stages, in increasingly many applications, by real-time electro-optical devices (e.g., arrays of ultrasonic light modulators forming a multichannel input are already in use).* The coherent optical computer will reach a certain maturity, becoming an even more powerful tool when the input, filtering, and output functions are performed electro-optically. A continuous stream of real-time data could flow into and out of such a device.

13.2.4 Phase Contrast

It was mentioned briefly in the last section that the reconstructed image could be altered by introducing a phase-shifting filter. Probably the best-known example of this technique dates back to 1934 and the work of the Dutch physicist Fritz Zernike, who invented the method of **phase contrast** and applied it in the *phase-contrast microscope*.

An object can be "seen" because it stands out from its surroundings—it has a color, tone, or lack of color, which provides contrast with the background. This kind of structure is known as an *amplitude object*, because it is observable by dint of variations that it causes in the amplitude of the lightwave. The wave that is either reflected or transmitted by such an object becomes *amplitude modulated* in the process. In contradistinction, it is often desirable to "see" *phase objects*, that is, ones that are transparent, thereby providing practically no contrast with their environs and altering only the phase of the detected wave. The optical thickness of such objects generally varies from point to point as either the refractive index or the actual thickness, or both, vary. Obviously, since the eye cannot detect phase variations, such objects are invisible. This is the problem that led biologists to develop techniques for staining transparent microscope specimens and in so doing to convert phase objects into amplitude objects. But this approach is unsatisfactory in many respects, for example, when the stain kills the specimen whose life processes are under study, as is all too often the case.

Recall that diffraction occurs when a portion of the surface of constant phase is obstructed in some way, that is, when a region of the wavefront is altered (either in amplitude or phase, i.e., shape). Suppose then that a plane wave passes through a transparent particle, which retards the phase of a region of the front. The emerging wave is no longer perfectly planar but contains a small indentation corresponding to the area retarded by the specimen; the wave is *phase modulated*.



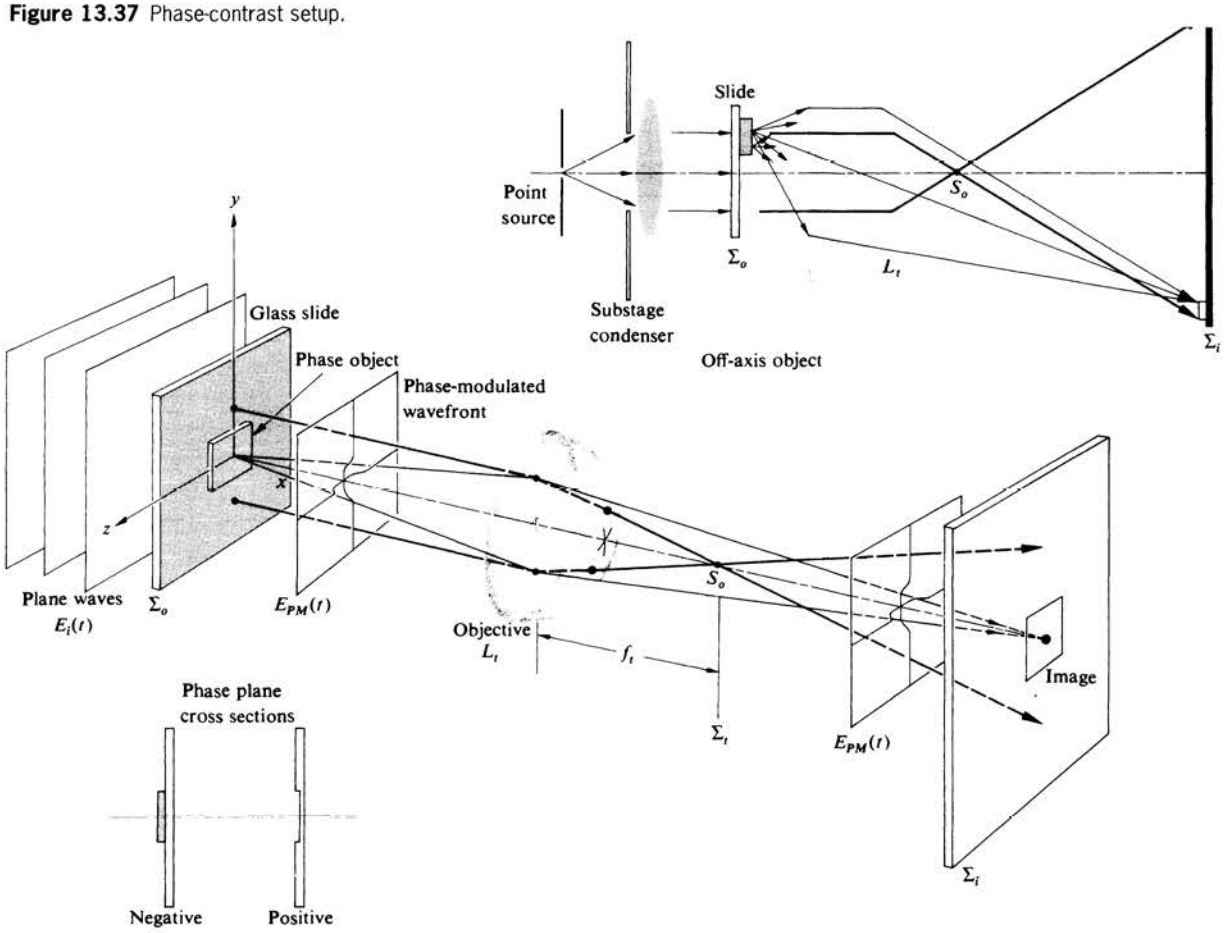
Fritz Zernike (1888-1966) Won the Nobel Prize for Physics in 1953.

Taking a rather simplistic view of things, we can imagine the phase-modulated wave $E_{PM}(\vec{r},t)$ (Fig. 13.37) to consist of the original incident plane wave $E_i(x, t)$ plus a localized disturbance $E_d(\vec{r}, t)$. (The symbol $\vec{\mathbf{r}}$ means that E_{PM} and E_d depend on x, y, and z; i.e., they vary over the yz-plane, whereas E_i is uniform and does not.) Indeed, if the phase retardation is very small, the localized disturbance is a wave of

*We have only touched on the subject of optical data processing; a more extensive discussion of these matters is given, for example, by Goodman in *Introduction to Fourier Optics*, Chapter 7. That text also includes a good reference list for further reading in the journal literature. Also see P. F. Mueller, "Linear Multiple Image Storage," *Appl. Opt.* **8**, 267 (1969). Here, as in much of modern Optics, the frontiers are fast moving, and obsolescence is a hard rider.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Figure 13.37 Phase-contrast setup.



very small amplitude, E_{0d}, lagging by just about $\lambda_0/4$, as in Fig. 13.38. There the difference between $E_{PM}(\vec{r}, t)$ and $E_i(x, t)$ is shown to be $E_d(\vec{r}, t)$. The disturbance $E_i(x, t)$ is called the direct or *zeroth-order wave*, while $E_d(\vec{r}, t)$ is the *diffracted wave*. The former produces a uniformly illuminated field at Σ_i, which is unaffected by the object, while the latter carries all of the information about the optical structure of the particle. After broadly diverging from the object, these higher-order spatial frequency terms (see Section 13.2.2) are caused to converge on the image plane. The direct and diffracted waves recombine out-of-phase by $\pi/2$, again forming the phase-modulated wave. Since the amplitude of the reconstructed wave $E_{PM}(\vec{r}, t)$ is everywhere the same on Σ_i, even though the phase varies from point to point, the flux density is uniform, and no image is perceptible. Likewise, the zeroth-order spectrum of a phase grating will be $\pi/2$ out-of-phase with the higher-order spectra.

If we could somehow shift the relative phase between the diffracted and direct beams by an additional $\pi/2$ prior to their recombination, they would still be coherent and could then interfere either constructively or destructively (Fig. 13.39). In either case, the reconstructed wavefront over the region of the image would then be amplitude modulated—the image would be visible.

We can see this in a very simple analytical way where

$$E_i(x, t)|_{x=0} = E_0 \sin \omega t$$

is the incoming monochromatic lightwave at Σ_o without the specimen in place. The particle will induce a position-dependent phase variation $\phi(y, z)$ such that the wave just leaving it is

$$E_{PM}(\vec{r}, t)|_{x=0} = E_0 \sin [\omega t + \phi(y, z)] \qquad (13.20)$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

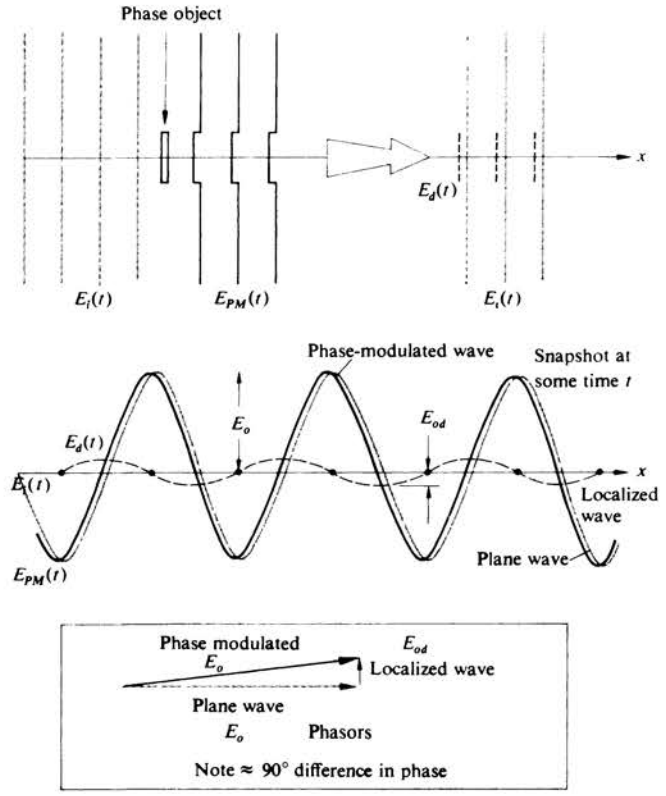


Figure 13.38 Wavefronts in the phase-contrast process.

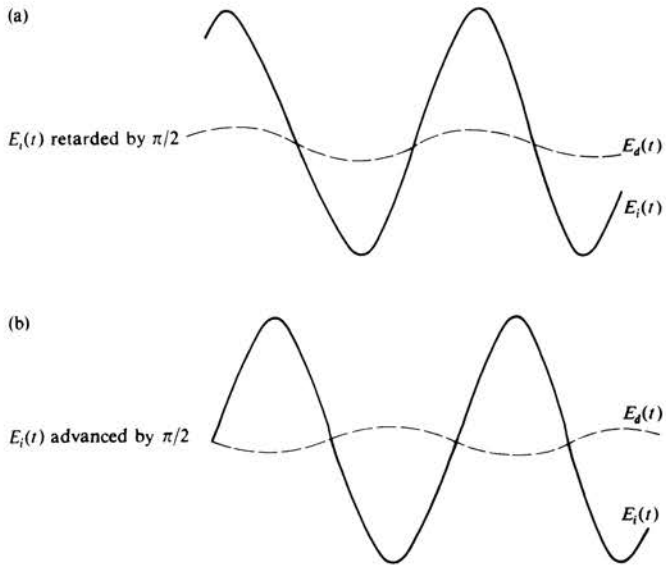


Figure 13.39 Effect of phase shifts.

This is a constant-amplitude wave, which is essentially the same on the conjugate image plane. That is, there are some losses, but if the lens is large and aberration-free and we neglect the orientation and size of the image, Eq. (13.20) will suffice to represent the PM wave on either Σ_o or Σ_i. Reformulating that disturbance as

$$E_{PM}(y, z, t) = E_0 \sin \omega t \cos \phi + E_0 \cos \omega t \sin \phi$$

and limiting ourselves to *very small values* of ϕ, we obtain

$$E_{PM}(y, z, t) = E_0 \sin \omega t + E_0 \phi(y, z) \cos \omega t$$

The first term is independent of the object, while the second term obviously isn't. Thus, as above, if we change their relative phase by $\pi/2$, that is, either change the cosine to sine or vice versa, we get

$$E_{AM}(y, z, t) = E_0[1 + \phi(y, z)] \sin \omega t \qquad (13.21)$$

which is an amplitude-modulated wave. Observe that $\phi(y, z)$ can be expressed in terms of a Fourier expansion, thereby introducing the spatial frequencies associated with the object. Incidentally, this discussion is precisely analogous to the one proposed in 1936 by E. H. Armstrong for converting AM radio waves to FM [$\phi(t)$ could be thought of as a frequency modulation wherein the zeroth-order term is the carrier]. An electrical bandpass filter was used to separate the carrier from the remaining information spectrum so that the $\pi/2$ phase shift could be accomplished. Zernike's method of doing essentially the same thing is as follows. He inserted a spatial filter in the transform plane Σ_i of the objective, which was capable of inducing the $\pi/2$ phase shift. Observe that the direct light actually forms a small image of the source on the optical axis at the location of Σ_i. The filter could then be a small circular indentation of depth d etched in a transparent glass plate of index n_g. Ideally, only the direct beam would pass through the indentation, and in so doing it would take on a *phase advance* with respect to the diffracted wave of $(n_g - 1)d$, which is made to equal $\lambda_0/4$. A filter of this sort is known as a ***phase plate***, and since its effect corresponds to Fig. 13.39*b*, that is, destructive interference, phase objects that are thicker or have higher indices appear dark against a bright background. If, instead, the phase plate had a small raised disk at its center, the opposite would be true. The former case is called *positive-phase contrast*; the latter, *negative-phase contrast*.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

In actual practice, a brighter image is obtained by using a broad, rather than a point, source along with a substage condenser. The emerging plane waves illuminate an annular diaphragm (Fig. 13.40), which, since it is the source plane, is conjugate to the transform plane of the objective. The zeroth-order waves, shown in the figure, pass through the object according to the tenets of geometrical optics. They then traverse the thin annular region of the phase plate located at Σ_t. That region of the plate is quite small, and so the cone of diffracted rays, for the most part, misses it. By making the annular region absorbing as well (a thin metal film will do), the very large uniform zeroth-order term (Fig. 13.41) is reduced with respect to the higher orders, and the contrast improves. Or, if you like, E_0 is reduced to a value comparable with that of the diffracted wave E_{0d}. Generally, a microscope will come with an assortment of these phase plates having different absorptions.

In the parlance of modern Optics (the still-blushing bride of communications theory), phase contrast is simply the process whereby we introduce a $\pi/2$ phase shift in the zeroth-order spectrum of the Fourier transform of a phase object (and perhaps attenuate its amplitude as well) through the use of an appropriate spatial filter.

The phase-contrast microscope, which earned Zernike the Nobel Prize in 1953, has found extensive applications (see photo), perhaps the most fascinating of which is the study of the life functions of otherwise invisible organisms.

13.2.5 The Dark-Ground and Schlieren Methods

Suppose we go back to Fig. 13.37, where we were examining a phase object, and this time rather than retard and attenuate the central zeroth order, we remove it completely with an opaque disk at S_o. Without the object in place, the image plane

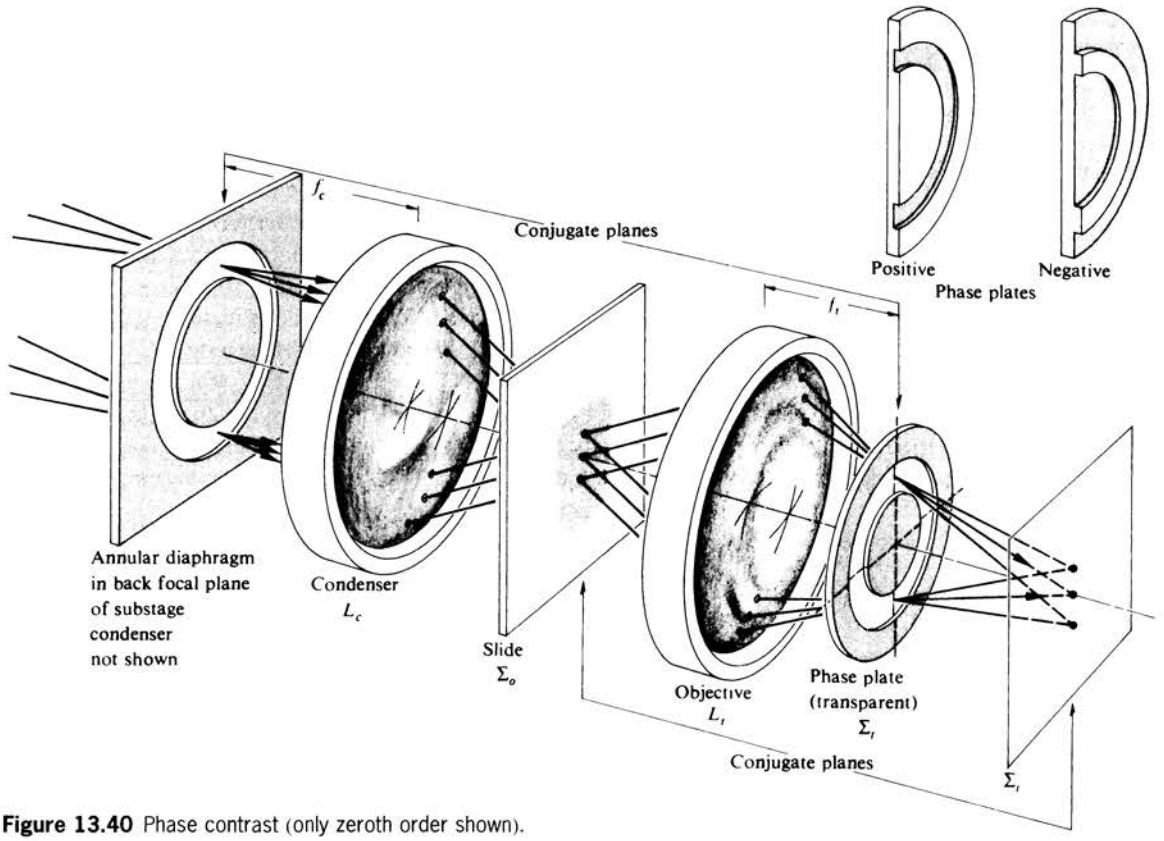


Figure 13.40 Phase contrast (only zeroth order shown).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

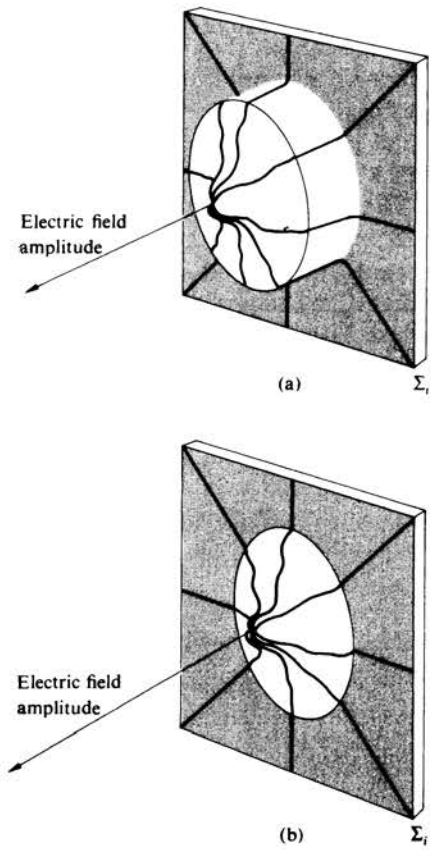


Figure 13.41 Field amplitude over a circular region on the image plane. In one case there is no absorption in the phase plate, and the irradiance would be a small ripple on a great plateau. With the zeroth order attenuated, the contrast increases.

will be completely dark—ergo the name ***dark ground***. With the object in position, only the localized diffracted wave will appear at Σ_i to form the image. (This can also be accomplished in microscopy by illuminating the object obliquely so that no direct light enters the objective lens.) Observe that by eliminating the dc contribution, the amplitude distribution (as in Fig. 13.41) will be lowered and portions that were near zero prior to filtering will become negative. Inasmuch as irradiance is proportional to the amplitude squared, this will result in somewhat of a contrast reversal from that which would have been seen in phase contrast (see Section 13.2.3). In general, this technique has not been as satisfactory as the phase-contrast method, which generates a flux-density distribution across the image that is directly proportional to the phase variations induced across the object.

In 1864 A. Toepler introduced a procedure for examining defects in lenses, which has come to be known as the **schlieren** method.* We will discuss it here because of the widespread current usage of the method in a broad range of fluid dynamics studies and furthermore because it is another beautiful example of the application of spatial filtering. Schlieren systems are particularly useful in ballistics, aerodynamics, and ultrasonic wave analysis (see photo)—indeed, wherever it is desirable to examine pressure variations as revealed by refractive-index mapping.

Suppose that we set up any one of the possible arrangements for viewing Fraunhofer diffraction (e.g., Fig. 10.3 or P.10.5). But now, instead of using an aperture of some sort as

*The word *Schlieren* in German means streaks or striae. It's frequently capitalized because all nouns are capitalized in German and not because there was a Mr. Schlieren.



(a) (b)

(*a*) A conventional photomicrograph of diatoms, fibers, and bacteria. (*b*) A phase photomicrograph of the same scene. (Photos by T. J. Lowery and R. Hawley.)

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



A schlieren photo of a spoon in a candle flame. (Photo by E. H.)

the diffracting amplitude object, we insert a phase object, for example, a gas-filled chamber (Fig. 13.42). Again a Fraunhofer pattern is formed in Σ_f, and if that plane is followed by the objective lens of a camera, an image of the chamber is formed on the film plane. We would then photograph any amplitude objects within the test area, but, of course, phase objects would still be invisible. Imagine that we now introduce a knife edge at Σ_f, raising it from below until it obstructs (sometimes only partially) the zeroth-order light and therefore all the higher orders on the bottom side as well. Just as in the dark-ground method, phase objects are then perceptible. Inhomogeneities in the test chamber windows and flaws in the lenses are also noticeable. For this reason and because of the large field of view usually required, mirror systems (Fig. 13.43) have not become commonplace.

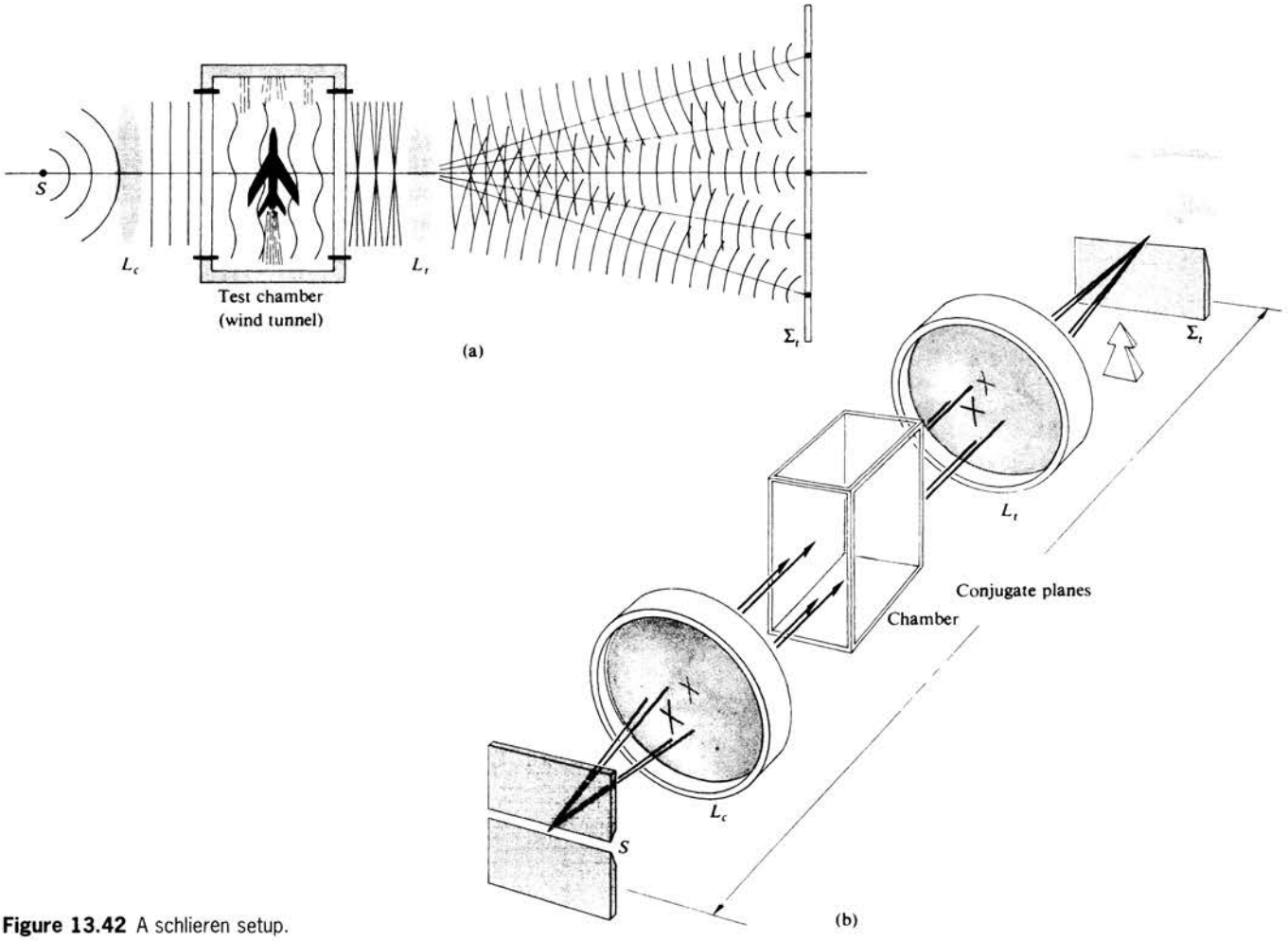


Figure 13.42 A schlieren setup.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

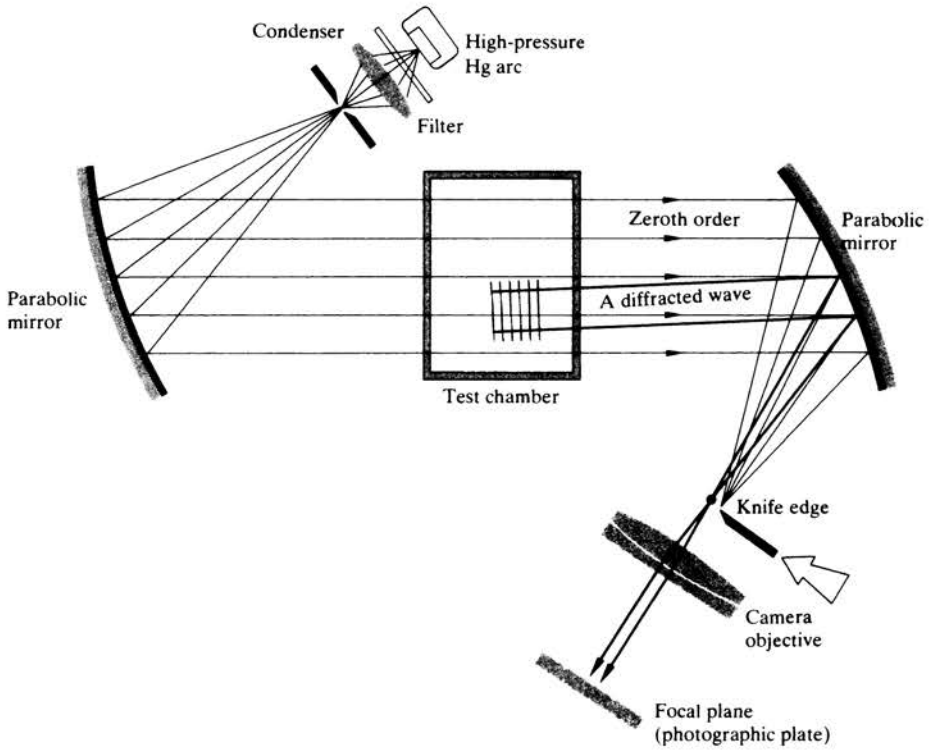


Figure 13.43 A schlieren setup using mirrors.

Quasimonochromatic illumination is generally made use of when resulting data are to be analyzed electronically, for example, with a photodetector. Sources with a broad spectrum, on the other hand, allow us to exploit the considerable color sensitivity of photographic emulsions, and a number of color schlieren systems have been devised.

