CHAPTER 1
Antennas

1.1 INTRODUCTION

An antenna is defined by Webster’s Dictionary as “a usually metallic device (as a rod or wire) for radiating or receiving radio waves.” The *IEEE Standard Definitions of Terms for Antennas* (IEEE Std 145–1983)* defines the antenna or aerial as “a means for radiating or receiving radio waves.” In other words the antenna is the transitional structure between free-space and a guiding device, as shown in Figure 1.1. The guiding device or transmission line may take the form of a coaxial line or a hollow pipe (waveguide), and it is used to transport electromagnetic energy from the transmitting source to the antenna, or from the antenna to the receiver. In the former case, we have a transmitting antenna and in the latter a receiving antenna.

A transmission-line Thevenin equivalent of the antenna system of Figure 1.1 in the transmitting mode is shown in Figure 1.2 where the source is represented by an ideal generator with impedance $Z_g$, the transmission line is represented by a line with characteristic impedance $Z_c$, and the antenna is represented by a load $Z_A = (R_L + R_r) + jX_A$ connected to the transmission line. The Thevenin and Norton circuit equivalents of the antenna are also shown in Figure 2.27. The load resistance $R_L$ is used to represent the conduction and dielectric losses associated with the antenna structure while $R_r$, referred to as the radiation resistance, is used to represent radiation by the antenna. The reactance $X_A$ is used to represent the imaginary part of the impedance associated with radiation by the antenna. This is discussed more in detail in Sections 2.13 and 2.14. Under ideal conditions, energy generated by the source should be totally transferred to the radiation resistance $R_r$, which is used to represent radiation by the antenna. However, in a practical system there are conduction-dielectric losses due to the lossy nature of the transmission line and the antenna, as well as those due to reflections (mismatch) losses at the interface between the line and the antenna. Taking into account the internal impedance of the source and neglecting line and reflection (mismatch) losses, maximum power is delivered to the antenna under conjugate matching. This is discussed in Section 2.13.

The reflected waves from the interface create, along with the traveling waves from the source toward the antenna, constructive and destructive interference patterns, referred to as standing waves, inside the transmission line which represent pockets of energy concentrations and storage, typical of resonant devices. A typical standing wave pattern is shown dashed in Figure 1.2, while another is exhibited in Figure 1.15. If the antenna system is not properly designed, the transmission line

---

could act to a large degree as an energy storage element instead of as a wave guiding and energy transporting device. If the maximum field intensities of the standing wave are sufficiently large, they can cause arcing inside the transmission lines.

The losses due to the line, antenna, and the standing waves are undesirable. The losses due to the line can be minimized by selecting low-loss lines while those of the antenna can be decreased by

![Figure 1.2](image)  
**Figure 1.2** Transmission-line Thevenin equivalent of antenna in transmitting mode.
reducing the loss resistance represented by $R_L$ in Figure 1.2. The standing waves can be reduced, and the energy storage capacity of the line minimized, by matching the impedance of the antenna (load) to the characteristic impedance of the line. This is the same as matching loads to transmission lines, where the load here is the antenna, and is discussed more in detail in Section 9.7. An equivalent similar to that of Figure 1.2 is used to represent the antenna system in the receiving mode where the source is replaced by a receiver. All other parts of the transmission-line equivalent remain the same. The radiation resistance $R_r$ is used to represent in the receiving mode the transfer of energy from the free-space wave to the antenna. This is discussed in Section 2.13 and represented by the Thevenin and Norton circuit equivalents of Figure 2.27.

In addition to receiving or transmitting energy, an antenna in an advanced wireless system is usually required to optimize or accentuate the radiation energy in some directions and suppress it in others. Thus the antenna must also serve as a directional device in addition to a probing device. It must then take various forms to meet the particular need at hand, and it may be a piece of conducting wire, an aperture, a patch, an assembly of elements (array), a reflector, a lens, and so forth.

For wireless communication systems, the antenna is one of the most critical components. A good design of the antenna can relax system requirements and improve overall system performance. A typical example is the TV for which the overall broadcast reception can be improved by utilizing a high-performance antenna. The antenna serves to a communication system the same purpose that eyes and eyeglasses serve to a human.

The field of antennas is vigorous and dynamic, and over the last 60 years antenna technology has been an indispensable partner of the communications revolution. Many major advances that occurred during this period are in common use today; however, many more issues and challenges are facing us today, especially since the demands for system performances are even greater. Many of the major advances in antenna technology that have been completed in the 1970s through the early 1990s, those that were under way in the early 1990s, and signals of future discoveries and breakthroughs were captured in a special issue of the *Proceedings of the IEEE* (Vol. 80, No. 1, January 1992) devoted to Antennas. The introductory paper of this special issue [1] provides a carefully structured, elegant discussion of the fundamental principles of radiating elements and has been written as an introduction for the nonspecialist and a review for the expert.

