

1

Introduction

Antennas have been around now for nearly 125 years. In those 125 years wireless communication has become increasingly important. Personal mobile communication applications are putting huge constraints on the antennas that need to be housed in limited spaces. Therefore the common practice of wireless engineers to consider the antenna as a black-box component is not valid anymore. The modern wireless engineer needs to have a basic understanding of antenna theory. Before we dive into the derivation of antenna characteristics, however, we will – in this chapter – present a brief overview of antenna history and the mechanisms of radiation. Thus, a solid foundation will be presented for understanding antenna characteristics and their derivations.

1.1 The Early History of Antennas

When James Clerk Maxwell, in the 1860s, united electricity and magnetism into electromagnetism, he described light as – and proved it to be – an electromagnetic phenomenon. He predicted the existence of electromagnetic waves at radio frequencies, that is at much lower frequencies than light. In 1886, Maxwell was proven right by Heinrich Rudolf Hertz who – without realizing it himself¹ – created the first ever radio system, consisting of a transmitter and a receiver, see Figure 1.1.

The transmitting antenna, connected to a spark gap at the secondary windings of a conduction coil, was a dipole antenna. The receiving antenna was a loop antenna ending in a second spark gap. Hertz, who conducted his experiments at frequencies around 50 MHz, was able to create electromagnetic waves and to transmit and receive these waves by using antennas. This immediately raises two questions:

1. What is an antenna?
2. How is electromagnetic radiation created?

¹ Hertz was not after creating wireless communication but proving the Maxwell equations experimentally.

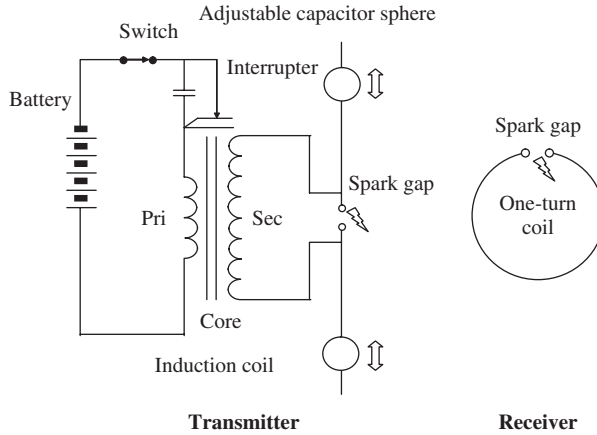


Figure 1.1 Hertz's radio system. With the receiving one-turn loop, small sparks could be observed when the transmitter discharged. From [1].

1.2 Antennas and Electromagnetic Radiation

From the previous it is obvious that

An antenna is a device for transmitting or receiving electromagnetic waves. An antenna converts electrical currents into electromagnetic waves (transmitting antenna) and vice versa (receiving antenna).

Before we describe this in detail, we will first take a closer look at the origin of electromagnetic radiation.

1.2.1 Electromagnetic Radiation

The source of electromagnetic radiation is accelerated (or decelerated) charge.

Let's start with a static, charged object and have a look at the electric field lines. These lines are the trajectories of a positively charged particle due to this static, charged object. Electric field lines are always directed perpendicular to the surface of a charged object and start and end on charged objects. Electric field lines due to single charged objects start at or extend towards infinity. For a positively charged object, the electric field lines start at the object and extend towards infinity, for a negatively charged object they start at infinity and end at the object.

For explaining the mechanisms of radiation, the direction of the electric field lines does not matter, therefore in Figure 1.2(a), where we show a uniformly moving particle at a certain instant of time, we do not indicate the direction of the field lines.

The uniformly charged particle is accelerated between $t = 0$ and $t = t_1$, see Figure 1.2(b), after which it continues its uniform movement. In Figure 1.2(b) we have indicated the position of the particle at the start ($t = 0$) and at the end ($t = t_1$) of the acceleration.

