SECOND EDITION

MODERN ANTENNA DESIGN

THOMAS A. MILLIGAN

MODERN ANTENNA DESIGN

Second Edition

THOMAS A. MILLIGAN

IEEE PRESS

WILEY-INTERSCIENCE

A JOHN WILEY & SONS, INC., PUBLICATION

MODERN ANTENNA DESIGN

MODERN ANTENNA DESIGN

Second Edition

THOMAS A. MILLIGAN

IEEE PRESS

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Milligan, Thomas A.
Modern antenna design / by Thomas A. Milligan.—2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN-13 978-0-471-45776-3 (cloth)
ISBN-10 0-471-45776-0 (cloth)
1. Antennas (Electronics)–Design and construction. I. Title.

TK7871.6.M54 2005
621.382′4—dc22 2004059098

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

To Mary, Jane, and Margaret

CONTENTS

PREFACE

I wrote this book from my perspective as a designer in industry, primarily for other designers and users of antennas. On occasion I have prepared and taught antenna courses, for which I developed a systematic approach to the subject. For the last decade I have edited the "Antenna Designer's Notebook" column in the IEEE antenna magazine. This expanded edition adds a combination of my own design notebook and the many other ideas provided to me by others, leading to this collection of ideas that I think designers should know.

The book contains a systematic approach to the subject. Every author would like to be read from front to back, but my own career assignments would have caused to me to jump around in this book. Nevertheless, Chapter 1 covers those topics that every user and designer should know. Because I deal with complete antenna design, which includes mounting the antenna, included are the effects of nearby structures and how they can be used to enhance the response. We all study ideal antennas floating in free space to help us understand the basics, but the real world is a little different.

Instead of drawing single line graphs of common relationships between two parameters, I generated scales for calculations that I perform over and over. I did not supply a set of computer programs because I seldom use collections supplied by others, and younger engineers find my programs quaint, as each generation learns a different computer language. You'll learn by writing your own.

IEEE Antennas and Propagation Society's digital archive of all material published from 1952 to 2000 has changed our approach to research. I have not included extensive bibliographies, because I believe that it is no longer necessary. The search engine of the archive can supply an exhaustive list. I referred only to papers that I found particularly helpful. Complete sets of the transactions are available in libraries, whereas the wealth of information in the archive from conferences was not. I have started mining this information, which contains many useful design ideas, and have incorporated some of them in this book. In this field, 40-year-old publications are still useful and we should not reinvent methods. Many clever ideas from industry are usually published

only once, if at all, and personally, I'll be returning to this material again and again, because all books have limited space.

As with the first edition, I enjoyed writing this book because I wanted to express my point of view of a rewarding field. Although the amount of information available is overwhelming and the mathematics describing it can cloud the ideas, I hope my explanations help you develop new products or use old ones.

I would like to thank all the authors who taught me this subject by sharing their ideas, especially those working in industry. On a personal note I thank the designers at Lockheed Martin, who encouraged me and reviewed material while I wrote: in particular, Jeannette McDonnell, Thomas Cencich, Donald Huebner, and Julie Huffman.

THOMAS A. MILLIGAN

1

PROPERTIES OF ANTENNAS

One approach to an antenna book starts with a discussion of how antennas radiate. Beginning with Maxwell's equations, we derive electromagnetic waves. After that lengthy discussion, which contains a lot of mathematics, we discuss how these waves excite currents on conductors. The second half of the story is that currents radiate and produce electromagnetic waves. You may already have studied that subject, or if you wish to further your background, consult books on electromagnetics. The study of electromagnetics gives insight into the mathematics describing antenna radiation and provides the rigor to prevent mistakes. We skip the discussion of those equations and move directly to practical aspects.

It is important to realize that antennas radiate from currents. Design consists of controlling currents to produce the desired radiation distribution, called its *pattern*. In many situations the problem is how to prevent radiation from currents, such as in circuits. Whenever a current becomes separated in distance from its return current, it radiates. Simply stated, we design to keep the two currents close together, to reduce radiation. Some discussions will ignore the current distribution and instead, consider derived quantities, such as fields in an aperture or magnetic currents in a slot or around the edges of a microstrip patch. You will discover that we use any concept that provides insight or simplifies the mathematics.

An antenna converts bound circuit fields into propagating electromagnetic waves and, by reciprocity, collects power from passing electromagnetic waves. Maxwell's equations predict that any time-varying electric or magnetic field produces the opposite field and forms an electromagnetic wave. The wave has its two fields oriented orthogonally, and it propagates in the direction normal to the plane defined by the perpendicular electric and magnetic fields. The electric field, the magnetic field, and the direction of propagation form a right-handed coordinate system. The propagating wave field intensity decreases by $1/R$ away from the source, whereas a static field

Modern Antenna Design, Second Edition, By Thomas A. Milligan

drops off by $1/R^2$. Any circuit with time-varying fields has the capability of radiating to some extent.

We consider only time-harmonic fields and use phasor notation with time dependence $e^{j\omega t}$. An outward-propagating wave is given by $e^{-j(kR-\omega t)}$, where k, the wave number, is given by $2\pi/\lambda$. λ is the wavelength of the wave given by c/f, where c is the velocity of light (3×10^8 m/s in free space) and f is the frequency. Increasing the distance from the source decreases the phase of the wave.

Consider a two-wire transmission line with fields bound to it. The currents on a single wire will radiate, but as long as the ground return path is near, its radiation will nearly cancel the other line's radiation because the two are 180° out of phase and the waves travel about the same distance. As the lines become farther and farther apart, in terms of wavelengths, the fields produced by the two currents will no longer cancel in all directions. In some directions the phase delay is different for radiation from the current on each line, and power escapes from the line. We keep circuits from radiating by providing close ground returns. Hence, high-speed logic requires ground planes to reduce radiation and its unwanted crosstalk.

1-1 ANTENNA RADIATION

Antennas radiate spherical waves that propagate in the radial direction for a coordinate system centered on the antenna. At large distances, spherical waves can be approximated by plane waves. Plane waves are useful because they simplify the problem. They are not physical, however, because they require infinite power.

The Poynting vector describes both the direction of propagation and the power density of the electromagnetic wave. It is found from the vector cross product of the electric and magnetic fields and is denoted **S**:

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}^* \qquad \text{W/m}^2$$

Root mean square (RMS) values are used to express the magnitude of the fields. $\mathbf{H}^*$ is the complex conjugate of the magnetic field phasor. The magnetic field is proportional to the electric field in the far field. The constant of proportion is η, the impedance of free space ($\eta = 376.73\,\Omega$):

$$|\mathbf{S}| = S = \frac{|\mathbf{E}|^2}{\eta} \qquad \text{W/m}^2 \tag{1-1}$$

Because the Poynting vector is the vector product of the two fields, it is orthogonal to both fields and the triplet defines a right-handed coordinate system: (**E, H, S**).

Consider a pair of concentric spheres centered on the antenna. The fields around the antenna decrease as $1/R$, $1/R^2$, $1/R^3$, and so on. Constant-order terms would require that the power radiated grow with distance and power would not be conserved. For field terms proportional to $1/R^2$, $1/R^3$, and higher, the power density decreases with distance faster than the area increases. The energy on the inner sphere is larger than that on the outer sphere. The energies are not radiated but are instead concentrated around the antenna; they are near-field terms. Only the $1/R^2$ term of the Poynting vector ($1/R$ field terms) represents radiated power because the sphere area grows as R^2 and

gives a constant product. All the radiated power flowing through the inner sphere will propagate to the outer sphere. The sign of the input reactance depends on the near-field predominance of field type: electric (capacitive) or magnetic (inductive). At resonance (zero reactance) the stored energies due to the near fields are equal. Increasing the stored fields increases the circuit Q and narrows the impedance bandwidth.

Far from the antenna we consider only the radiated fields and power density. The power flow is the same through concentric spheres:

$$4\pi R_1^2 S_{1,\text{avg}} = 4\pi R_2^2 S_{2,\text{avg}}$$

The average power density is proportional to $1/R^2$. Consider differential areas on the two spheres at the same coordinate angles. The antenna radiates only in the radial direction; therefore, no power may travel in the θ or ϕ direction. Power travels in flux tubes between areas, and it follows that not only the average Poynting vector but also every part of the power density is proportional to $1/R^2$:

$$S_1 R_1^2 \sin\theta \, d\theta \, d\phi = S_2 R_2^2 \sin\theta \, d\theta \, d\phi$$

Since in a radiated wave S is proportional to $1/R^2$, E is proportional to $1/R$. It is convenient to define radiation intensity to remove the $1/R^2$ dependence:

$$U(\theta, \phi) = S(R, \theta, \phi) R^2 \qquad \text{W/solid angle}$$

Radiation intensity depends only on the direction of radiation and remains the same at all distances. A probe antenna measures the relative radiation intensity (pattern) by moving in a circle (constant R) around the antenna. Often, of course, the antenna rotates and the probe is stationary.

Some patterns have established names. Patterns along constant angles of the spherical coordinates are called either *conical* (constant θ) or *great circle* (constant ϕ). The great circle cuts when $\phi = 0^\circ$ or $\phi = 90^\circ$ are the principal plane patterns. Other named cuts are also used, but their names depend on the particular measurement positioner, and it is necessary to annotate these patterns carefully to avoid confusion between people measuring patterns on different positioners. Patterns are measured by using three scales: (1) linear (power), (2) square root (field intensity), and (3) decibels (dB). The dB scale is used the most because it reveals more of the low-level responses (sidelobes).

Figure 1-1 demonstrates many characteristics of patterns. The half-power beamwidth is sometimes called just the beamwidth. The tenth-power and null beamwidths are used in some applications. This pattern comes from a parabolic reflector whose feed is moved off the axis. The vestigial lobe occurs when the first sidelobe becomes joined to the main beam and forms a shoulder. For a feed located on the axis of the parabola, the first sidelobes are equal.

1-2 GAIN

Gain is a measure of the ability of the antenna to direct the input power into radiation in a particular direction and is measured at the peak radiation intensity. Consider the

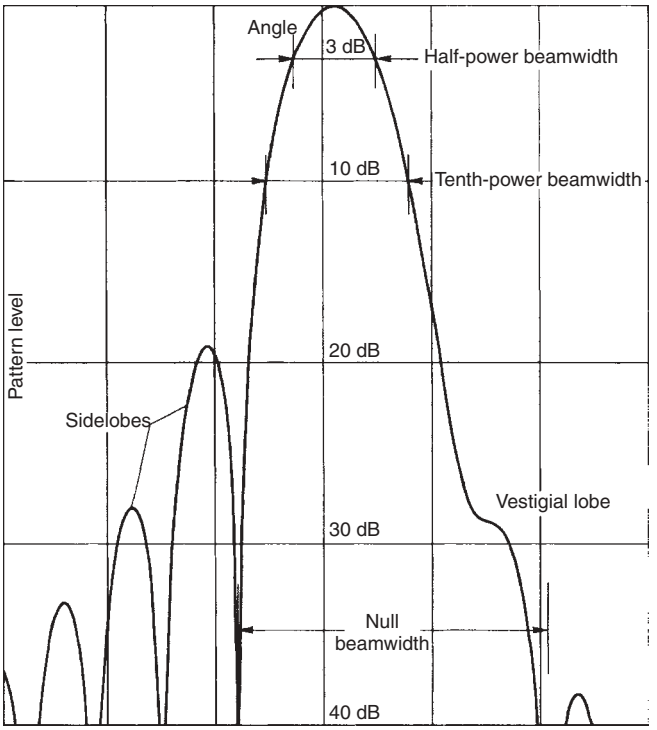


FIGURE 1-1 Antenna pattern characteristics.

power density radiated by an isotropic antenna with input power P_0 at a distance R: $S = P_0/4\pi R^2$. An isotropic antenna radiates equally in all directions, and its radiated power density S is found by dividing the radiated power by the area of the sphere $4\pi R^2$. The isotropic radiator is considered to be 100% efficient. The gain of an actual antenna increases the power density in the direction of the peak radiation:

$$S = \frac{P_0 G}{4\pi R^2} = \frac{|\mathbf{E}|^2}{\eta} \quad \text{or} \quad |\mathbf{E}| = \frac{1}{R}\sqrt{\frac{P_0 G \eta}{4\pi}} = \sqrt{S\eta} \tag{1-2}$$

Gain is achieved by directing the radiation away from other parts of the radiation sphere. In general, gain is defined as the gain-biased pattern of the antenna:

$$S(\theta, \phi) = \frac{P_0 G(\theta, \phi)}{4\pi R^2} \qquad \text{power density}$$

$$U(\theta, \phi) = \frac{P_0 G(\theta, \phi)}{4\pi} \qquad \text{radiation intensity} \tag{1-3}$$

The surface integral of the radiation intensity over the radiation sphere divided by the input power P_0 is a measure of the relative power radiated by the antenna, or the antenna efficiency:

$$\frac{P_r}{P_0} = \int_0^{2\pi} \int_0^{\pi} \frac{G(\theta, \phi)}{4\pi} \sin\theta \, d\theta \, d\phi = \eta_e \qquad \text{efficiency}$$

where P_r is the radiated power. Material losses in the antenna or reflected power due to poor impedance match reduce the radiated power. In this book, integrals in the equation above and those that follow express concepts more than operations we perform during design. Only for theoretical simplifications of the real world can we find closed-form solutions that would call for actual integration. We solve most integrals by using numerical methods that involve breaking the integrand into small segments and performing a weighted sum. However, it is helpful that integrals using measured values reduce the random errors by averaging, which improves the result.

In a system the transmitter output impedance or the receiver input impedance may not match the antenna input impedance. Peak gain occurs for a receiver impedance conjugate matched to the antenna, which means that the resistive parts are the same and the reactive parts are the same magnitude but have opposite signs. Precision gain measurements require a tuner between the antenna and receiver to conjugate-match the two. Alternatively, the mismatch loss must be removed by calculation after the measurement. Either the effect of mismatches is considered separately for a given system, or the antennas are measured into the system impedance and mismatch loss is considered to be part of the efficiency.

Example Compute the peak power density at 10 km of an antenna with an input power of 3 W and a gain of 15 dB.

First convert dB gain to a ratio: $G = 10^{15/10} = 31.62$. The power spreads over the sphere area with radius 10 km or an area of $4\pi(10^4)^2$ m^2. The power density is

$$S = \frac{(3\text{ W})(31.62)}{4\pi \times 10^8\text{ m}^2} = 75.5\text{ nW/m}^2$$

We calculate the electric field intensity using Eq. (1-2):

$$|\mathbf{E}| = \sqrt{S\eta} = \sqrt{(75.5 \times 10^{-9})(376.7)} = 5333\,\mu\text{V/m}$$

Although gain is usually relative to an isotropic antenna, some antenna gains are referred to a $\lambda/2$ dipole with an isotropic gain of 2.14 dB.

If we approximate the antenna as a point source, we compute the electric field radiated by using Eq. (1-2):

$$E(\theta, \phi) = \frac{e^{-jkR}}{R}\sqrt{\frac{P_0 G(\theta, \phi)\eta}{4\pi}} \tag{1-4}$$

This requires only that the antenna be small compared to the radial distance R. Equation (1-4) ignores the direction of the electric field, which we define as *polarization*. The units of the electric field are volts/meter. We determine the far-field pattern by multiplying Eq. (1-4) by R and removing the phase term e^{-jkR} since phase has meaning only when referred to another point in the far field. The far-field electric field E_{ff} unit is volts:

$$E_{\text{ff}}(\theta, \phi) = \sqrt{\frac{P_0 G(\theta, \phi)\eta}{4\pi}} \quad \text{or} \quad G(\theta, \phi) = \frac{1}{P_0}\left[E_{\text{ff}}(\theta, \phi)\sqrt{\frac{4\pi}{\eta}}\right]^2 \tag{1-5}$$

During analysis, we often normalize input power to 1 W and can compute gain easily from the electric field by multiplying by a constant $\sqrt{4\pi/\eta} = 0.1826374$.

1-3 EFFECTIVE AREA

Antennas capture power from passing waves and deliver some of it to the terminals. Given the power density of the incident wave and the effective area of the antenna, the power delivered to the terminals is the product

$$P_d = SA_{\text{eff}} \tag{1-6}$$

For an aperture antenna such as a horn, parabolic reflector, or flat-plate array, effective area is physical area multiplied by aperture efficiency. In general, losses due to material, distribution, and mismatch reduce the ratio of the effective area to the physical area. Typical estimated aperture efficiency for a parabolic reflector is 55%. Even antennas with infinitesimal physical areas, such as dipoles, have effective areas because they remove power from passing waves.

1-4 PATH LOSS [1, p. 183]

We combine the gain of the transmitting antenna with the effective area of the receiving antenna to determine delivered power and path loss. The power density at the receiving antenna is given by Eq. (1-3), and the received power is given by Eq. (1-6). By combining the two, we obtain the path loss:

$$\frac{P_d}{P_t} = \frac{A_2 G_1(\theta, \phi)}{4\pi R^2}$$

Antenna 1 transmits, and antenna 2 receives. If the materials in the antennas are linear and isotropic, the transmitting and receiving patterns are identical (reciprocal) [2, p. 116]. When we consider antenna 2 as the transmitting antenna and antenna 1 as the receiving antenna, the path loss is

$$\frac{P_d}{P_t} = \frac{A_1 G_2(\theta, \phi)}{4\pi R^2}$$

Since the responses are reciprocal, the path losses are equal and we can gather and eliminate terms:

$$\frac{G_1}{A_1} = \frac{G_2}{A_2} = \text{constant}$$

Because the antennas were arbitrary, this quotient must equal a constant. This constant was found by considering the radiation between two large apertures [3]:

$$\frac{G}{A} = \frac{4\pi}{\lambda^2} \tag{1-7}$$

We substitute this equation into path loss to express it in terms of the gains or effective areas:

$$\frac{P_d}{P_t} = G_1 G_2 \left(\frac{\lambda}{4\pi R}\right)^2 = \frac{A_1 A_2}{\lambda^2 R^2} \tag{1-8}$$

We make quick evaluations of path loss for various units of distance R and for frequency f in megahertz using the formula

$$\text{path loss(dB)} = K_U + 20\log(fR) - G_1(\text{dB}) - G_2(\text{dB}) \tag{1-9}$$

where K_U depends on the length units:

Unit	K_U
km	32.45
nm	37.80
miles	36.58
m	−27.55
ft	−37.87

Example Compute the gain of a 3-m-diameter parabolic reflector at 4 GHz assuming 55% aperture efficiency.

Gain is related to effective area by Eq. (1-7):

$$G = \frac{4\pi A}{\lambda^2}$$

We calculate the area of a circular aperture by $A = \pi(D/2)^2$. By combining these equations, we have

$$G = \left(\frac{\pi D}{\lambda}\right)^2 \eta_a = \left(\frac{\pi D f}{c}\right)^2 \eta_a \tag{1-10}$$

where D is the diameter and η_a is the aperture efficiency. On substituting the values above, we obtain the gain:

$$G = \left[\frac{3\pi(4 \times 10^9)}{0.3 \times 10^9}\right]^2 (0.55) = 8685 \quad (39.4\,\text{dB})$$

Example Calculate the path loss of a 50-km communication link at 2.2 GHz using a transmitter antenna with a gain of 25 dB and a receiver antenna with a gain of 20 dB.

$$\text{Path loss} = 32.45 + 20\log[2200(50)] - 25 - 20 = 88.3\,\text{dB}$$

What happens to transmission between two apertures as the frequency is increased? If we assume that the effective area remains constant, as in a parabolic reflector, the transmission increases as the square of frequency:

$$\frac{P_d}{P_t} = \frac{A_1 A_2}{R^2}\frac{1}{\lambda^2} = \frac{A_1 A_2}{R^2}\left(\frac{f}{c}\right)^2 = Bf^2$$

where B is a constant for a fixed range. The receiving aperture captures the same power regardless of frequency, but the gain of the transmitting antenna increases as the square of frequency. Hence, the received power also increases as frequency squared. Only for antennas, whose gain is a fixed value when frequency changes, does the path loss increase as the square of frequency.

1-5 RADAR RANGE EQUATION AND CROSS SECTION

Radar operates using a double path loss. The radar transmitting antenna radiates a field that illuminates a target. These incident fields excite surface currents that also radiate

to produce a second field. These fields propagate to the receiving antenna, where they are collected. Most radars use the same antenna both to transmit the field and to collect the signal returned, called a *monostatic* system, whereas we use separate antennas for *bistatic* radar. The receiving system cannot be detected in a bistatic system because it does not transmit and has greater survivability in a military application.

We determine the power density illuminating the target at a range R_T by using Eq. (1-2):

$$S_{\text{inc}} = \frac{P_T G_T(\theta, \phi)}{4\pi R_T^2} \tag{1-11}$$

The target's radar cross section (RCS), the scattering area of the object, is expressed in square meters or dBm^2: 10 log(square meters). The RCS depends on both the incident and reflected wave directions. We multiply the power collected by the target with its receiving pattern by the gain of the effective antenna due to the currents induced:

$$\text{RCS} = \sigma = \frac{\text{power}_{\text{reflected}}}{\text{power density incident}} = \frac{P_s(\theta_r, \phi_r, \theta_i, \phi_i)}{P_T G_T / 4\pi R_T^2} \tag{1-12}$$

In a communication system we call P_s the *equivalent isotropic radiated power* (EIRP), which equals the product of the input power and the antenna gain. The target becomes the transmitting source and we apply Eq. (1-2) to find the power density at the receiving antenna at a range R_R from the target. Finally, the receiving antenna collects the power density with an effective area A_R. We combine these ideas to obtain the power delivered to the receiver:

$$P_{\text{rec}} = S_R A_R = \frac{A_R P_T G_T \sigma(\theta_r, \phi_r, \theta_i, \phi_i)}{(4\pi R_T^2)(4\pi R_R^2)}$$

We apply Eq. (1-7) to eliminate the effective area of the receiving antenna and gather terms to determine the bistatic radar range equation:

$$\frac{P_{\text{rec}}}{P_T} = \frac{G_T G_R \lambda^2 \sigma(\theta_r, \phi_r, \theta_i, \phi_i)}{(4\pi)^3 R_T^2 R_R^2} \tag{1-13}$$

We reduce Eq. (1-13) and collect terms for monostatic radar, where the same antenna is used for both transmitting and receiving:

$$\frac{P_{\text{rec}}}{P_T} = \frac{G^2 \lambda^2 \sigma}{(4\pi)^3 R^4}$$

Radar received power is proportional to $1/R^4$ and to G^2.

We find the approximate RCS of a flat plate by considering the plate as an antenna with an effective area. Equation (1-11) gives the power density incident on the plate that collects this power over an area A_R:

$$P_C = \frac{P_T G_T(\theta, \phi)}{4\pi R_T^2} A_R$$

The power scattered by the plate is the power collected, P_C, times the gain of the plate as an antenna, G_P:

$$P_s = P_C G_P = \frac{P_T G_T(\theta_i, \phi_i)}{4\pi R_T^2} A_R G_P(\theta_r, \phi_r)$$

This scattered power is the effective radiated power in a particular direction, which in an antenna is the product of the input power and the gain in a particular direction. We calculate the plate gain by using the effective area and find the scattered power in terms of area:

$$P_s = \frac{P_T G_T 4\pi A_R^2}{4\pi R_T^2 \lambda^2}$$

We determine the RCS σ by Eq. (1-12), the scattered power divided by the incident power density:

$$\sigma = \frac{P_s}{P_T G_T / 4\pi R_T^2} = \frac{4\pi A_R^2}{\lambda^2} = \frac{G_R(\theta_i, \phi_i) G_R(\theta_r, \phi_r) \lambda^2}{4\pi} \tag{1-14}$$

The right expression of Eq. (1-14) divides the gain into two pieces for bistatic scattering, where the scattered direction is different from the incident direction. Monostatic scattering uses the same incident and reflected directions. We can substitute any object for the flat plate and use the idea of an effective area and its associated antenna gain. An antenna is an object with a unique RCS characteristic because part of the power received will be delivered to the antenna terminals. If we provide a good impedance match to this signal, it will not reradiate and the RCS is reduced. When we illuminate an antenna from an arbitrary direction, some of the incident power density will be scattered by the structure and not delivered to the antenna terminals. This leads to the division of antenna RCS into the antenna mode of reradiated signals caused by terminal mismatch and the structural mode, the fields reflected off the structure for incident power density not delivered to the terminals.

1-6 WHY USE AN ANTENNA?

We use antennas to transfer signals when no other way is possible, such as communication with a missile or over rugged mountain terrain. Cables are expensive and take a long time to install. Are there times when we would use antennas over level ground? The large path losses of antenna systems lead us to believe that cable runs are better.

Example Suppose that we must choose between using a low-loss waveguide run and a pair of antennas at 3 GHz. Each antenna has 10 dB of gain. The low-loss waveguide has only 19.7 dB/km loss. Table 1-1 compares losses over various distances. The waveguide link starts out with lower loss, but the antenna system soon overtakes it. When the path length doubles, the cable link loss also doubles in decibels, but an antenna link

TABLE 1-1 Losses Over Distance

Distance (km)	Waveguide Loss (dB)	Antenna Path Loss (dB)
2	39.4	88
4	78.8	94
6	118.2	97.6
10	197	102

increases by only 6 dB. As the distance is increased, radiating between two antennas eventually has lower losses than in any cable.

Example A 200-m outside antenna range was set up to operate at 2 GHz using a 2-m-diameter reflector as a source. The receiver requires a sample of the transmitter signal to phase-lock the local oscillator and signal at a 45-MHz difference. It was proposed to run an RG/U 115 cable through the power and control cable conduit, since the run was short. The cable loss was 36 dB per 100 m, giving a total cable loss of 72 dB. A 10-dB coupler was used on the transmitter to pick off the reference signal, so the total loss was 82 dB. Since the source transmitted 100 mW (20 dBm), the signal was −62 dBm at the receiver, sufficient for phase lock.

A second proposed method was to place a standard-gain horn (15 dB of gain) within the beam of the source on a small stand out of the way of the measurement and next to the receiver. If we assume that the source antenna had only 30% aperture efficiency, we compute gain from Eq. (1-10) ($\lambda = 0.15$ m):

$$G = \left(\frac{2\pi}{0.15}\right)^2 (0.3) = 526 \quad (27.2\,\text{dB})$$

The path loss is found from Eq. (1-9) for a range of 0.2 km:

$$32.45 + 20\log[2000(0.2)] - 27.2 - 15 = 42.3\,\text{dB}$$

The power delivered out of the horn is 20 dBm − 42.3 dB = −22.3 dBm. A 20-dB attenuator must be put on the horn to prevent saturation of the receiver (−30 dBm). Even with a short run, it is sometimes better to transmit the signal between two antennas instead of using cables.

1-7 DIRECTIVITY

Directivity is a measure of the concentration of radiation in the direction of the maximum:

$$\text{directivity} = \frac{\text{maximum radiation intensity}}{\text{average radiation intensity}} = \frac{U_{\max}}{U_0} \tag{1-15}$$

Directivity and gain differ only by the efficiency, but directivity is easily estimated from patterns. Gain—directivity times efficiency—must be measured.

The average radiation intensity can be found from a surface integral over the radiation sphere of the radiation intensity divided by 4π, the area of the sphere in steradians:

$$\text{average radiation intensity} = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} U(\theta,\phi)\sin\theta\, d\theta\, d\phi = U_0 \tag{1-16}$$

This is the radiated power divided by the area of a unit sphere. The radiation intensity $U(\theta,\phi)$ separates into a sum of co- and cross-polarization components:

$$U_0 = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} [U_C(\theta,\phi) + U_\times(\theta,\phi)]\sin\theta\, d\theta\, d\phi \tag{1-17}$$

Both co- and cross-polarization directivities can be defined:

$$\text{directivity}_{\text{C}} = \frac{U_{\text{C,max}}}{U_0} \qquad \text{directivity}_{\times} = \frac{U_{\times,\text{max}}}{U_0} \tag{1-18}$$

Directivity can also be defined for an arbitrary direction $D(\theta, \phi)$ as radiation intensity divided by the average radiation intensity, but when the coordinate angles are not specified, we calculate directivity at U_{max}.

1-8 DIRECTIVITY ESTIMATES

Because a ratio of radiation intensities is used to calculate directivity, the pattern may be referred to any convenient level. The most accurate estimate is based on measurements at equal angle increments over the entire radiation sphere. The average may be found from coarse measurements by using numerical integration, but the directivity measured is affected directly by whether the maximum is found. The directivity of antennas with well-behaved patterns can be estimated from one or two patterns. Either the integral over the pattern is approximated or the pattern is approximated with a function whose integral is found exactly.

1-8.1 Pencil Beam

By estimating the integral, Kraus [4] devised a method for pencil beam patterns with its peak at $\theta = 0^\circ$. Given the half-power beamwidths of the principal plane patterns, the integral is approximately the product of the beamwidths. This idea comes from circuit theory, where the integral of a time pulse is approximately the pulse width (3-dB points) times the pulse peak: $U_0 = \theta_1\theta_2/4\pi$, where θ_1 and θ_2 are the 3-dB beamwidths, in radians, of the principal plane patterns:

$$\text{directivity} = \frac{4\pi}{\theta_1\theta_2}(\text{rad}) = \frac{41{,}253}{\theta_1\theta_2}(\text{deg}) \tag{1-19}$$

Example Estimate the directivity of an antenna with E- and H-plane (principal plane) pattern beamwidths of 24° and 36°.

$$\text{Directivity} = \frac{41{,}253}{24(36)} = 47.75 \quad (16.8\,\text{dB})$$

An analytical function, $\cos^{2N}(\theta/2)$, approximates a broad pattern centered on $\theta = 0^\circ$ with a null at $\theta = 180^\circ$:

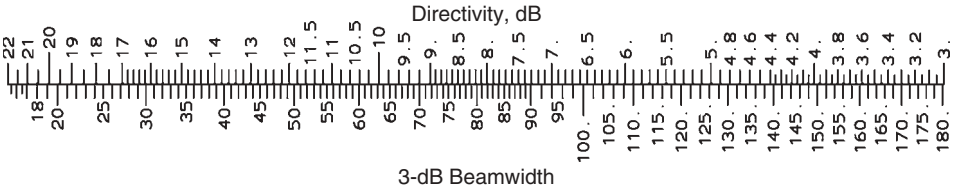
$$U(\theta) = \cos^{2N}(\theta/2) \quad \text{or} \quad E = \cos^{N}(\theta/2)$$

The directivity of this pattern can be computed exactly. The characteristics of the approximation are related to the beamwidth at a specified level, Lvl(dB):

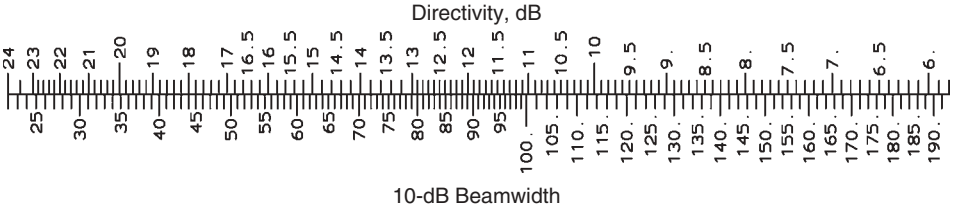
$$\text{beamwidth [Lvl(dB)]} = 4\cos^{-1}(10^{-\text{Lvl(dB)}/20N}) \tag{1-20a}$$

$$N = \frac{-\text{Lvl(dB)}}{20\log[\cos(\text{beamwidth}_{\text{Lvl(dB)}}/4)]} \tag{1-20b}$$

$$\text{directivity} = N + 1 \quad (\text{ratio}) \tag{1-20c}$$



SCALE 1-1 3-dB beamwidth and directivity relationship for $\cos^{2N}(\theta/2)$ pattern.



SCALE 1-2 10-dB beamwidth and directivity relationship for $\cos^{2N}(\theta/2)$ pattern.

Scales 1-1 and 1-2, which give the relationship between beamwidth and directivity using Eq. (1-20), are useful for quick conversion between the two properties. You can use the two scales to estimate the 10-dB beamwidth given the 3-dB beamwidth. For example, an antenna with a 90° 3-dB beamwidth has a directivity of about 7.3 dB. You read from the lower scale that an antenna with 7.3-dB directivity has a 159.5° 10-dB beamwidth. Another simple way to determine the beamwidths at different pattern levels is the square-root factor approximation:

$$\frac{\text{BW[Lvl 2(dB)]}}{\text{BW[Lvl 1(dB)]}} = \sqrt{\frac{\text{Lvl 2(dB)}}{\text{Lvl 1(dB)}}}$$

By this factor, $\text{beamwidth}_{10\,\text{dB}} = 1.826\ \text{beamwidth}_{3\,\text{dB}}$; an antenna with a 90° 3-dB beamwidth has a $(1.826)90° = 164.3°$ 10-dB beamwidth.

This pattern approximation requires equal principal plane beamwidths, but we use an elliptical approximation with unequal beamwidths:

$$U(\theta, \phi) = \cos^{2N_e}(\theta/2)\cos^2\phi + \cos^{2N_h}(\theta/2)\sin^2\phi \tag{1-21}$$

where N_e and N_h are found from the principal plane beamwidths. We combine the directivities calculated in the principal planes by the simple formula

$$\text{directivity (ratio)} = \frac{2 \cdot \text{directivity}_e \cdot \text{directivity}_h}{\text{directivity}_e + \text{directivity}_h} \tag{1-22}$$

Example Estimate the directivity of an antenna with E- and H-plane pattern beamwidths of 98° and 140°.

From the scale we read a directivity of 6.6 dB in the E-plane and 4.37 dB in the H-plane. We convert these to ratios and apply Eq. (1-22):

$$\text{directivity (ratio)} = \frac{2(4.57)(2.74)}{4.57 + 2.74} = 3.426 \quad \text{or} \quad 10\log(3.426) = 5.35\,\text{dB}$$

Many analyses of paraboloidal reflectors use a feed pattern approximation limited to the front hemisphere with a zero pattern in the back hemisphere:

$$U(\theta) = \cos^{2N}\theta \quad \text{or} \quad E = \cos^{N}\theta \qquad \text{for } \theta \le \pi/2(90^\circ)$$

The directivity of this pattern can be found exactly, and the characteristics of the approximation are

$$\text{beamwidth [Lvl(dB)]} = 2\cos^{-1}(10^{-\text{Lvl(dB)}/20N}) \tag{1-23a}$$

$$N = \frac{-\text{Lvl(dB)}}{20\log[\cos(\text{beamwidth}_{\text{Lvl(dB)}}/2)]} \tag{1-23b}$$

$$\text{directivity} = 2(2N+1) \quad \text{(ratio)} \tag{1-23c}$$

We use the elliptical model [Eq. (1-21)] with this approximate pattern and use Eq. (1-22) to estimate the directivity when the E- and H-plane beamwidths are different.

1-8.2 Butterfly or Omnidirectional Pattern

Many antennas have nulls at $\theta = 0^\circ$ with rotational symmetry about the z-axis (Figure 1-2). Neither of the directivity estimates above can be used with these patterns because they require the beam peak to be at $\theta = 0^\circ$. We generate this type of antenna pattern by using mode 2 log-periodic conical spirals, shaped reflectors, some higher-order-mode waveguide horns, biconical horns, and traveling-wave antennas. A formula similar to Kraus's can be found if we assume that all the power is between the 3-dB beamwidth angles θ_1 and θ_2:

$$U_0 = \frac{1}{2}\int_{\theta_1}^{\theta_2} \sin\theta\, d\theta = \frac{\cos\theta_1 - \cos\theta_2}{2}$$

Rotational symmetry eliminates integration over ϕ:

$$\text{directivity} = \frac{U_{\max}}{U_0} = \frac{2}{\cos\theta_1 - \cos\theta_2} \tag{1-24}$$

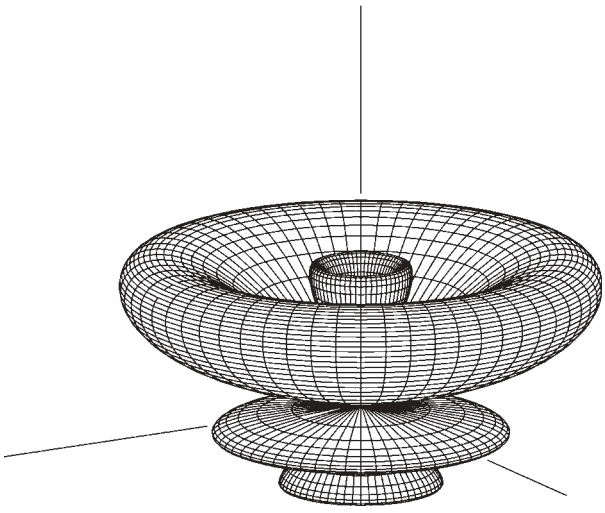
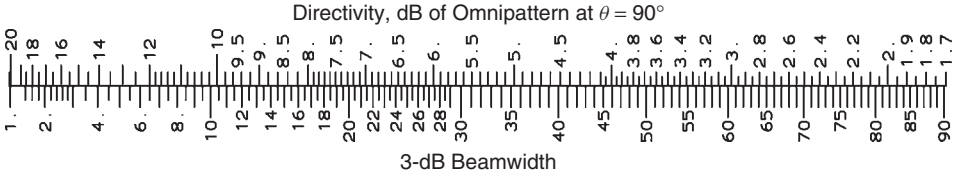


FIGURE 1-2 Omnidirectional antenna pattern with sidelobes scanned above the horizon.



SCALE 1-3 Relationship between 3-dB beamwidth of omnidirectional pattern and directivity.

Example A pattern with rotational symmetry has half-power points at 35° and 75°. Estimate the directivity.

$$\text{Directivity} = \frac{2}{\cos 35^\circ - \cos 75^\circ} = 3.57 \quad (5.5\,\text{dB})$$

If the pattern also has symmetry about the $\theta = 90^\circ$ plane, the integral for the average radiation intensity has limits from 0 to $\pi/2$. Equation (1-24) reduces to directivity $= 1/\cos\theta_1$.

Example A rotationally symmetric pattern with a maximum at 90° has a 45° beamwidth. Estimate the directivity.

$\theta_1 = 90^\circ - 45^\circ/2 = 67.5^\circ$, so

$$\text{directivity} = \frac{1}{\cos 67.5^\circ} = 2.61 \quad (4.2\,\text{dB})$$

The pattern can be approximated by the function

$$U(\theta) = B \sin^{2M}(\theta/2) \cos^{2N}(\theta/2)$$

but the directivity estimates found by integrating this function show only minor improvements over Eq. (1-24). Nevertheless, we can use the expression for analytical patterns. Given beam edges θ_L and θ_U at a level Lvl(dB), we find the exponential factors.

$$AA = \frac{\ln[\cos(\theta_U/2)] - \ln[\cos(\theta_L/2)]}{\ln[\sin(\theta_L/2)] - \ln[\sin(\theta_U/2)]} \quad \text{and} \quad TM_2 = \tan^{-1}\sqrt{AA}$$

$$N = \frac{-|\text{Lvl(dB)}|/8.68589}{AA\{\ln[\sin(\theta_L/2)] - \ln(\sin TM_2)\} + \ln(\cos(\theta_L/2)) - \ln(\cos TM_2)}$$

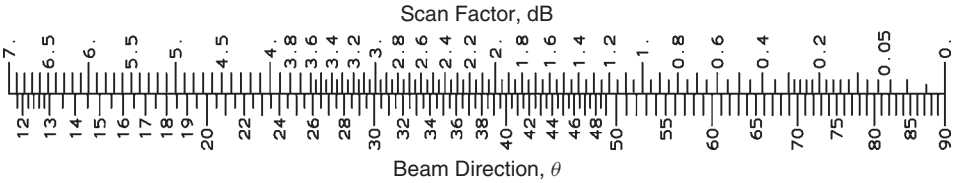
$$M = AA(N)$$

A second pattern model of an omnidirectional pattern based on the pattern function with minor sidelobes and a beam peak at θ_0 measured from the symmetry axis is

$$\frac{\sin[b(\theta_0 - \theta)]}{b(\theta_0 - \theta)}$$

We estimate the directivity from the half-power beamwidth (HPBW) and the beam peak θ_0 [5]:

$$\text{directivity(dB)} = 10 \log \frac{101}{(\text{HPBW} - 0.0027\text{HPBW}^2)\sin\theta_0} \tag{1-25}$$



SCALE 1-4 Additional directivity of omnidirectional pattern when scanned into conical pattern.

Scale 1-3 evaluates this formula for a beam at $\theta_0 = 90^\circ$ given HPBW, and Scale 1-4 gives the additional gain when the beam peak scans toward the axis.

The directivity of butterfly patterns with unequal beamwidths in the principal planes cannot be estimated directly from the foregoing formulas. Similarly, some pencil beam patterns have large sidelobes which decrease the directivity and cannot be estimated accurately from Eq. (1-19). Both problems are solved by considering the directivity as an estimate of the average radiation intensity.

Example A butterfly pattern peak is at 50° in both principal planes, but the beamwidths are 20° and 50°. Estimate the directivity.

The 3-dB pattern points are given by:

Cut 1 (40° and 60°):

$$U_{01} = \frac{\cos 40^\circ - \cos 60^\circ}{2} = 0.133$$

Cut 2 (25° and 75°):

$$U_{02} = \frac{\cos 25^\circ - \cos 75^\circ}{2} = 0.324$$

Average the two pattern integral estimates:

$$U_0 = \frac{0.133 + 0.324}{2} = 0.228$$

$$\text{directivity} = \frac{U_{\max}}{U_0} = \frac{1}{0.228} = 4.38 \quad (6.4\,\text{dB})$$

Suppose that the beams are at different levels on the same pattern. For example, the lobe on the right of the first pattern is the peak and the left lobe is reduced by 3 dB. The peaks of the second pattern are reduced by 1 dB. We can average on one pattern alone. Each lobe contributes $U_{\max}(\cos\theta_1 - \cos\theta_2)/4$ to the integral. The integral of the first pattern is approximated by

$$\frac{0.266 + 0.266 \times 10^{-3/10}}{4} = 0.100$$

The integral of the second pattern is reduced 1 dB from the peak. The average radiation intensity is found by averaging the two pattern averages:

$$U_0 = \frac{0.100 + 0.324 \times 10^{-0.1}}{2} = 0.178$$

$$\text{directivity} = \frac{1}{U_0} = 5.602 \quad (7.5\,\text{dB})$$

Pencil beam patterns with large sidelobes can be averaged in a similar manner: $U_p = 1/\text{directivity}$. By using Eq. (1-19) and assuming equal beamwidths, we have $U_p = \text{HPBW}^2/41,253$, where U_p is the portion of the integral due to the pencil beam and HPBW is the beamwidth in degrees.

Example Consider a pencil beam antenna with pattern beamwidths of 50° and 70° in the principal planes. The second pattern has a sidelobe at $\theta = 60°$ down 5 dB from the peak and a 30° beamwidth below the 5 dB. What is the effect of the sidelobe on the directivity estimate?

Without the sidelobe the directivity estimate is

$$\text{directivity} = \frac{41253}{50(70)} = 11.79 \quad (10.7\,\text{dB})$$

Consider each pattern separately:

$$U_{\text{P1}} = \frac{50^2}{41{,}253} = 0.0606 \qquad U_{\text{P2}} = \frac{70^2}{41{,}253} = 0.1188$$

The sidelobe adds to the second integral:

$$U_{\text{PS2}} = \frac{(\cos 45° - \cos 75°)10^{-5/10}}{4} = 0.0354$$

Averaging the integrals of the parts gives us 0.1074:

$$\text{directivity} = \frac{1}{U_0} = 9.31 \quad (9.7\,\text{dB})$$

If there had been a sidelobe on each side, each would have added to the integral. Estimating integrals in this manner has limited value. Remember that these are only approximations. More accurate results can be obtained by digitizing the pattern and performing numerical integration on each pattern by using Eq. (1-16) or (1-17).

1-9 BEAM EFFICIENCY

Radiometer system designs [6, p. 31–6] specify the antenna in terms of beam efficiency. For a pencil beam antenna with the boresight at $\theta = 0$, the beam efficiency is the ratio (or percent) of the pattern power within a specified cone centered on the boresight to the total radiated power. In terms of the radiation intensity U,

$$\text{beam efficiency} = \frac{\displaystyle\int_0^{\theta_1}\int_0^{2\pi} U(\theta,\phi)\sin\theta\, d\theta\, d\phi}{\displaystyle\int_0^{\pi}\int_0^{2\pi} U(\theta,\phi)\sin\theta\, d\theta\, d\phi} \tag{1-26}$$

where U includes both polarizations if necessary. Extended noise sources, such as radiometry targets, radiate noise into sidelobes of the antenna. Beam efficiency measures the probability of the detected target being located within the main beam ($\theta \le \theta_1$).

Sometimes we can calculate directivity more easily than the pattern everywhere required by the denominator of Eq. (1-26): for example, a paraboloidal reflector. We use Eqs. (1-15) and (1-16) to calculate the denominator integral:

$$\int_0^{\pi}\int_0^{2\pi} U(\theta,\phi)\sin\theta\, d\theta\, d\phi = \frac{4\pi U_{\max}}{\text{directivity}}$$

This reduces Eq. (1-26) to

$$\text{beam efficiency} = \frac{\text{directivity}\displaystyle\int_0^{\theta_1}\int_0^{2\pi} U(\theta,\phi)\sin\theta\, d\theta\, d\phi}{4\pi U_{\max}} \tag{1-27}$$

Equation (1-27) greatly reduces the pattern calculation requirements to compute beam efficiency when the directivity can be found without pattern evaluation over the entire radiation sphere.

1-10 INPUT-IMPEDANCE MISMATCH LOSS

When we fail to match the impedance of an antenna to its input transmission line leading from the transmitter or to the receiver, the system degrades due to reflected power. The input impedance is measured with respect to some transmission line or source characteristic impedance. When the two are not the same, a voltage wave is reflected, ρV, where ρ is the voltage reflection coefficient:

$$\rho = \frac{Z_A - Z_0}{Z_A + Z_0} \tag{1-28}$$

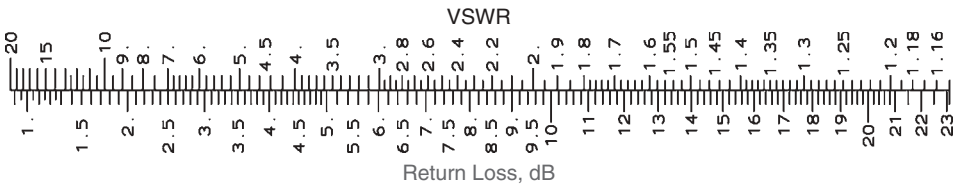
Z_A is the antenna impedance and Z_0 is the measurement characteristic impedance. On a transmission line the two traveling waves, incident and reflected, produce a standing wave:

$$V_{\max} = (1 + |\rho|)V_i \qquad V_{\min} = (1 - |\rho|)V_i \tag{1-29}$$

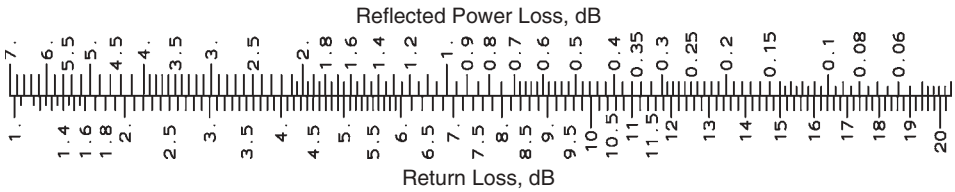
$$\text{VSWR} = \frac{V_{\max}}{V_{\min}} = \frac{1 + |\rho|}{1 - |\rho|} \tag{1-30}$$

VSWR is the voltage standing-wave ratio. We use the magnitude of ρ, a complex phasor, since all the terms in Eq. (1-28) are complex numbers. The reflected power is given by $V_i^2|\rho|^2/Z_0$. The incident power is V_i^2/Z_0. The ratio of the reflected power to the incident power is $|\rho|^2$. It is the returned power ratio. Scale 1-5 gives the conversion between return loss and VSWR:

$$\text{return loss} = -20\log|\rho| \tag{1-31}$$



SCALE 1-5 Relationship between return loss and VSWR.



SCALE 1-6 Reflected power loss due to antenna impedance mismatch.

The power delivered to the antenna is the difference between the incident and the reflected power. Normalized, it is expressed as

$$1 - |\rho|^2 \text{ or reflected power loss(dB)} = 10\log(1 - |\rho|^2) \tag{1-32}$$

The source impedance to achieve maximum power transfer is the complex conjugate of the antenna impedance [7, p. 94]. Scale 1-6 computes the power loss due to antenna impedance mismatch.

If we open-circuit the antenna terminals, the reflected voltage equals the incident voltage. The standing wave doubles the voltage along the transmission line compared to the voltage present when the antenna is loaded with a matched load. We consider the effective height of an antenna, the ratio of the open-circuit voltage to the input field strength. The open-circuit voltage is twice that which appears across a matched load for a given received power. We can either think of this as a transmission line with a mismatch that doubled the incident voltage or as a Thévenin equivalent circuit with an open-circuit voltage source that splits equally between the internal resistor and the load when it is matched to the internal resistor. Path loss analysis predicts the power delivered to a matched load. The mathematical Thévenin equivalent circuit containing the internal resistor does not say that half the power received by the antenna is either absorbed or reradiated; it only predicts the circuit characteristics of the antenna load under all conditions.

Possible impedance mismatch of the antenna requires that we derate the feed cables. The analysis above shows that the maximum voltage that occurs on the cable is twice that present when the cable impedance is matched to the antenna. We compute the maximum voltage given the VSWR using Eq. (1-29) for the maximum voltage:

$$V_{\max} = \frac{2\ \text{VSWR}(V_i)}{\text{VSWR} + 1} = \frac{2V_i}{1 + 1/\text{VSWR}} \tag{1-33}$$

1-11 POLARIZATION

The polarization of a wave is the direction of the electric field. We handle all polarization problems by using vector operations on a two-dimensional space using the far-field radial vector as the normal to the plane. This method is systematic and reduces chance of error. The spherical wave in the far field has only θ and ϕ components of the electric field: $\mathbf{E} = E_\theta\hat{\theta} + E_\phi\hat{\phi}$. E_θ and E_ϕ are phasor components in the direction of the unit vectors $\hat{\theta}$ and $\hat{\phi}$. We can also express the direction of the electric field in terms of a plane wave propagating along the z-axis: $\mathbf{E} = E_x\hat{\mathbf{x}} + E_y\hat{\mathbf{y}}$. The direction of propagation confines the electric field to a plane. Polarization is concerned with methods

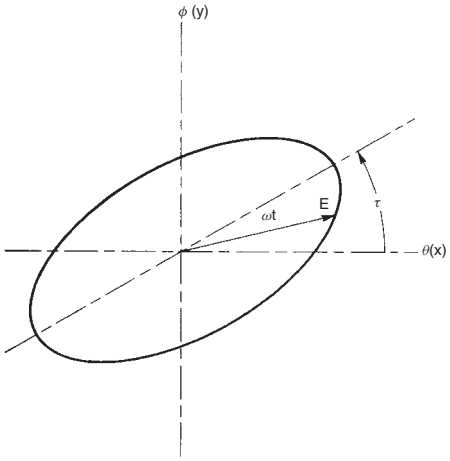


FIGURE 1-3 Polarization ellipse.

of describing this two-dimensional space. Both of the above are linear polarization expansions. We can rewrite them as

$$\mathbf{E} = E_\theta(\hat{\theta} + \hat{\rho}_L\hat{\phi}) \qquad \hat{\rho}_L = \frac{E_\phi}{E_\theta}$$

$$\mathbf{E} = E_x(\hat{\mathbf{x}} + \hat{\rho}_L\hat{\mathbf{y}}) \qquad \hat{\rho}_L = \frac{E_y}{E_x} \tag{1-34}$$

where $\hat{\rho}_L$ is the linear polarization ratio, a complex constant. If time is inserted into the expansions, and the tip of the electric field traced in space over time, it appears as an ellipse with the electric field rotating either clockwise (CW) or counter clockwise (CCW) (Figure 1-3). τ is the tilt of the polarization ellipse measured from the x-axis ($\phi = 0$) and the angle of maximum response. The ratio of the maximum to minimum linearly polarized responses on the ellipse is the axial ratio.

If $\hat{\rho}_L = e^{\pm j\pi/2}$, the ellipse expands to a circle and gives the special case of circular polarization. The electric field is constant in magnitude but rotates either CW (left hand) or CCW (right hand) at the rate ωt for propagation perpendicular to the page.

1-11.1 Circular Polarization Components

The two circular polarizations also span the two-dimensional space of polarization. The right- and left-handed orthogonal unit vectors defined in terms of linear components are

$$\hat{\mathbf{R}} = \frac{1}{\sqrt{2}}(\hat{\theta} - j\hat{\phi}) \quad \text{or} \quad \hat{\mathbf{R}} = \frac{1}{\sqrt{2}}(\hat{\mathbf{x}} - j\hat{\mathbf{y}}) \tag{1-35a}$$

$$\hat{\mathbf{L}} = \frac{1}{\sqrt{2}}(\hat{\theta} + j\hat{\phi}) \quad \text{or} \quad \hat{\mathbf{L}} = \frac{1}{\sqrt{2}}(\hat{\mathbf{x}} + j\hat{\mathbf{y}}) \tag{1-35b}$$

The electric field in the polarization plane can be expressed in terms of these new unit vectors:

$$\mathbf{E} = E_L\hat{\mathbf{L}} + E_R\hat{\mathbf{R}}$$

When projecting a vector onto one of these unit vectors, it is necessary to use the complex conjugate in the scalar (dot) product:

$$E_L = \mathbf{E} \cdot \hat{\mathbf{L}}^* \qquad E_R = \mathbf{E} \cdot \hat{\mathbf{R}}^*$$

When we project $\hat{\mathbf{R}}$ onto itself, we obtain

$$\hat{\mathbf{R}} \cdot \hat{\mathbf{R}}^* = \tfrac{1}{2}(\hat{\theta} - j\hat{\phi}) \cdot (\hat{\theta} + j\hat{\phi}) = \tfrac{1}{2}(1 - j \cdot j) = 1$$

Similarly,

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{R}}^* = \tfrac{1}{2}(\hat{\theta} + j\hat{\phi}) \cdot (\hat{\theta} + j\hat{\phi}) = \tfrac{1}{2}(1 + j \cdot j) = 0$$

The right- and left-handed circular (RHC and LHC) components are orthonormal.

A circular polarization ratio can be defined from the equation

$$\mathbf{E} = E_L(\hat{\mathbf{L}} + \hat{\rho}_c \hat{\mathbf{R}}) \qquad \hat{\rho}_c = \frac{E_R}{E_L} = \rho_c e^{j\delta_c}$$

Let us look at a predominately left-handed circularly polarized wave when time and space combine to a phase of zero for E_L. We draw the polarization as two circles (Figure 1-4). The circles rotate at the rate ωt in opposite directions (Figure 1-5), with the center of the right-handed circular polarization circle moving on the end of the vector of the left-handed circular polarization circle. We calculate the phase of the circular polarization ratio $\hat{\rho}_c$ from the complex ratio of the right- and left-handed circular components. Maximum and minimum electric fields occur when the circles

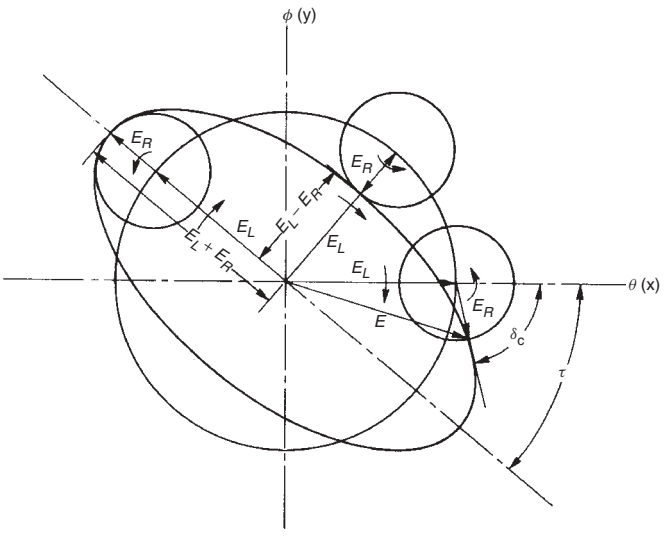


FIGURE 1-4 Polarization ellipse LHC and RHC components. (After J. S. Hollis, T. J. Lyons, and L. Clayton, *Microwave Antenna Measurements,* Scientific Atlanta, 1969, pp. 3–6. Adapted by permission.)

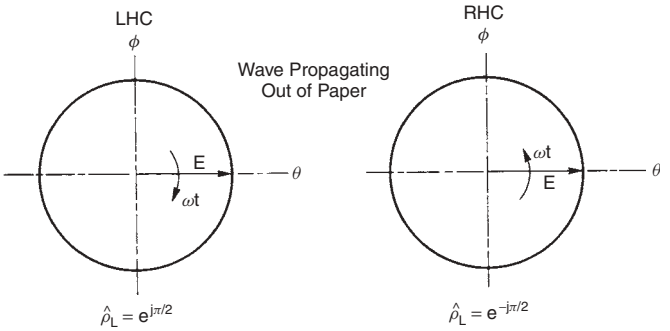
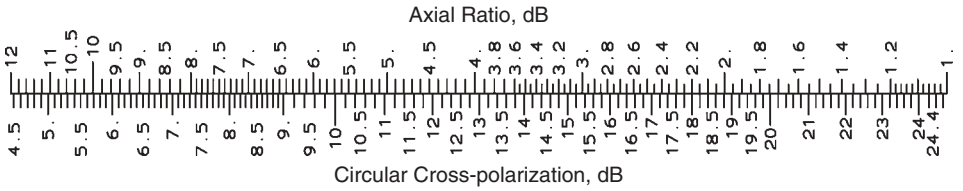


FIGURE 1-5 Circular polarization components. (After J. S. Hollis, T. J. Lyons, and L. Clayton, *Microwave Antenna Measurements,* Scientific Atlanta, 1969, pp. 3–5. Adapted by permission.)



SCALE 1-7 Circular cross-polarization/axial ratio.

alternately add and subtract as shown in Figure 1-4. Scale 1-7 shows the relationship between circular cross-polarization and axial ratio:

$$E_{\max} = (|E_L| + |E_R|)/\sqrt{2} \qquad E_{\min} = (|E_L| - |E_R|)/\sqrt{2}$$

$$\text{axial ratio} = \begin{cases} \dfrac{E_{\max}}{E_{\min}} = \dfrac{|E_L| + |E_R|}{|E_L| - |E_R|} = \dfrac{1 + |\hat{\rho}_c|}{1 - |\hat{\rho}_c|} & \text{LHC} \\ \dfrac{E_{\max}}{E_{\min}} = \dfrac{|E_R| + |E_L|}{|E_R| - |E_L|} = \dfrac{|\hat{\rho}_c| + 1}{|\hat{\rho}_c| - 1} & \text{RHC} \end{cases} \tag{1-36}$$

$$0 \le \begin{cases} |\hat{\rho}_c| < 1 & \text{LHC} \\ \left|\dfrac{1}{\hat{\rho}_c}\right| < 1 & \text{RHC} \end{cases}$$

$$\text{axial ratio(dB)} = 20 \log \frac{E_{\max}}{E_{\min}}$$

The tilt angle of the polarization ellipse τ is one-half δ_c, the phase of $\hat{\rho}_c$. Imagine time moving forward in Figure 1-5. When the LHC vector has rotated $\delta_c/2$ CW, the RHC vector has rotated $\delta_c/2$ CCW and the two align for a maximum.

1-11.2 Huygens Source Polarization

When we project the currents induced on a paraboloidal reflector to an aperture plane, Huygens source radiation induces aligned currents that radiate zero cross-polarization

in the principal planes. We separate feed antenna radiation into orthogonal Huygens sources for this case. To calculate the far-field pattern of a paraboloid reflector, we can skip the step involving currents and integrate over the Huygens source fields in the aperture plane directly. We transform the measured fields of the feed into orthogonal Huygens sources by

$$\begin{bmatrix} E_{\mathrm{c}} \\ E_{\mathrm{x}} \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} E_{\theta f} \\ E_{\phi f} \end{bmatrix} \tag{1-37}$$

where E_{c} is the $\phi = 0$ direction of polarization in the feed pattern and E_{x} is the $\phi = 90^{\circ}$ polarization. This division corresponds to Ludwig's third definition of cross-polarization [8]. The following matrix converts the Huygens source polarizations to the normal far-field components of spherical coordinates:

$$\begin{bmatrix} E_{\theta} \\ E_{\phi} \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} E_{\mathrm{c}} \\ E_{\mathrm{x}} \end{bmatrix} \tag{1-38}$$

1-11.3 Relations Between Bases

In problems with antennas at arbitrary orientations, circularly polarized components have an advantage over linear components. When the coordinate system is rotated, both the amplitude and phase change for $\hat{\rho}_L$, the linear polarization ratio, whereas the circular polarization ratio $\hat{\rho}_c$ magnitude is constant under rotations and only the phase changes. In other words, the ratio of the diameters of the circles (Figure 1-4) is constant.

The circular components can be found from linear polarization components by projection.

$$E_R = (E_\theta \hat{\theta} + E_\phi \hat{\phi}) \cdot \hat{\mathbf{R}}^* = \frac{1}{\sqrt{2}} (E_\theta \hat{\theta} + E_\phi \hat{\phi}) \cdot (\hat{\theta} + j\hat{\phi})$$

$$E_R = \frac{1}{\sqrt{2}} (E_\theta + jE_\phi) \tag{1-39}$$

Similarly,

$$E_L = \frac{1}{\sqrt{2}} (E_\theta - jE_\phi)$$

The linear polarizations can be found in terms of the circular components in the same manner:

$$E_\theta = \frac{1}{\sqrt{2}} (E_L + E_R) \qquad E_\phi = \frac{j}{\sqrt{2}} (E_L - E_R)$$

These relations enable the conversion between polarizations.

Good circularly polarized antennas over a wide bandwidth are difficult to build, but good linearly polarized antennas are obtained easily. After we measure the phase and amplitude of E_θ and E_ϕ component phasors, we compute the circular components from Eq. (1-39), the axial ratio by using Eq. (1-36), and the polarization ellipse tilt τ from one-half the phase of E_R/E_L. We employ a leveled phase-locked source to record two patterns with orthogonal linear sources (or the same linear source is rotated between patterns). Afterward, we use the equations given above to convert polarization

to any desired polarization components. We calculate the maximum and minimum linear components by projecting the linear components into the rotated coordinate system of the polarization ellipse:

$$E_{\max} = E_\theta \cos\tau + E_\phi \sin\tau$$
$$E_{\min} = -E_\theta \sin\tau + E_\phi \cos\tau$$

1-11.4 Antenna Polarization Response

The path loss formulas assume that the two antennas have matched polarizations. Polarization mismatch adds an extra loss. We determine polarization efficiency by applying the scalar (dot) product between normalized polarization vectors. An antenna transmitting in the z-direction has the linear components

$$\mathbf{E}_a = E_1(\hat{\mathbf{x}} + \hat{\rho}_{\mathrm{L}1}\hat{\mathbf{y}})$$

The incident wave on the antenna is given by

$$\mathbf{E}_i = E_2(\hat{\mathbf{x}} + \hat{\rho}_{\mathrm{L}2}\hat{\mathbf{y}})$$

where the wave is expressed in the coordinates of the source antenna. The z-axis of the source is in the direction opposite that of the antenna. It is necessary to rotate the coordinates of the receiving antenna wave. Rotating about the x-axis is equivalent to changing the sign of the tilt angle or taking the complex conjugate of $\mathbf{E}_a$.

The measurement antenna projects the incident wave polarization onto the antenna polarization. The antenna measures the incident field, but we need to normalize the antenna polarization to a unit vector to calculate polarization efficiency:

$$\mathbf{E}_2 \cdot \mathbf{E}_1^* = \frac{E_2 E_1^*(1 + \hat{\rho}_{\mathrm{L}2}\hat{\rho}_{\mathrm{L}1}^*)}{\sqrt{1 + |\hat{\rho}_{\mathrm{L}1}|^2}}$$

We normalize both the incident wave and antenna responses to determine loss due to polarization mismatch:

$$\frac{\mathbf{E}_i}{|\mathbf{E}_i|} = \frac{\hat{\mathbf{x}} + \hat{\rho}_{\mathrm{L}2}\hat{\mathbf{y}}}{\sqrt{1 + \hat{\rho}_{\mathrm{L}2}^*\hat{\rho}_{\mathrm{L}2}}} \qquad \frac{\mathbf{E}_a^*}{|\mathbf{E}_a|} = \frac{\hat{\mathbf{x}} + \hat{\rho}_{\mathrm{L}1}^*\hat{\mathbf{y}}}{\sqrt{1 + \hat{\rho}_{\mathrm{L}1}^*\hat{\rho}_{\mathrm{L}1}}}$$

The normalized voltage response is

$$\frac{\mathbf{E}_i \cdot \mathbf{E}_a^*}{|\mathbf{E}_i||\mathbf{E}_a|} = \frac{1 + \hat{\rho}_{\mathrm{L}1}^*\hat{\rho}_{\mathrm{L}2}}{\sqrt{1 + \hat{\rho}_{\mathrm{L}1}\hat{\rho}_{\mathrm{L}1}^*}\sqrt{1 + \hat{\rho}_{\mathrm{L}2}\hat{\rho}_{\mathrm{L}2}^*}} \tag{1-40}$$

When we express it as a power response, we obtain the polarization efficiency Γ:

$$\Gamma = \frac{|\mathbf{E}_i \cdot \mathbf{E}_a^*|^2}{|\mathbf{E}_i|^2|\mathbf{E}_a|^2} = \frac{1 + |\hat{\rho}_{\mathrm{L}1}|^2|\hat{\rho}_{\mathrm{L}2}|^2 + 2|\hat{\rho}_{\mathrm{L}1}||\hat{\rho}_{\mathrm{L}2}|\cos(\delta_1 - \delta_2)}{(1 + |\hat{\rho}_{\mathrm{L}1}|^2)(1 + |\hat{\rho}_{\mathrm{L}2}|^2)} \tag{1-41}$$

This is the loss due to polarization mismatch. Given that δ_1 and δ_2 are the phases of the polarization ratios of the antenna and the incident wave. As expressed in terms of linear polarization ratios, the formula is awkward because when the antenna is rotated to determine the peak response, both the amplitudes and phases change. A formula

using circular polarization ratios would be more useful, because only phase changes under rotation.

Two arbitrary polarizations are orthogonal ($\Gamma = 0$) only if

$$|\hat{\rho}_1| = \frac{1}{|\hat{\rho}_2|} \quad \text{and} \quad \delta_1 - \delta_2 = \pm 180^\circ \tag{1-42}$$

This can be expressed as vectors by using unit vectors: $\mathbf{a}_1 \cdot \mathbf{a}_2^* = 0$; $\mathbf{a}_1$ and $\mathbf{a}_2$ are the orthonormal generalized basis vectors for polarization. We can define polarization in terms of this basis with a polarization ratio ρ. By paralleling the analysis above for linear polarizations, we obtain the polarization efficiency for an arbitrary orthonormal polarization basis:

$$\Gamma = \frac{1 + |\hat{\rho}_1|^2|\hat{\rho}_2|^2 + 2|\hat{\rho}_1||\hat{\rho}_2|\cos(\delta_1 - \delta_2)}{(1 + |\hat{\rho}_1|^2)(1 + |\hat{\rho}_2|^2)} \tag{1-43}$$

It has the same form as Eq. (1-41) derived for linear polarizations.

We can use Eq. (1-43) with circular polarizations whose polarization ratio ρ_c magnitudes are constant with rotations of the antenna. The maximum and minimum polarization efficiencies occur when $\delta_1 - \delta_2$ equals 0° and 180°, respectively. The polarization efficiency becomes

$$\Gamma_{\max/\min} = \frac{(1 \pm |\hat{\rho}_1||\hat{\rho}_2|)^2}{(1 + |\hat{\rho}_1|^2)(1 + |\hat{\rho}_2|^2)} \tag{1-44}$$

In all other vector pair bases for polarization, the magnitude of the polarization ratio ρ changes under rotations.

Figure 1-6 expresses Eq. (1-44) as a nomograph. If we have fixed installations, we can rotate one antenna until the maximum response is obtained and realize minimum polarization loss. In transmission between mobile antennas such as those mounted on missiles or satellites, the orientation cannot be controlled and the maximum polarization loss must be used in the link analysis. Circularly polarized antennas are used in these cases.

Example A satellite telemetry antenna is RHC with an axial ratio of 7 dB. The ground station is RHC with a 1.5-dB axial ratio. Determine the polarization loss.

Because the orientation of the satellite is unknown, we must use the maximum polarization loss. To find it, use the RHC ends of the scales in Figure 1-6. Draw a line from 7 on the leftmost scale to 1.5 on the center scale. Read the loss on the scale between: 0.9 dB. The measured cross-polarization response of a linearly polarized antenna is the reciprocal of the axial ratio, the same absolute magnitude in decibels.

Example Suppose that the linear cross-polarization responses of two antennas in a stationary link are given as 10 and 20 dB. Compute the minimum polarization loss.

We rotate one of the antennas until the maximum response is found. The specification of cross-polarization response does not state whether an antenna is predominately left- or right-handed circularly polarized. It must be one or the other. Suppose that the 20-dB cross-polarization antenna is LHC. If the other antenna also is LHC, we use a line drawn from the lower portion of the center scale in Figure 1-6 to the rightmost LHC scale and read 0.2 dB of loss on the scale between the two. The second possibility is

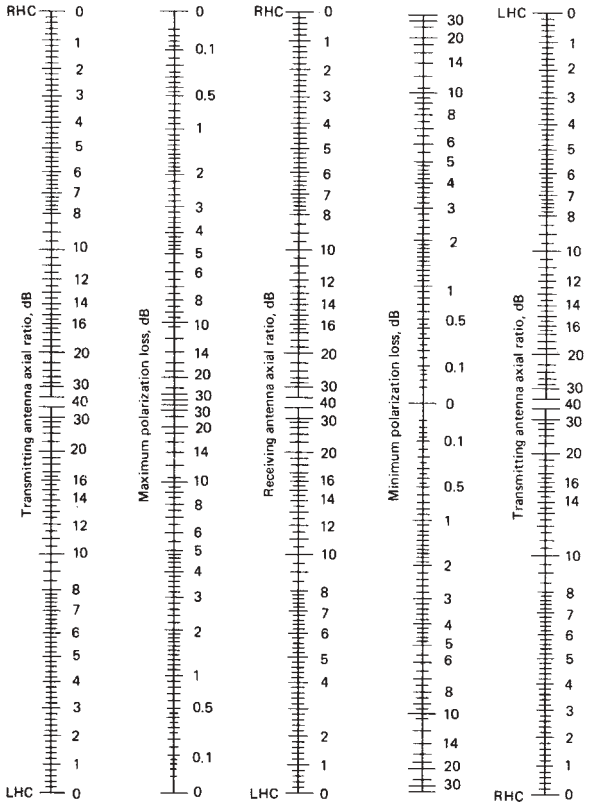


FIGURE 1-6 Maximum and minimum polarization loss. (After A. C. Ludwig, A simple graph for determining polarization loss, *Microwave Journal,* vol. 19, no. 9, September 1976, p. 63.)

that the antenna could be predominately RHC. On drawing a line to the RHC (lower) scale, we read 0.7 dB on the center scale. When polarization is expressed in terms of linearly polarized components, it is ambiguous to give only magnitudes and no information of the circular polarization sense.

1-11.5 Phase Response of Rotating Antennas

The polarization sense of an antenna can be determined from the phase slope of a rotating antenna. Before starting the phase measurement, determine that the setup is proper. Some older phase–amplitude receivers are ambiguous, depending on whether the local oscillator frequency was above or below the signal frequency. We use the convention that increased distance between antennas gives decreased phase. Move the antenna away from the source and observe decreasing phase or correct the setup. A rotating linearly polarized source field is given by

$$E_s = E_2(\cos\alpha\hat{\mathbf{x}} + \sin\alpha\hat{\mathbf{y}})$$

where α is clockwise rotation viewed from the direction of propagation (forward). A horizontally polarized linear antenna has the response $E_a = E_1\hat{\mathbf{x}}$. It responds to the

rotating linear source field, $E_1E_2\cos\alpha$. The phase is constant under rotation until the null is passed and it flips 180° through the null.

An RHC polarized antenna has the response $E_1(\hat{\mathbf{x}} - j\hat{\mathbf{y}})$. It responds to the rotating linear source field,

$$E_1E_2(\cos\alpha - j\sin\alpha) = E_1E_2e^{-j\alpha}$$

The magnitude remains constant, but the phase decreases with rotation. Phase increases when the antenna is LHC. By observing the phase slope, the sense of the predominant polarization can be determined: RHC = negative phase slope; LHC = positive phase slope. It is easily remembered by considering the basis vectors of circular polarization:

$$\hat{\mathbf{R}} = \frac{1}{\sqrt{2}}(\hat{\mathbf{x}} - j\hat{\mathbf{y}})$$

In rotation from the x-axis to the y-axis, the phase decreases 90°.

1-11.6 Partial Gain

If we measure the antenna gain to one polarization (e.g., RHC) and operate it in a link with an antenna also measured to one polarization, Eq. (1-44) fails to predict the response. Polarization efficiency assumes that the antenna gain was measured using a source field with matched polarization. Gains referred to a single polarization are partial gains. If we align the two polarization ellipses of the two antennas, the response increases. Similarly, when the ellipses are crossed, the link suffers polarization loss. To obtain the full gain, we add the factor

$$10\log(1 + |\rho|^2) \tag{1-45}$$

to the partial gain, an expression valid using ρ for either circular or linear polarization. In terms of axial ratio A for circular polarization, the conversion is

$$20\log\frac{\sqrt{2(1 + A^2)}}{1 + A}$$

When using measured partial gains for both antennas, the range of polarization efficiency is given by

$$\text{polarization efficiency } \Gamma = 20\log(1 \pm \rho_1\rho_2) \tag{1-46}$$

We can convert Eq. (1-46) to expressions that use the axial ratio of the two antennas:

$$\text{maximum polarization efficiency} = 20\log\frac{2(A_1A_2 + 1)}{(A_1 + 1)(A_2 + 1)}$$

$$\text{minimum polarization efficiency} = 20\log\frac{2(A_1 + A_2)}{(A_1 + 1)(A_2 + 1)}$$

1-11.7 Measurement of Circular Polarization Using Amplitude Only

The analyses given above assume that you can measure both amplitude and phase response of antennas, whereas in some cases only amplitude can be measured. If

you do not know the sense of circular polarization, it will be necessary to build two antennas that are identical except for their circular polarization sense. For example, you can build two identically sized counter-wound helical wire antennas. You determine polarization sense by using both sources and comparing measured levels. Once you establish the polarization sense, mount a linearly polarized measurement antenna with low cross-polarization. For a given pointing direction of the antenna under test, rotate the source antenna and record the maximum and minimum levels. The ratio of the maximum to the minimum is the axial ratio.

To measure gain, rotate the measurement linearly polarized antenna to determine the peak response. Replace the antenna under test with a linearly polarized gain standard (horn) and perform a gain comparison measurement. Given the antenna axial ratio A, you adjust the linearly polarized gain by the correction factor:

$$\text{gain correction factor(dB)} = 20 \log \frac{A+1}{\sqrt{2}A} \tag{1-47}$$

We obtain the RHC and LHC response from

$$E_R = \frac{1}{\sqrt{2}}(E_{\max} + E_{\min}) \quad \text{and} \quad E_L = \frac{1}{\sqrt{2}}(E_{\max} - E_{\min})$$

assuming that the antenna is predominately RHC.

1-12 VECTOR EFFECTIVE HEIGHT

The vector effective height relates the open-circuit voltage response of an antenna to the incident electric field. Although we normally think of applying effective height to a line antenna, such as a transmitting tower, the concept can be applied to any antenna. For a transmitting tower, effective height is the physical height multiplied by the ratio of the average current to the peak current:

$$V_{\text{OC}} = \mathbf{E}_i \cdot \mathbf{h}^* \tag{1-48}$$

The vector includes the polarization properties of the antenna. Remember from our discussion of antenna impedance mismatch that the open-circuit voltage V_{OC} is twice that across a matched load Z_L for a given received power: $V_{\text{OC}} = 2\sqrt{P_{\text{rec}} Z_L}$. The received power is the product of the incident power density S and the effective area of the antenna, A_{eff}. Gathering terms, we determine the open-circuit voltage from the incident field strength E and a polarization efficiency Γ:

$$V_{\text{OC}} = 2E\sqrt{\frac{Z_L A_{\text{eff}} \Gamma}{\eta}}$$

We calculate polarization efficiency by using the scalar product between the normalized incident electric field and the normalized vector effective height:

$$\Gamma = \frac{|\mathbf{E}_i \cdot \mathbf{h}^*|^2}{|\mathbf{E}_i|^2 |\mathbf{h}|^2} \tag{1-49}$$

Equation (1-49) is equivalent to Eq. (1-41) because both involve the scalar product between the incident wave and the receiving polarization, but the expressions have different normalizations. You can substitute vector effective height of the transmitting antenna for the incident wave in Eq. (1-49) and calculate polarization efficiency between two antennas. When an antenna rotates, we rotate **h**. We could describe polarization calculations in terms of vector effective height, which would parallel and repeat the discussion given in Section 1-11. We relate the magnitude of the effective height h to the effective area A_{eff} and the load impedance Z_L:

$$h = 2\sqrt{\frac{Z_L A_{\text{eff}}}{\eta}} \tag{1-50}$$

The mutual impedance in the far field between two antennas can be found from the vector effective heights of both antennas [9, p. 6–9]. Given the input current I_1 to the first antenna, we find the open-circuit voltage of the second antenna:

$$Z_{12} = \frac{(V_2)_{\text{OC}}}{I_1} = \frac{jk\eta e^{-jkr}}{4\pi r}\mathbf{h}_1 \cdot \mathbf{h}_2^* \tag{1-51}$$

When we substitute Eq. (1-50) into Eq. (1-51) and gather terms, we obtain a general expression for the normalized mutual impedance of an arbitrary pair of antennas given the gain of each in the direction of the other antenna as a function of spacing r:

$$\frac{Z_{12}}{\sqrt{Z_{L1}Z_{L2}}} = \frac{j\sqrt{G_1 G_2}}{kr} e^{-jkr} \frac{\mathbf{h}_1 \cdot \mathbf{h}_2^*}{|\mathbf{h}_1||\mathbf{h}_2|} \tag{1-52}$$

The magnitude of mutual impedance increases when the gain increases or the distance decreases. Of course, Eq. (1-52) is based on a far-field equation and gives only an approximate answer, but it produces good results for dipoles spaced as close as 1λ. Figure 1-7 gives a plot of Eq. (1-52) for isotropic gain antennas with matched polarizations which shows the $1/R$ amplitude decrease with distance and that resistance and

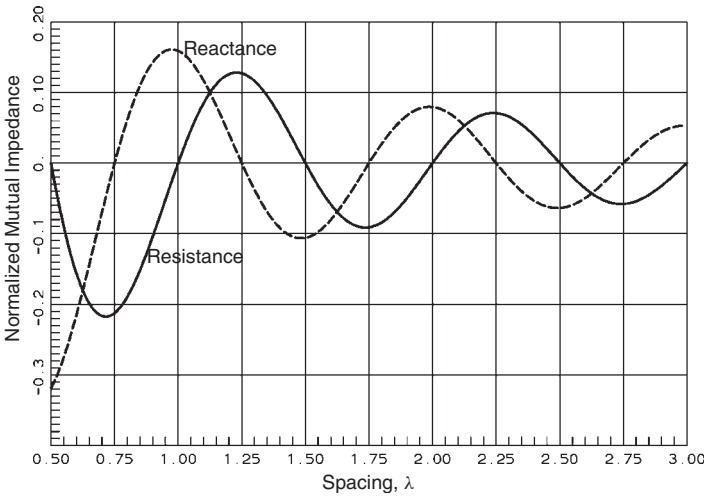


FIGURE 1-7 Normalized mutual impedance (admittance) from the vector effective length for two antennas with 0 dB gain along the line between them.

reactance curves are shifted out of phase. The cosine and sine factors of the complex exponential produce this effect. We multiply these curves by the product of the antenna gains, but the increased gain from larger antennas means that it is a greater distance to the far field. When we bring two antennas close together, the currents on each antenna radiate and excite additional currents on the other that modify the result given by Eq. (1-52). But as we increase the distance, these induced current effects fade. Equivalent height analysis can be repeated using magnetic currents (e.g., used with microstrip patches), and Eqs. (1-51) and (1-52) become mutual admittance. Figure 1-7 is also valid for these antennas when we substitute normalized mutual admittance for normalized mutual impedance. For antennas with pattern nulls directed toward each other, the mutual impedance decreases at the rate $1/R^2$, due to the polarization of current direction **h.**

1-13 ANTENNA FACTOR

The EMC community uses an antenna connected to a receiver such as a spectrum analyzer, a network analyzer, or an RF voltmeter to measure field strength E. Most of the time these devices have a load resistor Z_L that matches the antenna impedance. The incident field strength E_i equals antenna factor AF times the received voltage V_{rec}. We relate this to the antenna effective height:

$$\text{AF} = \frac{E_i}{V_{\text{rec}}} = \frac{2}{h} \tag{1-53}$$

AF has units meter^{-1} but is often given as dB(m^{-1}). Sometimes, antenna factor is referred to the open-circuit voltage and it would be one-half the value given by Eq. (1-53). We assume that the antenna is aligned with the electric field; in other words, the antenna polarization is the electric field component measured:

$$\text{AF} = \sqrt{\frac{\eta}{Z_L A_{\text{eff}}}} = \frac{1}{\lambda}\sqrt{\frac{4\pi}{Z_L G}}$$

This measurement may be corrupted by a poor impedance match to the receiver and any cable loss between the antenna and receiver that reduces the voltage and reduces the calculated field strength.

1-14 MUTUAL COUPLING BETWEEN ANTENNAS

The simplest approach for coupling between antennas is to start with a far-field approximation. We can modify Eq. (1-8) for path loss and add the phase term for the finite distance to determine the S-parameter coupling:

$$S_{21} = \sqrt{G_1 G_2}\frac{e^{-jkr}}{2kr}\frac{\mathbf{E}_1 \cdot \mathbf{E}_2^*}{|\mathbf{E}_1||\mathbf{E}_2|} \tag{1-54}$$

Equation (1-54) includes the polarization efficiency when the transmitted polarization does not match the receiving antenna polarization. We have an additional phase term

because the signal travels from the radiation phase center along equivalent transmission lines to the terminals of each antenna. Equations (1-52) and (1-54) have the same accuracy except that Eq. (1-54) eliminates the need to solve the two-port circuit matrix equation for transmission loss. These formulas assume that antenna size is insignificant compared to the distance between the antennas, and each produces approximately uniform amplitude and phase fields over the second element.

We can improve on Eq. (1-54) when we use the current distribution on one of the two antennas and calculate the near-field fields radiated by the second antenna at the location of these currents. Since currents vary across the receiving antenna, we use vector current densities to include direction: $\mathbf{J}_r$ electric and $\mathbf{M}_r$ magnetic. Although magnetic current densities are fictitious, they simplify the representation of some antennas. We compute coupling from reactance, an integral across these currents [see Eq. (2-34)]:

$$S_{21} = \frac{j}{2\sqrt{P_r P_t}} \iiint (\mathbf{E}_t \cdot \mathbf{J}_r - \mathbf{H}_t \cdot \mathbf{M}_r)\, dV \tag{1-55}$$

The input power to the transmitting antenna P_t produces fields $\mathbf{E}_t$ and $\mathbf{H}_t$. The power P_r into the receiving antenna excites the currents. The scalar product between the incident fields and the currents includes polarization efficiency. If we know the currents on the transmitting antennas, we calculate the near-field pattern response from them at the location of the receiving antenna. Similar to many integrals, Eq. (1-55) is notional because we perform the integral operations only where currents exist. The currents could be on wire segments or surfaces. A practical implementation of Eq. (1-55) divides the currents into patches or line segments and performs the scalar products between the currents and fields on each patch and sums the result. A second form of the reactance [see Eq. (2-35)] involves an integral over a surface surrounding the receiving antenna. In this case each antenna radiates its field to this surface, which requires near-field pattern calculations for both. Equation (1-55) requires adding the phase length between the input ports and the currents, similar to using Eq. (1-54). When we use Eq. (1-55), we assume that radiation between the two antennas excites insignificant additional currents on each other. We improve the answer by using a few iterations of physical optics, which finds induced currents from incident fields (Chapter 2).

We improve on Eq. (1-55) by performing a moment method calculation between the two antennas. This involves subdividing each antenna into small elements excited with simple assumed current densities. Notice the similarity between Eqs. (1-52) and (1-54) and realize that Eq. (1-55) is a near-field version of Eq. (1-54). We use reactance to compute the mutual impedance Z_{21} between the small elements as well as their self-impedance. For the moment method we calculate a mutual impedance matrix with a row and column for each small current element. We formulate a matrix equation using the mutual impedance matrix and an excitation vector to reduce coupling to a circuit problem. This method includes the additional currents excited on each antenna due to the radiation of the other.

1.15 ANTENNA NOISE TEMPERATURE [10]

To a communication or radar system, an antenna contributes noise from two sources. The antenna receives noise power because it looks out on the sky and ground. The

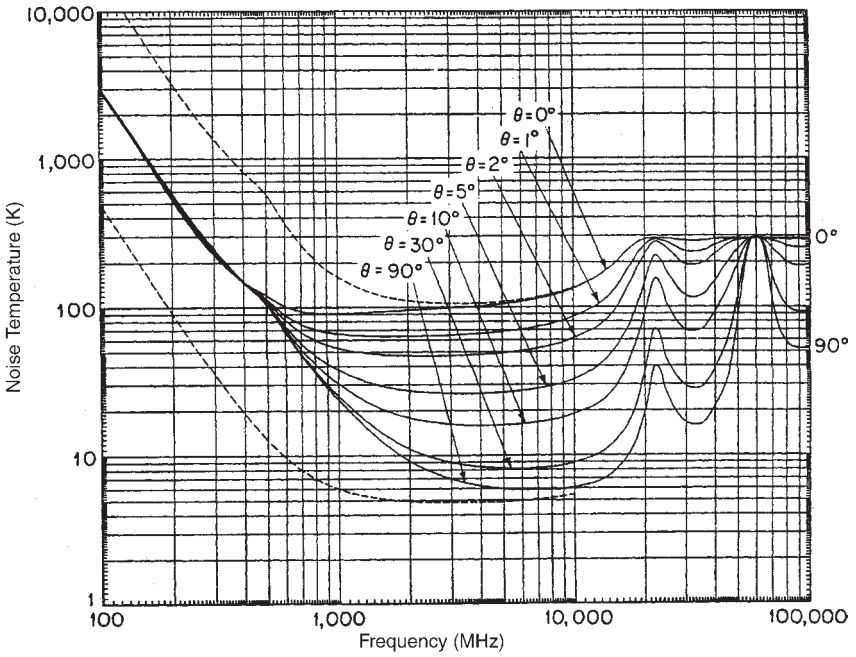


FIGURE 1-8 Antenna sky temperature. Noise temperature of an idealized antenna (lossless, no Earth-directed sidelobes) located at the Earth's surface, as a function of frequency, for a number of beam elevation angles. Solid curves are for geometric-mean galactic temperature, sun noise 10 times quiet level, sun in unity-gain sidelobe, cool temperate-zone troposphere, 2.7 K cosmic blackbody radiation, zero ground noise. The upper dashed curve is for maximum galactic noise (center of galaxy, narrow-beam antenna). Sun noise 100 times quiet level, zero elevation, other factors the same as solid curves. The lower dashed curve is for minimum galactic noise, zero sun noise, 90° elevation angle. (The bump in the curves at about 500 MHz is due to the sun-noise characteristic. The curves for low elevation angles lie below those for high angles at frequencies below 400 MHz because of reduction of galactic noise by atmospheric absorption. The maxima at 22.2 and 60 GHz are due to the water-vapor and oxygen absorption resonance.) (From L. V. Blake, A guide to basic pulse-radar maximum-range calculation, *Naval Research Laboratory Report 5868*, December 1962.)

ground generates noise because it is about 290 K and a portion of the antenna pattern falls on it. Similarly, the sky adds noise dependent on the elevation angle and the operating frequency. Figure 1-8 gives the sky temperature versus frequency and elevation angle. The frequency range of lowest noise occurs in the middle of microwave frequencies of 1 to 12 GHz. The graphs show a large variation between the dashed curves, which occurs because of antenna direction and the pointing relative to the galactic center. In the middle of microwaves the sky noise temperatures are around 50 K, whereas near zenith the temperature is under 10 K. Near the horizon it rises because of the noise from oxygen and water vapor. The exact value must be determined for each application. As frequency decreases below 400 MHz, the sky temperature rises rapidly and becomes independent of antenna pointing. The curve continues the rapid rise at the same slope for lower frequencies. Low-frequency sky temperatures are often given as decibels relative to 290 K.

An antenna receives this blackbody noise from the environment, but the value that affects the communication system depends both on the pattern shape and the direction of the main beam. We determine the antenna noise temperature by integrating the pattern times the environmental noise temperature distribution:

$$T_a = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} G(\theta, \phi) T_s(\theta, \phi) \sin\theta \, d\theta \, d\phi \tag{1-56}$$

where $G(\theta, \phi)$ is the antenna gain pattern and $T_s(\theta, \phi)$ is the angle-dependent blackbody radiation of the environment. Changing the antenna pointing changes T_a. Equation (1-56) is a weighted average of the environment noise temperature, usually referred to as the sky temperature. The second source of noise in the antenna is that of components that have both dissipative losses and reflection losses that generate noise.

A receiving system needs to maximize the signal-to-noise ratio for given resources. System considerations, such as bit error rate, establish the required S/N ratio. We determine the noise power from the product

$$N = k_0 B_n T_e \tag{1-57}$$

where k_0 is Boltzmann's constant (1.38×10^{-23} W/K · Hz $= -228.6$ dB) and B_n is the receiver bandwidth (Hz). T_e is the effective noise temperature (K). When referring noise temperature to other parts of the network, we increase or decrease it by the gain or loss, since it represents power and not a true temperature. Antenna gain is a measure of the signal level, since we can increase gain independent of the noise temperature, although the gain pattern is a factor by Eq. (1-56).

The antenna conductor losses have an equivalent noise temperature:

$$T_e = (L - 1) T_p \tag{1-58}$$

where T_p is the antenna physical temperature and L is the loss (a ratio > 1). From a systems point of view, we include the transmission line run to the first amplifier or mixer of the receiver. We do not include the current distribution losses (aperture efficiencies) that reduce gain in Eq. (1-58) because they are a loss of potential antenna gain and not noise-generating losses (random electrons). The antenna–receiver chain includes mismatch losses, but these do not generate random electrons, only reflected waves, and have a noise temperature of zero. We include them in a cascaded devices noise analysis as an element with loss only.

Noise characteristics of some receiver components are specified as the noise figure F_N (ratio), and cascaded devices' noise analysis can be analyzed using the noise figure, but we will use noise temperature. Convert the noise figures to noise temperature using

$$T_E = (F_N - 1) T_0 \tag{1-59}$$

T_0 is the standard reference temperature 290 K.

We calculate noise temperature for the entire receiver chain of devices at a particular point normally at the input to the first device. To calculate the S/N ratio we use the transmitter power, path loss (including antenna gain and polarization efficiency), and the gains (losses) of any devices for signal to the location in the receiver chain where

noise temperature is being calculated. We characterize a given antenna by the ratio G/T, a measure independent of transmitter power and path loss, but including the receiver noise characteristics. Using the input of the first device as the noise reference point, we calculate the input noise temperature from component noise temperatures and gains:

$$T = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 G_2} + \frac{T_4}{G_1 G_2 G_3} + \cdots \tag{1-60}$$

Equation (1-60) merely states that noise temperatures are powers that decrease when we pass backward through a device with gain G. Each noise term is referred to the input of the device, and we pass backward to all previous devices and reduce noise temperature by $1/G$. If we decided to locate the noise reference point at the input to the second device, the noise initially referred to the chain input would increase by the gain of the first device. The system noise temperature becomes $T_{(2)}$:

$$T_{(2)} = T_1 G_1 + T_2 + \frac{T_3}{G_2} + \frac{T_4}{G_2 G_3} + \cdots$$

The signal also passes through the first device and the new gain at the input to the second device becomes GG_1. The gain and the noise temperature change by the same factor G_1 and produce a constant ratio. By extending these operations to any location in the receiver chain, we show that G/T is constant through the receiver device chain.

It is easiest to illustrate G/T noise calculations with an example. A ground station has a 5-m-diameter paraboloid reflector with 60% aperture efficiency with the system operating at 2.2 GHz ($\lambda = 0.136$ m). We compute antenna directivity using the physical area and aperture efficiency:

$$\text{directivity} = 0.60\left(\frac{\pi \cdot \text{Dia}}{\lambda}\right)^2 = 0.60\left(\frac{5\pi}{0.136}\right)^2 = 7972 \quad (39\,\text{dB})$$

The reflector feed loss is 0.2 dB and it has a VSWR of 1.5 : 1. The cable between the feed and the first amplifier (LNA) has a 0.5-dB loss. These are elements under control of the antenna designer. We calculate the noise temperature of these by using Eq. (1-58) when we use a physical antenna temperature of 37.7°C (100°F) (310.8 K).

Feed loss: $T_1 = (10^{0.2/10} - 1)310.8 = 14.65\,\text{K}$
Feed mismatch: $T_2 = 0\,\text{K}$
Cable: $T_3 = (10^{0.5/10} - 1)310.8 = 37.92\,\text{K}$

The gains of these devices are $G_1 = 10^{-0.2/10} = 0.955$ (feed loss), $G_2 = 10^{-0.18/10} = 0.959$ (reflected power loss for 1.5 : 1 VSWR), and $G_3 = 10^{-0.5/10} = 0.891$ (cable loss). The antenna sees the environment that generates noise due to blackbody radiation from the sky and ground. A typical value for the antenna pointed at 5° elevation is 50 K. This is not a physical temperature but represents an equivalent received power. Remember that the 60% aperture efficiency has no noise or loss contribution, because it only represents the loss of potential gain, since no random electrons are generated.

We must consider the rest of the receiver chain when calculating the total input noise temperature. For this example we assume that the LNA has a noise figure of

2 dB with 20 dB gain. The final portion of the receiver includes the mixer and IF of the receiver, which we assume has a 10-dB noise figure. We use Eq. (1-59) to convert noise figure to noise temperature.

$$\text{LNA noise temperature } T_4 = 290(10^{2/10} - 1) = 162.62\,\text{K}$$

$$\text{Receiver noise temperature } T_5 = 290(10^{10/10} - 1) = 2610\,\text{K}$$

The 20-dB (100) LNA gain greatly reduces the effect of the 2610-K receiver. We calculate the contribution of each device to the input noise temperature by applying Eq. (1-60) to each device. We pass the noise temperature of the receiver through the four devices, and its temperature is reduced by the gain of each device:

$$T_{e5} = \frac{T_5}{G_1 G_2 G_3 G_4} = \frac{2610}{0.955(0.959)(0.891)(100)} = 31.98\,\text{K}$$

The gain of the LNA greatly reduced the effective noise of the receiver at the antenna input. This operation shows that cascading noise temperature involves passing each device's noise temperature through the gains of all preceding devices to the input and reducing it by the product of their gains. Similarly, we perform this operation on all the other noise temperatures.

$$T_{e4} = \frac{T_4}{G_1 G_2 G_3} = \frac{169.62}{0.955(0.959)(0.891)} = 207.86\,\text{K}$$

$$T_{e3} = \frac{T_3}{G_1 G_2} = \frac{37.92}{0.955(0.959)} = 41.40\,\text{K}$$

$$T_{e2} = \frac{T_2}{G_1} = \frac{0}{0.955} = 0$$

$$T_{e1} = T_1 = 14.65$$

These operations illustrate that the cascaded devices' noise temperature equation (1-60) is easily derived by considering the passage of noise temperature (power) through devices with gain to a common point where we can add the contributions.

The sky temperature is not an input noise temperature but the noise power delivered at the fictitious point called the *antenna directivity*, where gain = directivity. Since noise temperature represents power, we convert it to decibels and subtract it from directivity to compute G/T:

$$G/T\,(\text{dB}) = 39 - 10\log(345.9) = 13.6\,\text{dB}$$

This G/T is a measure of the antenna and receiver combined performance when the antenna is pointed to 5° elevation. Changing the pointing direction affects only the sky temperature added directly to the final result. We use G/T in the link budget of the communication system.

We can supply a single value for the antenna gain and noise temperature at the output port connected to the receiver. Recognize that the first three noise temperatures and the sky temperature are associated with the antenna. We moved the noise reference of each device to the input by dividing by the gain of the preceding devices. To move

to the output of the antenna, we increase the noise temperature and the antenna gain by the product of the gain for the devices:

$$\begin{aligned} T &= (T_{\text{sky}} + T_{e1} + T_{e2} + T_{e3})G_1G_2G_3 \\ &= (50 + 14.65 + 0 + 37.92)10^{-0.88/10} = 83.7\,\text{K} \\ \text{gain(dB)} &= \text{directivity(dB)} - 0.88\,\text{dB} = 39 - 0.88 = 38.12\,\text{dB} \end{aligned}$$

This reduces the antenna to a single component similar to the directivity and sky temperature that started our analysis.

1-16 COMMUNICATION LINK BUDGET AND RADAR RANGE

We illustrate communication system design and path loss by considering a sample link budget example. The 5-m-diameter reflector is pointing at a satellite in an orbit 370 km above the Earth with a telemetry antenna radiating 10 W at 2.2 GHz. Since the antenna pattern has to cover the visible Earth, its performance is compromised. Considering the orbit geometry and antenna pointing is beyond the scope of this discussion. The range from a satellite at 370 km to a ground station pointing at 5° is 1720 km. The satellite antenna pointing angle from nadir is 70.3°, and a typical antenna for this application would have gain = −2 dBiC (RHC gain relative to an isotropic antenna) and an axial ratio of 6 dB. Assume that the ground station antenna has a 2-dB axial ratio. We apply the nomograph of Figure 1-6 to read the maximum polarization loss of 0.85 dB since we cannot control the orientation of the polarization ellipses. The link budget needs to show margin in the system, so we take worst-case numbers. When we apply Eq. (1-9) for path loss, we leave out the antenna gains and add them as separate terms in the link budget (Table 1-2):

$$\text{free-space path loss} = 32.45 + 20\log[2200(1720)] = 164\,\text{dB}$$

The link budget shows a 4.4-dB margin, which says that the communication link will be closed. This link budget is only one possible accounting scheme of the system parameters. Everyone who writes out a link budget will separate the parameters differently. This budget shows typical elements.

Radar systems have similar link budgets or detection budgets that consider S/N:

$$S/N = \frac{P_{\text{rec}}}{KTB} = \frac{P_T G_T(\text{directivity})\lambda^2\sigma}{(4\pi)^3 R^4 KTB} = \frac{(\text{EIRP})\lambda^2(G/T)\sigma}{(4\pi)^3 R^4 KB}$$

The radar has a required S/N value to enable it to process the information required, which leads to the maximum range equation:

$$R = \left[\frac{(\text{EIRP})\lambda^2(G/T)\sigma}{(4\pi)^3(S/N)_{\text{req}}KB}\right]^{1/4} \tag{1-61}$$

Equation (1-61) clearly shows the role of the transmitter, EIRP; the receiver and antenna noise; G/T; and the requirement for signal quality, S/N_{req}, on the radar range for a given target size σ.

Equation (1-61) applies to CW radar, whereas most radars use pulses. We increase radar performance by adding many pulses. We ignore the aspects of pulse train encoding

TABLE 1-2 Link Budget

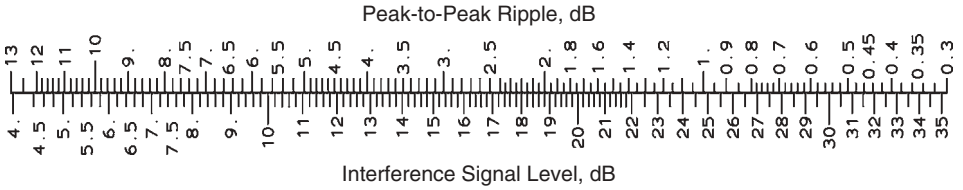
Frequency	2.2 GHz	Information only
Transmit power	10 dBW	10 log(10)
Transmit antenna gain	−2 dB	
EIRP (effective isotropic radiated power)	8 dBW	Transmit power dBW + antenna gain dB
Free-space path loss	164 dB	Isotropic antenna path loss
Polarization loss	0.85	Maximum for uncontrolled orientation
Atmospheric loss	0.30	5° elevation at 2.2 GHz
Rain loss	0.00	Little loss at this frequency
Pointing loss	0.00	
Receive antenna directivity	39 dB	Location in receiver chain for G/T calculation
G/T	13.6 dB	From preceding section
Boltzmann's constant	228.6 dB	
Carrier/noise (C/N) ratio (ignores bandwidth)	85 dB	EIRP + G/T − path loss − polarization loss −atmospheric loss − rain loss + 228.6
Bit rate: 8 Mb/s	69 dB	10 log(bit rate) bandwidth
E_b/N_0 (energy per bit/noise density)	16 dB	$E_b/N_0 = C/N - 10$ log(bit rate)
Implementation loss	2 dB	Groups extra system losses
E_b/N_0 required	9.6 dB	For bit error rate (BER) $= 10^{-5}$ in QPSK
Margin	4.4 dB	E_b/N_0 − required E_b/N_0 − implementation loss

that allow coherent addition. Radar range is determined by the total energy contained in the pulses summed. We replace EIRP with G_T(energy) since $P_T \times$ time = energy. It is the total energy that illuminates the target that determines the maximum detection range. Using antennas in radar leads to speaking of the radiated energy correct for pulsed systems, but when we do not integrate pulse shape times time, the antenna radiates power. To be correct we should call radiation that we integrate over angular space to find power, "power density." To say "energy radiated in the sidelobes" is poor physics unless it is a radar system, because it is power.

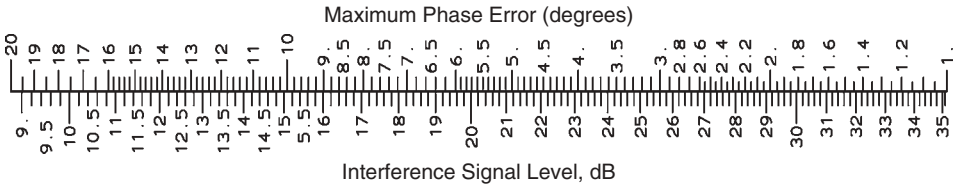
1-17 MULTIPATH

Multipath means that the field intensity at a particular point is the sum of a number of waves that arrive from different directions or from different sources. It arises from signal transmission paths such as edge reflections from the mounting structure around an antenna and general reflections from objects near the antenna. Nearby reflections only seem to modify the antenna pattern, while reflections from additional objects cause rapid ripple with changing pattern angle. In Section 3-1 we discuss how to use the ripple angular rate and pattern distribution to locate its source. Multipath causes degraded system performance or measurement errors. Of course, multipath can improve performance as well. In fact, we add nearby objects, such as ground planes, to improve antenna performance.

We specify pattern response in terms of the power response, but we add fields. An extra signal −20 dB relative to the main signal is 0.01 in power but 0.1 in field strength



SCALE 1-8 Signal peak-to-peak amplitude ripple due to multipath signal.



SCALE 1-9 Peak phase error due to multipath signal.

(voltage). Since the extra signal can have any phase relative to the main signal, it can add or subtract. Given an extra signal MP(dB), the pattern ripple is

$$\text{ripple(dB)} = 20 \log \frac{1 + 10^{\text{MP}/20}}{1 - 10^{\text{MP}/20}} \tag{1-62}$$

where MP(dB) has a negative sign. Scale 1-8 gives the relationship between peak-to-peak amplitude ripple and the level of the multipath signal. Equation (1-62) is numerically the same as the relationship between return loss and 20 log(VSWR). The multipath signal can change the phase when summed with the main signal over a range given by

$$\text{maximum phase error} = \pm \tan^{-1}(10^{\text{MP}/20}) \tag{1-63}$$

Scale 1-9 calculates the peak phase error due to a multipath signal.

1-18 PROPAGATION OVER SOIL

When we position antennas over soil and propagate the signal any significant distance, it will reflect from soil or water and produce a large multipath signal. Soil is a conductive dielectric that reflects horizontally and vertically polarized signals differently. Typical ground constants are listed in Table 1-3. Given the grazing angle ψ measured between the reflected ray and ground, the voltage reflection coefficients are

$$\rho_h = \frac{\sin\psi - \sqrt{\varepsilon_r - jx - \cos^2\psi}}{\sin\psi + \sqrt{\varepsilon_r - jx - \cos^2\psi}} \quad \text{and} \quad \rho_v = \frac{(\varepsilon_r - jx)\sin\psi - \sqrt{\varepsilon_r - jx - \cos^2\psi}}{(\varepsilon_r - jx)\sin\psi + \sqrt{\varepsilon_r - jx - \cos^2\psi}} \tag{1-64}$$

where $x = \sigma/\omega\varepsilon_0 = 17,975\sigma/\text{frequency(MHz)}$.

Figure 1-9 gives the reflection coefficient for the two polarizations versus grazing angle. Horizontal polarization reflects from soil about the same as a metal surface. Vertical polarization reflection produces a more interesting curve. The graph shows that the reflection is low over a region of grazing angles. The minimum reflection direction is called the *Brewster angle*. At this angle the reflected wave is absorbed

TABLE 1-3 Typical Ground Constants

Surface	Dielectric Constant	Conductivity (S)
Dry ground	4–7	0.001
Average ground	15	0.005
Wet ground	25–30	0.020
Fresh water	81	0.010
Seawater	81	5.0

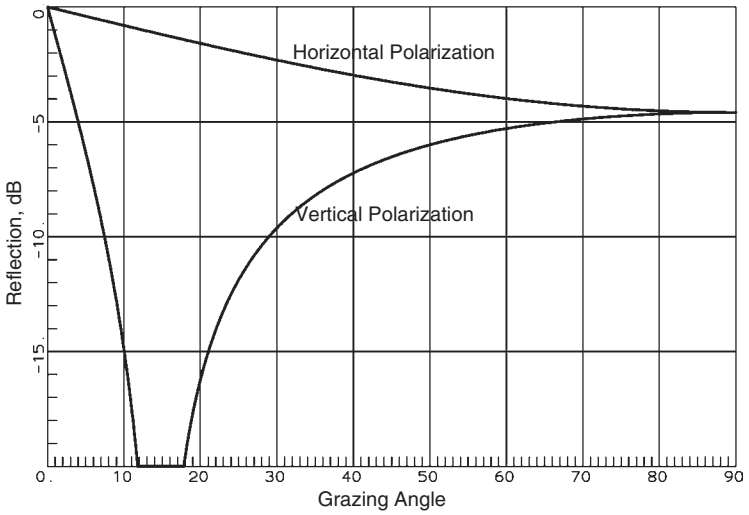


FIGURE 1-9 Average soil reflection for horizontal and vertical polarization.

into the soil. At high grazing angles ρ_h has a phase near 180° and ρ_v a phase of 0°. When the grazing angle decreases and becomes less than the Brewster angle, the vertical polarization reflection changes from 0° to 180°. Remember that for most general response nulls, the signal phase changes by 180° when passing through the transition. As the grazing angle approaches zero both reflection coefficients approach −1 and multipath is independent of polarization.

The electric field at the receiving antenna is the sum of the direct wave plus the reflected wave, which traveled along a longer path:

$$E = E_d[1 - \exp(-j\Delta\phi)] = E_d(1 - \cos\Delta\phi + j\sin\Delta\phi)$$

We compute the magnitude

$$|E| = |E_d|\sqrt{1 + \cos^2\Delta\phi + \sin^2\Delta\phi - 2\cos\Delta\phi} = 2|E_d|\sin(\Delta\phi/2)$$

for the small phase difference between the two equal-amplitude signals. The received power P_{rec} is proportional to E^2. The path loss for this multipath link is modified from the free-space equation:

$$P_{\text{rec}} = 4P_T\left(\frac{\lambda}{4\pi d}\right)^2 G_T G_R \sin^2\frac{2\pi h_T h_R}{\lambda d} \to P_T G_T G_R\left(\frac{h_T h_R}{d^2}\right)^2 \qquad (1\text{-}65)$$

Equation (1-65) states that the power received is proportional to $1/d^4$ and increases by h^2 for either antenna. We can approximate the propagation over soil by a region for closely spaced antennas when the results consist of the free-space transmission with $1/d^2$ average transmission with significant variation due to multipath and a second region proportional to $1/d^4$ with small multipath variations. The breakpoint between the two models occurs at a distance $d = 4h_T h_R/\lambda$.

Experiments at mobile telephone frequencies showed that Eq. (1-65) overestimates the received power when the receiving antenna height is less than 30 m and a more correct model modifies the exponent of h_R [11, p. 38]:

$$P_{\text{rec}} = P_T G_T G_R \frac{h_T^2 h_R^C}{d^4} \tag{1-66}$$

Below 10 m, $C = 1$ and the exponent varies linearly between 10 and 30 m: $C = h_R/20 + \frac{1}{2}$.

On a narrow-beam terrestrial propagation path, scattering from an object along a path an odd multiple of $\lambda/2$ produces a signal that reduces the main path signal. Given an obstacle at a distance h radial from the direct ray path and located d_T from the transmitter and a distance d_R from the receiving antenna, we determine the differential path length as

$$\Delta = \frac{h^2}{2}\frac{d_T + d_R}{d_T d_R} = n\frac{\lambda}{2} \quad \text{or} \quad \text{clearance height } h = \sqrt{\frac{n\lambda d_T d_R}{d_T + d_R}} \tag{1-67}$$

We call these Fresnel clearance zones of order n. The direct path should clear obstacles by at least one clearance zone distance h to prevent the scattered signal from having a negative impact on the communication link. The first Fresnel zone touches ground when $d_T = 4h_T h_R/\lambda$ is the breakpoint distance between $1/d^2$ and $1/d^4$ propagation models.

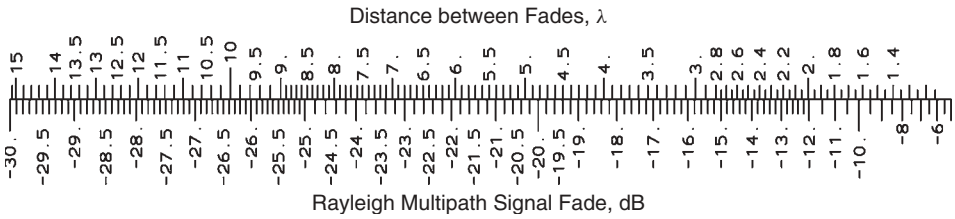
1-19 MULTIPATH FADING

Most mobile communication occurs when there is no direct path between the base station antennas and the mobile user. The signal reflects off many objects along the path between the two. This propagation follows a Rayleigh probability distribution about the mean signal level:

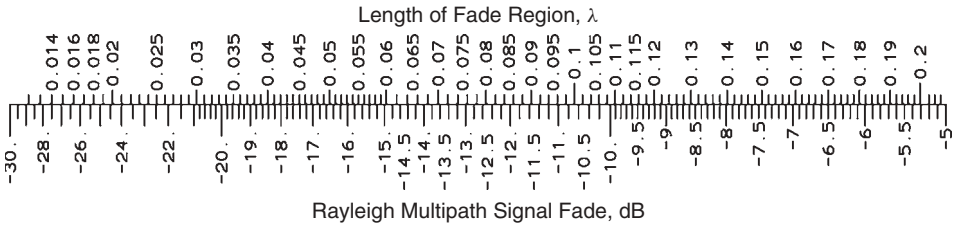
$$p_r(r) = \frac{r}{\alpha^2}\exp\left(-\frac{r^2}{2\alpha^2}\right) \quad \text{prob}[r < R] = P_R(R) = 1 - \exp\left(-\frac{R^2}{2\alpha^2}\right)$$

R is the signal level, α the value of the peak in the distribution, with mean $= \alpha\sqrt{\pi/2}$ and median $R_M = \alpha\sqrt{2\ln(2)} = 1.1774\alpha$. The median signal level is found by fitting measured data for various localities (town, small town, open country, etc.) into a prediction model. The signal will have large signal fades where the level drops rapidly. The Rayleigh model can be solved for the average distance between fades given the level. As a designer it is important to realize the magnitude of the problem [12, pp. 125–130]:

$$\text{average distance between fades} = \lambda\frac{2^{(R/R_M)^2}}{\sqrt{2\pi\ln(2)}(R/R_M)} \tag{1-68}$$



SCALE 1-10 Average distance between fades and depth of fade for a Rayleigh multipath.



SCALE 1-11 Average fade length and depth of fade for a Rayleigh multipath.

R is the fade level (ratio) and R_M is the median signal level found from a propagation model. Scale 1-10 shows the relationship between the average distance between fades and the depth of fade for Rayleigh multipath. A mobile channel operating at 1.85 GHz ($\lambda = 16.2$ cm) has a 15-dB fade every 2.75λ which equals 44.5 cm, while 10-dB fades occur every $1.62\lambda = 26.25$ cm. The communication system must overcome these fades. Fortunately, the deep fades occur over a short distance:

$$\text{average length of fade} = \lambda \frac{2^{(R/R_M)^2} - 1}{\sqrt{2\pi \ln(2)}(R/R_M)}$$

The signal fades and then recovers quickly for a moving user. Scale 1-11 shows the average fade length along a path given the depth of fade. For the 1.85-GHz channel the 15-dB fade occurs only over $0.06(16.2) = 0.97$ cm, and the 10-dB fade length is $0.109(16.2) = 1.76$ cm.

The solution to mobile communication multipath fading is found either in increasing the link margin with higher gain base station antennas or the application of diversity techniques. We use multiple paths between the user and the base station so that while one path experiences a fade, the other one does not. Diversity has no effect on the median signal level, but it reduces the effects of the nulls due to the Rayleigh distribution propagation.

REFERENCES

1. S. A. Schelkunoff and H. Friis, *Antenna Theory and Practice*, Wiley, New York, 1952.
2. R. F. Harrington, *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York, 1961.
3. H. Friis, A note on a simple transmission formula, *Proceedings of IRE*, vol. 34, May 1946, pp. 254–256.

4. J. D. Kraus, *Antennas*, McGraw-Hill, New York, 1950.
5. N. McDonald, Omnidirectional pattern directivity in the presence of minor lobes: revisited, *IEEE Antennas and Propagation Magazine*, vol. 41, no. 2, April 1999, pp. 63–65.
6. W. F. Croswell and M. C. Bailey, in R. C. Johnson and H. Jasik, eds., *Antenna Engineering Handbook*, McGraw-Hill, New York, 1984.
7. C. G. Montgomery, R. H. Dicke, and E. M. Purcell, *Principles of Microwave Circuits*, McGraw-Hill, New York, 1948.
8. A. C. Ludwig, The definition of cross polarization, *IEEE Transactions on Antennas and Propagation*, vol. AP-21, no. 1, January 1973, pp. 116–119.
9. K. P. Park and C. T. Tai, Receiving antennas, Chapter 6 in Y. T. Lo and S. W. Lee, eds., *Antenna Handbook*, Van Nostrand Reinhold, New York, 1993.
10. L. V. Blake, Prediction of radar range, Chapter 2 in *Radar Handbook*, M. Skolnik, ed., McGraw-Hill, New York, 1970.
11. K. Fujimoto and J. R. James, *Mobile Antenna Systems Handbook*, 2nd ed., Artech House, Boston, 2001.
12. J. D. Parsons, *The Mobile Radio Propagation Channel*, Wiley, New York, 1992.

2

RADIATION STRUCTURES AND NUMERICAL METHODS

Antenna analysis, an important part of design, requires a compromise between extensive calculations and the fabrication and measurement of prototypes, which depends on your working environment. You should minimize cost, which means reducing the time from the start of a design to completion of a working model. In some cases you should not rush to build a prototype. For example, when designing large and expensive antennas, such as paraboloidal reflectors, the high fabrication cost justifies the time required for analysis. Management will not let you proceed before knowing the design will work. You should develop a cost model for each design in which analysis is one factor.

Analyses allow optimization of a design. You can design a number of antennas and adjust the dimensions until you find the best one. Again, you should be considering the costs of your time. At some point the incremental improvements are not worth the extra time for further analyses. In any case, when you build the prototype, you can expect differences. You soon determine that you can achieve only limited knowledge about a design because fabrication and measurement errors mask the true response of the antenna. You are doing engineering, not a science project.

Textbooks contain many analyses of ideal antennas, and this book is no exception. You need to consider the application and the final antenna environment. The mounting structure has little effect on the pattern of a large antenna with narrow beamwidth because little radiation strikes it. The overall radiation characteristics of narrow- or wide-beam antennas depend significantly on the shape of the vehicle and how the antenna is mounted. In later chapters we discuss how to use antenna mounting to improve performance, so you can take advantage of it. The size of the mounting structure limits the type of analysis used.

In this chapter we discuss physical optics (PO) and geometric optics (GO) [geometric theory of diffraction (GTD)] for large structures. In physical optics we compute the current induced on the vehicle due to antenna radiation and include their radiation in the

Modern Antenna Design, Second Edition, By Thomas A. Milligan

overall pattern. But the PO analysis cost rises rapidly as the number of small current patches increases for larger structures. PO analysis works well with large antennas, such as paraboloidal reflectors, that produce focused beams. Geometric optics uses ray optics techniques whose computation cost is independent of the size of the vehicle and whose accuracy improves as structure size increases. GO provides insight because we can visualize the combination of direct, reflected, and diffracted (GTD) rays to calculate the pattern, but it requires the solution of difficult geometry problems.

Smaller structures allow the use of multiple methods. For example, the moment method divides the surroundings into small patches and uses an expansion of the current in predetermined basis functions. This method uses integral equations of the boundary conditions to calculate a matrix equation involving coefficients of the current expansion. Numerical methods invert the matrix to solve for the coefficients, but it is a costly numerical operation and limits the size of the problem that can be handled to a few wavelengths. The finite-difference time-domain (FDTD) technique computes the fields on the structure in the time domain. This method handles moderate-sized structures and readily includes complex material properties such as biological features. FDTD divides the region into cubic cells and when excited by pulse feeding functions, it produces wide frequency bandwidth responses. Finite-element methods (FEMs) also divide the problem into cubic cells, but the analysis is performed in the frequency domain. FEM analysis must be repeated at every frequency of interest. FDTD and FEMs require a program to divide the structure into a mesh before starting the solution. Both methods calculate currents on a boundary surface by using the equivalence theorem with the incident fields and then calculate the far-field radiation pattern from these boundary currents.

Most methods start by assuming a current distribution on the antenna or, equivalently, a distribution of fields on an aperture. The fields on the aperture can be reduced to a current distribution. The moment method uses a summation of assumed basis function currents and solves for the coefficients of the expansion, but it, too, starts with assumed currents over small regions. You will discover that the radiation pattern can be found with greater accuracy than the input impedance. For antennas constructed from wires, the moment method computes the input current for a given excitation voltage and we calculate impedance from the ratio. Interaction of an antenna with the currents induced on a structure has little effect on impedance for narrow-beam antennas. Even for wide-beam antennas, such as dipoles, the structure effect on impedance can be found by using source mutual coupling with its images. In the end, antenna impedance should be measured when mounted in the final configuration.

An antenna has both a radiation pattern bandwidth and an impedance bandwidth, but you must give the pattern primary consideration. Too many designs concentrate on the wideband impedance characteristics of an antenna when, in fact, the antenna pattern has changed over the frequency range of the impedance bandwidth. Your primary task should be to design for the radiation pattern desired. In Chapter 1 we detailed the system aspects of impedance mismatch (Section 1-10), and you may determine the overall system impact of small impedance mismatch.

2-1 AUXILIARY VECTOR POTENTIALS

We do not use vector potentials in design. It seems as though they would be useful, but only a few simple antennas fit their direct use. You cannot measure them

because they are not physical entities, so they seem artificial. Physical optics (PO) calculates the radiation directly from currents using dyadic Green's functions but uses long expressions. Nevertheless, many analysis techniques find them more efficient than PO expressions and you should be aware of them. We illustrate their use with a couple of simple antennas.

We use vector potentials to introduce a few antenna concepts. In the first example we apply the magnetic vector potential to calculate the radiation from a short-length current element (dipole) and show how to obtain the pattern. Integration of the radiation pattern power density (Section 1-2) determines the total power radiated. Because we know the input current and the total radiated power, the ratio of the power to the input current squared gives the radiation resistance. We combine the low radiation resistance with the material resistance to compute the antenna efficiency. Electric vector potentials used with fictitious magnetic currents illustrate analysis by duality. We apply this to the analysis of a small loop and show that it has the same pattern as that of a small dipole.

2-1.1 Radiation from Electric Currents

Normal electron currents radiate when time varying. The simplest example is a filamentary current on wire, but we include surface and volumetric current densities as well. We analyze them by using the magnetic vector potential. Far-field electric fields are proportional to the magnetic vector potential **A**:

$$\mathbf{E} = -j\omega\mathbf{A} \tag{2-1}$$

We determine the magnetic field from

$$|\mathbf{E}| = \eta|\mathbf{H}| \tag{2-2}$$

and realizing the cross product of the electric field with the magnetic field points in the direction of power flow, the Poynting vector. Since the electric field direction defines polarization, we usually ignore the magnetic field. We derive the magnetic vector potential from a retarded volume integral over the current density **J**:

$$\mathbf{A} = \mu \iiint \frac{\mathbf{J}(\mathbf{r}')e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}\, dV' \tag{2-3}$$

where $\mathbf{r}$ is the field measurement point radius vector, $\mathbf{r}'$ the source-point radius vector, μ the permeability ($4\pi \times 10^{-7}$ A/m in free space), and k, the wave number, is $2\pi/\lambda$. As written, Eq. (2-3) calculates the potential **A** everywhere: near and far field. The vector potential can be written in terms of a free-space Green's function:

$$g(R) = \frac{e^{-jkR}}{4\pi R} \qquad \text{where} \quad R = |\mathbf{r}-\mathbf{r}'|$$

$$\mathbf{A} = \mu \iiint g(R)\mathbf{J}(\mathbf{r}')\, dV' \tag{2-4}$$

Radiation Approximation When we are interested only in the far-field response of an antenna, we can simplify the integral [Eq. (2-3)]. An antenna must be large in terms

of wavelengths before it can radiate efficiently with gain, but at great distances it still appears as a point source. Consider the radiation from two different parts of an antenna. Far away from the antenna, the ratio of the two distances to the different parts will be nearly 1. The phase shift from each part will go through many cycles before reaching the observation point, and when adding the response from each part, we need only the difference in phase shift. In the radiation approximation we pick a reference point on the antenna and use the distance from that point to the far-field observation point for amplitudes, $1/R$, for all parts of the antenna. The direction of radiation defines a plane through the reference point. This plane is defined by the radius normal vector, given in rectangular coordinates by

$$\hat{\mathbf{r}} = \sin\theta\cos\phi\hat{\mathbf{x}} + \sin\theta\sin\phi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}}$$

We compute the phase difference to the far-field point by dropping a normal to the reference plane from each point on the antenna. This distance multiplied by k, the propagation constant, is the phase difference. Given a point on the antenna $\mathbf{r}'$, the phase difference is $k\mathbf{r}' \cdot \hat{\mathbf{r}}$. When we substitute these ideas into Eq. (2-3), the equation becomes

$$\mathbf{A} = \frac{e^{-jkr}}{4\pi r}\mu \iiint \mathbf{J}' e^{jk\mathbf{r}'\cdot\hat{\mathbf{r}}}\, dV' \tag{2-5}$$

In rectangular coordinates $k\mathbf{r}' \cdot \hat{\mathbf{r}}$ becomes

$$k(x'\sin\theta\cos\phi + y'\sin\theta\sin\phi + z'\cos\theta)$$

We can combine k and $\hat{\mathbf{r}}$ to form a $\mathbf{k}$-space vector:

$$\mathbf{k} = k\hat{\mathbf{r}} = k\sin\theta\cos\phi\hat{\mathbf{x}} + k\sin\theta\sin\phi\hat{\mathbf{y}} + k\cos\theta\hat{\mathbf{z}}$$

and the phase constant becomes $\mathbf{k} \cdot \mathbf{r}'$. Currents in filaments (wires) simplify Eq. (2-5) to a single line integral. Magnetic vector potentials and electric fields are in the same directions as the wires that limit the directions of current. For example, filamentary current along the z-axis produces z-directed electric fields. Spherical waves (far field) have only $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ components found from the projection of E_z onto those axes. Filamentary currents on the z-axis produce only z-directed electric fields with a null from $\hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{z}} = -\sin\theta$ at $\theta = 0$. In turn, x- or y-directed currents produce electric fields depending on the scalar products (projections) of the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ unit vectors onto the $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ vectors in the far field:

$$\hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{x}} = \cos\theta\cos\phi \qquad \hat{\boldsymbol{\phi}} \cdot \hat{\mathbf{x}} = -\sin\phi$$
$$\hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{y}} = \cos\theta\sin\phi \qquad \hat{\boldsymbol{\phi}} \cdot \hat{\mathbf{y}} = \cos\phi$$

By examining antenna structure you can discover some of its characteristics without calculations. Without knowing the exact pattern, we estimate the polarization of the waves by examining the directions of the wires that limit the current density. Consider various axes or planes of symmetry on an antenna: for example, a center-fed wire along the z-axis. If we rotate it about the z-axis, the problem remains the same, which means that all conical polar patterns (constant θ) must be circles; in other words, all great

circle patterns must be the same. An antenna with the same structure above and below the $x-y$ plane radiates the same pattern above and below the $x-y$ plane. Always look for axes and planes of symmetry to simplify the problem.

We can extend the magnetic vector potential [Eq. (2-1)] to determine near fields:

$$\mathbf{E} = -j\omega\mathbf{A} + \frac{\nabla(\nabla \cdot \mathbf{A})}{j\omega\varepsilon\mu}$$

$$\mathbf{H} = \frac{1}{\mu}\nabla \times \mathbf{A} \tag{2-6}$$

The electric field separates into far- and near-field terms, but the equation for the magnetic field, the defining equation of the potential, does not separate. If we substitute the free-space Green's function from Eq. (2-4) into Eq. (2-6), expand, and gather terms, we can determine the fields directly from the electric currents and eliminate the use of a vector potential.

$$\mathbf{E}(\mathbf{r}) = \frac{\eta k^2}{4\pi} \iiint_{V'} \left[\mathbf{J}(\mathbf{r}') \left(-\frac{j}{kR} - \frac{1}{k^2R^2} + \frac{j}{k^3R^3} \right) + [\mathbf{J}(\mathbf{r}') \cdot \hat{\mathbf{R}}]\hat{\mathbf{R}} \left(\frac{j}{kR} + \frac{3}{k^2R^2} - \frac{3j}{k^3R^3} \right) \right] e^{-jkR}\, dV' \tag{2-7}$$

$$\mathbf{H}(\mathbf{r}) = \frac{k^2}{4\pi} \iiint_{V'} \mathbf{J}(\mathbf{r}') \times \hat{\mathbf{R}} \left(\frac{j}{kR} + \frac{1}{k^2R} \right) e^{-jkR}\, dV' \tag{2-8}$$

$$\hat{\mathbf{R}} = \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} = \frac{\mathbf{r} - \mathbf{r}'}{R} \qquad \text{since} \quad R = |\mathbf{r} - \mathbf{r}'|$$

Terms with $1/R$ dependence are the far-field terms. The radiative near-field terms have $1/R^2$ dependence and near-field terms have $1/R^3$ dependence. The impedance of free space, η, is 376.7 Ω. We can rearrange Eqs. (2-7) and (2-8) so that they become the integral of the dot product of the current density $\mathbf{J}$ with dyadic Green's functions [1]. It is only a notation difference that leads to a logic expression. Except for a few examples given below, we leave the use of these expressions to numerical methods when designing antennas.

Example Use the magnetic vector potential to derive the far field of a short-length current element.

Assume a constant current on the wire. The current density is $Il\delta(r')$, where $\delta(r')$ *is* the Dirac delta distribution and l is the length over which the far-field phase is constant. The integral in Eq. (2-4) easily reduces to

$$A_z = \frac{\mu I l e^{-jkr}}{4\pi r}$$

The current element is so short that the phase distances from all parts of the element are considered to be equal; e^{-jkr} is the retarded potential phase term. The electric field is found from A_z using Eq. (2-1):

$$E_z = -j\omega\mu \frac{Il}{4\pi r} e^{-jkr}$$

$$E_\theta = E_z \hat{\mathbf{z}} \cdot \hat{\boldsymbol{\theta}} = j\omega\mu \frac{Il}{4\pi r} e^{-jkr} \sin\theta$$

Evaluate ω as $2\pi f$, split μ in $\sqrt{\mu}\sqrt{\mu}$, and divide and multiply by $\sqrt{\varepsilon}$:

$$E_\theta = \frac{jIl2\pi f\sqrt{\mu\varepsilon}}{4\pi r}\sqrt{\frac{\mu}{\varepsilon}}e^{-jkr}\sin\theta$$

The following terms can be recognized as

$$c = \frac{1}{\sqrt{\mu\varepsilon}} \qquad \frac{f}{c} = \frac{l}{\lambda} \qquad \eta = \sqrt{\frac{\mu}{\varepsilon}}$$

The far-field electric field becomes

$$E_\theta = \frac{jIl\eta}{2\lambda r}e^{-jkr}\sin\theta$$

The magnetic field is found from the electric field using Eq. (2-2):

$$H_\phi = \frac{E_\theta}{\eta} = \frac{jIl}{2\lambda r}e^{-jkr}\sin\theta$$

The term j can be evaluated as $e^{j\pi/2}$, a phase shift term. The power density S_r is

$$S_r = E_\theta H_\phi^* = \frac{|I|^2 l^2}{4\lambda^2 r^2}\eta\sin^2\theta$$

The normalized power pattern is equal to $\sin^2\theta$. Figure 2-1 gives the polar pattern of this antenna as a dashed plot. The dashed circle is the -3-dB pattern level. We measure the angular separation between the 3-dB points to determine the beamwidth (half-power beamwidth). For comparison, Figure 2-1 shows the pattern of a half-wavelength-long dipole as a solid curve. At a length about 5% shorter than a half wavelength, the reactive component of the impedance vanishes. The figure illustrates that a short dipole has about the same pattern as a long-resonant-length (reactance equals zero) dipole.

We determine directivity (Section 1-7) by calculating the average radiation intensity, often normalized to the peak of the power pattern:

$$U_{\text{avg}} = \int_0^{\pi/2}\sin^2\theta\sin\theta\,d\theta = \frac{2}{3}$$

$$U_{\max} = 1$$

$$\text{directivity} = \frac{U_{\max}}{U_{\text{avg}}} = 1.5 \quad (1.76\,\text{dB})$$

The resonant-length dipole ($\approx \lambda/2$) has a directivity of 2.15 dB or only 0.39 dB more than that of the very short dipole. The total power radiated by the antenna is found by integrating the Poynting vector magnitude over a sphere:

$$P_r = \int_0^{2\pi}\int_0^{\pi} S_r r^2\sin\theta\,d\theta\,d\phi$$

$$= \int_0^{2\pi}\int_0^{\pi}\left(\frac{|I|l}{2\lambda}\right)^2\eta\sin\theta\,d\theta\,d\phi$$

$$= \frac{2\pi}{3}\left(\frac{|I|l}{2\lambda}\right)^2\eta$$

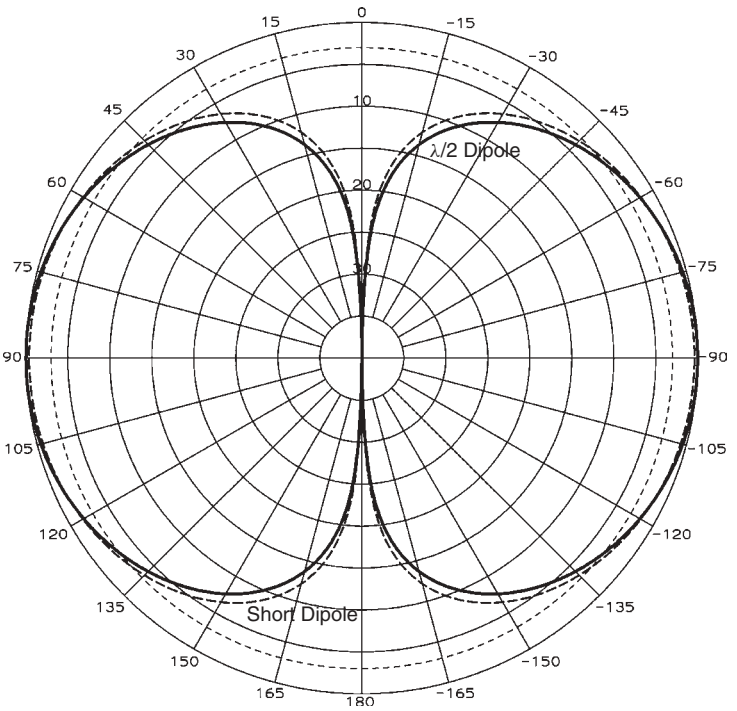


FIGURE 2-1 Pattern of a short current element and small loop (dashed curve) compared to a $\lambda/2$-long dipole (solid curve) located along the 0 to 180° axis.

We represent the radiated power as a radiation resistance at the input of the antenna:

$$R_R = \frac{P_r}{|I|^2} = \frac{2\pi}{3}\eta\left(\frac{l}{\lambda}\right)^2$$

While a short dipole with a length $\lambda/20$ has a radiation resistance of about 2 Ω, a resonant-length dipole has about a 50-Ω radiation resistance and is more efficient because the relative material resistance is low.

The input resistance of the antenna is the sum of the radiation resistance and the resistance due to material losses:

$$P_{\text{in}} = (R_R + R_L)|I|^2$$

The gain of an antenna is the ratio of the peak radiation intensity to the input power averaged over the radiation sphere:

$$\text{gain} = \frac{S_{r,\text{peak}} r^2}{\dfrac{P_{\text{in}}}{4\pi}} = \frac{U_{\max}}{\dfrac{P_{\text{in}}}{4\pi}}$$

By using the idea of radiation resistance, we rewrite this as

$$\text{gain} = \frac{4\pi U_{\max}}{(R_R + R_L)|I|^2}$$

Efficiency is the ratio of radiated power to input power:

$$\eta_e = \frac{P_r}{P_{\text{in}}} = \frac{R_r|I|^2}{(R_r + R_L)|I|^2} = \frac{R_r}{R_r + R_L}$$

Instead of integrating the pattern to calculate the total power radiated, we sometimes compute the input power of the antenna from currents induced on the antenna elements by given voltage sources on various terminals of the antenna in analysis:

$$P_{\text{in}} = \text{Re}(V_1 I_1^*) + \text{Re}(V_2 I_2^*) + \cdots + \text{Re}(V_N I_N^*)$$

The gain can be found from

$$\text{gain} = \frac{S_r(\theta, \phi)r^2}{\dfrac{P_{\text{in}}}{4\pi}} = \frac{U(\theta, \phi)}{\dfrac{P_{\text{in}}}{4\pi}}$$

This method is considerably easier than integrating the radiation intensity to compute directivity.

By integrating the pattern, we found only the input resistance of the short antenna, not the reactive component. A short antenna has a large capacitive reactance term that limits the impedance bandwidth when combined with a match network. The short antenna has a large pattern bandwidth but a narrow impedance bandwidth. Of course, an active network could be designed to impedance-match the antenna at any frequency, but the instantaneous bandwidth is narrow. The moment method of analysis gives us the currents for given input voltages and calculates the complete input impedance.

2-1.2 Radiation from Magnetic Currents

Magnetic currents are fictitious, but they enable slot radiation to be solved by the same methods as electric currents on dipoles by using duality. Slot radiation could be calculated from the surface currents around it, but it is easier to use magnetic currents to replace the electric field in the slot. Magnetic currents along the long axis of slots in ground planes replace the electric fields across the slots by application of the equivalence theorem. Similarly, current loops can be replaced by magnetic dipole elements to calculate radiation.

We use the electric vector potential $\mathbf{F}$ with magnetic currents. The far-field magnetic field is proportional to the electric vector potential:

$$\mathbf{H} = -j\omega\mathbf{F} \tag{2-9}$$

We determine the magnitude of the electric field from Eq. (2-2); it is perpendicular to $\mathbf{H}$. The electric vector potential is found from a retarded volume integral over the magnetic current density $\mathbf{M}$. Applying the radiation approximation, it is

$$\mathbf{F} = \frac{e^{-jkr}}{4\pi r}\varepsilon \iiint \mathbf{M}' e^{j\mathbf{k}\cdot\mathbf{r}'}\, dV' \tag{2-10}$$

where ε is the permittivity (8.854×10^{-12} F/m in free space). Equation (2-9) is the dual of Eq. (2-1), and Eq. (2-10) is the dual of Eq. (2-5). The dual of Eq. (2-3) is valid in both the near- and far-field regions.

The magnetic currents in a slot are perpendicular to the slot electric fields: $\mathbf{M} = \mathbf{E} \times \hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is the normal to the plane with the slot. The filamentary currents of thin slots reduce Eq. (2-10) to a line integral, and magnetic current direction limits the direction of the electric vector potential and the magnetic field. Since the electric field (far field) is orthogonal to the magnetic field, the electric field is in the same direction as the field across the slots. We use the direction of the electric field across the slots to estimate the polarization of the far field. As with filamentary electric currents, the far field is zero along the axis of the magnetic current.

The electric vector potential can also be used to derive the near field:

$$\mathbf{H} = -j\omega\mathbf{F} + \frac{\nabla(\nabla \cdot \mathbf{F})}{j\omega\mu\varepsilon}$$

$$\mathbf{E} = -\frac{1}{\varepsilon}\nabla \times \mathbf{F}$$

The magnetic field separates into near- and far-field terms in the electric vector potential; the electric field does not. We can determine the radiated fields directly in terms of the magnetic currents and avoid using the vector potential:

$$\mathbf{E}(\mathbf{r}) = -\frac{k^2}{4\pi}\iiint_{V'} \mathbf{M}(\mathbf{r}') \times \hat{\mathbf{R}}\left(\frac{j}{kR} + \frac{1}{k^2R}\right)e^{-jkR}\,dV' \tag{2-11}$$

$$\mathbf{H}(\mathbf{r}) = \frac{k^2}{4\pi\eta}\iiint_{V'}\left[\mathbf{M}(\mathbf{r}')\left(-\frac{j}{kR} - \frac{1}{k^2R^2} + \frac{j}{k^3R^3}\right)\right.$$
$$\left. + [\mathbf{M}(\mathbf{r}') \cdot \hat{\mathbf{R}}]\hat{\mathbf{R}}\left(\frac{j}{kR} + \frac{3}{k^2R^2} - \frac{3j}{k^3R^3}\right)\right]e^{-jkR}\,dV' \tag{2-12}$$

Equations (2-11) and (2-12) can be rearranged to find the dyadic Green's functions for magnetic currents. These differ from the dyadic Green's functions for electric currents by only constants.

Example Derive the fields radiated from a small constant-current loop.

We could use the magnetic vector potential and calculate over the currents in the wire but must account for changing current direction around the loop. Place the loop in the $x-y$ plane. The electric field radiated by the loop is in the ϕ direction because the currents in the loop can only be in the $\hat{\boldsymbol{\phi}}$ direction. When solving the integral for the magnetic vector potential, note that the direction of the current on the loop, $\hat{\boldsymbol{\phi}}'$ at a general point is not in the same direction as the field point, $\hat{\boldsymbol{\phi}}$, unit vector. The integral must be solved with a constant vector direction, one component at a time.

Although the magnetic vector potential can be computed, it is easier to replace the current loop with an incremental magnetic current element. The equivalent magnetic current element is

$$I_m l = j\omega\mu I A$$

where A is the area of the loop. The magnetic current density is

$$\mathbf{M} = I_m l\delta(r')\hat{\mathbf{z}} = j\omega\mu I A\delta(r')\hat{\mathbf{z}}$$

The electric vector potential is found using Eq. (2-10):

$$F_z = \frac{j\omega\mu\varepsilon I A}{4\pi r} e^{-jkr}$$

The magnetic field is found from this electric vector potential using Eq. (2-9):

$$H_z = -j\omega F_z = \frac{\omega^2\mu\varepsilon I A}{4\pi r} e^{-jkr}$$

We calculate H_θ by projection:

$$H_\theta = H_z \hat{\mathbf{z}} \cdot \hat{\boldsymbol{\theta}} = -\frac{\omega^2\mu\varepsilon I A}{4\pi r} e^{-jkr} \sin\theta$$

E_ϕ and H_θ are related in the far field because the wave propagates in the $\mathbf{r}$ direction:

$$E_\phi = -\eta H_\theta = \frac{\omega^2\mu\varepsilon I A\eta}{4\pi r} e^{-jkr} \sin\theta$$

The small current loop and small current element have the same pattern shape, $\sin\theta$, but opposite polarizations. The directivity is 1.5 (1.76 dB). Figure 2-1 uses a dashed curve to plot the response of the small loop, while the solid curve gives the pattern of a half-wavelength slot that radiates on both sides of the ground sheet.

2-2 APERTURES: HUYGENS SOURCE APPROXIMATION

Many antennas, such as horns or paraboloid reflectors, can be analyzed simply as apertures. We replace the incident fields in the aperture with a combination of equivalent electric and magnetic currents. We calculate radiation as a superposition of each source by using the vector potentials. Most of the time, we assume that the incident field is a propagating free-space wave whose electric and magnetic fields are proportional to one another. This gives us the Huygens source approximation and allows the use of integrals over the electric field in the aperture. Each point in the aperture is considered to be a source of radiation. The far field is given by a Fourier transform of the aperture field:

$$\mathbf{f}(k_x, k_y) = \iint_S \mathbf{E} e^{j\mathbf{k}\cdot\mathbf{r}'} \, ds' \tag{2-13}$$

This uses the vector propagation constant

$$\mathbf{k} = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}} + k_z\hat{\mathbf{z}}$$
$$k_x = k\sin\theta\cos\phi \qquad k_y = k\sin\theta\sin\phi \qquad k_z = k\cos\theta$$

where $\mathbf{f}(k_x, k_y)$ is the pattern in k-space. We multiply the Fourier transform far field by the pattern of the Huygens source:

$$\frac{je^{-jkr}}{2\lambda r}(1 + \cos\theta) \tag{2-14}$$

When apertures are large, we can ignore this pattern factor. In Eq. (2-13), $\mathbf{f}(k_x, k_y)$ is a vector in the same direction as the electric field in the aperture. Each component

is transformed separately. The far-field components E_θ and E_ϕ are found by projection (scalar products) from $\mathbf{f}(k_x, k_y)$ times the pattern factor of the Huygens source [Eq. (2-14)].

If we have a rectangular aperture in which the electric field is expressed as a product of functions of x and y only, the integral reduces to the product of two single integrals along each coordinate. The Fourier transform relationships provide insight into pattern shape along the two axes. Large apertures radiate patterns with small beamwidths. An antenna with long and short axes has a narrow-beamwidth pattern in the plane containing the long dimension and a wide beamwidth in the plane containing the short dimension. This is the same as the time and frequency dual normally associated with the Fourier transform.

We draw on our familiarity with signal processing to help us visualize the relationship between aperture distributions and patterns. Large apertures give small beamwidths, just as long time pulses relate to low-frequency bandwidths in normal time–frequency transforms. The sidelobes of the pattern correspond to the frequency harmonics of an equivalent time waveform under the Fourier transform and rapid transitions in the time response lead to high levels of harmonics in the frequency domain. Rapid amplitude transitions in the aperture plane produce high sidelobes (harmonics) in the far-field response (Fourier transform). Step transitions on the aperture edges produce high sidelobes, while tapering the edge reduces sidelobes and we relate the sidelobe envelope of peaks to the derivative of the distributions at the edges. To produce equal-level sidelobes, we need Dirac delta functions in the aperture that transform to a constant level in the pattern domain. Another example is periodic aperture errors that produce single high sidelobes. When we discuss aperture distribution synthesis, we see that the aperture extent in wavelengths limits our ability to control the pattern.

A uniform amplitude and phase aperture distribution produces the maximum aperture efficiency and gain that we determine from the following argument. An aperture collects power from a passing electromagnetic wave and maximum collectible power occurs at its peak amplitude response. If the amplitude response somewhere else in the aperture is reduced from the maximum, that portion will collect less power. The amplitude response can be reduced only by adding loss or reflecting power in reradiation. The antenna with the highest aperture efficiency reflects the least amount of power when illuminated by a plane wave. Similarly, if the phase shift from the collecting aperture to the antenna connector is different for different parts of the aperture, the voltages from the various parts will not add in phase. Gain is directly proportional to aperture efficiency [Eq. (1-10)]. Therefore, a uniform amplitude and phase aperture distribution has maximum gain. All this assumes that the input match on various aperture distribution antennas is the same.

For example, consider the pattern of a uniform aperture distribution in a rectangular aperture $a \times b$. We use the Fourier transform and ignore the polarization of the electric field in the aperture. (This assumes that the field has a constant polarization or direction.)

$$f(k_x, k_y) = E_0 \int_{-b/2}^{b/2} \int_{-a/2}^{a/2} e^{j\mathbf{k}\cdot\mathbf{r}'}\, dx\, dy$$

$$= E_0 \int_{-b/2}^{b/2} \int_{-a/2}^{a/2} e^{jk_x x} e^{jk_y y} dx\, dy$$

We separate the integral into a product of two integrals each with the form

$$\int_{-a/2}^{a/2} e^{jk_x x}\,dx = \frac{e^{jk_x a/2} - e^{-jk_x a/2}}{jk_x} = \frac{a\sin(k_x a/2)}{k_x a/2}$$

On combining the two similar integrals, we have

$$f(k_x, k_y) = \frac{ab\sin(k_x a/2)}{k_x a/2}\frac{\sin(k_y b/2)}{k_y b/2}$$

where $k_x = k\sin\theta\cos\phi$, $k_y = k\sin\theta\sin\phi$, and $k_z = k\cos\theta$ and $k = 2\pi/\lambda$. The pattern in both planes is given by a k-space function, $\sin u/u$. Figure 2-2 plots this pattern function as a solid curve using k_x-space $[(ka/2)\sin\theta]$ as the abscissa to produce a universal curve independent of aperture size a. The half-power points occur when

$$\frac{\sin u}{u} = \frac{1}{\sqrt{2}} \quad \text{or} \quad u = 1.39156$$

When we substitute for u, we have in the principal planes

$$\frac{\pi a}{\lambda}\sin\theta = 1.39156$$

By solving for θ, we compute the half-power beamwidth (HPBW):

$$\text{HPBW} = 2\sin^{-1}\frac{0.4429\lambda}{a}$$

By using the approximation $u = \sin u$ for small angles, the half-power beamwidth can be estimated as

$$\text{HPBW} = 50.76^\circ\frac{\lambda}{a}$$

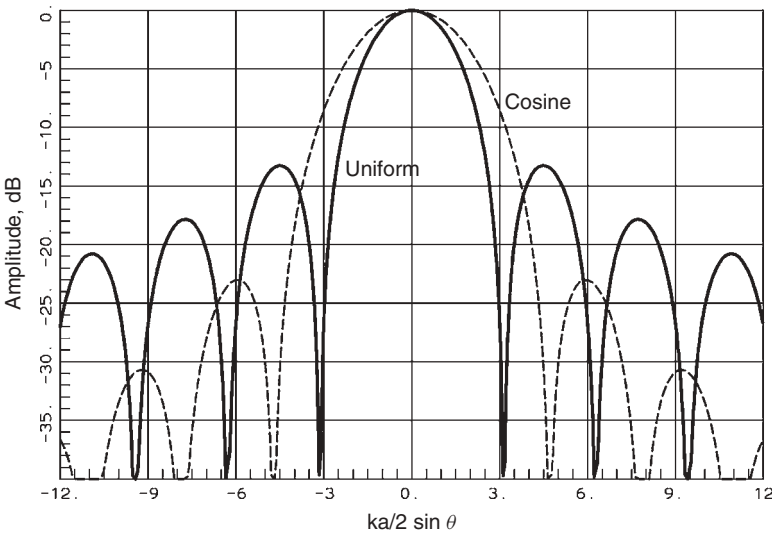


FIGURE 2-2 Universal k-space pattern for the radiation from uniform (solid curve) and cosine (dashed curve) aperture distributions.

Note that we have ignored the $(1 + \cos\theta)/2$ pattern of the Huygens source, which reduces the beamwidth for radiation from small apertures. We discover on Figure 2-2 the pattern nulls occur at integer multiples of π and the first sidelobe amplitude is 13.2 dB below the peak.

The gain of a uniform amplitude and phase aperture distribution is given by Eq. (1-7), where A is the area of the aperture. Chapter 4 develops amplitude taper efficiency for nonuniform aperture amplitude distributions to calculate the gain reduction. Phase error efficiency gives the gain reduction due to phase anomalies. Each of these is found from the distribution of aperture fields. Figure 2-2 plots the pattern of a half cosine aperture distribution as a dashed curve. The distribution peaks in the center and tapers linearly to zero at the edges. Tapering the aperture distribution widens the beamwidth and reduces both gain and sidelobe levels. The pattern beamwidth is 1.342 times wider than the uniform distribution beamwidth. A cosine distribution produces a -0.91-dB amplitude taper loss, and the distribution edge taper causes the sidelobes to fall off at a faster rate.

Example Compute the length of the aperture with a uniform distribution that will give a 10° beamwidth.

$$\frac{a}{\lambda} = \frac{50.76^\circ}{10^\circ} \simeq 5 \text{ wavelengths}$$

We can calculate radiated power by integrating the Poynting vector magnitude over the radiation sphere, but there is an easier way. We assumed that the aperture fields are free-space waves. The total power radiated is in the aperture,

$$P_r = \iint_{\text{aperture}} \frac{|E|^2}{\eta}\, ds \qquad P_{\text{avg}} = U_{\text{avg}} = \frac{P_r}{4\pi}$$

where η is the impedance of free space. The radiated electric field is

$$\mathbf{E} = j\frac{e^{-jkr}(1 + \cos\theta)}{2\lambda r}\mathbf{f}(k_x, k_y)$$

The Poynting vector magnitude is

$$S_r = \frac{|\mathbf{E}|^2}{\eta} = \frac{(1 + \cos\theta)^2}{4\lambda^2 r^2}|\mathbf{f}(k_x, k_y)|^2 \tag{2-15}$$

By combining Eqs. (2-14) and (2-15), we determine directivity:

$$\begin{aligned}\text{directivity}(\theta, \phi) &= \frac{U(\theta, \phi)}{U_{\text{avg}}} = \frac{S_r r^2}{P_r/4\pi} \\ &= \frac{\pi(1 + \cos\theta)}{\lambda^2}\frac{\left|\iint E e^{j\mathbf{k}\cdot\mathbf{r}'}\, ds'\right|^2}{\iint |E|^2 ds'}\end{aligned} \tag{2-16}$$

By considering electric and magnetic fields separately in the aperture, we eliminate the requirement that the ratio electric and magnetic fields are the same as free space used in the Huygens source approximation. Given the fields in an aperture, we can equate

them to magnetic and electric currents:

$$\mathbf{M}_s = \mathbf{E}_a \times \mathbf{n} \qquad \mathbf{J}_s = \mathbf{n} \times \mathbf{H}_a \tag{2-17}$$

where $\mathbf{E}_a$ and $\mathbf{H}_a$ are the aperture fields and $\mathbf{n}$ is the outward normal. The equivalence theorem [2, p. 113] results in exact solutions by using the total aperture field, incident and reflected. When using the equivalence theorem, we replace the total fields present with equivalent currents. The induction theorem equates currents only to the incident fields on the aperture, which ignores wave reflection and results in approximate solutions:

$$\mathbf{F} = \varepsilon \iint_s \frac{\mathbf{M}_s e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}\, ds' \qquad \mathbf{A} = \mu \iint_s \frac{\mathbf{J}_s e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}\, ds' \tag{2-18}$$

We derive the radiated fields from each distribution of currents by using vector potentials where $\mathbf{r}$ is the field point and $\mathbf{r}'$ is the source point in the aperture. These expressions are valid in the near and far fields. By integrating over only a finite aperture, we assume zero fields outside the aperture, while rigorous expressions require integrals over closed boundaries. A planar aperture must extend to infinity, but the fields outside the aperture are nearly zero and contribute little.

2-2.1 Near- and Far-Field Regions

The radiative near- and far-field regions are characterized by the approximations made to the integrals [Eq. (2-18)]. The radiative near-field region lies between the near field, with no approximations, and the far-field region. In both approximations the field (observation) distance r is substituted for $|\mathbf{r}-\mathbf{r}'|$ in the amplitude term. The vector potentials reduce to

$$\mathbf{F} = \frac{\varepsilon}{4\pi r} \iint_s \mathbf{M}_s e^{-jk|\mathbf{r}-\mathbf{r}'|}\, ds' \qquad \mathbf{A} = \frac{\mu}{4\pi r} \iint_s \mathbf{J}_s e^{-jk|\mathbf{r}-\mathbf{r}'|}\, ds' \tag{2-19}$$

We handle the phase term differently in the two regions. First, we expand the phase term in a Taylor series,

$$|\mathbf{r}-\mathbf{r}'| = \sqrt{r^2 + r'^2 - 2\mathbf{r}\cdot\mathbf{r}'} = r - \hat{\mathbf{r}}\cdot\mathbf{r}' + \frac{1}{2r}[r'^2 - (\hat{\mathbf{r}}\cdot\mathbf{r}')^2]\cdots$$

where $\hat{\mathbf{r}}$ is the unit vector in the field point direction. We retain the first two terms for the far-field approximation and the vector potentials become

$$\mathbf{F} = \frac{e^{-jkr}\varepsilon}{4\pi r} \iint_s \mathbf{M}_s e^{j\mathbf{k}\cdot\mathbf{r}'}\, ds', \qquad \text{etc.} \tag{2-20}$$

where we have combined k, the propagation constant, with the unit vector $\hat{\mathbf{r}}$:

$$\mathbf{k} = k\hat{\mathbf{r}} = k(\sin\theta\cos\phi\hat{\mathbf{x}} + \sin\theta\sin\phi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}})$$

The magnetic vector potential integral parallels Eq. (2-20) as in Eq. (2-19). In the radiative near-field zone approximation the terms in r'^2 are retained and we obtain the following integral for the electric vector potential:

$$\mathbf{F} = \frac{e^{-jkr}\varepsilon}{4\pi r} \iint \mathbf{M}_s \exp\left[j(\mathbf{k}\cdot\mathbf{r}') + \frac{(\mathbf{k}\cdot\mathbf{r}')^2}{2rk} - \frac{kr'^2}{2r} \right] ds' \tag{2-21}$$

No clear boundary between the three regions exists because the fields are continuous. Common boundaries are

$$\frac{r}{L} < 1 \qquad \text{near field}$$

$$1 < \frac{r}{L} < \frac{L}{\lambda} \qquad \text{radiative near field}$$

$$\frac{r}{L} > \frac{L}{\lambda} \qquad \text{far field}$$

where L is the maximum dimension on the aperture.

Example Determine the maximum difference between the radiative near- and far-field approximations at a point normal to the maximum aperture dimension when $r = L^2/\lambda$ and $r = 2L^2/\lambda$.

Normal to the maximum dimension, $\hat{\mathbf{r}} \cdot \mathbf{r}' = 0$. The phase difference is

$$\frac{kr'^2_{\max}}{2r} \qquad \text{where} \quad r'_{\max} = \frac{L}{2}$$

$$\text{phase difference } \phi = \frac{2\pi L^2}{8\lambda r}$$

$$\phi = \pi/4 \quad \text{at} \quad r = L^2/\lambda \quad \text{and} \quad \phi = \pi/8 \quad \text{at} \quad r = 2L^2/\lambda$$

The usual minimum distance used for antenna patterns is $2L^2/\lambda$, where L is the maximum dimension of the antenna. At that distance, the phase error across the aperture from a point source antenna is $\pi/8$. The distance is not sufficient for low-sidelobe antennas [3] because quadratic phase error raises the measured sidelobes.

We can use vector potentials in the aperture after determining equivalent currents, but we will find it more convenient to use the fields directly. Define the following integrals:

$$\mathbf{f} = \iint_s \mathbf{E}_a e^{j\mathbf{k}\cdot\mathbf{r}'}\, ds \qquad \mathbf{g} = \iint_s \mathbf{M}_a e^{j\mathbf{k}\cdot\mathbf{r}'}\, ds \tag{2-22}$$

using the far-field approximation. Near-field integrals require additional phase terms. Given an aperture, we calculate the vector potentials in terms of $\mathbf{E}_a$ and $\mathbf{H}_a$ through the currents by using either the equivalence or inductance theorems, and we use the integrals of Eq. (2-22) in the vector potentials. We combine the fields in the far field due to each partial source:

$$\mathbf{E} = -j\omega\mathbf{A} - j\eta\omega\mathbf{F} \times \hat{\mathbf{r}}$$

For an aperture in the x–y plane, we carry out these steps by using the inductance theorem and obtain the following far fields from the incident aperture fields

$$\begin{aligned} E_\theta &= \frac{jke^{-jkr}}{4\pi r}[f_x \cos\phi + f_y \sin\phi + \eta\cos\theta(-g_x \sin\phi + g_y \cos\phi)] \\ E_\phi &= \frac{-jke^{-jkr}}{4\pi r}[(f_x \sin\phi - f_y \cos\phi)\cos\theta + \eta(g_x \cos\phi + g_y \sin\phi)] \end{aligned} \tag{2-23}$$

where **f** and **g** have been expanded in terms of their x- and y-components and η is the impedance of free space.

2-2.2 Huygens Source

The Huygens source approximation is based on the assumption that the magnetic and electric fields are related as in a plane wave in the aperture:

$$\eta g_y = f_x \quad \text{and} \quad -\eta g_x = f_y$$

since

$$\eta H_y = E_x \quad \text{and} \quad -\eta H_x = E_y$$

With this approximation, the far field [Eq. (2-23)] becomes

$$\begin{aligned} E_\theta &= \frac{jke^{-jkr}}{4\pi r}(1+\cos\theta)(f_x\cos\phi + f_y\sin\phi) \\ E_\phi &= \frac{-jke^{-jkr}}{4\pi r}(1+\cos\theta)(f_x\sin\phi - f_y\cos\phi) \end{aligned} \tag{2-24}$$

The two-dimensional vector Fourier transform $\mathbf{f} = (f_x, f_y)$ of the aperture electric field in the $x-y$ plane determines the far-field components. We derive the radiated components by projecting (vector scalar product) this field onto the vectors $\hat{\theta}/\cos\theta$ and $\hat{\phi}$. The transform **f** expands the field in k-space [usually, (k_x, k_y)]. This normalizes the pattern and removes the direct dependence on aperture length.

We separate out all but **f** when we consider aperture distributions. We drop the terms for the radiation from a point source and the pattern of a Huygens point source [Eq. (2-14)] and limit our discussions to Huygens sources and far fields. General aperture fields require Eq. (2-23), and for any region other than the far field, additional phase terms are needed in the transforms [Eq. 2-21)].

2-3 BOUNDARY CONDITIONS

Material boundaries cause discontinuities in the electric and magnetic fields. The effects can be found by considering either vanishing small pillboxes or loops that span the boundary between the two regions. By using the integral form of Maxwell's equations on these differential structures, the integrals reduce to simple algebraic expressions. These arguments can be found in most electromagnetic texts and we give only the results. Conversely, we will discover that artificial boundaries such as shadow and reflection boundaries used in geometric optics (ray optics) cannot cause a discontinuity in the fields because they are not material boundaries. The idea that the fields remain continuous across the boundary leads to the necessity of adding terms to extend ray optics methods. We discuss these ideas when considering the uniform theory of diffraction (UTD) method used with ray optics.

Suppose that we have a locally plane boundary in space described by a point and a unit normal vector $\hat{\mathbf{n}}$ that points from region 1 to region 2. We compute the tangential fields from the vector (cross) product of the fields and the normal vector. The fields

can be discontinuous at the interface between the two regions if surface magnetic $\mathbf{M}_S$ or electric current $\mathbf{J}_S$ densities exist on the surface.

$$\hat{\mathbf{n}} \times (\mathbf{E}_2 - \mathbf{E}_1) = -\mathbf{M}_S \qquad \hat{\mathbf{n}} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{J}_S \tag{2-25a,b}$$

The normal components of the fields change due to the differing dielectric and magnetic properties of the materials and the charges induced on the surface:

$$\hat{\mathbf{n}} \cdot (\varepsilon_2 \mathbf{E}_2 - \varepsilon_1 \mathbf{E}_1) = \rho_S \qquad \hat{\mathbf{n}} \cdot (\mu_2 \mathbf{H}_2 - \mu_1 \mathbf{H}_1) = \tau_S \tag{2-26}$$

with ρ_S and τ_S given as electric and magnetic surface charge densities, respectively. Perfect dielectric and magnetic materials can have no currents, which reduces Eq. (2-25) to

$$\hat{\mathbf{n}} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0 \qquad \hat{\mathbf{n}} \times (\mathbf{H}_2 - \mathbf{H}_1) = 0 \tag{2-27}$$

Equation (2-27) means that the tangential fields are continuous across the boundary.

These boundary conditions are used in the method of moment analyses to determine currents. The method applies the boundary condition in integral equations to determine the coefficients of the expansion of currents in the sum of basis functions. The currents described as these sums do not satisfy the boundary conditions at all points but do when integrated over a region. This method leads to approximations that will converge as more terms are included in the expansions.

When doing analysis we find two types of surfaces convenient. We use these surfaces to reduce analysis effort by using planes of symmetry. The first one is the *perfect electric conductor* (PEC). A PEC surface causes the fields to vanish inside and to have electric currents induced on it:

$$\hat{\mathbf{n}} \times \mathbf{E}_2 = 0 \qquad \hat{\mathbf{n}} \times \mathbf{H}_2 = \mathbf{J}_S \qquad \text{(PEC)} \tag{2-28a,b}$$

A PEC surface is also called an *electric wall*. The second surface is the *perfect magnetic conductor* (PMC) and is a hypothetical surface. Whereas good conductors approximate PEC, there are no PMC materials. The PMC has no internal fields like the PEC and forces the tangential magnetic field to be zero:

$$\hat{\mathbf{n}} \times \mathbf{E}_2 = -\mathbf{M}_S \qquad \hat{\mathbf{n}} \times \mathbf{H}_2 = 0 \qquad \text{(PMC)} \tag{2-29}$$

A PMC surface supports the hypothetical magnetic current density $\mathbf{M}_S$. We find that the magnetic wall (PMC) concept simplifies analysis.

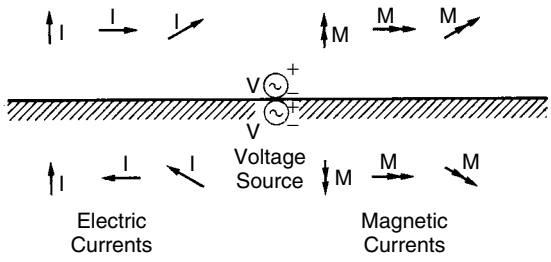


FIGURE 2-3 Ground-plane images.

We use images of currents to include material boundaries in analysis. Figure 2-3 illustrates ground-plane images. When we analyze radiation from currents in the presence of a boundary, we include the actual antenna and its image to compute the fields. The figure shows an infinite ground plane, but a finite ground-plane image can be used in the angular region where a reflected wave occurs in the finite plane. We consider this idea further when discussing geometric optics. We can use images in dielectric boundaries provided that we calculate the polarization sensitive reflection coefficients to adjust the magnitude and phase of the image.

2-4 PHYSICAL OPTICS

Physical optics uses things that can be measured. We can measure both currents and fields, but auxiliary vector potentials have no physical reality, only mathematical artifacts that simplify Maxwell's equations. Nevertheless, the auxiliary vector potentials provide simple models for problems that enable simple mental pictures, as shown earlier, but we cannot easily formulate them into a systematic analysis tool for antenna problems.

The physical optics analysis method combines the use of Green's functions to calculate fields radiated by a given distribution of currents and then uses boundary conditions to determine the currents induced on objects due to incident fields. We compute the effects of a mounting structure by inducing currents on it and adding their radiation to the antenna pattern. The method assumes that radiation from the induced currents on the structure does not change the initial currents.

We start analyses from either currents or incident fields and work from those. The resonant structure of many antennas determines the approximate current distribution that we normalize to the radiated power. We calculate the fields from these currents. Physical optics can use an iterative technique to calculate incremental currents induced on the original radiators and improve the solution, but we usually just sum the radiation from the original currents to the radiation from the induced currents. The second starting point for physical optics can be incident fields. These could be plane waves or could be fields found from the measured radiation patterns of antennas: for example, the pattern of a reflector feed. We add the radiation from the induced currents to the incident waves.

2-4.1 Radiated Fields Given Currents

The radiated fields can be found from distribution of the electric and magnetic currents by the use of dyadic Green's functions that contain source and field coordinates. We sometimes refer to the Green's functions as *vector propagators* or *transfer functions* between currents and fields. We calculate the fields from integrals over the source points of the dot (scalar) product between the dyadic and current densities. The dyadic Green's function contains both near- and far-field terms and requires slightly different expressions for the electric and magnetic fields. The general propagator from electric and magnetic currents has separate terms for electric and magnetic currents, which when used with surface patch currents can be reduced to short subroutines or procedures easily programmed [1]:

$$\mathbf{E}(\mathbf{r}) = \int \mathbf{G}_{\mathrm{EJ}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}(\mathbf{r}')\, dV' + \int \mathbf{G}_{\mathrm{EM}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{M}(\mathbf{r}')\, dV' \tag{2-30}$$

$$\mathbf{H}(\mathbf{r}) = \int \mathbf{G}_{\mathrm{HJ}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}(\mathbf{r}')\, dV' + \int \mathbf{G}_{\mathrm{HM}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{M}(\mathbf{r}')\, dV' \tag{2-31}$$

These expressions integrate over the currents located at source points $\mathbf{r}'$ for a dyadic Green's function that changes at each field point $\mathbf{r}$ and source point $\mathbf{r}'$. Although these Green's functions are valid at all field points in space both near and far field, they are singular at a source point. Only retaining terms with $1/R$ dependence for the far field greatly simplifies the expressions.

When fields are incident on a perfect electric conductor (PEC), the combination of incident and reflected tangential magnetic fields induces an electric current density on the surface. The fields inside the conductor are zero. We assume locally plane surfaces on patches and compute currents that satisfy the boundary condition. Given the local unit normal $\hat{\mathbf{n}}$ to the surface, the induced current density is given by

$$\begin{aligned} \mathbf{J}_S &= \hat{\mathbf{n}} \times (\mathbf{H}_{\text{incident}} + \mathbf{H}_{\text{reflected}}) \\ \mathbf{H}_{\text{incident}} &= \mathbf{H}_{\text{reflected}} \\ \mathbf{J}_S &= 2\hat{\mathbf{n}} \times \mathbf{H}_{\text{incident}} \end{aligned} \tag{2-32}$$

The reflected magnetic field equals the incident magnetic field because the field reflects from the conductive surface. The sum of the tangential electric fields must be zero. Because the reflected wave changes direction, the vector (cross) product of the electric and magnetic fields must change direction. The reflected tangential electric field changes direction by 180°, so the tangential magnetic field must not change direction because the Poynting vector changed its direction. Equation (2-32) is the magnetic field equation applied on a PEC. Equation (2-25*b*) is the general magnetic field equation at a boundary.

Physical optics starts with a given current distribution that radiates, or the measured pattern of an antenna. When an object is placed in the radiated field, the method calculates induced current on the object to satisfy the internal field condition. For example, PEC or PMC have zero fields inside. When we use simple functions such as constant-current surface patches, the sum of the radiation from the incident wave and the scattered fields from induced surface currents produces only approximately zero fields inside. As the patch size decreases, the method converges to the correct solution. To obtain the radiated field everywhere, we sum the incident wave and scattered waves. The fields radiated by the induced currents produce the shadow caused by the object. With geometric optics techniques such as UTD, the object blocks the incident wave and we determine the fields in the shadow regions from separate diffraction waves. In physical optics the incident wave continues as though the object were not present. Only geometric optics techniques use blockage.

We can calculate the fields radiated from antennas in free space or measure them in an anechoic chamber that simulates free space, but we mount the antenna on finite ground planes, handsets, vehicles, over soil, and so on, when we use them. Physical optics is one method of accounting for the scattering. We show in later chapters that the mounting configuration can enhance the patterns.

2-4.2 Applying Physical Optics

In this book we do not discuss how to develop numerical techniques, but it is important to understand how to apply methods. Whether you develop your own codes or use commercial codes, certain rules should be applied. Consider Eq. (2-32). The normal to

the surface points in the direction of the incident wave: outward. If the normal pointed inward, the sign of induced electric current density would change. Most codes have made the assumption that the normal points outward, but some codes may check on the direction of the normal relative to the incident wave and make the necessary sign change. We must keep track of the direction of the normal, and it may be necessary to rotate the normal depending on the expected direction of the incident wave. If an object can have radiation from both sides, it may be necessary to use two objects in the analysis.

Many codes store each object as a separate entity in a disk file. In some cases we need to store an object multiple times. Take, for example, a Cassegrain dual reflector. The feed antenna illuminates the subreflector and induces currents on it. These currents radiate and excite currents on the main reflector. When the main reflector–induced currents radiate, the subreflector intercepts or blocks part of the fields. We account for this blockage by using a second subreflector object on which the code calculates a new set of induced currents by using the main reflector currents as the source. We could add these currents to the existing disk file object or merely keep the second object. We want to keep the second object separate so that we can calculate additional currents induced on the main reflector using these currents as sources. These currents will be reduced from the initial set, but they are an important contribution to the fields radiated behind the reflector. This example illustrates iterative PO. When objects face each other significantly, iterative PO is necessary to calculate correct patterns. The method converges rapidly in most cases.

Figure 2-4 illustrates the geometry of a corner reflector. A half-wavelength-long dipole is placed between two metal plates usually bent to form a 90° angle. We can use other angular orientations between the plates, but this is the usual design. The figure does not show the feed line to the dipole, which usually starts at the juncture of the two plates and runs up to the dipole. This feed line contains the balun discussed in Section 5-15. Although the figure shows the plates as solid, many implementations use metal rods to reduce weight and wind loading.

The analysis starts with assumed currents on the dipole. We divide the plates analytically into small rectangular patches, which can be small ($\approx \lambda/8$ to $\lambda/4$) on a side since it takes only a few to cover the plates. You should repeat the analysis with

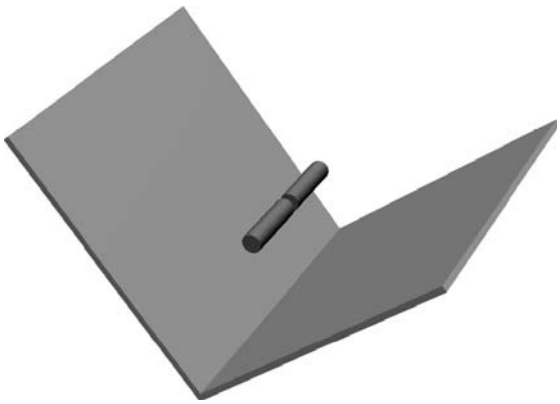


FIGURE 2-4 Corner reflector with a dipole located between two flat plates.

different-sized current patches to determine if the analysis has converged. In a similar manner, we break down the current on the dipole into short linear segments, each with constant amplitude. By using a near-field version of Eq. (2-31), we calculate the magnetic field incident on each patch on the plates. This field induces electric currents on the plates calculated from Eq. (2-32). Remember that we combine radiation from the source dipole with that radiated from the induced currents to reduce the radiation behind the antenna. The currents were induced to satisfy the boundary condition of the plate, but only with both radiations present. Figure 2-5*a* illustrates this process of inducing currents. Figure 2-6 shows the antenna pattern calculated using these currents. The E-plane pattern drawn as a solid line produces a null at 90° because the dipole pattern has this null. The plates cause the narrowing of the beam in the H-plane. The plates reduced the back radiation to -22 dB relative to the forward radiation, called the *front-to-back ratio* (F/B). The gain has increased from the 2.1 dB expected from a dipole to 9.3 dB. An equivalent geometric optics analysis uses two images in the plates, as shown in Figure 2-5*b*, for the analysis.

If you look at Figure 2-4 or 2-5, you should notice that the two plates face each other. Currents on one plate will radiate toward the other plate and induce another set of currents on it. We could ignore these induced currents if the radiation was insignificant, but to produce correct patterns we must include them. The solution to this problem calls for an iterative technique where we calculate the radiation from the currents on the first plate and induce incremental currents on the second plate. These incremental currents produce further radiation that induces additional currents on the other plate. The method converges rapidly. Figure 2-7 gives the antenna pattern after the iterations have been completed and we include radiation from all currents. The actual F/B ratio of the antenna is 29 dB, and the additional currents increased the gain by 0.7 dB to 10 dB. Adding the two plates in the original analysis increased the gain by 7.2 dB, whereas the iterative technique had a much smaller effect. Figure 2-8 illustrates the iterative technique and shows that the equivalent geometric optics analysis adds a third image to represent the reflection between the plates. Remember when you mount the antenna in an application, the structure will change the realized pattern, but the high F/B ratio reduces this effect. The mounting structure used when measuring the antenna changes the pattern as well, which limits our knowledge of the real pattern.

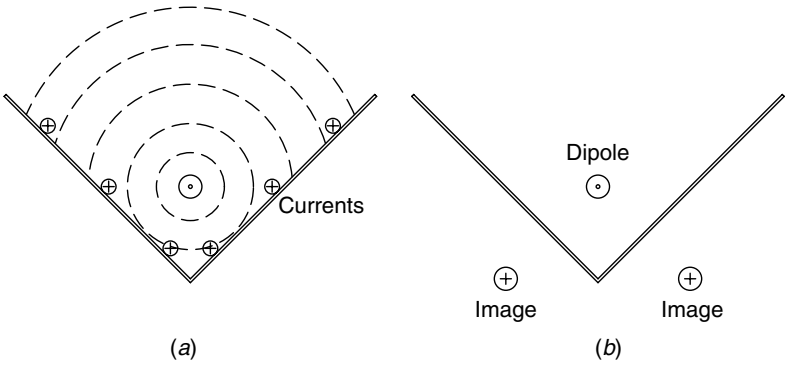


FIGURE 2-5 Cross-sectional view of a corner reflector: (*a*) magnetic field radiated from a dipole induces currents on plates; (*b*) plate currents replaced with image dipoles.

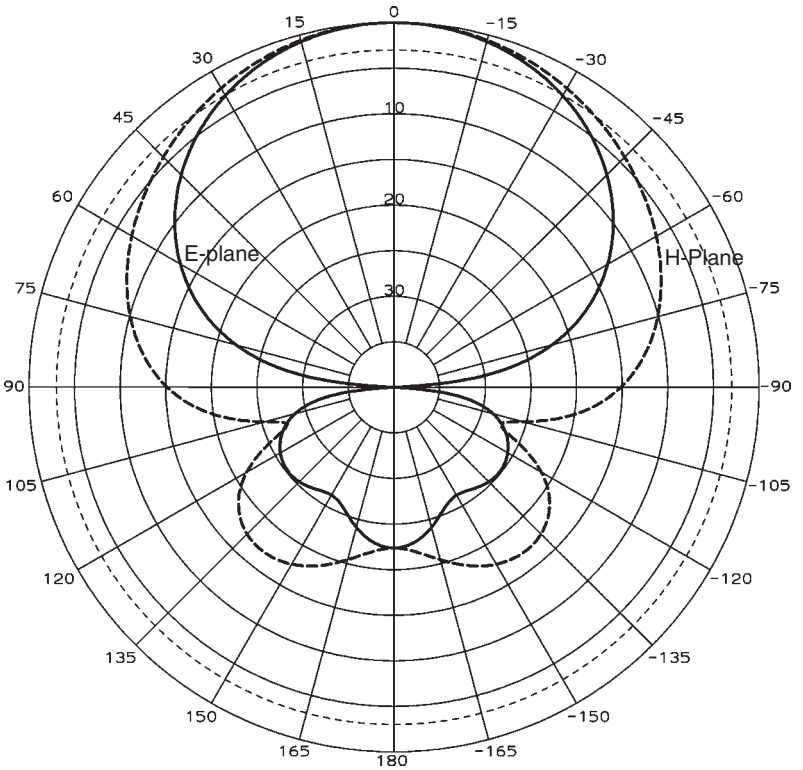


FIGURE 2-6 Pattern calculated from a combination of dipole and plate currents in a corner reflector with $1 \times 0.9\lambda$ plates without induced current iteration.

Physical optics can determine the impedance effects of the limited images in the ground planes, such as the corner reflector. The local nature of impedance effects allows the use of images to calculate the mutual impedance effects of ground planes. We use impedance calculations not only to determine the bounds of ground-plane effects on input impedance, but to calculate the total power radiated by the antenna. The images (excited currents on ground planes) radiate but do not receive input power. A ground plane at least $\lambda/2$ on a side located about $\lambda/4$ away from the antenna produces nearly the same impedance effects as an infinite ground plane, but the ground plane alters the radiation pattern greatly because it restricts possible radiation directions.

It has commonly been thought that physical optics could compute the field only in the main beam pattern direction of a paraboloidal reflector. The method can determine this pattern region accurately by using only a few patches, each one being many wavelengths on a side. As the processing power of computers increases, the patch size can be shrunk until PO can calculate the pattern in every direction, including behind the reflector. It is important to remember to include the feed pattern behind the reflector even though its radiation is obviously blocked by the main reflector. Physical optics uses induced currents to cancel the fields inside objects when the incident fields and the radiation from the induced currents are added. We can calculate the pattern behind a reflector using UTD (GTD), the uniform (geometric) theory of diffraction. This

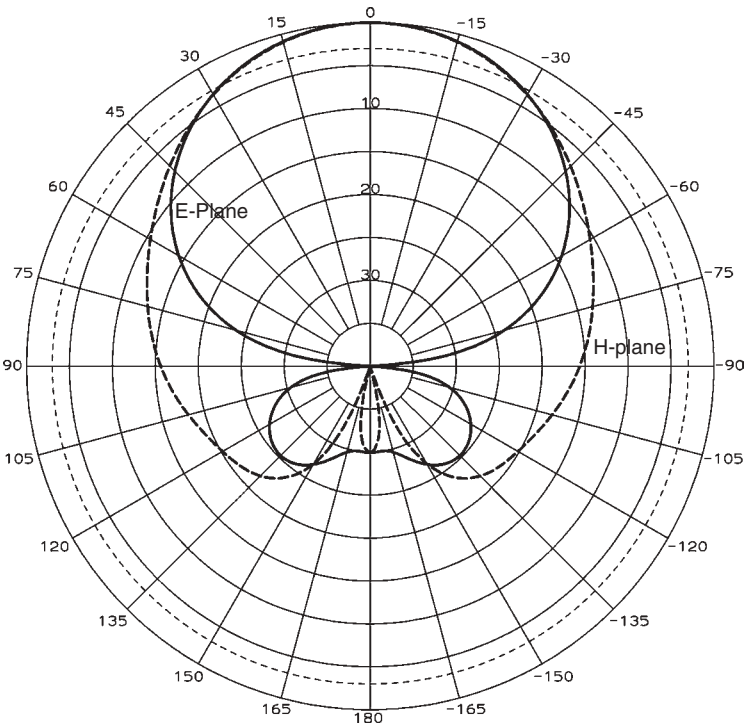


FIGURE 2-7 Pattern of corner reflector with $1 \times 0.9\lambda$ plates with induced current iteration equivalent to multiple reflectors between the plates.

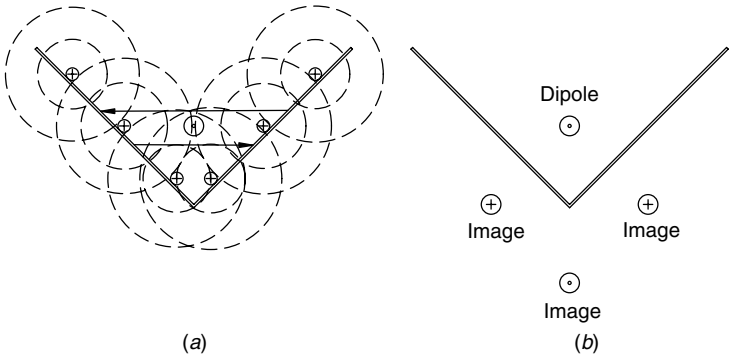


FIGURE 2-8 (*a*) Wall currents on plates radiate magnetic fields that induce additional currents on facing plates; (*b*) added induced currents equivalent to additional image dipole.

geometric optics-based method blocks the radiation from the feed and uses diffractions from the rim edge to calculate the pattern behind the reflector. We discuss UTD in Section 2-7. A comparison of UTD and physical optics calculations [1,4] of the pattern behind shows that the two methods match.

The dashed curve of Figure 2-9 plots the results of the PO analysis of a 20λ-diameter centrally fed paraboloidal reflector. The feed antenna radiation tapers to -12 dB at the

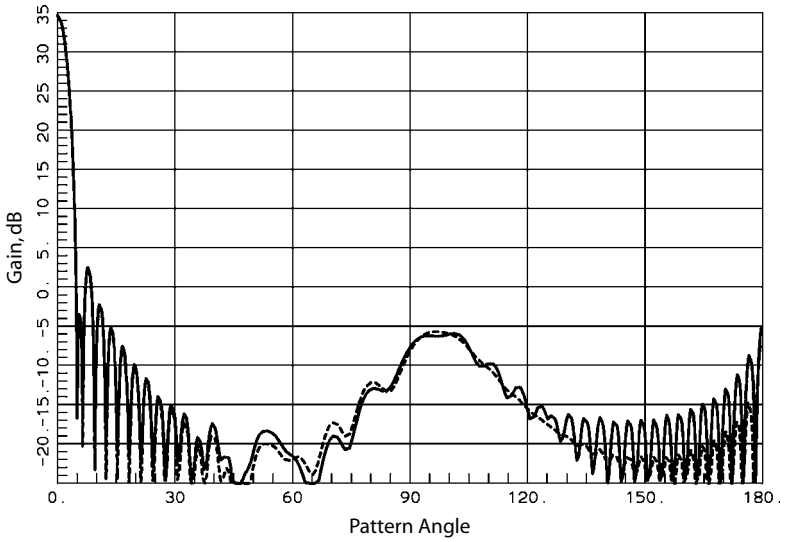


FIGURE 2-9 Physical optics analysis of a 20λ-aperture-diameter paraboloidal reflector (dashed curve) compared to analysis that includes PTD (solid curve).

reflector rim. Figure 2-9 shows the feed power spillover peaking at angles off the boresight near 100°. PO analysis computes the currents on a patch by assuming that it is embedded in an infinite plate. The reflector rim violates this assumption and we need extra terms to calculate the pattern behind the reflector accurately. Adding PTD (the physical theory of diffraction) to PO improves the match between the two methods behind the reflector as shown by the solid curve on Figure 2-9. PTD handles caustic regions of PO in a manner similar to the equivalent current method based on diffraction coefficients of UTD with geometric optics for shadow and reflection boundaries. For this example, the additional PTD currents add with the same phase because of the symmetry of the reflector geometry and produce the maximum effect. The PTD currents on the rim of an offset reflector will not add and produce a peak effect behind the reflector but will produce a more defuse effect. We only need PTD over a limited pattern angular range to reduce error, and the cost of implementing the fix may exceed the necessity of knowing the pattern in these regions. Similarly, UTD needs the addition of edge currents for accurate calculation of the radiation near 180°, behind the reflector. Although any model for the feed pattern can be used with PO, results matching UTD exactly occur only over all regions of the back radiation when the feed satisfies Maxwell's equations in the near and far fields [4, p. 212]. One such feed is the Gaussian beam approximation. Again, like PTD fixes, the small errors when using other feed antenna approximations occur only at limited pattern regions that may be unimportant.

2-4.3 Equivalent Currents

We can relate the concept of equivalent currents to physical optics. In this case we generate an artificial surface that covers a source of radiation. The incident fields generate surface electric and magnetic current whose radiation cancels the internal

fields and generates the external pattern. We use these at the apertures of antennas such as horns. By using the dyadic Green's functions we can calculate the near-field patterns and the coupling between antennas when the assumption is made that the presence of a second antenna does not alter the aperture fields. Given the outward normal $\hat{\mathbf{n}}$, we calculate the equivalent currents by

$$\hat{\mathbf{n}} \times \mathbf{E}_{\text{incident}} = -\mathbf{M}_S \qquad \hat{\mathbf{n}} \times \mathbf{H}_{\text{incident}} = \mathbf{J}_S \tag{2-33}$$

We must use both electric and magnetic current densities on the surface to replace the internal fields. If the ratio of the electric field to the magnetic field equals the impedance of free space (376.7 Ω), the combination of the two currents produces the radiation of the Huygens aperture source when used with the dyadic Green's function. We use equivalent currents for a variety of analyses over flat apertures such as horns and paraboloidal reflectors, but they can also be used with curved structures or apertures.

We can, for example, use equivalent currents for calculation of the effects of radomes. Locally, we assume that the incident waves are plane waves and use boundary conditions to calculate reflected and transmitted waves. It is necessary to separate the incident wave into parallel and perpendicular polarizations, the ray-fix representation discussed in Section 2-7.8. These polarizations have differing reflection and transmission coefficients. We generate one surface on the inside of the radome and another on the outside. We use locally free-space waves for the reflected and transmitted waves lying outside the radome. Both these waves can be replaced with equivalent currents. The equivalent currents produce null fields inside the radome when combined with the incident wave radiation [4, p. 155]. Including these equivalent currents in a PO analysis, we add the effect of the radome.

Equivalent currents can also be used with lenses. We use the incident waves combined with the idea of locally plane waves to calculate reflected and transmitted waves at each surface and replace them with equivalent currents. We include the dielectric constant of the lens in the dyadic Green's functions for the internal radiation of the lens to calculate the fields at the second surface. We apply locally plane waves at the second surface to determine the transmitted and reflected rays and then replace them with equivalent currents. Because the lens has internal reflections, we need to apply an iterative PO analysis to calculate the multiple reflections between the two surfaces. The method converges rapidly because the internal reflections are small.

2-4.4 Reactance Theorem and Mutual Coupling

In Section 1-14 we discussed how the coupling between two antennas can be found from reactance. Given a transmitting antenna that generates a field at the receiving antenna, the reactance is described by an integral equation [5]:

$$\text{reactance} = \iiint (\mathbf{E}_t \cdot \mathbf{J}_r - \mathbf{H}_t \cdot \mathbf{M}_r)\, dV = \langle t, r \rangle \tag{2-34}$$

The volume integral is over the receiving antenna currents, but it is often reduced to a surface or line integral. A second form of Eq. (2-34) uses the fields radiated by

both antennas. Given a surface that surrounds the receiving antenna, the integral for reactance is taken over this surface:

$$\text{reactance} = \iint_{S_r} (\mathbf{E}_r \times \mathbf{H}_t - \mathbf{E}_t \times \mathbf{H}_r) \cdot d\mathbf{s} = \langle t, r \rangle \tag{2-35}$$

The differential normal $d\mathbf{s}$ is pointed away from the receiving antenna.

When we represent the two antennas and the transmission between them as an impedance matrix, it implies that we know the input currents to both antennas. By expressing the coupling as an impedance matrix, we compute mutual impedance from the reactance integral:

$$Z_{12} = \frac{-1}{I_1 I_2} \langle t, r \rangle \tag{2-36}$$

Antennas that we describe by input currents only have electric current densities excited on their surfaces. The mutual impedance formula using reactance reduces to

$$Z_{12} = \frac{-1}{I_1 I_2} \iiint_{V_r} \mathbf{E}_t \cdot \mathbf{J}_r \, dV \tag{2-37}$$

The volume integral reduces to a line integral in most cases.

Antennas with given input voltages such as slots can be described using magnetic currents and we use a mutual admittance matrix for the antenna pair:

$$Y_{12} = \frac{1}{V_t V_r} \cdot \text{reactance} = \frac{-1}{V_t V_r} \iiint_{V_r} \mathbf{H}_t \cdot \mathbf{M}_r \, dV \tag{2-38}$$

By using reciprocity antennas made of linear, isotropic materials, we have equal cross-matrix terms:

$$Z_{12} = Z_{21} \quad \text{and} \quad Y_{12} = Y_{21} \tag{2-39}$$

We calculate self-impedance terms by integrating over the surface of the antenna: for example, the radius of a dipole with the source of the field located at the center of wires or slots.

2-5 METHOD OF MOMENTS

The method of moments (MOM) [6] expands the currents on an antenna (or scattering object) in a linear sum of simple basis functions. The approximate solution is a finite series of these basis functions:

$$f_a = \sum_{i=1}^{N} a_i f_i \tag{2-40}$$

We compute the coefficients by solving integral equations to satisfy boundary conditions on the surface of the antenna (or object). The integral equation can be expressed in the form $Lf_a = g$, where L is a linear operator, usually a scalar product using an integral, f_a the unknown currents given by Eq. (2-40), and g the known excitation

or source function. We substitute the summation of Eq. (2-40) into the linear operator equation and use the scalar product integral to calculate the terms in a matrix equation. The solution of the matrix equation determines the coefficients of current expansion. The MOM produces filled matrices that require time-consuming numerical methods for inversion. The art of the MOM is in choosing basis functions and deriving efficient expressions for evaluating the fields using the basis function currents. Common basis functions are simple staircase pulses, overlapping triangles, trigonometric functions, or polynomials.

The method does not satisfy boundary conditions at every point, only over an integral average on the boundaries. By increasing the number of basis functions, the method will converge to the correct solution. We need to judge how many terms are required for an adequate engineering evaluation. Spending excessive time on the solution cannot be justified if it greatly exceeds our ability to measure antenna performance accurately using real hardware.

2-5.1 Use of the Reactance Theorem for the Method of Moments

We can use the reactance theorem to generate a moment method solution to the currents on a thin-wire antenna. Thin-wire solutions assume that there are no circumferential currents and reduces the problem to filamentary currents. An electric field integral equation (EFIE) satisfies the boundary condition of Eq. (2-25*a*), a zero tangential field at the surface of the wires, but it does not seem explicit in the derivation. The reactance theorem produces an impedance matrix whose inversion yields the coefficients of the current expansion [7]. Similar to many other methods, the Green's function has been solved explicitly to reduce run time. This method [7] uses overlapping sinusoidal currents on V-dipoles as basis function currents and uses the Green's function to calculate the radiation from one V-dipole at the location of a second V-dipole. Both the radiating and receiving dipoles use the same expansion function. Galerkin's method uses the same weighting (or testing) function as the basis function and yields the most stable solutions. The reactance equation (2-37) calculates the mutual impedance between the two dipoles when each has unity current. We compute self-impedance by spacing a second V-dipole one radius away and by using the reactance theorem to calculate mutual impedance, a technique equivalent to the induced EMF method.

The scalar (dot) product between the incident vector electric field and the current density along the dipole reduces the vectors to scalars that can be integrated. The current density acts as the testing or weighting function for the method of moments. Performing the integration means that the current density only satisfies the zero tangential electric field boundary condition in an average sense. If series impedances are placed in the V-dipole, their impedance is added to the diagonal elements of the mutual impedance matrix. To excite the structure, we place a delta voltage source in series with the V-dipole terminals. The solution for the currents can be found by inverting the matrix equation and using the voltage excitation vector starting with the matrix equation

$$[Z_{mn}][I_m] = [V_n] \tag{2-41}$$

After computing the matrix inverse and specifying the input voltage vector, the complex current values are found on the structure:

$$[I_n] = [Z_{mn}]^{-1}[V_m] \tag{2-42}$$

Given the input voltage and the solution for the currents, the input impedance can be calculated. Similarly, the far- and near-field patterns can be calculated by using Eqs. (2-30) and (2-31) of the dyadic Green's function.

The code must satisfy Kirchhoff's current law at the junction between groups of V-dipoles, which adds a constraint to the currents. Because an overlapping sinusoidal basis function closely follows the actual currents normally excited on dipoles, the segments can be on the order of a quarter-wavelength long or more and yield acceptable results. Basis functions that closely follow expected current distributions are sometimes called *entire domain functions*. These reduce the size of the matrix to be inverted but require more complicated calculations for matrix terms and radiation. Although the concept of a V-dipole was expanded to a V rectangular plate [8], the method is only a subset of general integral equation solutions. This approach generates a simple impedance matrix formulation easily understood from an engineering point of view.

2-5.2 General Moments Method Approach

The method of moments can solve other types of electromagnetic problems: for example, electrostatic problems involving charges and dielectrics [9]. These solutions can determine the characteristic impedance of transmission lines useful in the design of antenna feeders. All moment method solutions are found from the solution of integral equations over boundary conditions. The boundary conditions can be either the tangential electric field (EFIE) or magnetic field (MFIE) conditions given by Eq. (2-25*a*,*b*) or a combination applied using an integral scalar product. We need a combination for closed bodies near an internal resonance frequency (resonant cavity) because the solutions exhibit resonances that make the solution invalid over a narrow frequency range. The method of moments can be applied to dielectric bodies when we use the constitutive relations of Eqs. (2-25) and (2-26), where the formulations for dielectric bodies use either volume or surface integrals [9].

Consider the use of the electric field integral equation (EFIE) with metal surfaces. We expand the currents on the objects using basis functions $\mathbf{B}_m(\mathbf{r}')$ with coefficients I_m:

$$\mathbf{J}(\mathbf{r}') = \sum I_m \mathbf{B}_m(\mathbf{r}') \tag{2-43}$$

The basis functions can be applied over a limited range of the structure in piecewise linear functions, which can be staircase pulses, overlapping triangular functions, or sinusoidal basis functions, whereas multiple functions can be applied over the whole or part of the structure for entire domain basis functions. For example, these could be a sum of sinusoidal functions which form a Fourier series representation.

On a PEC surface the tangential electric field vanishes [Eq. (2-28*a*)]. At field point $\mathbf{r}$ along the surface S,

$$\begin{aligned} &\hat{\mathbf{n}} \times [\mathbf{E}_{\text{incident}}(\mathbf{r}) + \mathbf{E}_{\text{scattered}}(\mathbf{r})] = 0 \\ &\mathbf{E}_{\text{scattered}} = \sum_{m=1}^{M} I_m \iint_{s'} \mathbf{B}_m(\mathbf{r}') \cdot \mathbf{G}(\mathbf{r}, \mathbf{r}')\, ds' \end{aligned} \tag{2-44}$$

We can only satisfy Eq. (2-44) using a finite sum in the average sense of an integral. Since the integral and summation operate on a linear function, we can interchange them.

We introduce weighting (or testing) vector functions tangent to the surface $\mathbf{W}_n(\mathbf{r})$ and take the scalar (dot) product of this vector with the sum of electric fields. This limits the result to the tangential component of the electric field:

$$\iint_s [\mathbf{E}_{\text{incident}}(\mathbf{r}) \cdot \mathbf{W}_n(\mathbf{r}) + \mathbf{E}_{\text{scattered}}(\mathbf{r}) \cdot \mathbf{W}_n(\mathbf{r})]\, ds = 0 \tag{2-45}$$

We identify the weighted integral of the incident field with the source and weighted integral of the field radiated by the basis functions (scattered field) as the impedance matrix terms. The integrals over the boundaries are one form of scalar product represented by $\langle \cdot \rangle$ notation. Using unity current on each basis function, we calculate the matrix terms by using the scalar product:

$$Z_{mn} = \left\langle \iint_{s'} \mathbf{B}_m \cdot \mathbf{G}(\mathbf{r}, \mathbf{r}')\, ds', \mathbf{W}_n(\mathbf{r}) \right\rangle = \iint_s \iint_{s'} \mathbf{B}_m \cdot \mathbf{G}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{W}_n(\mathbf{r})\, ds'\, ds \tag{2-46}$$

$$V_n = -\langle \mathbf{E}_{\text{incident}}(\mathbf{r}), \mathbf{W}_n(\mathbf{r}) \rangle = -\iint_s \mathbf{E}_{\text{incident}}(\mathbf{r}) \cdot \mathbf{W}_n(\mathbf{r})\, ds \tag{2-47}$$

The combination of Eqs. (2-46) and (2-47) when integrated over each portion of the source gives a matrix equation:

$$[Z_{mn}][I_m] = [V_n] \tag{2-48}$$

The weighting functions could be as simple as pulse functions, overlapping triangular functions on lines or surfaces (rooftop), piecewise sinusoidal functions, or others. The type of basis functions determines the convergence more than the weighting (testing) functions, which only determine the averaging. Realize that the moment method converges to the exact solution when we increase the number of basis functions, but it is a matter of engineering judgment to determine how many terms give acceptable answers.

Equation (2-47) defines the source voltage occurring over a segment when the formulation uses a piecewise function expansion. The incident voltage is the weighted integral of the incident electric field. For example, the NEC formulation applies an excitation voltage across one segment. The reaction integral formulation of Section 2-5.1 applies a voltage source at the end of a segment. The modeling of sources is an important part of the art in the method of moments.

The expansion of Eq. (2-44) is only one possible moment method solution. We could use the boundary condition on the magnetic field, a combination of the electric and magnetic field conditions on a PEC. If the surface has finite conductivity, the boundary conditions are modified. The moment method is a general method that computes approximate solutions to the currents. Unlike physical optics, the currents do not have to be assumed beforehand but are found as a finite series approximation.

Antenna designers discover that adequate codes are available for most problems. Moment method solutions are typically limited to objects only one or two wavelengths in size, although any method can be stretched. Analysis of large structures becomes intractable because of the large amount of computer memory required and the length

of time needed to calculate the solution. Coarse models may not give totally accurate results but can be useful in determining trends. Given these ideas, remember that physical models can be built that solve the electromagnetic problem instantaneously. We found that it takes considerable time to learn any code, and a new code has to offer considerable advantages or solve problems that the present one cannot solve before we invest our time.

2-5.3 Thin-Wire Moment Method Codes

Thin-wire codes that assume only filamentary currents are readily available. We have experience with NEC, the Richmond code (ASAP), and AWAS [10], a commercial code. All have advantages, but they take time to learn. A commercial code with a graphical interface makes the input and output easier: for example, for NEC. These pay for themselves quickly by saving time. NEC can include plates, but since it uses a MFIE (magnetic field integral equation) for them, it is limited to closed bodies. When accuracy becomes important, it is necessary to decrease the segment length and increase their number. These codes use matrix inversion with calculation time proportional to N^3 and a matrix fill time proportional to N^2. Run time increases enormously as the number of segments increases.

The commercial code AWAS determines the segmentation, while the user of NEC must specify it. The rule is to use at least 10 segments per wavelength, but initial analysis can tolerate the errors due to using fewer segments. The segments should be longer than the diameter, and care must be taken that the segments do not overlap because the radius of the wires is too large. Solid objects, such as plates, can be modeled as wire frames, with the rule that the perimeter of the wire equal the spacing between the wires [11]. This rule can be violated, but a test of the convergence should be made. When we model slots in a solid object, we cannot apply the perimeter equal-spacing rule because the slot will disappear. These codes compute the radiation pattern more accurately than the input impedance due to simplistic source models, and we may have to build the antenna to determine the true input impedance. Of course, an antenna with a good input impedance response that does not have the required pattern is useless.

We can reduce NEC run time if the antenna has symmetry with multiple inputs. The code reduces input by allowing the user to specify symmetry. For example, a multiarm spiral analysis requires only the input of one arm. The various mode voltages are entered after the basic structure impedance matrix has been solved. If an object has M-way symmetry, the matrix fill time is reduced by M^2 and the solution time by M^3. The various voltage modes can be applied afterward. If we add another wire segment after specifying symmetry, the symmetry is destroyed and the program uses the full matrix. The only advantage we gain is in specifying the model because the program solves the full matrix instead of the reduced matrix.

2-5.4 Surface and Volume Moment Method Codes

Antennas made of plates or containing finite plate ground planes can be solved by using wire meshing of a thin-wire code. The method of moments code has been extended to plates [12,13] using a rooftop basis function on both rectangular and triangular patches. The number of basis functions (i.e., matrix size) grows rapidly. One solution is to use entire domain basis functions. These require more complicated integrals, but

they reduce the matrix size. Dielectric portions of the problem lead to either volumetric integrals or various forms of surface integrals that use equivalent currents to replace the internal fields [9,14]. These problems lead to a variety of boundary conditions solved using a finite series of basis function and integral equations to satisfy those boundary conditions approximately.

MOM analysis of antennas mounted on dielectric substrates requires special techniques. Commercial codes determine the currents flowing on these antennas while accounting for the dielectric. Often, Green's functions are found numerically, which increases the execution time. Since the currents are located on the surface and the integrals of the boundary conditions are over the same surface, the singularity of the Green's function causes a numerical problem. For example, the free-space Green's function has the term $1/|\mathbf{r} - \mathbf{r}'|$, which becomes infinite on the surface. Spectral domain methods remove the singularity by using a sum of current sheets on the surface as an entire domain basis function. A uniform plane wave propagating at an angle to the surface excites the current sheet. The actual current flowing on the metal portions is expanded as a sum of these current sheets [15, p. 208ff; 16]. The uniform current sheets are expanded in a spatial Fourier transform as well as the Green's function, and the MOM problem is solved. The Fourier-transformed Green's function no longer has the singularity. When the metallization can be expressed as an infinite periodic structure, the current is expanded as a Fourier series. The infinite periodic structure is used with frequency-selective surfaces and infinite arrays. In this case the fields and currents are expanded in Floquet modes (harmonics).

2-5.5 Examples of Moment Method Models

Figure 2-10 demonstrates the use of a wire mesh to replace a solid plate. We located a resonant ($\approx \lambda/2$) dipole $\lambda/4$ distance over a λ-wide ground plane in the H-plane and offset $3/8\lambda$ from one edge. This is repeated in Figure 2-20 using GTD analysis. The rods only run parallel to the dipole because cross wires do not have currents induced on them in the ideal world of analysis. The circumference of the rods equals the spacing between the rods and forms an equivalent solid plate. An actual antenna could use smaller-diameter rods and work as effectively as the solid plate and would reduce weight and wind loading. NEC analysis produces the same pattern as the GTD analysis of Section 2-7.2, except that the E-plane size of the rods alters the backlobe predicted by GTD to some extent, because that analysis assumes infinite-length rods.

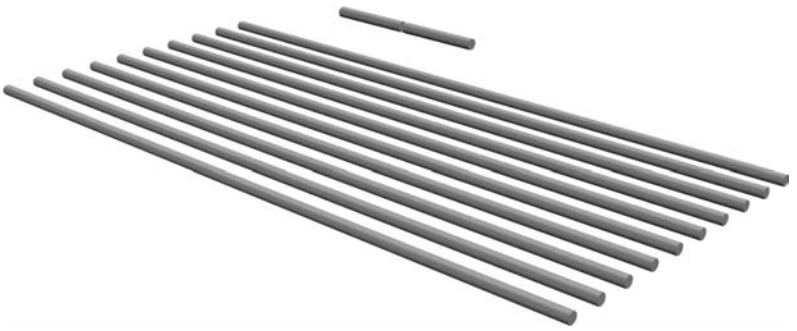


FIGURE 2-10 Use of a wire mesh to replace a solid plate for dipole over a ground plane in a MOM calculation.

Over most of the pattern angles the two analyses produce identical results. The NEC analysis accounts for the mutual impedance between the dipole and its image in the finite ground plane. For impedance calculations a small ground plane gives almost the same reaction to the antenna as an infinite ground plane.

Figure 2-11 shows a wire frame model of a cell phone. The model contains more wires than necessary for $\lambda/10$ spacing, but more wires improve the geometry match. When using crossed wires that shield both polarizations, we reduce the wire circumference in half since the wires approach the squares from four sides. The small wire antenna must be connected to the wire grid of the model to generate proper currents on the box. Either we restrict possible locations of the antenna or we must distort the wire grid locally. You should write an automatic grid generator if you use this analysis often. Consider that you need to specify whether an edge wire should be generated when two plates share the same edge. The hand holding the cell phone and the head nearby have significant effect on the antenna performance. The model given in Figure 2-11 has limited use. We need either a moment method analysis, such as WIPL-D, which includes volume dielectric structures, or FDTD, which can include complex material structures to model the head and produce good results.

