Chapter 10: Antennas and Radiation

10.1 Radiation from charges and currents

10.1.1 Introduction to antennas and radiation

An antenna is a device that couples currents to electromagnetic waves for purposes of radiation or reception. The process by which antennas radiate can be easily understood in terms of the way in which accelerating charged particles or time-varying currents radiate, which is discussed in Section 10.1. The expressions for radiated electromagnetic fields derived in Section 10.1.4 are simple extensions of those derived in Sections 10.1.2 and 10.1.3 for the fields produced by static charges and currents, respectively.

Using the basic expressions for radiation derived in Section 10.1, simple short dipole antennas are shown in Section 10.2 to have stable directional properties far from the antenna (the antenna far field), and different directional properties closer than $\sim\lambda/2\pi$ (the antenna near field). In Section 10.3 these properties are related to basic metrics that characterize each antenna, such as gain, effective area, and impedance. These metrics are then related to the performance of various communications systems. Antenna arrays are discussed in Section 10.4, followed by aperture and more complicated wire antennas in Sections 11.1 and 11.2, respectively.

10.1.2 Electric fields around static charges

One simple way to generate electromagnetic waves is to vibrate electric charges, creating time-varying current. The equation characterizing this radiation is very similar to that characterizing the electric fields produced by a single static charge, which is developed below. Section 10.1.3 extends this result to magnetic fields produced by moving charges.

Faraday's and Gauss's laws for static charges in vacuum are:

$$\nabla \times \bar{E} = 0 \tag{10.1.1}$$

$$\nabla \bullet \varepsilon_o \bar{E} = \rho \tag{10.1.2}$$

Since the curl of $\bar{E}$ is zero, $\bar{E}$ can be the gradient of any arbitrary scalar function $\Phi(\bar{r})$ and still satisfy (10.1.1). That is:

$$\bar{E}(\bar{r}) = -\nabla\Phi(\bar{r}) \tag{10.1.3}$$

where Φ is the *scalar electric potential* and is in units of *Volts*. The negative sign is consistent with $\bar{E}$ pointing away from regions of high potential and toward lower potentials. Note that (10.1.3) satisfies (10.1.1) because $\nabla \times (-\nabla\Phi) \equiv 0$ is an identity, and that a simple three-dimensional scalar field Φ fully characterizes the three-dimensional vector electric field $\bar{E}(\bar{r})$. It

is therefore often easiest to find the electric potential $\Phi(\bar{r})$ before computing the electric field produced by static source charges.

If the charge q [Coulombs] is spherically symmetric, both Φ and $\bar{E}$ must also be spherically symmetric. The only way a vector field can be spherically symmetric is for it to be directed radially, so:

$$\bar{E} = \hat{r}E_r(r) \tag{10.1.4}$$

where r is the radius from the origin where the charge is centered and $E_r(r)$ is the radial field. We can now relate $\bar{E}$ to q by applying Gauss's divergence theorem (2.4.6) to the volume integral of Gauss's law (10.1.2):

$$\begin{aligned} \iiint_V \left(\nabla \bullet \varepsilon_o \bar{E}\right) dv = \iiint_V \rho \; dv = q = \iint_A \varepsilon_o \bar{E} \bullet \hat{n} \; da \\ = \iint_A \varepsilon_o \hat{n} \bullet \hat{r}E_r(r) \; da = 4\pi r^2 \varepsilon_o E_r(r) \end{aligned} \tag{10.1.5}$$

Therefore the electric field produced by a charge q at the origin is:

$$\bar{E}(\bar{r}) = \hat{r}E_r(r) = \hat{r}q/4\pi\varepsilon_o r^2 = -\nabla\Phi \quad \left[\mathrm{Vm}^{-1}\right] \tag{10.1.6}$$

To find the associated scalar potential Φ we integrate (10.1.6) using the definition of the *gradient operator*:

$$\nabla\Phi \equiv \left(\hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z}\right)\Phi \qquad \text{(gradient in Cartesian coordinates)} \tag{10.1.7}$$

$$= \left[\hat{r}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial\theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right]\Phi \qquad \text{(gradient in spherical coordinates)} \tag{10.1.8}$$

Since the spherically symmetric potential Φ (10.1.6) is independent of θ and ϕ, it follows that $\partial/\partial\theta = \partial/\partial\phi = 0$ and Equation (10.1.8) becomes:

$$\nabla\Phi = \hat{r}\partial\Phi/\partial r \tag{10.1.9}$$

This mathematical simplification occurs only in spherical coordinates, not Cartesian. Substitution of (10.1.9) into (10.1.6), followed by integration of (10.1.6) with respect to radius r, yields:

$$\Phi(\bar{r}) = \int \left(q/4\pi\varepsilon_o r^2\right) dr = \Phi_o + q/4\pi\varepsilon_o r = q/\left(4\pi\varepsilon_o |\bar{r}|\right) \tag{10.1.10}$$

where we define as zero the electric potential Φ_o contributed by any charge infinitely far away.

The solution for the electric potential Φ due to charge q at some position $\bar{r}_q$ other than the origin follows from (10.1.10):

$$\Phi(\bar{r}) = q/\left(4\pi\varepsilon_o\left|\bar{r}-\bar{r}_q\right|\right) = q/\left(4\pi\varepsilon_o r_{pq}\right) \quad [V] \tag{10.1.11}$$

which can alternatively be written using subscripts p and q to refer to the locations $\bar{r}_p$ and $\bar{r}_q$ of the person (or observer) and the charge, respectively, and r_{pq} to refer to the distance $\left|\bar{r}_p - \bar{r}_q\right|$ between them.

If we replace the charge q with a charge density ρ_q in the infinitesimal volume dv, then we can integrate (10.1.11) over the source region to obtain the total static *electric potential* produced by an arbitrary charge distribution ρ_q:

$$\Phi_p = \iiint_{V_q}\left[\rho_q/\left(4\pi\varepsilon_o r_{pq}\right)\right]dv \quad [V] \qquad \text{(scalar Poisson integral)} \tag{10.1.12}$$

This integration to find Φ_p can be performed because Maxwell's equations are linear so that superposition applies. Thus we have a simple way to compute Φ_p and $\bar{E}$ for any arbitrary static charge density distribution ρ_q. This *scalar Poisson integral* for the potential function Φ is similar to that found for dynamic charge distributions in the next section. The integral (10.1.12) is also a solution to the *Poisson equation*:

$$\nabla^2\Phi = -\rho/\varepsilon_o \qquad \text{(Poisson equation)} \tag{10.1.13}$$

which follows from computing the divergence of Gauss's law:

$$\nabla\bullet\{\nabla\Phi = -\bar{E}\} \Rightarrow \nabla^2\Phi = -\nabla\bullet\bar{E} = -\rho/\varepsilon_o \tag{10.1.14}$$

Poisson's equation reduces to *Laplace's equation,* $\nabla^2\Phi = 0$, when $\rho = 0$.

10.1.3 Magnetic fields around static currents

Maxwell's equations governing static magnetic fields in vacuum are:

$$\nabla\times\bar{H} = \bar{J} \qquad \text{(static Ampere's law)} \tag{10.1.15}$$

$$\nabla\bullet\mu_o\bar{H} = 0 \qquad \text{(Gauss's law)} \tag{10.1.16}$$

Because the divergence of $\bar{H}$ is always zero, we can define the magnetic flux density in vacuum as being:

$$\overline{B} = \mu_o \overline{H} = \nabla \times \overline{A} \tag{10.1.17}$$

where $\overline{A}$ is defined as the *magnetic vector potential*, which is a vector analog to Φ. This very general expression for $\overline{B}$ always satisfies Gauss's law: $\nabla \bullet (\nabla \times \overline{A}) \equiv 0$.

Substituting (10.1.17) into Ampere's law (10.1.15) results in:

$$\nabla \times (\nabla \times \overline{A}) = \mu_o \bar{J} \tag{10.1.18}$$

This can be simplified using the vector identity:

$$\nabla \times (\nabla \times \overline{A}) \equiv \nabla (\nabla \bullet \overline{A}) - \nabla^2 \overline{A} \tag{10.1.19}$$

where we note that $\nabla \bullet \overline{A}$ is arbitrary and does not impact any of our prior equations; therefore we set it equal to zero. Then (10.1.18) becomes the *vector Poisson equation*:

$$\nabla^2 \overline{A} = -\mu_o \bar{J} \qquad \text{(vector Poisson equation)} \tag{10.1.20}$$

The three vector components of (10.1.20) are each scalar Poisson equations identical to (10.1.13) except for the constant, so the solution is nearly identical to (10.1.12) once the constants have been reconciled; this solution is:

$$\overline{A}_p = \oiiint_{V_q} \left[\mu_o \bar{J}_q / (4\pi r_{pq}) \right] dv \quad \left[V\ s\ m^{-1} \right] \tag{10.1.21}$$

Thus we have a simple way to compute $\overline{A}$ and therefore $\overline{B}$ for any arbitrary static current distribution $\bar{J}_q$.

10.1.4 Electromagnetic fields produced by dynamic charges

In the static case of Section 10.1.2 it was very helpful to define the potential functions $\overline{A}$ and Φ, and the time-dependent Maxwell's equations for vacuum permit us to do so again:

$$\nabla \times \overline{E} = -\partial \overline{B} / \partial t \qquad \text{(Faraday's law)} \tag{10.1.22}$$

$$\nabla \times \overline{H} = \bar{J} + \partial \overline{D} / \partial t \qquad \text{(Ampere's law)} \tag{10.1.23}$$

$$\nabla \bullet \overline{E} = \rho / \varepsilon_o \qquad \text{(Gauss's law)} \tag{10.1.24}$$

$$\nabla \bullet \overline{B} = 0 \qquad \text{(Gauss's law)} \tag{10.1.25}$$

Although the curl of $\overline{E}$ is no longer zero so that $\overline{E}$ no longer equals the gradient of some potential Φ, we can satisfy $\nabla \bullet \overline{B} = 0$ if we define a vector potential $\overline{A}$ such that:

$$\overline{B} = \nabla \times \overline{A} = \mu_o \overline{H} \tag{10.1.26}$$

This definition of $\overline{A}$ always satsifies Gauss's law: $\nabla \bullet (\nabla \times \overline{A}) \equiv 0$. Substituting $\nabla \times \overline{A}$ for $\overline{B}$ in Faraday's law yields:

$$\nabla \times \overline{E} = -\partial (\nabla \times \overline{A}) / \partial t \tag{10.1.27}$$

Rearranging terms yields:

$$\nabla \times \left(\overline{E} + \partial \overline{A} / \partial t \right) = 0 \tag{10.1.28}$$

which implies that the quantity $\left(\overline{E} + \partial \overline{A} / \partial t \right)$ can be the gradient of any potential function Φ:

$$\overline{E} + \partial \overline{A} / \partial t = -\nabla \Phi \tag{10.1.29}$$

$$\overline{E} = -\left(\partial \overline{A} / \partial t + \nabla \Phi \right) \tag{10.1.30}$$

Thus dynamic electric fields have two components—one due to the instantaneous value of $\Phi(t)$, and one proportional to the time derivative of $\overline{A}$.

We can now use the vector identity (10.1.19) to simplify Ampere's law after $(\nabla \times \overline{A}) / \mu_o$ replaces $\overline{H}$:

$$\nabla \times (\nabla \times \overline{A}) = \mu_o \left(\overline{J} + \partial \overline{D} / \partial t \right) = \nabla (\nabla \bullet \overline{A}) - \nabla^2 \overline{A} \tag{10.1.31}$$

By substituting (10.1.30) in (10.1.31) for $\overline{D} = \varepsilon_o \overline{E}$ and grouping terms we obtain:

$$\nabla^2 \overline{A} - \nabla \left(\nabla \bullet \overline{A} + \mu_o \varepsilon_o \partial \Phi / \partial t \right) - \mu_o \varepsilon_o \partial^2 \overline{A} / \partial t^2 = -\mu_o \overline{J} \tag{10.1.32}$$

In the earlier static case we let $\nabla \bullet \overline{A} = 0$ because specifying the curl of a vector field $(\overline{B} = \nabla \times \overline{A})$ does not constrain its divergence, which can be independently chosen[51]. Here we can let:

$$\nabla \bullet \overline{A} = -\mu_o \varepsilon_o \partial \Phi / \partial t \tag{10.1.33}$$

---

[51] Let $\overline{A} = \nabla \Phi + \nabla \times \overline{N}$; then $\nabla \times \overline{A} = \nabla \times (\nabla \times \overline{N})$ and $\nabla \bullet \overline{A} = \nabla^2 \Phi$, so $\nabla \times \overline{A}$ and $\nabla \bullet \overline{A}$ can be chosen independently simply by choosing $\overline{N}$ and Φ independently.

This reduces (10.1.32) to a simple equation by eliminating its second term, yielding:

$$\nabla^2 \overline{A} - \mu_o \varepsilon_o \partial^2 \overline{A} / \partial t^2 = -\mu_o \overline{J} \tag{10.1.34}$$

which is called the inhomogeneous vector *Helmholtz equation* (the homogeneous version has no source term on the right hand side; $\overline{J} = 0$). It is a wave equation for $\overline{A}$ driven by the current source $\overline{J}$.

A similar inhomogeneous wave equation relating the electric potential Φ to the charge distribution ρ can also be derived. Substituting (10.1.30) into Gauss's law (10.1.24) yields:

$$\nabla \bullet \overline{E} = -\nabla \bullet \left(\partial \overline{A} / \partial t + \nabla \Phi \right) = -\partial \left(\nabla \bullet \overline{A} \right) / \partial \tau - \nabla^2 \Phi \tag{10.1.35}$$

Replacing $\nabla \bullet \overline{A}$ using (10.1.33) then produces:

$$\nabla \bullet \overline{E} = \mu_o \varepsilon_o \partial^2 \Phi / \partial t^2 - \nabla^2 \Phi = \rho / \varepsilon_o \tag{10.1.36}$$

which is more commonly written as the inhomogeneous scalar Helmholtz equation:

$$\nabla^2 \Phi - \mu_o \varepsilon_o \partial^2 \Phi / \partial t^2 = -\rho / \varepsilon_o \tag{10.1.37}$$

analogous to the vector version (10.1.34) for $\overline{A}$. These inhomogeneous scalar and vector Helmholtz equations, (10.1.34) and (10.1.37), permit us to calculate the electric and magnetic potentials and fields produced anywhere in vacuum as a result of arbitrary source charges and currents, as explained below.

