CAMBRIDGE

Cambridge University Press
978-1-107-17556-3 — Analysis and Modeling of Radio Wave Propagation
Christopher John Coleman
Excerpt
More Information

1 Basic Concepts

This chapter introduces the fundamental ideas of radio waves and radio systems. It is designed to give a brief introduction to radio technology for those without a background in this area. The chapter includes an introduction to a variety of propagation phenomena as a motivation for the more detailed analysis in later chapters.

1.1 Waves

The concept of a wave is something for which it is very difficult to find a clear definition in the literature. Before proceeding, however, it is important that we have a good understanding of what we mean by a wave. In this regard, it is instructive to start with the surface water wave, a phenomenon that gives us one of the best practical illustrations of wave phenomena in general. Water waves are something that most of us have experienced and that exhibit many of the important features of waves and their propagation. As children, we have nearly all generated waves by throwing stones into a pond. Before the stone lands, the surface of the pond (the propagation medium) is calm. After impact, however, there is a ripple that travels radially outward from the point of impact. The ripple forms a circular band of disturbance that expands at a finite speed. Within the band the ripple maintains its shape but with amplitude that reduces as the radius of the band increases. As the ripple travels outward, it might encounter a floating object and then cause it to bob up and down. This motion can be used to extract energy from the wave, energy that was originally supplied by the stone's impact (the wave source). Further, the vertical motion of the object provides a means of detecting the passage of a wave.

Water waves illustrate several important features that are common to all wave phenomena. First, the wave can transport energy from one point (the source) to another (the detector), the energy being transported at a finite speed. Second, after the passage of the wave, the medium returns to its undisturbed state. This last point leads on to another important property of wave phenomena, the ability to make arbitrarily shaped waves. Instead of causing the wave by casting a single stone, we could simply drive the water up and down in an arbitrary fashion (by a sequence of impacts of varying force). The ripples would now constitute a record of the driving sequence and these would then be reproduced some distance away in a detector (our floating object bobbing up and down in sympathy with the wave as it passes). In this way, we could transfer information

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-17556-3 — Analysis and Modeling of Radio Wave Propagation
Christopher John Coleman
Excerpt
More Information

between the source and the detector by use of a suitable code. Not only can information be transmitted in such a fashion by water waves but, just as importantly, by sound and radio waves. For the generation of arbitrarily shaped waves, the important property is that disturbances of the propagation medium should allow for discontinuous behaviour across a surface in space and time, the wavefront. In the case of a uniform medium, such a surface is of the form $R - ct =$ constant where R is the distance from the source, c is the speed of wave propagation and t is time. The partial differential equations that admit solutions with the requisite property are known as hyperbolic equations, and wave phenomena satisfy equations of this form.

1.2 Electromagnetic Waves

Electromagnetic fields satisfy hyperbolic partial differential equations and therefore exhibit wave phenomena. Electromagnetic waves are generated when charge (the source of electromagnetic fields) accelerates. Static isolated charges cause electric fields that decay at least as fast as $1/R^2$ where R is the distance from the charge. When the charges accelerate, however, they cause a field that only decays as $1/R$. This is a wavelike field that can carry energy and information over vast distances. For example, consider a rearrangement of a system of charge that takes place over a short period of time. Before and after the rearrangement, the charge is static and its field falls off as $1/R^2$. During the rearrangement, however, the field will only fall off as $1/R$. Like the ripples in the pond, the effect of the rearrangement (the $1/R$ field) will travel outward as a ripple in the electromagnetic field (see Figure 1.1). The speed of propagation for this ripple is the speed of light ($c = 3 \times 10^8$ m/s). (Indeed, light is an example of electromagnetic waves.) Such a ripple could be detected through the motions that it induces in a second system of charge.

