

Chapter 11: Common Antennas and Applications

11.1 Aperture antennas and diffraction

11.1.1 Introduction

Antennas couple circuits to radiation, and vice versa, at wavelengths that can extend into the infrared region and beyond. The output of an antenna is a voltage or field proportional to the input field strength $\bar{E}(t)$ and at the same frequency. By this definition, devices that merely amplify, detect, or mix signals are not antennas because they do not preserve phase and frequency, although they generally are connected to the outputs of antennas. For example, some sensors merely sense the increased temperature and heating caused by incoming waves. Chapter 10 introduced short-dipole and small-loop antennas, and arrays thereof. Chapter 11 continues with an introductory discussion of aperture antennas and diffraction in Section 11.1, and of wire antennas in 11.2. Applications are then discussed in Section 11.4 after surveying the basics of wave propagation and thermal emission in Section 11.3. These applications include communications, radar and lidar, radio astronomy, and remote sensing. Most optical applications are deferred to Chapter 12.

11.1.2 Diffraction by apertures

Plane waves passing through finite openings emerge propagating in all directions by a process called *diffraction*. Antennas that radiate or receive plane waves within finite apertures are *aperture antennas*. Examples include the parabolic reflector antennas used for radio astronomy, radar, and receiving satellite television signals, as well as the lenses and finite apertures employed in cameras, microscopes, telescopes, and many optical communications systems. As in the case of dipole antennas; we assume reciprocity and knowledge of the source fields or equivalent currents.

Since we have already derived expressions for fields radiated by arbitrary current distributions, one approach to finding aperture-radiated fields is to determine current distributions equivalent to the given aperture fields. Then these equivalent currents can be replaced by a continuous array of Hertzian dipoles for which we know the radiated far fields.

Consider a uniform current sheet \bar{J} [$A\ m^{-1}$] occupying the x-z plane, as illustrated in Figure 11.1.1. Maxwell's equations are then satisfied by:

$$\bar{E} = \hat{z}E_0 e^{-jky} \quad \bar{H} = \hat{x}(E_0/\eta_0) e^{-jky} \quad (\text{for } y > 0) \quad (11.1.1)$$

$$\bar{E} = \hat{z}E_0 e^{+jky} \quad \bar{H} = -\hat{x}(E_0/\eta_0) e^{+jky} \quad (\text{for } y < 0) \quad (11.1.2)$$

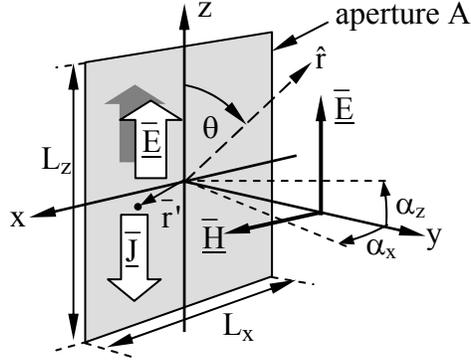


Figure 11.1.1 Aperture radiation from an equivalent current sheet.

The electric field $\hat{z}E_o \equiv \bar{E}_o$ must satisfy the boundary condition (2.6.11) that:

$$\bar{J}_s = \hat{y} \times [\bar{H}(y=0_+) - \bar{H}(y=0_-)] = \hat{y} \times \left[\hat{x} \frac{E_o}{\eta_o} + \hat{x} \frac{E_o}{\eta_o} \right] \quad (11.1.3)$$

$$\bar{J}_s = -\hat{z} 2 \frac{E_o}{\eta_o} = -2 \frac{\bar{E}_o}{\eta_o} \left[\text{Am}^{-1} \right] \quad (11.1.4)$$

Therefore we can consider any plane wave emerging from an aperture as emanating from an equivalent current sheet \bar{J}_s given by (11.1.4) provided we neglect radiation from the charges and currents induced at the aperture edges. They can generally be neglected if the aperture is large compared to a wavelength and if we remain close to the y axis in the direction of propagation, because then the aperture area dominates the observable radiating area. This approximation (11.1.4) for a finite aperture is valid even if the strength of the plane wave varies across the aperture slowly relative to a wavelength.

The equivalent current sheet (11.1.4) radiates according to (11.1.5) [from (10.2.8)], where we represent the current sheet by an equivalent array of Hertzian dipoles of length dz and current $\bar{I} = \bar{J}_s dx$:

$$\bar{E} = \hat{\theta} \frac{jkI d\eta_o}{4\pi r} e^{-jkr} \sin \theta \quad (\text{far-field radiation}) \quad (11.1.5)$$

The far fields radiated by the z-polarized current sheet $\bar{J}_s(x,z)$ in the aperture A are then:

$$\begin{aligned} \bar{E}_{ff}(\theta, \phi) &\cong \hat{\theta} \frac{j\eta_o}{2\lambda r} \sin \theta \int_A J_z(x,z) e^{-jkr(x,z)} dx dz \\ &\cong -\hat{\theta} \frac{j}{\lambda r} \sin \theta \int_A E_o(x,y) e^{-jkr(x,z)} dx dz \end{aligned} \quad (11.1.6)$$

To simplify the integral we can assume all rays are parallel by using the Fraunhofer far-field approximation:

$$e^{-jk\mathbf{r}(x,z)} \cong e^{-jk r_0} e^{+jk\hat{\mathbf{r}} \cdot \bar{\mathbf{r}}'} \quad (\text{Fraunhofer approximation}) \quad (11.1.7)$$

where we define position within the aperture $\bar{\mathbf{r}}' \equiv x\hat{x} + z\hat{z}$, and the distance $r_0 = (x^2 + y^2 + z^2)^{0.5}$. The Fraunhofer approximation is generally used when $r_0 > 2L^2/\lambda$. Then:

$$\bar{\mathbf{E}}_{\text{ff}}(\theta, \phi) \cong -\hat{\theta} \frac{j}{\lambda r} e^{-jk r_0} \sin \theta \int_A \underline{\mathbf{E}}_{\text{oz}}(x, z) e^{+jk\hat{\mathbf{r}} \cdot \bar{\mathbf{r}}'} dx dz \quad (11.1.8)$$

Those points in space too close to the aperture for the Fraunhofer approximation to apply lie in the *Fresnel region* where $r < \sim 2L^2/\lambda$, as shown in (11.1.4). If we restrict ourselves to angles close to the y axis we can define the angles α_x and α_z from the y axis in the x and z directions, respectively, as illustrated in Figure 11.1.1, so that:

$$\hat{\mathbf{r}} \cdot \bar{\mathbf{r}}' \cong x \sin \alpha_x + z \sin \alpha_z \cong x\alpha_x + z\alpha_z \quad (11.1.9)$$

Therefore, close to the y axis (11.1.8) can be approximated⁵⁵ as:

$$\bar{\mathbf{E}}_{\text{ff}}(\alpha_x, \alpha_z) \cong -\hat{\theta} \frac{j}{\lambda r} e^{-jk r_0} \int_A \underline{\mathbf{E}}_{\text{oz}}(x, z) e^{+j2\pi(x\alpha_x + z\alpha_z)/\lambda} dx dz \quad (11.1.10)$$

which is the Fourier transform of the aperture field distribution $\underline{\mathbf{E}}_{\text{oz}}(x, z)$, times a factor that depends on distance r and wavelength λ . Unlike the usual Fourier transform for converting signals between the time and frequency domains, this reversible transform in (11.1.10) is between the aperture spatial domain and the far-field angular domain.

For reference, the *Fourier transform* for signals is:

$$\underline{\mathbf{S}}(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi ft} dt \quad (11.1.11)$$

$$s(t) = \int_{-\infty}^{\infty} \underline{\mathbf{S}}(f) e^{j2\pi ft} df \quad (11.1.12)$$

The Fourier transform (11.1.11) has exactly the same form as the integral of (11.1.10) if we replace the aperture coordinates x and z with their wavelength-normalized equivalents x/λ and z/λ , analogous to time t ; α is analogous to frequency f .

Assume the aperture of Figure 11.1.1 is z-polarized, has dimensions $L_x \times L_z$, and is uniformly illuminated with amplitude $\underline{\mathbf{E}}_0$. Then its far fields can be computed using (11.1.10):

⁵⁵ In the *Huygen's approximation* a factor of $(1 + \cos\alpha)/2$ is added to improve the accuracy, but this has little impact near the y axis. In this expression α is the angle from the direction of propagation (y axis) in any direction.

$$\bar{E}_{\text{ff}}(\alpha_x, \alpha_z) \cong -\hat{\theta} \frac{j}{\lambda r} e^{-jk r_0} \int_{-L_z/2}^{+L_z/2} e^{+j2\pi\alpha_z z/\lambda} \int_{-L_x/2}^{+L_x/2} \underline{E}_0(x, z) e^{+j2\pi\alpha_x x/\lambda} dx dz \quad (11.1.13)$$

The inner integral yields:

$$\begin{aligned} \int_{-L_x/2}^{+L_x/2} \underline{E}_0(x, z) e^{+j2\pi\alpha_x x/\lambda} dx &= \underline{E}_0 \frac{\lambda}{j2\pi\alpha_x} [e^{+j\pi\alpha_x L_x/\lambda} - e^{-j\pi\alpha_x L_x/\lambda}] \\ &= \underline{E}_0 \frac{\sin(\pi\alpha_x L_x/\lambda)}{\pi\alpha_x/\lambda} \end{aligned} \quad (11.1.14)$$

The outer integral yields a similar result, so the far field is:

$$\bar{E}_{\text{ff}}(\theta, \phi) \cong -\hat{\theta} \frac{j}{\lambda r} \underline{E}_{0z} e^{-jk r_0} \bullet L_x L_z \frac{\sin(\pi\alpha_x L_x/\lambda)}{\pi\alpha_x L_x/\lambda} \frac{\sin(\pi\alpha_z L_z/\lambda)}{\pi\alpha_z L_z/\lambda} \quad (11.1.15)$$

The total power P_t radiated through the aperture is simply $A |\underline{E}_0|^2 / 2\eta_0$, where $A = L_x L_z$, so the antenna gain $G(\alpha_x, \alpha_z)$ given by (10.3.1) is:

$$G(\alpha_x, \alpha_z) \cong \frac{|\bar{E}_{\text{ff}}(\alpha_x, \alpha_z)|^2 / 2\eta_0}{P_t / 4\pi r^2} \quad (11.1.16)$$

$$\cong A \frac{4\pi}{\lambda^2} \left(\frac{\sin^2(\pi\alpha_x L_x/\lambda)}{(\pi\alpha_x L_x/\lambda)^2} \right) \left(\frac{\sin^2(\pi\alpha_z L_z/\lambda)}{(\pi\alpha_z L_z/\lambda)^2} \right) \quad (11.1.17)$$

The function $(\sin x)/x$ appears so often in electrical engineering that it has its own symbol 'sinc(x)'. Note that $\text{sinc}(0) = 1$ since $\sin(x) \cong x - (x^3/6)$ for $x \ll 1$. This gain pattern is plotted in Figure 11.1.2. The first nulls occur when $\pi\alpha_i L_i/\lambda = \pi$ ($i = x$ or z), and therefore $\alpha_{\text{null}} = \lambda/L$, where a narrower beamwidth α corresponds to a wider aperture L . The on-axis gain is:

$$G(0, 0) = \frac{4\pi}{\lambda^2} A \quad (\text{gain of uniformly illuminated aperture area } A) \quad (11.1.18)$$

Equation (11.1.18) applies to any uniformly illuminated aperture antenna, and such antennas have on-axis effective areas $A(\theta, \phi)$ that approach their physical areas A , and have peak gains $G_0 = 4\pi A/\lambda^2$. The antenna pattern of Figure 11.1.2 vaguely resembles that of circular apertures as well, and the same nominal angle to first null, λ/L , roughly applies to all. Such diffraction patterns largely explain the limiting angular resolution of telescopes, cameras, animal eyes, and photolithographic equipment used for fabricating integrated circuits.

