1051-455-20073, Physical Optics

1 Laboratory 6: Diffraction and Imaging

References:*Introduction to Optics*, Pedrotti & Pedrotti, Chapters 16,18; *Optics*, E. Hecht, Chapter 10

We are able to determine locations of images and their magnifications by using the concept of light as a "ray" (geometrical optics) via the equation:

$$\frac{1}{z_1} + \frac{1}{z_2} = \frac{1}{\mathbf{f}}$$

where z_1 and z_2 are the distances from object to lens and from lens to image and $\mathbf{f}$ is the focal length of the lens, which measures its ability to deviate rays. The model of light as a "wave" adds a fundamental constraint to an imaging system via the inherent process of "diffraction." The concepts/processes of "interference" and "diffraction" actually are different manifestations of the same phenomenon, differing only in the number of sources involved (interference $\Longrightarrow$ few sources, say 2 - 10 and the light is recombined deliberately; diffraction $\Longrightarrow$ many sources, up to an infinite number, and the light is recombined naturally). Diffraction is the source of the theoretical limit to the performance of an imaging system.

1.1 Theory

This lab will investigate the diffraction patterns generated from apertures of different shapes and observed at different distances. As we have mentioned, the physical process that results in observable diffraction patterns is identical to that responsible for interference. In the latter case, we generally speak of the intensity patterns generated by light after passing through a few apertures whose size(s) generally are smaller than the distance between them. The term *diffraction* usually is applied to the process either for a single large aperture, or (equivalently) a large number of small (usually infinitesmal) contiguous apertures and the light recombines "naturally" as a part of its propagation.

In studies of both interference and diffraction, the patterns are most obvious if the illumination is *coherent*, which means that the phase of each sinusoidal electric field is strictly deterministic. In other words, knowledge of the phase of the field at some point in space and/or time determines the phase at other points in space and/or time. Coherence has two flavors: *spatial* and *temporal.* For *spatially coherent* light, the phase difference $\Delta\phi \equiv \phi_1 - \phi_2$ of the electric field measured at any pair of locations in space by a vector distance $\Delta\underline{\mathbf{r}}$ at the instant of time ($\Delta t = 0$) is the same. In equations, we could write this as:

$$\Phi\left[\underline{\mathbf{r}_1}, t_1\right] - \Phi\left[\underline{\mathbf{r}_1} + \Delta\underline{\mathbf{r}}, t_1\right] = \Phi\left[\underline{\mathbf{r}_2}, t_2\right] - \Phi\left[\underline{\mathbf{r}_2} + \Delta\underline{\mathbf{r}}, t_2\right] \Longrightarrow \textit{spatial coherence}$$

If the phase difference measured at the SAME location at two different times separated by $\Delta t \equiv t_1 - t_2$ is the same for all points in space, the light is *temporally coherent*:

$$\Phi\left[\underline{\mathbf{r}_1}, t_1\right] - \Phi\left[\underline{\mathbf{r}_1}, t_1 + \Delta t\right] = \Phi\left[\underline{\mathbf{r}_2}, t_2\right] - \Phi\left[\underline{\mathbf{r}_2}, t_2 + \Delta t\right] \Longrightarrow \textit{temporal coherence}$$

Light from a laser may be considered to be BOTH spatially and temporally coherent. The properties of coherent light allow phase differences of light that has traveled different paths to be made visible, since the phase difference is constant with time. In intereference, the effect often results in a sinusoidal fringe pattern in space. In diffraction, the phase difference of light from different points in the same large source can be seen as a similar pattern of dark and bright fringes, though not (usually) with sinusoidal spacing.

Observed diffraction patterns from the same object usually have very different forms at different distances from the object to the observation plane. If viewed very close to the aperture (in the *Rayleigh-Sommerfeld* diffraction region), then Huygens' principle says that the amplitude of the electric field is the summation (integral) of the spherical wavefronts generated by each point in the aperture. The resulting amplitude pattern may be quite complicated to evaluate. If observed somewhat farther from the aperture, the spherical wavefronts may be accurately approximated by paraboloidal wavefronts. The approximation applies in the *near field*, or the *Fresnel diffraction region.* If viewed at a large distance compared to the extent of the object, the light from different locations in the aperture may be accurately modeled as *plane waves* with different wavefront tilts. This occurs in the *Fraunhofer diffraction region.*

1.1.1 Rayleigh-Sommerfeld Diffraction: Spherical Waves

A spherical wave of light with frequency ν_0 emitted by a source at the origin and observed at $\underline{\mathbf{r}} = [x, y, z]$ has the form:

$$\begin{aligned} f[x,y,z,t] &= \frac{A_0}{r}\exp\left[+i\left(kr-\omega_0 t\right)\right] \\ &= \frac{A_0}{\sqrt{x^2+y^2+z^2}}\exp\left[+2\pi i\left(\frac{\sqrt{x^2+y^2+z^2}}{\lambda_0}-\nu_0 t\right)\right] \end{aligned}$$

