2

RADIATION STRUCTURES AND NUMERICAL METHODS

Antenna analysis, an important part of design, requires a compromise between extensive calculations and the fabrication and measurement of prototypes, which depends on your working environment. You should minimize cost, which means reducing the time from the start of a design to completion of a working model. In some cases you should not rush to build a prototype. For example, when designing large and expensive antennas, such as paraboloidal reflectors, the high fabrication cost justifies the time required for analysis. Management will not let you proceed before knowing the design will work. You should develop a cost model for each design in which analysis is one factor.

Analyses allow optimization of a design. You can design a number of antennas and adjust the dimensions until you find the best one. Again, you should be considering the costs of your time. At some point the incremental improvements are not worth the extra time for further analyses. In any case, when you build the prototype, you can expect differences. You soon determine that you can achieve only limited knowledge about a design because fabrication and measurement errors mask the true response of the antenna. You are doing engineering, not a science project.

Textbooks contain many analyses of ideal antennas, and this book is no exception. You need to consider the application and the final antenna environment. The mounting structure has little effect on the pattern of a large antenna with narrow beamwidth because little radiation strikes it. The overall radiation characteristics of narrow- or wide-beam antennas depend significantly on the shape of the vehicle and how the antenna is mounted. In later chapters we discuss how to use antenna mounting to improve performance, so you can take advantage of it. The size of the mounting structure limits the type of analysis used.

In this chapter we discuss physical optics (PO) and geometric optics (GO) [geometric theory of diffraction (GTD)] for large structures. In physical optics we compute the current induced on the vehicle due to antenna radiation and include their radiation in the

overall pattern. But the PO analysis cost rises rapidly as the number of small current patches increases for larger structures. PO analysis works well with large antennas, such as paraboloidal reflectors, that produce focused beams. Geometric optics uses ray optics techniques whose computation cost is independent of the size of the vehicle and whose accuracy improves as structure size increases. GO provides insight because we can visualize the combination of direct, reflected, and diffracted (GTD) rays to calculate the pattern, but it requires the solution of difficult geometry problems.

Smaller structures allow the use of multiple methods. For example, the moment method divides the surroundings into small patches and uses an expansion of the current in predetermined basis functions. This method uses integral equations of the boundary conditions to calculate a matrix equation involving coefficients of the current expansion. Numerical methods invert the matrix to solve for the coefficients, but it is a costly numerical operation and limits the size of the problem that can be handled to a few wavelengths. The finite-difference time-domain (FDTD) technique computes the fields on the structure in the time domain. This method handles moderate-sized structures and readily includes complex material properties such as biological features. FDTD divides the region into cubic cells and when excited by pulse feeding functions, it produces wide frequency bandwidth responses. Finite-element methods (FEMs) also divide the problem into cubic cells, but the analysis is performed in the frequency domain. FEM analysis must be repeated at every frequency of interest. FDTD and FEMs require a program to divide the structure into a mesh before starting the solution. Both methods calculate currents on a boundary surface by using the equivalence theorem with the incident fields and then calculate the far-field radiation pattern from these boundary currents.

Most methods start by assuming a current distribution on the antenna or, equivalently, a distribution of fields on an aperture. The fields on the aperture can be reduced to a current distribution. The moment method uses a summation of assumed basis function currents and solves for the coefficients of the expansion, but it, too, starts with assumed currents over small regions. You will discover that the radiation pattern can be found with greater accuracy than the input impedance. For antennas constructed from wires, the moment method computes the input current for a given excitation voltage and we calculate impedance from the ratio. Interaction of an antenna with the currents induced on a structure has little effect on impedance for narrow-beam antennas. Even for wide-beam antennas, such as dipoles, the structure effect on impedance can be found by using source mutual coupling with its images. In the end, antenna impedance should be measured when mounted in the final configuration.

An antenna has both a radiation pattern bandwidth and an impedance bandwidth, but you must give the pattern primary consideration. Too many designs concentrate on the wideband impedance characteristics of an antenna when, in fact, the antenna pattern has changed over the frequency range of the impedance bandwidth. Your primary task should be to design for the radiation pattern desired. In Chapter 1 we detailed the system aspects of impedance mismatch (Section 1-10), and you may determine the overall system impact of small impedance mismatch.

2-1 AUXILIARY VECTOR POTENTIALS

We do not use vector potentials in design. It seems as though they would be useful, but only a few simple antennas fit their direct use. You cannot measure them

because they are not physical entities, so they seem artificial. Physical optics (PO) calculates the radiation directly from currents using dyadic Green's functions but uses long expressions. Nevertheless, many analysis techniques find them more efficient than PO expressions and you should be aware of them. We illustrate their use with a couple of simple antennas.

We use vector potentials to introduce a few antenna concepts. In the first example we apply the magnetic vector potential to calculate the radiation from a short-length current element (dipole) and show how to obtain the pattern. Integration of the radiation pattern power density (Section 1-2) determines the total power radiated. Because we know the input current and the total radiated power, the ratio of the power to the input current squared gives the radiation resistance. We combine the low radiation resistance with the material resistance to compute the antenna efficiency. Electric vector potentials used with fictitious magnetic currents illustrate analysis by duality. We apply this to the analysis of a small loop and show that it has the same pattern as that of a small dipole.

2-1.1   Radiation from Electric Currents

Normal electron currents radiate when time varying. The simplest example is a filamentary current on wire, but we include surface and volumetric current densities as well. We analyze them by using the magnetic vector potential. Far-field electric fields are proportional to the magnetic vector potential $\mathbf{A}$:

$$\mathbf{E} = -j\omega\mathbf{A} \tag{2-1}$$

We determine the magnetic field from

$$|\mathbf{E}| = \eta|\mathbf{H}| \tag{2-2}$$

and realizing the cross product of the electric field with the magnetic field points in the direction of power flow, the Poynting vector. Since the electric field direction defines polarization, we usually ignore the magnetic field. We derive the magnetic vector potential from a retarded volume integral over the current density $\mathbf{J}$:

$$\mathbf{A} = \mu \iiint \frac{\mathbf{J}(\mathbf{r}')e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \, dV' \tag{2-3}$$

where $\mathbf{r}$ is the field measurement point radius vector, $\mathbf{r}'$ the source-point radius vector, μ the permeability ($4\pi \times 10^{-7}$ A/m in free space), and k, the wave number, is $2\pi/\lambda$. As written, Eq. (2-3) calculates the potential $\mathbf{A}$ everywhere: near and far field. The vector potential can be written in terms of a free-space Green's function:

$$g(R) = \frac{e^{-jkR}}{4\pi R} \qquad \text{where} \quad R = |\mathbf{r} - \mathbf{r}'|$$

$$\mathbf{A} = \mu \iiint g(R)\mathbf{J}(\mathbf{r}') \, dV' \tag{2-4}$$

Radiation Approximation   When we are interested only in the far-field response of an antenna, we can simplify the integral [Eq. (2-3)]. An antenna must be large in terms

of wavelengths before it can radiate efficiently with gain, but at great distances it still appears as a point source. Consider the radiation from two different parts of an antenna. Far away from the antenna, the ratio of the two distances to the different parts will be nearly 1. The phase shift from each part will go through many cycles before reaching the observation point, and when adding the response from each part, we need only the difference in phase shift. In the radiation approximation we pick a reference point on the antenna and use the distance from that point to the far-field observation point for amplitudes, $1/R$, for all parts of the antenna. The direction of radiation defines a plane through the reference point. This plane is defined by the radius normal vector, given in rectangular coordinates by

$$\hat{\mathbf{r}} = \sin\theta\cos\phi\hat{\mathbf{x}} + \sin\theta\sin\phi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}}$$

We compute the phase difference to the far-field point by dropping a normal to the reference plane from each point on the antenna. This distance multiplied by k, the propagation constant, is the phase difference. Given a point on the antenna $\mathbf{r}'$, the phase difference is $k\mathbf{r}' \cdot \hat{\mathbf{r}}$. When we substitute these ideas into Eq. (2-3), the equation becomes

$$\mathbf{A} = \frac{e^{-jkr}}{4\pi r}\mu \iiint \mathbf{J}'e^{jk\mathbf{r}'\cdot\hat{\mathbf{r}}}\,dV' \tag{2-5}$$

In rectangular coordinates $k\mathbf{r}' \cdot \hat{\mathbf{r}}$ becomes

$$k(x'\sin\theta\cos\phi + y'\sin\theta\sin\phi + z'\cos\theta)$$

We can combine k and $\hat{\mathbf{r}}$ to form a **k**-space vector:

$$\mathbf{k} = k\hat{\mathbf{r}} = k\sin\theta\cos\phi\hat{\mathbf{x}} + k\sin\theta\sin\phi\hat{\mathbf{y}} + k\cos\theta\hat{\mathbf{z}}$$

and the phase constant becomes $\mathbf{k} \cdot \mathbf{r}'$. Currents in filaments (wires) simplify Eq. (2-5) to a single line integral. Magnetic vector potentials and electric fields are in the same directions as the wires that limit the directions of current. For example, filamentary current along the z-axis produces z-directed electric fields. Spherical waves (far field) have only $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ components found from the projection of E_z onto those axes. Filamentary currents on the z-axis produce only z-directed electric fields with a null from $\hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{z}} = -\sin\theta$ at $\theta = 0$. In turn, x- or y-directed currents produce electric fields depending on the scalar products (projections) of the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ unit vectors onto the $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ vectors in the far field:

$$\hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{x}} = \cos\theta\cos\phi \qquad \hat{\boldsymbol{\phi}} \cdot \hat{\mathbf{x}} = -\sin\phi$$
$$\hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{y}} = \cos\theta\sin\phi \qquad \hat{\boldsymbol{\phi}} \cdot \hat{\mathbf{y}} = \cos\phi$$

By examining antenna structure you can discover some of its characteristics without calculations. Without knowing the exact pattern, we estimate the polarization of the waves by examining the directions of the wires that limit the current density. Consider various axes or planes of symmetry on an antenna: for example, a center-fed wire along the z-axis. If we rotate it about the z-axis, the problem remains the same, which means that all conical polar patterns (constant θ) must be circles; in other words, all great

circle patterns must be the same. An antenna with the same structure above and below the $x-y$ plane radiates the same pattern above and below the $x-y$ plane. Always look for axes and planes of symmetry to simplify the problem.

We can extend the magnetic vector potential [Eq. (2-1)] to determine near fields:

$$\mathbf{E} = -j\omega\mathbf{A} + \frac{\nabla(\nabla \cdot \mathbf{A})}{j\omega\varepsilon\mu}$$

$$\mathbf{H} = \frac{1}{\mu}\nabla \times \mathbf{A} \tag{2-6}$$

The electric field separates into far- and near-field terms, but the equation for the magnetic field, the defining equation of the potential, does not separate. If we substitute the free-space Green's function from Eq. (2-4) into Eq. (2-6), expand, and gather terms, we can determine the fields directly from the electric currents and eliminate the use of a vector potential.

$$\mathbf{E}(\mathbf{r}) = \frac{\eta k^2}{4\pi} \iiint_{V'} \left[\mathbf{J}(\mathbf{r}') \left(-\frac{j}{kR} - \frac{1}{k^2R^2} + \frac{j}{k^3R^3} \right) \right.$$

$$\left. + [\mathbf{J}(\mathbf{r}') \cdot \hat{\mathbf{R}}]\hat{\mathbf{R}} \left(\frac{j}{kR} + \frac{3}{k^2R^2} - \frac{3j}{k^3R^3} \right) \right] e^{-jkR} \, dV' \tag{2-7}$$

$$\mathbf{H}(\mathbf{r}) = \frac{k^2}{4\pi} \iiint_{V'} \mathbf{J}(\mathbf{r}') \times \hat{\mathbf{R}} \left(\frac{j}{kR} + \frac{1}{k^2R} \right) e^{-jkR} \, dV' \tag{2-8}$$

$$\hat{\mathbf{R}} = \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} = \frac{\mathbf{r} - \mathbf{r}'}{R} \quad \text{since} \quad R = |\mathbf{r} - \mathbf{r}'|$$

Terms with $1/R$ dependence are the far-field terms. The radiative near-field terms have $1/R^2$ dependence and near-field terms have $1/R^3$ dependence. The impedance of free space, η, is 376.7 Ω. We can rearrange Eqs. (2-7) and (2-8) so that they become the integral of the dot product of the current density $\mathbf{J}$ with dyadic Green's functions [1]. It is only a notation difference that leads to a logic expression. Except for a few examples given below, we leave the use of these expressions to numerical methods when designing antennas.

Example   Use the magnetic vector potential to derive the far field of a short-length current element.

Assume a constant current on the wire. The current density is $Il\delta(r')$, where $\delta(r')$ *is* the Dirac delta distribution and l is the length over which the far-field phase is constant. The integral in Eq. (2-4) easily reduces to

$$A_z = \frac{\mu Il e^{-jkr}}{4\pi r}$$

The current element is so short that the phase distances from all parts of the element are considered to be equal; e^{-jkr} is the retarded potential phase term. The electric field is found from A_z using Eq. (2-1):

$$E_z = -j\omega\mu \frac{Il}{4\pi r} e^{-jkr}$$

$$E_\theta = E_z \hat{\mathbf{z}} \cdot \hat{\boldsymbol{\theta}} = j\omega\mu \frac{Il}{4\pi r} e^{-jkr} \sin\theta$$

Evaluate ω as $2\pi f$, split μ in $\sqrt{\mu}\sqrt{\mu}$, and divide and multiply by $\sqrt{\varepsilon}$:

$$E_\theta = \frac{jIl2\pi f \sqrt{\mu\varepsilon}}{4\pi r}\sqrt{\frac{\mu}{\varepsilon}}e^{-jkr}\sin\theta$$

The following terms can be recognized as

$$c = \frac{1}{\sqrt{\mu\varepsilon}} \qquad \frac{f}{c} = \frac{l}{\lambda} \qquad \eta = \sqrt{\frac{\mu}{\varepsilon}}$$

The far-field electric field becomes

$$E_\theta = \frac{jIl\eta}{2\lambda r}e^{-jkr}\sin\theta$$

The magnetic field is found from the electric field using Eq. (2-2):

$$H_\phi = \frac{E_\theta}{\eta} = \frac{jIl}{2\lambda r}e^{-jkr}\sin\theta$$

The term j can be evaluated as $e^{j\pi/2}$, a phase shift term. The power density S_r is

$$S_r = E_\theta H_\phi^* = \frac{|I|^2 l^2}{4\lambda^2 r^2}\eta\sin^2\theta$$

The normalized power pattern is equal to $\sin^2\theta$. Figure 2-1 gives the polar pattern of this antenna as a dashed plot. The dashed circle is the -3-dB pattern level. We measure the angular separation between the 3-dB points to determine the beamwidth (half-power beamwidth). For comparison, Figure 2-1 shows the pattern of a half-wavelength-long dipole as a solid curve. At a length about 5% shorter than a half wavelength, the reactive component of the impedance vanishes. The figure illustrates that a short dipole has about the same pattern as a long-resonant-length (reactance equals zero) dipole.

We determine directivity (Section 1-7) by calculating the average radiation intensity, often normalized to the peak of the power pattern:

$$U_{\text{avg}} = \int_0^{\pi/2} \sin^2\theta \sin\theta \, d\theta = \frac{2}{3}$$
$$U_{\text{max}} = 1$$
$$\text{directivity} = \frac{U_{\text{max}}}{U_{\text{avg}}} = 1.5 \quad (1.76\,\text{dB})$$

The resonant-length dipole ($\approx \lambda/2$) has a directivity of 2.15 dB or only 0.39 dB more than that of the very short dipole. The total power radiated by the antenna is found by integrating the Poynting vector magnitude over a sphere:

$$P_r = \int_0^{2\pi}\int_0^\pi S_r r^2 \sin\theta \, d\theta \, d\phi$$
$$= \int_0^{2\pi}\int_0^\pi \left(\frac{|I|l}{2\lambda}\right)^2 \eta\sin\theta \, d\theta \, d\phi$$
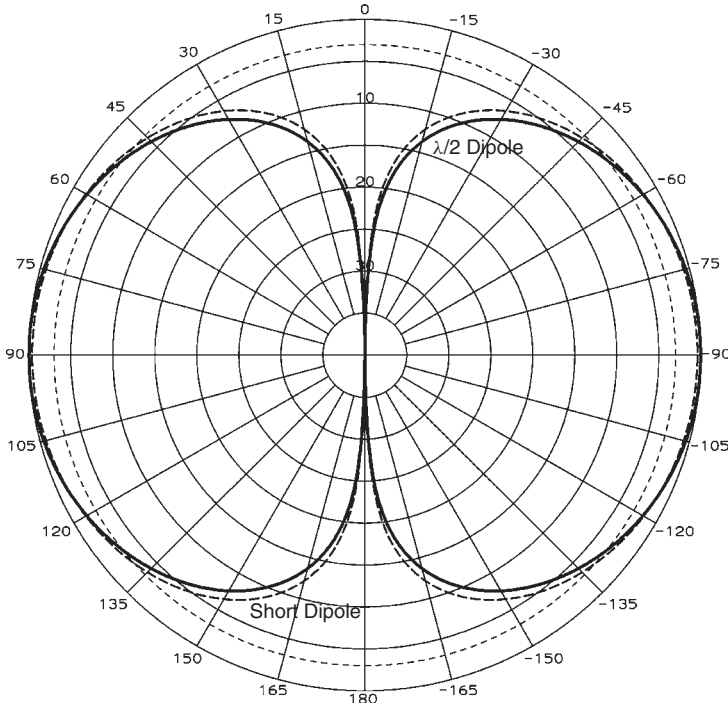$$= \frac{2\pi}{3}\left(\frac{|I|l}{2\lambda}\right)^2 \eta$$

FIGURE 2-1   Pattern of a short current element and small loop (dashed curve) compared to a $\lambda/2$-long dipole (solid curve) located along the 0 to $180°$ axis.

We represent the radiated power as a radiation resistance at the input of the antenna:

$$R_R = \frac{P_r}{|I|^2} = \frac{2\pi}{3}\eta\left(\frac{l}{\lambda}\right)^2$$

While a short dipole with a length $\lambda/20$ has a radiation resistance of about 2 Ω, a resonant-length dipole has about a 50-Ω radiation resistance and is more efficient because the relative material resistance is low.

   The input resistance of the antenna is the sum of the radiation resistance and the resistance due to material losses:

$$P_{\text{in}} = (R_R + R_L)|I|^2$$

The gain of an antenna is the ratio of the peak radiation intensity to the input power averaged over the radiation sphere:

$$\text{gain} = \frac{S_{r,\text{peak}}r^2}{\dfrac{P_{\text{in}}}{4\pi}} = \frac{U_{\text{max}}}{\dfrac{P_{\text{in}}}{4\pi}}$$

By using the idea of radiation resistance, we rewrite this as

$$\text{gain} = \frac{4\pi U_{\text{max}}}{(R_R + R_L)|I|^2}$$

Efficiency is the ratio of radiated power to input power:

$$\eta_e = \frac{P_r}{P_{\text{in}}} = \frac{R_r|I|^2}{(R_r + R_L)|I|^2} = \frac{R_r}{R_r + R_L}$$

Instead of integrating the pattern to calculate the total power radiated, we sometimes compute the input power of the antenna from currents induced on the antenna elements by given voltage sources on various terminals of the antenna in analysis:

$$P_{\text{in}} = \text{Re}(V_1 I_1^*) + \text{Re}(V_2 I_2^*) + \cdots + \text{Re}(V_N I_N^*)$$

The gain can be found from

$$\text{gain} = \frac{S_r(\theta, \phi)r^2}{\dfrac{P_{\text{in}}}{4\pi}} = \frac{U(\theta, \phi)}{\dfrac{P_{\text{in}}}{4\pi}}$$

This method is considerably easier than integrating the radiation intensity to compute directivity.

By integrating the pattern, we found only the input resistance of the short antenna, not the reactive component. A short antenna has a large capacitive reactance term that limits the impedance bandwidth when combined with a match network. The short antenna has a large pattern bandwidth but a narrow impedance bandwidth. Of course, an active network could be designed to impedance-match the antenna at any frequency, but the instantaneous bandwidth is narrow. The moment method of analysis gives us the currents for given input voltages and calculates the complete input impedance.

2-1.2   Radiation from Magnetic Currents

Magnetic currents are fictitious, but they enable slot radiation to be solved by the same methods as electric currents on dipoles by using duality. Slot radiation could be calculated from the surface currents around it, but it is easier to use magnetic currents to replace the electric field in the slot. Magnetic currents along the long axis of slots in ground planes replace the electric fields across the slots by application of the equivalence theorem. Similarly, current loops can be replaced by magnetic dipole elements to calculate radiation.

We use the electric vector potential **F** with magnetic currents. The far-field magnetic field is proportional to the electric vector potential:

$$\mathbf{H} = -j\omega\mathbf{F} \tag{2-9}$$

We determine the magnitude of the electric field from Eq. (2-2); it is perpendicular to **H**. The electric vector potential is found from a retarded volume integral over the magnetic current density **M**. Applying the radiation approximation, it is

$$\mathbf{F} = \frac{e^{-jkr}}{4\pi r}\varepsilon \iiint \mathbf{M}' e^{j\mathbf{k}\cdot\mathbf{r}'}\, dV' \tag{2-10}$$

where ε is the permittivity (8.854×10^{-12} F/m in free space). Equation (2-9) is the dual of Eq. (2-1), and Eq. (2-10) is the dual of Eq. (2-5). The dual of Eq. (2-3) is valid in both the near- and far-field regions.

The magnetic currents in a slot are perpendicular to the slot electric fields: $\mathbf{M} = \mathbf{E} \times \hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is the normal to the plane with the slot. The filamentary currents of thin slots reduce Eq. (2-10) to a line integral, and magnetic current direction limits the direction of the electric vector potential and the magnetic field. Since the electric field (far field) is orthogonal to the magnetic field, the electric field is in the same direction as the field across the slots. We use the direction of the electric field across the slots to estimate the polarization of the far field. As with filamentary electric currents, the far field is zero along the axis of the magnetic current.