The solutions to the Helmholtz equations must reduce to: a) the traveling-wave solutions [e.g., (2.2.9)] for the wave equation [e.g., (2.2.7)] when the source terms are zero, and b) the static solutions (10.1.10) and (10.1.21) when $\partial / \partial t = 0$. The essential feature of solutions to wave equations is that their separate dependences on space and time must have the same form because their second derivatives with respect to space and time are identical within a constant multiplier. These solutions can therefore be expressed as an arbitrary function of a single argument that sums time and space, e.g. $(z - ct)$ or $(t - r_{pq}/c)$. The solutions must also have the form of the static solutions because they reduce to them when the source is static. Thus the solutions to the Helmholtz inhomogeneous equations are the static solutions expressed in terms of the argument $(t - r_{pq}/c)$:

$$\Phi_p = \iiint_{V_q} \left[\rho_q \left(t - r_{pq}/c \right) / \left(4\pi \varepsilon_o r_{pq} \right) \right] dv \quad [V] \tag{10.1.38}$$

$$\overline{A}_p = \iiint_{V_q} \left[\mu_o \overline{J}_q \left(t - r_{pq}/c \right) / \left(4\pi r_{pq} \right) \right] dv \quad \left[Vsm^{-1} \right] \tag{10.1.39}$$

These solutions are the dynamic scalar *Poisson integral* and the dynamic vector Poisson integral, respectively. Note that Φ_p and $\overline{A}_p$ depend on the state of the sources at some time in the past, not on their instantaneous values. The delay r_{pq}/c is the ratio of the distance r_{pq} between the source and observer, and the velocity of light c. That is, r_{pq}/c is simply the propagation time between source and observer.

10.2 Short dipole antennas

10.2.1 Radiation from Hertzian dipoles

Since Maxwell's equations are linear, superposition applies and therefore the electromagnetic field produced by an arbitrary current distribution is simply the integral of the fields produced by each infinitesimal element. Thus the electromagnetic field response to an infinitesimal current element is analogous to the impulse response of a linear circuit, and comparably useful for calculating responses to arbitrary stimuli.

The simplest infinitesimal radiating element, called a *Hertzian dipole*, is a current element of length d carrying I(t) amperes. Conservation of charge requires charge reservoirs at each end of the current element containing ±q(t) coulombs, where I = dq/dt, as illustrated in Figure 10.2.1(a). The total charge is zero. If we align the z axis with the direction of the current and assume the cross-sectional area of the current element is A_c [m^2], then the current density within the element is:

$$\bar{J}_q(t) = \hat{z} I(t) \Big/ A_c \quad \left[Am^{-2}\right] \tag{10.2.1}$$

Substituting this current density into the expression (10.1.39) for vector potential yields:

$$\overline{A}_p = \hat{z} \iiint_V \left[\mu_o I\left(t - \frac{r_{pq}}{c} \right) \Big/ \left(A_c 4\pi r_{pq} \right) \right] dv = \hat{z} \frac{\mu_o d}{4\pi r_{pq}} I\left(t - \frac{r_{pq}}{c} \right) \ [Vs/m] \tag{10.2.2}$$

where integration over the volume V of the current element yielded a factor of $A_c d$.

To obtain simple expressions for the radiated electric and magnetic fields we must now switch to: 1) time-harmonic representations because radiation is frequency dependent, and 2) polar coordinates because the symmetry of the radiation is polar, not cartesian, as suggested in Figure 10.2.1(b). The time harmonic form of $I(t - r_{pq}/c)$ is $Ie^{-jkr_{pq}}$ and the polar form of $\hat{z}$ is $\hat{r}\cos\theta - \hat{\theta}\sin\theta$, so (10.2.2) becomes:

$$\underline{\overline{A}}_p = \left(\hat{r}\cos\theta - \hat{\theta}\sin\theta \right) \mu_o \underline{I} d e^{-jkr_{pq}} \Big/ 4\pi r_{pq} \quad \left[Vsm^{-1}\right] \tag{10.2.3}$$

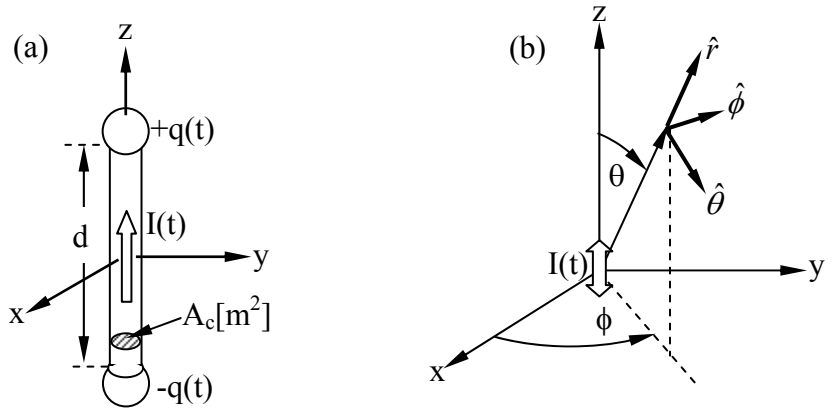

Figure 10.2.1 Hertzian dipole in spherical coordinates.

To find $\overline{\underline{H}}$ and $\overline{\underline{E}}$ radiated by this current element, we need to compute the curl of $\overline{\underline{A}}$ in spherical coordinates:

$$\overline{\underline{H}} = \left(\nabla \times \overline{\underline{A}}\right)/\mu_o = \left(\mu_o r^2 \sin\theta\right)^{-1} \det \begin{vmatrix} \hat{r} & r\hat{\theta} & r\sin\theta\hat{\phi} \\ \partial/\partial r & \partial/\partial\theta & \partial/\partial\phi \\ A_r & rA_\theta & r\sin\theta A_\phi \end{vmatrix} \tag{10.2.4}$$

Since $\overline{\underline{A}}$ is independent of position ϕ (so $\partial/\partial\phi = 0$) and has no ϕ component, (10.2.4) becomes:

$$\overline{\underline{H}} = \hat{\phi}\left(jk\underline{I}d/4\pi r\right)e^{-jkr}\left[1+\left(jkr\right)^{-1}\right]\sin\theta \tag{10.2.5}$$

After some computation the radiated electric field can be found from (10.2.5) using Ampere's law (2.3.17):

$$\begin{aligned} \overline{\underline{E}} &= \left(\nabla \times \overline{\underline{H}}\right)/j\omega\varepsilon_o \\ &= j\frac{k\underline{I}d\eta_o}{4\pi r}e^{-jkr}\left\{\hat{r}\left[\frac{1}{jkr}+\frac{1}{\left(jkr\right)^2}\right]2\cos\theta + \hat{\theta}\left[1+\frac{1}{jkr}+\frac{1}{\left(jkr\right)^2}\right]\sin\theta\right\} \end{aligned} \tag{10.2.6}$$

These solutions (10.2.5–6) for the Hertzian dipole are fundamental because they permit us to calculate easily the radiation from arbitrary current sources. It suffices to know the source current distribution because it uniquely determines the charge distribution via conservation of charge (2.1.19), and therefore the charge does not radiate independently.

These solutions for $\overline{E}$ and $\overline{H}$ are polynomials in 1/jkr, so they have two asymptotes—one for large values of kr and one for small values. When kr is very large the lowest order terms dominate, so $kr = 2\pi r/\lambda >> 1$, or:

$$r >> \lambda/2\pi \qquad \text{(far field)} \qquad (10.2.7)$$

which we call the *far field* of the dipole.

In the far field the expressions (10.2.5) and (10.2.6) for $\overline{H}$ and $\overline{E}$ simplify to:

$$\overline{\underline{E}} = \hat{\theta}\,\frac{jk\underline{I}d\eta_o}{4\pi r}\,e^{-jkr}\sin\theta \qquad \text{(far-field electric field)} \qquad (10.2.8)$$

$$\overline{\underline{H}} = \hat{\phi}\left(jk\underline{I}de^{-jkr}\sin\theta\right)\Big/4\pi r \qquad \text{(far-field magnetic field)} \qquad (10.2.9)$$

These expressions are identical, except that $\overline{E}$ points in the θ direction while $\overline{H}$ points in the orthogonal ϕ direction; the radial components are negligible in the far field. Also:

$$|\overline{\underline{E}}| = |\overline{\underline{H}}|\,\eta_o \qquad (10.2.10)$$

where the impedance of free space $\eta_o = \sqrt{\mu_o/\varepsilon_o} \cong 377$ ohms. We found similar orthogonality and proportionality for uniform plane waves in Sections 2.3.2 and 2.3.3.

We can calculate the radiated intensity in the far field using (2.7.41) and the field expressions (10.2.8) and (10.2.9):

$$\langle\overline{S}(t)\rangle = 0.5R_e\left[\overline{S}\right] = 0.5R_e\left[\overline{\underline{E}}\times\overline{\underline{H}}^*\right] \qquad (10.2.11)$$

$$\langle\overline{S}(t)\rangle = \hat{r}\,|\overline{\underline{E}}_o|^2\big/2\eta_o = \hat{r}\,(\eta_o/2)|kId/4\pi r|^2\sin^2\theta \quad \left[W\,m^{-2}\right] \qquad (10.2.12)$$

The *radiation pattern* $\langle S(t,\theta)\rangle$ for a Hertzian dipole is therefore a donut-shaped figure of revolution about its z axis, as suggested in Figure 10.2.2(b).

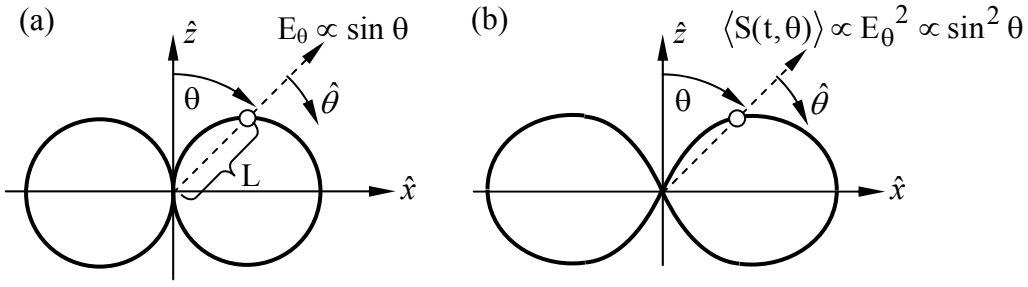

Figure 10.2.2 Electric field strength E_θ and power $\langle S(t,\theta)\rangle$ radiated by a Hertzian dipole.

Hertzian dipoles preferentially radiate laterally, with zero radiation along their axis. The electric field strength E_θ varies as sin θ, which yields a circle in a polar plot, as illustrated in Figure 10.2.2(a). The distance L in the plot corresponds to $|\underline{E}_\theta|$. Since $\sin^2 45^o = 0.5$, the width of the beam between half-power points in the θ direction is 90^o.

The total power P_R radiated by a Hertzian dipole can be calculated by integrating the radial component $\langle S_r \rangle$ of $\langle \bar{S}(t,\theta) \rangle$ over all directions:

$$\begin{aligned} P_R &= \int_0^{2\pi} d\phi \int_0^{\pi} \langle S_r \rangle r^2 \sin\theta \, d\theta = \pi\eta_o |k\underline{I}d/4\pi|^2 \int_0^{\pi} \sin^3 \theta d\theta \\ &= (\eta_o/12\pi)|k\underline{I}d|^2 \cong 395 \, (Id/\lambda)^2 \quad [W] \end{aligned} \tag{10.2.13}$$

Thus the radiated fields increase linearly with $\underline{I}d/\lambda$ and the total radiated power increases as the square of this factor, i.e. as $|\underline{I}d/\lambda|^2$. Since the electric and magnetic fields are in phase with each other in the far field, the imaginary power $I_m\{\bar{\underline{S}}\} = 0$ there.

Example 10.2A

Equation (10.2.13) says the current I input to a Hertzian dipole radiates P_R watts. What value resistor R_r would dissipate the same power for the same I?

Solution: $P_R = (\eta_o/12\pi)|k\underline{I}d|^2 = |\underline{I}|^2 R_r/2 \Rightarrow R_r = (2\pi\eta_o/3)(d/\lambda)^2$ ohms; this quantity is often called the radiation resistance of the radiator.

10.2.2 Near fields of a Hertzian dipole

If we examine the *near fields* radiated by a Hertzian dipole close to the origin where kr << 1, then the $(kr)^{-2}$ terms in the expression (10.2.7) for $\bar{\underline{E}}$ dominate all other terms for both $\bar{\underline{E}}$ (10.2.6) and $\bar{\underline{H}}$ (10.2.5), so we are left primarily with $\bar{\underline{E}}$:

$$\bar{\underline{E}} \cong \left(\underline{I}d/j\omega 4\pi\varepsilon_o r^3\right)\left(\hat{r}\, 2\cos\theta + \hat{\theta}\sin\theta\right) \tag{10.2.14}$$

This is simply the electric field produced by a static *electric dipole* of length d with a *dipole moment* of $\underline{p} = \underline{q}d$, where the charges at the ends of the dipole are ± q and $\underline{I} = j\omega\underline{q}$ to conserve charge.

Substituting this definition of $\underline{p}$ into (10.2.14) yields:

$$\bar{\underline{E}} \cong \left(\underline{p}/4\pi\varepsilon_o r^3\right)\left(\hat{r}\, 2\cos\theta + \hat{\theta}\sin\theta\right) \qquad \text{(near-field electric field)} \tag{10.2.15}$$

The dominant term for $\underline{\overline{H}}$ in the near field is:

$$\underline{\overline{H}} \cong \hat{\phi}\left(j\omega p/4\pi r^2\right)\sin\theta \qquad \text{(near-field magnetic field)} \quad (10.2.16)$$

Because $\underline{\overline{S}} = \underline{\overline{E}} \times \underline{\overline{H}}^*$ is purely negative imaginary in the near fields, these fields correspond to reactive power and stored electric energy. Integrating the exact expressions for $\underline{\overline{S}}$ over 4π steradians yields a real part that is independent of r; that is, the total power radiated is the same (10.2.13) regardless of the radius r at which we integrate, even in the near field.

A simple expression for $\overline{H}$ in the near field of the source is called the *Biot-Savart law*; it easily follows from the expression (10.2.5) for magnetic fields close to a current element $\underline{I}d\hat{z}$ when kr << 1:

$$\underline{\overline{H}} = \hat{\phi}(\underline{I}d\sin\theta)/4\pi r^2 \qquad (10.2.17)$$

The Biot-Savart law relates arbitrary current distributions $\overline{J}(\bar{r}',t)$ to $\overline{H}(\bar{r},t)$ when the distance $r = |\bar{r}' - \bar{r}| << \lambda/2\pi$; it therefore applies to static current distributions as well. But $\underline{I}d$ here is simply $\int_V \underline{J}\, dx\, dy\, dz$, where the current $\underline{I}$ and current density $\underline{J}$ are in the z direction and V is the volume of the sources of $\bar{r}'$. With these substitutions (10.2.17) becomes:

$$\underline{\overline{H}}(\bar{r}) = \iiint_V \frac{\underline{J}\hat{\phi}\sin\theta}{4\pi|\bar{r} - \bar{r}'|^2}\, dx\, dy\, dz \qquad (10.2.18)$$

We can also use the definition of vector cross product to replace $\hat{\phi}\underline{J}\sin\theta$:

$$\hat{\phi}\underline{J}\sin\theta = \frac{\overline{J} \times (\bar{r} - \bar{r}')}{|\bar{r} - \bar{r}'|} \qquad (10.2.19)$$

Substituting (10.2.19) into (10.2.18) in their time-domain forms yields the Biot-Savart law:

$$\overline{H}(\bar{r},t) = \iiint_{V'} \frac{\overline{J} \times (\bar{r} - \bar{r}')}{4\pi|\bar{r} - \bar{r}'|^3}\, dx\, dy\, dz \qquad (10.2.20)$$

Equation (10.2.20) has been generalized to $\overline{H}(\bar{r},t)$ because $\overline{H}$ is independent of frequency if $|r - r'| << \lambda/2\pi$.