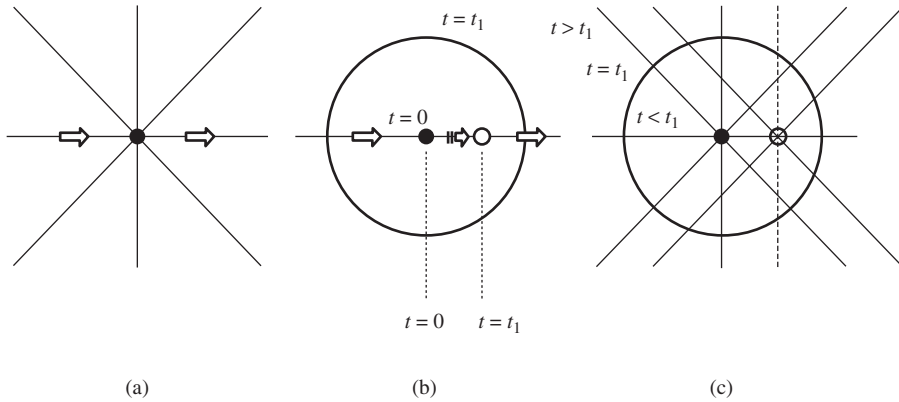


Figure 1.2 Electric field lines of a charged particle. (a) Field lines at a certain moment of time for a uniformly moving charged particle. (b) The particle is accelerated between $t = 0$ and $t = t_1$. The position of an observer, traveling with the speed of light along an electric field line at $t = t_1$ is indicated with the circle. (c) Electric field lines at $t = 0$ and $t = t_1$.

Also indicated is the position of an observer that has moved with the speed of light along a static electric field line from the particle, for the duration of the acceleration (t_1).

In Figure 1.2(c) we repeat Figure 1.2(b), where we now also indicate static electric field lines associated with the particle at $t = 0$ and at $t = t_1$.

We now think of ourselves positioned anywhere on the ‘observer circle’ and accept that nothing can move faster than the speed of light. Then, everywhere from the ‘observer circle’ to infinity, the static field lines must follow those associated with the particle position at $t = 0$. Everywhere inside the circle, the static field lines must follow those associated with the particle position at $t = t_1$. Since electric field lines must be continuous, so-called *kinks* must exist at the observer position to make the electric field lines connect, see Figure 1.3.²

Having explained the construction of electric field lines for an accelerated charged particle, we can now take a closer look at the electric field lines as a function of time. In Figure 1.4 we look at the electric field lines at different times within the acceleration time interval.

When we take the disturbances, that is the transverse components of the electric field, taken at the subsequent moments and add them in one graph, as in Figure 1.4, we see that these disturbances move out from the accelerated charge at the speed of light. Associated with the changing electric field is a changing magnetic field. Both fields are *in phase*³ since they are due to a single event. The electric and magnetic fields travel along in phase, their directions being perpendicular to each other. This is what we call an *electromagnetic wave*.

²The continuous electric field lines are shown a little displaced to clarify the construction from the initial and end position static electric field lines.

³As opposed to the situation for a coil or a capacitor.

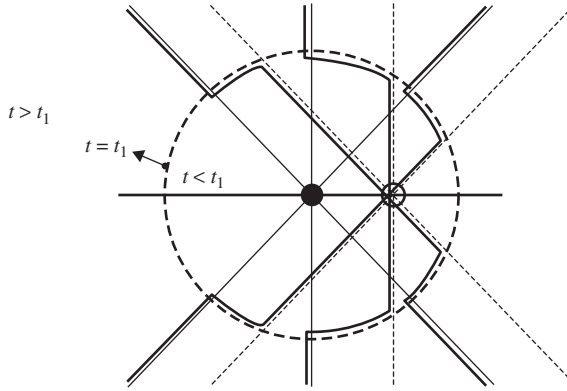


Figure 1.3 The electric field lines of a briefly accelerated charged particle must form *kinks* in order to connect the field lines associated with the initial and end position of the particle, thus forming continuous electric field lines.

