Chapter 1
Foundations of Scalar Diffraction Theory

Light can be described by two very different approaches: classical electrodynamics and quantum electrodynamics. In the classical treatment, electric and magnetic fields are continuous functions of space and time, and light comprises co-oscillating electric and magnetic wave fields. In the quantum treatment, photons are elementary particles with no mass nor charge, and light comprises one or more photons. There is rigorous theory behind each approach, and there is experimental evidence supporting both. Neither approach can be dismissed, which leads to the wave-particle duality of light. Generally, classical methods are used for macroscopic properties of light, while quantum methods are used for submicroscopic properties of light.

This book describes macroscopic properties, so it deals entirely with classical electrodynamics. When the wavelength λ of an electromagnetic wave is very small, approaching zero, the waves travel in straight lines with no bending around the edges of objects. That is realm of geometric optics. However, this book treats many situations in which geometric optics are inadequate to describe observed phenomena like diffraction. Therefore, the starting point is classical electrodynamics with solutions provided by scalar diffraction theory. Geometric optics is treated briefly in Sec. 6.5.

1.1 Basics of Classical Electrodynamics

Classical electrodynamics deals with relationships between electric fields, magnetic fields, static charge, and moving charge (i.e., current) in space and time based on the macroscopic properties of the materials in which the fields exist. We define each quantity here along with some basic relationships. This introduces the reader to the quantities in Maxwell's equations, which describe how electrically charged particles and objects give rise to electric and magnetic fields. Maxwell's equations are introduced here in their most general form, and then the discussion focuses on a specific case and solutions for oscillating electric and magnetic fields, which light comprises.

1.1.1 Sources of electric and magnetic fields

Electric charge, measured in coulombs, is a fundamental property of elementary particles and bulk materials. Classically, charge may be positive, negative, or zero. Further, charge is quantized, specifically the smallest possible nonzero amount of charge is the *elementary charge* $e = 1.602 \times 10^{-19}$ C. All nonzero amounts of charge are integer multiples of e. For bulk materials, the integer may be very large so that total charge can be treated as continuous rather than discrete. We denote the volume density of free charge, measured in coulombs per cubic meter, by $\rho(\mathbf{r}, t)$, where $\mathbf{r}$ is a three-dimensional spatial vector, and t is time. Moving charge density is called free volume current density $\mathbf{J}(\mathbf{r}, t)$. Volume current density is measured in Ampères per square meter (1 A = 1 C/s). This represents the time rate at which charge passes through a surface of unit area. Finally, charge is conserved, meaning that the total charge of any system is constant. This is mathematically stated by the continuity equation

$$\nabla \cdot \mathbf{J}(\mathbf{r}, t) + \frac{\partial \rho(\mathbf{r}, t)}{\partial t} = 0. \tag{1.1}$$

Almost every material we encounter in life is composed of many, many atoms each with many positive and negative charges. Usually, the numbers of positive and negative charges are equal or nearly equal so that the whole material is electrically neutral. Still, such a material can give rise to electric or magnetic fields when the total charge and free current are zero. If the distribution of charge is not homogeneous or if the charges are circulating in tiny current loops, fields could be present.

The separation of charge is described by the electric dipole moment, which is the amount of separated charge times the separation distance. If a bulk material has its charge arranged in many tiny dipoles, it is said to be electrically polarized. The volume polarization density $\mathbf{P}(\mathbf{r}, t)$ is the density of electric dipole moments per unit volume, measured in coulombs per square meter.

Magnetization is a similar concept for moving charge. Charge circulating in a tiny current loop is described by magnetic dipole moment, which is the circulating current times the area of the loop. When a bulk material has internal current arranged in many tiny loops, it is said to be magnetized. The volume magnetization density $\mathbf{M}(\mathbf{r}, t)$ is the density of magnetic dipole moments per unit volume, measured in Ampères per meter.

