Chapter 6
Fresnel Diffraction in Vacuum

The goal of this chapter is to develop methods for modeling near-field optical-wave propagation with high fidelity and some flexibility, which is considerably more challenging than for far-field propagation. This chapter uses the same coordinate convention as in Fig. 1.2. It begins with a discussion of different forms of the Fresnel diffraction integral. These different forms can be numerically evaluated in different ways, each with benefits and drawbacks. Then, to emphasize the different mathematical operations in the notation, operators are introduced that are used throughout Chs. 6–8. The rest of this chapter develops basic algorithms for wave propagation in vacuum and other simulation details.

The quadratic phase factor inside the Fresnel diffraction integral is not bandlimited, so it poses some challenges related to sampling. There are two different ways to evaluate the integral: as a single FT or as a convolution. This chapter develops both basic methods as well as more sophisticated versions that provide some flexibility. There are different types of flexibility that one might need. For example, Delen and Hooker present a method that is particularly useful for simulating propagation in integrated optical components. Because the interfaces in these components are often slanted or offset and the angles are not always paraxial, they developed a Rayleigh-Summerfeld propagation method that can handle propagation between arbitrarily oriented planes with good accuracy.[28,29]

In contrast, the applications discussed in this book involve parallel source and observation planes, and the paraxial approximation is a very good one. When long propagation distances are involved, beams can spread to be much larger than their original size. Accordingly, some algorithms discussed in this chapter provide the user with the flexibility to choose the scaling between the observation- and source-plane grid spacings. Many authors have presented algorithms with this ability including Tyler and Fried,[30] Roberts,[31] Coles,[32] Rubio,[33] Deng *et al.*,[34] Coy,[35] Rydberg and Bengtsson,[36] and Voelz and Roggemann.[37] Most of these methods are mathematically equivalent to each other. However, one unique algorithm was presented by Coles[32] and later augmented by Rubio[33] in which a diverging spherical coordinate system was used by an angular grid with constant angular grid spacing. This was done specifically because the source was a point source, which naturally diverges spherically. Rubio augmented this basic concept to allow for very long

propagation distances. When the grid grows too large to adequately sample the field, Rubio's method is to extract a central portion and interpolate it to a finer grid.

In this chapter, two flexible propagation methods are presented. The first uses two steps of evaluating the Fresnel diffraction integral, with the grid spacings adjusted by the distances of the two propagations. The second method uses some algebraic manipulation of the convolution form of the Fresnel diffraction integral. The manipulation introduces a free parameter that directly sets the observation-plane grid spacing.

6.1 Different Forms of the Fresnel Diffraction Integral

We start with the Fresnel diffraction integral, which is repeated here for convenience:

$$U(x_2, y_2) = \frac{e^{ik\Delta z}}{i\lambda\Delta z} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(x_1, y_1)\, e^{i\frac{k}{2\Delta z}\left[(x_2-x_1)^2+(y_2-y_1)^2\right]}\, dx_1\, dy_1. \quad (6.1)$$

Also, we define spatial and spatial-frequency vectors

$$\mathbf{r}_1 = x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}} \quad (6.2)$$

$$\mathbf{r}_2 = x_2\hat{\mathbf{i}} + y_2\hat{\mathbf{j}} \quad (6.3)$$

$$\mathbf{f}_1 = f_{x1}\hat{\mathbf{i}} + f_{y1}\hat{\mathbf{j}}, \quad (6.4)$$

where $\mathbf{r}_1$ is in the source plane, and $\mathbf{r}_2$ is in the observation plane. This is used throughout the chapter. Table 6.1 summarizes these quantities and others that are important to this development.

We want to use the Fresnel diffraction integral to compute the observation-plane optical field from knowledge of the source-plane field. Sections 6.3 and 6.4 deal with numerically evaluating this equation. There are two forms of Eq. (6.1) that are used for numerical evaluation. The first comes about by expanding the squared terms in the exponential and factoring portions out of the integral. This yields

$$U(x_2, y_2) = \frac{e^{ik\Delta z}}{i\lambda\Delta z}\, e^{i\frac{k}{2\Delta z}\left(x_2^2+y_2^2\right)} \times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(x_1, y_1)\, e^{i\frac{k}{2\Delta z}\left(x_1^2+y_1^2\right)} e^{-i\frac{2\pi}{\lambda\Delta z}(x_2x_1+y_2y_1)}\, dx_1\, dy_1, \quad (6.5)$$

which can be evaluated as an FT as discussed in Sec. 6.3. The second form of Eq. (6.1) comes about by noting that it is a convolution of the source-plane field with the free-space amplitude spread function so that

$$U(x_2, y_2) = U(x_1, y_1) \otimes \left[\frac{e^{ik\Delta z}}{i\lambda\Delta z}\, e^{i\frac{k}{2\Delta z}\left(x_1^2+y_1^2\right)}\right]. \quad (6.6)$$

Then, the convolution theorem is used to evaluate Eq. (6.6) via two FTs.

Table 6.1 Definition of symbols for Fresnel propagation.

symbol	meaning
$\mathbf{r}_1 = (x_1, y_1)$	source-plane coordinates
$\mathbf{r}_2 = (x_2, y_2)$	observation-plane coordinates
δ_1	grid spacing in source plane
δ_2	grid spacing in observation plane
$\mathbf{f}_1 = (f_{x1}, f_{y1})$	spatial-frequency of source plane
δ_{f1}	grid spacing in source-plane spatial frequency
z_1	location of source plane along the optical axis
z_2	location of observation plane along the optical axis
Δz	distance between source plane and observation plane
m	scaling factor from source plane to observation plane

6.2 Operator Notation

Operator notation is useful in Fresnel diffraction computations for writing the equations compactly without explicit integral notation. Using operators places the emphasis on operations that are taking place. The notation used here is adapted from that described by Nazarathy and Shamir,[38] who also incorporated it with ray matrices to describe diffraction through optical systems.[39] The key difference is that we specify the domains in which they operate. These operators are defined by:

$$\mathcal{Q}\left[c, \mathbf{r}\right]\left\{U\left(\mathbf{r}\right)\right\} \equiv e^{i\frac{k}{2}c|\mathbf{r}|^2} U\left(\mathbf{r}\right) \tag{6.7}$$

$$\mathcal{V}\left[b, \mathbf{r}\right]\left\{U\left(\mathbf{r}\right)\right\} \equiv b\, U\left(b\mathbf{r}\right) \tag{6.8}$$

$$\mathcal{F}\left[\mathbf{r}, \mathbf{f}\right]\left\{U\left(\mathbf{r}\right)\right\} \equiv \int_{-\infty}^{\infty} U\left(\mathbf{r}\right) e^{-i2\pi\mathbf{f}\cdot\mathbf{r}}\, d\mathbf{r} \tag{6.9}$$

$$\mathcal{F}^{-1}\left[\mathbf{f}, \mathbf{r}\right]\left\{U\left(\mathbf{f}\right)\right\} \equiv \int_{-\infty}^{\infty} U\left(\mathbf{f}\right) e^{i2\pi\mathbf{f}\cdot\mathbf{r}}\, d\mathbf{f} \tag{6.10}$$

$$\mathcal{R}\left[d, \mathbf{r}_1, \mathbf{r}_2\right]\left\{U\left(\mathbf{r}_1\right)\right\} \equiv \frac{1}{i\lambda d} \int_{-\infty}^{\infty} U\left(\mathbf{r}_1\right) e^{i\frac{k}{2d}|\mathbf{r}_2 - \mathbf{r}_1|^2}\, d\mathbf{r}_1. \tag{6.11}$$

The operators' parameters are given in square brackets, and the operand is given in curly braces. Note that in Eqs. (6.9) and (6.10), the domain of the operand is listed as the first parameter, and the domain of the result is listed as the second parameter. See Refs. 38 and 39 for relations betweens these operators. Finally, we define one more quadratic-phase exponential operator

$$\mathcal{Q}_2\left[d, \mathbf{r}\right]\left\{U\left(\mathbf{r}\right)\right\} \equiv e^{i\pi^2 \frac{2d}{k}|\mathbf{r}|^2} U\left(\mathbf{r}\right). \tag{6.12}$$

The operator $\mathcal{Q}_2[d, \mathbf{r}]$ is not defined by Nazarathy and Shamir. In fact, it can be written in terms of the operator $\mathcal{Q}$ as

$$\mathcal{Q}_2[d, \mathbf{r}] = \mathcal{Q}\left[\frac{4\pi^2}{k^2} d, \mathbf{r}\right]. \tag{6.13}$$

However, it is just a convenient definition for use in Sec. 6.4.

6.3 Fresnel-Integral Computation

This section describes two methods of implementing the Fresnel diffraction integral in the form of Eq. (6.5). The first method evaluates this integral once as a single FT, which is the most straightforward. This method is desirable because of its computational efficiency. The second method evaluates the Fresnel integral twice, which adds some flexibility in the grid spacing at the cost of performing a second FT.

6.3.1 One-step propagation

Figure 1.2 shows the geometry of propagation from the source plane to the observation plane. The Fresnel integral can be used via Eq. (6.5) to compute the observation-plane field $U(x_2, y_2)$ directly, given the source-plane field $U(x_1, y_1)$ and the propagation geometry. We write Eq. (6.5) in operator notation as

$$U(\mathbf{r}_2) = \mathcal{R}[\Delta z, \mathbf{r}_1, \mathbf{r}_2]\{U(\mathbf{r}_1)\} \tag{6.14}$$

$$= \mathcal{Q}\left[\frac{1}{\Delta z}, \mathbf{r}_2\right] \mathcal{V}\left[\frac{1}{\lambda \Delta z}, \mathbf{r}_2\right] \mathcal{F}[\mathbf{r}_1, \mathbf{f}_1]\, \mathcal{Q}\left[\frac{1}{\Delta z}, \mathbf{r}_1\right]\{U(\mathbf{r}_1)\}. \tag{6.15}$$

The order of operation is right to left. In general, these operators do not commute; only certain combinations commute. It is clear that the observation-plane field is computed by (reading right to left) multiplying the source field by a quadratic phase ($\mathcal{Q}$), Fourier transforming ($\mathcal{F}$), scaling by a constant [$\mathcal{V}$ transforms from spatial frequency to spatial coordinates with $\mathbf{f}_1 = \mathbf{r}_2/(\lambda \Delta z)$], and multiplying by another quadratic phase factor ($\mathcal{Q}$). An intuitive explanation is that propagation can be represented as an FT between confocal spheres centered at the source and observation planes. The spheres' common radius of curvature is Δz.