13.3 Holography

The technology of photography has been with us for a long time, and we've all grown accustomed to seeing the three-dimensional world compressed into the flatness of a scrapbook page. The depthless television pitchman who smiles out of a myriad of phosphorescent flashes, although inescapably there, seems no more palpable than a postcard image of the Eiffel Tower. Both share the severe limitation of being simply irradiance mappings. In other words, when the image of a scene is ordinarily reproduced, by whatever traditional means, what we ultimately see is not an accurate reproduction of the light field that once inundated the object, but rather a point-by-point record of just the square of the field's amplitude. The light reflecting off a photograph carries with it information about the irradiance but nothing about the phase of the wave that once emanated from the object. Indeed, if both the amplitude and phase of the original wave could be reconstructed somehow, the resulting light field (assuming the frequencies are the same) would be indistinguishable from the original. This means that you would then see (and could photograph) the reformed image in perfect three-dimensionality, exactly as if the object were there before you, actually generating the wave.

13.3.1 Methods

Dennis Gabor had been thinking along these lines for a number of years prior to 1947, when he began conducting his now famous experiments in holography at the Research Laboratory of the British Thomson–Houston Company. His original setup, depicted in Fig. 13.44, was a two-step lensless imaging process in which he first



Dennis Gabor (1900–1979) Hungarian-born British physicist won the Nobel Prize in 1971.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

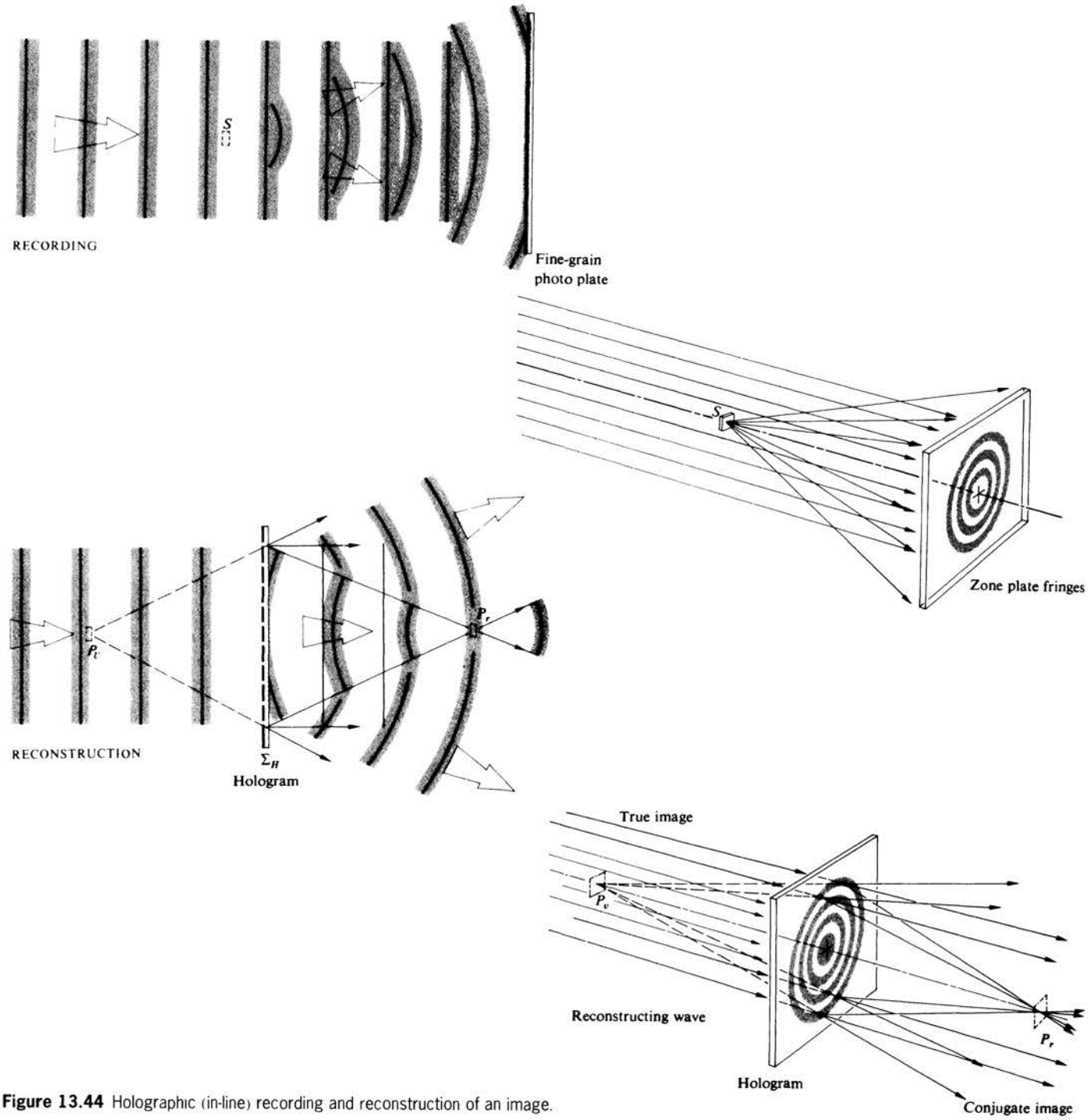


Figure 13.44 Holographic (in-line) recording and reconstruction of an image.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

photographically recorded an interference pattern, generated by the interaction of scattered quasimonochromatic light from an object and a coherent reference wave. The resulting pattern was something he called a **hologram**, after the Greek word *holos*, meaning whole. The second step in the procedure was the *reconstruction* of the optical field or image, and this was done through the diffraction of a coherent beam by a transparency, which was the developed hologram. In a way quite reminiscent of Zernike's phase-contrast technique (Section 13.2.4), the hologram was formed when the unscattered *background* or *reference wave* interfered with the diffracted wave from the small semitransparent object, S—which was, in those early days, often a piece of microfilm. The key point is that the interference pattern or hologram contains, by way of the fringe configuration, information corresponding to both the amplitude and phase of the wave scattered by the object.

Admittedly, it's not at all obvious that by now shining a plane wave through the processed hologram one could reconstruct an image of the original object. Suffice it to say for the moment that if the object were very small, the scattered wave would be nearly spherical, and the interference pattern a series of concentric rings (centered about an axis through the object and normal to the plane wave). Except for the fact that the circular fringes would vary gradually in irradiance from one to the next, the resulting flux-density distribution would correspond to a conventional Fresnel zone plate (Section 10.3.5). Recall that a zone plate functions somewhat like a lens in that it diffracts collimated light into a beam converging to a real focal point, P_r. In addition, it produces a diverging wave, which appears to come from the point P_r and constitutes a virtual image. Thus we can imagine, albeit rather simplistically, that each point on an extended object generates its own zone plate displaced from the others and that the ensemble of all such partially overlapping zone plates forms the hologram.*

During the reconstruction step, each constituent zone plate forms both a real and virtual image of a single object point, and in this way, point by point, the hologram regenerates the original light field. When the reconstructing beam has the same wavelength as the initial recording beam (which need not necessarily be the case, and quite often isn't), the virtual image is undistorted and appears at the location formerly occupied by the object. Thus it is the virtual image field that actually corresponds to the original object field. As such, the virtual image is sometimes spoken of as the *true image*, while the other is the real or, perhaps more fittingly, the *conjugate image*. In any event, we envision the hologram as a composite of interference patterns, and at least for this very simple configuration, those patterns resemble zone plates. As we will see presently, the sinusoidal grating is an equally fundamental fringe system making up complex holograms.

Gabor's research, which won him the 1971 Nobel Prize in Physics, had as its motivation an improvement in electron microscopy. His work initially generated some interest, but all in all it remained in a state of quasi-unnoticed oblivion for about 15 years. In the early 1960s there was a resurgence of interest in Gabor's **wavefront reconstruction** process and, in particular, in its relation to certain radar problems. Soon, aided by an abundance of the new coherent laserlight and extended by a number of technological advances, holography became a subject of widespread research and tremendous promise. This rebirth had its origin in the Radar Laboratory of the University of Michigan, with the work of Emmett N. Leith and Juris Upatnieks. Among other things, they introduced an improved arrangement for generating holograms, which is illustrated in Fig. 13.45 on page 626. Unlike Gabor's *in line*-configuration, where the conjugate image was inconveniently located in front of the true image, the two were now satisfactorily separated off-axis, as shown in the diagram. Once again, the hologram is an interference pattern arising from a coherent reference wave and a wave scattered from the object (this type is sometimes referred to as a **side-band Fresnel hologram**). Figure 13.46 shows the equivalent arrangement for producing side-band Fresnel holograms from transparent objects.

What's happening here can be appreciated in two ways—an essentially pictorial, Fourier-optical way and, alternatively, a direct mathematical way. We will look from both perspectives, because they complement each other. First, this is at heart an interference (or, if you like, a diffraction) problem, and we can again return to the notion of the complicated object wavefront being composed of Fourier-component plane waves (Figs. 7.43 and 10.7*d*) traveling in directions associated with the different spatial frequencies of the object's light field, reflected or transmitted. Each one of these Fourier plane waves interferes with the reference wave on the photographic plate and thus preserves the information associated with that particular spatial frequency in the form of a characteristic fringe pattern.

*See M. P. Givens, "Introduction to Holography," *Am. J. Phys.* **35**, 1056 (1967).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

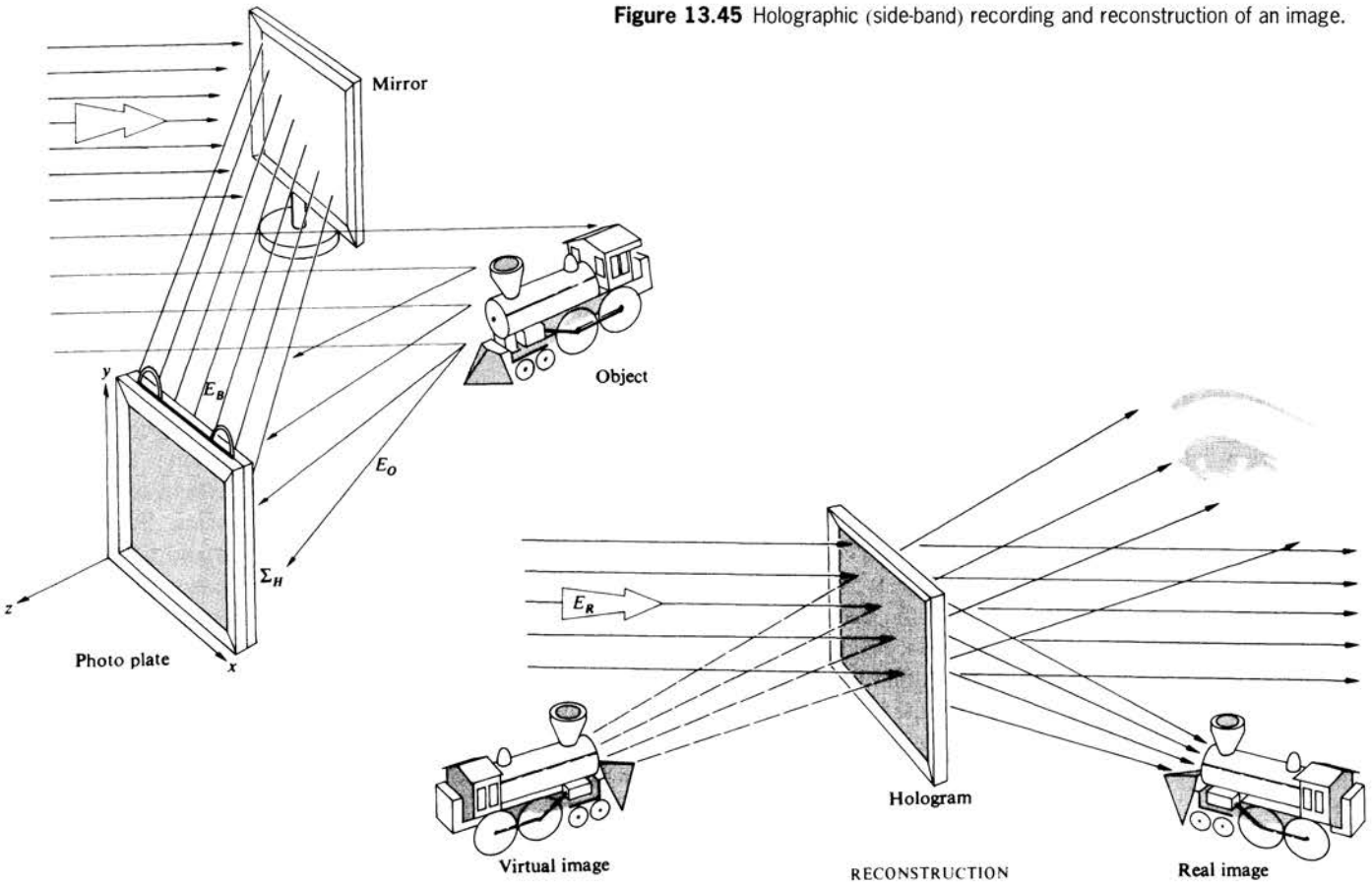


Figure 13.45 Holographic (side-band) recording and reconstruction of an image.

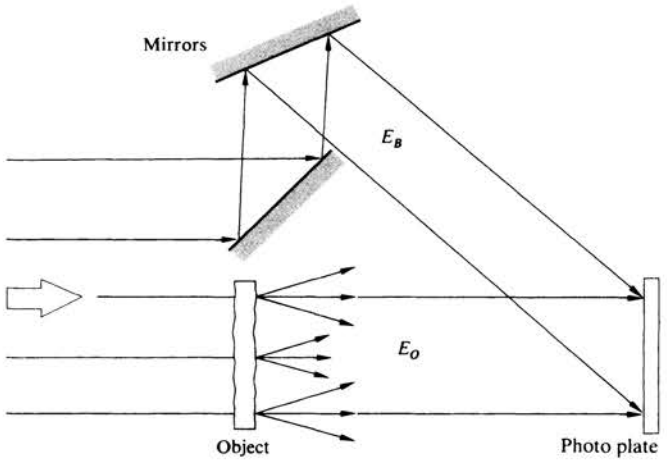


Figure 13.46 A side-band Fresnel holographic setup for a transparent object.

To see how this occurs, examine the simplified two-wave version depicted in Fig. 13.47. At the moment shown, the reference wave happens to have a crest along the face of the film plane, and the scattered object wavelet, coming in at an angle θ, similarly has crests at points A, B, and C. These correspond to points where interference maxima will occur at the moment shown. But as both waves progress to the right, they will remain in-phase at these points, trough will overlap trough, and the maxima will remain fixed at A, B, and C. Similarly, between these points, trough overlaps crest, and minima exist. The relative phase (ϕ) of these two waves, which varies from point to point along the film, can be written as a function of x. Since ϕ changes by 2π as x goes the length of $\overline{AB}$, $\phi/2\pi = x/\overline{AB}$. Notice that $\sin\theta = \lambda/\overline{AB}$, and so getting rid of the specific length $\overline{AB}$, the phase in general becomes

$$\phi(x) = (2\pi x \sin\theta)/\lambda \qquad (13.22)$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

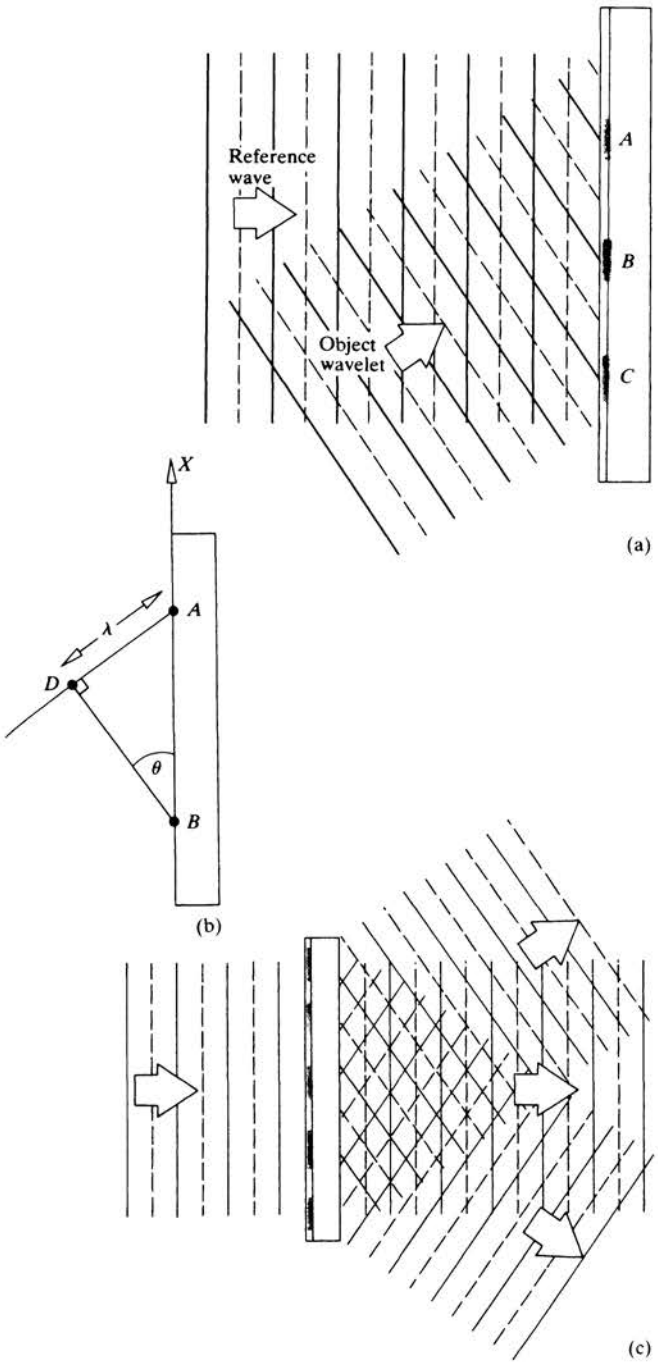


Figure 13.47 The interference of two plane waves to create a cosine grating.

If the two waves are assumed to have the same amplitude E_0, the resultant field follows from Eq. (7.17):

$$E = 2E_0 \cos \tfrac{1}{2}\phi \sin (\omega t - kx - \tfrac{1}{2}\phi)$$

and the irradiance distribution, which is proportional to the field amplitude squared, by way of Eq. (3.44), has the form

$$I(x) = \tfrac{1}{2}c\epsilon_0(2E_0 \cos \tfrac{1}{2}\phi)^2 = 2c\epsilon_0 E_0^2 \cos^2 \tfrac{1}{2}\phi$$

or

$$I(x) = 2c\epsilon_0 E_0^2 + 2c\epsilon_0 E_0^2 \cos \phi. \qquad (13.23)$$

What we have is a cosinusoidal irradiance distribution across the film plane with a spatial period of $\overline{AB}$ and a spatial frequency $(1/\overline{AB})$ of sin θ/λ.

Upon processing the film so that the amplitude transmission profile corresponds to $I(x)$, the result is a cosinusoidal grating. When this simple hologram (which essentially corresponds to a structureless object with no information) is illuminated by a plane wave identical to the original reference wave (Fig. 13.47*c*) three beams will emerge: one zeroth and two first order. One of these first-order beams will travel in the direction of the original object beam and corresponds to its reconstructed wavefront.

Now suppose we go one step beyond this most basic hologram and examine an object that has some optical structure. Accordingly, let's use as the object a transparency with a simple periodic structure that has a single spatial frequency—a cosine grating. A slightly idealized representation (which leaves out the weak higher-order terms due to the finite size of the beam and grating) is depicted in Fig. 13.48, which shows the illuminated grating, the three transmitted beams, and the reference beam. What results is three slightly different versions of Fig. 13.47, where each of the three transmitted waves makes a slightly different angle (θ) with the reference wave. Consequently, each of the three overlap areas will correspond to a set of cosine fringes of a slightly different spatial frequency, from Eq. (13.22). Again when we play back the resulting hologram, Fig. 13.48*a* and *b*, we have three pieces of business: the undiffracted wave, the virtual image, and the real image. Observe that it is only where the three beams come together to contribute their spatial frequency content that images of the original grating are formed.

When a still more complex object is used, we can anticipate that the relative phase between the object and reference waves (ϕ) will vary from point to point in a complicated way, thereby modulating the basic carrier signal (Fig. 13.49) produced

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

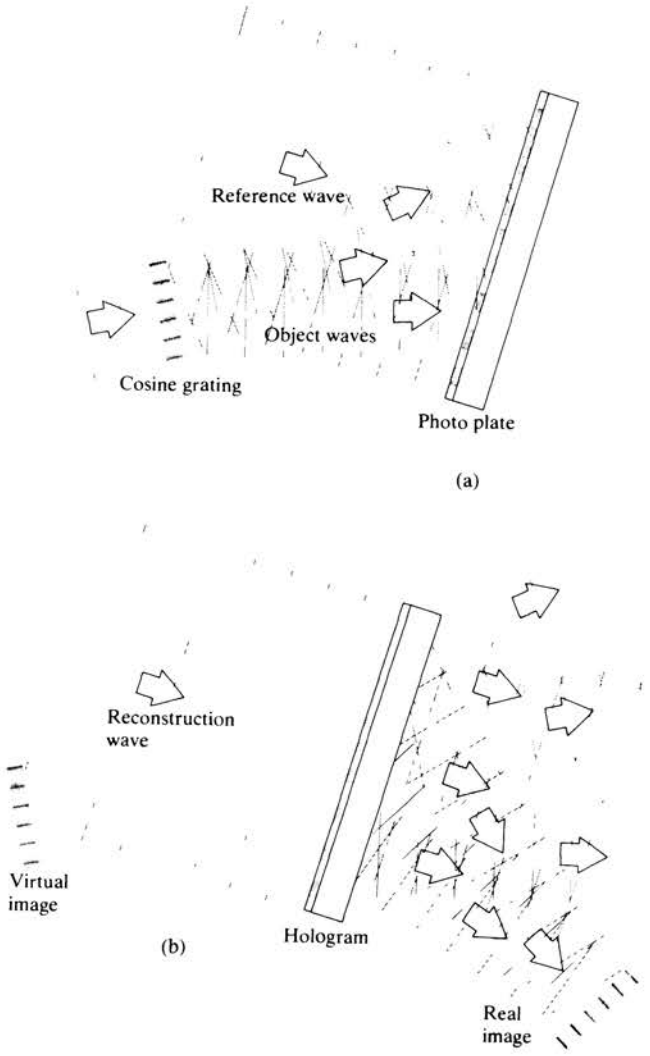


Figure 13.48 Notice that there are three regions with different spatial frequencies. Each of these on the reilluminated hologram generates three waves.

by two plane waves when no object is present. We can generalize from Fig. 13.48 and conclude that the phase-angle difference ϕ (which varies with θ) is encoded in the configuration of the fringes. Furthermore, had the amplitudes of the reference and object waves been different, the irradiance of those fringes would have been altered accordingly. Thus we can guess that the amplitude of the object wave at every point on the film plane will be encoded in the visibility of the resulting fringes.

The process depicted in Fig. 13.45 can be treated analytically as follows. Suppose that the xy-plane is the plane of the hologram, Σ_H. Then

$$E_B(x, y) = E_{0B} \cos [2\pi\nu t + \phi(x, y)] \qquad (13.24)$$

describes the planar background or reference wave at Σ_H, overlooking considerations of polarization. Its amplitude, E_{0B}, is constant, while the phase is a function of position. This just means that the reference wavefront is tilted in some known manner with respect to Σ_H. For example, if the wave were oriented such that it could be brought into coincidence with Σ_H by a single rotation through an angle of θ about y, the phase at any point on the hologram plane would depend on its value of x. Thus ϕ would again have the form

$$\phi = \frac{2\pi}{\lambda} x \sin \theta = kx \sin \theta$$

being, in that particular case, independent of ν and varying linearly with x. For the sake of simplicity, we'll just write it, quite generally, as $\phi(x, y)$ and keep in mind that it's a simple known function. The wave scattered from the object can, in turn, be expressed as

$$E_O(x, y) = E_{0O}(x, y) \cos [2\pi\nu t + \phi_O(x, y)] \qquad (13.25)$$

where both the amplitude and phase are now complicated functions of position corresponding to an irregular wavefront. From the communications-theoretic point of view, this is an amplitude- and phase-modulated carrier wave bearing all of the available information about the object. Note that this information is encoded in spatial rather than temporal variations of the wave. The two disturbances E_B and E_O superimpose and interfere to form an irradiance distribution, which is recorded by the photographic emulsion. The resulting irradiance, except for a multiplicative constant, is $I(x, y) = \langle (E_B + E_O)^2 \rangle_T$, which, from Section 9.1, is given by

$$I(x, y) = \frac{E_{0B}^2}{2} + \frac{E_{0O}^2}{2} + E_{0B}E_{0O} \cos (\phi - \phi_O) \qquad (13.26)$$

Observe once again that the phase of the object wave determines the location on Σ_H of the irradiance maxima and minima. Moreover, the contrast or fringe visibility

$$\mathcal{V} \equiv (I_{max} - I_{min})/(I_{max} + I_{min}) \qquad [12.1]$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

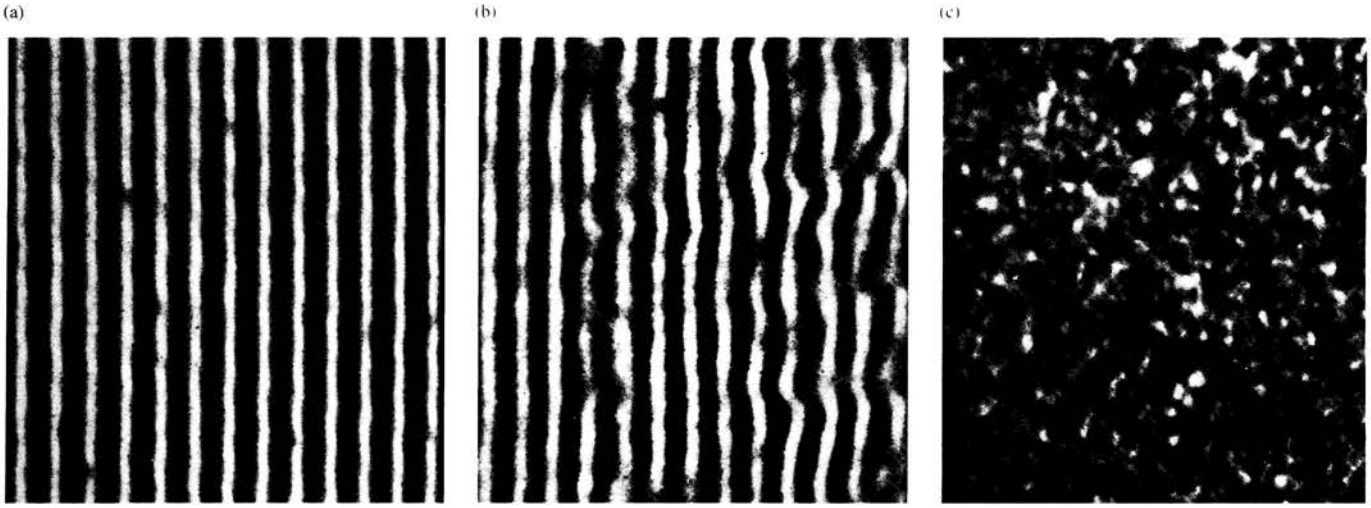


Figure 13.49 Various degrees of modulation of hologram fringes. (Photo courtesy Emmett N. Leith and *Scientific American*.

across the hologram plane, which is

$$\mathcal{V} = 2E_{0B}E_{0O}/(E_{0B}^2 + E_{0O}^2) \tag{13.27}$$

contains the appropriate information about the object wave's amplitude.

Once more, in the parlance of communications theory, we might observe that the film plate serves as both the storage device and detector or mixer. It produces, over its surface, a distribution of opaque regions corresponding to a modulated spatial waveform. Accordingly, the third or difference frequency term in Eq. (13.27) is both amplitude and phase modulated by way of the position dependence of $E_{0O}(x, y)$ and $\phi_O(x, y)$.

Figure 13.49*b* is an enlarged view of a portion of the fringe pattern that constitutes the hologram for a simple, essentially two-dimensional, semitransparent object. Were the two interfering waves perfectly planar (as in Fig. 13.49*a*), the evident variations in fringe position and irradiance, which represent the information, would be absent, yielding the traditional Young's pattern (Section 9.3). The sinusoidal transmission-grating configuration (Fig. 13.49*a*) may be thought of as the carrier waveform, which is then modulated by the signal. Furthermore, we can imagine that the coherent superposition of countless zone-plate patterns, one arising from each point on a large object, have metamorphosed into the modulated fringes of Fig. 13.49*b*. When the amount of modulation is further greatly increased, as it would be for a large, three-dimensional, diffusely reflecting object, the fringes lose the kind of symmetry still discernible in Fig. 13.49*b* and become considerably more complicated. Incidentally, holograms are often covered with extraneous swirls and concentric ring systems that arise from diffraction by dust and the like on the optical elements.