If emitted at a different location (say, $\underline{\mathbf{r}}_0 = [x_0, y_0, z_0]$, then the distances in the denominator and the exponent must be adjusted:

$$f[x,y,z,t;x_0,y_0,z_0] = \frac{A_0[x_0,y_0,z_0]}{\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}}\exp\left[+2\pi i\left(\frac{\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}}{\lambda_0}-\nu_0 t\right)\right]$$

The amplitude at an observation point generated by light from a number of such sources is the integral over the source function:

$$\begin{aligned} g[x,y,z,t] &= \iiint_{\text{source}} f[x,y,z,t;x_0,y_0,z_0]\; dx_0\; dy_0\; dz_0 \\ &= \iiint_{\text{source}} \frac{A_0[x_0,y_0,z_0]}{\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}} e^{+2\pi i\left(\frac{\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}}{\lambda_0}-\nu_0 t\right)}\; dx_0\; dy_0\; dz_0 \end{aligned}$$

Often, the source is constrained to a 2-D plane $[x_0, y_0]$ and observed at a 2-D plane $[x, y; z_1]$ so the integral is 2-D as well:

$$g[x,y,t;z_1,z_0] = e^{-2\pi i\nu_0 t}\iint_{\text{source}} \frac{A_0[x_0,y_0;z_0]}{\sqrt{(x-x_0)^2+(y-y_0)^2+(z_1-z_0)^2}} e^{+2\pi i\frac{\sqrt{(x-x_0)^2+(y-y_0)^2+(z_1-z_0)^2}}{\lambda_0}}\; dx_0\; dy_0$$

We often discard the time dependence and specify the source as a 2-D function $A_0[x_0, y_0; z_0] \equiv f[x_0, y_0]$ in the plane $z = z_0$:

$$g[x,y;z_1,z_0] = \iint_{-\infty}^{+\infty} \frac{f[x_0,y_0]}{\sqrt{(x-x_0)^2+(y-y_0)^2+(z_1-z_0)^2}} e^{+2\pi i\frac{\sqrt{(x-x_0)^2+(y-y_0)^2+(z_1-z_0)^2}}{\lambda_0}}\; dx_0\; dy_0$$

The mathematical solution of this integral is difficult for all but the simplest source functions, so we will not consider this case further.

1.1.2 Fresnel Diffraction

If the distance between the object and the observation plane is assumed to be large compared to the 2-D extent of a planar object, then the diffraction integral is significantly simpler (though perhaps still challenging). Under these conditions in the *Fresnel* diffraction region, the wavefront emitted by each source point in the 2-D object plane is assumed to have a paraboloidal shape (rather than spherical) and its amplitude is assumed to be constant at all locations in the observation plane; in other words, the inverse square law does not apply to different points in the observation plane. The integral over the source object is the superposition of the individual paraboloidal waves, which have the form of quadratic-phase factors:

$$g\left[x,y;z_1,z_0\right]=\iint\limits_{-\infty}^{+\infty}\frac{f\left[x_0,y_0\right]}{z_1}e^{+2\pi i\frac{1}{\lambda_0}\left(z_1+\frac{(x-x_0)^2+(y-y_0)^2}{z_1}\right)}\,dx_0\;dy_0$$

which may be rewritten as a convolution with an impulse response $h\left[x,y;z_1\right]$

$$\begin{aligned}
g\left[x,y\right] &= \frac{1}{i\lambda_0 z_1}\exp\left[+\frac{2\pi i z_1}{\lambda_0}\right]\iint_{-\infty}^{+\infty} f\left[x-\alpha,y-\beta\right]\exp\left[-\frac{i\pi(\alpha^2+\beta^2)}{\lambda_0 z_1}\right]\,d\alpha\;d\beta\\
&= \frac{1}{i\lambda_0 z_1}\exp\left[+\frac{2\pi i z_1}{\lambda_0}\right]\left(f\left[x,y\right]*\exp\left[-\frac{i\pi(x^2+y^2)}{\lambda_0 z_1}\right]\right)\\
&= f\left[x,y\right]*\left(\frac{1}{i\lambda_0 z_1}\exp\left[+\frac{2\pi i z_1}{\lambda_0}\right]\exp\left[-\frac{i\pi(x^2+y^2)}{\lambda_0 z_1}\right]\right)\\
&\equiv f\left[x,y\right]*h\left[x,y;z_1\right]\\
h\left[x,y;z_1\right] &\equiv \left(\frac{1}{i\lambda_0 z_1}\exp\left[+\frac{2\pi i z_1}{\lambda_0}\right]\exp\left[-\frac{i\pi(x^2+y^2)}{\lambda_0 z_1}\right]\right)
\end{aligned}$$