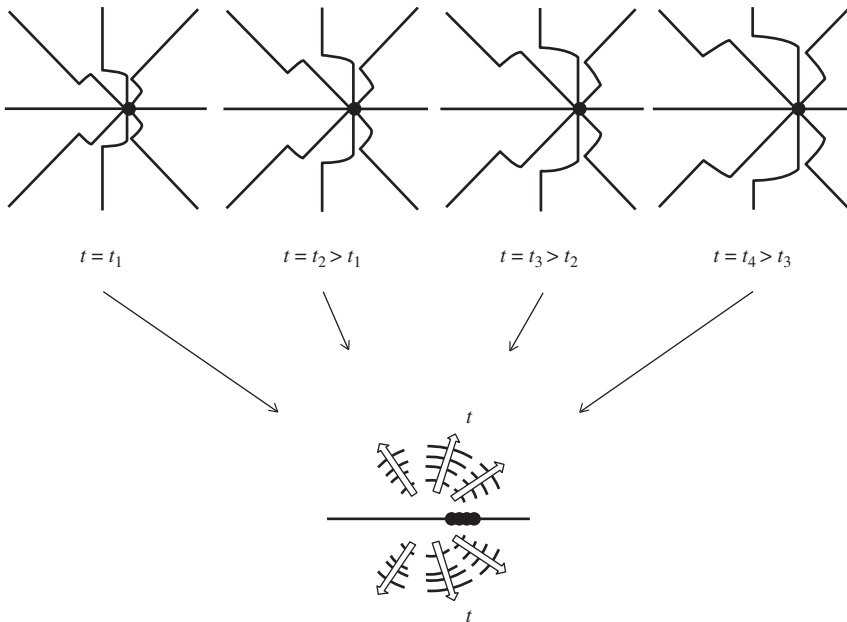


Figure 1.4 The electric field lines of a briefly accelerated charged particle at subsequent instances of time, and the resulting transverse field moving out at the speed of light.

Accelerating (or decelerating) charges may be found in electrically conducting wires at positions where the wire is bent, curved, discontinuous or terminated. Before we discuss the radiation from a wire dipole antenna in detail, we note that, see Figure 1.4, radiation does *not* take place in directions along the charged particle acceleration.

Next, we will take a look at the radiation from a short dipole antenna.

1.2.2 Short Wire Dipole Radiation

We consider two short – that is much shorter than a wavelength – electrically conducting wires, each folded back 90 degrees, and connected to an AC source. We will look at the electric field around this structure at different instances of time within one half of the period T , see Figure 1.5.

- $t = 0^+$, see **Figure 1.5(a)**. The time is a short while after $t = 0$. The source has been turned on and charge is accelerated from the source to the wire ends. Because of the accelerating charges at the feed point, a transverse electric field component is traveling outward, in a direction perpendicular to the wires. Since field lines have to be continuous and start and end perpendicular to a charged body, the electric field line takes the form as shown. Underneath the dipole, the current is shown as a function of time; the time of the snapshot (0^+) is indicated with a black dot.
- $t = (\frac{T}{4})^-$, see **Figure 1.5(b)**. At this moment, the current has reached its maximum value, its change with time has become zero. The electric field lines are as shown in the figure. The transverse electric field component that was created at $t = 0^+$ has traveled a distance of a quarter of a wavelength. New transverse electric field components have been created after the creation of this first one.
- $t = (\frac{T}{4})^+$, see **Figure 1.5(c)**. The current has become less than the maximum value and the time change of the current has changed sign. Charges are now accelerated into

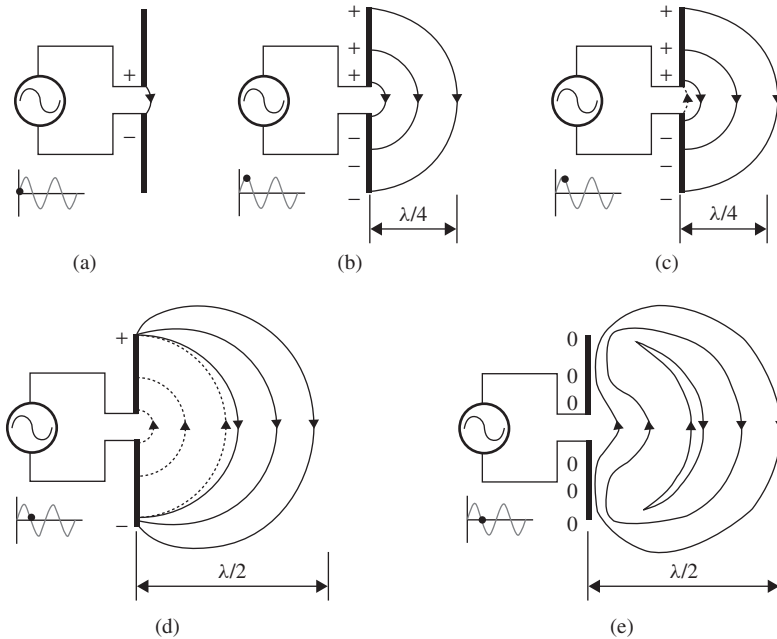


Figure 1.5 Electromagnetic radiation by charges in oscillatory acceleration. (a) $t = 0^+$. (b) $t = (\frac{T}{4})^-$. (c) $t = (\frac{T}{4})^+$. (d) $t = (\frac{T}{2})^-$. (e) $t = (\frac{T}{2})^+$.