The electric vector potential can also be used to derive the near field:

$$\mathbf{H} = -j\omega\mathbf{F} + \frac{\nabla(\nabla \cdot \mathbf{F})}{j\omega\mu\varepsilon}$$

$$\mathbf{E} = -\frac{1}{\varepsilon}\nabla \times \mathbf{F}$$

The magnetic field separates into near- and far-field terms in the electric vector potential; the electric field does not. We can determine the radiated fields directly in terms of the magnetic currents and avoid using the vector potential:

$$\mathbf{E}(\mathbf{r}) = -\frac{k^2}{4\pi}\iiint\limits_{V'} \mathbf{M}(\mathbf{r'}) \times \hat{\mathbf{R}}\left(\frac{j}{kR} + \frac{1}{k^2R}\right)e^{-jkR}\,dV' \qquad (2\text{-}11)$$

$$\mathbf{H}(\mathbf{r}) = \frac{k^2}{4\pi\eta}\iiint\limits_{V'} \left[\mathbf{M}(\mathbf{r'})\left(-\frac{j}{kR} - \frac{1}{k^2R^2} + \frac{j}{k^3R^3}\right)\right.$$

$$\left. +[\mathbf{M}(\mathbf{r'}) \cdot \hat{\mathbf{R}}]\hat{\mathbf{R}}\left(\frac{j}{kR} + \frac{3}{k^2R^2} - \frac{3j}{k^3R^3}\right)\right]e^{-jkR}\,dV' \qquad (2\text{-}12)$$

Equations (2-11) and (2-12) can be rearranged to find the dyadic Green's functions for magnetic currents. These differ from the dyadic Green's functions for electric currents by only constants.

Example Derive the fields radiated from a small constant-current loop.

We could use the magnetic vector potential and calculate over the currents in the wire but must account for changing current direction around the loop. Place the loop in the x–y plane. The electric field radiated by the loop is in the ϕ direction because the currents in the loop can only be in the $\hat{\boldsymbol{\phi}}$ direction. When solving the integral for the magnetic vector potential, note that the direction of the current on the loop, $\hat{\boldsymbol{\phi}}'$ at a general point is not in the same direction as the field point, $\hat{\boldsymbol{\phi}}$, unit vector. The integral must be solved with a constant vector direction, one component at a time.

Although the magnetic vector potential can be computed, it is easier to replace the current loop with an incremental magnetic current element. The equivalent magnetic current element is

$$I_m l = j\omega\mu I A$$

where A is the area of the loop. The magnetic current density is

$$\mathbf{M} = I_m l\delta(r')\hat{\mathbf{z}} = j\omega\mu I A\delta(r')\hat{\mathbf{z}}$$

The electric vector potential is found using Eq. (2-10):

$$F_z = \frac{j\omega\mu\varepsilon I A}{4\pi r} e^{-jkr}$$

The magnetic field is found from this electric vector potential using Eq. (2-9):

$$H_z = -j\omega F_z = \frac{\omega^2 \mu\varepsilon I A}{4\pi r} e^{-jkr}$$

We calculate H_θ by projection:

$$H_\theta = H_z \hat{\mathbf{z}} \cdot \hat{\boldsymbol{\theta}} = -\frac{\omega^2 \mu\varepsilon I A}{4\pi r} e^{-jkr} \sin\theta$$

E_ϕ and H_θ are related in the far field because the wave propagates in the **r** direction:

$$E_\phi = -\eta H_\theta = \frac{\omega^2 \mu\varepsilon I A \eta}{4\pi r} e^{-jkr} \sin\theta$$

The small current loop and small current element have the same pattern shape, $\sin\theta$, but opposite polarizations. The directivity is 1.5 (1.76 dB). Figure 2-1 uses a dashed curve to plot the response of the small loop, while the solid curve gives the pattern of a half-wavelength slot that radiates on both sides of the ground sheet.

2-2   APERTURES: HUYGENS SOURCE APPROXIMATION

Many antennas, such as horns or paraboloid reflectors, can be analyzed simply as apertures. We replace the incident fields in the aperture with a combination of equivalent electric and magnetic currents. We calculate radiation as a superposition of each source by using the vector potentials. Most of the time, we assume that the incident field is a propagating free-space wave whose electric and magnetic fields are proportional to one another. This gives us the Huygens source approximation and allows the use of integrals over the electric field in the aperture. Each point in the aperture is considered to be a source of radiation. The far field is given by a Fourier transform of the aperture field:

$$\mathbf{f}(k_x, k_y) = \iint\limits_S \mathbf{E} e^{j\mathbf{k}\cdot\mathbf{r}'} \, ds' \tag{2-13}$$

This uses the vector propagation constant

$$\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{z}}$$

$$k_x = k \sin\theta \cos\phi \qquad k_y = k \sin\theta \sin\phi \qquad k_z = k \cos\theta$$

where $\mathbf{f}(k_x, k_y)$ is the pattern in k-space. We multiply the Fourier transform far field by the pattern of the Huygens source:

$$\frac{j e^{-jkr}}{2\lambda r}(1 + \cos\theta) \tag{2-14}$$

When apertures are large, we can ignore this pattern factor. In Eq. (2-13), $\mathbf{f}(k_x, k_y)$ is a vector in the same direction as the electric field in the aperture. Each component

is transformed separately. The far-field components E_θ and E_ϕ are found by projection (scalar products) from $\mathbf{f}(k_x, k_y)$ times the pattern factor of the Huygens source [Eq. (2-14)].

If we have a rectangular aperture in which the electric field is expressed as a product of functions of x and y only, the integral reduces to the product of two single integrals along each coordinate. The Fourier transform relationships provide insight into pattern shape along the two axes. Large apertures radiate patterns with small beamwidths. An antenna with long and short axes has a narrow-beamwidth pattern in the plane containing the long dimension and a wide beamwidth in the plane containing the short dimension. This is the same as the time and frequency dual normally associated with the Fourier transform.

We draw on our familiarity with signal processing to help us visualize the relationship between aperture distributions and patterns. Large apertures give small beamwidths, just as long time pulses relate to low-frequency bandwidths in normal time–frequency transforms. The sidelobes of the pattern correspond to the frequency harmonics of an equivalent time waveform under the Fourier transform and rapid transitions in the time response lead to high levels of harmonics in the frequency domain. Rapid amplitude transitions in the aperture plane produce high sidelobes (harmonics) in the far-field response (Fourier transform). Step transitions on the aperture edges produce high sidelobes, while tapering the edge reduces sidelobes and we relate the sidelobe envelope of peaks to the derivative of the distributions at the edges. To produce equal-level sidelobes, we need Dirac delta functions in the aperture that transform to a constant level in the pattern domain. Another example is periodic aperture errors that produce single high sidelobes. When we discuss aperture distribution synthesis, we see that the aperture extent in wavelengths limits our ability to control the pattern.

A uniform amplitude and phase aperture distribution produces the maximum aperture efficiency and gain that we determine from the following argument. An aperture collects power from a passing electromagnetic wave and maximum collectible power occurs at its peak amplitude response. If the amplitude response somewhere else in the aperture is reduced from the maximum, that portion will collect less power. The amplitude response can be reduced only by adding loss or reflecting power in reradiation. The antenna with the highest aperture efficiency reflects the least amount of power when illuminated by a plane wave. Similarly, if the phase shift from the collecting aperture to the antenna connector is different for different parts of the aperture, the voltages from the various parts will not add in phase. Gain is directly proportional to aperture efficiency [Eq. (1-10)]. Therefore, a uniform amplitude and phase aperture distribution has maximum gain. All this assumes that the input match on various aperture distribution antennas is the same.

For example, consider the pattern of a uniform aperture distribution in a rectangular aperture $a \times b$. We use the Fourier transform and ignore the polarization of the electric field in the aperture. (This assumes that the field has a constant polarization or direction.)

$$f(k_x, k_y) = E_0 \int_{-b/2}^{b/2} \int_{-a/2}^{a/2} e^{j\mathbf{k} \cdot \mathbf{r}'} \, dx \, dy$$

$$= E_0 \int_{-b/2}^{b/2} \int_{-a/2}^{a/2} e^{jk_x x} e^{jk_y y} dx \, dy$$

We separate the integral into a product of two integrals each with the form

$$\int_{-a/2}^{a/2} e^{jk_x x}\, dx = \frac{e^{jk_x a/2} - e^{-jk_x a/2}}{jk_x} = \frac{a\sin(k_x a/2)}{k_x a/2}$$

On combining the two similar integrals, we have

$$f(k_x, k_y) = \frac{ab\sin(k_x a/2)}{k_x a/2} \frac{\sin(k_y b/2)}{k_y b/2}$$

where $k_x = k\sin\theta\cos\phi$, $k_y = k\sin\theta\sin\phi$, and $k_z = k\cos\theta$ and $k = 2\pi/\lambda$. The pattern in both planes is given by a k-space function, $\sin u/u$. Figure 2-2 plots this pattern function as a solid curve using k_x-space [$(ka/2)\sin\theta$] as the abscissa to produce a universal curve independent of aperture size a. The half-power points occur when

$$\frac{\sin u}{u} = \frac{1}{\sqrt{2}} \quad \text{or} \quad u = 1.39156$$

When we substitute for u, we have in the principal planes

$$\frac{\pi a}{\lambda}\sin\theta = 1.39156$$

By solving for θ, we compute the half-power beamwidth (HPBW):

$$\text{HPBW} = 2\sin^{-1}\frac{0.4429\lambda}{a}$$

By using the approximation $u = \sin u$ for small angles, the half-power beamwidth can be estimated as

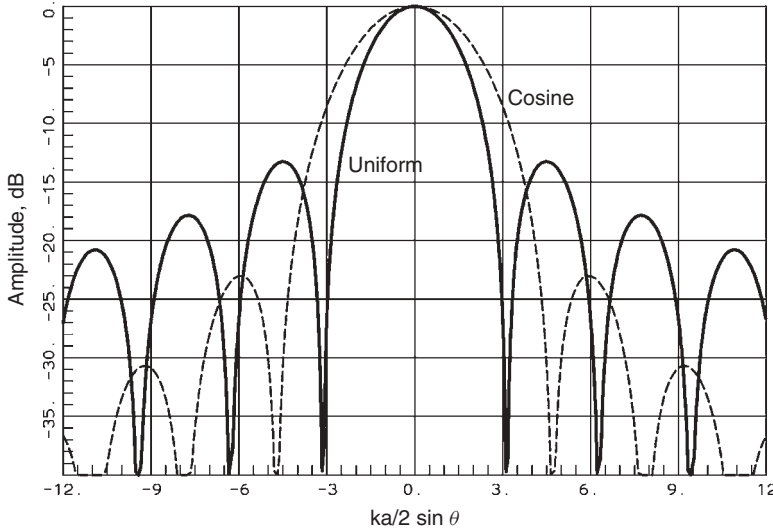$$\text{HPBW} = 50.76°\frac{\lambda}{a}$$

FIGURE 2-2 Universal k-space pattern for the radiation from uniform (solid curve) and cosine (dashed curve) aperture distributions.

Note that we have ignored the $(1 + \cos\theta)/2$ pattern of the Huygens source, which reduces the beamwidth for radiation from small apertures. We discover on Figure 2-2 the pattern nulls occur at integer multiples of π and the first sidelobe amplitude is 13.2 dB below the peak.

The gain of a uniform amplitude and phase aperture distribution is given by Eq. (1-7), where A is the area of the aperture. Chapter 4 develops amplitude taper efficiency for nonuniform aperture amplitude distributions to calculate the gain reduction. Phase error efficiency gives the gain reduction due to phase anomalies. Each of these is found from the distribution of aperture fields. Figure 2-2 plots the pattern of a half cosine aperture distribution as a dashed curve. The distribution peaks in the center and tapers linearly to zero at the edges. Tapering the aperture distribution widens the beamwidth and reduces both gain and sidelobe levels. The pattern beamwidth is 1.342 times wider than the uniform distribution beamwidth. A cosine distribution produces a -0.91-dB amplitude taper loss, and the distribution edge taper causes the sidelobes to fall off at a faster rate.

Example   Compute the length of the aperture with a uniform distribution that will give a $10°$ beamwidth.

$$\frac{a}{\lambda} = \frac{50.76°}{10°} \simeq 5 \text{ wavelengths}$$

We can calculate radiated power by integrating the Poynting vector magnitude over the radiation sphere, but there is an easier way. We assumed that the aperture fields are free-space waves. The total power radiated is in the aperture,

$$P_r = \iint\limits_{\text{aperture}} \frac{|E|^2}{\eta}\, ds \qquad P_{\text{avg}} = U_{\text{avg}} = \frac{P_r}{4\pi}$$

where η is the impedance of free space. The radiated electric field is

$$\mathbf{E} = j\frac{e^{-jkr}(1 + \cos\theta)}{2\lambda r}\mathbf{f}(k_x, k_y)$$

The Poynting vector magnitude is

$$S_r = \frac{|\mathbf{E}|^2}{\eta} = \frac{(1 + \cos\theta)^2}{4\lambda^2 r^2}|\mathbf{f}(k_x, k_y)|^2 \qquad\qquad (2\text{-}15)$$

By combining Eqs. (2-14) and (2-15), we determine directivity:

$$\text{directivity}(\theta, \phi) = \frac{U(\theta, \phi)}{U_{\text{avg}}} = \frac{S_r r^2}{P_r/4\pi}$$

$$= \frac{\pi(1 + \cos\theta)}{\lambda^2}\frac{\left|\iint E e^{j\mathbf{k}\cdot\mathbf{r}'}\, ds'\right|^2}{\iint |E|^2 ds'} \qquad\qquad (2\text{-}16)$$

By considering electric and magnetic fields separately in the aperture, we eliminate the requirement that the ratio electric and magnetic fields are the same as free space used in the Huygens source approximation. Given the fields in an aperture, we can equate

them to magnetic and electric currents:

$$\mathbf{M}_s = \mathbf{E}_a \times \mathbf{n} \qquad \mathbf{J}_s = \mathbf{n} \times \mathbf{H}_a \qquad (2\text{-}17)$$

where $\mathbf{E}_a$ and $\mathbf{H}_a$ are the aperture fields and $\mathbf{n}$ is the outward normal. The equivalence theorem [2, p. 113] results in exact solutions by using the total aperture field, incident and reflected. When using the equivalence theorem, we replace the total fields present with equivalent currents. The induction theorem equates currents only to the incident fields on the aperture, which ignores wave reflection and results in approximate solutions:

$$\mathbf{F} = \varepsilon \iint\limits_s \frac{\mathbf{M}_s e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}\, ds' \qquad \mathbf{A} = \mu \iint\limits_s \frac{\mathbf{J}_s e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}\, ds' \qquad (2\text{-}18)$$

We derive the radiated fields from each distribution of currents by using vector potentials where $\mathbf{r}$ is the field point and $\mathbf{r}'$ is the source point in the aperture. These expressions are valid in the near and far fields. By integrating over only a finite aperture, we assume zero fields outside the aperture, while rigorous expressions require integrals over closed boundaries. A planar aperture must extend to infinity, but the fields outside the aperture are nearly zero and contribute little.

2-2.1 Near- and Far-Field Regions

The radiative near- and far-field regions are characterized by the approximations made to the integrals [Eq. (2-18)]. The radiative near-field region lies between the near field, with no approximations, and the far-field region. In both approximations the field (observation) distance r is substituted for $|\mathbf{r}-\mathbf{r}'|$ in the amplitude term. The vector potentials reduce to

$$\mathbf{F} = \frac{\varepsilon}{4\pi r} \iint\limits_s \mathbf{M}_s e^{-jk|\mathbf{r}-\mathbf{r}'|}\, ds' \qquad \mathbf{A} = \frac{\mu}{4\pi r} \iint\limits_s \mathbf{J}_s e^{-jk|\mathbf{r}-\mathbf{r}'|}\, ds' \qquad (2\text{-}19)$$

We handle the phase term differently in the two regions. First, we expand the phase term in a Taylor series,

$$|\mathbf{r}-\mathbf{r}'| = \sqrt{r^2 + r'^2 - 2\mathbf{r}\cdot\mathbf{r}'} = r - \hat{\mathbf{r}}\cdot\mathbf{r}' + \frac{1}{2r}[r'^2 - (\hat{\mathbf{r}}\cdot\mathbf{r}')^2]\cdots$$

where $\hat{\mathbf{r}}$ is the unit vector in the field point direction. We retain the first two terms for the far-field approximation and the vector potentials become

$$\mathbf{F} = \frac{e^{-jkr}\varepsilon}{4\pi r} \iint\limits_s \mathbf{M}_s e^{j\mathbf{k}\cdot\mathbf{r}'}\, ds', \qquad \text{etc.} \qquad (2\text{-}20)$$

where we have combined k, the propagation constant, with the unit vector $\hat{\mathbf{r}}$:

$$\mathbf{k} = k\hat{\mathbf{r}} = k(\sin\theta\cos\phi\hat{\mathbf{x}} + \sin\theta\sin\phi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}})$$

The magnetic vector potential integral parallels Eq. (2-20) as in Eq. (2-19). In the radiative near-field zone approximation the terms in r'^2 are retained and we obtain the following integral for the electric vector potential:

$$\mathbf{F} = \frac{e^{-jkr}\varepsilon}{4\pi r} \iint \mathbf{M}_s \exp\left[j(\mathbf{k}\cdot\mathbf{r}') + \frac{(\mathbf{k}\cdot\mathbf{r}')^2}{2rk} - \frac{kr'^2}{2r} \right] ds' \qquad (2\text{-}21)$$

No clear boundary between the three regions exists because the fields are continuous. Common boundaries are

$$\frac{r}{L} < 1 \qquad\qquad \text{near field}$$

$$1 < \frac{r}{L} < \frac{L}{\lambda} \qquad \text{radiative near field}$$

$$\frac{r}{L} > \frac{L}{\lambda} \qquad\qquad \text{far field}$$

where L is the maximum dimension on the aperture.

Example   Determine the maximum difference between the radiative near- and far-field approximations at a point normal to the maximum aperture dimension when $r = L^2/\lambda$ and $r = 2L^2/\lambda$.

   Normal to the maximum dimension, $\hat{\mathbf{r}} \cdot \mathbf{r}' = 0$. The phase difference is

$$\frac{kr_{\max}'^2}{2r} \qquad \text{where} \quad r_{\max}' = \frac{L}{2}$$

$$\text{phase difference } \phi = \frac{2\pi L^2}{8\lambda r}$$

$$\phi = \pi/4 \quad \text{at} \quad r = L^2/\lambda \quad \text{and} \quad \phi = \pi/8 \quad \text{at} \quad r = 2L^2/\lambda$$

The usual minimum distance used for antenna patterns is $2L^2/\lambda$, where L is the maximum dimension of the antenna. At that distance, the phase error across the aperture from a point source antenna is $\pi/8$. The distance is not sufficient for low-sidelobe antennas [3] because quadratic phase error raises the measured sidelobes.

   We can use vector potentials in the aperture after determining equivalent currents, but we will find it more convenient to use the fields directly. Define the following integrals:

$$\mathbf{f} = \iint_s \mathbf{E}_a e^{j\mathbf{k}\cdot\mathbf{r}'}\, ds \qquad \mathbf{g} = \iint_s \mathbf{M}_a e^{j\mathbf{k}\cdot\mathbf{r}'}\, ds \qquad\qquad (2\text{-}22)$$

using the far-field approximation. Near-field integrals require additional phase terms. Given an aperture, we calculate the vector potentials in terms of $\mathbf{E}_a$ and $\mathbf{H}_a$ through the currents by using either the equivalence or inductance theorems, and we use the integrals of Eq. (2-22) in the vector potentials. We combine the fields in the far field due to each partial source:

$$\mathbf{E} = -j\omega\mathbf{A} - j\eta\omega\mathbf{F} \times \hat{\mathbf{r}}$$

For an aperture in the $x-y$ plane, we carry out these steps by using the inductance theorem and obtain the following far fields from the incident aperture fields

$$E_\theta = \frac{jke^{-jkr}}{4\pi r}[f_x \cos\phi + f_y \sin\phi + \eta \cos\theta(-g_x \sin\phi + g_y \cos\phi)]$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2\text{-}23)$$

$$E_\phi = \frac{-jke^{-jkr}}{4\pi r}[(f_x \sin\phi - f_y \cos\phi)\cos\theta + \eta(g_x \cos\phi + g_y \sin\phi)]$$

where **f** and **g** have been expanded in terms of their x- and y-components and η is the impedance of free space.

2-2.2   Huygens Source

The Huygens source approximation is based on the assumption that the magnetic and electric fields are related as in a plane wave in the aperture:

$$\eta g_y = f_x \quad \text{and} \quad -\eta g_x = f_y$$

since

$$\eta H_y = E_x \quad \text{and} \quad -\eta H_x = E_y$$

With this approximation, the far field [Eq. (2-23)] becomes

$$E_\theta = \frac{jke^{-jkr}}{4\pi r}(1 + \cos\theta)(f_x \cos\phi + f_y \sin\phi)$$

$$E_\phi = \frac{-jke^{-jkr}}{4\pi r}(1 + \cos\theta)(f_x \sin\phi - f_y \cos\phi)$$

$$(2\text{-}24)$$

The two-dimensional vector Fourier transform $\mathbf{f} = (f_x, f_y)$ of the aperture electric field in the $x-y$ plane determines the far-field components. We derive the radiated components by projecting (vector scalar product) this field onto the vectors $\hat{\theta}/\cos\theta$ and $\hat{\phi}$. The transform $\mathbf{f}$ expands the field in k-space [usually, (k_x, k_y)]. This normalizes the pattern and removes the direct dependence on aperture length.

We separate out all but $\mathbf{f}$ when we consider aperture distributions. We drop the terms for the radiation from a point source and the pattern of a Huygens point source [Eq. (2-14)] and limit our discussions to Huygens sources and far fields. General aperture fields require Eq. (2-23), and for any region other than the far field, additional phase terms are needed in the transforms [Eq. 2-21)].