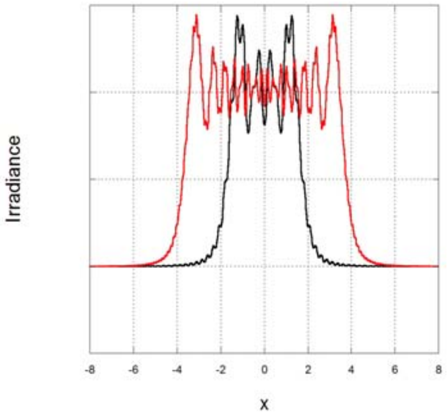
The "impulse response of light propagation" is:

$$h\left[x,y;z_1\right]\equiv\left(\frac{1}{i\lambda_0 z_1}\exp\left[+\frac{2\pi i z_1}{\lambda_0}\right]\exp\left[-\frac{i\pi(x^2+y^2)}{\lambda_0 z_1}\right]\right)$$

for wavelength λ_0 and axial distance z_1. The calculation has the form of the mathematical operation of *convolution* of the object pattern and a *quadratic-phase pattern* (which represents the paraboloidal shape of the individual waves). The convolution operation involves a translation of the reversed input function in the integration coordinates $[\alpha,\beta]$, followed by multiplication by the quadratic-phase factor $\exp\left[-\frac{i\pi(\alpha^2+\beta^2)}{\lambda_0 z_1}\right]$ and then evaluation of the area for each value of the output coordinates $[x,y]$. Thus it is "complicated" and computationally intensive.

The resulting diffraction pattern resembles the original object with "fuzzy" or "ringing" edges. The extents of the object and the diffraction pattern are (in some sense) proportional; if the object size is increased, so will be the extent of the Fresnel diffraction pattern. Though less obvious to show mathematically, the distance between the "ringing" fringes increases as the distance to the observation plane increases.

A 1-D model of Fresnel diffraction for square input apertures with different widths is shown in the figure:



Fresnel diffraction patterns from slits of different widths observed at the same propagation distance; the width of the pattern is proporitional to the width of the slit.

1.1.3 Fraunhofer Diffraction

At large distances from the object plane, the diffraction is in the *far field* or *Fraunhofer* diffraction region. Here, the pattern of diffracted light usually does not resemble the object at all. The size of the observed pattern varies in proportion to the *reciprocal* of the object dimension, i.e., the larger the object, the smaller the diffraction pattern. Note that increasing the size of the object also produces a brighter diffraction pattern, because more light reaches the observation plane. The mathematical relation between the shape and size of the output relative to that of the input is a *Fourier transform*, which is a mathematical coordinate transformation that was "discovered" by Baron Jean-Baptiste Joseph de Fourier in the early 1800s. For the same input pattern $f[x,y]$, the diffraction pattern in the Fraunhofer region has the form:

$$\begin{aligned}\mathcal{F}_2\{f[x,y]\} &\equiv F[\xi,\eta] = \iint_{-\infty}^{+\infty} f[x,y]\left(\exp\left[+2\pi i\left(\xi x+\eta y\right)\right]\right)^* \, dx\, dy \\ &= \iint_{-\infty}^{+\infty} f[x,y]\exp\left[-2\pi i\left(\xi x+\eta y\right)\right]\, dx\, dy\end{aligned}$$

In words, the input function $f[x,y]$ is transformed into the equivalent function $F[\xi,\eta]$, where the coordinates ξ,η are spatial frequencies measured in cycles per unit length, e.g., cycles per mm. In optical propagation, the end result is a function of the original 2-D coordinates $[x,y]$, which means that the coordinates $[\xi,\eta]$ are "mapped" back to the space domain via a scaling factor. Since the coordinates of the transform have dimensions of $(\text{length})^{-1}$ and the coordinates of the diffracted light have dimensions of length, the scale factor applied to ξ and η must have dimensions of $(\text{length})^2$. It is easy to show that the scaling factor is the product of the two length parameters available in the problem: the wavelength λ_0 and the propagation distance z_1. The pattern of diffracted light in the Fraunhofer diffraction region is:

$$g[x,y] \propto \mathcal{F}_2\{f[x,y]\}|_{\lambda_0 z_1\xi\to x,\lambda_0 z_1\eta\to y} \equiv \iint_{-\infty}^{+\infty} f[\alpha,\beta]\exp\left[-2\pi i\left(\alpha\frac{x}{\lambda_0 z_1}+\beta\frac{y}{\lambda_0 z_1}\right)\right]\, d\alpha\, d\beta$$

In mathematical terms, this is a "linear, shift-variant" operation; it is linear because if the input amplitude is scaled by a constant factor, the output amplitude is scaled by the same factor. It is shift variant because a translation of the input does not produce a corresponding transformation of the output. Because Fraunhofer diffraction is shift variant, it may NOT be represented as a single convolution. However, once the Fourier transform is understood, it is very easy to visualize Fraunhofer diffraction patterns of many kinds of objects.