1.1.2 Electric and magnetic fields

When a hypothetical charge, called a test charge, passes near a bulk material that has non-zero ρ, $\mathbf{J}$, $\mathbf{P}$, or $\mathbf{M}$, the charge experiences a force. This interaction is characterized by two vectors $\mathbf{E}$ and $\mathbf{B}$. The electromagnetic force $\mathbf{F}$ on a test particle at a given point and time is a function of these vector fields and the particle's charge q and velocity $\mathbf{v}$. The Lorentz force law describes this interaction as

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{1.2}$$

If this empirical statement is valid (and, of course, countless experiments over the course of centuries have shown that it is), then two vector fields $\mathbf{E}$ and $\mathbf{B}$ are thereby defined throughout space and time, and these are called the "electric field" and "magnetic induction."[1]

Eq. (1.2) can be examined in a little more detail to provide more intuitive definitions of these fields. The electric field is the amount of force per unit of test charge when the test charge is stationary, given by

$$\mathbf{E} = \lim_{q \to 0^+} \left. \frac{\mathbf{F}}{q} \right|_{\mathbf{v}=\mathbf{0}} . \tag{1.3}$$

This is called a push-and-pull force because the force is in either the same or opposite direction as the field, depending on the sign of the charge. Electric field is measured in units of volts per meter (1 V = 1 N m/C). The magnetic field is related to the amount of force per unit test charge given by

$$\mathbf{v} \times \mathbf{B} = \lim_{q \to 0^+} \left. \frac{\mathbf{F} - q\mathbf{E}}{q} \right|_{\mathbf{v} \neq \mathbf{0}} . \tag{1.4}$$

The force due to a magnetic field is called deflective because it is perpendicular to the particle's velocity, which deflects its trajectory. Magnetic field is measured in units of Tesla [1 T = 1 N s/(C m)].

With this understanding of the fields, they now need to be related to the sources. This was accomplished through centuries of experimental measurements and theoretical and intuitive insight, resulting in

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = \mathbf{0} \tag{1.5}$$

$$\nabla \times \mathbf{B} - \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \left(\mathbf{J} + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M} \right) . \tag{1.6}$$

These are two of Maxwell's equations, the former being Faraday's law and the latter being Ampère's law with Maxwell's correction. In Eq. (1.6), the sources on the right hand side include the free current $\mathbf{J}$ and two terms due to bound currents. These are the polarization current $\partial \mathbf{P} / \partial t$ and the magnetization current $\nabla \times \mathbf{M}$.

These equations can be written in a more functionally useful form. Eq. (1.6) can be rewritten as

$$\nabla \times \left(\frac{\mathbf{B}}{\mu_0} - \mathbf{M} \right) = \mathbf{J} + \frac{\partial}{\partial t} \left(\epsilon_0 \mathbf{E} + \mathbf{P} \right) . \tag{1.7}$$

Making the definitions

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \tag{1.8}$$

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M} \tag{1.9}$$

introduces the concepts of electric displacement $\mathbf{D}$ and magnetic field $\mathbf{H}$, which are fields that account for the medium's response to the applied fields. Now, the working form of these Maxwell equations becomes

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.10}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}. \tag{1.11}$$

Further, when these are combined with conservation of charge expressed in Eq. (1.1), this leads to

$$\nabla \cdot \nabla \times \mathbf{H} = \nabla \cdot \mathbf{J} + \frac{\partial}{\partial t} \nabla \cdot \mathbf{D} \tag{1.12}$$

$$= -\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial t} \nabla \cdot \mathbf{D} \tag{1.13}$$

$$= 0. \tag{1.14}$$

Focusing on the right-hand side,

$$\frac{\partial}{\partial t} (\nabla \cdot \mathbf{D} - \rho) = 0 \tag{1.15}$$

$$\nabla \cdot \mathbf{D} - \rho = f(\mathbf{r}), \tag{1.16}$$

where $f(\mathbf{r})$ is an unspecified function of space but not time. Causality requires that $f(\mathbf{r}) = 0$ before the source is turned on, yielding Coulomb's law:

$$\nabla \cdot \mathbf{D} = \rho. \tag{1.17}$$

Similar manipulations yield

$$\nabla \cdot \mathbf{B} = 0. \tag{1.18}$$

This indicates that magnetic monopole charges do not exist. Finally, Eqs. (1.10), (1.11), (1.17), and (1.18) constitute Maxwell's equations.[1]