10.2.3 Short dipole antennas

Antennas transform freely propagating electromagnetic waves into circuit voltages for reception, and also transform such voltages into free-space waves for transmission. They are used for wireless communications, power transmission, or surveillance at wavelengths ranging from micrometers (infrared and visible wavelengths) to hundreds of kilometers. Their sophistication and performance continue to increase as improved computational and fabrication methods are developed, although simple structures still dominate today.

Determining the fields and currents associated with a given antenna can be difficult using traditional approaches to boundary value problems because many waves must usually be superimposed in order to match boundary conditions, even when well chosen orthogonal wave expansions other than plane-waves are used. Fortunately modern computer software tools can handle most antenna problems. Here we take a traditional alternative approach to antenna analysis that yields acceptable solutions for most common configurations by making one key assumption—that the current distribution on the antenna is known. Determination of antenna current distributions is discussed in Sections 11.1–2.

Arbitrary antenna current distributions can be approximated by superimposing infinitesimal Hertzian dipole radiators that have constant current $\bar{I}$ over an infinitesimal length d. The electric far fields each dipole radiates are given by (10.2.8), and the total radiated field $\bar{E}$ is simply the sum of these differential contributions. Superposition of these fields is valid because Maxwell's equations are linear. $\bar{H}$ can then be readily found using Faraday's law or direct integration. This is the approach taken here; the far fields of the short dipole antenna are found by integrating the contributions from each infinitesimal element of that dipole. From these fields the antenna gain, effective area, and circuit properties can then be found, as discussed in Section 10.3. Many practical antennas, such as those used in many cars for the ~1-MHz Amplitude-Modulated (AM) band, are approximately short-dipole antennas with lengths less than a few percent of the associated wavelength λ. Their simple behavior provides an easy introduction to antenna analysis.

Consider the *short-dipole antenna* illustrated in Figure 10.2.3. It has length $d << \lambda$ and is driven by a current source with $\underline{I}_o$ amperes at angular frequency ω [radians/second]. Complex notation is used here for simplicity because antenna characteristics are frequency-dependent. As discussed later, the wires, or "feed-lines", providing current to the dipole do not alter the dipole's radiated fields because those wires are perpendicular to the antenna fields and also do not radiate.

The electric far field $\bar{\underline{E}}_{ff}'$ radiated by an infinitesimal current element $\underline{I}$ of length δ is given by (10.2.8):

$$\bar{\underline{E}}_{ff}' = \hat{\theta}\frac{jk\underline{I}\delta\eta_o}{4\pi r}e^{-jkr}\sin\theta \qquad \text{(far-field electric field)} \quad (10.2.21)$$

This expression can be integrated over the contributions from all infinitesimal elements of the current distribution on the short dipole to find the total radiated electric far field $\bar{\underline{E}}_{ff}$:

$$\underline{\overline{E}}_{ff} = \int_{-d/2}^{d/2} \underline{\overline{E}}_{ff}{}'(r',\theta)\,dz = \frac{jk\eta_o}{4\pi}\int_{-d/2}^{d/2} \hat{\theta}'\left[\underline{I}(z)e^{-jkr'}\frac{\sin\theta'}{r'}\right]dz \qquad (10.2.22)$$

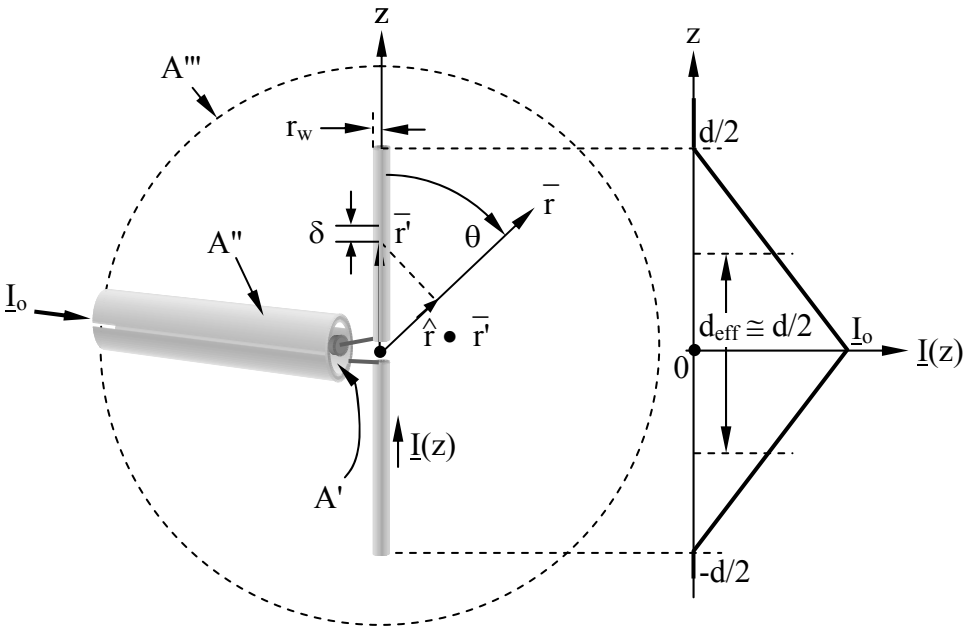

Figure 10.2.3 Short-dipole antenna.

This integral can be simplified if the observer is far from the antenna relative to its length d so that $\theta' = \theta$, $\hat{\theta}' \cong \hat{\theta}$, and $r'^{-1} \cong r^{-1}$. In addition, if $d << \lambda/2\pi$ for all z, then:

$$e^{-jkr'} \cong e^{-jkr} \qquad (10.2.23)$$

$$\begin{aligned}\underline{\overline{E}}_{ff} &\cong \hat{\theta} j\frac{k\eta_o}{4\pi r}\sin\theta e^{-jkr}\int_{-d/2}^{d/2}\underline{I}(z)\,dz \\ &= \hat{\theta} j\frac{k\eta_o \underline{I}_o d_{eff}}{4\pi r}\sin\theta\, e^{-jkr}\end{aligned} \qquad \text{(far-field radiation)} \quad (10.2.24)$$

where the *effective length* of the dipole d_{eff} is illustrated in Figure 10.2.3 and is defined as:

$$d_{eff} \equiv \underline{I}_o^{-1}\int_{-d/2}^{d/2}\underline{I}(z)\,dz \qquad \text{(effective length of short dipole)} \quad (10.2.25)$$

Because both these far fields $\underline{\overline{E}}_{ff}$ and the near fields are perpendicular to the x-y plane where the feed line is located, they are consistent with the boundary conditions associated with a sufficiently small conducting feed line located in that plane, and no reflected waves are

produced. Moreover, if the feed line is a coaxial cable with a conducting sheath, Poynting's vector on its outer surface is zero so it radiates no power.

The far fields radiated by a short dipole antenna are thus radially propagating θ-polarized plane waves with ϕ-directed magnetic fields $\overline{H}$ of magnitude $|\overline{E}|/\eta_o$. The time-average intensity $\overline{P}$ of these radial waves is given by Poynting's vector:

$$\overline{P} = \frac{1}{2} R_e \{\overline{S}\} = \frac{1}{2} R_e \left\{\overline{E} \times \overline{H}^*\right\} = \hat{r} \frac{|\overline{E}_{ff}|^2}{2\eta_o} \quad [\mathrm{Wm}^{-2}] \tag{10.2.26}$$

$$\overline{P} = \hat{r} \frac{\eta_o}{2} \left(\frac{k |I_o d_{eff}|}{4\pi r} \right)^2 \sin^2\theta = \hat{r} \frac{\eta_o}{2} \left| \frac{I_o d_{eff}}{\lambda 2r} \right|^2 \sin^2\theta \tag{10.2.27}$$

This angular distribution of radiated power is illustrated in Figure 10.2.4.

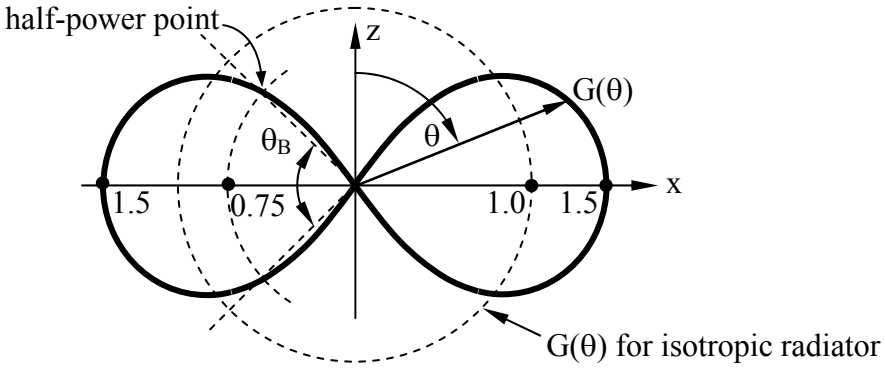

Figure 10.2.4 Antenna gain G(θ) for a short dipole or Hertzian antenna.

The total power radiated is the integral of this intensity over 4π steradians[52]:

$$P_T = \int_{4\pi} \left[\overline{P}(r,\theta) \bullet \hat{r}\right] r^2 \sin\theta \, d\theta \, d\phi = \frac{\eta_o \pi}{3} \left| \frac{I_o d_{eff}}{\lambda} \right|^2 [\mathrm{W}] \qquad \text{(radiated power)} \tag{10.2.28}$$

10.3 Antenna gain, effective area, and circuit properties

10.3.1 Antenna directivity and gain

The far-field intensity $\overline{P}(r,\theta)$ [W m^{-2}] radiated by any antenna is a function of direction, as given for a short dipole antenna by (10.2.27) and illustrated in Figure 10.2.4. *Antenna gain*

[52] Recall $\int_0^{\pi/2} \sin^n x \, dx = [2\bullet4\bullet6...(n-1)]/[1\bullet3\bullet5...(n)]$ for n odd; $(\pi/2)[1\bullet3\bullet5...(n-1)]/[2\bullet4\bullet6...(n)]$ for n even.

$G(\theta,\phi)$ is defined as the ratio of the intensity $P(\theta,\phi,r)$ to the intensity [Wm^{-2}] that would result if the same total power available at the antenna terminals, P_A [W], were radiated isotropically over 4π steradians. $G(\theta,\phi)$ is often called "gain over isotropic" where:

$$G(\theta,\phi) \equiv \frac{P(r,\theta,\phi)}{\left(P_A/4\pi r^2\right)} \qquad \text{(antenna gain definition)} \qquad (10.3.1)$$

A related quantity is *antenna directivity* $D(\theta,\phi)$, which is normalized to the total power radiated P_T rather than to the power P_A available at the antenna terminals:

$$D(\theta,\phi) \equiv \frac{P(r,\theta,\phi)}{\left(P_T/4\pi r^2\right)} \qquad \text{(antenna directivity definition)} \qquad (10.3.2)$$

The transmitted power is less than the available power if the antenna is mismatched or lossy. Since the total power radiated is $P_T = r^2 \int_{4\pi} P(r,\theta,\phi) \sin\theta \, d\theta \, d\phi$, a useful relation follows from (10.3.2):

$$\oint_{4\pi} D(\theta,\phi) \sin\theta \, d\theta \, d\phi = 4\pi \qquad (10.3.3)$$

Equation (10.3.3) says that if the directivity or gain is large in one direction, it must be correspondingly diminished elsewhere, as suggested in Figure 10.2.4, where the pattern is plotted relative to an isotropic radiator and exhibits its "main lobe" in the direction $\theta = 90°$. This pattern is independent of ϕ. The half-power *antenna beamwidth* in the θ direction is the angle θ_B between two directions where the radiated power is half that radiated at the peak, as illustrated. Thus (10.3.3) and the figure also suggest that high directivity antennas have narrower beamwidths θ_B, or are more "directive".

The ratio P_T/P_A is that fraction of the power available at the antenna terminals (P_A) that is radiated; it is defined as the *radiation efficiency* η_R:

$$\eta_R \equiv P_T/P_A \qquad \text{(radiation efficiency)} \qquad (10.3.4)$$

$$G(\theta,\phi) \equiv \eta_R D(\theta,\phi) \qquad (10.3.5)$$

The radiation efficiency is usually near unity because the resistive losses and the reflective losses due to impedance mismatches are small in most systems. Typical exceptions to the rule $\eta_R \cong 1$ include most short dipoles and antennas that are used over bandwidths much greater than an octave; their impedances are difficult to match.

The directivity of a short dipole antenna is given by substituting (10.2.27) and (10.2.28) into (10.3.2):

$$D(\theta,\phi) = \frac{(\eta_o/2)|\underline{I}_o d/\lambda 2r|^2 \sin^2\theta}{(\eta_o \pi/3)|\underline{I}_o d/\lambda|^2 / 4\pi r^2} = 1.5\sin^2\theta \quad \text{(short dipole directivity)} \qquad (10.3.6)$$

Lossless matched short dipole antennas have gain:

$$G(\theta,\phi) = 1.5\sin^2\theta \qquad \text{(short-dipole antenna gain)} \qquad (10.3.7)$$

Example 10.3A

What is the maximum solid angle Ω_B [steradians] over which a lossless matched antenna can have constant gain G_o = 40 dB? If the beam is circular, approximately what is its diameter θ_B? How much transmitter power P_T is required to yield $\underline{E}_o$ = 1 volt per meter at 10 kilometers?

Solution: Since $G(\theta,\phi) = D(\theta,\phi)$ for a lossless matched antenna, and $\int_{4\pi} D(\theta,\phi)\,d\Omega = 4\pi$, it follows that $G_o\Omega_B = 4\pi$ since the maximum gain results when all sidelobes have G = 0. Therefore $\Omega_B = 4\pi\times10^{-4}$, corresponding to $\pi\theta_B^2/4 \cong \Omega_B \Rightarrow \theta_B \cong 2(\Omega_B/\pi)^{0.5} \cong 2(4\pi\times10^{-4}/\pi)^{0.5} \cong 0.04$ radians $\cong 2.4°$. $G_oP_T/4\pi r^2 = |\underline{E}_o|^2/2\eta_o \Rightarrow P_T = 4\pi r^2|\underline{E}_o|^2/2\eta_oG_o = 4\pi(10^4)^2\times1^2/(2\times377\times10^4) \cong 166$ [W].

