

5.4 Fraunhofer Diffraction

In this section, several important examples of Fraunhofer diffraction are considered in detail. Recall that, in the Fraunhofer region behind an aperture, the wave emitted by each secondary source in the aperture is essentially a plane wave by the time it reaches the observation plane. Implications are that the relationship between the field distribution in the aperture and the observation plane follows the simple Fourier transform relationship of Eq. (5.49), which is repeated here

$$U_0(x_0, y_0) = -\frac{je^{jkz_0}}{\lambda z_0} \exp\left[j\frac{k}{2z_0}(x_0^2 + y_0^2)\right] \mathbf{F}_{\eta=\frac{y_0}{\lambda z_0}} \mathbf{F}_{\xi=\frac{x_0}{\lambda z_0}} \left[U_S^+(x_S, y_S)\right], \quad (5.49)$$

where z_0 is the distance from the aperture plane to the observation plane, $k = 2\pi/\lambda$, (x_0, y_0) are coordinates of the observation plane, (x_S, y_S) are coordinates of the aperture plane, and $U_S^+(x_S, y_S)$ is the field transmitted through the aperture. Tables of Fourier transform pairs and relationships are found in Appendix B.

An equivalent expression, when $U_S^+(x_S, y_S)$ is rotationally symmetric in the $x_S y_S$ plane, follows from application of the Hankel transform of Eq. (B.14), which is repeated here

$$G(\rho, \phi) = G_p(\rho) = 2\pi \int_0^\infty r g_R(r) J_0(2\pi r \rho) dr = \mathbf{B}_\rho [g_R(r)], \quad (B.14)$$

where ρ is the radial coordinate of the transform space, J_0 is a zero-order Bessel function, and r is the radial coordinate of the integrand. $g_R(r)$ is the radial component of the function under transformation. If we set

$$\rho_0 = \sqrt{x_0^2 + y_0^2} \quad (5.132)$$

and

$$\rho_S = \sqrt{x_S^2 + y_S^2}, \quad (5.133)$$

then Eq. (5.49) becomes

$$\begin{aligned} U_0(\rho_0) &= -\frac{je^{jkz_0}}{\lambda z_0} \exp\left(j\frac{k\rho_0^2}{2z_0}\right) 2\pi \int_0^\infty \text{circ}\left(\frac{\rho_S}{2a}\right) U_S^+(\rho_S) J_0\left(\frac{2\pi\rho_0\rho_S}{\lambda z_0}\right) \rho_S d\rho_S \\ &= -\frac{je^{jkz_0}}{\lambda z_0} \exp\left(j\frac{k\rho_0^2}{2z_0}\right) \mathbf{B}_{\sigma_0=\frac{\rho_0}{\lambda z_0}} \left[\text{circ}\left(\frac{\rho_S}{2a}\right) U_S^+(\rho_S)\right], \end{aligned} \quad (5.134)$$

where a is the radius of the circular aperture. Tables of Hankel transforms and properties are found in Appendix B.

The usefulness of Eqs. (5.49) and (5.134) cannot be overstated. When an observation plane is sufficiently far from an aperture illuminated by a laser, a simple Fourier or Hankel transform determines the distribution in the observation plane. This relationship is almost always true in optical imaging systems between the object and the entrance pupil, and between the exit pupil and the image. It can be recognized in many other practical applications as well, where the observation distance z_0 satisfies Eq. (5.47), which is repeated here

$$z_0 \gg \frac{k}{2} (x_1^2 + y_1^2)_{\max} \quad . \quad (5.47)$$

Equivalently, the distance is specified when the Fresnel number of the observation plane is such that

$$N_f = \frac{a^2}{\lambda L} \ll 1 \quad , \quad (5.135)$$

where L is the effective propagation distance defined by Eq. (5.101). As shown in Fig. 5.56, if $N_f < 0.1$, the primary changes in the pattern are changes in scale, rather than changes in shape.

Sections 5.4.1 through 5.4.4 describe Fraunhofer diffraction from a circular aperture, the exit pupil in an imaging system, a rectangular aperture and slits, respectively.

5.4.1 Circular aperture

Fraunhofer diffraction from a circular aperture is one of the most important problems to understand in optics, because it determines the fundamental limits of performance in simple imaging systems. It also determines the focused spot size when a laser beam is brought to focus with a lens.

Imagine a unit amplitude plane wave normally incident on a circular aperture. Since this problem is circularly symmetric, Eq. (5.134) is applied so that

$$U_0(\rho_0) = -\frac{je^{jkz_0}}{\lambda z_0} \exp\left(j \frac{k\rho_0^2}{2z_0}\right) \mathbf{B}_{\sigma_0 = \frac{\rho_0}{\lambda z_0}} \left[\text{circ}\left(\frac{\rho_s}{2a}\right) \right] \quad . \quad (5.136)$$

The transform indicated in Appendix B implies

$$U_0(\rho_0) = -\frac{je^{jkz_0}}{\lambda z_0} \exp\left(j \frac{k\rho_0^2}{2z_0}\right) \pi a^2 \text{somb}\left(\frac{2a\rho_0}{\lambda z_0}\right) \quad , \quad (5.137)$$

and the irradiance associated with Eq. (5.137) is¹

$$I_0(\rho_0) = \left(\frac{1}{\lambda z_0} \right)^2 (\pi a^2)^2 \text{somb}^2 \left(\frac{2a\rho_0}{\lambda z_0} \right) \quad (5.138)$$

Notice that Eq. (5.137) is proportional to the aperture area πa^2 , and the irradiance in Eq. (5.138) is proportional to the square of the aperture area and inversely proportional to the wavelength and the square of the distance between the aperture plane and the observation plane. Equation (5.138) is known as the *Airy disk* and is displayed in Fig. 5.57. The $\text{somb}^2(\rho_0)$ portion of the Airy disk is displayed in detail in Fig. B.6.

5.4.2 Exit pupil of an imaging system

The fundamental limit of resolution in an optical system is determined by diffraction from the exit pupil to the image plane. As shown in Fig. 5.58, the function of a lens is to map an object point, like an expanding spherical wave, into an image point, like a converging spherical wave. When the converging spherical wave passes through the limiting aperture of the system (the exit pupil), diffraction occurs. As a straightforward application of Fresnel diffraction theory, we now analyze this problem with respect to a perfect lens that changes only wavefront curvature.

We will assume that the converging spherical wavefront at the exit pupil is well described by the parabolic approximation

$$\phi_{EXP}(\rho_{EXP}) = -\frac{2\pi}{\lambda} \frac{\rho_{EXP}^2}{2z_0} = -\frac{k}{2z_0} \rho_{EXP}^2 \quad (5.139)$$

where ρ_{EXP}^2 is the radial coordinate in the plane of the pupil. Application of the Fresnel formula in the form of Eq. (5.44) with $x_S = x_{EXP}$ and $y_S = y_{EXP}$ and $x_0 y_0$ defining the image plane yields

$$\begin{aligned} U_0(x_0, y_0) &= -\frac{j e^{jkz_0}}{\lambda z_0} \exp\left(j \frac{k\rho_0^2}{2z_0}\right) \mathbf{B}_{\frac{\rho_0}{\lambda z_0}} \left[\text{circ}\left(\frac{\rho_{EXP}}{2a}\right) e^{-j \frac{k}{2z_0} \rho_{EXP}^2} e^{j \frac{k}{2z_0} \rho_{EXP}^2} \right] \\ &= -\frac{j e^{jkz_0}}{\lambda z_0} \exp\left(j \frac{k\rho_0^2}{2z_0}\right) \mathbf{B}_{\frac{\rho_0}{\lambda z_0}} \left[\text{circ}\left(\frac{\rho_{EXP}}{2a}\right) \right] \\ &= \frac{j e^{jkz_0}}{\lambda z_0} \exp\left(j \frac{k\rho_0^2}{2z_0}\right) \pi a^2 \text{somb}\left(\frac{2a\rho_0}{\lambda z_0}\right) \quad , \end{aligned} \quad (5.140)$$

¹ In this section, the leading factors $\left(\frac{1}{2} cn\epsilon_0^2\right)$ relating the magnitude square of the electric field to the irradiance are ignored for clarity.

which is identical to Eq. (5.137). The quadratic phase factor of the converging wave exactly cancels the Fresnel phase factor at the image plane. That is, the field at the image plane of a perfect imaging system resulting from a point-source object is the Fourier transform of the exit pupil. This result makes intuitive sense from the discussion of Fresnel diffraction in Section 5.3.2.4, where $N_f \rightarrow 0$ at the center of wavefront curvature. The Fourier transform relationship between exit pupil and point-source image is also true for other aperture shapes, like rectangles. Notice that Fresnel ring patterns resulting from the circular aperture are observed in the focus cone of the lens, as displayed in Figs. 5.37 and 5.38. The Fourier transform relationship is strictly valid only at the image plane.

The Airy disk irradiance distribution of the focus spot has particular characteristics that define resolution of the image. For example, *Rayleigh's criterion* states that images resulting from two incoherent point sources are barely resolved if the principle maximum of one spot lies on the first minimum of the other. In the perfect imaging system described by Fig. 5.58, that minimum resolvable distance is approximately

$$\rho_{RES} = 1.22 \frac{\lambda z_0}{2a} = 0.61 \frac{\lambda}{NA} = 1.22 \lambda f / \# \quad , \quad (5.141)$$

where NA is the *numerical aperture* of the image cone and $f/\#$ is the f -number. Other resolution criteria lead to similar results. For example, the more restrictive *Sparrow criterion* suggests that the minimum resolvable distance is

$$\rho_{RES-SPARROW} = 0.94 \frac{\lambda z_0}{2a} = 0.47 \frac{\lambda}{NA} = 0.94 \lambda f / \# \quad . \quad (5.142)$$

Examples of both criteria are shown in Fig. 5.59. In practice, the resolution criterion of a particular instrument depends on many factors, such as noise and illumination conditions, so the definitions of Eqs. (5.141) and (5.142) are often only a starting point in determining the minimum resolvable distance. For example, if the point-source objects are coherent or partially coherent with respect to each other, the mutual coherence function must be used to calculate the image and determine resolvability. Either resolution criteria can also be mapped into the object space by simply applying the transverse magnification factor m_T to Eq. (5.141) or (5.142).

A second consequence of diffraction from the exit pupil is that minimum spot size for the system is determined by the Airy disk for a uniformly filled pupil. It is physically impossible to achieve a smaller spot without applying some type of filtering in the pupil that leads to decreased peak irradiance. Also, a Gaussian-amplitude distribution illuminating the pupil, like from a laser beam, changes the focus distribution slightly, as shown in Fig. 5.60. The “softer” edge of the Gaussian pupil effectively lowers the NA slightly, which produces a slightly larger spot diameter and less energy in the rings.