The amplitude transmission profile of the processed hologram can be made proportional to $I(x, y)$. In that case, the *final emerging wave*, $E_F(x, y)$, is proportional to the product $I(x, y)E_R(x, y)$, where $E_R(x, y)$ is the *reconstructing wave* incident on the hologram. Thus if the reconstructing wave, of frequency ν, is incident obliquely on Σ_H, as was the background wave, we can write

$$E_R(x, y) = E_{0R} \cos [2\pi\nu t + \phi(x, y)] \tag{13.28}$$

The final wave (except for a multiplicative constant) is the product of Eqs. (13.26) and (13.28):

$$\begin{aligned} E_F(x, y) = {} & \tfrac{1}{2}E_{0R}(E_{0B}^2 + E_{0O}^2) \cos [2\pi\nu t + \phi(x, y)] \\ & + \tfrac{1}{2}E_{0R}E_{0B}E_{0O} \cos (2\pi\nu t + 2\phi - \phi_O) \\ & + \tfrac{1}{2}E_{0R}E_{0B}E_{0O} \cos (2\pi\nu t + \phi_O) \end{aligned} \tag{13.29}$$

Three terms describe the light issuing from the hologram; the

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

first can be rewritten as

$$\tfrac{1}{2}(E_{0B}^2 + E_{0O}^2)E_R(x, y)$$

and is an amplitude-modulated version of the reconstructing wave. In effect, each portion of the hologram functions as a diffraction grating, and this is again the *zeroth-order*, undeflected, direct beam. Since it contains no information about the phase of the object wave, ϕ_O, it is of little concern here.

The next two or *side-band waves* are the sum and difference terms, respectively. These are the two *first-order waves* diffracted by the grating-like hologram. The first of these (i.e., the sum term) represents a wave that, except for a multiplicative constant, has the same amplitude as the object wave $E_{0O}(x, y)$. Moreover, its phase contains a $2\phi(x, y)$ contribution, which, as you recall, arose from tilting the background and reconstructing wavefronts with respect to Σ_H. It's this phase factor that provides the angular separation between the real and virtual images. Furthermore, rather than containing the phase of the object wave, the sum term contains its negative. Thus it's a wave carrying all of the appropriate information about the object but in a way that is not quite right. Indeed, this is the real image formed in converging light in the space beyond the hologram, that is, between it and the viewer. The negative phase is manifest in an inside-out image something like the pseudoscopic effect occurring when the elements of a photographic stereo pair are interchanged. Bumps appear as indentations, and object points that were in front of and nearer to Σ_H are now imaged nearer to but beyond Σ_H. Thus a point on the original subject closest to the observer appears farthest away in the real image. The scene is turned in on itself along one axis in a way that perhaps must be seen to be appreciated.

For example, imagine you are looking down the holographic conjugate image of a bowling alley. The "back" row of pins, even though partially obscured by the "front" rows, are nonetheless imaged closer to the viewer than is the one-pin. Despite this, bear in mind that it's not as if you were looking at the array from behind. No light from the very backs of the pins was ever recorded—you're seeing an inside-out front view. As a consequence, the conjugate image is usually of limited utility, although it can be made to have a normal configuration by forming a second hologram with the real image as the object.

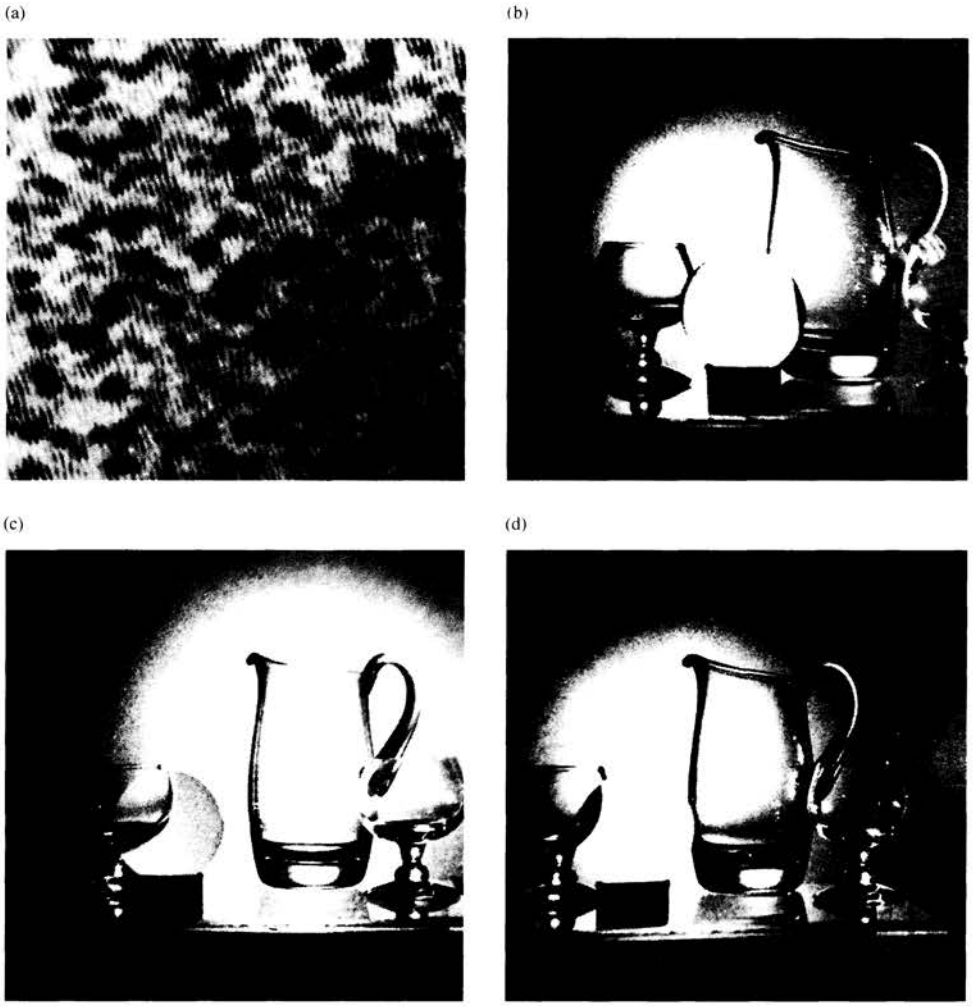
The difference term in Eq. (13.29), except for a multiplicative constant, has precisely the form of the object wave $E_{0O}(x, y)$. If you were to peer into (not at) the illuminated hologram, as if it were a window looking out onto the scene beyond, you would "see" the object exactly as if it were truly sitting there. You could move your head a bit and look around an item in the foreground in order to see the view it had previously been obstructing. In other words, in addition to complete three-dimensionality, parallax effects are apparent as they are in no other reproducing technique (see photo). Imagine that you are viewing the holographic image of a magnifying glass focused on a page of print. As you move your eye with respect to the hologram plane, the words being magnified by the lens (which is itself just an image) actually change, just as they would in "real" life with a "real" lens and "real" print. In the case of an extended scene having considerable depth, your eyes would have to refocus as you viewed different regions of it at various distances. In precisely the same way, a camera lens would have to be readjusted if you were photographing different regions of the virtual image (see photo on page 633).

Holograms display other extremely important and interesting features. For example, if you were standing close to a window, you could obscure all of it with, say, a piece of cardboard, except for a tiny area through which you could then peer and still see the objects beyond. The same is true of a hologram, since each small fragment of it contains information about the entire object, at least as seen from the same vantage point, and each fragment can reproduce, albeit with diminishing resolution, the entire image.

Figure 13.50 summarizes pictorially much of what's been said so far while also providing a convenient setup for actually making and viewing a hologram. Here the photographic emulsion is shown having some depth, as compared with Fig. 13.47, where it was treated as though it were purely two-dimensional. Of course, any emulsion must certainly have a finite thickness. Typically, it would be about 10 μm thick, as compared with the spatial period of the fringes, which might average around 1 μm or so. Figure 13.51*a* is closer to the point, showing the kind of three-dimensional fringes that actually exist throughout the emulsion. For plane waves, these straight parallel fringe-planes are oriented so as to bisect the angle between the reference and object waves. Realize that all the holograms considered up to now have been viewed by looking through them; they're all **transmission holograms**, and in each case they were made by causing the reference wave and the object wave to traverse the film from the same side.

Something similar happens when the reference and object

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



Parts (*b*) through (*d*) are three different views photographed from the same holographic image generated by the hologram in (*a*). (Photos from Smith, *Principles of Holography*.)

waves traverse the emulsion from opposite sides, as in Fig. 13.51*b*. If for simplicity we again let both waves be planar, the resulting pattern can be visualized by sliding two pencils along with the fronts; it should then be clear that the fringes are straight bands (planes) lying parallel to the face of the film plate. When an actual, highly contorted, object wave is made to overlap a planar, coherent, reference wave, these fringes become modulated with the information describing the object. The corresponding three-dimensional diffraction grating is called a **reflection hologram**. During playback it scatters the reilluminating beam back out toward the viewer, and one sees a virtual image behind the hologram (as if looking into a mirror).

The zone-plate interpretation has been applicable to the various holographic schemes we've considered thus far, and this regardless of whether the diffracted wave was of the *near*- or *far*-field variety (i.e., whether we had Fresnel or Fraunhofer holograms, respectively). Indeed, it applies generally where the interferogram results from the superpositioning of the scattered spherical wavelets from each object point and a coherent plane or even spherical reference wave (provided the latter's curvature is different from that of the wavelets). An inherent

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

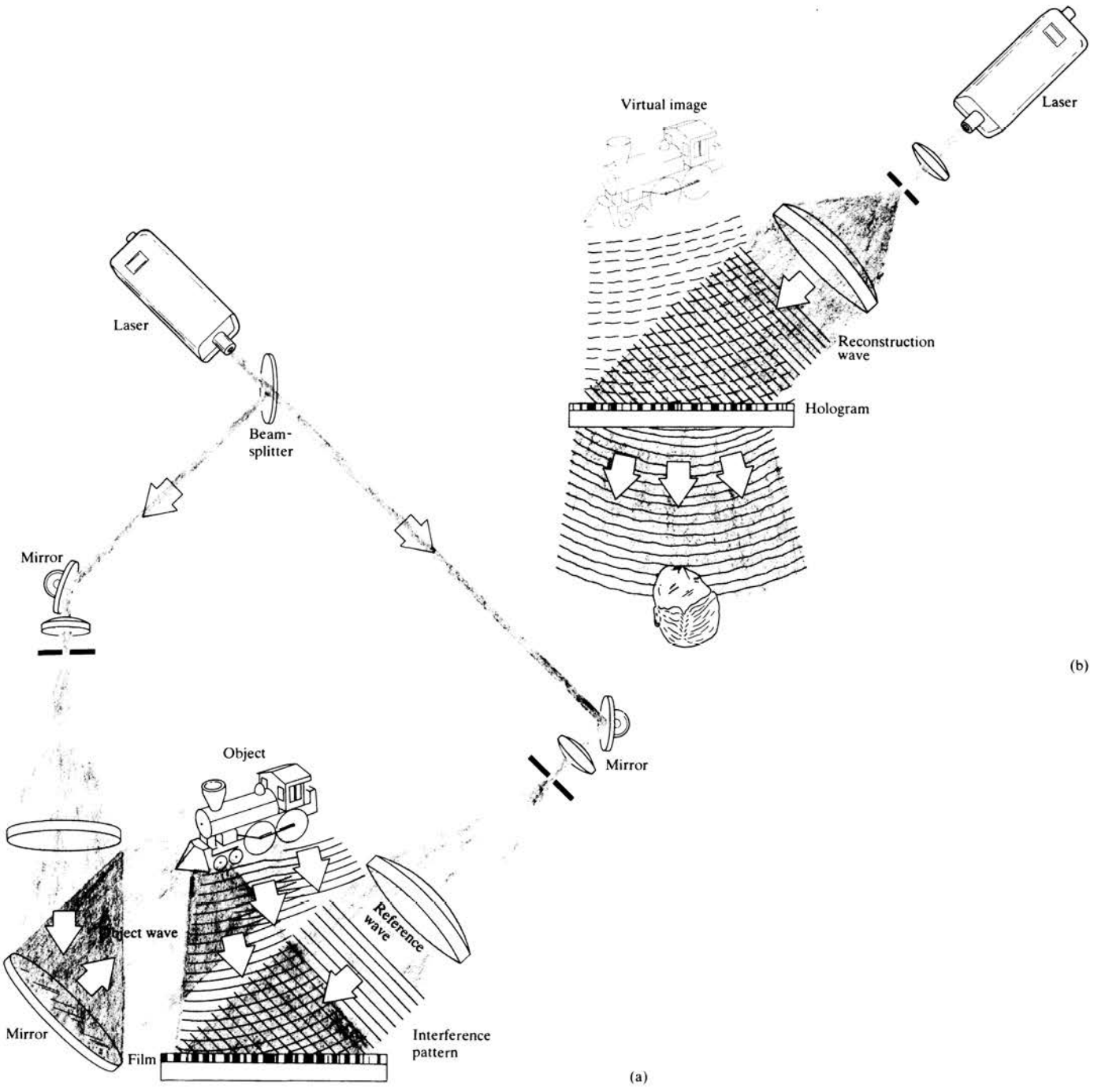
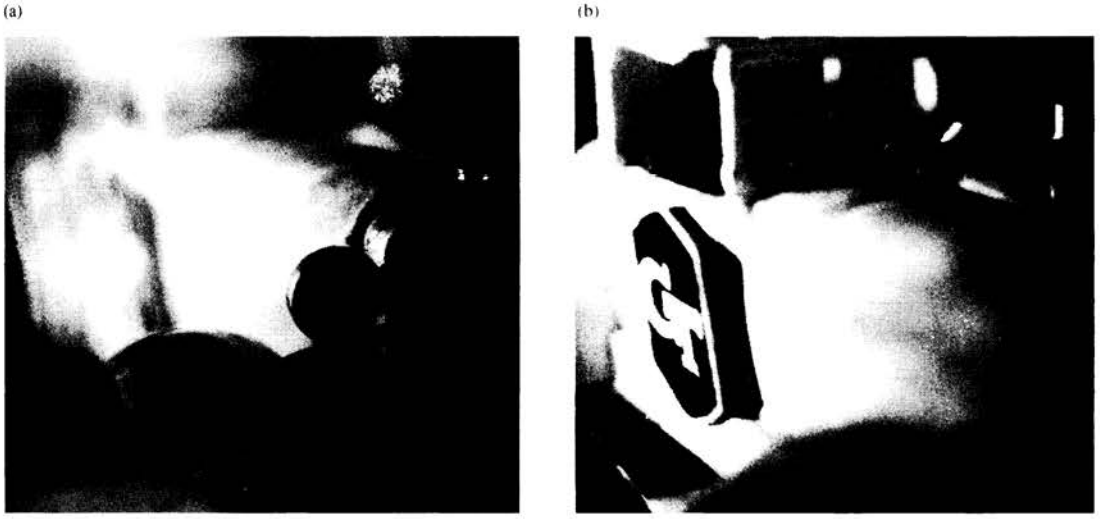


Figure 13.50 (*a*) The creation of a transmission hologram of a toy locomotive. (*b*) Replay of a transmission hologram.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



A reconstructed holographic image of a model automobile. The camera position and plane of focus were changed between (*a*) and (*b*). (Photos from O'Shea, Callen, and Rhodes, *An Introduction to Lasers and Their Applications.*)

problem, which these schemes therefore have in common, arises from the fact that the zone-plate radii, R_m, vary as $m^{1/2}$ from Eq. (10.91). Thus the zone fringes are more densely packed farther from the center of each zone lens (i.e., at larger values of m). This is tantamount to an increasing spatial frequency of bright and dark rings, which must be recorded by the photographic plate. The same thing can be appreciated in the cosine-grating representation, where the spatial frequency increases with θ. Since film, no matter how fine-grained, is limited in its spatial frequency response, there will be a cutoff beyond which it cannot record data. All of this represents a built-in limitation on resolution. In contrast, if the mean fre-

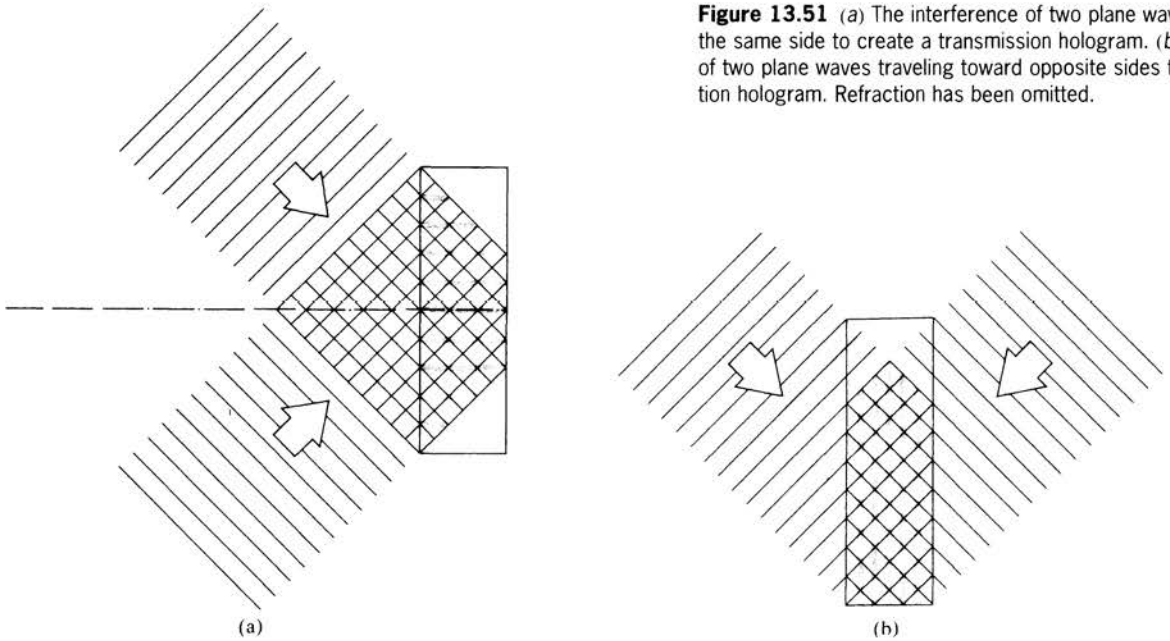


Figure 13.51 (*a*) The interference of two plane waves traveling toward the same side to create a transmission hologram. (*b*) The interference of two plane waves traveling toward opposite sides to create a reflection hologram. Refraction has been omitted.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

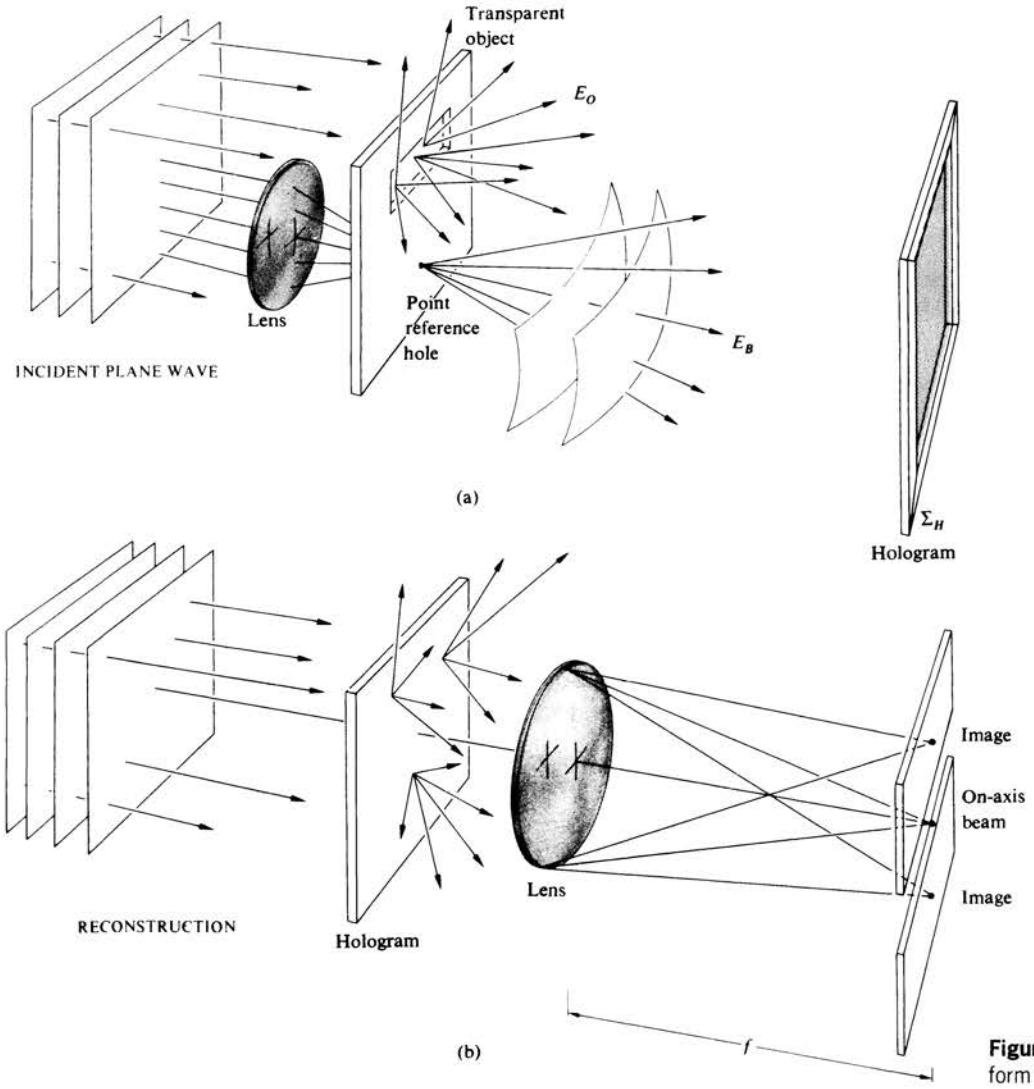


Figure 13.52 Lensless Fourier-transform holography (a transparent object).

quency of the fringes could be made constant, the limitations imposed by the photographic medium would be considerably reduced, and the resolution correspondingly increased. As long as it could record the average spatial fringe frequency, even a coarse emulsion, such as Polaroid P/N, could be used without extensive loss of resolution. Figure 13.52 shows an arrangement that accomplishes just this by having the diffracted object wavelets interfere with a spherical reference wave of about the same curvature. The resulting interferogram is known as a **Fourier-transform** hologram (in this specific instance, it's of the high-resolution *lensless* variety). This scheme is designed to have the reference wave cancel the quadratic (zone-lens type) dependence of the phase with position on Σ_H. But that will occur precisely only for a planar two-dimensional object. In the case of a three-dimensional object (Fig. 13.53) this only happens over one plane, and the resulting hologram is therefore a composite of both types, that is, a zone lens and Fourier transform. Unlike the other arrange-

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

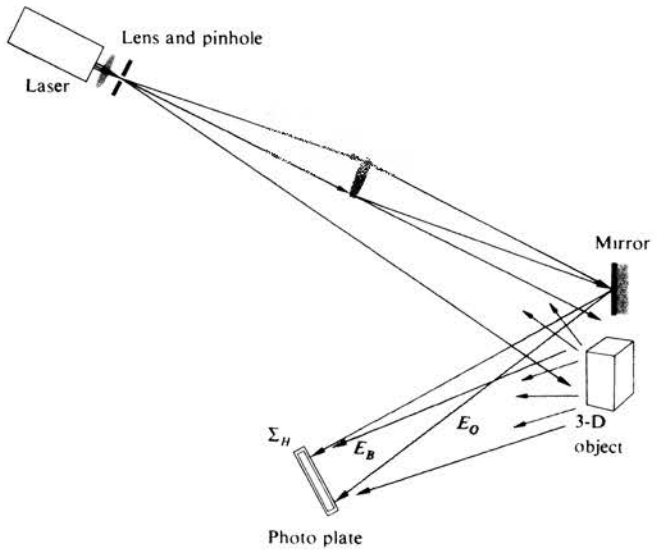
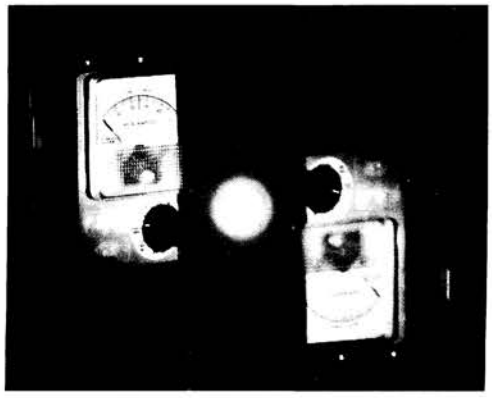


Figure 13.53 Lensless Fourier-transform holography (an opaque object).



A reconstruction of a Fourier-transform hologram. [From G. W. Stroke, D. Brumm, and A. Funkhauser, *J. Opt. Soc. Am.* **55**, 1327 (1965).]

ments, both images generated by a Fourier-transform hologram are virtual, in the same plane, and oriented as if reflected through the origin (see photo).

The grating-like nature of all previous holograms is evident here as well. In fact, if you look through a Fourier-transform hologram at a small white-light source (a flashlight in a dark room works beautifully), you see the two mirror images, but they are extremely vague and surrounded by bands of spectral colors. The similarity with white light that has passed through a grating is unmistakable.*

13.3.2 Developments and Applications

For years holography was an invention in search of application, that notwithstanding certain obvious possibilities, such as the all too inevitable 3-D billboard. Fortunately, several significant technological developments have in recent times begun what will surely be an ongoing extension of the scope and utility of holography. The early efforts in the field were typified by countless images of toy cars and trains, chess pieces and statuettes—small objects resting on giant blocks of granite. They had to be small because of limited laser power and coherence length, while the ever-present massive granite platform served to isolate the slightest vibrations that might blur the fringes and thereby degrade or obliterate the stored data. A loud sound or gust of air could result in deterioration of the reconstructed image by causing the photographic plate, object, or mirrors to shift several millionths of an inch during the exposure, which itself might last on the order of a minute or so. That was the still-life era of holography. But now, with the use of new, more sensitive films and the short duration ($\approx$40 ns) high-power light flashes from a single-mode pulsed ruby laser, even portraiture and stop-action holography have become a reality† (see photo on page 636).

Throughout the 1960s and much of the 1970s, the emphasis in the field was on the obvious visual wonders of holography. This continued in the 1980s with the mass production of over a hundred million inexpensive plastic reflection holograms (bonded to credit cards; tucked in candy packages; decorating magazine covers, jewelry, and record albums). The development of a photopolymer that is stable, cheap, and able to produce high-quality images has stimulated the manufacture of even more of these throwaway holograms. Still there is

*See DeVelis and Reynolds, *Theory and Applications of Holography*; Stroke, *An Introduction to Coherent Optics and Holography*; Goodman, *Introduction to Fourier Optics*; Smith, *Principles of Holography*; or perhaps *The Engineering Uses of Holography*, edited by E. R. Robertson and J. M. Harvey.

†L. D. Siebert, *Appl. Phys. Letters* **11**, 326 (1967), and R. G. Zech and L. D. Siebert, *Appl. Phys. Letters* **13**, 417 (1968).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



A reconstruction of a holographic portrait. (Photo courtesy L. D. Siebert.)

now a widespread recognition of the potential of holography as a nonpictorial instrumentality, and that new direction is finding increasingly important applications.

Volume Holograms

Yuri Nikolayevitch Denisyuk of the Soviet Union, in 1962, introduced a scheme for generating holograms that was conceptually similar to the early (1891) color photographic process of Gabriel Lippmann. In brief, the object wave is reflected from the subject and propagates backward, overlapping the incoming coherent background wave. In so doing, the two waves set up a three-dimensional pattern of standing waves, as in Fig. 13.51. The spatial distribution of fringes is recorded by the photoemulsion throughout its entire thickness to form what has become known as a **volume hologram**. Several variations have since been introduced, but the basic ideas are the same; rather than generating a two-dimensional grating-like scattering structure, the volume hologram is a three-dimensional grating. In other words, it's a three-dimensional, modulated, periodic array of phase or amplitude objects, which represent the data. It can be recorded in several media, for example, in thick photoemulsions wherein the amplitude objects are grains of deposited silver; in photochromic glass; with halogen crystals, such as KBr, which respond to irradiation via color-center variations; or with a ferroelectric crystal, such as lithium niobate, which undergoes local alterations in its index of refraction, thus forming what might be called a phase volume hologram. In any event, one is left with a volume array of data, however stored in the medium, which in the reconstruction process behaves very much like a crystal being irradiated by X-rays. It scatters the incident (reconstructing) wave according to Bragg's Law (p. 485). This isn't very surprising, since both the scattering centers and λ have simply been scaled up proportionately.