The study of Fourier transforms allows us to infer some important (and possibly counterintuitive) properties of Fraunhofer diffraction:

1. Scaling Theorem: If the scale factor of the aperture function $f[x,y]$ increases, then the resulting diffraction pattern becomes brighter and "smaller," i.e., the scale factor is proportional to the reciprocal of the scale factor of the input function.
2. Shift Theorem: Translation of $f[x,y]$ adds a linear term to the phase of the diffraction pattern, which is not visible in the irradiance. Thus translation of the input has no visible effect on the diffraction pattern.
3. Modulation Theorem: If an aperture can be expressed as the product of two functions, the amplitude of the diffraction pattern is their convolution.
4. Filter Theorem: if an aperture pattern is the convolution of two patterns, the amplitude of the resulting diffraction pattern is the product of the amplitudes of the individual patterns.

Example: Consider Fraunhofer diffraction of a simple 2-D rectangular object:

$$f[x,y] = RECT\left[\frac{x}{a},\frac{y}{b}\right] \equiv \begin{cases} 1 & if \quad |x|<\frac{a}{2} \text{ and } |y|<\frac{b}{2} \\ \frac{1}{2} & if \quad |x|=\frac{a}{2} \text{ and } |y|<\frac{b}{2} \text{ or } |x|<\left|\frac{b}{2}\right| \text{ and } |y|=\frac{b}{2} \\ \frac{1}{4} & if \quad |x|=\frac{a}{2} \text{ and } y=\frac{b}{2} \\ 0 & if \quad |x|>\frac{a}{2} \text{ and } >\frac{b}{2} \end{cases}$$

The integral evaluates rather easily:

$$g\left[x,y\right]\propto\left|\iint_{-\infty}^{+\infty}RECT\left[\frac{\alpha}{a},\frac{\beta}{b}\right]\exp\left[-\frac{2\pi i(x\alpha+y\beta)}{\lambda_0 z}\right]\,d\alpha\;d\beta\right|^2$$

$$\int_{y=-\frac{b}{2}}^{y=+\frac{b}{2}}\int_{x=-\frac{a}{2}}^{x=+\frac{a}{2}}\exp\left[-\frac{2\pi ix\alpha}{\lambda_0 z}\right]\exp\left[-\frac{2\pi iy\beta}{\lambda_0 z}\right]\,d\alpha\;d\beta$$

$$=\int_{x=-\frac{a}{2}}^{x=+\frac{a}{2}}\exp\left[\left(-\frac{2\pi ix}{\lambda_0 z}\right)\alpha\right]\,d\alpha\cdot\int_{y=-\frac{b}{2}}^{y=+\frac{b}{2}}\exp\left[\left(-\frac{2\pi iy}{\lambda_0 z}\right)\beta\right]\,d\beta$$

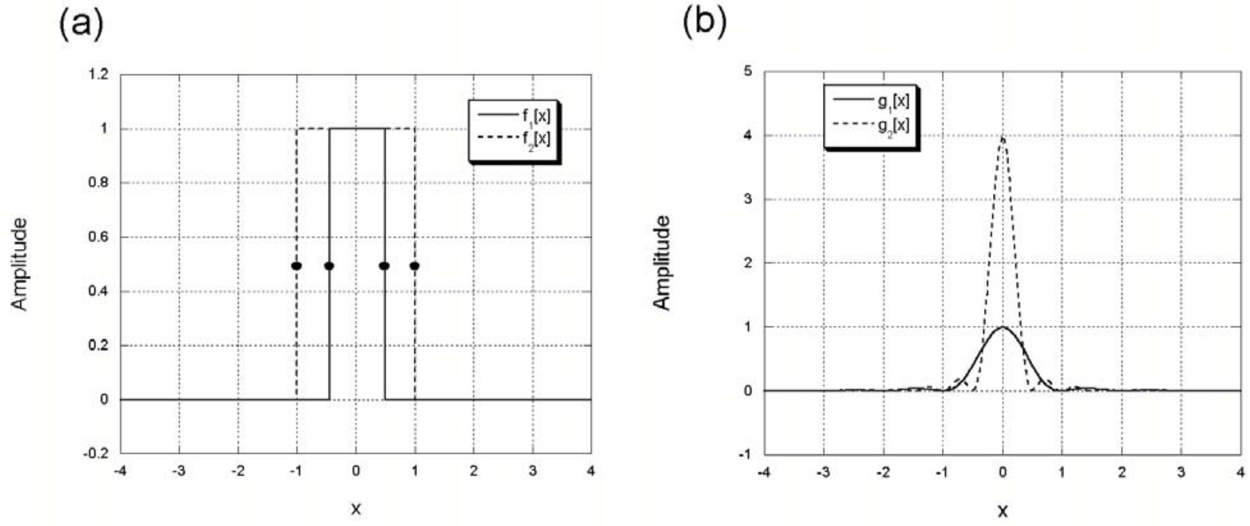
$$=\left.\frac{\exp\left[\left(-\frac{2\pi ix}{\lambda_0 z}\right)\alpha\right]}{\left(-\frac{2\pi ix}{\lambda_0 z}\right)}\right|_{\alpha=-\frac{\alpha}{2}}^{\alpha=+\frac{\alpha}{2}}\cdot\left.\frac{\exp\left[\left(-\frac{2\pi iy}{\lambda_0 z}\right)\beta\right]}{\left(-\frac{2\pi iy}{\lambda_0 z}\right)}\right|_{\beta=-\frac{\beta}{2}}^{\beta=+\frac{\beta}{2}}$$