To evaluate the Fresnel integral on a computer, again we must use a sampled version of the source-plane optical field $U(\mathbf{r}_1)$. Let the spacing in the source plane be δ_1. As before, the spacing in the frequency domain is $\delta_{f1} = 1/(N\delta_1)$, so then the spacing in the observation plane is

$$\delta_2 = \frac{\lambda \Delta z}{N \delta_1}. \tag{6.16}$$

Listing 6.1 gives the MATLAB function `one_step_prop` that numerically evaluates Eq. (6.5).

Listing 6.1 Code for evaluating the Fresnel diffraction integral in MATLAB using a single step.

```
function [x2 y2 Uout] ...
    = one_step_prop(Uin, wvl, d1, Dz)
% function [x2 y2 Uout] ...
%     = one_step_prop(Uin, wvl, d1, Dz)

    N = size(Uin, 1);   % assume square grid
    k = 2*pi/wvl;    % optical wavevector
    % source-plane coordinates
    [x1 y1] = meshgrid((-N/2 : 1 : N/2 - 1) * d1);
    % observation-plane coordinates
    [x2 y2] = meshgrid((-N/2 : N/2-1) / (N*d1)*wvl*Dz);
    % evaluate the Fresnel-Kirchhoff integral
    Uout = 1 / (i*wvl*Dz) ...
        .* exp(i * k/(2*Dz) * (x2.^2 + y2.^2)) ...
        .* ft2(Uin .* exp(i * k/(2*Dz) ...
        * (x1.^2 + y1.^2)), d1);
```

Listing 6.2 gives example usage of `one_step_prop` for a square aperture. Figure 6.1 shows the numerical result along with the analytic result, and it is clear

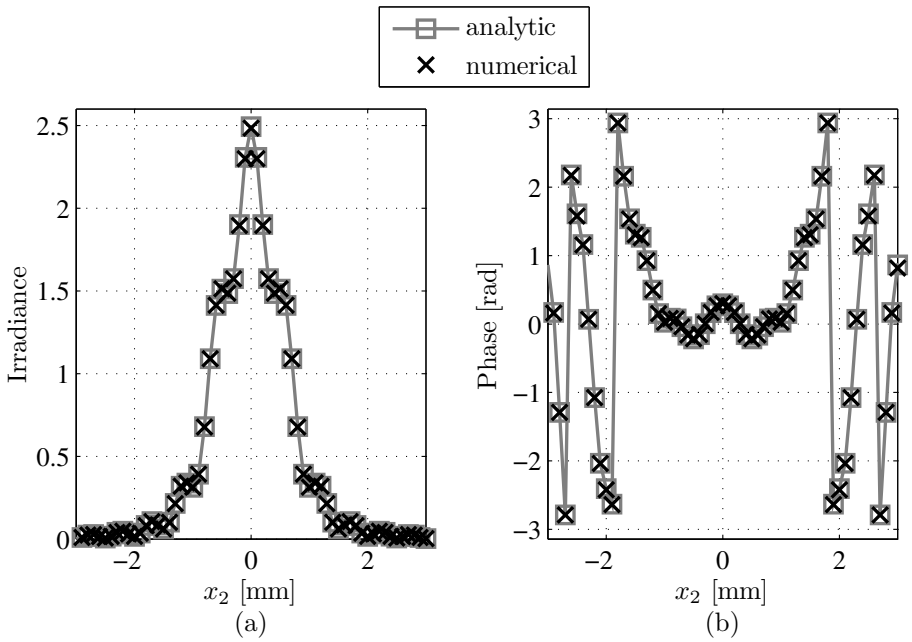

Figure 6.1 Fresnel diffraction from a square aperture, simulation and analytic: (a) observation-plane irradiance and (b) observation-plane phase.

Listing 6.2 Example of evaluating the Fresnel diffraction integral in MATLAB using a single step.

```
% example_square_prop_one_step.m

N = 1024;     % number of grid points per side
L = 1e-2;    % total size of the grid [m]
delta1 = L / N;  % grid spacing [m]
D = 2e-3;    % diameter of the aperture [m]
wvl = 1e-6;  % optical wavelength [m]
k = 2*pi / wvl;
Dz = 1;      % propagation distance [m]

[x1 y1] = meshgrid((-N/2 : N/2-1) * delta1);
ap = rect(x1/D) .* rect(y1/D);
[x2 y2 Uout] = one_step_prop(ap, wvl, delta1, Dz);

% analytic result for y2=0 slice
Uout_an ...
    = fresnel_prop_square_ap(x2(N/2+1,:), 0, D, wvl, Dz);
```

that the comparison is very close.

Obviously, we have no control over spacing in the final grid without changing the geometry because Eq. (6.16) gives a fixed grid spacing in the observation plane. What if we have an application where the fixed value of δ_2 does not sample the observation-plane field adequately? We could obtain finer sampling in the observation plane by increasing N. Typically, we would prefer not to increase N due to the longer execution time of the simulation, though.

6.3.2 Two-step propagation

To choose the observation-plane grid spacing, we must introduce a new scaling parameter $m = \delta_2/\delta_1$. For one-step propagation [compute $U(x_2, y_2)$ directly from $U(x_1, y_1)$], there is little freedom to choose m as indicated in Eq. (6.16). Typically, λ and Δz are fixed for a given problem, so N and δ_1 must be adjusted to select a desired value of m. There must be a trade-off between the source and observation grids. A finer source grid produces a coarser observation grid and vice-versa. We could adjust N to help, but there is a practical limit to the number of grid points that can be simulated, and increasing N increases the simulation's execution time, which is typically not desirable.

Coy[35] and Rydberg and Bengtsson[36] presented a method that has more flexibility in selecting the grids. In this method, $U(x_1, y_1)$ propagates from the source plane at z_1 to an intermediate plane located at z_{1a} and then propagates to the observation plane at z_2, so that we can choose z_{1a} such that m (equivalently δ_2) has

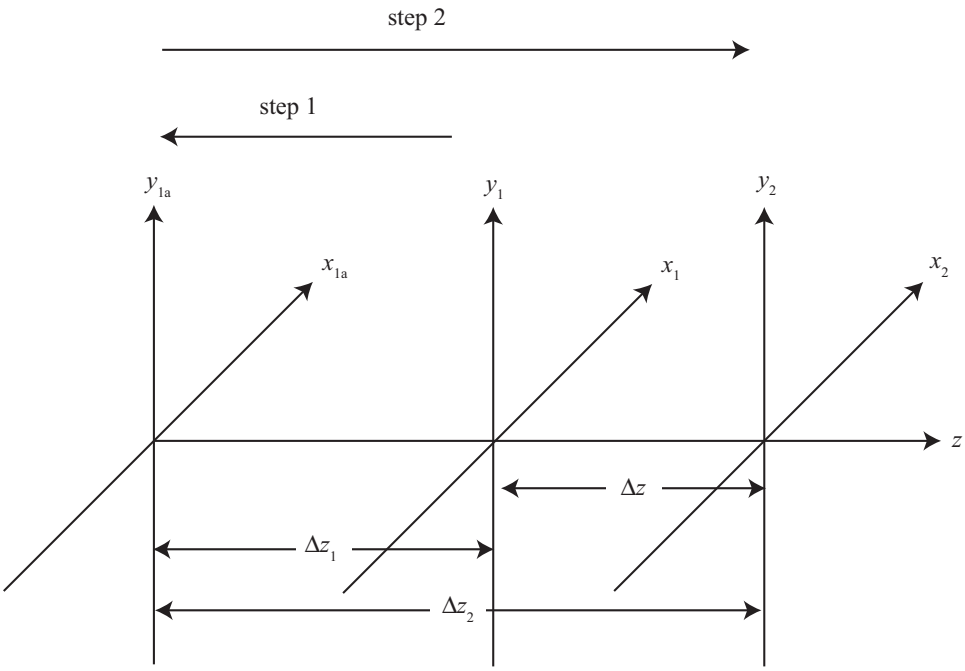

Figure 6.2 Two-step propagation geometry in which the intermediate plane is not between the source and observation planes.

the desired value. The following development follows Rydberg and Bengtsson's algorithm description with Coy's analysis of the grid spacings.

This is called two-step propagation as specified below. To keep the notation clear, the following definitions are still used: the source plane is at $z = z_1$ ($\mathbf{r}_1$ coordinates), and the observation plane is at $z = z_2$ ($\mathbf{r}_2$ coordinates) with $\Delta z = z_2 - z_1$ and scaling parameter of $m = \delta_2/\delta_1$. We define the intermediate plane at $z = z_{1a}$ [$\mathbf{r}_{1a} = (x_{1a}, y_{1a})$ coordinates] such that the distance of the first propagation is $\Delta z_1 = z_{1a} - z_1$ and the distance of the second is $\Delta z_2 = z_2 - z_{1a}$. As discussed below, there are two possible intermediate planes that yield a given scaling parameter after the two-step propagation. These two different geometries are shown in Figs. 6.2 and 6.3. In one case, the intermediate plane is far from the source and observation planes. In the other, the intermediate plane is between the source and observation planes.

