Chapter 3

Image Evaluation

3.1 Point Spread Function

The diffraction integral, in the Fresnel limit, has the form

$$u_2(x,y) = \frac{e^{jkz_{12}}}{j\lambda z_{12}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u_1(\alpha,\beta) \exp\left[j\frac{\pi}{\lambda z_{12}}((x-\alpha)^2+(y-\beta)^2)\right] d\alpha d\beta. \tag{3.1}$$

based on the geometry shown in Fig. 3.1.

The quadratic phase term can be expanded to give

$$u_2(x,y) = \frac{e^{jkz_{12}}}{j\lambda z_{12}} \exp\left[j\frac{\pi}{\lambda z_{12}}(x^2+y^2)\right] U_1\left(\frac{x}{\lambda z_{12}}, \frac{y}{\lambda z_{12}}\right) \tag{3.2}$$

where

$$U_1(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u_1(\alpha,\beta) \exp\left[j\frac{\pi}{\lambda z_{12}}(\alpha^2+\beta^2)\right] \exp\left[j2\pi(\alpha x+\beta y)\right] d\alpha d\beta \tag{3.3}$$

Usually $u_1(\alpha,\beta)$ will be a converging spherical wave and therefore include a phase term canceling the quadratic phase factor in the integral.

In polar coordinates

$$U_1(r) = 2\pi \int_0^{\infty} u_1(\rho) \exp\left[j\frac{\pi}{\lambda z_{12}}\rho^2\right] J_0(2\pi\rho r)\rho d\rho \tag{3.4}$$

The diffraction integral in polar coordinates can be written using normalized (unit radius) pupil coordinates and normalized image coordinates (units of ϵ_0). The results are

$$u_2(r) = \frac{\pi}{2} \int_0^1 \exp(2\pi j W(\rho)) J_0(2\pi r\rho)\rho d\rho \tag{3.5}$$

where $W(\rho)$ is the wavefront. Then

$$I_2(r) = |\frac{4}{\pi} u_2(r)|^2 \tag{3.6}$$

is the normalized diffraction irradiance. The normalization gives $I_2(0) = 1$ for $W(r) = 0$.

Figure 3.1: Geometry of diffraction (from Gaskill).

Example 1 The rectangular aperture of width a and height b

$$u_1(\alpha,\beta) = A\exp\left[-j\frac{\pi}{\lambda z_{12}}(\alpha^2+\beta^2)\right]\mathrm{rect}\left(\frac{x}{a},\frac{y}{b}\right) \tag{3.7}$$

has diffracted intensity

$$I_2(x,y) = |A|^2\left(\frac{ab}{\lambda z_{12}}\right)^2\mathrm{sinc}^2\left(\frac{ax}{\lambda z_{12}},\frac{by}{\lambda z_{12}}\right), \tag{3.8}$$

Example 2 The circular aperture of diameter d

$$u_1(\rho) = A\exp\left[-j\frac{\pi}{\lambda z_{12}}\rho^2\right]\mathrm{cyl}\left(\frac{\rho}{d}\right) \tag{3.9}$$

has a diffracted intensity profile of

$$I_2(r) = |A|^2\left(\frac{\pi d^2}{4\lambda z_{12}}\right)^2\mathrm{somb}^2\left(\frac{dr}{\lambda z_{12}}\right). \tag{3.10}$$

Fig. 3.2 shows a perspective plot of the diffraction pattern for a circular aperture.

Example 3 The effects of aberration on the diffraction pattern of a point object illuminating a circular aperture can be demonstrated by considering a simple defocus wavefront. The aberration function $W(\rho)$ for defocus is given by

$$W(\rho) = a_z\rho^2$$

Then Eq. 3.5 and Eq. 3.6 are evaluated numerically and the relative irradiance is plotted as a function of image radius (in units of ϵ_o). Different cases are obtained by changing the

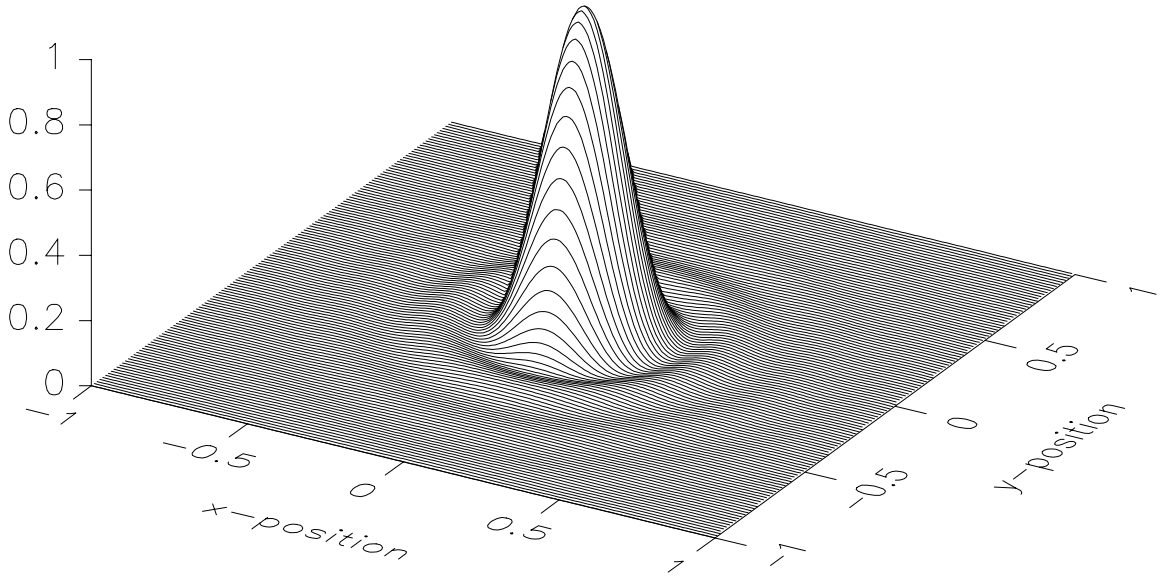


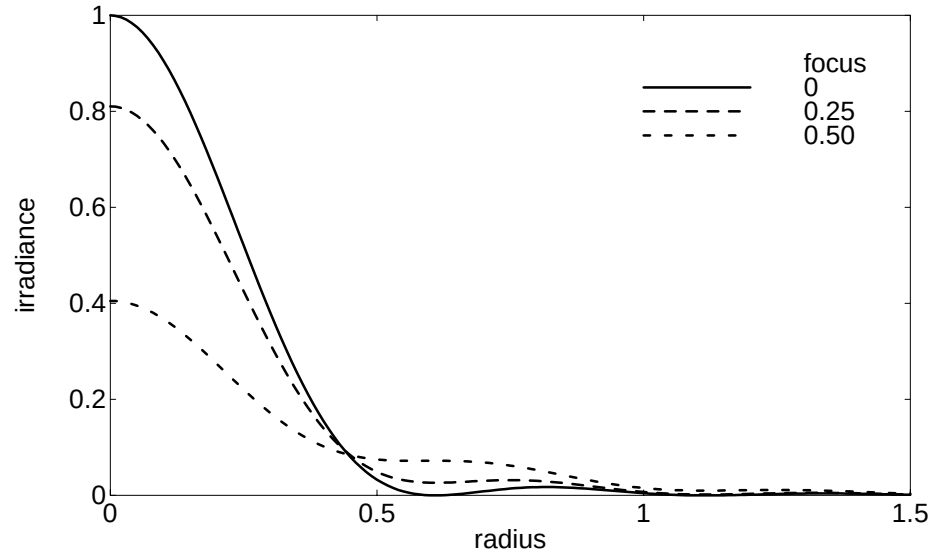
Figure 3.2: Diffraction of circular aperture.

value of a_z. Fig 3.3(a) shows the results for small focus shifts. The zero focus case is just the Airy diffraction pattern. It provides a check on the scaling parameters since the peak irradiance should be unity and the first zero in irradiance should appear at $r = 0.61\epsilon_o$.

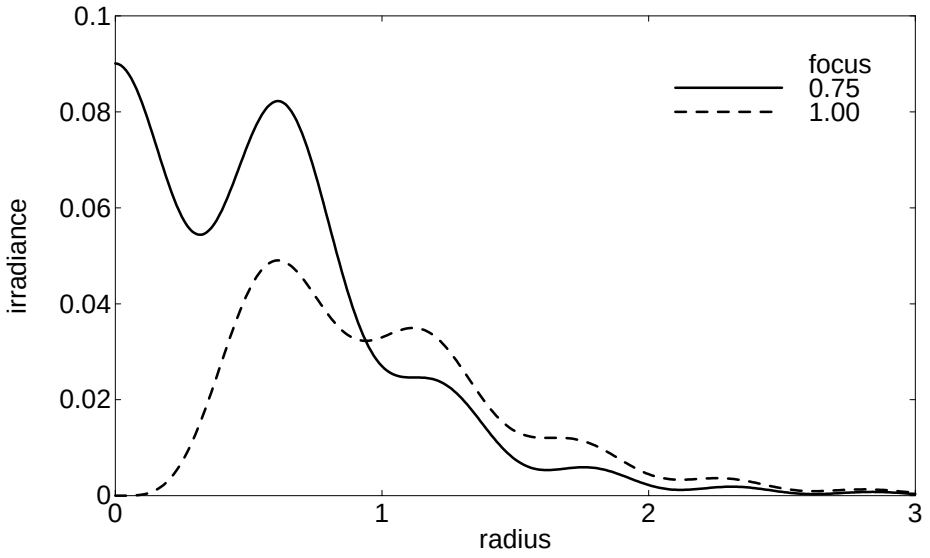
The quarterwave defocus example corresponds to the Raleigh criterion for a diffraction-limited image. It also matches the Maréchal criterion of a Strehl ratio of at least 0.8. For small focal shifts, and small aberration values in general, the effect on the image is a decrease in the magnitude of the central diffraction peak and an increase of the magnitude of the secondary ring structures. One could conclude that any point spread function in which the central peak is a dominant structure is diffraction-limited.

As the coeffcient for focal shift increases, the central peak becomes less prominent, and eventually is superceded by the ring structure. This is shown in Fig. 3.3(b). Note that the vertical scale is expanded by a factor of ten to show the detail better. For focal shifts given by integer values of a_z, there is a zero in irradiance on axis. This phenomenon was important historically in demonstrating the wave nature of light.