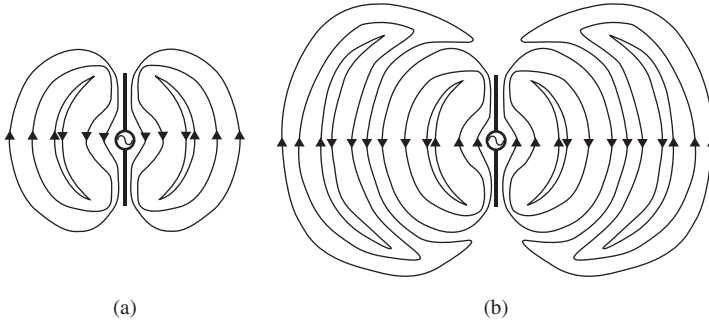


Figure 1.6 Detachment of electric field lines from a dipole antenna at different times. (a) $t = (\frac{T}{2})^+$. (b) $t = T^+$.

the opposite direction and new electric field lines, oppositely directed relative to the existing ones, may be thought of as being created.

- $t = (\frac{T}{2})^-$, see **Figure 1.5(d)**. The current amplitude has become very small and excess charges are only present at the dipole tips. Additional, upward-directed, transverse field lines have been created since $t = (\frac{T}{4})^+$. The first one of these has traveled a distance of nearly a quarter of a wavelength.
- $t = (\frac{T}{2})$, see **Figure 1.5(e)**. Both halves of the dipole antenna have become charge free. No excess charge is present and the current has become zero. The electric field lines do not need to be perpendicular to the conductors anymore, since these conductors have become charge-free. As a consequence, the field lines form closed loops and detach from the conductors.

For clarity reasons we have shown the mechanism of radiation from a dipole in a plane and only at one side of the dipole. Of course, in this plane, the radiation takes place at both sides, see **Figure 1.6**. In three dimensions, the field line pattern is rotationally symmetric around the axis of the dipole antenna. For the same clarity reasons we have left out magnetic field lines in our explanation. The magnetic field lines form closed loops around and perpendicular to the electric field lines.

Most wire antennas may be thought to consist of an infinite number of elementary (that is: infinitely small) dipole antennas.

With electromagnetic radiation and dipole antenna radiation now explained, we can continue with our short overview of antenna history.

1.3 The Modern History of Antennas

Guglielmo Marconi grasped the potential of Hertz's equipment and started experimenting with wireless telegraphy. In 1895 he hit upon a new arrangement of his equipment that suddenly allowed him to transmit and receive over distances that progressively increased up to and beyond 1.5 km [2, 3]. Marconi had enlarged the antenna. His monopole antenna,⁴

⁴ A monopole forms a dipole antenna together with its image in the ground.

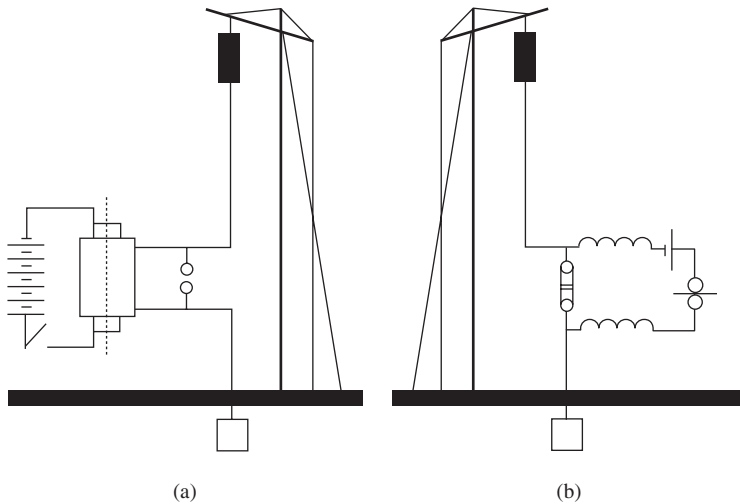


Figure 1.7 Marconi's antennas in 1895. (a) Scheme of the transmitter used by Marconi at Villa Griffone. (b) Scheme of the receiver used by Marconi at Villa Griffone. From [1].