In operator notation, two steps of Fresnel-integral propagation are given by

$$U(\mathbf{r}_2) = \mathcal{R}[\Delta z_2, \mathbf{r}_{1a}, \mathbf{r}_2]\, \mathcal{R}[\Delta z_1, \mathbf{r}_1, \mathbf{r}_{1a}]\, \{U(\mathbf{r}_1)\} \tag{6.17}$$

$$= \mathcal{Q}\left[\frac{1}{\Delta z_2}, \mathbf{r}\right] \mathcal{V}\left[\frac{1}{\lambda \Delta z_2}\right] \mathcal{F}[\mathbf{r}_2, \mathbf{f}_{1a}]\, \mathcal{Q}\left[\frac{1}{\Delta z_2}, \mathbf{r}_{1a}\right] \tag{6.18}$$

$$\times\, \mathcal{Q}\left[\frac{1}{\Delta z_1}, \mathbf{r}_{1a}\right] \mathcal{V}\left[\frac{1}{\lambda \Delta z_1}\right] \mathcal{F}[\mathbf{r}_1, \mathbf{f}_1]\, \mathcal{Q}\left[\frac{1}{\Delta z_1}, \mathbf{r}_1\right] \{U(\mathbf{r}_1)\}.$$

If we examine the spacings δ_{1a} in the intermediate plane and δ_2 in the observation

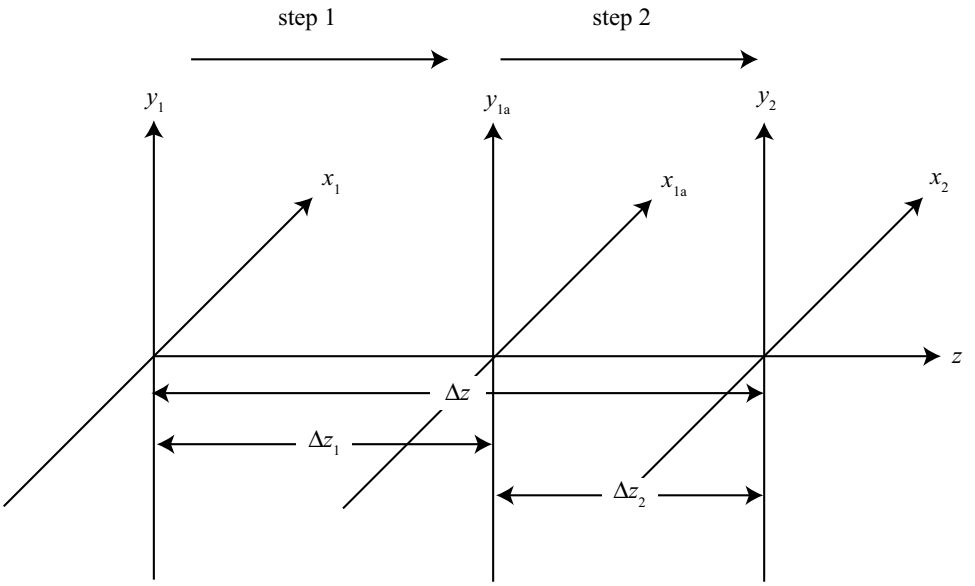

Figure 6.3 Two-step propagation geometry in which the intermediate plane is between the source and observation planes.

plane, we find

$$\delta_{1a} = \frac{\lambda\,|\Delta z_1|}{N\delta_1} \qquad \text{with} \qquad \Delta z_1 = z_{1a} - z_1 \tag{6.19}$$

$$\delta_2 = \frac{\lambda\,|\Delta z_2|}{N\delta_{1a}} \tag{6.20}$$

$$= \frac{\lambda\,|\Delta z_2|}{N\left(\frac{\lambda|\Delta z_1|}{N\delta_1}\right)} \tag{6.21}$$

$$= \left|\frac{\Delta z_2}{\Delta z_1}\right| \delta_1 \tag{6.22}$$

$$= m\delta_1, \tag{6.23}$$

which is expected given the definition of scaling parameter $m = \delta_2/\delta_1$.

Thus, a choice of m (which directly sets the sizes of the grids) defines the location of the intermediate plane, i.e., from above

$$m = \left|\frac{z_2 - z_{1a}}{z_{1a} - z_1}\right| = \left|\frac{\Delta z_2}{\Delta z_1}\right|, \tag{6.24}$$

which has solutions for the choice of z_{1a} (constrained such that $\Delta z_1 + \Delta z_2 = \Delta z$) given by

$$\Delta z_1 = z_{1a} - z_1 = \Delta z \left(\frac{1}{1 \pm m}\right) \tag{6.25}$$

Table 6.2 Examples of scaling parameter values for two-step Fresnel integral computation.

m	$\Delta z_1^+/\Delta z$	$\Delta z_2^+/\Delta z$	$\Delta z_1^-/\Delta z$	$\Delta z_2^-/\Delta z$
	$\frac{1}{(1+m)}$	$\frac{m}{(1+m)}$	$\frac{1}{(1-m)}$	$\frac{-m}{(1-m)}$
2	$1/3$	$2/3$	-1	2
1	$1/2$	$1/2$	$\pm\infty$	$\mp\infty$
$1/2$	$2/3$	$1/3$	2	-1

$$z_{1a} = z_1 + \Delta z \left(\frac{1}{1 \pm m} \right) \tag{6.26}$$

$$\Delta z_2 = z_2 - z_{1a} = \Delta z \left(\frac{\pm m}{1 \pm m} \right) \tag{6.27}$$

$$z_{1a} = z_2 - \Delta z \left(\frac{\pm m}{1 \pm m} \right) \tag{6.28}$$

$$z_{1a} = z_2 + \Delta z \left(\frac{\mp m}{1 \pm m} \right). \tag{6.29}$$

This has a very simple proof:

$$\left| \frac{\Delta z_2}{\Delta z_1} \right| = \left| \frac{\Delta z \left(\frac{\pm m}{1 \pm m} \right)}{\Delta z \left(\frac{1}{1 \pm m} \right)} \right| = |\pm m| = m. \tag{6.30}$$

Table 6.2 gives some example values of m with the corresponding intermediate plane locations. The Δz_1^- and Δz_2^- columns correspond to Fig. 6.2, and the Δz_1^+ and Δz_2^+ columns correspond to Fig. 6.3. Note that for unit scaling parameter, the intermediate plane is either located halfway between the source and observation planes or infinitely far away.

Listing 6.3 gives the MATLAB function `two_step_prop` that numerically evaluates Eq. (6.18). Listing 6.4 shows example usage by simply repeating the previous MATLAB example but with the two-step propagation algorithm. Figure 6.4 shows the numerical and analytic results. Note that the simulation results are identical to the analytic results again.

6.4 Angular-Spectrum Propagation

This section evaluates the convolution form of the Fresnel diffraction integral given in Eq. (6.6). We can rewrite it using the convolution theorem in operator notation as

$$U(\mathbf{r}_2) = \mathcal{F}^{-1}[\mathbf{r}_2, \mathbf{f}_1] H(\mathbf{f}_1) \mathcal{F}[\mathbf{f}_1, \mathbf{r}_1] \{U(\mathbf{r}_1)\}, \tag{6.31}$$

where $H(\mathbf{f})$ is the transfer function of free-space propagation given by

$$H(\mathbf{f}_1) = e^{ik\Delta z} e^{-i\pi\lambda\Delta z\left(f_{x1}^2 + f_{y1}^2\right)}. \tag{6.32}$$

Listing 6.3 Code for evaluating the Fresnel diffraction integral in MATLAB using two-step propagation.

```
function [x2 y2 Uout] ...
    = two_step_prop(Uin, wvl, d1, d2, Dz)
% function [x2 y2 Uout] ...
%     = two_step_prop(Uin, wvl, d1, d2, Dz)

    N = size(Uin, 1);   % number of grid points
    k = 2*pi/wvl;    % optical wavevector
    % source-plane coordinates
    [x1 y1] = meshgrid((-N/2 : 1 : N/2 - 1) * d1);
    % magnification
    m = d2/d1;
    % intermediate plane
    Dz1 = Dz / (1 - m);  % propagation distance
    d1a = wvl * abs(Dz1) / (N * d1);    % coordinates
    [x1a y1a] = meshgrid((-N/2 : N/2-1) * d1a);
    % evaluate the Fresnel-Kirchhoff integral
    Uitm = 1 / (i*wvl*Dz1) ...
        .* exp(i*k/(2*Dz1) * (x1a.^2+y1a.^2)) ...
        .* ft2(Uin .* exp(i * k/(2*Dz1) ...
        * (x1.^2 + y1.^2)), d1);
    % observation plane
    Dz2 = Dz - Dz1;  % propagation distance
    % coordinates
    [x2 y2] = meshgrid((-N/2 : N/2-1) * d2);
    % evaluate the Fresnel diffraction integral
    Uout = 1 / (i*wvl*Dz2) ...
        .* exp(i*k/(2*Dz2) * (x2.^2+y2.^2)) ...
        .* ft2(Uitm .* exp(i * k/(2*Dz2) ...
        * (x1a.^2 + y1a.^2)), d1a);
```

Equation (6.31) is known as the angular-spectrum form of the Fresnel diffraction integral, and it has been discussed and applied by many authors specifically for numerical evaluation.[28,31,32,37,40–44] Section 3.1 in this book already covers discrete convolution, which could be applicable here, but we cannot simply use the `myconv2` function from Sec. 3.1 as-is. If we did, we would have no control over the grid spacing δ_2 in the observation plane. We would be stuck with $\delta_1 = \delta_2$, corresponding to $m = 1$.

To introduce the scaling parameter m, we must go back to Eq. (6.1) and rewrite

Listing 6.4 Example of evaluating the Fresnel diffraction integral in MATLAB using two-step propagation.

```
% example_square_prop_two_step.m

N = 1024;     % number of grid points per side
L = 1e-2;    % total size of the grid [m]
delta1 = L / N;  % grid spacing [m]
D = 2e-3;    % diameter of the aperture [m]
wvl = 1e-6;  % optical wavelength [m]
k = 2*pi / wvl;
Dz = 1;      % propagation distance [m]

[x1 y1] = meshgrid((-N/2 : N/2-1) * delta1);
ap = rect(x1/D) .* rect(y1/D);
delta2 = wvl * Dz / (N*delta1);
[x2 y2 Uout] = two_step_prop(ap, wvl, delta1, delta2, Dz);

% analytic result for y2=0 slice
Uout_an ...
    = fresnel_prop_square_ap(x2(N/2+1,:), 0, D, wvl, Dz);
```