The exit pupil may not be a physical aperture, like that shown in Fig. 5.58. A more general description is that the image pupil is the image of the limiting aperture as seen by

an observer at the image plane. The physical limiting aperture is the *aperture stop* and is usually located somewhere inside the system at a location based on some desirable characteristic. Likewise, the *entrance pupil* is the image of the aperture stop as seen by an observer from the object.

A general lens system is sketched in Fig. 5.61. The first-order description is determined from locations of the principle planes (P in object space and P' in image space), aperture stop S , object distance o , and image distance i . Transverse magnification of the system is $m_T = -i/o$. Exit pupil radius is a . Locations of entrance and exit pupils are determined from lens prescriptions between object and aperture stop and between aperture stop and image, respectively. Notice that locations of pupils and principle planes are not limited to the region containing the physical lens. Their locations vary substantially between lenses, depending on the lens prescription. Good assumptions for most systems are that the entrance pupil is in the Fraunhofer region of the object and that the exit pupil field distribution is related to the entrance pupil field distribution through a simple magnification and power normalization calculation. For image-plane calculations, the field distribution in the exit pupil and the propagation distance z_0 must be known.

5.4.3 Aberrations

Real lens systems are not perfect, and the converging wavefront in image space from a point object is usually distorted to some degree. Distortion of the wavefront is called *aberration*. It is characterized by function $W(\bar{\rho}_{EXP}, \theta_{EXP})$ with polar coordinates, where $\bar{\rho}_{EXP}$ is the normalized exit pupil radius given by

$$\bar{\rho}_{EXP} = \frac{\rho_{EXP}}{a}, \quad (5.143)$$

a is the exit pupil radius and θ_{EXP} is the angle from the y_{EXP} axis. According to Hopkins,² the function $W(\bar{\rho}_{EXP}, \theta_{EXP})$ for a symmetrical optical system can be written to third order as

$$\begin{aligned} W(\bar{\rho}_{EXP}, \theta_{EXP}) &= W(\bar{y}_0, \bar{\rho}_{EXP}, \theta_{EXP}) \\ &= W_{020} \bar{\rho}_{EXP}^2 + W_{111} \bar{y}_0 \bar{\rho}_{EXP} \cos \theta_{EXP} + W_{040} \bar{\rho}_{EXP}^4 + W_{131} \bar{y}_0 \bar{\rho}_{EXP}^3 \cos \theta_{EXP} + \\ &\quad W_{222} \bar{y}_0^2 \bar{\rho}_{EXP}^2 \cos^2 \theta_{EXP} + W_{220} \bar{y}_0^2 \bar{\rho}_{EXP}^2 + W_{311} \bar{y}_0^3 \bar{\rho}_{EXP} \cos \theta_{EXP}, \end{aligned} \quad (5.144)$$

where \bar{y}_0 is the geometrical field height at the observation plane without aberrations normalized to the maximum field height. Subscripts of the W_{mnl} coefficients refer to field height, pupil radius and pupil angle, respectively, and coefficients are typically specified in units of wavelengths of OPD. The first two terms describe defocus and tilt about the x

² H.H. Hopkins, Wave Theory of Aberrations, Clarendon Press, Oxford (1950), p. 48

axis. The remaining five terms are the Seidel aberrations: spherical, coma, astigmatism, field curvature and distortion, respectively. The expansion in Eq. (5.144) has physical significance, because the terms relate directly to commonly observed properties of optical systems.

An alternative method for describing $W(\bar{\rho}_{EXP}, \theta_{EXP})$ is Zernike polynomials,³ which are convenient because they can be orthogonalized easily when using measured data. A table of the first 15 Zernike terms as described by Malacara is shown in Table 5.2, along with their monomial representation and meaning.⁴ Notice that terms often contain components of mixed power. Lower-order components balance W to minimize variance.

The effects of defocus (W_{020}), spherical (W_{040}), coma (W_{131}) and astigmatism (W_{222}) are to significantly modify the focused spot distribution. Tilt (W_{111}) simply moves the center of the spot away from the optical axis, field curvature (W_{220}) is a field-dependent defocus and distortion (W_{331}) is a field-dependent tilt. Defocus can be thought of as a change in observation-plane location Δz_0 away from the paraxial image plane, where

$$\Delta z_0 \approx -2\lambda \left(\frac{z_0}{a} \right)^2 W_{020} , \quad (5.145)$$

and z_0 is the distance from the exit pupil to the paraxial focus. In Eq. (5.145), Δz_0 is small compared to z_0 . Interestingly, Δz_0 corresponds directly to a change in Fresnel number N_f as defined in Section 5.3.2.4, because both are quadratic functions of aperture radius. The relationship is

$$N_f = 2W_{020} , \quad (5.146)$$

where W_{020} is specified in waves. Graphical representations of W_{020} , W_{040} , W_{131} and W_{222} are shown in Figs. 5.62 through 5.65, respectively. Further discussion of aberration effects is presented in Chapter 9.

Table 5.2 Zernike Polynomial Expansion.

Zernike Polynomial	Monomial Representation	Meaning
1	1	Piston (constant term)
$\bar{\rho}_{EXP} \sin \theta_{EXP}$	\bar{x}_{EXP}	Tilt about \bar{x}_{EXP} axis
$\bar{\rho}_{EXP} \cos \theta_{EXP}$	\bar{y}_{EXP}	Tilt about \bar{y}_{EXP} axis
$\bar{\rho}_{EXP}^2 \sin 2\theta_{EXP}$	$2\bar{x}_{EXP}\bar{y}_{EXP}$	Astigmatism with axis at 45°
$\bar{\rho}_{EXP}^2 - 1$	$-1 + 2\bar{y}_{EXP}^2 + 2\bar{x}_{EXP}^2$	Focus shift with piston

³ D. Malacara, *Optical Shop Testing, Second Ed.*, John Wiley and Sons, Inc., New York (1992), p. 461

⁴ There are several other types of Zernike expansions, and the reader is cautioned to understand the particular expansion being used when evaluating Zernike coefficients.

$\bar{\rho}_{EXP}^2 \cos 2\theta_{EXP}$	$\bar{y}_{EXP}^2 - \bar{x}_{EXP}^2$	Astigmatism with axis at 0° or 90°
$\bar{\rho}_{EXP}^3 \sin 3\theta_{EXP}$	$3\bar{x}_{EXP}\bar{y}_{EXP}^2 - \bar{x}_{EXP}^3$	Triangular astigmatism with base on \bar{x}_{EXP} axis.
$(3\bar{\rho}_{EXP}^3 - 2\bar{\rho}_{EXP})\sin \theta_{EXP}$	$-2\bar{x}_{EXP} + 3\bar{x}_{EXP}\bar{y}_{EXP}^2 + 3\bar{x}_{EXP}^3$	Third-order coma along \bar{x}_{EXP} axis with compensating tilt
$(3\bar{\rho}_{EXP}^3 - 2\bar{\rho}_{EXP})\cos \theta_{EXP}$	$-2\bar{y}_{EXP} + 3\bar{y}_{EXP}^3 + 3\bar{x}_{EXP}^2\bar{y}_{EXP}$	Third-order coma along \bar{y}_{EXP} axis with compensating tilt
$\bar{\rho}_{EXP}^3 \cos 3\theta_{EXP}$	$\bar{y}_{EXP}^3 - 3\bar{x}_{EXP}^2\bar{y}_{EXP}$	Triangular astigmatism with base on \bar{y}_{EXP} axis
$\bar{\rho}_{EXP}^4 \sin 4\theta_{EXP}$	$4\bar{y}_{EXP}^3\bar{x}_{EXP} - 4\bar{x}_{EXP}^3\bar{y}_{EXP}$	
$(4\bar{\rho}_{EXP}^4 - 3\bar{\rho}_{EXP}^2)\sin 2\theta_{EXP}$	$-6\bar{x}_{EXP}\bar{y}_{EXP} + 8\bar{y}_{EXP}^3\bar{x}_{EXP} + 8\bar{x}_{EXP}^3\bar{y}_{EXP}$	
$6\bar{\rho}_{EXP}^4 - 6\bar{\rho}_{EXP}^2 + 1$	$1 - 6\bar{y}_{EXP}^2 - 6\bar{x}_{EXP}^2 + 6\bar{y}_{EXP}^4 + 12\bar{x}_{EXP}^2\bar{y}_{EXP}^2 + 6\bar{x}_{EXP}^4$	Third-order spherical aberration with compensating defocus and tilt
$(4\bar{\rho}_{EXP}^4 - 3\bar{\rho}_{EXP}^2)\cos 2\theta_{EXP}$	$-3\bar{y}_{EXP}^2 + 3\bar{x}_{EXP}^2 + 4\bar{y}_{EXP}^4 - 4\bar{x}_{EXP}^4$	
$\bar{\rho}_{EXP}^4 \cos 4\theta_{EXP}$	$\bar{y}_{EXP}^4 - 6\bar{x}_{EXP}^2\bar{y}_{EXP}^2 + \bar{x}_{EXP}^4$	

5.4.4 Wide-angle diffraction to a spherical surface

Several important concepts in diffraction theory applied to optical imaging systems require that the restrictions imposed in development of plane-to-plane Fraunhofer diffraction be removed. For example, objects often diffract light over very large angles, which are not strictly allowed in the Fraunhofer development of Section 5.2.6. In this section, these restrictions are removed by developing a diffraction formula for propagating from a plane to a spherical surface.