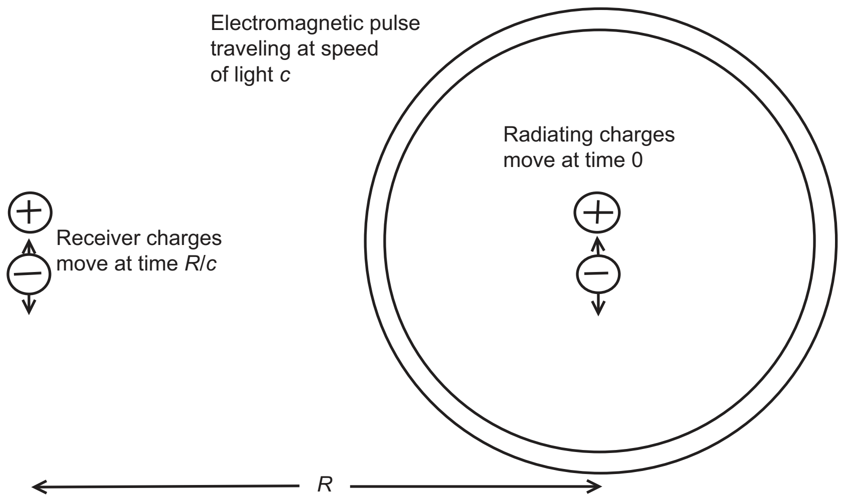

Figure 1.1 Propagation of an electromagnetic pulse.

© in this web service Cambridge University Press www.cambridge.org

Cambridge University Press
978-1-107-17556-3 — Analysis and Modeling of Radio Wave Propagation
Christopher John Coleman
Excerpt
More Information

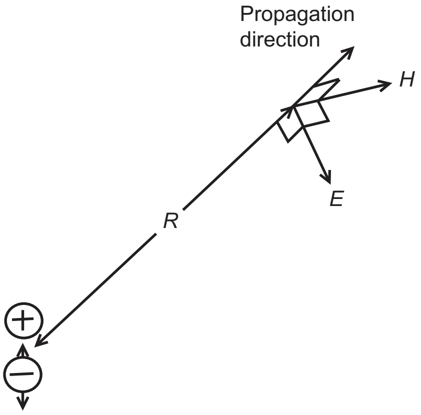

Figure 1.2 The electromagnetic fields caused by accelerating charge.

Unlike the case of static fields, there is always a magnetic wave field H associated with the electric wave field E. The associated magnetic field is proportional to the electric field

$$H = \frac{E}{\eta_o} \tag{1.1}$$

where η_o is the impedance of the free space (377 Ω).The magnetic field is perpendicular to the electric field and also to the direction of wave propagation (see Figure 1.2). Further, the electric wave is also perpendicular to the direction of propagation. Ignoring the static field (this falls off much faster than the wave field), the electric field behaves as

$$E = \frac{K(t - R/c)}{4\pi R} \tag{1.2}$$

where $K(t)$ is a function that depends on the motion of the source charge during rearrangement. Consequently, $K(t)$ will only be non-zero over the period of charge rearrangement, i.e the effect of this rearrangement will be a pulse of electromagnetic field that travels radially outward. Importantly, by suitably controlling the acceleration of the source charge, we can create an arbitrary $K(t)$ and hence transfer information to a distant observer due to the slow fall off in the field of the accelerating charge. This is the basis of radio communication.

If the charges in a system oscillate, they will nearly always be accelerating and so there will nearly always be a wave field. For an oscillation frequency ω (radians per second or $f = \omega/2\pi$ in terms of hertz)

$$K(t) = A\cos(\omega t + \phi) \tag{1.3}$$

where A determines the amplitude of the field and ϕ its phase. The waves produced by such a system will contain virtually no information and so the system needs *modulation* if it is to be used for transferring information. Modulation is achieved by arbitrary variations of A, ϕ and ω. Variations in A are known as amplitude modulation (AM), ω as frequency modulation (FM) and ϕ as phase modulation (PM). Many other forms

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-107-17556-3 — Analysis and Modeling of Radio Wave Propagation
Christopher John Coleman
Excerpt
More Information

of modulation can be created by suitable combinations of these basic forms of modulation. Modulated sinusoidal signals are the basis of most radio communications, but modulation causes the signal to spread in frequency around that of the unmodulated sinusoidal *carrier* signal. It is possible, however, for many communications systems to coexist by operating at different frequencies with the variations in modulation limited in order to prevent the signals overlapping. Detection systems are designed to select a particular band of frequencies (the extent of the band being known as the *bandwidth* of the system) when the signal is *modulated*.