2-3   BOUNDARY CONDITIONS

Material boundaries cause discontinuities in the electric and magnetic fields. The effects can be found by considering either vanishing small pillboxes or loops that span the boundary between the two regions. By using the integral form of Maxwell's equations on these differential structures, the integrals reduce to simple algebraic expressions. These arguments can be found in most electromagnetic texts and we give only the results. Conversely, we will discover that artificial boundaries such as shadow and reflection boundaries used in geometric optics (ray optics) cannot cause a discontinuity in the fields because they are not material boundaries. The idea that the fields remain continuous across the boundary leads to the necessity of adding terms to extend ray optics methods. We discuss these ideas when considering the uniform theory of diffraction (UTD) method used with ray optics.

Suppose that we have a locally plane boundary in space described by a point and a unit normal vector $\hat{\mathbf{n}}$ that points from region 1 to region 2. We compute the tangential fields from the vector (cross) product of the fields and the normal vector. The fields

can be discontinuous at the interface between the two regions if surface magnetic $\mathbf{M}_S$ or electric current $\mathbf{J}_S$ densities exist on the surface.

$$\hat{\mathbf{n}} \times (\mathbf{E}_2 - \mathbf{E}_1) = -\mathbf{M}_S \qquad \hat{\mathbf{n}} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{J}_S \qquad (2\text{-}25a,b)$$

The normal components of the fields change due to the differing dielectric and magnetic properties of the materials and the charges induced on the surface:

$$\hat{\mathbf{n}} \cdot (\varepsilon_2 \mathbf{E}_2 - \varepsilon_1 \mathbf{E}_1) = \rho_S \qquad \hat{\mathbf{n}} \cdot (\mu_2 \mathbf{H}_2 - \mu_1 \mathbf{H}_1) = \tau_S \qquad (2\text{-}26)$$

with ρ_S and τ_S given as electric and magnetic surface charge densities, respectively. Perfect dielectric and magnetic materials can have no currents, which reduces Eq. (2-25) to

$$\hat{\mathbf{n}} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0 \qquad \hat{\mathbf{n}} \times (\mathbf{H}_2 - \mathbf{H}_1) = 0 \qquad (2\text{-}27)$$

Equation (2-27) means that the tangential fields are continuous across the boundary.

   These boundary conditions are used in the method of moment analyses to determine currents. The method applies the boundary condition in integral equations to determine the coefficients of the expansion of currents in the sum of basis functions. The currents described as these sums do not satisfy the boundary conditions at all points but do when integrated over a region. This method leads to approximations that will converge as more terms are included in the expansions.

   When doing analysis we find two types of surfaces convenient. We use these surfaces to reduce analysis effort by using planes of symmetry. The first one is the *perfect electric conductor* (PEC). A PEC surface causes the fields to vanish inside and to have electric currents induced on it:

$$\hat{\mathbf{n}} \times \mathbf{E}_2 = 0 \qquad \hat{\mathbf{n}} \times \mathbf{H}_2 = \mathbf{J}_S \qquad (\text{PEC}) \qquad (2\text{-}28a,b)$$

A PEC surface is also called an *electric wall*. The second surface is the *perfect magnetic conductor* (PMC) and is a hypothetical surface. Whereas good conductors approximate PEC, there are no PMC materials. The PMC has no internal fields like the PEC and forces the tangential magnetic field to be zero:

$$\hat{\mathbf{n}} \times \mathbf{E}_2 = -\mathbf{M}_S \qquad \hat{\mathbf{n}} \times \mathbf{H}_2 = 0 \qquad (\text{PMC}) \qquad (2\text{-}29)$$

A PMC surface supports the hypothetical magnetic current density $\mathbf{M}_S$. We find that the magnetic wall (PMC) concept simplifies analysis.
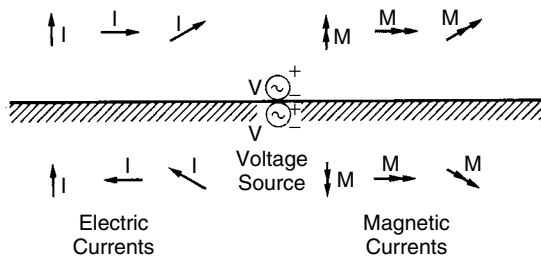

FIGURE 2-3   Ground-plane images.

We use images of currents to include material boundaries in analysis. Figure 2-3 illustrates ground-plane images. When we analyze radiation from currents in the presence of a boundary, we include the actual antenna and its image to compute the fields. The figure shows an infinite ground plane, but a finite ground-plane image can be used in the angular region where a reflected wave occurs in the finite plane. We consider this idea further when discussing geometric optics. We can use images in dielectric boundaries provided that we calculate the polarization sensitive reflection coefficients to adjust the magnitude and phase of the image.

2-4   PHYSICAL OPTICS

Physical optics uses things that can be measured. We can measure both currents and fields, but auxiliary vector potentials have no physical reality, only mathematical artifacts that simplify Maxwell's equations. Nevertheless, the auxiliary vector potentials provide simple models for problems that enable simple mental pictures, as shown earlier, but we cannot easily formulate them into a systematic analysis tool for antenna problems.

The physical optics analysis method combines the use of Green's functions to calculate fields radiated by a given distribution of currents and then uses boundary conditions to determine the currents induced on objects due to incident fields. We compute the effects of a mounting structure by inducing currents on it and adding their radiation to the antenna pattern. The method assumes that radiation from the induced currents on the structure does not change the initial currents.

We start analyses from either currents or incident fields and work from those. The resonant structure of many antennas determines the approximate current distribution that we normalize to the radiated power. We calculate the fields from these currents. Physical optics can use an iterative technique to calculate incremental currents induced on the original radiators and improve the solution, but we usually just sum the radiation from the original currents to the radiation from the induced currents. The second starting point for physical optics can be incident fields. These could be plane waves or could be fields found from the measured radiation patterns of antennas: for example, the pattern of a reflector feed. We add the radiation from the induced currents to the incident waves.

2-4.1   Radiated Fields Given Currents

The radiated fields can be found from distribution of the electric and magnetic currents by the use of dyadic Green's functions that contain source and field coordinates. We sometimes refer to the Green's functions as *vector propagators* or *transfer functions* between currents and fields. We calculate the fields from integrals over the source points of the dot (scalar) product between the dyadic and current densities. The dyadic Green's function contains both near- and far-field terms and requires slightly different expressions for the electric and magnetic fields. The general propagator from electric and magnetic currents has separate terms for electric and magnetic currents, which when used with surface patch currents can be reduced to short subroutines or procedures easily programmed [1]:

$$\mathbf{E}(\mathbf{r}) = \int \mathbf{G}_{\mathrm{EJ}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}(\mathbf{r}') \, dV' + \int \mathbf{G}_{\mathrm{EM}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{M}(\mathbf{r}') \, dV' \qquad (2\text{-}30)$$

$$\mathbf{H}(\mathbf{r}) = \int \mathbf{G}_{\mathrm{HJ}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}(\mathbf{r}') \, dV' + \int \mathbf{G}_{\mathrm{HM}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{M}(\mathbf{r}') \, dV' \qquad (2\text{-}31)$$

These expressions integrate over the currents located at source points $\mathbf{r}'$ for a dyadic Green's function that changes at each field point $\mathbf{r}$ and source point $\mathbf{r}'$. Although these Green's functions are valid at all field points in space both near and far field, they are singular at a source point. Only retaining terms with $1/R$ dependence for the far field greatly simplifies the expressions.

When fields are incident on a perfect electric conductor (PEC), the combination of incident and reflected tangential magnetic fields induces an electric current density on the surface. The fields inside the conductor are zero. We assume locally plane surfaces on patches and compute currents that satisfy the boundary condition. Given the local unit normal $\hat{\mathbf{n}}$ to the surface, the induced current density is given by

$$\mathbf{J}_S = \hat{\mathbf{n}} \times (\mathbf{H}_{\text{incident}} + \mathbf{H}_{\text{reflected}})$$

$$\mathbf{H}_{\text{incident}} = \mathbf{H}_{\text{reflected}}$$

$$\mathbf{J}_S = 2\hat{\mathbf{n}} \times \mathbf{H}_{\text{incident}} \tag{2-32}$$

The reflected magnetic field equals the incident magnetic field because the field reflects from the conductive surface. The sum of the tangential electric fields must be zero. Because the reflected wave changes direction, the vector (cross) product of the electric and magnetic fields must change direction. The reflected tangential electric field changes direction by $180°$, so the tangential magnetic field must not change direction because the Poynting vector changed its direction. Equation (2-32) is the magnetic field equation applied on a PEC. Equation (2-25b) is the general magnetic field equation at a boundary.

Physical optics starts with a given current distribution that radiates, or the measured pattern of an antenna. When an object is placed in the radiated field, the method calculates induced current on the object to satisfy the internal field condition. For example, PEC or PMC have zero fields inside. When we use simple functions such as constant-current surface patches, the sum of the radiation from the incident wave and the scattered fields from induced surface currents produces only approximately zero fields inside. As the patch size decreases, the method converges to the correct solution. To obtain the radiated field everywhere, we sum the incident wave and scattered waves. The fields radiated by the induced currents produce the shadow caused by the object. With geometric optics techniques such as UTD, the object blocks the incident wave and we determine the fields in the shadow regions from separate diffraction waves. In physical optics the incident wave continues as though the object were not present. Only geometric optics techniques use blockage.

We can calculate the fields radiated from antennas in free space or measure them in an anechoic chamber that simulates free space, but we mount the antenna on finite ground planes, handsets, vehicles, over soil, and so on, when we use them. Physical optics is one method of accounting for the scattering. We show in later chapters that the mounting configuration can enhance the patterns.

2-4.2   Applying Physical Optics

In this book we do not discuss how to develop numerical techniques, but it is important to understand how to apply methods. Whether you develop your own codes or use commercial codes, certain rules should be applied. Consider Eq. (2-32). The normal to

the surface points in the direction of the incident wave: outward. If the normal pointed inward, the sign of induced electric current density would change. Most codes have made the assumption that the normal points outward, but some codes may check on the direction of the normal relative to the incident wave and make the necessary sign change. We must keep track of the direction of the normal, and it may be necessary to rotate the normal depending on the expected direction of the incident wave. If an object can have radiation from both sides, it may be necessary to use two objects in the analysis.

Many codes store each object as a separate entity in a disk file. In some cases we need to store an object multiple times. Take, for example, a Cassegrain dual reflector. The feed antenna illuminates the subreflector and induces currents on it. These currents radiate and excite currents on the main reflector. When the main reflector–induced currents radiate, the subreflector intercepts or blocks part of the fields. We account for this blockage by using a second subreflector object on which the code calculates a new set of induced currents by using the main reflector currents as the source. We could add these currents to the existing disk file object or merely keep the second object. We want to keep the second object separate so that we can calculate additional currents induced on the main reflector using these currents as sources. These currents will be reduced from the initial set, but they are an important contribution to the fields radiated behind the reflector. This example illustrates iterative PO. When objects face each other significantly, iterative PO is necessary to calculate correct patterns. The method converges rapidly in most cases.

Figure 2-4 illustrates the geometry of a corner reflector. A half-wavelength-long dipole is placed between two metal plates usually bent to form a 90° angle. We can use other angular orientations between the plates, but this is the usual design. The figure does not show the feed line to the dipole, which usually starts at the juncture of the two plates and runs up to the dipole. This feed line contains the balun discussed in Section 5-15. Although the figure shows the plates as solid, many implementations use metal rods to reduce weight and wind loading.

The analysis starts with assumed currents on the dipole. We divide the plates analytically into small rectangular patches, which can be small ($\approx \lambda/8$ to $\lambda/4$) on a side since it takes only a few to cover the plates. You should repeat the analysis with
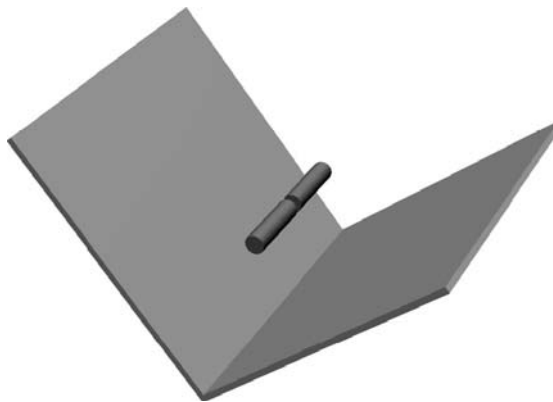

FIGURE 2-4   Corner reflector with a dipole located between two flat plates.

different-sized current patches to determine if the analysis has converged. In a similar manner, we break down the current on the dipole into short linear segments, each with constant amplitude. By using a near-field version of Eq. (2-31), we calculate the magnetic field incident on each patch on the plates. This field induces electric currents on the plates calculated from Eq. (2-32). Remember that we combine radiation from the source dipole with that radiated from the induced currents to reduce the radiation behind the antenna. The currents were induced to satisfy the boundary condition of the plate, but only with both radiations present. Figure 2-5a illustrates this process of inducing currents. Figure 2-6 shows the antenna pattern calculated using these currents. The E-plane pattern drawn as a solid line produces a null at $90°$ because the dipole pattern has this null. The plates cause the narrowing of the beam in the H-plane. The plates reduced the back radiation to -22 dB relative to the forward radiation, called the *front-to-back ratio* (F/B). The gain has increased from the 2.1 dB expected from a dipole to 9.3 dB. An equivalent geometric optics analysis uses two images in the plates, as shown in Figure 2-5b, for the analysis.

If you look at Figure 2-4 or 2-5, you should notice that the two plates face each other. Currents on one plate will radiate toward the other plate and induce another set of currents on it. We could ignore these induced currents if the radiation was insignificant, but to produce correct patterns we must include them. The solution to this problem calls for an iterative technique where we calculate the radiation from the currents on the first plate and induce incremental currents on the second plate. These incremental currents produce further radiation that induces additional currents on the other plate. The method converges rapidly. Figure 2-7 gives the antenna pattern after the iterations have been completed and we include radiation from all currents. The actual F/B ratio of the antenna is 29 dB, and the additional currents increased the gain by 0.7 dB to 10 dB. Adding the two plates in the original analysis increased the gain by 7.2 dB, whereas the iterative technique had a much smaller effect. Figure 2-8 illustrates the iterative technique and shows that the equivalent geometric optics analysis adds a third image to represent the reflection between the plates. Remember when you mount the antenna in an application, the structure will change the realized pattern, but the high F/B ratio reduces this effect. The mounting structure used when measuring the antenna changes the pattern as well, which limits our knowledge of the real pattern.
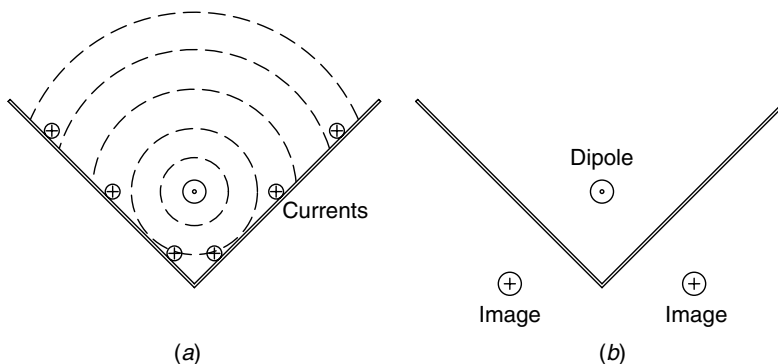

FIGURE 2-5   Cross-sectional view of a corner reflector: (*a*) magnetic field radiated from a dipole induces currents on plates; (*b*) plate currents replaced with image dipoles.
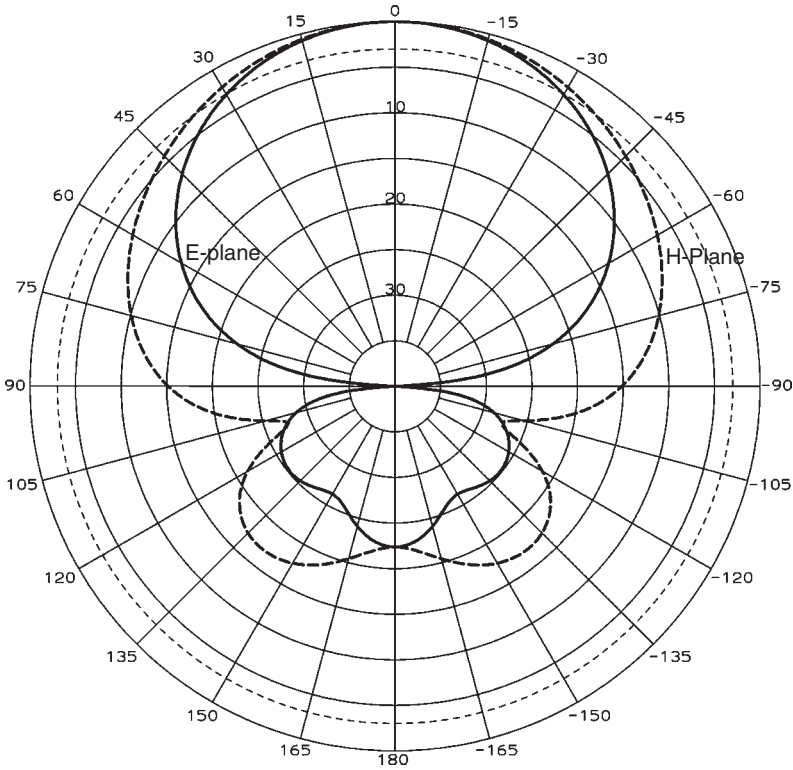
FIGURE 2-6 Pattern calculated from a combination of dipole and plate currents in a corner reflector with $1 \times 0.9\lambda$ plates without induced current iteration.

Physical optics can determine the impedance effects of the limited images in the ground planes, such as the corner reflector. The local nature of impedance effects allows the use of images to calculate the mutual impedance effects of ground planes. We use impedance calculations not only to determine the bounds of ground-plane effects on input impedance, but to calculate the total power radiated by the antenna. The images (excited currents on ground planes) radiate but do not receive input power. A ground plane at least $\lambda/2$ on a side located about $\lambda/4$ away from the antenna produces nearly the same impedance effects as an infinite ground plane, but the ground plane alters the radiation pattern greatly because it restricts possible radiation directions.

It has commonly been thought that physical optics could compute the field only in the main beam pattern direction of a paraboloidal reflector. The method can determine this pattern region accurately by using only a few patches, each one being many wavelengths on a side. As the processing power of computers increases, the patch size can be shrunk until PO can calculate the pattern in every direction, including behind the reflector. It is important to remember to include the feed pattern behind the reflector even though its radiation is obviously blocked by the main reflector. Physical optics uses induced currents to cancel the fields inside objects when the incident fields and the radiation from the induced currents are added. We can calculate the pattern behind a reflector using UTD (GTD), the uniform (geometric) theory of diffraction. This
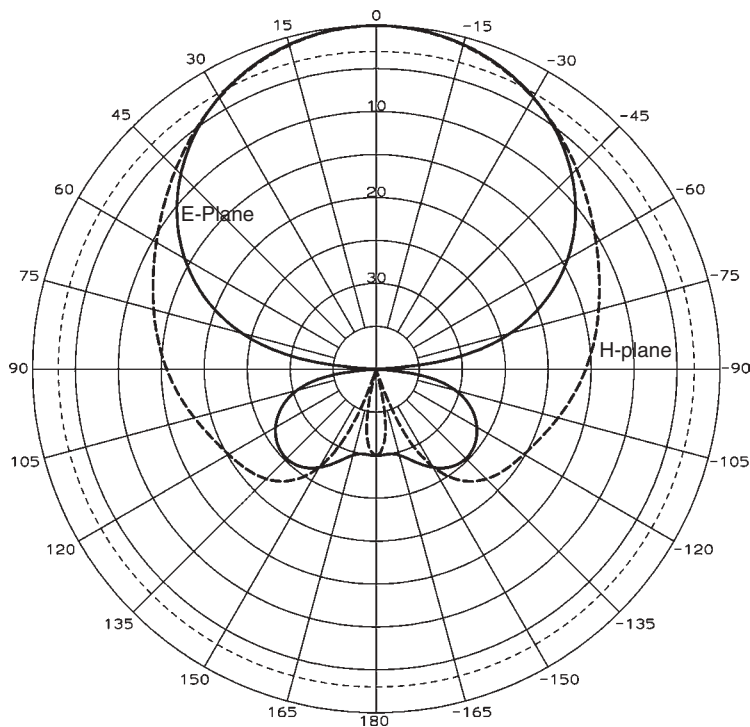
FIGURE 2-7   Pattern of corner reflector with $1 \times 0.9\lambda$ plates with induced current iteration equivalent to multiple reflectors between the plates.
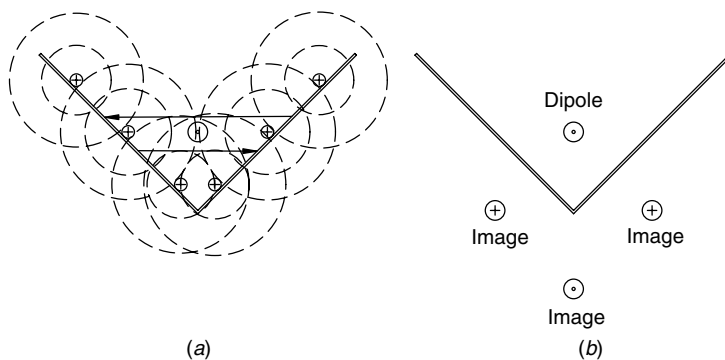

FIGURE 2-8   (*a*) Wall currents on plates radiate magnetic fields that induce additional currents on facing plates; (*b*) added induced currents equivalent to additional image dipole.

geometric optics-based method blocks the radiation from the feed and uses diffractions from the rim edge to calculate the pattern behind the reflector. We discuss UTD in Section 2-7. A comparison of UTD and physical optics calculations [1,4] of the pattern behind shows that the two methods match.