First, we return to the Raleigh-Sommerfeld diffraction formula of Eq. (5.20), which is restricted only by requiring that $r_{0S} \gg \lambda$. This equation is rewritten as

$$U_0(\mathbf{r}_0) = -\frac{j}{\lambda} \iint_{\infty} \gamma_z U_S^+(\boldsymbol{\rho}_S) \frac{e^{jk r_{0S}}}{r_{0S}} ds \quad . \quad (5.146)$$

The observation surface is now a sphere of radius r that has its center of curvature at $x_S = 0$ and $y_S = 0$, as shown in Fig. 5.66. The vector

$$\boldsymbol{\rho}_S = x_S \hat{\mathbf{x}} + y_S \hat{\mathbf{y}} \quad (5.147)$$

defines location in the aperture plane, and

$$\mathbf{r} + \mathbf{r}_{0S} = \boldsymbol{\rho}_S \quad (5.148)$$

with

$$\mathbf{r} = r(\alpha \hat{\mathbf{x}} + \beta \hat{\mathbf{y}} + \gamma \hat{\mathbf{z}}) \quad (5.149)$$

The scalar value r_{0S} is written as

$$\begin{aligned} r_{0S} &= |\boldsymbol{\rho}_S - \mathbf{r}| \\ &= r \sqrt{\left(\alpha - \frac{x_S}{r}\right)^2 + \left(\beta - \frac{y_S}{r}\right)^2 + \gamma^2} \\ &\approx r - (\alpha x_S + \beta y_S) + \frac{x_S^2 + y_S^2}{2r} + \dots \end{aligned} \quad (5.150)$$

where the first two terms in the square root expansion are sufficient for most practical problems and require only that $x_S \ll r$ and $y_S \ll r$. The third and higher-order terms in the expansion are conceptually interesting, in that they lead to *diffraction aberration*, as defined by Harvey and Shack.⁵ However, they are not considered further in this section.

The $1/r_{0S}$ term in the integrand and in $\gamma_z = z_0 / r_{0S}$ can be approximated by r without severe consequences on the accuracy of the result. Also, since $z_0 = \gamma r$, $\gamma_z \approx \gamma$. With these minor approximations, Eq. (5.146) becomes simply

$$\begin{aligned} U_0(\xi, \eta) &= -\frac{j\gamma e^{jkr}}{\lambda r} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_S^+(x_S, y_S) e^{j\frac{k}{2r}(x_S^2 + y_S^2)} e^{-j2\pi\left(\frac{\alpha}{\lambda}x_S + \frac{\beta}{\lambda}y_S\right)} dx_S dy_S \Bigg|_{\substack{\alpha=\xi, \\ \beta=\eta}} \\ &= -\frac{j\gamma e^{jkr}}{\lambda r} \mathbf{F}_{\frac{\beta}{\lambda}=\eta} \mathbf{F}_{\frac{\alpha}{\lambda}=\xi} \left[U_S^+(x_S, y_S) e^{j\frac{k}{2r}(x_S^2 + y_S^2)} \right] \end{aligned} \quad (5.151)$$

Notice that the quadratic phase that multiplies the transmitted field is a Fresnel-phase like factor that is cancelled completely when a converging wavefront illuminates the aperture with its center of curvature at the observation sphere vertex. Also, $U_0(\xi, \eta)$ is nearly equal to the angular spectrum $A_z(\xi, \eta; z_0)$ with this illumination condition, except for the leading constant factors and γ 's dependence on (ξ, η) . If

⁵ J. E. Harvey R. V. Shack, "Aberrations of diffracted wave fields," *Appl. Opt.* **17**(18), p. 3003 (1978).

$$\frac{k}{2r}(x_s^2 + y_s^2) \ll 1 \quad , \quad (5.151)$$

the quadratic phase factor can be ignored without regard to illumination, and

$$U_0(\xi, \eta) = -\frac{j\gamma e^{jkr}}{\lambda r} \mathbf{F}_{\beta=\eta} \mathbf{F}_{\alpha=\xi} \left[U_s^+(x_s, y_s) \right] \quad . \quad (5.152)$$

Note that the observation coordinate system is defined on a hemisphere, where

$$\begin{aligned} x_0 &= \lambda \xi r \\ y_0 &= \lambda \eta r \\ z_0 &= \gamma r \quad . \end{aligned} \quad (5.153)$$

Equation (5.152) is an important result, because now the entrance pupil can be considered to contain a reference sphere onto which is diffracted the Fourier spectrum of the object, without restriction on angular extent of the diffraction. That is, we consider the aperture in Fig. 5.66 to be the object and the sphere to be located with its vertex in the entrance pupil. The field distribution on the entrance pupil reference sphere is mapped onto a converging reference sphere in the exit pupil, where image-space and object-space direction cosines are related by

$$\alpha' = \frac{\alpha}{m_T} \quad (5.154)$$

and

$$\beta' = \frac{\beta}{m_T} \quad , \quad (5.155)$$

respectively, for an imaging system obeying the Abbe sine condition and image and object spaces in air. The diameter of the exit-pupil reference sphere is limited by the diameter of the exit pupil $2a$. In addition, the total power on the exit pupil reference sphere must be normalized to the total power passing the entrance pupil. Figure 5.67 illustrates this relationship.

5.4.5 Rectangular aperture

Fraunhofer diffraction from a rectangular aperture is also an important problem, because they are routinely encountered in laser systems. In this case,

$$U_s^+(x_s, y_s) = A \operatorname{rect}\left(\frac{x_s}{2a}\right) \operatorname{rect}\left(\frac{y_s}{2b}\right) \quad , \quad (5.156)$$

where A is the electric field amplitude and $2a$ and $2b$ are x_s and y_s widths of the aperture, respectively. Application of Eq. (5.49) results in

$$\begin{aligned} U_0(x_0, y_0) &= -\frac{j e^{jkz_0}}{\lambda z_0} \exp\left[\frac{k}{2z_0}(x_0^2 + y_0^2)\right] \mathbf{F}_{\eta=\frac{y_0}{\lambda z_0}} \mathbf{F}_{\xi=\frac{x_0}{\lambda z_0}} \left[A \operatorname{rect}\left(\frac{x_s}{2a}\right) \operatorname{rect}\left(\frac{y_s}{2b}\right) \right] \\ &= -\frac{j A e^{jkz_0}}{\lambda z_0} \exp\left[\frac{k}{2z_0}(x_0^2 + y_0^2)\right] 4ab \operatorname{sinc}\left(\frac{4ax_0}{\lambda z_0}\right) \operatorname{sinc}\left(\frac{4by_0}{\lambda z_0}\right) . \end{aligned} \quad (5.157)$$

The irradiance at the observation plane (or the image plane of an image system with a rectangular aperture) is

$$I_0(x_0, y_0) = \left(\frac{A}{\lambda z_0}\right)^2 (4ab)^2 \operatorname{sinc}^2\left(\frac{2ax_0}{\lambda z_0}\right) \operatorname{sinc}^2\left(\frac{2by_0}{\lambda z_0}\right) . \quad (5.158)$$

Notice that, like the case for the Airy spot with a circular aperture, the irradiance is proportional to the square of the aperture area. The irradiance from a rectangular aperture with $a = 2b$ is shown in Fig. 5.68. The width between zeros of the central spot is approximately $s_x = \frac{\lambda z_0}{a} = \frac{\lambda}{NA_x}$ in the x direction and $s_y = \frac{\lambda z_0}{b} = \frac{\lambda}{NA_y}$ in the y direction.

Notice that central spot width is inversely proportional to aperture dimension.

5.4.6 Diffraction from slits

Fraunhofer diffraction from slits is another common occurrence in laser systems. A slit is simply a long but narrow rectangle. From Eq. (5.158), if $a \gg b$, the resulting diffraction pattern is long in the y_0 direction and narrow in the x_0 direction. Combination of several slits results in addition of the fields from a single slit with appropriate linear phase factors due to shifts of the slits. We will assume $a \gg b$ and that analysis is performed only along the $x_0 = 0$ axis. That is, we use a one-dimensional treatment versus y_0 . Illumination consists of a normally incident plane wave with amplitude A .

Mathematically, the single-slit diffraction pattern of a slit is given by Eq. (5.158), where

$$I_0(0, y_0) = \left(\frac{A}{\lambda z_0}\right)^2 (4ab)^2 \operatorname{sinc}^2\left(\frac{2by_0}{\lambda z_0}\right) , \quad (5.159)$$

as shown in Fig. 5.69. For two symmetrical slits,

$$U_s^+(x_0, y_0) = A \operatorname{rect}\left(\frac{x_s}{2a}\right) \operatorname{rect}\left(\frac{y_s}{2b}\right) * [\delta(y_s - d/2) + \delta(y_s + d/2)] , \quad (5.160)$$

where d is the slit separation and the slits are symmetrical about the $x_0 = 0$ axis. The Fraunhofer pattern resulting from the symmetrical slits is

$$\begin{aligned}
 U_0(0, y_0) &= -\frac{je^{jkz_0}}{\lambda z_0} \exp\left(\frac{ky_0^2}{2z_0}\right) \mathbf{F}_{\eta=\frac{y_0}{\lambda z_0}} \mathbf{F}_{\xi=\frac{x_0}{\lambda z_0}} \left\{ A \operatorname{rect}\left(\frac{x_s}{2a}\right) \operatorname{rect}\left(\frac{y_s}{2b}\right) * [\delta(y_s + d/2) + \delta(y_s - d/2)] \right\} \Big|_{x_0=0} \\
 &= -\frac{jAe^{jkz_0}}{\lambda z_0} \exp\left(\frac{ky_0^2}{2z_0}\right) 4ab \operatorname{sinc}\left(\frac{2by_0}{\lambda z_0}\right) 2 \cos\left(\frac{2\pi dy_0}{2\lambda z_0}\right) ,
 \end{aligned}
 \tag{5.161}$$

and

$$\begin{aligned}
 I_0(0, y_0) &= 4 \left(\frac{A}{\lambda z_0}\right)^2 (4ab)^2 \operatorname{sinc}^2\left(\frac{2by_0}{\lambda z_0}\right) \cos^2\left(\frac{2\pi dy_0}{2\lambda z_0}\right) \\
 &= 4 \left(\frac{A}{\lambda z_0}\right)^2 (4ab)^2 \operatorname{sinc}^2\left(\frac{2by_0}{\lambda z_0}\right) \left[\frac{1}{2} + \frac{1}{2} \cos\left(\frac{2\pi dy_0}{\lambda z_0}\right) \right] .
 \end{aligned}
 \tag{5.162}$$

Irradiance of the single-slit pattern is now modulated by the cosine term, as shown in Fig. 5.69 for the case where the slits are separated by twice the slit width ($d/b = 4$). Notice that peak irradiance is now four times higher than with a single slit, due to the additional transmission area at the aperture plane. The cosine modulation produces multiple peaks inside the envelope of the sinc function. As the d/b ratio increases, the number of peaks inside the sinc envelope also increases. Figure 5.70 displays characteristics of the diffraction pattern as d/b increases by increasing the separation of the slits d and leaving the width b constant. The number of peaks inside the central lobe of the sinc envelope increases with increasing d/b . In fact, the number of peaks in the central lobe is equal to d/b for odd values and $d/b-1$ for even values.