In a practical communications system, radio waves will be produced by an electronic source that drives varying current into a structure (usually metal) known as an antenna. Variations in current constitute the accelerating charge that will cause radio waves. A simple antenna, known as a *dipole*, consists of a rod that is driven at its center. Waves travel radially outward from the antenna, and the power density in the waves ($E^2/2\eta_0$ where η_0 is the impedance of free space) will fall away as $1/R^2$. Furthermore, the power will not be uniform in all directions.

The effectiveness of the antenna in a particular direction can be described by its *directivity* in that direction:

$$\text{Directivity} = \frac{\text{power radiated in a particular direction}}{\text{average of power radiated in all directions}} \tag{1.4}$$

Antennas, however, lose some of their energy as heat (in their structure and its surroundings) and so an important quantity is the *antenna efficiency*

$$\text{efficiency } \eta = \frac{\text{total power radiated}}{\text{total power supplied}} \tag{1.5}$$

As a consequence, a more realistic measure of effectiveness is the *gain*

$$\text{gain} = \eta \times \text{ directivity} \tag{1.6}$$

No antenna exists that has a uniform gain in all directions and so we quite often represent this variation as a gain pattern. This is a 3D surface whose distance from the origin is the gain of the antenna in that direction. For the dipole, the pattern will be rotationally symmetric about the axis along the dipole's length and so it is sufficient to represent the pattern as a 2D slice through the axis. Figure 1.3 shows a dipole together with its gain pattern and from which it will be noted that there is no gain along the dipole axis.

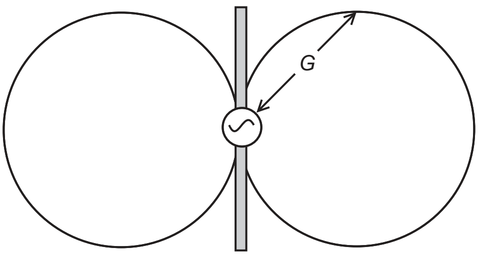

Figure 1.3 A dipole antenna and its gain pattern.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-17556-3 — Analysis and Modeling of Radio Wave Propagation
Christopher John Coleman
Excerpt
More Information

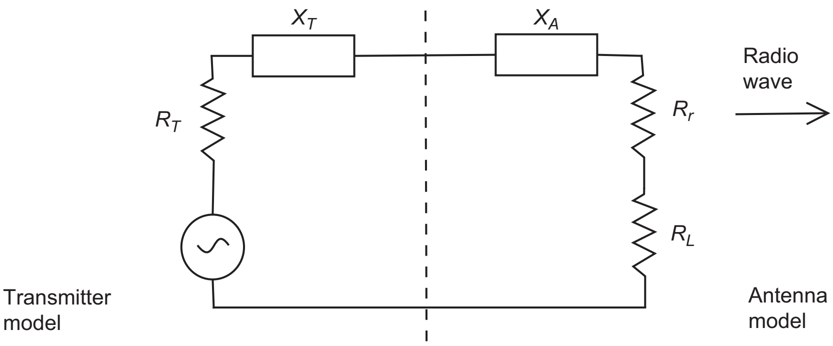

Figure 1.4 Circuit model of a transmitter system.

Radiation is maximum in directions perpendicular to the axis and, for a dipole with an efficiency of 1, this gain will be about 3/2.