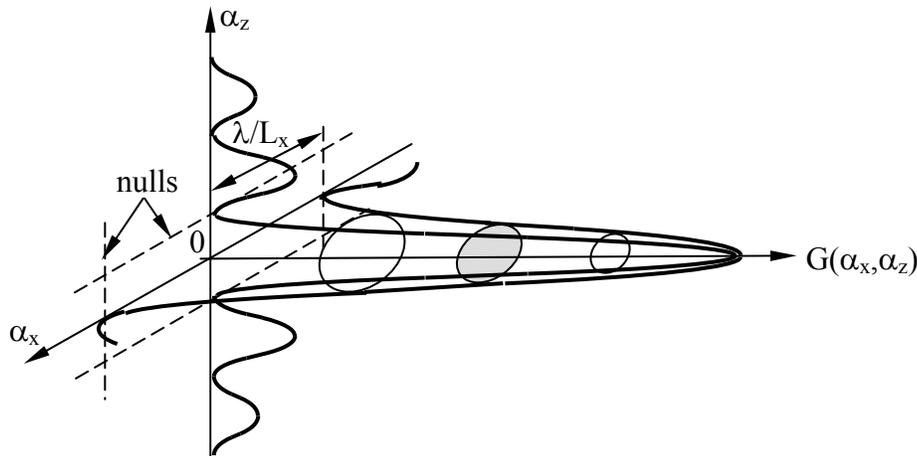


Figure 11.1.2 Antenna gain for uniformly illuminated rectangular aperture.

The coupling between two facing aperture antennas having effective areas A_1 and A_2 is:

$$P_{r_2} = \frac{P_{t_1}}{4\pi r^2} G_1 A_2 = \frac{P_{t_1}}{\lambda^2 r^2} A_1 A_2 = P_{t_1} \left(\frac{\lambda}{4\pi r} \right)^2 G_1 G_2 \quad (11.1.19)$$

where P_{r_2} and P_{t_1} are the power received by antenna 2 and the total power transmitted by antenna 1, respectively. For (11.1.19) to be valid, $r^2 \lambda^2 \gg A_1 A_2$; if $A_1 = A_2 = D^2$, then we require $r \gg D^2/\lambda$ for validity. Otherwise (11.1.19) could predict that more power would be received than was transmitted.

Example 11.1A

What is the angle between the first nulls of the diffraction pattern for a visible laser ($\lambda = 0.5$ microns) illuminating a 1-mm square aperture (about the size of a human iris)? What is the approximate diffraction-limited angular resolution of the human visual system? How does this compare to the maximum angular diameters of Venus, Jupiter, and the moon (~ 1 , ~ 1 , and ~ 30 arc minutes in diameter, respectively)?

Solution: The first null occurs at $\phi = \sin^{-1}(\lambda/L) \cong 5 \times 10^{-7}/10^{-3} = 5 \times 10^{-4}$ radians = $0.029^\circ \cong 1.7$ arc minutes. This is 70 percent larger than the planets Venus and Jupiter at their points of closest approach to Earth, and ~ 6 percent of the lunar diameter. Cleverly designed neuronal connections in the human visual system improve on this for linear features, as can a dark-adapted iris, which has a larger diameter.

Example 11.1B

A cell-phone dipole antenna radiates one watt toward a uniformly illuminated square aperture antenna of area $A =$ one square meter. If $P_r = 10^{-9}$ watts are required by the receiver for satisfactory link performance, how far apart r can these two terminals be? Does this depend on the shape of the aperture antenna if A remains constant?

Solution: $P_r = AP_t G_t / 4\pi r^2 \Rightarrow r = (AP_t G_t / 4\pi P_r)^{0.5} = (1 \times 1 \times 1.5 / 4\pi 10^{-9})^{0.5} \cong 10.9$ km The on-axis gain G_t of a uniformly illuminated constant-phase aperture antenna is given by (11.1.18). The denominator depends only on the power transmitted through the aperture A , not on its shape. The numerator depends only on the on-axis far field \bar{E}_{ff} , given by (11.1.10), which again is independent of shape because the phase term in the integral is unity over the entire aperture. Since the on-axis gain is independent of aperture shape, so is the effective area A since $A(\theta, \phi) = G(\theta, \phi)\lambda^2 / 4\pi$. The on-axis effective area of a uniformly illuminated aperture approximates its physical area.

11.1.3 Common aperture antennas

Section 11.1.2 derived the basic equation (11.1.10) that characterizes the far fields radiated by aperture antennas excited with z-polarized electric fields $\bar{E}_o(x, z) = \hat{z}E_{oz}(x, z)$ in the x-z aperture plane:

$$\bar{E}_{ff}(\theta, \phi) \cong \hat{\theta} \frac{j}{\lambda r} \sin\theta e^{-jkr} \iint_A E_{oz}(x, z) e^{jk\hat{r} \cdot \bar{r}'} dx dz \quad (11.1.20)$$

The unit vector \hat{r} points from the antenna toward the receiver and \bar{r}' is a vector that locates $\bar{E}_o(\bar{r}')$ within the aperture. This expression assumes the receiver is sufficiently far from the aperture that a single unit vector \hat{r} suffices for the entire aperture and that the receiver is therefore in the *Fraunhofer region*. The alternative is the near-field Fresnel region where $r < 2D^2/\lambda$, as discussed in Section 11.1.4; D is the aperture diameter. It also assumes the observer is close to the axis perpendicular to the aperture, say within $\sim 40^\circ$. The Huygen's approximation extends this angle further by replacing $\sin\theta$ with $(1 + \cos\beta)/2$, where θ is measured from the polarization axis and β is measured from the y axis:

$$\bar{E}_{ff}(\theta, \phi) \cong \hat{\theta} \frac{j}{2\lambda r} (1 + \cos\beta) e^{-jkr} \iint_A E_{oz}(x, z) e^{-jk\hat{r} \cdot \bar{r}'} dx dz \quad (\text{Huygen's approximation}) \quad (11.1.21)$$

Evaluating the on-axis gain of a uniformly excited aperture of physical area A having $\bar{E}_{oz}(x, z) = E_{oz}$ is straightforward when using (11.1.21) because the exponential factor in the integral is unity within the entire aperture. The gain follows from (11.1.16). The results are:

$$\bar{E}_{ff}(0, 0) \cong \hat{\theta} \frac{j}{\lambda r} e^{-jkr} \iint_A E_{oz} dx dz \quad (\text{on-axis field}) \quad (11.1.22)$$

$$G(0, 0) = \frac{|\bar{E}_{ff}(0, 0)|^2 / 2\eta_o}{P_T / 4\pi r^2} \quad (\text{on-axis gain}) \quad (11.1.23)$$

But the total power P_T transmitted through the aperture area A can be evaluated more easily than the alternative of integrating the radiated intensity $I(\theta, \phi)$ over all angles. The intensity I within the aperture is $|\overline{E}_{oz}|^2/2\eta_0$, therefore:

$$P_T = \frac{|\overline{E}_{oz}|^2}{2\eta_0} A \quad (11.1.24)$$

Then substitution of (11.1.22) and (11.1.24) into (11.1.23) yields the gain of a uniformly illuminated lossless aperture of physical area A :

$$G(0,0) = \frac{(\lambda r)^{-2} (E_{oz} A)^2 / 2\eta_0}{(E_{oz}^2 / 2\eta_0) A / 4\pi r^2} = \frac{\lambda^2}{4\pi} A \quad (\text{gain of uniform aperture}) \quad (11.1.25)$$

The off-axis gain of a uniformly illuminated aperture depends on its shape, although the on-axis gain does not.

Perhaps the most familiar radio aperture antennas are parabolic dishes having a point feed that radiates energy toward a *parabolic mirror* so as to produce a planar wave front for transmission, as suggested in Figure 11.1.3(a). Conversely, incoming radiation is focused by the mirror on the *antenna feed*, which intercepts and couples it to a transmission line connected to the receiver. Typical focal lengths (labeled “ f ” in the figure) are \sim half the diameter D for radio systems, and are often much longer for optical mirrors that produce images.

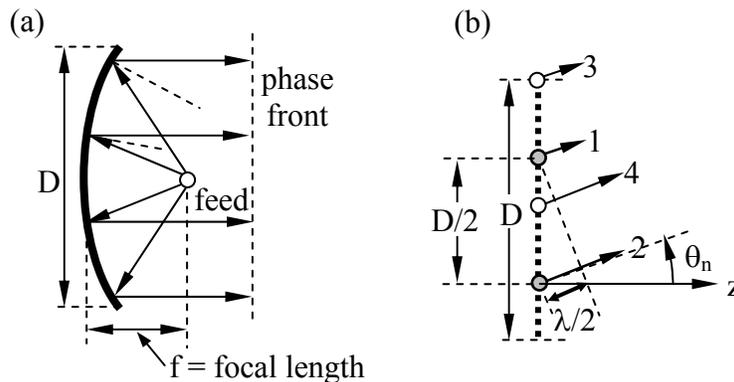


Figure 11.1.3 Aperture antennas and angle of first null.

Figure 11.1.3(b) suggests the angle θ_n at which the first null of a uniformly illuminated rectangular aperture of width D occurs; it is the angle at which all the phasors emanating from each point on the aperture integrate in (11.1.10) to zero. In this case it is easy to pair the phasors originating $D/2$ apart so each pair cancels at θ_n . For example, radiation from aperture element 2 has to travel $\lambda/2$ farther than radiation from element 1 and therefore they cancel each other. Similarly radiation from elements 3 and 4 cancel, and the sum of all such pairs cancel at the null angle:

$$\theta_n = \sin^{-1}(\lambda/D) \text{ [radians]} \quad (11.1.26)$$

where $\theta_n \cong \lambda/D$ for $\lambda/D \ll 1$.

Approximately the same null angle results for uniformly illuminated circular apertures, for which integration yields $\theta_n \cong 1.2\lambda/D$. Consider the human eye, which has a pupil that normally is ~ 2 mm in diameter, but can dilate to ~ 1 cm in the dark. For a wavelength of 5×10^{-7} meters, we find the normal diffraction-limited angular resolution of the eye is $\sim \lambda/D = 5 \times 10^{-7} / (2 \times 10^{-3}) = 2.5 \times 10^{-4}$ radians or ~ 0.014 degrees, or ~ 0.9 arc minutes. For comparison, the planets Venus and Jupiter are approximately 1 arc minute in diameter at closest approach, and the moon and sun are approximately 30 arc minutes in diameter.

A large astronomical telescope like the 200-inch system at Palomar has a nominal diffraction limit of $\lambda/D \cong 5 \times 10^{-7} / 5.08 \cong 10^{-7}$ radians or ~ 0.02 arc seconds, where there are 60 arc seconds in an arc minute, and 60 arc minutes in a degree. This is adequate to resolve an automobile on the moon. Unfortunately mirror surface imperfections, focus misplacement, and atmospheric turbulence limit the actual angular resolution of Palomar to ~ 1 arc second on the very best nights; normal daytime turbulence is far worse.

Practical issues generally shape the design of parabolic radio antennas. First, mechanical (gravity and wind) and thermal issues (temperature gradients) usually limit their angular resolution to ~ 1 arc minute; most antennas are too small relative to λ to achieve this resolution, however. Second, the antenna feed that illuminates the parabola tends to spray its radiation in a broad pattern that extends past the edge of the reflector creating backlobes. Third, the finite extent of the aperture results in an antenna pattern with sidelobes and unwanted responsiveness to directions beyond the main lobe.