In this model of macroscopic electrodynamics, Eqs. (1.10) and (1.11) are two independent vector equations. With three scalar components each, these are six independent scalar equations. Unfortunately, given knowledge of the sources, there are four unknown vector fields $\mathbf{D}$, $\mathbf{B}$, $\mathbf{H}$, and $\mathbf{E}$. Each has three scalar components for a total of twelve unknown scalars. With so many more unknown field components than equations, this is a poorly posed problem.

The key is to understand the medium in which the fields exist. This produces a means of relating $\mathbf{P}$ to $\mathbf{E}$ and $\mathbf{M}$ to $\mathbf{H}$, which amount to six more scalar equations. For example, in simple media (linear, homogeneous, and isotropic),

$$\mathbf{P} = \epsilon_0 \chi_e \mathbf{E} \tag{1.19}$$

$$\mathbf{M} = \chi_m \mathbf{H}, \tag{1.20}$$

where χ_e is the electric susceptibility of the medium and χ_m is its magnetic susceptibility. Substituting these into Eqs. (1.8) and (1.9) yields

$$\begin{aligned} \mathbf{D} &= \epsilon_0 \mathbf{E} + \mathbf{P} && (1.21) \\ &= \epsilon_0 (1 + \chi_e) \mathbf{E} && (1.22) \\ &= \epsilon \mathbf{E} && (1.23) \end{aligned}$$

and

$$\begin{aligned} \mathbf{B} &= \mu_0 (\mathbf{H} + \mathbf{M}) && (1.24) \\ &= \mu_0 (1 + \chi_m) \mathbf{H} && (1.25) \\ &= \mu \mathbf{H}, && (1.26) \end{aligned}$$

where $\epsilon = (1 + \chi_e) \epsilon_0$ is the electric permittivity and $\mu = (1 + \chi_m) \mu_0$ is the magnetic permeability of the medium. Now, this simplifies Eqs. (1.10) and (1.11) so that

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} \tag{1.27}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \epsilon \frac{\partial \mathbf{E}}{\partial t}. \tag{1.28}$$

Now, there are still six equations but only six unknowns (as long as the free current density $\mathbf{J}$ is known). Finally, with a proper understanding of the materials, this is a well posed problem.

1.2 Simple Traveling-Wave Solutions to Maxwell's Equations

There are many solutions to Maxwell's equations, but there are only a few that can be written in closed form without an integral. This section begins with transforming Maxwell's four equations into two uncoupled wave equations. It continues with a few specific simple solutions such as the infinite-extent plane wave. A more general solution is left to the next section.

1.2.1 Obtaining a wave equation

This book deals with optical wave propagation through linear, isotropic, homogeneous, nondispersive, dielectric media in the absence of source charges and currents. In this case, the media discussed throughout the remainder of this book have

$$\epsilon = \text{a scalar, independent of } \lambda, \mathbf{r}, t \tag{1.29}$$

$$\mu = \mu_0 \tag{1.30}$$

$$\rho = 0 \tag{1.31}$$

$$\mathbf{J} = \mathbf{0}. \tag{1.32}$$

Taking the curl of Eq. (1.27) yields

$$\nabla \times (\nabla \times \mathbf{E}) = -\mu_0 \frac{\partial}{\partial t} (\nabla \times \mathbf{H}) . \tag{1.33}$$

Then, substituting in Eq. (1.28) gives

$$\nabla \times (\nabla \times \mathbf{E}) = -\mu_0 \epsilon \frac{\partial^2}{\partial t^2} \mathbf{E}. \tag{1.34}$$