### 1.2 TYPES OF ANTENNAS

We will now introduce and briefly discuss some forms of the various antenna types in order to get a glance as to what will be encountered in the remainder of the book.

#### 1.2.1 Wire Antennas

Wire antennas are familiar to the layman because they are seen virtually everywhere—on automobiles, buildings, ships, aircraft, spacecraft, and so on. There are various shapes of wire antennas such as a straight wire (dipole), loop, and helix which are shown in Figure 1.3. Loop antennas need not only be circular. They may take the form of a rectangle, square, ellipse, or any other configuration. The circular loop is the most common because of its simplicity in construction. Dipoles are discussed in more detail in Chapter 4, loops in Chapter 5, and helices in Chapter 10.

#### 1.2.2 Aperture Antennas

Aperture antennas may be more familiar to the layman today than in the past because of the increasing demand for more sophisticated forms of antennas and the utilization of higher frequencies. Some forms of aperture antennas are shown in Figure 1.4. Antennas of this type are very useful for aircraft and spacecraft applications, because they can be very conveniently flush-mounted on the skin of
ANTENNAS

Figure 1.3  Wire antenna configurations.

Figure 1.4  Aperture antenna configurations.
the aircraft or spacecraft. In addition, they can be covered with a dielectric material to protect them from hazardous conditions of the environment. Waveguide apertures are discussed in more detail in Chapter 12 while horns are examined in Chapter 13.

1.2.3 Microstrip Antennas

Microstrip antennas became very popular in the 1970s primarily for spaceborne applications. Today they are used for government and commercial applications. These antennas consist of a metallic patch on a grounded substrate. The metallic patch can take many different configurations, as shown in Figure 14.2. However, the rectangular and circular patches, shown in Figure 1.5, are the most popular because of ease of analysis and fabrication, and their attractive radiation characteristics, especially low cross-polarization radiation. The microstrip antennas are low profile, conformable to planar and nonplanar surfaces, simple and inexpensive to fabricate using modern printed-circuit technology, mechanically robust when mounted on rigid surfaces, compatible with MMIC designs, and very versatile in terms of resonant frequency, polarization, pattern, and impedance. These antennas can be mounted on the surface of high-performance aircraft, spacecraft, satellites, missiles, cars, and even mobile devices. They are discussed in more detail in Chapter 14.

1.2.4 Array Antennas

Many applications require radiation characteristics that may not be achievable by a single element. It may, however, be possible that an aggregate of radiating elements in an electrical and geometrical

(a) Rectangular

(b) Circular

Figure 1.5 Rectangular and circular microstrip (patch) antennas.
arrangement (an array) will result in the desired radiation characteristics. The arrangement of the array may be such that the radiation from the elements adds up to give a radiation maximum in a particular direction or directions, minimum in others, or otherwise as desired. Typical examples of arrays are shown in Figure 1.6. Usually the term array is reserved for an arrangement in which the individual radiators are separate as shown in Figures 1.6(a–c). However the same term is also used to describe an assembly of radiators mounted on a continuous structure, shown in Figure 1.6(d).

### 1.2.5 Reflector Antennas

The success in the exploration of outer space has resulted in the advancement of antenna theory. Because of the need to communicate over great distances, sophisticated forms of antennas had to be used in order to transmit and receive signals that had to travel millions of miles. A very common antenna form for such an application is a parabolic reflector shown in Figures 1.7(a) and (b). Antennas of this type have been built with diameters of 305 m or even larger. Such large dimensions are needed to achieve the high gain required to transmit or receive signals after millions of miles of travel. Another form of a reflector, although not as common as the parabolic, is the corner reflector, shown in Figure 1.7(c). These antennas are examined in detail in Chapter 15.

### 1.2.6 Lens Antennas

Lenses are primarily used to collimate incident divergent energy to prevent it from spreading in undesired directions. By properly shaping the geometrical configuration and choosing the appropriate material of the lenses, they can transform various forms of divergent energy into plane waves. They can be used in most of the same applications as are the parabolic reflectors, especially at
higher frequencies. Their dimensions and weight become exceedingly large at lower frequencies.

Lens antennas are classified according to the material from which they are constructed, or according to their geometrical shape. Some forms are shown in Figure 1.8 [2].

In summary, an ideal antenna is one that will radiate all the power delivered to it from the transmitter in a desired direction or directions. In practice, however, such ideal performances cannot be achieved but may be closely approached. Various types of antennas are available and each type can take different forms in order to achieve the desired radiation characteristics for the particular application. Throughout the book, the radiation characteristics of most of these antennas are discussed in detail.

### 1.3 RADIATION MECHANISM

One of the first questions that may be asked concerning antennas would be “how is radiation accomplished?” In other words, how are the electromagnetic fields generated by the source, contained and guided within the transmission line and antenna, and finally “detached” from the antenna to form a free-space wave? The best explanation may be given by an illustration. However, let us first examine some basic sources of radiation.