Example 4 Studying the effects of general aberrations on the diffraction point function requires the repeated evaluation of the Fourier transform integral in Eq. 3.3. The Fast Fourier Transform (FFT) algorithm can be used for this purpose. As an example, we sampled the amplitude function for a circular pupil with five waves of defocus, using a 32 x 32 sample grid. This grid was inserted into a complex array with 256 x 256 elements which had previously been filled with zero values. The FFT of this complex array was taken. Then each element



(a) small values of focal shift.



(b) larger values of focal shift (expanded vertical scale).

Figure 3.3: Radial profiles of the diffraction of circular aperture with varying focus shift (given in waves).

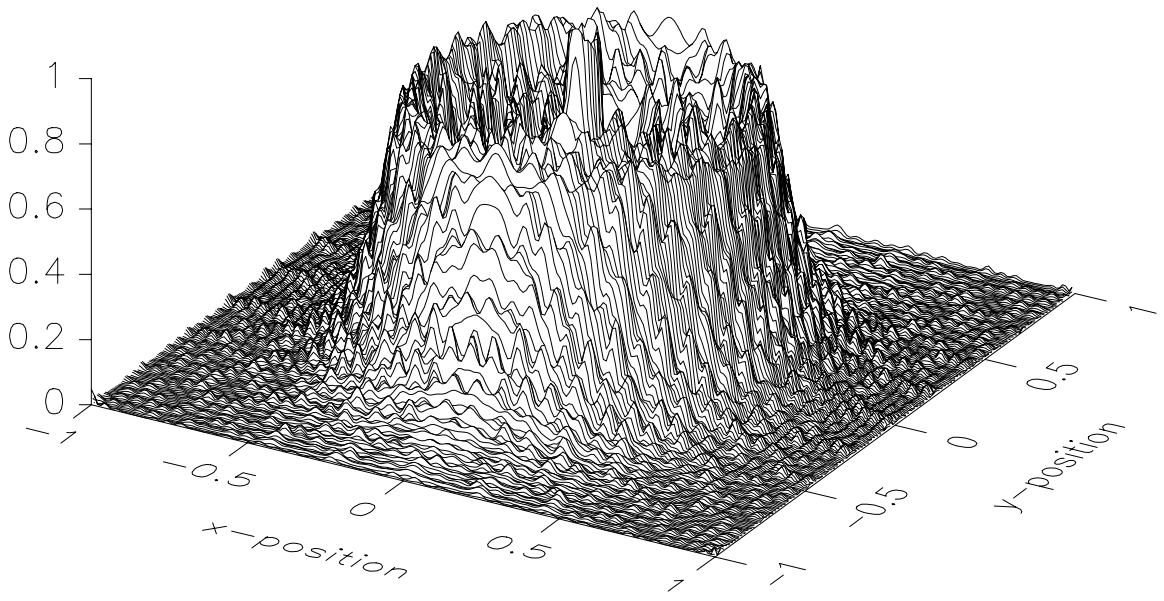


Figure 3.4: Diffraction of circular aperture with 5 waves of defocus.

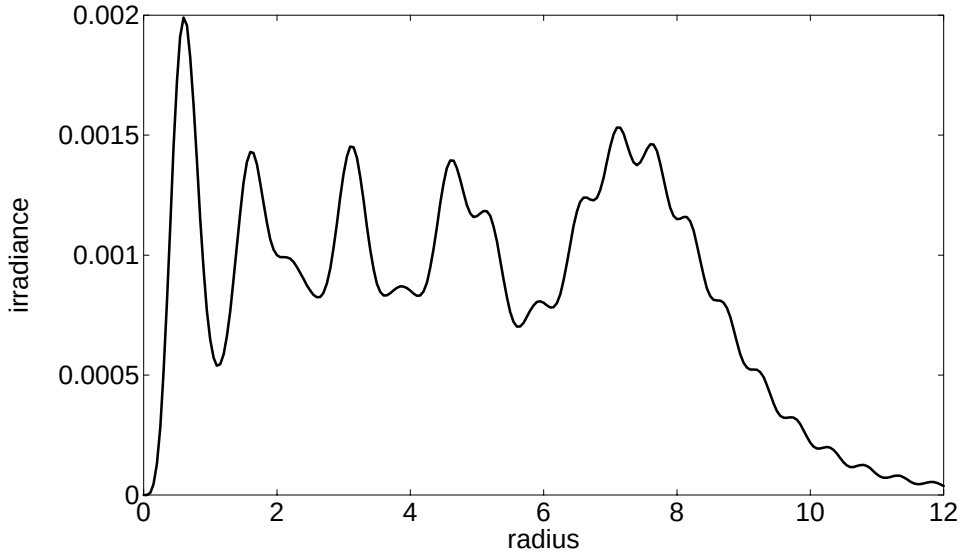
was multiplied by its complex conjugate. The result is a sampled representation of the diffraction point spread function.

Fig. 3.4 shows a perspective plot of the resulting diffraction pattern. The peak irradiance (0.002) has been normalized to unity. Fig. 3.5(a) shows a radial profile calculated using numerical integration of Eq. 3.5. Fig. 3.5(b) shows a cross-section of the image array calculated using the FFT. Note that the gross features in the image-plane array match those obtained using direct integration, but that the smaller features are only approximate matches. Note further the variation in irradiance along the top of the outer ring in Fig. 3.4. These are sampling artifacts due to the discrete nature of the pupil grid.

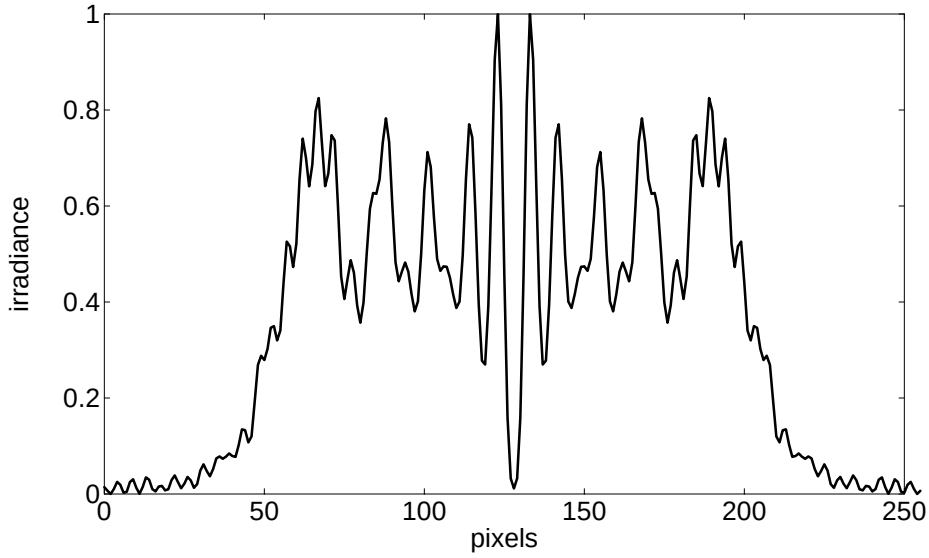
3.2 Energy Distribution Functions

Energy distribution functions determine the total energy from a point source, passing through the optical system, and collected according to specific geometrical constraints. A simple energy distribution measurement, for example, might determine the fraction of total available energy falling on a rectangular detector element in the image plane. Optical system specifications may also be estabished related to energy distribution. For example, an optical system could be required deliver 80% of the available energy into a circular area whose radius is twice the Airy radius. Such a system, of course, would be nearly diffraction limited in performance.

The energy distribution functions most commonly used are the radial energy distribution



(a) radial profile calculated from numerical integration.



(b) cross-section of image array calculated from the FFT.

Figure 3.5: Radial profiles of the diffraction of circular aperture with focus shift of five waves.

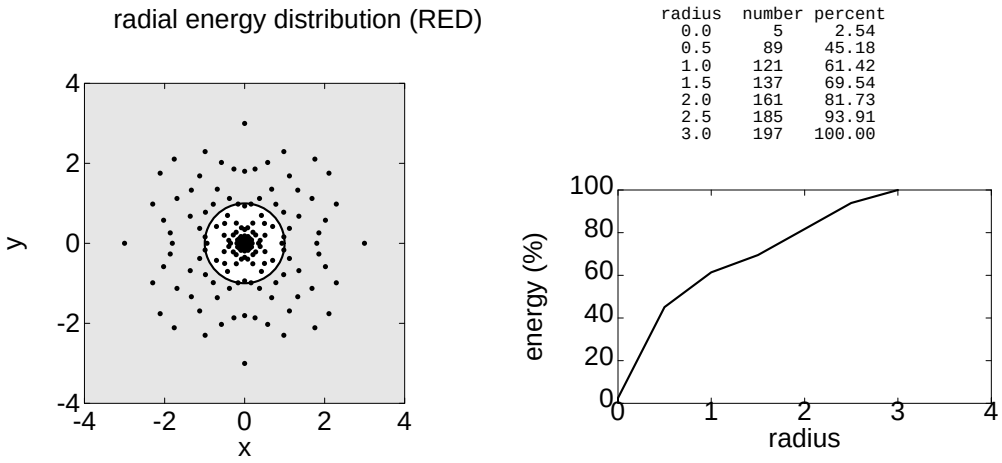


Figure 3.6: Construction of the radial energy distribution function.

function (RED) and the knife edge distribution function (KED). The RED function, shown in Fig. 3.6, measures the fraction of total energy in the point spread function that lies within the radius r. The RED function is sometimes called the encircled energy distribution function. The KED function, shown in Fig. 3.7, measures the fraction of total energy that passes by a knife edge oriented in a given direction as a function of its transverse position. Both functions are projections, reducing the full two-dimensional point spread function to a one-dimensional energy distribution function. The RED averages over the angular coordinate, whereas the KED averages linearly in a direction parallel to the knife edge.

The radial energy distribution function is calculated as

$$E_r(r) = (1/E_m) \int_0^{2\pi} \int_0^r h(\rho, \theta) \rho d\rho d\theta \tag{3.11}$$

where E_m is the total energy and $h(\rho, \theta)$ is the point spread function in image plane coordinates (ρ, θ). It is most relevant in systems for which the point spread function is circularly symmetric. Otherwise there can be considerable uncertainty on the appropriate location for the center of the coordinate system.