10.3.2 Circuit properties of antennas

Antennas connect to electrical circuits, and therefore it is important to understand the circuit properties of antennas. The linearity of Maxwell's equations applies to antennas, so they can therefore be modeled by a *Thevenin equivalent circuit* consisting of a *Thevenin equivalent impedance* $\underline{Z}_A$ in series with a *Thevenin voltage source* $\underline{V}_{Th}$. This section evaluates the Thevenin equivalent impedance $\underline{Z}_A$, and Section 10.3.3 evaluates $\underline{V}_{Th}$. The frequency dependence of these circuit equivalents usually does not map neatly into that of inductors, capacitors, and resistors, and so we simply use complex notation and a generalized $\underline{Z}_A(\omega)$ instead, where:

$$\underline{Z}_A(\omega) = R(\omega) + jX(\omega) \qquad (10.3.8)$$

$R(\omega)$ is the resistive part of the impedance corresponding to the total power dissipated and radiated, and $X(\omega)$ is the reactive part, corresponding to near-field energy storage.

To find $\underline{Z}_A(\omega)$ we can use the integral form of Poynting's theorem (2.7.23) for a volume V bounded by surface area A to relate the terminal voltage $\underline{V}$ and current $\underline{I}$ to the near and far fields of any antenna:

$$\oiint_A \left(\overline{\underline{E}}\times\overline{\underline{H}}^*\right)\bullet\hat{n}\;da = -\iiint_V \left\{\overline{\underline{E}}\bullet\overline{\underline{J}}^* + j\omega\left(\overline{\underline{H}}^*\bullet\overline{\underline{B}} - \overline{\underline{E}}\bullet\overline{\underline{D}}^*\right)\right\}dv \qquad (10.3.9)$$

For example, the short dipole antenna in Figure 10.2.3 is shown surrounded by a surface area A = A' + A" + A"', where A' is the cross-sectional area of the TEM feed line, A" is the outer surface of the coaxial feed line, and A"' is far from the antenna and intercepts only radiated fields.

These three contributions (A', A", and A"') to the surface integral on the left-hand side of (10.3.9) are given by the next three equations:

$$\frac{1}{2}\iint_{A'}\left(\underline{\overline{E}}\times\underline{\overline{H}}^{*}\right)\bullet\hat{n}\,da=-\frac{1}{2}\underline{V}\underline{I}^{*}=-\frac{1}{2}\underline{Z}\left|\underline{I}_{o}\right|^{2}\ \ [W] \tag{10.3.10}$$

Equation (10.3.10) simply expresses in two different ways the power flowing away from the antenna through the TEM feed line; the negative sign results because Poynting's vector here is oriented outward and the current flow $\underline{I}$ is oriented inward. Because no power flows perpendicular to the conducting sheath of the feed line, we have:

$$\iint_{A"}\left(\underline{\overline{E}}\times\underline{\overline{H}}^{*}\right)\bullet\hat{n}\,da=0 \tag{10.3.11}$$

The third integral over the far fields A"' captures the total power radiated by the antenna, which must equal the real power into the antenna associated with radiation, or $R_r\left|\underline{I}_o\right|^2/2$, where (10.3.12) defines the *radiation resistance* R_r of an antenna. In the far field the left-hand side is purely real:

$$\frac{1}{2}\iint_{A'''}\left(\underline{\overline{E}}\times\underline{\overline{H}}^{*}\right)\bullet\hat{n}\,da=P_T\equiv\frac{1}{2}\left|\underline{I}_o\right|^2 R_r\ \ [W]\qquad\text{(radiation resistance)} \tag{10.3.12}$$

By combining the expression for $\underline{Z}(\omega)$ in (10.3.10) with equations (10.3.9–12) we obtain:

$$\underline{Z}(\omega)=R+jX=R_r+\iiint_V\left\{\left[\underline{\overline{E}}\bullet\underline{\overline{J}}^{*}+j\omega\left(\underline{\overline{H}}^{*}\bullet\underline{\overline{B}}-\underline{\overline{E}}\bullet\underline{\overline{D}}^{*}\right)\right]\Big/\left|\underline{I}_o\right|^2\right\}dv \tag{10.3.13}$$

$$R(\omega)=R_r+\iiint_V jR_e\left\{\left[\underline{\overline{E}}\bullet\underline{\overline{J}}^{*}+\ \omega\left(\underline{\overline{H}}^{*}\bullet\underline{\overline{B}}\ -\ \underline{\overline{E}}\bullet\underline{\overline{D}}^{*}\right)\right]\Big/\left|\underline{I}_o\right|^2\right\}dv=R_r+R_d \tag{10.3.14}$$

$$X(\omega)=\iiint_V I_m\left\{\left[\underline{\overline{E}}\bullet\underline{\overline{J}}^{*}+j\omega\left(\underline{\overline{H}}^{*}\bullet\underline{\overline{B}}-\underline{\overline{E}}\bullet\underline{\overline{D}}^{*}\right)\right]\Big/\left|\underline{I}_o\right|^2\right\}dv \tag{10.3.15}$$

$X(\omega)$ is the antenna reactance, and the integral in (10.3.14) is the dissipative component $R_d(\omega)$ of antenna resistance $R(\omega)$. If the average near-field magnetic energy storage exceeds the electric energy storage, then the *antenna reactance* X is positive and inductive; if the energy stored is predominantly electric, then X is negative and capacitive. In practice the real part of the $j\omega$ term in (10.3.14) is usually zero, as is the imaginary part of the $\underline{\overline{E}}\bullet\underline{\overline{J}}^{*}$ term in (10.3.15), but there can be exceptions. The R and X of antennas are seldom computed analytically, but are usually determined by experiment or computational tools.

The radiation resistance R_r of short dipole antennas can be estimated using (10.3.12) and (10.2.28); the dissipative resistance R_d in short wires given by (10.3.14) is usually negligible:

$$R_r = \frac{2P_T}{|\underline{I}_o|^2} = \frac{2\eta_o \pi}{3}\left(\frac{d_{eff}}{\lambda}\right)^2 \text{ ohms} \qquad \text{(radiation resistance, short dipole)} \quad (10.3.16)$$

The effective length d_{eff} of a short dipole is approximately half its physical length [see (10.2.25) and Figure 10.2.3].

The reactance X of a short dipole antenna can be found using (10.3.15); it results primarily from the energy stored in the near fields. The near-field energy for short or Hertzian dipoles is predominantly electric, since the near-field $\overline{E} \propto r^{-3}$ (10.2.15) while the near-field $\overline{H} \propto r^{-2}$ (10.2.16), and $r \to 0$. Since the electric term of (10.3.15) is much greater than the magnetic term, X is negative.

Example 10.3B

A certain matched antenna radiates one watt (P_r) when driven with voltage $\underline{V}_o$ = 10 volts. What is the antenna radiation resistance R_r?

Solution: $P_r = |\underline{V}_o|^2 / 2R_r \Rightarrow R_r = |\underline{V}_o|^2 / 2P_r = 10^2/(2\times 1) = 50\Omega$.

10.3.3 Receiving properties of antennas

Because Maxwell's equations are linear in field strength, antennas have equivalent circuits consisting of a Thevenin equivalent impedance $\underline{Z}_A(\omega)$, given by (10.3.13), in series with a Thevenin voltage source $\underline{V}_{Th}(\omega)$ that we can now evaluate. Non-zero voltages appear when antennas receive signals, where these voltages depend upon the direction, polarization, and strength of the intercepted waves.

Figure 10.3.1(a) illustrates the Thevenin equivalent circuit for any antenna, and Figure 10.3.1(b) illustrates the electric fields and equipotentials associated with a short dipole antenna intercepting a uniform plane wave polarized parallel to the dipole axis. When the wavelength λ greatly exceeds d and other local dimensions of interest, i.e. $\lambda \to \infty$, then Maxwell's equations become:

$$\nabla \times \overline{\underline{E}} = -j(2\pi c/\lambda)\overline{\underline{B}} \to 0 \qquad \text{for } \lambda \to \infty \qquad (10.3.17)$$

$$\nabla \times \overline{\underline{H}} = \overline{\underline{J}} + j(2\pi c/\lambda)\overline{\underline{D}} \to \overline{\underline{J}} \quad \text{for } \lambda \to \infty \qquad (10.3.18)$$

But these limits are the equations of electrostatics and magnetostatics. Therefore we can quickly sketch the electric field lines near the short dipole of Figure 10.3.1 using a three-dimensional version of the quasistatic field mapping technique of Section 4.6.2.

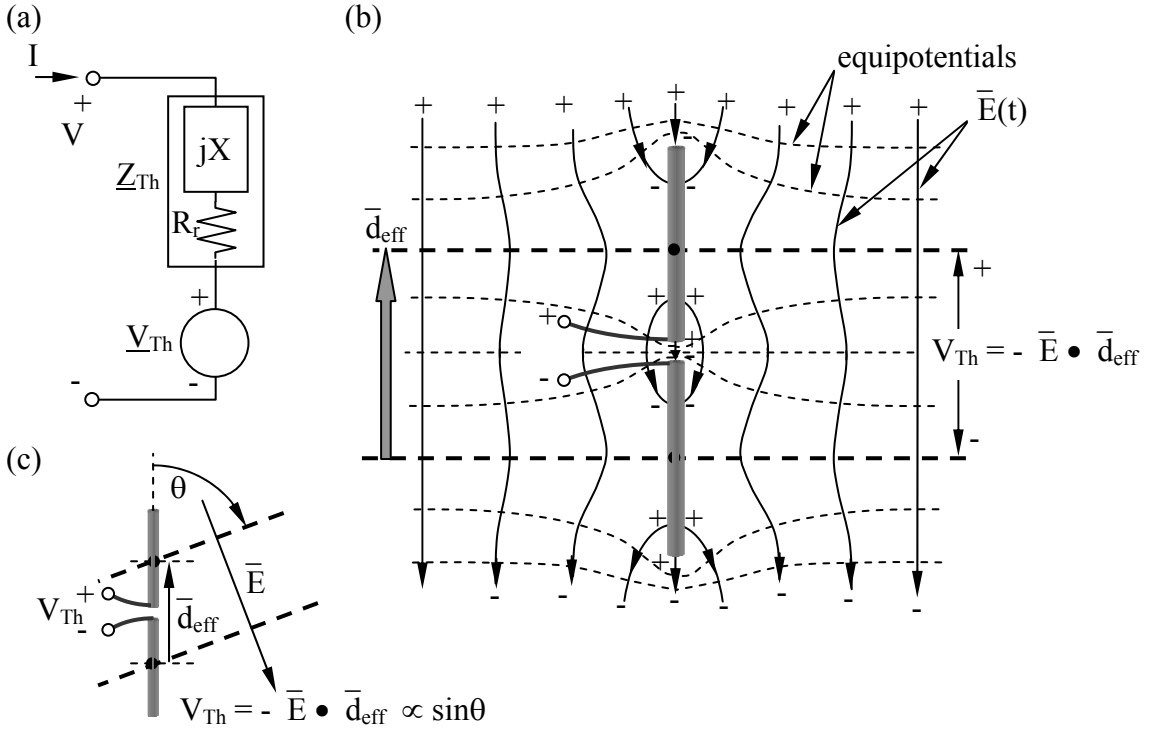

Figure 10.3.1 Thevenin voltage induced on a short dipole antenna.

Far from the dipole the field lines $\bar{E}$ in Figure 10.3.1(b) are those of the quasistatic incident plane wave, i.e., uniform and parallel to the dipole. Close to the conducting dipole $\bar{E}$ is distorted to match the boundary conditions: 1) $\bar{E}_{||} = 0$, and 2) each half of the dipole is an equipotential, intercepting only one equipotential line (boldface, dashed). If the wires comprising the short dipole are very thin, the effects of each wire on the other are negligible. Under these assumptions symmetry dictates the form for three of the equipotentials in Figure 10.3.1—the equipotentials through the center of the dipole and through each of its two halves are straight lines. The other equipotentials sketched with dashed lines curve around the conductors. The field lines $\bar{E}$ are sketched with solid lines locally perpendicular to the equipotentials. The field lines terminate at charges on the surface of the conductors and possibly at infinity, as governed by Gauss's law: $\hat{n} \bullet \bar{D} = \sigma_s$.

Figures 10.3.1(b) and (c) suggest why the open-circuit voltage V_{Th} of the short dipole antenna equals the potential difference between the centers of the two halves of this ideal dipole:

$$V_{Th} \equiv -\bar{E} \bullet \bar{d}_{eff} \qquad \text{(voltage induced on dipole antenna)} \quad (10.3.19)$$

The effective length of the dipole, $\bar{d}_{eff}$, is defined by (10.3.19), and is the same as the effective length defined in terms of the current distribution (10.2.25) for infinitesimally thin straight wires of length $d << \lambda$. Generally $d_{eff} \cong d/2$, which is the distance between the centers of the two conductors. Each conductor is essentially sampling the electrostatic potential in its vicinity and conveying that to the antenna terminals. The orientation of $\bar{d}_{eff}$ is that of the dipole current flow that would be driven by external sources having the defined terminal polarity.

The maximum power an antenna can deliver to an external circuit of impedance $\underline{Z}_L$ is easily computed once the antenna equivalent circuit is known. To maximize this transfer it is first necessary to add an external load reactance, $-jX_L$, in series to cancel the antenna reactance $+jX$ (X is negative for a short dipole antenna because it is capacitive). Then the resistive part of the load R_L must match that of the antenna, i.e., $R_L = R_r$. Maximum power transfer occurs when the impedances match so incident waves are not reflected. In this conjugate-match case ($Z_L = Z_A^*$), the antenna Thevenin voltage $\underline{V}_{Th}$ is divided across the two resistors R_r and R_L so that the voltage across R_L is $\underline{V}_{Th}/2$ and the power received by the short dipole antenna is:

$$P_r = \frac{1}{2R_r}\left|\frac{\underline{V}_{Th}}{2}\right|^2 \ [W] \qquad \text{(received power)} \quad (10.3.20)$$

Substitution into (10.3.20) of R_r (10.3.16) and V_{Th} (10.3.19) yields the received power:

$$P_r = \frac{3}{4\eta_o \pi (d/\lambda)^2}\left|\frac{\underline{\bar{E}} d_{eff} \sin\theta}{2}\right|^2 = \frac{|\underline{\bar{E}}|^2}{2\eta_o}\frac{\lambda^2}{4\pi}(1.5\sin^2\theta) \qquad (10.3.21)$$

$$P_r = I(\theta,\varphi)\frac{\lambda^2}{4\pi}G(\theta,\varphi) = I(\theta,\varphi)A(\theta,\varphi) \ [W] \qquad \text{(power received)} \quad (10.3.22)$$

where $I(\theta,\phi)$ is the power intensity $[Wm^{-2}]$ of the plane wave arriving from direction (θ,ϕ), $G(\theta,\phi) = D(\theta,\phi) = 1.5 \sin^2\theta$ is the antenna gain of a lossless short-dipole antenna (10.3.7), and $A(\theta,\phi)$ is the *antenna effective area* as defined by the equation $P_r \equiv I(\theta,\phi)\, A(\theta,\phi)$ [W] for the power received. Section 10.3.4 proves that the simple relation between gain $G(\theta,\phi)$ and effective area $A(\theta,\phi)$ proven in (10.3.22) for a short dipole applies to essentially all[53] antennas:

$$A(\theta,\varphi) = \frac{\lambda^2}{4\pi}G(\theta,\varphi) \quad \left[m^2\right] \qquad \text{(antenna effective area)} \quad (10.3.23)$$

Equation (10.3.23) says that the effective area of a matched short-dipole antenna is equivalent to a square roughly $\lambda/3$ on a side, independent of antenna length. A small wire structure ($<< \lambda/3$) can capture energy from this much larger area if it has a conjugate match, which generally

[53] This expression requires that all media near the antenna be reciprocal, which means that no magnetized plasmas or ferrites should be present so that the permittivity and permeabiliy matrices ε and μ everywhere equal their own transposes.

requires a high-Q resonance, large field strengths, and high losses. In practice, short-dipole antennas generally have a reactive mismatch that reduces their effective area below optimum.