The dashed curve of Figure 2-9 plots the results of the PO analysis of a 20λ-diameter centrally fed paraboloidal reflector. The feed antenna radiation tapers to $-12\,\text{dB}$ at the
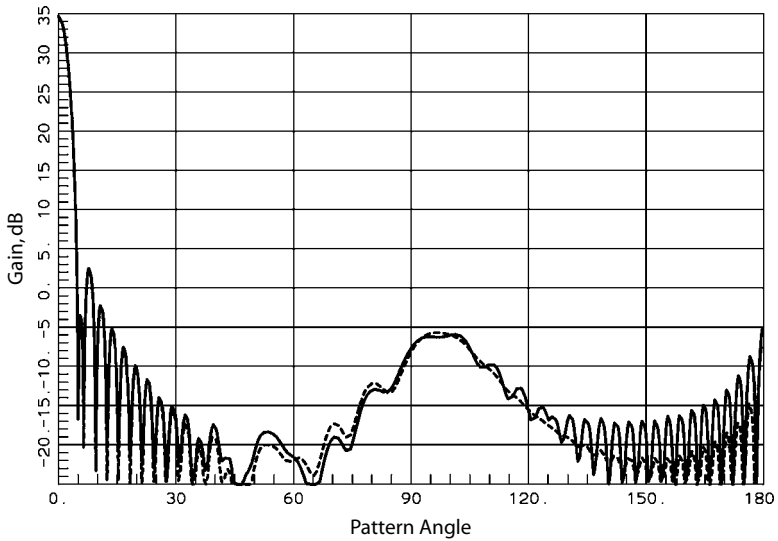
FIGURE 2-9 Physical optics analysis of a 20λ-aperture-diameter paraboloidal reflector (dashed curve) compared to analysis that includes PTD (solid curve).

reflector rim. Figure 2-9 shows the feed power spillover peaking at angles off the boresight near 100°. PO analysis computes the currents on a patch by assuming that it is embedded in an infinite plate. The reflector rim violates this assumption and we need extra terms to calculate the pattern behind the reflector accurately. Adding PTD (the physical theory of diffraction) to PO improves the match between the two methods behind the reflector as shown by the solid curve on Figure 2-9. PTD handles caustic regions of PO in a manner similar to the equivalent current method based on diffraction coefficients of UTD with geometric optics for shadow and reflection boundaries. For this example, the additional PTD currents add with the same phase because of the symmetry of the reflector geometry and produce the maximum effect. The PTD currents on the rim of an offset reflector will not add and produce a peak effect behind the reflector but will produce a more defuse effect. We only need PTD over a limited pattern angular range to reduce error, and the cost of implementing the fix may exceed the necessity of knowing the pattern in these regions. Similarly, UTD needs the addition of edge currents for accurate calculation of the radiation near 180°, behind the reflector. Although any model for the feed pattern can be used with PO, results matching UTD exactly occur only over all regions of the back radiation when the feed satisfies Maxwell's equations in the near and far fields [4, p. 212]. One such feed is the Gaussian beam approximation. Again, like PTD fixes, the small errors when using other feed antenna approximations occur only at limited pattern regions that may be unimportant.

2-4.3 Equivalent Currents

We can relate the concept of equivalent currents to physical optics. In this case we generate an artificial surface that covers a source of radiation. The incident fields generate surface electric and magnetic current whose radiation cancels the internal

fields and generates the external pattern. We use these at the apertures of antennas such as horns. By using the dyadic Green's functions we can calculate the near-field patterns and the coupling between antennas when the assumption is made that the presence of a second antenna does not alter the aperture fields. Given the outward normal $\hat{\mathbf{n}}$, we calculate the equivalent currents by

$$\hat{\mathbf{n}} \times \mathbf{E}_{\text{incident}} = -\mathbf{M}_S \qquad \hat{\mathbf{n}} \times \mathbf{H}_{\text{incident}} = \mathbf{J}_S \qquad (2\text{-}33)$$

We must use both electric and magnetic current densities on the surface to replace the internal fields. If the ratio of the electric field to the magnetic field equals the impedance of free space (376.7 Ω), the combination of the two currents produces the radiation of the Huygens aperture source when used with the dyadic Green's function. We use equivalent currents for a variety of analyses over flat apertures such as horns and paraboloidal reflectors, but they can also be used with curved structures or apertures.

We can, for example, use equivalent currents for calculation of the effects of radomes. Locally, we assume that the incident waves are plane waves and use boundary conditions to calculate reflected and transmitted waves. It is necessary to separate the incident wave into parallel and perpendicular polarizations, the ray-fix representation discussed in Section 2-7.8. These polarizations have differing reflection and transmission coefficients. We generate one surface on the inside of the radome and another on the outside. We use locally free-space waves for the reflected and transmitted waves lying outside the radome. Both these waves can be replaced with equivalent currents. The equivalent currents produce null fields inside the radome when combined with the incident wave radiation [4, p. 155]. Including these equivalent currents in a PO analysis, we add the effect of the radome.

Equivalent currents can also be used with lenses. We use the incident waves combined with the idea of locally plane waves to calculate reflected and transmitted waves at each surface and replace them with equivalent currents. We include the dielectric constant of the lens in the dyadic Green's functions for the internal radiation of the lens to calculate the fields at the second surface. We apply locally plane waves at the second surface to determine the transmitted and reflected rays and then replace them with equivalent currents. Because the lens has internal reflections, we need to apply an iterative PO analysis to calculate the multiple reflections between the two surfaces. The method converges rapidly because the internal reflections are small.

2-4.4   Reactance Theorem and Mutual Coupling

In Section 1-14 we discussed how the coupling between two antennas can be found from reactance. Given a transmitting antenna that generates a field at the receiving antenna, the reactance is described by an integral equation [5]:

$$\text{reactance} = \iiint (\mathbf{E}_t \cdot \mathbf{J}_r - \mathbf{H}_t \cdot \mathbf{M}_r) \, dV = \langle t, r \rangle \qquad (2\text{-}34)$$

The volume integral is over the receiving antenna currents, but it is often reduced to a surface or line integral. A second form of Eq. (2-34) uses the fields radiated by

both antennas. Given a surface that surrounds the receiving antenna, the integral for reactance is taken over this surface:

$$\text{reactance} = \iint\limits_{S_r} (\mathbf{E}_r \times \mathbf{H}_t - \mathbf{E}_t \times \mathbf{H}_r) \cdot d\mathbf{s} = \langle t, r \rangle \tag{2-35}$$

The differential normal $d\mathbf{s}$ is pointed away from the receiving antenna.

When we represent the two antennas and the transmission between them as an impedance matrix, it implies that we know the input currents to both antennas. By expressing the coupling as an impedance matrix, we compute mutual impedance from the reactance integral:

$$Z_{12} = \frac{-1}{I_1 I_2} \langle t, r \rangle \tag{2-36}$$

Antennas that we describe by input currents only have electric current densities excited on their surfaces. The mutual impedance formula using reactance reduces to

$$Z_{12} = \frac{-1}{I_1 I_2} \iint\limits_{V_r} \mathbf{E}_t \cdot \mathbf{J}_r \, dV \tag{2-37}$$

The volume integral reduces to a line integral in most cases.

Antennas with given input voltages such as slots can be described using magnetic currents and we use a mutual admittance matrix for the antenna pair:

$$Y_{12} = \frac{1}{V_t V_r} \cdot \text{reactance} = \frac{-1}{V_t V_r} \iint\limits_{V_r} \mathbf{H}_t \cdot \mathbf{M}_r \, dV \tag{2-38}$$

By using reciprocity antennas made of linear, isotropic materials, we have equal cross-matrix terms:

$$Z_{12} = Z_{21} \quad \text{and} \quad Y_{12} = Y_{21} \tag{2-39}$$

We calculate self-impedance terms by integrating over the surface of the antenna: for example, the radius of a dipole with the source of the field located at the center of wires or slots.

2-5   METHOD OF MOMENTS

The method of moments (MOM) [6] expands the currents on an antenna (or scattering object) in a linear sum of simple basis functions. The approximate solution is a finite series of these basis functions:

$$f_a = \sum_{i=1}^{N} a_i f_i \tag{2-40}$$

We compute the coefficients by solving integral equations to satisfy boundary conditions on the surface of the antenna (or object). The integral equation can be expressed in the form $L f_a = g$, where L is a linear operator, usually a scalar product using an integral, f_a the unknown currents given by Eq. (2-40), and g the known excitation

or source function. We substitute the summation of Eq. (2-40) into the linear operator equation and use the scalar product integral to calculate the terms in a matrix equation. The solution of the matrix equation determines the coefficients of current expansion. The MOM produces filled matrices that require time-consuming numerical methods for inversion. The art of the MOM is in choosing basis functions and deriving efficient expressions for evaluating the fields using the basis function currents. Common basis functions are simple staircase pulses, overlapping triangles, trigonometric functions, or polynomials.

The method does not satisfy boundary conditions at every point, only over an integral average on the boundaries. By increasing the number of basis functions, the method will converge to the correct solution. We need to judge how many terms are required for an adequate engineering evaluation. Spending excessive time on the solution cannot be justified if it greatly exceeds our ability to measure antenna performance accurately using real hardware.

2-5.1 Use of the Reactance Theorem for the Method of Moments

We can use the reactance theorem to generate a moment method solution to the currents on a thin-wire antenna. Thin-wire solutions assume that there are no circumferential currents and reduces the problem to filamentary currents. An electric field integral equation (EFIE) satisfies the boundary condition of Eq. (2-25a), a zero tangential field at the surface of the wires, but it does not seem explicit in the derivation. The reactance theorem produces an impedance matrix whose inversion yields the coefficients of the current expansion [7]. Similar to many other methods, the Green's function has been solved explicitly to reduce run time. This method [7] uses overlapping sinusoidal currents on V-dipoles as basis function currents and uses the Green's function to calculate the radiation from one V-dipole at the location of a second V-dipole. Both the radiating and receiving dipoles use the same expansion function. Galerkin's method uses the same weighting (or testing) function as the basis function and yields the most stable solutions. The reactance equation (2-37) calculates the mutual impedance between the two dipoles when each has unity current. We compute self-impedance by spacing a second V-dipole one radius away and by using the reactance theorem to calculate mutual impedance, a technique equivalent to the induced EMF method.

The scalar (dot) product between the incident vector electric field and the current density along the dipole reduces the vectors to scalars that can be integrated. The current density acts as the testing or weighting function for the method of moments. Performing the integration means that the current density only satisfies the zero tangential electric field boundary condition in an average sense. If series impedances are placed in the V-dipole, their impedance is added to the diagonal elements of the mutual impedance matrix. To excite the structure, we place a delta voltage source in series with the V-dipole terminals. The solution for the currents can be found by inverting the matrix equation and using the voltage excitation vector starting with the matrix equation

$$[Z_{mn}][I_m] = [V_n] \tag{2-41}$$

After computing the matrix inverse and specifying the input voltage vector, the complex current values are found on the structure:

$$[I_n] = [Z_{mn}]^{-1}[V_m] \tag{2-42}$$

Given the input voltage and the solution for the currents, the input impedance can be calculated. Similarly, the far- and near-field patterns can be calculated by using Eqs. (2-30) and (2-31) of the dyadic Green's function.

The code must satisfy Kirchhoff's current law at the junction between groups of V-dipoles, which adds a constraint to the currents. Because an overlapping sinusoidal basis function closely follows the actual currents normally excited on dipoles, the segments can be on the order of a quarter-wavelength long or more and yield acceptable results. Basis functions that closely follow expected current distributions are sometimes called *entire domain functions*. These reduce the size of the matrix to be inverted but require more complicated calculations for matrix terms and radiation. Although the concept of a V-dipole was expanded to a V rectangular plate [8], the method is only a subset of general integral equation solutions. This approach generates a simple impedance matrix formulation easily understood from an engineering point of view.

2-5.2 General Moments Method Approach

The method of moments can solve other types of electromagnetic problems: for example, electrostatic problems involving charges and dielectrics [9]. These solutions can determine the characteristic impedance of transmission lines useful in the design of antenna feeders. All moment method solutions are found from the solution of integral equations over boundary conditions. The boundary conditions can be either the tangential electric field (EFIE) or magnetic field (MFIE) conditions given by Eq. (2-25*a,b*) or a combination applied using an integral scalar product. We need a combination for closed bodies near an internal resonance frequency (resonant cavity) because the solutions exhibit resonances that make the solution invalid over a narrow frequency range. The method of moments can be applied to dielectric bodies when we use the constitutive relations of Eqs. (2-25) and (2-26), where the formulations for dielectric bodies use either volume or surface integrals [9].

Consider the use of the electric field integral equation (EFIE) with metal surfaces. We expand the currents on the objects using basis functions $\mathbf{B}_m(\mathbf{r}')$ with coefficients I_m:

$$\mathbf{J}(\mathbf{r}') = \sum I_m \mathbf{B}_m(\mathbf{r}') \tag{2-43}$$

The basis functions can be applied over a limited range of the structure in piecewise linear functions, which can be staircase pulses, overlapping triangular functions, or sinusoidal basis functions, whereas multiple functions can be applied over the whole or part of the structure for entire domain basis functions. For example, these could be a sum of sinusoidal functions which form a Fourier series representation.

On a PEC surface the tangential electric field vanishes [Eq. (2-28*a*)]. At field point $\mathbf{r}$ along the surface S,

$$\hat{\mathbf{n}} \times [\mathbf{E}_{\text{incident}}(\mathbf{r}) + \mathbf{E}_{\text{scattered}}(\mathbf{r})] = 0$$

$$\mathbf{E}_{\text{scattered}} = \sum_{m=1}^{M} I_m \iint\limits_{s'} \mathbf{B}_m(\mathbf{r}') \cdot \mathbf{G}(\mathbf{r}, \mathbf{r}') \, ds' \tag{2-44}$$

We can only satisfy Eq. (2-44) using a finite sum in the average sense of an integral. Since the integral and summation operate on a linear function, we can interchange them.

We introduce weighting (or testing) vector functions tangent to the surface $\mathbf{W}_n(\mathbf{r})$ and take the scalar (dot) product of this vector with the sum of electric fields. This limits the result to the tangential component of the electric field:

$$\iint\limits_{s} [\mathbf{E}_{\text{incident}}(\mathbf{r}) \cdot \mathbf{W}_n(\mathbf{r}) + \mathbf{E}_{\text{scattered}}(\mathbf{r}) \cdot \mathbf{W}_n(\mathbf{r})] \, ds = 0 \tag{2-45}$$

We identify the weighted integral of the incident field with the source and weighted integral of the field radiated by the basis functions (scattered field) as the impedance matrix terms. The integrals over the boundaries are one form of scalar product represented by $\langle \cdot \rangle$ notation. Using unity current on each basis function, we calculate the matrix terms by using the scalar product:

$$Z_{mn} = \left\langle \iint\limits_{s'} \mathbf{B}_m \cdot \mathbf{G}(\mathbf{r}, \mathbf{r}') \, ds', \mathbf{W}_n(\mathbf{r}) \right\rangle = \iint\limits_{s} \iint\limits_{s'} \mathbf{B}_m \cdot \mathbf{G}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{W}_n(\mathbf{r}) \, ds' \, ds \tag{2-46}$$

$$V_n = -\langle \mathbf{E}_{\text{incident}}(\mathbf{r}), \mathbf{W}_n(\mathbf{r}) \rangle = -\iint\limits_{s} \mathbf{E}_{\text{incident}}(\mathbf{r}) \cdot \mathbf{W}_n(\mathbf{r}) \, ds \tag{2-47}$$

The combination of Eqs. (2-46) and (2-47) when integrated over each portion of the source gives a matrix equation:

$$[Z_{mn}][I_m] = [V_n] \tag{2-48}$$

The weighting functions could be as simple as pulse functions, overlapping triangular functions on lines or surfaces (rooftop), piecewise sinusoidal functions, or others. The type of basis functions determines the convergence more than the weighting (testing) functions, which only determine the averaging. Realize that the moment method converges to the exact solution when we increase the number of basis functions, but it is a matter of engineering judgment to determine how many terms give acceptable answers.

Equation (2-47) defines the source voltage occurring over a segment when the formulation uses a piecewise function expansion. The incident voltage is the weighted integral of the incident electric field. For example, the NEC formulation applies an excitation voltage across one segment. The reaction integral formulation of Section 2-5.1 applies a voltage source at the end of a segment. The modeling of sources is an important part of the art in the method of moments.

The expansion of Eq. (2-44) is only one possible moment method solution. We could use the boundary condition on the magnetic field, a combination of the electric and magnetic field conditions on a PEC. If the surface has finite conductivity, the boundary conditions are modified. The moment method is a general method that computes approximate solutions to the currents. Unlike physical optics, the currents do not have to be assumed beforehand but are found as a finite series approximation.

Antenna designers discover that adequate codes are available for most problems. Moment method solutions are typically limited to objects only one or two wavelengths in size, although any method can be stretched. Analysis of large structures becomes intractable because of the large amount of computer memory required and the length

of time needed to calculate the solution. Coarse models may not give totally accurate results but can be useful in determining trends. Given these ideas, remember that physical models can be built that solve the electromagnetic problem instantaneously. We found that it takes considerable time to learn any code, and a new code has to offer considerable advantages or solve problems that the present one cannot solve before we invest our time.

2-5.3   Thin-Wire Moment Method Codes

Thin-wire codes that assume only filamentary currents are readily available. We have experience with NEC, the Richmond code (ASAP), and AWAS [10], a commercial code. All have advantages, but they take time to learn. A commercial code with a graphical interface makes the input and output easier: for example, for NEC. These pay for themselves quickly by saving time. NEC can include plates, but since it uses a MFIE (magnetic field integral equation) for them, it is limited to closed bodies. When accuracy becomes important, it is necessary to decrease the segment length and increase their number. These codes use matrix inversion with calculation time proportional to N^3 and a matrix fill time proportional to N^2. Run time increases enormously as the number of segments increases.

The commercial code AWAS determines the segmentation, while the user of NEC must specify it. The rule is to use at least 10 segments per wavelength, but initial analysis can tolerate the errors due to using fewer segments. The segments should be longer than the diameter, and care must be taken that the segments do not overlap because the radius of the wires is too large. Solid objects, such as plates, can be modeled as wire frames, with the rule that the perimeter of the wire equal the spacing between the wires [11]. This rule can be violated, but a test of the convergence should be made. When we model slots in a solid object, we cannot apply the perimeter equal-spacing rule because the slot will disappear. These codes compute the radiation pattern more accurately than the input impedance due to simplistic source models, and we may have to build the antenna to determine the true input impedance. Of course, an antenna with a good input impedance response that does not have the required pattern is useless.

We can reduce NEC run time if the antenna has symmetry with multiple inputs. The code reduces input by allowing the user to specify symmetry. For example, a multiarm spiral analysis requires only the input of one arm. The various mode voltages are entered after the basic structure impedance matrix has been solved. If an object has M-way symmetry, the matrix fill time is reduced by M^2 and the solution time by M^3. The various voltage modes can be applied afterward. If we add another wire segment after specifying symmetry, the symmetry is destroyed and the program uses the full matrix. The only advantage we gain is in specifying the model because the program solves the full matrix instead of the reduced matrix.

2-5.4   Surface and Volume Moment Method Codes

Antennas made of plates or containing finite plate ground planes can be solved by using wire meshing of a thin-wire code. The method of moments code has been extended to plates [12,13] using a rooftop basis function on both rectangular and triangular patches. The number of basis functions (i.e., matrix size) grows rapidly. One solution is to use entire domain basis functions. These require more complicated integrals, but

they reduce the matrix size. Dielectric portions of the problem lead to either volumetric integrals or various forms of surface integrals that use equivalent currents to replace the internal fields [9,14]. These problems lead to a variety of boundary conditions solved using a finite series of basis function and integral equations to satisfy those boundary conditions approximately.

MOM analysis of antennas mounted on dielectric substrates requires special techniques. Commercial codes determine the currents flowing on these antennas while accounting for the dielectric. Often, Green's functions are found numerically, which increases the execution time. Since the currents are located on the surface and the integrals of the boundary conditions are over the same surface, the singularity of the Green's function causes a numerical problem. For example, the free-space Green's function has the term $1/|\mathbf{r} - \mathbf{r}'|$, which becomes infinite on the surface. Spectral domain methods remove the singularity by using a sum of current sheets on the surface as an entire domain basis function. A uniform plane wave propagating at an angle to the surface excites the current sheet. The actual current flowing on the metal portions is expanded as a sum of these current sheets [15, p. 208ff; 16]. The uniform current sheets are expanded in a spatial Fourier transform as well as the Green's function, and the MOM problem is solved. The Fourier-transformed Green's function no longer has the singularity. When the metallization can be expressed as an infinite periodic structure, the current is expanded as a Fourier series. The infinite periodic structure is used with frequency-selective surfaces and infinite arrays. In this case the fields and currents are expanded in Floquet modes (harmonics).

2-5.5   Examples of Moment Method Models

Figure 2-10 demonstrates the use of a wire mesh to replace a solid plate. We located a resonant ($\approx \lambda/2$) dipole $\lambda/4$ distance over a λ-wide ground plane in the H-plane and offset $3/8\lambda$ from one edge. This is repeated in Figure 2-20 using GTD analysis. The rods only run parallel to the dipole because cross wires do not have currents induced on them in the ideal world of analysis. The circumference of the rods equals the spacing between the rods and forms an equivalent solid plate. An actual antenna could use smaller-diameter rods and work as effectively as the solid plate and would reduce weight and wind loading. NEC analysis produces the same pattern as the GTD analysis of Section 2-7.2, except that the E-plane size of the rods alters the backlobe predicted by GTD to some extent, because that analysis assumes infinite-length rods.
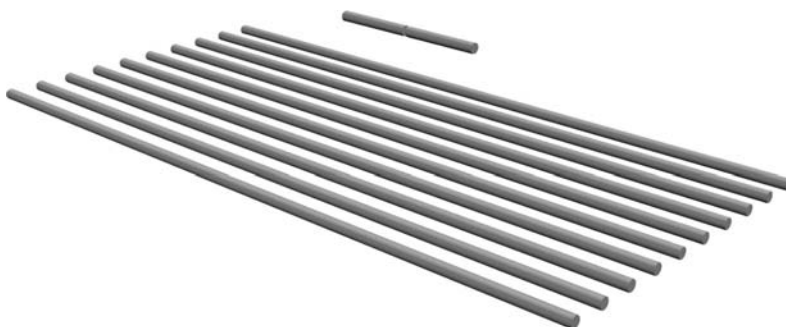

FIGURE 2-10   Use of a wire mesh to replace a solid plate for dipole over a ground plane in a MOM calculation.