Figure 2-12 illustrates a wire frame model of an airplane used for low-frequency analysis. Antennas mounted on free-flying models such as airplanes or spacecraft will excite the structure. Electrically, small antennas can excite the entire vehicle as an antenna. For example, a small antenna mounted on a large ground plane that would produce vertical polarization can excite the wings or fuselage and the entire system will

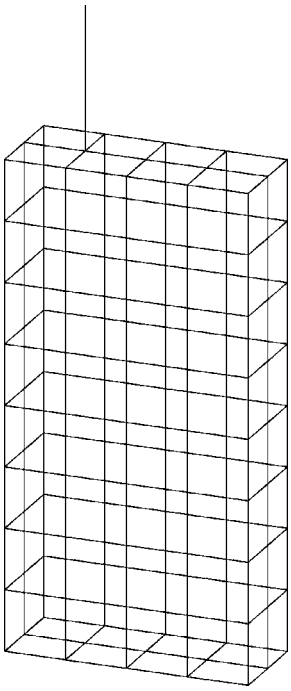


FIGURE 2-11 Wire frame MOM model of a cellular telephone handset with an antenna connected to the mesh.

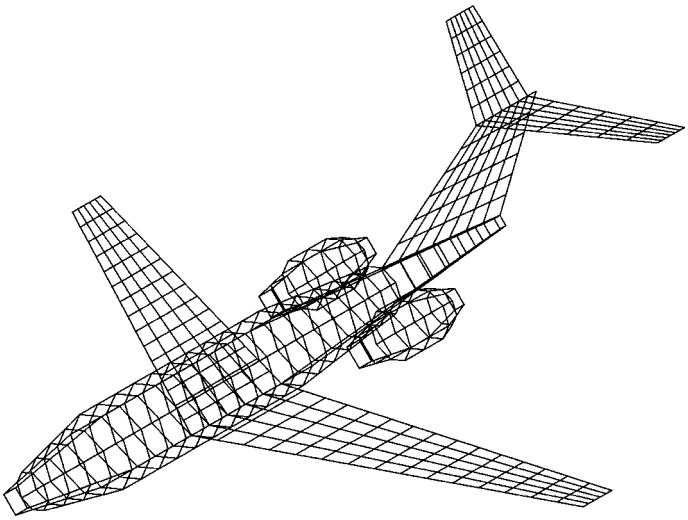


FIGURE 2-12 Wire frame MOM model of an airplane.

radiate horizontal polarization. Models similar to Figure 2-12 can eliminate surprises. The model restricts antenna mounting locations to the wire positions and may require local distortions of the grid.

Moment methods can include solid plates. Figure 2-13 shows an open waveguide horn analysis that uses a combination of plates and a single-feed wire monopole [12]. Locating the monopole or a small dipole inside the waveguide produces excitation of the waveguide mode that feeds the horn. Even though the model does not necessarily produce accurate impedance information, the model accurately calculates the pattern generated by the currents excited in the walls. We can either use an aperture method

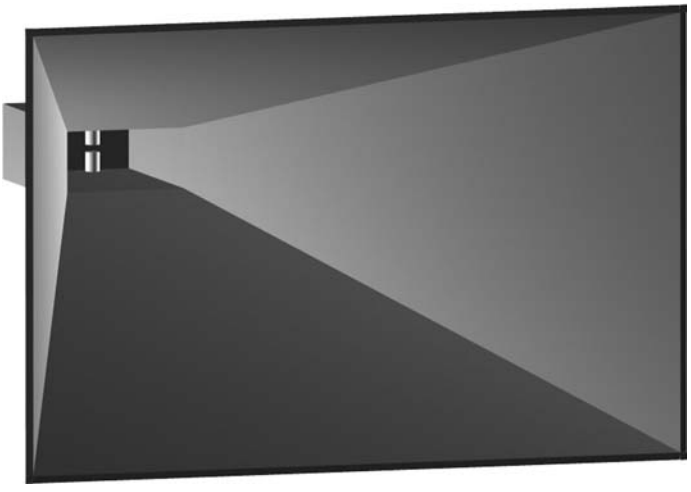


FIGURE 2-13 MOM model of a pyramidal horn using flat plates fed by a small dipole. (From [14, p. 229].)

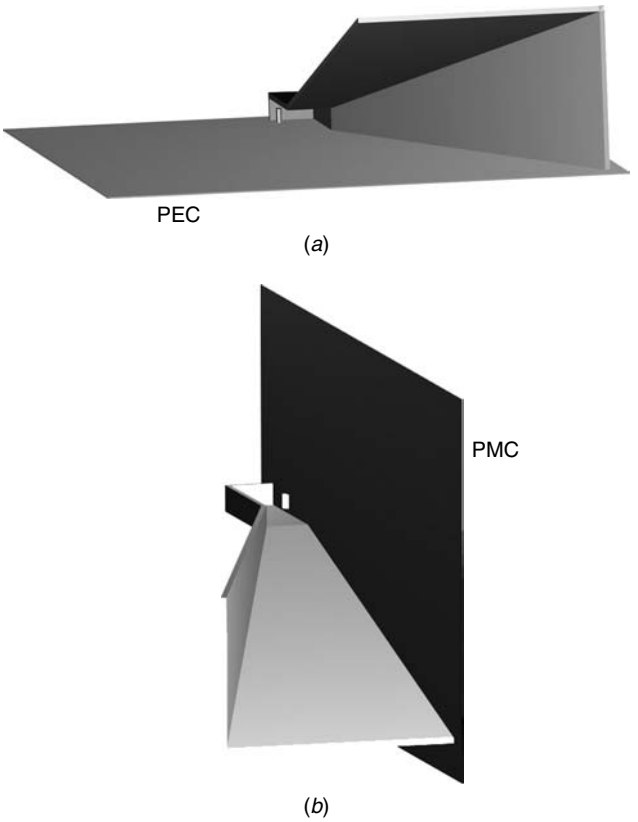


FIGURE 2-14 Use of electric and magnetic walls to reduce the model size in MOM analysis of a pyramidal horn: (*a*) PEC wall divides the horn; (*b*) PMC wall divides the horn.

for the horn that replaces the aperture fields or use the currents excited in the walls to calculate the pattern. Either method works for the front lobe. The moment method calculation requires significantly greater calculation time but produces results that better match measurements in all directions. Figure 2-14 demonstrates how to reduce calculation time by using planes of symmetry in a moment method analysis. In this case the small dipole feed is separated by two equally fed closely spaced dipoles. The right–left symmetry of the antenna allows reduction of the model by half. A vertical PMC wall divides the antenna into two parts, with only one remaining in the analysis. A horizontal PEC conductor divides the remaining model in half because halfway between the dipole feed is a virtual short circuit. Figure 2-14 contains only one-fourth the size of the original problem. Since matrix inversion requires N^3 calculations for an $N \times N$ matrix, dividing the analytical model down to one-fourth size reduces this calculation by a 64 : 1 factor. This also reduces the matrix element (fill time) calculations by 16 : 1. Reducing the model by using symmetry planes enables the solution of larger problems and reduces calculation time.

Analyses in later chapters use the moment method to predict antenna performance. Wire frame and plate analyses determine vehicle and mounting structure pattern effects. The moment method produces excellent analyses because it determines the approximate

current distribution as a sum of simple basis functions and we need not start with an assumed current distribution on the antenna.

2-6 FINITE-DIFFERENCE TIME-DOMAIN METHOD

The finite-difference time-domain (FDTD) method solves the coupled Maxwell's curl equations directly in the time domain by using finite time steps over small cells in space. The method reduces the differential equations to difference equations that can be solved by sets of simple equations. The method alternates between the electric and magnetic fields solved at locations a half-step apart because central differences are used to approximate derivatives. A 1966 paper by Yee [17] described the basic method that many authors have improved upon, but the original method remains the approach of choice.

FDTD can solve many types of electromagnetic problems, of which antenna analyses are only one type. Computer memory and speed limit the size of problems that can be solved, but larger and larger problems can be solved as the cost of computing keeps reducing. Besides antenna problems, the method is applied to microwave circuits, biological interaction with electromagnetic waves, optics, and radar cross-section problems. The number of uses expands every day. The method allows each cell to be made of different materials, leading to the solution of volumetric complex structures. The solution of the equations is robust and the errors are well understood.

Currently, the method solves moderately small antenna problems on the order of a few wavelengths. Of course, faster and larger computers can solve larger problems, especially if the analyst has patience. FDTD handles microstrip antennas with their complex layering of dielectrics, including a finite ground plane without the use of complex Green's functions required of frequency-domain solutions. The interaction of antennas with the near environment, such as the effect of the head on cellular telephone handsets, can be solved. In this case the complex electromagnetic properties of the head can be described as cells each with different electrical properties. In addition to giving a solution to the radiation pattern and allowing characterization of the communication system, it can provide insight into the radiation safety concerns of users. The method handles the solution of the interaction of antennas with the human body in a straightforward manner for prediction of biomedical applications, such as electromagnetic heating for cancer treatment.

Learning to apply the technique, whether formulating your own routines or using a commercial code, will yield insight for design. The method can produce time-domain animated displays of the fields that show radiation centers and where the fields propagate, but the user must learn to interpret these new displays. It will be worth your effort to learn this task. The time-domain responses using impulse signals can produce solutions over a wide band of frequencies when converted to the frequency domain using the discrete Fourier transform (DFT). The only drawback is the computer run time required.

2-6.1 Implementation

By using a direct implementation of Maxwell's curl equations in the time domain, you do little analytical processing of the equations. No vector potential or Green's

functions are developed as in frequency-domain methods. Although the antenna may be volumetrically complex and contain many different materials, the method yields sparse matrices rather than the dense matrices produced by moment methods. It is a direct solution that does not require the inversion of large matrices and includes only nearest-neighbor interactions. Having only nearest-neighbor interactions means that it is possible to run problems on parallel machines.

You need to embed the antenna in a rectangular region and divide it into rectangular cubical cells with sizes ranging from 10 to 20 samples per wavelength at the highest frequency where analysis is desired. The outer surfaces contain absorbing boundaries to eliminate reflections that would produce errors. Formulating absorbing boundary conditions has been a significant part of the method. You need to locate a solution surface between the absorbing boundaries and the antenna outer surface where we compute currents by using the equivalence theorem. The DFT of the time response determines the radiation pattern at a given frequency after the equivalent currents are found. If you need the pattern amplitude in only a few directions, the time-domain radiation can be found directly: for example, the gain in one direction.

We can formulate some problems in one or two dimensions if they possess symmetry instead of the three-dimensional rectangular cube. The solution time is reduced dramatically, and the time-animated presentation may provide sufficient insight when the radiation pattern is found in two dimensions. Because this is a time-domain analysis, we need to excite the structure with a pulse. You use the pulse frequency power response to normalize the patterns and compute gain. When the formulation includes the material losses, the efficiency of the antenna can be found since the dissipation in the inner cells prevents the radiation from reaching the outer surface.

2-6.2 Central Difference Derivative

Numerical derivatives have greater potential for errors than integrals, but FDTD uses them to reduce Maxwell's differential curl equations to simple difference equations. A second-order accurate formula for a derivative can be found by using central differences instead of using the difference between the value at a location or time and the value at the next point in a sequence of evenly spaced points:

$$\frac{\partial f}{\partial u} = \frac{f(u_0 + \Delta u/2) - f(u_0 - \Delta u/2)}{\Delta u} + O(\Delta u)^2 \tag{2-49}$$

We can use finite differences to solve the curl equations provided that we use electric and magnetic fields spaced at half intervals because each is related to the derivatives of the other field and we want to use central differences to reduce error. Because Maxwell's equations involve time derivatives, we need to calculate the electric and magnetic fields at interspersed half time intervals.

2-6.3 Finite-Difference Maxwell's Equations

Consider Maxwell's curl equations in the time domain, including lossy materials:

$$\begin{aligned} \frac{\partial \mathbf{H}}{\partial t} &= -\frac{1}{\mu}(\nabla \times \mathbf{E} - \mathbf{M} + \sigma^* \mathbf{H}) \\ \frac{\partial \mathbf{E}}{\partial t} &= -\frac{1}{\varepsilon}(\nabla \times \mathbf{H} - \mathbf{J} + \sigma \mathbf{E}) \end{aligned} \tag{2-50}$$

Equations (2-50) contain the source currents **J** and **M** and include losses due to conducting dielectric material σ and magnetic material losses σ^*. Both equations have the same form, with only an interchange of symbols. Expanding the curl operator, we get the following equation for the x-component of the magnetic field:

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu}\left(\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} - M_x - \sigma^* H_x\right) \tag{2-51}$$

The x-component of the electric field has the same form but with the interchanges $H \to E$, $E \to H$, $M \to J$, and $\sigma^* \to \sigma$. You obtain the equations for the y- and z-components by a cyclic variation (repeating pattern of interchanges) $x \to y \to z \to x \to y$, and so on. For example, the equations are reduced to two dimensions by leaving out the y-component.

FDTD calculates the field at discrete times and locations on a grid. The fields can be represented as an indexed function using integers:

$$f(i\ \Delta x, j\ \Delta y, k\ \Delta z, n\ \Delta t) = f(i, j, k, n)$$

Because we use central differences [Eq. (2-49)], for derivatives, and the magnetic (electric) field is found from the space derivative of the electric (magnetic) field, the magnetic and electric fields need to be spaced a half-space interval apart. The time derivative becomes

$$\frac{\partial f(i, j, k, n)}{\partial t} = \frac{f(i, j, k, n + \frac{1}{2}) - f(i, j, k, n - \frac{1}{2})}{\Delta t}$$

and means that the electric and magnetic components are interspersed at $\Delta t/2$ times that which produces a leapfrog algorithm. We substitute these ideas into Eq. (2-51) to derive the time-stepping equation for one component:

$$\begin{aligned} H_x(i - \tfrac{1}{2}, j, k, n + 1) = {} & \frac{1 - \sigma^*(i - \frac{1}{2}, j, k)\ \Delta t/2\mu(i - \frac{1}{2}, j, k)}{1 + \sigma^*(i - \frac{1}{2}, j, k)\ \Delta t/2\mu(i - \frac{1}{2}, j, k)} H_x(i - \tfrac{1}{2}, j, k, n) \\ & + \frac{\Delta t/\mu(i - \frac{1}{2}, j, k)}{1 + \sigma^*(i - \frac{1}{2}, j, k)\ \Delta t/2\mu(i - \frac{1}{2}, j, k)} \\ & \left[\frac{E_y(i - \frac{1}{2}, j, k + \frac{1}{2}, n + \frac{1}{2}) - E_y(i - \frac{1}{2}, j, k - \frac{1}{2}, n + \frac{1}{2})}{\Delta z}\right. \\ & - \frac{E_z(i - \frac{1}{2}, j + \frac{1}{2}, k, n + \frac{1}{2}) - E_z(i - \frac{1}{2}, j - \frac{1}{2}, k, n + \frac{1}{2})}{\Delta y} \\ & \left. - M_x(i - \tfrac{1}{2}, j, k, n + \tfrac{1}{2})\right] \end{aligned} \tag{2-52}$$

FDTD uses similar equations for the other components [18,19].

Yee's Cell Figure 2-15 shows one cubic cell and the components of the fields. When we consider the upper face, we see that the magnetic field components are spaced a

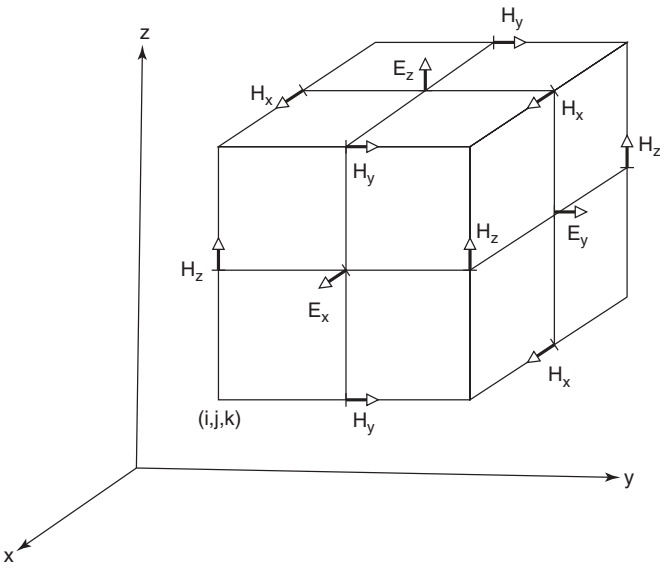


FIGURE 2-15 Unit cell of a Yee space lattice showing time and space separation of electric and magnetic fields in a cell. (From [15], Fig. 1, © 1966 IEEE.)

half space interval from the central electrical field and the arrows show the direction of fields. Although it would appear that the electric field is different on the upper and lower face along the z-axis, the method assumes that the field is constant throughout the cell. The magnetic fields shown are at the center of adjoining cells.

A leapfrog solution uses stored values of the electric fields to calculate the magnetic fields at a half time interval later and stores these values. In the second step the solution takes another half time step and uses the stored values of the magnetic fields to calculate the electric fields. The method gains stability by using the half time steps and by solving for both electric and magnetic fields. Although the fields are a half time step out of synch, we can average between the two half time steps to produce simultaneous fields at a point, but we only need to do this when calculating equivalent currents on the surface used for far-field pattern calculations.

2-6.4 Time Step for Stability

You need to pick the time step to produce a stable solution. Consider a plane wave traveling through the cubes. If the time step is too large, the wave can pass through more than one cell for each time step. At that point the solution cannot follow the actual wave propagation and fails. We must reduce the time step until it is less than the Courant condition or the wave propagation rate. Consider the fastest-moving wave in the problem, usually free space, and for equal sides to the cube, we compute the time step from the velocity and cell length:

$$v\ \Delta t \leq \frac{\Delta x}{\sqrt{d}} \tag{2-53}$$

The cell length is Δx and the number of dimensions is d. The time step must be lower for conducting materials ($\sigma > 0$) to produce a stable solution. The magic step uses

the equality and produces the most stable solutions. If you pick unequal sides to the rectangular cell, Eq. (2-53) is modified.

2-6.5 Numerical Dispersion and Stability

FDTD analyses produce solutions that fail to propagate through the cells at the proper phase velocity in all directions. The propagation velocity depends on the cell size in wavelengths; it has a frequency-dependent component. You need to consider this numerical dispersion because it affects accuracy. Because the waves travel at different velocities in different directions, the dispersion problem increases for large structures where many time steps must be taken. After many steps, signals disperse because they have taken different routes and fail to add together in the correct phase. Finer cells solve the problem, but the computation requirements grow rapidly. The equation for the propagation constant can be found from considering the FDTD formulation to produce the following equation for three-dimensional problems:

$$\left(\frac{1}{c\ \Delta t}\sin\frac{\omega\ \Delta t}{2}\right)^2 = \left(\frac{1}{\Delta x}\sin\frac{k'_x\ \Delta x}{2}\right)^2 + \left(\frac{1}{\Delta y}\sin\frac{k'_y\ \Delta y}{2}\right)^2 + \left(\frac{1}{\Delta z}\sin\frac{k'_z\ \Delta z}{2}\right)^2 \tag{2-54}$$

The factor k'_x is the FDTD propagation constant in the cells along the x-axis, only approximately the same as k_x, the actual propagation constant in the structure. The y- and z-axes have similar problems. If you take the limit as cell length approaches zero, $u \to 0$, and so on, then $\sin(au)/u \to a$. Because $\Delta t \to 0$ as the cell size shrinks for the solution still to satisfy the Courant limit, Eq. (2-54) reduces to the expression

$$\left(\frac{\omega}{c}\right)^2 = k_x'^2 + k_y'^2 + k_z'^2 \tag{2-55}$$

Equation (2-55) is the normal propagation constant equation for a plane wave in space and shows that the cell propagation constants converge to the correct values as the cell size shrinks. If you formulate a problem in one or two dimensions, you remove terms from the right side of Eq. (2-54) to determine the dispersion relationship.

Absorbing boundary conditions (ABCs) can cause numerical instabilities. ABCs approximate infinite space to simulate radiation by the antenna into space. FDTD problems must be placed in a finite number of cells because each cell requires computer storage. Every FDTD problem uses a finite number of cells for the ABCs with more cells required in the directions of maximum radiation. ABCs degrade as the number of time steps increases and eventually leads to numerical instabilities. A lively research on ABCs has produced good ones, but be aware that most have been found to produce problems at some point. If you write your own analyses, you will need to find appropriate ones. Commercial codes will give their limitations.

At one time, ABCs limited solution dynamic range, but ABCs are now available that produce reflection coefficients from 10^{-4} to 10^{-6}. Numerical dispersion limits the dynamic range as well. Remember that the antenna will be modeled with small cubes that limit the resolution of the results. The errors of modeling lead to solution errors that limit the dynamic range.

2-6.6 Computer Storage and Execution Times

The antenna to be analyzed is modeled by a set of cubic cells. Choosing an appropriate number is an art. Similarly, it will be necessary to have a meshing program. Using a

two-dimensional model will greatly reduce computer storage and run time. Remember that our purpose should be to gain insight unavailable from measurements. The calculations require the storage of three components of both the electric and magnetic fields in each Yee cell. Because we solve the problem in the time domain, the components are only real numbers, unlike frequency responses, which use complex numbers for each component. The material properties of the cells can be indicated with short 1-byte integers provided that there are no more than 256 different ones. Single-precision storage of the components requires 30 bytes for each cell; double-precision storage requires 54 bytes. A three-dimensional problem with 200 cells on a side contains 8 M cells and would need 240 Mbytes of storage for single-precision and 432 Mbytes for double-precision components.

At each time step approximately 10 floating-point operations (flops) are needed for each component in each cell. We must run the time steps until the input pulse has peaked and died out in each cell. This takes about 10 times the number of cells in the longest direction (maximum number along one axis). The three-dimensional problem with 200 cells on a side runs for 2000 time steps and requires 60 flops times the number of cells. The solution needs $2000 \times 8\,\text{M} \times 60$ flops = 960 Gflops for completion.

2-6.7 Excitation

We specify the excitation of an antenna in the time domain since FDTD operates in the time domain. If all we need is a single-frequency solution, a ramped sinusoidal waveform can be applied. The waveform is tapered from zero in about three cycles and the FDTD solution steps continue until a steady state is reached. It is more efficient to use a waveform that gives a wide-frequency-range response after performing a discrete Fourier transform on the radiating boundary to compute equivalent currents used at a given frequency. The computer storage and run times are the same for the wideband response as the single-frequency response.

A suitable wide-bandwidth excitation is the differentiated Gaussian pulse shown in Figure 2-16:

$$V_{\text{inc}}(t) = -V_0 \frac{t}{\tau_p} \exp\left[-\frac{(t/\tau_p)^2 - 1}{2}\right] \tag{2-56}$$

We calculate the frequency response of the differentiated Gaussian pulse from the Fourier transform of Eq. (2-56):

$$V_{\text{inc}}(\omega) = -j\omega\sqrt{2\pi}\ \tau_p^2 V_0 \exp\left[-\frac{(\omega\tau_p)^2 - 1}{2}\right] \tag{2-57}$$

The spectrum of Eq. (2-57) peaks for $\omega_p = 1/\tau_p$. Figure 2-17 gives the normalized frequency response and shows that the −20-dB-level normalized frequency extends from 0.06 to 2.75. For example, if we wanted to center the frequency response at 10 GHz, the normalizing pulse time is easily found:

$$\tau_p = \frac{1}{2\pi(10 \times 10^9)} = 1.592 \times 10^{-11}\,\text{s} = 15.92\ \text{ps}$$

A check of Figure 2-17 shows that the antenna frequency response could be found from 2 to 22 GHz with only a 10-dB loss in dynamic range compared to the response at 10 GHz. A single time response computation yields a wide-frequency-range response.

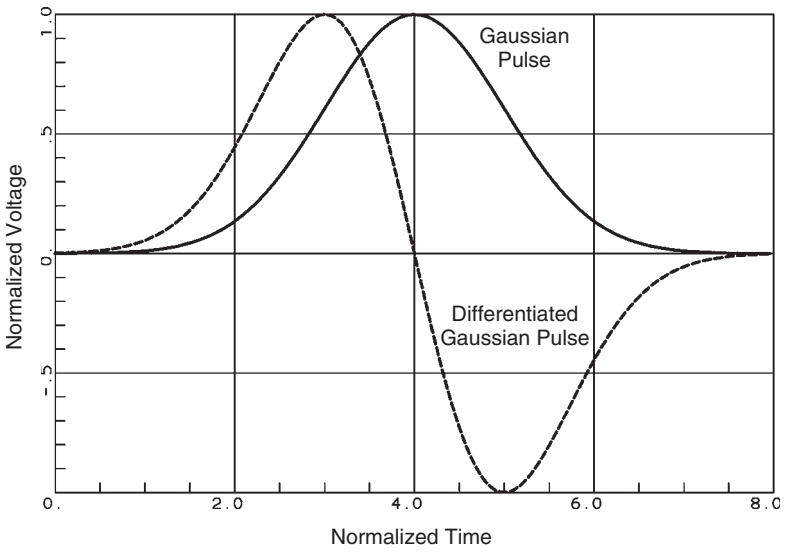


FIGURE 2-16 Differentiated Gaussian pulse time response used in FDTD analysis.

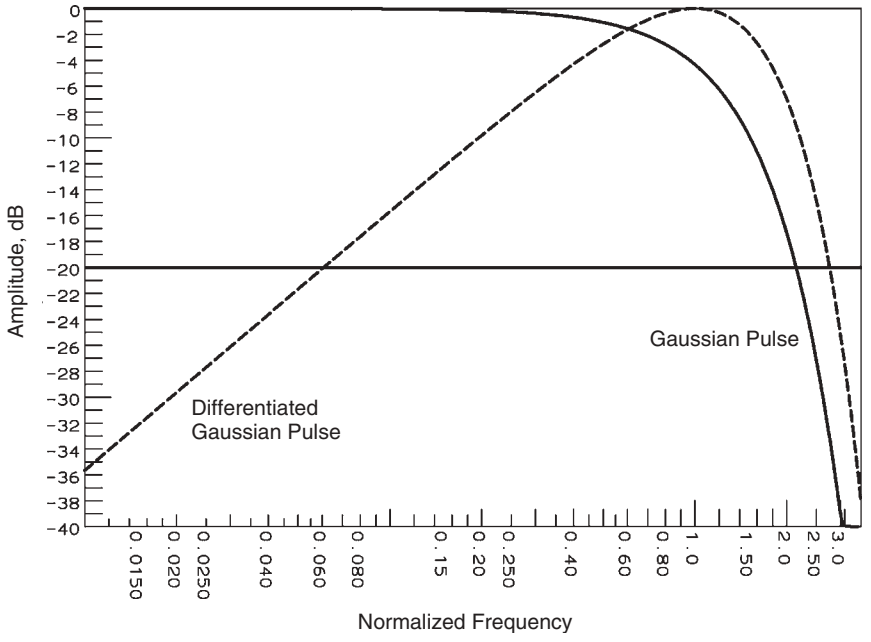


FIGURE 2-17 Differentiated Gaussian pulse normalized frequency response.

A sinusoidal modulated Gaussian pulse produces a narrow-bandwidth excitation useful in visualization because the bandwidth of the pulse can be controlled:

$$V_{\text{inc}}(t) = V_0 \exp\left[-\frac{(t/\tau_p)^2}{2}\right] \sin \omega_0 t \tag{2-58}$$

The unmodulated Gaussian pulse shown in Figure 2-16 has a low-pass frequency response:

$$V_{\text{inc}}(\omega) = \sqrt{2\pi}\tau_p V_0 \exp\left[-\frac{(\omega\tau_p)^2}{2}\right] \tag{2-59}$$

Figure 2-17 gives the low-pass frequency response of the Gaussian pulse with a −4.37-dB response at $\omega_p = 1/\tau_p$. The sinusoidal modulation centers the frequency response of the Gaussian pulse around ω_0, and the convolution of the two frequency responses produces a two-sided response of the Gaussian pulse.

2-6.8 Waveguide Horn Example [19]

The literature contains solutions for the patterns of a number of antennas. Figure 2-18 shows the meshing of a commercial standard gain horn analyzed and compared to measurement. The horn operates from 8.2 to 12.4 GHz. The horn has a radiating aperture that is 110 mm wide and 79 mm high and a bell length of 228 mm. The 51-mm length of the input waveguide and the details of the feed probe were included in the model.

Placing a perfectly magnetic conductor through the midsection of the horn uses symmetry to halve the number of cells to a uniform mesh of 519 × 116 × 183 Yee cells. Ten cells were used on the outside for the ABCs around the sides of the horn and 40 cells for the front ABCs in the maximum radiation direction. The model placed 20 cells between the edge of the horn and the equivalent current surface used for pattern calculations. The longest side of the grid determined the number of time steps at 10 times the number of cells = 5190 time steps. The model contains approximately 11 M Yee cells that require 330 Mbytes of computer storage. Assuming that the problem

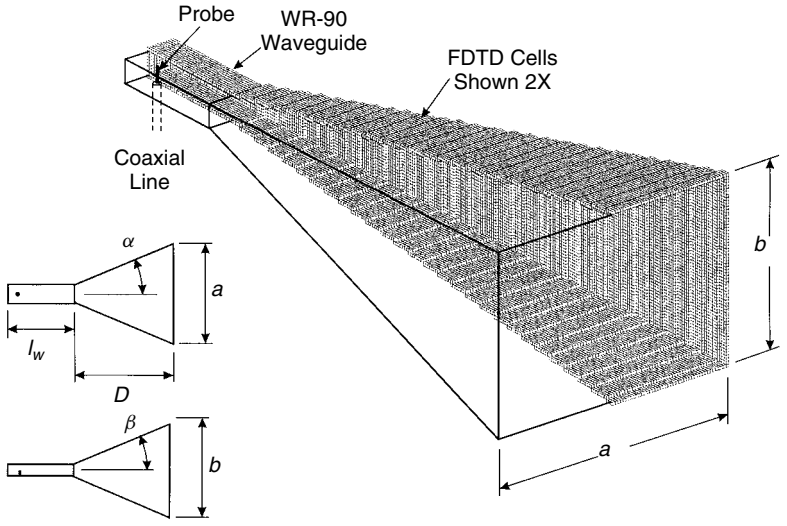


FIGURE 2-18 FDTD model of a standard gain horn. (From [17], Fig. 7.17, © 1998 Artech House, Inc.)

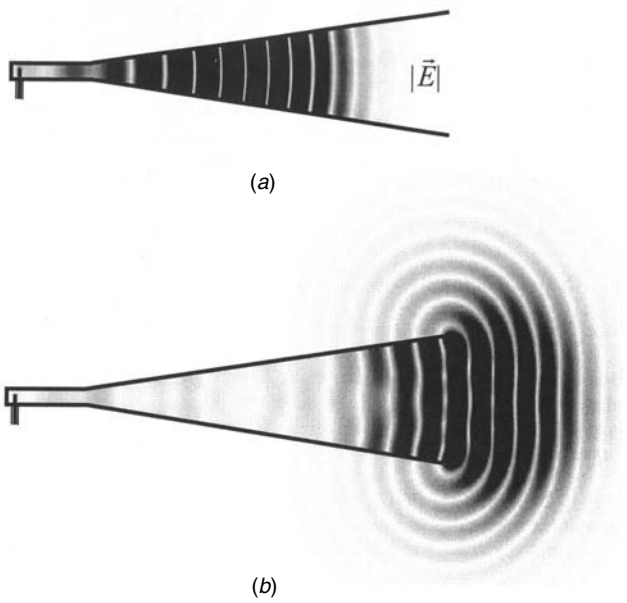


FIGURE 2-19 FDTD calculated electric field in the vertical symmetrical plane of a standard gain horn: (*a*) early time with a pulse in the throat; (*b*) pulse leaving the mouth of the horn. (From [17], Fig. 7.20, © 1998 Artech House, Inc.)

takes 60 flops per cell for each time step, the solution required 3.43 Tflops of computer calculations.

The initial calculation used a differentiated Gaussian pulse excitation with $\tau_p = 15.9$ ps that centered the response at 10 GHz. The calculation produced patterns that matched measurements. A second calculation used a sinusoidal modulated Gaussian pulse with the time constant 79.6 ps. This pulse time constant gives a normalized frequency of 2 GHz for the Gaussian pulse. The −3-dB frequency is 0.83 times the normalizing frequency. The pulse is centered at 10 GHz with a 3-dB bandwidth of 3.32 GHz. Figure 2-19 shows the fields when the pulse reached the horn aperture. Note the high fields in front of the horn and the amount of fields still radiating beyond and behind the aperture. By using a sinusoidal modulated pulse, the visual display contains nulls that improve its clarity.

2-7 RAY OPTICS AND THE GEOMETRIC THEORY OF DIFFRACTION

Ray optics can give you a good physical feel for radiation and spur design ideas, but we need to question the accuracy of their use. Ray optics or geometric optics (GO) methods come from the design of lens and optical reflectors where the wavelength is very short compared to the size of the object being analyzed, whereas we may be interested in analyzing or designing an antenna on a structure only a few wavelengths in size. Below we show that GO is essentially correct over most of the radiation sphere and that by using elements of the geometric theory of diffraction [GTD (UTD)], the pattern prediction can be improved. In this case improvement means that we will increase the area of the radiation pattern that becomes more accurate. You will discover that it takes

an increasing amount of effort to improve small areas of the pattern prediction, and at some point you should decide that it fails to give enough improvement to justify the work. Your real design purpose is to determine antenna dimensions that produce the desired antenna response. Of course, as the expense of the antenna increases, your customer may demand better predictions of the final result, and then the cost of a better analysis is justified. You need to accept a new approach. Even though a part of the pattern prediction shows errors, obvious discontinuities, it only means that the pattern is inaccurate in directions near them and that over most of the radiation sphere the prediction is essentially correct.

Discussion of this method begins with simple examples given in two-dimensional space that introduce the ideas behind GO and GTD. These examples can ignore the details of rotation of polarization directions because the waves are either polarized with the electric field normal to the page or located in the plane of the page. We consider radiation blockage by objects, the reflection of rays by the objects, and the diffraction of rays around edges that fills in the pattern in the shadow regions and across the boundary of the last reflected ray.

After the discussion of simple examples, the key points of GTD will be given for use in three-dimensional problems. This involves the rotation of coordinate systems so that ray polarizations line up with planes of incidence for reflections, with edges for diffraction and curvature directions on curved surfaces that shed rays around the object into the shadow. You will need to investigate the references if you want to develop your own routines, but this discussion will introduce you to the topic and give you an appreciation of the method so that you can use available computer programs and understand their limitations.

GO uses ray methods to approximate electromagnetics. It is exact only in the limit of zero wavelength (infinite frequency), but we gain useful insight from it at any frequency. It will not give good results close to physical boundaries; but when we include the GTD, the results are accurate down to one-wavelength sizes and are useful at $\lambda/4$ sizes. GO gives us physical insight when we deal with reflectors. We must consider three aspects to use GO fully: (1) ray reflections, (2) polarization, and (3) amplitude variations along the ray path and through reflections.

2-7.1 Fermat's Principle

Rays travel through a medium at the speed of light determined by the index of refraction: $n = \sqrt{\varepsilon_r \mu_r}$. We define the optical path length as $\int_C n\, dl$, where C is a prescribed path in space. Fermat's principle determines the paths of rays between two points. It states that the optical path length is stationary along a valid ray path. An expression is stationary when its first derivatives are zero and the optical path is a minimum (or maximum). We use Fermat's principle to trace ray paths through reflection or refraction by searching for the minimum optical path lengths. We can find more than one possible ray between points because Fermat's principle requires only a local minimum. When we exclude the boundaries of lenses, regions of homogeneous medium, rays travel in straight lines.

2-7.2 *H*-Plane Pattern of a Dipole Located Over a Finite Strip

Figure 2-20 illustrates the geometry of this problem and the various regions of the analysis. The diagram shows the end of the dipole rod with the two rods located

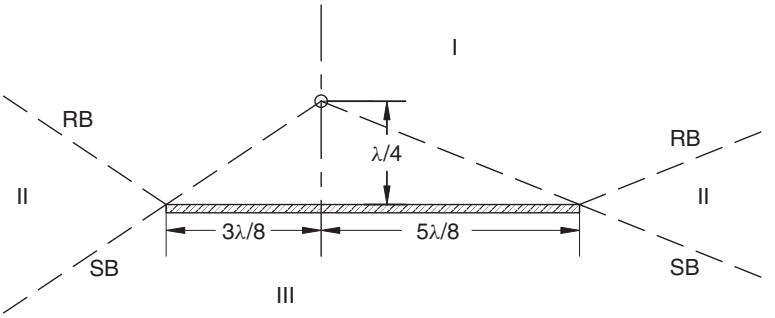


FIGURE 2-20 GTD example using a two-dimensional model of a dipole located over an asymmetrical ground plane.

normal to the page. The dipole pattern is omnidirectional in the page with the electric field directed normal to the page. When we trace rays from the dipole to the finite strip, we discover two significant directions on both sides of the strip. The dashed boundaries labeled RB (reflection boundary) are the directions of the last rays reflected from the strip. Similarly, the dashed boundaries labeled SB (shadow boundary) are the last rays of radiation not blocked by the strip. The radiation in region I results from the sum of the direct radiation from the dipole plus the radiation reflected by the strip. Only direction radiation from the dipole occurs in the two parts of region II. Finally, region III is totally blocked from any radiation by a direct or reflected ray. This region receives rays diffracted around the edges.

If we add the direct and reflected rays in an analysis, we obtain the pattern given in Figure 2-21, which also traces the actual pattern. The pattern, using only the direct and reflected rays, accounts for the phasing between the direct radiation from the dipole and an image dipole located below the strip. If you compare the two traces on Figure 2-21, you see that the two patterns are similar near $\theta = 0$, but the direct plus reflected ray pattern has discontinuities at the SBs and RBs. Figure 2-22 gives the results for the same analysis, but using a 5λ-wide ground-plane strip. When using the larger strip, the two patterns match to about 80°, and in the second case the simple analysis is correct over most of the forward semicircle. Simple geometric optics gives good results for large objects provided that you realize the patterns will contain discontinuities.

Removing the discontinuities requires extra effort. A discontinuity in the pattern cannot exist because shadow and reflection boundaries occur in free space. It takes a material boundary to produce a discontinuous field. But, for example, the tangential electric field must be continuous across even material boundaries. Edge diffraction solves the discontinuity problem. Figure 2-23 gives the pattern of the edge diffraction for both edges normalized to the total pattern. The edge diffraction has matching discontinuities to the sum of the direct and reflected rays at the SBs and RBs. The UTD (uniform theory of diffraction) technique [20, p. 55] calculated these diffractions. When these diffractions are added to the direct and reflected ray radiation, the total pattern given in Figure 2-21 is obtained. The dipole, its image in the ground plane, and the two edge diffractions form a four-element array where each element has a unique pattern. Adding edge diffractions to the geometric optics fields removes the discontinuities and allows calculation of the pattern behind the strip ground plane.

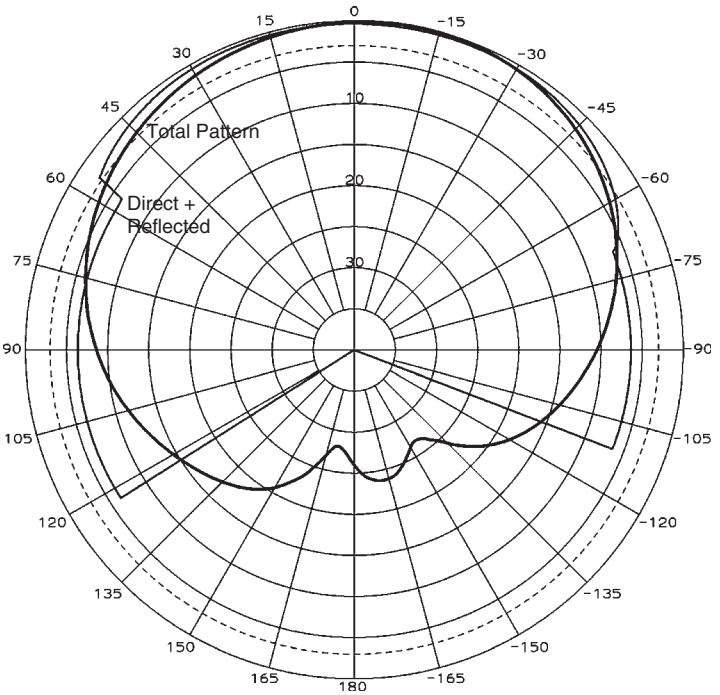


FIGURE 2-21 H-plane pattern of a dipole over asymmetrical ground using direct and reflected rays compared only to a full solution for the 1λ ground plane of Figure 2-20.

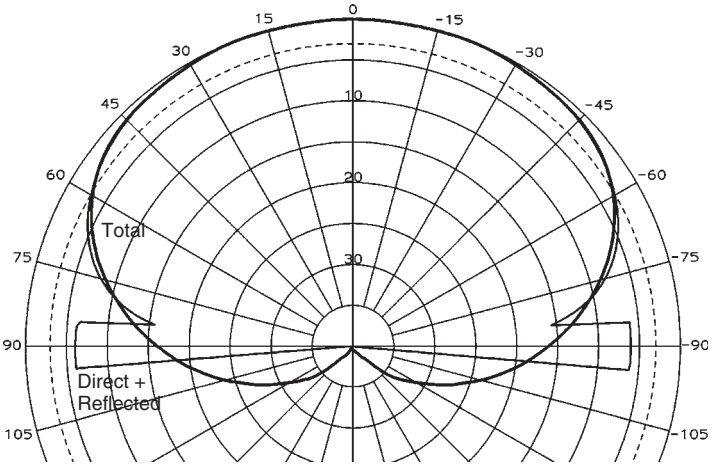


FIGURE 2-22 H-plane pattern of a dipole over symmetrical ground using direct and reflected rays compared only to a full solution for a 5λ ground plane.

2-7.3 E-Plane Pattern of a Rectangular Horn

Figure 2-24 illustrates the cross section of a horn or, in this case, a two-dimensional approximation to a horn. The waveguide feeds the horn and produces a uniform aperture distribution in the E-plane. In this model the direct GO radiation is a constant wedge signal as shown in Figure 2-25 ranging between -15° and $+15^\circ$. The reflected pattern

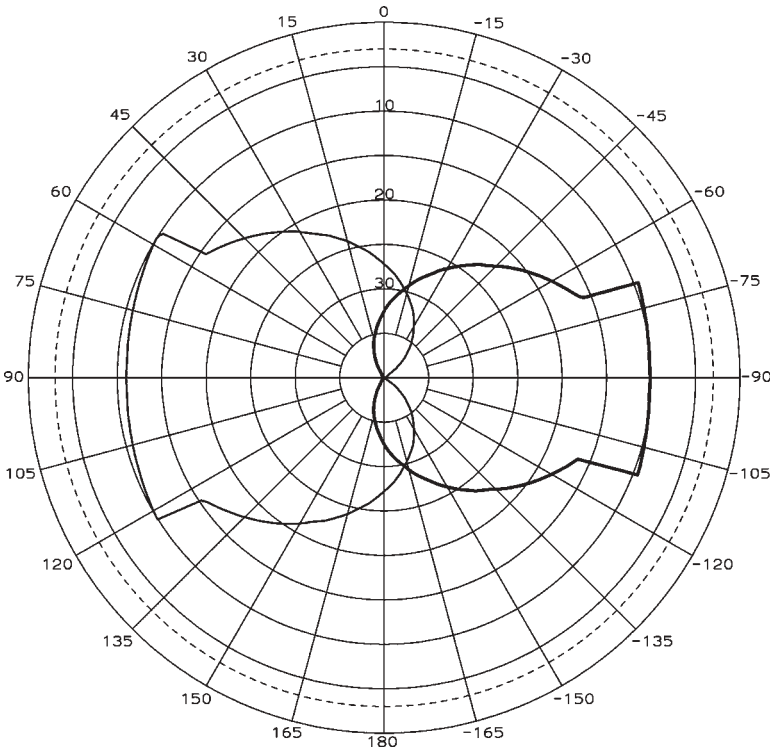


FIGURE 2-23 GTD edge diffraction of an H-plane pattern for an asymmetrical 1λ ground plane under a dipole.

combines with the direct radiation and produces the same pattern. Figure 2-25 also shows the diffraction patterns from the two edges. These peak along the plates and exhibit a discontinuity at the same angle as the GO field. Each diffraction pattern has a discontinuity on one side at 90° because the mouth of the horn blocks the diffraction from the opposite edge. When we add the diffracted fields to the GO field, the pattern shown in Figure 2-26 is obtained. By just adding the three components, we obtain an accurate pattern of the horn over most of the angles of the plot. At 90° we see discontinuities in the pattern caused by not considering enough terms in the GTD calculation. You need to realize that these discontinuities only cause pattern errors at nearby angles. The majority of the pattern is correct.

We need another term to correct the pattern near 90°. The blockage of the diffraction from one edge by the mouth of the horn causes a secondary diffraction at that edge. We call this *double diffraction*. Some available programs do not implement double diffraction because the general three-dimensional double diffraction takes considerable calculation due to the extensive ray tracing required. In these cases you must accept the pattern discontinuities. Some programs calculate double diffraction as an option, but turning on this option will slow the calculations. Figure 2-27 gives the pattern when double diffraction is included. Double diffraction reduces the discontinuity at 90°, but a small discontinuity remains. Adding triple diffraction would reduce this further, but the pattern area affected by the small discontinuity has shrunk. A new discontinuity near 60° appeared in the pattern after adding double diffraction at the

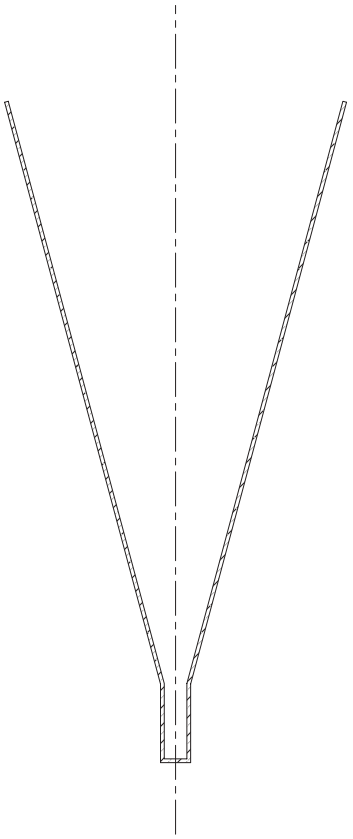


FIGURE 2-24 Geometry of a two-dimensional model of a rectangular horn used for GTD analysis.

mouth of the horn. We could continue to add another term to remove this one or just accept it.

2-7.4 *H*-Plane Pattern of a Rectangular Horn

The tangential electric fields vanish at the walls of the two-dimensional horn in the H-plane. This affects the GO field and produces the following equation for them:

$$E^{\mathrm{GO}} = \cos\frac{\pi\tan\theta}{2\tan\alpha}\frac{e^{-jkR}}{\sqrt{R}} \tag{2-60}$$

Equation (2-60) includes the phasing term and square-root spreading factor of a two-dimensional field. The horn walls tilt from the centerline by the angle α. Figure 2-28 plots the GO field and shows that it vanishes at the walls. We do not expect edge diffraction because the field vanishes at the edges, but Figure 2-28 shows diffraction patterns that peak in the direction of the walls.

We call this new term *slope diffraction*. This new type requires another set of coefficients not identical to the edge (or wedge) diffraction coefficients. While the amplitude of the edge diffraction is proportion to the field incident on the edge, the amplitude of slope diffraction is proportional to the derivative of the field in the direction normal

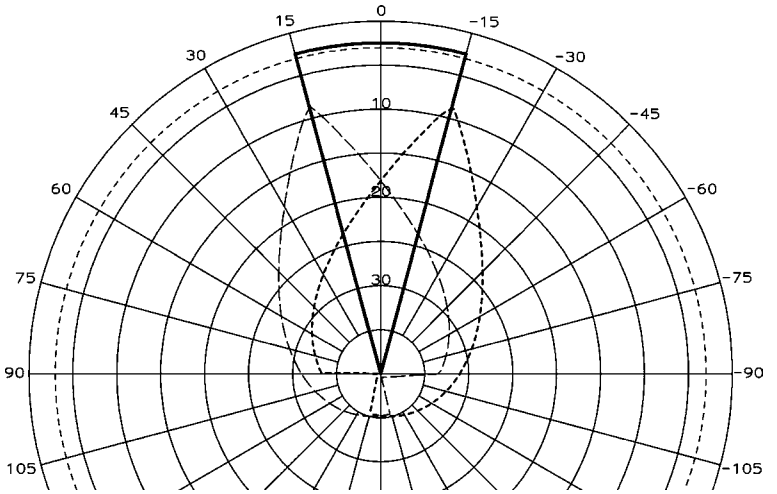


FIGURE 2-25 E-plane pattern of a rectangular horn with a GO term (solid curve) and edge diffractions (dashed curves).

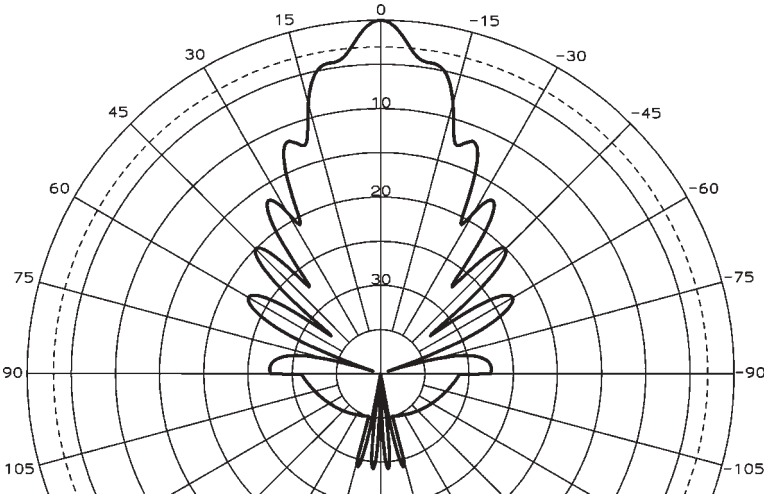


FIGURE 2-26 Combination of GO and edge diffractions in the E-plane pattern of a rectangular horn.

to the edge. We figure the same geometric factors for both edge and slope diffraction but now must calculate the normal derivative of the incident electric field. Figure 2-29 plots the H-plane pattern of the horn. The pattern fails to predict a pattern behind it. The E-plane diffraction produces a back hemisphere pattern for a real horn, but our two-dimensional model does not include the E-plane.

2-7.5 Amplitude Variations Along a Ray

Power decreases in a general ray as the distance from the source increases. If we expand the constant-phase surface (eikonal) about the ray in a Taylor series, we obtain a surface described by its radii of curvature [20, p. 55]. The maximum and minimum

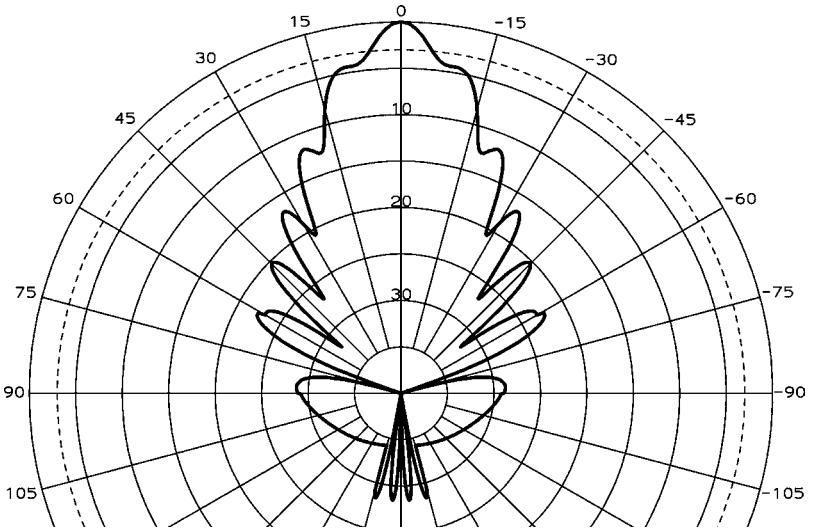


FIGURE 2-27 E-plane pattern of a rectangular horn combining GTD terms of direct GO, edge diffractions, and double diffractions between edges.

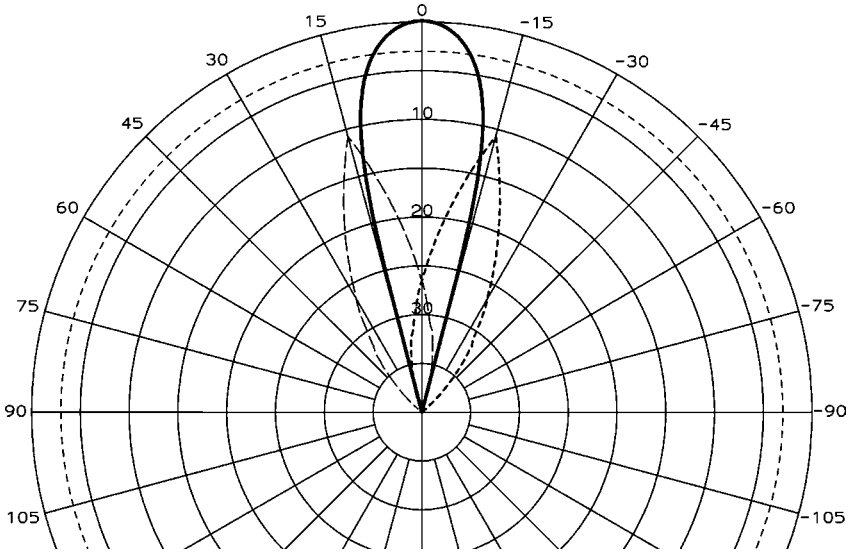


FIGURE 2-28 H-plane pattern of a rectangular horn with a GO term (solid curve) and edge slope diffractions (dashed curves).

values lie in the orthogonal principal planes. These radii of curvature determine the amplitude spread of the wave from point to point on the ray. We compute the ratio of differential areas about the ray at two locations as

$$\frac{dA_2}{dA_1} = \frac{\rho_1 \rho_2}{(\rho_1 + d)(\rho_2 + d)} \tag{2-61}$$

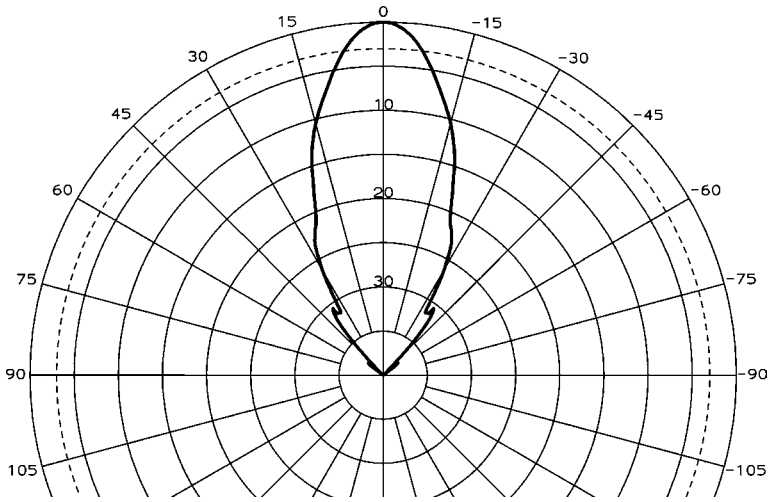


FIGURE 2-29 H-plane pattern of a rectangular horn by GTD analysis by combining direct GO field and edge slope diffraction.

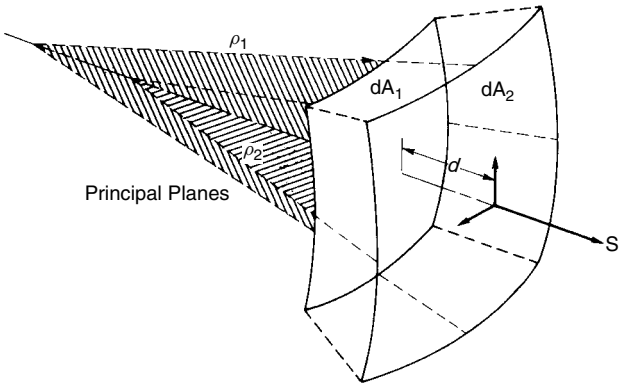


FIGURE 2-30 Astigmatic ray.

where ρ_1 and ρ_2 are the principal radii of curvature and d is the distance between two points on the ray (Figure 2-30). The electric field variation along the ray becomes

$$E_0 e^{-jkd} \sqrt{\frac{\rho_1 \rho_2}{(\rho_1 + d)(\rho_2 + d)}} \tag{2-62}$$

for the astigmatic ray spreading from unequal radii of curvature. When $d = -\rho_1$ or $d = -\rho_2$, GO fails because it predicts an infinite power density. We call these locations *caustics*. Remember that the ray always has differential area and never has any real area as implied by Figure 2-30. We have three special cases of the astigmatic ray:

1. Spherical wave, $\rho_1 = \rho$:

$$E_0 e^{-jkd} \frac{\rho}{\rho + d} \tag{2-63}$$

2. Cylindrical wave, $\rho_1 = \infty$:

$$E_0 e^{-jkd} \sqrt{\frac{\rho}{\rho + d}} \tag{2-64}$$

3. Plane wave, $\rho_1 = \rho_2 = \infty$:

$$E_0 e^{-jkd} \tag{2-65}$$

The plane wave does not spread but has constant amplitude as distance changes. Both cylindrical and plane waves require infinite power, and they are therefore nonphysical, but we find them convenient mathematically.

2-7.6 Extra Phase Shift Through Caustics

We cannot determine the ray amplitude at a caustic but can determine its amplitude and phase on either side. Passage through a caustic causes an extra phase shift to the ray [21, p. 31]. The denominator factors in the square root of Eq. (2-62) produce a 180° sign change when the ray distance factor d passes through either ρ_1 or ρ_2. The square root changes 180° to +90° ($e^{j\pi/2}$) or −90° ($e^{-j\pi/2}$), depending on the direction of movement along the ray. When tracing a ray moving through a caustic in the direction of propagation, you multiply by $e^{j\pi/2}$. The field is multiplied by $e^{-j\pi/2}$ for a ray traced in the opposite direction of propagation.

2-7.7 Snell's Laws and Reflection

We derive Snell's laws of reflection and refraction from Fermat's principle. The two laws of reflection are given as:

1. The incident ray, the reflected ray, and the normal of the reflecting surface at the point of reflection lie in the same plane.
2. The incident and reflected rays make equal angles with the surface normal.

Implicit in Snell's laws is the idea that locally the wavefront behaves like a plane wave and that the reflector can be treated as a plane. Given the direction of the incident ray $\mathbf{S}_1$, reflected ray $\mathbf{S}_2$, and the reflector normal $\mathbf{n}$, Snell's laws of reflection can be expressed vectorially [22]:

$$\mathbf{n} \times (\mathbf{S}_2 - \mathbf{S}_1) = 0 \qquad \mathbf{n} \cdot (\mathbf{S}_1 + \mathbf{S}_1) = 0 \tag{2-66}$$

We combine Eq. (2-66) to determine the ray directions before or after reflection:

$$\mathbf{S}_1 = \mathbf{S}_2 - 2(\mathbf{S}_2 \cdot \mathbf{n})\mathbf{n} \qquad \mathbf{S}_2 = \mathbf{S}_1 - 2(\mathbf{S}_1 \cdot \mathbf{n})\mathbf{n} \tag{2-67}$$

Snell's law of refraction can also be expressed vectorially as

$$\mathbf{n} \times (n_2 \mathbf{S}_2 - n_1 \mathbf{S}_1) = 0 \tag{2-68}$$

where n_1 and n_2 are the index of refractions in the two mediums.

2-7.8 Polarization Effects in Reflections

The electric field is orthogonal to the ray direction (a free-space wave) and is described by a two-dimensional polarization space (Section 1-11). We can describe polarization in any conveniently rotated two-dimensional basis vectors in the plane with the ray vector as its normal. We will use a ray-fixed coordinate system that changes direction after a reflection:

$$\mathbf{E}_i = \mathbf{a}^i_{\parallel} E_{i\parallel} + \mathbf{a}^i_{\perp} E_{i\perp} \tag{2-69}$$

where $\mathbf{a}^i_{\parallel}$ is a unit vector in the plane of incidence and $\mathbf{a}^i_{\perp}$ is perpendicular to the plane of incidence. We compute $\mathbf{a}^i_{\perp}$ from the normal to the plane $\mathbf{n}$ at the reflection point and the incident ray unit vector $\mathbf{S}_i$:

$$\begin{aligned} \mathbf{a}^i_{\perp} &= \frac{\mathbf{S}_i \times \mathbf{n}}{|\mathbf{S}_i \times \mathbf{n}|} \\ \mathbf{a}^i_{\parallel} &= \mathbf{a}^i_{\perp} \times \mathbf{S}_i \end{aligned} \tag{2-70}$$

After reflection, we calculate the output ray-fixed polarization vectors using the output ray $\mathbf{S}_r$:

$$\mathbf{a}^r_{\perp} = \mathbf{a}^i_{\perp} \quad \text{and} \quad \mathbf{a}^r_{\parallel} = \mathbf{a}^r_{\perp} \times \mathbf{S}_r$$

$E_{i\parallel}$ is the incident electric field in the direction of $\mathbf{a}^i_{\parallel}$ and $E_{i\perp}$ is in the direction of $\mathbf{a}^i_{\perp}$. Of course, the direction of unit vector $\mathbf{a}_{\parallel}$ changes from incident to reflected rays. The electric field parallel to the reflector surface must vanish on the conductor surface:

$$E_{r\perp} = -E_{i\perp} \tag{2-71}$$

where $E_{r\perp}$ is the reflected field along $\mathbf{a}^i_{\perp}$. We calculate the reflection properties of $E_{\parallel}$ from the corresponding magnetic fields parallel to the surface:

$$H_{r\parallel} = H_{i\parallel} \tag{2-72}$$

By combining Eqs. (2-71) and (2-72), we obtain the dyadic relation for the ray-fixed coordinate system:

$$\begin{bmatrix} E_{r\parallel} \\ E_{r\perp} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} E_{i\parallel} \\ E_{i\perp} \end{bmatrix} \tag{2-73}$$

where $E_{r\parallel}$ and $E_{i\perp}$ are the reflected field components. At each reflection we rotate the polarizations to align $\mathbf{a}_{i\perp}$ with the normal to the plane of incidence. We can express Eq. (2-73) as a dyadic in terms of the incident and reflected wave polarization vectors $\vec{\mathbf{R}} = \mathbf{a}^i_{\parallel}\mathbf{a}^r_{\parallel} - \mathbf{a}^i_{\perp}\mathbf{a}^r_{\perp}$. Of course, the alternative method is to describe polarizations in a fixed three-dimensional coordinate system, but it requires a 3×3 reflection matrix.

2-7.9 Reflection from a Curved Surface

A wave reflected from a curved surface changes its radii of curvature and principal planes. The field along the reflected ray is given by

$$\mathbf{E}_r(s) = \mathbf{E}_{i0} \cdot \vec{\mathbf{R}} \sqrt{\frac{\rho^r_1 \rho^r_2}{(\rho^r_1 + s)(\rho^r_2 + s)}} e^{-jks} \tag{2-74}$$

where s is the distance along the ray from the reflection, ρ_1 and ρ_2 the reflected ray radii of curvature, and $\vec{\mathbf{R}}$ the reflection dyadic. $\mathbf{E}_{i0}$ the incident ray electric field. For a flat surface we use images of the incident ray caustics for ρ_1^r and ρ_2^r, but in general, ρ_1^r and ρ_2^r become

$$\frac{1}{\rho_1^r} = \frac{1}{2}\left(\frac{1}{\rho_1^i} + \frac{1}{\rho_2^i}\right) + \frac{1}{f_1} \quad \frac{1}{\rho_2^r} = \frac{1}{2}\left(\frac{1}{\rho_1^i} + \frac{1}{\rho_2^i}\right) + \frac{1}{f_2} \tag{2-75}$$

where f_1 and f_2 are generalized focal lengths of the surface. The spreading factor of Eq. (2-74) simplifies in the far field:

$$\sqrt{\frac{\rho_1^r \rho_2^r}{(\rho_1^r + s)(\rho_2^r + s)}} \approx \frac{\sqrt{\rho_1^r \rho_2^r}}{s}$$

Kouyoumjian and Pathak [23] derived formulas for the focal lengths of a surface. We start with a surface with principal radii of curvature R_1 and R_2 with directions $\mathbf{u}_1$ and $\mathbf{u}_2$ at the point of reflection. For an incident ray with principal axes defined by unit vectors $\mathbf{x}_1^i$ and $\mathbf{x}_2^i$, we define a matrix relation between the incident ray and surface principal curvature directions:

$$\theta = \begin{bmatrix} \mathbf{x}_1^i \cdot \mathbf{u}_1 & \mathbf{x}_1^i \cdot \mathbf{u}_2 \\ \mathbf{x}_2^i \cdot \mathbf{u}_1 & \mathbf{x}_2^i \cdot \mathbf{u}_2 \end{bmatrix} \tag{2-76}$$

where the determinant is $|\theta| = (\mathbf{x}_1^i \cdot \mathbf{u}_1)(\mathbf{x}_2^i \cdot \mathbf{u}_2) - (\mathbf{x}_2^i \cdot \mathbf{u}_1)(\mathbf{x}_1^i \cdot \mathbf{u}_2)$. Given the angle of incidence θ^i, the following are the focal lengths:

$$\begin{aligned} \frac{1}{f_{1,2}} = {} & \frac{\cos\theta^i}{|\theta|^2}\left(\frac{\theta_{22}^2 + \theta_{12}^2}{R_1} + \frac{\theta_{21}^2 + \theta_{11}^2}{R_2}\right) \\ & \pm \frac{1}{2}\left\{\left(\frac{1}{\rho_1^i} - \frac{1}{\rho_2^i}\right)^2 + \left(\frac{1}{\rho_1^i} - \frac{1}{\rho_2^i}\right)\frac{4\cos\theta^i}{|\theta|^2}\left(\frac{\theta_{22}^2 - \theta_{12}^2}{R_1} + \frac{\theta_{21}^2 - \theta_{11}^2}{R_2}\right)\right. \\ & \left. + \frac{4\cos^2\theta^i}{|\theta|^4}\left[\left(\frac{\theta_{22}^2 + \theta_{12}^2}{R_1} + \frac{\theta_{21}^2 + \theta_{11}^2}{R_2}\right)^2 - \frac{4|\theta|^2}{R_1 R_2}\right]\right\}^{1/2} \end{aligned} \tag{2-77}$$

With a single reflection, we need not compute the direction of the principal axes. We need only the focal lengths. Multiple reflections require knowledge of the reflected-ray principal plane directions. Define the following matrices to determine the directions of the principal axes after reflection:

$$Q_0^i = \begin{bmatrix} \frac{1}{\rho_1^i} & 0 \\ 0 & \frac{1}{\rho_2^i} \end{bmatrix} \qquad C_0 = \begin{bmatrix} \frac{1}{R_1} & 0 \\ 0 & \frac{1}{R_2} \end{bmatrix}$$

$$Q^r = Q_0^i + 2(\theta^{-1})^{\mathrm{T}} C_0 \theta^{-1} \cos\theta_i$$

$$\mathbf{b}_1^r = \mathbf{x}_1^i - 2(\mathbf{n} \cdot \mathbf{x}_1^i)\mathbf{n} \qquad \mathbf{b}_2^r = \mathbf{x}_2^i - 2(\mathbf{n} \cdot \mathbf{x}_2^i)\mathbf{n}$$

where **n** is the surface normal at the reflection point. One principal-axis direction is

$$\mathbf{x}_1^r = \frac{\left(Q_{22}^r - 1/\rho_1^r\right)\mathbf{b}_1^r - Q_{12}^r\mathbf{b}_2^r}{\sqrt{\left(Q_{22}^r - 1/\rho_1^r\right)^2 + (Q_{12}^r)^2}} \tag{2-78}$$

We derive the other from the cross product of Eq. (2-78) and the reflected ray unit vector:

$$\mathbf{x}_2^r = -\mathbf{S}_r \times \mathbf{x}_1^r \tag{2-79}$$

We must reapply Eqs. (2-75) through (2-79) for every reflection.

We use Eqs. (2-75) through (2-79) for analysis, but except for computer optimizations, they cannot be applied directly to synthesis. If we limit the reflectors to figures of rotation, the radii of curvature are given by the meridians and parallels and these problems reduce to two dimensions. Similarly, a cylindrical reflector fed with a cylindrical wave [Eq. (2-64)] reduces the problem to two dimensions. The incident and reflected waves remain in the single plane chosen for the reflector analysis.

2-7.10 Ray Tracing

Tracing rays through a reflector system is conceptually straightforward. Where a ray strikes a reflector, we compute the normal to the surface. By using Eq. (2-67), we solve for the reflected-ray direction. Equation (2-73) determines the polarization effects when we express the incident and reflected rays in the ray-fixed coordinates. We use geometric arguments to determine the amplitude variation along the ray through the reflection instead of the general expressions given above. We experience difficulty when we try to discover the reflection points for given field and source points. No analytical expressions exist for calculating the reflection point of a general surface. The usual computer routines search for the minimum optical path length (Fermat's principle) without using Eq. (2-67), since a local minimum will satisfy this equation.

2-7.11 Edge Diffraction

Keller [24] extended the idea of reflection to edge diffraction by applying a generalized Fermat's principle to the rays. Figure 2-31 illustrates the rays in edge diffraction and the associated polarization directions. The figure shows the edge vector at the diffraction point. The vector cross product between the edge vector and the incident ray points in the direction of the incident plane normal. We measure the diffraction angle of incidence in this plane between the incident ray and the edge normal. Because diffraction obeys a generalized Fermat's principle, the diffracted ray exits at the same angle, similar to the reflected ray angles. The diffracted rays lie in a cone with the edge vector as its axis. The diffracted rays spread the incident power into a cone. Figure 2-31 shows a particular diffracted ray and how we determine the diffracted ray exiting plane.

We define diffracted ray polarization in terms of the incident and diffracted planes. The vectors are parallel and perpendicular to the two planes. Given the edge unit vector **e**, you compute the incident perpendicular polarization vector:

$$\mathbf{a}_{\phi'} = \frac{\mathbf{e} \times \mathbf{S}'}{\sin\beta_0} \tag{2-80}$$

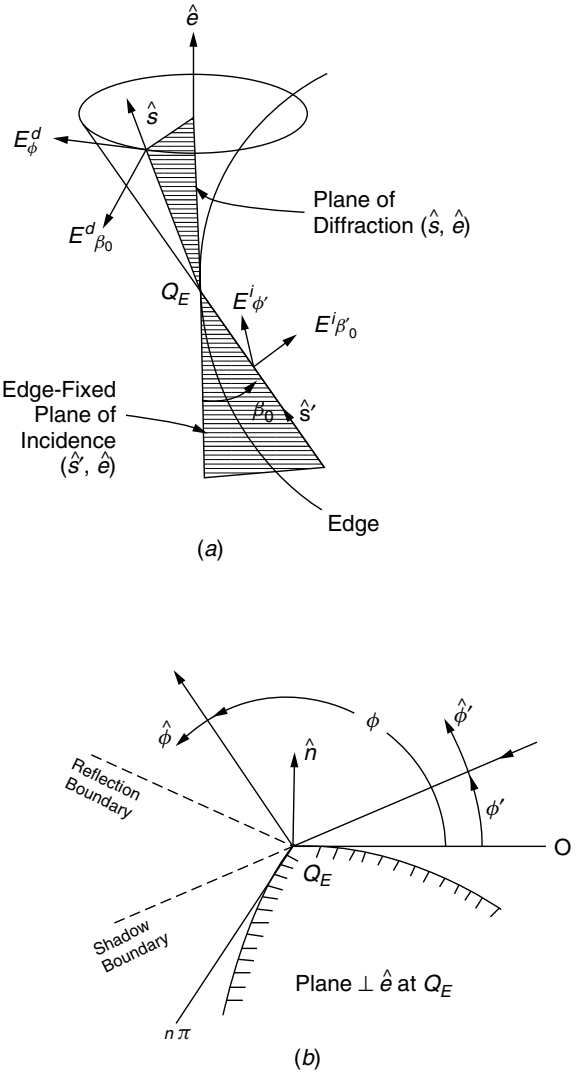


FIGURE 2-31 Ray-fixed coordinates related to edge-fixed coordinates at the edge diffraction point on a curved edge by showing planes of incidence and diffraction. (From [25], Fig. 5, © 1974 IEEE.)

where $\mathbf{S}'$ is the incident ray and β_0 is the angle between the edge tangent and the incident ray. The diffracted ray perpendicular polarization is similar to the incident ray

$$\mathbf{a}_\phi = -\frac{\mathbf{e} \times \mathbf{S}}{\sin \beta_0} \tag{2-81}$$

where $\mathbf{S}$ is the diffracted ray unit vector. We have the following vector relations for diffraction:

$$|\mathbf{e} \times \mathbf{S}| = |\mathbf{e} \times \mathbf{S}'| \quad \text{and} \quad \mathbf{e} \cdot \mathbf{S} = \mathbf{e} \cdot \mathbf{S}' \tag{2-82}$$

We determine the parallel polarization vector along the ray-fixed coordinates by the following cross products:

$$\mathbf{a}_{\phi'} \times \mathbf{a}_{\beta_0'} = \mathbf{S}' \quad \text{and} \quad \mathbf{a}_{\phi} \times \mathbf{a}_{\beta_0} = \mathbf{S} \tag{2-83}$$

By using ray-fixed coordinates, the diffraction matrix reduces to 2×2.

When $\beta_0 = \pi/2$, the parallel polarization components are parallel to the edge and the electric field vanishes: $E_{\beta_0'} + E_{\beta_0} = 0$. Acoustics calls this the *soft boundary condition* (Dirichlet); it operates on the parallel polarization components. The perpendicular components satisfy the *hard boundary condition* (Neumann). At a diffraction point Q_e we describe diffraction by the matrix equation

$$\begin{bmatrix} E^d_{\beta_0}(s) \\ E^d_{\phi}(s) \end{bmatrix} = \begin{bmatrix} -D_s & 0 \\ 0 & -D_h \end{bmatrix} \begin{bmatrix} E^i_{\beta_0'}(Q_e) \\ E^i_{\phi'}(Q_e) \end{bmatrix} \sqrt{\frac{\rho}{s(s+\rho)}} e^{-jks} \tag{2-84}$$

where s is the distance from the diffraction point. Diffraction locates one caustic on the diffraction point. We compute the second caustic distance ρ from the incident ray radius of curvature in the plane of incidence ρ^i_e and the edge curvature unit vector $\hat{\mathbf{n}}_e$:

$$\frac{1}{\rho} = \frac{1}{\rho^i_e} - \frac{\hat{\mathbf{n}}_e \cdot (\hat{\mathbf{s}}' - \hat{\mathbf{s}})}{a \sin^2 \beta_0} \tag{2-85}$$

where a is the edge radius of curvature. When $a \to \infty$ (straight edge), the second term of Eq. (2-85) vanishes.

A number of factors determine the wedge diffraction coefficients. The diffracting edge factors include (1) the angle between the faces, (2) the edge curvature, and (3) the curvature of the faces. The ray angle factors are (1) the incident angle relative to the edge tangent, (2) the diffraction angle to the shadow boundary, and (3) the angle to the reflection boundary. The diffraction coefficients peak at the shadow and reflection boundaries. UTD formulation uses characteristic lengths associated with incident and diffracted ray caustics. These many factors are beyond the current discussion.

2-7.12 Slope Diffraction

The spatial rate of change of the field normal to the edge produces slope diffraction, an added field component. This ray optics term also satisfies the generalized Fermat's principle with geometry determined by Eqs. (2-80) through (2-83), and (2-85). The slope diffraction equation has the same form as Eq. (2-84):

$$\begin{bmatrix} E^d_{\beta_0}(s) \\ E^d_{\phi}(s) \end{bmatrix} = \begin{bmatrix} -e_s & 0 \\ 0 & -e_h \end{bmatrix} \begin{bmatrix} E^i_{\beta_0'}(Q_e) \\ E^i_{\phi'}(Q_e) \end{bmatrix} \sqrt{\frac{\rho}{s(s+\rho)}} e^{-jks} \tag{2-86}$$

where the diffraction coefficients $e_{s,h}$ are related to the field derivative normal to the surface:

$$e_{s,h} = \frac{1}{jk \sin \beta_0} \left(\frac{\partial D_{s,h}}{\partial \phi'} \frac{\partial}{\partial n'} \right) \tag{2-87}$$

The term $\partial/\partial n'$ of Eq. (2-87) indicates the derivative of the incident fields given in the vector of Eq. (2-86). Equation (2-87) has the term $\partial D_{s,h}/d\phi'$ for the soft and hard slope diffraction terms returned from a subroutine; it is only a notational derivative.

2-7.13 Corner Diffraction

Every structural discontinuity diffracts waves. We derive edge diffraction from an infinite wedge where the wedge terminations (corners) produce diffracted rays. Recall from Section 2-4.2 that PTD added currents at edges to handle the effect of not having an infinite surface; the formulation for corner diffraction uses equivalent currents to derive these coefficients. We handle edge diffraction from each edge as always. Since each corner arises from two edges, we compute separate corner diffraction for each edge, two terms per corner.

Whereas edge diffraction is bound to a cone, corner diffraction radiates in all directions. The edge must be visible from both the source and receive points before corner diffraction contributes. We must include corner diffraction in any three-dimensional problem. As the source and receiver become farther and farther away from the object, corner diffraction contributions dominate over edge diffractions since it is derived from equivalent currents.

2-7.14 Equivalent Currents

GTD fails to predict fields at caustics. In many cases we consider these points unimportant, but for those cases where we need the fields, equivalent currents provide the answer. We derive equivalent currents from edge diffraction, which then replaces it and we use them instead of edge diffraction for all pattern points. The use of currents reduces the problem to a PO solution and line integrals are required.

We relate the incident fields expressed in the ray fixed to equivalent currents:

$$I = \frac{2j}{\eta k} E^i_{\beta'_0} D_s \sqrt{2\pi k}\ e^{j\pi/4} \tag{2-88}$$

$$M = \frac{2j}{k} E^i_{\phi'_0} D_h \sqrt{2\pi k}\ e^{j\pi/4} \tag{2-89}$$

The soft and hard diffraction coefficients $D_{s,h}$ depend on the source and receiver positions. Since we calculate the fields using vector potentials or dyadic Green's functions, the formulation has no caustics. They are only associated with a geometric optics solution.

Equivalent currents allow the calculation of the fields directly behind a reflector near the axis. The GTD solution produces a caustic as all points along the rim "light up" for an axisymmetrical design. PTD uses equivalent currents in a similar but different way to calculate correct fields in the same region. Equivalent currents derived from the diffraction coefficients produce the entire solution, since the reflector blocks the incident field. In PO we continue to include the direct field and the induced current radiation on the reflector, but add the PTD current radiation. Realize that slope diffraction also adds to the equivalent currents.

2-7.15 Diffraction from Curved Surfaces [26, 27]

In one analytical approach to surface-wave radiation we postulate waves bound to a surface that radiate only from discontinuities. Surface waves on infinite structure do not radiate but attenuate exponentially away from the surface, because they are bound to it. We can formulate GTD as radiating from discontinuities, and this produces an approach

for fields radiated on the shadowed side of a curved body. The continuous discontinuity of the curved surface causes power to be radiated at every point in the shadow region. These waves radiate tangentially from a wave traveling along a geodesic and bound to the surface. Surface waves require a dielectric coating or a corrugated surface to slow and bind the wave to the surface. The surface curvature slows and binds the wave to the surface without the need for a dielectric or corrugated surface coating. The wave that propagates along the surface sheds power in rays tangentially to it.

The rays travel along a surface geodesic from the attachment point to the radiation point. The geodesic curve is a minimum distance path on the surface between two points. In differential geometry it has a broader meaning, but for our purpose, the minimum distance definition will serve. The curved surface diffraction satisfies a generalized Fermat's principle (minimum distance) as do all other terms of GTD. The best approach uses another ray-fixed coordinate along the surface where the vectors are normal and tangential to the surface at both the attachment and radiation (shedding) points.

Curved surface diffraction considers three types of problems with different formulations. Two of them start with an antenna mounted on the surface. We either calculate the pattern in the presence of the curved object or calculate the coupling to a second antenna also mounted on the curved object. The third case determines the field scattered for a source located off the surface. All three use the ray-fixed coordinates. We start with the surface normal $\hat{\mathbf{n}}$ and the tangent vector $\hat{\mathbf{t}}$ directed along the geodesic path. A vector cross product defines the third direction of the local coordinate system. We use the surface binormal $\hat{\mathbf{b}}$, and the three vectors form a triad: $\hat{\mathbf{n}} \times \hat{\mathbf{b}} = \hat{\mathbf{t}}$. On a general surface all three vectors change direction as the wave moves along the geodesic. We use the term *torsion* for a path with a changing binormal. A soft dyadic diffraction coefficient is used with fields aligned with the attachment point binormal and the tangential shedding point binormal. We apply the hard dyadic diffraction coefficient fields aligned along the normal vectors. No formulas exist for computing the attachment and shedding points on a general curved surface given the source and receive points. We usually start with a known diffraction and find other points by incrementing along the curve by small steps.

REFERENCES

1. L. Diaz and T. A. Milligan, *Antenna Engineering Using Physical Optics*, Artech House, Boston, 1996.
2. B. F. Harrington, *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York, 1961.
3. P. S. Hacker and H. E. Schrank, Range distance requirements for measuring low and ultralow sidelobe antenna patterns, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 5, September 1982, pp. 956–966.
4. K. Pontoppidan, ed., *Technical Description of Grasp 8*, Ticra, Copenhagen, 2000 (self-published and available at *www.ticra.com*).
5. J. H. Richmond, A reaction theorem and its application to antenna impedance calculations, *IRE Transactions on Antennas and Propagation*, vol. AP-9, no. 6, November 1961, pp. 515–520.
6. R. F. Harrington, *Field Computation by Moment Methods*, Macmillan, New York, 1968; reprinted by IEEE Press, New York, 1993.

7. J. H. Richmond, Radiation and scattering by thin-wire structures in homogeneous conducting medium, *IEEE Transactions on Antennas and Propagation*, vol. AP-22, no. 2, March 1974, p. 365 (see also ASAP wire code).
8. N. N. Wang, J. H. Richmond, and M. C. Gilreath, Sinusoidal reactance formulation for radiation from conducting structures, *IEEE Transactions on Antennas and Propagation*, vol. AP-23, no. 3, May 1975, pp. 376–382.
9. B. M. Kolundzija and A. R. Djordjevic, *Electromagnetic Modeling of Composite Metallic and Dielectric Structures*, Artech House, Boston, 2002.
10. A. R. Djordjevic et al., *AWAS for Windows Version 2.0: Analysis of Wire Antennas and Scatterers*, Artech House, Boston, 2002.
11. A. C. Ludwig, Wire grid modeling of surface, *IEEE Transactions on Antennas and Propagation*, vol. AP-35, no. 9, September 1987, pp. 1045–1048.
12. A. W. Glisson and D. R. Wilton, Simple and efficient numerical methods for problems of electromagnetic radiation and scattering from surfaces, *IEEE Transactions on Antennas and Propagation,* vol. AP-28, no. 5, September 1980, pp. 563–603.
13. S. M. Rao, D. R. Wilton, and A. W. Glisson, Electromagnetic scattering by surfaces of arbitrary shape, *IEEE Transactions on Antennas and Propagation,* vol. AP-30, no. 3, May 1982, pp. 409–418.
14. B. M. Kolundzija et al., *WIPL-D: Electromagnetic Modeling of Composite Metallic and Dielectric Structures*, Artech House, Boston, 2001.
15. P. -S. Kildal, *Foundations of Antennas*, Studentlitteratur, Lund, Sweden, 2000.
16. C. Scott, *The Spectral Domain Method in Electromagnetics*, Artech House, Boston, 1989.
17. K. S. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Transactions on Antennas and Propagation*, vol. 14, no. 3, May 1966, pp. 302–307.
18. K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*, CRC Press, Boca Raton, FL, 1993.
19. J. G. Maloney and G. S. Smith, Modeling of antennas, Chapter 7 in A. Taflove, ed., *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Artech House, Boston, 1998.
20. D. J. Struik, *Differential Geometry*, Addison-Wesley, Reading, MA, 1950.
21. D. A. McNamara, C. W. I. Pistorius, and J. A. G. Malherbe, *Introduction to the Uniform Geometrical Theory of Diffraction*, Artech House, Boston, 1990.
22. F. S. Holt, in R. E. Collin and F. J. Zucker, eds., *Antenna Theory,* Part 2, McGraw-Hill, New York, 1969.
23. R. Kouyoumjian and P. Pathak, The dyadic diffraction coefficient for a curved edge, *NASA CR-2401*, June 1974.
24. J. B. Keller, Geometrical theory of diffraction, *Journal of the Optical Society of America*, vol. 52, 1962, pp. 116–130.
25. R. G. Kouyoumjian and P. H. Pathak, A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface, *Proceedings of IEEE,* vol. 62, no. 11, November 1974, pp. 1448–1461.
26. P. H. Pathak, W. D. Burnside, and R. J. Marhefka, A uniform GTD analysis of the diffraction of electromagnetic waves by a smooth convex surface, *IEEE Transactions on Antennas and Propagation*, vol. AP-28, no. 5, September 1980.
27. P. H. Pathak, N. Wang, W. D. Burnside, and R. G. Kouyoumjian, A uniform GTD solution for the radiation from sources on a convex surface, *IEEE Transactions on Antennas and Propagation*, vol. AP-29, no. 4, July 1981.

3

ARRAYS

We begin with arrays of antennas before discussing particular antenna elements to show the relationship between antenna size and shape and the resulting pattern characteristics. We ignore the feed network design initially and assume that the proper array feed distribution will be obtained. At first, we assume a distribution of point sources and compute the approximate array pattern. Working with simple models provides insight rather than accuracy, and later we consider element pattern and interaction. In Chapter 12 we discuss feed network design and analysis in the discussion of phased arrays.

The chapter begins with a mathematic description of an array and gives various assumptions used to simplify the expressions. We analyze a simple two-element array to gain insight into the radiation phenomenon and how far-field patterns can be found with simple arguments. The discussion of a uniformly spaced linear array shows the Fourier series relationship between array layout and the pattern space given in sin(angle) space. The principal idea is that pattern beamwidth shrinks as the array length increases. If we space the elements too far apart, multiple beam peaks or grating lobes form in the pattern, and we show how to control these grating lobes and their relationship to maximum scan angle, array layout, and element spacing.

Phased arrays scan the beam by controlling the relative phasing between the elements. We extend the linear array to planar layouts that produce narrow beams in both principal planes. The planar array design is unchanged from the methods for linear arrays, but the grating lobe analysis shows their unique properties, as they sometimes form outside the plane of scan. We can divide the phased array into pieces to form multiple scanning beams, but the beam shape is determined by the segment size and shape used for each beam. By adding amplitude control the phased array can form multiple beams with beamwidths determined by the entire size of the array.

Each element in an array receives a portion of the power radiated by the other elements on transmittal, or scatters power into neighboring elements in reception. The

Modern Antenna Design, Second Edition, By Thomas A. Milligan

radiation from each antenna excites currents on its neighboring elements that also radiate, and we associate the total pattern with the antenna input. In an array the effective element patterns change due to this scattering. Because of reciprocity, which says that transmit and received patterns are identical, we can analyze the problem either way. This leads to mutual coupling, which we describe and analyze by mutual impedance (admittance, or scattering) matrices. This phenomenon causes the input impedance of the elements to change as we scan the array. The mutual coupling can lead to scan blindness when the feed reflection coefficient grows due to mutual coupling, and the array totally reflects the signal into the feed network. If we want the exact pattern designed for, we must compensate the feeding coefficients for the mutual coupling.

A discussion of array gain gives two methods of calculation. First, the effective area and the associated gain of a planar array cannot exceed its area when we include the extra half-element spacing area provided by the edge elements. When we space the elements so that their individual effective areas no longer overlap, array gain is the element gain multiplied by the number of elements. We can calculate gain by adding up the input power instead of integrating the pattern to compute total radiated power. We relate input power of elements to the self- and mutual resistances to determine gain of linear and planar arrays using realistic elements. The chapter ends with a discussion of three-dimensional arrays using arbitrarily oriented elements. We add this analysis to the simple array formula to handle the polarization of rotated antennas. Related to this problem is the pointing of an antenna on a positioner. We apply rotation matrices to both problems.

An array radiates or receives from two or more antennas at the same frequency. To calculate the field radiated from arrays we add the electric fields radiated from each element. The amplitudes and phases of each antenna, determined by the feed network, give us extra degrees of freedom to shape the pattern and design shifts from radiating elements to the feed network.

A single antenna radiates an electric field with both polarization components:

$$\mathbf{E} = E_\theta(\theta, \phi)\hat{\theta} + E_\phi(\theta, \phi)\hat{\phi}$$

where E_θ and E_ϕ are the two complex components (amplitude and phase) referred to some point on the antenna. If we move the antenna or the phase reference point, we only change the antenna radiated phase. We assume that the movement is small enough that the radiation approximation can still be used. Given r' as the location of the antenna relative to the phase reference point, the added phase component is $e^{j\mathbf{k}\cdot\mathbf{r}'}$, where $\mathbf{k} = 2\pi/\lambda(\sin\theta\cos\phi\hat{\mathbf{x}} + \sin\theta\sin\phi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}})$ and $\mathbf{r}' = x'\hat{\mathbf{x}} + y'\hat{\mathbf{y}} + z'\hat{\mathbf{z}}$ is the location of the antenna; $\mathbf{k}\cdot\mathbf{r}'$ is the phase distance from the antenna to the reference plane through the reference point and is defined by the radiation (receiving) direction. The electric field radiated from the moved antenna becomes

$$[E_\theta(\theta, \phi)\hat{\theta} + E_\phi(\theta, \phi)\hat{\phi}]e^{j\mathbf{k}\cdot\mathbf{r}'}$$

We assume that nearby objects do not alter the patterns in the movement, but we can alter element patterns if necessary.

Suppose that we have an array of antennas located at points $\mathbf{r}'_1$, $\mathbf{r}'_2$, and so on. We obtain the total pattern by adding the electric fields radiated from each:

$$\mathbf{E} = \sum_{i=1}^{N}[E_{\theta i}(\theta, \phi)\hat{\theta} + E_{\phi i}(\theta, \phi)\hat{\phi}]e^{j\mathbf{k}\cdot\mathbf{r}'_i} \tag{3-1}$$

Bringing the antennas close together will change the patterns of each because every antenna will block the radiation of the others and the distribution of currents on the elements may be changed. The shape of small resonant antennas limits the possible distribution of currents, but the magnitude and phase may be changed due to the coupling.

We make various approximations to Eq. (3-1). Changes in the patterns due to nearby antennas are ignored, and isolated element patterns are used. We assume initially a certain amplitude and phase distribution on the elements and ignore the problem of the feed network. Polarization reduces to a single term for equally polarized elements, such as dipoles, slots, or horns. If the antennas have identical element patterns, we can separate Eq. (3-1) into a product.

$$\mathbf{E} = [E_\theta(\theta, \phi)\hat{\theta} + E_\phi(\theta, \phi)\hat{\phi}] \sum E_i e^{j\mathbf{k}\cdot\mathbf{r}'_i} \tag{3-2}$$

where E_θ and E_ϕ are the normalized patterns of the single element. E_i is the electric field of the ith element, including the amplitude and phase of the feed distribution.

Equation (3-2) describes pattern multiplication that separates the pattern into an element pattern and an array factor. The method requires that all antennas have the same pattern and be orientated in the same direction. The array factor represents the pattern from an array of isotropic pattern antennas. Because array factors can be calculated by hand, we find them useful for gaining insight. We leave calculations using Eqs. (3-1) and (3-2) to the computer. The element patterns themselves could be arrays and we could use pattern multiplication to synthesize planar and volumetric arrays from linear arrays.

3-1 TWO-ELEMENT ARRAY

Consider two elements lying on the z-axis and spaced a distance d centered on the origin (Figure 3-1). If we rotate the isotropic pattern antennas around the z-axis, the problem remains unchanged, which means that all great-circle (constant ϕ) patterns are identical. On the z-axis, the element phase constant becomes $e^{jkz'\cos\theta}$. For simple line arrays we can locate pattern nulls and peaks by simple arguments.

Example Two elements are spaced $\lambda/2$ and have equal amplitudes and phases. Locate the nulls and peaks.

The phase reference planes can be placed at any convenient point. Consider the pattern at $\theta = 90^\circ$. We place the reference plane through the axis of the array. The added phase factor is zero for both elements and we just add components. The equal element phases add to give a beam peak. If we place a second reference plane through the top element, the wave radiated from the bottom element travels across the array $\lambda/2$ to the reference plane. Increasing the distance propagated decreases phase and it changes by -180°. The two out-of-phase signals cancel to produce a pattern null. The array has symmetry about the $x-y$ plane, which means that the array will have the same pattern above and below the symmetry plane. We denote this configuration an even-mode array. Figure 3-2 plots this pattern with a solid line. You should repeat the example for an odd-mode array (phases 0° and 180°) and convince yourself that the null occurs at $\theta = 90^\circ$ and the beam peak occurs at $\theta = 0^\circ$ (180°), plotted in

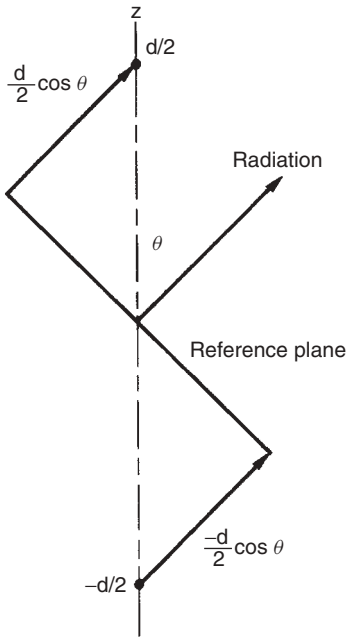


FIGURE 3-1 Two-element array on a z-axis.

Figure 3-2 as a short-dashed curve. The solid and short-dashed curves have the same directivity.

Example Suppose that the two elements are spaced $\lambda/4$, with the top element phase -90° and the bottom element phase 0°. Locate the beam peak and pattern null.

We start by placing a reference plane through the top element. The wave radiated from the bottom element travels across the array, and its phase decreases by 90°. Both radiated waves have the same phase (-90°) at the reference plane and add in phase for a beam peak. Consider a second plane through the bottom element. The wave from the top element loses 90° propagating across the array and the two waves are 180° out of phase and cancel for a null.

The second example is an end-fire array. Figure 3-2 illustrates the end-fire pattern with a long-dashed curve. All three patterns on the figure have the same directivity. The phase distribution of an end-fire array matches those of a wave traveling in the direction of the maximum. In these examples unequal amplitudes would limit the null depth to the difference. Varying the element phases while maintaining equal amplitudes changes the null directions.

Consider a general two-element array with equal amplitudes and a phase difference between them. We split the phase shift into equal parts. The top-element phase is $-\delta/2$ and the bottom-element phase is $\delta/2$. When we apply Eq. (3-2) with an isotropic element pattern, we obtain the following electric field using Euler's identity:

$$E(\theta) = 2E_0 \cos\left(\frac{\pi d}{\lambda}\cos\theta - \frac{\delta}{2}\right)\frac{e^{-jkr}}{r} \tag{3-3}$$

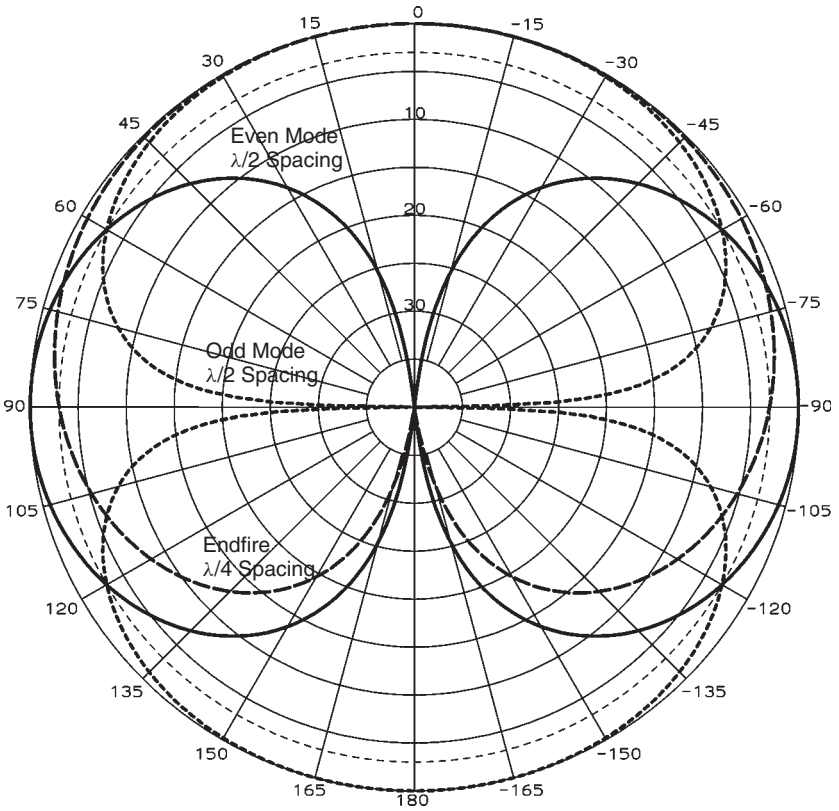


FIGURE 3-2 Two isotropic element array pattern: even-mode λ/2 spacing (solid curve); odd-mode λ/2 spacing (short-dashed curve); end-fire λ/4 spacing (long-dashed curve).

θ is measured from the z-axis. If we spaced the elements along the x-axis and found the pattern in the x–z plane, we substitute sin θ for cos θ in Eq. (3-3). In Chapter 4 we sample continuous distributions and position the elements along the x- or y-axis. Pattern peaks occur when the argument of the cosine is $n\pi$, the nulls when it is $(2n-1)\pi/2$.

$$\cos\theta_{\max} = \left(n\pi + \frac{\delta}{2}\right)\frac{\lambda}{\pi d} \tag{3-4}$$

$$\cos\theta_{\text{null}} = \left[(2n-1)\frac{\pi}{2} + \frac{\delta}{2}\right]\frac{\lambda}{\pi d} \tag{3-5}$$

If we subtract either Eq. (3-4) or (3-5) evaluated at two peaks or nulls, we get the same equation:

$$\cos\theta_1 - \cos\theta_2 = (n_1 - n_2)\frac{\lambda}{d} \tag{3-6}$$

Figure 3-3 illustrates the pattern of an equally phased two-isotropic-pattern-element array spaced 5λ along the z-axis. Because array symmetry makes the patterns on the right and left sides the same, we consider only one side. The wide element spacing

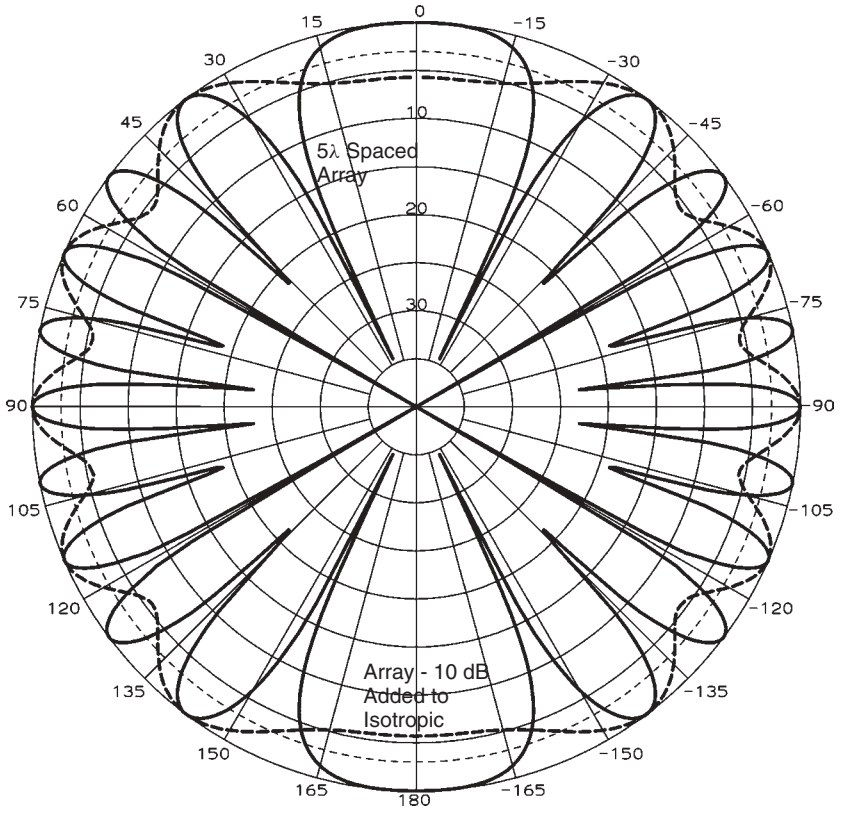


FIGURE 3-3 Two-isotropic-element array-spaced 5λ pattern (solid curve); added central element 10 dB higher power than array (dashed curve).

allows six solutions to Eq. (3-4) from 0 to 90° for the pattern peaks and five solutions for Eq. (3-5) over the same range for the nulls because the magnitude of cos θ is limited to 1. We call the multiple beams grating lobes. We usually choose the main beam and call the others grating lobes, but they are just all lobes of the array. Figure 3-3 shows that we must space the elements close together to prevent grating lobes. With a greater number of elements in the array, the amount of beam movement due to element phasing adds another factor to the prediction of when grating lobes form. The amount of phase scanning determines the maximum spacing allowed without the formation of grating lobes. The $n = 0$ lobe forms at $\theta = 90°$ and we compute the $n = 1$ mode direction from Eq. (3-4): $\theta = \cos^{-1}(\frac{1}{5}) = 78.46°$. When we substitute these angles into Eq. (3-3), we calculate a relative phase of 180° between them. The lobes have a phase of zero for n even and 180° phase for n odd in the far-field approximation. Remember we remove the exponential and $1/R$ factors from Eq. (3-3) for the far-field pattern. The actual phase of any real point depends on the distance from the center of the array.

The dashed curve in Figure 3-3 shows what happens if we add the array pattern to an isotropic radiator in the center. For a peak response of the array −10 dB relative to the isotropic antenna, we get the 5.7-dB peak-to-peak ripple shown by using Scale 1-8. The array pattern either adds or subtracts from the isotropic radiator pattern. The angular ripple rate is half that of the array lobes. Below we see that a two-element array

spaced at an integer multiple of λ/2 has a 3-dB greater gain than a single element. We feed half the power of the array into each element. By adding these factors we calculate the array element level to be −16 dB below the main central radiating antenna.

When we mount an antenna over a finite ground plane, the diffraction from the edges creates a two-element array. A 5λ-wide ground plane would produce the same pattern ripple angular rate as shown in Figure 3-3. You will often observe a similar-amplitude ripple in measured antenna patterns. Note the minimum angular distance between the peak and minimum responses in the pattern. The extra signals occur along the line in the pattern plane perpendicular to this direction. Use Eq. (3-6) to determine the distance between the array elements and you should be able to identify the structure causing the ripple. The scattering point could be on the test fixture. Consider whether the mounting structure will be different in the final configuration. You can calculate the effect from a single diffraction point by forming an array using the baseline of the primary radiator and the diffraction point. Both configurations produce the same angular ripple rate. The ripple peak occurs along that array axis, but Figure 3-3 shows that the angular ripple rate will be reduced along this end-fire direction of the $\theta = 0$ axis. If you make a careful consideration of the angular rates, in various pattern planes, you should be able to discover the cause. Always consider unexpected sources of diffraction.

You can consider the ripple using its beamwidth. To produce a symmetrical pattern about zero, we use $\sin\theta$ instead of $\cos\theta$ in Eq. (3-3), which means that the array lies along the x-axis. The −3-dB angle for the two-element uniform amplitude array can be found from Eq. (3-3):

$$\frac{\pi d}{\lambda}\sin\theta_{3\,\mathrm{dB}} = \frac{\pi}{4} \qquad \theta_{3\,\mathrm{dB}} = \sin^{-1}\frac{\lambda}{4d} \tag{3-7}$$

The beamwidth is twice Eq. (3-7). For large d we can approximate $\sin X \approx X$ and beamwidth $= \lambda/2d$. The 5λ spaced array has a beamwidth of 5.7° (0.1 rad). We can look at a 5λ-wavelength ground-plane example that has a large-amplitude element compared to the edge diffraction as two 2.5λ-spaced two-element arrays where one element has a high amplitude. Each two-element array produces a pattern with an 11.4° beamwidth the value of the composite pattern in Figure 3-3. We often mount an antenna in the center of a ground plane for measurement and observe patterns similar to Figure 3-3. If in the actual application the antenna is mounted off center, we need to add the patterns of arrays formed on both sides of the finite ground plane. The final pattern will be the composite pattern from each array and be more complicated than the simple case given above.

We calculate average radiation intensity by an integral:

$$U_{\mathrm{avg}} = \frac{4E_0^2}{\eta}\int_0^{\pi/2}\cos^2\left(\frac{\pi d}{\lambda}\cos\theta - \frac{\delta}{2}\right)\sin\theta\, d\theta$$

The directivity is

$$\frac{U_{\max}}{U_{\mathrm{avg}}} = \frac{|2E_{\max}|^2}{1 + \sin(2\pi d/\lambda)\cos\delta/(2\pi d/\lambda)} \tag{3-8}$$

where $E_{\max} = \cos[(\pi d/\lambda)\cos\theta_{\max} - \delta/2]$. When $d \geq \lambda/2$, $E_{\max} = 1$. Figure 3-4 shows the directivity versus spacing for the special cases $\delta = 0°$ and $\delta = 180°$ (even and odd modes). The directivity varies because each antenna receives power from the other. The

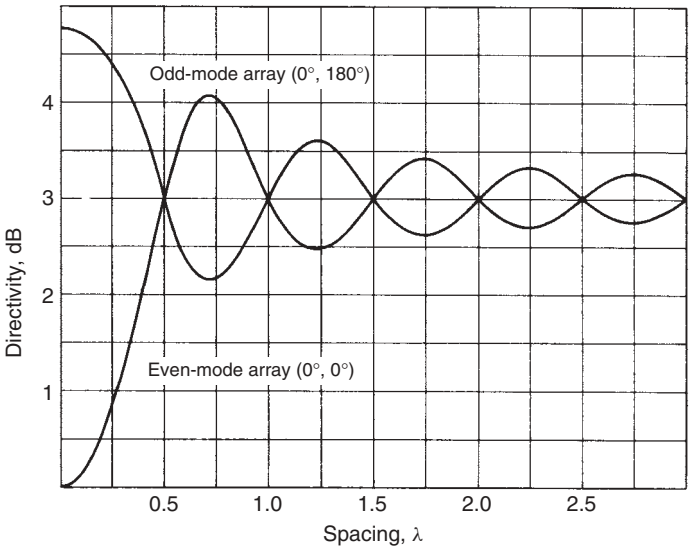


FIGURE 3-4 Directivity of even- and odd-mode two-isotropic-element arrays.

combination of the input power and the power transferred between elements changes with spacing.

3-2 LINEAR ARRAY OF N ELEMENTS

Suppose that there are N isotropic radiators equally spaced along the z-axis and fed with equal amplitudes. We assign a fixed phase shift δ between progressive elements. The array factor field is

$$\frac{\sin(N\psi/2)}{N\sin(\psi/2)} \tag{3-9}$$

where $\psi = kd\cos\theta + \delta$ [1, p. 258]. We use this to plot a universal radiation pattern for the array (Figure 3-5) for two to 10 elements. The abscissa ψ is plotted in degrees (360° is substituted for 2π in k). Both ends of the plot are lines of symmetry. The plot is periodic (period 360°). We see that the level of the first sidelobe ($N = 2$ has no sidelobe) decreases as N decreases but approaches a limit of 13.3 dB of the continuous aperture.

Figure 3-6 demonstrates the periodic pattern for $N = 6$ and shows a projection to a polar pattern when the progressive phase between elements is zero and the elements are spaced $\lambda/2$. We can plot similar curves for other array distributions; all have a period of 360°. Figure 3-6 illustrates the use of a circle diagram, a method of constructing a polar pattern from the universal pattern such as Eq. (3-9) for the uniform-amplitude distribution. An array can be analyzed as a sampling of the continuous distribution that produces a Fourier series of the distribution. A Fourier series has multiple responses. In Chapter 4 we design large arrays by sampling continuous distributions. The pattern angle of an array is measured either from the axis using cosine of pattern angle or is measured broadside using sine. You should become comfortable with either notation since the sine and cosine of angles involves only a complementary operation of the angles.

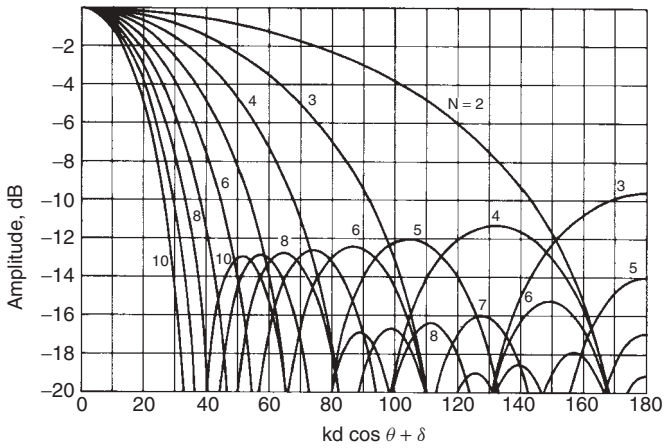


FIGURE 3-5 ψ-space pattern of linear arrays with a uniform amplitude distribution.

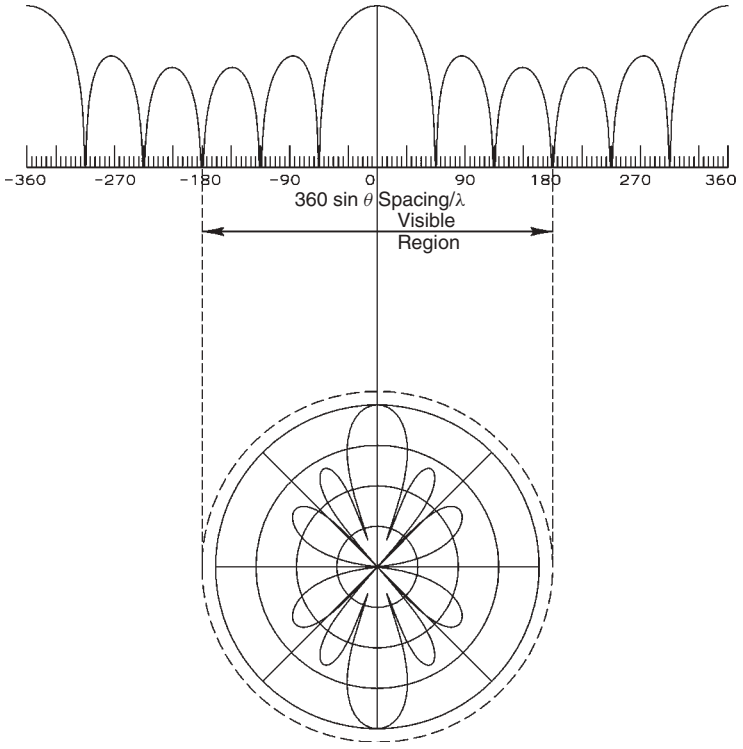


FIGURE 3-6 Circle diagram of a six-element uniform-amplitude array with λ/2 spacing.

Since $\cos\theta$ (or $\sin\theta$) is limited to ± 1, the region along the abscissa of the universal pattern used (the visible region) is found from the range of ψ:

$$\frac{-360^\circ d}{\lambda} + \delta \quad \text{to} \quad \frac{360^\circ d}{\lambda} + \delta$$

The circle diagram is constructed by first drawing a circle the same diameter as the visible region below the universal diagram centered at δ, the progressive phase shift between elements. Figure 3-6 has a $\delta = 0$. Since the element spacing is $\lambda/2$, the range is $\pm 180^\circ$. The polar pattern radius equals the amplitude of the universal pattern. Both the universal pattern and the polar pattern use a logarithmic (dB) scale from 0 to -40 dB. Projecting points vertically from the universal pattern to the visible region performs the cosine or sine operation, and the polar pattern becomes the real pattern in space. We project each point vertically until it intersects the dashed visible region circle in two places and then draw lines from these points to the center. After you project the nulls and peaks of the universal pattern to the dashed circle, it is easy to sketch the polar pattern. The circle diagram helps us visualize patterns and the effects of scanning, but no one would do serious design with it. Second, it is useful only for small arrays because large arrays produce unwieldy diagrams.

When the spacing between elements is greater than $\lambda/2$, the visible region widens to include more than one periodic main lobe and the array has multiple beams. To have a beam centered at θ_1, set the progressive phase difference between elements:

$$\delta = \frac{-360^\circ d}{\lambda} \cos \theta_1 \tag{3-10}$$

End fire ($\theta_1 = 0$) occurs when

$$\delta = \frac{-360^\circ d}{\lambda} \tag{3-11}$$

We can use Figure 3-5 to compute beamwidth angles of arrays. Table 3-1 is a list of the ψ-space angles of the 3- and 10-dB levels.

Example A six-element equally spaced uniform array has spacings of $\lambda/2$ and zero progressive phase shift between elements ($\delta = 0^\circ$). Calculate the 3-dB beamwidth.

We read from Table 3-1 the value $\psi_{3\,\text{dB}} = 26.90^\circ$. Because the pattern is symmetrical in ψ space (Figure 3-6), the second $\psi_{3\,\text{dB}}$ is -26.9°.

$$kd \cos \theta_{1,2} + \delta = \pm \psi_{3\,\text{dB}}$$

$$\frac{360^\circ}{\lambda} \frac{\lambda}{2} \cos \theta_{1,2} = \pm 26.90^\circ \qquad \cos \theta_{1,2} = \frac{\pm 26.9^\circ}{180^\circ}$$

$$\theta_1 = 81.4^\circ \qquad \theta_2 = 98.6^\circ$$

TABLE 3-1 ψ-Space Angles of 3- and 10-dB Levels of an Equal-Amplitude Distribution Array (deg)

N	3 dB	10 dB	N	3 dB	10 dB	N	3 dB	10 dB
2	90.00	143.13	11	14.55	24.21	20	7.980	13.29
3	55.90	91.47	12	13.33	22.18	24	6.649	11.08
4	40.98	67.63	13	12.30	20.47	28	5.698	9.492
5	32.46	53.75	14	11.42	19.00	32	4.985	8.305
6	26.90	44.63	15	10.65	17.74	36	4.431	7.382
7	22.98	38.18	16	9.98	16.62	40	3.988	6.643
8	20.07	33.36	17	9.39	15.64	50	3.190	5.314
9	17.81	29.62	18	8.87	14.77	64	2.492	4.152
10	16.02	26.64	19	8.40	14.00	100	1.595	2.657

Remember that θ is measured from the axis of the array (z-axis) and the 3-dB beamwidth is the difference (17.2°). On Figure 3-6 the visible region ranges between $-180°$ and $+180°$. There are four sidelobes in the visible region (Figure 3-6). Since an array samples a continuous aperture distribution, the continuous distribution is Nd long. We can estimate beamwidth by using a uniform amplitude distribution:

$$\text{HPBW} = 50.76° \frac{\lambda}{Nd} = 16.92°$$

This formula approximates the array beamwidth reasonably.

Example A six-element array has a progressive phase shift δ of 90° between elements. Compute the 10-dB beam edge angles for $\lambda/2$ spacing.

Figure 3-7 shows the circle diagram analysis of this example. The line to the center of the polar pattern has been shifted to 90° and the pattern spans 360° of the linear scale. By projecting the nulls and peaks to the circle below, the pattern can easily be sketched.

$$\psi_{10\,\text{dB}} = \pm 44.63° \quad \text{(Table 3-1)}$$
$$kd \cos\theta_{1,2} = \pm\psi_{10\,\text{dB}} - \delta$$

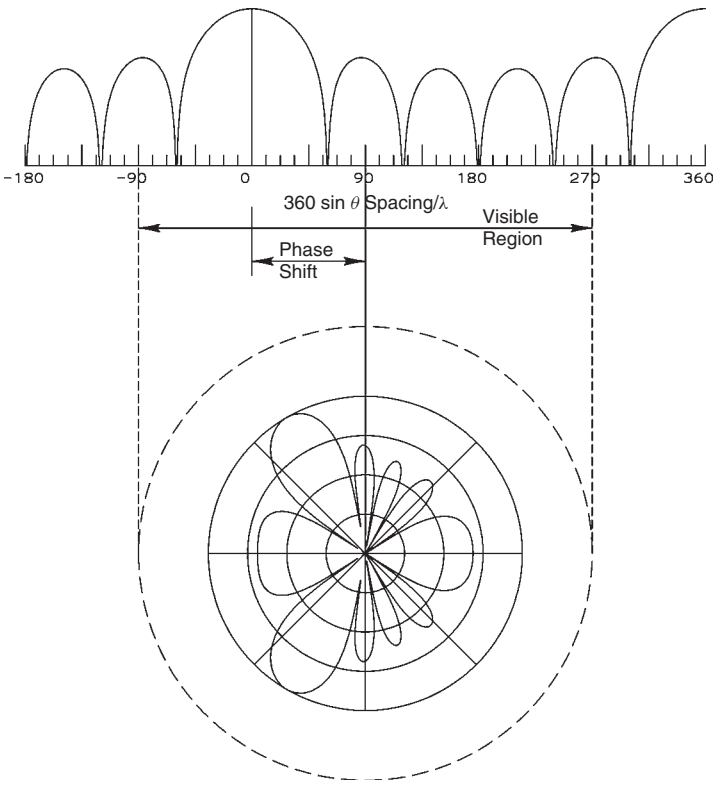


FIGURE 3-7 Six-element uniform-amplitude array with $\lambda/2$ spacing scanned with 90° progressive phase shift between elements.

Solving for $\cos\theta_{1,2}$, we have

$$\cos\theta_{1,2} = \frac{\pm\psi_{10\,\text{dB}} - \delta}{kd} = \frac{\pm 44.63^\circ - 90^\circ}{180^\circ}$$

$$\theta_1 = 104.6^\circ \qquad \theta_2 = 138.4^\circ \qquad \text{beamwidth} = 33.8^\circ$$

There are five sidelobes in the visible region (Figure 3-7). Equation (3-9) gives the beam maximum direction:

$$\cos\theta_0 = \frac{-\delta}{kd} = \frac{-90^\circ}{180^\circ} \qquad \theta_0 = 120^\circ$$

The main beam is no longer symmetrical about the beam peak. The 3-dB pattern angles are 110.5° and 130.5°. The beamwidth (3-dB beamwidth = 20°) increases with scan angle. What element spacing would result in this beamwidth for broadside radiation ($\delta = 0^\circ$)?

$$\frac{360^\circ}{\lambda} d\cos\theta_1 = 26.90^\circ \quad \text{(Table 3-1)}$$

On solving for spacing, we have

$$\frac{d}{\lambda} = \frac{26.9^\circ}{360^\circ\cos\theta_1}$$

Remember that the beam is centered on $\theta = 90^\circ$, so that $\theta_1 = 90 - 20/2 = 80^\circ$.

$$\frac{d}{\lambda} = \frac{26.9^\circ}{360^\circ\cos 80^\circ} = 0.431$$

The effective spacing has been reduced by approximately the cosine of the scan angle from $\theta = 90^\circ$, broadside:

$$\frac{d}{\lambda}\cos 30^\circ = 0.433$$

The accuracy of the cosine relation increases with more elements.

Example Determine the progressive phase shift between elements for an end-fire array with 0.3λ element spacing and compute beamwidth for a uniform distribution array with five elements.

Figure 3-8 illustrates this example using the circle diagram. End fire occurs when [Eq. (3-11)]

$$\delta = \frac{-360^\circ(0.3\lambda)}{\lambda} = -108^\circ$$

This is the progressive phase shift for all distributions with 0.3λ element spacing for an end-fire pattern. Table 3-1 gives the ψ-space angle, $\psi_{3\,\text{dB}} = \pm 32.46^\circ$. Substituting in the expression for ψ, we have

$$\frac{360^\circ(0.3\lambda)}{\lambda}\cos\theta_{1,2} = \pm 32.46^\circ$$

$$\cos\theta_{1,2} = \frac{\pm 32.46^\circ + 108^\circ}{360^\circ(0.3)}$$

$$\cos\theta_1 = \frac{140.46}{108} = 1.301 \qquad \cos\theta_2 = \frac{75.46}{108} = 0.699$$

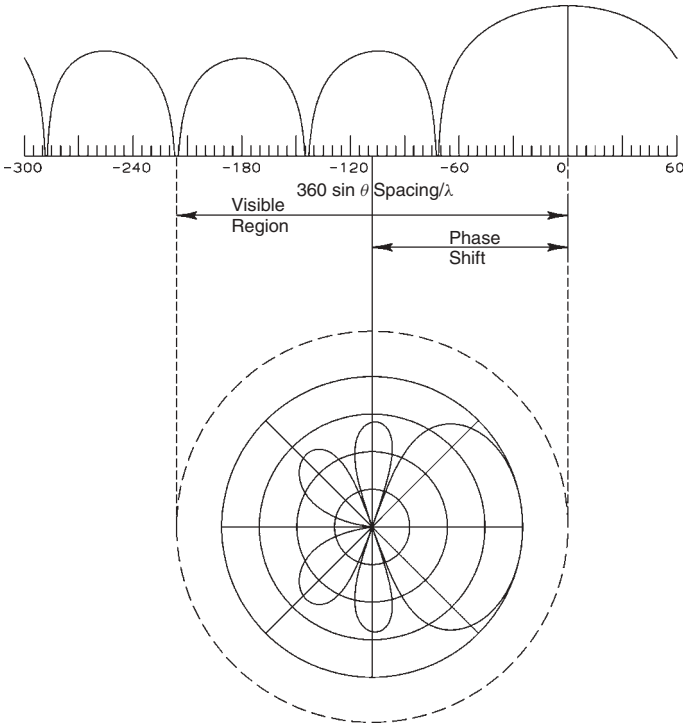


FIGURE 3-8 Five-element uniform-amplitude array scanned to end fire.

θ_1 is in invisible space, since $|\cos\theta| \leq 1$; $\theta_2 = 45.6^\circ$. Symmetry about the z-axis supplies us with the second angle $\theta_1 = -45.6^\circ$ and beamwidth is the difference: 91.2°. The end-fire array samples a traveling-wave distribution. The continuous uniform distribution phased for end fire with the same length has a 90.4° beamwidth.

Remember that we have been dealing with isotropic pattern antennas. For example, broadcast towers, seen from above, approximate isotropic antennas in the horizontal plane. The patterns of the individual antennas modify the results of isotropic antenna arrays. In small arrays the element pattern is quite significant, but the beamwidths of large arrays are determined mainly by the array factor. The beamwidths calculated for array factors approximate the actual beamwidths only when the elements have significant patterns. We must rely on computer solutions of specific cases, including the element pattern, for better results.

3-3 HANSEN AND WOODYARD END-FIRE ARRAY [2]

The end-fire array directivity increases if the sum of the progressive phase shifts between elements is decreased by approximately π. The equivalent traveling-wave velocity slows in the structure relative to free space. The progressive phase shift between elements becomes

$$\delta = -kd - \frac{2.94}{N} \simeq -kd - \frac{\pi}{N} \qquad \text{rad} \tag{3-12}$$

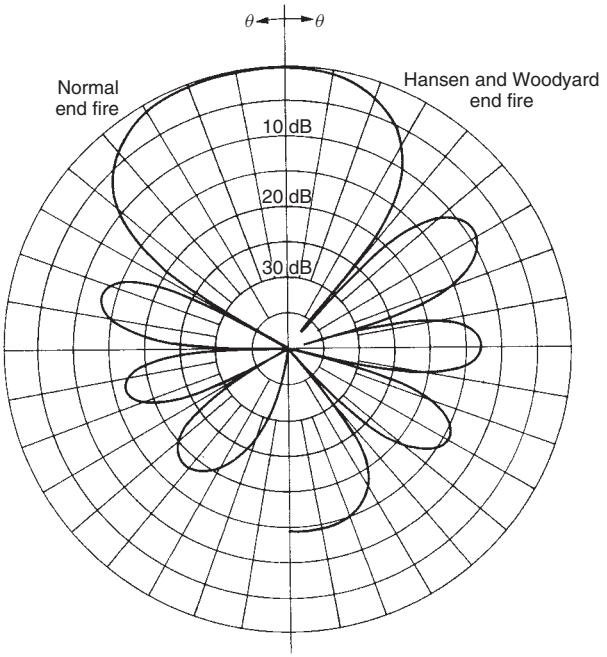


FIGURE 3-9 Patterns of a normal end fire and a Hansen and Woodyard end-fire array of isotropic elements.

where N is the number of elements in the array. The normal end-fire progressive phase shift between elements, $\delta = -kd$, places one edge of the visible region at the origin of ψ space. This method shifts the edge to a lower portion of the curve. The universal radiation curve peak (Figure 3-5) shifts into invisible space and the sidelobes rise in proportion to the new beam peak, but the beamwidth narrows. Equation (3-12) holds strictly only for large arrays, but the directivity increases for all arrays when it is applied.

Example Suppose that eight elements are spaced $\lambda/4$ apart with a uniform amplitude distribution. Compare the two endfire designs. The two patterns are compared in Figure 3-9.

The results are as follows: The beamwidth decreases, and the directivity increases by 2.5 dB. The sidelobes rise to 9 dB from 13 dB.

3-4 PHASED ARRAYS

Suppose that a wave approaches at an angle to the axis of an array located on the z axis (Figure 3-10). The wave reaches the top element first and progresses down the array in succession. If the signals are added directly, they will cancel each other to some extent because they have a progression of phases. Figure 3-10 shows the results of adding a series of time delays to equalize the path lengths in the lines where the position z_i along the axis determines the time delay τ_i for incident angle θ_0:

$$\tau_i = \frac{z_i}{c} \cos \theta_0 + \tau_0$$

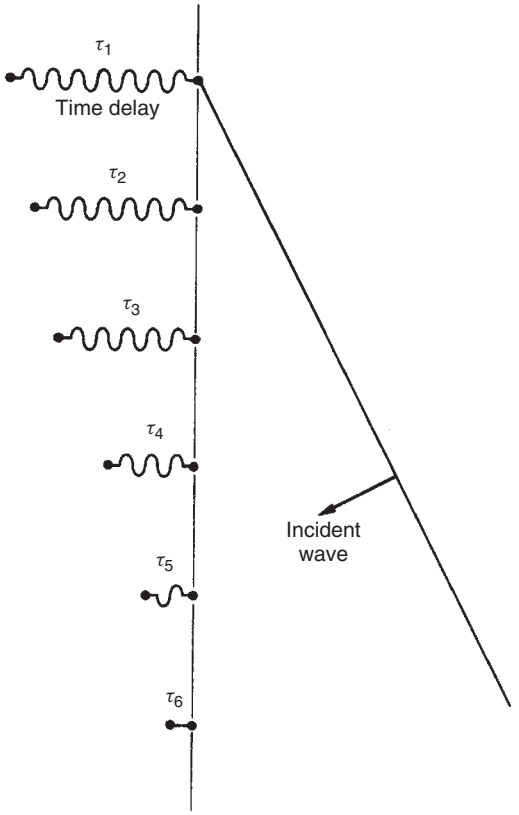


FIGURE 3-10 Linear array scanned with time-delay networks.

and velocity of light c. We add an arbitrary time delay τ_0 to keep all time delays, τ_i, positive. This feed network is frequency independent, as we vary the progression of time delays to scan the beam.

Phase shifters replace the time-delay networks in phased arrays. They provide equivalent beam scanning at a single frequency. To scan to an angle θ_0, the required phase shift is

$$-\frac{2\pi}{\lambda} z \cos\theta_0 \qquad \text{modulo } 2\pi \qquad \text{(rad)}$$

$$\frac{-360^\circ}{\lambda} z \cos\theta_0 \quad \text{modulo } 360^\circ \qquad \text{(deg)}$$

for elements located on the z-axis. For a general space array we must counteract the phase difference to the reference plane, $e^{j\mathbf{k}\cdot\mathbf{r}'}$, for the direction of scan so that the phases of all elements are zero. To scan in the direction (θ_0, ϕ_0), we must add a phase factor to every element, depending on its position. The phase factor on each element of a general space array is

$$e^{-j\mathbf{k}_0 \cdot \mathbf{r}'} \tag{3-13}$$

where

$$\mathbf{k}_0 = \frac{2\pi}{\lambda}(\sin\theta_0 \cos\phi_0 \hat{\mathbf{x}} + \sin\theta_0 \sin\phi_0 \hat{\mathbf{y}} + \cos\theta_0 \hat{\mathbf{z}})$$

is the vector propagation constant in the direction of the beam and **r** is the element location. Adding this phase factor to the element phases causes the product of the exponential factors [Eq. (3-2)] to be 1 at the scan angle, and the components E_i add in the scan direction.

Using phase shifters limits the frequency bandwidth. Given a fixed phase shift over a small frequency range, increasing the frequency scans the beam toward broadside:

$$\Delta\theta = \frac{f_2 - f_1}{f_2} \tan\left(\frac{\pi}{2} - \theta_0\right) \qquad \text{rad} \tag{3-14}$$

where θ_0 is the scan angle [3]. Limiting the allowable scanning with frequency to plus or minus one-fourth of the local beamwidth defines the bandwidth of the array. When the beam is scanned to 30° off the axis, the bandwidth is related directly to the beamwidth at broadside ($\theta = 90°$):

$$\text{bandwidth}(\%) \simeq \text{beamwidth (deg) at } \theta_0 = 30^\circ$$

The beam shifts less with frequency near broadside, since the tangent factor in Eq. (3-14) approaches zero. A general estimate is given by

$$\text{bandwidth}(\%) \simeq \frac{\text{beamwidth (deg)}}{2\cos\theta_0} \tag{3-15}$$

where the broadside beamwidth is used.

Example Given an array with 100 elements spaced at λ/2, determine the bandwidth when scanned to 45°.

The beamwidth is estimated from the aperture width:

$$\text{HPBW} = \frac{50.76^\circ}{100(\frac{1}{2})} \simeq 1^\circ$$

$$\text{bandwidth}(\%) \simeq \frac{1}{2\cos 45^\circ} = 0.7\%$$

Any radar antenna would have a broader beamwidth because the sidelobes need to be reduced, but this is a good first estimate.

The bandwidth can be increased by feeding subarrays with time-delay networks. The subarrays continue to be scanned with phase shifters. Only a few time-delay networks are needed, and the subarray beamwidth determines the bandwidth. In Chapter 12 we discuss the problems caused by using subarrays.

3-5 GRATING LOBES

Phased arrays vary the progressive phase by Eq. (3-13) to scan the beam. When the array element spacing is greater than λ/2, the appearance of secondary beam peaks (grating lobes) limits the scan angle. The grating lobe attains full amplitude when

$$\frac{d}{\lambda}(1 + \cos\theta_{\text{gr}}) = 1 \qquad \theta_{\text{gr}} = \cos^{-1}\left(\frac{\lambda}{d} - 1\right) \tag{3-16}$$

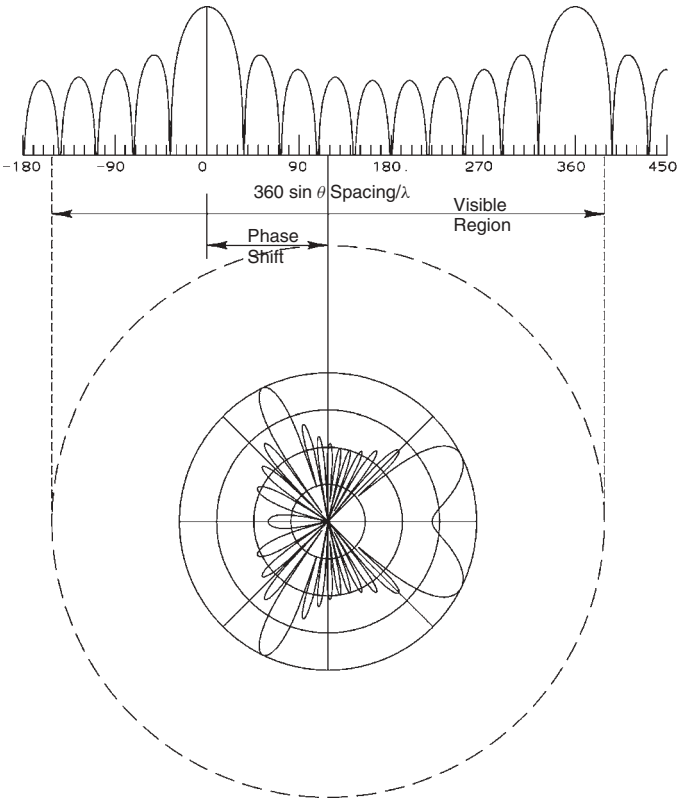


FIGURE 3-11 Ten-element array with 3λ/4 spacing scanned to 26°, showing the onset of a grating lobe.

Example The spacing of the elements of an array is 0.75λ. Determine the scan angle when the grating lobe is full amplitude.

$$\theta_{gr} = \cos^{-1}\left(\frac{4}{3} - 1\right) = 70.5^\circ$$

At this point the grating lobe is the same amplitude as the main beam. The lobe does not appear suddenly, but it grows as the visible region shifts and starts including the second periodic main lobe. Figure 3-11 shows the grating lobe formation for an array with 0.75λ element spacings on a circle diagram. The dashed circle of the visible region spans more than one beam of the universal radiation pattern of the uniform amplitude array.

Arrays with element spacing greater than λ always have grating lobes (multiple main beams), but the pattern of the antenna elements may reduce the grating lobes to acceptable levels and allow a wide element spacing.

3-6 MULTIPLE BEAMS

An array can form multiple beams. Equation (3-13) gives the phase coefficients to multiply each element feed voltage E_i to scan it to a given angle. The array will form

a second beam if we add a second distribution: $E_i e^{-j\mathbf{k}_2\cdot\mathbf{r}'_i}$. The distribution E_i remains constant for both beams. We add the two distributions to obtain both beams:

$$E_i(e^{-j\mathbf{k}_1\cdot\mathbf{r}'_i} + e^{-j\mathbf{k}_2\cdot\mathbf{r}'_i}) \tag{3-17}$$

This multiplies the distribution E_i by a second distribution whose amplitudes and phases are functions of the antenna position and the scan angles of the two beams. Each beam uses the entire array to form its beam. In a phased array both phase and amplitude must be varied to achieve multiple beams. An array, which can only vary phase, must be divided into subarrays to form multiple beams, but its beamwidths will depend on the subarray widths.

We can produce unequal beams with different amplitude distributions and pattern shapes if needed. We can add as many beams as necessary by including the distribution element factors with the scanning phase coefficients in Eq. (3-17). The element feeding coefficients become the sum.

Example Compute the feed coefficients of a 15-element array with $\lambda/2$ spacings and a uniform distribution scanned to 45° and 120° from the z-axis.

First center the array on the z-axis. The elements are located at

$$z_i = \frac{(-8+i)\lambda}{2}$$

To scan to 45°, the element phase factors are

$$\exp(-jkz_i\cos 45^\circ) = \exp\left[-j\frac{360^\circ}{\lambda}\frac{1}{\sqrt{2}}\frac{(-8+i)\lambda}{2}\right]$$

To scan to 120°, the element phase factors are $e^{j90^\circ(-8+i)}$. The ninth-element ($z_9 = \lambda/2$) phase factors are $e^{-j127.3^\circ}$ and e^{j90°. Assume a voltage magnitude of one-half for each uniform-amplitude-distribution beam so that the center element has a magnitude of 1. We sum the distributions to calculate the feeding coefficient of the ninth element:$0.32e^{j161.4^\circ}$. When converted to decibel ratios, Table 3-2 lists the feeding coefficients for the array. We can estimate both beamwidths from Table 3-1. $\psi_{3\,\mathrm{dB}} = 10.65^\circ$:

$$\cos\theta_{1,2} = \begin{cases} \dfrac{\pm 10.65^\circ + 127.28^\circ}{180^\circ} & \theta_1 = 40^\circ, \theta_2 = 49.6^\circ \\ \dfrac{\pm 10.65^\circ - 90^\circ}{180^\circ} & \theta_1 = 116.2^\circ, \theta_2 = 124^\circ \end{cases}$$

The pattern (Figure 3-12) has these beams.

The gain of each beam depends on the feed network. If a single input supplies power to two beams, each beam can receive only half the input power and gain reduced 3 dB for both beams. Butler matrices [4] and Blass beamforming networks [5] supply an input for each beam. The inputs are isolated from each other and the transmitter power in each port feeds only one beam, and therefore the full array gain is available to each input. Similarly, we can place a receiver on each port and use the full effective area for each.

TABLE 3-2 Feeding Coefficients for a Dual-Beam 15-Element Array

Element	Beamwidth (dB)	Angle (deg)	Element	Beamwidth (dB)	Angle (deg)
1	−2.38	130.48	9	−9.91	161.36
2	−8.59	111.84	10	−1.99	142.72
3	−0.01	−86.80	11	−1.64	−55.92
4	−11.49	74.56	12	−11.49	−74.56
5	−1.64	55.92	13	−0.01	86.80
6	−1.99	−142.72	14	−8.59	−111.84
7	−9.91	161.36	15	−2.38	−130.48
8	0.00	0.0			

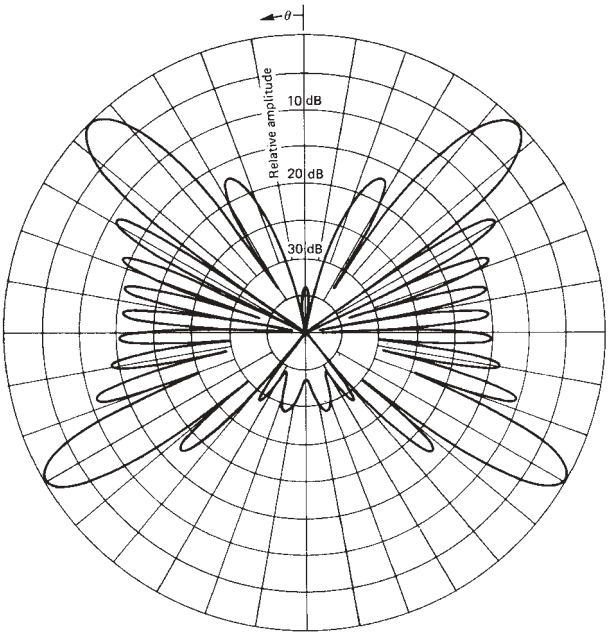


FIGURE 3-12 Fifteen-element linear array pattern with simultaneous beams at $\theta = 45^\circ$ and 120°.

We will delay the important topics of array synthesis and sidelobe reduction until after we have discussed aperture distributions. A trade-off is made between the beamwidth and the sidelobe levels. The beamwidth narrows only by putting more power into the sidelobes.

3-7 PLANAR ARRAY

The linear array only controls the pattern in one plane; it depends on the element pattern to control the beam in the other plane. Planar arrays can control the beam shape in both planes and form pencil beams. Whereas a linear array can only scan in

a single plane, a planar array can scan to any angle in the upper hemisphere. Most planar arrays rely on the element pattern or ground plane to eliminate the backlobe on the opposite side of the plane. The planar array has $N - 1$ nulls that can be used to control the pattern where N is the total number of elements.

A simple feed distribution uses the product of two linear arrays. This eliminates many degrees of freedom of the array because an $M \times N$ array would be determined by $M - 1 + N - 1$ nulls when we could have used $M \times N - 1$ possible nulls. Figure 3-13 shows the spherical pattern of a uniformly spaced 8×8 planar array where all elements are fed the same amplitude where a 90° beamwidth element eliminates the backlobe. The pattern along either principal axis shows the steady sidelobe reduction that starts with -13.2 dB. Diagonal plane sidelobes are the product of the sidelobes in the principal planes. The first sidelobe in the diagonal plane is down 26.4 dB. An array feed distribution, not a product of two linear arrays, can yield more equal sidelobes in all pattern planes.

Figure 3-14 illustrates the pattern of the rectangular array when the element feeding coefficients are phased to scan the beam along one principal plane. The main beam broadens in the plane of scan as the effective array length is reduced but stays narrow in the orthogonal plane. More sidelobes appear behind the main beam. We see a large sidelobe growing on the horizon that will become a grating lobe when the array is scanned further. The sidelobes in the plane orthogonal to the plane of scan move with the main beam but roll into a cone that becomes tighter with increased scan.

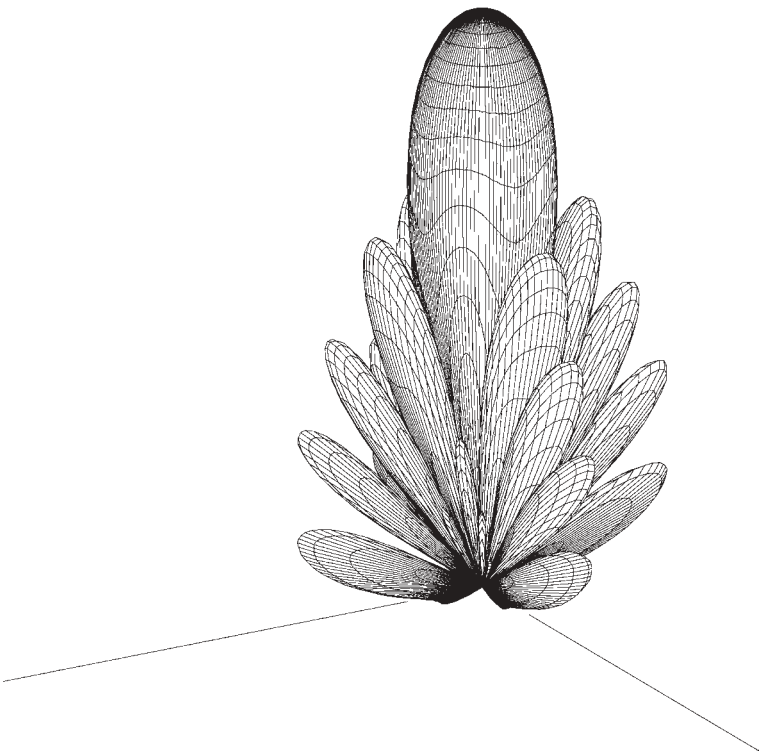


FIGURE 3-13 Spherical radiation pattern of an 8×8-element uniform-amplitude and spaced square planar array.

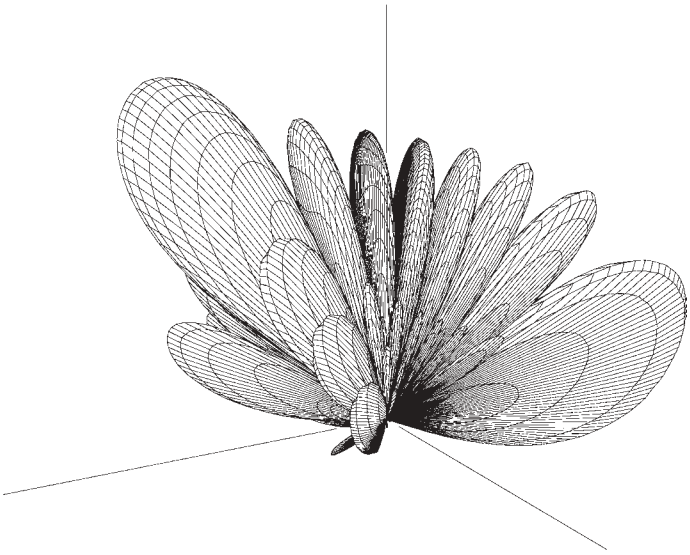


FIGURE 3-14 Spherical radiation pattern of an 8 × 8-element square-planar array scanned along a principal plane.

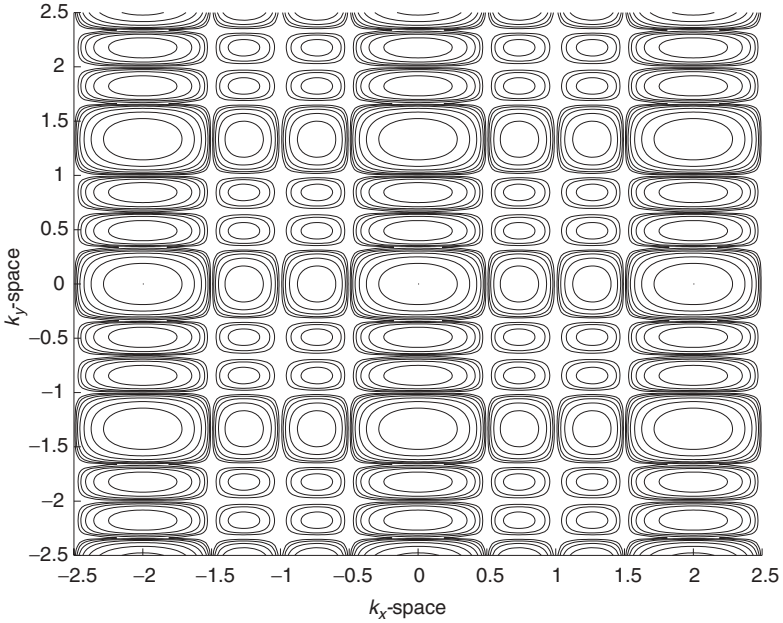


FIGURE 3-15 Contour plot of the pattern of a 4 × 4-element square array in $k_x k_y$-space showing multiple beams and sidelobes.

Figure 3-15 shows a contour plot of the universal pattern of a 4 × 4 element rectangular array. We denote this universal pattern $k_x k_y$ space because the principal axes have sin θ factors similar to the universal pattern of a linear array. The array for Figure 3-15 has its y-axis element spacing 1.5 times wider than the x-axis spacing. The diagram axes extend until multiple beams show on the figure. The main beams correspond

to the center of the large "squares." The visible region on the figure is a unit circle with its center at the negative scan direction $(-k_{x0}, -k_{y0})$. This technique mirrors the circle diagram of the linear array where the visible region is given by a linear region centered at the negative scan direction. You should notice that the diagonal sidelobes have smaller amplitudes than the principal plane sidelobes. We move the unit circle as the array scans and the diagram shows those locations of scan that have multiple beams (grating lobes). A grating analysis simplifies the diagram of Figure 3-15 to the main beam locations.

When we place the two axes of the planar array at an angle instead of orthogonal, we form a triangular array. Figure 3-16 gives the positions of a hexagon array made with equilateral triangles. We derive the characteristics of this array from a linear transformation of the rectangular array [6, p. 11-23ff]. Because the array has six-way symmetry, Figure 3-17 the pattern of a uniform-amplitude 61-element hexagon array shows the same six-way symmetry in the ring sidelobe around the main beam. If we collapsed the hexagonal distribution to a line in one plane, the distribution has a taper that reduces the sidelobes. The sidelobe amplitudes of the uniform hexagonal array are lower that the principal-plane sidelobes of the rectangular uniform array. Figure 3-18 plots the spherical pattern of the hexagon array when scanned to 36°. The first ring sidelobe has a distorted six-fold symmetry. Similar to the scanned rectangular array (Figure 3-14), the hexagon array moves more sidelobes into visible space in the area opposite the scanned main beam. Figure 3-14 showed a grating lobe entering visible space, but the hexagon array pattern in Figure 3-18 does not. The grating lobes of a rectangular array can be found from a linear array when it is scanned along one of the principal axes, but the hexagon array requires a more elaborate analysis. When we scan the rectangular array off the principal axes, we can no longer use the grating lobe analysis of linear arrays.

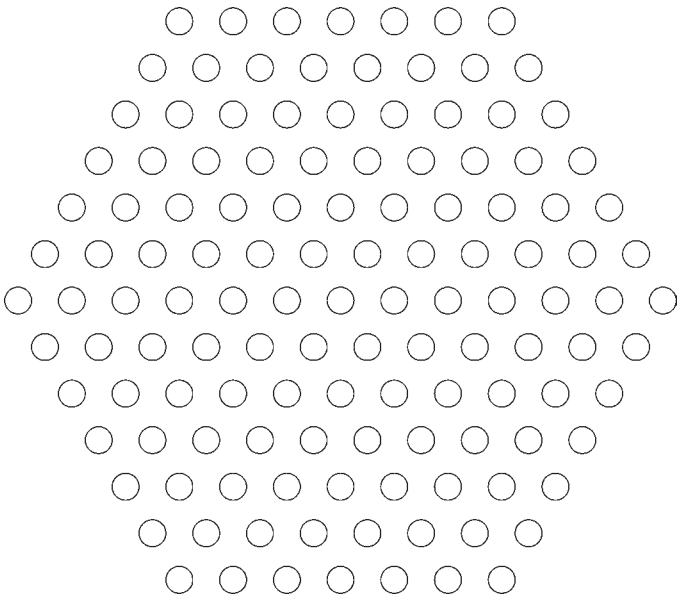


FIGURE 3-16 Position of elements in a hexagonal planar array.

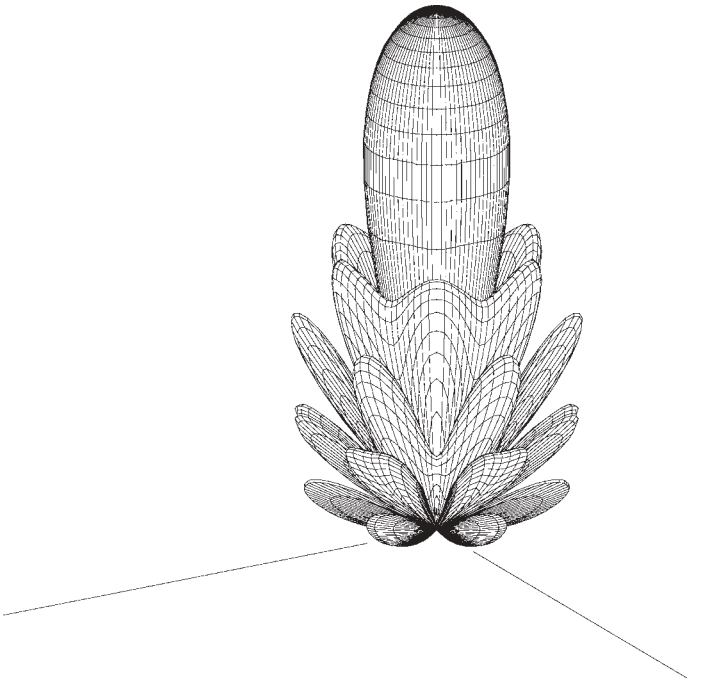


FIGURE 3-17 Spherical radiation pattern of a 61-element hexagonal array.

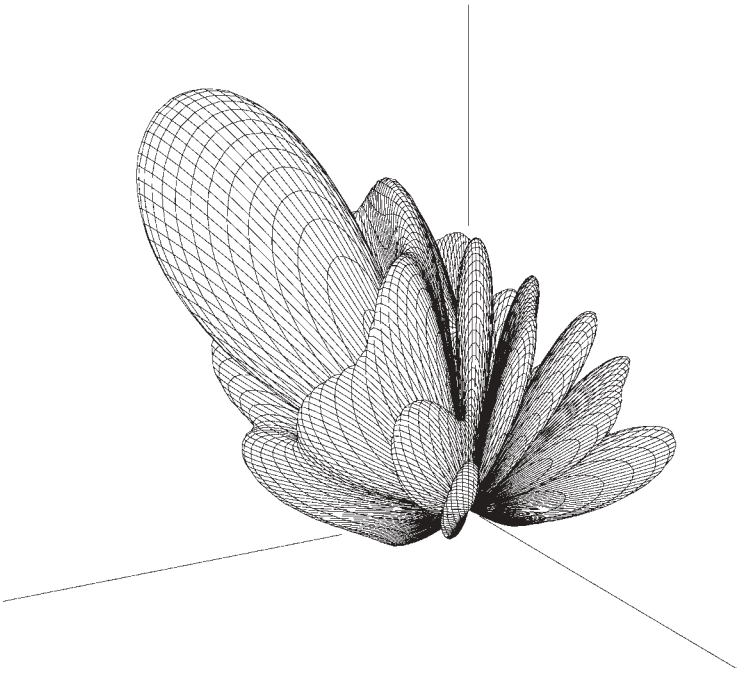


FIGURE 3-18 Spherical radiation pattern of a 61-element hexagonal array scanned along a principal plane.

3-8 GRATING LOBES IN PLANAR ARRAYS

The circle diagrams of the linear array can be used in principal planes of a rectangular array to compute grating lobes. For the planar array we use a $\sin\theta$ pattern space to see the periodicity of the grating lobes and to analyze scanning in planes other than the principal axes. The visible region is now limited to a unit circle in $k_x k_y$-space where $k_x = \sin\theta\cos\phi$ and $k_y = \sin\theta\sin\phi$. k_x is the pattern in the $x-z$ plane, and k_y is the pattern in the $y-z$ plane. Figure 3-19*a* shows the array layout, and Figure 3-19*b* shows the corresponding $k_x k_y$-plane grating lobe diagram. We reduce a contour plot of the pattern response similar to Figure 3-15 to only the main beams for analysis of grating lobes. The full contour plot is too busy. The narrower x-axis array spacing compared to the y-axis spacing leads to wider-spaced grating lobes in the k_x-plane than in the k_y-plane.

Beam scanning corresponds to movement of the unit circle in $k_x k_y$-space. Each small circle in $k_x k_y$-space is a main beam in pattern space. When the unit circle encloses more than one $k_x k_y$-circle, the pattern has multiple main beams or grating lobes. The $k_x k_y$-plane diagram could also include sidelobe peaks or the contour plot of the array pattern and illustrate the pattern change with scan. The $k_x k_y$-plane is the two-dimensional Fourier transform of the distribution that becomes the periodic two-dimensional Fourier series because the distribution is discrete. Increasing frequency or relative spacing between elements increases the unit circle diameter on an existing $k_x k_y$-diagram in a manner similar to the circle diagram.

When we scan the beam, we move the unit-circle center in $k_x k_y$-space. We use Eq. (3-13) to locate the unit-circle center on the diagram: $k(\sin\theta_0\cos\phi_0, \sin\theta_0\sin\phi_0)$. The off-center circle on Figure 3-19 corresponds to a scanned beam and encloses two main beams. In this case the grating lobe does not lie in the scan plane and would fail to show in a simple pattern cut through the scan plane. A rectangular array produces a rectangular grating lobe diagram, while other periodic arrays lead to more complicated grating lobe diagrams.

Figure 3-20 shows the layout of the hexagonal array and the corresponding grating lobe diagram. The hexagonal array (or equilateral triangular array) can be found from a linear transformation of the rectangular array. The grating lobe diagram can be found from the transformation as well. The spacing along the x-axis A_1 corresponds to the

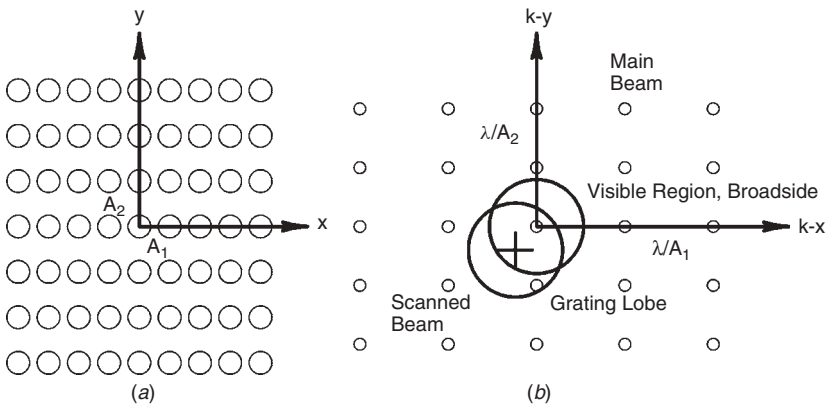


FIGURE 3-19 (*a*) Grating lobe diagram of a rectangular array; (*b*) distribution in k-space.

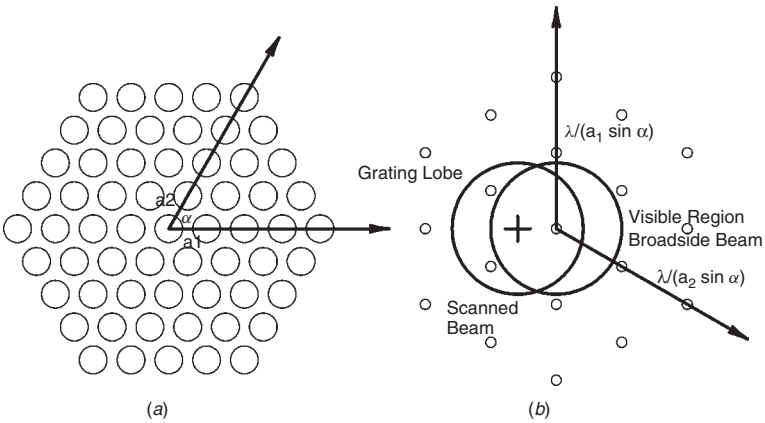


FIGURE 3-20 (*a*) Grating lobe diagram of a hexagonal array; (*b*) distribution diagram.

vertical spacing B_2 on the grating lobe diagram, and the spacing along the diagonal of each diagram are related. In both cases the corresponding axes on the two diagrams are perpendicular:

$$B_2 = \frac{\lambda}{A_1 \sin \alpha} \quad \text{and} \quad B_1 = \frac{\lambda}{A_2 \sin \alpha}$$

The angle between the triangular axes α is 60° for the hexagonal array. By allowing a grating lobe when the beam is scanned to 90°, we can determine the maximum element spacing without grating lobes:

$$\frac{\lambda}{A_1 \sin 60^\circ} = 2 \quad \text{or} \quad \frac{A_1}{\lambda} = \frac{1}{2 \sin 60^\circ} = 0.577$$

Figure 3-20 shows the visible region unit circle with the beam broadside to the plane and then scanned to 36° for an element spacing of λ. When scanned, the unit circle encloses three lobes. The three lobes do not lie in a plane. Figure 3-21 gives the

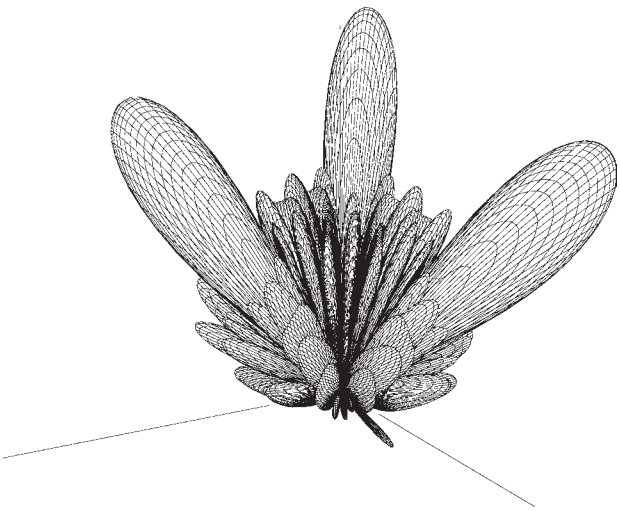


FIGURE 3-21 Spherical radiation pattern of hexagon-array grating lobes.

spherical pattern of the array when scanned to 36° and shows the three lobes in the pattern. The array sidelobes have been reduced by sampling a circular Taylor distribution with the array so that the lobes show clearly. In Chapter 4 we discuss the use of continuous aperture distributions to determine the feed amplitudes of planar arrays.

3-9 MUTUAL IMPEDANCE

Antennas in an array couple to each other because they receive a portion of the power radiated from nearby elements. This affects the input impedance seen by each element, which depends on the array excitation. We scan a phased array by changing the feeding coefficients, and this changes the element input impedance called the *scan impedance*. To first order, the coupling or mutual impedance is proportional to the element pattern level along the array face, and we reduce coupling by using narrower-beamwidth elements. Mutual coupling can be represented by an impedance, admittance, or scattering parameter matrix.

The first element of an N-element array has the impedance equation

$$V_1 = Z_{11}I_1 + Z_{12}I_2 + Z_{13}I_3 + \cdots + Z_{1N}I_N$$

If we know the radiation amplitudes, we calculate the ratio of the currents:

$$V_1 = I_1\left(Z_{11} + \frac{I_2}{I_1}Z_{12} + \frac{I_3}{I_1}Z_{13} + \cdots + \frac{I_N}{I_1}Z_{1N}\right)$$

The effective or scan impedance of the first element is

$$Z_1 = \frac{V_1}{I_1} = Z_{11} + \frac{I_2}{I_1}Z_{12} + \frac{I_3}{I_1}Z_{13} + \cdots + \frac{I_N}{I_1}Z_{1N} \tag{3-18}$$

It depends on the self-impedance and the excitation of all the other antennas. *Scan impedance* was formerly called *active impedance*, but this led to confusion. The power into the first element is

$$P_1 = \mathrm{Re}(V_1 I_1^*) = I_1 I_1^* \mathrm{Re}\left(Z_{11} + \frac{I_2}{I_1}Z_{12} + \frac{I_3}{I_1}Z_{13} + \cdots + \frac{I_N}{I_1}Z_{1N}\right) \tag{3-19}$$

By knowing the feeding coefficients and the mutual impedances, we can compute the total input power and gain. In general, every antenna in the array has different input impedances. As the feeding coefficients change in a phased array to scan the beam, so will the impedance of elements. The scan impedance change with scan angle causes problems with the feed network. We can repeat the same arguments for slots using mutual conductance, since magnetic currents are proportional to the voltage across each slot.

3-10 SCAN BLINDNESS AND ARRAY ELEMENT PATTERN [7, pp. 339–355; 8, pp. 365–366]

Large arrays made from elements with wide beamwidths can exhibit scan blindness. When a phased array is scanned, at certain angles the input reflection coefficient of every element rapidly increases to 1. The array fails to radiate and forms a pattern

null. Mutual coupling between elements causes the change of the scan impedance, which leads to scan blindness, which is complicated and difficult to predict accurately except where the array structure supports a surface wave. One approach says that scan blindness occurs when a grating lobe first enters from invisible space and radiates along the surface of the array. In this case we solve Eq. (3-16) for the scan angle of the grating lobe:

$$|\cos\theta_{\text{gr}}| = \frac{\lambda}{d} - 1 \tag{3-20}$$

The angle in Eq. (3-20) is measured from the array plane (or axis). Scan blindness occurs approximately at this angle, but it can be reduced to only a dip in the pattern if the array is small or the mutual coupling between the elements is small, because they have narrow beams.

The grating lobe causes a large increase in mutual coupling. Arrays made with antennas that can support surface waves, such as microstrip patches on dielectric substrates, can exhibit scan blindness when the electrical distance between the elements equals the surface-wave propagation phase shift:

$$|\cos\theta_{\text{gr}}| = \frac{\lambda}{d} - \frac{k_{\text{sw}}}{k} = \frac{\lambda}{d} - P \tag{3-21}$$

P is the relative propagation constant with a value > 1 for a surface wave (Section 10-1). Scan blindness will occur at an angle near this value because of the complicated nature of the coupling addition in the array.

We can build a small portion of the array and determine where scan blindness will occur. Feed the center element and load all others with the feeder resistance. Each element in an array will couple to its neighboring elements and we can associate the combination of the element radiation and the coupled radiation of the neighbors when loaded to the element. We call this the *array element pattern* or *scan element pattern* (formerly called the *active pattern*). Elements near the edges will have different effective patterns, but in a first-order solution we assume the pattern of the center element for all and calculate the total pattern as the product of the element pattern and the array factor. The array element pattern will exhibit dips where scan blindness will occur in the full array. Because it is only a small portion of the array, the full scan blindness will not occur. You should build a small array and test for scan blindness whenever it is a possibility. For example, arrays that scan to large angles off broadside using broad-beamwidth elements need to be tested with a small array before building the complete array.

3-11 COMPENSATING ARRAY FEEDING FOR MUTUAL COUPLING

Mutual coupling (impedance) is a measure of how much one antenna receives radiation from its neighbors in the array. Each element radiation changes the effective excitation on its neighbors. In a large array not requiring exact patterns, the effects average out. But when the array is small or you try to achieve low sidelobes, mutual coupling must be compensated for in the array. Small antenna elements such as dipoles or slots are resonant structures that radiate in only one mode. Mutual coupling only changes the element excitation, not the shape of the current distribution on the element. In this case we measure or calculate the mutual coupling matrix and use it to compute element

excitation to achieve the desired excitation [9]. Find the coupling matrix by adding the identity matrix to the S-parameter matrix of the antenna coupling:

$$\mathbf{C} = \mathbf{I} + \mathbf{S} \tag{3-22}$$

We compute the new feed excitation from the desired excitation and matrix inverse of Eq. (3-22):

$$\mathbf{V}_{\text{required}} = \mathbf{C}^{-1}\mathbf{V}_{\text{desired}} \tag{3-23}$$

Because we assumed a single mode distribution on the antenna elements, **S** is independent of scanning and Eq. (3-23) gives the compensation for all scan angles. The compensation can be applied to the received signal in an adaptive array by matrix multiplication in digital signal processing. The effects of these operations have been illustrated [10]. Without compensation adaptive arrays, such as the MUSIC algorithm, only generate small peaks, whereas compensation produces the expected large peaks.

Compensation for multimode elements starts with a moment solution [11] and uses the pattern characteristics to solve for the feeding coefficients. We use the pattern desired to compensate the feeding coefficients. We start with a matrix between the pattern response and the currents on all the antenna elements found from a moment method solution, where each array element has multiple current segments:

$$\mathbf{A}(\mathbf{k}) = \mathbf{FI} \tag{3-24}$$

$\mathbf{A}(\mathbf{k}_i)$ is an element of the column matrix that gives the pattern response at an angle given by $\mathbf{k}_i = \hat{\mathbf{x}}\sin\theta_i\cos\phi_i + \hat{\mathbf{y}}\sin\theta_i\sin\phi_i + \hat{\mathbf{z}}\cos\theta_i$ or a given pattern angle (θ_i, ϕ_i). The elements of the matrix **F** are the isotropic element phase terms, $e^{j\mathbf{k}_i\cdot\mathbf{r}_j}$, and **I** is the column vector of the currents on the segments. We calculate excitation voltages by inverting the mutual impedance matrix:

$$\mathbf{I} = \mathbf{Z}^{-1}\mathbf{V} \tag{3-25}$$

We substitute Eq. (3-25) into Eq. (3-24) and note that the matrix **V** has only q nonzero terms corresponding to the feed points. We specify q pattern points, which reduces **F** to $q \times M$ for M current segments. The vector **V** has $M - q$ zero elements and we delete the corresponding columns in the matrix product $\mathbf{FZ}^{-1}$. This reduces the matrix to $q \times q$, denoted **B:**

$$\mathbf{A}(\mathbf{k}) = \mathbf{BV}'$$

This uses the nonzero element $\mathbf{V}'$. We solve for the feeding coefficients by inverting the matrix **B** found from q pattern points:

$$\mathbf{V}' = \mathbf{B}^{-1}\mathbf{A}(\mathbf{k}) \tag{3-26}$$

Choosing good pattern points is an art that requires pattern evaluation to verify whether the final pattern is acceptable.

3-12 ARRAY GAIN

We can use the mutual impedance concept to determine the effective input power of every element and thereby avoid having to integrate the pattern to calculate average

radiation intensity. We represent the circuit relations of two antennas by a two-port impedance matrix:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}$$

Symmetrical elements across the diagonal of the matrix are equal for antennas satisfying reciprocity. The total input power is given by

$$P_{\text{in}} = \text{Re}(V_1 I_1^*) + \text{Re}(V_2 I_2^*)$$

The general N-element array has an $N \times N$ matrix and N terms in the input power sum. Given the feed coefficients, we have a relation between different I_i. For our two elements,

$$I_2 = I_1 e^{j\delta} \quad \text{and} \quad V_1 = (Z_{11} + Z_{12} e^{j\delta}) I_1$$

The power into the first element is

$$\text{Re}(Z_{11} + Z_{12} e^{j\delta}) I_1 I_1^*$$

By symmetry, the power into the second antenna is the same. The total input power to the array is

$$P_{\text{in}} = 2\ \text{Re}(Z_{11}) I_1 I_1^* \left[1 + \frac{\text{Re}(Z_{12} e^{j\delta})}{\text{Re}(Z_{11})} \right]$$

The factor $\text{Re}(Z_{11}) I_1 I_1^*$ is the power into an isolated element: $4\pi E_0^2/\eta$. The average radiation intensity (100% efficient antenna) is $P_{\text{in}}/4\pi$:

$$\text{gain} = \text{directivity} = \frac{|2E(\theta_{\max})|^2/\eta}{P_{\text{in}}/4\pi} = \frac{|2E(\theta_{\max})|^2}{1 + [\text{Re}(Z_{12} e^{j\delta})]/[\text{Re}(Z_{11})]} \tag{3-27}$$

By comparing Eqs. (3-27) and (3-8), we can identify

$$\frac{\text{Re}(Z_{12} e^{j\delta})}{\text{Re}(Z_{11})} = \frac{R_{12} \cos\delta}{R_{11}} = \frac{\sin(2\pi d/\lambda) \cos\delta}{2\pi d/\lambda}$$

$$\frac{R_{12}(d)}{R_{11}} = \frac{\sin(2\pi d/\lambda)}{2\pi d/\lambda}$$

We can use this mutual impedance ratio to compute directivity of arrays of isotropic elements of any number.

Example Calculate the directivity of a linear array of three equally spaced isotropic elements with equal amplitudes and phases.

The powers into the elements are

$$P_1 = P_3 = \frac{4\pi E_0^2}{\eta} \left[1 + \frac{R_{12}(d)}{R_{11}} + \frac{R_{12}(2d)}{R_{11}} \right]$$

$$P_2 = \frac{4\pi E_0^2}{\eta} \left[1 + \frac{2R_{12}(d)}{R_{11}} \right] \qquad U_{\max} = \frac{3^2 E_0^2}{\eta}$$

The total power into the array is found from the sum:

$$P_t = P_1 + P_2 + P_3 = \frac{4\pi E_0^2}{\eta}\left[1 + \frac{4R_{12}(d)}{R_{11}} + \frac{2R_{12}(2d)}{R_{11}}\right]$$

$$\text{directivity} = \frac{U_{\max}}{P_t/4\pi} = \frac{9}{3 + [4R_{12}(d)/R_{11}] + [2R_{12}(2d)/R_{11}]}$$

The directivity of the general N-element equally spaced linear array, excited by equal-amplitude and equal-phase signals, is easily found by extending the development:

$$\text{directivity} = \frac{N^2(\text{element directivity})}{N + 2\sum_{M=1}^{N-1}(N-M)[R_{12}(Md)/R_{11}]} \tag{3-28}$$

The directivity attained in an array depends on the particular mutual impedance terms of the radiators. The equation above only handles uniform-amplitude linear arrays. We can extend the idea of mutual resistance to calculate input power to a general planar array consisting of identical elements and determine gain.

By using a two-element array spaced along the x-axis we can integrate the pattern to compute directivity and from that determine the ratio of mutual resistance to self-resistance of the elements versus element spacing:

$$\frac{R_{12}(x)}{R_{11}} = \frac{\text{element directivity}}{2\pi}\int_0^{2\pi}\int_0^{\pi} E_e^2(\theta,\phi)\cos^2\left(\frac{\pi x}{\lambda}\cos\phi\sin\theta\right)\sin\theta\, d\theta\, d\phi - 1 \tag{3-29}$$

Equation (3-29) uses the normalized element pattern in the integral. By using an axisymmetrical element pattern, we calculate the ratio of resistances at a number of different distances and interpolate on the table for the directivity (gain) analysis of a planar (linear) array. If the element pattern is not symmetrical, the normalized resistance must be calculated for a number of ϕ. Given the element excitations E_i with elements located at the vector locations $\boldsymbol{x}_i$, we can derive an equation similar to Eq. (3-28) for directivity of a planar array:

$$\text{directivity} = \frac{\left|\sum_{i=1}^{N} E_i\right|^2 (\text{element directivity})}{\sum_{i=1}^{N}\sum_{j=1}^{N}[R_{12}(|\mathbf{x}_i - \mathbf{x}_j|)/R_{11}]\text{Re}[E_j/E_i]|E_i|^2} \tag{3-30}$$

Figure 3-22 illustrates the directivity calculated from Eq. (3-30) for linear arrays with realistic elements, such as a microstrip patch with 90° beamwidths, as the element spacing is varied. The graph shows directivity reduction when the element spacing exceeds λ and grating lobes form a more pronounced characteristic as the number of elements increases. When the second grating lobe occurs for wider element spacing, the directivity exhibits only minor variations. Increasing the element directivity (decreased beamwidth) reduces variation because the element pattern reduces the grating lobe.

We use Eq. (3-30) with a planar array to obtain Figure 3-23. This array consists of 217 elements arranged in a hexagonal pattern, with amplitudes found from sampling a circular Taylor distribution (Sections 4-18 and 4-19) to lower the sidelobes. The 30-dB circular Taylor distribution reduces the gain by 0.6 dB relative to a uniform distribution

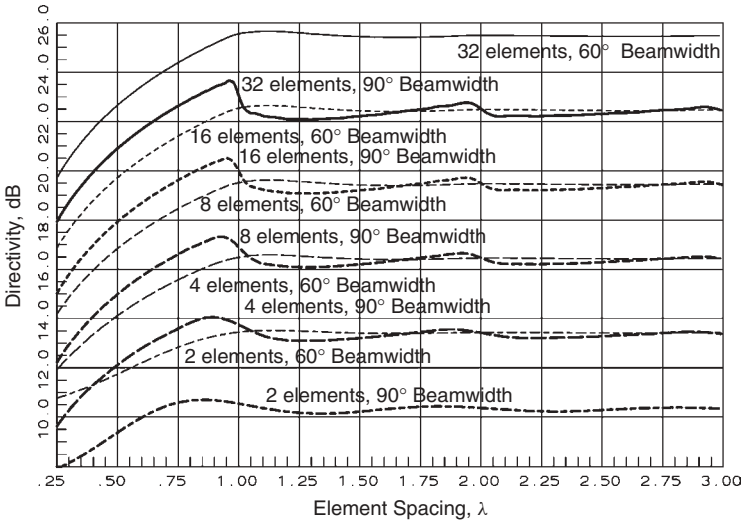


FIGURE 3-22 Directivity of a uniform-amplitude line array versus element spacing for 60° and 90° beamwidth elements.

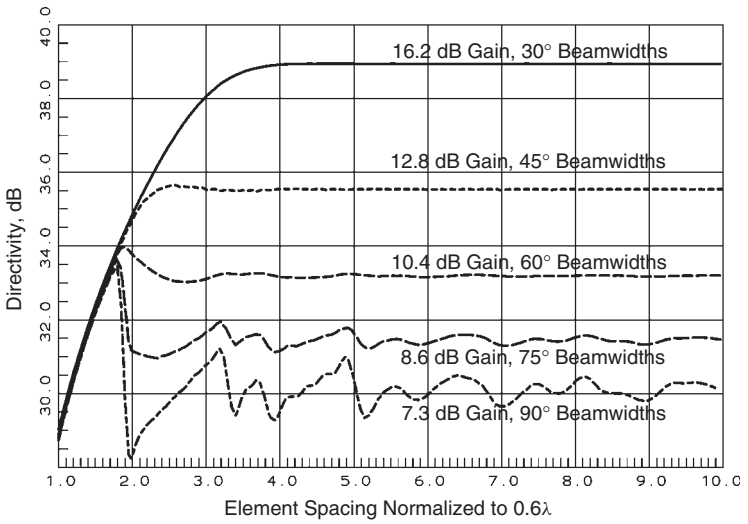


FIGURE 3-23 Directivity of a 217-element uniform-amplitude hexagonal array for various element beamwidths versus spacing.

due to the amplitude taper across the array. Initially, the element spacing is 0.6λ, but the element spacing has been allowed to grow in Figure 3-23 to show the effect. Figure 3-23 also illustrates the effect of increasing the element gain on the gain of a planar array. When we space the elements less than λ, increasing the element gain has no effect on array gain because the effective area of an antenna with a 90° beamwidth exceeds the area between elements and collects all power incident on the array. If we increase the element gain, the effective areas of the elements overlap and they share the incident power. On Figure 3-23 the curves overlap for element spacing less than

about λ. At the lower end of Figure 3-23 the gain increases by 6 dB as the element spacing doubles. This shows that increasing the element gain will have no effect on array gain when the present element covers the area associated with it.

When the grating lobe enters from invisible space as the element spacing increases, those arrays with narrower-beam elements suppress the lobes and continue the general gain increase. The array directivity (gain) drops as the element spacing increases for the wide-beamwidth element with 90° beamwidth because of the grating lobe. At a large element spacing the array gain becomes N times the element gain. We determine array gain from the array area and the amplitude taper for closely spaced elements. For wide element spacing, we calculate gain from the product of the number of elements and the element gain. Figure 3-23 shows a smooth transition between the two regions.

3-13 ARRAYS USING ARBITRARILY ORIENTED ELEMENTS

When we mount arrays on vehicles, the elements are pointed in arbitrary directions. Although Eq. (3-1) will calculate the pattern of any array, the element patterns are usually measured in a coordinate system in a different orientation than in the array. The idea of an array factor times the element pattern collapses and an analysis must rotate the pattern direction into the coordinates of each element. We will use coordinate rotations on the elements not only to specify them, but to calculate the pattern of the array. In a later chapter we use the same concept to point a feed antenna at a reflector.

We rotate the pointing direction into the coordinate system of the orientated antenna to determine what direction angles to use for the element pattern. We do this by using a 3×3 rotation matrix on rectangular components:

$$[X_{\text{rotated}} \quad Y_{\text{rotated}} \quad Z_{\text{rotated}}] = [\text{rotation matrix}] \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{3-31}$$

A similar problem is rotating an object. Both cases use the same matrix. To rotate an object we multiply the vector by the rotation matrix from the left to compute the rotated coordinates. Rotating a position is given by the equation

$$[X_{\text{old}} \quad Y_{\text{old}} \quad Z_{\text{old}}][\text{rotation matrix}] = \begin{bmatrix} X_{\text{rotated}} \\ Y_{\text{rotated}} \\ Z_{\text{rotated}} \end{bmatrix} \tag{3-32}$$

The rotation matrix can be found from the directions of the unit vectors when rotated. It is given by

$$\text{rotation matrix} = \begin{bmatrix} \text{rotated } X\text{-axis} \\ \text{rotated } Y\text{-axis} \\ \text{rotated } Z\text{-axis} \end{bmatrix} \tag{3-33}$$

The method uses 3×3 matrices to perform the rotation by a multiplication with a position or direction vector. Rotation about the X-axis is given by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos A & \sin A \\ 0 & -\sin A & \cos A \end{bmatrix}$$

rotation about the Y-axis is given by

$$\begin{bmatrix} \cos B & 0 & -\sin B \\ 0 & 1 & 0 \\ \sin B & 0 & \cos B \end{bmatrix}$$

and rotation about the Z-axis is given by

$$\begin{bmatrix} \cos C & \sin C & 0 \\ -\sin C & \cos C & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

We use products of these axis rotations to reorient an object or pointing direction. Consider the rotation of a position by the product of three rotation matrices:

$$[X_{\text{old}} \quad Y_{\text{old}} \quad Z_{\text{old}}]\mathbf{R}_1\mathbf{R}_2\mathbf{R}_3 = \begin{bmatrix} X_{\text{rotated}} \\ Y_{\text{rotated}} \\ Z_{\text{rotated}} \end{bmatrix}$$

The logical approach is to multiply the 3×3 matrices, $\mathbf{R}_1$, $\mathbf{R}_2$, and $\mathbf{R}_3$, before multiplying by the position vector. When we postmultiply $\mathbf{R}_1$ by $\mathbf{R}_2$, it rotates the axis of rotation of $\mathbf{R}_1$. The postmultiplication by $\mathbf{R}_3$ rotates the rotation axis $\mathbf{R}_2$ and $\mathbf{R}_1$ is rotated once more. We can take the rotations one by one from left to right and use the rotation matrices about each of the principal axes provided that we convert the column vector back to a row vector after each multiplication.

A convenient way to define the orientation of objects in space is to use spherical coordinate angles, since they are the same as pattern angles. We line up the matrices from right to left in this case. When rotating the coordinate system about an axis, the other axes change direction. The next rotations are about these new axes. The three rotations are often called the *Euler angles*. We use the following three rotations for spherical coordinate pointing:

1. Z-axis rotation $= \phi$
2. New Y-axis rotation $= \theta$
3. New Z-axis rotation: aligns the polarization of the antenna

The last rotation takes some thought because the first two rotations have altered the orientation of the antenna.

When calculating the pattern of the array for a particular direction, first compute rectangular components of the direction vector and the two polarization vectors. Multiply the direction vector by $k(2\pi/\lambda)$ and take the dot (scalar) product with the position vector to calculate phasing of a particular element. You need to determine the pattern direction in the rotated antenna's coordinate system found by using Eq. (3-31). Multiply the rotation matrix by the unit direction vector placed to the right. When you convert the output vector to spherical coordinates, you obtain pattern coordinates of the rotated antenna. Both the pattern components of the rotated element and the unit polarization vectors are needed. In the next operation you rotate the prime coordinate polarization unit vectors into the rotated element coordinate system using the same operation as the direction vector.

You calculate final radiated components by projecting the rotated prime coordinate polarization vectors on the element pattern unit polarization vectors:

$$E_\theta = E_{\theta,\text{element}}\hat{\theta}_{\text{element}} \cdot \hat{\theta}_{\text{rotated}} + E_{\phi,\text{element}}\hat{\phi}_{\text{element}} \cdot \hat{\theta}_{\text{rotated}}$$
$$E_\phi = E_{\theta,\text{element}}\hat{\theta}_{\text{element}} \cdot \hat{\phi}_{\text{rotated}} + E_{\phi,\text{element}}\hat{\phi}_{\text{element}} \cdot \hat{\phi}_{\text{rotated}} \quad (3\text{-}34)$$

Since we measure element patterns on antenna positioners, it is convenient to consider positioners as a series of coordinate system rotations.

REFERENCES

1. C. A. Balanis, *Antenna Theory, Analysis and Design*, 2nd ed., Wiley, New York, 1997.
2. W. W. Hansen and J. R. Woodyard, A new principle in directional antenna design, *Proceedings of IRE*, vol. 26, March 1938, pp. 333–345.
3. T. C. Cheston and J. Frank, Array antennas, Chapter 11 in M. I. Skolnik, ed., *Radar Handbook*, McGraw-Hill, New York, 1970.
4. J. L. Butler, Digital, matrix, and intermediate frequency scanning, Section 3 in R. C. Hansen, ed., *Microwave Scanning Antennas*, Vol. III, Academic Press, New York, 1966.
5. J. Blass, The multidirectional antenna: a new approach to stacked beams, *IRE International Convention Record*, vol. 8, pt. 1, 1960, pp. 48–50.
6. Y. T. Lo and S. W. Lee, eds., *Antenna Handbook*, Vol. II, Van Nostrand Reinhold, New York, 1993.
7. R. J. Mailloux, *Phased Array Antenna Handbook*, Artech House, Boston, 1994.
8. P.-S. Kildal, *Foundations of Antennas*, Studentlitteratur, Lund, Sweden, 2000.
9. H. Steyskal and J. S. Herd, Mutual coupling compensation in small array antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-38, no. 12, December 1990, pp. 1971–1975.
10. A. G. Derneryd, Compensation of mutual coupling effects in array antennas, *IEEE Antennas and Propagation Symposium*, 1996, pp. 1122–1125.
11. B. J. Strait and K. Hirasawa, Array design for a specified pattern by matrix methods, *IEEE Transactions on Antennas and Propagation*, vol. AP-18, no. 1, January 1971, pp. 237–239.

4

APERTURE DISTRIBUTIONS AND ARRAY SYNTHESIS

Continuous apertures and arrays share similar characteristics. We compute the radiation pattern of the aperture by using the Fourier transform. Array sampling of an aperture distribution leads to a Fourier series analysis for its pattern. We rely on our familiarity with signal processing to give us insights into these processes and their characteristics. We apply aperture theory to the analysis of horns, lens, and reflector antennas, but it also describes array antennas. Since we can design antennas only approximately to produce particular aperture distributions, we often realize them by sampling with an array.

We start with aperture efficiencies developed from the Huygens source approximation of Section 2-2. We apply this method to horns, lens, and reflector antennas for both synthesis and tolerance analysis. The uniform and cosine distributions occur naturally in horns and simple resonant antennas. We use aperture distributions to realize bounds on antenna characteristics given size and excitation distribution.

Taylor developed an aperture distribution based on Dolph's use of the Chebyshev polynomials to produce the narrowest beamwidth for a specified sidelobe level for an array. The Chebyshev array design produces equal-amplitude sidelobes that we discover to be undesirable for large arrays because the equivalent aperture distribution peaks at the ends and the average value of the sidelobes limits the directivity to 3 dB above the sidelobe level. Large edge peaking of the distribution requires a feed network containing a large ratio of coupling values. Mutual coupling between elements causes unwanted excitation for a large ratio of element amplitudes and we lose control. Our usual practice is to sample a Taylor distribution for large arrays. The distribution has limited edge peaking, and large arrays can realize high gains.

Aperture distribution synthesis involves manipulating pattern nulls to achieve desired characteristics. Taylor used the zeros of the Chebyshev array to alter the positions of the inner nulls of the uniform distribution to lower sidelobe levels. Elliott extended this

Modern Antenna Design, Second Edition, By Thomas A. Milligan

idea to iterate the positions of these nulls to produce a linear aperture that radiates individually specified sidelobes. Schelkunoff developed a transformation between the pattern of an array and a polynomial where we combine the roots (or zeros) of the array polynomial in the complex plane with a mapped pattern variable that traverses the unit circle to analyze array patterns. We synthesize arrays by manipulating these polynomial zeros in the complex plane. Similar to Elliott's method of null positioning for the continuous linear aperture, Orchard (and Elliott) developed an iterative method applied to array polynomial zeros to synthesize arrays. The method allows us to specify sidelobes individually and to shape the main beam pattern by moving some zeros off the unit circle. When designing shaped beams, improved synthesis by the Orchard method reduces our use of both array sampling of the Woodward continuous aperture method and direct Fourier series synthesis for linear arrays, but both earlier methods give us insight. We consider the design of series feeding where elements are fed directly from a transmission line for a linear array or continuous linear aperture. This requires specification of the couplers or loading of the transmission line along the array because a portion of the power is extracted at each position with the remaining power dissipated in a load.

We repeat aperture analysis for circular apertures to show limitations of large reflector antennas and for use in sampling with an array. For planar arrays, we reduce many rectangular apertures to the product of two linear distributions. A Chebyshev-type planar array with equal sidelobes can be designed so that the sidelobes in the diagonal planes are not reduced unnecessarily. Convolution synthesis of planar arrays allows manipulation of the pattern zeros in groups of smaller arrays similar to the Schelkunoff method. Finally, we consider aperture blockage and phase errors that lead to gain reduction and increased sidelobes.

4-1 AMPLITUDE TAPER AND PHASE ERROR EFFICIENCIES

When we use the Huygens source approximation, we calculate power radiated by summing (integrating) the magnitude squared of the electric field in the aperture and dividing by the impedance of free space. The average radiation intensity is the radiated power divided by the area of a unit sphere, 4π. To complete the calculation, we compute the maximum radiation intensity by dividing the maximum of the magnitude squared of Eq. (2-24) by the impedance of free space and directivity ($U_{\max}/U_{\text{avg}}$) becomes

$$\frac{\pi(1+\cos\theta)^2}{\lambda^2}\frac{\left|\iint\limits_s E e^{j\mathbf{k}\cdot\mathbf{r}'}\,ds'\right|^2_{\max}}{\iint\limits_s |E|^2\,ds'} \tag{4-1}$$

Equation (4-1) can be used for directivity in any pattern direction, including the maximum of the numerator integral.

An aperture with a uniform amplitude and phase distribution has directivity $4\pi A/\lambda^2$, where A is the area. We separate directivity reductions into individual terms due to

aperture field amplitude and phase variations, and we express the general aperture directivity as

$$\text{directivity} = \frac{4\pi A}{\lambda^2} \cdot \text{ATL} \cdot \text{PEL}$$

where ATL is the amplitude taper efficiency (loss) and PEL is the phase error efficiency (loss). Only amplitude variations contribute to ATL, and only phase variations determine PEL.

We start with a uniform phase distribution in the aperture where the beam peak occurs normal to the aperture ($\theta = 0^\circ$) and PEL = 1. We obtain uniform phase fields by using $|E|$ in Eq. (4-1):

$$\text{directivity} = \frac{4\pi}{\lambda^2} \frac{\left(\iint_s |E|\, ds \right)^2}{\iint_s |E|^2\, ds} = \frac{4\pi A}{\lambda^2} \cdot \text{ATL}$$

where $k_x = k_y = 0$ on the boresight ($\theta = 0^\circ$). On solving for ATL, we derive

$$\text{ATL} = \frac{\left(\iint_s |E|\, ds \right)^2}{A \iint |E|^2\, ds} \tag{4-2}$$

We have forced a constant phase everywhere in the aperture to separate out the amplitude taper effects. We account for nonuniform phase with PEL. The phase error efficiency can be found from

$$\text{PEL}(\theta, \phi) = \frac{\text{directivity}(\theta, \phi)}{(4\pi A/\lambda^2) \cdot \text{ATL}}$$

where we use directivity (θ, ϕ) and PEL (θ, ϕ) depends on the pattern direction $(\theta,\ \phi)$:

$$\text{PEL}(\theta, \phi) = \frac{(1 + \cos\theta)^2}{4} \frac{\left| \iint_s E e^{j\mathbf{k}\cdot\mathbf{r}'}\, ds \right|^2}{\left(\iint_s |E|\, ds \right)^2} \tag{4-3}$$

$$\mathbf{k} = k(\sin\theta\cos\phi\hat{\mathbf{x}} + \sin\theta\sin\phi\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}})$$

For an aperture in the $x-y$ plane,

$$\mathbf{k} \cdot \mathbf{r}' = k(x' \sin\theta\cos\phi + y' \sin\theta\sin\phi)$$

We determine maximum PEL to relate it and ATL to directivity. Traditionally, we use the boresight value ($\theta = 0^\circ$) and Eq. (4-3) reduces to

$$\text{PEL} = \frac{\left| \iint_s E\, ds \right|^2}{\left(\iint_s |E|\, ds \right)^2} \tag{4-4}$$

Unless specified, PEL will be Eq. (4-4) and we use Eq. (4-3) for scanned apertures.

Equations (4-2) and (4-4) separate the effects of amplitude and phase variations in the aperture on the directivity at the boresight. If these efficiencies are expressed in decibels, the directivity becomes

$$\text{directivity(dB)} = 10 \log \frac{4\pi A}{\lambda^2} + \text{ATL}_{\text{dB}} + \text{PEL}_{\text{dB}}$$

Expressed in decibels, the efficiencies are called *losses:* amplitude taper loss (ATL) and phase error loss (PEL). It is important to remember that these are the losses at the boresight. A linear phase taper across the aperture scans the beam, but Eq. (4-4) predicts the boresight loss, which could be a null of the pattern. ATL is independent of phase variations that cause squinting of the beam.

4-1.1 Separable Rectangular Aperture Distributions

If the distribution in a rectangular aperture is separable,

$$E(x, y) = E_1(x) E_2(y)$$

the efficiencies also are separable.

$$\text{ATL} = \text{ATL}_x \text{ATL}_y \quad \text{and} \quad \text{PEL} = \text{PEL}_x \text{PEL}_y \tag{4-5}$$

Given a rectangular aperture with an x-axis excursion of $\pm a/2$,

$$\text{ATL}_x = \frac{\left[\int_{-a/2}^{a/2} |E_1(x)|\, dx \right]^2}{a \int_{-a/2}^{a/2} |E_1(x)|^2 dx} \tag{4-6}$$

$$\text{PEL}_x = \frac{\left| \int_{-a/2}^{a/2} E_1(x)\, dx \right|^2}{\left[\int_{-a/2}^{a/2} |E_1(x)|\, dx \right]^2} \tag{4-7}$$

The formulas for the y-axis are the same except for the substitution of y for x.

4-1.2 Circularly Symmetrical Distributions

If a circular aperture has a circularly symmetrical distribution, we easily reduce Eqs. (4-2) and (4-4) to

$$\text{ATL} = \frac{2\left[\int_0^a |E(r)|r\,dr\right]^2}{a^2\int_0^a |E(r)|^2 r\,dr} \tag{4-8}$$

$$\text{PEL} = \frac{\left|\int_0^a E(r)r\,dr\right|^2}{\left[\int_0^a |E(r)|r\,dr\right]^2} \tag{4-9}$$

where a is the radius.

We need a short word on formulas using integrals. They look formidable and seem to have little immediate practical use. In the catalog of distributions to follow, results will be given. A general distribution must be solved by numerical integration. One of the Newton–Cotes methods, such as Simpson's rule or the Rhomberg integration, can be used when evenly spaced values are known. With a known function for the distribution, we use the Gauss–Legendre technique, whereby the method selects the required function values. It is sometimes easier to calculate the integrals numerically instead of writing routines for special functions that arise with circular apertures. Exact expressions are ideal; but unless a distribution is forced by a mode on the structure, it is difficult to achieve the exact distribution. We need only approximations to the accuracy of practical interest.

4-2 SIMPLE LINEAR DISTRIBUTIONS

We assume that rectangular apertures have separable distributions so that we can deal with one coordinate at a time. We compute the pattern in the plane containing the line. By drawing the pattern in k_x (or k_y)-space, we can calculate patterns independent of the aperture size in a way similar to that used for arrays in Chapter 3. In Chapter 2 we derived the k_x-space pattern for a uniform distribution:

$$\frac{a\sin(k_x a/2)}{k_x a/2} \tag{4-10}$$

where a is the aperture width and $k_x = k\sin\theta\cos\phi$. We suppress $\cos\phi$ and consider only patterns in the $\phi = 0^\circ$ plane. Figure 4-1 shows the $k\sin\theta$ space pattern of a uniform distribution. The pattern does not repeat at 2π intervals (radians) as the array does, but the sidelobes continue to decrease at a rate of $1/x$. The first sidelobe is 13.2 dB below the peak. The aperture size a, along with the scanning variable $\sin\theta_0$, determines the visible region in Figure 4-1. It ranges between $\pm ka/2$ centered on $ka/2\sin\theta_0$, since the maximum value of $\sin\theta = 1$.

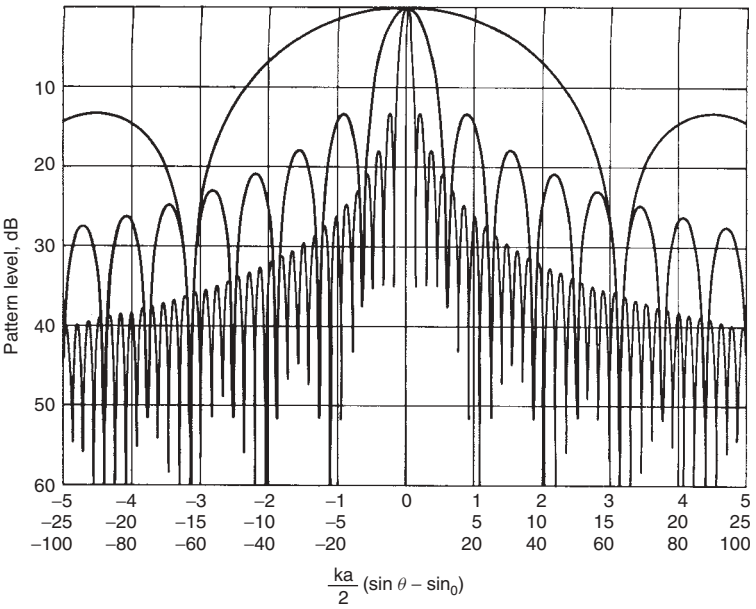


FIGURE 4-1 k_x-space pattern of uniform line-source distribution.

Example An aperture is four wavelengths long. Determine the number of sidelobes between $\theta = \pm 90^\circ$ when $\sin \theta_0 = 0$ (boresight).

The maximum value in ($k \sin \theta$)-space is

$$\frac{2\pi}{\lambda}\frac{4\lambda}{2} = 4\pi \quad \text{or} \quad 12.57$$

There are three sidelobes on each side of the main beam (Figure 4-1) in the visible region. The first sidelobe occurs when $ka/2 \sin \theta_1 = 4.5$, or

$$\theta_1 = \sin^{-1} \frac{4.5\lambda}{a\pi} = 21^\circ$$

We found the half-power beamwidth in Chapter 2:

$$\text{HPBW} = \sin^{-1} \frac{0.4429\lambda}{a} \tag{4-11}$$

valid when we ignore the obliquity factor, $(1 + \cos\theta)/2$. When we approximate $x = \sin x$ (radians) for small angles, we obtain

$$\text{HPBW} = 50.76^\circ \frac{\lambda}{a} \tag{4-12}$$

We use this as the standard and describe other HPBW by their beamwidth factors. The beamwidth factor of the uniform distribution is 1.00. We also consider the null

beamwidth (BW_{null}) of the distribution. The first null occurs at $\pm\pi$ in the $k \sin\theta$ pattern:

$$BW_{null} = 2\sin^{-1}\frac{\lambda}{a} \approx 114.59^\circ\frac{\lambda}{a}$$

We also establish a beamwidth factor for the null beamwidth. When we scan the beam to a direction θ_0, the visible region centers at $\pi a/\lambda \sin\theta_0$ in $(k \sin\theta)$-space.

Example Compute beam edges when $\theta_0 = 30^\circ$ and $a = 6\lambda$ for a uniform distribution.

$$\frac{a}{\lambda}(\sin\theta_{1,2} - \sin\theta_0) = \pm 0.4429$$

$$\sin\theta_{1,2} = \frac{\pm 0.4429}{6} + 0.5$$

$$\theta_1 = 35.02^\circ \qquad \theta_2 = 25.23^\circ$$

The beamwidth is the difference, 9.79°. If we take the beam center as the average between the 3-dB beam edges, we get 30.12° for the beam center. By using the cosine of the beam center times the aperture size, we get 5.19λ, the projected aperture dimension. On substituting this in Eq. (4-11), we calculate HPBW = 9.79°. The actual beam peak is at $\theta = 30^\circ$, but the pattern is asymmetrical about θ_0.

Other simple geometrical distributions on a linear aperture follow the same Fourier transform relation as the uniform distributions with differing transforms in $(k \sin\theta)$-space. Table 4-1 lists the properties of some common distributions.

Example Compute the beamwidth of a 7λ aperture with a cosine distribution.

From Table 4-1, the beamwidth factor = 1.342. The taper increases the beamwidth over that of a uniform distribution:

$$\text{HPBW} = 1.342\frac{50.76\lambda}{a} = 9.73^\circ \quad \text{or} \quad \text{HPBW} = 2\sin^{-1}\left(1.342\frac{0.4429\lambda}{a}\right) = 9.74^\circ$$

We can add distributions and calculate the pattern from the sum of the transforms. Adding a pedestal (uniform distribution) to the cosine-squared distribution decreases

TABLE 4-1 Common Linear Distribution Characteristics

Distribution	f_x	First Sidelobe (dB)	HPBW Factor	BW_{null} Factor	ATR (dB)
Uniform	$\dfrac{\sin(k_x a/2)}{k_x a/2}$	13.2	1.000	1.000	0
Triangular	$\left[\dfrac{\sin(k_x a/4)}{k_x a/4}\right]^2$	26.5	1.439	2.000	1.25
Cosine	$2\pi\dfrac{\cos(k_x a/2)}{\pi^2 - (k_x a)^2}$	23.0	1.342	1.5	0.91
Cosine^2	$\dfrac{\sin(k_x a/2)}{(k_x a/2)[1 - (a/2\lambda)]^2}$	31.5	1.625	2.000	1.76

the beamwidth and the sidelobes of the cosine-squared distribution. The aperture distribution is given by

$$E(x) = \text{PD} + (1 - \text{PD}) \cos^2 \frac{\pi x}{a} \qquad |x| \leq \frac{a}{2}$$

where PD is the voltage pedestal level. The first sidelobe of the uniform distribution lies within the null beamwidth of the cosine-squared distribution. The phase of sidelobes with respect to the main beam alternates between 180° and 0°, and the sidelobe of the pedestal subtracts from the main lobe. The second sidelobe of the pedestal occurs in almost exactly the same k-space location as the first sidelobe of the cosine-squared distribution. These lobes cancel each other to some extent. Table 4-2 gives the required pedestal measured relative to the peak of the distribution for a given maximum sidelobe level. The minimum sidelobes (43.2 dB) occur for a pedestal level of −22.3 dB. At lower pedestal levels, the sidelobes rise and the beamwidth factor increases at a constant rate as the pedestal level decreases.

The amplitude taper efficiency of the cosine squared on a pedestal is

$$\text{ATL} = \frac{2(1 + \text{PD})^2}{3 + 2\text{PD} + 3\text{PD}^2} \qquad \text{ratio} \tag{4-13}$$

Amplitude distributions based on simple functions have limited use. The uniform and cosine distributions or close approximations occur naturally, but the others must be forced on an aperture. An array can sample a distribution to achieve results similar to those for an aperture. A sampled cosine squared on a pedestal is handy for quick tolerance studies of array feed networks, but is far from optimum. Table 4-2 lists the pedestal to achieve a given sidelobe level for this distribution. We consider distributions that allow close control of sidelobes and achieve minimum beamwidths.

The rate of decrease of the far-out sidelobe depends on the functional relation of the distribution at the edges [1]. If α is the exponent of the distribution approximation x_e^{α}, where x_e is the distance from the edge, then the sidelobes decay as $U^{-(1+\alpha)}$, where U

TABLE 4-2 Pedestal Level to Achieve a Given Maximum Sidelobe Level for a Cosine Squared on a Pedestal Distribution

Sidelobe (dB)	Pedestal (dB)	Beamwidth Factor	ATL (dB)
30	−12.9	1.295	0.79
32	−14.2	1.325	0.89
34	−15.7	1.357	0.99
36	−17.3	1.390	1.10
38	−18.7	1.416	1.18
40	−20.0	1.439	1.25
42	−21.4	1.463	1.32
42.7[a]	−21.9	1.471	1.34
43.2[b]	−22.3	1.476	1.36

[a] Hamming distribution.
[b] Minimum sidelobe level.

is a linear function of the k-space variable. Both the triangular and cosine distributions have $\alpha = 1$, and the far-out sidelobes decay as $1/U^2$. The cosine-squared distribution sidelobes decay as $1/U^3$, since $\alpha = 2$. In the case of a cosine squared on a pedestal, the edge functional relation is a step (pedestal, $\alpha = 0$) and the far-out sidelobes decay as $1/U$. The sidelobes of the pedestal eventually overtake the cosine-squared distribution sidelobes, decreasing as $1/U^3$. To achieve uniform sidelobes, α must be -1, which occurs only when the distribution edges are Dirac delta functions, which requires infinite energy in the aperture or design reduction to discrete sources (an array).

We must accept a trade-off between radiated power in the main beam and in the sidelobes. When we narrow the main beam in a fixed size aperture, more power radiates in the sidelobes. We achieve minimum beamwidth in the main beam when all the sidelobes radiate the same power (maximum radiated power in the sidelobes for a given level) and all sidelobes are at the same level. This case leads to the Dolph–Chebyshev array [2], impossible to duplicate in a continuous aperture.