One important feature of volume holograms is the interdependence [via Bragg's Law, $2d \sin \theta = m\lambda$, Eq. (10.71)] of the wavelength and the scattering angle; that is, only a given color light will be diffracted at a particular angle by the hologram. Another significant property is that by successively altering the incident angle (or the wavelength), a single-volume medium can store a great many coexisting holograms at one time. This latter property makes such systems extremely appealing as densely packed memory devices. For example, an 8-mm-thick hologram has been used to store 550 pages of information, each individually retrievable. In theory, a single lithium niobate crystal is capable of easily storing thousands of holograms, and any one of them could be replayed by addressing the crystal with a laserbeam at the appropriate angle. Current research is also focusing on potassium tantalate niobate (KTN) as a potential photorefractive crystal-storage medium. Imagine a 3-D holographic motion picture; a library; or everyone's vital statistics—beauty marks, credit cards, taxes, bad habits, income, life history, and so on, all recorded on a handful of small transparent crystals.

Multicolored reconstructions have been formed using (black and white) volume holographic plates. Two, three, or more different colored and mutually incoherent overlapping laserbeams are used to generate separate, cohabitating, component holograms of the object, and this can be done one at a time or all at once. When these are illuminated simultaneously by the various constituent beams, a multicolored image results.

Another important and highly promising scheme, devised by G. W. Stroke and A. E. Labeyrie, is known as **white-light reflection holography**. Here, the reconstructing wave is an ordinary white-light beam from, say, a flashlight or projector, having a wavefront similar to the original quasimonochromatic background wave. When illuminated on the same side as the viewer, only the specific wavelength that enters the volume hologram at the proper Bragg angle is reflected off to form a reconstructed 3-D virtual image. Thus, if the scene were recorded in red laserlight, only red light would presumably be reflected as an image. It is of pedagogical interest to point out,

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

however, that the emulsion may shrink during the fixing process, and if it is not swollen back to its original form chemically (with say triethylnolamine), the spacing of the Bragg planes, d, decreases. That means that at a given angle θ, the reflected wavelength will decrease proportionately. Hence, a scene recorded in He–Ne red might play back in orange or even green when reconstructed by a beam of white light.

If several overlapping holograms corresponding to different wavelengths are stored, a multicolored image will result. The advantages of using an ordinary source of white light to reconstruct full-color 3-D images are obvious and far-reaching.

Holographic Interferometry

One of the most innovative and practical of recent holographic advances is in the area of interferometry. Three distinctive approaches have proved to be quite useful in a wealth of nondestructive testing situations where, for example, one might wish to study microinch distortions in an object resulting from strain, vibration, heat, and so on. In the *double exposure* technique, one simply makes a hologram of the undisturbed object and then, before processing, exposes the hologram for a second time to the light coming from the now distorted object. The ultimate result is two overlapping reconstructed waves, which proceed to form a fringe pattern indicative of the displacements suffered by the object, that is, the changes in optical path length (see photo). Variations in index such as those arising in wind tunnels and the like will generate the same sort of pattern.

In the *real-time* method, the subject is left in its original position throughout; a processed hologram is formed, and the resulting virtual image is made to overlap the object precisely (Fig. 13.54). Any distortions that arise during subsequent testing show up, on looking through the hologram, as a system of fringes, which can be studied as they evolve in real time. The method applies to both opaque and transparent objects. Motion pictures can be taken to form a continuous record of the response.

The third method is the *time-average* approach and is particularly applicable to rapid, small-amplitude, oscillatory systems. Here the film plate is exposed for a relatively long duration, during which time the vibrating object has executed a number of oscillations. The resulting hologram can be thought of as a superposition of a multiplicity of images, with the effect that a standing-wave pattern emerges. Bright areas reveal undeflected or stationary nodal regions, while contour lines trace out areas of constant vibrational amplitude.



Double exposure holographic interferogram. [From S. M. Zivi and G. H. Humberstone, "Chest Motion Visualized by Holographic Interferometry," *Medical Research Eng.* p. 5 (June 1970).] Compare this with the radar photo on page 437.

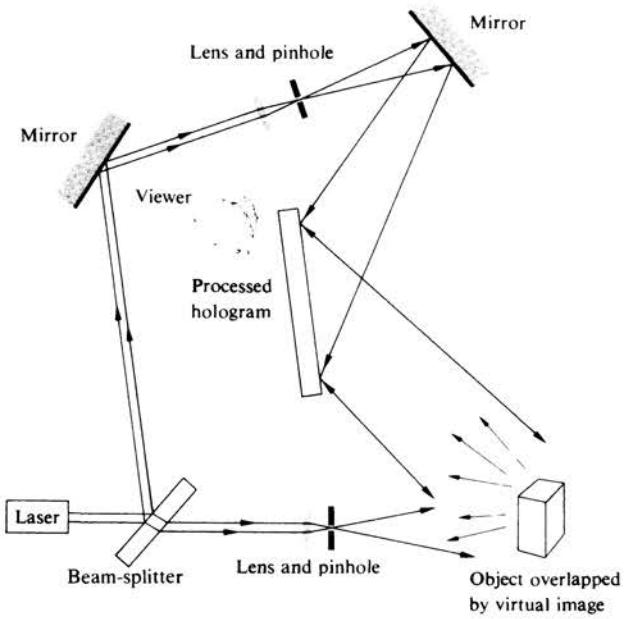


Figure 13.54 Real-time holographic interferometry.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Today holographic testing of mechanical systems is a well-established practice in industry. It continues to serve in a broad range of applications, from noise reduction in automobile transmissions to routine jet engine inspections.

Acoustical Holography

In acoustical holography, an ultra-high-frequency sound wave (ultrasound) is used to create the hologram initially, and a laserbeam then serves to form a recognizable reconstructed image. In one application, the stationary ripple pattern on the surface of a water body produced by submerged coherent transducers corresponds to a hologram of the object beneath (Fig. 13.55). Photographing it creates a hologram that can be illuminated optically to form a visual image. Alternatively, the ripples can be irradiated from above with a laserbeam to produce an instantaneous reconstruction in reflected light.

The advantages of acoustical techniques reside in the fact that sound waves can propagate considerable distances in dense liquids and solids where light cannot. Thus acoustical holograms can record such diverse things as underwater submarines and internal body organs.* In the case of Fig. 13.55, one would see something that resembled an X-ray motion picture of the fish. The accompanying photo is the image of a penny formed via acoustical holography using ultrasound at a frequency of 48 MHz. In water, that corresponds to a wavelength of roughly 30 μm, and so each fringe contour reveals a change in elevation of $\frac{1}{2}\lambda$ or 15 μm.

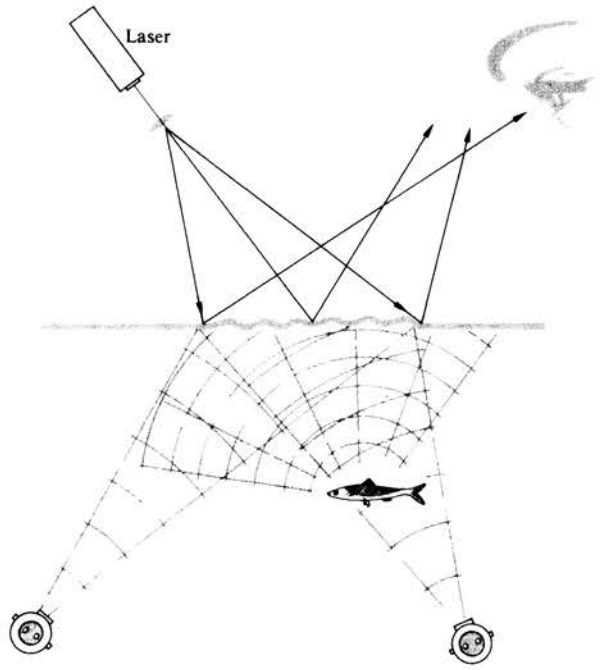


Figure 13.55 Acoustical holography.



Interferometric image of a penny via acoustical holography. (Photo courtesy Holosonics, Inc.)

Holographic Optical Elements

Evidently, when two plane waves overlap, as in Fig. 13.47, they produce a cosine grating. This suggests the rather obvious notion that holography can be used for nonpictorial purposes, like making diffraction gratings. Indeed, the *holographic optical element* (HOE) is any diffractive device consisting of a "fringe" system (i.e., a distribution of diffracting amplitude or phase objects) created either directly by interferometry or by computer simulation thereof. Holographic diffraction gratings, both blazed and sinusoidal, are available commercially (with up to around 3600 lines/mm). Although generally less efficient than ruled gratings, they do produce far less stray light, which can be important in many applications.

Suppose we record the interference pattern of a converging

*See A. F. Metherell, "Acoustical Holography," *Sci. Am.* **36**, 221, (October 1969). Refer to A. L. Dalisa et al., "Photoanodic Engraving of Holograms on Silicon," *Appl. Phys. Letters* **17**, 208 (1970), for another interesting use of surface relief patterns.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

beam using a planar reference wave. Upon reilluminating the resulting transmission hologram with a matching plane wave, out will come a recreated converging wave—the hologram will function like a lens (see Fig. 13.44). Similarly, if the reference beam is a diverging wave from a point source and the object is a plane wave, the resulting hologram, reilluminated by the point source, will play back a plane wave. In this way a holographic optical element can perform the tasks of a complex lens with the added benefit of allowing for an inexpensive, lightweight, compact system design.

Holographic optical elements are already in use inside supermarket check-out scanners that automatically read the bar patterns of the Universal Product Code (UPC) on merchandise. A laserbeam passes through a rotating disk composed of a number of holographic lens-prism facets. These rapidly refocus, shift, and scan the beam across a volume of space, ensuring that the code will be read on the first pass across the device. HOEs are used in so-called heads-up displays in airplane cockpits. These allow reflected data to appear on an otherwise transparent screen in front of the pilot's face and yet not obscure the view. They're also in office copy machines and solar concentrators.

As *matched spatial filters*, HOEs are used in optical correlators to spot defects in semiconductors and tanks in reconnaissance pictures. In such cases the HOE is a hologram formed using the Fourier transform of the target (e.g., a picture of a tank or perhaps a printed word) as the object. Suppose the problem is to find a word on a printed page automatically, using an optical computer like that in Fig. 13.30, that is, to cross-correlate the word and the page of words. The target-transform hologram is placed in the transform plane and illuminated with the transform of an entire page of print. The field amplitude emerging from this HOE-filter will then be proportional to the product of the transforms of the page and the word. The transform of this product, generated by the last lens and displayed on the image plane, is the desired cross-correlation (recall the Wiener–Khintchine theorem). If the word is on the page, there will be a high correlation, and a bright spot of light will appear superimposed in the final image everywhere the target word occurs.*

It is possible to synthesize, point by point, a hologram of a fictitious object. In other words, in the most direct approach holograms can be produced by calculating, with a digital computer, the irradiance distribution that would arise were some object appropriately illuminated in a hypothetical recording session. A computer-controlled plotter drawing or cathode ray tube read out of the interferogram is then photographed, thence to serve as the actual hologram. The result upon illumination is a three-dimensional reconstructed image of an object that never had any real existence in the first place. More practically, computer-generated HOEs are now routinely being produced, often to serve as references for optical testing. Since this mating of technologies can in principle generate wavefronts otherwise essentially impossible to produce, the future is very promising.

*See A. Ghatak and K. Thyagarajan, *Contemporary Optics*, p. 214.

13.4 Nonlinear Optics

Generally, the domain of *nonlinear optics* is understood to encompass those phenomena for which electric and magnetic field intensities of higher powers than the first play a dominant role. The Kerr Effect (Section 8.11.3), which is a quadratic variation of refractive index with applied voltage, and thereby electric field, is typical of several long-known nonlinear effects.

The usual classical treatment of the propagation of light—superposition, reflection, refraction, and so forth—assumes a linear relationship between the electromagnetic light field and the responding atomic system constituting the medium. But just as an oscillatory mechanical device (e.g., a weighted spring) can be overdriven into nonlinear response through the application of large enough forces, so too we might anticipate that an extremely intense beam of light could generate appreciable nonlinear optical effects.

The electric fields associated with light beams from ordinary or, if you will, traditional sources are far too small for such behavior to be easily observable. It was for this reason, coupled with an initial lack of technical prowess, that the subject had to await the advent of the laser in order that sufficient brute force could be brought to bear in the optical region of the spectrum. As an example of the kinds of fields readily obtainable with the current technology, consider that a good lens can focus a laserbeam down to a spot having a diameter of about 10^{-3} inch or so, which corresponds to an area of roughly 10^{-9} m^2. A 200-megawatt pulse from, say, a Q-switched ruby laser would then produce a flux density of 20×10^{16} W/m^2. It follows (Problem 13.27) that the corresponding electric field

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

amplitude is given by

$$E_0 = 27.4\left(\frac{I}{n}\right)^{1/2} \tag{13.30}$$

In this particular case, for $n \approx 1$, the field amplitude is about 1.2×10^8 V/m. This is more than enough to cause the breakdown of air (roughly 3×10^6 V/m) and just several orders of magnitude less than the typical fields holding a crystal together, the latter being roughly about the same as the cohesive field on the electron in a hydrogen atom (5×10^{11} V/m). The availability of these and even greater (10^{12} V/m) fields has made possible a wide range of important new nonlinear phenomena and devices. We shall limit this discussion in the consideration of several nonlinear phenomena associated with passive media (i.e., media that act essentially as catalysts without making their own characteristic frequencies evident). Specifically, we'll consider optical rectification, optical harmonic generation, frequency mixing, and self-focusing of light. In contrast, Stimulated Raman, Rayleigh, and Brillouin Scattering exemplify nonlinear optical phenomena arising in active media that do impose their characteristic frequencies on the lightwave.*

As you may recall, the electromagnetic field of a lightwave propagating through a medium exerts forces on the loosely bound outer or valence electrons. Ordinarily, these forces are quite small, and in a linear isotropic medium the resulting electric polarization is parallel with and directly proportional to the applied field. In effect, the polarization follows the field; if the latter is harmonic, the former will be harmonic as well. Consequently, one can write

$$P = \epsilon_0 \chi E \tag{13.31}$$

where χ is a dimensionless constant known as the electric susceptibility, and a plot of P versus E is a straight line. Quite obviously in the extreme case of very high fields, we can expect that P will become saturated; in other words, it simply cannot increase linearly indefinitely with E (just as in the familiar case of ferromagnetic materials, where the magnetic moment becomes saturated at fairly low values of H). Thus we can anticipate a gradual increase of the ever-present, but usually insignificant, nonlinearity as E increases. Since the directions of $\vec{\mathbf{P}}$ and $\vec{\mathbf{E}}$ coincide in the simplest case of an isotropic medium, we can express the polarization more effectively as a series expansion:

$$P = \epsilon_0(\chi E + \chi_2 E^2 + \chi_3 E^3 + \cdots) \tag{13.32}$$

The usual linear susceptibility, χ, is much greater than the coefficients of the nonlinear terms χ_2, χ_3, and so on, and hence the latter contribute noticeably only at high-amplitude fields. Now suppose that a lightwave of the form

$$E = E_0 \sin \omega t$$

is incident on the medium. The resulting electric polarization

$$P = \epsilon_0 \chi E_0 \sin \omega t + \epsilon_0 \chi_2 E_0^2 \sin^2 \omega t + \epsilon_0 \chi_3 E_0^3 \sin^3 \omega t + \cdots \tag{13.33}$$

can be rewritten as

$$P = \epsilon_0 \chi E_0 \sin \omega t + \frac{\epsilon_0 \chi_2}{2} E_0^2 (1 - \cos 2\omega t) + \frac{\epsilon_0 \chi_3}{4} E_0^3 (3 \sin \omega t - \sin 3\omega t) + \cdots \tag{13.34}$$

As the harmonic lightwave sweeps through the medium, it creates what might be thought of as a polarization wave, that is, an undulating redistribution of charge within the material in response to the field. If only the linear term were effective, the electric polarization wave would correspond to an oscillatory current following along with the incident light. The light thereafter reradiated in such a process would be the usual refracted wave generally propagating with a reduced speed v and having the same frequency as the incident light. In contrast, the presence of higher-order terms in Eq. (13.33) implies that the polarization wave does not have the same harmonic profile as the incident field. In fact, Eq. (13.34) can be likened to a Fourier series representation of the distorted profile of $P(t)$.

13.4.1 Optical Rectification

The second term in Eq. (13.34) has two components of great interest. First there is a *dc* or *constant bias polarization* varying as E_0^2. Consequently, if an intense plane-polarized beam traverses an appropriate (piezoelectric) crystal, the presence of the quadratic nonlinearity will, in part, be manifest by a constant electric polarization of the medium. A voltage difference, proportional to the beam's flux density, will accordingly

*For a more extensive treatment than is possible here, see N. Bloembergen, *Nonlinear Optics*, or G. C. Baldwin, *An Introduction to Nonlinear Optics*.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

appear across the crystal. This effect, in analogy to its radiofrequency counterpart, is known as **optical rectification**.

13.4.2 Harmonic Generation

The cos $2\omega t$ term [Eq. (13.34)] corresponds to a variation in electric polarization at twice the fundamental frequency (i.e., at twice that of the incident wave). The reradiated light that arises from the driven oscillators also has a component at this same frequency, 2ω, and the process is spoken of as **second-harmonic generation**, or SHG for short. In terms of the photon representation, we can envision two identical photons of energy $\hbar\omega$ coalescing within the medium to form a single photon of energy $\hbar 2\omega$. Peter A. Franken and several coworkers at the University of Michigan in 1961 were the first to observe SHG experimentally. They focused a 3-kW pulse of red (694.3 nm) ruby laserlight onto a quartz crystal. Just about one part in 10^8 of this incident wave was converted to the 347.15-nm ultraviolet second harmonic.

Notice that, for a given material, if $P(E)$ is an odd function, that is, if reversing the direction of the $\vec{\mathbf{E}}$-field simply reverses the direction of $\vec{\mathbf{P}}$, the even powers of E in Eq. (13.32) must vanish. But this is just what happens in an isotropic medium, such as glass or water—there are no special directions in a liquid. Moreover, in crystals like calcite, which are so structured as to have what's known as a *center of symmetry* or an *inversion center*, a reversal of all of the coordinate axes must leave the interrelationships between physical quantities unaltered. Thus no even harmonics can be produced by materials of this sort. Third-harmonic generation (THG), however, can exist and has been observed in several materials (p. 598) including calcite. The requirement for SHG that a crystal not have inversion symmetry is also necessary for it to be piezoelectric. Under pressure a piezoelectric crystal [such as quartz, potassium dihydrogen phosphate (KDP), or ammonium dihydrogen phosphate (ADP)] undergoes an asymmetric distortion of its charge distribution, thus producing a voltage. Of the 32 crystal classes, 20 are of this kind and may therefore be useful in SHG. The simple scalar expression [Eq. (13.32)] is actually not an adequate description of a typical dielectric crystal. Things are a good deal more complicated, because the field components in several different directions in a crystal can affect the electric polarization in any one direction. A complete treatment requires that $\vec{\mathbf{P}}$ and $\vec{\mathbf{E}}$ be related not by a single scalar but by a group of quantities arranged in the particular form of a tensor, namely, the susceptibility tensor.*

A major difficulty in generating copious amounts of second-harmonic light arises from the frequency dependence of the refractive index, that is, dispersion. At some initial point where the incident or ω-wave, generates the second-harmonic or 2ω-wave, the two are coherent. As the ω-wave propagates through the crystal, it continues to generate additional contributions of second-harmonic light, which all combine totally constructively only if they maintain a proper phase relationship. Yet the ω-wave travels at a phase velocity v_ω, which is ordinarily different from the phase velocity, $v_{2\omega}$, of the 2ω-wave. Thus the newly emitted second harmonic periodically falls out-of-phase with some of the previously generated 2ω-waves. When the irradiance of the second harmonic, $I_{2\omega}$, emerging from a plate of thickness ℓ is computed,† it turns out to be

$$I_{2\omega} \propto \frac{\sin^2 [2\pi(n_\omega - n_{2\omega})\ell/\lambda_0]}{(n_\omega - n_{2\omega})^2} \tag{13.35}$$

(see Fig. 13.56). This yields the result that $I_{2\omega}$ has its maximum value when $\ell = \ell_c$, where

$$\ell_c = \frac{1}{4}\frac{\lambda_0}{|n_\omega - n_{2\omega}|} \tag{13.36}$$

This is commonly known as the *coherence length* (although a different name would be better), and it's usually of the order of only about $20\lambda_0$. Despite this, efficient SHG can be accomplished by a procedure known as *index matching*, which negates the undesirable effects of dispersion; in short, one arranges things so that $n_\omega = n_{2\omega}$. A commonly used SHG material is KDP. It is piezoelectric, transparent, and also negatively uniaxially birefringent. Furthermore, it has the interesting property that if the fundamental light is a linear polarized *ordinary wave*, the resulting second harmonic will be an *extraordinary wave*. As can be seen from Fig. 13.57, if light propagates within a KDP crystal at the specific angle θ_0 with

*Incidentally, there is nothing extraordinary about this kind of behavior—it comes up all the time. There are inertia tensors, demagnetization coefficient tensors, stress tensors, and so forth.

†See, for example, B. Lengyel, *Introduction to Laser Physics*, Chapter VII. This is a fine elementary treatment.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

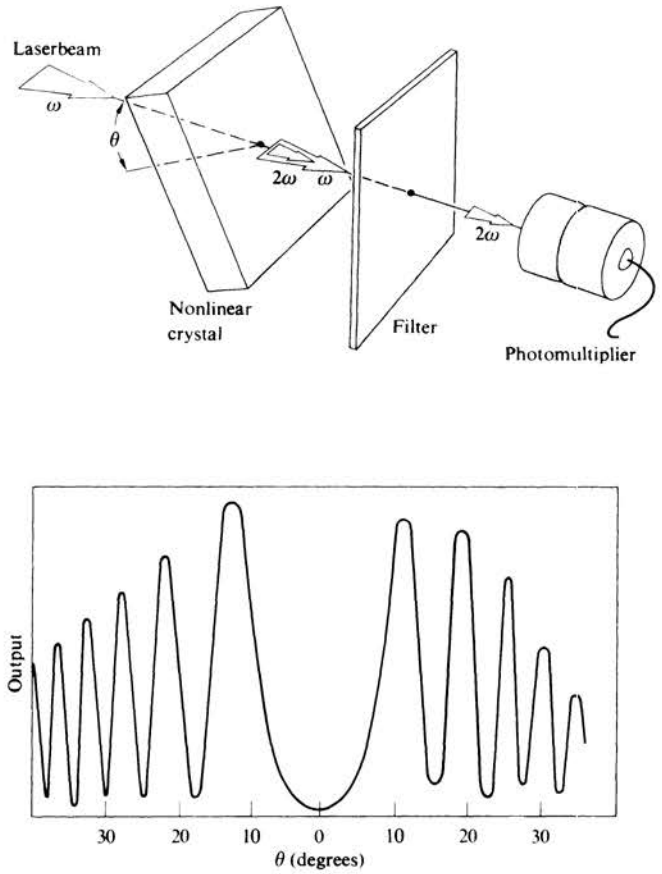


Figure 13.56 Second-harmonic generation as a function of θ for a 0.78-mm thick quartz plate. Peaks occur when the effective thickness is an even multiple of ℓ_c. [From P. D. Maker, R. W. Terhune, M. Nisenoff, and C. M. Savage, *Phys. Rev. Letters* **8**, 21 (1962).]

respect to the optic axis, the index, $n_{0\omega}$, of the ordinary fundamental wave will precisely equal the index of the extraordinary second harmonic $n_{e2\omega}$. The second-harmonic wavelets will then interfere constructively, thereupon increasing the conversion efficiency by several orders of magnitude. Second-harmonic generators, which are simply appropriately cut and oriented crystals, are available commercially, but do keep in mind that θ_0 is a function of λ, and each such device performs at one frequency. Not long ago, a continuous 1-W second-harmonic beam at 532.3 nm was obtained by placing a barium sodium niobate crystal within the cavity of a 1-W 1.06μ laser. The fact that the ω-wave sweeps back and forth through the crystal increases the net conversion efficiency.

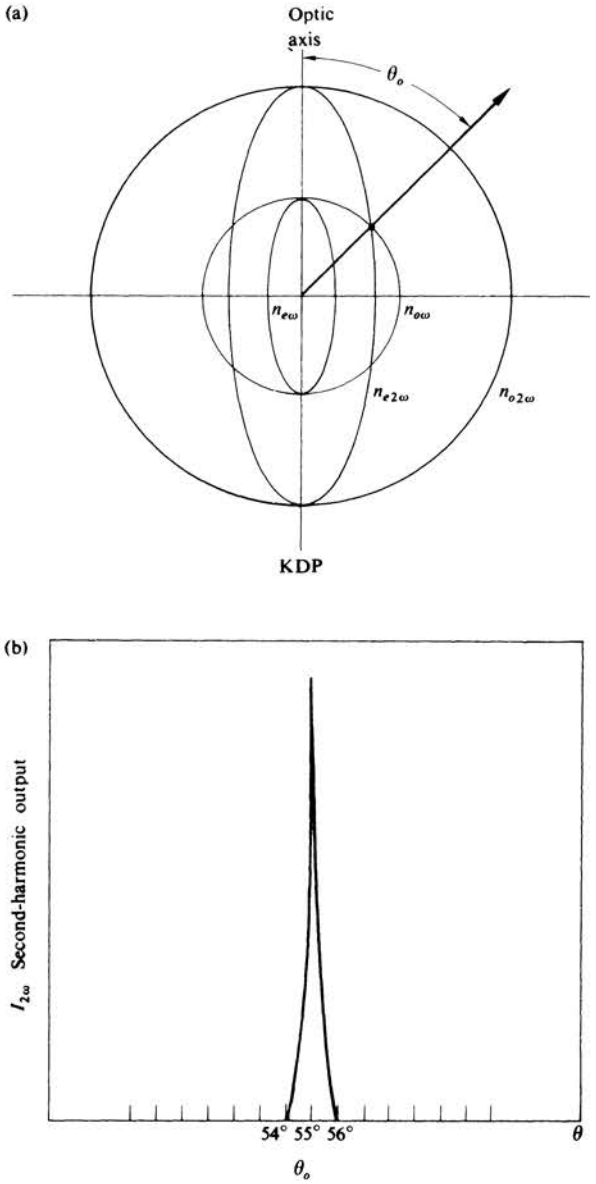
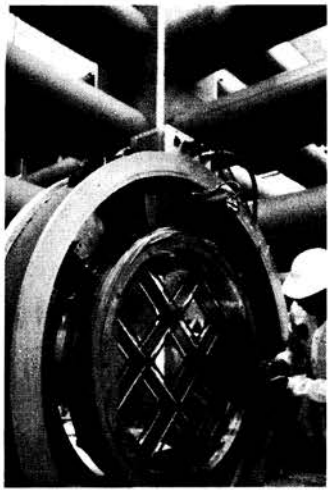


Figure 13.57 Refractive index surface for KDP. (*b*) $I_{2\omega}$ versus crystal orientation in KDP. (From Maker et al.)

Optical harmonic generation soon lost its initial exotic quality and became a routine commercial process by the early 1980s. Still, there were exciting technical accomplishments, such as the 74-cm–diameter harmonic conversion array (see photo) built for the Nova laser-fusion program, which led to the Omega's frequency-tripling system a decade later (p. 598).

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027



The KDP frequency converter for the Nova laser. (Photo courtesy Lawrence Livermore National Laboratory.)

Its function was to convert upwards of 80% of the infrared (1.05 μm) emission from the neodymium–glass laser into more efficient high-frequency radiation. Because of its great size, the converter was an aligned mosaic of smaller KDP single-crystal panels forming two layers, one behind the other. To generate the second harmonic (green light at 0.53 μm), the array is positioned so that each layer functions independently to produce two overlapping frequency-shifted components. These arise one from each crystal layer and are orthogonally polarized. The third harmonic (blue light at 0.35 μm) is created by reorienting the assembly to the appropriate phase-matching angle so as to shift about two-thirds of the beam energy into the second harmonic as it traverses the first crystal layer. The second layer mixes the remaining IR and the second-harmonic green light to produce third-harmonic blue.

13.4.3 Frequency Mixing

Another situation of considerable practical interest involves the *mixing* of two or more primary beams of different frequencies within a nonlinear dielectric. The process can most easily be appreciated by substituting a wave of the form

$$E = E_{01} \sin \omega_1 t + E_{02} \sin \omega_2 t \qquad (13.37)$$

into the simplest expression for P given by Eq. (13.32). The second-order contribution is then

$$\epsilon_0 \chi_2 (E_{01}^2 \sin^2 \omega_1 t + E_{02}^2 \sin^2 \omega_2 t + 2E_{01}E_{02} \sin \omega_1 t \sin \omega_2 t)$$

The first two terms can be expressed as functions of $2\omega_1$ and $2\omega_2$, respectively, while the last quantity gives rise to sum and difference terms, $\omega_1 + \omega_2$ and $\omega_1 - \omega_2$.