Now, applying the vector identity $\nabla \times (\nabla \times \mathbf{E}) = \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}$ leads to

$$\nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu_0 \epsilon \frac{\partial^2}{\partial t^2} \mathbf{E}. \tag{1.35}$$

Finally, substituting in Eqs. (1.17) and (1.23), and keeping in mind that ϵ is independent of position results in a wave differential equation:

$$\nabla^2 \mathbf{E} - \mu_0 \epsilon \frac{\partial^2}{\partial t^2} \mathbf{E} = \mathbf{0}. \tag{1.36}$$

Similar manipulations beginning with the curl of Eq. (1.28) yield

$$\nabla^2 \mathbf{B} - \mu_0 \epsilon \frac{\partial^2}{\partial t^2} \mathbf{B} = \mathbf{0}. \tag{1.37}$$

When the Laplacian is used on the Cartesian components of $\mathbf{E}$ and $\mathbf{B}$, the result is six uncoupled but identical equations of the form

$$\left(\nabla^2 - \mu_0 \epsilon \frac{\partial^2}{\partial t^2} \right) U(x, y, z) = 0, \tag{1.38}$$

where the scalar $U(x, y, z)$ stands for any of the x-, y- or z- directed components of the vector fields $\mathbf{E}$ and $\mathbf{B}$.

At this point, we can define index of refraction

$$n = \sqrt{\frac{\epsilon}{\epsilon_0}} \tag{1.39}$$

and the vacuum speed of light

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \tag{1.40}$$

so that

$$\left(\nabla^2 - \frac{n^2}{c^2} \frac{\partial^2}{\partial t^2} \right) U(x, y, z) = 0. \tag{1.41}$$

The electric and magnetic fields that compose light are traveling wave fields. Therefore, fields with harmonic time dependence $\exp(-i2\pi\nu t)$ (where ν is the wave frequency) are the types of solutions sought for the purposes of this book. When this is substituted into Eq. (1.41), the result is

$$\left[\nabla^2 + \left(\frac{2\pi n\nu}{c}\right)^2\right] U = 0. \tag{1.42}$$

Typically, the wavelength is given by $\lambda = c/\nu$, and the wavenumber is defined as $k = 2\pi/\lambda$ so that

$$\left[\nabla^2 + k^2 n^2\right] U = 0. \tag{1.43}$$

This is the Helmholtz equation, and it appears in many other branches of physics including thermodynamics and quantum mechanics. At this point, we can dispense with the time dependence since it is the same for all solutions of the Helmholtz equation. From this point forward, the field $U(x, y, z)$ refers to the phasor portion of the optical field (i.e, no time dependence). Further, we define the units of $U(x, y, z)$ to be square-root watts per meter (1 W = 1 J/s = 1 N m/s) so that optical irradiance $I = |U|^2$ is in units of watts per meter squared. The value of the electric field or magnetic induction can always be obtained by a simple conversion of units.

1.2.2 Simple traveling-wave fields

There are several simple traveling-wave fields that are useful in this book. These are planar, spherical, and Gaussian-beam waves. With each of these solutions, the field at all points always maintains its planar, spherical, or Gaussian-beam form, and parameters like radius of curvature change in a simple manner as the wave propagates. The next section on scalar diffraction theory handles more general cases.

A planar wave is the simplest possible traveling wave. It has uniform amplitude and phase in any plane perpendicular to its direction of propagation. More generally, when the optical axis is not along the direction of propagation, a planar wave field is given by

$$U_P(\mathbf{r}) = A \exp(i\mathbf{k} \cdot \mathbf{r}), \tag{1.44}$$

where A is the amplitude of the wave and

$$\mathbf{k} = \frac{2\pi}{\lambda}(\alpha\hat{\mathbf{x}} + \beta\hat{\mathbf{y}} + \gamma\hat{\mathbf{z}}) \tag{1.45}$$

is the wavevector with direction cosines given by α, β, and γ. Then, making the direction cosines more explicit,

$$U_P(\mathbf{r}) = A \exp\left[i\frac{2\pi}{\lambda}(\alpha x + \beta y + \gamma z)\right]. \tag{1.46}$$

This wave travels at an angle $\cos^{-1}\alpha$ from the x-axis and $\cos^{-1}\beta$ from the y-axis as shown in Fig. 1.1.