Equation (11.1.10) showed how the angular dependence of the far-fields of an aperture was proportional to the Fourier transform of the aperture excitation function. For example, (11.1.17) and Figure 11.1.2 showed the radiation pattern of a uniformly illuminated aperture measuring L_x by L_z . Significant energy was radiated beyond the first nulls at $\alpha_x = \lambda/L_x$ and $\alpha_z = \lambda/L_z$. A finite aperture necessarily radiates something at all angles, just as a finite voltage pulse in a circuit has at least some energy at all frequencies; the sharper the pulse edges, the more high-frequency content they have. Therefore, reducing the sharp discontinuities in field strength at the aperture edge, a strategy called tapering, can reduce diffraction sidelobes. Antenna feeds are typically designed to reduce field strengths by factors of 2-4 at the mirror edges for this reason, but the resulting effective reduction in aperture diameter produces a slightly broader main lobe, just as the Fourier transform of a narrower pulse produces a broader spectral band.

A final consideration is sometimes important when designing aperture antennas, and that involves aperture blockage, which results when transmitted radiation reflected from the mirror is blocked or scattered by the antenna feed at the focus of the parabola. Not only does the scattered radiation contribute to side or back lobes, but it also is lost to the main beam. Example 11.1C illustrates these issues.

Example 11.1C

A uniformly illuminated square aperture is 1000 wavelengths long on each side. What is its antenna gain $G(\alpha_x, \alpha_z)$ for $\alpha \ll 1$? What is the gain G_o' if the center of this fully illuminated aperture is blocked by a square absorber 100 wavelengths on a side? What is the extent and approximate magnitude of the sidelobes introduced by the blockage?

Solution: The on-axis gain $G_o = A4\pi/\lambda^2 = 1000^2 \times 4\pi$. The angular dependence is proportional to the square of the far-field, $|E(\alpha_x, \alpha_z)|^2$, where the far field is the Fourier transform of the aperture field distribution. The full solution for $G(\alpha_x, \alpha_z)$ is developed in Equations (11.1.13–17). If the blocked portion of the aperture is illuminated so the energy there is absorbed, then the total transmitted power P_t in the expression (11.1.16) for gain is unchanged, while the area over which \bar{E}'_{ff} is integrated in (11.1.13) is reduced by the 1 percent blockage (100^2 is 1 percent of 1000^2). Therefore $|\bar{E}'_{ff}(0,0)|^2$, the numerator of (11.1.16), and $G_o'(0,0)$ are all reduced by a factor of $0.99^2 \cong 0.98$. Thus $G_o' \cong 0.98 G_o$. If the blocked portion of the aperture were not illuminated so as to avoid the one percent absorption, then $G_o'(0,0)$ would be reduced by only 1 percent: $G_o' = A'4\pi/\lambda^2$. The sidelobes for the blocked aperture follow from the Fourier transform (11.1.13), where the aperture excitation $E_o(x,z)$ is the sum of a positive square “boxcar” function 1000λ on a side, and a negative square boxcar 100λ on a side. Since this transform is linear, $\bar{E}'_{ff}(\alpha_x, \alpha_z)$ is the sum of the transforms of the positive and negative boxcar functions, and the antenna sidelobes therefore have contributions from each. Most important is the main lobe of the diffraction pattern of the smaller “blockage” boxcar, which has magnitude $\sim 0.01^2$ that of G_o' , and a half-power beamwidth θ_{BB} that is 10 times greater than the main lobe of the larger boxcar: $\theta_{BB} \cong \lambda/D_B = \lambda/100\lambda$. The total antenna pattern is the square of the summed transforms and more complicated; the innermost few sidelobes are approximately those of the original antenna, while the blockage-induced sidelobes are more important at greater angles.

11.1.4 Near-field diffraction and Fresnel zones

Often receivers are sufficiently close to the source that the Fraunhofer parallel-ray approximation of (11.1.7) is invalid. Then the *Huygen's approximation* (11.1.21) can be used:

$$\bar{E}_{ff} \cong \hat{\theta} \frac{j}{2\lambda r} (1 + \cos\beta) \iint_A E_{oz}(x,z) e^{-jk r(x,y)} dx dz \quad (\text{Huygen's approximation}) \quad (11.1.27)$$

for which the distance between the receiver and the point x,y in the aperture is defined as $r(x,y)$. This region close to a source or obstacle where the Fraunhofer approximation is invalid is called the *Fresnel region*.

If the phase of \underline{E}_{oz} in the source aperture is constant everywhere, then contributions to $\underline{E}_{ff}(0,0)$ from some parts of the aperture will tend to cancel contributions from other parts because they are out of phase. For example, contributions from the central circular zone where $r(x,y)$ ranges from r_0 to $r_0 + \lambda/2$ will largely cancel the contributions from the surrounding ring where $r(x,y)$ ranges from $r_0 + \lambda/2$ to $r_0 + \lambda$; it is easily shown that these two rings have approximately the same area, as do all such rings over which the delay varies by $\lambda/2$.⁵⁶ Such rings are illustrated in Figure 11.1.4(a).

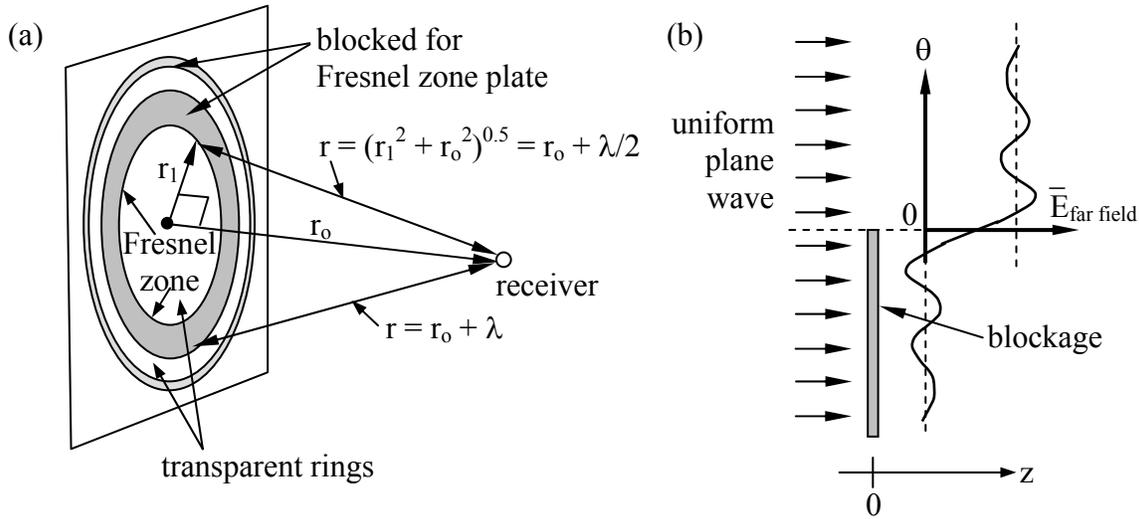


Figure 11.1.4 Fresnel zone plate.

One technique for maximizing diffraction toward an observer is therefore simply to physically block radiation from those alternate zones contributing negative fields, as suggested in Figure 11.1.4(a). Such a blocking device is called a *Fresnel zone plate*. The central ring having positive phase is called the *Fresnel zone*. Note that if only the central zone is permitted to pass, the received intensity is maximum, and if the first two zones pass, the received intensity is nearly zero because they have approximately the same area. The second zone is weaker, however, because r and θ are larger. Three zones can yield nearly the same intensity as the first zone alone because two of the three zones nearly cancel, and so on. By blocking alternate zones the received intensity can be many times greater than if there were no blockage at all. Thus a multi-ring zone plate acts as a lens by focusing energy received over a much larger area than would be intercepted by the receiver alone. This type of lens is particularly valuable for focusing very short-wave radiation such as x-rays which are difficult to reflect or diffract using traditional mirrors or lenses.

Another advantage of zone plate lenses is that they can be manufactured lithographically, and their critical dimensions are usually many times larger than the wavelengths involved. For example, an x-ray zone plate designed to operate at $\lambda = 10^{-8}$ [m] at a distance r_0 of one centimeter

⁵⁶ The area of the inner circle (radius a) is $\pi a^2 = \pi[(r_0 + \lambda/2)^2 - r_0^2] \cong \pi r_0 \lambda$ if $\lambda \ll 2r_0$. The area of the immediately surrounding Fresnel ring (radius b) is $\pi(b^2 - a^2) = \pi[(r_0 + \lambda)^2 - r_0^2] - \pi[(r_0 + \lambda/2)^2 - r_0^2] \cong \pi r_0 \lambda$, subject to the same approximation. Similarly, all other Fresnel rings can be shown to have approximately the same area if $\lambda \ll 2r_0$.

will have a central zone of diameter $2[(r_0 + \lambda/2)^2 - r_0^2]^{0.5} \cong 2(r_0\lambda)^{0.5} = 2 \times 10^{-5}$ meters, a dimension easily fabricated using modern semiconductor lithographic techniques.

Another example of diffraction is wireless communications in urban environments, which often involves line-of-sight reception of waves past linear obstacles slightly to one side or slightly obscuring the source. Again Huygen's equation (11.1.27) can be used to determine the result. Referring to Figure 11.1.4(b), if there is no blockage, traditional equations can be used to compute the received intensity. If exactly half the path is blocked by a wall obscuring the bottom half of the illuminated aperture, for example, then the integral in (11.1.27) will yield exactly half the previous value of E_{ff} , and the power (proportional to E_{ff}^2) will be reduced by a factor of four, or ~ 6 dB. If the observer moves up or down less than \sim half the radius of the Fresnel zone, then the received power will vary only modestly. For example, an FM radio (say 10^8 Hz) about 100 meters beyond a tall wide metal wall can have a line of sight that passes through the wall a distance of $\sim (r\lambda)^{0.5}/2 = 17$ meters below its top without suffering great loss; $(r\lambda)^{0.5}$ is the radius of the Fresnel zone. Conversely, a line-of-sight that passes less than ~ 17 meters above the top of the wall will also experience modest diffractive effects.

The Fresnel region approximately begins when the central ray arrives at distance r_0 , more than $\lambda/16$ ahead of rays from the perimeter of an aperture of diameter D . That is:

$$\sqrt{\left(\frac{D}{2}\right)^2 + r_0^2} - r_0 \gtrsim \frac{\lambda}{16} \quad (11.1.28)$$

For $D \ll R$ this becomes:

$$r_0 \left(\sqrt{\left(\frac{D}{2r_0}\right)^2 + 1} - 1 \right) \cong \frac{D^2}{8R} \gtrsim \frac{\lambda}{16} \quad (11.1.29)$$

Therefore the *Fresnel region* is:

$$r_0 \lesssim \frac{2D^2}{\lambda} \quad (\text{Fresnel region}) \quad (11.1.30)$$

11.2 Wire antennas

11.2.1 Introduction to wire antennas

Exact solution of Maxwell's equations for antennas is difficult because antennas typically have complex shapes for which it is difficult to match boundary conditions. Often complex wave expansions with many degrees of freedom are required, and even modern software tools can be challenged. Fortunately, most common *wire antennas* permit their current distributions to be guessed accurately relative to the given terminal current, as explained in Section 11.2.2. Once the current distribution is known everywhere, the radiated fields, radiation and dissipative

resistance, antenna gain, and antenna effective area can be calculated. If the antenna is used at a frequency far from resonance, the reactance can also be estimated.