10.3.4 Generalized relation between antenna gain and effective area

Section 10.3.3 proved for a short-dipole antenna the basic relation (10.3.23) between antenna gain $G(\theta,\phi)$ and antenna effective area $A(\theta,\phi)$:

$$A(\theta,\phi) = \frac{\lambda^2}{4\pi} G(\theta,\phi) \tag{10.3.24}$$

This relation can be proven for any arbitrary antenna provided all media in and near the antenna are *reciprocal media,* i.e., their complex permittivity, permeability, and conductivity matrices $\underline{\varepsilon}$, $\underline{\mu}$, and $\underline{\sigma}$ are all symmetric:

$$\underline{\varepsilon} = \underline{\varepsilon}^t, \quad \underline{\mu} = \underline{\mu}^t, \quad \underline{\sigma} = \underline{\sigma}^t \tag{10.3.25}$$

where we define the transpose operator t such that $\underline{A}^t_{ij} = \underline{A}_{ji}$. Non-reciprocal media are rare, but include magnetized plasmas and magnetized ferrites; they are not discussed in this text. Media characterized by matrices are discussed in Section 9.5.1.

To prove (10.3.24) we characterize a general linear 2-port network by its impedance matrix:

$$\underline{\underline{Z}} = \begin{bmatrix} \underline{Z}_{11} & \underline{Z}_{12} \\ \underline{Z}_{21} & \underline{Z}_{22} \end{bmatrix} \qquad \text{(impedance matrix)} \tag{10.3.26}$$

$$\underline{\overline{V}} = \underline{\underline{Z}}\,\underline{\overline{I}} \tag{10.3.27}$$

where $\underline{\overline{V}}$ and $\underline{\overline{I}}$ are the two-element voltage and current vectors $[\underline{V}_1, \underline{V}_2]$ and $[\underline{I}_1, \underline{I}_2]$, and $\underline{V}_i$ and $\underline{I}_i$ are the voltage and current at terminal pair i. This matrix $\underline{\underline{Z}}$ does not depend on the network to which the 2-port is connected. If the 2-port system is a *reciprocal network,* then $\underline{\underline{Z}} = \underline{\underline{Z}}^t$, so $\underline{Z}_{12} = \underline{Z}_{21}$.

Since Maxwell's equations are linear, $\underline{V}$ is linearly related to $\underline{I}$, and we can define an antenna impedance $\underline{Z}_{11}$ consisting of a real part (10.3.14), typically dominated by the radiation resistance R_r (10.3.12), and a reactive part jX (10.3.15). Thus $\underline{Z}_{11} = R_1 + jX_1$, where R_1 equals the sum of the dissipative resistance R_{d1} and the radiation resistance R_{r1}. For most antennas $R_d << R_r$.

Figure 10.3.2 illustrates an unknown reciprocal antenna (1) that communicates with a short-dipole test antenna (2) that is aimed at antenna (1). Because the relations between the voltages and currents at the terminals are determined by electromagnetic waves governed by the linear Maxwell equations, the two antennas constitute a two-port network governed by (10.3.26) and

(10.3.27) and the complex impedance matrix $\underline{\underline{Z}}$. Complex notation is appropriate here because antennas are frequency dependent. This impedance representation easily introduces the reciprocity constraint to the relation between $G(\theta,\phi)$ and $A(\theta,\phi)$. We assume each antenna is matched to its load $\underline{Z}_L = R_r - jX$ so as to maximize power transfer.

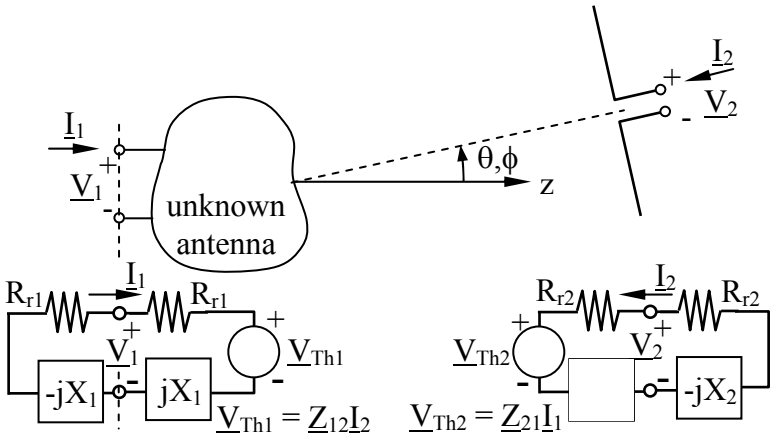

Figure 10.3.2 Coupled reciprocal antennas for relating $G(\theta,\phi)$ to $A(\theta,\phi)$.

The power P_r received by each antenna and dissipated in the load can be expressed in two equivalent ways—in terms of antenna *mutual impedance* $\underline{Z}_{ij}$ and in terms of antenna gain and effective area:

$$P_{r1} = \frac{|\underline{V}_{Th1}|^2}{8R_{r1}} = \frac{|\underline{Z}_{12}\underline{I}_2|^2}{8R_{r1}} = \frac{G_2 P_{t2}}{4\pi r^2} A_1 \tag{10.3.28}$$

$$P_{r2} = \frac{|\underline{V}_{Th2}|^2}{8R_{r2}} = \frac{|\underline{Z}_{21}\underline{I}_1|^2}{8R_{r2}} = \frac{G_1 P_{t1}}{4\pi r^2} A_2 \tag{10.3.29}$$

Taking the ratio of these two equations in terms of G and A yields:

$$\frac{P_{r2}}{P_{r1}} = \frac{G_1 A_2 P_{t1}}{G_2 A_1 P_{t2}} \tag{10.3.30}$$

$$\therefore \frac{A_1}{G_1} = \frac{A_2}{G_2}\frac{P_{t1}P_{r1}}{P_{t2}P_{r2}} \tag{10.3.31}$$

But the ratio of the same equations in terms of $\underline{Z}_{ij}$ also yields:

$$\frac{P_{r1}}{P_{r2}} = \frac{|\underline{Z}_{12}\underline{I}_2|^2 R_{r2}}{|\underline{Z}_{21}\underline{I}_1|^2 R_{r1}} = \frac{|\underline{Z}_{12}|^2 P_{t2}}{|\underline{Z}_{21}|^2 P_{t1}} \tag{10.3.32}$$

Therefore if reciprocity applies, so that $|\underline{Z}_{12}|^2 = |\underline{Z}_{21}|^2$, then (10.3.23) for a short dipole and substitution of (10.3.32) into (10.3.31) proves that all reciprocal antennas obey the same A/G relationship:

$$\frac{A_1(\theta,\phi)}{G_1(\theta,\phi)} = \frac{A_2}{G_2} = \frac{\lambda^2}{4\pi} \qquad \text{(generalized gain-area relationship)} \quad (10.3.33)$$

10.3.5 Communication links

We now can combine the transmitting and receiving properties of antennas to yield the power that can be transmitted from one place to another. For example, the intensity $I(\theta,\phi)$ at distance r that results from transmitting P_t watts from an antenna with gain $G_t(\theta,\phi)$ is:

$$I(\theta,\phi) = G(\theta,\phi)\frac{P_t}{4\pi r^2} \quad \left[\mathrm{W/m^2}\right] \qquad \text{(radiated intensity)} \quad (10.3.34)$$

The power received by an antenna with effective area $A(\theta,\phi)$ in the direction θ,ϕ from which the signal arrives is:

$$P_r = I(\theta,\phi)A(\theta,\phi) \quad [\mathrm{W}] \qquad \text{(received power)} \quad (10.3.35)$$

where use of the same angles θ,ϕ for the transmission and reception implies here that the same ray is being both transmitted and received, even though the transmitter and receiver coordinate systems are typically distinct. Equation (10.3.33) says:

$$A(\theta,\phi) = \frac{\lambda^2}{4\pi}G_r(\theta,\phi) \qquad (10.3.36)$$

where G_r is the gain of the receiving antenna, so the power received (10.3.35) becomes:

$$P_r = \frac{P_t}{4\pi r^2}G_t(\theta,\phi)\frac{\lambda^2}{4\pi}G_r(\theta,\phi) = P_t G_t(\theta,\phi)G_r(\theta,\phi)\left(\frac{\lambda}{4\pi r}\right)^2 \quad [\mathrm{W}] \qquad (10.3.37)$$

Although (10.3.37) suggests the received power becomes infinite as $r \to 0$, this would violate the far-field assumption that $r >> \lambda/2\pi$.

Example 10.3C

Two wireless phones with matched short dipole antennas having d_{eff} equal one meter communicate with each other over a ten kilometer unobstructed path. What is the maximum power P_A available to the receiver if one watt is transmitted at f = 1 MHz? At 10 MHz? What is P_A at 1 MHz if the two dipoles are 45° to each other?

Solution: P_A = AI, where A is the effective area of the receiving dipole and I is the incident wave intensity [W m^{-2}]. $P_A = A\left(P_t G_t/4\pi r^2\right)$ where $A = G_r\lambda^2/4\pi$ and $G_t \le 1.5$; $G_r \le 1.5$. Thus $P_A = \left(G_r\lambda^2/4\pi\right)\left(P_t G_t/4\pi r^2\right) = P_t\left(1.5\lambda/4\pi r\right)^2 = P_t\left(1.5c/4\pi rf\right)^2$ $= 1\left(1.5\times 3\times 10^8/4\pi 10^4\times 10^6\right)^2 \cong 1.3\times 10^{-5}\,[\mathrm{W}]$. At 10 MHz the available power out is ~1.3×10^{-7} [W]. If the dipoles are 45° to each other, the receiving cross section is reduced by a factor of $\sin^2 45° = 0.5 \Rightarrow P_A \cong 6.4\times 10^{-6}\,[\mathrm{W}]$.

Example 10.3D

In terms of the incident electric field $\underline{E}_o$, what is the maximum Thevenin equivalent voltage source $\underline{V}_{Th}$ for a small N-turn loop antenna operating at frequency f? A loop antenna is made by winding N turns of a wire in a flat circle of diameter D, where D << λ. If N = 1, what must D be in order for this loop antenna to have the same maximum $\underline{V}_{Th}$ as a short dipole antenna with effective length d_{eff}?

Solution: The open-circuit voltage $\underline{V}_{Th}$ induced at the terminals of a small wire loop (D << λ) follows from Ampere's law: $\underline{V}_{Th} = \int_C \underline{\bar{E}}\bullet d\bar{s} = -N\iint j\omega\mu_o \underline{\bar{H}}\bullet d\bar{a}$ $= -Nj\omega\mu_o \underline{H}\pi D^2/4 = -Nj\omega\mu_o \underline{E}\pi D^2/4\eta_o$. But $\omega\mu_o\pi/4\eta_o = f\pi^2/2c$, so $|\underline{V}_{Th}| =$ $Nf\pi^2\,|\underline{E}_o|\,D^2/2c$. For a short dipole antenna the maximum $|\underline{V}_{Th}| = d_{eff}\,|\underline{E}_o|$, so $D = \left(2cd_{eff}/f\pi^2 N\right)^{0.5} = \left(2\lambda d_{eff}/\pi^2 N\right)^{0.5} \cong 0.45\left(d_{eff}\lambda/N\right)^{0.5}$.

10.4 Antenna arrays

10.4.1 Two-dipole arrays

Although some communications services such as mobile phones use nearly omnidirectional electric or magnetic dipole antennas (short-dipole and loop antennas), most fixed services such as point-to-point, broadcast, and satellite services benefit from larger antenna gains. Also, some applications require rapid steering of the antenna beam from one point to another, or even the ability to observe or transmit in multiple narrow directions simultaneously. *Antenna arrays* with two or more dipoles can support all of these needs. Arrays of other types of antennas can similarly boost performance.

Since the effective area of an antenna, A(θ,ϕ), is simply related by (10.3.36) to antenna gain G(θ,ϕ), the gain of a dipole array fully characterizes its behavior, which is determined by the array current distribution. Sometimes some of the dipoles are simply mirrored images of others.

In every case the total radiated field is simply the superposition of the fields radiated by each contributing dipole in proportion to its strength, and delayed in proportion to its distance from the observer. For two-dipole arrays, the differential path length to the receiver can lead to reinforcement if the two sinusoidal waves are in phase, cancellation if they are 180° out of phase and equal, and intermediate strength otherwise.

It is convenient to represent the signals as phasors since the patterns are frequency dependent, so the total observed electric field $\overline{\underline{E}} = \sum_i \overline{\underline{E}}_i$, where $\overline{\underline{E}}_i$ is the observed contribution from short-dipole i, including its associated phase lag of γ_i radians due to distance traveled. Consider first the two-dipole array in Figure 10.4.1(a), where the dipoles are z-axis oriented, parallel, fed in phase, and spaced distance L apart laterally in the y direction. Any observer in the x-z plane separating the dipoles receives equal in-phase contributions from each dipole, thereby doubling the observed far-field $\overline{\underline{E}}_{ff}$ and quadrupling the power intensity P [Wm^{-2}] radiated in that direction θ relative to what would be transmitted by a single dipole.

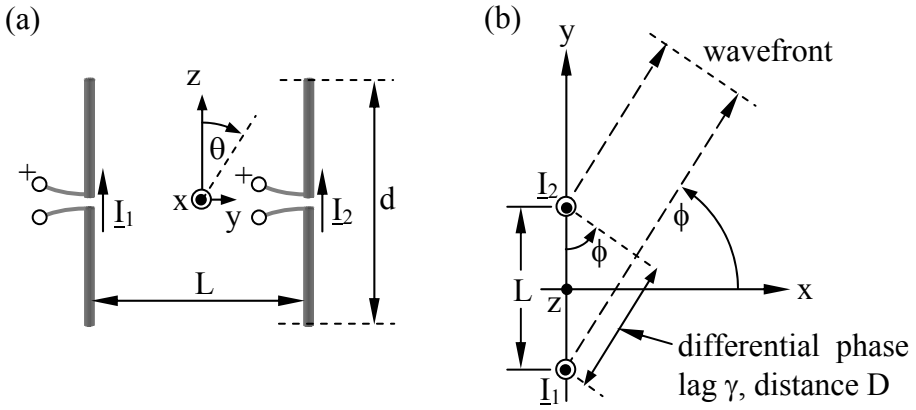

Figure 10.4.1 Two-dipole array.