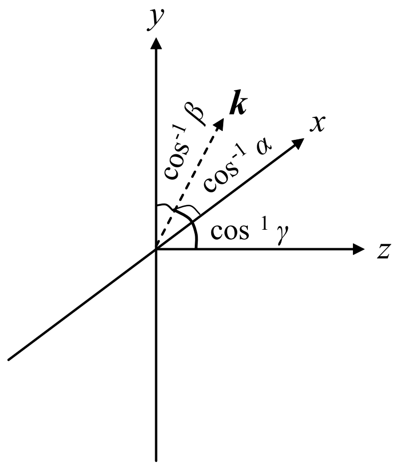

Figure 1.1 Depiction of direction cosines α, β, and γ.

A spherical wave is the next simplest wave field. It has a wavefront that is spherical in shape, and it is either diverging or converging. The energy of the wave is spread uniformly over a spherical surface with area given by $4\pi R^2$, where R is the wavefront radius of curvature. Conservation of energy requires that the amplitude is accordingly proportional to R^{-1}. A spherical wave is given by

$$U_S(\mathbf{r}) = A\frac{\exp[ikR(\mathbf{r})]}{R(\mathbf{r})}. \tag{1.47}$$

If the center of the sphere is located at $\mathbf{r}_c = (x_c, y_c, z_c)$, then at an observation point $\mathbf{r} = (x, y, z)$, the radius of curvature is given by

$$R(\mathbf{r}) = \sqrt{(x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2}. \tag{1.48}$$

Often in optics, attention is restricted to regions of space that are very close to the optical axis. This is called the paraxial approximation, and assuming propagation in the positive z direction, this approximation is mathematically written as

$$\cos^{-1}\alpha \ll 1 \tag{1.49}$$

$$\cos^{-1}\beta \ll 1. \tag{1.50}$$

With this approximation, we eliminate the square root by expanding it as a Taylor series and keeping only the first two terms, yielding

$$R(\mathbf{r}) \simeq \Delta z\left[1 + \frac{1}{2}\left(\frac{x - x_c}{\Delta z}\right)^2 + \frac{1}{2}\left(\frac{y - y_c}{\Delta z}\right)^2\right], \tag{1.51}$$

where we have defined $\Delta z = |z - z_c|$. With the paraxial approximation, a spherical wave is approximately

$$U_S(\mathbf{r}) \simeq A\frac{e^{ik\Delta z}}{\Delta z}e^{i\frac{k}{2\Delta z}\left[(x-x_c)^2+(y-y_c)^2\right]}. \tag{1.52}$$

One final simple traveling wave often encountered in optics is the Gaussian-beam wave. It has a Gaussian amplitude profile and "paraxially spherical" wavefront. The full derivation of the Gaussian-beam solution invokes the paraxial approximation along the way. Such a derivation can be found in common laser textbooks like Refs. 2–3. This solution is given by

$$U_G(\mathbf{r}) = \frac{A}{q(z)} \exp\left[ik \frac{x^2+y^2}{2q(z)}\right], \tag{1.53}$$

where

$$\frac{1}{q(z)} = \frac{1}{R(z)} + \frac{i\lambda}{\pi W^2(z)} \tag{1.54}$$

and the beam radius and wavefront radius of curvature are given by

$$W^2(z) = W_0^2 \left[1 + \left(\frac{\lambda z}{\pi W_0^2}\right)^2\right] \tag{1.55}$$

$$R(z) = z\left[1 + \left(\frac{\pi W_0^2}{\lambda z}\right)^2\right], \tag{1.56}$$

where W_0 is the minimum spot radius. At any point along the z axis, $W(z)$ is the $1/e$ radius of the field amplitude. Also, by this convention, $W(0) = W_0$ so that the minimum spot radius is located at $z = 0$.