As for the quantum picture, the photon of frequency $\omega_1 + \omega_2$ simply corresponds to a coalescing of the two original photons into a new photon, just as it did in the case of SHG, where both quanta had the same frequency. The energy and momentum of the annihilated photons are carried off by the created sum photon. The generation of an $\omega_1 - \omega_2$ difference-photon is a little more involved. Conservation of energy and momentum requires that on interacting with an ω_2-photon, only the higher-frequency ω_1-photon vanishes, thereby creating two new quanta, one an ω_2-photon and the other a difference-photon.

As an application of this phenomenon, suppose we beat, within a nonlinear crystal a strong wave of frequency ω_p, called the *pump light*, with a weak *signal wave* of lower frequency ω_s, which is to be amplified. Pump light is thereby converted into both signal light and a difference wave, called *idler light*, of frequency $\omega_i = \omega_p - \omega_s$. If the idler light is then made to beat with the pump light, the latter is converted into additional amounts of idler and signal light. In this way both the signal and idler waves are amplified. This is actually an extension into the optical-frequency region of the well-known concept of *parametric amplification*, whose use in the microwave spectrum dates back to the late 1940s. The first *optical-parametric oscillator*, which was operated in 1965, is depicted in Fig. 13.58. The flat parallel end faces of a nonlinear crystal (lithium niobate) were coated to form an optical Fabry–Perot cavity. The signal and idler frequencies (both about 1000 nm) corresponded to two of the resonant frequencies of the cavity. When the flux density of the pumping light was high enough, energy was transferred from it into the signal and idler oscillatory modes, with the consequent buildup of those modes and emission of coherent radiant energy at those frequencies. This transfer of energy from one wave to another within a lossless medium typifies parametric processes. By changing the refractive index of the crystal (via temperature, electric field, etc.), the oscillator becomes tunable. Various oscillator configurations have since evolved, with other nonlinear materials used as well, such as barium sodium niobate. The optical parametric oscillator is a laser-like, broadly tunable source of coherent radiant energy in the IR to the UV.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

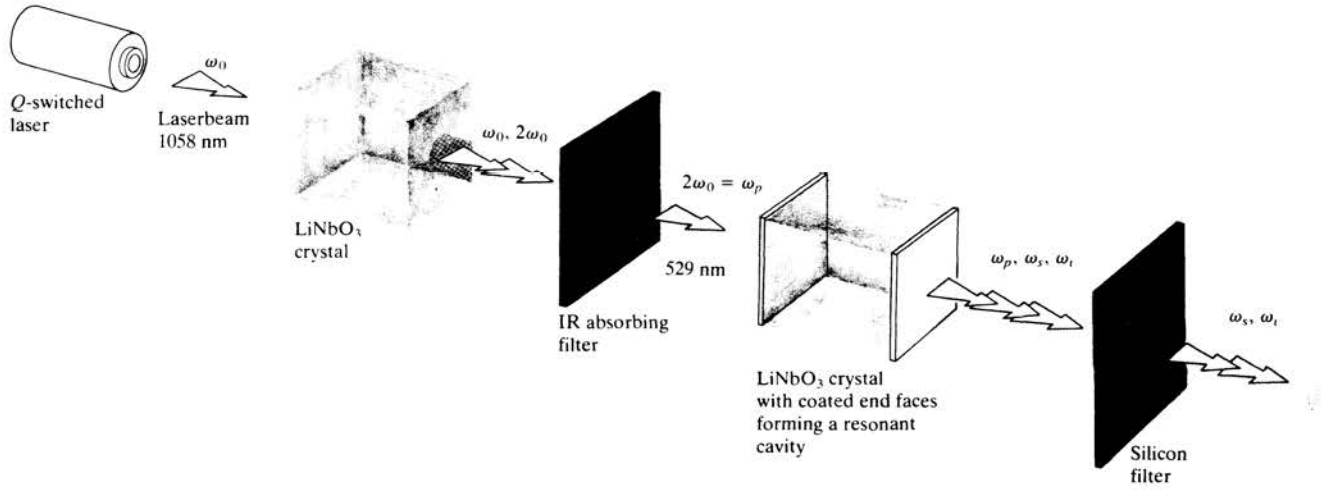


Figure 13.58 An optical parametric oscillator. [After J. A. Giordmaine and R. C. Miller, *Phys. Rev. Letters* **4**, 973 (1965).]

13.4.4 Self-Focusing of Light

When a dielectric is subjected to an electric field that varies in space, in other words, when there is a gradient of the field parallel to $\vec{\mathbf{P}}$, an internal force will result. This has the effect of altering the density, changing the permittivity, and thereby varying the refractive index, and this in both linear and nonlinear isotropic media. Suppose then that we shine an intense laserbeam with a transverse Gaussian flux-density distribution onto a specimen. The induced refractive-index variations will cause the medium in the region of the beam to function much as if it were a positive lens. Accordingly, the beam contracts, the flux density increases even more, and the contraction continues in a process known as **self-focusing**. The effect can be sustained until the beam reaches a limiting filament diameter (of about 5×10^{-6} m), being totally internally reflected as if it were in a fiberoptic element embedded within the medium.*

*See J. A. Giordmaine, "Nonlinear Optics." *Phys. Today*, **39** (January 1969).

PROBLEMS

Complete solutions to all problems—except those with an asterisk—can be found in the back of the book.

13.1* After a while, a cube of rough steel (10 cm on a side) reaches equilibrium inside a furnace at a temperature of 400°C. Knowing that its total emissivity is 0.97, determine the rate at which the cube radiates energy from each face.

13.2 A somewhat typical person has a total naked area of about 1.4 m^2 and an average skin temperature of 33°C. Determine the net power radiated per unit area, the irradiance or more precisely the exitance, if the person's total emissivity is 97% and the environment is room temperature (20°C). How much energy does that body radiate per second?

13.3 Suppose that we measure the emitted exitance from a small hole in a furnace to be 22.8 W/cm^2, using an optical pyrometer of some sort. Compute the internal temperature of the furnace.

13.4 The temperature of an object resembling a blackbody is raised from 200 K to 2000 K. By how much does the amount of energy it radiates increase?

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

13.5* Your average skin temperature is about 33°C. Assuming you radiate as does a blackbody at that temperature, at what wavelength do you emit the most energy?

13.6* What is the wavelength that carries away the most energy when an object resembling a blackbody radiates energy into a room-temperature (20°C) environment?

13.7* The surface temperature of a class O blue-white star is around 40×10^3 K. At what frequency will it radiate most of its energy?

13.8* When the Sun's spectrum is photographed, using rockets to range above the Earth's atmosphere, it is found to have a peak in its spectral exitance at roughly 465 nm. Compute the Sun's surface temperature, assuming it to be a blackbody. This approximation yields a value that is about 400 K too high.

13.9* An object resembling a blackbody emits a maximum amount of energy per unit wavelength in the red end of the visible spectrum ($\lambda = 680$ nm). What's its surface temperature?

13.10* The energy per unit area per unit time per wavelength interval emitted by a blackbody at a temperature T is given by

$$I_\lambda = \frac{2\pi hc^2}{\lambda^5}\left[\frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1}\right]$$

At a specific temperature, the total power radiated per unit area of the blackbody is equal to the area under the corresponding I_λ versus λ curve. Use this to derive the Stefan–Boltzmann Law. [*Hint*: To clean up the exponential, change variables in the integral so that

$$x = \frac{hc}{\lambda k_B T}$$

Use the fact that $\int_0^\infty x^n \frac{dx}{e^x - 1} = \Gamma(n+1)\,\zeta(n+1)$ where the gamma function is given by $\Gamma(n+1) = n!$ and the Riemann zeta function for $n = 3$ is $\zeta(4) = \frac{\pi^4}{90}$.]

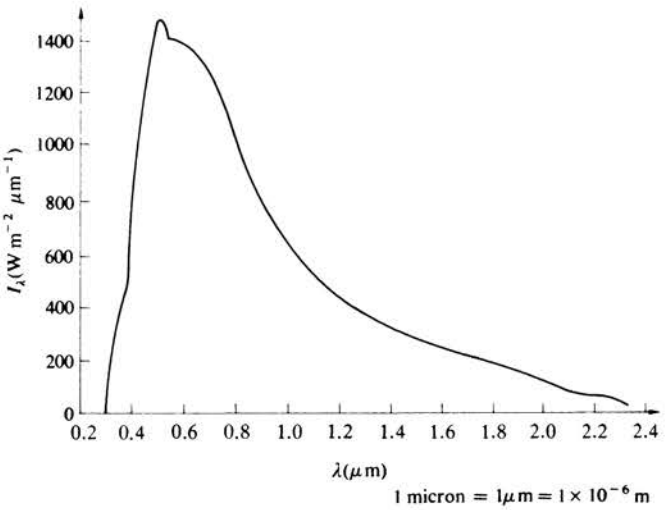
13.11* In the atomic domain, energy is often measured in electron-volts. Arrive at the following expression for the energy of a light-quantum in eV when the wavelength is in nanometers:

$$\mathscr{E} = \frac{1239.8\ \text{eV}\cdot\text{nm}}{\lambda}$$

What is the energy of a quantum of 600-nm light?

13.12 Figure P.13.12 shows the *spectral irradiance* impinging on a horizontal surface, for a clear day, at sea level, with the Sun at the zenith. What is the most energetic photon we can expect to encounter (in eV and in J)?

Figure P.13.12



13.13* Suppose we have a 100-W yellow lightbulb (550 nm) 100 m away from a 3-cm diameter shuttered aperture. Assuming the bulb to have a 2.5% conversion to radiant power, how many photons will pass through the aperture if the shutter is opened for $\frac{1}{1000}$ s?

13.14 The *solar constant* is the radiant flux density at a spherical surface centered on the Sun having a radius equal to that of the Earth's mean orbital radius; it has a value of 0.133–0.14 W/cm^2. If we assume an average wavelength of about 700 nm, how many photons at most will arrive on each square meter per second of a solar cell panel just above the atmosphere?

13.15 A 50.0-cm^3 chamber is filled with argon gas to a pressure of 20.3 Pa at a temperature of 0°C. All but a negligible number of these atoms are initially in their ground states. A flash tube surrounding the sample energizes 1.0% of the atoms into the same excited state having a mean life of 1.4×10^{-8} s. What is the maximum rate at which photons are subsequently emitted by the gas, of course it falls off with time? Assume both that spontaneous emission is the only mechanism at work and that the medium is an ideal gas.

13.16* Show that for a system of atoms and photons in equilibrium at a temperature T the ratio of the transition rates of stimulated to

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

spontaneous emission is given by

$$\left[\frac{1}{e^{\frac{h\nu}{k_B T}} - 1}\right]$$

13.17* A system of atoms in thermal equilibrium is emitting and absorbing 2.0-eV light photons. Determine the ratio of the transition rates of stimulated emission to spontaneous emission at a temperature of 300 K. Discuss the implications of your answer. [*Hint*: See the previous problem.]

13.18 Redo the previous problem for a temperature of 30.0×10^3 K and compare the results of both calculations.

13.19* For a system of atoms (in equilibrium) having two energy levels, show that at high temperatures where $k_B T >> \mathscr{E}_j - \mathscr{E}_i$, the number densities of the two states tend to become equal. [*Hint*: Form the ratio of the transition rates for total emission to absorption.]

13.20* Radiation at 21 cm pours down on the Earth from outer space. Its origin is great clouds of hydrogen gas. Taking the background temperature of space to be 3.0 K, determine the ratio of the transition rates of stimulated emission to spontaneous emission and discuss the result.

13.21* The beam (λ = 632.8 nm) from a He–Ne laser, which is initially 3.0 mm in diameter, shines on a perpendicular wall 100 m away. Given that the system is diffraction limited, how large is the circle of light on the wall?

13.22* Make a rough estimate of the amount of energy that can be delivered by a ruby laser whose crystal is 5.0 mm in diameter and 0.050 m long. Assume the pulse of light lasts 5.0×10^{-6} s. The density of aluminum oxide (Al_2O_3) is 3.7×10^3 kg/m^3. Use the data in the discussion of Fig. 13.6 and the fact that the chromium ions make a 1.79 eV lasing transition. How much power is available per pulse?

13.23 What is the transition rate for the neon atoms in a He–Ne laser if the energy drop for the 632.8 nm emission is 1.96 eV and the power output is 1.0 mW?

13.24* Given that a ruby laser operating at 694.3 nm has a frequency bandwidth of 50 MHz, what is the corresponding linewidth?

13.25* Determine the frequency difference between adjacent axial resonant cavity modes for a typical gas laser 25 cm long ($n \approx 1$).

13.26* A He–Ne c-w laser has a Doppler-broadened transition bandwidth of about 1.4 GHz at 632.8 nm. Assuming $n = 1.0$, determine the maximum cavity length for single-axial mode operation. Make a sketch of the transition linewidth and the corresponding cavity modes.

13.27 Show that the maximum electric-field intensity, E_{max}, that exists for a given irradiance I is

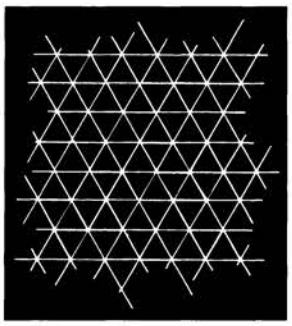
$$E_{max} = 27.4\left(\frac{I}{n}\right)^{1/2} \text{ in units of V/m}$$

where n is the refractive index of the medium.

13.28* A He–Ne laser operating at 632.8 nm has an internal beam-waist diameter of 0.60 mm. Calculate the full-angular width, or divergence, of the beam.

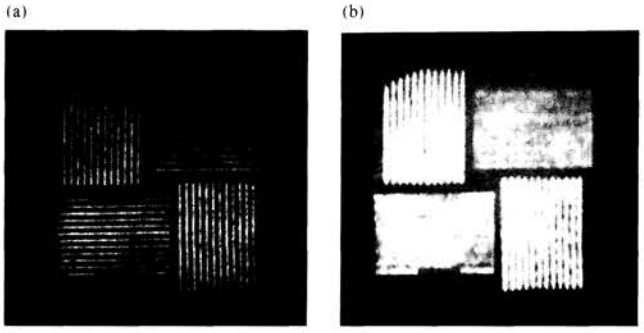
13.29 What would the pattern look like for a laserbeam diffracted by the three crossed gratings of Fig. P.13.29?

Figure P.13.29



13.30 Make a rough sketch of the Fraunhofer diffraction pattern that would arise if a transparency of Fig. P.13.30*a* served as the object. How would you filter it to get Fig. P.13.30*b*?

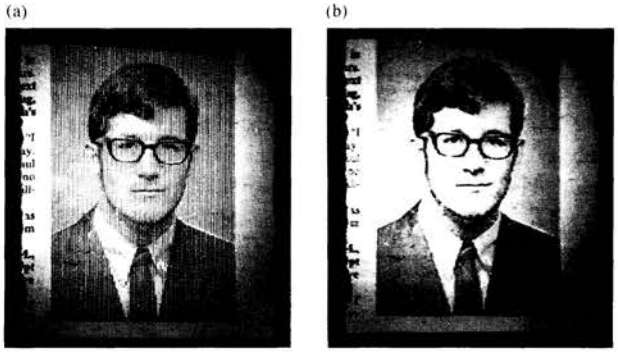
Figure P.13.30



Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

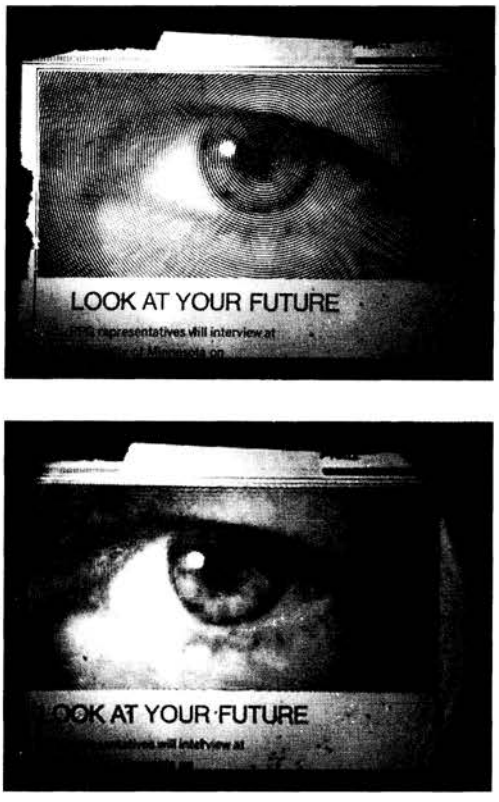
13.31 Repeat the previous problem using Fig. P.13.31 instead.

Figure P.13.31 (Photos courtesy R. A. Phillips.)



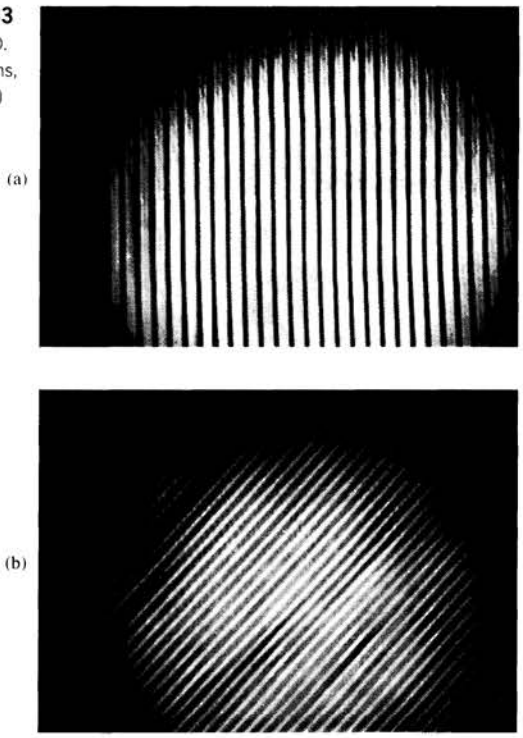
13.32* Repeat the previous problem using Fig. P.13.32 this time.

Figure P.13.32 (Photos courtesy R. A. Phillips.)



13.33 Returning to Fig. 13.32, what kind of spatial filter would produce each of the patterns shown in Fig. P.13.33?

Figure P.13.33 (Photos courtesy D. Dutton, M. P. Givens, and R. E. Hopkins.)



13.34 With Fig. 13.31 in mind, show that the transverse magnification of the system is given by $-f_t/f_t$ and draw the appropriate ray diagram. Draw a ray up through the center of the first lens at an angle θ with the axis. From the point where that ray intersects Σ_t, draw a ray downward that passes through the center of the second lens at an angle Φ. Prove that $\Phi/\theta = f_t/f_t$. Using the notion of spatial frequency, from Eq. (11.64), show that k_O at the object plane is related to k_I at the image plane by

$$k_I = k_O(f_t/f_t)$$

What does this mean with respect to the size of the image when $f_t > f_t$? What can then be said about the spatial periods of the input data as compared with the image output?

13.35 A diffraction grating having a mere 50 grooves per cm is the object in the optical computer shown in Fig. 13.31. If it is coherently illuminated by plane waves of green light (543.5 nm) from a He–Ne laser and each lens has a 100-cm focal length, what will be the spacing of the diffraction spots on the transform plane?

13.36* Imagine that you have a cosine grating (i.e., a transparency whose *amplitude* transmission profile is cosinusoidal) with a spatial

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

period of 0.01 mm. The grating is illuminated by quasimonochromatic plane waves of $\lambda = 500$ nm, and the setup is the same as that of Fig. 13.31, where the focal lengths of the transform and imaging lenses are 2.0 m and 1.0 m, respectively.

a) Discuss the resulting pattern and design a filter that will pass *only* the first-order terms. Describe it in detail.

b) What will the image look like on Σ_i with that filter in place?

c) How might you pass only the dc term, and what would the image look like then?

13.37 Suppose we insert a mask in the transform plane of the previous problem, which obscures everything but the $m = +1$ diffraction contribution. What will the reformed image look like on Σ_i? Explain your reasoning. Now suppose we remove *only* the $m = +1$ or the $m = -1$ term. What will the re-formed image look like?

13.38* Referring to the previous two problems with the cosine grating oriented horizontally, make a sketch of the electric-field amplitude along y' with no filtering. Plot the corresponding image irradiance distribution. What will the electric field of the image look like if the dc term is filtered out? Plot it. Now plot the new irradiance distribution. What can you say about the spatial frequency of the image with and without the filter in place? Relate your answers to Fig. 11.13.

13.39 Replace the cosine grating in the previous problem with a "square" bar grating, that is, a series of many fine alternating opaque and transparent bands of equal width. We now filter out all terms in the transform plane but the zeroth and the two first-order diffraction spots. These we determine to have relative irradiances of 1.00, 0.36, and 0.36: compare them with Figs. 7.32*a* and 7.33. Derive an expression for the general shape of the irradiance distribution on the image plane—make a sketch of it. What will the resulting fringe system look like?

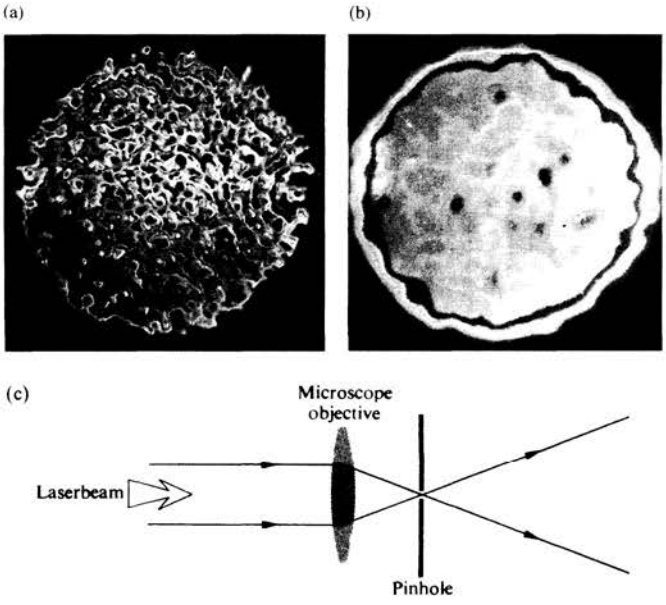
13.40 A fine square wire mesh with 50 wires per cm is placed vertically in the object plane of the optical computer of Fig. 13.30. If the lenses each have 1.00-m focal lengths, what must be the illuminating wavelength, if the diffraction spots on the transform plane are to have a horizontal and vertical separation of 2.0 mm? What will be the mesh spacing as it appears on the image plane?

13.41* Imagine that we have an opaque mask into which are punched an ordered array of circular holes, all of the same size, located as if at the corners of the boxes of a checkerboard. Now suppose our robot puncher goes mad and makes an additional batch of holes essentially randomly all across the mask. If this screen is now made the object in Problem 13.39, what will the diffraction pattern look like? Given that the ordered holes are separated from their nearest neighbors on the object by 0.1 mm, what will be the spatial frequency of the corresponding dots in the image? Describe a filter that will remove the random holes from the final image.

13.42* Imagine that we have a large photographic transparency on which there is a picture of a student made up of a regular array of small circular dots, all of the same size, but each with its own density, so that it passes a spot of light with a particular field amplitude. Considering the transparency to be illuminated by a plane wave, discuss the idea of representing the electric-field amplitude just beyond it as the product (on average) of a regular two-dimensional array of top-hat functions (Fig. 11.4, p. 523) and the continuous two-dimensional picture function: the former like a dull bed of nails, the latter an ordinary photograph. Applying the frequency convolution theorem, what does the distribution of light look like on the transform plane? How might it be filtered to produce a continuous output image?

13.43* The arrangement shown in Fig. P.13.43 is used to convert a collimated laserbeam into a spherical wave. The pinhole cleans up the beam; that is, it eliminates diffraction effects due to dust and the like on the lens. How does it manage it?

Figure P.13.43 (*a*) and (*b*) A high-power laserbeam before and after spatial filtering. (Photos courtesy Lawrence Livermore National Laboratory.)



13.44 What would happen to the speckle pattern if a laserbeam were projected onto a suspension such as milk rather than onto a smooth wall?

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Appendix 1

Electromagnetic Theory

MAXWELL'S EQUATIONS IN DIFFERENTIAL FORM

The set of integral expressions that have come to be known as Maxwell's Equations are

$$\oint_C \vec{\mathbf{E}} \cdot d\vec{\ell} = -\iint_A \frac{\partial \vec{\mathbf{B}}}{\partial t} \cdot d\vec{\mathbf{S}} \qquad [3.5]$$

$$\oint_C \frac{\vec{\mathbf{B}}}{\mu} \cdot d\vec{\ell} = \iint_A \left(\vec{\mathbf{J}} + \epsilon \frac{\partial \vec{\mathbf{E}}}{\partial t} \right) \cdot d\vec{\mathbf{S}} \qquad [3.13]$$

$$\oiint_A \epsilon \vec{\mathbf{E}} \cdot d\vec{\mathbf{S}} = \iiint_V \rho \, dV \qquad [3.7]$$

and

$$\oiint_A \vec{\mathbf{B}} \cdot d\vec{\mathbf{S}} = 0 \qquad [3.9]$$

where the units, as usual, are SI.

Maxwell's Equations can be written in a differential form, which is more useful for deriving the wave aspects of the electromagnetic field. This transition can readily be accomplished by making use of two theorems from vector calculus, namely, Gauss's Divergence Theorem,

$$\oiint_A \vec{\mathbf{F}} \cdot d\vec{\mathbf{S}} = \iiint_V \nabla \cdot \vec{\mathbf{F}} \, dV \qquad (A1.1)$$

and Stokes's Theorem

$$\oint_C \vec{\mathbf{F}} \cdot d\vec{\ell} = \iint_A \nabla \times \vec{\mathbf{F}} \cdot d\vec{\mathbf{S}} \qquad (A1.2)$$

Here the quantity $\vec{\mathbf{F}}$ is not one fixed vector but a function that depends on the position variables. It is a rule that associates a single vector, for example, in Cartesian coordinates, $\vec{\mathbf{F}}(x, y, z)$, with each point (x, y, z) in space. Vector-valued functions of this kind, such as $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$, are known as vector fields.

Applying Stokes's Theorem to the electric field intensity, we have

$$\oint \vec{\mathbf{E}} \cdot d\vec{\ell} = \iint \nabla \times \vec{\mathbf{E}} \cdot d\vec{\mathbf{S}} \qquad (A1.3)$$

If we compare this with Eq. (3.5), it follows that

$$\iint \nabla \times \vec{\mathbf{E}} \cdot d\vec{\mathbf{S}} = -\iint \frac{\partial \vec{\mathbf{B}}}{\partial t} \cdot d\vec{\mathbf{S}} \qquad (A1.4)$$

This result must be true for all surfaces bounded by the path C. This can only be the case if the integrands are themselves equal. that is, if

$$\nabla \times \vec{\mathbf{E}} = -\frac{\partial \vec{\mathbf{B}}}{\partial t} \qquad (A1.5)$$

A similar application of Stokes's Theorem to $\vec{\mathbf{B}}$, using Eq. (3.13), results in

$$\nabla \times \vec{\mathbf{B}} = \mu \left(\vec{\mathbf{J}} + \epsilon \frac{\partial \vec{\mathbf{E}}}{\partial t} \right) \qquad (A1.6)$$

Gauss's Divergence Theorem applied to the electric field intensity yields

$$\oiint \vec{\mathbf{E}} \cdot d\vec{\mathbf{S}} = \iiint \nabla \cdot \vec{\mathbf{E}} \, dV \qquad (A1.7)$$

If we make use of Eq. (3.7), this becomes

$$\iiint_V \nabla \cdot \vec{\mathbf{E}} \, dV = \frac{1}{\epsilon} \iiint_V \rho \, dV \qquad (A1.8)$$

and since this is to be true for any volume (i.e., for an arbitrary closed

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

domain), the two integrands must be equal. Consequently, at any point (x, y, z, t) in space–time

$$\nabla \cdot \vec{\mathbf{E}} = \frac{\rho}{\epsilon} \tag{A1.9}$$

In the same fashion, Gauss's Divergence Theorem applied to the $\vec{\mathbf{B}}$-field and combined with Eq. (3.9) yields

$$\nabla \cdot \vec{\mathbf{B}} = 0 \tag{A1.10}$$

Equations (A1.5), (A1.6), (A1.9), and (A1.10) are Maxwell's Equations in differential form. Refer back to Eqs. (3.18) through (3.21) for the simple case of Cartesian coordinates and *free space* ($\rho = J = 0$, $\epsilon = \epsilon_0$, $\mu = \mu_0$).