Fig. 3.6 shows a spot diagram for 1 wave of fourth-order spherical aberration and -0.5 wave focal shift. Each spot represents the same fraction of the total energy passing through the exit pupil. The table shows the number of spots inside a circle of the specified radius. The corresponding RED plot is shown below the data table.

The knife edge distribution is produced by first rotating the image plane coordinate system from (x, y) to (s, t) so that the s direction is transverse or perpendicular to the knife edge, and t is parallel to the knife edge. Then the distribution function itself is calculated from

$$E_k(s) = (1/E_m) \int_{-\infty}^{\infty} \int_s^{\infty} h(s, t) ds dt \tag{3.12}$$

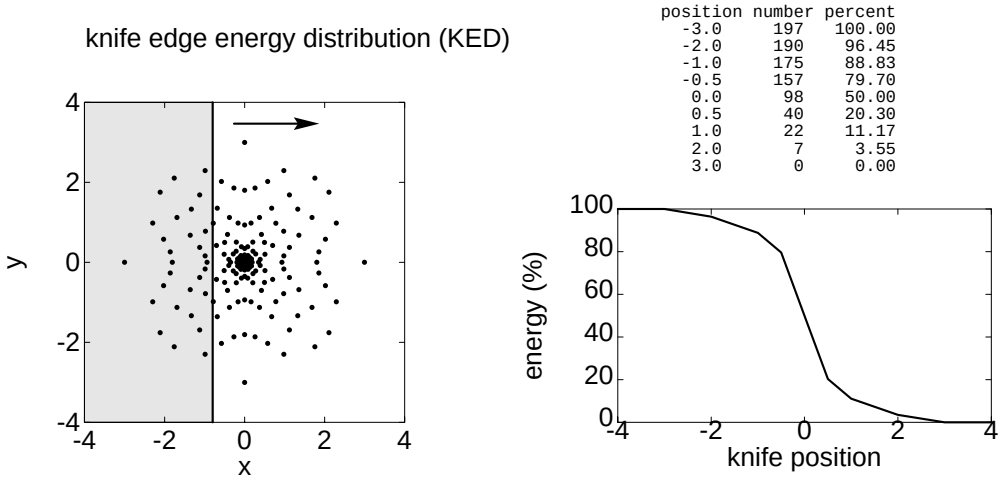


Figure 3.7: Construction of the knife edge energy distribution function.

where $h(s,t)$ is the point spread function in rotated image plane coordinates.

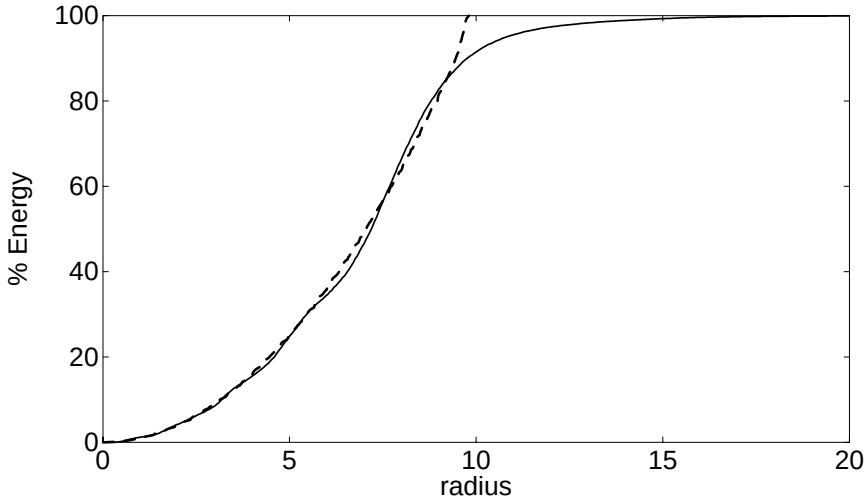
Fig. 3.7 shows a spot diagram with a knife edge traveling along the x-axis, as shown by the arrow. The table shows the number of spots to the right of the knife edge as a function of its position. The corresponding KED plot is shown below the data table.

Energy distribution diagrams for diffraction spread functions are obtained much like those derived from geometrical spot diagrams. Instead of counting spots within a given radius, for example, we numerically integrate the irradiance within the given radius. If the point spread function is not diffraction-limited, there is a good correspondence between an energy distribution diagram obtained from the geometrical spot diagram and one obtained from the diffraction spread function. This is illustrated in Fig. 3.8, which shows RED and KED functions for a circular pupil with 5 waves of focal shift. The solid curve was derived from the point spread function shown in Fig. 3.2. The dashed curve was derived from the geometrical spot diagram.

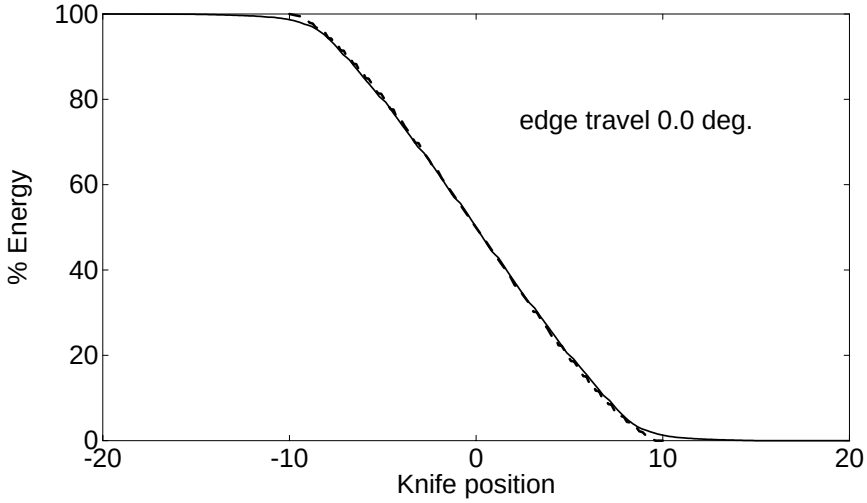
Example 1 The radial energy distribution of the Airy pattern of a circular aperture is given by

$$E_r(r) = 1 - j_0(2\pi r)^2 - j_1(2\pi r)^2 \tag{3.13}$$

where the radius is in units of ϵ_o. This function is shown in Fig 3.9. A diffraction-limited spot has a signficant fraction of energy (16.2%) located outside of the central peak. For the dark rings, shown by the dashed lines in Fig. 3.9, $j_1(2\pi r) = 0$. The total energy outside of a dark ring is given by $j_0(2\pi r)^2$.



(a) radial energy distribution function.



(b) knife edge energy distribution function.

Figure 3.8: Energy distribution plots for a circular pupil with 5 waves of focal shift. The solid curves were obtained from the diffraction point function, and the dashed curves from the geometrical spot diagram.

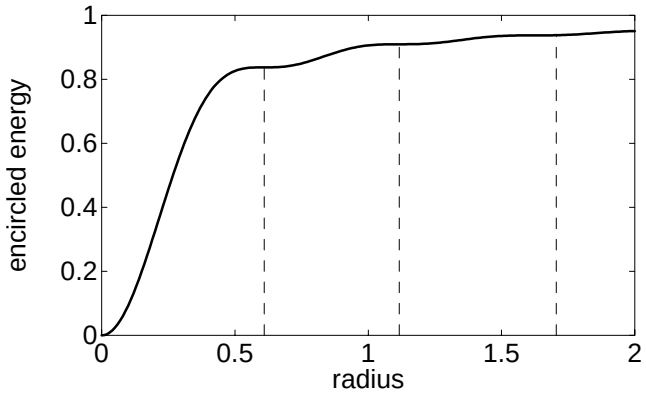


Figure 3.9: Radial energy distribution for perfect optical system. The radius is in units of ϵ_o. The dashed lines show the location of the dark Airy rings.

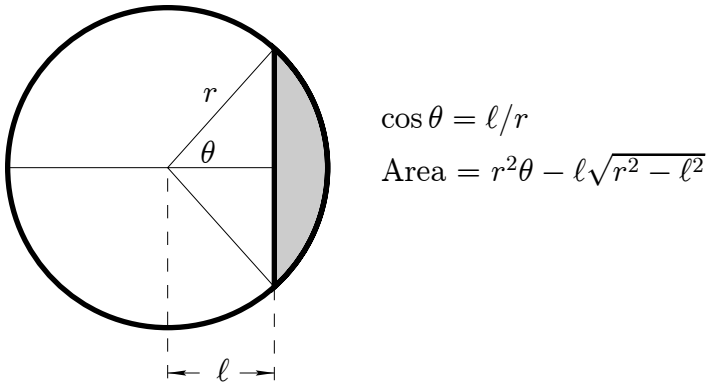


Figure 3.10: Area of segment of circle.

Example 2 The knife-edge energy distribution of a cylindrical point spread function of unit radius is given by

$$E_k(x) = \left[\cos^{-1}(x) - x(1.0 - x^2)^{1/2}\right]/\pi \tag{3.14}$$

where x is the transverse position of the knife edge relative to the center of the cylinder. Eq. 3.14 is derived from the formula for the area of a segment of a circle, as shown in Fig. 3.10. A plot of the KED function for a cylinder is shown in Fig. 3.11.

3.3 Edge and Line Response Functions

A sharp edge placed in the object plane of a diffraction-limited optical system will produce an edge image blurred by diffraction. The width of the edge will be approximately an Airy

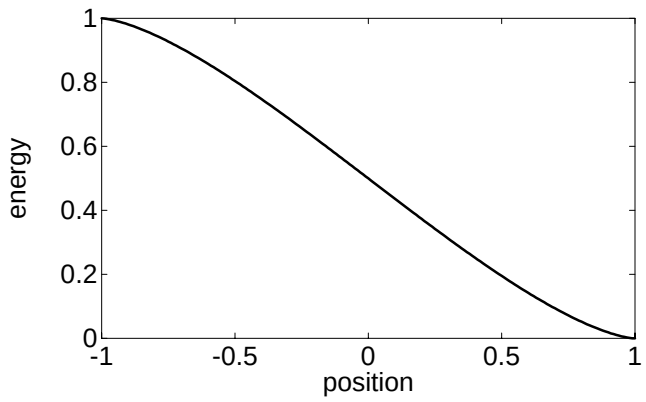


Figure 3.11: Knife-edge energy distribution for cylindrical point spread function.

diameter. The image irradiance profile associated with the image of the edge is called the *edge response function*.

