Contents

Chapter 8

Diffraction

Version 1008.1.K, 12 Nov 2008
Please send comments, suggestions, and errata via email to kip@caltech.edu or on paper to Kip Thorne, 350-17 Caltech, Pasadena CA 91125

Box 8.1
Reader's Guide

- This chapter depends substantially on Secs. 6.1–6.4 of Chap. 6 (Geometric Optics)
- In addition, Sec. 8.6 of this chapter (diffraction at a caustic) depends on Sec. 6.5 of Chap. 6.
- Chapters 7 and 8 depend substantially on Secs. 8.1–8.5 of this chapter
- Nothing else in this book relies on this chapter.

8.1 Overview

The previous chapter was devoted to the classical mechanics of wave propagation. We showed how a classical wave equation can be solved in the short wavelength approximation to yield Hamilton's dynamical equations. We showed that, when the medium is time-independent (as we shall require in this chapter), the frequency of a wave packet is constant. And for time-independent medica, we imported a result from classical mechanics, the principle of stationary action, to show that the true geometric-optics rays coincide with paths along which the action or the integral of the phase is stationary [Eq. (6.39) and associated discussion]. Our physical interpretation of this result was that the waves do indeed travel along every path, from some source to a point of observation, where they are added together but they only give a significant net contribution when they can add coherently in phase, i.e. along the true rays. This is, essentially, Huygens' model of wave propagation, or, in modern language, a *path integral*.

Huygens' principle asserts that every point on a wave front acts as a source of secondary waves that combine so that their envelope constitutes the advancing wave front. This principle must be supplemented by two ancillary conditions, that the secondary waves are only formed in the forward direction, not backward, and that a $\pi/2$ phase shift be introduced into the secondary wave. The reason for the former condition is obvious, that for the latter, less so. We shall discuss both together with the formal justification of Huygens' construction below.

We begin our exploration of the "wave mechanics" of optics in this chapter, and we shall continue it in Chapters 8 and 9. Wave mechanics differs increasingly from geometric optics as the reduced wavelength λ increases relative to the length scales $\mathcal{R}$ of the phase fronts and $\mathcal{L}$ of the medium's inhomogeneities. The number of paths that can combine constructively increases and the rays that connect two points become blurred. In quantum mechanics, we recognize this phenomenon as the uncertainty principle and it is just as applicable to photons as to electrons.

Solving the wave equation exactly is very hard except in very simple circumstances. Geometric optics is one approximate method of solving it — a method that works well in the short wavelength limit. In this chapter and the next two, we shall develop approximate techniques that work when the wavelength becomes longer and geometric optics fails.

In this book, we shall make a somewhat artificial distinction between phenomena that arise when an effectively infinite number of paths are involved, which we call *diffraction* and which we describe in this chapter, and those when a few paths, or, more correctly, a few tight bundles of rays are combined, which we term *interference*, and whose discussion we defer to the next chapter.

In Sec. 8.2, we shall present the Fresnel-Helmholtz-Kirchhoff theory that underlies most elementary discussions of diffraction, and we shall then distinguish between Fraunhofer diffraction (the limiting case when spreading of the wavefront mandated by the uncertainty principle is very important), and Fresnel diffraction (where wavefront spreading is a modest effect and geometric optics is beginning to work, at least roughly). In Sec. 8.3, we shall illustrate Fraunhofer diffraction by computing the expected angular resolution of the Hubble Space Telescope, and in Sec. 8.4, we shall analyze Fresnel diffraction and illustrate it using lunar occultation of radio waves and zone plates.

Many contemporary optical devices can be regarded as linear systems that take an input wave signal and transform it into a linearly related output. Their operation, particularly as image processing devices, can be considerably enhanced by processing the signal in the Fourier domain, a procedure known as spatial filtering. In Sec. 8.5 we shall introduce a tool for analyzing such devices: *paraxial Fourier optics* — a close analog of the paraxial geometric optics of Sec. 6.4. We shall use paraxial Fourier optics in Sec. 8.5 to analyze the phase contrast microscope and develop the theory of Gaussian beams — the kind of light beam produced by lasers when their optically resonating cavities have spherical mirrors. Finally, in Sec. 8.6 we shall analyze diffraction near a caustic of a wave's phase field, a location where geometric optics predicts a divergent magnification of the wave (Sec. 6.5 of the preceeding chapter). As we shall see, diffraction makes the magnification finite and produces an oscillating intensity pattern (interference fringes).

8.2 Helmholtz-Kirchhoff Integral

In this section, we shall derive a formalism for describing diffraction. We shall restrict attention to the simplest (and, fortunately, the most widely useful) case: a monochromatic scalar wave

$$\boxed{\Psi = \psi(\mathbf{x})e^{-i\omega t}} \tag{8.1a}$$

with field variable ψ that satisfies the Helmholtz equation

$$\boxed{\nabla^2\psi + k^2\psi = 0 \quad \text{with } k = \omega/c\,,} \tag{8.1b}$$

except at boundaries. Generally Ψ will represent a real valued physical quantity, but for mathematical convenience we give it a complex representation and take the real part of Ψ when making contact with physical measurements. This is in contrast to a quantum mechanical wave function satisfying the Schrödinger equation which is an intrinsically complex function. We shall assume that the wave (8.1) is monochromatic (constant ω) and nondispersive, and the medium is isotropic and homogeneous (constant phase and group speed c) so k is constant. Each of these assumptions can be relaxed, but with some technical penalty.

The scalar formalism that we shall develop based on Eq. (8.1b) is fully valid for weak sound waves in a fluid, e.g. air (Chap. 15). It is also fairly accurate, but not precisely so, for the most widely used application of diffraction theory: electromagnetic waves in vacuo or in a medium with homogeneous dielectric constant. In this case ψ can be regarded as one of the Cartesian components of the electric field vector, e.g. E_x (or equally well a Cartesian component of the vector potential or the magnetic field vector). In vacuo or in a homogeneous dielectric medium, Maxwell's equations imply that this $\psi = E_x$ satisfies the scalar wave equation and thence, for fixed frequency, the Helmholtz equation (8.1b). However, when the wave hits a boundary of the medium (e.g. the edge of an aperture, or the surface of a mirror or lens), its interaction with the boundary can couple the various components of $\mathbf{E}$, thereby invalidating the simple scalar theory we shall develop. Fortunately, this polarizational coupling is usually very weak in the paraxial (small angle) limit, and also under a variety of other circumstances, thereby making our simple scalar formalism quite accurate.[1]

The Helmholtz equation (8.1b) is an elliptic, linear, partial differential equation, and we can thus express the value $\psi_{\mathcal{P}}$ of ψ at any point $\mathcal{P}$ inside some closed surface $\mathcal{E}$ as an integral over $\mathcal{E}$ of some linear combination of ψ and its normal derivative; see Fig. 8.1. To derive such an expression, we first augment the actual wave ψ in the interior of $\mathcal{E}$ with a second solution of the Helmholtz equation, namely

$$\psi_0 = \frac{e^{ikr}}{r}\,. \tag{8.2}$$

This is a spherical wave originating from the point $\mathcal{P}$, and r is the distance from $\mathcal{P}$ to the point where ψ_0 is evaluated. Next we apply Gauss's theorem, Eq. (1.71a), to the vector field

[1] For a formulation of diffraction that takes account of these polarization effects, see, e.g., Chap. 11 of Born and Wolf (1999).

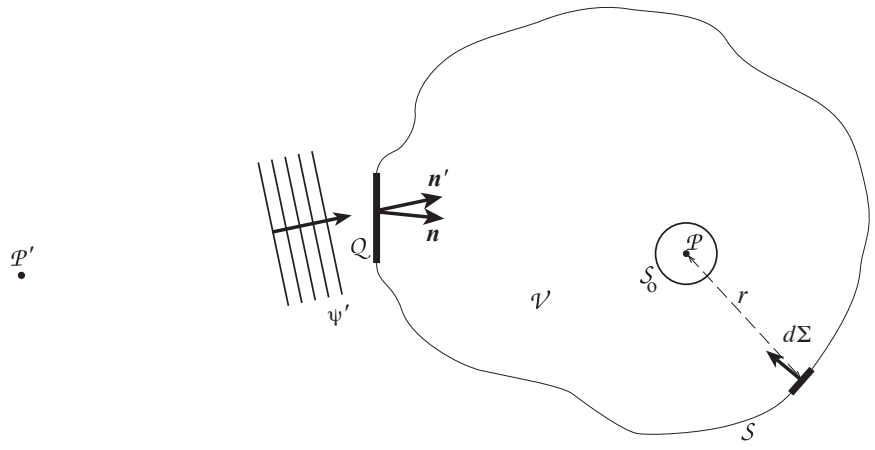

Fig. 8.1: Geometry for Helmholtz-Kirchhoff Integral, which expresses the field $\psi_{\mathcal{P}}$ at the point $\mathcal{P}$ in terms of an integral of the field and its normal derivative on the surrounding surface $\mathcal{E}$. The small sphere $\mathcal{E}_o$ surrounds the observation point $\mathcal{P}$, and $\mathcal{V}$ is the volume bounded by $\mathcal{E}$ and $\mathcal{E}_o$. The aperture $\mathcal{Q}$, the vectors $\mathbf{n}$ and $\mathbf{n}'$ at the aperture, the incoming wave ψ', and the point $\mathcal{P}'$ are irrelevant to the formulation of the Helmholtz-Kirchoff integral, but appear in subsequent applications.

$\psi\boldsymbol{\nabla}\psi_0 - \psi_0\boldsymbol{\nabla}\psi$ and invoke Eq. (8.1b), thereby arriving at Green's theorem:

$$\begin{aligned}\int_{\mathcal{E}+S_o}(\psi\boldsymbol{\nabla}\psi_0 - \psi_0\boldsymbol{\nabla}\psi)\cdot d\boldsymbol{\Sigma} &= -\int_{\mathcal{V}}(\psi\nabla^2\psi_0 - \psi_0\nabla^2\psi)dV \\ &= 0 \end{aligned} \tag{8.3}$$

Here we have introduced a small sphere $\mathcal{S}_o$ of radius r_o surrounding $\mathcal{P}$ (Fig. 8.1); $\mathcal{V}$ is the volume between the two surfaces $\mathcal{S}_o$ and $\mathcal{S}$; and for future convenience we have made an unconventional choice of direction for the integration element $d\boldsymbol{\Sigma}$: it points into $\mathcal{V}$ instead of outward thereby producing the minus sign in the second expression in Eq. (8.3). As we let the radius r_o decrease to zero, we find that, $\psi\boldsymbol{\nabla}\psi_0 - \psi_0\boldsymbol{\nabla}\psi \to -\psi(0)/r_o^2 + O(1/r_o)$ and so the integral over $\mathcal{E}_o$ becomes $4\pi\psi(\mathcal{P}) \equiv 4\pi\psi_{\mathcal{P}}$. Rearranging, we obtain

$$\boxed{\psi_{\mathcal{P}} = \frac{1}{4\pi}\int_{\mathcal{E}}\left(\psi\boldsymbol{\nabla}\frac{e^{ikr}}{r} - \frac{e^{ikr}}{r}\boldsymbol{\nabla}\psi\right)\cdot d\boldsymbol{\Sigma}\ .} \tag{8.4}$$

Equation (8.4), known as the ***Helmholtz-Kirchhoff integral****, is the promised expression for the field ψ at some point $\mathcal{P}$ in terms of a linear combination of its value and normal derivative on a surrounding surface.* The specific combination of ψ and $d\boldsymbol{\Sigma}\cdot\boldsymbol{\nabla}\psi$ that appears in this formula is perfectly immune to contributions from any wave that might originate at $\mathcal{P}$ and pass outward through $\mathcal{S}$ (any "outgoing wave"). The integral thus is influenced only by waves that enter $\mathcal{V}$ through $\mathcal{E}$, propagate through $\mathcal{V}$, and then leave through $\mathcal{E}$. [There cannot be sources inside $\mathcal{E}$, except conceivably at $\mathcal{P}$, because we assumed ψ satisfies the source-free Helmholtz equation throughout $\mathcal{V}$.] If $\mathcal{P}$ is many wavelengths away from the boundary $\mathcal{E}$, then, to high accuracy, the integral is influenced by the waves ψ only when they are entering

through $\mathcal{E}$ (when they are incoming), and not when they are leaving (outgoing). This fact is important for applications, as we shall see.

8.2.1 Diffraction by an Aperture

Next, let us suppose that *some aperture $\mathcal{Q}$ of size much larger than a wavelength but much smaller than the distance to $\mathcal{P}$ is illuminated by a distant wave source* (Fig. 8.1). (If the aperture were comparable to a wavelength in size, or if part of it were only a few wavelengths from $\mathcal{P}$, then polarizational coupling effects at the aperture would be large[1]; our assumption avoids this complication.) Let the surface $\mathcal{E}$ pass through $\mathcal{Q}$, and denote by ψ' the wave incident on $\mathcal{Q}$. We assume that the diffracting aperture has a local and linear effect on ψ'. More specifically, we suppose that the wave transmitted through the aperture is given by

$$\psi_{\mathcal{Q}} = \mathfrak{t}\,\psi' \,, \tag{8.5}$$

where $\mathfrak{t}$ is a complex transmission function that varies over the aperture. In practice, $\mathfrak{t}$ is usually zero (completely opaque region) or unity (completely transparent region). However $\mathfrak{t}$ can also represent a variable phase factor when, for example, the aperture comprises a medium (lens) of variable thickness and of different refractive index from that of the homogeneous medium outside the aperture — as is the case in microscopes, telescopes, and other optical devices.

What this formalism does not allow, though, is that $\psi_{\mathcal{Q}}$ at any point on the aperture be influenced by the wave's interaction with other parts of the aperture. For this reason, *not only the aperture, but any structure that it contains must be many wavelengths across.* To give a specific example of what might go wrong, suppose that electromagnetic radiation is normally incident upon a wire grid. A surface current will be induced in each wire by the wave's electric field, and that current will produce a secondary wave that cancels the primary wave immediately behind the wire, thereby "eclipsing" the wave. If the secondary wave from the current flowing in the next wire is comparable with the first wire's secondary wave, then the transmitted net wave field will get modified in a complex, polarization-dependent manner. Such modifications are negligble if the wires are many wavelengths apart.