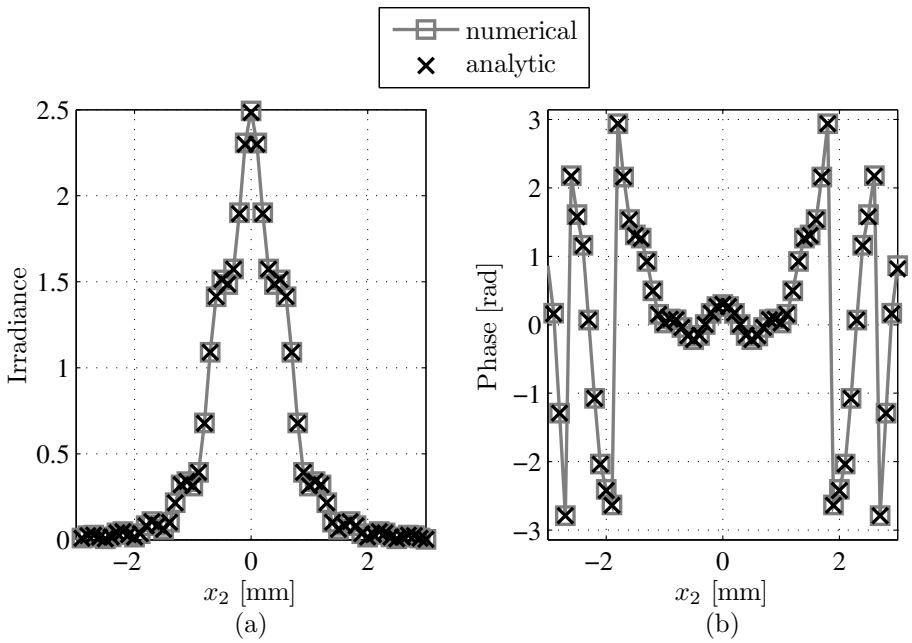

Figure 6.4 Fresnel diffraction from a square aperture, two-step simulation and analytic: (a) observation-plane irradiance and (b) observation-plane phase.

it using $\mathbf{r}_1$ and $\mathbf{r}_2$ as

$$U(\mathbf{r}_2) = \frac{1}{i\lambda\Delta z}\int_{-\infty}^{\infty} U(\mathbf{r}_1)\, e^{i\frac{k}{2\Delta z}|\mathbf{r}_2-\mathbf{r}_1|^2}\, d\mathbf{r}_1. \tag{6.33}$$

Tyler and Fried[30] and Roberts[31] are the only authors who discuss this scaling factor. Following their approach, we manipulate the exponential to introduce m:

$$|\mathbf{r}_2-\mathbf{r}_1|^2 = r_2^2 - 2\mathbf{r}_2\cdot\mathbf{r}_1 + r_1^2 \tag{6.34}$$

$$= \left(r_2^2 + \frac{r_2^2}{m} - \frac{r_2^2}{m}\right) - 2\mathbf{r}_2\cdot\mathbf{r}_1 + \left(r_1^2 + mr_1^2 - mr_1^2\right) \tag{6.35}$$

$$= \frac{r_2^2}{m} + \left(1-\frac{1}{m}\right)r_2^2 - 2\mathbf{r}_2\cdot\mathbf{r}_1 + \left[mr_1^2 + (1-m)\,r_1^2\right] \tag{6.36}$$

$$= m\left[\left(\frac{r_2}{m}\right)^2 - 2\left(\frac{\mathbf{r}_2}{m}\right)\cdot\mathbf{r}_1 + r_1^2\right] + \left(1-\frac{1}{m}\right)r_2^2 + (1-m)\,r_1^2 \tag{6.37}$$

$$= m\left|\frac{\mathbf{r}_2}{m} - \mathbf{r}_1\right|^2 - \left(\frac{1-m}{m}\right)r_2^2 + (1-m)\,r_1^2. \tag{6.38}$$

Then, we can substitute it back into Eq. (6.33) to get

$$U(\mathbf{r}_2) = \frac{1}{i\lambda\Delta z}\int_{-\infty}^{\infty} U(\mathbf{r}_1)\, e^{i\frac{k}{2\Delta z}\left[m\left|\frac{\mathbf{r}_2}{m}-\mathbf{r}_1\right|^2 - \left(\frac{1-m}{m}\right)r_2^2 + (1-m)r_1^2\right]}\, d\mathbf{r}_1 \tag{6.39}$$

$$= \frac{e^{-i\frac{k}{2\Delta z}\left(\frac{1-m}{m}\right)r_2^2}}{i\lambda\Delta z}\int_{-\infty}^{\infty} U(\mathbf{r}_1)\, e^{i\frac{k}{2\Delta z}(1-m)r_1^2} e^{i\frac{km}{2\Delta z}\left|\frac{\mathbf{r}_2}{m}-\mathbf{r}_1\right|^2}\, d\mathbf{r}_1. \tag{6.40}$$

We start on the path back to obtaining a convolution integral by defining

$$U''(\mathbf{r}_1) \equiv \frac{1}{m}U(\mathbf{r}_1)\, e^{i\frac{k}{2\Delta z}(1-m)r_1^2}, \tag{6.41}$$

and substitute it into Eq. (6.40) to get

$$U(\mathbf{r}_2) = \frac{e^{-i\frac{k}{2\Delta z}\left(\frac{1-m}{m}\right)r_2^2}}{i\lambda\Delta z}\int_{-\infty}^{\infty} mU''(\mathbf{r}_1)\, e^{i\frac{km}{2\Delta z}\left|\frac{\mathbf{r}_2}{m}-\mathbf{r}_1\right|^2}\, d\mathbf{r}_1. \tag{6.42}$$

Then, defining the scaled coordinate and distance

$$\mathbf{r}_2' = \frac{\mathbf{r}_2}{m} \tag{6.43}$$

$$\Delta z' = \frac{\Delta z}{m}, \tag{6.44}$$

we obtain

$$U\left(m\mathbf{r}_2'\right) = \frac{e^{-i\frac{k}{2\Delta z'}(1-m)\left(r_2'\right)^2}}{i\lambda\Delta z'} \int_{-\infty}^{\infty} U''\left(\mathbf{r}_1\right) e^{i\frac{k}{2\Delta z'}\left|\mathbf{r}_2'-\mathbf{r}_1\right|^2} d\mathbf{r}_1. \tag{6.45}$$

Finally, this is in the form of a convolution so that

$$U\left(m\mathbf{r}_2'\right) = e^{-i\frac{k}{2\Delta z'}(1-m)\left(r_2'\right)^2} \int_{-\infty}^{\infty} U''\left(\mathbf{r}_1\right) h\left(\mathbf{r}_2'-\mathbf{r}_1\right)\, d\mathbf{r}_1, \tag{6.46}$$

$$\text{with}\;\; h\left(\mathbf{r}_1\right) = \frac{1}{i\lambda\Delta z'} e^{i\frac{k}{2\Delta z'}r_1^2}. \tag{6.47}$$

Once again, propagation can be treated as a linear system with a known impulse response (amplitude spread function). The FT of the impulse response is the amplitude transfer function, given by

$$\mathcal{F}\left[\mathbf{r}_1,\mathbf{f}_1\right] h\left(\mathbf{r}_1\right) = H\left(\mathbf{f}_1\right) \tag{6.48}$$

$$= e^{-i\pi\lambda\Delta z' f_1^2}. \tag{6.49}$$

At this point, we could evaluate Eq. (6.46) numerically using `myconv2`. However, using the convolution theorem and substituting back to original coordinates allows us to keep all of the details of this algorithm manifest and thereby make some simplifications in later chapters. Applying the convolution theorem leads to

$$\begin{aligned}
U'\left(m\mathbf{r}_2'\right) &= \mathcal{F}^{-1}\left[\mathbf{f}_1,\mathbf{r}_2'\right] e^{-i\pi\lambda\Delta z' f_1^2} \mathcal{F}\left[\mathbf{r}_1,\mathbf{f}_1\right]\left\{U''\left(\mathbf{r}_1\right)\right\} \\
U'\left(\mathbf{r}_2\right) &= \mathcal{F}^{-1}\left[\mathbf{f}_1,\frac{\mathbf{r}_2}{m}\right] e^{-i\pi\lambda\frac{\Delta z}{m} f_1^2} \mathcal{F}\left[\mathbf{r}_1,\mathbf{f}_1\right]\left\{U''\left(\mathbf{r}_1\right)\right\} \\
U\left(\mathbf{r}_2\right) &= e^{-i\frac{k}{2\Delta z} m(1-m)\left(\frac{\mathbf{r}_2}{m}\right)^2} \mathcal{F}^{-1}\left[\mathbf{f}_1,\frac{\mathbf{r}_2}{m}\right] e^{-i\frac{\pi\lambda\Delta z}{m} f_1^2} \\
&\quad \times \mathcal{F}\left[\mathbf{r}_1,\mathbf{f}_1\right]\left\{\frac{1}{m} U\left(\mathbf{r}_1\right) e^{i\frac{k}{2\Delta z}(1-m) r_1^2}\right\} \\
&= e^{-i\frac{k}{2\Delta z}\frac{1-m}{m} r_2^2} \mathcal{F}^{-1}\left[\mathbf{f}_1,\frac{\mathbf{r}_2}{m}\right] e^{-i\frac{\pi\lambda\Delta z}{m} f_1^2} \\
&\quad \times \mathcal{F}\left[\mathbf{r}_1,\mathbf{f}_1\right]\left\{\frac{1}{m} U\left(\mathbf{r}_1\right) e^{i\frac{k}{2\Delta z}(1-m) r_1^2}\right\} \\
&= \mathcal{Q}\left[\frac{m-1}{m\Delta z},\mathbf{r}_2\right] \mathcal{F}^{-1}\left[\mathbf{f}_1,\frac{\mathbf{r}_2}{m}\right] \mathcal{Q}_2\left[-\frac{\Delta z}{m},\mathbf{f}_1\right] \\
&\quad \times \mathcal{F}\left[\mathbf{r}_1,\mathbf{f}_1\right] \mathcal{Q}\left[\frac{1-m}{\Delta z},\mathbf{r}_1\right] \frac{1}{m}\left\{U\left(\mathbf{r}_1\right)\right\}.
\end{aligned} \tag{6.50}$$

Now that we have an expression of angular-spectrum propagation in terms of operators, we can examine grid spacings δ_1 in the source plane, δ_{f1} in the spatial-frequency plane, and δ_2 in the observation plane:

$$\delta_{f1} = \frac{1}{N\delta_1} \quad \text{from}\;\; \mathcal{F}\left[\mathbf{r}_1,\mathbf{f}_1\right] \tag{6.51}$$

$$\delta_2 = \frac{m}{N\delta_{f1}} \quad \text{from } \mathcal{F}^{-1}\left[\mathbf{f}_1, \mathbf{r}_2/m\right] \tag{6.52}$$

$$= \frac{m}{N\left(\frac{1}{N\delta_1}\right)} \tag{6.53}$$

$$= m\delta_1. \tag{6.54}$$

This last equation is a consistency check. Also, we can determine two other relationships:

$$\frac{1}{1-m} = \frac{1}{1-\frac{\delta_2}{\delta_1}} = \frac{\delta_1}{\delta_1 - \delta_2} \tag{6.55}$$

$$\frac{m}{1-m} = \frac{\frac{\delta_2}{\delta_1}}{1-\frac{\delta_2}{\delta_1}} = \frac{\delta_2}{\delta_1 - \delta_2}. \tag{6.56}$$

These relationships are used later in Sec. 8.2.