4-3 TAYLOR ONE-PARAMETER LINEAR DISTRIBUTION [3]

The uniform distribution has k-space zeros at (Figure 4-1) $\pm n\pi$, $n = 1, 2, 3, \ldots$. Taylor defines a new variable U to replace $k \sin\theta$:

$$\frac{\sin \pi U}{\pi U} \tag{4-14}$$

where $U = (a/\lambda)(\sin\theta - \sin\theta_0)$ and a is the aperture width. The nulls (zeros) are then located at integer values of U. Taylor adjusted the inner zeros of the uniform distribution to lower the sidelobes while retaining the outer zeros at their locations in the uniform distribution. The zeros are modified by a parameter B the boundary between the two regions in U-space:

$$U_n = \sqrt{n^2 + B^2} \tag{4-15}$$

The pattern has different expressions in two regions:

$$F(U) = \begin{cases} \dfrac{\sinh \pi\sqrt{B^2 - U^2}}{\pi\sqrt{B^2 - U^2}} & |U| \le B \qquad (4\text{-}16a) \\ \dfrac{\sin \pi\sqrt{U^2 - B^2}}{\pi\sqrt{U^2 - B^2}} & |U| \ge B \qquad (4\text{-}16b) \end{cases}$$

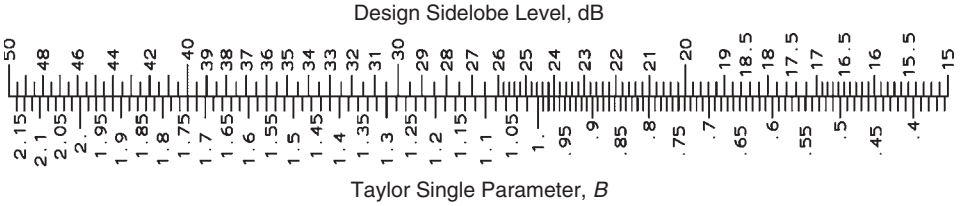
The high value of Eq. (4-16a) at the boresight depresses the sidelobes of the uniform distribution, and the parameter B controls all the parameters of the distribution. We compute B from the desired sidelobe level (SLR) by an iterative solution of the equation

$$\text{SLR} = 13.26 + 20 \log \frac{\sinh \pi B}{\pi B} \tag{4-17}$$

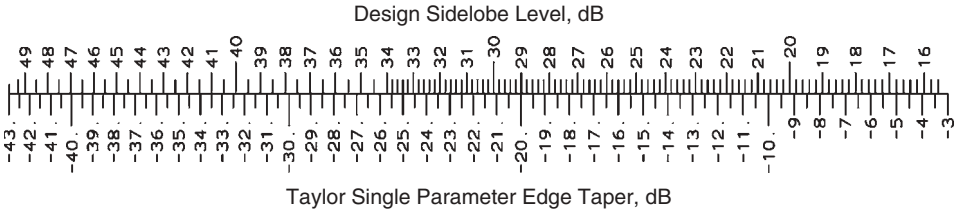
Scale 4-1 gives the Taylor single-parameter distribution B for a given sidelobe level.

The aperture distribution over the range -0.5 to 0.5 is given by the equation

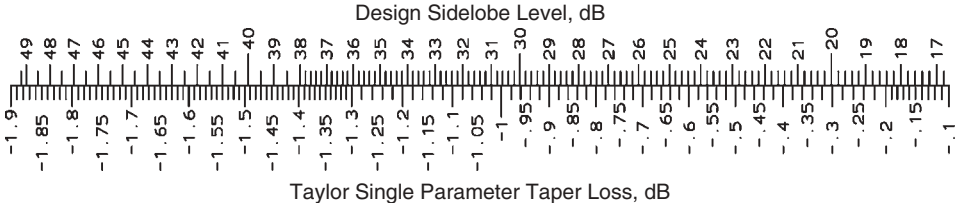
$$\frac{I_0[\pi B\sqrt{1 - (2x)^2}]}{I_0(\pi B)} \tag{4-18}$$



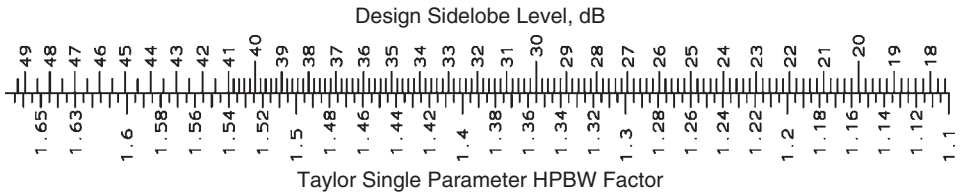
SCALE 4-1 Taylor single-parameter B for a sidelobe level.



SCALE 4-2 Taylor single-parameter edge taper for a given sidelobe level.



SCALE 4-3 Taylor single-parameter amplitude taper loss for a given sidelobe level.



SCALE 4-4 Taylor single-parameter HPBW factor for a given sidelobe level.

Using Eq. (4-18), we calculate aperture edge taper as a function of sidelobe level, given by Scale 4-2. By inserting the expression for the aperture distribution [Eq. (4-18)] into Eq. (4-6), we calculate amplitude taper loss as a function of sidelobe level (Scale 4-3). The HPBW factor can be found from Eq. (4-16) or read easily from Scale 4-4.

Figure 4-2 compares the U-space patterns of the Taylor one-parameter and uniform distributions. Synthesizing aperture distributions and arrays concentrates on the placement of pattern nulls. The one-parameter distribution scaled the locations of the nulls (zeros) by using Eq. (4-15). You should notice that the nulls approach those of the

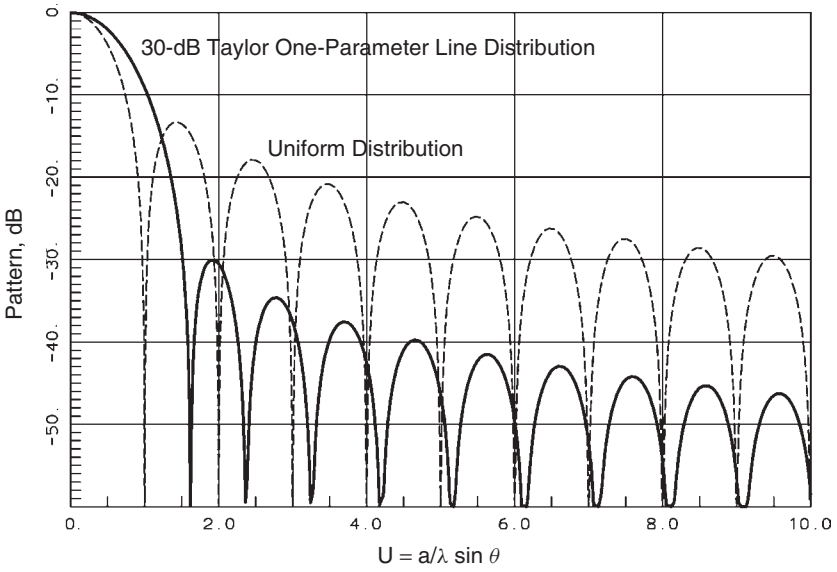


FIGURE 4-2 U-space pattern of 30-dB Taylor one-parameter linear distribution versus uniform distribution.

uniform distribution as U increases. Except for a shift near $U = 0$, the pattern falls off at a $1/U$ rate for far-out sidelobes.

You can use the one-parameter Taylor distribution to estimate the characteristics of a linear distribution for a given sidelobe level. A comparison of this distribution to the cosine squared on a pedestal (Table 4-3) shows that it is not as efficient for moderate sidelobe levels. The cosine squared on a pedestal distribution achieves low sidelobes by canceling sidelobes in two distributions and cannot be extended to any sidelobe level, whereas the one-parameter distribution can produce designs for any sidelobe level. More important, it demonstrates the systematic use of U-space pattern null placement for design. Taylor improved on this distribution by considering the zeros of the Dolph–Chebyshev array to flatten the first few sidelobes of the pattern response and achieved a more efficient distribution.

TABLE 4-3 Comparison Between the Taylor One-Parameter Distribution and Cosine Squared on Pedestal Linear Distribution for Selected Sidelobe Levels

Distribution	Pedestal (dB)	ATR (dB)	HPBW Factor
30-dB one-parameter	−21.13	0.96	1.355
30-dB $\cos^2$ + pedestal	−12.9	0.79	1.295
36-dB one-parameter	−28.49	1.30	1.460
36-dB $\cos^2$ + pedestal	−17.3	1.10	1.390
40-dB one-parameter	−32.38	1.49	1.524

4-4 TAYLOR $\overline{n}$ LINE DISTRIBUTION [1]

The Taylor $\overline{n}$ line-source distribution modifies the location of the inner pattern zeros (nulls) of a uniform distribution to approximate the Dolph–Chebyshev array. The distribution contains a pedestal $\alpha = 0$ and retains the $1/U$ fall-off for the far-out sidelobes. We can modify any number of inner zeros of the pattern to approximate the uniform sidelobe-level array, but we force the aperture voltage to peak at the ends in approximating the Dirac delta functions. We limit the number of altered zeros to keep the distribution practical. After a point, shifting more zeros reduces beamwidth negligibly.

We manipulate the location of pattern zeros to obtain desired patterns. Both aperture and array syntheses depend on zero locations. The number of array elements determines the number of independent zeros ($\overline{n} - 1$), but a continuous aperture has an infinite number of independent zeros. Practical consideration of the distribution edge shape limits the number, but we are free to move zeros. For a given aperture size, we can move zeros out of the invisible region into the visible region and narrow the main beam as much as we want while maintaining low sidelobes. The invisible region represents stored energy in the aperture. When a zero moves out of the invisible region, the amount of stored energy and the Q of the antenna increase. The overall efficiency of the antenna decreases while the antenna becomes more and more narrowband. We call these arrays superdirective because their directivity exceeds that of a uniform distribution. The Taylor line-source distribution retains the zeros in the invisible region and prevents superdirectivity. There is no limit to the directivity achievable on paper for a given aperture, but the theoretical distributions are unrealizable except for very small levels of superdirectivity. The costs of superdirectivity are decreased bandwidth and efficiency.

We will modify the location of the first $\overline{n} - 1$ pairs of inner nulls to lower the sidelobes. Choosing the zeros symmetrically about the origin of U-space gives us a constant phase distribution. We remove the inner zeros by dividing them out of the uniform distribution U-space pattern:

$$\frac{\sin \pi U}{\pi U \prod_{N=1}^{\overline{n}-1} \left(1 - U^2/N^2\right)}$$

We then add new nulls U_n without becoming superdirective:

$$F(U) = \frac{\sin \pi U \prod_{N=1}^{\overline{n}-1} \left(1 - \dfrac{U^2}{U_N^2}\right)}{\pi U \prod_{N=1}^{\overline{n}-1} \left(1 - \dfrac{U^2}{N^2}\right)} \tag{4-19}$$

Because we want to approximate the DoIph–Chebyshev response, we choose the inner zeros from the array:

$$U_N = \overline{n} \frac{\sqrt{A^2 + (N - \frac{1}{2})^2}}{\sqrt{A^2 + (\overline{n} - \frac{1}{2})^2}} \qquad N = 1, 2, \ldots, \overline{n} - 1 \tag{4-20}$$

where A relates to the maximum sidelobe level:

$$\cosh \pi A = b \tag{4-21}$$

in which $20 \log b =$ sidelobe level. Equation (4-19) gives us the U-space (k-space) pattern of the distribution with modified zeros. We determine the aperture distribution by expanding it in a Fourier cosine series:

$$E(x) = \sum_{m=0}^{\infty} B_m \cos 2m\pi x \qquad |x| \le 0.5 \tag{4-22}$$

where the aperture size has been normalized. We calculate the pattern of the distribution from the Fourier transform:

$$f(k_x) = \int_{-1/2}^{1/2} E(x) e^{jk_x x}\, dx \quad \text{or} \quad f(U) = \int_{-1/2}^{1/2} E(x) e^{j2\pi U x}\, dx$$

We substitute Eq. (4-22) for $E(x)$ and reverse the order of summation and integration:

$$f(U) = \sum_{m=0}^{\infty} B_m \int_{-1/2}^{1/2} \cos 2m\pi x \cos 2\pi U x\, dx \tag{4-23}$$

Since the aperture function is an even function, the odd-function part of the integral is zero, as reflected in Eq. (4-23). We calculate coefficients B_m by matching the patterns at integer values of U. The integral [Eq. (4-23)] is zero unless $U = m$:

$$B_0 = f(0) \qquad \frac{B_m}{2} = f(m) \qquad m = 1, 2, \ldots, \overline{n} - 1$$

Since we have only modified the location of the first $\overline{n} - 1$ zeros of the U-space pattern, $f(m) = 0$ for $m \ge \overline{n}$ and the Fourier cosine series has only $\overline{n}$ components:

$$E(x) = f(0) + 2 \sum_{m=1}^{\overline{n}-1} f(m) \cos 2m\pi x \tag{4-24}$$

The coefficients are given by

$$f(0) = 1$$

$$f(m) = \frac{(-1)^m \prod_{N=1}^{\overline{n}-1} \left(1 - m^2/U_N^2\right)}{-2 \prod_{N=1, N\neq m}^{\overline{n}-1} \left(1 - m^2/N^2\right)} \qquad m = 1, 2, \ldots, \overline{n} - 1 \tag{4-25}$$

Equation (4-19) computes the U-space pattern of the Taylor distribution but requires L'Hospital's rule at integer values of U. The finite number of coefficients B_m makes Eq. (4-23) more convenient since the integral is easily solved:

$$f(U) = B_0 \frac{\sin(\pi U)}{\pi U} + \frac{1}{2} \sum_{i=1}^{\overline{n}-1} B_i \left[\frac{\sin[\pi(U - i)]}{\pi(U - i)} + \frac{\sin[\pi(U + i)]}{\pi(U + i)} \right] \tag{4-26}$$

Example Design the Taylor line-source distribution with 30-dB maximum sidelobes and $\overline{n} = 6$.

We use Eq. (4-21) to calculate A:

$$b = 10^{30/20} = 31.6228$$

$$A = \frac{\cosh^{-1} b}{\pi} = 1.3200$$

We substitute this constant into Eq. (4-20) to compute the five $(\overline{n} - 1)$ nulls:

No.	1	2	3	4	5
Null U_N	1.4973	2.1195	2.9989	3.9680	4.9747

The first null value gives us the BW_{null} factor (1.4973). The null beamwidth has been increased almost 50% relative to the uniform distribution. The coefficients of the Fourier cosine aperture distribution are found from Eqs. (4-24) and (4-25) (Table 4-4).

Coefficients of the series are normalized so that the distribution is 1 at $x = 0$, and the amplitude distribution is found by plotting the Fourier cosine series. We calculate the U-space pattern by using Eq. (4-26). We calculate the half-power point and compare it to the uniform distribution to determine the HPBW factor, 1.2611. By using Eq. (4-6), we calculate ATL = 0.66 dB for the distribution.

The U-space plot (Figure 4-3) of the example above shows the 30-dB sidelobe level. The first sidelobe is at 30 dB, and lobes after that fall away from 30 dB. With a higher value of $\overline{n}$, the first unchanged zero, more sidelobes would be nearer 30 dB. The dashed curve gives the pattern of a uniform distribution. Notice that the inner five nulls have been shifted to lower the sidelobes. At the sixth null and higher, the Taylor distribution has the same nulls as the uniform distribution. The $\overline{n}$ distribution has a narrower beamwidth than the one-parameter distribution (Figure 4-2) and a higher taper efficiency of 0.66 dB versus 0.96 dB. Figure 4-4 shows the normalized aperture voltage for 30-dB-maximum sidelobe Taylor distributions. The one-parameter design produces a lower pedestal than the two $\overline{n}$ designs. The $\overline{n} = 20$ design voltage peaks as it approaches the edge. This peaks because the Taylor $\overline{n}$ distribution approximates the Dolph–Chebyshev array that peaks at the edge of the array.

The amplitude taper efficiency was calculated for a number of designs and is given in Table 4-5. The corresponding beamwidth factors are listed in Table 4-6 together with

TABLE 4-4 Fourier Cosine Series Coefficients for Taylor Distribution: 30 dB, $\overline{n} = 6$

No.	B_m	B_m Normalized	Function
1	1.0000	0.64672	1
2	0.5733	0.37074	$\cos 2\pi x$
3	−0.0284	−0.01838	$\cos 4\pi x$
4	−0.000213	−0.000138	$\cos 6\pi x$
5	0.005561	0.003597	$\cos 8\pi x$
6	−0.003929	−0.002541	$\cos 10\pi x$

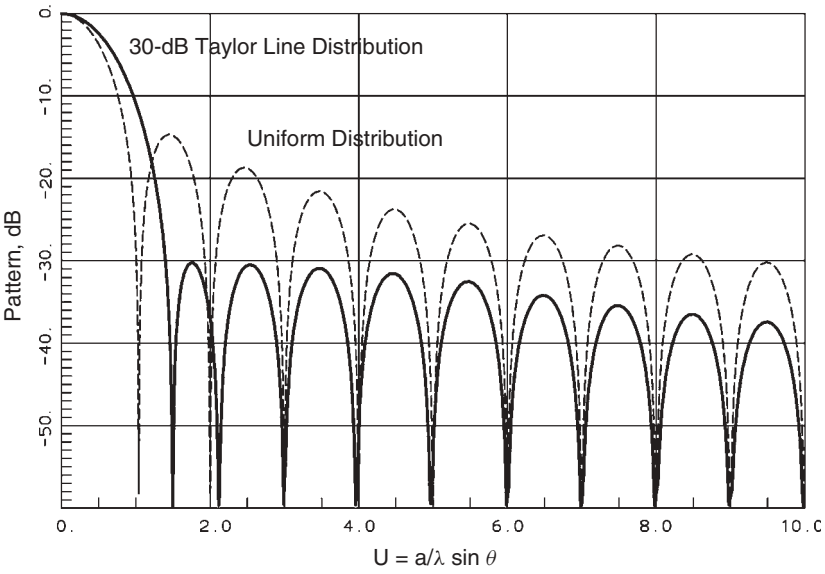


FIGURE 4-3 U-space pattern of 30-dB Taylor $\overline{n} = 6$ linear distribution versus uniform distribution.

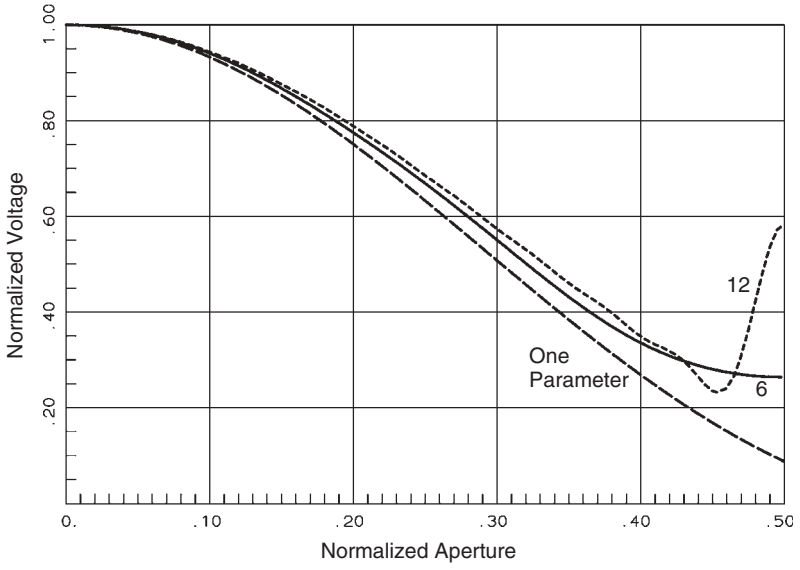


FIGURE 4-4 A 30-dB Taylor linear aperture distribution comparison.

the null beamwidth factors (location of first zero in U-space) in Table 4-7. ATL depends on the sidelobe level (Table 4-5) more than the number of modified zeros. Both the 20- and 25-dB sidelobe levels show that there is an optimum number of zeros. The edge of the distribution peaks toward the Dirac delta function and reduces the amplitude taper efficiency. More than three modified zeros are needed to reduce the sidelobes

TABLE 4-5 Amplitude Taper Losses of Taylor Line-Source Distributions

	Sidelobe Level (dB)						
$\overline{n}$	20	25	30	35	40	45	50
4	0.17	0.43	0.69	0.95			
5	0.15	0.41	0.68	0.93	1.16		
6	0.15	0.39	0.66	0.92	1.15	1.37	
7	0.15	0.37	0.65	0.91	1.15	1.36	1.56
8	0.16	0.36	0.63	0.90	1.14	1.36	1.55
10	0.19	0.34	0.61	0.88	1.13	1.35	1.55
12	0.24	0.34	0.59	0.86	1.11	1.34	1.54
16	0.35	0.35	0.57	0.84	1.09	1.32	1.53
20	0.46	0.27	0.56	0.82	1.07	1.30	1.51

TABLE 4-6 Beamwidth Factor of Taylor Line-Source Distribution

	Sidelobe Level (dB)						
$\overline{n}$	20	25	30	35	40	45	50
4	1.1043	1.1925	1.2696	1.3367			
5	1.0908	1.1837	1.2665	1.3404	1.4065		
6	1.0800	1.1752	1.2611	1.3388	1.4092	1.4733	
7	1.0715	1.1679	1.2555	1.3355	1.4086	1.4758	1.5377
8	1.0646	1.1617	1.2504	1.3317	1.4066	1.4758	1.5400
10	1.0545	1.1521	1.2419	1.3247	1.4015	1.4731	1.5401
12	1.0474	1.1452	1.2353	1.3189	1.3967	1.4695	1.5379
16	1.0381	1.1358	1.2262	1.3103	1.3889	1.4628	1.5326
20	1.0324	1.1299	1.2203	1.3044	1.3833	1.4576	1.5280

TABLE 4-7 Null Beamwidth Factor of Taylor Line-Source Distributions

	Sidelobe Level (dB)						
$\overline{n}$	20	25	30	35	40	45	50
4	1.1865	1.3497	1.5094	1.6636			
5	1.1696	1.3376	1.5049	1.6696	1.8302		
6	1.1566	1.3265	1.4973	1.6671	1.8347	1.9990	
7	1.1465	1.3172	1.4897	1.6632	1.8337	2.0031	2.1699
8	1.1386	1.3095	1.4828	1.6569	1.8306	2.0032	2.1739
10	1.1270	1.2978	1.4716	1.6471	1.8231	1.9990	2.1740
12	1.1189	1.2894	1.4632	1.6392	1.8161	1.9934	2.1705
16	1.1086	1.2783	1.4518	1.6277	1.8051	1.9835	2.1623
20	1.1023	1.2714	1.4444	1.6200	1.7975	1.9760	2.1553

below 40 dB; hence, the blanks represent unrealizable designs. The beamwidth factor (Table 4-6) reduces with increasing $\overline{n}$, but it depends mainly on the sidelobe level.

Example Compute beamwidths and ATL of an 8λ-wide aperture with $\overline{n} = 8$, 40-dB sidelobes, and a Taylor line-source distribution design.

From Table 4-6,

$$\text{HPBW} = \frac{1.4066(50.76^\circ)}{8} = 8.92^\circ$$

From Table 4-7,

$$\text{BW}_{\text{null}} = \frac{1.8306(114.59^\circ)}{8} = 26.22^\circ$$

From Table 4-5, ATL = 1.14 dB. A square aperture with the same distribution in both directions has

$$\text{directivity} = 10 \log \frac{4\pi A}{\lambda^2} - 2\text{ATL} = 26.77 \text{ dB}$$

4-5 TAYLOR LINE DISTRIBUTION WITH EDGE NULLS

Rhodes [4] has shown that it is impossible to have a step discontinuity of the fields at the edge of a physical aperture. Given the radius of curvature of the edge, ρ, the field varies as

$$E_d \sim \frac{C_2}{\rho} d \qquad \text{polarized perpendicular to the aperture edge}$$

$$E_s \sim \frac{C_1}{\rho} d^2 \qquad \text{polarized parallel to the aperture edge}$$

where C_1 and C_2 are constants and d is the distance from the edge. Without the possibility of an edge pedestal, a traditional Taylor line source cannot be realized with a physical aperture. We can sample the distribution with an array or closely approximate it, but we cannot achieve the exact distribution. A Taylor distribution with a null at the edge can be realized in an aperture.

Rhodes [5] extended the Taylor line source by modifying the U-space pattern zeros of the cosine distribution. Since $\alpha = 1$, the far-out sidelobes drop off as $1/U^2$ and the distribution is zero on the edges. The zeros of the cosine distribution occur at

$$(N + 1/2)\pi \qquad N = 1, 2, 3, \ldots \qquad k \text{ space}$$

When the Taylor U-space variable is used, the modified U-space pattern becomes

$$f(U) = \frac{\cos \pi U \prod_{N=1}^{\overline{n}-1} \left(1 - U^2/U_N^2\right)}{[1 - (2U)^2] \prod_{N=1}^{\overline{n}-1} (1 - U^2/(N + \frac{1}{2})^2)} \tag{4-27}$$

We remove the inner $\overline{n} - 1$ zeros at $N + \frac{1}{2}$ and substitute new ones given by

$$U_N = \pm(\overline{n} + \tfrac{1}{2}) \frac{\sqrt{A^2 + (N - \frac{1}{2})^2}}{\sqrt{A^2 + (\overline{n} - \frac{1}{2})^2}} \qquad N = 1, 2, \ldots, \overline{n} - 1 \tag{4-28}$$

When we compare Eqs. (4-28) and (4-20), we see that the nulls are shifted by $(\overline{n} + \frac{1}{2})/\overline{n}$ between the two Taylor distributions. When $\overline{n}$ is large, the nulls are close to the same for the two distributions.

To determine the amplitude distribution in the aperture, we expand the aperture fields in a Fourier cosine series,

$$E(x) = \sum B_m \cos(2m+1)\pi x \qquad |x| \le 0.5 \tag{4-29}$$

Like the Taylor line source, there are only $\overline{n}$ terms in the series whose coefficients are found by equating the pattern from the Fourier transform of Eq. (4-29) to Eq. (4-27). The coefficients are given by

$$B_0 = \frac{2\prod_{N=1}^{\overline{n}-1}\left(1 - \frac{1}{4}/U_N^2\right)}{\prod_{N=1}^{\overline{n}-1}\left[1 - \frac{1}{4}/(N+\frac{1}{2})^2\right]}$$

$$B_m = \frac{(-1)^m(m+\frac{1}{2})\prod_{N=1}^{\overline{n}-1}\left[1 - (m+\frac{1}{2})^2/U_N^2\right]}{[1-(2m+1)^2]\prod_{N=1,N\neq m}^{\overline{n}-1}\left[1-(m+\frac{1}{2})^2/(N+\frac{1}{2})^2\right]} \quad m = 1, 2, \ldots, \overline{n}-1 \tag{4-30}$$

The U-space pattern can be found using the coefficients B_m:

$$f(U) = C_0 \sum_{i=0}^{\overline{n}-1} B_i \left[\frac{\sin[\pi(U-i-\frac{1}{2})]}{\pi(U-i-\frac{1}{2})} + \frac{\sin[\pi(U+i+\frac{1}{2})]}{\pi(U+i+\frac{1}{2})}\right]$$
$$C_0 = 2\sum_{i=0}^{\overline{n}-1} B_i \frac{\sin[\pi(i+\frac{1}{2})]}{\pi(i+\frac{1}{2})} \tag{4-31}$$

Example Design the Taylor line-source distribution with edge nulls for 30-dB maximum sidelobes and $\overline{n} = 6$.

We use Eq. (4-21) to calculate A:

$$b = 10^{30/20} = 31.6228$$
$$A = \frac{\cosh^{-1} b}{\pi} = 1.32$$

the same as the pedestal edge Taylor line-source distribution. We substitute this constant into Eq. (4-28) to compute the five modified nulls:

No.	1	2	3	4	5
Null U_N	1.6221	2.2962	3.2488	4.2987	5.3892

The null locations have increased by $(\overline{n} + \frac{1}{2})/\overline{n} = 6.5/6 = 1.0833$ from the pedestal Taylor line-source design. The null beamwidth factor has also increased by this factor

as well. The coefficients of the Fourier cosine aperture distribution are found from Eq. (4-30) Table 4-8.

The normalized coefficients sum to 1 at $x = 0$. Equation (4-27) determines the U-space pattern given the nulls. On finding the half-power point and comparing it with the uniform distribution half-power point, we compute the beamwidth factor: 1.3581. Tables 4-9 to 4-11 give results for this Taylor line source. As $\overline{n}$ increases, the results approach the result of the pedestal Taylor line source. Since the maximum sidelobe of the cosine distribution is 23 dB, a distribution must have peaking toward the edges to raise the sidelobes above that level. In all distributions the voltage approaches zero linearly at the edges.

TABLE 4-8 Fourier Cosine Series Coefficients for Taylor Distribution with Edge Nulls: 30 dB, $\overline{n} = 6$

No.	B_m	B_m Normalized	Function
1	0.50265	0.94725	$\cos \pi x$
2	0.023087	0.04351	$\cos 3\pi x$
3	0.017828	0.02220	$\cos 5\pi x$
4	−0.010101	−0.02075	$\cos 7\pi x$
5	0.007374	0.01390	$\cos 9\pi x$
6	−0.003245	−0.006116	$\cos 11\pi x$

TABLE 4-9 Amplitude Taper Losses of a Taylor Line Source with Edge Null Distribution

	Sidelobe Level (dB)					
$\overline{n}$	25	30	35	40	45	50
4	0.86	1.13	1.36	1.55	1.71	1.84
6	0.67	0.97	1.24	1.47	1.66	1.84
8	0.56	0.87	1.14	1.39	1.60	1.79
12	0.45	0.74	1.02	1.28	1.51	1.71
16	0.41	0.68	0.96	1.22	1.45	1.66
20	0.39	0.64	0.92	1.17	1.41	1.62

TABLE 4-10 Beamwidth Factor of a Taylor Line Source with Edge Null Distribution

	Sidelobe Level (dB)					
$\overline{n}$	25	30	35	40	45	50
4	1.3559	1.4092	1.4815	1.5443	1.5991	1.6470
6	1.2666	1.3581	1.4407	1.5153	1.5831	1.6448
8	1.2308	1.3242	1.4097	1.4882	1.5608	1.6280
12	1.1914	1.2850	1.3716	1.4522	1.5276	1.5984
16	1.1705	1.2635	1.3500	1.4308	1.5068	1.5785
20	1.1576	1.2502	1.3363	1.4170	1.4930	1.5649

TABLE 4-11 Null Beamwidth Factor of a Taylor Line Source with Edge Null Distribution

	Sidelobe Level (dB)					
$\overline{n}$	25	30	35	40	45	50
4	1.5184	1.6980	1.8715	2.0374	2.1949	2.3433
6	1.4371	1.6221	1.8060	1.9875	2.1656	2.3395
8	1.3913	1.5755	1.7604	1.9450	2.1284	2.3097
12	1.3431	1.5242	1.7075	1.8918	2.0765	2.2610
16	1.3182	1.4971	1.6786	1.8616	2.0455	2.2298
20	1.3031	1.4805	1.6605	1.8424	2.0254	2.2091

4-6 ELLIOTT'S METHOD FOR MODIFIED TAYLOR DISTRIBUTION AND ARBITRARY SIDELOBES [6, pp. 162–165]

Elliott's method separates the distribution nulls into right- and left-hand values in U-space that allows different sidelobe levels in the two regions. By applying a differential expression, the null positions in U-space can be found from the solution of a set of linear equations to produce designs with arbitrary sidelobes. Consider Eq. (4-19) and factor the null location term:

$$1 - \frac{U^2}{U_N^2} = \left(1 + \frac{U}{U_N}\right)\left(1 - \frac{U}{U_N}\right) = \left(1 + \frac{U}{U_{NL}}\right)\left(1 - \frac{U}{U_{NR}}\right)$$

We associate U_{NL} with a nulls on the left side of the origin and U_{NR} with the right side or a positive pattern angle. If we also separate the term in the denominator of Eq. (4-19), we can independently pick the number of nulls to be moved on either side of the pattern:

$$F(U) = C_0 \frac{\sin \pi U \prod_{N=1}^{\overline{n}_L - 1} (1 + U/U_N) \prod_{N=1}^{\overline{n}_R - 1} (1 - U/U_N)}{\pi U \prod_{N=1}^{\overline{n}_L - 1} (1 + U/N) \prod_{N=1}^{\overline{n}_R - 1} (1 - U/N)} \tag{4-32}$$

Equation (4-32) allows different Taylor distributions on the two sides. We add a normalization factor C_0 when we use different distributions. The pattern peak will shift off zero for unbalanced distributions. Since the two sides are not independent, a simple selection of the two levels will not produce the desired sidelobes. Table 4-12 lists the U-space locations of the pattern peaks and sidelobe level for a design with 35- and 30-dB sidelobes. The left distribution lowered the sidelobes on the right and the right one raised the left sidelobes. A few manual iterations produced suitable left and right distributions to give the desired sidelobes. The main beam shifts a little bit. A linear progressive phase shift across the aperture can shift the pattern to broadside.

We expand the aperture distribution in a complex exponential series similar to Eq. (4-22):

$$E(x) = \sum_{i=-\overline{n}_L+1}^{\overline{n}_R - 1} B_i e^{-j2\pi i x} \qquad |x| \le 0.5 \tag{4-33}$$

TABLE 4-12 Modified Taylor Distribution Sidelobes for Independent Left and Right Sidelobe Design Using $\overline{n} = 6$ for Both Sides

Left-Side−35-dB U-Space	Right Side−30-dB Sidelobe (dB)	Left-Side−36-dB U-Space	Right-Side−28.6-dB Sidelobe (dB)
−5.4849	−35.70	−5.4874	−36.00
−4.4905	−35.06	−4.4976	−35.44
−3.5275	−34.66	−3.5399	−35.11
−2.6313	−34.44	−2.6510	−34.97
−1.8997	−34.39	−1.9293	−34.99
−0.0511	0	−0.0758	0
1.7546	−31.05	1.7141	−30.05
2.5372	−31.45	2.5119	−30.55
3.4697	−32.00	3.4544	−31.19
4.4584	−32.74	4.4498	−32.02
5.4711	−33.75	5.4676	−33.13

We calculate the coefficients by the same method used for Eq. (4-25):

$$B(0) = 1$$

$$B(m) = \frac{(-1)^{|m|} \prod_{N=1}^{\overline{n}_L - 1} (1 + m/U_N) \prod_{N=1}^{\overline{n}_R - 1} (1 - m/U_N)}{-\prod_{N=1, N\neq m}^{\overline{n}_L - 1} (1 + m/N) \prod_{N=1, N\neq m}^{\overline{n}_R - 1} (1 - m/N)}$$

$$m = -\overline{n}_L + 1, \ldots, -1, 1, 2, \ldots, \overline{n}_R - 1 \qquad (4\text{-}34)$$

We derive the pattern from the integral of the finite complex exponential:

$$f(U) = C_0 \sum_{i=-\overline{n}_L+1}^{\overline{n}_R - 1} B_i \frac{\sin[\pi(U - i)]}{\pi(U - i)} \qquad (4\text{-}35)$$

We include the normalization factor C_0 for unequal left and right sidelobes.

We control the sidelobes by adjusting the location of the nulls in the U-space pattern. We can iterate the null positions to produce individually selected sidelobes. The peak of each sidelobe given in Table 4-12 was found through a one-dimensional search between pairs of nulls. A search based on the Fibonacci numbers [7, p. 280] computes the peak with the minimum number of evaluations of the pattern using Eq. (4-35). We denote the pattern peaks by U_m^p starting with the peak between $-\overline{n}_L$ and $U_{-\overline{n}_L+1}$, those between nulls, the peak near 0, and the last peak between $U_{\overline{n}_R-1}$ and $\overline{n}_R$ for $\overline{n}_L + \overline{n}_R - 1$.

We adjust the U-space nulls by the differentials δU_N found from the solution of a matrix equation. The terms of the matrix are the differential term of a Taylor series expansion of the numerator of Eq. (4-32) evaluated at the pattern peaks:

$$a_{m,n} = \frac{U_m^p / U_N^2}{1 - U_m^p / U_N} \qquad N = -\overline{n}_L + 1, \ldots, -1, 1, \ldots, \overline{n}_R - 1$$

$$a_{m,0} = 1 \qquad (4\text{-}36)$$

The vector of differential nulls is

$$\delta \mathbf{U} = [\delta U_{-\overline{n}_L+1}, \ldots, \delta U_{-1}, \delta C/C_0, \delta U_1 \ldots, \delta U_{\overline{n}_R-1}]^{\mathrm{T}}$$

where $\delta C/C_0$ is the change in the pattern normalization. We form a vector using the ratio of the desired pattern peak $f_d(U)$ to the actual pattern $f_a(U)$ with terms $[f_d(U_m^p)/f_a(U_m^p)] - 1$. We solve the matrix equation for the null shifts:

$$[a_{m,n}]^{-1} \left[\frac{f_d(U_m^p)}{f_a(U_m^p)} - 1 \right]^{\mathrm{T}} = [\delta U_N] \tag{4-37}$$

We calculate new distribution nulls $U_N + \delta U_N$, substitute the new nulls into Eq. (4-34) to determine the new expansion coefficients B_m, and evaluate the pattern using Eq. (4-35) between the new nulls to compute new pattern peaks. We iterate the process until the sidelobe levels are satisfactory. Notice that $f(U)$ is a voltage.

The Taylor linear distribution produces a pattern with only approximately equal sidelobes. Table 4-13 lists the iteration to produce a distribution with a pattern that has five 30-dB sidelobes. The solution starts with a 30-dB, $\overline{n} = 6$ Taylor distribution. In two iterations the method found a distribution with exactly the desired sidelobes. Table 4-14 gives the results of repeating the example of Table 4-12 of the design for 35- and 30-dB sidelobes. This method can produce a linear distribution with individually

TABLE 4-13 Iteration of Distribution Nulls for a Pattern with 30-dB Sidelobes

Taylor Distribution			First Iteration			Second Iteration		
Null	U-Space	Sidelobe	Null	U-Space	Sidelobe	Null	U-Space	Sidelobe
1.4973	1.7557	−30.22	1.4708	1.7258	−30.00	1.4729	1.7284	−30.00
2.1195	2.5387	−30.46	2.0827	2.4987	−29.99	2.0859	2.5027	−30.00
2.9989	3.4709	−30.89	2.9490	3.4215	−29.96	2.9541	3.4274	−30.00
3.9680	4.4591	−31.53	3.9075	4.4072	−29.86	3.9152	4.4147	−30.00
4.9747	5.4718	−32.48	4.9145	5.4424	−29.63	4.9242	5.4471	−29.99

TABLE 4-14 Iteration for 35- and 30-dB Sidelobes in Linear Distribution

Left-Side−36-dB	Right-Side−28.6-dB		Second Iteration		
U-Space	Sidelobe (dB)	Null	U-Space	Sidelobe (dB)	Null
−5.4874	−36.00	−4.9964	−5.4798	−35.00	−4.9778
−4.4976	−35.44	−4.0169	−4.4859	−35.00	−4.0043
−3.5399	−35.11	−3.0845	−3.5348	−35.00	−3.0829
−2.6510	−34.97	−2.2583	−2.6555	−35.00	−2.2670
−1.9293	−34.99	−1.7008	−1.9410	−35.00	−1.7143
−0.0758	0		−0.1037	0	
1.7141	−30.05	1.4495	1.6615	−30.00	1.4023
2.5119	−30.55	2.0883	2.4475	−30.00	2.0249
3.4544	−31.19	2.9801	3.3839	−30.00	2.9049
4.4498	−32.02	3.9574	4.3839	−30.00	3.8780
5.4676	−33.13	4.9699	5.4313	−29.98	4.9001

selected sidelobes. It may be necessary to design an intermediate distribution if the change in sidelobes is too great for the simple iteration scheme to converge. We will design linear arrays by a similar technique of manipulating pattern nulls to produce arbitrary sidelobes.

4-7 BAYLISS LINE-SOURCE DISTRIBUTION [8]

The Bayliss distribution produces a pattern null on a boresight while controlling the height of the sidelobes. The second dashed curve in Figure 4-5 below is a Bayliss difference pattern also designed to give 30-dB sidelobes when combined with the Taylor distribution. As in the Taylor distribution, the first few sidelobes are nearly the same height, to minimize the beamwidth of the two beams split about the boresight.

A monopulse tracking system uses an auxiliary pattern with a boresight null coincident with the beam peak of the main pattern. The tracking system drives the antenna positioner until the signal in this difference channel nulls so that the main channel (sum) points at the emitter or radar target. The accuracy of the pointing angle is improved, since a null is a more exact direction than the broad sum pattern peak. Noise and receiver sensitivity, along with the slope of the difference pattern, limit the tracking accuracy. Stronger signals can be tracked farther into the null. Because the phase of a pattern shifts by 180° when passing through a null, phase relative to the sum pattern (a reference signal) can be used to give direction. Without monopulse or some other sequential lobing technique, such as conical scan, radar cannot track effectively.

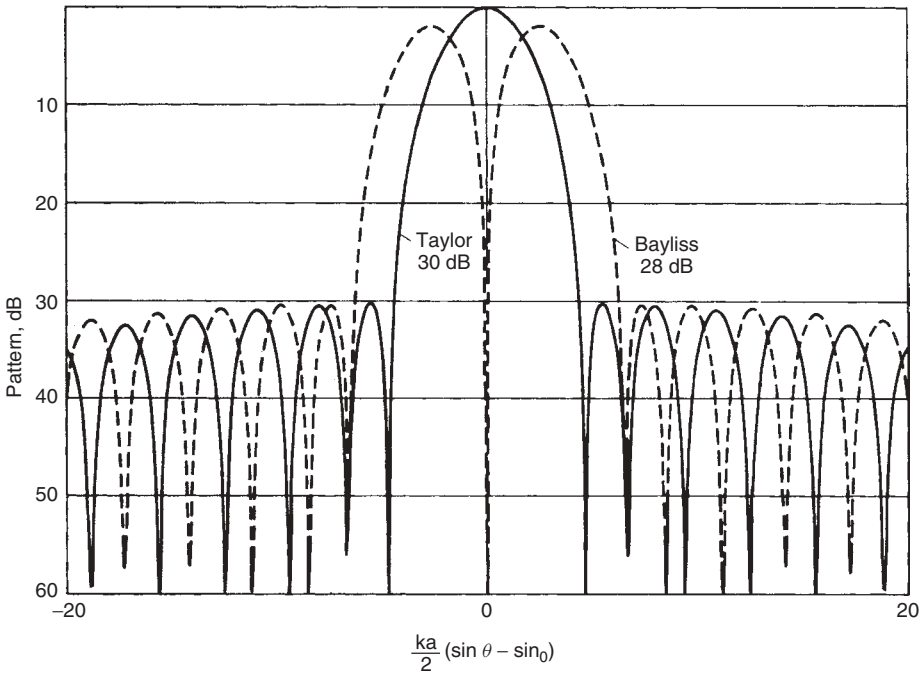


FIGURE 4-5 Taylor and Bayliss line distributions to give 30-dB sidelobes ($\overline{n} = 6$).

Any odd-function distribution produces a null on a boresight. A uniform distribution that switches phase by 180° in the center has the best amplitude taper efficiency but high sidelobes (10 dB). These high sidelobes allow interfering or noise signals to enter the receiver. The Bayliss distribution adjusts the inner nulls of the U-space pattern to lower the sidelobes. Adjusting the zeros to correspond to the Dolph–Chebyshev array does not lower the sidelobes to the same level as it did in the Taylor distribution. Further adjustments of the four inner zeros are required. Bayliss found the proper location through a computer search. We locate the zeros by

$$U_N = \begin{cases} (\overline{n} + \frac{1}{2})\sqrt{\dfrac{\xi_N^2}{A^2 + \overline{n}^2}} & N = 1, 2, 3, 4 \\ (\overline{n} + \frac{1}{2})\sqrt{\dfrac{A^2 + N^2}{A^2 + \overline{n}^2}} & N = 5, 6, \ldots, \overline{n} - 1 \end{cases} \tag{4-38}$$

By using the U-space pattern, we have

$$f(U) = U \cos \pi U \frac{\prod_{N=1}^{\overline{n}-1} \left(1 - U^2/U_N^2\right)}{\prod_{N=0}^{\overline{n}-1} \left[1 - U^2/(N + \frac{1}{2})^2\right]} \tag{4-39}$$

The coefficients were fitted to polynomials depending on the sidelobe level. Given $S = |\text{sidelobe level(dB)}|$:

$$A = 0.3038753 + S(0.05042922 + S(-0.00027989 + S(0.343 \times 10^{-5} - S(0.2 \times 10^{-7})))) \tag{4-40a}$$

$$\xi_1 = 0.9858302 + S(0.0333885 + S(0.00014064 + S(-0.19 \times 10^{-5} + S(0.1 \times 10^{-7})))) \tag{4-40b}$$

$$\xi_2 = 2.00337487 + S(0.01141548 + S(0.0004159 + S(-0.373 \times 10^{-5} + S(0.1 \times 10^{-7})))) \tag{4-40c}$$

$$\xi_3 = 3.00636321 + S(0.00683394 + S(0.00029281 + S(-0.161 \times 10^{-5}))) \tag{4-40d}$$

$$\xi_4 = 4.00518423 + S(0.00501795 + S(0.00021735 + S(-0.88 \times 10^{-6}))) \tag{4-40e}$$

The location of the pattern peak was also fitted to a polynomial:

$$U_{\max} = 0.4797212 + S(0.01456692 + S(-0.00018739 + S(0.218 \times 10^{-5} + S(-0.1 \times 10^{-7})))) \tag{4-41}$$

We obtain the aperture distribution by a Fourier sine series having only $\overline{n}$ terms:

$$E(x) = \sum B_m \sin(m + \tfrac{1}{2})2\pi x \qquad |x| \le 0.5 \tag{4-42}$$

where

$$B_m = \frac{(-1)^m (m+\frac{1}{2})^2}{2j} \frac{\prod_{N=1}^{\overline{n}-1} \left[1-(m+\frac{1}{2})^2/U_N^2\right]}{\prod_{N=0, N\neq m}^{\overline{n}-1} \left[1-(m+\frac{1}{2})^2/(N+\frac{1}{2})^2\right]} \tag{4-43}$$

The phase constant $(-j)$ has little effect on the coefficient B_m except to balance the phase $\pm 90^\circ$ about the null.

Example Design a Bayliss distribution with 30-dB sidelobes and $\overline{n} = 6$.

Use Eq. (4-40) to compute the coefficients.

$$A = 1.64126 \qquad \xi_1 = 2.07086 \qquad \xi_2 = 2.62754 \qquad \xi_3 = 3.43144 \qquad \xi_4 = 4.32758$$

We substitute these constants into Eq. (4-38) to calculate the five $(\overline{n}-1)$ nulls:

No.	1	2	3	4	5
Null U_N	2.1639	2.7456	3.5857	4.5221	5.4990

Equation (4-41) computes the beam peak of the split-beam pattern in U space:

$$U_{\max} = 0.7988 \qquad \frac{ka}{2} \sin\theta_{\max} = \pi U_{\max} = 2.5096$$

where a is the aperture width. We substitute these zeros into Eq. (4-43) to determine the coefficients of the Fourier sine series of the aperture distribution (Table 4-15). By evaluating the series across the aperture, the coefficients can be normalized to give a maximum aperture voltage of 1.

We use Eq. (4-39), after substituting the zeros, to evaluate the pattern. The 3-dB pattern points can be found by searching the pattern:

$$\frac{ka}{2} \sin\theta_1 = 1.27232 \qquad \frac{ka}{2} \sin\theta_2 = 4.10145$$

Figure 4-5 contains the plot of a Bayliss distribution $(\overline{n} = 6)$ designed to have sidelobes 30 dB below the Taylor distribution with 30-dB sidelobes. The losses to the difference pattern are about 2 dB higher than the sum pattern. We design the Bayliss

TABLE 4-15 Fourier Cosine Series Coefficients for Bayliss Distribution: 30 dB, $\overline{n} = 6$

No.	B_m	B_m Normalized	Function
0	0.13421	0.85753	$\sin \pi x$
1	0.081025	0.51769	$\sin 3\pi x$
2	−0.0044151	−0.028209	$\sin 5\pi x$
3	0.001447	0.0092453	$\sin 7\pi x$
4	−0.0003393	−0.0021679	$\sin 9\pi x$
5	−0.000014077	−0.00008994	$\sin 11\pi x$

distribution to have 28-dB sidelobes. If designed for 30-dB sidelobes as in the example above, then, relative to the sum Taylor distribution, the sidelobes would be 32 dB down from the sum pattern peak. The last nulls show that the unmodified zeros of the Taylor distribution occur at $\pm n\pi$, whereas the unmodified zeros of the Bayliss distribution occur at $\pm(n + \frac{1}{2})\pi$.

By using Eq. (4-6), we calculate amplitude taper efficiency of the pattern at the beam peak. When we evaluate the phase error efficiency by using Eq. (4-7), the result is zero because of the boresight null. We use Eq. (4-3) to evaluate the phase error efficiency at the beam peak:

$$\text{PEL} = \frac{\left| \int_{-a/2}^{a/2} E(x) e^{jk \sin \theta_{\max} x} \, dx \right|^2}{\left[\int_{-a/2}^{a/2} |E(x)| \, dx \right]^2} \tag{4-44}$$

Table 4-16 lists results of calculations on Bayliss distributions with $\overline{n} = 10$ for various sidelobe levels. Lower sidelobe levels produce higher distribution losses and push the beam peak out. The position of the beam peak is independent of $\overline{n}$, since the first four zeros are fixed by Eq. (4-40). Like the Taylor distribution, the sidelobe level determines most of the parameters of the Bayliss distribution. Changing $\overline{n}$ has less effect than it has for the Taylor distribution. The values of parameters for distribution with $\overline{n} \neq 10$ will differ little from those in Table 4-16.

Example Compute the beam peak and beam edges for an 8λ-wide aperture excited in a Bayliss distribution with $\overline{n} = 10$ and 30-dB sidelobes.

$$\frac{2\pi}{\lambda} \sin \theta_{\max} \frac{8\lambda}{2} = 2.5096$$

$$\sin \theta_{\max} = \frac{2.5096}{8\pi}$$

$$\sin \theta_1 = \frac{1.263}{8\pi} \qquad \sin \theta_2 = \frac{4.071}{8\pi}$$

$$\theta_{\max} = 5.73^\circ \qquad \theta_1 = 2.88^\circ \qquad \theta_2 = 9.32^\circ$$

TABLE 4-16 Characteristics of a Bayliss Line-Source Distribution with $\overline{n} = 10$ Parameters

Sidelobe Level (dB)	Beam Peak, $ka/2 \sin \theta_{\max}$	3-dB Edge		ATL (dB)	PEL (dB)
		$ka/2 \sin \theta_1$	$ka/2 \sin \theta_2$		
20	2.2366	1.140	3.620	0.50	1.81
25	2.3780	1.204	3.855	0.54	1.90
30	2.5096	1.263	4.071	0.69	1.96
35	2.6341	1.318	4.270	0.85	2.01
40	2.7536	1.369	4.455	1.00	2.04

4-8 WOODWARD LINE-SOURCE SYNTHESIS [9]

In the preceding sections, methods to determine distributions that give the minimum beamwidth for specified sidelobe levels were discussed. Some applications require shaped beams extending over a range of angles. The Woodward synthesis samples the desired k-space pattern at even intervals to determine the aperture distribution. No integrals are required to compute coefficients.

The technique is based on the scanned pattern of a uniform amplitude distribution. Express the pattern in terms of U-space so that when scanned to U_0, it becomes

$$\frac{\sin \pi(U - U_0)}{\pi(U - U_0)}$$

with the nulls of the pattern occurring at integer values of $U - U_0$.

$$U = \frac{a}{\lambda} \sin\theta \qquad U_0 = \frac{a}{\lambda} \sin\theta_0$$

The visible region extends between $+a$ and $-a$, centered about U_0.

Figure 4-6 shows two patterns, scanned to $U_0 = 1$ and $U_0 = 2$. The peak of the curve scanned at $U_0 = 2$ occurs at one of the nulls of the pattern scanned to $U_0 = 1$. If we allow only integer values of U_0, the pattern scanned to U_0 solely determines the pattern at the point U_0 in U-space. The two curves (Figure 4-6) in the regions below $U = 0$ and above $U = 3$ cancel each other to some extent when the distributions are

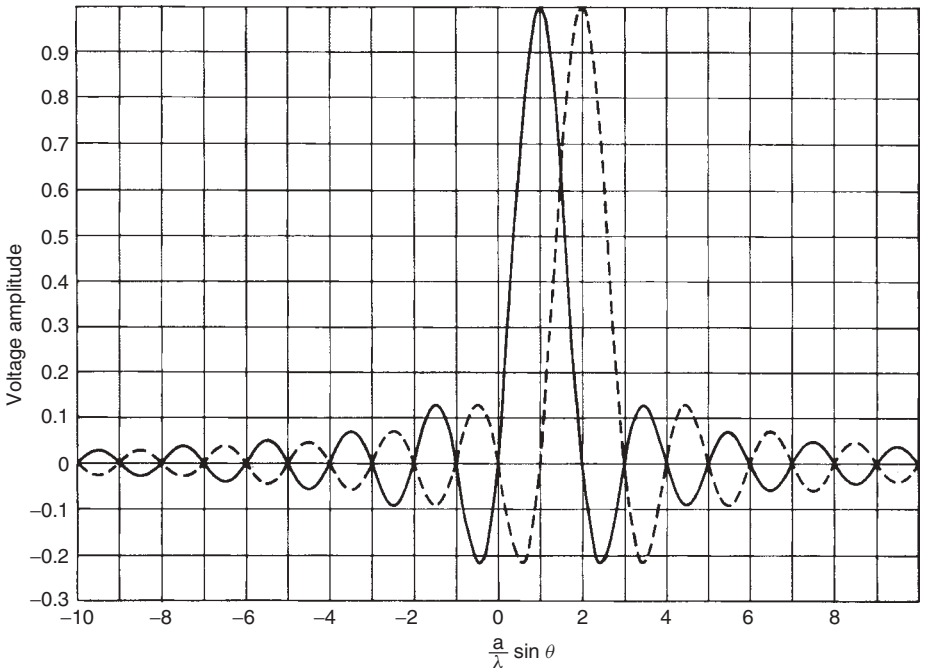


FIGURE 4-6 Scanned uniform distributions: $U_0 = 1$ and $U_0 = 2$.

added. We form the aperture distribution from a sum of $2a/\lambda + 1$ independent sample points of scanned apertures:

$$E(x) = \sum_{i=-N}^{N} E_i e^{-j(i/a)x} \tag{4-45}$$

where $N =$ integer part (a/λ). Each term is a uniform amplitude distribution scanned to an integer value of U. The amplitudes E_i are determined by the sample values of the U-space pattern at those points.

Example Design a 10λ aperture with a constant beam between $\theta = 0^\circ$ and $\theta = 30^\circ$. The nonzero portion of the U-space pattern extends from $U_1 = 10 \sin 0^\circ = 0$ to $U_2 = 10 \sin 30^\circ = 5$.

When we sample the U-space pattern, we discover six nonzero terms:

i	0	1	2	3	4	5
E_i	0.5	1.0	1.0	1.0	1.0	0.5

At $U_1 = 0$ and $U_2 = 5$, we use the average value. The aperture distribution is

$$0.5 + e^{-jx/a} + e^{-j2x/a} + e^{-j3x/a} + e^{-j4x/a} + 0.5e^{-j5x/a}$$

The U-space pattern of this distribution (Figure 4-7) shows some ripple in the beam and the reduction to 6 dB at the beam edges. If we increased the sample level at the edges, $U = 0$ and $U = 5$, the pattern would increase to that level.

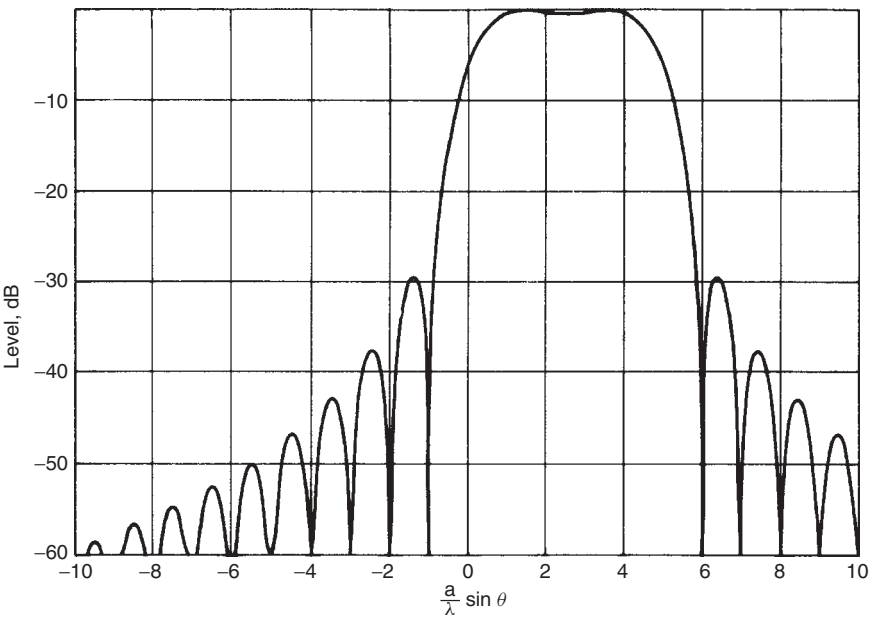


FIGURE 4-7 U-space pattern of Woodward–Lawson sampling for constant beam from 0 to 30° (10λ aperture).

A cosecant-squared power pattern can be designed by the same method as in the preceding example. When an antenna with this pattern on the ground points its maximum toward the horizon, it delivers a constant signal to an aircraft that maintains a constant altitude. The pattern falloff matches the range decrease as the aircraft flies toward the antenna. The voltage pattern is given by

$$E = E_0 \frac{\sin \theta_{\max}}{\sin \theta}$$

where $\theta_{\max}$ is the angle of the pattern maximum. In U-space this becomes

$$E(U) = E_0 \frac{U_m}{U}$$

The amplitudes of the scanned apertures decrease as $1/U$.

Example Design a 10λ aperture with a cosecant-squared pattern from $\theta = 5^\circ$ to $\theta = 70^\circ$ with the maximum at 5°.

There are $2a/\lambda + 1$ possible sample points (21). The nonzero portion of the U-space pattern extends from $U_{\min} = 10 \sin 5^\circ = 0.87$ to $U_{\max} = 10 \sin 70^\circ = 9.4$. We sample only at integer values of U, which gives us nine nonzero terms: $U_m = 0.8716$. The coefficients are given in Table 4-17.

The sum [Eq. (4-45)] for this distribution contains nine terms.

$$E(x) = \sum_{i=1}^{9} E_i e^{-j(i/a)x}$$

Figure 4-8 shows the amplitude and phase of this aperture distribution. The pattern obtained by summing the scanned aperture distributions (Figure 4-9) shows ripple about the desired pattern. Increasing the aperture size increases the number but does not change the level of ripples. The aperture distribution (Figure 4-8) has a negative phase slope to scan the beam off broadside.

4-9 SCHELKUNOFF'S UNIT-CIRCLE METHOD [10]

Schelkunoff's unit-circle method consists of the manipulation of the zeros (nulls) of the array pattern to achieve a desired pattern for a line array. The method is similar to designing networks by specifying the placement of poles and zeros in the complex plane, but the array has only zeros to manipulate. We can use the representation to describe any uniformly spaced array.

TABLE 4-17 Woodward Synthesis Coefficients of 10λ Cosecant-Squared Pattern

i	E_i	i	E_i	i	E_i
1	0.8716	4	0.2179	7	0.1245
2	0.4358	5	0.1743	8	0.1089
3	0.2905	6	0.1453	9	0.0968

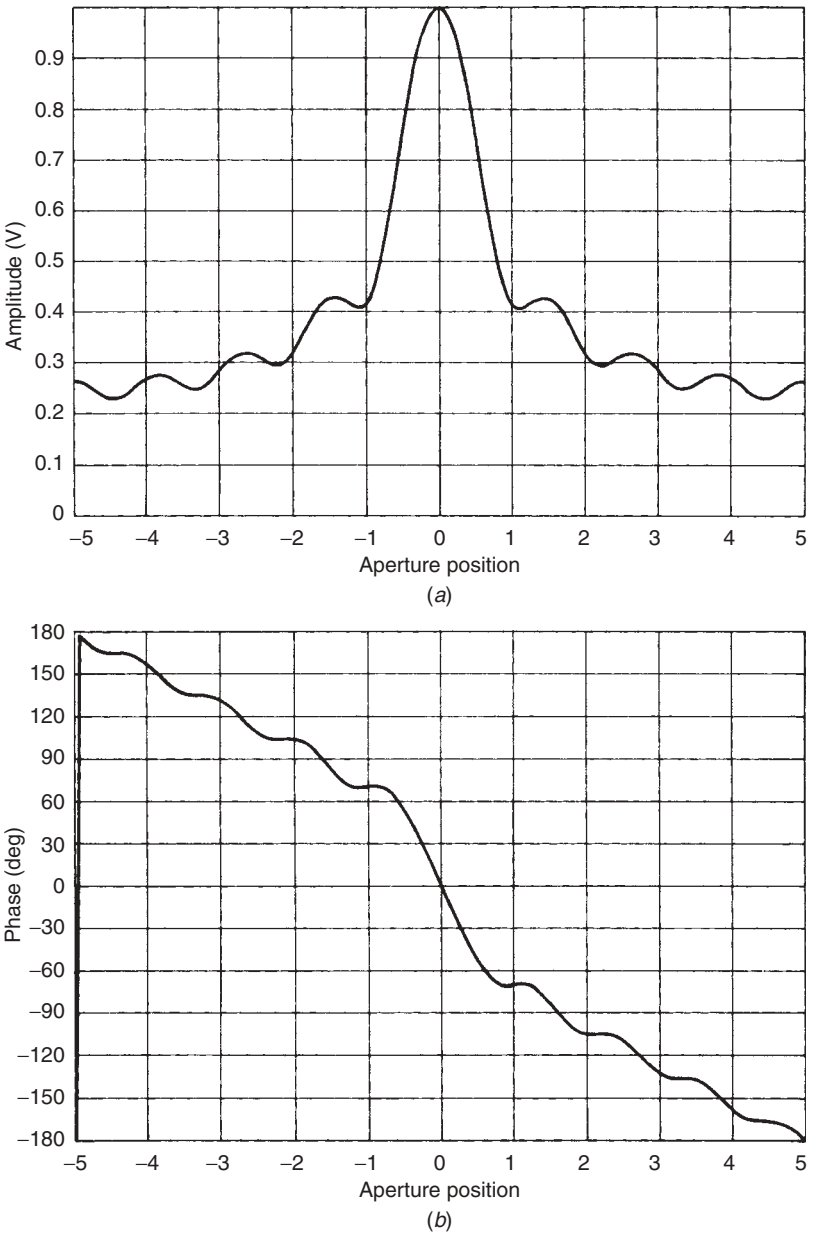


FIGURE 4-8 Aperture distribution of Woodward–Lawson sampling for cosecant-squared pattern (10λ aperture): (*a*) aperture amplitude distribution; (*b*) aperture phase distribution.

Consider a uniformly spaced array along the z-axis with the pattern angle θ measured from the axis. The array response will be symmetrical about the z-axis. If we define the variable $\psi = kd\cos\theta + \delta$, where δ is a fixed progressive phase shift between elements, d the element spacing, and k the wave number $(2\pi/\lambda)$, the pattern of the array is given by

$$E = I_0 + I_1 e^{j\psi} + I_2 e^{j2\psi} + I_3 e^{j3\psi} \cdots \tag{4-46}$$

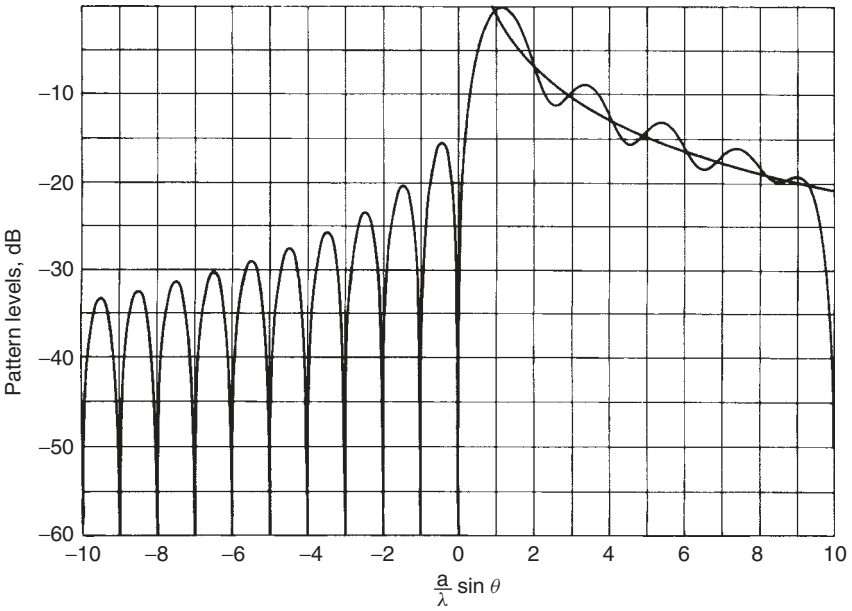


FIGURE 4-9 U-space pattern of 10λ aperture Woodward–Lawson sampling for a cosecant-squared pattern.

where I_i, a phasor, is the excitation of the ith element in the array. We simplify the notation further by defining

$$W = e^{j\psi} \tag{4-47}$$

We then write Eq. (4-46) as

$$E = I_0 + I_1 W + I_2 W^2 + I_3 W^3 + \cdots + I_{N-1} W^{N-1} \tag{4-48}$$

where N is the number of elements in the array. We use the first element as our phase reference point. This array factor (isotropic elements) is a polynomial with $N - 1$ roots (zeros) for N elements.

We denote the roots as W_i and rewrite Eq. (4-48) as

$$E = E_0(W - W_1)(W - W_2)\cdots(W - W_{N-1})$$

We can ignore the normalization E_0 and compute array pattern magnitude as

$$|E(W)| = |W - W_1||W - W_2|\cdots|W - W_{N-1}|$$

where $|W - W_i|$ is the distance from the root W_i to W in the complex plane. W is limited to the unit circle [Eq. (4-47)] because it always has unit value. Both the spacing of the elements and the progressive phase shift δ determine the limits of the phase of W:

$$\begin{aligned} &\theta = 0^\circ \qquad \psi_s = kd + \delta \qquad \text{start} \\ &\theta = 180^\circ \qquad \psi_f = -kd + \delta \qquad \text{finish} \end{aligned} \tag{4-49}$$

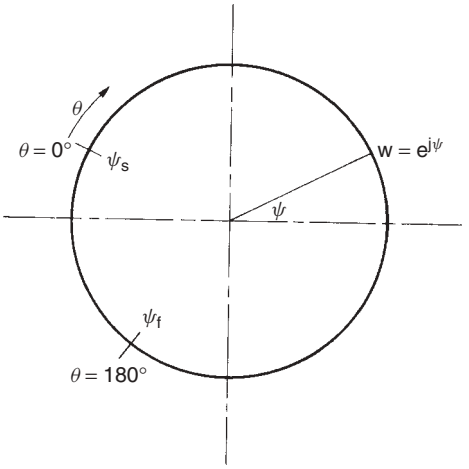


FIGURE 4-10 Unit circle in the W-plane.

As θ increases, ψ decreases and W progresses in a clockwise rotation along the unit circle (Figure 4-10). We have no 2π limitation on either ψ_s or ψ_f. The element spacing determines the number of times W cycles the unit circle as θ varies from 0 to 180°. If $\psi_s - \psi_f$, $2kd$, exceeds 2π, there is a possibility of more than one main beam (grating lobes).

The zeros W_i, suppress the pattern when W moves close to one or more of them. The pattern rises to form a lobe when W is far from the zeros. The main-beam peak occurs at the point with the maximum product of the distances from the zeros. Whenever W passes through that point, another main beam forms. A uniformly fed array has the W-space polynomial

$$f(W) = \frac{1 - W^N}{1 - W} \qquad \text{for } N \text{ elements}$$

The zeros of $f(W)$ are the N zeros of $W^N = 1$ with the zero at $W = 1$ removed: $W_i = e^{j2\pi i/N}$. These are spaced uniformly on the unit circle.

Figure 4-11 shows the unit circle diagram of a 10-element array fed with uniform phase and amplitude. W starts at -1 since $d = \lambda/2$, and it progresses clockwise around the unit circle one revolution to the same point as θ varies from 0 to 180°. At $\theta = 90°$, the product of the distances from the zeros is a maximum. A lobe forms within the space between each pair of zeros. As W moves from the start to the main beam at $W = 1$, it starts at a zero and passes through four additional zeros. These zeros W_i correspond to the nulls in the pattern from $\theta = 0°$ to $\theta = 90°$. An equal number of nulls occur as W moves through the range $\theta = 90$ to 180°.

A uniform-amplitude end-fire array can be represented on the same unit-circle diagram. With antenna elements spaced $\lambda/4$, the excursion from start ψ_s *to* finish ψ_f *is* only $\pi(2kd)$. A progressive phase shift δ of $-kd$ through the array forms an end-fire pattern. From Eq. (4-49), $\psi_s = 0°$ and $\psi_f = 180°$. The end-fire array pattern has only five nulls, including the null at $\theta = 180°$ as θ ranges from 0 to 180°, since only five zeros occur in the visible region.

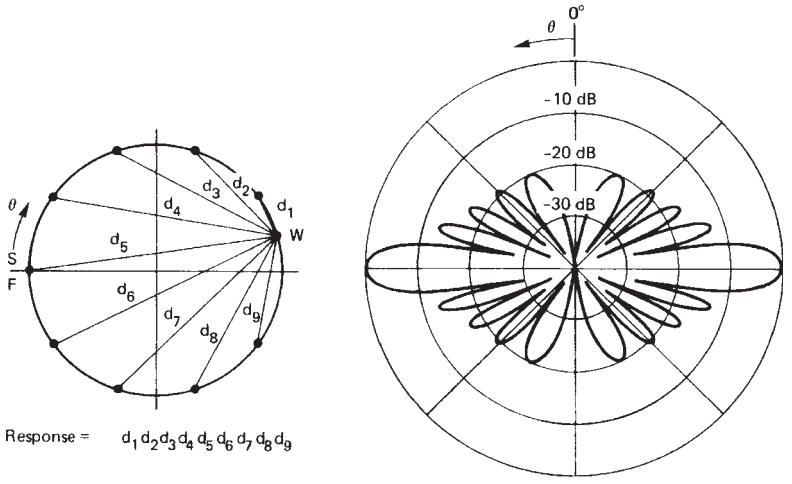


FIGURE 4-11 Unit-circle representation of a 10-element array with λ/2 spacings.

The Hansen and Woodyard increased-directivity end-fire array corresponds to a shift in the start and stop locations on the unit circle. The excursion from start to finish remains π determined by element spacing. Equation (4-49) calculates the start: $\psi_s = 90^\circ - 108^\circ = -18^\circ$. The pattern has five nulls from $\theta = 0$ to 180°.

A binomial array has all its zeros at $W = -1$ and its pattern has no sidelobes, since they occur for points on the unit circle between zeros. Only one beam forms as W traverses the unit circle. The W-space polynomial is $f(W) = (W + 1)^{N-1}$. For an array of given size we can manipulate the location of the nulls either to reduce sidelobes or to place pattern nulls. We reduce a sidelobe by moving the zeros on both sides of it closer together, but either the main-lobe beamwidth increases or the other sidelobes rise. We form a null in the array pattern by moving one of the zeros to that point on the unit circle corresponding to W at the null angle. Given a desired null θ_n,

$$W_i = e^{j(kd\cos\theta_n + \delta)} \tag{4-50}$$

Equation (4-50) gives the phase angle $kd\cos\theta_n + \delta$ of the zero required on the unit circle in W-space.

In the case of an end-fire array in which the spacing between elements is less than λ/2, we can shift zeros from invisible space into visible space to narrow the beam and reduce sidelobes. We thereby form large lobes in invisible space that represent energy storage in the array. The large energy storage reduces the bandwidth and efficiency of the array. This super-directivity method has limited success, although we can produce beautiful patterns on paper.

Example Design a four-element array of broadcast towers to give nearly uniform coverage for $\theta = \pm 45^\circ$ with nulls at $\theta = 270^\circ$ and 135° [11, p. 69].

We will align the array with $\theta = 0^\circ$ to obtain symmetry for the $\pm 45^\circ$ requirement. We actually need only three elements, since only two nulls are specified. Using λ/4 spacings, we set $\delta = -90^\circ$ to get an end-fire array. Equation (4-50) gives the zeros of

the polynomial required for the pattern nulls.

$$W_1: \qquad \psi = \frac{360^\circ}{\lambda}\frac{\lambda}{4}\cos(270^\circ) - 90^\circ = -90^\circ$$

$$W_2: \qquad \psi = \frac{360^\circ}{\lambda}\frac{\lambda}{4}\cos(135^\circ) - 90^\circ = -153.64^\circ$$

We determine the polynomial from the roots:

$$\begin{aligned} f(W) &= (W - e^{-j90^\circ})(W - e^{-j153.64^\circ}) \\ &= W^2 + 1.6994We^{j53.18^\circ} + e^{j116.36^\circ} \end{aligned}$$

We normalize the phase to the first element of the array [constant term of $f(W)$]:

$$f(W) = W^2e^{-j116.36^\circ} + 1.6994We^{-j148.18^\circ} + 1$$

At this point the polynomial representation of the array $f(W)$ does not include the progressive phase factor $\delta = -90^\circ$. We add the factor to the polynomial by adding -90° to the phase of the second element (W term) and -180° to the third element (W^2 term):

$$f(W) = W^2e^{-296.36^\circ} + 1.6994We^{-j148.18^\circ} + 1$$

The coefficients of the polynomial are the voltage (or current) components of the array. No null develops at $\theta = 180^\circ$ because the two available nulls ($N - 1$) were used. Adding the fourth element gives us the freedom to improve the response flatness in the $\pm 45^\circ$ region of θ. Figure 4-12 shows a unit-circle representation and pattern to give a nearly equal ripple response between $\pm 45^\circ$ and the required nulls. We increase

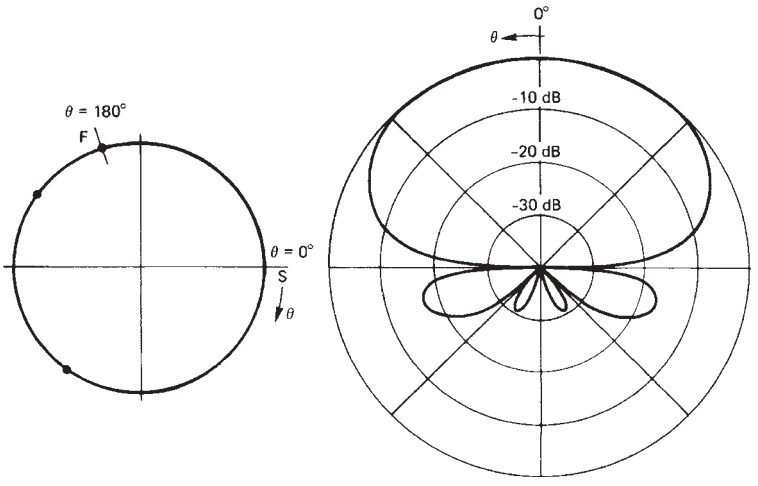


FIGURE 4-12 Four-element linear array with pattern nulls at $\theta = 90$, 135, and 180°. The elements are spaced at 0.35λ to give a flat response $\pm 45^\circ$.

TABLE 4-18 Four-Element 0.35λ Spaced-Array Coefficients for Uniform Beam ±45°

No.	Amplitude (dB)	Phase (deg)
1	−9.50	0.0
2	−4.11	−103.3
3	−4.11	138.4
4	−9.11	35.1

the spacing to 0.35λ and place the pattern nulls at 90°, 135°, and 180°. W starts at +1 on the unit circle or $\psi_s = 0^\circ$ and determines δ:

$$\psi_s = 0 = kd + \delta \quad \text{or} \quad \delta = -kd = -\frac{360^\circ}{\lambda} 0.35\lambda = -126^\circ$$

We compute the phase of the zeros from Eq. (4-50):

$$\psi_1 = 360^\circ (0.35) \cos(90^\circ) - 126^\circ = -126^\circ$$
$$\psi_2 = 360^\circ (0.35) \cos(135^\circ) - 126^\circ = -215.1^\circ \quad (144.9^\circ)$$
$$\psi_3 = 360^\circ (0.35) \cos(180^\circ) - 126^\circ = -252^\circ \quad (108^\circ)$$

By following the same steps as above, we compute the phase and amplitude of the array elements (Table 4-18).

4-10 DOLPH–CHEBYSHEV LINEAR ARRAY [2]

The Chebyshev polynomials have equal ripples in the region $x = \pm 1$, and the amplitude varies between +1 and −1. Outside that region the polynomial value rises exponentially:

$$T_m(x) = \begin{cases} (-1)^m \cosh(m \cosh^{-1} |x|) & x < -1 \\ \cos(m \cos^{-1} x) & -1 \le x \le 1 \\ \cosh(m \cosh^{-1} x) & x > 1 \end{cases}$$

The order of the polynomial m equals the number of roots. Dolph devised a method of relating the Chebyshev polynomials to the array factor polynomial for a broadside array. We scale the polynomial to make the equal-ripple portion the sidelobes and the exponential increase beyond $x = 1$ becomes the main beam. Take an array fed symmetrically about the centerline that has either $2N + 1$ or $2N$ elements. We expand the array factor in a polynomial with factors $\cos(\psi/2)$, where $\psi = kd \cos\theta + \delta$. The beam peak occurs when $\psi = 0$. If we make this correspond to a value x_0, where the Chebyshev polynomial has a value R, the sidelobes will be equal to the ripple at the level $1/R$. By substituting $x = x_0 \cos(\psi/2)$, we use the Chebyshev polynomial for the array polynomial with

$$T_m(x_0) = R \quad \text{or} \quad x_0 = \cosh \frac{\cosh^{-1} R}{m} \tag{4-51}$$

where 20 log R is the desired sidelobe level in decibels. The zeros of $T_m(x)$ are given by

$$x_p = \pm \cos \frac{(2p-1)\pi}{2m} \tag{4-52}$$

By using the equation $x_p = x_0 \cos(\psi/2) = x_0(e^{j\psi/2} + e^{-j\psi/2})$, we calculate the angles of the symmetrical zeros in the W-plane:

$$\psi_p = \pm 2 \cos^{-1} \frac{x_p}{x_0} \tag{4-53}$$

Both values of x_p [Eq. (4-52)] give the same ψ_p pair. Given the zeros in the W-plane, we multiply out the root form of the polynomial to calculate feeding coefficients of the array.

Example Design a 10-element array with 25-dB sidelobes.

The array has nine nulls, so we pick $m = 9$ for the Chebyshev polynomial.

$$\text{[Eq. (4-51)]} \qquad R = 10^{25/20} = 17.7828 \qquad x_0 = 1.0797$$

We need only the first five zeros, since they are symmetrical about zero. We calculate them from Eq. (4-52), divide them by x_0, and use Eq. (4-53) for their angles on the unit circle of the W-plane (Table 4-19). We multiply out the root form of the polynomial for the voltage (current) feeding coefficients of the array. Because the roots are symmetrical about the real axis, all phase angles are zero. We obtain the following coefficients:

Nos.	1, 10	2, 9	3, 8	4, 7	5, 6
Coefficient (dB)	−8.07	−5.92	−2.84	−0.92	0.0

Figure 4-13 shows the unit-circle representation and pattern of the array with λ/2 spacing.

We can estimate the beamwidth of a Chebyshev array by using a beamwidth broadening factor and the beamwidth of a same-length uniformly fed array [12]. The beamwidth broadening factor is given by

$$f = 1 + 0.632 \left[\frac{2}{R} \cosh \sqrt{(\cosh^{-1} R)^2 - \pi^2} \right]^2 \tag{4-54}$$

TABLE 4-19 Chebyshev Polynomial Roots and W-Plane Roots for 10-Element 25-dB Sidelobe Array

p	X_p	ψ_p (deg)
1	0.9848	±48.41
2	0.6428	±106.93
3	0.3420	±143.06
4	0.0	180.00

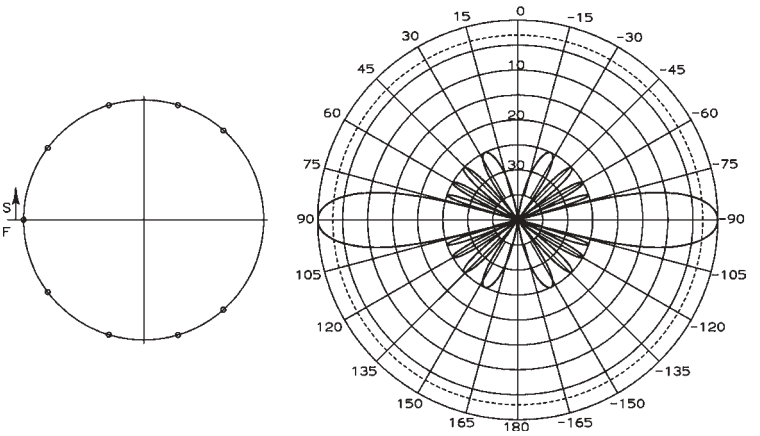


FIGURE 4-13 Ten-element Chebyshev array designed for 25-dB sidelobes.

Equation (4-54) is valid in the range of sidelobe levels from 20 to 60 dB and for scanning near broadside.

Example Compute the broadside beamwidth of a Dolph–Chebyshev array with 61 elements, a 30-dB sidelobe level, and $\lambda/2$ spacings.

Equation (4-54) gives the value 1.144 for f using $R = 10^{30/20}$.

We estimate the beamwidth of the uniform array from $\text{HPBW} = 50.76^\circ\lambda/Nd = 1.66^\circ$, where d is the element spacing:

$$\text{HPBW}_{\text{array}} = (f)\text{HPBW}_{\text{uniform}} = 1.144(1.66^\circ) = 1.90^\circ$$

We use the beam-broadening factor to estimate the array directivity:

$$D = \frac{2R^2}{1 + (R^2 - 1)f\lambda/Nd} \tag{4-55}$$

Example We calculate the directivity of the 61-element array above from Eq. (4-55). $D = 52.0$ (17.2 dB).

If we take its limit as $Nd \to \infty$, Eq. (4-55) becomes $2R^2$. An infinite Dolph–Chebyshev array has a gain 3 dB more than the sidelobe level.

4-11 VILLENEUVE ARRAY SYNTHESIS [13]

Villeneuve devised a method similar to the Taylor distribution that modifies the $\overline{n} - 1$ inner zeros of a uniform amplitude array to lower sidelobes. Since the positions of the outer zeros remain fixed, the outer pattern sidelobes decrease as $1/U$. The uniform distribution W-plane zeros are located uniformly around the unit circle except for $W = 1$:

$$\psi_p = \frac{2\pi p}{N_e} \tag{4-56}$$

The inner zeros correspond to the Chebyshev zeros [Eq. (4-53)] except that we multiply them by a constant factor α dependent on the number of elements, the sidelobe level, and $\overline{n}$:

$$\alpha = \frac{\overline{n}\pi}{N_e \cos^{-1}\left\{(1/x_0)\cos\left[(2\overline{n}-1)\pi/2m\right]\right\}} \tag{4-57}$$

The order of the Chebyshev polynomial $m = N_e/2$. We use Eq. (4-51) to compute x_0.

Example Design a 10-element Villeneuve array containing 10 elements for 25-dB sidelobes and $\overline{n} = 4$.

We determine $\alpha = 1.00653$ using Eq. (4-57). The inner three W-plane zeros are found by multiplying the Chebyshev zeros by α, which occur in pairs, and the next three zeros are found from the uniform amplitude array using Eq. (4-56):

$$\psi_p: \quad \pm 48.73 \qquad \pm 73.82 \qquad \pm 107.63 \qquad \pm 144 \qquad 180$$

Figure 4-14 illustrates the W-plane and pattern of the 10-element Villeneuve array:

Nos.	1, 10	2, 9	3, 8	4, 7	5, 6
Coefficient (dB)	−8.44	−5.85	−2.91	−0.91	0.0

The sidelobes drop off instead of staying constant: −25.08, −25.19, −25.43, −26.14.

4-12 ZERO SAMPLING OF CONTINUOUS DISTRIBUTIONS [14]

We sample continuous distributions, such as the Taylor line source, for large arrays. By using that method, we avoid the numerical difficulties of multiplying out long polynomials. When a small array samples an aperture distribution, its pattern fails to follow the pattern of the distribution. We improve the pattern by matching the zeros of the array to the distribution nulls. The array ψ-space pattern repeats at 2π intervals, but the k-space pattern of the aperture has no repeat. We space elements by $\lambda/2$ to

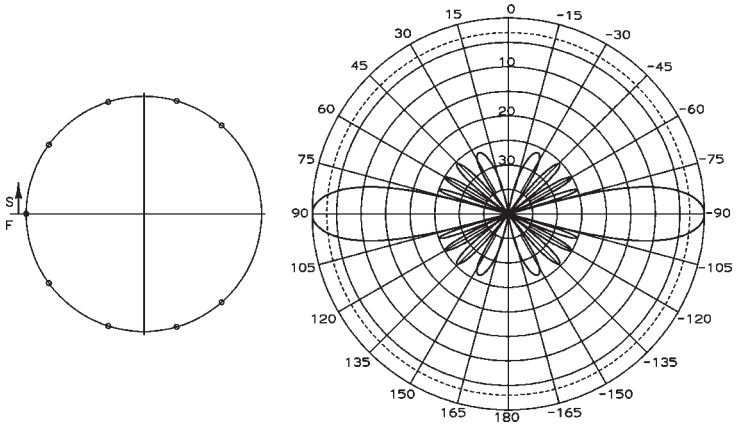


FIGURE 4-14 Ten-element Villeneuve array designed for 25-dB sidelobes, $\overline{n} = 4$.

span the total ψ-space nonrepeating region. We then equate an array with $\lambda/2$ spacings to an aperture of the same length regardless of the actual spacings between elements. Since the array samples a continuous distribution, the aperture is Nd long, where d is the distance between array elements and we consider the array element to be sampling $d/2$ on both sides of its location.

Consider the U-space pattern of a uniform aperture distribution: $\sin \pi U / \pi U$. The aperture zeros occur at integer values of U. The corresponding zeros of the uniformly fed array are $W_i = e^{j2\pi i/N}$, where $i = 1, 2, \ldots, N-1$. The Taylor distribution modifies the location of the zeros of the uniform distribution to U_i, and the sampled zeros of the array must move to follow this pattern:

$$W_i = e^{j2\pi U_i/N} \tag{4-58}$$

Example Given a Taylor line source with 30-dB sidelobes and $\overline{n} = 6$, compute the zeros of an array with 12 elements to sample the distribution.

The array spans 12/2 in U-space. We calculate zeros of the distribution from Section 4-4 and the angles of the array zeros from Eq. (4-58):

U_i	± 1.473	± 2.1195	± 2.9989	± 3.9680	± 4.9747	6
ψ_i	± 44.19	± 63.58	± 89.97	± 119.04	± 149.24	180

We multiply out the root form of the polynomial to compute the array feeding coefficients. The array has 30-dB first sidelobes. A straight sampling of the distribution gives an array whose sidelobes exceed 30 dB.

Figure 4-15 shows the unit-circle diagram of a zero-sampled Taylor line source with 25-dB sidelobes and $\overline{n} = 5$. The method places the zeros on the unit circle close enough together to limit the sidelobe peaks to less than 25 dB when W for a given pattern direction lies between the zeros. The array has higher sidelobes than the equivalent

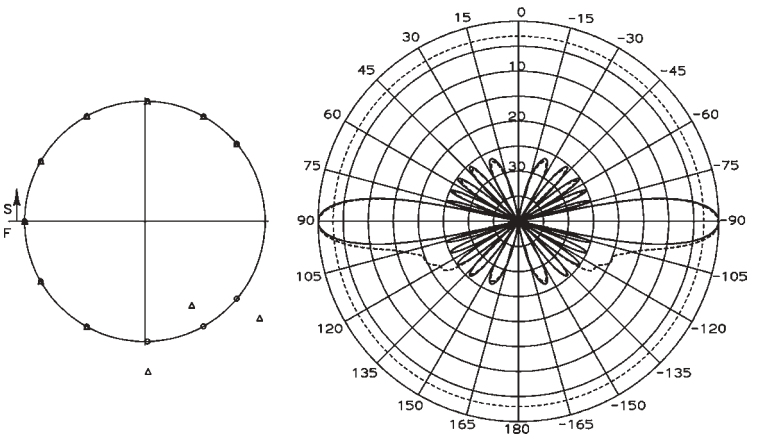


FIGURE 4-15 Twelve-element array designed by zero-sampling 25-dB Taylor distribution: pattern of normal array (solid curve); pattern with null filling by moving three zeros off unit circle (dashed curve and triangles).

aperture, but closer to the specified 25 dB, because the finite array cannot control sidelobes as well as the continuous aperture.

Aperture	25.29	25.68	26.39	27.51	29.63
12-Element Array	25.03	25.07	25.18	25.44	26.41

The dashed plot of Figure 4-15 illustrates pattern behavior when W-space zeros are moved off the unit circle. We can fill pattern nulls and generally shape the pattern. When we place all zeros on the unit circle in the complex plane, it can be proved that the array excitations will have amplitude symmetry about the centerline. Moving the zeros off the unit circle disturbs this symmetry. We can eliminate all pattern nulls by moving all the W-plane zeros off the unit circle. If we start with a uniformly fed array and move all the zeros to the same radius, the distribution taper across the array will be linear in decibels. In the next two sections we explore techniques for moving the zeros systematically to produce shaped patterns from an array.

4-13 FOURIER SERIES SHAPED-BEAM ARRAY SYNTHESIS

The preceding methods seek the narrowest beamwidths for a given sidelobe level. Arrays can also produce shaped beams. We discussed the Woodward line-source method for shaped beams in Section 4-8. We obtain good approximations by sampling the line-source distribution with an array. Beyond sampling a line source, we can apply Fourier series to design an array directly. An array for a shaped beam must be much larger than is required for the beamwidth. The extra size of the array gives us the degrees of freedom necessary for beam shaping. Increasing the array size increases the match between the specified and the actual beam shape.

Because the array pattern is periodic in k-space, we can expand the pattern in a Fourier series. The array pattern for a symmetrically fed array is given by either

$$f(\psi) = 1 + 2\sum_{n=1}^{m} \frac{I_n}{I_0} \cos\frac{2n\psi}{2} \qquad N \text{ odd} \tag{4-59}$$

or

$$f(\psi) = 2\sum_{n=1}^{m} \frac{I_n}{I_0} \cos\frac{(2n-1)\psi}{2} \qquad N \text{ even} \tag{4-60}$$

where $m = (N-1)/2$ (odd) or $m = N/2$ (even) with $\psi = kd\cos\theta + \delta$. Equations (4-59) and (4-60) are Fourier series expansions of the pattern in ψ-space. The elements farthest from the centerline produce the highest harmonics in the series.

In an asymmetrically fed array, we express Eqs. (4-59) and (4-60) as a sum of exponential terms:

$$f(\psi) = \begin{cases} \displaystyle\sum_{n=-m}^{m} a_n e^{jn\psi} & N \text{ odd} \\ \displaystyle\sum_{n=1}^{m} a_n e^{j(2n-1)\psi/2} + a_{-n} e^{-j(2n-1)\psi/2} & N \text{ even} \end{cases} \tag{4-61}$$

Suppose that we have a desired pattern in k-space given by $f_d(\psi)$. We expand it in an infinite Fourier series of the same form as Eq. (4-61) with $m = \infty$. We equate the first m coefficients of the two Fourier series to approximate the desired pattern. As in any Fourier series method, we solve for the coefficients by using the orthogonality of the expansion functions when integrated over a period:

$$a_n = \begin{cases} \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} f_d(\psi) e^{-jn\psi}\, d\psi & N \text{ odd} \\ \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} f_d(\psi) e^{-j(2n-1)\psi/2}\, d\psi & N \text{ even} \end{cases} \tag{4-62}$$

$$a_{-n} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_d(\psi) e^{j(2n-1)\psi/2}\, d\psi \qquad N \text{ even} \tag{4-63}$$

We determine the array coefficients directly from the Fourier series coefficients.

Example Design a 21-element array with $\lambda/2$ element spacing with a constant beam $2b$ wide centered in ψ-space.

We use Eq. (4-62) to compute coefficients a_n:

$$a_n = \frac{1}{2\pi} \int_{-b}^{b} e^{-jn\psi}\, d\psi = \frac{\sin nb}{\pi n}$$

Suppose that the constant beam is 45° at broadside: $67.5^\circ \le \theta \le 112.5^\circ$. Then

$$b = \frac{360^\circ}{\lambda} \frac{\lambda}{2} \cos 67.5^\circ = 68.88^\circ$$

We can ignore the constant factor i/π and expand to compute the array coefficients (Table 4-20).

The method fails to some extent when we try it on arrays with spacings greater than $\lambda/2$. The integral does not cover the total visible region. We can, however, use it with

TABLE 4-20 Fourier Series Synthesis Coefficients of 21-Element Array for Pattern of Figure 4-16

n	a_n	Amplitude (dB)	Phase (deg)
0	1.0000	0.00	0
±1	0.9328	−0.60	0
±2	0.3361	−9.47	0
±3	−0.1495	−16.50	180
±4	−0.2488	−12.08	180
±5	−0.0537	−25.40	180
±6	0.1336	−17.48	0
±7	0.1209	−18.35	0
±8	−0.0240	−32.40	180
±9	−0.1094	−19.22	180
±10	−0.0518	−25.72	180

spacings less than $\lambda/2$ with good results. As we increase the number of elements in the array, the match to the desired pattern improves. Of course, tapering the desired pattern reduces the higher harmonics and the subsequent need for more elements.

Example Suppose that we want to scan the beam of the 21-element array with $\lambda/2$ element spacing to 60° with a 45° beamwidth. The beam edges are 37.5° and 82.5°.

We could calculate coefficients by integrating Eq. (4-62) directly with this requirement, but we can use δ, the progressive phase shift between elements, to simplify the problem. The beam edges in ψ-space are

$$\begin{array}{cc} 180^\circ \cos(37.5^\circ) + \delta & \text{and} \quad 180^\circ \cos(82.5^\circ) + \delta \\ 142.8^\circ + \delta & 23.49^\circ + \delta \end{array}$$

We pick δ to center the beam in ψ-space: $b = 142.8^\circ + \delta$, $-b = 23.49^\circ + \delta$. On solving, we have $\delta = -83.15^\circ$ and $b = 59.65^\circ$. We use the formula $\sin(nb)/\pi n$ to compute coefficients of the array and then add the progressive phase shift through the array. Figure 4-16 shows the array pattern.

When we scan the beam to end fire, we must account for the symmetry about $\theta = 0^\circ$. Because we limit the spacings to less than $\lambda/2$ to prevent grating lobes, we have an unspecified region of ψ-space that we can choose in any convenient manner.

Example Design a 21-element end-fire array with a 90° beamwidth and 0.30λ spacings.

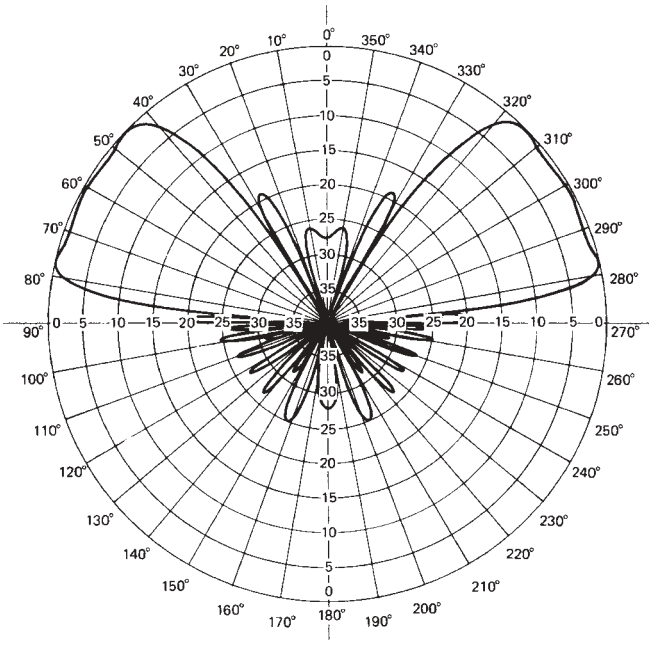


FIGURE 4-16 Twenty-one-element array designed by the Fourier series method to scan to 60° with a 45° beamwidth.

For an end-fire array we pick $\delta = -kd = -108^\circ$. This places the edge of the visible region on the ψ-space origin. We are free to specify the invisible region that will be included in the integral [Eq. (4-62)]. We specify the invisible region as the mirror image of the portion in the visible region and solve for b:

$$-b = 360^\circ(0.3)\cos(45^\circ) - 108^\circ = -31.63^\circ$$

We use the $\sin(nb)/\pi n$ formula to calculate array coefficients and then apply the progressive phase shift δ to the coefficients obtained to get the proper phase to scan to end fire.

We cannot control the sidelobes of an array designed using Fourier series expansion. The initial specification calls for no sidelobes. Sampling a Woodward linear aperture with an array also fails to give control of the sidelobes. The Woodward linear distribution cannot control sidelobes; it provides only ease of design. In the next section we explore a method with direct control of sidelobes of an array.

4-14 ORCHARD METHOD OF ARRAY SYNTHESIS [15]

In Section 4-6 we manipulate the nulls of a continuous linear distribution to control the sidelobes of the radiated pattern individually. In Section 4-9 we show that the nulls of the linear aperture pattern can be related directly to the roots of Schelkunoff polynomial representation of the linear array pattern in W-space. The unit circle method gives us a tool for array synthesis expanded in the Orchard method for the design of arrays with arbitrary patterns. We apply an iterative technique on the W-space zeros to produce the pattern desired. We control all the sidelobes individually and produce shaped patterns for the main beam. The finite size of the array limits the control of the main beam shape as we saw in the Fourier series expansion method. Each array element corresponds to a term in the Fourier series expansion.

We start with the Schelkunoff transformation of the array pattern:

$$f(W) = C_0 \prod_{n=1}^{N} (W - W_n) \tag{4-64}$$

A normalization constant C_0 has been added. We write $W_n = \exp(a_n + jb_n)$. Expansion of Eq. (4-64) produces the feeding coefficients of an array with $N + 1$ elements:

$$W = e^{j\psi} \qquad \text{with} \quad \psi = kd\cos\theta + \delta$$

θ is measured from the array axis. The effect of δ on the unit circle method is to rotate the starting and finishing points when varying W to calculate the pattern using Eq. (4-64). An equally valid method is to rotate the zeros about the origin of the complex plane, which leaves the ψ-space pattern shape unchanged. When designing a shaped beam, we need to rotate the main beam peak to the proper location to calculate the amplitudes because our specification will be in terms of the pattern angle θ relative to the peak.

Figure 4-11 illustrates that the pattern amplitude is the product of distances from each zero to the pattern W point. Expansion of Eq. (4-64) in terms of the product of distances to W gives

$$|f(W)|^2 = C_0^2 \prod_{n=1}^{N} [1 - 2e^{a_n} \cos(\psi - b_n) + e^{2a_n}] \tag{4-65}$$

The Orchard method requires the specification of each sidelobe and additional values located at the minimum ripple points in the shaped region. For a single-beam unshaped pattern, we only specify sidelobes, and all a_n will be zero since all zeros W_n will be on the unit circle. We restrict the array to $\lambda/2$ spaced elements when applying the method so that the entire unit circle is used in the pattern. An array with N zeros has N pattern peaks which lie between the zeros in the W-plane. When we include the normalization constant C_0 to specify the main beam peak and all the zeros, we have $N + 1$ unknowns to find. Without loss of generality we specify the last zero as $W_N = -1$ or $\psi = \pi$ to reduce the number of unknowns to N. Since we rotate the zeros after we determine the proper zero spacing for specified sidelobes, we place the main beam between W_{N-1} and $W_N = -1$. Before starting the iteration technique, we generate a list of sidelobe levels with the main beam as the last one.

The method expands the pattern in a multiple-variable Taylor series using b_n, a_n, and the normalization constant as variables. To facilitate calculating the partial derivatives, we express Eq. (4-65) in decibels:

$$G = \sum_{n=1}^{N-1} \frac{10}{\ln(10)} \ln[1 - 2e^{a_n} \cos(\psi - b_n) + e^{2a_n}] + 10 \log_{10}[2(1 + \cos\psi)] + C \tag{4-66}$$

The second term of Eq. (4-66) is due to the zero $W_N = -1$ and C is the normalization constant of the main beam. The logarithm to the base 10 has been expressed as a natural logarithm for the calculation of derivatives:

$$\frac{\partial G}{\partial a_n} = \frac{M e^{a_n}[e^{a_n} - \cos(\psi - b_n)]}{1 - 2e^{a_n} \cos(\psi - b_n) + e^{2a_n}} \tag{4-67}$$

$$\frac{\partial G}{\partial b_n} = -\frac{M e^{a_n} \sin(\psi - b_n)}{1 - 2e^{a_n} \cos(\psi - b_n) + e^{2a_n}} \tag{4-68}$$

$$\frac{\partial G}{\partial C} = 1 \tag{4-69}$$

The variable $M = 20/\ln(10)$. The multiple-variable Taylor series involves three types of terms:

$$G(b_n, a_n, C) = G_0(b_{n0}, a_{n0}, C_0) + \sum_{n=1}^{N-1} \frac{\partial G}{\partial b_n}(b_n - b_{n0}) + \sum_{n=1}^{N-1} \frac{\partial G}{\partial a_n}(a_n - a_{n0}) + (C - C_0) \tag{4-70}$$

Every nonzero value of a_n fills in the pattern null at $\psi = b_n$. If we specify the desired pattern amplitude at every sidelobe peak, the main beam, and at points between the sidelobe peaks equal to the number of nonzero a_n, we form a square matrix equation. The solution gives the changes in b_n, a_n, and C. Since we expanded Eq. (4-66) as a linear approximation, the solution of Eq. (4-70) gives only an approximate solution. In a few iterations the method converges and we obtain an acceptable pattern.

Suppose that the shaped pattern is limited to a range in W-space so that there are only L nonzero a_n. Given the desired pattern $S_m(\psi_m)$ at ψ_m and the current pattern $G_0(\psi_m)$, one row of the matrix is

$$\left[\frac{\partial G(\psi_m)}{\partial b_1}, \ldots, \frac{\partial G(\psi_m)}{\partial b_{N-1}}, \frac{\partial G(\psi_m)}{\partial a_1}, \ldots, \frac{\partial G(\psi_m)}{\partial a_L}, 1\right]$$

We need $N + L$ rows or pattern points to solve Eq. (4-70) for changes in b_n, a_n, and C:

$$[\delta b_1, \ldots, \delta b_N, \delta a_1, \ldots, \delta a_L, \delta C]^{\mathrm{T}}$$

We require a search routine to locate the pattern peaks between the pattern nulls or minima between peaks in the shaped region for given values of b_n and a_n after we normalize to the current pattern peak. We subtract these from the levels desired:

$$[S(\psi_1) - G_0(\psi_1), \ldots, S(\psi_{N+L}) - G_0(\psi_{N+L})]^{\mathrm{T}}$$

After solving the square matrix equation, we update the W-plane zeros:

$$\begin{aligned} b_1 &= b_1 + \delta b_1 \\ &\vdots \\ b_{N-1} &= b_{N-1} + \delta b_{N-1} \\ &\vdots \\ a_1 &= a_1 + \delta a_1 \\ &\vdots \\ a_L &= a_L + \delta a_L \\ C &= C_0 + \delta C \end{aligned}$$

The iteration alters the beam peak and its location. The pattern peak is normalized after iteration, and for a shaped pattern a new zero rotation is found to line up the beam peak for the pattern-shaping function.

Example Design an eight-element array with its beam peak at 90° and specified sidelobes before the peak of 25, 30, and 25 dB and 20, 25, and 30 dB after the peak.

The sidelobes values begin with the first sidelobe after the peak and rotate to the peak:

$$-20 \qquad -25 \qquad -30 \qquad -25 \qquad -30 \qquad -25 \qquad 0$$

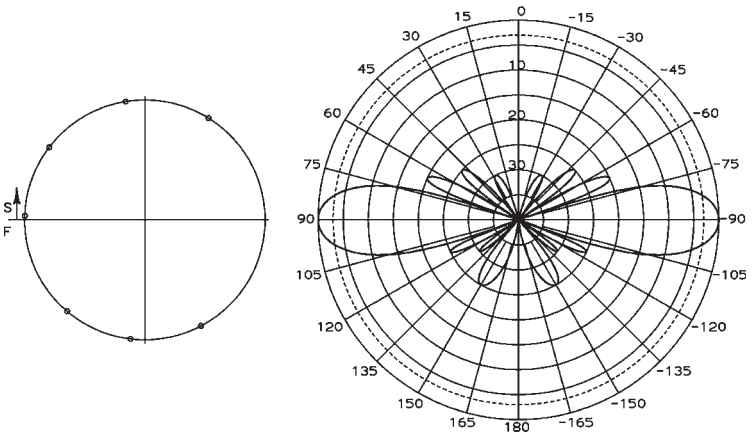


FIGURE 4-17 Eight-element array designed using Orchard synthesis for individually specified sidelobes, $\lambda/2$ spacings.

The solution converges in four iterations after starting with uniformly spaced zeros on the unit circle. Figure 4-17 shows the unit-circle zeros on the left and the corresponding pattern on the right with $\lambda/2$ element spacing:

W-Space Zero (deg)	178.14	142.72	99.26	58.01	−62.18	−96.89	−130.37
Pattern Null (deg)	8.25	37.54	56.53	71.20	110.21	122.57	136.41

The feeding coefficients for the final design are given in Table 4-21.

Although the Orchard method requires the elements to be spaced $\lambda/2$ during synthesis, the completed design can be used at another element spacing. Figure 4-18 gives the unit-circle diagram of the same array with a 0.7λ element spacing. The range of W now exceeds 2π and the sidelobe regions of the unit circle have been used more than once. Sidelobes 3 and 4 occur twice in the pattern. Of course, if we scan the array too far, the pattern would have grating lobes. Figure 4-19 plots the pattern of an end-fire array with $\lambda/4$-element spacing using the same zeros. Only a portion of the unit circle is used, and not all sidelobes are realized. Figure 4-20 illustrates the end-fire case with the elements spaced so that the final position of W occurs at a null. The pattern contains all six sidelobes. The unit-circle analysis mirrors that of

TABLE 4-21 Coefficients of Eight-Element Array of Figure 4-17 Designed by Orchard Synthesis

Element	Magnitude (dB)	Phase (deg)	Element	Magnitude (dB)	Phase (deg)
1	−8.69	8.70	5	0	3.79
2	−3.90	3.22	6	−1.06	7.41
3	−1.06	1.29	7	−3.90	5.48
4	0	4.91	8	−8.69	0

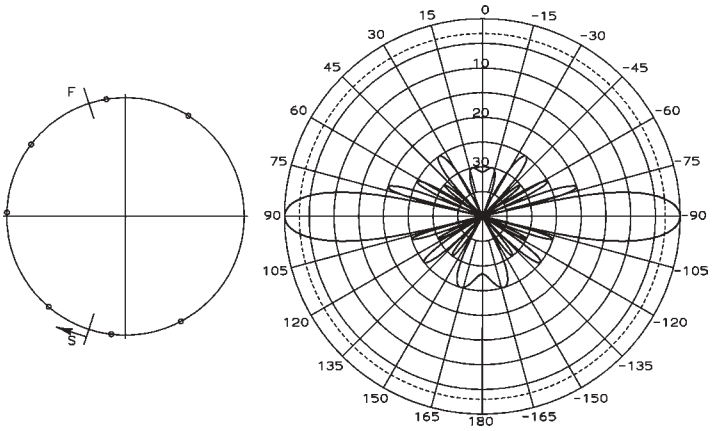


FIGURE 4-18 Eight-element array designed using Orchard synthesis for individually specified sidelobes, 0.7λ spacings.

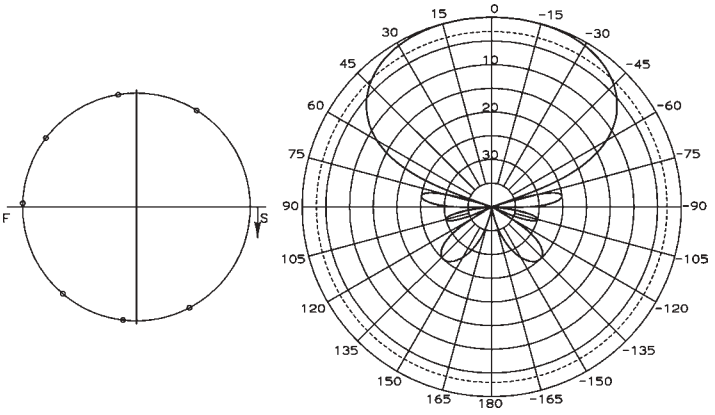


FIGURE 4-19 Eight-element array designed using Orchard synthesis for individually specified sidelobes, λ/4 spacings scanned to end fire.

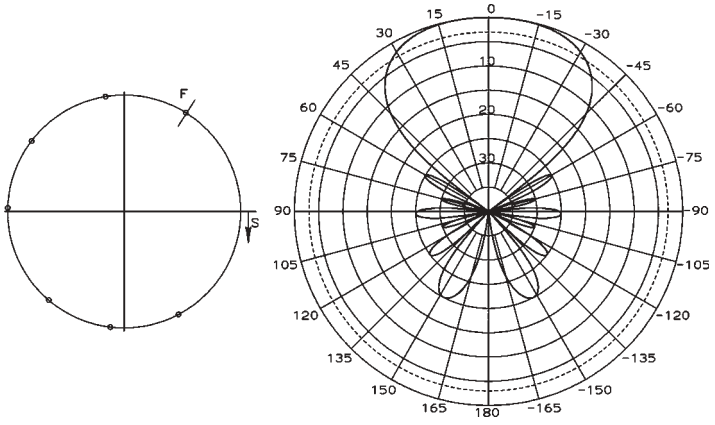


FIGURE 4-20 Eight-element array designed using Orchard synthesis for individually specified sidelobes, 0.42λ spacings scanned to end fire.

the circle diagram in Chapter 3, where increasing the element spacing increases the visible region. In this case the visible region corresponds to rotation about the unit circle. Expansion of Eq. (4-64) produces the array feeding coefficients independent of element spacing, and the progressive phase shift between elements δ affects phase but not amplitude. The four examples given in Figures 4-17 to 4-20 have the same sequence of feed magnitudes.

We can use Orchard synthesis to generate a difference pattern similar to the Bayliss line distribution and control all the sidelobes. A difference pattern has two main beams. Using the same example of an eight-element array, we modify the sidelobe list to include side-by-side main beams. We eliminate the −25-dB lobe next to the original main beam from the values above:

$$-20 \qquad -25 \qquad -30 \qquad -25 \qquad -30 \qquad 0 \qquad 0$$

When we apply the synthesis by placing the last main beam at 90°, we obtain a pattern with two main beams with the null between them at 101.6°, corresponding to a W-plane null at −36.3°. We rotate all W-plane zeros by 36.3° to place the null between the two main beams at 90°. Figure 4-21 shows the W-plane and polar pattern for the final design. Note the placement of the W-plane zero at $W = +1$. Table 4-22 lists the feeding coefficients.

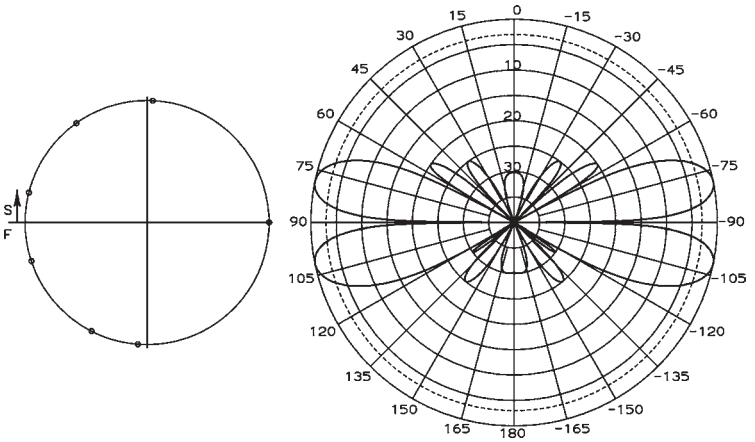


FIGURE 4-21 Difference pattern array using eight-elements designed by Orchard synthesis.

TABLE 4-22 Coefficients of Eight-Element Difference Pattern Array of Figure 4-21 Designed by Orchard Synthesis

Element	Magnitude (dB)	Phase (deg)	Element	Magnitude (dB)	Phase (deg)
1	−6.32	5.39	5	−6.91	178.35
2	−0.35	1.28	6	0.0	184.6
3	0.0	0.8	7	−0.35	184.12
4	−6.91	7.05	8	−6.32	180

We fill in the null between the different lobes to form a flat-topped beam for the eight-element array and use a constant-amplitude shaping function for the pattern desired. The beamwidth of the flat lobe is determined by the lobe spacing, and only certain sizes are possible. Remember that an array is a Fourier series approximation to the pattern desired. With only eight elements the match is poor between the pattern desired and the approximate pattern. We use one nonzero a_n to move the W-plane zero off the unit circle that forms the pattern null between the two beams and add another pattern specification:

$$-20 \qquad -25 \qquad -30 \qquad -25 \qquad -30 \qquad 0 \qquad 0 \qquad -1$$

The last number gives the pattern level at the null relative to the shaped pattern level. This last term uses Eq. (4-67) for its columns. The constant beam design uses a 22°-wide beam centered at 90° for the pattern shape function. We start with $a_n = 0.01$ before iterating. The iteration using the matrix equation computes $a_1 = 0.4435$, which can be either positive or negative without changing the pattern. Rotation of the W-plane zeros placed the zero for minimum ripple along the positive real axis and produced a symmetrical pattern about $\theta = 90°$. Figure 4-22 contains the final design W-space zeros and polar pattern. The iterations produced the sidelobe levels specified (Table 4-23).

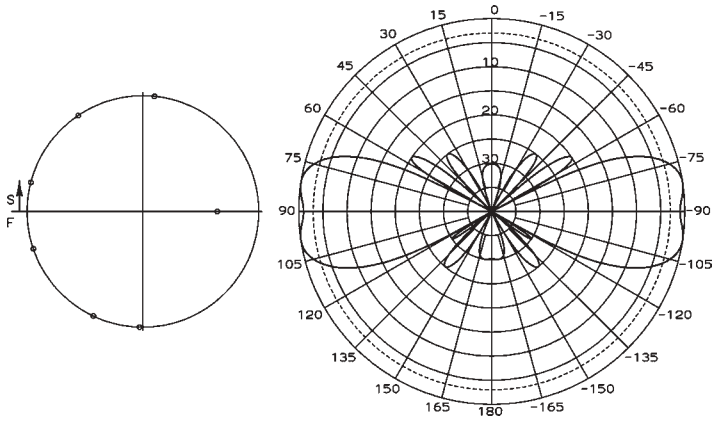


FIGURE 4-22 Flat-topped beam eight-element array designed by Orchard synthesis.

TABLE 4-23 W-Plane Zeros of Eight-Element Flat-Topped Beam of Figure 4-22 Designed by Orchard Synthesis

W-Space Zero (deg)	W-Space Radius	Pattern Null (deg)
165.51	1.0	23.15
123.99	1.0	46.46
84.33	1.0	62.06
0	0.6418	90
−91.65	1.0	120.61
−115.39	1.0	129.87
−161.18	1.0	153.57

TABLE 4-24 Coefficients of Eight-Element Array for Flat-Topped Beam of Figure 4-22 Designed by Orchard Synthesis

Element	Magnitude (dB)	Phase (deg)	Element	Magnitude (dB)	Phase (deg)
1	−12.95	−174.39	5	−1.15	1.54
2	−10.78	178.85	6	0.0	4.37
3	−24.69	167.92	7	−2.26	3.95
4	−7.90	−1.47	8	−9.10	0.0

The radius of the fourth term could be $1/0.6418 = 1.5581$ without affecting the pattern result. Inserting the zeros into Eq. (4-64) and expanding the polynomial produces the feeding coefficients (Table 4-24).

The Fourier series example for a constant beam centered at 60° with a 45° beamwidth using 21 elements spaced $\lambda/2$ (Figure 4-16) was repeated using Orchard synthesis. Fourier series synthesis could not control the sidelobes. First, we need to figure out how many array lobes cover the shaped pattern region. Place the zeros uniformly around the unit circle in the W-plane and determine how many of the roots are within the beam. For a 21-element array six beams and five zeros lie in the $\psi = \pi \cos\theta$ angular region of the constant beam found using Eq. (4-71):

$$\text{beams} = \frac{N(\cos\theta_{\min} - \cos\theta_{\max})}{2} \tag{4-71}$$

The solution to Eq. (4-71) is an integer given N as the number of W-plane zeros. All sidelobes were set at −30 dB and the ripple at −0.9 dB below the constant beam:

Lobes	1–14	15–20	21–25
Sidelobe (dB)	−30	0.0	−0.9

Figure 4-23 gives the final result of the synthesis, an improvement over Figure 4-16, with its uncontrolled sidelobes.

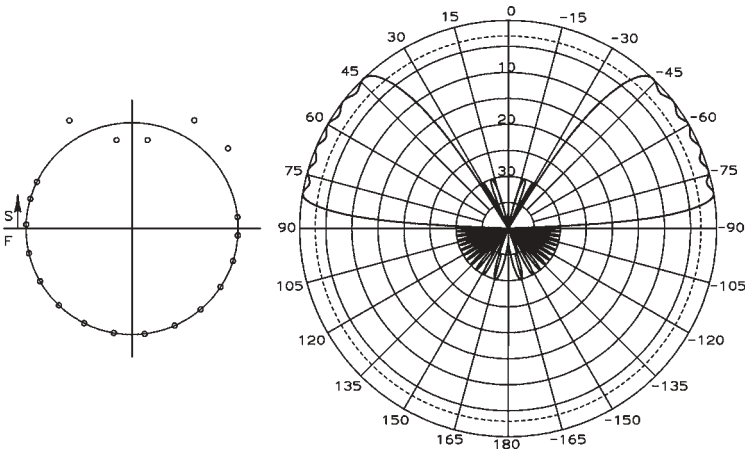


FIGURE 4-23 Twenty-one-element array designed by Orchard synthesis to scan to 60° with a 45° beamwidth.

We must consider the element excitations. The zeros not lying on the unit circle can be either inside or outside the circle and produce the same pattern. Different combinations of zero locations lead to different element amplitudes. The last example has five zeros displaced from the unit circle, which produces $2^5 = 32$ combinations. We need to check the amplitude distribution that results from each case (Table 4-25). Arrays with a large range of amplitudes are difficult to produce. In some cases the range of amplitude available is limited, such as waveguide slot arrays. Mutual coupling between elements also makes it difficult to achieve the desired low amplitudes on some elements because nearby elements will excite them, and compensation for mutual coupling may prove difficult. Figure 4-23 shows one of the combinations of root placements that produced the minimum amplitude variation in the array.

The Fourier series synthesis gave an amplitude variation of 32.4 dB, whereas the Orchard synthesis variation is 13.29 dB. This synthesis produced better patterns with less amplitude variation. Decreasing the ripple depth increases the amplitude variation of the array.

Csc²θ cos θ Pattern This pattern produces constant round-trip signals versus the elevation angle for radar. The pattern from the array axis is given by $\csc^2(\theta - 90^\circ)\cos(\theta - 90^\circ)$. The peak occurs beyond 90° and decreases for greater angles. The shaped pattern function requires the rotation of the W-plane zeros at each step so that the pattern peak calculated from the zeros occurs at the proper angle. The changing zero locations move the beam peak location at each iteration.

Example Design a 16-element $\csc^2(\theta - 90^\circ)\cos(\theta - 90^\circ)$ beam array to operate from 100 to 140° and have 30-dB sidelobes.

Equation (4-71) determines that five beams cover the pattern region and sets the number of nonzero a_n as 4. The 16-element array has 15 zeros, with the first 10 specified as −30 dB, five for the shaped-beam region, and four for the minima between the shaped-beam peaks. We specify the shaped-beam lobes relative to the shape levels. The last lobe is the beam peak.

TABLE 4-25 Coefficients of 21-Element Array for Flat-Topped Beam of Figure 4-23 Designed by Orchard Synthesis

Element	Amplitude (dB)	Phase (deg)	Element	Amplitude (dB)	Phase (deg)
1	−11.85	−65.47	12	−3.48	−115.31
2	−6.89	−146.85	13	−3.01	−158.73
3	−6.23	126.55	14	−2.97	134.01
4	−11.34	4.64	15	−4.93	49.59
5	−5.12	−158.19	16	−7.51	−40.53
6	0.0	108.08	17	−9.64	−111.80
7	−0.01	23.97	18	−9.44	−159.41
8	−6.33	−65.70	19	−8.06	142.99
9	−9.73	86.93	20	−9.26	73.66
10	−2.09	−4.11	21	−13.29	0.0
11	−1.75	−70.12			

Lobe	11	12	13	14	15	16	17	18	19
Amplitude (dB)	1.0	0.8	0.6	0.4	0.2	−0.9	−0.7	−0.5	−0.2

Allowing the ripple to increase in the lower levels of the shaped pattern region decreases the range of element amplitudes. The method converged in 11 iterations to the design given in Figure 4-24. All $2^4 = 16$ combinations of a_n placements inside and outside the unit circle were checked (Table 4-26). The amplitude variation ranged from 11.47 to 25.47 dB.

Figure 4-25 illustrates the design repeated with eight elements (Table 4-27). Although the sidelobes could be controlled at −30 dB, the shaped pattern region shows less pattern control than with 16 elements.

Extensions to the Orchard method make various improvements. By adding balancing zeros inside and outside the unit circle in the W-plane, the feeding coefficients of the array can be made real with only 0 or 180° phases [16]. This adds elements to the array and changes the shape of the beam somewhat. The coefficients are real only if the pattern is symmetrical about $\theta = 90°$. To implement the method you add a term

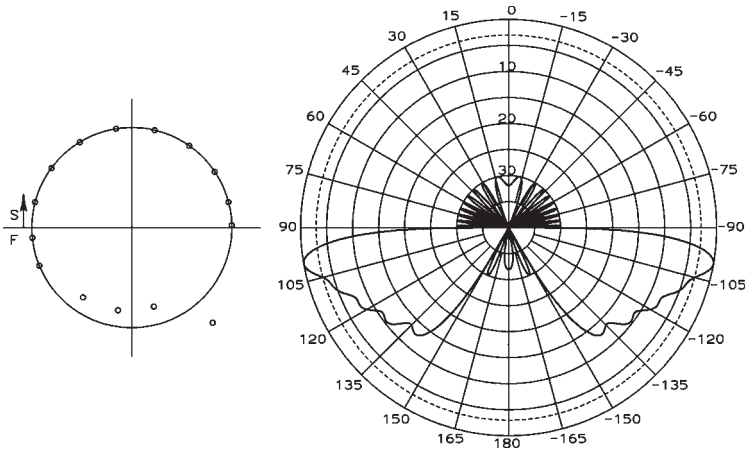


FIGURE 4-24 Sixteen-element array with $\csc^2\theta \cos\theta$ pattern designed by Orchard synthesis.

TABLE 4-26 Coefficients of 16-Element Array for $\csc^2\theta \cos\theta$ Beam of Figure 4-24 Designed by Orchard Synthesis

Element	Magnitude (dB)	Phase (deg)	Element	Magnitude (dB)	Phase (deg)
1	−11.47	−149.27	9	−1.19	149.09
2	−9.84	−100.16	10	−2.71	−177.54
3	−8.07	−69.72	11	−3.53	−131.64
4	−4.79	−40.25	12	−7.82	−52.50
5	−2.65	−0.56	13	−9.46	100.34
6	−2.04	34.20	14	−4.30	−157.29
7	−0.82	65.82	15	−4.55	−76.99
8	0.0	107.07	16	−8.72	0.0

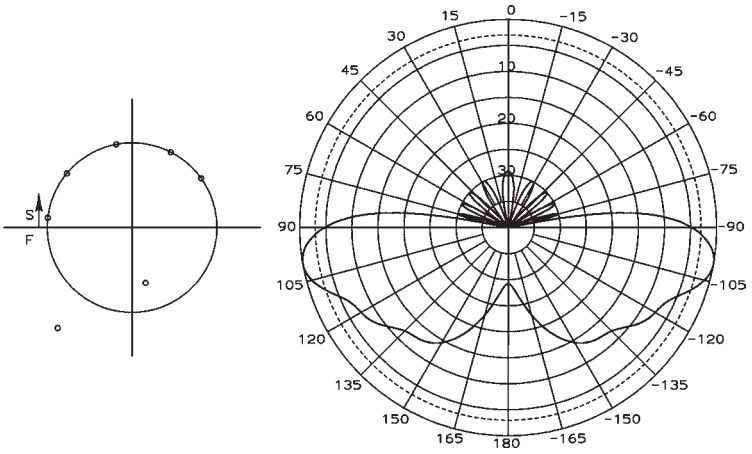


FIGURE 4-25 Eight-element array with $\csc^2\theta \cos\theta$ pattern designed by Orchard synthesis.

TABLE 4-27 Coefficients of Eight-Element Array for $\csc^2\theta \cos\theta$ Beam of Figure 4-24 Designed by Orchard Synthesis

Element	Magnitude (dB)	Phase (deg)	Element	Magnitude (dB)	Phase (deg)
1	−11.62	130.02	5	0.0	−122.98
2	−8.61	−170.78	6	−2.05	−79.05
3	−12.62	179.21	7	−5.73	−44.44
4	−2.61	−173.92	8	−11.54	0.00

to Eqs. (4-67) and (4-68) for the extra elements located off the unit circle. A design of a flat-topped beam centered at 90° using the balanced zeros produced a design with more than 30 dB of variation between the elements similar to a Fourier series expansion that had about the same range of amplitudes. The range of amplitudes in the array can be reduced by placing all the zeros off the unit circle [17, p. 124]. We give up the nulls between the lobes and must now search a large set of possible solutions to select a design with the least amplitude variation. A genetic algorithm sorts through the large set of zero combinations inside/outside all that satisfy the pattern requirements to discover the best design.

4-15 SERIES-FED ARRAY AND TRAVELING-WAVE FEED SYNTHESIS

A series-fed array uses couplers along a line that distribute power to the elements from a single transmission line. A single wave travels along the line with each element removing a portion of the power. A matched load absorbs the remaining power at the end to prevent the reflection of a wave traveling toward the source end. A second backward traveling wave would produce another beam with reduced amplitude indistinguishable from a sidelobe. The coupling could be a physical coupler or it could be just a series or shunt load across the transmission line. Waveguide slots are an example of loads on a transmission line. An array using couplers can have phase shifters between

the couplers and the elements to form a phased array. A second configuration for a phased array places phase shifters in the transmission line between the couplers. This case uses the simple control of identical phase shifters set to the same value to scan the beam. The phase shifters are the progressive phase δ along the array used for scanning.

The array distribution is given by the sequence of radiated powers, P_i. The traveling wave or nonresonant array dissipates a ratio of the input power R in the load:

$$\sum_{i=1}^{N} P_i = P_{\text{in}}(1 - R)$$

We normalize the distribution to the sum of radiated power: $P_0 = \sum_{i=1}^{N} P_i$. The element power becomes $P_i(1 - R)/P_0$ and we use the normalized power distribution to calculate coupling values:

$$C_1 = P_1 \qquad \text{remaining power} = 1 - P_1$$

The coupling to the second element removes power from the remaining power:

$$C_2 = \frac{P_2}{1 - P_1} \qquad \text{remaining power} = 1 - P_1 - P_2$$

The general expression is

$$C_i = \frac{P_i}{1 - \sum_{j=1}^{i-1} P_j} \tag{4-72}$$

If the element electrical model consists of a shunt conductance on a transmission line, such as waveguide slots, the power radiated by each slot $= |V_{\text{inc}}|^2 g_i$ and the normalized $g_i = C_i$. Similarly, an electrical model of an element as a series resistance on a transmission line can be solved in a similar manner. Power radiated $= |I_{\text{inc}}|^2 r_i$ and the normalized $r_i = C_i$.

Some array feeders have significant losses between the elements and we must account for these losses when designing the couplers. Suppose that the feeder has identical losses $L_f = 1 - 10^{-\text{attenuation}/10}$ between couplers. The power balance equation becomes

$$\underset{\text{total}}{P_{\text{in}}} = \underset{\text{load}}{RP_{\text{in}}} + \underset{\text{losses to load}}{(N-1)L_f R P_{\text{in}}} + \underset{\text{antennas}}{\sum_{i=1}^{N} P_i} + \underset{\text{losses to antennas}}{L_f \sum_{j=2}^{N} (j-1)P_j}$$

$$P_{\text{in}} = \frac{L_f \sum_{j=2}^{N} (j-1)P_j + \sum_{i=1}^{N} P_i}{1 - R - (N-1)L_f R}$$

As before, we must normalize the power at each element to the input power P_i/P_{in}. The coupling to the first element is $C_1 = P_1$ and the power left is $1 - P_1$. The transmission medium attenuates the signal between the first and second elements and we compute the power at the second element $= (1 - P_1)(1 - L_f)$. We determine the coupling value from the ratio

$$C_2 = \frac{P_2}{(1 - P_1)(1 - L_f)}$$

and the element removes P_2 power. The remaining power travels to the next element but is attenuated by $(1 - L_f)$. The power removed, P_2, is subtracted from the power at that point and the remaining power is attenuated before reaching the next extraction:

$$C_3 = \frac{P_3}{[(1 - P_1)(1 - L_f) - P_2](1 - L_f)}$$

$$C_4 = \frac{P_4}{\{[(1 - P_1)(1 - L_f) - P_2](1 - L_f) - P_3\}(1 - L_f)} \qquad \text{etc.}$$

The total loss due to attenuation is found from the sum of the normalized powers:

$$\text{loss(dB)} = 10 \log \left(\sum_{i=1}^{N} P_i \right)$$

Continuous Traveling Wave As the wave propagates along the antenna, it loses power continuously. The slots or holes must radiate more and more of the remaining power if the distribution is to be uniform. In general, the holes or slots must load the waveguide increasingly as the wave travels to the termination. The power at any point in the guide is

$$P(z) = P_0 \exp \left[-2 \int_0^z \alpha(z)\, dz \right] \tag{4-73}$$

where P_0 is the power at $z = 0$ and $\alpha(z)$ is the attenuation distribution (nepers/length). Suppose that we have a desired amplitude distribution, $A(z)$ (voltage):

$$P_{\text{in}} = \int_0^L |A(z)|^2\, dz + \int_0^L \rho_L(z)\, dz + P_{\text{load}} \tag{4-74}$$

where P_{load} is the power lost in the termination, $|A(z)|^2$ the radiated power distribution, and $\rho_L(z)$ the ohmic loss in the walls. Let the power into the termination be a ratio of the input power $P_{\text{load}} = R P_{\text{in}}$; then

$$P_{\text{in}} = \frac{1}{1 - R} \int_0^L \left[|A(z)|^2 + \rho_L(z) \right] dz \tag{4-75}$$

The power anywhere along the leaky wave antenna is

$$P(z) = P_{\text{in}} - \int_0^L |A(z)|^2 + \rho_L(z)\, dz \tag{4-76}$$

We differentiate this to get

$$\frac{dP(z)}{dz} = -[|A(z)|^2 + \rho_L(z)] \tag{4-77}$$

We differentiate Eq. (4-73) to relate $\alpha(z)$ to $P(z)$:

$$\frac{1}{P(z)} \frac{dP(z)}{dz} = -2\alpha(z) \tag{4-78}$$

We substitute Eq. (4-75) into Eq. (4-76) for P_{in}. By combining Eqs. (4-76) and (4-77) into Eq. (4-78), we derive the required attenuation distribution [18, p. 153]:

$$\alpha(z) = \frac{\frac{1}{2}|A(z)|^2}{[1/(1-R)]\displaystyle\int_0^L |A(z)|^2 + \rho_L(z)dz - \int_0^z |A(z)|^2 + \rho_L(z)dz} \tag{4-79}$$

If we assume a lossless transmission line, $\rho_L(z) = 0$ and Eq. (4-79) simplifies.

Example Design the attenuation distribution for a uniform distribution along a lossless transmission-line leaky wave antenna.

Substitute $A(z) = 1$ and $\rho_L(z) = 0$ into Eq. (4-79) and perform the integrations:

$$\alpha(z) = \frac{\frac{1}{2}}{[L/(1-R)] - z} = \frac{\frac{1}{2}(1-R)}{L[1 - z(1-R)/L]}$$

Given $R = 0.05$ (5% of the power into the load) for a structure with length 10λ. The initial and final attenuation constants are

$$\alpha_i(0) = \frac{0.95}{20} = 0.0475 \text{ Np}/\lambda \quad \text{or} \quad 0.413\,\text{dB}/\lambda$$

$$\alpha_f(L) = \frac{0.95}{2LR} = 0.95 \text{ Np}/\lambda \quad \text{or} \quad 8.25\,\text{dB}/\lambda$$

We reduce the variation between the initial and final values by dissipating more power in the termination. Given $R = 0.1$,

$$\alpha_i(0) = 0.045 \text{ Np}/\lambda \quad \text{or} \quad 0.39\,\text{dB}/\lambda$$

$$\alpha_f(L) = 0.45 \text{ Np}/\lambda \quad \text{or} \quad 3.9\,\text{dB}/\lambda$$

If we take the ratio of the attenuations at the ends, we have $\alpha(L)/\alpha(0) = 1/R$.

We can normalize Eq. (4-79) to the interval ± 2 and use the linear distributions given above where $x = z/L$ and $\rho_L(z) = 0$. Figure 4-26 shows the attenuation distribution for a Taylor distribution with 30-dB sidelobes and $\overline{n} = 8$ for various levels of power dissipation in the load. Table 4-28 lists the bounds on $\alpha(x)L$ for various Taylor distributions. Changing the number of modified zeros has only a minor effect on the bounds. A $\cos^2$ on a pedestal distribution with a 30-dB sidelobe level has very similar bounds on the attenuation. The 40-dB sidelobe level design requires a greater variation of attenuation than the 30-dB cases. Long structures may not be able to provide the low levels of radiation above the ohmic losses for an effective design. In all cases we decrease the attenuation range on an antenna by decreasing the antenna efficiency though absorbing more power in the termination.

4-16 CIRCULAR APERTURES

Many common apertures conform to circles. The two-dimensional Fourier transform relation for the pattern holds for any aperture rim shape and becomes for the circle

$$f(\theta, \phi) = \int_0^{2\pi} \int_0^a E(r', \phi') e^{jkr' \sin\theta \cos(\phi - \phi')} r'\, dr'\, d\phi' \tag{4-80}$$

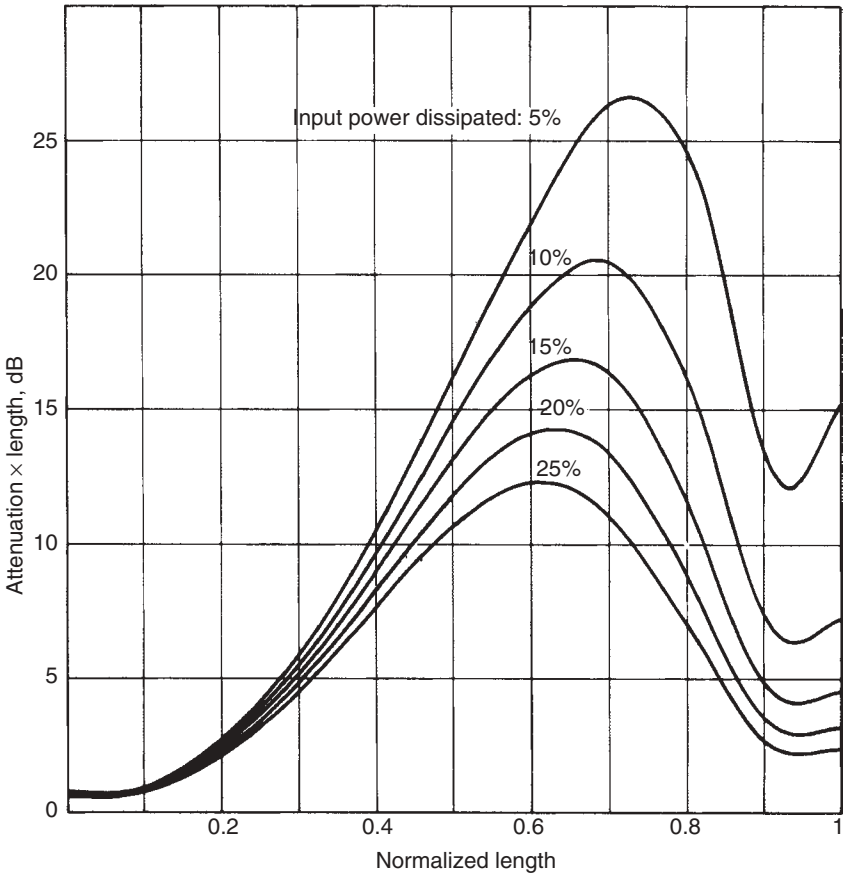


FIGURE 4-26 Leaky wave attenuation distribution for Taylor distribution with 30-dB sidelobes, $\overline{n} = 8$.

TABLE 4-28 Maximum and Minimum Normalized Attenuation $\alpha(z)L$ of a Leaky Wave Taylor Distribution

Termination Power (%)	30 dB				40 dB, $\overline{n} = 8$	
	$\overline{n} = 6$		$\overline{n} = 12$			
	Maximum	Minimum	Maximum	Minimum	Maximum	Minimum
5	27.08	0.59	26.04	0.63	31.61	0.12
6	25.38	0.58	24.54	0.63	29.63	0.12
8	22.70	0.57	22.06	0.61	26.52	0.12
10	20.66	0.56	20.08	0.60	24.12	0.11
12	19.00	0.55	18.46	0.59	22.18	0.11
15	17.00	0.53	16.52	0.57	19.81	0.11
20	14.42	0.50	14.04	0.53	16.78	0.10
25	12.42	0.46	12.09	0.50	14.44	0.10

where a is the radius, r' the radial coordinate, and ϕ' the angle coordinate of the aperture point. The integral leads to a k_r-space.

When the distribution has circular symmetry, the ϕ' integral can be evaluated easily, which reduces Eq. (4-80) to

$$f(k_r) = 2\pi \int_0^a E(r') J_0(kr' \sin\theta) r' \, dr' \tag{4-81}$$

where $J_0(x)$ is the zeroth-order Bessel function of the first kind. All great-circle patterns (constant ϕ) are identical. For a uniform distribution,

$$f(k_r) = \frac{2J_1(ka \sin\theta)}{ka \sin\theta}$$

plotted in Figure 4-27. The zeros occur at the zeros of $J_1(x)$. The 3-dB pattern point of the uniform distribution is

$$\begin{aligned} ka \sin\theta_1 &= 1.6162 \qquad \sin\theta_1 = 0.5145\frac{\lambda}{D} \\ \text{HPBW} &= 2\sin^{-1}\frac{0.5145\lambda}{D} \end{aligned} \tag{4-82}$$

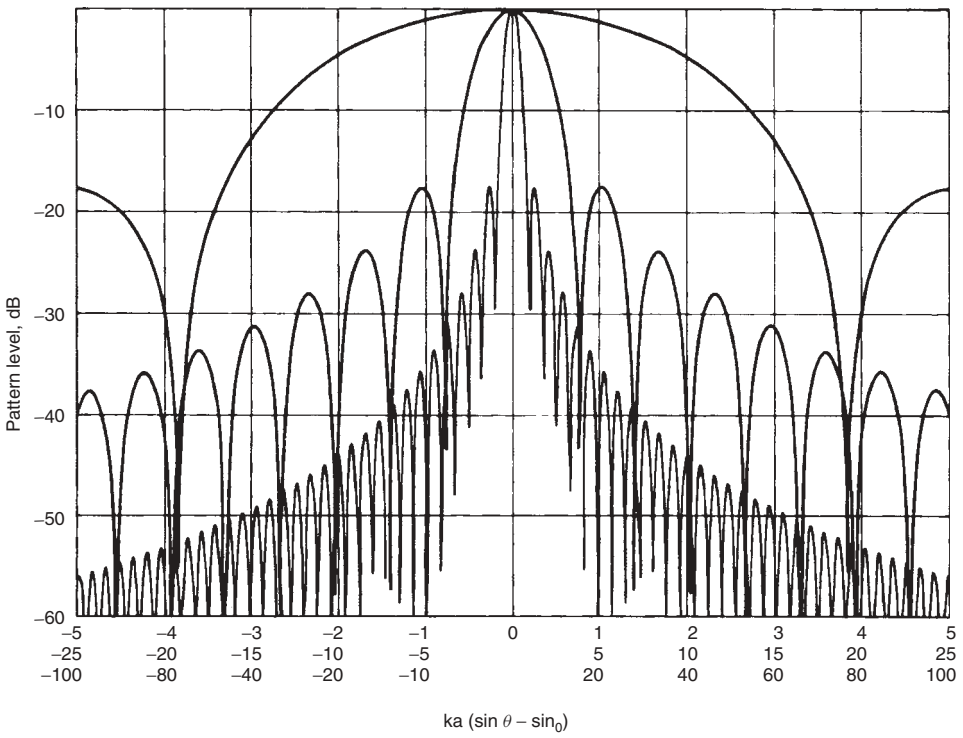


FIGURE 4-27 k_r-space pattern of uniform circular aperture distribution.

where D is the diameter. For large apertures we can approximate $\sin\theta$ by θ (rad). Converted to degrees, the half-power beamwidth becomes

$$\text{HPBW} = 58.95^\circ \frac{\lambda}{D} \tag{4-83}$$

Example Compute the beamwidth of a uniform distribution circular aperture with 10.5λ diameter.

The beamwidths are found from Eqs. (4-82) and (4-83):

$$\text{HPBW} = 2\sin^{-1}\frac{0.5145}{10.5} = 5.62^\circ$$

$$\text{HPBW} = \frac{58.95^\circ}{10.5} = 5.61^\circ$$

The first zero of $J_1(x)$ gives the k-space pattern null point.

$$\begin{aligned} ka\sin\theta_{\text{null}} &= 3.8317 \\ \text{BW}_{\text{null}} &= 2\sin^{-1}\frac{1.2197\lambda}{D} \simeq 139.76^\circ\frac{\lambda}{D} \end{aligned} \tag{4-84}$$

We can also define a null beamwidth factor and relate the beams of other distributions to the uniform circular distribution beamwidth [Eq. (4-84)]. All other circular distributions relate to Eq. (4-82) or (4-83) through a beamwidth factor. The uniform distribution has a unity beamwidth factor.

4-17 CIRCULAR GAUSSIAN DISTRIBUTION [19]

A truncated Gaussian distribution has a simple functional relation:

$$E(r) = e^{-\rho r^2} \qquad |r| \le 1 \tag{4-85}$$

We can easily calculate the edge taper through the conversion between logarithms:

$$\text{edge taper(dB)} = 8.686\rho \tag{4-86}$$

We determine amplitude taper efficiency by substituting Eq. (4-85) into Eq. (4-8) and carrying out the integrations:

$$\text{ATL} = \frac{2(1-e^{-\rho})^2}{\rho(1-e^{-2\rho})} \tag{4-87}$$

Table 4-29 lists designs for various sidelobe levels in terms of the single parameter: edge taper. Equation (4-86) relates the parameter ρ to the edge taper.

Example Estimate the beamwidth of the pattern radiated from a circular distribution with a 13-dB edge taper and radius of three wavelengths.

TABLE 4-29 Circular-Aperture Gaussian Distribution, $e^{-\rho r^2}$ ($|r| < 1$)

Sidelobe Level (dB)	Edge Taper (dB)	ATL (dB)	Beamwidth Factor
20	4.30	0.09	1.0466
22	7.18	0.24	1.0800
24	9.60	0.41	1.1109
25	10.67	0.50	1.1147
26	11.67	0.59	1.1385
28	13.42	0.76	1.1626
30	14.93	0.92	1.1839
32	16.23	1.06	1.2028
34	17.32	1.18	1.2188
35	17.81	1.23	1.2263
36	18.75	1.34	1.2405
38	21.43	1.65	1.2820
40	24.42	2.00	1.3296

We use linear interpolation in Table 4-29 to determine the beamwidth factor. From Eq. (4-82),

$$\text{HPBW} = 2\sin^{-1}\frac{1.1568(0.5145)}{6} = 11.38^\circ$$

From Eq. (4-83),

$$\text{HPBW} = 58.95^\circ\left(\frac{1.1568}{6}\right) = 11.36^\circ$$

The amplitude taper efficiency is calculated from Eq. (4-87):

$$\rho = \frac{13}{8.686} = 1.497$$

$$\text{ATL} = \frac{2(1-e^{-1.497})^2}{1.497(1-e^{-2.993})} = 0.847 \quad (-0.72\,\text{dB})$$

We obtain the same value by interpolating in Table 4-29.

Sidelobes below 40 dB are difficult to obtain with this distribution. The inner sidelobes continue to decrease with a decreasing edge level, but the outer lobes fail to reduce and dominate over the first few sidelobes. Table 4-29 results from a search because no direct method exists for computing the edge taper for a specified sidelobe level.

4-18 HANSEN SINGLE-PARAMETER CIRCULAR DISTRIBUTION [20, 21]

This distribution leads directly from sidelobe level to a single parameter H that relates through closed-form expressions to all other distribution parameters. The pattern of a uniform distribution is modified close in to the main beam. By using the U-space

variable of Taylor, we have $U = (2a/\lambda)\sin\theta$, where a is the radius. The pattern has different expressions in two regions:

$$f(U) = \begin{cases} \dfrac{2I_1(\pi\sqrt{H^2 - U^2})}{\pi\sqrt{H^2 - U^2}} & |U| \le H \qquad (4\text{-}88a) \\ \dfrac{2J_1(\pi\sqrt{U^2 - H^2})}{\pi\sqrt{U^2 - H^2}} & |U| \ge H \qquad (4\text{-}88b) \end{cases}$$

$I_1(x)$ is the first-order modified Bessel function of the first kind.

The high function value of Eq. (4-88a) at the boresight reduces the sidelobes of the uniform distribution [Eq. (4-88b)], 17.57 dB, below the level at $U = H$. The sidelobe level is

$$\text{SLR} = 17.57 + 20\log\frac{2I_1(\pi H)}{\pi H} \qquad (4\text{-}89)$$

Given the sidelobe level [positive (dB)], we use Eq. (4-89) in an iteration scheme to determine H.

The aperture distribution is given by

$$E(r) = I_0(\pi H\sqrt{1 - r^2}) \qquad |r| \le 1 \qquad (4\text{-}90)$$

where I_0 is the zeroth-order modified Bessel function of the first kind. Equation (4-8) can be integrated for this circularly symmetrical distribution [Eq. (4-90)] to derive the amplitude taper efficiency:

$$\text{ATL} = \frac{4I_1^2(\pi H)}{\pi^2 H^2[I_0^2(\pi H) - I_1^2(\pi H)]} \qquad (4\text{-}91)$$

Table 4-30 lists the parameters of the Hansen distribution for various sidelobe levels. At the top, Tables 4-29 and 4-30 are very similar. Any sidelobe level can be achieved with this distribution, subject to tolerance problems generated by any low-sidelobe design. The distribution is not optimum, but it is convenient.

4-19 TAYLOR CIRCULAR-APERTURE DISTRIBUTION [22]

Similar to the line source, the Taylor circular-aperture distribution modifies inner zeros of the uniform amplitude and phase circular-aperture k-space pattern to approximate the Dolph–Chebyshev distribution. By use of the variable $\pi U = ka\sin\theta$ the uniform distribution pattern is found to be $J_1(\pi U)/\pi U$. We remove $\overline{n} - 1$ inner zeros and add those of the Dolph–Chebyshev distribution:

$$f(U) = \frac{J_1(\pi U)\prod_{N=1}^{\overline{n}-1}(1 - U^2/U_N^2)}{\pi U\prod_{N=1}^{\overline{n}-1}(1 - U^2/S_N^2)} \qquad (4\text{-}92)$$

TABLE 4-30 Hansen Single-Parameter Circular-Aperture Distribution

Sidelobe Level (dB)	H	Edge Taper (dB)	ATL (dB)	Beamwidth Factor
20	0.48717	4.49	0.09	1.0484
22	0.66971	7.79	0.27	1.0865
24	0.82091	10.87	0.48	1.1231
25	0.88989	12.35	0.60	1.1409
26	0.95573	13.79	0.72	1.1584
28	1.08027	16.59	0.96	1.1924
30	1.19770	19.29	1.19	1.2252
32	1.30988	21.93	1.42	1.2570
34	1.41802	24.51	1.64	1.2876
35	1.47084	25.78	1.75	1.3026
36	1.52295	27.04	1.85	1.3174
38	1.62525	29.53	2.05	1.3462
40	1.72536	31.98	2.24	1.3742
45	1.96809	38.00	2.68	1.4410
50	2.20262	43.89	3.08	1.5039

Given a zero of $J_1(x)$, $J_1(x_{1N}) = 0$, let $x_{1N} = \pi S_N$. By retaining approximately the same number of zeros in the visible region as in the uniform distribution, we avoid superdirectivity. The new zeros U_N are modified zeros of the uniform distribution:

$$U_N = S_{\overline{n}} \frac{\sqrt{A^2 + (N - \frac{1}{2})^2}}{\sqrt{A^2 + (\overline{n} - \frac{1}{2})^2}} \tag{4-93}$$

where A relates to the maximum sidelobe level, $\cosh \pi A = b$ and $20 \log b =$ sidelobe level(dB). Equation (4-93) is the same as Eq. (4-20) except for the scaling constant $S_{\overline{n}}$, the $\overline{n}$th zero of $J_1(x)$ divided by π.

Equation (4-92) gives the U-space pattern of the new distribution. We expand the aperture distribution in a Fourier–Bessel series:

$$E(r) = \sum_{m=0}^{\overline{n}-1} B_m J_0(\pi S_m r) \qquad r \le 1 \tag{4-94}$$

We compute coefficients B_m by transforming the Fourier–Bessel series [Eq. (4-94)] into U-space and comparing the far-field pattern with Eq. (4-92). As indicated in Eq. (4-94), the series contains only $\overline{n}$ nonzero terms:

$$\begin{aligned} B_0 &= 1 \\ B_m &= \frac{-\prod_{N=1}^{\overline{n}-1} (1 - S_m^2/U_N^2)}{J_0(\pi S_m) \prod_{N=1, N\neq m}^{\overline{n}-1} (1 - S_m^2/S_N^2)} \qquad m = 1, 2, \ldots, \overline{n} - 1 \end{aligned} \tag{4-95}$$

Example Design a Taylor circular-aperture distribution with 30 dB maximum sidelobes and $\overline{n} = 6$.

We use Eq. (4-21) to calculate the constant A:

$$b = 10^{30/20} = 31.6228$$
$$A = \frac{\cosh^{-1} b}{\pi} = 1.32$$

We substitute this value into Eq. (4-93) to compute the five nulls:

No.	1	2	3	4	5
Null U_N	1.5582	2.2057	3.1208	4.1293	5.1769

The first null of the uniform distribution occurs at

$$x_{11} = 3.83171 \qquad S_1 = \frac{x_{11}}{\pi} = 1.2197$$

We use this with the location of the first zero to determine the null beamwidth factor:

$$\text{BW}_{\text{null}} = \frac{U_1}{S_1} = \frac{1.5582}{1.2197} = 1.2775$$

The coefficients of the Fourier series [Eq. (4-95)] are given in Table 4-31. Figure 4-28 contains the k-space pattern.

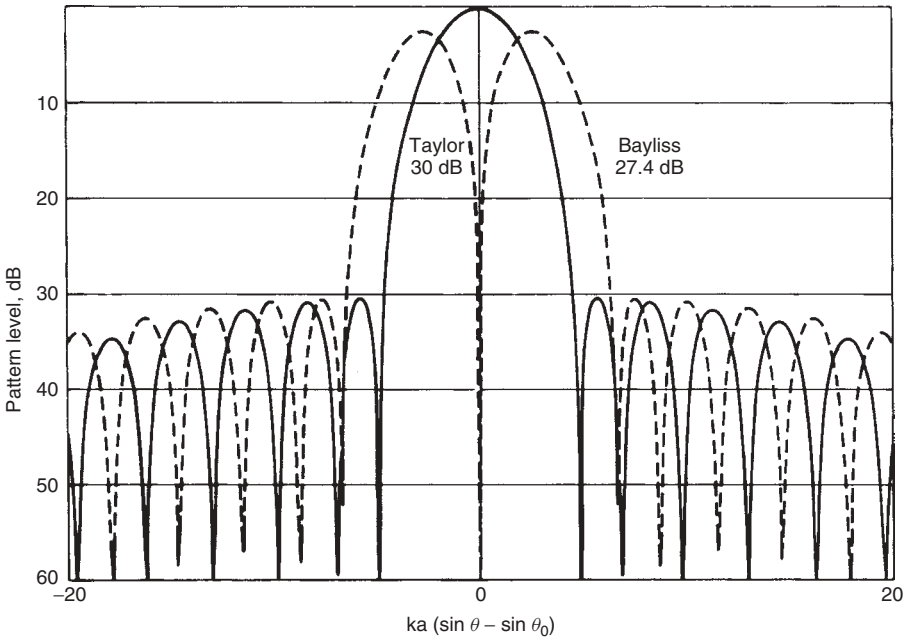


FIGURE 4-28 Taylor and Bayliss circular aperture distributions to give 30-dB sidelobes ($\overline{n} = 6$).

Tables 4-32 to 4-34 list the characteristics for a few designs of the circular Taylor distribution. Table 4-32 shows that for each sidelobe level there is an optimum $\overline{n}$. As the sidelobes are lowered, the optimum value of $\overline{n}$ increases. The blanks are unsuitable designs. The beamwidth factor (Table 4-33) and the null beamwidth factor (Table 4-34) continue to decrease as $\overline{n}$ increases at a given sidelobe level. In all three tables the values depend primarily on the sidelobe level.

TABLE 4-31 Fourier–Bessel Series Coefficients for Taylor Distribution: 30 dB, $\overline{n} = 6$

No.	B_m	B_m Normalized	Function
0	1.0000	0.53405	1
1	0.93326	0.49841	$J_0(x_{11}r)$
2	0.038467	0.01808	$J_0(x_{12}r)$
3	−0.16048	−0.08570	$J_0(x_{13}r)$
4	0.16917	0.09035	$J_0(x_{14}r)$
5	−0.10331	−0.05517	$J_0(x_{15}r)$

TABLE 4-32 Amplitude Taper Losses of Taylor Circular-Aperture Distribution (dB)

	Sidelobe Level (dB)					
$\overline{n}$	25	30	35	40	45	50
4	0.30	0.71	1.14	1.51	1.84	
6	0.28	0.59	1.03	1.48	1.88	2.23
8	0.43	0.54	0.94	1.40	1.82	2.21
12	1.03	0.62	0.86	1.28	1.71	2.12
16	1.85	0.86	0.87	1.22	1.64	2.05
20		1.20	0.94	1.20	1.60	2.01

TABLE 4-33 Beamwidth Factor of Taylor Circular-Aperture Distribution

	Sidelobe Level (dB)					
$\overline{n}$	25	30	35	40	45	50
4	1.0825	1.1515	1.2115	1.2638	1.3095	
6	1.0504	1.1267	1.1957	1.2581	1.3149	1.3666
8	1.0295	1.1079	1.1796	1.2457	1.3067	1.3632
12	1.0057	1.0847	1.1580	1.2262	1.2899	1.3499
16	0.9927	1.0717	1.1451	1.2137	1.2782	1.3391
20		1.0634	1.1367	1.2054	1.2701	1.3314

TABLE 4-34 Null Beamwidth Factor of Taylor Circular-Aperture Distribution

	Sidelobe Level (dB)					
$\overline{n}$	25	30	35	40	45	50
4	1.1733	1.3121	1.4462	1.5744	1.6960	
6	1.1318	1.2775	1.4224	1.5654	1.7056	1.8426
8	1.1066	1.2530	1.4001	1.5470	1.6928	1.8370
12	1.0789	1.2244	1.3716	1.5197	1.6680	1.8162
16	1.0643	1.2087	1.3552	1.5029	1.6514	1.8003
20		1.1989	1.3442	1.4920	1.6402	1.7890

4-20 BAYLISS CIRCULAR-APERTURE DISTRIBUTION [8]

We can also design a Bayliss distribution (difference pattern) for circular apertures. This gives us the pattern necessary for monopulse tracking along one axis. The U-space pattern has modified zeros to produce nearly equal sidelobes close in to the main lobes:

$$f(U,\phi) = \cos\phi\,\pi U J_1'(\pi U)\frac{\prod_{N=1}^{\overline{n}-1}(1-U^2/U_N^2)}{\prod_{N=0}^{\overline{n}-1}(1-U^2/\mu_N^2)} \tag{4-96}$$

where U_N are the new zeros and $\pi\mu_N$ are zeros of $J_1'(\pi U)$. Bayliss lists those zeros μ_N (Table 4-35). The inner zeros have been removed and replaced by new ones, U_N. We compute the zeros in a manner similar to that used for a linear distribution (Section 4-7):

$$U_N = \begin{cases} \mu_{\overline{n}}\sqrt{\dfrac{\xi_N^2}{A^2+\overline{n}^2}} & N = 1,2,3,4 \\ \mu_{\overline{n}}\sqrt{\dfrac{A^2+N^2}{A^2+\overline{n}^2}} & N = 5,6,\ldots,\overline{n}-1 \end{cases} \tag{4-97}$$

The four inner zeros had to be adjusted to achieve the desired sidelobe level. Bayliss found these through a computer search. The values for ξ_N and A can be found through the polynomial approximations [Eq. (4-40)].

TABLE 4-35 Bessel Function Zeros, $J_1'(\pi\mu_N)$

N	μ_N	N	μ_N	N	μ_N	N	μ_N
0	0.5860670	5	5.7345205	10	10.7417435	15	15.7443679
1	1.6970509	6	6.7368281	11	11.7424475	16	16.7447044
2	2.7171939	7	7.7385356	12	12.7430408	17	17.7450030
3	3.7261370	8	8.7398505	13	13.7435477	18	18.7452697
4	4.7312271	9	9.7408945	14	14.7439856	19	19.7455093

Like the Taylor circular aperture distribution, the aperture distribution is expanded in a finite-length Fourier–Bessel series:

$$E(r,\phi) = \cos\phi' \sum_{m=0}^{\overline{n}-1} B_m J_1(\pi\mu_m r) \qquad r \le 1 \tag{4-98}$$

where the coefficients are found by transforming Eq. (4-98) and comparing it with a U-space pattern [Eq. (4-96)]. The coefficients are given by

$$B_m = \frac{\mu_m^2}{j J_1(\pi\mu_m)} \frac{\prod_{N=1}^{\overline{n}-1}(1-\mu_m^2/U_N^2)}{\prod_{N=0,N\neq m}^{\overline{n}-1}(1-\mu_m^2/\mu_N^2)} \tag{4-99}$$

Example Design a Bayliss circular-aperture distribution with 30-dB sidelobes and $\overline{n} = 6$.

We start with Eq. (4-40) to compute coefficients A and ξ_N:

$$A = 1.64126 \qquad \xi_1 = 2.07086 \qquad \xi_2 = 2.62754 \qquad \xi_3 = 3.43144 \qquad \xi_4 = 4.32758$$

We substitute these constants into Eq. (4-97) along with the zeros from Table 4-35 to calculate the modified zeros:

N	1	2	3	4	5
U_N	2.2428	2.8457	3.7163	4.6868	5.6994

We use the zeros in Eq. (4-96) to calculate the pattern. The U-space pattern peak can be found by using Eq. (4-41):

$$U_{\max} = 0.7988 \qquad ka\sin\theta_{\max} = \pi U_{\max} = 2.5096$$

where a is the aperture radius. The coefficients of the Fourier–Bessel series are found from Eq. (4-99) (Table 4-36). The normalized coefficients give an aperture distribution peak of 1. The 3-dB pattern points can be found by searching the pattern:

$$ka\sin\theta_1 = 1.3138 \qquad ka\sin\theta_2 = 4.2384$$

TABLE 4-36 Fourier–Bessel Series Coefficients for Bayliss Distribution: 30 dB, $\overline{n} = 6$

No.	B_m	B_m Normalized	Function
0	0.62680	1.2580	$J_1(\pi\mu_0 r)$
1	0.50605	1.0157	$J_1(\pi\mu_1 r)$
2	−0.06854	−0.03415	$J_1(\pi\mu_2 r)$
3	−0.0028703	−0.005761	$J_1(\pi\mu_3 r)$
4	0.014004	0.028106	$J_1(\pi\mu_4 r)$
5	−0.011509	−0.02310	$J_1(\pi\mu_5 r)$

TABLE 4-37 Characteristics of a Bayliss Circular-Aperture Distribution, $\overline{n} = 10$

Sidelobe Level (dB)	Beam Peak, $ka/2 \sin \theta_{\max}$	3-dB Beam Edge		ATL (dB)	PEL (dB)
		$ka/2 \sin \theta_1$	$ka/2 \sin \theta_2$		
20	2.2366	1.165	3.700	1.47	1.80
25	2.3780	1.230	3.940	1.15	1.89
30	2.5096	1.290	4.160	1.32	1.96
35	2.6341	1.346	4.363	1.62	2.01
40	2.7536	1.399	4.551	1.95	2.05

Figure 4-28 contains a plot of a Bayliss circular-aperture distribution ($\overline{n} = 6$) designed to have sidelobes 30 dB below those of the Taylor distribution with 30-dB sidelobes. The losses to the difference pattern are about 2.6 dB higher than the sum pattern. The amplitude taper efficiency is calculated from

$$ATL = \frac{\left[\int_0^1 4 \left| \sum_{m=0}^{\overline{n}-1} B_m J_1(\pi \mu_m r) \right| r \, dr \right]^2}{\pi^2 \int_0^1 4 \left| \sum_{m=0}^{\overline{n}-1} B_m J_1(\pi \mu_m r) \right|^2 r \, dr} \tag{4-100}$$

where the integrals over ϕ' have been separated and evaluated. An integral expression for the phase error efficiency can be found similarly by evaluating the separable $\cos \phi'$ integrals along the coordinate $\phi = 0$, the peak:

$$\text{PEL}(U) = \frac{\left| 2\pi \int_0^1 \sum_{m=0}^{\overline{n}-1} B_m J_1(\pi \mu_m r) J_1(\pi U r) r \, dr \right|^2}{\left[\int_0^1 4 \left| \sum_{m=0}^{\overline{n}-1} B_m J_1(\pi \mu_m r) \right| r \, dr \right]^2} \tag{4-101}$$

Table 4-37 lists the parameters of Bayliss circular-aperture distributions with $\overline{n} = 10$ and various sidelobe levels. The optimum design for $\overline{n} = 10$ occurs for 25-dB sidelobes.

4-21 PLANAR ARRAYS

We design planar arrays with nearly circular boundaries by sampling circular distributions. Given enough sample points in the array, a distribution such as the circular Taylor will be modeled adequately to produce a similar pattern. We can use pattern multiplication to combine the designs for linear arrays into a planar array, but in the special case of a square array, a true Chebyshev design can be obtained in all planes. A technique has been developed to allow the synthesis from pattern nulls provided that some of the possible nulls are not specified. We are still left with the problem of specifying the numerous nulls possible with a planar array.

Chebyshev Array [23] When we combine two Dolph–Chebyshev linear arrays through pattern multiplication, it produces a pattern that has lower sidelobes than

those specified in all planes except the principal ones along the axes. These designs give beamwidths in the diagonal planes that are wider than necessary. The pattern deviates from the optimum because sidelobes are suppressed more than necessary.

We use a technique on a square array to produce equal sidelobes in all constant ϕ cuts around the array. The array is square in the number of elements, but different spacings along the axes can produce a rectangular array. We expand the pattern in a single Chebyshev polynomial:

$$T_{L-1}(x_0 \cos \psi_1 \cos \psi_2) \tag{4-102}$$

where $\psi_1 = kd_x \cos\theta \cos\phi + \delta_1$ and $\psi_2 = kd_y \cos\theta \sin\phi + \delta_2 L = 2N$ or $L = 2N + 1$ for L^2 elements in the array. We compute x_0 from Eq. (4-51) for a given sidelobe level. The pattern for an odd number of elements in each row and column is

$$E(\theta, \phi) = \sum_{m=1}^{N+1} \sum_{n=1}^{N+1} \varepsilon_m \varepsilon_n I_{mn} \cos 2(m-1)\psi_1 \cos 2(n-1)\psi_2 \qquad L = 2n + 1$$

where $\varepsilon_m = 1$ for $m = 1$ and $\varepsilon_m = 2$ for $m \neq 1$. Similarly,

$$E(\theta, \phi) = 4 \sum_{m=1}^{N} \sum_{n=1}^{N} I_{mn} \cos(2m-1)\psi_1 \cos(2n-1)\psi_2 \qquad L = 2n$$

The element excitations I_{mn} are given by

$$I_{mn} = \left(\frac{2}{L}\right)^2 \sum_{p=1}^{N+1} \sum_{q=1}^{N+1} \varepsilon_p \varepsilon_q T_{L-1}\left[x_0 \cos\frac{(p-1)\pi}{L} \cos\frac{(q-1)\pi}{L}\right] \times \cos\frac{2\pi(m-1)(p-1)}{L} \cos\frac{2\pi(n-1)(q-1)}{L} \qquad L = 2N+1 \tag{4-103}$$

or

$$I_{mn} = \left(\frac{4}{L}\right)^2 \sum_{p=1}^{N} \sum_{q=1}^{N} T_{L-1}\left[x_0 \cos\frac{\left(p-\frac{1}{2}\right)\pi}{L} \cos\frac{\left(q-\frac{1}{2}\right)\pi}{L}\right] \times \cos\frac{2\pi\left(m-\frac{1}{2}\right)\left(p-\frac{1}{2}\right)}{L} \cos\frac{2\pi\left(n-\frac{1}{2}\right)\left(q-\frac{1}{2}\right)}{L} \qquad L = 2N \tag{4-104}$$

4-22 CONVOLUTION TECHNIQUE FOR PLANAR ARRAYS

We may synthesize a desired pattern through multiplication of two or more simpler patterns. Because patterns derive from Fourier transforms of distributions in space, the distribution to produce the product of two simpler patterns is the convolution of the

simpler distributions [24, p. 30]. We find it easier to synthesize by using a few elements and then build up patterns through multiplication.

Consider the convolution of a linear array with another linear array on the same axis. We describe the array as a distribution consisting of weighted impulse functions, $\delta(x - x_i)$:

$$A_1(x) = \sum_{i=1}^{N_1} a_{1i}\delta(x - x_i)$$

where a_{1i} are the feeding coefficients and x_i are the locations for an N_1-element array. To determine the array that gives the product of two array patterns, we convolve the second array, $A_2(x)$, with the first:

$$A_1(x) * A_2(x) = \int A_1(\tau)A_2(x - \tau)\, d\tau \tag{4-105}$$

We evaluate a function at the argument of the impulse function when we convolve the two arrays [25, p. 237]. Equation (4-105) reduces to

$$A_1(x) * A_2(x) = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} a_{1i}a_{2i}\delta(x - x_i - x_j) \tag{4-106}$$

Example Consider the two 2-element arrays in Figure 4-29 and the graphical solution of the convolution. Figure 4-29*a* shows the location of the elements in the arrays on the x-axis. To perform the convolution, we reflect the x-axis of one array and move it across the other array while performing the integral at each location x, the coordinate of the convolution. We have a net result to the integral only when two impulse functions are aligned, $x = x_i + x_j$. We have four elements in the resulting array (Figure 4-29*c*). If the elements are equally spaced in the two arrays, two elements will sum into one.

Patterns are the result of a three-dimensional Fourier transform. For a general array with element locations $\mathbf{r}_i$, we must perform a convolution along all three axes to find the distribution that gives the product of the patterns of two simpler distributions. For the general array, Eq. (4-106) becomes

$$A_1(\mathbf{r}) * A_2(\mathbf{r}) = \sum\sum a_{1i}a_{2j}\delta(\mathbf{r} - \mathbf{r}_i - \mathbf{r}_j) \tag{4-107}$$

where $\mathbf{r}$ is the location vector and $\mathbf{r}_i$ and $\mathbf{r}_j$ are the locations of elements in the two arrays.

A rectangular array can be described as the convolution of a linear array on the x-axis with a linear array on the y-axis. When $y = y_j$ there is a string of values $x = x_i$ that satisfy the impulse argument [Eq. (4-107)]. These are the locations of the elements. We step through all values of y_j until the entire array is formed. Equation (4-107) gives the feeding coefficient of each element $a_{1i}a_{2j}$ since no two elements of the convolution are in the same place. The pattern of the rectangular array is the product of the linear array patterns along the axes.

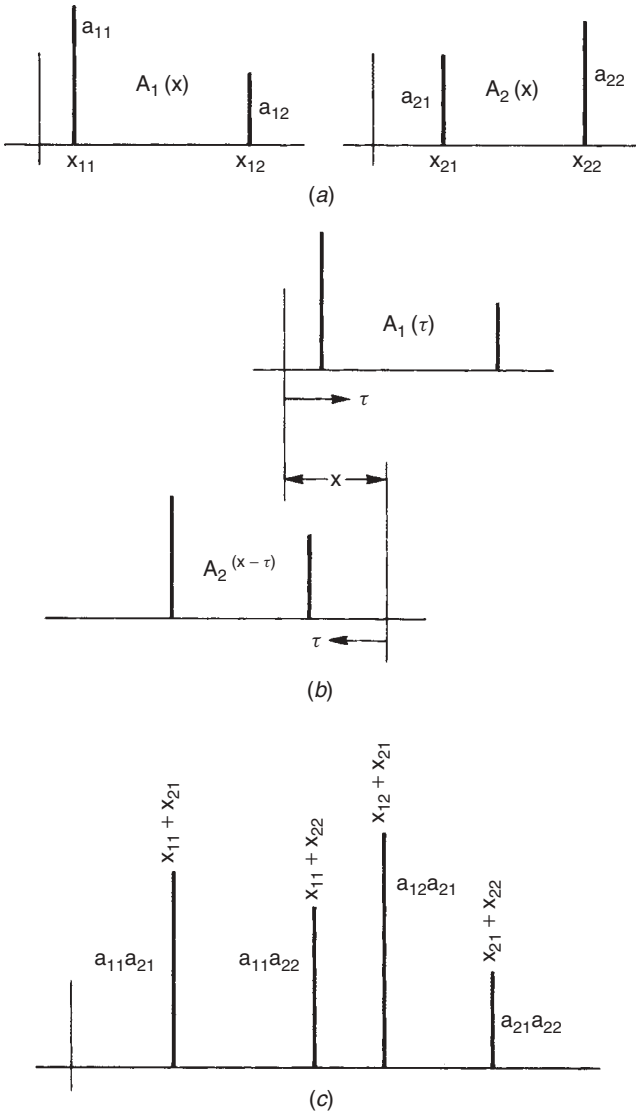


FIGURE 4-29 Convolution of two linear arrays: (*a*) separate arrays; (*b*) graphical convolution; (*c*) convolution.

Given an array, we compute the pattern from a Fourier transform containing N terms each of which corresponds to one element. Ignoring the element pattern, we have

$$E = \sum_{i=1}^{N} a_i e^{j\mathbf{k}\cdot\mathbf{r}_i'} \tag{4-108}$$

The array has $N - 1$ independent nulls (zeros) in the pattern. Given the set of nulls $\mathbf{k}_j$ we can substitute them into Eq. (4-108) to form a matrix equation in $N - 1$ unknowns a_i. We must normalize one coefficient, $a_i = 1$, to solve the set of equations for the

feeding coefficients:

$$[B]\begin{bmatrix} a_2 \\ a_3 \\ \vdots \\ a_N \end{bmatrix} = -\begin{bmatrix} e^{j\mathbf{k}_1\cdot\mathbf{r}_1} \\ e^{j\mathbf{k}_2\cdot\mathbf{r}_1} \\ \vdots \\ e^{j\mathbf{k}_{N-1}\cdot\mathbf{r}_1} \end{bmatrix} \tag{4-109}$$

where $b_{ij} = e^{j\mathbf{k}_i\cdot\mathbf{r}_{j+1}}$. We find the direct solution of Eq. (4-109) unwieldy for a large array. We can subdivide the array into smaller arrays whose convolution is the total array and use pattern multiplication. We reduce the number of nulls we need to specify in the synthesis of an array.

Convolution can be used in the synthesis of planar arrays by using a rhombic array (four elements) as the basic building block [26] (Figure 4-30). If we convolve two identically shaped rhombic arrays, we obtain a nine-element (three on a side) array (Figure 4-30*b*). By continuing to convolve the resulting array with other rhombic arrays, we can build up a large array in the shape of the rhombus. Each rhombic array has three pattern nulls without the symmetry of the linear array about some axis. The rhombic array has symmetry only about the plane of the rhombus. We build up an array of $N + 1$ by $N + 1$ elements through the convolution of N rhombic arrays. The original array has $(N + 1)(N + 1) - 1$ independent nulls. The convolution of N rhombic arrays reduces the number of independent nulls to $3N$. Similarly, when we use the convolution of two linear arrays to form a square array, $N + 1$ by $N + 1$, the number of independent nulls is $2N$, or N for each array.

We denote a single rhombic array as RA_1 and the convolution of two rhombic arrays as RA_2. The number of elements on each side of an RA_N array is $N + 1$. We can convolute a rhombic array with a linear array to form an $M \times N$ array ($M > N$). Denote the linear array by L_N, where the array has $N + 1$ elements. The planar array $PA_{M,N}$ becomes

$$PA_{M,N} = L_{M-N} * RA_{N-1} \tag{4-110}$$

We specify $3(N - 1)$ nulls in space for the rhombic arrays and $M - N$ nulls about the axis of the array. Like all convolutions, the pattern is the product of the individual array patterns.

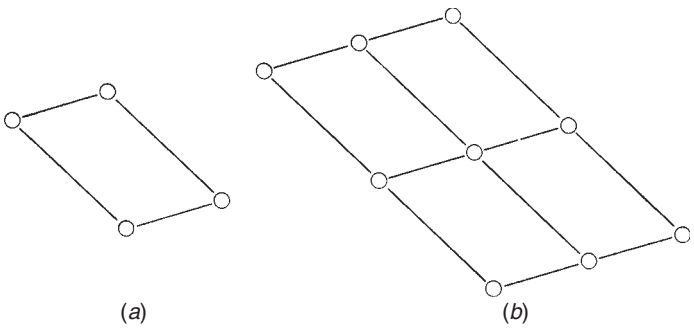


FIGURE 4-30 Rhombic array with its convolution: (*a*) rhombic array RA_1; (*b*) convolution of two rhombic arrays RA_2.

This method allows the specification of nulls in space with other than linear symmetry. Second, it reduces the required specification of nulls. Third, it provides a method for synthesis of triangularly or hexagonally spaced elements.

Example Consider the six-element rectangular array shown in Figure 4-31*a*. It can be broken down into the convolution of a four-element rectangular array (rhombic) and a two-element linear array from Eq. (4-110):

$$PA_{3,2} = L_1 * RA_1$$

Pick the three nulls of the rhombic array at

θ	90°	90°	90°
ϕ	110°	-60°	180°

We measure the pattern nulls from the normal of the plane containing the rhombus and the x-axis (ϕ). For an array of broadcast towers, the nulls point toward the horizon. We restrict θ to less than or equal to 90°. We substitute the positions of the elements (Figure 4-31*b*) and the nulls into Eq. (4-109) to solve for the feeding coefficients of the rhombic array (Table 4-38).

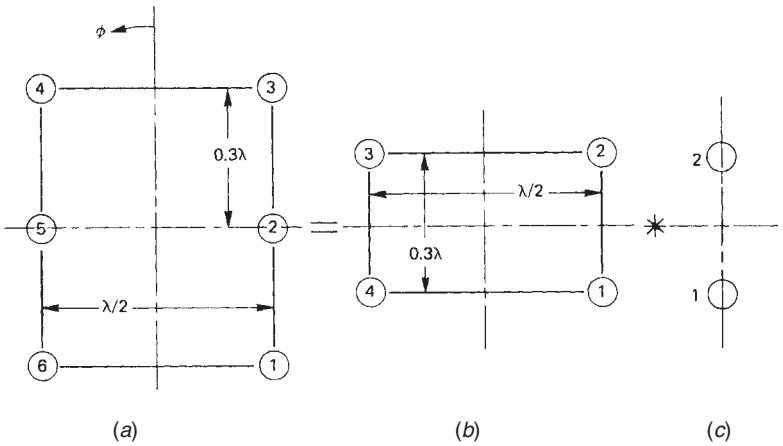


FIGURE 4-31 Rectangular array from convolution of rhombic and linear arrays: (*a*) six-element rectangular array; (*b*) rhombic array; (*c*) linear array.

TABLE 4-38 Coefficients of Rhombic Array for Horizon ($\theta = 90^\circ$) Nulls at $\phi = 100^\circ, -60^\circ$, and 180°

Element	Amplitude (dB)	Phase (deg)
1	0.00	0.0
2	4.12	−79.2
3	0.00	−109.2
4	4.12	−30.2

TABLE 4-39 Coefficients of Six-Element Rectangular Array with Pattern of Figure 4-32*b*

Element	Amplitude (dB)	Phase (deg)
1	0.00	0.0
2	8.13	−88.5
3	4.12	177.2
4	0.00	147.2
5	8.13	−124.3
6	4.12	−30.1

We pick the single null of the two-element array at 135°. This null has symmetry about the axis of the array. With the first element at zero phase, we pick the element phase to cancel the first element voltage when $\theta = 135°$:

$$\text{phase} = 180° - 360°(0.3)\cos 135° = 256.37°$$

When we convolute the two arrays, we obtain the feeding coefficients from Eq. (4-107) (Table 4-39). The elements in the center that result from two convolutions have summed feeding coefficients producing a six-element array. Figure 4-32 shows the patterns of the convolution. We obtain the six-element array pattern (Figure 4-32*b*) by multiplication of the patterns of the individual subarrays (Figure 4-32*a*).

4-23 APERTURE BLOCKAGE

Blocking an aperture reduces the gain and raises the sidelobes. The blockage either scatters the aperture power in unwanted directions in a broad pattern or is just an area without fields. Scattered blockage causes higher sidelobes and greater loss than the

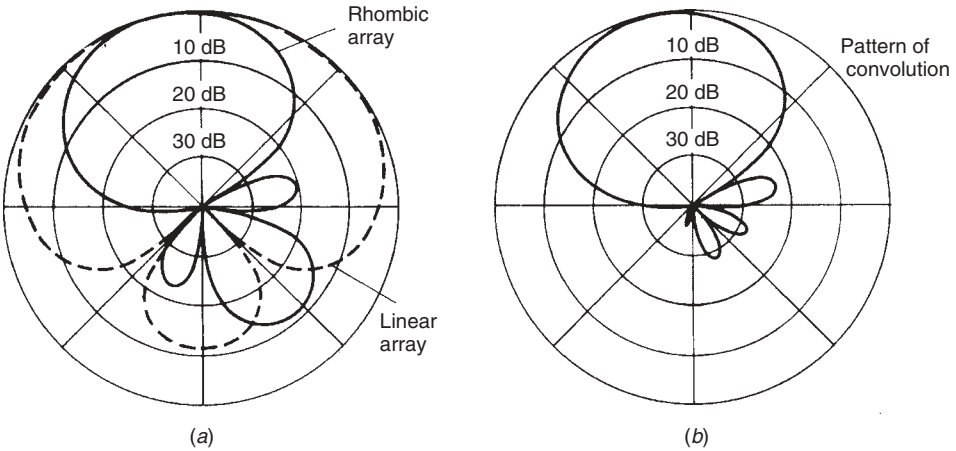


FIGURE 4-32 Patterns of the convolution of a rhombic and a linear array to form the six-element rectangular array of Figure 4-31.

nonexcitation blockage. Scattered blockage has the same power input as the unblocked aperture, but fields scattered off the blockage do not contribute significantly to the maximum field. Compared to the unblocked aperture, the blockage efficiency becomes

$$\text{blockage efficiency} = \frac{\left| \iint_{\text{blocked}} E e^{j\mathbf{k}\cdot\mathbf{r}'} \, ds' \right|^2_{\max}}{\left| \iint_{\text{unblocked}} E e^{j\mathbf{k}\cdot\mathbf{r}'} \, ds' \right|^2_{\max}} \qquad \text{scattered} \tag{4-111}$$

We use Eq. (2-16) to compute the directivity of each distribution by using the total power radiated from the unblocked aperture [denominator of Eq. (2-16)] for the blocked aperture. A centrally blocked circular aperture with a uniform distribution has the blockage efficiency $(1 - b^2)^2$ (scattered), where b is the normalized blockage radius. The blockage of a circular Gaussian distribution has a simple blockage function:

$$\text{blockage efficiency} = \left(\frac{e^{-\rho b} - e^{-\rho}}{1 - e^{-\rho}} \right)^2$$

The second type of blockage is an area without fields. The blockage does not waste power in the aperture. When we take the ratio of the two directivities, we must account for the power in each aperture:

$$\text{blockage efficiency} = \frac{\left| \iint_{\text{blocked}} E e^{j\mathbf{k}\cdot\mathbf{r}'} \, ds' \right|^2_{\max} \iint_{\text{unblocked}} |E|^2 \, ds'}{\left| \iint_{\text{unblocked}} E e^{j\mathbf{k}\cdot\mathbf{r}'} \, ds' \right|^2_{\max} \iint_{\text{blocked}} |E|^2 \, ds'} \qquad \text{nonexcitation} \tag{4-112}$$

The blockage of a uniformly excited centrally blocked circular aperture where the center is not excited reduces the directivity only by the area lost from the aperture, $1 - b^2$ (nonexcitation). In a sense, nonexcitation blockage is not a true loss; it is a loss of potential radiation aperture.

Table 4-40 lists the blockage losses of centrally blocked circular apertures calculated by Eq. (4-111), the more severe case. The uniformly excited aperture is affected least by blockage. All points are equally important. The tapered distributions suffer more loss with increased taper toward the edge. The lists for different tapered distributions track each other fairly closely, and any one of them gives a good estimate of the blockage loss. Blockage causes sidelobes. In the case of scattered blockage the exact sidelobes cannot be found without an analysis of the scatterer. A Cassegrain reflector would need a geometric theory of diffraction (GTD) analysis to locate the directions of scattering from the subreflector. We can handle the nonexcitation blockage in a general fashion. Consider the aperture to be broken into two radiating apertures. The first is the unblocked aperture; the second is the blockage. If we take the blockage aperture to be 180° out of phase with respect to the unblocked aperture distribution, the sum gives us the blocked distribution.

We use this analysis as an approximation to scattered blockage with the realization that scattering may produce unpredicted lobes.

TABLE 4-40 Blockage Losses of Circular-Aperture Distributions (dB)

Central Blockage (%)	Uniform	Gaussian 12-dB Edge	Taylor 30 dB, $\overline{n} = 6$	Taylor 40 dB, $\overline{n} = 6$	Hansen 30 dB	Hansen 40 dB
5	0.02	0.04	0.04	0.05	0.05	0.07
6	0.03	0.06	0.06	0.08	0.07	0.09
7	0.04	0.08	0.08	0.11	0.09	0.13
8	0.06	0.10	0.10	0.14	0.12	0.17
9	0.07	0.13	0.13	0.18	0.16	0.21
10	0.09	0.16	0.16	0.22	0.19	0.26
11	0.11	0.19	0.20	0.26	0.23	0.32
12	0.13	0.23	0.24	0.31	0.28	0.38
13	0.15	0.27	0.28	0.37	0.33	0.44
14	0.17	0.32	0.32	0.43	0.38	0.51
15	0.20	0.36	0.37	0.49	0.43	0.59
16	0.22	0.41	0.42	0.56	0.49	0.67
17	0.26	0.47	0.48	0.63	0.56	0.76
18	0.29	0.52	0.54	0.71	0.63	0.85
19	0.32	0.58	0.60	0.79	0.70	0.95
20	0.36	0.65	0.67	0.88	0.77	1.06
21	0.39	0.71	0.74	0.97	0.86	1.17
22	0.43	0.78	0.81	1.07	0.94	1.28
23	0.47	0.86	0.88	1.17	1.03	1.40
24	0.52	0.94	0.96	1.27	1.12	1.53
25	0.56	1.02	1.05	1.38	1.22	1.66

We can calculate an upper bound on the sidelobes easily. Assume that the blockage distribution is uniform and compared to the main aperture, produces a broad, flat beam. Since the blockage aperture fields are 180° out of phase from the unblocked aperture fields, their radiation subtracts from the main beam and adds sidelobes 180° out of phase with respect to the main lobe. The sidelobe due to the blockage is proportional to the area: sidelobe level $= 20 \log b$. This formula estimates values much higher than are realized. Table 4-41 lists the sidelobes of a centrally blocked Taylor circular aperture distribution with 40-dB design sidelobes. They are far less than predicted by the upper bound.

TABLE 4-41 Sidelobe Level Due to Central Blockage of a Circular Aperture with Taylor Distribution (40 dB, $\overline{n} = 6$)

Blockage (% of Diameter)	Sidelobe Level (dB)	Blockage (% of Diameter)	Sidelobe Level (dB)	Blockage (% of Diameter)	Sidelobe Level (dB)
7	34.5	13	26.1	19	21.1
8	32.8	14	25.6	20	20.4
9	31.3	15	24.2	21	19.7
10	29.8	16	23.3	22	19.1
11	28.5	17	21.7	23	18.5
12	27.3	18	21.7	24	18.0

Ludwig [27] has found distributions to reduce the sidelobes of blocked apertures. The first sidelobe can be reduced only a little, but the outer sidelobe levels can be controlled. In many applications one high sidelobe next to the main beam is acceptable. A Taylor distribution for circular apertures with a zero edges value, like Section 4-5 for linear apertures, reduces the far-out sidelobes. A second aperture function with a doughnut distribution also reduces all but the first sidelobe. Reducing the edge taper of the blockage distribution lowers the blockage-caused sidelobes.

Sachidananda and Ramakrishna [28] use a numerical optimization technique to reduce the sidelobes of a blocked aperture for both the sum and difference patterns of a monopulse excitation. They start with the Taylor and Bayliss circular-aperture distribution functions [Eqs. (4-94) and (4-98)]. The coefficients B_m are determined through the numerical optimization, which restrains the sidelobes while optimizing the monopulse tracking coefficients and sum pattern gain.

4-24 QUADRATIC PHASE ERROR

A linear phase error function scans the aperture beam with some loss of gain because of the shrinking of the projected aperture in the direction of the main beam. Quadratic phase error (order 2) does not scan the beam but causes loss and a change in the sidelobe levels and the depth of the nulls between them. This phase error arises mainly from defocusing when the source of radiation appears as a point source. A feed axially displaced from the focus of a parabolic reflector produces quadratic phase error in the aperture. The flare angle of a horn changes the distance from the assumed point source in the throat to different points in the aperture at the end of the flare. We can approximate the phase distribution as quadratic.

We express the quadratic phase error in a line-source aperture as

$$\text{linear:} \quad e^{-j2\pi S(2x/a)^2} \qquad |x/a| \le 0.5 \tag{4-113a}$$

where S is a dimensionless constant, cycles and a is the aperture width. Similarly, the circular-aperture phase is

$$\text{circular:} \quad e^{-j2\pi Sr^2} \qquad r \le 1 \tag{4-113b}$$

where r is the normalized radius. We use Eq. (4-7) with the linear-source aperture phase error [Eq. (4-113a)] and use Eq. (4-9) with the quadratic phase error [Eq. (4-113b)] in a circularly symmetrical aperture distribution to compute phase error loss:

$$\text{PEL}_x = \frac{\left| \int_{-a/2}^{a/2} E(x) e^{-j2\pi S(2x/a)^2}\, dx \right|^2}{\left[\int_{-a/2}^{a/2} |E(x)|\, dx \right]^2} \qquad \text{linear} \tag{4-114}$$

$$\text{PEL} = \frac{\left| \int_0^1 E(r) e^{-j2\pi Sr^2} r\, dr \right|^2}{\left[\int_0^1 |E(r)| r\, dr \right]^2} \qquad \text{circular} \tag{4-115}$$

A few distributions have simple formulas for the phase error efficiency when excited with quadratic phase error [29]:

$$\text{uniform linear:} \quad \text{PEL}_x = \frac{1}{2S}\left[C^2(2\sqrt{S}) + S^2(2\sqrt{S})\right] \tag{4-116}$$

where $C(t)$ and $S(t)$ are the Fresnel integrals, tabulated functions:

$$\text{uniform circular: PEL} = \left(\frac{\sin \pi S}{\pi S}\right)^2 \tag{4-117}$$

$$\text{Circular Gaussian}(e^{-\rho r^2}): \quad \text{PEL} = \frac{\rho^2[1 - 2e^{-\rho}\cos(2\pi S) + e^{-2\rho}]}{[\rho^2 + (2\pi S)^2](1 - e^{-\rho})^2} \tag{4-118}$$

We use numerical integration for the general distribution.

Table 4-42 lists quadratic phase error losses for various linear-aperture distributions. We will use the lists for uniform and cosine distributions to evaluate the gains of rectangular horns. The effect of quadratic phase error decreases as the distribution taper increases. Table 4-43 lists results for a few circularly symmetrical aperture distributions. Quadratic phase error raises the sidelobes of low-sidelobe antennas. Figure 4-33 shows the effects on a circular Taylor distribution with 35-dB design sidelobes. The first lobe increases, and the null between the main lobe and the first sidelobe disappears as the quadratic phase error increases. A source antenna spaced a finite distance, as on

TABLE 4-42 Quadratic Phase Error Loss of Linear-Aperture Distributions (dB)

Cycles, S	Uniform	Cosine	Cosine^2	Cosine^2 + 19.9-dB Pedestal
0.05	0.04	0.02	0.01	0.02
0.10	0.15	0.07	0.04	0.07
0.15	0.34	0.16	0.09	0.16
0.20	0.62	0.29	0.16	0.28
0.25	0.97	0.45	0.25	0.44
0.30	1.40	0.65	0.36	0.63
0.35	1.92	0.88	0.49	0.84
0.40	2.54	1.14	0.64	1.08
0.45	3.24	1.43	0.80	1.34
0.50	4.04	1.75	0.97	1.62
0.55	4.93	2.09	1.16	1.90
0.60	5.91	2.44	1.36	2.19
0.65	6.96	2.82	1.57	2.48
0.70	8.04	3.20	1.79	2.76
0.75	9.08	3.58	2.01	3.04
0.80	9.98	3.95	2.23	3.29
0.85	10.60	4.31	2.46	3.52
0.90	10.87	4.65	2.69	3.73
0.95	10.80	4.97	2.91	3.92
1.00	10.50	5.25	3.13	4.09

TABLE 4-43 Quadratic Phase Error Loss of Circular-Aperture Distributions (dB)

Cycles, S	Uniform	Gaussian 12-dB Edge	Taylor 30 dB	Taylor 40 dB	Hansen 30 dB	Hansen 40 dB
0.05	0.04	0.03	0.04	0.03	0.03	0.02
0.10	0.14	0.13	0.15	0.11	0.11	0.08
0.15	0.32	0.29	0.33	0.26	0.25	0.19
0.20	0.58	0.53	0.59	0.46	0.45	0.34
0.25	0.91	0.82	0.93	0.72	0.70	0.53
0.30	1.33	1.20	1.36	1.03	1.01	0.76
0.35	1.83	1.64	1.86	1.41	1.38	1.03
0.40	2.42	2.16	2.46	1.84	1.81	1.34
0.45	3.12	2.76	3.16	2.33	2.30	1.69
0.50	3.92	3.44	3.95	2.87	2.85	2.08
0.55	4.86	4.22	4.86	3.47	3.46	2.50
0.60	5.94	5.08	5.88	4.11	4.16	2.95
0.65	7.20	6.04	7.01	4.79	4.85	3.43
0.70	8.69	7.10	8.25	5.50	5.63	3.94
0.75	10.46	8.24	9.56	6.21	6.43	4.46
0.80	12.62	9.44	10.87	6.91	7.26	4.98
0.85	15.39	10.66	12.01	7.56	8.09	5.51
0.90	19.23	11.81	12.80	8.14	8.88	6.03
0.95		12.75		8.62	9.60	6.53
1.00		13.36		8.99	10.20	6.99

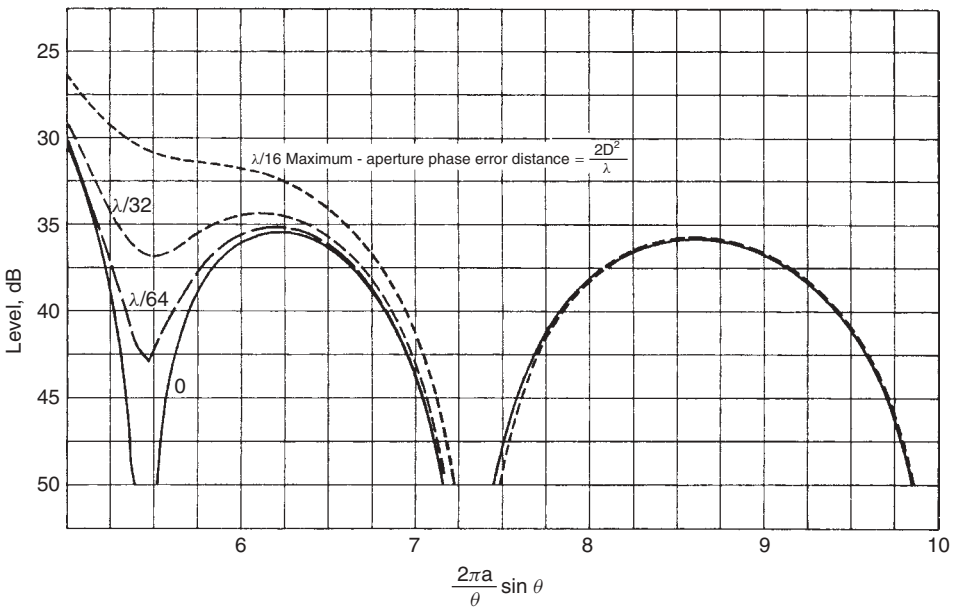


FIGURE 4-33 Effects of quadratic phase error on 35-dB circular Taylor distribution ($\overline{n} = 6$).

an antenna measurement range, feeds the aperture with a quadratic phase error. The source would have to be spaced $8D^2/\lambda$ to measure the sidelobe level within 0.5 dB. Low-sidelobe antennas require greater distances than the usual $2D^2/\lambda$ for accurate sidelobe measurement [30].

4-25 BEAM EFFICIENCY OF CIRCULAR APERTURES WITH AXISYMMETRIC DISTRIBUTION

From Eq. (1-27) we can derive an approximate formula for axisymmetric distributions that depends on the normalized variable k_r (or U). For large apertures we can approximate $\cos\theta \approx 1$ in the main beam, integrate the ϕ integral to obtain 2π, and incorporate the $(ka)^2$ directivity factor into the integral:

$$\text{beam efficiency} = \frac{\text{ATL} \cdot \text{PEL} \int_0^{k_{r1}} |f(k_r)|^2 k_r \, dk_r}{2|f(0)|^2} \tag{4-119}$$

$$= \frac{\text{ATL} \cdot \text{PEL} \int_0^{u_1} |f(U)|^2 U \, dU}{2|f(0)|^2} \tag{4-120}$$

where k_r is the factor $(2\pi a \sin\theta)/\lambda$, U (the Taylor distribution factor) is $(2a \sin\theta)/\lambda$, and a is the aperture radius. U_1 and k_{r1} correspond to the beam edge θ_1. The integrals of Eqs. (4-119) and (4-120) cause underestimations of beam efficiency for small apertures when we ignore the $\cos\theta$ factor, which should divide the argument of the integral.

Table 4-44 lists beam edges in k_r-space $(2\pi a \sin\theta)/\lambda$ for various distributions along with the beam efficiency at the null beam edge. We can use it to determine the aperture size required for a given beam efficiency beamwidth specification.

Example Calculate the aperture radius to give a 90% beam efficient beamwidth of 5° for the distribution: parabolic on 12-dB pedestal.

TABLE 4-44 Beam Efficiencies of Axisymmetric Circular-Aperture Distributions

			$k_r = 2\pi a \sin\theta/\lambda$ Specified Beam Efficiency (%)			
Distribution	Null, k_r	Beam Efficiency at Null (%)	80	85	90	95
Uniform	3.83	83.7	2.82	4.71	5.98	
Parabolic	5.14	98.2	2.81	3.03	3.31	3.75
Parabolic + 12-dB pedestal	4.58	96.4	2.60	2.81	3.10	3.64
Taylor						
30 dB, $\overline{n} = 6$	4.90	96.2	2.65	2.88	3.19	3.82
30 dB, $\overline{n} = 10$	4.74	91.4	2.76	3.06	3.65	
40 dB, $\overline{n} = 6$	6.00	99.5	2.90	3.13	3.42	3.85
Hansen						
30 dB	5.37	99.3	2.79	3.01	3.28	3.69
40 dB	6.64	99.9	3.17	3.42	3.73	4.19

From Table 4-43,

$$k_r(90\%) = 3.10 = \frac{2\pi a}{\lambda} \sin \frac{5^\circ}{2}$$

$$\frac{a}{\lambda} = \frac{3.10}{2\pi \sin(5^\circ/2)} = 11.31$$

The beam edge has $\cos 2.5^\circ = 0.999$, which justifies the approximation in Eq. (4-119).

REFERENCES

1. T. T. Taylor, Design of line source antennas for narrow beamwidth and low sidelobes, *IEEE Transactions on Antennas and Propagation*, vol. AP-4, no. 1, January 1955, pp. 16–28.
2. C. L. Dolph, A current distribution for broadside arrays which optimizes the relationship between beamwidth and sidelobe level, *Proceedings of IEEE*, vol. 34, June 1946, pp. 335–348.
3. R. C. Hansen, Linear arrays, Chapter 9 in A. W. Rudge et al., eds., *The Handbook of Antenna Design*, Vol. 2, Peter Peregrinus, London, 1982.
4. D. R. Rhodes, On a new condition for physical realizability of planar antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-19, no. 2, March 1971, pp. 162–166.
5. D. R. Rhodes, On the Taylor distribution, *IEEE Transactions on Antennas and Propagation*, vol. AP-20, no. 2, March 1972, pp. 143–145.
6. R. S. Elliott, *Antenna Theory and Design*, Prentice-Hall, Englewood Cliffs, NJ, 1981.
7. D. A. Pierre, *Optimization Theory with Applications*, Wiley, New York, 1969.
8. E. T. Bayliss, Design of monopulse antenna difference patterns with low sidelobes, *Bell System Technical Journal*, vol. 47, May–June 1968, pp. 623–650.
9. P. M. Woodward, A method of calculating the field over a plane aperture required to produce a given polar diagram, *Proceedings of IEE*, vol. 93, pt. IIIA, 1947, pp. 1554–1558.
10. S. A. Schelkunoff, A mathematical theory of linear arrays, *Bell System Technical Journal*, vol. 22, 1943, pp. 80–107.
11. J. Kraus, *Antennas*, McGraw-Hill, New York, 1950.
12. R. S. Elliott, Beamwidth and directivity of large scanning arrays, *Microwave Journal*, vol. 6, no. 12, December 1963, pp. 53–60.
13. A. T. Villeneuve, Taylor patterns for discrete arrays, *IEEE Transactions on Antennas and Propagation*, vol. AP-32, no. 10, October 1984, pp. 1089–1092.
14. R. S. Elliott, On discretizing continuous aperture distributions, *IEEE Transactions on Antennas and Propagation*, vol. AP-25, no. 5, September 1977, pp. 617–621.
15. H. J. Orchard, R. S. Elliott, and G. J. Stern, Optimising the synthesis of shaped beam antenna patterns, *IEE Proceedings*, vol. 132, pt. H, no. 1, February 1985, pp. 63–68.
16. Y. U. Kim and R. S. Elliott, Shaped-pattern synthesis using pure real distributions, *IEEE Transactions on Antennas and Propagation*, vol. AP-36, no. 11, November 1988, pp. 1645–1648.
17. F. Ares-Pena, Application of genetic algorithms and simulated annealing to some antenna problems, Chapter 5 in Y. Rahamat-Samii and E. Michielssen, eds., *Electromagnetic Optimization by Genetic Algorithms*, Wiley, New York, 1999.
18. C. H. Walters, *Traveling Wave Antennas*, Dover, New York, 1970.
19. G. Doundoulakis and S. Gethin, Far field patterns of circular paraboloidal reflectors, *IRE National Convention Record*, pt. 1, 1959, pp. 155–173.

20. R. C. Hansen, Circular aperture distribution with one parameter, *Electronic Letters*, vol. 11, no. 8, April 17, 1975, p. 184.
21. R. C. Hansen, A one-parameter circular aperture with narrow beamwidth and low sidelobes, *IEEE Transactions on Antennas and Propagation*, vol. AP-24, no. 4, July 1976, pp. 477–480.
22. T. T. Taylor, Design of circular apertures for narrow beamwidth and low side lobes, *IEEE Transactions on Antennas and Propagation*, vol. AP-8, no. 1, January 1960, pp. 17–22.
23. F. I. Tseng and D. K. Cheng, Optimum scannable planar arrays with an invariant side-lobe level, *Proceedings of IEEE*, vol. 56, no. 11, November 1968, pp. 1771–1778.
24. D. Steinberg, *Principles of Aperture and Array System Design*, Wiley, New York, 1976.
25. D. K. Cheng, *Analysis of Linear Systems*, Addison-Wesley, Reading, MA., 1959.
26. S. R. Laxpati, Planar array synthesis with prescribed pattern nulls, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 6, November 1982, pp. 1176–1183.
27. A. C. Ludwig, Low sidelobe aperture distributions for blocked and unblocked circular apertures, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 5, September 1982, pp. 933–946.
28. M. Sachidananda and S. Ramakrishna, Constrained optimization of monopulse circular aperture distribution in the presence of blockage, *IEEE Transactions on Antennas and Propagation*, vol. AP-31, no. 2, March 1983, pp. 286–293.
29. A. W. Love, Quadratic phase error loss in circular apertures, *Electronics Letters*, vol. 15, no. 10, May 10, 1979, pp. 276, 277.
30. P. S. Hacker and H. E. Schrank, Range requirements for measuring low and ultralow sidelobe antenna patterns, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 5, September 1982, pp. 956–966.

5

DIPOLES, SLOTS, AND LOOPS

A dipole is a conductive rod usually split in the center and fed from a balanced transmission line that carries equal and oppositely flowing currents. Not all dipoles are split and fed in the center because currents can be excited on it electromagnetically or it can be shunt fed. The dipole length determines possible current distributions in modes, and when we place a continuous rod near an antenna radiating a linear polarization component directed along the rod, it excites a standing-wave current on the rod. The amount excited on the rod depends on how close its length is to resonance and the antenna spacing. Of course, the continuous rod loads the fed antenna through mutual coupling. We can feed the continuous rod from a coax line by attaching the outer conductor to the center and then connecting the center conductor away from the center in a shunt feed.

A slot is a narrow-width opening in a conductive sheet. When excited by a voltage across the narrow dimension it appears to radiate from an equivalent magnetic current flowing along the long dimension that replaces the voltage (or electric field) across it. Most slots, similar to dipoles, have a finite length with either short or open circuits at both ends. The voltage along the slot forms a standing wave. Of course, magnetic currents are fictitious, and real electric currents flow in the conductive sheet around the slot. These currents do not have a simple distribution and are difficult to use for analysis, so we use simpler magnetic currents, although when analyzing a slot using the method of moments, we model the conductors around the slot and calculate patterns, reaction, and so on, from these real currents. Initial slot calculations assume that the conductive sheet is infinite, similar to the analysis of dipoles situated in free space. Complete analysis of the dipole requires analysis in the presence of the mounting configuration. Similarly, full analysis of slots includes the effects of the finite sheet and scattering from the objects around it.