#### 1.3.1 Single Wire

Conducting wires are material whose prominent characteristic is the motion of electric charges and the creation of current. Let us assume that an electric volume charge density, represented by $q_v$, ...
(coulombs/m³), is distributed uniformly in a circular wire of cross-sectional area \( A \) and volume \( V \), as shown in Figure 1.9. The total charge \( Q \) within volume \( V \) is moving in the \( z \) direction with a uniform velocity \( v_z \) (meters/sec). It can be shown that the current density \( J_z \) (amperes/m²) over the cross section of the wire is given by [3]

\[
J_z = q_v v_z
\]

(1-1a)

If the wire is made of an ideal electric conductor, the current density \( J_s \) (amperes/m) resides on the surface of the wire and it is given by

\[
J_s = q_s v_z
\]

(1-1b)
where \( q_s \) (coulombs/m\(^2\)) is the surface charge density. If the wire is very thin (ideally zero radius), then the current in the wire can be represented by

\[
I_z = q_s v_z
\]  
(1-1c)

where \( q_l \) (coulombs/m) is the charge per unit length.

Instead of examining all three current densities, we will primarily concentrate on the very thin wire. The conclusions apply to all three. If the current is time varying, then the derivative of the current of (1-1c) can be written as

\[
\frac{dI_z}{dt} = q_l \frac{dv_z}{dt} = q_l a_z
\]  
(1-2)

where \( \frac{dv_z}{dt} = a_z \) (meters/sec\(^2\)) is the acceleration. If the wire is of length \( l \), then (1-2) can be written as

\[
\frac{l}{dI_z/dt} = l q_l \frac{dv_z}{dt} = l q_l a_z
\]  
(1-3)

Equation (1-3) is the basic relation between current and charge, and it also serves as the fundamental relation of electromagnetic radiation [4], [5]. It simply states that to create radiation, there must be a time-varying current or an acceleration (or deceleration) of charge. We usually refer to currents in time-harmonic applications while charge is most often mentioned in transients. To create charge acceleration (or deceleration) the wire must be curved, bent, discontinuous, or terminated [1], [4]. Periodic charge acceleration (or deceleration) or time-varying current is also created when charge is oscillating in a time-harmonic motion, as shown in Figure 1.17 for a \( \lambda/2 \) dipole. Therefore:

1. If a charge is not moving, current is not created and there is no radiation.
2. If charge is moving with a uniform velocity:
   a. There is no radiation if the wire is straight, and infinite in extent.
   b. There is radiation if the wire is curved, bent, discontinuous, terminated, or truncated, as shown in Figure 1.10.
3. If charge is oscillating in a time-motion, it radiates even if the wire is straight.

A qualitative understanding of the radiation mechanism may be obtained by considering a pulse source attached to an open-ended conducting wire, which may be connected to the ground through a discrete load at its open end, as shown in Figure 1.10(d). When the wire is initially energized, the charges (free electrons) in the wire are set in motion by the electrical lines of force created by the source. When charges are accelerated in the source-end of the wire and decelerated (negative acceleration with respect to original motion) during reflection from its end, it is suggested that radiated fields are produced at each end and along the remaining part of the wire, [1], [4]. Stronger radiation with a more broad frequency spectrum occurs if the pulses are of shorter or more compact duration while continuous time-harmonic oscillating charge produces, ideally, radiation of single frequency determined by the frequency of oscillation. The acceleration of the charges is accomplished by the external source in which forces set the charges in motion and produce the associated field radiated. The deceleration of the charges at the end of the wire is accomplished by the internal (self) forces associated with the induced field due to the buildup of charge concentration at the ends of the wire. The internal forces receive energy from the charge buildup as its velocity is reduced to zero at the ends of the wire. Therefore, charge acceleration due to an exciting electric field and deceleration due
to impedance discontinuities or smooth curves of the wire are mechanisms responsible for electromagnetic radiation. While both current density \( \mathbf{J}_c \) and charge density \( q_v \) appear as source terms in Maxwell’s equation, charge is viewed as a more fundamental quantity, especially for transient fields. Even though this interpretation of radiation is primarily used for transients, it can be used to explain steady state radiation [4].

### 1.3.2 Two-Wires

Let us consider a voltage source connected to a two-conductor transmission line which is connected to an antenna. This is shown in Figure 1.11(a). Applying a voltage across the two-conductor transmission line creates an electric field between the conductors. The electric field has associated with it electric lines of force which are tangent to the electric field at each point and their strength is proportional to the electric field intensity. The electric lines of force have a tendency to act on the free electrons (easily detachable from the atoms) associated with each conductor and force them to be displaced. The movement of the charges creates a current that in turn creates a magnetic field intensity. Associated with the magnetic field intensity are magnetic lines of force which are tangent to the magnetic field.
Figure 1.11 Source, transmission line, antenna, and detachment of electric field lines.