The radiated power P(r,ϕ) in Figure 10.4.1 depends on the differential phase lag γ between the contributions from the two antennas. When the two dipoles are excited equally $(\underline{I}_1 = \underline{I}_2 = \underline{I})$ and are spaced L = λ/2 apart, the two rays add in phase everywhere in the x-z plane perpendicular to the array axis, but are λ/2 (180°) out of phase and cancel along the array (y) axis. The resulting G(ϕ) is sketched in Figure 10.4.2(a) for the x-y plane. If L = λ as illustrated in Figure 10.4.2(b), then the two rays add in phase along both the x-z plane and the y axis, but cancel in the x-y plane at $\phi_{null} = 30^o$ where the differential delay between the two rays is λ/2, as suggested by the right triangle in the figure.

Figure 10.4.2(c) illustrates how a non-symmetric pattern can be synthesized by exciting the two dipoles out of phase. In this case the lower dipole leads the upper dipole by 90 degrees, so that the total phase difference between the two rays propagating in the negative y direction is 180 degrees, producing cancellation; this phase difference is zero degrees for radiation in the +y direction, so the two rays add. Along the ±x axis the two rays are 90 degrees out of phase so the total E is $\sqrt{2}$ greater than from a single dipole, and the intensity is doubled. When the two phasors are in phase the total E is doubled and the radiated intensity is 4 times that of a single dipole; thus the intensity radiated along the ±x axis is half that radiated along the +y axis. Figure 10.4.2(d) illustrates how a null-free pattern can be synthesized with non-equal excitation of the two dipoles. In this case the two dipoles are driven in phase so that the radiated phase difference is 180 degrees along the ±y axis due to the λ/2 separation of the dipoles. Nulls are avoided by

exciting either dipole with a current that is ~40 percent of the other so that the ratio of maximum gain to minimum gain is ~$[(1 + 0.4)/(1 - 0.4)]^2 = 5.44$, and the pattern is vaguely rectangular.

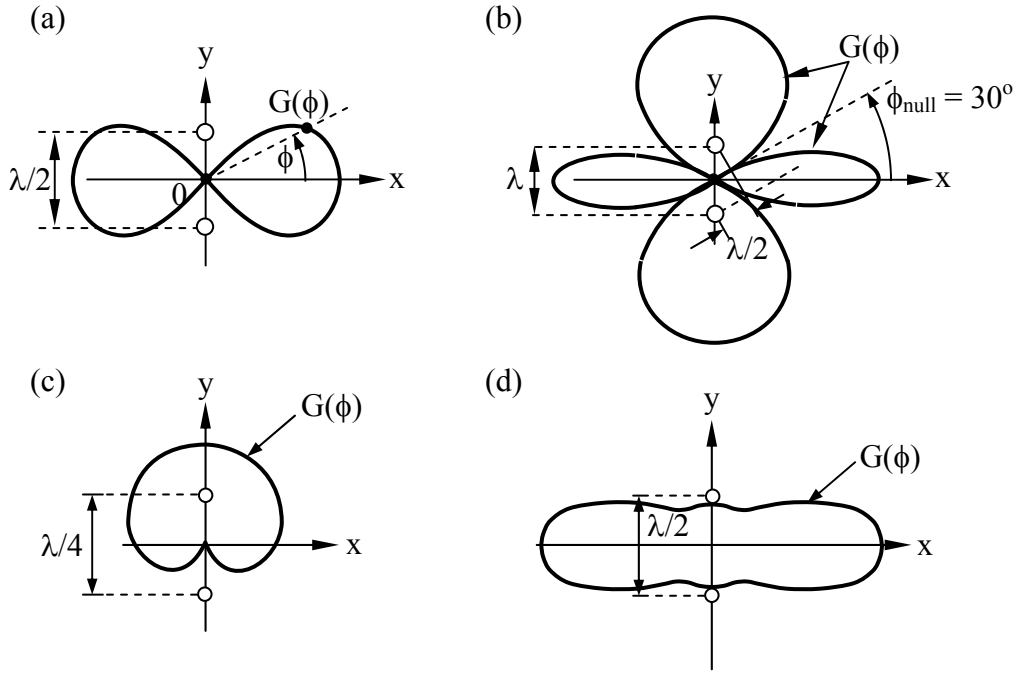

Figure 10.4.2 Gain G(ϕ) in the x-y plane orthogonal to two z-oriented dipoles.

A mathematical expression for the gain pattern can also be derived. Superimposing (10.2.8) for $\underline{I}_1$ and $\underline{I}_2$ yields:

$$\begin{aligned}\overline{\underline{E}}_{ff} &\cong \hat{\theta}\, j\left(k\eta_o d_{eff}/4\pi r\right)\sin\theta\left(\underline{I}_1 e^{-jkr_1} + \underline{I}_2 e^{-jkr_2}\right) \\ &\cong \hat{\theta}\, j\left(\eta_o d_{eff}/2\lambda r\right)\sin\theta\ \underline{I}\ e^{-jkr}\left(e^{+0.5jkL\sin\phi} + e^{-0.5jkL\sin\phi}\right) \qquad (10.4.1)\\ &\cong \hat{\theta}\, j\left(\eta_o \underline{I} d_{eff}/\lambda r\right)\sin\theta\, e^{-jkr}\cos\left(\pi L\lambda^{-1}\sin\phi\right) \qquad (10.4.2)\end{aligned}$$

where we have used the identities $e^{j\alpha} + e^{-j\alpha} = 2\cos\alpha$ and $k = 2\pi/\lambda$.

Example 10.4A

If the two dipoles of Figure 10.4.1 are fed in phase and their separation is $L = 2\lambda$, at what angles ϕ in the x-y plane are there nulls and peaks in the gain $G(\phi)$? Are these peaks equal? Repeat this analysis for $L = \lambda/4$, assuming the voltage driving the dipole at $y > 0$ has a 90° phase lag relative to the other dipole.

Solution: Referring to Figure 10.4.1, there are nulls when the phase difference γ between the two rays arriving at the receiver is π or 3π, or equivalently, $D = \lambda/2$ or $3\lambda/2$, respectively. This happens at the angles $\phi = \pm \sin^{-1}(D/L) =$

$\pm\sin^{-1}[(\lambda/2)/2\lambda] = \pm \sin^{-1}(0.25) \cong \pm 14°$, and $\phi = \pm \sin^{-1}(0.75) \cong \pm 49°$. There are also nulls, by symmetry, at angles 180° away, or at $\phi \cong \pm 194°$ and $\pm 229°$. There are gain peaks when the two rays are in phase ($\phi = 0°$ and 180°) and when they differ in phase by 2π or 4π, which happens when $\phi = \pm \sin^{-1}(\lambda/2\lambda) = \pm 30°$ and $\phi = \pm 210°$, or when $\phi = \pm 90°$, respectively. The gain peaks are equal because they all correspond to the two rays adding coherently with the same magnitudes. When $L = \lambda/4$ the two rays add in phase at $\phi = 90°$ along the +y axis because in that direction the phase lag balances the 90° delay suffered by the ray from the dipole on the -y axis. At $\phi = 270°$ these two 90-degree delays add rather than cancel, so the two rays cancel in that direction, producing a perfect null.

10.4.2 Array antennas with mirrors

One of the simplest ways to boost the gain of a short dipole antenna is to place a mirror behind it to as to reinforce the radiation in the desired forward direction and cancel it behind. Figure 10.4.3 illustrates how a short current element I placed near a perfectly conducting planar surface will behave as if the mirror were replaced by an image current an equal distance behind the mirror and pointed in the opposite direction parallel to the mirror but in the same direction normal to the mirror. The fields in front of the mirror are identical with and without the mirror if it is sufficiently large. Behind the mirror the fields approach zero, of course. Image currents and charges were discussed in Section 4.2.

Figure 10.4.3(a) illustrates a common way to boost the forward gain of a dipole antenna by placing it $\lambda/4$ in front of a planar mirror and parallel to it.

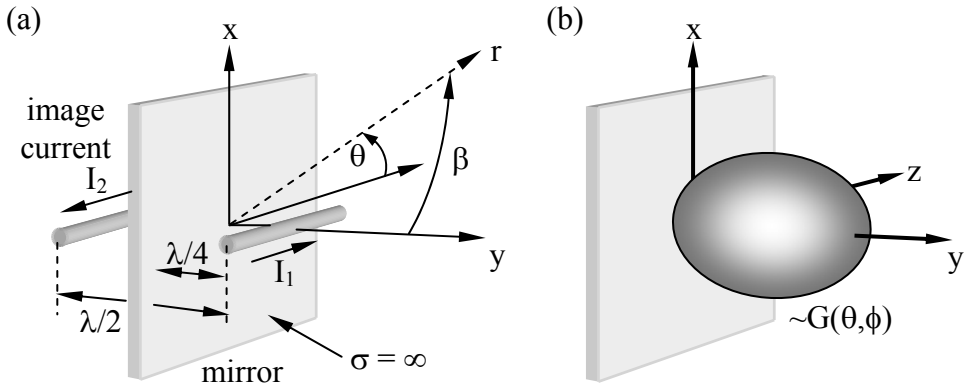

Figure 10.4.3 Half-wave dipole antenna $\lambda/4$ in front of a mirror.

The image current is 180 degrees out of phase, so the $\lambda/2$ delay suffered by the image ray brings it into phase coherence with the direct ray, effectively doubling the far field E_{ff} and quadrupling the intensity and gain G_o relative to the absence of the mirror. In all directions more nearly parallel to the mirror the source and image are more nearly out of phase, so the gain in those directions is diminished relative to the absence of the mirror. The resulting antenna gain $G(\theta,\phi)$ is sketched in Figure 10.4.3(b), and has no backlobes.

For the case where the dipole current $\underline{I}_2 = -\underline{I}_1$ and $kr_1 = kr - (\pi/2)\cos\beta$, the far-field in the forward direction is the sum of the contributions from $\underline{I}_1$ and $\underline{I}_2$, as given by (10.4.1):

$$\begin{aligned}\overline{\underline{E}}_{ff} &= \hat{\theta}\left(j\eta_o d_{eff}/2\lambda r\right)\sin\ \theta\left(\underline{I}_1 e^{-jkr_1} + \underline{I}_2 e^{-jkr_2}\right) \\ &= \hat{\theta}\left(j\eta_o d_{eff}/2\lambda r\right)\sin\ \theta e^{-jkr}\underline{I}_1\left(e^{j(\pi/2)\cos\beta} - e^{-j(\pi/2)\cos\beta}\right) \qquad (10.4.3)\\ &= -\hat{\theta}\left(\eta_o d_{eff}/\lambda r\right)\sin\ \theta\ e^{-jkr}\underline{I}_1 \sin\left[(\pi/2)\cos\beta\right] \qquad (10.4.4)\end{aligned}$$

This expression reveals that the antenna pattern has no sidelobes and is pinched somewhat more in the θ direction than in the β direction (these directions are not orthogonal). An on-axis observer will receive a z-polarized signal.

Mirrors can also be parabolic and focus energy at infinity, as discussed further in Section 11.1. The sidelobe-free properties of this dipole-plus-mirror make it a good antenna feed for radiating energy toward much larger parabolic reflectors.

Example 10.4B

Automobile antennas often are thin metal rods ~1-meter long positioned perpendicular to an approximately flat metal surface on the car; the rod and flat surface are electrically insulated from each other. The rod is commonly fed by a coaxial cable, the center conductor being attached to the base of the rod and the sheath being attached to the adjacent car body. Approximately what is the radiation resistance and pattern in the 1-MHz radio broadcast band, assuming the flat plate is infinite?

Solution: Figure 4.2.3 shows how the image of a current flowing perpendicular to a conducting plane flows in the same direction as the original current, so any current flowing in the rod has an image current that effectively doubles the length of this antenna. The wavelength at 1 MHz is ~300 meters, much longer than the antenna, so the short-dipole approximation applies and the current distribution on the rod and its image resembles that of Figure 10.2.3; thus $d_{eff} \cong 1$ meter and the pattern above the metal plane is the top half of that illustrated in Figure 10.2.4. The radiation resistance of a normal short dipole antenna (10.3.16) is $R_r = 2P_T/|I_o|^2 = 2\eta_o\pi(d_{eff}/\lambda)^2/3 = 0.0088$ ohms for $d_{eff} = 1$ meter. Here, however, the total power radiated P_T is half that radiated by a short dipole of length 2 meters because there is no power radiated below the conducting plane, so $R_r = 0.0044$ ohms. The finite size of an automobile effectively warps and shortens both the image current and the effective length of the dipole, although the antenna pattern for a straight current is always dipolar above the ground plane.

10.4.3 Element and array factors

The power radiated by dipole arrays depends on the directional characteristics of the individual dipole antennas as well as on their spacing relative to wavelength λ. For example, (10.4.3) can be generalized to N identically oriented but independently positioned and excited dipoles:

$$\underline{\bar{E}}_{ff} \cong \left[\hat{\theta}\frac{j\eta_o d_{eff}}{2\lambda r}\sin\theta\right]\left[\sum_{i=1}^{N}\underline{I}_i e^{-jkr_i}\right] = \underline{\bar{\varepsilon}}(\theta,\phi)\underline{F}(\theta,\phi) \qquad (10.4.5)$$

The *element factor* $\underline{\bar{\varepsilon}}(\theta,\phi)$ for the dipole array represents the behavior of a single element, assuming the individual elements are identically oriented. The *array factor,* $\underline{F}(\theta,\phi) = \sum_i^N \underline{I}_i e^{-jkr_i}$, represents the effects of the relative strengths and placement of the elements. The distance between the observer and each element i of the array is r_i, and the phase lag $kr_i = 2\pi r_i/\lambda$.

Consider the element factor in the x-y plane for the two z-oriented dipoles of Figure 10.4.4(a).

