25

Some Practical Aspects of Electromagnetic Waves

25.1 Introduction

Applications of electromagnetic waves are numerous, ranging from cooking food to controlling a faraway spacecraft and receiving information from it. Electromagnetic waves cover a very broad frequency (wavelength) spectrum, and it is impossible in this text to even attempt to cover applications in all regions of the spectrum. Therefore, we confine ourselves mainly to waves used in the versatile area of communications, as the readers of this text are likely to spend a part of their professional lives dealing with this subject. The word “communications” refers broadly to sending and receiving a signal that contains some useful information. This signal might be sent along a coaxial cable or through a waveguide, or radiated or received by an antenna, or propagated along an optical fiber, for example. We will describe some issues related to these different ways of communicating, but we will not deal with the information itself. At the end of the chapter, some other common applications of electromagnetic waves, such as cooking, are described briefly.
25.2 Power Attenuation of Electromagnetic Waves

When one sends or receives information using electromagnetic (radio) waves, the maximum distance at which this can be done is limited by the amount of power available at the sending end, and the loss of the wave energy by the time it gets to the receiving end, assuming a certain receiver sensitivity. The path loss varies with the medium through which the wave is propagating, as well as the frequency (wavelength) of the wave. So let us calculate the loss per unit distance for a few different cases, and then do a performance comparison. In the following examples, we calculate the loss in a coaxial cable, a rectangular waveguide, an optical fiber, and a line-of-sight radio link (Fig. 25.1), referring to knowledge we gained in previous chapters. In each case, we consider a link where the power at the transmitting (sending) end is $P_T$ and the received power $P(r)$ is a function of the distance $r$ between the transmitter and receiver. So, at a point $r$ away from the transmitter, the received power is $P(r) = P_T f(r) < P_T$. What is this function $f(r)$ in different cases?

Example 25.1—Attenuation of an electromagnetic wave in a coaxial cable. In a coaxial cable, if the losses are not large, the power along the line as a function of distance $r$ from the generator can be expressed as

$$P(r) = P_T e^{-2a r}$$  \hspace{1cm} (25.1)

![Figure 25.1](image)

(a) A coaxial cable, (b) a rectangular waveguide, (c) an optical fiber, and (d) antennas in a line-of-sight radio link are used in communications as ways of transferring information contained in an electromagnetic wave. The coaxial cable was discussed in detail in Chapter 18, the rectangular waveguide in Chapter 23, and the line-of-sight link in Chapter 24.
(see section 18.4, and note that power is the product of the voltage and current, giving the factor of 2 in the exponent). Therefore,

\[
-\frac{dP(r)}{dr} = 2\alpha P(re^{-2\alpha r}) = 2\alpha P(r) = \frac{dP_{\text{losses}}(r)}{dr}.
\] (25.2)

The attenuation coefficient \(\alpha\) comes from losses in the conductor, described by the resistance per unit length, \(R'\), and losses in the dielectric, described by the conductance per unit length, \(G'\), as described in Fig. 18.10. Usually the conductive losses are dominant, and

\[
\frac{dP_{\text{losses}}(r)}{dr} = R'|I(r)|^2.
\]

The power transmitted to a point \(r\) away from the line's beginning can be expressed in terms of the current \(I(r)\) and the characteristic impedance (assuming no losses) as \(P(r) = Z_0|I(r)|^2\), where \(Z_0 \simeq \sqrt{L'/C'}\). So

\[
\alpha = \frac{R'}{2Z_0} = \frac{R'}{2\sqrt{\frac{C'}{L'}}}.
\] (25.3)

As one example, let us calculate the loss at 1 MHz in a coaxial cable made of copper, filled with a dielectric of permittivity \(\varepsilon_r = 3\), and of dimensions such that the inner conductor radius \(a = 0.45\) mm and inner radius of the outer conductor \(b = 1.47\) mm. The skin depth is \(\delta = 0.067\) mm (Example 20.1), which means that the current is not distributed through the entire cross section. We obtain in this case that \(R' \simeq \rho/(2\pi a\delta) = 0.093\) \(\Omega/m\), and \(Z_0 = 50\) \(\Omega\). This gives \(\alpha = 0.00093\) Np/m = 0.016 dB/m, so that \(f(r) = e^{-0.0019r}\), where \(r\) is in meters. This just corresponds to the loss in the inner conductor. In the outer conductor, the losses are lower (see problem P25.5).

What happens at higher frequencies with the loss in coaxial cables? As another example, let us look at losses at 0.1, 1, and 10 GHz in a high-frequency 50-\(\Omega\) RG-58 cable, made of copper with polyethylene dielectric, with \(a = 0.45\) mm and \(b = 1.47\) mm. The loss can usually be found in manufacturer's data sheets, and for this cable at 0.1 GHz, the loss is about 0.2 dB/m, and at 1 and 10 GHz it increases to 0.66 and 2.6 dB/m, respectively. This means that at 10 GHz almost half of the power (which would be 3 dB) is lost after only 1 m of propagation. In this case, the function \(f(r) = e^{-0.6r}\), where \(r\) is in meters. Why is the loss this high?

**Example 25.2—History: the transatlantic telegraphy cable. Reduction of losses along lines by increasing inductance per unit length.** When the first transatlantic cable was laid in the Atlantic Ocean, the engineers did not understand that the loss over several thousand kilometers would make the cable impractical (see problem P25.1—calculate the loss for \(r = 5000\) km at 10 kHz for practice). The famous British physicist Oliver Heaviside had warned the engineers about losses, but they did not listen. Later, Mihailo Pupin, a Columbia University professor, noticed that in practice, the first term of \(\alpha\) in parentheses in Eq. (18.36), repeated here for convenience,

\[
\alpha \simeq \frac{1}{2}\left(R'\sqrt{\frac{C'}{L'}} + G'\sqrt{\frac{L'}{C'}}\right).
\] (18.36)

is much greater (several orders of magnitude) than the second, due to the relatively large value of \(R'\). He next realized that it is possible to reduce this term considerably by increasing \(L'\),
without making the second term prohibitively large. He then proposed to reduce $\alpha$ by placing series inductive coils along the cable at regular distances. These are today called Pupin coils, and they enabled transmission of signals along transmission lines of great lengths, including the transatlantic cable.