If the antenna is small compared to a wavelength λ then its current distribution \underline{I} and the open-circuit voltage \underline{V}_{Th} can be determined using the quasistatic approximation. If the current distribution is known, then the radiated far-fields $\underline{\bar{E}}_{ff}$ can be computed using (10.2.8) by integrating the contributions $\Delta \underline{\bar{E}}_{ff}$ from each short current element $\underline{I}d$ (d is the element length and is replaced by ds in the integral), where:

$$\Delta \underline{\bar{E}}_{ff} = \hat{\theta} \frac{jk \underline{I} d \eta_0}{4\pi r} e^{-jkr} \sin \theta \quad (11.2.1)$$

$$\underline{\bar{E}}_{ff} \cong \frac{jk \eta_0}{4\pi r} \int_S \hat{\theta} \underline{I}(s) e^{-jkr} \sin \theta ds \quad (11.2.2)$$

For antennas small compared to λ the factor before the integral of (11.2.2) is nearly constant over the integrated length S , so average values suffice. If the wires run in more than one direction, the definition of $\hat{\theta}$ and θ must change accordingly; θ is defined by the local angle between $\underline{\bar{I}}$ and \hat{r} , where \hat{r} is the unit vector pointing from the antenna to the observer, as suggested in Figure 10.2.3. Equation (11.2.2), not surprisingly, reduces to (11.2.1) for a short straight wire carrying constant current \underline{I} over a distance $d \ll \lambda$.

Once the radiated fields are known for a given antenna input current \underline{I} , the radiated intensity can be integrated over a sphere surrounding the antenna to yield the total power radiated P_R and the radiation resistance R_r , which usually dominates the resistive component of the antenna impedance and corresponds to power lost through radiation (10.3.16). The radiation resistance is simply related to P_R :

$$R_r = \frac{2P_R}{|\underline{I}|^2} \text{ [ohms]} \quad (\text{radiation resistance}) \quad (11.2.3)$$

The open-circuit voltage can also be easily estimated for wire antennas small compared to λ . For example, the open-circuit voltage induced across a short dipole antenna shown in Figure 10.3.1 is simply the projection of the incident electric field $\underline{\bar{E}}$ on the electrical centers of the two metallic structures comprising the dipole, and Example 10.3D showed how the open-circuit voltage across a loop antenna was proportional to the time derivative of the magnetic flux through it. In both cases the open-circuit voltage reveals the directional properties of the antenna. Computation of the *radiation resistance* requires knowledge of the radiated fields and integration of the radiated power over all angles, however. Equation (10.3.16) showed that the radiation resistance of a short dipole antenna of length d is $(2\pi\eta_0/3)(d/\lambda)^2$ ohms. Slightly more complicated integrals over angles yield the radiation resistance for half-wave dipoles of length d and N -turn loop antennas of diameter $d \ll \lambda$: ~ 73 ohms and $\sim 1.9 \times 10^4 N^2 (d/\lambda)^4$ ohms, respectively. The higher radiation resistance of loop antennas often makes them the antenna of

choice when space is limited relative to wavelength, particularly when they are wound on a ferrite core ($\mu \gg \mu_0$) that increases their magnetic dipole moment.

Most wire antennas are not small compared to a wavelength, however, and the methods of the next section are then often used.

11.2.2 Current distribution on wires

The current distribution on wires is governed by Maxwell's equations, which are most easily solved for simple geometries such as that of a coaxial cable, as discussed in Example 7.1B. The fields for a TEM wave in a coaxial cable are cylindrically symmetric and a function of radius r :

$$\bar{\mathbf{E}}(r,z) = \hat{r}E_0(z)/r \text{ [V m}^{-1}\text{]} \quad (\text{coaxial cable electric field}) \quad (11.2.4)$$

$$\bar{\mathbf{H}}(r,z) = \hat{\theta}H_0(z)/r \text{ [A m}^{-1}\text{]} \quad (\text{coaxial cable magnetic field})^{57} \quad (11.2.5)$$

The energy density and Poynting's vector are proportional to field strength squared, so they decay as r^{-2} . Therefore the electromagnetic behavior of the line is dominated by the geometry near the central conductor where most of the electromagnetic energy is located, and the outer conductor can be deformed substantially before the fields near the center are significantly perturbed. For example, two-thirds of the power propagates within 10 cm of a 1-mm wire centered within a 1-meter outer cylinder, even though this represents only one-percent of the volume. This is easily shown by integrating the energy density from radius a to radius b , $\int_a^b E_0^2 r^{-2} 2\pi r dr = 2\pi E_0^2 \ln(b/a)$, and comparing the results for different sub-volumes.

Therefore the fields near the axis of the coaxial cable illustrated in Figure 11.2.1(a) are altered but little if the outer conductor is replaced by a ground plane as illustrated in Figure 11.2.1(b), or even by a second wire, as shown in Figure 11.2.1(c). The significance of Figure 11.2.1 is therefore that current distributions on thin wire antennas closely resemble those on equivalent TEM lines, provided the lines are not so many wavelengths long that the energy is lost before it reaches the end, or so tightly bent that the segments induce strong voltages on their neighbors. This TEM approximation is valid for understanding the examples of this section.

A widely used antenna is the half-wave dipole, illustrated in Figure 11.2.1(d), which exhibits essentially no reactive impedance because the electric and magnetic energy storages approximately balance. The radiation resistance for any half-wave dipole in free space is ~ 73 ohms. Section 7.4.2 discusses how these energies balance in any TEM structure of length $D = n\lambda/2$ where n is an integer. Typical bandwidths $\Delta\omega$ of a half-wave dipole are $\Delta\omega/\omega_0 = 1/Q \cong 0.1$, where $Q = \omega_0 W_T/P_d$, as discussed in Section 7.4.3 and (7.4.4).

⁵⁷ The magnetic field around a central wire, $H = I/2\pi r$, was given in (1.4.3).

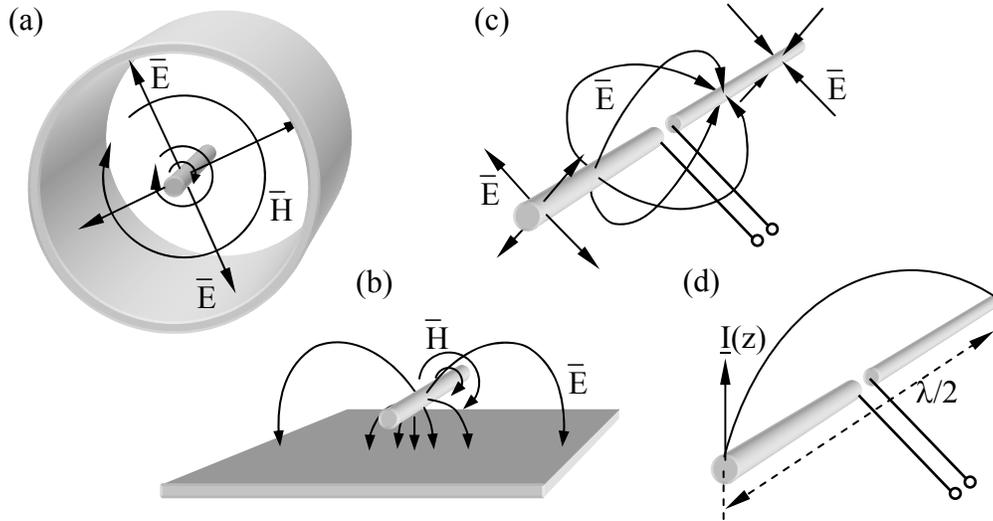


Figure 11.2.1 Fields near wire antennas resemble fields in TEM coaxial cables.

Figure 11.2.2 illustrates nominal current distributions on several antenna structures; these currents are consistent with those on comparable TEM lines propagating signals at the speed of light. The current distributions in the figure represent instantaneous distributions at the moment of current maximum.

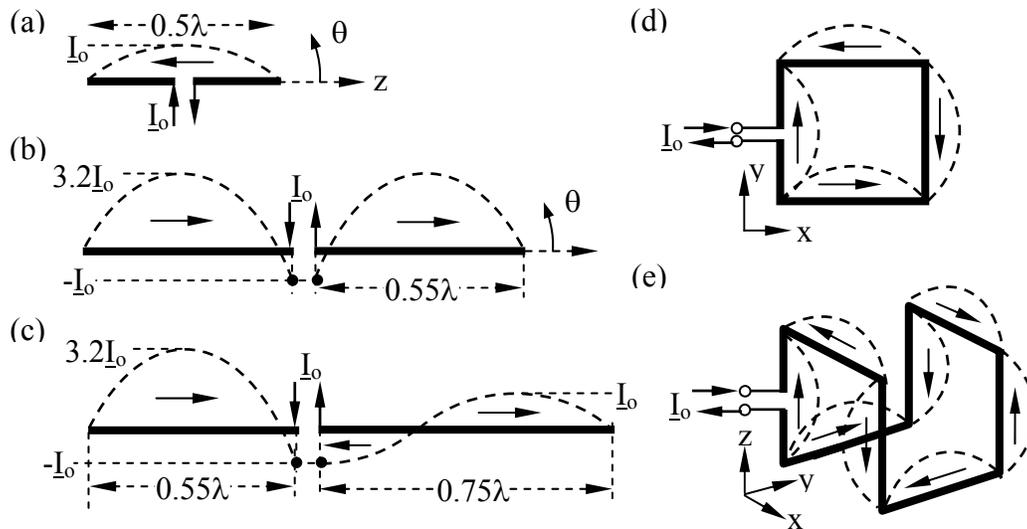


Figure 11.2.2 Current distributions on wire antenna structures.

In these idealized cases the currents everywhere on the antenna approach zero one-quarter cycle later as the energy all converts from magnetic to electric. The voltage distributions when the currents are zero resemble those on the equivalent TEM lines, and are offset spatially by $\lambda/4$; at

resonance the voltage peaks coincide with the current nulls. For example, the voltages and currents for Figure 11.2.2(a) resemble those of the open-circuited TEM resonator of Figure 7.4.1(a). The actual current and voltage distributions are slightly different from those pictured because radiation tends to weaken the currents farther from the antenna terminals, and because such free-standing or bent wires are not true TEM lines.

Figure 11.2.2(b) illustrates how terminal currents (I_o) can be made less than one-third the peak currents ($3.2 I_o$) flowing on the antenna simply by lengthening the two arms so they are each slightly longer than $\lambda/2$ so the current is close to a null at the terminals. Because smaller terminal currents thus correspond to larger antenna currents and radiated power, the effective radiation resistance of this antenna is increased well above the nominal 73 ohms of the half-wave dipole of (a). The reactance is slightly capacitive, however, and should be canceled with an inductor. Figure 11.2.2(c) illustrates how the peak currents can be made different in the two arms; note that the currents fed to the two arms must be equal and opposite, and this fact forces the two peak currents in the arms to differ. Figures (d) and (e) show more elaborate configurations, demonstrating that wire antennas do not have to lie in a straight line. The patterns for these antennas are discussed in the next section.