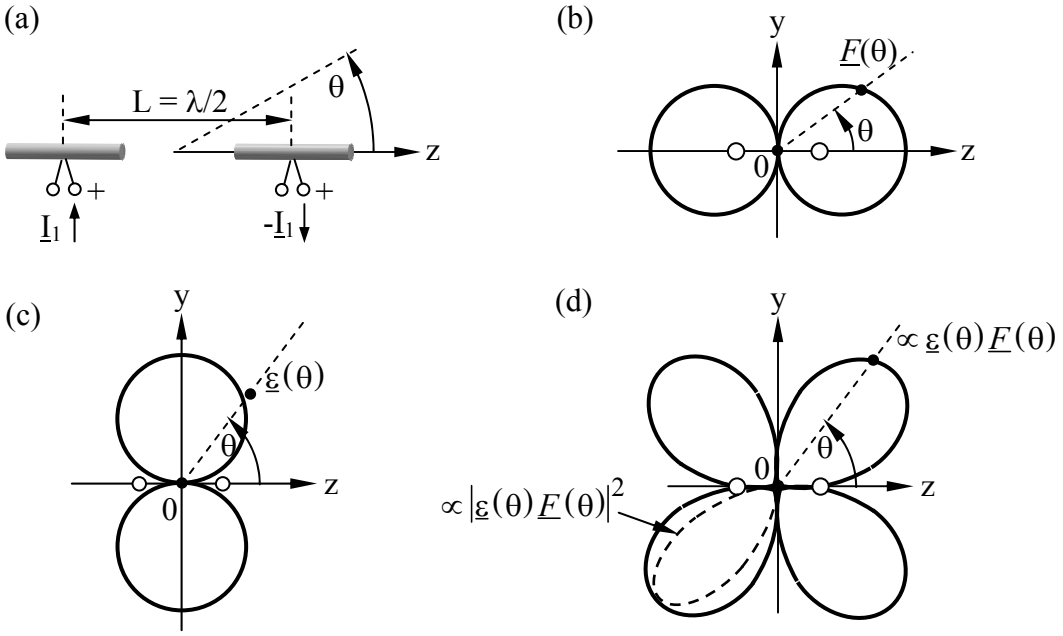

Figure 10.4.4 Normalized array and element factors for dipole arrays.

This element factor $\underline{\bar{\varepsilon}}(\theta,\phi)$ is constant, independent of ϕ, and has a circular pattern. The total antenna pattern is $\left|\underline{\bar{\varepsilon}}(\theta,\phi)\underline{F}(\theta,\phi)\right|^2$ where the array factor $\underline{F}(\theta,\phi)$ controls the array antenna pattern in the x-y plane for these two dipoles. The resulting antenna pattern $\left|\underline{F}(\theta,\phi)\right|^2$ in the x-y plane is plotted in Figure 10.4.4(a) and (b) for the special cases L = λ/2 and L = λ, respectively.

Both the array and element factors contribute to the pattern for this antenna in the x-z plane, narrowing its beamwidth (not illustrated).

Figure 10.4.4 illustrates a case where both the element and array factors are important; L = $\lambda/2$ here and the dipoles are fed 180^o out of phase. In this case the out-of-phase signals from the two dipoles cancel everywhere in the x-y plane and add in phase along the z axis, corresponding to the array factor plotted in Figure 10.4.4(b) for the y-z plane. Note that when $\theta = 60^o$ the two phasors are 45^o out of phase and $\left|\underline{F}(\theta,\phi)\right|^2$ has half its peak value. The element factor in the y-z plane appears in Figure 10.4.4(c), and the dashed antenna pattern $\left|\underline{\varepsilon}(\theta)\underline{F}(\theta)\right|^2 \propto G(\theta)$ in Figure 10.4.4(d) shows the effects of both factors (only one of the four lobes is plotted). This antenna pattern is a figure of revolution about the z axis and resembles two wide rounded cones facing in opposite directions.

Example 10.4C

What are the element and array factors for the two-dipole array for the first part of Example 10.4A?

Solution: From (10.4.5) the element factor for such dipoles is $\hat{\theta}\mathrm{j}\left(\eta_o d_{eff}/2\lambda r\right)\sin\theta$. The last factor of (10.4.5) is the array factor for such two-dipole arrays:

$$\underline{F}(\theta,\phi) = \underline{I}\left(e^{+0.5(j2\pi 2\lambda/\lambda)\sin\phi} + e^{-0.5(j2\pi 2\lambda/\lambda)\sin\phi}\right)$$

$$= \left(e^{2\pi j\sin\phi} + e^{-2\pi j\sin\phi}\right) = 2\underline{I}\cos(2\pi\sin\phi)$$

10.4.4 Uniform dipole arrays

Uniform dipole arrays consist of N identical dipole antennas equally spaced in a straight line. Their current excitation $\underline{I}_i$ has equal magnitudes for all i, and a phase angle that uniformly increases by ψ radians between adjacent dipoles. The fields radiated by the array can be determined using (10.4.5):

$$\overline{\underline{E}}_{ff} \cong \left[\hat{\theta}\mathrm{j}\left(\eta_o d_{eff}/2\lambda r\right)\sin\theta\right]\left[\sum_{i=1}^{N}\underline{I}_i e^{-\mathrm{j}kr_i}\right] = \overline{\underline{\varepsilon}}(\theta,\phi)\underline{F}(\theta,\phi) \qquad (10.4.6)$$

The z axis is defined by the orientation of the dipoles, which are all parallel to it. The simplest arrangement of the dipoles is along that same z axis, as in Figure 10.4.5, although (10.4.6) applies equally well if the dipoles are spaced in any arbitrary direction. Figure 10.4.1(a) illustrates the alternate case where two dipoles are spaced along the y axis, and Figure 10.4.2 shows the effects on the patterns.

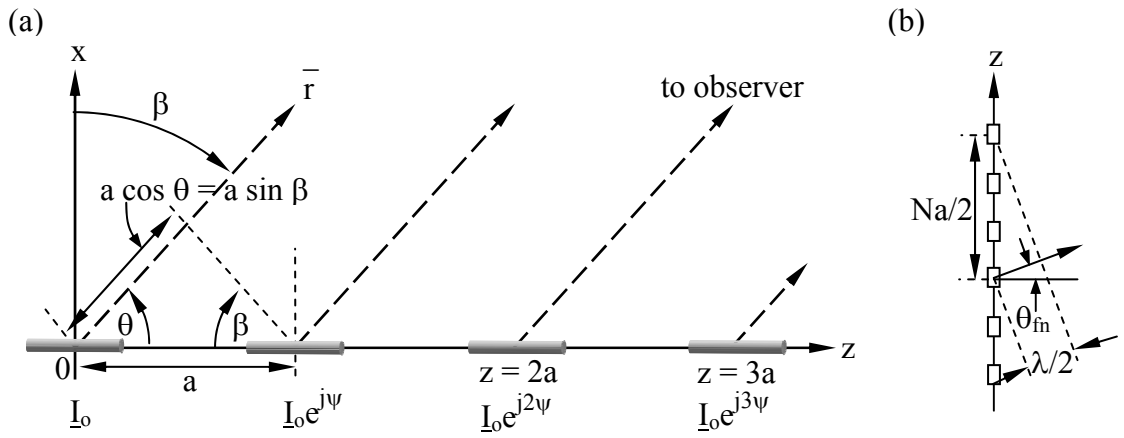

Figure 10.4.5 Uniform dipole array.

Consider the N-element array for Figure 10.4.5(a). The principal difference between these two-dipole cases and N-element uniform arrays lies in the array factor:

$$\underline{F}(\theta,\phi) = \sum_{i=1}^{N} \underline{I}_i e^{-jkr_i} = \underline{I}_o e^{-jkr} \sum_{i=0}^{N-1} e^{ji\psi} e^{jika\cos\theta} = \underline{I}_o e^{-jkr} \sum_{i=0}^{N-1} \left[e^{j(\psi + ka\cos\theta)} \right]^i \qquad (10.4.7)$$

The geometry illustrated in Figure 10.4.5(a) yields a phase difference of (ψ + ka cosθ) between the contributions from adjacent dipoles.

Using the two identities:

$$\sum_{i=0}^{N-1} x^i = (1 - x^N)/(1 - x) \qquad (10.4.8)$$

$$1 - e^{jA} = e^{jA/2}\left(e^{-jA/2} - e^{+jA/2}\right) = -2je^{jA/2}\sin(A/2) \qquad (10.4.9)$$

(10.4.7) becomes:

$$\begin{aligned} \underline{F}(\theta,\phi) &= \underline{I}_o e^{-jkr} \frac{1 - e^{jN(\psi + ka\cos\theta)}}{1 - e^{j(\psi + ka\cos\theta)}} \\ &= \underline{I}_o e^{-jkr} \times \frac{e^{jN(\psi + ka\cos\theta)/2}}{e^{j(\psi + ka\cos\theta)/2}} \frac{\sin\left[N(\psi + ka\cos\theta)/2\right]}{\sin\left[(\psi + ka\cos\theta)/2\right]} \end{aligned} \qquad (10.4.10)$$

Since the element factor is independent of ϕ, the antenna gain has the form:

$$G(\theta) \propto \left|\underline{F}(\theta,\phi)\right|^2 \propto \frac{\sin^2\left[N(\psi + ka\cos\theta)/2\right]}{\sin^2\left[(\psi + ka\cos\theta)/2\right]} \tag{10.4.11}$$

If the elements are excited in phase ($\psi = 0$), then the maximum gain is broadside with $\theta = 90^0$, because only in that direction do all N rays add in perfect phase. In this case the first nulls $\theta_{\text{first null}}$ bounding the main beam occur when the numerator of (10.4.11) is zero, which happens when:

$$\frac{N}{2} ka\cos\theta_{\text{first null}} = \pm\pi \tag{10.4.12}$$

Note that the factor $ka = 2\pi a/\lambda$ is in units of radians, and therefore $\cos\theta_{\text{first null}} = \pm\lambda/Na$. If $\theta_{\text{first null}} \equiv \pi/2 \pm \theta_{fn}$, where θ_{fn} is the null angle measured from the x-y plane rather than from the z axis, then we have $\cos\theta_{\text{first null}} = \sin\theta_{fn} \cong \theta_{fn}$ and:

$$\theta_{fn} \cong \pm\lambda/Na \quad [\text{radians}] \tag{10.4.13}$$

The following simple geometric argument yields the same answer. Figure 10.4.5(b) shows that the first null of this 6-dipole array occurs when the rays from the first and fourth dipole element cancel, for then the rays from the second and fifth, and the third and sixth will also cancel. This total cancellation occurs when the delay between the first and fourth ray is $\lambda/2$, which corresponds to the angle $\theta_{fn} = \pm\sin^{-1}[(\lambda/2)/(aN/2)] \cong \pm\lambda/aN$.

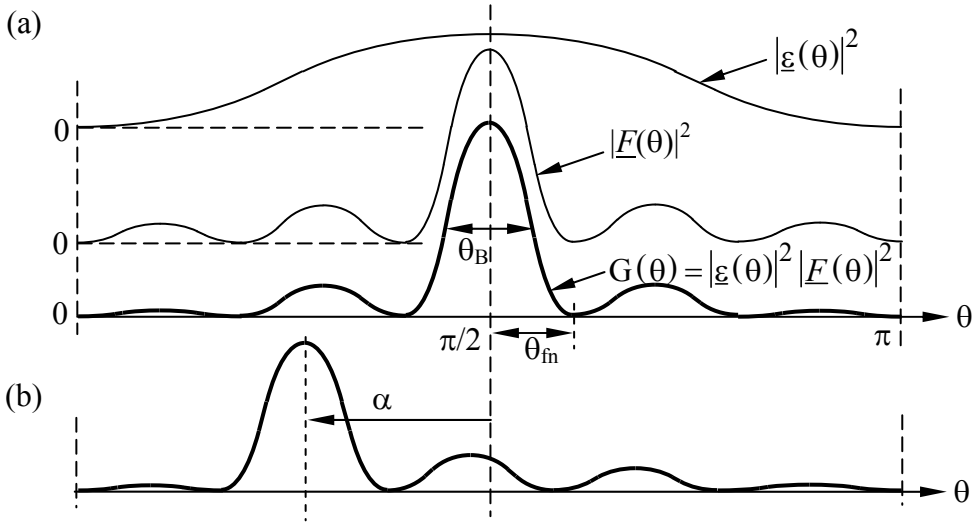

Figure 10.4.6 Antenna pattern for N-element linear dipole array.

The angle θ_{fn} between the beam axis and the first null is approximately the half-power beamwidth θ_B of an N-element antenna array. The antenna gain $G(\theta)$ associated with (10.4.11) for N = 6, $\psi = 0$, and $a = \lambda/2$ is sketched in Figure 10.4.6(a), together with the squares of the

array factor [from (10.4.10)] and element factor [from (10.4.6)]. In this case $\theta_{fn} = \pm\sin^{-1}(2/N) \cong \pm 2/N$ radians $\cong \theta_B$.

If $\psi \neq 0$ so that the excitation phase varies linearly across the array, then the main beam and the rest of the pattern is "squinted" or "scanned" to one side by angle α. Since a phase delay of ψ is equivalent to a path delay of δ, where $\psi = k\delta$, and since the distance between adjacent dipoles is a, it follows that adjacent rays for a scanned beam will be in phase at angle $\theta = \pi/2 + \sin^{-1}(\delta/a) = \pi/2 + \alpha$, where:

$$\alpha = \sin^{-1}(\delta/a) = \sin^{-1}(\psi\lambda/2\pi a) \qquad \text{(scan angle)} \quad (10.4.14)$$

as sketched in Figure 10.4.6(b) for the case $\psi = 2$ radians, $a = \lambda/2$, and $\alpha \cong 40^o$.

Note that larger element separations a can produce multiple main lobes separated by smaller ones. Additional main lobes appear when the argument $(\psi + ka\cos\theta)/2$ in the denominator of (10.4.11) is an integral multiple of π so that the denominator is zero; the numerator is zero at the same angles, so the ratio is finite although large. To preclude multiple main lobes the spacing should be $a < \lambda$, or even $a < \lambda/2$ if the array is scanned.

Example 10.4D

A uniform row of 100 x-oriented dipole antennas lies along the z axis with inter-dipole spacing $a = 2\lambda$. At what angles θ in the y-z plane is the gain maximum? See Figure 10.4.5 for the geometry, but note that the dipoles for our problem are x-oriented rather than z-oriented. What is the angle Δ between the two nulls adjacent to $\theta \cong \pi/2$? What is the gain difference $\Delta G(dB)$ between the main lobe at $\theta = \pi/2$ and its immediately adjacent sidelobes? What difference in excitation phase ψ between adjacent dipoles is required to scan these main lobes 10° to one side?