see Figure 1.7, was resonant at a wavelength much larger than any that had been studied before and it was this creation of long-wavelength electromagnetic waves that turned out to be the key to his success. It was also Marconi who, in 1909, introduced the term *antenna* for the device that was formerly referred to as an *areal* or an *elevated wire* [3, 4].

The invention of the *thermionic valve* or *diode* by Fleming in 1905 and the *audion* or *triode* by Lee de Forest in 1907 paved the way for a reliable detection, reception and amplification of radio signals. From 1910 onwards broadcasting experiments were conducted that resulted in Europe in 1922 in the forming of the British Broadcasting Corporation (BBC) [5].

In the 1930s a return of interest to the higher end of the radio spectrum took place. This interest intensified with the outbreak of World War II with the immediate need for compact communication equipment as well as compact (airborne), high-resolution radar. Antenna design became a new specialism. At the end of World War II, antenna theory was mature to a level that made the analysis possible of, among others, free-standing dipole, horn and reflector antennas, monopole antennas, slots in waveguides and arrays thereof. The end of the War also saw the beginning of the development of electronic computers. The introduction of the IBM-PC⁵ in 1981 considerably helped in the development of numerical electromagnetic analysis software. The 1980s may be seen as the decade of numerical microwave circuit and planar antenna theory development. In this period the numerical electromagnetics code (NEC), for the analysis of wire antennas, was commercially distributed. The 1990s, however, may be seen as the decade of the numerical electromagnetic-based microwave circuit and (planar, integrated) antenna design. In 1989 Sonnet started distribution, followed in 1990 by HP

⁵ 4.77 MHz, 16 kb RAM, no hard drive.

(now Agilent), High Frequency Structure Simulator (HFSS).⁶ [6] These two numerical electromagnetic analysis tools were followed by Zeland's IE3D, Remcom's XFDTD, Agilent's Momentum, CST's Microwave Studio, FEKO from EM Software & Systems, and others.

Despite the diversity of the numerical electromagnetic analysis software commercially available today, it is the author's strong belief that there exists, and will continue to exist, a need to develop an understanding of electromagnetics and antenna theory. At least a basic understanding needs to be present to be able to evaluate the results obtained with numerical electromagnetic analysis software. And a bit more than a basic understanding is necessary for designing new antennas. In the design process, the choice for a particular type of antenna is mainly dictated by the volume available for the antenna, the frequency (directly related to this volume) and the distance over which wireless communication needs to be performed.

1.4 Frequency Spectrum and Antenna Types

The radio frequency (RF) radiation of electromagnetic waves is used for frequencies that lie roughly between 30 Hz and 300 GHz. Table 1.1 [7] lists a number of frequency bands, associated wavelengths⁷ and applications for these bands.

In the table we have made use of the IEEE-defined frequency band designations [8]. We stop at a frequency of 300 GHz, where infrared starts, followed by visible light from 400 THz upwards.

Of the many antenna types that exist, only a few basic ones will be discussed in detail in this book. Figure 1.8 shows these antenna types.

Other antenna types may be seen as combinations of these basic antennas (e.g. Yagi-Uda antennas are combinations of active and short-circuited dipole antennas placed in parallel) or derivatives of one of the basic antennas (e.g. a microstrip patch antenna may be seen as consisting of two rectangular aperture antennas). The detailed discussion of most of these antennas is beyond the scope of this short course in antenna theory and may be found in specialist textbooks, see for example [9].

1.4.1 Dipole Antennas

The short dipole antenna consists of two wires or circular tubes having a length much shorter than the wavelength and placed along the same axis, see Figures 1.5 and 1.6. The dipole antenna is voltage-excited in the small gap between the two dipole halves. This may be accomplished through a transmission line, connecting the gap to the voltage source. Short dipoles are used for radio broadcasting systems at VHF frequencies and below.