11.2.3 Antenna patterns

Once the current distributions on wire antennas are known, the antenna patterns can be computed using (11.2.2). Consider first the dipole antenna of Figure 11.2.2(a) and let its length be d , its terminal current be I_o' , and its maximum current be I_o . Then (11.2.2) becomes:

$$\bar{E}_{ff} \cong \frac{jk\eta_o}{4\pi r} \int_{-d/2}^{d/2} \hat{\theta} I(s) e^{-jkr} \sin\theta ds \quad (11.2.6)$$

$$\bar{E}_{ff} \cong \hat{\theta} \frac{j\eta_o I_o e^{-jkr}}{2\pi r \sin\theta} \left[\cos\left(\frac{kd}{2} \cos\theta\right) - \cos\left(\frac{kd}{2}\right) \right] \quad (11.2.7)$$

This expression, which requires some effort to derive, applies to symmetric dipole antennas of any modest length d ; I_o is the maximum current, which is not necessarily the terminal current. The common half-wave dipole has $d = \lambda/2$, so (11.2.7) reduces to:

$$\bar{E}_{ff} \cong \hat{\theta} \left(j\eta_o I_o e^{-jkr} / 2\pi r \sin\theta \right) \cos\left[(\pi/2) \cos\theta \right] \quad (\text{half-wave dipole}) \quad (11.2.8)$$

The antenna of Figure 11.2.2(b) can be considered to be a two-element antenna array (see Section 10.4.1) for which the two radiated phasors add in some directions and cancel in others, depending on the differential phase lag between the two rays. Antenna (b) has its peak gain at $\theta = \pi/2$, but its beamwidth is less than for (a) because rays from the two arms of the dipole are increasingly out of phase for propagation directions closer to the z axis, even more than for the half-wave dipole; thus the gain of (b) modestly exceeds that of (a). Whether one determines patterns numerically or by using the more intuitive phasor addition approach of Sections 10.4.1

and 10.4.5 is a matter of choice. Antenna (c) has very modest nulls for θ close to the $\pm z$ axis. The nulls are weak because the electric field due to $3.2I_0$ is only slightly reduced by the contributions from the phase-reversed segment carrying I_0 .

Simple inspection of the current distribution for the antenna of Figure 11.2.2(d) and use of the methods of Section 10.4.1 reveal that its pattern has peaks in gain along the $\pm x$ and $\pm y$ axes, and a null along the $\pm z$ axes. Extending simple superposition and phase cancellation arguments to other angular directions makes it possible to guess the form of the complete antenna pattern $G(\theta, \phi)$, and therefore to check the accuracy of any integration using (11.2.6) for all antenna arms. Similar simple phase addition/cancellation analysis reveals that the more complicated antenna (e) has gain peaks along the $\pm x$ and $\pm y$ axes, and nulls along the $\pm z$ axes, although the polarization of each peak is somewhat different, as discussed in an example. Exact determination of pattern (e) is confused by the fact that these wires are sufficiently close to each other to interact, so the current distribution may be modified relative to the nominal TEM assumption sketched in the figure.

Example 11.2A

Determine the relative gains and polarizations along the $\pm x$, $\pm y$, and $\pm z$ axes for the antenna illustrated in Figure 11.2.2(e).

Solution: The two x-oriented wires do not radiate in the $\pm x$ direction. The four z-oriented wires emit radiation that cancels in that direction (one pair cancels the other), while the two y-oriented wires radiate in-phase y-polarized radiation in the $\pm x$ direction with relative total electric field strength $E_y = 2$. We assume each $\lambda/2$ segment radiates a relative electric field of unity. Similarly, the two y-oriented wires do not radiate in the $\pm y$ direction. The four z-oriented wires emit radiation that cancels in that direction (one pair cancels the other), while the two x-oriented wires radiate in-phase x-polarized radiation in the $\pm y$ direction with relative total electric field strength $E_x = 2$. The four z-oriented wires do not radiate in the $\pm z$ direction, and the two out-of-phase pairs of currents in the x and y directions also cancel in that direction, yielding a perfect null. Thus the gains are equal in the x and y directions (but with x polarization along the y axis, and y-polarization along the x axis), and the gain is zero on the z axis.

11.3 Propagation of radio waves and thermal emission

11.3.1 Multipath propagation

Electromagnetic waves can be absorbed, refracted, and scattered as they propagate through linear media. One result of this is that beams from the same transmitter can arrive at a receiver from multiple directions simultaneously with differing delays, strengths, polarizations, and Doppler shifts. These separate phasors add constructively or destructively to yield an enhanced or diminished total response that is generally frequency dependent. Since cellular telephones are

mobile and seldom have a completely unobstructed propagation path, they often exhibit strong fading and multipath effects.

Consider first the simple case where a single beam arrives via a direct path and a reflected beam with one-quarter the power of the first arrives along a reflected path that is 100λ longer. If the powers of these two beams are constant, then the total received power will fluctuate with frequency. If the voltage received for the direct beam is \underline{V} and that of the second beam is $\underline{V}/2$ corresponding to quarter power, then when they are in phase the total received power is $1.5^2|\underline{V}|^2/2R$, where R is the circuit impedance. When they are 180° out of phase the power is $0.5^2|\underline{V}|^2/2R$, or one-ninth the maximum. This shift between maximum and minimum occurs each time the relative delay between the two paths changes by $\lambda/2$. Because the differential delay D is $\sim 100\lambda$, this represents a frequency change Δf of only one part in 200; $\Delta f/f = \lambda/2D$. Note that reflections can enhance or diminish the main signal, and clever antenna arrays can always compensate for the differential delays experienced from different directions so as to enhance the result.

Since cellular phones can have path differences of ~ 1 km at wavelengths of ~ 10 cm, their two-beam frequency maxima can be separated by as little as $10^{-4}f$, where f can be $\sim 10^9$ Hz. Fortunately this separation of 10^5 Hz is large compared to typical voice bandwidths. Alternatively, cellular phone signals can be coded to cover bandwidths large compared to fading bandwidths so the received signal strength is averaged over multiple frequency nulls and peaks and is therefore more stable.

Multipath also produces nulls in space if the rays arrive from different directions. For example, if two rays A and B of wavelength λ and arrive from angles separated by a small angle γ , then the distance D between intensity maxima and minima along a line roughly perpendicular to the direction of arrival will be $\sim \lambda/(2\sin\gamma)$. The geometry is sketched in Figure 11.3.1. Three or more beams can be analyzed by similar phasor addition methods. Sometimes one of the beams is reflected from a moving surface, or the transmitter or receiver are moving, so these maxima and minima can vary rapidly with time.

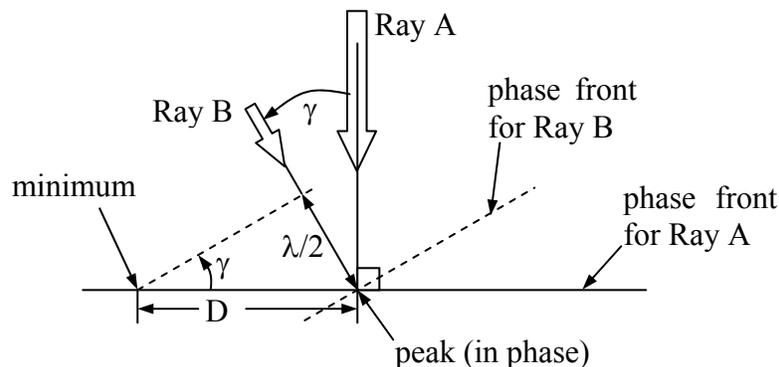


Figure 11.3.1 Maxima and minima created by multipath.

Example 11.3A

Normal broadcast NTSC television signals have 6-MHz bandwidth. If a metal building reflects perfectly a signal that travels a distance L further than the direct beam before the two equal-strength beams sum at the receiving antenna, how large can L be and still ensure that there are not two nulls in the 6-MHz passband between 100 and 106 MHz?

Solution: The differential path L is L/λ wavelengths long. If this number of wavelengths increases by one, then $L/\lambda' = L/\lambda + 1$ as λ decreases to λ' ; this implies $\lambda/\lambda' = 1 + \lambda/L = f'/f = 1.06$. When the direct and reflected signals sum, the 2π phase change over this frequency band will produce one null, or almost two nulls if they fall at the band edges. Note that only the differential path length is important here. Therefore $L = \lambda/(1.06 - 1) \cong 16.7\lambda = 16.7c/f \cong 16.7 \times 3 \times 10^8 / 10^8 = 50.0$ meters.

11.3.2 Absorption, scattering, and diffraction

The terrestrial *atmosphere* can absorb, scatter, and refract electromagnetic radiation. The dominant gaseous absorbers at radio and microwave wavelengths are water vapor and oxygen. At submillimeter and infrared wavelengths, numerous trace gases such as ozone, NO, CO, OH, and others also become important. At wavelengths longer than 3 mm only the oxygen absorption band $\sim 50 - 70$ GHz is reasonably opaque. Horizontal attenuation at some frequencies 57-63 GHz exceeds 10 dB/km, and vertical attenuation can exceed 100 dB. The water vapor band 20-24 GHz absorbs less than 25 percent of radiation transmitted toward zenith or along a ~ 2 -km horizontal path.

More important at low frequencies is the *ionosphere*, which reflects all radiation below its plasma frequency f_o , as discussed in Section 9.5.3. Radio waves transmitted vertically upward at frequency f are reflected directly back if any ionospheric layer has a plasma frequency $f_p < f$, where f_p is given by (9.5.25) and is usually below 15 MHz. The ionosphere generally extends from ~ 70 to ~ 700 km altitude, with electron densities peaking ~ 300 km and exhibiting significant drops below ~ 200 km at night when solar radiation no longer ionizes the atmosphere fast enough to overcome recombination.

Above the plasma frequency f_p radio waves are also perfectly reflected at an angle of reflection θ_r equal to the angle of incidence θ_i if θ_i exceeds the critical angle $\theta_c(f)$ (9.2.30) for any ionospheric layer. The critical angle $\theta_c(f) = \sin^{-1}(\epsilon_{ion}/\epsilon_o)$, where the permittivity of the ionosphere $\epsilon_{ion}(f) = \epsilon_o[1 - (f_p/f)^2]$. Since $\epsilon_o > \epsilon_{ion}$ at any finite frequency, there exists a grazing angle of incidence θ_i where waves are perfectly reflected from the ionosphere at frequencies well above f_p . The curvature of the earth precludes grazing incidence with $\theta_i \rightarrow 90^\circ$ unless the bottom surface of the ionosphere is substantially tilted. Therefore the maximum frequency at which radio waves can bounce around the world between the ionosphere and the surface of the earth is limited to $\sim 2f_p$, depending on the height of the ionosphere for the frequency of interest.

The most important non-gaseous atmospheric absorbers are clouds and rain, where the latter can attenuate signals 30 dB or more. Rain is a major absorber for centimeter-wavelength satellite dishes, partly in the atmosphere and partly as the rain accumulates on the antennas. At

longer wavelengths most systems have enough sensitivity to tolerate such attenuation. In comparison, clouds are usually not a problem except for through-the-air optical communication systems.

Atmospheric refraction is dominated by water vapor at radio wavelengths and by atmospheric density at optical wavelengths. These effects are not trivial. The radio sun can appear to set almost one solar diameter later on a very humid summer day (the sun emits strong radio waves too), and weak scattering from inhomogeneities in atmospheric humidity was once used as a major long-distance radio communications technique that avoided reliance on signals reflected from the ionosphere, as well as providing bandwidths of several GHz. Refraction by the ionosphere is even more extreme, and the angles of refraction can be computed using the properties of plasmas noted in Section 9.5.3 and Snell's law (9.2.26).

It is often convenient to model urban multipath and diffractive communications links by a power law other than r^{-2} . One common model is $r^{-3.8}$, which approximates the random weakening of signals by sequences of urban obstacles as signals 1.5-5 GHz propagate further. In any study of wireless communications systems propagation effects such as these must always be considered.

11.3.3 Thermal emission

A final effect impacting wireless communications systems is thermal noise arising from the environment, plus other forms of interference. Usually the thermal noise is considered interference too, but in radio astronomy and remote sensing it is the signal of interest. *Thermal noise* arises from electromagnetic radiation emitted by electrons colliding randomly with other particles in thermal equilibrium at temperature T. These collisions cause electrons to accelerate in random directions and therefore radiate. Thus every material object or medium radiates thermal noise provided that object or medium is coupled to the radiation field to any degree at all. Decoupled media perfectly reflect or transmit electromagnetic radiation without loss and are rare.