For N slits that are equally spaced with pitch d ,

$$\begin{aligned}
U_0(0, y_0) &= -\frac{j e^{jkz_0}}{\lambda z_0} \exp\left(\frac{ky_0^2}{2z_0}\right) \mathbf{F}_{\eta=\frac{y_0}{\lambda z_0}} \mathbf{F}_{\xi=\frac{x_0}{\lambda z_0}} \left\{ A \operatorname{rect}\left(\frac{x_s}{2a}\right) \operatorname{rect}\left(\frac{y_s}{2b}\right) * \left[\sum_{n=-N/2+1/2}^{N/2-1/2} \delta(y_s - nd) \right] \right\} \Big|_{x_0=0} \\
&= -\frac{j A e^{jkz_0}}{\lambda z_0} \exp\left(\frac{ky_0^2}{2z_0}\right) 4ab \operatorname{sinc}\left(\frac{2by_0}{\lambda z_0}\right) \sum_{n=-N/2+1/2}^{N/2-1/2} e^{-j2\pi n \frac{dy_0}{\lambda z_0}} \\
&= -\frac{j A e^{jkz_0}}{\lambda z_0} \exp\left(\frac{ky_0^2}{2z_0}\right) 4ab \operatorname{sinc}\left(\frac{2by_0}{\lambda z_0}\right) e^{-j2\pi(-N/2+1/2)\frac{dy_0}{\lambda z_0}} \sum_{n=0}^{N-1} e^{-j2\pi n \frac{dy_0}{\lambda z_0}} \\
&= -\frac{j A e^{jkz_0}}{\lambda z_0} \exp\left(\frac{ky_0^2}{2z_0}\right) 4ab \operatorname{sinc}\left(\frac{2by_0}{\lambda z_0}\right) e^{j2\pi(N/2-1/2)\frac{dy_0}{\lambda z_0}} \left(\frac{1 - e^{-j2\pi N \frac{dy_0}{\lambda z_0}}}{1 - e^{-j2\pi \frac{dy_0}{\lambda z_0}}} \right) \\
&= -\frac{j A e^{jkz_0}}{\lambda z_0} \exp\left(\frac{ky_0^2}{2z_0}\right) 4ab \operatorname{sinc}\left(\frac{2by_0}{\lambda z_0}\right) e^{j2\pi N \frac{dy_0}{2\lambda z_0}} \frac{e^{-j2\pi N \frac{dy_0}{2\lambda z_0}} - e^{-j2\pi \frac{dy_0}{2\lambda z_0}}}{e^{j2\pi \frac{dy_0}{2\lambda z_0}} - e^{-j2\pi \frac{dy_0}{2\lambda z_0}}} \\
&= -\frac{j A e^{jkz_0}}{\lambda z_0} \exp\left(\frac{ky_0^2}{2z_0}\right) 4ab \operatorname{sinc}\left(\frac{2by_0}{\lambda z_0}\right) \frac{\sin\left(2\pi N \frac{dy_0}{2\lambda z_0}\right)}{\sin\left(2\pi \frac{dy_0}{2\lambda z_0}\right)},
\end{aligned} \tag{5.163}$$

where the identity

$$1 + \rho + \rho^2 + \dots + \rho^{N-1} = \frac{1 - \rho^N}{1 - \rho} \tag{5.164}$$

is used to simplify the summation. Irradiance associated with Eq. (5.163) is

$$I_0(0, y_0) = \left(\frac{A}{\lambda z_0}\right)^2 (4ab)^2 \operatorname{sinc}^2\left(\frac{2by_0}{\lambda z_0}\right) \left[\frac{\sin\left(2\pi N \frac{dy_0}{2\lambda z_0}\right)}{\sin\left(2\pi \frac{dy_0}{2\lambda z_0}\right)} \right]^2 . \quad (5.165)$$

As the number of slits increases, the diffraction pattern of each peak narrows, as shown in Fig. 5.71, where the slit width $2b$ is very small compared to the spacing d so that many peaks are present inside the central lobe. Notice that the number of secondary peaks between the primary peaks equals $N - 2$.

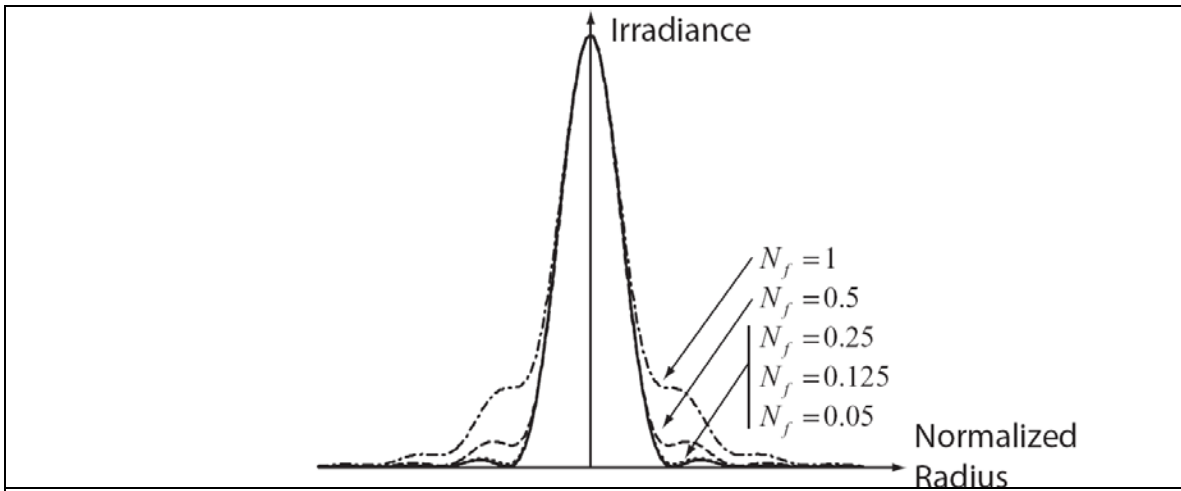


Figure 5.56. Profile of the irradiance diffracted from a circular hole for different Fresnel numbers. The profile is drawn versus radius normalized to the propagation distance.

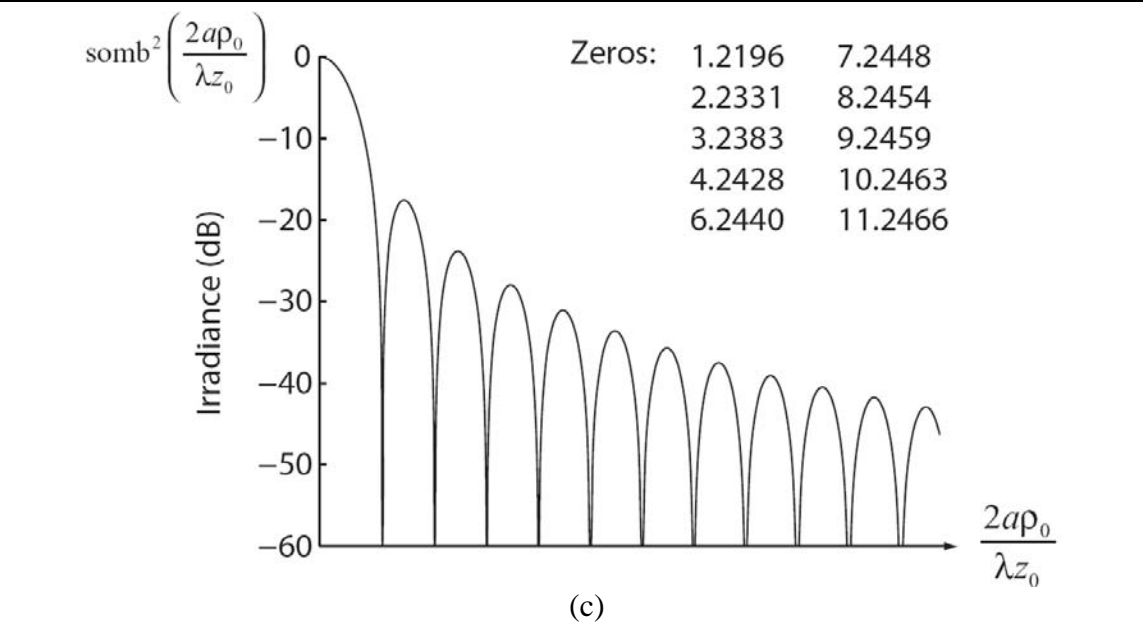
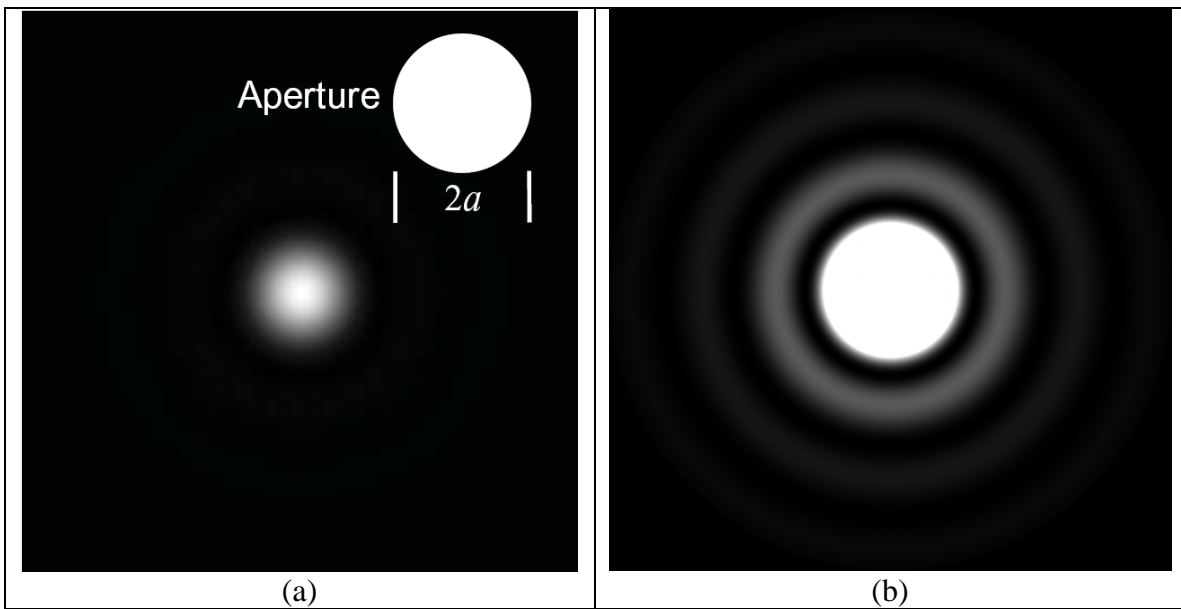


Figure 5.57. The Airy Disk. a) Linear gray-scale irradiance; b) Saturated central peak illustrating ring energy; and c) Irradiance profile of $\text{somb}^2\left(\frac{2a\rho_0}{\lambda z_0}\right)$ in dB $[10\log(I)]$ and table of zeros.