Let us consider a typical coaxial line and estimate the first and second terms in parentheses of Eq. (18.36). Let the line dielectric have a relative permittivity $\epsilon_r = 3$, permeability $\mu_0$, and conductivity $10^{-12} \text{S/m}$. Let the line conductors be copper, of conductivity $56 \times 10^6 \text{S/m}$, and let the ratio of conductor radii (see Table 18.1) be $b/a = e = 2.71828$.

Using the expressions for $C'$, $G'$, $L'$, and $R'$ in Table 18.1, we find that $C' \simeq 167 \text{pF/m}$, $G' \simeq 0.3 \text{pS/m}$, $L' \simeq 0.2 \mu\text{H/m}$, and $R' \simeq 0.0055 \Omega/\text{m}$. With these values, the first term in the expression for $\alpha$ (including $1/2$) is about $0.8 \times 10^{-4} \text{Np/m}$, and the second term is about $5 \times 10^{-12} \text{Np/m}$. The first term, due to imperfect cable conductors, is indeed much greater than the second, due to imperfect cable dielectric.

By increasing $L'$ artificially, as Pupin did by connecting lumped series coils in the cable, the first term can be substantially reduced, and therefore $\alpha$ made smaller. Note that the attenuation of the wave is proportional to $e^{-\alpha r}$, so a 10-fold increase in inductance per unit length results in about a 24-fold decrease in signal attenuation. This means that if a cable with no Pupin coils has an attainable range of 100 km, with a 10-fold artificial increase of the line series inductance per unit length the range is increased to about 2400 km.

**Example 25.3—Attenuation of electromagnetic waves in a rectangular waveguide for the dominant mode.** Losses in hollow metal waveguides depend on the mode propagating in the waveguide, the type of metal and dielectric used to make the waveguide, the geometry of the waveguide, and frequency. It can be shown (see, e.g., S. Ramo et al., *Fields and waves in communication electronics*, 3d ed., J. Wiley & Sons, 1993, p. 423) that the attenuation coefficient for the dominant TE$_{10}$ mode in a rectangular waveguide of sides $a$ and $b$ is given by

$$\alpha = R_s \sqrt{\frac{\varepsilon}{\mu} \frac{a/b + 2 f_c^2}{a} \frac{1}{f_c^2}}$$

(TE$_{10}$),

where $R_s$ is the surface resistance of the metal walls, Eq. (20.10). Consequently, the losses depend on the metal conductivity and frequency. For a typical X-band (8.2 to 12.4 GHz) waveguide ($a = 25.4 \text{mm}$, $b = 12.7 \text{mm}$), made of brass plated with silver and a rhodium anticorrosion coating, the conductivity is $6.17 \times 10^7 \text{S/m}$, and the skin depth at 10 GHz is $\delta = 0.64 \mu\text{m}$, yielding $R_s = 2.53 \Omega$ and $\alpha = 0.0883 \text{Np/m} = 0.767 \text{dB/m}$.

Note that at very high frequencies, when the skin depth is on the order of the conductor surface imperfections (the surface cannot be made absolutely flat), the surface resistance becomes substantially larger than its theoretical value for a perfectly flat surface.

It should be noted that some other kinds of metallic waveguides have field profiles that result in lower losses than in the previous case. An example is a TM$_{11}$ mode in a waveguide with circular cross section, with a typical $\alpha = 0.01 \text{dB/m}$. Initially there was a large development effort, mostly in Bell Labs, to use this kind of waveguide for the entire phone network across the United States, but before the network was built, optical fibers were shown to have lower loss and cost, so the waveguide technology was never implemented. In Socorro, New Mexico, however, there is a large radio telescope [Very Large Array (VLA), nicely shown in the 1997 movie *Contact*] in which the signals received from 27 large dish antennas (each 25 m in diameter) formerly were propagated at 44 GHz through a circular waveguide to the operations center that is about 60 km away from the telescope. Recently, the circular waveguide was replaced by optical fibers, but the waveguide is still used as the pipe for the fibers.
Figure 25.2 Sketch of wave propagation along a dielectric slab

Example 25.4—Dielectric waveguides and optical fibers. Optical fibers are used as waveguides for electromagnetic waves in the visible and infrared part of the frequency spectrum, with wavelengths between roughly 300 nm and 10 μm (optical engineers usually think in terms of wavelength, whereas radio engineers think in terms of frequency). Fibers are so-called dielectric waveguides. The simplest dielectric waveguide is a flat dielectric slab. Because the permittivity of the slab is always greater than $\varepsilon_0$, a possibility exists that the wave propagating in the slab is totally reflected at the interface.

If we excite in the slab a plane wave incident on one slab face at an angle greater than the critical angle, it will be reflected totally at the same angle toward the other face, then reflected totally from that other face, etc. So the wave will bounce between the two slab faces, and the slab will serve as a guiding medium of the wave, as sketched in Fig. 25.2.

The principle of optical fibers is essentially the same, although the wave types propagating along them are more complicated. Thanks to total reflection, however, these waves also are restricted to the domain of the fiber. The fiber is made of an inhomogeneous dielectric (quartz), and roughly speaking it has a core and an outer cladding layer, sketched in Fig. 25.3. The permittivity of the core is typically a fraction of a percent higher than that of the outer part (it is germanium doped). The cladding has a permittivity of about $\varepsilon_r = 2.1$ (an optical index of $n = \sqrt{\varepsilon_r} = 1.46$). The typical attenuation of a so-called single-moded fiber at a wavelength of 1.55 μm is 1 dB/km = 0.001 dB/m (Corning specification sheets, 1998), and for very specialized low-loss fibers it can be as low as 0.1 dB/km.