ELECTROMAGNETIC WAVES

To derive the electromagnetic wave equation in its most general form, we must again consider the presence of some medium. We saw in Section 3.5.1 that there is a need to introduce the *polarization* vector $\vec{\mathbf{P}}$, which is a measure of the overall behavior of the medium, in that it is the resultant electric dipole moment per unit volume. Since the field within the material has been altered, we are led to define a new field quantity, the *displacement* $\vec{\mathbf{D}}$:

$$\vec{\mathbf{D}} = \epsilon_0 \vec{\mathbf{E}} + \vec{\mathbf{P}} \tag{A1.11}$$

Clearly then,

$$\vec{\mathbf{E}} = \frac{\vec{\mathbf{D}}}{\epsilon_0} - \frac{\vec{\mathbf{P}}}{\epsilon_0}$$

The internal electric field $\vec{\mathbf{E}}$ is the difference between the field $\vec{\mathbf{D}}/\epsilon_0$, which would exist in the absence of polarization, and the field $\vec{\mathbf{P}}/\epsilon_0$ arising from polarization.

For a homogeneous, linear, isotropic dielectric, $\vec{\mathbf{P}}$ and $\vec{\mathbf{E}}$ are in the same direction and are mutually proportional. It follows that $\vec{\mathbf{D}}$ is therefore also proportional to $\vec{\mathbf{E}}$:

$$\vec{\mathbf{D}} = \epsilon \vec{\mathbf{E}} \tag{A1.12}$$

Like $\vec{\mathbf{E}}$, $\vec{\mathbf{D}}$ extends throughout space and is in no way limited to the region occupied by the dielectric, as is $\vec{\mathbf{P}}$. The lines of $\vec{\mathbf{D}}$ begin and end on free, movable charges. Those of $\vec{\mathbf{E}}$ begin and end on either *free* charges or bound polarization charges. If no free charge is present, as might be the case in the vicinity of a polarized dielectric or in free space, the lines of $\vec{\mathbf{D}}$ close on themselves.

Since in general the response of optical media to $\vec{\mathbf{B}}$-fields is only slightly different from that of a vacuum, we need not describe the process in detail. Suffice it to say that the material will become polarized. We can define a *magnetic polarization* or *magnetization* vector $\vec{\mathbf{M}}$ as the magnetic dipole moment per unit volume. In order to deal with the influence of the magnetically polarized medium, we introduce an auxiliary vector $\vec{\mathbf{H}}$, traditionally known as the *magnetic field intensity*

$$\vec{\mathbf{H}} = \mu_0^{-1} \vec{\mathbf{B}} - \vec{\mathbf{M}} \tag{A1.13}$$

For a homogeneous, linear (nonferromagnetic), isotropic medium, $\vec{\mathbf{B}}$ and $\vec{\mathbf{H}}$ are parallel and proportional:

$$\vec{\mathbf{H}} = \mu^{-1} \vec{\mathbf{B}} \tag{A1.14}$$

Along with Eqs. (A1.12) and (A1.14), there is one more *constitutive equation*,

$$\vec{\mathbf{J}} = \sigma \vec{\mathbf{E}} \tag{A1.15}$$

Known as *Ohm's Law*, it is a statement of an experimentally determined rule that holds for conductors at constant temperatures. The electric field intensity, and therefore the force acting on each electron in a conductor, determines the flow of charge. The constant of proportionality relating $\vec{\mathbf{E}}$ and $\vec{\mathbf{J}}$ is the conductivity of the particular medium, σ.

Consider the rather general environment of a linear (nonferroelectric and nonferromagnetic), homogeneous, isotropic medium, which is physically at rest. By making use of the constitutive relations, we can rewrite Maxwell's Equations as

$$\nabla \cdot \vec{\mathbf{E}} = \frac{\rho}{\epsilon} \tag{A1.9}$$

$$\nabla \cdot \vec{\mathbf{B}} = 0 \tag{A1.10}$$

$$\nabla \times \vec{\mathbf{E}} = -\frac{\partial \vec{\mathbf{B}}}{\partial t} \tag{A1.5}$$

and

$$\nabla \times \vec{\mathbf{B}} = \mu\sigma \vec{\mathbf{E}} + \mu\epsilon \frac{\partial \vec{\mathbf{E}}}{\partial t} \tag{A1.16}$$

If these expressions are somehow to yield a wave equation (2.61), we had best form some second derivatives with respect to the space variables. Taking the curl of Eq. (A1.16), we obtain

$$\nabla \times (\nabla \times \vec{\mathbf{B}}) = \mu\sigma(\nabla \times \vec{\mathbf{E}}) + \mu\epsilon \frac{\partial}{\partial t}(\nabla \times \vec{\mathbf{E}}) \tag{A1.17}$$

where, since $\vec{\mathbf{E}}$ is assumed to be a well-behaved function, the space and time derivatives can be interchanged. Equation (A1.5) can be substituted to obtain the needed second derivative with respect to time:

$$\nabla \times (\nabla \times \vec{\mathbf{B}}) = -\mu\sigma \frac{\partial \vec{\mathbf{B}}}{\partial t} - \mu\epsilon \frac{\partial^2 \vec{\mathbf{B}}}{\partial t^2} \tag{A1.18}$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

The vector triple product can be simplified by making use of the operator identity

$$\nabla \times (\nabla \times) = \nabla(\nabla \cdot) - \nabla^2 \tag{A1.19}$$

so that

$$\nabla \times (\nabla \times \vec{\mathbf{B}}) = \nabla(\nabla \cdot \vec{\mathbf{B}}) - \nabla^2 \vec{\mathbf{B}}$$

where in Cartesian coordinates

$$(\nabla \cdot \nabla)\vec{\mathbf{B}} = \nabla^2 \vec{\mathbf{B}} \equiv \frac{\partial^2 \vec{\mathbf{B}}}{\partial x^2} + \frac{\partial^2 \vec{\mathbf{B}}}{\partial y^2} + \frac{\partial^2 \vec{\mathbf{B}}}{\partial z^2}$$

Since the divergence of $\vec{\mathbf{B}}$ is zero, Eq. (A1.18) becomes

$$\nabla^2 \vec{\mathbf{B}} = \mu\epsilon \frac{\partial^2 \vec{\mathbf{B}}}{\partial t^2} - \mu\sigma \frac{\partial \vec{\mathbf{B}}}{\partial t} = 0 \tag{A1.20}$$

A similar equation is satisfied by the electric field intensity. Following essentially the same procedure as above, take the curl of Eq. (A1.5):

$$\nabla \times (\nabla \times \vec{\mathbf{E}}) = -\frac{\partial}{\partial t}(\nabla \times \vec{\mathbf{B}})$$

Eliminating $\vec{\mathbf{B}}$ this becomes

$$\nabla \times (\nabla \times \vec{\mathbf{E}}) = -\mu\sigma \frac{\partial \mathbf{E}}{\partial t} - \mu\epsilon \frac{\partial^2 \vec{\mathbf{E}}}{\partial t^2}$$

and then by making use of Eq. (A1.19), we arrive at

$$\nabla^2 \vec{\mathbf{E}} - \mu\epsilon \frac{\partial^2 \vec{\mathbf{E}}}{\partial t^2} - \mu\sigma \frac{\partial \vec{\mathbf{E}}}{\partial t} = \nabla(\rho/\epsilon)$$

having utilized the fact that

$$\nabla(\nabla \cdot \vec{\mathbf{E}}) = \nabla(\rho/\epsilon)$$

For an uncharged medium ($\rho = 0$) and

$$\nabla^2 \vec{\mathbf{E}} - \mu\epsilon \frac{\partial^2 \vec{\mathbf{E}}}{\partial t^2} - \mu\sigma \frac{\partial \vec{\mathbf{E}}}{\partial t} = 0 \tag{A1.21}$$

Equations (A1.20) and (A1.21) are known as the *equations of telegraphy.**

In nonconducting media $\sigma = 0$, and these equations become

$$\nabla^2 \vec{\mathbf{B}} - \mu\epsilon \frac{\partial^2 \vec{\mathbf{B}}}{\partial t^2} = 0 \tag{A1.22}$$

$$\nabla^2 \vec{\mathbf{E}} - \mu\epsilon \frac{\partial^2 \vec{\mathbf{E}}}{\partial t^2} = 0 \tag{A1.23}$$

and similarly

$$\nabla^2 \vec{\mathbf{H}} - \mu\epsilon \frac{\partial^2 \vec{\mathbf{H}}}{\partial t^2} = 0 \tag{A1.24}$$

and

$$\nabla^2 \vec{\mathbf{D}} - \mu\epsilon \frac{\partial^2 \vec{\mathbf{D}}}{\partial t^2} = 0 \tag{A1.25}$$

In the special nonconducting medium of a vacuum (free space) where

$$\rho = 0 \qquad \sigma = 0 \qquad K_e = 1 \qquad K_m = 1$$

these equations become simply

$$\nabla^2 \vec{\mathbf{E}} = \mu_0\epsilon_0 \frac{\partial^2 \vec{\mathbf{E}}}{\partial t^2} \tag{A1.26}$$

and

$$\nabla^2 \vec{\mathbf{B}} = \mu_0\epsilon_0 \frac{\partial^2 \vec{\mathbf{B}}}{\partial t^2} \tag{A1.27}$$

Both of these expressions describe coupled space- and time-dependent fields, and both have the form of the differential wave equation (see Section 3.2 for further discussion).

*For a pair of parallel wires that might serve as a telegraph line, the finite wire resistance results in a power loss and joule heating. An electromagnetic wave advancing along the line has less and less energy available to it. The first-order time derivatives in Eqs. (A1.20) and (A1.21) arise from the conduction current and lead to the dissipation or damping.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Appendix 2

The Kirchhoff Diffraction Theory

To solve the Helmholtz Equation [Eq. (10.113)], suppose that we have two scalar functions U_1 and U_2 for which Green's theorem is

$$\iiint_V (U_1\nabla^2 U_2 - U_2\nabla^2 U_1)dV$$

$$= \oiint_S (U_1\nabla U_2 - U_2\nabla U_1)\cdot d\vec{\mathbf{S}} \qquad \text{(A2.1)}$$

It is clear that if U_1 and U_2 are solutions of the Helmholtz Equation, that is, if

$$\nabla^2 U_1 + k^2 U_1 = 0$$

and

$$\nabla^2 U_2 + k^2 U_2 = 0$$

then

$$\oiint_S (U_1\nabla U_2 - U_2\nabla U_1)\cdot d\vec{\mathbf{S}} = 0 \qquad \text{(A2.2)}$$

Let $U_1 = \tilde{\mathscr{E}}$, the space portion of an unspecified scalar optical disturbance [Eq. (10.112)]. And let

$$U_2 = \frac{e^{ikr}}{r}$$

where r is measured from a point P. Both of these choices clearly satisfy the Helmholtz Equation. There is a singularity at point P, where $r = 0$, so that we surround it by a small sphere in order to exclude P from the region enclosed by S (see Fig. A2.1). Equation (A2.2) now becomes

$$\oiint_S \left[\tilde{\mathscr{E}}\nabla\left(\frac{e^{ikr}}{r}\right) - \frac{e^{ikr}}{r}\nabla\tilde{\mathscr{E}}\right]\cdot d\vec{\mathbf{S}}$$

$$+ \oiint_{S'} \left[\tilde{\mathscr{E}}\nabla\left(\frac{e^{ikr}}{r}\right) - \frac{e^{ikr}}{r}\nabla\tilde{\mathscr{E}}\right]\cdot d\vec{\mathbf{S}} = 0 \qquad \text{(A2.3)}$$

Now expand out the portion of the integral corresponding to S'. On the small sphere, the unit normal $\hat{\mathbf{n}}$ points toward the origin at P, and

$$\nabla\left(\frac{e^{ikr}}{r}\right) = \left(\frac{1}{r^2} - \frac{ik}{r}\right)e^{ikr}\hat{\mathbf{n}}$$

since the gradient is directed radially outward. In terms of the solid angle ($dS = r^2 d\Omega$) measured at P, the integral over S' becomes

$$\oiint_{S'} \left(\tilde{\mathscr{E}} - ik\tilde{\mathscr{E}}r + r\frac{\partial\tilde{\mathscr{E}}}{\partial r}\right)e^{ikr}\, d\Omega \qquad \text{(A2.4)}$$

where $\nabla\tilde{\mathscr{E}}\cdot d\vec{\mathbf{S}} = -(\partial\tilde{\mathscr{E}}/\partial r)r^2 d\Omega$. As the sphere surrounding P shrinks, $r \to 0$ on S' and $\exp(ikr) \to 1$. Because of the continuity of $\tilde{\mathscr{E}}$ its value at any point on S' approaches its value at P, that is, $\tilde{\mathscr{E}}_p$. The last two terms in Eq. (A2.4) go to zero, and the integral becomes $4\pi\tilde{\mathscr{E}}_p$. Finally then, Eq. (A2.3) becomes

$$\tilde{\mathscr{E}}_p = \frac{1}{4\pi}\left[\oiint_S \frac{e^{ikr}}{r}\nabla\tilde{\mathscr{E}}\cdot d\vec{\mathbf{S}} - \oiint_S \tilde{\mathscr{E}}\nabla\left(\frac{e^{ikr}}{r}\right)\cdot d\vec{\mathbf{S}}\right] \qquad \text{[10.114]}$$

which is known as the *Kirchhoff Integral Theorem.*



Page 658

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Table 1

THE SINC FUNCTION

(sin u)/u

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	1.000000	0.999983	0.999933	0.999850	0.999733	0.999583	0.999400	0.999184	0.998934	0.998651
0.1	0.998334	0.997985	0.997602	0.997186	0.996737	0.996254	0.995739	0.995190	0.994609	0.993994
0.2	0.993347	0.992666	0.991953	0.991207	0.990428	0.989616	0.988771	0.987894	0.986984	0.986042
0.3	0.985067	0.984060	0.983020	0.981949	0.980844	0.979708	0.978540	0.977339	0.976106	0.974842
0.4	0.973546	0.972218	0.970858	0.969467	0.968044	0.966590	0.965105	0.963588	0.962040	0.960461
0.5	0.958851	0.957210	0.955539	0.953836	0.952104	0.950340	0.948547	0.946723	0.944869	0.942985
0.6	0.941071	0.939127	0.937153	0.935150	0.933118	0.931056	0.928965	0.926845	0.924696	0.922518
0.7	0.920311	0.918076	0.915812	0.913520	0.911200	0.908852	0.906476	0.904072	0.901640	0.899181
0.8	0.896695	0.894182	0.891641	0.889074	0.886480	0.883859	0.881212	0.878539	0.875840	0.873114
0.9	0.870363	0.867587	0.864784	0.861957	0.859104	0.856227	0.853325	0.850398	0.847446	0.844471
1.0	0.841471	0.838447	0.835400	0.832329	0.829235	0.826117	0.822977	0.819814	0.816628	0.813419
1.1	0.810189	0.806936	0.803661	0.800365	0.797047	0.793708	0.790348	0.786966	0.783564	0.780142
1.2	0.776699	0.773236	0.769754	0.766251	0.762729	0.759188	0.755627	0.752048	0.748450	0.744833
1.3	0.741199	0.737546	0.733875	0.730187	0.726481	0.722758	0.719018	0.715261	0.711488	0.707698
1.4	0.703893	0.700071	0.696234	0.692381	0.688513	0.684630	0.680732	0.676819	0.672892	0.668952
1.5	0.664997	0.661028	0.657046	0.653051	0.649043	0.645022	0.640988	0.636942	0.632885	0.628815
1.6	0.624734	0.620641	0.616537	0.612422	0.608297	0.604161	0.600014	0.595858	0.591692	0.587517
1.7	0.583332	0.579138	0.574936	0.570725	0.566505	0.562278	0.558042	0.553799	0.549549	0.545291
1.8	0.541026	0.536755	0.532478	0.528194	0.523904	0.519608	0.515307	0.511001	0.506689	0.502373
1.9	0.498053	0.493728	0.489399	0.485066	0.480729	0.476390	0.472047	0.467701	0.463353	0.459002
2.0	0.454649	0.450294	0.445937	0.441579	0.437220	0.432860	0.428499	0.424137	0.419775	0.415414
2.1	0.411052	0.406691	0.402330	0.397971	0.393612	0.389255	0.384900	0.380546	0.376194	0.371845
2.2	0.367498	0.363154	0.358813	0.354475	0.350141	0.345810	0.341483	0.337161	0.332842	0.328529
2.3	0.324220	0.319916	0.315617	0.311324	0.307036	0.302755	0.298479	0.294210	0.289947	0.285692
2.4	0.281443	0.277202	0.272967	0.268741	0.264523	0.260312	0.256110	0.251916	0.247732	0.243556
2.5	0.239389	0.235231	0.231084	0.226946	0.222817	0.218700	0.214592	0.210495	0.206409	0.202334
2.6	0.198270	0.194217	0.190176	0.186147	0.182130	0.178125	0.174132	0.170152	0.166185	0.162230
2.7	0.158289	0.154361	0.150446	0.146546	0.142659	0.138786	0.134927	0.131083	0.127253	0.123439
2.8	0.119639	0.115854	0.112084	0.108330	0.104592	0.100869	0.097163	0.093473	0.089798	0.086141
2.9	0.082500	0.078876	0.075268	0.071678	0.068105	0.064550	0.061012	0.057492	0.053990	0.050506
3.0	0.047040	0.043592	0.040163	0.036753	0.033361	0.029988	0.026635	0.023300	0.019985	0.016689
3.1	0.013413	0.010157	0.006920	0.003704	0.000507	−0.002669	−0.005825	−0.008960	−0.012075	−0.015169
3.2	−0.018242	−0.021294	−0.024325	−0.027335	−0.030324	−0.033291	−0.036236	−0.039160	−0.042063	−0.044943
3.3	−0.047802	−0.050638	−0.053453	−0.056245	−0.059014	−0.061762	−0.064487	−0.067189	−0.069868	−0.072525
3.4	−0.075159	−0.077770	−0.080358	−0.082923	−0.085465	−0.087983	−0.090478	−0.092950	−0.095398	−0.097823
3.5	−0.100224	−0.102601	−0.104955	−0.107285	−0.109591	−0.111873	−0.114131	−0.116365	−0.118575	−0.120761
3.6	−0.122922	−0.125060	−0.127173	−0.129262	−0.131326	−0.133366	−0.135382	−0.137373	−0.139339	−0.141282
3.7	−0.143199	−0.145092	−0.146960	−0.148803	−0.150622	−0.152416	−0.154186	−0.155930	−0.157650	−0.159345
3.8	−0.161015	−0.162661	−0.164281	−0.165877	−0.167448	−0.168994	−0.170515	−0.172011	−0.173482	−0.174929
3.9	−0.176350	−0.177747	−0.179119	−0.180466	−0.181788	−0.183086	−0.184358	−0.185606	−0.186829	−0.188027

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

TABLE 1 (CONTINUED)

(sin u)/u

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
4.0	−0.189201	−0.190349	−0.191473	−0.192573	−0.193647	−0.194698	−0.195723	−0.196724	−0.197700	−0.198652
4.1	−0.199580	−0.200483	−0.201361	−0.202216	−0.203046	−0.203851	−0.204633	−0.205390	−0.206124	−0.206833
4.2	−0.207518	−0.208179	−0.208817	−0.209430	−0.210020	−0.210586	−0.211128	−0.211647	−0.212142	−0.212614
4.3	−0.213062	−0.213487	−0.213888	−0.214267	−0.214622	−0.214955	−0.215264	−0.215550	−0.215814	−0.216055
4.4	−0.216273	−0.216469	−0.216642	−0.216793	−0.216921	−0.217028	−0.217112	−0.217174	−0.217214	−0.217232
4.5	−0.217229	−0.217204	−0.217157	−0.217089	−0.217000	−0.216889	−0.216757	−0.216604	−0.216430	−0.216235
4.6	−0.216020	−0.215784	−0.215527	−0.215250	−0.214953	−0.214635	−0.214298	−0.213940	−0.213563	−0.213166
4.7	−0.212750	−0.212314	−0.211858	−0.211384	−0.210890	−0.210377	−0.209846	−0.209296	−0.208727	−0.208140
4.8	−0.207534	−0.206911	−0.206269	−0.205609	−0.204932	−0.204236	−0.023524	−0.202794	−0.202046	−0.201282
4.9	−0.200501	−0.199702	−0.198887	−0.198056	−0.197208	−0.196344	−0.195464	−0.194568	−0.193656	−0.192728
5.0	−0.191785	−0.190826	−0.189853	−0.188864	−0.187860	−0.186841	−0.185808	−0.184760	−0.183699	−0.182622
5.1	−0.181532	−0.180428	−0.179311	−0.178179	−0.177035	−0.175877	−0.174706	−0.173522	−0.172326	−0.171117
5.2	−0.169895	−0.168661	−0.167415	−0.166158	−0.164888	−0.163607	−0.162314	−0.161010	−0.159695	−0.158369
5.3	−0.157032	−0.155684	−0.154326	−0.152958	−0.151579	−0.150191	−0.148792	−0.147384	−0.145967	−0.144540
5.4	−0.143105	−0.141660	−0.140206	−0.138744	−0.137273	−0.135794	−0.134307	−0.132812	−0.131309	−0.129798
5.5	−0.128280	−0.126755	−0.125222	−0.123683	−0.122137	−0.120584	−0.119024	−0.117459	−0.115887	−0.114310
5.6	−0.112726	−0.111137	−0.109543	−0.107943	−0.106338	−0.104728	−0.103114	−0.101495	−0.099871	−0.098243
5.7	−0.096611	−0.094976	−0.093336	−0.091693	−0.090046	−0.088396	−0.086743	−0.085087	−0.083429	−0.081768
5.8	−0.080104	−0.078438	−0.076770	−0.075100	−0.073428	−0.071755	−0.070080	−0.068404	−0.066726	−0.065048
5.9	−0.063369	−0.061689	−0.060009	−0.058329	−0.056648	−0.054967	−0.053287	−0.051606	−0.049927	−0.048248
6.0	−0.046569	−0.044892	−0.043216	−0.041540	−0.039867	−0.038195	−0.036524	−0.034856	−0.033189	−0.031525
6.1	−0.029863	−0.028203	−0.026546	−0.024892	−0.023240	−0.021592	−0.019947	−0.018305	−0.016667	−0.015032
6.2	−0.013402	−0.011775	−0.010152	−0.008533	−0.006919	−0.005309	−0.003703	−0.002103	−0.000507	0.001083
6.3	0.002669	0.004249	0.005824	0.007393	0.008956	0.010514	0.012066	0.013612	0.015151	0.016684
6.4	0.018211	0.019731	0.021244	0.022751	0.024250	0.025743	0.027228	0.028706	0.030177	0.031640
6.5	0.033095	0.034543	0.035983	0.037414	0.038838	0.040253	0.041661	0.043059	0.044449	0.045831
6.6	0.047203	0.048567	0.049922	0.051268	0.052604	0.053931	0.055249	0.056558	0.057857	0.059146
6.7	0.060425	0.061695	0.062955	0.064204	0.065444	0.066673	0.067892	0.069101	0.070299	0.071487
6.8	0.072664	0.073830	0.074986	0.076130	0.077264	0.078386	0.079498	0.080598	0.081688	0.082765
6.9	0.083832	0.084887	0.085930	0.086962	0.087982	0.088991	0.089987	0.090972	0.091945	0.092906
7.0	0.093855	0.094792	0.095717	0.096629	0.097530	0.098418	0.099293	0.100157	0.101008	0.101846
7.1	0.102672	0.103485	0.104286	0.105074	0.105849	0.106611	0.107361	0.108098	0.108822	0.109533
7.2	0.110232	0.110917	0.111589	0.112249	0.112895	0.113528	0.114149	0.114756	0.115350	0.115931
7.3	0.116498	0.117053	0.117594	0.118122	0.118637	0.119138	0.119627	0.120102	0.120563	0.121012
7.4	0.121447	0.121869	0.122277	0.122673	0.123055	0.123423	0.123779	0.124121	0.124449	0.124765
7.5	0.125067	0.125355	0.125631	0.125893	0.126142	0.126378	0.126600	0.126809	0.127005	0.127188
7.6	0.127358	0.127514	0.127658	0.127788	0.127905	0.128009	0.128100	0.128178	0.128243	0.128295
7.7	0.128334	0.128360	0.128373	0.128373	0.128361	0.128335	0.128297	0.128247	0.128183	0.128107
7.8	0.128018	0.127917	0.127803	0.127677	0.127539	0.127388	0.127224	0.127049	0.126861	0.126661
7.9	0.126448	0.126224	0.125988	0.125739	0.125479	0.125207	0.124923	0.124627	0.124320	0.124000

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

TABLE 1 (CONTINUED)

(sin u)/u

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
8.0	0.123670	0.123328	0.122974	0.122609	0.122232	0.121845	0.121446	0.121036	0.120615	0.120183
8.1	0.119739	0.119286	0.118821	0.118345	0.117859	0.117363	0.116855	0.116338	0.115810	0.115272
8.2	0.114723	0.114165	0.113596	0.113018	0.112429	0.111831	0.111223	0.110605	0.109978	0.109341
8.3	0.108695	0.108040	0.107376	0.106702	0.106019	0.105327	0.104627	0.103918	0.103200	0.102473
8.4	0.101738	0.100994	0.100243	0.099483	0.098714	0.097938	0.097154	0.096362	0.095562	0.094755
8.5	0.093940	0.093117	0.092287	0.091450	0.090606	0.089755	0.088896	0.088031	0.087159	0.086280
8.6	0.085395	0.084503	0.083605	0.082701	0.081790	0.080874	0.079951	0.079023	0.078089	0.077149
8.7	0.076203	0.075253	0.074296	0.073335	0.072369	0.071397	0.070421	0.069439	0.068453	0.067463
8.8	0.066468	0.065468	0.064465	0.063457	0.062445	0.061429	0.060410	0.059386	0.058359	0.057328
8.9	0.056294	0.055257	0.054217	0.053173	0.052127	0.051077	0.050025	0.048970	0.047913	0.046853
9.0	0.045791	0.044727	0.043660	0.042592	0.041521	0.040449	0.039375	0.038300	0.037223	0.036145
9.1	0.035066	0.033985	0.032904	0.031821	0.030738	0.029654	0.028569	0.027484	0.026399	0.025313
9.2	0.024227	0.023141	0.022055	0.020970	0.019884	0.018799	0.017714	0.016630	0.015547	0.014464
9.3	0.013382	0.012301	0.011222	0.010143	0.009066	0.007990	0.006916	0.005843	0.004772	0.003703
9.4	0.002636	0.001570	0.000507	−0.000554	−0.001612	−0.002669	−0.003722	−0.004774	−0.005822	−0.006868
9.5	−0.007911	−0.008950	−0.009987	−0.011021	−0.012051	−0.013078	−0.014101	−0.015121	−0.016138	−0.017150
9.6	−0.018159	−0.019164	−0.020165	−0.021161	−0.022154	−0.023142	−0.024126	−0.025106	−0.026081	−0.027051
9.7	−0.028017	−0.028977	−0.029933	−0.030884	−0.031830	−0.032771	−0.033707	−0.034637	−0.035562	−0.036482
9.8	−0.037396	−0.038304	−0.039207	−0.040104	−0.040995	−0.041881	−0.042760	−0.043633	−0.044500	−0.045361
9.9	−0.046216	−0.047064	−0.047906	−0.048741	−0.049570	−0.050392	−0.051208	−0.052017	−0.052819	−0.053614
10.0	−0.054402	−0.055183	−0.055957	−0.056724	−0.057484	−0.058237	−0.058982	−0.059720	−0.060450	−0.061173
10.1	−0.061888	−0.062596	−0.063296	−0.063988	−0.064673	−0.065350	−0.066019	−0.066680	−0.067333	−0.067978
10.2	−0.068615	−0.069244	−0.069865	−0.070477	−0.071082	−0.071678	−0.072266	−0.072845	−0.073416	−0.073979
10.3	−0.074533	−0.075078	−0.075615	−0.076143	−0.076663	−0.077174	−0.077677	−0.078170	−0.078655	−0.079131
10.4	−0.079599	−0.080057	−0.080507	−0.080947	−0.081379	−0.081802	−0.082216	−0.082620	−0.083016	−0.083403
10.5	−0.083781	−0.084149	−0.084509	−0.084859	−0.085200	−0.085532	−0.085855	−0.086169	−0.086473	−0.086768
10.6	−0.087054	−0.087331	−0.087599	−0.087857	−0.088106	−0.088346	−0.088576	−0.088797	−0.089009	−0.089212
10.7	−0.089405	−0.089589	−0.089764	−0.089929	−0.090085	−0.090232	−0.090370	−0.090498	−0.090617	−0.090727
10.8	−0.090827	−0.090919	−0.091001	−0.091073	−0.091137	−0.091191	−0.091236	−0.091272	−0.091299	−0.091316
10.9	−0.091324	−0.091324	−0.091314	−0.091295	−0.091267	−0.091229	−0.091183	−0.091128	−0.091064	−0.090990
11.0	−0.090908	−0.090817	−0.090717	−0.090608	−0.090490	−0.090364	−0.090228	−0.090084	−0.089931	−0.089770
11.1	−0.089599	−0.089420	−0.089233	−0.089037	−0.088832	−0.088619	−0.088397	−0.088167	−0.087929	−0.087682
11.2	−0.087427	−0.087163	−0.086891	−0.086612	−0.086324	−0.086027	−0.085723	−0.085411	−0.085091	−0.084763
11.3	−0.084426	−0.084083	−0.083731	−0.083371	−0.083004	−0.082630	−0.082247	−0.081857	−0.081460	−0.081055
11.4	−0.080643	−0.080223	−0.079796	−0.079362	−0.078921	−0.078473	−0.078017	−0.077555	−0.077086	−0.076609
11.5	−0.076126	−0.075636	−0.075140	−0.074637	−0.074127	−0.073611	−0.073088	−0.072559	−0.072023	−0.071481
11.6	−0.070934	−0.070379	−0.069819	−0.069253	−0.068681	−0.068103	−0.067519	−0.066929	−0.066334	−0.065733
11.7	−0.065127	−0.064515	−0.063898	−0.063275	−0.062647	−0.062014	−0.061376	−0.060733	−0.060084	−0.059431
11.8	−0.058773	−0.058111	−0.057443	−0.056771	−0.056095	−0.055414	−0.054728	−0.054039	−0.053345	−0.052646
11.9	−0.051944	−0.051238	−0.050528	−0.049814	−0.049096	−0.048375	−0.047650	−0.046921	−0.046189	−0.045453