Let us now use the Helmholtz-Kirchoff formula (8.4) to compute the field at $\mathcal{P}$ due to the wave $\psi_{\mathcal{Q}} = \mathfrak{t}\psi'$ transmitted through the aperture. Let the surface $\mathcal{E}$ of Fig. 8.1 comprise the aperture $\mathcal{Q}$, a sphere of radius $R \gg r$ centered on $\mathcal{P}$, and the linear extension of the aperture to meet the sphere; and assume that the only incoming waves are those which pass through the aperture. Then, as noted above, when the incoming waves subsequently pass on outward through $\mathcal{E}$, they contribute negligibly to the integral (8.4), so the only contribution is from the aperture itself.[2]

On the aperture, because $kr \gg 1$, we can write $\nabla(e^{ikr}/r) \simeq -ik\mathbf{n}e^{ikr}/r$ where $\mathbf{n}$ *is a unit vector pointing towards* $\mathcal{P}$ (Fig. 8.1). Similarly, we write $\boldsymbol{\nabla}\psi \simeq ik\mathfrak{t}\,\mathbf{n}'\psi'$, where $\mathbf{n}'$ *is a unit*

[2]Actually, the incoming waves will diffract around the edge of the aperture onto the back side of the screen that bounds the aperture, i.e. the side facing $\mathcal{P}$; and this diffracted wave will contribute to the Helmholtz-Kirchhoff integral in a polarization-dependent way; see Chap. 11 of Born and Wolf (1999). However, because the diffracted wave decays along the screen with an e-folding length of order a wavelength, its contribution will be negligible if the aperture is many wavelengths across and $\mathcal{P}$ is many wavelengths away from the edge of the aperture, as we have assumed.

vector along the direction of propagation of the incident wave (and where our assumption that anything in the aperture varies on scales long compared to $\lambda = 1/k$ permits us to ignore the gradient of $\mathbf{t}$). Inserting these gradients into the Helmholtz-Kirchoff formula, we obtain

$$\boxed{\psi_{\mathcal{P}} = -\frac{ik}{2\pi}\int_{\mathcal{Q}} d\mathbf{\Sigma}\cdot\left(\frac{\mathbf{n}+\mathbf{n}'}{2}\right)\frac{e^{ikr}}{r}\mathbf{t}\,\psi'.} \tag{8.6}$$

Equation (8.6) can be used to compute the wave from a small aperture at any point $\mathcal{P}$ in the far field. It has the form of an integral transform of the incident field variable, ψ', where the integral is over the area of the aperture. The kernel of the transform is the product of several factors. There is a factor $1/r$. This guarantees that the flux falls off as the inverse square of the distance to the aperture as we might have expected. There is also a phase factor $-ie^{ikr}$ which advances the phase of the wave by an amount equal to the optical path length between the element $d\mathbf{\Sigma}$ of the aperture and $\mathcal{P}$, minus $\pi/2$ (the phase of $-i$). The amplitude and phase of the wave ψ' can also be changed by the transmission function $\mathbf{t}$. Finally there is the geometric factor $d\hat{\mathbf{\Sigma}}\cdot(\mathbf{n}+\mathbf{n}')/2$ (with $d\hat{\mathbf{\Sigma}}$ the unit vector normal to the aperture). This is known as the *obliquity factor*, and it ensures that the waves from the aperture propagate only forward with respect to the original wave and not backward (not in the direction $\mathbf{n} = -\mathbf{n}'$). More specifically, this factor prevents the backward propagating secondary wavelets in Huygens construction from reinforcing each other to produce a back-scattered wave. When dealing with paraxial Fourier optics (Sec. 8.5), we can usually set the obliquity factor to unity.

It is instructive to specialize to a point source seen through a small diffracting aperture. If we *suppose that the source has unit strength and is located at $\mathcal{P}'$, a distance r' before $\mathcal{Q}$* (Fig. 8.1), then $\psi' = -e^{ikr'}/4\pi r'$, and $\psi_{\mathcal{P}}$ can be written in the symmetric form

$$\boxed{\psi_{\mathcal{P}} = \int\left(\frac{e^{ikr}}{4\pi r}\right) i\mathbf{t}\,(\mathbf{k}'+\mathbf{k})\cdot d\mathbf{\Sigma}\left(\frac{e^{ikr'}}{4\pi r'}\right).} \tag{8.7}$$

We can think of this expression as the Greens function response at $\mathcal{P}$ to a δ-function source at $\mathcal{P}'$. Alternatively, we can regard it as a ***propagator*** *from $\mathcal{P}'$ to $\mathcal{P}$ by way of the aperture.*

8.2.2 Spreading of the Wavefront: Fresnel and Fraunhofer Regions

Equation (8.6) [or (8.7)] gives a general prescription for computing the diffraction pattern from an illuminated aperture. It is commonly used in two complementary limits, called "Frauhnhofer" and "Fresnel".

Suppose that the aperture has linear size a (as in Fig. 8.2) and is roughly centered on the geometric ray from the source point $\mathcal{P}'$ to the field point $\mathcal{P}$. Consider the variations of the phase φ of the contributions to $\psi_{\mathcal{P}}$ that come from various places in the aperture. Using elementary trigonometry, we can estimate that locations on the aperture's opposite sides produce phases at $\mathcal{P}$ that differ by $\Delta\varphi = k(\rho_2 - \rho_1) \sim ka^2/2\rho$, where ρ_1 and ρ_2 are the distances of $\mathcal{P}$ from the two edges of the aperture and ρ is the distance from the center

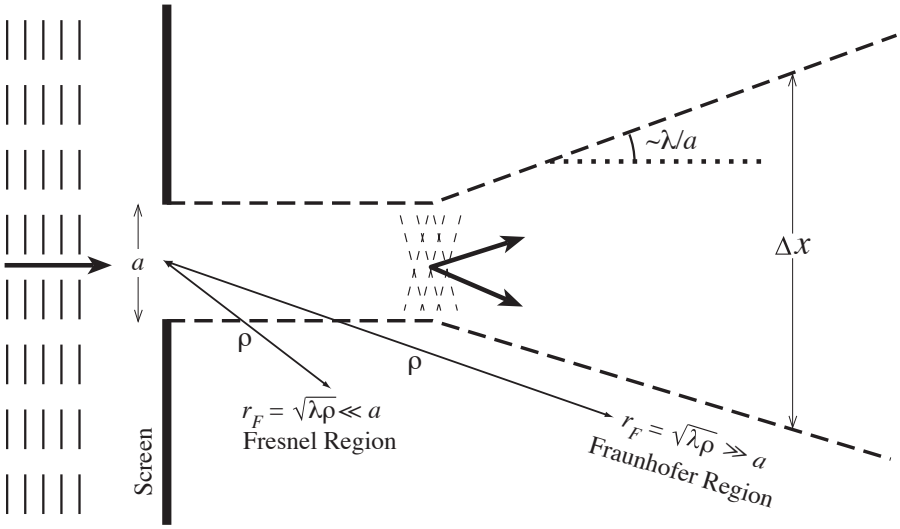

Fig. 8.2: Fraunhofer and Fresnel Diffraction.

of the aperture. There are two limiting regions for ρ depending on whether $\mathcal{P}$'s so-called *Fresnel length*

$$\boxed{r_F \equiv \left(\frac{2\pi\rho}{k}\right)^{1/2} = (\lambda\rho)^{1/2}.} \tag{8.8}$$

(a surrogate for the distance ρ) is large or small compared to the aperture. When $r_F \gg a$ (field point far from the aperture), the phase variation across the aperture, $\Delta\varphi \sim ka^2/2r$, is $\ll \pi$ and can be ignored, so the contributions at $\mathcal{P}$ from different parts of the aperture are essentially in phase with each other. This is the *Fraunhofer* region. When $r_F \ll a$ (near the aperture), the phase variation is $\Delta\varphi \gg \pi$ and therefore is of upmost importance in determining the observed intensity[3] pattern $F \propto |\psi_{\mathcal{P}}|^2$. This is the *Fresnel* region; see Fig. 8.2.

We can use an argument familiar, perhaps, from quantum mechanics to deduce the qualitative form of the intensity patterns in these two regions. For simplicity, let the incoming wave be planar (r' huge) and let it propagate perpendicular to the aperture as shown in Fig. 8.2. Then geometric optics (photons treated like classical particles) would predict that an opaque screen will cast a sharp shadow; the wave leaves the aperture plane as a beam with a sharp edge. However, wave optics insists that the transverse localization of the wave into a region of size $\Delta x \sim a$ must produce a spread in its transverse wave vector, $\Delta k_x \sim 1/a$ (a momentum uncertainty $\Delta p_x = \hbar\Delta k_x \sim \hbar/a$ in the language of the Heisenberg uncertainty principle). This uncertain transverse wave vector produces, after propagating a distance ρ, a corresponding uncertainty $(\Delta k_x/k)\rho \sim r_F^2/a$ in the beam's transverse size. This uncertainty superposes incoherently on the aperture-induced size a of the beam to produce

[3]In optics it is conventional to use the word *intensity* to mean energy flux, $F = dE/dAdt$. Astronomers often mean by "intensity" $I = dE/dAdtd\Omega$. The phrase "specific intensity" means, pretty universally, $I_\nu = dE/dAdtd\Omega d\nu$ (d Energy / d everything). In Chaps. 7, 8 and 9 we shall follow the optics convention: intensity is the same as energy flux.

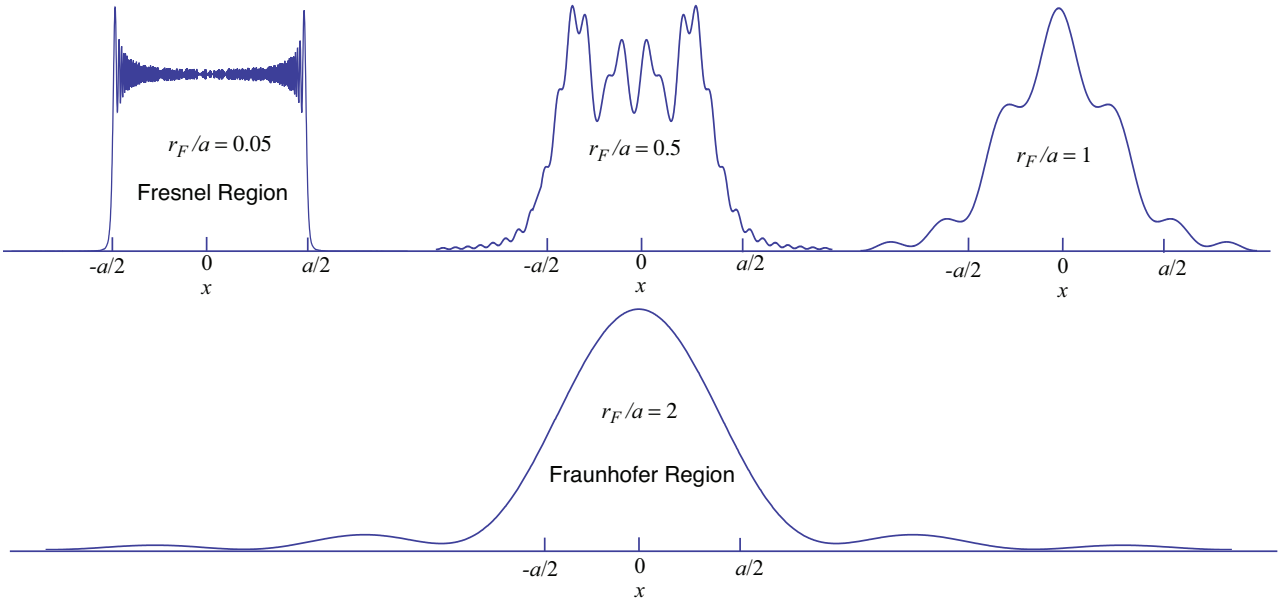

Fig. 8.3: The one-dimensional intensity diffraction pattern $|\psi|^2$ produced by a slit, $\mathfrak{t}(x) = 1$ for $|x| < a/2$ and $\mathfrak{t}(x) = 0$ for $|x| > a/2$. Four patterns are shown, each for a different value of $r_F/a = \sqrt{\lambda z}/a$. For $r_F/a = 0.05$ (very near the slit; extreme Fresnel region), the intensity distribution resembles the slit itself: sharp edges at $x = \pm a/2$, but with damped oscillations (interference fringes) near the edges. For $r_F/a = 2$ (beginning of Fraunhofer region) there is a bright central peak and low-brightess, oscillatory side bands. As r_F/a increeases 0.05 to 2, the pattern transitions (quite rapidly between $\alpha = 2$ and 0.5) from the Fraunhofer pattern to the Fresnel pattern. These intensity distributions are derived in Ex. 8.8.

a net transverse beam size

$$\begin{aligned} \Delta x &\sim \sqrt{a^2 + (r_F^2/a)^2} \\ &\sim a \qquad \text{if} \quad r_F \ll a\ , \quad \text{i.e., } \rho \ll a^2/\lambda \text{ (Fresnel region)} \\ &\sim \left(\frac{\lambda}{a}\right)\rho \quad \text{if} \quad r_F \gg a\ , \quad \text{i.e., } \rho \gg a^2/\lambda \text{ (Fraunhofer region).} \end{aligned} \tag{8.9}$$

In the nearby, Fresnel region, the aperture creates a beam whose edges will have the same shape and size as the aperture itself, and will be reasonably sharp (but with some oscillatory blurring, associated with the wave-packet spreading, that we shall analyze below); see Fig. 8.3. Thus, in the Fresnel region the field behaves approximately as one would predict using geometric optics. By contrast, in the more distant, Fraunhofer region, wave-front spreading will cause the transverse size of the entire beam to grow linearly with distance; and, as illustrated in Fig. 8.3, the intensity pattern will differ markedly from the aperture's shape. We shall analyze the distant, Fraunhofer region in Sec. 8.3, and the near, Fresnel region in Sec. 8.4.

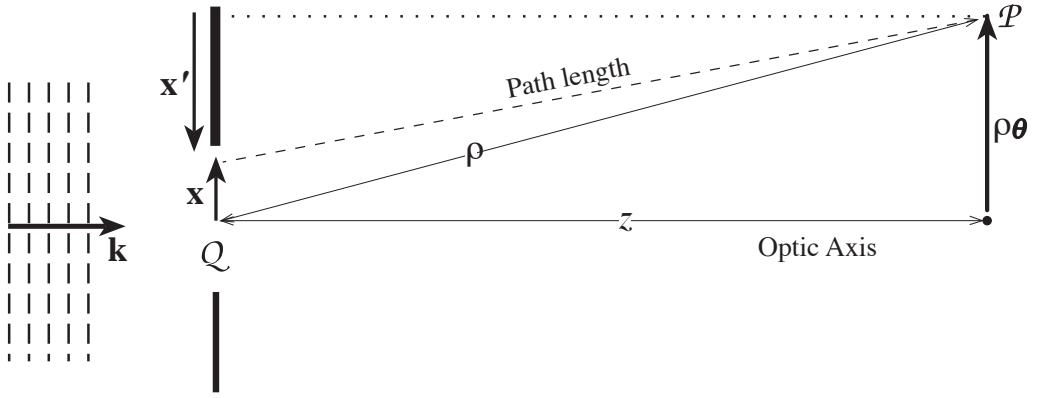

Fig. 8.4: Geometry for computing the path length between a point $\mathcal{Q}$ in the aperture and the point of observation $\mathcal{P}$. The transverse vector $\mathbf{x}$ is used to identify $\mathcal{Q}$ in our Fraunhofer analysis (Sec. 8.3), and $\mathbf{x}'$ is used in our Fresnel analysis (Sec. 8.4).