Example 25.5—Attenuation of electromagnetic waves in a line-of-sight radio link through a vacuum. In a line-of-sight radio link (which means that the radio wave travels between two antennas directly, with no reflections), the power loss function $f(r)$ is given by the Friis transmission formula:

$$P(r) = P_T \frac{G_T A_R}{4\pi r^2} = P_T \frac{A_T A_R}{\lambda^2 r^2},$$

(25.5)

Figure 25.3 Sketch of wave propagation along an optical fiber, based on total internal reflection
where $G_T$ is the gain of the transmitting antenna, and $A_R$ and $A_T$ are effective areas of the receiving and transmitting antennas used in the link. We can measure these effective areas in terms of the operating wavelength $\lambda$ used in the link. If the two antennas are equal, $n\lambda^2$ large, and we assume they are well designed so that the effective areas are roughly equal to their geometric areas, we get

$$\frac{P(r)}{P_T} = f(r) = \frac{n^2\lambda^2}{r^2}. \quad (25.6)$$

This means that the larger the antennas are (measured in wavelengths), the lower the loss of power between the transmitter and receiver. Large antennas, of course, correspond to high directivity (gain). As an example, a standard X-band horn antenna measures about 2 by 2.6 wavelengths at 10 GHz, yielding $f(r) = 0.025/r^2$. If instead of this horn, 3-m round dishes are used, $f(r) = 55,000/r^2$. In the second case, a much larger distance can be spanned with the same receiver sensitivity.

In the preceding examples we have seen that in coaxial cable, waveguides, and optical fiber, the power decays exponentially away from the transmitter, with very different decay constants (on the order of 1 dB/m in a coaxial cable, 0.1 dB/m in a waveguide, and 0.01 dB/m in a fiber). If antennas are used, however, the power drops as $1/r^2$, which is a function that decays more slowly than an exponential for large distances. A plot that shows the attenuation as a function of distance is shown in Fig. 25.4, for the attenuation coefficient values calculated in Examples 25.1 to 25.5. The lower attenuation for long distances is one of the reasons that antennas are used.

*Figure 25.4* The attenuation function $\log f(r)$ for coaxial cable, rectangular waveguide, and a 3-m diameter dish antenna line-of-sight link at 10 GHz. Attenuation in an optical fiber at 1.55-μm wavelength is also shown. Note the logarithmic scale.
for communications. Another reason is that in many cases, for example in aircraft guidance, satellite communications, portable phones, and pagers, it would not be practical to use cables.

**Example 25.6—Curvature of the earth and effective earth radius in line-of-sight links.** AM broadcasting systems rely on surface wave transmission between two points on the earth’s surface. Shortwave radio systems bounce waves off the ionosphere and the earth’s surface. The UHF and VHF radio waves used for communications by airplanes, as well as microwaves in radio relay links, propagate along a direct path. As mentioned, this is called line-of-sight propagation, illustrated in Fig. 25.5. The figure shows how the range is limited by the curvature of the earth. That is why almost all radio relay stations are put on high peaks, even though the weather conditions at these places often complicate the design (and very few people want to work there). From the figure,

\[(R + h)^2 = R^2 + r^2,\]  

(25.7)

where \(R\) is the radius of the earth, \(h\) is the height of the antenna above ground, and \(r\) is the range of the communication link. If we solve for \(r\),

\[r = \sqrt{h^2 + 2Rh} \approx \sqrt{2Rh}\]  

(25.8)

since \(R \gg h\).

Equation (25.8) predicts shorter ranges than the ones achievable in reality. The reason is the change in the refractive index of the atmosphere, so that the waves really follow a curved path, not a straight one. This is sketched in Fig. 25.6. The waves bend toward the denser (lower) layers, and this gives longer ranges. This effect varies quantitatively depending on the location on the earth and the hour of the day. A reasonable approximation is to use an effective radius for the earth, \(R_{\text{eff}}\), somewhat larger than the real radius. It turns out that if this radius is taken to be about 4/3 of the actual radius, or about 8500 km, the wave refraction is fairly accurately taken into account. If both the receiving and transmitting antennas are above ground, the line-of-sight approximate range formula becomes

\[r = \sqrt{2R_{\text{eff}}h_1} + \sqrt{2R_{\text{eff}}h_2}.\]  

(25.9)

**Figure 25.5** Line-of-sight path limit on the curved earth
Figure 25.6 The variation of the refractive index of the atmosphere makes the paths of the waves longer.

When deriving the antenna link path loss, we assumed no extra attenuation due to the atmosphere. This is a good assumption in clear weather close to the earth’s surface, but only in a certain range of frequencies. Rain and snow degrade the path loss significantly, but even clear atmosphere has a frequency-dependent attenuation curve that has strong peaks due to specific properties of oxygen, the hydroxide (OH) radical, water vapor, and other constituents of the atmosphere. This dependence dictates the frequencies used for specific purposes, as described later in this chapter. In satellite communications, the waves pass through the ionosphere, a layer of the atmosphere that has unique properties and that also significantly affects wave propagation. We will be able to analyze the effects of the ionosphere in more detail after we develop an understanding of plane wave propagation through ionized gases, described in the next section.

Questions and problems: Q25.1 to Q25.5, P25.1 to P25.8

25.3 Effects of the Ionosphere on Wave Propagation

The upper layer of the atmosphere, between about 50 and 500 km above the earth’s surface, is a highly rarefied ionized gas. This ionized layer of the atmosphere is known as the ionosphere. It has a pronounced influence on the propagation of electromagnetic waves in a wide frequency range. It is therefore important to understand this influence when dealing with any kind of radio communications in which the waves travel through the ionosphere. The influence of the ionosphere on radio-wave propagation is of great interest starting from very low frequencies (10 to 100 kHz), to short waves (up to 30 MHz), but also for higher frequencies than these.