The resonant dipole antenna is an antenna where the length of the two wires or tubes together is a multiple (most often *one*) of half the wavelength. Half-wavelength dipole antennas are used for small-band applications at low GHz frequencies.

⁶ Currently Ansoft HFSS.

⁷ $\lambda = c/f$, where λ is the wavelength, c is the velocity of light and f is the frequency.

Table 1.1 RF Frequency band designations, wavelengths and applications

Frequency band	Frequencies	Wavelengths	Applications
extremely low frequency (ELF)	30–300 Hz	1000–10,000 km	Submarine communications
infralow frequency (ILF)	0.3–3 kHz	100–1000 km	–
very low frequency (VLF)	3–30 kHz	10–100 km	Navigation Weather
low frequency (LF)	30–300 kHz	1–10 km	Navigation Maritime communications Information and weather systems Time systems
medium frequency (MF)	0.3–3 MHz	0.1–1 km	Navigation AM radio Mobile radio
high frequency (HF)	3–30 MHz	10–100m	Citizen’s band (CB) Shortwave radio Mobile radio Maritime radio
very high frequency (VHF)	30–300 MHz	1–10m	Amateur radio VHF TV FM radio Mobile satellite Mobile radio Fixed radio
ultra high frequency (UHF)	0.3–3 GHz	0.1–1m	Microwave Satellite UHF TV Paging Cordless telephony Cellular telephony Wireless LAN
super high frequency (SHF)	3–30 GHz	1–10 cm	Microwave Satellite Wireless LAN
extremely high frequency (EHF)	30–300 GHz	1–10 mm	Microwave Satellite Radiolocation

1.4.2 Loop Antennas

Small loop antennas may be considered to be *magnetic* dipole antennas. The fields from a small loop antenna are similar to those from a small dipole antenna with the electric and magnetic fields interchanged, as we will see in a later chapter. Loop antennas are used, among others, in direction finding systems.

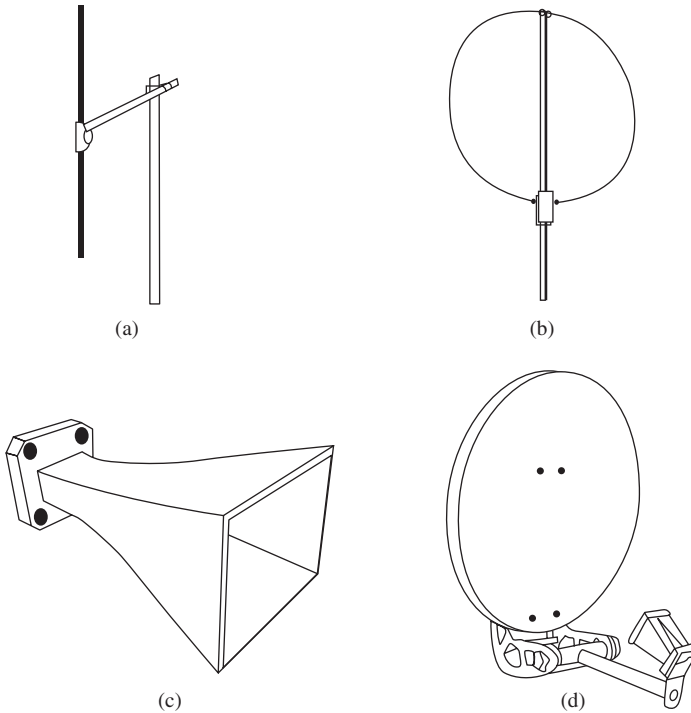


Figure 1.8 Basic antenna types. (a) Dipole antenna. (b) Loop antenna. (c) Horn antenna. (d) Parabolic reflector antenna.

1.4.3 Aperture Antennas

An aperture antenna consists of an ‘opening’ in a metallic surrounding. The fields across this opening, that is the aperture, radiate into free space. An electromagnetic horn is an antenna where the radiating aperture is matched to the waveguiding system that supports the excitation signal. This matching is accomplished through properly shaping the transition from waveguiding structure to aperture. Aperture antennas are used in the GHz range of frequencies.