Thermal radiation propagating in a single-mode transmission line has intensity:

$$I[\text{W/Hz}] = \frac{hf}{e^{hf/kT} - 1} \cong kT \quad \text{for } hf \ll kT \quad ^{58} \quad (\text{thermal intensity}) \quad (11.3.1)$$

Because there is a one-to-one relationship between intensity I and the corresponding brightness temperature T, the *brightness temperature* $T[\text{K}] = I/k$ often replaces I because of its more natural physical significance. T is the temperature of a matched load ($R = Z_0$) that would naturally radiate the same intensity $I = kT$ Watts/Hz for $hf \ll kT$. This *Rayleigh-Jeans approximation* for I is valid at temperatures T above 50K for all frequencies f below ~100 GHz.

Thus the Thevenin equivalent circuit of a resistor at temperature T includes a voltage source producing a generally observable gaussian white voltage $v_{Th}(t)$ called *Johnson noise*. This

⁵⁸ $e^\delta = 1 + \delta + \delta^2/2! + \dots$ for $\delta \ll 1$.

source voltage $v_{Th}(t)$ radiates kTB [W] down a matched transmission line within the bandwidth B [Hz]. This Johnson noise voltage $v_{Th}(t)$ also divides across the Thevenin resistance R and its matched load $Z_o = R$ to produce the propagating line voltage $v_+(t, z=0) = v_{Th}/2$. But the radiated power is:

$$P_+ = \frac{\langle v_+^2 \rangle}{Z_o} = \frac{\langle (v_{Th}/2)^2 \rangle}{Z_o} = kTB \text{ [W]} \quad (\text{thermal noise power}) \quad (11.3.2)$$

Therefore within bandwidth B the root-mean-square open-circuit thermal voltage v_{Thrms} across a resistor R at temperature T is:

$$v_{Thrms} = \sqrt{4kTB\Delta f} \text{ [V]} \quad (\text{Johnson noise}) \quad (11.3.3)$$

A TEM line of impedance Z_o does not add any Johnson noise to that of the resistor if the line is lossless and therefore decoupled from the radiation.

Any antenna matched to its TEM transmission line therefore receives thermal noise power $kT_A B$ [W] from the environment, where T_A is defined as the *antenna temperature*. T_A is the gain-weighted average of the brightness temperature T_B of the environment over 4π steradians:

$$T_A = \frac{1}{4\pi} \int_{4\pi} T_B(\theta, \phi) G(\theta, \phi) d\Omega \quad (\text{antenna temperature}) \quad (11.3.4)$$

If the entire field of view has brightness temperature $T_B = T_o$, and if the antenna is lossless so that $G(\theta, \phi) = D(\theta, \phi)$, then $T_A = T_o$ since $\int_{4\pi} D(\theta, \phi) d\Omega = 4\pi$ (10.3.3).

11.3.4 Radio astronomy and remote sensing

An antenna looking down at the earth sees a brightness temperature T_B , which is the sum of thermal radiation emitted by the earth plus downward propagating power that is then reflected from the same surface: $T_B = \xi T + |\Gamma|^2 T_B'$, where the emissivity of the earth $\xi = 1 - |\Gamma|^2$, Γ is the wave reflection coefficient of the earth, and T_B' is the brightness temperature of the radiation reflected from the earth into the antenna beam. The radiation from space at microwave frequencies has a brightness temperature near 2.7K arising from the "big bang" that occurred at the birth of the universe, and reaches temperatures over 7000K in the direction of the sun and certain astronomical objects, depending on frequency. The science of *radio astronomy* involves the study of such celestial radio waves.

The emissivity $1 - |\Gamma|^2$ of the terrestrial surface is typically 0.85-0.98 over land and $\lesssim 0.3$ over ocean. Since most communications antennas point horizontally, about half their beam intercepts the earth ($\sim 260K$) and half intercepts space ($\sim 4K$ at microwave frequencies), so the thermal noise from the environment typically adds $\sim 132K$ to the antenna temperature and total system noise.

The study of natural radio, infrared, and visible emission from the earth is called *remote sensing*, although one can also remotely sense biological, manufacturing, and other systems. Today many satellites in polar and geostationary orbits routinely observe the earth at tens to thousands of wavelengths across the radio and optical spectrum for meteorological and other geophysical purposes. For example, a satellite observing in the opaque 53-67 GHz oxygen resonance band can not see much lower than 70 km altitude at the very centers of the strongest spectral lines, and therefore those channels observe the temperature of the atmosphere at those high altitudes. At nearby frequencies where the atmosphere is more transparent these sensors see the air temperatures at lower altitudes. Combinations of such observations yield the temperature profile of the atmosphere all over the globe, enabling better numerical weather predictions. Channels near the centers of water vapor, ozone, and other spectral lines can similarly measure their abundance and altitude profiles for similar purposes. Channels in the more transparent bands see closer to the terrestrial surface and permit estimates to be made of rain rate, surface winds, soil moisture, and other parameters.

Communications, radioastronomy, and remote sensing systems all receive non-thermal *radio interference* as well. Man-made interference comes from other transmitters in the same or nearby bands, automobiles, microwave ovens, motors, power supplies, corona around power lines, and other electrical devices. Each unshielded wire in any electrical device is a small antenna that radiates. For example, computers can emit highly structured signals that reveal the state of the computation and, in special cases, even the contents of registers. Poorly shielded power supplies often radiate at very high harmonics of their fundamental operating frequencies. Fortunately, regulations increasingly restrict radio emissions from modern electrical and electronic systems. Natural non-thermal emission arises from lightning, solar bursts, the planet Jupiter, and other sources.

11.4 Applications

11.4.1 Wireless communications systems

Section 11.4.1 introduces simple communications systems without using Maxwell's equations and Section 11.4.2 then discusses radar and lidar systems used for surveillance and research. Optical communications is deferred to Chapter 12, while the design, transformation, and switching of the communications signals themselves are issues left to other texts.

Wireless communications systems have a long history, beginning with wireless telegraph systems installed several years after Hertz's laboratory demonstrations of wireless links late in the nineteenth century. These systems typically used line-of-sight propagation paths, and sometimes inter-continental ionospheric reflections. Telephone, radio, and television systems followed. In the mid-twentieth century, the longer interstate and international wireless links were almost entirely replaced by more capable and reliable coaxial cables and multi-hop microwave links. These were soon supplemented by satellite links typically operating at frequencies up to ~14 GHz; today frequencies up to ~100 GHz are used. At century's end, these longer microwave links were then largely replaced again, this time by optical fibers with bandwidths of Terahertz. At the same time many of the shorter links are being replaced or supplemented by wireless cellular technology, which was made practical by the development of

inexpensive r.f. integrated circuits. Each technical advance markedly boosted capacity and market penetration, and generally increased performance and user mobility while reducing costs.

Most U.S. homes and offices are currently served by twisted pairs of telephone wires, each capable of conveying ~50 kbs - 1.5 Mbps, although coaxial cables, satellite links, and wireless services are making significant inroads. The most common wireless services currently include cell phones, wireless phones (within a home or office), wireless internet connections, wireless intra-home and intra-office connections, walkie-talkies (dedicated mobile links), satellite links, microwave tower links, and many specialized variations designed for private or military use. In addition, optical or microwave line-of-sight links between buildings offer instant broadband connectivity for the “last mile” to some users; the last mile accounts for a significant fraction of all installed plant cost. Weather generally restricts optical links to very short hops or to weather-independent optical fibers. Specialized wireless medical devices, such as RF links to video cameras inside swallowed pills, are also being developed.

Broadcast services now include AM radio near 1 MHz, FM radio near 100 MHz and higher frequencies, TV in several bands between 50 and 600 MHz for local over-the-air service, and TV and radio delivered by satellite at ~4, ~12, and ~20 GHz. Shortwave radio below ~30 MHz also offers global international broadcasts dependent upon ionospheric conditions, and is widely used by radio hams for long-distance communications.

Wireless services are so widespread today that we may take them for granted, forgetting that a few generations ago the very concept of communicating by invisible silent radio waves was considered magic. Despite the wide range of services already in use, it is reasonable to assume that over the next few decades numerous other wireless technologies and services will be developed by today’s engineering students.

Communications systems convey information between two or more nodes, usually via wires, wireless means, or optical fibers. After a brief discussion relating signaling rates (bits per second) to the signal power required at the wireless receiver, this section discusses in general terms the launching, propagation, and reception of electromagnetic signals and messages in wired and wireless systems.

Information is typically measured in bits. One *bit of information* is the information content of a single yes-no decision, where each outcome is equally likely. A string of M binary digits (equiprobable 0’s or 1’s) conveys M bits of information. An analog signal measured with an accuracy of one part in 2^M also conveys M bits because a unique M-bit binary number corresponds to each discernable analog value. Thus both analog and digital signals can be characterized in terms of the bits of information they convey. All wireless receivers require that the energy received per bit exceed a rough minimum of $w_0 \cong 10^{-20}$ Joules/bit, although most practical systems are orders of magnitude less sensitive.⁵⁹

⁵⁹ Most good communications systems can operate with acceptable probabilities of error if $E_b/N_0 \gg 10$, where E_b is the energy per bit and $N_0 = kT$ is the noise power density [$W Hz^{-1}$] = [J]. Boltzmann's constant $k \cong 1.38 \times 10^{-23}$ [$J ^\circ K^{-1}$], and T is the system noise temperature, which might approximate 100K in a good system at RF frequencies. Thus the nominal minimum energy E_b required to detect each bit of information is $\sim 10N_0 \cong 10^{-20}$ [J].

To convey N bits per second [b s^{-1}] of information therefore requires that at least $\sim Nw_o$ watts [W] be intercepted by the receiver, and that substantially more power be transmitted. Note that $[\text{W}] = [\text{J s}^{-1}] = [\text{J b}^{-1}][\text{b s}^{-1}]$. Wireless communications is practical because so little power P_r is actually required at the receiver. For example, to communicate 100 megabits per second (Mbps) requires as little as one picowatt (10^{-12}W) at the receiver if $w_o = 10^{-20}$; that is, we require $P_r > Nw_o \cong 10^8 \times 10^{-20} = 10^{-12}$ [W].

It is fortunate that radio receivers are so sensitive, because only a tiny fraction of the transmitted power usually reaches them. In most cases the path loss between transmitter and receiver is primarily geometric; the radiation travels in straight lines away from the transmitting antenna with an intensity I [W m^{-2}] that grows weaker with distance r as r^{-2} . For example, if the transmitter is isotropic and radiates its power P_t equally in all 4π directions, then $I(\theta, \phi, r) = P_t/4\pi r^2$ [W m^{-2}]. The power P_r intercepted by the receiving antenna is proportional to the incident wave intensity $I(\theta, \phi)$ and the receiving *antenna effective area* $A(\theta, \phi)$ [m^2], or “capture cross-section”, where the power P_r received from a plane wave incident from direction θ, ϕ is:

$$P_r = I(\theta, \phi, r) A(\theta, \phi) \quad [\text{W}] \quad (\text{antenna gain}) \quad (11.4.1)$$

The power received from an isotropic transmitting antenna is therefore $P_r = (P_t/4\pi r^2)A(\theta, \phi)$, so in this special case the line-of-sight path loss between transmitter and receiver is $P_r/P_t = A(\theta, \phi)/4\pi r^2$, or that fractional area of a sphere of radius r represented by the receiving antenna cross-section A . Sometimes additional propagation losses due to rain, gaseous absorption, or scattering must be recognized too, as discussed in Section 11.3.2.