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

TABLE 1 (CONTINUED)

(sin u)yu

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
12.0	−0.044714	−0.043972	−0.043227	−0.042479	−0.041727	−0.040973	−0.040216	−0.039456	−0.038694	−0.037929
12.1	−0.037161	−0.036391	−0.035618	−0.034844	−0.034067	−0.033288	−0.032506	−0.031723	−0.030938	−0.030152
12.2	−0.029363	−0.028573	−0.027781	−0.026988	−0.026193	−0.025398	−0.024600	−0.023802	−0.023003	−0.022202
12.3	−0.021401	−0.020599	−0.019796	−0.018992	−0.018188	−0.017384	−0.016578	−0.015773	−0.014967	−0.014161
12.4	−0.013355	−0.012549	−0.011743	−0.010937	−0.010131	−0.009326	−0.008521	−0.007716	−0.006912	−0.006109
12.5	−0.005306	−0.004504	−0.003702	−0.002902	−0.002103	−0.001304	−0.000507	0.000289	0.001083	0.001877
12.6	0.002668	0.003459	0.004248	0.005035	0.005820	0.006603	0.007385	0.008164	0.008942	0.009717
12.7	0.010491	0.011262	0.012030	0.012797	0.013560	0.014321	0.015080	0.015836	0.016589	0.017339
12.8	0.018087	0.018831	0.019572	0.020311	0.021046	0.021778	0.022506	0.023231	0.023953	0.024671
12.9	0.025386	0.026097	0.026804	0.027507	0.028207	0.028903	0.029594	0.030282	0.030966	0.031645
13.0	0.032321	0.032992	0.033658	0.034321	0.034978	0.035632	0.036281	0.036925	0.037564	0.038199
13.1	0.038829	0.039454	0.040075	0.040690	0.041300	0.041905	0.042506	0.043101	0.043690	0.044275
13.2	0.044854	0.045428	0.045996	0.046559	0.047117	0.047669	0.048215	0.048756	0.049291	0.049820
13.3	0.050344	0.050861	0.051373	0.051879	0.052379	0.052873	0.053361	0.053843	0.054319	0.054788
13.4	0.055252	0.055709	0.056160	0.056605	0.057043	0.057476	0.057901	0.058321	0.058733	0.059140
13.5	0.059540	0.059933	0.060320	0.060700	0.061073	0.061440	0.061800	0.062154	0.062500	0.062840
13.6	0.063174	0.063500	0.063820	0.064132	0.064438	0.064737	0.065029	0.065314	0.065593	0.065864
13.7	0.066128	0.066385	0.066636	0.066879	0.067115	0.067344	0.067566	0.067781	0.067989	0.068190
13.8	0.068384	0.068570	0.068750	0.068922	0.069087	0.069245	0.069396	0.069540	0.069677	0.069806
13.9	0.069929	0.070044	0.070152	0.070253	0.070346	0.070433	0.070512	0.070584	0.070649	0.070707
14.0	0.070758	0.070801	0.070838	0.070867	0.070889	0.070904	0.070912	0.070913	0.070907	0.070893
14.1	0.070873	0.070846	0.070811	0.070770	0.070721	0.070666	0.070603	0.070534	0.070457	0.070374
14.2	0.070284	0.070186	0.070082	0.069971	0.069854	0.069729	0.069598	0.069460	0.069315	0.069163
14.3	0.069005	0.068840	0.068668	0.068490	0.068305	0.068114	0.067916	0.067712	0.067501	0.067283
14.4	0.067060	0.066829	0.066593	0.066350	0.066101	0.065845	0.065584	0.065316	0.065042	0.064762
14.5	0.064476	0.064183	0.063885	0.063581	0.063271	0.062954	0.062633	0.062305	0.061971	0.061632
14.6	0.061287	0.060936	0.060580	0.060218	0.059851	0.059478	0.059100	0.058717	0.058328	0.057933
14.7	0.057534	0.057129	0.056719	0.056304	0.055884	0.055459	0.055029	0.054594	0.054154	0.053710
14.8	0.053260	0.052806	0.052347	0.051884	0.051416	0.050944	0.050467	0.049985	0.049500	0.049010
14.9	0.048516	0.048017	0.047515	0.047008	0.046497	0.045983	0.045464	0.044942	0.044416	0.043886
15.0	0.043353	0.042815	0.042275	0.041730	0.041183	0.040632	0.040077	0.039520	0.038959	0.038395
15.1	0.037828	0.037257	0.036684	0.036108	0.035529	0.034948	0.034363	0.033776	0.033187	0.032595
15.2	0.032000	0.031403	6.030803	0.030202	0.029598	0.028992	0.028383	0.027773	0.027161	0.026547
15.3	0.025931	0.025313	0.024693	0.024072	0.023450	0.022825	0.022199	0.021572	0.020944	0.020314
15.4	0.019683	0.019051	0.018418	0.017783	0.017148	0.016512	0.015875	0.015237	0.014599	0.013960
15.5	0.013320	0.012680	0.012040	0.011399	0.010758	0.010116	0.009475	0.008833	0.008191	0.007549
15.6	0.006907	0.006266	0.005624	0.004983	0.004342	0.003702	0.003062	0.002422	0.001783	0.001145
15.7	0.000507	−0.000130	−0.000766	−0.001401	−0.002035	−0.002668	−0.003300	−0.003931	−0.004561	−0.005190
15.8	−0.005817	−0.006443	−0.007067	−0.007690	−0.008311	−0.008931	−0.009549	−0.010166	−0.010780	−0.011393
15.9	−0.012004	−0.012613	−0.013219	−0.013824	−0.014427	−0.015027	−0.015625	−0.016221	−0.016814	−0.017405

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

TABLE 1 (CONTINUED)

(sin u)/u

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
16.0	−0.017994	−0.018580	−0.019163	−0.019744	−0.020322	−0.020898	−0.021470	−0.022040	−0.022607	−0.023170
16.1	−0.023731	−0.024289	−0.024843	−0.025395	−0.025943	−0.026488	−0.027030	−0.027568	−0.028103	−0.028634
16.2	−0.029162	−0.029686	−0.030207	−0.030724	−0.031237	−0.031747	−0.032252	−0.032754	−0.033252	−0.033746
16.3	−0.034236	−0.034722	−0.035204	−0.035682	−0.036156	−0.036626	−0.037091	−0.037552	−0.038009	−0.038461
16.4	−0.038909	−0.039352	−0.039792	−0.040226	−0.040656	−0.041081	−0.041502	−0.041918	−0.042330	−0.042737
16.5	−0.043139	−0.043536	−0.043928	−0.044315	−0.044698	−0.045076	−0.045448	−0.045816	−0.046179	−0.046536
16.6	−0.046889	−0.047236	−0.047578	−0.047915	−0.048247	−0.048574	−0.048895	−0.049212	−0.049522	−0.049828
16.7	−0.050128	−0.050423	−0.050713	−0.050997	−0.051275	−0.051548	−0.051816	−0.052078	−0.052335	−0.052586
16.8	−0.052831	−0.053071	−0.053306	−0.053535	−0.053758	−0.053975	−0.054187	−0.054393	−0.054594	−0.054789
16.9	−0.054978	−0.055161	−0.055339	−0.055511	−0.055677	−0.055837	−0.055992	−0.056141	−0.056284	−0.056421
17.0	−0.056553	−0.056678	−0.056798	−0.056912	−0.057021	−0.057123	−0.057220	−0.057310	−0.057395	−0.057474
17.1	−0.057548	−0.057615	−0.057677	−0.057732	−0.057782	−0.057826	−0.057865	−0.057897	−0.057924	−0.057944
17.2	−0.057959	−0.057968	−0.057972	−0.057969	−0.057961	−0.057947	−0.057927	−0.057902	−0.057870	−0.057833
17.3	−0.057790	−0.057742	−0.057688	−0.057628	−0.057562	−0.057491	−0.057414	−0.057331	−0.057243	−0.057149
17.4	−0.057049	−0.056944	−0.056834	−0.056717	−0.056596	−0.056468	−0.056336	−0.056197	−0.056054	−0.055905
17.5	−0.055750	−0.055590	−0.055425	−0.055254	−0.055078	−0.054897	−0.054710	−0.054518	−0.054321	−0.054119
17.6	−0.053912	−0.053699	−0.053481	−0.053258	−0.053031	−0.052798	−0.052560	−0.052317	−0.052069	−0.051816
17.7	−0.051558	−0.051296	−0.051028	−0.050756	−0.050479	−0.050198	−0.049911	−0.049620	−0.049324	−0.049024
17.8	−0.048719	−0.048410	−0.048096	−0.047778	−0.047455	−0.047128	−0.046796	−0.046461	−0.046121	−0.045776
17.9	−0.045428	−0.045075	−0.044718	−0.044358	−0.043993	−0.043624	−0.043251	−0.042875	−0.042494	−0.042110
18.0	−0.041722	−0.041330	−0.040934	−0.040535	−0.040132	−0.039726	−0.039316	−0.038902	−0.038485	−0.038065
18.1	−0.037642	−0.037215	−0.036785	−0.036351	−0.035915	−0.035475	−0.035033	−0.034587	−0.034139	−0.033687
18.2	−0.033233	−0.032775	−0.032315	−0.031853	−0.031387	−0.030919	−0.030449	−0.029976	−0.029500	−0.029022
18.3	−0.028541	−0.028059	−0.027574	−0.027086	−0.026597	−0.026105	−0.025612	−0.025116	−0.024619	−0.024119
18.4	−0.023618	−0.023114	−0.022610	−0.022103	−0.021594	−0.021085	−0.020573	−0.020060	−0.019546	−0.019030
18.5	−0.018512	−0.017994	−0.017474	−0.016953	−0.016431	−0.015908	−0.015384	−0.014859	−0.014333	−0.013806
18.6	−0.013278	−0.012750	−0.012220	−0.011691	−0.011160	−0.010629	−0.010098	−0.009566	−0.009033	−0.008501
18.7	−0.007968	−0.007435	−0.006901	−0.006368	−0.005834	−0.005301	−0.004767	−0.004234	−0.003701	−0.003168
18.8	−0.002635	−0.002102	−0.001570	−0.001038	−0.000507	0.000024	0.000554	0.001083	0.001612	0.002140
18.9	0.002668	0.003194	0.003720	0.004245	0.004769	0.005292	0.005813	0.006334	0.006853	0.007371
19.0	0.007888	0.008404	0.008918	0.009431	0.009942	0.010452	0.010960	0.011466	0.011971	0.012474
19.1	0.012976	0.013475	0.013973	0.014468	0.014962	0.015454	0.015944	0.016431	0.016917	0.017400
19.2	0.017881	0.018360	0.018836	0.019310	0.019782	0.020251	0.020717	0.021181	0.021643	0.022102
19.3	0.022558	0.023011	0.023462	0.023910	0.024355	0.024797	0.025236	0.025672	0.026105	0.026535
19.4	0.026962	0.027386	0.027807	0.028224	0.028638	0.029049	0.029457	0.029861	0.030262	0.030659
19.5	0.031053	0.031444	0.031831	0.032214	0.032594	0.032970	0.033342	0.033711	0.034076	0.034437
19.6	0.034794	0.035148	0.035497	0.035843	0.036185	0.036522	0.036856	0.037186	0.037512	0.037833
19.7	0.038151	0.038464	0.038774	0.039079	0.039379	0.039676	0.039968	0.040256	0.040540	0.040820
19.8	0.041095	0.041365	0.041632	0.041893	0.042151	0.042404	0.042652	0.042896	0.043135	0.043370
19.9	0.043600	0.043826	0.044047	0.044263	0.044475	0.044682	0.044885	0.045082	0.045275	0.045464

Adapted from L. Levi, *Applied Optics*.

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Solutions to Selected Problems

Chapter 2

2.1 $(0.003)(2.54 \times 10^{-2})/580 \times 10^{-9}$ = number of waves = 131; $c = \nu\lambda$, $\lambda = c/\nu = 3 \times 10^8/10^{10}$, $\lambda = 3$ cm. Waves extend 3.9 m.

2.6 $v = \nu\lambda = 1498$ m/s $= (440\text{ Hz})\lambda$; $\lambda = 3.40$ m.

2.16 $\psi = A \sin 2\pi(\kappa x - \nu t)$, $\psi_1 = 4 \sin 2\pi(0.2x - 3t)$

(a) $\nu = 3$ (b) $\lambda = 1/0.2$ (c) $\tau = 1/3$
(d) $A = 4$ (e) $v = 15$ (f) positive x

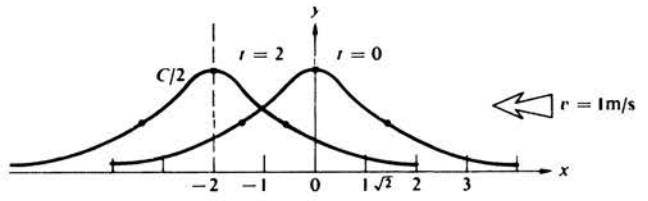
$\psi = A \sin(kx + \omega t)$, $\psi_2 = (1/2.5)\sin(7x + 3.5t)$

(a) $\nu = 3.5/2\pi$ (b) $\lambda = 2\pi/7$ (c) $\tau = 2\pi/3.5$
(d) $A = 1/2.5$ (e) $v = \frac{1}{2}$ (f) negative x

2.21 $v_y = -\omega A \cos(kx - \omega t + \varepsilon)$, $a_y = -\omega^2 y$. Simple harmonic motion since $a_y \propto y$.

2.22 $\tau = 2.2 \times 10^{-15}$ s; therefore $\nu = 1/\tau = 4.5 \times 10^{14}$ Hz; $v = \nu\lambda$, 3×10^8 m/s $= (4.5 \times 10^{14}\text{ Hz})\lambda$; $\lambda = 6.6 \times 10^{-7}$ m and $k = 2\pi/\lambda = 9.5 \times 10^6\text{ m}^{-1}$. $\psi(x, t) = (10^3\text{ V/m}) \cos[9.5 \times 10^6\text{ m}^{-1} \times (x + 3 \times 10^8\text{ m/s}\, t)]$. It's cosine because $\cos 0 = 1$.

2.23 $y(x, t) = C/[2 + (x + vt)^2]$.



2.25 No, not twice differentiable (in a nontrivial way) and not a solution of the differential wave equation.

2.27 $\frac{d\psi}{dt} = \frac{\partial\psi}{\partial x}\frac{dx}{dt} + \frac{\partial\psi}{\partial y}\frac{dy}{dt}$ and let $y = t$, whereupon $\frac{d\psi}{dt} = \frac{\partial\psi}{\partial x}(\pm v) + \frac{\partial\psi}{\partial t} = 0$ and the desired result follows immediately.

2.28 $\frac{d\varphi}{dt} = \frac{\partial\varphi}{\partial x}\frac{dx}{dt} + \frac{\partial\varphi}{\partial t} = 0 = k\frac{dx}{dt} - kv$, and this is zero provided $\frac{dx}{dt} = \pm v$, as it should be. For the particular wave of Problem 2.20, $\frac{d\varphi}{dt} = \frac{\partial\varphi}{\partial y}(\pm v) + \frac{\partial\varphi}{\partial t} = \pi^3 \times 10^6(\pm v) + \pi^9 \times 10^{14} = 0$ and the speed is -3×10^8 m/s.

2.30

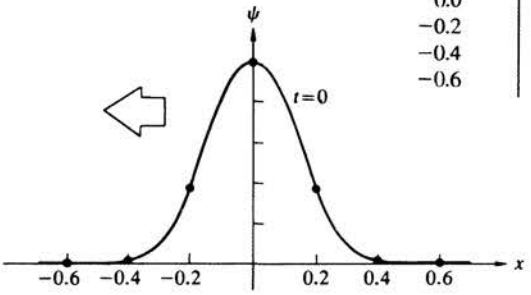
$$\psi(z, 0) = A \sin(kz + \varepsilon);$$
$$\psi(-\lambda/12, 0) = A \sin(-\pi/6 + \varepsilon) = 0.866$$
$$\psi(\lambda/6, 0) = A \sin(\pi/3 + \varepsilon) = 1/2$$
$$\psi(\lambda/4, 0) = A \sin(\pi/2 + \varepsilon) = 0$$
$$A \sin(\pi/2 + \varepsilon) = A(\sin \pi/2 \cos\varepsilon + \cos\pi/2 \sin\varepsilon)$$
$$= A \cos\varepsilon = 0,\ \varepsilon = \pi/2$$
$$A \sin(\pi/3 + \pi/2) = A \sin(5\pi/6) = 1/2$$

therefore $A = 1$, hence $\psi(z, 0) = \sin(kz + \pi/2)$.

2.31 Both (a) and (b) are waves since they are twice differentiable functions of $(z - vt)$ and $(x + vt)$, respectively. Thus for (a) $\psi = a^2(z - bt/a)^2$ and the velocity is b/a in the positive z-direction. For (b) $\psi = a^2(x + bt/a + c/a)^2$ and the velocity is b/a in the negative x-direction.

2.33 $\psi(x, t) = 5.0 \exp[-a(x + \sqrt{b/a}t)^2]$, the propagation direction is negative x; $v = \sqrt{b/a} = 0.6$ m/s. $\psi(x, 0) = 5.0 \exp(-25x^2)$;

x	ψ
0.6	0.0006
0.4	0.09
0.2	1.8
0.0	5.0
−0.2	1.8
−0.4	0.09
−0.6	0.0006



Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

2.35 30° corresponds to $\frac{1}{12}\lambda$ or $(1/12)3 \times 10^8/6 \times 10^{14} = 42$ nm.

2.36
$$\psi = A \sin 2\pi\left(\frac{z}{\lambda} \pm \frac{t}{\tau}\right)$$
$$\psi = 60 \sin 2\pi\left(\frac{z}{400 \times 10^{-9}} - \frac{t}{1.33 \times 10^{-15}}\right)$$
$$\lambda = 400 \text{ nm}$$
$$v = 400 \times 10^{-9}/1.33 \times 10^{-15} = 3 \times 10^8 \text{ m/s}$$
$$\nu = (1/1.33) \times 10^{+15} \text{ Hz}, \quad \tau = 1.33 \times 10^{-15} \text{ s}.$$

2.39
$$\psi = A \exp i(k_x x + k_y y + k_z z)$$
$$k_x = k\alpha \qquad k_y = k\beta \qquad k_z = k\gamma$$
$$|\vec{\mathbf{k}}| = [(k\alpha)^2 + (k\beta)^2 + (k\gamma)^2]^{1/2} = k[\alpha^2 + \beta^2 + \gamma^2]^{1/2}$$

2.41 $\lambda = h/mv = 6.6 \times 10^{-34}/6(1) = 1.1 \times 10^{-34}$ m

2.42 $\vec{\mathbf{k}}$ can be constructed by forming a unit vector in the proper direction and multiplying it by k. The unit vector is

$$[(4-0)\hat{\mathbf{i}} + (2-0)\hat{\mathbf{j}} + (1-0)\hat{\mathbf{k}}]/\sqrt{4^2 + 2^2 + 1^2}$$
$$= (4\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}})/\sqrt{21}$$

and $\vec{\mathbf{k}} = k(4\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}})/\sqrt{21}$.

$$\vec{\mathbf{r}} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$$

hence $\psi(x, y, z, t) = A \sin [(4k/\sqrt{21})x$
$$+ (2k/\sqrt{21})y + (k/\sqrt{21})z - \omega t].$$

2.44
$$\psi(\vec{r}_1, t) = \psi[\vec{r}_2 - (\vec{r}_2 - \vec{r}_1), t] = \psi(\vec{k} \cdot \vec{r}_1, t)$$
$$= \psi[\vec{k} \cdot \vec{r}_2 - \vec{k} \cdot (\vec{r}_2 - \vec{r}_1), t]$$
$$= \psi(\vec{k} \cdot \vec{r}_2, t) = \psi(\vec{r}_2, t)$$

since $\vec{k} \cdot (\vec{r}_2 - \vec{r}_1) = 0$

Chapter 3

3.1
$$E_y = 2 \cos [2\pi \times 10^{14}(t - x/c) + \pi/2]$$
$$E_y = A \cos [2\pi\nu(t - x/v) + \pi/2] \quad \text{from Eq. (2.26)}$$

(a) $\nu = 10^{14}$ Hz, $v = c$, and $\lambda = c/\nu = 3 \times 10^8/10^{14} = 3 \times 10^{-6}$ m, moves in positive x-direction, $A = 2$ V/m, $\varepsilon = \pi/2$ linearly polarized in y-direction.

(b) $B_x = 0, \quad B_y = 0, \quad B_z = \frac{2}{c} \cos [2\pi \times 10^{14}(t - x/c) + \pi/2]$.

3.2 $E_z = 0$, $E_y = E_x = E_0 \sin(kz - \omega t)$ or cosine; $B_z = 0$, $B_y = -B_x = E_y/c$, or if you like,

$$\vec{\mathbf{E}} = \frac{E_0}{\sqrt{2}}(\hat{\mathbf{i}} + \hat{\mathbf{j}}) \sin(kz - \omega t), \quad \vec{\mathbf{B}} = \frac{E_0}{c\sqrt{2}}(\hat{\mathbf{j}} - \hat{\mathbf{i}}) \sin(kz - \omega t).$$

3.6 The field is linearly polarized in the y-direction and varies sinusoidally from zero at $z = 0$ to zero at $z = z_0$. Using the wave equation

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial y^2} + \frac{\partial^2 E_y}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 E_y}{\partial t^2} = 0$$
$$\left[-k^2 - \frac{\pi^2}{z_0^2} + \frac{\omega^2}{c^2}\right] E_0 \sin\frac{\pi z}{z_0} \cos(kx - \omega t) = 0$$

and since this is true for all x, z, and t, each term must equal zero

and so $k = \frac{\omega}{c}\sqrt{1 - \left(\frac{c\pi}{\omega z_0}\right)^2}$

Moreover, $v = \frac{\omega}{k} = \frac{c}{\sqrt{1 - \left(\frac{c\pi}{\omega z_0}\right)^2}}$

3.10 $\langle \cos^2(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t)\rangle = \frac{1}{T}\int_t^{t+T} \cos^2(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t')\, dt'$.

Let $\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - vt' = x$; then

$$\langle \cos^2(\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t)\rangle = \frac{1}{-\omega T}\int \cos^2 x\, dx$$
$$= \frac{1}{-\omega T}\int \frac{1 + \cos 2x}{2}\, dx$$
$$= -\frac{1}{\omega T}\left[\frac{x}{2} + \frac{\sin 2x}{4}\right]_{\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega t}^{\vec{\mathbf{k}} \cdot \vec{\mathbf{r}} - \omega(t+T)}$$

3.17 $\vec{\mathbf{E}}_0 = (-E_0/\sqrt{2})\hat{\mathbf{i}} + (E_0/\sqrt{2})\hat{\mathbf{j}}$; $\vec{\mathbf{k}} = (2\pi/\lambda)(\hat{\mathbf{i}}/\sqrt{2} + \hat{\mathbf{j}}\sqrt{2})$, hence $\vec{\mathbf{E}} = (1/\sqrt{2})(-10\hat{\mathbf{i}} + 10\hat{\mathbf{j}}) \cos [(\sqrt{2}\pi/\lambda)(x + y) - \omega t]$ and $I = \frac{1}{2}c\epsilon_0 E_0^2 = 0.13$ W/m^2.

3.18

(a) $l = c\Delta t = (3.00 \times 10^8 \text{ m/s})(2.00 \times 10^{-9} \text{ s}) = 0.600$ m.

(b) The volume of one pulse is $(0.600 \text{ m})(\pi R^2) = 2.945 \times 10^{-6}$ m^3; therefore $(6.0 \text{ J})/(2.945 \times 10^{-6} \text{ m}^3) = 2.0 \times 10^6$ J/m^3.

3.20
$$u = \frac{(\text{power})(t)}{(\text{volume})} = \frac{(10^{-3} \text{ W})(t)}{(\pi r^2)(ct)} = \frac{10^{-3} \text{ W}}{\pi(10^{-3})^2(3 \times 10^8)}$$
$$u = \frac{10^{-5}}{3\pi} \text{ J/m}^3 = 1.06 \times 10^{-6} \text{ J/m}^3$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

3.22 $h = 6.63 \times 10^{-34}$, $E = h\nu$

$$\frac{I}{h\nu} = \frac{19.88 \times 10^{-2}}{(6.63 \times 10^{-34})(100 \times 10^{6})}$$

$$= 3 \times 10^{24} \text{ photons/m}^2 \text{ s}$$

All photons in volume V cross the unit area in one second

$$V = (ct)(1 \text{ m}^2) = 3 \times 10^8 \text{ m}^3$$

$$3 \times 10^{24} = V(\text{density})$$

$$\text{density} = 10^{16} \text{ photons/m}^3$$

3.24 $P_e = iV = (0.25)(3.0) = 0.75$ W. This is the electrical power dissipated. The power available as light is $P_l = (0.01)P_e = 75 \times 10^{-4}$ W.

(a) Photon flux

$$= P_l/h\nu = 75 \times 10^{-4}\,\lambda/hc$$

$$= 75 \times 10^{-4}(550 \times 10^{-9})/(6.63 \times 10^{-34})3 \times 10^8$$

$$= 2.08 \times 10^{16} \text{ photons/s}$$

(b) There are 2.08×10^{16} in volume $(3 \times 10^8)(1\text{s})(10^{-3} \text{ m}^2)$;

$$\therefore \quad \frac{2.08 \times 10^{16}}{3 \times 10^5} = \text{photons/m}^3 = 0.69 \times 10^{11}$$

(c) $I = 75 \times 10^{-4} \text{ W}/10 \times 10^{-4} \text{ m}^2 = 7.5 \text{ W/m}^2$

3.26 Imagine two concentric cylinders of radius r_1 and r_2 surrounding the wave. The energy flowing per second through the first cylinder must pass through the second cylinder; that is, $\langle S_1 \rangle 2\pi r_1 = \langle S_2 \rangle 2\pi r_2$, and so $\langle S \rangle 2\pi r$ = constant and $\langle S \rangle$ varies inversely with r. Therefore, since $\langle S \rangle \propto E_0^2$, E_0 varies as $\sqrt{1/r}$.

3.28

$$\left\langle \frac{dp}{dt} \right\rangle = \frac{1}{c}\left\langle \frac{dW}{dt} \right\rangle$$

A = area.

$$\langle \mathcal{P} \rangle = \frac{1}{A}\left\langle \frac{dp}{dt} \right\rangle = \frac{1}{Ac}\left\langle \frac{dW}{dt} \right\rangle = \frac{I}{c}$$

3.30

$$\mathscr{E} = 300 \text{ W}(100 \text{ s}) = 3 \times 10^4 \text{ J}$$

$$p = \mathscr{E}/c = 3 \times 10^4/3 \times 10^8 = 10^{-4} \text{ kg} \cdot \text{m/s}$$

3.31

(a) $\langle \mathcal{P} \rangle = 2\langle S \rangle/c = 2(1.4 \times 10^3 \text{ W/m}^2)/(3 \times 10^8 \text{ m/s}) = 9 \times 10^{-6} \text{ N/m}^2$.