Over most of the pattern angles the two analyses produce identical results. The NEC analysis accounts for the mutual impedance between the dipole and its image in the finite ground plane. For impedance calculations a small ground plane gives almost the same reaction to the antenna as an infinite ground plane.

Figure 2-11 shows a wire frame model of a cell phone. The model contains more wires than necessary for $\lambda/10$ spacing, but more wires improve the geometry match. When using crossed wires that shield both polarizations, we reduce the wire circumference in half since the wires approach the squares from four sides. The small wire antenna must be connected to the wire grid of the model to generate proper currents on the box. Either we restrict possible locations of the antenna or we must distort the wire grid locally. You should write an automatic grid generator if you use this analysis often. Consider that you need to specify whether an edge wire should be generated when two plates share the same edge. The hand holding the cell phone and the head nearby have significant effect on the antenna performance. The model given in Figure 2-11 has limited use. We need either a moment method analysis, such as WIPL-D, which includes volume dielectric structures, or FDTD, which can include complex material structures to model the head and produce good results.

Figure 2-12 illustrates a wire frame model of an airplane used for low-frequency analysis. Antennas mounted on free-flying models such as airplanes or spacecraft will excite the structure. Electrically, small antennas can excite the entire vehicle as an antenna. For example, a small antenna mounted on a large ground plane that would produce vertical polarization can excite the wings or fuselage and the entire system will
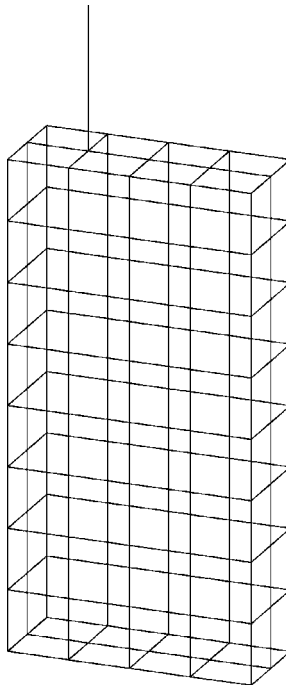

FIGURE 2-11   Wire frame MOM model of a cellular telephone handset with an antenna connected to the mesh.
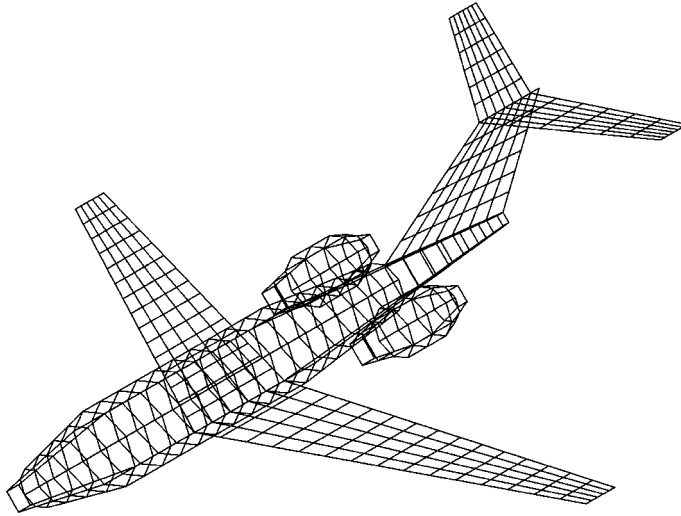
FIGURE 2-12   Wire frame MOM model of an airplane.

radiate horizontal polarization. Models similar to Figure 2-12 can eliminate surprises. The model restricts antenna mounting locations to the wire positions and may require local distortions of the grid.

Moment methods can include solid plates. Figure 2-13 shows an open waveguide horn analysis that uses a combination of plates and a single-feed wire monopole [12]. Locating the monopole or a small dipole inside the waveguide produces excitation of the waveguide mode that feeds the horn. Even though the model does not necessarily produce accurate impedance information, the model accurately calculates the pattern generated by the currents excited in the walls. We can either use an aperture method
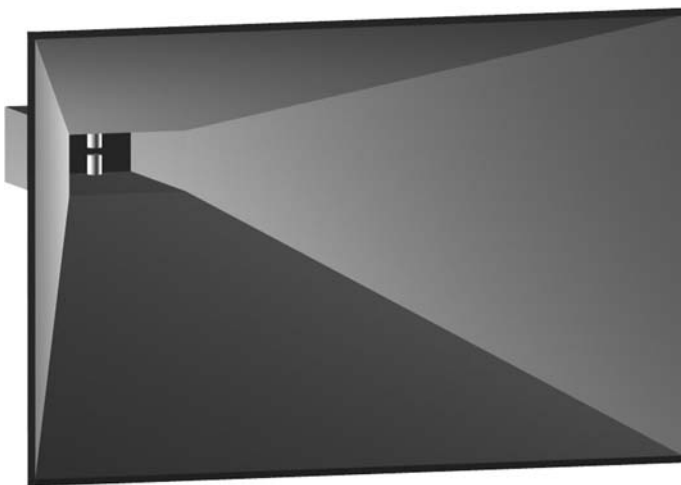
FIGURE 2-13   MOM model of a pyramidal horn using flat plates fed by a small dipole. (From [14, p. 229].)
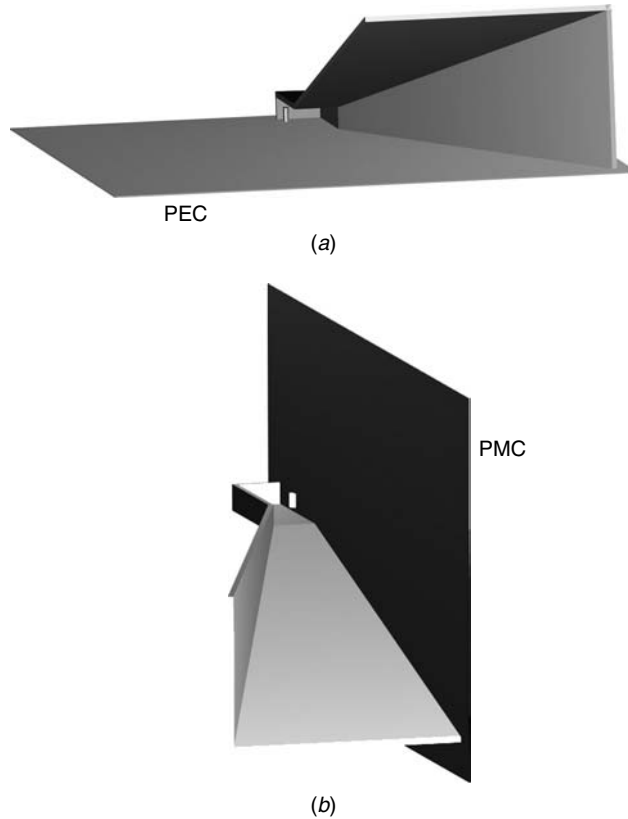
FIGURE 2-14 Use of electric and magnetic walls to reduce the model size in MOM analysis of a pyramidal horn: (*a*) PEC wall divides the horn; (*b*) PMC wall divides the horn.

for the horn that replaces the aperture fields or use the currents excited in the walls to calculate the pattern. Either method works for the front lobe. The moment method calculation requires significantly greater calculation time but produces results that better match measurements in all directions. Figure 2-14 demonstrates how to reduce calculation time by using planes of symmetry in a moment method analysis. In this case the small dipole feed is separated by two equally fed closely spaced dipoles. The right–left symmetry of the antenna allows reduction of the model by half. A vertical PMC wall divides the antenna into two parts, with only one remaining in the analysis. A horizontal PEC conductor divides the remaining model in half because halfway between the dipole feed is a virtual short circuit. Figure 2-14 contains only one-fourth the size of the original problem. Since matrix inversion requires N^3 calculations for an $N \times N$ matrix, dividing the analytical model down to one-fourth size reduces this calculation by a 64 : 1 factor. This also reduces the matrix element (fill time) calculations by 16 : 1. Reducing the model by using symmetry planes enables the solution of larger problems and reduces calculation time.

Analyses in later chapters use the moment method to predict antenna performance. Wire frame and plate analyses determine vehicle and mounting structure pattern effects. The moment method produces excellent analyses because it determines the approximate

current distribution as a sum of simple basis functions and we need not start with an assumed current distribution on the antenna.

2-6   FINITE-DIFFERENCE TIME-DOMAIN METHOD

The finite-difference time-domain (FDTD) method solves the coupled Maxwell's curl equations directly in the time domain by using finite time steps over small cells in space. The method reduces the differential equations to difference equations that can be solved by sets of simple equations. The method alternates between the electric and magnetic fields solved at locations a half-step apart because central differences are used to approximate derivatives. A 1966 paper by Yee [17] described the basic method that many authors have improved upon, but the original method remains the approach of choice.

FDTD can solve many types of electromagnetic problems, of which antenna analyses are only one type. Computer memory and speed limit the size of problems that can be solved, but larger and larger problems can be solved as the cost of computing keeps reducing. Besides antenna problems, the method is applied to microwave circuits, biological interaction with electromagnetic waves, optics, and radar cross-section problems. The number of uses expands every day. The method allows each cell to be made of different materials, leading to the solution of volumetric complex structures. The solution of the equations is robust and the errors are well understood.

Currently, the method solves moderately small antenna problems on the order of a few wavelengths. Of course, faster and larger computers can solve larger problems, especially if the analyst has patience. FDTD handles microstrip antennas with their complex layering of dielectrics, including a finite ground plane without the use of complex Green's functions required of frequency-domain solutions. The interaction of antennas with the near environment, such as the effect of the head on cellular telephone handsets, can be solved. In this case the complex electromagnetic properties of the head can be described as cells each with different electrical properties. In addition to giving a solution to the radiation pattern and allowing characterization of the communication system, it can provide insight into the radiation safety concerns of users. The method handles the solution of the interaction of antennas with the human body in a straightforward manner for prediction of biomedical applications, such as electromagnetic heating for cancer treatment.

Learning to apply the technique, whether formulating your own routines or using a commercial code, will yield insight for design. The method can produce time-domain animated displays of the fields that show radiation centers and where the fields propagate, but the user must learn to interpret these new displays. It will be worth your effort to learn this task. The time-domain responses using impulse signals can produce solutions over a wide band of frequencies when converted to the frequency domain using the discrete Fourier transform (DFT). The only drawback is the computer run time required.

2-6.1   Implementation

By using a direct implementation of Maxwell's curl equations in the time domain, you do little analytical processing of the equations. No vector potential or Green's

functions are developed as in frequency-domain methods. Although the antenna may be volumetrically complex and contain many different materials, the method yields sparse matrices rather than the dense matrices produced by moment methods. It is a direct solution that does not require the inversion of large matrices and includes only nearest-neighbor interactions. Having only nearest-neighbor interactions means that it is possible to run problems on parallel machines.

You need to embed the antenna in a rectangular region and divide it into rectangular cubical cells with sizes ranging from 10 to 20 samples per wavelength at the highest frequency where analysis is desired. The outer surfaces contain absorbing boundaries to eliminate reflections that would produce errors. Formulating absorbing boundary conditions has been a significant part of the method. You need to locate a solution surface between the absorbing boundaries and the antenna outer surface where we compute currents by using the equivalence theorem. The DFT of the time response determines the radiation pattern at a given frequency after the equivalent currents are found. If you need the pattern amplitude in only a few directions, the time-domain radiation can be found directly: for example, the gain in one direction.

We can formulate some problems in one or two dimensions if they possess symmetry instead of the three-dimensional rectangular cube. The solution time is reduced dramatically, and the time-animated presentation may provide sufficient insight when the radiation pattern is found in two dimensions. Because this is a time-domain analysis, we need to excite the structure with a pulse. You use the pulse frequency power response to normalize the patterns and compute gain. When the formulation includes the material losses, the efficiency of the antenna can be found since the dissipation in the inner cells prevents the radiation from reaching the outer surface.

2-6.2 Central Difference Derivative

Numerical derivatives have greater potential for errors than integrals, but FDTD uses them to reduce Maxwell's differential curl equations to simple difference equations. A second-order accurate formula for a derivative can be found by using central differences instead of using the difference between the value at a location or time and the value at the next point in a sequence of evenly spaced points:

$$\frac{\partial f}{\partial u} = \frac{f(u_0 + \Delta u/2) - f(u_0 - \Delta u/2)}{\Delta u} + O(\Delta u)^2 \tag{2-49}$$

We can use finite differences to solve the curl equations provided that we use electric and magnetic fields spaced at half intervals because each is related to the derivatives of the other field and we want to use central differences to reduce error. Because Maxwell's equations involve time derivatives, we need to calculate the electric and magnetic fields at interspersed half time intervals.

2-6.3 Finite-Difference Maxwell's Equations

Consider Maxwell's curl equations in the time domain, including lossy materials:

$$\frac{\partial \mathbf{H}}{\partial t} = -\frac{1}{\mu}(\nabla \times \mathbf{E} - \mathbf{M} + \sigma^* \mathbf{H})$$

$$\frac{\partial \mathbf{E}}{\partial t} = -\frac{1}{\varepsilon}(\nabla \times \mathbf{H} - \mathbf{J} + \sigma \mathbf{E}) \tag{2-50}$$

Equations (2-50) contain the source currents $\mathbf{J}$ and $\mathbf{M}$ and include losses due to conducting dielectric material σ and magnetic material losses σ^*. Both equations have the same form, with only an interchange of symbols. Expanding the curl operator, we get the following equation for the x-component of the magnetic field:

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu}\left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} - M_x - \sigma^* H_x\right) \tag{2-51}$$

The x-component of the electric field has the same form but with the interchanges $H \rightarrow E$, $E \rightarrow H$, $M \rightarrow J$, and $\sigma^* \rightarrow \sigma$. You obtain the equations for the y- and z-components by a cyclic variation (repeating pattern of interchanges) $x \rightarrow y \rightarrow z \rightarrow x \rightarrow y$, and so on. For example, the equations are reduced to two dimensions by leaving out the y-component.

FDTD calculates the field at discrete times and locations on a grid. The fields can be represented as an indexed function using integers:

$$f(i\ \Delta x, j\ \Delta y, k\ \Delta z, n\ \Delta t) = f(i, j, k, n)$$

Because we use central differences [Eq. (2-49)], for derivatives, and the magnetic (electric) field is found from the space derivative of the electric (magnetic) field, the magnetic and electric fields need to be spaced a half-space interval apart. The time derivative becomes

$$\frac{\partial f(i, j, k, n)}{\partial t} = \frac{f(i, j, k, n + \frac{1}{2}) - f(i, j, k, n - \frac{1}{2})}{\Delta t}$$

and means that the electric and magnetic components are interspersed at $\Delta t/2$ times that which produces a leapfrog algorithm. We substitute these ideas into Eq. (2-51) to derive the time-stepping equation for one component:

$$H_x(i - \tfrac{1}{2}, j, k, n + 1) = \frac{1 - \sigma^*(i - \tfrac{1}{2}, j, k)\ \Delta t/2\mu(i - \tfrac{1}{2}, j, k)}{1 + \sigma^*(i - \tfrac{1}{2}, j, k)\ \Delta t/2\mu(i - \tfrac{1}{2}, j, k)} H_x(i - \tfrac{1}{2}, j, k, n)$$

$$+ \frac{\Delta t/\mu(i - \tfrac{1}{2}, j, k)}{1 + \sigma^*(i - \tfrac{1}{2}, j, k)\ \Delta t/2\mu(i - \tfrac{1}{2}, j, k)}$$

$$\left[\frac{E_y(i - \tfrac{1}{2}, j, k + \tfrac{1}{2}, n + \tfrac{1}{2}) - E_y(i - \tfrac{1}{2}, j, k - \tfrac{1}{2}, n + \tfrac{1}{2})}{\Delta z}\right.$$

$$- \frac{E_z(i - \tfrac{1}{2}, j + \tfrac{1}{2}, k, n + \tfrac{1}{2}) - E_z(i - \tfrac{1}{2}, j - \tfrac{1}{2}, k, n + \tfrac{1}{2})}{\Delta y}$$

$$\left. - M_x(i - \tfrac{1}{2}, j, k, n + \tfrac{1}{2})\right] \tag{2-52}$$

FDTD uses similar equations for the other components [18,19].

Yee's Cell   Figure 2-15 shows one cubic cell and the components of the fields. When we consider the upper face, we see that the magnetic field components are spaced a
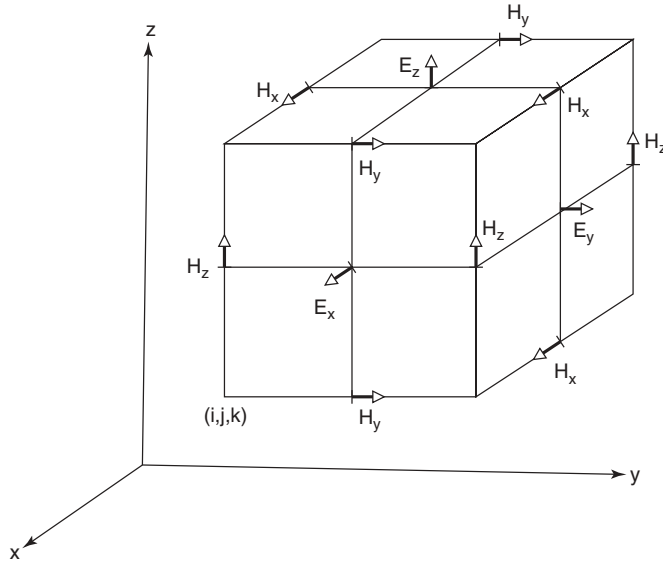
FIGURE 2-15 Unit cell of a Yee space lattice showing time and space separation of electric and magnetic fields in a cell. (From [15], Fig. 1, © 1966 IEEE.)

half space interval from the central electrical field and the arrows show the direction of fields. Although it would appear that the electric field is different on the upper and lower face along the z-axis, the method assumes that the field is constant throughout the cell. The magnetic fields shown are at the center of adjoining cells.

A leapfrog solution uses stored values of the electric fields to calculate the magnetic fields at a half time interval later and stores these values. In the second step the solution takes another half time step and uses the stored values of the magnetic fields to calculate the electric fields. The method gains stability by using the half time steps and by solving for both electric and magnetic fields. Although the fields are a half time step out of synch, we can average between the two half time steps to produce simultaneous fields at a point, but we only need to do this when calculating equivalent currents on the surface used for far-field pattern calculations.

2-6.4 Time Step for Stability

You need to pick the time step to produce a stable solution. Consider a plane wave traveling through the cubes. If the time step is too large, the wave can pass through more than one cell for each time step. At that point the solution cannot follow the actual wave propagation and fails. We must reduce the time step until it is less than the Courant condition or the wave propagation rate. Consider the fastest-moving wave in the problem, usually free space, and for equal sides to the cube, we compute the time step from the velocity and cell length:

$$v \, \Delta t \leq \frac{\Delta x}{\sqrt{d}} \tag{2-53}$$

The cell length is Δx and the number of dimensions is d. The time step must be lower for conducting materials ($\sigma > 0$) to produce a stable solution. The magic step uses

the equality and produces the most stable solutions. If you pick unequal sides to the rectangular cell, Eq. (2-53) is modified.

2-6.5 Numerical Dispersion and Stability

FDTD analyses produce solutions that fail to propagate through the cells at the proper phase velocity in all directions. The propagation velocity depends on the cell size in wavelengths; it has a frequency-dependent component. You need to consider this numerical dispersion because it affects accuracy. Because the waves travel at different velocities in different directions, the dispersion problem increases for large structures where many time steps must be taken. After many steps, signals disperse because they have taken different routes and fail to add together in the correct phase. Finer cells solve the problem, but the computation requirements grow rapidly. The equation for the propagation constant can be found from considering the FDTD formulation to produce the following equation for three-dimensional problems:

$$\left(\frac{1}{c\,\Delta t}\sin\frac{\omega\,\Delta t}{2}\right)^2 = \left(\frac{1}{\Delta x}\sin\frac{k_x'\,\Delta x}{2}\right)^2 + \left(\frac{1}{\Delta y}\sin\frac{k_y'\,\Delta y}{2}\right)^2 + \left(\frac{1}{\Delta z}\sin\frac{k_z'\,\Delta z}{2}\right)^2$$
(2-54)

The factor k_x' is the FDTD propagation constant in the cells along the x-axis, only approximately the same as k_x, the actual propagation constant in the structure. The y- and z-axes have similar problems. If you take the limit as cell length approaches zero, $u \to 0$, and so on, then $\sin(au)/u \to a$. Because $\Delta t \to 0$ as the cell size shrinks for the solution still to satisfy the Courant limit, Eq. (2-54) reduces to the expression

$$\left(\frac{\omega}{c}\right)^2 = k_x'^2 + k_y'^2 + k_z'^2$$
(2-55)

Equation (2-55) is the normal propagation constant equation for a plane wave in space and shows that the cell propagation constants converge to the correct values as the cell size shrinks. If you formulate a problem in one or two dimensions, you remove terms from the right side of Eq. (2-54) to determine the dispersion relationship.