8.3 Fraunhofer Diffraction

Consider the Fraunhofer region of strong wavefront spreading, $r_F \gg a$, and for simplicity specialize to the case of an incident plane wave with wave vector $\mathbf{k}$ orthogonal to the aperture plane; see Fig. 8.4. Regard the line along $\mathbf{k}$ through the center of the aperture $\mathcal{Q}$ as the "optic axis"; identify points in the aperture by their transverse two-dimensional vectorial separation $\mathbf{x}$ from that axis; identify $\mathcal{P}$ by its distance ρ from the aperture center and its 2-dimensional transverse separation $\rho\boldsymbol{\theta}$ from the optic axis; and restrict attention to small-angle diffraction $|\boldsymbol{\theta}| \ll 1$. Then the geometric path length between $\mathcal{P}$ and a point $\mathbf{x}$ on $\mathcal{Q}$ [the length denoted r in Eq. (8.6)] can be expanded as

$$\text{Path length } = r = (\rho^2 - 2\rho\mathbf{x}\cdot\boldsymbol{\theta} + x^2)^{1/2} \simeq \rho - \mathbf{x}\cdot\boldsymbol{\theta} + \frac{x^2}{2\rho} + \ldots \tag{8.10}$$

cf. Fig. 8.4. The first term in this expression, ρ, just contributes an $\mathbf{x}$-independent phase $e^{ik\rho}$ to the $\psi_{\mathcal{P}}$ of Eq. (8.6). The third term, $x^2/2\rho$, contributes a phase variation that is $\ll 1$ here in the Fraunhofer region (but that will be important in the Fresnel region, Sec. 8.4 below). Therefore, in the Fraunhofer region we can retain just the second term, $-\mathbf{x}\cdot\boldsymbol{\theta}$ and write Eq. (8.6) in the form

$$\boxed{\psi_{\mathcal{P}}(\boldsymbol{\theta}) \propto \int e^{-ik\mathbf{x}\cdot\boldsymbol{\theta}} \mathfrak{t}\,(\mathbf{x}) d\Sigma \equiv \tilde{\mathfrak{t}}(\boldsymbol{\theta})\ ,} \tag{8.11a}$$

where $d\Sigma$ is the surface area element in the aperture plane and we have dropped a constant phase factor and constant multiplicative factors. Thus, $\psi_{\mathcal{P}}(\boldsymbol{\theta})$ in the Fraunhofer region is given by the two-dimensional Fourier transform, denoted $\tilde{\mathfrak{t}}(\boldsymbol{\theta})$, of the transmission function $\mathfrak{t}(\mathbf{x})$, with $\mathbf{x}$ made dimensionless in the transform by multiplying by $k = 2\pi/\lambda$.

The intensity distribution $F = dE/dAdt$ of the diffracted waves is

$$\boxed{F(\boldsymbol{\theta}) = \overline{(\Re[\psi_{\mathcal{P}}(\boldsymbol{\theta})e^{-i\omega t}])^2} = \tfrac{1}{2}|\psi_{\mathcal{P}}(\boldsymbol{\theta})|^2 \propto |\tilde{\mathfrak{t}}(\boldsymbol{\theta})|^2\ ,} \tag{8.11b}$$

where $\Re$ means take the real part, and the bar means average over time.

As an example, the bottom curve in Fig. 8.3 above shows the intensity distribution from a slit

$$\mathfrak{t}(x) = H_1(x) \equiv \begin{cases} 1 & |x| < a/2 \\ 0 & |x| > a/2 \ , \end{cases} \tag{8.12a}$$

for which

$$\psi_{\mathcal{P}}(\theta) \propto \tilde{H}_1 \propto \int_{-a/2}^{a/2} e^{ikx\theta} dx \propto \mathrm{sinc}\left(\tfrac{1}{2}ka\theta\right) \ , \tag{8.12b}$$

$$F(\theta) \propto \mathrm{sinc}^2\left(\tfrac{1}{2}ka\theta\right) \ . \tag{8.12c}$$

Here $\mathrm{sinc}(\xi) \equiv \sin(\xi)/\xi$. The bottom intensity curve is almost but not quite described by Eq. (8.12c); the differences (e.g., the not-quite-zero value of the minimum between the central peak and the first side lobe) are due to the field point not being fully in the Fraunhofer region, $r_F/a = 2$ rather than $r_F/a \gg 1$.

It is usually uninteresting to normalise Fraunhofer diffraction patterns. On those rare occasions when the absolute value of the observed flux is needed, rather than just the angular shape of the diffraction pattern, it typically can be derived most easily from conservation of the total wave energy. This is why we ignore the proportionality factors in in the above diffraction patterns.

All of the techniques for handling Fourier transforms (which should be familiar from quantum mechanics and elsewhere) can be applied to derive Fraunhofer diffraction patterns. In particular, the *convolution theorem* turns out to be very useful. It says that *the Fourier transform of the convolution*

$$f_2 \otimes f_1 \equiv \int_{-\infty}^{+\infty} f_2(\mathbf{x} - \mathbf{x}') f_1(\mathbf{x}') d\Sigma' \tag{8.13}$$

of two functions f_1 and f_2 is equal to the product $\tilde{f}_2(\boldsymbol{\theta})\tilde{f}_1(\boldsymbol{\theta})$ of their Fourier transforms, and conversely. [Here and throughout this chapter we use the optics version of a Fourier transform in which two-dimensional transverse position $\mathbf{x}$ is made dimensionless via the wave number k; Eq. (8.11a) above.]

As an example of the convolution theorem's power, we shall compute the diffraction pattern produced by a diffraction grating:

8.3.1 Diffraction Grating

A diffraction grating can be modeled as a finite series of alternating transparent and opaque, long, parallel stripes. Let there be N transparent and opaque stripes each of width $a \gg \lambda$ (Fig. 8.5 a), and idealize them as infinitely long so their diffraction pattern is one-dimensional. We shall outline how to use the convolution theorem to derive their Fraunhofer diffraction pattern. The details are left as an exercise for the reader (Ex. 8.4).

Our idealized N-slit grating can be regarded as an infinite series of δ-functions with separation $2a$ convolved with the transmission function H_1 [Eq. (8.12a)] for a single slit of

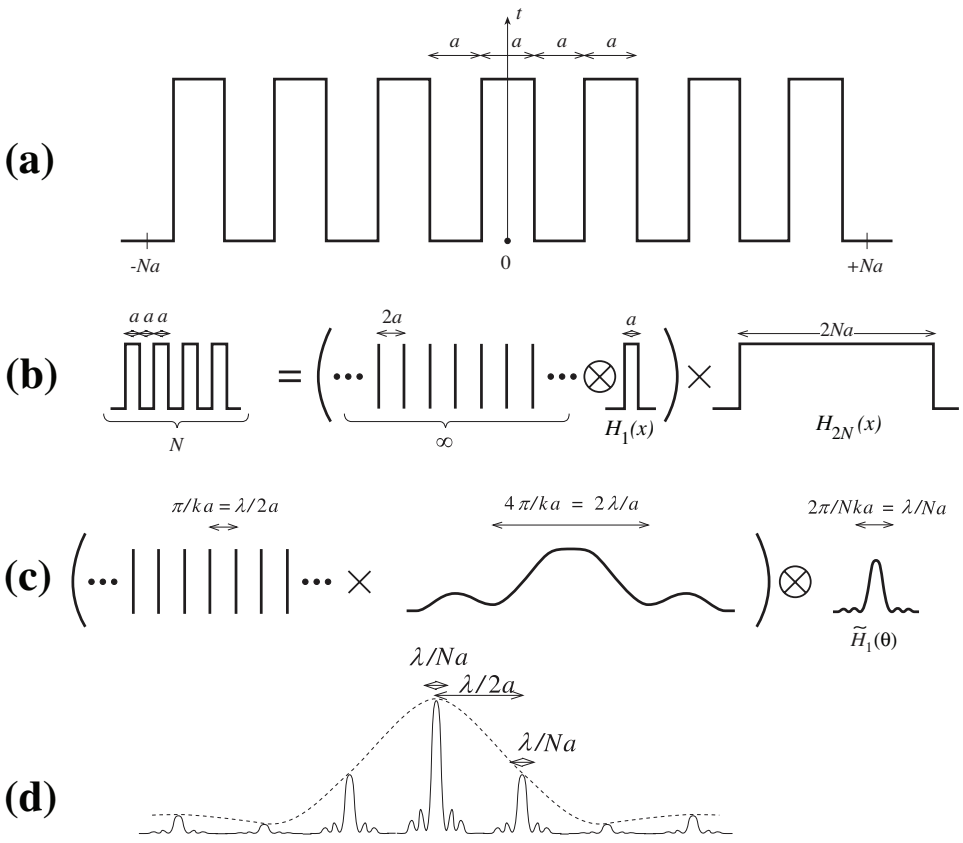

Fig. 8.5: (a) Diffraction grating $\mathfrak{t}(x)$ formed by N alternating transparent and opaque stripes each of width a. (b) Decomposition of this finite grating into an infinite series of equally spaced δ-functions that are convolved (the symbol $\otimes$) with the shape of an individual transparent stripe and then multiplied (the symbol $\times$) by a large aperture function covering N such stripes; cf. Eq. (8.16) (c) The resulting Fraunhofer diffraction pattern $\tilde{\mathfrak{t}}(\theta)$ shown schematically as the Fourier transform of a series of delta functions multiplied by the Fourier transform of the large aperture and then convolved with the transform of a single stripe. (d) The intensity $F \propto |\tilde{\mathfrak{t}}(\theta)|^2$ of this diffraction pattern.

width a,

$$\int_{-\infty}^{\infty} \left[\sum_{n=-\infty}^{+\infty} \delta(\xi - 2an) \right] H_1(x-\xi) d\xi \tag{8.14}$$

and then multiplied by an aperture function with width $2Na$

$$H_{2N}(x) \equiv \begin{cases} 1 & |x| < Na \\ 0 & |x| > Na \ . \end{cases} \tag{8.15}$$

More explicitly,

$$\mathfrak{t}(x) = \left(\int_{-\infty}^{\infty} \left[\sum_{n=-\infty}^{+\infty} \delta(y - 2an) \right] H_1(x-\xi) d\xi \right) H_{2N}(x) \ , \tag{8.16}$$

which is shown graphically in Fig. 8.5(b).

Let us apply convolution theorem to expression (8.16) for our transmission grating. The diffraction pattern of the infinite series of δ-functions with spacing $2a$ is itself an infinite series of δ-functions with reciprocal spacing $2\pi/(2ka) = \lambda/2a$ (see the hint in Ex. 8.4). This must be multiplied by the Fourier transform $\tilde{H}_1(\theta) \propto \mathrm{sinc}(\frac{1}{2}ka\theta)$ of the single narrow slit, and then convolved with the Fourier transform $\tilde{H}_{2N}(\theta) \propto \mathrm{sinc}(Nka\theta)$ of the wide slide. The result is shown schematically in Fig. 8.5(c). (Each of the transforms is real, so the one-dimensional functions shown in the figure fully embody them.)

The resulting diffracted intensity, $F \propto |\mathfrak{t}(\theta)|^2$ (as computed in Ex. 8.4), is shown in Fig. 8.5(d). The grating has channeled the incident radiation into a few equally spaced beams with directions $\theta = \pi p/ka$, where p is an integer known as the *order* of the beam. Each of these beams has a shape given by $|\tilde{H}_{2N}(\theta)|^2$: a sharp central peak with half width (distance from center of peak to first null of the intensity) $\lambda/2Na$, followed by a set of *side lobes* whose intensities are $\propto N^{-1}$.

The fact that the deflection angles $\theta = \pi p/ka = p\lambda/2a$ of these beams are proportional to λ underlies the use of diffraction gratings for spectroscopy. It is of interest to ask what the wavelength resolution of such an idealized grating might be. If one focuses attention on the p'th order beams at two wavelengths λ and $\lambda + \delta\lambda$ (which are located at $\theta = p\lambda/2a$ and $p(\lambda + \delta\lambda)/2a$, then one can distinguish the beams from each other when their separation $\delta\theta = p\delta\lambda/2a$ is at least as large as the angular distance $\lambda/2Na$ between the maximum of each beam's diffraction pattern and its first minimum, i.e., when

$$\frac{\lambda}{\delta\lambda} \lesssim \mathcal{R} \equiv Np \; . \tag{8.17}$$

$\mathcal{R}$ is called the grating's *chromatic resolving power*.

Real gratings are not this simple. First, they usually work not by modulating the amplitude of the incident radiation in this simple manner, but instead by modulating the phase. Second, the manner in which the phase is modulated is such as to channel most of the incident power into a particular order, a technique known as *blazing*. Third, gratings are often used in reflection rather than transmission. Despite these complications, the principles of a real grating's operation are essentially the same as our idealized grating. Manufactured gratings typically have $N \gtrsim 10,000$, giving a wavelength resolution for visual light that can be as small as $\lambda/10^5 \sim 10\,\mathrm{pm}$, i.e. 10^{-11}m.

8.3.2 Airy Pattern of a Circular Aperture: Hubble Space Telescope

The Hubble Space Telescope was launched in April 1990 to observe planets, stars and galaxies above the earth's atmosphere. One reason for going into space is to avoid the irregular refractive index variations in the earth's atmosphere, known, generically, as *seeing*, which degrade the quality of the images. (Another reason is to observe the ultraviolet part of the spectrum, which is absorbed in the earth's atmosphere.) Seeing typically limits the angular resolution of Earth-bound telescopes at visual wavelengths to $\sim 0.5''$. We wish to

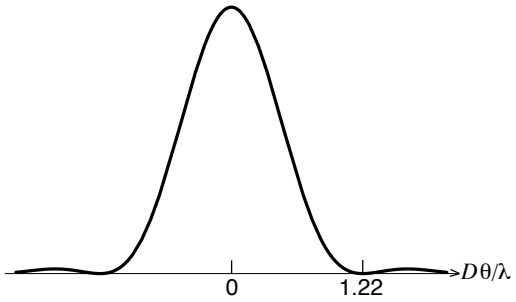

Fig. 8.6: Airy diffraction pattern produced by a circular aperture.

compute how much the angular resolution improves by going into space. As we shall see, the computation is essentially an exercise in Fraunhofer diffraction theory.

The essence of the computation is to idealise the telescope as a circular aperture with diameter equal to the diameter of the primary mirror. Light from this mirror is actually reflected onto a secondary mirror and then follows a complex optical path before being focused onto a variety of detectors. However, this path is irrelevant to the angular resolution. The purpose of the optics is merely to bring the Fraunhofer-region light to a focus close to the mirror, in order to produce an instrument that is compact enough to be launched and to match the sizes of stars' images to the pixel size on the detector. In doing so, however, the optics leaves the angular resolution unchanged; *the resolution is the same as if we were to observe the light, which passes through the primary mirror's circular aperture, far beyond the mirror, in the Fraunhofer region.*

If the telescope aperture were very small, for example a pin hole, then the light from a point source (a very distant star) would create a broad diffraction pattern, and the telescope's angular resolution would be correspondingly poor. As we increase the diameter of the aperture, we still see a diffraction pattern, but its angular width diminishes.

Using these considerations, we can compute how well the telescope can distinguish neighboring stars. We do not expect it to resolve them fully if they are closer together on the sky than the angular width of the diffraction patttern. Of course, optical imperfections and pointing errors in a real telescope may degrade the image quality even further, but this is the best that we can do, limited only by the uncertainty principle.