Another solution for the angular-spectrum formulation can be found. Let us start at Eq. (6.34) to manipulate $|\mathbf{r}_2 - \mathbf{r}_1|^2$ a little differently:

$$|\mathbf{r}_2 - \mathbf{r}_1|^2 = r_2^2 - 2\mathbf{r}_2 \cdot \mathbf{r}_1 + r_1^2 \tag{6.57}$$

$$= \left(r_2^2 + \frac{r_2^2}{m} - \frac{r_2^2}{m}\right) - 2\mathbf{r}_2 \cdot \mathbf{r}_1 + \left(r_1^2 + mr_1^2 - mr_1^2\right) \tag{6.58}$$

$$= -\frac{r_2^2}{m} + \left(1 + \frac{1}{m}\right) r_2^2 - 2\mathbf{r}_2 \cdot \mathbf{r}_1 - mr_1^2 + (1+m)\, r_1^2 \tag{6.59}$$

$$= -\frac{r_2^2}{m} - 2\mathbf{r}_2 \cdot \mathbf{r}_1 - mr_1^2 + \left(1 + \frac{1}{m}\right) r_2^2 + (1+m)\, r_1^2 \tag{6.60}$$

$$= -m\left(\left|\frac{\mathbf{r}_2}{m}\right|^2 + 2\left(\frac{\mathbf{r}_2}{m}\right) \cdot \mathbf{r}_1 + r_1^2\right) + \left(1 + \frac{1}{m}\right) r_2^2 + (1+m)\, r_1^2 \tag{6.61}$$

$$= -m\left|\frac{\mathbf{r}_2}{m} + \mathbf{r}_1\right|^2 + \left(\frac{1+m}{m}\right) r_2^2 + (1+m)\, r_1^2. \tag{6.62}$$

With a substitution of $m' = -m$,

$$= m'\left|\frac{\mathbf{r}_2}{-m'} + \mathbf{r}_1\right|^2 + \left(\frac{1-m'}{-m'}\right) r_2^2 + \left(1-m'\right) r_1^2 \tag{6.63}$$

$$= m'\left|\frac{\mathbf{r}_2}{m'} - \mathbf{r}_1\right|^2 - \left(\frac{1-m'}{m'}\right) r_2^2 + \left(1-m'\right) r_1^2, \tag{6.64}$$

it is obvious that this is identical to Eq. (6.38) with the use of m' rather than m.

Now with the realization that $\pm m$ may be used in the angular-spectrum form of diffraction, there are two possible equations:

$$U\left(\mathbf{r}_2\right) = \mathcal{Q}\left[\frac{m-1}{m\Delta z}, \mathbf{r}_2\right] \mathcal{F}^{-1}\left[\mathbf{f}_1, \frac{\mathbf{r}_2}{m}\right] \mathcal{Q}_2\left[-\frac{\Delta z}{m}, \mathbf{f}_1\right]$$

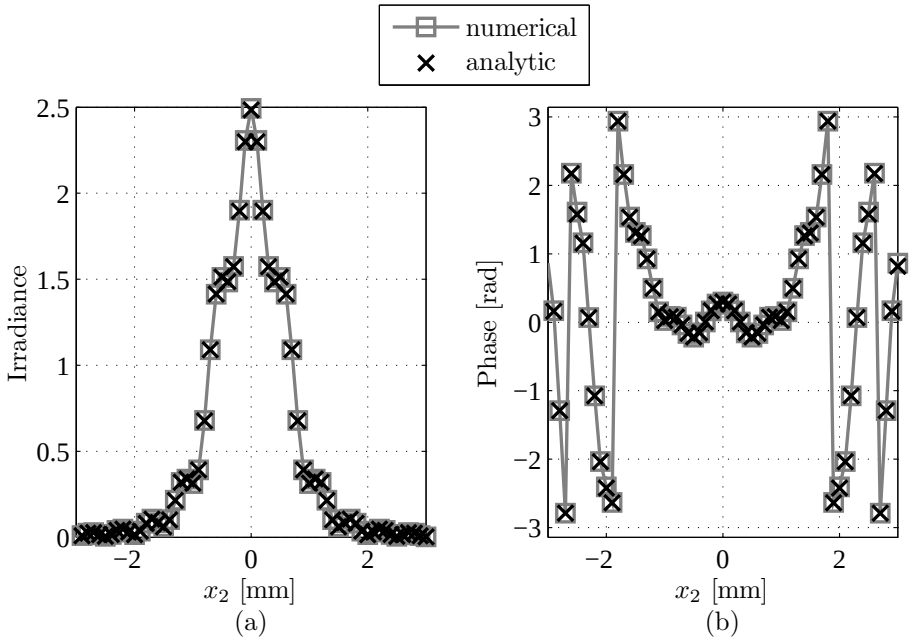

Figure 6.5 Fresnel diffraction from a square aperture, angular-spectrum simulation and analytic: (a) observation-plane irradiance and (b) observation-plane phase.

$$\times \mathcal{F}\left[\mathbf{r}_1, \mathbf{f}_1\right] \mathcal{Q}\left[\frac{1-m}{\Delta z}, \mathbf{r}_1\right] \frac{1}{m} \left\{U\left(\mathbf{r}_1\right)\right\} \tag{6.65}$$

$$= \mathcal{Q}\left[-\frac{m-1}{m\Delta z}, \mathbf{r}_2\right] \mathcal{F}^{-1}\left[\mathbf{f}_1, \frac{\mathbf{r}_2}{m}\right] \mathcal{Q}_2\left[\frac{\Delta z}{m}, \mathbf{f}_1\right]$$
$$\times \mathcal{F}\left[\mathbf{r}_1, \mathbf{f}_1\right] \mathcal{Q}\left[-\frac{1-m}{\Delta z}, \mathbf{r}_1\right] \left(\frac{-1}{m}\right) \left\{U\left(\mathbf{r}_1\right)\right\} \tag{6.66}$$

This can be written more compactly as

$$U\left(\mathbf{r}\right) = \mathcal{Q}\left[\frac{m \pm 1}{m\Delta z}, \mathbf{r}_2\right] \mathcal{F}^{-1}\left[\mathbf{f}_1, \mp\frac{\mathbf{r}_2}{m}\right] \mathcal{Q}_2\left[\pm\frac{\Delta z}{m}, \mathbf{f}_1\right]$$
$$\times \mathcal{F}\left[\mathbf{r}_1, \mathbf{f}_1\right] \mathcal{Q}\left[\frac{1 \pm m}{\Delta z}, \mathbf{r}_1\right] \left(\mp\frac{1}{m}\right) \left\{U\left(\mathbf{r}_1\right)\right\}, \tag{6.67}$$

where the top sign corresponds to Eq. (6.66), and the bottom sign corresponds to Eq. (6.65).

Listing 6.5 gives the MATLAB function `ang_spec_prop` that numerically evaluates Eq. (6.65). Figure 6.5 shows the results of repeating the previous MATLAB examples using angular-spectrum propagation. The code that produced Fig. 6.5 is not shown here because it is identical to Listing 6.4 except for line 14, which calls the function `ang_spec_prop` given in Listing 6.5. Note that the numerical results are identical to the analytic results again.

Listing 6.5 Example of evaluating the Fresnel diffraction integral in MATLAB using the angular-spectrum method.

```
function [x2 y2 Uout] ...
    = ang_spec_prop(Uin, wvl, d1, d2, Dz)
% function [x2 y2 Uout] ...
%     = ang_spec_prop(Uin, wvl, d1, d2, Dz)

    N = size(Uin,1);   % assume square grid
    k = 2*pi/wvl;    % optical wavevector
    % source-plane coordinates
    [x1 y1] = meshgrid((-N/2 : 1 : N/2 - 1) * d1);
    r1sq = x1.^2 + y1.^2;
    % spatial frequencies (of source plane)
    df1 = 1 / (N*d1);
    [fX fY] = meshgrid((-N/2 : 1 : N/2 - 1) * df1);
    fsq = fX.^2 + fY.^2;
    % scaling parameter
    m = d2/d1;
    % observation-plane coordinates
    [x2 y2] = meshgrid((-N/2 : 1 : N/2 - 1) * d2);
    r2sq = x2.^2 + y2.^2;
    % quadratic phase factors
    Q1 = exp(i*k/2*(1-m)/Dz*r1sq);
    Q2 = exp(-i*pi^2*2*Dz/m/k*fsq);
    Q3 = exp(i*k/2*(m-1)/(m*Dz)*r2sq);
    % compute the propagated field
    Uout = Q3.* ift2(Q2 .* ft2(Q1 .* Uin / m, d1), df1);
```

6.5 Simple Optical Systems

Most of the wave propagation simulations in this book are through either vacuum or weakly refractive media like atmospheric turbulence. Moreover, the whole formalism presented up to this point can be extended to simple refractive and reflective optical systems. The effect of such simple systems is described through geometric optics by the use of paraxial ray matrices.[45]

Ray matrices describe how a refractive element transforms the location and direction of paraxial rays. In this framework, rays are represented by their ray height y_1 (distance from the optical axis at a certain z location), ray slope y_1', and the refractive index n_1 of the medium that contains the ray. Usually rays are confined to the marginal ($y - z$) plane. As a ray passes through a simple optical system, the system's effect on the ray is represented by a system of two coupled linear equations:

$$y_2 = A\, y_1 + B\, n_1\, y_1' \tag{6.68}$$

$$n_2\, y_2' = C\, y_1 + D\, n_1\, y_1', \tag{6.69}$$

where y_2, y_2', and n_2 are the ray height, slope, and refractive index, respectively, after the optical system. This way, the system is characterized by the values of A, B, C, and D. This can be written in matrix-vector notation as

$$\begin{pmatrix} y_2 \\ n_2\, y_2' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} y_1 \\ n_1\, y_1' \end{pmatrix}. \tag{6.70}$$

Note that ray matrices are always written so that $AD - BC = 1$.