25.3.1 Plane Wave Propagation Through the Ionosphere (an Ionized Gas)

This section is aimed at presenting the basic theory of propagation of uniform plane electromagnetic waves in ionized gases. To simplify the analysis, the collisions of moving charged particles with neutral gas molecules, resulting in wave attenuation, will be ignored.
Consider a uniform plane electromagnetic wave of angular frequency $\omega$ propagating in an ionized gas. Let there be $N$ ions per unit volume of charge $Q$ and mass $m$. Assume that at a fixed point inside the gas the electric field strength of the wave varies in time as $E(t) = E_m \cos \omega t$. The equation of motion of a single ion under the influence of the electric and magnetic field of the wave has the form

$$m \frac{dv}{dt} = QE_m \cos \omega t + Qv \times (\mu_0 H_m) \cos \omega t. \quad (25.10)$$

We know that for a uniform plane wave $H_m = \sqrt{\epsilon_0 / \mu_0 E_m}$, and that $\sqrt{\epsilon_0 \mu_0} = 1/c_0$. Therefore, the second term on the right side of this equation is approximately proportional to the first term multiplied by the ratio $v/c_0$. Because the velocities that ions can acquire in a time-harmonic electric field are much smaller than the velocity of light in a vacuum, the second term can be ignored. If we multiply the equation thus obtained by $dt$ and then integrate, we obtain

$$v = \frac{Q}{\omega m} E_m \sin \omega t. \quad (25.11)$$

The integration constant, a time-constant velocity, is zero because a time-harmonic electric field cannot produce a steady drift of ions.

Knowing the velocity of ions, we know also the current density that they produce,

$$J = NQv = \frac{NQ^2}{\omega m} E_m \sin \omega t. \quad (25.12)$$

The second Maxwell's equation now becomes

$$\nabla \times H = J + \epsilon_0 \frac{\partial E}{\partial t} = \left( \frac{NQ^2}{\omega m} - \epsilon_0 \omega \right) E_m \sin \omega t, \quad (25.13)$$

or

$$\nabla \times H = -\omega \left( \epsilon_0 - \frac{NQ^2}{\omega^2 m} \right) E_m \sin \omega t. \quad (25.14)$$

For $N = 0$ (a vacuum), the second term in the parentheses does not exist. Therefore, the presence of the ions can be represented by an equivalent reduction in permittivity. Because this reduction is proportional to $Q^2/m$, we conclude that the sign of the ions is unimportant, and that those ions having the largest ratio $Q^2/m$ have the most pronounced influence. Because this ratio is the largest for free electrons, they are the dominant factor for plane wave propagation in an ionized gas.

Let us define the equivalent (or effective) permittivity of an ionized gas by

$$\epsilon' = \epsilon_0 \left( 1 - \frac{NQ^2}{\omega^2 \epsilon_0 m} \right) = \epsilon_0 \left( 1 - \frac{\omega_0^2}{\omega^2} \right), \quad (25.15)$$
where
\[ \omega_c = \sqrt{\frac{NQ^2}{\varepsilon_0 m}} \] (25.16)
is known as the critical angular frequency, and \( f_c = \frac{\omega_c}{2\pi} \) as the critical frequency, of the ionized gas.

The propagation coefficient can now be written as
\[ \beta = \omega \sqrt{\varepsilon' \mu_0} = \frac{\omega}{c_0} \sqrt{1 - \frac{\omega_c^2}{\omega^2}}, \] (25.17)
and the phase velocity of the wave is given by
\[ v_{ph} = \frac{\omega}{\beta} = \frac{c_0}{\sqrt{1 - \omega_c^2/\omega^2}}. \] (25.18)

If \( \omega > \omega_c \), the expression under the square root is positive, so that \( v_{ph} > c_0 \). We know that this is only a geometrical velocity, and that the velocity of propagation of a signal, or of energy, is less than \( c_0 \) (see Example 21.6).

If \( \omega < \omega_c \), however, the expression under the square root is negative and \( \beta \) becomes imaginary, which means that waves of angular frequencies less than \( \omega_c \) cannot propagate in this ionized gas. This is why \( \omega_0 \) is called "the critical angular frequency," and \( f_c " \text{the critical frequency.}"

Example 25.7—Penetration of plane waves through the ionosphere. The critical frequency of the ionosphere varies greatly with the distance from the earth’s surface, as well as with the hour of the day and month of the year, and with the sun’s activity. Roughly, for a plane wave propagating vertically, this frequency ranges from about 3 to 8 MHz. Therefore, no wave of a frequency below about 3 MHz can escape the earth, nor can such a wave reach us from outer space. Therefore, for communications via satellites we must use higher frequencies than these.

It is interesting that at very low frequencies (less than about 100 Hz), the wavelength is much larger than the ionosphere thickness, and such waves can penetrate the ionosphere.

25.3.2 REFLECTION AND REFRACTION OF PLANE WAVES IN THE IONOSPHERE

We shall now see what happens if a plane wave is emitted from the earth’s surface toward the ionosphere at an arbitrary angle. The ionosphere is an ionized layer of the atmosphere, and we have shown earlier that free electrons have the most pronounced influence on wave propagation. The concentration of electrons changes with the height and depends greatly on the hour of the day, and significantly on the season of the year, the latitude, and the activity of the sun.

During the day, the variation of the concentration of electrons with height above the earth’s surface shows certain regularities, resembling four blurred layers. Starting from the surface of the earth, the layers are designated as \( D \), \( E \), \( F_1 \), and \( F_2 \). The corresponding heights are 50 to 70 km (the \( D \) layer), 100 to 150 km (the \( E \) layer),
about 200 km (the $F_1$ layer), and between 250 and 300 km (the $F_2$ layer). During the night, the $D$ and $E$ layers practically disappear, and the layers $F_1$ and $F_2$ merge into a single layer, $F$, between about 250 and 400 km above earth. The critical frequency of layers $E$, $F_1$, $F_2$, and $F$ are about 3 to 4 MHz, 4 to 5 MHz, 6 to 8 MHz, and 3 to 5 MHz, respectively. The lowest layer, $D$, in which collisions of electrons with neutral atoms and molecules are most pronounced, dominates the attenuation of waves propagating through the ionosphere. This is why attenuation of waves reflected from the ionosphere is the largest during the day, when this layer is present.