The calculation of the Fraunhofer amplitude far from the aperture is straightforward (Fig. 8.5):

$$\begin{aligned} \psi(\theta) &\propto \int_{\text{Disk with diameter } D} e^{-ik\mathbf{x}\cdot\boldsymbol{\theta}} d\Sigma \\ &\propto \text{jinc}\left(\frac{kD\theta}{2}\right) \end{aligned} \tag{8.18}$$

where D is the diameter of the aperture (i.e., of the telescope's primary mirror), $\theta \equiv |\boldsymbol{\theta}|$ is angle from the optic axis, and $\text{jinc}(x) \equiv J_1(x)/x$ with J_1 the Bessel function of order one. The flux from the star observed at angle θ is therefore $\propto \text{jinc}^2(kD\theta/2)$. This intensity pattern, known as the *Airy pattern*, is shown in Fig. 8.6. There is a central "Airy disk" surrounded by a circle where the flux vanishes, and then further surrounded by a series of

concentric rings whose flux diminishes with radius. Only 16 percent of the total light falls outside the central Airy disk. The angular radius θ_A of the Airy disk, i.e. the radius of the dark circle surrounding it, is determined by the first zero of $J_1(kD\theta/2)$:

$$\boxed{\theta_A = 1.22\lambda/D\ .} \tag{8.19}$$

A conventional, though essentially arbitrary, criterion for angular resolution is to say that two point sources can be distinguished if they are separated in angle by more than θ_A. For the Hubble Space Telescope, $D = 2.4$m and $\theta_A \sim 0.04''$ at visual wavelengths, which is over ten times better than is achievable on the ground with conventional (non-adaptive) optics.

Initially, there was a serious problem with Hubble's telescope optics. The hyperboloidal primary mirror was ground to the wrong shape, so rays parallel to the optic axis did not pass through a common focus after reflection off a convex hyperboloidal secondary mirror. This defect, known as *spherical aberration*, created blurred images. However, it was possible to correct this error in subsequent instruments in the optical train, and the Hubble Space Telescope became the most successful telescope of all time, transforming our view of the Universe.

8.3.3 Babinet's Principle

Suppose that monochromatic light falls normally onto a large circular aperture with diameter D. At distances $z \lesssim D^2/\lambda$ (i.e., $r_F \lesssim D$), the transmitted light will be collimated into a beam with diameter D, and at larger distances, the beam will become conical with opening angle λ/D and flux distribution given by the Airy diffraction pattern of Fig. 8.6.

Now, place into this aperture a significantly smaller object (size $a \ll D$; Fig. 8.7) with transmissivity $\mathfrak{t}_1(\mathbf{x})$ — for example an opaque star-shaped object. This object will produce a Fraunhofer diffraction pattern with opening angle $\lambda/a \gg \lambda/D$ that extends well beyond the large aperture's beam. Outside that beam, the diffraction pattern will be insensitive to the shape and size of the large aperture because only the small object can diffract light to these large angles; so the diffracted flux will be $F_1(\boldsymbol{\theta}) \propto |\tilde{\mathfrak{t}}_1(\theta)|^2$.

Suppose, next, that we replace the small object by one with a *complentary transmissivity* $\mathfrak{t}_2$, complementary in the sense that

$$\boxed{\mathfrak{t}_1(\mathbf{x}) + \mathfrak{t}_2(\mathbf{x}) = 1\ .} \tag{8.20a}$$

For example, we replace a small, opaque star-shaped object by an opaque screen that fills the original, large aperture except for a star-shaped hole. This new, complementary object will produce a diffraction pattern $F_2(\boldsymbol{\theta}) \propto |\tilde{\mathfrak{t}}_2(\theta)|^2$. Outside the large aperture's beam, this pattern again is insensitive to the size and shape of the large aperture, i.e., insensitive to the 1 in $\mathfrak{t}_2 = 1 - \mathfrak{t}_1$; so at these large angles, $\tilde{\mathfrak{t}}_2(\boldsymbol{\theta}) = -\tilde{\mathfrak{t}}_1(\boldsymbol{\theta})$, which implies that *the intensity diffraction pattern of the original object and the new, complementary object will be the same, outside the large aperture's beam*:

$$\boxed{F_2(\boldsymbol{\theta}) \propto |\tilde{\mathfrak{t}}_2(\boldsymbol{\theta})|^2 = |\tilde{\mathfrak{t}}_1(\boldsymbol{\theta})|^2 \propto F_1(\boldsymbol{\theta})\ .} \tag{8.20b}$$

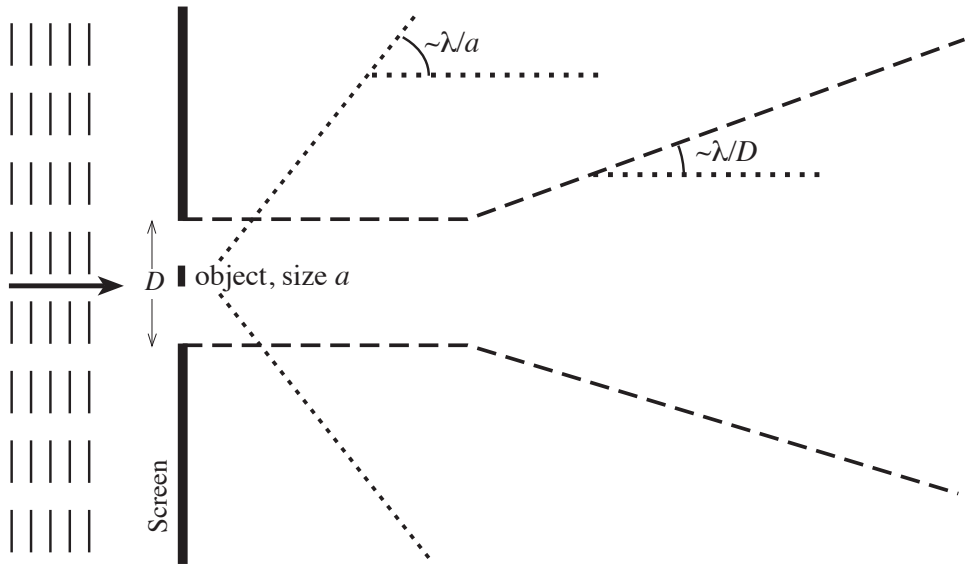

Fig. 8.7: Geometry for Babinet's Principle. The beam produced by the large aperture D is confined between the long-dashed lines. Outside this beam, the intensity pattern $F(\boldsymbol{\theta}) \propto |\mathfrak{t}(\boldsymbol{\theta})|^2$ produced by a small object (size a) and its complement are the same, Eqs. (8.20).

This is called *Babinet's Principle*

****

EXERCISES

Exercise 8.1 *Practice: Convolutions and Fourier Transforms*

(a) Calculate the one-dimensional Fourier transforms of the functions $f_1(x) \equiv e^{-x^2/2\sigma^2}$, and $f_2 \equiv 0$ for $x < 0$, $f_2 \equiv e^{-x/h}$ for $x \geq 0$.

(b) Take the inverse transforms of your answers to part (a) and recover the original functions.

(c) Convolve the exponential function f_2 with the Gaussian function f_1 and then compute the Fourier transform of their convolution. Verify that the result is the same as the product of the Fourier transforms of f_1 and f_2.

Exercise 8.2 *Problem: Pointilist Painting*

The neo-impressionist painter George Seurat was a member of the pointillist school. His paintings consisted of an enormous number of closely spaced dots of pure pigment (of size ranging from ~ 0.4mm in his smaller paintings to ~ 4mm in his largest paintings such as *A Sunday afternoon on the island of La Grande Jatte*, Fig. 8.8). The illusion of color mixing was produced only in the eye of the observer. How far from the painting should one stand in order to obtain the desired blending of color?

Fig. 8.8: Left: George Seurat's painting *A Sunday afternoon on the Island of La Grande Jatte*. When viewed from sufficient distance, adjacent dots of paint with different colors blend together in the eye to form another color. Right: Enlargement of the woman at the center of the painting. In this enlargement one sees clearly the individual dots of paint.

Exercise 8.3 *Problem: Thickness of a Human Hair*

Conceive and carry out an experiment using light diffraction to measure the thickness of a hair from your head, accurate to within a factor ~ 2. [Hint: make sure the source of light that you use is small enough that its finite size has negligible influence on your result.]

Exercise 8.4 *Derivation: Diffraction Grating*

Use the convolution theorem to carry out the calculation of the Fraunhofer diffraction pattern from the grating shown in Fig. 8.5. [*Hint:* To show that the Fourier transform of the infinite sequence of equally spaced delta functions is a similar sequence of delta functions, perform the Fourier transform to get $\sum_{n=\infty}^{+\infty} e^{i2kan\theta}$ (aside from a multiplicative factor); then use the formulas for a Fourier *series* expansion, and its inverse, for any function that is periodic with period π/ka to show that $\sum_{n=-\infty}^{+\infty} e^{i2kan\theta}$ is a sequence of delta functions.]

Exercise 8.5 *Derivation: Airy Pattern*

Derive and plot the Airy diffraction pattern (8.18) and show that 84 percent of the light is contained within the Airy disk.

Exercise 8.6 *Problem: Triangular Diffraction Grating*

Sketch the Fraunhofer diffraction pattern you would expect to see from a diffraction grating made from three groups of parallel lines aligned at angles of 120° to each other (Fig. 8.9).

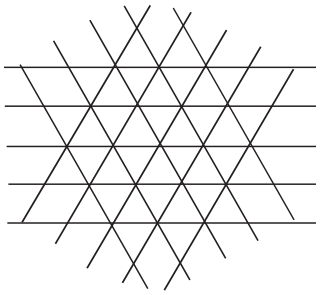

Fig. 8.9: Diffraction grating formed from three groups of parallel lines.

Exercise 8.7 *Problem: Light Scattering by Particles*

Consider the scattering of light by an opaque particle of size $a \gg 1/k$. One component of the scattered radiation is due to diffraction around the particle. This component is confined to a cone with opening angle $\Delta\theta \sim \pi/ka \ll 1$ about the incident wave direction. It contains power $P_S = FA$, where F is the incident intensity and A is the cross sectional area of the particle perpendicular to the incident wave.

(a) Give a semi-quantitative derivation of $\Delta\theta$ and P_S using Babinet's principle.

(b) Explain why the total "extinction" (absorption plus scattering) cross section is equal to $2A$ independent of the shape of the opaque particle.

****

8.4 Fresnel Diffraction

We next turn to the Fresnel region of observation points $\mathcal{P}$ with $r_F = \sqrt{\lambda\rho}$ much smaller than the aperture. In this region, the field at $\mathcal{P}$ arriving from different parts of the aperture has significantly different phase $\Delta\varphi \gg 1$. We again specialize to incoming wave vectors that are approximately orthogonal to the aperture plane and to small diffraction angles so that we can ignore the obliquity factor. By contrast with the Fraunhofer case, however, we identify $\mathcal{P}$ by its distance z from the aperture plane instead of its distance ρ from the aperture center, and we use as our integration variable in the aperture $\mathbf{x}' \equiv \mathbf{x} - \rho\boldsymbol{\theta}$ (cf. Fig. 8.4.), thereby writing the dependence of the phase at $\mathcal{P}$ on $\mathbf{x}$ in the form

$$\Delta\varphi \equiv k \times [(\text{path length from } \mathbf{x} \text{ to } \mathcal{P}) - z] = \frac{k\mathbf{x}'^2}{2z} + O\left(\frac{kx'^4}{z^3}\right). \tag{8.21}$$

In the Fraunhofer region (Sec. 8.3 above), only the linear term $-k\mathbf{x}\cdot\boldsymbol{\theta}$ in $k\mathbf{x}'^2/2z \simeq k(\mathbf{x} - r\boldsymbol{\theta})^2/r$ was significant. In the Fresnel region the term quadratic in $\mathbf{x}$ is also significant (and we have changed variables to $\mathbf{x}'$ so as to simplify it), but the $O(x'^4)$ term is negligible.

Let us consider the Fresnel diffraction pattern formed by a simple aperture of arbitrary shape, illuminated by a normally incident plane wave. It is convenient to introduce transverse Cartesian coordinates (x', y') and to define

$$\boxed{\sigma = \left(\frac{k}{\pi z}\right)^{1/2} x' \ , \quad \tau = \left(\frac{k}{\pi z}\right)^{1/2} y' \ .} \tag{8.22a}$$

[Notice that $(k/\pi z)^{1/2}$ is $\sqrt{2}/r_F$; cf. Eq. (8.8).] We can thereby rewrite Eq. (8.6) (setting the obliquity factor to one) in the form

$$\boxed{\psi_{\mathcal{P}} = -\frac{ike^{ikz}}{2\pi z}\int_{\mathcal{Q}} e^{i\Delta\varphi}\psi_{\mathcal{Q}}dx'dy' = -\frac{i}{2}\int\int_{\mathcal{Q}} e^{i\pi\sigma^2/2}e^{i\pi\tau^2/2}\psi_{\mathcal{Q}}e^{ikz}d\sigma d\tau \ .} \tag{8.22b}$$

We shall use this rather general expression in Sec. 8.5, when discussing Paraxial Fourier optics.

In this section we shall focus on the details of the Fresnel diffraction pattern for an incoming plane wave that falls perpendicularly on the aperture, so $\psi_{\mathcal{Q}}$ is constant over the aperture.

8.4.1 Rectangular Aperture, Fourier Integrals and Cornu Spiral

For simplicity, we initially confine attention to a rectangular aperture with edges along the x' and y' directions. Then the two integrals have limits that are independent of each other and the integrals can be expressed in the form $\mathcal{E}(\sigma_{max}) - \mathcal{E}(\sigma_{min})$ and $\mathcal{E}(\tau_{max}) - \mathcal{E}(\tau_{min})$, so

$$\boxed{\psi_{\mathcal{P}} = \frac{-i}{2}[\mathcal{E}(\sigma_{max}) - \mathcal{E}(\sigma_{min})][\mathcal{E}(\tau_{max}) - \mathcal{E}(\tau_{min})]\psi_{\mathcal{Q}}e^{ikz} \equiv \frac{-i}{2}\Delta\mathcal{E}_\sigma\Delta\mathcal{E}_\tau\psi_{\mathcal{Q}}e^{ikz} \ ,} \tag{8.23a}$$

where the arguments are the limits of integration and where

$$\boxed{\mathcal{E}(\xi) \equiv \int_0^\xi e^{i\pi\sigma^2/2}d\sigma \equiv C(\xi) + iS(\xi) \ .} \tag{8.23b}$$

Here

$$\boxed{C(\xi) \equiv \int_0^\xi d\sigma \cos(\pi\sigma^2/2) \ , \quad S(\xi) \equiv \int_0^\xi d\sigma \sin(\pi\sigma^2/2) \ .} \tag{8.23c}$$

are known as *Fresnel Integrals*, and are standard functions tabulated in many books and known to Mathematica and Maple. Notice that the intensity distribution is

$$\boxed{F \propto |\psi_{\mathcal{P}}|^2 \propto |\Delta\mathcal{E}_\sigma|^2|\Delta\mathcal{E}_\tau|^2 \ .} \tag{8.23d}$$

It is convenient to exhibit the Fresnel integrals graphically using a *Cornu spiral*, Fig. 8.10. This is a graph of the parametric equation $[C(\xi), S(\xi)]$, or equivalently a graph of $\mathcal{E}(\xi) =$

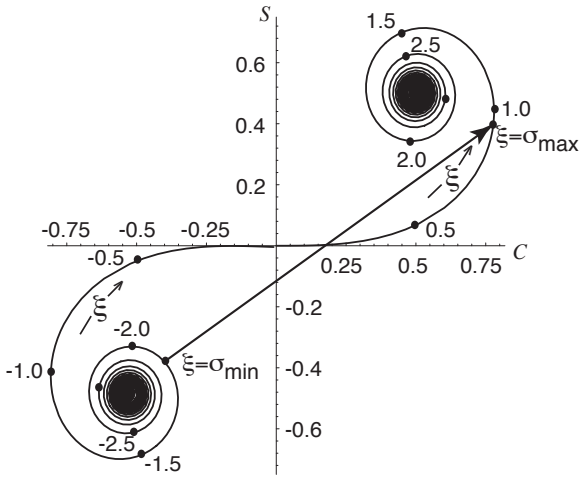

Fig. 8.10: Cornu Spiral in the complex plane; the real part of $\mathcal{E}(\xi) = C(\xi) + iS(\xi)$ is plotted horizontally and the imaginary part vertically; the point $\xi = 0$ is at the origin, positive ξ in the upper right quadrant, and negative ξ in the lower left quadrant. The diffracted intensity is proportional to the square of the arrow reaching from $\xi = \sigma_{\rm min}$ to $\xi = \sigma_{\rm max}$.