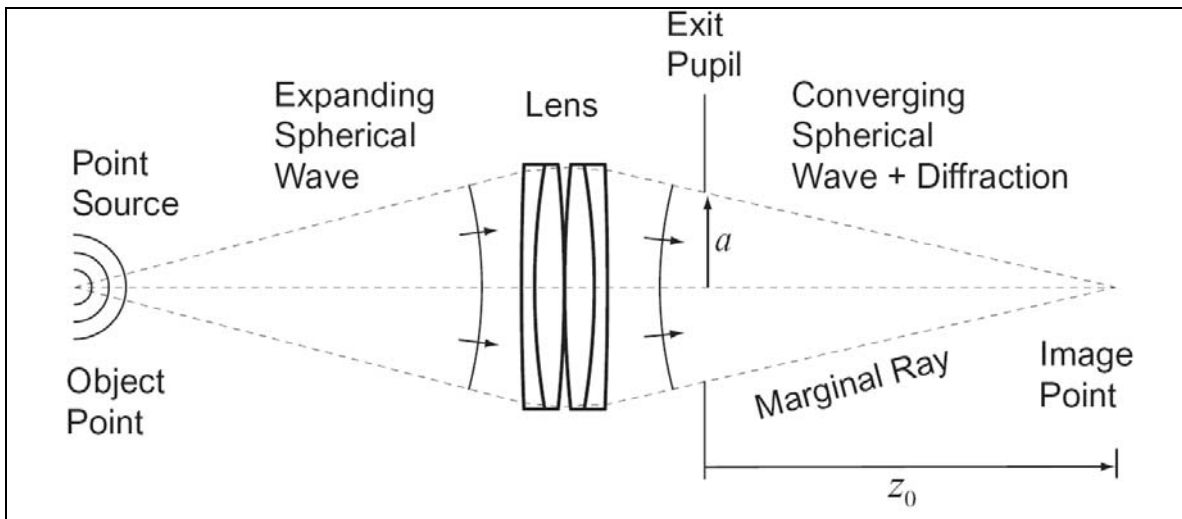


Figure 5.58. Diffraction from the exit pupil of a lens. A lens transforms an expanding spherical wave into a converging spherical wave. The Fraunhofer diffraction pattern of the exit pupil is created at the image plane due to light passing through the exit pupil of the lens.

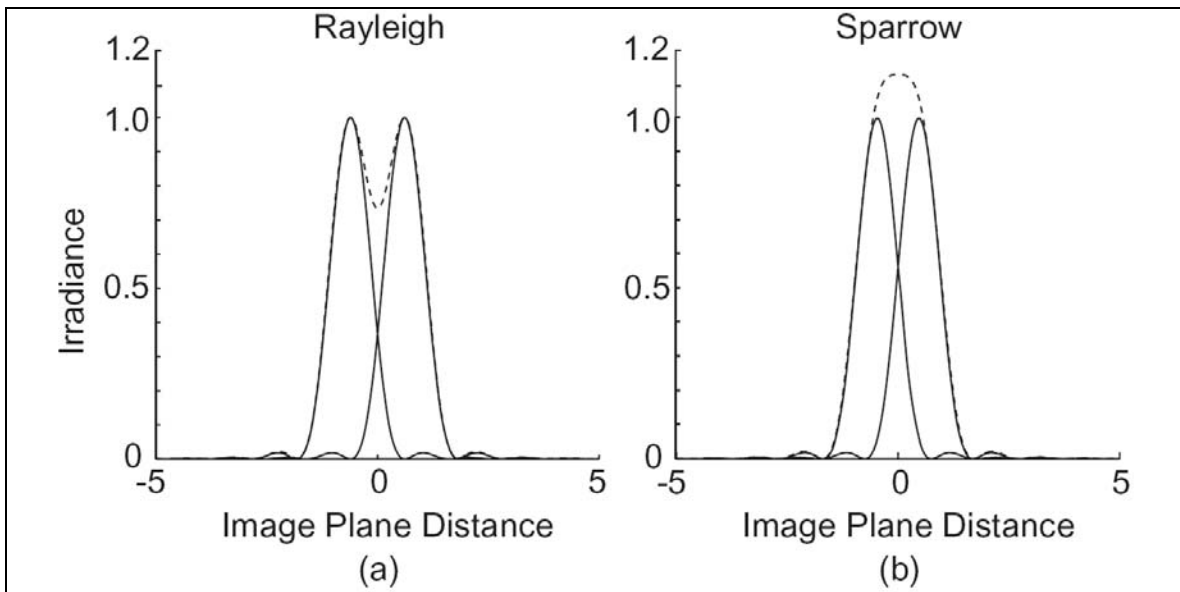


Figure 5.59. Rayleigh and Sparrow resolution criteria. Two incoherent source points are imaged by an optical system. The image-plane irradiance is shown for each Airy spot separately in terms of $\text{somb}^2(r)$, where r is the image plane distance. The dashed lines indicate addition of the irradiance profiles; a) Image points are separated by the Rayleigh criterion, where separation is $0.61\lambda/NA$; and b) Image points are separated by the Sparrow criterion, where separation is $0.47\lambda/NA$.

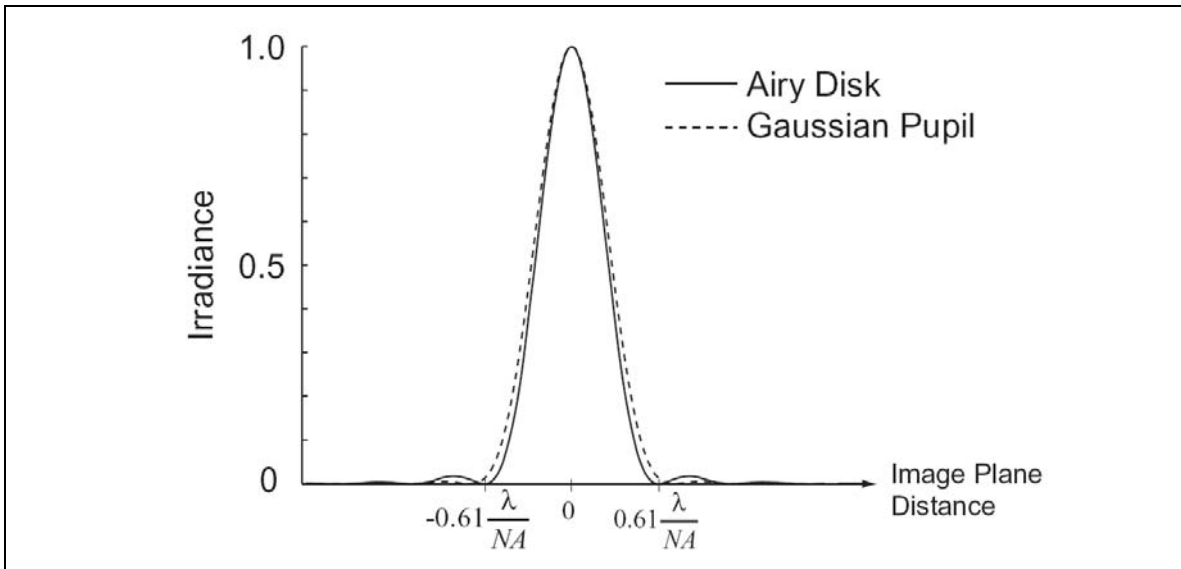


Figure 5.60. Comparison of Airy spots resulting from uniform amplitude and Gaussian amplitude filling of the pupil. For the Gaussian, the $1/e^2$ full-width-at-half maximum coincides with the pupil diameter.

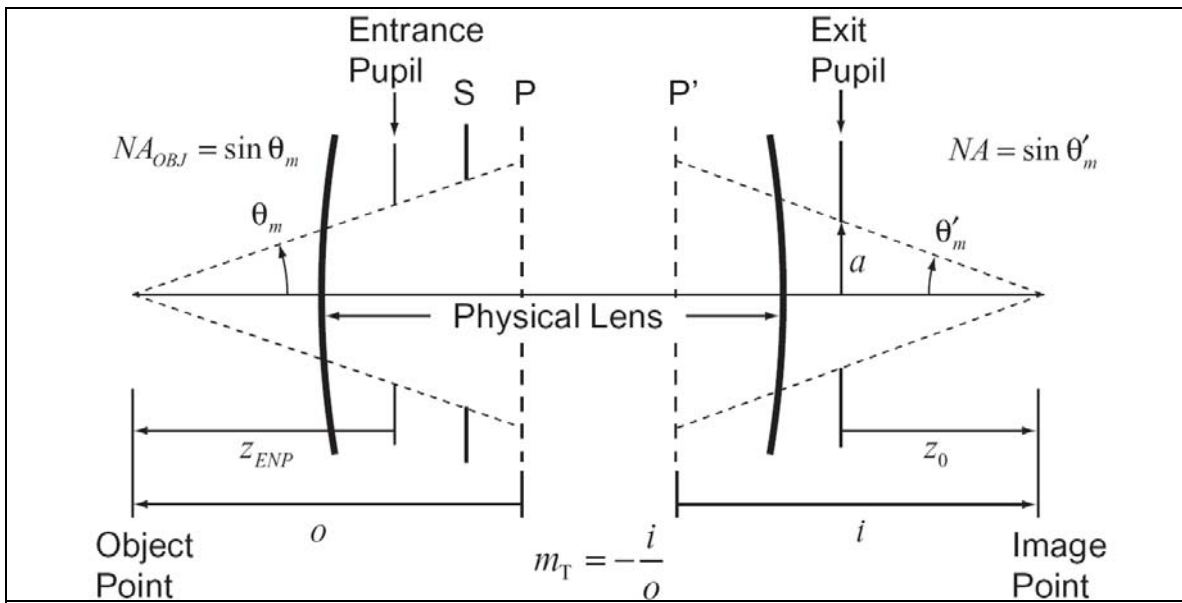


Figure 5.61. First-order description of a general optical system. P is the principle plane in object space, P' is the principle plane in image space, S is the aperture stop, object distance is o , and image distance is i . Transverse magnification of the system is $m_T = -i/o$. Exit pupil radius is a .