Assume that the ionosphere critical frequency versus height $h$ above the earth’s surface is as in Fig. 25.7, with a maximum critical frequency $f_{c\text{ max}}$ at a certain height. Assume that an antenna radiates a plane wave of frequency $f$ so that it is incident at an angle $\theta_0$ on the lower boundary of the ionosphere (which we assume to be plane), as in Fig. 25.7. We know that in the case of a homogeneous ionized gas, it can be considered as a medium of equivalent apparent permittivity from Eq. (25.15):

$$
\varepsilon' = \varepsilon_0 \left( 1 - \frac{\omega^2}{\omega_c^2} \right) = \varepsilon_0 \left( 1 - \frac{f_c^2}{f^2} \right).
$$

(25.19)

We can imagine the ionosphere consisting of many thin homogeneous layers of slightly different critical frequencies, as indicated in Fig. 25.7. The incident wave then

![Diagram](image)

Figure 25.7 A model of the ionosphere critical frequency versus the height, $h$, above earth’s surface (diagram on the left), and a sketch of propagation of two waves incident on the ionosphere. The solid line represents a wave of frequency less than the maximum critical frequency. The dashed line represents a wave of frequency greater than the maximum critical frequency.
refracts on the first layer, with no reflected wave because the apparent permittivity of that layer is almost the same as \( \varepsilon_0 \). It is next incident at a new angle, \( \theta_1 \), on the next layer, and so on. By applying Snell’s law, we obtain the following sequence of equations:

\[
\frac{\sin \theta_0}{\sin \theta_1} = \sqrt{\frac{\varepsilon'_1}{\varepsilon_0}}, \quad \frac{\sin \theta_1}{\sin \theta_2} = \sqrt{\frac{\varepsilon'_2}{\varepsilon'_1}}, \quad \frac{\sin \theta_2}{\sin \theta_3} = \sqrt{\frac{\varepsilon'_3}{\varepsilon'_2}}, \quad \ldots \quad (25.20)
\]

where \( \varepsilon'_1, \varepsilon'_2, \ldots \) are effective permittivities of successive layers. Let \( \theta' \) be the incident angle at any desired height \( h \), where the effective permittivity is \( \varepsilon' \). The angle \( \theta' \) can be calculated if we multiply together all the Eqs. (25.20) up to the angle \( \theta' \). It is easily seen that the result of this multiplication is

\[
\frac{\sin \theta_0}{\sin \theta'} = \sqrt{\frac{\varepsilon'}{\varepsilon_0}}. \quad (25.21)
\]

From this equation and Eq. (25.15) we obtain

\[
\sin \theta' \sqrt{1 - \frac{f_c^2}{f^2}} = \sin \theta_0. \quad (25.22)
\]

If the frequency \( f \) of the wave and the initial incident angle \( \theta_0 \) are such that \( \theta' < \pi/2 \) at the height at which the critical frequency \( f_c = f_c \text{ max} \), the wave will be bent, but will pass through the ionosphere, and leave it at exactly the angle \( \theta_0 \) (case 1 in Fig. 25.7).

If, however, \( f \) and \( \theta_0 \) are such that \( \theta' = \pi/2 \) before the wave reaches the layer of maximum critical frequency, the wave is reflected from the ionosphere (case 2 in Fig. 25.7), leaving the ionosphere in the downward direction also at an angle \( \theta_0 \). The wave reaches the height at which the ionization is such that

\[
\sqrt{1 - \frac{f_c^2}{f^2}} = \sin \theta_0, \quad (25.23)
\]

which, after simple manipulations, becomes

\[
f_c = f \cos \theta_0. \quad (25.24)
\]

**Example 25.8—Waves incident normally on the ionosphere.** According to Eq. (25.23), for \( \theta_0 = 0 \) (normal incidence on the ionosphere), \( f_c = f \). So, with a wave incident perpendicularly on the ionosphere, all waves of frequencies less than \( f_c \text{ max} \) will be reflected back. However, all waves of frequencies greater than \( f_c \text{ max} \) will go through the ionosphere. This conclusion can be used for the experimental determination of \( f_c \text{ max} \). One would use a variable-frequency transmitter radiating waves vertically, send wave packages of increasing frequency, and listen to the echo. The frequency at which the echo disappears is the maximum critical frequency of the ionosphere at the time and site of the probing.

**Example 25.9—Waves incident obliquely on the ionosphere.** For \( \theta_0 > 0 \), Eq. (25.24) tells us that all the waves of frequencies less than \( f_c \text{ max} / \cos \theta_0 \) will be reflected back, and those of
frequencies greater than \( f_{c_{\text{max}}} / \cos \theta_0 \) will go through the ionosphere. This means that for larger initial angles of incidence, \( \theta_0 \), higher-frequency waves are reflected from the same ionosphere.

As mentioned, \( f_{c_{\text{max}}} \) is between roughly 3 and 5 MHz, so no wave of a frequency below about 3 MHz can pass through the ionosphere. However, for perpendicular incidence only, waves of frequencies greater than \( f_{c_{\text{max}}} \) pass through it. If the wave is incident at some other angle, frequencies higher than \( f_{c_{\text{max}}} \) will also be reflected. So the ionosphere and the surface of the earth can be used as a kind of duct, or waveguide, along which waves of appropriate frequencies propagate by bouncing back and forth between the ionosphere and earth’s surface. This is used in AM radio broadcasting, and long-range short-wave radio links.

Obviously, for communications with satellites we must utilize frequencies so high that at practically no angle of incidence is the wave reflected from the ionosphere.

Questions and problems: Q25.6 to Q25.8

25.4 Choice of Wave Frequencies and Guiding Medium for Different Applications

At lower frequencies, we have seen that the losses in cables are relatively low, and in applications in which the two ends can be physically connected, coaxial cables are often used. An example is cable television, which is distributed over a 75-\( \Omega \) coax between about 54 and 88 MHz (channels 2 to 6), and about 100 to 700 MHz (channels 7 to 99). Each analog channel uses 6 MHz of bandwidth. At these relatively low frequencies, waveguides cannot be used in practice—they would be too large (see problem P25.7).

In a cable TV distribution system, as in Fig. 25.8, in each neighborhood a head end station receives the broadcast signal. Antenna links are used for broadcasting in the frequency range of the channels (VHF and UHF). After the signals are received by

![Figure 25.8 Sketch of a cable TV distribution system](Image)
one or more antennas, they are first distributed over a trunk cable, and then branched to distribution lines where there is a tap-off for each customer. The cable has loss, so every 40 to 70 m a 20-dB amplifier boosts the signal. As we have seen earlier, the cable will have higher loss at the higher frequencies, so a device called an equalizer is used to equalize the power in all the channels.