Modern Antenna Design, Second Edition, By Thomas A. Milligan

After considering ideal cases, we analyze the effects of finite ground planes, nearby scatterers, and the interaction between dipoles and slots. The batwing antenna presents an unusual case where the antenna at first glance looks like a dipole but actually radiates from a combination of a slot and a finite dipole structure. Another interesting case is the waveguide slot. Currents flow on the inside surfaces of a waveguide, and the finite current skin depth prevents it from reaching the outside. The metal walls shield the currents and prevent the loss of power by radiation. When we cut a slot in the wall, the internal currents flow out the slot and onto the outside of the waveguide and radiate. The excitation and length of the slot relative to the internal currents determine the amount radiated. Similarly, the slots load the waveguide as a transmission line because of the loss of power.

Our analysis starts with a dipole in free space or a slot on an infinite conductive sheet. The two problems are duals. Dipoles radiate from a standing-wave electric (real) current, whereas the slot radiates from a standing-wave magnetic current. We use the same mathematics for both patterns. By the Babinet–Booker principle of complementary structures, we relate the input impedance of one to the other. Both structures radiate the same pattern but differ in polarization. Dipoles and slots share the same analysis through duality, so we develop them together. Singly and in arrays, they satisfy many antenna needs. Although they share a dual analysis, they have unique feeding requirements. We discuss baluns for dipoles and waveguide slot excitations as practical implementations.

In Chapter 2 we presented the analysis of a small loop excited with a uniform current (Section 2-1.2). The loop current was replaced with a small magnetic current element flowing along the normal to the plane of the loop. Multiple turns and ferrite loading increase the efficiency of loops and produce a more useful antenna. Exciting a uniform current on a loop is a difficult task that offers little practical benefit. The loops discussed will have standing-wave electric currents excited on them determined by feeding methods. The natural balun used to excite a small loop produces a standing-wave current with zero current at the point where the two sides are connected to form the loop. A resonant length loop of about one wavelength perimeter radiates a dipole pattern from a standing-wave current. The quadrifilar helix consists of two loops twisted around a common axis. The twist produces currents that radiate circular polarization from each loop. Analysis shows that the currents are standing wave.

Feeding a dipole or loop requires a balun to prevent current flow either along the outside of a coaxial feeder or excitation of unbalanced currents along a two-wire line. The current flowing along the outside of the coax or unbalanced currents on the two-wire line radiate in unwanted directions or radiate undesired polarization. When we design an antenna without considering or knowing its final mounting, we produce an uncontrolled situation without a balun. Our initial configuration may work without a balun, but the antenna may fail to produce the desired pattern in the final location. If you control the installation completely, you can reduce your design effort and may be able to eliminate the balun.

5-1 STANDING-WAVE CURRENTS

Think of a dipole as a diverging two-wire transmission line. The characteristic impedance increases as the wave approaches the open-circuited ends. The slot is the

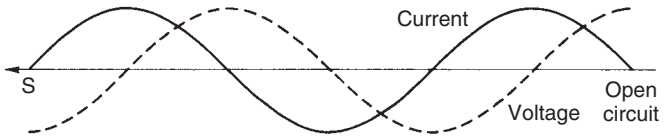


FIGURE 5-1 Standing wave.

dual of a strip dipole. A voltage excited across the slot propagates along a slotline toward short-circuited ends. Each type of transmission line reflects the incident wave from the terminations. The combination of two waves traveling in opposite directions creates a standing wave on the line. The current and voltage are 90° out of phase and 90° out of space phase (Figure 5-1). Current and voltage change places on the short-circuited termination of the slot.

The dipole is not a uniform transmission line, but we can approximate the current as a standing wave with the current vanishing on the ends. The slot voltage is a standing wave also vanishing on the ends. The standing waves for a center-fed dipole or slot are expressed as follows:

$$\begin{array}{cccc} \text{Dipole} & & \text{Slot} & \\ I = I_0 \sin k\left(\frac{L}{2} - z\right) & & V = V_0 \sin k\left(\frac{L}{2} - z\right) & z \geq 0 \\ I = I_0 \sin k\left(\frac{L}{2} + z\right) & & V = V_0 \sin k\left(\frac{L}{2} + z\right) & z \leq 0 \end{array} \tag{5-1}$$

The voltage distribution on the slot is equivalent to a magnetic current.

We calculate radiation from the linear sinusoidal current distributions by the vector potentials: electric (slot) (Section 2-1.2) and magnetic (dipole) (Section 2-1.1). Figure 5-2 gives typical sinusoidal distributions for various lengths. The currents match at the feed point and vanish on the ends. Consider the pattern of the 2λ dipole at $\theta = 90°$. We can assume that it is a continuous array and sum the fields from each

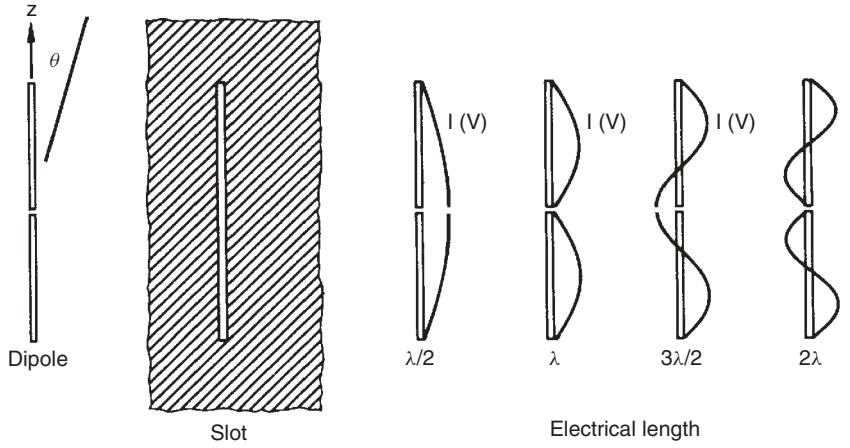


FIGURE 5-2 Sinusoidal distributions.

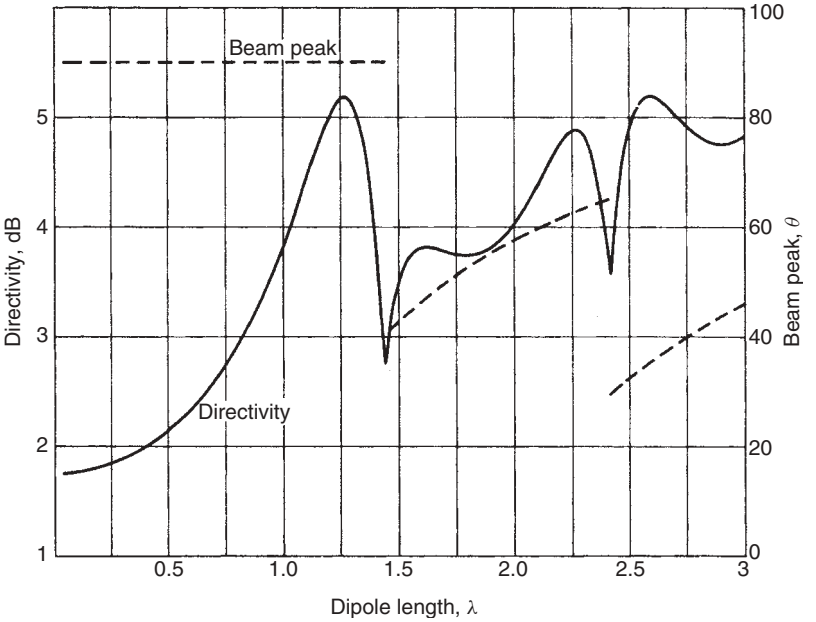


FIGURE 5-3 Dipole (slot) directivity and beam peak versus length.

portion along the axis. The equal positive and negative portions of the standing-wave current sum to zero and produce a pattern null normal to the axis. By integrating Eqs. (2-5) and (2-10), we compute far fields for radiators centered on the z-axis through the far-field conversion [Eqs. (2-1) and (2-9)] [1, p. 82]:

$$E_\theta = j\eta \frac{I_0}{2\pi r} e^{-jkr} \frac{\cos(kL/2\cos\theta) - \cos(kL/2)}{\sin\theta} \qquad \text{dipole} \tag{5-2}$$

where L is the total dipole length. Using the $Y = 0$ plane as the slot ground plane, the far-field magnetic field is found as

$$H_\theta = \frac{\pm jV_0}{\eta 2\pi r} e^{-jkr} \frac{\cos(kL/2\cos\theta) - \cos(kL/2)}{\sin\theta} \qquad \text{slot} \tag{5-3}$$

where L is the total slot length. We apply the upper sign for $Y > 0$ and the lower sign for $Y < 0$. The electric field of the slot is found from $E_\phi = -\eta H_\theta$. Equations (5-2) and (5-3) have the same pattern shape and directivity. We integrate the magnitude squared of Eqs. (5-2) and (5-3) to determine the average radiation intensity. Joined with the maximum radiation intensity, we calculate directivity (Figure 5-3) versus length.

5-2 RADIATION RESISTANCE (CONDUCTANCE)

The far-field power densities, Poynting vectors, are given by

$$S_r = \begin{cases} \dfrac{|E_\theta|^2}{\eta} & \text{dipole} \\ |H_\theta|^2 \eta & \text{slot} \end{cases}$$

where η is the impedance of free space (376.7 Ω). When these are integrated over the radiation sphere to compute the power radiated, the results contain either $|I_0|^2$ (dipole) or $|V_0|^2$ (slot), the maximum sinusoidal current (voltage). We define the radiation resistance (conductance) as

$$\begin{aligned} R_r &= \frac{P_r}{|I_0|^2} \qquad \text{dipole} \\ G_r &= \frac{P_r}{|V_0|^2} \qquad \text{slot} \end{aligned} \tag{5-4}$$

Figure 5-4 is a plot of the radiation resistance of each versus length [2, p. 157]. The input resistance differs from the radiation resistance because it is the ratio of the input current (voltage) to the power radiated:

$$\begin{aligned} I_i &= I_0 \sin \frac{kL}{2} \qquad \text{dipole} \\ V_i &= V_0 \sin \frac{kL}{2} \qquad \text{slot} \end{aligned} \tag{5-5}$$

Combining Eqs. (5-4) and (5-5), we find that

$$\begin{aligned} R_i &= \frac{R_r}{\sin^2(kL/2)} \qquad \text{dipole} \\ G_i &= \frac{G_r}{\sin^2(kL/2)} \qquad \text{slot} \end{aligned} \tag{5-6}$$

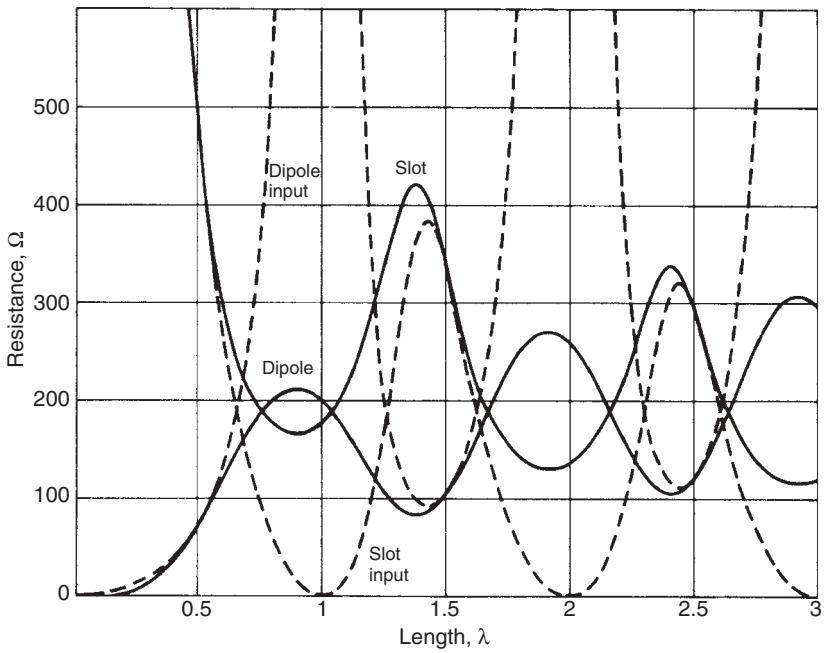


FIGURE 5-4 Dipole and slot radiation and center-fed input resistances.

The input resistances (Figure 5-4) differ from the radiation resistances by Eq. (5-6). The input resistance of a one-wavelength dipole is large but not infinite, as shown; it depends greatly on the diameter and input region. If we take the product of the radiation or input resistances, we determine that

$$R_{\text{dipole}} R_{\text{slot}} = \frac{\eta^2}{4} \tag{5-7}$$

one of the consequences of the Babinet–Booker principle [3].

The input resistance depends on the current at the input [Eq. (5-6)]. When the standing-wave current is high and the voltage is low, the input resistance is moderate. A center-fed half-wavelength dipole has the same input resistance as radiation resistance, since the current maximum occurs as the input. On the other hand, a center-fed half-wavelength slot has a current minimum (voltage maximum) at its input, which gives it high input resistance. When both are a full wavelength long, the dipole standing-wave current is at a minimum and the slot standing-wave current is at a maximum (Figure 5-2). The dipole has a high input resistance and the slot has a low input resistance. We can lower the input resistance by feeding at a high current point, but we may excite a distribution different from that expected.

A short dipole looks like a capacitor at the input. As the length increases, the radiation resistance grows and the capacitance decreases. Just before the length reaches $\lambda/2$, the capacitance becomes zero. The exact length at which the antenna resonates (zero reactance) depends on the diameter of the elements and the input gap. A good starting point is 95% of a half wavelength. Beyond the resonant length, the dipole becomes inductive. The impedance of a thin half-wavelength dipole is $73 + j42.2\,\Omega$, whereas the resonant-length dipole resistance is about $67\,\Omega$. The slot looks like an inductor when short. Think of it as a short-length short-circuited shunt slotline stub. The inductance increases as its length increases and the slot resonates like the dipole, just short of $\lambda/2$. Additional resonances occur at longer lengths. Increasing the frequency is equivalent to increasing the length for the thin dipole.

5-3 BABINET–BOOKER PRINCIPLE [3; 4, p. 337]

A strip dipole and a slot are complementary antennas. The solution for the slot can be found from the solution to an equivalent dipole by an interchange of the electric and magnetic fields. Not only the pattern but also the input impedance can be found. Figure 5-5 shows two such complementary structures. Babinet's principle of optical screens (scalar fields) states that given the solutions to the diffraction patterns of a screen, F_i, and the screen's complement, F_c, the sum equals the pattern without the screen. Booker extended Babinet's principle to vector electromagnetic fields. Strict complementation of an electric conductor requires a nonexistent magnetic conductor. Booker solved this problem by using only perfectly conducting infinitesimally thin screens and by interchanging the electric and magnetic fields between the screen and its complement. If we take two such complementary screens and perform line integrals over identical paths to compute the impedance of each, we obtain the result

$$Z_1 Z_c = \frac{\eta^2}{4} \tag{5-8}$$

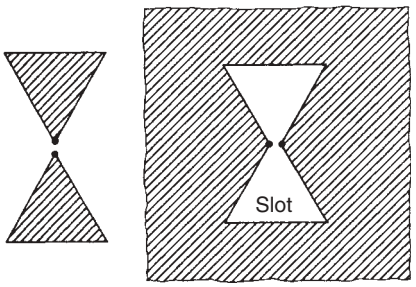


FIGURE 5-5 Complementary screens.

where Z_1 is the input impedance of the structure, Z_c the input impedance of the complementary structure, and η the impedance of free space (376.7 Ω). Equation (5-8) extends Eq. (5-7) to the total impedance and includes mutual impedances as well as self-impedances.

Certain antennas, such as flat spirals, are self-complementary—an exchange of the spaces and conductors leaves the structure unchanged except for rotation. For a two-arm structure,

$$Z_0^2 = \frac{\eta^2}{4} \quad \text{or} \quad Z_0 = 188\,\Omega$$

Rumsey [5, p. 28] extended these ideas to antennas with more than two conductors to determine the input impedances in various feeding modes.

We must relate flat-strip dipoles to normal round-rod dipoles to use the available results for round dipoles. The diameter of an equivalent round rod equals one-half the strip width of the flat structure. Consider a thin dipole with its near λ/2 resonance of 67 Ω. We calculate equivalent slot impedance from Eq. (5-8):

$$Z_{\text{slot}} = \frac{376.7^2}{4(67)} = 530\,\Omega$$

A half-wavelength slot impedance is

$$Z_{\text{slot}} = \frac{376.7^2}{4(73 + j42.5)} = 363 - j211\,\Omega$$

The λ/2 dipole is inductive when it is longer than a resonant length, whereas the slot is capacitive.

5-4 DIPOLES LOCATED OVER A GROUND PLANE

We analyze a dipole over a ground plane as a two-element array of the dipole and its image. The ground plane more than doubles the gain of the element by limiting the radiation directions. We can expect a change in the input impedance as the dipole interacts with its image. A vertical dipole excites currents in the ground plane, when transmitting, equivalent to its image. The image is vertical (Figure 5-6) and has the same phase as the dipole (even mode). The impedance of the dipole becomes

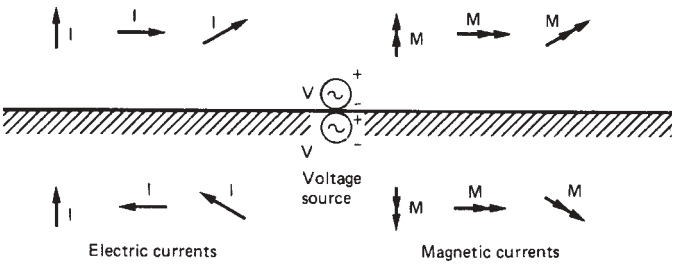


FIGURE 5-6 Ground-plane images.

$Z = Z_{11} + Z_{12}$. Z_{12} is the mutual impedance between the dipole and its image spaced $2H$, where H is the center height of the dipole over the ground plane. The array radiates its maximum in the direction of the ground plane. The dipole also radiates its maximum pattern along the ground plane given by

$$U_{\max} = \frac{\eta |I_0|^2}{(2\pi)^2} \left(1 - \cos \frac{kL}{2}\right)^2 \tag{5-9}$$

where L is the dipole length. The radiated power of the single dipole is

$$P_{\text{in}} = R_{11} |I_0|^2 \left(1 + \frac{R_{12}}{R_{11}}\right)$$

The two-element array increases the field over a single element by 2 and the radiation intensity by 4:

$$\text{directivity} = \frac{4U_{d,\max}}{P_{\text{in}}/4\pi} = \frac{4\eta[1 - \cos(kL/2)]^2}{(R_{11} + R_{12})\pi}$$

We used only the power into the dipole, since no source is connected to the image. Figure 5-7 is a plot of the directivity of a vertical dipole versus height over the ground plane.

A horizontal dipole and its image (Figure 5-6) form an odd-mode two-element array (Section 3-1). The input impedance of the dipole becomes $Z_{11} - Z_{12}$ for the odd-mode array. The value of the mutual impedance Z_{12} approaches that of the self-impedance Z_{11} as the two dipoles move close together. The input impedance approaches zero as the distance from the dipole to ground plane shrinks. The input impedance of all odd-mode array elements decreases as the elements approach each other. The two-element odd-mode array produces a null along the ground plane. The beam peak occurs normal to the ground plane ($\theta = 0^\circ$) when the distance between the dipole and its image is less than $\lambda/2$ or $H \leq \lambda/4$. The pattern bifurcates after that height is exceeded. The maximum radiation from the array is

$$U_{A,\max} = \begin{cases} 4 \sin^2 \dfrac{2\pi H}{\lambda} & H \leq \dfrac{\lambda}{4} \\ 4 & H \geq \dfrac{\lambda}{4} \qquad \theta_{\max} = \cos^{-1} \dfrac{\lambda}{4H} \end{cases}$$

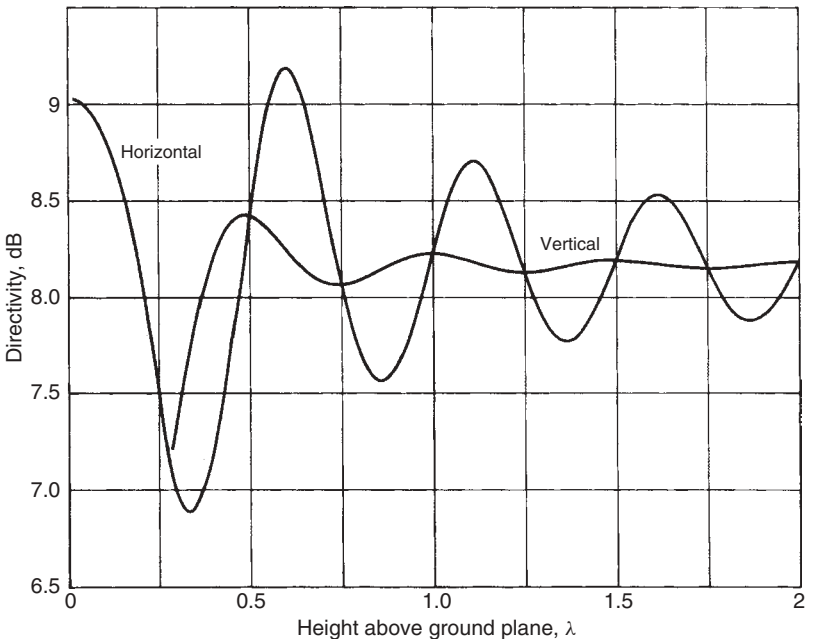


FIGURE 5-7 Directivity of half-wavelength dipoles over a ground plane.

The dipole pattern [Eq. (5-2)] increases the radiation intensity. The total input power into the single dipole becomes

$$P_{\text{in}} = |I_0|^2(R_{11} - R_{12})$$

$$\text{directivity} = \frac{U_{A,\max} U_{d,\max}}{P_{\text{in}}/4\pi}$$

After inserting the various terms, we obtain the directivity of a horizontal dipole over ground:

$$\text{directivity} = \begin{cases} \dfrac{4\eta \sin^2(2\pi H/\lambda)[1 - \cos(kL/2)]^2}{\pi(R_{11} - R_{12})} & H \le \dfrac{\lambda}{4} \\ \dfrac{4\eta[1 - \cos(kL/2)]^2}{\pi(R_{11} - R_{12})} & H \ge \dfrac{\lambda}{4} \end{cases}$$

Its plot is included in Figure 5-7.

5-5 DIPOLE MOUNTED OVER FINITE GROUND PLANES

Most configurations have a dipole mounted over a finite ground plane. You can calculate the final pattern by using GTD, PO, or MOM, or you can measure the pattern using the actual ground plane. Analyses produce idealized patterns, and measurements contain errors due to the presence of the positioner mounting. If the final system requires

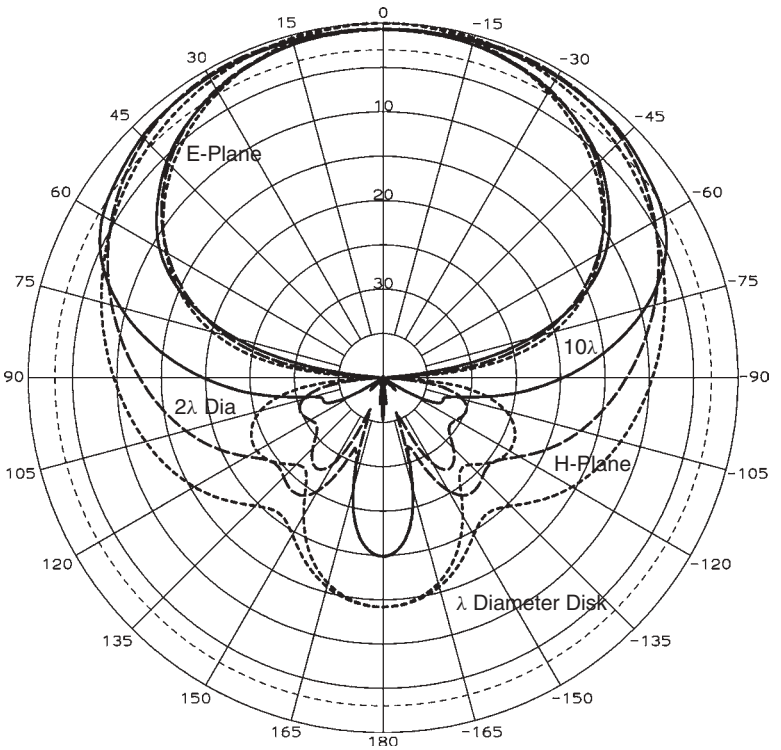


FIGURE 5-8 Dipole spaced λ/4 over disk ground planes with 1λ, 2λ, and 10λ diameters.

exacting patterns, it has no margin and will fail. In this section we consider dipoles with idealized ground planes and give you ideas about the final performance or spur you to use the ground plane purposely as a design parameter.

Figure 5-8 shows the result of a PO analysis of a dipole mounted $\lambda/4$ above finite disks 1, 2, and 10λ in diameter. The E-plane pattern contains a pattern null at 90° due to the dipole pattern. The ground plane restricts the broad H-plane pattern as pattern angles approach 90° and reduces the backlobe more and more as it increases in size. At one wavelength the disk increases the gain of the antenna from the 7.5 dB given in Figure 5-7 to 8.1 dB. We can size the ground plane to produce small gain increases.

We can analyze flat-plate reflectors from three perspectives. In the first, plates restrict radiation directions and thereby increase directivity. Waves polarized parallel with the surface must vanish on the reflector surface and cause a greater restriction of the beam. We see this effect in Figure 5-7, which shows horizontal dipoles having greater directivities than vertical dipoles for close spacing over a ground plane. In the second method we use aperture theory to analyze the reflector by using an aperture plane and integrate the fields or evaluate illumination losses. If the phase of the fields on the aperture varies rapidly, we must either take fine increments in numerical integration or evaluate only around areas of stationary phase. Third, we can replace the reflector with images and restrict the valid pattern region. In GTD this method is combined with diffractions to smooth the field across shadow and reflection boundaries.

In Section 5-4 we analyzed the pattern and gain of a dipole mounted over an infinite ground plane by the method of images. The antenna and its image formed a two-element

TABLE 5-1 Results of a GTD Analysis of a Horizontal λ/2 Dipole λ/4 Over a Limited Square Ground Plane (*H*-Plane)

Ground-Plane Size (λ)	Directivity (dB)	Front-to-Back Ratio (dB)	*H*-Plane Pattern Level at 90° (dB)	*H*-Plane Beamwidth (deg)	Phase Center (λ)
0.5	5.37	8.4	−6.3	108.5	0.18
0.6	6.32	10.3	−7.6	104.0	0.15
0.7	7.08	12.0	−8.8	100.9	0.14
0.8	7.68	13.5	−9.8	97.8	0.12
0.9	8.14	14.8	−10.6	95.1	0.10
1.0	8.34	16.0	−11.2	93.2	0.08
1.2	8.65	17.8	−12.0	93.3	0.04
1.4	8.45	19.1	−12.3	99.4	0.01
1.6	7.96	20.0	−12.2	108.4	0.0
1.8	7.39	21.1	−12.3	112.4	0.0
2.0	6.95	22.3	−12.4	113.1	0.0
2.5	7.13	25.0	−12.7	115.8	0.0
3.0	7.74	28.3	−13.8	111.4	0.0
4.0	7.28	32.8	−14.8	116.1	0.0
5.0	7.56	35.4	−16.2	118.0	0.0
10	7.41	36	−19.1	121.3	0.0

array, but with real power into only one element. The imaging method gives limited information that can be filled with GTD methods. Table 5-1 lists the results of a GTD analysis of a half-wavelength horizontal dipole located λ/4 over a limited square ground plane. An infinite ground plane and dipole combination has an infinite front-to-back (F/B) ratio with the fields vanishing in the ground-plane direction. By using the methods of Section 3-3, we calculate a 120° half-power beamwidth for the two-element half-wavelength spaced array of the dipole and its image. The F/B ratio increases as the reflector (ground plane) size increases. Unfortunately, F/B is only the ratio of two pattern angles. We could tune the size of the ground plane to produce a high F/B ratio for a nonsquare ground plane, but it holds for only a small range of angles. Figure 5-8 illustrates the general increase in F/B as the size of the ground plane increases. We expect zero fields at $\theta = 90^\circ$ on an infinite ground plane, and Table 5-1 shows a decrease of the fields with an increase of the ground plane. The half-power beamwidth cycles about 120° as the ground plane increases in size.

Phase center is the apparent radiation center placed at the focus of a paraboloidal reflector when used as a feed. The phase center of the equivalent two-element array is located on the ground plane. As we decrease the ground plane, the effect of the image decreases and causes the phase center to move toward the dipole. In the limit of no ground plane, the phase center is on the dipole.

Table 5-1 shows the small gain changes that occur as the relative phase of the ground-plane scattered fields and the dipole direct fields add in the far field. The small ground plane at λ/2 square fails to significantly limit radiation and gain drops. Peak gain occurs when the ground plane is 1.2λ square, but this result would not necessarily hold for a circular ground plane. In most applications the dipole cannot be mounted directly above the ground-plane center, but we can add a small ground plane to control

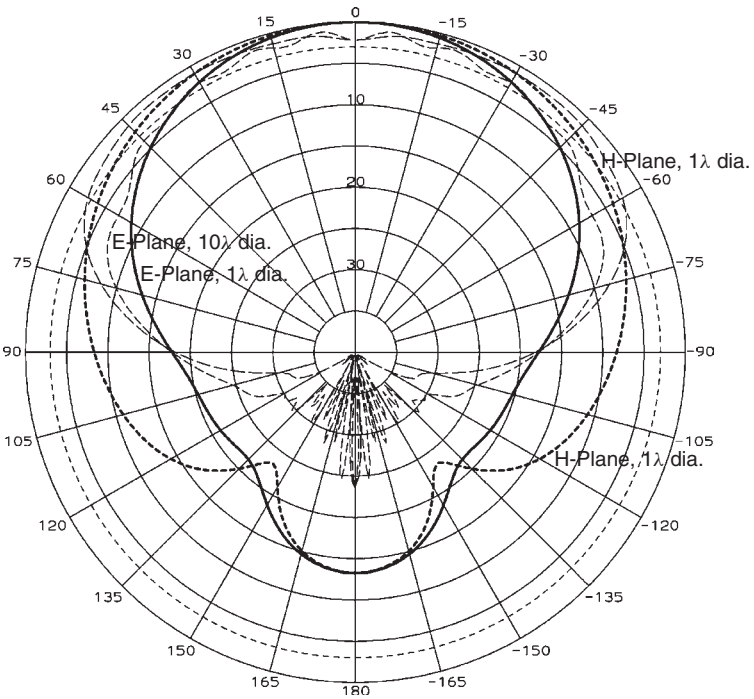


FIGURE 5-9 V-dipole spaced 0.35λ over and tilted 35° toward 1λ- and 10λ-diameter ground planes.

the pattern and then place the combination on a pedestal over the larger ground plane. Most cases should be analyzed or measured in the final configuration.

The dipole *E*-plane pattern null can be reduced by tilting the two poles down toward the ground plane. Figure 5-9 illustrates the calculated pattern of a tilted element dipole above a finite disk ground plane. The feed point of the dipole has been raised to 0.35λ to allow for the 35° tilt to the poles. Tilt and ground-plane height give additional parameters to control the pattern of the dipole mounted over a finite ground plane. For example, a horizontal dipole located $\lambda/2$ over an infinite ground plane forms an odd-mode (0°, 180°) two-element array using the dipole and its image. The simple ray-tracing argument given in Section 3-1 predicts a pattern null at zenith. But when placed over a finite ground plane, the fainter image fails to produce a complete null.

We sometimes mount a dipole spaced away from a metal cylinder that provides a ground plane to restrict radiation. The curved ground plane allows greater radiation around the cylinder when rays spread as they scatter from it. Figure 5-10 shows the horizontal plane pattern for a vertical dipole mounted near a 1λ-diameter cylinder for spacing of 0.25λ, 0.4λ, 0.5λ, and 0.75λ. When we space a dipole $\lambda/2$ above a large flat ground plane, the pattern has a null normal to the plane. The cylinder is unable to generate a full image of the dipole to produce this null, but the pattern does dip 11.2 dB from the peak. A dipole spaced $3\lambda/4$ over a ground plane produces a three-lobed pattern that we can see in Figure 5-10 except that the cylinder can produce only 8-dB dips. If we mount the dipole over a 2λ-diameter cylinder, the pattern is similar to Figure 5-10 except that F/B increases and the nulls have greater depths. Table 5-2 summarizes pattern results for vertical dipoles mounted over small cylinders.

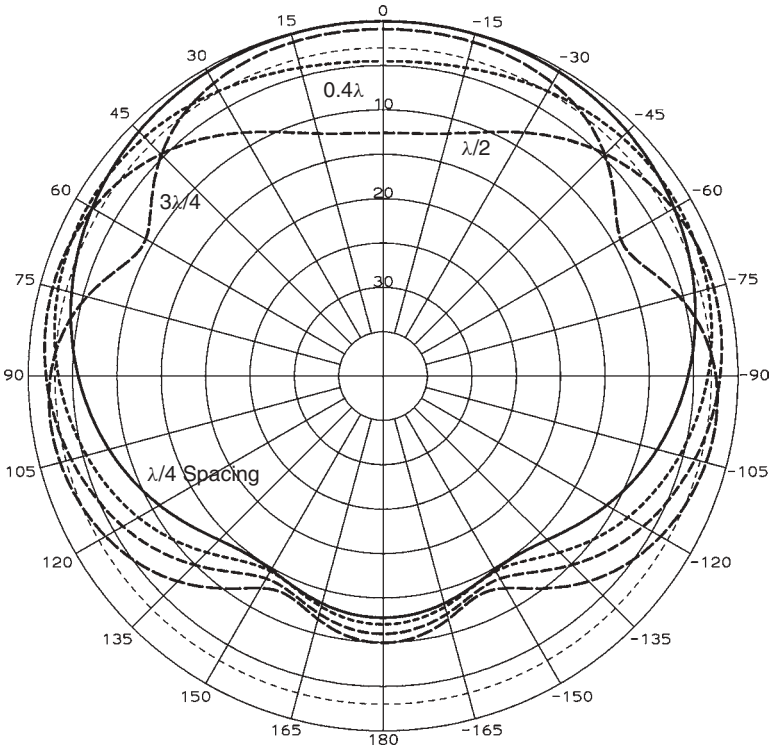


FIGURE 5-10 Horizontal plane pattern for a vertical dipole mounted near a 1λ-diameter cylinder at 0.25λ, 0.4λ, 0.5λ, and 0.75λ distances.

TABLE 5-2 Dipole Mounted Over a Cylinder Aligned with a Cylinder

Height (λ) Over Cylinder	Cylinder Diameter (λ)	Gain (dB) At 0°	Gain (dB) At 180°	Gain Peak	Peak Angle
0.25	0.25	3.5	−2.1	3.6	0
	0.50	6.1	−2.7	6.1	0
	1.0	6.7	−6.1	6.7	0
	2.0	7.3	−10.7	7.3	0
0.4	0.25	3.2	0.3	4.9	64
	0.50	3.6	−1.3	6.0	62
	1.0	2.2	−5.3	5.1	60
	2.0	2.2	−9.7	6.0	54
0.5	0.25	0.5	0.9	5.2	80
	0.50	−2.9	−1.8	4.8	80
	1.0	−5.9	−4.2	5.2	76
	2.0	−8.7	−8.5	5.9	70
0.75	0.25	3.3	−0.2	3.4	102
	0.50	5.0	−0.9	4.7	102
	1.0	5.1	−3.2	4.6	98
	2.0	6.4	−6.8	5.2	90

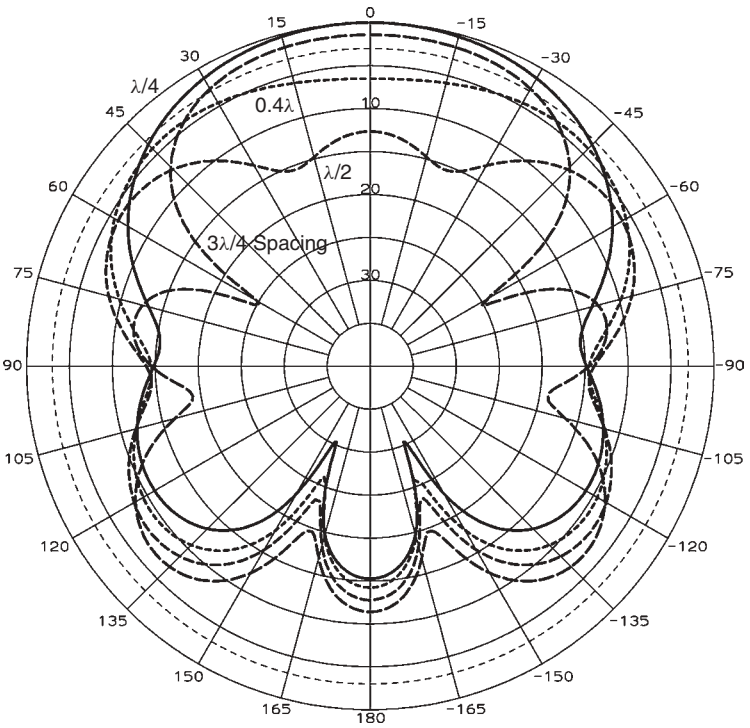


FIGURE 5-11 Horizontal dipole mounted over a vertical 1λ diameter pole at 0.25λ, 0.4λ, 0.5λ, and 0.75λ distances.

TABLE 5-3 Dipole Mounted Over a Cylinder Perpendicular to a Cylinder

Height (λ) Over Cylinder	Cylinder Diameter (λ)	Gain (dB) At 0°	Gain (dB) At 180°	Gain Peak Perpendicular to Plane	Peak Angle Perpendicular to Plane
0.25	0.25	3.8	−2.6	3.8	0
	0.50	6.6	−4.4	6.6	0
	1.0	7.1	−8.2	7.1	0
	2.0	7.0	−8.6	7.4	30
0.4	0.25	3.0	−0.2	4.9	46
	0.50	2.9	−2.5	6.5	48
	1.0	0.6	−7.1	6.4	50
	2.0	1.2	−7.3	7.3	54
0.5	0.25	0.0	0.5	5.1	54
	0.50	−4.2	−2.8	5.2	54
	1.0	−5.6	−5.7	6.6	56
	2.0	−6.7	−6.3	7.3	60
0.75	0.25	3.5	−0.4	2.7	66
	0.50	5.3	−1.6	4.3	66
	1.0	5.7	−4.3	5.1	66
	2.0	6.4	−5.7	5.8	70

To complete the analysis, the dipole was rotated so that its axis is perpendicular to the pole. Figure 5-11 illustrates the patterns calculated for a horizontally polarized dipole mounted above a vertical pole. We expect a pattern null at 90° in this horizontal plane due to the dipole polarization null, but the dipole induces curved currents on the cylinder that radiate and fill in these nulls. The null due to the dipole does narrow the pattern in the horizontal plane compared to Figure 5-10, and in many cases peak radiation occurs in the vertical plane. Table 5-3 lists the characteristics of the horizontal dipole mounted over a vertical pole for various dipole spacing above the pole and its diameter.

5-6 CROSSED DIPOLES FOR CIRCULAR POLARIZATION

We produce a circularly polarized antenna by placing two dipoles along the x- and y-axes over a ground plane and feeding them with equal amplitudes and quadrature phase (0° and −90° for RHC). Without the ground plane the combination radiates LHC in the $-z$ direction. The ground plane changes the sense of circular polarization of the wave radiated in the $-z$-direction and it adds with the direct radiated wave. The dipoles are fed from either dual folded baluns that produce two separate inputs or by a split coax balun connecting both dipoles in shunt. The shunt connection requires differing lengths for the dipoles to produce the 90° phase difference that we call the *turnstile configuration*.

The dual-feed antenna uses either a quadrature hybrid equal-amplitude power divider to feed the two ports or an equal phase and amplitude power divider with an extra line length on one of the two ports. The hybrid power divider feed produces an antenna with a wide impedance and axial ratio bandwidth. The hybrid power divider has two inputs that provide ports for both RHC and LHC polarizations. The signals reflected from the two equal-length dipoles when fed from one port of the hybrid reflect into the second port due to the phasing in the hybrid coupler. When measuring at one port of the hybrid, the impedance bandwidth is quite broad because the reflected power is dissipated in the load on the other port. This dissipated power lowers the efficiency of the antenna, a hidden loss unless you measure the coupling between the inputs of the hybrid. The second configuration, using the extra line length, produces an antenna with a narrowed axial ratio bandwidth and a wider impedance bandwidth compared to a single dipole. The extra 180° round-trip total signal path in one arm causes the equal reflections to cancel. Figure 5-12 gives the circularly polarized pattern from a pair of crossed dipoles over a ground plane with a perfect feed. The E-plane dipole null limits the angular range of good circular polarization. We improve the circular polarization by raising the dipoles a little and tilting them down to widen the E-plane beamwidth. Figure 5-12 shows the pattern for the tilted dipole pair and illustrates the improved cross polarization and the wider beamwidth. The placement on a finite ground plane complicates this result somewhat and will require extra design effort.

Turnstile feeding exploits the impedance properties of the dipole to shift the relative phase between two different dipoles when shunt connected to the same port. When we shorten a dipole below resonance, its impedance is capacitive and its current has positive phase relative to the resonant-length dipole, while the lengthened dipole has an inductive reactance and a negatively phased current. We determine the lengths of

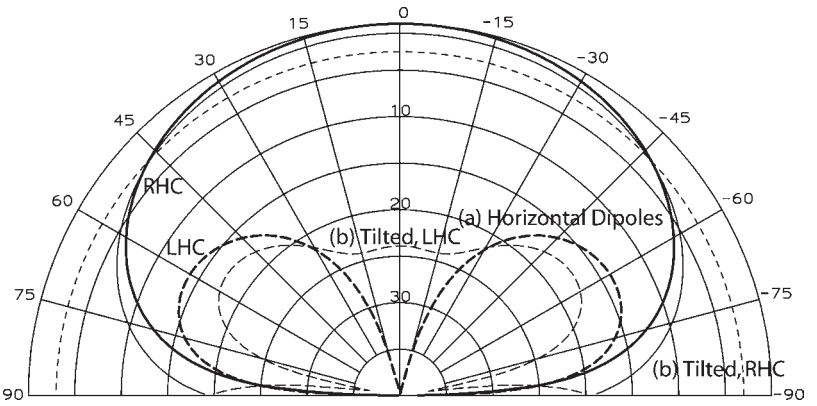


FIGURE 5-12 Crossed dipoles fed for circular polarization: (*a*) λ/4 height and 0° tilt; (*b*) 0.3λ height and 30° tilt.

the two dipoles by a perturbation technique using the Q of the resonant circuit of the dipole. Q is related to the VSWR bandwidth:

$$\text{BW} = \frac{\text{VSWR} - 1}{Q\sqrt{\text{VSWR}}} \qquad Q = \frac{\text{VSWR} - 1}{\text{BW}\sqrt{\text{VSWR}}} \tag{5-10}$$

We derive the lengths of the two dipoles in terms of the resonant (zero reactance)-length dipole, L_0:

$$L_x = \frac{L_0}{\sqrt{1 + 1/Q}} \qquad L_y = L_0\sqrt{1 + \frac{1}{Q}} \qquad \text{RHC polarization} \tag{5-11}$$

A dipole of 0.014λ diameter located 0.3λ above a ground plane and tilted down 30° has a resonant length of 0.449λ. The 2 : 1 VSWR bandwidth for 70 Ω is 18.3% or a Q of 3.863 by using Eq. (5-10). When we insert this Q in Eq. (5-11), we calculate the two lengths for a turnstile design: $L_x = 0.400\lambda$ and $L_y = 0.504\lambda$ for RHC polarization. The $+x$ and $+y$ poles are fed from the same port. Figure 5-13 plots the Smith chart of this design. The trace on a Smith chart rotates clockwise for increasing frequency. The cusp in the trace is the frequency with the best axial ratio, which did not occur at the frequency of best match. Nevertheless, the 2 : 1 VSWR bandwidth of the antenna has increased to 41.5% because the combined reactance of the two dipoles cancels over a large frequency range. At center frequency the pattern is similar to Figure 5-12 except that the patterns in the two planes have slightly different beamwidths due to the dipole lengths. When the frequency shifts off center, the axial ratio degrades. The axial ratio bandwidth is far less than the impedance bandwidth, and the design gives a 16.4% 6-dB axial ratio bandwidth. An axial ratio of 6 dB produces 0.5-dB polarization loss similar to the 0.5 dB reflected power loss of 2 : 1 VSWR. This illustrates the importance of considering not only the impedance bandwidth but also the pattern characteristics over the frequency band.

We can increase the beamwidth of the turnstile dipole located over a ground plane by adding a notched cone under it. Figure 5-14 illustrates the arrangement of the slightly

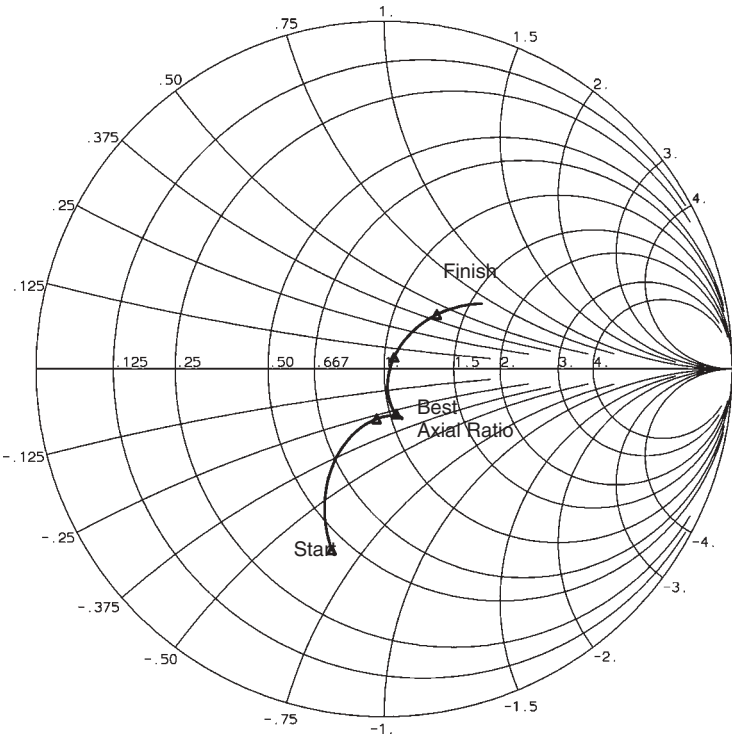


FIGURE 5-13 Smith chart response of a turnstile dipole pair $L_x = 0.400\lambda$ and $L_y = 0.504\lambda$ mounted 0.30λ over a ground plane with 30° tilt.

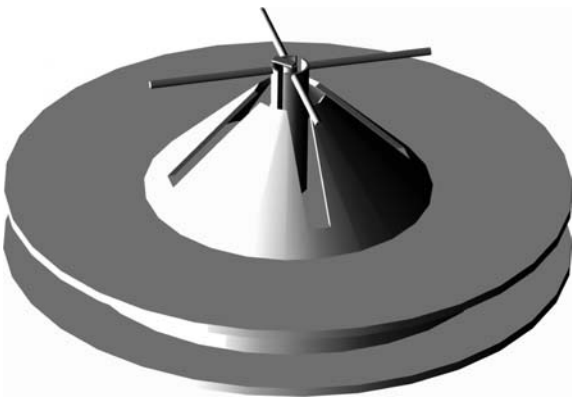


FIGURE 5-14 Turnstile dipole mounted over a notched cone on a finite circular ground plane with radial line chokes to reduce the backlobe.

less than λ/4-long notches in a 45° cone with the turnstile dipoles located about λ/4 above the ground plane. A split-tube coaxial balun feeds the two dipoles sized as a turnstile with dipoles of longer and shorter length. The upper feed jumper excites RHC radiation. The dipoles excite magnetic currents in the slots that radiate a broad pattern to fill in the E-plane nulls of the dipoles. On an infinite ground plane the horizontal

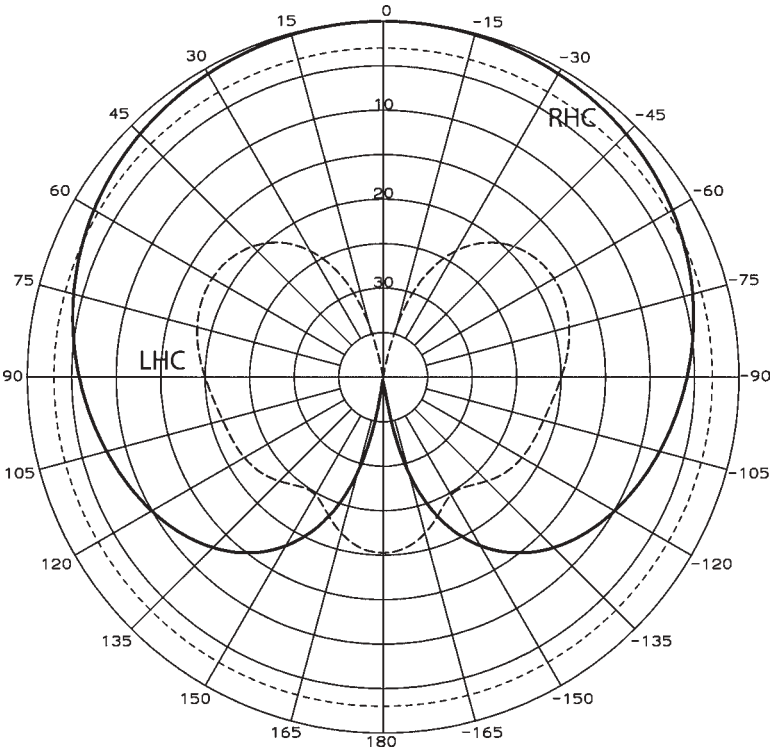


FIGURE 5-15 Pattern of a turnstile dipole mounted over a notched cone with a 0.75λ ground plane and two radial chokes.

polarization must vanish along the ground plane, and the RHC and LHC components would be equal at 90° similar to the pattern shown in Figure 5-12. By using a finite-size ground plane, the horizontal component does not vanish, and a wide beamwidth is obtained with circular polarization at 90° as shown in Figure 5-15, which uses a 0.75λ-diameter ground plane and 0.5λ-base-diameter cone. To reduce the backlobe below the ground plane, two short-circuited radial transmission line chokes were placed around the edge to form a soft surface. We size the inner radius so that the transmission line produces an open-circuit impedance at the outer rim that reduces the edge diffraction and the backlobe [6, p. 88]. From a PO perspective the radial line choke is a slot that supports a magnetic current loop. This example illustrates that slots or notches can be used to shape the patterns of small antennas.

5-7 SUPER TURNSTILE OR BATWING ANTENNA [7]

The super turnstile or batwing antenna was developed for TV transmitter antennas. The antenna combines a slot with a dipole batwing to produce an antenna with a wide impedance bandwidth. Figure 5-16 shows the normal configuration, with four wings placed around a central support metal mast. Each wing connects to the mast at the top and bottom with a metal-to-metal connection. The inner vertical rod and the support

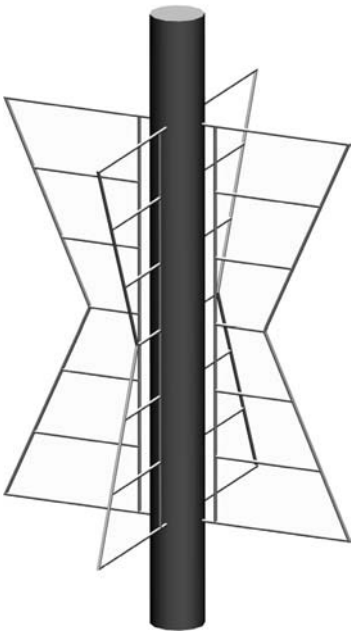


FIGURE 5-16 Super turnstile or batwing antenna using an open rod construction.

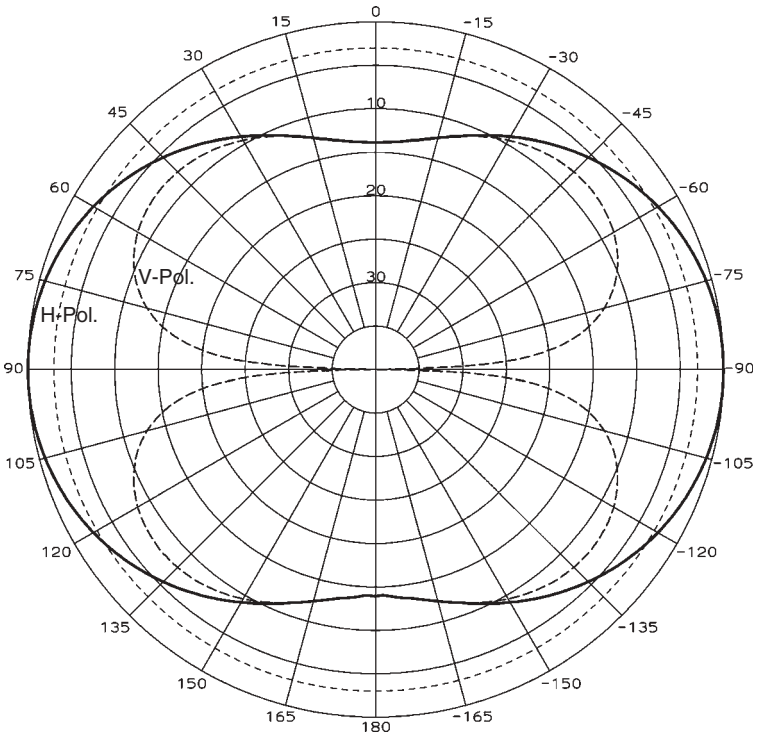


FIGURE 5-17 Elevation pattern of a single bay of a super turnstile antenna showing horizontal and vertical polarization components.

mast form a two-line slot fed by a jumper located at the center of each wing. To produce an omnidirectional pattern about the mast, a feed power divider located inside the mast phases the inputs for circular polarization (0°, 90°, 180°, 270°). The antenna radiates horizontal polarization in the horizontal plane but radiates cross-polarization that increases with elevation (depression) angle as shown in Figure 5-17. A four-wing antenna produces a horizontal plane pattern ripple of about 1.5 dB. Adding more wing antennas around a larger central mast reduces the ripple.

The extraordinary characteristic of the antenna is its impedance bandwidth. Figure 5-18 gives the return loss frequency response for a wire frame antenna. The 1.1 : 1 VSWR bandwidth is about 35%; if adjusted to 1.25 : 1 VSWR, the antenna has a 51% bandwidth. You make small adjustments to the spacing between the mast and the inner rod to tune the VSWR. Table 5-4 lists the parameters of batwing antennas with both wire frame and solid panel wings. The solid panels lower the input impedance to 75 Ω from the 100 Ω of the wire frame antenna. Using an antenna with only two wings

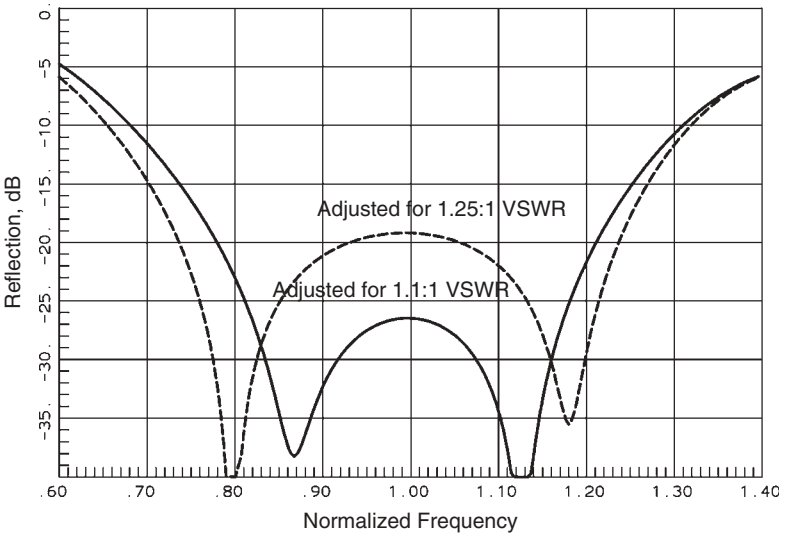


FIGURE 5-18 Super turnstile wire frame antenna return-loss response adjusted for 1.1 : 1 and 1.25 : 1 VSWR.

TABLE 5-4 Dimensions of a Super Turnstile Antenna in Wavelengths for Four Wings Center-Fed for Circular Polarization

Parameter	Wire Frame	Solid Wing
Impedance (Ω)	100	75
Height	0.637	0.637
Wing upper	0.2254	0.229
Wing middle	0.0830	0.0847
Gap	0.0169	0.0216
Rod diameter	0.0508	
Mast diameter	0.0847	0.0847

changes the input impedance from the value for an antenna with four wings because the close coupling between the wings alters the impedance. It depends on the feeding mode. This holds for any antenna with close coupling: for example, a spiral antenna. You must feed it in the operating mode to measure the correct input impedance. The transmitter antenna will consist of a number of these antennas stacked vertically to produce a narrow pattern directed at the horizon.

5-8 CORNER REFLECTOR [8, p. 328]

The usual corner reflector (Figure 5-19) has a dipole located between two flat plates that limit directions of radiation. The angle between the reflectors can be any value, but 90° seems to be the most effective. On paper, decreased angles give better results, but only marginally. We could consider the flat plate as a limiting case. The tangential electric fields must vanish at the surface of the flat plates. We discover a greater restriction, since the fields can only decrease gradually in the limited space between the ground planes and the dipole. Most of the power is concentrated in lower-order spherical modes. In the limit of zero vertex distance, the single mode possible restricts the beamwidth to 45° in the H-plane.

We analyze the 90° corner reflector as an array by using the three images of the dipole in the ground planes (Figure 5-19) plus the real dipole. The array factor of the array of dipole and images is

$$\left|e^{jkd\cos\theta} + e^{-jkd\cos\theta} - (e^{jkd\sin\theta\sin\phi} + e^{-jkd\sin\theta\sin\phi})\right|^2$$

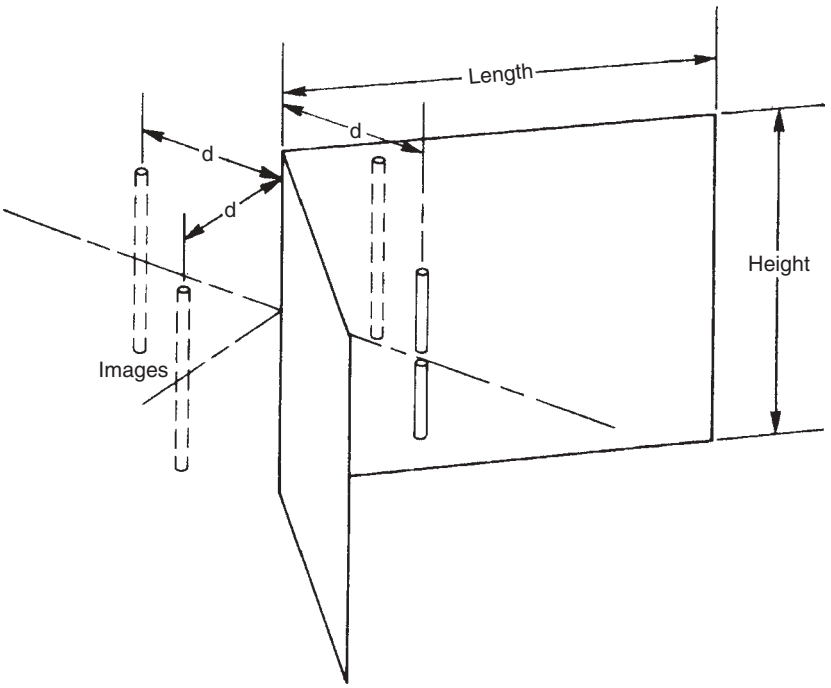


FIGURE 5-19 A 90° corner reflector.

In the H-plane, $\phi = 90^\circ$ and we evaluate terms to get

$$4[\cos(kd\cos\theta) - \cos(kd\sin\theta)]^2$$

where d is the distance from the vertex to the dipole and θ is the H-plane pattern angle from the axis. We must multiply this by the pattern of the dipole to obtain the radiation intensity. We consider only the H-plane, where the maximum radiation intensity is found from Eq. (5-9):

$$U = 4[\cos(kd\cos\theta) - \cos(kd\sin\theta)]^2 \frac{\eta|I_0|^2}{(2\pi)^2}\left(1 - \cos\frac{kL}{2}\right)^2 \tag{5-12}$$

where η is the impedance of free space, I_0 the dipole current, and L the dipole length.

The radiated power of the single dipole is

$$P_{\text{in}} = |I_0|^2[R_{11} + R_{12}(2d) - 2R_{12}(\sqrt{2}\,d)] \tag{5-13}$$

where R_{11} is the self-resistance of the dipole and $R_{12}(x)$ is the mutual resistance function between the dipole and its images. The directivity is found from

$$\text{directivity}(\theta) = \frac{4\pi U(\theta)}{P_{\text{in}}} \tag{5-14}$$

We combine Eqs. (5-12) and (5-13) into Eq. (5-14) to compute directivity of the 90° corner reflector with infinite sides:

$$\text{directivity}(\theta) = \frac{4\eta[1 - \cos(kL/2)]^2[\cos(kd\cos\theta) - \cos(kd\sin\theta)]^2}{R_{11} + R_{12}(2d) - 2R_{12}(\sqrt{2}\,d)} \tag{5-15}$$

Table 5-5 gives the directivity, beamwidth, and impedance of a 90° corner reflector fed from a dipole 0.42λ long and 0.02λ in diameter. We must shorten the dipole further than a free-space dipole length at resonance to compensate for the mutual coupling between dipoles. Directivity increases as the vertex distance decreases, but the effects of superdirectivity cause the efficiency and gain to fall as the vertex is approached. The antenna has a 50-Ω input impedance for $d = 0.37\lambda$. This point shifts when we increase the dipole's diameter to increase its bandwidth.

Kraus gives the following guidelines for the size of the sides. Each plate should be at least twice the length of the dipole-to-vertex distance, and the plate height (the dipole direction) should be at least 0.6λ. To evaluate those guidelines, a GTD analysis was performed on various combinations (Table 5-6) with $d = 0.37\lambda$. The H-plane beamwidth decreases with an increase in plate length. After about 1.5λ sides, the H-plane beamwidth fluctuates about 45° as the sides increase. Even with 5λ sides the beamwidth is below 45°. The E-plane beamwidth fluctuates with the plate height. The directivity was estimated from the beamwidths. In one case—1.5λ sides and 1.5λ high—the estimated directivity exceeds the directivity of the infinite-side case. The edge diffractions add to the reflected and direct radiation of the rest of the antenna.

Refer to Section 2-4.2 for an example using PO to analyze a corner reflector. Similar to infinite plate analysis, the reaction of the image dipoles in the finite plates can be used to find the input impedance and gain of the antenna. When we analyze the corner

TABLE 5-5 Characteristics of a 90° Corner Reflector with Infinite Sides and 0.42λ Dipole

Vertex Distance (λ)	Directivity (dB)	Beamwidth (deg)	Input Impedance (Ω)
0.30	12.0	44.7	$29.1 - j1.1$
0.32	12.0	44.6	$34.9 + j0.4$
0.34	11.9	44.5	$40.9 + j1.1$
0.36	11.9	44.3	$47.0 + j0.8$
0.37	11.9	44.2	$50.0 + j0.3$
0.38	11.8	44.1	$53.0 - j0.5$
0.40	11.8	43.9	$58.8 - j2.8$
0.42	11.7	43.6	$64.1 - j6.0$
0.44	11.7	43.3	$68.8 - j10.0$
0.46	11.6	42.9	$72.7 - j14.9$
0.48	11.5	42.4	$75.7 - j20.3$
0.50	11.4	41.8	$77.7 - j26.2$
0.52	11.4	41.1	$78.6 - j32.2$
0.54	11.3	40.2	$78.4 - j38.4$
0.56	11.2	39.2	$77.0 - j44.3$
0.58	11.1	38.1	$74.6 - j49.8$
0.60	10.9	36.8	$71.3 - j54.8$

TABLE 5-6 Results of a GTD Analysis of a 90° Corner Reflector with Finite Sides and Vertex Distance 0.37λ

Side Length (λ)	Plate Height (λ)	Beamwidth		F/B (dB)	Estimated Directivity (dB)
		E-Plane	*H*-Plane		
0.75	0.75	70.4	97.4	18.4	7.7
1.00	0.75	73.6	72.4	17.3	8.8
1.50	0.75	72.6	50.8	18.2	10.0
0.75	1.00	60.2	91.6	23.4	8.5
1.00	1.00	61.0	62.8	22.7	10.1
1.50	1.00	58.5	46.0	23.8	11.4
0.75	1.50	53.4	81.6	34.0	9.3
1.00	1.50	51.6	60.0	39.0	11.0
1.50	1.50	48.2	42.6	46.3	12.6
5.00	5.00	68.8	43.4	63.5	10.8

reflector using GTD, the method does not determine input impedance and gain must be estimated from the patterns. We can use the method of moments to analyze the corner reflector. One preferred construction method is to use rods for the reflector so that the antenna has minimum wind loading. Figure 5-20 illustrates a corner reflector made with only six rods on each side. Figure 5-21 gives the pattern of this antenna from a moment method calculation. This small antenna produces excellent results.

We can use the angle of the sides as a design parameter. A geometric optics analysis that uses images restricts the angle, but nothing stops the antenna from working for

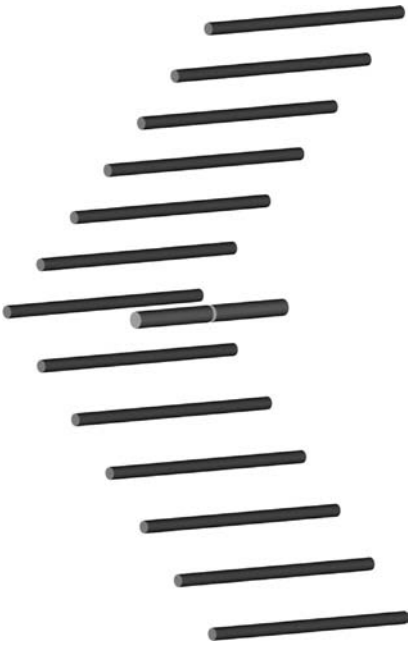


FIGURE 5-20 Corner reflector constructed from 0.6λ-long rods spaced 1/6λ with a dipole spaced 0.37λ from the vertex.

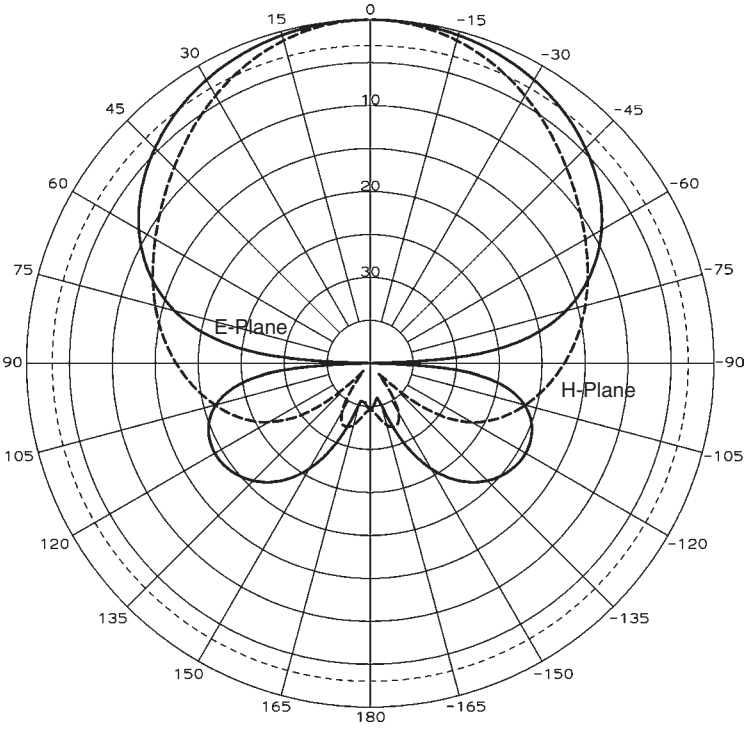


FIGURE 5-21 Pattern of a corner reflector made from 0.6λ-long rods spaced 1/6λ with dipole 0.37λ from vertex.

TABLE 5-7 Corner Reflector with Varying Angle H-Plane 0.9λ-Wide Plates Connected to a 0.2λ Central Plate, 1λ E-Plane Width, Dipole 0.3λ Above the Central Plate

Side Angle	Beamwidth		Gain (dB)	F/B (dB)	Side Angle	Beamwidth		Gain (dB)	F/B (dB)
	E-Plane	H-Plane				E-Plane	H-Plane		
60	58.1	59.4	9.1	21.7	15	59.2	65.9	9.5	19.3
55	57.1	56.0	9.9	22.6	10	61.6	83.3	8.7	26.3
50	56.3	52.3	10.5	23.0	5	64.7	99.2	7.8	25.1
45	55.8	49.1	10.8	23.4	0	67.8	108.6	7.5	23.6
40	55.4	46.8	11.1	22.8	−5	70.2	117.0	7.2	22.0
35	55.4	45.5	11.2	24.3	−10	71.5	125.8	6.9	16.5
30	55.6	45.6	11.0	24.9	−15	71.8	135	6.6	19.0
25	56.2	47.9	10.7	25.5	−20	71.7	143.8	6.3	17.8
20	57.4	53.7	10.2	20.3	−25	71.8	152.2	5.4	16.1

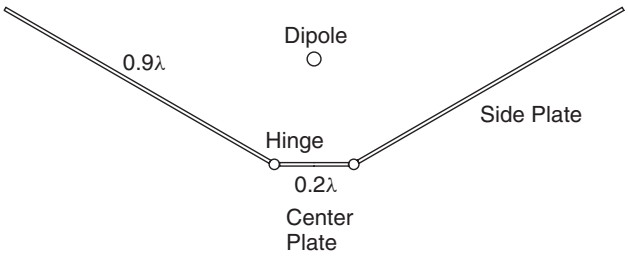


FIGURE 5-22 Corner reflector with variable-angle side plates and a center flat plate.

arbitrary side angles. It is convenient to have a small plate between the tilted sides for the mounting brackets, and these side plates could be mounted on hinges and rotated to vary the H-plane beamwidth. Table 5-7 lists the parameters of a corner reflector 1λ along the E-plane, a central plate 0.2λ wide in the H-plane, and sides 0.9λ long where the side angle is varied. The dipole is located 0.3λ above the central plate. We measure the side plate angle from the plane containing the small central ground plane; zero corresponds to a flat plane ground plane and 45° the usual corner reflector. Negative side-plate angle means that the side plates are tilted behind the central plate away from the dipole. Figure 5-22 illustrates the H-plane cross section of this corner reflector with 30° side plates.

We should not design corner reflectors with large sides since the gain is limited. The gain of paraboloid reflectors of the same size soon exceeds that of a corner reflector. A 2λ-diameter paraboloid reflector at 50% efficiency has a gain of 13 dB, and its gain exceeds that of a corner reflector. Any corner reflector with a vertex angle given by $180°/N$, where N is an integer, can be analyzed by the method of images. Corner reflectors with N greater than 2 have only marginally higher gains. The 90° corner reflector gives the best result for a given amount of material. Elkamchouchi [9] adds a cylindrical surface between the plates centered on the vertex. This surface adds another set of images within the cylinder. The images increase the gain by about 2 dB and decrease the frequency dependence of the impedance.

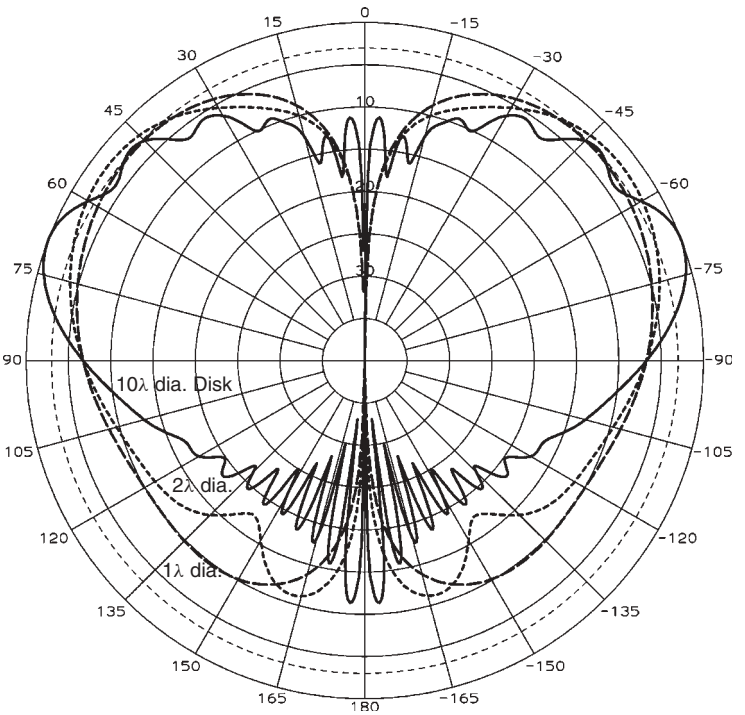


FIGURE 5-23 A λ/4 monopole located on 1λ-, 2λ-, and 10λ-diameter disk ground planes.

5-9 MONOPOLE

A monopole consists of a single conductor fed out of a ground plane from the center conductor of a coax. When we include its image (Figure 5-6), the monopole equates to a dipole for analysis. The fields vanish below the ground plane and restricting the fields to the upper hemisphere doubles the gain over a dipole, since only half the input power of the dipole is needed to produce the same field strength.

The input impedance decreases to half that of the equivalent dipole. We can form the image of the voltage source feeding the monopole in the ground plane. The voltage across the input of the equivalent dipole is twice that of the monopole to produce the same current. Therefore, the impedance of the monopole is half the impedance of the dipole.

The large value of edge diffraction greatly limits the F/B ratio of a monopole when it is placed on a finite ground plane. Figure 5-23 shows the pattern of a monopole when placed on 1λ-, 2λ-, and 10λ-diameter circular ground planes. The back radiation can be reduced by placing the monopole over a ground plane with circular corrugations that forms a soft surface at the edge when the corrugations are slightly deeper than λ/4 [10]. When the corrugations are less than λ/4, the ground plane can support surface waves.

5-10 SLEEVE ANTENNA [8, p. 422; 11, Chap. 5; 12; 13, p. 278]

A sleeve around the monopole (Figure 5-24) moves the virtual antenna feed up the monopole. The bandwidth increases because the current at the feed point remains nearly

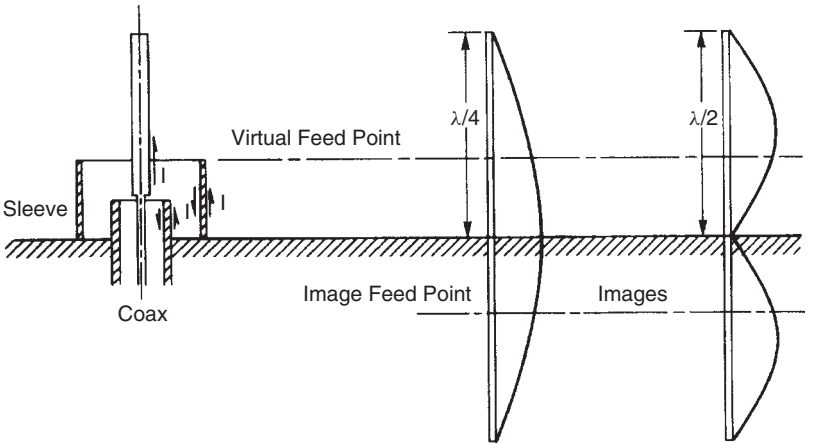


FIGURE 5-24 Sleeve monopole and current distributions.

constant over a wide band. Currents at the input for the case when the monopole is a quarter-wavelength long and when it is a half-wavelength long are about the same (Figure 5-24). The input resistance remains constant as the frequency changes.

The sleeve shields possible radiation from the internal currents while the currents on the outside of the sleeve radiate. The pattern changes little from that of an unshielded monopole. The internal structure is available as a series-matching stub and a transformer to broadband the antenna. Design consists of adjusting the parts until a suitable compromise input impedance match is achieved over the band.

Dipole sleeve antennas (Figure 5-25) require symmetrical sleeves on the arms to maintain the symmetry of the currents. It is equivalent to feeding the antenna in two places. The balun is made an integral part of the base. In both antennas, strips or rods can replace the total coaxial sleeve [14]. The currents on the rods cancel the radiation from the currents on the internal feeder. Figure 5-26 illustrates an open-sleeve dipole using two rods designed to be mounted over a ground plane. The antenna is fed from a folded balun that consists of a grounded vertical coax with one pole connected to the outer shield and a matching tube connected to the second pole. The center conductor jumps across the gap to the second pole. Following are the design dimensions in wavelengths normalized to the lower-frequency band edge:

Dipole length	0.385	Dipole diameter	0.0214
Sleeve length	0.2164	Sleeve diameter	0.0214
Dipole-to-sleeve spacing	0.0381	Dipole height above ground	0.1644
Input taper	0.056		

Figure 5-27 plots the return-loss response of the antenna for various configurations and models of the antenna. The dipole without the sleeves has its best return loss over a narrow band centered at a normalized frequency of 1.05. The sleeves have little effect on this response at the low-frequency end. Adding sleeves produces a second resonance, which combines with the lower one to produce a broad bandwidth. An initial method of moments analysis used constant-diameter rods for the antenna, and Figure 5-27 shows the poor impedance match response of the antenna. A key element of the experimental

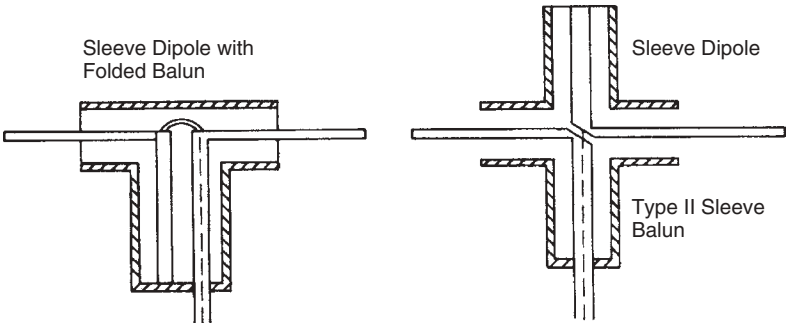


FIGURE 5-25 Sleeve dipoles.

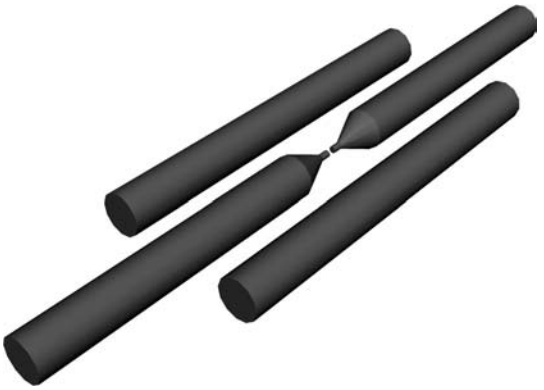


FIGURE 5-26 Open-sleeve dipole with conical input taper.

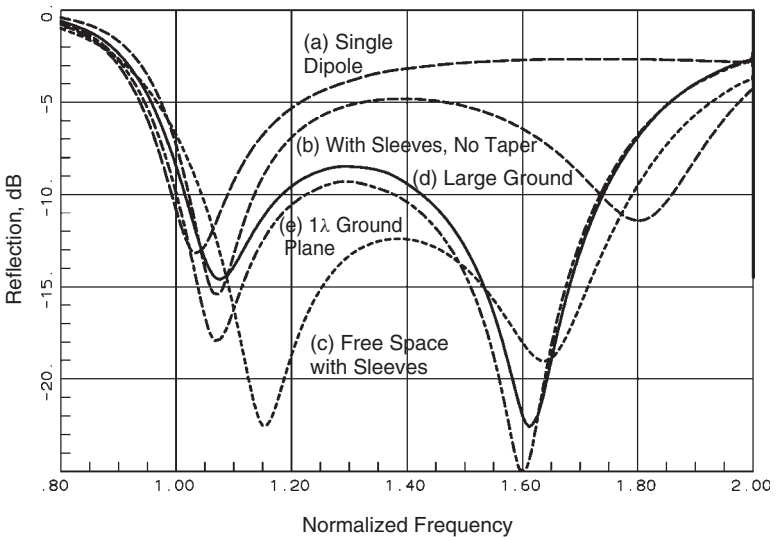


FIGURE 5-27 Return-loss response of an open sleeve dipole: (*a*) dipole without sleeves; (*b*) open sleeve antenna; (*c*) open sleeve antenna with tapered input; (*d*) open sleeve antenna with tapered input located $\lambda/4$ over ground plane; (*e*) open sleeve antenna over a 1λ-diameter ground plane.

antenna is the tapered input. Adding this feature to the model produced the improved broad response of the experimental antenna. The constant-diameter model response shows a notable capacitive term on a Smith chart, and the tapered input produced the necessary inductance to reduce this effect. If the antenna is located in free space, the impedance response improves as shown. Figure 5-27 points out the importance of analyzing an antenna in its operating environment. The dot-dashed curve illustrates the response when the antenna was mounted over a one-wavelength-square ground plane. The finite ground plane produces a small but noticeable change in the input impedance. The effects of small changes in the analytical model warn us that we cannot expect antennas to match their models exactly and that small mechanical details can be used to improve performance.

An open-sleeve antenna can be made using a wire cage. Since the diameters of the dipole and sleeve rods are large, the weight can be reduced by using a circular array of wires for each conductor. The effective diameter of the cage, d_{eff}, is given as

$$d_{\text{eff}} = d\left(\frac{nd_0}{d}\right)^{1/n} \quad \text{or} \quad \frac{d_0}{d} = \frac{1}{n}\left(\frac{d_{\text{eff}}}{d}\right)^{n} \tag{5-16}$$

The diameter of the individual wires is d_0, the cage diameter is d, and n is the number of wires.

5-11 CAVITY-MOUNTED DIPOLE ANTENNA

A dipole can be placed in a cup, and the assembly can be flush-mounted in a ground plane. The antenna shown in Figure 5-28 has disk sleeves located above and below the dipoles to stretch the bandwidth over a 1.8 : 1 range [15]. Following are the dimensions normalized to the dipole length:

$$\frac{D}{L} = 2.57 \qquad \frac{H}{L} = 0.070 \qquad \frac{S}{L} = 0.505$$
$$\frac{T}{L} = 0.68 \qquad \frac{G}{L} = 0.40$$

The operating range is $0.416\lambda \le L \le 0.74\lambda$. The antenna cavity ranged from 0.28λ to 0.50λ deep and can no longer be considered thin. The cup antenna has a nearly constant gain (± 0.5 dB) of 10.5 dB over the band. Mounting the antenna in a cavity opens up new possibilities, because extra parameters are added to the design. At the low-frequency end, the cavity diameter is 1.07λ, which grows to 1.90λ at the high end.

We can use a dipole in a cup as a reflector feed. Excellent pattern and impedance response is obtained with the dipole mounted in a truncated cone cup with a 0.88λ aperture diameter, a 0.57λ-diameter base, and a 0.44λ depth [16, pp. 106–108]. The dipole is foreshortened to 0.418λ for an element diameter of 0.013λ and mounted 0.217λ above the base to achieve a 21% 2 : 1 VSWR bandwidth for a single element. When we use a cross-polarized pair fed from a hybrid coupler to radiate CP, the impedance match at the input port improves. The signals reflected from the two dipoles add in phase at the isolated port and cancel at the input port. The load dissipates the reflected power, and the antenna through the hybrid presents an excellent impedance match.

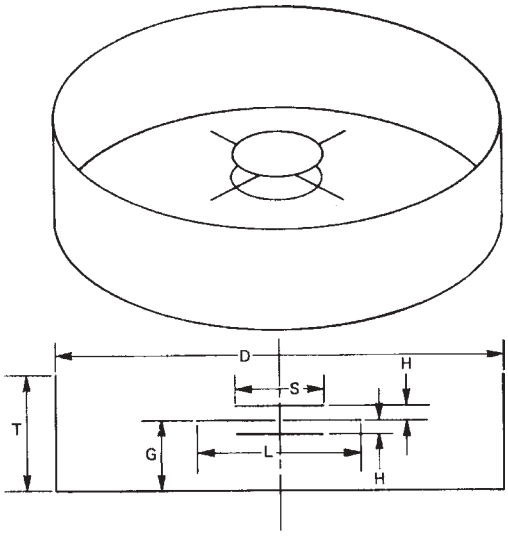


FIGURE 5-28 Cavity-mounted sleeve dipole antenna.

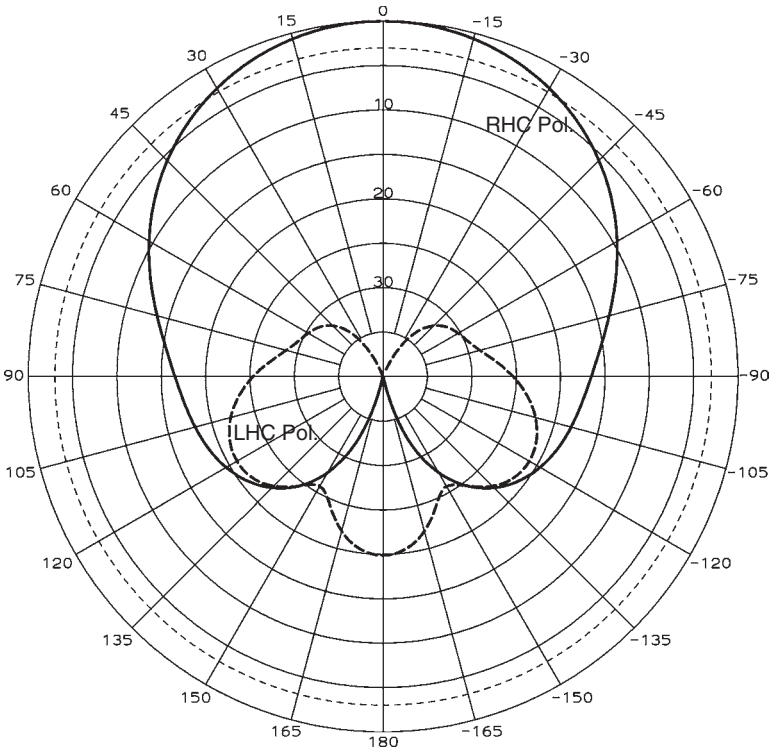


FIGURE 5-29 Circular polarization response of a crossed dipole mounted 0.217λ above the bottom of a truncated cone 0.44λ deep with a 0.88λ aperture and a 0.57λ base.

TABLE 5-8 Illumination Losses When Pattern of Figure 5-29 Feeds a Paraboloidal Reflector

	Loss (dB)	
f/D	Average	Maximum
0.36	1.69	1.74
0.38	1.60	1.66
0.40	1.54	1.65
0.42	1.50	1.65
0.44	1.49	1.68
0.46	1.50	1.72
0.48	1.52	1.77
0.50	1.55	1.83
0.52	1.60	1.91

Figure 5-29 plots its pattern when excited for CP. The cross-polarization is about 30 dB below the peak co-polarization response over the entire 10-dB beamwidth cone. It has the following illumination losses when the antenna is used as a paraboloidal reflector feed (see Section 8-2); for $f/D = 0.44$ and averaged over the 21% bandwidth:

$$\text{spillover loss} = 0.72\,\text{dB} \qquad \text{amplitude taper loss} = 0.65\,\text{dB}$$

$$\text{cross-polarization loss} = 0.12\,\text{dB}$$

Table 5-8 demonstrates the broad optimum reflector f/D for a phase center 0.02λ inside the aperture plane, where we position it at the reflector focus.

5-12 FOLDED DIPOLE

A half-wavelength folded dipole increases the input impedance of a normal dipole fourfold while radiating the pattern of a single dipole. With the two elements closely coupled, we analyze the antenna using even and odd modes (Figure 5-30). The even mode divides the antenna into separate dipoles because the magnetic wall halfway between them is a virtual open circuit. The input current to the even mode becomes

$$I_e = \frac{V}{2(Z_{11} + Z_{12})}$$

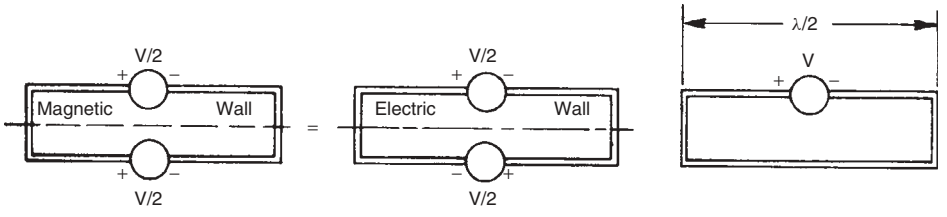


FIGURE 5-30 Folded dipole analysis modes.

where Z_{11} is the self-impedance of one of the dipoles and Z_{12} is the mutual impedance between the closely coupled dipoles. The odd mode reduces the antenna to the series connection of two nonradiating $\lambda/4$ stubs:

$$I_o = \frac{V}{jZ_0 \tan(kL/2)}$$

where Z_0 is the characteristic impedance between the two rods. The input current is the sum of the even- and odd-mode currents. Near $L = \lambda/2$, the odd-mode current is quite small because its input impedance is an open circuit, and the input impedance is then determined by the even mode only:

$$Z_{\text{in}} = \frac{V}{I_e} = 2(Z_{11} + Z_{12})$$

For closely coupled lines, $Z_{11} = Z_{12}$ and the input impedance becomes $Z_{\text{in}} = 4Z_{11}$, where Z_{11} is the self-impedance of the dipole. Higher input impedance levels can be obtained by adding more elements.

A second method of altering the step ratio from 4 is to use unequal feed and shorted element diameters [17,18]. Given a driven element radius a_1, parasitic element radius a_2, and center-to-center spacing b, Hansen [18] gives a convenient formula for the step-up ratio $(1 + \gamma^2)$:

$$\gamma = \frac{\cosh^{-1}[(v^2 - u^2 + 1)/2v]}{\cosh^{-1}[(v^2 + u^2 - 1)/2uv]} \qquad \text{where} \quad u = a_2/a_1 \text{ and } v = b/a_1 \qquad (5\text{-}17)$$

5-13 SHUNT FEEDING [19, p. 118]

Shunt feeding grows out of the folded dipole. The T-match (Figure 5-31) starts as a folded dipole when the taps are at the ends. As the taps move toward the center, the impedance of the dipole dominates at first, since the admittance of the shunt stub in the odd mode is small and the input impedance is capacitive. At some point, as the taps move toward the center, the inductive admittance of the stub will cancel the capacitive admittance of the dipole and produce antiresonance with its high input resistance. The location and magnitude of this peak resistance depends on the diameters of the rods in the T-match section and the diameter of the radiator. The input resistance decreases as we continue to move the tap point toward the center after the feed location passes the antiresonance point. The input impedance is inductive and match is achieved by using symmetrical series capacitors. The T-match is fed from a balanced line.

The center short on the dipole allows the direct connection of the dipole to ground. Direct connection of broadcast towers (monopoles) to ground gives some lightning protection because the transmitter is capacitively connected to the tower. Shunt feeding with a T-match enables solid conductors, such as the skin of an aircraft, to be excited as a dipole. Horizontal shunt-fed dipoles can be connected directly to vertical towers with a metal-to-metal connection to increase the strength of the antenna to withstand adverse weather conditions.

A gamma match (Figure 5-31) can be fed from an unbalanced coax line. The shield of the coax connects to the shorted center of the dipole while the center conductor

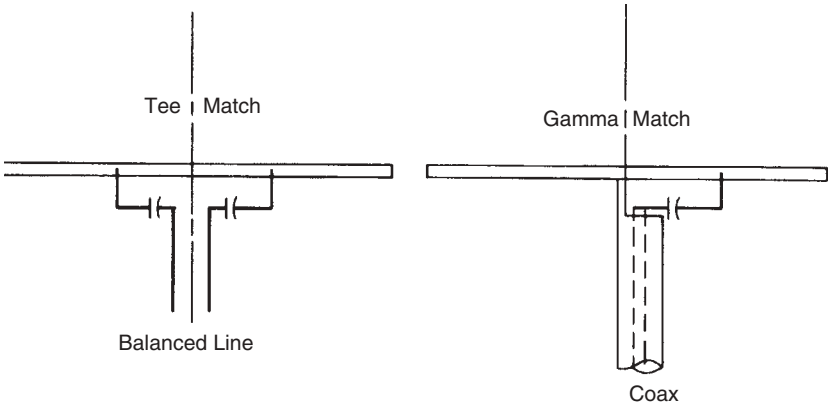


FIGURE 5-31 Shunt-fed dipoles.

taps into one side of the solid rod. Moving the tap away from the center increases the input resistance. The inductive reactance is series-tuned with a capacitor. Both of these connections reduce the bandwidth of the antenna as the input impedance is raised because the combination of the series capacitor and the shunt inductive stub increases the stored energy and Q of the antenna.

5-14 DISCONE ANTENNA

The discone antenna (Figure 5-32) is a modification of the dipole where the upper pole becomes a disk and the lower pole turns into a cone. We feed the antenna by locating a coax in the center of the cone and by connecting its outer shield to the lower cone at its top while we extend the coax center conductor and connect it to the disk. We obtain an antenna with a wide impedance bandwidth and a dipolelike pattern. As frequency increases the pattern peak moves toward the cone and gives a downward-pointing pattern. Figure 5-33 shows the pattern of a discone antenna at the design frequency and at two, three, and four times this frequency. The antenna produces less useful patterns as frequency increases. The antenna that gives the patterns in Figure 5-33 has a VSWR less than 3 : 1 from 1 to 10 times the design frequency. The cone upper diameter determines the high-frequency end of good impedance match. Typical slant length dimensions versus cone angle are as follows [20, pp. 128–130]:

Total Cone Angle	25	35	60	70	90
Slant Length (λ)	0.318	0.290	0.285	0.305	0.335

The upper disk diameter equals 0.7 times the lower cone diameter. The spacing between the top of the cone and the upper disk equals 0.3 times the diameter of the upper cone. The diameter of the upper cone determines the upper frequency limit, but practice shows that the antenna patterns are good only over a 4 : 1 to 4.5 : 1 frequency range. The impedance bandwidth is much wider than the pattern bandwidth. To reduce weight and wind loading, the cone and disk can be made from rods, with a typical implementation having at least eight.

FIGURE 5-32 Discone antenna with coaxial feed with a center conductor connected to the upper disk and a shield connected to the lower cone.

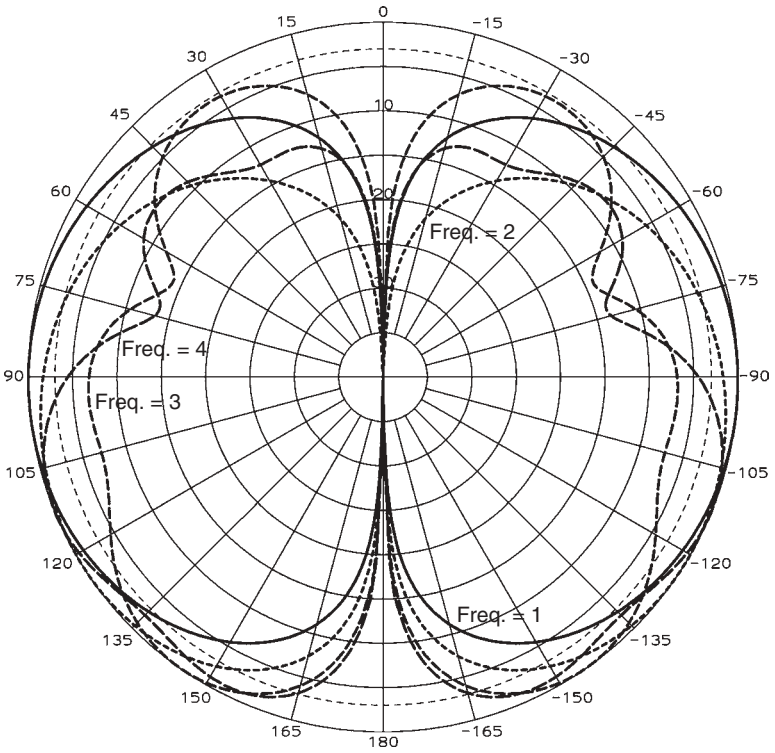


FIGURE 5-33 Elevation pattern of a 60° discone antenna at normalized frequencies = 1, 2, 3, and 4.

5-15 BALUNS [21; 22, pp. 167–180]

A balun properly connects a balanced transmission line to an unbalanced transmission line. Simple arguments about impedances to the balanced and unbalanced modes of the three-wire transmission lines explain its operation. Considering one of the lines of a transmission line as ground misleads us. A ground plane under the transmission-line feeder becomes the third conductor of a three-wire line. Currents flowing in the ground plane can unbalance the currents in the feeder. A balanced three-wire transmission-line mode carries equal and opposite currents in the feeder lines. The capacitances per unit length of the two lines to ground are the same. Coax is an example of an unbalanced line structure (Figure 5-34). The inner conductor has no direct capacitance to ground. The two-wire line shown in Figure 5-34 is a balanced line having equal capacitances to ground, but we must judge a balanced line by the currents, not just the physical structure.

Before we analyze baluns, we must consider the fundamental modes of a three-wire transmission line. Figure 5-35 shows circuit representations of the modes without showing the ground conductor. Equal loads terminate ports 3 and 4. The even mode applies equal voltages on ports 1 and 2 and forms a magnetic wall between the conductors where the magnetic field vanishes to produce a virtual open circuit. The unbalanced mode—equal current directions—is associated with the even mode. Equal and opposite voltages on ports 1 and 2 form the odd mode and set up an electric wall between the conductors. The electric wall is a virtual short circuit. The odd mode excites equal and opposite currents—balanced mode—on the two lines. When the loads on ports 3 and 4 are unequal, the modes separate according to the voltages, even and odd, or the currents, unbalanced and balanced. Dipoles present loads between the lines and not to ground.

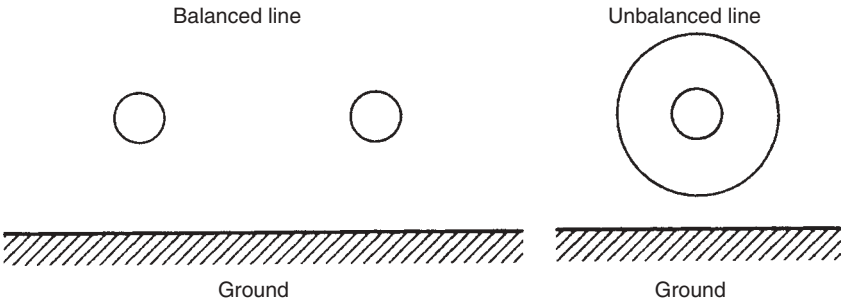


FIGURE 5-34 Physically balanced and unbalanced transmission lines.

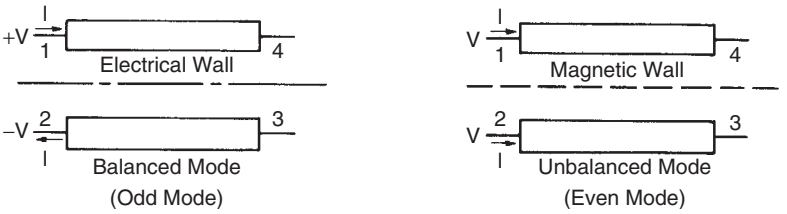


FIGURE 5-35 Balanced and unbalanced modes on a three-wire transmission line.

Unbalanced mode circuits radiate. Only closely spaced equal and opposite currents, the balanced mode, cancel the far-field radiation from the currents on the feed lines. The radiating feeder line adds radiation components to the antenna. These components can radiate unwanted polarizations and redirect the beam peak of the antenna (squint). In reception, the unwanted currents excited on the feeder by passing electromagnetic waves reach the receiver terminals without a balun to block them. We analyze baluns by using either transmitting or receiving antennas, depending on convenience, because reciprocity applies to baluns as well as antennas.

We detect balance problems from pattern squint and cross polarization. An impedance-measuring setup can detect some balance problems. Radiating unbalanced currents cause changes in the impedance. The radiation shows when the impedance changes as fingers are run over the coax line from the equipment. If we feed a dipole from a coax without a balun, the current on the outer conductor splits between the dipole conductor and the outside of the conductor. Patterns and impedance measurements detect this current. Unbalanced currents on the arms of the dipole and feeder currents cause pattern squint, but the cross-polarization radiated is usually a greater concern.

5-15.1 Folded Balun

A folded balun (Figure 5-36) allows the direct connection of a coax line to the dipole. A dummy coax outer conductor is connected to the pole fed from the center conductor. It runs alongside the feeder coax for $\lambda/4$ and connects to ground. The other pole connects directly to the shield of the feeder coax. The outer conductor of the coax and the extra line are two lines in a three-wire line with ground. We analyze the structure by using balanced (odd) and unbalanced (even) modes. Unbalanced-mode excitation

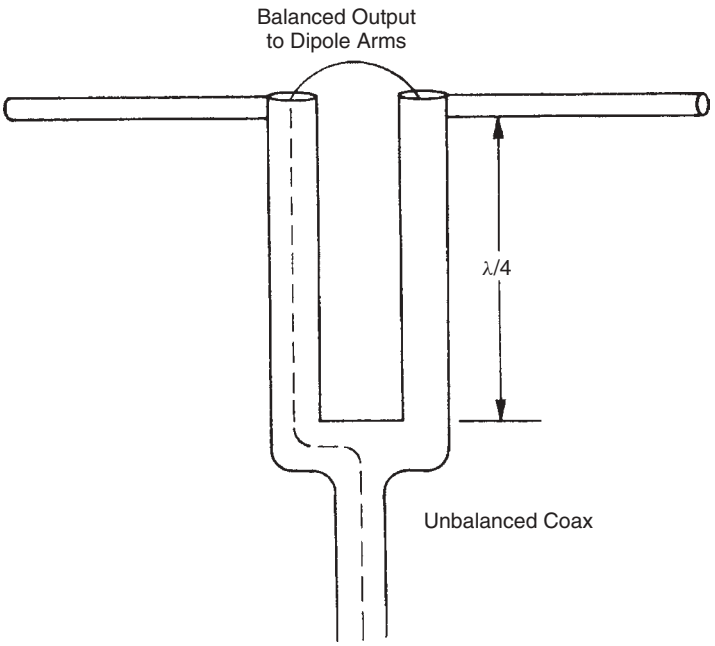


FIGURE 5-36 Folded balun.

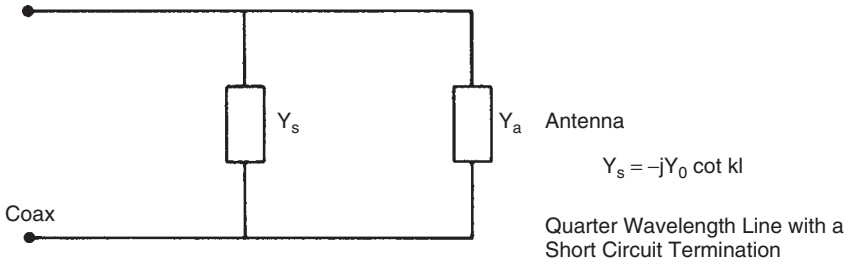


FIGURE 5-37 Folded balun equivalent circuit (balanced mode).

at the dipole forms a magnetic wall through the ground connection between the two coax shields. The circuit reduces to a single line with an open circuit at the ground connection. The open circuit transforms through the quarter-wavelength line to a short circuit at the dipole. Any unbalanced currents induced on the dipole or the coax outer conductor are shorted at the input. Balanced-mode excitation at the dipole forms an electric wall through the ground connection. The balanced-mode circuit of the two coax shields is a λ/4 short-circuited stub connected in shunt with the dipole (Figure 5-37). We analyze the frequency response from Figure 5-37. The bandwidth of the balun, although narrow, exceeds the bandwidth of the dipole. The Roberts balun [23] design adds an open-circuited stub λ/4 long inside the dummy coax of the folded balun. Instead of connecting the center conductor of feeding coax to the outer shield, we connect it to the open-circuited stub. The equivalent circuit for the balanced mode includes the short-circuited stub of the folded balun plus the open-circuited stub. The two reactances shift in opposite directions as frequency changes and produce a dual resonance we see as a loop on the Smith chart plot of impedance. The frequency bandwidth increases to almost 3 : 1, a more suitable choice for wide-bandwidth antennas.

5-15.2 Sleeve or Bazooka Baluns

An outer jacket shields the outer conductor of the coax feeder in a sleeve balun (Figure 5-38). The sleeve and outer conductor of the coax form a series stub between the coax feeder and ground when the cup is short circuited to the coax outer conductor. The λ/4 stub presents a high impedance to the unbalanced currents at the top of the cup (Figure 5-39). A second sleeve below the first one and directed away from the dipole further prevents currents excited on the coax from reaching the input. When the frequency shifts, the connection to ground through the sleeve unbalances the transmission line. This balun is inherently narrowband.

Adding a stub to the center conductor (Figure 5-40) increases the bandwidth because the stubs track each other when the frequency changes. Figure 5-39 demonstrates the circuit diagrams of the two types of sleeve baluns. The type II sleeve balun has matching series stubs on the outputs. The lines remain balanced at all frequencies, but the stubs limit the bandwidth of efficient operation. Marchand [21] adds an open-circuited λ/4 stub inside the matching type II extra shorted stub of the sleeve balun and connects it to the coax center conductor in the same manner as the Roberts balun. The Roberts balun is a folded balun version of the Marchand compensated sleeve balun.

The coaxial dipole is a variation of the sleeve or bazooka balun. We rotate the right pole in Figure 5-38 until it is vertical and remove the left pole. We turn over the

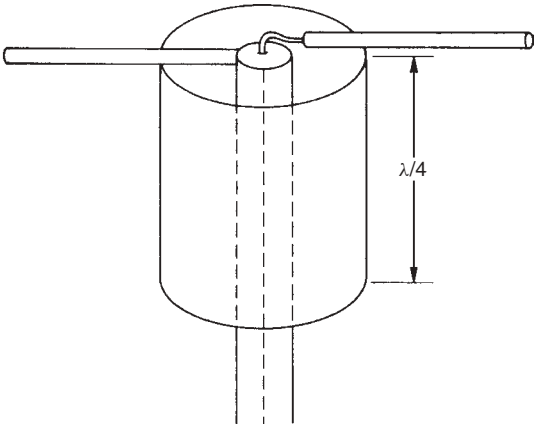


FIGURE 5-38 Sleeve or bazooka balun.

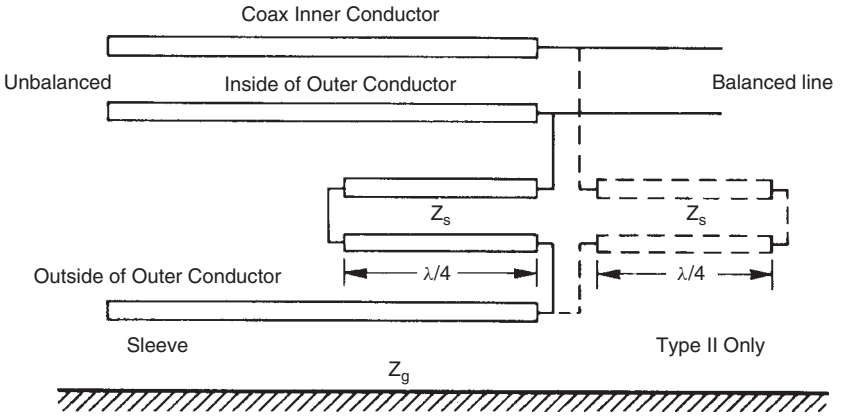


FIGURE 5-39 Schematic of types I and II sleeve or bazooka baluns.

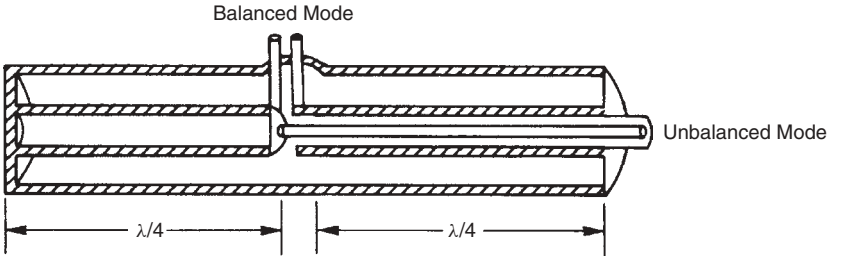


FIGURE 5-40 Type II sleeve balun.

sleeve and connect the short-circuit end to the outer conductor of the coax. The sleeve becomes the second pole of the dipole. The short-circuited stub at the bottom of the dipole between the outer conductor of the coax and sleeve transforms to an open-circuit impedance at the end of the lower pole. This prevents current flow farther down the coax. Some references call this a sleeve dipole, which should not be confused with the

sleeve dipole used to increase the impedance bandwidth. The coaxial dipole has the inherently narrow bandwidth of the bazooka balun, but is a convenient construction.

5-15.3 Split Coax Balun [24, p. 245]

A split coax balun allows the connection of both arms of a dipole to the outer shield of the coax that maintains symmetry to the dipole arms. Its rigidity helps to overcome vibration problems. Slots cut in the outer shield (Figure 5-41) enable the coax line to support two modes and make it equivalent to a three-wire line. A shorting pin excites the TE_{11} mode in the slotted coax (Figure 5-42) to feed the dipole in the balanced mode.

Analysis of a split coax balun is similar to that of a folded balun. The ends of the slots are equivalent to the ground connection of the two coax shields of the folded balun. A virtual open circuit forms at the ends of the slots in the unbalanced (even) mode. It transforms to a short circuit at the dipole and shorts the unbalanced mode at the input. The virtual short circuit at the end of the slots in the balanced mode transforms to an open circuit at the input. Figure 5-37 gives its circuit diagram.

Symmetry improves the performance of a split coax balun over a folded balun. The shorting pin is used only to excite the TE_{11} mode to feed the dipole arms. The extra wire length of the center conductor jumper of the folded balun introduces phase shift to the second arm and squints the beam. For that reason, the split coax balun is a better high-frequency balun. The phase shift problem of the jumper also occurs with the "infinite" balun of the log-periodic antenna.

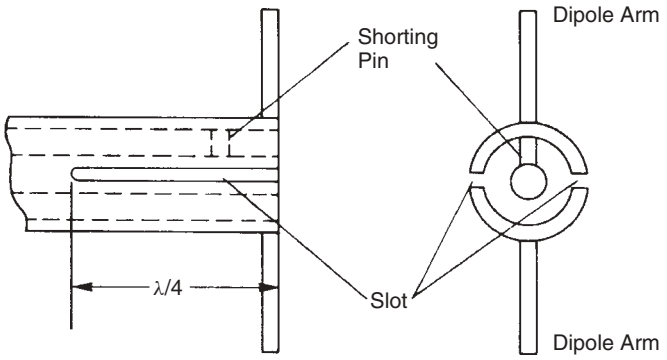


FIGURE 5-41 Split coax balun. (From [24], Fig. 8-5, © 1948 McGraw-Hill.)

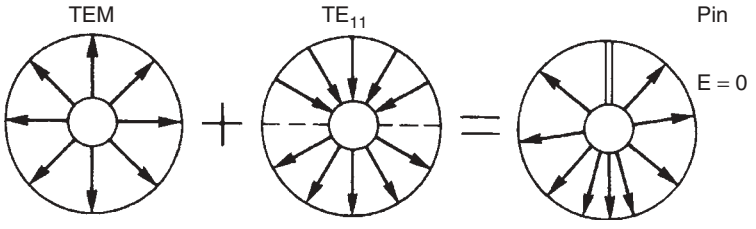


FIGURE 5-42 Coaxial transmission-line modes in a split coax balun. (From [24], Fig. 8-6, © 1948 McGraw-Hill.)

5-15.4 Half-Wavelength Balun

A half-wavelength balun (Figure 5-43) works by cancellation of the unbalanced-mode currents at the input to the coax. The impedance transforms by a factor of 4 from unbalanced- to balanced-mode ports. In the unbalanced (even) mode, equal voltages are applied to the two output ports. When the voltage wave on the upper line propagates through $\lambda/2$, its phase changes by 180°. This signal cancels the signal connected directly to the coax center conductor.

A load across a balanced-mode transmission line has a virtual short circuit halfway through it. The load on each balanced-mode line is $2Z_0$, where Z_0 is the coax characteristic impedance. The load on the end of the $\lambda/2$-long line is transformed by the transmission line to the identical impedance when it circles the entire Smith chart. The two loads, each $2Z_0$, are connected in shunt at the coax input and combine to Z_0. A balanced-mode impedance of $4Z_0$ transforms to Z_0 at the coax input. The $\lambda/2$-long cable can be rolled up for low frequencies. The balun transforms 300-Ω input impedances of folded dipoles to 75 Ω by using RG-59 cable (75 Ω).

5-15.5 Candelabra Balun

A candelabra balun (Figure 5-44) transforms the unbalanced-mode impedance fourfold to the balanced-mode port. The coax cables on the balanced-mode side connect in series, whereas those on the unbalanced-mode side connect in parallel. We can divide the balanced-mode impedance in two and connect each half to a $2Z_0$ impedance transmission line. These lines then connect in shunt at the unbalanced-mode port. The unbalanced-mode currents short out at the input to the $2Z_0$ coax lines in the same manner as does the folded balun. More lines can be stacked in series and higher-impedance transformations obtained, but construction becomes more difficult.

5-15.6 Ferrite Core Baluns

Ferrite cores can be used to increase the load impedance to unbalanced-mode currents and reduce them. At low frequencies (<100 MHz) ferrite has high permeability. As the frequency increases, the permeability drops, but the losses to internal magnetic fields increase. The increased inductance of transmission lines is used at low frequencies, and the increased loss is used at high frequencies to inhibit currents.

Ferrite bazooka or sleeve balun ferrite cores, placed on the outside of a coax line (Figure 5-45), increase the impedance to ground for currents on the outside of the

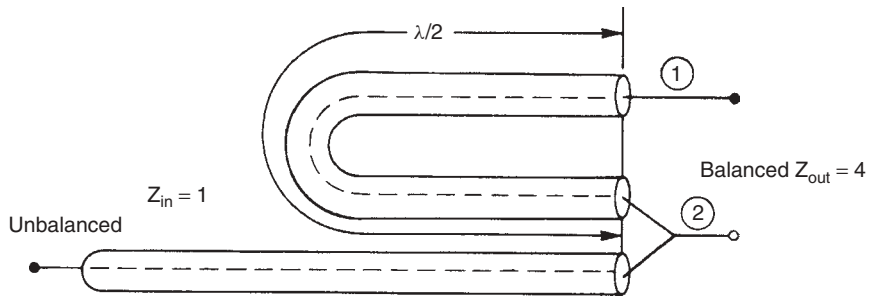


FIGURE 5-43 Half-wavelength balun.

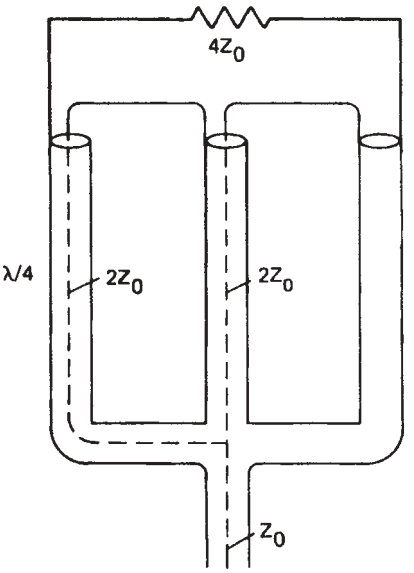


FIGURE 5-44 Candelabra balun.

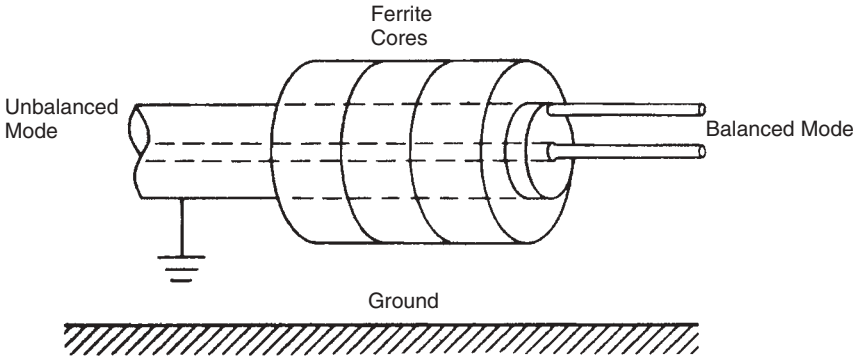


FIGURE 5-45 Ferrite core bazooka balun.

shield and inhibit unbalanced currents between the ground and the outer shield. The ferrite balun can work over many decades. The ferrite material provides high impedance through inductance at low frequencies. As the ferrite material impedance drops when frequency increases, the transmission line approaches λ/4 of the bazooka balun. Any ferrite core balun is a compromise design between low-frequency response controlled by the amount of ferrite and high-frequency response controlled by the length of transmission line.

Bifilar wire windings on a ferrite core form the 1 : 1 balun (Figure 5-46*a*). The bifilar wire closely approximates 50- to 100-Ω characteristic impedance transmission lines. This balun can work satisfactorily from 100 kHz to 1 GHz. When there are balanced currents in the windings, there is no net magnetic field in the ferrite. In the unbalanced mode, the fields add in the core and give a high series impedance due to the high inductance (low frequency) or high resistance (high frequency). The amplitude at port

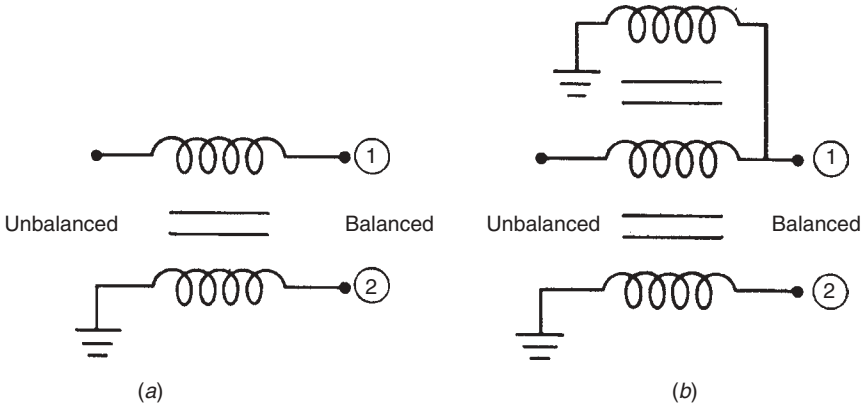


FIGURE 5-46 Bifilar wire ferrite core baluns: (*a*) type I; (*b*) type II.

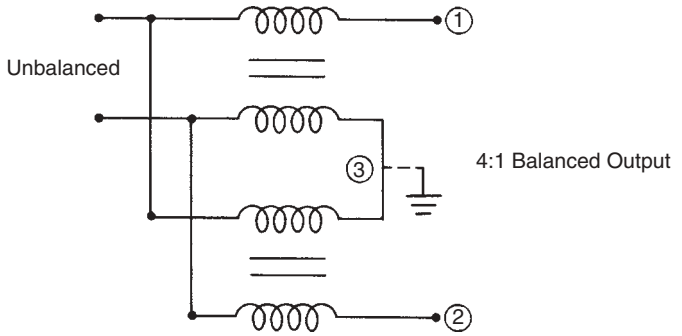


FIGURE 5-47 Ferrite core candelabra balun.

2 (Figure 5-46*a*) is lower than at port 1 because of the extra losses introduced in the ground lead by the ferrite core. This can be corrected by adding an extra winding to ground (Figure 5-46*b*), which produces a type II bazooka balun. The extra winding balances the outputs by adding loss to port 1 without any increase in bandwidth due to the ferrite loading.

5-15.7 Ferrite Candelabra Balun

We can make a parallel-to-series ferrite balun that transforms the unbalanced-mode input impedance fourfold to the balanced-mode output (Figure 5-47). As in the coax version, the characteristic impedance between the wires in the cores should be twice the unbalanced-mode input impedance. Point 3 (Figure 5-47) is a virtual short. Connecting it to ground sometimes helps the balance. Both windings can be wound on the same core, such as a binocular core.

5-15.8 Transformer Balun

A transformer balun has no transmission-line equivalent; it is merely a transformer (Figure 5-48). The balanced-mode output impedance is fourfold that of the input

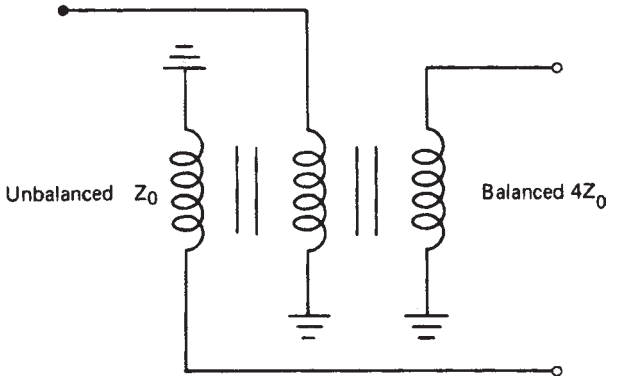


FIGURE 5-48 Ferrite transformer balun.

unbalanced mode. Wound with trifilar wire, the output has twice the number of turns compared to the input, and each output line has the same impedance to ground through the transformer. No transmission-line effects extend the upper band edge beyond the transformer action of the windings and ferrite, but the transformer balun is a fine low-frequency device. Use of No. 36 and 38 wires in ferrite core baluns limits the power-handling capability of the baluns to receive-only levels. The ferrite bazooka balun (Figure 5-45) carries the power in coax, which allows higher power levels.

5-15.9 Split Tapered Coax Balun [25]

A split tapered coax balun starts with inherently unbalanced coax. Moving toward the balanced end, an outer conductor slot opens and exposes more and more of the center conductor (Figure 5-49). At the point where the size of the outer conductor is reduced to that of the inner conductor, we connect a balanced twin line to the two conductors. The impedance must be raised from input to output, since the two-wire line, spaced the radius of the coax, has higher impedance than the one at the coax input. The balance depends on reducing the reflected wave in the transformer. Any suitable tapered transformer, such as Dolph–Chebyshev or exponential, can be used, and design return loss is the level of the unbalanced mode. The balun can be constructed in microstrip. The ground plane tapers until it and the upper conductor are the same size. This balun can operate over decades of bandwidth, since the tapered transformer determines the bandwidth.

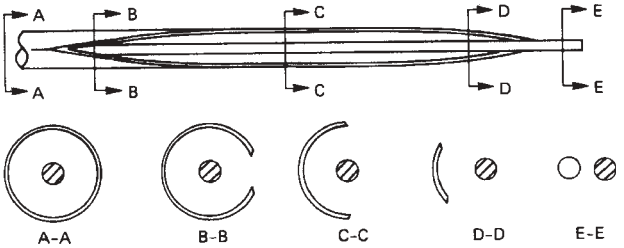


FIGURE 5-49 Split tapered coax balun. (From [16], Fig. 1, © 1960 IEEE.)

5-15.10 Natural Balun [26, p. 821]

A natural balun feeds the coax through a loop antenna to the feed point where the outer shield is split and the center conductor jumps the gap to connect to the outer shield of the coax. At this point the currents flow on to the outer shield and radiate. By moving an equal distance along the coax until the two halves meet, we can connect the feed coax and not have current flow down the outside. The currents flow in opposite directions along the loop and cancel at the connection. From a circuit point of view, the connection point is a virtual short circuit to the balanced mode similar to a folded balun at its connection point. In a similar manner, on a folded dipole we can connect the feed coax to the middle of the shorted dipole and form a natural balun.

We have not exhausted the number of balun designs. The "infinite" balun of log-periodic antennas will be discussed as part of the construction of such antennas. When a broad-beam antenna is designed, sometimes a little squint in the beam and a little cross-polarization are acceptable and the antenna may be fed without a balun.

5-16 SMALL LOOP

In Section 2-1.2 we discuss the radiation from a small constant-current loop. For a small loop the current is approximately constant and has the pattern of a short magnetic dipole located along the axis of the loop. Similar to the short dipole, we calculate the series resistance and radiation resistance to calculate efficiency. We increase efficiency by adding closely coupled turns and ferrite rods to increase the magnetic field. A multiturn loop with N turns and loaded with ferrite with an effective permeability μ_{eff} and area A has a radiation resistance:

$$R_{\text{loop}} = 320N^2\mu_{\text{eff}}\pi^4\frac{A^2}{\lambda^4}$$

The wire adds a series loss resistance R_L to the input resistance of the multiturn loop, but it is proportional to N instead of N^2 and the wire surface resistance R_s given the wire conductivity σ:

$$R_L = \frac{(\text{loop length})NR_s}{\text{perimeter of wire cross section}}$$

$$R_s = \sqrt{\frac{\omega\mu_0}{2\sigma}}$$

When we have a circular loop radius b and a wire diameter of $2a$, we find a series inductance from the loop:

$$L_{\text{loop}} = \mu_0\mu_{\text{eff}}N^2 b\ln\frac{b}{a}$$

$$R_L = \frac{b}{a}R_s$$

The radiation efficiency of the loop is calculated by using the series-loss resistance and the radiation resistance:

$$\eta_e = \frac{P_r}{P_{\text{in}}} = \frac{R_{\text{loop}}}{R_{\text{loop}} + R_L}$$

Adding turns and ferrite material increases the radiation efficiency.

We compute the mean effective permeability by integrating along the ferrite rod the permeability distribution $\mu_c(x)$ of the core and dividing by its length [27, p. 6-20]. The following approximates its distribution for a core of length l:

$$\mu_c(x) = \mu_{cs}(1 + 0.106\overline{x} - 0.988\overline{x}^2)$$
$$\overline{x} = \frac{2|x|}{l}$$

The factor μ_{cs} depends on the ferrite geometry. For a cylindrical rod with diameter D, we calculate μ_{cs} from

$$\mu_{cs} = \frac{\mu}{1 + (\mu - 1)(D/l)^2(\ln(l/D)\{0.5 + 0.7[1 - \exp(-\mu \times 10^{-3})]\} - 1)}$$

For a rectangular cross section of height h and width w ($w \geq h$), we find μ_{cs} from

$$\mu_{cs} = \frac{\mu}{1 + (\mu - 1)(4wh/\pi l^2)\{\ln[\beta l/(w + h)] - 1\}}$$
$$\beta = 4 - 0.732\left[1 - \exp\left(-5.5\frac{w}{h}\right)\right] - 1.23\exp(-\mu \times 10^{-3})$$

When the axis of the loop is along the z-axis, the effective height $\mathbf{h}$ is determined by the area:

$$\mathbf{h} = -j\overline{\mu}_{\text{eff}}kA\sin\theta\hat{\boldsymbol{\theta}}$$

The ferrite loop antenna finds use as a receiving antenna at low frequencies where the sky noise is very high, and the added noise of the antenna due to poor efficiency has little effect on the overall G/T value.

5-17 ALFORD LOOP [28]

An Alford loop feeds two dipoles curved into a loop that radiates an omnidirectional pattern with horizontal polarization when located horizontally over a ground plane. Figure 5-50 shows a configuration fed from coax where it feeds two parallel-plate transmission lines connected in shunt. The flat dipoles and the sides of the transmission line are offset by a central substrate (not shown). The opposite direction of the dipoles produces a 180° phase shift between them. The odd-mode feed produces a pattern null along the coax axis that reduces current excitation on the outside of the coax and eliminates the need for a balun. We space the loops so that the circumference is approximately 1λ and adjust the parallel-plate line impedance to transform the dipole impedance to 100 Ω where the two sides are connected in shunt.

The horizontal pattern improves when we place the Alford loop in a slotted cylinder [29]. Figure 5-51 illustrates the positioning of the loop in the slotted cylinder when viewed from below before it is attached to a ground plane. The 0.38λ-diameter cylinder has four slots each 0.5λ long with open-circuited ends. We cut an opening about 0.2λ along the circumference about 0.12λ along the cylinder axis to create an open circuit for the slot where the cylinder attaches to the ground plane. Of course,

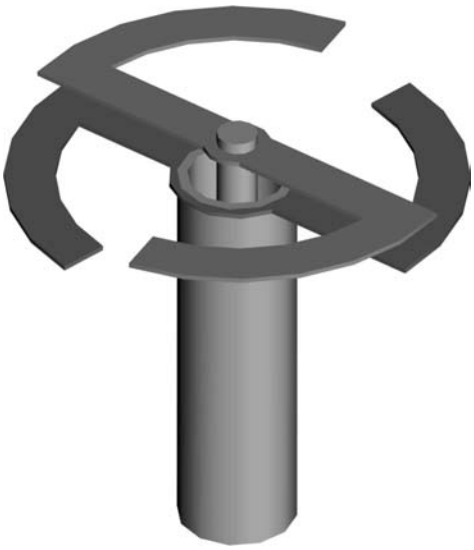


FIGURE 5-50 Alford loop fed by coax into a parallel-plate transmission line.

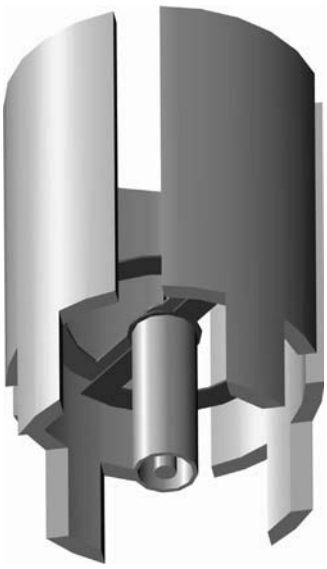


FIGURE 5-51 Alford loop feeding slotted cylinder viewed from below where connected to ground plane.

the coax runs through the ground plate to its connector. We point the parallel-plate transmission line halfway between two slots so that each curved dipole feeds two slots. Figure 5-52 gives the typical pattern of the antenna on a ground plane. This antenna illustrates another example of the interaction of dipoles and slots used to improve patterns.

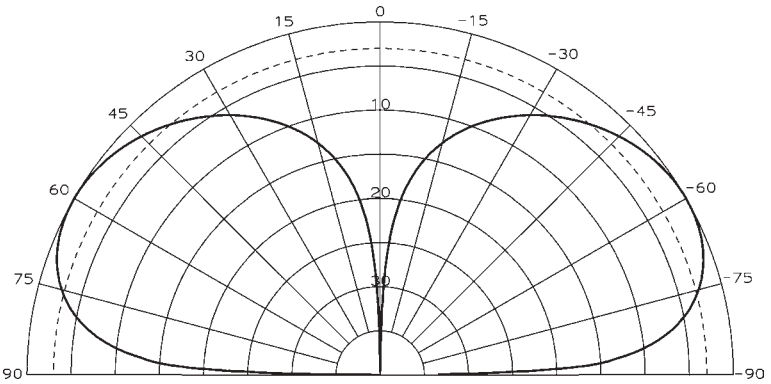


FIGURE 5-52 Horizontally polarized pattern of Alford loop feeding slotted cylinder.

5-18 RESONANT LOOP [19, p. 158; 30]

A folded dipole has a perimeter of one wavelength, and any other shaped loop will resonate when its perimeter is near one wavelength. The sinusoidal current distribution of the folded dipole remains on the loop. The folded dipole has a dipole pattern with its null in the direction of the voltage across the feed. The loop retains this pattern property when it is symmetrical about the feed point. The loop shape that opens in the H-plane distorts the normal perfect-circle H-plane pattern of the dipole and gives 3- to 4-dB peaks in the directions normal to the loop compared with the directions in the plane of the loop. The E-plane null is filled in with a cross-polarized pattern at about 20 dB below the beam peak.

Since the current distribution is sinusoidal on the loop, opposite the feed and halfway around the loop perimeter is a virtual short-circuit point. The current reaches maximums at the virtual short circuit and at the feed. We can expect a moderate input resistance because the standing-wave current is high. A circular loop has an input resistance of about 130 Ω when the loop is 1.08 wavelengths in perimeter. If the loop is a parallelogram, the resonant input resistance depends on the angle between the wires at the feed. The resistance starts at about 300 Ω for the folded dipole and decreases for decreasing angles. At 120° between the lines, the resistance is about 250 Ω and drops to 50 Ω when the angle is 60°. Four common loop shapes are: (1) circle, (2) square (quad), (3) parallelogram, and (4) triangle. Changes in the shape affect the input resistance at resonance and, to a slight degree, the resonant perimeter. The Q of the antenna is about the same as that of a half-wavelength dipole. The gain equals that of the one-wavelength dipole, 3.8 dB.

Figure 5-53 illustrates a resonant loop used as a ground plane for a dipole spaced at $\lambda/4$ above it. Each square is $\lambda/8$ on a side with a center square with mounting holes for the balun. The ground plane consists of two resonant loops, because the center ring is eight squares around, while the outer ring has 16 squares for a 2λ loop. For this antenna each pole consists of two rods: one horizontal and one tilted at 30°. Two versions of this antenna landed on Mars in 1976 [31]. Not only is the ground plane extremely lightweight, but it gives an excellent F/B ratio for a ground plane only 1λ across. We can easily add a resonant ring to an antenna—whether a dipole or another antenna—and expect an improvement.

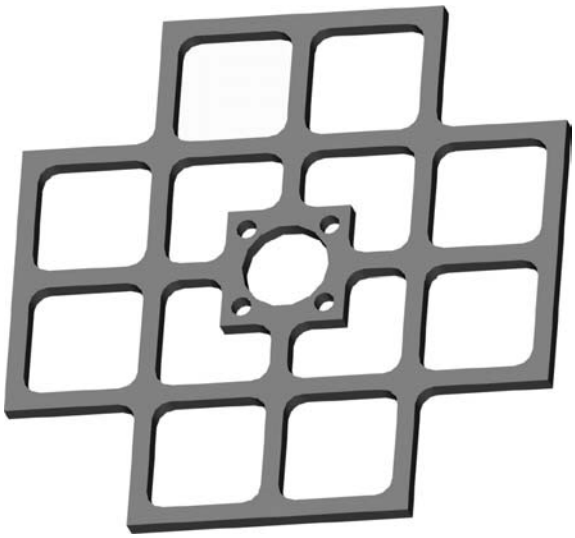


FIGURE 5-53 Combined 1λ and 2λ resonant loops ground plane for crossed dipoles using λ/8-side squares.

5-19 QUADRIFILAR HELIX [32, 33]

A quadrifilar helix consists of two interwound resonant loops twisted into a helical shape. Although the antenna can be made using two open U-shaped wires, the usual implementation has two loops. The loop quadrifilar helix is resonant when each loop has a perimeter slightly greater than one wavelength, similar to a planar resonant loop. The feed produces a standing-wave current distribution that peaks at the feed and at the center of the shorting wire that joins the two ends of the helix. Nulls in the current distribution occur halfway along the helical section. The lower half of the antenna can be removed at the null points to form a dual open U-shaped antenna. An antenna using a half-turn in each of the four arms of the helix with a diameter of 0.174λ and a height of 0.243λ produces a circularly polarized pattern with a 120° beamwidth when fed from two equal-amplitude feeds phased in quadrature.

If we consider a single twisted loop oriented with the z-axis along the helix axis, we discover the unique radiation characteristics by doing a MOM analysis of the wire loop. The analysis shows equal and opposite currents located on the feed and shorting line that reduces the radiation from the closely spaced straight sections. The currents along the helical section have a progressive traveling-wave phasing except for the 180° phase shift through the null. This traveling-wave current radiates circular polarization.

Consider a loop twisted into a right-hand helix. The loop radiates a pattern with lobes along the $+z$- and $-z$-axes, both with left-hand circular polarization. If we rotate the helix end to end, the helix remains right-hand and the problem has not changed. Whether we feed the antenna at the top or bottom in the center of the straight wire, the current distribution on the antenna is the same and the pattern has the same polarization. When we feed both loops with phasing for left-hand circular polarization (x-axis 0° and y-axis 90°) on the right-hand helix, the two left-hand circularly polarized lobes from the two loops add along the z-axis, while the left-hand lobes along the $-z$-axis

cancel because the feeding phases are for RHC along the $-z$-axis. Of course, we use a left-hand helix and right-hand feeding phases for RHC polarization.

Figure 5-54 shows a left-hand half-turn quadrifilar helix that radiates RHC. Normal practice uses a dual folded balun to feed the antenna, with the two ports connected to a hybrid coupler to produce circular polarization. The balun short circuit is a disk located $\lambda/4$ from the feed through which all four coax cables pass. We solder the coax lines to the disk to complete the folded balun structure. We add two upper jumper wires at the feed points. We can use the balun structure to support the two helixes as shown in Figure 5-54. A second construction uses self-supporting helixes fed at the lower end by a dual balun. Out of band this antenna also suffers from poor efficiency even though the antenna is well matched at the hybrid port, because the equal reflections from the two loops are routed to the load on the hybrid coupler. Although mutual coupling between the two twisted loops modifies the input impedance, we discover that each individual loop has nearly the same impedance as the total antenna.

We can feed a quadrifilar helix from a turnstile feed if we use unequal perimeter twisted loops to produce a phase difference caused by the impedance changes when the loop is longer or shorter than a resonant length. A development similar to Section 5-6 gives us the dimensional changes of the loops required to produce a circularly polarized pattern. For a half-turn helix the perimeter length determines the resonant frequency:

$$\text{perimeter} = \sqrt{(2 \cdot \text{height})^2 + (\pi \cdot \text{diameter})^2} + 2 \cdot \text{diameter} \tag{5-18}$$

By Section 5-6 the ratio of the two perimeters that will produce a circularly polarized pattern is related to antenna bandwidth (Q). If we keep the same diameter (D) for both helixes, we modify the heights (H):

$$\frac{\sqrt{(2H_1)^2 + (\pi D)^2} + 2D}{\sqrt{(2H_2)^2 + (\pi D)^2} + 2D} = 1 + \frac{1}{Q} \tag{5-19}$$

FIGURE 5-54 Half-turn quadrifilar helix fed from two folded baluns.

This allows the feed to be a split-tube balun when the two loops are joined to the same coax balun located in the center. We split Eq. (5-19) in the same manner as Eq. (5-11) to compute the two new heights:

$$\begin{aligned}\sqrt{(2H_x)^2+(\pi D)^2}+2D &= \frac{\sqrt{(2H_0)^2+(\pi D)^2}+2D}{\sqrt{1+1/Q}} \qquad \text{for RHC} \\ \sqrt{(2H_y)^2+(\pi D)^2}+2D &= \left[\sqrt{(2H_0)^2+(\pi D)^2}+2D\right]\sqrt{1+1/Q}\end{aligned} \tag{5-20}$$

A half-turn quadrifilar helix with diameter = 0.174λ and height = 0.243λ has a 3.2% 2 : 1 VSWR bandwidth. We determine antenna Q by using Eq. (5-10) to be equal to 22.1. We use Eq. (5-18) to calculate the perimeter of the unmodified helix to be 1.079λ. When we substitute the Q and the perimeter into Eq. (5-20), we can easily solve for the two heights: $H_x = 0.2248\lambda$ and $H_y = 0.2608\lambda$. This assumes that the antenna has a left-hand helix for RHC polarization. When we make this antenna with the correct dimensions, the measured Smith chart of the design will have a small loop at the frequency with the least cross-polarization. The two shorted loops do not meet at the point opposite the feed but pass over and under each other. We can design a turnstile quadrifilar helix with the same height for the two loops by using Eq. (5-20) to compute the two diameters.

5-20 CAVITY-BACKED SLOTS

A slot that radiates only on one side of the ground plane is the dual of a monopole. As in the case of the monopole, restricting the radiation to above the ground plane doubles the gain. The voltage across the slot determines the field strength. Since the radiated power is only half that of the slot radiating on both sides and having the same peak fields, the input impedance doubles. The already high slot impedance becomes even higher. The cavity must present an open circuit at the slot, or its susceptance must combine with the slot susceptance to resonate. Normally, it is a quarter-wavelength deep. Since many cavities form a box, the waveguide mode determines the propagation constant (wavelength) used to determine the depth.

5-21 STRIPLINE SERIES SLOTS

Stripline consists of a center strip equally spaced between two flat ground planes. It supports a coaxial-type TEM-mode wave between the central strip and the two ground planes. The ground plane currents match the currents flowing in the central strip. A waveguide has axial currents that flow along the axis and transverse currents that flow in the direction of the sidewalls. Any slots cut in the ground plane can only interrupt axial currents and present series loads to the transmission line because the TEM wave has no transverse currents. The load that a slot presents to the transmission line is a parallel combination of a radiation conductance and an energy storage susceptance. Low values of inductive reactance shunt power around the high resistance of short slots. The inductance increases with increasing electrical length and supports higher voltages across the slot radiation resistance. This increases the radiated power. The

inductance increases to an antiresonance near λ/2, where further increases in electrical length decrease the capacitive reactance.

Oliner [34] gives an expression for slot conductance normalized to the stripline characteristic impedance. When unnormalized, it becomes

$$G = \frac{8\sqrt{\varepsilon_r}}{45\pi^2}\left(\frac{a'}{\lambda}\right)^2\left[1 - 0.374\left(\frac{a'}{\lambda}\right)^2 + 0.130\left(\frac{a'}{\lambda}\right)^4\right] \tag{5-21}$$

where a' is the length of the slot and ε_r is the dielectric constant of the stripline boards. More complete expressions for full series admittance are available [35], but near resonance Eq. (5-21) suffices.

Most striplines are etched on dielectric substrates. The dielectric fills the slot and reduces the resonant length. The effective dielectric constant in the slot is [36]

$$\varepsilon_r' = \frac{2\varepsilon_r}{1 + \varepsilon_r} \tag{5-22}$$

Slot length determines the radiation conductance. Decreasing the resonant length increases radiation resistance at resonance.

Example A set of woven Teflon fiberglass ($\varepsilon_r = 2.55$) dielectric boards supports a resonant-length slot in a stripline circuit. Compute resonant length and center-fed radiation conductance for a slot that resonates when $a' = 0.48\lambda$ in air.

The effective dielectric constant from Eq. (5-22) is 1.44. The effective dielectric constant reduces the resonant length:

$$\frac{a'}{\lambda} = \frac{0.48}{\sqrt{1.44}} = 0.40$$

Equation (5-21) finds the resonant conductance as 3.27 mS or 306 Ω resistance. The high impedance requires an offset feed to match the slot to a stripline. We locate the offset feed from the slot center to reduce the input impedance:

$$\xi = \frac{a'}{2} - \frac{\lambda}{2\pi\sqrt{\varepsilon_r'}}\sin^{-1}\sqrt{\frac{Z_{\text{in}}}{Z_c}} \tag{5-23}$$

We can determine the 50- and 100-Ω feed points of the slot:

$$50\,\Omega: \quad \xi = \frac{0.40\lambda}{2} - \frac{\lambda}{2\pi\sqrt{1.44}}\sin^{-1}\sqrt{\frac{50}{306}} = 0.145\lambda$$

$$100\,\Omega: \quad \xi = \frac{0.40\lambda}{2} - \frac{\lambda}{2\pi\sqrt{1.44}}\sin^{-1}\sqrt{\frac{100}{306}} = 0.119\lambda$$

At 2 GHz, the dimensions of the slot and the feed locations become:

$$a' = 6\,\text{cm}: \quad \xi = 2.17\,\text{cm}\ (50\,\Omega) \qquad 0.83\,\text{cm from edge}$$

$$\xi = 1.79\,\text{cm}\ (100\,\Omega) \qquad 1.21\,\text{cm from edge}$$

Increasing the slot width decreases the impedance below that of the thin-width slot result given above and will require experimental work to find the exact feed point. Measuring the center-fed impedance allows the use of Eq. (5-23) to compute the approximate offset feed point.

Figure 5-55 illustrates a typical stripline-fed slot. Shorting the center strip to ground at the location of the slot creates a current maximum at the slot for feeding. An open circuit a quarter wavelength beyond the slot creates the same standing-wave current maximum at the slot. Convenience determines the feeding method. The slot interrupts currents only in the top ground plane. Unequal current flow on the two ground planes unbalances the stripline and excites a parallel-plate mode between the ground planes. Waveguide wall slots also excite higher-order modes, but these cannot propagate because they are below their cutoff frequencies. The parallel-plate mode is another TEM mode with no low-frequency cutoff. Power in this mode propagates away from the slot and couples into the other slots in an undesirable manner or radiates from edges.

Shorting pins between the two ground planes contain the parallel-plate mode around the slot. By placing the rows of pins parallel with the axis of the slot and a quarter wavelength away from the slot, the rows of pins reflect an open-circuit impedance at the slot. The side rows of pins complete the box and convert the parallel-plate mode into a waveguide TE_{10} mode. The box formed by the rows of shorting pins and the two ground planes form a resonant cavity in shunt with the slot admittance. The resonant cavity places a standing-wave current null on a centered slot and does not excite it. Only the normal stripline currents feed the slot.

From an impedance point of view, the cavity is a second parallel resonant circuit that increases the stored energy of the antenna. Q increases and bandwidth decreases. Because only a portion of the available power would be converted to the parallel-plate mode by the slot discontinuity, we analyze the cavity as a circuit coupled through a transformer to the input. The transformer increases the impedance of the resonant cavity at the input and controls the division of power between the slot and

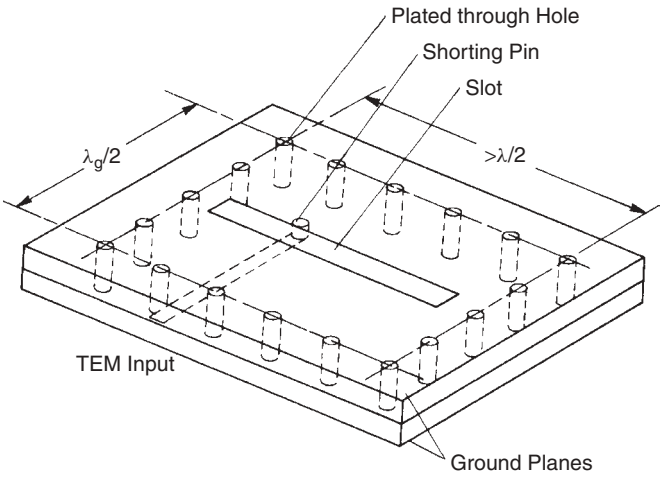


FIGURE 5-55 Stripline series slot.

the cavity. The cavity reactance slope limits the bandwidth of the stripline-fed slot to a few percent. Increasing the impedance of the waveguide cavity transmission line reduces the reactance slope contributed by the cavity. We increase the bandwidth by using greater distances between the ground planes and thereby increase the waveguide transmission-line impedance. In general, greater volumes for an antenna increase the impedance bandwidth.

Rotating the slot relative to the stripline feeding line reduces its load on the transmission line. The waveguide top wall series slot relation [Eq. (5-34)] applies in this case. The slot maintains its polarization while the nonradiating stripline center conductor approaches the slot at an angle. Rotated slots in waveguide must be paired symmetrically to reduce cross-polarization. A longitudinal array [37] can be made by placing all the slots on the centerline of a boxed stripline. Either edge plating or a series of plated-through holes forms a waveguide structure that supports only the TE_{10} mode. Slots placed on the centerline (as in Figure 5-59, slot c) fail to interrupt the waveguide mode currents. The stripline meanders below and varies the excitation by changing the angle between the slot and the stripline center conductor. The slight loading of each slot excites very little of the parallel-plate mode that causes unwanted slot coupling. Both traveling-wave and resonant linear arrays are possible. See Section 5-26 for a discussion of slot arrays.

5-22 SHALLOW-CAVITY CROSSED-SLOT ANTENNA

We can feed the slot in Figure 5-55 by exciting the cavity in an odd mode from two points on opposite sides of the slot. To be able to excite both polarizations, we divide the slot in two and rotate the two parts in opposite directions by 45° to form a cross. We use a square cavity to maintain symmetry and replace the shorting pins with solid walls (Figure 5-56*c*). Since we feed across the diagonal between the crossed slots, we excite both slots. The sum of the fields radiated from the two slots is polarized in the direction of the diagonal. We increase the radiation conduction by lengthening the crossed slots to the maximum, which lowers the Q (increased bandwidth). The cavity compensates for the slot susceptance to obtain resonance. A crossed-slot antenna was built [38] with the following dimensions:

Cavity edge	0.65λ
Cavity depth	0.08λ
Slot length	0.915λ

The measured 2 : 1 VSWR bandwidth was 20.8%. The bandwidth exceeded that of a microstrip patch of the same thickness by about $\sqrt{2}$. Lindberg [39] found that the resonant length of the slot depends on the cavity depth and requires some experimental adjustment.

King and Wong [41] added ridges (Figure 5-56*b*) to increase the bandwidth. Antennas with ridges need a larger cavity width and a longer slot than the unridged design. The ridges can be stepped as shown to increase the bandwidth. Adding ridges gives us extra parameters to adjust for best input match performance. The following design with uniform ridges produces a 58.7% 2.5 : 1 VSWR bandwidth with a double resonance curve.

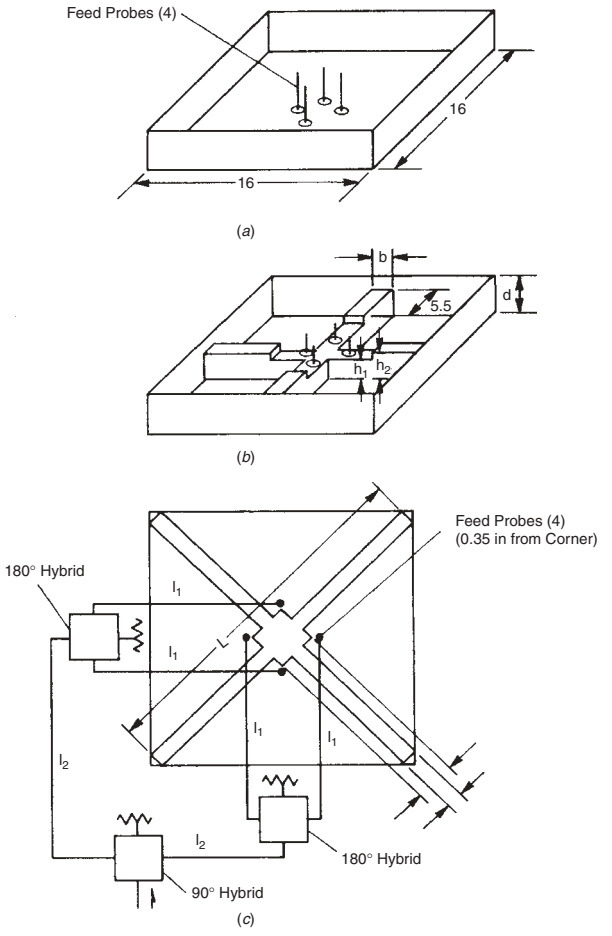


FIGURE 5-56 Shallow-cavity crossed-slot antenna: (*a*) cavity with ridge; (*b*) cavity with ridge; (*c*) typical slot configuration. All dimensions are in inches. (From [38], Fig. 2, © 1975 IEEE.)

Cavity edge	0.924λ	Slot width, W_2	0.058λ
Slot length	1.3λ	Ridge height	0.076λ
Ridge width	0.087λ	Feed width, W_1	0.144λ
Cavity thickness	0.115λ		

Both the ridge and slot shapes can be varied to improve the performance. As fed in Figure 5-56*c*, the antenna radiates circular polarization on a boresight. Near the horizon (90° from the boresight), the polarization reduces to linear as we enter the null of one of the slots.

5-23 WAVEGUIDE-FED SLOTS [24, p. 291; 40, p. 95]

Waveguide is an ideal transmission line for feeding slots. Although its impedance cannot be defined uniquely, all possible candidates—voltage and current, power and

current, or power and voltage—yield high values that match the high values of impedance of half-wavelength slots. Waveguide provides a rigid structure with shielded fields. The slots couple to the internal fields and allow the easy construction of linear arrays fed from traveling waves or standing waves in the waveguide. By controlling the position of the slots in the walls, the amplitude of the slot excitation can be controlled.

The waveguide fields excite a slot when the slot interrupts the waveguide wall currents. When excited, the slot loads the waveguide transmission line. We make the following assumptions about the wall slots.

1. The slot width is narrow. When a slot grows in width, we must either consider it to be an aperture in the wall or assume that it is excited by interrupting currents in two coordinate directions.
2. The slot is a resonant length and its length is near $\lambda/2$. The waveguide environment, the wall thickness, and the position in the wall all affect the resonant length. In most cases, experiments must determine the resonant length.
3. The electric field is directed across the narrow width of the slot and varies sinusoidally along its length and is independent of the excitation fields. This reiterates assumptions 1 and 2. An aperture radiates the polarization of the incident fields, but resonant-length slots can be excited only with a sinusoidal voltage standing wave. The slot direction determines polarization.
4. The waveguide walls are perfectly conducting and infinitely thin. Even though the walls have thickness, the difference has a small effect on the general form of the slot excitation formulas. As in the case of the resonant length, experiments determine a few values from which the rest must be interpolated, or the values provide the constants for more elaborate models.

5-24 RECTANGULAR-WAVEGUIDE WALL SLOTS

The lowest-order mode (TE_{10}) in a rectangular waveguide has the following fields [41, p. 69]:

$$\begin{aligned} E_y &= E_0 \sin(k_c x) e^{-jk_g z} \\ H_x &= -\frac{k_g E_0}{\omega\mu} \sin(k_c x) e^{-jk_g z} \\ H_z &= -\frac{k_c E_0}{j\omega\mu} \cos(k_c x) e^{-jk_g z} \end{aligned} \tag{5-24}$$

where $k_c = \pi/a$, $k_g^2 = k_c^2 - k^2$, and a is the guide width with cutoff wavelength $\lambda_c = 2a$. We can separate TE_{10}-mode rectangular waveguide fields into two plane waves that propagate at an angle to the axis and reflect from the two narrow walls. We denote as ξ the angle of the waves measured from the centerline of the waveguide or with respect to the wall. We relate the waveguide propagation to this angle:

$$\xi = \sin^{-1}(\lambda/\lambda_c) \tag{5-25}$$

At high frequencies, $\xi \to 0$ and the waves travel straight through the guide as though the walls are not there. As the wavelength approaches cutoff, $\xi \to 90^\circ$ and the waves

reflect back and forth between the sidewalls instead of propagating down the guide. This angle factors into the expressions for slot loading to the waveguide transmission line and can be related to propagation:

$$\text{guide wavelength, } \lambda_g = \frac{\lambda}{\cos\xi} = \frac{\lambda}{\sqrt{1-(\lambda/\lambda_c)^2}}$$

$$\text{phase velocity } V_{\text{ph}} = \frac{c}{\cos\xi} \quad \text{and} \quad \text{group velocity} = c\cos\xi$$

$$\text{relative propagation constant } P = \frac{\lambda}{\lambda_g} = \cos\xi$$

For analysis we divide the fields bouncing down the waveguide into z-directed fields of the axial wave moving down the guide and x-directed fields of the transverse wave, a standing wave between the two narrow walls. A standing wave causes a 90° separation of the voltage and currents in a transmission line as shown in Figure 5-1. The phase of the currents excited in the waveguide walls due to the fields will be 90° relative to the electric field.

The wall currents $\mathbf{J}_s$ are determined by $\mathbf{J} = \mathbf{n} \times \mathbf{H}$, where $\mathbf{n}$ is the unit normal to the wall. When we apply this boundary condition to the walls, we obtain the following wall currents:

Sidewalls:

$$J_y = -j\frac{E_0 k_c}{\omega\mu}e^{-jk_g z}$$

Bottom wall ($y = 0$):

$$J_s = \frac{E_0}{\omega\mu}e^{-jk_g z}[k_g \sin(k_c x)\hat{\mathbf{z}} + jk_c \cos(k_c x)\hat{\mathbf{x}}] \tag{5-26}$$

Top wall ($y = b$):

$$J_s = \frac{-E_0}{\omega\mu}e^{-jk_g z}[k_g \sin(k_c x)\hat{\mathbf{z}} + jk_c \cos(k_c x)\hat{\mathbf{x}}]$$

Equation (5-26) shows that transverse wave currents are 90° out of phase with respect to the electric field E_0. The current alternates between the two types of current as the wave propagates down the waveguide. In the case of a standing wave along the z-axis caused by a short circuit, the axial wave currents are 90° out of phase with the electric field across the waveguide (Figure 5-1). The peak amplitude of the transverse wave currents occurs at the same point as the electric field in a standing wave along the z-axis, since both are 90° out of phase with the axial wave currents. The sidewalls J_y have only transverse wave currents. The top and bottom broad walls have both x-directed transverse wave and z-directed axial wave currents. Figure 5-57*a* shows the direction and amplitude distribution of these transverse waves. Slots interrupting these

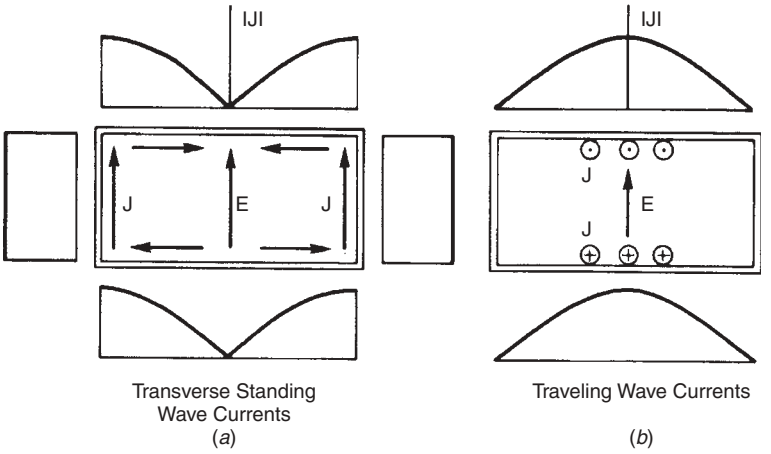


FIGURE 5-57 TE_{10}-mode rectangular waveguide wall currents: (*a*) transverse wave currents; (*b*) axial wave currents.

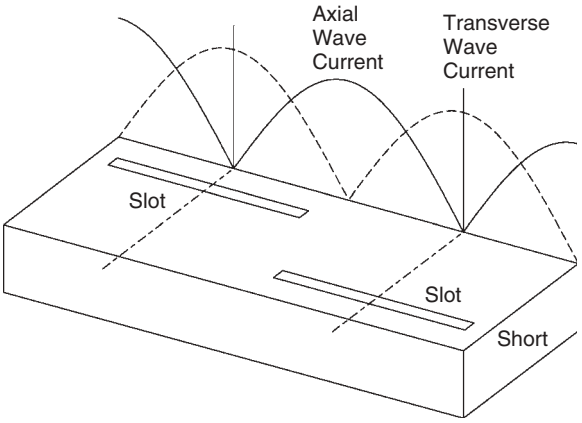


FIGURE 5-58 Short-circuited waveguide axial and transverse wave currents and the location of longitudinal wall slots.

currents are shunt loads to the waveguide. In an axial wave along the z-axis, these transverse waves propagate in the z-axis direction.

Equation (5-26) shows that the transverse wave currents are 90° phase with respect to the axial wave currents. Figure 5-58 shows the two types of currents along the z-axis when the guide has a short circuit at its end. When measuring slots that interrupt transverse wave currents, we need to place the waveguide short at $\lambda_g/4$ or $3\lambda_g/4$ away from the slot. This locates the peak of the transverse wave currents flowing around the waveguide walls at the slot shown in Figure 5-58 because the axial wave currents are at a minimum. The second consideration is the shunt load on the waveguide. The $\lambda_g/4$ section of waveguide transforms a short circuit on the end of the waveguide (to the axial wave currents) to an open circuit at the slot. From a voltage point of view the susceptance of the shorted stub is at a minimum. We place the short circuit at

$\lambda_g/2$ from the last slot for a series loading slot that interrupts the axial wave currents. This locates the current maximum at the slot and causes maximum interaction with the waveguide fields. Figure 5-58 illustrates the placement of the next slot $\lambda_g/2$ down the guide at the next current maximum. Figure 5-57*a* indicates the transverse wave current flow and we see that the currents flow toward the centerline, producing currents 180° out of phase on the two sides of the centerline. The two slots in Figure 5-58 are excited by oppositely directed currents that add 180° phase shift between the slots. This phase shift compensates for internal standing-wave current phasing of 180° due to the $\lambda_g/2$ spacing.

Longitudinal top and bottom wall slots cut x-directed transverse shunt currents. The central slot c, located at a current null, fails to be excited. We use this nonradiating slot to insert a traveling probe to measure VSWR. When moved off center, slots d and e cut x-directed currents and are excited. The shunt conductance has the relation

$$g = g_1 \sin^2 \frac{\pi x'}{a} \tag{5-27}$$

where x' is the distance from the guide centerline. Shunt currents on either side of the centerline of the top or bottom wall (Figure 5-57*a*) have different directions. Besides any traveling-wave phase, slots d and e (Figure 5-59) are 180° out of phase. Top-wall longitudinal slots generate no cross-polarization, since all maintain the same orientation. We relate the peak conductance g_1 to the direction of the waves in the guide [41]:

$$g_1 = 2.09 \frac{a}{b} \frac{\cos^2[(\pi/2)\cos\xi]}{\cos\xi} \tag{5-28}$$

Equation (5-28) indicates that the conductance increases for a given spacing off the centerline as the frequency approaches cutoff and $\xi \to \pi/2$.

We cannot use Eqs. (5-27) and (5-28) for design because they do not include the wall thickness and we need to determine the exact length for resonance. The resonant length depends on the spacing from the centerline. Fortunately, the coupling between longitudinal slots is small enough that measurements can be made on single slots. Elliott suggests a measurement plan for longitudinal slots [3]. We build a series of slotted waveguides each containing a single slot at different distances from the centerline. Seven cases are sufficient to generate a curve for design. We need

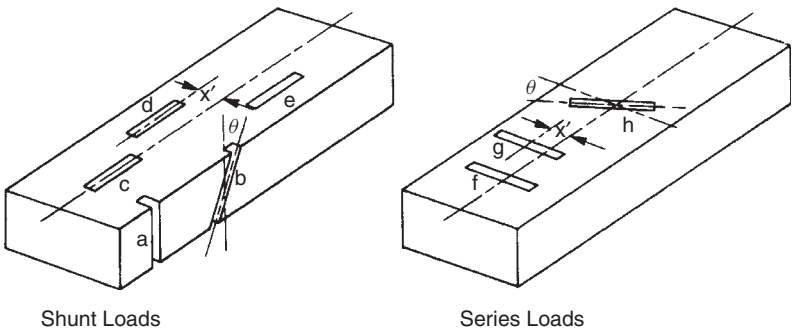


FIGURE 5-59 TE_{10}-mode rectangular waveguide wall slots.

to locate a sliding short circuit farther down the waveguide and adjust it until the standing-wave current peaks at the slot to produce maximum radiation and conductance. With a network analyzer we measure the conductance normalized with respect to the waveguide impedance. Initially, we machine the slots too short, measure the results, and then machine longer slots using the same guides and remeasure until they pass through resonance.

Since the manufacturing cost of test slots is high and they require careful measurements, analytical methods of determining slot parameters become attractive. FEM programs can model the details of the slot, the waveguide, and the wall thickness. A number of runs similar to those of the measurements allows design curves to be created.

Sidewall slots (Figure 5-59) interrupt shunt transverse waves. Slot a fails to cut surface currents and is not excited. By tilting slot b, currents are cut. The sidewall slot conductance is given for $\theta < 30^\circ$ by

$$g = g_0 \sin^2 \theta \tag{5-29}$$

where g_0 is the peak conductance. Note that the sidewall slots must cut into the top and bottom walls to achieve a resonant length. The peak conductance can be related to the direction of the waves in the waveguide [Eq. (5-25)] [1, p. 82]:

$$g_0 = 2.09 \frac{a \sin^4 \xi}{b \cos \xi} \tag{5-30}$$

Equation (5-30) shows the relationship of the slot load conductance versus the frequency. As frequency increases, ξ decreases and the conductance falls off as the fourth power of the sine of the angle. The complete theory of Stevenson gives the conductance for an arbitrary tilt [42]:

$$g = 2.09 \frac{a \sin^4 \xi}{b \cos \xi} \left[\frac{\sin \theta \cos[(\pi/2) \cos \xi \sin \theta]}{1 - \cos^2 \xi \sin^2 \theta} \right]^2 \tag{5-31}$$

Tilting the slots to interrupt currents introduces cross-polarization components in the array pattern. We alternate the direction of tilt to reduce cross-polarization. Two things prevent the total cancellation of cross-polarization. First, the amplitude taper of the array changes the amplitude from element to element and the fields do not cancel. Alternating the tilt of the slots symmetrically about the centerline in an array with an even number of elements prevents cross-polarization on the boresight. Off the boresight, the array effect of the spaced elements introduces a cross-polarization pattern, since cross-polarization is not canceled at each element.

Although Eqs. (5-29) and (5-30) give the slot conductance, they cannot be used for design. They assume an infinitely thin wall and ignore the high level of radiation along the waveguide wall. These slots readily couple to neighboring slots. The effective conductance needs to include the mutual conductance. For these slots we build a series of slotted waveguides containing a group of slots all tilted to the same angle and cut so that they are a resonant length. This means that we will first need to build the slots about 5% shorter than resonance length, make measurements, and then machine the slots longer and repeat the measurements to find the resonant length. We space the slots at the same distance as will be used in the final design and either place a

short-circuit beyond the last slot to produce a maximum current at all slots or load the waveguide to form a nonresonant array. We measure the load of the group of slots on the waveguide transmission line using a network analyzer and divide the conductance by the number of slots to get an incremental conductance. This conductance is larger than the one measured on a single slot. We fit the group of measurements to a curve that replaces Eq. (5-29) for design.

Axial z-directed waves (Figure 5-57*b*) peak in the center of the broad walls and taper to zero at the edges. They remain zero on the sidewalls. When centered, transverse slots f and g (Figure 5-59) interrupt the maximum current. When moved off center, g, their series loading to the waveguide drops:

$$R = R_0 \cos^2 \frac{\pi x'}{a} \tag{5-32}$$

The maximum resistance is related to the direction of the waves in the waveguide:

$$R_0 = 2.09 \frac{a}{b} \frac{\sin^2 \xi}{\cos^3 \xi} \cos^2 \left(\frac{\pi}{2} \sin \xi \right) \tag{5-33}$$

An evaluation of Eq. (5-33) shows that the resistance increases as frequency approaches cutoff for a given location of the slot, a result similar to that for other slot configurations. The mutual coupling between these series slots is high. We perform incremental resistance experiments similar to the procedure used for sidewall slots to discover the true values of resistance versus offset.

Rotating the broadwall transverse slot, h, reduces the z-axis directed current interrupted. When the slot is centered, equal and opposite shunt currents are cut by the slot and the slot fails to present a load to shunt currents:

$$R = R_0 \cos^2 \theta \tag{5-34}$$

We can excite slots a and c by probe coupling into the waveguide. A probe placed next to the slot and extending into the guide feeds the slot. The longer the probe, the more it disturbs the waveguide fields to excite the slot. Probes placed on opposite sides of the slots induce fields 180° out of phase with respect to each other.

5-25 CIRCULAR-WAVEGUIDE SLOTS

Figure 5-60 shows the transverse wave and axial wave currents of the circular waveguide TE_{11} dominant mode. Slots may be placed successfully only at the current maximums without affecting the polarization of the internal wave. A longitudinal slot placed halfway between the current maximums, 45°, interrupts only shunt transverse waves. Since any polarization is possible in the circular waveguide, analytically we divide the incident wave into two waves. One is polarized in the direction of the slot; the other is polarized perpendicular to the slot axis. The wave polarized perpendicular to the slot location has its current maximum at the slot and it removes power from the wave. The other wave produces a current null on the slot. When we combine the two fields after the slot, the unloaded wave is larger and the combined wave rotates its polarization toward the slot. Circumferential slots interrupting axial wave currents

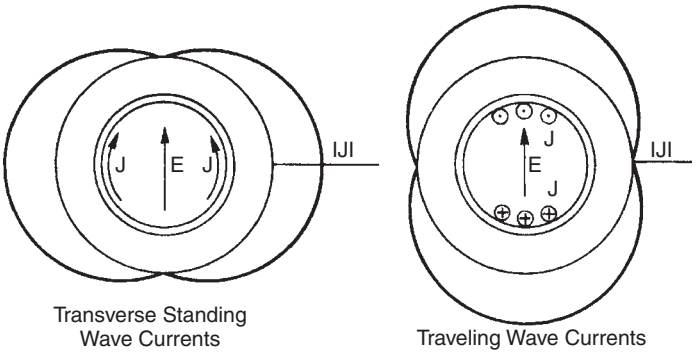


FIGURE 5-60 TE_{11}-mode circular waveguide wall currents.

also cause polarization rotation of the wave when not centered 90° from the electric field direction.

Slots placed at the maximum of transverse currents cut them when rotated about the axis of the waveguide. Like rectangular-waveguide sidewall slots, the slots oriented perpendicular to the guide axis, circumferential, do not load the waveguide. Rotating the slot increases the shunt load on the waveguide. Slots placed at the maximum of the axial wave cut z-directed currents. Field probes can monitor the internal fields of the waveguide through a longitudinal slot without causing radiation from the slot. When the slot is rotated away from the axis direction, it interrupts series axial wave currents, loads the waveguide, and radiates.

Coaxial TEM-mode transmission line and TM_{01}-mode circular waveguide have the same outer wall currents (Figure 5-61). Slots can be excited and load the waveguide only by interrupting these axial wave currents. In Figure 5-61, slot *a* fails to cut currents and is not excited. VSWR measuring probes use this slot. Slots *b* and *c* interrupt the currents and series-load the guide. Slot *c*, whose total length is resonant, is excited by the small portion in the center cutting z-directed currents. We can probe feed slot *a*, but the probe shunt loads the waveguide or TEM coax that would be series loads on the waveguide if they directly interrupted the axial wave currents.

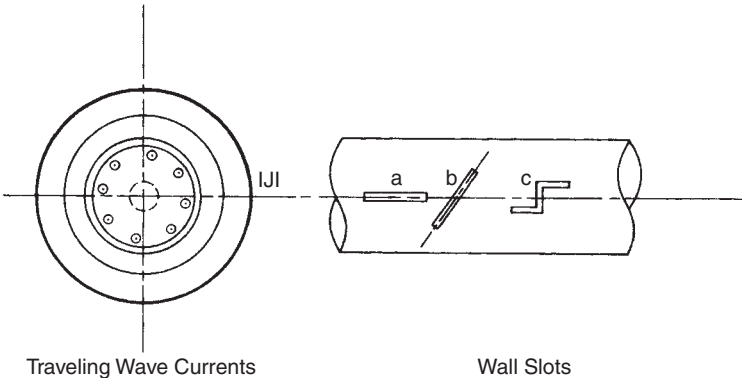


FIGURE 5-61 Coax or TM_{01}-mode circular waveguide wall currents and slots.

5-26 WAVEGUIDE SLOT ARRAYS [4, p. 402]

Waveguide slot arrays can produce low sidelobe antennas for pencil beams with good aperture efficiency. Array fabrication requires close manufacturing tolerances to achieve the desired amplitude distribution because random errors in manufacture produce unwanted sidelobes and raise the general sidelobe level. Producing these arrays is an art requiring careful analysis of all slot interactions, slot dimensioning determined from models and measurements, and precision machining and assembly.

An array consists of a set of waveguides loaded with slots and joined with a corporate feed into the total array. The corporate feed can also be a slotted array feeding the individual waveguides that contain the radiating slots. Aperture size and distribution determine the beamwidth and sidelobes in the various planes. We divide arrays of slots into two groups: nonresonant, excited by traveling waves, and resonant, excited by standing waves. Waves either travel along the guide into a terminating load or reflect from a short and set up standing waves along the z-axis (Figure 5-58). Traveling-wave currents excite the slots as they pass, and slots may be placed anywhere relative to the load. The distance between slots and the propagation constant determine the relative phases. Standing waves set up a fixed sinusoidal current pattern along the waveguide axis at a given frequency. The standing-wave phase is either 0° or 180°. Slots placed in the current nulls of standing waves interrupt no currents and fail to be excited by the waveguide. We can vary the amplitude by the z-axis placement of the slots. The termination determines the array type. Do not confuse transverse waves that produce shunt currents and z-axis standing waves caused by a short-circuit termination. Both traveling and standing waves on the z-axis have shunt currents.

Standing waves (resonant array) produce beams normal to the array axis. A resonant array maintains its beam direction when frequency changes, but the standing-wave pattern shifts and changes the excitation of the slots (Figure 5-62). The amplitudes of the slots farthest from the short circuit change the most, since the standing waves have shifted farther. The length of the resonant array determines its bandwidth. The pattern shape changes because distribution and input impedance change as the loads change when the standing-wave currents shift.

Nonresonant array (traveling-wave) beam directions are functions of the propagation constant of the wave exciting the slots. Changing the frequency shifts the beam direction. If the load on the end reflects a wave, another beam forms from the reflected traveling wave. The second beam appears at the same angle to the axis of the waveguide

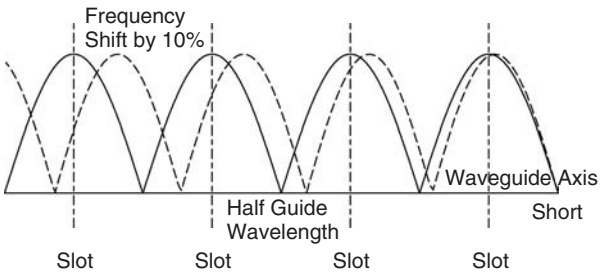


FIGURE 5-62 Standing-wave currents in resonant array relative to slots and after 10% frequency shift.

as the first but measured from the $-z$-axis. The first-pass radiated power and return loss of the load determine the level of this second beam relative to the first.

Both resonant and nonresonant waveguide slot arrays use resonant-length slots. We space the slots $\lambda_g/2$ apart in the resonant array, as shown in Figure 5-58. We place the slots at alternating positions about the centerline of the broadwall or at alternating tilt angles in the sidewall to give the additional 180° phase shift to produce a broadside beam. The admittances of the slots of the resonant array add at the input because the $\lambda_g/2$ spacing produces a complete rotation around the Smith chart. In nonresonant arrays a traveling wave is used to excite the slots. We space the slots at other than $\lambda_g/2$ distances and terminate the waveguide with a load. We assume a matched system throughout the antenna in a first-order analysis suitable for most designs. The beam of most nonresonant slot arrays is designed to backfire at an angle to broadside.

5-26.1 Nonresonant Array [43]

In a nonresonant waveguide, slot array resonant-length slots are used in a traveling-wave antenna terminated at the end with a load. The antenna radiates at an angle to the normal of the waveguide face determined by wave velocity and slot spacing. We vary the slot loading along the waveguide so that each slot radiates the proper amount of the remaining power. A termination absorbs the power remaining after the last slot. With a mismatched termination the reflected power radiates a second lower-amplitude beam as the wave travels to the source.

We design with either shunt- or series-loading slots. A shunt slot radiates the power $|V|^2 g_i/2$, where g_i is the normalized slot conductance. Similarly, a series slot radiates the power $|I|^2 r_i/2$, where r_i is the normalized slot resistance. We normalize the conductance or resistance to a per unit length function: $g(z)$ or $r(z)$. The attenuation equation (4-78) becomes

$$\frac{1}{P(z)}\frac{dP}{dz} = -g(z) \quad \text{or} \quad -r(z) \tag{5-35}$$

Equation (5-35) modifies the normalized attenuation equation (4-79) [24, p. 291]:

$$g(z)L = \frac{|A(z)|^2}{[1/(1-R)]\displaystyle\int_0^L |A(z)|^2\,dz - \int_0^z |A(z)|^2\,dz} \tag{5-36}$$

where the aperture runs $\pm L/2$ and R is the ratio of the input power absorbed by the termination. $A(z)$ is the normalized aperture distribution on the interval $\pm\frac{1}{2}$. We change to $r(z)L$ in Eq. (5-36) for series-loading slots.

Equation (5-36) assumes light loading by the slots so that the waveguide transmission line is matched at all points. This approximation improves as the length increases. Equation (5-36) is the same as Eq. (4-79) except for a constant. We divide the values in Table 4-28 or Figure 4-26 by 4.34 to calculate normalized conductance (resistance) of shunt (series) slots times the array length. Each slot provides the loading over the spacing between slots:

$$g_i = \int_{-d/2}^{d/2} g(z)\,dz \simeq g(z_i)\,d$$

where d is the spacing of the slot at z_i.

We space the slots at other than $\lambda_g/2$. At $\lambda_g/2$ spacings, all reflections from the mismatches (slots) add in phase at the input. The small mismatches from each slot add with various phase angles for element spacing different from $\lambda_g/2$ and cancel each other to some extent to give a good input match over a reasonable bandwidth. When we increase the array length, we can no longer ignore the waveguide losses. The slot conductances become very small and radiate power on the same order as the losses. We modify Eq. (5-36) to include the losses as in Eq. (4-79), and the slot conductance increases to compensate for the ohmic losses in the walls. A small slot conductance is difficult to achieve with longitudinal broadwall slots because one edge of the slot must be over the centerline of the waveguide wall and the results become unpredictable. The achievable conductances limit possible distributions in a slotted waveguide array. Mutual coupling between slots changes the distribution and we must modify the slot offsets to account for mutual coupling using Eq. (3-23).

If we specify the radiating power of each slot in a discrete sequence P_i, we modify Eq. (5-36). The integrals become summations, since

$$|A(z)|^2 = \delta(z - id)P_i$$

where d is the slot spacing, $\delta(x)$ the Dirac delta (impulse) function, and P_i the power coefficient of the ith slot. The power radiated is

$$\sum_{i=1}^{N} P_i = P_{\text{in}}(1 - R) = \int_0^L |A(z)|^2 \, dz$$

The integral $\int_0^z |A(z)|^2 \, dz$ is the power radiated by the preceding slots. Equation (5-36) reduces to

$$g_i = r_i = \frac{P_i}{1 - \sum_{n=1}^{i-1} P_n} \tag{5-37}$$

Dissipating more power in the termination decreases each P_i and the required conductance (resistance) range of the slots.

We alternate the locations of longitudinal slots about the centerline of the broadwall to add 180° phase shift between elements. Similarly, sidewall slot directions are alternated along the array. The additional phase shifts cause backfire of the beam in most cases. The element spacing, as well as the traveling-wave phase velocity, determines the beam direction. The phasing equation in the array factor for beam peak becomes $kd \cos\theta + 2n\pi = Pkd - \pi$, where θ is measured from the array axis, P is the relative propagation constant ($P < 1$), and n is an arbitrary integer. We solve for the beam peak direction and the necessary spacing to get a particular beam direction:

$$\theta = \cos^{-1}\left[P - \frac{(n + \frac{1}{2})\lambda}{d}\right] \tag{5-38}$$

$$\frac{d}{\lambda} = \frac{n + \frac{1}{2}}{P - \cos\theta_{\max}} \tag{5-39}$$

We usually work with $n = 0$ because using $n > 0$ produces multiple beams.

Example Compute slot spacing to produce a beam at $\theta = 135^\circ$ in a waveguide of width 0.65λ. Calculate the relative propagation constant from the general equation for a waveguide.

$$P = \sqrt{1 - \left(\frac{\lambda}{\lambda_c}\right)^2}$$

For $\lambda_c = 2a$,

$$P = \sqrt{1 - \left(\frac{1}{1.3}\right)^2} = 0.640 = \frac{\lambda}{\lambda_g}$$

From Eq. (5-39), using $n = 0$, we determine spacing in free space: $d/\lambda = 0.371$. The waveguide spacing is given by

$$\frac{d}{\lambda_g} = \frac{d}{\lambda} P = 0.371(0.640) = 0.237$$

If we use $n = 1$, then $d/\lambda = 1.11$, which radiates an additional beam at $\theta = 79^\circ$ for $n = 0$ [Eq. (5-38)].

Beams enter visible space at $\cos\theta = -1$ (180°) and move toward end fire ($\theta = 0$) as the spacing increases. We calculate the region of single-beam operation from Eq. (5-39). The minimum d/λ occurs when $\theta = 180^\circ$ for $n = 0$, and the maximum occurs when $\theta = 180^\circ$ for $n = 1$:

$$\frac{0.5}{1 + P} \le \frac{d}{\lambda} \le \frac{1.5}{1 + P} \tag{5-40}$$

We substitute the upper bound into Eq. (5-38) and use $n = 1$ to derive the minimum angle of single-beam operation:

$$\theta_{\min} = \cos^{-1}\left(P - \frac{1 + P}{3}\right) \tag{5-41}$$

Example Determine the minimum scan angle (toward end fire) for $P = 0.6$, 0.7, 0.8, and 0.9 that has a single beam. We substitute these values into Eq. (5-41) to find:

P	0.6	0.7	0.8	0.9
$\theta_{\min}$	86.2°	82.3°	78.5°	74.5°

If we scan to $\theta = 90^\circ$, the spacing becomes $\lambda_g/2$ and the mismatches from each slot add to the input and produce a resonant array. The array with a forward firing beam has a slot spacing greater than $\lambda_g/2$. Given a waveguide with $P = 0.8$, we use Eq. (5-39) to calculate spacing to give beams at 80° and 100°:

$$\frac{d}{\lambda} = \frac{0.5}{0.8 - \cos 80^\circ} = 0.798 \quad \text{and} \quad \frac{d}{\lambda} = \frac{0.5}{0.8 - \cos 100^\circ} = 0.514$$

$$\frac{d}{\lambda_g} = \frac{d}{\lambda} P = 0.639 \quad \text{and} \quad \frac{d}{\lambda_g} = \frac{d}{\lambda} P = 0.411$$

A nonresonant array has a backfire beam that scans toward broadside as frequency (and P) increases. Hansen [44, p. 90] gives the slope of the beam shift with frequency change:

$$f\frac{d\sin\theta}{df} = \frac{1}{P} - \sin\theta \tag{5-42}$$

where f is the frequency.

5-26.2 Resonant Array

We space the slots at $\lambda_g/2$ and terminate the waveguide end with a short circuit either $\lambda_g/4$ or $3\lambda_g/4$ from the last one for shunt loading slots in a resonant array. The beam radiates broadside to the array. The 2 : 1 VSWR bandwidth of the array is approximately $50\%/N$ for N elements in the array. The antenna is narrowband. The admittances of all elements add at the input. To have a matched input, $\sum_{i=1}^{N} g_i = 1$, where g_i is the normalized slot conductance. If we define P_i as the normalized power radiated by the ith slot, then

$$g_i = P_i \qquad \text{where} \quad \sum_{i=1}^{N} P_i = 1$$

5-26.3 Improved Design Methods

The methods given above ignore the interaction of slots and their effect on the transmission line. We can describe the array as a loaded transmission line and consider the interactions of the slots by accounting for the transmission-line mismatches [45, pp. 9–11]. We ignore the mutual coupling for longitudinal broadwall slots because it is small, but sidewall slots have high mutual coupling and require an adjustment of the effective slot impedance. We use an incremental admittance, found from the measured change in admittance, when one slot is added to the array or total conductance of the array divided by the number. This accounts somewhat for the mutual coupling.

Elliott and Kurtz [46] relate the self-admittance of a longitudinal broad-wall slot, measured or calculated, to the mutual admittance of the array of slots found from equivalent dipoles. They use Babinet's principle and the mutual impedance of equivalent dipoles. The method requires solution of a set of $2N$ equations in the location and length of the slots to give the desired excitation while accounting for mutual coupling. Their formulation ignores slot interaction in the waveguide beyond the first-order mode. Elliott [47] extends this method to the analysis and design of nonresonant arrays. Of course, when we design a planar array, the slots between waveguide sticks couple readily and we need to account for the mutual coupling between them. The voltage excitation needs to be adjusted to account for this coupling or the desired distribution will not be achieved.

Dielectric loaded waveguide arrays require additional analysis because the approximation of a piecewise sinusoidal distribution, such as dipole current, fails to model the slot distribution adequately. Elliott [48] uses a slot distribution

$$E(x) = \cos\frac{\pi x}{2b}$$

where b is the length. Mutual impedances between dipoles that have the wrong distribution are not used; instead, the active admittances are found from forward and

back scattering between the slots directly. The method still requires the solution of $2N$ equations for the slot lengths and locations.

REFERENCES

1. R. F. Harrington, *Time-Harmonic Electromagnetic Fields*, McGraw-HiIl, New York, 1961.
2. C. A. Balanis, *Antenna Theory, Analysis and Design*, 2nd ed., Wiley, New York, 1997.
3. H. G. Booker, Slot aerials and their relation to complementary wire aerials, *Proceedings of IEE*, vol. 92, pt. IIIA, 1946, pp. 620–626.
4. R. S. Elliott, *Antenna Theory and Design*, Prentice-Hall, Englewood Cliffs, NJ, 1981.
5. V. H. Rumsey, *Frequency Independent Antennas*, Academic Press, New York, 1966.
6. W. H. Watson, *Wave Guide Transmission and Antenna Systems*, Oxford University Press, London, 1947.
7. R. W. Masters, Super-turnstile antenna, *Broadcast News*, vol. 42, January 1946.
8. J. D. Kraus, *Antennas*, McGraw-Hill, New York, 1950.
9. H. M. Elkamchouchi, Cylindrical and three-dimensional corner reflector antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-31, no. 3, May 1983, pp. 451–455.
10. S. Maci et al., Diffraction at artificially soft and hard surfaces by using incremental diffraction coefficients, *IEEE AP-S Symposium*, 1994, pp. 1464–1467.
11. E. L. Bock, J. A. Nelson, and A. Dome, *Very High Frequency Techniques*, McGraw-Hill, New York, 1947. Chapter 5.
12. A. J. Poggio and P. E. Mayes, Pattern bandwidth optimization of the sleeve monopole antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-14, no. 5, September 1966, pp. 623–645.
13. W. L. Stutzman and G. A. Thiele, *Antenna Theory and Design*, Wiley, New York, 1981.
14. H. E. King and J. L. Wong, An experimental study of a balun-fed open-sleeve dipole in front of a metallic reflector, *IEEE Transactions on Antennas and Propagation*, vol. AP-19, no. 2, March 1972, pp. 201–204.
15. J. L. Wong and H. E. King, A cavity-backed dipole antenna with wide bandwidth characteristics, *IEEE Transactions on Antennas and Propagation*, vol. AP-21, no. 5, September 1973, pp. 725–727.
16. A. Kumar and H. D. Hristov, *Microwave Cavity Antennas*, Artech House, Boston, 1989.
17. Y. Mushiake, An exact step-up impedance ratio chart of folded antenna, *IRE Transactions on Antennas and Propagation*, vol. AP-2, 1954, p. 163.
18. R. C. Hansen, Folded and T-match dipole transformation ratio, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 1, January 1982, pp. 161–162.
19. *The ARRL Antenna Book*, American Radio Relay League, Inc., Newington, CT, 1974.
20. R. A. Burberry, *VHF and UHF Antennas*, Peter Peregrinus, London, 1992.
21. N. Marchand, Transmission-line conversion transformers, *Electronics*, December 1941, pp. 142–145.
22. W. L. Weeks, *Antenna Engineering*, McGraw-Hill, New York, 1968.
23. W. K. Roberts, A new wide-band balun, *Proceedings of IRE*, vol. 45, December 1957, pp. 1628–1631.
24. S. Silver, ed., *Microwave Antenna Theory and Design*, McGraw-Hill, New York, 1949.
25. J. W. Duncan and V. P. Minerva, 100 : 1 Bandwidth balun transformer, *Proceedings of IRE*, vol. 48, February 1960, pp. 156–164.
26. B. A. Munk, Baluns, Chapter 23 in J. D. Kraus and R. J. Marhefka, *Antennas*, McGraw-Hill, New York, 2002.

27. P. K. Park and C. T. Tai, Receiving antennas, Chapter 6 in Y. T. Lo and S. W. Lee, eds., *Antenna Handbook*, Van Nostrand Reinhold, New York, 1993.
28. A. Alford and A. G. Kandoian, Ultrahigh frequency loop antennas, *AIEE Transactions,* vol. 59, 1940, pp. 843–848.
29. A. J. Fenn, Arrays of horizontally polarized loop-fed slotted cylinder antennas, *IEEE Transactions on Antennas and Propagation,* vol. AP-33, no. 4, April 1985, pp. 375–382.
30. T. Tsukiji and S. Tou, On polygonal loop antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-28, no. 4, July 1980, p. 571.
31. W. C. Wilkinson et al., Two communication antennas for the Viking lander spacecraft, *1974 IEEE Antennas and Propagation Symposium Digest*, vol. 12, June 1974, pp. 214–216.
32. C. C. Kilgus, Multielement fractional turn helices, *IEEE Transactions on Antennas and Propagation*, vol. AP-16, no. 4, July 1968, pp. 499–500.
33. C. C. Kilgus, Resonant quadrifilar helix, *IEEE Transactions on Antennas and Propagation*, vol. AP-17, no. 3, May 1969, pp. 349–351.
34. A. A. Oliner, Equivalent circuits for discontinuities in balanced strip transmission line, *IRE Transactions on Microwave Theory and Techniques*, vol. MTT-3, March 1955, pp. 134–143.
35. J. S. Rao and B. N. Ras, Impedance of off-centered stripline fed series slot, *IEEE Transactions on Antennas and Propagation*, vol. AP-26, no. 6, November 1978, pp. 893, 894.
36. J. Van Bladel, Small hole in waveguide wall, *Proceedings of IEE*, vol. 118, January 1971, pp. 43–50.
37. P. K. Park and R. S. Elliott, Design of collinear longitudinal slot arrays fed by boxed stripline, *IEEE Transactions on Antennas and Propagation*, vol. AP-29, no. 1, January 1981, pp. 135–140.
38. H. E. King and J. L. Wong, A shallow ridged-cavity cross-slot antenna for the 240- to 400-MHz frequency range, *IEEE Transactions on Antennas and Propagation*, vol. AP-23, no. 5, September 1975, pp. 687–689.
39. C. A. Lindberg, A shallow-cavity UHF crossed-slot antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-17, no. 5, September 1969, pp. 558–563.
40. R. C. Hansen, ed., *Microwave Scanning Antennas*, Vol. II, Academic Press, New York, 1966.
41. R. F. Harrington, *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York, 1961, p. 69.
42. A. F. Stevenson, Theory of slots in rectangular waveguides, *Journal of Applied Physics*, vol. 19, January 1948, pp. 24–38.
43. A. Dion, Nonresonant slotted arrays, *IRE Transactions on Antennas and Propagation*, vol. AP-7, October 1959, pp. 360–365.
44. R. C. Hansen, in A. W. Rudge et al., eds., *The Handbook of Antenna Design*, Vol. II, Peter Peregrinus, London, 1983.
45. M. J. Ehrlich, in H. Jasik, ed., *Antenna Engineering Handbook*, McGraw-Hill, New York, 1961.
46. R. S. Elliott and L. A. Kurtz, The design of small slot arrays, *IEEE Transactions on Antennas and Propagation*, vol. AP-26, no. 2, March 1978, pp. 214–219.
47. R. S. Elliott, On the design of traveling-wave-fed longitudinal shunt slot arrays, *IEEE Transactions on Antennas and Propagation*, vol. AP-27, no. 5, September 1979, pp. 717–720.
48. R. S. Elliott, An improved design procedure for small arrays of shunt slots, *IEEE Transactions on Antennas and Propagation*, vol. AP-31, no. 1, January 1983, pp. 48–53.

6

MICROSTRIP ANTENNAS

Microstrip antennas are planar resonant cavities that leak from their edges and radiate. We can utilize printed circuit techniques to etch the antennas on soft substrates to produce low-cost and repeatable antennas in a low profile. The antennas fabricated on compliant substrates withstand tremendous shock and vibration environments. Manufacturers for mobile communication base stations often fabricate these antennas directly in sheet metal and mount them on dielectric posts or foam in a variety of ways to eliminate the cost of substrates and etching. This also eliminates the problem of radiation from surface waves excited in a thick dielectric substrate used to increase bandwidth.

As electronic devices continue to shrink in size, the antenna designer is pushed to reduce the antenna size as well. Cavity antennas use valuable internal volume, but we have the conflict that restricting the volume limits impedance bandwidth. Bandwidths widen with increased circuit losses (material losses) or by efficient use of the restricted volume. Bounds on bandwidth can be found by enclosing the antenna in a sphere and expanding the fields into TE and TM spherical modes [1,2]. Each mode radiates, but it requires more and more stored energy as the mode number increases. Decreasing the volume increases the Q value of each mode and a sum, weighted by the energy in each mode, determines the overall Q value. Antennas that use the spherical volume efficiently and reduce power in the higher-order modes have the greatest bandwidths. A single lowest-order mode puts an upper bound on bandwidth, given the size of the enclosing sphere. Greater volumes have potential for greater bandwidth provided that the energy in higher-order spherical modes is restricted. Increasing material losses or adding small resistors increases bandwidth beyond the single-mode bound [2]. We discover that increasing the volume of flush antennas increases the impedance bandwidth provided that the radiation mode on the structure can be maintained. Thicker substrates develop greater bandwidths, but they increase the possibility of higher-order-mode excitation and surface-wave losses. Losses limit the lower bound of bandwidth

Modern Antenna Design, Second Edition, By Thomas A. Milligan

as we reduce the thickness because efficiency degrades to a point where the bandwidth remains constant.

Microstrip consists of a metal strip on a dielectric substrate covered by a ground plane on the other side. Unlike stripline, the single ground plane shields the circuit on only one side, but normal packaged microstrip—within a receiver, for example—has a second shielding ground plane to reduce circuit interactions. The dielectric substrate retains most of the power because the shielding ground plane is spaced a few substrate thicknesses away. Removing the shield in antenna applications allows radiation from resonant cavities. We also discover feeding circuits etched on the substrate radiate to some extent, but their radiation is comparatively small.

Arrays of antennas can be photoetched on the substrate, along with their feeding networks, and microstrip provides easy connections to active devices and allows placement of preamps or distributed transmitters next to the antenna elements. Diode phase-shifter circuits etched in the microstrip form single-board phased arrays. Microstrip circuits make a wide variety of antennas possible through the use of the simple photoetching techniques.

The vast literature on microstrip antennas concentrates on the microwave circuit analysis of the internal parts of the antenna used to control the internal modes. Designers have increased the bandwidth of the antenna by coupling to multiple resonators, such as vertically stacked or coplanar coupled patches or by using internal slots and apertures. These multiple resonators increase the impedance bandwidth, and in the best cases the antenna continues to radiate the same pattern. As antenna designers we need to concentrate first on obtaining the desired pattern while working to increase the impedance bandwidth. Simple microstrip antennas have much larger pattern bandwidths than impedance bandwidths, but as more resonators are added to increase the impedance bandwidth, spreading in the horizontal plane alters the radiated pattern and we must return to concentrate on the pattern.

Microstrip patch antennas consist of metal patches large with respect to normal transmission-line widths. A patch radiates from fringing fields around its edges. Impedance match occurs when a patch resonates as a resonant cavity. When matched, the antenna achieves peak efficiency. A normal transmission line radiates little power because the fringing fields are matched by nearby counteracting fields. Power radiates from open circuits and from discontinuities such as corners, but the amount depends on the radiation conductance load to the line relative to the patches. Without proper matching, little power radiates.

The edges of a patch appear as slots whose excitations depend on the internal fields of the cavity. A general analysis of an arbitrarily shaped patch considers the patch to be a resonant cavity with metal (electric) walls of the patch and the ground plane and magnetic or impedance walls around the edges. The radiating edges and fringing fields present loads along the edges. In one analysis [3] the patch effective size is increased to account for the capacitive susceptance of fringing fields, and the radiation admittance is ignored to calculate resonant frequency. The far field is integrated to compute radiated power and the equivalent radiation conductance. The second method [4] is to retain the patch size but satisfy boundary conditions into a loaded wall whose load is determined by radiation and fringing fields. Assuming a constant electric field from the ground plane to the substrate allows solutions in terms of modes TM to the substrate thickness. Boundary conditions determine possible modes and correspond to the dual TE modes of waveguides having electric walls. Patches in the shape of standard coordinate system

axes, such as rectangular and circular, give solutions in terms of tabulated functions. Numerical techniques used for arbitrarily shaped waveguides can be applied to patches with nonstandard shapes. We consider only rectangular and circular patches.

6-1 MICROSTRIP ANTENNA PATTERNS

We start our discussion of patches with their pattern characteristics. It is difficult to separate a discussion of pattern from the internal construction consideration, but we will only briefly discuss the internal structures that affect the pattern. The small size of microstrip antennas limits control of the pattern and we must use arrays of patches to control its pattern seriously. Rectangular and circular are the most common shapes for microstrip antennas and they radiate similar broad patterns. When we load the cavity to shrink its size, it radiates wider beamwidth patterns that lower directivity (gain). Antennas that couple to coplanar patches to increase the impedance bandwidth will radiate narrower beams, but the basic patch has a wide beamwidth. If we couple to multiple coplanar patches, we can expect the pattern to narrow or vary its shape as the mixture of modes on the various patches changes over the frequency range of operation.

Patches consist of metal plates suspended over large ground planes. We excite the cavity in a variety of ways that we discuss later. Electric currents flow on the plate and on the ground plane around the antenna, and these radiate. If we use vertical probes to excite the antenna from coaxial lines, the currents flowing on these radiate and add to the pattern. We can reduce the antenna size by adding vertical shorting plates (quarter-wave patches) or shorting pins near the feed pins (compact patches), and these also radiate from the current flow on them. Remember that the patch radiates from real electric currents, although the distribution is complicated.

We simplify the problem of computing patch radiation by using magnetic currents along the edges. Figure 6-1 illustrates the fringing electric fields around the edges of square and circular patch antennas excited in the lowest-order cavity modes. The arrow sizes indicate the magnitude of the fields. The square patch has nearly uniform fields along two edges we call the width, and a sinusoidal variation along the other two edges, called the resonant length. The fields vanish along a virtual electrically short-circuited plane halfway across the patches. On either side of the short-circuit plane, the fields are directed in opposite directions. Looking from above the fields along the width, both edges are in the same direction. The circular patch fringing fields distribution varies as $\cos \phi$, where the angle ϕ along the rim is measured from the peak electric field. Magnetic currents found from the fringing electric fields can replace the electric currents located on the patch and the surrounding ground plane for pattern analysis. Figure 6-2 shows the distribution of magnetic currents around the edges, with the size of the arrowhead indicating magnitude.

Our use of magnetic currents around the patch perimeter reduces the pattern calculation to equivalent slots. A two-element array consisting of slots with equivalent uniform magnetic currents produces the E-plane radiation of a rectangular patch. To first order, the slots are spaced $\lambda/2\sqrt{\varepsilon_r}$ and we can determine the pattern from the equivalent two-element array. The magnetic currents along the resonant length sides individually cancel because the current changes direction halfway across the edge. The currents also cancel from side to side. These cancellations eliminate pattern contributions to the E- and H-plane patterns. The slot length determines the H-plane pattern.

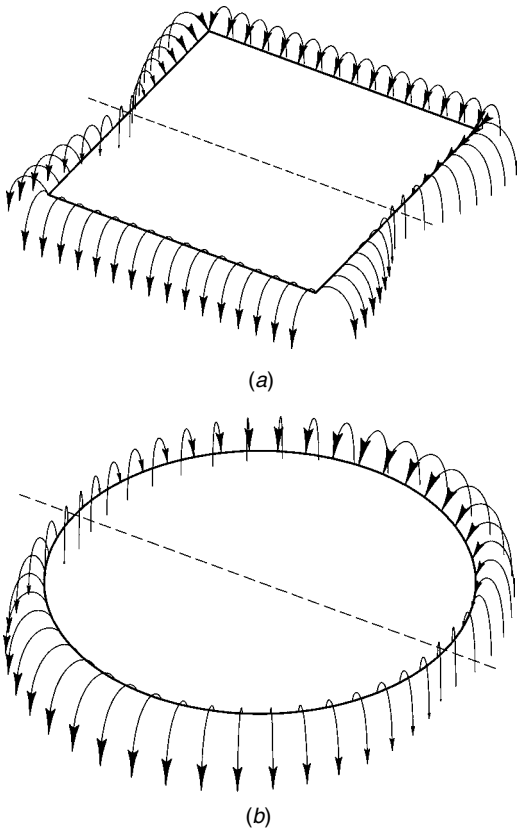


FIGURE 6-1 Fringing electric fields around microstrip patches: (*a*) square; (*b*) circular. (From L. Diaz and T. A. Milligan, *Antenna Engineering Using Physical Optics,* Figs. 3.12 and 3.19, © 1996 Artech House, Inc.).

The H-plane of the slot has the same pattern as the E-plane of a dipole and produces a null along its axis. Figure 6-3 illustrates the pattern of a patch on an infinite ground plane using a free-space substrate. The two-element slot array in the E-plane generates a null along the ground plane because the elements are spaced $\lambda/2$. The H-plane dashed curve shows the null along the ground plane due to the polarization of the slots. The light curves give the Huygens source polarization (Section 1-11.2) patterns in the diagonal planes. The antenna radiates cross-polarization (dashed curve) in this plane from the combination of separated magnetic currents along the resonant-length sides and from the unbalance in the beamwidths in the principal planes.

When we design a microstrip patch on a dielectric substrate, the size reduction brings the two slots closer together and widens the E-plane beamwidth and eliminates its null along the ground plane. Figure 6-4 illustrates the pattern of a patch designed for a substrate with $\varepsilon_r = 2.2$. The H-plane pattern retains its null along the ground plane due to the slot pattern. The cross-polarization of the Huygens source in the diagonal plane increases because of the increased difference between the beamwidths of the principal plane patterns. Table 6-1 gives the directivity of a square and circular patch on an infinite ground plane found by integrating the pattern. The range of directivity

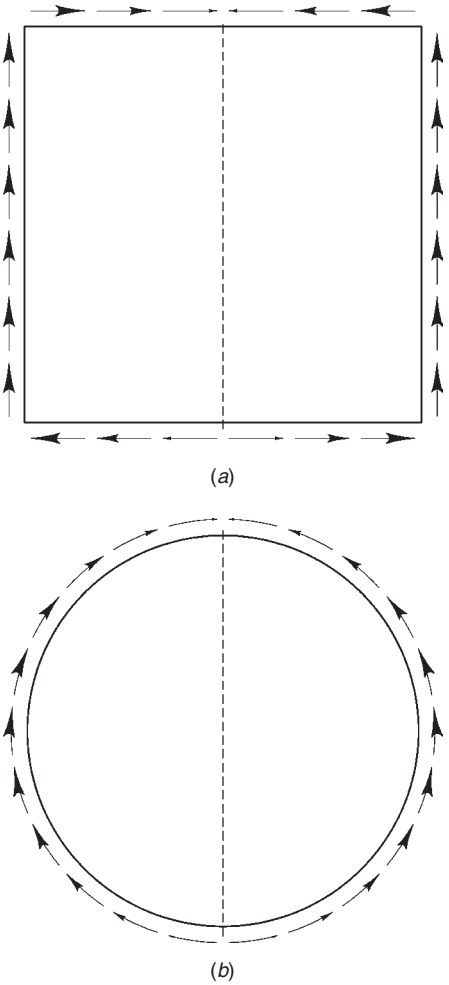


FIGURE 6-2 Equivalent magnetic currents on the edges of microstrip patches: (*a*) square; (*b*) circular.

of a patch is limited. Increasing the width of a rectangular patch increases directivity by shrinking the H-plane beamwidth.