$$=\frac{\exp\left[-i\frac{\pi ax}{\lambda_0 z}\right]-\exp\left[+i\frac{\pi ax}{\lambda_0 z}\right]}{\left(-\frac{2\pi ix}{\lambda_0 z}\right)}\cdot\frac{\exp\left[-i\frac{\pi yb}{\lambda_0 z}\right]-\exp\left[+i\frac{\pi yb}{\lambda_0 z}\right]}{\left(-\frac{2\pi ib}{\lambda_0 z}\right)}$$

$$=|a|\left(\frac{\sin\left[\frac{\pi ax}{\lambda_0 z}\right]}{\left(\frac{\pi ax}{\lambda_0 z}\right)}\right)\cdot|b|\left(\frac{\sin\left[\frac{\pi by}{\lambda_0 z}\right]}{\left(\frac{\pi by}{\lambda_0 z}\right)}\right)\equiv|ab|\,SINC\left[\frac{x}{\left(\frac{\lambda_0 z}{a}\right)},\frac{y}{\left(\frac{\lambda_0 z}{b}\right)}\right]$$

$$g\left[x,y\right]\propto(ab)^2\left(SINC\left[\frac{x}{\left(\frac{\lambda_0 z}{a}\right)},\frac{y}{\left(\frac{\lambda_0 z}{b}\right)}\right]\right)^2$$

Examples are shown in the figure:



Fraunhofer diffraction patterns from two slits of different widths; a wider slit has a narrower and taller Fraunhofer diffraction pattern.

Where the 2-D "SINC" function is defined as the orthogonal product of two 1-D SINC functions:

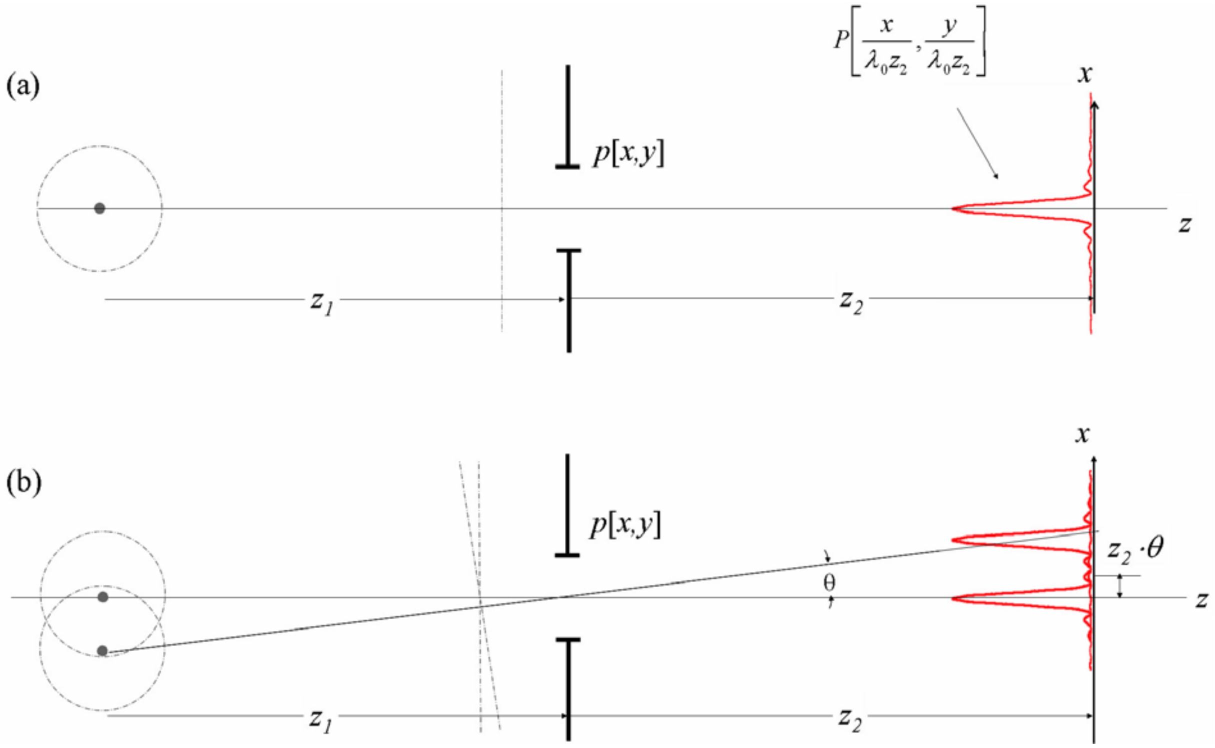
$$SINC\left[x,y\right]\equiv SINC\left[x\right]\cdot SINC\left[y\right]\equiv\frac{\sin\left[\pi x\right]}{\pi x}\cdot\frac{\sin\left[\pi y\right]}{\pi y}$$

which has the pattern shown in the figure.