An antenna will act as a load to the electronics that drives it; Figure 1.4 shows a circuit model of the transmit antenna and its driver. A dipole is resonant (no reactance X_A in its impedance) when its length is approximately $0.48\,\lambda$ (λ is the wavelength c/f at the operating frequency f). For this reason, a resonant dipole is often known as a half wave dipole. At resonance, the radiation resistance R_r is about 73 Ω and the loss resistance R_L is usually negligible. Dipoles that are much shorter than a wavelength, however, have a much smaller radiation resistance and a large capacitive reactance. For such dipoles, R_L and R_r can often be comparable and this can make them very inefficient.

In general, the amplitude of the electric field of an antenna that is sinusoidally excited has the form

$$E = \frac{\omega\mu_0 I h_{eff}}{4\pi R} \tag{1.7}$$

where I is the current in the feed, μ_0 is the permeability of free space and h_{eff} is the *effective antenna length*. In the case of a resonant dipole, $h_{eff} \approx 0.64 l \sin\theta$ where l is the dipole length and θ is the angle that the propagation direction makes with the dipole axis.

An antenna can be used to extract energy from an electric field since an incident electromagnetic wave will set charges in motion and hence cause current to flow on the antenna structure. In this receive mode, an electric field E will induce an open circuit voltage $h_{eff}E$ in the antenna terminals. An antenna will act as a source to the electronics that acts as the radio receiver; Figure 1.5 shows a circuit model of the receive antenna and its receiver load. It will be noted that the antenna exhibits the same impedance $R_r + R_L + jX_A$ in both receive and transmit functions.

1.3 Communications Systems

Radio waves are used to communicate information over both large and small distances without the use of wires, hence the term *wireless*. The effectiveness of a communication

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-107-17556-3 — Analysis and Modeling of Radio Wave Propagation
Christopher John Coleman
Excerpt
More Information

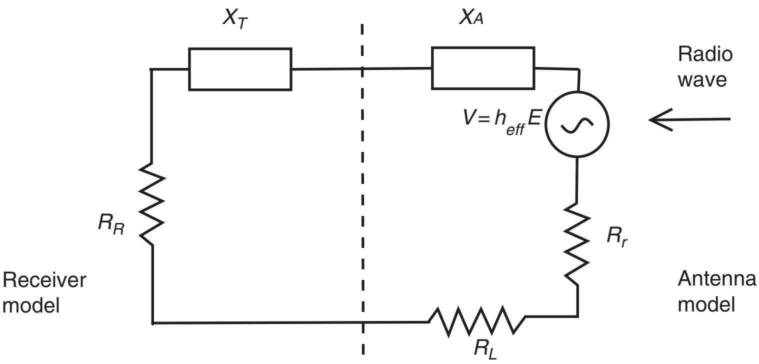

Figure 1.5 Circuit model of a receiver system.

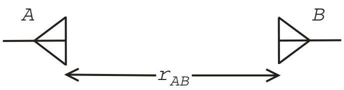

Figure 1.6 Communication system with two stations.

system is often calculated using the *Friis* equation. If the transmitter and a receiver are distance r_{AB} apart, the received power P_B is related to the transmitted power P_A by

$$P_B = P_A G_A G_B \left(\frac{\lambda}{4\pi r_{AB}} \right)^2 \tag{1.8}$$

where G_A and G_B are the gains of the receive and transmit antennas, respectively. The last term represents the decay in the power of the wave as it spreads out from the source, and this *spreading loss* is defined to be $L_{sprd} = (4\pi r_{AB}/\lambda)^2$ (i.e. $P_B = P_A G_A G_B / L_{sprd}$). It should be noted that this loss is often quoted in terms of decibels, i.e $10 \log_{10} \left(L_{sprd} \right)$.