We gain some control of the pattern by placing the patch on a finite ground plane. Figure 6-5 shows the pattern of a square patch on a 2.21 dielectric constant substrate when located on circular disks 5λ, 2λ, and 1λ in diameter. On a 5λ ground plane, edge diffraction adds ripple to the pattern. As the ground plane increases, the angular separation between the ripples decreases, due to the increased array size of the radiation from the two edges. The H-plane pattern widens significantly for 1λ- and 2λ-diameter ground planes, as the limited ground plane can no longer support the currents that make the patch edge radiate like a slot. Although the principal-plane beamwidths are more nearly equal for the patch on the 2λ-diameter disk, the cross-polarization in the diagonal plane increases relative to the pattern on the infinite ground plane. The 1λ ground plane increases the gain of the patch by about 1 dB relative to the patch on an

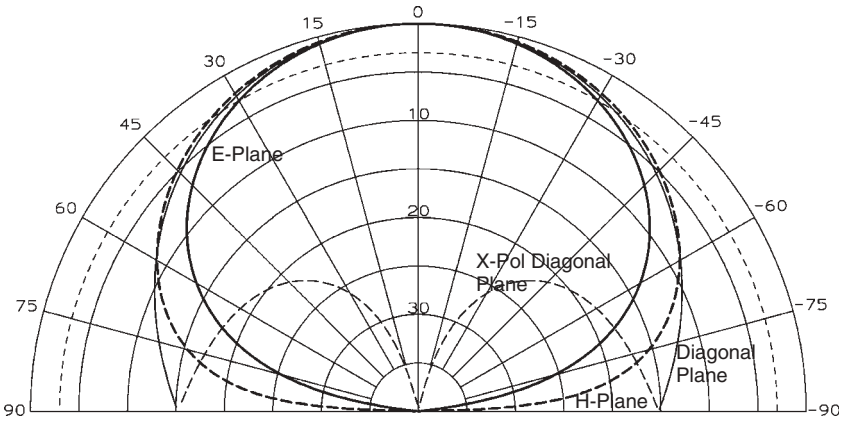


FIGURE 6-3 Patterns of microstrip patch on a free-space substrate mounted on an infinite ground plane.

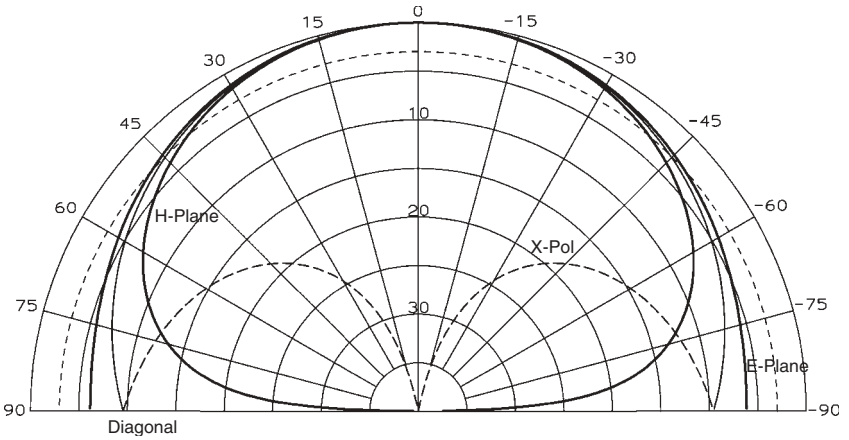
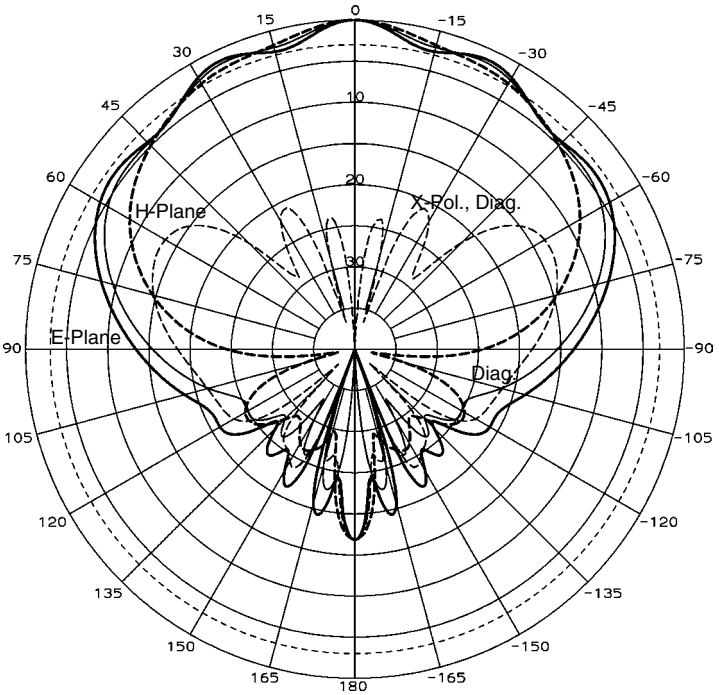


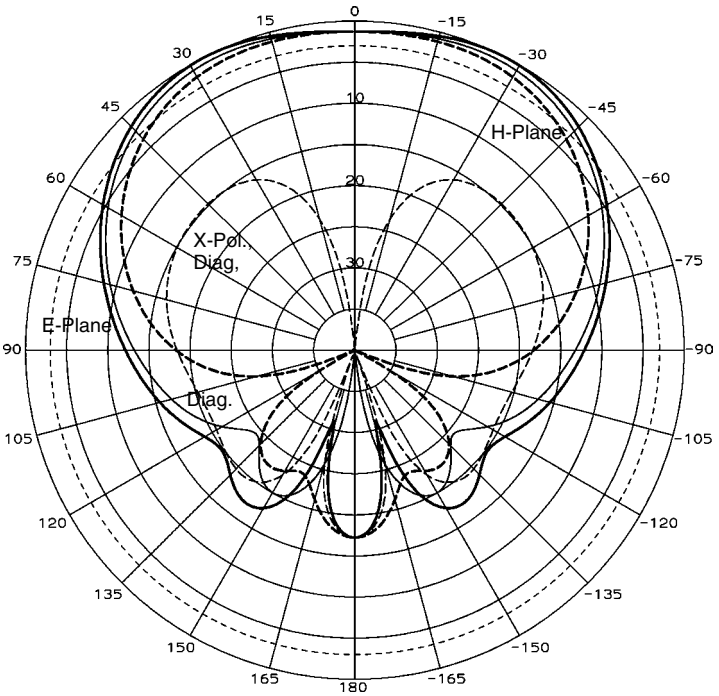
FIGURE 6-4 Patterns of microstrip patch on a dielectric substrate $\varepsilon_r = 2.2$ over an infinite ground plane.

TABLE 6-1 Estimated Directivity of Square and Circular Microstrip Patches on a Large Ground Plane

Dielectric Constant	Square Patch (dB)	Circular Patch (dB)
1.0	8.4	9.8
2.0	7.7	7.6
3.0	7.2	6.7
4.0	7.0	6.2
6.0	6.7	5.8
8.0	6.5	5.5
10.0	6.4	5.4
16.0	6.3	5.1

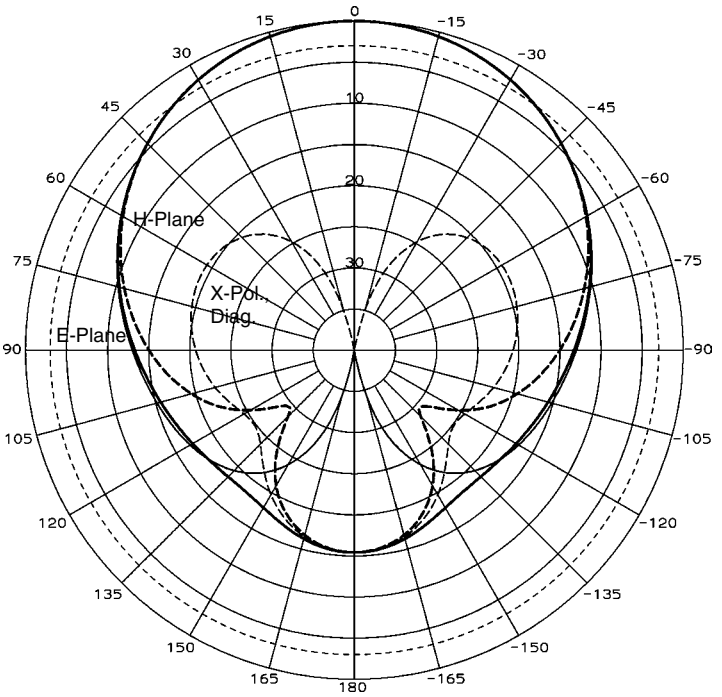


(*a*)



(*b*)

FIGURE 6-5 Patterns of microstrip patches with dielectric substrate $\varepsilon_r = 2.2$ mounted over finite circular ground planes: (*a*) 5λ diameter; (*b*) 2λ diameter. (*continued*)



(*c*)

FIGURE 6-5 (*continued*) (*c*) 1λ diameter.

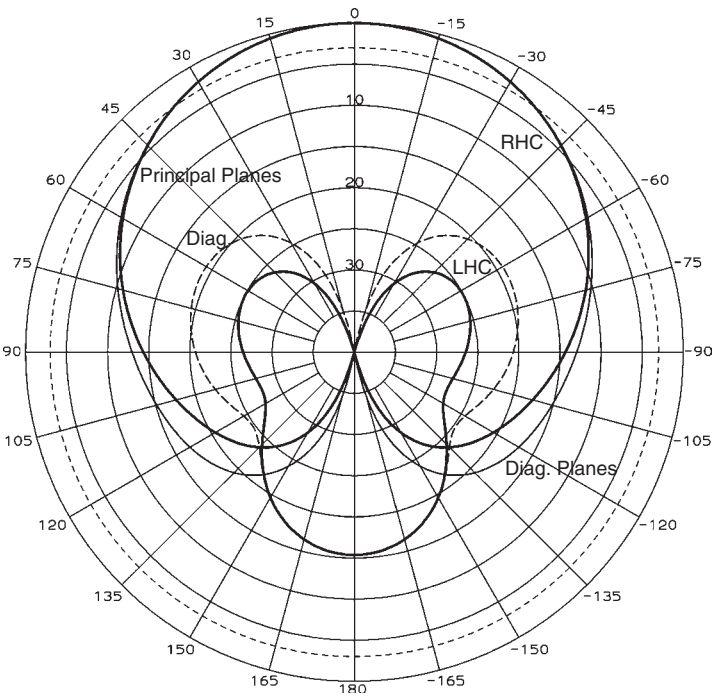


FIGURE 6-6 Circularly polarized patch mounted on a 1λ-diameter ground plane.

infinite ground plane. In this case the edge diffractions add constructively to narrow the beamwidths. We can take advantage of the nearly equal E- and H-plane patterns in the forward hemisphere to produce a pattern with excellent circular polarization over the entire hemisphere when we feed the patch for circular polarization. Figure 6-6 gives the circular polarization pattern when the patch is fed in two spots with equal signals phased 90° apart. The cross-polarization is 13 dB below the co-polarization at $\theta = 90^\circ$ in the principal planes and -7 dB in the diagonal plane. We retain these excellent polarization characteristics over a large ground plane if we place the finite ground plane on a 1λ or greater pedestal above the ground plane.

6-2 MICROSTRIP PATCH BANDWIDTH AND SURFACE-WAVE EFFICIENCY

Microstrip patches radiate from the currents induced on the patch or equivalently, the magnetic currents around the periphery of the patch and from surface waves induced in the dielectric slab. The surface waves radiate when they reach the edges of the substrate and their radiation contributes to the normal patch radiation. The fringing fields from the patch to the ground plane readily excite the lowest-order surface-wave TM_0 mode that has no low frequency cutoff. Any thickness dielectric slab supports this mode. We can control the surface-wave radiation by limiting the substrate area or by adding etched photonic bandgap patterns to the open areas of the substrate, but generally, surface waves are undesirable. As the substrate thickness or dielectric constant increases, the ratio of the power in surface waves increases. When we calculate the microstrip patch antenna impedance bandwidth, we must include the directly radiated power and the surface-wave power. For most cases we consider surface-wave radiation as reducing radiation efficiency, but for a single patch on a substrate with limited area, its radiation can add constructively. We eliminate surface waves by using metal plate patches without dielectric substrates or low-density foam supports of the patch. Surface waves are bound to the dielectric similar to any transmission line except that the field decays exponentially in the direction normal to the surface. Because the surface wave is excited along the finite edges of the patch, it spreads in the horizontal plane. The radiation spreads like a two-dimensional wave and the fields decay as $1/\sqrt{r}$, where r is the horizontal distance from the edge. This is a far-field approximation, and close to the edge it is a near-field problem. Unfortunately, these surface waves increase the coupling between patches fabricated on the same substrate.

Simple formulas have been derived for the impedance bandwidth of rectangular patches that include the surface-wave loss [5]. Since substrates can be both electric and magnetic, we define the index of refraction of a patch substrate that includes both parameters: $n = \sqrt{\varepsilon_r \mu_r}$. The idea is that the ratio of space-wave radiation to surface-wave radiation can be found for any small antenna mounted on the substrate and we can then apply it to a patch. By integrating the power density in the radiation from a horizontal Hertzian (incremental) dipole spaced the substrate thickness over a ground plane, we obtain the space-wave radiated power in closed form given the substrate thickness h and the free-space propagation constant k:

$$P_R^h \simeq k^2 (kh)^2 \cdot 20\,\mu_r^2 C_1$$
$$C_1 = 1 - \frac{1}{n^2} + \frac{0.4}{n^4} \tag{6-1}$$

We express the current on the patch as an integral of Hertzian dipoles. The surface-wave power generated in the substrate by the Hertzian dipole can be simplified when the substrate is thin:

$$P_{\mathrm{SW}}^{h} = k^2(kh)^3 \cdot 15\pi\mu_r^3 \left(1 - \frac{1}{n^2}\right)^3 \tag{6-2}$$

We define the surface-wave radiation efficiency by the ratio of radiated power to total power:

$$\eta_{\mathrm{SW}} = \frac{P_r^h}{P_R^h + P_{\mathrm{SW}}^h} = \frac{4C_1}{4C_1 + 3\pi kh\mu_r(1 - 1/n^2)^3} \tag{6-3}$$

We relate the power radiated by a patch to a Hertzian dipole by integrating the surface current on the patch consisting of a distribution of small dipoles to calculate the total space-wave power of the patch:

$$P_R = P_R^h m_{\mathrm{eq}}^2 = P_R^h \left(\iint_S J_S \, dx \, dy \right)^2 \tag{6-4}$$

For a rectangular patch the ratio of P_R to $P_R^h m_{\mathrm{eq}}^2$, p, can be approximated by a simple formula given the resonant length L, the width W, and the propagation constant k:

$$p = 1 - \frac{0.16605(kW)^2}{20} + \frac{0.02283(kW)^4}{560} - \frac{0.09142(kL)^2}{10} \tag{6-5}$$

The 2 : 1 VSWR of the rectangular patch is related to the quality factor Q that includes the space- and surface-wave radiations:

$$\mathrm{BW} = \frac{1}{\sqrt{2}\,Q} = \frac{16C_1 p}{3\sqrt{2}\,\eta_{\mathrm{SW}}} \frac{1}{\varepsilon_r} \frac{h}{\lambda_0} \frac{W}{L} \tag{6-6}$$

Figure 6-7 plots the 2 : 1 VSWR bandwidth given by Eq. (6-6) for common substrates versus the free-space thickness in wavelengths and includes the radiation due to surface waves. The surface-wave radiation found using Eq. (6-3) becomes a significant part of the total radiation as the substrate thickness increases or the dielectric constant increases, as shown in Figure 6-8 of the surface-wave loss.

For a single resonator circuit model for a patch, Eq. (6-6) computes bandwidth from the Q and the allowable input VSWR:

$$\mathrm{BW} = \frac{\mathrm{VSWR} - 1}{Q\sqrt{\mathrm{VSWR}}} \quad \text{or} \quad Q = \frac{\mathrm{VSWR} - 1}{\mathrm{BW}\sqrt{\mathrm{VSWR}}} \tag{6-7}$$

We determine bandwidth at different VSWR levels by manipulating Eq. (6-7):

$$\frac{\mathrm{BW}_2}{\mathrm{BW}_1} = \frac{(\mathrm{VSWR}_2 - 1)\sqrt{\mathrm{VSWR}_1}}{(\mathrm{VSWR}_1 - 1)\sqrt{\mathrm{VSWR}_2}} \tag{6-8}$$

Quality factor Q is another way of expressing efficiency. The Q used in Eq. (6-6) is the combination of the space-wave radiation Q_R and the surface-wave radiation Q_{SW}:

$$\frac{1}{Q} = \frac{1}{Q_{\mathrm{rad}}} = \frac{1}{Q_R} + \frac{1}{Q_{\mathrm{SW}}} = \frac{P_R + P_{\mathrm{SW}}}{\omega W_T}$$

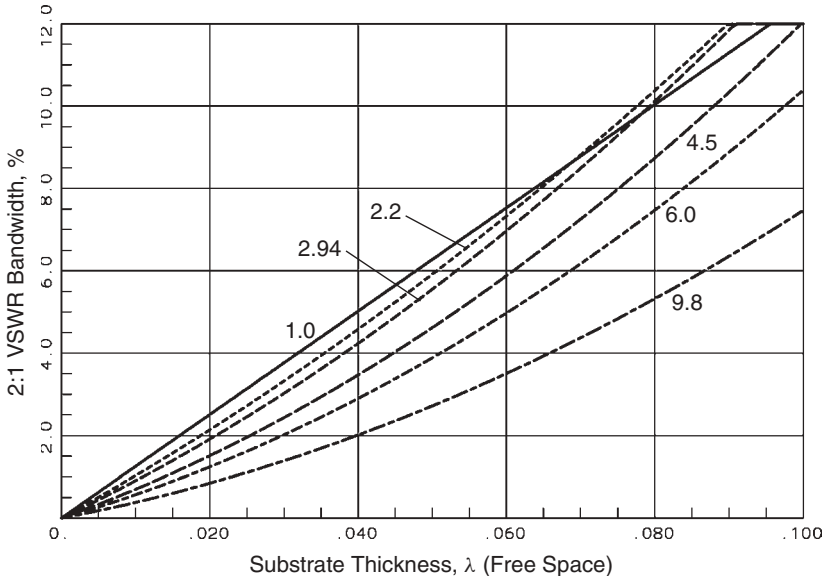


FIGURE 6-7 2:1 VSWR bandwidth of square microstrip patches versus substrate thickness in free-space wavelengths, including surface-wave radiation.

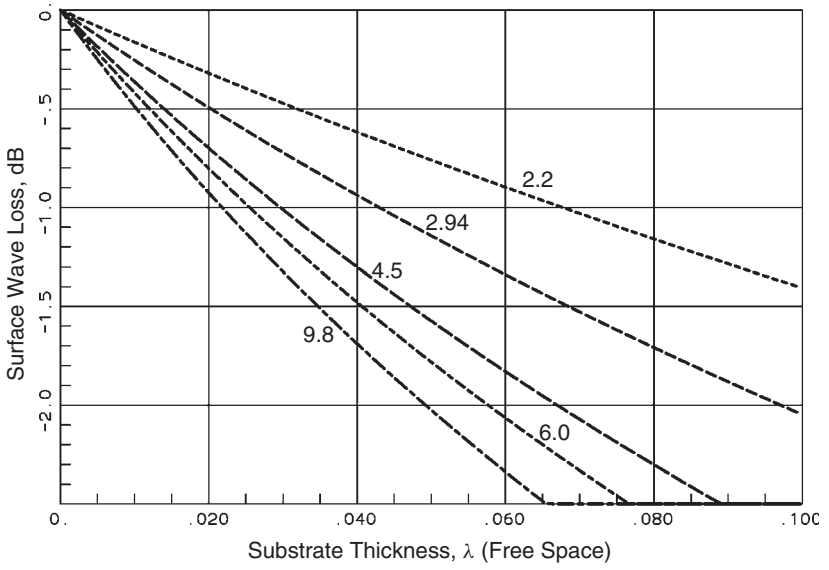


FIGURE 6-8 Surface-wave loss of microstrip patches versus substrate thickness for common substrate dielectric constants.

W_T is the energy stored in the patch and the surface wave and $\omega = 2\pi f$, the radian frequency. Equation (6-3) can be expressed in terms of Q:

$$\eta_{\text{SW}} = \frac{Q}{Q_R} = \frac{Q_{\text{rad}}}{Q_R}$$

The surface wave is not a dissipation loss, but potentially an uncontrolled radiation. Dielectric and conductor losses increase the impedance bandwidth of the patch, but reduce its gain. We express these losses as Q to compute patch efficiency. Given the dielectric loss tangent, $\tan\delta$, and the patch conductivity σ, we have two more Q terms that reduce the overall Q of the patch in terms of impedance bandwidth:

$$Q_d = \frac{1}{\tan\delta} \quad \text{and} \quad Q_c = h\sqrt{\pi f \mu_0 \sigma} \tag{6-9}$$

The total quality factor Q_T is found from the sum of the inverses:

$$\frac{1}{Q_T} = \frac{1}{Q_R} + \frac{1}{Q_{\text{SW}}} + \frac{1}{Q_d} + \frac{1}{Q_c} \tag{6-10}$$

If we attempt to fabricate a patch on a thin substrate, Q_d and Q_c become commensurate with the radiation Q_{rad} and efficiency suffers. The impedance bandwidth increases due to the dissipation in the microstrip patch. Figure 6-7 does not include these losses.

Dielectric Slab Surface Wave We consider the dielectric slab surface wave because it can be excited not only by a microstrip patch but by any wave that passes across it. The slab binds a portion of the wave and releases it when it diffracts at its edges. The surface-wave device slows the wave velocity of this wave relative to the space-wave signal, and when it radiates from the edges it no longer adds in phase with the space wave. The surface-wave fields decrease exponentially in the direction normal to the surface, and the exponential rate increases as the binding increases and the wave propagates more slowly.

A dielectric slab on a ground plane will support TM modes when thin and TE modes when thick. The TM mode is polarized normal to the slab surface, whereas the TE mode is polarized parallel to the slab surface. A TM mode requires an inductive surface such as a corrugated ground plane to bind the wave. While corrugations prevent propagation between the slots, the wave propagates in the dielectric slab by bouncing between the two interfaces at an angle with respect to the surfaces. The second surface can be either free space or a conductor. To solve for the fields, we equate not only the wave impedance at the boundary but the propagation constants in the two regions as well.

We deduce the grounded dielectric slab solution from a slab twice as thick in free space that has an odd-mode electric field excitation on the slab sides. The center becomes a virtual short circuit for the odd-mode excitation. We divide the space around the slab into three regions: 1 above the slab, 2 in the slab, and 3 below the slab and then derive the fields from potential functions [6, p. 129]:

$$\begin{aligned}
\psi_1 &= A_1 \exp\left(\frac{-2\pi b x}{\lambda}\right) \exp(-jk_z z) \\
\psi_2 &= A_2 \begin{Bmatrix} \sin \dfrac{2\pi p_x x}{\lambda} \\ \cos \dfrac{2\pi p_x x}{\lambda} \end{Bmatrix} \exp(-jk_z z) \\
\psi_3 &= \pm A_1 \exp\left(\frac{2\pi b x}{\lambda}\right) \exp(-jk_z z)
\end{aligned} \tag{6-11}$$

where the sign of ψ_3 depends on satisfying continuous tangential fields across the lower slab boundary. The center of the coordinate normal to the slab (x) is the slab center. Equating the propagation constants and x-directed wave impedances produces transcendental equations in the transverse propagation constant in the slab p_x:

$$\sqrt{\left(\frac{\omega a}{2}\right)^2(\varepsilon_1\mu_1-\varepsilon_0\mu_0)-\left(\frac{\pi p_x a}{\lambda}\right)^2}=\pm B_0\frac{\pi p_x a}{\lambda}\left\{\begin{matrix}\tan\dfrac{\pi p_x a}{\lambda}\\ \cot\dfrac{\pi p_x a}{\lambda}\end{matrix}\right\} \tag{6-12}$$

where $B_0=\mu_0/\mu_1$ for TE waves. $B_0=\varepsilon_0/\varepsilon_1$ for TM waves, ω is the radian frequency ($2\pi f$), a is the slab thickness, and ε_1 and μ_1 are the permittivity and permeability of the slab. We solve for p_x [Eq. (6-12)] numerically or graphically and use

$$\frac{\pi b}{\lambda}a=\sqrt{\left(\frac{\omega a}{2}\right)^2(\varepsilon_1\mu_1-\varepsilon_0\mu_0)-\left(\frac{\pi p_x}{\lambda}\right)^2} \tag{6-13}$$

to determine attenuation constant b and the relative propagation constant P of the slab surface wave:

$$k_z=Pk=k\sqrt{1+b^2}\quad\text{or}\quad P=\sqrt{1+b^2} \tag{6-14}$$

For the TM_0 mode we can use an approximate expression for P instead of solving Eq. (6-12) numerically when the slab is thin [7]:

$$P^2=1+\frac{(\varepsilon_r\mu_r-1)^2}{(\varepsilon_r\mu_r)^2}(2ka)^2 \tag{6-15}$$

Equation (6-12) has an infinite number of solutions, corresponding to the multiple values of the tangent and cotangent functions. Order 0 corresponds to the tangent function from 0 to 90°; order 1 corresponds to the cotangent function from 90 to 180°; and so on. Even-mode orders use the tangent function, and odd-mode orders use the cotangent function. We define the cutoff frequency as the point where $\alpha=0$, the transition between attached and detached waves:

$$\lambda_c=\frac{2a}{n}\sqrt{\frac{\varepsilon_1\mu_1}{\varepsilon_0\mu_0}-1} \tag{6-16}$$

The cutoff frequency for the zeroth-order mode is zero. Only the TM_0 mode has odd symmetry, required for the grounded slab. The grounded slab supports even-order TM modes and odd-order TE modes. Equation (6-12) coupled to Eq. (6-13) has been solved numerically to generate Tables 6-2 and 6-3. Table 6-4 lists the thicknesses of a slab in free space supporting the TM_0 mode for a given P. The grounded slab is one-half the thickness of the values in Table 6-2. Similarly, Table 6-3 lists the thicknesses for the TE_1 mode. Equation (6-16) can be solved for the minimum thickness to support the TE_1 mode. Below that thickness the waves do not bind to the surface.

Besides microstrip patches, we feed these surfaces from either a small horn or a parallel-plate transmission line. We match the feed polarization to the mode on the slab, but the slab binds only part of the power. The rest radiates directly from the feed or reflects to the feed input. We can feed an ungrounded slab by centering it in a waveguide. The TE_{10} waveguide mode excites the TE_0 slab mode when the mode velocity determining thickness is in the H-plane. Like the grounded slab with the TM_0

TABLE 6-2 Thickness (λ_0) of a Dielectric Slab Supporting a TM_0 Mode[a]

	Dielectric Constant				
P	2.21	2.94	4.50	6.00	9.80
1.001	0.02699	0.02152	0.01831	0.01839	
1.002	0.03672	0.03041	0.02574	0.02420	
1.005	0.05792	0.04784	0.04032	0.03744	0.03446
1.01	0.08162	0.06713	0.05623	0.05195	0.04710
1.02	0.1147	0.09355	0.07746	0.07094	0.06316
1.04	0.1607	0.1289	0.1046	0.09444	0.08180
1.06	0.1956	0.1545	0.1231	0.1099	0.09331
1.08	0.2253	0.1752	0.1374	0.1215	0.1015
1.10	0.2520	0.1930	0.1491	0.1307	0.1078
1.12	0.2770	0.2088	0.1590	0.1384	0.1129
1.14	0.3012	0.2233	0.1677	0.1450	0.1171
1.16	0.3251	0.2369	0.1756	0.1508	0.1208
1.18	0.3493	0.2499	0.1827	0.1560	0.1240
1.20	0.3741	0.2625	0.1894	0.1607	0.1269
1.25	0.4426	0.2934	0.2045	0.1712	0.1329
1.30	0.5282	0.3250	0.2182	0.1803	0.1380
1.35	0.6492	0.3593	0.2314	0.1887	0.1424
1.40		0.3986	0.2444	0.1966	0.1463

[a] Use half-thickness for a slab on a ground plane.

TABLE 6-3 Thickness (λ_0) of a Dielectric Slab Supporting a TE_1 Mode[a]

	Dielectric Constant				
P	2.21	2.94	4.50	6.00	9.80
1.001	0.4469	0.3689	0.2743	0.2260	0.1701
1.002	0.4720	0.3709	0.2774	0.2272	0.1705
1.005	0.4829	0.3765	0.2770	0.2302	0.1717
1.01	0.4961	0.3843	0.2810	0.2330	0.1736
1.02	0.5164	0.3962	0.2873	0.2373	0.1761
1.04	0.5494	0.4150	0.2968	0.2438	0.1797
1.06	0.5790	0.4313	0.3049	0.2492	0.1825
1.08	0.6078	0.4465	0.3122	0.2540	0.1851
1.10	0.6368	0.4613	0.3191	0.2585	0.1874
1.15	0.7140	0.4982	0.3356	0.2690	0.1928
1.20	0.8046	0.5372	0.3518	0.2790	0.1978
1.25	0.9182	0.5802	0.3683	0.2890	0.2026
1.30	1.0712	0.6291	0.3856	0.2992	0.2073

[a] Use half-thickness for a slab on a ground plane.

mode, the TE_0 mode has no cutoff frequency for a free-space slab. Table 6-4 lists the slab thicknesses for a given relative propagation constant for the TE_0 mode.

The surface-wave power was found in terms of the relative propagation constant P[7]:

$$P_{SW} = \frac{15\pi k^2 n^2 \mu_r^3 (P^2 - 1)}{n^2 \left(\dfrac{1}{\sqrt{P^2 - 1}} + \dfrac{\sqrt{P^2 - 1}}{n^2 - P^2} \right) + kh \left[1 + \dfrac{n^4 (P^2 - 1)}{n^2 - P^2} \right]} \tag{6-17}$$

TABLE 6-4 Thickness (λ_0) of a Dielectric Slab Supporting a TE_0 Mode

	Dielectric Constant				
P	2.21	2.94	4.50	6.00	9.80
1.001	0.01274	0.00994			
1.002	0.01684	0.01174	0.00661	0.00410	0.00250
1.005	0.02649	0.01666	0.00920	0.00638	0.00364
1.01	0.03772	0.02341	0.01289	0.00905	0.00514
1.02	0.05409	0.03345	0.01846	0.01286	0.00729
1.04	0.07872	0.04823	0.02640	0.01839	0.01040
1.06	0.09935	0.06027	0.03275	0.02276	0.01284
1.08	0.1184	0.07104	0.03832	0.02656	0.01494
1.10	0.1368	0.08113	0.04343	0.03002	0.01684
1.15	0.1833	0.1051	0.05507	0.03779	0.02106
1.20	0.2348	0.1290	0.06595	0.04489	0.02483
1.25	0.2968	0.1542	0.07661	0.05168	0.02835

We combine Eq. (6-1) for the space-wave power with Eq. (6-17) for the surface-wave power to calculate efficiency in the same manner as Eq. (6-3). The results are similar.

6-3 RECTANGULAR MICROSTRIP PATCH ANTENNA

Although design equations will be given below for single-layer rectangular and circular patches, serious design work should use one of the excellent available commercial design codes [8]. Their use reduces the need to modify the final dimensions using a knife to remove metal or metal tape to increase the patches. Antennas can be built with tuning tabs, but the labor to trim these increases cost. Tuning tabs are unsuitable for arrays when the input port to individual antennas cannot be accessed. As we add layers to increase bandwidth, a cut-and-try method becomes extremely difficult, and numerical methods are a necessity.

Rectangular patch antennas can be designed by using a transmission-line model [9] suitable for moderate bandwidth antennas. Patches with bandwidths of less than 1% or greater than 4% require a cavity analysis for accurate results, but the transmission-line model covers most designs. The lowest-order mode, TM_{10}, resonates when the effective length across the patch is a half-wavelength. Figure 6-9 demonstrates the patch fed below from a coax along the resonant length. Radiation occurs from the fringing fields. These fields extend the effective open circuit (magnetic wall) beyond the edge. The extension is given by [10]

$$\frac{\Delta}{H} = 0.412 \frac{\varepsilon_{\text{eff}} + 0.300}{\varepsilon_{\text{eff}} - 0.258} \frac{W/H + 0.262}{W/H + 0.813} \tag{6-18}$$

where H is the substrate thickness, W the patch nonresonant width, and ε_{eff} the effective dielectric constant of a microstrip transmission line the same width as the patch.

A suitable approximation for ε_{eff} is given by [11]

$$\varepsilon_{\text{eff}} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2} \left(1 + \frac{10H}{W}\right)^{-1/2} \tag{6-19}$$

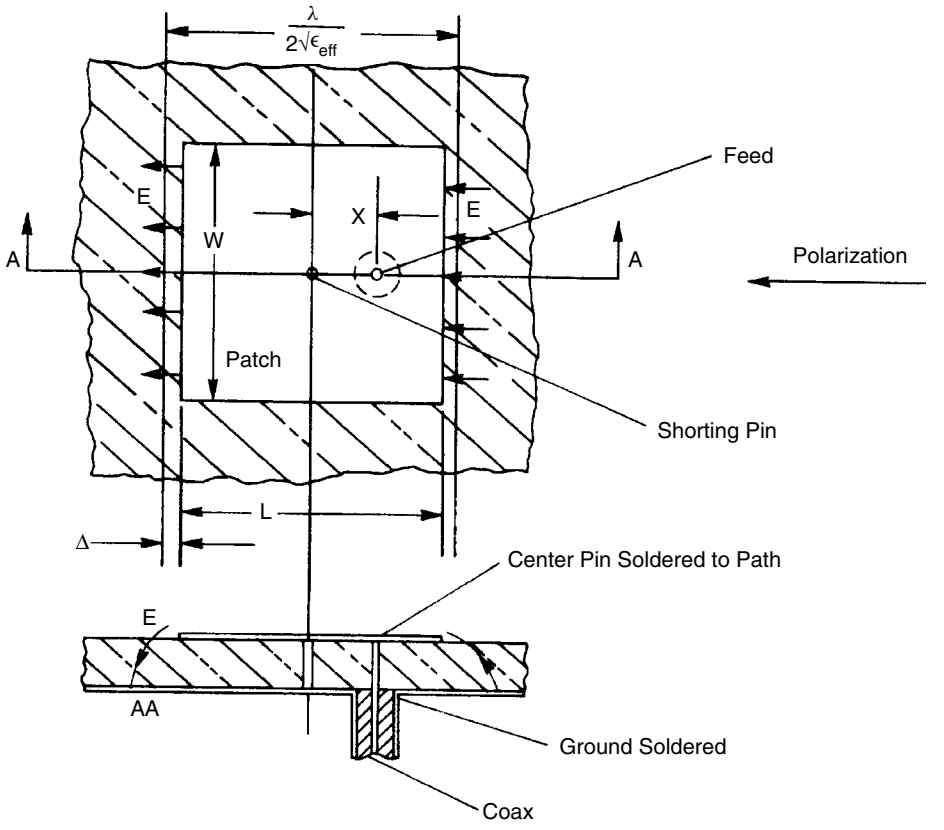


FIGURE 6-9 Coax-fed microstrip patch antenna.

where ε_r is the substrate dielectric constant. The transmission-line model represents the patch as a low-impedance microstrip line whose width determines the impedance and effective dielectric constant. A combination of parallel-plate radiation conductance and capacitive susceptance loads both radiating edges of the patch.

Harrington [6, p. 183] gives the radiation conductance for a parallel-plate radiator as

$$G = \frac{\pi W}{\eta \lambda_0}\left[1 - \frac{(kH)^2}{24}\right] \tag{6-20}$$

where λ_0 is the free-space wavelength. The capacitive susceptance relates to the effective strip extension:

$$B = 0.01668\frac{\Delta}{H}\frac{W}{\lambda}\varepsilon_{\text{eff}} \tag{6-21}$$

Example Design a square microstrip patch antenna at 3 GHz on a 1.6-mm substrate with a dielectric constant of 2.55 (woven Teflon fiberglass). The patch will be approximately a half-wavelength long in the dielectric.

Assume at first that the width is λ/2.

$$W = \frac{c}{2f\sqrt{\varepsilon_r}} = 31.3\,\text{mm}$$

By Eq. (6-19), $\varepsilon_{\text{eff}} = 2.405$. On substituting that value into Eq. (6-18), we obtain the effective cutback on each edge; $\Delta = 0.81\,\text{mm}$. The resonant length is

$$L = \frac{c}{2f\sqrt{\varepsilon_{\text{eff}}}} - 2\Delta = 30.62\,\text{mm}$$

When we use this length as the width (square patch) to calculate the effective dielectric constant, we obtain 2.403, very close to the initial value. We can iterate it once more and obtain 30.64 mm for the resonant length. The input conductance of the patch fed on the edge will be twice the conductance of one of the edge slots [Eq. (6-20)]:

$$G = \frac{30.64\,\text{mm}}{120(100\,\text{mm})}\left\{1 - \frac{[2\pi(1.6)/100]^2}{24}\right\} = 2.55\,\text{mS}$$

$$R = \frac{1}{2G} = 196\,\Omega$$

A microstrip feeding line can be attached to the center of one of the radiating edges but 50-Ω transmission lines become inconveniently wide on low-dielectric-constant substrates. More convenient, 100-Ω narrower lines have about the same low loss and are generally used in feed networks. To transform the 196-Ω input resistance of the example above to 100 Ω, we use a 140-Ω quarter-wavelength transformer. The bandwidth of the transformer far exceeds that of the antenna.

In the example above, we have a square patch. Why doesn't the antenna radiate from the other two edges? We can equally well say that the patch is a transmission line in the other direction. The equal distances from the feed point to the nonradiating edges produce equal fields from the patch to ground. Equal fields on the edges set up a magnetic wall (virtual open circuit) through the centered feed line and create a poor impedance match to the feed.

We expand the radiating edge fields in an odd mode, since the power traveling across the patch loses 180° of phase. The odd mode places a virtual short circuit halfway through the patch. A shorting pin through the center (Figure 6-9) has no effect on radiation or impedance, but it allows a low-frequency grounding of the antenna. The patch can be fed by a coax line from underneath (Figure 6-9). The impedance varies from zero in the center to the edge resistance approximately as

$$R_i = R_e \sin^2 \frac{\pi x}{L} \qquad 0 \le x \le \frac{L}{2} \tag{6-22}$$

where R_i is the input resistance, R_e the input resistance at the edge, and x the distance from the patch center. The feed location does not significantly affect the resonant frequency. By using Eq. (6-22), we locate the feed point given the desired input impedance:

$$x = \frac{L}{\pi} \sin^{-1} \sqrt{\frac{R_i}{R_e}} \tag{6-23}$$

Compute the 50-Ω feed point in the example above:

$$x = \frac{30.64}{\pi} \sin^{-1} \sqrt{\frac{50}{196}} = 5.16\,\text{mm}$$

The feed pin currents add to the pattern by radiating a monopole pattern. Figure 6-10 shows this radiation for a patch using a free-space substrate where the E-plane radiating edges are spaced $\lambda/2$. The pattern of Figure 6-10 has a null along the ground plane in the E-plane, but the monopole radiation increases the radiation along the ground plane. On one side the radiation adds and on the other it subtracts from the E-plane pattern to form a null tilted above the ground plane. The H-plane pattern now contains cross-polarization. We can reduce the monopole radiation by feeding the patch at a second port located an equal distance from the center on the opposite side. This requires an external feed network that divides the power equally between the two ports with a 180° phase difference. The problem with this feed arrangement is that significant power is coupled between the two feeds in the equivalent microwave circuit of the patch. The estimated value of −6 dB coupling between the ports causes a portion of the input power to be dissipated in the second port. At this level the patch efficiency drops 1.25 dB. We can reduce the monopole radiation by coupling to a second short-circuited probe to the patch instead of directly feeding it. The gap between the second probe and the patch is adjusted until the antenna radiates minimum cross-polarization in the H-plane. This uses the microstrip patch as the feed network, and the second probe has no resistive load to dissipate power.

The feed probe across the microstrip patch substrate is a series inductor at the input. Higher-order modes excited in the patch by this feeding method add to the inductive component of the antenna. Below resonance, the antenna is inductive and has near-zero resistance. As the frequency increases, the inductance and resistance grow as the parallel resonance is approached. Above the resonant frequency, the antenna is capacitive as the impedance sweeps clockwise around the Smith chart (Figure 6-11) and finally back to a slight inductive component near a short circuit. Increasing the

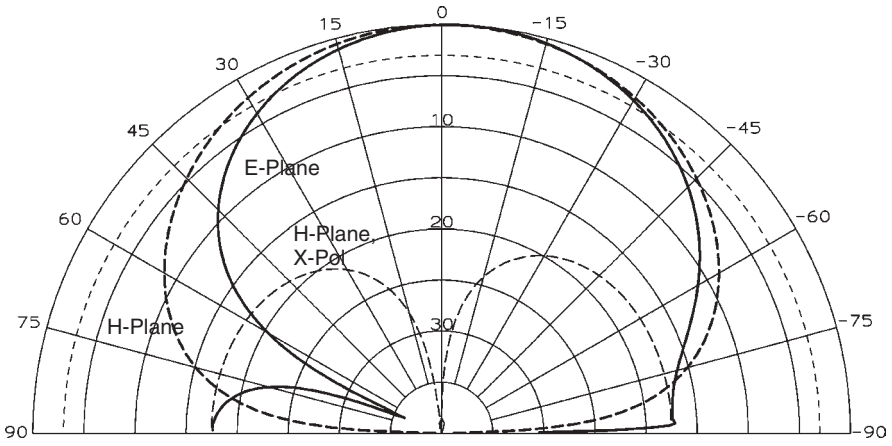


FIGURE 6-10 Pattern of coax-fed, microstrip patch including feed pin radiation for free-space substrate.

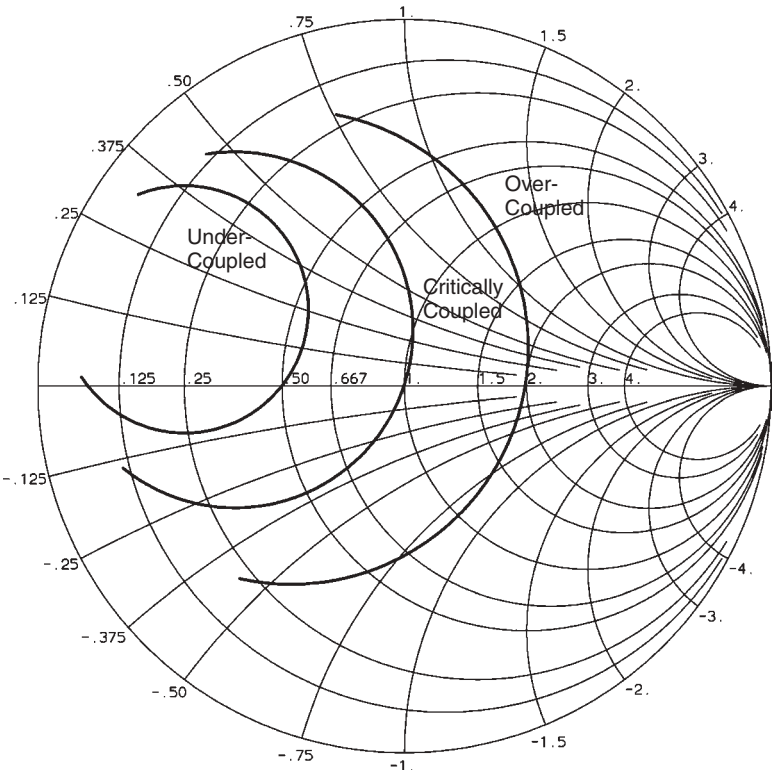


FIGURE 6-11 Smith chart frequency response of under-, critically, and overcoupled patches as the feed point moves toward one radiating edge of a rectangular patch.

input resistance by changing the feed point causes the resonant frequency response circle to grow on the Smith chart and cross the resistance line at a higher level. We call the left-hand curve the undercoupled case because the sweep of the curve fails to enclose the center of the chart. The center curve is critically coupled and the right curve is the overcoupled case. This general impedance response also holds for circular patches. We use these terms for all resonance curves when they sweep around or toward the Smith chart center from any peripheral point.

Figure 6-12 plots the Smith chart for a design with a patch on a 0.05λ-thick substrate with dielectric constant 1.1 that includes the inductance of the feed pin. The response locus lies above the real axis and is always inductive. We can tune this impedance locus by adding a series capacitor at the input with a reactance $-j50$ at the center frequency. The series capacitor moves the locus down until it sweeps around the center of the chart in an overcoupled response. Figure 6-13 shows implementation of the capacitor as a disk on the end of the feed pin. The pin passes through a hole in the patch so that the only connection is through the capacitor disk. The disk can be placed below the patch on a separate substrate in a multiple-layer construction. Other configurations use an annular ring capacitor etching in the patch at the feed point for small capacitors. Adding to this a series inductor and adjusting the series capacitor improves the impedance match over a larger frequency range, as shown in Figure 6-14, where the locus encircles the origin [12]. The patch with the single added

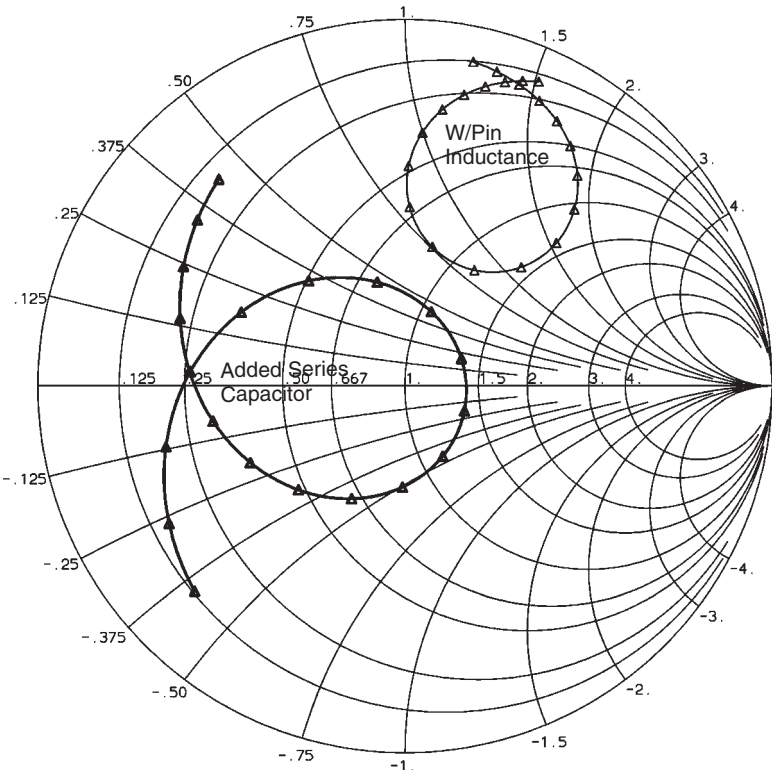


FIGURE 6-12 Impedance improvement by adding a series capacitor to a patch on a thick substrate.

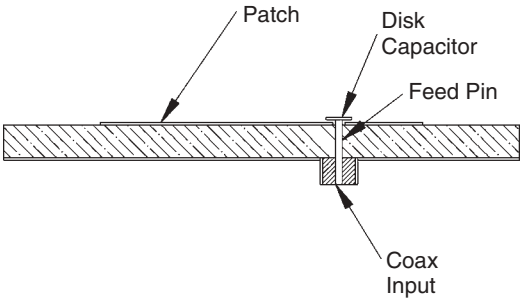


FIGURE 6-13 Cross section of a probe-fed patch with an added series capacitor.

series capacitor has a 9.1% 10-dB return-loss bandwidth while adjusting the series capacitor, and adding a series inductor increases the impedance bandwidth to 15.4%. Matching networks have limited ability to add resonances to broadband the impedance match, but construction becomes difficult. Later, we will obtain extra resonances by adding antenna elements.

We can feed patches from the edge by using an inset microstrip line as shown in Figure 6-15, where the gap on either side of the microstrip line equals its width. A FDTD analysis shows that the inset disturbs the transmission line or cavity model

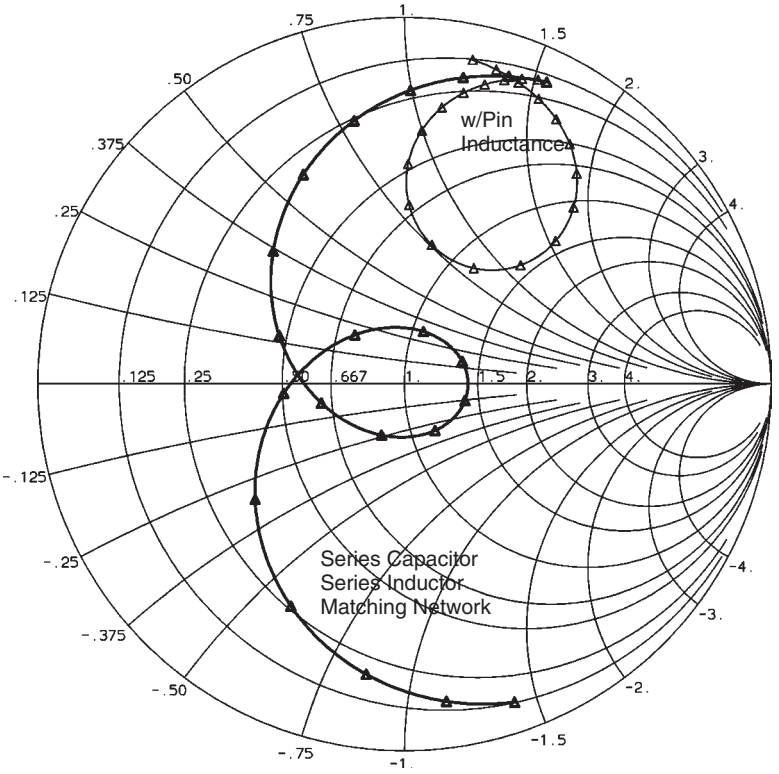


FIGURE 6-14 Impedance response of a patch with a two-element matching network.

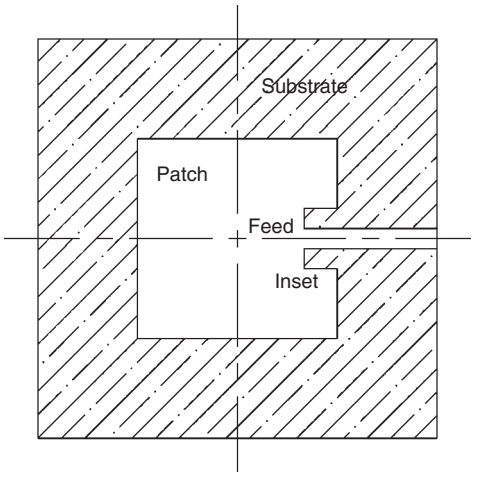


FIGURE 6-15 Inset-fed square patch.

and increases the impedance variation with distance compared to a coaxial probe feed given a patch resonant length L and feed position x from the center [13]:

$$R_i = R_e \sin^4 \frac{\pi x}{L} \qquad 0 \le x \le \frac{L}{2} \tag{6-24}$$

Equation (6-24) is an approximate solution because at $x = 0$, the resistance remains finite. We locate the feed from the equation using a radian angle measure:

$$x = \frac{L}{\pi} \sin^{-1} \left(\frac{R_i}{R_e} \right)^{1/4} \tag{6-25}$$

Compute the 50-Ω feed point in the example above:

$$x = \frac{30.64}{\pi} \sin^{-1} \left(\frac{50}{196} \right)^{0.25} = 7.71 \text{ mm}$$

The inset distance (7.3 mm) is less than the distance of the probe (9.8 mm) from the edge.

Aperture Feed [14,15] A microstrip patch is a planar resonant cavity with open-circuited sidewalls that leak power in radiation. We can also think of the rectangular patch operating in the lowest-order mode as a low-impedance transmission line with end susceptance and radiation conductance. Both models predict a resonant structure with significant Q. Resonant cavities are readily excited by coupling to a transmission line through an aperture or by direct feeding from a transmission line. The Q of the resonant cavity limits the excitation fields to one of the modes. We can expand the excitation in the cavity modes, but the lowest-order mode is usually the most significant and contains most of the stored energy. We generally consider the voltage distribution in a patch with its null plane located halfway across the patch through the center. Whether we consider it as a cavity or a transmission line the standing-wave voltage has a standing-wave current associated with it. This current is out of phase with the voltage and its peak occurs along the virtual short circuit through the centerline. Along the resonant length the current has a cosinusoidal distribution that vanishes at the radiating edges in a single half-cycle for the lowest-order mode. The current has a uniform distribution along the patch width.

We produce maximum coupling to a patch through a slot by distorting the currents in the ground plane of the patch where they are maximum in the center of the patch. To first order the currents flow along the resonant length. This means that we align the slot perpendicular to the current flow for maximum excitation in the same manner as slots in waveguides (Section 5-24). To excite the slot we pass a microstrip transmission line across it perpendicularly. This leads to a three-layer structure. The patch is located on the top layer. Its ground plane contains a coupling aperture usually placed under the center of the patch for maximum coupling. The third layer contains a microstrip transmission using the same ground plane as the patch and located under the center of the slot to maximize coupling. Figure 6-16*a* shows an exploded view of the patch, ground with its aperture, and the microstrip transmission line flipped over relative to the patch. Figure 6-16*b* gives the general parameters associated with the slot aperture. Although x_{os} and y_{os} are usually zero to maximize coupling, the patch current distribution tells us how the coupling varies with slot location. Because the current in the ground plane is uniform across the patch width W, coupling is independent of x_{os} until the slot starts to overlap the edge of the patch. The cosinusoidal distribution current distribution along the resonant length direction L means that the coupling falls off slowly as y_{os} is moved off zero. The sign of y_{os} does not matter because the distribution is an even function. The slow variation of current near the patch center means that the slot location has a loose tolerance.

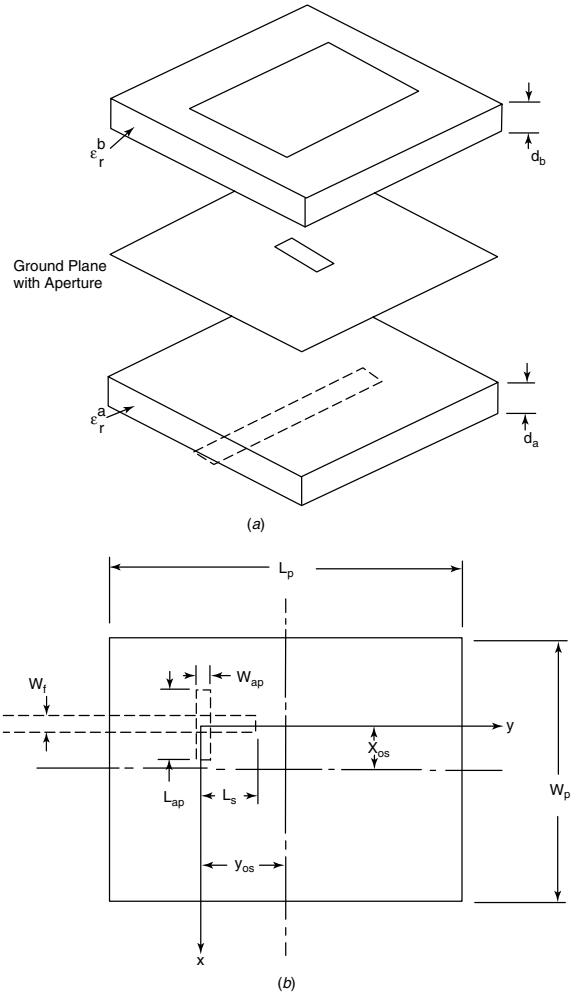


FIGURE 6-16 Aperture feed of square patch. (From [15], Fig. 1, © 1986 IEEE.)

The microstrip transmission line excites the slot (aperture) from a standing-wave with its maximum current located at the slot. We maximize the standing-wave current by either using a shorting via from the microstrip line to the ground plane or by using a quarter-wave open-circuited transmission line stub of length L_s. L_s will be less than a quarter-wave in the effective dielectric constant of the microstrip line because the open-circuit end has fringing capacitance and its capacitance must overcome the higher-order modes of the microstrip patch, which load the input inductively. The reactance of the stub, a series load to the input, is given by the equation

$$Z_S = -jZ_0 \cot(k_{\text{eff}} L_S)$$

where Z_0 is the characteristic impedance of the microstrip feed line, k_{eff} the effective propagation constant in the microstrip substrate, and L_s the stub length $\approx 0.22\lambda_{\text{eff}}$.

We increase the coupling to the patch resonant cavity by increasing the aperture size. Figure 6-17 shows the Smith chart variation with aperture size as the coupling

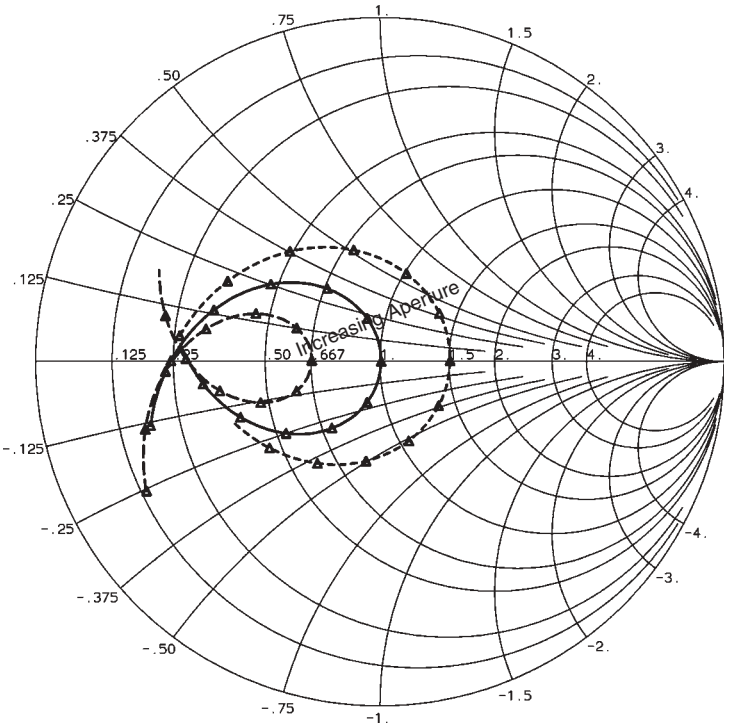


FIGURE 6-17 Effect of aperture size on coupling to a patch where larger openings move the response to the right.

varies left to right as undercoupled, critically coupled, and overcoupled. When we increase the bandwidth, we lower the Q, and the coupling aperture size must increase. Waterhouse [8] suggests starting with a slot about one-half the patch width and using a commercial code to analyze the response while adjusting dimensions before fabrication. We control the rotational position on the Smith chart by varying the open-circuited stub length. Shorter lengths, below $\lambda/4$, increase the capacitive reactance and the coupling loop will rotate around a constant-resistance circle with its diameter determined by the aperture size, as shown in Figure 6-18.

Figure 6-19 gives aperture shapes in order of increasing coupling. The longer slot of (*b*) compared to slot (*a*) increases coupling. Widening the aperture as in (*c*) increases coupling relative to (*a*). The H-shaped slot has a more uniform distribution along the horizontal slot and increased coupling. The bowtie and hourglass apertures increase coupling from a consideration of increased path length around the opening. The smooth curve of the hourglass reduces current discontinuities at the edges and increases coupling [16, pp. 158–159].

Aperture feeding eliminates the vertical pin structure in the microstrip patch and eases construction but at the cost of a multiple-layer etching. The elimination of the vertical pin removes the added monopole pattern, which increases cross-polarization. When the patch is edge fed, whether directly or inset, the substrate for good patch radiation does not match the one needed for good microstrip lines. With an aperture-fed patch, each structure can use its optimum substrate, because they are independent and connected only through the aperture. As we try to feed broadband patches, the

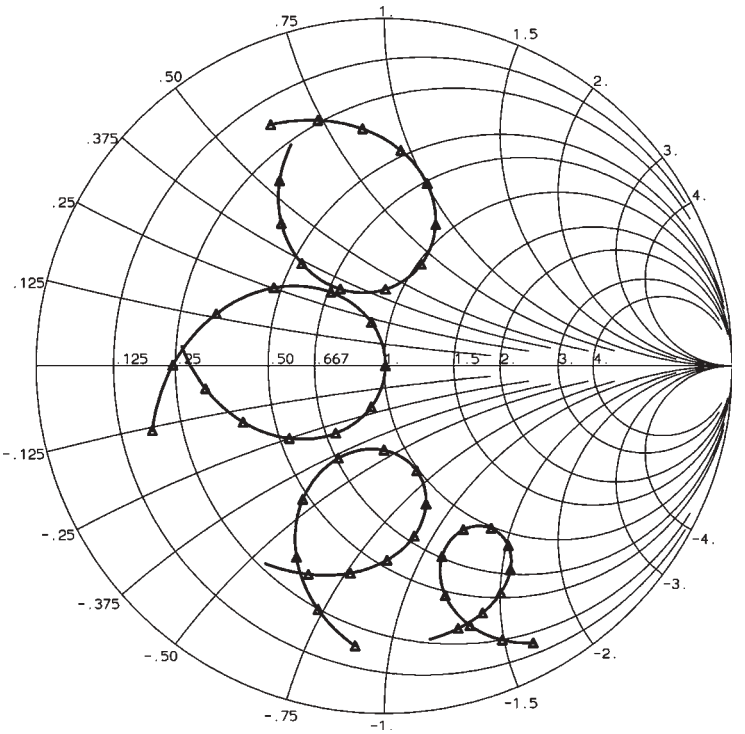


FIGURE 6-18 Effect of varying length of an open-circuited stub in an aperture-fed patch when critically coupled.

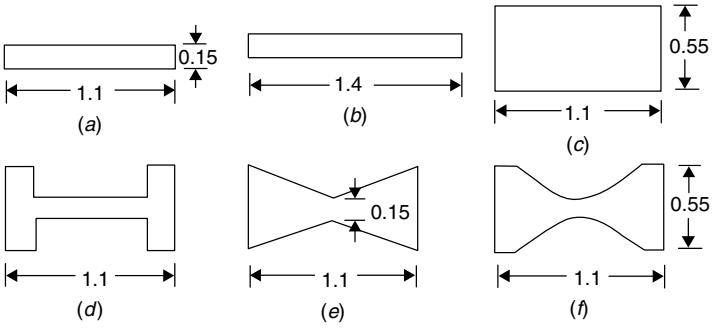


FIGURE 6-19 Aperture shapes to increase coupling and bandwidth. (From [16], Fig. 4-29, © 2003 Artech House, Inc.)

Q decreases and the aperture size grows. This slot, although below a resonant size, increases its radiation and decreases the front-to-back ratio because it radiates equally on both sides. One solution is to enclose the microstrip line in a box to prevent slot radiation on the back side. If we use a high-dielectric-constant substrate for the microstrip line, the coupling through the aperture remains high, but the second ground plane of the microstrip will reduce the coupling. The slot aperture adds a pole to the patch circuit that can be used to broadband the impedance response. To use this pole effectively, we must increase the aperture size until it becomes a significant radiator.

6-4 QUARTER-WAVE PATCH ANTENNA

When operation is in the lowest mode, a virtual short circuit forms through a plane centered between the two radiating edges. We can make an antenna by using half the patch and supplying the short circuit (Figure 6-20). The E-plane pattern broadens to that of a single slot. The resonant length is about a quarter-wavelength in the dielectric of the substrate. We use the effective dielectric constant ε_{eff} of a microstrip line of patch width W and Δ given by Eq. (6-18) to determine the resonant length L of the quarter-wave patch:

$$\frac{L}{2} = \frac{\lambda}{4\sqrt{\varepsilon_{\text{eff}}}} - \Delta \tag{6-26}$$

We can implement the short circuit with a series of pins or etched vias between the ground plane and the patch. These add an inductive component to the transmission-line model of the antenna. The effective shorting plane occurs further along the transmission line. The equivalent extra length Δl is found from the parallel-plate circuit model of a row of evenly spaced pins [17, p. 104]. Given the pin center spacing S, their radius r, and the wavelength in the dielectric $\lambda_d = \lambda_0/\sqrt{\varepsilon_r}$, we compute the patch-length reduction from the equation

$$\Delta l = \frac{S}{2\pi}\left[\ln\frac{S}{2\pi r} - \left(\frac{2\pi r}{S}\right)^2 + 0.601\left(\frac{S}{\lambda_d}\right)^2\right] \tag{6-27}$$

We have only the conductance and susceptance of a single edge that doubles the resonant resistance at the edge as compared with the full patch. It becomes difficult to feed the antenna from microstrip because this raises the quarter-wavelength transformer impedance and requires narrower lines. We can increase the edge width to reduce the edge input resistance, but the antenna is usually fed from underneath. Equation (6-23) gives the approximate feed location measured from the short circuit. The resonant frequency shifts slightly as the feed point moves. During tuning for impedance match, the length of the cavity will have to be adjusted to maintain the desired resonant frequency. Quarter-wave and full-patch antennas have the same Q. A half-patch antenna has half

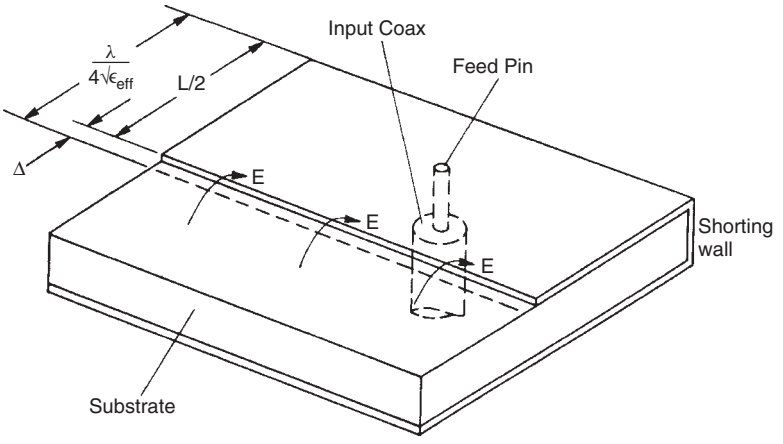


FIGURE 6-20 Quarter-wave patch.

the radiation conductance but only half the stored energy of a full-patch antenna. Its bandwidth is approximately the same as that of the full patch.

Example Design a half-patch antenna at 5 GHz on a 0.8-mm-thick substrate ($\varepsilon_r = 2.21$) with a radiating width of 0.75λ.

The edge width is $0.75(300\,\text{mm})/5 = 45\,\text{mm}$. By using Eq. (6-19), we compute the effective dielectric constant in the cavity: $\varepsilon_{\text{eff}} = 2.16$. Equation (6-18) gives us the cutback for fringing fields: $\Delta = 0.42\,\text{mm}$. The resonant length becomes

$$\frac{L}{2} = \frac{\lambda}{4\sqrt{\varepsilon_{\text{eff}}}} - \Delta = 10.20 - 0.42 = 9.78\,\text{mm}$$

The radiation conductance from the single edge is [Eq. (6-20)]

$$G = \frac{45}{120(60)} = 6.25\,\text{mS} \quad \text{or} \quad R = 160\ \Omega$$

The 50-Ω feed point is found from Eq. (6-23):

$$x = \frac{19.56}{\pi}\sin^{-1}\sqrt{\frac{50}{160}} = 3.69\,\text{mm}$$

where x is the distance from the short.

The short circuit of this antenna is quite critical. The low impedance of the microstrip cavity raises the currents in the short circuit. Without a good low-impedance short, the antenna will detune and have spurious radiation. If the antenna is made from a machined cavity, careful attention must be paid to the junction between the top plate and the cavity to assure good electrical contact.

Figure 6-21 shows the calculated pattern of a quarter-wave patch on a free-space substrate 0.04λ thick on an infinite ground plane. The antenna radiates primarily from the single edge located opposite the short-circuited edge. A vertical probe feeds the

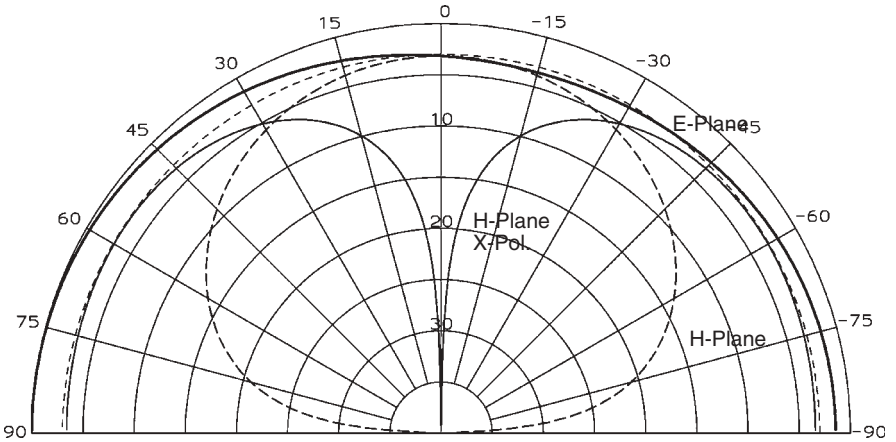


FIGURE 6-21 Pattern of a quarter-wave patch on a free-space substrate.

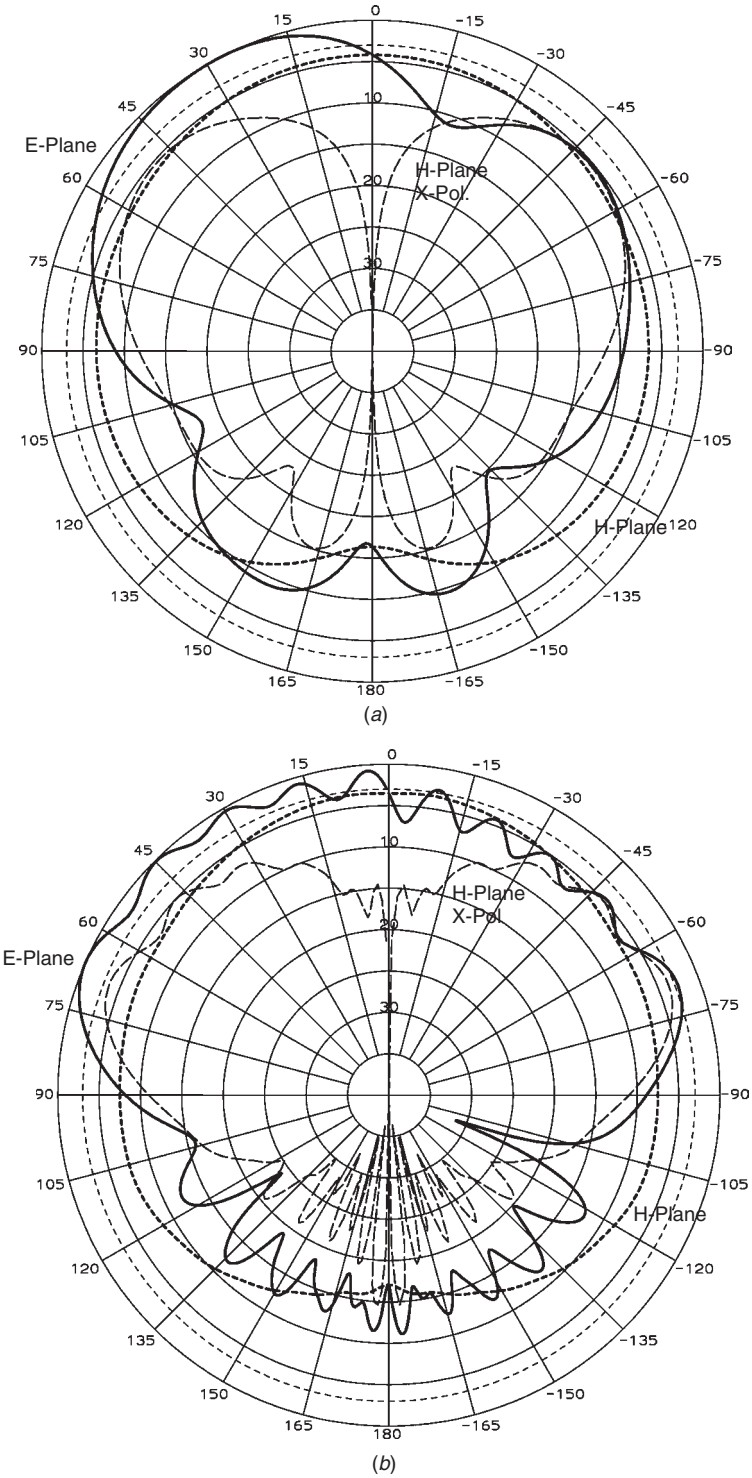


FIGURE 6-22 Pattern of a quarter-wave patch mounted on (*a*) 2λ- and (*b*) 10λ-diameter ground planes.

antenna directly. The E-plane has a broad, nearly constant pattern. Radiation from current on the probe and shorting pins adds to positive angle radiation of the equivalent single slot while subtracting in the opposite direction. The H-plane pattern (dashed curve) retains its null along the ground plane. The light-line curve gives the cross-polarization in the H-plane. The feed probe and shorting pin currents produce a pattern similar to that from a monopole. The model that uses equivalent magnetic currents fails to predict the high radiation from these currents. When the quarter-wave patch is mounted on a finite ground plane, it exhibits behavior similar to that of a monopole. Figure 6-22*a,b* plots the pattern when it is mounted on 2λ- and 10λ-diameter ground planes. These show a monopole-type pattern, where radiation spreads readily behind the ground plane. Currents flowing in the feed pin and shorting wall distort the E-plane and cause asymmetry. The magnetic currents flowing along the side slots no longer cancel as in the square patch and increase cross-polarization.

If we close off the nonradiating edges with metal walls, the walls convert the parallel-plate line into a waveguide and we use the waveguide propagation constant to calculate the quarter-wavelength cavity depth. The slot fields vanish on the ends and establish a sinusoidal slot distribution. We can offset the feed toward both the back wall and the sidewall to reduce the input impedance. The peak voltage (minimum current and peak resistance) occurs at the slot center. Figure 6-23 illustrates the pattern of the waveguide quarter-wave patch on an infinite ground plane. The sidewalls reduce the monopole radiation, and the H-plane cross-polarization is reduced compared to a quarter-wave patch. When mounted on a 2λ-diameter disk, centered on the feed pin, the pattern (Figure 6-24) exhibits lower-level radiation in the backlobe because the monopole pattern has been reduced. The high radiation level at the disk edges still causes considerable edge diffraction in the E-plane.

6-5 CIRCULAR MICROSTRIP PATCH

In some applications, a circular patch fits in the available space better than a rectangular one. In a triangularly spaced array, they maintain a more uniform element

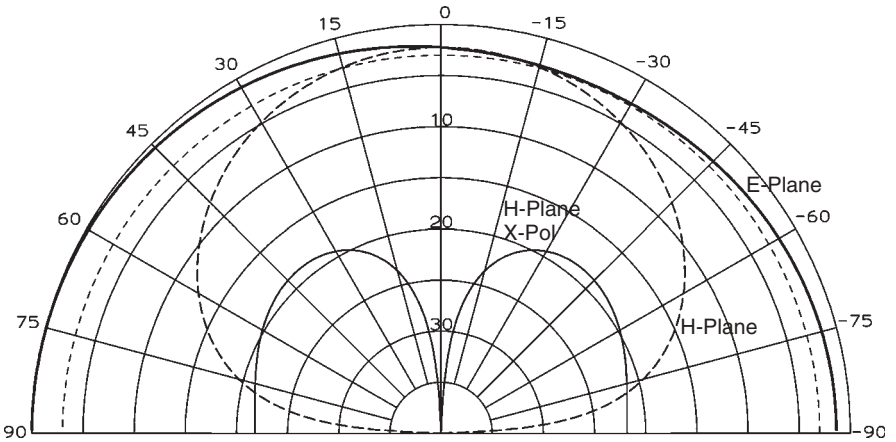


FIGURE 6-23 Pattern of a quarter-wave waveguide patch.

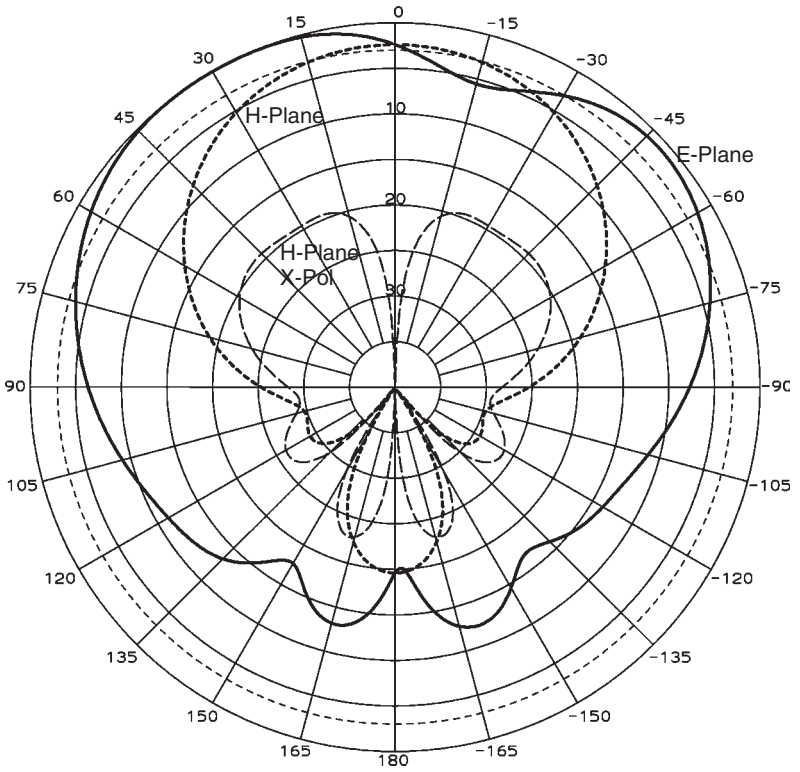


FIGURE 6-24 Quarter-wave waveguide patch mounted on a 2λ-diameter disk.

environment. No suitable transmission-line model presents itself, and the cavity model must determine the resonant frequency and bandwidth. The cutoff frequencies of TE modes of circular waveguides give the resonant frequencies of circular patch antennas. The patch with its magnetic walls and TM modes is the dual of the waveguide. The resonant frequencies are given by

$$f_{np} = \frac{X'_{np} c}{2\pi a_{\text{eff}} \sqrt{\varepsilon_r}} \tag{6-28}$$

where X'_{np} are the zeros of the derivative of the Bessel function $J_n(x)$ of order n, as is true of TE-mode circular waveguides. The term a_{eff} is an effective radius of the patch [18]:

$$a_{\text{eff}} = a\sqrt{1 + \frac{2H}{\pi a \varepsilon_r}\left(\ln\frac{\pi a}{2H} + 1.7726\right)} \tag{6-29}$$

where a is the physical radius and H is the substrate thickness. Using the effective radius gives the resonant frequency within 2.5%.

We combine Eqs. (6-28) and (6-29) to determine radius to give a particular resonant frequency:

$$a_{\text{eff}} = \frac{X'_{np} c}{2\pi f_{np} \sqrt{e_r}} \tag{6-30}$$

Since a and a_{eff} are nearly the same, we can iterate Eq. (6-29) to compute a, the physical radius [19, p. 119]:

$$a = \frac{a_{\text{eff}}}{\sqrt{1 + 2H/\pi a\varepsilon_r \left[\ln(\pi a/2H) + 1.7726\right]}} \tag{6-31}$$

We start by using a_{eff} for a in Eq. (6-31), which converges rapidly. The lowest-order mode, TM_{11}, uses X'_{11} (1.84118) and produces a linearly polarized field similar to a square patch. The TM_{01} mode ($X'_{01} = 3.83171$) produces a monopole-type pattern from a uniform edge fringing field.

Example Design a circular microstrip patch antenna (TM_{11} mode) at 3 GHz on a 1.6-mm substrate that has a dielectric constant of 2.55 (woven Teflon fiberglass).

We calculate the effective radius from Eq. (6-30):

$$a_{\text{eff}} = \frac{1.84118(300 \times 10^9 \text{ mm/s})}{2\pi(3 \times 10^9 \text{ Hz})\sqrt{2.55}} = 18.35 \text{ mm}$$

The physical radius will be slightly less. By using a_{eff} in the denominator of Eq. (6-31), we obtain a physical radius: $a = 17.48$ mm. We can then substitute this back into Eq. (6-31) and obtain $a = 17.45$ mm. Equation (6-31) converges in two iterations to a reasonable tolerance, since another iteration gives the same value. Actually, a single iteration gives the value within 0.2% on a formula accurate to only 2.5%.

The fields of the TM_{11} mode produce a virtual short circuit at the center of the patch. We can reinforce the short circuit with a pin soldered between the patch and ground. The radial line along which the feed is placed determines the direction of the linear polarization. The nonuniform radiation along its edge gives a larger edge impedance than the square patch. Experience shows that the 50-Ω feed point is located from the center at about one-third the radius. Experiments, actual or numerical, will be required to locate the proper point. Use a network analyzer with a Smith chart display to measure the input impedance. If the resonance circle swings around the origin, the impedance is too high (overcoupled). Move the feed toward the center. A scalar return-loss display cannot give you the direction of movement required. Like the rectangular patch, mismatching the impedance at center frequency to about 65Ω increases the bandwidth slightly. Derneryd [20] gives an approximate expression for the radial impedance variation:

$$R_{\text{in}} = R_e \frac{J_1^2(k_\varepsilon \rho)}{J_1^2(k_\varepsilon a)} \tag{6-32}$$

where R_e is the edge resistance, ρ the radial distance, and J_1 the Bessel function of the first kind. k_ε is the propagation constant in the substrate dielectric constant: $k_\varepsilon = k\sqrt{\varepsilon_r}$. Figure 6-25 gives the 2 : 1 VSWR bandwidth of a circular patch on various substrates as a function of the substrate thickness. It has a slightly smaller bandwidth than that of a square patch because it has a smaller volume. The curves on Figure 6-25 include surface-wave radiation (or losses).

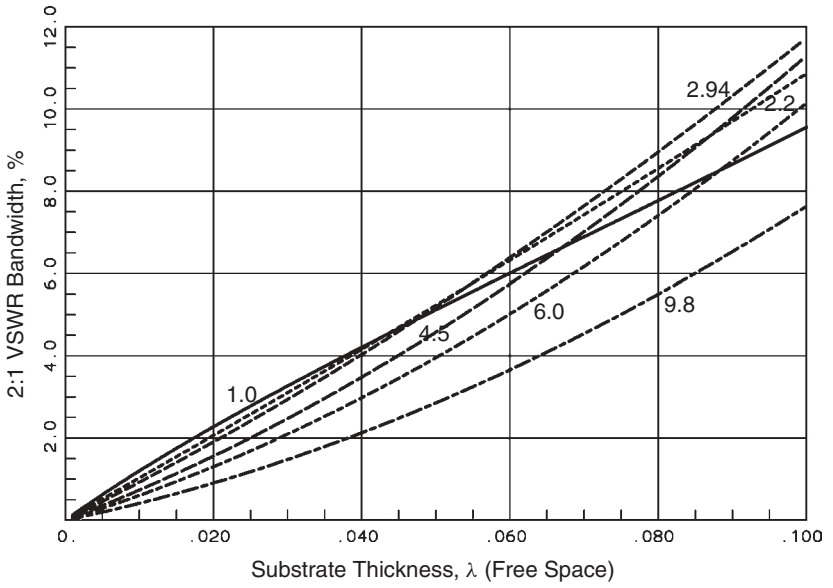


FIGURE 6-25 2 : 1 VSWR bandwidth of circular microstrip patches versus substrate thickness in free-space wavelengths, including surface-wave radiation.

6-6 CIRCULARLY POLARIZED PATCH ANTENNAS

Figure 6-26 show methods of achieving circular polarization with square patches fed with two inputs. The patches are fed by equal signals 90° out of phase. The branchline hybrid (Figure 6-26*a*) consists of four transmission lines connected in a square. The hybrid shown (100-Ω system) produces equal outputs 90° out of phase at center frequency. The two inputs produce patterns with opposite senses of circular polarization. Both the VSWR and axial ratio bandwidths far exceed the singly fed patch bandwidth. Reflections due to the patch mismatch are routed to the opposite input. Patch input reflections, undetected at the input, reduce the efficiency of the antenna by the same amount as the singly fed patch mismatches. The antenna can be fed from below in two places by using a coupled line hybrid, but it suffers from the same efficiency problem.

The cross-fed antenna (Figure 6-26*b*) splits the signal to feed both edges. A quarter-wavelength-longer line provides the extra 90° phase shift to give circular polarization. Shifting the impedance from one input through a quarter-wavelength line before adding the two in shunt cancels some of the reflection from the second line and increases the impedance bandwidth. The impedance bandwidth approximately doubles compared to the singly fed patch. The 6-dB axial ratio bandwidth roughly equals the singly fed square-patch bandwidth. The polarization loss (0.5 dB) of a 6-dB axial ratio equals the 2 : 1 VSWR mismatch loss.

The antennas in Figure 6-27 use asymmetries to perturb the resonance frequencies of two possible modes and achieve circular polarization [21]. The approximately square patches have been divided into two groups: type A, fed along the centerline, and type B, fed along the diagonal. All these antennas radiate RHC. We can understand the operation of these patches from an analysis of the turnstile dipole antenna (Figure 6-28). The orthogonal dipoles could be of equal length and fed from a 90° hybrid to achieve

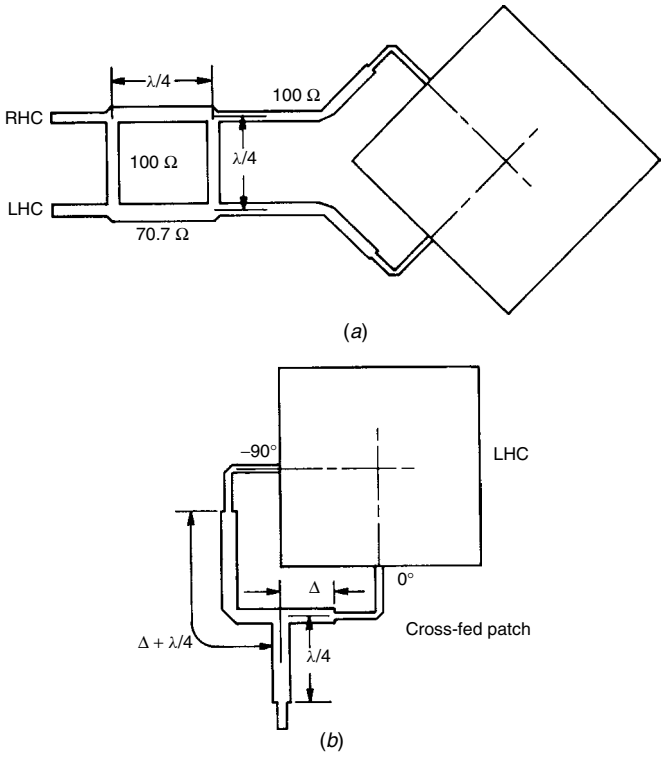


FIGURE 6-26 Dual-fed circularly polarized patch antennas: (*a*) branchline hybrid fed; (*b*) cross-fed patch.

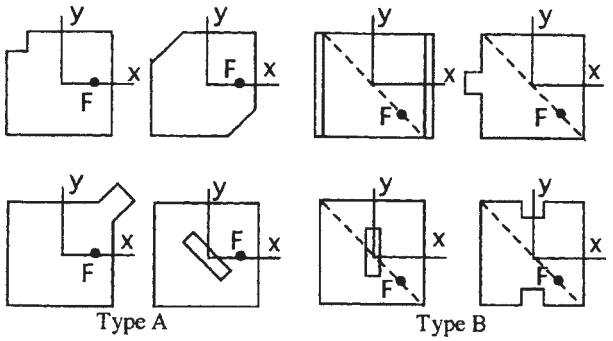


FIGURE 6-27 Classes of perturbed microstrip patches to generate circular polarization from a single feed. (From R. Garg et al., *Microstrip Patch Handbook*, Fig. 8-15, © 1999 Artech House, Inc.)

circular polarization (like the patch in Figure 6-26*a*). Instead, the lengths are changed to shift the phase of each dipole by 45° at resonance. If we lengthen the dipole beyond resonance, the input impedance becomes inductive. The current becomes

$$I = \frac{V}{R_2 + jX_2} = \frac{V(R_2 - jX_2)}{R_2^2 + X_2^2}$$

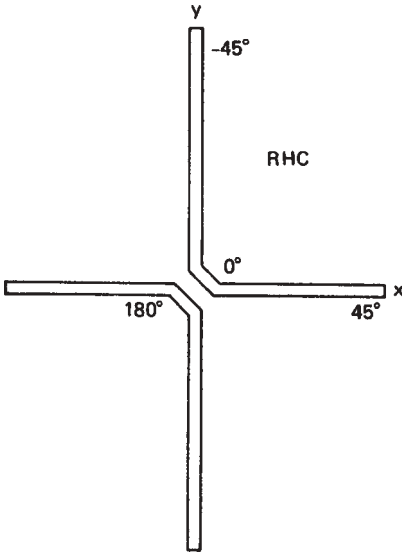


FIGURE 6-28 Turnstile dipole antenna.

The radiated field phase decreases relative to the resonant-length dipole. Shortening the dipole from resonance increases the far-field phase. We adjust the lengths until the phase difference of the radiated fields is 90° and the susceptances from the two dipoles cancel at center frequency. The combination of the two modes produces a Smith chart response with a small loop or kink (see Figure 5-13). The best circular polarization occurs at the frequency of the kink, and the response degrades below and above this frequency. The axial ratio bandwidth is far less than the impedance bandwidth, because the combination of the two modes causes a cancellation of transmission-line reflections from the two modes and increases the impedance bandwidth. The phase required for good circular polarization changes rapidly.

We denote the total change in area ΔS to achieve two resonances for a normal patch area of S and it is proportional to the Q. A type A patch, fed along the square patch axis, requires less area change than a type B patch, fed along the diagonal:

$$\text{type A:}\frac{\Delta S}{S} = \frac{1}{2Q} \qquad \text{type B:}\frac{\Delta S}{S} = \frac{1}{Q} \tag{6-33a,b}$$

We achieve the same effect with a patch by perturbing the lengths of a square patch and feeding both polarizations. An input along the diagonal (type B) feeds all edges in two separate resonances. The ratio of the edge lengths is found in terms of Q by a perturbation technique [4]. We rearrange Eq. (6-33*b*) to derive the ratio of these lengths:

$$\frac{b}{a} = 1 + \frac{1}{Q} \tag{6-34}$$

We calculate resonant frequencies for the two lengths from Eq. (6-34):

$$f_1 = \frac{f_0}{\sqrt{1 + 1/Q}} \qquad f_2 = f_0\sqrt{1 + \frac{1}{Q}} \tag{6-35}$$

Q is related to the VSWR bandwidth by Eq. (6-7). The 3-dB axial ratio bandwidth of the antenna is limited to $35\%/Q$ or 35% of the frequency difference between f_1 and f_2.

Example Compute resonant lengths for a corner-fed patch on a 1.6-mm substrate with $\varepsilon_r = 2.55$ at 3 GHz.

We have $\lambda = 100$ mm and thickness/$\lambda = 0.016$. From Figure 6-7 we read the 2 : 1 VSWR bandwidth: 1.61%. From Eq. (6-7) we calculate Q:

$$Q = \frac{1}{0.0161\sqrt{2}} = 43.9$$

We use Eq. (6-35) to determine the resonant frequencies:

$$f_1 = \frac{3}{\sqrt{1 + 1/43.9}} = 2.966\,\text{GHz}$$

$$f_2 = 3\sqrt{1 + 1/43.9} = 3.034\,\text{GHz}$$

By using the techniques of Section 6-3, we calculate the resonant lengths: $a = 30.27$ mm, $b = 31.01$ mm.

All perturbations by small areas in a circular patch can only be type A feeding. The perturbation equations are related to the circular patch separation constant X'_{11} (1.84118):

$$\text{type A:} \frac{\Delta S}{S} = \frac{1}{X'_{11} Q} \qquad \text{type B:} \frac{\Delta S}{S} = \frac{2}{X'_{11} Q} \tag{6-36}$$

A circular patch perturbed into an elliptical patch radiates circular polarization when fed on a 45° diagonal from the major or minor axis and produces type B feeding. The ratio of major to minor axes is related to Q [4]:

$$\frac{b}{a} = 1 + \frac{1.0887}{Q}$$

with resonant frequencies

$$f_1 = \frac{f_0}{\sqrt{1 + 1.0887/Q}} \quad \text{and} \quad f_2 = f_0\sqrt{1 + \frac{1.0887}{Q}} \tag{6-37}$$

We compute Q by using Eq. (6-7) and read the bandwidths from Figure 6-25 for circular patches. Use the techniques of Section 6-5 to calculate the physical radius of the major and minor axes from the frequencies [Eq. (6-35)].

6-7 COMPACT PATCHES

The desire to produce small patches for cellular telephone handset use has lead to the development of compact designs. The ideal antenna is one whose location the user is unaware of and which is as small as possible. Because most signals arrive at the user

after many bounces and edge diffractions, polarization is arbitrary. We do not need to control the radiation pattern or its polarization carefully and it opens up a range of possibilities. Shorting pins placed close to the feed pin reduce the patch size to about $\frac{1}{8}\lambda$ on a side, but its polarization is poorly controlled. If we can force the current to take a longer path along the resonant-length path, we can shrink the overall size. We etch notches in the patch to make the current wander or use various spiral-wound networks on a flat substrate. Three-dimensional solutions consist of folding a patch by using the vertical direction or some sort of winding around a cylinder. Many variations on these ideas appear in the literature and in collections of these ideas [16,22,23].

Adding a shorting pin closely spaced to the vertical feed pin (Figure 6-29) greatly reduces the resonant frequency of a given-size patch and produces a compact patch [24]. The idea is to make the current flow over a longer path from the feed point to the radiation site; in other words, the transmission line has been folded to make the path longer in the resonant cavity. We use this concept for all compact patches. In this configuration the resonant wavelength is found from the patch perimeter. Given the width W and the length L of the patch on a dielectric substrate ε_r, the resonant wavelength is given by

$$\lambda_0 = 4\sqrt{\varepsilon_r}(L + W) \tag{6-38}$$

which reduces to a square patch $\lambda/8$ on a side. This patch has one-half the length and one-fourth the area of a quarter-wave path, with its short circuit along an entire edge. The circular shorting pin compact patch resonant diameter equals $0.14\lambda_0/\sqrt{\varepsilon_r}$. Making a patch this small produces highly inductive input impedance, which we can see by looking at the Smith chart of a coaxial probe-fed patch (Figure 6-12). The curve sweeps clockwise as the frequency increases. At low frequency (or small size) the patch is highly inductive. Figure 6-12 shows that using thicker substrates to increase bandwidth makes the patch impedance even more inductive. The shorting pin next to the feed pin forms a transmission line with it and adds a capacitive component to the input impedance that counteracts the patch and feed pin inductance. As the shorting pin is moved farther away from the feed pin, the capacitance decreases and the shorting pin becomes an inductive component, as it is in the quarter-wave patch.

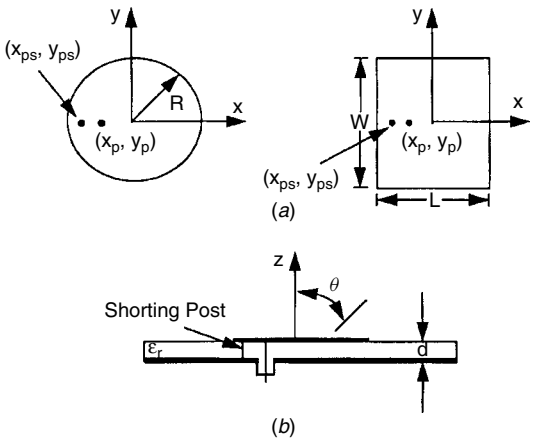


FIGURE 6-29 Compact patch with a shorting pin near the feed. (From [24], Fig. 1, © 1998 IEEE.)

The recommended position of the shorting pin is 80 to 90% of the distance from the center to the outer edge and a diameter of 0.008λ. You will need to iterate the position of the feed probe, usually one-half the diameter of the shorting pin, to achieve an impedance match. Table 6-5 lists the bandwidth achieved versus substrate thickness on foam, $\varepsilon_r = 1.07$ [25, p. 207].

Figure 6-30 gives the calculated pattern of a shorting pin compact patch on 0.034λ free-space substrate. The broad E-plane pattern has 10-dB dip on the broadside matched by the H-plane E_ϕ component. The large current in the shorting pin produces a significant monopole pattern seen in the E_θ radiation in the H-plane. This small antenna is a combination of a top-loaded monopole and a patch. Thinner antennas have a lower pattern dip broadside to the substrate because the monopole is shorter.

The planar inverted F antenna (PIFA) is similar electrically to the shorting pin compact patch. We move the shorting pin to one corner and often make it a small shorting plate. We locate the feed pin close to the small shorting plate to again form a transmission line whose capacitance with the feed pin counteracts the inductive component of the small patch. We use Eq. (6-38) to determine its resonant wavelength. If we rotate the coordinates so that the shorting plate and diagonal lie on the x-axis, we obtain the pattern response of Figure 6-30. Since there is practically no difference between the two antennas, Table 6-5 gives the bandwidth of the PIFA versus thickness [26].

TABLE 6-5 Bandwidth of a Single Shorting Pin Compact Patch

Thickness (λ_0)	Bandwidth, 2 : 1 VSWR (%)	Feed-to-Pin Center Distance (λ_0)
0.01	1.6	0.0071
0.02	2.2	0.0076
0.03	2.7	0.0081
0.04	3.4	0.0085
0.05	4.3	0.0101
0.06	5.7	0.0135

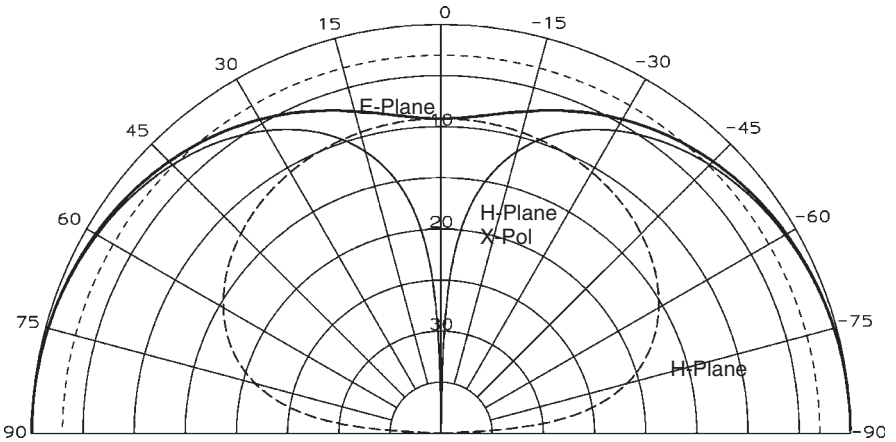


FIGURE 6-30 Pattern of a compact patch.

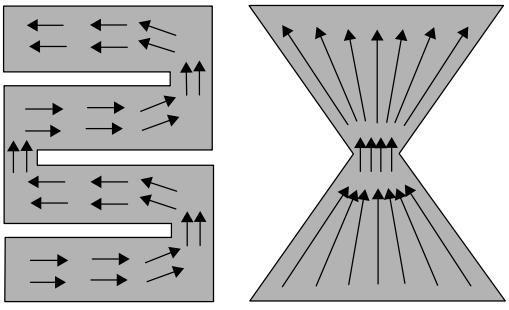


FIGURE 6-31 Reduced-size microstrip patches using meandered current paths. (From [22], Fig. 1-3, © 2002 John Wiley & Sons, Inc.)

Modest size reduction can be obtained by making the currents flow along a longer path along the resonant length. Figure 6-31 shows two planar antennas where slits cut from the width sides and disrupting the resonant-length path cause wandering of the current. The bowtie patch also makes the current path longer. These antennas radiate normal patch patterns with broader beamwidths in the E-plane because the notches bring the radiating edges closer together. The antennas in Figure 6-32 shrink the resonant length by folding the antennas vertically. The total length along the path is approximately $\lambda/2$, but the radiating edges are closer together. A large number of variations using slots have been investigated and offer interesting approaches to both shrink the patch size and produce dual-frequency antennas by using both the patch mode and slot radiation [22].

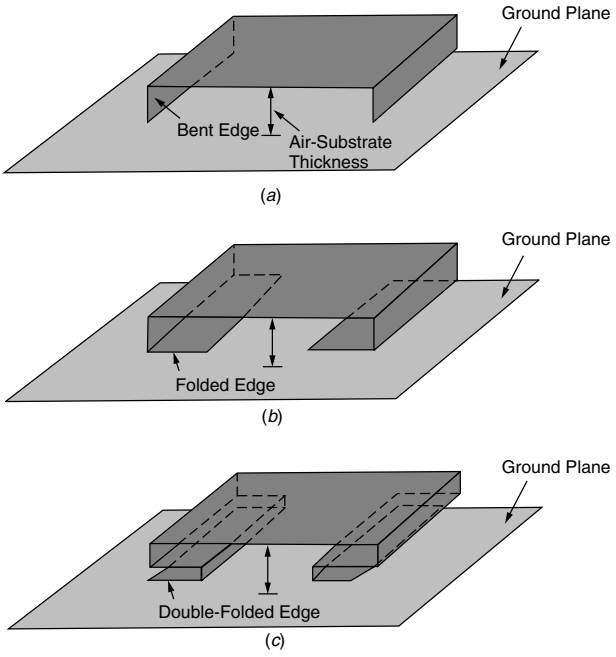


FIGURE 6-32 Folding microstrip patches to reduce size. (From [22], Fig. 1-4, © 2002 John Wiley & Sons, Inc.)

6-8 DIRECTLY FED STACKED PATCHES

Figure 6-7 illustrates the limited impedance bandwidth achievable from a single-resonator microstrip patch. When we increase the substrate thickness to widen the bandwidth, the antenna excites more surface waves (Figure 6-8) difficult to control, and we accept them as losses. In Section 6-3 we discussed the use of external circuit elements to improve the impedance response. These have limited usefulness, although the simple series capacitor input to overcome the inductance of a long feed probe and the inductive nature of the higher-order modes is easily implemented. These external elements add poles to the resonant circuit to increase bandwidth. We can increase the number of poles by adding antenna elements instead. One solution is to couple to additional patches located around the fed patch on the same substrate surface. This increases the antenna size and reduces pattern beamwidths. This solution is difficult to use in an array because the large spacing between elements produces grating lobes. Stacking patches vertically above the driven patch and coupling to them electromagnetically produces the best solution in terms of pattern response. The disadvantage of this approach is the additional fabrication cost. Our discussion of aperture-coupled patches in Section 6-3 points out that large apertures also add resonant poles that can increase the bandwidth. These added resonant elements complicate the design and call for the application of analytical tools instead of a cut-and-try approach.

Although either patch in a two-element stacked patch design can be fed, feeding the lower element produces a design with minimum feed pin inductance. Aperture coupling through the ground plane feeds the lower patch directly as well. If we use an edge feed, we want the input transmission line to be as narrow as possible to reduce radiation by feeding the lower patch. Initially, we consider the probe-fed stacked patch [27]. A coaxial probe feeds directly a lower substrate of thickness d_1 and dielectric constant ε_{r1} through a hole in the ground plane. Figure 6-12 shows that the feed probe adds inductance for a thick substrate and the resonant loop is located on the upper inductive portion of the Smith chart. When we couple the lower patch to an upper patch with thickness d_2 and dielectric constant ε_{r2}, its circuit response becomes more inductive. We need to start with the impedance locus of the lower patch to be capacitive without the upper patch. This can be achieved by using an overcoupled feed. Figure 6-11 illustrates the overcoupled patch whose impedance locus sweeps around the origin of the Smith chart. The inductance of the feed probe rotates these curves clockwise around the center of the chart and the overcoupled response has significant capacitive reactance when it sweeps around the origin. If we matched the lower patch critically, upward movement of the locus due to the coupled patch would reduce the impedance bandwidth. Figure 6-33 illustrates these design steps. Figure 6-33*b* shows that increasing the lower patch thickness leads to a longer feed probe that sweeps across the center from a more inductive portion of the Smith chart. Adding the second patch fails to increase the bandwidth relative to the thinner optimum lower patch. The thickness of the upper patch substrate d_2 controls the tightness of the resonant loop. A greater thickness d_2 produces a tighter loop in the Smith chart response that leads to a lower VSWR over a narrower bandwidth. Remember that we cannot use Eq. (6-7) to determine the bandwidth for different VSWR levels because we now have multiple resonators.

If we use a foam upper substrate, the dielectric constant and thickness of the lower substrate determines the surface-wave efficiency. Waterhouse [27] used a dielectric

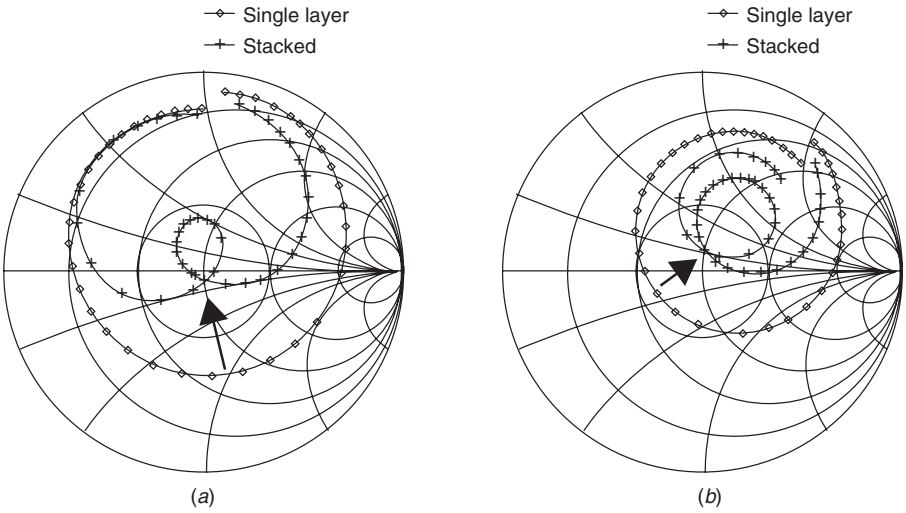


FIGURE 6-33 Effect of coupling to a second patch: (*a*) overcoupled single lower patch response forms resonant loop with the second coupled patch; (*b*) increasing lower patch thickness causes rotation on a Smith chart and lower bandwidth. (From [27], Fig. 3, © 1999 IEEE.)

constant of 2.2 for the lower substrate with a thickness of $0.04\lambda_0$ and a foam upper substrate $0.06\lambda_0$ to achieve optimum bandwidth with acceptable surface-wave losses. The lower patch was overcoupled so that it swept through the 250-Ω resistance point at resonance. Since the impedance locus sweeps clockwise on the Smith chart as frequency increases, this resonant point should be slightly below the lower end of the desired frequency band. We adjust the second substrate thickness to move the resonant loop on the Smith chart in the vertical direction. As we increase the size of the upper patch, the loop moves around an arc in the clockwise direction, which we use to center the impedance response on the Smith chart for optimum bandwidth. This method produces impedance bandwidths of around 25%. The pattern bandwidth exceeds this bandwidth and we expect little change in pattern over this frequency range.

Another successful stacked patch fed from a coaxial probe is the *hi-lo configuration*, in which a high dielectric substrate ($\varepsilon_{r1} = 10.4$) is used for the lower substrate and a foam ($\varepsilon_{r2} = 1.07$) for the upper substrate [25, pp. 178–182]. The upper patch captures the surface wave of the lower patch and greatly improves the overall efficiency by radiating this power in a space wave. Although the two patches have different sizes, the coupling remains sufficient to produce a broadband antenna with impedance bandwidths approaching 30%. In this design the lower patch is designed for the high dielectric of the lower substrate with little consideration for the upper patch except for making it a little overcoupled. The upper patch can be designed using the substrate thickness and dielectric constant assuming that the high dielectric substrate acts as the ground plane. When we mount the upper patch over the smaller lower patch, small adjustments must be made to the dimensions to achieve a 50-Ω impedance match. The example given used a lower substrate thickness of $0.032\lambda_0$ with $\varepsilon_{r1} = 10.4$, and by Figure 6-8 would have −1.3-dB surface-wave loss. Locating the second patch on a $0.067\lambda_0$-thick foam substrate directly over the first patch reduced the surface-wave loss to better than −0.7 dB over the entire band.

6-9 APERTURE-COUPLED STACKED PATCHES

The discussion on aperture feeding of a patch in Section 6-3 stated that we can utilize the aperture as another resonator to broadband the antenna. Figure 6-34 shows the stacked patch antenna fed from an aperture. In this implementation we make the coupling slot long enough to be one of the resonators, which increases the number of resonators to three: the aperture, the lower patch, and the upper patch. We must use element spacing to control coupling because frequencies control resonator sizes. By careful control of parameters two loops will form in the Smith chart response of impedance and be made to wrap tightly around the center of the chart [28] as shown in Figure 6-35*b*. We form these loops by coupling resonators. Undercoupling produces small tight loops; overcoupling produces large loops.

Figure 6-35 illustrates the effect of aperture size. The left Smith chart shows undercoupling between the aperture and the lower patch by the small left loop. We increase the coupling by increasing the aperture slot length (Figure 6-35*b*) or by increasing the

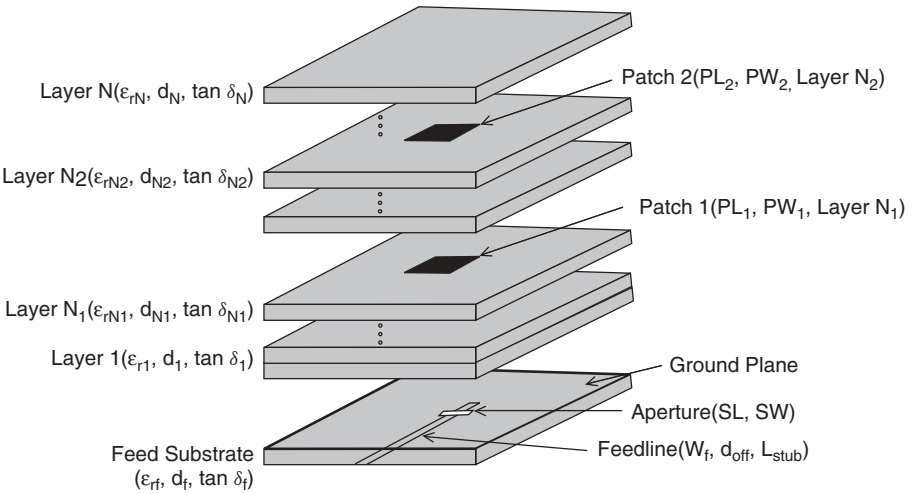


FIGURE 6-34 Construction of a resonant aperture coupled dual patch in exploded view. (From [28], Fig. 1, © 1998 IEEE.)

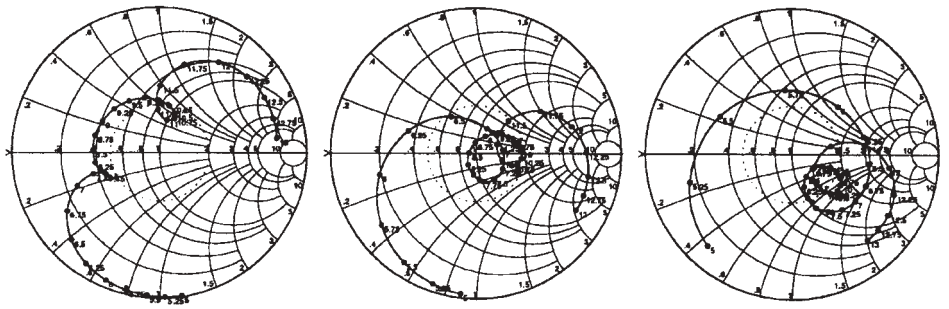


FIGURE 6-35 Effect on increasing slot length, SL, of apertures in stacked dual patches: (*a*) SL = 8 mm; (*b*) SL = 10 mm; (*c*) SL = 12 mm. (From [28], Fig. 4, © 1998 IEEE.)

lower patch size or reducing the lower patch thickness. The best results have been obtained by having the lower-frequency (left) loop determined by the lower patch and aperture. Overcoupling the aperture to the lower patch produces the impedance locus of Figure 6-35*c*. We control the upper loop size by varying the upper patch size, the relative size between the two patches, and the upper substrate thickness. The lower patch size is a critical parameter because it affects the coupling and size of both loops while shifting their center frequencies. For fixed sizes of the other two resonators, decreasing the lower patch size decreases the coupling to the aperture while increasing the coupling to the upper patch. Increasing aperture size increases coupling to the aperture and decreases coupling to the upper patch. By remembering that overcoupling produces larger Smith chart loops, we determine in which direction to change parameters by observing changes in analytical results on the Smith chart to produce optimum designs.

Because the slot aperture is one of the three resonators, we cannot vary its length to determine coupling to the lower patch. The overcoupled large slot produces high resistance at the microstrip input. We can lower this impedance by offset feeding the slot or by using a wide transmission line. A single offset line will unbalance the fields in the slot and lead to unbalanced excitation of the patches. This unbalanced excitation on the patches increases cross-polarization. The dual balanced offset feeding shown in Figure 6-36, where we join the two lines in a reactive power divider, both lowers the resistance and balances the patch excitation.

A design using rectangular patches for a single linear polarization achieved a 67% 2 : 1 VSWR bandwidth [28]. The only significant problem with the design is the poor front-to-back ratio, which is reduced to 6 dB at the upper frequencies as the aperture radiation increases. Placing a reflector patch below the microstrip feed line, it can be sized to reduce the F/B ratio by forming a Yagi–Uda antenna with the stacked patch [29]. Figure 6-37 illustrates an exploded view of a dual polarized aperture stacked patch. The potential bandwidth shrinks because we lose width as a parameter with square patches to optimize impedance. The key element of this design is the feed crossed slot [30]. The crossed-slot feeding aperture is located on a ground plane shared by microstrip networks located below and above the aperture. Each network consists of a reactive power divider to raise the impedance of the feed lines and allow offset feeding of the slot for each polarization. The balanced feed reduces cross-polarization and cross coupling between the two ports that would occur in both the crossed slot

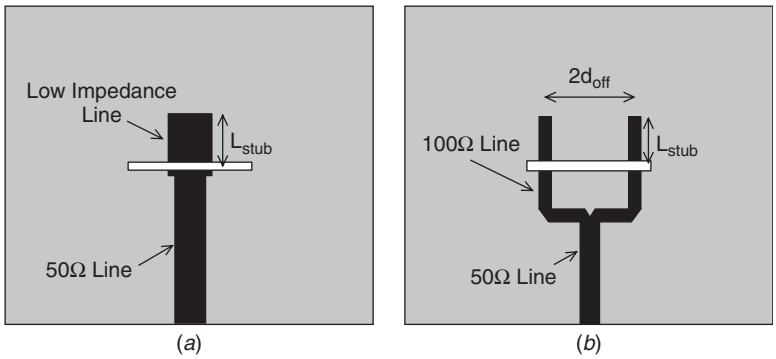


FIGURE 6-36 Impedance matching for resonant aperture dual stacked patches: (*a*) wide transmission line; (*b*) dual offset feed. (From [28], Fig. 3, © 1998 IEEE.)

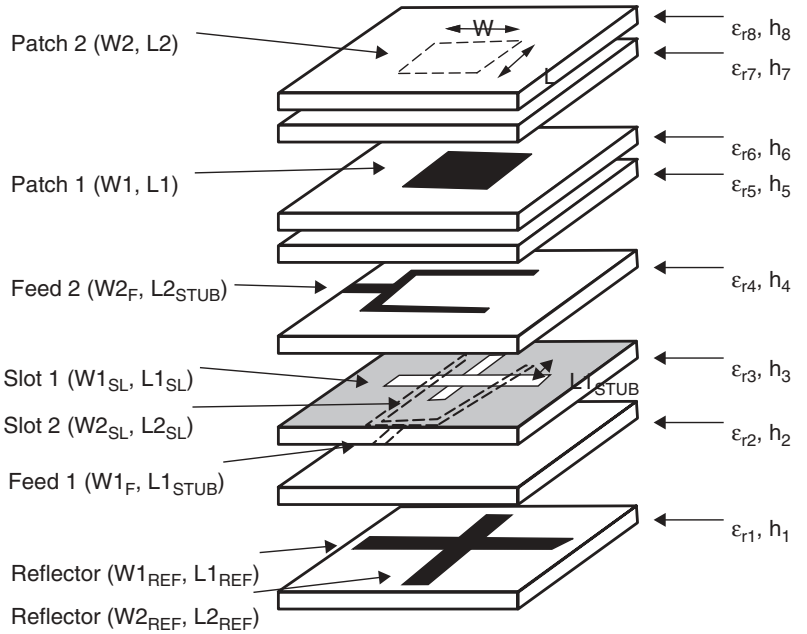


FIGURE 6-37 Exploded view of construction of a dual polarized aperture-fed stacked patch utilizing a crossed strip reflector. (From [8], Fig. 3.6.22, © 2003 Kluwer Academic Publishers.)

and the patch elements. This shows that the slot that couples to a patch resonator can be fed by a microstrip line located either below or above the slot. The ground plane between the two networks for each polarization eliminates direct coupling between the microstrip networks and symmetrical feeding reduces coupling in the slot.

Because we use long slots to feed the lower microstrip patch in an overcoupled excitation, direct coupling of the upper microstrip to the lower patch is minimal in comparison. We use thin substrates of moderate dielectric constant ($\varepsilon_r = 2.2$) to support the etched patches and foam layers between to separate the patches to increase bandwidth and control coupling. Figure 6-37 shows a crossed dipole used as a reflector element below the microstrip feed lines to reflect direct radiation from the crossed slot that reduces the F/B ratio.

6-10 PATCH ANTENNA FEED NETWORKS

Patch antenna arrays may be fed from below (Figure 6-9) by using a stripline distribution network. The connections between the boards greatly complicate the assembly. A connection made vertically from the center strip of a stripline unbalances the fields and induces parallel-plate modes. Shorting pins between the ground planes suppress this mode. It is far easier to etch the feed network on the microstrip and use either edge feeds or aperture feeds with the network located below the patch layer. Feed networks radiate very little in comparison with the patches when etched on the same substrate because radiation from fringing fields on the two sides of the microstrip lines cancel each other except at discontinuities (corners and steps).

Consider the equally fed array (Figure 6-38). Equal amplitude and phase feeding generates virtual magnetic walls between the patches as shown. We can join the edges

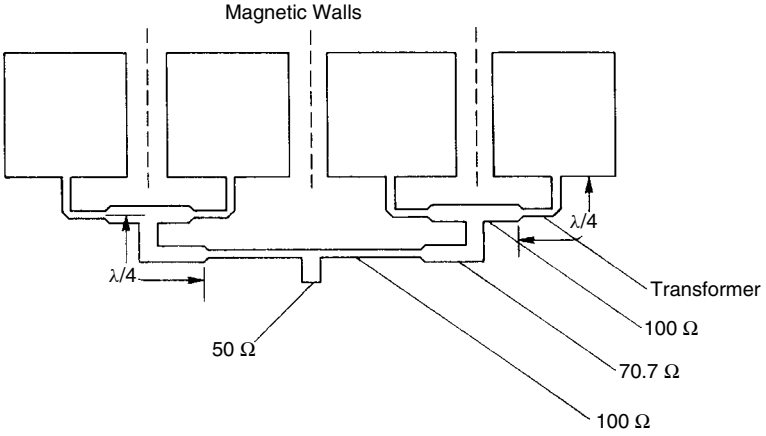


FIGURE 6-38 Equally fed microstrip patch array.

between the patches without effect, since the midpoint remains a virtual open circuit and the separate patches join into a continuous strip. The feeds must be spaced close enough together to prevent grating lobes and to provide uniform amplitude along the edges. These antennas can be wrapped around missiles to provide omnidirectional coverage about the roll axis. To eliminate pattern ripple, feeds must be spaced about every 0.75λ in a circular array. The resistance at each feed at resonance will be the combination of the radiation conductances from the portions of the edges between the magnetic walls.

Figure 6-38 illustrates an equally fed four-element array. Starting from the patch, a quarter-wavelength transformer reduces the roughly 200-Ω impedance to 100Ω. Two 100-Ω lines join in shunt to 50Ω at their juncture. A 70.7-Ω quarter-wavelength line transforms the 50Ω back to 100Ω. We continue this sequence for any 2^N-array for reactive power dividers at each junction. Equal path lengths from the input excite them with equal phases. Arrays with the number of elements different from 2^N-are possible, but they require more difficult feed networks. A 100-Ω system was picked because 50-Ω lines on low-dielectric-constant substrates are quite wide.

The reactive power divider (Figure 6-38) has more bandwidth than the patch while it is matched at the input but not at its outputs. The network can be analyzed by using even and odd modes and shows that the output return loss is 6 dB, and it provides only 6 dB of isolation between outputs. The power reflected from a damaged antenna distributes to the other elements of the array and produces an effect greater than that of just a missing element. Making power dividers with isolation resistors reduces this problem, but we cannot justify the added difficulty of mounting resistors when both good etchings and low probability of damage make them unnecessary.

We must be wary of coupling between different parts of the feed network. We want to pack the feed network into the smallest area, but coupled signals between the lines produce unexpected anomalies. Distinguishing direct radiation from the feed and coupling redistribution is difficult. Although couplings are predictable, they appear as random errors when we cannot perform a full analysis. Unfortunately, the coupling between microstrip lines falls off quite slowly. Table 6-6 lists the coupling and peak errors for 100-Ω lines; those of 50-Ω lines are very similar. We read the amplitude and phase errors from Scales 1-8 and 1-9.

TABLE 6-6 Peak Feed Errors Due to Microstrip Coupling for 100-Ω Lines ($\varepsilon_r = 2.4$)

Spacing/Substrate Thickness	Coupling (dB)	Amplitude Error (dB)	Phase Error (deg)
1.0	16	1.5	9.0
2.0	23	0.7	4.0
3.0	28	0.4	2.3
4.0	32	0.22	1.5
5.0	35	0.12	1.0

6-11 SERIES-FED ARRAY

If we reduce the width of the patch, the radiation conductance is insufficient to match the input. We can use the microstrip patch as a transmission line and connect a line opposite the feed to lead to other patches (Figure 6-39). If we space the patches by half-wavelengths, the impedances of the patches will add in phase at the input, because it rotates once around the Smith chart in $\lambda/2$. The characteristic impedance of the connecting lines has no effect at center frequency. The junction of the transmission-line feeder and the patch introduces excess phase shift. In arrays of a few elements, the extra phase shift can be ignored, but arrays with a large number of elements, or when we design for critical amplitude taper, must account for δ. Of course, traveling-wave or resonant arrays can be designed. The frequency dispersion of the traveling-wave array can be used to frequency-scan the beam.

Various experimental methods have been devised to measure the parameters of the series array. Metzler [31] performed experiments on uniform-width element arrays to determine the radiation conductance and excess phase shift. Measuring the transmission loss through the array as a network with input and output connectors determines the radiation conductance of the patches. An empirical equation was obtained:

$$G = 0.0162\left(\frac{W}{\lambda_0}\right)^{1.757} \qquad 0.033 \le \frac{W}{\lambda} \le 0.254 \tag{6-39}$$

where G is the total radiation conductance of each patch, with half from each edge. Measurement of the beam direction of the uniform traveling-wave array determines the excess phase shift in each patch.

Jones et al. [32] model the patch (Figure 6-39) with extensions Δ due to the fringing fields as a transmission line: $L + 2\Delta$ long. The other excess phase shift, due to the step, is modeled as extensions to the input lines (δ). Jones et al. perform measurements on single elements to establish these lengths. Δ is found from the resonant frequency of the patch: $L + 2\Delta = \lambda/2\sqrt{\varepsilon_{\text{eff}}}$, where ε_{eff} is given by Eq. (6-19). When the transmission-line phase is measured through the patch at resonance, the excess phase beyond π is equated to a phase shift length in the narrow feeder lines:

$$2\delta = \frac{\lambda_N}{2\pi}\phi_{\text{excess}}$$

where λ_N is the wavelength in the narrow line.

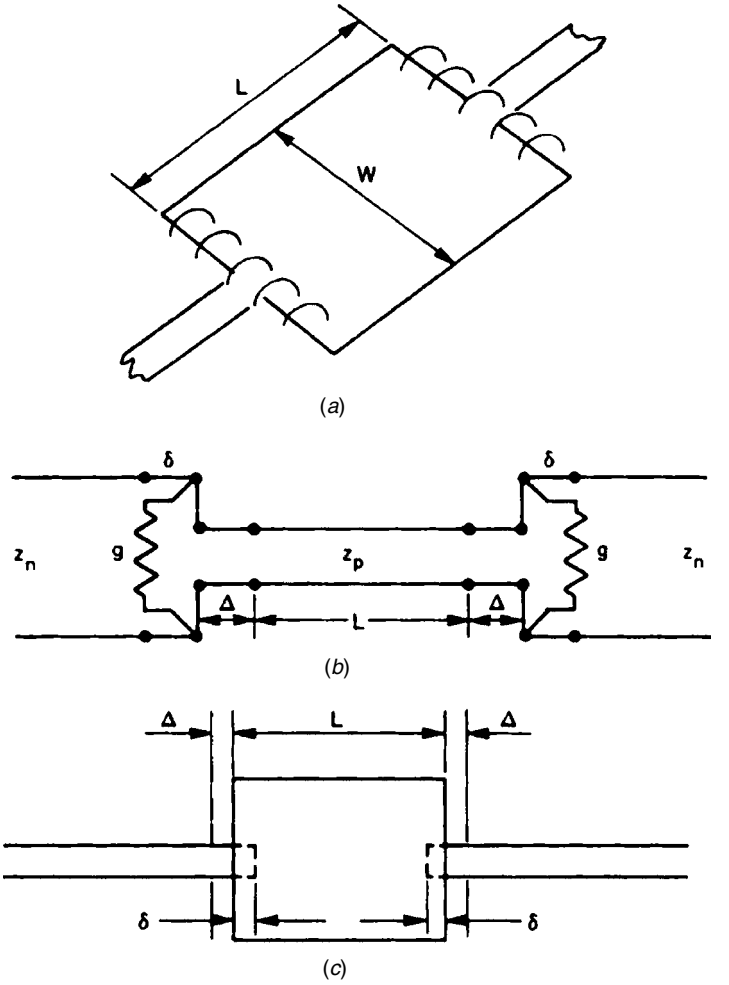


FIGURE 6-39 Series-fed patch and its equivalent circuit. (From [32], Fig. 2, © 1982 IEEE.)

When designing the array, we vary the widths of the patches to achieve the desired amplitude taper. The voltage distribution at each patch is given by $V\sqrt{g}$, where g is the patch conductance. Standing-wave (resonant) arrays require that the sum of the conductances be equal to the input conductance desired. We have some latitude when we feed the array through a quarter-wavelength transformer. The nonresonant array requires a matched load on the end to prevent standing waves. We must pick the ratio of the power dissipated to the radiated power that gives us an extra parameter with which to optimize the design. We control the beam direction by spacing the elements to achieve the phase shift required.

6-12 MICROSTRIP DIPOLE [33]

As the width W of a patch narrows, the input impedance increases. When the width approaches that of a microstrip feed line, either the patch fails to be a resonator or the

feed line becomes very narrow in trying to transform the impedance. The microstrip dipole solves these problems by having a coupled line feeder. The dipole is a half-wavelength strip whose width equals that of a microstrip feed line. A line etched on a substrate below feeds the dipole by coupling into the strip (Figure 6-40*a*). The equivalent circuit (Figure 6-40*b*) transforms the high impedance of the dipole through the unequal coupled lines. By varying the coupling, we can change the input impedance at resonance. Best results occur for quarter-wavelength overlap where the equivalent stubs (Figure 6-40*b*) do not contribute reactance. We vary the coupling by changing the thickness of the substrate between the strips or by offsetting the lower strip.

The dipole radiates as a narrow patch and not as a dipole. No pattern nulls appear along the axis of the strip, but they occur more strongly in the direction of the equivalent magnetic currents of the edges. The H-plane pattern becomes quite broad for the narrow strip width. The feed distribution circuit is etched on the substrate below the dipoles. With the feed circuit on a separate level, we have greater freedom in the feed network design to excite desired distributions. Also, because the dipoles are small, we can use density tapering of the dipoles to that end. Proper design requires measurement [34]

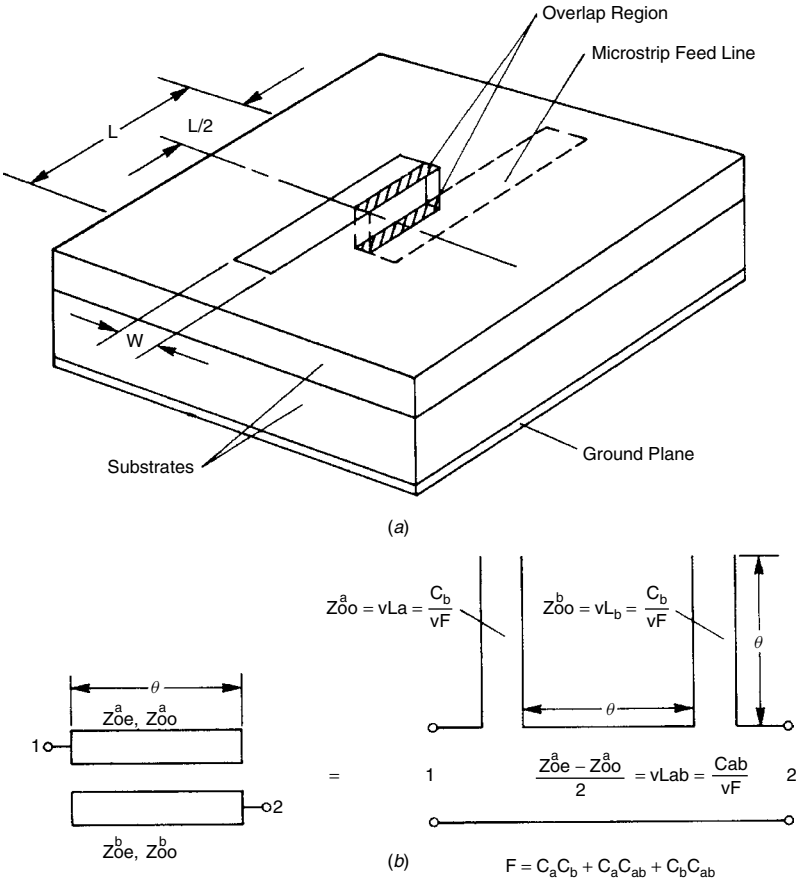


FIGURE 6-40 (*a*) Microstrip dipole; (*b*) equivalent circuit. [(*b*) From G. L. Matthaei et al., *Microwave Filters, Impedance Matching Networks, and Coupling Structures,* © 1980 Artech House, Inc.]

to obtain the desired effect, since mutual coupling will change the distribution by changing the active impedance of each dipole. The feed network must compensate for the coupling.

6-13 MICROSTRIP FRANKLIN ARRAY [35]

An electrically long line with a standing wave on it fails to radiate on broadside because the many cycles cancel each other. We obtain a pattern with many nulls and lobes. By folding the lines with out-of-phase standing-wave currents close together, we can prevent their radiation. The other portions are free to radiate (Figure 6-41*a*). The Franklin array consists of straight sections $\lambda/2$ long connected by $\lambda/4$ shorted stubs. The standing-wave currents on the straight portions add in phase.

We can construct a microstrip version (Figure 6-41*b*). Half-wavelength lines act as radiators (patches). We connect them with half-wavelength lines folded into stubs so that the counteracting standing-wave currents do not radiate. The straight lines are narrow patches. The total radiation conductance of each strip is

$$G = \frac{1}{45}\left(\frac{W}{\lambda}\right)^2 \tag{6-40}$$

for narrow strip widths W, where λ is the free-space wavelength. Using lines for the stubs whose impedance is twice the radiating strip impedance reduces unwanted internal reflections. The two stubs add in shunt. Since the antenna is quite narrowband and the length of the lines between patches is a half-wavelength long, the impedance of these connecting arms has a secondary effect.

Example Design an eight-wavelength array at 10 GHz. There are 16 patches in the array.

The radiation conductances add for elements spaced at $\lambda/2$ intervals. For a 100-Ω input, each patch supplies a conductance 0.01/16. We solve Eq. (6-40) for the width:

$$W = \lambda\sqrt{\frac{0.01(45)}{16}} = 0.168\lambda$$

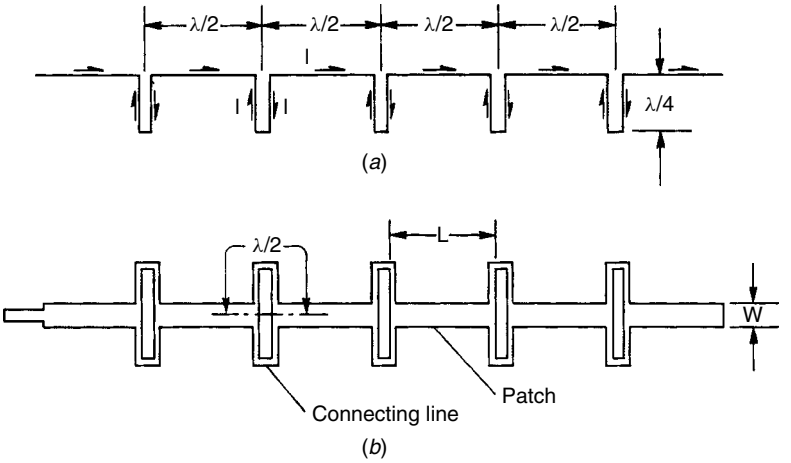


FIGURE 6-41 (*a*) Dipole and (*b*) microstrip Franklin arrays.

If we use Eq. (6-39) from the series patch, we obtain $W = 0.157\lambda$, within the range of the empirical formula. For 10 GHz, $W = 4.71$ mm. On an 0.8-mm substrate ($\varepsilon_r = 2.21$), $W/H = 5.89$ and the impedance of the strip radiator $Z_0 = 44.01\ \Omega$. We need to find the effective dielectric constant of the strip to determine the patch length and impedance. From Eq. (6-19), $\varepsilon_r = 1.97$.

We calculate the cutback from each end by using Eq. (6-18); $\Delta = 0.40$ mm. Each radiating strip is

$$L = \frac{300 \times 10^9}{10^{10}(2)\sqrt{1.97}} - 2(0.40) = 9.88 \text{ mm}$$

The radiating-strip impedance is $Z_0/\sqrt{\varepsilon_{\text{eff}}} = 31.3\ \Omega$. We need 62.6-Ω connecting lines in the stubs to achieve the broadest bandwidth. With so few radiators, we could use 100-Ω connecting arms with little change in bandwidth and have more reasonable connecting arm widths: 0.71 mm.

The example shows that the microstrip Franklin array works best for high frequencies or long arrays. The elements are narrow, and the interconnecting arms are thin.

6-14 MICROSTRIP ANTENNA MECHANICAL PROPERTIES

A microstrip patch antenna has very desirable mechanical properties. It can withstand tremendous shock and vibration. Because the antenna is on a solid substrate, the patch cannot flex, and small changes in the substrate thickness have only a minor effect on the resonant frequency. The commonly used soft substrate (Teflon and fiberglass) has a good damped resilience. Microstrip patch antennas have been used to telemeter data from artillery shells and high-velocity rockets, which have high shock and vibration levels. The repeatability of the dimensions of the patches depends only on the etcher's art. Complicated shapes and feed networks are produced as cheaply as simple ones.

The antennas can withstand exposure to high temperatures when covered by a radome made of the same soft dielectric as the substrate. The cover protects the metal patches but has only a minor effect on the resonant frequency [36]. High temperatures on the surface of the radome or ablation fail to change the resonance significantly because the radome itself has only a minor effect. Variation in the dielectric constant of the substrate from lot to lot causes problems with repeatability. The narrowband antennas require measurement of the dielectric constant of each lot, and sometimes of each sheet, to get the center frequency desired. A series of etching masks can be made to cover the expected range. The antennas can be tuned with inductive shorting pins or capacitive screws, but tuning is prohibitive when the number of elements in an array is large. Careful quality control of the dielectric constant is the answer. Close monitoring of the etching process may also be needed to prevent excessive undercutting.

Temperature variations can be a problem with thin substrates when the bandwidth is narrow. The patch and substrate size grow when the temperature rises, but they are overshadowed by the change in dielectric constant of soft substrates. Instead of decreasing the resonant frequency because of the increased patch size, a lowered dielectric constant raises the center frequency.

Whenever we need more bandwidth than a microstrip patch can provide, we must turn to cavity antennas. We increase the antenna volume by penetrating the vehicle for the cavity, but we gain a design parameter.

REFERENCES

1. R. F. Harrington, Effects of antenna size on gain, bandwidth, and efficiency, *Journal of Research, NBS, D, Radio Propagation*, vol. 64D, January–February 1960, pp. 1–12.
2. R. C. Hansen, Fundamental limitations in antennas, *Proceedings of IEEE*, vol. 69, no. 2, February 1981, pp. 169–173.
3. Y. T. Lo et al., Study of microstrip antenna elements, arrays, feeds, losses, and applications, *Final Report, RADC-TR-81-98*, Rome Air Development Center, Rome, NY, June 1981.
4. K. R. Carver and E. L. Coffey, Theoretical investigation of the microstrip antenna, *Technical Report PT-00929*, Physical Science Laboratory, New Mexico State University, Las Cruces, NM, January 1979.
5. D. R. Jackson and N. G. Alexopoulas, Simple approximate formulas for input resistance, bandwidth, and efficiency of a resonant rectangular patch, *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 3, March 1991, pp. 407–410.
6. R. F. Harrington, *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York, 1961.
7. D. M. Pozar, Rigorous closed-form expressions for the surface wave loss of printed antennas, *Electronics Letters*, vol. 26, no. 13, June 21, 1990, pp. 954–956.
8. R. B. Waterhouse, ed., *Microstrip Patch Antennas, A Designer's Guide*, Kluwer Academic, Boston, 2003.
9. R. E. Munson, Conformal microstrip antennas and microstrip phased arrays, *IEEE Transactions on Antennas and Propagation*, vol. AP-22, no. 1, January 1974, pp. 74–78.
10. E. O. Hammerstad, Equations for microstrip circuit design, *Proceedings of the 5th European Micro-strip Conference*, Hamburg, Germany, September 1975, pp. 268–272.
11. M. V. Schneider, Microstrip lines for microwave integrated circuits, *Bell System Technical Journal*, vol. 48, May–June 1969, pp. 1421–1444.
12. D. A. Paschen, Practical examples of integral broadband matching of microstrip antenna elements, *Proceedings of the 1986 Antenna Applications Symposium*, Monticello, IL.
13. T. Samaras, A. Kouloglou, and J. N. Sahalos, A note on impedance variation of a rectangular microstrip patch antenna with feed position, *IEEE Antennas and Propagation Magazine*, vol. 46, no. 2, April 2004.
14. D. M. Pozar, Microstrip antenna aperture-coupled to a microstrip-line, *Electronics Letters*, vol. 21, no. 2, January 1985, pp. 49–50.
15. P. L. Sullivan and D. H. Schaubert, Analysis of an aperture coupled microstrip antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-34, no. 8, August 1986, pp. 977–984.
16. G. Kumar and K. P. Ray, *Broadband Microstrip Antennas*, Artech House, Boston, 2003.
17. J. R. James, P. S. Hall, and C. Wood, *Microstrip Antennas: Theory and Design*, Peter Peregrinus, London, 1981.
18. L. C. Shen et al., Resonant frequency of a circular disk, printed circuit antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-25, no. 4, July 1977, pp. 595–596.
19. I. J. Bahl and P. Bhartia, *Microstrip Antennas*, Artech House, Dedham, MA, 1980.
20. A. G. Derneryd, Analysis of the microstrip disk antenna element, *IEEE Transactions on Antennas and Propagation*, vol. AP-27, no. 5, September 1979, pp. 660–664.
21. J. L. Kerr, Microstrip antenna developments, *Proceedings of the Workshop on Printed Circuit Antennas*, New Mexico State University, Las Cruces, NM, October 1979, pp. 3.1–3.20.
22. K.-L. Wong, *Compact and Broadband Microstrip Antennas*, Wiley, New York, 2002.
23. K.-L. Wong, *Planar Antennas for Wireless Communications*, Wiley, New York, 2003.

24. R. B. Waterhouse, S. D. Targonski, and D. M. Kokotoff, Design and performance of small printed antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-46, no. 11, November 1998, pp. 1629–1633.
25. R. B. Waterhouse, ed., *Microstrip Patch Antennas: A Designer's Guide*, Kluwer Academic, Boston, 2003.
26. T. Taga and K. Tsunekawa, Performance analysis of a built-in planar inverted-F antenna for 800 MHz band portable radio units, *IEEE Transactions on Selected Areas in Communication*, vol. SAC-5, no. 5, June 1987, pp. 921–929.
27. R. B. Waterhouse, Design of probe-fed stacked patches, *IEEE Transactions on Antennas and Propagation*, vol. AP-47, no. 12, December 1999, pp. 1780–1784.
28. S. D. Targonski, R. B. Waterhouse, and D. M. Pozar, Design of wide-band aperture-stacked patch microstrip antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-46, no. 9, September 1998, pp. 1245–1251.
29. S. D. Targonski and R. B. Waterhouse, Reflector elements for aperture and aperture coupled microstrip antennas, *IEEE Antennas and Propagation Symposium Digest*, Montreal, Quebec, Canada, July 1997, pp. 1840–1843.
30. J. R. Sanford and A. Tengs, A two substrate dual polarized aperture coupled patch, *IEEE Antennas and Propagation Symposium Digest*, 1996, pp. 1544–1547.
31. T. Metzler, Microstrip series arrays, *Proceedings of the Workshop on Printed Circuit Antennas*, New Mexico State University, Las Cruces, NM, October 1979, pp. 20.1–20.16.
32. B. B. Jones, F. V. M. Chow, and A. W. Seeto, The synthesis of shaped patterns with series-fed microstrip patch arrays, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 6, November 1982, pp. 1206–1212.
33. D. A. Huebner, An electrically small dipole planar array, *Proceedings of the Workshop on Printed Circuit Antennas*, New Mexico State University, Las Cruces, NM, October 1979, pp. 17.1–17.16.
34. R. S. Elliott and G. J. Stern, The design of microstrip dipole arrays including mutual coupling, parts I and II, *IEEE Transactions on Antennas and Propagation*, vol. AP-29, no. 5, September 1981, pp. 757–765.
35. K. Solbach, Microstrip-Franklin antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 4, July 1982, pp. 773–775.
36. I. J. Bahl et al., Design of microstrip antennas covered with a dielectric layer, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 2, March 1982, pp. 314–318.

7

HORN ANTENNAS

Horn antennas have a long history, traced in part in the collection of papers by Love [1] together with papers on every other horn topic. Horns have a wide variety of uses, from small-aperture antennas to feed reflectors to large-aperture antennas used by themselves as medium-gain antennas. Horns can be excited in any polarization or combination of polarizations. The purity of polarization possible and the unidirectional pattern make horns good laboratory standards and ideal reflector feeds. Horns also closely follow the characteristics predicted by simple theories.

Horns are analyzed using a variety of techniques. Barrow and Chu [2] analyzed a sectoral horn, flaring in only one plane, by solving the boundary value problem in the wedge. They expanded the fields in terms of Hankel functions in cylindrical coordinates. The fields form an equiphase surface over a cylindrical cap to which the Kirchhoff–Huygens equivalent current method [Eq. (2-23)] can be applied to compute the pattern. Similarly, Schorr and Beck [3] use spherical Hankel and Legendre functions to analyze conical horns. The integration surface consists of a spherical cap. Schelkunoff and Friis [4] use the mouth of the horn as the aperture and approximate the phase distribution as quadratic. Both aperture theories have the same valid pattern range. The method predicts patterns accurately in the area in front of the aperture. The error increases as the plane of the aperture is approached. The predicted pattern remains continuous and gives no indication of its increasing error. GTD methods [5] predict the pattern both in back and in front of the aperture while providing estimates of the error in the predictions. Most of the details needed for design can be obtained from the aperture theory. Only GTD predicts sidelobes accurately, since no assumption of zero fields outside the horn aperture is made.

Figure 7-1 shows the general horn geometry. The input waveguide can be either rectangular or circular (elliptical). W is the width of a rectangular aperture, and a is the radius of a circular aperture. The distance from the junction of the projected sides

Modern Antenna Design, Second Edition, By Thomas A. Milligan

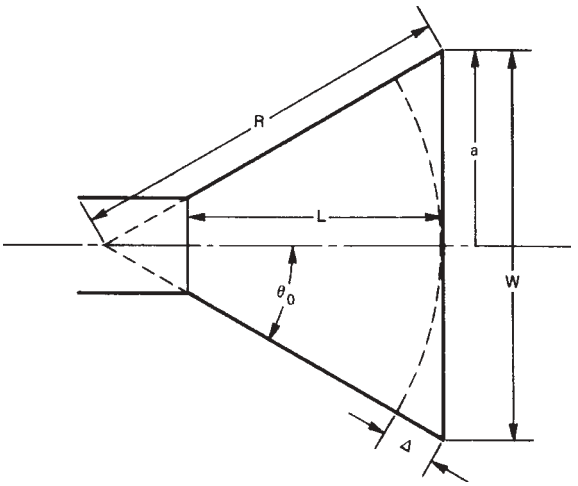


FIGURE 7-1 General geometry of a horn.

to the aperture is the slant radius R. The distance along the centerline from the aperture to the waveguide is the axial length. We derive the aperture field amplitude from the input waveguide mode while the phase distribution is approximately quadratic across the aperture. We assume that the aperture fields radiate in spherical waves from the projected juncture of the sides, and the extra distance along the sides compared with the distance to the center of the aperture is given by

$$\begin{aligned}\Delta &= R - \sqrt{R^2 - a^2} \\ &= R\left(1 - \sqrt{1 - \frac{a^2}{R^2}}\right) \\ &\approx R\left[1 - \left(1 - \frac{a^2}{2R^2}\right)\right] = \frac{a^2}{2R} = \frac{W^2}{8R}\end{aligned}$$

We divide by wavelength to obtain the dimensionless constant S of the quadratic phase distribution:

$$S = \frac{\Delta}{\lambda} = \frac{W^2}{8\lambda R} = \frac{a^2}{2\lambda R} \tag{7-1}$$

Since the semiflare angle θ_0 of most practical horns is small, we use the quadratic phase error approximation.

7-1 RECTANGULAR HORN (PYRAMIDAL)

The rectangular horn flares out of a rectangular or square waveguide with flat metal walls. Figure 7-2 shows the horn geometry. The slant radiuses along the sides will be unequal, in general. The input waveguide dimensions are width a and height b.

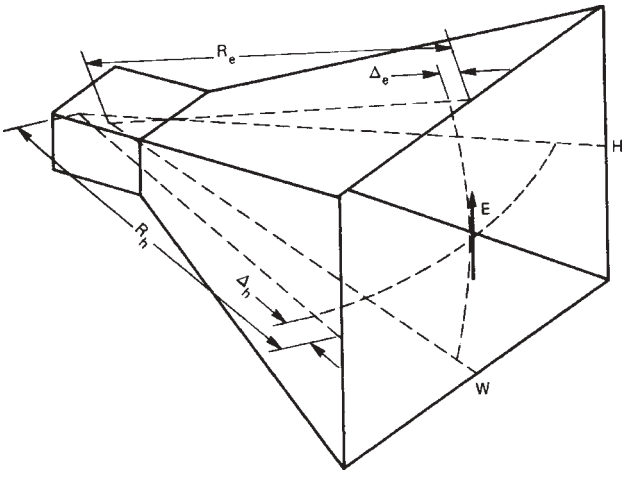


FIGURE 7-2 Rectangular horn geometry.

The aperture has width W in the H-plane and height H in the E-plane. Each aperture coordinate has its own quadratic phase distribution constant:

$$S_e = \frac{H^2}{8\lambda R_e} \qquad S_h = \frac{W^2}{8\lambda R_h} \tag{7-2}$$

The TE_{10} mode of the lowest-order waveguide mode has the field distribution

$$E_y = E_0 \cos \frac{\pi x}{a}$$

Combining these ideas, the aperture electric field is approximated by

$$E_y = E_0 \cos \frac{\pi x}{W} \exp \left\{ -j2\pi \left[S_e \left(\frac{2y}{H} \right)^2 + S_h \left(\frac{2x}{W} \right)^2 \right] \right\} \tag{7-3}$$

The ratio of the electric and magnetic fields approaches the impedance of free space for large apertures. In this case we use the Huygens source approximation and need only the electric field with Eq. (2-24) to find the pattern. Small-aperture horns require Eq. (2-23) with an arbitrary ratio of the magnetic and electric fields.

We compute the E-plane pattern by using a uniform aperture distribution and the H-plane pattern from a cosine distribution. Both have a quadratic phase error. Figures 7-3 and 7-4 plot the E- and H-plane universal patterns in U-space of the Taylor distribution with S as a parameter. We can use them to determine the pattern of a general rectangular horn.

Example Compute the pattern level at $\theta = 15^\circ$ in the E- and H-planes of a horn with the following measured dimensions:

Aperture: $W(H\text{-plane}) = 28.9\,\text{cm}$, $H(E\text{-plane}) = 21.3\,\text{cm}$
Input waveguide: width $a = 3.50\,\text{cm}$, height $b = 1.75\,\text{cm}$

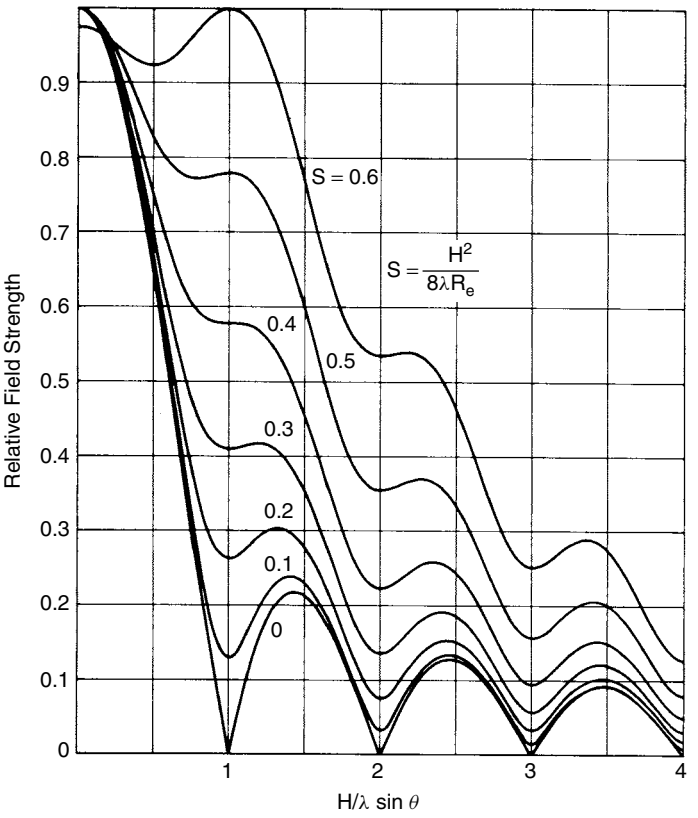


FIGURE 7-3 E-plane universal pattern of a rectangular, TE_{10} mode.

The slant distance from the aperture to the waveguide along the center of each plate of the flare was measured: $D_h = 44.8\,\text{cm}$ and $D_e = 44.1\,\text{cm}$. We calculate the slant radius from similar triangles:

$$\frac{R_h}{D_h} = \frac{W}{W - a} \qquad \frac{R_e}{D_e} = \frac{H}{H - b} \tag{7-4}$$

Slant radius: $R_h = 50.97\,\text{cm}$, $R_e = 48.05\,\text{cm}$

The frequency is 8 GHz ($\lambda = 3.75\,\text{cm}$). Using Eq. (7-2), we compute $S_h = 0.55$ and $S_e = 0.31$. We use Figures 7-3 and 7-4 to determine the universal pattern field intensity (voltage):

$$\frac{W}{\lambda}\sin\theta = 2.0 \qquad \frac{H}{\lambda}\sin\theta = 1.47$$

The fields from the figures are 0.27 (H-plane) and 0.36 (E-plane). We must include the obliquity factor of the Huygens source element pattern: $(1 + \cos\theta)/2$ to obtain the proper pattern level. At $\theta = 15^\circ$, the obliquity factor is 0.983. We calculate the pattern level in decibels from 20 times the logarithm of the product of the field intensity from

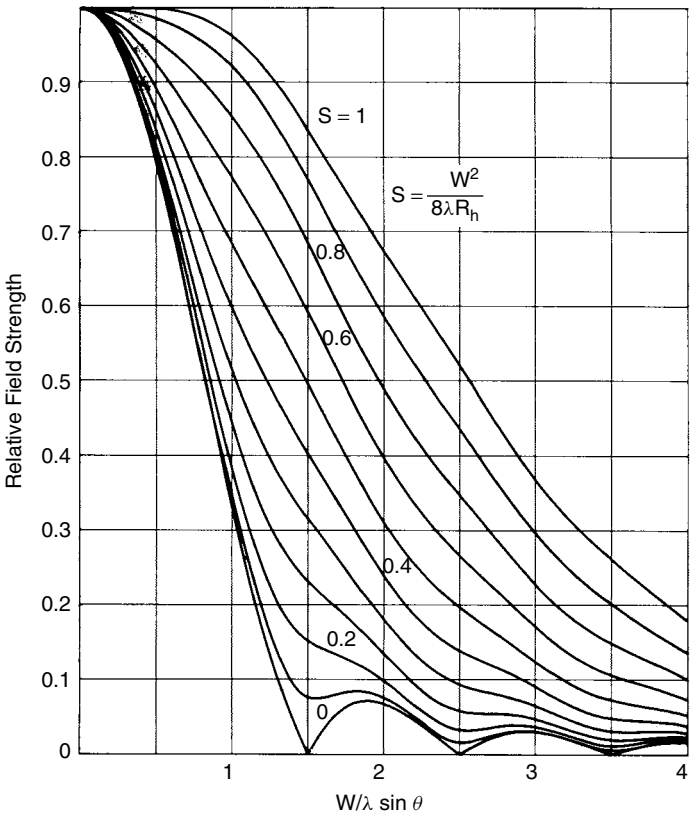


FIGURE 7-4 H-plane universal pattern of a rectangular, TE_{10} mode.

the figures and the obliquity factor:

$$H\text{-plane}: -11.5\,\text{dB} \qquad E\text{-plane}: -9\,\text{dB}$$

We can calculate gain of this horn by using aperture efficiencies:

$$H\text{-plane (cosine) (Table 4-1)}: \ 0.91\,\text{dB} \qquad E\text{-plane (uniform)}: \ 0\,\text{dB}$$

These values hold for all rectangular horns excited by the TE_{10} mode. The quadratic phase distributions give us the phase error loss. From Table 4-42 we interpolate these losses:

$$S_h = 0.55 \text{ cosine distribution} \qquad \text{PEL} = 2.09\,\text{dB}$$
$$S_e = 0.31 \text{ uniform distribution} \qquad \text{PEL} = 1.50\,\text{dB}$$

The directivity is given by

$$\text{directivity} = 10\log\frac{4\pi WH}{\lambda^2} - \text{ATL}_h - \text{ATL}_e - \text{PEL}_h - \text{PEL}_e = 22.9\,\text{dB} \qquad (7\text{-}5)$$

The aperture efficiency is 35.5%, since $\text{ATL} + \text{PEL}_h + \text{PEL}_e = 4.5\,\text{dB}$.

We usually equate gain to directivity, since the wall losses are very small. Of course, for millimeter-wave horns we must include wall losses. Although we can use Table 4-42 along with the fixed-amplitude taper loss of 0.91 dB to determine the aperture efficiency of a rectangular horn, Schelkunoff and Friis [4] give the following closed-form equation for the directivity:

$$\text{directivity} = \frac{8R_hR_e}{WH}\{[C(u) - C(v)]^2 + [S(u) - S(v)]^2\}[C^2(z) + S^2(z)]$$

where

$$u = \frac{1}{\sqrt{2}}\left(\frac{\sqrt{\lambda R_h}}{W} + \frac{W}{\sqrt{\lambda R_h}}\right) \qquad v = \frac{1}{\sqrt{2}}\left(\frac{\sqrt{\lambda R_h}}{W} - \frac{W}{\sqrt{\lambda R_h}}\right) \qquad z = \frac{H}{\sqrt{2\lambda R_e}} \tag{7-6}$$

and

$$C(x) = \int_0^x \cos\frac{\pi t^2}{2}dt \qquad S(x) = \int_0^x \sin\frac{\pi t^2}{2}dt$$

are the Fresnel integrals. Closed-form expressions for these integrals are available [6].

7-1.1 Beamwidth

We can use Figures 7-3 and 7-4 to compute beamwidths. The 3-dB point corresponds to 0.707 and the 10-dB point to 0.316 on the graphs. Table 7-1 lists the 3- and 10-dB points (values of U) for different quadratic phase constants S in the H-plane. Similarly, Table 7-2 lists the E-plane points. The tables are more convenient than the graphs. Because we remove the obliquity factor to get a universal pattern, we must modify the beamwidths found by using the tables. We find the beamwidth, but then we must add the obliquity factor to the beamwidth level. The beamwidth is found for

TABLE 7-1 Rectangular-Horn H-Plane Beamwidth Points, TE_{10} Mode

	$(W/\lambda)\sin\theta$			$(W/\lambda)\sin\theta$	
S	3 dB	10 dB	S	3 dB	10 dB
0.00	0.5945	1.0194	0.52	0.8070	1.8062
0.04	0.5952	1.0220	0.56	0.8656	1.8947
0.08	0.5976	1.0301	0.60	0.9401	1.9861
0.12	0.6010	1.0442	0.64	1.0317	2.0872
0.16	0.6073	1.0652	0.68	1.1365	2.2047
0.20	0.6150	1.0949	0.72	1.2445	2.3418
0.24	0.6248	1.1358	0.76	1.3473	2.4876
0.28	0.6372	1.1921	0.80	1.4425	2.6246
0.32	0.6526	1.2700	0.84	1.5320	2.7476
0.36	0.6716	1.3742	0.88	1.6191	2.8618
0.40	0.6951	1.4959	0.92	1.7071	2.9744
0.44	0.7243	1.6123	0.96	1.7991	3.0924
0.48	0.7609	1.7143	1.00	1.8970	3.2208

TABLE 7-2 Rectangular-Horn *E*-Plane Beamwidth Points, TE_{10} Mode

	$(H/\lambda)\sin\theta$			$(H/\lambda)\sin\theta$	
S	3 dB	10 dB	S	3 dB	10 dB
0.00	0.4430	0.7380	0.24	0.4676	1.4592
0.04	0.4435	0.7405	0.28	0.4793	1.5416
0.08	0.4452	0.7484	0.32	0.4956	1.6034
0.12	0.4482	0.7631	0.36	0.5193	1.6605
0.16	0.4527	0.7879	0.40	0.5565	1.7214
0.20	0.4590	0.8326	0.44	0.6281	1.8004

a lower pattern level than specified. Since the beamwidth levels are close, we use a relation by Kelleher [7, p. 12–5] with good results:

$$\frac{\text{BW}_2}{\text{BW}_1} = \sqrt{\frac{\text{level}_2(\text{dB})}{\text{level}_1(\text{dB})}} \tag{7-7}$$

Example Compute 3- and 10-dB beamwidths for the horn in the preceding example. We have $S_h = 0.55$ and $S_e = 0.31$. From Tables 7-1 and 7-2,

$$\frac{W}{\lambda}\sin\theta_{h3} = 0.851 \qquad \frac{H}{\lambda}\sin\theta_{e3} = 0.4915$$

$$\frac{W}{\lambda}\sin\theta_{h10} = 1.8726 \qquad \frac{H}{\lambda}\sin\theta_{e10} = 1.588$$

$$\frac{W}{\lambda} = \frac{28.9}{3.75} = 7.707 \qquad \frac{H}{\lambda} = \frac{21.3}{3.75} = 5.68$$

$$\theta_{h3} = 6.34 \qquad \theta_{h10} = 14.06 \qquad \theta_{e3} = 4.96 \qquad \theta_{e10} = 16.24$$

We consider the obliquity factor, $(1 + \cos\theta)/2$, at these angles, and apply Eq. (7-7) to reduce the beamwidths found.

$$\text{BW}_{h3} = 12.68^\circ \text{at } 3.03\,\text{dB} \qquad \text{BW}_{h3}x = 12.62^\circ \text{at } 3.01\,\text{dB}$$

$$\text{BW}_{e3} = 9.92^\circ \text{at } 3.02\,\text{dB} \qquad \text{BW}_{e3} = 9.89^\circ \text{at } 3.01\,\text{dB}$$

$$\text{BW}_{h10} = 28.12^\circ \text{at } 10.13\,\text{dB} \qquad \text{BW}_{h10} = 27.94^\circ \text{at } 10\,\text{dB}$$

$$\text{BW}_{e10} = 32.48^\circ \text{at } 10.18\,\text{dB} \qquad \text{BW}_{10} = 32.2^\circ \text{at } 10\,\text{dB}$$

Including the obliquity factor has a very small effect on the results, but the effect increases with larger beamwidths (smaller apertures).

Aperture theory fails for small horns because the beam is determined more by edge diffraction than the aperture fields. Empirical data were collected and reduced to simple formulas for small rectangular horns based on aperture size only [8, p. 46–22]:

$$\text{BW}_{10e} = 88^\circ \frac{\lambda}{H} \quad \text{and} \quad \text{BW}_{10h} = 31^\circ + 79^\circ \frac{\lambda}{W}$$

7-1.2 Optimum Rectangular Horn

A rectangular horn has extra parameters, which we can use to design various optimum horns. Given a desired gain, we can design any number of horns with the same gain. Any optimum design depends on the requirements. Without any particular requirements, we will pick an antenna with equal E- and H-plane 3-dB beamwidths [9], but even this does not determine the design totally.

If we pick a constant slant radius and vary the aperture width, the gain increases with increasing aperture width, but the quadratic phase error loss increases faster and produces a maximum point. The maximum occurs in the two planes at approximately constant phase deviations independent of the slant radius:

$$S_h = 0.40 \qquad S_e = 0.26 \tag{7-8}$$

At these points we read the 3-dB points from Tables 7-1 and 7-2:

$$\frac{W}{\lambda}\sin\theta = 0.6951 \qquad \frac{H}{\lambda}\sin\theta = 0.4735$$

On dividing these equations to eliminate the constant $\sin\theta$ in the two planes, we derive the ratio of width to height to give a constant 3-dB beamwidth in the two planes for this optimum point:

$$\frac{H}{W} = 0.68 \tag{7-9}$$

The ratio depends on the beamwidth level. For 10-dB beamwidths,

$$\frac{H}{W} = 1.00 \tag{7-10}$$

These values of S determine the efficiency of the optimum horn. We read the PEL of the quadratic phase distribution from Table 4-42 by using a cosine distribution for the H-plane and a uniform distribution for the E-plane. The H-plane distribution has an ATL of 0.91 dB.

$$\text{PEL}_h = 1.14\,\text{dB} \qquad \text{PEL}_e = 1.05\,\text{dB} \qquad \text{ATL} = 0.91\,\text{dB}$$

or an efficiency of 49%:

$$\text{gain} = \frac{4\pi HW}{\lambda^2}0.49$$

We solve for H and W for a given gain, since we know the ratio between them [Eq. (7-9)]:

$$\frac{W}{\lambda} = \sqrt{\frac{\text{gain}}{4\pi(0.68)(0.49)}} \qquad \frac{H}{\lambda} = \sqrt{\frac{\text{gain}(0.68)}{4\pi(0.49)}}$$

$$\frac{W}{\lambda} = 0.489\sqrt{\text{gain}} \qquad \frac{H}{\lambda} = 0.332\sqrt{\text{gain}} \tag{7-11}$$

We combine Eqs. (7-8) and (7-11) to calculate slant radiuses:

$$\frac{R_h}{\lambda} = 0.0746 \cdot \text{gain} \qquad \frac{R_e}{\lambda} = 0.0531 \cdot \text{gain} \tag{7-12a, b}$$

If, given gain, we use Eqs. (7-11) and (7-12) to design a horn, the dimensions will not be practical with an arbitrary input waveguide. The axial lengths from the waveguide to the aperture must be equal in the E- and H-planes, so the horn will meet the waveguide in a single plane. Given waveguide dimensions a and b, the axial lengths are

$$L_h = \frac{W-a}{W}\sqrt{R_h^2 - \frac{W^2}{4}} \qquad L_e = \frac{H-a}{H}\sqrt{R_e^2 - \frac{H^2}{4}} \tag{7-13a, b}$$

We have a choice between retaining the E- or H-plane slant radius given by Eq. (7-12) and forcing the other radius to give the same axial length. The primary factor affecting gain is the aperture dimensions, which we retain from Eq. (7-11). The slant radius is secondary. We retain the H-plane radius calculated from Eq. (7-12) and modify the E-plane radius. Modifying the H-plane radius would give us a second optimum design:

$$R_e = \frac{H}{H-b}\sqrt{L^2 + \frac{(H-b)^2}{4}} \tag{7-14}$$

To obtain the proper gain, we must iterate, since we cannot use both Eq. (7-12). Design the horn by using Eqs. (7-11), (7-12a), (7-13a), and (7-14). Calculate the gain from the dimensions and obtain a new design gain from

$$G_{d,\text{new}} = \frac{G_{\text{required}} G_{d,\text{old}}}{G_{\text{actual}}} \tag{7-15}$$

where G_{required} is the required gain, G_{actual} is the actual gain, and $G_{d,\text{old}}$ is the old design gain.

Example Design a horn fed from WR-90 waveguide to have 22 dB of gain at 10 GHz.

The waveguide dimensions are 2.286 cm × 1.016 cm (0.9 in. × 0.4 in.) and $G_{\text{req}} = G_d = 10^{22/10} = 158.5$.

On substituting in Eq. (7-11), we calculate aperture dimensions: $W = 18.47$ cm and $H = 12.54$ cm. From Eqs. (7-12a) and (7-13a), $R_h = 35.47$ cm and $L = 30.01$ cm. To get the same axial length in the E-plane [Eq. (7-14)], $R_e = 33.25$ cm. We now calculate gain and compare it with the gain required. The amplitude taper loss and phase error loss in the H-plane remain constant, since S_h is fixed.

$$\text{ATL} = 0.91\,\text{dB} \qquad \text{PEL}_h = 1.14\,\text{dB} \quad \text{at} \quad S_h = 0.40$$

Calculate S_e:

$$S_e = \frac{H^2}{8\lambda R_e} = 0.197 \qquad \text{PEL}_e = 0.60\,\text{dB} \quad \text{(Table 4-42)}$$

$$G_{\text{actual}}(\text{dB}) = 10\log\frac{4\pi HW}{\lambda^2} - \text{ATL} - \text{PEL}_h - \text{PEL}_e = 22.45\,\text{dB}$$

$$= 175.8$$

We must pick a new design gain [Eq. (7-15)]:

$$G_{d,\text{new}} = \frac{158.5(158.5)}{175.8} = 142.9$$

A second iteration with this gain gives the following dimensions:

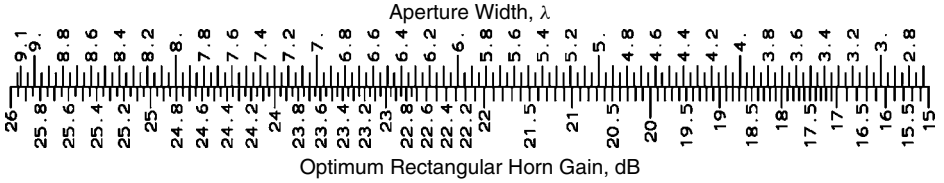
$$W = 17.54\,\text{cm} \qquad H = 11.91\,\text{cm} \qquad L = 26.75\,\text{cm} \qquad R_h = 31.98\,\text{cm}$$

$$R_e = 29.84\,\text{cm} \qquad S_e = 0.198 \qquad \text{PEL}_e = 0.60$$

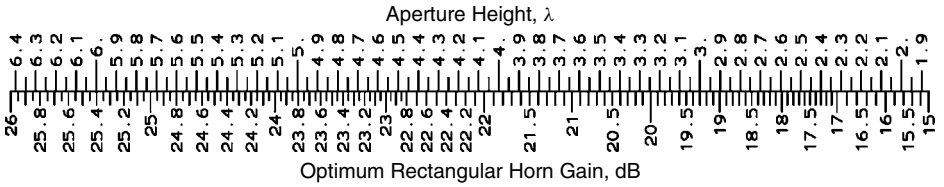
$$\text{gain} = 10\log\frac{4\pi HW}{\lambda^2} - 0.91 - 1.14 - 0.60 = 22.00\,\text{dB}$$

We obtain the gain desired, but the 3-dB beamwidths are only approximately equal: H-plane: 13.66°, E-plane: 13.28°.

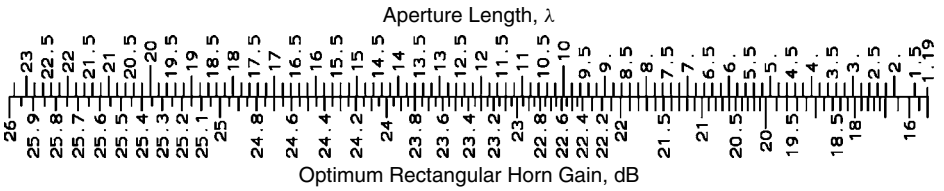
Scales 7-1 to 7-3 provide the dimensions of the optimum rectangular horn for a given gain. During their generation, a waveguide with a 2 : 1 aspect was used, but they are close to the proper values for nearby aspects. They design horns to within 0.1 dB. These scales produce short rapidly flaring horns for low-gain antennas. In these cases it is better to deviate from the optimum design that gives the lightest horn for a given gain and design a horn with a small value of S. Scales 7-4 to 7-6 give designs for $S = 0.1$ to be used for low-gain designs.



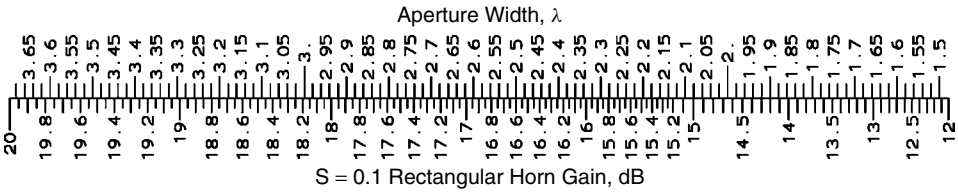
SCALE 7-1 Aperture width of an optimum pyramidal horn for a 2 : 1-aspect waveguide.



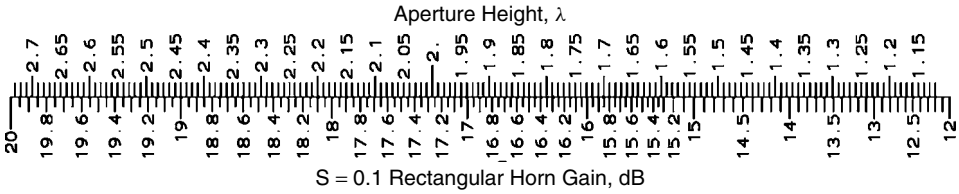
SCALE 7-2 Aperture height of an optimum pyramidal horn for a 2 : 1-aspect waveguide.



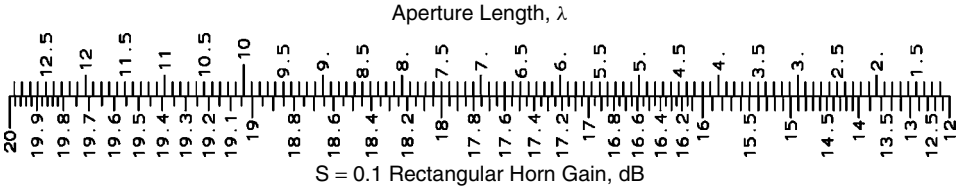
SCALE 7-3 Axial length of an optimum pyramidal horn for a 2 : 1-aspect waveguide.



SCALE 7-4 Aperture width of a pyramidal horn with $S = 0.1$ connected to a 2 : 1 waveguide.



SCALE 7-5 Aperture height of a pyramidal horn with $S = 0.1$ connected to a 2 : 1 waveguide.



SCALE 7-6 Axial length of a pyramidal horn with $S = 0.1$ connected to a 2 : 1 waveguide.

7-1.3 Designing to Given Beamwidths

The beamwidths in the two planes of a rectangular horn can be designed independently. The axial lengths in the two planes must be equal if the design is to be realizable, but the aperture width and height can be adjusted to give the desired beamwidths. We pick S in one plane and then vary S in the other plane to produce the required beamwidth and the same axial length as in the first plane. Since the first S is arbitrary, the design is not unique, but in only a limited range of S will designs be realizable.

Example Design a rectangular horn for the following 10-dB beamwidths: 30° H-plane and 70° E-plane at 7 GHz using a 3.5-cm × 1.75-cm waveguide.

Since the H-plane has the narrower beamwidth and therefore the wider aperture, we use it to determine length. Pick $S_h = 0.20$ (an arbitrary choice). The obliquity factor at 15° is 0.15 dB. When using Table 7-1 we must design for wider than a 30° beamwidth to compensate for the obliquity factor:

$$\mathrm{BW}_d = \sqrt{\frac{10.15}{10}}(30^\circ) = 30.22^\circ$$

The horn width to provide that beamwidth is

$$\frac{W}{\lambda} = \frac{1.0949}{\sin(\mathrm{BW}_d/2)} = 4.200$$

$$W = 18.00\,\text{cm at } 7\,\text{GHz} \qquad R_h = 47.25 \quad \text{[Eq. (7.2)]}$$

Use Eq. (7-13*a*) to determine the axial length: $L_h = L = 37.36\,\text{cm}$. Because the E-plane beamwidth is wider than the H-plane beamwidth, the E-plane aperture will be smaller. We try a smaller value of S_e than S_h for our initial guess: $S_e = 0.04$. The obliquity factor at 35° adds 0.82 dB to the pattern loss and requires a larger design beamwidth.

$$\text{BW}_d = \sqrt{\frac{10.82}{10}}70^\circ = 72.82^\circ$$

$$\frac{H}{\lambda} = \frac{0.7405}{\sin(\text{BW}_d/2)} = 1.248 \quad \text{(Table 7-2)} \qquad H = 5.246$$

We use Eqs. (7-2) and (7-13*b*) to calculate slant radius and axial length: $R_e = 20.84\,\text{cm}$ and $L_e = 13.90\,\text{cm}$. The axial lengths in the two planes do not match, so we pick a smaller S_e because the E-plane is shorter than the H-plane beamwidth. At $S_e = 0.02$, $H = 5.337\,\text{cm}$, $R_e = 41.54\,\text{cm}$, and $L_e = 27.86\,\text{cm}$. L_e has doubled when S_e changes from 0.04 to 0.02, but H changes by only 0.01 cm. We change our method and pick $H = 5.33\,\text{cm}$ and force R_e to give the same axial length as the H-plane: $R_e = 55.69\,\text{cm}$ [from Eq. (7-14)] or $S_e = 0.0149$.

7-1.4 Phase Center

We define the phase center as the point from which it appears that an antenna radiates spherical waves. Measurements show that the phase center is seldom a unique point in a plane, but depends on the pattern angle. The E- and H-plane phase centers will also be unequal in general. Usually, they are extremes, and the axial position varies elliptically between the planes. Even with the phase-center location fuzzy, it is a useful point. We place the phase center of a feed at the focus of a parabolic reflector to minimize the reflector aperture phase error loss.

An aperture with a quadratic phase distribution appears to be radiating from a point behind the aperture. Without quadratic phase error ($S = 0$), the phase center is located at the aperture plane. Increasing S moves the phase center toward the apex of the horn. Muehldorf [10] has calculated the phase-center location as a function of S, and Table 7-3 summarizes his results. The phase center located inside the aperture is given as a ratio of the slant length.

TABLE 7-3 Phase-Center Axial Location of a Rectangular Horn (TE_{10} Mode) Behind the Aperture as a Ratio of the Slant Radius

S	H-Plane L_{ph}/R_h	E-plane L_{ph}/R_e	S	H-plane L_{ph}/R_h	E-plane L_{ph}/R_e
0.00	0.0	0.0	0.28	0.258	0.572
0.04	0.0054	0.011	0.32	0.334	0.755
0.08	0.022	0.045	0.36	0.418	
0.12	0.048	0.102	0.40	0.508	
0.16	0.086	0.182	0.44	0.605	
0.20	0.134	0.286	0.48	0.705	
0.24	0.191	0.416	0.52	0.808	

Example Calculate the phase-center location for the beamwidth design example above.

$$S_h = 0.20 \qquad R_h = 47.25\,\text{cm}$$
$$S_e = 0.015 \qquad R_e = 55.69\,\text{cm}$$

From Table 7-3, we interpolate

$$H\text{-plane phase center} = 0.134(47.25\,\text{cm}) = 6.33\,\text{cm}$$
$$E\text{-plane phase center} = 0.004(55.69\,\text{cm}) = 0.22\,\text{cm}$$

The difference in the phase-center location in the two planes is 1.43λ.

As in the example, antennas with widely differing beamwidths will have widely separated phase centers.

7-2 CIRCULAR-APERTURE HORN

With a circular-aperture horn, we lose independent control of the beamwidths in the principal planes. The circular waveguide can support any orientation of the electric field and thereby allows any polarization in the horn. We use the same aperture field method as with the rectangular horn and the waveguide mode determines the aperture amplitude. The cone of the horn projects to a point in the feed waveguide where we assume a point source radiating to the aperture. The aperture phase is approximately quadratic. The waveguide fields are given by

$$E_\rho = \frac{E_0}{\rho} J_1\left(\frac{x'_{11}\rho}{a}\right)\cos\phi_c$$
$$E_{\phi_c} = -\frac{E_0 x'_{11}}{a} J'_1\left(\frac{x'_{11}\rho}{a}\right)\sin\phi_c \tag{7-16}$$

where J_1 is the Bessel function, ρ the radial component in the waveguide, a the radius, and ϕ_c the cylindrical coordinate. x'_{11} (1.841) is the first zero of $J'_1(x)$. Equation (7-16) has its maximum electric field directed along the $\phi_c = 0$ plane.

We add the quadratic phase factor to Eq. (7-16) and calculate the Fourier transform on the circular aperture to determine the far field. The direction of the electric field changes from point to point in the aperture. For a given direction (θ, ϕ_c) we must project the fields in the aperture onto the $\hat{\theta}$ and $\hat{\phi}$ directions before integrating over the aperture:

$$E_\theta = E_0 \int_0^{2\pi}\int_0^a \left[\frac{J_1(x'_{11}\rho/a)}{\rho}\cos\phi_c \frac{\hat{\theta}\bullet\hat{\rho}}{\cos\theta} - \frac{x'_{11}}{a} J'_1\left(\frac{x'_{11}\rho}{a}\right)\sin\phi_c \frac{\hat{\theta}\bullet\hat{\phi}_c}{\cos\theta}\right]$$
$$\times\, \rho \exp\left\{j\left[k\rho\sin\theta\cos(\phi-\phi_c) - 2\pi S\left(\frac{\rho}{a}\right)^2\right]\right\} d\rho\, d\phi_c \tag{7-17}$$

$$E_\phi = E_0 \int_0^{2\pi} \int_0^a \left[\frac{J_1(x'_{11}\rho/a)}{\rho} \cos\phi_c \hat{\phi} \cdot \hat{\rho} - \frac{x'_{11}}{a} J'_1 \left(\frac{x'_{11}\rho}{a} \right) \sin\phi_c \hat{\phi} \cdot \hat{\phi}_c \right]$$
$$\times \rho \exp\left\{ j \left[k\rho \sin\theta \cos(\phi - \phi_c) - 2\pi S \left(\frac{\rho}{a} \right)^2 \right] \right\} d\rho \, d\phi_c \tag{7-18}$$

$$\hat{\theta} \cdot \hat{\rho} = \cos\theta(\cos\phi \cos\phi_c + \sin\phi \sin\phi_c)$$
$$\hat{\theta} \cdot \hat{\phi}_c = \cos\theta(\sin\phi \cos\phi_c - \cos\phi \sin\phi_c)$$
$$\hat{\phi} \cdot \hat{\rho} = \cos\phi \sin\phi_c - \sin\phi \cos\phi_c$$
$$\hat{\phi} \cdot \hat{\phi}_c = \cos\phi \cos\phi_c + \sin\phi \sin\phi_c$$

By a suitable change of variables in the integrals, universal radiation patterns can be generated for the E-and H-planes (Figures 7-5 and 7-6). The equality of S in the two planes ties the curves together. The axis is the k-space variable. We can calculate a few pattern points for a given horn with those curves if we remember to add the obliquity factor to the values taken from the curves.

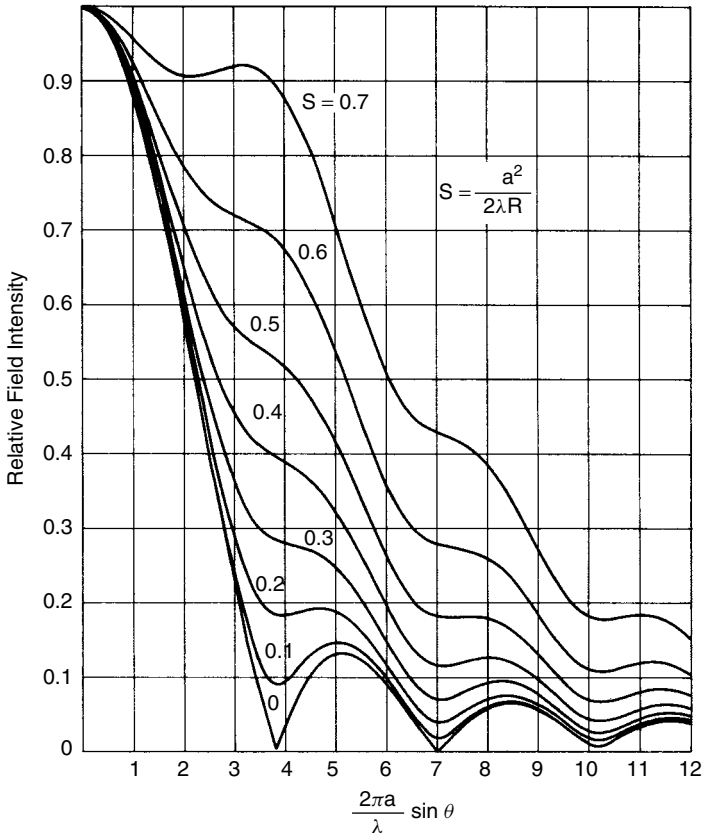


FIGURE 7-5 E-plane universal pattern of a circular, TE_{11} mode. (From T. Milligan, Universal patterns ease circular horn design, *Microwaves,* vol. 20, no. 3, March 1981, p. 84.)

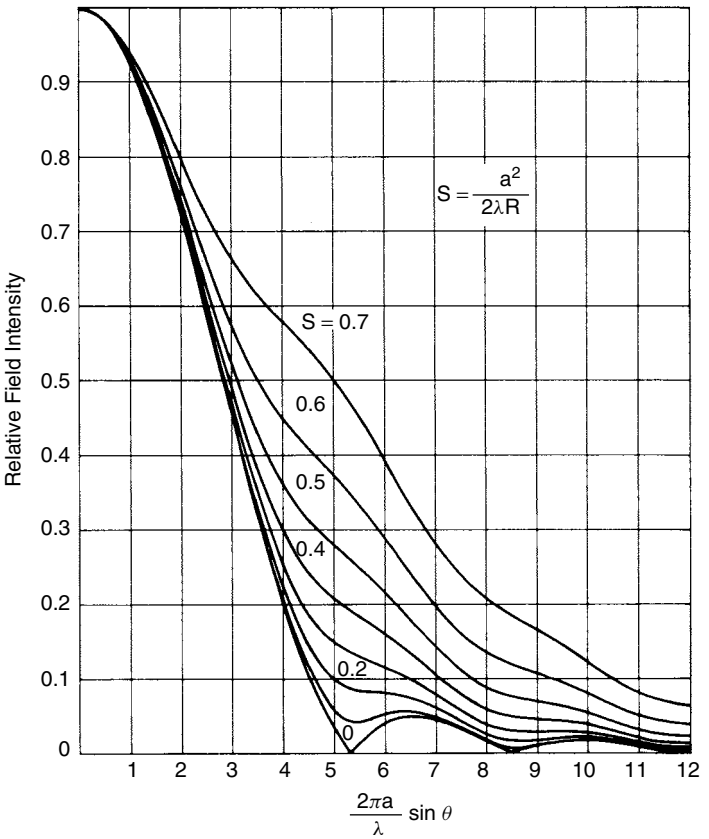


FIGURE 7-6 H-plane universal pattern of a circular, TE_{11} mode. (From T. Milligan, Universal patterns ease circular horn design, *Microwaves,* vol. 20, no. 3, March 1981, p. 84.)

Example A horn has an aperture radius of 12 cm and a slant radius of 50 cm. Compute the pattern level at $\theta = 20^\circ$ at 5 GHz.

From Figures 7-5 and 7-6 we interpolate the pattern voltage level:

$$H\text{-plane level} = 0.18 \quad (-14.7\,\text{dB}) \qquad E\text{-plane level} = 0.22 \quad (-13.1\,\text{dB})$$

The obliquity factor is $20\log[(1 + \cos 20^\circ)/2] = -0.3\,\text{dB}$, and the plane level at 20° becomes

$$H\text{-plane level} = -15\,\text{dB} \qquad E\text{-plane level} = -13.4\,\text{dB}$$

7-2.1 Beamwidth

Table 7-4 lists the 3- and 10-dB points from Figures 7-5 and 7-6. We can use them to compute beamwidths from dimensions.

Example Calculate 10-dB beamwidths of the horn in the example above. $S = 0.24$, $a = 12\,\text{cm}$, and $\lambda = 6\,\text{cm}$.

TABLE 7-4 Circular-Horn Beamwidths, TE_{11} Mode

	$(2\pi a/\lambda)\sin\theta$				
	3 dB		10 dB		ATL + PEL
S	E-Plane	H-Plane	E-Plane	H-Plane	(dB)
0.00	1.6163	2.0376	2.7314	3.5189	0.77
0.04	1.6175	2.0380	2.7368	3.5211	0.80
0.08	1.6212	2.0391	2.7536	3.5278	0.86
0.12	1.6273	2.0410	2.7835	3.5393	0.96
0.16	1.6364	2.0438	2.8296	3.5563	1.11
0.20	1.6486	2.0477	2.8982	3.5799	1.30
0.24	1.6647	2.0527	3.0024	3.6115	1.54
0.28	1.6855	2.0592	3.1757	3.6536	1.82
0.32	1.7123	2.0676	3.5720	3.7099	2.15
0.36	1.7471	2.0783	4.6423	3.7863	2.53
0.40	1.7930	2.0920	5.0492	3.8933	2.96
0.44	1.8552	2.1100	5.3139	4.0504	3.45
0.48	1.9441	2.1335	5.5375	4.2967	3.99
0.52	2.0823	2.1652	5.7558	4.6962	4.59
0.56	2.3435	2.2089	6.0012	5.2173	5.28
0.60	3.4329	2.2712	6.3500	5.6872	5.98
0.64	4.3656	2.3652	7.6968	6.0863	6.79
0.68	4.8119	2.5195	8.4389	6.4622	7.66
0.72	5.1826	2.8181	8.8519	6.8672	8.62

From Table 7-4 we read the k-space values at 10 dB:

$$H\text{-plane } k\text{-space} = \frac{2\pi a}{\lambda}\sin\theta_h = 3.6115$$

$$E\text{-plane } k\text{-space} = \frac{2\pi a}{\lambda}\sin\theta_e = 3.0024$$

$$\text{BW}_h = 2\sin^{-1}\frac{3.6115}{4\pi} = 33.40^\circ \qquad \text{BW}_e = 2\sin^{-1}\frac{3.0024}{4\pi} = 27.65^\circ$$

We must add the obliquity factor to the 10-dB universal pattern level:

$$\frac{1+\cos 16.7^\circ}{2}: \quad 0.18\,\text{dB} \qquad H\text{-plane}$$

$$\frac{1+\cos 13.8^\circ}{2}: \quad 0.13\,\text{dB} \qquad E\text{-plane}$$

$$\text{BW}_{h10} = \sqrt{\frac{10}{10.18}}33.40^\circ = 33.10^\circ \qquad \text{BW}_{e10} = \sqrt{\frac{10}{10.13}}27.65^\circ = 27.48^\circ$$

We can also use Table 7-4 to design a horn to a given beamwidth, but we can design to only one plane. Any number of horns can be designed to a given beamwidth, since S is an independent parameter. Table 7-4 lists the combined amplitude taper loss and phase error loss as a function of S for the circular horn. With this table we can easily estimate the gain of a given horn or design a horn to a given gain.

Example Compute the gain of a horn with a 12-cm aperture radius and 50-cm slant radius at 5 GHz.

From the examples above we read $S = 0.24$ and $\lambda = 6\,\text{cm}$:

$$\begin{aligned}\text{gain} &= 20\log\frac{\pi D}{\lambda} - \text{GF} \qquad \text{where} \quad \text{GF} = \text{ATL} + \text{PEL (dB)} \\ &= 20\log\frac{24\pi}{6} - 1.54 = 20.4\,\text{dB}\end{aligned} \tag{7-19}$$

Example Design a circular horn at 8 GHz with a gain of 22 dB.

The quadratic phase distribution constant S is arbitrary. Pick $S = 0.20$. Rearrange Eq. (7-19) to find the diameter:

$$\begin{aligned}D &= \frac{\lambda}{\pi}\cdot 10^{(\text{gain}+\text{GF})/20} \\ &= \frac{3.75}{\pi}\cdot 10^{(22+1.30)/20} = 17.45\,\text{cm} \\ R &= \frac{D^2}{8\lambda S} = 50.77\,\text{cm}\end{aligned} \tag{7-20}$$

We can determine an optimum circular horn in the sense of minimizing the slant radius at a given gain. When we plot gain as a function of aperture radius for a fixed slant radius, we discover a broad region in which the gain peaks. By plotting a series of these lines with a voltage gain ordinate, we see that a single line corresponding to $S = 0.39$ can be drawn through the peaks. This is our optimum with GF = 2.85 dB (ATL + PEL).

Example Design an optimum horn at 8 GHz with gain of 22 dB.

From Eq. (7-20),

$$D = \frac{3.75}{\pi}\cdot 10^{(22+2.85)/20} = 20.86\,\text{cm} \qquad R = \frac{D^2}{8\lambda S} = \frac{20.86^2}{8(3.75)(0.39)} = 37.2\,\text{cm}$$

The optimum is quite broad. A horn designed with $S = 0.38$ has a 0.07-cm-longer slant radius and a 0.25-cm-smaller aperture diameter.

7-2.2 Phase Center

The phase-center location behind the aperture plane is a function of S. Table 7-5 lists the phase-center location as a ratio of the slant radius. As S increases, the phase center migrates toward the horn apex and the difference between the phase-center locations in the E- and H-planes increases.

Example Use Table 7-5 to compute phase-center locations in the E- and H-planes for the circular horns of the preceding two examples.

$$\begin{aligned}R = 50.77\,\text{cm} \quad S = 0.20 \\ H\text{-plane phase center} = 0.117(50.77) = 5.94\,\text{cm} \\ E\text{-plane phase center} = 0.305(50.77) = 15.48\,\text{cm}\end{aligned}$$

TABLE 7-5 Phase-Center Axial Location of a Circular-Waveguide Horn TE_{11} Mode Behind the Aperture as a Ratio of the Slant Radius

S	H-Plane L_{ph}/R_h	E-Plane L_{ph}/R_e	S	H-Plane L_{ph}/R_h	E-Plane L_{ph}/R_e
0.00	0.0	0.0	0.28	0.235	0.603
0.04	0.0046	0.012	0.32	0.310	0.782
0.08	0.018	0.048	0.36	0.397	0.801
0.12	0.042	0.109	0.40	0.496	0.809
0.16	0.075	0.194	0.44	0.604	0.836
0.20	0.117	0.305	0.48	0.715	0.872
0.24	0.171	0.416			

The phase centers differ by 2.5 wavelengths at 8 GHz. The optimum horn has the dimensions

$$R = 37.2\,\text{cm} \quad S = 0.39$$

$$H\text{-plane phase center} = 0.471(37.2) = 17.5\,\text{cm}$$

$$E\text{-plane phase center} = 0.807(37.2) = 30.0\,\text{cm}$$

The phase centers of the optimum horn differ by 3.3 wavelengths. The difference increases with increasing S.

7-3 CIRCULAR (CONICAL) CORRUGATED HORN

Normal smooth-wall horns present problems that can be eliminated by corrugating the walls. Many applications require dual linear or circular polarizations. The horn aperture must be square or circular and the beamwidths in the two planes are unequal. When the smooth-wall horn feeds a reflector, we have astigmatism (unequal phase centers in orthogonal planes). A horn has higher sidelobes in the E-plane than in the H-plane. Finally, the diffraction off E-plane walls causes backlobes. The aperture theory fails to predict them, but measurement or a GTD analysis shows them. Corrugating the walls can prevent all these problems.

Figure 7-7 shows the cross sections of two types of corrugated horns. The small-flare horn (*a*) is nominally the corrugated horn, and the wide-flare horn (*b*) is the scalar horn of Simmons and Kay [11]. Many papers on corrugated horns appear in Section VI of Love's [1] collection of papers. Thomas [12] provides a good design summary in a topic with many papers. The corrugations that extend circumferentially should be cut normal to the slant radius as in (*b*), but they may be cut normal to the axis (*a*) for small flare angles. Horns can be built either way, but when cut normal to the axis, the depth is different on the two sides.

The corrugated wall presents the same boundary conditions to the electric and magnetic fields when it is capacitive (slots $\lambda/4$ to $\lambda/2$ deep). When excited in the transition between the smooth-wall waveguide and the corrugated-wall cone, the TE_{11} and TM_{11} waveguide modes, have equal phase velocities. The combination of these

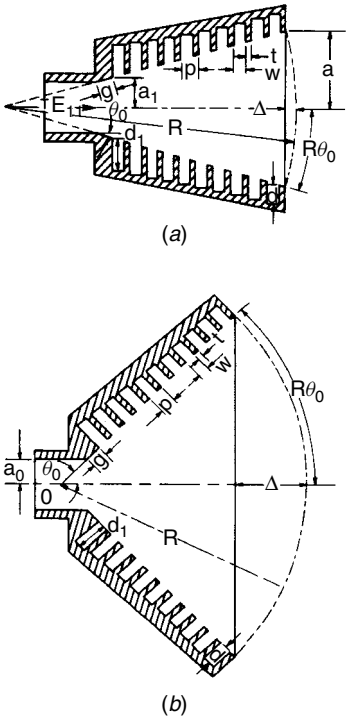


FIGURE 7-7 (*a*) Corrugated horn; (*b*) scalar horn. (From [12], © 1978 IEEE.)

modes forms the hybrid mode HE_{11} when the mode phases are equal. When the modes are out of phase by 180°, the hybrid mode is denoted EH_{11}. The ratio of the modes is called γ, and $\gamma = 1$ for the balanced HE_{11} mode. $\gamma = 0$ corresponds to having only the TM_{11} mode and $\gamma = \infty$ to having only the TE_{11} mode. $\gamma = 1$ occurs when the corrugation depth is $\lambda/4$, but the horn parameters vary slowly with changing γ [13]. We consider only $\gamma = 1$. Changing γ has its biggest effect on the cross-polarization [12,14].

When $\gamma = 1$, the amplitude of the aperture fields is given by [15]

$$E_\rho = J_0\left(\frac{x_{01}\rho}{a}\right)\cos\phi_c$$
$$E_\phi = -J_0\left(\frac{x_{01}\rho}{a}\right)\sin\phi_c \tag{7-21}$$

where $x_{01} = 2.405$ is the first zero of $J_0(x)$, the Bessel function. The fields vanish at the walls and prevent edge diffractions that produce sidelobes and backscatter. The lower-order slope diffractions still produce small sidelobes and backlobes, but we get H-plane-type lobes in all planes. In amplitude the aperture fields are symmetrical about the axis and all patterns through the cone axis are identical.

A Huygens source analysis of the aperture fields with a quadratic phase distribution produces Figure 7-8, valid when the 10-dB beamwidth is less than 74° [12]. For greater beamwidths the flange changes the beamwidths of the small-aperture horn in the two planes and we should use the scalar horn. Table 7-6 lists the 3-, 10-, and 20-dB points

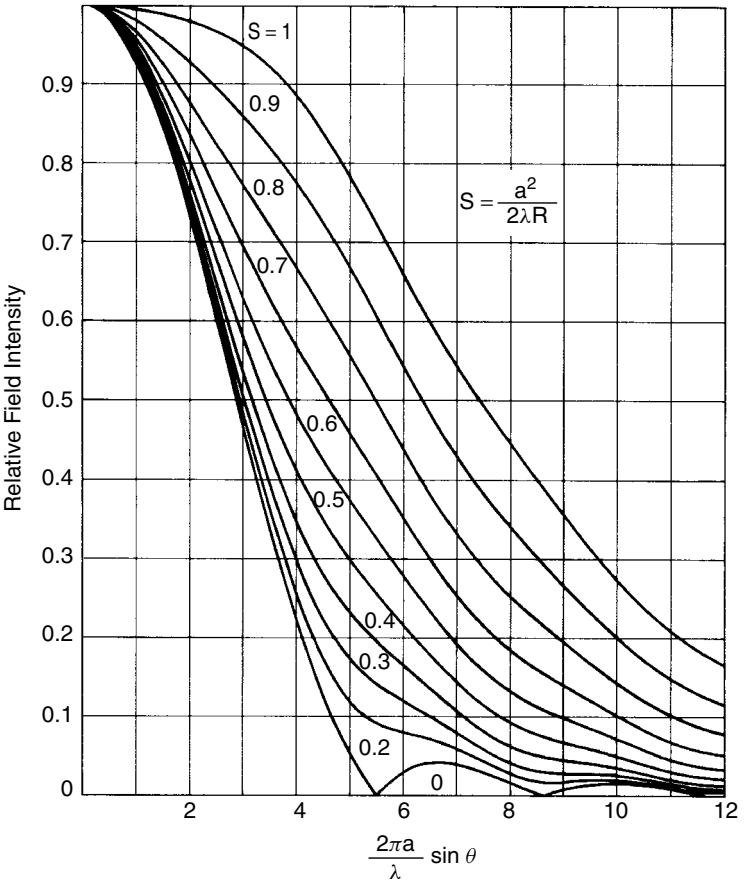


FIGURE 7-8 Universal pattern of a circular corrugated horn: HE_{11} mode.

TABLE 7-6 Resonant Circular Corrugated Horn Beamwidth Points $(2\pi a/\lambda)\sin\theta$, HE_{11} Mode

S	3 dB	10 dB	20 dB	ATL + PEL (dB)	S	3 dB	10 dB	20 dB	ATL + PEL (dB)
0.00	2.0779	3.5978	4.6711	1.60	0.52	2.3688	4.9532	7.9936	4.04
0.04	2.0791	3.6020	4.6878	1.62	0.56	2.4411	5.2720	8.4261	4.44
0.08	2.0827	3.6150	4.7405	1.66	0.60	2.5317	5.5878	8.9472	4.86
0.12	2.0887	3.6371	4.8387	1.73	0.64	2.6469	5.8913	9.4352	5.31
0.16	2.0974	3.6692	5.0061	1.83	0.68	2.7966	6.1877	9.8514	5.79
0.20	2.1088	3.7129	5.3052	1.96	0.72	2.9946	6.4896	10.2337	6.30
0.24	2.1234	3.7699	5.8451	2.12	0.76	3.2597	6.8134	10.6250	6.83
0.28	2.1415	3.8433	6.3379	2.30	0.80	3.6061	7.1788	11.0735	7.39
0.32	2.1637	3.9372	6.6613	2.52	0.84	4.0189	7.6042	11.6356	7.96
0.36	2.1906	4.0572	6.9179	2.76	0.88	4.4475	8.0852	12.2658	8.54
0.40	2.2231	4.2112	7.1534	3.04	0.92	4.8536	8.5773	12.8236	9.13
0.44	2.2624	4.4090	7.3939	3.34	0.96	5.2331	9.0395	13.3059	9.72
0.48	2.3103	4.6578	7.6633	3.68	1.00	5.5984	9.4701	13.7706	10.29

from Figure 7-8. We use the table to find beamwidths of given horns or design to given beamwidths. Table 7-6 also lists the sum of ATL and PEL(GF) for various S. We estimate gain or design to a given gain with this listing.

Example Calculate 10-dB beamwidth and gain of a corrugated conical horn with an aperture radius of 12 cm and a slant radius of 50 cm at 5 GHz:

$$S = \frac{a^2}{2\lambda R} = \frac{12^2}{2(6)(50)} = 0.24$$

From Table 7-7 we read the k-space value at 10 dB for $S = 0.24$:

$$\frac{2\pi a}{\lambda} \sin\theta = 3.7699 \quad \text{or} \quad \theta = 17.46^\circ$$

We include the obliquity factor, since the pattern loss will be greater than 10 dB at $\theta = 17.46^\circ$:

$$(1 + \cos 17.46^\circ)/2 : -0.20\,\text{dB} \qquad \text{BW}_{10} = \sqrt{\frac{10}{10.20}}34.92^\circ = 34.57^\circ$$

A smooth-wall horn with the same dimensions has a similar H-plane beamwidth (33.4°).

We calculate gain from the distribution losses and aperture area:

$$\text{GF} = \text{ATL} + \text{PEL} = 2.12\,\text{dB} \qquad \text{gain} = 20\log\frac{\pi D}{\lambda} - \text{GF} = 19.86\,\text{dB}$$

The smooth-wall horn has about 0.5 dB more gain (20.4 dB).

Example Design a circular corrugated-wall horn at 8 GHz with a gain of 22 dB.

We use Eq. (7-20) with the GF from Table 7-6. Choose $S = 0.20$ (arbitrary):

$$D = \frac{3.75\,\text{cm}}{\pi} \cdot 10^{(22+1.96)/20} = 18.83\,\text{cm}$$

$$R = \frac{D^2}{8\lambda S} = 59.10\,\text{cm}$$

TABLE 7-7 Phase-Center Axial Location of a Circular Corrugated Horn (HE_{11} Mode) Behind the Aperture as a Ratio of the Slant Length

S	L_p/R	S	L_p/R
0.00	0.0	0.36	0.386
0.04	0.005	0.40	0.464
0.08	0.020	0.44	0.542
0.12	0.045	0.48	0.614
0.16	0.080	0.52	0.673
0.20	0.124	0.56	0.718
0.24	0.178	0.60	0.753
0.28	0.240	0.64	0.783
0.32	0.310	0.68	0.811

The phase center, like that of other horns, starts at the aperture for $S = 0$ ($R = \infty$) and moves toward the apex as S increases. Table 7-7 lists the phase-center location as a ratio of the slant radius. Because the aperture distribution is the same along all radial lines of the aperture, the phase center is the same in all planes. We measure some variation, since the balance between modes will not be perfect, and some higher-order modes will be generated. The phase center moves least over a frequency band for long horns (small S). A wide-flare-angle horn has its phase center at the apex and is better described as a scalar horn.

7-3.1 Scalar Horn

A scalar horn has a wide half-flare angle, θ_0. Its beamwidth depends on the half-flare angle. Since the phase distribution in the aperture is large, there is an optimum diameter for a given flare angle. Table 7-8 lists the optimum diameter versus the flare angle. The beamwidth is approximately a linear function of the half-flare angle, θ_0, for the optimum horn:

$$\begin{aligned} \mathrm{BW}_{3\,\mathrm{dB}} &= 0.74\theta_0 \\ \mathrm{BW}_{10\,\mathrm{dB}} &= 1.51\theta_0 \\ \mathrm{BW}_{20\,\mathrm{dB}} &= 2.32\theta_0 \end{aligned} \tag{7-22}$$

A scalar horn has a wider bandwidth as a reflector feed than that of a small-flare-angle corrugated horn, because the phase center is fixed at the horn apex.