In general, however, the transmitting antenna is not isotropic, but is designed to radiate power preferentially in the direction of the receivers. We define *antenna gain* $G(\theta, \phi)$, often called “gain over isotropic”, as the ratio of the intensity $I(\theta, \phi, r)$ [W m^{-2}] of waves transmitted in the direction θ, ϕ (spherical coordinates) at distance r , to the intensity that would be transmitted by an isotropic antenna. That is:

$$G(\theta, \phi) \equiv \frac{I(\theta, \phi, r)}{P_t/4\pi r^2} \quad (\text{antenna gain}) \quad (11.4.2)$$

If the radiated power is conserved, then the integral of wave intensity over a spherical surface enclosing the antenna is independent of the sphere’s radius r . Therefore the angular distribution of power and $G(\theta, \phi)$ plotted in spherical coordinates behave much like a balloon that must push out somewhere when it is pushed inward somewhere else, as suggested in Figure 11.4.1. The maximum gain G_o often defines the z axis and is called the on-axis gain. The angular width θ_B of the main beam at the half-power points where $G(\theta, \phi) \cong G_o/2$ is called the *antenna beamwidth* or “half-power beamwidth”. Other local peaks in gain are called *sidelobes*, and those sidelobes behind the antenna are often called backlobes. Angles at which the gain is nearly zero are called nulls.

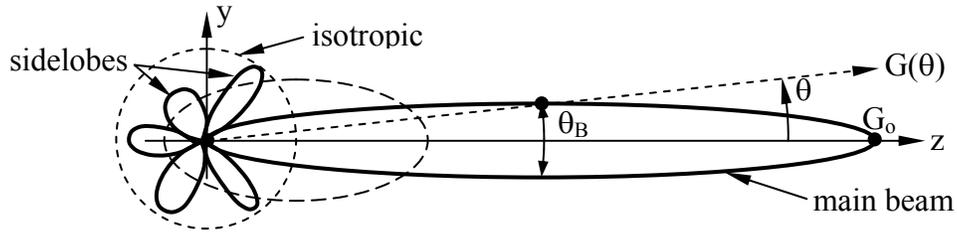


Figure 11.4.1 Isotropic and directive antenna gain patterns.

Antennas with $G(\theta, \phi) > 1$ generally focus their radiated energy by using lenses, mirrors, or multiple radiators phased so their radiated contributions add in phase in the desired direction, and largely cancel otherwise. Typical gains for most wire antennas range from ~ 1.5 to ~ 100 , and large aperture antennas such as parabolic dishes or optical systems can have gains of 10^8 or more. The directionality or gain of a mirror or any antenna system is generally the same whether it is transmitting or receiving.⁶⁰ The fundamentals of transmission and reception are presented in more detail in Section 10.3.1.

Consider the following typical example. A television station transmits 100 kW at ~ 100 MHz toward the horizon with an antenna gain of ~ 10 . Because the gain is much greater than unity in the desired horizontal direction, it is therefore less than unity for most other downward and upward directions where users are either nearby or absent. The intensity I [W m^{-2}] sensed by users on the horizon at 100-km range follows from (11.4.2):

$$I \cong G \frac{P_t}{4\pi r^2} = 10 \times \frac{10^5}{4\pi(10^5)^2} \cong 10^{-5} \text{ [W/m}^2\text{]} \quad (11.4.3)$$

Whether this intensity is sufficient depends on the properties of the receiving antenna and receiver. For the example of Equation (11.4.3), a typical TV antenna with an effective area $A \cong 2$ [m^2] would capture $IA \cong 10^{-5}[\text{W/m}^2] \times 2[\text{m}^2] = 2 \times 10^{-5}$ [W]. If the received power is $\langle v^2(t) \rangle / R \cong 2 \times 10^{-5}$ [W], and the receiver has an input impedance R of 100 ohms, then the root-mean-square (rms) voltage $v_{\text{rms}} \equiv \langle v^2(t) \rangle^{0.5}$ would be $(0.002)^{0.5} \cong 14$ mv, much larger than typical noise levels in TV receivers (~ 10 μv).⁶¹

⁶⁰ The degree of focus is the same whether the waves are transmitted or received. That is, if we reverse the direction of time for a valid electromagnetic wave solution to Maxwell's equations, the result is also a valid solution if the system is lossless and reciprocal. Reciprocity requires that the complex matrices characterizing $\underline{\epsilon}$, $\underline{\mu}$, and $\underline{\sigma}$ near the antenna equal their own transposes; this excludes magnetized plasmas such as the ionosphere, and magnetized ferrites, as discussed further in Section 10.3.4.

⁶¹ Typical TV receivers might have a superimposed noise voltage of power $N = kTB$ [W], where the system noise temperature T might be $\sim 10^4$ K (much is interference), Boltzmann's constant $k = 1.38 \times 10^{-23}$, and B is bandwidth [Hz]. $B \cong 6$ MHz for over-the-air television. Therefore $N \cong 1.38 \times 10^{-23} \times 10^4 \times 6 \times 10^6 \cong 8 \times 10^{-13}$ watts, and a good TV signal-to-noise ratio S/N of $\sim 10^4$ requires only $\sim 8 \times 10^{-9}$ watts of signal S . Since $N \cong n_{\text{rms}}^2 / R$, the rms noise voltage $\cong (NR)^{0.5}$, or ~ 10 μv if the receiver input impedance $R = 100$ ohms.

Because most antennas are equally focused whether they are receiving or transmitting, their effective area $A(\theta,\phi)$ and gain $G(\theta,\phi)$ are closely related:

$$G(\theta,\phi) = \frac{4\pi}{\lambda^2} A(\theta,\phi) \quad (11.4.4)$$

Therefore the on-axis gain $G_o = 4\pi A_o/\lambda^2$. This relation (11.4.4) was proven for a short dipole antenna in Section 10.3.3 and proven for other types of antenna in Section 10.3.4, although the proof is not necessary here. This relation is often useful in estimating the peak gain of aperture antennas like parabolic mirrors or lenses because their peak effective area A_o often approaches their physical cross-section A_p within a factor of two; typically $A_o \cong 0.6 A_p$. This approximation does not apply to wire antennas, however. Thus we can easily estimate the on-axis gain of such aperture antennas:

$$G_o = 0.6 \times \frac{4\pi}{\lambda^2} A_o \quad (11.4.5)$$

Combining (11.4.1) and (11.4.3) yields the *link expression* for received power:

$$P_r = G_t \frac{P_t}{4\pi r^2} A_r [\text{W}] \quad (\text{link expression}) \quad (11.4.6)$$

where G_t is the gain of the transmitting antenna and A_r is the effective area of the receiving antenna. The data rate R associated with this received power is : $R = P_r/E_b$ [bits s^{-1}].

A second example illustrates how a communications system might work. Consider a *geosynchronous communications satellite*⁶² transmitting 12-GHz high-definition television (HDTV) signals at 20 Mbps to homes with 1-meter dishes, and assume the satellite antenna spreads its power P_t roughly equally over the eastern United States, say $3 \times 10^6 \text{ km}^2$. Then the intensity of the waves falling on the U.S. is: $I \cong P_t/(3 \times 10^{12})$ [W m^{-2}], and the power P_r received by an antenna with effective area $A_o \cong 0.6$ [m²] is:

$$P_r = A_o I = 0.6 \frac{P_t}{3 \times 10^{12}} = 2 \times 10^{-13} P_t [\text{W}] \quad (11.4.7)$$

If $E_b = 10^{-20}$ Joules per bit suffices, then an $R = 20$ -Mbps HDTV signal requires:

$$P_r = E_b R = 10^{-20} \times (2 \times 10^7) = 2 \times 10^{-13} [\text{W}] \quad (11.4.8)$$

The equality of the right-hand parts of (11.4.7) and (11.4.8) reveals that one watt of transmitter power P_t in this satellite could send a digital HDTV signal to all the homes and businesses in the

⁶² A satellite approximately 35,000 km above the equator circles the earth in 24 hours at the same rate at which the earth rotates, and therefore can remain effectively stationary in the sky as a communications terminal serving continental areas. Such satellites are called “geostationary” or “geosynchronous”.

eastern U.S. Since a 20-dB margin⁶³ for rain attenuation, noisy receivers, smaller or poorly pointed home antennas, etc. is desirable, 100-watt transmitters might be used in practice.

We can also estimate the physical area A_p of the aperture antenna on the satellite. If we know P_t and I at the earth, then we can determine the satellite gain G using $I = GP_t/4\pi r^2$ (11.4.2), where $r \cong 40,000$ km in the northern U.S; here we have $I \cong 3.3 \times 10^{-12}$ when $P_t = 1$ watt. The wavelength λ at 12 GHz is 2.5 cm ($\lambda = c/f$). But $A_p \cong 1.5A_o$, where A_o is related to G by (11.4.4). Therefore we obtain the reasonable result that a 2.5-meter diameter parabolic dish on the satellite should suffice:

$$A_p \cong 1.5A_o = (1.5\lambda^2/4\pi)G = (1.5\lambda^2/4\pi)(4\pi r^2 I/P_t) \cong 5 \text{ [m}^2\text{]} \quad (11.4.9)$$

The same result could have been obtained by determining the angular extent of the U.S. coverage area as seen from the satellite and then, as discussed in Section 11.1.2, determining what diameter antenna would have a diffraction pattern with that same beamwidth.

Thus we can design digital communications systems for a *data rate* R [b s^{-1}] if we know the range r , wavelength λ , and receiver sensitivity (Joules required per bit). For analog systems we also need to know the desired signal-to-noise ratio (SNR) at the receiver and the noise power N . Table 11.4.1 lists typical data rates R for various applications, and Table 11.4.2 lists typical SNR values required for various types of analog signal.

Table 11.4.1 Digital data rates for typical applications and source coding techniques⁶⁴.

Applications	Data rate R after source coding	R before coding
Intelligible voice	$\gg 1200$ bps	~ 64 kbps
Good voice	$\gg 4.8 - 9.6$ kbps	~ 128 kbps
Excellent voice	$\gg 16$ kbps	~ 256 kbps
CD-quality music	2×128 kbps	~ 1.4 Mbps
Talking head, lip read	$\gg 64$ kbps	~ 1.4 Mbps
Good video conference	$\gg 128-384$ kbps	~ 12 Mbps
VHS video	$\gg 1.5$ Mbps	~ 30 Mbps
NTSC studio video	$\gg 6$ Mbps	~ 256 Mbps
HDTV video	$\gg 18$ Mbps	~ 1 Gbps

⁶³ Decibels (dB) are defined for a ratio R such that $\text{dB} = 10 \log_{10} R$ and $R = 10^{(\text{dB})/10}$; thus $20 \text{ dB} \rightarrow R = 100$.

⁶⁴ Source coding reduces the number of bits to be communicated by removing redundancies and information not needed by the user. The table lists typical data rates before and after coding.

Table 11.4.2 Signal-to-noise ratios⁶⁵ for typical wireless applications.

Application	Desired SNR $\geq \sim$
Digital communications at ~ 1 bps/Hz	10 dB E_b/N_o
Digital communications at $> \sim 4$ bps/Hz	20 dB E_b/N_o
Amplitude modulated (AM) signals (20 kHz typical)	30 dB S/N
Frequency modulated (FM) signals (100 kHz typical)	20 dB S/N
NTSC broadcast television (6 MHz typical)	35 dB S/N
CD-quality music (55-dB SNR + 40-dB dynamic range)	95 dB S/N

Example 11.4A

A parabolic reflector antenna of 2-meter diameter transmits $P_t = 10$ watts at 3 GHz from beyond the edge of the solar system ($R \cong 10^{10}$ km) to a similar antenna on earth of 50-m diameter at a maximum data rate N bits/sec. What is N if the receiver requires 10^{-20} Joules bit⁻¹?