1.4.4 Reflector Antennas

A reflector antenna is also an aperture antenna. A primary radiator (dipole or horn antenna) in the focal point of a parabolic reflector illuminates the reflector. The aperture formed by this reflector then radiates into free space. Since the radiated waves are concentrated into a beam, the width of which is inversely proportional to the size of the aperture as we will see in a later chapter, reflector antennas offer a convenient way to concentrate radiation. This allows long distance communication in the low and high GHz range of frequencies. Parabolic reflector antennas are used, among others, for satellite television transmission and reception, and for radar (radio detection and ranging).

1.4.5 Array Antennas

Antennas may be combined with similar or other types of antenna to form an array antenna. The antennas within the array we designate as *antenna elements* or *array elements* or, for short, *elements*. The combination of all the elements we designate *array antenna*, *antenna array* or *array*. Commonly, similar elements are positioned at regular intervals on a line (linear array antenna) or in a plane (planar array antenna). By forming an array, a radiation beam may be created having a small beamwidth. By electronically controlling the phase differences between the elements, we may electronically direct the beam in different directions without physically rotating the antenna. This possibility of electronic beam steering or *scanning* makes (phased) array antennas particularly attractive.

1.4.6 Modern Antennas

In this book we will briefly touch upon the subject of modern antennas. By ‘modern’ we mean antennas that can be considered as derivatives of the basic antennas that are dealt with in detail. Of all modern antennas we will, briefly, discuss the printed monopole antenna, the inverted F antenna (IFA) and the microstrip patch antenna, see Figure 1.9. These antennas may be encountered in today’s wireless devices.

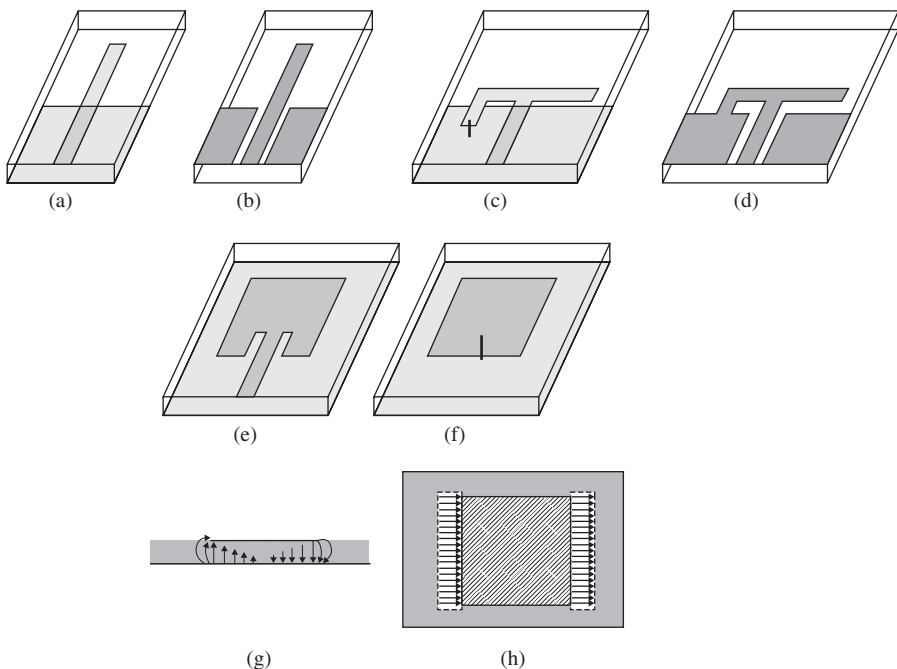


Figure 1.9 Modern antennas. (a) Microstrip printed monopole antenna. (b) Coplanar waveguide (CPW) printed monopole antenna. (c) Microstrip IFA. (d) CPW IFA. (e) Microstrip excited microstrip patch antenna. (f) Probe excited microstrip patch antenna. (g) Side view of electric field lines patch antenna. (h) Top view of electric field lines patch antenna.

The printed monopole antenna, either in microstrip technology (with a separate trace and a ground layer, see Figure 1.9(a) or in CPW technology (with a combined trace and ground layer, see Figure 1.9(b)), may be regarded as an asymmetric printed dipole antenna, the arms being the monopole and the ground. The asymmetry causes differences in the current distributions in the two arms which results in some disturbances in the dipole radiation pattern. To avoid the existence of multiple lobes, the length of the ground plane should remain smaller than a quarter of a wavelength.