We have accepted that electric field lines start on positive charges and end on negative charges. They also can start on a positive charge and end at infinity, start at infinity and end on a negative charge, or form closed loops neither starting or ending on any charge. Magnetic field lines always form closed loops encircling current-carrying conductors because physically there are no magnetic charges. In some mathematical formulations, it is often convenient to introduce equivalent magnetic charges and magnetic currents to draw a parallel between solutions involving electric and magnetic sources.

The electric field lines drawn between the two conductors help to exhibit the distribution of charge. If we assume that the voltage source is sinusoidal, we expect the electric field between the conductors to also be sinusoidal with a period equal to that of the applied source. The relative magnitude of the electric field intensity is indicated by the density (bunching) of the lines of force with the arrows showing the relative direction (positive or negative). The creation of time-varying electric and magnetic fields between the conductors forms electromagnetic waves which travel along the transmission line, as shown in Figure 1.11(a). The electromagnetic waves enter the antenna and have associated
with them electric charges and corresponding currents. If we remove part of the antenna structure, as shown in Figure 1.11(b), free-space waves can be formed by “connecting” the open ends of the electric lines (shown dashed). The free-space waves are also periodic but a constant phase point $P_0$ moves outwardly with the speed of light and travels a distance of $\frac{\lambda}{2}$ (to $P_1$) in the time of one-half of a period. It has been shown [6] that close to the antenna the constant phase point $P_0$ moves faster than the speed of light but approaches the speed of light at points far away from the antenna (analogous to phase velocity inside a rectangular waveguide). Figure 1.12 displays the creation and travel of free-space waves by a prolate spheroid with $\frac{\lambda}{2}$ interfocal distance where $\lambda$ is the wavelength.

The free-space waves of a center-fed $\frac{\lambda}{2}$ dipole, except in the immediate vicinity of the antenna, are essentially the same as those of the prolate spheroid.

The question still unanswered is how the guided waves are detached from the antenna to create the free-space waves that are indicated as closed loops in Figures 1.11 and 1.12. Before we attempt to explain that, let us draw a parallel between the guided and free-space waves, and water waves [7] created by the dropping of a pebble in a calm body of water or initiated in some other manner. Once the disturbance in the water has been initiated, water waves are created which begin to travel outwardly. If the disturbance has been removed, the waves do not stop or extinguish themselves but continue their course of travel. If the disturbance persists, new waves are continuously created which lag in their travel behind the others. The same is true with the electromagnetic waves created by an electric disturbance. If the initial electric disturbance by the source is of a short duration, the created electromagnetic waves travel inside the transmission line, then into the antenna, and finally are radiated as free-space waves, even if the electric source has ceased to exist (as was with the water waves and their generating disturbance). If the electric disturbance is of a continuous nature, electromagnetic waves exist continuously and follow in their travel behind the others. This is shown in Figure 1.13 for a biconical antenna. When the electromagnetic waves are within the transmission line and antenna, their existence is associated with the presence of the charges inside the conductors. However, when the waves are radiated, they form closed loops and there are no charges to sustain
1.3.3 Dipole

Now let us attempt to explain the mechanism by which the electric lines of force are detached from the antenna to form the free-space waves. This will again be illustrated by an example of a small dipole antenna where the time of travel is negligible. This is only necessary to give a better physical interpretation of the detachment of the lines of force. Although a somewhat simplified mechanism, it does allow one to visualize the creation of the free-space waves. Figure 1.14(a) displays the lines of force created between the arms of a small center-fed dipole in the first quarter of the period during which time the charge has reached its maximum value (assuming a sinusoidal time variation) and the lines have traveled outwardly a radial distance $\lambda/4$. For this example, let us assume that the number of lines formed are three. During the next quarter of the period, the original three lines travel an additional $\lambda/4$ (a total of $\lambda/2$ from the initial point) and the charge density on the conductors begins to diminish. This can be thought of as being accomplished by introducing opposite charges which at the end of the first half of the period have neutralized the charges on the conductors. The lines of force created by the opposite charges are three and travel a distance $\lambda/4$ during the second quarter of the first half, and they are shown dashed in Figure 1.14(b). The end result is that there are three lines of force pointed upward in the first $\lambda/4$ distance and the same number of lines directed downward in the second $\lambda/4$. Since there is no net charge on the antenna, then the lines of force must have been forced to detach themselves from the conductors and to unite together to form closed loops. This is shown in Figure 1.14(c). In the remaining second half of the period, the same procedure is followed but in the opposite direction. After that, the process is repeated and continues indefinitely and electric field patterns, similar to those of Figure 1.12, are formed.

1.3.4 Computer Animation-Visualization of Radiation Problems

A difficulty that students usually confront is that the subject of electromagnetics is rather abstract, and it is hard to visualize electromagnetic wave propagation and interaction. With today’s advanced numerical and computational methods, and animation and visualization software and hardware, this dilemma can, to a large extent, be minimized. To address this problem, we have developed and
included in this chapter computer programs to animate and visualize three radiation mechanisms. Descriptions of the computer programs are found in the website created by the publisher for this book. Each problem is solved using the Finite-Difference Time-Domain (FD-TD) method [8]–[10], a method which solves Maxwell’s equations as a function of time in discrete time steps at discrete points in space. A picture of the fields can then be taken at each time step to create a video which can be viewed as a function of time. Other animation and visualization software, referred to as applets, are included in the book website.