TV stations can also be received from satellites, in which case the antennas on both ends need to be directional. The channel frequencies are relatively low, so a directional antenna would be quite large (see problem P25.9) and hard to mount on a satellite. In addition, more than one antenna would probably be needed to cover the entire range. As mentioned earlier, to propagate a signal through the ionosphere, a high enough frequency needs to be used so that at practically no angle of incidence is the wave reflected from the ionosphere. The solution to all these problems is to use a higher frequency for the wave transmitted from the satellite to the head station. This is done in such a way that TV channels are translated in frequency, modulating some much higher frequency, which is then radiated from an antenna. On the receiving end, the frequencies are translated back down to the original range. Several properties determine the satellite frequency: size of the antennas, properties of the atmosphere, and available bandwidth.

We discussed antenna size for a given directivity earlier. In satellites, very narrow beam (high-directivity) antennas are used. The satellite is usually several hundred kilometers above the earth's surface, and a narrow beamwidth (corresponding to a small footprint on the surface) translates to electrically large antennas (see problem P25.10). In order for the antenna to fit on a satellite (which is typically a cylinder several meters in diameter and several meters tall), high frequencies (small wavelengths) have to be used.

In addition to frequencies dictated by the ionosphere, the rest of the earth's atmosphere has a pronounced effect on wave propagation. The measured attenuation as a function of frequency at sea level and at a height of 40 km is shown in Fig. 25.9. It can be seen that there are some regions with clearly lower attenuation up to about 20 GHz, around 30 to 40 GHz, and around 90 GHz. These are called the atmospheric windows. The peaks in the attenuation that define these windows are due to material properties of the different constituents of the atmosphere, as indicated in the figure. Typical frequencies used in satellite communication for TV worldwide are 1.7 to 3 GHz (S band), 3.7 to 4.2 GHz (C band), 10.9 to 11.75 GHz (so-called Ku1 band, although this is really in X band), 11.75 to 12.5 GHz (Ku2 band), 12.5 to 12.75 GHz (Ku3 band), and 18 to 20 GHz (Ka band). Other satellite communications use the regions around 30 GHz and 44 GHz, and some military applications use the 90-GHz region (W band).

In some cases, the high attenuation around 60 GHz is used on purpose. For example, communication between satellites can be done at this frequency with no interference with ground stations. Other examples are wireless local area networks (LANs), which use this frequency because it gives natural cell boundaries due to the high attenuation, as well as some collision avoidance radar systems, which only need to see the nearest obstacles on the road.

The available bandwidth in satellite links determines the amount of information that can be sent. For example, a 6-GHz link with a 5% (300-MHz) bandwidth can
accommodate 50 analog or 25 digital channels. At 30 GHz, a 5% (1500-MHz) bandwidth would accommodate 250 analog or 125 digital channels. For comparison, in an optical fiber at 1.55-μm wavelength (a frequency of about 200 THz), a 5% bandwidth is very large—over 10 GHz. This large available bandwidth is the major reason for using optical fibers. A standard in 1998 for channels over fibers is a bandwidth of 2.5 Gbit/sec with up to 40 channels (a total bandwidth of over 100 GHz). Another advantage is that one does not have to worry about the atmosphere. The specific wavelength is chosen because very low-loss and low-dispersion fibers can be made at this wavelength.

A number of commercial applications exist for cellular telephony, mostly around 900 MHz and 2 GHz. There are several reasons for choosing these frequencies. The lower part of the spectrum is very noisy due to man-made noise. As an example, the level of man-made noise at 100 MHz is 30 dB lower than at 10 MHz and continues to decrease with frequency. On the upper end, the atmospheric loss due to rain and snowfall increases dramatically above about 3 GHz (as an example, the attenuation in heavy rain is about 0.02 dB/m at 3 GHz, and 2 dB/m at 20 GHz, and a typical cell size is 5 to 10 km). An interesting part of man-made noise is so-called emissions noise from cars: the spark that ignites the combustible mixture of gasoline vapor and air is very nonlinear and has very high harmonics with power levels that are quite high around 2 MHz, but about 40 dB lower at 100 MHz.

In a cellular system, the propagation path is often not direct because the waves bounce off buildings, the ground, and other objects. The situation is also complicated by the fact that users are mobile. Often several waves reach the receiver at the same time, and they might interfere so that they add up or possibly subtract, depending on their relative phases. We will see in the following simple example that with only one
reflector (the ground), one finds periodic positions where a receiver will detect peaks and nulls. This is called multipath fading and exists in all realistic wireless systems, being especially pronounced in mobile systems.

**Example 25.10—Line-of-sight link with real ground: a simple multipath fading model.** Consider the line-of-sight link shown in Fig. 25.10. The transmitting and receiving antennas are separated by a distance \( r \) and are at heights \( H \) and \( h \) above ground, which is assumed to be a perfect conductor. The receiving antenna receives not only the direct signal but also signals radiated by the transmitting antenna toward the ground that reflect toward the receiver. The effect of the ground can be taken into account by an image of the transmitting antenna, as shown in the figure. At a mobile receiver, the phases of the direct and reflected signals can differ by an even number of half wavelengths, which amounts to the waves adding up or canceling out periodically as the receiver moves away or toward the transmitter. This multipath fading becomes significantly more complicated when other reflective bodies, such as buildings and vehicles, are part of the propagation path.

We use free-space wave propagation every day in a number of places without really thinking about it, for example garage door openers (which typically work at 140 or 450 MHz), or remote controls for home entertainment equipment (which operate in the infrared region with wavelengths between 780 and 860 nm). But sometimes we would like to send information using waves that propagate through a medium other than air, and the issues involved can be very different, which Example 25.11 illustrates.

**Example 25.11—Radio communication with submarines.** For radio communications in normal circumstances we use frequencies greater than 100 kHz. Assume that we would like to establish a radio link with a submerged submarine. Table 20.1 tells us that this is not possible. Therefore, submarines use much lower frequencies (on the order of 10 kHz) for radio communications. Even this is not sufficiently low, for a submarine must be quite close to the surface in
order to make use of even such low frequencies. The low frequency implies a small bandwidth for communication, which means that very few words per minute can be transmitted.