7-3.2 Corrugation Design

The corrugations present a capacitive reactance to the passing wave. When a corrugated surface is inductive, it will support surface waves. The depth of corrugations must be between $\lambda/4$ and $\lambda/2$. Less than $\lambda/4$ or greater than $\lambda/2$, it is inductive. Between $3\lambda/2$ and λ it will be capacitive again, but this second passband is seldom used. A quarter-wavelength corrugation depth balances the two modes and gives the best results. The corrugations need be only $\lambda/4$ at the aperture. Before the aperture we find it better to deepen the slots. Quarter-wavelength-deep corrugations mismatch the horn in the transition region, where the TM_{11} mode is generated from the TE_{11} mode and depths approaching $\lambda/2$ have the least effect on match.

TABLE 7-8 Optimum Diameter of a Scalar Horn

Half-Flare Angle, θ_0 (deg)	Aperture Diameter (λ)	Half-Flare Angle, θ_0 (deg)	Aperture Diameter (λ)
15	10.5	45	3.5
20	8.0	50	3.0
25	6.4	55	2.8
30	5.2	60	2.6
35	4.5	65	2.4
40	3.9	70	2.3

Source: [16, p. 429].

To design for a particular band, limited to about 1.5 : 1 for a good match, we design with tapered corrugation depths. We make the corrugations $\lambda/4$ deep at the aperture at the low-frequency end. The high-frequency band edge determines the corrugation depth at just short of $\lambda/2$ in the throat. The horn needs at least four corrugations per wavelength along the slant radius. The high-frequency end determines the corrugation spacing. The first few corrugations can be used to match the horn to the waveguide, and we can improve the match by shaping the corrugation widths [14]. The slots should be as wide as practical spacers will allow. Mechanical considerations, such as shock and vibration, will determine the limits on the thinness of spacers, but corrugations greatly increase the strength of the bell.

The circular geometry of the horn changes the corrugation depth necessary for the balanced HE_{11} mode from $\lambda/4$. An empirical formula for the depth is given by [17]

$$d = \frac{\lambda}{4} \exp\left(\frac{1}{2.5\ ka}\right) \qquad ka > 2 \tag{7-23}$$

We increase the depth slightly at the horn aperture.

7-3.3 Choke Horns

We can place corrugations in the flanges of small-aperture horns and design wide-beamwidth antennas with good pattern symmetry and low cross-polarization. The choke horn (Figure 7-9) is the limit of a scalar horn opened to $\theta_0 = 90°$. The corrugations consist of concentric rings about the aperture and are generally a quarter-wavelength deep. The best location for the corrugated rings may not be in the same plane as the aperture but instead somewhat behind as reported for a feed for $f/D = 0.3$ reflector [18;19, pp. 200–209].

The design for the reflector feed [18] used four corrugations. James [20] and Kumer [21] show that using only one corrugation is quite effective. More corrugations improve the design but only add marginally. Small apertures need the corrugations to reduce the cross-polarization that peaks in the diagonal planes between the E- and H-planes.

We usually assume a Huygens source in the aperture plane of the horn. This approximation collapses as we shrink the aperture to achieve wide beamwidths. In the limit

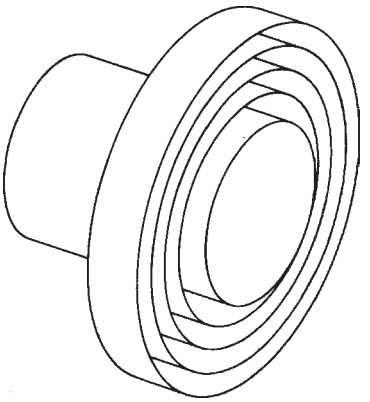


FIGURE 7-9 Choke horn.

we have only a slot analyzed from magnetic currents replacing the electric field in the aperture. The magnetic field is ignored in the slot while a Huygens source assumption is that the ratio of the electric to magnetic field is the same as the impedance of free space. Waveguides have high wave impedances, which implies small magnetic fields. To calculate the far field, Eq. (2-23) must be used with the actual ratio of fields in the aperture instead of Eq. (2-24), with its Huygens source approximation. Restricting the aperture dimensions to achieve wide beamwidths will limit the bandwidth as well as the cross-polarization isolation, because reducing volume raises Q.

7-3.4 Rectangular Corrugated Horns

We can design rectangular horns with corrugated walls, but we only need to cut corrugations in the E-plane walls to produce a cosine distribution in the E-plane. Only for dual polarization do we need corrugations in the H-plane walls. We analyze the horn as having an H-plane distribution (cosine) in both planes and use the results of Section 7-1. The larger aperture dimension in the diagonal plane decreases the beamwidth slightly, but the rectangular horn still provides an acceptable design. Both planes have a linear amplitude taper loss of 0.91 dB. We use the cosine column of Table 4-42 for the quadratic phase error loss. We design beamwidths using Table 7-1. Equalizing the distributions in both planes of square horns results in equal phase centers, given by Table 7-3 (H-plane).

Example Compute the gain of a square corrugated horn with an aperture width of 24 cm and a slant radius of 50 cm at 5 GHz.

From Eq. (7-1),

$$S = \frac{W^2}{8\lambda R} = \frac{24^2}{8(6)(50)} = 0.24$$

We use the cosine column of Table 4-42 for the phase error loss: $\text{PEL}_x = \text{PEL}_y = 0.42$ (cosine). The amplitude taper loss is the same in both planes: 0.91 dB.

$$\begin{aligned}\text{Gain} &= 10\log\frac{4\pi W^2}{\lambda^2} - \text{ATL}_x - \text{ATL}_y - \text{PEL}_x - \text{PEL}_y \\ &= 23.03 - 0.91 - 0.91 - 0.42 - 0.42 = 20.4\,\text{dB}\end{aligned}$$

A circular corrugated horn with a diameter equal to the width and having the same slant radius has a gain of 19.9 dB or 0.5 dB less. The larger aperture area increases the gain over the circular horn.

7-4 CORRUGATED GROUND PLANE

The corrugated surface (Figure 7-10) supports surface waves (TM) when the slot depth is less than $\lambda/4$ (inductive). As with the corrugated horn, we assume many slots per wavelength along the direction of propagation. The fields attenuate exponentially above the ends of the corrugations in a surface wave. We derive the fields from a potential function

$$\psi = A_1 \exp\left(\frac{-2\pi bx}{\lambda}\right)\exp(-jk_z z) \tag{7-24}$$

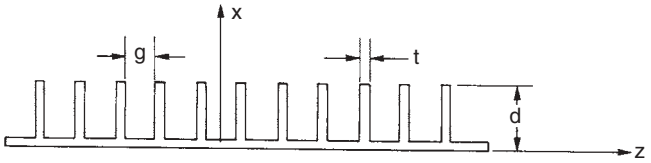


FIGURE 7-10 Corrugated ground plane.

above the corrugations, where A_1 is an amplitude constant, x the distance out of the corrugations, and $\alpha(= 2\pi b/\lambda)$ the attenuation constant of the fields above the corrugations. We expand the fields and take the ratio of the z-directed electric field to the y-directed magnetic field to find the wave impedance into the corrugated surface:

$$E_z = \frac{1}{j\omega\varepsilon_0}(k^2 - k_z^2)\psi = -\left(\frac{2\pi b}{\lambda}\right)^2 \frac{A_1}{j\omega\varepsilon_0} \exp\left(-\frac{2\pi b}{\lambda}\right) \exp(-jk_z z)$$

$$H_y = -\frac{\partial\psi}{\partial x} = \frac{2\pi b}{\lambda} A_1 \exp\left(-\frac{2\pi b}{\lambda}\right) \exp(-jk_z z)$$

$$Z_{-x} = \frac{E_z}{H_y} = \frac{j(2\pi b/\lambda)}{\omega\varepsilon_0} = \frac{j(2\pi b/\lambda)\sqrt{\mu_0}}{\omega\sqrt{\varepsilon_0\mu_0}\sqrt{\varepsilon_0}} = \frac{j(kb)}{k}\eta = jb\eta \tag{7-25}$$

where η is the impedance of free space and b is related to α [Eq. (7-24)]. The structure must present this impedance to the wave. The corrugated surface is a parallel-plate transmission line to E_z, and it presents a per unit length impedance of

$$Z_c = j\eta \tan kd \tag{7-26}$$

where d is the corrugation depth. We equate Eqs. (7-25) and (7-26) to determine the constant b:

$$b = \tan kd \tag{7-27}$$

We use Eq. (7-27) in Eq. (10-16) for the relative propagation constant:

$$P = \sqrt{1 + b^2} = \sqrt{1 + \tan^2 kd} \tag{7-28}$$

We include the effect of the corrugation thickness by averaging between the parallel-plate impedance and the zero impedance along the corrugation edges. Equation (7-28) becomes

$$P = \sqrt{1 + \left(\frac{g}{g+t}\right)^2 \tan^2 kd} \tag{7-29}$$

where g is the gap distance and t is the corrugation thickness. P increases without bound as the depth d approaches $\lambda/4$. The fields become tightly bound to the surface and attenuate rapidly to zero above the corrugations—the normal electric field vanishes as in a corrugated horn E-plane wall. We design the depth of the corrugations by using

$$d = \frac{\lambda}{2\pi} \tan^{-1} \frac{g+t}{g\sqrt{P^2 - 1}} \tag{7-30}$$

When the corrugation depth approaches $\lambda/4$, the surface impedance [Eq. (7-26)] approaches infinity and the tangential magnetic field vanishes on the surface to create an artificial PMC (Section 2-3) for waves polarized along the z-axis. The reflection coefficient is $+1$ instead of -1 for the PEC surface. Waves polarized along the y-axis encounter closely spaced corrugations that approximately produce a PEC surface with the usual metal wall reflection coefficient of -1. Whereas a PEC reflects an incident circularly polarized wave with opposite sense of circular polarization, the artificial PMC (soft) surface [22, pp. 276–280] reflects the wave with the same sense of polarization. We can use these surfaces to shape the pattern of a wide-beamwidth circularly polarized antenna to narrow the beamwidth without generating the opposite sense polarization, which would be generated by metal walls.

A ground plane covered with circular coaxial corrugations $\lambda/4$ deep reduces the edge diffraction that produces a large backlobe for a monopole antenna mounted in the center (Figure 5-23). The artificial soft wall causes the reduction of circumferential magnetic fields and the associated GTD diffraction (Section 2-7.11) [23]. This increases the forward gain by reducing the backlobe. It is not necessary to corrugate the entire top surface. Figure 7-11 illustrates a surface with only two coaxial corrugations around the outer rim. These reduce the backlobe for a dipole mounted over the ground plane without generating significant cross-polarization from a pair of orthogonal dipoles fed for a circular polarization. Corrugating the entire surface would cause radiation of cross-polarization because the region below the dipole pair radiates oppositely sensed circular polarization. The corrugated surface reflects the same sense of circular polarization as incident. The PEC surface reverses the sense of circular polarization of the reflected wave and both waves add. The corrugations only reduce the backlobe. The choke horn uses the same type of corrugations to reduce the backlobe radiated from the small-diameter horn aperture.

The corrugations can be placed radially below the ground plane by using short-circuited radial transmission lines (Figure 7-12) and also reduce the backlobe. We

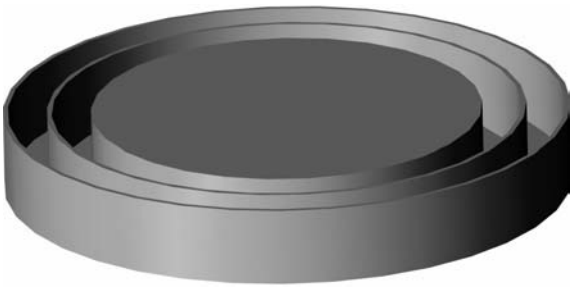


FIGURE 7-11 Ground plane with two coaxial corrugations to reduce edge diffraction.

FIGURE 7-12 Ground plane with short-circuited radial transmission-line corrugations.

TABLE 7-9 Radial Transmission Outer Choke Depth at Resonance

Outer Radius (λ)	Depth (λ)	Outer Radius (λ)	Depth (λ)
0.25	0.188	0.70	0.230
0.30	0.199	0.80	0.233
0.35	0.208	1.0	0.236
0.40	0.213	2.0	0.243
0.50	0.222	4.0	0.247
0.60	0.227		

compute the reactance at the outer radius from the following equation, which uses Bessel and Neumann functions:

$$X = j\frac{\eta b}{2\pi r}\frac{N_0(kr)J_0(kr_o) - J_0(kr)N_0(kr_o)}{J_1(kr)N_0(kr_o) - N_1(kr)J_0(kr_o)}$$

The short-circuited wall is located at radius r_o and the spacing between the plates is b. Reactance X grows rapidly as r approaches resonance. For a large outer radius the difference $r - r_o$ approaches λ/4 but is less for a small radius. Table 7-9 gives the difference in radius versus the outer radius for resonance.

Corrugations on the upper surface are more effective than radial corrugations, but the radial line chokes fit easily behind a small ground plane. In both cases the corrugations enhance radiation behind the ground plane at frequencies below resonance λ/4 depths, because a surface wave is generated along the corrugations. Corrugated surfaces are useful structures because they can be used to enhance or reduce radiation, depending on their depth.

7-5 GAUSSIAN BEAM

Corrugated horns and simple reflector feeds can be approximated with Gaussian beams. An infinite circularly symmetrical Gaussian aperture distribution located in the $x-y$ plane radiates a Gaussian beam along the z-axis. The radial exponent of the Gaussian distribution determines the spread of the wave as it propagates along the z-axis. We use the distribution to calculate the radiation pattern and then add the Huygens source (Section 2-2.2) for polarization. The analysis is divided into far- and near-field approximations. The near-field approximation consists of a paraxial wave. The Gaussian beam satisfies Maxwell's equations by using the free-space Helmholtz equation and produces correct patterns when applied with physical optics (PO).

The free-space Green's function satisfies the Helmholtz equation: e^{-jkR}/R. We derive the Gaussian beam from a point source placed at a complex position along the z-axis: $z_0 = -jb$. A source at this position produces a Gaussian distribution in the $z = 0$ plane.

$$\exp\left(\frac{-\rho^2}{W_0^2}\right) \quad \text{with} \quad \rho^2 = x^2 + y^2$$

W_0 is the beam waist radius where amplitude has dropped by $1/e$. We relate the waist radius W_0 to the position b by [24, pp. 80–90]

$$W_0^2 = \frac{2b}{k} \qquad \text{where} \quad k = \frac{2\pi}{\lambda}$$

As the wave propagates along the z-axis, its amplitude retains the Gaussian distribution in the radial direction ρ but the waist spreads:

$$W^2(z) = W_0^2 \left[1 + \left(\frac{z}{b}\right)^2\right]$$

The waist surface is a hyperboloid with a ring focus at radius b located at $z = 0$. The wave amplitude reduces by the ratio of the waists and combines with the radial Gaussian distribution:

$$\frac{W_0}{W(z)} \exp\left[-\frac{\rho^2}{W^2(z)}\right]$$

The phase of the paraxial (near-field) wave has two terms. The first is the normal z-directed wave phase $\exp(-jkz)$ and the second is a quadratic phase term that arises from the complex location of the point source at $z = -jb$. The quadratic phase term slant radius depends on the location along the z-axis:

$$R_c(z) = z\left[1 + \left(\frac{b}{z}\right)^2\right]$$

The paraxial Gaussian beam has an additional slippage phase term $\zeta(z) = \tan^{-1}(z/b)$. The paraxial Gaussian beam phase term is the sum

$$\exp\left[-jkz - jk\frac{\rho^2}{2R_c(z)} + j\zeta(z)\right]$$

The constant phase (eikonal) surfaces between the hyperboloid amplitude surfaces are ellipsoids with a ring focus at radius b located at $z = 0$. At $z = 0$ the eikonal surface is planar. We combine the amplitude and phase terms for the complete paraxial Gaussian beam equation:

$$-jE_0 \cos^2\frac{\theta}{2}\frac{W_0}{W(z)} \exp\left[-\frac{\rho^2}{W^2(z)}\right] \exp\left[-jkz - jk\frac{\rho^2}{2R_c(z)} + j\zeta(z)\right] \times (\hat{\theta}\cos\phi - \hat{\phi}\sin\phi) \tag{7-31}$$

The Huygens source polarization for an x-directed wave [Eq. (1-38)] and the obliquity factor [Eq. (2-14)] have been added to Eq. (7-31). We determine the constant E_0 by equating the radiation between this paraxial beam and the far-field expression for a Gaussian beam with a given input power. The recommended distance to equate the two representations is $z = 200W_0^2/\lambda$.

We calculate the far-field Gaussian beam by substituting the point source position into e^{-jkR}/R and approximating R with a far-field expression [25, pp. 96–106]:

$$R = \sqrt{x^2 + y^2 + z^2 - b^2 + j2bz} = \sqrt{r^2 - b^2 + j2br\cos\theta} \tag{7-32}$$

In the far field we can ignore b^2 and expand Eq. (7-32) in a Taylor series and retain the first two terms, which reduces e^{-jkR}/R to $e^{kb\cos\theta}e^{-jkr}/r$. We combine this term with the Huygens source radiation to produce the far-field Gaussian beam equation for an x-directed linear polarization in the aperture normalized at $\theta = 0$ to directivity:

$$\mathbf{E}(r, \theta, \phi) = \sqrt{\frac{P_0 \cdot \text{directivity} \cdot \eta}{4\pi}} \cos^2 \frac{\theta}{2} e^{kb(\cos\theta - 1)} (\hat{\theta} \cos\phi - \hat{\phi} \sin\phi) \frac{e^{-jkr}}{r} \quad (7\text{-}33)$$

The directivity is found by integrating the pattern of Eq. (7-33):

$$\text{directivity} = \frac{4(2\,kb)^2}{2(2\,kb) - 2 + 1/(2\,kb) - e^{-2(2\,kb)}/(2\,kb)} \quad (7\text{-}34)$$

Scale 7-7 gives the relationship between gain and the 10-dB beamwidth for a Gaussian beam.

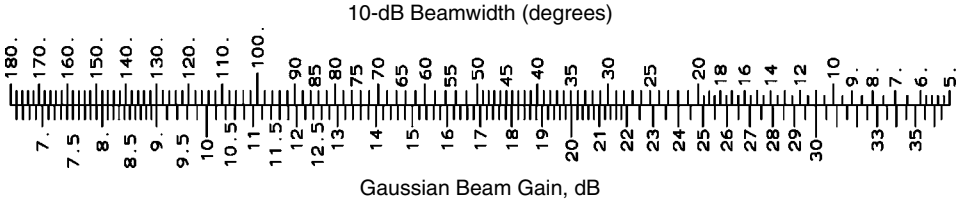
Given the beamwidth (BW) at a given level L(dB), we solve Eq. (7-33) for the complex-plane point source position b:

$$b = \frac{2\log[\cos(\text{BW}/4)] + |L/20|}{k[1 - \cos(\text{BW}/2)]\log e} \quad (7\text{-}35)$$

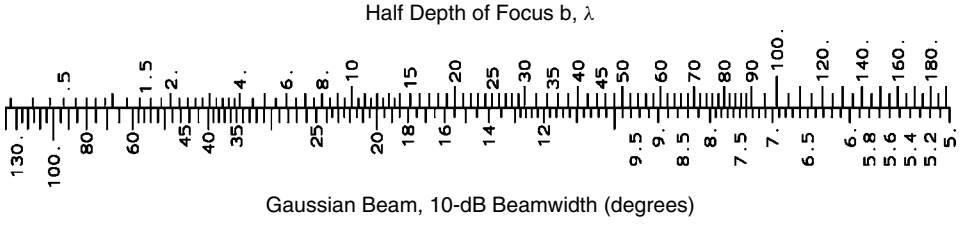
Scale 7-8 relates Gaussian beam half-depth of focus, b, to its 10-dB beamwidth, and Scale 7-9 gives the minimum waist diameter.

We simplify the expression for the Gaussian beam for small angles by expanding $\cos\theta$ in a Taylor series $\cos\theta \approx 1 - \theta^2/2$, which reduces Eq. (7-33):

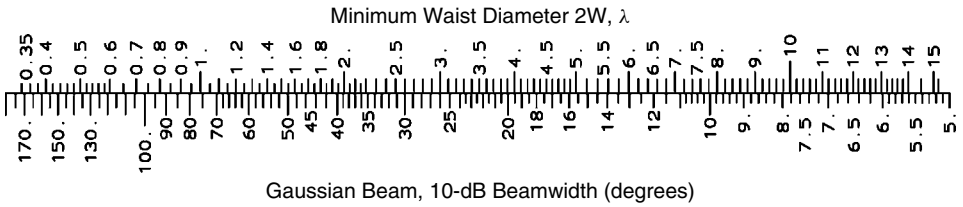
$$\mathbf{E}(r, \theta, \phi) = E_0 \cos^2 \frac{\theta}{2} e^{-(\theta/\theta_0)^2} (\hat{\theta} \cos\phi - \hat{\phi} \sin\phi) \frac{e^{-jkr}}{r} \quad (7\text{-}36)$$



SCALE 7-7 Gaussian beam gain compared to a 10-dB beamwidth.



SCALE 7-8 Gaussian beam half-depth of focus, b, compared to a 10-dB beamwidth.



SCALE 7-9 Gaussian beam minimum waist diameter compared to a 10-dB beamwidth.

The angle θ_0 is the beam divergence [24, pp. 80–90], given by

$$\theta_0 = \frac{2}{kW_0} = \sqrt{\frac{2}{kb}}$$

Equation (7-36) cannot be used beyond θ_0 because it is based on a small-angle approximation.

We can use a Gaussian beam to approximate the pattern of a corrugated horn [26, pp. 170–176]. The minimum waist is located behind the horn aperture L_p, the phase-center distance given the aperture radius a and the slant radius R:

$$L_p = \frac{R}{1 + [2R/k(0.644a)^2]^2} \tag{7-37}$$

L_p is the location of $z = 0$ of the Gaussian beam. The minimum waist radius W_0 is given by

$$W_0 = \frac{0.644a}{1 + [k(0.644a)^2/2R]^2} \qquad b = \frac{W_0^2 k}{2} \tag{7-38}$$

For a 22 dB-gain corrugated horn, Eq. (7-38) produces a Gaussian beam with the same gain as the horn for $S = 0.134$. For different values of S, Eq. (7-38) gives only approximate Gaussian beams to match the gain of corrugated horns. The Gaussian beam has a 10-dB beamwidth of 27.5° and the corrugated horn has a beamwidth of 27.2°. The phase center of the Gaussian beam given by Eq. (7-37) is 2.44λ behind the aperture and the actual horn phase center is at 0.89λ. The Gaussian beam approximation finds the near-field pattern of the corrugated horn because it includes the finite waist size instead of assuming a point source at the phase center of the horn. A PO analysis using the equivalent currents in the aperture [27, pp. 141–156] also finds the near-field pattern but requires greater calculation effort.

7-6 RIDGED WAVEGUIDE HORNS

Inserting ridges in the E-plane of a waveguide lowers the cutoff frequency compared to a waveguide of the same width. The ridges raise the cutoff frequencies of the next two higher modes and can produce a waveguide that operates over a 10 : 1 frequency range or more. If we use this as the input waveguide to a horn and taper the ridges until they do not block the horn aperture, the horn radiates a pattern similar to the smooth-wall horn. Near the aperture the horn can support many higher-order waveguide modes

as frequency increases. The horn generates some higher-mode content to the fields which distorts the pattern over narrow frequency ranges, but for many applications such distortions are acceptable. Initial designs [28] used dual ridges for a single linear polarization, while later designs increased the number of ridges to four (quad-ridge) to allow for dual linear (or circular) polarization.

Design concentrates on the input waveguide dimensions. We apply transverse resonance to the waveguide to calculate its cutoff frequencies. A rectangular waveguide with the electric field parallel to the narrow wall can be considered as a parallel-plate transmission line with the wave traveling between the two narrow wall shorts at cutoff (see Section 5-24). The parallel-plate transmission-line impedance is ηb for a height of b (meters). The lowest-order mode cutoff frequency for a normal rectangular waveguide occurs when the width $a = \lambda/2$. The transverse resonance method considers half the width as a transmission line and cutoff occurs when the impedance at the centerline is an open circuit (odd-order mode) or a short circuit (even-order mode) (i.e., $a/2 = N\lambda/4$) for mode TE_{N0}. Of course, we ignore the impedance of the parallel-plate line because it is uniform.

Figure 7-13*a* shows the cross section of a dual-ridged waveguide. The diagram illustrates feeding the waveguide with a coaxial line running through the center of one ridge. The center conductor extends across the gap to feed the second ridge. The center pin does not need to touch the second ridge but can be coupled capacitively. The transverse resonance circuit of a dual-ridged horn used to determine cutoff frequencies consists of two transmission-line segments with a shunt capacitor due to the step. The capacitance depends on the ratio of the heights $\alpha = b_2/b_1$, where $b_2 < b_1$ [29]:

$$C = \frac{\varepsilon_0}{\pi}\left(\frac{\alpha^2+1}{\alpha}\cosh^{-1}\frac{1+\alpha^2}{1-\alpha^2} - 2\ln\frac{4\alpha}{1-\alpha^2}\right) \tag{7-39}$$

For the dual-ridged waveguide we analytically place a ground plane halfway across the waveguide E-plane and divide the waveguide into two half-height waveguides. Later we will consider the impedance, and the total impedance of the guide is these two

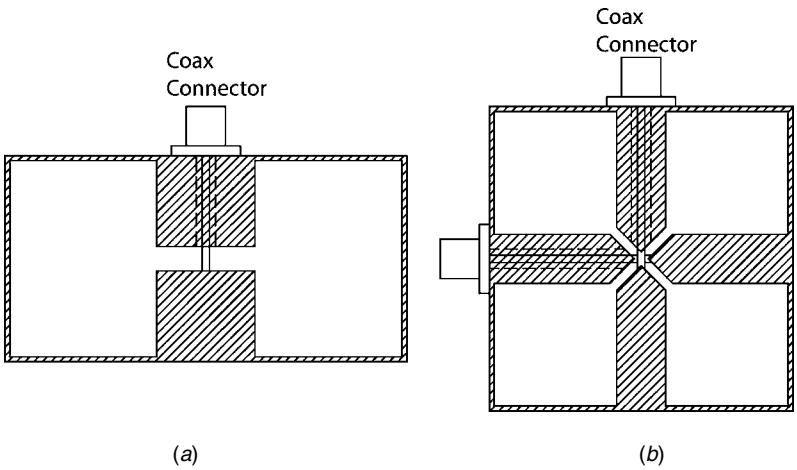


FIGURE 7-13 Coaxial feeds of ridged waveguides: (*a*) dual ridge; (*b*) quad ridge.

half-height guides in series. Given the waveguide width a_1 and height $2b_1$, and the ridge width a_2 with gap $2b_2$, we solve for the cutoff frequency using a transcendental equation in admittance at the transition point between the two half-height waveguides. Odd-order TE modes have a virtual open circuit in the center of the ridge and a short circuit at the wall. Cutoff occurs for $k_c = 2\pi/\lambda_c = 2\pi f_c/c$ for c equal to the speed of light [30]:

$$\frac{\tan(k_c a_2/2)}{\eta b_2} + k_c c C - \frac{\cot[k_c(a_1 - a_2)/2]}{\eta b_1} = 0 \tag{7-40}$$

We solve Eq. (7-40) numerically for k_c for odd-order modes. The even modes have a virtual short circuit in the center, which leads to a similar equation for the cutoff number k_c:

$$-\frac{\cot(k_c a_2/2)}{\eta b_2} + k_c c C - \frac{\cot[k_c(a_1 - a_2)/2]}{\eta b_1} = 0 \tag{7-41}$$

We use Eq. (7-40) to calculate the cutoff wavelengths of modes TE_{10} and TE_{30} and Eq. (7-41) to compute the cutoff wavelength of mode TE_{20} for given dimensions.

We design the waveguide to have a suitable low-frequency cutoff with an impedance equal to the input coax, whose outer conductor is connected to one ridge, with the center conductor jumping the gap to feed the other. The impedance at an infinite frequency is given by the equations

$$\begin{aligned} Y_\infty = \frac{1}{k\eta b_2}\Bigg\{ &\frac{ka_2}{4} + \frac{\sin ka_2}{4} + \frac{b_2}{b_1}\frac{\cos^2(ka_2/2)}{\sin^2(ka_1/2)}\left(\frac{ka_1}{4} - \frac{\sin ka_1}{4}\right) \\ &+ \frac{2b_2}{\lambda}\ln\left[\csc\left(\frac{\pi}{2}\frac{b_2}{b_1}\right)\cos^2\frac{ka_2}{2}\right]\Bigg\} \\ Z_\infty = \frac{1}{Y_\infty} \end{aligned} \tag{7-42}$$

The impedance at a finite frequency increases:

$$Z_0 = \frac{Z_\infty}{\sqrt{1 - (f_c/f)^2}} \tag{7-43}$$

An approximate value for the gap can be found from the impedance of microstrip. The infinite impedance equals slightly less than twice the impedance of microstrip the same width as the ridge with one-half the gap. The extra fringing capacitance between the sides of the ridges lowers the impedance compared to microstrip. You can use a microstrip line design program to find an approximate gap and a few evaluations of Eq. (7-42) to determine the correct gap. Design for Z_∞ because the impedance approaches Z_∞ rapidly as frequency increases by Eq. (7-43) and ridged horns operate over a large bandwidth.

Quad-ridged waveguide, illustrated in Figure 7-13*b*, requires modifications at the input to a horn. To achieve $Z_\infty = 50\,\Omega$, the gap must be reduced and the ridges made with a rooftop shape so that they fit within each other. The capacitance between the ridges for one polarization is a series combination of the two capacitors to the ridges for the second polarization. Similar to the dual-ridged waveguide, we divide the

waveguide along the centerline through the second set of ridges and analyze a single-ridged waveguide. Given a square waveguide with width w, ridge width s, and gap g between the ridges of different polarizations, the equivalent single-ridged waveguide has parameters given by the expressions

$$a_1 = w + s(\sqrt{2} - 1) \qquad a_2 = s\sqrt{2}$$
$$b_1 = w - s/2 \qquad b_2 = g \tag{7-44}$$

For the quad-ridged waveguide the infinite impedance equals slightly less than four times the impedance of microstrip the same width as the equivalent ridge a_2 with one-half the gap. We use the parameters of Eq. (7-44) in Eqs. (7-39) through (7-43) to find the parameters of quad-ridged waveguide. Figure 7-13*b* demonstrates that the feed pin of one coax passes over the other to reduce coupling between them. The difference in distance to the waveguide shorting wall for the two coaxial lines produces different impedances for the two inputs.

We can use the expressions above for circular waveguides. We design with a width equal to the diameter. The infinite impedance is lower by the factor $\pi/4$. The cutoff frequency is about 1.25 times the cutoff frequency of the equivalent square waveguide [31].

Figure 7-14 gives a cross-sectional drawing of a ridged horn and demonstrates the key elements of design. A coax is fed through the center of one ridge and the center conductor jumps the gap and feeds the second ridge. We locate the coax close to the end of the ridge truncated before it reaches the waveguide back wall short circuit, leaving a small gap. Without ridges the waveguide is cutoff at the low-frequency end of the horn operation. Operating the waveguide below its cutoff frequency does not

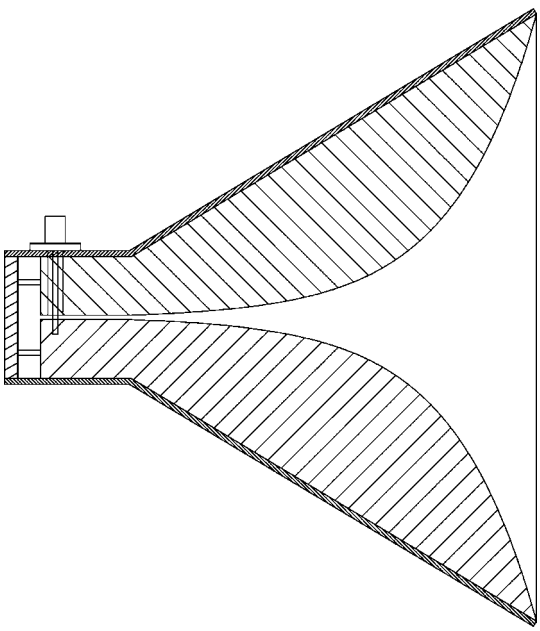


FIGURE 7-14 Dual-ridged waveguide horn cross section.

prevent the wave from reaching the back wall because the distance is short. The original horns [28] used waveguides in cutoff at the feed point at the lowest frequencies. By tapering the sidewalls the waveguide operates above cutoff in a short distance from the feed and the waves propagated to that region. Cutoff only means that a wave will not propagate in a long waveguide, but it attenuates as it moves along the waveguide. Figure 7-14 shows optional shorting pins between the back wall and the ridges. These prevent an additional resonance in impedance that may arise at a frequency when the height of a single ridge approaches $\lambda/2$. Not all designs need these pins.

We space the ridges to form a transmission line matched to the feed coax at the input. A uniform section of ridged waveguide extends to the throat of the horn. The horn shown in Figure 7-14 uses an exponentially tapered ridge that has an additional linear taper with slope 0.02 [28] empirically found to improve the impedance match. It would seem that designing a classical tapered impedance transformer would give a better impedance match, but the simple exponential physical taper produces an excellent impedance match. The gain of the horn falls short of the equivalent open horn because multiple modes are excited and beamwidth broadens. In a dual-ridged horn the power concentrates between the ridges in the E-plane, and we can replace the H-plane sidewalls with a few rods. We space the rods close enough to block radiation at the lower frequencies and allow high-frequency radiation through the spaces. Since the fields are concentrated between the ridges at high frequencies, the side H-plane walls have little effect on the pattern. A quad-ridged horn requires solid walls.

A circular quad-ridged horn was measured as a possible feed for a Cassegrain reflector from 6 to 18 GHz. The horn has a 13.2-cm aperture diameter and a 37.6-cm slant radius and operates from 2 to 18 GHz. Figure 7-15 plots the measured E- and H-plane 10-dB beamwidths along with the beamwidths of both smooth and corrugated wall horns of the same size. Neither smooth wall nor corrugated wall horns could be designed to operate over this wide bandwidth; they are shown only for comparison. The

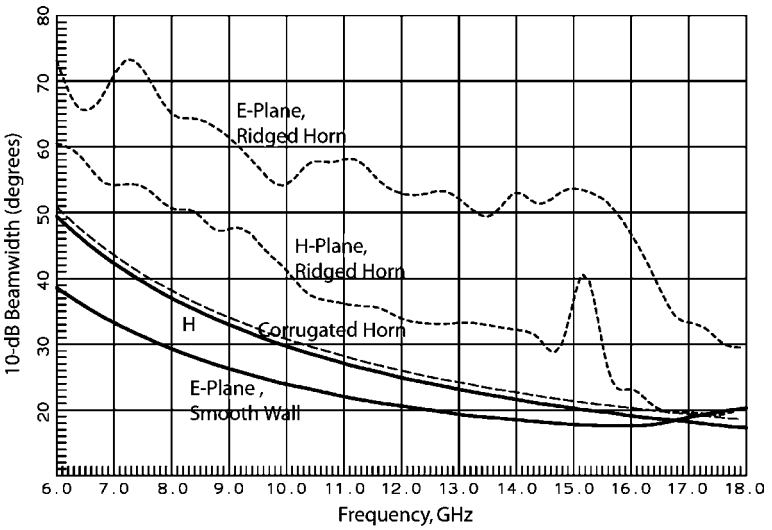


FIGURE 7-15 Measured 10-dB beamwidths of a circular quad-ridged horn compared to the calculated beamwidths of smooth- and corrugated-wall horns.

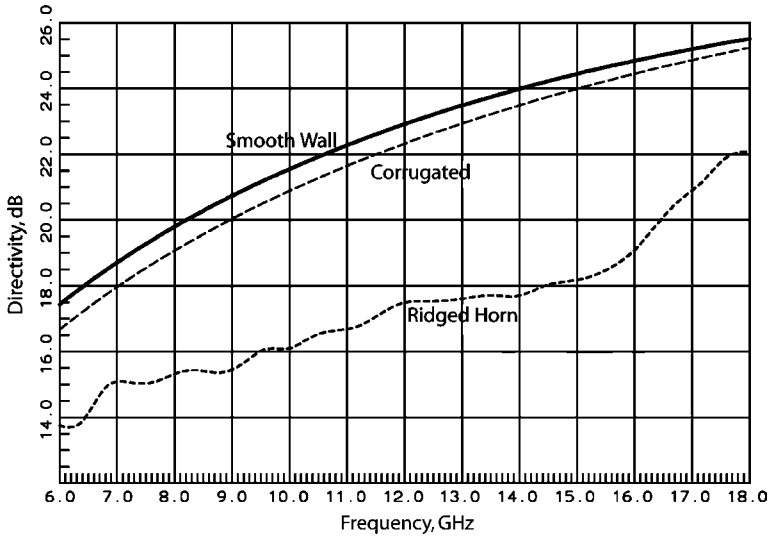


FIGURE 7-16 Measured directivity of a circular quad-ridged horn compared to those of smooth- and corrugated-wall horns.

quad-ridged horn has wider beamwidths in both planes compared to the other horns. This reduces the gain shown in Figure 7-16. Similar to the corrugated horn, the quad-ridged horn operates with multiple modes. We can determine the circular waveguide modes radiated by using physical optics analysis on the measured pattern. We radiate a plane wave into a circular aperture plane equal to the physical horn aperture and placed at the average phase center. Each plane wave, weighted by the pattern level and sin θ, excites Huygens source currents on the patches that cover the aperture by using Eq. (2-33). We normalize the currents to 1 watt and project the currents for each mode of a circular waveguide horn onto the incident wave currents by integrating over the aperture to determine their excitation levels b_m:

$$b_m = \iint_S \mathbf{J}_a \cdot \mathbf{J}_m^* \, dS \tag{7-45}$$

We use the aperture currents $\mathbf{J}_a$ and mode currents $\mathbf{J}_m$ in Eq. (7-45), where we take the complex conjugate of the vector for projection in the same manner as polarization calculations (section 1-11). We operate on the electric currents only because the magnetic currents are proportional to the electric currents for Huygens sources. Figure 7-17 plots the levels of the TE_{11}, TM_{11}, and diagonally oriented TE_{21} modes. TE_{11} and TM_{11} modes are also excited in a corrugated horn, but the level of the TM_{11} mode is approximately -5 dB relative to the TE_{11} mode. Further measurements of the horn show that it has approximately equal power in the TE_{11} and TM_{11} modes, all the way down to 2.7 GHz. Below that frequency the horn aperture will not support the TM_{11} mode and the pattern reverts to the TE_{11} mode only, which narrows the beamwidth. Analysis shows that the diagonally oriented TE_{21} mode peaks at an angle halfway around from the ridges and increases cross-polarization in the diagonal plane. The unmatched beamwidths in the E- and H-planes also increases the Huygens source

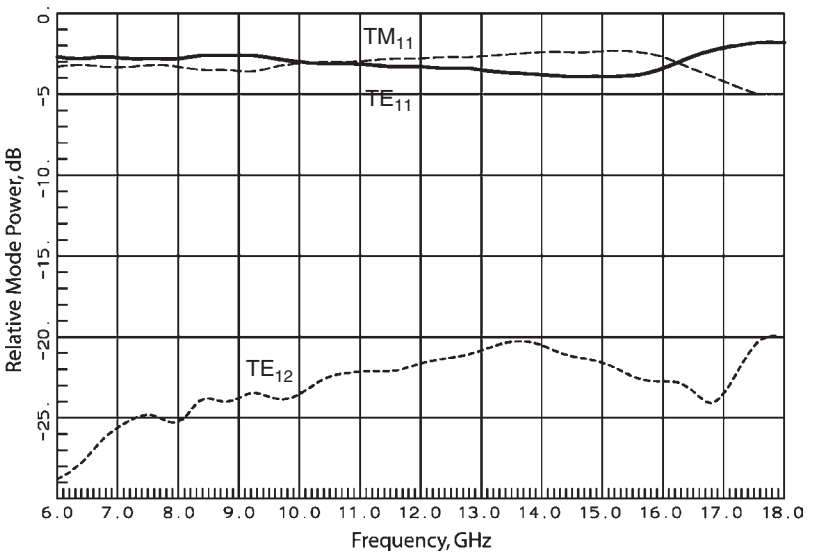


FIGURE 7-17 Modal decomposition into circular waveguide modes of the measured pattern of a circular quad-ridged horn.

cross-polarization (section 1-11.2) in the diagonal plane. Square quad-ridged horns have similar modes. Measurements on a 63.5-cm-square aperture horn with a 140-cm slant length produced nearly equal levels of TE_{10} and TM_{12} modes, similar to the TE_{11} and TM_{11} circular modes in field distribution. The TE_{10} and TM_{12} modes have approximately the same phase. The horn radiated the TE_{12} mode at the higher-frequency end of the band, which caused pattern distortion over a narrow frequency range. Both the TM_{12} and TE_{12} modes are excited by the electric field between the ridges. The interplay of these three modes causes rapid changes in the beam shape as frequency changes. The horn exhibits these changes at the high end of the frequency band when all three modes exist with nearly equal power. Measurements on a dual-ridged horn produced patterns that reduced to the same three dominant modes as the square quad-ridged horn radiated and produced similar results.

We fail to obtain a close match with the measured pattern of the quad-ridged horn by using the aperture currents beyond the 10-dB beamwidth for an aperture small in wavelengths. If we include currents excited along the outside of the horn bell in the physical optics analysis, we better match the measured pattern. This illustrates that the pattern of a horn is determined not only by aperture fields but also by the currents that flow down the bell. Figure 7-18 shows the measured E- and H-plane patterns and the cross-polarization in the diagonal plane. The three-dimensional measured pattern plot in Figure 7-19 at 6 GHz shows the four cross-polarization lobes in the diagonal planes.

The average pattern beamwidth matches a reflector with $f/D = 1$ and has an average illumination loss of 3 dB, with the value ranging from 2.5 to 4 dB. The average taper loss is 1.07 dB and the average spillover loss 1.08 dB. The cross-polarization exhibited in Figure 7-19 contributes an average 0.7 dB of loss. The phase-center location measurements show that the horn has up to 2λ astigmatism, which contributes 0.4 dB of loss when used as a reflector feed.

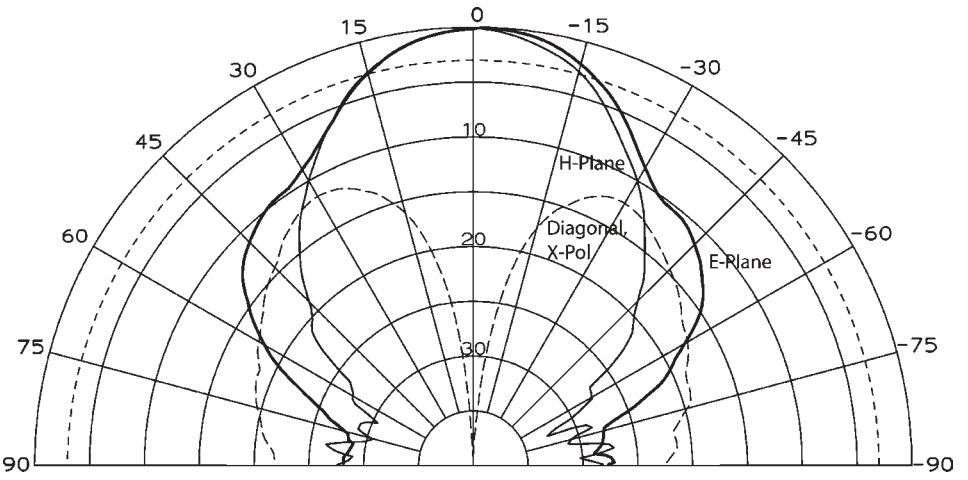


FIGURE 7-18 Measured pattern of a circular quad-ridged horn.

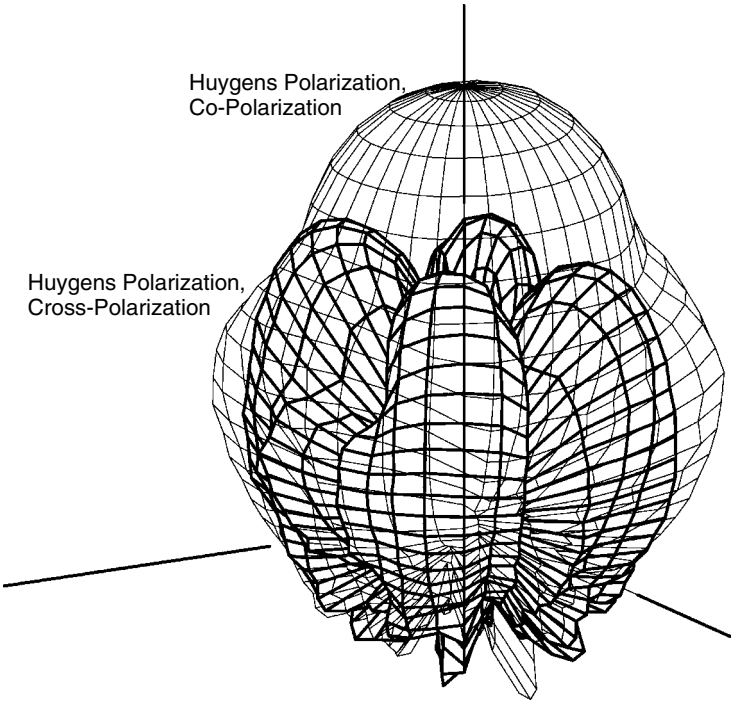


FIGURE 7-19 Spherical radiation pattern of a circular quad-ridged horn showing four-way symmetry of cross-polarization in diagonal planes.

7-7 BOX HORN [32, pp. 377–380]

With a box horn (Figure 7-20), multiple waveguide modes are used to decrease the H-plane amplitude taper loss and axial length of the horn. We add the TE_{30} mode to

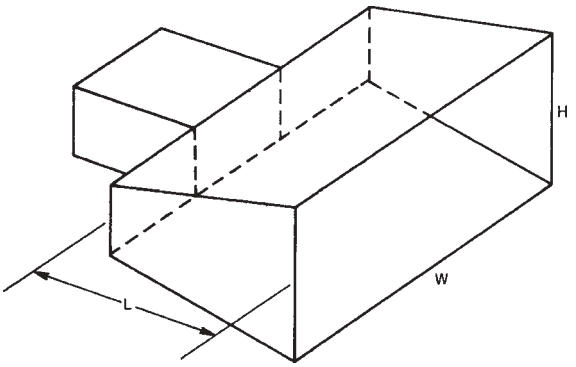


FIGURE 7-20 Box horn.

the TE_{10} mode to reduce the cosine distribution taper of the H-plane. By phasing the modes 180° out of phase in the center of the aperture, the $\cos 3\pi x$ distribution subtracts from the TE_{10}-mode distribution in the center and adds in the region near the edges.

A step in the width of a waveguide generates TE_{N0} modes when fed by the TE_{10} mode. Any modes not cut off by the waveguide will propagate to the aperture. If we maintain symmetry about the axis of the waveguide, only odd-order modes (TE_{30}, TE_{50}, etc.) will be generated. The width W of the waveguide (Figure 7-20) beyond the step determines the possible propagating modes: $\lambda_c = 2W/N$, where N is the mode number. If we limit the modes to the TE_{10} and TE_{30} modes in the aperture, the cutoff wavelength of the TE_{50} mode determines the maximum width: $W_{\max} = 2.5\lambda$. The TE_{30}-mode cutoff wavelength establishes the minimum width: $W_{\min} = 1.5\lambda$. Within this range, short horns with good aperture efficiency can be designed. We can flare the E-plane to increase its aperture (Figure 7-20), but the limited axial length of the horn bounds the possible flare without an excessive phase error loss. The H-plane can also be flared, but flaring it complicates the design for the proper length L. The step generates smaller amplitudes of higher-order modes with each increase in N. Small amounts of higher-order modes (TE_{50}, TE_{70}, etc.) will decrease the efficiency only marginally, since the mode amplitudes are small.

The step generates modes in phase with the input TE_{10} mode, since the higher-order modes must peak in the center and subtract from the TE_{10} fields on the back wall of the larger waveguide section. The aperture distribution is a sum of TE_{10} and TE_{30} modes:

$$E_y(x) = a_1 \cos \frac{\pi x}{W} \exp(-jk_{10}L) + a_3 \cos \frac{3\pi x}{W} \exp(-jk_{30}L) \tag{7-46}$$

where k_{10} and k_{30} are the propagation constants of the two modes. The amplitude distribution in the H-plane will be more nearly uniform if the phase between the modes is 180°. The modes travel from the step with different phase velocities, depending on their cutoff frequencies. We adjust the length L to give a 180° phase difference between the modes:

$$(k_{10} - k_{30})L = \pi$$

where $k_{10} = k\sqrt{1 - (\lambda/2\,W)^2}$ and $k_{30} = k\sqrt{1 - (3\lambda/2\,W)^2}$. We solve for the length:

$$L = \frac{\lambda/2}{\sqrt{1 - (\lambda/2\,W)^2} - \sqrt{1 - (3\lambda/2\,W)^2}} \tag{7-47}$$

TABLE 7-10 Box Horn Characteristics

TE_{30}/TE_{10} (Voltages)	Ratio of Input Waveguide to Aperture	Linear ATL_w (dB)	$(W/\lambda)\sin\theta$	
			3 dB	10 dB
0.00	1.000	0.91	0.594	1.019
0.05	0.940	0.78	0.575	0.981
0.10	0.888	0.67	0.558	0.947
0.15	0.841	0.58	0.544	0.917
0.20	0.798	0.52	0.530	0.890
0.25	0.758	0.48	0.518	0.866
0.30	0.719	0.46	0.507	0.844
0.35	0.682	0.46	0.496	0.824
0.40	0.645	0.47	0.487	0.806
0.45	0.609	0.50	0.479	0.790
0.50	0.573	0.54	0.471	0.775
0.55	0.537	0.60	0.463	0.761
0.60	0.500	0.66	0.456	0.749
0.65	0.462	0.74	0.450	0.737
0.70	0.424	0.82	0.444	0.726

The ratio of the modes generated by the step can be found from mode matching on the input waveguide aperture of width a:

$$\frac{a_N}{a_1} = \frac{\displaystyle\int_{-a/2}^{a/2} \cos(\pi x/a)\cos(N\pi x/W)dx}{\displaystyle\int_{-a/2}^{a/2} \cos(\pi x/a)\cos(\pi x/W)dx} \tag{7-48}$$

where a_N is the TE_{N0} mode amplitude. Table 7-10 lists the step dimensions needed to design to a given ratio of modes. The amplitude taper loss is a minimum at $a_3/a_1 = 0.32$. The possible 3-dB beamwidths with a single mode, TE_{10}, range from 20 to 44°.

Example Design a box horn with an H-plane 10-dB beamwidth of 50°.

We pick $a_3/a_1 = 0.35$. From Table 7-10, $(W/\lambda)\sin\theta = 0.824$. The obliquity factor at 25° adds 0.42 dB of loss. We must design with a wider 10-dB beamwidth. This is within the permissible range for only two modes in the aperture. We calculate the length to phase the modes by 180° by using Eq. (7-47): $L = 1.451\lambda$. The horn is shorter than the aperture width.

7-8 T-BAR-FED SLOT ANTENNA

A T-bar-fed slot antenna (Figure 7-21) looks more like an open-circuited waveguide to coax transition than a slot. Like a slot, its pattern is very broad. The antenna has been designed experimentally [33, pp. 184–190] and those dimensions provide a good starting point. Table 7-11 lists two designs [33] referred to Figure 7-21. The aperture admittance is a combination of the radiation admittance and a capacitive

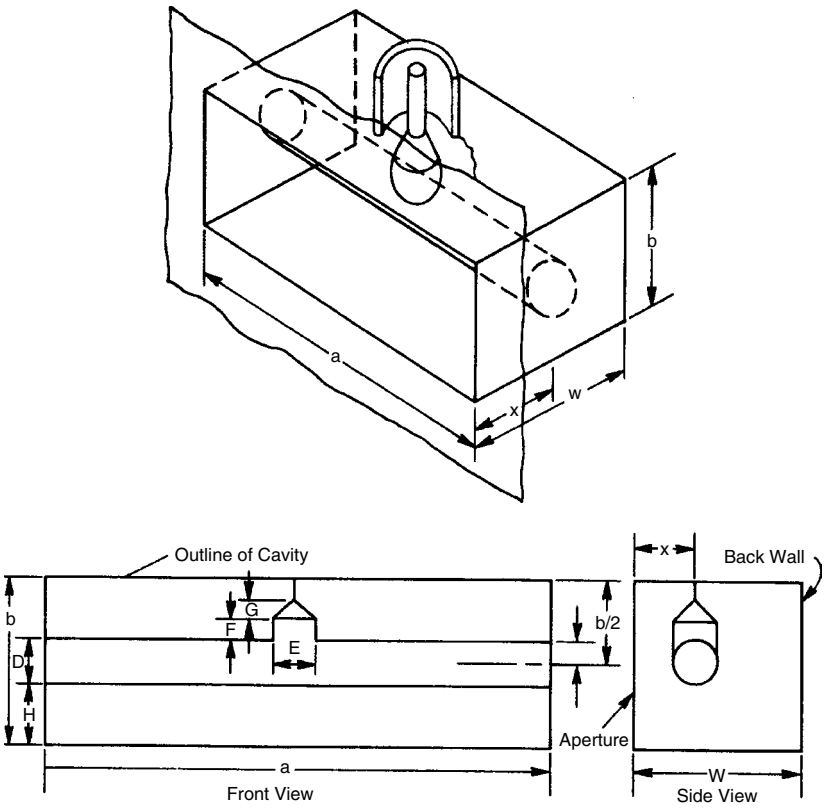


FIGURE 7-21 T-bar-fed cavity slot antenna. (From [34], © 1975 IEEE.)

TABLE 7-11 Dimensions for Two Antenna Designs

Dimensions	Antenna 1	Antenna 2
b/a	0.323	0.226
W/a	0.323	0.295
x/a	0.118	0.113
D/a	0.118	0.090
I/a	0.059	0.045
E/a	0.118	0.090
F/a	0.057	0.045

susceptance. Behind the feed point, the length of short-circuited waveguide adds an inductive susceptance that grows as frequency decreases. The horizontal bar produces a capacitive susceptance at the input to counteract the back-wall susceptance. These susceptances track with frequency changes; each one decreases to maintain the sum near resonance.

Later experimental work [34] revealed further properties of the antenna. Measurements on antenna 1 show that the lower-end 2 : 1 VSWR band edge occurs when $a = 0.57\lambda$ and the upper end when $a = 0.9\lambda$. The bandwidth is about 1.6 : 1.

Antenna 2 was reported [33] as having less bandwidth than antenna 1. When the round rod was replaced with a flat strip, whose width across the guide was the same as the diameter of the rod, almost identical results were obtained. We have a choice. The flat strip is an easier construction, but the round rod gives better mechanical support in all axes to withstand shock and vibration.

The flat strip adds to the freedom of design. The bandwidth potential increases when H is decreased while I is held constant. Newman and Thiele [34] found that when H was decreased, the nominal impedance level was raised. When the input impedance is plotted on the Smith chart, the locus is centered about a higher resistance. By adding a broadband impedance transformer on the input, we can achieve the higher bandwidth potential. Newman and Thiele achieved a nearly 2 : 1 VSWR bandwidth from $a = 0.52\lambda$ to $a = 1.12\lambda$, or 2.3 : 1 bandwidth.

7-9 MULTIMODE CIRCULAR HORN [35]

A step in the diameter of a circular waveguide generates a TM_{11} mode to satisfy the boundary conditions. The fields of the TM_{11} mode can be phased to cancel the fields from the TE_{11} mode at the edges of the aperture in the E-plane. The tapering of the fields in the aperture reduces the E-plane sidelobes while broadening the beamwidth. Equalizing the field distributions in the two planes helps to bring the E- and H-plane phase centers closer together.

The modes generated by the step are more complicated than those for the box horn. Symmetry eliminates generation of the unwanted modes: TM_{01}, TE_{21}, and TE_{01}. The step transition shifts the phase of the TM_{11} mode relative to the TE_{11} mode [36]. Since the waveguide modes have different phase velocities, they can be phased to produce the desired field at the aperture. Although calculated information [36] is helpful, the designs must be completed empirically. The required phasing to achieve field cancellation limits the bandwidth, but for narrowband applications a stepped horn is cheaper than a corrugated horn.

Satoh [37] loads the flare of a conical horn with a conical dielectric step to generate the TM_{11} mode. Symmetry prevents the excitation of unwanted modes. He places the step at a diameter where the TM_{11} mode can propagate. By using two steps, the bandwidth can be increased because the lengths can be adjusted to give perfect mode cancellation at two frequencies. We can replace the dielectric cone by metal steps each of which generates the TM_{11} mode and thereby achieve good results, in theory, at multiple frequencies.

7-10 BICONICAL HORN [4]

A biconical horn consists of two cones with a common vertex. The angle of the generating lines of the cones is measured from a common axis. The cones of the usual antenna have angles that sum to 180°. Spherical modes describe the fields between the cones, but we can use approximations with good results. The lowest-order mode is TEM between the cones and is easily excited by a coax line. The outer conductor connects to one cone, and the second cone feeds out of the center conductor. The electric field of the TEM mode is polarized in the direction of the axis. The first higher-order mode

has a circulating electric field with the magnetic field in the direction of the axis. It can be excited either from a TE_{01}-mode circular waveguide or by an array of slots on a cylinder. The distance between cones must be at least $\lambda/2$ at the point of excitation of the TE_{01} biconical mode.

We approximate the distribution of the zeroth-order mode, TEM, as uniform along the axis. The first-order mode, TE_{01}, distribution is approximately cosine along the axis. We calculate gain by using aperture distribution losses. We describe the horn with a slant radius along the generating line and a height between the ends of the cones. The expansion in spherical modes requires integration over a spherical cap at the aperture if a constant phase surface is used. We obtain good results by using a cylindrical aperture and a quadratic phase distribution. The antenna has circular symmetry about the z-axis that bounds the directivity to $2L/\lambda$. We use linear distribution efficiencies to compute directivity (gain):

$$\text{gain} = 10 \log \frac{2L}{\lambda} - \text{ATL}_x - \text{PEL}_x \tag{7-49}$$

The TEM mode has a uniform distribution, so we use the "uniform" column of Table 4-42 to calculate phase error loss. The uniform distribution has no amplitude taper loss. The cosine distribution of the first-order mode requires an ATL = 0.91 dB and use of the cosine distribution quadratic phase error loss of Table 4-42. Given the height between the ends of the cones, H, and the slant radius R, we determine the quadratic phase distribution constant from

$$S = \frac{H^2}{8\lambda R} \tag{7-50}$$

Example Compute the gain of a biconical horn with a slant radius of 10λ and cone angles of 75° and 105°. $H = 2R \cos 75^\circ = 5.176\lambda$ and $S = 0.33$.

From Table 4-42, we read

$$\text{PEL}_{\text{TEM}} = -1.76\,\text{dB} \quad \text{PEL}_{\text{TE}-01} = -0.79\,\text{dB}$$

Vertical mode, TEM: gain $= 10 \log[2(5.176)] + \text{PEL}_{\text{TEM}}$:

$$10.15\,\text{dB} - 1.76\,\text{dB} = 8.4\,\text{dB}$$

Horizontal mode, TE_{01}: gain $= 10 \log[2(5.176)] + \text{PEL}_{\text{TE}-01} + \text{ATL}_{\text{cosine}}$:

$$10.15\,\text{dB} - 0.79\,\text{dB} - 0.91\,\text{dB} = 8.45\,\text{dB}$$

We can calculate beamwidths by using the results of the rectangular horn, where we measure the angles from $\theta = 90^\circ$ for the complementary-angled biconical horn.

Example Compute the 3-dB beamwidths of the horn above. $S = 0.33$ and $H = 5.176\lambda$.

Use Table 7-2 with the TEM mode and α as the angle from $\theta = 90^\circ$:

$$\frac{H}{\lambda} \sin \alpha = 0.5015 \quad \alpha = 5.56^\circ$$

The obliquity factor is insignificant.

$$\text{HPBW} = 11.1^\circ \qquad \text{TEM mode}$$

Use Table 7-1 with the TE_{01} mode.

$$\frac{H}{\lambda}\sin\alpha = 0.6574 \qquad \text{HPBW} = 14.6^\circ \qquad TE_{01} \text{ mode}$$

The two modes have about the same gain, but the TE_{01} mode has a greater beamwidth. When we refer to Figures 7-3 and 7-4, we see that the TEM-mode horn has about 7-dB sidelobes and the TE_{01}-mode horn has practically no sidelobes. The sidelobes reduce the gain of the TEM mode with its narrower beamwidth.

REFERENCES

1. A. W. Love, ed., *Electromagnetic Horn Antennas*, IEEE Press, New York, 1976.
2. W. L. Barrow and L. J. Chu, Theory of electromagnetic horn, *Proceedings of IRE*, vol. 27, January 1939, pp. 51–64.
3. M. C. Schorr and E. J. Beck, Electromagnetic field of the conical horn, *Journal of Applied Physics*, vol. 21, August 1950, pp. 795–801.
4. S. A. Schelkunoff and H. Friis, *Antenna Theory and Practice*, Wiley, New York, 1952.
5. P. M. Russo et al., A method of computing E-plane patterns of horn antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-13, no. 2, March 1965, pp. 219–224.
6. J. Boersma, Computation of Fresnel integrals, *Mathematics of Computation*, vol. 14, 1960, p. 380.
7. K. S. Kelleher, in H. Jasik, ed., *Antenna Engineering Handbook*, McGraw-Hill, New York, 1961.
8. D. G. Bodnar, Materials and design data, Chapter 46 in R. C. Johnson, ed., *Antenna Engineering Handbook*, 3rd ed., McGraw-Hill, New York, 1993.
9. E. H. Braun, Gain of electromagnetic horns, *Proceedings of IRE*, vol. 41, January 1953, pp. 109–115.
10. E. I. Muehldorf, The phase center of horn antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-18, no. 6, November 1970, pp. 753–760.
11. A. J. Simmons and A. F. Kay, The scalar feed: a high performance feed for large paraboloid reflectors, Design and Construction of Large Steerable Aerials, *IEE Conference Publication* 21, 1966, pp. 213–217.
12. B. M. Thomas, Design of corrugated conical horns, *IEEE Transactions on Antennas and Propagation*, vol. AP-26, no. 2, March 1978, pp. 367–372.
13. T. S. Chu and W. E. Legg, Gain of corrugated conical horn, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 4, July 1982, pp. 698–703.
14. G. L. James, TE_{11} to HE_{11} mode converters for small angle corrugated horns, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 6, November 1982, pp. 1057–1062.
15. P. J. B. Clarricoats and P. K. Saha, Propagation and radiation of corrugated feeds, *Proceedings of IEE*, vol. 118, September 1971, pp. 1167–1176.
16. A. W. Rudge et al., eds., *The Handbook of Antenna Design*, Vol. 1, Peter Peregrinus, London, 1982.

17. G. L. James and B. M. Thomas, TE_{11}-to-HE_{11} corrugated cylindrical waveguide mode converters using ring-loaded slots, *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-30, no. 3, March 1982, pp. 278–285.
18. R. Wohlleben, H. Mattes, and O. Lochner, Simple small primary feed for large opening angles and high aperture efficiency, *Electronics Letters*, vol. 8, September 21, 1972, pp. 474–476.
19. A. D. Olver et al., *Microwave Horns and Feeds*, IEEE Press, New York, 1994.
20. G. L. James, Radiation properties of 90° conical horns, *Electronics Letters*, vol. 13, no. 10, May 12, 1977, pp. 293–294.
21. A. Kumer, Reduce cross-polarization in reflector-type antennas, *Microwaves*, March 1978, pp. 48–51.
22. P.-S. Kildal, *Foundations of Antennas*, Studentlitteratur, Lund, Sweden, 2000.
23. S. Maci et al., Diffraction at artificially soft and hard edges by using incremental theory of diffraction, *IEEE Antennas and Propagation Symposium*, 1994, pp. 1464–1467.
24. B. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, Wiley, New York, 1991.
25. K. Pontoppidan, ed., *Technical Description of Grasp 8*, Ticra, Copenhagen, 2000 (self published and available at www.ticra.com).
26. P. F. Goldsmith, *Quasioptical Systems*, IEEE Press, New York, 1998.
27. L. Diaz and T. A. Milligan, *Antenna Engineering Using Physical Optics*, Artech House, Boston, 1996.
28. J. L. Kerr, Short axial length broadband horns, *IEEE Transactions on Antennas and Propagation*, vol. AP-21, no. 5, September 1973, pp. 710–714.
29. J. R. Whinnery and H. W. Jamieson, Equivalent circuits for discontinuities in transmission lines, *Proceedings of IRE*, vol. 32, no. 2, February 1944, pp. 98–114.
30. S. B. Cohn, Properties of ridge waveguide, *Proceedings of IRE*, vol. 35, no. 8, August 1947, pp. 783–788.
31. M. H. Chen, G. N. Tsandoulas, and F. W. Willwerth, Modal characteristics of quadruple-ridged circular and square waveguide, *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-21, August 1974, pp. 801–804.
32. S. Silver, ed., *Microwave Antenna Theory and Design*, McGraw-Hill, New York, 1949.
33. A. Dome and D. Lazarno, Radio Research Laboratory Staff, *Very High Frequency Techniques*, McGraw-Hill, New York, 1947, pp. 184–190.
34. E. H. Newman and G. A. Thiele, Some important parameters in the design of T-bar fed slot antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-23, no. 1, January 1975, pp. 97–100.
35. P. D. Potter, A new horn antenna with suppressed sidelobes and equal beamwidths, *Microwaves*, vol. 6, June 1963, pp. 71–78.
36. W. J. English, The circular waveguide step-discontinuity mode transducer, *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-21, no. 10, October 1973, pp. 633–636.
37. T. Satoh, Dielectric loaded horn antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-20, no. 2, March 1972, pp. 199–201.

8

REFLECTOR ANTENNAS

The importance of reflector antennas cannot be overstated. Large-aperture antennas can be built only with reflectors or arrays and reflectors are far simpler than arrays. The arrays give us more degrees of freedom than is necessary in many applications. With plenty of room and slow scan rates, a reflector becomes a better design than an array. Of course, there can be many valid reasons for using an array in an application, but a reflector should always be considered. An array needs an elaborate feed network, whereas a reflector uses a simple feed and free space as its feed network.

Most reflector designs require extensive calculations together with full characterization of the feed antenna. Many types of analysis have been developed. As with horn antennas, Love [1] has collected significant papers on reflector antennas. In his classic book, Silver [2] provides the foundation for an analysis based on aperture theory and physical optics (induced currents on the reflector). Aperture theory or physical optics reduced to aperture theory is still used for most designs. Rusch and Potter fully develop aperture and physical optics theories for the design and analysis of both prime focus and dual-reflector (Cassegrain) antennas [3]. Other methods have been developed either to increase the range of valid patterns or to decrease the pattern calculation time so that optimization techniques can be applied. Wood [4] collects ideas for designing by using spherical wave expansions that allow for an overall system optimization using only a few terms. GTD methods [5,6] find increasing applications as an analysis technique suitable for a full pattern analysis except at boresight. Improved methods of calculating the secondary pattern have been developed using aperture fields, such as FFT methods [7] and Jacobi–Bessel series [8]. Many of these techniques and hardware implementations of reflectors are summarized in a handbook [9, Chaps. 2 and 3]. Although all these methods are available, aperture theory and physical optics remain the main techniques of reflector design and analysis.

Modern Antenna Design, Second Edition, By Thomas A. Milligan

8-1 PARABOLOIDAL REFLECTOR GEOMETRY

Figure 8-1 shows the geometry of the parabolic reflector. We form the reflector by rotating the figure about its axis or by moving the figure along an axis out of the paper to form a cylindrical reflector. Because the cylindrical reflector requires a line source, it is less important than the circularly symmetrical reflector fed from a single point source. A paraboloidal reflector transforms a spherical wave radiated by the feed located at its focus into a plane wave. Although the feed wave spreads from the focus, which reduces its amplitude, geometric optics predicts a plane wave reflection that remains constant. The reflected wave does not remain a plane wave but spreads because the fields must be continuous across the reflection boundary of the beam plane wave column because fields can be discontinuous only across physical boundaries. Nevertheless, we will use the aperture theory on the projected diameter to predict its performance. Since the reflected rays are parallel, we can place the aperture plane anywhere along the axis, but somewhat close in front of the reflector. The equations for the reflector surface are

$$\underset{\substack{\text{rectangular}\\\text{coordinates}}}{r^2 = 4f(f+z)} \qquad \underset{\substack{\text{polar}\\\text{coordinates}}}{\rho = \frac{f}{\cos^2(\psi/2)}} \tag{8-1}$$

where f is the focal length, D the diameter, ρ the distance from the focus to the reflector, and ψ the feed angle from the negative z-axis. The reflector depth from the rim to the center is $z_0 = D^2/16f$.

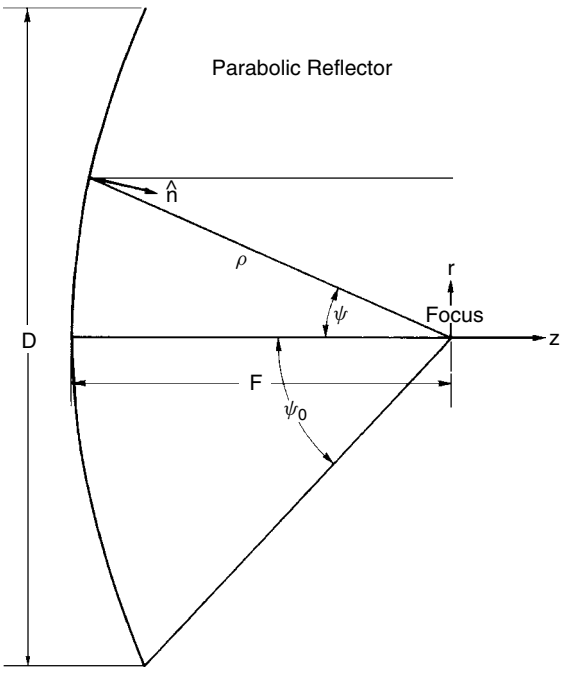
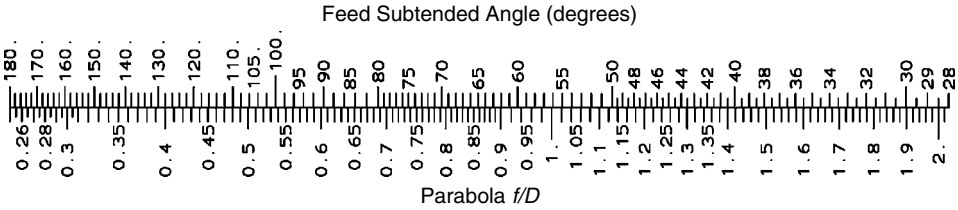


FIGURE 8-1 Geometry of a parabolic reflector.



SCALE 8-1 Parabola f/D compared to a feed total subtended angle.

We eliminate the dimensions of the reflector by using the ratio f/D. The half subtended angle of the reflector, ψ_0, relates to f/D by

$$\psi_0 = 2\tan^{-1}\frac{1}{4f/D} \tag{8-2}$$

Scale 8-1 computes the total feed subtended angle from reflector f/D. When we place the aperture plane at the focus, the ray path distance becomes

$$\rho + \rho\cos\psi = 2\rho\cos^2\frac{\psi}{2} = 2f$$

all ray path lengths are equal, and the aperture plane is a constant-phase surface (eikonal).

The normal unit vector at a point on the reflector (r, z) is found from the feed angle:

$$\hat{\mathbf{n}} = -\sin\frac{\psi}{2}\hat{\mathbf{r}} + \cos\frac{\psi}{2}\hat{\mathbf{z}}$$

At this point we need the radius of curvatures in the principal planes to apply Eq. (2-77) reflection from a curved surface: R_1 in the $r-z$ plane and R_2 in the $\phi-z$ plane:

$$R_1 = \frac{2f}{\cos^3(\psi/2)} \quad \text{and} \quad R_2 = \frac{2f}{\cos(\psi/2)}$$

The spherical wave spreads from the feed as $1/\rho$. At the surface of the reflector the wave curvature changes to a plane wave and propagates to the aperture plane at a constant amplitude. The spherical wave spreading multiplies the feed distribution by [Eq. (8-1)] $\cos^2(\psi/2)$ in the aperture. Then

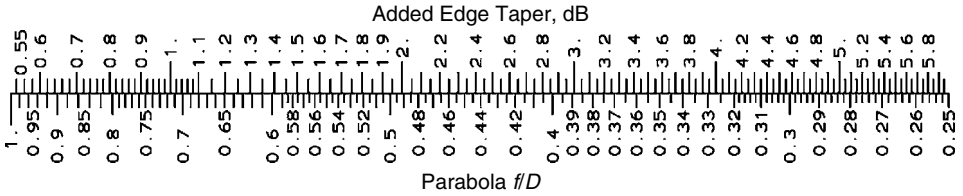
$$\text{added edge taper} = \cos^2\frac{\psi_0}{2} \qquad \text{voltage} \tag{8-3}$$

Deeper reflectors (smaller f/D) have greater edge tapers than shallow reflectors (larger f/D). Scale 8-2 provides a quick calculation of the added edge taper due to spherical wave spreading.

Example Calculate the edge taper of a paraboloidal reflector for $f/D = 0.5$ and an isotropic feed.

From Eq. (8-2), $\psi_0 = 2\tan^{-1}\frac{1}{2} = 53.13°$. The edge taper is [Eq. (8-3)]

$$\text{edge taper} = 20\log\cos^2\frac{53.13°}{2} = -1.94\,\text{dB}$$



SCALE 8-2 Added edge taper due to a spherical wave from feed.

If the feed has its 10-dB pattern point directed toward the reflector edge, the aperture edge taper is 11.9 dB.

8-2 PARABOLOIDAL REFLECTOR APERTURE DISTRIBUTION LOSSES

We manipulate Eq. (4-2) for ATL to eliminate the dimensions and relate the integrals to the feed pattern:

$$\text{ATL} = \frac{\left[\int_0^{2\pi}\int_b^a \left|E_a(r',\phi')\right| r'\,dr'\,d\phi'\right]^2}{\pi a^2 \int_0^{2\pi}\int_b^a \left|E_a(r',\phi')\right|^2 r'\,dr'\,d\phi'} \tag{8-4}$$

where a is the aperture radius, b the central blockage radius, and $E_a(r', \phi')$ the aperture field. We make the following substitutions into Eq. (8-4):

$$\begin{aligned} r' &= \rho \sin\psi = 2\sin\frac{\psi}{2}\cos\frac{\psi}{2}\frac{f}{\cos^2(\psi/2)} = 2f\tan\frac{\psi}{2} \\ dr' &= f\sec^2\frac{\psi}{2} = \rho\, d\psi \end{aligned} \tag{8-5}$$

The aperture field is related to the feed pattern by

$$E_a(r',\phi') = \frac{E(\psi',\phi')}{\rho}$$

These substitutions eliminate dimensions in Eq. (8-4):

$$\text{ATL} = \frac{\left[\int_0^{2\pi}\int_{\psi_b}^{\psi_0} |E(\psi,\phi)|\tan(\psi/2)\,d\psi\,d\phi\right]^2}{\pi[\tan^2(\psi_0/2) - \tan^2(\psi_b/2)]\int_0^{2\pi}\int_{\psi_b}^{\psi_0} |E(\psi,\phi)|^2 \sin\psi\,d\psi\,d\phi} \tag{8-6}$$

where $\psi_b = 2\tan^{-1}[b/(2f)]$. When we substitute the relations in Eq. (8-5) into Eq. (4-9) to eliminate dimensions in the integrals, we obtain an expression with only the feed

pattern:

$$\text{PEL} = \frac{\left| \int_0^{2\pi} \int_{\psi_b}^{\psi_0} E(\psi, \phi) \tan(\psi/2)\, d\psi\, d\phi \right|^2}{\left[\int_0^{2\pi} \int_{\psi_b}^{\psi_0} |E(\psi, \phi)| \tan(\psi/2)\, d\psi\, d\phi \right]^2} \tag{8-7}$$

PEL is the efficiency at the boresight. We modify Eq. (8-7) when we scan the beam to give off-boresight values as in Eq. (4-3).

The amplitude taper efficiency (ATL) of Eq. (8-6) and the phase error efficiency (PEL) of Eq. (8-7) do not account for the total directivity loss of the aperture. The reflector does not intercept all the power radiated by the source and some of it spills over the edge. Spillover adds little to the pattern except as sidelobes, since usual feeds have small backlobes. We consider this spilled-over power as a loss (SPL):

$$\text{SPL} = \frac{\int_0^{2\pi} \int_{\psi_b}^{\psi_0} |E(\psi, \phi)|^2 \sin\psi\, d\psi\, d\phi}{\int_0^{2\pi} \int_0^{\pi} |E(\psi, \phi)|^2 \sin\psi\, d\psi\, d\phi} \tag{8-8}$$

This expression for spillover includes the scattered portion of the central blockage efficiency, but not the loss of potential aperture. We include the remainder in the directivity calculation.

We have ignored the cross-polarized power radiated by the source. We define cross-polarization efficiency (XOL) as

$$\text{XOL} = \frac{\int_0^{2\pi} \int_0^{\pi} |E_C(\psi, \phi)|^2 \sin\psi\, d\psi\, d\phi}{\int_0^{2\pi} \int_0^{\pi} (|E_C(\psi, \phi)|^2 + |E_X(\psi, \phi)|^2) \sin\psi\, d\psi\, d\phi} \tag{8-9}$$

where E_c is the co-polarized field and E_x is the cross-polarized field. These polarizations correspond to Ludwig's [10] third definition of cross-polarization. A Huygens source produces straight reflector surface currents when projected to the aperture plane. Including the cross-polarization efficiency gives us the true average radiation intensity as in Eq. (1-17).

If we express the efficiencies as ratios, the directivity will be found from

$$\text{directivity} = \left(\frac{\pi}{\lambda}\right)^2 (D_r^2 - D_b^2)\text{SPL} \cdot \text{ATL} \cdot \text{PEL} \cdot \text{XOL} \quad \text{(ratio)} \tag{8-10}$$

where D_r is the reflector diameter and D_b is the diameter of the central blockage. Equation (8-10) includes the nonscattered blockage loss of potential aperture. Equation (8-10) can be expressed in terms of decibel ratios:

$$\begin{aligned} \text{directivity} = 10\log\left[\left(\frac{\pi}{\lambda}\right)^2 (D_r^2 - D_b^2)\right] + \text{SPL(dB)} + \text{ATL(dB)} \\ + \text{PEL(dB)} + \text{XOL(dB)} \end{aligned} \tag{8-11}$$

Of course, all the decibel ratios of the efficiencies will be negative and subtract from the directivity calculated from the area.

When measuring an actual feed, we can ignore the cross-polarized power. We measure the efficiency as the difference between directivity and gain. Actual directivity includes the co-polarizations and cross-polarizations in the average radiation intensity. If we ignore the cross-polarization, the measured efficiency decreases by the cross-polarization loss because the measured and true directivity differ by that loss. We must measure the cross-polarization pattern distribution of the feed if we want to calculate the cross-polarized secondary (reflector) pattern. When the cross-polarization pattern is not required, we save time without loss of accuracy by measuring only the co-polarized feed pattern.

Equations (8-8) and (8-9) are by no means unique. We could include the cross-polarized power in the spillover calculation [Eq. (8-8)] and limit the integration limits in Eq. (8-9) to the reflector. A set of efficiency relations is correct when the equations account for all the power radiated by the feed. When we use calculated feed patterns, we must determine cross-polarization efficiency, since we can only estimate the efficiency due to material losses. The cross-polarization efficiency cannot be included as it is in measurements, and the division of cross-polarized power between Eqs. (8-8) and (8-9) is arbitrary.

8-3 APPROXIMATE SPILLOVER AND AMPLITUDE TAPER TRADE-OFFS

We use the approximate pattern $\cos^{2N}(\psi/2)$ for a feed pattern to establish trends. Of course, if the actual feed pattern distribution is available, we should use Eqs. (8-6) to (8-9). We obtain closed-form expressions when we substitute this pattern into Eqs. (8-6) and (8-8). Ignoring any central blockage, we get

$$\text{spillover efficiency} = 1 - u^{2(N+1)} \tag{8-12}$$

$$\text{amplitude taper efficiency} = \frac{4(N+1)(1-u^N)^2}{N^2[1-u^{2(N+1)}]} \cot^2 \frac{\psi_0}{2} \tag{8-13}$$

where $u = \cos(\psi_0/2)$. We combine Eqs. (8-12) and (8-13) and plot their combination to find the beamwidth for minimum loss. In Figure 8-2 the loss versus the 10-dB beamwidth for various f/D values is plotted. At narrow beamwidths little feed power spills over the reflector edge, but the reflector is underilluminated. Increasing the beamwidth improves the illumination but increases the spillover. The efficiency peaks when the feed 10-dB beamwidth is approximately the subtended angle of the reflector. Figure 8-2 shows a broad peak for any given f/D. Small changes in the beamwidth near the peak have no practical effect on the reflector's gain. Scale 8-3 relates the average illumination loss reduction given the feed pattern level in the direction of the reflector rim for typical antennas.

Example Estimate the amplitude taper loss for a reflector with $f/D = 0.5$ whose feed has a 10-dB edge taper.

Compare the loss with that of the circular Gaussian and the Hansen single-parameter distributions: $\psi_0 = 2\tan^{-1}\frac{1}{2} = 53.13^\circ$. The 10-dB beamwidth of the feed is then 106.26°. We modify Eq. (1-20) to compute the exponent N of the $\cos^{2N}(\psi/2)$ feed

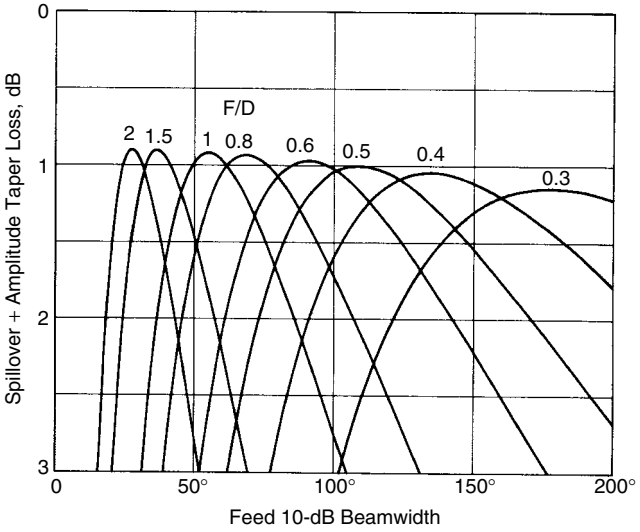
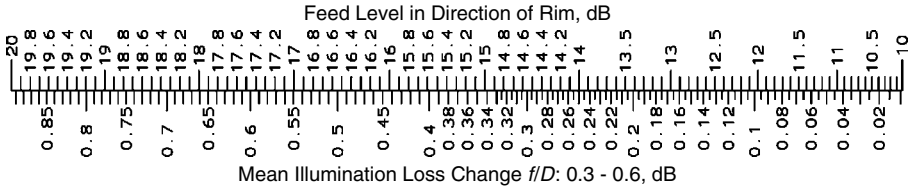


FIGURE 8-2 Sum of spillover and amplitude taper losses versus feed 10-dB beamwidth.



SCALE 8-3 Mean illumination loss change of a reflector given the feed pattern level in the rim direction.

pattern approximation,

$$N = \frac{\log 0.1}{2 \log \cos(106.26^\circ/4)} = 10.32$$

From Eq. (8-13), $u = \cos(53.13^\circ/2) = 0.894$:

$$\text{ATL(dB)} = \frac{4(1 - 0.894^{10.32})^2(11.32)}{10.32^2(1 - 0.894^{22.64})} \cot^2(53.13^\circ/2) = 0.864$$
$$= 10 \log 0.864 = -0.63 \text{ dB}$$

The extra distance from the feed to the reflector edge compared with the center distance adds 1.94 dB and increases the aperture amplitude taper to 11.94 dB. We interpolate Table 4-29 for the circular Gaussian distribution and Table 4-30 for the Hansen single-parameter distribution to find the following data:

Gaussian	Hansen
ATL(dB) = −0.62 dB	ATL(dB) = −0.57 dB
Sidelobe level = 26.3 dB	Sidelobe level = 24.7 dB
Beamwidth factor = 1.142	Beamwidth factor = 1.136

We multiply Eq. (4-83) by the beamwidth factor to estimate the reflector beamwidth:

$$\text{HPBW} = 67.3^\circ \frac{\lambda}{D} \quad \text{and} \quad \text{HPBW} = 67^\circ \frac{\lambda}{D}$$

These compare well with the approximation, $\text{HPBW} = 70^\circ\lambda/D$ for a parabolic reflector. An integration of the aperture distribution for the far-field pattern gives the following results:

$$\text{HPBW} = 67.46^\circ \frac{\lambda}{D} \qquad \text{sidelobe level} = 27\,\text{dB}$$

8-4 PHASE ERROR LOSSES AND AXIAL DEFOCUSING

All rays starting at the reflector focus travel the same distance through reflection to the aperture plane. The aperture plane is any convenient plane in front of the dish whose normal is the axis of the reflector. If we could build a feed with a unique phase center and place it at the focus of a perfect paraboloidal reflector, we would eliminate phase error loss in the aperture plane because it would have a constant phase. The feed, the positioning of the feed, and the reflector surface all contribute to the phase error loss.

We discussed techniques for obtaining unique phase centers in the various planes for horns. Unlike smooth-wall horns, corrugated horns can have equal phase centers in all planes through the axis, but even their position will wander with changes in frequency. We measure the feed pattern distribution (amplitude and phase) to predict the contribution of the feed to the overall efficiency. From those measurements we define the practical phase center as the point on the feed leading to the minimum phase error loss when placed at the focus. The random and systematic phase error contributions can be measured directly on the feed and calculated numerically using Eq. (8-7).

The feed phase center cannot always be placed at the focus. The phase-center location wanders with changes in frequency, and in any wideband application we expect axial defocusing. For example, the location of the phase center of a log-periodic antenna moves toward the apex as frequency increases. Figure 8-3 is a plot of the phase error loss due to axial defocusing. Each feed has its 10-dB beamwidth equal to the reflector subtended angle. Axial defocusing affects deep dishes (lower f/D) more than shallow dishes. We can estimate the axial defocusing phase error loss by approximating the distribution with a quadratic aperture phase distribution. Given z as the axial defocusing, the maximum phase deviation in cycles is

$$S = \frac{z}{\lambda}\left[1 - \cos\left(2\tan^{-1}\frac{1}{4f/D}\right)\right] \tag{8-14}$$

We combine this with the quadratic phase error loss of the circular Gaussian distribution to estimate the loss. With $z = \lambda$ we obtain a scaling factor for S (Scale 8-4) given z from Eq. (8-14). The scaling factor decreases with increasing f/D.

Example Estimate the phase error loss for $z = 2\lambda$ when $f/D = 0.6$ and the feed 10-dB beamwidth equals the reflector subtended angle.

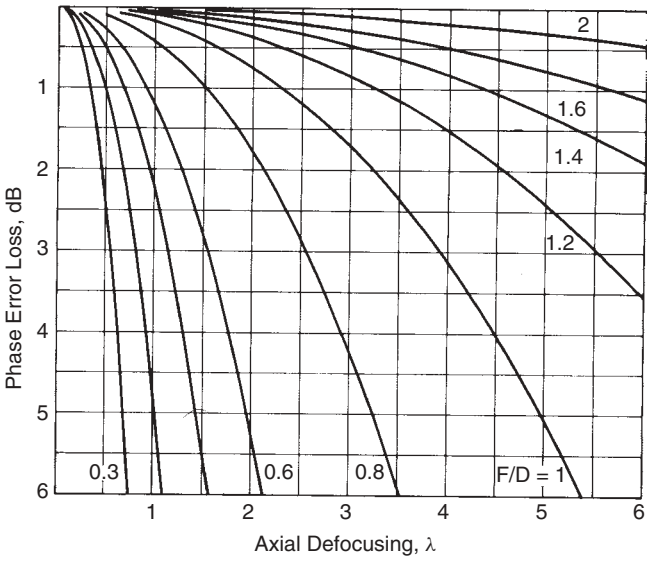
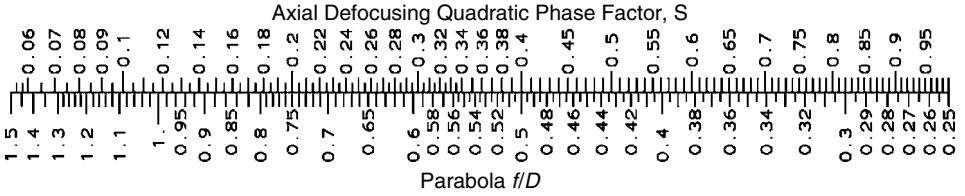


FIGURE 8-3 Paraboloidal reflector phase error loss due to axial defocusing of the feed.



SCALE 8-4 Quadratic phase factor S for axial defocusing of a paraboloidal reflector.

From Scale 8-4, $S = 0.30(2) = 0.60$. We use Eq. (8-3) to compute the edge taper:

$$\psi_0 = 2\tan^{-1}\frac{1}{2.4} = 45.2^\circ$$

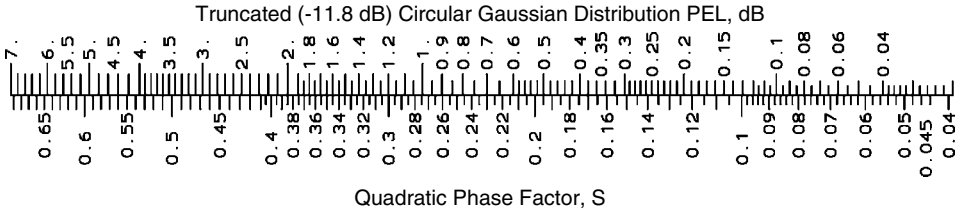
$$\text{edge taper} = 20\log\cos^2\frac{\psi_0}{2} = -1.4\,\text{dB}$$

An equivalent truncated Gaussian aperture distribution tapers to

$$10\,\text{dB} + 1.4\,\text{dB} = 11.4\,\text{dB} \qquad \rho = \frac{11.4}{8.69} = 1.31$$

We use Eq. (4-118) to calculate phase error efficiency of the truncated Gaussian distribution: PEL = 0.305 or PEL (dB) = −5.2 dB. This matches the value from Figure 8-3 found by integration of the actual distribution. The optimum feed beamwidth produces an average aperture edge taper of 11.8 dB. Scale 8-5 evaluates Eq. (4-118) for this taper.

We detect axial defocusing by looking at the patterns of the reflector. Axial defocusing fills-in nulls between sidelobes. We adjust the feed location to maximize the null



SCALE 8-5 Truncated circular Gaussian distribution (−11.8 dB taper) phase error loss given S.

depth, but antenna range errors and receiver sensitivity limit our ability to eliminate this defocusing.

8-5 ASTIGMATISM [11]

Both the feed and the reflector can have astigmatism: unequal phase centers in different planes. We measure the feed by itself to discover its astigmatism. When the feed is mounted in the reflector, we detect astigmatism by the depth of nulls in the various pattern planes. A series of measurements can separate the feed and reflector astigmatism, but the feed must be able to move along the reflector axis and to rotate by 90° during the measurements. Move the feed along the axis to find the locations that give maximum nulls. The extrema of the reflector focuses may not occur in the E- and H-planes and will require a search in the other planes. At this point we cannot separate the feed astigmatism from the reflector astigmatism. We rotate the feed and repeat the measurements. The feed phase center locations shift, and the reflector focuses remain fixed. Simple manipulation of the data from the two measurements separates the two sources of astigmatism. The reflector can be shimmed to remove its astigmatism, or the feed phase centers can be matched to the reflector focuses. Figure 8-4 shows the magnitude

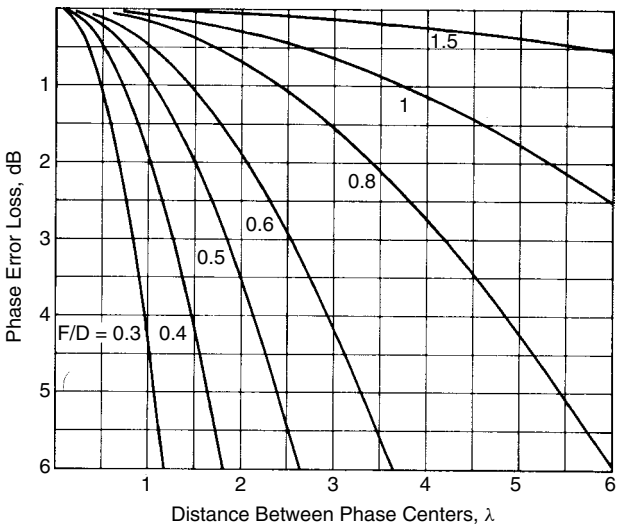


FIGURE 8-4 Paraboloidal reflector phase error loss due to feed astigmatism.

of phase error losses due to feed astigmatism. Astigmatism loss is not as severe as axial defocusing because in two planes the feed phase center is at the reflector focus. As is true of axial defocusing loss, deep dishes are affected more than shallow reflectors.

8-6 FEED SCANNING

Moving the phase center of the feed off axis laterally scans the reflector beam to a limited extent without severe pattern problems. Figure 8-5 shows the k-space pattern effects of feed scanning. The sidelobes show the effects of coma (cubic phase errors) where the sidelobes on the boresight side grow and the sidelobes on the other side decrease. We call these *coma lobes*, although no new lobes are generated. In fact, we see one lobe disappearing as a vestigial lobe with increased scan (Figure 8-5). Suppose that the feed is offset from the axis by a distance d. We measure the offset angle ψ_S from the axis to a line from the feed to the reflector vertex: $d = f \tan \psi_S$. We ignore the slight amplitude distribution change due to the small lateral offset. Referred to the focus, the movement produces a phase factor in the feed pattern: $-kd \sin \psi \cos \phi_c$ when the feed is moved along the negative x-axis.

Equation (8-7) predicts only the boresight phase error loss. Like Eq. (4-3), we must calculate the phase error efficiency at any angle to determine the loss at the pattern peak:

$$\mathrm{PEL}(\theta, \phi) = \frac{\left| \int_0^{2\pi} \int_0^{\psi_0} E(\psi, \phi_c) \tan(\psi/2) e^{jk2f \tan(\psi/2) \sin\theta \cos(\phi - \phi_c)} \, d\psi \, d\phi_c \right|^2}{\left[\int \int |E(\psi, \phi_c)| \tan(\psi/2) \, d\psi \, d\phi_c \right]^2} \tag{8-15}$$

When we include the offset along $\phi = 0$, the phase factor becomes

$$\exp\left[jkf \cos\phi_c \left(2 \tan \frac{\psi}{2} \sin\theta - \tan\psi_S \sin\psi \right) \right]$$

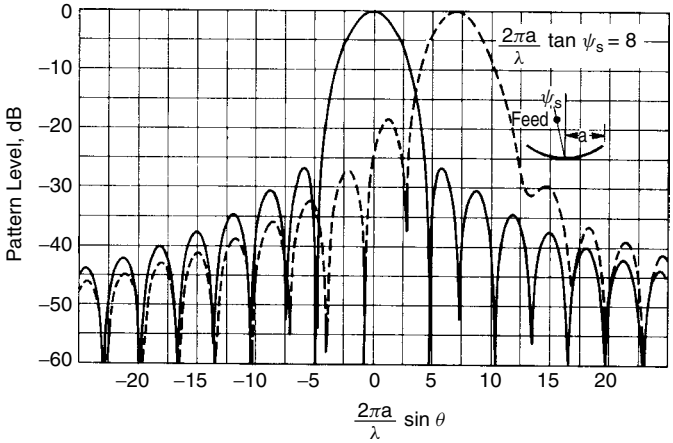


FIGURE 8-5 Feed-scanned paraboloidal reflector $f/D = 0.5$ and feed beamwidth $= 60^\circ$.

For large reflectors we make the approximations $\psi_S \approx \tan\psi_S$ and $\theta \approx \sin\theta$. The pattern scale and the offset phase factor become $ka\theta$ and $ka\psi_S$.

A flat plate would reflect the ray at an equal angle on the other side of the axis for an offset feed, but a curved reflector modifies that result slightly. The offset factor in Figure 8-5 is 8, and the beam peak is at 7. We call the ratio of the beam maximum to offset angle the beam deviation factor (BDF) [12]:

$$\text{BDF} = \frac{\theta_m}{\psi_S} = \frac{7}{8} \qquad \theta_m = \text{BDF} \cdot \psi_S$$

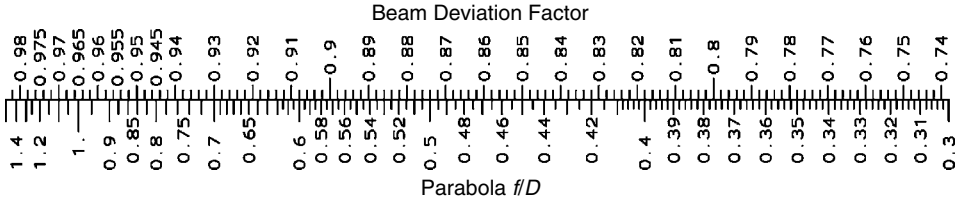
The BDF varies from less than 1 for a concave reflector to greater than 1 for a convex reflector. BDF equals 1 for a flat reflector. Table 8-1 lists the BDF values for various f/D and Scale 8-6 gives the relationship. The BDF approaches 1 as f/D approaches infinity (flat plate). The approximate expression for BDF is

$$\text{BDF} = \frac{(4f/D)^2 + 0.36}{(4f/D)^2 + 1} \tag{8-16}$$

Feed scanning increases the phase error loss. When normalized to beamwidths of scan, a single loss curve can be drawn for each f/D (Figure 8-6). Scanning also raises the sidelobes. Table 8-2 gives the approximate level of the peak coma lobe for a given scan loss. It is almost independent of f/D.

TABLE 8-1 Feed-Scanned Paraboloidal Reflector Beam Deviation Factor

f/D	BDF	f/D	BDF
0.30	0.724	0.80	0.945
0.35	0.778	0.85	0.951
0.40	0.818	0.90	0.957
0.45	0.850	1.00	0.965
0.50	0.874	1.10	0.970
0.55	0.893	1.20	0.975
0.60	0.908	1.40	0.981
0.65	0.921	1.60	0.986
0.70	0.930	1.80	0.989
0.75	0.938	2.00	0.991



SCALE 8-6 Feed-scanned reflector beam deviation factor given f/D.

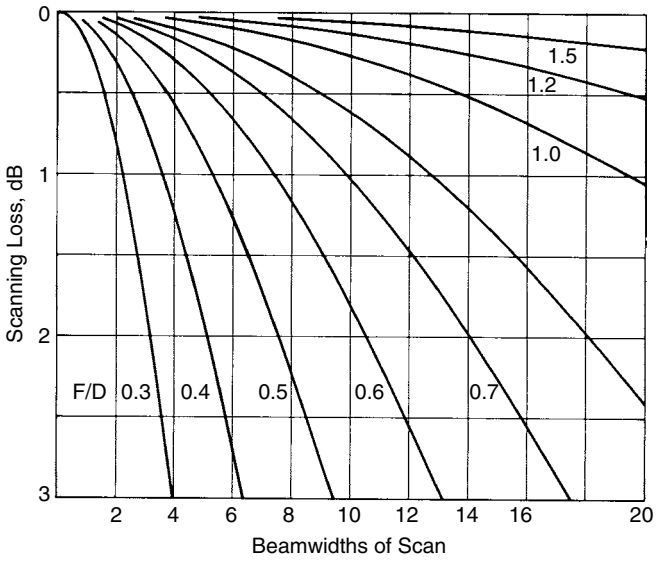


FIGURE 8-6 Feed-scanning loss of a paraboloidal reflector.

TABLE 8-2 Sidelobe Level of a Feed-Scanned Paraboloidal Reflector

Scanning Loss (dB)	Sidelobe Level (dB)	Scanning Loss (dB)	Sidelobe Level (dB)
0.50	14.1	1.75	10.1
0.75	12.9	2.0	9.7
1.00	11.9	2.5	9.0
1.25	11.2	3.0	8.5
1.50	10.6		

Example A reflector with a 50λ diameter is feed-scanned to 6°. Compute the offset distance and scanning loss when $f/D = 0.6$.

Use the approximation $\text{HPBW} = 70^\circ\lambda/D = 1.4^\circ$. The reflector is scanned $6/1.4 = 4.3$ beamwidths:

$$\text{scanning loss (Figure 8-6)} = 0.4\,\text{dB}$$

$$\text{sidelobe level (Table 8-2)} = 14.6\,\text{dB}$$

The angle between the axis and the feed point to vertex must be greater than the scan angle, since the reflector is concave:

$$\psi_S = \frac{\theta_S}{\text{BDF}} = \frac{6^\circ}{0.908} = 6.61^\circ \quad \text{(Table 8-1)}$$

The offset distance is $f \tan 6.61^\circ = 0.6(50\lambda) \tan 6.61^\circ = 3.48\lambda$.

The scalar analysis of this section gives only approximate results. Large feed scanning produces higher-order aberrations other than coma [13–15]. The optimum gain point moves off the focal plane but fails to follow the curve predicted from optics for reflectors extremely large in wavelengths [14]. The reflector f/D and illumination taper determine the maximum gain contour for feed-scanning a reflector. Vector analysis improves the match between calculated and measured results [15].

8-7 RANDOM PHASE ERRORS

Reflector anomalies reduce the gain predicted from the feed analysis. We must specify reasonable manufacturing tolerances for the frequency of operation. It would appear that gain can be increased without bound by increasing the reflector diameter, but the tolerance problems of large reflectors limit the maximum gain. We consider only surface anomalies so small that on average the reflector retains its basic shape. The surface imperfections change the optical path length from the feed to the reflector aperture plane by $\delta(r, \phi)$, which gives us

$$\text{PEL} = \frac{\left| \int_0^{2\pi} \int_0^a E(r, \phi) e^{j\delta(r,\phi)} r \, dr \, d\phi \right|^2}{\left[\int_0^{2\pi} \int_0^a |E(r, \phi)| r \, dr \, d\phi \right]^2} \tag{8-17}$$

Cheng [16] bounds the phase error loss by using a limit on the integrals. Given a peak phase error of m (radians), the change in gain is bounded:

$$\frac{G}{G_0} \geq \left(1 - \frac{m^2}{2}\right)^2 \tag{8-18}$$

This gain loss estimate is too conservative, but it is useful as an upper bound.

Ruze [17] improved the random surface error loss estimate by using a Gaussian distributed error correlated over regions. Dents or segments making up the reflector are correlated with the errors over a nearby region. The error at a point depends on the location of nearby points in the correlation region. The phase error efficiency becomes an infinite series:

$$\text{PEL} = \exp(-\overline{\delta}^2) + \frac{1}{\eta} \left(\frac{2C}{D}\right)^2 \exp(-\overline{\delta}^2) \sum_{n=1}^{\infty} \frac{(\overline{\delta}^2)^n}{n \cdot n!} \tag{8-19}$$

where C is the correlation distance, D the diameter, and η the aperture efficiency (ATL). $\overline{\delta}^2$ is the mean-square phase deviation, given by

$$\overline{\delta}^2 = \frac{\int_0^{2\pi} \int_0^a |E(r, \phi)| \delta^2(r, \phi) r \, dr \, d\phi}{\int_0^{2\pi} \int_0^a |E(r, \phi)| r \, dr \, d\phi} \tag{8-20}$$

If we include the correlation distance, PEL decreases. The infinite series [Eq. (8-19)] converges rapidly. When the correlation distance is small compared with the diameter, the phase error efficiency becomes

$$\text{PEL} = \exp\left(\frac{-4\pi\varepsilon_0}{\lambda}\right)^2 = \exp(-\bar{\delta}^2) \tag{8-21}$$

where ε_0 is the effective reflector tolerance. We use 4π instead of 2π because the wave travels to and from the reflector and the phase distance is twice the reflector tolerance. From Eq. (8-20) we derive the effective RMS tolerance:

$$\varepsilon_o^2 = \frac{\displaystyle\int_0^{2\pi}\int_0^a |E(r,\phi)|\varepsilon^2(r,\phi)r\,dr\,d\phi}{\displaystyle\int_0^{2\pi}\int_0^a |E(r,\phi)|r\,dr\,d\phi} \tag{8-22}$$

Ruze gives the distance ε in terms of the z-axis deviation Δz and the surface normal

$$\varepsilon = \frac{\Delta z}{1+(r/2f)^2} \qquad \varepsilon = \frac{\Delta n}{\sqrt{1+(r/2f)^2}} \tag{8-23}$$

We evaluate the constants in Eq. (8-21) and convert to a decibel ratio:

$$\text{PEL(dB)} = -685.8\left(\frac{\varepsilon_0}{\lambda}\right)^2 \tag{8-24}$$

Example Compute the required reflector tolerance at 30 GHz to limit the RMS surface tolerance phase error loss to 1 dB.

Using Eq. (8-24), we get

$$\frac{\varepsilon_0}{\lambda} = \sqrt{\frac{1}{685.8}} = 0.038$$

At 30 GHz, $\lambda = 1$ cm and $\varepsilon_0 = 0.38$ mm. We can also use Eq. (8-18), which gives the upper bound on surface error loss:

$$m = \frac{4\pi\varepsilon_0}{\lambda} = \sqrt{2\left(1-\sqrt{\frac{G}{G_0}}\right)} = 0.466 \text{ at } 1\text{ dB}$$

$\varepsilon_0 = 0.037\lambda$ or $\varepsilon_0 = 0.37$ mm at 30 GHz. Both methods give about the same answer in this case.

Zarghamee [18] extended tolerance theory to include the effects of the surface error distribution. Some antennas have better support and construction in some areas and are more accurate in those areas. This improves the reflector performance. Zarghamee defined a second variation of surface deviations by

$$\eta_0^4 = \frac{\displaystyle\int_0^{2\pi}\int_0^a |E(r,\phi)|[\varepsilon^2(r,\phi)-\varepsilon_0^2]r\,dr\,d\phi}{\displaystyle\int_0^{2\pi}\int_0^a |E(r,\phi)|r\,dr\,d\phi}$$

The phase error efficiency becomes

$$\text{PEL} = \exp\left(\frac{-4\pi\varepsilon_0}{\lambda}\right)^2 \exp\left(\frac{\pi\eta_0}{\lambda}\right)^4$$

The correlation of random errors increases the probable sidelobe level. The sidelobe level increases with the size of the correlation interval and decreases for larger aperture diameters. Increasing the amplitude taper of the distribution makes the aperture pattern more susceptible to random-error sidelobes, since increasing the taper is somewhat equivalent to decreasing the aperture diameter. Blockage and feed diffraction also limit the achievable sidelobe level in a reflector. A simple feed cannot carefully control the aperture distribution necessary for low sidelobes. Hansen [19, p. 74] discusses sidelobe limitations caused by random phase error in some detail.

Paraboloidal reflectors can be made in an umbrella shape where the ribs are parabolic and wire mesh is stretched between them [20]. The gore shape causes phase error loss and their periodicity produces extra sidelobes. Given the number of gores N_G and the focal length of the ribs f_r, the surface is given by

$$f(\psi) = f_r \frac{\cos^2(\pi/N_G)}{\cos^2\psi}$$

where ψ is measured from the centerline between the ribs. We calculate the average focal length by integrating across the gore half-angle π/N_G and dividing by π/N_G:

$$f_{\text{av}} = f_r \frac{\sin(2\pi/N_G)}{2\pi/N_G} \tag{8-25}$$

We use Eq. (8-25) to calculate the rib focal length given the average focal length of the reflector.

The peak sidelobe due to the periodic gores occurs at an angle θ_p found from the number of gores and the diameter D:

$$\theta_p = \sin^{-1}\left(1.2N_G\frac{\lambda}{\pi D}\right) \tag{8-26}$$

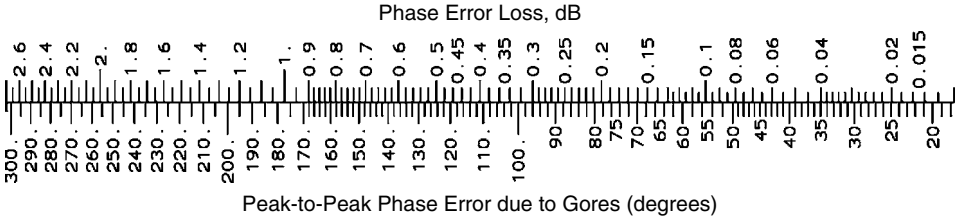
Given the average f/D of the reflector, we determine the peak-to-peak phase deviation across the gore by the approximate equation

$$\Delta = \frac{800 - 500(f/D - 0.4)}{N_G^2}\frac{D}{\lambda} \tag{8-27}$$

Scale 8-7 lists the phase error loss for a feed edge taper of 10 dB. Increasing the feed taper decreases the phase error loss due to gore construction. When we use a 20-dB feed taper, the values given by Scale 8-7 reduce by 0.16 dB for 0.5 dB of loss, 0.31 dB for 1 dB, and 0.45 dB for 1.5 dB. The gain losses due to underillumination by the 20-dB edge taper feed exceed these values.

Example Given a reflector with $D/\lambda = 35$ with a limit of 0.5 dB loss due to gore construction for $f/D = 0.34$, we discover that the allowable peak-to-peak phase error from Scale 8-7 is 124°. Using Eq. (8-27), we solve for the number of ribs:

$$N_G^2 = \frac{830}{124}35 = 234 \quad \text{or} \quad N_G = 16$$



SCALE 8-7 Phase error loss due to gore construction of a paraboloidal reflector.

We use Eq. (8-26) to compute the angle of the peak gore sidelobe, $\theta_p = 10.0^\circ$. Equation (8-27) shows that the phase deviation Δ is proportional to frequency. If the frequency increases by 1.5 times, then by Eq. (8-27), 124° increases to 186° and we read 1.1 dB of loss from Scale 8-7 while the peak gore sidelobe becomes $\theta_p = 6.7^\circ$.

8-8 FOCAL PLANE FIELDS

We improve the efficiency and pattern response of a reflector if we match the feed fields to the focal plane fields. GO assumes a point focus, but an actual focus is extended. We determine the reflector and feed efficiency from the field match over the focal plane. When the reflector f/D value is large, we use the diffraction pattern of a circular aperture, the Airy function:

$$E = \frac{J_1(kr\psi_0)}{kr\psi_0} \tag{8-28}$$

where ψ_0 is the half subtended angle of the reflector (radians), r the radial coordinate, k the propagation constant, and J_1 the Bessel function.

In a more exact method the currents induced on the reflector ($2\mathbf{n} \times \mathbf{H}$) and the magnetic vector potential are used to calculate the focal plane fields. As f/D decreases, the currents on the reflector interact and modify their distribution, but it is a secondary effect [21]. Iterative physical optics analysis (section 2-4) can find these current modifications. We calculate the reflector efficiency from the field match of the focal plane fields ($\mathbf{E}_1$, $\mathbf{H}_1$) and the fields of the feed ($\mathbf{E}_2$, $\mathbf{H}_2$) using Robieux's theorem [4]:

$$\eta = \frac{\left| \iint_S (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1) \cdot dS \right|^2}{4P_1P_2} \tag{8-29}$$

where P_1 and P_2 are the input powers to produce the fields and η is the efficiency. Equation (8-29) is the magnitude squared of Eq. (2-35), the reactance equation equivalence applied to Eq. (1-55) for the coupling between two antennas S_{21}. The finite size of the feed causes spillover. The extent of amplitude and phase mismatch between the two fields determines the efficiency. By illuminating the reflector with a cross-polarized wave, we compute the cross-polarization radiation level through its field match [Eq. (8-29)].

We maximize efficiency [Eq. (8-29)] by conjugate-matching the focal plane fields with the feed fields. Corrugated horns can be designed by expanding the focal plane

fields in axial hybrid modes of the horn and mode matching [22,23]. Wood [4] expands the reflector and feed fields in spherical harmonics and matches them at a boundary. Both sets of fields can be approximated very well by just a few terms, and this method can handle dual-reflector and offset reflector systems as well as axisymmetric prime focus reflectors.

We can feed the reflector with an array to match the focal plane fields [24–26]. The array samples the focal plane field and conjugate-matches it so that the powers sum in phase. The array can form multiple beams and also correct reflector aberrations [24]. By using the multiple feeds of the array, coma can be reduced for scanned beams and efficiency improved. However, quantization of the array element locations and excitations, amplitude, and phase reduces efficiency and raises the sidelobe level [27].

We apply Eq. (1-55) for the coupling between two antennas to determine the feeding coefficients of an array feed for a dish. Assume an incident field distribution on the reflector that includes the incident wave direction and the desired aperture distribution for the reflector. Using physical optics, we calculate the currents induced on the reflector surface. If the reflector has significant curvature so that the patches face each other, iterative PO can be used to account for their interaction. We calculate the fields radiated by each feed on the reflector surface and apply Eq. (1-55) to calculate coupling. This method applies the feed pattern to the calculation instead of the point matching used in a focal plane solution. Similar to scanning of an array, we use conjugative matching for the feed array elements to produce the beam desired. This method can determine array feed element amplitude and phase for any composite reflector aperture distribution that includes aperture distribution to control sidelobes or include multiple beams. The method reduces coma to the minimum possible with a given array.

Analysis finds the array distribution desired, but we do not achieve this distribution merely by designing the feed network to produce these amplitudes and phases because the feed elements have significant mutual coupling. We need to include the effect of the paraboloidal reflector when computing mutual coupling because the field radiated by one feed induces currents on the reflector that couple to other feed elements. Below we show that the effect of the reflector diminishes as the reflector diameter increases. If the mutual coupling is significant whether direct or due to the reflector, we need to apply the corrections given in Section 3-11 to adjust the feeding coefficients of the array.

8-9 FEED MISMATCH DUE TO THE REFLECTOR

The feed receives some of its transmitted power because it reflects from the parabola and returns as a mismatch at the feed terminals. We calculate the reflected field at the feed by using surface currents and the magnetic vector potential. The only significant contribution comes from areas near where the normal of the reflector points at the feed. Around every other point, the phase of the reflection varies rapidly and cancels and we need to consider only points of stationary phase. We calculate the reflection from each point of stationary phase from [2]

$$\Gamma = -j\frac{G_f(\rho_0)}{4k\rho_0}\sqrt{\frac{\rho_1\rho_2}{(\rho_1+\rho_0)(\rho_2+\rho_0)}}e^{-j2k\rho_0} \tag{8-30}$$

where Γ is the reflection coefficient, ρ_0 the distance to the stationary phase point, $G_f(\rho_0)$ the feed gain in the direction of ρ_0, and ρ_1 and ρ_2 the radiuses of curvature of the reflector at ρ_0. The vertex is the only point of stationary phase on a paraboloidal reflector: $\rho_1 = \rho_2 = -2f$ and $\rho_0 = f$. Equation (8-30) reduces to

$$\Gamma = -j\frac{G_f(0)}{2kf}e^{-j2kf} \tag{8-31}$$

Example Suppose that we have a reflector with $f/D = 0.40$. Compute reflector mismatch for a source with its 10-dB beamwidth equal to the reflector subtended angle.

Half subtended angle [Eq. (8-2)] $\psi_0 = 2\ \tan(1/1.6) = 64^\circ$. By using the feed approximation $\cos^{2N}(\theta/2)$, we have

$$N = \frac{\log 0.1}{2\log\cos(64^\circ/2)} = 6.98$$

The feed gain at the boresight is $N + 1$ [Eq. (1-20*c*)]:

$$\text{[Eq. (8-31)]} \qquad |\Gamma| = \frac{8\lambda}{4\pi f} = 1.59\frac{\lambda}{D}$$

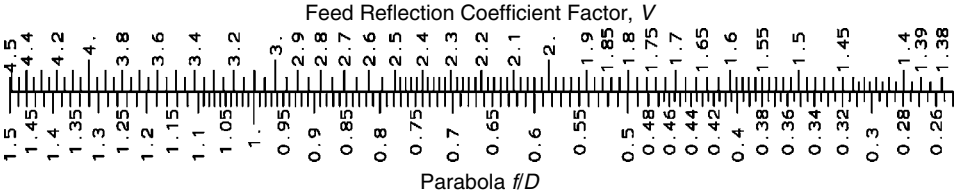
Increasing the reflector diameter in wavelengths decreases the reaction of the reflector on the feed. For example, given a 3-m reflector at 4 GHz, we calculate reflector reflection coefficient as 0.04, or VSWR = 1.08.

We can express the reflector reflection of a paraboloidal reflector as

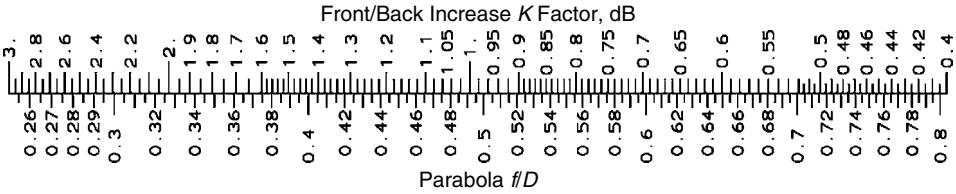
$$|\Gamma| = V\frac{\lambda}{D} \tag{8-32}$$

and calculate Scale 8-8 of V versus f/D for feeds with 10-dB beamwidths equal to the reflector subtended angle. Higher reflector f/D values produce larger feed reflections, since the feed gain increases faster than the reduced area of the reflector seen from the feed.

Narrowband corrections to these reflections can be designed by using a vertex plate (Silver [2]) or by designing sets of concentric ring ridges in the reflector (Wood [4]). The rings can match the feed at more than one frequency. By any of these methods, the free-space mismatch of the feed could be corrected for, but, of course, the feed itself can be mismatched to compensate for the reflector reaction.



SCALE 8-8 Feed reflection scale factor V given f/D.



SCALE 8-9 Paraboloidal reflector front-to-back ratio increase K given f/D.

8-10 FRONT-TO-BACK RATIO

Figure 2-9 illustrates the pattern response of a paraboloidal reflector and shows that the pattern behind the reflector peaks along the axis. The diffractions from all points along the rim add in-phase along the axis and produce a pattern peak. We can reduce this rim diffraction by using a rolled, serrated, or castellated edge to reduce diffraction. An absorber-lined cylindrical shroud extending out to enclose the feed will greatly reduce back radiation, including spillover, and allows the close spacing of terrestrial microwave antennas with reduced crosstalk.

For a normal truncated circular reflector rim, the following equation estimates the front-to-back ratio given the reflector gain G, the feed taper T, and feed gain G_f [28]:

$$\text{F/B} = G + T + K - G_f \qquad \text{dB} \tag{8-33}$$

The constant K, given by Scale 8-9, is related to f/D:

$$K = 10 \log \left[1 + \frac{1}{(4f/D)^2} \right] \tag{8-34}$$

Example Estimate F/B for a reflector with $f/D = 0.34$ and 40 dB of gain.

We read the feed subtended angle from Scale 8-1 to be 143°. A 10-dB edge taper feed has a gain of about 8.1, found from Scale 1-2. Using Eq. (8-33), we estimate $\text{F/B} = 40 + 10 + 1.9 - 8.1 = 43.8$ dB.

8-11 OFFSET-FED REFLECTOR

Moving the feed out of the aperture eliminates some of the problems with axisymmetrical reflectors. Blockage losses and diffraction-caused sidelobes and cross-polarization disappear. We can increase the size of the feed structure and include more if not all of the receiver with the feed. For example, the reflector may be deployed from a satellite, with the feed mounted on the main satellite body.

Figure 8-7 shows the offset-fed reflector geometry. We form the reflector out of a piece of a larger paraboloid. Every piece of the paraboloidal reflector converts spherical waves from the focus into a plane wave moving parallel with its axis. We point the feed toward the center of the reflector to reduce the spillover, but we still locate the feed phase center at the focus of the reflector. The aperture plane projects to a circle, although the rim shape is an ellipse. ψ_0 is the angle from the axis of the parabola to

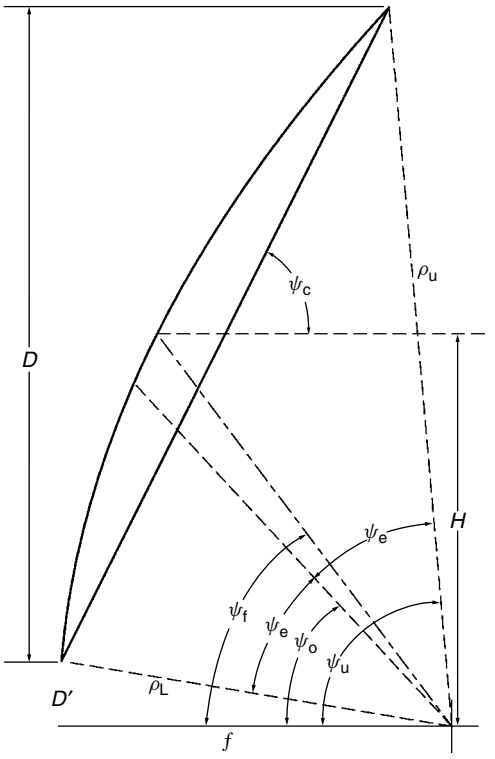


FIGURE 8-7 Parameters of an offset-fed parabolic reflector.

the center of the cone of the reflector, and the reflector subtends an angle $2\psi_e$ about this centerline. Given the aperture plane diameter D and the height H of the center, we find the lower rim offset $D' = H - D/2$. From these parameters we determine the angle of the center of the rim cone from the z-axis:

$$\psi_0 = \tan^{-1} \frac{16fH}{16f^2 + D^2 - 4H^2} = \tan^{-1} \frac{2f(D + 2D')}{4f^2 - D'(D + D')} \tag{8-35}$$

The half cone angle defines the rim:

$$\psi_e = \tan^{-1} \frac{8fD}{16f^2 + 4H^2 - D^2} = \tan^{-1} \frac{2fD}{4f^2 + D'(D + D')} \tag{8-36}$$

We direct the feed an angle ψ_f from the z-axis to the center of the projected diameter different from the angle ψ_0 of the rim cone axis:

$$\psi_f = 2 \tan^{-1} \frac{H}{2f} = 2 \tan^{-1} \frac{2D' + D}{4f} \tag{8-37}$$

The rim lies in a plane at an angle ψ_c with respect to the z-axis:

$$\psi_c = \tan^{-1} \frac{2f}{H} = \tan^{-1} \frac{4f}{2D' + D} \tag{8-38}$$

The rim is an ellipse in this plane with major and minor axes given by

$$a_e = \frac{D}{2\sin\psi_c} \quad \text{and} \quad b_e = \frac{D}{2} \tag{8-39}$$

The offset angle modifies the f/D of the reflector:

$$\frac{f}{D} = \frac{\cos\psi_e + \cos\psi_0}{4\sin\psi_e} \tag{8-40}$$

We calculate the rim offset from the cone angles:

$$D' = 2f\tan\frac{\psi_0 - \psi_e}{2} \tag{8-41}$$

Manufacturing an offset reflector requires specification of the reflector when laid on its rim in the $x-y$ plane so that the mold can be machined. We center the major axis of the reflector elliptical rim $L = 2a_e$ along the x-axis and the minor axis D along the y-axis. In this position the reflector depth $d(x, y)$ is found from the expression [29]

$$d(x,y) = \frac{2fL^3}{D(L^2-D^2)}\left[\sqrt{1 + \frac{xD^2\sqrt{L^2-D^2}}{fL^3} + \frac{D^2(L^2-D^2)}{4f^2L^4}\left(\frac{D^2}{4} - y^2\right)} - 1 - \frac{xD^2\sqrt{L^2-D^2}}{2fL^3}\right] \tag{8-42}$$

The deepest point of the reflector $d_{\max}$ occurs along the x-axis at x_b:

$$x_b = -\frac{D^2\sqrt{L^2-D^2}}{16fL} \qquad \text{where} \quad d_{\max} = \frac{D^3}{16fL} \tag{8-43}$$

After measuring D, L, and $d_{\max}$, we determine the offset focal length from the equation

$$f = \frac{D^3}{16Ld_{\max}} \tag{8-44}$$

We calculate the center height of the offset from

$$H = 2f\sqrt{\frac{L^2}{D^2} - 1} \tag{8-45}$$

We calculate the reflector half cone angle ψ_e and the cone axis angle from the z-axis and ψ_0 from the focal length f, ellipse major diameter L, and minor diameter D:

$$\begin{bmatrix}\psi_e \\ \psi_0\end{bmatrix} = \tan^{-1}\left(\sqrt{\frac{L^2}{D^2} - 1} + \frac{D}{4f}\right) \mp \tan^{-1}\left(\sqrt{\frac{L^2}{D^2} - 1} - \frac{D}{4f}\right) \tag{8-46}$$

To align the reflector, we use the angle of the reflector rim major axis $\psi_c = \sin^{-1}(D/L)$ with respect to the z-axis and the radial distances from the lower and upper edges of the reflector in the offset plane, since the center offset H is not a distinguishable point:

$$\begin{bmatrix} \rho_U \\ \rho_L \end{bmatrix} = \frac{fL^2}{D^2} + \frac{D^2}{16f} \pm D\left(\frac{L^2}{D^2} - 1\right) \tag{8-47}$$

We analyze the offset reflector with the same tools as those used with the axisymmetric reflector: aperture field, physical optics, and GTD. The asymmetry of the reflector to feed geometry introduces anomalies. Huygens sources no longer eliminate cross-polarization, because the source must be tilted. Symmetry prevents cross-polarization in the plane containing the x-axis (Figure 8-8), but cross-polarization for linear polarization increases in the plane containing the y-axis (symmetry plane) as f/D decreases

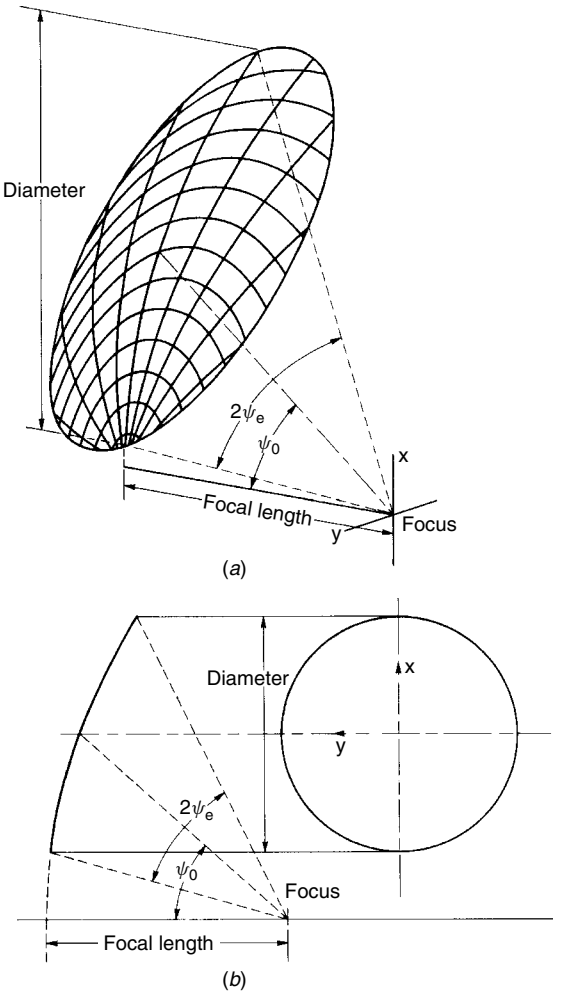


FIGURE 8-8 Offset-fed paraboloidal reflector geometry: (*a*) perspective; (*b*) orthographic representation.

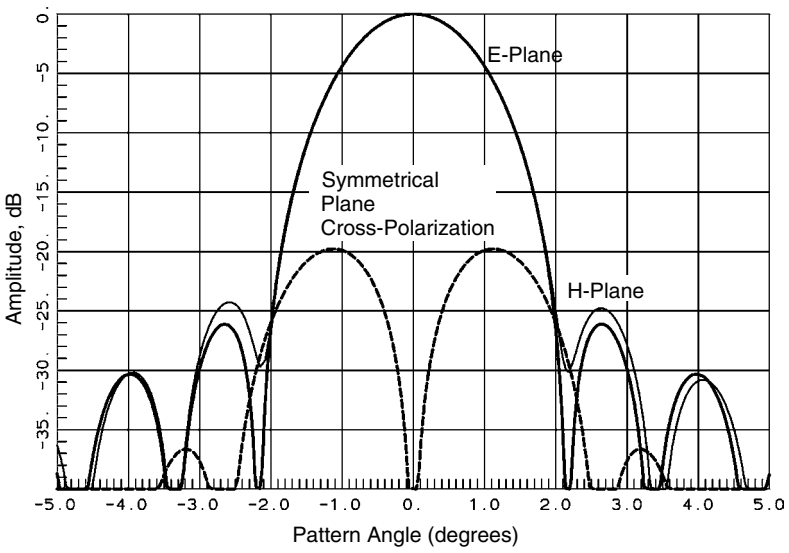


FIGURE 8-9 Pattern of an offset-fed reflector with linearly polarized feed.

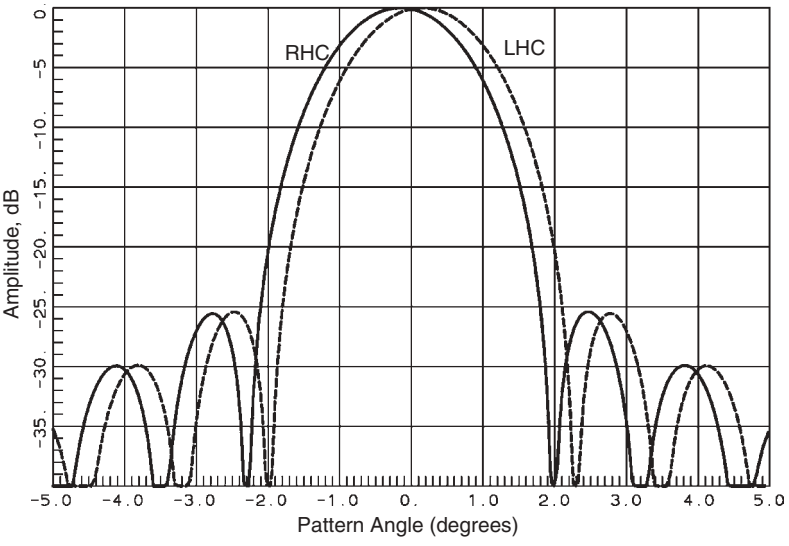


FIGURE 8-10 Pattern of an offset-fed reflector with circularly polarized feed.

(Figure 8-9). The Condon lobes move off the diagonal planes and into the plane containing the y-axis. The asymmetry along the x-axis tapers the amplitude distribution from a symmetrical feed, since the spherical wave travels farther to the outer edge of the reflector than to the lower edge. The offset-fed reflector geometry squints circularly polarized pattern peaks in the symmetrical (y-axis) plane without generating cross-polarization (Figure 8-10). An approximate formula for the squint is [9]

$$\psi_s = \sin^{-1} \frac{\lambda \sin \psi_0}{4\pi f} \tag{8-48}$$

where ψ_s is the squint angle. Opposite senses of circular polarization squint in opposite directions and cause a problem with dual circularly polarized feed systems. In all cases increasing the f/D or the effective f/D through a subreflector reduces these problems. The subreflector should be kept out of the aperture of the main reflector. We can feed-scan the offset-fed reflector by moving the feed laterally along a line that lies perpendicular to the boresight of the feed (the line defined by ψ_0). We must modify the beam deviation factor (BDF):

$$\text{BDF}_{\text{offset fed}} = \text{BDF}_{\text{center fed}} \frac{(f/D)_{\text{offset}}}{(f/D)_{\text{center fed}}} \tag{8-49}$$

Example Given an offset-fed reflector with $\psi_0 = 45^\circ$ and $\psi_e = 40^\circ$, compute the beam deviation factor.

From Eq. (8-49),

$$\left(\frac{f}{D}\right)_{\text{center fed}} = \frac{\cos 40^\circ + 1}{4 \sin 40^\circ} = 0.687$$

$$\left(\frac{f}{D}\right)_{\text{offset fed}} = \frac{\cos 40^\circ + \cos 45^\circ}{4 \sin 40^\circ} = 0.573$$

From Table 8-1 we interpolate $\text{BDF}_{\text{center fed}} = 0.928$, and we substitute the values into Eq. (8-49) to calculate $\text{BDF}_{\text{offset fed}} = 0.774$. We must laterally offset the feed farther than with a center-fed reflector to achieve the same feed scanning.

Periscope Configuration The periscope consists of an offset paraboloidal reflector with $\psi_0 = 90^\circ$ with a long focal length fed by a paraboloidal reflector located at the focus. This eliminates the need to run a transmission line up a tower. Periscope antennas can be made using a flat-plate reflector, but the long focal length means that the parabolic splash plate antenna has only a small deviation from flat. The flat plate is limited to a gain of only 6 dB more than the feed reflector for optimum conditions with a large plate. The gain of the offset paraboloidal reflector is determined by the diameter of the splash plate, not the feed reflector. Because the splash plate is in the near field of the feeding reflector, gain is reduced by phase error, whereas spillover and amplitude taper losses also contribute to gain loss.

Design starts with determining the splash reflector center height H required to clear obstacles along the transmission path. We calculate the splash reflector aperture diameter from the required gain and beamwidth. The periscope configuration contributes to gain loss, but with proper selection of the feed paraboloidal reflector these losses are minor and can be compensated for by using a larger splash reflector. Having the splash reflector directly overhead corresponds to a parent reflector design with $f/D_p = 0.25$ and $f = H/2$. An analysis using a radial parabolic aperture distribution in the feed reflector determined that the optimum feed reflector diameter is found from the ratio of height to projected splash reflector aperture diameter D_s [30]:

$$D_f = \frac{2\lambda H}{D_s} = \frac{\alpha \lambda H}{D_s} \quad \text{or} \quad \alpha = \frac{D_s D_f}{\lambda H} = \frac{D_s D_f F}{Hc} \tag{8-50}$$

Whereas $\alpha = 2$ is the optimum dimensions at a particular frequency, we account for shift from the optimum with this factor. The parameter α is the frequency response

factor for frequency F and speed of light c. The illumination efficiency is the product of the feed reflector illumination efficiency and the periscope efficiency factor η_p:

$$\eta_p = \frac{4\left[1 - (1 - K)J_0(m) - K(2/m)J_1(m)\right]}{m^2(1 - K/2)^2} \tag{8-51}$$

J_0 and J_1 are Bessel functions, $m = \alpha\pi/2$, and $K = 1 - 10^{-[\mathrm{ET(dB)}/20]}$ for the feed reflector edge illumination taper ET(dB). Table 8-3 lists the added illumination loss of a periscope, given geometry using α in Eq. (8-50) for a 12-dB edge taper in a feeding reflector.

Example A periscope antenna system placed a 3-m projected aperture splash reflector 30 m above the feed reflector to operate at 12 GHz ($\lambda = 0.05$ m). Using Eq. (8-50), we calculate the feed reflector diameter to be 0.5 m for $\alpha = 2$. If we assume that the feed reflector has an efficiency of 60% (−2.22 dB), the efficiency of the splash reflector will be −2.6 dB:

$$\text{gain (dB)} = 20\log\frac{\pi D_s}{\lambda} - 2.6 = 20\log\frac{3\pi}{0.05} - 2.6 = 42.9$$

The focal length f of the splash reflector is $H/2 = 15$ m. Since the angle of the splash reflector rim is 45°, $L = D_s/\sin(45^\circ) = 4.24$ m. We determine the maximum depth of the reflector by using Eq. (8-43) to be 2.65 cm located 2.65 cm off center. The splash reflector increased the gain relative to the feed reflector by 7.4 dB.

8-12 REFLECTIONS FROM CONIC SECTIONS

We use reflectors made from conic sections other than the parabola as subreflectors. The ellipse and hyperbola rotated about their axes to form solid figures that reflect incident spherical waves into spherical waves with different caustics (focal points). Reflectors formed by moving the figure along a line change the caustics of cylindrical waves. We consider only spherical waves, but we need only convert to cylindrical waves for cylindrical reflectors.

All conic-section reflectors convert spherical waves from one focus into spherical waves directed toward the other focus. The ellipse has its two focuses located within the

TABLE 8-3 Added Illumination Loss of a Periscope, Given Geometry Using α in Eq. (8-50) for 12-dB Edge Taper in a Feeding Reflector (dB)

α	η_p	α	η_p
1.0	3.17	2.2	0.45
1.2	2.06	2.4	0.68
1.4	1.28	2.6	1.04
1.6	0.77	2.8	1.52
1.8	0.48	3.0	2.11
2.0	0.38	3.2	2.79
		3.4	3.53

figure. As we let one focus approach infinity, the ellipse transforms into a parabola. If the focus is pushed through infinity to the negative axis, the figure becomes a hyperbola located between the two focuses. Figure 8-11 shows the ray tracing for axisymmetrical conic section reflectors. A spherical source at one focus is reflected to the second focus by the reflector, although it is virtual (not actually reached) in some cases.

We describe all conic sections with the same polar equation:

$$\rho = \frac{eP}{1 - e\cos\theta} \tag{8-52}$$

where P is the distance between the origin, the focus, to a line called the *directrix* (Figure 8-12). The eccentricity e is the ratio of the distance from the origin to a point on the curve to the distance from the same point to the directrix: $r_1 = er_2$. In an ellipse, $e < 1$; in a parabola, $e = 1$; and in a hyperbola, $e > 1$. The distance between the focuses is

$$2c = \frac{2Pe^2}{1 - e^2} \tag{8-53}$$

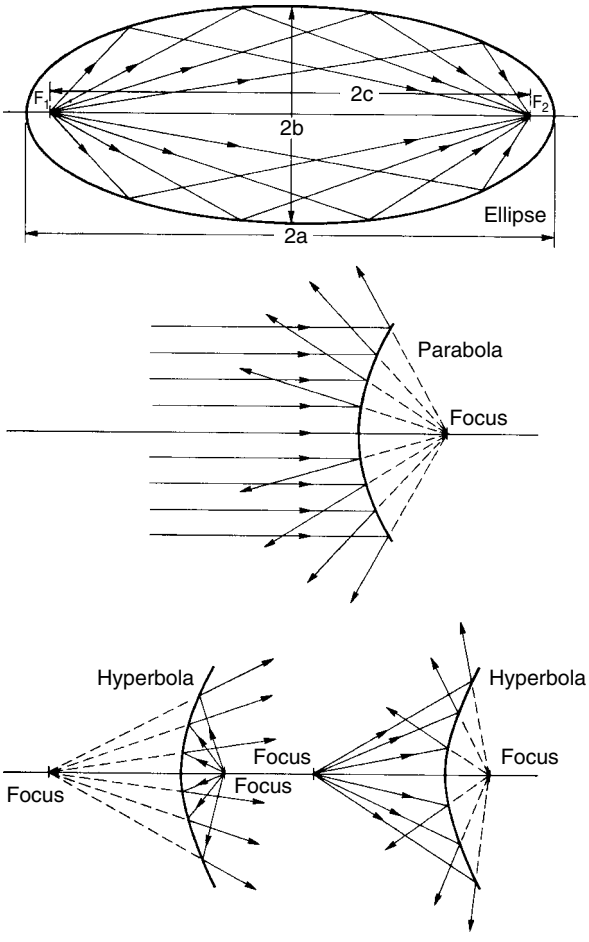


FIGURE 8-11 Reflections from conic-section reflectors.

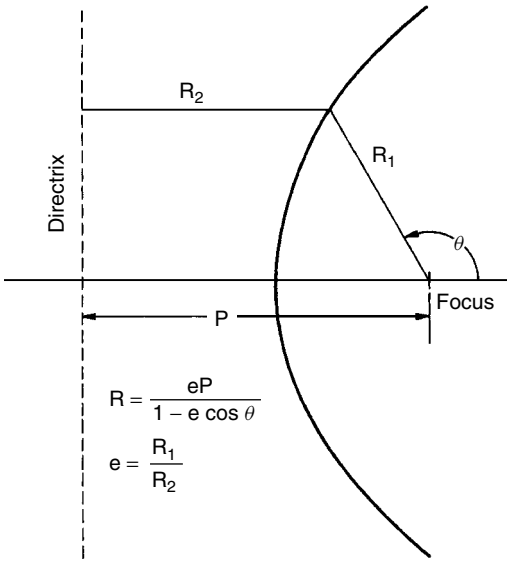


FIGURE 8-12 Conic-section geometry.

A hyperbola with its axis containing the two focuses along the z-axis, located at $\pm c$, intersects the z-axis at $\pm a$ and satisfies the equation

$$\frac{z^2}{a^2} - \frac{r^2}{b^2} = 1 \qquad \text{where} \quad b^2 = c^2 - a^2 \quad \text{and} \quad e = \frac{c}{a} > 1 \tag{8-54}$$

When we take the portion of the hyperbola along the $+z$-axis that intersects the axis at $+a$, we define the angles from the two focuses from the line between them because we place the feed at the left focus and locate a parabola focus at the right hyperbola focus. The left angle θ is the feed angle, and the right angle ψ is the parabola angle in a dual-reflector antenna. Given a point on the hyperbola, the distance from the left focus is ρ_1 and the distance from the right focus is ρ_2:

$$\rho_1 = \frac{a(e^2 - 1)}{e\cos\theta - 1} = \frac{b^2}{e\cos\theta - 1} \quad \text{and} \quad \rho_2 = \frac{a(e^2 - 1)}{e\cos\psi + 1} = \frac{b^2}{e\cos\psi + 1} \tag{8-55}$$

We determine the radial position off the axis from either polar equation:

$$r = \rho_1 \sin\theta = \rho_2 \sin\psi \tag{8-56}$$

The two angles are related by the eccentricity e:

$$(e + 1)\tan\frac{\theta}{2} = (e - 1)\tan\frac{\psi}{2} \tag{8-57}$$

At a given point on the hyperbola, the angle of the normal u relative to the radial line ρ_1 is half the sum of the two angles, $u = (\theta + \psi)/2$. We need the radius of curvatures in the principal planes to apply Eq. (2-77) for reflection from a curved surface: R_1 in

the $r-z$ plane and R_2 in the $\phi-z$ plane:

$$R_1 = \frac{b^2}{a\cos^2 u} \quad \text{and} \quad R_2 = \frac{b^2}{a\cos u} \tag{8-58}$$

When we offset-feed a hyperboloid, a cone determines the rim and it lies in a planar ellipse. Similar to the offset paraboloid, we center the cone at an angle θ_0 from the axis and define the cone by the half feed edge angle θ_e. We compute the distances from the focus to the upper and lower rim along the major axis of the rim ellipse:

$$\rho_L = \frac{b^2}{e\cos(\theta_0 - \theta_e) - 1} \quad \text{and} \quad \rho_U = \frac{b^2}{e\cos(\theta_0 + \theta_e) - 1}$$

We determine the major axis diameter ($2a_e$) of the elliptical rim from the triangle with sides ρ_L and ρ_U and angle $2\theta_e$ between them:

$$2a_e = \sqrt{\rho_L^2 + \rho_U^2 - 2\rho_L\rho_U\cos 2\theta_e} \tag{8-59}$$

The minor axis diameter ($2b_e$) is given by the equation

$$2b_e = \sqrt{(2a_e)^2 - (\rho_L - \rho_U)^2} \tag{8-60}$$

An ellipsoid can also be used as a subreflector in a dual-reflector antenna. Its equations are similar to the hyperboloid:

$$\frac{z^2}{a^2} + \frac{r^2}{b^2} = 1 \quad \text{and} \quad b^2 = a^2 - c^2 \quad \text{with} \quad e = \frac{c}{a} < 1 \tag{8-61}$$

We locate the ellipse with its major axis along the z-axis and use the portion that intersects the $+z$-axis in a dual-reflector antenna. We place a feed at the left focus and the focal point of a paraboloid at the right ellipse focus. The radial distances to the two focuses are given by the equations

$$\rho_1 = \frac{a(e^2-1)}{1 - e\cos\theta} = \frac{b^2}{1 - e\cos\theta} \quad \text{and} \quad \rho_2 = \frac{a(e^2-1)}{1 + e\cos\psi} = \frac{b^2}{1 + e\cos\psi} \tag{8-62}$$

The ellipse uses Eqs. (8-56) and (8-57) without change, but the angle of the normal u relative to the vector ρ_1 is the average difference between the two angles, $u = (\theta - \psi)/2$. Using the new angle u, Eq. (8-58) gives the principal radiuses of curvature at the point on an ellipsoid.

When we offset an ellipsoid by determining the rim as the intersection of an offset cone with it, we use the left equation of Eq. (8-62) for the upper and lower radiuses. The rim is a planar ellipse whose major and minor diameters are found using Eqs. (8-59) and (8-60), a consequence of Eq. (8-52) that defines both an ellipse and a hyperbola.

8-13 DUAL-REFLECTOR ANTENNAS

We derive dual-reflector antennas, Cassegrain and Gregorian, from their optical telescope counterparts. Each increases the effective focal length. In the Cassegrain dual-reflector antenna a hyperbolic subreflector is used (Figure 8-13), and in the Gregorian

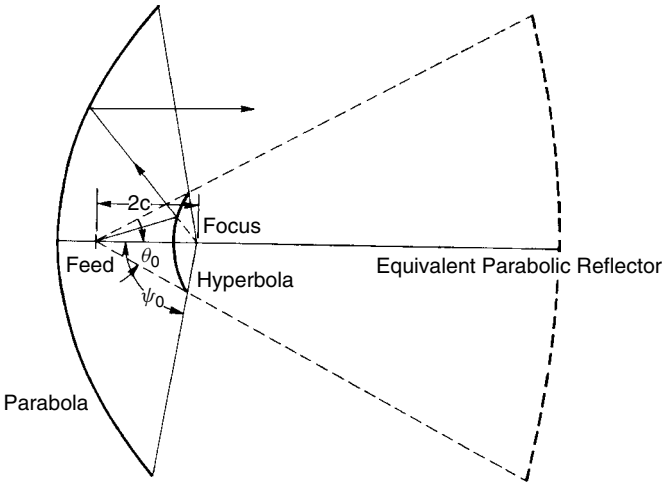


FIGURE 8-13 Cassegrain reflector antenna.

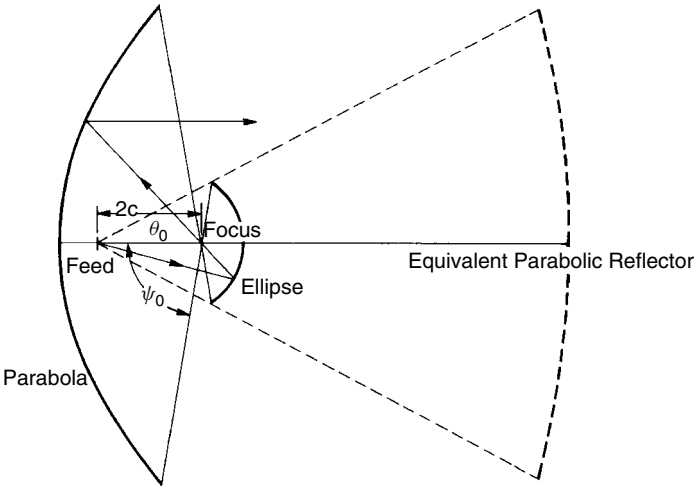


FIGURE 8-14 Gregorian reflector antenna.

dual-reflector antenna an elliptical subreflector is used (Figure 8-14). We locate one focus of the subreflector at the focus of the main paraboloidal reflector and place the second focus of the subreflector at the feed antenna phase center. The subreflector changes the curvature of waves coming from one focus into waves with their caustic at the second subreflector focus.

The number of Cassegrain designs exceeds the number of Gregorian dual reflectors. The Gregorian design requires a larger subreflector support because it extends farther from the main reflector vertex. The subreflector edge curvature increases the diffraction and reduces the control of the field incident on the main reflector, but by shaping the subreflector of the Cassegrain, we can increase overall efficiency. The inversion of the fields in the Gregorian reflector complicates such procedures.

In Figures 8-13 and 8-14 the subtended angle of the main reflector is $2\psi_0$, but the effective subtended angle at the feed is $2\theta_0$. We can calculate subreflector eccentricity from those angles.

$$\begin{array}{cc} \text{Cassegrain} & \text{Gregorian} \\ e = \dfrac{\sin\frac{1}{2}(\psi_0 + \theta_0)}{\sin\frac{1}{2}(\psi_0 - \theta_0)} & e = \dfrac{\sin\frac{1}{2}(\psi_0 - \theta_0)}{\sin\frac{1}{2}(\psi_0 + \theta_0)} \end{array} \tag{8-63}$$

We can analyze the dual-reflector system by using an equivalent parabola (Figures 8-13 and 8-14) from the subreflector subtended angle $2\theta_0$. We define the magnification factor as the ratio of the effective focal length to the actual focal length of the main reflector: $M = f_e/f$. The subreflector eccentricity can be calculated from M:

$$\begin{array}{cc} \text{Cassegrain} & \text{Gregorian} \\ e = \dfrac{M+1}{M-1} & e = \dfrac{M-1}{M+1} \end{array} \tag{8-64}$$

Equation (8-53) gives the distance between the focuses of the subreflectors:

$$\begin{array}{cc} \text{Cassegrain} & \text{Gregorian} \\ 2c = \dfrac{2Pe^2}{e^2-1} & 2c = \dfrac{2Pe^2}{1-e^2} \end{array} \tag{8-65}$$

We gain the design parameter P and with it some freedom in the placement of the feed. We easily solve for the length P in terms of the distance between focuses $2c$:

$$\begin{array}{cc} \text{Cassegrain} & \text{Gregorian} \\ P = \dfrac{2c(e^2-1)}{2e^2} & P = \dfrac{2c(1-e^2)}{2e^2} \end{array} \tag{8-66}$$

The subreflector diameter varies with P:

$$D_s = \frac{2eP\sin(\pi - \psi_0)}{1 - e\cos(\pi - \psi_0)} \tag{8-67}$$

We easily manipulate the geometry to find the distance from the main reflector vertex to the feed focus L_m and from the feed focus to the subreflector L_s for both the Cassegrain and Gregorian reflectors. $L_m = f - 2c$ and $L_s = a + c = c\,(1 + 1/e)$. Additional sets of equations are available to obtain the geometry of the dual reflector for various sets of specified input parameters [31].

8-13.1 Feed Blockage

The increased effective focal length requires feeds with narrow beamwidths, and we can no longer consider the feed as a point source. It projects a shadow into the center of the reflector (Figure 8-15) and causes a central blockage. The subreflector also blocks the center. As we reduce the subreflector diameter to reduce blockage, the feed antenna moves closer to the subreflector and its projected shadow increases. The optimum occurs when the projected feed blockage diameter equals the subreflector diameter.

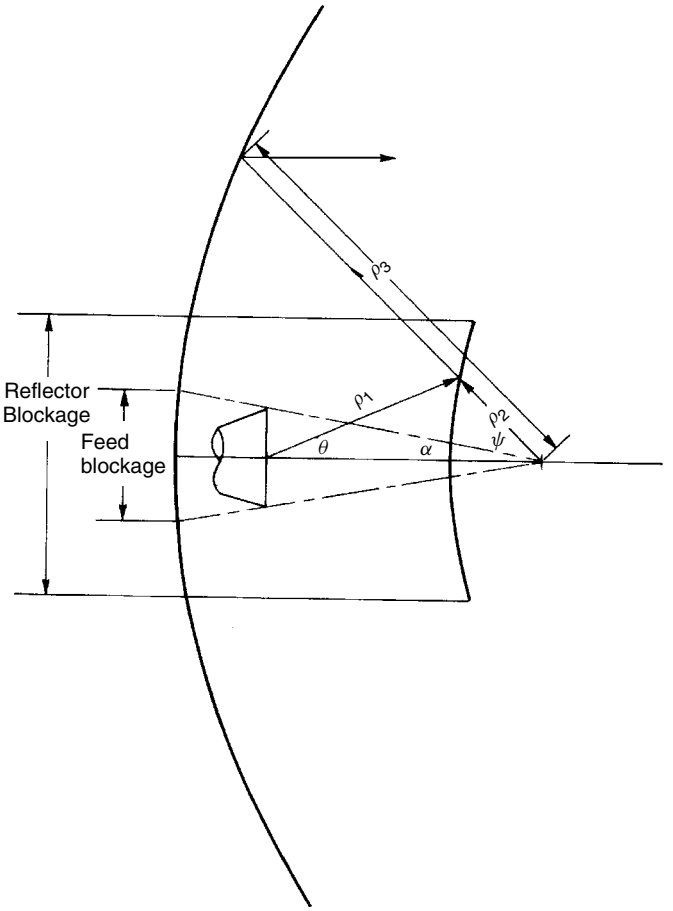


FIGURE 8-15 Cassegrain central blockage.

The feed size depends on the frequency of operation and the effective f/D value, whereas the subreflector diameter depends only on geometry. We cannot determine the optimum independent of frequency.

Example Design a 10-m-diameter main reflector Cassegrain antenna with $f/D = 0.3$ and an effective $f/D = 1.5$ to operate at 3.9 GHz. Minimize the aperture blockage.

$$M = \frac{1.5}{0.3} = 5 \quad \text{[Eq. (8-64)]} \qquad e = \frac{5+1}{5-1} = 1.5$$

From Eq. (8-2),

$$\theta_0 = 2\tan^{-1}\frac{1}{4(1.5)} = 18.9^\circ$$

We feed the reflector with a circular corrugated horn with a 10-dB beamwidth equal to the subtended angle of the subreflector. We use the methods of Section 7-3 to design the horn. When we include the thickness of the corrugations, we calculate a horn diameter

of 0.415 m. We determine the blockage angle α (Figure 8-15): $\alpha = \tan^{-1}(0.415/4c)$. From Eq. (8-56), $2c = 3.6P$. We use Eq. (8-67) with the parabola P_0 and α to compute the projected feed blockage:

$$P_0 = 2\frac{f}{D}D = 6\,\text{m} \qquad \frac{0.415}{4c} = \frac{0.0577}{P}$$

$$D_{\text{fb}} = \frac{2P_0 \sin(\pi - \alpha)}{1 - \cos(\pi - \alpha)} = \frac{12 \sin \tan^{-1}(0.0577/P_0)}{1 + \cos \tan^{-1}(0.0577/P_0)}$$

The subreflector blockage is given by Eq. (8-67) also:

$$\psi_0 = 2 \tan^{-1} \frac{1}{4(0.3)}$$

$$D_s = \frac{2(1.5)P_0 \sin(\pi - \psi_0)}{1 - 1.5\cos(\pi - \psi_0)} = 2.322P_0$$

We equate these blockages and solve the transcendental equation for P numerically:

$$P_0 = 0.385\,\text{m}$$

$$D_s = 0.894\,\text{m} \qquad \text{subreflector diameter}$$

$$2c = 1.386\,\text{m} \qquad \text{distance between subreflector foci}$$

We read the blockage loss from Table 4-40 using the Gaussian distribution, 0.14 dB. A Gregorian subreflector designed by parallel steps has a diameter about 6 cm larger and has about 0.02 dB more blockage loss.

Since the phase centers of most horns lie inside the bell along the axis, we move the feed horn toward the subreflector and increase the shadow blockage of the feed on the main reflector. We find the parameters of the reflector system to equalize feed shadowing to subreflector blockage by accounting for the phase-center distance D_{pc} and the diameter of the feed horn D_f by solving a quadratic equation. Given the main reflector diameter D, focal length f, and subreflector half subtended angle θ_0, we solve the quadratic equation for roots X_1 [32]:

$$[8fD - \sigma \tan\theta_0(16f^2 - D^2)]X_1^2 - 16D_{\text{pc}} \tan\theta_0 fDX_1 - 16D_f f^2 D \tan\theta_0 = 0$$

The parameter σ equals -1 for a Cassegrain dual reflector and $+1$ for a Gregorian system. We use X_1 to solve for c, the focal length of the conic-section subreflector:

$$c = X_1 \frac{8fD - \sigma \tan\theta_0(16f^2 - D^2)}{32fD \tan\theta_0}$$

We compute e from the magnification factor from the effective f/D [Eq. (8-64)] or from the feed half-subtended angle θ_0 and the main half-subtended angle ψ_0 [Eq. (8-63)]. The rest of the parameters of the dual-reflector antenna follow from these parameters. The dual-reflector antenna geometry can be found from various other specified parameters [32].

8-13.2 Diffraction Loss [3, 6, 9, 33]

We cannot decrease the subreflector size to reduce the central blockage without penalty. We based the subreflector design on geometric optics (GO), which assumes large reflectors in terms of wavelengths. We size the subreflector to have all the spillover and none from the main reflector based on GO. The finite subreflector size produces diffractions that cause main-reflector spillover, cross-polarized subreflector reflections, phase error losses, and additional amplitude taper losses. We lump these excess losses into a diffraction loss term.

Rusch [3,33] finds these losses by using vector diffraction theory of PO where he calculates the currents on the subreflector. Similar results can be calculated by using GTD methods [6]. The results of GTD calculations for a circularly polarized feed are listed in Table 8-4. They are similar to the results given for a TE_{11}-mode circular horn feed [9]. The loss depends on the effective focal length and the diameter of the subreflector. Increasing the feed taper lowers the diffraction loss but it has a broad flat optimum at about a 12-dB edge taper. Equalizing the feed and subreflector blockage may not lead to the optimum (highest-gain) design. We must trade off diffraction loss and blockage loss.

Example Optimize the 10-m-diameter Cassegrain dual reflector for the sum of diffraction and blockage losses at 3.9 GHz.

TABLE 8-4 Diffraction Loss of a Cassegrain Antenna with a 10- and a 15-dB Feed Edge Amplitude Taper (dB)

Subreflector Diameter (λ)	Effective f/D					
	0.75	1.00	1.5	2.0	2.5	3.0
			For 10 dB			
6	0.67	0.81	1.02	1.43	1.85	2.28
8	0.55	0.68	0.78	1.09	1.41	1.75
10	0.48	0.58	0.64	0.89	1.14	1.40
12	0.43	0.51	0.55	0.76	0.97	1.17
14	0.39	0.46	0.48	0.67	0.84	1.01
16	0.36	0.42	0.43	0.60	0.75	0.88
20	0.31	0.37	0.36	0.50	0.62	0.72
30	0.24	0.27	0.25	0.37	0.45	0.49
40	0.20	0.24	0.20	0.30	0.36	0.39
60	0.16	0.20	0.12	0.22	0.27	0.28
100	0.11	0.15	0.09	0.17	0.20	0.22
			For 15 dB			
6	0.53	0.66	0.91	1.29	1.72	2.17
8	0.42	0.54	0.68	0.96	1.29	1.64
10	0.36	0.45	0.55	0.77	1.02	1.30
12	0.32	0.39	0.46	0.65	0.85	1.07
16	0.26	0.31	0.36	0.49	0.63	0.78
20	0.22	0.27	0.29	0.40	0.51	0.62
50	0.12	0.15	0.13	0.19	0.22	0.24
100	0.07	0.10	0.11	0.12	0.14	0.14

TABLE 8-5 Subreflector Diameter Trade-off

Diameter (λ)	Blockage (dB)	Diffraction Loss (dB)	Sum (dB)
11.6	0.14	0.57	0.71
14.0	0.20	0.48	0.68
16.0	0.26	0.43	0.69
18.0	0.33	0.40	0.73

The subreflector diameter is 0.894 m, or 11.62λ. We can increase the subreflector diameter without affecting the results of the preceding example, and the feed blockage will be less than the subreflector blockage. We generate Table 8-5 from Tables 4-40 and 8-4. The optimum occurs at about 15λ diameter, or 1.154 m. The table shows a broad optimum, with variations of ±λ having no practical effect. We move the feed back to illumine the larger subreflector:

$$2c = 3.6P = \frac{3.6D_s}{2.322} = 1.789\,\text{m}$$

Kildal [34] has derived a simple formula for the optimum subreflector size to minimize the combination of blockage and diffraction losses:

$$\frac{d}{D} = \left[\frac{\cos^4(\theta_0/2)}{(4\pi)^2 \sin\psi_0} E \frac{\lambda}{D}\right]^{1/5}$$

where d is the subreflector diameter, D the main-reflector diameter, and E the power illumination at the edge of the subreflector (0.1 for 10 dB). The approximate product of the efficiencies (blockage and diffraction) at the optimum d/D is

$$\eta \approx \left\{1 - C_b\left[1 + 4\sqrt{1 - \frac{d}{D}}\right]\left(\frac{d}{D}\right)^2\right\}^2$$

where $C_b = -\ln\sqrt{E}/(1 - \sqrt{E})$. This analysis gives a good first estimate of the proper subreflector size and associated loss. Beyond this rule, extending the subreflector diameter by 1λ to 2λ beyond the GO design reduces the diffraction loss of a Cassegrain significantly even though the blockage loss increases. The reduction in diffraction is much greater than the increased blockage loss. This approach does not work with a Gregorian reflector, because it already has a lower diffraction loss, due to the convex subreflector.

8-13.3 Cassegrain Tolerances [9, 35]

To calculate the feed movement effects, reduce them by the magnification factor M and use the f/D value of the main reflector. The result is the reduction of the tight location tolerances normally required of the deep (small f/D) main reflector. For feed scanning we must multiply the offset required for the main reflector by the magnification factor to compute the necessary lateral offset. Translation of the hyperboloid is nearly equivalent

TABLE 8-6 Scale Factor for RMS Surface Deviations of a Dual Reflector

Eccentricity, e	Main Reflector, f/D					
	0.25	0.30	0.35	0.40	0.45	0.50
0.5	1.259	1.183	1.136	1.105	1.083	1.068
0.6	1.203	1.143	1.107	1.082	1.066	1.053
0.7	1.148	1.105	1.078	1.060	1.048	1.039
0.8	1.096	1.068	1.051	1.039	1.031	1.026
0.9	1.046	1.033	1.024	1.019	1.015	1.012
1.0	1.000	1.000	1.000	1.000	1.000	1.000
1.2	0.919	0.941	0.956	0.965	0.972	0.977
1.5	0.821	0.868	0.900	0.921	0.937	0.948
2.0	0.707	0.780	0.830	0.865	0.891	0.910
2.5	0.632	0.718	0.780	0.824	0.857	0.882
3.0	0.580	0.674	0.743	0.794	0.832	0.860

Source: [36].

TABLE 8-7 Scale Factor for Random Normal Surface Deviations Δn of a Dual Reflector

Eccentricity, e	Main Reflector f/D					
	0.25	0.30	0.35	0.40	0.45	0.50
0.5	0.872	0.898	0.918	0.933	0.944	0.953
0.6	0.834	0.868	0.894	0.914	0.929	0.940
0.7	0.796	0.839	0.871	0.895	0.913	0.928
0.8	0.759	0.811	0.849	0.877	0.899	0.915
0.9	0.725	0.784	0.828	0.860	0.885	0.904
1.0	0.693	0.759	0.808	0.844	0.872	0.893
1.2	0.637	0.715	0.772	0.815	0.847	0.872
1.5	0.569	0.660	0.727	0.778	0.816	0.846
2.0	0.490	0.592	0.671	0.730	0.777	0.812
3.0	0.402	0.512	0.600	0.670	0.725	0.768

to a movement of the virtual feed but it is reduced by the factor $(M-1)/M$. Rotating the subreflector laterally offsets the virtual feed by the factor $2c\beta/(M+1)$, where $2c$ is the distance from the feed to the virtual feed and β is the rotation (radians). Since all factors are small, we can add them.

The surface tolerance of the subreflector adds another phase error loss term in the same form as the main reflector. The tolerance loss factor is less for a hyperbola than for a parabola because the angles of incidence are higher [36]. The ellipsoid of the Gregorian reflector has higher losses. We express the loss at each reflector as

$$\text{PEL} = e^{-A(4\pi\Delta\varepsilon/\lambda)^2} \tag{8-68}$$

where $\Delta\varepsilon$ is the RMS deviation to the perfect reflector surface and A is a constant depending on the eccentricity of the reflector and the main reflector f/D. This factor

is listed in Table 8-6. The main reflector has $e = 1$ (parabola). An equation matching Eq. (8-24) is

$$\text{PEL(dB)} = -685.8A\left(\frac{\Delta\varepsilon}{\lambda}\right)^2 \tag{8-69}$$

The factor $\Delta\varepsilon$ is the change in path length along a ray due to the surface deviation [Eq. (8-23)]. It is easier to specify the surface deviation in the normal direction for a subreflector. Table 8-7 gives the scale factor in terms of the normal directed surface deviation.

8-14 FEED AND SUBREFLECTOR SUPPORT STRUT RADIATION

Feed or subreflector support struts block the aperture of a centrally fed reflector and reduce gain. Because the passing waves induce currents on the struts that radiate, the effect of struts can be larger than their area. Thin struts, like thin dipoles, have a significant effective area as an antenna. The radiation from struts generates sidelobes and cross-polarization that degrade the reflector pattern. We approach strut analysis either by increasing their effective area due to the induced currents or by using physical optics to add the induced currents to the pattern, or use a ray optics method, such as GTD, to find the total pattern.

Strut losses and stray radiation should be small in a properly designed reflector, but mechanical considerations may lead to less than ideal electrical configurations. If possible, attach the struts to the outer rim of the reflector. When mounted partway out of the radius, the struts can support the feed or subreflector with thinner elements, but they block the feed (or subreflector) radiation and cast a shadow on the main reflector. Of course, when struts pass through the main reflector, they can be mounted to a smaller support frame located behind the reflector and the final design will be a compromise. Symmetrically oriented struts reduce boresight cross-polarization. For linear polarization, use four struts aligned parallel and perpendicular to the polarization vector to reduce cross-polarization at all angles. Thorough analysis considers the currents induced on the struts due to the feed radiation. These strut currents radiate and illuminate the main or subreflector along with the feed because the primary strut effect is blocking the nearly plane wave radiated by the main reflector.

Kay [37] introduced the induced field ratio (IFR) to analyze space metal frame radomes. The IFR hypothesis assumes that the currents induced on the strut are the same as those on an infinite strut illuminated by a plane wave. This reduces the problem to two dimensions. Currents induced on a thin strut increase the effective blockage area by the IFR factor. We sum these blockage areas projected on the aperture plane without regard to phase. IFR is the forward-scattered field divided by the radiation from a plane wave incident on the same area (i.e., an aperture). This reduces the complicated radiation from the strut to a simple blocked area. IFR depends on the incident polarization, and since it is a two-dimensional problem, we have only TM (E-field parallel to strut axis) or TE (H-field parallel to axis) waves.

The TM wave induces currents directed along the axis that we compute either by using a canonical problem with a closed-form solution or by using the method of moments with an incident plane. Given the surface current density $J_s(\phi)'$, a function

of the angle from the incident plane wave ϕ' on a strut cross section $\rho(\phi')$, the TM wave IFE_E can be found from an integral around the strut perimeter:

$$\text{IFR}_E = -\frac{\eta}{2wE_0}\int_{S_1} J_s(\phi')e^{-jk\rho(\phi')}\rho(\phi')\,d\phi' \tag{8-70}$$

The strut has a projected width w normal to the incident plane wave with electric field E_0 and η is the impedance of free space.

The TE wave induces currents that flow around the perimeter. We describe the current by the z-directed magnetic field on the strut. Given a unit vector $\mathbf{a}_f$ in the forward direction and the surface normal $\mathbf{n}(\phi')$, IFR_H is found from an integral around the strut:

$$\text{IFR}_H = \frac{1}{2wH_0}\int_{S_1} [\mathbf{a}_f \cdot \mathbf{n}(\phi')]H_z(\phi')e^{-jk\rho(\phi')}\rho(\phi')\,d\phi' \tag{8-71}$$

We calculate effective strut blockage area by multiplying the projected area of the strut in the aperture by IFR and use Eq. (4-111) for the loss. For a linearly polarized wave at an angle γ_i to the projected strut, we use an elliptical addition of the area:

$$\text{area} = \sum_{i=1}^{N} w_i(\text{IFR}_{Ei}\cos^2\gamma_i + \text{IFR}_{Hi}\sin^2\gamma_i) \tag{8-72}$$

We use the mean $\text{IFR}_M = (\text{IFR}_E + \text{IFR}_H)/2$ for circular polarization. The difference $\text{IFR}_D = (\text{IFR}_E - \text{IFR}_H)/2$ is a useful quantity for estimating the sidelobes and cross-polarization.

Given a reflector of radius r_0 and symmetrically placed N struts which reduces cross-polarization, we define the strut factor A_{co} as IFR times the ratio of the projected area of one strut to the reflector area:

$$A_{\text{co}} = |\text{IFR}_i|\frac{N_p w' r_0}{\pi r_0^2} \tag{8-73}$$

with $w' = w\sin\theta_0$, the width projected along the plane wave,

$$N_p = \begin{cases} 1 & \text{for } N \text{ odd} \\ 2 & \text{for } N \text{ even} \end{cases}$$

and

$$|\text{IFR}_i| = \begin{cases} |\text{IFR}_M| & \text{for circular polarization} \\ |\text{IFR}_M - \text{IFR}_D\cos 2\gamma_i| & \text{for linear polarization} \end{cases}$$

Blockage co-polarization sidelobes due to the struts radiate a broad pattern that reaches an asymptotic level [38]:

$$\text{sidelobe(dB)} = \begin{cases} 20\log(A_{\text{co}}) & \text{for uniform aperture} \\ 20\log(A_{\text{co}}) - 2 & \text{for } -20\text{-dB taper aperture} \end{cases} \tag{8-74}$$

The maximum cross-polarization within the 3-dB beamwidth can be found from the difference IFR_D. Define the cross-polarization strut factor:

$$A_{\text{xp}} = |\text{IFR}_D| \frac{N_p w' r_0}{\pi r_0^2}$$

For three struts the cross-polarization(dB) $= 20\log(A_{\text{xp}}) + 2.5$, and for four struts cross-polarization(dB) $= 20\log(A_{\text{xp}}) - 8.5$ for a uniform distribution and $20\log(A_{\text{xp}}) - 7$ with a 20-dB aperture taper [38].

Closed-form equations for IFR exist for circular struts [39] that depend on the angle between the strut axis and the incident plane wave θ_0, the strut radius a, and the propagation constant k:

$$\begin{aligned}
\text{IFR}_E &= -\frac{1}{ka\sin\theta_0} \sum_{n=0}^{\infty} \varepsilon_n \frac{J_n(ka\sin\theta_0)}{H_n^{(2)}(ka\sin\theta_0)} \\
\text{IFR}_H &= -\frac{1}{ka\sin\theta_0} \sum_{n=0}^{\infty} \varepsilon_n \frac{J_n'(ka\sin\theta_0)}{H_n^{(2)'}(ka\sin\theta_0)} \\
\varepsilon_n &= \begin{cases} 1 & \text{for } n = 0 \\ 2 & \text{for } n > 0 \end{cases}
\end{aligned} \tag{8-75}$$

where J_n is the Bessel function and $H_n^{(2)}$ is the outward-traveling Hankel function, with J_n' the derivative of J_n, and so on. Table 8-8 lists these factors for both polarizations versus strut radius.

Using only physical optics we can determine accurately the blockage effects of struts that are at least 3λ in diameter. PO excites currents only on the visible half of the struts. For smaller-diameter struts, currents creep to the far side and alter the results. No matter how thin the struts, currents will be excited on them and affect the pattern. The physical optics analysis of a dual reflector includes currents excited on the struts a number of times. Assume that the antenna is transmitting. The feed illuminates the subreflector and the struts. The current excited on the struts also radiates a field that illuminates the subreflector. If the strut blocks the path between the feed and the subreflector, PO analysis uses strut current to calculate its blockage. At this point we use the currents

TABLE 8-8 IFR_E and IFR_H for a Circular Strut

$a\sin\theta_0$	$\text{Re}(\text{IFR}_E)$	$\text{Im}(\text{IFR}_E)$	$\text{Re}(\text{IFR}_H)$	$\text{Im}(\text{IFR}_H)$
0.005	−5.148	14.088	−0.0001	−0.0198
0.010	−3.645	6.786	−0.0005	−0.050
0.020	−2.712	3.964	−0.004	−0.103
0.050	−1.982	2.003	−0.054	−0.272
0.10	−1.641	1.225	−0.292	−0.448
0.20	−1.414	0.758	−0.552	−0.374
0.50	−1.215	0.381	−0.781	−0.258
1.00	−1.145	0.255	−0.858	−0.188
2.00	−1.092	0.160	−0.914	−0.126

on the subreflector to compute additional currents excited on the struts. These add to the first set of strut currents. Radiation from the subreflector, all currents on the struts, and stray feed illumination add to illuminate the main reflector. The radiation from the main reflector current excites additional current on the subreflector and a third set of currents on the struts. We apply the far-field Green's function on the sum of all currents to calculate the pattern.

PO currents are modified on thin struts to account for creeping-wave currents. We multiply the PO strut currents by the induced current ratio (ICR) to obtain equivalent currents suitable for predictions. The factor ICR includes a strut current distribution and a complex value:

$$\mathbf{J}_s = 2\mathbf{n} \times \mathbf{H}_{\text{inc}} \cdot \text{ICR}(a, \theta_0, \phi') \tag{8-76}$$

ICR depends on the strut radius a, the incidence angle θ_0 with respect to the strut axis, and the angle around the strut ϕ' from the direction of the plane wave and the incident wave polarization. We find the incident magnetic field $\mathbf{H}_{\text{inc}}$ at the point where the plane wave touches the strut and use the current excited at this point to calculate the current in a ring around the strut. Remember that the primary effect of the strut is to block the radiation from the main reflector that approximates a plane wave in the near field where the struts are located.

We solve for ICR by considering two-dimensional scattering of the strut cross section by a plane wave. By applying moment methods to a two-dimensional scattering problem, we can solve for the current distribution on any strut cross section, but here we consider only circular struts that have a closed-form solution [40, pp. 209–219]. To simplify the problem, consider a strut lying along the z-axis. For actual analysis you will need to rotate the strut into place and rotate the incident wave into the strut coordinate system to use ICR to calculate the current distribution. In two-dimensional space the incident wave is either TM or TE with respect to the z-axis. The TM wave has its electric field in the plane containing the strut axis. A TE wave has its magnetic field in this plane. The TM case produces the following equation from the scattering of a plane wave:

$$\text{ICR}_E \hat{\mathbf{z}} = \text{ICR}_{\text{TM}}(a, \theta_0, \phi')\hat{\mathbf{z}} = \frac{\hat{\mathbf{z}} e^{-jka\cos\phi'}}{\pi ka \sin\theta_0} \sum_{m=0}^{\infty} \frac{j^m \varepsilon_m \cos m\phi'}{H_m^{(2)}(ka\sin\theta_0)} \tag{8-77}$$

The phase factor $e^{-jka\cos\phi'}$ shifts the reference plane from the strut center to the attachment point. Equation (8-77) expands the current in a cos $m\phi'$ Fourier series around the strut. ICR_E has a complex value because $H_m^{(2)}$ has a complex value. As the strut radius a increases, ICR approaches 1 at the location $\phi' = 0$. For practical purposes ICR is 1 for $a/\lambda > 1.5$.

Equation (8-77) gives the current distribution on the circular strut relative to the current excited at the initial contact point of the incident plane wave. Table 8-9 lists ICR_E evaluated when $\phi' = 0$. The constant term grows rapidly as $a \to 0$, with its imaginary part growing faster than the real part because it approximates the vector potential of a filamentary current element with its $-j$ factor between the current and the field [Eq. (2-1)]. Small struts have nearly constant current around their periphery. Equation (8-77) requires more and more terms as the strut diameter increases, and finally, a simple PO solution produces the same results. Table 8-9 lists the ICR factors versus the strut radius.

TABLE 8-9 ICR_E and ICR_H PO Current Multipliers When $\phi' = 0$ for a Circular Strut

$a \sin \theta_0$	$\text{Re}(ICR_E)$	$\text{Im}(ICR_E)$	$\text{Re}(ICR_H)$	$\text{Im}(ICR_H)$
0.002	3.660	−7.932	0.500	0.003
0.004	2.735	−4.530	0.500	0.013
0.01	2.018	−2.228	0.499	0.033
0.02	1.687	−1.353	0.500	0.071
0.05	1.376	−0.751	0.546	0.195
0.10	1.200	−0.479	0.748	0.275
0.50	1.030	−0.140	0.948	0.109
1.00	1.010	−0.076	0.982	0.068
1.50	1.005	−0.052	0.991	0.049

A TE incident wave produces similar results for scattering from a circular cylinder but has co- and cross-polarization terms:

$$ICR_H = ICR_{TE}(a, \theta_0, \phi')\hat{\boldsymbol{\phi}} = \frac{j\hat{\boldsymbol{\phi}}e^{-jka\cos\phi'}}{\pi ka \sin\theta_0} \sum_{m=0}^{\infty} \frac{j^m \varepsilon_m \cos m\phi'}{H_m^{(2)'}(ka \sin\theta_0)} \tag{8-78}$$

Equation (8-78) has the same form as Eq. (8-77) and is expanded in the even function $\cos m\phi'$ as Eq. (8-77) with coefficients using the derivative of the Hankel function.

When the incident wave approaches the strut at an angle θ_0 other than 90°, the strut scatters cross-polarization for a TE incident wave:

$$JCR_H = JCR_{TE}\hat{\mathbf{z}} = \frac{j\cos\theta_0 \hat{\mathbf{z}}e^{-jka\cos\phi'}}{(ka\sin\theta_0)^2} \sum_{m=-\infty}^{\infty} \frac{mj^m e^{jm\phi'}}{H_m^{(2)'}(ka\sin\theta_0)} \tag{8-79}$$

JFR_H is an odd function around the perimeter of the strut with a zeroth term of zero. We can expand Eq. (8-79) in terms of $\sin m\phi'$:

$$JCR_H = \frac{-2j\cos\theta_0 \hat{\mathbf{z}}e^{-jka\cos\phi'}}{(ka\sin\theta_0)^2}(I_1' \sin\phi' + j2I_2' \sin 2\phi' - 3I_3' \sin 3\phi' - \cdots)$$

$$I_m' = \frac{1}{H_m^{(2)'}(ka\sin\theta_0)} \tag{8-80}$$

Consider a plane wave incident on a straight strut at an angle θ_0 to its axis. As the wave sweeps across the strut it excites current whose phase velocity is $c/\cos\theta_0$ with respect to the strut axis. This is the same situation as a waveguide with two waves traveling back and forth between the sidewalls that produces a central phase velocity greater than c (Section 5-24). A thin strut with its constant current distribution around the circumference radiates a cone-shaped pattern peaked at an angle determined by the current phase velocity. The current is a fast or leaky wave radiator that radiates in a cone at an angle θ_0 from the axis, while the length of the strut in wavelengths determines the narrowness of the radiation beamwidth. As the diameter of the strut increases, the peripheral current distribution alters the radiation level around the cone, but the peak radiation occurs along the cone determined by the incident angle.

We use incident plane waves to derive strut blockage and scattering and to modify the formulation for spherical wave incidence. First divide the struts into coin sections. For a wave incident from a given point, we trace a ray from the point to the strut axis through a given coin section. We determine the incident magnetic field and calculate the surface current density at the point of intersection of this ray and the strut. The intersection point is $\phi' = 0$. We apply ICR_E, ICR_H, and JCR_H to calculate the currents around the coin cross section. This near-field case does not radiate a strut cone pattern because it is not a plane wave incident on the strut. As the strut diameter grows, this method leads directly to a PO formulation for strut scattering.

8-15 GAIN/NOISE TEMPERATURE OF A DUAL REFLECTOR

Collins [41] has developed a procedure for calculating the noise temperature of Cassegrain antennas pointed near the horizon. First, the diffraction pattern of the feed and subreflector combination is calculated. Some of the diffraction is added to the main-reflector spillover. At low elevation angles the antenna points about one-half of the spillover on the ground. It is a major noise temperature contribution, $\frac{1}{2}(1 - \text{SPL})T_G$, where T_G is the ground temperature and SPL is the spillover efficiency (ratio). The scattered portion of the blockage produces wide-angle sidelobes, half of which see the ground. The gain is reduced by the spillover loss, and a uniform distribution for the blockage is assumed (ATL = 1):

$$\frac{1}{2}\frac{S_b}{S_a}(\text{SPL})T_G$$

where S_b is the blocked area and S_a is the total potential aperture. The main beam points toward the sky and collects noise, $\text{SPL}\eta_b\eta_m T_S$, where η_b is the blockage efficiency, η_m the ratio of the power in the main beam and the first few sidelobes ($\eta_m \approx 0.99$), and T_s the sky temperature. We include a group of minor contributors:

$$\frac{1}{2}(1 - \text{SPL})T_s \qquad \frac{1}{2}\frac{S_b}{S_a}(\text{SPL})T_s \qquad \frac{1}{2}\text{SPL}\eta_b(1 - \eta_m)(T_G + T_s)$$

Equation (1-56) can be used when the temperature distribution is known, but the procedure of Collins gives good, although slightly conservative results. Refer to Section 1-15 to calculate the gain noise temperature of the receiving system.

8-16 DISPLACED-AXIS DUAL REFLECTOR

A displaced-axis dual reflector uses a paraboloidal main reflector with a ring focus that transforms the vertex into a ring. GO rays reflected from the main paraboloid miss the subreflector and reduce the blockage loss to a nonexcitation area instead of scattered blockage. This reflector achieves high aperture efficiency by using a subreflector that directs the higher feed radiation at the boresight to the outer rim of the parabola, where the differential area is the largest.

For the moment, consider Figure 8-14 of the Gregorian dual reflector in two dimensions. The parabola and ellipse retain their reflecting properties because we can extend them out of the page into cylindrical reflectors and use a linear array as a feed. Mentally, remove the lower half of the parabola and the upper half of the ellipse. Rays

from the left ellipse focus (feed) reflect from the remaining half of the lower ellipse to the upper half of the parabola, which transforms them into plane waves. If we fix the right ellipse focus at the focus of the parabola, we can rotate the ellipse axis about the right focus without changing the ray tracing from a feed at the left focus. We use a slightly different portion of the ellipse determined by the rays traced to the edges of the remaining half parabola.

Place a horizontal axis at the lower edge of the ellipse and rotate the ellipse axis until the feed focus is on this axis. Rotate the two-dimensional figure about this horizontal axis to form a three-dimensional reflector, and it becomes a displaced-axis dual reflector. Both the focus and vertex of the main reflector have become rings. The subreflector has a matching ring focus at the same diameter as that of the main vertex and a point focus at the feed (Figure 8-16). Now rays from the upper portion of the subreflector reflect to the upper portion of the parabola. Rays from the center of the subreflector terminate on the outer edge of the main reflector, while outer subreflector edge rays reflect to the ring vertex of the main reflector [42].

The reflector geometry has been found in closed form [43]. Given the main reflector diameter D, focal length f, diameter of the subreflector D_s, and feed half-subtended angle θ_0, the distance along the reflector axis from the vertex to the feed L_m is

$$L_m = \frac{f\,D}{D - D_s} - \frac{D_s}{4}\frac{\cos\theta_0 + 1}{\sin\theta_0} \tag{8-81}$$

We tilt the axis of the ellipse ϕ to collapse the ring focus to a point at the feed:

$$\tan\phi = \frac{2}{(\cos\theta_0 + 1)/\sin\theta_0 - 4f/(D - D_s)} \tag{8-82}$$

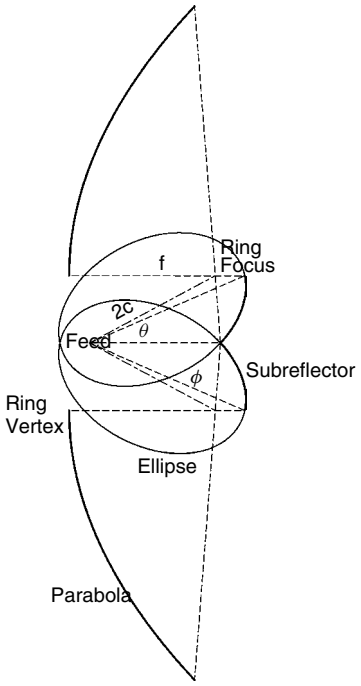


FIGURE 8-16 Displaced-axis reflector antenna.

The parameters of the ellipse are given by the equations

$$c = \frac{D_s}{4\sin\phi} \quad \text{and} \quad a = \frac{D_s}{8}\left(\frac{\cos\theta_0 + 1}{\sin\theta_0} + \frac{4f}{D - D_s}\right) \tag{8-83}$$

The half-subtended angle of the main reflector ψ_0 is found from the normal parabola with the subreflector removed and the ring focus collapsed to a point [Eq. (8-2)]:

$$\psi_0 = 2\tan^{-1}\frac{D - D_s}{4f}$$

We compute the distance between the feed and the subreflector along the axis L_s from the geometry [45]:

$$L_s = 2c\cos\phi + \frac{D_s}{2\tan\psi_0}$$

We determine aperture power distribution $A(r')$ by tracing rays from the feed to the aperture radius r' of the main reflector and by equating power in differential areas:

$$P(\theta)\sin\theta d\theta = A(r')\,dr' \tag{8-84}$$

Figure 8-17 gives the aperture distribution for a displaced-axis reflector designed for a main reflector $f/D = 0.27$ and an effective $f_{\text{eff}}/D = 1.2$ from the feed ($\theta_0 = 23.54^\circ$) for various feed edge tapers. The plot shows that increasing the feed edge taper increases the aperture power at larger radiuses but reduces the center amplitude. The center 20% of the diameter of the aperture is not excited, but this corresponds to only 4% lost area, or -0.18 dB. When calculating the amplitude taper loss [Eq. (4-8)], we

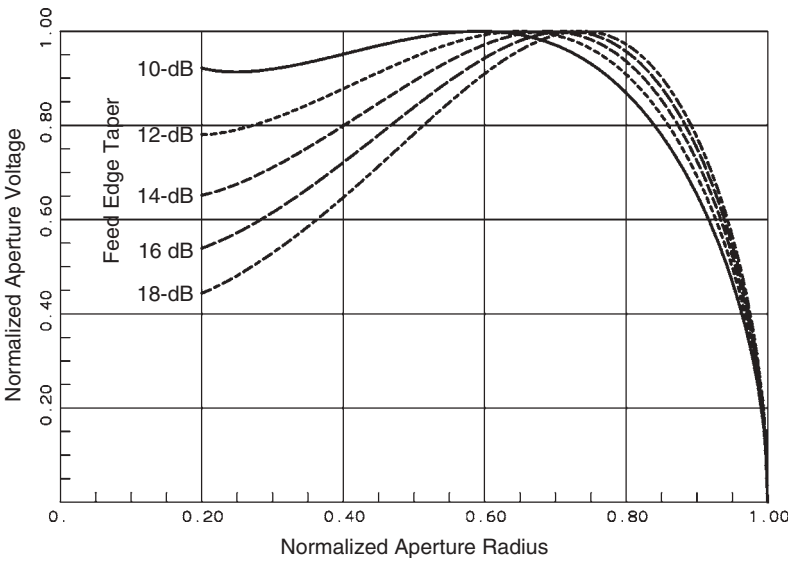


FIGURE 8-17 Aperture distribution in a displaced-axis reflector given feed edge taper for a particular antenna. (From [43], Fig. 3, © 1997 IEEE.)

TABLE 8-10 Illumination Losses of a Displaced-Axis Dual Reflector, $f/D = 0.27$, $f_{\text{eff}}/D = 1.2$ for $D_s = 0.2D$ and $0.1D$ Versus a Feed Edge Taper

	D_s/D (%)							
	20				10			
Edge Taper (dB)	SPL (dB)	ATL (dB)	Total (dB)	Sidelobe	SPL (dB)	ATL (dB)	Total (dB)	Sidelobe
10	0.434	0.476	0.910	15.2	0.434	0.411	0.845	18.3
11	0.341	0.455	0.796	14.9	0.341	0.377	0.718	17.8
12	0.268	0.442	0.710	14.5	0.268	0.350	0.619	17.3
13	0.212	0.434	0.646	14.2	0.212	0.330	0.542	16.8
14	0.167	0.432	0.600	14.0	0.167	0.316	0.483	16.4
15	0.132	0.436	0.568	13.7	0.132	0.308	0.440	16.0
16	0.105	0.444	0.548	13.5	0.105	0.304	0.409	15.6
17	0.083	0.456	0.539	13.3	0.083	0.306	0.388	15.3
18	0.066	0.472	0.537	13.1	0.066	0.311	0.377	14.9
19	0.052	0.491	0.543	12.9	0.052	0.320	0.372	14.6
20	0.041	0.513	0.555	12.7	0.041	0.333	0.374	14.3

use the full radius $a = D/2$, which accounts for the lost aperture center area. Table 8-10 lists the illumination losses of this reflector. The antenna has 88.2% aperture efficiency, including blockage loss for $D_s/D = 0.2$ and 91.8% for $D_s/D = 0.1$. These numbers do not include diffraction loss due to the subreflector and main reflector size in wavelengths or strut blockage. Similar to a Cassegrain reflector, increasing the subreflector diameter beyond the GO design by 1 to 2λ decreases the diffraction loss.

Figure 8-17 shows the complete taper of the aperture to zero voltage at the edge. We can increase the aperture efficiency slightly by designing the antenna with an effective main reflector diameter slightly larger than the real diameter and produce a finite aperture edge taper at the cost of increased spillover past the main reflector. Table 8-11 lists the illumination losses for designs the same as Table 8-10 except that the effective main reflector is 2% greater.

Four versions of displaced-axis reflectors have been derived from Gregorian and Cassegrain antennas [44]. One other case, the double-offset Cassegrain, crosses the feed illumination so that boresight feed amplitude reflects to the outer rim of the main reflector. This antenna, similar to the case covered above, has a high aperture efficiency, whereas the other two cases have modest aperture efficiencies. Equations to specify all four antennas are available [45]. The normal displaced-axis dual reflector has less sensitivity to feed axial defocusing than does a normal Cassegrain or Gregorian antenna, but it is more sensitive to lateral offset of the feed [46].

8-17 OFFSET-FED DUAL REFLECTOR

When we offset-feed a dual reflector, we can eliminate subreflector central blockage of the Cassegrain or Gregorian reflectors. This design adds parameters to give more convenient packaging that fits in the available space, such as on a spacecraft. More important, by rotating the subreflector axis relative to the main reflector axis, we can

TABLE 8-11 Illumination Losses of a Displaced-Axis Dual Reflector, $f/D = 0.27$, $f_{\text{eff}}/D = 1.2$ for $D_s = 0.2D$, and $0.1D$ Versus a Feed Edge Taper (Effective Main *Diameter* = 102% Actual)

	D_s/D (%)							
	20				10			
Edge Taper (dB)	SPL (dB)	ATL (dB)	Total (dB)	Aperture Taper (dB)	SPL (dB)	ATL (dB)	Total (dB)	Aperture Taper (dB)
10	0.455	0.371	0.816	10.6	0.443	0.304	0.746	13.5
11	0.352	0.352	0.705	10.4	0.350	0.271	0.621	12.5
12	0.280	0.340	0.621	10.1	0.278	0.246	0.524	11.5
13	0.225	0.335	0.559	9.9	0.222	0.227	0.449	10.6
14	0.181	0.334	0.515	9.6	0.178	0.215	0.393	10.4
15	0.147	0.339	0.486	9.5	0.144	0.208	0.352	10.2
16	0.120	0.349	0.469	9.3	0.117	0.206	0.323	10.0
17	0.099	0.363	0.462	9.1	0.096	0.209	0.305	9.8
18	0.083	0.380	0.463	9.0	0.079	0.216	0.295	9.7
19	0.070	0.401	0.471	8.8	0.066	0.227	0.293	9.5
20	0.060	0.425	0.485	8.7	0.056	0.241	0.297	9.4

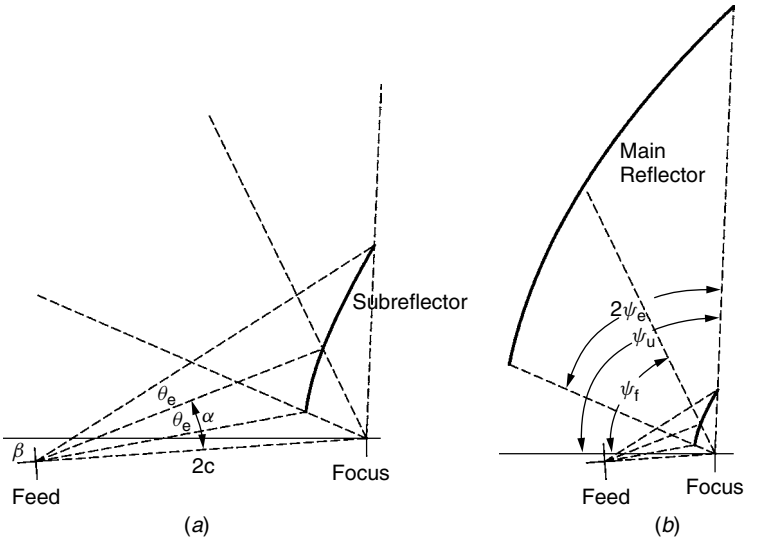


FIGURE 8-18 Dual offset-fed Cassegrain reflector, including Mizugutch feed axis tilt: (*a*) feed and subreflector geometry; (*b*) dual reflector.

greatly reduce cross-polarization or beam squint of dual circularly polarized feeds in the offset reflector.

Figure 8-18 illustrates the geometry of an offset-fed Cassegrain reflector, and Figure 8-19 shows the offset-fed Gregorian geometry. Refer to Figure 8-7 for the parameters of the offset main reflector. We point the feed at the subreflector center to reduce spillover and to equalize the amplitude distribution in the aperture of the

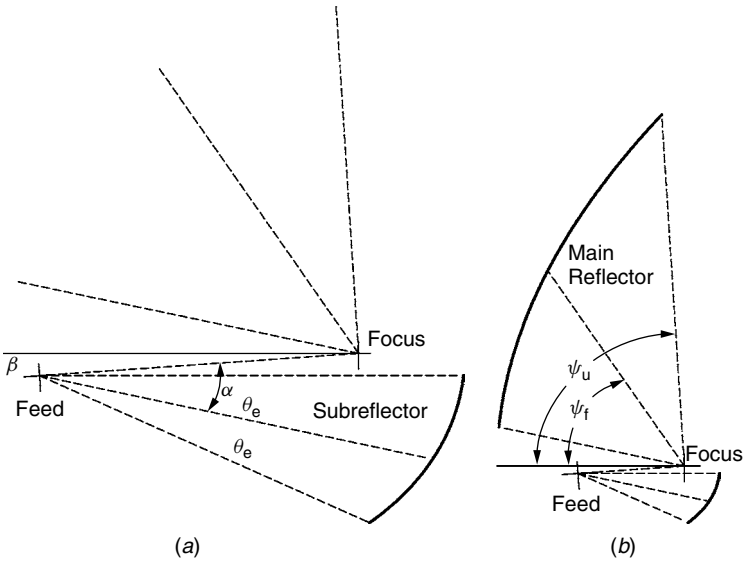


FIGURE 8-19 Dual offset-fed Gregorian reflector, including Mizugutch feed axis tilt: (*a*) feed and subreflector geometry; (*b*) dual reflector.

main reflector. Similar to a displaced-axis dual reflector, we tilt the axis between the focuses of the subreflector relative to the main reflector axis β (α in Ticra [48, App. B]). The small amount of tilt to the subreflector axis converts the equivalent parabola of the dual reflector to an axisymmetric geometry [47].

It takes five parameters to specify the antenna if the Mizugutch angle requirement is applied to the feed tilt angle α (ψ_0 Ticra) relative to the subreflector axis determined from the magnification M given the subreflector eccentricity e [Eq. (8-64)]:

$$M \tan\frac{\beta}{2} = \tan\frac{\alpha}{2} \quad \text{or} \quad \left(M \tan\frac{\alpha}{2} = \tan\frac{\psi_0}{2}\right)^{\dagger} \qquad \text{for } M = \frac{e+1}{e-1} \tag{8-85}$$

The equation for M in Eq. (8-85) is used for both Cassegrain and Gregorian reflectors. We start the design with the diameter of the main reflector D because it determines the gain and beamwidth. Ticra [48, App. B] uses main reflector focal length f, the half distance between focuses of subreflector c, subreflector eccentricity e, and axis tilt β. Granet [49,50] supplies equations to calculate the reflector dimensions for 17 different sets of five input parameters. These sets of equations allow the direct application of various mechanical constraints to the design or electrical constraints, such as subreflector size to limit diffraction loss. All of Granet's sets apply the Mizugutch relationship, because this small change should be applied to all designs.

By tracing rays through the reflectors, the center offset H is found:

$$H = -2f\frac{\tan(\beta/2) - M\tan(\alpha/2)}{1 + M\tan(\beta/2)\tan(\alpha/2)} \tag{8-86}$$

Given H, D, and f, we compute main reflector parameters from Eqs. (8-35) to (8-47). We compute the half subtended angle of feed θ_e by tracing the ray to the upper rim of

the main reflector with feed angles ψ_U and α:

$$\psi_U = -2\tan^{-1}\frac{2H+D}{4f} \quad \text{and} \quad \theta_e = \left|2\tan^{-1}\left(\frac{1}{M}\tan\frac{\psi_U-\beta}{2}\right)-\alpha\right| \tag{8-87}$$

The feed subtended angle of the subreflector is $2\theta_e$. We calculate the rim ellipse of the subreflector determined by the cone axis with angle α and cone angle θ_e using Eq. (8-59) and (8-60).

8-18 HORN REFLECTOR AND DRAGONIAN DUAL REFLECTOR

The horn reflector shown in Figure 8-20 consists of a pyramidal or conical input section excited with a rectangular or circular waveguide mode, respectively, that feeds an offset paraboloidal reflector. The beam exits horizontally. The horn reflector geometry is an offset reflector with offset angle $\psi_0 = 90^\circ$ and center offset $H = 2f$ the same as the periscope configuration. Figure 8-21 gives the pattern of a 3-m-diameter reflector operating at 6 GHz (diameter $= 60\lambda$) with $f = 3.215$ m ($\psi_e = 15^\circ$). The antenna radiates cross-polarization -23 dB relative to the beam peak in the horizontal plane. It cannot be used for two channels with different polarizations, because similar to all offset-fed reflectors, circularly polarized beams squint right and left in this horizontal plane. The antenna radiates a significant sidelobe 90° from the boresight in the horizontal plane that can be controlled using serrated-edge blinders [51].

A Dragonian dual reflector uses a hyperbola subreflector that curves toward the main reflector in a Cassegrain system. This produces a dual reflector with magnification

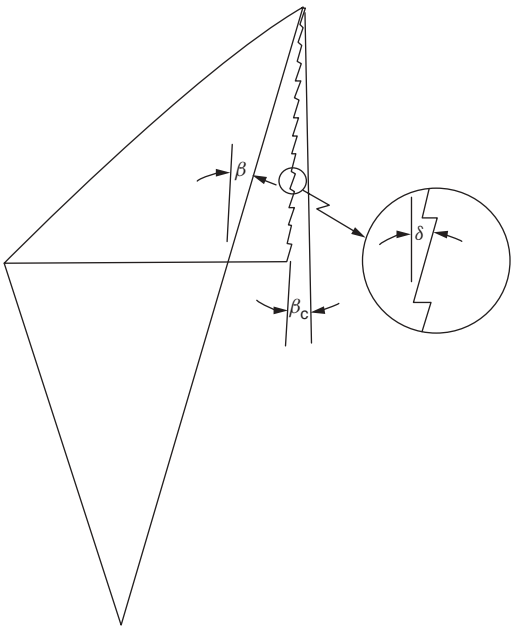


FIGURE 8-20 Horn reflector with serrated side blinders. (From [51], Fig. 3, © 1973 IEEE.)

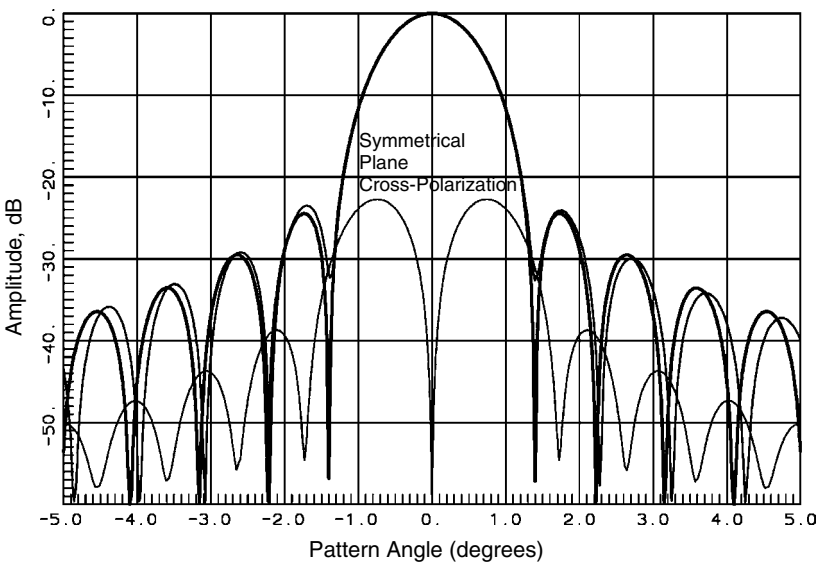


FIGURE 8-21 Pattern of a 3-m horn reflector at 6 GHz.

$M < 1$ and we need a feed antenna with a wider beamwidth than would be required for efficient feed of the main reflector. We use long focal lengths for the main reflector, which flattens its curve, so that the feed beamwidth can remain small. Jones and Kelleher [52] applied this Cassegrain arrangement to a horn reflector and located the feed horn in the middle of the paraboloidal main reflector. Dragone [53] derived the generalized Mizugutch criterion for multiple reflectors and showed that it could be applied to this Cassegrain system to eliminate cross-polarization. Figure 8-22 shows a Dragonian dual reflector fed by a corrugated horn feed located beyond the rim of the paraboloidal reflector to replace the 3-m horn reflector. Pattern analysis produces the same curves as in Figure 8-21 except that cross-polarization in the horizontal plane is eliminated.

The design given in Figure 8-22 has $f = 9.8$ m, $D = 3$ m, $\theta_e = 20^\circ$, and tilts the subreflector axis by -73° to place the main reflector and subreflector in different quadrants. The Mizugutch criterion between the subreflector axis tilt and the feed tilt locates the feed axis at -97.5° relative to the main reflector axis and -24.5° relative to the subreflector axis. The parameters of the reflector were adjusted so that the plane wave radiated from the main reflector misses both the feed corrugated horn and the subreflector. All dimensions can be found using available equation sets [54] that require various sets of five inputs to totally specify the dual reflector. By using these equations, we discover that the hyperboloidal subreflector rim is an ellipse with 2.05- and 2.60-m diameters. To use the hyperbola close to the feed, specify a negative eccentricity and the equations curve the hyperbola toward the main reflector. Similarly, the equation for magnification produces a value of less than 1 for a negative eccentricity. Given the $e = -1.832$ for the reflector of Figure 8-22,

$$M = \frac{e+1}{e-1} = \frac{-1.832+1}{-1.832-1} = 0.2938$$

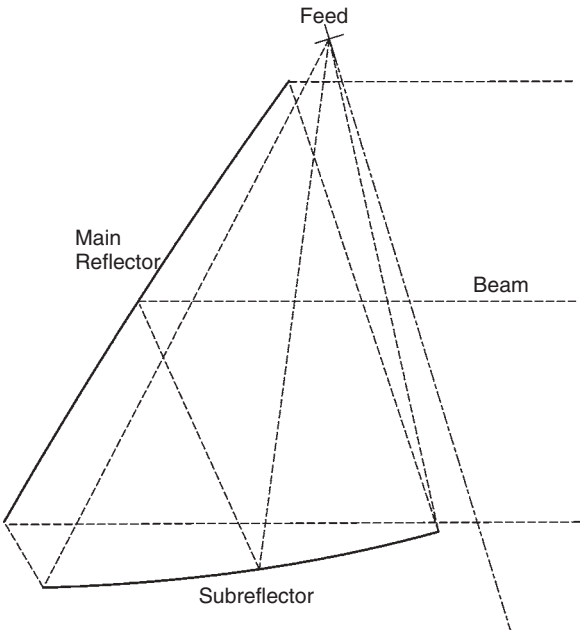


FIGURE 8-22 Dragonian dual-reflector geometry.