To say that an ideal optical system is a linear, shift-invariant system means that

1. Any point in the object field maps into a corresponding image point without distortion.
2. An object point produces the same energy distribution (point spread function) independent of its location in the object field.
3. The irradiance at any point in the image is the linear superposition of irradiances from the spread functions of every point of an extended source.

In such a system, the image is obtained as

$$E(x,y) = \int\int E_i(\alpha,\beta)h(x-\alpha, y-\beta)d\alpha d\beta \tag{3.15}$$

where $h(x,y)$ is the point spread function and $E_i(x,y)$ is the ideal image irradiance obtained by mapping object space into image space. Generally this mapping implies a change in scale or *magnification*. The range of integration extends over the dummy image space coordinates (α, β). Eq. 3.15 says that the actual image is the convolution of the ideal image with the system point spread function,

$$E(x,y) = E_i(x,y) * h(x,y). \tag{3.16}$$

Similarly, a thin slit or line placed in the object plane of a diffraction-limited optical system will produce a line image blurred by diffraction. The line width will be approximately an Airy diameter. The image irradiance profile associated with the image of the line is called the *line response function*.

The edge response function and line response function are closely related. An ideal edge is given by the step function, step(x), which describes a vertical edge located at the origin. An ideal line is the derivative of the step function,

$$\text{line}(x) = \frac{d}{dx}\text{step}(x) = \delta(x) \tag{3.17}$$

where $\delta(x)$ is the two-dimensional line delta-function. The line response function is the integral of the point spread function,

$$\ell(x) = \int h(x, y)dy, \tag{3.18}$$

along a direction parallel to the line itself.

The edge and lines response functions for an ideal, diffraction-limited system are shown in Fig. 3.12. Both functions were evaluated numerically using the Airy function as the point spread function. The line spread function is normalized to a central value of unity.

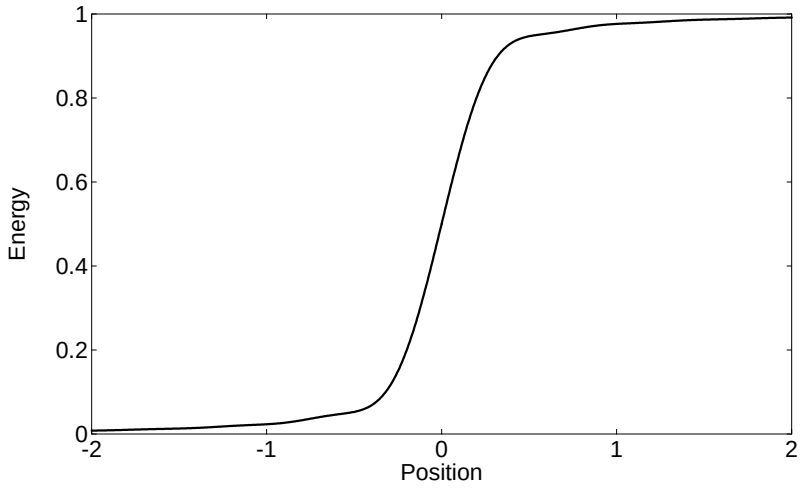
There is a close association between the edge response function and the knife-edge energy distribution (KED) function. The edge response function is determined by placing an edge in the object plane and scanning a point detector perpendicular to the edge. It could also be determined by keeping the detector fixed, and scanning the edge in the object plane. The KED function is determined by using a point object and scanning an edge in the image plane. For a strictly linear shift-invariant system, all three measurements should produce the same response curve. In a real optical system, each measurement would yield a different result, because the point spread functions involved in each measurement are generally different. Thus the image of a vertical edge, centered horizontally in the object field, will have a different profile near the top of the field than at the middle of the field. In many optical systems, however, there will be regions of the object field over which the point spread function is essentially the same. Such a region is called an *isoplanatic patch.*

3.4 Isoplantism and the Sine Condition

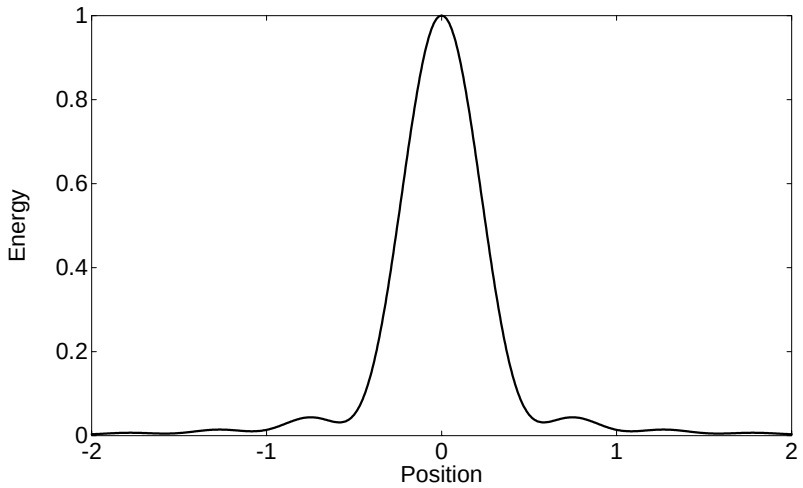
An optical system is said to be isoplanatic if a displacement of the object point produces no change in the aberrations of the corresponding shifted image point. This concept normally applies to a limited region of the field, perhaps even a differential area. An *aplanatic* optical system is also free from the primary optical aberrations. The term aplanatism was used by writers in the 1800's to refer to the abscence of spherical aberration. In present day usage, it usually means freedom from spherical aberration, linear coma, and primary astigmatism.

The axial symmetry in an optical system makes it possible to determine the departure from isoplanatism near the axis by very simple calculations involving only axial rays. Consider two axial points P and P' which are conjugates and two associated meridional rays with angles U and U'. Let P be displaced a small distance y_s in the sagital direction. The *optical sine theorem* then states

$$ny_s \sin U = n'y'_s \sin U' \tag{3.19}$$



(a) Edge response function



(b) Line response function

Figure 3.12: Edge and line response of diffraction-limited optical system.

This is the skew equivalent of the Lagrange invariant. It gives a magnification for small distances imaged by rays at large angles, whereas the Lagrange invariant is for large distances with rays at small angles in the meridional plane.

We have already seen that coma is the result of a change in magnification as a function of aperture. Spherical aberration is the result of a focal shift as a function of aperture. If the spherical aberration at the edge of the aperture is zero, then the coma will also be zero if

$$\frac{\sin U}{u} = \frac{\sin U'}{u'}. \tag{3.20}$$

This equation is the simplest form of the *Sine Condition*. Early lens designers made extensive use of a dimensionless ratio called the *offense against the sine condition* (OSC), calculated from

$$\text{OSC} = \frac{\sin U}{u}\frac{u'}{\sin U'}\left(\frac{R}{R-\delta\ell}\right) - 1, \tag{3.21}$$

where R is the radius of the reference sphere and $\delta\ell$ is the longitudinal spherical aberration at the finite aperture angle U. Isoplanatism is destroyed by significant OSC residuals.

A more complete discussion of the sine condition, isoplanatism and related topics, especially in non-symmetric systems, may be found in Chapter 8 of Welford [2].

3.5 Resolving Power

The resolving power of an optical system indicates how close two objects can be and still be recognized as separate objects. The resolving power of a grating, for example, is the minimum separation necessary to distinguish two narrowly space spectral lines. Similarly, the resolving power of a telescope is the minimum angular separation needed to distinguish two stars. Fig. 3.13 shows a cluster of point objects, including some which are just resolved and some which are not. Generally the concept of resolving power is applied to situations where the objects are self-luminous and mutually incoherent, that is, such that there are no interference effects (fringes) between the objects. Then the irradiance distribution in the combined image is the sum of the irradiance patterns of the individual objects.

Under ideal conditions, where the diffraction patterns are symmetric and noise-free, even a small separation of two points may be distinguished by the slight asymmetry imparted to the combined pattern. In practice, however, several criteria are commonly used to define resolving power. In Fig. 3.14, we see three possible situations, corresponding to separations of $0.4\epsilon_o$, $0.5\epsilon_o$, and $0.6\epsilon_o$. The first case is not resolved because there is no evidence of two peaks. The second case is barely resolved, and the third case is adequately resolved.

One criterion, attributed to Lord Rayleigh, is that the central diffraction maximum of one image should be no closer than the first diffraction pattern zero of the second image. Rayleigh's criterion applied to point images demands that the images be separated by $0.61\epsilon_o$. Rayleigh's criterion applied to two uniformly illuminated disks, as shown in Fig. 3.15, means that the disks must be separated by one radius ($0.5\epsilon_o$ if the disk diameter is ϵ_o).

Figure 3.13: Image showing a cluster of point objects, some of which are just resolved.

Another criterion, less conservative than Rayleigh's criterion, is Sparrow's criterion which states that two image points can just be resolved if the irradiance is a constant along a line separating the image centers. Mathematically, Sparrow's criterion means that the second derivative of the irradiance be zero at the midpoint. For a circular aperture, Sparrow's criterion becomes a separation of $0.475\epsilon_o$. The case of $0.5\epsilon_o$ shown in Fig. 3.14 is just larger than Sparrow's criterion.

Fig 3.16 shows the same three conditons plotted as cross-sections in Fig 3.14.

3.6 Resolution Bar Targets

An optical systems is typically judged by its resolving power, or ability to reproduce fine detail in an image. The size of the finest detail which can be resolved is specified by the *limiting resolution*.

Three-bar resolution charts represent the most common form of target used for testing limiting resolution. The basic pattern, shown in Fig. 3.17, consists of three bars. The bars are each 1 unit wide and 5 units long and are placed 1 unit apart, so that the pattern is 5 x 5 units square. Each spatial frequency is represented by a set of three horizontal and three vertical bars. The sets are usually arranged in groups of six (shown as columns in Fig. 3.17). Each group is one-half the size of the previous group. Adjacent members in the same group are reduced in size by a factor of the sixth root of two (0.8909). The largest member of a group is then related by this same factor to the bottom member of the previous group.

The basic unit of size is the cycle or line-pair, shown in Fig. 3.18. Spatial frequency is the number of cycles per mm. and angular frequency the number of cycles per milliradian. In television or array-based systems, the resolution limit is determined by the number of lines or pixels per line. There are two television lines per optical cycle or line-pair. An image array with pixels separated by 0.01 mm can resolve 50 cycles/mm.