Absorbing boundary conditions (ABCs) can cause numerical instabilities. ABCs approximate infinite space to simulate radiation by the antenna into space. FDTD problems must be placed in a finite number of cells because each cell requires computer storage. Every FDTD problem uses a finite number of cells for the ABCs with more cells required in the directions of maximum radiation. ABCs degrade as the number of time steps increases and eventually leads to numerical instabilities. A lively research on ABCs has produced good ones, but be aware that most have been found to produce problems at some point. If you write your own analyses, you will need to find appropriate ones. Commercial codes will give their limitations.

At one time, ABCs limited solution dynamic range, but ABCs are now available that produce reflection coefficients from 10^{-4} to 10^{-6}. Numerical dispersion limits the dynamic range as well. Remember that the antenna will be modeled with small cubes that limit the resolution of the results. The errors of modeling lead to solution errors that limit the dynamic range.

2-6.6 Computer Storage and Execution Times

The antenna to be analyzed is modeled by a set of cubic cells. Choosing an appropriate number is an art. Similarly, it will be necessary to have a meshing program. Using a

two-dimensional model will greatly reduce computer storage and run time. Remember that our purpose should be to gain insight unavailable from measurements. The calculations require the storage of three components of both the electric and magnetic fields in each Yee cell. Because we solve the problem in the time domain, the components are only real numbers, unlike frequency responses, which use complex numbers for each component. The material properties of the cells can be indicated with short 1-byte integers provided that there are no more than 256 different ones. Single-precision storage of the components requires 30 bytes for each cell; double-precision storage requires 54 bytes. A three-dimensional problem with 200 cells on a side contains 8 M cells and would need 240 Mbytes of storage for single-precision and 432 Mbytes for double-precision components.

At each time step approximately 10 floating-point operations (flops) are needed for each component in each cell. We must run the time steps until the input pulse has peaked and died out in each cell. This takes about 10 times the number of cells in the longest direction (maximum number along one axis). The three-dimensional problem with 200 cells on a side runs for 2000 time steps and requires 60 flops times the number of cells. The solution needs $2000 \times 8\,\text{M} \times 60$ flops $= 960$ Gflops for completion.

2-6.7 Excitation

We specify the excitation of an antenna in the time domain since FDTD operates in the time domain. If all we need is a single-frequency solution, a ramped sinusoidal waveform can be applied. The waveform is tapered from zero in about three cycles and the FDTD solution steps continue until a steady state is reached. It is more efficient to use a waveform that gives a wide-frequency-range response after performing a discrete Fourier transform on the radiating boundary to compute equivalent currents used at a given frequency. The computer storage and run times are the same for the wideband response as the single-frequency response.

A suitable wide-bandwidth excitation is the differentiated Gaussian pulse shown in Figure 2-16:

$$V_{\text{inc}}(t) = -V_0 \frac{t}{\tau_p} \exp\left[-\frac{(t/\tau_p)^2 - 1}{2}\right] \tag{2-56}$$

We calculate the frequency response of the differentiated Gaussian pulse from the Fourier transform of Eq. (2-56):

$$V_{\text{inc}}(\omega) = -j\omega\sqrt{2\pi}\ \tau_p^2 V_0 \exp\left[-\frac{(\omega\tau_p)^2 - 1}{2}\right] \tag{2-57}$$

The spectrum of Eq. (2-57) peaks for $\omega_p = 1/\tau_p$. Figure 2-17 gives the normalized frequency response and shows that the -20-dB-level normalized frequency extends from 0.06 to 2.75. For example, if we wanted to center the frequency response at 10 GHz, the normalizing pulse time is easily found:

$$\tau_p = \frac{1}{2\pi(10 \times 10^9)} = 1.592 \times 10^{-11}\,\text{s} = 15.92\ \text{ps}$$

A check of Figure 2-17 shows that the antenna frequency response could be found from 2 to 22 GHz with only a 10-dB loss in dynamic range compared to the response at 10 GHz. A single time response computation yields a wide-frequency-range response.
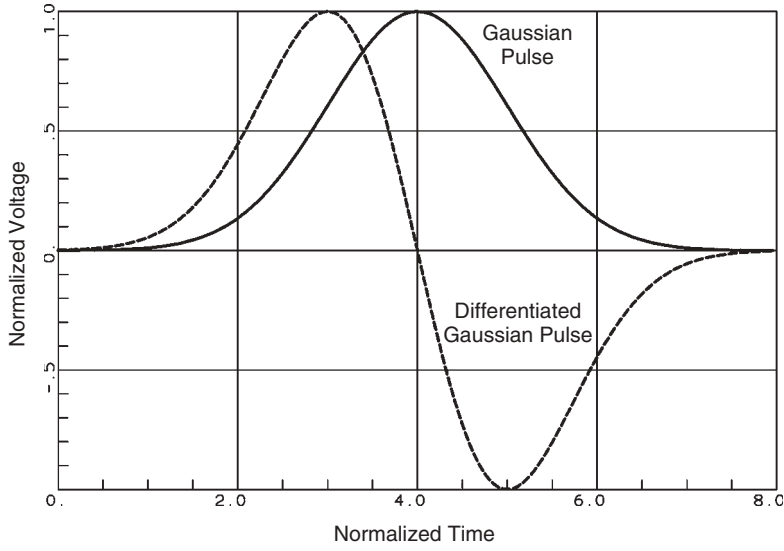
FIGURE 2-16   Differentiated Gaussian pulse time response used in FDTD analysis.
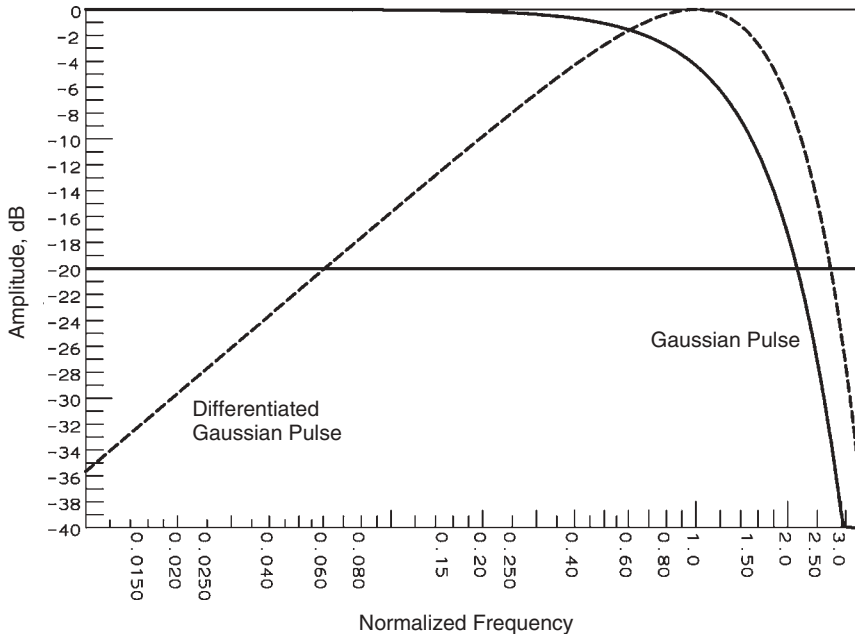

FIGURE 2-17   Differentiated Gaussian pulse normalized frequency response.

A sinusoidal modulated Gaussian pulse produces a narrow-bandwidth excitation useful in visualization because the bandwidth of the pulse can be controlled:

$$V_{\text{inc}}(t) = V_0 \exp\left[-\frac{(t/\tau_p)^2}{2}\right] \sin \omega_0 t \qquad (2\text{-}58)$$

The unmodulated Gaussian pulse shown in Figure 2-16 has a low-pass frequency response:

$$V_{\text{inc}}(\omega) = \sqrt{2\pi}\,\tau_p\,V_0 \exp\left[-\frac{(\omega\tau_p)^2}{2}\right] \tag{2-59}$$

Figure 2-17 gives the low-pass frequency response of the Gaussian pulse with a -4.37-dB response at $\omega_p = 1/\tau_p$. The sinusoidal modulation centers the frequency response of the Gaussian pulse around ω_0, and the convolution of the two frequency responses produces a two-sided response of the Gaussian pulse.

2-6.8   Waveguide Horn Example [19]

The literature contains solutions for the patterns of a number of antennas. Figure 2-18 shows the meshing of a commercial standard gain horn analyzed and compared to measurement. The horn operates from 8.2 to 12.4 GHz. The horn has a radiating aperture that is 110 mm wide and 79 mm high and a bell length of 228 mm. The 51-mm length of the input waveguide and the details of the feed probe were included in the model.

Placing a perfectly magnetic conductor through the midsection of the horn uses symmetry to halve the number of cells to a uniform mesh of $519 \times 116 \times 183$ Yee cells. Ten cells were used on the outside for the ABCs around the sides of the horn and 40 cells for the front ABCs in the maximum radiation direction. The model placed 20 cells between the edge of the horn and the equivalent current surface used for pattern calculations. The longest side of the grid determined the number of time steps at 10 times the number of cells $= 5190$ time steps. The model contains approximately 11 M Yee cells that require 330 Mbytes of computer storage. Assuming that the problem
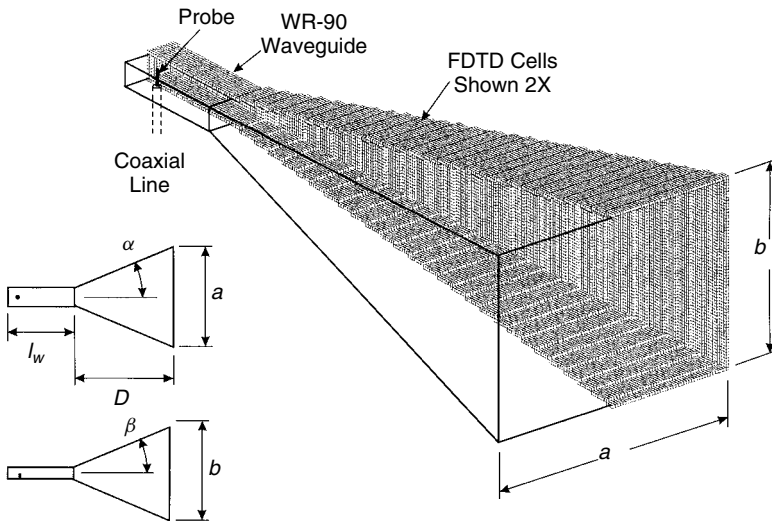

FIGURE 2-18   FDTD model of a standard gain horn. (From [17], Fig. 7.17, © 1998 Artech House, Inc.)
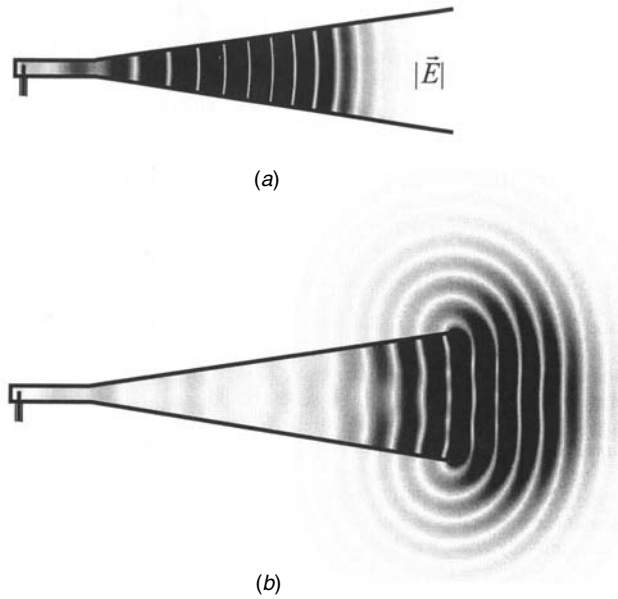
(a)

(b)

FIGURE 2-19   FDTD calculated electric field in the vertical symmetrical plane of a standard gain horn: (a) early time with a pulse in the throat; (b) pulse leaving the mouth of the horn. (From [17], Fig. 7.20, © 1998 Artech House, Inc.)

takes 60 flops per cell for each time step, the solution required 3.43 Tflops of computer calculations.

The initial calculation used a differentiated Gaussian pulse excitation with $\tau_p = 15.9$ ps that centered the response at 10 GHz. The calculation produced patterns that matched measurements. A second calculation used a sinusoidal modulated Gaussian pulse with the time constant 79.6 ps. This pulse time constant gives a normalized frequency of 2 GHz for the Gaussian pulse. The -3-dB frequency is 0.83 times the normalizing frequency. The pulse is centered at 10 GHz with a 3-dB bandwidth of 3.32 GHz. Figure 2-19 shows the fields when the pulse reached the horn aperture. Note the high fields in front of the horn and the amount of fields still radiating beyond and behind the aperture. By using a sinusoidal modulated pulse, the visual display contains nulls that improve its clarity.

2-7   RAY OPTICS AND THE GEOMETRIC THEORY OF DIFFRACTION

Ray optics can give you a good physical feel for radiation and spur design ideas, but we need to question the accuracy of their use. Ray optics or geometric optics (GO) methods come from the design of lens and optical reflectors where the wavelength is very short compared to the size of the object being analyzed, whereas we may be interested in analyzing or designing an antenna on a structure only a few wavelengths in size. Below we show that GO is essentially correct over most of the radiation sphere and that by using elements of the geometric theory of diffraction [GTD (UTD)], the pattern prediction can be improved. In this case improvement means that we will increase the area of the radiation pattern that becomes more accurate. You will discover that it takes

an increasing amount of effort to improve small areas of the pattern prediction, and at some point you should decide that it fails to give enough improvement to justify the work. Your real design purpose is to determine antenna dimensions that produce the desired antenna response. Of course, as the expense of the antenna increases, your customer may demand better predictions of the final result, and then the cost of a better analysis is justified. You need to accept a new approach. Even though a part of the pattern prediction shows errors, obvious discontinuities, it only means that the pattern is inaccurate in directions near them and that over most of the radiation sphere the prediction is essentially correct.

Discussion of this method begins with simple examples given in two-dimensional space that introduce the ideas behind GO and GTD. These examples can ignore the details of rotation of polarization directions because the waves are either polarized with the electric field normal to the page or located in the plane of the page. We consider radiation blockage by objects, the reflection of rays by the objects, and the diffraction of rays around edges that fills in the pattern in the shadow regions and across the boundary of the last reflected ray.

After the discussion of simple examples, the key points of GTD will be given for use in three-dimensional problems. This involves the rotation of coordinate systems so that ray polarizations line up with planes of incidence for reflections, with edges for diffraction and curvature directions on curved surfaces that shed rays around the object into the shadow. You will need to investigate the references if you want to develop your own routines, but this discussion will introduce you to the topic and give you an appreciation of the method so that you can use available computer programs and understand their limitations.

GO uses ray methods to approximate electromagnetics. It is exact only in the limit of zero wavelength (infinite frequency), but we gain useful insight from it at any frequency. It will not give good results close to physical boundaries; but when we include the GTD, the results are accurate down to one-wavelength sizes and are useful at $\lambda/4$ sizes. GO gives us physical insight when we deal with reflectors. We must consider three aspects to use GO fully: (1) ray reflections, (2) polarization, and (3) amplitude variations along the ray path and through reflections.

2-7.1 Fermat's Principle

Rays travel through a medium at the speed of light determined by the index of refraction: $n = \sqrt{\varepsilon_r \mu_r}$. We define the optical path length as $\int_C n \, dl$, where C is a prescribed path in space. Fermat's principle determines the paths of rays between two points. It states that the optical path length is stationary along a valid ray path. An expression is stationary when its first derivatives are zero and the optical path is a minimum (or maximum). We use Fermat's principle to trace ray paths through reflection or refraction by searching for the minimum optical path lengths. We can find more than one possible ray between points because Fermat's principle requires only a local minimum. When we exclude the boundaries of lenses, regions of homogeneous medium, rays travel in straight lines.

2-7.2 *H*-Plane Pattern of a Dipole Located Over a Finite Strip

Figure 2-20 illustrates the geometry of this problem and the various regions of the analysis. The diagram shows the end of the dipole rod with the two rods located
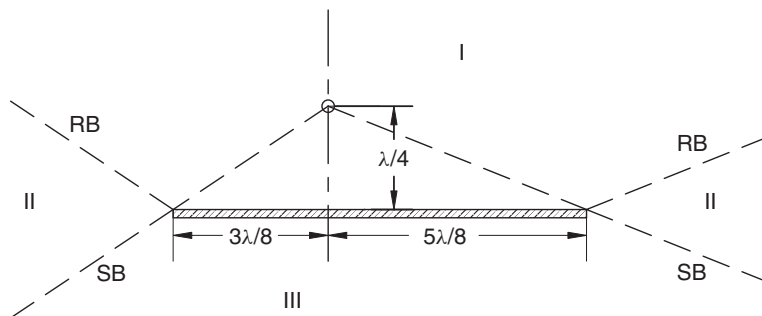
FIGURE 2-20   GTD example using a two-dimensional model of a dipole located over an asymmetrical ground plane.

normal to the page. The dipole pattern is omnidirectional in the page with the electric field directed normal to the page. When we trace rays from the dipole to the finite strip, we discover two significant directions on both sides of the strip. The dashed boundaries labeled RB (reflection boundary) are the directions of the last rays reflected from the strip. Similarly, the dashed boundaries labeled SB (shadow boundary) are the last rays of radiation not blocked by the strip. The radiation in region I results from the sum of the direct radiation from the dipole plus the radiation reflected by the strip. Only direction radiation from the dipole occurs in the two parts of region II. Finally, region III is totally blocked from any radiation by a direct or reflected ray. This region receives rays diffracted around the edges.

If we add the direct and reflected rays in an analysis, we obtain the pattern given in Figure 2-21, which also traces the actual pattern. The pattern, using only the direct and reflected rays, accounts for the phasing between the direct radiation from the dipole and an image dipole located below the strip. If you compare the two traces on Figure 2-21, you see that the two patterns are similar near $\theta = 0$, but the direct plus reflected ray pattern has discontinuities at the SBs and RBs. Figure 2-22 gives the results for the same analysis, but using a 5λ-wide ground-plane strip. When using the larger strip, the two patterns match to about $80°$, and in the second case the simple analysis is correct over most of the forward semicircle. Simple geometric optics gives good results for large objects provided that you realize the patterns will contain discontinuities.

Removing the discontinuities requires extra effort. A discontinuity in the pattern cannot exist because shadow and reflection boundaries occur in free space. It takes a material boundary to produce a discontinuous field. But, for example, the tangential electric field must be continuous across even material boundaries. Edge diffraction solves the discontinuity problem. Figure 2-23 gives the pattern of the edge diffraction for both edges normalized to the total pattern. The edge diffraction has matching discontinuities to the sum of the direct and reflected rays at the SBs and RBs. The UTD (uniform theory of diffraction) technique [20, p. 55] calculated these diffractions. When these diffractions are added to the direct and reflected ray radiation, the total pattern given in Figure 2-21 is obtained. The dipole, its image in the ground plane, and the two edge diffractions form a four-element array where each element has a unique pattern. Adding edge diffractions to the geometric optics fields removes the discontinuities and allows calculation of the pattern behind the strip ground plane.
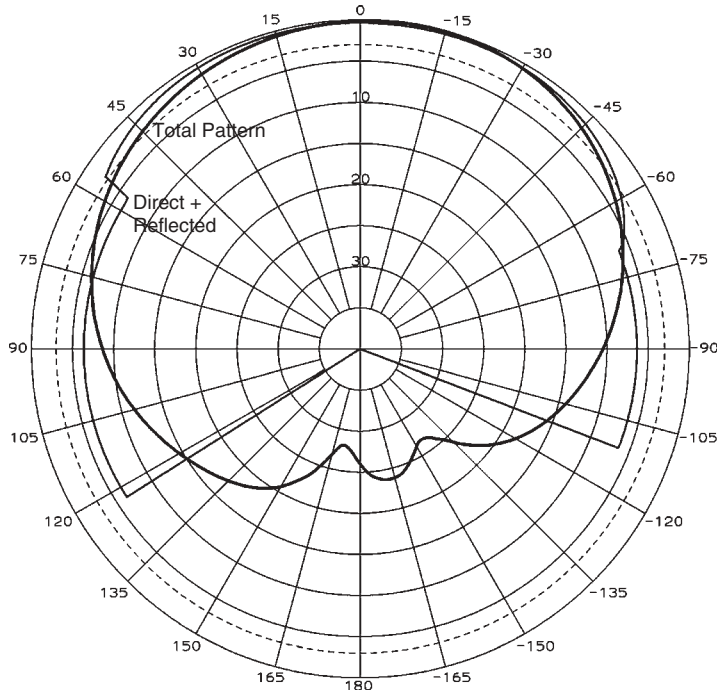
FIGURE 2-21 H-plane pattern of a dipole over asymmetrical ground using direct and reflected rays compared only to a full solution for the 1λ ground plane of Figure 2-20.
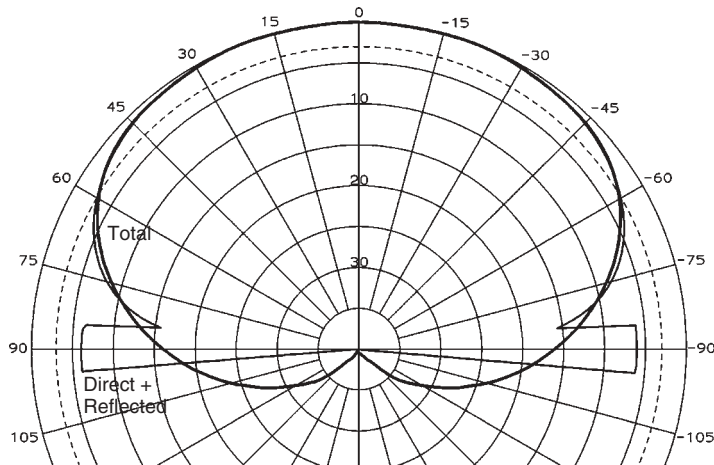

FIGURE 2-22 H-plane pattern of a dipole over symmetrical ground using direct and reflected rays compared only to a full solution for a 5λ ground plane.