There are two elementary ray matrices: that for ray transfer and that for refraction. Ray transfer simply refers to pure propagation, and refraction means that the ray encounters a surface that forms the interface between two materials of unlike refractive index. With ray transfer, the ray slope remains the same, and the ray height increases according to the ray slope and propagation distance so that[45]

$$\begin{pmatrix} y_2 \\ n_2\, y_2' \end{pmatrix} = \begin{pmatrix} 1 & \Delta z/n_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ n_1\, y_1' \end{pmatrix}. \tag{6.71}$$

With refraction, the ray height remains the same, but the ray slope changes according to the paraxial version of Snell's law so that

$$\begin{pmatrix} y_2 \\ n_2\, y_2' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{n_1 - n_2}{R} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ n_1\, y_1' \end{pmatrix}, \tag{6.72}$$

where R is the surface's radius of curvature.[45]

Without regard to vignetting, optical systems can be modeled as the successive application of ray transfer and refraction matrices written right-to-left. For example, a light ray passing from air just before the front face of a singlet lens of index n to just after the back end of the lens encounters refraction at the first surface, transfer through the lens, and then refraction at the back interface, represented by the system matrix

$$\mathsf{S} = \begin{pmatrix} 1 & 0 \\ \frac{n-1}{R_2} & 1 \end{pmatrix} \begin{pmatrix} 1 & \Delta z/n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1-n}{R_1} & 1 \end{pmatrix}. \tag{6.73}$$

In this equation, R_1 and R_2 are the radii of curvature of the two lens faces. If the lens is thin enough that $\Delta z \approx 0$, then the lens matrix simplifies to

$$\mathsf{S} = \begin{pmatrix} 1 & 0 \\ (1-n)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) & 1 \end{pmatrix}. \tag{6.74}$$

Now, the lensmaker's equation gives the focal length f_l of a lens in terms its radii and index according to

$$\frac{1}{f_l} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right). \tag{6.75}$$

When this is used, the lens matrix becomes

$$\mathsf{S} = \begin{pmatrix} 1 & 0 \\ -1/f_l & 1 \end{pmatrix}. \tag{6.76}$$

Diffraction calculations account for simple optical systems through the generalized Huygens-Fresnel integral given by[15, 34, 46–48]

$$U(x_2, y_2) = \frac{e^{ikz}}{i\lambda B} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x_1, y_1)\, e^{i\frac{k}{2B}\left(Dr_2^2 - 2\mathbf{r}_1\cdot\mathbf{r}_2 + Ar_1^2\right)}\, dx_1 dy_1. \tag{6.77}$$

Note that this is valid only for optical systems possessing azimuthal symmetry, such as circular lenses with spherical radii of curvature on each face. Eq. (6.77) can be easily generalized for non-symmetric systems like square apertures, cylindrical lenses, and toroidal lenses.[47] This integral is closely related to the fractional Fourier transform.[49] Numerical implementations have been implemented numerically by several authors.[34, 50–52]

There are two particularly interesting cases to note here. For pure ray transfer, $A = D = 1$, $C = 0$, and $B = \Delta z$ so that Eq. (6.77) reduces to the free-space Fresnel diffraction integral in Eq. (6.1), as it should. When the light propagates from the front face of a spherical lens to its back focal plane, $A = 0$, $B = f_l$, $C = -f_l^{-1}$, and $D = 1$ so that Eq. (6.77) reduces to a scaled FT, much like in Eq. (4.9).

The generalized Huygens-Fresnel integral is more complicated than the Fresnel diffraction integral, and at first glance it may not appear like a convolution integral. However, Lambert and Fraser showed that simple substitutions can transform it into a convolution so that the computational methods discussed in the previous sections of this chapter may be applied.[47] Following their method, we substitute

$$\alpha = \frac{A}{\lambda B} \text{ and } \beta = \frac{AC}{\lambda} \tag{6.78}$$

and recall that $AD - BC = 1$ to obtain[47]

$$U(A\mathbf{r}_2) = \frac{1}{i\lambda B} e^{i\pi\beta r_2^2} \int_{-\infty}^{\infty} U(\mathbf{r}_1)\, e^{i\pi\alpha|\mathbf{r}_2 - \mathbf{r}_1|^2}\, d\mathbf{r}_1. \tag{6.79}$$

This is clearly a convolution, and we can write it explicitly as

$$U(A\mathbf{r}_2) = \frac{1}{i\lambda B} e^{i\pi\beta r_2^2} \left[U(\mathbf{r}_1) \otimes e^{i\pi\alpha r_1^2} \right]. \tag{6.80}$$

Further, we can see that the transfer function for the optical system is

$$H(\mathbf{f}) = \frac{i}{\alpha} e^{-i\frac{\pi}{\alpha}\left(f_x^2 + f_y^2\right)}. \tag{6.81}$$

Listing 6.6 Code for evaluating the Fresnel diffraction integral in MATLAB using the angular-spectrum method with an ABCD ray matrix.

```
function [x2 y2 Uout] ...
    = ang_spec_propABCD(Uin, wvl, d1, d2, ABCD)
% function [x2 y2 Uout] ...
%     = ang_spec_propABCD(Uin, wwl, d1, d2, ABCD)

    N = size(Uin,1);   % assume square grid
    k = 2*pi/wvl;    % optical wavevector
    % source-plane coordinates
    [x1 y1] = meshgrid((-N/2 : 1 : N/2 - 1) * d1);
    r1sq = x1.^2 + y1.^2;
    % spatial frequencies (of source plane)
    df1 = 1 / (N*d1);
    [fX fY] = meshgrid((-N/2 : 1 : N/2 - 1) * df1);
    fsq = fX.^2 + fY.^2;
    % scaling parameter
    m = d2/d1;
    % observation-plane coordinates
    [x2 y2] = meshgrid((-N/2 : 1 : N/2 - 1) * d2);
    r2sq = x2.^2 + y2.^2;
    % optical system matrix
    A = ABCD(1,1); B = ABCD(1,2); D = ABCD(2,2);
    % quadratic phase factors
    Q1 = exp(i*pi/(wvl*B)*(A-m)*r1sq);
    Q2 = exp(-i*pi*wvl*B/m*fsq);
    Q3 = exp(i*pi/(wvl*B)*(D-1/m)*r2sq);
    % compute the propagated field
    Uout = Q3.* ift2(Q2 .* ft2(Q1 .* Uin / m, d1), df1);
```

Recall that this algorithm does not account for vignetting of the rays due to finite-extent apertures in the optical system. The most straightforward way to handle this is to simulate propagation from aperture to aperture, setting the vignetted portions to zero at each aperture. However, the reader is directed to Coy for a more detailed and efficient method of accounting for vignetting in simulations.[35]

Listing 6.6 gives the MATLAB function `ang_spec_propABCD` that evaluates Eq. (6.79). Figure 6.6 shows the results of repeating the previous MATLAB examples using angular-spectrum propagation, using an ABCD ray matrix to represent the free space. The code that produced Fig. 6.6 is given in Listing 6.7. Note that the numerical results are identical to the analytic results again.

Listing 6.7 Example of propagating light from a square aperture using the ABCD ray-matrix simulation method.

```
% example_square_prop_ang_specABCD.m

N = 1024;    % number of grid points per side
L = 1e-2;   % total size of the grid [m]
delta1 = L / N;  % grid spacing [m]
D = 2e-3;   % diameter of the aperture [m]
wvl = 1e-6;  % optical wavelength [m]
k = 2*pi / wvl;
Dz = 1;     % propagation distance [m]
f = inf;    % source field radius of curvature [m]

[x1 y1] = meshgrid((-N/2 : N/2-1) * delta1);
ap = rect(x1/D) .* rect(y1/D);
delta2 = wvl * Dz / (N*delta1);

ABCD = [1 Dz; 0 1] * [1 0 ; -1/f 1];
[x2 y2 Uout] ...
    = ang_spec_propABCD(ap, wvl, delta1, delta2, ABCD);
```

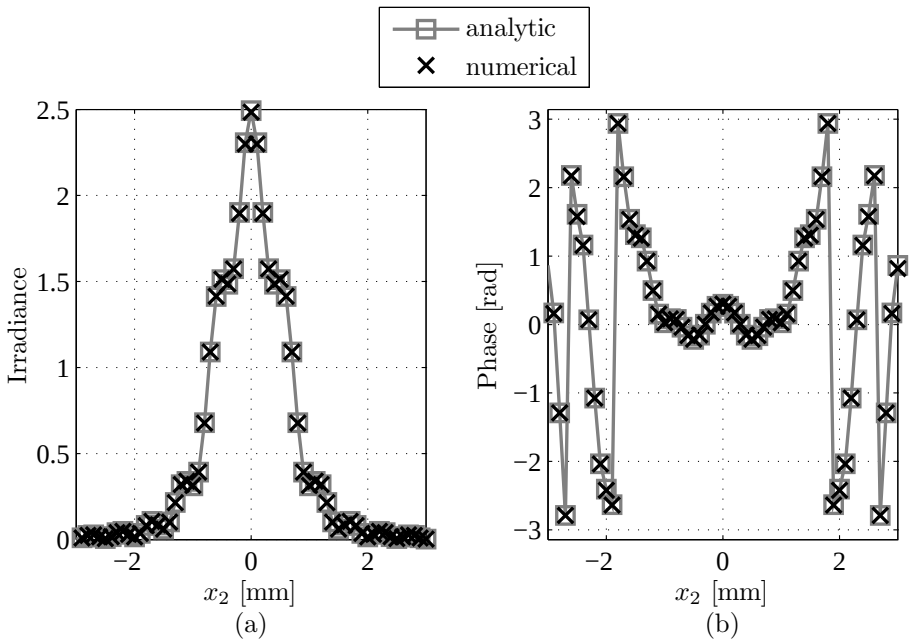

Figure 6.6 Observation-plane field resulting from square-aperture source with a diverging spherical wavefront. This simulation used the ABCD ray-matrix method of propagation.