The three radiation problems that are animated and can be visualized using the computer program of this chapter and included in the book website are:

a. *Infinite length line source (two-dimensional) excited by a single Gaussian pulse and radiating in an unbounded medium.*

b. *Infinite length line source (two-dimensional) excited by a single Gaussian pulse and radiating inside a perfectly electric conducting (PEC) square cylinder.*

c. *E-plane sectoral horn (two-dimensional form of Figure 13.2) excited by a continuous sinusoidal voltage source and radiating in an unbounded medium.*
In order to animate and then visualize each of the three radiation problems, the user needs MATLAB [11] and the MATLAB M-file, found in the publisher’s website for the book, to produce the corresponding FD-TD solution of each radiation problem. For each radiation problem, the M-File executed in MATLAB produces a video by taking a picture of the computational domain every third time step. The video is viewed as a function of time as the wave travels in the computational space.

A. Infinite Line Source in an Unbounded Medium (tm_open)
The first FD-TD solution is that of an infinite length line source excited by a single time-derivative Gaussian pulse, with a duration of approximately 0.4 nanoseconds, in a two-dimensional TM\textsuperscript{z} computational domain. The unbounded medium is simulated using a six-layer Berenger Perfectly Matched Layer (PML) Absorbing Boundary Condition (ABC) [9], [10] to truncate the computational space at a finite distance without, in principle, creating any reflections. Thus, the pulse travels radially outward creating a traveling type of a wavefront. The outward moving wavefronts are easily identified using the coloring scheme for the intensity (or gray scale for black and white monitors) when viewing the video. The video is created by the MATLAB M-File which produces the FD-TD solution by taking a picture of the computational domain every third time step. Each time step is 5 picoseconds while each FD-TD cell is 3 mm on a side. The video is 37 frames long covering 185 picoseconds of elapsed time. The entire computational space is 15.3 cm by 15.3 cm and is modeled by 2500 square FD-TD cells (50 \times 50), including 6 cells to implement the PML ABC.

B. Infinite Line Source in a PEC Square Cylinder (tm_box)
This problem is simulated similarly as that of the line source in an unbounded medium, including the characteristics of the pulse. The major difference is that the computational domain of this problem is truncated by PEC walls; therefore there is no need for PML ABC. For this problem the pulse travels in an outward direction and is reflected when it reaches the walls of the cylinder. The reflected pulse, along with the radially outward traveling pulse, interfere constructively and destructively with each other and create a standing type of a wavefront. The peaks and valleys of the modified wavefront can be easily identified when viewing the video, using the colored or gray scale intensity schemes. Sufficient time is allowed in the video to permit the pulse to travel from the source to the walls of the cylinder, return back to the source, and then return back to the walls of the cylinder. Each time step is 5 picoseconds and each FD-TD cell is 3 mm on a side. The video is 70 frames long covering 350 picoseconds of elapsed time. The square cylinder, and thus the computational space, has a cross section of 15.3 cm by 15.3 cm and is modeled using an area 50 by 50 FD-TD cells.

C. E-Plane Sectoral Horn in an Unbounded Medium (te_horn)
The E-plane sectoral horn is excited by a cosinusoidal voltage (CW) of 9.84 GHz in a TE\textsuperscript{z} computational domain, instead of the Gaussian pulse excitation of the previous two problems. The unbounded medium is implemented using an eight-layer Berenger PML ABC. The computational space is 25.4 cm by 25.4 cm and is modeled using 100 by 100 FD-TD cells (each square cell being 2.54 mm on a side). The video is 70 frames long covering 296 picoseconds of elapsed time and is created by taking a picture every third frame. Each time step is 4.23 picoseconds in duration. The horn has a total flare angle of 52° and its flared section is 2.62 cm long, is fed by a parallel plate 1 cm wide and 4.06 cm long, and has an aperture of 3.56 cm.

1.4 CURRENT DISTRIBUTION ON A THIN WIRE ANTENNA
In the preceding section we discussed the movement of the free electrons on the conductors representing the transmission line and the antenna. In order to illustrate the creation of the current distribution on a linear dipole, and its subsequent radiation, let us first begin with the geometry of a
lossless two-wire transmission line, as shown in Figure 1.15(a). The movement of the charges creates a traveling wave current, of magnitude $I_0/2$, along each of the wires. When the current arrives at the end of each of the wires, it undergoes a complete reflection (equal magnitude and 180° phase reversal). The reflected traveling wave, when combined with the incident traveling wave, forms in each wire a pure standing wave pattern of sinusoidal form as shown in Figure 1.15(a). The current in each wire undergoes a 180° phase reversal between adjoining half-cycles. This is indicated in Figure 1.15(a) by the reversal of the arrow direction. Radiation from each wire individually occurs because of the time-varying nature of the current and the termination of the wire.