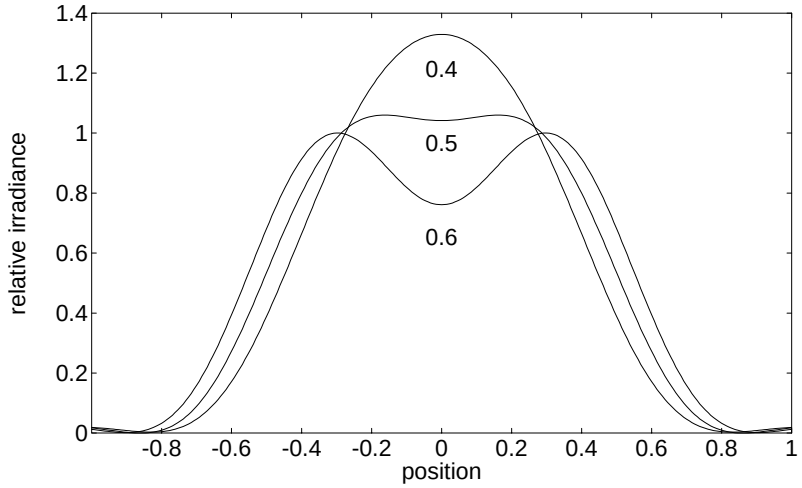
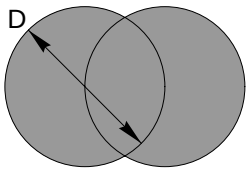
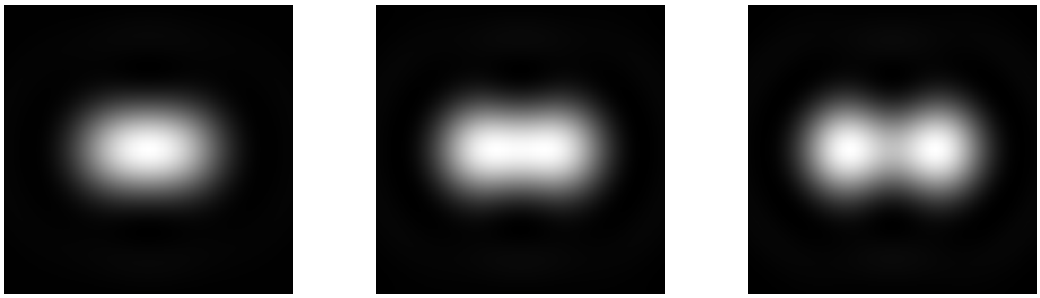


Figure 3.14: Cross-section plots of irradiance for ideal circular point images separated by distances varying from unresolved to clearly resolved.



Separation is D/2

Figure 3.15: Resolution of two uniformly illuminated disks.



$0.4\epsilon_o$ separation $0.5\epsilon_o$ separation $0.6\epsilon_o$ separation

Figure 3.16: Computer-generated images for two ideal circular point images separated by small distances.

As the size of a structured object decreases, its spatial frequency content increases. The spatial frequency of each group in Fig. 3.17 is a factor of two larger than the group to the left.

The effect on the optical system on the image of the resolution chart is to blur the edges of the bars. Fig. 3.18 shows a cylindrical point spread function being scanned across a three bar target. The diameter of the cylinder function is 1.5 units. The maximum and minimum brightness levels in the image are reduced from those in the object because the point spread function is larger than a single bar. This reduces the contrast of the image. As the size and separation of the bars decreases, the blurred edges begin to blend together, and this blending continues until individual bars are no longer distiguishable.

Limiting resolution of an optical system is tested by visual examination of resolution-chart images. The tests should be done at sufficient magnification and illumination to minimize limitations imposed by the eye of the observer. The limiting resolution is the smallest set of bars for which individual bars may be recognized. This is, of course, a subjective measurement and should be determined statistically using a number of different observers.

An objective measure of resolution quality is the *contrast* or *modulation* M, defined by

$$M = \frac{V_{\max} - V_{\min}}{V_{\max} + V_{\min}} \tag{3.22}$$

where $V_{\max}$ and $V_{\min}$ are the image brightness levels shown in Fig. 3.18. A plot of modulation as a function of spatial frequency is the *contrast function* (CF). A hypothetical CF is shown as the solid curve in Fig. 3.19. We distinguish here between modulation determined for a three-bar target and that determined for a sinuosoidal target. The response to sinusoidal targets of varying spatial frequency is the *modulation transfer function* (MTF). The curve indicating the smallest modulation detectable by a sensor or observer is the *threshold modulation function* TM. A typical threshold curve is shown as the dashed line in Fig. 3.19. Note that The threshold curve usually rises with spatial frequency. The intersection of the

Figure 3.17: Simple three-bar resolution chart.

threshold curve with the contrast function yields the limiting frequency of the optical system. TM curves are derived empirically from repeated experiments with different observers, batches of photographic emulsions, or sensor samples. The final TM curve represents an average of many measurements and should really include uncertainty bands to indicate the expected spread or uncertainty in the values.

The biggest advantage of using the threshold modulation curve to determine limiting modulation is that it separates measurements involving observers or the sensing system from those involving the image-forming system. There are many technical limitations to the approach [3], but in general the use of a threshold curve is helpful in understanding the nature of the imaging process.

3.7 Modulation Transfer Function

The idea of measuring the contrast of periodic images can be generalized by analogy to the Fourier methods used for signal analysis in Electrical Engineering. Instead of a three-bar target, we use a sine-wave target of variable spatial frequency, as shown in Fig. 3.20. The ideal target should be of infinite extent and produce a varying irradiance pattern given by $\cos(2\pi f_x x)$. The *modulation transfer function* (MTF) is the ratio of the contrast of the actual image relative to that of an ideal image (object pattern scaled by the magnification) as a function of the spatial frequency f_x.

If the optical system is a true linear, shift-invariant (LSI) system, then its response to a

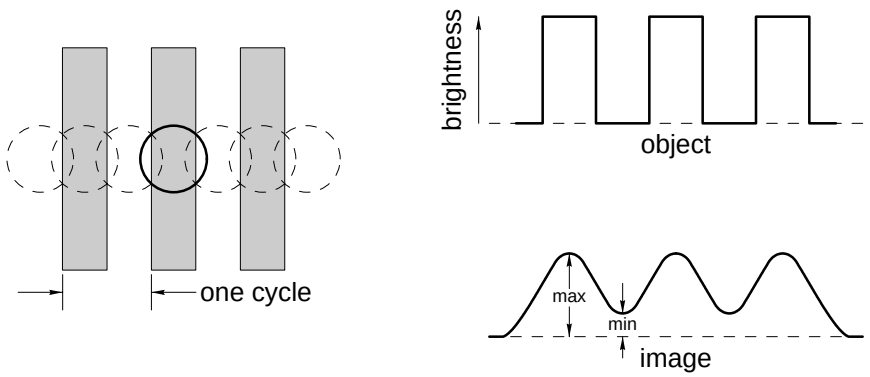


Figure 3.18: Contrast of a three-bar pattern.

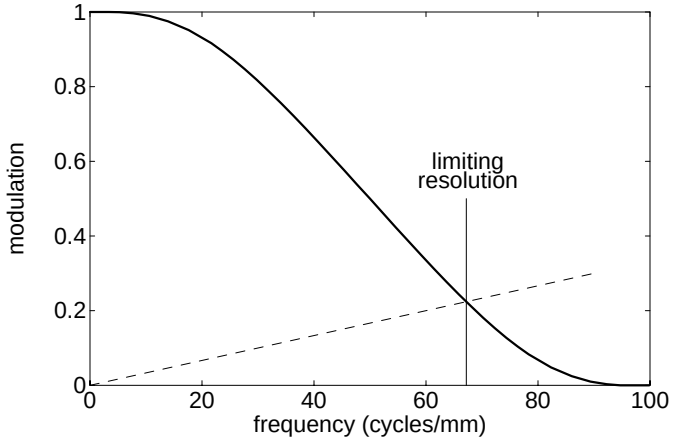


Figure 3.19: Modulation transfer function (solid curve) and the threshold modulation curve (dashed line) intersecting at the limiting resolution of the optical system.

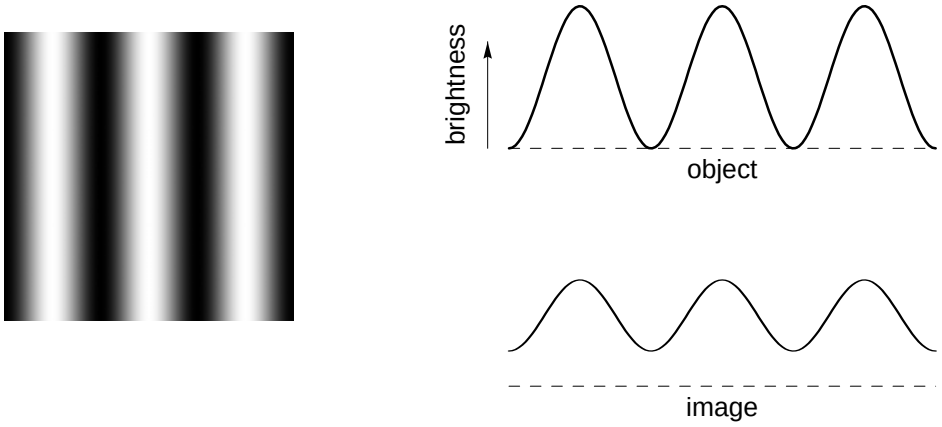


Figure 3.20: Contrast of a sinusoidal pattern.

sine-wave input pattern will be a sine-wave output pattern of the same frequency but different amplitude (and possibly phase). This is a consequence of sine waves being eigenfunctions of LSI systems. Bar targets are popular test charts because it is much easier to make a chart with a binary irradiance profile (white-black) than it is to make a continuous gray-level chart with the proper transmittance or reflectance properties. The mathematical form of the response to a bar target, however, depends on the spatial frequency of the target and on the mathematical form of the point spread function. A sine-wave target, on the other hand, will produce a sine-wave response no matter what the form of the point spread function.