1.3 Scalar Diffraction Theory

Often, the optical source is not a simple planar, spherical, nor Gaussian-beam wave. For more general cases, we must use more sophisticated means to solve the scalar Helmholtz equation. This means taking advantage of Green's theorem with clever use of boundary conditions. This process is not discussed in detail here, but the interested reader should consult books like Refs. 4–5 for a detailed treatment.

The geometry for this more general case is shown in Fig. 1.2. In this figure, the coordinates are $\mathbf{r}_1 = (x_1, y_1)$ in the source plane and $\mathbf{r}_2 = (x_2, y_2)$ in the observation plane. The distance between the two planes is Δz. The figure illustrates the basic problem: given the source-plane optical field $U(x_1, y_1)$, what is the observation-plane field $U(x_2, y_2)$? The solution is given by the Fresnel diffraction integral

$$U(x_2, y_2) = \frac{e^{ik\Delta z}}{i\lambda\Delta z} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(x_1, y_1)\, e^{i\frac{k}{2\Delta z}\left[(x_1-x_2)^2+(y_1-y_2)^2\right]}\, dx_1 dy_1. \tag{1.57}$$

Note that this is not the most general solution. In fact, it is a paraxial approximation, but it is general enough and accurate enough for the purposes of this book.

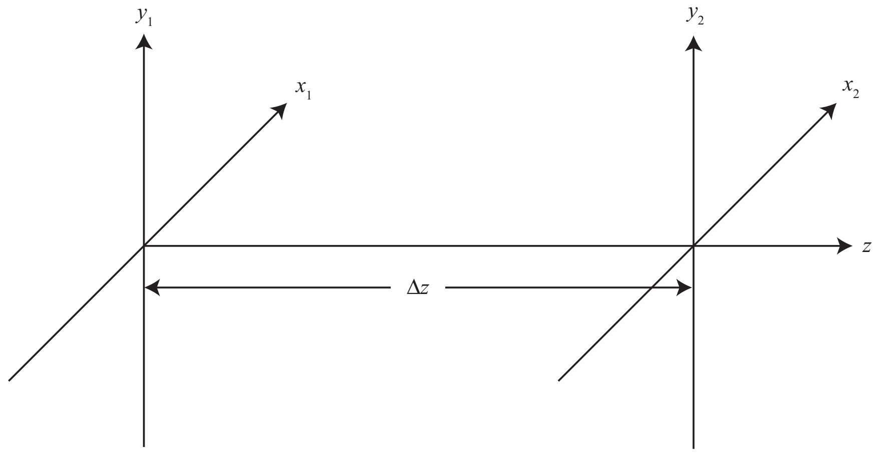

Figure 1.2 Coordinate systems for optical-wave propagation.

There are only a handful of analytic solutions to Eq. (1.57). Particularly, Fresnel diffraction from a rectangular aperture is used many times as an example in Chs. 6–8. Because few other Fresnel diffraction problems have an analytic answer, this one is used to compare against numerical results in several example simulations. When the source field is

$$U(x, y) = \mathrm{rect}\left(\frac{x_1}{D}\right)\mathrm{rect}\left(\frac{y_1}{D}\right), \tag{1.58}$$

(for the definition of the rect function, see Appendix A) the diffracted field in the observation plane a distance Δz away is given by

$$U(x_2, y_2) = \frac{e^{ik\Delta z}}{i\lambda\Delta z}\int_{-D/2}^{D/2}\int_{-D/2}^{D/2} e^{i\frac{k}{2\Delta z}\left[(x_1-x_2)^2+(y_1-y_2)^2\right]}\,dx_1\,dy_1. \tag{1.59}$$