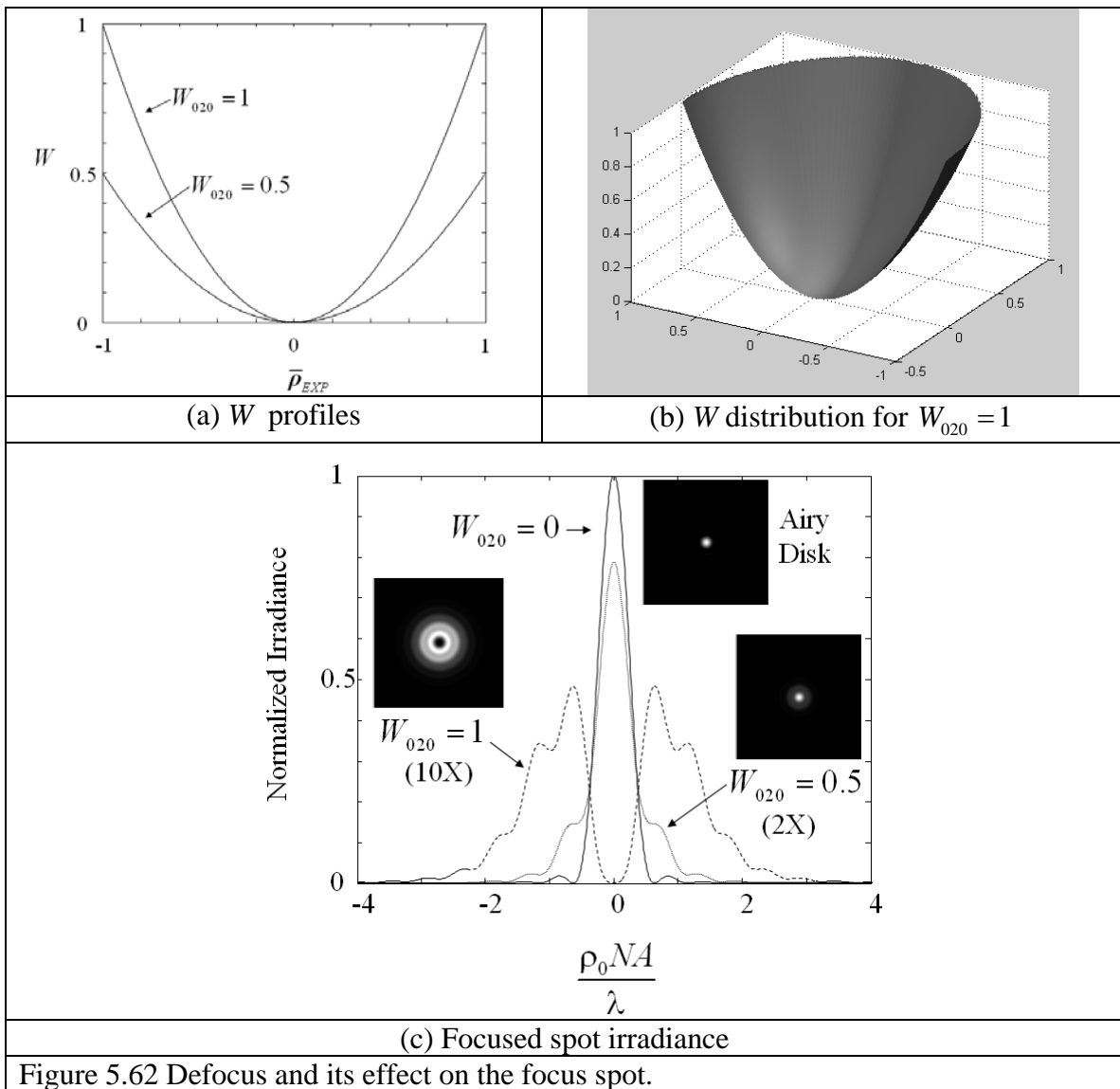


Figure 5.62 Defocus and its effect on the focus spot.

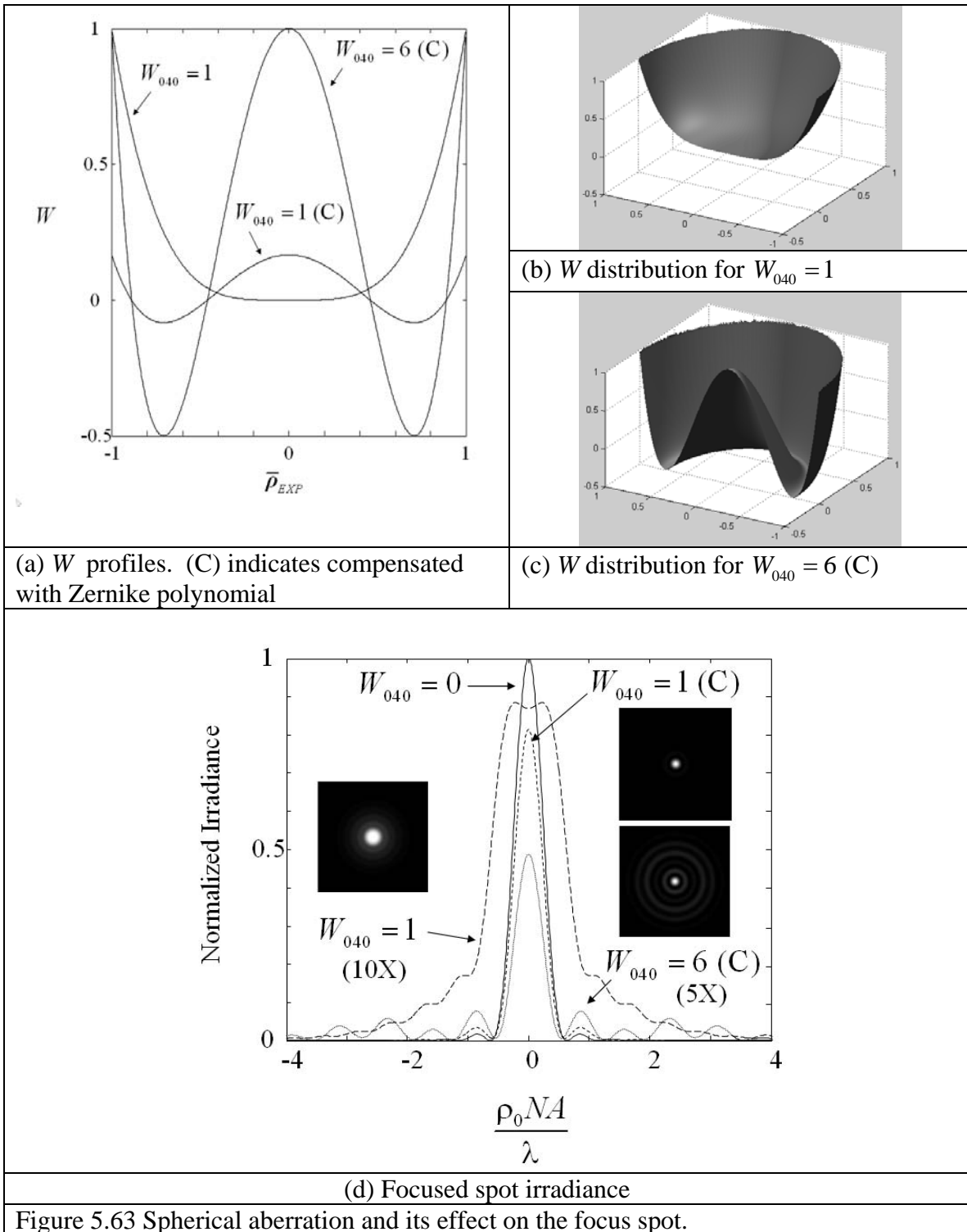


Figure 5.63 Spherical aberration and its effect on the focus spot.