Questions and problems: Q25.9, P25.9, and P25.10

25.5 Radar

Radars are essentially a type of wireless communication link, where the transmitter and receiver are located at the same place, as in Fig. 25.11. The transmitter sends a wave, which eventually reflects off some object (called a target, or scatterer) partly in the direction from which the wave came. This ("scattered") wave is received at the position of the transmitter, and from it some conclusions can be made about the object that caused the reflected wave.

Radar was invented for military purposes by the British in the Second World War and contributed greatly to the victory of the Allied forces. The word radar is an acronym for RA dio Detection And Ranging. Today, there are a number of commercial radar applications, such as weather radar for meteorology, mapping radar, police radar, anticollision radar for cars, and space-imaging radar.

The basic principle of a radar is as follows. The radar transmitter sends a wave with power $P_T$ toward a target. At the target, the power density is $P_T D / (4\pi r^2)$, where $r$ is the distance to the target, and $D$ is the radar antenna directivity. The target scatters the wave proportionally to a quantity called the radar scattering cross section, usually denoted by $\sigma(\theta, \phi)$, which is essentially the effective area of the target acting as a receiving antenna. When it reflects the wave, the target acts as a transmitting antenna with a directivity of $4\pi \sigma / (\lambda^2)$. Now the Friis formula can be applied one more time to obtain the power received by the radar receiver:

$$P_{\text{rec}} = P_T \frac{D^2(\theta, \phi)\sigma^2(\theta, \phi)}{16\pi^2 r^4}.$$  \hspace{1cm} (25.25)

This was derived assuming the radar uses the same antenna for transmission and reception, which is commonly but not always the case (see problem P25.11).

![Figure 25.11 A simple schematic of radar operation](image.png)
received signal in radar is very small, as it falls off as the fourth power of the distance from the target. However, much can be deduced from these signals when properly amplified. Two cases are described in Example 25.12.

Example 25.12—FM ranging radar and Doppler radar. In a type of ranging radar, the frequency of transmission is changed linearly from $f_1$ to $f_2$, as in Fig. 25.12. The transmitted signal in this case is said to be frequency modulated (FM). If $f_1$ is transmitted, by the time the wave at this frequency reflects back and reaches the receiver, the transmitter is transmitting a different frequency $f < f_2$. In the radar circuit, a so-called beat signal is made, corresponding to the difference $f - f_1$. This difference is obviously dependent on how far the target is, or how long it takes a wave traveling at the speed of light to get there and back. This time is exactly the same time it takes the transmitter to get from $f_1$ to $f$, and is known. So the distance from the target (the range) is

$$r = \frac{cT}{2} \frac{f - f_1}{f_2 - f_1}.$$  (25.26)

Figure 25.12 (a) Sketch of an FM ranging radar and (b) Doppler radar for measuring speed.
This is true for a stationary target. However, if the target is moving, it shifts the received frequency due to the Doppler effect. Using this Doppler shift, the speed of the target can be determined with a radar, as in Fig. 25.12b. In this case, a wave at frequency $f$ is transmitted, and the wave reflected off the target is $f \pm f_D$, where $f_D$ is the Doppler shift, and the sign in front of it depends on whether the target is moving toward or away from the radar. In the receiver circuit, the difference between the two frequencies is measured, and using that, the speed of the target is calculated. Police radars that monitor speed on the roads operate in this way, usually using frequencies around either 10 GHz or 30 GHz.

Questions and problems:  Q25.10 to Q25.12, P25.11 to P25.13

25.6 Some Electromagnetic Effects in Digital Circuits

We have so far not mentioned wave effects in computers or other digital systems. As the clocks that determine processing speed in computers become faster, electromagnetic effects such as radiation and coupling become more pronounced.

Digital circuits often have printed microstrip transmission lines connecting pins of two chips, possibly through some extra interconnects. The designer wants to make sure that a “one” is indeed a “one” when it reaches the second chip, and the same for a “zero” level. As rise times increase, depending on the logic family, transmission-line effects like overshoot, undershoot, ringing, reflections, and cross talk, can all become critical to maintaining noise margins. For example, in transistor-transistor logic (TTL), the values for a “one” are between 2.7 and 2 V, for a “zero” they are between 0.5 and 0.8 V, and the noise margin is 0.7 V (with a 10 to 90% rise time of 4 to 10 ns). In very fast gallium arsenide (GaAs) digital circuits (with a rise time of 0.2 to 0.4 ns), the values for a “one” are between −0.2 and −0.9 V, for a “zero” they are between −1.6 and −1.9 V, and the noise margin is 0.7 V. From these numbers it is seen that digital circuit margins are quite forgiving. For example, in the latter case, a 0.7-V undershoot on a 1.7-V signal is a 45% undershoot, which is considerable. However, it is easy to have a 45% variation in a signal if there is a discontinuity (impedance mismatch) in

![Diagram](image_url)

Figure 25.13 Near- and far-end cross-talk measurements in the case of two adjacent printed-circuit board traces
a transmission-line trace on a printed-circuit board. A special type of mismatch is coupling between adjacent traces, which are in effect parasitic inductances or capacitances. This is illustrated in Fig. 25.13 with two lines. If the line labeled “signal line” is excited by a step function, some of the voltage will get coupled to the next closest line even if the line is open-circuited. The coupled signal will appear at both ends of the line, and this is called near- and far-end cross talk. For example, the cross talk could be as high as 25% for two parallel traces, and resistors and trace bends can cause up to 15% and 5% reflections, respectively. All of these could be easily measured with time domain reflectometry (TDR) during the design of the digital backplane (printed-circuit board containing all the traces for chip interconnects).

25.7 Cooking with Electromagnetic Waves: Conventional Ovens and Microwave Ovens

In conventional ovens, heating of food is done principally by infrared radiation from the heaters. The infrared electromagnetic region of the spectrum is roughly between 900 nm and 10 μm. Another way to cook food is with a lower frequency in the microwave region with a wavelength on the order of centimeters. The two cooking mechanisms are quite different because of different skin depths of most foods in the two frequency ranges.