The inverted F antenna (IFA) may be realized in microstrip technology, see Figure 1.9(c), or in CPW technology, see Figure 1.9(d). The IFA may be regarded as a printed monopole antenna where the top section has been folded down. The folded-down part, being parallel to the ground plane, introduces capacitance to the input impedance. This additional capacitance is compensated for by a short circuited stub. The realization of the short circuit is straight forward in CPW technology, but requires a via to the ground plane in microstrip technology, see Figure 1.9(c).

A microstrip patch antenna, see Figures 1.9(e), 1.9(f), may be considered as a cavity with electrical conducting top and bottom and magnetically conducting side walls. The fields inside the ‘cavity’ may be excited by, among others, a microstrip transmission line, see Figure 1.9(e), or a probe, see Figure 1.9(f). Since the walls of this ‘cavity’ are not perfectly conducting, the electric fields will ‘fringe’ at the edges, see Figure 1.9(g), and the horizontal components of these so-called ‘fringe fields’ are responsible for the radiation. If the length of the patch is chosen to be half a wavelength in the dielectric sheet, the radiation may be thought to originate from two in-phase slots as depicted in Figure 1.9(h).

1.5 Organization of the Book

Before we dive into the field calculations of different antenna types, we will first introduce, in Chapter 2, the so-called *performance parameters* of antennas. These parameters serve to evaluate antennas, offer a means to compare them to each other and offer the possibility of including antenna performance in a high-level system evaluation. Since a thorough understanding of these parameters is considered to be of paramount importance for the practical antenna engineer, we designate a large portion of this book to this topic.

With the antenna parameters well understood, the derivation of these parameters may be undertaken for a selected group of antennas. Before we do so, however, we think it is wise to first give a short refresher course in vector algebra. This is the topic of Chapter 3. Then, in Chapter 4, we introduce the calculation of far-fields from general current distributions and introduce the concept of reciprocity. In Chapter 5, we will use the concepts developed in Chapter 4 to calculate the fields of an elementary dipole antenna and a half-wave dipole antenna. In Chapter 6 we will analyze the loop antenna. Chapter 7 is devoted to aperture antennas, and the theory developed there will be applied to a horn antenna, a parabolic reflector antenna and a rectangular microstrip patch antenna. Finally, in Chapter 8 we will introduce array antennas.

Most chapters are concluded with a more practically oriented section wherein the acquired knowledge is used for designing a modern (mobile) antenna using CST Microwave Studio[®] (CST MWS), a commercially available specialist tool for the

3D EM simulation of high frequency components [10]. The steps, outcomes and resulting next iterations are explained so that having access to this software suite is not absolutely necessary.

1.6 Problems

- 1.1 What is an antenna?
- 1.2 What is the source of electromagnetic radiation?
- 1.3 Why do point source electromagnetic radiators not exist in reality?

References

1. H.J. Visser, *Array and Phased Array Antenna Basics*, John Wiley & Sons, Chichester, UK, 2005.
2. G. Masini, *Marconi*, Marsilio Publishers, New York, 1995.
3. G. Marconi, 'Wireless Telegraphic Communications', *Nobel Lectures in Physics*, 1901–21, Elsevier, 1967.
4. G. Pelosi, S. Selleri and B. Valotti, 'Antennae', *IEEE Antennas and Propagation Magazine*, Vol. 42, No. 1, February 2000, pp. 61–63.
5. J. Hamilton (ed.), *They Made Our World, Five Centuries of Great Scientists and Inventors*, Broadside Books, London, 1990, pp. 125–132.
6. J.C. Rautio, 'Planar Electromagnetic Analysis', *IEEE Microwave Magazine*, March 2003, pp. 35–41.
7. R. Horak, *Webster's New World Telecom Dictionary*, John Wiley & Sons, New York, 2007.
8. IEEE Standards Board, 'IEEE Standard Letter Designations for Radar-Frequency Bands', *IEEE Standard IEEE Std 521-1984*, March 1984.
9. C.A. Balanis, *Antenna Theory, Analysis and Design, third edition*, John Wiley & Sons, New York, 2005.
10. CST Computer Simulation Technology: <http://www.cst.com>