In a radio system, the intentional radio signals must compete with *noise* (unwanted interfering signals) that is generated within the electronics of the system and within the propagation environment. Even a simple resistor will create a noise voltage v_n due to the random thermal motion of its electrons. Such noise has rms voltage

$$\overline{v_n^2} = 4kTBR \tag{1.9}$$

where T (in Kelvin) is the absolute temperature, B (in hertz) is the bandwidth of the radio channel, R (in ohms) is the resistance and k is the Boltzmann constant (1.38×10^{-23} joules per Kelvin). Semiconductors are the source of many different sorts of noise, and so radio receivers will have a variety of contributions to their *internal noise*. In a real radio receiving system, *external noise* is just as important as the internal variety. This can arise from manmade sources (ignition interference, for example) and natural sources (lightning, for example). Consequently, the input signal will normally need to be at a level above that of the combined internal and external noise. In the case of an antenna, the external noise can be considered as that due to resistance $R_r + R_L$, but at a temperature T_A (the *antenna temperature*) that is possibly different from the ambient

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-107-17556-3 — Analysis and Modeling of Radio Wave Propagation
Christopher John Coleman
Excerpt
More Information

temperature (around 290 K). External noise is the ultimate constraint, and, for best performance, a radio receiver should be *externally noise limited* (i.e. the internal noise is below the level of external noise).

The crucial quantity in calculating radio system performance is the signal to noise ratio (SNR) that is required for detection. SNR is defined by

$$\mathrm{SNR} = \frac{\mathrm{S}}{\mathrm{N}} = \frac{\text{signal power}}{\text{noise power}} \tag{1.10}$$

It is necessary to make sure that the strength of the received is sufficient to make the SNR greater than that required for detection. Obviously, this will require the correct choice of transmit power and/or antenna gains.

1.4 Cellular Radio

Although we can accommodate many users by dividing the radio spectrum into many channels (slices of spectrum with limited bandwidth), this still fails to satisfy the modern demands for high volume personnel communications (including video and internet). A solution has been found by limiting the range coverage of a channel so its frequency can be reused at some other location. The full communication region is divided into small cells within which the transmit power is limited so that the users within any one cell can only communicate with the radio base station (RBS) for that cell. In this manner the same set of frequencies can be used in many different cells, provided that they are sufficiently isolated. The RBSs within the network are all connected to each other, and the fixed network, through a mobile switching center (MSC). When a user passes from one cell to another, control is passed to the RBS associated with the new cell and the frequency will appropriately change. Such a system is known as *cellular radio*.

A cellular radio system is designed around a *cluster* of cells, each cell in the cluster having a unique frequency set. A typical cluster consists of seven cells; Figure 1.7 shows a cellular system based on such a cluster (frequency sets are labeled A to G). It can be seen that the cluster topology allows reuse of the frequency sets within separate clusters. The reuse, however, has the potential to cause intercellular interference

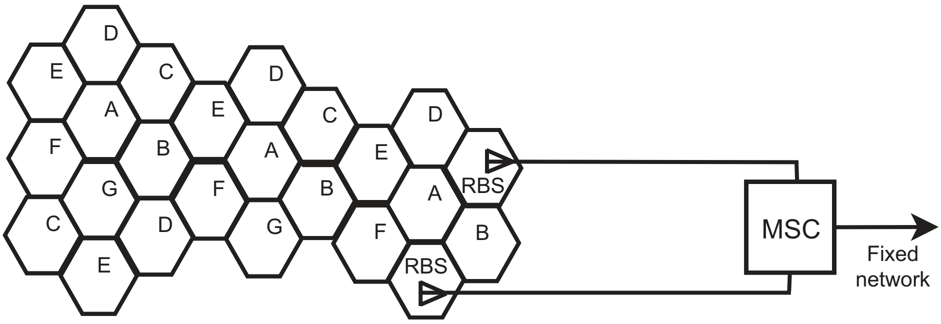

Figure 1.7 A cellular radio system.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-17556-3 — Analysis and Modeling of Radio Wave Propagation
Christopher John Coleman
Excerpt
More Information