We compute feed tilt relative to the subreflector axis to satisfy the Mizugutch criterion:

$$\alpha = 2\tan^{-1}\left(M\tan\frac{\beta}{2}\right) = 2\tan^{-1}\left(0.2938\tan\frac{-73}{2}\right) = -24.5^{\circ}$$

8-19 SPHERICAL REFLECTOR

When we feed-scan a paraboloidal reflector, the pattern sidelobes develop coma and the beam shape generally degrades. Feed scanning is limited. In a spherical reflector a feed moved in an arc from the center of the sphere and sees the same reflector geometry if we discount the edge effects. Greater scanning is possible, but the spherical reflector fails to focus an incident plane wave to a point and requires more elaborate feeds.

We can design many types of feeds for the spherical reflector. The reflector can be fed from a point source for large f/D by assuming that it is a distorted parabola [55,56]. It can be fed with a line source to follow the axis fields. Corrector subreflectors can be designed to correct the spherical aberrations [58]. Like the parabolic reflector, we can design arrays [24] to compensate for spherical aberrations and give multiple beams.

Figure 8-23 shows the geometry and ray tracing of a spherical reflector illuminated by a plane wave. All rays intersect a radial line of the sphere (the axis) in the direction of the incident wave because the reflector has circular symmetry about all axes. The diagram traces rays hitting the outer portion of the reflector as passing through the axis closer to the vertex than do the rays reflected from areas closer to the axis. The reflector has a line focus. A distorted paraboloidal reflector with a line focus exhibits spherical aberration because the focal length depends on the radial distance from the axis of the reflection point. The spherical reflector has a cusplike caustic where GO

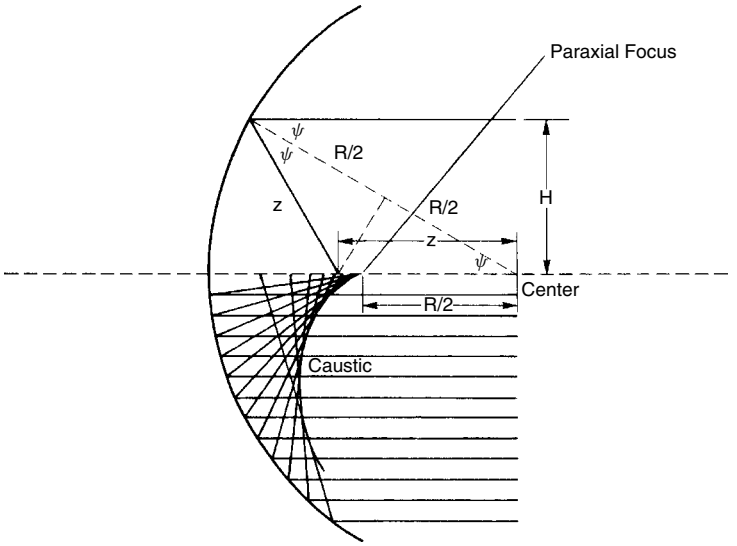


FIGURE 8-23 Ray tracing in a spherical reflector.

predicts infinite fields. The second side of Figure 8-23 traces a single ray. We can easily solve the isosceles triangle for the results:

$$z = \frac{R/2}{\sqrt{1 - H^2/R^2}} \tag{8-88}$$

$$H^2 = R^2\left(1 - \frac{R^2}{4z^2}\right) \tag{8-89}$$

where z is the location of the focus for a given ray. As H approaches zero, with rays near the axis, the reflected ray passes through the paraxial focus ($z = R/2$).

We use Eq. (8-89) to find the power distribution on the axis by using the conservation of power. The power in a differential area of the plane wave reflects into a differential length on the axis: $dA = 2\pi H\, dH$. We differentiate Eq. (8-89) implicitly:

$$2H\, dH = \frac{R^2}{2z^3}\, dz$$

The power distribution along the axis is

$$P_z = \frac{P_0 R^3}{8z^3} \tag{8-90}$$

where P_0 is the power at the paraxial focus. The peak power occurs at the paraxial focus and drops by one-eighth (−9 dB) at the vertex. We determine the required length of the line source feed from the rotation angle ψ of the illuminated portion of the reflector:

$$\text{feed length} = \frac{R(1/\cos\psi - 1)}{2} \tag{8-91}$$

Example If the half-rotation angle of the illuminated region is 30°, the feed length is $0.0774R$ from Eq. (8-91). The amplitude decreases by

$$[\text{Eq. (8-90)}] \quad \frac{P_z}{P_0} = \frac{R^3}{8(R/2 + 0.0774R)^3} = 0.65 \quad (-1.9\,\text{dB})$$

The rays intersecting the axis are not at constant phase. The path length from the aperture plane through the reflector origin is

$$\text{path length} = \frac{R^2}{2z} + z \tag{8-92}$$

We can approximate Eq. (8-92) by a linear function if the feed length is short.

Example The feed length is $0.0774R$ long; calculate the phase change required along the feed.

The feed starts at the paraxial focus ($z = R/2$):

$$\text{path length} = \frac{R^2}{R} + \frac{R}{2} = 1.5R$$

At $z = R/2 + 0.0774R$, the path length $= 1.443R$. The phase change is $(2\pi/\lambda)R(0.0566)$. If we plot the phase change over the region of the feed, we can approximate the phase change by a linear function very accurately.

The spherical reflector can be fed from a point source when the f/D is large [55]. The center of the reflector approximates a parabola. The optimum focal point is

$$f = \frac{1}{4}\left[R + \sqrt{R^2 - (D/2)^2}\right] \tag{8-93}$$

The maximum phase path length error is [56]

$$\frac{\Delta L}{\lambda} = \frac{1}{2048}\frac{D}{\lambda}\frac{1}{(f/D)^3} \tag{8-94}$$

The approximate gain loss is

$$\frac{\Delta G}{G} = 3.5092\left(\frac{\Delta L}{\lambda}\right)^2$$

or

$$\text{PEL(dB)} = 10\log\left[1 - 3.5092\left(\frac{\Delta L}{\lambda}\right)^2\right] \tag{8-95}$$

A path length deviation of 0.25λ reduces the gain by 1.08 dB.

Example Determine the f/D value of a spherical reflector to limit ΔL to $1/16\lambda$ for a reflector diameter of 50λ.

By rearranging Eq. (8-94), we find that

$$\frac{f}{D} = \frac{16(50)}{2048} = 0.73$$

8-20 SHAPED REFLECTORS

Shaped reflectors spread cylindrical or spherical waves into a desired pattern that depends on geometric optics. Shaped reflectors do not radiate patterns exactly as prescribed by GO. In all cases we must apply techniques such as aperture diffraction, induced currents, or geometric theory of diffraction (GTD) to compute the actual pattern. We consider only the first-order GO for design, although analysis requires more elaborate techniques.

We use two principles to design shaped reflectors. The first is GO reflection expressed as a differential equation. The second is the conservation of power in ray tubes, which can be expressed either in terms of differential areas or integrals over sections of the feed and reflection patterns. We define two angles for the GO reflection equation. The feed points toward the reflector, and we measure its pattern angle ψ from an axis pointed toward the reflector. The reflector reradiates the incident feed pattern in a far-field pattern whose angle θ is measured from the axis pointing away from the reflector. The differential equation of reflection is [2]

$$\tan\frac{\theta + \psi}{2} = \frac{d\rho}{\rho d\psi} \tag{8-96}$$

where ρ is the distance of the reflector from the feed. The edges of the reflector are defined by angles ψ_1 and ψ_2 measured from the feed axis and reflect in directions θ_1 and θ_2. We integrate this differential equation for a solution:

$$\ln\frac{\rho(\psi)}{\rho_0(\psi_1)} = \int_{\psi_1}^{\psi} \tan\frac{\theta(\psi) + \psi}{2}\, d\psi \tag{8-97}$$

where ψ_1 is some initial angle of the feed and $\rho(\psi_1)$ is the initial radius vector locating the reflector at ψ_1.

GO, the zero-wavelength approximation, is consistent at any size. All parabolic reflectors collimate spherical waves radiated from the focus regardless of size. Only by considering diffraction or currents induced on the reflector can we compute gain and beamwidth of the antenna.

Example From Eq. (8-97), determine the reflector surface to give $\theta(\psi) = 0$.

$$\ln\frac{\rho(\psi)}{\rho_0(\psi_1)} = \int_{\psi_1}^{\psi} \tan\frac{\psi}{2}\, d\psi = -2\left(\ln\cos\frac{\psi}{2} - \ln\cos\frac{\psi_1}{2}\right)$$

By the properties of the ln function, this becomes

$$\ln\frac{\rho(\psi)}{\rho_0(\psi_1)} = \ln\frac{\cos^2(\psi_1/2)}{\cos^2(\psi/2)}$$

By taking the exponential of each side, we get the polar equation of the reflector:

$$\rho(\psi) = \rho_0(\psi_1)\frac{\cos^2(\psi_1/2)}{\cos^2(\psi/2)}$$

We let $\psi_1 = 0$ and set $\rho(\psi) = f$, and we obtain the equation for the parabola [Eq. (8-1)].

The differential equation of reflection tells us only the shape of the reflector locally to produce a reflection in a direction θ for an incident angle ψ. We must still find the power density in various directions. The ratio of differential areas gives us these power densities. Given the pattern of the feed $G_f(\psi, \phi)$ and the pattern of the reflection $P(\theta, \phi)$,

$$KP(\theta, \phi)\, dA(\theta, \phi) = G_f(\psi, \phi)\, dA_f(\psi, \phi) \tag{8-98}$$

where $dA(\theta, \phi)$ is the differential area of the reflection pattern, $dA_f(\psi, \phi)$ the differential area of the feed pattern, and K a constant found by equating the total incident and reflected powers. Equation (8-98) is based on the assumption that reflections are 1 : 1 with the feed pattern.

8-20.1 Cylindrical Reflector Synthesis

We feed cylindrical reflectors with line sources. The reflector determines the beam shape in one plane and the line-source distribution in the other. The problem reduces to designing a single two-dimensional curve moved along an axis to define the reflector. The power radiated by the feed is given by $G_f(\psi)\, d\psi$. The reflector directs this power at an angle θ whose power density is $P(\theta)\, d\theta$. These are proportional [Eq. (8-98)]:

$$K\, P(\theta)\, d\theta = G_f(\psi)\, d\psi \tag{8-99}$$

At the limits of the reflector, feed angles ψ_1 and ψ_2 reflect to angles θ_1 and θ_2. We calculate the constant K by equating the power in each pattern:

$$K = \frac{\displaystyle\int_{\psi_1}^{\psi_2} G_f(\psi)\, d\psi}{\displaystyle\int_{\theta_1}^{\theta_2} P(\theta)\, d\theta} \tag{8-100}$$

We integrate Eq. (8-99) to derive a formal solution usually arrived at numerically. By combining Eqs. (8-99) and (8-100), we eliminate K:

$$\frac{\displaystyle\int_{\theta_1}^{\theta} P(\theta)\, d\theta}{\displaystyle\int_{\theta_1}^{\theta_2} P(\theta)\, d\theta} = \frac{\displaystyle\int_{\psi_1}^{\psi} G_f(\psi)\, d\psi}{\displaystyle\int_{\psi_1}^{\psi_2} G_f(\psi)\, d\psi} \tag{8-101}$$

We use Eq. (8-101) with a known feed pattern $G_f(\psi)$ and a desired pattern function $P(\theta)$ to establish the relation $\theta(\psi)$. We insert $\theta(\psi)$ into the differential equation for

the reflection [Eq. (8-97)] and determine radial distance as a function of ψ:

$$\rho(\psi) = \rho_0(\psi_1) \exp\left[\int_{\psi_1}^{\psi} \tan\frac{\theta(\psi)+\psi}{2}\, d\psi\right] \tag{8-102}$$

With Eq. (8-102) we calculate the reflector coordinates to within a scale factor $\rho_0(\psi_1)$.

8-20.2 Circularly Symmetrical Reflector Synthesis

The synthesis of circularly symmetrical reflectors is readily reduced to a two-dimensional problem. We must assume that the feed pattern is also axisymmetrical. We describe the feed pattern by $G_f(\psi)$ and the reflector pattern by $P(\theta)$. The differential areas are $\sin\psi\, d\psi$ and $\sin\theta\, d\theta$. Equation (8-98) becomes

$$KP(\theta)\sin\theta\, d\theta = G_f(\psi)\sin\psi\, d\psi \tag{8-103}$$

We integrate Eq. (8-103) to find the function $\theta(\psi)$:

$$\frac{\displaystyle\int_{\theta_1}^{\theta} P(\theta)\sin\theta\, d\theta}{\displaystyle\int_{\theta_1}^{\theta_2} P(\theta)\sin\theta\, d\theta} = \frac{\displaystyle\int_{\psi_1}^{\psi} G_f(\psi)\sin\psi\, d\psi}{\displaystyle\int_{\psi_1}^{\psi_2} G_f(\psi)\sin\psi\, d\psi} \tag{8-104}$$

We use Eq. (8-102) from the reflection differential equation with $\theta(\psi)$ to determine the polar equation of the reflector. The design of a shaped reflector can best be explained with an example. The cylindrical reflector synthesis follows parallel steps.

Example Design a reflector to transform the feed pattern of Figure 8-24*a* into the pattern of Figure 8-24*b*.

The required pattern drops by about 9 dB from 50° to 0°. We will use the feed pattern from 4° to 54° and design a reflector with a ring caustic. The feed pattern at 4° is reflected to 50°, and the feed pattern at 54° is reflected to 0°. The geometric optics rays cross somewhere in front of the reflector. We have the following limits:

Feed: $\psi_1 = 4^\circ$, $\psi_2 = 54^\circ$
Reflection: $\theta_1 = 50^\circ$, $\theta_2 = 0$

We insert the patterns of Figure 8-24 into both sides of Eq. (8-104) and evaluate the ratio of the integrals. Table 8-12 gives the results of these integrals for θ and ψ as normalized in Eq. (8-104). Given ψ, we find θ by equating integrals in the table. For example, follow the line from $\psi = 28^\circ$ (feed) to its integral value, match the values of the integrals, and determine the reflection angle $\theta = 42^\circ$. We trace a number of these through interpolation to generate Table 8-13 of reflection angles $\theta(\psi)$ for given feed angles.

We use Table 8-13 of $\theta(\psi)$ in the integral of Eq. (8-102) to calculate the normalized polar equation of the reflector listed in Table 8-14. Figure 8-25 shows the reflector shape as well as the geometry of axisymmetric-shaped reflectors. There is a hole in the center because the procedure fails to specify that region. Note in Table 8-12 how

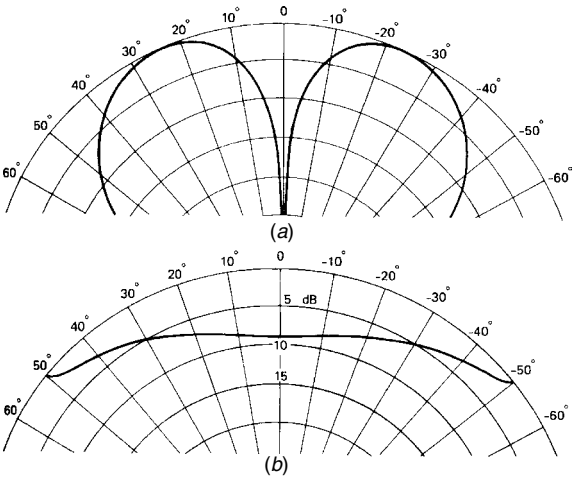


FIGURE 8-24 Axisymmetrical reflector pattern transformation: (*a*) feed pattern; (*b*) reflector pattern.

TABLE 8-12 Use of Normalized Integrals of Feed and Pattern Power to Determine $\theta(\psi)$ in Shaped Reflector Synthesis

Feed Angle, ψ (deg)	Feed Normalized Integral	Reflection Normalized Integral	Reflection Angle, θ (deg)
4	0.000	0.000	50
6	0.002	0.143	48
8	0.006	0.259	46
10	0.016	0.360	44
12	0.031	0.449	42
14	0.054	0.527	40
16	0.086	0.596	38
18	0.128	0.657	36
20	0.178	0.711	34
22	0.237	0.757	32
24	0.302	0.798	30
26	0.373	0.833	28
28	0.445	0.864	26
30	0.518	0.889	24
32	0.590	0.911	22
34	0.658	0.930	20
36	0.720	0.946	18
38	0.777	0.959	16
40	0.827	0.970	14
42	0.869	0.978	12
44	0.905	0.985	10
46	0.934	0.991	8
48	0.958	0.995	6
50	0.976	0.998	4
52	0.999	0.999	2
54	1.000	1.000	0

TABLE 8-13 Shaped Reflector Synthesis Reflection Angles for Given Feed Angles, $\theta(\psi)$

Feed Angle, ψ (deg)	Reflection Angle, θ (deg)	Feed Angle, ψ (deg)	Reflection Angle, θ (deg)
4	50.0	30	40.2
6	50.0	32	38.2
8	49.9	34	36.0
10	49.8	36	33.6
12	49.6	38	31.1
14	49.3	40	28.4
16	48.8	42	25.6
18	48.2	44	22.6
20	47.4	46	19.5
22	46.4	48	16.2
24	45.2	50	12.6
26	43.7	52	8.4
28	42.1	54	0.0

TABLE 8-14 Shaped Reflector Synthesis Normalized Polar Equation of a Reflector, $\rho(\psi)$

Feed Angle, ψ (deg)	Normalized Radius, ρ	Feed Angle, ψ (deg)	Normalized Radius, ρ	Feed Angle, ψ (deg)	Normalized Radius, ρ
4	1.000	22	1.208	40	1.503
6	1.018	24	1.238	42	1.539
8	1.038	26	1.268	44	1.575
10	1.058	28	1.299	46	1.611
12	1.080	30	1.332	48	1.648
14	1.103	32	1.365	50	1.684
16	1.128	34	1.398	52	1.719
18	1.153	36	1.433	54	1.753
20	1.180	38	1.468		

much of the inner portion of the reflector must reflect rays near 50° to achieve the high gain required. We could repeat the example and design without a ring caustic where rays from the feed near 54° reflect to 50° and would produce a flatter reflector.

Because diffraction effects spread the pattern, we could approximate the pattern of Figure 8-24*b* by designing a reflector to point the beam in a cone about the axis. If we take Eq. (8-97) and let $\theta(\psi) = \theta_0$, a constant, we get the surface

$$\rho(\psi) = \rho_0 \frac{\cos^2[(\psi_0 + \theta_0)/2]}{\cos^2[(\psi + \theta_0)/2]} \tag{8-105}$$

Example Estimate the directivity of a 40λ-diameter reflector shaped by Eq. (8-105) to scan in a cone to $\theta_0 = 50°$.

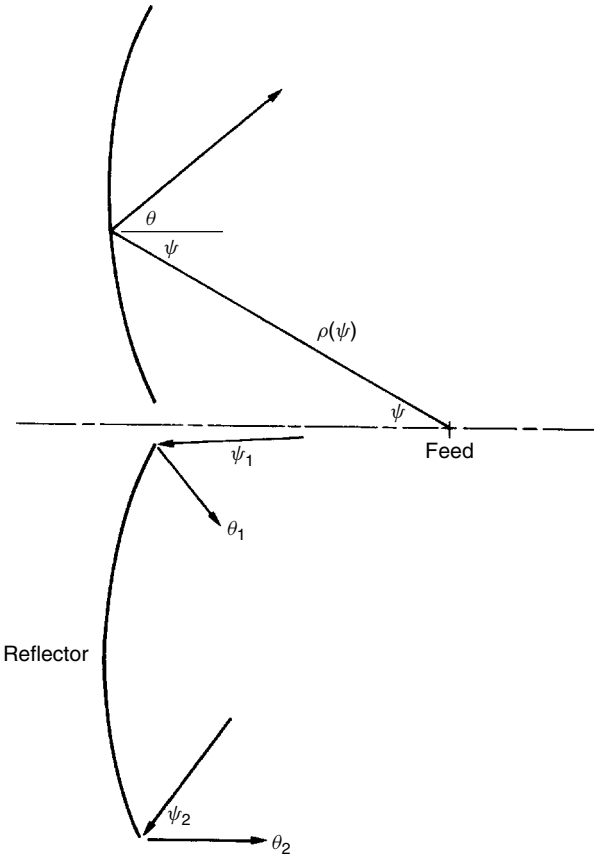


FIGURE 8-25 Circularly symmetrical reflector designed with a caustic reflector.

Only half of the diameter is used for each side. The effective scanned aperture width becomes $(40\lambda/2)\cos 50^\circ = 12.8\lambda$. If we assume a uniform-amplitude aperture distribution, we obtain an upper bound. From Eq. (4-83), HPBW $= 59^\circ/12.8 = 4.6^\circ$. We use Eq. (1-24) to estimate the directivity:

$$\text{directivity} = \frac{2}{\cos(50^\circ - 2.3^\circ) - \cos(50^\circ + 2.3^\circ)} = 32.5 \quad (15\,\text{dB})$$

The boresight gain of the aperture with a uniform distribution is 42 dB. Spreading the reflection into a cone greatly reduces gain. The shaped reflector above will have even less directivity because it has a greater edge taper.

8-20.3 Doubly Curved Reflector for Shaped Beams

It is a common radar requirement to have a narrow beam in one plane and a shaped beam in the other. Such beams can be obtained from shaped cylindrical reflectors, but it is simpler to replace the line source with a single feed. We only specify the pattern in the principal planes denoted: θ_V, the shaped pattern coordinate, and θ_H, the pencil

beam [2,59]. Similarly, we specify the feed antenna pattern in terms of ψ_V and ψ_H. For a given feed angle ψ_V, the reflected wave angle is θ_V. The only θ_H value allowed is zero. The reflector collimates the wave in the horizontal plane. This collimation requires a symmetrical reflector made from parabolic curves in the horizontal plane. We design the vertical curve only through the center of the reflector.

For a given feed angle ψ_V (Figure 8-26*a*), all incoming rays at an angle θ_V must be reflected into the feed. The incoming rays form the $x-z'$ plane in Figure 8-26, and the reflector collimates these to the feed by a parabola in the plane. We call this parabola a *rib* of the reflector. Figure 8-26*b* shows the plane and two rays reflecting into the feed from a wave arriving at an angle θ_V in the $x-z'$ plane. For the beam to focus, the optical path lengths must be equal:

$$\overline{BP} + \overline{PO} = \overline{AN} + \overline{NO} \tag{8-106}$$

Equation (8-106) establishes the curve of the rib in the $x-z'$ plane as a parabola with focal length

$$f = \rho_c(\psi_V)\cos^2\frac{\theta_V(\psi_V) + \psi_V}{2} \tag{8-107}$$

with the focus located on the z'-axis. Using the parabolic ribs reduces the problem to the design of the central curve $\rho_c(\psi_V)$.

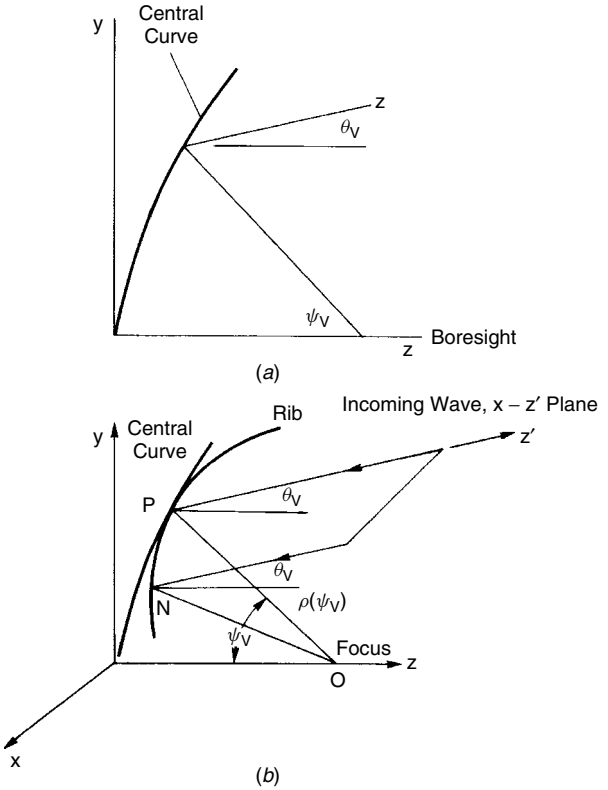


FIGURE 8-26 Doubly curved shaped reflector.

The reflected and feed power densities modify Eq. (8-98) to

$$K\,P(\theta_V)d\theta_V\,\rho_c(\psi_V)\,d\psi_H = G_f(\psi_V)\,d\psi_V\,d\psi_H \tag{8-108}$$

We integrate Eq. (8-108) and normalize to the total power:

$$\frac{\displaystyle\int_{\theta_1}^{\theta_V} P(\theta_V)\,d\theta_V}{\displaystyle\int_{\theta_1}^{\theta_2} P(\theta_V)\,d\theta_V} = \frac{\displaystyle\int_{\psi_1}^{\psi_V} [G_f(\psi_V)/\rho_c(\psi_V)]\,d\psi_V}{\displaystyle\int_{\psi_1}^{\psi_2} [G_f(\psi_V)/\rho_c(\psi_V)]\,d\psi_V} \tag{8-109}$$

Equation (8-109) is similar to Eqs. (8-101) and (8-104) except that the feed pattern integral value depends on the radial distance to the central rib. We must know $\rho_c(\psi_V)$ before we can determine $\theta_V(\psi_V)$, which will be required to compute $\rho_c(\psi_V)$ from the reflection differential equation [Eq. (8-102)]. The solution can be found only by an iterative process.

We must assume a $\rho_c(\psi_V)$, solve for $\theta_V(\psi_V)$, and use the result to compute a new $\rho_c(\psi_V)$. After a few iterations, the values of $\rho_c(\psi_V)$ converge. We use the normalized ρ_c with the foregoing ratio of integrals. We start with a parabola:

$$\frac{\rho_c(\psi_V)}{\rho_c(\psi_1)} = \frac{\cos^2(\psi_1/2)}{\cos^2(\psi_V/2)}$$

The surface generated by following the method may not be defined uniquely. We pick a constant width for the reflector in the horizontal plane. We define the surface with a continuous series of parabolas each in a x–z' plane determined by the reflection angle θ_V, which changes direction along the central rib. We must plot the curve of the vertical coordinate of the edge versus ψ_V to see if it is monotonic. If there are loops in the curve, the surface defined is not unique.

Given the width x, we calculate the vertical coordinate of the edge by the following development. The location of the rib on the central curve is given by $\rho_c(\psi_V)\sin\psi_V$. The rib is a parabola in the x–z' plane with its focus given by Eq. (8-107). The z'-coordinate at the edge is $z' = x^2/4f(\psi_V)$. We determine the vertical dimension by projecting this point onto the y-axis: $y = \rho_c(\psi_V)\sin\psi_V$. Elliott [60, p. 500] points out that by following this method, one does not get the proper slope for reflection at all points, but we will get the desired pattern when we design for only small deviations from a pencil beam. The surface can be designed with or without a caustic depending on the reflection angles at the edges. Reflectors designed with caustic edge reflections have a better chance of being unique [60]. Carberry [61] presents a method of analysis that involves physical optics. When we apply these methods, we must subdivide the reflector into many patches because the phases of the currents change rapidly with position on the reflector, and the analysis must be repeated with finer and finer patches until the result converges.

8-20.4 Dual Shaped Reflectors

We can design a dual-reflector antenna to produce an arbitrary phase and amplitude in the aperture plane by shaping both reflectors. By using both the conservation of power

and the differential equations of reflection on the two surfaces, Galindo [62] derived a pair of differential equations in terms of the aperture radius. Runge–Kutta or any other suitable numerical method can be used to solve the simultaneous differential equations instead of an integration of the power equation. Williams [63] finds a solution to Cassegrain antennas within the restriction of equal amplitude and phase in the aperture plane by integration of the power equation. Collins [64] considers using a parabolic reflector for the main reflector, since the difference between the shaped main reflector and a parabola is small. He accepts a quadratic phase error in the aperture. Existing large reflectors can be retrofitted with a shaped subreflector to improve performance. For the method to work, an axisymmetric feed such as a corrugated horn is required.

Galindo-Israel and Mittra [65] use a pair of reflectors offset from each other to transform a spherical wave from a feed antenna into a second spherical wave with a modified pattern amplitude. This combination of a feed with two reflectors can illuminate either prime focus paraboloidal reflectors or Cassegrain systems without modification of the existing reflector surfaces. For example, a $\sec^4(\theta/2)$ pattern can be realized from an ordinary pattern source to increase the aperture efficiency of the overall reflector system. The reflectors maintain equal GO path lengths for all rays, but they only approximate the desired pattern amplitude from the virtual focus. The procedure can be used to determine the contours of the reflectors along radial lines through numerical solution of differential equations. The equations develop from simplifying assumptions that depend on the extra degree of freedom introduced by the second reflector. In most cases, solution of the equations produces usable designs, although the method is not exact.

Lee et al. [66] developed a method to shape offset-fed dual reflectors that reduces to the solution of a differential equation similar to that of the single-reflector design given above. The reflection properties of the subreflector determine the main reflector amplitude distribution to first order. This method does not produce exact results but is close enough for engineering purposes. We start with a desired aperture power distribution $P(r, \phi_c)$ and a known feed power pattern $G_f(\theta', \phi_c)$ given in the radial direction ϕ_c. Most cases use distributions independent of ϕ_c, but the design is performed along these planes. For a circularly symmetric design we only need to solve the differential equation along one plane, but the general case requires solutions along enough planes to allow splines along the coordinate ϕ_c to find every point on both reflectors.

A differential expression relates the feed power to the aperture power:

$$G_f(\theta', \phi_c) \sin\theta' \, d\theta' = P(r, \phi_c) r \, dr$$

This leads to a ratio of integrals:

$$\frac{\displaystyle\int_{-\theta_e}^{\theta} G_f(\theta') \sin\theta' \, d\theta'}{\displaystyle\int_{-\theta_e}^{\theta_e} G_f(\theta') \sin\theta' \, d\theta'} = \frac{\displaystyle\int_{R_1}^{R} P(r') r' \, dr'}{\displaystyle\int_{R_1}^{R_2} P(r') r' \, dr'} \tag{8-110}$$

Equation (8-110) covers the general case where the offset subreflector directs power from a lower angle $-\theta_e$ to an offset radius R_1 that changes for each plane ϕ_c. For a circularly symmetric design, $-\theta_e = 0$ and $R_1 = 0$. Although many designs attempt to generate a uniform aperture distribution for the main reflector, we can substitute

any distribution, such as a circular Taylor distribution to control the sidelobes into Eq. (8-110). Given the aperture distribution and the feed pattern, we calculate a table similar to Table 8-12 for each plane ϕ_c that gives the feed angle as a function of aperture radius. We interpolate on this table to determine every value.

We start at the center of the subreflector described in spherical coordinates $(\rho_0, 0, 0)$ relative to the axis of the subreflector centered at the feed focus. The subreflector axis may be tilted relative to the main reflector axis. The rectangular coordinates of the subreflector are $(\rho \sin\theta \cos\phi_c, \rho \sin\theta \sin\phi_c, \rho \cos\theta)$. The incident wave reflects to a point on the main reflector: $(H \pm R\cos\phi_c, \pm R \sin\phi_c, z)$ using + Cassegrain, −Gregorian. We calculate the unit vector between the subreflector point and the main reflector. The normal vector on the subreflector is expressed as a differential:

$$\mathbf{n} = \frac{1}{\Delta}\left(\mathbf{a}_\rho - \frac{1}{\rho}\frac{\partial\rho}{\partial\theta}\mathbf{a}_\theta - \frac{1}{\rho\sin\theta}\frac{\partial\rho}{\partial\phi_c}\mathbf{a}_{\phi_c}\right) \tag{8-111}$$

where

$$\Delta = \sqrt{1 + \left(\frac{1}{\rho}\frac{\partial\rho}{\partial\theta}\right)^2 + \left(\frac{1}{\rho\sin\theta}\frac{\partial\rho}{\partial\phi_c}\right)^2}$$

We apply both equations of Snell's law [Eq. (2-67)] to the subreflector reflection and gather terms to form a pair of differential equations:

$$\frac{\partial\rho}{\partial\theta} = \frac{QV}{Q^2+U^2} \quad \text{and} \quad \frac{\partial\rho}{\partial\phi_c} = \frac{UV\sin\theta}{Q^2+U^2} \tag{8-112}$$

The terms of Eq. (8-112) are given by the expressions

$$\begin{aligned}
Q &= \frac{a\cos\theta\cos\phi_c + b\sin\theta\sin\phi_c - c\sin\theta}{\rho} \\
U &= \frac{b\cos\phi_c - a\sin\phi_c}{\rho} \\
V &= L + a\sin\theta\cos\phi_c + b\sin\theta\sin\phi_c + c\sin\theta \\
a &= H \pm R\cos\phi_c - \rho\sin\theta\cos\phi_c \\
b &= \pm R\sin\phi_c - \rho\sin\theta\sin\phi_c \\
c &= z - \rho\cos\theta
\end{aligned} \tag{8-113}$$

where the vector (a, b, c) is from the subreflector to the main reflector and $L = \sqrt{a^2+b^2+c^2}$.

We choose $z = 0$ as the aperture and equate path lengths along every ray. This gives an equation for the z-position of the main reflector:

$$\overline{OL} = \rho_0 + L_0 - z_0 = \rho + L - z$$

We solve for z:

$$z = \frac{a^2+b^2}{2(\rho\cos\theta - \rho + \overline{OL})} + \frac{1}{2}(\rho\cos\theta - \rho + \overline{OL}) \tag{8-114}$$

To solve for the reflector surfaces, we choose a starting point, usually the center of the reflector as the first ray from the feed to the subreflector, and calculate the initial distance L_0 between the subreflector and the main reflector to find the path length. We select a polar plane ϕ_c and solve the left differential equation [Eq. (8-112)] for both surfaces using a Runge–Kutta numerical solution. We repeat this in a sufficient number of planes ϕ_c to specify the surface totally. If the antenna is circularly symmetric, we solve the equation only once. For an offset dual reflector we can improve the cross-polarization by computing an equivalent subreflector using least squares and use its eccentricity to calculate the Mizugutch subreflector axis rotation.

8-21 OPTIMIZATION SYNTHESIS OF SHAPED AND MULTIPLE-BEAM REFLECTORS

Silver [2] discusses using a linear array feed to shape the beam of a paraboloidal reflector. The method is quite empirical and involves the addition of a number of offset beams. A similar technique is used in three-dimensional radar, but the feeds are kept separate so that multiple beams can scan a larger area in a given time. An array feed provides the best solution to beam shaping in many cases. The number of elements in the array limits the number of variables to a finite set to which optimization techniques can be applied.

A second method uses optimization to shape the reflector and possible subreflectors. This requires distortion functions on the reflectors. We start with conic-section reflectors and add distortions. These distortions can be global Zernike functions defined over the total surface, or they could be localized functions such as B-splines [67]. A B-spline uses a grid of points on the reflector, but the spline coefficients apply only over a limited area. In both cases we obtain a set of coefficients used in the optimization algorithms. We have the choice of combining these coefficients, or we can iterate between different sets of coefficients. Optimization is an art.

Because the reflector is an aperture antenna, we pick a set of directions in $(u, v) = (\sin\theta\cos\phi, \sin\theta\sin\phi)$ space to evaluate the pattern. The number of points should exceed the number of coefficients and be spaced close enough to fully describe the main-beam pattern:

$$\Delta u \text{ and } \Delta v \sim \frac{0.5\lambda}{D} \text{ to } \frac{0.25\lambda}{D} \tag{8-115}$$

We calculate the pattern power $P_m(u, v)$ at these points and compare them to the desired pattern $P_m^d(u,v)$ using a suitable cost function. We weight each pattern direction ω_m and use a summation cost [68] with a gradient minimization technique:

$$F(\mathbf{x}) = \sum_{m=1}^{M} |\omega_m(P_m(u_m, v_m)) - P_m^d(u_m, v_m)|^2 \tag{8-116}$$

A second choice is a min–max optimization [69]. This algorithm minimizes the maximum error:

$$\max[\omega_m(P_m(u_m, v_m)) - P_m^d(u_m, v_m)] \tag{8-117}$$

If we optimize the reflector shape, we express the distortion as B-splines specified at evenly spaced points across the aperture with the number determined by the maximum pattern angle $\theta_{\max}$ and the reflector diameter D [69]:

$$N_x = N_y = \frac{\pi D \sin\theta_{\max}}{\lambda} + 2 \tag{8-118}$$

Given a Zernike polynomial expansion with maximum azimuthal mode expansion $M_{\max}$ and maximum polar mode index $N_{\max}$, we have similar mode number requirements:

$$M_{\max} = N_{\max} = \frac{\pi D \sin\theta_{\max}}{\lambda} + 2 \tag{8-119}$$

Shaping starts with a paraboloid main reflector whose beamwidth may be so narrow that a portion of the specified $u-v$ space area may lie in the sidelobe region. In this case the optimization may become trapped because it cannot satisfy this area when changes effecting the main beam region positively affect the sidelobe region negatively. We must distort the main reflector before starting the optimization [69]. First surround the $u-v$ space area of specified points with an ellipse centered at (u_0, v_0) with major radius ω_1 and minor radius ω_2 tilted an angle α. Given a paraboloid with diameter D, focal length f, and center offset (x_0, y_0), we define rotated coordinates on the aperture.

$$x' = (x - x_0)\cos\alpha + (y - y_0)\sin\alpha$$
$$y' = -(x - x_0)\sin\alpha + (y - y_0)\cos\alpha$$

Using these coordinates, we alter the z-axis position of the reflector:

$$\Delta z = -\left(\frac{1}{2} + \frac{x^2 + y^2}{8f^2}\right)\left[\frac{\omega_1 x'^2 + \omega_2 y'^2}{D} + u_0(x - x_0) + v_0(y - y_0)\right] \tag{8-120}$$

We have a choice with ω_1 and ω_2 because they can be both either positive or negative. Positive values flatten the reflector while negative values cause a caustic reflection to broaden the beam.

REFERENCES

1. A. W. Love, *Reflector Antennas*, IEEE Press, New York, 1978.
2. S. Silver, ed., *Microwave Antenna Theory and Design*, McGraw-Hill, New York, 1949.
3. W. V. T. Rusch and P. D. Potter, *Analysis of Reflector Antennas*, Academic Press, New York, 1970.
4. P. J. Wood, *Reflector Antenna Analysis and Design*, Peter Peregrinus, London, 1980.
5. C. A. Mentzer and L. Peters, A GTD analysis of the far-out sidelobes of Cassegrain antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-23, no. 5, September 1975, pp. 702–709.
6. S. W. Lee et al., Diffraction by an arbitrary subreflector: GTD solution, *IEEE Transactions on Antennas and Propagation*, vol. AP-27, no. 3, May 1979, pp. 305–316.
7. A. D. Craig and P. D. Simms, Fast integration techniques for reflector antenna pattern analysis, *Electronics Letters*, vol. 18, no. 2, January 21, 1982, pp. 60–62.

8. V. Galindo-Israel and R. Mittra, A new series representation for the radiation integral with application to reflector antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-25, no. 5, September 1977, pp. 631–641.
9. A. W. Rudge et al., eds., *The Handbook of Antenna Design*, Vol. 1, Peter Peregrinus, London, 1982.
10. A. C. Ludwig, The definition of cross polarization, *IEEE Transactions on Antennas and Propagation*, vol. AP-21, no. 1, January 1973, pp. 116–119.
11. J. R. Cogdell and J. H. Davis, Astigmatism in reflector antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-21, no. 4, July 1973, pp. 565–567.
12. Y. T. Lo, On the beam deviation factor of a parabolic reflector, *IEEE Transactions on Antennas and Propagation*, vol. AP-8, no. 3, May 1960, pp. 347–349.
13. J. Ruze, Lateral-feed displacement in a paraboloid, *IEEE Transactions on Antennas and Propagation*, vol. AP-13, no. 5, September 1965, pp. 660–665.
14. W. V. T. Rusch and A. C. Ludwig, Determination of the maximum scan-gain contours of a beam scanned paraboloid and their relation to the Petzval surface, *IEEE Transactions on Antennas and Propagation*, vol. AP-21, no. 2, March 1973, pp. 141–147.
15. W. A. Imbriale, P. G. Ingerson, and W. C. Wong, Large lateral feed displacements in a parabolic reflector, *IEEE Transactions on Antennas and Propagation*, vol. AP-22, no. 6, November 1974, pp. 742–745.
16. D. K. Cheng, Effect of arbitrary phase errors on the gain and beamwidth characteristics of radiation pattern, *IEEE Transactions on Antennas and Propagation*, vol. AP-3, no. 4, July 1955, pp. 145–147.
17. J. Ruze, Antenna tolerance theory: a review, *Proceedings of IRE*, vol. 54, no. 4, April 1966, pp. 633–640.
18. M. S. Zarghamee, On antenna tolerance theory, *IEEE Transactions on Antennas and Propagation*, vol. AP-15, no. 6, November 1967, pp. 777–781.
19. R. C. Hansen, ed., *Microwave Scanning Antennas*, Academic Press, New York, 1964.
20. W. V. T. Rusch and R. D. Wanselow, Boresight gain loss and gore related sidelobes of an umbrella reflector, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 1, January 1982, pp. 153–157.
21. W. H. Watson, The field distribution in the focal plane of a paraboloidal reflector, *IEEE Transactions on Antennas and Propagation*, vol. AP-12, no. 5, September 1964, pp. 561–569.
22. T. B. Vu, Optimization of efficiency of reflector antennas: approximate method, *Proceedings of IEE*, vol. 117, January 1970, pp. 30–34.
23. B. M. Thomas, Theoretical performance of prime focus paraboloids using cylindrical hybrid modes, *Proceedings of IEE*, vol. 118, November 1971, pp. 1539–1549.
24. N. Amitay and H. Zucker, Compensation of spherical reflector aberrations by planar array feeds, *IEEE Transactions on Antennas and Propagation*, vol. AP-20, no. 1, January 1972, pp. 49–56.
25. V. Galindo-Israel, S. W. Lee, and R. Mittra, Synthesis of laterally displaced cluster feed for a reflector antenna with application to multiple beams and contoured patterns, *IEEE Transactions on Antennas and Propagation*, vol. AP-26, no. 2, March 1978, pp. 220–228.
26. B. Popovich et al., Synthesis of an aberration corrected feed array for spherical reflector antennas, *IEEE/APS Symposium Digest*, May 1983.
27. V. Mrstik, Effect of phase and amplitude quantization errors on hybrid phased-array reflector antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 6, November 1982, pp. 1233–1236.
28. C. M. Knop, On the front to back ratio of a parabolic dish antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-24, no. 1, January 1976, pp. 109–111.

29. M. Uhm, A. Shishlov, and K. Park, Offset-paraboloid geometry: relations for practical use, *IEEE Antennas and Propagation Magazine*, vol. 38, no. 3, June 1996, pp. 77–79.

30. R. F. H. Yang, Illuminating curved passive reflector with defocused parabolic antenna, *1958 IRE Wescon Convention Record*, August 1958, pp. 260–265.

31. C. Granet, Designing axially symmetric Cassegrain and Gregorian dual-reflector antennas from combinations of prescribed geometric parameters, *IEEE Antennas and Propagation Magazine*, vol. 40, no. 2, April 1998, pp. 76–82.

32. C. Granet, Designing axially symmetric Cassegrain and Gregorian dual-reflector antennas from combinations of prescribed geometric parameters, part 2: minimum blockage condition while taking into account the phase-center of the feed, *IEEE Antennas and Propagation Magazine*, vol. 40, no. 3, June 1998, pp. 82–85.

33. W. V. T. Rusch, Phase error and associated cross polarization effects in Cassegrainian-fed microwave antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-14, no. 3, May 1966, pp. 266–275.

34. P.-S. Kildal, The effects of subreflector diffraction on the aperture efficiency of a conventional Cassegrain antenna: an analytical approach, *IEEE Transactions on Antennas and Propagation*, vol. AP-31, no. 6, November 1983, pp. 903–909.

35. A. M. Isber, Obtaining beam-pointing accuracy with Cassegrain antennas, *Microwaves*, August 1967, pp. 40–44.

36. W. V. T. Rusch and R. Wohlleben, Surface tolerance loss for dual-reflector antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 4, July 1982, pp. 784–785.

37. A. F. Kay, Electrical design of metal space frame radomes, *IEEE Transactions on Antennas and Propagation*, vol. AP-13, no. 2, March 1965, pp. 188–202.

38. P.-S. Kildal, E. Olsen, and J. A. Aas, Losses, sidelobes, and cross polarization caused by feed-support struts in reflector antennas: design curves, *IEEE Transactions on Antennas and Propagation*, vol. AP-36, no. 2, February 1988, pp. 182–190.

39. W. V. T. Rusch et al., Forward scattering from square cylinders in the resonance region with application to aperture blockage, *IEEE Transactions on Antennas and Propagation*, vol. AP-24, no. 2, March 1976, pp. 182–189.

40. G. T. Ruck, ed., *Radar Cross Section Handbook*, Vol. 1, Plenum Press, New York, 1970.

41. G. W. Collins, Noise temperature calculations from feed system characteristics, *Microwave Journal*, vol. 12, December 1969, pp. 67–69.

42. B. E. Kinber, On two-reflector antennas, *Radioengineering and Electronics*, vol. 7, no. 6, 1962, pp. 973–980.

43. A. P. Popov and T. A. Milligan, Amplitude aperture-distribution control in displaced-axis two reflector antennas, *IEEE Antennas and Propagation Magazine*, vol. 39, no. 6, December 1997, pp. 58–63.

44. S. P. Morgan, Some examples of generalized Cassegrainian and Gregorian antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-12, no. 6, November 1964, pp. 685–691.

45. C. Granet, A simple procedure for the design of classical displaced-axis dual-reflector antennas using a set of geometric parameters, *IEEE Antennas and Propagation Magazine*, vol. 41, no. 6, December 1999, pp. 64–72.

46. T. A. Milligan, The effects of feed movement on the displaced-axis dual reflector, *IEEE Antennas and Propagation Magazine*, vol. 40, no. 3, June 1998, pp. 86–87.

47. Y. Mizugutch, M. Akagawa, and H. Yokoi, Offset dual reflector antenna, *IEEE Symposium on Antennas and Propagation Digest*, 1976, pp. 2–5.

48. P. H. Nielson and S. B. Sørensen, *Grasp8 Software Users Manual*, Ticra, Copenhagen, 2001.

49. C. Granet, Designing classical offset Cassegrain or Gregorian dual-reflector antennas from combinations of prescribed geometric parameters, *IEEE Antennas and Propagation Magazine*, vol. 44, no. 3, June 2002, pp. 114–123.
50. C. Granet, Designing classical offset Cassegrain or Gregorian dual-reflector antennas from combinations of prescribed geometric parameters, part 2: feed-horn blockage conditions, *IEEE Antennas and Propagation Magazine*, vol. 45, no. 6 December 2003, pp. 86–89.
51. D. T. Thomas, Design of multiple-edge blinders for large horn reflector antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-21, no. 2, March 1973, pp. 153–158.
52. S. R. Jones and K. S. Kelleher, A new low noise, high gain antenna, *IEEE International Convention Record*, March 1963, pp. 11–17.
53. C. Dragone, Offset multireflector antennas with perfect pattern symmetry and polarization discrimination, *Bell System Technical Journal*, vol. 57, no. 7, September 1978, pp. 2663–2684.
54. C. Granet, Designing classical dragonian offset dual-reflector antennas from combinations of prescribed geometric parameters, *IEEE Antennas and Propagation Magazine*, vol. 43, no. 6, December 2001, pp. 100–107.
55. T. Li, A study of spherical reflectors as wide-angle scanning antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-7, no. 4, July 1959, p. 223–226.
56. R. Woo, A multiple-beam spherical reflector antenna, *JPL Quarterly Technical Review*, vol. 1, no. 3, October 1971, pp. 88–96.
57. A. W. Love, Spherical reflecting antennas with corrected line sources, *IEEE Transactions on Antennas and Propagation*, vol. AP-10, no. 5, September 1962, pp. 529–537.
58. F. S. Bolt and E. L. Bouche, A Gregorian corrector for spherical reflectors, *IEEE Transactions on Antennas and Propagation*, vol. AP-12, no. 1, January 1964, pp. 44–47.
59. A. S. Dunbar, Calculation of doubly curved reflectors for shaped beams, *Proceedings of IRE*, vol. 36, no. 10, October 1948, pp. 1289–1296.
60. R. S. Elliott, *Antenna Theory and Design*, Prentice-Hall, Englewood Cliffs, NJ, 1981.
61. T. F. Carberry, Analysis theory for the shaped beam doubly curved reflector antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-17, no. 2, March 1969, pp. 131–138.
62. V. Galindo, Design of dual reflector antennas with arbitrary phase and amplitude distributions, *IEEE Transactions on Antennas and Propagation*, vol. AP-12, no. 4, July 1964. pp. 403–408.
63. W. F. Williams, High efficiency antenna reflector, *Microwave Journal*, vol. 8, July 1965, pp. 79–82.
64. C. Collins, Shaping of subreflectors in Cassegrainian antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-21, no. 3, May 1973, pp. 309–313.
65. V. Galindo-Israel and R. Mittra, Synthesis of offset dual shaped subreflector antennas for control of Cassegrain aperture distributions, *IEEE Transactions on Antennas and Propagation*, vol. AP-32, no. 1, January 1984, pp. 86–92.
66. J. J. Lee, L. I. Parad, and R. S. Chu, A shaped offset-fed dual-reflector antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-27, no. 2, March 1979, pp. 165–171.
67. M. E. Mortenson *Geometric Modeling*, Wiley, New York, 1985.
68. C. C. Han and Y. Hwang, Satellite antennas, Chapter 21 in Y. T. Lo and S. W. Lee, eds., *Antenna Handbook*, Van Nostrand Reinhold, New York, 1993.
69. H.-H. Viskum, S. B. Sørensen, and M. Lumholt, *User's Manual for POS4*, Ticra, Copenhagen, 2003.

9

LENS ANTENNAS

In lenses, as in parabolic reflectors, we utilize free space as a feed network to excite a large aperture. Because we locate the feed behind the aperture, the configuration eliminates aperture blockage and allows direct connection of the feed to the transmitter or receiver. When frequencies are above microwaves, this feeding method removes lossy transmission lines that increase system noise.

Lenses have only half the tolerance requirements of reflectors because the wave passes by the anomaly only once. In a reflector the wave path deviates by twice the distance as the wave travels to and from the reflector. At low microwave frequencies the lens is prohibitively heavy, but zoning and the use of artificial dielectrics reduce this problem. Both zoning and artificial dielectrics present mechanical stability problems and narrow the bandwidth.

We organize the design of lenses by the available degrees of freedom. A single lens with a uniform dielectric has two surfaces and is equivalent to a dual reflector because each surface is a degree of freedom. We start our discussion with single-surface lenses where we eliminate one degree of freedom by making the second surface match either the incoming or outgoing wave. Shaping both surfaces lets us correct one lens anomaly. We can remove either coma to improve the feed scanning or design to convert a given feed pattern to a desired aperture distribution. Bootlace lenses have three possible degrees of freedom. They consist of back-to-back arrays with cables connecting the sides. Normally, we use the degrees of freedom of the bootlace lens to increase the number of focal points. We give up degrees of freedom in many designs to simplify the mechanical layout. Finally, we discuss the use of a variable index of refraction in the Luneburg lens.

We base the design of lenses on geometric optics. Like parabolic reflectors, lenses have no inherent frequency bandwidth limitation. We are limited by the feeds and mechanical problems of large sizes. Because we borrow from optics, lenses have great high-frequency potential.

Modern Antenna Design, Second Edition, By Thomas A. Milligan

9-1 SINGLE REFRACTING SURFACE LENSES

Single-surface lenses convert the wave type, such as spherical to planar, at one surface through refraction. The constant-phase surface (eikonal) of the wave type determines the shape of the second surface of the lens. The common lens converts an incident spherical or cylindrical wave to a plane wave. Conversion to a cylindrical wave requires a line source feed and cylindrical surfaces on the lens. Spherical waves use point feeds and axisymmetrical surfaces. Like the reflector, which also converts spherical waves from the feed to plane waves by geometric optics (GO), diffraction from an aperture determines the far-field pattern.

Consider the second or nonrefracting surface. If the surface toward the feed converts the wave type, the wave exits the second surface as a plane wave and it is a plane. Similarly, when the surface away from the feed converts the exiting wave to a plane wave, the inner surface toward the feed follows the incident wave eikonal, a cylindrical or spherical surface. Figure 9-1 shows the two types of single refracting surface lenses. We can determine the refracting surface shape by either of two different approaches. Snell's law can be applied to the refracting surface, and the surface slope can be determined for each feed angle. Equivalently, we can apply Fermat's principle to equalize the optical path length from the feed through the lens to an aperture plane. The designs are easily found [1,2]. For Figure 9-1*a*,

$$\rho(\psi) = \frac{(n-1)f}{n\cos\psi - 1} \tag{9-1}$$

where n is the index of refraction, given by

$$n = \sqrt{\varepsilon_r \mu_r} \tag{9-2}$$

and ε_r and μ_r are the relative permittivity and permeability of the lens medium. When $n > 1$, Eq. (9-1) describes a hyperbola with the feed at one focus. The distance from the feed to the hyperbola along the axis is f. The asymptotes of the hyperbola limit the collimated portion of the feed radiation:

$$\psi_a = \cos^{-1}\frac{1}{n} \tag{9-3}$$

We must limit the lens edge to angles less than ψ_a because the asymptotes imply an infinite aperture size. Similar to the paraboloidal reflector, we have feed spillover, considered to be lost in sidelobes. But, for example, placing the lens at the aperture of a horn eliminates spillover. We calculate the lens diameter from

$$D = 2\rho\sin\psi_e = \frac{2(n-1)f\sin\psi_e}{n\cos\psi_e - 1} \tag{9-4}$$

where ψ_e is the edge angle, subject to the restriction of Eq. (9-3). The surfaces of the lens in Figure 9-l*b* have the polar equations

$$\rho_1 = \text{constant} \qquad \rho_2(\psi) = \frac{(n-1)f}{n - \cos\psi} \tag{9-5}$$

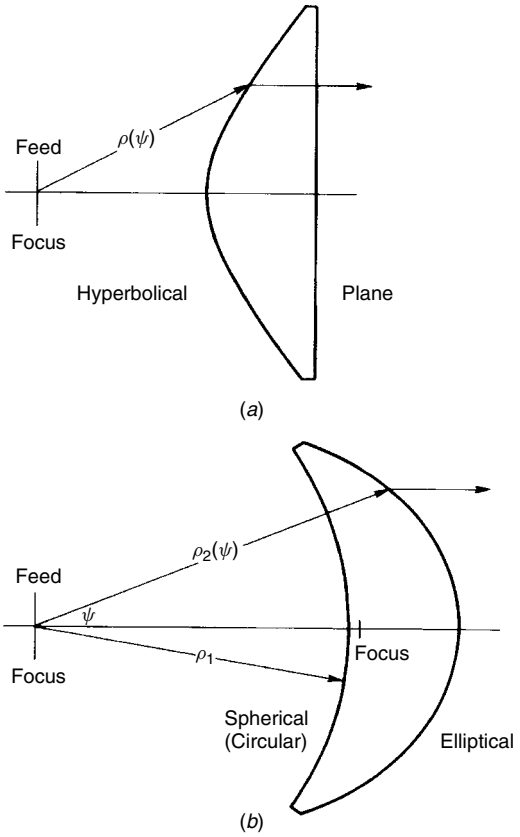


FIGURE 9-1 Single-surface lenses.

The inner surface must either be a circular cylinder (cylindrical lens) or be spherical (axisymmetrical lens). The outer surface $\rho_2(\psi)$ is elliptical for $n > 1$. The junction of the circle and ellipse determines the feed angle limitation:

$$\cos \psi_e = n - \frac{(n-1)f}{\rho_1} \tag{9-6}$$

We can, of course, truncate the lens before the two curves meet. Equation (9-6) gives the limitation on ρ_1 at the lens edge as well:

$$\rho_1 \leq \frac{(n-1)f}{n - \cos \psi_e} \tag{9-7}$$

Example Compute f and ρ_1 at the edge for an elliptical lens (Figure 9-l*b*) with $D = 10\lambda$, $n = 1.6$ (polystyrene), and $\psi_e = 50^\circ$.

Solve Eq. (9-5) for f/D:

$$\frac{f}{D} = \frac{n - \cos \psi_e}{2 \sin \psi_e (n-1)} \qquad f = 10.41\lambda \tag{9-8}$$

$$\rho_1 \leq \frac{D}{2 \sin \psi_e} \tag{9-9}$$

If ρ_1 remained constant to the center of the lens, it would be 3.88λ thick. In narrowband applications we can remove the multiple wavelength thicknesses by zoning and reduce weight and dielectric material loss.

Lenses change the amplitude distribution of the feed in the aperture plane. We relate the feed pattern to the aperture distribution through the conservation of power in differential areas. For an axisymmetrical lens:

$$\underset{\text{feed power}}{F(\psi, \phi) \sin \psi \, d\psi \, d\phi} = \underset{\text{aperture power}}{A(r, \phi) r \, dr \, d\phi} \tag{9-10}$$

where ψ is the feed angle and r, $\rho \sin \psi$, is the aperture radial distance; $F(\psi, \phi)$ is the feed power pattern and $A(r, \phi)$ is the aperture power distribution:

$$\frac{A(r, \phi)}{F(\psi, \phi)} = \frac{\sin \psi}{r} \frac{d\psi}{dr} \tag{9-11}$$

For a cylindrical lens, we also equate differential area multiplied by the feed or aperture power:

$$\begin{aligned} F(\psi, y) \, d\psi \, dy &= A(r, y) \, dr \, dy \\ \frac{A(r, y)}{F(\psi, \phi)} &= \frac{d\psi}{dr} \end{aligned} \tag{9-12}$$

We substitute Eqs. (9-11) and (9-12) into Eq. (9-1) for the hyperbolical lens to calculate the aperture distribution relative to the feed power pattern.

$$\underset{\text{Axisymmetrical}}{\frac{A(r, \phi)}{F(\psi, \phi)} = \frac{(n \cos \psi - 1)^3}{f^2 (n-1)^2 (n - \cos \psi)}} \qquad \underset{\text{Cylindrical}}{\frac{A(r, y)}{F(\psi, y)} = \frac{(n \cos \psi - 1)^2}{f (n-1)(n - \cos \psi)}} \tag{9-13a, b}$$

The field distribution is the square root of Eq. (9-13).

We substitute Eqs. (9-11) and (9-12) into Eq. (9-5) for the elliptical lens:

$$\underset{\text{Axisymmetrical}}{\frac{A(r, \phi)}{F(\psi, \phi)} = \frac{(n - \cos \psi)^3}{f^2 (n-1)(n \cos \psi - 1)}} \qquad \underset{\text{Cylindrical}}{\frac{A(r, y)}{F(\psi, \phi)} = \frac{(n - \cos \psi)^2}{f (n-1)(n \cos \psi - 1)}} \tag{9-14a, b}$$

The hyperbolical and elliptical lenses concentrate the aperture power in different ways. The hyperbolical lens reduces the feed power directed toward the edges and produces an additional aperture taper. On the other hand, an elliptical single-surface lens increases the power toward the edges as compared with the center.

Example For axisymmetrical lenses with $\psi_e = 50^\circ$ and $n = 1.6$, compute the edge taper due to the lenses.

We divide Eq. (9-13*b*) with $\psi_e = \psi$ by the same equation with $\psi = 0$ to determine the ratio of power at the edge to that at the center of the aperture (assuming an isotropic feed). We do the same calculation with Eq. (9-14*a*):

$$\frac{A_e}{A_c} = \begin{cases} \dfrac{(n\cos\psi_e - 1)^3}{(n-1)^2(n-\cos\psi_e)} & \text{hyperbolical lens} \quad (9\text{-}15) \\ \dfrac{(n-\cos\psi_e)^3}{(n-1)^2(n\cos\psi_e - 1)} & \text{elliptical lens} \quad (9\text{-}16) \end{cases}$$

By substituting ψ_e and n, we compute the edge taper.

hyperbolical lens: 0.038 (−14.2 dB) elliptical lens : 7.14 (8.5 dB)

The increased taper of the hyperbolical lens reduces sidelobes, and the elliptical lens increases aperture efficiency by compensating for some of the feed antenna pattern taper to make the aperture distribution more uniform.

9-2 ZONED LENSES

Lenses designed by the methods of Section 9-1 have bandwidth limitations determined only by the invariability of the dielectric constant. Zoning removes multiples-of-wavelength path lengths in the lens to reduce weight, to reduce the lens-induced amplitude taper, or to thin the lens. The act of changing dimensions by wavelengths implies narrowing the frequency bandwidth.

We step the lens in either the nonrefracting or refracting surface. Stepping the nonrefracting surface (Figure 9-2*a, b*) has the least effect. The edges of the steps, parallel with the waves, will diffract waves and cause some change in the aperture fields, but GO predicts no effect. Stepping the refracting surface introduces losses either as misdirected feed power (Figure 9-2*c, d*) or as unexcited aperture (Figure 9-2*e, f*). But stepping the refracting surface reduces the lens-induced aperture taper by changing the focal lengths in various zones. Figure 9-2 shows the limits in the two types of refracting surface steps, since we could compromise between the directions and have both feed spillover and unexcited aperture.

We can easily calculate the step dimensions in Figure 9.2*a* and *b*. We equate the path lengths inside and outside the dielectric along the step with a difference of Δ or some integer multiple of λ:

$$\underset{\text{inside}}{n\Delta} = \underset{\text{outside}}{\Delta + \lambda}$$

The step Δ becomes

$$\Delta = \frac{\lambda}{n-1} \tag{9-17}$$

In Figure 9-2*c–f* we determine the change in focal lengths instead of the step dimensions. Zoning affects the optical path lengths in the center of the lens. We calculate the edge focal length from the unzoned case. The focal length increases by Eq. (9-17)

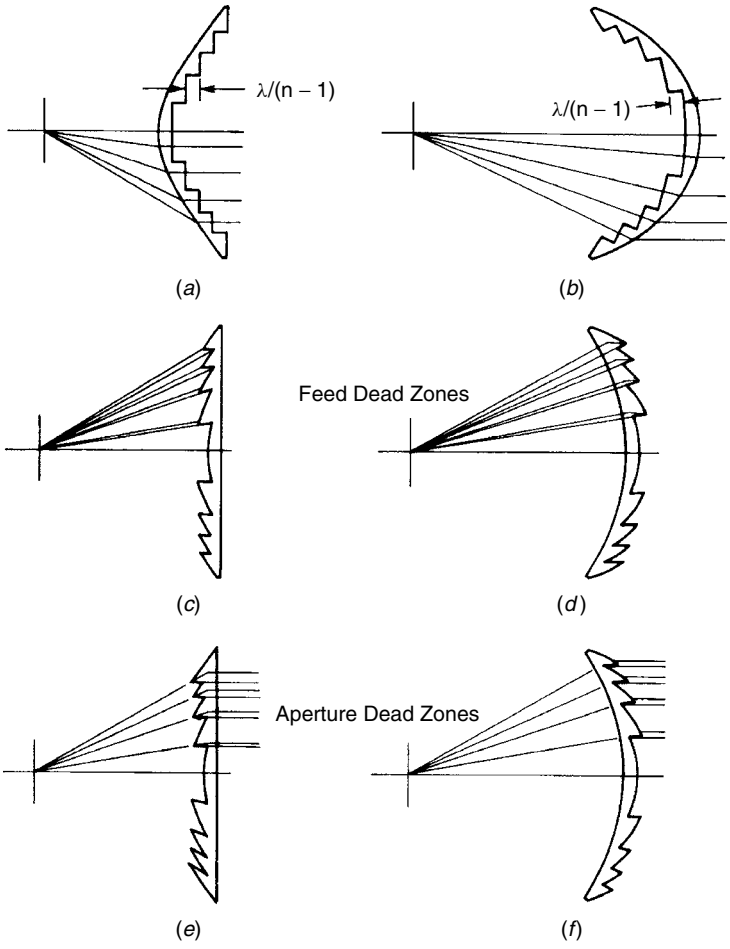


FIGURE 9-2 Zoning of single-surface lenses.

for each inner step of hyperbolical lenses and decreases by the same amount for each inner step of elliptical lenses. We derive the lens-induced taper relative to the center by using ratios of Eq. (9-13) or (9-14). For axisymmetrical lenses,

$$\frac{A(r,\phi)}{A_c} = \begin{cases} \dfrac{f_c^2(n\cos\psi - 1)^3}{f^2(n-1)^2(n-\cos\psi)} & \text{hyperbolical} \quad (9\text{-}18) \\ \dfrac{f_c^2(n-\cos\psi)^3}{f^2(n-1)^2(n\cos\psi - 1)} & \text{elliptical} \quad (9\text{-}19) \end{cases}$$

where f_c is the focal length in the center, f the focal length in the feed direction ψ, and A_c the central amplitude.

Example Design an axisymmetrical hyperbolical lens ($n = 1.6$) by using the three types of zoning shown in Figure 9-2, using an aperture diameter of 30λ and a maximum feed angle of 35° with a 70° 10-dB beamwidth feed.

The minimum allowable thickness is 0.5λ with a 0.3λ edge thickness. By working through the geometry, we obtain dimensions for the following cases:

Figure 9-2a, nonrefracting surface zoning, we compute steps $= \lambda/(n-1) = 1.67\lambda$.

Step	1	2	3	4
Aperture Radius (λ)	12.418	10.009	7.403	4.256

We estimate the feed spillover loss from Eq. (8-12) for the 10-dB feed edge taper, 0.41 dB. Equation (9-15) gives the edge taper (9.72 dB), since the refracting surface is not zoned. Combined with the feed edge taper of 10 dB, we have a 19.72-dB amplitude taper in the aperture plane. We use Eq. (4-8) to calculate the amplitude taper loss (1.8 dB) for this axisymmetric distribution.

Figure 9-2c, zoning in the refracting surface of the hyperbolical lens (parallel with outgoing waves), we compute the dimensions starting with the edge focal length found from a rearrangement of Eq. (9-4) (Table 9-1). The changing focal lengths in the zones alter the aperture amplitude distribution. The edge taper becomes [Eq. (9-18)] 6.24 dB and reduces the amplitude taper loss [Eq. (4-8)] to 1.19 dB. The portion of the feed directed to the dead zones refracts out of the aperture and forms sidelobes that reduce the aperture efficiency. We consider this as a second spillover loss (0.81 dB).

Figure 9-2e, zoning in the refracting surface of the hyperbolical lens (parallel with the feed wave rays), again we start with the edge focal length and increase it by $\lambda/(n-1)$ at each step and determine the feed angles where the change in the focal length will give the minimum allowable thickness. The dimensions given in Table 9-2 were obtained. Since the focal lengths are the same as for Fig. 9-2*c*, the lens-induced edge taper is 6.24 dB. The dead zones in the aperture distribution increase the amplitude taper loss to 3.10 dB. These ring dead zones can be considered as radiating and adding to the fully excited aperture pattern. They radiate patterns with closely spaced lobes that raise the near sidelobes of the total antenna. The three designs are compared in Table 9-3.

Example Similar to the example above, we can design zoned elliptical axisymmetrical lenses that have the same problems of feed angle or aperture dead zones.

The edge taper of the elliptical lens counteracts some of the feed taper and reduces amplitude taper loss. The losses calculated for those designs are given in Table 9-4.

TABLE 9-1 Zoned Hyperbolical Lens, Figure 9-2*c*

Zone	Focal Length (λ)	Aperture Radius of Step (λ)	Thickness (λ)	Feed Dead Zone Angles (deg)
1	20.21	0	1.52	
2	18.54	5.12	2.09	13.57–14.02
3	16.87	8.42	1.98	21.64–23.10
4	15.21	10.84	1.90	27.06–28.68
5	13.54	12.89	1.83	31.28–32.95

TABLE 9-2 Zoned Hyperbolical Lens, Figure 9-2*e*

Zone	Focal Length (λ)	Feed Angle (deg)	Thickness (λ)	Aperture Dead Zones (λ)
1	20.21	0	1.51	0
2	18.54	13.57	2.25	4.70–5.12
3	16.87	21.64	2.41	7.66–8.42
4	15.21	27.06	2.60	9.77–10.84
5	13.54	31.28	2.83	11.48–12.89

TABLE 9-3 Aperture Illumination Losses of Three Hyperbolical Lenses, Figure 9-2

Design (Figure 9-2)	*SPL* (dB)	*ATL* (dB)	Sum (dB)
(*a*)	0.41	1.80	2.21
(*c*)	1.22	1.19	2.41
(*e*)	0.41	3.10	3.51

TABLE 9-4 Aperture Illumination Losses of Three Elliptical Lenses, Figure 9-2

Design (Figure 9-2)	*SPL* (dB)	*ATL* (dB)	Sum (dB)
(*b*)	0.41	0.06	0.47
(*d*)	1.43	0.14	1.57
(*f*)	0.41	1.43	1.84

Zoning reduces the frequency bandwidth. At the center frequency the optical path length difference between the central ray and edge ray is $K - 1$ for K zones. A common maximum allowed deviation of feed to aperture path is 0.125λ, which leads to a bandwidth of

$$B \simeq \frac{25\%}{K - 1} \tag{9-20}$$

For the five zone lenses of the examples above, Eq. (9-20) gives a 6% bandwidth. We determine the loss at band edge by tracing rays from the feeds-to-aperture plane and calculating the phase error loss using Eq. (4-9). The loss at band edge is about 0.3 dB for all the designs. Bandwidth is greatly underestimated by Eq. (9-20) if a greater phase error loss is allowed. The 1-dB phase error loss bandwidth is $45\%/(K - 1)$.

9-3 GENERAL TWO-SURFACE LENSES

Optical lens designs use either flat or spherical surfaces, an approximation useful for long focal lengths. We design lenses exactly. In Section 9-1 we discussed lenses where the rays refracted at only one surface. The shape curve of these two lenses could be

derived easily. In this section we pick the curve for one lens surface and use a numerical search for the second surface that generates a table of (r, z) values. We generate a cubic spline from the table and use it to calculate the surface normal vector and the radius of curvature useful in manufacture. The initial numerical search generates ray paths through the lens that produce a table of feed angle versus aperture radius. We apply Eq. (9-11) to compute the aperture distribution after we generate the cubic spline whose evaluation includes the derivative.

If we specify the surface on the feed side, we start at the feed and trace rays to the inner surface. We know the inner surface normal vector because it is a specified surface. We compute the direction of the ray inside the lens given the incident medium index of refraction n_i, n_o inside the lens, the surface unit normal $\mathbf{n}$ directed into the lens, and the incident ray unit vector $\mathbf{S}_i$. First determine if the ray will exit the incident ray medium, because if $n_i > n_o$, it can act as a prism and have total reflection:

$$R_a = n_o^2 - n_i^2[1 - (\mathbf{S}_i \cdot \mathbf{n})^2]$$

If $R_a < 1$, the ray is totally reflected. For $R_a > 1$, we determine the direction of the output ray unit vector $\mathbf{S}_o$ from the following operations [3, p. 355]:

$$\gamma = \sqrt{R_a} - n_i \ (\mathbf{n} \cdot \mathbf{S}_i) \qquad \text{find the vector } \mathbf{t} = n_i \mathbf{S}_i + \gamma \mathbf{n} \qquad \mathbf{S}_o = \frac{\mathbf{t}}{|\mathbf{t}|} \tag{9-21}$$

For our case the normal vector, incident, and refracted rays are in two-dimensional space (z, r), because the lens has axisymmetry. To start, specify the z-axis distance from the feed to the outer rim of the lens, the initial radius, and the rim thickness along the initial internal ray direction. The feed-side position of the lens is (z_1, r_1). Apply Eq. (9-21) and find the outer lens point given the rim thickness t_r, by tracing along the vector $\mathbf{t}$ to point (z_2, r_2). Since the lens collimates the beam, we trace the ray to a plane $z = z_3$ whose normal is the z-axis. Figure 9-3 illustrates the rim and internal ray paths for a lens with a spherical inner surface. We calculate the electrical path length from the feed located at $z = z_f$ to the output plane by including the index

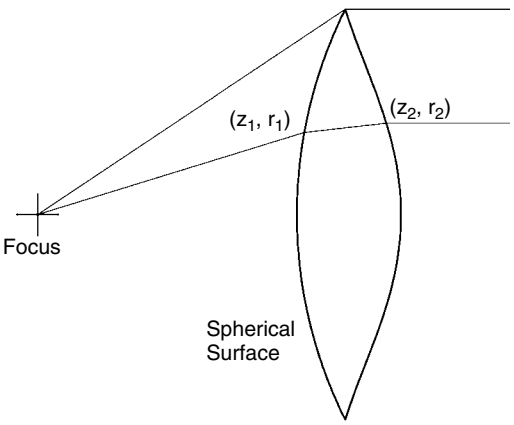


FIGURE 9-3 Ray tracing in a single-surface lens with a specified spherical inner surface.

of refraction n of the lens for the ray path length (PL) through the lens, and the input and output ray lengths:

$$\text{PL} = \sqrt{r_1^2 + (z_1 - z_f)^2} + n\sqrt{(r_2 - r_1)^2 + (z_2 - z_1)^2} + (z_3 - z_2) \tag{9-22}$$

The design consists of stepping in r_1, tracing the ray to the inner lens surface, computing the direction of the ray internal by Eq. (9-21), and determining thickness t_r when Eq. (9-22) minus the initial path length is zero. By Fermat's principle of equal path length, the outer surface refraction direction adds to the inner surface refraction to produce parallel output rays. This procedure generates a table of (z_2, r_2) pairs that we convert to a cubic spline. By using the cubic spline, we produce an evenly spaced table of (z_2, r_2) for machining, and if necessary, a table of radius of curvature from the second derivative to assist machining operations. The next step is to calculate a cubic spline between the aperture radius r_2 and the feed angle ψ because its output includes $d\psi/dr_2$. By rearranging Eq. (9-11), we compute the aperture power distribution given the feed power pattern:

$$A(r_2, \phi) = \frac{F(\psi, \phi) \sin \psi}{r_2} \frac{d\psi}{dr_2} \tag{9-23}$$

The design steps for a lens with the outer surface specified are similar except that we trace rays from the output plane backward to the feed point. Again, we start at the lens rim, use Eq. (9-21) to calculate the internal ray direction by using the $-z$-directed ray and the known surface normal to travel along this ray by the rim thickness to the inner surface. Equation (9-22) gives the electrical path length for this lens, as well. We repeat the root-finding procedure used above to determine a series of points (z_1, r_1) along the surface by equating all electrical path lengths. We generate the same series of cubic splines to obtain machining dimensions and differential $d\psi/dr_2$ used in Eq. (9-23) for the aperture power distribution. Figure 9-4 shows the aperture distributions for various lenses with focus points located 1.5 times the radius below the lens. The curves include lenses of Section 9-1.

The focal spot is not a singularity as drawn on the figures using geometric optics, but spreads due to the finite wavelength. We use Gaussian beams to evaluate the size of focal spots. For a lens with a collimated output, we assume a Gaussian beam on the output with minimum waist diameter $2W_0 = D$, the lens diameter, radiating into free space, and a matching Gaussian beam on the feed side which tapers to the focal plane. The lens transforms one Gaussian beam into another. The focal length $f = z_f$ and we determine the diameter of the focal spot $2W_0'$ and the half depth of focus b from the lens F-number $F_\# = f/D$ [4, pp. 91–95]:

$$2W_0' = \frac{4\lambda}{\pi} F_\# = \frac{4\lambda}{\pi} \frac{f}{D} \qquad b = \frac{8\lambda}{\pi} F_\#^2 \tag{9-24}$$

A small modification to Eq. (9-22) allows the design of lenses with a second focus at a finite position $z = z_3$ on the axis:

$$\text{PL} = \sqrt{r_1^2 + (z_1 - z_f)^2} + n\sqrt{(r_2 - r_1)^2 + (z_2 - z_1)^2} + \sqrt{r_2^2 + (z_3 - z_2)^2} \tag{9-25}$$

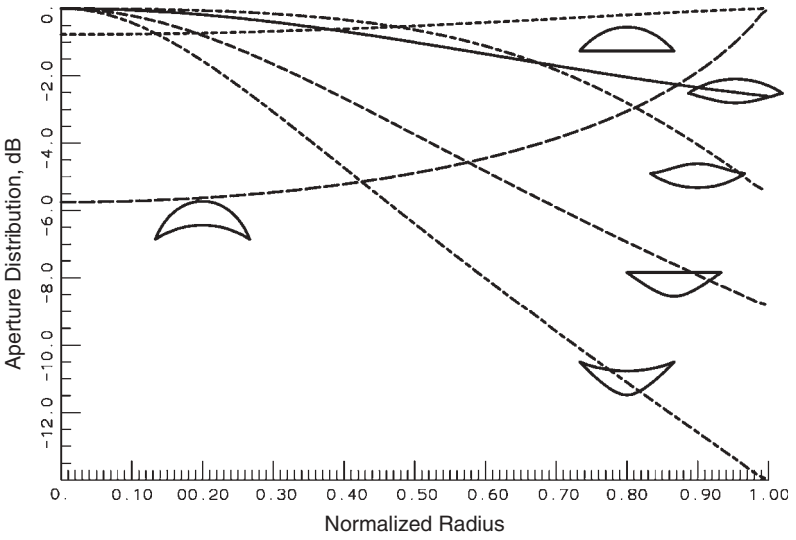


FIGURE 9-4 Aperture distribution for various single-surface lenses with $f/D = 1.5$.

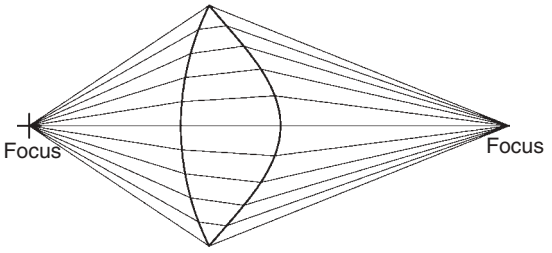


FIGURE 9-5 Ray tracing in lens designed for two finite focuses.

We follow the same steps as in the design above except that we need to generate a table of feed angles given the output ray angle with respect to the z-axis. Figure 9-5 illustrates a typical design and shows the ray tracing.

The lens designs noted above narrow the beamwidth. We design lenses to spread the beam by using a virtual focus located behind the output surface of the lens, as shown in Figure 9-6. Because the rays trace backward to the virtual focus, we change the sign of the last term in Eq. (9-25):

$$\text{PL} = \sqrt{r_1^2 + (z_1 - z_f)^2} + n\sqrt{(r_2 - r_1)^2 + (z_2 - z_1)^2} - \sqrt{r_2^2 + (z_3 - z_2)^2} \qquad (9\text{-}26)$$

The iterative design procedure should be modified to start in the center of the lens because the concave lens has a significant outer thickness. We compute focal length f of the lens by the distance of the two focuses from the lens:

$$\frac{1}{f} = \frac{1}{z_f} + \frac{1}{z_3} \qquad (9\text{-}27)$$

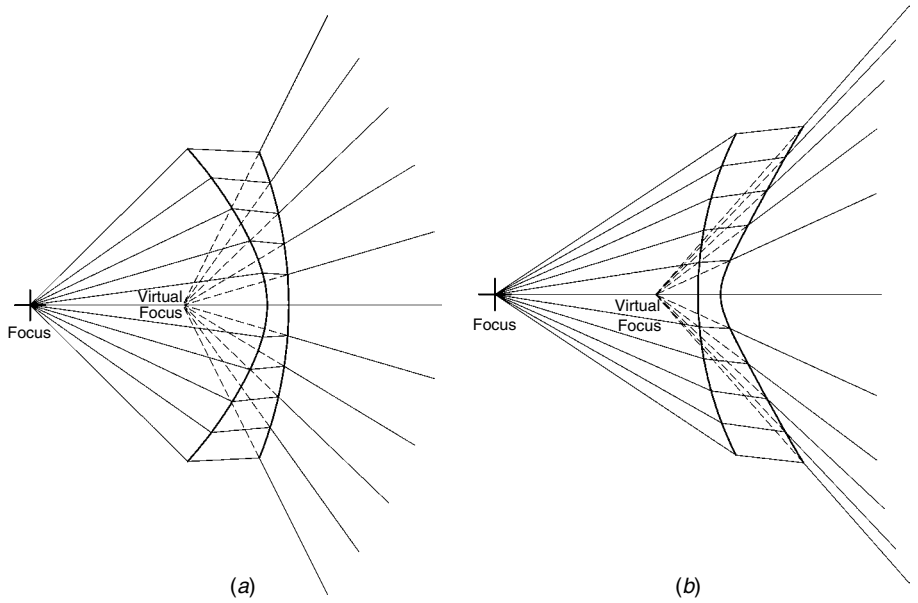


FIGURE 9-6 Lenses designed with virtual focuses to widen beam: (*a*) spherical outer surfaces; (*b*) spherical inner surfaces.

To calculate the pattern we trace rays through the lens either to a planar surface output side of the lens or the actual surface. By using a cubic spline between the feed angle and the aperture position we calculate the amplitude due to spreading [Eq. (9-23)]. We replace the fields with currents and use physical optics to calculate the pattern. A second simpler approach uses a Gaussian beam approximation for lenses that accounts for electrical size. Both the input and output Gaussian beam have the same waist at the lens plane. Each Gaussian beam decreases in a hyperbola to the minimum waist at the location of the phase center or focus. The output does not pass through a narrow caustic, as shown in Figure 9-5, but reaches a finite-diameter waist related to the feed beamwidth and the lens diameter.

At this point we design the lenses scaled to wavelengths. Given the feed beamwidth, we calculate the half depth of focus b by using Eq. (7-35). The lens magnifies the output beam waist compared to the input beam waist by M. First, we calculate an input magnification factor M_r:

$$M_r = \left| \frac{f}{z_f - f} \right| \tag{9-28}$$

With b given in wavelengths, we compute the ratio of b to the shift of the feed relative to the focus:

$$r = \frac{b}{z_f - f} \tag{9-29}$$

We combine Eqs. (9-28) and (9-29) to calculate the lens magnification M:

$$M = \frac{M_r}{\sqrt{1 + r^2}} \tag{9-30}$$

The output Gaussian beam propagates from the second focus with new parameters: b', the half depth of focus, waist W_0', and divergence angle θ_0' [4]:

$$b' = M^2 b \qquad W_0' = M W_0 \quad \theta_0' = \frac{\theta_0}{M} \tag{9-31}$$

Example The lens in Figure 9-5 was designed for a diameter of 20λ, $z_f = 15\lambda$, and $z_3 = 25\lambda$. We calculate the focal length from Eq. (9-27) as $f = 9.375\lambda$. A feed with a 68° 10-dB beamwidth is approximated by a Gaussian beam with $b = 0.99\lambda$ (Scale 7-8) and $2W_0 = 1.12\lambda$ (Scale 7-9). We use b to find the magnification M by using Eqs. (9-28) to (9-30), $M = 1.64$. The Gaussian beam at the second focus has a minimum waist $2W_0' = 1.64(1.12) = 1.84\lambda$. Using Scales 7-7 to 7-9, we read the Gaussian beam output values: 10-dB beamwidth = 42°(gain = 18.4 dB). Scale 7-7 gives the gain of the feed as 14.3 dB. The half depth of focus increased from 0.99λ to 3.39λ.

Example When we repeat the lens calculations for the lens of Figure 9-6*a* for a diameter = 20λ, $z_f = 15\lambda$, and $z_3 = -5\lambda$, we calculate $f = -7.5$. Starting with a feed 10-dB beamwidth = 80°, we use Scale 7-8 to find $b = 0.7\lambda$ and waist diameter $2W_0 = 0.94\lambda$. The negative focal length produces a magnification below 1: $M = 0.333$. This decreases b to 0.0775λ and $2W_0' = 0.314$. The output Gaussian beam gain drops to 6.6 dB from a feed gain of 12.9 dB.

9-4 SINGLE-SURFACE OR CONTACT LENSES

We can alter the pattern of antennas with planar surfaces such as spirals and microstrip patches by placing a lens directly on the planar surface. The lens can be spaced a small distance away to avoid potential damage and have little pattern impact. The lens modifies the original pattern of the antenna by using the refraction at the single output surface. Because the lens contacts the antenna, the pattern inside the lens is the same as radiated into free space except for the dielectric loading on the antenna. This loading shifts the operating frequency and in the case of a spiral improves its efficiency and widens its beamwidth (see Section 11-5.1).

We design these antennas by first generating a mapping between the feed angle ψ and the output angle θ. This could be as simple as $\theta(\psi) =$ constant or a function to generate a shaped beam, in the same manner as a shaped reflector (see Section 8-20). The relationship $\theta\ (\psi)$ enables calculation of the surface normal at every point along the outer surface. The surface normal is found from the gradient of the radial vector, and equating the two values produces a differential equation between the radius r and the feed angle ψ. Given the feed angle and the output angle at a point on the lens, we calculate the surface normal from $\mathbf{n} = n\mathbf{S}_i - \mathbf{S}_o$, where n is the index of refraction, $\mathbf{S}_i$ the incident unit vector, and $\mathbf{S}_o$ the exiting ray unit vector. We normalize $\mathbf{n}$ to a vector $\mathbf{v}$. The gradient of $r(\psi)$ gives a second expression for the normal vector:

$$\nabla r(\psi) = \mathbf{a}_r + \mathbf{a}_\psi \frac{1}{r(\psi)} \frac{\partial r(\psi)}{\partial \psi} \tag{9-32}$$

The coefficient of $\mathbf{a}_\psi$ in Eq. (9-32) is the tangent of the angle α between the normal vector and the radial vector (incident ray). The tangent can be found from the unit

vector $\mathbf{v}$ and the incident ray unit vector $\mathbf{S}_i$:

$$\tan\alpha = \frac{(\mathbf{S}_i \times \mathbf{v})}{\mathbf{S}_i \cdot \mathbf{v}} = \frac{1}{r(\psi)} \frac{\partial r(\psi)}{\partial \psi} \tag{9-33}$$

The cross product between the incident ray and the unit vector $\mathbf{v}$ only has a z-axis component, because the incident ray and the surface normal lie in the $x-y$ plane. We design the lens surface by solving the differential equation (9-33). A numerical technique such as the Runge–Kutta method easily solves the equation when we start at one feed angle and an arbitrary lens radius and step through feed angles. The method determines only the shape of the lens to an arbitrary size that we scale to the diameter desired.

Figure 11-9*a* illustrates a lens designed to redirect all feed rays to $\theta = 0$. Figure 9-7 illustrates the shape of a contact designed to redirect all rays between feed angles 0 through 60° to output rays at 30°. For a 3λ lens diameter and a feed 12-dB beamwidth of 120°, the lens spreads the beam to form a flat-topped output beam shown in Figure 9-8. Contact lenses can greatly modify radiation from a feed with electrically small lenses.

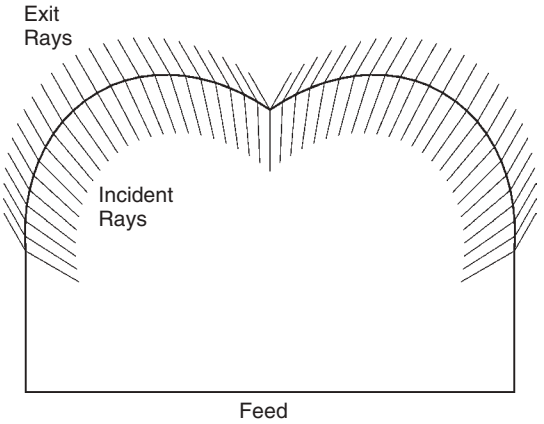


FIGURE 9-7 Contact lens designed to direct the beam in a cone at 30°.

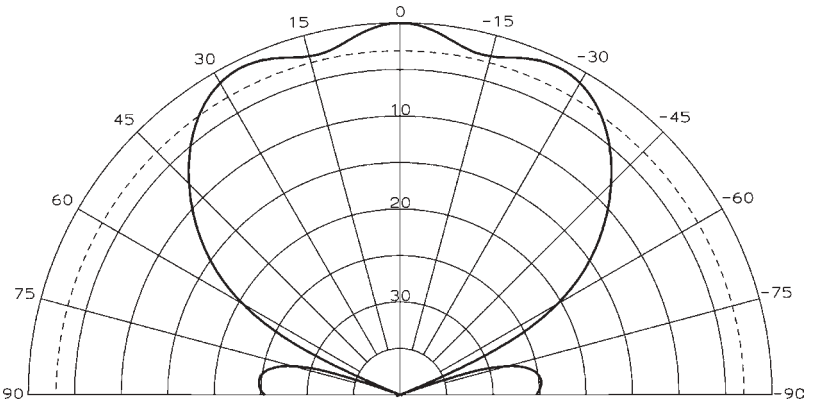


FIGURE 9-8 Pattern of contact 3λ-diameter lens for a feed with a 12-dB beamwidth of 120°.

9-5 METAL PLATE LENSES

The phase velocity of a wave in waveguide exceeds that of a wave in free space and produces a medium with an effective refractive index of less than 1. We can make a microwave lens by spacing parallel metal plates and feeding the lens with a wave polarized in the direction of the plates. For plates spaced a distance a, the index of refraction is

$$n = \sqrt{1 - \left(\frac{\lambda}{2a}\right)^2} \tag{9-34}$$

where λ is the wavelength in the medium between the plates. The index of refraction is frequency dependent. The lens can be made polarization independent by forming an egg crate of orthogonal plates. We divide an arbitrary polarization into orthogonal polarizations normal to each set of plates.

If we substitute n from Eq. (9-34) into Eq. (9-1), we obtain the equation of a front single-surface lens:

$$\rho(\psi) = \frac{(1-n)f}{1 - n\cos\psi} \tag{9-35}$$

Equation (9-35) is an ellipse with f as the distance from the far focus of the ellipse to the center of the lens front surface. This surface refracts waves parallel with the axis and determines the second surface: a plane. The parallel plates constrain the waves parallel with the axis and prevent the design of an outer single-surface lens.

The cutoff wavelength $2a$ and the possibility of higher-order modes restrain the range of n. The second-order-mode cutoff occurs when $\lambda = a$, and it limits n to 0.866 [Eq. (9-34)]. At cutoff, $\lambda = 2a$ and n equals zero. Reasonable values lie between 0.3 and 0.7.

The variation of n versus frequency limits bandwidth. When the phase variation in the aperture is limited to $\lambda/8$, the bandwidth is approximately [1]:

$$\text{bandwidth}(\%) = \frac{25n}{1-n}\frac{\lambda}{(1-n)t} \tag{9-36}$$

where n is the center-frequency index of refraction and t is the maximum thickness. An acceptable bandwidth is underestimated by Eq. (9-36) through restriction of the band edge phase error.

The elliptical surface increases the aperture distribution toward the edges. When we substitute Eq. (9-35) into Eqs. (9-11) and (9-12), we obtain the aperture amplitude distribution relative to the feed pattern:

$$\frac{A(r,\phi)}{F(\psi,\phi)} = \frac{(1-n\cos\psi)^3}{f^2(1-n)^2(\cos\psi - n)} \qquad \text{axisymmetrical}$$

$$\frac{A(r,y)}{F(\psi,y)} = \frac{(1-n\cos\psi)^2}{f(1-n)(\cos\psi - n)} \qquad \text{cylindrical} \tag{9-37a, b}$$

Example Design an axisymmetric parallel-plate lens with a diameter of 30λ, maximum feed angle of 35°, $n = 0.625$, and minimum thickness of λ.

$$\rho(35^\circ) = \frac{30\lambda}{2\sin 35^\circ} = 26.15\lambda$$

Rearrange Eq. (9-35) to calculate the focal length at the edge:

$$f = \frac{(1 - n\cos 35^\circ)\rho(35^\circ)}{1 - n} = 34.03\lambda$$

Equation (9-34) gives us the plate separation with a slight rearrangement:

$$a = \frac{1}{2\sqrt{1 - n^2}} = 0.64\lambda$$

The amplitude variation from the center to the edge caused by the ellipse is given by Eq. (9-37*a*):

$$\frac{A(\psi_e)}{A(0)} = \frac{(1 - n\cos\psi_e)^3}{(1 - n)^2(\cos\psi_e - n)} = 4.26 \quad (6.3\,\text{dB})$$

A feed with its 10-dB beamwidth equal to the subtended angle of the lens at the feed produces −3.7 dB edge taper in the aperture. The edge thickness is given by

$$t = f + 1 - \rho(35^\circ)\cos(35^\circ) = 13.61\lambda$$

Equation (9-36) predicts a bandwidth of 1.9%. A detailed analysis using ray tracing through the lens and varying n with changes in frequency predicts a 0.2-dB loss at this band edge. The 1-dB bandwidth is about 4.5%.

Zoning a parallel-plate lens increases its bandwidth by limiting the maximum thickness, since the variation of the optical path length due to the varying n exceeds that due to the zoning. Figure 9-9 gives the central curve of the three possible types of zoning. The lens in Figure 9-9*a* only suffers loss due to diffractions from edges. The other two zoned lenses (Figure 9-9*b, c*) have dead zones in the aperture. These dead zones produce additional amplitude taper loss and high close-in sidelobes.

The feed-side zoning has different focal lengths in each zone. The outer zone remains the same as the unzoned lens. The focal lengths of the inner zones are reduced by $\lambda/(1 - n)$ at each step and the stepping reduces the amplitude taper of the lens by varying f:

$$\frac{A(\psi)}{A(0)} = \frac{f_c^2(1 - n\cos\psi)^3}{f^2(1 - n)^2(\cos\psi - n)} \tag{9-38}$$

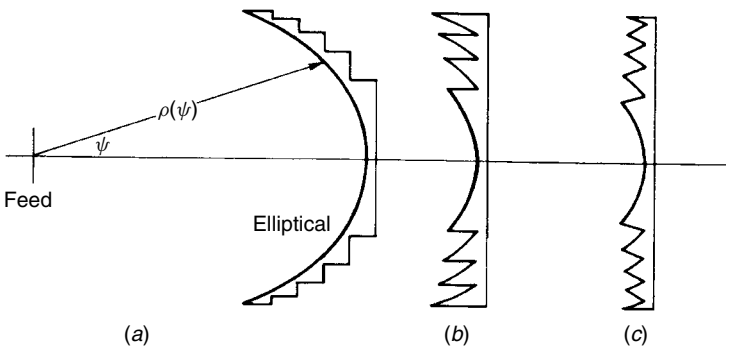


FIGURE 9-9 Central section of zoned parallel metal plate lenses.

where f_c is the central focal length and f is the focal length of the ellipse at ψ. The zoned lens bandwidth is approximately [1]

$$\text{BW} = \frac{25\%}{K - 1 + [(1+n)(1-n)t/n]} \tag{9-39}$$

where K is the number of zones and t is the maximum thickness. Equation (9-39) also is an underestimation of an acceptable loss-level bandwidth.

Example The lens of the example above was zoned as in Figure 9-9*a* and *c* with five zones. The maximum thickness is 3.4λ. Equation (9-39) predicts a bandwidth of 3.4%. By tracing rays through the lens and applying Eq. (4-9) to calculate phase error loss, we predict a 10% bandwidth for a 1-dB loss.

Zoning the nonrefracting surface has no effect on the aperture distribution except for edge diffractions that modify the fields slightly. Zoning the refracting surfaces causes aperture dead zones and reduces the lens-induced amplitude taper. The focal length of the ellipse at the edge remains at 34.03λ. The focal length of the center ellipse is reduced by $4\lambda/(1-n)$ from the edge ellipse to 23.36λ. The edge taper [Eq. (9.38)] becomes 2.01 (3 dB). The aperture dead zones increase the loss by 2 dB.

The bandwidth of a parallel-plate lens can be increased by a method of compounding lenses into a doublet [5]. We can make a lens by using a uniform waveguide length between the input and output surfaces and placing a phase shifter in each line to compensate for the optical path-length differences and produce an eikonal at the aperture plane. If we combine a refracting surface waveguide plate lens with a differential phase shift lens, we can match the aperture phase at two frequencies. This matching broadbands the antenna like an optical achromatic doublet.

9-6 SURFACE MISMATCH AND DIELECTRIC LOSSES

The reflection coefficient of a wave normally incident on a dielectric is

$$\Gamma = \frac{1-n}{1+n} \tag{9-40}$$

valid for both dielectric and metal lenses. The actual reflection coefficient of any ray depends on the angle of incidence and the polarization. Both surfaces of the lens have reflections, and these interact to produce the actual reflection. With plane surfaces, such as those assumed for radomes, we can analyze the combination of reflections using transmission-line mismatch equivalence. Since the reflections from second surfaces may not return to the same point as the incident waves and may have their caustic distances changed by the curved surfaces, transmission-line models of the two surfaces have limited use for lenses. Equation (9-40) gives us a reasonable approximation to expected mismatch, since one surface of the single refracting surface lens will be normal to the wave incident from the feed and reflect into the feed. The second surface fails to focus power back to the feed and has a minor effect.

Example A lens with $n = 1.6$ or $n = 0.625$ has a reflection coefficient magnitude of 0.23 [Eq. (9-40)] at one surface that focuses into the feed. This gives a feed mismatch if

$$\text{VSWR} = \frac{1 + 0.23}{1 - 0.23} = 1.6$$

the same as n ($n > 1$) or $1/n$ ($n < 1$). The mismatch loss becomes $1 - |\Gamma|^2 = 0.95$ (0.2 dB).

The wave may be matched to the surface by a quarter-wavelength transformer with an index of refraction $n^{1/2}$, but adding matching transformers narrows the bandwidth. The surface, which reflects power back to the feed, should be matched first, since the second surface has a minor effect on feed mismatch. Also, the primary reflecting surface has normally incident waves and does not suffer from the need to vary the thickness to match waves off normal incidence. Transformers to match waves off normal incidence are polarization-sensitive.

Simple methods can be used to reduce lens-caused feed mismatch [2]. The lens can be tilted to cause the reflection to miss the feed. Offsetting half the lens by $\lambda/4$ causes cancellation of the reflection from the two halves. Tilting does not reduce the mismatch loss but does produce backlobe power in the pattern. Similarly, the reflected power from the hyperbolical surface forms backlobes. These reflections reduce the antenna efficiency below that predicted by aperture theory alone.

Cohn and Morita [2,6] developed methods of matching the surface of the lenses by removing some of the dielectric for a quarter wavelength. The surfaces are either corrugated, have arrays of holes, or have arrays of rods. With this method, the lens can be made from a single dielectric slab. The design depends on the angle of incidence and the polarization of the waves. The lens dissipates power by the attenuation constant of the material:

$$\alpha = \frac{27.3n \tan \delta}{\lambda} \qquad \text{dB/length} \tag{9-41}$$

where $\tan \delta$ is the loss tangent of the dielectric. Waveguide losses reduce the power transmitted through metal plate lenses. Zoning eliminates material and its associated loss to improve efficiency, but for most materials, this effect is small.

Artificial dielectrics [2] reduce excessive weight and material losses of lenses. We make them by embedding metal particles or plated microspheres in foam with a dielectric constant near 1. The metal parts may be strips or disks made from metal foil. Similarly, solid metal parts can be hollow. Since the effective dielectric constant depends on the size of the metal particles in wavelengths, lenses made from artificial dielectrics will be narrowband if the particles are large, but the use of plated microspheres dispersed in the foam reduces this problem.

9-7 FEED SCANNING OF A HYPERBOLOIDAL LENS [7]

The hyperboloidal lens has no cross-polarization when fed from an electric dipole source. Kreutel [7] analyzed the effects of off-axis dipole sources on the pattern of the hyperboloidal lens. The coma increases more rapidly for the lens than for a paraboloidal reflector for the same scanning. Like the paraboloidal reflector, the hyperboloidal lens beam scans less than the deviation angle of the feed relative to the vertex and axis and has a beam deviation factor (Table 9-5). The scanning loss (Table 9-6) decreases with increasing n. The peak coma lobe (Table 9-7) limits the possible scan before unusable patterns are obtained. The paraboloidal reflector can be scanned further (Table 8-2) for the same coma.

TABLE 9-5 Beam Deviation Factor for a Feed-Scanned Hyperboloidal Lens

f/D	$n=\sqrt{2}$	$n=2$	f/D	$n=\sqrt{2}$	$n=2$	f/D	$n=\sqrt{2}$	$n=2$
0.8	0.75	0.84	1.4	0.86	0.91	2.0	0.93	0.95
1.0	0.80	0.87	1.6	0.89	0.92	2.5	0.95	0.96
1.2	0.83	0.89	1.8	0.92	0.94	3.0	0.97	0.98

TABLE 9-6 Scanning Loss for a Hyperboloidal Lens (dB)

Beamwidth of Scan	$n=\sqrt{2}$		$n=2$	
	$f/D=1$	$f/D=2$	$f/D=1$	$f/D=2$
0.5	0.03	0.00	0.01	0.00
1.0	0.06	0.01	0.04	0.01
1.5	0.12	0.03	0.07	0.02
2.0	0.23	0.05	0.12	0.04
2.5	0.36	0.09	0.20	0.06
3.0	0.51	0.13	0.28	0.08
3.5	0.69	0.19	0.37	0.11
4.0	0.90	0.24	0.49	0.14
4.5	1.09	0.31	0.60	0.18
5.0		0.38	0.75	0.22
5.5		0.44		0.25

TABLE 9-7 Coma Sidelobe Level for a Scanned Hyperboloidal Lens (dB)

Beamwidth of Scan	$n=\sqrt{2}$		$n=2$
	$f/D=1$	$f/D=2$	$f/D=1$
0	20.9	18.5	19.5
1	17.6	17.5	17.6
2	15.2	16.5	15.7
3	13.1	15.5	14.2
4	11.3	14.8	12.9
5	9.8	14.0	11.5
6	8.8	13.3	10.4

9-8 DUAL-SURFACE LENSES

The second surface of the lens offers an additional degree of freedom that can be used to control the pattern characteristics. Ruze [8] developed methods to reduce coma for feed-scanned cylindrically shaped metal plate lenses, which constrain the wave parallel with the axis. Both surfaces are used to satisfy focusing requirements. We will develop a method for axisymmetrical dielectric lenses to eliminate coma for small feed

displacements. In a second design we can also use the second surface shape to control the amplitude distribution in the aperture plane.

9-8.1 Coma-Free Axisymmetric Dielectric Lens [9]

The design of the coma-free axisymmetric antenna reduces to the numerical solution of a differential equation with side conditions to produce a collimated beam and satisfy the Abbe sine condition [10, p. 157]. Successful designs require a number of iterations, since the ultimate lens shape depends heavily on the initial conditions. Solution of the differential equation will sometimes diverge into unrealizable designs or fail to continue to satisfy the side conditions.

A lens satisfying the Abbe sine condition is free of coma aberrations for small deviations of the feed from the axis. The deviations produce higher-order aberrations that eventually distort the beam with continued scanning, but coma is removed. For a lens focused at infinity, the Abbe sine condition requires that the surface which refracts the waves parallel with the axis must be spherical with its center on the effective focus of the lens. The dielectric lens refracts waves parallel with the axis on the outer surface (away from the feed). Given the aperture radial component r,

$$r = f_e \sin \psi \tag{9-42}$$

where f_e is the effective focus and ψ is the feed angle. A waveguide plate lens satisfies the Abbe sine condition by having a spherical inner surface [6], since the waves are parallel with the axis in the lens because the waveguide plates constrain the wave to be parallel with the axis.

The second surface must produce the conditions for a uniform phase in the aperture plane. The waveguide plate lens only has to equalize path lengths. In a dielectric lens, the inner surface must refract the waves in the proper direction to satisfy the Abbe sine condition, and it must be so placed as to equalize the path lengths from the feed to the aperture plane. The locations of both surfaces along the axis are varied to equalize the path lengths.

Figure 9-10 shows the coordinates of the coma-free dielectric lens. The polar equation $\rho\ (\psi)$ describes the inner surface and ψ' is the angle of the refracted wave with the axis. The distance from the feed to the center of the lens inner surface is f, and T is the thickness. The coordinates (r, z) describe the outer lens surface, where z is the axis dimension and r is the aperture radial component. Snell's law reduces to a differential equation at the inner surface:

$$\frac{d\rho}{d\psi} = \frac{n \sin(\psi - \psi')\rho}{n \cos(\psi - \psi') - 1} \tag{9-43}$$

where

$$\tan \psi' = \frac{r - \rho \sin \psi}{z - \rho \cos \psi} \tag{9-44}$$

By use of Eq. (9-42), this reduces to

$$\tan \psi' = \frac{(f_e - \rho) \sin \psi}{z - \rho \cos \psi} \tag{9-45}$$

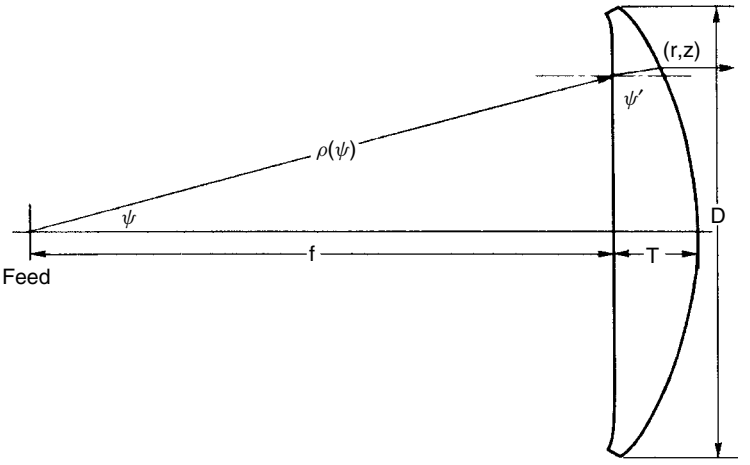


FIGURE 9-10 Coma-corrected dual-surface axisymmetric lens; $n = 1.6$, $D = 35$, $f = 45$, $T = 6.5$, $f_e = 49$.

The requirement for equal optical path lengths to the aperture determines a quadratic equation in z:

$$Az^2 + Bz + C = 0 \tag{9-46}$$

where

$$A = n^2 - 1$$
$$B = 2(\rho - K) - 2n^2\rho\cos\psi$$
$$C = n^2\rho^2\cos^2\psi + n^2(f_e - \rho)^2\sin^2\psi - (\rho - K)^2$$
$$K = T(n-1)$$
$$z = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$$

Design consists of the numerical solution of the differential equation (9-43) subject to the conditions of Eqs. (9-45) and (9-46). Realizable solutions depend on the initial conditions. Most failures to produce a usable design occur in Eq. (9-46), which satisfies the requirement of equal aperture phase.

Example Figure 9-10 shows a scale drawing of a realizable design for $n = 1.6$, focal distance $f = 45$, diameter $D = 35$, center thickness $T = 6.5$, and effective focal length $f_e = 49$. Table 9-8 lists a few points of the solution obtained by a Runge–Kutta numerical method for the differential equation (9-43).

The example above contains only relative dimensions. The solution is size and frequency independent, since it is obtained by geometric optics. We can zone the lens for a given frequency by cutting along ray paths. Each step is $\lambda/(n-1)$. Table 9-8, as in the example above, determines the ray paths through the lens. Zoning will produce either feed or aperture dead zones. The reduction of weight must be balanced with

TABLE 9-8 Design of Figure 9-11 for a Coma-Free Lens

Feed Angle, ψ (deg)	Inner Surface, $\rho(\psi)$	Horizontal Distance, z	Radius, r	Thickness Along Ray, T
0	45.00	51.50	0	6.50
5	45.18	51.19	4.27	6.20
10	45.70	50.27	8.51	5.29
15	46.59	48.71	12.68	3.77
20	47.63	46.28	16.76	1.59
20.92	47.61	45.40	17.50	1.06

the loss in efficiency to achieve some compromise. Because we used the degrees of freedom of the second surface to satisfy the Abbe sine condition, we lose control of the aperture distribution through the lens surfaces. Most practical designs produce an amplitude taper near that of the feed antenna. We must achieve low sidelobes, if required, through a tapered illumination from the feed. The feed pattern plays no part in the design and gives us degrees of freedom for amplitude taper. An antenna designed and built with a diameter of 32 wavelengths [9] showed no coma in a scanning of ± 2 beamwidths. Like the paraboloidal reflectors, increasing the focal length for a given diameter allows greater scanning without significant coma.

9-8.2 Specified Aperture Distribution Axisymmetric Dielectric Lens [11]

We use the desired aperture amplitude distribution to specify the relation between the aperture radius r, and the feed angle ψ. Earlier, the Abbe sine condition established this relation. Given a feed power pattern $F(\psi)$ and a required aperture distribution $A(r)$, we relate the two through differential areas:

$$F(\psi) \sin \psi \, d\psi = A(r) r \, dr$$

where an axisymmetrical pattern is assumed. We derive the relation between ψ and r through normalized integrals as in Section 8-20:

$$\frac{\displaystyle\int_0^{\psi} F(\psi) \sin \psi \, d\psi}{\displaystyle\int_0^{\psi_m} F(\psi) \sin \psi \, d\psi} = \frac{\displaystyle\int_0^{r} A(r) r \, dr}{\displaystyle\int_0^{r_m} A(r) r \, dr} \tag{9-47}$$

In any particular design we generate a table, such as Table 8-12, of the feed angle versus its normalized feed pattern integral and the aperture radius versus its normalized aperture distribution integral. For a given feed angle ψ we equate the normalized integrals to compute the corresponding aperture radius. We generate a table of aperture radius versus feed angle, such as Table 8-13, using interpolation techniques. The design is very dependent on feed pattern, because changing the feed pattern alters the table. Low sidelobe aperture distributions require close tolerances and a good specification of the feed pattern.

Once we have the relation between ψ and r, the design follows steps similar to those taken in designing the lens for the Abbe sine condition. We solve the differential

equation (9-43) numerically. We specify the aperture radius by the table generated for the transformation of the feed pattern into the aperture distribution. The requirement for equal path length through the lens determines the axis location z of the outer surface.

$$
\begin{aligned}
&Ax^2 + Bx + C = 0 \qquad x = z - f \\
&A = n^2 - 1 \\
&B = 2[n^2(f - \rho\cos\psi)] + \rho - K - f \\
&C = [n(\rho\cos\psi - f)]^2 + (r - \rho\sin\psi)^2 - (f + K - \rho)^2 \\
&K = T(n-1) \qquad z = f + \frac{-B + \sqrt{B^2 - 4AC}}{4A}
\end{aligned}
\tag{9-48}
$$

where T is the central thickness and f is the axis focal length.

A successful design requires a number of iterations, starting with different initial conditions. The differential equation solution will diverge to shapes unable to satisfy the equal-path-length side requirement with poor initial conditions. Each design has a narrow range of satisfactory initial conditions. In most cases, increasing the thickness increases the chance for a successful design.

Example A lens was designed to transform the feed pattern of a conical corrugated horn into a circular Taylor distribution with 40-dB sidelobes ($\overline{n} = 8$). The initial conditions were $n = 1.6$, diameter $D = 32$, focal distance $f = 35$, central thickness $T = 9$, and maximum feed angle $\psi_m = 20^\circ$. A table of the feed angle and its normalized power pattern integral, along with the aperture radius and its normalized power distribution integral, was generated. A table of feed angle and the corresponding aperture radius follows from equating normalized integrals. This table is independent of the lens thickness but not the feed pattern.

A Runge–Kutta numerical method is used to solve the differential equation (9-43) subject to the conditions imposed by the aperture radius table and Eq. (9-48) for equal optical path lengths. Figure 9-11 shows a design for a 36° 10-dB beamwidth feed horn (12-dB feed edge taper). The horn dimensions are aperture radius $= 1.90\lambda$ and slant radius $= 9\lambda$ for a maximum quadratic phase deviation $S = 0.2$. A few points of the design are listed in Table 9-9. Other antennas designed with small changes in the feed pattern beamwidth show significant changes in the lens shape near the edges for a constant center thickness. Axisymmetric dielectric lenses can be designed to be independent of frequency because only relative sizes are specified. The lenses tend to be thick to allow room to satisfy the requirement for an equal optical path length.

Zoning lens reduces weight while decreasing bandwidth. Low-sidelobe designs are inherently narrowband, since small changes in the beamwidth of the feed alter the aperture distribution and sidelobe levels. Zoning may not reduce the bandwidth significantly. An antenna designed and tested using the technique above revealed a number of requirements on the design [12]. The sidelobe levels exceeded the design specification for three main reasons. First, the feed pattern was specified as $\sin(\pi U)/\pi U$, an oversimplification of the actual feed pattern. Realistic feed patterns must be used because small changes in the feed pattern require new designs. Second, the surfaces must be matched with quarter-wavelength sections to prevent reflections, unaccounted

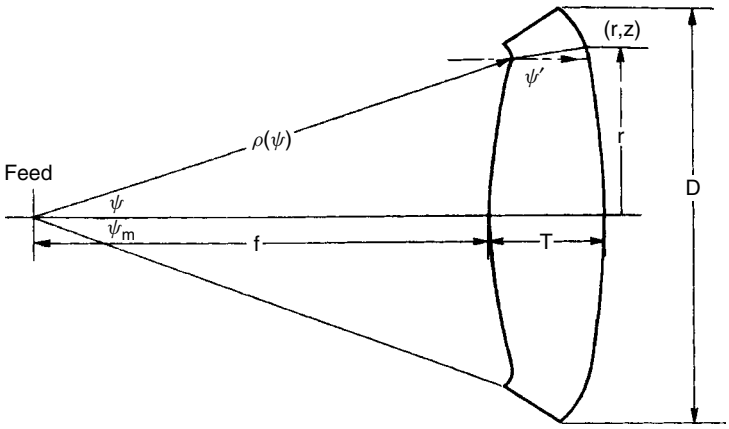


FIGURE 9-11 Dual-surface axisymmetric lens for circular Taylor aperture distribution (40-dB, $\hat{n} = 8$). Lens: $n = 1.6$, $D = 32$, $f = 35$, $T = 9$, $\psi_m = 20°$. Feed: conical corrugated horn, 36° 10-dB beamwidth. $S = 0.2$.

TABLE 9-9 Design of Figure 9-12 for a Specified Aperture Distribution Axisymmetric Dielectric Lens

Feed Angle, ψ (deg)	Inner Surface, $\rho(\psi)$	Horizontal Distance, z	Radius, r	Thickness Along Ray, T
0	35.00	44.00	0	9.00
5	35.31	43.95	3.27	8.78
10	36.14	43.67	6.59	8.08
15	37.69	43.25	10.11	6.86
20	38.71	40.58	16.00	5.03

for in the design, that change the aperture distribution. Third, diffractions from the edges affect the distribution. Increasing the aperture diameter or using a low edge illumination feed reduces these effects.

As designed, the lens exhibits severe coma when scanned by feed lateral offset. Since most lenses are quite thick, zoning can be used to approximate the Abbe sine condition on the inner surface. The lens refracts most of the rays parallel with the axis by the inner surface. If the zones approximate a spherical surface on the average, coma is reduced for the scanned beams. These coma-corrected lenses are useful for multibeam applications when each beam is fed from an offset feed.

9-9 BOOTLACE LENS

The bootlace lens consists of a set of receiving antennas on a surface connected by cables to a set of transmitting antennas on the second surface (Figure 9-12). The cables constrain the path through the lens. We have three degrees of freedom with this lens: (1) input surface, (2) output surface, and (3) cable length. We can change the lens characteristics dynamically by placing phase shifters and/or attenuators in the lines

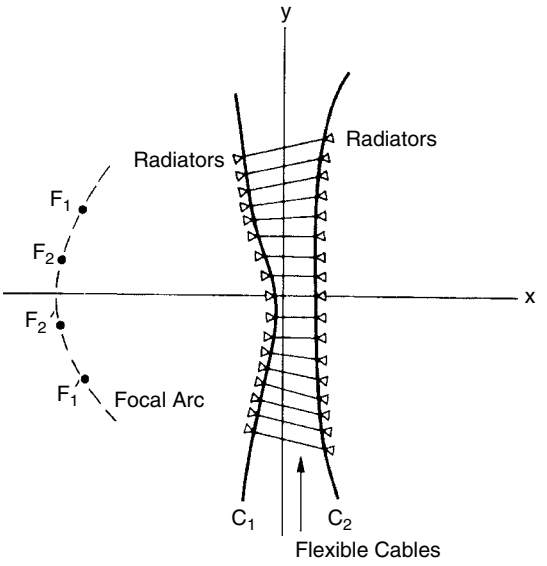


FIGURE 9-12 Bootlace lens. (From [13], Fig. 3, ©1965 IEEE.)

between the input and output radiators and scan one or more beams. The input and output surfaces are arrays, and we can generate multiple beams by placing more than one feed on a focal arc determined by the lens geometry.

The simplest bootlace lens consists of a spherical input surface connected to a plane output surface by equal-length cables. This lens converts spherical waves radiated by the feed into plane waves at the output surface. The lens uses true time delay, which removes bandwidth limitations. Most bootlace lenses are line sources or two-dimensional lenses fed by line sources. The general lens can have four focal points [13] placed symmetrically about the axis of the symmetric structure. The focal arc is chosen on a curve through the focal points to minimize defocusing when feeds are placed off the focal points. A feed at each point along the focal arc produces an output beam in a different direction. Because each feed uses the full aperture, it achieves the full array gain less the loss of projection of the aperture length in the beam direction.

The number of focal points is reduced to three when the lens is further restrained. The Ruze design for a metal plate lens [8] has three focal points, since the waveguides between the surfaces travel in straight lines. There is one central axis focal point and two symmetrically placed focal points. The Rotman [14] lens loses one possible focal point because the output surface is limited to a straight line. A parallel-plate structure in the Rotman lens leads from the possible feed locations to the feed-side surface, which normally is excited by probes in the parallel-plate guide. The lens becomes a feed network that produces multiple beams whose directions depend on the location of the feed on the focal arc. Although perfect focusing is achieved at only three points, the phase error loss associated with points between them is small. Because the Rotman lens feed network is a true time-delay array feed network, we can achieve bandwidths greater than an octave from it.

Rao [15] extends the design of bootlace lenses to three dimensions and shows that the number of focal points cannot be extended beyond four. Because the lens is not

axisymmetric, it has different scanning capabilities in orthogonal planes. Rao designs lenses with two, three, and four focal points on a focal line. Decreasing the number of focal points in one plane increases the scanning capability in the orthogonal plane for a given phase error level.

9-10 LUNEBURG LENS [16, p. 545]

A Luneburg lens, a spherically symmetric lens with a variable index of refraction, radiates a beam in any direction for a feed located opposite the beam. We place the feed phase center either on the surface of the lens or a short distance away. We form multiple beams by feeding the lens at a number of places. Our only restriction is the blockage due to other feeds or support structures. We can rapidly scan a beam by moving a lightweight feed around the sphere or by switching between multiple feeds.

When we place the feed on the outer surface of the sphere, the required index of refraction is

$$n = \sqrt{2 - \left(\frac{r}{a}\right)^2} \tag{9-49}$$

where a is the lens outer radius and r is the inner radius. The dielectric constant n^2 must vary between 2 in the center and 1 on the outer surface. Few feeds have their phase centers on a surface that can be mounted against a sphere. We can move the feed away from the surface by changing the variation of the index of refraction from that given by Eq. (9-49), but the required center index of refraction decreases as we move the feed away from the lens surface. We calculate the variation of the index of refraction from an integral equation, and the curves follow the general shape of Eq. (9-49). For the feed-to-sphere radius of 1.1, the proper center dielectric constant is 1.83 and it varies smoothly to 1 at the lens surface. Similarly, the center dielectric constant starts at 1.68 for the feed-to-sphere radius of 1.2.

The lens changes the amplitude distribution in the aperture compared with the feed. Given the ratio of feed radius to lens radius r_i, the aperture plane power distribution becomes

$$A(r) = \frac{F(\psi)}{r_1^2 \cos \psi} \tag{9-50}$$

where ψ is the feed angle, $F(\psi)$ the feed power pattern, and $A(r)$ the aperture power distribution. Equation (9-50) shows that the lens refracts power toward the edge of the aperture. Lenses have been made by using a series of concentric spherical shells each with a constant dielectric. A minimum of 10 shells is needed for an adequate approximation of the required variation of the dielectric constant.

REFERENCES

1. J. R. Risser, Chapter 11 in S. Silver, ed., *Microwave Antenna Theory and Design*, McGraw-Hill, New York, 1948.
2. S. B. Cohn, Chapter 14 in H. Jasik, ed., *Antenna Engineering Handbook*, McGraw-Hill, New York, 1961.
3. S. Cornbleet, *Microwave Optics*, Academic Press, London, 1976.

4. B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, Wiley, New York, 1991.
5. A. R. Dion, A broadband compound waveguide lens, *IEEE Transactions on Antennas and Propagation*, vol. AP-26, no. 5, September 1978, pp. 751–755.
6. T. Morita and S. B. Cohn, Microwave lens matching by simulated quarter-wave transformers, *IEEE Transactions on Antennas and Propagation*, vol. AP-4, no. 1, January 1956, pp. 33–39.
7. R. W. Kreutel, The hyperboloidal lens with laterally displaced dipole feed, *IEEE Transactions on Antennas and Propagation*, vol. AP-28, no. 4, July 1980, pp. 443–450.
8. J. Ruze, Wide-angle metal-plate optics, *Proceedings of IRE*, vol. 38, no. 1, January 1950, pp. 53–59.
9. J. J. Lee, Numerical methods make lens antennas practical, *Microwaves*, vol. 21, no. 9, September 1982, pp. 81–84.
10. R. Kingslake, *Lens Design Fundamentals*, Academic Press, New York, 1978.
11. J. J. Lee, Dielectric lens shaping and coma-correcting zoning, part I: analysis, *IEEE Transactions on Antennas and Propagation*, vol. AP-31, no. 1, January 1983, pp. 211–216.
12. J. J. Lee and R. L. Carlise, A coma-corrected multibeam shaped lens antenna, part II: experiments, *IEEE Transactions on Antennas and Propagation* vol. AP-31, no. 1, January 1983, pp. 216–220.
13. M. L. Kales and R. M. Brown, Design considerations for two dimensional symmetric bootlace lenses, *IEEE Transactions on Antennas and Propagation*, vol. AP-13, no. 4, July 1965, pp. 521–528.
14. W. Rotman and R. F. Turner, Wide-angle microwave lens for line source applications, *IEEE Transactions on Antennas and Propagation*, vol. AP-11, no. 6, November 1963, pp. 623–632.
15. J. B. L. Rao, Multifocal three-dimensional bootlace lenses, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 6, November 1982, pp. 1050–1056.
16. R. S. Elliott, *Antenna Theory and Design*, Prentice-Hall, Englewood Cliffs, NJ, 1981.

10

TRAVELING-WAVE ANTENNAS

Traveling-wave antennas consist of transmission-line structures that radiate. We develop a unified theory for end-fire line antennas because length and propagation constant along the structure determine most of their properties. To first order, length determines gain and bandwidth. The size and shape of the structure produce secondary effects such as polarization nulls and narrower beamwidths. Most of these structures are slow wave-transmission structures that bind waves to it and radiate at discontinuities. We use surface-wave structures to radiate end-fire beams and leaky wave structures to radiate beams at an angle to the axis of the line source. In both cases there are planar configurations that have their uses, but in this chapter we concentrate on long, thin geometries. We combine leaky wave line-source radiators, such as slotted rectangular waveguides, into planar arrays, but the line source remains the building block.

We make traveling-wave antennas from structures that guide waves. Surface-wave structures bind the power to the transmission line and radiate from discontinuities such as bends or dimensional changes. In some cases we analyze the surface wave as radiating throughout its extent on the transmission line. Both methods provide insight. Leaky wave antennas carry waves internally, such as a waveguide, and radiate at openings that allow power to escape. The radiation mechanism differs in the two cases, but we use similar mathematics to describe both types. We may have trouble distinguishing the radiation mode because the structures may be similar because with small changes in structure, some antennas can radiate in either mode. We separate traveling-wave antennas from other antennas by the presence of a wave traveling along the structure, with most of its power propagating in a single direction.

We divide antennas by their structure: line and planar. We usually analyze planar structures as being infinite in the direction normal to the wave propagation. Similarly, we usually ignore the diameter of line sources in a first-order analysis. The diameter is important for determining the mode structure, but to first order we calculate patterns

Modern Antenna Design, Second Edition, By Thomas A. Milligan

based on a thin line source since length and propagation constant determine the pattern and bandwidth. The width of a planar structure determines the pattern beamwidth in that plane. Increasing the diameters of the rods of line sources will decrease the pattern beamwidth and increase gain, but the effect is secondary. Only when we include the diameter can we make the transition to aperture-type structures considered to be radiating from the end.

In this chapter we must consider unusual transmission-line structures. Properly designed dimensions provide the proper phase velocity to establish a single end-fire beam for slow-wave antennas or to point the beam of a leaky wave antenna. We calculate some of the dimensions by analysis (an ever-expanding list), but we can also measure the velocities and leakage and proceed to design.

10-1 GENERAL TRAVELING WAVES

A wave traveling in a single direction has a field representation:

$$E = E_0(z)e^{-kPz} \tag{10-1}$$

where z is the direction of propagation, k the free-space propagation constant (wave number) $2\pi/\lambda$, and P the relative propagation constant. $E_0(z)$ describes the amplitude variation:

$$\begin{aligned} P &> 1 \quad \text{surface waves} \\ P &< 1 \quad \text{leaky waves} \end{aligned} \tag{10-2}$$

For a planar structure in the $y-z$ plane, we consider separable distributions:

$$E = E_0(z)E_1(y)e^{-jkPz}$$

We compute the pattern from

$$f = \int_0^L \int_{-a}^{a} E_0(z)E_1(y)e^{-jkPz}e^{jk_z z}e^{jk_y y}dz\,dy \tag{10-3}$$

where $k_z = k\cos\theta$. Similarly, for circular distributions we have

$$E = E_0(z)E_1(\phi)e^{-jkPz}$$

and

$$f = \int_0^L E_0(z)e^{-jkPz}e^{jk_z z}dz \int_0^{2\pi} E_1(\phi_c)ae^{jka\sin\theta\cos(\phi-\phi_c)}d\phi_c \tag{10-4}$$

where a is the radius. The second integral includes vector dot products to project the ring aperture fields on to the far-field polarizations (see Section 7-2). We consider only the term along the z-axis, and we can consider the effect of the other coordinate separately. The pattern response is

$$f = \int_0^L E_0(z)e^{-jk_z(P-\cos\theta)}dz \tag{10-5}$$

We use the results of Chapter 4 with these separable distributions. Maximum gain comes from a uniform distribution reduced by the amplitude taper efficiency for tapered distributions. In Eq. (10-3) the y-axis distribution and size determine the gain factor as a product for aperture area. Equation (10-4) has a separable ϕ distribution that separates directivity into a product. We ignore these factors for now and concentrate on the z-axis pattern and associated directivity. Linear-rod antennas have increased directivity because of dipole ϕ distributions and their increased radius.

A traveling wave with a uniform distribution has pattern response

$$\frac{\sin(\psi/2)}{\psi/2} \qquad \text{where} \quad \psi = kL(P - \cos\theta) \tag{10-6}$$

for θ measured from the z-axis. The y or ϕ distribution determines the pattern in the other coordinate. For $P > 1$, a slow wave, the beam peak approaches $\theta = 0$ when $P \to 1$. The length bounds the range P for an end-fire pattern peak. Leaky waves, $P < 1$, have a pattern peak when $P = \cos\theta$, or

$$\theta_{\max} = \cos^{-1} P \tag{10-7}$$

The pattern peak approaches end fire ($\theta = 0$) as $P \to 1$. By increasing P beyond 1, the directivity increases and reaches maximum value for a given P, depending on the length [1]

$$P = 1 + \frac{0.465}{L} \tag{10-8}$$

Equation (10-8) is the Hansen and Woodyard criterion for increased directivity of a long end-fire structure commonly approximated by [2]

$$P = 1 + \frac{1}{2L} \tag{10-9}$$

The phase increase of 180° [Eq. (10-9)] along the length gives the maximum directivity for a long structure with a uniform distribution. The amplitude distribution for most surface-wave devices ($P > 1$) peaks near the input and the taper reduces the gain by the amplitude taper efficiency [Eq. (4-8)]. We reduce the relative propagation constant from that given by Eq. (10-8) depending on the length [3]:

$$P = 1 + \frac{1}{RL} \tag{10-10}$$

where $R = 6$ at $L = \lambda$, diminishing to 3 from $L = 3\lambda$ to $L = 8\lambda$ and tapering to 2 [Eq. (10-9)] at $L = 20\lambda$. Zucker [4] uses $R = 6$ for the amplitude, which peaks by 3 dB at the input for all lengths. Equations (10-8) and (10-10) give designs with only small differences in gain.

The value of P controls one edge of the visible region. Setting $P = 0$ centers the visible region about $\psi = 0$. End fire occurs at $P = 1$. As we increase P beyond 1, the beam peak of the distribution in ψ space moves into invisible space and the sidelobe level increases. A progression of the distribution sidelobes becomes beam peaks as P increases. Since the amplitude difference between sidelobes decreases as the sidelobe

number increases, the sidelobe level of the pattern increases as the pattern degrades because P exceeds the value given by Eq. (10-9).

Figure 10-1 shows the effects on directivity of varying P on an axisymmetrical traveling-wave antenna with a uniform amplitude distribution. For broadside radiation ($P = 0$) and for P near end fire, the directivity is a constant value with scan:

$$\text{directivity} = \frac{2L}{\lambda} \tag{10-11}$$

When the broadside conical beam is scanned until the cone joins into a single end-fire beam, the directivity increases. For end fire, $P = 1$:

$$\text{directivity} = \frac{4L}{\lambda} \quad \text{end fire} \tag{10-12}$$

The directivity peaks for P given by Eq. (10-8):

$$\text{directivity} = \frac{SL}{\lambda} \tag{10-13}$$

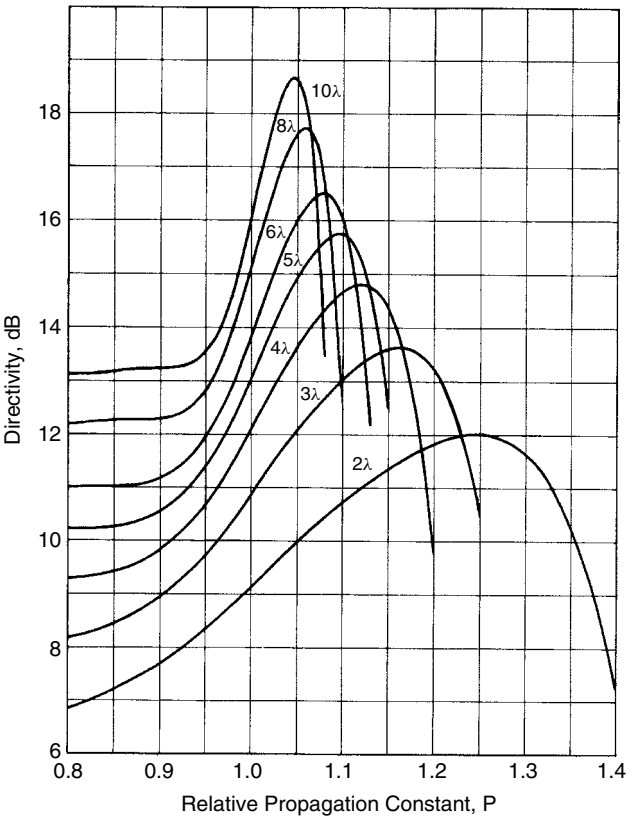


FIGURE 10-1 Directivity of an axisymmetrical uniform-distribution traveling wave.

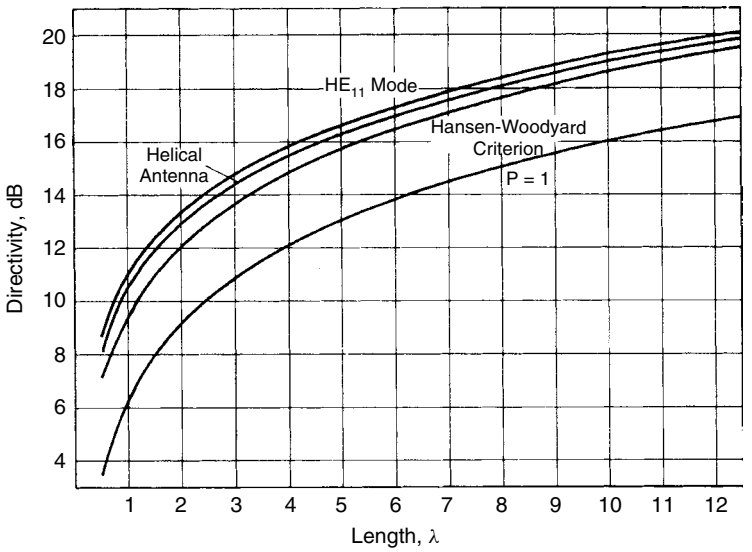


FIGURE 10-2 Directivity of an end-fire traveling-wave antenna.

L/λ	2	4	6	10	20
S	7.92	7.58	7.45	7.33	7.25

Figure 10-2 plots the maximum directivity of end-fire structures versus length. For the case $P = 1$, Eq. (10-12) gives the directivity on the curve. The Hansen and Woodyard criterion increases the directivity as shown in Figure 10-2 for an infinitesimal-diameter structure. The distribution on the finite diameter of the helical wire antenna in the axial mode increases the directivity over that for the Hansen and Woodyard increased directivity criterion. The hybrid mode with its linear polarization has a dipole null normal to the traveling-wave axis whose elemental pattern also increases directivity. Figure 10-3 is a plot of the corresponding beamwidths of those structures. Figure 10-2 sets an upper bound to the possible directivity of a small-diameter end-fire traveling-wave structure of given length.

10-1.1 Slow Wave

A slow wave exists on an open transmission-line structure that binds the wave by slowing a passing wave and bending it in the direction of the structure. In the same manner, a lens bends waves toward regions of higher index of refraction (increased slowing). We designate x as the direction normal to a planar structure and the radial coordinate ρ as the direction normal to the cylindrical slow-wave structure. The relation between propagation constants in various directions is found in any electromagnetics text [5]:

$$k_z^2 + k_x^2 = k^2 \quad \text{or} \quad k_z^2 + k_\rho^2 = k^2 \tag{10-14}$$

Since x (or ρ) is unbounded, the waves must attenuate exponentially from the surface:

$$\alpha = jk_x \quad \text{or} \quad \alpha = jk_\rho \tag{10-15}$$

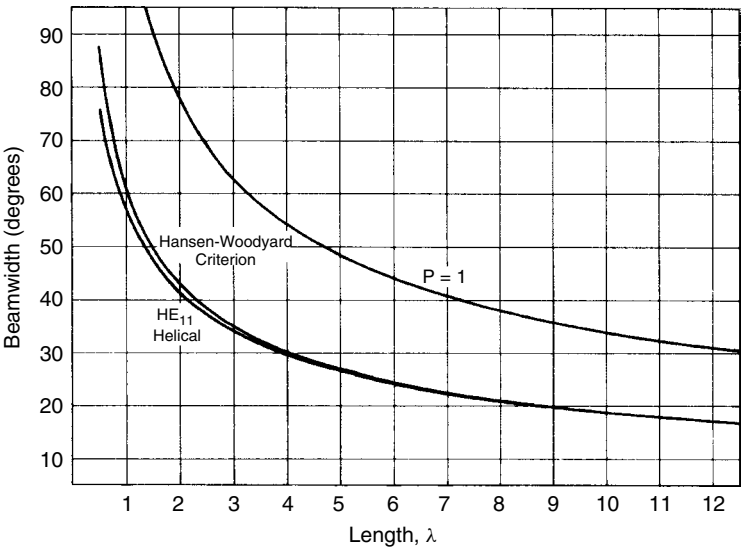


FIGURE 10-3 Beamwidth of a traveling-wave end-fire antenna.

The z-directed propagation constant becomes

$$k_z^2 = k^2 + \alpha^2 = P^2 k^2$$

where

$$P = \sqrt{1 + \frac{\alpha^2}{k^2}} = \sqrt{1 + \left(\frac{\lambda\alpha}{2\pi}\right)^2} \tag{10-16}$$

P, the relative propagation constant, becomes a measure of the wave binding to the surface. We rearrange Eq. (10-16):

$$\alpha = \frac{2\pi}{\lambda}\sqrt{P^2 - 1} \text{ (Np/}\lambda\text{)} = 8.63\frac{2\pi}{\lambda}\sqrt{P^2 - 1} \text{ (dB/}\lambda\text{)}$$

As P increases, the wave is more tightly bound to the surface. Figure 10-4 is a plot of the distances normal to the surface of constant-field contours versus P. The fields attenuate rapidly normal to the surface. For $P \to 1$, the slow-wave structure only diffracts passing plane waves without capturing the power. This is the sense of a cutoff frequency for the structure.

Most surface-wave antennas consist of three regions. The feed region launches the wave on the structure with P between 1.2 and 1.3 [4]. The structure tapers in a short section until P suitable for the length is reached. We design for a given phase shift along the entire length. For example, a long antenna would be designed so that the wave on the structure has an excess phase shift of 180° [Eq. (10-9)] over the traveling wave in free space. Near the end we sometimes taper the structure to reduce the end reflection given approximately by [4] $P^2 - 1$ (power). This end taper can be quite short and achieve good results.

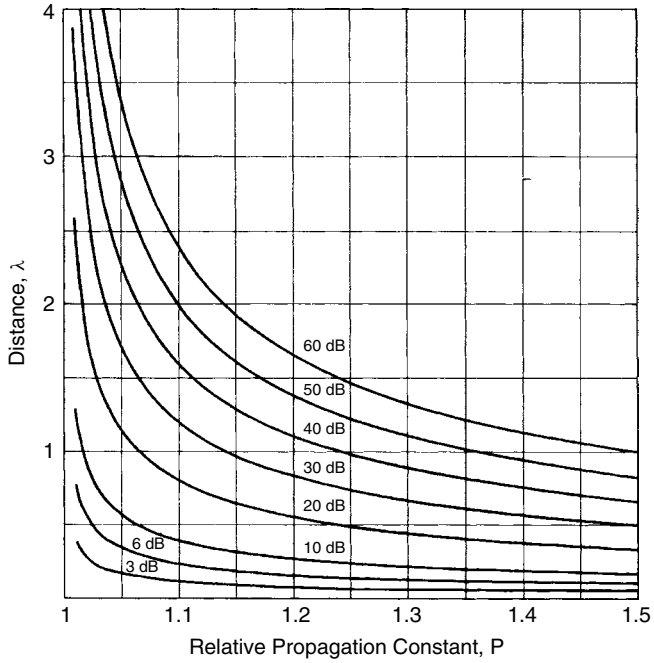


FIGURE 10-4 Constant field contours off the surface of a surface-wave structure.

10-1.2 Fast Waves (Leaky Wave Structure)

Only closed structures such as waveguides support fast waves. An open structure requires a negative α [Eq. (10-16)] for fast waves, which implies an exponentially increasing wave away from the structure. The structure soon radiates all its power and no longer guides the wave. By limiting α, the leakage, we can extend the length of the radiating structure. We include the attenuation due to leakage in the z-axis propagation constant, in general, but we ignore it when α is small: $\theta_{\max} = \cos^{-1}(k_z/k)$.

Example A rectangular waveguide has $k_x = 2\pi/2a$ for the TE_{10} mode, where a is the guide width. We determine the z-directed propagation constant

$$k_z = \sqrt{k^2 - \left(\frac{\pi}{a}\right)^2}$$

which gives us

$$P = \sqrt{1 - \left(\frac{\lambda}{2a}\right)^2} = \cos\theta_{\max}$$

The waveguide propagation constant determines the direction of radiation from a traveling wave leaking out of the guide at a slow rate. We compute the guide width to give radiation in a given direction from

$$a = \frac{\lambda}{2\sqrt{1 - \cos^2\theta_{\max}}}$$

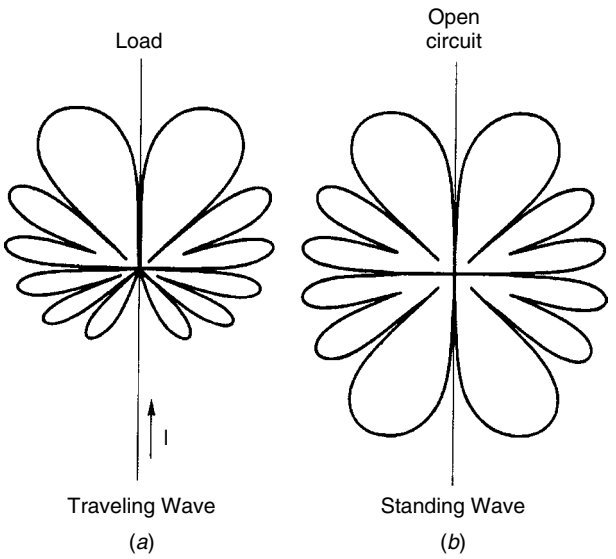


FIGURE 10-5 Patterns of long-wire antennas (3λ long): (*a*) traveling wave; (*b*) standing wave.

10-2 LONG WIRE ANTENNAS

One of the simplest traveling-wave antennas consists of a terminated wire. A standing wave can be divided into two waves traveling in opposite directions. By terminating the wire, we eliminate or reduce the reflected wave and its radiation. The uniform current traveling on a wire is given by $I = I_0 e^{-jkPz}$. We insert this current into Eq. (2-3) to compute the magnetic vector potential and use Eq. (2-1) to calculate the electric field. The resulting radiation intensity is

$$U = \frac{\eta |I_0|^2}{(2\pi)^2} \sin^2 \theta \left[\frac{\sin[kL(p - \cos\theta)/2]}{P - \cos\theta} \right]^2 \tag{10-17}$$

Equation (10-17) separates into two pattern factors: $\sin^2 \theta$, the pattern of an incremental current element, and Eq. (10-6) for the uniform distribution. The null at $\theta = 0$ due to the current element pushes the beam peak off the axis of the traveling wave. Figure 10-5 shows the patterns of traveling-wave and standing-wave currents using a 40-dB scale for $L = 3\lambda$. The patterns are axisymmetrical about the wire. Table 10-1 lists the beam peak direction and the directivity of the pattern for various lengths. The beam peak approaches the wire as the length increases. If we remove the termination, a wave reflects from the end and forms a beam in the opposite direction (Figure 10-5*b*). The small backlobes due to the forward-traveling wave have little effect on the main beam from the backward-traveling wave.

10-2.1 Beverage Antenna [6]

A Beverage antenna consists of a wire strung horizontally a fraction of a wavelength over ground (Figure 10-6). The antenna must be fed relative to ground, and its ground

TABLE 10-1 Characteristics of a Traveling-Wave Current on a Straight Wire

Length (λ)	Directivity (dB)	Beam Peak (deg)	Length (λ)	Directivity (dB)	Beam Peak (deg)
0.5	3.55	64.3	5.5	11.32	20.2
1.0	5.77	47.2	6.0	11.61	19.4
1.5	7.06	38.9	6.5	11.88	18.7
2.0	8.00	33.7	7.0	12.13	17.9
2.5	8.71	30.1	7.5	12.37	17.4
3.0	9.30	27.5	8.0	12.59	16.8
3.5	9.81	25.2	8.5	12.80	16.3
4.0	10.25	23.8	9.0	13.00	15.8
4.5	10.64	22.3	9.5	13.18	15.4
5.0	11.00	21.3	10.0	13.35	15.0

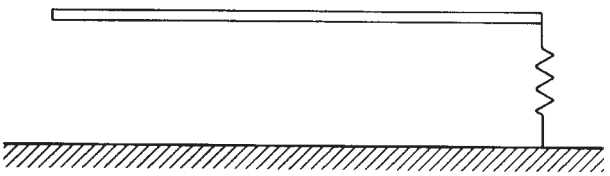


FIGURE 10-6 Beverage antenna.

plane image of the horizontally polarized wire cancels much of the far-field pattern of the antenna. Table 10-1 gives the length for a given takeoff angle for sky waves. The antenna can have good directivity but poor efficiency, since a great deal of the power is absorbed by the load.

We make a vertically polarized Beverage antenna by elevating the feed point on a tower and tilting the wire to the load on the ground. The beam due to the wire and its image add to form a horizontal beam when the wire tilt equals the beam direction given by Table 10-1 for a given length. The vertically polarized antenna needs a more conductive ground plane than the horizontally polarized antenna because soil reflects horizontally polarized waves better than vertically polarized waves. We may vary the tilt angle to produce a takeoff angle, but the beams of the wire and its image may no longer add. It is difficult to feed the tilted antenna because it is fed relative to ground at an elevated point.

10-2.2 V Antenna

We separate two Beverage antennas by an angle and feed from a balanced line to form a V antenna. When the separation angle is twice that given by Table 10-1, the beam peaks of the individual wires add. The balanced line input eases the feed problem of the Beverage antenna, since the antenna no longer feeds against ground. We place the input on a single insulating tower and stretch sloping wires to ground (Figure 10-7) to produce a beam. The beam is horizontal when the tilt angle equals the beam peak angle of Table 10-1. By varying the tilt angle, we change the elevation angle of the beam. A balanced feed produces horizontal polarization on the V antenna and reduces the requirement for a ground screen. The antenna impedance is about 800 Ω. The

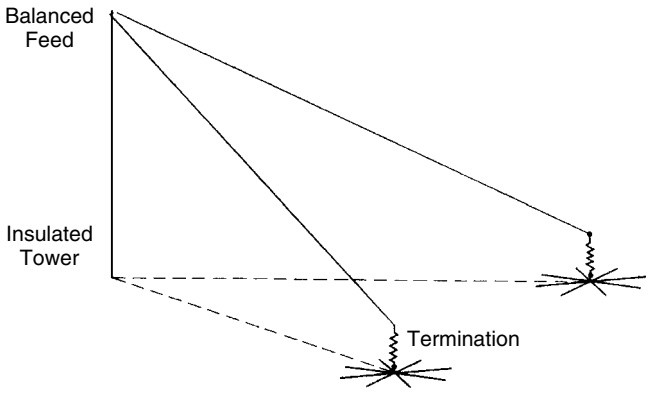


FIGURE 10-7 Sloping V antenna.

pattern contains high sidelobes with an overall low efficiency, but the bandwidth of the terminated antenna approaches an octave.

We can also make an unterminated V antenna. One method is to connect $\lambda/4$ rods, which extend in the same direction as the V elements, to the normally grounded end of the loads on the V antenna [7]. The open-circuited quarter-wavelength lines reflect a short circuit to the loads over a limited bandwidth. The terminating resistors reduce the backlobe from that of the unterminated V antenna while eliminating the need for grounding the loads. In a second design a moment method solution is used to determine the angle between the elements to optimize the directivity of a beam halfway between the directions of the expanding elements [8]. The angle between the elements is fitted to a polynomial:

$$\alpha = -149.3\left(\frac{L}{\lambda}\right)^3 + 603.4\left(\frac{L}{\lambda}\right)^2 - 809.5\frac{L}{\lambda} + 443.6 \text{ degrees} \quad \text{for } 0.5 \le L/\lambda \le 1.5$$

or

$$\alpha = 13.36\left(\frac{L}{\lambda}\right)^2 - 78.27\frac{L}{\lambda} + 169.77 \qquad \text{for} \qquad 1.5 \le L/\lambda \le 3.0 \tag{10-18}$$

where L is the length of each arm. The maximum directivity at α given by Eq. (10-18) is

$$\text{directivity} = 2.94\frac{L}{\lambda} + 1.15 \qquad \text{dB} \tag{10-19}$$

10-2.3 Rhombic Antenna [9]

The rhombic antenna consists of two V antennas. The second V antenna brings the two sides back together and makes it possible to connect the balanced lines to a single terminating resistor (Figure 10-8). Using a single terminating resistor eliminates the grounding problem of V-antenna terminations when the antenna transmits, although the load absorbs up to one-half of the transmitter power. Figure 10-8 shows the patterns of the individual wire radiators on the elements and the combined pattern in the horizontal

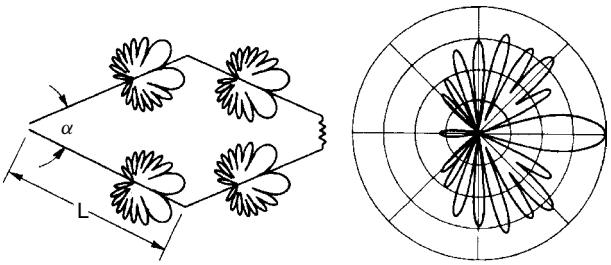


FIGURE 10-8 Rhombic antenna.

TABLE 10-2 V Angle α for Maximum Output of a Rhombic Antenna Given the Elevation Angle

Arm Length,	Elevation Angle (deg)						
L/λ	0	5	10	15	20	25	30
1.5	90	89	88	86	84	80	76
2.0	77	76	75	73	70	66	60
2.5	69	68	66	64	60	56	48
3.0	62	61	60	57	52	48	40
3.5	58	57	55	52	47	42	34
4.0	54	53	50	47	42	36	28
4.5	51	50	47	43	38	32	24
5.0	48	47	44	40	35	28	20
6.0	44	43	40	36	30	22	14
7.0	41	38	36	32	25	18	8
8.0	38	37	34	28	22	14	
10.0	34	32	29	23	16		
15.0	28	26	22	15			
20.0	24	22	17				

Source: [9].

plane for the antenna in free space that unfortunately has high sidelobes. The peak pattern output occurs when the angle α is approximately twice the peak radiation angle of the individual wires. Table 10-2 lists the angle α for the maximum output for a given elevation angle when we mount the antenna parallel with ground. We raise the antenna a height H over ground to control its beam elevation angle:

$$H = \frac{\lambda n}{4 \sin \Delta} \qquad n = 1, 3, 5, \ldots \tag{10-20}$$

where Δ is the elevation angle measured from the horizon and n is an odd integer. The elevated beam forms from a combination of the antenna and its ground image radiations. Because the antenna radiates horizontal polarization, the requirement for a ground-plane screen is minimal.

The terminating resistor should be about 600 Ω, and the input impedance varies from 600 to 900 Ω over an octave bandwidth, the actual value depending on frequency, height, and load resistor. We can use multiple wires that spread apart at the corners to reduce the impedance variation over the band and increase the power-handling capability of smaller wires. We build an inverted V antenna by mounting half a rhombic antenna over a ground plane. We use a single insulated tower; and by controlling the wire tilt angle with respect to ground, we combine the beams for the wire and its ground image into a horizontal beam. The inverted V places both the input and the terminating load on the ground. Because the antenna is vertically polarized, we must provide a good ground screen. In outdoor applications we ground and feed the antenna through a transformer to give some lightning protection and to match the high input impedance of the antenna.

10-3 YAGI–UDA ANTENNAS [10]

A Yagi–Uda antenna uses mutual coupling between standing-wave current elements to produce a traveling-wave unidirectional pattern. It uses parasitic elements around the feed element for reflectors and directors to produce an end-fire beam. Because the antenna can be described as a slow wave structure [11], the directivity of a traveling-wave antenna (Figure 10-2) is bounded when we include the directivity due to the element pattern. Maximum directivity depends on length along the beam direction and not on the number of elements.

Consider two broadside-coupled dipoles. We describe the circuit relation between them by a mutual impedance matrix:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12} & Z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{10-21}$$

where the diagonal elements of the matrix are equal from reciprocity. If we feed one element and load the other, we can solve for the input impedance of the feed antenna:

$$Z_{\text{in}} = \frac{V_1}{I_1} = Z_{11} - \frac{Z_{12}^2}{Z_{22} + Z_2} \tag{10-22}$$

where Z_2 is the load on the second antenna. We short the second antenna ($Z_2 = 0$) to maximize the induced standing-wave current and eliminate power dissipation:

$$Z_{\text{in}} = Z_{11} - \frac{Z_{12}^2}{Z_{22}} \tag{10-23}$$

The mutual impedance between broadside-coupled dipoles (Z_{12}) approaches the self-impedance (Z_{11}) as we move the dipoles close together and causes the input impedance [Eq. (10-23)] to approach zero.

The second equation of Eq. (10-21) for a shorted antenna relates the currents in the two dipoles:

$$0 = Z_{12} I_1 + Z_{22} I_2 \quad \text{or} \quad I_2 = -\frac{Z_{12} I_1}{Z_{22}} \tag{10-24}$$

Since $Z_{12} \approx Z_{22}$, the current in the shorted dipole is opposite the current in the feed element, and radiation from the induced current reduces the fields around the dipoles. Given the current on the parasitic element, we solve for the far field by array techniques. When the elements are spaced a distance d, with the parasitic element on the z-axis and the feed element at the origin, the normalized pattern response is

$$E = 1 + I_r e^{j(kd\cos\theta+\delta)}$$

where $I_r e^{j\alpha} = I_2/I_1$ is the current of the parasitic element relative to the feed element. If we take the power pattern difference between the pattern at $\theta = 0$ and $\theta = 180^\circ$, we get

$$|\Delta E|^2 = -2I_r \sin\delta \sin kd \tag{10-25}$$

Case 1. $\delta = 180^\circ$, $\Delta E = 0$, and we have equal pattern levels in both directions with a null at $\theta = 90^\circ$.

Case 2. $180^\circ < \delta < 360^\circ$, $\Delta E > 0$. The parasitic element is a director, and the pattern in its direction will be higher ($\theta = 0$) than at $\delta = 180^\circ$.

Case 3. $0^\circ < \delta < 180^\circ$, $\Delta E < 0$. The parasitic element is a reflector because the pattern away from it ($\theta = 180^\circ$) is higher than at $\theta = 0^\circ$. We look at the phase of the relative currents to determine whether a parasitic element is a director or a reflector.

The mutual impedance between dipoles has been reduced to equations for a variety of configurations [12–16]. By use of these equations, Figure 10-9 was generated to show the phasing between a half-wavelength dipole and a parasitic dipole as the length and spacing are varied. A parasitic dipole of given length can be either a director or a reflector for different element spacing. Generally, a director is somewhat shorter and a reflector is somewhat longer than the feed element. If we reduce the length of the feed element or increase the element's diameter, the dividing line between a director and a reflector shifts upward. Figure 10-9 also shows the decreased element length at the transition point for additional spaced elements. Figure 10-10 illustrates a three-element Yagi–Uda dipole antenna having one reflector and one director around the feed element. The design is a compromise between various characteristics. With a 50-Ω source its response is as follows:

Gain = 7.6 dB	Front/back ratio = 18.6 dB
Input impedance = $33 - j7.5$	VSWR = 1.57 (50-Ω system)
E-plane beamwidth = 64°	H-plane beamwidth = 105°

With input matching we could increase the gain by 0.2 dB. A 3-dB gain bandwidth is 15% and a 1-dB gain bandwidth is 10%. At the 3-dB band edges the F/B ratio drops to 5.5 dB. As in many designs, the peak gain does not occur at the peak F/B value. The gain rises by 0.2 dB to a 50-Ω source at a point 3% higher in frequency than the point of maximum F/B. The maximum gain with input impedance matching (8.6 dB) occurs at a point 7% above the center frequency. The dipole element pattern narrows the E-plane beamwidth and produces a null at 90° from the boresight. The traveling wave alone forms the H-plane beam. A design to optimize the gain would have a phase progression along the elements to match Eq. (10-10) for short traveling-wave

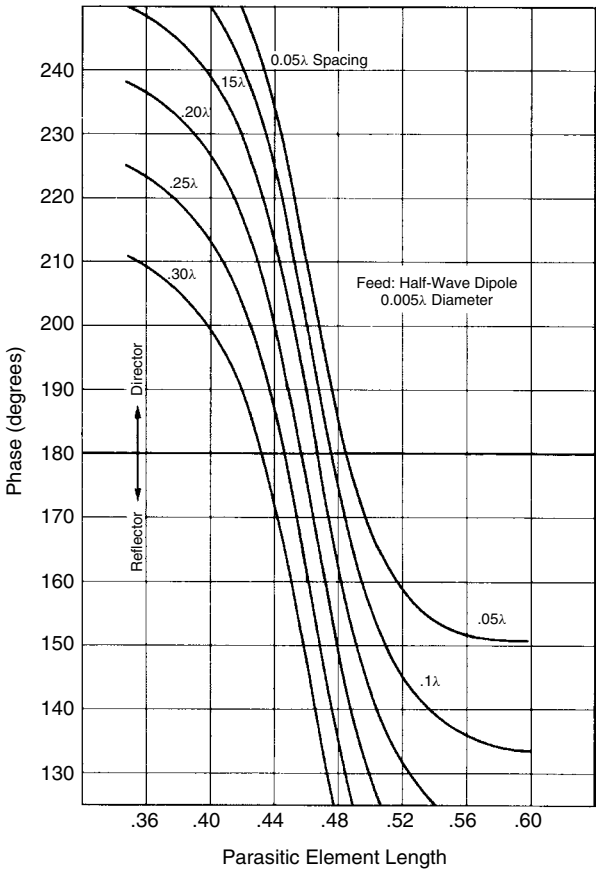


FIGURE 10-9 Phase of current on a parasitic dipole relative to current on a driven dipole.

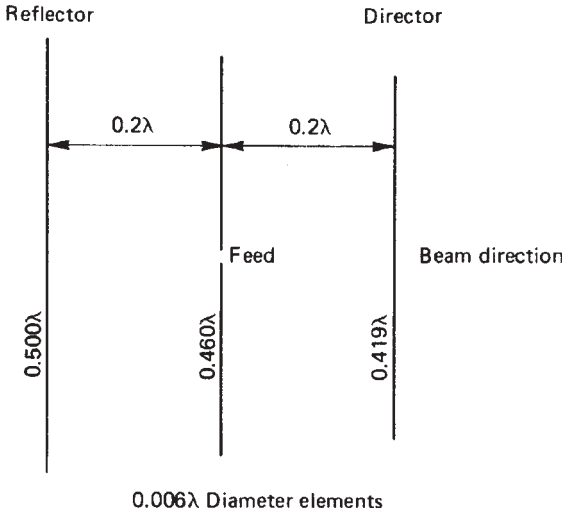


FIGURE 10-10 Three-element Yagi–Uda dipole antenna.

antennas. The array of dipoles samples an aperture 0.6λ long. Figure 10-3 gives an H-plane beamwidth of about 80° for this optimum design. With only three elements we find difficulty in matching the phase distribution required and will produce a design with a low F/B value.

We analyze Yagi–Uda antennas by using the moment method [17]. First, calculate the mutual impedance matrix:

$$[V] = [Z][I] \tag{10-26}$$

The input voltage vector $[V]$ has only one nonzero term (the feed). By solving the linear equations (10-26), we compute the currents at the base of each element. We assume a sinusoidal current distribution on each element [Eq. (5-1)] and solve for the pattern response from the array of dipoles. We calculate the input impedance of the array by using the moment method; and by retaining the current levels on the dipoles for a known input power, we can calculate gain directly.

The moment method allows reasonably quick calculation of the antenna characteristics so that optimization techniques can be applied. Cheng and Chen [18,19] use a perturbation technique that alternates between element spacings and lengths to find an optimum design with rapid convergence. Running a single iteration of each produces the design. No optimization technique can assure that a global optimum can be found. Designs will converge to local maximums. By limiting the search variables, the method of Cheng and Chen avoids some local optimums. Most optimizations work at a single frequency and produce designs with high gains that fall off rapidly with increasing frequency because the wave relative propagation constant increases with frequency. The gain curves in Figure 10-1 decrease rapidly after the peak point as the sidelobes increase. All traveling-wave end-fire antennas follow this pattern.

Kajfez [20] developed a method for optimizing over a band of frequencies. His method is to reduce the peak gain while increasing the bandwidth and reducing some of the tolerance requirements. We fix the band limits at frequencies f_1 and f_2 and optimize a composite cost function containing both the average directivity and the ripple or deviation from the average. Maximize

$$C = D_{\text{avg}} - wd \tag{10-27}$$

where D_{avg} is the average directivity and d is the RMS deviation of the directivity from D_{avg} over the band weighted by w:

$$D_{\text{avg}} = \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} D(f)df$$

$$d = \sqrt{\frac{1}{f_2 - f_1} \int_{f_1}^{f_2} [D(f) - D_{\text{avg}}]^2 df} \tag{10-28}$$

$D(f)$ is the directivity function (a ratio, not dB). We vary w to stress either high directivity or design flatness. If w is made two or more times greater than the directivity expected, an optimization routine will design an antenna with a flat response. We calculate the integrals by using Gauss–Legendre quadrature, which requires only a few evaluations of the pattern response (five to eight).

We usually use only one reflector element. More elements can be used, but they add little. The gain of a two-element pair—a feed dipole and a reflector dipole—rises with increasing separation in a broad peak starting about 0.15λ, peaking about 0.20λ, and falling slowly with greater spacing. The reflector primarily affects F/B. As we move it closer to the feed element, the impedance level of the antenna drops. The response of the antenna is less sensitive to the diameter of the reflectors than to the director diameters, and it requires smaller adjustments to the length to compensate the design for the change.

We add length to the antenna to increase gain by adding director elements. Length along the axis, not the number of elements, produces gain provided that element currents are phased correctly. Beyond 0.3λ to 0.4λ spacings, the coupling drops and reduces control of the current phasing on the directors. Changing the diameters of directors requires changing the element lengths to retune the antenna. In one method, computer optimization routines are used to vary the dimensions to optimize some cost function. The lengths of the feed and first director elements have the greatest effect on gain. If we use a gradient-type search, we observe that most of the design concentrates on varying those element lengths. In long antennas with many directors the elements far from the feed change little from initial lengths by a gradient search. We can also adjust the lengths to match the current phasing of the elements to a distribution such as that given by Eq. (10-10). Manual tuning requires a retuning of the two elements closest to the feed and the feed element after tuning each director.

When adjusting the lengths to maximize bandwidth [Eqs. (10-27) and (10-28)], we should optimize gain to a given source impedance. We can optimize a single-frequency antenna to any reasonable impedance level and design a matching network, but broadband antennas may produce difficult matching network problems when we try to achieve design gain. A moment method gives us the peak gain to a conjugate matched source, but to compute gain to a given source impedance Z_s, we multiply the voltage gain by

$$\frac{2\sqrt{\text{Re}(Z_I)\text{Re}(Z_s)}}{|Z_I + Z_s|} \tag{10-29}$$

where Z_I is the antenna input impedance. Equation (10-29) allows optimizations to a given source impedance.

Table 10-3 gives the dimensions of an antenna designed by using Eqs. (10-27) and (10-28) for a maximum gain to a 50-Ω source [Eq. (10-29)]. The normalized frequency limits are 0.95 to 1.05, with a weighting w of 15 on the RMS deviation d.

TABLE 10-3 Six-Element Yagi–Uda Dipole Antenna with 0.01λ-Diameter Elements

Element Type	Element Length (λ)	Location Along Boom (λ)
Reflector	0.484	0
Feed	0.480	0.250
Director	0.434	0.400
	0.432	0.550
	0.416	0.700
	0.400	0.850

TABLE 10-4 Normalized Frequency Response of the Yagi–Uda Dipole Antenna of Table 10-3

Normalized Frequency	Gain, 50-Ω Source (dB)	Max Gain (dB)	F/B (dB)	Input Impedance (Ω)	VSWR, 50-Ω Source
0.90	4.8	6.3	1.4	$29.6 - j45.2$	3.36
0.92	7.0	7.7	3.6	$32.5 - j30.8$	2.29
0.94	8.4	8.7	6.5	$39.3 - j19.6$	1.65
0.96	9.0	9.2	9.9	$42.7 - j13.2$	1.39
0.98	9.4	9.5	14.4	$39.5 - j7.6$	1.34
1.00	9.7	10.0	22.7	$31.2 + j2.8$	1.61
1.02	9.5	10.6	21.1	$22.0 + j19.8$	2.70
1.04	8.1	11.1	12.2	$15.2 + j42.5$	5.79
1.06	5.5	10.8	7.2	$12.9 + j69.1$	11.4
1.08	2.5	8.9	3.7	$16.0 + j97.2$	15.2

Table 10-4 lists the calculated response of the antenna. Like many designs, gain falls off more rapidly on the high-frequency end. In this case maximum gain actually occurs at maximum F/B. Pattern calculations of a uniform traveling-wave distribution, satisfying the Hansen and Woodyard criterion, gives peak F/B when it is an odd integer multiple of λ/4, of which Yagi–Uda antennas are one example.

Higher gains can be attained for six elements. Chen and Cheng [19] achieved 13.4 dB, but for a restricted bandwidth. We achieve bandwidth by limiting gain, and with each added element we can improve flatness at a higher gain level. Kajfez shows a nearly constant gain–bandwidth product for different designs. The gradient search method used in the design in Table 10-3 is sensitive to initial conditions and often converges to a local maximum and misses the global maximum, because it cannot escape a local optimum. A better design may be achieved by using a new starting point.

Table 10-5 lists the design of a 16-element Yagi–Uda dipole antenna. In this case the optimization routine manipulated elements 4 through 16 as a group to reduce the number of different-sized elements to four. This technique reduces tooling cost and causes little loss in possible performance. Since the size of the last few directors has only a minor effect on the total gain, most of the iterations in an optimization routine are spent changing the elements around the driven dipole and only slowly changing those elements. We adjust the reflector, feed element, and first director to match into the uniform surface wave structure of equally spaced elements.

TABLE 10-5 Sixteen-Element Yagi–Uda Dipole Antenna with 0.006λ-Diameter Elements

Element	Element Length (λ)	Element Spacing (λ)
Reflector	0.4836	0.2628
Feed	0.4630	0.2188
Director 1	0.4448	0.2390
Directors 2 to 13	0.4228	0.2838

The frequency response (Table 10-6) has a peak gain equal to the value given by Figure 10-2 for the hybrid mode and at this overall length. Like many Yagi–Uda dipole antennas, this antenna matches best to a low input impedance (30 Ω). The gain rises slowly below the resonant frequency of the antenna and falls rapidly for higher frequencies, as predicted by Figure 10-1, which shows gain versus relative propagation constant along a uniform surface-wave structure.

The element diameter affects the optimum length of the Yagi–Uda dipole elements. A number of optimum Yagi–Uda dipole antennas of various overall lengths were built with different-diameter elements and adjusted to regain optimum performance [21]. Table 10-7 summarizes the length changes required of directors when their diameters are changed. Similarly, Table 10-8 lists the changes required for reflectors. Tables 10-7 and 10-8 require little or no change in the feeder element length to regain the optimum design. We use the tables to modify element lengths when changing

TABLE 10-6 Normalized Frequency Response of a 16-Element Yagi–Uda Dipole Antenna

Normalized Frequency	Gain, 30-Ω Source (dB)	F/B (dB)	Input Impedance (Ω)	VSWR, 30-Ω Source
0.95	12.1	12.2	$26.7 - j36.1$	3.34
0.96	13.4	12.2	$28.4 - j26.4$	2.40
0.97	14.4	12.2	$31.6 - j17.7$	1.77
0.98	15.0	13.3	$34.3 - j12.1$	1.49
0.99	15.5	17.3	$32.2 - j7.3$	1.28
1.00	15.9	35.1	$27.0 + j3.4$	1.17
1.01	15.2	15.6	$27.3 + j20.8$	2.05
1.02	14.0	10.6	$40.4 + j33.0$	2.61
1.03	13.7	11.8	$36.0 + j23.5$	2.06
1.04	8.2	16.4	$22.8 + j51.5$	5.77
1.05	5.4	4.7	$54.1 + j68.2$	5.03

TABLE 10-7 Length Changes of Directors in a Yagi–Uda Dipole Antenna for Various Diameters

Dipole Diameter (λ)	Dipole Length Change (λ)	Dipole Diameter (λ)	Dipole Length Change (λ)
0.001	0.030	0.008	0.006
0.0012	0.029	0.009	0.002
0.0015	0.027	0.010	0.000
0.002	0.025	0.012	−0.004
0.0025	0.023	0.015	−0.010
0.003	0.021	0.020	−0.018
0.004	0.017	0.025	−0.024
0.005	0.014	0.030	−0.029
0.006	0.011	0.040	−0.038
0.007	0.008		

Source: [21].

TABLE 10-8 Length Changes of Reflectors in a Yagi–Uda Dipole Antenna for Various Diameters

Dipole Diameter (λ)	Dipole Length Change (λ)	Dipole Diameter (λ)	Dipole Length Change (λ)
0.001	0.011	0.008	0.002
0.002	0.008	0.010	0.000
0.003	0.006	0.020	−0.003
0.004	0.005	0.030	−0.005
0.006	0.003	0.040	−0.006

Source: [21].

TABLE 10-9 Length Changes to Increase Elements of a Yagi–Uda Dipole Antenna to Account for Boom Diameter

Boom Diameter (λ)	Dipole Length Change (λ)	Boom Diameter (λ)	Dipole Length Change (λ)
0.002	0.0010	0.022	0.0158
0.004	0.0022	0.024	0.0173
0.006	0.0034	0.026	0.0189
0.008	0.0048	0.028	0.0205
0.010	0.0064	0.030	0.0220
0.012	0.0084	0.032	0.0236
0.014	0.0095	0.034	0.0252
0.016	0.0111	0.036	0.0265
0.018	0.0127	0.038	0.0283
0.020	0.0142	0.040	0.0299

Source: [21].

diameters. Element length tolerance is 0.003λ. The boom supporting the elements also affects element length for optimum designs. Table 10-9 lists element length increase versus boom diameter.

Example Determine the length adjustment necessary when building the design of Table 10-5 with 0.002λ elements.

From Table 10-7, the length adjustment for 0.006λ is 0.010λ and the adjustment for 0.002λ-diameter elements is 0.025λ. We increase the lengths of the directors by the difference, 0.015λ. Similarly, from Table 10-8 we increase the reflector length by 0.005λ. For a boom diameter of 0.02λ we also increase the lengths of all the elements by 0.0142λ (Table 10-9).

10-3.1 Multiple-Feed Yagi–Uda Antennas

We further improve the response of Yagi–Uda antennas by directly feeding more than one element. The convenience of Yagi–Uda antennas lies in the single feed point. With many feed points a log-periodic dipole antenna (Section 11-12) produces a wider band design. Figure 10-11 illustrates a four-element design with a crisscross feeder. Table 10-10 lists the calculated response of the antenna. The gain response is flat with good F/B and input VSWR over the band for this short antenna (0.3λ long), where

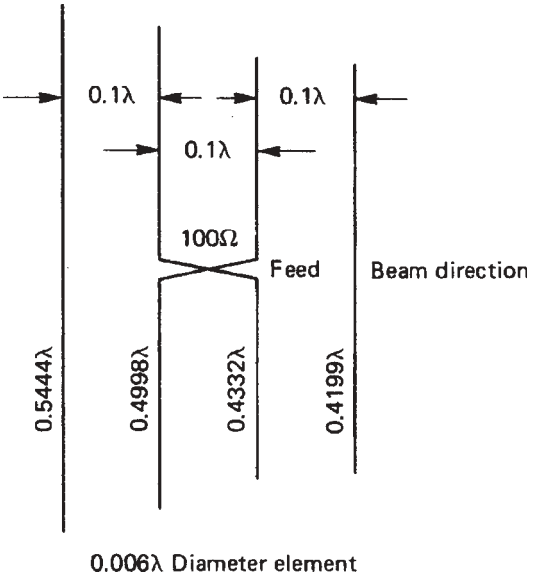


FIGURE 10-11 Multiple-feed Yagi–Uda dipole antenna.

TABLE 10-10 Normalized Frequency Response of a Four-Element Yagi–Uda Dipole Antenna with a Crisscross Feeder Between Two Elements (Figure 10-11)

Normalized Frequency	Gain, 50-Ω Source (dB)	F/B (dB)	Input Impedance (Ω)	VSWR, 50-Ω Source
0.90	7.0	12.1	$30.1 + j14.0$	1.82
0.92	7.1	17.5	$42.6 + j4.0$	1.20
0.94	7.0	21.9	$44.0 - j3.0$	1.15
0.96	6.9	25.9	$43.9 - j5.8$	1.20
0.98	7.0	31.3	$44.4 - j7.3$	1.22
1.00	7.0	53.0	$45.2 - j9.4$	1.25
1.02	7.1	30.7	$44.8 - j13.5$	1.35
1.04	7.2	23.6	$40.0 - j18.8$	1.60
1.06	7.0	19.0	$29.0 - j20.5$	2.12
1.08	6.5	15.3	$16.5 - j14.2$	3.29
1.10	5.0	11.9	$8.0 - j2.3$	6.30

the crisscross feeder adds an additional 180° phase shift along the length to produce a backfire pattern from a fast wave.

The simple antenna of Figure 10-12 uses only two dipole elements with a crisscross feed. Like many Yagi–Uda dipole antennas, this antenna, matches best to a low-impedance (25-Ω) source. Although its gain and bandwidth (Table 10-11) are less than those of the four-element multiple-feed antenna, its bandwidth is better than that of a singly fed equivalent antenna. The length along the boom still determines the ultimate gain of the antenna, but the backfire crisscross feed improves F/B.

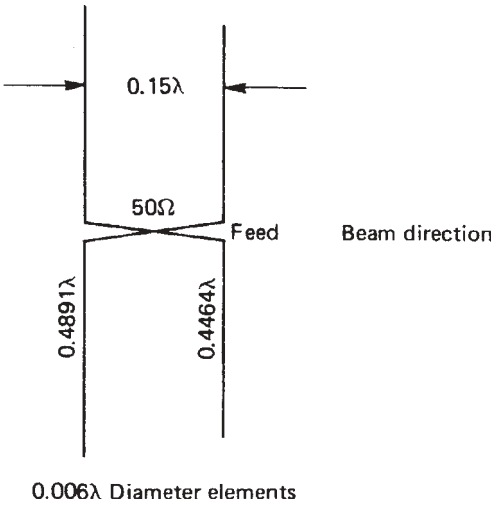


FIGURE 10-12 Dipole two-element backfire antenna.

TABLE 10-11 Normalized Frequency Response of a Two-Element Crisscross-Fed Antenna

Normalized Frequency	Gain (dB), 25-Ω Source	F/B (dB)	Input Impedance (Ω)	VSWR, 25-Ω Source
0.95	5.0	13.1	$12.8 - j1.6$	1.96
0.96	5.2	14.8	$13.6 - j0.5$	1.84
0.97	5.4	17.0	$14.4 + j0.8$	1.74
0.98	5.6	20.0	$15.2 + j2.3$	1.67
0.99	5.8	24.6	$16.0 + j3.9$	1.62
1.00	5.9	34.1	$17.0 + j5.9$	1.61
1.01	6.0	29.0	$18.2 + j8.2$	1.64
1.02	6.1	22.0	$19.7 + j10.8$	1.71
1.03	6.1	18.1	$21.7 + j14.0$	1.84
1.04	6.1	15.3	$24.3 + j17.6$	2.02
1.05	6.1	13.2	$27.9 + j22.0$	2.36

We can no longer just use Eq. (10-26), the matrix equation of mutual impedances, to analyze a multiple-feed antenna. We invert the mutual impedance matrix and add a 2×2 admittance matrix of the transmission-line feeders between the dipole center point nodes:

$$[I] = [Y_a + Y_f][V] \tag{10-30}$$

where Y_a is the dipole mutual admittance matrix and Y_f is the feeder admittance matrix. Since most of the elements are short-circuited, we reduce the matrix [Eq. (10-30)] to only the terms involving unshorted dipoles and solve for the input voltages for a current vector with one nonzero term (the input). We return to Eq. (10-26) and use the voltage vector obtained from the solution of Eq. (10-30) to calculate dipole base currents.

Multiple feeds can also be used to reduce sidelobes. A backward-wave excited array (crisscross feeder) feeding a Yagi–Uda dipole antenna reduced the antenna sidelobes [22]. The extra feed points add degrees of freedom for a part of design independent of the number of elements.

10-3.2 Resonant Loop Yagi–Uda Antennas

We can make Yagi–Uda arrays by using the resonant loop element (Section 5-18). The loop radiates its maximum signal normal to the plane of the loop (along the axis) and has a linearly polarized wave in the direction of the voltage across the feed. A two-element loop parasitic array, reported in 1942, was built to eliminate corona problems at high altitudes in Ecuador [23]. The maximum standing-wave voltage points occur at $\lambda/4$ from the feed along the loop and not at the ends of rods, where air breakdown can occur. The symmetry of the loop reduces the effects of nearby structures on the antenna and they can be mounted at the voltage null opposite the feed, along a metal rod in a coaxial array with little effect.

Loop shape has only a minor effect on input impedance at a resonant perimeter length of about 1.1λ. The larger size (and gain) of the basic element raises the gain of a parasitic (Yagi–Uda) array from an array made with dipoles. A simple antenna consists of two square loops (Figure 10-13) with a feed element and a parasitic reflector. The reflector perimeter is somewhat larger than the feed loop. Such an antenna, commonly called a *cubic quad*, is usually made by stretching wires on a frame. The direction of the voltage across the feed point determines the linear polarization sense. The antenna of Figure 10-13 radiates horizontal polarization as fed.

Table 10-12 lists the dimensions of antennas designed to be resonant (zero reactance) and have maximum F/B ratios with a wire radius of 0.0002λ. The maximum F/B occurs for a loop spacing of 0.163λ. Table 10-13 lists the frequency response of the antenna spaced at 0.15λ. As with a Yagi–Uda dipole antenna, the design must be adjusted as the wire radius changes to obtain the same response. Table 10-14 lists the design changes on a logarithmic scale. The greater change occurs to the reflector perimeter.

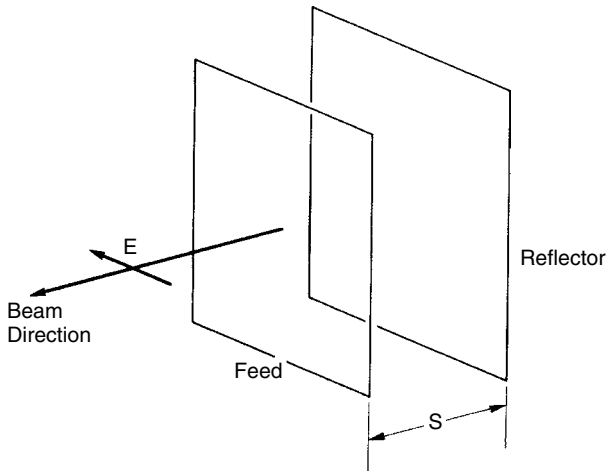


FIGURE 10-13 Cubic-quad traveling-wave resonant loop antenna.

TABLE 10-12 Characteristics of Resonant Cubic-Quad Antennas, 0.0002λ-Diameter Wire

Element Spacing (λ)	Feed Perimeter (λ)	Reflector Perimeter (λ)	Gain (dB)	F/B (dB)	Input Resistance
0.10	1.000	1.059	7.2	17.5	76
0.15	1.010	1.073	7.1	32.8	128
0.16	1.013	1.075	7.1	46.1	137
0.163	1.014	1.0757	7.1	59.6	140
0.17	1.016	1.077	7.1	38.1	145
0.18	1.018	1.079	7.0	31.0	153
0.20	1.025	1.082	6.9	24.6	166

TABLE 10-13 Normalized Frequency Response of a Cubic-Quad Antenna Resonant at a Spacing of 0.15λ

Normalized Frequency	Gain (dB)	F/B (dB)	Input Impedance (Ω)
0.96	7.3	2.9	$38.5 - j140.9$
0.98	7.8	11.0	$72.9 - j55.4$
0.99	7.5	17.9	$100.5 - j22.4$
1.00	7.1	32.8	128.0
1.01	6.8	19.9	$150.7 + j13.3$
1.02	6.8	14.8	$167.0 + j24.2$
1.03	6.2	12.1	$178.0 + j35.1$
1.04	6.0	10.4	$185.4 + j47.3$
1.06	5.7	8.2	$195.0 + j77$
1.08	5.5	6.8	$202.0 + j113.4$

TABLE 10-14 Cubic-Quad Antenna Adjustments for Changing Wire Diameter

Wire Diameter (λ)	Reflector Perimeter (λ)	Feeder Perimeter (λ)
0.00005	−0.013	−0.002
0.0001	−0.007	−0.001
0.0002	0	0
0.0004	0.009	0.002
0.0008	0.019	0.003
0.0016	0.033	0.005
0.0032	0.052	0.008

When close to the final design, we adjust the reflector for maximum F/B and the feed loop for resonance independently. Each has little effect on the other.

We approach the design of coaxial loop arrays from traveling-wave phasing. Table 10-15 lists the traveling-wave relative propagation constant along an infinite array of loops. For a given array length we use Eq. (10-8) or (10-10) to determine the required mode relative propagation constant to maximize gain.

TABLE 10-15 Relative Propagation Constant Along a Coaxial Array of Circular Loops

Circumference,	*S*/*b*		
$2\pi b/\lambda$	0.25	0.50	1.00
0.74	1.05	1.03	1.01
0.76	1.06	1.04	1.016
0.78	1.08	1.05	1.02
0.80	1.09	1.06	1.03
0.82	1.12	1.08	1.04
0.84	1.14	1.11	1.06
0.86	1.17	1.13	1.07
0.88	1.20	1.16	1.09
0.90	1.24	1.20	1.12
0.92	1.30	1.26	1.16
0.94	1.37	1.32	1.23
0.96	1.47	1.40	1.32
0.98	1.60	1.55	1.44

Source: [24].

Example Compute the loop size for an array 2λ long with the ratio of loop spacing to loop radius = 0.5 for maximum gain.

We use Eq. (10-10) with $R = 4.5$ to calculate $P = 1 + 1/9 = 1.111$. By using Table 10-15, we have $2\pi b/\lambda = 0.84$ and $b = 0.134\lambda$. The spacing between loops is $0.5(0.134) = 0.0668\lambda$, and 30 loops are required for an antenna length of 2λ.

Table 10-15 suggests closely spaced elements. With further tabular data [25], designs with loop spacing up to 0.3λ are possible. The reflector element is spaced 0.1λ away from the exciter with a perimeter of 1.05λ and we adjust the perimeter of the feed element to resonate the antenna.

In the design above, the Hansen–Woodyard criterion for maximum gain for the relative propagation constant was not used. The gain in Figure 10-1 falls rapidly for an increase in P (increased frequency), and we obtain a better pattern bandwidth by designing with less than the maximum gain value of P. The possible gain bandwidth shrinks (Figure 10-1) when we increase length to obtain more gain. This is a general property of all traveling-wave end-fire structures. Similarly, the beam of a leaky wave antenna ($P < 1$) scans toward end fire with increasing frequency because P increases.

We can combine loops and dipoles in the same array. In some designs, dipoles are used for far-out directors in front of loops. Similarly, a parasitic loop can provide an effective reflector element for a dipole or crossed dipoles. A loop spaced 0.25λ from a resonant dipole (0.47λ long) increased gain to 5.9 dB with an F/B value of 21.7 dB for a loop circumference of 1.15λ.

10-4 CORRUGATED ROD (CIGAR) ANTENNA

A corrugated rod traveling-wave structure consists of disks attached to a central metal rod (Figure 10-14). The rod supports a TM axisymmetrical mode with readily calculable parameters [2] that has a pattern null on axis. The hybrid mode HE_{11} (Section 7-3)

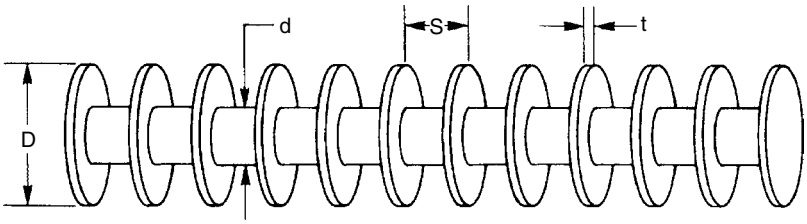


FIGURE 10-14 Corrugated rod (cigar) traveling-wave structure.

TABLE 10-16 Measured Relative Propagation Constant on a Corrugated Rod (Cigar) Antenna[a]

$(D-d)/\lambda$	P	$(D-d)/\lambda$	P
0.15	1.03	0.275	1.23
0.175	1.05	0.30	1.31
0.20	1.08	0.325	1.47
0.225	1.12	0.35	1.67
0.25	1.16	0.375	1.92

Source: [26].

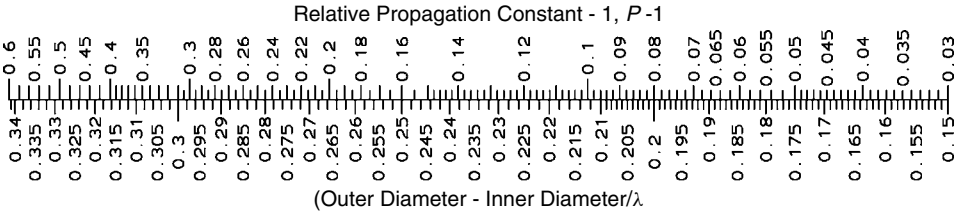
[a] $0.15 \leq$ disk spacing/$\lambda \leq 0.21$; $0.15 \leq$ central rod diameter/$\lambda \leq 0.21$; $0.018 \leq$ disk thickness/$\lambda \leq 0.025$.

propagates on the rod to produce a pattern with its peak on the axis with linear polarization. The hybrid mode is the sum of the TE_{11} and TM_{11} modes, and the surface wave has exponentially decaying radial fields.

We design this antenna empirically by measuring phase velocities along an excited structure. The gain increases slightly for an increased diameter, but it is determined primarily by length. Table 10-16 lists measured data for the range of disk spacing from 0.15λ to 0.21λ. In this range, disk spacing has a negligible effect on the wave velocity. We use this table with a given length to determine the necessary relative propagation constant P.

We can excite the corrugated rod from a number of feeding structures. A circular waveguide propagating the TE_{11} mode will excite the rod in the hybrid HE_{11} mode [26] when tapered into the guide. A resonant ring will excite the rod when backed by a suitable reflector ring or disk. Wong and King [27] excite the rod from an open-sleeve dipole in a cavity (Figure 5-28). The length of the rod controls the pattern beamwidth, and the feed structure controls the impedance. Each may be adjusted somewhat separately.

The traveling-wave disk on rod radiates high sidelobes of about 10 dB, which can be reduced by placing the antenna in a cone [27]. The rod runs the length of the cone and extends beyond the end by about 0.72λ. This antenna uses a corrugated rod 3.39λ long at the center frequency and is placed in a cone with a 2.94λ diameter and a height of 2.58λ. The rod was under the limit given above with a diameter of 0.074λ and average disk diameter of 0.311λ. Using the scale, we read $P = 1.139$, close to the optimum value of 1.137 for the given rod length. A disk spacing of 0.24λ shows that we can build successful antennas beyond the limits given above and that the scale



SCALE 10-1 Corrugated rod relative propagation constant P given dimensions.

can still be used. When placed in the cone, the sidelobes dropped to about −30 dB, and when fed by the disk sleeve dipole shown in Figure 5-28, the antenna has a 34% 2 : 1 VSWR bandwidth. At a peak gain of 16.5 dB, the antenna exceeds the gain of a traveling-wave antenna by about 1 dB for its given length, but falls short of the gain of a horn by about 1 dB. Table 10-16 was interpolated to produce Scale 10-1 for the design of a corrugated rod.

Example Design a corrugated rod 4λ long at 10 GHz (30 mm wavelength).

By using Eq. (10-13), we compute directivity = 7.58(4) = 30.32 or 14.8 dB for P given by Eq. (10-8). $P = 1 + 0.465/4 = 1.116$. We use $P - 1$ with Scale 10-1 to determine the difference between the central rod diameter and the disk diameter: $(D - d)/\lambda = 0.2224$. If we select the midpoint in the range of values for the dimensions, we obtain the dimensions of the corrugated rod in the center constant-dimension region. The disk spacing $S = 0.18\lambda = 5.4$ mm. The central rod diameter $d = 0.18\lambda = 5.4$ mm. The outer disk diameter $D = 0.2224\lambda + 0.18\lambda = 0.4024\lambda = 12.07$ mm. The disk thickness $t = 0.022\lambda = 0.66$ mm. The corrugated rod should start with $P = 1.3$ at the feed point. We read $(D - d)/\lambda = 0.2973$, which calculates to an outer disk diameter $D = (0.2973 + 0.18)\lambda = 14.32$ mm. The initial taper can be three disk elements long. At the end it is a good idea to taper the last few disk diameters to reduce end reflection.

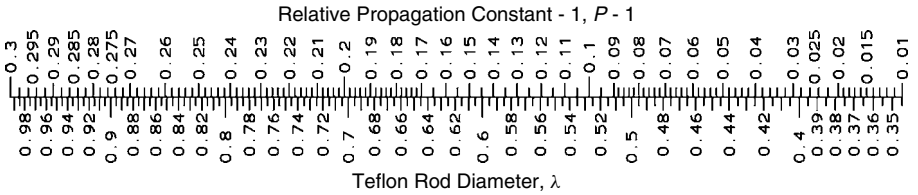
10-5 DIELECTRIC ROD (POLYROD) ANTENNA

A dielectric rod will support an HE_{11} hybrid mode. Inside the rod we describe the fields with Bessel function J_n. Outside the rod the fields fall off exponentially and we use the modified Bessel function K_n. The hybrid mode consists of the sum of TE_{11} and TM_{11} modes. To determine the mode velocity, we equate the propagation constants of the internal and external waves. This equation contains two constants and we eliminate one by equating the radially directed wave impedance at the boundary. The result is a transcendental equation that must be solved by graphical or numerical methods [2]. Table 10-17 summarizes the results of these calculations for common dielectrics. Interpolation of the table produces Scales 10-2 and 10-3 for design using Telfon or Delrin with a dielectric constant of 3.1.

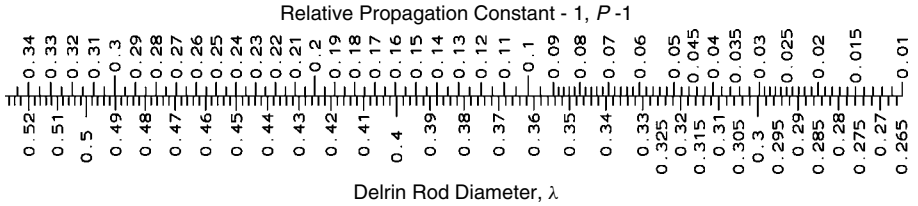
Figure 10-15 shows a common feeding arrangement for the polyrod antenna. The rod protrudes from a circular waveguide supporting the TE_{11} mode, which excites the hybrid mode HE_{11} on the rod. At the waveguide exit we use a rod diameter to give P from 1.2 to 1.3 so that the wave will be closely bound to the rod. The feeding

TABLE 10-17 Diameter of a Dielectric Rod Supporting an HE_{11} Mode (λ)

	Dielectric Constant				
P	2.08	2.32	2.55	3.78	10
1.01	0.345	0.316	0.296	0.240	
1.02	0.378	0.345	0.322	0.257	0.1780
1.03	0.403	0.366	0.340	0.270	0.1824
1.04	0.425	0.384	0.356	0.279	0.1860
1.05	0.444	0.400	0.369	0.287	0.1888
1.06	0.462	0.414	0.381	0.294	0.1912
1.07	0.479	0.427	0.393	0.300	0.1933
1.08	0.495	0.440	0.404	0.306	0.1951
1.10	0.527	0.465	0.424	0.317	0.1983
1.12	0.559	0.489	0.444	0.327	0.2010
1.14	0.592	0.513	0.463	0.336	0.2034
1.16	0.627	0.538	0.482	0.344	0.2055
1.18	0.663	0.563	0.501	0.353	0.2074
1.20	0.703	0.590	0.521	0.361	0.2092
1.25	0.823	0.664	0.575	0.381	0.2133
1.30	0.994	0.758	0.638	0.402	0.2170
1.35	1.283	0.885	0.716	0.424	0.2205
1.40		1.081	0.820	0.448	0.2238



SCALE 10-2 Teflon rod HE_{11}-mode relative propagation constant.



SCALE 10-3 Delrin rod HE_{11}-mode relative propagation constant.

guide (Figure 10-15) has a quarter-wavelength choke to reduce the backfire lobe due to direct radiation from the transition [28]. The choke region can also be flared in a short horn [29].

The second region of the rod tapers either to a uniform diameter section to produce maximum gain or to a tapered section to reduce sidelobes. At the end of the antenna we taper the rod rapidly in a terminating section to bring the relative propagation constant

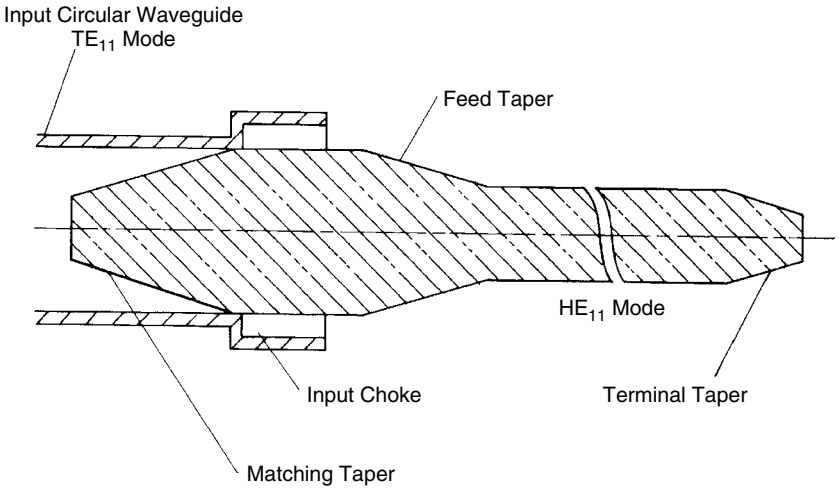


FIGURE 10-15 Dielectric rod (polyrod) antenna.

of the surface wave near 1, to reduce reflection from the end. We calculate P along the guide and adjust the uniform section diameter or tapered section length to satisfy the total extra phase shift condition for maximum end-fire radiation.

Example Design a dielectric rod antenna 5λ long using Teflon and compare it to a design using Delrin.

The relative propagation constant for peak gain is independent of the material, which we compute from Eq. (10-8): $P = 1 + 0.465/5 = 1.093$. By using Scales 10-2 and 10-3, we read the rod diameters: 0.516λ for Teflon and 0.356λ for Delrin. At the point where the rod exits from the feeding waveguide, a suitable relative propagation constant is 1.25. We use the scales to find the rod diameters: 0.822λ for Teflon and 0.456λ for Delrin. These diameters are proportional to the free-space wavelength, not the wavelength in the rod.

The loss tangent for Delrin is 0.005, whereas Teflon is only 0.0012; we need to consider the loss in the antenna. The loss for a wave propagating through a dielectric is given by

$$\frac{20\pi\sqrt{\varepsilon_r}\tan\delta}{\lambda \ln 10} = \frac{27.3\sqrt{\varepsilon_r}\tan\delta}{\lambda} \qquad \text{dB}/\lambda$$

Consider that only a portion of the power in the wave propagates in the dielectric. We use the effective dielectric constant to determine a filling factor that reduces the loss:

$$\text{filling factor} = \text{QF} = \frac{1 - 1/\varepsilon_{\text{eff}}}{1 - 1/\varepsilon_r}$$

We relate the effective dielectric constant to P: $\varepsilon_{\text{eff}} = P^2 = (1.093)^2 = 1.195$:

$$\text{loss(dB)} = 27.3\text{QF}\sqrt{\varepsilon_r}\tan\delta[\text{Length}(\lambda)]$$

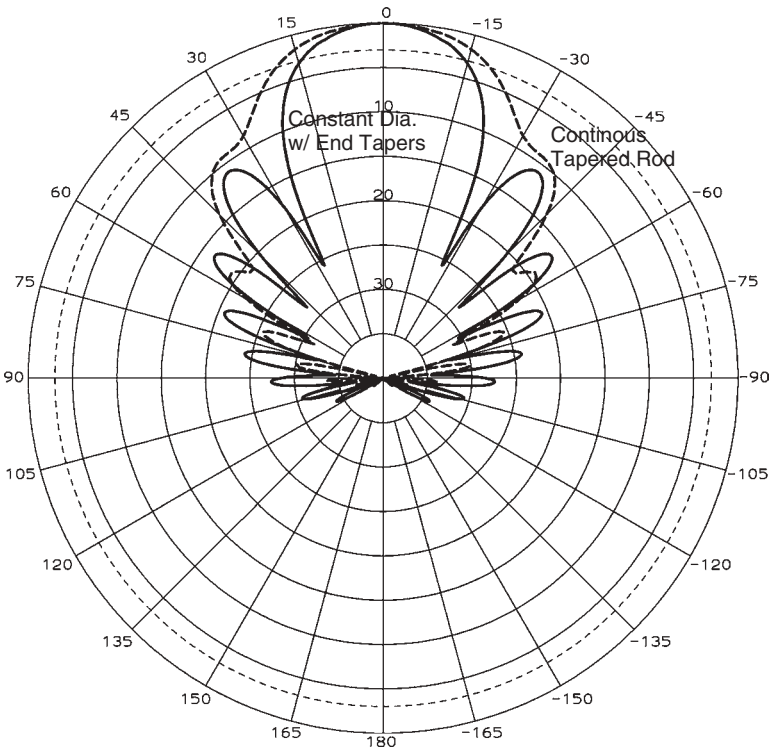


FIGURE 10-16 Teflon rod antenna: constant diameter with short end tapers (solid line); continuous tapered rod (dashed line).

The filling factors are easily calculated: QF = 0.3138 (Teflon) and QF = 0.2405 (Delrin). When we use the filling factor, we calculate the losses: 0.074 dB (Teflon) and 0.29 dB (Delrin). The filling factor reduces the losses if the wave is traveling totally in the waveguide: 0.24 dB (Teflon) and 1.20 dB in Delrin. Because a large proportion of the power travels outside the rod, it is reasonable to use a lossy dielectric for the antenna.

Figure 10-16 shows the patterns of 5λ-long Teflon rod antennas for two designs. The solid curve illustrates the pattern for an antenna with an initial taper 0.4λ long between the 0.822λ diameter and the constant 0.516λ rod and a final taper 0.15λ long to a 0.42λ diameter. The dash curve is the pattern of an antenna with a continuous taper along the length from 0.822λ to 0.42λ. You can see that the first null fills in and the beamwidth increases. The Delrin rod design shows similar results. For short dielectric rod antennas we can use a continuous taper. The pattern for a design 6λ long distorts completely with a continuous taper design.

10-6 HELICAL WIRE ANTENNA [13, Chap. 7]

We can excite a single wire wound in a helix to radiate an end-fire pattern on its axis for a circumference around one wavelength. The axial beam is circularly polarized in the same sense as the helix. In low-gain applications the antenna works over a 1.7 : 1

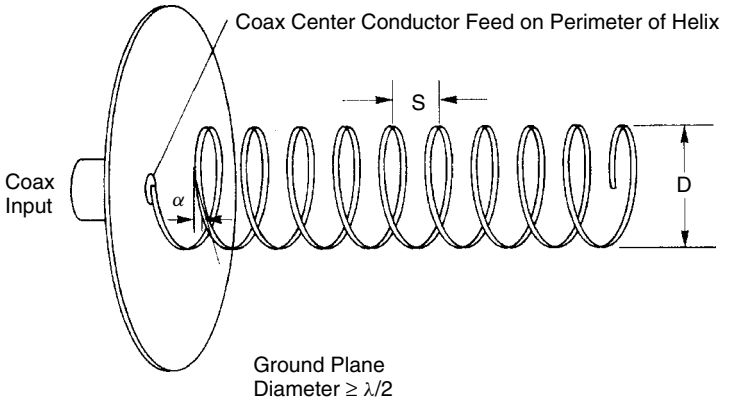


FIGURE 10-17 Axial mode helical wire antenna (RHC).

bandwidth that decreases as gain increases. We see this result from Figure 10-1 for any traveling-wave antenna. Figure 10-17 illustrates the parameters of a helical antenna. The helix has a pitch angle α with spacing S on a diameter D. These parameters are interdependent.

$$\begin{aligned} C &= \pi D && \text{circumference} \\ \tan\alpha &= \frac{S}{C} = \frac{S}{\pi D} \\ L &= \sqrt{S^2 + C^2} = \frac{\pi D}{\cos\alpha} && \text{length of turn} \end{aligned} \tag{10-31}$$

10-6.1 Helical Modes

We arrive at an understanding of antenna operation by considering the modes of the helical transmission line. For diameters small in wavelengths, the wave travels along the wire at the free-space velocity in the T_0 mode. This mode has equal phase points occurring on separate turns. Traveling-wave tube amplifiers use this mode to couple power from an electron beam. The helix slows the axial velocity of the wave by sin α.

The second mode, T_1, occurs when the circumference of the helix approaches one wavelength. The entire phase variation cycles on one turn. The velocity on the wire adjusts for the circumference in this mode. The adjustment gives a velocity whose axial component closely matches the Hansen and Woodyard criterion for increased end-fire directivity over a range of pitch angles.

The third helical transmission-line mode, T_2, has two cycles on each turn. The wave slows as compared with the T_1 mode, but near a circumference of one wavelength, the T_1 mode predominates. We express the general mode distribution by $I_m e^{\pm jm\phi}$, where m is an integer, ϕ the polar angle of the developing helix, and the sign determines the polarization sense ($-$ = right hand).

The T_1 mode radiates its beam peak on the helix axis. The T_0 mode radiates the normal mode, a dipole-shaped pattern, from a combination of small loops and short dipoles arrayed along the axis. It produces circular polarization over a narrow band of

frequencies. The T_2 mode radiates a pattern with a shape like that of the traveling-wave current (Figure 10-5*a*) with a null on its axis.

10-6.2 Axial Mode

Antennas radiating from the T_1 mode have good axial ratio, bandwidth, and the possibility of high gains over narrow bandwidths. Transitions such as the feed and helix end excite other helical transmission-line modes detected in their patterns, but proper construction reduces these anomalies. We analyze the pattern response from a linear array of single-turn helical antennas spaced at the pitch S. Each loop is a traveling-wave current element $I_1 e^{\pm j\phi}$, from $I_m e^{\pm jm\phi}$. We calculate the total pattern from pattern multiplication of the single-turn helical antenna and a uniform-amplitude linear array with the pattern

$$E_0 \frac{\sin(N\psi/2)}{N\psi/2} \tag{10-32}$$

where $\psi = kS\cos\theta + \delta$, δ is the phase shift between array elements, and N is the number of turns. The antenna currents satisfy the Hansen and Woodyard criterion in the range of circumferences 0.78λ to 1.33λ. We can excite the T_1 mode over that entire band for short helixes (lengths less than 2λ), but we can obtain it cleanly for only a limited bandwidth on long helixes. The Hansen–Woodyard criterion determines the axial phase shift: $-\delta = kS + (\pi/N)$. Since the wave travels around the wire, we add 2π to match the T_1 mode. Along the wire we have

$$PkL = Pk\sqrt{C^2 + S^2} = kS + \frac{\pi}{N} + 2\pi \quad \text{or} \quad \frac{PL}{\lambda} = \frac{S}{\lambda} + 1 + \frac{1}{2N} \tag{10-33}$$

Given $L = \pi D/\cos\alpha = C/\cos\alpha$, we define $C_\lambda = C/\lambda$ and substitute it into Eq. (10-33) to obtain the relative propagation constant along the wire a result verified by experiments:

$$\begin{aligned} \frac{PC_\lambda}{\cos\alpha} &= C_\lambda \tan\alpha + \frac{2N+1}{2N} \\ P &= \sin\alpha + \frac{[(2N+1)/2N]\cos\alpha}{C_\lambda} \end{aligned} \tag{10-34}$$

The pattern of a single turn determines polarization along the axis while the array determines the pattern shape. Figure 10-18 gives the pattern of a single turn of a helical antenna with the T_1 traveling-wave current for a five-turn helix. We see a cross-polarization level on a boresight of less than 25 dB, but little front to back when we include both polarizations. The end-fire array forms the beam to produce the total pattern (Figure 10-19). The five-turn helical antenna has the same boresight polarization as the individual turn. The boresight axial ratio improves when we add turns by

$$\text{axial ratio} = \frac{2N+1}{2N} \tag{10-35}$$

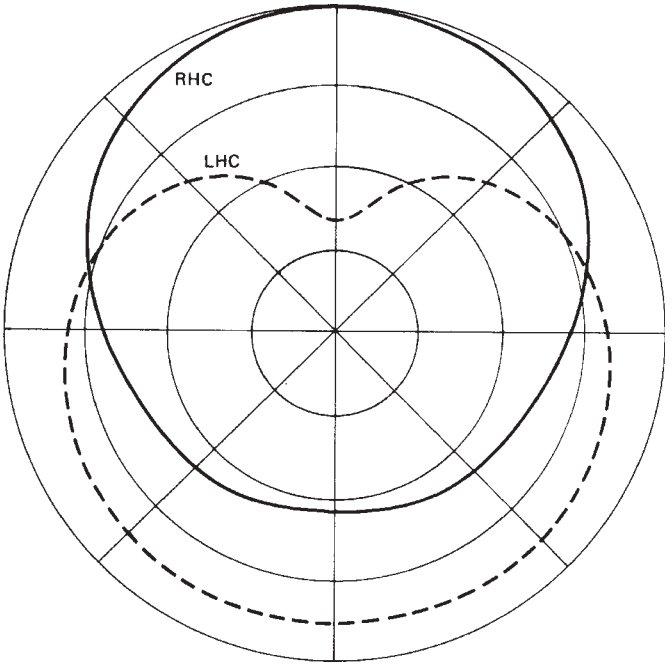


FIGURE 10-18 Single turn of helical wire antenna with a circumference of 0.9λ and $\alpha = 13^\circ$. The relative propagation constant $P = 1.416$ on wire.