(b) S, and therefore $\mathcal{P}$, drops off with the inverse square of the distance, and hence $\langle S \rangle = [(0.7 \times 10^9 \text{ m})^{-2}/(1.5 \times 10^{11} \text{ m})^{-2}] \times (1.4 \times 10^3 \text{ W/m}^2) = 6.4 \times 10^7 \text{ W/m}^2$, and $\langle \mathcal{P} \rangle = 0.21 \text{ N/m}^2$.

3.34

$$\langle S \rangle = 1400 \text{ W/m}^2$$

$$\langle \mathcal{P} \rangle = 2(1400 \text{ W/m}^2/3 \times 10^8 \text{ m/s}) = 9.3 \times 10^{-6} \text{ N/m}^2$$

$$\langle F \rangle = A\langle \mathcal{P} \rangle = 2000 \text{ m}^2(9.3 \times 10^{-6} \text{ N/m}^2) = 1.9 \times 10^{-2} \text{ N}$$

3.35

$$\langle S \rangle = (200 \times 10^3 \text{ W})(500 \times 2 \times 10^{-6} \text{ s})/A(1s)$$

$$\langle F \rangle = A\langle \mathcal{P} \rangle = A\langle S \rangle/c = 6.7 \times 10^{-7} \text{ N}$$

3.36

$$\langle F \rangle = A\langle \mathcal{P} \rangle = A\langle S \rangle/c = \frac{10 \text{ W}}{3 \times 10^8} = 3.3 \times 10^{-8} \text{ N}$$

$$a = 3.3 \times 10^{-8}/100 \text{ kg} = 3.3 \times 10^{-10} \text{ m/s}^2$$

$$v = at = \tfrac{1}{3} \times 10^{-9}(t) = 10 \text{ m/s}$$

$$t = 3 \times 10^{10} \text{ s} \qquad 1 \text{ year} = 3.2 \times 10^7 \text{ s}$$

3.37 $\vec{\mathbf{B}}$ surrounds $\vec{\mathbf{v}}$ in circles, and $\vec{\mathbf{E}}$ is radial; hence $\vec{\mathbf{E}} \times \vec{\mathbf{B}}$ is tangent to the sphere, and no energy radiates outward from it.

3.42 $n = c/v = (2.998 \times 10^8 \text{ m/s})/(1.245 \times 10^8 \text{ m/s}) = 2.41$

3.47 Thermal agitation of the molecular dipoles causes a marked reduction in K_e but has little effect on n. At optical frequencies n is predominantly due to electronic polarization, rotations of the molecular dipoles having ceased to be effective at much lower frequencies.

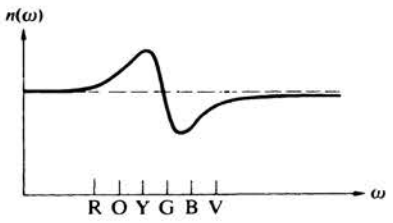
3.48 From Eq. (3.70), for a single resonant frequency we get

$$n = \left[1 + \frac{Nq_e^2}{\epsilon_0 m_e}\left(\frac{1}{\omega_0^2 - \omega^2}\right)\right]^{1/2}$$

since for low-density materials $n \approx 1$, the second term is $<< 1$, and we need only retain the first two terms of the binomial expansion of n. Thus $\sqrt{1 + x} \approx 1 + x/2$ and

$$n = 1 + \frac{1}{2}\frac{Nq_e^2}{\epsilon_0 m_e}\left(\frac{1}{\omega_0^2 - \omega^2}\right)$$

3.50 The normal order of the spectrum for a glass prism is R, O, Y, G, B, V, with red (R) deviated the least and violet (V) deviated the most. For a fuchsin prism, there is an absorption band in the green, and so the indices for yellow and blue on either side (n_Y and n_B) of it are extremes, as in Fig. 3.26; that is, n_Y is the maximum, n_B the minimum, and $n_Y > n_O > n_R > n_V > n_B$. Thus the spectrum in order of increasing deviation is B, V, black band, R, O, Y.



Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

3.52 With ω in the visible, $(\omega_0^2 - \omega^2)$ is smaller for lead glass and larger for fused silica. Hence $n(\omega)$ is larger for the former and smaller for the latter.

3.54 C_1 is the value that n approaches as λ gets larger.

3.55 The horizontal values of $n(\omega)$ approached in each region between absorption bands increase as ω decreases.

Chapter 4

4.1 $E_{0s} \propto \dfrac{VE_{0i}}{r} = K\dfrac{VE_{0i}}{r}$; thus $\dfrac{VK}{r}$ must be unitless, and so K has units of $(length)^{-2}$. The only quantity unaccounted for is λ, and so we conclude that $K = \lambda^{-2}$, and $I_i/I_s \propto K^2 \propto \lambda^{-4}$.

4.4 $x_0(-\omega^2 + \omega_0^2 + i\gamma\omega) = (q_eE_0/m_e)e^{i\alpha} = (q_eE_0/m_e)(\cos\ \alpha + i\sin\alpha)$; squaring both sides yields $x_0^2[(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2] = (q_eE_0/m_e)^2(\cos^2\alpha + \sin^2\alpha)$. x_0 follows immediately. As for α, divide the imaginary parts of both sides of the first equation above, namely, $x_0\gamma\omega = (q_eE_0/m_e)\sin\alpha$, by the real parts, $x_0(\omega_0^2 - \omega^2) = (q_eE_0/m_e)\cos\alpha$ to get $\alpha = \tan^{-1}[\gamma\omega/(\omega_0^2 - \omega^2)]$. α ranges continuously from 0 to $\pi/2$ to π.

4.5 The phase angle is retarded by an amount $(n\ \Delta y\ 2\pi/\lambda) - \Delta y 2\pi/\lambda$ or $(n-1)\,\Delta y\,\omega/c$. Thus

$$E_p = E_0 \exp i\omega[t - (n-1)\Delta y/c - y/c]$$

or $$E_p = E_0 \exp[-i\omega(n-1)\Delta y/c] \exp i\omega(t - y/c)$$

if $n \approx 1$ or $\Delta y << 1$. Since $e^x \approx 1 + x$ for small x,

$$\exp[-i\omega(n-1)\Delta y/c] \approx 1 - i\omega(n-1)\Delta y/c$$

and since $\exp(-i\pi/2) = -i$,

$$E_p = E_u + \frac{\omega(n-1)\Delta y}{c} E_u e^{-i\pi/2}$$

4.9

$$n_i \sin\theta_i = n_t \sin\theta_t$$
$$\sin 30° = 1.52 \sin\theta_t$$
$$\theta_t = \sin^{-1}(1/3.04)$$
$$\theta_t = 19°\ 13'$$

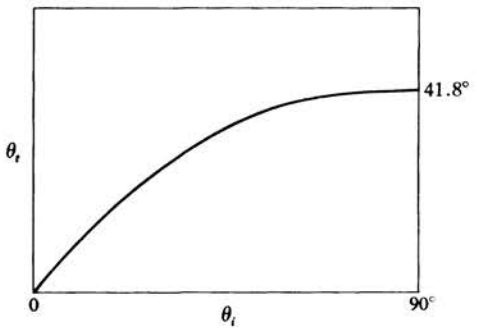
4.14 $n_{ti} = \dfrac{n_t}{n_i} = \dfrac{c/v_t}{c/v_i} = \dfrac{v_i}{v_t} = \dfrac{\nu\lambda_i}{\nu\lambda_t} = \dfrac{\lambda_i}{\lambda_t}$

therefore $\lambda_t = \lambda_i 3/4 = 9$ cm

$$\sin\theta_i = n_{ti}\sin\theta_t$$
$$\sin^{-1}[\tfrac{3}{4}(0.707)] = \theta_t = 32°$$

4.18

θ_i (degrees)	θ_t (degrees)
0	0
10	6.7
20	13.3
30	19.6
40	25.2
50	30.7
60	35.1
70	38.6
80	40.6
90	41.8



4.26 The number of waves per unit length along $\overline{AC}$ on the interface equals $(\overline{BC}/\lambda_i)/\overline{BC}\sin\theta_i) = (\overline{AD}/\lambda_t)(\overline{AD}/\sin\theta_t)$. Snell's Law follows on multiplying both sides by c/ν.

4.28 Let τ be the time for the wave to move along a ray from b_1 to b_2, from a_1 to a_2, and from a_1 to a_3. Thus $\overline{a_1a_2} = \overline{b_1b_2} = v_i\tau$ and $\overline{a_1a_3} = v_t\tau$.

$$\sin\theta_i = \overline{b_1b_2}/\overline{a_1b_2} = v_i/\overline{a_1b_2}$$
$$\sin\theta_t = \overline{a_1a_3}/\overline{a_1b_2} = v_t/\overline{a_1b_2}$$
$$\sin\theta_r = \overline{a_1a_2}/\overline{a_1b_2} = v_i/\overline{a_1b_2}$$
$$\frac{\sin\theta_i}{\sin\theta_t} = \frac{v_i}{v_t} = \frac{n_t}{n_i} = n_{ti} \quad \text{and} \quad \theta_i = \theta_r$$

4.29

$$n_i \sin\theta_i = n_t \sin\theta_t$$
$$n_i(\hat{\mathbf{k}}_i \times \hat{\mathbf{u}}_n) = n_t(\hat{\mathbf{k}}_t \times \hat{\mathbf{u}}_n)$$

where $\hat{\mathbf{k}}_i$, $\hat{\mathbf{k}}_t$ are unit propagation vectors. Thus

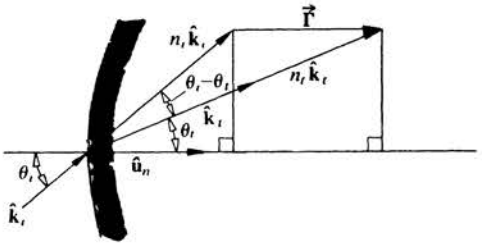
$$n_t(\hat{\mathbf{k}}_t \times \hat{\mathbf{u}}_n) - n_i(\hat{\mathbf{k}}_i \times \hat{\mathbf{u}}_n) = 0$$
$$(n_t\hat{\mathbf{k}}_t - n_i\hat{\mathbf{k}}_i) \times \hat{\mathbf{u}}_n = 0$$

Let $n_t\hat{\mathbf{k}}_t - n_i\hat{\mathbf{k}}_i = \vec{\Gamma} = \Gamma\hat{\mathbf{u}}_n$

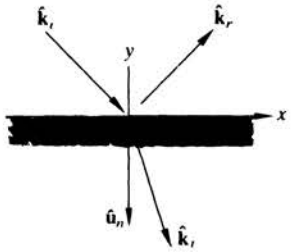
Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

Γ is often referred to as the *astigmatic constant*: $\vec{\Gamma}$ = the difference between the projections of $n_t\hat{\mathbf{k}}_t$ and $n_i\hat{\mathbf{k}}_i$ on $\hat{\mathbf{u}}_n$; in other words, take dot product $\vec{\Gamma} \cdot \hat{\mathbf{u}}_n$:

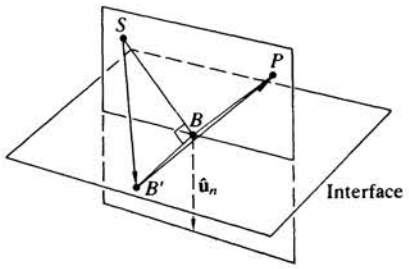
$$\Gamma = n_t \cos\theta_t - n_i \cos\theta_i$$

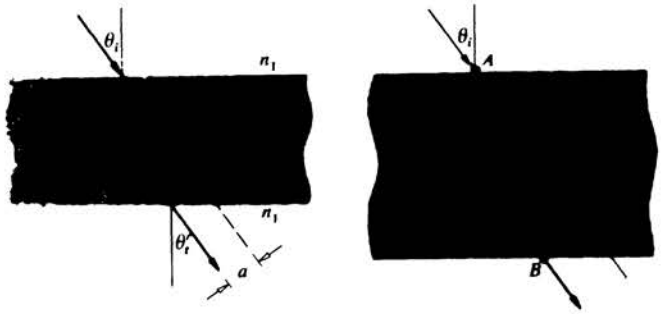


4.30 Since $\theta_i = \theta_r$, $\hat{\mathbf{k}}_{ix} = \hat{\mathbf{k}}_{rx}$ and $\hat{\mathbf{k}}_{iy} = -\hat{\mathbf{k}}_{ry}$, and since $(\hat{\mathbf{k}}_i \cdot \hat{\mathbf{u}}_n)\hat{\mathbf{u}}_n = \hat{\mathbf{k}}_{iy}$, $\hat{\mathbf{k}}_i - \hat{\mathbf{k}}_r = 2(\hat{\mathbf{k}}_i \cdot \hat{\mathbf{u}}_n)\hat{\mathbf{u}}_n$.



4.31 Since $\overline{SB'} > \overline{SB}$ and $\overline{B'P} > \overline{BP}$, the shortest path corresponds to B' coincident with B in the plane-of-incidence.





4.34

$$n_1 \sin\theta_i = n_2 \sin\theta_t \qquad \theta_t = \theta_i'$$

$$n_2 \sin\theta_i' = n_1 \sin\theta_t'$$

$$n_1 \sin\theta_i = n_1 \sin\theta_t' \quad \text{and} \quad \theta_i = \theta_t'$$

$$\cos\theta_t = d/\overline{AB}$$

$$\sin(\theta_i - \theta_t) = a/\overline{AB}$$

$$\sin(\theta_i - \theta_t) = \frac{a}{d}\cos\theta_t$$

$$\frac{d\sin(\theta_i - \theta_t)}{\cos\theta_t} = a$$

4.36 Rather than propagating from $\vec{\mathbf{B}}$ point S to point P in a straight line, the ray traverses a path that crosses the plate at a sharper angle. Although in so doing the path lengths in air are slightly increased, the decrease in time spent within the plate more than compensates. This being the case, we might expect the displacement a to increase with n_{21}. As n_{21} gets larger for a given θ_i, θ_t decreases, $(\theta_i - \theta_t)$ increases, and from the results of Problem 4.30, a clearly increases.

4.37

$$\oint_C \vec{\mathbf{E}} \cdot d\vec{\ell} = -\iint_A \frac{\partial \vec{\mathbf{B}}}{\partial t} \cdot d\vec{\mathbf{S}} \qquad [3.5]$$

This reduces in the limit to $E_{2x}(\overline{BC}) - E_{1x}(\overline{AD}) = 0$, since Area $\to 0$ and $\partial\vec{\mathbf{B}}/\partial t$ is finite. Thus $E_{2x} = E_{1x}$.

4.38 From Eq. (4.40)

$$r_{\parallel} = \frac{1.52\cos 30° - \cos 19°13'}{\cos 19°13' + 1.52\cos 30°}$$

where from Problem 4.9 $\theta_t = 19°13'$. Similarly

$$t_{\parallel} = \frac{2\cos 30}{\cos 19°13' + 1.52\cos 30°}$$

$$r_{\parallel} = \frac{1.32 - 0.944}{0.944 + 1.32} = 0.165$$

$$t_{\parallel} = \frac{1.732}{0.944 + 1.32} = 0.766$$

4.39 Starting with Eq. (4.34), divide top and bottom by n_i and replace n_{it} with $\sin\theta_i/\sin\theta_t$ to get

$$r_{\perp} = \frac{\sin\theta_t\cos\theta_i - \sin\theta_i\cos\theta_t}{\sin\theta_t\cos\theta_i + \sin\theta_i\cos\theta_t}$$

which is equivalent to Eq. (4.42). Equation (4.44) follows in exactly the same way. To find $r_{\parallel}$ start the same way with Eq. (4.40) and get

$$r_{\parallel} = \frac{\sin\theta_i\cos\theta_i - \cos\theta_t\sin\theta_t}{\cos\theta_t\sin\theta_t + \sin\theta_i\cos\theta_i}$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

There are several routes that can be taken now: one is to rewrite $r_\parallel$ as

$$r_\parallel = \frac{(\sin\theta_i\cos\theta_t - \sin\theta_t\cos\theta_i)(\cos\theta_i\cos\theta_t - \sin\theta_i\sin\theta_t)}{(\sin\theta_i\cos\theta_t + \sin\theta_t\cos\theta_i)(\cos\theta_i\cos\theta_t + \sin\theta_i\sin\theta_t)}$$

and so $r_\parallel = \dfrac{\sin(\theta_i - \theta_t)\cos(\theta_i + \theta_t)}{\sin(\theta_i + \theta_t)\cos(\theta_i - \theta_t)} = \dfrac{\tan(\theta_i - \theta_t)}{\tan(\theta_i + \theta_t)}$

We can find $t_\parallel$, which has the same denominator, in a similar way.

4.50 $[E_{0r}]_\perp + [E_{0i}]_\perp = [E_{0t}]_\perp$; tangential field in incident medium equals that in transmitting medium,

$$[E_{0t}/E_{0i}]_\perp - [E_{0r}/E_{0i}]_\perp = 1, \qquad t_\perp - r_\perp = 1$$

Alternatively, from Eqs. (4.42) and (4.44),

$$\frac{+\sin(\theta_i - \theta_t) + 2\sin\theta_t\cos\theta_i}{\sin(\theta_i + \theta_t)} \stackrel{?}{=} 1$$

$$\frac{\sin\theta_i\cos\theta_t - \cos\theta_i\sin\theta_t + 2\sin\theta_t\cos\theta_i}{\sin\theta_i\cos\theta_t + \cos\theta_i\sin\theta_t} = 1$$

4.59 From Eq. (4.73) we see that the exponential will be in the form $k(x - vt)$, provided that we factor out $k_t \sin\theta_i/n_{ti}$, leaving the second term as $\omega n_{ti}t/k_t \sin\theta_i$, which must be $v_t t$. Hence $\omega n_t/(2\pi/\lambda_t)n_i \times \sin\theta_i = v_t$ and so $v_t = c/n_i \sin\theta_t = v_i/\sin\theta_t$.

4.60 From the defining equation $\beta = k_t[(\sin^2\theta_i/n_{ti}^2) - 1]^{1/2} = 3.702 \times 10^6\ \mathrm{m}^{-1}$, and since $y\beta = 1$, $y = 2.7 \times 10^{-7}$ m.

4.61 The beam scatters off the wet paper and is mostly transmitted until the critical angle is attained, at which point the light is reflected back toward the source. $\tan\theta_c = (R/2)/d$, and so $n_{ti} = 1/n_i = \sin[\tan^{-1}(R/2d)]$.

4.62 $1.000\,29 \sin 88.7° = n \sin 90°$

$(1.000\,29)(0.999\,74) = n; \qquad n = 1.000\,03$

4.64

$$\theta_i + \theta_t = 90° \text{ when } \theta_i = \theta_p$$

$$n_i \sin\theta_p = n_t \sin\theta_t = n_t \cos\theta_p$$

$$\tan\theta_p = n_t/n_i = 1.52, \qquad \theta_p = 56°40' \qquad [8.25]$$

4.66

$$\tan\theta_p = n_t/n_i = n_2/n_1$$

$$\tan\theta_p' = n_1/n_2, \qquad \tan\theta_p = 1/\tan\theta_p'$$

$$\frac{\sin\theta_p}{\cos\theta_p} = \frac{\cos\theta_p'}{\sin\theta_p'} \therefore \sin\theta_p\sin\theta_p' - \cos\theta_p\cos\theta_p' = 0$$

$$\cos(\theta_p + \theta_p') = 0, \qquad \theta_p + \theta_p' = 90°$$

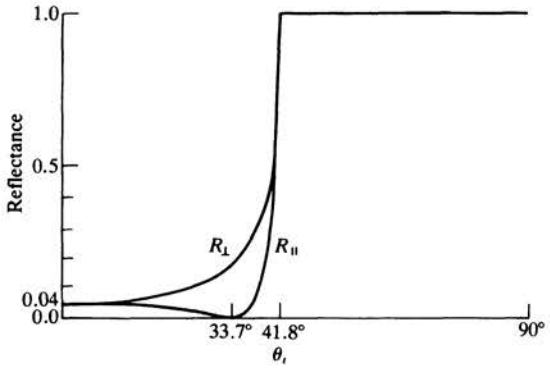
4.67 From Eq. (4.92)

$$\tan\gamma_r = r_\perp[E_{0i}]_\perp/r_\parallel[E_{0i}]_\parallel = \frac{r_\perp}{r_\parallel}\tan\gamma_i$$

and from Eqs. (4.42) and (4.43)

$$\tan\gamma_r = -\frac{\cos(\theta_i - \theta_t)}{\cos(\theta_i + \theta_t)}\tan\gamma_i$$

4.69



4.70 $T_\perp = \left(\dfrac{n_t\cos\theta_t}{n_i\cos\theta_i}\right)t_\perp^2$. From Eq. (4.44) and Snell's Law,

$$T_\perp = \left(\frac{\sin\theta_i\cos\theta_t}{\sin\theta_t\cos\theta_i}\right)\left(\frac{4\sin^2\theta_t\cos^2\theta_i}{\sin^2(\theta_i + \theta_t)}\right) = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2(\theta_i + \theta_t)}$$

Similarly for $T_\parallel$.

4.72 If Φ_i is the incident radiant flux or power and T is the transmittance across the first air–glass boundary, the transmitted flux is the $T\Phi_i$. From Eq. (4.68) at normal incidence the transmittance from glass to air is also T. Thus a flux $T\Phi_i T$ emerges from the first slide, and $\Phi_i T^{2N}$ from the last one. Since $T = 1 - R$, $T_t = (1 - R)^{2N}$ from Eq. (4.67).

$$R = (0.5/2.5)^2 = 4\%, \qquad T = 96\%$$

$$T_t = (0.96)^6 \approx 78.3\%$$

4.73 $T = \dfrac{I(y)}{I_0} = e^{-\alpha y}, \qquad T_1 = e^{-\alpha}, \qquad T = (T_1)^y$

$$T_t = (1 - R)^{2N}(T_1)^d$$

4.74 At $\theta_i = 0$, $R = R_\parallel = R_\perp = \left(\dfrac{n_t - n_i}{n_t + n_i}\right)^2$ [4.67]

As $n_{ti} \to 1$, $n_t \to n_i$ and clearly $R \to 0$.
At $\theta_i = 0$,

$$T = T_\parallel = T_\perp \frac{4n_t n_i}{(n_t + n_i)^2}$$

and since $n_t \to n_i$, $\lim_{n_{ti}\to 1} T = 4n_i^2/(2n_i)^2 = 1$.
From Problem 4.61, and the fact that as $n_t \to n_i$ Snell's Law says that $\theta_t \to \theta_i$, we have

$$\lim_{n_{ti}\to 1} T_\parallel = \frac{\sin^2 2\theta_i}{\sin^2 2\theta_i} = 1, \qquad \lim_{n_{ti}\to 1} T_\perp = 1$$

Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027

From Eq. (4.43) and the fact that $R_\parallel = r_\parallel^2$ and $\theta_t \to \theta_i$, $\lim_{n_{ti}\to 1} R_\parallel = 0$. Similarly from Eq. (4.42) $\lim_{n_{ti}\to 1} R_\perp = 0$.

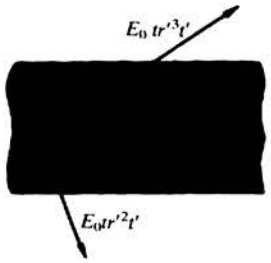
4.76 For $\theta_i > \theta_c$, Eq. (4.70) can be written

$$r_\perp = \frac{\cos\theta_i - i(\sin^2\theta_i - n_{ti}^2)^{1/2}}{\cos\theta_i + i(\sin^2\theta_i - n_{ti}^2)^{1/2}}$$

$$r_\perp r_\perp^* = \frac{\cos^2\theta_i + \sin^2\theta_i - n_{ti}^2}{\cos^2\theta_i + \sin^2\theta_i - n_{ti}^2} = 1$$

Similarly $r_\parallel r_\parallel^* = 1$.

4.77



$$t_\parallel = \frac{2\sin\theta_2\cos\theta_1}{\sin(\theta_1+\theta_2)\cos(\theta_1-\theta_2)}$$

$$t'_\parallel = \frac{2\sin\theta_1\cos\theta_2}{\sin(\theta_1+\theta_2)\cos(\theta_2-\theta_1)}$$

$$t_\parallel t'_\parallel = \frac{\sin 2\theta_1 \sin 2\theta_2}{\sin^2(\theta_1+\theta_2)\cos^2(\theta_1-\theta_2)}$$

$$= T_\parallel \text{ from Eq. (4.98)}$$

Similarly $t_\perp t'_\perp = T_\perp$

$$r_\parallel^2 = \left[\frac{\tan(\theta_1-\theta_2)}{\tan(\theta_1+\theta_2)}\right]^2 = \left[\frac{-\tan(\theta_2-\theta_1)}{\tan(\theta_1+\theta_2)}\right]^2$$

$$r'^2_\parallel = \left[\frac{\tan(\theta_2-\theta_1)}{\tan(\theta_1+\theta_2)}\right]^2 = r_\parallel^2 = R_\parallel$$

4.79 From Eq. (4.45)

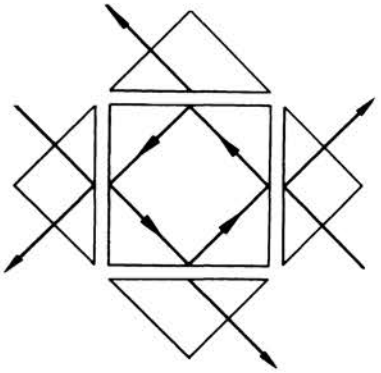
$$t'_\parallel(\theta'_p)t_\parallel(\theta_p) = \left[\frac{2\sin\theta_p\cos\theta'_p}{\sin(\theta_p+\theta'_p)\cos(\theta'_p-\theta_p)}\right] \times \left[\frac{2\sin\theta'_p\cos\theta_p}{\sin(\theta_p+\theta'_p)\cos(\theta_p-\theta'_p)}\right]$$

$$= \frac{\sin 2\theta'_p \sin 2\theta_p}{\cos^2(\theta_p-\theta'_p)} \quad \text{since} \quad \theta_p + \theta'_p = 90°$$

$$= \frac{\sin^2 2\theta_p}{\cos^2(\theta_p-\theta'_p)} \quad \text{since} \quad \sin 2\theta'_p = \sin 2\theta_p$$

$$= \frac{\sin^2 2\theta_p}{\cos^2(2\theta_p-90°)} = 1$$

4.80 Can be used as mixer to get various proportions of the two incident waves in the emitted beams. This could be done by adjusting gaps. [For some further remarks, see H. A. Daw and J. R. Izatt, *J. Opt. Soc. Am.* **55**, 201 (1965).]



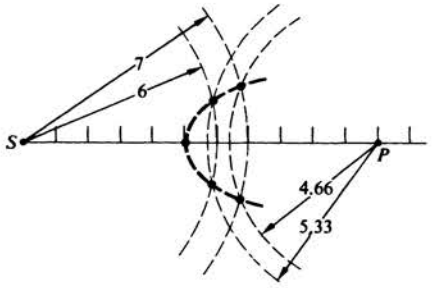
4.81 From Fig. 4.62 the obvious choice is silver. Note that in the vicinity of 300 nm, $n_I \approx n_R \approx 0.6$, in which case Eq. (4.83) yields $R \approx 0.18$. Just above 300 nm n_I increases rapidly, while n_R decreases quite strongly, with the result that $R \approx 1$ across the visible and then some.

4.82 Light traverses the base of the prism as an evanescent wave, which propagates along the adjustable coupling gap. Energy moves into the dielectric film when the evanescent wave meets certain requirements. The film acts like a waveguide, which will support characteristic vibration configurations or modes. Each mode has associated with it a given speed and polarization. The evanescent wave will couple into the film when it matches a mode configuration.

Chapter 5

5.1 All *OPL*s from S to P must be equal; therefore $\ell_o n_1 + \ell_i n_2 = s_o n_1 + s_i n_2 =$ constant; drop a perpendicular from A to the optical axis, the point where it touches is B. $\overline{BP} = s_o + s_i - x$ and the rest follows from the Pythagorean Theorem.

5.2 Using $\ell_o n_1 + \ell_i n_2 =$ constant, $\ell_o + \ell_i 3/2 =$ constant, $5 + (6)\ 3/2 = 14$. Therefore $2\ell_o + 3\ell_i = 28$ when $\ell_o = 6$, $\ell_i = 5.3$, $\ell_o = 7$, $\ell_i = 4.66$. Note that the arcs centered on S and P have to intercept for physically meaningful values of ℓ_o and ℓ_i.



Giesecke+Devrient GmbH,
Giesecke+Devrient Currency Technology GmbH,
and Papierfabrik Louisenthal GmbH. - Ex. 1027