(the dominant source of unwanted interfering signals in cellular systems). From Figure 1.7, it will be noted that the minimum distance between cells with the same frequency set is approximately $4.583R$ where R is the cell radius. On the assumption that all transmitters use the same power level, and that the power decays as $(1/distance)^n$, the *signal to interference ratio* (SIR) is given by

$$\text{SIR} = \frac{\text{minimum power within a cell}}{6 \times \text{maximum power between cells}} = \frac{R^{-n}}{6(4.583R)^{-n}} = \frac{4.583^n}{6} \tag{1.11}$$

From the Friis equation, it will be noted that a system in free space will have a value of 2 for n. There are, however, many factors that can modify free space propagation, and it is found, in practice, that a value of n around 4 is more appropriate. For the configuration of Figure 1.7 this would suggest an SIR of greater than 10 dB.

1.5 Radar Systems

One of the major non-communications applications of radio waves is *radar* (radio detection and ranging). In classical radar, the transmitted signal is interrupted by a *target* from which a small amount of energy is re-radiated back to a receiver. The receiver will normally ascertain the direction of the target using a steerable beam (mechanical or electronic steering) and the time of flight of the signal will then provide the target range. For radar systems, the power returned P_R is related to that transmitted P_T by the *radar equation*,

$$P_R = P_T G_R G_T \left(\frac{\lambda}{4\pi R_T}\right)^2 \left(\frac{\lambda}{4\pi R_R}\right)^2 \frac{4\pi\sigma}{\lambda^2} \tag{1.12}$$

where R_T and R_R are the ranges of the target from the transmitter and receiver, respectively, and G_T and G_R are the gains of the transmit and receive antennas, respectively (see Figure 1.8). σ is the *radar cross section* of the target and represents the amount of power reflected when a field with unit power per unit area is incident. The radar equation can be regarded as the double application of the Friis equation with the target acting as both receiver and transmitter. The term $4\pi\sigma/\lambda^2$ is effectively the product of the target receive and transmit gains. Radar cross sections can be quite complex, often

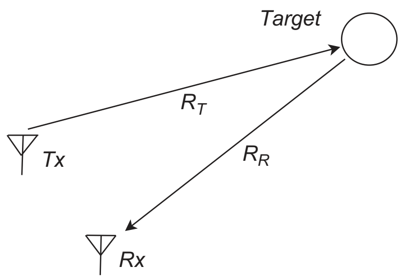

Figure 1.8 A general radar configuration.

CAMBRIDGE
Cambridge University Press
978-1-107-17556-3 — Analysis and Modeling of Radio Wave Propagation
Christopher John Coleman
Excerpt
More Information

depending on both the direction of the illuminator and the direction of the receiver. For a metallic sphere the back-scatter cross section is approximately direction independent with a value πa^2 when the radius a is greater than a wavelength.

One of the major problems with a radar system is that surfaces such as a rough sea and ground can also return a considerable amount of power, known as *clutter*, which can mask a radar target. This interference will be in addition to the noise we have discussed for communications systems. Consequently, the signal to clutter ratio (SCR) can be just as important as SNR in determining radar performance. Fortunately, the motion of the target will cause a shift in frequency for the return signal from the target and so the target can be discriminated from the clutter in the frequency domain. The frequency shift is known as the *Doppler shift* and is related to the target dynamics through

$$\Delta f = -\frac{f}{c}\left(\frac{dR_T}{dt} + \frac{dR_R}{dt}\right) \tag{1.13}$$

Radar systems can effectively be regarded as radio systems in which the environment modulates the signal, hence allowing a radar operator to glean information about the environment. The most obvious application of radar is in the detection of ships and aircraft. However, we increasingly use radar to monitor things such as weather (wind profiling radar, for example) and underground conditions (ground penetrating radar, for example).