Listing 6.8 Example of propagating a sinc model point source in MATLAB using the angular-spectrum method.

```
% example_pt_source.m

D = 8e-3;   % diameter of the observation aperture [m]
wvl = 1e-6;  % optical wavelength [m]
k = 2*pi / wvl; % optical wavenumber [rad/m]
Dz = 1;     % propagation distance [m]
arg = D/(wvl*Dz);
delta1 = 1/(10*arg); % source-plane grid spacing [m]
delta2 = D/100; % observation-plane grid spacing [m]
N = 1024;        % number of grid points
% source-plane coordinates
[x1 y1] = meshgrid((-N/2 : N/2-1) * delta1);
[theta1 r1] = cart2pol(x1, y1);
A = wvl * Dz;   % sets field amplitude to 1 in obs plane
pt = A * exp(-i*k/(2*Dz) * r1.^2) * arg^2 ...
    .* sinc(arg*x1) .* sinc(arg*y1);
[x2 y2 Uout] = ang_spec_prop(pt, wvl, delta1, delta2, Dz);
```

6.6 Point Sources

Point sources are especially challenging to model. Recall from Ch. 1 that a true point source $U_{pt}(\mathbf{r}_1)$ is represented by a Dirac delta function via

$$U_{pt}(\mathbf{r}_1) = \delta(\mathbf{r}_1 - \mathbf{r}_c), \tag{6.82}$$

where $\mathbf{r}_c = (x_c, y_c)$ is the location of the point source in the $x_1 - y_1$ plane. The field $U_{pt}(\mathbf{r}_1)$ has a Fourier spectrum that is constant across all spatial frequencies. This means that it has infinite spatial bandwidth, which is unusual because most optical sources are spatially bandlimited. The infinite spatial bandwidth is a problem for the discretely sampled and finite-sized grid that we must use in computer simulations. If a propagation grid has spacing δ_1 in the source plane, then the highest spatial frequency represented on that grid without aliasing is $1/(2\delta_1)$. Therefore, a bandlimited version of a point source must suffice. The point source in the simulation must have a finite spatial extent.

Various point-source models have been used in the literature. To simulate propagation though turbulence, Martin and Flatté[44] and Coles[32] used a narrow Gaussian function with a quadratic phase. Martin and Flatté's model point source is given by

$$\exp\left(-\frac{r^2}{2\sigma^2}\right)\exp\left(-i\frac{r^2}{2x_0^2}\right). \tag{6.83}$$

The parameters σ and x_0 were equal to the grid spacing. This is similar to the example from Sec. 2.5.3. With use of absorbing boundaries in the simulation (discussed

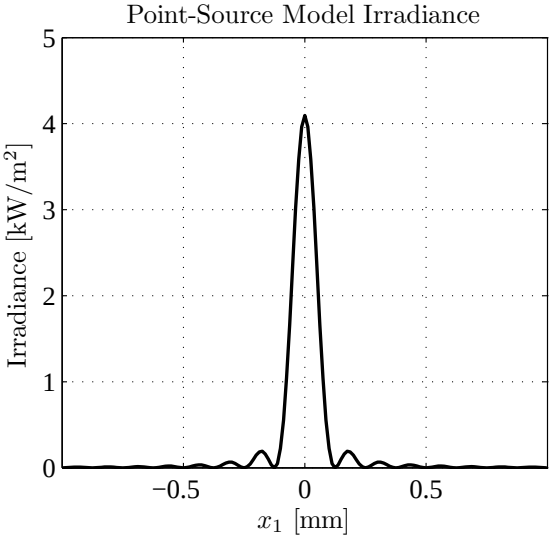

Figure 6.7 Irradiance of a sinc model of a point source (source plane).

in Sec. 8.1), this model produced an observation-plane field that was approximately flat across the central one-third of their propagation grid and tapered to zero toward the edge. Later, Flatté *et al.*[53] used a model point-source field given by

$$\exp\left(-\frac{r^2}{2\sigma^2}\right)\cos^2\left(\frac{r^2}{2\rho^2}\right), \tag{6.84}$$

where σ and ρ are nearly equal to the grid spacing. This model also produced a field that was approximately flat across the central one-third of their observation-plane grid and tapered to zero toward the edge.

Here, we take a different approach and seek a good model by analytically computing the desired observation-plane field. If we observe the field in the $x_2 - y_2$ plane a distance Δz away from the source, we can easily evaluate Eq. (6.1), (6.5), or (6.18) to obtain the field, given by

$$U\left(\mathbf{r}_2\right) = \frac{e^{ik\Delta z}}{i\lambda\Delta z} e^{i\frac{k}{2\Delta z}\left|\mathbf{r}_2-\mathbf{r}_c\right|^2}. \tag{6.85}$$

This result is the paraxial approximation to a spherical wave. It has constant amplitude across the $x_2 - y_2$ plane and a parabolic phase.

Our goal is to obtain good agreement between the simulation and potential experiments. Any camera or wavefront sensor that we might use occupies only a finite region of the $x_2 - y_2$ plane. Therefore, our source model is valid if our simulation obtains good agreement over the detector area. Then, let us work with a field $\widetilde{U}\left(\mathbf{r}_2\right)$ that has finite spatial extent, given by

$$\widetilde{U}\left(\mathbf{r}_2\right) = \frac{e^{ik\Delta z}}{i\lambda\Delta z} W\left(\mathbf{r}_2 - \mathbf{r}_c\right) e^{i\frac{k}{2\Delta z}\left|\mathbf{r}_2-\mathbf{r}_c\right|^2}, \tag{6.86}$$

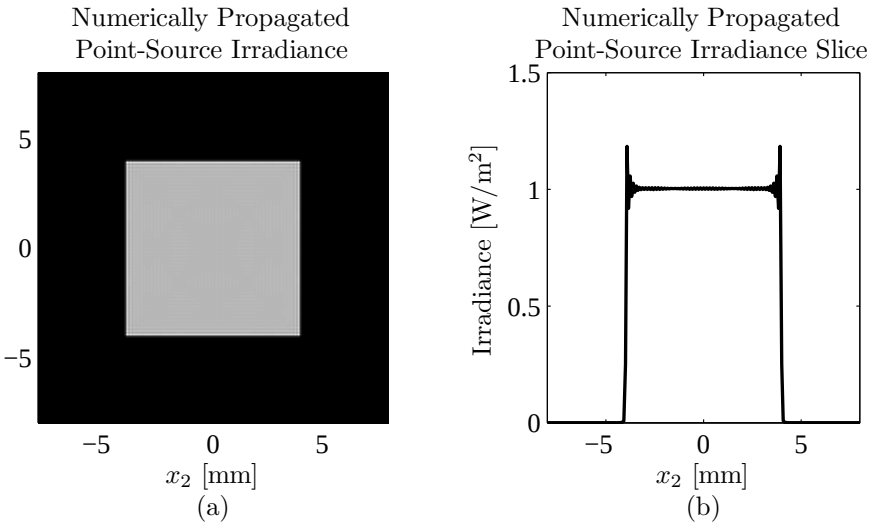

Figure 6.8 Fresnel diffraction irradiance from a sinc model of a point source (observation plane).

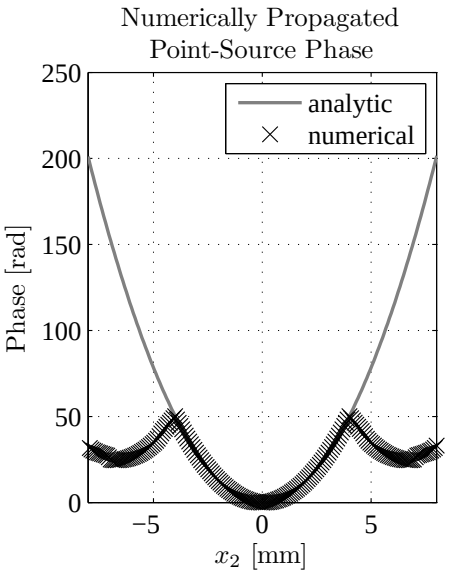

Figure 6.9 Fresnel diffraction phase from a sinc model of a point source (observation plane).

where $W(\mathbf{r}_2)$ is a "window" function that is nonzero over only a finite region of space. The extent of $W(\mathbf{r}_2)$ must be at least as large as the detector, but smaller than the propagation grid. For example, it might be a two-dimensional rect or circ function.

Let us represent our point-source model by $\widetilde{U}_{pt}(\mathbf{r}_1)$, substitute it into the Fres-

nel diffraction integral, and set the result equal to $\widetilde{U}(\mathbf{r}_2)$:

$$\begin{aligned}\widetilde{U}(\mathbf{r}_2) &= \frac{e^{ik\Delta z}}{i\lambda\Delta z}\, e^{i\frac{k}{2\Delta z}r_2^2} \int_{-\infty}^{\infty} \widetilde{U}_{pt}(\mathbf{r}_1)\, e^{i\frac{k}{2\Delta z}r_1^2} e^{-i\frac{2\pi}{\lambda\Delta z}\mathbf{r}_1\cdot\mathbf{r}_2}\, d\mathbf{r}_1 \\ &= \frac{e^{ik\Delta z}}{i\lambda\Delta z}\, e^{i\frac{k}{2\Delta z}r_2^2} \mathcal{F}\left\{\widetilde{U}_{pt}(\mathbf{r}_1)\, e^{i\frac{k}{2\Delta z}r_1^2}\right\}_{\mathbf{f}_1=\frac{\mathbf{r}_2}{\lambda\,\Delta z}}. \end{aligned} \tag{6.87}$$

Then, we can solve for the point-source model given by

$$\widetilde{U}_{pt}(\mathbf{r}_1) = i\lambda\Delta z\, e^{-ik\Delta z} e^{-i\frac{k}{2\Delta z}r_1^2} \mathcal{F}^{-1}\left\{\widetilde{U}(\lambda\Delta z\mathbf{f}_1)\, e^{-i\pi\lambda\Delta z f_1^2}\right\}. \tag{6.88}$$