2-7.3 E-Plane Pattern of a Rectangular Horn

Figure 2-24 illustrates the cross section of a horn or, in this case, a two-dimensional approximation to a horn. The waveguide feeds the horn and produces a uniform aperture distribution in the E-plane. In this model the direct GO radiation is a constant wedge signal as shown in Figure 2-25 ranging between $-15°$ and $+15°$. The reflected pattern
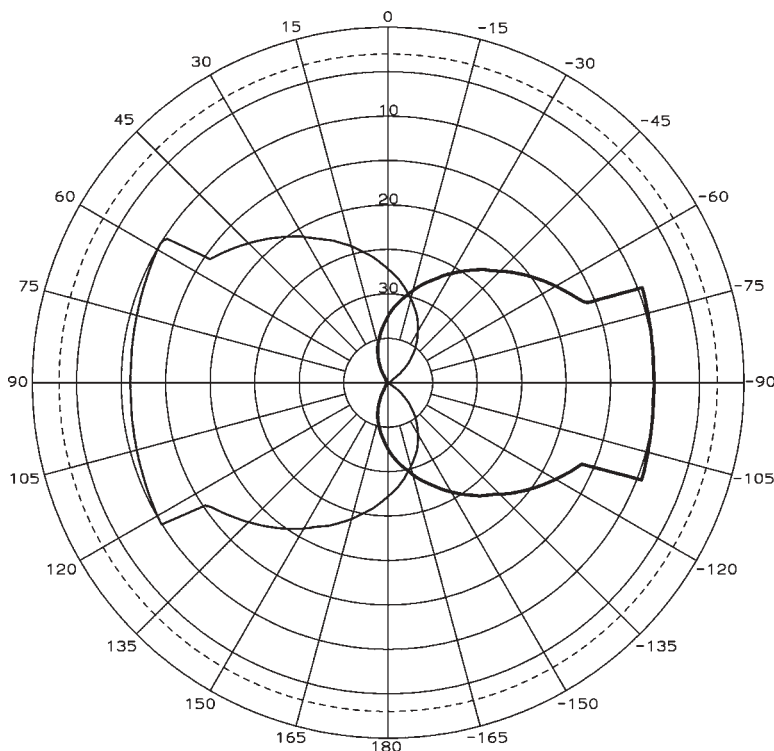
FIGURE 2-23   GTD edge diffraction of an *H*-plane pattern for an asymmetrical 1λ ground plane under a dipole.

combines with the direct radiation and produces the same pattern. Figure 2-25 also shows the diffraction patterns from the two edges. These peak along the plates and exhibit a discontinuity at the same angle as the GO field. Each diffraction pattern has a discontinuity on one side at 90° because the mouth of the horn blocks the diffraction from the opposite edge. When we add the diffracted fields to the GO field, the pattern shown in Figure 2-26 is obtained. By just adding the three components, we obtain an accurate pattern of the horn over most of the angles of the plot. At 90° we see discontinuities in the pattern caused by not considering enough terms in the GTD calculation. You need to realize that these discontinuities only cause pattern errors at nearby angles. The majority of the pattern is correct.

   We need another term to correct the pattern near 90°. The blockage of the diffraction from one edge by the mouth of the horn causes a secondary diffraction at that edge. We call this *double diffraction*. Some available programs do not implement double diffraction because the general three-dimensional double diffraction takes considerable calculation due to the extensive ray tracing required. In these cases you must accept the pattern discontinuities. Some programs calculate double diffraction as an option, but turning on this option will slow the calculations. Figure 2-27 gives the pattern when double diffraction is included. Double diffraction reduces the discontinuity at 90°, but a small discontinuity remains. Adding triple diffraction would reduce this further, but the pattern area affected by the small discontinuity has shrunk. A new discontinuity near 60° appeared in the pattern after adding double diffraction at the
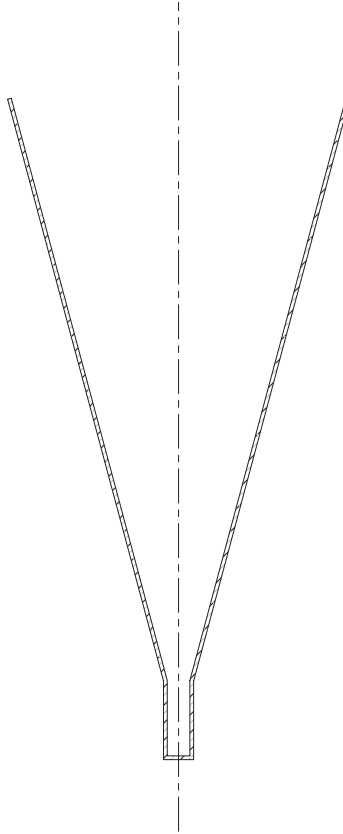
FIGURE 2-24   Geometry of a two-dimensional model of a rectangular horn used for GTD analysis.

mouth of the horn. We could continue to add another term to remove this one or just accept it.

2-7.4   *H*-Plane Pattern of a Rectangular Horn

The tangential electric fields vanish at the walls of the two-dimensional horn in the *H*-plane. This affects the GO field and produces the following equation for them:

$$E^{\text{GO}} = \cos \frac{\pi \tan \theta}{2 \tan \alpha} \frac{e^{-jkR}}{\sqrt{R}} \tag{2-60}$$

Equation (2-60) includes the phasing term and square-root spreading factor of a two-dimensional field. The horn walls tilt from the centerline by the angle α. Figure 2-28 plots the GO field and shows that it vanishes at the walls. We do not expect edge diffraction because the field vanishes at the edges, but Figure 2-28 shows diffraction patterns that peak in the direction of the walls.

We call this new term *slope diffraction*. This new type requires another set of coefficients not identical to the edge (or wedge) diffraction coefficients. While the amplitude of the edge diffraction is proportion to the field incident on the edge, the amplitude of slope diffraction is proportional to the derivative of the field in the direction normal
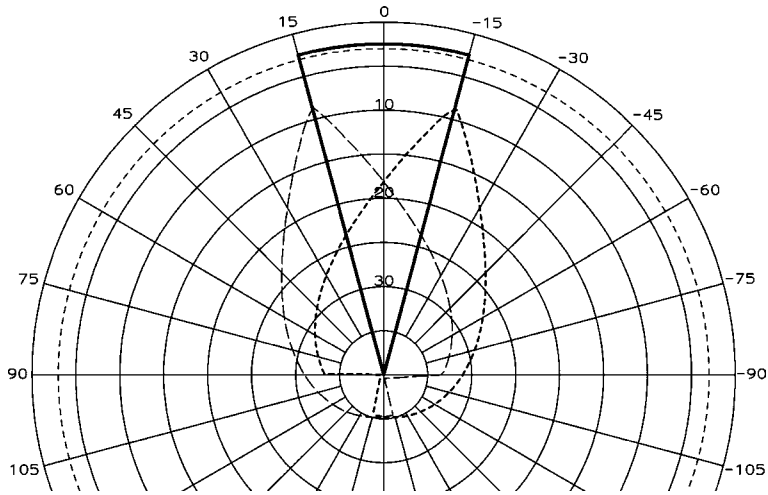
FIGURE 2-25   *E*-plane pattern of a rectangular horn with a GO term (solid curve) and edge diffractions (dashed curves).
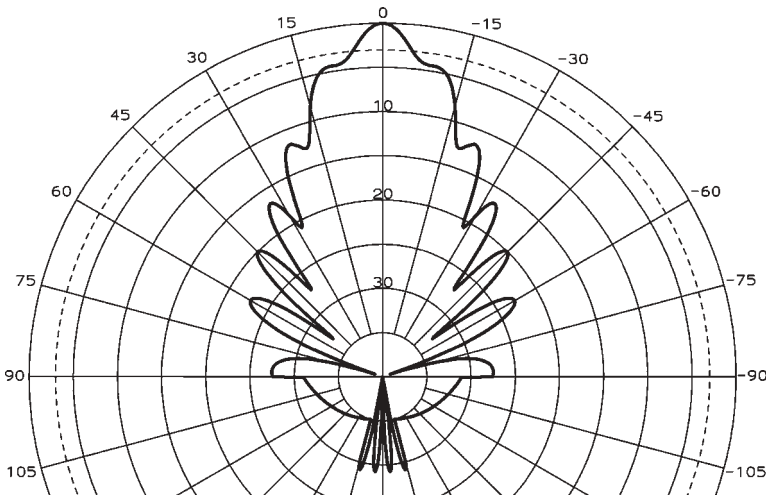

FIGURE 2-26   Combination of GO and edge diffractions in the *E*-plane pattern of a rectangular horn.

to the edge. We figure the same geometric factors for both edge and slope diffraction but now must calculate the normal derivative of the incident electric field. Figure 2-29 plots the *H*-plane pattern of the horn. The pattern fails to predict a pattern behind it. The *E*-plane diffraction produces a back hemisphere pattern for a real horn, but our two-dimensional model does not include the *E*-plane.

2-7.5   Amplitude Variations Along a Ray

Power decreases in a general ray as the distance from the source increases. If we expand the constant-phase surface (eikonal) about the ray in a Taylor series, we obtain a surface described by its radii of curvature [20, p. 55]. The maximum and minimum
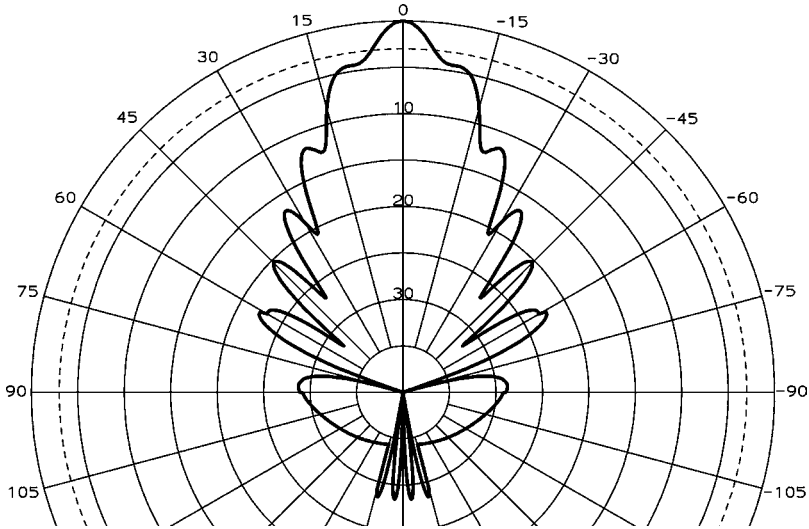
FIGURE 2-27 *E*-plane pattern of a rectangular horn combining GTD terms of direct GO, edge diffractions, and double diffractions between edges.
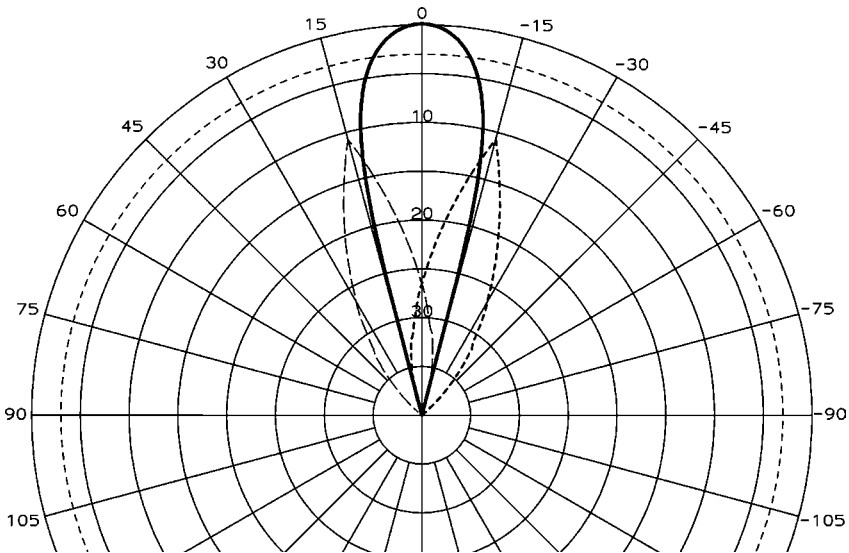

FIGURE 2-28 *H*-plane pattern of a rectangular horn with a GO term (solid curve) and edge slope diffractions (dashed curves).

values lie in the orthogonal principal planes. These radii of curvature determine the amplitude spread of the wave from point to point on the ray. We compute the ratio of differential areas about the ray at two locations as

$$\frac{dA_2}{dA_1} = \frac{\rho_1 \rho_2}{(\rho_1 + d)(\rho_2 + d)} \qquad (2\text{-}61)$$
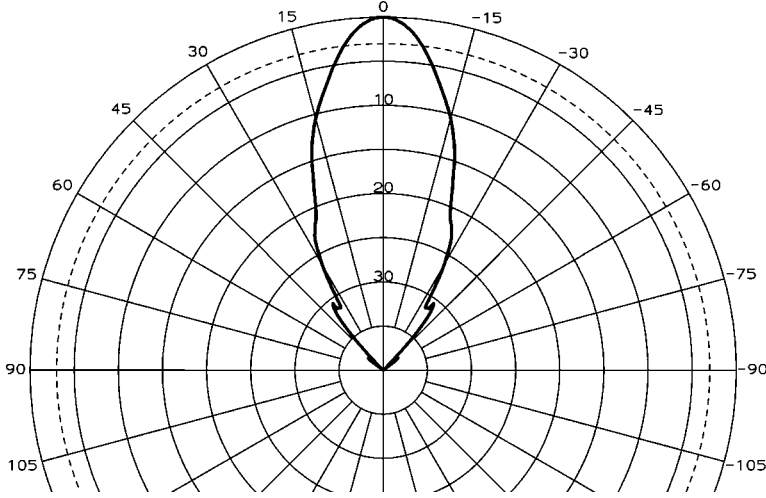
FIGURE 2-29 H-plane pattern of a rectangular horn by GTD analysis by combining direct GO field and edge slope diffraction.
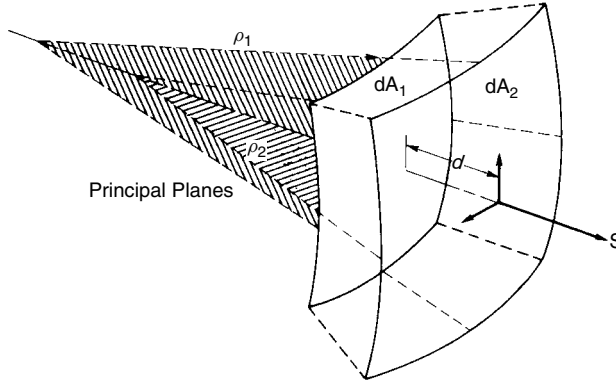

FIGURE 2-30   Astigmatic ray.

where ρ_1 and ρ_2 are the principal radii of curvature and d is the distance between two points on the ray (Figure 2-30). The electric field variation along the ray becomes

$$E_0 e^{-jkd} \sqrt{\frac{\rho_1 \rho_2}{(\rho_1 + d)(\rho_2 + d)}} \tag{2-62}$$

for the astigmatic ray spreading from unequal radii of curvature. When $d = -\rho_1$ or $d = -\rho_2$, GO fails because it predicts an infinite power density. We call these locations *caustics*. Remember that the ray always has differential area and never has any real area as implied by Figure 2-30. We have three special cases of the astigmatic ray:

1. Spherical wave, $\rho_1 = \rho$:

$$E_0 e^{-jkd} \frac{\rho}{\rho + d} \tag{2-63}$$

2. Cylindrical wave, $\rho_1 = \infty$:

$$E_0 e^{-jkd} \sqrt{\frac{\rho}{\rho + d}} \tag{2-64}$$

3. Plane wave, $\rho_1 = \rho_2 = \infty$:

$$E_0 e^{-jkd} \tag{2-65}$$

The plane wave does not spread but has constant amplitude as distance changes. Both cylindrical and plane waves require infinite power, and they are therefore nonphysical, but we find them convenient mathematically.

2-7.6 Extra Phase Shift Through Caustics

We cannot determine the ray amplitude at a caustic but can determine its amplitude and phase on either side. Passage through a caustic causes an extra phase shift to the ray [21, p. 31]. The denominator factors in the square root of Eq. (2-62) produce a $180°$ sign change when the ray distance factor d passes through either ρ_1 or ρ_2. The square root changes $180°$ to $+90°$ ($e^{j\pi/2}$) or $-90°$ ($e^{-j\pi/2}$), depending on the direction of movement along the ray. When tracing a ray moving through a caustic in the direction of propagation, you multiply by $e^{j\pi/2}$. The field is multiplied by $e^{-j\pi/2}$ for a ray traced in the opposite direction of propagation.

2-7.7 Snell's Laws and Reflection

We derive Snell's laws of reflection and refraction from Fermat's principle. The two laws of reflection are given as:

1. The incident ray, the reflected ray, and the normal of the reflecting surface at the point of reflection lie in the same plane.
2. The incident and reflected rays make equal angles with the surface normal.

Implicit in Snell's laws is the idea that locally the wavefront behaves like a plane wave and that the reflector can be treated as a plane. Given the direction of the incident ray $\mathbf{S}_1$, reflected ray $\mathbf{S}_2$, and the reflector normal $\mathbf{n}$, Snell's laws of reflection can be expressed vectorially [22]:

$$\mathbf{n} \times (\mathbf{S}_2 - \mathbf{S}_1) = 0 \qquad \mathbf{n} \cdot (\mathbf{S}_1 + \mathbf{S}_1) = 0 \tag{2-66}$$

We combine Eq. (2-66) to determine the ray directions before or after reflection:

$$\mathbf{S}_1 = \mathbf{S}_2 - 2(\mathbf{S}_2 \cdot \mathbf{n})\mathbf{n} \qquad \mathbf{S}_2 = \mathbf{S}_1 - 2(\mathbf{S}_1 \cdot \mathbf{n})\mathbf{n} \tag{2-67}$$

Snell's law of refraction can also be expressed vectorially as

$$\mathbf{n} \times (n_2 \mathbf{S}_2 - n_1 \mathbf{S}_1) = 0 \tag{2-68}$$

where n_1 and n_2 are the index of refractions in the two mediums.

2-7.8   Polarization Effects in Reflections

The electric field is orthogonal to the ray direction (a free-space wave) and is described by a two-dimensional polarization space (Section 1-11). We can describe polarization in any conveniently rotated two-dimensional basis vectors in the plane with the ray vector as its normal. We will use a ray-fixed coordinate system that changes direction after a reflection:

$$\mathbf{E}_i = \mathbf{a}_{\|}^i E_{i\|} + \mathbf{a}_{\perp}^i E_{i\perp} \tag{2-69}$$

where $\mathbf{a}_{\|}^i$ is a unit vector in the plane of incidence and $\mathbf{a}_{\perp}^i$ is perpendicular to the plane of incidence. We compute $\mathbf{a}_{\perp}^i$ from the normal to the plane $\mathbf{n}$ at the reflection point and the incident ray unit vector $\mathbf{S}_i$:

$$\mathbf{a}_{\perp}^i = \frac{\mathbf{S}_i \times \mathbf{n}}{|\mathbf{S}_i \times \mathbf{n}|}$$
$$\mathbf{a}_{\|}^i = \mathbf{a}_{\perp}^i \times \mathbf{S}_i \tag{2-70}$$

After reflection, we calculate the output ray-fixed polarization vectors using the output ray $\mathbf{S}_r$:

$$\mathbf{a}_{\perp}^r = \mathbf{a}_{\perp}^i \quad \text{and} \quad \mathbf{a}_{\|}^r = \mathbf{a}_{\perp}^r \times \mathbf{S}_r$$

$E_{i\|}$ is the incident electric field in the direction of $\mathbf{a}_{\|}^i$ and $E_{i\perp}$ is in the direction of $\mathbf{a}_{\perp}^i$. Of course, the direction of unit vector $\mathbf{a}_{\|}$ changes from incident to reflected rays. The electric field parallel to the reflector surface must vanish on the conductor surface:

$$E_{r\perp} = -E_{i\perp} \tag{2-71}$$

where $E_{r\perp}$ is the reflected field along $\mathbf{a}_{\perp}^i$. We calculate the reflection properties of $E_{\|}$ from the corresponding magnetic fields parallel to the surface:

$$H_{r\|} = H_{i\|} \tag{2-72}$$

By combining Eqs. (2-71) and (2-72), we obtain the dyadic relation for the ray-fixed coordinate system:

$$\begin{bmatrix} E_{r\|} \\ E_{r\perp} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} E_{i\|} \\ E_{i\perp} \end{bmatrix} \tag{2-73}$$

where $E_{r\|}$ and $E_{i\perp}$ are the reflected field components. At each reflection we rotate the polarizations to align $\mathbf{a}_{i\perp}$ with the normal to the plane of incidence. We can express Eq. (2-73) as a dyadic in terms of the incident and reflected wave polarization vectors $\vec{\mathbf{R}} = \mathbf{a}_{\|}^i \mathbf{a}_{\|}^r - \mathbf{a}_{\perp}^i \mathbf{a}_{\perp}^r$. Of course, the alternative method is to describe polarizations in a fixed three-dimensional coordinate system, but it requires a 3×3 reflection matrix.