Solution: Recall that the on-axis effective area A of a circular aperture antenna equals ~ 0.6 times its physical area (πr^2), and it has gain $G = 4\pi A/\lambda^2 = (2\pi r/\lambda)^2$. The received power is $P_{rec} = P_t G_t A_r / 4\pi R^2$ (11.4.6); therefore:

$$R \cong P_{rec}/E_b = \left[P_t (0.6)^2 (2\pi r_t/\lambda)^2 \pi r_r^2 \right] / \left[4\pi R^2 E_b \right]$$

$$= \left[10 \times (0.6)^2 (2\pi \times 1/0.1)^2 \pi 25^2 \right] / \left[4\pi 10^{26} \times 10^{-20} \right] \cong 2.2 \text{ bps}$$

11.4.2 Radar and lidar

Radar (RADio Direction and Range finding) and *lidar* (LIght Direction and Range finding) systems transmit signals toward targets of interest and receive echoes. They typically determine: 1) target distance using the round-trip propagation delay, 2) target direction using echo strength relative to antenna orientation, 3) target radial velocity using the observed Doppler shift, and 4) target size or scattering properties using the maximum echo strength. Figure 11.4.2 illustrates the most common radar configuration.

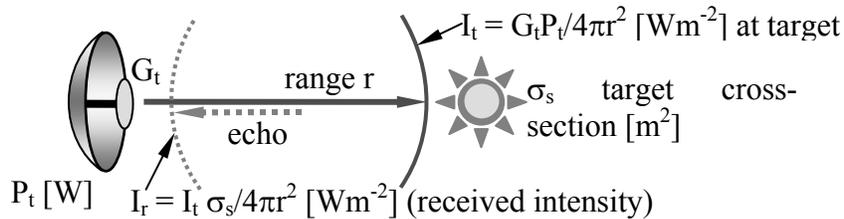


Figure 11.4.2 Radar signals reflected from a target.

⁶⁵ For digital signals the dimensionless signal-to-noise ratio (SNR) given here is the energy-per-bit E_b divided by the noise power density N_o [$W Hz^{-1}$], where $N_o = kT$ and T is the noise temperature, say $100-10^4 K$ typically. For analog signals, S and N are the total signal and noise powers, respectively, where $N = kTB$ and B is signal bandwidth [Hz].

To compute the received power, we first compute the intensity I_t of radiation at the target at range r for a transmitter power and antenna gain of P_t and G , respectively:

$$I_t = G \frac{P_t}{4\pi r^2} \text{ [W/m}^2\text{]} \quad (\text{intensity at target}) \quad (11.4.10)$$

The target then scatters this radiation in some pattern and absorbs the rest. Some of this scattered radiation reaches the receiver with intensity I_r , where:

$$I_r = I_t \frac{\sigma_s}{4\pi r^2} \text{ [W/m}^2\text{]} \quad (\text{intensity at radar}) \quad (11.4.11)$$

where σ_s is the *scattering cross-section* of the target and is defined by (11.4.11). That is, σ_s is the capture cross-section [m²] at the target that would produce I_r if the target scattered incident radiation isotropically. Thus targets that preferentially scatter radiation toward the transmitter can have scattering cross-sections substantially larger than their physical cross-sections.

The received power P_r is then simply $I_r A_r$ [W], where A_r is the effective area of the receiving antenna. That is:

$$P_r = I_r A_r = \frac{I_t \sigma_s}{4\pi r^2} A_r = G \frac{P_t \sigma_s}{(4\pi r^2)^2} A_r \quad (11.4.12)$$

$$P_r = P_t \frac{\sigma_s}{4\pi} \left(\frac{G\lambda}{4\pi r^2} \right)^2 \text{ [W]} \quad (\text{radar equation}) \quad (11.4.13)$$

where we used $A_r = G\lambda^2/4\pi$, and where (11.4.13) is often called the *radar equation*. The dependence of received power on the fourth power of range and the square of antenna gain often control radar system design.

Atmospheric attenuation is often included in the radar equation by means of a round-trip attenuation factor $e^{-2\alpha r}$, where α is the average atmospheric attenuation coefficient (m⁻¹) and r is range. Atmospheric attenuation is discussed in Section 11.3.2 and below 200 GHz is due principally to oxygen, water vapor, and rain; it is usually not important below ~3 GHz. Oxygen absorption occurs primarily in the lowest 10 km of the atmosphere ~50-70 GHz and near 118 GHz, water vapor absorption occurs primarily in the lowest 3 km of the atmosphere above ~10 GHz, and rain absorption occurs up to ~15 km in the largest rain cells above ~3 GHz.

Lidar systems also obey the radar equation, but aerosol scattering by clouds, haze, or smoke becomes more of a concern. Also the phase fronts of optical beams are more easily disturbed by refractive inhomogeneities in the atmosphere that can modulate received echoes on time scales of milliseconds with random fading of ten dB or more.

A simple example illustrates use of the radar equation (11.4.13). Suppose we wish to know the range r at which we can detect dangerous asteroids having diameters over ~300m that are

approaching the earth. Assume the receiver has additive noise characterized by the system noise temperature T_s , and that the radar bandwidth is one Hertz because the received sinusoid will be averaged for approximately one second. Detectable radar echos must have $P_r > kT_s B$ [W], where k is Boltzmann's constant ($k = 1.38 \times 10^{-23}$) and B is the system bandwidth (~ 1 Hz); this implies $P_r \cong 1.4 \times 10^{-23} T_s$ watts. We can estimate σ_s for a 300-meter asteroid by assuming it reflects roughly as well as the earth, say fifteen percent, and that the scattering is roughly isotropic; then $\sigma_s \cong 10^4$ [m²]. If we further assume our radar is using near state-of-the-art components, then we might have $P_t \cong 1$ Mw, $G_t \cong 10^8$, $\lambda = 0.1$ m, and $T_s \cong 10$ K. The radar equation then yields:

$$r \cong \left[P_t \sigma_s (G_t \lambda)^2 / (4\pi)^3 P_r \right]^{0.25} \cong 5 \times 10^7 \text{ km} \quad (11.4.14)$$

This range is about one-third of the distance to the sun and would provide about 2-3 weeks warning.

Optical systems with a large aperture area A might perform this task better because their antenna gain $G = A4\pi/\lambda^2$, and λ for lidar is typically 10^{-5} that of a common radar. For antennas of the same physical aperture and transmitter power, 1-micron lidar has an advantage over 10-cm radar of $\sim 10^{10}$ in P_r/P_t .

Radar suffers because of its dependence on the fourth power of range for targets smaller than the antenna beamwidth. If the radar can place all of its transmitted energy on target, then it suffers only the range-squared loss of the return path. The ability of lidar systems to strongly focus their transmitting beam totally onto a small target often enables their operation in the highly advantageous r^{-2} regime rather than r^{-4} .

Equations (11.4.13) and (11.4.14) can easily be revised for the case where all the radar energy intercepts the target. The radar equation then becomes:

$$P_r = P_t R G (\lambda / 4\pi r)^2 \text{ [W]} \quad (11.4.15)$$

where the target retro-reflectivity R is defined by (11.4.15) and is the dimensionless ratio of back-scattered radiation intensity at the radar to what would be back scattered if the radiation were scattered isotropically by the target. For the same assumptions used before, asteroids could be detected at a range r of $\sim 3 \times 10^{12}$ km if $R \cong 0.2$, a typical value for icy rock. The implied detection distance is now dramatically farther than before, and reaches outside our solar system. However, the requirement that the entire radar beam hit the asteroid would be essentially impossible even for the very best optical systems, so this approach to boosting detection range is usually not practical for probing small distant objects.

Radar systems often use phased arrays of antenna elements, as discussed in Section 10.4, to focus their energy on small spots or to look in more than one direction at once. In fact a single moving radar system, on an airplane for example, can coherently receive sequential reflected radar pulses and digitally reassemble the signal over some time period so as to synthesize the equivalent of a phased array antenna that is far larger than the physical antenna. That is, a small

receiving antenna can be moved over a much larger area A , and by combining its received signals from different locations in a phase-coherent way, can provide the superior angular resolution associated with area A . This is called *synthetic aperture radar* (SAR) and is not discussed further here.

Example 11.4B

A radar with 1-GHz bandwidth and 40-dB gain at 10 GHz views the sun, which has angular diameter 0.5 degrees and brightness temperature $T_B = 10,000\text{K}$. Roughly what is the antenna temperature T_A and the power received by the radar from the sun if we ignore any radar reflections?

Solution: The power received is the intensity at the antenna port I [W/Hz] times the bandwidth B [Hz], where $I \cong kT_A$ (11.3.1), and T_A is the antenna temperature given by the integral in (11.3.4). This integral is trivial if $G(\theta, \phi)$ is nearly constant over the solid angle Ω_S of the sun; then $T_A \cong G_0 T_B \Omega_S / 4\pi$. Constant gain across the sun requires the antenna beamwidth $\theta_B \gg 0.5$ degrees. We can roughly estimate θ_B by approximating the antenna gain as a constant G_0 over solid angle Ω_B , and zero elsewhere; then (10.3.3) yields $\int_{4\pi} G(\theta, \phi) d\Omega = 4\pi = G_0 \Omega_B = 10^4 \Omega_B$. Therefore $\Omega_B \cong 4\pi / 10^4 \cong \pi(\theta_B/2)^2$, and $\theta_B \cong 0.04$ radians $\cong 2.3$ degrees, which is marginally greater than the solar diameter required for use of the approximation $\theta_B \gg 0.5$ in a rough estimate. It follows that $T_A \cong G_0 T_B \Omega_S / 4\pi = 10^4 \times 10^4 \times \pi(\theta_S/2)^2 / 4\pi \cong 480$ degrees Kelvin, somewhat larger than the noise temperature of good receivers. The power received is $kT_A B \cong 1.38 \times 10^{-23} \times 480 \times 10^9 \cong 6.6 \times 10^{-12}$ watts. This is a slight overestimate because the gain is actually slightly less than G_0 at the solar limb.

Example 11.4C

What is the scattering cross-section σ_s of a small distant flat plate of area F oriented so as to reflect incident radiation directly back toward the transmitter?

Solution: The radar equation (11.4.12) relates the transmitted power P_t to that received, P_r , in terms of σ_s . A similar relation can be derived by treating the power reflected from the flat plate as though it came from an aperture uniformly illuminated with intensity $P_t G_t / 4\pi r^2$ [W m⁻²]. The power P_r received by the radar is then the power radiated by the flat-plate aperture, $F P_t G_t / 4\pi r^2$ [W], inserted as P_t into the link expression (11.4.6): $P_r = (F P_t G_t / 4\pi r^2) G_f A_r / 4\pi r^2$. The gain of the flat plate aperture is $G_f = F 4\pi / \lambda^2$, and $A_r = G \lambda^2 / 4\pi$, so $G_f A_r = GF$. Equating P_r in this expression to that in the radar equation yields: $P_t \sigma_s (G \lambda / 4\pi r^2)^2 / 4\pi = (F P_t G_t / 4\pi r^2) GF / 4\pi r^2$, so $\sigma_s = F(4\pi F / \lambda^2)$. Note $\sigma_s \gg F$ if $F \gg \lambda^2 / 4\pi$. *Corner reflectors* (three flat plates at right angles intersecting so as to form one corner of a cube) reflect plane waves directly back toward their source if the waves impact the concave portion of the reflector from any angle. Therefore the corner reflector becomes a very area-efficient radar target if its total projected area F is larger than λ^2 .

MIT OpenCourseWare
<http://ocw.mit.edu>

6.013 Electromagnetics and Applications
Spring 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.