1.1.4 Fraunhofer Diffraction in Optical Imaging Systems

A monochromatic point object located a long distance away from an imaging system produces a set of wavefronts that are "regularly" spaced (separated by the wavelength) and are approximately planar. The entrance pupil of the optical system (the image of the aperture stop in object space) collects a section of the plane wavefront and the optical elements convert it to a spherical wave that converges to an image "point." We can use the concept of Fraunhofer diffraction to define the "angular resolution" of the imaging system.

As an introduction, consider an optical system that consists of *only* the entrance pupil (which coincides with the aperture stop because no other optics are involved), as shown in the figure. The "pieces" of the object wavefronts that are collected by the stop will continue to propagate "downstream." If observed a long distance from the stop, the irradiance would be the Fraunhofer diffraction pattern of the stop; the smaller the stop, the larger the diffraction pattern and vice versa. Of course, the observed irradiance is the time average of the squared magnitude of the amplitude, and thus is nonnegative.



The system may be modeled in three stages:

1. Propagation from the input object $f[x,y]$ to the Fraunhofer diffraction region over the distance z_1,
2. multiplication by the (possibly complex-valued) transmittance function $t[x,y] = |t[x,y]|\ \exp[+i\Phi_t[x,y]]$ that specifies the aperture (or *pupil*), and
3. a second propagation over the distance z_2 into the Fraunhofer diffraction region (determined from the aperture).

To eliminate an awkward notation, we will substitute the notation $p[x,y]$ for the magnitude of the pupil function $|t[x,y]|$. In this example, we assume that the pupil has no phase component, so that $\Phi_t[x,y] = 0$, though solution of the more general case is straightforward. The 2-D input function $f[x, y; z = 0]$ is illuminated by a unit amplitude monochromatic plane wave with wavelength λ_0. The light propagates into the Fraunhofer diffraction region at a distance z_1, where the resulting amplitude pattern is:

$$E[x, y; z_1] = \frac{\mathcal{E}_0}{i\lambda_0 z_1} \exp\left[+2\pi i \frac{z_1}{\lambda_0}\right] \exp\left[+i\pi \frac{(x^2 + y^2)}{\lambda_0 z_1}\right] F\left[\frac{x}{\lambda_0 z_1}, \frac{y}{\lambda_0 z_1}\right]$$

This pattern illuminates the 2-D aperture function $p[x, y]$ and then propagates the distance z_2 into the Fraunhofer diffraction region (determined by the support of p). A second application produces the amplitude at the observation plane:

$$\begin{aligned} E[x, y; z_1 + z_2] = \mathcal{E}_0 &\left(\frac{1}{i\lambda_0 z_1} e^{+2\pi i \frac{z_1}{\lambda_0}} e^{+i\pi \frac{\left(x^2+y^2\right)}{\lambda_0 z_1}}\right) \left(\frac{1}{i\lambda_0 z_2} e^{+2\pi i \frac{z_2}{\lambda_0}} e^{+i\pi \frac{\left(x^2+y^2\right)}{\lambda_0 z_2}}\right) \\ &\cdot \mathcal{F}_2\left\{F\left[\frac{x}{\lambda_0 z_1}, \frac{y}{\lambda_0 z_1}\right] \cdot p[x, y]\right\}\Bigg|_{\xi = \frac{x}{\lambda_0 z_2}, \frac{y}{\lambda_0 z_2}} \end{aligned}$$

$$\begin{aligned} &= \mathcal{E}_0 \left(-\frac{1}{\lambda_0^2 z_1 z_2}\right) e^{+2\pi i \frac{z_1+z_2}{\lambda_0}} e^{+i\pi \frac{\left(x^2+y^2\right)}{\lambda_0}\left(\frac{1}{z_1}+\frac{1}{z_2}\right)} \\ &\qquad \cdot (\lambda_0 z_1)^2 \left(f[-\lambda_0 z_1 \xi, -\lambda_0 z_1 \eta] * P[\xi, \eta]\right)\big|_{\xi = \frac{x}{\lambda_0 z_2}, \frac{y}{\lambda_0 z_2}} \end{aligned}$$

$$\begin{aligned} &= \mathcal{E}_0 \left(-\frac{z_1}{z_2}\right) e^{+2\pi i \frac{z_1+z_2}{\lambda_0}} e^{+i\pi \frac{\left(x^2+y^2\right)}{\lambda_0}\left(\frac{1}{z_1}+\frac{1}{z_2}\right)} \left(f\left[\left(-\frac{z_1}{z_2}\right) x, \left(-\frac{z_1}{z_2}\right) y\right] * P\left[\frac{x}{\lambda_0 z_2}, \frac{y}{\lambda_0 z_2}\right]\right) \\ &= \frac{\mathcal{E}_0}{M_T} e^{+2\pi i \frac{z_1+z_2}{\lambda_0}} e^{+i\pi \frac{\left(x^2+y^2\right)}{\lambda_0}\left(\frac{1}{z_1}+\frac{1}{z_2}\right)} \left(f\left[\frac{x}{M_T}, \frac{y}{M_T}\right] * P\left[\frac{x}{\lambda_0 z_2}, \frac{y}{\lambda_0 z_2}\right]\right) \end{aligned}$$

where the theorems of the Fourier transform and the definition of the transverse magnification from geometrical optics, $M_T = -\frac{z_2}{z_1}$, have been used. Note that if the propagation distances z_1 and z_2 must both be positive in Fraunhofer diffraction, which requires that $M_T < 0$ and the image is "reversed."