$C(\xi) + iS(\xi)$ in the complex plane. The two terms $\Delta\mathcal{E}_\sigma$ and $\Delta\mathcal{E}_\tau$ in Eq. (8.23b) can be represented in amplitude and phase by arrows in the (C, S) plane reaching from $\xi = \sigma_{min}$ on the Cornu spiral to $\xi = \sigma_{max}$ (Fig. 8.10), and from $\xi = \tau_{min}$ to $\xi = \tau_{\rm max}$. Correspondingly, *the intensity, F [Eq. (8.23d)], is proportional to the product of the squared lengths of these two vectors.*

8.4.2 Unobscured Plane Wave

The simplest illustration is the totally unobscured plane wave. In this case, the limits of both integrations extend from $-\infty$ to $+\infty$, which as we see in Fig. 8.10 is an arrow of length $2^{1/2}$ and phase $\pi/4$. Therefore, $\psi_{\mathcal{P}}$ is equal to $(2^{1/2}e^{i\pi/4})^2(-i/2)\psi_Q e^{ikz} = \psi_Q e^{ikz}$, as we could have deduced simply by solving the Helmholtz equation (8.1b) for a plane wave.

This unobscured-wavefront calculation elucidates three issues that we have already met. First, it illustrates our interpretation of Fermat's principle in geometric optics. In the limit of short wavelength, the paths that contribute to the wave field are just those along which the phase is stationary to small variations in path. Our present calculation shows that, *because of the tightening of the Cornu spiral as one moves toward a large argument, the paths that contribute significantly to $\psi_{\mathcal{P}}$ are those that are separated from the geometric-optics path by less than a few Fresnel lengths at $\mathcal{Q}$.* (For a laboratory experiment with light and $z \sim 2$m, a Fresnel length is typically $\sqrt{\lambda z} \sim 1$mm.)

A second, and related, point is that *in computing the Fresnel diffraction pattern from a more complicated aperture, we need only perform the integral (8.6) in the immediate vicinity of the geometric-optics ray.* We can ignore the contribution from the extension of the aperture $\mathcal{Q}$ to meet the "sphere at infinity" (the surface $\mathcal{E}$ in Fig. 8.1) even when the wave is unobstructed there. The rapid phase variation makes the contribution from $\mathcal{E}$ sum to zero.

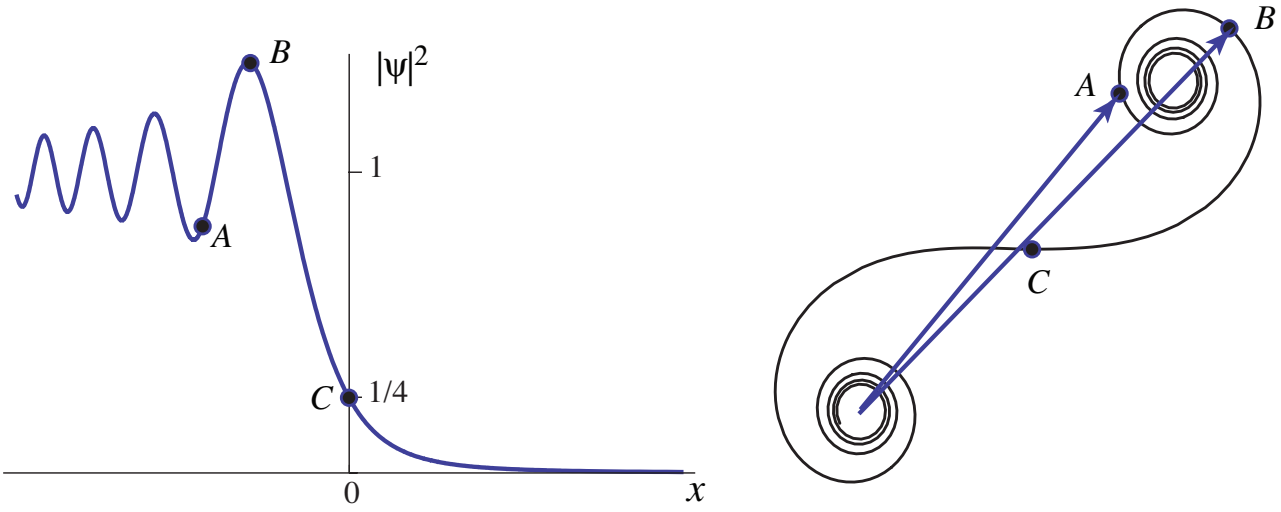

Fig. 8.11: Intensity diffraction pattern formed by a straight edge, and graphical interpretation using Cornu Spiral. The intensity $|\psi|^2 \propto |\Delta\mathcal{E}_\sigma|^2$ is proportional to the squared length of the vector whose tail is at the center of the lower left spiral and whose tip moves along the spiral curve.

Third, in integrating over the whole area of the wave front at $\mathcal{Q}$, we have summed contributions with increasingly large phase differences that add in such a way that the total has a net extra phase of $\pi/2$, relative to the geometric-optics ray. This phase factor cancels exactly the prefactor $-i$ in the Fresnel-Kirchhoff integral, Eq. (8.6). (This phase factor is unimportant in the limit of geometric optics.)

8.4.3 Fresnel Diffraction by a Straight Edge: Lunar Occultation of a Radio Source

The next simplest case of Fresnel diffraction is the pattern formed by a straight edge. As a specific example, consider a cosmologically distant source of radio waves that is occulted by the moon. If we treat the lunar limb as a straight edge, then the radio source will create a changing diffraction pattern as it passes behind the moon, and the diffraction pattern can be measured by a radio telescope on earth. We orient our coordinates so the moon's edge is along the y' direction (t direction). Then in Eq. (8.23a) $\Delta\mathcal{E}_\tau \equiv \mathcal{E}(\tau_{\rm max}) - \mathcal{E}(\tau_{\rm min}) = \sqrt{2i}$ is constant, and $\Delta\mathcal{E}_\sigma \equiv \mathcal{E}(\sigma_{\rm max}) - \mathcal{E}(\sigma_{\rm min})$ is described by the Cornu spiral.

Long before the occultation, $\Delta\mathcal{E}_\sigma$ will be given by the arrow from $(-1/2, -1/2)$ to $(1/2, 1/2)$, i.e. $\Delta\mathcal{E}_\sigma = \sqrt{2i}$. The observed wave amplitude, Eq. (8.23a), is therefore $\psi_Q e^{ikz}$. When the moon starts to occult the radio source, the upper bound on the Fresnel integral begins to diminish from $\sigma_{max} = +\infty$, and the complex vector on the Cornu spiral begins to oscillate in length (e.g., from A to B in Fig. 8.11) and in phase. The observed flux will also oscillate, more and more strongly as geometric occultation is approached. At the point of geometric occultation (point C in Fig. 8.11), the complex vector extends from $(-1/2, -1/2)$ to $(0, 0)$ and so the observed wave amplitude is one half the unocculted value, and the intensity is reduced to one fourth. As the occultation proceeds, the length of the complex

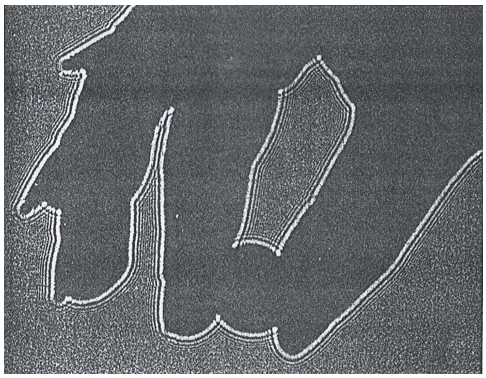

Fig. 8.12: Fresnel diffraction pattern in the shadow of *Mary's hand holding a dime* — a photograph by Eugene Hecht, from Fig. 10.1 of Hecht (1998).

vector and the observed flux will decrease monotonically to zero, while the phase continues to oscillate.

Historically, diffraction of a radio source's waves by the moon led to the discovery of quasars—the hyperactive nuclei of distant galaxies. In the early 1960s, a team of British radio observers led by Cyril Hazard knew that the moon would occult a powerful radio source named 3C273, so they set up their telescope to observe the development of the diffraction pattern as the occultation proceeded. From the pattern's observed times of ingress (passage into the moon's shadow) and egress (emergence from the moon's shadow), Hazard determined the coordinates of 3C273 on the sky. These coordinates enabled Maarten Schmidt at the 200-inch telescope on Palomar Mountain to identify 3C273 optically and discover (from its optical redshift) that it was surprisingly distant and consequently had an unprecedented luminosity.

In Hazard's occultation measurements, the observing wavelength was $\lambda \sim 0.2$ m. Since the moon is roughly $z \sim 400,000$ km distant, the Fresnel length was about $r_F = \sqrt{\lambda z} \sim 10$ km. The moon's orbital speed is $v \sim 200$ m s^{-1}, so the diffraction pattern took a time $\sim 5r_F/v \sim 4$ min to pass through the telescope.

The straight-edge diffraction pattern of Fig. 8.11 occurs universally along the edge of the shadow of any object, so long as the source of light is sufficiently small and the shadow's edge bends on lengthscales long compared to the Fresnel length $r_F = \sqrt{\lambda z}$. Examples are the diffraction patterns on the two edges of a slit's shadow in the upper left curve in Fig. 8.3, and the diffraction pattern along the edge of a shadow cast by a person's hand in Fig. 8.12.

8.4.4 Circular Apertures: Fresnel Zones and Zone Plates

We have shown how the Fresnel diffraction pattern for a plane wave can be thought of as formed by waves that derive from a patch a few Fresnel lengths in size. This notion can be made quantitatively useful by reanalyzing the unobstructed wave front in circular polar coordinates. More specifically: consider, a plane wave incident on an aperture $\mathcal{Q}$ that is infinitely large (no obstruction), and define $\varpi \equiv |\mathbf{x}'|/r_F = \sqrt{\frac{1}{2}(\sigma^2 + \tau^2)}$. Then the phase

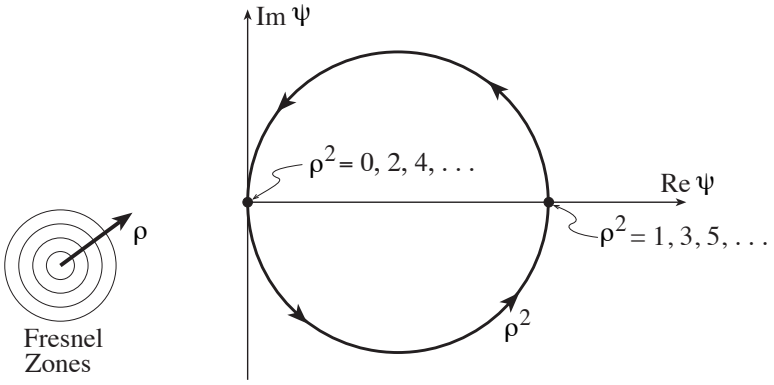

Fig. 8.13: Amplitude-and-phase diagram for an unobstructed plane wave front, decomposed into Fresnel zones; Eq. (8.24).

factor in Eq. (8.22b) is $\Delta\varphi = \pi\varpi^2$ and the observed wave will thus be given by

$$\begin{aligned} \psi_{\mathcal{P}} &= -i\int_0^{\varpi} \pi\varpi d\varpi e^{i\pi\varpi^2}\psi_{\mathcal{Q}}e^{ikz} \\ &= (1-e^{i\pi\varpi^2})\psi_{\mathcal{Q}}e^{ikz}\ . \end{aligned} \tag{8.24}$$

Now, this integral does not appear to converge as $\varpi \to \infty$. We can see what is happening if we sketch an amplitude-and-phase diagram (Fig. 8.13). Adding up the contributions to $\psi_{\mathcal{P}}$ from each annular ring, we see that as we integrate outward from $\varpi = 0$, the complex vector has the initial phase retardation of $\pi/2$ but then moves on a semi-circle so that by the time we have integrated out to a radius of r_F ($\varpi = 1$), the contribution to the observed wave is $\psi_{\mathcal{P}} = 2\psi_{\mathcal{Q}}$ in phase with the incident wave. Then, when the integration has been extended onward to $\sqrt{2}\,r_F$, ($\varpi = \sqrt{2}$), the circle has been completed and $\psi_{\mathcal{P}} = 0$! The integral continues on around the same circle as the upper-bound radius is further increased.

Of course, the field must actually have a well-defined value, despite this apparent failure of the integral to converge. To understand how the field becomes well-defined, imagine splitting the aperture $\mathcal{Q}$ up into concentric annular rings, known as *Fresnel half-period zones*, of radius $\sqrt{n}\,r_F$, where $n = 1, 2, 3\ldots$. The integral fails to converge because the contribution from each odd-numbered ring cancels that from an adjacent even-numbered ring. However, the thickness of these rings decreases as $1/\sqrt{n}$, and eventually we must allow for the fact that the incoming wave is not exactly planar; or, equivalently and more usefully, we must allow for the fact that the wave's distant source has some finite angular size. The finite size causes different pieces of the source to have their Fresnel rings centered at slightly different points in the aperture plane, and this causes our computation of $\psi_{\mathcal{P}}$ to begin averaging over rings. This averaging forces the tip of the complex vector to asymptote to the center of the circle in Fig. 8.13. Correspondingly, due to the averaging, the observed intensity asymptotes to $|\psi_{\mathcal{Q}}|^2$ (Eq. (8.24) with the exponential going to zero).

Although this may not seem to be a particularly wise way to decompose a plane wave front, it does allow a particularly striking experimental verification of our theory of diffraction. Suppose that we fabricate an aperture (called a *zone plate*) in which, for a chosen

observation point $\mathcal{P}$ on the optic axis, alternate half-period zones are obscured. Then the wave observed at $\mathcal{P}$ will be the linear sum of several diameters of the circle in Fig. 8.13, and therefore will be far larger than $\psi_{\mathcal{Q}}$. This strong amplification is confined to our chosen spot on the optic axis; most everywhere else the field's intensity is reduced, thereby conserving energy. Thus, the zone plate behaves like a lens. The lens's focal length is $f = kA/2\pi^2$, where A (typically chosen to be a few mm^2 for light) is the common area of each of the half-period zones.

Zone plates are only good lenses when the radiation is monochromatic, since the focal length is wavelength-dependent, $f \propto \lambda^{-1}$. They have the further interesting property that they possess secondary foci, where the fields from $3, 5, 7, \ldots$ contiguous zones add up coherently (Ex. 8.9).