The details of the steps involved in solving this integral are given in Fourier-optics textbooks like Goodman (Ref. 5). The solution, making use of Fresnel sine and cosine integrals is given by

$$\begin{aligned} U(x_2, y_2) = \frac{e^{ik\Delta z}}{2i} &\{[C(\alpha_2) - C(\alpha_1)] + i[S(\alpha_2) - S(\alpha_1)]\} \\ &\times \{[C(\beta_2) - C(\beta_1)] + i[S(\beta_2) - S(\beta_1)]\}, \end{aligned} \tag{1.60}$$

where

$$\alpha_1 = -\sqrt{\frac{2}{\lambda\Delta z}}\left(\frac{D}{2} + x_2\right) \tag{1.61}$$

$$\alpha_2 = \sqrt{\frac{2}{\lambda\Delta z}}\left(\frac{D}{2} - x_2\right) \tag{1.62}$$

$$\beta_1 = -\sqrt{\frac{2}{\lambda \Delta z}} \left(\frac{D}{2} + y_2 \right) \tag{1.63}$$

$$\beta_2 = \sqrt{\frac{2}{\lambda \Delta z}} \left(\frac{D}{2} - y_2 \right). \tag{1.64}$$

In Eq. (1.60), $S(x)$ and $C(x)$ are the Fresnel sine and cosine integrals given by

$$S(x) = \int_0^x \sin\left(\frac{\pi t^2}{2} \right) dt \tag{1.65}$$

$$C(x) = \int_0^x \cos\left(\frac{\pi t^2}{2} \right) dt, \tag{1.66}$$

respectively. MATLAB code for evaluating this solution is given in Appendix B.

Numerically evaluating the Fresnel diffraction integral with accurate results poses some interesting challenges. These challenges are due to using discrete samples on a finite-sized grid, which is required to evaluate this integral on a digital computer. Basic analysis of these issues is discussed in Ch. 2, which actually focuses on Fourier transforms because they arise so often in scalar diffraction theory. In fact, Eq. (1.57) can be written in terms of a Fourier transform, which is desirable because discrete Fourier transforms can be computed with great efficiency.

After Ch. 2 discusses discrete Fourier transforms, Ch. 3 discusses several basic computations that can be written in terms of Fourier transforms. Chapter 4 presents this book's first application of discrete Fourier transforms to optics by studying situations with very far propagation distances through free space and situations with lenses. These conditions allow simplifications to Eq. (1.57). For example, when we assume that the propagation distance Δz is very far, we can approximate the quadratic phase factor in Eq. (1.57) as being flat. Specifically, we must have $\Delta z > 2D^2/\lambda$, where D is the maximum spatial extent of the source-plane field[5]. This is the Fraunhofer approximation, which leads to the Fraunhofer diffraction integral

$$U(x_2, y_2) = \frac{e^{ik\Delta z} e^{i \frac{k}{2\Delta z}\left(x_2^2 + y_2^2\right)}}{i\lambda \Delta z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x_1, y_1)\, e^{-i\frac{k}{\Delta z}(x_1 x_2 + y_1 y_2)}\, dx_1 dy_1. \tag{1.67}$$

As an example of a Fraunhofer diffraction pattern, consider a planar wave passing through a two-slit aperture in an opaque screen. With two rectangular slits, the field just after the screen is

$$U(x_1, y_1) = \left[\mathrm{rect}\left(\frac{x_1 - \Delta x/2}{D_x} \right) + \mathrm{rect}\left(\frac{x_1 + \Delta x/2}{D_x} \right) \right] \mathrm{rect}\left(\frac{y_1}{D_y} \right), \tag{1.68}$$

where the slits are D_x wide in the x_1 direction and D_y wide in the y_1 direction and $\Delta x > D_x$ is the distance between the slits' centers. The resulting observation-plane field is