The irradiance of the image is proportional to the squared magnitude of the amplitude:

$$|E[x, y; z_1 + z_2]|^2 = \left|\frac{\mathcal{E}_0}{M_T}\right|^2 \left|f\left[\frac{x}{M_T}, \frac{y}{M_T}\right] * P\left[\frac{x}{\lambda_0 z_2}, \frac{y}{\lambda_0 z_2}\right]\right|^2$$

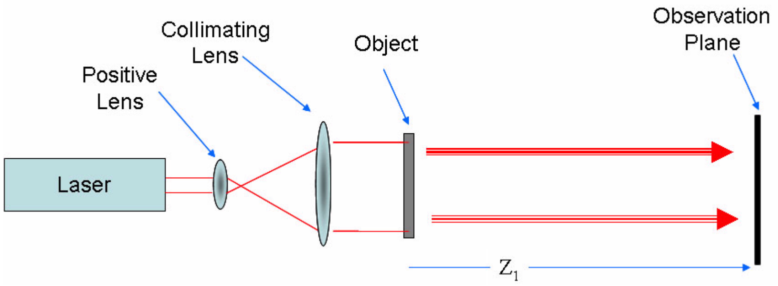
In words, the output amplitude created by this imaging "system" is the product of some constants, a quadratic-phase function of $[x, y]$, and the convolution of the input amplitude scaled by the transverse magnification and the scaled replica of the spectrum of the aperture function, $P\left[\frac{x}{\lambda_0 z_2}, \frac{y}{\lambda_0 z_2}\right]$. Since the output is the result of a convolution, we identify the spectrum as the impulse response of a shift-invariant convolution that is composed of two shift-variant Fourier transforms and multiplication by a quadratic-phase factor of $[x, y]$. This system does not satisfy the strict conditions for shift invariance because of the leading quadratic-phase factor and the fact that the input to the convolution is a scaled and reversed replica of the input to the system. That said, these details are often ignored to allow the process to considered to be shift invariant. We will revisit this conceptual imaging system after considering the mathematical models for optical elements.

1.2 Equipment:

1. He:Ne laser
2. microscope objective to expand the beam; larger power gives larger beam in shorter distance;
3. pinhole aperture to "clean up" the beam;
4. positive lens with diameter $d \cong 50\,\mathrm{mm}$ and focal length $f \lessapprox 600\,\mathrm{mm}$, to collimate the beam;
5. positve lens with diameter $d \cong 50\,\mathrm{mm}$ and focal length $f > 200\,\mathrm{mm}$, to compute the Fourier transform;
6. aluminum foil, needles, and razor blades to make your own objects for diffraction;
7. set of Metrologic transparencies;
8. digital camera to record diffraction patterns.

1.3 Procedure

1. Set up the experimental bench as in the figure with the observing screen close to the aperture (within a foot or so) to examine the results in the *Fresnel diffraction region*. Measure and record the relevant distances. A number of apertures are available for use, including single and multiple slits of different spacings, single and multiple circular apertures, needles (both tips and eyes), razor blades, *etc.*. In addition, aluminum foil and needles are availble to make your own apertures.



Apparatus for viewing Fresnel diffraction patterns (and Fraunhofer patterns if z_1 is sufficiently large).

   (a) Begin with a single slit, a square aperture, or a circular aperture. Note the form of the diffraction pattern. For example, sketch how its "brightness" changes with position and note the sizes and locations of any features. For a slit or circular aperture, you should note light and dark regions in the pattern; measure the positions of some maxima and minima (at least 5). Use the data to derive a scale of the pattern. Sketch the pattern noting the scale.

   (b) Repeat the previous step with a "wider" slit or aperture. Note the difference in the results.

   (c) Vary the distance between the screen and the diffracting object. Repeat measurements. What is the relation between the shange in distance and the change in scale of the pattern? Repeat for 5 different distances where the character of the pattern remains the same.