****

EXERCISES

Exercise 8.8 *Exercise: Diffraction Pattern from a Slit*

Derive a formula for the intensity diffraction pattern $F(x)$ of a slit with width a, as a function of distance x from the center of the slit, in terms of Fresnel integrals. Plot your formula for various distances z from the slit's plane, i.e. for various values of $r_F/a = \sqrt{\lambda z/a^2}$ (using, e.g., Mathematica or Maple), and compare with Fig. 7.3.

Exercise 8.9 *Problem: Zone Plate*

(a) Use an amplitude-and-phase diagram to explain why a zone plate has secondary foci at distances of $f/3, f/5, f/7 \ldots$.

(b) An opaque, perfectly circular disk of diameter D is placed perpendicular to an incoming plane wave. Show that, at distances r such that $r_F \ll D$, the disk casts a rather sharp shadow, but at the precise center of the shadow there should be a bright spot.[4] How bright?

Exercise 8.10 *Example: Seeing in the atmosphere.*

Stars viewed through the atmosphere appear to have angular diameters of order an arc second and to exhibit large amplitude fluctuations of flux with characteristic frequencies that can be as high as 100Hz. Both of these phenomena are a consequence of irregular variations in the refractive index of the atmosphere. An elementary model of this effect consists of a thin phase-changing screen, about a km above the ground, on which the rms phase variation is $\Delta\varphi \gtrsim 1$ and the characteristic spatial scale, on which the phase changes by $\sim \Delta\varphi$, is a.

[4]Poisson predicted the existence of this spot as a consequence of Fresnel's wave theory of light, in order to demonstrate that Fresnel's theory was wrong. However, Dominique Arago quickly demonstrated experimentally that the bright spot existed.

(a) Explain why the rays will be irregularly deflected through a scattering angle $\Delta\theta \sim (\lambda/a)\Delta\varphi$. Strong intensity variation requires that several rays deriving from points on the screen separated by more than a, combine at each point on the ground. These rays combine to create a diffraction pattern on the ground with scale b.

(b) Show that the Fresnel length in the screen is $\sim \sqrt{ab}$. Now the time variation arises because winds in the upper atmosphere with speeds $u \sim 30\text{m s}^{-1}$ blow the irregularities and the diffraction pattern past the observer. Use this information to estimate the Fresnel length, r_F, the atmospheric fluctuation scale size a, and the rms phase variation $\Delta\varphi$. Do you think the assumptions of this model are well satisfied?

Exercise 8.11 *Challenge: Multi-Conjugate Adaptive Optics*

The technique of *Adaptive Optics* can be used to improve the quality of the images observed by a telescope. Bright artificial "laser stars" are created by shining several lasers at layers of sodium atoms in the upper atmosphere and observing the scattered light. The wavefronts from these "stars" will be deformed at the telescope due to inhomogeneities in the lower atmosphere, and the deformed wavefront shapes can be measured across the image plane. The light from a much dimmer adjacent astronomical source can then be processed, e.g. using a deformable reflecting surface, so as to remove its wavefront distortions. Discuss some of the features that an effective adaptive optics system needs. Assume the atmospheric model discussed in Ex. 8.10.

Exercise 8.12 *Problem: Spy Satellites*

Telescopes can also look down through the same atmospheric irregularities as those discussed in the previous example. In what important respects will the optics differ from that for telescopes looking upward?

****

8.5 Paraxial Fourier Optics

We have developed a linear theory of wave optics which has allowed us to calculate diffraction patterns in the Fraunhofer and Fresnel limiting regions. That these calculations agree with laboratory measurements provides some vindication of the theory and the assumptions implicit in it. We now turn to practical applications of these ideas, specifically to the *acquisition and processing of images by instruments operating throughout the electromagnetic spectrum.* As we shall see, these instruments rely on an extension of paraxial geometric optics (Sec. 6.4) to situations where diffraction effects are important. Because of the central role played by Fourier transforms in diffraction [e.g. Eq. (8.11a)], the theory underlying these instruments is called paraxial Fourier optics, or just Fourier optics.

Although the conceptual framework and mathematical machinery for image processing by Fourier optics were developed over a century ago, Fourier optics has only been widely exploited during the past thirty years. This maturation has been driven in part by a growing recognition of similarities between optics and communication theory — for example, the realization that a microscope is simply an image processing system. The development of electronic computation has also triggered enormous strides; computers are now seen as extensions of optical devices, and vice versa. It is a matter of convenience, economics and practicality to decide which parts of the image processing are carried out with mirrors, lenses, etc., and which parts are performed numerically.

One conceptually simple example of optical image processing would be an improvement in one's ability to identify a faint star in the Fraunhofer diffraction rings ("fringes") of a much brighter star. As we shall see below [Eq. (8.30) and subsequent discussion], the bright image of a source in a telescope's or microscope's focal plane has the same Airy diffraction pattern as we met in Eq. (8.18) and Fig. 8.6. If the shape of that image could be changed from the ring-endowed Airy pattern to a Gaussian, then it would be far easier to identify a nearby feature or faint star. One way to achieve this would be to attenuate the incident radiation at the telescope aperture in such a way that, immediately after passing through the aperture, it has a Gaussian profile instead of a sharp-edged profile. Its Fourier transform (the diffraction pattern in the focal plane) would then also be a Gaussian. Such a Gaussian-shaped attenuation is difficult to achieve in practice, but it turns out—as we shall see–that there are easier options.

Before exploring these options, we must lay some foundations, beginning with the concept of coherent illumination in Sec. 8.5.1, and then point spread functions in Sec. 8.5.2.

8.5.1 Coherent Illumination

If the radiation that arrives at the input of an optical system derives from a single source, e.g. a point source that has been collimated into a parallel beam by a converging lens, then the radiation is best described by its complex amplitude ψ (as we are doing in this chapter). An example might be a biological specimen on a microscope slide, illuminated by an external point source, for which the phases of the waves leaving different parts of the slide are strongly correlated with each other. This is called *coherent illumination*. If, by contrast, the source is self luminous and of non-negligible size, with the atoms or molecules in its different parts radiating independently—for example a cluster of stars—then the phases of the radiation from different parts are uncorrelated, and it may be the radiation's intensity, not its complex amplitude, that obeys well-defined (non-probabilistic) evolution laws. This is called *incoherent illumination*. In this chapter we shall develop Fourier optics for a coherently illuminating source (the kind of illumination tacitly assumed in previous sections of the chapter) . A parallel theory with a similar vocabulary can be developed for incoherent sources, and some of the foundations for it will be laid in Chap. 8. In Chap. 8 we shall also develop a more precise formulation of the concept of *coherence*.

8.5.2 Point Spread Functions

In our treatment of paraxial geometric optics (Sec. 6.4), we showed how it is possible to regard a group of optical elements as a sequence of linear devices and relate the output rays to the input by linear operators, i.e. matrices. This chapter's theory of diffraction is also linear and so a similar approach can be followed. As in Sec. 6.4, we will restrict attention to small angles relative to some optic axis ("paraxial Fourier optics"). We shall describe the wave field at some distance z_j along the optic axis by the function $\psi_j(\mathbf{x})$, where $\mathbf{x}$ is a two dimensional vector perpendicular to the optic axis as in Fig. 8.4. If we consider a single linear optical device, then we can relate the output field ψ_2 at z_2 to the input ψ_1 at z_1 using a Green's function denoted $P_{21}(\mathbf{x}_2, \mathbf{x}_1)$:

$$\boxed{\psi_2(\mathbf{x}_2) = \int P_{21}(\mathbf{x}_2, \mathbf{x}_1) d\Sigma_1 \psi_1 \ .} \tag{8.25}$$

If ψ_1 were a δ-function, then the output would be simply given by the function P_{21}, up to normalization. For this reason, P_{21} is usually known as the *Point Spread Function*. Alternatively, we can think of it as a *propagator*. If we now combine two optical devices sequentially, so the output of the first device ψ_2 is the input of the second, then the point spread functions combine in the natural manner of any linear propagator to give a total point spread function

$$\boxed{P_{31}(\mathbf{x}_3, \mathbf{x}_1) = \int P_{32}(\mathbf{x}_3, \mathbf{x}_2) d\Sigma_2 P_{21}(\mathbf{x}_2, \mathbf{x}_1) \ .} \tag{8.26}$$

Just as the simplest matrix for paraxial, geometric-optics propagation is that for free propagation through some distance d, so also the simplest point spread function is that for free propagation. From Eq. (8.22b) we see that it is given by

$$\boxed{P_{21} = \frac{-ik}{2\pi d} e^{ikd} \exp\left(\frac{ik(\mathbf{x}_1 - \mathbf{x}_2)^2}{2d}\right) \quad \text{for free propagation through a distance } d = z_2 - z_1 \ .} \tag{8.27}$$

Note that this P_{21} depends upon only on $\mathbf{x}_1 - \mathbf{x}_2$ and not on $\mathbf{x}_1$ or $\mathbf{x}_2$ individually, as it should because there is translational invariance in the $\mathbf{x}_1, \mathbf{x}_2$ planes.

A thin lens adds or subtracts an extra phase $\Delta\varphi$ to the wave, and $\Delta\varphi$ depends quadratically on distance $|\mathbf{x}|$ from the optic axis, so the angle of deflection, which is proportional to the gradient of the phase, will depend linearly on $\mathbf{x}$. Correspondingly, the point-spread function for a thin lens is

$$\boxed{P_{21} = \exp\left(\frac{-ik|\mathbf{x}_1|^2}{2f}\right) \delta(\mathbf{x}_2 - \mathbf{x}_1) \quad \text{for a thin lens with focal length } f \ .} \tag{8.28}$$

For a converging lens, f is positive; for a diverging lens, it is negative.

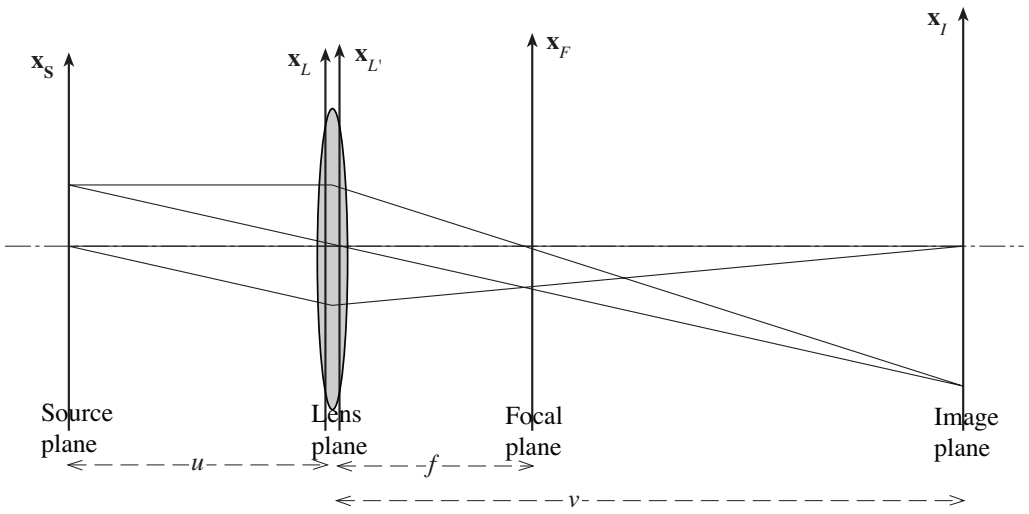

Fig. 8.14: Wave theory of a single converging lens. The focal plane is a distance f =(lens focal length) from the lens plane; and the impage plane is a distance $v = fu/(u-f)$ from the lens plane.

8.5.3 Abbé's Description of Image Formation by a Thin Lens

We can use these two point spread functions to give a wave description of the production of images by a single converging lens, in parallel to the geometric-optics description of Figs. 6.5 and 6.7. We shall do this in two stages. First, we shall propagate the wave from the source plane S a distance u in front of the lens, through the lens L, to its focal plane F a distance f behind the lens (Fig. 8.14). Then we shall propagate the wave a further distance $v-f$ from the focal plane to the image plane. We know from geometric optics that $v = fu/(u-f)$ [Eq. (6.60)]. We shall restrict ourselves to $u > f$ so v is positive and the lens forms a real image.

Using Eqs. (8.26), (8.27), (8.28), we obtain for the propagator from the source plain to the focal plane

$$\begin{aligned} P_{FS} &= \int P_{FL'} d\Sigma_{L'} P_{L'L} d\Sigma_L L P_{LS} \\ &= \int \frac{ik}{2\pi f} e^{ikf} \exp\left(\frac{ik(\mathbf{x}_F - \mathbf{x}'_L)^2}{2f}\right) d\Sigma_{L'} \delta(\mathbf{x}_{L'} - \mathbf{x}_L) \exp\left(\frac{-ik|\mathbf{x}_L|^2}{2f}\right) \\ &\quad \times d\Sigma_L \frac{-ik}{2\pi u} e^{iku} \exp\left(\frac{ik(\mathbf{x}_L - \mathbf{x}_S)^2}{2u}\right) \\ &= \frac{-ik}{2\pi f} e^{ik(f+u)} \exp\left(-\frac{ikx_F^2}{2(v-f)}\right) \exp\left(-\frac{ik\mathbf{x}_F \cdot \mathbf{x}_S}{f}\right) . \end{aligned} \tag{8.29}$$

Here we have extended all integrations to $\pm\infty$ and have used the values of the Fresnel integrals at infinity, $\mathcal{E}(\pm\infty) = \pm(1+i)/2$ to get the expression on the last line. The wave in the focal plane is given by $\psi_F(\mathbf{x}_F) = \int P_{FS} d\Sigma_S \psi_S(\mathbf{x}_S)$, which integrates to

$$\boxed{\psi_F(\mathbf{x}_F) = -\frac{ik}{2\pi f} e^{ik(f+u)} \exp\left(-\frac{ikx_F^2}{2(v-f)}\right) \tilde{\psi}_S(\mathbf{x}_F/f) .} \tag{8.30}$$

Here

$$\tilde{\psi}_S(\boldsymbol{\theta}) = \int d\Sigma_S \psi_S(\mathbf{x}_S) e^{-ik\boldsymbol{\theta}\cdot\mathbf{x}_S} \ . \tag{8.31}$$

Thus, we have shown that the field in the focal plane is, apart from an unimportant phase factor, proportional to the Fourier transform of the field in the source plane; in other words, *the focal-plane field is the Fraunhofer diffraction pattern of the input wave.* That this has to be the case can be understood from Fig. 8.14. The focal plane F is where the converging lens brings parallel rays from the source plane to a focus. By doing so, *the lens in effect brings in from "infinity" the Fraunhofer diffraction pattern of the source, and places it into the focal plane.*

It now remains to propagate the final distance from the focal plane to the image plane. We do so with the free-propagation point-spread function of Eq. (8.27): $\psi_I = \int P_{IF} d\Sigma_F \psi_F$, which integrates to

$$\boxed{\psi_I(\mathbf{x}_I) = -\left(\frac{u}{v}\right) e^{ik(u+v)} \exp\left(\frac{ikx_I^2}{2(v-f)}\right) \psi_S(\mathbf{x}_S = -\mathbf{x}_I u/v) \ .} \tag{8.32}$$

This says that (again ignoring a phase factor) *the wave in the image plane is just a magnified version of the wave in the source plane*, as we might have expected from geometric optics. In words, *the lens acts by taking the Fourier transform of the source and then takes the Fourier transform again to recover the source structure.*[5]

The focal plane is a convenient place to process the image by altering its Fourier transform—a process known as *spatial filtering.* One simple example is a *low-pass filter* in which a small circular aperture or "stop" is introduced into the focal plane, thereby allowing only the low-order spatial Fourier components to be transmitted to the image plane. This will lead to considerable smoothing of the wave. An application is to the output beam from a laser (Chap. 9), which ought to be smooth but has high spatial frequency structure on account of noise and imperfections in the optics. A low-pass filter can be used to clean the beam. In the language of Fourier transforms, if we multiply the transform of the source, in the focal plane, by a small-diameter circular aperture function, we will thereby convolve the image with a broad Airy-disk smoothing function. Conversely, we can exclude the low spatial frequencies with a high-pass filter, e.g. by placing an opaque circular disk in the focal plane, centered on the optic axis. This will have the effect of accentuating boundaries and discontinuities in the source and can be used to highlight features where the gradient of the brightness is large. Another type of filter is used when the image is pixellated and thus has unwanted structure with wavelength equal to pixel size: a narrow range of frequencies centered around this spatial frequency is removed by putting an appropriate filter in the focal plane.