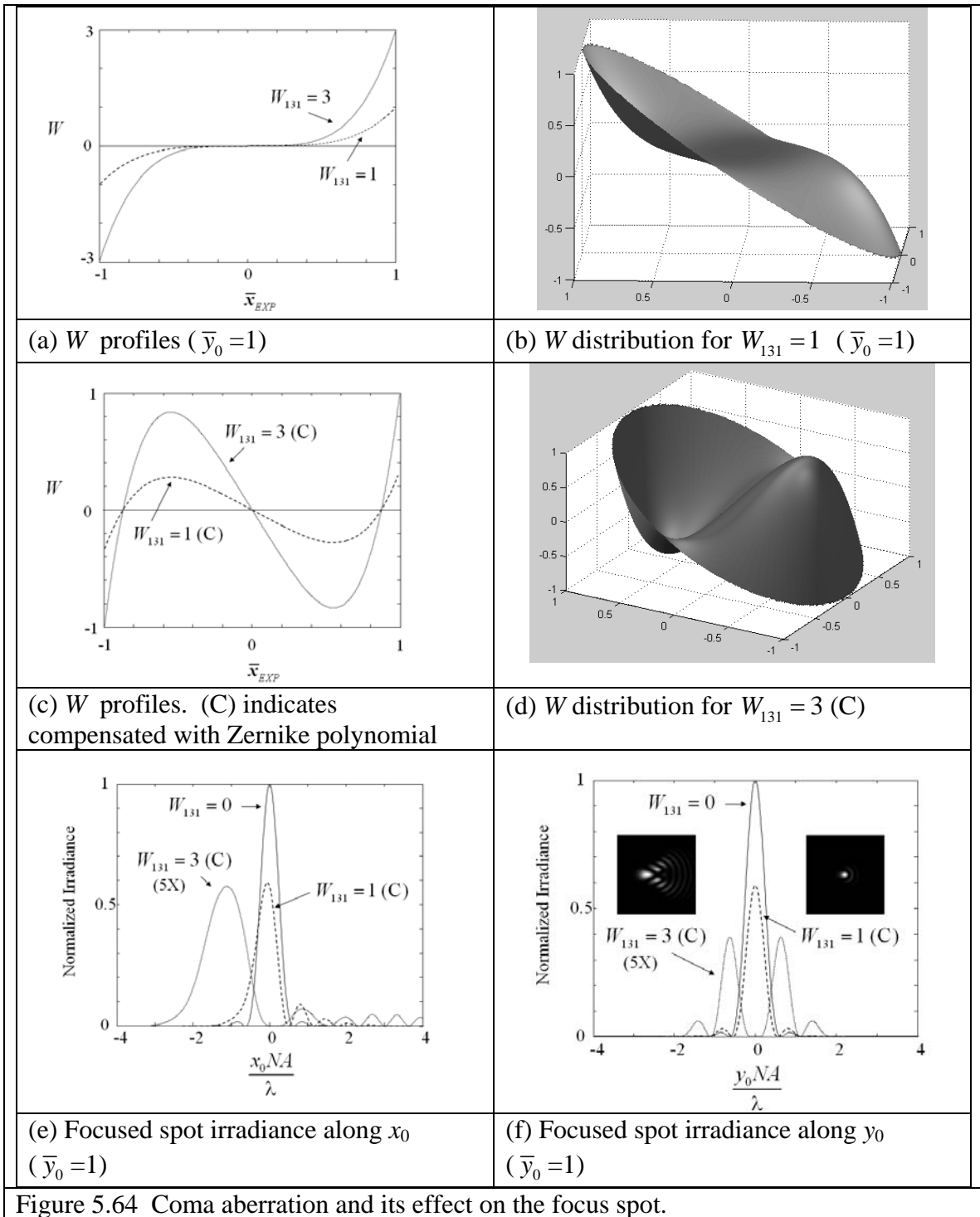
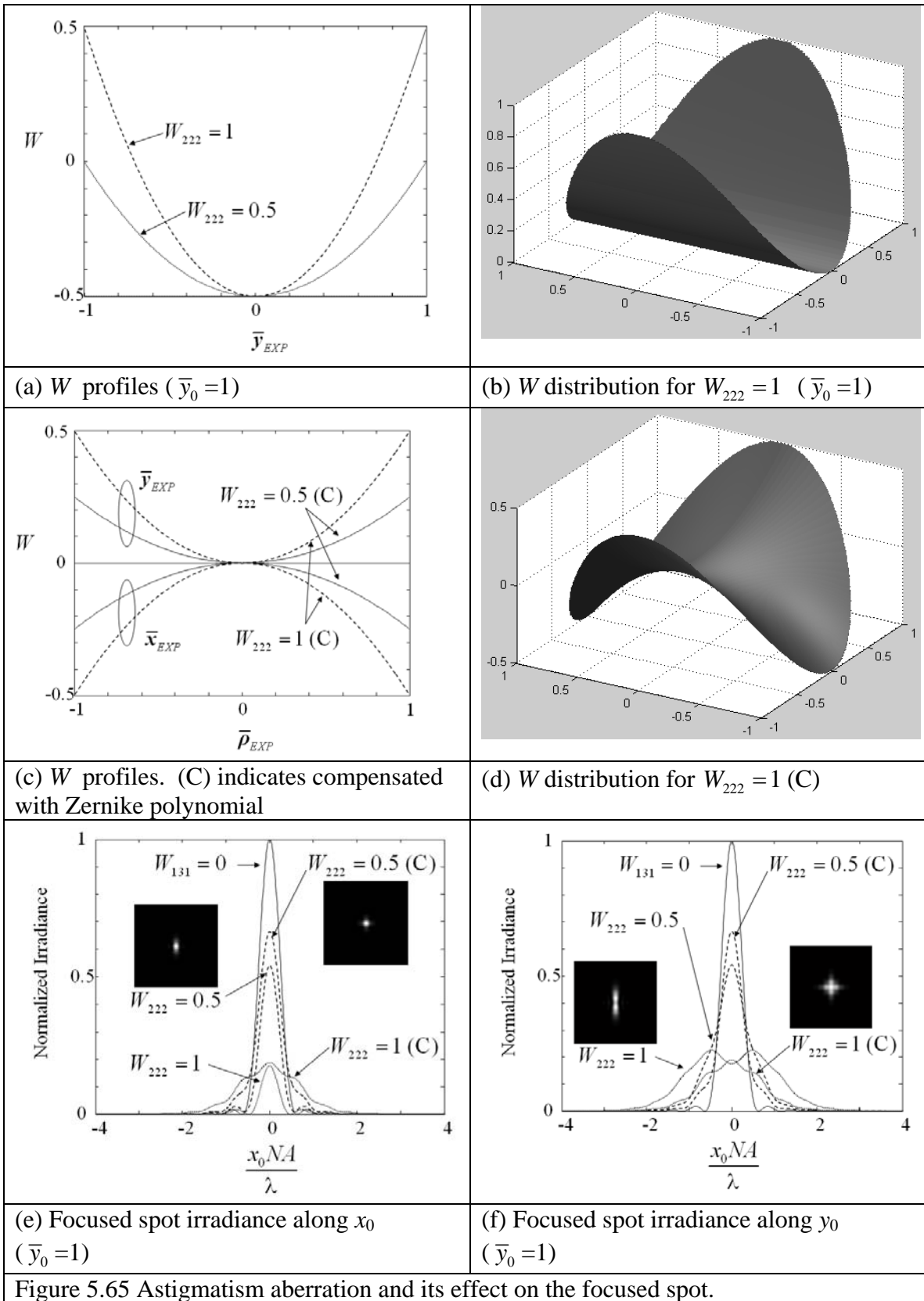


Figure 5.64 Coma aberration and its effect on the focus spot.



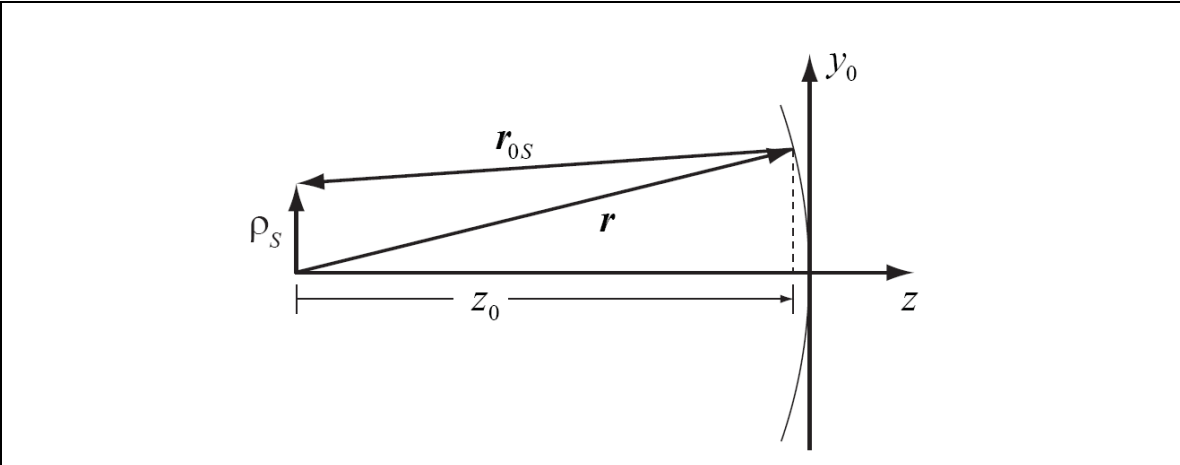


Figure 5.66. Diffraction from the aperture (or an object) to a spherical surface.

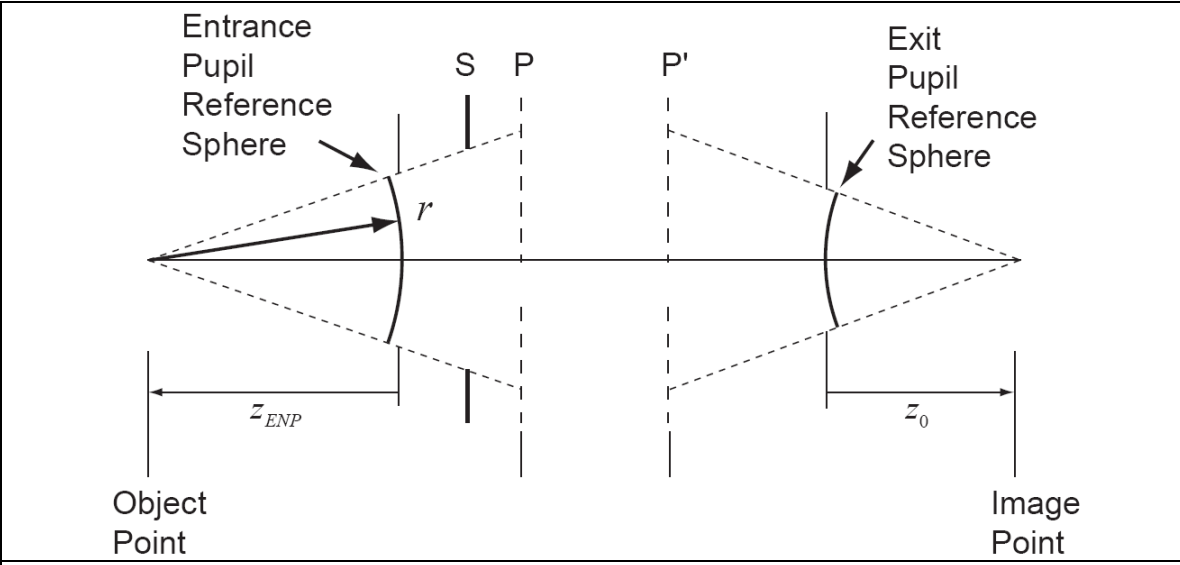


Figure 5.67 Entrance and exit pupil reference spheres.