For the two-wire balanced (symmetrical) transmission line, the current in a half-cycle of one wire is of the same magnitude but 180° out-of-phase from that in the corresponding half-cycle of the other wire. If in addition the spacing between the two wires is very small ($s \ll \lambda$), the fields radiated by the current of each wire are essentially cancelled by those of the other. The net result is an almost ideal (and desired) nonradiating transmission line.
As the section of the transmission line between \(0 \leq z \leq l/2\) begins to flare, as shown in Figure 1.15(b), it can be assumed that the current distribution is essentially unaltered in form in each of the wires. However, because the two wires of the flared section are not necessarily close to each other, the fields radiated by one do not necessarily cancel those of the other. Therefore, ideally, there is a net radiation by the transmission-line system.

Ultimately the flared section of the transmission line can take the form shown in Figure 1.15(c). This is the geometry of the widely used dipole antenna. Because of the standing wave current pattern, it is also classified as a standing wave antenna (as contrasted to the traveling wave antennas which will be discussed in detail in Chapter 10). If \(l < \lambda\), the phase of the current standing wave pattern in each arm is the same throughout its length. In addition, spatially it is oriented in the same direction as that of the other arm as shown in Figure 1.15(c). Thus the fields radiated by the two arms of the dipole (vertical parts of a flared transmission line) will primarily reinforce each other toward most directions of observation (the phase due to the relative position of each small part of each arm must also be included for a complete description of the radiation pattern formation).

If the diameter of each wire is very small \((d \ll \lambda)\), the ideal standing wave pattern of the current along the arms of the dipole is sinusoidal with a null at the end. However, its overall form depends on the length of each arm. For center-fed dipoles with \(l \ll \lambda\), \(l = \lambda/2\), \(\lambda/2 < l < \lambda\) and \(\lambda < l < 3\lambda/2\), the current patterns are illustrated in Figures 1.16(a–d). The current pattern of a very small dipole (usually \(\lambda/50 < l \leq \lambda/10\)) can be approximated by a triangular distribution since \(\sin(kl/2) \simeq kl/2\) when \(kl/2\) is very small. This is illustrated in Figure 1.16(a).

![Figure 1.16](image)

**Figure 1.16** Current distribution on linear dipoles.
Because of its cyclical spatial variations, the current standing wave pattern of a dipole longer than \( \lambda (l > \lambda) \) undergoes \( 180^\circ \) phase reversals between adjoining half-cycles. Therefore the current in all parts of the dipole does not have the same phase. This is demonstrated graphically in Figure 1.16(d) for \( \lambda < l < 3\lambda/2 \). In turn, the fields radiated by some parts of the dipole will not reinforce those of the others. As a result, significant interference and cancelling effects will be noted in the formation of the total radiation pattern. See Figure 4.11 for the pattern of a \( \lambda/2 \) dipole and Figure 4.7 for that of a 1.25\( \lambda \) dipole.

For a time-harmonic varying system of radian frequency \( \omega = 2\pi f \), the current standing wave patterns of Figure 1.16 represent the maximum current excitation for any time. The current variations, as a function of time, on a \( \lambda/2 \) center-fed dipole, are shown in Figure 1.17 for \( 0 \leq t \leq T/2 \) where \( T \) is the period. These variations can be obtained by multiplying the current standing wave pattern of Figure 1.16(b) by \( \cos(\omega t) \).

### 1.5 HISTORICAL ADVANCEMENT

The history of antennas [12] dates back to James Clerk Maxwell who unified the theories of electricity and magnetism, and eloquently represented their relations through a set of profound equations best known as Maxwell’s Equations. His work was first published in 1873 [13]. He also showed that light was electromagnetic and that both light and electromagnetic waves travel by wave disturbances of the same speed. In 1886, Professor Heinrich Rudolph Hertz demonstrated the first wireless electromagnetic system. He was able to produce in his laboratory at a wavelength of 4 m a spark in the gap of a transmitting \( \lambda/2 \) dipole which was then detected as a spark in the gap of a nearby loop. It was not until 1901 that Guglielmo Marconi was able to send signals over large distances. He performed, in 1901, the first transatlantic transmission from Poldhu in Cornwall, England, to St. John’s Newfoundland. His transmitting antenna consisted of 50 vertical wires in the form of a fan connected to ground through a spark transmitter. The wires were supported horizontally by a guyed wire between two 60-m wooden poles. The receiving antenna at St. John’s was a 200-m wire pulled and supported by a kite. This was the dawn of the antenna era.

From Marconi’s inception through the 1940s, antenna technology was primarily centered on wire related radiating elements and frequencies up to about UHF. It was not until World War II that modern antenna technology was launched and new elements (such as waveguide apertures, horns, reflectors)
were primarily introduced. Much of this work is captured in the book by Silver [14]. A contributing factor to this new era was the invention of microwave sources (such as the klystron and magnetron) with frequencies of 1 GHz and above.