8.5.4 Phase Contrast Microscopy

"Phase contrast microscopy" (Fig. 8.15) is a useful technique for studying small objects, such as transparent biological specimens, that modify the phase of coherent illuminating light but not its amplitude. Suppose that the phase change in the specimen, $\varphi(\mathbf{x})$, is small, $|\varphi| \ll 1$,

[5]This description of image formation was developed by Ernst Abbé in 1873.

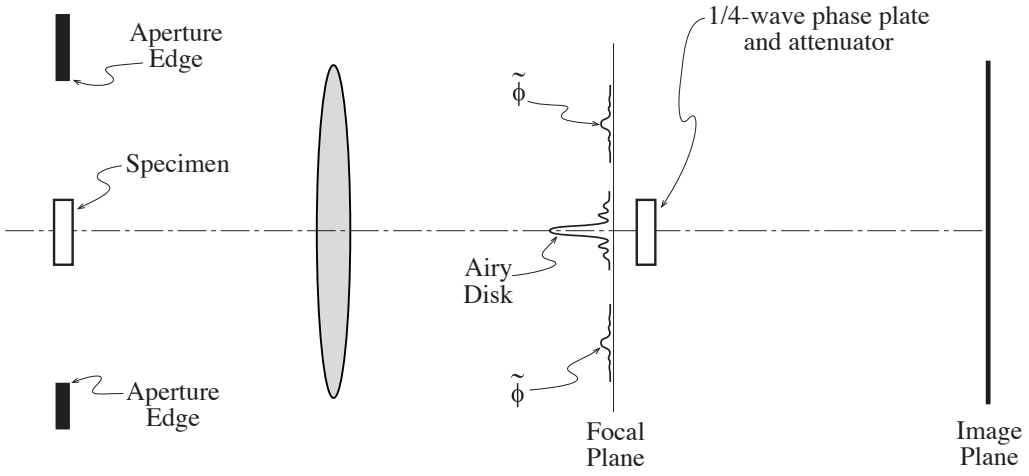

Fig. 8.15: Schematic Phase Contrast Microscope.

as often is the case for biological specimens. We can then write the field just after it passes through the specimen as

$$\psi_S(\mathbf{x}) = H(\mathbf{x})e^{i\varphi(\mathbf{x})} \simeq H(\mathbf{x}) + i\varphi(\mathbf{x})H(\mathbf{x}) \; ; \tag{8.33}$$

Here H is the microscope's aperture function, unity for $|\mathbf{x}| < D/2$ and zero for $|\mathbf{x}| > D/2$, with D the aperture diameter. The intensity is not modulated, and therefore the effect of the specimen on the wave is very hard to observe unless one is clever.

Equation (8.33) and the linearity of the Fourier transform imply that the wave in the focal plane is the sum of (i) the Fourier transform of the aperture function, i.e. an Airy function (bright spot with very small diameter), and (ii) the transform of the phase function convolved with that of the aperture (in which the fine-scale variations of the phase function dominate and push $\tilde{\varphi}$ to large radii in the focal plane, Fig. 8.15, and the aperture has little influence):

$$\psi_F \sim \mathrm{jinc}\left(\frac{kD|\mathbf{x}_F|}{2f}\right) + i\tilde{\varphi}\left(\frac{k\mathbf{x}_F}{f}\right) \; . \tag{8.34}$$

If a high pass filter is used to remove the Airy disk completely, then the remaining wave field in the image plane will be essentially φ magnified by v/u. However, the intensity $F \propto (\varphi v/u)^2$ will be quadratic in the phase and so the contrast in the image will still be small. A better technique is to phase shift the Airy disk in the focal plane by $\pi/2$ so that the two terms in Eq. (8.34) are in phase. The intensity variations, $F \sim (1 \pm \varphi)^2 \simeq 1 \pm 2\varphi$, will now be linear in the phase φ. An even better procedure is to attenuate the Airy disk until its amplitude is comparable with the rms value of φ and also phase shift it by $\pi/2$ (as indicated by the "1/4 wave phase plate and attenuation" in Fig. 8.15). This will maximise the contrast in the final image. Analogous techniques are used in communications to interconvert amplitude-modulated and phase-modulated signals.

8.5.5 Gaussian Beams: Interferometric Gravitational-Wave Detectors

The mathematical techniques of Fourier optics enable us to analyze the structure and propagation of light beams that have Gaussian profiles. (Such Gaussian beams are the natural output of ideal lasers, they are the real output of spatially filtered lasers, and they are widely used for optical communications, interferometry and other practical applications. Moreover, they are the closest one can come in the real world of wave optics to the idealization of a geometric-optics pencil beam.)

Consider a beam that is precisely plane-fronted, with a Gaussian profile, at location $z = 0$ on the optic axis,

$$\psi_0 = \exp\left(\frac{-\varpi^2}{\sigma_0^2}\right) \; ; \tag{8.35}$$

here $\varpi = |\mathbf{x}|$ is radial distance from the optic axis. The form of this same wave at a distance z further down the optic axis can be computed by folding this ψ_0 into the point spread function (8.27) (with the distance d replaced by z). The result is

$$\boxed{\psi_z = \frac{\sigma_0}{\sigma_z}\exp\left(\frac{-\varpi^2}{\sigma_z^2)}\right)\exp\left[i\left(\frac{k\varpi^2}{2R_z} + kz - \tan^{-1}\frac{z}{z_0}\right)\right],} \tag{8.36a}$$

where

$$\boxed{z_0 = \frac{k\sigma_0^2}{2} = \frac{\pi\sigma_0^2}{\lambda}\;, \quad \sigma_z = \sigma_0(1 + z^2/z_0^2)^{1/2}\;, \quad R_z = z(1 + z_0^2/z^2)\;.} \tag{8.36b}$$

These equations for the freely propagating Gaussian beam are valid for negative z as well as positive.

From these equations we learn the following properties of the beam:

- *The beam's cross sectional intensity distribution* $F \propto |\psi_z|^2 \propto \exp(-\varpi^2/\sigma_z^2)$ *remains a Gaussian as the wave propagates, with a beam radius* $\sigma_z = \sigma_0\sqrt{1 + z^2/z_0^2}$ *that is a minimum at* $z = 0$ *(the beam's* waist*) and grows away from the waist, both forward and backward, in just the manner one expects from our uncertainty-principle discussion of wave-front spreading* [Eq. (8.9)]. At distances $|z| \ll z_0$ from the waist location (corresponding to a Fresnel length $r_F = \sqrt{\lambda|z|} \ll \sqrt{\pi}\sigma_0$), the beam radius is nearly constant; this is the Fresnel region. At distances $z \gg z_0$ ($r_F \gg \sqrt{\pi}\sigma_0$), the beam radius increases linearly, i.e. the beam spreads with an opening angle $\theta = \sigma_0/z_0 = \lambda/(\pi\sigma_0)$; this is the Fraunhofer region.

- The beam's wave fronts (surfaces of constant phase) have $\varphi = k\varpi^2/2R_z + kz - \tan^{-1}(z/z_0) =$ constant. The tangent term (called the wave's *Guoy phase*) varies far far more slowly with changing z than does the kz term, so the wave fronts are almost precisely $z = \varpi^2/2R_z$+constant, which is a segment of a sphere of radius R_z. Thus, *the wave fronts are spherical, with radii of curvature* $R_z = z(1 + z_0^2/z^2)$*, which is infinite (flat phase fronts) at the waist* $z = 0$*, increases to* $2z_0$ *at* $z = z_0$ *(boundary between*

Fresnel and Fraunhofer regions and beginning of substantial wave front spreading), and then decreases as z_0^2/z (gradually flattening of spreading wave fronts) as one moves deep into the Fraunhofer region.

- *The Gaussian beam's form (8.36) at some arbitrary location is fully characterized by three parameters:* the wavelength $\lambda = 2\pi/k$, the distance z to the waist, and the beam radius at the waist σ_0 [from which one can compute the local beam radius σ_z and the local wave front radius of curvature R_z via Eqs. (8.36b).

One can easily compute the effects of a thin lens on a Gaussian beam by folding the ψ_z at the lens's location into the lens point spread function (8.28). The result is a phase change that preserves the general Gaussian form of the wave, but alters the distance z to the waist and the radius σ_0 at the waist. Thus, by judicious placement of lenses (or, equally well curved mirrors), and with judicious choices of the lenses' and mirrors' focal lengths, one can tailor the parameters of a Gaussian beam to fit whatever optical device one is working with. For example, if one wants to send a Gaussian beam into a self-focusing optical fiber (Exs. 6.7 and 8.14), one should place its waist at the entrance to the fiber, and adjust its waist size there to coincide with that of the fiber's Gaussian mode of propagation (the mode analyzed in Ex. 8.14). The beam will then enter the fiber smoothly, and will propagate steadily along the fiber, with the effects of the transversely varying index of refraction continually compensating for the effects of diffraction so as to keep the phase fronts flat and the waist size constant.

Gaussian beams are used (among many other places) in *interferometric gravitational-wave detectors*, such as LIGO (the Laser Interferometer Gravitational-wave Observatory). We shall learn how these *GW interferometers* work in Sec. 8.5. For the present, all we need to know is that a GW interferometer entails an *optical cavity* formed by mirrors facing each other, as in Fig. 6.8 of Chap. 6. A Gaussian beam travels back and forth between the two mirrors, with its light superposing on itself coherently after each round trip, i.e. the light *resonates* in the cavity formed by the two mirrors. Each mirror hangs from an overhead support, and when a gravitational wave passes, it pushes the hanging mirrors back and forth with respect to each other, causing the cavity to lengthen and shorten by a very tiny fraction of a light wavelength. This puts a tiny phase shift on the resonating light, which is measured by allowing some of the light to leak out of the cavity and interfere with light from another, similar cavity. See Sec. 8.5.

In order for the light to resonate in the cavity, the mirrors' surfaces must coincide with the Gaussian beam's wave fronts. Suppose that the mirrors are identical, with radii of curvature R, and are separated by a distance $L = 4$km, as in LIGO. Then the beam must be symmetric around the center of the cavity, so its waist must be half-way between the mirrors. What is the smallest that the beam radius can be, at the mirrors' locations $z = \pm L/2 = \pm 2$km? From $\sigma_z = \sigma_0(1 + z^2/z_0^2)^{1/2}$ together with $z_0 = \pi\sigma_0^2/\lambda$, we see that $\sigma_{L/2}$ is minimized when $z_0 = L/2 = 2$km. If the wavelength is $\lambda = 1.06\mu$m (Nd:YAG laser light) as in LIGO, then the beam radii at the waist and at the mirrors are $\sigma_0 = \sqrt{\lambda z_0/\pi} = \sqrt{\lambda L/2\pi} = 2.6$cm, and $\sigma_z = \sqrt{2}\sigma_0 = 3.7$cm, and the mirrors' radii of curvature are $R_{L/2} = L = 4$km. This was approximately the regime of parameters used for LIGO's initial GW interferometers, which carried out a two-year-long search for gravitational waves from autumn 2005 to autumn

2007.

A new generation of GW interferometers, called "Advanced LIGO", is in preparation. In these GW interferometers, the spot sizes on the mirrors will be made much larger, so as to reduce thermal noise by averaging over a much larger spatial sampling of thermal fluctuations of the mirror surfaces (cf. Sec. 10.5 and Exs. 5.8 and 10.14). How can the spot sizes on the mirrors be enlarged? From Eqs. (8.36b) we see that, in the limit $z_0 = \pi\sigma_0^2/\lambda \to 0$, the mirrors' radii of curvature approach the cavity half-length, $R_{L/2} \to L/2$, and the beam radii on the mirrors diverge as $\sigma_{L/2} \to L\lambda/(2\pi\sigma_0) \to \infty$. This is the same instability as we discovered, in the geometric optics limit, in Ex. 6.11. Advanced LIGO takes advantage of this instability by moving toward the near-unstable regime, causing the beams on the mirrors to enlarge. The mirrors' radiiof curvature are set at $R_{L/2} = 2.079$km, just 4 per cent above the unstable point $R = L/2 = 2$km; and Eqs. (8.36b) then tell us that $\sigma_0 = 1.16$cm, $z_0 = 0.399$km $\ll L/2 = 2$km, and σ_z has been pushed up by nearly a factor two, to $\sigma_z = 5.93$cm. The mirrors are deep into the Fraunhofer, wave-front-spreading region.

****

EXERCISES

Exercise 8.13 *Problem: Convolution via Fourier Optics*

(a) Suppose that you have two thin sheets with transmission functions $t = g(x, y)$ and $t = h(x, y)$, and you wish to compute via Fourier optics the convolution

$$g \otimes h(x_o, y_o) \equiv \int\int g(x, y)h(x + x_o, y + y_o)dxdy \; . \tag{8.37}$$

Devise a method for doing so using Fourier optics. [Hint: use several lenses and a projection screen with a pinhole through which passes light whose intensity is proportional to the convolution; place the two sheets at strategically chosen locations along the optic axis, and displace one of the two sheets transversely with respect to the other.]

(b) Suppose you wish to convolve a large number of different one-dimensional functions simultaneously, i.e. you want to compute

$$g_j \otimes h_j(x_o) \equiv \int g_j(x)h_j(x + x_o)dx \tag{8.38}$$

for $j = 1, 2, \ldots$. Devise a way to do this via Fourier optics using appropriately constructed transmissive sheets and cylindrical lenses.

Exercise 8.14 *Problem: Guided Gaussian Beams*

Consider a self-focusing optical fiber discussed in Sec. 6.7, in which the refractive index is

$$\mathfrak{n}(\mathbf{x}) = \mathfrak{n}_0(1 - \alpha^2\varpi^2)^{1/2} \; , \tag{8.39}$$

where $\varpi = |\mathbf{x}|$.

(a) Write down the Helmholtz equation in cylindrical polar coordinates and seek an axisymmetric mode for which $\psi = R(\varpi)Z(z)$, where R, Z are functions to be determined and z measures distance along the fiber. In particular show that there exists a mode with a Gaussian radial profile that propagates along the fiber without spreading.

(b) Compute the group and phase velocities along the fiber for this mode.