3.8 Optical Transfer Function

The image of an LSI optical system is the convolution of the point spread function with the ideal image,

$$f(x,y) = h(x,y) * f_i(x,y) \tag{3.23}$$

where $h(x,y)$ is the point spread function and $f_i(x,y)$ the ideal image irradiance function. The Fourier transform operation allows us to equate the spatial-frequency spectrum of the image to the product of the spectrum of the object and the *Optical Transfer Function* (OTF),

$$F(f_x,f_y) = H(f_x,f_y)F_i(f_x,f_y) \tag{3.24}$$

where $H(f_x,f_y)$ is the OTF and $F_i(f_x,f_y)$ the spectrum of the ideal image.

The modulation transfer function, discussed in the previous section, is the modulus of a cross-section of the optical transfer function.

There are essentially three different functions involved in calculating the OTF. These are the pupil function, point spread function, and optical transfer function. The relationships among these functions are diagrammed in Fig. 3.21

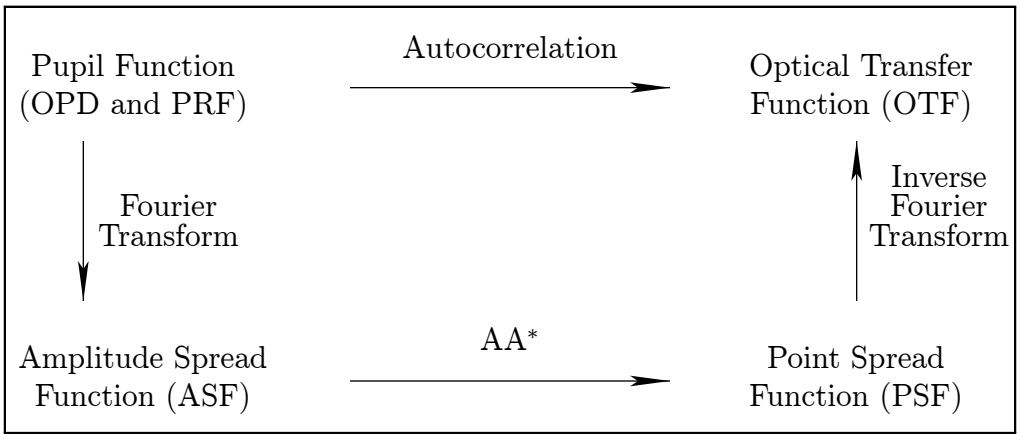


Figure 3.21: Optical Functions

The pupil function describes the magnitude and phase of the light in the pupil plane. The modulus squared of the pupil function, or pupil radiance function (PRF), gives the radiance distribution of the light over the pupil. The phase of the pupil function is proportional to the wavefront OPD.

The amplitude spread function (ASF) is the Fourier transform of the pupil function. It describes the magnitude and phase of the light at the focal plane. The point spread function (PSF) is the square modulus of the amplitude spread function and describes the distribution of energy at the focal plane. The point spread function is generally considered to be the impulse response of an optical system, the actual image produced by a point object. The PSF of a uniform circular pupil with no aberrations is an Airy function. The PSF can be used to obtain the value for the Strehl ratio. It can also be used to produce a diffraction-based energy distributions in a similar way that a spot diagram is used to generate a geometrical energy distributions.

The optical transfer function (OTF) is the autocorrelation of the pupil function. It is also the inverse Fourier transform of the point spread function. The OTF delineates the response of the optical system to spatial frequencies present in the object. Multiplication of the OTF by the Fourier transform of the object gives the Fourier transform of the image.

The OTF of an ideal optical system having a circular aperture with uniform transmittance can be derived from the autocorrelation of the pupil function, as shown in Fig. 3.22. If the pupil radius is a, then the radius of the autocorrelation function is $2a$, and the maximium spatial frequency imaged by the optical system (in air) is

$$f_m = \frac{2a}{\lambda R} = \frac{2}{\epsilon_o} \tag{3.25}$$

The modulation is proportional to the shaded area of overlap in the shifted pupils, at a spatial frequency proportional to the distance s between the centers. The overlap area is

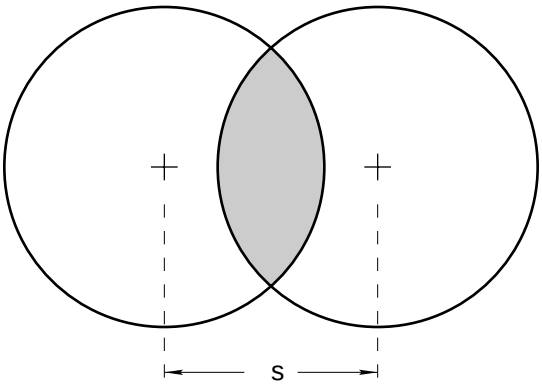


Figure 3.22: Autocorrelation of ideal circular pupil is given by shaded area as function of s.

twice that of the segment shown in Fig. 3.10. The OTF of the ideal circular pupil is therefore

$$H(x) = \frac{2}{\pi}\left(\cos^{-1}(x) - x\sqrt{1-x^2}\right) \tag{3.26}$$

where $x = f/f_m$. If the pupil function The OTF for this case is independent of the orientation of the spatial frequency component.

The maximum spatial frequency can also be expressed as

$$f_m = \frac{1}{\lambda F_\#} \tag{3.27}$$

where $F_\#$ is the f-number. For example, if the wavelength is 550 nm and the system is f/2.8, the maximum spatial frequency is 645 cycles/mm.

3.9 Focal Shift and Spurious Resolution

Focal shift is an aberration of the form $W(r) = A_z r^2$. The resulting geometrical point spread function is a uniformly illuminated circular spot, that is

$$U(r) = \text{cyl}(r/d) \tag{3.28}$$

where d is the blur diameter of the geometrical spot. The OTF is the Fourier transform of the point spread function,

$$H(f) = \text{somb}(df) \tag{3.29}$$

where f is the spatial frequency for any orientation. This is the geometrical approximation of the Optical Transfer function.

If a geometrical representation of the point spread function is used, the resulting OTF is the geometrical optical transfer function (GOTF). If a diffraction integral is used to calculate the point spread function, the corresponding OTF is the diffraction optical transfer function.

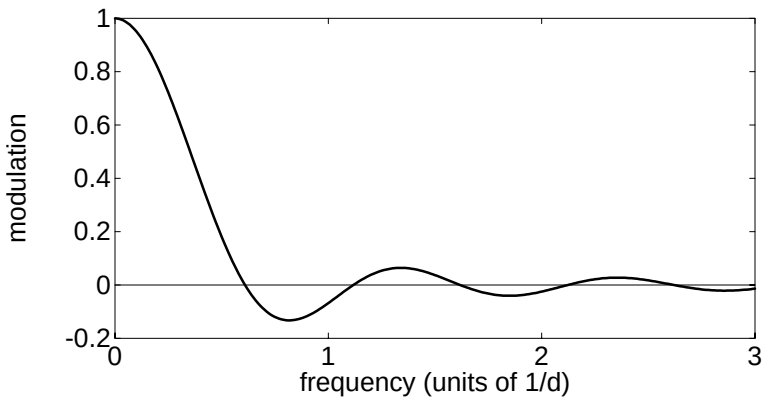


Figure 3.23: Geometrical Optical Transfer Functions for a cylindrical point spread function of diameter d.

The pupil function of an optical system is space limited and often sharply bounded by the edge of an aperture. The associated point spread function, according to physical optics and Fourier transform theory, must be infinite in extent. The OTF, on the other hand, is frequency limited because it is given by the autocorrelation of a function of limited extent.

In the case of geometrical optics, however, the geometrical spread function is space limited, so that its Fourier transform, the GOTF is unlimited in the frequency domain. Any analysis involving the GOTF must be restricted to spot sizes large enough that there is no significant modulation at frequencies higher than the maximum allowed by physical optics. For example, the first zero of the GOTF for a uniform spot is given by $df = 1.22$. If this zero is to correspond to a spatial frequency less than $f_m = 2/\epsilon_o$, then d must be less than 0.61 ϵ_o. In other words, the geometrical spot must be at least as large in diameter as the central peak of the Airy function.

Fig. 3.23 shows the GOTF for a cylindrical point spread function. Notice that the modulation switches from positive to negative and back again as the frequency increases. The ordinary definition of modulation does not allow for the existence of negative values. Our interpretation of modulation must be extended to include not only the magnitude of the modulation but also the *phase* of the modulation. Negative modulation values indicate a region of phase reversal, in which white bars become black and vice versa. This phenomenon is called *spurious resolution*. An example of spurious resolution is shown in Fig. 3.24. The image on the left is a fully resolved sector target, for which the angular spatial frequencies increase toward the center of the target. The image on the right has been convolved with a cylindrical point spread function. Image contrast decreases from right to left as the angular spatial frequency increases, finally blurring to zero, but then increasing again with reversed contrast. Several cycles of contrast reversal may be observed if the range in spatial frequencies is large enough.

Fig. 3.25 shows what happens during contrast reversal. Circular point spread functions of

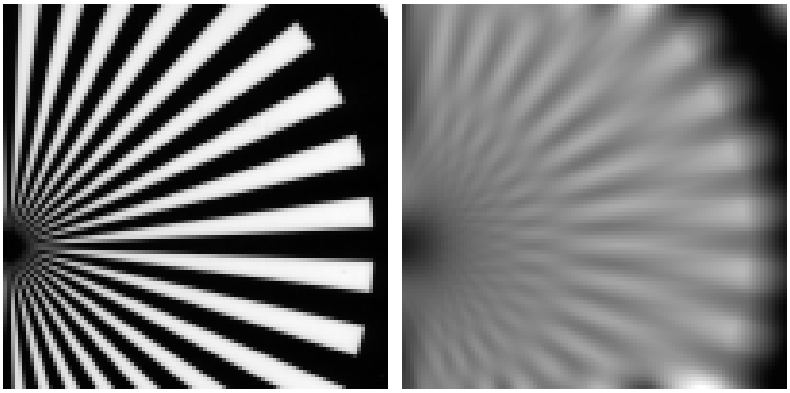


Figure 3.24: Spurious resolution in a sector target.