While World War II launched a new era in antennas, advances made in computer architecture and technology during the 1960s through the 1990s have had a major impact on the advance of modern antenna technology, and they are expected to have an even greater influence on antenna engineering into the twenty-first century. Beginning primarily in the early 1960s, numerical methods were introduced that allowed previously intractable complex antenna system configurations to be analyzed and designed very accurately. In addition, asymptotic methods for both low frequencies (e.g., Moment Method (MM), Finite-Difference, Finite-Element) and high frequencies (e.g., Geometrical and Physical Theories of Diffraction) were introduced, contributing significantly to the maturity of the antenna field. While in the past antenna design may have been considered a secondary issue in overall system design, today it plays a critical role. In fact, many system successes rely on the design and performance of the antenna. Also, while in the first half of this century antenna technology may have been considered almost a “cut and try” operation, today it is truly an engineering science. Analysis and design methods are such that antenna system performance can be predicted with remarkable accuracy. In fact, many antenna designs proceed directly from the initial design stage to the prototype without intermediate testing. The level of confidence has increased tremendously.

The widespread interest in antennas is reflected by the large number of books written on the subject [15]. These have been classified under four categories: Fundamental, Handbooks, Measurements, and Specialized. This is an outstanding collection of books, and it reflects the popularity of the antenna subject, especially since the 1950s. Because of space limitations, only a partial list is included here [2], [5], [7], [16]–[39], including the first, second and third editions of this book in 1982, 1997, 2005. Some of other books are now out of print.

### 1.5.1 Antenna Elements

Prior to World War II most antenna elements were of the wire type (long wires, dipoles, helices, rhombuses, fans, etc.), and they were used either as single elements or in arrays. During and after World War II, many other radiators, some of which may have been known for some and others of which were relatively new, were put into service. This created a need for better understanding and optimization of their radiation characteristics. Many of these antennas were of the aperture type (such as open-ended waveguides, slots, horns, reflectors, lenses), and they have been used for communication, radar, remote sensing, and deep space applications both on airborne and earth-based platforms. Many of these operate in the microwave region and are discussed in Chapters 12, 13, 15 and in [40].

Prior to the 1950s, antennas with broadband pattern and impedance characteristics had bandwidths not much greater than about 2:1. In the 1950s, a breakthrough in antenna evolution was created which extended the maximum bandwidth to as great as 40:1 or more. Because the geometries of these antennas are specified by angles instead of linear dimensions, they have ideally an infinite bandwidth. Therefore, they are referred to as frequency independent. These antennas are primarily used in the 10–10,000 MHz region in a variety of applications including TV, point-to-point communications, feeds for reflectors and lenses, and many others. This class of antennas is discussed in more detail in Chapter 11 and in [41].

It was not until almost 20 years later that a fundamental new radiating element, which has received a lot of attention and many applications since its inception, was introduced. This occurred in the early 1970s when the microstrip or patch antennas was reported. This element is simple, lightweight, inexpensive, low profile, and conformal to the surface. These antennas are discussed in more detail in Chapter 14 and in [42].

Major advances in millimeter wave antennas have been made in recent years, including integrated antennas where active and passive circuits are combined with the radiating elements in one compact unit (monolithic form). These antennas are discussed in [43].
Specific radiation pattern requirements usually cannot be achieved by single antenna elements, because single elements usually have relatively wide radiation patterns and low values of directivity. To design antennas with very large directivities, it is usually necessary to increase the electrical size of the antenna. This can be accomplished by enlarging the electrical dimensions of the chosen single element. However, mechanical problems are usually associated with very large elements. An alternative way to achieve large directivities, without increasing the size of the individual elements, is to use multiple single elements to form an array. An array is a sampled version of a very large single element. In an array, the mechanical problems of large single elements are traded for the electrical problems associated with the feed networks of arrays. However, with today’s solid-state technology, very efficient and low-cost feed networks can be designed.

Arrays are the most versatile of antenna systems. They find wide applications not only in many spaceborne systems, but in many earthbound missions as well. In most cases, the elements of an array are identical; this is not necessary, but it is often more convenient, simpler, and more practical. With arrays, it is practical not only to synthesize almost any desired amplitude radiation pattern, but the main lobe can be scanned by controlling the relative phase excitation between the elements. This is most convenient for applications where the antenna system is not readily accessible, especially for spaceborne missions. The beamwidth of the main lobe along with the side lobe level can be controlled by the relative amplitude excitation (distribution) between the elements of the array. In fact, there is a trade-off between the beamwidth and the side lobe level based on the amplitude distribution. Analysis, design, and synthesis of arrays are discussed in Chapters 6 and 7. However, advances in array technology are reported in [44]–[48].