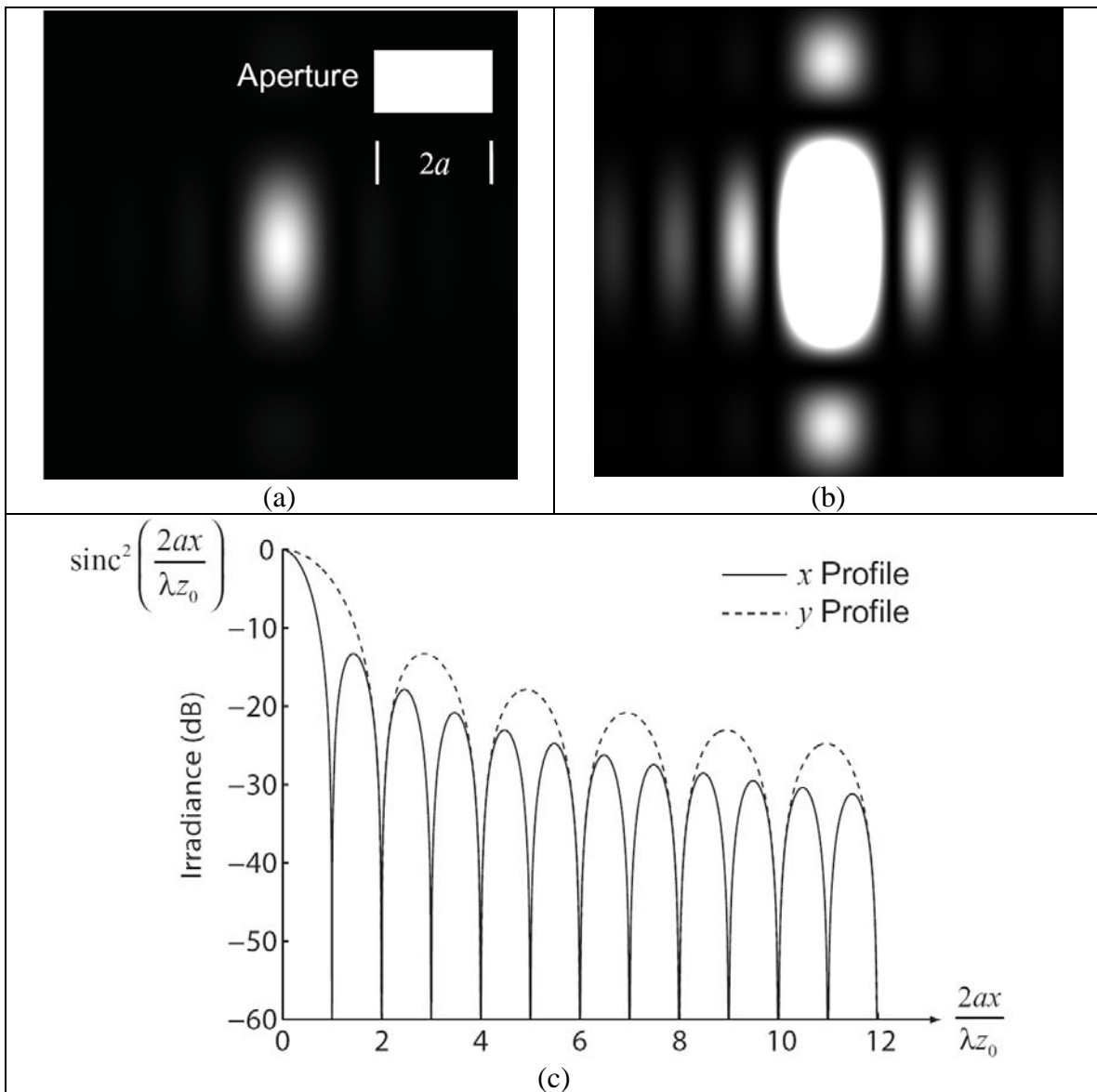


Figure 5.68. Fraunhofer diffraction for a rectangular aperture. a) Linear gray-scale irradiance; b) Saturated central peak illustrating side lobe energy; and c) Irradiance profiles of $\text{sinc}^2\left(\frac{2ax}{\lambda z_0}\right)$ in dB $[10\log(I)]$ and table of zeros. In this case, the y -direction aperture width is $\frac{1}{2}$ the width in the x direction.

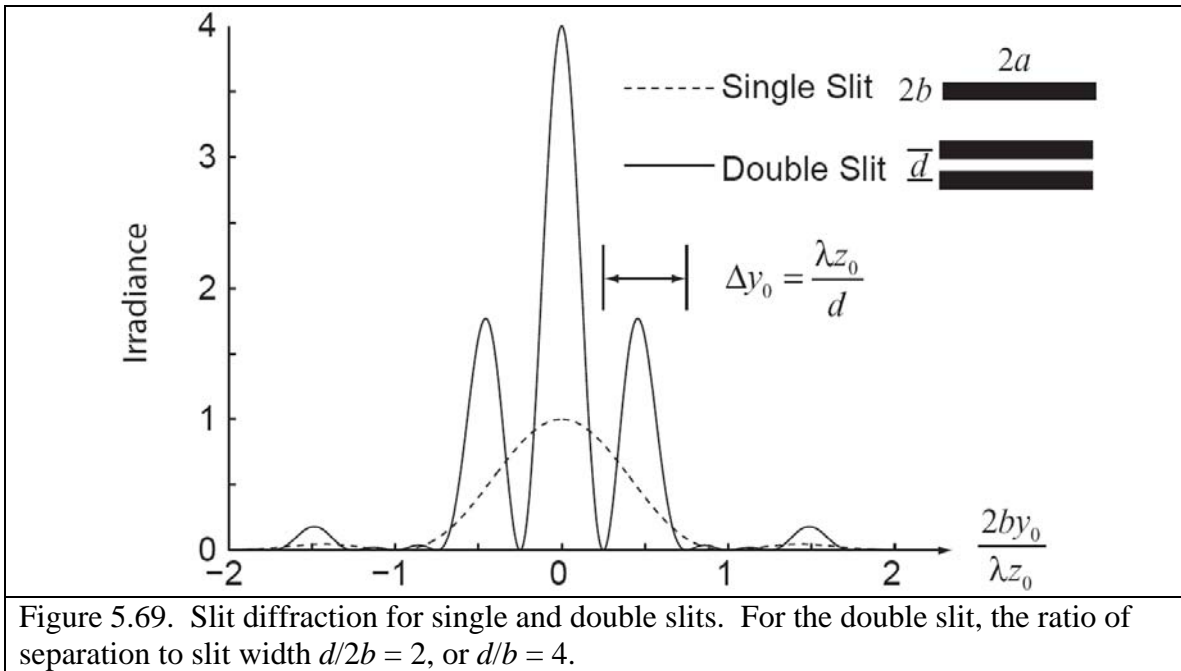


Figure 5.69. Slit diffraction for single and double slits. For the double slit, the ratio of separation to slit width $d/2b = 2$, or $d/b = 4$.

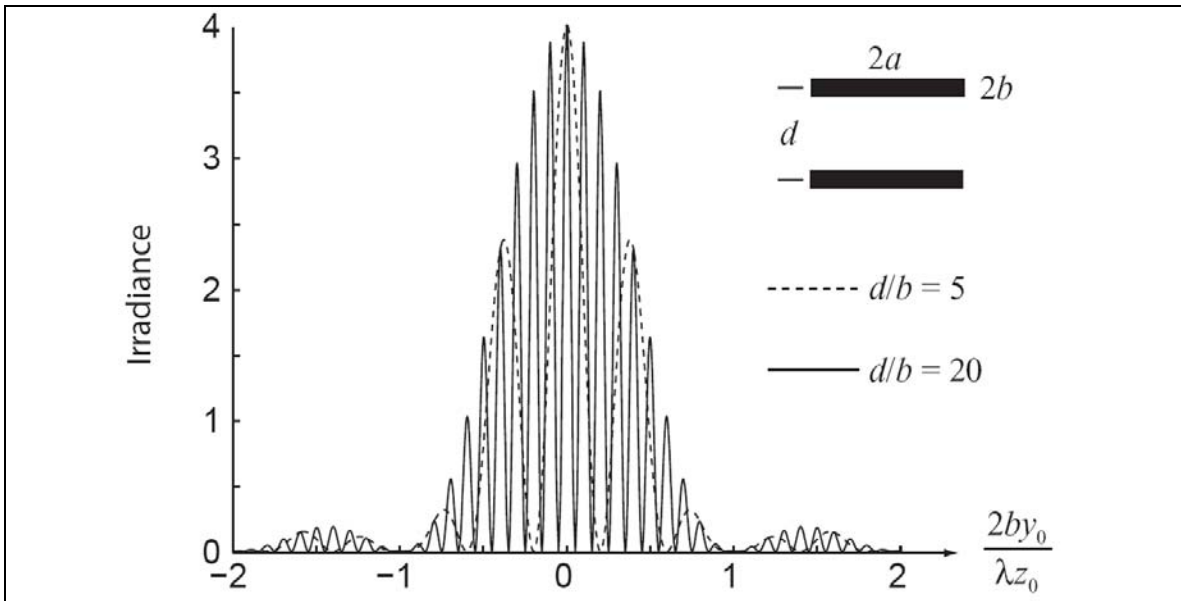


Figure 5.70. Increasing the slit separation. As the slit separation d increases while the width $2b$ remains constant, the number of peaks inside the central lobe increases.

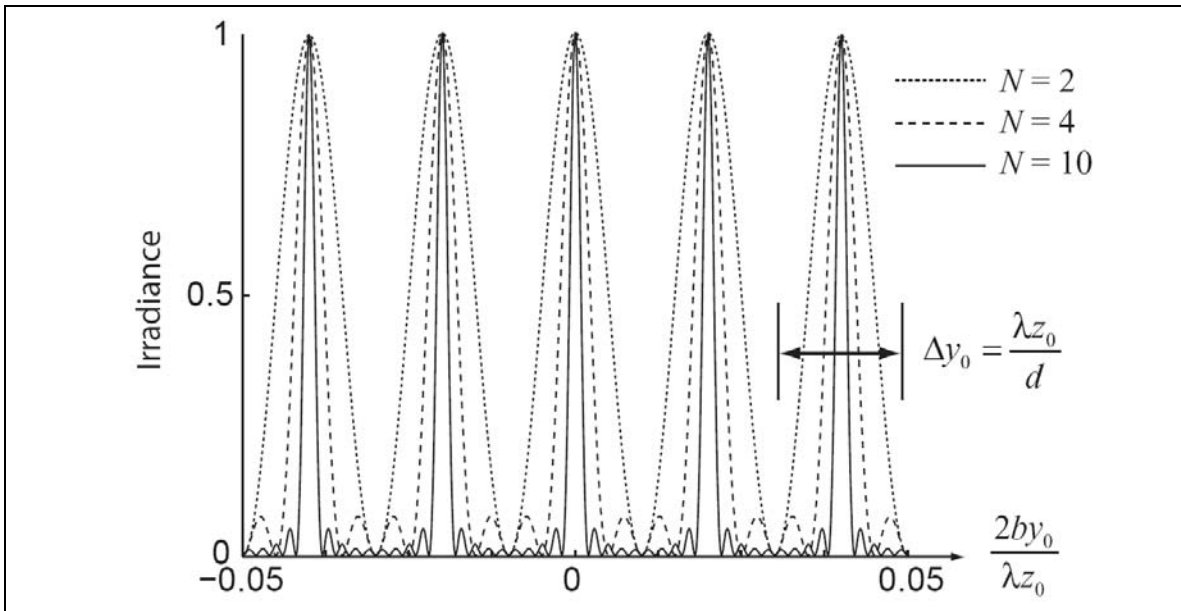


Figure 5.71. Increasing the number of slits N . The number of secondary peaks between the primary peaks equals $N - 2$. In this figure, the peaks are normalized to unity irradiance. In reality, the peak irradiance increases as N^2 . Also, $d/b = 100$.