The frequency of infrared radiation is much higher than the highest frequency in Table 20.1. Most often, the food baked in the oven has a conductivity less than that of seawater, but for infrared frequencies the skin depth remains extremely small. We see therefore that a regular oven heats up a very thin surface layer, and then this heat is transferred by thermal conduction to the deeper layers. The thermal conductivity of most foods is not high. Therefore, cooking in regular ovens takes a lot of time, in particular if large chunks of food are being cooked. To expedite the process, we use higher temperatures, which result in some drying and browning of the food (at least the parts close to the surface).

In microwave ovens, the standard frequency is 2.45 GHz. Table 20.1 tells us that at that frequency, the skin depth for food (with conductivity usually less, and often significantly less, than that of seawater) is still relatively large (at least 1 cm, and often much larger). Consequently, the microwave oven instantly starts to heat most of the volume of the objects in it. Therefore, preparing food is much faster in a microwave oven than in a regular oven, but the food may not brown on the outside. If the cooking time is excessive, much of the water from the food can evaporate, and it can be first dried, then burned (as most of us have probably noticed).

**Questions**

Q25.1. Explain what the physical origin of loss in coaxial waveguides is.

Q25.2. Explain what the physical origin of loss in metallic waveguides is, and why the loss can be smaller than in coaxial cables.
Q25.3. Explain what the physical origin of loss in optical fiber is, and why the loss can be smaller than in metallic structures.

Q25.4. Explain what the physical origin of loss in a line-of-sight antenna link is.

Q25.5. What is the range in a line-of-sight link limited by?

Q25.6. Explain in your own words why there is attenuation in an ionized medium with neutral gas molecules.

Q25.7. A wave of frequency higher than the highest critical frequency for the ionosphere needs to be used for communication between two points of the earth. Is this possible? Explain.

Q25.8. A wave of extremely low frequency (e.g., below 100 Hz) coming from outer space penetrates through the ionosphere and reaches the earth’s surface. Explain.

Q25.9. Imagine a line-of-sight link in a hallway with conducting walls on top and bottom, and absorbing walls on the sides. How many waves can contribute to the received signal? How would you construct antenna images that approximate the influence of the walls?

Q25.10. Derive the radar equation (25.25).

Q25.11. Consider a Doppler radar at 10 GHz. The received signal from one car is in the audio range and can be between 300 Hz and 4 kHz. What is the range of speeds this radar can detect?

Q25.12. Consider an FM ranging radar in which the frequency varies linearly from \( f_1 = 10 \text{ GHz} \) to \( f_2 \) in \( T = 10 \mu s \). How would you choose \( f_2 \) in order to be able to detect targets 1 km away, if the radar bandwidth is 500 MHz?

---

**PROBLEMS**

P25.1. Calculate how much power is received in England if 1 MW is sent from Boston along a transatlantic 50-Ω cable at 10 kHz. You can assume that the main loss in the cable is due to conductor loss, and that \( R' = 0.005 \Omega/m \).

P25.2. What value of Pupin coils would you choose and how would you place them to reduce the loss in the cable in Example 25.1?

P25.3. Calculate the skin depth and attenuation coefficient of a rectangular waveguide with dimensions \( a = 23 \text{ mm} \) and \( b = 10 \text{ mm} \), at 10 GHz, if the waveguide is made of (1) copper, (2) aluminum, (3) silver, or (4) gold. What do you think are the engineering problems associated with each metal? Can you think of any combined solution?

P25.4. Calculate the skin depth of gold in the optical domain, at wavelengths of 500 nm, 830 nm, 1.33 μm, and 1.55 μm. How thin would one need to make a sheet of gold to see through?

P25.5. Compare the loss in the inner conductor and outer conductor of a coaxial cable at 1 MHz. Assume the conductors are made of copper, that the cable is filled with a dielectric of permittivity \( \varepsilon_r = 3 \), and that the dimensions are such that the inner conductor radius \( a = 0.45 \text{ mm} \) and inner radius of the outer conductor \( b = ae \).

P25.6. Plot the power attenuation in dB versus distance from 1 m to 1000 km on a logarithmic scale for: coaxial cable at 10GHz with \( \alpha = 0.5 \text{ dB/m} \), waveguide with \( \alpha = 0.1 \text{ dB/m} \), 1.55-μm single-mode optical fiber with \( \alpha = 0.1 \text{ dB/km} \), and a free space
link at 10 GHz with a horn antenna with 20-dB directivity and a 1-m diameter dish antenna.

P25.7. Calculate the dimensions for a rectangular waveguide with a dominant TE_{10} mode at cable TV frequencies between 100 and 600 MHz.

P25.8. A UHF radio system for communication between airplanes uses antennas with a directivity of 2. What is the maximum line-of-sight range between two airplanes at an altitude of 10 km? If the required received power is 10 pW, what is the minimum transmitted power \( P_t \) required for successful transmission at 100 MHz, 300 MHz, and 1 GHz?

P25.9. Calculate the effective area of a dish antenna for TV that requires a 1-degree beamwidth in both \( \theta \) and \( \phi \) planes, assuming one of the standard cable frequencies (e.g., 225 MHz). Is this a practical antenna? (Note: you can use an approximate formula for the maximal directivity given the beamwidths, \( \alpha_1 \) and \( \alpha_2 \), in the two planes, \( D \approx 32,000/(\alpha_1 \alpha_2) \), where the beamwidths are given in degrees.)

P25.10. If a satellite is 1000 km above the earth’s surface, and has a 0.1-degree beamwidth in both planes, calculate the corresponding directivity using the approximate formula in the previous problem. Find the size of the footprint on the earth’s surface, and the effective area of the antenna at a satellite frequency of 4 GHz.

P25.11. Derive the radar equation (25.25) for a radar that uses two antennas, one for transmitting and another for receiving.

P25.12. Assuming a 10-GHz police radar uses an antenna with a directivity of 20 dB (standard horn), and your car has a scattering cross section of 100 \( \lambda^2 \), plot the received power as a function of target distance, for a transmitted power of 1 W. If the receiver sensitivity is 10 nW, how close to the radar would you need to slow down to avoid getting a speeding ticket?

P25.13. How large is the dynamic range of the radar from problem P25.12? (The dynamic range is the ratio of the largest to smallest signal power detected, expressed in decibels.)