A new antenna array design referred to as smart antenna, based on basic technology of the 1970s and 1980s, is sparking interest especially for wireless applications. This antenna design, which combines antenna technology with that of digital signal processing (DSP), is discussed in some detail in Chapter 16.

1.5.2 Methods of Analysis

There is plethora of antenna elements, many of which exhibit intricate configurations. To analyze each as a boundary-value problem and obtain solutions in closed form, the antenna structure must be described by an orthogonal curvilinear coordinate system. This places severe restrictions on the type and number of antenna systems that can be analyzed using such a procedure. Therefore, other exact or approximate methods are often pursued. Two methods that in the last four decades have been preeminent in the analysis of many previously intractable antenna problems are the Integral Equation (IE) method and the Geometrical Theory of Diffraction (GTD).

The Integral Equation method casts the solution to the antenna problem in the form of an integral (hence its name) where the unknown, usually the induced current density, is part of the integrand. Numerical techniques, such as the Moment Method (MM), are then used to solve for the unknown. Once the current density is found, the radiation integrals of Chapter 3 are used to find the fields radiated and other systems parameters. This method is most convenient for wire-type antennas and more efficient for structures that are small electrically. One of the first objectives of this method is to formulate the IE for the problem at hand. In general, there are two type of IE’s. One is the Electric Field Integral Equation (EFIE), and it is based on the boundary condition of the total tangential electric field. The other is the Magnetic Field Integral Equation (MFIE), and it is based on the boundary condition that expresses the total electric current density induced on the surface in terms of the incident magnetic field. The MFIE is only valid for closed surfaces. For some problems, it is more convenient to formulate an EFIE, while for others it is more appropriate to use an MFIE. Advances, applications, and numerical issues of these methods are addressed in Chapter 8 and in [3] and [49].

When the dimensions of the radiating system are many wavelengths, low-frequency methods are not as computationally efficient. However, high-frequency asymptotic techniques can be used to analyze many problems that are otherwise mathematically intractable. One such method that has
received considerable attention and application over the years is the GTD/UTD, which is an extension of geometrical optics (GO), and it overcomes some of the limitations of GO by introducing a diffraction mechanism. The Geometrical/Uniform Theory of Diffraction is briefly discussed in Section 12.10. However, a detailed treatment is found in Chapter 13 of [3] while recent advances and applications are found in [50] and [51].

For structures that are not convenient to analyze by either of the two methods, a combination of the two is often used. Such a technique is referred to as a hybrid method, and it is described in detail in [52]. Another method, which has received a lot of attention in scattering, is the Finite-Difference Time-Domain (FDTD). This method has also been applied to antenna radiation problems [53]–[57]. A method that has gained a lot of momentum in its application to antenna problems is the Finite Element Method [58]–[63].

1.5.3 Some Future Challenges

Antenna engineering has enjoyed a very successful period during the 1940s–1990s. Responsible for its success have been the introduction and technological advances of some new elements of radiation, such as aperture antennas, reflectors, frequency independent antennas, and microstrip antennas. Excitement has been created by the advancement, utilization, and proliferation of Computational Electromagnetics (CEM) software that provides students, engineers, and scientists with versatile and indispensable tools for modeling, visualizing, animating, and interpreting EM phenomena and characteristics. In addition, with such tools, electrically large structures that are complex and may otherwise be intractable can be designed and analyzed to gain insight into the performance of systems in order to advance and improve their efficiency. Today, antenna engineering is a science based on fundamental principles.

Although a certain level of maturity has been attained, there are many challenging opportunities and problems to be solved. Phased array architecture integrating monolithic MIC technology is still a most challenging problem. Integration of new materials, such as metamaterials [64], artificial magnetic conductors and soft/hard surfaces [65], into antenna technology offers many opportunities to control, discipline, harness and manipulate the EM waves to design devices with desired and functional characteristics, and improved performance. Computational electromagnetics using supercomputing and parallel computing capabilities will model complex electromagnetic wave interactions, in both the frequency and time domains. Innovative antenna designs, such as those using smart antennas [66], and multifunction, multiband, ultra wide band, reconfigurable antennas and antenna systems [67], to perform complex and demanding system functions remain a challenge. New basic elements are always welcomed and offer refreshing opportunities. New applications include, but are not limited to nanotechnology, wireless communications, direct broadcast satellite systems, global positioning satellites (GPS), high-accuracy airborne navigation, security systems, global weather, earth resource systems, and others. Because of the many new applications, the lower portion of the EM spectrum has been saturated and the designs have been pushed to higher frequencies, including the millimeter wave and terahertz frequency bands.

1.6 MULTIMEDIA

In the publisher’s website for this book, the following multimedia resources related to this chapter are included:

a. Java-based interactive questionnaire with answers.

b. Three Matlab-based animation-visualization programs designated
   - tm_open
ANTENNAS

- \texttt{tm\_box}
- \texttt{te\_horn}

which are described in detail in Section 1.3.4 and the corresponding README file in the book website.

c. Power Point (PPT) viewgraphs.

REFERENCES