1.6 Complex Propagation

The Friis equation assumes there to be no interaction of the radio waves with their environment. At a minimum, however, the ground will reflect radio waves and these reflected waves provide additional paths for the waves between the antennas of a communication system (see Figure 1.8). The waves on the direct and reflected paths will travel different distances and so there will be the potential for interference at the receiver. If the antennas are at heights H_A and H_B above the reflection point, the Friis equation will require modification to

$$P_B = P_A G_A G_B \left(\frac{\lambda}{4\pi r_{AB}}\right)^2 \left|1 + R\exp\left(-2j\frac{\beta H_A H_B}{r_{AB}}\right)\right|^2 \tag{1.14}$$

where R is the reflection coefficient of the ground (approximately -1 at frequencies above a few hundred MHz). From this expression it can be seen that there can be constructive, or destructive, interference between direct and reflected signals that will

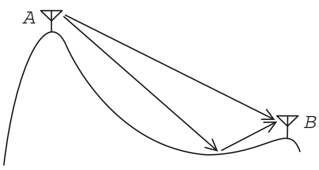

Figure 1.9 Propagation with ground reflections.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-17556-3 — Analysis and Modeling of Radio Wave Propagation
Christopher John Coleman
Excerpt
More Information

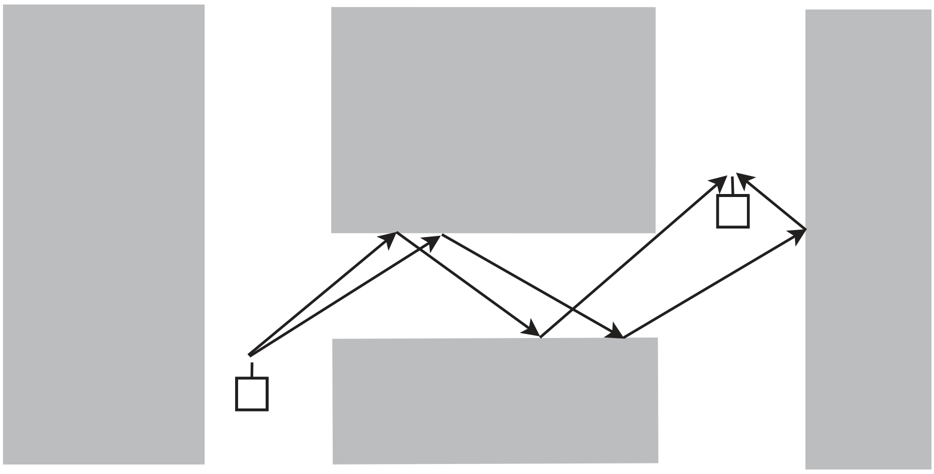

Figure 1.10 Propagation through an urban environment.

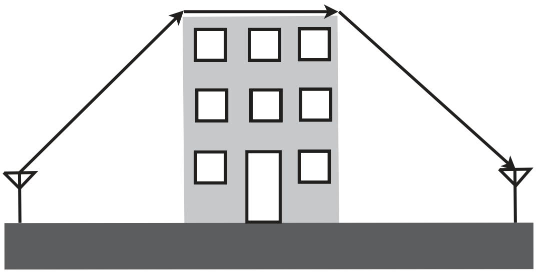

Figure 1.11 Diffractive propagation over a building.

vary with distance between transmitter and receiver. The effect is often pronounced in mobile communications where it can cause the signal to experience *flutter* as the vehicle moves. This can get worse in urban environments where the paths can be more tortuous and numerous (see Figure 1.10). The reflected signals, however, can have a positive effect in that they can allow communication into a *shadow* region (a region where no line-of-sight path is available).

Even without reflections, non-line-of-sight propagation is still possible through a mechanism known as *diffraction*. Consider the propagation of radio waves over a building, as illustrated in Figure 1.11. A small amount of energy will be diffracted into the *shadow* region at the top of the building and then a small amount of this power will be diffracted into the shadow region over the other side of the building. Such power can sometimes be sufficient for communication purposes.

Diffraction, and many other propagation phenomena, can be explained through *Huygens' principle*:

Each point on the wavefront of a general wave can be considered as the source of a secondary spherical wave. A subsequent wavefront of the general wave can then be constructed as the envelope of secondary wavefronts.

© in this web service Cambridge University Press www.cambridge.org