$$U\left(x_2,y_2\right)=\frac{e^{ik\Delta z}e^{i\frac{k}{2\Delta z}\left(x_2^2+y_2^2\right)}}{i\lambda\Delta z}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left[\mathrm{rect}\left(\frac{x_1-\Delta x/2}{D_x}\right)+\mathrm{rect}\left(\frac{x_1+\Delta x/2}{D_x}\right)\right]$$

$$\times\,\mathrm{rect}\left(\frac{y_1}{D_y}\right)e^{-i\frac{k}{\Delta z}\left(x_1x_2+y_1y_2\right)}\,dx_1dy_1 \tag{1.69}$$

$$=\frac{e^{ik\Delta z}e^{i\frac{k}{2\Delta z}\left(x_2^2+y_2^2\right)}}{i\lambda\Delta z}\left[\int\limits_{-(\Delta x+D_x)/2}^{(-\Delta x+D_x)/2}e^{-i\frac{k}{\Delta z}x_1x_2}\,dx_1+\int\limits_{(\Delta x-D_x)/2}^{(\Delta x+D_x)/2}e^{-i\frac{k}{\Delta z}x_1x_2}\,dx_1\right]$$

$$\times\int\limits_{-D_y/2}^{D_y/2}e^{-i\frac{k}{\Delta z}y_1y_2}\,dy_1 \tag{1.70}$$

$$=e^{ik\Delta z}e^{i\frac{k}{2\Delta z}\left(x_2^2+y_2^2\right)}\frac{2D_xD_y}{i\lambda\Delta z}\cos\left(\frac{\pi\Delta x\,x_2}{\lambda\Delta z}\right)\mathrm{sinc}\left(\frac{D_xx_2}{\lambda\Delta z}\right)\mathrm{sinc}\left(\frac{D_yy_2}{\lambda\Delta z}\right). \tag{1.71}$$

While fully coherent illumination was used here, two-slit apertures like this are useful for studying partially coherent sources.[6]

Further problems involving Fraunhofer (Ch. 4) and Fresnel (Chs. 6–8) diffraction are studied and simulated later in the book.

1.4 Problems

1. Using Maxwell's equations, show that

$$\mathbf{E}=-\frac{c^2}{2\pi\nu}\mathbf{k}\times\mathbf{B} \tag{1.72}$$

for a planar wave propagating through vacuum.

2. Using Maxwell's equations, show that

$$\mathbf{B}=\frac{1}{2\pi\nu}\mathbf{k}\times\mathbf{E} \tag{1.73}$$

for a planar wave propagating through vacuum.

3. A diverging spherical wave is the result of a Dirac delta-function source. Show that when the source field $U\left(\mathbf{r}_1\right)=\delta\left(\mathbf{r}_1\right)$ is substituted into the Fresnel diffraction integral, the observation-plane field $U\left(\mathbf{r}_2\right)$ is a paraxial spherical wave.

4. Write the scalar wave equation in spherical coordinates and show that the spherical wave is a solution.

5. Suppose that a spherical wave given by

$$U\left(\mathbf{r}_1\right) = A\frac{e^{ikR_1}}{R_1}e^{i\frac{k}{2R_1}\left[(x-x_c)^2+(y-y_c)^2\right]} \tag{1.74}$$

   is the optical field in the source plane. Substitute this into Eq. (1.57) to compute the optical field $U\left(\mathbf{r}_2\right)$ in the observation plane.

6. Suppose that a monochromatic, uniform-amplitude planar wave has passed through an annular circular aperture, and immediately after the aperture, the field is given by

$$U\left(\mathbf{r}_1\right) = \mathrm{circ}\left(\frac{2r_1}{D_{out}}\right) - \mathrm{circ}\left(\frac{2r_1}{D_{in}}\right), \tag{1.75}$$

   where $D_{out} > D_{in}$. Use the Fraunhofer diffraction integral to compute the observation-plane field (far away).