Exercise 8.15 *Exercise: Noise Due to Scattered Light in LIGO*

In LIGO and other GW interferometers, one potential source of noise is scattered light: When the Gaussian beam in one of LIGO's cavities reflects off a mirror, a small portion of the light gets scattered toward the walls of the cavity's vacuum tube. Some of this scattered light can reflect or scatter off the tube wall and then propagate toward the distant mirror, where it scatters back into the Gaussian beam; see Fig. 8.16 (without the baffles that are shown dashed). This is troublesome because the tube wall vibrates due to sound-wave excitations and seismic excitations, and those vibrations put a phase shift on the scattered light. Although the fraction of all the light that scatters in this way is tiny, the phase shift is huge compared to that produced in the Gaussian beam by gravitational waves; and when the tiny amount of scattered light with its huge oscillating phase shift recombines into the Gaussian beam, it produces a net Gausian-beam phase shift that can be large enough to mask a gravitational wave. This exercise will explore some aspects of this scattered noise and its control.

(a) The scattering of Gaussian-beam light off the mirror is caused by bumps in the mirror surface (imperfections). Denote by $h(\mathbf{x})$ the height of the mirror surface, relative to the desired shape (a segment of a sphere with radius of curvature that matches the Gaussian beam's wave fronts). Show that, if the Gaussian-beam field emerging from a perfect mirror is $\psi^G(\mathbf{x})$ [Eq. (8.36)] at the mirror plane, then the beam emerging from the actual mirror is $\psi'(\mathbf{x}) = \psi^G(\mathbf{x}) \exp[-i2kh(\mathbf{x})]$. The magnitude of the mirror irregularities is very small compared to a wavelength, so $|2kh| \ll 1$, and the wave field emerging from the mirror is $\psi'(\mathbf{x}) = \psi^G(\mathbf{x})[1 - i2kh(\mathbf{x})]$. Explain why the factor 1 does not contribute at all to the scattered light (where does its light go?), so the scattered light field, emerging from the mirror, is

$$\psi^S(\mathbf{x}) = -i\psi^G(\mathbf{x})2kh(\mathbf{x}) \ . \tag{8.40}$$

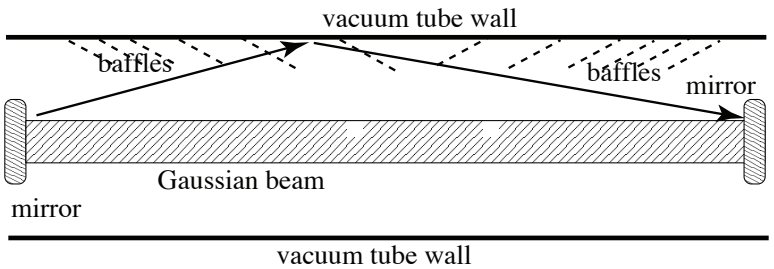

Fig. 8.16: Scattered light in LIGO's beam tube.

(b) Assume that when, arriving at the vacuum-tube wall, the scattered light is in the Fraunhofer region. You will justify this below. Then at the tube wall, the scattered light field is given by the Fraunhofer formula

$$\psi^S(\boldsymbol{\theta}) \propto \int \psi^G(\mathbf{x}) k h(\mathbf{x}) e^{ik\mathbf{x}\cdot\boldsymbol{\theta}} \, . \tag{8.41}$$

Show that the light that hits the tube wall at an angle $\theta = |\boldsymbol{\theta}|$ to the optic axis arises from irregularities in the mirror that have spatial wavelengths $\lambda_{\rm mirror} \sim \lambda/\theta$. The radius of the beam tube is $\mathcal{R} = 60$cm in LIGO and the length of the tube (distance between cavity mirrors) is $L = 4$km. What is the spatial wavelength of the mirror irregularities which scatter light to the tube wall at distances $z \sim L/2$ (which can then reflect or scatter off the wall toward the distant mirror and there scatter back into the Gaussian beam)? Show that for these irregularities, the tube wall is, indeed, in the Fraunhofer region. (Hint: the irregularities have a coherence length of only a few wavelengths $\lambda_{\rm mirror}$.)

(c) In the initial LIGO interferometers, the mirrors' scattered light consisted of two components: one peaked strongly toward small angles so it hit the distant tube wall, e.g. at $z \sim L/2$, and the other roughly isotropically distributed. What was the size of the irregularities that produced the isotropic component?

(d) To reduce substantially the amount of scattered light reaching the distant mirror via reflection or scattering from the tube wall, a set of *baffles* was installed in the tube, in such a way as to hide the wall from scattered light (dashed lines in Fig. 8.16). The baffles have an angle of 35° to the tube wall, so when light hits a baffle, it reflects at a steep angle, $\sim 70^0$ toward the opposite tube wall and after a few bounces gets absorbed. However, a small portion of the scattered light can now *diffract* off the top of each baffle and then propagate to the distant mirror and scatter back into the main beam. Especially troublesome is the case of a mirror in the center of the beam tube's cross section, because light that scatters off such a mirror travels nearly the same total distance from mirror to the top of some baffle and then to the distant mirror, independent of the azimuthal angle ϕ on the baffle at which it diffracts. There is then a danger of *coherent superposition* of all the scattered light that diffracts off all angular locations around any given baffle—and coherent superposition means a much enlarged net noise. To protect against any such coherence, the baffles in the LIGO beam tubes are serrated, i.e. they have saw-tooth edges, and the heights of the teeth are drawn from a random (Gaussian) probability distribution. The typical tooth heights are large enough to extend through about six Fresnel zones. *Questions for which part (e) may be helpful:* How wide is each Fresnel zone at the baffle location, and correspondingly, how high must be the typical baffle tooth? By approximately how much do the random serrations reduce the light-scattering noise, relative to what it would be with no serrations and with coherent scattering?

(e) To aid you in answering part (d), show that the propagator (point spread function) for light that begins at the center of one mirror, travels to the edge of a baffle [at a

radial distance $R(\phi)$ from the beam-tube axis, where ϕ is azimuthal angle around the beam tube, and at a distance ℓ down the tube from the scattering mirror] and that then Propagates to the center of the distant mirror is

$$P =\propto \exp\left(\frac{ikR^2(\phi)}{2\ell_{\rm red}}\right) d\phi \ , \quad \text{where } \frac{1}{\ell_{\rm red}} = \frac{1}{\ell} + \frac{1}{L-\ell} \ . \tag{8.42}$$

Note that $\ell_{\rm red}$ is the "reduced baffle distance" by analogy with the "reduced mass" in a binary system. One can show that the time-varying part of the scattered-light amplitude (i.e. the part whose time dependence is produced by baffle vibrations) is proportional to this propagator. Explain why this is plausible. Then explain how the baffle serrations, embodied in the ϕ dependence of $R(\phi)$, produce the reduction of scattered-light amplitude in the manner described in part (c).

****

8.6 Diffraction at a Caustic

In Sec. 6.5, we described how caustics can be formed in general in the geometric-optics limit—e.g., on the bottom of a swimming pool when the water's surface is randomly rippled, or behind a gravitational lens. We chose as an example a simple phase changing screen illuminated by a point source and observed from some fixed distance r, and we showed how a pair of images would merge as the transverse distance x of the observer from the caustic decreases to zero. We expanded the phase in a Taylor series, $\varphi(s,x) = as^3/3 - bxs$, where the coefficients a, b are constant and s is a transverse coordinate in the screen (cf. Fig. 6.11). We were then able to show that the magnification of the images diverged $\propto x^{-1/2}$ [Eq. (6.75)] as the caustic was approached, then crashed to zero just past the caustic (the two images disappeared). This singular behavior raised the question of what happens when we take into account the finite wavelength of the wave.

We are now in a position to answer this question. We simply use the Helmholtz-Kirchhoff integral (8.6) to write the expression for the amplitude measured at position x in the form

$$\psi(x) \propto \frac{1}{\lambda r}\int ds e^{i\varphi(s,x)} = \frac{1}{\lambda r}\int ds(\cos\varphi + i\sin\varphi) \ , \tag{8.43}$$

ignoring multiplicative constants and constant phase factors. The phase φ varies rapidly with s at large $|s|$, so we can treat the limits of integration as $\pm\infty$. Because $\varphi(s,x)$ is odd in s, the sin term integrates to zero, and the integral turns out to be the Airy function

$$\psi \propto \frac{1}{\lambda r}\int_{-\infty}^{\infty} ds\cos(as^3/3 - bxs) = \frac{1}{\lambda r}\frac{2\pi}{a^{1/3}}{\rm Ai}(-bx/a^{1/3}) \ . \tag{8.44}$$

${\rm Ai}(\xi)$ is displayed in Fig. 8.17.

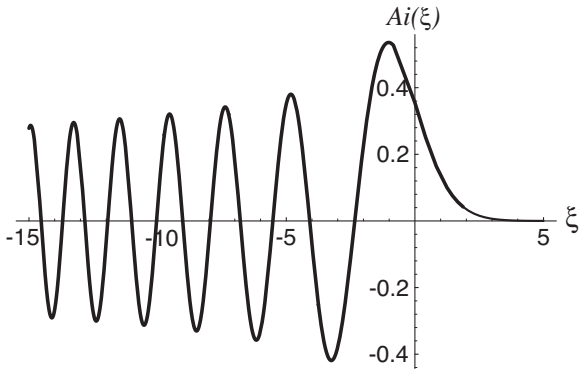

Fig. 8.17: The Airy Function Ai(z) describing diffraction at a caustic. The argument is $z = -bx/a^{1/3}$ where x is distance from the caustic and a, b are constants.

The asymptotic behavior of Ai(ξ) is

$$\begin{aligned} \mathrm{Ai}(\xi) &\sim \pi^{-1/2}\xi^{-1/4}\sin(2\xi^{3/2}/3+\pi/4), \quad \xi\to-\infty \\ &\sim \frac{e^{-2\xi^{3/2}/3}}{2\pi^{1/2}\xi^{1/4}}, \quad \xi\to\infty\ . \end{aligned} \tag{8.45}$$

We see that the amplitude ψ remains finite as the caustic is approached instead of diverging as in the geometric-optics limit, and it decreases smoothly toward zero when the caustic is past, instead of crashing instantaneously to zero. For $x > 0$ ($\xi = -bx/a^{1/3} < 0$; left part of Fig. 8.17), where an observer sees two geometric-optics images, the envelope of ψ diminishes $\propto x^{-1/4}$, so the intensity $|\psi|^2$ decreases $\propto x^{-1/2}$ just as in the geometric-optics limit. The peak magnification is $\propto a^{-2/3}$. What is actually seen is a series of bands alternating dark and light with spacing calculable using $\Delta(2\xi^{3/2}/3) = \pi$ or $\Delta x \propto x^{-1/2}$. At sufficient distance from the caustic, it will not be possible to resolve these bands and a uniform illumination of average intensity will be observed, so we recover the geometric-optics limit.

The near-caustic scalings derived above and others in Ex. 8.16, like the geometric-optics scalings [text following Eq. (6.75)] are a universal property of this type of caustic (the simplest caustic of all, the "fold").

There is a helpful analogy, familiar from quantum mechanics. Consider a particle in a harmonic potential well in a very excited state. Its wave function is given in the usual way using Hermite polynomials of large order. Close to the classical turning point, these functions change from being oscillatory to an exponential decay, just like the Airy function (and if we were to expand about the turning point, we would recover Airy functions). What is happening, of course, is that the probability density of finding the particle close to its turning point diverges classically because it is moving vanishingly slowly at the turning point; the oscillations are due to interference between waves associated with the particle moving in opposite directions at the turning point.

For light near a caustic, If we consider the motions of photons transverse to the screen, then we have essentially the same problem. The field's oscillations are associated with interference of the waves associated with the motions of the photons in two geometric-optics

beams coming from slightly different directions and thus having slightly different transverse photon speeds.

This is our first illustration of the formation of large-contrast interference fringes when only a few beams are combined. We shall meet other examples of such interference in the following chapter.

****

EXERCISES

Exercise 8.16 *Problem: Wavelength scaling at a caustic*

Assume that the phase variation introduced at the screen in Fig. 6.12 is non-dispersive so that the $\varphi(s, x)$ in Eq. (8.43) is $\varphi \propto \lambda^{-1}$. Show that the peak magnification of the interference fringes at the caustic scales with wavelength $\propto \lambda^{-4/3}$. Also show that the spacing of the fringes at a given observing position is $\propto \lambda$.

****

Box 8.2
Important Concepts in Chapter 7

- Helmholtz equation for a propagating, monochromatic wave – Eq. (8.1b)
- Helmholtz-Kirchhoff integral – Eq. (8.4)
- Complex transmission function – Eq. (8.5)
- Helmholtz-Kirchilff for wave propagating through an aperture – Eqs. (8.6), (8.7)
- Fresnel and Fraunhofer Diffraction compared:
  - Fresnel length and criteria for Fraunhofer and Fresnel Regions – Sec. 8.2.2
  - Qualitative forms of diffraction in Fresnel and Fraunhofer regions – Sec. 8.2.2
  - Wavefront spreading, at angle $\theta \sim \lambda/a$, in Fresnel regions – Sec. 8.2.2
- Fraunhofer Diffraction – Sec. 8.3
  - Diffracted field as Fourier transform of transmission function $\mathfrak{t}(\boldsymbol{\theta})$ – Eq. (8.11a)
  - Diffracted intensity $F \propto |\mathfrak{t}(\boldsymbol{\theta})|^2$ – Eq. (8.11b)
  - Diffraction patterns for a slit and a circular aperture – bottom curve in Fig. 8.3, Eqs. (8.12), Sec. 8.3.2
  - Use of convolution theorem to analyze diffraction grating – Sec. 8.3.1
  - Babinet's principle – Sec. 8.3.3
- Fresnel Diffraction – Sec. 8.4
  - As integral over the aperture with quadratically varying phase – Eqs. 8.22
  - For rectangular aperture, slit, and straight edge, in terms of Fresnel integrals and Cornu spiral – Secs. 8.4.1 and 8.4.3.
- Paraxial Fourier Optics
  - Point spread functions, as propagators down the optic axis – Sec. 8.5.2
  - Thin lens: field at focal plane as Fourier transform of source field (Fraunhofer region brought to focus) – Eq. (8.30)
  - Thin lens: field at image plane as inverted and magnified source field – Eq. (8.32)
  - Image and signal processing by optical techniques (e.g., phase contrast microscope) – last paragraph of Sec. 8.5.3, plus Sec. 8.5.4
- Gaussian beams – Sec. 8.5.5
  - Evolution of beam radius and phase front curvature – Eqs. (**??**)
  - Manipulating Gaussian beams with lenses and mirrors – Sec. 8.5.5
- Diffraction at a caustic: Airy pattern – Sec. 8.6

Bibliographic Note

Hecht (1998) has an excellent treatment of diffraction at roughly the same level as this chapter, but much more detailed. For a more advanced treatment, including mixing of polarizations by interaction of an electromanetic wave with the edges of apertures and other objects, see Born and Wolf (1999). Other good texts are listed in the bibliography:

Bibliography

Berry, M. V. & Upstill, C. 1980 *Prog. Optics* **18**, 257

Born, M. & Wolf, E. 1999 *Principles of Optics*, Seventh Edition, Cambridge: Cambridge University Press

Goodman, J. W. *Introduction to Fourier Optics*, New York: McGraw-Hill

Hecht, E. 1998 *Optics*, Third Edition, New York: Addison Wesley

Longhurst, R. S. 1973 *Geometrical and Physical Optics*, London: Longmans

Welford, W. T. 1988 *Optics*, Oxford: Oxford University Press