Now, substituting Eq. (6.86) for $\widetilde{U}(\lambda\Delta z\mathbf{f}_1)$ yields

$$\widetilde{U}_{pt}(\mathbf{r}_1) = e^{-i\frac{k}{2\Delta z}r_1^2} e^{i\frac{k}{2\Delta z}r_c^2} \mathcal{F}^{-1}\left\{W(\lambda\Delta z\mathbf{f}_1 - \mathbf{r}_c)\, e^{-i2\pi\mathbf{r}_c\cdot\mathbf{f}_1}\right\}. \tag{6.89}$$

For example, if a square region of width D is being used,

$$W(\mathbf{r}_2 - \mathbf{r}_c) = A\ \mathrm{rect}\left(\frac{x_2 - x_c}{D}\right) \mathrm{rect}\left(\frac{y_2 - y_c}{D}\right) \tag{6.90}$$

(where A is an amplitude factor) so that we have a model point source given by

$$\begin{aligned}\widetilde{U}_{pt}(\mathbf{r}_1) &= A\, e^{-i\frac{k}{2\Delta z}r_1^2} e^{i\frac{k}{2\Delta z}r_c^2} \qquad (6.91)\\ &\quad \times \mathcal{F}^{-1}\left\{\mathrm{rect}\left(\frac{\lambda\Delta z f_x - x_c}{D}\right) \mathrm{rect}\left(\frac{\lambda\Delta z f_y - y_c}{D}\right) e^{-i2\pi\mathbf{r}_c\cdot\mathbf{f}_1}\right\} \\ &= A\, e^{-i\frac{k}{2\Delta z}r_1^2} e^{i\frac{k}{2\Delta z}r_c^2} e^{-i\frac{k}{\Delta z}\mathbf{r}_c\cdot\mathbf{r}_1} \qquad (6.92)\\ &\quad \times \left(\frac{D}{\lambda\Delta z}\right)^2 \mathrm{sinc}\left[\frac{D(x_1 - x_c)}{\lambda\Delta z}\right] \mathrm{sinc}\left[\frac{D(y_1 - y_c)}{\lambda\Delta z}\right]. \end{aligned}$$

An example use of a point source is given in Listing 6.8. The point-source model used in the code is shown in Fig. 6.7. The grid spacing is set so that there are ten grid points across the central lobe. This may not seem very point-like, but actually this is only 0.125 mm in diameter. This is much narrower than the window function, which is 8.0 mm across as can be seen in the plot of the propagated irradiance shown in Fig. 6.8. The propagated phase is shown in Fig. 6.9. The effect of the window is clearly visible in both plots, and the model point source is producing exactly what we want in the observation plane region of interest. Later when this model is used for turbulent simulations in Sec. 9.5, the parameter D in the model point source is set to be four times larger than the observing telescope diameter. This ensures that the turbulent fluctuations never cause the window edge to be observed by the telescope.

Unfortunately, Fig. 6.9 does show aliasing outside the region of interest. Perhaps a modification of the point-source model could mitigate some of the aliasing.

Listing 6.9 Example of propagating a sinc-Gaussian model point source in MATLAB using the angular-spectrum method.

```
% example_pt_source_gaussian.m

D = 8e-3;   % diameter of the observation aperture [m]
wvl = 1e-6;  % optical wavelength [m]
k = 2*pi / wvl; % optical wavenumber [rad/m]
Dz = 1;     % propagation distance [m]
arg = D/(wvl*Dz);
delta1 = 1/(10*arg); % source-plane grid spacing [m]
delta2 = D/100; % observation-plane grid spacing [m]
N = 1024;        % number of grid points
% source-plane coordinates
[x1 y1] = meshgrid((-N/2 : N/2-1) * delta1);
[theta1 r1] = cart2pol(x1, y1);
A = wvl * Dz;   % sets field amplitude to 1 in obs plane
pt = A * exp(-i*k/(2*Dz) * r1.^2) * arg^2 ...
    .* sinc(arg*x1) .* sinc(arg*y1) ...
    .* exp(-(arg/4*r1).^2);
[x2 y2 Uout] ...
    = ang_spec_prop(pt, wvl, delta1, delta2, Dz);
```

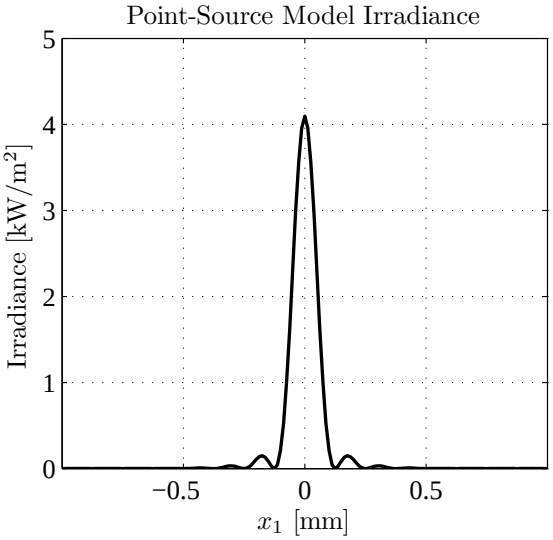

Figure 6.10 Irradiance of a sinc-Gaussian model of a point source (source plane).

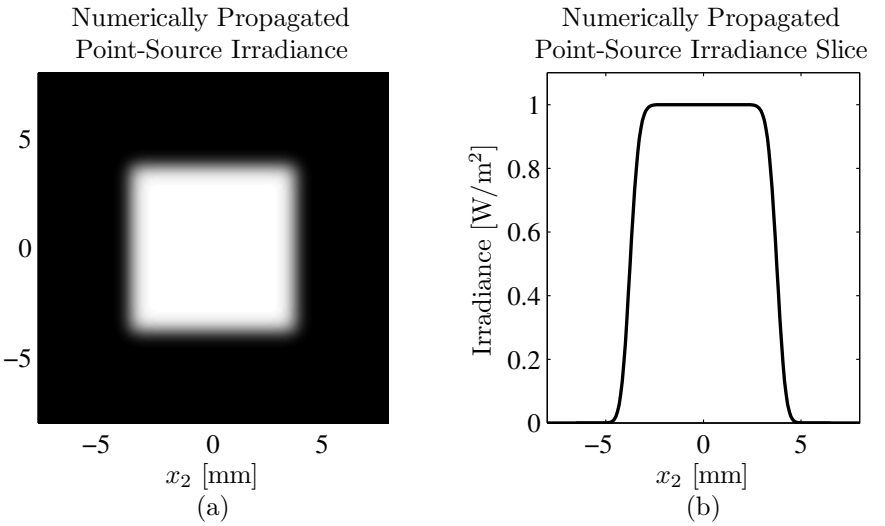

Figure 6.11 Fresnel diffraction irradiance from a sinc-Gaussian model of a point source (observation plane).

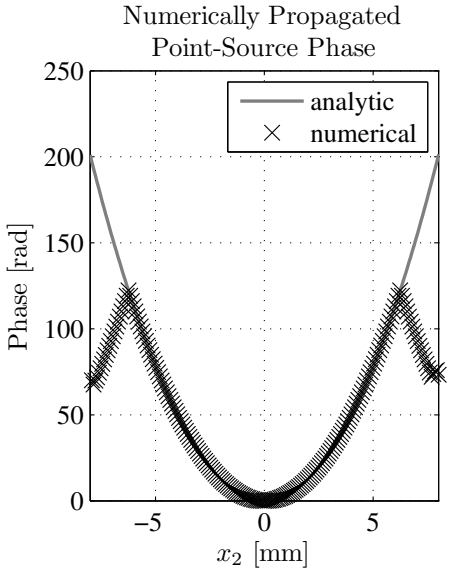

Figure 6.12 Fresnel diffraction phase from a sinc-Gaussian model of a point source (observation plane).

The approaches of Martin and Flatté and Flatté *et al.* do not have such a problem with aliasing because of the Gaussian model they use. Combining the sinc and Gaussian point-source models does, in fact, reduce the phase aliasing slightly. To illustrate, Listing 6.9 implements this. The code is very similar to Listing 6.8, but the model point source is multiplied by a Gaussian function in line 17.

The sinc-Gaussian model point source and resulting observation-plane field

are shown in Figs. 6.10–6.12. It is obvious by comparing Figs. 6.7 and 6.10 that the Gaussian factor reduces the side lobes in the model point source and thereby smooths the irradiance profile in the observation-plane field. Further, the computed observation-plane phase shown in Fig. 6.12 matches the analytic phase much better toward the edges of the grid.

6.7 Problems

1. Adjust the example in Listing 6.2 to propagate a Gaussian laser beam using the angular-spectrum method. In the source plane, let the laser beam be at its waist, i.e., $w = w_0 = 1$ mm and $R = \infty$, and let the observation plane be at $z_2 = 4$ m. Use $\lambda = 1$ μm, 512 grid points, a 1-cm grid in the source plane, and a 1.5-cm grid in the observation plane. Show separate plots of the irradiance and phase for the $y_2 = 0$ slice in the observation plane. Include the simulated and analytic results on the same plot for comparison.

2. Adjust the example in Listing 6.2 to propagate a focused beam with a circular aperture using the angular-spectrum method. Let the observation plane be the beam's focal plane. Use $\lambda = 1$ μm, $D = 1$ cm, $f_l = 16$ cm, 1024 grid points, a 2-cm grid in the source plane, and set the grid spacing in the observation plane to be one hundredth of the diffraction-limited spot diameter. Show a plot of the irradiance for the $y_2 = 0$ slice in the focal plane. Include the simulated and analytic results on the same plot for comparison.

3. Adjust the example in Listing 6.2 to simulate Talbot imaging using the angular-spectrum method. Let there be an amplitude grating with amplitude transmittance equal to

$$t_A(x_1, y_1) = \frac{1}{2}\left[1 + \cos\left(2\pi x_1/d\right)\right] \tag{6.93}$$

   in the source plane, and let the observation plane be the first Talbot-image plane. Use $\lambda = 1$ μm, $d = 0.5$ mm, 1024 grid points, a 2 cm grid in both the source plane and observation plane. Show images of the irradiance in the Talbot-image plane (You only need to display the central 10 periods). Display the simulated and analytic results side-by-side for comparison.

4. Compute the model point source if the region of interest is rectangular with widths D_x and D_y in the x_2 and y_2 directions, respectively.

5. Compute the model point source if the region of interest is circular with diameter D.