2-7.9   Reflection from a Curved Surface

A wave reflected from a curved surface changes its radii of curvature and principal planes. The field along the reflected ray is given by

$$\mathbf{E}_r(s) = \mathbf{E}_{i0} \cdot \vec{\mathbf{R}} \sqrt{\frac{\rho_1^r \rho_2^r}{(\rho_1^r + s)(\rho_2^r + s)}} e^{-jks} \tag{2-74}$$

where s is the distance along the ray from the reflection, ρ_1 and ρ_2 the reflected ray radii of curvature, and $\bar{\mathbf{R}}$ the reflection dyadic. $\mathbf{E}_{i0}$ the incident ray electric field. For a flat surface we use images of the incident ray caustics for ρ_1^r and ρ_2^r, but in general, ρ_1^r and ρ_2^r become

$$\frac{1}{\rho_1^r} = \frac{1}{2}\left(\frac{1}{\rho_1^i} + \frac{1}{\rho_2^i}\right) + \frac{1}{f_1} \quad \frac{1}{\rho_2^r} = \frac{1}{2}\left(\frac{1}{\rho_1^i} + \frac{1}{\rho_2^i}\right) + \frac{1}{f_2} \tag{2-75}$$

where f_1 and f_2 are generalized focal lengths of the surface. The spreading factor of Eq. (2-74) simplifies in the far field:

$$\sqrt{\frac{\rho_1^r \rho_2^r}{(\rho_1^r + s)(\rho_2^r + s)}} \approx \frac{\sqrt{\rho_1^r \rho_2^r}}{s}$$

Kouyoumjian and Pathak [23] derived formulas for the focal lengths of a surface. We start with a surface with principal radii of curvature R_1 and R_2 with directions $\mathbf{u}_1$ and $\mathbf{u}_2$ at the point of reflection. For an incident ray with principal axes defined by unit vectors $\mathbf{x}_1^i$ and $\mathbf{x}_2^i$, we define a matrix relation between the incident ray and surface principal curvature directions:

$$\theta = \begin{bmatrix} \mathbf{x}_1^i \cdot \mathbf{u}_1 & \mathbf{x}_1^i \cdot \mathbf{u}_2 \\ \mathbf{x}_2^i \cdot \mathbf{u}_1 & \mathbf{x}_2^i \cdot \mathbf{u}_2 \end{bmatrix} \tag{2-76}$$

where the determinant is $|\theta| = (\mathbf{x}_1^i \cdot \mathbf{u}_1)(\mathbf{x}_2^i \cdot \mathbf{u}_2) - (\mathbf{x}_2^i \cdot \mathbf{u}_1)(\mathbf{x}_1^i \cdot \mathbf{u}_2)$. Given the angle of incidence θ^i, the following are the focal lengths:

$$\frac{1}{f_{1,2}} = \frac{\cos\theta^i}{|\theta|^2}\left(\frac{\theta_{22}^2 + \theta_{12}^2}{R_1} + \frac{\theta_{21}^2 + \theta_{11}^2}{R_2}\right)$$

$$\pm \frac{1}{2}\left\{\left(\frac{1}{\rho_1^i} - \frac{1}{\rho_2^i}\right)^2 + \left(\frac{1}{\rho_1^i} - \frac{1}{\rho_2^i}\right)\frac{4\cos\theta^i}{|\theta|^2}\left(\frac{\theta_{22}^2 - \theta_{12}^2}{R_1} + \frac{\theta_{21}^2 - \theta_{11}^2}{R_2}\right)\right.$$

$$\left. + \frac{4\cos^2\theta^i}{|\theta|^4}\left[\left(\frac{\theta_{22}^2 + \theta_{12}^2}{R_1} + \frac{\theta_{21}^2 + \theta_{11}^2}{R_2}\right)^2 - \frac{4|\theta|^2}{R_1 R_2}\right]\right\}^{1/2} \tag{2-77}$$

With a single reflection, we need not compute the direction of the principal axes. We need only the focal lengths. Multiple reflections require knowledge of the reflected-ray principal plane directions. Define the following matrices to determine the directions of the principal axes after reflection:

$$Q_0^i = \begin{bmatrix} \dfrac{1}{\rho_1^i} & 0 \\ 0 & \dfrac{1}{\rho_2^i} \end{bmatrix} \qquad C_0 = \begin{bmatrix} \dfrac{1}{R_1} & 0 \\ 0 & \dfrac{1}{R_2} \end{bmatrix}$$

$$Q^r = Q_0^i + 2(\theta^{-1})^{\mathrm{T}} C_0 \theta^{-1} \cos\theta_i$$

$$\mathbf{b}_1^r = \mathbf{x}_1^i - 2(\mathbf{n}\cdot\mathbf{x}_1^i)\mathbf{n} \qquad \mathbf{b}_2^r = \mathbf{x}_2^i - 2(\mathbf{n}\cdot\mathbf{x}_2^i)\mathbf{n}$$

where **n** is the surface normal at the reflection point. One principal-axis direction is

$$\mathbf{x}_1^r = \frac{\left(Q_{22}^r - 1/\rho_1^r\right)\mathbf{b}_1^r - Q_{12}^r\mathbf{b}_2^r}{\sqrt{\left(Q_{22}^r - 1/\rho_1^r\right)^2 + (Q_{12}^r)^2}} \tag{2-78}$$

We derive the other from the cross product of Eq. (2-78) and the reflected ray unit vector:

$$\mathbf{x}_2^r = -\mathbf{S}_r \times \mathbf{x}_1^r \tag{2-79}$$

We must reapply Eqs. (2-75) through (2-79) for every reflection.

We use Eqs. (2-75) through (2-79) for analysis, but except for computer optimizations, they cannot be applied directly to synthesis. If we limit the reflectors to figures of rotation, the radii of curvature are given by the meridians and parallels and these problems reduce to two dimensions. Similarly, a cylindrical reflector fed with a cylindrical wave [Eq. (2-64)] reduces the problem to two dimensions. The incident and reflected waves remain in the single plane chosen for the reflector analysis.

2-7.10   Ray Tracing

Tracing rays through a reflector system is conceptually straightforward. Where a ray strikes a reflector, we compute the normal to the surface. By using Eq. (2-67), we solve for the reflected-ray direction. Equation (2-73) determines the polarization effects when we express the incident and reflected rays in the ray-fixed coordinates. We use geometric arguments to determine the amplitude variation along the ray through the reflection instead of the general expressions given above. We experience difficulty when we try to discover the reflection points for given field and source points. No analytical expressions exist for calculating the reflection point of a general surface. The usual computer routines search for the minimum optical path length (Fermat's principle) without using Eq. (2-67), since a local minimum will satisfy this equation.

2-7.11   Edge Diffraction

Keller [24] extended the idea of reflection to edge diffraction by applying a generalized Fermat's principle to the rays. Figure 2-31 illustrates the rays in edge diffraction and the associated polarization directions. The figure shows the edge vector at the diffraction point. The vector cross product between the edge vector and the incident ray points in the direction of the incident plane normal. We measure the diffraction angle of incidence in this plane between the incident ray and the edge normal. Because diffraction obeys a generalized Fermat's principle, the diffracted ray exits at the same angle, similar to the reflected ray angles. The diffracted rays lie in a cone with the edge vector as its axis. The diffracted rays spread the incident power into a cone. Figure 2-31 shows a particular diffracted ray and how we determine the diffracted ray exiting plane.

We define diffracted ray polarization in terms of the incident and diffracted planes. The vectors are parallel and perpendicular to the two planes. Given the edge unit vector **e**, you compute the incident perpendicular polarization vector:

$$\mathbf{a}_{\phi'} = \frac{\mathbf{e} \times \mathbf{S}'}{\sin \beta_0} \tag{2-80}$$
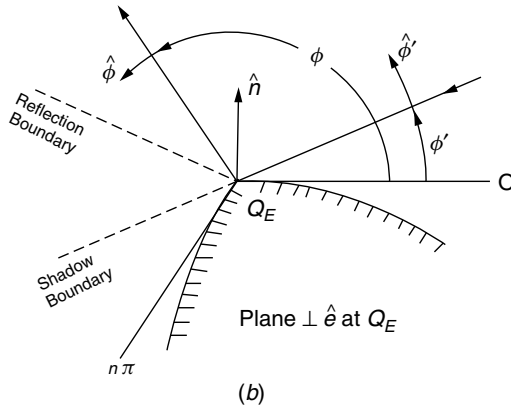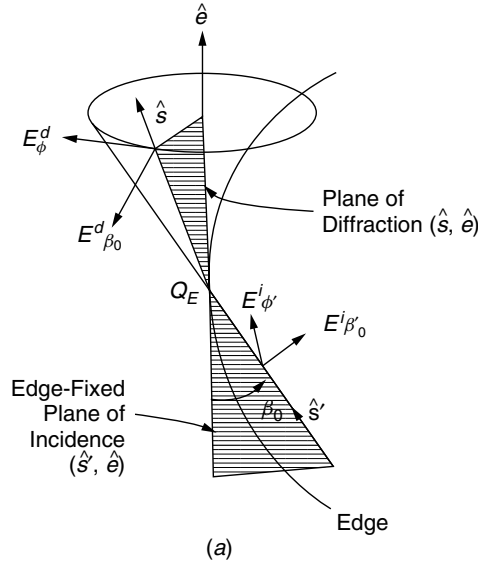
FIGURE 2-31 Ray-fixed coordinates related to edge-fixed coordinates at the edge diffraction point on a curved edge by showing planes of incidence and diffraction. (From [25], Fig. 5, © 1974 IEEE.)

where $\mathbf{S}'$ is the incident ray and β_0 is the angle between the edge tangent and the incident ray. The diffracted ray perpendicular polarization is similar to the incident ray

$$\mathbf{a}_\phi = -\frac{\mathbf{e} \times \mathbf{S}}{\sin \beta_0} \tag{2-81}$$

where $\mathbf{S}$ is the diffracted ray unit vector. We have the following vector relations for diffraction:

$$|\mathbf{e} \times \mathbf{S}| = |\mathbf{e} \times \mathbf{S}'| \quad \text{and} \quad \mathbf{e} \cdot \mathbf{S} = \mathbf{e} \cdot \mathbf{S}' \tag{2-82}$$

We determine the parallel polarization vector along the ray-fixed coordinates by the following cross products:

$$\mathbf{a}_{\phi'} \times \mathbf{a}_{\beta_0'} = \mathbf{S}' \quad \text{and} \quad \mathbf{a}_\phi \times \mathbf{a}_{\beta_0} = \mathbf{S} \tag{2-83}$$

By using ray-fixed coordinates, the diffraction matrix reduces to 2×2.

When $\beta_0 = \pi/2$, the parallel polarization components are parallel to the edge and the electric field vanishes: $E_{\beta_0'} + E_{\beta_0} = 0$. Acoustics calls this the *soft boundary condition* (Dirichlet); it operates on the parallel polarization components. The perpendicular components satisfy the *hard boundary condition* (Neumann). At a diffraction point Q_e we describe diffraction by the matrix equation

$$\begin{bmatrix} E_{\beta_0}^d(s) \\ E_\phi^d(s) \end{bmatrix} = \begin{bmatrix} -D_s & 0 \\ 0 & -D_h \end{bmatrix} \begin{bmatrix} E_{\beta_0'}^i(Q_e) \\ E_{\phi'}^i(Q_e) \end{bmatrix} \sqrt{\frac{\rho}{s(s+\rho)}} e^{-jks} \tag{2-84}$$

where s is the distance from the diffraction point. Diffraction locates one caustic on the diffraction point. We compute the second caustic distance ρ from the incident ray radius of curvature in the plane of incidence ρ_e^i and the edge curvature unit vector $\hat{\mathbf{n}}_e$:

$$\frac{1}{\rho} = \frac{1}{\rho_e^i} - \frac{\hat{\mathbf{n}}_e \cdot (\hat{\mathbf{s}}' - \hat{\mathbf{s}})}{a \sin^2 \beta_0} \tag{2-85}$$

where a is the edge radius of curvature. When $a \to \infty$ (straight edge), the second term of Eq. (2-85) vanishes.

A number of factors determine the wedge diffraction coefficients. The diffracting edge factors include (1) the angle between the faces, (2) the edge curvature, and (3) the curvature of the faces. The ray angle factors are (1) the incident angle relative to the edge tangent, (2) the diffraction angle to the shadow boundary, and (3) the angle to the reflection boundary. The diffraction coefficients peak at the shadow and reflection boundaries. UTD formulation uses characteristic lengths associated with incident and diffracted ray caustics. These many factors are beyond the current discussion.

2-7.12   Slope Diffraction

The spatial rate of change of the field normal to the edge produces slope diffraction, an added field component. This ray optics term also satisfies the generalized Fermat's principle with geometry determined by Eqs. (2-80) through (2-83), and (2-85). The slope diffraction equation has the same form as Eq. (2-84):

$$\begin{bmatrix} E_{\beta_0}^d(s) \\ E_\phi^d(s) \end{bmatrix} = \begin{bmatrix} -e_s & 0 \\ 0 & -e_h \end{bmatrix} \begin{bmatrix} E_{\beta_0'}^i(Q_e) \\ E_{\phi'}^i(Q_e) \end{bmatrix} \sqrt{\frac{\rho}{s(s+\rho)}} e^{-jks} \tag{2-86}$$

where the diffraction coefficients $e_{s,h}$ are related to the field derivative normal to the surface:

$$e_{s,h} = \frac{1}{jk \sin \beta_0} \left(\frac{\partial D_{s,h}}{\partial \phi'} \frac{\partial}{\partial n'} \right) \tag{2-87}$$

The term $\partial/\partial n'$ of Eq. (2-87) indicates the derivative of the incident fields given in the vector of Eq. (2-86). Equation (2-87) has the term $\partial D_{s,h}/d\phi'$ for the soft and hard slope diffraction terms returned from a subroutine; it is only a notational derivative.

2-7.13 Corner Diffraction

Every structural discontinuity diffracts waves. We derive edge diffraction from an infinite wedge where the wedge terminations (corners) produce diffracted rays. Recall from Section 2-4.2 that PTD added currents at edges to handle the effect of not having an infinite surface; the formulation for corner diffraction uses equivalent currents to derive these coefficients. We handle edge diffraction from each edge as always. Since each corner arises from two edges, we compute separate corner diffraction for each edge, two terms per corner.

Whereas edge diffraction is bound to a cone, corner diffraction radiates in all directions. The edge must be visible from both the source and receive points before corner diffraction contributes. We must include corner diffraction in any three-dimensional problem. As the source and receiver become farther and farther away from the object, corner diffraction contributions dominate over edge diffractions since it is derived from equivalent currents.

2-7.14 Equivalent Currents

GTD fails to predict fields at caustics. In many cases we consider these points unimportant, but for those cases where we need the fields, equivalent currents provide the answer. We derive equivalent currents from edge diffraction, which then replaces it and we use them instead of edge diffraction for all pattern points. The use of currents reduces the problem to a PO solution and line integrals are required.

We relate the incident fields expressed in the ray fixed to equivalent currents:

$$I = \frac{2j}{\eta k} E^i_{\beta'_0} D_s \sqrt{2\pi k} \; e^{j\pi/4} \tag{2-88}$$

$$M = \frac{2j}{k} E^i_{\phi'_0} D_h \sqrt{2\pi k} \; e^{j\pi/4} \tag{2-89}$$

The soft and hard diffraction coefficients $D_{s,h}$ depend on the source and receiver positions. Since we calculate the fields using vector potentials or dyadic Green's functions, the formulation has no caustics. They are only associated with a geometric optics solution.

Equivalent currents allow the calculation of the fields directly behind a reflector near the axis. The GTD solution produces a caustic as all points along the rim "light up" for an axisymmetrical design. PTD uses equivalent currents in a similar but different way to calculate correct fields in the same region. Equivalent currents derived from the diffraction coefficients produce the entire solution, since the reflector blocks the incident field. In PO we continue to include the direct field and the induced current radiation on the reflector, but add the PTD current radiation. Realize that slope diffraction also adds to the equivalent currents.

2-7.15 Diffraction from Curved Surfaces [26, 27]

In one analytical approach to surface-wave radiation we postulate waves bound to a surface that radiate only from discontinuities. Surface waves on infinite structure do not radiate but attenuate exponentially away from the surface, because they are bound to it. We can formulate GTD as radiating from discontinuities, and this produces an approach

for fields radiated on the shadowed side of a curved body. The continuous discontinuity of the curved surface causes power to be radiated at every point in the shadow region. These waves radiate tangentially from a wave traveling along a geodesic and bound to the surface. Surface waves require a dielectric coating or a corrugated surface to slow and bind the wave to the surface. The surface curvature slows and binds the wave to the surface without the need for a dielectric or corrugated surface coating. The wave that propagates along the surface sheds power in rays tangentially to it.

The rays travel along a surface geodesic from the attachment point to the radiation point. The geodesic curve is a minimum distance path on the surface between two points. In differential geometry it has a broader meaning, but for our purpose, the minimum distance definition will serve. The curved surface diffraction satisfies a generalized Fermat's principle (minimum distance) as do all other terms of GTD. The best approach uses another ray-fixed coordinate along the surface where the vectors are normal and tangential to the surface at both the attachment and radiation (shedding) points.

Curved surface diffraction considers three types of problems with different formulations. Two of them start with an antenna mounted on the surface. We either calculate the pattern in the presence of the curved object or calculate the coupling to a second antenna also mounted on the curved object. The third case determines the field scattered for a source located off the surface. All three use the ray-fixed coordinates. We start with the surface normal $\hat{\mathbf{n}}$ and the tangent vector $\hat{\mathbf{t}}$ directed along the geodesic path. A vector cross product defines the third direction of the local coordinate system. We use the surface binormal $\hat{\mathbf{b}}$, and the three vectors form a triad: $\hat{\mathbf{n}} \times \hat{\mathbf{b}} = \hat{\mathbf{t}}$. On a general surface all three vectors change direction as the wave moves along the geodesic. We use the term *torsion* for a path with a changing binormal. A soft dyadic diffraction coefficient is used with fields aligned with the attachment point binormal and the tangential shedding point binormal. We apply the hard dyadic diffraction coefficient fields aligned along the normal vectors. No formulas exist for computing the attachment and shedding points on a general curved surface given the source and receive points. We usually start with a known diffraction and find other points by incrementing along the curve by small steps.

REFERENCES

1. L. Diaz and T. A. Milligan, *Antenna Engineering Using Physical Optics*, Artech House, Boston, 1996.

2. B. F. Harrington, *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York, 1961.

3. P. S. Hacker and H. E. Schrank, Range distance requirements for measuring low and ultralow sidelobe antenna patterns, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 5, September 1982, pp. 956–966.

4. K. Pontoppidan, ed., *Technical Description of Grasp 8*, Ticra, Copenhagen, 2000 (self-published and available at *www.ticra.com*).

5. J. H. Richmond, A reaction theorem and its application to antenna impedance calculations, *IRE Transactions on Antennas and Propagation*, vol. AP-9, no. 6, November 1961, pp. 515–520.

6. R. F. Harrington, *Field Computation by Moment Methods*, Macmillan, New York, 1968; reprinted by IEEE Press, New York, 1993.

7. J. H. Richmond, Radiation and scattering by thin-wire structures in homogeneous conducting medium, *IEEE Transactions on Antennas and Propagation*, vol. AP-22, no. 2, March 1974, p. 365 (see also ASAP wire code).

8. N. N. Wang, J. H. Richmond, and M. C. Gilreath, Sinusoidal reactance formulation for radiation from conducting structures, *IEEE Transactions on Antennas and Propagation*, vol. AP-23, no. 3, May 1975, pp. 376–382.

9. B. M. Kolundzija and A. R. Djordjevic, *Electromagnetic Modeling of Composite Metallic and Dielectric Structures*, Artech House, Boston, 2002.

10. A. R. Djordjevic et al., *AWAS for Windows Version 2.0: Analysis of Wire Antennas and Scatterers*, Artech House, Boston, 2002.

11. A. C. Ludwig, Wire grid modeling of surface, *IEEE Transactions on Antennas and Propagation*, vol. AP-35, no. 9, September 1987, pp. 1045–1048.

12. A. W. Glisson and D. R. Wilton, Simple and efficient numerical methods for problems of electromagnetic radiation and scattering from surfaces, *IEEE Transactions on Antennas and Propagation,* vol. AP-28, no. 5, September 1980, pp. 563–603.

13. S. M. Rao, D. R. Wilton, and A. W. Glisson, Electromagnetic scattering by surfaces of arbitrary shape, *IEEE Transactions on Antennas and Propagation,* vol. AP-30, no. 3, May 1982, pp. 409–418.

14. B. M. Kolundzija et al., *WIPL-D: Electromagnetic Modeling of Composite Metallic and Dielectric Structures*, Artech House, Boston, 2001.

15. P. -S. Kildal, *Foundations of Antennas*, Studentlitteratur, Lund, Sweden, 2000.

16. C. Scott, *The Spectral Domain Method in Electromagnetics*, Artech House, Boston, 1989.

17. K. S. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Transactions on Antennas and Propagation*, vol. 14, no. 3, May 1966, pp. 302–307.

18. K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*, CRC Press, Boca Raton, FL, 1993.

19. J. G. Maloney and G. S. Smith, Modeling of antennas, Chapter 7 in A. Taflove, ed., *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Artech House, Boston, 1998.

20. D. J. Struik, *Differential Geometry*, Addison-Wesley, Reading, MA, 1950.

21. D. A. McNamara, C. W. I. Pistorius, and J. A. G. Malherbe, *Introduction to the Uniform Geometrical Theory of Diffraction*, Artech House, Boston, 1990.

22. F. S. Holt, in R. E. Collin and F. J. Zucker, eds., *Antenna Theory,* Part 2, McGraw-Hill, New York, 1969.

23. R. Kouyoumjian and P. Pathak, The dyadic diffraction coefficient for a curved edge, *NASA CR-2401*, June 1974.

24. J. B. Keller, Geometrical theory of diffraction, *Journal of the Optical Society of America*, vol. 52, 1962, pp. 116–130.

25. R. G. Kouyoumjian and P. H. Pathak, A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface, *Proceedings of IEEE,* vol. 62, no. 11, November 1974, pp. 1448–1461.

26. P. H. Pathak, W. D. Burnside, and R. J. Marhefka, A uniform GTD analysis of the diffraction of electromagnetic waves by a smooth convex surface, *IEEE Transactions on Antennas and Propagation*, vol. AP-28, no. 5, September 1980.

27. P. H. Pathak, N. Wang, W. D. Burnside, and R. G. Kouyoumjian, A uniform GTD solution for the radiation from sources on a convex surface, *IEEE Transactions on Antennas and Propagation*, vol. AP-29, no. 4, July 1981.