Solution: The gain is maximum when the rays from adjacent dipoles add in phase, and therefore all rays add in phase. This occurs at $\theta = 0, \pm\pi/2$, and $\pm\sin^{-1}(\lambda/a) \cong 30°$ [see Figure 10.4.5(b) for the approximate geometry, where we want a phase lag of λ to achieve a gain maximum]. The nulls nearest $\theta = \pi/2$ occur at that θ_{fn} when the rays from the first and 51st dipoles first cancel [see text after (10.4.13)], or when $\theta_{fn} = \frac{\pi}{2} \pm \sin^{-1}\left(\frac{\lambda/2}{aN/2}\right) = \frac{\pi}{2} \pm \sin^{-1}\left(\frac{\lambda/2}{2\lambda N/2}\right) \cong \frac{\pi}{2} \pm \frac{1}{2N}$; thus $\Delta = 1/N$ radians $\cong 0.57^o$. The array factors for this problem and Figure 10.4.5(a) are the same, so (10.4.10) applies. Near $\theta \cong \pi/2$ the element factor is approximately constant and can therefore be ignored because we seek only gain ratios. We define $\beta \equiv \pi/2 - \theta$ so $\cos\theta$ becomes $\sin\beta$. Therefore (10.4.11) becomes $G_o(\theta) \propto |F(\theta,\phi)|^2 \propto \sin^2(Nk\lambda\sin\beta)/\sin^2(k\lambda\sin\beta)$ where $\psi = 0$. $\beta << 1$, so $\sin\beta \cong \beta$. Similarly, $k\lambda\beta << 1$ so $\sin(k\lambda\beta) \cong k\lambda\beta$. Thus $G_o(\beta = 0) \propto \sim (Nk\lambda\beta)^2/(k\lambda\beta)^2 = N^2$, and the first

adjacent peak in gain occurs when $Nk\lambda\sin\beta_{\text{first peak}} = 1$, so $G\left(\beta = \beta_{\text{first peak}}\right) \propto \sim \left(k\lambda\beta_{\text{first peak}}\right)^{-2}$. The numerator is unity when $Nk\lambda\beta_{\text{first peak}} \cong 3\pi/2$, or $\beta_{\text{first peak}} \cong 3\pi/2Nk\lambda = 3/(4N)$. Therefore $G\left(\beta = \beta_{\text{first peak}}\right) \propto\sim \left(2\pi 3/4N\right)^{-2} \cong 0.045N^2$, which is $10\log_{10}(0.045) = -13.5$ dB relative to the peak N^2. A 10° scan angle requires the rays from adjacent dipoles to be in phase at that angle, and therefore the physical lag δ meters between the two rays must satisfy $\sin\beta_{\text{scan}} = \delta/a = \delta/2\lambda$. The corresponding phase lag in the leading dipole is $\psi = k\delta = (2\pi/\lambda)(2\lambda\sin\beta_{\text{scan}}) = 4\pi\sin(10°)$ radians $= 125°$.

10.4.5 Phasor addition in array antennas

Phasor addition can be a useful tool for analyzing antennas. Consider the linear dipole array of Figure 10.4.5, which consists of N identical z-oriented dipole antennas spaced at distance a equally along the z-axis. In direction θ from the z axis the array factor is the sum of the phasors emitted from each dipole. Figure 10.4.6(a) shows this sum A for the x-y plane (θ = 90°) when the dipoles are all excited in phase and N = 8. This yields the maximum possible gain for this antenna. As θ departs from 90° (broadside radiation) the phasors each rotate differently and add to form a progressively smaller sum B. When the total phasor B corresponds to Δϕ = 5-degree lag for each successive contribution, then $\theta = \cos^{-1}(\frac{5}{360}\frac{\lambda}{a})$. Figure 10.4.6(b, c, and d) show the sum B when Δϕ is 45°, 72°, and 90°, respectively. The antenna gain is proportional to $|B|^2$. Figures (b) and (d) correspond to radiation angles θ that yield nulls in the pattern ($|B| = 0$), while (c) is near a local maximum in the antenna pattern. Because |C| is ~0.2|A|, the gain of this sidelobe is ~0.04 times the maximum gain ($|C|^2 \cong 0.04|A|^2$), or ~14 dB weaker.[54] The spatial angles θ corresponding to (a) - (d) depend on the inter-dipole distance 'a'.

If $a = 2\lambda$, then angles θ from the z axis that correspond to phasor A in Figure 10.4.7 (a) are 0°, 60°, 90°, 120°, and 180°; the peaks at 0 and 180° fall on the null of the element factor and can be ignored. The angle from the array axis is θ, and 'a' is the element spacing, as illustrated in Figure 10.4.5. The angles θ = 0° and 180° correspond to $\cos^{-1}(2\lambda/2\lambda)$, while θ = 60° and θ = 120° correspond to $\cos^{-1}(\lambda/2\lambda)$, and θ = 90° corresponds to $\cos^{-1}(0/2\lambda)$; the numerator in the argument of $\cos^{-1}$ is the lag distance in direction θ, and the denominator is the element spacing 'a'. Thus this antenna has three equal peaks in gain: θ = 60, 90, and 120°, together with numerous smaller sidelobes between those peaks.

[54] dB ≡ 10 $\log_{10}$N, so N = 0.04 corresponds to ~-14 dB.

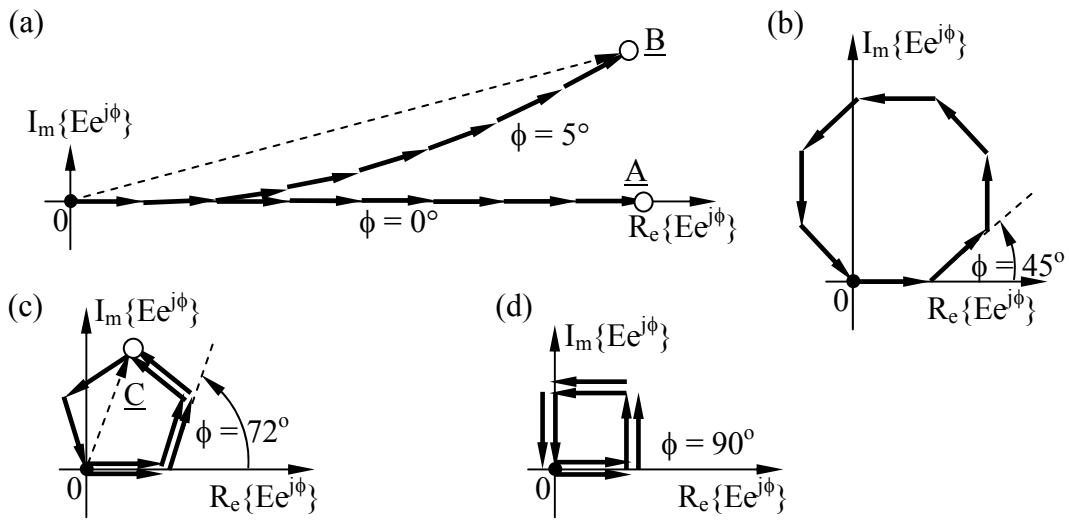

Figure 10.4.7 Phasor addition for an 8-element linear dipole array.

Four small sidelobes occur between the adjacent peaks at 60°, 90°, and 120°. The first sidelobe occurs in each case for $\phi \cong 70°$ as illustrated in Figure 10.4.2(c), i.e., approximately half-way between the nulls at $\phi = 45°$ [Figure 10.4.2(b)] and $\phi = 90°$ [Figure 10.4.2(d)], and the second sidelobe occurs for $\phi \cong 135°$, between the nulls for $\phi = 90°$ [Figure 10.4.2(d)] and $\phi = 180°$ (not illustrated). Consider, for example, the broadside main lobe at $\theta = 90°$; for this case $\phi = 0°$. As θ decreases from 90° toward zero, ϕ increases toward 45°, where the first null occurs as shown in (b); the corresponding $\theta_{null} = \cos^{-1}[(\phi\lambda/360)/2\lambda] = 86.4°$. The denominator 2λ in the argument is again the inter-element spacing. The first sidelobe occurs when $\phi \cong 72°$ as shown in (c), and $\theta \cong \cos^{-1}[(\phi\lambda/360)/2\lambda] = 84.4°$. The next null occurs at $\phi = 90°$ as shown in (d), and $\theta_{null} = \cos^{-1}[(\phi\lambda/360)/2\lambda] = 82.8°$. The second sidelobe occurs for $\phi \cong 135°$, followed by a null when $\phi = 180°$. The third and fourth sidelobes occur for $\phi \cong 225°$ and 290° as the phasor patterns repeat in reverse sequence: (d) is followed by (c) and then (b) and (a) as θ continues to decline toward the second main lobe at $\theta = 60°$. The entire gain pattern thus has three major peaks at 60, 90, and 120°, typically separated by four smaller sidelobes intervening between each major pair, and also grouped near $\theta = 0°$ and 180°.

Example 10.4E

What is the gain G_S of the first sidelobe of an n-element linear dipole array relative to the main lobe G_o as $n \to \infty$?

Solution: Referring to Figure 10.4.7(c), we see that as $n \to \infty$ the first sidelobe has an electric field E_{ffS} that is the diameter of the circle formed by the n phasors when $\sum_{i=1}^{n} |\underline{E}_i| = E_{ffo}$ is ~1.5 times the circumference of that circle, or $\sim 1.5 \times \pi E_{ffS}$. The ratio of the gains is therefore $G_S/G_o = |E_{ffS}/E_{ffo}|^2 = (1/1.5\pi)^2 = 0.045$, or -13.5 dB.

10.4.6 Multi-beam antenna arrays

Some antenna arrays are connected so as to produce several independent beams oriented in different directions simultaneously; phased array radar antennas and cellular telephone base stations are common examples. When multiple antennas are used for reception, each can be filtered and amplified before they are added in as many different ways as desired. Sometimes these combinations are predetermined and fixed, and sometimes they are adjusted in real time to place nulls on sources of interference or to place maxima on transmitters of interest, or to do both.

The following cellular telephone example illustrates some of the design issues. The driving issue here is the serious limit to network capacity imposed by the limited bandwidth available at frequencies suitable for urban environments. The much broader spectrum available in the centimeter and millimeter-wave bands propagates primarily line-of-sight and is not very useful for mobile applications; lower frequencies that diffract well are used instead, although the available bandwidth is less. The solution is to reuse the same low frequencies multiple times, even within the same small geographic area. This is accomplished using array antennas that can have multiple inputs and outputs.

A typical face of a cellular base station antenna has 3 or 4 elements that radiate only into the forward half-space. They might also have a combining circuit that forms two or more desired beams. An alternate way to use these arrays based on switching is described later. Three such faces, such as those illustrated in Figure 10.4.8(a) with four elements spaced at 3λ, might be arranged in a triangle and produce two sets of antenna lobes, for example, the $\phi = 0$ set and the $\phi = \pi$ set indicated in (b) by filled and dashed lines, respectively.

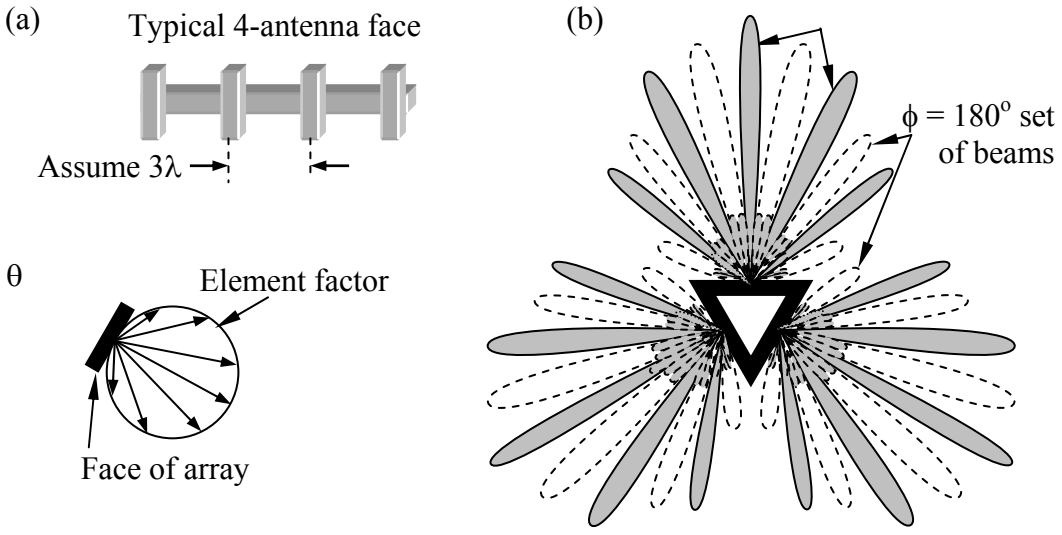

Figure 10.4.8 Cellular base station antenna patterns with frequency reuse.

As before, ϕ is the phase angle difference introduced between adjacent antenna elements. Inter-antenna separations of 3λ result in only 5 main lobes per face, because the two peaks in the plane

of each face are approximately zero for typical element factors. Between each pair of peaks there are two small sidelobes, approximately 14 dB weaker as shown above.

These two sets ($\phi = 0$, $\phi = \pi$) can share the same frequencies because digital communication techniques can tolerate overlapping signals if one is more than ~10-dB weaker. Since each face of the antenna can be connected simultaneously to two independent receivers and two independent transmitters, as many as six calls could simultaneously use the same frequency band, two per face. A single face would not normally simultaneously transmit and receive the same frequency, however. The lobe positions can also be scanned in angle by varying ϕ so as to fill any nulls. Designing such antennas to maximize frequency reuse requires care and should be tailored to the distribution of users within the local environment. In unobstructed environments there is no strong limit to the number of elements and independent beams that can be used per face, or to the degree of frequency reuse. Moreover, half the beams could be polarized one way, say right-circular or horizontal, and the other half could be polarized with the orthogonal polarization, thereby doubling again the number of possible users of the same frequencies. Polarization diversity works poorly for cellular phones, however, because users orient their dipole antennas as they wish.

In practice, most urban cellular towers do not currently phase their antennas as shown above because many environments suffer from severe *multipath* effects where reflected versions of the same signals arrive at the receiving tower from many angles with varying delays. The result is that at each antenna element the phasors arriving from different directions with different phases and amplitudes will add to produce a net signal amplitude that can be large or small. As a result one of the elements facing a particular direction may have a signal-to-interference ratio that is more than 10 dB stronger than another for this reason alone, even though the antenna elements are only a few wavelengths away in an obstacle-free local environment. Signals have different differential delays at different frequencies and therefore their peak summed values at each antenna element are frequency dependent. The antenna-use strategy in this case is to assign users to frequencies and single elements that are observed to be strong for that user, so that another user could be overlaid on the same frequency while using a different antenna element pointed in the same direction. The same frequency-reuse strategy also works when transmitting because of reciprocity.

That signal strengths are frequency dependent in multipath environments is easily seen by considering an antenna receiving both the direct line-of-sight signal with delay t_1 and a reflected second signal with comparable strength and delay t_2. If the differential lag $c(t_2 - t_1) = n\lambda = D$ for integer n, then the two signals will add in phase and reinforce each other. If the lag $D = (2n + 1)\lambda/2$, then they will partially or completely cancel. If $D = 10\lambda$ and the frequency f increases by 10 percent, then the lag measured in wavelengths will also change 10 percent as the sum makes a full peak-to-peak cycle with a null between. Thus the gap between frequency nulls is $\sim\Delta f = f(\lambda/D) = c/D$ Hz. The depth of the null depends on the relative magnitudes of the two rays that interfere. As the number of rays increases the frequency structure becomes more complex. This phenomenon of signals fading in frequency and time as paths and frequencies change is called *multipath fading*.

6.013 Electromagnetics and Applications
Spring 2009