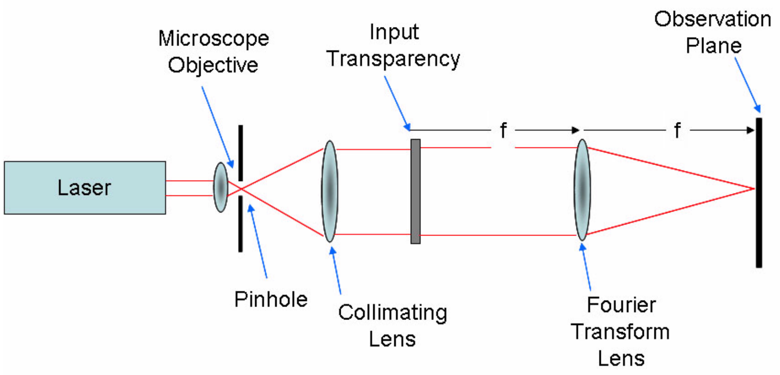


Figure 1: Apparatus for viewing Fraunhofer diffraction patterns

   (d) Repeat the procedure with a knife edge as the object. Sketch the pattern observed. You will see that the intensity distribution near the edge of the geometric shadow is not a sharp transition, but rather an undulatory pattern; a magnifying lens, microscope, or digital camera may be helpful to view the pattern, but BE SURE THAT THE LASER LIGHT HAS BEEN ATTENUATED SUFFICIENTLY.

2. Now observe the diffraction pattern far from the aperture (several feet away for a small aperture, a proportionally larger distance for a larger aperture) to examine *Fraunhofer diffraction*. You may "fold" the pattern with one or two mirrors or you may use a lens to "image" the pattern, i.e., to bring the image of the pattern created "a long distance away" much closer to the object. Whichever method you use, be sure to use the same setup for all measurements.

   (a) For an aperture of a known fixed (small) size, increase the distance to the observation plane as much as you can. Estimate the location of the transition between the Fresnel and Fraunhofer diffraction regions (this will certainly be ill-defined and "fuzzy"). Record and justify your measurement.

   (b) Add another lens to the system as shown below to "bring infinity closer"

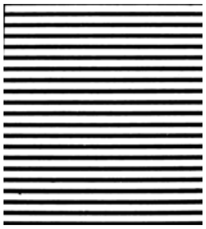
      i. Set up the "spatial filter" consisting of a microscope objective (or positive lens with a short focal length of about $f \cong 15\,\text{mm}$) and a "pinhole" that is large enough to pass the main beam but small enough to block stray light.

      ii. Add a positive lens after the input transparency with a large aperture and focal length of about $f \cong 150 - 200\,\text{mm}$. Ideally, the lens should be located one focal length after the transparency, though this is not critical for the current application. The observation plane should be located one focal length after the lens, which is the location of the smallest image.

   (c) Observe Fraunhofer diffraction from apertures of the same shape but different sizes. Measure the size of observable features and repeat this measurements using the other slits and then the other apertures. What is the influence of the physical dimension of the diffracting objects on the pattern?

   (d) Make some of your own patterns by punching holes in aluminum foil with a needle. For example, try to make two holes close together of about the same size. Observe the pattern. Repeat after enlarging these same holes and after creating new holes somewhate farther apart. Relate the observations to the laboratory on interference by division of wavefront.

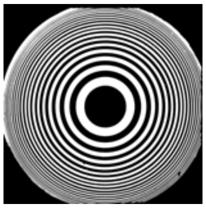
   (e) Repeat the procedure using a periodic structure (diffraction grid or grating) as the object. Among these, sketch the diffraction patterns of specific transparencies available in the Metrologic set, including #4 (parallel lines with wide spacing), #5 (parallel lines with medium spacing), #6 (parallel lines with narrow spacing, #7 (concentric circles with wide spacing), #8 (concentric circles with medium spacing), #9 (concentric circles with narrow spacing), and “crossed” gratings #10 (wide spacing), #11 (medium spacing), and #12 (narrow spacing).



(f) Now overlay a periodic structure (grid) with a circular aperture and observe the pattern. The overlaying of the two slides produces the product of the two patterns (also called the *modulation* of one pattern by the other).

(g) Examine the image and diffraction pattern of the transparency *Albert* (Metrologic slide #18). Note the features of the diffraction pattern and relate them to the features of the transparency.

(h) Examine the pattern generated by a Fresnel Zone Plate (Metrologic slide #13) at different distances. The FZP is a circular grating whose spacing decreases with increasing distance from the center. Sketch a side view of the FZP and indicate the diffraction angle for light incident at different distances from the center of symmetry. You might also overlap another transparency (such as a circular aperture) and the FZP and record the result. I guarantee that this result will not resemble that of part d.



(i) If time permits, you can also find the diffraction patterns of other objects, such as the tip and/or the eye of the needle.

2 Questions

1. This experiment demonstrates that interaction of light with an obstruction will spread the light. For example, consider Fresnel diffraction of two identical small circular apertures that are separated by a distance d. How will diffraction affect the ability to distinguish the two sources? Comment on the result as lens diameter d is made smaller.
2. The Fresnel Zone Plate (Metrologic slide #13) may be viewed as a circularly symmetric grating with variable period that decreases in proportion to the radial distance from the center. It is possible to use the FZP as an imaging element (i.e., as a lens). Use the model of diffraction from a constant-period grating to describe how the FZP may be used to “focus” light in an optical imaging system. This may be useful for wavelengths (such as x rays) where imaging lenses do not exist.