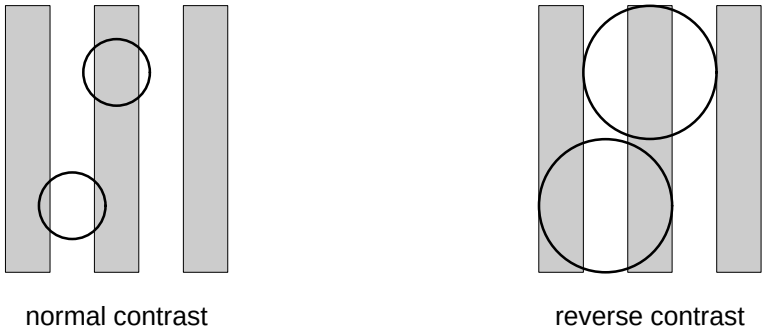


Figure 3.25: Spurious resolution.

different sizes are shown centered on the white or black bars of a three-bar target. The actual image irradiance is product of the point spread function and the ideal image, integrated over the area of the the point spread function. In this case both functions are binary functions, so the image irradiance is the difference between the area covered by white and the area covered by black. If the point spread function is small compared to the bar spacing, then the center of a black bar will be imaged as black and that of a white bar will be imaged as white. If the point spread function is large enough, however, the area outside of the bar on which the cylinder is centered will be larger than the area inside, resulting in contrast reversal.

3.10 Optical Transfer Function from Knife Edge Distribution

The optical transfer function may be calculated by generating a knife edge energy distribution table at n equal energy increments and then transforming that data.

Let $E(s)$ represent the fractional energy exposed by a knife edge as a function of distance s measured along the line of travel. The optical transfer function is given by

$$H(f) = \int_{-\infty}^{+\infty} \frac{dE(s)}{ds} e^{j2\pi f s} ds \tag{3.30}$$

For example, $E(s)$ might be given in the tabular form,

$E(s)$	s
0	s_0
.01	s_1
.02	s_2
$\vdots$	$\vdots$
1.0	s_{100}

where n = 100. The function $E(s)$ can be approximated by a sequence of line segments joining the tabular points. Then between s_{i-1} and s_i

$$\frac{dE(s)}{ds} = \frac{\Delta E}{\Delta_i} \tag{3.31}$$

where $\Delta E = 1/n$ is fixed and

$$\Delta_i = s_i - s_{i-1} \tag{3.32}$$

Then the integral reduces to

$$H(f) = (1/n) \sum_{i=1}^{n} \text{sinc}(f\Delta_i) e^{j2\pi f \hat{s}_i} \tag{3.33}$$

where

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \tag{3.34}$$

and

$$\hat{s}_i = (s_i + s_{i-1})/2 \tag{3.35}$$

3.11 Geometrical Optical Transfer Function

The geometrical point spread function $h_g(x, y)$ is the ray density in the image plane and is inversely proportional to the Jacobian J given by

$$J = \frac{\partial(x_e, y_e)}{\partial(x, y)} = \begin{vmatrix} \frac{\partial^2 W}{\partial x^2} & \frac{\partial^2 W}{\partial x \partial y} \\ \frac{\partial^2 W}{\partial x \partial y} & \frac{\partial^2 W}{\partial y^2} \end{vmatrix} \tag{3.36}$$

where $(x,\ y)$ are pupil coordinates and $(x_e,\ y_e)$ are image plane coordinates. We can now write the GOTF as the Fourier transform of the geometrical point spread function,

$$H(f_x, f_y) = \int\int \frac{\partial(x,y)}{\partial(x_e,y_e)} \exp\left[2\pi j(f_x x_e + f_y y_e)\right] dx_e dy_e \tag{3.37}$$

The Jacobian in the above expression can be used to change the variables of integration, which gives

$$H(f_x, f_y) = \frac{1}{A}\int\int \exp\left[-2\pi j \epsilon_o \left(f_x \frac{\partial W}{\partial x} + f_y \frac{\partial W}{\partial y}\right)\right] dxdy \tag{3.38}$$

where A is the area of the exit pupil.

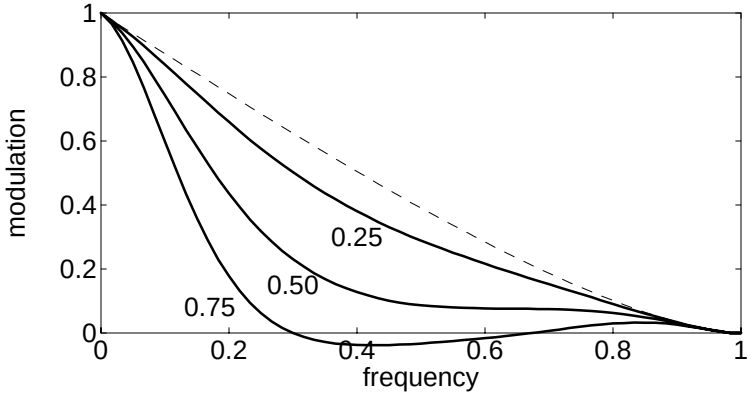
3.12 Effect of Aberrations on the OTF

The effect of aberrations on the OTF is to reduce the contrast. There is no combination of optical aberrations that will increase contrast at any spatial frequency above the diffraction limit at that frequency. Wavefront aberration is a phase function, which means that area of overlap for the pupil autocorrelation of the aberrated pupil must be smaller than that of the unaberrated pupil.

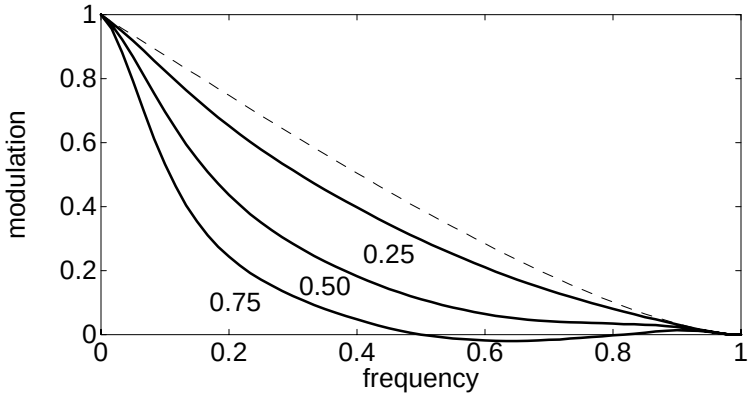
The effect of small aberrations on the point spread function is to reduce the magnitude of the central peak, with the extra energy being distributed into the diffraction rings. The modulation at higher spatial frequencies remains near diffraction-limited so long as a sharp central diffraction peak remains a dominant feature of the diffraction pattern. The contrast at low spatial frequencies, however, is reduced by the larger effective spot size due to the ring structure. For larger aberrations, however, the geometrical OTF approximates the corresponding physical OTF. The correspondence includes having zero-crossings at frequencies inversely proportional to the geometrical spot size and regions of spurious resolution.

As examples, consider the following three aberrations: focal shift $W(r) = A_1 r^2$, fourth-order spherical aberration at the paraxial focus $W(r) = A_2 r^4$, and balanced fourth-order spherical aberration (mid-focus) $W(r) = A_3(r^4 - r^2)$. Fig. 3.26 shows OTF curves for each case with peak-to-valley (p-v) wavefront values of 0.25, 0.50, and 0.75 waves. All curves shown were calculated from diffraction-based point spread functions. The dashed curve is the OTF for the ideal optical system. For aberrations of 0.25 and 0.50 the OTF remains non-zero up to the diffraction cut-off frequency. For the 0.75 wavefront curves, the modulation crosses zero before the diffraction cut-off. The 0.75-wave p-v curve for defocus already resembles the form expected from geometrical optics. Fig. 3.27 shows diffraction and geometrical OTF curves for a focal shift of 2.5 waves. The geometrical OTF is the function somb$(10x)$. The basic features of the two curves are the same, although the location of modulation zero crossings are not.

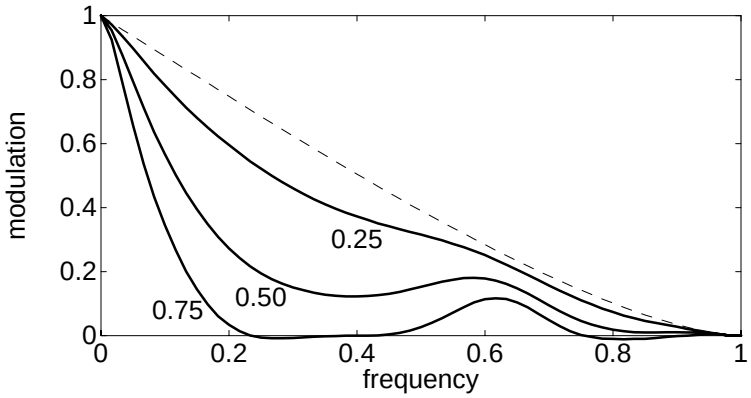
Fig. 3.28 shows the OTF plots for optical systems with 0.5 wave p-v of the three aberrations used in Fig. 3.26. All have approximately the same Strehl ratio and rms spot size, so which is the preferred system? None of the OTF curves are superior for all frequencies. The



(a) Focal shift $W(r) = A_1 r^2$.



(b) Spherical aberration (paraxial focus) $W(r) = A_2 r^4$.



(c) Spherical aberration (mid focus) $W(r) = A_3(r^4 - r^2)$.

Figure 3.26: Optical transfer functions for small amounts of wavefront aberration. The labels are the peak-valley waves of aberration for each curve. The dashed line is the OTF for the ideal system.

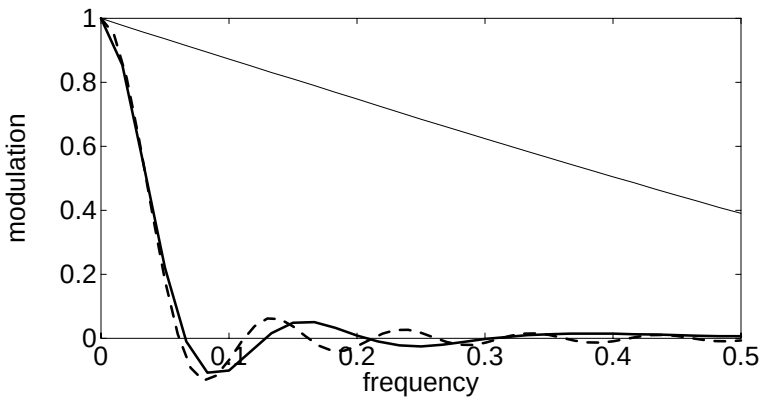


Figure 3.27: Optical transfer functions for focal shift of 2.5 waves. The solid curve is the diffraction OTF and the dashed curve the geometrical OTF. The thin solid line is the OTF for the ideal system.

answer depends on the weight given to the response for different spatial frequencies, and this depends on the application for which the optical system is intended. The problem of how to trade-off resolution among frequency bands does not seem to come with easy solutions.

The aberrations and point spread functions considered so far have been radially symmetric. For the nonsymmetric aberrations such as astigmatism and coma, the modulation transfer of spatial sine patterns depends not only on the spatial frequency but also on the orientation of the pattern. Optical tests involving sector targets, such as shown in Fig. 3.24, are very useful in such situations because all target orientations are represented. In optical design calculations, the common practice is to find the OTF for horizontal and vertical pattern orientations. In some situtations, however, this practice is not sufficient to characterize the system. For example, consider fourth-order astigmatism at mid focus, which is given by a wavefront $W(x, y) = A_4(x^2 - y^2)$. Fig 3.29 shows a projection plot of the MTF. The full OTF function $H(f_x, f_y)$ is complex Hermitian (real part even, imaginary part odd) because it is the Fourier transform of a real function. Note that the cross-sections along the x- and y-axes are the same, but that the cross-sections along the 45-degree directions are different.

The OTF for astigmatism is a real, even function because the point spread function is real and even in both x and y. In the case of coma the point spread function has bilateral symmetry but is neither even or odd, so that the OTF is complex valued. This means there there is a non-trivial phase component to the OTF. The phase shift has the effect, for example, of distorting the relative position of image features such as edges rather than merely softening them. Fig. 3.30 shows the modulation transfer function for coma and Fig. 3.31 shows a contour plot of the imaginary part of the OTF for the same coma wavefront. The dashed curves are negative contours and the solid curves are positive contours. Coma should be small in a well-designed optical system, because of its effect on isoplanatism, so that the imaginary part of the optical transfer function should also be small. This does not exclude

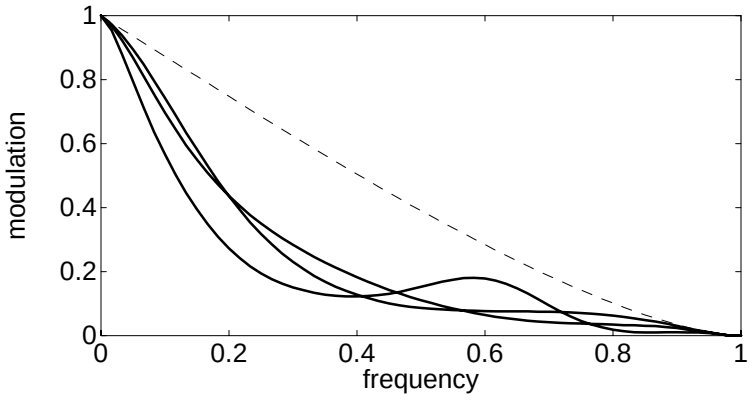


Figure 3.28: Optical transfer functions for 0.5 wave peak-valley of three different aberrations. The dashed line is the OTF for the ideal system.

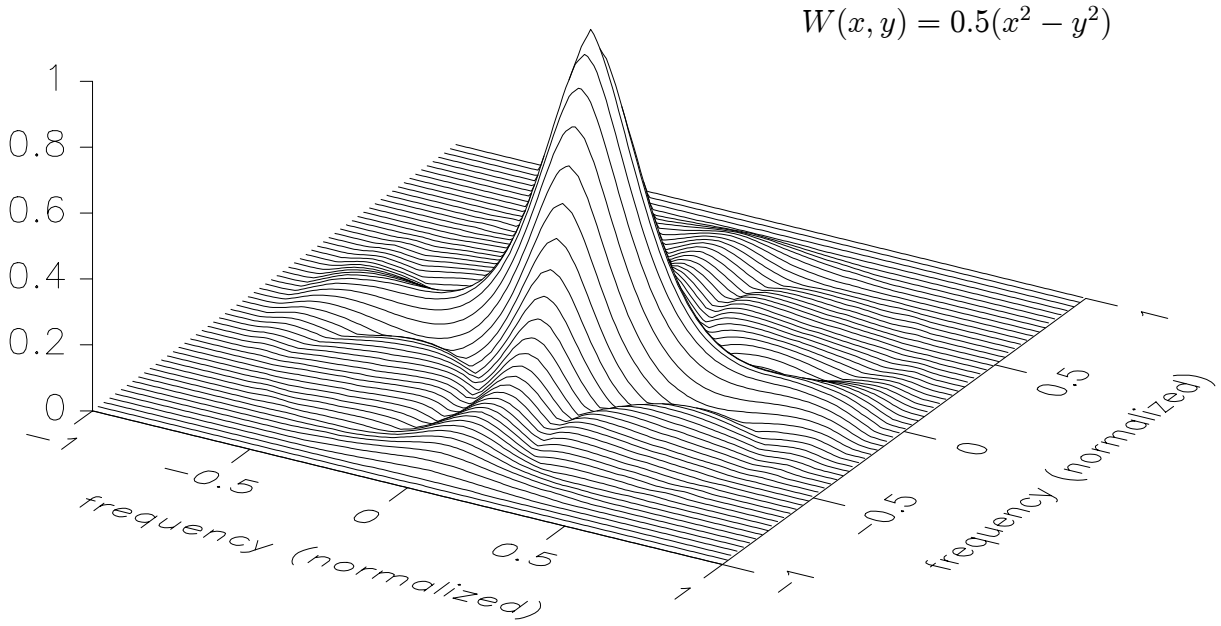


Figure 3.29: Modulation Transfer Function for astigmatism at mid-focus.

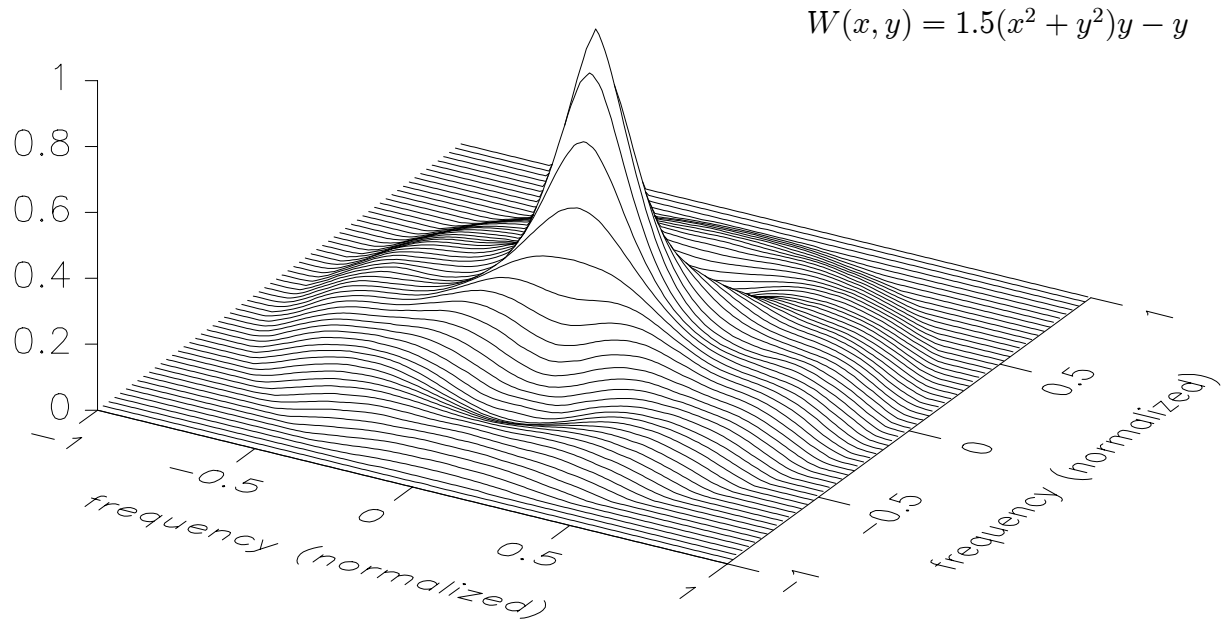


Figure 3.30: Modulation Transfer Function for coma.

effects of spurious resolution, which can be present in any optical system.

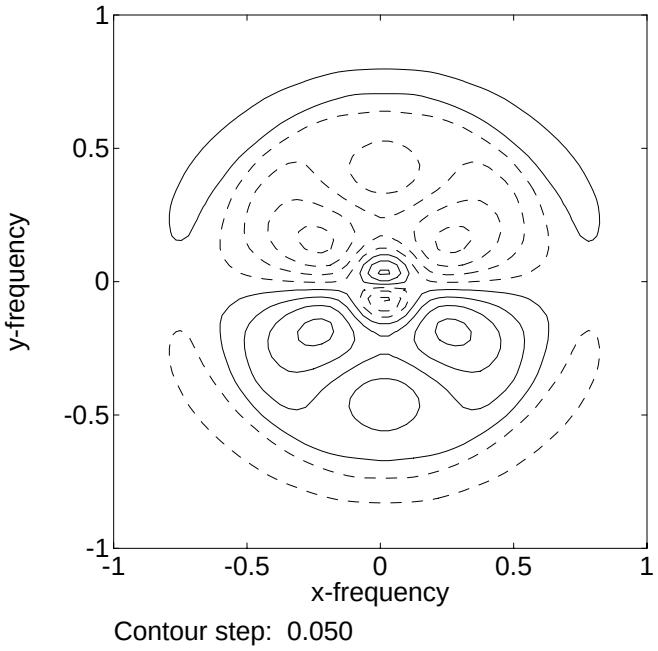


Figure 3.31: Imaginary part of the Optical Transfer Function for coma. The dashed curves are negative contour values.

Bibliography

[1] Jack D. Gaskill, *Linear Systems, Fourier Transforms, and Optics*, (John Wiley & Sons, 1978).

[2] W. T. Welford, *Aberrations of Optical Systems*, (Adam Hilger Ltd, 1986).

[3] William B. Wetherwell, "The Calculation of Image Quality," in *Applied Optics and Optical Engineering*, Volume VIII, Edited by Robert R. Shannon and James C. Wyant, (Academic Press, 1980).

[4] C. S. Williams and O. A. Becklund, *Introduction to the Optical Transfer Function*, (John Wiley & Sons, 1989).