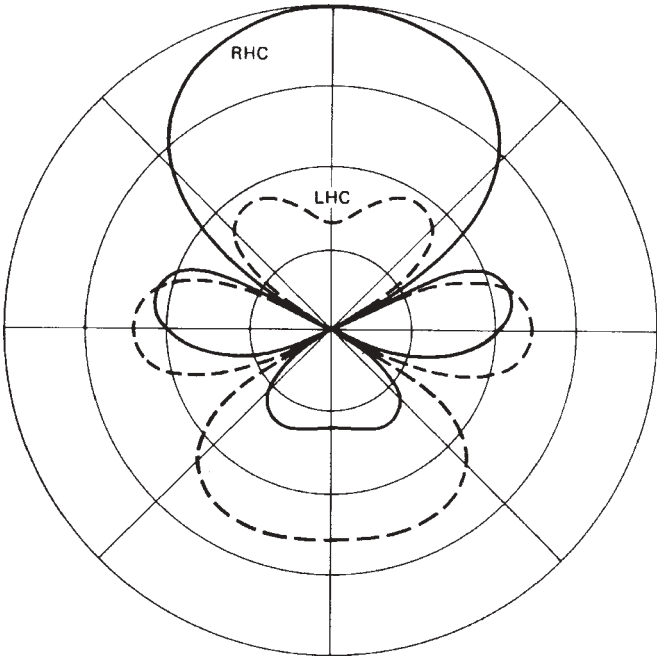


FIGURE 10-19 Five-turn helical wire antenna with a circumference of 0.9λ and $\alpha = 13^\circ$. The relative propagation constant $P = 1.416$ on wire.

for current that follows the increased directivity criterion. The current wave rotates on the helix at nearly the radian frequency that approximately matches a circularly polarized wave. The helical antenna radiates the same sense of polarization as its screw sense. When looking into (from the beam peak) a helical antenna radiating RHC, the currents rotate counterclockwise.

The helical antenna operates best over the following limits:

$$\text{pitch angle:} \quad 12^\circ \le \alpha \le 18^\circ \qquad \text{circumference}/\lambda: \quad 0.78 \le C_\lambda \le 1.33$$

Helical antennas can be built successfully outside the limits of α. The diameter affects the possible directivity slightly, but we determine the directivity of the T_1 mode mainly from the helix length. Figure 10-2 includes a plot of helix directivity. Exciting only the T_1 mode becomes a problem as length increases because the feed region excites other helical transmission-line modes.

10-6.3 Feed of a Helical Antenna

We feed short helical antennas from a coax mounted on a ground plane that should be at least $\lambda/2$ in diameter to achieve a good transition. The pattern does not depend on the ground plane to achieve good F/B because the five-turn helical antenna does not contain a ground plane. The antenna can be fed successfully from a coax twisted in a helix with the outer conductor tapered away in the manner of the split tapered coax balun (Section 5-15.9) until the center conductor alone feeds the helical antenna and a good F/B is achieved [30]. We place the coax feed on the edge of the helix and not on its axis, where an extra length transition must be made to the diameter. When the wire bends into the axis, it generates additional modes on the helix.

Kraus [13] gives an approximate formula for the helical antenna impedance: $R = 141C_\lambda$. The actual value varies about it. We reduce this impedance by soldering a flat strip on the wire from the feed of the helix [31]. The combination of the strip and the ground plane form a parallel-plate line that we space close together to form a tapered impedance transformer to produce a broadband low-impedance input. Empirical adjustments, by distorting the last half- to quarter-turn to taper wire spacing close to the ground plane, can also match the antenna to 50 Ω in a manner similar to the flat plate on the wire. We can design the transition by considering a flat sheet transmission line over a ground plane. The thickness of a dielectric sheet h and line width w necessary to match to the input impedance Z_0 is given by [32, p. 234].

$$h = \frac{w}{(377/\sqrt{\varepsilon_r}Z_0) - 2}$$

For narrowband applications we can use a gamma match feed (Section 5-13) to improve overall system performance. The long helical antenna is grounded and shunt fed. This prevents static charge buildup on the antenna and discharge through the receiver. The gamma match feed narrowbands the antenna and makes it act as a filter.

The feed configuration generates modes other than the T_1 to satisfy the boundary conditions. Similarly, conditions on the end of the helical antenna also generate extra modes. Tapering the last two turns to a diameter 65% of the helix diameter reduces the mode generation from the end [33]. Tapering the diameter in the middle of the helix can also control the antenna's characteristics as we add degrees of freedom.

10-6.4 Long Helical Antenna

A long helical antenna can achieve the directivity predicted in Figure 10-2 over a limited bandwidth. The limited bandwidth comes not so much from the restrictions of Figure 10-1 (because the velocity on the helical transmission line adjusts to the length through coupling) as from the feed structure. Its structure generates extra modes whose radiation greatly reduces directivity. The normal flat ground-plane feed is limited to lengths of about 2λ. Beyond that length the feed generates the T_0 mode that radiates a broad pattern at $\theta = 90^\circ$.

A circular cup around the feed point (Figure 10-20) greatly reduces the excitation of unwanted modes. We determine the cup dimensions empirically. The dimensions given in Table 10-18 have been built or reported. In both cases the directivity fell by 0.2 dB from the curve for the Hansen and Woodyard criterion of Figure 10-2. The cup cannot fully eliminate the generation of extra modes on the helical antenna, since the pattern has dropped by about 0.5 dB from prediction because of the radiation of these extra modes. The 35-turn helical antenna has a 15% bandwidth, and the 50-turn helix has only a 10% bandwidth.

The helical transmission line fails to maintain the Hansen and Woodyard phase criterion over the total range of C_λ as the length increases. The lower values of C_λ retain this property, and the range of acceptable values shrinks as the number of turns increases. As the helix approaches infinite length, the range of acceptable C_λ values approaches 0.78 [35]. At 50 turns on a 13° pitch helical wire antenna, the proper phase velocity is maintained over the range $0.78 \leq C_\lambda \leq 1.0$.

Based on empirical data the peak gain can be expressed as an approximate formula in decibels [34]:

$$G_p = 8.3 \left(\frac{\pi D}{\lambda_p}\right)^{\sqrt{N+2}-1} \left(\frac{NS}{\lambda_p}\right)^{0.8} \left(\frac{\tan 12.5^\circ}{\tan \alpha}\right)^{\sqrt{N}/2}$$

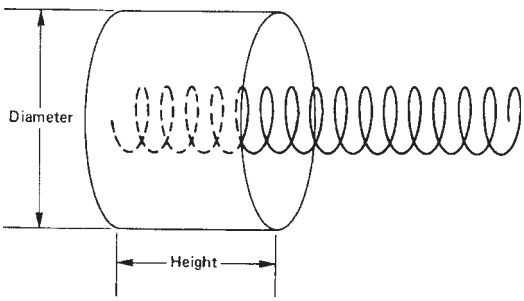


FIGURE 10-20 Cup feed for a long axial mode helical wire antenna.

TABLE 10-18 Cup Feed Dimensions for Long Helical Antennas

N	α	C	L	Cup Diameter	Cup Height
35	12.8°	1.07λ	8.5λ	0.82λ	0.40λ
50	14°	0.90λ	11.2λ	0.97λ	0.85λ

Source: [34]

The peak gain occurs when $\pi D/\lambda \sim 1.135$. Given the ratio of gain to peak gain G/G_p, the frequency range ratio of high/low frequency is expressed as another empirical formula:

$$\frac{f_h}{f_l} = 1.07\left(\frac{0.91}{G/G_p}\right)^{4/(3\sqrt{N})}$$

The beamwidth of a helical wire antenna has been expressed as an empirical formula [36]:

$$\text{HPBW} = \frac{K_B[2N/(N+5)]^{0.6}}{(\pi D/\lambda)^{\sqrt{N}/4}(NS/\lambda)^{0.7}}\left(\frac{\tan\alpha}{\tan 12.5^\circ}\right)^{\sqrt{N}/4}$$

Usual helixes have a constant $K_B = 61.5^\circ$, but helixes constructed on a dielectric rod require a different constant because the dielectric slows the wave on the helix.

10-6.5 Short Helical Antenna [37]

We can mount a short helical wire in a square cavity to make a conformal antenna. As shown in Figure 10-21, the helix has only two turns and the cavity depth is just sufficient for the helix height. The number of turns has minor effects on the pattern, determined primarily by the cavity width. Helixes with pitch angles from 12 to 14° give the best results when the cavity width is at least 0.5λ. The pattern axial ratio improves for larger cavities, and the best results are obtained with an antenna on a flat ground plane. A cavity width of 0.75λ gives good results, with its beamwidths ranging from 45 to 60°.

In the last quarter-turn we reduce the pitch angle to between 2 and 6° to form the tapered transformer to match the feed to 50 Ω. As with the longer helical antennas, we place the feed connection on the perimeter of the helix. Both the impedance match

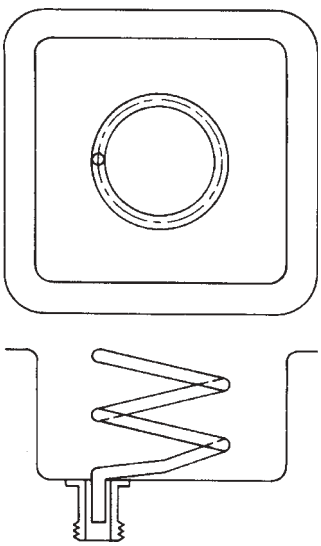


FIGURE 10-21 Two-turn helical wire antenna in a cavity. [From [37], Fig. 1, © 1956 IRE (now IEEE).]

and pattern axial ratio improve compared with the central axis feed. The VSWR and pattern band extend from 0.9 to about 1.7 times the design frequency when the helix circumference is one wavelength. A two-turn helical antenna with a pitch angle of 12° has a cavity depth given by $2\lambda \tan 12^\circ = 0.425\lambda$ for a one-wavelength helix circumference. This antenna is somewhat thicker than the antennas of Chapter 6, but the increased volume increases the bandwidth to nearly an octave.

10-7 SHORT BACKFIRE ANTENNA

A short backfire antenna consists of a round disk $\approx 2\lambda$ diameter with a $\lambda/4$ rim. A dipole (or crossed dipole) is mounted $\lambda/4$ above the disk, and a second small disk $\approx \lambda/2$ diameter is located $\lambda/2$ above the lower disk [38]. The smaller disk acts as a semitransparent obstacle that reflects a portion of the radiation from the dipole. The large disk reflects the signal from the small disk and the ringing between the two increases the effective length of the antenna. This increases the gain to 12 to 14 dB. The general backfire consists of a slow-wave radiator pointed into a ground plane that causes ringing with the reflector element, for example, in a Yagi–Uda dipole antenna. Unfortunately, the short backfire has a bandwidth that ranges between 3 and 5%. Changing the large flat disk into a shallow cone and retaining the outer rim increases the bandwidth but reduces the gain a little [39]. Adding an extra director disk gives the extra parameter needed for impedance matching over a wider band.

Figure 10-22 shows side views of the normal and conical short backfire antennas. Often, the dipole is fed with a sleeve that attaches to a central tube used to support the director disk. We place the balun inside the tube and feed one or two dipoles. The sleeves cannot increase the bandwidth, but they raise the input resistance. Similar to a Yagi–Uda dipole antenna, the director and reflector disks lower the input

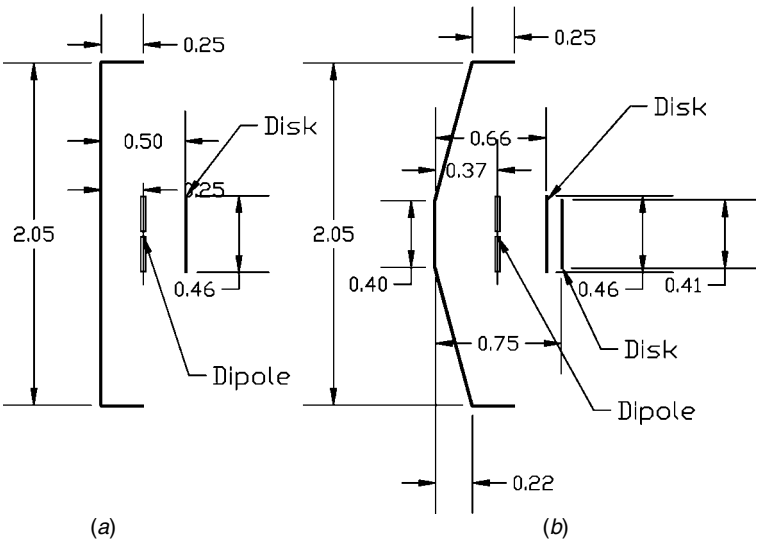


FIGURE 10-22 Short backfire antenna with dimensions in λ: (*a*) original narrowband design; (*b*) conical wideband design.

impedance. The conical reflector increases the input resistance because the cone lowers the interaction between the two disks by reducing one reflection.

We could say that the conical back reflector approximates a parabola and explain its operation in that way, but the director disk is too small to produce an effective ground plane for the feed dipole. Figure 10-23 plots the pattern of the normal short backfire of Figure 10-22*a* when fed from a crossed dipole excited for RHC polarization in the direction of the director disk. The crossed dipole radiates LHC toward the 2.05λ-diameter rimmed reflector that converts LHC to RHC upon reflection. Gain peaks at 14.2 dB, as shown, with a narrow beamwidth determined by the 2.05λ diameter with an F/B ratio of 12 dB (LHC backlobe) and low sidelobes. The conical reflector (15° cone, Figure 10-22*b*) pattern (Figure 10-24) has the same beamwidth as Figure 10-23, but has higher sidelobes and greater radiation in LHC polarization, which produced a pattern with 2 dB less directivity. Figure 10-25 plots the Smith chart of a conical short backfire, showing the dual resonance that produces a 1.6 : 1 VSWR bandwidth of 29%. For comparison a 0.8λ-diameter disk was used as the feed dipole ground plane. Figure 10-26 gives the pattern of the comparison paraboloidal reflector with diameter 2λ fed by a crossed dipole located over a 0.8λ-diameter disk. We read Table 5-1 to determine a suitable reflector f/D (0.34) given the beamwidth and place the reflector focus at the phase center 0.12λ above the ground plane. Gain drops to 10.8 dB because the feed and reflector combination radiates large spread-out sidelobes and significant cross-polarization, but its beamwidth matches that of short backfire antennas.

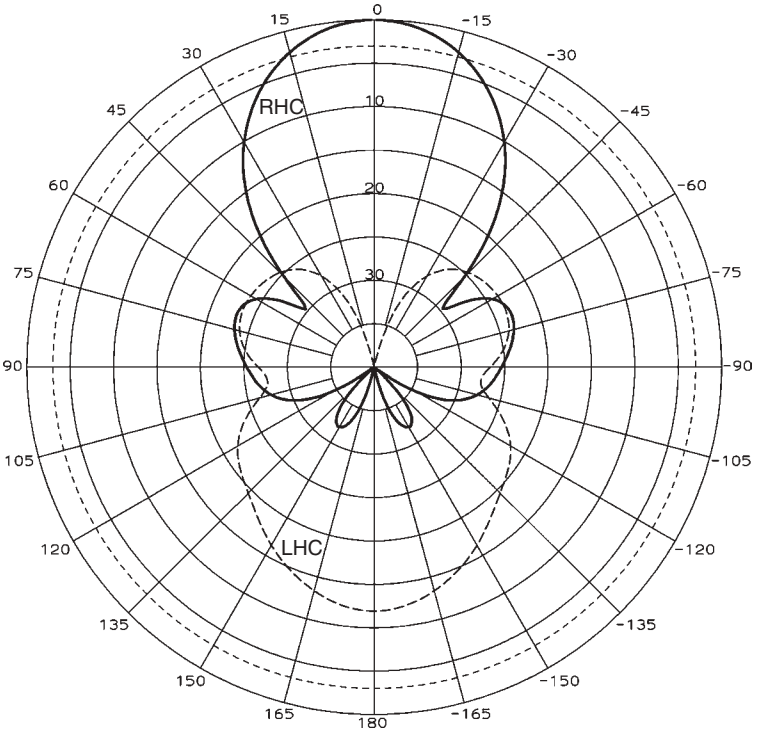


FIGURE 10-23 Pattern of a short backfire antenna of Figure 10-22*a* with RHC crossed-dipole feed.

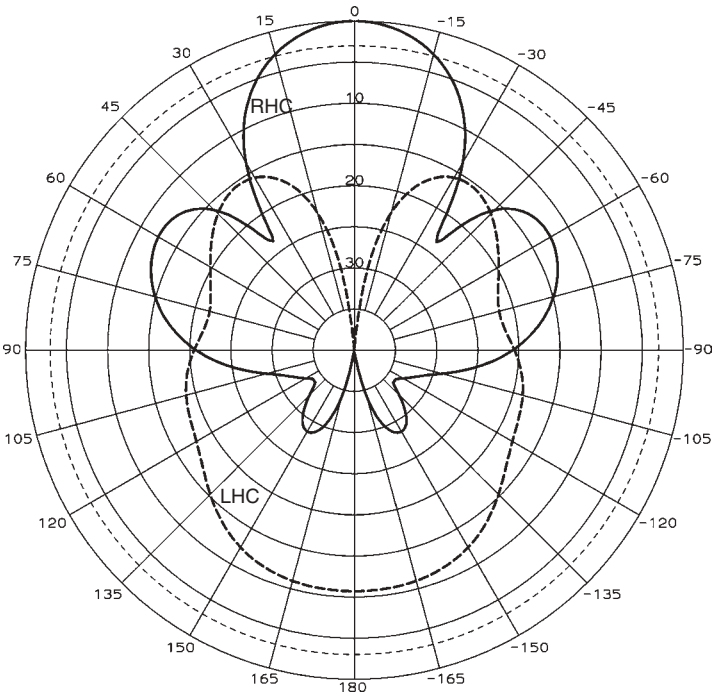


FIGURE 10-24 Pattern of a conical short backfire antenna of Figure 10-22*b* with RHC crossed-dipole feed.

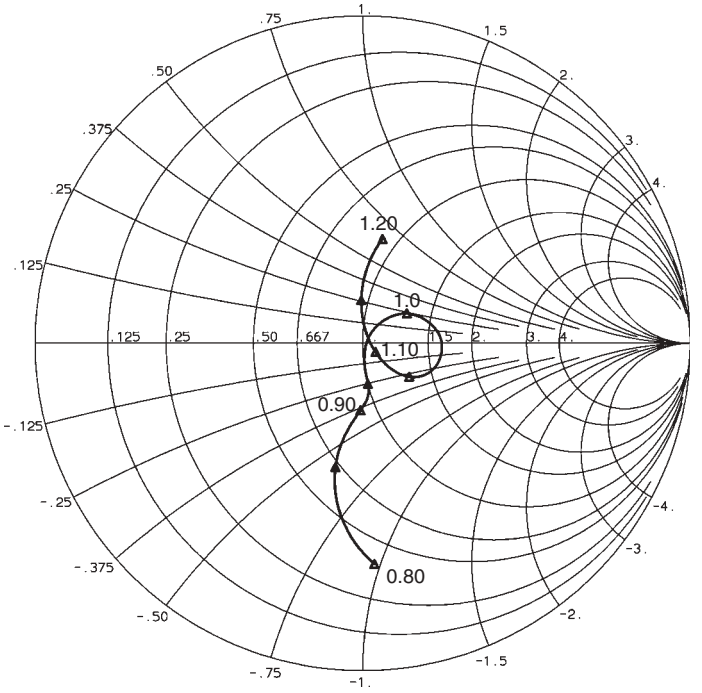


FIGURE 10-25 Normalized impedance response of a conical short backfire antenna (Figure 10-22*b*).

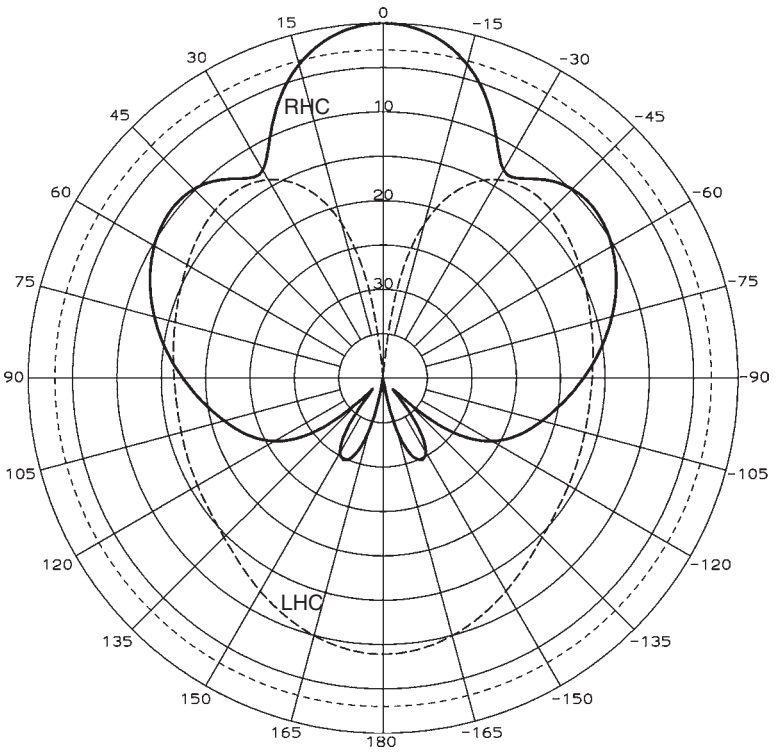


FIGURE 10-26 Pattern of a 2λ-diameter paraboloidal reflector $f/D = 0.34$ with RHC crossed-dipole feed located over a 0.8λ-diameter disk.

We cannot make significant changes to the pattern obtained from a short backfire antenna by varying its parameters. We increase gain by arraying these antennas, but we cannot scan the array because the narrow element beamwidth reduces gain rapidly with scan. Of course, the larger ground plane of the array reduces the backlobe, but in Figure 10-23 it only reduces gain by 0.28 dB. The simple construction is its advantage.

10-8 TAPERED SLOT ANTENNAS

A tapered slot antenna uses a flared slot line etched on a dielectric substrate to produce an end-fire pattern from a surface wave. A single element radiates this end-fire pattern over a wide bandwidth. When we combine a number of these elements in an array, their mutual coupling improves the impedance match when we space them closely. This unique property allows the construction of wide bandwidth arrays with both good impedance properties and the suppression of grating lobes.

Figure 10-27 illustrates four types of tapered slot antennas. The input at the bottom of the figure consists of a either a coaxial-to-slotline transition or a fin-line in a waveguide that feeds into the slotline. In the slotline region the small gap binds the power to the transmission line, and as the transmission line slot widens, it radiates. Similar to the dielectric rod antenna, where we design the initial portion to have waves tightly bound to the rod, the slotline binds waves tightly to it. In the dielectric rod we

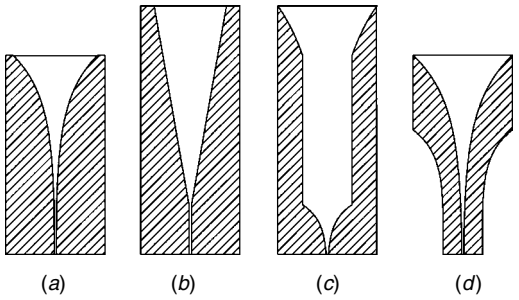


FIGURE 10-27 Tapered slot antennas: (*a*) exponentially tapered slot antenna; (*b*) linear tapered slot antenna; (*c*) continuous-width slot antenna; (*d*) dual exponentially tapered slot antenna.

use a slow wave to attach the waves since values of $P > 1$, the relative propagation constant, cause a rapid attenuation of the fields for movement away from the surface. Instead of slowing the wave, we use the narrow gap of the slotline to bind the wave. Because the tapered slot is etched on a dielectric substrate, the wave is slowed on the slotline. Opening the slot increases radiation.

An exponentially tapered slot antenna (ETSA) [40], also called a *Vivaldi antenna*, radiates nearly equal E- and H-plane beamwidths that change only slightly as frequency increases. The input impedance match is good and the antenna radiates when the final width of the slot is $\geq\lambda/2$. At the lowest frequency the exponential taper is 0.72λ long for an antenna etched on an alumina substrate. The antenna radiates a pattern with a 180° H-plane beamwidth and a 70° E-plane beamwidth for a $\lambda/2$ aperture. For a λ aperture width it radiates E- and H-plane beamwidths of 60° and 70°. The two beamwidths are essentially the same for 1.5λ and larger apertures:

Vivaldi antenna beamwidths, alumina substrate

Aperture (λ)	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6
Beamwidths (deg)	50	42	38	33	31	30	30	30	32	35

The effective radiation region length combined with the substrate slowing determines the H-plane beamwidth of this slow-wave structure. Figure 10-2 shows that the antenna with an aperture of 1.5λ and a length of 2.2λ has its wave slowed by the dielectric to achieve a 50° beamwidth because without dielectric, slowing the beamwidth would be 75°. The flat and increasing beamwidth as frequency increases indicates that the radiation region shrinks. This means that the large slot widths radiate all the power in the wave before the end.

The antenna has high E-plane sidelobes on the order of -5 dB. It appears that the antenna generates a surface wave in the dielectric because it has high cross-polarization that contains 20 to 30% of the radiated power, with its peak cross-polarization occurring in the diagonal plane. We can improve the cross-polarization by removing the dielectric in the slot region, but the beamwidth broadens because the dielectric no longer slows the wave. Even removing a portion, such as a rectangular notch, helps.

We excite this type of tapered slot antenna by using a microstrip-to-slotline transition. We etch a microstrip transmission line on the side opposite the slotted ground plane and run it across the slot. We can either connect the line to a via that runs to the other side of the slotted ground plane or connect it to a fan pattern that produces

a wide-bandwidth short circuit. The fan reduces fabrication cost because it is only an etched element. The open-circuited end of the slotline is terminated with a circular opening to produce a wide-bandwidth open circuit. Unfortunately, this connection generates a surface wave in the dielectric that radiates cross-polarization, due to the imbalance between the feed and the ground plane.

A linear tapered slot antenna (LTSA) radiates a higher gain pattern than that of an ETSA because it relies on length to narrow its beamwidth [41]. The opening angles of these antennas range from 5 to 12°. We use the development in Section 10-1 to determine the gain and beamwidth. The dielectric substrate slows the wave in the slot and increases gain. We calculate an effective dielectric thickness from the equation

$$\frac{t_{\text{eff}}}{\lambda} = (\sqrt{\varepsilon_r} - 1)\frac{t}{\lambda}$$

The optimum values for t_{eff}/λ range between 0.005 and 0.03. Thinner substrates produce insufficient slowing, and thicker substrates produce too much slowing, which breaks up the main beam. Figure 10-28 plots the pattern of an LTSA with and without the dielectric in the tapered region. The pattern of the antenna containing the dielectric has broken up and has high sidelobes. Removing the dielectric increases the beamwidth while decreasing the sidelobes.

A continuous-width slot antenna (CWSA), is similar to a dielectric rod antenna. An initial short taper opens the slot to a uniform-width region where most of the radiation occurs. Sometimes the uniform region then opens to a wider region (Figure 10-27) that terminates the slotline. We design this antenna similar to a dielectric rod and determine its length by the desired gain. This antenna has the narrowest beamwidth and highest gain of the tapered slot antennas for a given length.

The final configuration in Figure 10-27 uses a dual exponential tapered slot (DETSA) to improve the impedance match. The bunny-ear version of this antenna achieves a 2 : 1 VSWR from 0.5 to 18 GHz [42]. This antenna uses a balanced slotline to reduce the cross-polarization. A balanced slotline has the same pattern etched on both sides of the substrate. The feed region consists of a stripline transition, which crosses both slots and is terminated in a centrally located fan to generate the short circuit. The balanced configuration prevents generation of a surface wave in the dielectric. The narrowed-width slotline ground plane improves the impedance match. Both the ground plane increases and the slot widens in the radiation region. The size and pattern of this antenna are similar to those of the Vivaldi, with a reduced mutual coupling because the substrate surface wave is eliminated.

Other feeding configurations of the tapered slot antenna include the antipodal Vivaldi antenna [43] and the balanced antipodal antenna [44]. These are not slot antennas but are similar in shape and performance to the Vivaldi antenna. The antipodal antenna places the two sides of the Vivaldi exponential taper on opposite sides of the substrate. We feed one side directly from the microstrip line and rapidly taper the ground plane symmetrically to form a twin plate line. To form the radiation portion, the twin plate line is spread into the dual exponentially radiator similar to a bunny-ear antenna. Unfortunately, the offset between the two sides causes significant cross-polarization, higher than that of the Vivaldi antenna, etched on a substrate. The advantage of the antenna is the easy transition from microstrip to radiator. By using a balanced antipodal design fed from a stripline, cross-polarization can be reduced while retaining the simple construction. This three-layer structure feeds one ear from the center conductor and

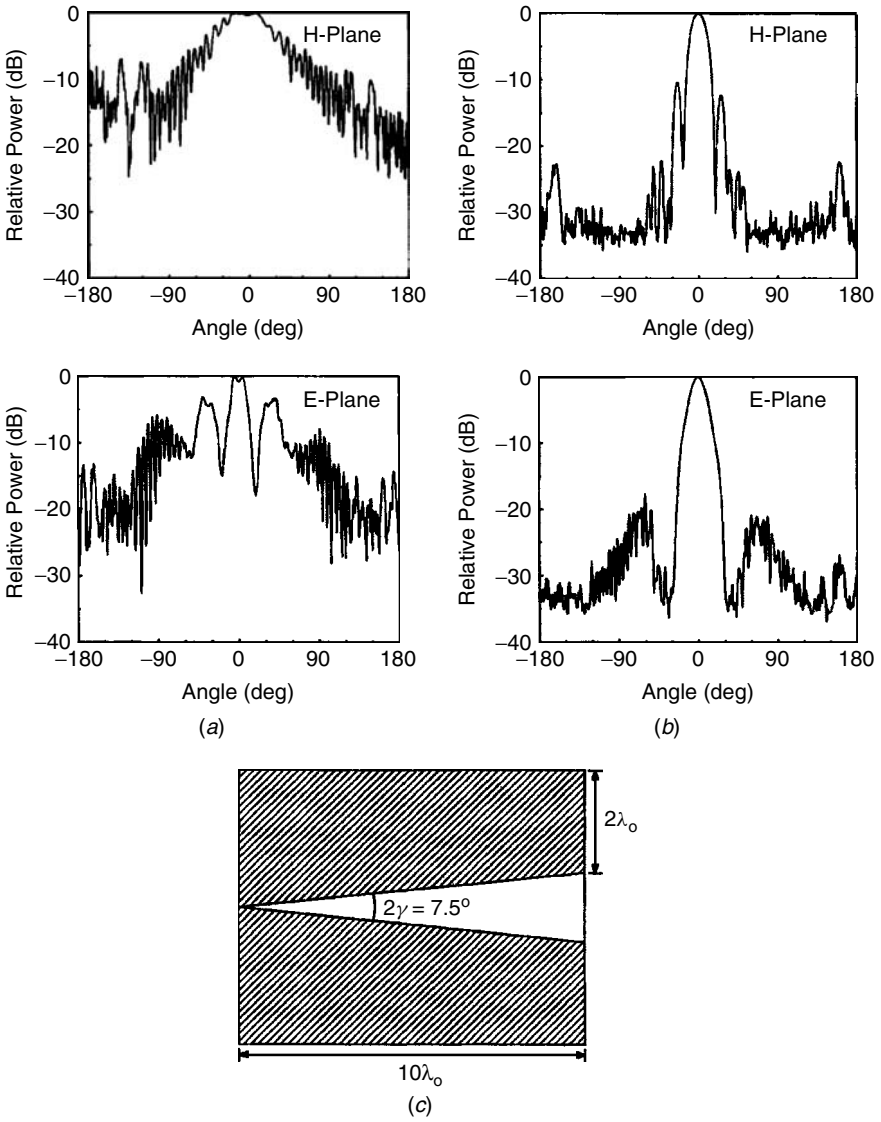


FIGURE 10-28 LTSA patterns: (*a*) with $\varepsilon_r = 10$ substrate in slot; (*b*) without dielectric in slot; (*c*) outline of the antenna. (From [41], Fig. 5, © 1989 IEEE.)

balanced ears located on top of each other from the two ground planes. The design tapers the two ground planes until a balanced triplate line is formed before widening the center conductor and ground planes by using the dual exponential design. The balanced configuration reduces cross-polarization by 15 to 20 dB.

Tapered slot antennas have unique properties useful for wide bandwidth arrays. Although a single-element radiator requires the slot opening to be at least $\lambda/2$, in an array the mutual coupling allows this to be reduced to 0.1λ [45]. The scan impedance produces a low-VSWR array measured to an array scan angle of 50° over a 5.9 : 1 bandwidth. Figure 10-29 shows an antenna built with tapered slot antennas in an egg

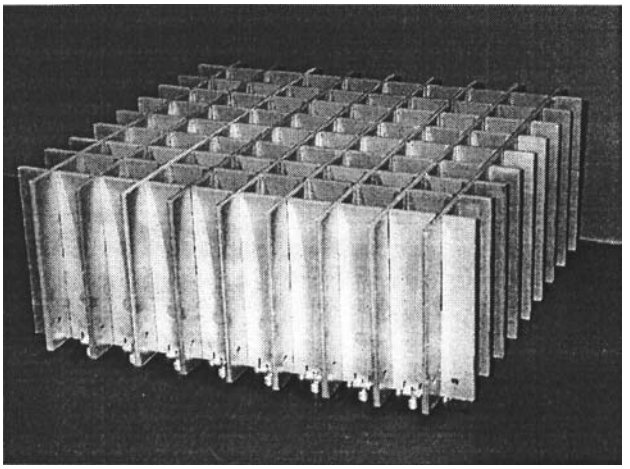


FIGURE 10-29 Square array of 64 tapered slot antennas. (From [45], Fig. 1, © 2000 IEEE.)

crate configuration to radiate dual linear polarization. To prevent grating lobes, the elements must be limited to a little larger than a $\lambda/2$ separation at the highest operating frequency. Remember that the length of the antennas will not increase the element gain because it cannot exceed the effective area associated with each element. The length is to improve the impedance match and to increase the mutual coupling that allows the small slot widths.

Figure 10-30 illustrates construction of an ETSA element in the array. It uses a balanced slotline construction that has matching slots on both sides of the substrate. Halfway between the two sides a strip transmission line feeds the two slots when terminated by a fan that produces a short circuit. A circular open circuit etched in the two sides terminates the slotline. The stripline feeder contains a tapered transformer to transition between the 50-Ω input and the 82-Ω balanced slotline. The mutual coupling between the elements produces narrow-frequency bandwidth resonances in the array that can be suppressed by adding vias between the two sides of the balanced slotline along the edges and around the stripline feeder.

10-9 LEAKY WAVE STRUCTURES

A leaky wave structure allows for the separate specification of the amplitude distribution and beam direction. We position the beam direction by adjusting the relative propagation constant [Eq. (10-7)]. The surface-wave structure permits only small modifications of the amplitude distribution for the pattern scanned to end fire. The uniform distribution on an axisymmetrical leaky wave structure has a directivity of $2L/\lambda$ independent of scan angle until the pattern approaches end fire. Near end fire the cone beamwidths join to produce higher directivities. Figure 10-1 shows the modification of directivity as these join ($P \to 1$). The directivity is nearly constant below $P = 0.9$ for lengths 8λ and 10λ.

Tapered amplitude distributions reduce directivity by the amplitude taper efficiency [Eq. (4-6)]. The relative propagation constant along the structure determines the beam

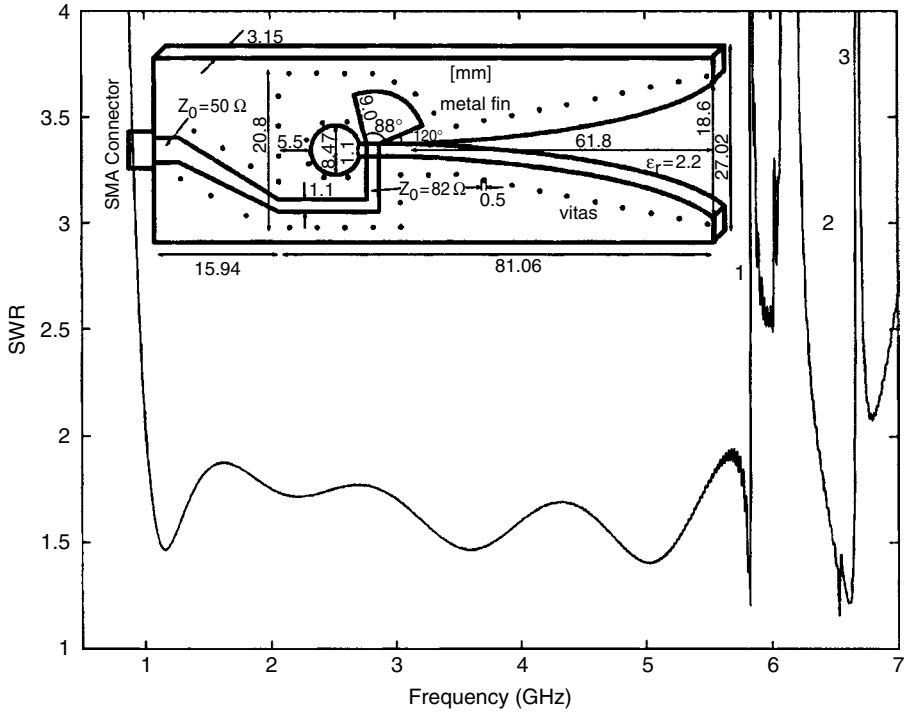


FIGURE 10-30 Numerically calculated VSWR for an infinite array comprised of "contour" via elements scanned to 50° in the H-plane. The first grating lobe appears at 6.29 GHz. (From [45], Fig. 3, © 2000 IEEE).

direction [Eq. (10-7)] for small values of attenuation per unit length. We use Eq. (4-44) for the phase error efficiency due to the varying P along the antenna by replacing $\sin\theta_{\max}$ by $\cos\theta_{\max}$:

$$\text{PEL} = \frac{\left|\int_0^L E(z)e^{jk\cos\theta_{\max}z}dz\right|^2}{\left[\int_0^L |E(z)|dz\right]^2} \tag{10-36}$$

where $E(z)$ is the voltage distribution along the line source. We can use values other than $\theta_{\max}$ in Eq. (10-36) and evaluate the pattern at any θ. When we join a number of leaky wave line sources into a planar array, scanning the beam causes a reduction in the effective aperture length. We reduce the gain by a projection of the length in the scan direction.

Fast-wave structures have cutoff frequencies that modify the relative propagation constant:

$$P = \sqrt{1 - \left(\frac{f_c}{f}\right)^2} = \sqrt{1 - \left(\frac{\lambda}{\lambda_c}\right)^2} \tag{10-37}$$

where f_c is the cutoff frequency and λ_c is the cutoff wavelength. As frequency increases, P approaches 1 and scans the beam toward end fire. A waveguide can be

made into a leaky wave antenna by openings that cut the wall currents. For example, a slot in the center of the broadwall of a rectangular waveguide does not radiate because it cuts no net current. If we vary the distance of a continuous slot from the centerline, we can control the leakage rate. We form other leaky wave antennas by cutting closely spaced holes or slots. We analyze these discrete radiators as continuous structures whose radiation rate is controlled by the size of the holes or the placement of the slots.

REFERENCES

1. W. W. Hansen and J. R. Woodyard, A new principle in antenna design, *Proceedings of IRE*, vol. 26, no. 3, March 1938, pp. 333–345.
2. C. H. Walters, *Traveling Wave Antennas*, Dover, New York, 1970.
3. H. W. Ehrenspeck and H. Poehler, A new method for obtaining maximum gain from Yagi antennas, *IRE Transactions on Antennas and Propagation*, vol. AP-7, no. 4, October 1959, p. 379.
4. F. J. Zucker, Chapter 16 in H. Jasik, ed., *Antenna Engineering Handbook*, McGraw-Hill, New York, 1961.
5. R. F. Harrington, *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York, 1961.
6. H. H. Beverage, C. W. Rice, and E. W. Kellogg, The wave antenna: a new type of highly directive antenna, *AIEE Transactions*, vol. 42, 1923, p. 215.
7. K. Lizuka, The traveling-wave V-antenna and related antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-15, no. 2, March 1967, pp. 236–243.
8. G. A. Thiele and E. P. Ekelman, Design formulas for vee dipoles, *IEEE Transactions on Antennas and Propagation*, vol. AP-28, no. 4, July 1980, pp. 588–590.
9. A. E. Harper, *Rhombic Antenna Design*, Van Nostrand, New York, 1941.
10. H. Yagi, Beam transmission of ultra short waves, *Proceedings of IRE*, vol. 26, June 1928, pp. 715–741.
11. R. J. Mailloux, Antenna and wave theories of infinite Yagi–Uda arrays, *IEEE Transactions on Antennas and Propagation*, vol. AP-13, no. 4, July 1965, pp. 499–506.
12. P. S. Carter, Circuit relations in radiating systems and applications to antenna problems, *Proceedings of IRE*, vol. 20, 1932, p. 1004.
13. J. D. Kraus, *Antennas*, McGraw-Hill, New York, 1950.
14. R. C. Hansen, ed., *Microwave Scanning Antennas*, Vol. II, Academic Press, New York, 1966.
15. H. E. King, Mutual impedance of unequal length antennas in echelon, *IRE Transactions on Antennas and Propagation*, vol. AP-5, July 1957, pp. 306–313.
16. J. H. Richmond, Coupled linear antennas with skew orientation, *IEEE Transactions on Antennas and Propagation*, vol. AP-18, no. 5, September 1970, pp. 694–696.
17. C. A. Thiele, Analysis of Yagi–Uda type antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-17, no. 1, January 1969, pp. 24–31.
18. D. K. Cheng and C. A. Chen, Optimum element spacings for Yagi–Uda arrays, *IEEE Transactions on Antennas and Propagation*, vol. AP-21, no. 5, September 1973, pp. 615–623.
19. C. A. Chen and D. K. Cheng, Optimum element lengths for Yagi–Uda arrays, *IEEE Transactions on Antennas and Propagation*, vol. AP-23, no. 1, January 1975, pp. 8–15.
20. D. Kajfez, Nonlinear optimization extends the bandwidth of Yagi antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-23, no. 2, March 1975, pp. 287–289.
21. P. P. Viezbicke, Yagi antenna design, *NBS Technical Note 688*, U.S. Department of Commerce/National Bureau of Standards, December 1976.

22. W. K. Kahn, Double ended backward-wave Yagi hybrid antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-29, no. 3, May 1981, pp. 530–532.
23. J. E. Lindsay, A parasitic endfire array of circular loop elements, *IEEE Transactions on Antennas and Propagation*, vol. AP-15, no. 5, September 1967, pp. 697–698.
24. L. C. Shen and G. W. Raffoul, Optimum design of Yagi array of loops, *IEEE Transactions on Antennas and Propagation*, vol. AP-22, no. 6, November 1974, pp. 829–831.
25. A. Shoamanesh and L. Shafai, Design data for coaxial Yagi array of circular loops, *IEEE Transactions on Antennas and Propagation*, vol. AP-27, no. 5, September 1979, pp. 711–713.
26. S. A. Brunstein and R. F. Thomas, Characteristics of a cigar antenna, *JPL Quarterly Technical Review*, vol. 1, no. 2, July 1971, pp. 87–95.
27. J. L. Wong and H. E. King, A wide-band low-sidelobe Disc-o-Cone antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-31, no. 1, January 1983, pp. 183, 184–.
28. V. C. Smits, Rear gain control of a dielectric rod antenna, *Microwave Journal*, vol. 11, no. 12, December 1968, pp. 65–67.
29. S. Kobayashi, R. Mittra, and R. Lampe, Dielectric tapered rod antennas for millimeter applications, *IEEE Transactions on Antennas and Propagation*, vol. AP-30, no. 1, January 1982, pp. 54–58.
30. B. A. Munk and L. Peters, A helical launcher for the helical antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-16, no. 3, May 1968, pp. 362–363.
31. J. D. Kraus, A 50-ohm input impedance for helical beam antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-25, no. 6, November 1977, p. 913.
32. J. D. Kraus and R. J. Marhefka, *Antennas for All Applications*, 3rd ed., McGraw-Hill, New York, 2003.
33. J. L. Wong and H. E. King. Broadband quasi-taper helical antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-27, no. 1, January 1979, pp. 72–78.
34. H. E. King and J. L. Wong, Characteristics of 1 to 8 wavelength uniform helical antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-28, no. 3, March 1980, pp. 291–296.
35. T. S. M. Maclean and R. G. Kouyoumjian, The bandwidth of helical antennas, *IRE Transactions on Antennas and Propagation, Symposium Supplement*, vol. AP-7, December 1959, pp. S379–S386.
36. J. L. Wong and H. E. King, Empirical helical antenna design, *Digest of the International Symposium on Antennas and Propagation*, 1982, pp. 366–369.
37. A. Bystrom and D. G. Berntsen, An experimental investigation of cavity-mounted helical antennas, *IRE Transactions on Antennas and Propagation*, vol. AP-4, no. 1, January 1956, pp. 53–58.
38. H. W. Ehrenspeck, The short backfire antenna, *Proceedings of IEEE*, vol. 53, no. 4, August 1965, pp. 1138–1140.
39. S. Ohmori et al., An improvement in electrical characteristics of a short backfire antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-31, no. 4, July 1983, pp. 644–646.
40. P. J. Gibson, The Vivaldi aerial, *Proceedings of the 9th European Microwave Conference*, Brighton, East Sussex, England, 1979, pp. 101–105.
41. K. S. Yngvesson et al., The tapered slot antenna: a new integrated element for millimeter-wave applications, *IEEE Transactions on Microwave Theory and Techniques*, vol. 37, no. 2, February 1989, pp. 365–374.
42. J. J. Lee and S. Livingston, Wide band bunny-ear radiating element, *IEEE AP-S Symposium*, 1993, pp. 1604–1607.

43. E. Gazit, Improved design of the Vivaldi antenna, *IEE Proceedings*, vol. 135, pt. H, no. 2, April 1988, pp. 89–92.

44. J. D. S. Langley, P. S. Hall, and P. Newham, Novel ultrawide-bandwidth Vivaldi antenna with low cross polarization, *Electronics Letters*, vol. 29, no. 23, November 11, 1993, pp. 2004–2005.

45. H. Holter, T. -H. Chio, and D. H. Schaubert, Experimental results of 144-element dual-polarized endfire tapered-slot phased arrays, *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 11, November 2000, pp. 1707–1718.

11

FREQUENCY-INDEPENDENT ANTENNAS

We derive the idea of self-scaling or frequency-independent antennas from the principle of frequency scaling used in model measurements. As we decrease the wavelength (increase frequency), we decrease the model size in the same proportion. To build wideband antennas, we need structures that can be their own scale models. One approach is to remove any characteristic length by specifying the antenna only in terms of angles [1]. This method leads to the continuously scaled spiral antennas. A second approach is to include antenna parts that scale a portion of the antenna exactly at discrete frequency intervals. We scale these parts logarithmically so that the intervals between frequencies of perfect scaling grow with frequency. These log-periodically scaled antennas have varying characteristics between the points of scaling whose ripple decreases as the scaling constant approaches 1 (continuous scaling), but the number of parts increases.

A continuous or log-periodically scaled structure has no ends, but we must be able to truncate a successful frequency-independent antenna with little effect on the pattern. A self-scaling antenna must be a transmission-line structure that delivers power to an active region where we feed the high-frequency end, and it serves as a transmission line for the lower-frequency portion. The currents must decay after a radiating active region so that the structure can be stopped without adversely affecting antenna properties. We identify the finite active region by truncation constants that we use to size the design. Although the spiral radiates most of the input power in a finite active region, we improve the pattern by loading the ends of the arms to prevent radiation of currents flowing in the reverse direction and accept loss.

The active region radiates most of the power for a particular frequency. A true frequency-independent antenna has a constant beamwidth over its band, although we expect small variations between the frequencies of exact scalings (log periodic). We

Modern Antenna Design, Second Edition, By Thomas A. Milligan

obtain constant beamwidths only if the active-region dimensions scale with wavelength. The truncation requirement affects the pattern. A self-scaling antenna cannot radiate in the direction of the expanding structure. If the antenna did radiate in that direction, portions of the structure would be excited in higher-order modes beyond the normal truncation point. Log-periodic and conical logarithmic spirals backfire toward the feed point.

We can make a structure with logarithmically scaled radiating parts along a transmission line and still not achieve a successful broadband antenna. The parts must couple electromagnetically, not just through the connection of the feeder. We place dipoles of the log-periodic dipole antenna close together to produce the coupling needed for rapid attenuation. Similarly, we closely space the turns of a spiral so that the arms couple and there is sufficient length in the active region along which to radiate. Usually, we can account for the rapid attenuation of currents through loss of power in radiation by considering a single mode.

A successful self-scaling antenna structure satisfies these requirements [2]:

1. The antenna contains its own scale model parts—continuous or discrete—that can be scaled to an infinitesimal size.
2. The antenna radiates most of the power in a finite active region so that it can be terminated with minimal effect.
3. Fed from the high-frequency end, the antenna must be a transmission line to carry power to the low-frequency end.
4. The dimensions of the active region must scale with wavelength.
5. The antenna must not radiate in the direction of expanding structure.
6. The parts must have significant direct coupling outside the transmission-line feeder.

SPIRAL ANTENNAS

Spiral antennas consist of a thin metal foil spiral pattern etched on a substrate, usually fed from the center, and located over a cavity. The etching contains a symmetrical pattern of at least two arms, but we build spiral antennas with more arms to radiate in multiple modes or to suppress unwanted modes. The two-arm version can be fed using a simple balanced line requiring a balun. With more arms we need a feed network called a *beamformer*, which contains an output port for each spiral arm and a separate input for each spiral mode. This network ideally divides the power into equal-amplitude outputs with a linear phase progression between them. The phase progression of each spiral mode cycles one or more times through all phases, and the complex summation of any set of mode voltages equals zero. The number of complete rotations through all phases equals the mode number; mode 1 has one cycle, 2π radians; mode 2 has 2 cycles, 4π radians; and so on.

Figure 11-1 plots the measured pattern of an eight-arm equiangular spiral operating in mode 1 obtained from single-arm measurements added using an ideal feed network (beamformer). The pattern shows RHC polarization as co-polarization and LHC polarization as cross-polarization. This is the most commonly used mode. When we change the number of arms and feed them properly with mode 1, we obtain similar patterns.

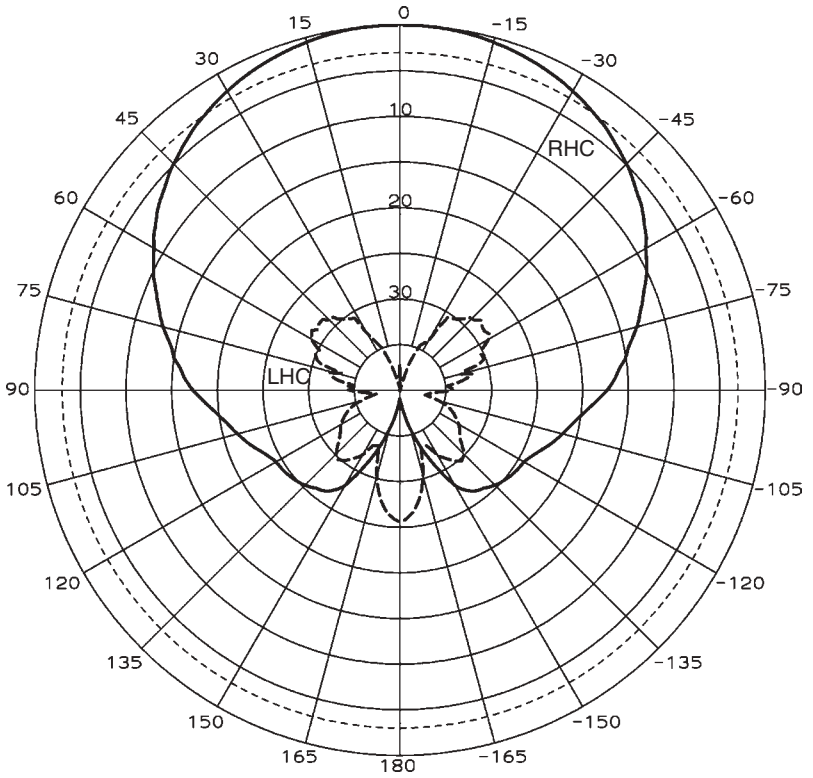


FIGURE 11-1 Measured pattern of an eight-arm spiral operating in mode 1 using an ideal beamformer.

The extra arms aid the pattern symmetry, because they suppress radiation of higher-order modes that distort the pattern. By building a spiral with more than two arms we can intentionally excite radiation of higher-order modes. Figure 11-2 gives the patterns of the eight-arm spiral radiating mode 2 when fed from an ideal beamformer. When we plot the pattern for mode 3, it is similar to the mode 2 except that the beam peak occurs at an angle farther from broadside. All higher-order modes of the spiral have a beam shape similar to mode 2 except that the beam peak angle continues to increase with mode number.

Each spiral mode radiation has a phase rotation of a full 2π radian cycles for conical patterns, which equals the mode number. A conical pattern measurement rotates ϕ about the spiral plane normal axis while holding θ constant. For example, we could measure patterns at $\theta = 45^\circ$ on the patterns of Figures 11-1 and 11-2 near the peaks of the higher-order modes. We determine the mode radiating by the phase slope. RHC polarization produces a negative slope as ϕ increases [counterclockwise (CCW) rotation]. We use the convention that positive modes radiate RHC and negative modes radiate LHC and place the negative sign in the mode expressions. Increasing the angle is similar to increasing the distance in terms of the radiation phase. By exciting the spiral in adjacent mode numbers, the phase difference between them can be used to determine the ϕ angle of arrival when one mode supplies the reference signal for phase measurement.

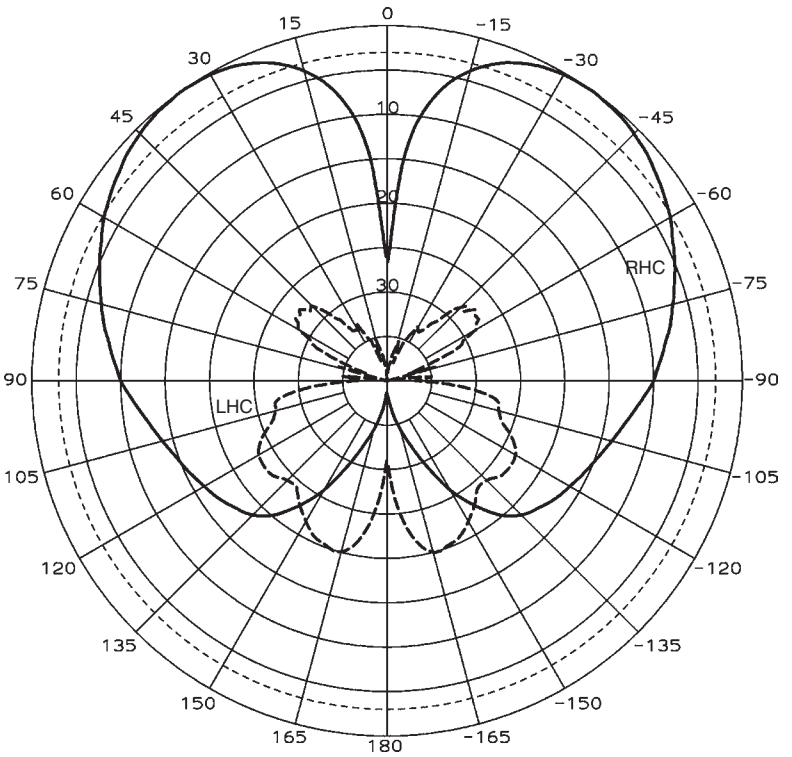


FIGURE 11-2 Measured pattern of an eight-arm spiral operating in mode 2 using an ideal beamformer.

We use multiple modes on the spiral for an angle of arrival (AOA) system by comparing two modes. The amplitude difference between modes 1 and 2 determines the angle off the axis of the spiral plane. Off-axis, where the higher-order modes radiate, we measure a phase progression equal to the mode; for example, mode 2 radiation changes by 720° as we rotate the antenna one revolution in a conic pattern. Given the mode number, the phase varies by $-m360^\circ$ during one revolution. If we use mode 1 as a phase reference, the phase of mode 2 relative to it changes -360° during one revolution about the spiral plane axis. By using both amplitude and phase, the two angles of arrival can be determined. Although a three-arm spiral will support both modes 1 and 2, necessary for AOA, we use a four-arm spiral because it uses a simpler feed network [3].

11-1 MODAL EXPANSION OF ANTENNA PATTERNS

We use a modal or Fourier series expansion of conical patterns for the analysis and measurement of spirals. A conical pattern varies ϕ while holding θ constant in the spherical coordinate system and is the AUT or head axis of a model tower positioner. Each term in the expansion includes the modal phasing of a circularly polarized signal. For example, the mode 1 phase has a single rotation -2π in the full CCW rotation of a conical pattern, mode 2 has a $-2(2\pi)$ phase rotation, and so on.

The mode number of a spiral refers to the number of 2π (radians) or 360° (degrees) cycles that occur in the feed phasing when progressing through the arms CCW. Mode 1 phases in a two-arm spiral are 0° and 180°. Cycling to the first input adds another 180°, to give 360° around the spiral arms. The phase difference moving CCW between arms is found from the mode number m and the number of arms N:

$$\text{phase} = -\frac{2\pi m}{N} \quad \text{or} \quad -\frac{360^\circ m}{N} \tag{11-1}$$

The spiral radiates RHC polarization for $m = 1$ using the notation of Eq. (11-1).

An axially symmetrical antenna such as a spiral can radiate these modes when we phase the feeding of the ports to match the phase rotation of the mode. We must add spiral arms to feed higher-order modes. Modes $+1$ and -1 produce the same phasing at the feed points of a two-arm spiral: 0° and 180°, and the spiral wrap direction determines the polarization radiated. All odd-order ($\ldots, -3, -1, 1, 3, 5, \ldots$) modes have the same phasing on two feeds, which means that the two-arm spiral will radiate these modes efficiently if current flows on the arms where the spiral circumference is the same integer number of wavelengths. The number of arms equals the number of independent modes, although the zero mode is difficult to use. You should expand Eq. (11-1) for various numbers of arms and modes. The sense of the spiral wrap and the direction of current flow determine the circular polarization sense. On an N-arm structure the modes have a cyclic variation, so that, for example, mode 3 of a four-arm spiral has phasing equivalent to $m = -1$ (i.e., mode 1 LHC polarization). In other words, the modes are modulo 4 with mode 3 equal mode -1. When feeding the right-hand four-arm spiral with mode 3 from the center, the spiral fails to radiate from the currents flowing out the arms because each arm radiates RHC while the feed phasing is LHC, and they cancel. The spiral acts as a circular polarization filter. These currents reflect from the ends of the open-circuited arms, travel inward, and radiate LHC polarization. On a two-arm spiral, whenever the circumference is an odd-integer multiple of a wavelength, the currents radiate. Increasing the number of arms to four from two reduces the number of modes radiating because modes 3, 7, and so on, phases no longer match the feeding phases of the arms. Only modes 5, 9, ... phases match mode 1.

When you apply Eq. (11-1) to an eight-arm spiral you will discover that for a rotation of five-cycles (mode 5) the phase shift between ports -225° equals the phasing of mode -3 135° and shows the modulo characteristic of modes. An eight-arm spiral has the following equivalences: mode -3 = mode 5, mode -2 = mode 6, mode -1 = mode 7. Although we can divide the feed voltages on an N-arm spiral into N orthogonal modes, the antenna can radiate any mode. For example, if we feed the center of a two-arm spiral in mode 1, it will radiate a large portion of the power when its circumference is 1λ. When the circumference is 2λ, the spiral would radiate mode 2 except that the 180° phase difference between the two arms cancels this radiation. The antenna radiates some power in mode 2 if the feed voltage balance and phasing is not perfect. Mode 3 radiates when the circumference is 3λ and radiates part of the remaining power on the two spiral arms because the feed phases on the arms do not cancel. If the spiral is large enough to have a circumference of 5λ, mode 5 will radiate from the two-arm spiral. Mode 4 radiation is canceled by the arm phasing. Given a spiral with N arms, the modes that have significant radiation are a multiple of the number of arms when fed with a perfect beamformer network [3]:

$$m_{\text{radiated}} = m + kN \qquad k = \ldots, -2, -1, 0, 1, 2, \ldots \tag{11-2}$$

N-arm spiral suppresses $N-1$ modes between possible modes given that the spiral circumference is large enough to support a particular mode. For example, a six-arm spiral excited in mode 1 will radiate modes 1, 7, 13, ..., −5, −11, ... and when excited in mode 2 it will radiate modes 2, 8, 14, ..., −4, −10, ..., and so on. If the antenna radiates sufficient power in the lower-order modes, little is left for the higher-order mode radiation and the patterns will be fine. Increasing the number of arms reduces the number of spiral modes radiated.

Given a conical pattern $F(\phi)$, we expand it in the Fourier series of modes and easily compute the expansion coefficients from an integral performed numerically on a measured pattern:

$$F(\phi) = \sum_{m=-\infty}^{m=\infty} E_m e^{-jm\phi} \qquad \text{where} \quad E_m = \frac{1}{2\pi}\int_0^{2\pi} F(\phi) e^{jm\phi}\, d\phi \tag{11-3}$$

Each conical pattern has its own set of modal coefficients $E_m(\theta)$ for each polarization. The polarizations are pairs of orthogonal polarizations, such as (E_θ, E_ϕ) or (E_{RHC}, E_{LHC}). We apply these modal coefficients when testing or analyzing spiral antennas as a measure of performance. We use an integral over the entire radiation sphere to determine the relative power in each mode:

$$P_m = \frac{\displaystyle\int_0^{\pi} \left[\left| \int_0^{2\pi} E_{\text{RHC}}(\theta,\phi) e^{jm\phi}\, d\phi \right|^2 + \left| \int_0^{2\pi} E_{\text{LHC}}(\theta,\phi) e^{jm\phi}\, d\phi \right|^2 \right] \sin\theta\, d\theta}{\displaystyle\int_0^{\pi}\int_0^{2\pi} [|E_{\text{RHC}}(\theta,\phi)|^2 + |E_{\text{LHC}}(\theta,\phi)|^2] \sin\theta\, d\phi\, d\theta} \tag{11-4}$$

Equation (11-4) is written in terms of the RHC and LHC polarizations, but we can substitute any other pair of orthogonal polarizations and use the same formula. The denominator is proportional to the total power radiated by the antenna.

11-2 ARCHIMEDEAN SPIRAL [4, 5]

Although Archimedean and exponential spirals have different equations defining them, practice shows that their characteristics do not differ by very much. The Archimedean spiral arm lengths can be long and produce high circuit losses at low frequencies. The wrap angle of an Archimedean spiral changes from a high value in the center to a low value on the outside. The high wrap rate in the center excites more higher-order modes at high frequencies. The low wrap rate at the outer diameter improves pattern shape at low frequencies. Although an exponential spiral has more uniform characteristics over the entire frequency range, an Archimedean spiral is useful.

The Archimedean spiral radius increases uniformly with angle:

$$r = r_0 + a\phi \tag{11-5}$$

where r_0 is the initial radius and a is the growth rate. We cannot scale the structure to an infinitesimal size by using Eq. (11-5), one of the requirements of frequency-independent antennas. Figure 11-3 shows two shapes of Archimedean spirals. We

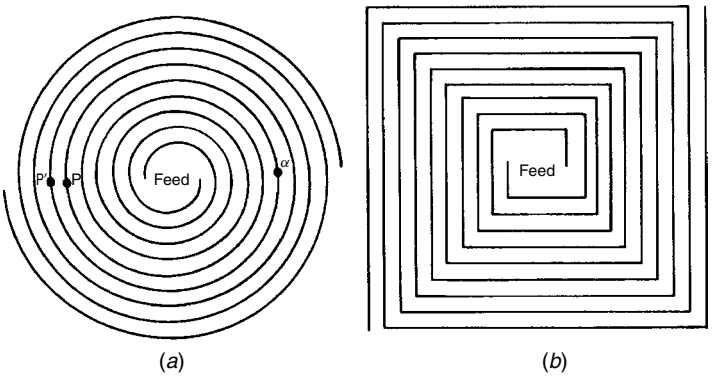


FIGURE 11-3 Archimedean spirals: (*a*) LHC; (*b*) RHC.

usually make the antenna complementary—the uniform-width metal strip equals the spacing between strips. A two-arm infinite structure has an impedance of 188 Ω from the Babinet–Booker principle (Section 5-3) for a self-complementary structure.

A balanced line feeds the spiral from the center. The radiations from the nearly equal and opposite currents at the feed point cancel in the far field. The growing spiral arms separate the currents. When the perimeter of the turn approaches one wavelength, the out-of-phase currents at P and Q (Figure 11-3) become in phase at points P and P', and radiation from the currents no longer cancel in the far field. This condition starts somewhat before the 1λ perimeter point and continues for some distance after it. To radiate efficiently the antenna should have a perimeter of 1.25λ at the lowest operating frequency. The upper-frequency truncation size is determined by a requirement to limit the spacing between feed points to less than $\lambda/4$, although to reduce radiation of higher-mode modes, they should be closer. For higher-mode operation, we increase the outer diameter to $(m + 0.25)\lambda/\pi$, and the spacing between the feed points can be increased to $m\lambda/4$.

A spiral radiates RHC polarization on one side and LHC polarization on the other side. We mount the antenna over a cavity to eliminate the unwanted polarization. The balun producing the balanced feed for a two-arm spiral can be mounted in the cavity. It converts the antenna from a coax input to prevent pattern squint and to limit higher-order mode radiation. Hand rules determine the sense of circular polarization. Let your fingers roll in the direction of the spiral (tips toward increasing radius) and the thumb points to the pattern maximum.

11-3 EQUIANGULAR SPIRAL

Rumsey [2] states that an antenna shape determined entirely by angles will be frequency independent because it is invariant to a change of scale. The biconical antenna satisfies the angle requirement but fails the truncation requirement for frequency independence because the current remains constant along its length and it fails the truncation requirement. An equiangular spiral antenna (Figure 11-4) defined by

$$r = r_0 e^{a\phi} \tag{11-6}$$

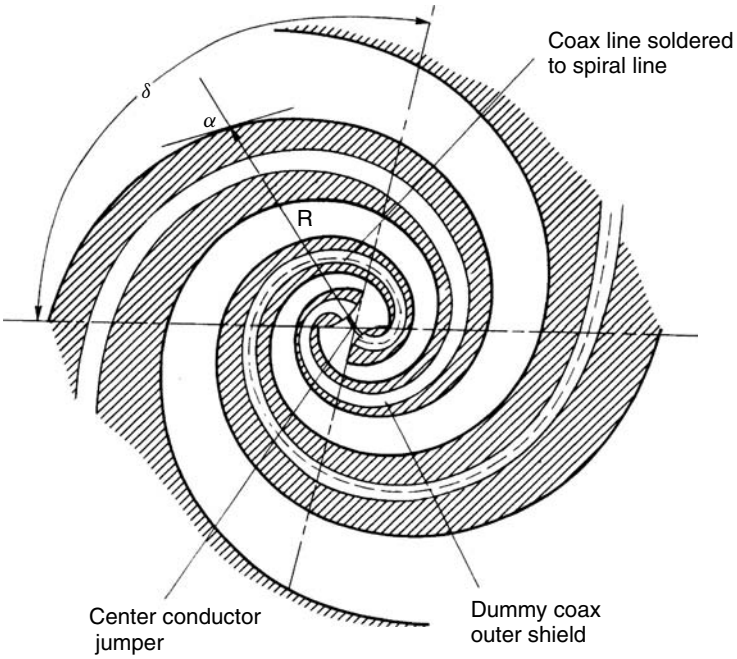


FIGURE 11-4 Equiangular spiral (RHC) with an infinite balun feeder.

is defined only by angles, since the inner radius can be related to an angle $r_0 = e^{a\phi_0}$ and satisfies the requirement for an antenna completely determined by angles. The wrap angle α (Figure 11-4) relates to the growth rate a of the spiral by

$$a = \frac{1}{\tan \alpha} \tag{11-7}$$

Another way of specifying the curves, the expansion factor (EF) specifies the ratio of radius increase in one turn; it is similar to the growth rate. The direct relationship to geometry makes it easy to specify:

$$a = \frac{\ln (\text{EF})}{2\pi} \quad \text{or} \quad \text{EF} = e^{2\pi a} \tag{11-8}$$

$$\text{EF} = \left(\frac{r_o}{r_i}\right)^{1/\text{turns}} \quad \text{or} \quad \text{Turns} = \frac{\ln (r_o/r_i)}{\ln (\text{EF})}$$

Angles also determine the width of the arms. We rotate the spiral by δ to generate the other edge of the spiral arm. We shift the angle ϕ by 180° (π) to generate the second arm. We also specify the spiral arms by the arm/gap ratio, so we can use noncomplementary arms as an impedance transformer. For δ specified in radians (substitute 360° for 2π if degrees) and N arms in the spiral, we can derive simple formulas between the two terms:

$$\delta = \frac{2\pi}{N\,(1 + \text{gap/arm})} \quad \text{or} \quad \frac{\text{gap}}{\text{arm}} = \frac{2\pi}{N\delta} - 1 \tag{11-9}$$

The input impedance of the antenna depends on the number of arms and the mode of operation. For complementary arms, or arm/gap = 1, an extension of the Babinet–Booker principle determines the impedance for a free-space spiral [6]:

$$Z_m = \frac{\eta_0/4}{\sin(\pi|m|/N)} \tag{11-10}$$

Equation (11-10) uses $\eta_0 = 376.73\ \Omega$ and positive m for the mode number where $m = 1, 2, \ldots, N-1$, whether RHC or LHC. Table 11-1 lists the characteristic impedance of N-arm multiterminal complementary structures in free space.

Table 11-1 does not give the impedance of a real spiral antenna because it will be mounted over a metallic cavity that will be either empty or partially filled with a multilayer sheet absorber. Second, the dielectric substrate on which the spiral is etched slows the waves on the arms and reduces the impedance by $1/\sqrt{\varepsilon_{r,\text{eff}}}$. Although these factors change the impedance, the table does illustrate the relative magnitudes of impedances for the various modes. We need to measure the spiral input impedance in its final configuration for its modes, which can be done using single-arm measurements of S_{11} and coupling to the other arms S_{21}. See Section 11-6 for this method.

We can impedance-match the antenna either for a single mode or for a compromise between modes by etching a tapered transformer at the spiral inputs. To vary the impedance we change the arm/gap ratio. The width of the spiral arms has only a minor effect on the pattern. An approach to this design uses the impedance of a coplanar strip transmission line for the complementary and noncomplementary structure to scale the impedances of the spiral [7]:

$$Z_{m,\text{noncomp}} = Z_{m,\text{comp}} \frac{Z_{\text{cp,noncomp}}}{Z_{\text{cp,comp}}} \tag{11-11}$$

Z_{cp} is the impedance of a coplanar stripline found from a model that includes multilayer substrates of lossy dielectrics, such as the absorber sheets [8, p. 70; 9]. The subscripts "noncomp" and "comp" refer to the noncomplementary and complementary line widths in the coplanar stripline and spiral. We generate a table by varying stripwidth/gap and calculate the impedance in the coplanar stripline and relate it to arm/gap of the spiral to determine equivalent impedance. This method starts with the impedance measurement

TABLE 11-1 Characteristic Impedance of N-Arm Multiterminal Complementary Structures in Free Space

Number of Arms	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6	Mode 7
2	94.2						
3	108.8	108.8					
4	133.2	94.2	133.2				
5	160.2	99	99	160.2			
6	188.4	108.8	94.2	108.8	188.4		
7	217.1	120.5	96.6	96.6	120.5	217.1	
8	246.1	133.2	101.9	94.2	101.9	133.2	246.1

Source: [6].

for various modes of a complementary spiral in its final configuration. After generating a table relating arm/gap to spiral impedance, we design a tapered transformer using standard techniques and interpolating on the table for dimensions. When you vary the arm/gap ratio along the spiral arm, divide δ [Eq. (11-9)] into two parts ($\pm\delta/2$) and vary both sides of the strip centered on the central curve [Eq. (11-6)] so that the etched spiral pattern retains symmetry.

We integrate Eq. (11-6) to calculate arm length:

$$L = (r - r_0)\sqrt{1 + \tan^2 \alpha} = (r - r_0)\sqrt{1 + \frac{1}{a^2}} \tag{11-12}$$

We estimate the loss in the spiral by modeling the arm as a coplanar transmission line and calculating the transmission-line loss by including the finite conductive and dielectric loss tangent of the substrate and the nearest absorber layer in the cavity [8, p. 70]. Increasing the gap between the substrate and the first absorber layer reduces transmission-line loss. A few calculations using the multilayer model [9] will determine a suitable gap for the design. Feeding the spiral from the outside to radiate the opposite sense of circular polarization, a dual polarized design, can produce a design with long transmission-line lengths with high losses for the high frequencies whose active regions occur in the center of the spiral. Increasing the expansion factor will reduce the transmission-line losses of the spiral.

11-4 PATTERN ANALYSIS OF SPIRAL ANTENNAS

Measurement of spiral patterns shows that they do not have a null in the plane of the spiral, as predicted by a current sheet model, and that the antenna radiates cross-polarization in the upper hemisphere. A simple analysis model uses a traveling-wave loop in free space. This current distribution has constant amplitude on the loop, but its phase progresses linearly around the loop in integer cycles. Regardless of the actual circumference of the loop, a mode 1 current progresses through -360° (RHC) while mode 2 current phase changes -720° and, in general, mode m current rolls through $-m360^\circ$ for CCW movement along the loop. Mode 1 radiates from a loop 1λ in circumference and the general mode radiates from a loop $m\lambda$ in circumference when the current propagates with the free-space velocity. The loop size determines the pattern beamwidth and the effective loop size decreases when the spiral is dielectrically loaded by either the substrate or by dielectric slabs (lenses) and its beamwidth increases.

The traveling-wave loop model radiates only one mode because both the RHC and LHC signals have integer cycle phase progression in conic patterns. Modal expansion on the model patterns fails to indicate the level of cross-polarization of actual spirals when we use the ratio of integrals of the power in the co-polarized signal only to the sum of the power of both polarizations to determine cross-polarization loss. We need a better model for that characteristic and a way to predict higher mode radiation.

A method of moments (MOM) model of the spiral using a wire code such as NEC can predict the multiple-mode radiation of a spiral. The wire diameter in the model has little effect on the pattern predictions. Since it is difficult to model the absorber-loaded cavity in a MOM code, accurate impedance values cannot be obtained and fabricating antennas is the cost-effective method for finding impedance. You should

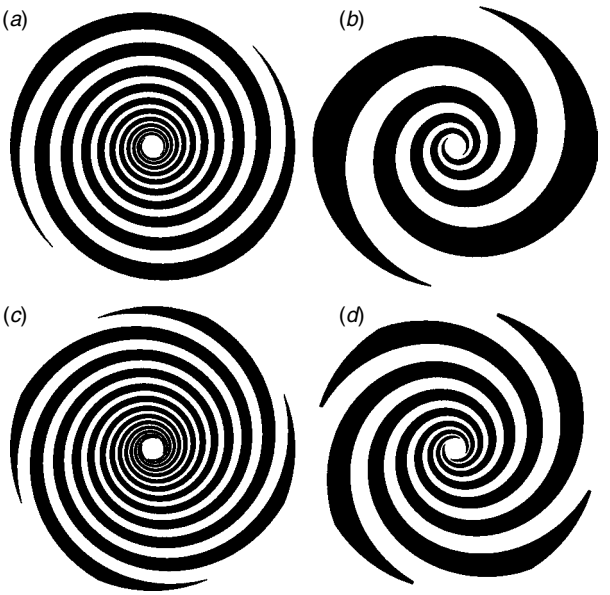


FIGURE 11-5 Equiangular spirals with differing expansion factors: (*a*) two-arm EF = 1.66; (*b*) two-arm EF = 3.32; (*c*) four-arm EF = 2.32; (*d*) four-arm EF = 4.64.

use the rotational symmetry capability of the code to reduce the model to a single-arm input so that the matrix is reduced by the number of arms N. We place the model in free space, which models the cavity as perfectly absorbing, and consider only the upper hemisphere.

Figure 11-5 illustrates the faces of two sets of two- and four-arm spirals designed to operate over a 10 : 1 frequency range. The inner diameter is 0.254λ. The left two-arm spiral contains five turns, with an expansion factor = 1.66 ($\alpha = 85.4^\circ$), while the looser-wrapped spiral on the right has 2.1 turns and EF = 3.32 ($\alpha = 79.15^\circ$). As we decrease the wrap angle (increased EF) the outer circumference of the spiral must be increased to support the lowest frequency. To illustrate the effect of the outer circumference, Table 11-2 lists the modal response of the five-turn (EF = 1.66) spiral versus outer circumference, and Table 11-3 lists the same results for the 2.1-turn (EF = 3.32) spiral. Scale 1-7 shows that the difference between mode 1 and mode -1 must be 9.6 dB to achieve a 6-dB axial ratio. The five-turn spiral needs an outer circumference of about 1.4λ at the lowest frequency for this value, while the 2.1-turn spiral should be 1.9λ. The looser-wrapped spiral requires a larger diameter to achieve the same

TABLE 11-2 Modal Response of a Two-Arm Exponential Spiral with *EF* = 1.66 ($\alpha = 85.4^\circ$), Five Turns

Circumference (λ)	Mode 1 (dB)	Mode -1 (dB)	Circumference (λ)	Mode 1 (dB)	Mode -1 (dB)
1.00	−1.96	−4.40	1.60	−0.16	−14.52
1.20	−1.09	−6.55	1.80	−0.05	−19.82
1.40	−0.46	−10.00	2.00	−0.04	−20.60

TABLE 11-3 Modal Response of a Two-Arm Exponential Spiral with $EF = 3.32$ ($\alpha = 79.16°$), 2.1 Turns

Circumference (λ)	Mode 1 (dB)	Mode −1 (dB)	Circumference (λ)	Mode 1 (dB)	Mode −1 (dB)
1.00	−2.34	−3.82	1.80	−0.53	−9.44
1.20	−1.81	−4.69	2.00	−0.34	−11.26
1.40	−1.25	−6.02	2.20	−0.21	−13.38
1.60	−0.82	−7.68	2.40	−0.12	−16.07

low-frequency axial ratio. These spirals have open-circuited arms. We reduce the axial ratio by loading the ends of the arms, which reduces the reflected mode radiation by the return loss of the loads relative to the spiral arm impedance. The efficiency decreases but the axial ratio improves.

An Archimedean spiral with 10.5 turns covers the same 10 : 1 frequency range. Table 11-4 gives the modal response. The tighter wrap of the spiral near the outer diameter reduces the diameter required to produce a lower axial ratio design. An outer circumference of 1.17λ is sufficient for a 6-dB axial ratio. The additional spiral arm length will reduce efficiency at the low-frequency end due to the transmission-line loss.

Increasing the number of turns will decrease the levels of the over-modes radiated. The two-arm spiral suppresses the even modes but allows radiation of modes 3, 5, 7, ... [Eq. (11-2)]. The moment method analysis predicts the over-mode levels given in Figure 11-6. The higher-order modes occur because the antenna is large enough to support these modes at higher frequencies but not at lower frequencies. Power not radiated in the mode 1 region near a 1λ circumference travels along the arms until it reaches the 3λ and 5λ circumferences and radiates. Over-modes alter conical patterns by adding amplitude and phase ripple. The phase ripple is added to the mode single 360° linear distribution. Mode 3 adds a single cycle ripple to conical patterns of mode 1, and mode 5 will add a two-cycle ripple on top of the single mode 3 cycle. At normalized frequency 9.5 for EF = 1.66, mode 3 at −10 dB relative to mode 1 produces an average peak-to-peak ripple of 5.7 dB (Scale 1-8) while the −23-dB mode 5 produces 1.2 dB of ripple. Amplitude and phase ripple is not a significant problem in most applications because we cannot use the antenna for an AOA system without two modes. Table 11-5 lists the mode 3 amplitude on the Archimedean spiral, which says that we must reduce the inner diameter to limit the over-mode level. For example, to limit it to −16.4 dB (2.6 dB of ripple), we need to reduce the inner diameter to 80% of 0.254λ (0.20λ). The Archimedean spiral requires a smaller feed diameter than the exponential spiral for the same over-mode level.

TABLE 11-4 Modal Response of a Two-Arm Archimedean Spiral with 10.5 Turns

Circumference (λ)	Mode 1 (dB)	Mode −1 (dB)	Circumference (λ)	Mode 1 (dB)	Mode −1 (dB)
1.00	−1.81	−4.68	1.20	−0.31	−11.68
1.05	−1.36	−5.70	1.25	−0.14	−14.92
1.10	−0.94	−7.13	1.30	−0.05	−19.34
1.15	−0.57	−9.11	1.35	−0.02	−24.04

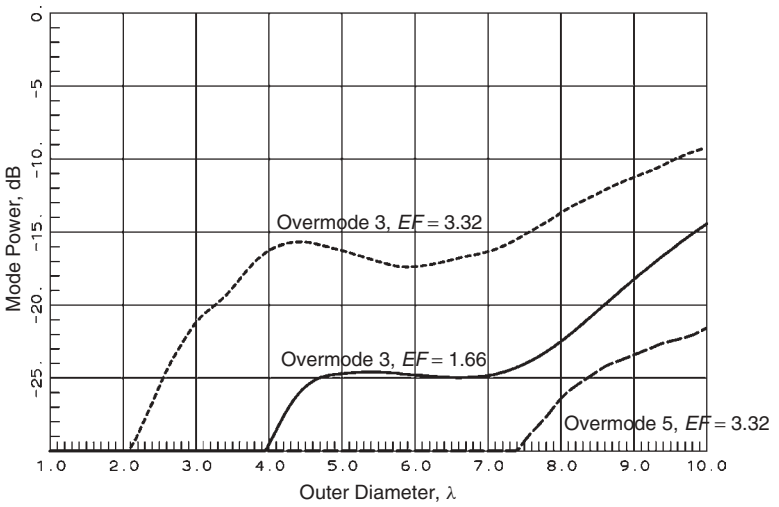


FIGURE 11-6 Over-modes of a two-arm spiral designed for a 10 : 1 frequency range.

TABLE 11-5 Mode 3 Relative Power for a 10 : 1 Archimedean Spiral with an 0.254λ Inner Diameter

Frequency	Power (dB)	Frequency	Power (dB)
7.2	−19.3	8.8	−14.0
7.6	−17.8	9.2	−13.0
8.0	−16.4	9.6	−12.0
8.4	−15.2	10.0	−11.0

The two examples of four-arm spirals exhibit results similar to those of two-arm spirals. Decreasing the expansion factor decreases high-order modes, and the phasing of the feed network eliminates additional modes. The four-arm spiral radiates mode 5 with a power of −13 dB relative to the total and the extra signal adds and subtracts with the mode 1 radiation to produce the four-way symmetry of the curves to the pattern deviation. The pattern ripple number equals the difference in mode number when two modes interact. The −13 dB extra radiation causes a 4-dB amplitude ripple (Scale 1-8) and 13° peak-to-peak phase variation (Scale 1-9). The similar levels of mode 6 shown in Figure 11-7 for mode 2 excitation produce pattern ripple. The AOA system uses the relative phase and amplitude between the two modes, and these extra signals cause the pattern variation to rotate in angular position as frequency changes. If we do not control the levels of the extra modes, AOA accuracy will suffer. In other applications the pattern fluctuation would be acceptable. Table 11-6 lists the modal response of a four-arm exponential spiral when fed for mode 2 and gives the size required to suppress mode −2 radiating cross-polarization.

To some extent we can feed the outer arms to radiate the opposite-sense circular polarization. The signals reach the circumference of higher-mode radiation before reaching the circumference of lower-order radiation. The spiral radiates substantial power in the higher-order modes which reduces the signal traveling inward to the

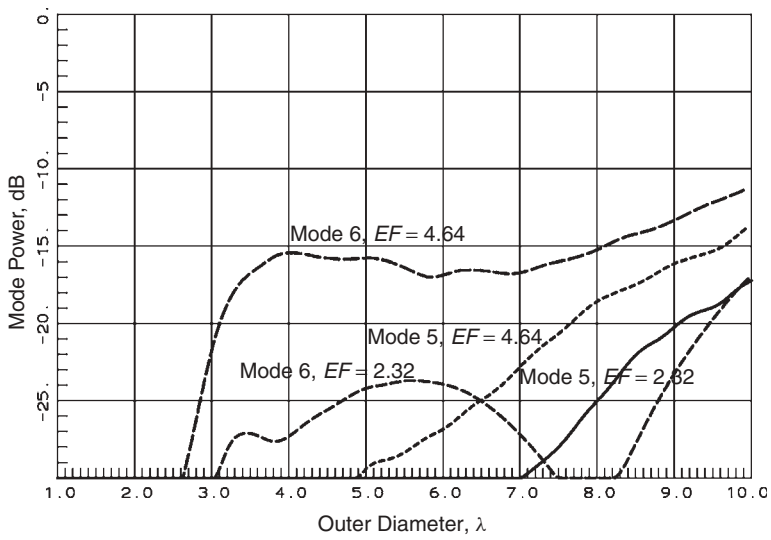


FIGURE 11-7 Over-modes of a four-arm spiral designed for a 10 : 1 frequency range.

TABLE 11-6 Modal Response of a Four-Arm Exponential Spiral with $EF = 2.07$ ($\alpha = 83.4^\circ$), 3.5 Turns Fed for Mode 2

Circumference (λ)	Mode 2 (dB)	Mode −2 (dB)	Circumference (λ)	Mode 2 (dB)	Mode −2 (dB)
2.00	−1.91	−4.49	2.80	−0.25	−12.6
2.20	−1.33	−5.78	3.00	−0.13	−15.4
2.40	−0.82	−7.64	3.20	−0.07	−18.1
2.60	−0.46	−10.0	3.40	−0.04	−20.3

lower-order-mode radiation circumference. We increase the number of arms to suppress the higher-mode radiation through phase cancellation produced by the feed network. Figure 11-8 shows the radiation levels of mode −1 for different number of arms. The outer circumference is 1λ at the lowest frequency for mode −1. We see an initial efficiency loss for the two-arm spiral in mode −1. Higher order modes have more restricted frequency regions.

Locating the spiral wire above a ground plane models a reflective cavity antenna to some extent, but it does not include the effects of the cylinder walls. The reflection of the waves from the ground plane will excite higher-mode currents farther out in the spiral, which increases the modal content of the pattern and limits the usable frequency band of the antenna. For example, a two-arm spiral fed for mode 1 with an initial depth of λ/8 has a mode 3 level at −8 dB when the cavity depth is λ/2. When the cavity depth is λ/2, the mode 1 pattern has a null at broadside. To produce an antenna with a wider bandwidth, we need a shallower cavity. The closer ground plane increases the power reflected into higher modes and the over-modes problem increases. We reduce the over-modes by increasing the number of arms to cancel modes through the feeding phases.

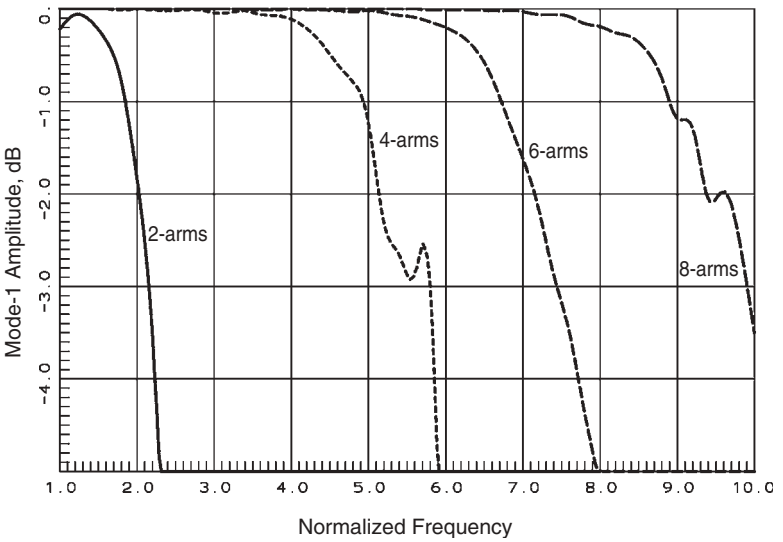


FIGURE 11-8 Efficiency of outside arm feeding of spirals designed for a 10 : 1 frequency range for mode −1 for different numbers of arms.

11-5 SPIRAL CONSTRUCTION AND FEEDING

11-5.1 Spiral Construction

We etch the spiral pattern on a thin dielectric sheet. We should specify a circuit board with a low-loss tangent because the spiral operates as a transmission line between the arms whose length becomes significant for tightly wrapped spirals. This transmission can be analyzed as a coplanar strip transmission line for losses, and the equivalent dielectric constant of the transmission line loads the spiral and reduces the effective loop radiator size. We can further load the antenna by placing dielectric sheets above the spiral or by placing a contact lens on the spiral face and shrink the spiral diameter.

Although it is not necessary to place a resistive load on the end of the spiral arms, it will improve the polarization at low frequencies by reducing reflections that radiate the opposite sense of circular polarization. Reflected currents travel inward and pass through the radiation regions of the spiral. Spirals fed from both inside and outside contain transmission-line feeds that load each arm; otherwise, we use a resistive paste or film on the last turn, but the antenna may produce acceptable patterns without loads on the arms.

We can use a reflective cavity under the spiral to prevent radiation of the opposite sense of circular polarization, but it limits the bandwidth. Waves reflected from the cavity base couple back into the spiral arms beyond the lower-mode radiation region. These waves excite spiral currents that travel to the next radiation region or into the loads on the ends. Without these loads the currents would reflect and travel inward to radiate in the first active region of an oppositely sensed circularly polarized mode.

We load the base of the cavity with absorber to build a wideband antenna. A tapered or stepped loaded absorber prevents reflection over a wideband. You should space the absorber away from the spiral face so that it does not significantly load the transmission line of the spiral arms. We analyze this absorber loading as an element in a coplanar

strip transmission line containing multiple lossy dielectric layers. A foam or dielectric honeycomb spacer between the absorber and the spiral circuit board prevents thermal or mechanical stress movement that severs the board connections.

Simple analysis says that half the power will be radiated into the absorber and reduce the gain by 3 dB. Unfortunately, the wave impedance of the absorber loaded cavity is much less than the impedance of free space. In an admittance model of the radiation in free space and the cavity, the power divides between the shunt loads of the two regions and the lower cavity impedance causes more than half of the power to be dissipated in the cavity, which further reduces the gain. We can recover some of this lost power by loading the spiral with a hemispherical contact lens that lowers the effective radiation impedance in the two-shunt-load model. The hemispherical lens has little effect on pattern shape since the outer surface does not refract rays that pass through normal to it. A lens with a dielectric constant in the range 2 to 3 increases gain by 1.5 to 2 dB, but it reduces the efficiency of mode radiation in the active region and leads to increased cross-polarization and over-mode levels.

A contact lens can decrease the beamwidth of a spiral besides decreasing the cavity absorber loss. Applying the technique of Section 9-4 and solving for a lens surface, we obtain the lens in Figure 11-9*a*, which shows the traced ray refraction for $\varepsilon_r = 2.55$. The lens surface can only refract rays straight upward out to an angle of 51°, but the side cylinder continues to refract rays from the feed upward. Figure 11-9*b* and *c* shows the measured pattern for a four-arm spiral operating in modes 1 and 2 with a 1.25λ-diameter contact lens on it. The lens reduced the beamwidth of mode 1 from 70° to 32° and mode 2 from 42° to 25° when we compare these patterns with Figures 11-1 and 11-2. The mode 2 beam peak moved from 35° to 22°. The lens design uses a point source, but the spiral radiates from rings approximately 0.33λ (mode 1) and 0.66λ (mode 2) in diameter. Nevertheless, the lens produces remarkable results with a mode 2 ring current one-half its diameter. The lens continued to shrink the beamwidths in a linear function as frequency increased. Similar results were reported using a hemispherical lens raised above the spiral [3]. This example demonstrates that small lenses have a significant effect on antenna patterns, whereas a reflector of this size would be useless.

11-5.2 Balun Feed

We feed a two-arm spiral with a balanced line, whereas spirals with more arms require a beamformer network. A balanced line feeder contains equal and oppositely phased currents when fed from a balun. The two common baluns used for spirals are the Marchand balun, a compensated sleeve balun (Section 5-15.2), and Bawer and Wolfe [11] version of the Roberts balun (Section 5-15.1), a compensated folded balun. We place the Marchand balun below the cavity to position its balance wire output in the center and we feed a two-wire line across the cavity to feed the spiral. We locate the two sleeves of the Marchand balun parallel to the cavity base. The Bawer and Wolfe balun is constructed on two sides of a printed circuit board oriented along the vertical axis. The balanced output is a short distance from the feed point and we feed two wires through to the etched board of the spiral. The printed circuit board causes a small asymmetry in the cavity and the absorber must be removed around the board so that it does not load it. Because mode 1 impedance is greater than 50 Ω, we can use a tapered microstrip balun similar to the split-tube coaxial balun to feed the two-arm spiral.

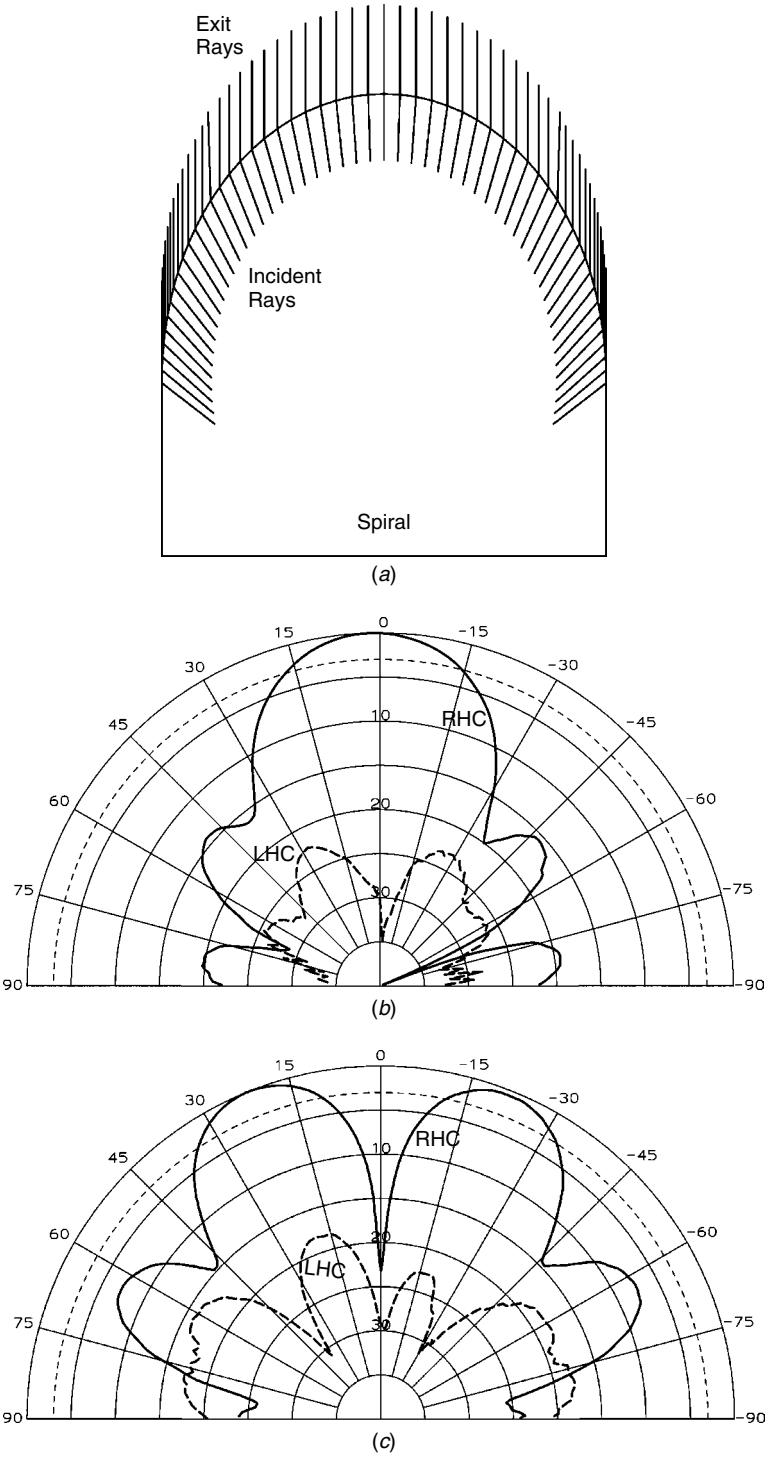


FIGURE 11-9 Measured patterns of a 1.25λ-diameter contact lens mounted over a spiral: (*a*) ray trace in polyethylene lens; (*b*) mode 1 pattern; (*c*) mode 2 pattern.

11-5.3 Infinite Balun

We can make a balun by using the truncation property of the spiral where the current attenuates rapidly beyond the active region. The balun prevents currents excited on the outside of a coax from reaching the input. In Figure 11-4 we solder the coax feeder to one of the arms. Since the antenna does not radiate in the direction of the expanding arms, it does not excite currents on the structure after the active region; conversely, no currents excited on the structure beyond the active region will reach the input. The coax feeder outer shield becomes part of the antenna. We solder a dummy coax on the second arm to maintain symmetry.

Because the balun uses the active region limitation on currents of the antenna, its bandwidth matches that of the antenna. We use the same balun structure for log-periodic dipole antennas when similar truncation requirements can be used to form the balun. This construction requires wide spiral traces for soldering, a difficult operation to perform over the long arm length. To shorten the arm length, we use spirals with low wrap angles, but they have poor radiation characteristics.

11-5.4 Beamformer and Coaxial Line Feed

Spirals with more than two arms require the direct feeding of each arm with a coaxial line. We feed phase-matched cables through the cavity and connect them to a microwave circuit that generates modal excitations. The beamformer provides a separate input port for each mode desired. The coax center conductors feed the spiral arms while we join the outer conductors. The spiral mode currents sum to zero, which means that the currents on the outer conductors also sum to zero. When feeding the inside, we strap the outer conductors together by using a ribbon connected to all shields or route the cables through a metal cylinder. The cylinder stops short of the spiral circuit board but is connected to the cavity bottom. For perfect modal feeding the current in the cylinder is zero. When we feed the outer arms of the spiral, we etch a shorting disk on the opposite side of the spiral face. We connect each coaxial cable outer conductor to this ring. By connecting the shorting disk to the outer wall of the cavity, we produce another path for the currents to sum to zero.

The design and construction of beamformers is beyond our discussion. We use a Butler matrix as a beamformer for spirals with 2^N arms (2, 4, 8, etc.), which provides a separate input port for the complete set of modes, while beamformers for spirals with 3, 5, 6, etc. arms present a challenge. Complete analysis of the spiral involves measurements of the spiral arms and the beamformer.

11-6 SPIRAL AND BEAMFORMER MEASUREMENTS [12]

We measure antenna patterns of the assembly of the spiral and the beamformer and determine the modal content by applying Eq. (11-4). Of course, we extract the normal pattern parameters of gain, beamwidth, beam direction, cross-polarization, and possibly phase center from the measurements. The final antenna can radiate in many modes, whereas the beamformer N_{arm} outputs can operate only in N_{arm} modes. During construction we gain insight and can correct problems by making single-arm measurements of the spiral and single port measurements of the beamformer.

The beamformer outputs can be expanded in N_{arm} modes, where each mode has equal amplitude on the outputs and a different phase progression in cycles of 2π radians. The feeding coefficients on arm N, for mode m, are given by

$$V_N = \frac{\exp[-j2\pi m(N-1)/N_{\text{arm}}]}{\sqrt{N_{\text{arm}}}} \tag{11-13}$$

where we number the arms CCW when looking at the spiral face. We calculate the modal expansion levels by multiplying network analyzer measurements of the arm responses b_n by the complex conjugate of Eq. (11-13) and adding over the arms. This is a summation version of Eq. (11-4) used for antenna patterns:

$$b_m = \sum_n V^*_{n,m} b_n \tag{11-14}$$

We repeat these measurements for each spiral mode of the beamformer and expand them in modes given by Eq. (11-14) to detect beamformer construction problems.

We measure the antenna pattern for each spiral arm separately while placing resistive loads on the other ports. This is the scan (active) pattern of each arm in a co-located array. The close arm spacing generates high mutual coupling that excites currents on all arms even though only one arm is fed. Measurement using an automated pattern system allows digital storage of all patterns that can be added using the measured outputs of the beamformer or an ideal beamformer to determine the final pattern. We apply Eq. (11-4) to the combined single-arm measurements to calculate the mode levels radiated for each input mode. Except for a ϕ rotation, all single-arm measurements should be the same. We gain insight into the construction of each arm by duplicating a single-arm measurement to produce N_{arm} copies and rotating $N_{\text{arm}} - 1$ of them, applying an ideal or measured beamformer responses, and calculating the resulting pattern to determine modal content. Construction differences between the arms become readily apparent. Of course, we may reduce the initial measurement effort during development by measuring a single arm and assuming ideal construction between arms. The other arms must be present so that the single arm couples to them and excites currents on them.

***S*-Parameter and Impedance Measurements** We use a network analyzer to measure the coupling to the other arms as well as the input reflection of each arm. We compute the input impedance for each mode by combining the reflection coefficient of a single arm with the others loaded (S-parameter) and the mutual coupling to the other arms weighted by the mode voltages [Eq. (11-13)]. Our analysis of the multiarm spiral as an array allows the use of the scan impedance. If we assume construction symmetry initially, we measure with only one arm as input. We load the outputs of the spiral arms if this is the final configuration; otherwise, we leave them with open-circuited terminations. The resulting reflection coefficient is found from

$$\Gamma_1 = \frac{b_1}{a_1} = S_{11} + S_{12}\frac{a_2}{a_1} + S_{13}\frac{a_3}{a_1} + \cdots + S_{1N}\frac{a_N}{a_1} \tag{11-15}$$

The coefficients a_i are the arm modal coefficients whose sum of magnitude squared equals 1, and S_{ij} are the mutual coupling values.

For spirals with arms loaded on the ends, we replace the loads with connectors and measure the coupling from the inputs to the loads. The power dissipated in each load depends on the mode. We denote the arm output ports as $N_{\text{arm}} + 1$ to $2N_{\text{arm}}$ in CCW order and measure the output wave on each arm:

$$\begin{aligned} b_{N+1} &= S_{N+1,1}a_1 + S_{N+1,2}a_2 + S_{N+1,3}a_3 + \cdots + S_{N+1,N}a_N \\ b_{N+2} &= S_{N+2,1}a_1 + S_{N+2,2}a_2 + S_{N+2,3}a_3 + \cdots + S_{N+2,N}a_N \\ &\vdots \\ b_{2N} &= S_{2N,1}a_1 + S_{2N,2}a_2 + S_{2N,3}a_3 + \cdots + S_{2N,N}a_N \end{aligned} \tag{11-16}$$

We sum the magnitude squared of b_{N+1} to b_{2N} to calculate the power dissipated in the loads [Eq. (11-16)] after we apply the modal coefficients a_i for a given mode. Equation (11-15) calculates the reflected power $1 - |\Gamma|^2$ and Eq. (11-16) determine the power dissipated in loads. We separate the antenna input power into the terms: (1) reflected, (2) dissipated in loads, (3) radiation, (4) circuit losses, and (5) power absorbed in the cavity. The power absorbed in the cavity and the circuit losses are only mildly frequency dependent. The sum of the first two terms will indicate the frequency range where the antenna has correct dimensions for efficient radiation and we can determine truncation constants with these bench measurements.

11-7 FEED NETWORK AND ANTENNA INTERACTION [8, pp. 347–353; 13]

The analysis above assumes an ideal feed network with perfect isolation between the output ports and impedance-matched ports on the antenna, the spiral arms, but the real feed network has limited isolation and the antenna ports are often mismatched. We encounter the same problem when dealing with a phased array because the input impedance of each element changes when we scan the beam. We use the subdomain growth method to solve these problems. The antenna has N ports matched by the same number of outputs on the feed network. We measure the $N \times N$ mutual coupling matrix of the antenna S_{RR} and the complete S-parameter matrix between the outputs of the feed network S_{QQ}, also $N \times N$. The normal connection matrix between the input and the feed network outputs S_{PQ} is $1 \times N$; the analysis in Section 11-6 used this matrix. We measure the input reflection coefficient S_{pp}, a single-element matrix, and then combine the matrices into an overall $(2N + 1) \times (2N + 1)$ matrix:

$$\begin{bmatrix} b_P \\ b_Q \\ b_R \end{bmatrix} = \begin{bmatrix} S_{PP} & S_{PQ} & [0] \\ S_{PQ}^{\mathrm{T}} & S_{QQ} & [0] \\ [0] & [0] & S_{RR} \end{bmatrix} \begin{bmatrix} a_P \\ a_Q \\ a_R \end{bmatrix} \tag{11-17}$$

The term [0] is an $N \times N$ null matrix.

We form a $2N \times 2N$ connection matrix between the antenna ports R and the feed network output ports Q using an $N \times N$ identity matrix $[I]$:

$$\begin{bmatrix} -S_{QQ} & [I] \\ [I] & S_{RR} \end{bmatrix}$$

We invert this matrix by partitioning it and obtain the matrix

$$\begin{bmatrix} S_{RR}(I - S_{QQ}S_{RR})^{-1} & (I - S_{RR}S_{QQ})^{-1} \\ (I - S_{QQ}S_{RR})^{-1} & S_{QQ}(I - S_{RR}S_{QQ})^{-1} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{11-18}$$

Since we have no direct inputs to the antenna elements, the analysis reduces to

$$[a_Q] = M_{11}S_{QP}a_P \quad \text{and} \quad [a_R] = M_{21}S_{QP}a_P \quad \text{with} \quad [a_Q] = [b_R] \tag{11-19}$$

The element a_P is the input to the feed network, and $[a_R]$ and $[a_Q]$ are the inputs to the antenna ports and the feed network output ports. We compute the input voltage vector to the antenna elements as $[V] = [a_R] + [b_R] = [a_R] + [a_Q]$. The input power to the elements is $P = [a_R]^{\mathrm{T}}[a_R]^* - [a_Q]^{\mathrm{T}}[a_Q]^*$ and the input reflection is $b_P = S_{PP} + S_{PQ}[a_Q]$. With this method we determine the effects of a non-perfect-feed network on the spiral or phased-array patterns when we use these excitations to calculate the pattern.

11-8 MODULATED ARM WIDTH SPIRAL [3, 14]

We have limited ability to feed a spiral from the outside for radiation of the opposite sense of circular polarization, LHC, from a right-hand wound spiral because we need many arms to suppress undesirable radiation (Figure 11-8) over a significant bandwidth. Modes -2 and -3 have a more restrictive bandwidth because the current reaches a higher-mode radiation point before the inner circumference of a particular mode and radiates. The modulated arm width (MAW) spiral solves these problems. We feed this antenna from the center to obtain the negative modes assuming that the spiral is wound right-hand. If we feed a right-hand spiral with left-hand modes, the current flows through the spiral with little radiation and in a normal spiral current reflects from the open-circuited arm ends and travels inward. Traveling backward on the spiral, the currents radiate LHC polarization at integer multiples of a wavelength circumference not suppressed by the phasing of multiple arms. Unfortunately, the complete trip of the current through the spiral to the ends of the arms and back to the radiation circumference adds to the transmission-line loss of the antenna.

Figure 11-10 shows the construction of this antenna for four, six, and eight-arms. We modulate the arm width to form a bandstop choke (filter) that reflects currents whose center wavelength increases at the same rate as the spiral diameter. The number of modulation cycles around one spiral turn equals the number of arms. For analysis we consider the spiral as a transmission line in the nonradiating regions whose arm/gap ratio determines the characteristic impedance. When the length of the modulation section approaches $\lambda/4$, the impedance mismatch at one discontinuity adds with the next one, and the numerous reflections build to form a bandstop choke response that increases with the line-width ratio. When the distances between the steps are substantially less than $\lambda/4$, the reflections fail to add coherently and cancel.

A four-arm MAW spiral has eight steps around one turn that locates a high reflection point at a 2λ circumference. If we feed the center of a four-arm spiral with mode 3, which equals mode -1 (LHC), the current travels outward on a right-hand-wound spiral with little radiation. The current reflects where the spiral circumference is 2λ

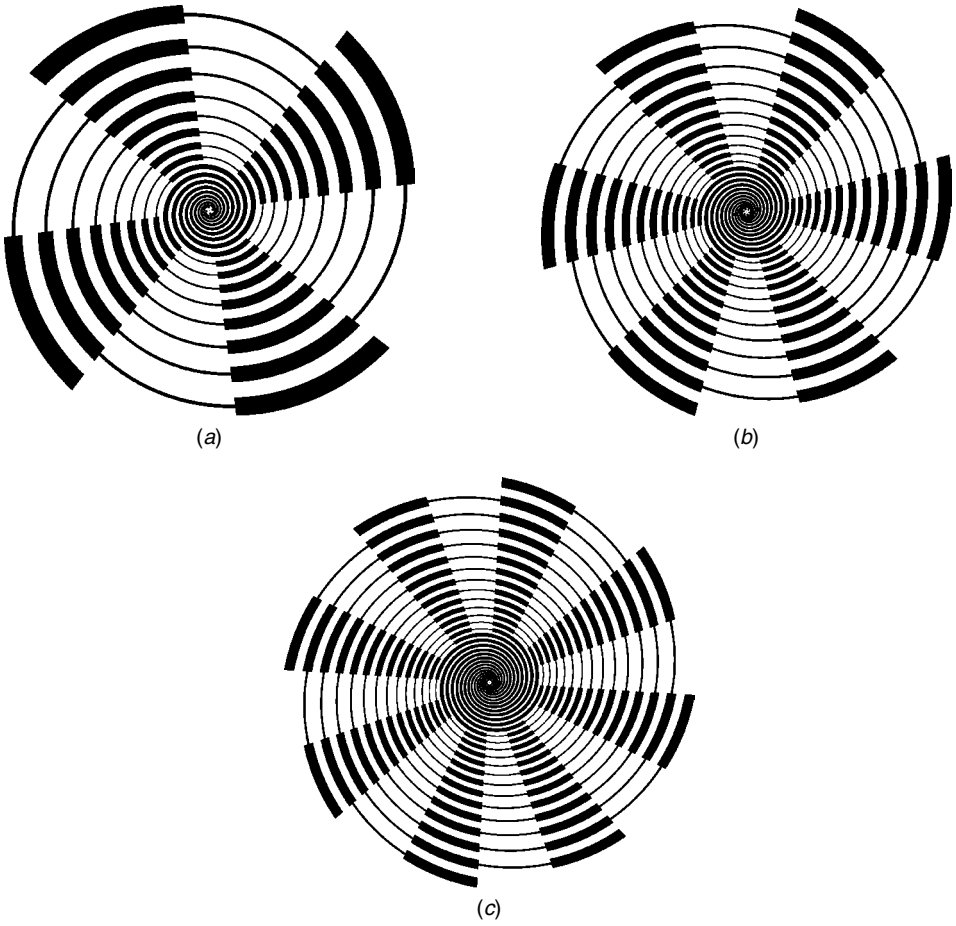


FIGURE 11-10 Modulated arm spirals: (*a*) four-arm; (*b*) six-arm; (*c*) eight-arm.

due to the bandstop chokes before it reaches the 3λ circumference and counterrotates to the 1λ circumference, where it radiates LHC polarization in mode −1. On the other hand, mode 1 radiates from the currents before they reach the bandstop region at a 2λ circumference, and the modulations have little effect. Of course, residual power left on the spiral after the first radiation region causes pattern ripple that varies over frequency, but the bandstop choke reduces radiation of mode 2 from the four-arm MAW spiral because the current reflects in the middle of the 2λ radiation region.

In a similar manner, a six-arm MAW spiral reflects the current at 3λ circumference and eight-arm MAW spiral at 4λ. The six-arm spiral with its reflection point at 3λ circumference can support modes −1 (5) and −2 (4). Both the six-arm feed phasing and the 3λ circumference bandstop choke suppresses mode −3 radiation when we feed it with mode −1 (5). The MAW spiral radiates mode 2 at a 2λ circumference before the current reaches the bandstop filter region at 3λ. Whereas a normal six-arm spiral can radiate mode 3, the MAW spiral cuts off its radiation because the bandstop region occurs at a 3λ circumference located at the middle of its active region. At this circumference both mode 3 and mode −3 exist and radiate to produce a linearly

polarized wave. In a similar manner the eight-arm spiral with its 4λ circumference bandstop filter can support modes 1, 2, 3, -1 (7), -2 (6), and (-3) 5. The eight-arm feeding suppresses mode -3 radiation [Eq. (11-2)] when fed mode 7 (-1), although the currents pass through the 3λ circumference point.

It would seem that we must construct a larger antenna to support the radiation of these reflected modes. The four-arm spiral effectively radiates only modes 1 and -1, and the bandstop region occurs at a circumference of 2λ. At the low-frequency end of the band, we can use the reflection from the open circuits on the ends of the arms instead of the bandstop filter. It is unnecessary to start modulating the arm width in the spiral center until approaching the bandstop filter region of the highest frequency. This eases construction problems and allows normal connection in the center to the many arms. We can also use this center section to construct a tapered transformer for impedance matching by varying the arm/gap ratio.

11-9 CONICAL LOG SPIRAL ANTENNA [15, 16]

When we form the equiangular spiral on a cone, the antenna radiates predominantly toward the vertex and we gain some control of the beamwidth. The antenna projected on a cone continues to satisfy the truncation condition: Radiation is reduced along the structure. We modify the beamwidth by varying both the cone angle θ_0 and the wrap angle δ (Figure 11-11). The antenna is a slow-wave structure from the feed on

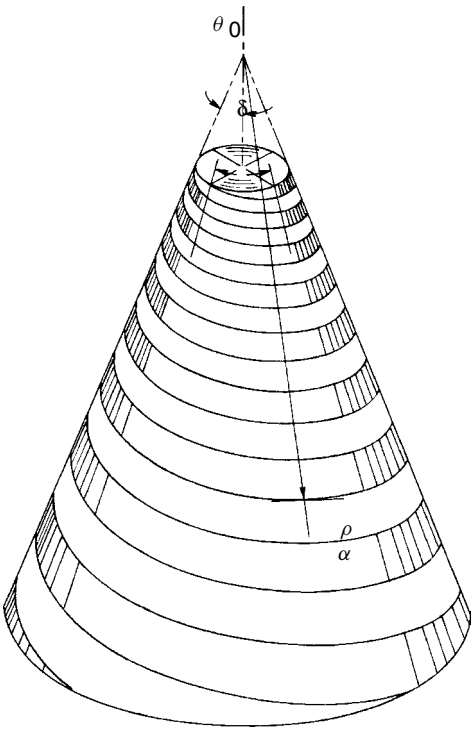


FIGURE 11-11 Two-arm conical log spiral antenna (RHC).

the upper diameter to the active region, but in the active region the antenna changes to a fast-wave structure to radiate a backfire pattern toward the cone vertex. Because the spiral on the cone radiates a unidirectional pattern, it reduces the radiation of one circular polarization sense. The flat spiral radiates equally on both sides, but because the pattern has a null in the direction of increasing structure, we can bend it downward on the cone and decrease the radiation of one sense of circular polarization. Figure 11-12 shows the calculated pattern of a conical spiral and illustrates the reduced back lobe and reduced cross-polarization caused by the conical shape.

We describe the spiral arms by the radius from the cone apex:

$$\rho = \rho_0 e^{b\phi} \qquad \text{where} \quad b = \frac{\sin\theta_0}{\tan\alpha} \tag{11-20}$$

We measure the angle of the spiral α with respect to the radius ρ along the cone. The angle δ determines the stripwidth, since Eq. (11-20) describes every edge of the spiral within an offset angle of ϕ. The length of the spiral strip edge is

$$L = (\rho - \rho_0)\sqrt{1 + \frac{1}{b^2}} = (\rho - \rho_0)\sqrt{1 + \frac{\tan^2\alpha}{\sin^2\theta_0}} \tag{11-21}$$

As θ_0 approaches $\pi/2$ (90°), we have the flat equiangular spiral.

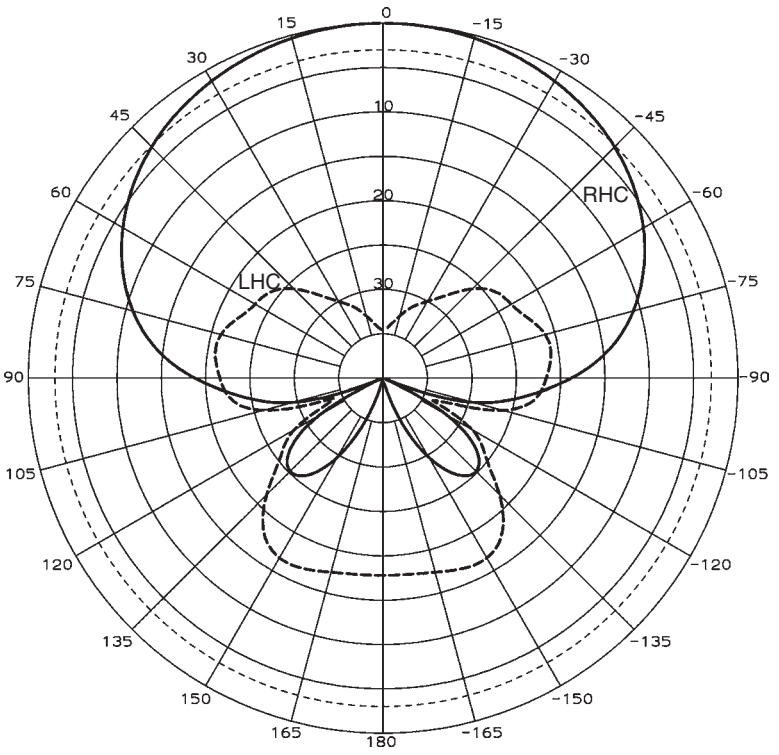


FIGURE 11-12 Calculated pattern of a two-arm conical spiral with a 30° total cone angle, $\alpha = 80°$, 11 turns.

We can describe the conical spiral using an expansion factor. Equation for the radius from the axis r is the same as the radius from the virtual apex ρ:

$$r = \rho_0 \sin\theta_0 e^{b\phi} = r_0 e^{b\phi}$$

In terms of the expansion factor, we have

$$b = \frac{\sin\theta_0}{\tan\alpha} = \frac{\ln(\text{EF})}{2\pi} \text{ or } \text{EF} = \exp\left(\frac{2\pi\sin\theta_0}{\tan\alpha}\right)$$

We rearrange these equations to compute the wrap angle α and the number of turns:

$$\alpha = \tan^{-1}\frac{2\pi\sin\theta_0}{\ln(\text{EF})} \quad \text{and} \quad \text{turns} = \frac{\tan\alpha\ln(r_o/r_i)}{2\pi\sin\theta_0}$$

Dyson [16] measured a large number of antennas to determine their properties, and we reduce his results to design tables. Table 11-7 gives the average beamwidth of the conical spiral. If we increase the wrap angle (tighter spiral) or decrease the cone angle (longer antenna), the beamwidth decreases and directivity increases. As we decrease the wrap angle to increase the beamwidth, the variation in beamwidth of various pattern cuts through the cone vertex increases because there are too few turns to maintain pattern symmetry, due to over-modes caused by an inefficient active region.

We specify the active region by upper and lower cone diameters. Dyson found a correlation between the band edges and the level of near-field probed currents. The current peaks in the active region, and in terms of design we can remove those portions of the antenna where the current drops by 3 dB on the high-frequency (small) end and 15 dB on the lower-frequency (large) end without affecting performance. These near-field current points give us upper and lower truncation diameters to scale the design. If we allow small changes in the beamwidth at the low-frequency end, we can use the 10-dB point to determine the lower diameter from the lowest operating frequency and make a smaller antenna. Tables 11-8 and 11-9 list the radiuses of the circles truncating

TABLE 11-7 Average Half-Power Beamwidth of a Two-Arm Conical Log Spiral Antenna ($\delta = 90°$)

Wrap Angle, α (deg)	Twice Cone Angle					
	$2\theta_0 = 2°$	$2\theta_0 = 5°$	$2\theta_0 = 10°$	$2\theta_0 = 15°$	$2\theta_0 = 20°$	$2\theta_0 = 30°$
90	36	49	55	60	65	70
85	37	50	58	64	68	74
80	38	53	63	70	74	81
75	41	56	70	78	83	90
70	44	60	79	88	95	103
65	47	65	89	100	108	119
60	52	71	102	114	127	139
55	57	79	115	132		
50	63	89				
45	69	106				

TABLE 11-8 Two-Arm Conical Log Spiral Antenna Upper Radius of Active Region (a_3^-/λ), Where the Current Drops 3 dB from the Peak

Wrap Angle, α (deg)	Twice Cone Angle						
	$2\theta_0 = 2^\circ$	$2\theta_0 = 5^\circ$	$2\theta_0 = 10^\circ$	$2\theta_0 = 15^\circ$	$2\theta_0 = 20^\circ$	$2\theta_0 = 30^\circ$	$2\theta_0 = 45^\circ$
85	0.119	0.111	0.106			0.091	
80	0.101	0.096	0.090	0.084	0.080	0.071	0.067
75	0.089	0.084	0.078	0.074	0.069	0.067	
70	0.078	0.074	0.069	0.066	0.060	0.057	
65	0.071	0.067	0.062	0.058	0.052	0.053	
60	0.063	0.059	0.054	0.050	0.045	0.046	
55	0.057	0.053	0.049	0.043	0.039		
50	0.052	0.048	0.043	0.035	0.036		
45	0.046	0.043		0.031	0.032		

TABLE 11-9 Two-Arm Conical Log Spiral Antenna Lower Radius (a_{10}^+/λ) of Active Region (Slightly Degraded Pattern), Where the Current Drops 10 dB from the Peak

Wrap Angle, α (deg)	Twice Cone Angle					
	$2\theta_0 = 2^\circ$	$2\theta_0 = 5^\circ$	$2\theta_0 = 10^\circ$	$2\theta_0 = 15^\circ$	$2\theta_0 = 20^\circ$	$2\theta_0 = 30^\circ$
85	0.136	0.144	0.150			0.174
80	0.117	0.128	0.132	0.147	0.156	0.164
75	0.106	0.120	0.132	0.144	0.156	0.172
70	0.100	0.118	0.130	0.144	0.159	0.185
65	0.096	0.117	0.131	0.145	0.168	0.215
60	0.095	0.116	0.132	0.150	0.178	0.250
55	0.095	0.116	0.134	0.156	0.186	
50	0.096	0.116		0.166	0.200	
45	0.098	0.117		0.180	0.215	

the cone at the end of the active region as functions of wrap angle and cone angle to be used to scale a design.

Example Design a conical log spiral with cone angle of 10° and wrap angle of 75° from 1 to 3 GHz.

$2\theta_0 = 20^\circ$. From Table 11-8, the upper truncation constant $a_3^-/\lambda = 0.069$. From Table 11-9, the lower truncation constant $a_{10}^+/\lambda = 0.156$. We use the radius a_3^-/λ with 3 GHz to determine the upper cone diameter and the radius a_{10}^+/λ with 1 GHz to determine the lower cone diameter:

$$\text{upper diameter} = 1.38\,\text{cm} \qquad \text{lower diameter} = 9.36\,\text{cm}$$

We compute the cone height from the projected central trapezoid:

$$\text{height} = \frac{D_L - D_u}{2\tan\theta_0} = 22.63\,\text{cm}$$

TABLE 11-10 Average Front/Back Ratio (dB) for a Conical Spiral Designed for the 10 : 1 Frequency Range Given the Cone Angle and Winding Wrap Angle

Cone Angle, 2θ	Wrap Angle (deg)								
	60	62.5	65	67.5	70	72.5	75	77.5	80
20	13.5	15.7	17.7	19.6	21.5	23.0	24.5	26.3	28.7
30	6.1	6.6	7.0	10.1	10.2	11.6	13.1	15.0	17.1
40					5.3	6.5	7.8	9.1	10.5

We determine the sense of circular polarization from the projection of the spiral on a plane ($\theta_0 = 90^\circ$) and by using hand rules with radiation toward the vertex. We use the same mode theory as the flat spiral to describe the radiation modes. The two-arm conical spiral radiates from the mode 1 with its peak on the boresight (on the axis). Table 11-10 lists the average front-to-back ratios for conical spirals averaged over a 10 : 1 frequency range and shows that long, thin cones produce the best F/B values.

Paraboloid Reflector Feed We can use the conical log spiral antenna as a broad bandwidth circularly polarized feed for a paraboloid reflector. The phase center moves along the cone axis when the frequency changes, but the illumination loss due to phase error is minor. Analysis of a number of possible feeds shows that the phase-center location depends almost entirely on cone geometry, not on wrap angle. The optimum location is dependent on the reflector f/D to a minor extent, but the location given by Table 11-11 is close to the optimum. Flatter cones have the phase center near the 1λ circumference, and as we narrow the cone it moves toward the virtual apex. Figure 11-12 for the conical spiral has a wider beamwidth than the flat spiral of Figure 11-1 because the phase center of the conical spiral occurs at 0.88λ circumference, whereas the spiral active region is centered at 1λ. The active region of this conical spiral has little axial length, which does decrease the beamwidth of the long, thin antennas (Table 11-7).

Table 11-12 lists the parameters of three conical spirals operating in mode 1 that produce the lowest average illumination losses as a paraboloid feed for designs that cover a 10 : 1 frequency range. Each design uses a wrap angle of 80° to reduce the pattern variation over the frequency range and to increase the F/B ratio to reduce spillover loss. Except for the lowest frequency range, the antennas have nearly constant beamwidth and produce nearly constant spillover and amplitude taper losses. Table 11-13 illustrates that varying reflector f/D has a slowly changing effect on the total average losses.

TABLE 11-11 Phase-Center Location of a Mode 1 Conical Spiral Relative to the Virtual Apex

Cone Angle, $2\theta_0$ (deg)	Axial Distance from Virtual Apex (λ)	Circumference (λ)
20	0.746	0.826
30	0.522	0.880
40	0.408	0.933

TABLE 11-12 Optimum Conical Spiral Mode 1 Feeds for a Paraboloid Reflector for a 10 : 1 Frequency Range

Cone Angle, $2\theta_0$ (deg)	f/D	Phase Error Loss (dB)		Taper Loss (dB)	Spillover Loss (dB)	Cross-Polarization Loss (dB)	Average Total (dB)
		Average	Maximum				
20	0.46	0.36	1.00	0.43	0.54	0.09	1.42
30	0.42	0.22	0.60	0.42	0.63	0.16	1.45
40	0.38	0.19	0.55	0.50	0.60	0.28	1.56

TABLE 11-13 Effect of Reflector f/D on Total Average Illumination Losses of the Designs of Table 11-12 (dB)

Cone Angle, $2\theta_0$ (deg)	Reflector f/D								
	0.34	0.36	0.38	0.40	0.42	0.44	0.46	0.48	0.50
20	2.24	1.97	1.76	1.61	1.51	1.44	1.42	1.42	1.44
30	1.68	1.55	1.48	1.45	1.45	1.47	1.53	1.60	1.68
40	1.63	1.58	1.56	1.58	1.63	1.70	1.79	1.89	2.00

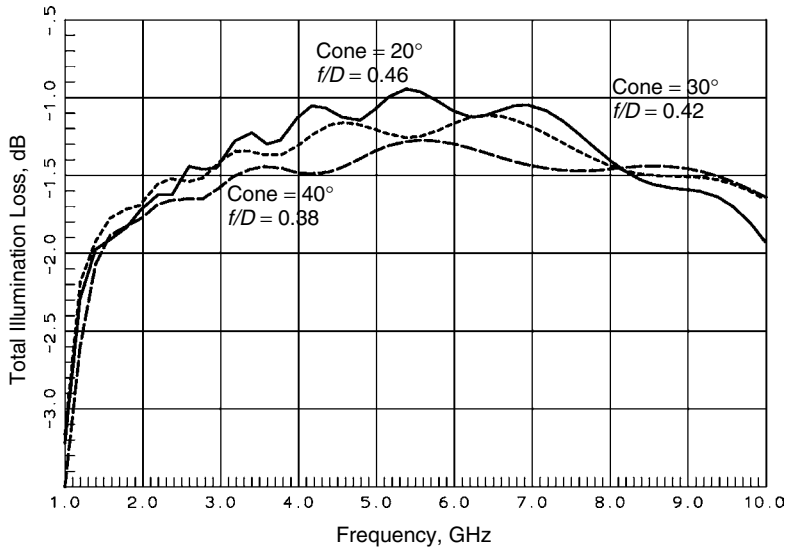


FIGURE 11-13 Total illumination losses for two-arm spirals designed for 1 to 10 GHz using Tables 11-8 and 11-9 feeding a paraboloid reflector.

Figure 11-13 gives the frequency response of the three reflector feeds. The narrower 20° design produces better reflector illumination over a narrower frequency range than that for the wider cones because the antennas were designed using truncation constants from Table 11-9, where the fields drop by 10 dB instead of 15 dB and the low-frequency pattern degrades.

11-10 MODE 2 CONICAL LOG SPIRAL ANTENNA

Conical log spiral antennas have limited available design information [17,18,19]. The beams off the boresight have half-power beamwidths ranging from 48 to 60°. Higher wrap angles and smaller cone angles decrease the beamwidth, but within this limited range. We cannot increase the axial length of the active region beyond a certain point, which limits the achievable gain. Although we have limited control of beamwidth, we can control the beam direction by the wrap angle. Table 11-14 lists the approximate beam peak given wrap angle on a 10° cone log spiral. Antennas built with cone angles in the range $2\theta_0 = 20$ to 40° follow Table 11-14 closely.

We must increase the diameters of the bottom and top of the four-arm mode 2 spiral from those calculated for the two-arm mode 1 conical log spiral. We obtain a suitable lower truncation constant from Dyson's results (Tables 11-8 and 11-9) if we multiply the lower truncation diameter by 1.42. The upper truncation constant multiplier varies linearly from 4 at $\alpha = 65°$ to 2.3 at $\alpha = 80°$. Since a little extra length on the top of the antenna will not degrade the pattern or increase the height significantly, we use the lower value for all designs.

Example Design a four-arm mode 2 conical log spiral antenna to point the beam at 50° over the frequency range 500 to 1500 MHz on a 10° cone.

Table 11-14 determines the wrap angle to scan the beam to 50°, $\alpha = 67°$. Table 11-8 gives us the upper truncation radius of the two-arm mode 1 spiral (0.055). Multiply by 2.3 for the mode 2 spiral:

$$\frac{a_3^-}{\lambda} = 2.3(0.055) = 0.126 \qquad \text{mode 2}$$

Similarly, Table 11-9 gives us the 10-dB lower truncation radius (0.165). We multiply it by 1.42 to calculate the lower truncation radius of the mode 2 spiral:

$$\frac{a_{10}^+}{\lambda} = 1.42(0.165) = 0.234 \qquad \text{mode 2}$$

We use a_3^-/λ with the highest frequency (1500 MHz) to determine the upper radius and use a_{10}^+/λ with the lowest frequency (500 MHz) to determine the lower radius of

TABLE 11-14 Beam Direction of a Mode 2 Conical Log Spiral Antenna for $2\theta_0 = 20°$

Wrap Angle, α (deg)	Beam Peak, θ (deg)	Wrap Angle, α (deg)	Beam Peak, θ (deg)
38	82	62	58
42	80	66	52
46	76	70	46
50	74	74	41
54	70	78	38
58	64		

Source: [19].

the truncated cone of the antenna:

$$D_U = 5.04\,\text{cm} \qquad D_L = 28.08\,\text{cm}$$
$$\text{height} = \frac{D_L - D_U}{2\tan\theta_0} = 65.33\,\text{cm}$$

11-11 FEEDING CONICAL LOG SPIRALS

We must feed the two-arm mode 1 spiral with a balanced line. The infinite balun consisting of coax soldered to the windings uses the truncation properties of the antenna to prevent currents induced on the outside of the coax from reaching the input. On the top we connect the center conductor of the coax to the second winding, which contains a dummy coax, to maintain symmetry. The coax length and associated loss become prohibitive for many spirals. We shorten the coax length by using a split tapered coax balun along the cone axis. Its bandwidth matches the antenna bandwidth.

It is difficult to tell if the four-arm mode 2 conical spiral antenna needs a balun. The pattern null on axis reduces the induced currents on the outside of a coax and we achieve suitable patterns without a balun. In some cases we see narrowband pattern distortion caused by the interaction of the outer-shield currents and higher-order modes. These cause pattern ripple in roll plane patterns (constant θ) near the beam peak. We can feed the antenna from four coax lines along the axis. We feed each winding separately and obtain cancellation of currents among the four shields soldered together.

Dyson says that we obtain the best patterns from complementary antennas—spacing equal stripwidth. The Babinet–Booker principle (Section 5-3) predicts an impedance of 188 Ω for a flat two-arm complementary spiral and 94 Ω for the four-arm spiral. Forming the spiral on a cone lowers the input impedance to about 150 Ω for the two-arm spiral and 85 Ω for the four-arm spiral. We can vary the stripwidth in the same way as a flat spiral to impedance-match the antenna with minor effects on the pattern.

LOG-PERIODIC ANTENNAS

All continuously scaled antennas radiate circular polarization. The point of constant beamwidth rotates with frequency. We can build linearly polarized self-scaling antennas only with structures that scale at discrete frequency intervals. The pattern characteristics will ripple between exact scaling frequencies, but with closely spaced scalings the antenna is practically frequency independent.

Every log-periodic structure has a basic scaling cell where we scale every dimension throughout the antenna by a constant:

$$\frac{f_1}{f_2} = \frac{\lambda_2}{\lambda_1} = \tau \qquad \text{scaling constant} \quad \tau < 1$$

The antenna will scale exactly at the sequence of frequencies: $f_n = f_0/\tau^n$. We make the antenna periodic in the logarithm of frequency with every dimension scaled by τ from element to element.

Log-periodic antennas were developed in the late 1950s out of modifications to the conical spiral concept of an antenna specified by angles. We will depart from an historical development and discuss the log-periodic dipole antenna first.

11-12 LOG-PERIODIC DIPOLE ANTENNA [20–23]

The design of a log-periodic antenna proceeds in two parts. First, the desired pattern characteristics determine the required number of elements in the active region and the element spacing. Second, we determine truncation points from the current levels on the antenna to establish the number of elements required for a given frequency range. As is true of the conical spiral, only a limited range of gains is possible because the aperture length is limited.

Figure 11-14 shows the log-periodic dipole antenna with a crisscross feeder line. We denote the longest dipole length by L_1. The element ends lie along lines that meet at the virtual apex. We measure the distance from the virtual apex to the dipole by R_n. The distance between elements is d_n. Starting with initial dimensions L_1, R_1, and d_1, we iterate all other dimensions by using the scaling constant τ:

$$L_2 = \tau L_1 \qquad R_2 = \tau R_1 \qquad d_2 = \tau d_1 \qquad L_3 = \tau L_2 = \tau^2 L_1 \qquad \text{etc.}$$

In general,

$$L_n = \tau^{n-1} L_1 \qquad R_n = \tau^{n-1} R_1 \qquad d_n = \tau^{n-1} d_1 \tag{11-22}$$

Note: d_n is not an independent variable, since

$$d_n = R_n - R_{n+1} = R_n(1 - \tau) \tag{11-23}$$

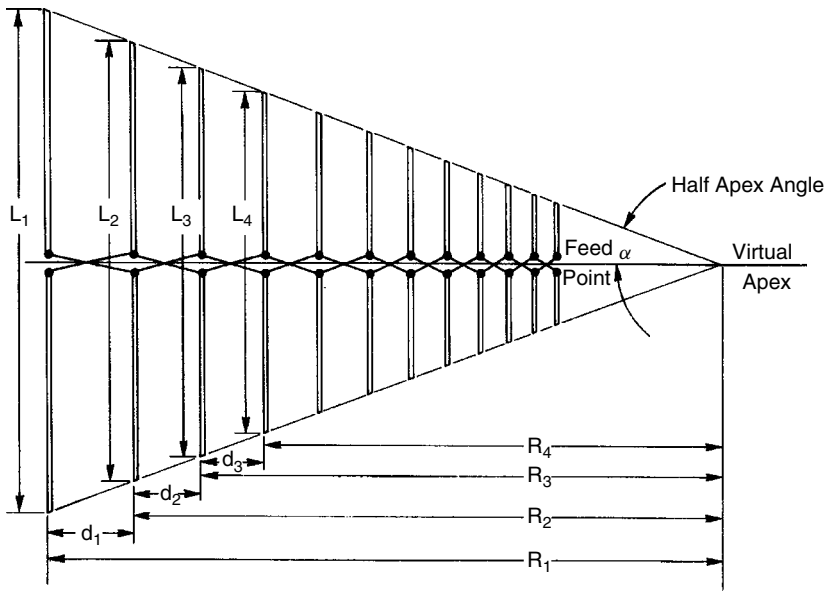


FIGURE 11-14 Log-periodic dipole antenna.

The angle between the dipole endpoints and the centerline—α, the half apex angle—is given by

$$\alpha = \tan^{-1} \frac{L_n}{2R_n} \tag{11-24}$$

Carrel introduced a spacing constant σ as a second constant to describe the antenna:

$$\sigma = \frac{d_n}{2L_n} \tag{11-25}$$

We specify the log-periodic dipole antenna by the constants τ and σ. We can compute α, the half apex angle, from τ and σ:

$$\alpha = \tan^{-1} \frac{1 - \tau}{4\sigma} \tag{11-26}$$

A method of moment analysis combining coupling between dipoles and transmission-line networks (Section 10-3.1) [23] was used to calculate the frequency response of a number of designs. These results were averaged over the operating frequency to obtain the values in tables below. Log-periodic dipole responses contain narrow-frequency-range regions where the pattern becomes distorted and the analysis summaries disregarded these regions. Nevertheless, the antenna response varies significantly. Table 11-15 lists the average gain for length/element diameter as 70. The element diameters were scaled by τ to maintain the same ratio (L/d) throughout the antenna. Tables 11-16 and 11-17 list the E- and H-plane average beamwidths over a suitable range of parameters τ and σ. The beamwidths vary between scaled frequency points. We use Tables 11-15 to 11-17 to determine suitable design constants. Although these tables give the average values for gain and beamwidth, the values have considerable ripple about the average that increases as τ is reduced.

We combine the desired operating frequency range with upper and lower truncation constants determined for τ and σ to compute the length of the longest element and

TABLE 11-15 Calculated Average Gain of a Log-Periodic Dipole Antenna for Length/Diameter = 70

	Scaling Constant, τ								
σ	0.80	0.82	0.84	0.86	0.88	0.90	0.92	0.94	0.96
0.06	6.0	6.5	6.0	6.0	6.9	7.2	7.9	8.6	9.7
0.07	5.2	6.1	6.9	6.8	6.9	7.6	8.2	9.0	10.1
0.08	5.5	5.6	6.7	7.2	7.1	7.8	8.4	9.2	10.4
0.09	6.0	5.7	6.0	7.3	7.7	7.8	8.8	9.6	10.7
0.10	6.6	6.4	6.1	6.5	7.7	8.2	8.8	9.7	10.9
0.12	6.5	6.9	7.3	7.5	7.7	8.3	9.3	10.1	11.4
0.14	6.3	6.7	7.1	7.5	8.0	8.6	9.5	10.4	11.7
0.16	6.7	7.1	7.6	8.0	8.4	8.7	9.4	10.6	11.9
0.18	6.3	6.8	7.5	8.1	8.8	9.3	9.8	10.6	12.1
0.20	5.7	5.9	6.4	7.2	8.1	9.0	10.0	10.8	12.1
0.22	5.3	5.3	5.7	6.3	7.2	8.3	9.6	10.8	12.1

TABLE 11-16 Calculated Average *E*-Plane Beamwidth of a Log-Periodic Dipole Antenna for Length/Diameter = 70

	Scaling Constant, τ								
σ	0.80	0.82	0.84	0.86	0.88	0.90	0.92	0.94	0.96
0.06	60	59	86	78	61	69	66	62	58
0.07	76	61	57	73	69	67	64	61	56
0.08	83	76	61	63	71	61	64	60	55
0.09	75	83	72	60	62	69	62	58	54
0.10	57	76	81	72	63	62	62	58	53
0.12	66	61	60	63	69	65	59	57	51
0.14	73	69	67	65	63	63	61	54	50
0.16	64	65	64	63	62	63	61	54	49
0.18	69	66	66	64	60	58	56	55	48
0.20	78	80	80	76	71	64	57	53	48
0.22	81	81	86	84	79	71	62	55	48

TABLE 11-17 Calculated Average *H*-Plane Beamwidth of a Log-Periodic Dipole Antenna for Length/Diameter = 70

	Scaling Constant, τ								
σ	0.80	0.82	0.84	0.86	0.88	0.90	0.92	0.94	0.96
0.06	157	127	118	150	118	120	104	92	78
0.07	171	146	111	122	124	107	97	87	74
0.08	166	156	123	101	122	98	98	83	70
0.09	115	159	135	106	96	108	90	80	68
0.10	103	124	142	122	99	100	88	77	65
0.12	108	99	95	106	113	95	82	74	62
0.14	121	115	107	99	95	100	82	71	59
0.16	107	106	100	96	95	92	82	68	58
0.18	121	109	99	90	82	77	74	70	56
0.20	135	131	123	111	98	86	73	66	56
0.22		149	136	126	116	100	82	67	56

determine the number of elements required. The longest element length is given by

$$L_1 = K_1 \lambda_L \tag{11-27}$$

where λ_L is the longest operating wavelength and K_1 is the lower truncation constant.

We determine K_1 from the empirical equation [22]

$$K_1 = 1.01 - 0.519\tau \tag{11-28}$$

Equation (11-28) overestimates K_1 for $\tau > 0.95$, and the lower band edge will be extended slightly. We calculate the upper truncation constant from

$$K_2 = 7.08\tau^3 - 21.3\tau^2 + 21.98\tau - 7.30 + \sigma(21.82 - 66\tau + 62.12\tau^2 - 18.29\tau^3) \tag{11-29}$$

another empirical equation. The shortest element length is $L_U = K_2\lambda_U$, where λ_U is the shortest operating wavelength. We use the truncation constants and the frequency band edges to determine the number of dipoles in the antenna:

$$N = 1 + \frac{\log(K_2/K_1) + \log(f_L/f_U)}{\log \tau} \tag{11-30}$$

For a given frequency, $f_L = f_U = f$, and we compute the number of elements in the active region from Eq. (11-30):

$$N_a = 1 + \frac{\log(K_2/K_1)}{\log \tau} \tag{11-31}$$

Increasing the number of elements in the active region increases gain. We combine Eqs. (11-23) and (11-25) to determine the virtual apex distance:

$$R_n = \frac{2L_n\sigma}{1-\tau} \tag{11-32}$$

The axial length of the antenna is the difference between R_1 and R_N:

$$\text{length} = R_1 - R_N = R_1(1-\tau^{N-1}) = \frac{2L_1\sigma(1-\tau^{N-1})}{1-\tau} \tag{11-33}$$

We compute the dimensions of the antenna from the equations above by using an integer number of dipoles [Eq. (11-30)].

Example Design a log-periodic dipole antenna to operate from 100 to 1000 MHz.

Use $\tau = 0.9$ and $\sigma = 0.15$. We estimate the E- and H-plane average beamwidths from Tables 11-16 and 11-17.

$$E\text{-plane beamwidth} = 63^\circ \qquad H\text{-plane beamwidth} = 96^\circ$$

We use Eqs. (11-28) and (11-29) to compute the truncation constants: $K_1 = 0.54$ and $K_2 = 0.32$. We calculate the length L_1 by using K_1 and the lowest operating wavelength [Eq. (11-27)]: $L_1 = K_1\lambda_{100\,\text{MHz}} = 162\,\text{cm}$. We determine the number of elements from Eq. (11-30), $N = 28$, and substitute N into Eq. (11-33) to determine the total length, 457.7 cm. We rearrange Eq. (11-25) to calculate the first spacing, $d_1 = 2\sigma L_1 = 48.6\,\text{cm}$. We compute the virtual apex distance from Eq. (11-32):

$$R_1 = \frac{2L_1\sigma}{1-\tau} = 486\,\text{cm}$$

We use these dimensions and the scaling constant to iterate the rest of the antenna dimensions by using Eq. (11-22). For example,

$$L_2 = \tau L_1 = 145.8\,\text{cm} \qquad R_2 = \tau R_1 = 437.4\,\text{cm} \qquad d_2 = \tau d_1 = 43.74\,\text{cm}$$

We can estimate the gain from the length of the active region. We calculate the half apex angle from Eq. (11-26): 9.46°. The axial length of the active region

is $(K_1 - K_2)/\tan\alpha = 1.32\lambda$, and we use Eq. (10-12) for the directivity of a linear end-fire antenna:

$$\text{directivity} = \frac{4L}{\lambda} = 5.28 \ (7.2\,\text{dB})$$

The dipole pattern of the elements decreases the average radiation intensity and increases directivity by 1.4 dB, somewhat less than the 2.1 dB of the dipole because effective radiation length had already shaped the pattern. We read a gain of 8.6 dB from Table 11-15.

Table 11-15 has multiple points with similar gain. The length of the active region determines gain, but increasing the number of elements reduces the ripple in the frequency response. We can also reduce the ripple by increasing the element diameters or, equivalently, by using flat trapezoidal teeth for elements instead of thin dipoles. We substitute a dipole for a flat strip by using a diameter equal to half the strip width. Three antennas were designed to cover the frequency range 100 to 1000 MHz: (1) $\tau = 0.8$, $\sigma = 0.1$, and $L/d = 140$; (2) $\tau = 0.8$, $\sigma = 0.1$, and $L/d = 25$; (3) $\tau = 0.9$, $\sigma = 0.06$, and $L/d = 140$. The first two designs contain 16 elements with a boom length of 1.78 m, and the second one has 26 elements along a 1.96-m boom.

Figure 11-15 plots the frequency response of gain for the three designs. All three plots show narrow frequency ranges where the gain drops due to the combination of adverse coupling and feed network. This analysis ignores the currents induced on the feed lines, which produce additional dropout regions. When we use the lower truncation constant to design the longest element, the antenna has its full gain at the lower-frequency edge. The solid and long-dashed curves plot the responses of the antennas with $\tau = 0.8$ and $\sigma = 0.1$ and demonstrate a similar ripple period, whereas the design with more elements ($\tau = 0.9$ and $\sigma = 0.06$), plotted with short dashes, shows little ripple. The antenna with thin elements produces a design with more than 1 dB of gain ripple. The thicker elements also reduce the level of gain dropouts in the response.

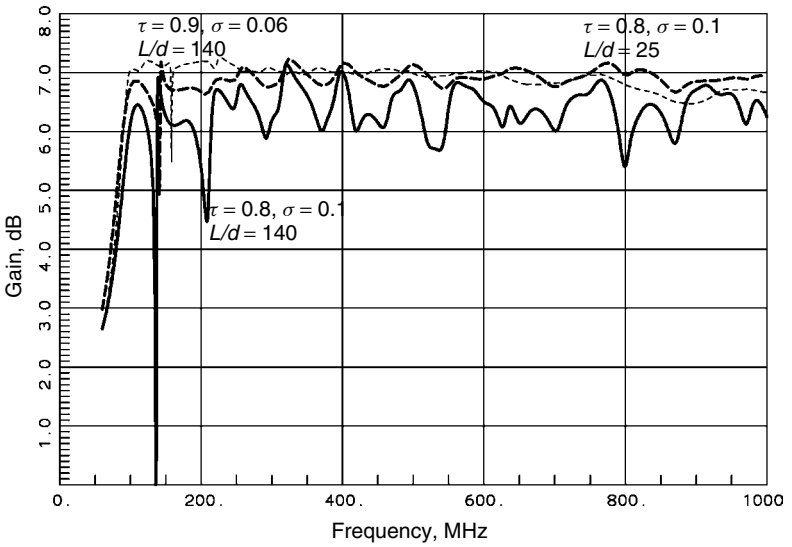


FIGURE 11-15 Frequency response of 16-element log-periodic dipole antennas designed for a 10 : 1 frequency range.

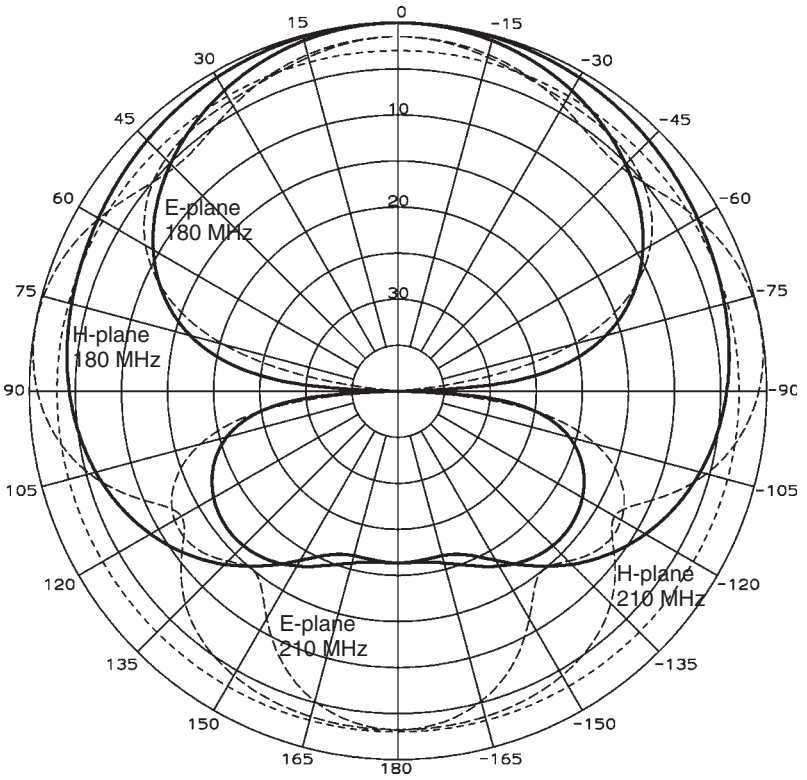


FIGURE 11-16 Patterns of a 16-element log-periodic dipole antenna with $\tau = 0.80$ and $\sigma = 0.1$.

Figure 11-16 gives the E- and H-plane patterns of case 1 at a normal frequency (180 MHz, solid) and at a frequency in the dropout region (210 MHz, dashed). The E-plane pattern contains a null at 90° due to the dipoles. The H-plane pattern shows how the pattern broadens and no longer has a traveling wave phasing along the elements in the active region to produce a significant front-to-back ratio. The pattern almost reduces to that of a single dipole. The front-to-back ratio is a more telling indicator than boresight gain of phasing problems in the log-periodic antenna. Antennas using thin dipoles and few elements in the active region have many regions with poor response. Increasing the diameter of the elements improves the F/B ratio. Increasing τ both improves F/B and increases the ripple rate.

11-12.1 Feeding a Log-Periodic Dipole Antenna

The antenna of Figure 11-14 must be fed from a balanced line; antennas designed for HF frequencies use that type of feed. If we can run a coax along the center of the antenna, we can use an infinite balun. We alternate the direction of the elements (Figure 11-17) on the outer shield of the coax feeder with a dummy coax to achieve the crisscross. The combination of the feed coax and dummy coax form a two-wire transmission line that feeds the elements. The truncation property of the antenna inhibits the flow of induced current on the feeder beyond the active region from reaching the input.

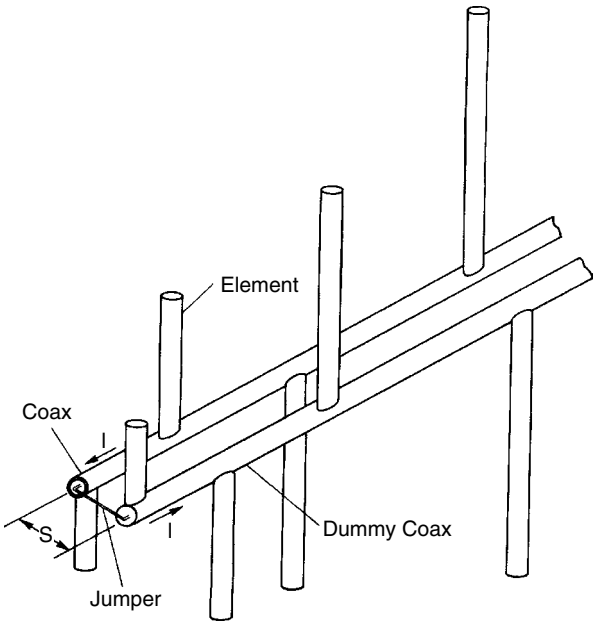


FIGURE 11-17 Log-periodic dipole antenna feed with an infinite balun.

We use a crisscross feed to increase the phase velocity on the feeder in the active region. If we used straight feeders between dipoles, the phase delay would equal that necessary for end fire in the direction of the feeder currents. This would produce a pattern in the direction of increasing structure and violate the truncation requirement. The extra 180° phase shift between elements produces a backfiring fast wave in the active region.

The region before the active region is a transmission line loaded with small open-circuited stubs (the dipoles). These capacitively load the line and reduce the effective characteristic impedance of the two-wire line. Each short dipole has capacitance

$$Z = -jZ_a \cot \frac{kL_i}{2} \tag{11-34}$$

where L_i is the total dipole length, k the wave number $(2\pi/\lambda)$, and Z_a the average characteristic impedance of the dipole:

$$Z_a = 120\left(\ln \frac{L_i}{2a} - 2.25\right) \tag{11-35}$$

for a dipole radius a. We use Eq. (11-35) with a constant length/diameter ratio in Eq. (11-34) to compute an effective added capacitance per unit length along the feeder due to the nonresonant dipoles. With a little manipulation we reduce this to an expression in terms of the log-periodic dipole antenna parameters. The effective feeder impedance R_0 is related to the unloaded two-wire impedance Z_0 by

$$R_0 = \frac{Z_0}{\sqrt{1 + (\sqrt{\tau}Z_0)/\sigma Z_a}} \tag{11-36}$$

If the length-to-diameter ratio of the dipoles remains constant, R_0 is constant along the antenna length. Even with a changing Z_0 or Z_a, the feeder acts as a tapered transmission-line transformer. We may expect impedance variations about the nominal value of Eq. (11-36) with a cycle of τ. The currents on the feeder radiate, but because they are close enough together and are nearly equal and opposite, they cancel in the far field. The feeder currents and the jumper at the feed limit the cross-polarization response to about 20 dB, which ripples with frequency changes.

The jumper between the center conductor of the coax feeder and the dummy coax line will squint the beam toward the dummy coax at high frequencies. We can represent the jumper as a series inductor in the transmission line, and construction difficulties with this jumper connection limit the high-frequency operation of log-periodic dipole antennas.

11-12.2 Phase Center

We expect the antenna phase center to be in the middle of the active region. If the antenna were made from many half-wavelength dipoles each resonant at a single frequency, we would look to the one $\lambda/2$ long at a given frequency to be the phase center. In the preceding example the active region dipoles ranged from 0.32λ to 0.54λ. We expect the phase center to be in front of the $\lambda/2$ element (or location along the triangle of the element ends of a possible element), since a great deal of the active region has elements of less than $\lambda/2$.

Table 11-18 lists the approximate E- and H-plane phase-center locations measured from the virtual apex relative to a $\lambda/2$ element. We may not have a $\lambda/2$ element at a given frequency; but given the envelope of elements defined by the apex angle, we compute the location of a possible element. Table 11-18 shows the astigmatism of the antenna with the H-plane phase center behind the E-plane phase center. The phase-center distance from the virtual apex increases linearly with λ.

Example Compute E- and H-plane phase centers of the antenna designed above for the range 100 to 1000 MHz ($\tau = 0.9$) at 600 MHz.

The half apex angle is 9.46°. The apex distance of the half-wavelength element is

$$R_g = \frac{\lambda}{4 \tan \alpha} = \frac{50}{4 \tan 9.46^\circ} = 75.02 \text{ cm}$$

TABLE 11-18 Log-Periodic Dipole Antenna Phase Center R_p Measured from the Virtual Apex Relative to the Location of a $\lambda/2$ Element $R_g = \lambda/(4 \tan \alpha)$

Scaling Constant, τ	R_p/R_g		Scaling Constant, τ	R_p/R_g	
	E-Plane	H-Plane		E-Plane	H-Plane
0.80	0.959	0.997	0.90	0.862	0.874
0.82	0.939	0.968	0.92	0.849	0.859
0.84	0.928	0.941	0.94	0.842	0.849
0.86	0.897	0.916	0.96	0.840	0.844
0.88	0.878	0.893			

Source: [22].

From Table 11-18 we read $R_p/R_g(E\text{-plane}) = 0.862$ and $R_p/R_g(H\text{-plane}) = 0.874$. The phase center distances from the virtual apex are

$$R_p(E - \text{plane}) = 64.67\,\text{cm} \quad R_p(H - \text{plane}) = 65.57\,\text{cm}$$

This astigmatism of 0.018λ produces an insignificant loss as a paraboloidal reflector feed (Figure 8-4).

11-12.3 Elevation Angle

The dipole elements must be connected alternately to the coax feeder and the dummy coax, but the feeder lines can diverge (Figure 11-18). We direct the feeder lines at angles $\pm\psi$ from the antenna axis. For the antenna to remain frequency independent, it is necessary for the projection of the two feeders to intersect at the virtual apex. Spacing the sides at angles $\pm\psi$ decreases the H-plane beamwidth since the aperture size increases in that plane. The elevation angle moves the antenna phase center toward the virtual apex and reduces movement with frequency changes. We must analyze the feeder line as a tapered transmission line. Moving the sides apart will increase the

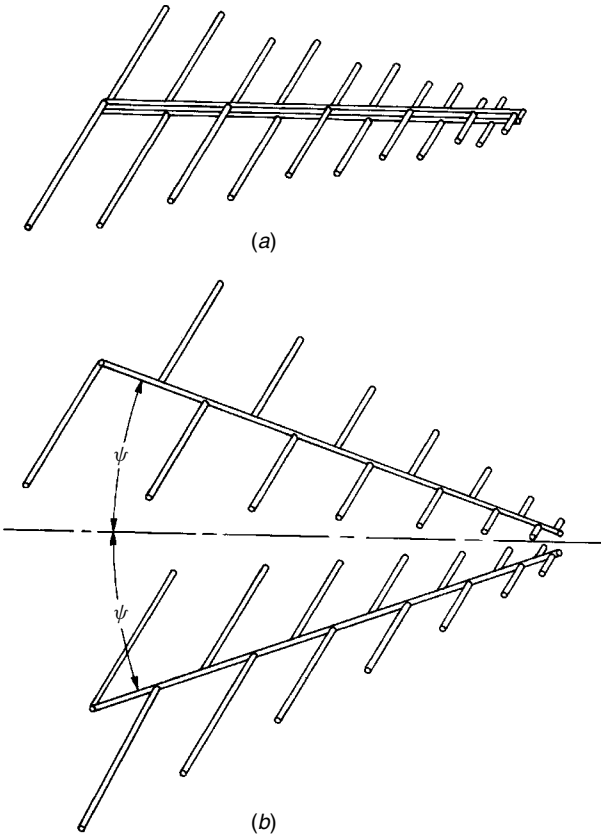


FIGURE 11-18 Log-periodic dipole antennas with the feeder diverged by elevation angle.

backlobe of the pattern because in the limit of $\psi = 90^\circ$, the front and back lobes are equal.

11-12.4 Arrays of Log-Periodic Dipole Antennas [24]

We can make broadband arrays with log-periodic antennas. Like the elevation angle of the single antenna, we have frequency-independent arrays only if the virtual apexes of all the elements of the array are coincident. The elements must also have the same τ and σ. Figure 11-19 shows E- and H-plane arrays. The relative phasing between elements can be changed in a frequency-independent manner. If an antenna is turned over, its far-field phase changes by 180°. In an array of two elements this will produce a null on the axis between them, the effect of placing a horizontally polarized antenna over a ground plane. On a particular antenna, if we multiply every element by the scaling constant, the far-field phase shifts by 180°. Multiplying every element by τ is equivalent to turning the antenna over (somewhere in the middle of the frequency band). Adding elements at the feed end does not change the location of the phase center. We can change the phase arbitrarily by multiplying the antenna dimensions by $\tau^{\gamma/180^\circ}$, where γ is the phase shift. Changing the phase by $\tau^{\gamma/180^\circ}$ has meaning only in an array.

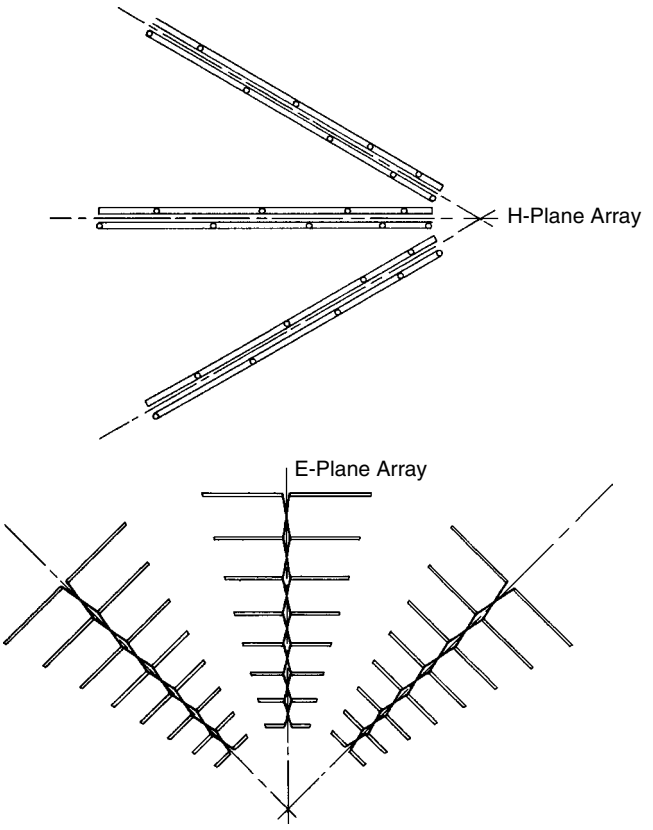


FIGURE 11-19 Arrays of log-periodic antennas.

We can build a frequency-independent circularly polarized antenna from two log-periodic antennas. We orient the two antennas at right angles with one of them scaled by $\tau^{1/2}$. When fed in phase, the pair combines 90° out of phase and radiates circular polarization. When we array log-periodic antennas, they may develop narrowband gain dropouts [25]. These dropouts occur in a sequence of frequencies with the scaling constant τ of the antenna. Antennas arrayed in either the E- or H-plane show these dropouts, as do antennas arrayed orthogonally. If we mismatch the scaling constants between the antennas in the array, each has dropouts scaled by its own scaling constant. Single antennas also can develop gain dropouts, although they seldom occur in a sequence of frequencies.

The phenomenon points to the importance of swept gain measurements because the location of the frequencies of dropouts remains somewhat unpredictable except for the sequence. Moment method models of the antenna will predict dropouts, but a large number of cases must be run to locate the frequencies. Unbalanced currents on the feeder lines interact with the elements to produce cross-polarization and large amounts of off-boresight radiation that lowers gain. These unwanted currents on the feeder are produced by either asymmetry in the antenna or interactions between antennas.

11-13 OTHER LOG-PERIODIC TYPES [26, 27]

Many different types of log-periodic antennas have been built. If we discover a structure that satisfies the self-scaling antenna requirements and has the desired polarization, we will have a new frequency-independent antenna. The first log-periodic antennas were teeth cut into the sides of bifin antennas. These radiated equally on both sides with polarization rotated parallel to the teeth instead of along the feed line. Isbell folded the sides at elevation angles to produce a unidirectional pattern. DuHamel and Ore [27] straighten the teeth to form the trapezoidal tooth antenna (Figure 11-20). We separate

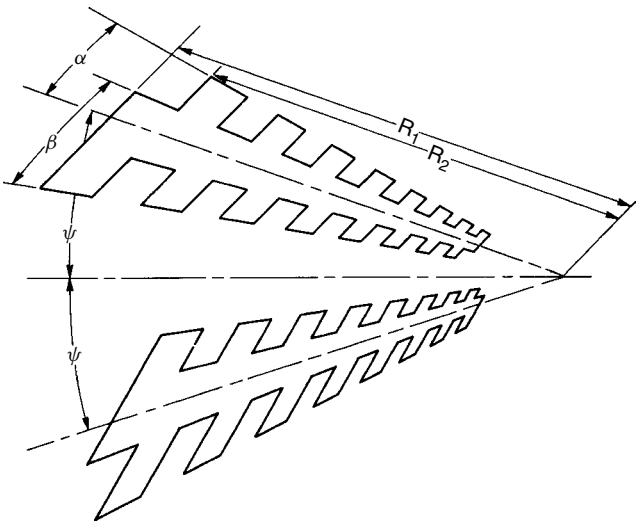


FIGURE 11-20 Trapezoidal tooth log-periodic antenna.

the sides by an angle 2ψ, and they project to the virtual apex. This is a good high-frequency antenna because it can be largely self-supporting. Figure 11-20 shows the teeth with the same width as the spacings. We can reduce the width of the teeth, but we continue to measure the distances R_n to the bottom of the teeth. As we continue to reduce the width of the teeth, the antenna transforms into a log-periodic dipole antenna.

When we remove material to form the teeth of a wire outline trapezoidal tooth antenna (Figure 11-21), it has only minor effects on the pattern. The shape of the teeth is not too important as long as the teeth scale with τ because the triangular tooth wire outline antenna (Figure 11-22) also works well. This tooth shape reduces some of the

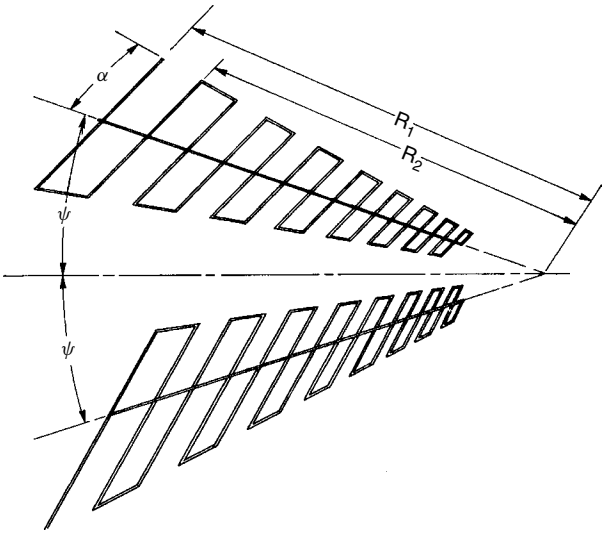


FIGURE 11-21 Trapezoidal tooth wire outline log-periodic antenna.

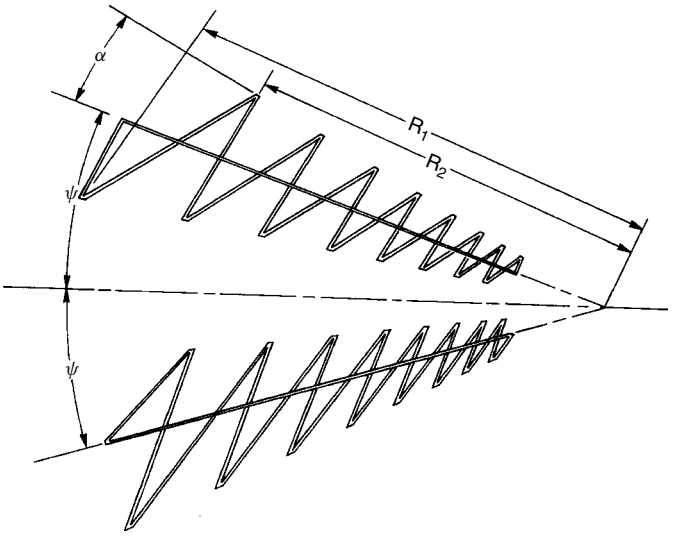


FIGURE 11-22 Triangular tooth wire outline log-periodic antenna.

TABLE 11-19 Trapezoidal Tooth Wire Outline Log-Periodic Antenna Designs

Scaling Constant, τ	Elevation Angle, ψ	Half Apex Angle, α	Beamwidth		Directivity (dB)	Sidelobes (dB)
			E-Plane	H-Plane		
0.63	15	30	85	153	5.0	12
0.63	15	37.5	74	155	5.6	12.4
0.71	15	30	70	118	7.0	17.7
0.71	15	37.5	66	126	7.0	17.0
0.63	22.5	30	86	112	6.3	8.6
0.63	22.5	37.5	72	125	6.6	11.4
0.71	22.5	30	71	95	7.9	14.0
0.71	22.5	37.5	67	106	7.6	14.9
0.77	22.5	30	67	85	8.6	15.8
0.84	22.5	22.5	66	66	9.8	12.3
0.84	22.5	30	64	79	9.1	15.8
0.63	30	30	87	87	7.4	7.0
0.63	30	37.5	73	103	7.4	8.6
0.71	30	30	71	77	8.8	9.9
0.71	30	37.5	68	93	8.1	12.8

Source: [27].

construction problems of the trapezoidal tooth design, especially at low frequencies. Because the wider teeth have greater coupling than dipoles, we can use smaller scaling constants τ and achieve good designs. Antennas have been built with $\tau = 0.63$, and the log-periodic dipole has a lower limit of about 0.80 for a good response. Table 11-19 lists the parameters of successful trapezoidal tooth wire outline antennas [27] in terms of log-periodic dipole parameters. Increasing the elevation angle decreases the H-plane beamwidth and raises the sidelobes. The H-plane beamwidth decreases as we increase the half apex angle, but the dipole-type pattern of the E-plane is its predominant factor. In the range of scaling constants given, the directivity increases with increasing scaling constant.

11-14 LOG-PERIODIC ANTENNA FEEDING PARABOLOIDAL REFLECTOR

The combination of a log periodic antenna feeding a paraboloidal reflector produces a wideband antenna with high gain. The significant problem is the movement of the phase center of the antenna over the frequency range, which causes phase error loss. For a given frequency range we locate the weighted-average phase center at the focus of the reflector. You determine the phase center by analysis or measurement at the lowest and highest frequencies of operation and calculate the best position to locate at the reflector focus:

$$PC_{\text{avg}} = \frac{PC_L F_L + PC_U F_U}{F_L + F_U} \tag{11-37}$$

This location of the phase center, PC_{avg}, produces the same loss at the lower frequency, F_L, and the upper frequency, F_U. You can extend Eq. (11-37) to include all frequencies for calculation of the average phase center as a weighted average. If you

build a log-periodic antenna using dipoles, having unequal beamwidths in the E- and H-planes generates cross-polarization in the diagonal plane when converted to dual-polarized Huygens sources. Each Huygens source excites uniformly polarized currents on the reflector. The nearly constant beamwidth of the antenna at all frequencies produces a constant sum of the spillover and taper loss over the frequency range of the feed operation.

Parameters for both broad- and narrowband designs were found to optimize the gain of the reflector using Carrell's analysis method [28]. Table 11-20 lists the design parameters of a wideband design given by the scaling constant τ and the half apex angle α. Antennas designed to the parameters of Table 11-20 to operate over a 10 : 1 frequency range were analyzed as feeds for a paraboloid reflector. We use Eq. (11-37) to locate the phase center so that we have the same phase error loss at the upper and lower frequencies. Table 11-21 lists the results.

TABLE 11-20 Parameters of Wide-Bandwidth Log-Periodic Dipole Feed for a Reflector

Reflector, f/D	Scaling Constant, τ	Half Apex Angle, α	Average Gain (dB)	Beamwidth	
				E-Plane	H-Plane
0.25	0.855	41.7	6.1	61.3	143.3
0.30	0.867	40	6.0	73.0	141.4
0.35	0.869	37.5	6.1	72.2	139.4
0.40	0.900	32	6.5	69.3	126.0
0.45	0.914	27	6.8	67.0	124.1
0.50	0.923	22	7.3	66.6	112.2
0.55	0.928	20	7.4	66.9	112.4
0.60	0.934	17.5	7.8	65.5	105.6
0.65	0.940	16	8.0	64.2	101.0
0.70	0.944	13.5	8.4	62.7	94.2
0.75	0.947	12	8.8	61.6	90.0

TABLE 11-21 Reflector Illumination Losses of Log-Periodic Dipole Antennas Designed to Operate over a 10 : 1 Frequency Range

Reflector, f/D	Number of Elements	Maximum PEL (dB)	Average ATL (dB)	Average SPL (dB)	Cross-Polarization Loss	Total (dB)		
						Minimum	Maximum	Average
0.25	19	0.60	1.52	0.46	0.40	2.32	3.37	2.64
0.30	21	0.44	0.95	0.73	0.36	1.93	2.94	2.20
0.35	23	0.53	0.97	0.71	0.35	1.97	2.95	2.23
0.40	26	0.38	0.44	1.22	0.32	2.05	2.60	2.21
0.45	30	0.40	0.36	1.52	0.28	2.10	2.70	2.30
0.50	33	0.44	0.26	1.63	0.23	2.04	2.80	2.28
0.55	36	0.38	0.20	1.90	0.21	2.19	3.08	2.44
0.60	40	0.38	0.16	2.07	0.17	2.30	3.10	2.53
0.65	43	0.34	0.12	2.33	0.15	2.48	3.30	2.73
0.70	47	0.36	0.11	2.40	0.11	2.50	3.40	2.74
0.75	50	0.36	0.09	2.55	0.09	2.61	3.50	2.85

Repeating the analysis of the antennas in Table 11-21 over a 20 : 1 frequency range for f/D between 0.30 and 0.50 increased the maximum loss by only 0.1 dB. The wideband antennas use short lengths to limit the phase-center movement as the operating frequency changes and radiate wide H-plane beamwidths from their short active region. The ideal feed for a reflector has nearly equal beamwidths to reduce spillover and reduce cross-polarization in the diagonal planes. Table 11-22 gives a second set of designs that produce optimum illumination from the log-periodic dipole antenna for a narrow band of frequencies. If we consider an additional 0.5 dB loss, similar to the 2 : 1 VSWR, the bandwidth of these antennas is approximately 70%, or equivalently, a 2 : 1 frequency range. Of course, the antenna needs to be designed with sufficient elements to cover the frequency range.

We reduce the number of elements required by using a trapezoidal tooth design (Figure 11-20). These designs work best with a reflector f/D between 0.40 and 0.55. We separate the sides by an angle 2ψ to decrease the H-plane beamwidth. Table 11-23 lists the results from a method of moments analysis for varying ψ on an antenna designed to operate over a 10 : 1 frequency range (16 elements) with $\tau = 0.80$, $\sigma = 0.125$ ($\alpha = 21.8^\circ$), and tooth width/element spacing $= \frac{1}{3}$. The MOM analysis includes the currents flowing on the feeder lines that radiate increasing cross-polarization as the sides separate. This design peaks at $f/D = 0.48$. Table 11-24 shows that the illumination losses change very slowly for different f/D and it is difficult to say that the feed peaks for a particular f/D.

A trapezoidal tooth log-periodic antenna was built and tested as a feed for a reflector. Table 11-25 lists the average and maximum sum of the illumination losses calculated

TABLE 11-22 Optimum Narrowband Log-Periodic Dipole Antennas Reflector Feeds

f/D	Scaling Constant, τ	α	Gain (dB)	Beamwidth (deg)		ATL (dB)	SPL (dB)	Cross-Polarization Loss	Minimum Total (dB)
				E-Plane	H-Plane				
0.30	0.920	18.5	7.7	66.6	108.1	1.35	0.27	0.18	1.79
0.35	0.928	15	8.1	64.1	99.8	1.00	0.33	0.13	1.46
0.40	0.940	12	8.7	61.8	90.1	0.78	0.39	0.10	1.23
0.45	0.944	10	9.1	60.3	85.5	0.61	0.47	0.08	1.10
0.50	0.947	8.2	9.6	58.4	79.6	0.50	0.52	0.06	1.02
0.55	0.952	7	10	56.5	74.3	0.42	0.58	0.04	0.99
0.60	0.952	5.3	10.7	53.8	68.0	0.34	0.54	0.03	0.96

TABLE 11-23 Trapezoidal Tooth Log-Periodic Antenna Feed for a Reflector with $f/D = 0.48$, 10 : 1 Frequency Range (16 Elements) with $\tau = 0.80$, $\sigma = 0.125$ ($\alpha = 21.8^\circ$)

ψ	Beamwidth (deg)		Maximum PEL (dB)	Average ATL (dB)	Average SPL (dB)	Average XOL (dB)	F/B (dB)	Total (dB)	
	10-dB E-Plane	10-dB H-Plane						Average	Maximum
5	108.1	180.4	0.54	0.44	1.13	0.23	20	1.98	2.51
10	108.0	158.1	0.53	0.45	0.93	0.26	15	1.82	2.37
15	110.2	143.3	0.57	0.42	0.83	0.38	13	1.87	2.52
20	113.5	124.7	0.49	0.52	0.84	0.60	12	2.15	2.80

TABLE 11-24 Illumination Loss Variation for Trapezoidal Tooth Log-Periodic Antenna Reflector Feed Over a 10 : 1 Frequency Range (16 Elements) with $\tau = 0.80$, $\sigma = 0.125$ ($\alpha = 21.8^\circ$), and $\psi = 15^\circ$

	f/D							
Loss (dB)	0.40	0.42	0.44	0.46	0.48	0.50	0.52	0.54
Average	1.94	1.88	1.86	1.85	1.87	1.91	1.95	2.01
Maximum	2.80	2.68	2.60	2.54	2.52	2.52	2.53	2.55
Minimum	1.58	1.54	1.53	1.55	1.58	1.67	1.67	1.73

TABLE 11-25 Measured Illumination Losses for Trapezoidal Tooth Log-Periodic Antenna Reflector Feed Over a 6 : 1 Frequency Range with $\tau = 0.80$, $\sigma = 0.1$ ($\alpha = 26.6^\circ$), and $\psi = 6^\circ$

	f/D							
Loss (dB)	0.40	0.42	0.44	0.46	0.48	0.50	0.52	0.54
Average	1.78	1.69	1.64	1.63	1.64	1.68	1.74	1.82
Maximum	2.73	2.53	2.38	2.28	2.24	2.40	2.47	2.56

At $f/D = 0.46$:

Beamwidth (deg)		Maximum	Average	Average	Average	Average	Maximum
10-dB E-Plane	10-dB H-Plane	PEL (dB)	ATL (dB)	SPL (dB)	XOL (dB)	Total (dB)	Total (dB)
107.6	147.0	0.58	0.49	0.63	0.28	1.63	2.28

from measurements over a 6 : 1 frequency range. An infinite balun feed was made by attaching the feeding coax to the central plate of one side and connecting the coax center conductor to a coax outer shield attached to the second arm. Small movements between the two sides rapidly alter the input impedance, and a few experiments involving changing their spacing produce an acceptable return loss. Table 11-25 shows results similar to Table 11-24 in that the antenna has a broad optimum affected very little by the reflector parameters. This antenna has an improved performance relative to the preceding design (Tables 11-23 and 11-24) because it operates over a 6 : 1 bandwidth instead of 10 : 1.

Increasing α to 30° ($\sigma = 0.0866$) makes the antenna shorter, and its performance peaks for $f/D = 0.42$ for the same 10 : 1 frequency range and $\tau = 0.8$. Table 11-26 lists the parameters of this design. Although this design peaks for $f/D = 0.42$, the losses increase only slowly when we change f/D, similar to the design in Table 11-24. For higher f/D we need narrower beamwidths to reduce spillover, which requires a design with a longer active region. The longer antenna will have a greater phase-center movement over the frequency range. A design that peaks at about $f/D = 0.54$ uses $\tau = 0.84$ and $\sigma = 0.125$ ($\alpha = 17.7^\circ$). Table 11-27 lists the results for this antenna.

A paper by DuHamel and Ore [29] reports the results of measurements of trapezoidal tooth log-periodic feeds for reflectors. These antennas use small scaling constants τ,

TABLE 11-26 Trapezoidal Tooth Log-Periodic Antenna Feed for a Reflector with $f/D = 0.42$

	Beamwidth (deg)		Maximum	Average	Average	Average		Total (dB)	
	10-dB	10-dB	PEL	ATL	SPL	XOL	F/B		
ψ	E-Plane	H-Plane	(dB)	(dB)	(dB)	(dB)	(dB)	Average	Maximum
5	116.1	203.8	0.38	0.44	1.09	0.29	23	1.96	2.16
10	114.4	186.5	0.38	0.49	0.90	0.28	18	1.80	2.04
15	114.9	179.2	0.36	0.52	0.82	0.31	17	1.78	2.01
20	116.1	166.0	0.35	0.56	0.79	0.41	13	1.90	2.15

TABLE 11-27 Trapezoidal Tooth Log-Periodic Antenna Feed for a Reflector with $f/D = 0.54$, 10 : 1 Frequency Range (16 Elements) with $\tau = 0.84$, $\sigma = 0.125$ ($\alpha = 17.7^\circ$)

	Beamwidth (deg)		Maximum	Average	Average	Average		Total (dB)	
	10-dB	10-dB	PEL	ATL	SPL	XOL	F/B		
ψ	E-Plane	H-Plane	(dB)	(dB)	(dB)	(dB)	(dB)	Average	Maximum
5	107.2	158.5	0.56	0.30	1.27	0.21	20	1.97	2.45
10	107.5	139.5	0.54	0.32	1.04	0.33	16	1.87	2.57
15	110.6	118.4	0.53	0.36	0.95	0.60	14	2.09	3.00
20	114.4	97.3	0.49	0.47	1.04	0.95	8	2.64	3.82

TABLE 11-28 Trapezoidal Tooth Log-Periodic Antenna Feed for a Reflector with $f/D = 0.46$

	Beamwidth (deg)		Maximum	Average	Average	Average		Total (dB)	
	10-dB	10-dB	PEL	ATL	SPL	XOL	F/B		
ψ	E-Plane	H-Plane	(dB)	(dB)	(dB)	(dB)	(dB)	Average	Maximum
10	102.5	181.4	0.32	0.49	1.24	0.76	13	2.62	3.64
22.5	105.4	154.7	0.36	0.56	1.17	1.14	8	3.01	4.55

Source: [29].

wide half apex angles α, and widely separated sides ψ to reduce the reflector feed phase error loss due to axial defocusing. The paper reports results over a large range of parameters. Two of the better ones were analyzed as a feed for a reflector with $f/D = 0.46$. These antennas have a scaling constant $\tau = 0.707$ and a half apex angle $\alpha = 30^\circ$ ($\sigma = 0.127$). Table 11-28 lists analysis results of two designs; both produce high cross-polarization loss due to feed line radiation and unequal E- and H-plane beamwidths.

11-15 V LOG-PERIODIC ARRAY [30]

A V log-periodic antenna sweeps the dipole elements of the log-periodic antenna shown in Figure 11-14 forward to form a series of V dipoles. This antenna can be operated

over a large frequency range such that the dipoles operate in higher-order modes. The dipoles operate in $\lambda/2$ mode at the lowest frequencies and are then scattered among a series of transition frequency bands of poor performance. The antenna operates again in $3\lambda/2$, $5\lambda/2$, $7\lambda/2$, and so on, modes of the dipoles. Dipoles, resonant in the higher-order modes, radiate multiple lobe patterns. In a log-periodic antenna these lobes produce high sidelobes. To reduce these sidelobes the elements are swept forward so that the pattern nulls are aligned with the poles that reduce them. The antenna does not need the

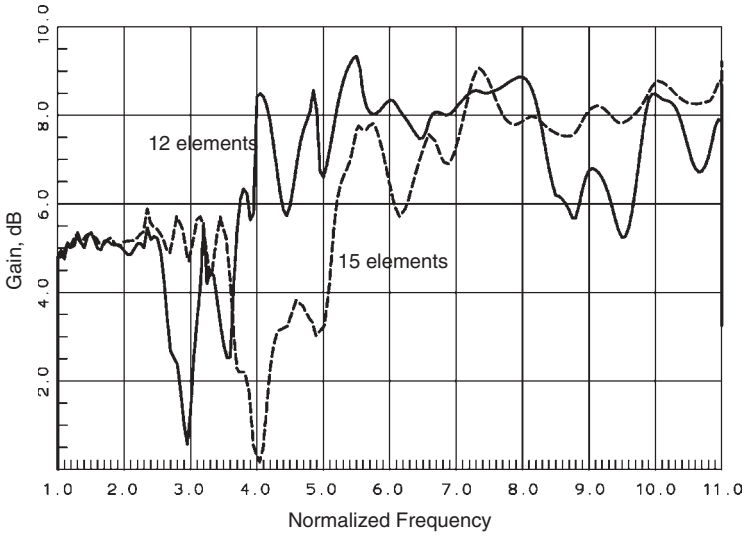


FIGURE 11-23 V-dipole log-periodic antenna gain for $\tau = 0.9$, $\sigma = 0.05$, $\psi = 45°$.

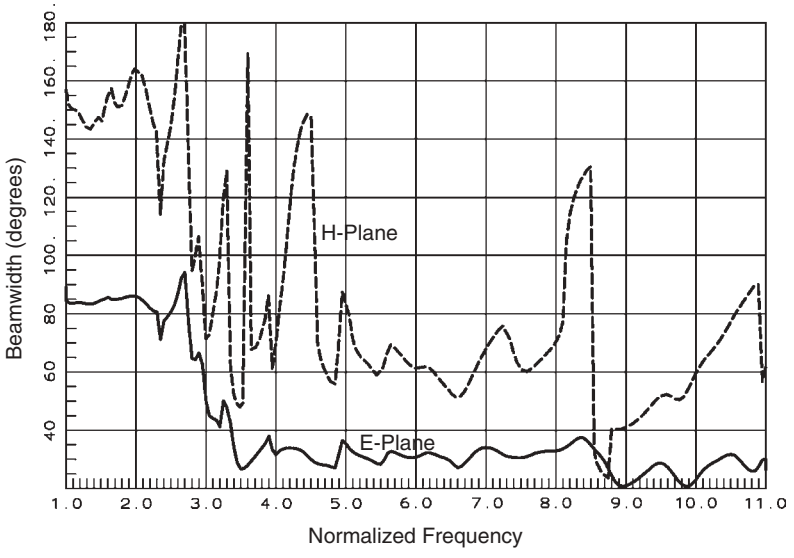


FIGURE 11-24 V-dipole log-periodic antenna beamwidth for $\tau = 0.9$, $\sigma = 0.05$, $\psi = 45°$: 12 elements (solid curve); 15 elements (dashed curve).

small dipoles required for high frequencies in normally designed log-periodic dipole antennas, which eases construction problems.

Designs use scaling constants $\tau \geq 0.9$ and small spacing constants $\sigma \leq 0.05$. Two antennas were designed with 12 and 15 elements using these factors and analyzed. The 12-element antenna requires a dipole length ratio of 3.2 across the elements and has a length of 0.37λ at the lowest frequency; the 15-element antenna dipole length ratio is 4.37, with a boom length 0.42λ at the lowest frequency. Figure 11-23 gives the gain versus normalized frequency found by reducing directivity by the reflected power loss due to impedance mismatch. The poor impedance match reduces gain in the band 2.7 to 3.7 for the 12-element antenna and 3.7 to 5.2 for the 15-element antenna. The return loss is better than -6 dB (3 : 1 VSWR) over the remaining frequency range. Figure 11-24 shows that the E-plane beamwidth decreases drastically when the antenna operates in the higher-order dipole modes. The H-plane pattern exhibits considerable variation and at some frequency ratios has large sidelobes that reduce gain. We cannot use the beamwidths to estimate directivity, but must integrate the entire pattern. The 12-element antenna has its shortest elements 2λ long at the highest frequency.

11-16 CAVITY-BACKED PLANAR LOG-PERIODIC ANTENNAS

We use some of the original planar log-periodic antennas that have wide beamwidth patterns on both sides and eliminate one side by placing it on a cavity in the same manner as a spiral. By using elements morphed into arcs, we interleave two or more antennas to build an antenna to radiate dual linear polarization, or we can combine the two linear polarizations with a hybrid power divider (0° and 90° phases) to obtain RHC and LHC polarization. Similar to a spiral, we can design antennas with multiple spiral modes by interleaving more than two antennas in a single aperture.

Figure 11-25 shows a layout of the interlog antenna [31] that contains two planar log-periodic antennas interleaved. We design such antennas using normal log-periodic antenna design procedures and roll the poles into arcs. We feed the antenna from the center and mount it over an absorber-loaded cavity. The antenna can radiate either dual linear or dual circular polarizations, depending on the feed network. The typical

FIGURE 11-25 Interlog antenna.

predicted pattern of the antenna when fed for RHC polarization is similar to a spiral. The log-periodic nature of the antenna produces a lower amplitude pattern at $\theta = 90^\circ$ than a spiral. Although the antenna has four arms, we cannot form a mode 2 spiral beam because generating circular polarization consumes a degree of freedom. The antenna shown in Figure 11-25 has a 1.54λ circumference at the lowest frequency. Interleaving causes problems because the coupling produces pattern distortion periodic with frequency. The cross-polarization for an antenna fed RHC increases and beamwidth broadens over narrow frequency ranges. We reduce this effect by decreasing the interleaving, but antenna size grows.

The second example of a planar log-periodic antenna over a cavity is a sinuous antenna [32,33]. The triangular tooth antenna has been morphed into arcs using a sinuous curve to reduce reflections from the corners (Figure 11-26). We use the same sinuous curve to define both edges of each arm by rotating it about the center axis. Similar to the interlog antenna, we can feed the four-arm antenna for dual linear or dual circular polarizations. The sinuous antenna radiates a pattern similar to a spiral. To design these antennas we make the sweep of the curved triangles $\lambda/4$, although it is one-half the dipole of a log-periodic antenna. Given the angle of the triangle sweep of the sinuous antenna, $\alpha + \delta$ (rad), we compute its length from the center distance r:

$$\text{length} = r(\alpha + \delta) = \frac{K_1 \lambda}{2} \tag{11-38}$$

We determine the factor K_1 from Eq. (11-28) for the lower truncation constant given τ, the scaling constant. A recommended value for $\alpha + \delta$ is 67.5° for a four-arm antenna that produces a self-complementary structure. We compute the circumference from the lowest frequency and K_1:

$$\text{circumference} = \frac{\pi K_1 \lambda_L}{\alpha + \delta} \text{ (rad)} = \frac{180^\circ K_1 \lambda_L}{\alpha + \delta} \approx \frac{\pi \lambda_L}{2(\alpha + \delta)} \tag{11-39}$$

For $\tau = 0.8$, Eq. (11-28) gives $K_1 = 0.595$. When we use $\alpha + \delta = 67.5^\circ$ for the four-arm antenna and insert these values in Eq. (11-40), we compute circumference $= 1.58\lambda$, a value similar to that for the interlog antenna.

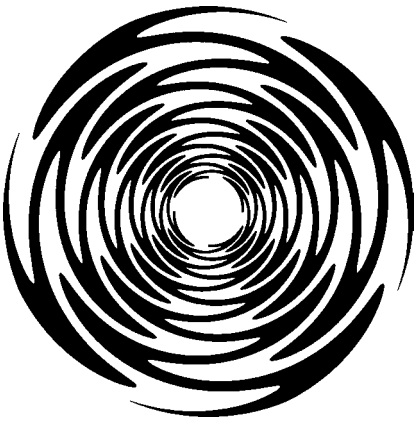


FIGURE 11-26 Sinuous antenna.

When we use a sinuous antenna for dual linear polarization, we discover that the polarization direction oscillates about $\pm 3^\circ$ when we vary frequency. The two directions track each other for both pairs of a four-arm structure, and cross-polarization is reduced when we feed it to radiate circular polarization.

REFERENCES

1. V. H. Rumsey, Frequency independent antennas, *1957 IRE National Convention Record*, pt. 1, pp. 114–118.
2. V. H. Rumsey, *Frequency Independent Antennas*, Academic Press, New York, 1966.
3. R. G. Corzine and J. A. Mosko, *Four-Arm Spiral Antennas*, Artech House, Boston, 1990.
4. E. M. Turner, Spiral slot antenna, *Technical Note WCLR-55-8*, Wright Air Development Center, Dayton, OH, June 1955.
5. J. A. Kaiser, The Archimedean two-wire spiral antenna, *IRE Transactions on Antennas and Propagation*, vol. AP-8, May 1960, pp. 312–323.
6. G. A. Deschamps, Impedance properties of complementary multi-terminal planar structures, *IRE Transactions on Antennas and Propagation*, vol. AP-7, special supplement, December 1959, pp. S371–S378.
7. J. A. Huffman and T. Cencich, Modal impedances of planar, non-complementary, N-fold symmetric antenna structures, *IEEE Antennas and Propagation Magazine*, vol. 47, no. 1, February 2005.
8. K. C. Gupta, R. Garg, and R. Chadha, *Computer Aided Design of Microwave Circuits*, Artech House, Boston, 1981.
9. J. Svacina, A simple quasi-static determination of basic parameters of multilayer microstrip and coplanar waveguide, *IEEE Microwave and Guided Wave Letters*, vol. 2, no. 10, October 1992, pp. 385–387.
10. J. D. Dyson, The equiangular spiral antenna, *IRE Transactions on Antennas and Propagation*, vol. AP-7, no. 2, April 1959, pp. 181–186.
11. R. Bawer and J. J. Wolfe, A printed circuit balun for use with spiral antennas, *IRE Transactions on Microwave Theory and Techniques*, vol. MTT-8, May 1960, pp. 319–325.
12. T. A. Milligan, Parameters of a multiple-arm spiral antenna from single-arm measurements, *IEEE Antennas and Propagation Magazine*, vol. 40, no. 6, December 1998, pp. 65–69.
13. V. A. Monaco and P. Tiberio, Automatic scattering matrix computation of microwave circuits, *Alta Frequenza*, vol. 39, February 1970, pp. 59–64.
14. P. G. Ingerson, Modulated arm width spiral antenna, U.S. patent 3,681,772, August 1, 1972.
15. J. D. Dyson, The unidirectional equiangular spiral antenna, *IRE Transactions on Antennas and Propagation*, vol. AP-7, no. 5, October 1959, pp. 329–334.
16. J. D. Dyson, The characteristics and design of the conical log-spiral antenna, *IEEE Transactions on Antennas and Propagation*, vol. AP-13, no. 4, July 1965, pp. 488–498.
17. C. S. Liang and Y. T. Lo, A multiple-field study for the multiarm log-spiral antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-16, no. 6, November 1968, pp. 656–664.
18. G. A. Deschamps and J. D. Dyson, The logarithmic spiral in a single-aperture multi-mode antenna system, *IEEE Transactions on Antennas and Propagation*, vol. AP-19, no. 1, January 1971, pp. 90–96.
19. A. E. Atia and K. K. Mei, Analysis of multiple-arm conical log-spiral antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-19, no. 3, May 1971, pp. 320–331.

20. D. E. Isbell, Log periodic dipole arrays, *IRE Transactions on Antennas and Propagation*, vol. AP-8, no. 3, May 1960, pp. 260–267.

21. R. L. Carrel, The design of log-periodic dipole antennas, *1961 IRE National Convention Record*, pt. 1, pp. 61–75.

22. C. E. Smith, *Log Periodic Design Handbook*, Smith Electronics, Cleveland, OH, 1966.

23. R. L. Carrel, Analysis and design of the log-periodic dipole antenna, *Antenna Laboratory Technical Report 52*, University of Illinois, Urbana, IL, October 1961.

24. R. H. DuHamel and D. G. Berry, Logarithmically periodic antenna arrays, *1958 IRE National Convention Record*, pt. 1, pp. 161–174.

25. K. G. Balmain and J. N. Nkeng, Asymmetry phenomenon of log-periodic dipole antennas, *IEEE Transactions on Antennas and Propagation*, vol. AP-24, no. 4, July 1976, pp. 402–410.

26. R. H. DuHamel and D. E. Isbell, Broadband logarithmically periodic antenna structures, *1957 IRE National Convention Record*, pt. 1, pp. 119–128.

27. R. H. DuHamel and F. R. Ore, Logarithmically periodic antenna designs, *1958 IRE National Convention Record*, pt. 1, pp. 139–151.

28. W. A. Imbriale, Optimum designs of broad and narrow band parabolic reflector antennas fed with log-periodic dipole arrays, *IEEE AP-S Symposium Digest*, vol. 12, June 1974, pp. 262–265.

29. R. H. DuHamel and F. R. Ore, Log periodic feeds for lens and reflectors, *1959 IRE National Convention Record*, March 1959, pp. 128–137.

30. P. E. Mayes and R. L. Carrel, Logarithmically periodic resonant-V arrays, *IRE WESCON Convention Record*, pt. 1, 1961.

31. D. A. Hofer, O. B. Kesler, and L. L. Loyet, A compact multi-polarized broadband antenna, *IEEE Antennas and Propagation 1990 Symposium Digest*, pp. 522–525.

32. R. H. DuHamel, Dual polarized sinuous antenna, U.S. patent 4,658,262, April 1987.

33. R. H. DuHamel and J. P. Scherer, Frequency-independent antennas, Chapter 14 in R. C. Johnson, ed., *Antenna Handbook*, 3rd ed., McGraw-Hill, New York, 1993.

12

PHASED ARRAYS

Chapter 3 covers the basics of phased arrays. We place phase shifters or time-delay networks on every element or a combination of these in a feed network to shift the beam direction. In this chapter we discuss the particular problems of phased arrays when we consider the effects of phase shifters and architecture of the feed network. In Chapter 4 we discuss array pattern synthesis using direct methods and array characteristics of those that sample aperture distributions. We do not repeat those topics in this chapter.

The pattern properties desired determine the number of array elements and their layout, from which we calculate gain as discussed in Section 3-12. To prevent grating lobes (Sections 3-5 and 3-8), we must place an element approximately every $\lambda/2$ for an array that scans to 90°. Of course, if we limit scan, we may place elements farther apart. By using a uniform linear aperture beamwidth approximation to a line array, its beamwidth is 100 (HPBW factor)/N_x in degrees for the general amplitude distribution when elements are spaced $\lambda/2$. We initially ignore the HPBW factor and assume a uniform distribution (minimum beamwidth). Given the beamwidth in the principal planes, θ_x and θ_y, the number of elements in the array is

$$N = N_x N_y = \frac{10{,}000}{\theta_x \theta_y} \tag{12-1}$$

When we use a tapered distribution to lower sidelobes, the required number of elements increases by the HPBW factors in the two planes. Section 3-12 showed that array gain is limited by its area for closely spaced elements. The area of $\lambda/2$ spaced elements is $\lambda^2/4$, which relates gain to the number of elements by effective area:

$$\text{gain} = \pi N \tag{12-2}$$

This assigns an effective gain of π to each element. When we scan the beam, the effective area (length in the scan plane) drops by cos θ_0, which causes gain to drop

Modern Antenna Design, Second Edition, By Thomas A. Milligan

and beamwidth to broaden:

$$\text{gain} = \pi N \cos\theta_0 \quad \text{and} \quad \theta_x(\text{scanned}) = \frac{\theta_x}{\cos\theta_0} \tag{12-3}$$

Because the gain drops by $\cos\theta_0$ as the beam scans, it leads to the idea of an element pattern:

$$E_e(\theta) = \sqrt{\cos\theta} \tag{12-4}$$

The element pattern alters the array factor pattern as the element is scanned. In general, narrow beamwidth elements greatly alter the pattern when the array factor scans across the element pattern, and it limits the practical scan range. To scan to wide angles, we need broad-beamwidth array elements.

Aperture theory uses a Huygens source along the array with a pattern given by

$$E(\theta) = \cos^2\frac{\theta}{2}$$

a function with values almost identical to those of the ideal element pattern [Eq. (12-4)]. The Huygens source is the combination of electric and magnetic incremental sources whose ratio equals the free-space impedance $\eta = 376.73$. A Huygens source array element impedance matched at broadside has a resistance η. Scanning the array in the E-plane foreshortens the electric field by $\cos\theta_0$, which reduces the resistance by the same factor:

$$\eta_E = \frac{E\cos\theta_0}{H} = \eta\cos\theta_0 \tag{12-5}$$

In the same manner the magnetic field is foreshortened for scan in the H-plane. The scan resistance becomes

$$\eta_H = \frac{E}{H\cos\theta_0} = \frac{\eta}{\cos\theta_0} \tag{12-6}$$

Equations (12-5) and (12-6) predict the general nature of array element impedance when it scans. E-plane scanning reduces input resistance while H-plane scanning increases resistance, a consequence of the infinite current sheet model [1,2]. A finite array has an impedance response generally similar to Eqs. (12-5) and (12-6) but dependent on element position. General scanning has a resistance response relative to the broadside value R_0 given by

$$\eta(\theta,\phi) = \frac{(1-\sin^2\theta\cos^2\phi)R_0}{\cos\theta} \tag{12-7}$$

In Section 3-9 we discuss how to use self- and mutual impedance to find the scan impedance for a finite array, and in Section 11-7 we discuss the interaction of the element impedance with the feed network.

12-1 FIXED PHASE SHIFTERS (PHASERS)

In a phased array every antenna element is connected to a phase shifter or time-delay network. We scan the beam by commanding every network to compensate for the extra time it takes the signal to propagate to a plane normal to the beam direction. Phase shifters operate in terms of a single frequency in a modulo 2π radians fashion

to compensate for the kd_i phase distance to the reference plane for the ith element, where d_i is the distance and $k = 2\pi/\lambda$.

The ideal phased array uses switched time-delay networks because the beam remains at a fixed direction as frequency changes. We make time-delay networks by varying lengths of transmission lines. In large phased arrays that use fixed phase differences instead of time-delay networks, we use time-delay networks near the input to limit the size of the phase-shift-controlled subarrays to increase bandwidth. Switching the length of a transmission line changes the insertion loss, and the switched time delay modulates the element amplitudes. We must add amplitude control in the time delay to compensate for the varying insertion loss to maintain the design aperture distribution. Another approach modulates a laser beam with RF signal, uses low-loss fiber optics to switch time delays, and detects the optical signal to recover the RF before transmission. Fortunately, the losses are so low when switching fiber optics that amplitude control is unnecessary. Unfortunately, the modulator losses must be overcome to make this approach useful.

Some phase shifters (also called *phasers*) are made by switching short transmission-line sections, but the modulo nature of their phase shift limits bandwidth. The most common feed architecture has equal path lengths from the single input to the multiple radiating elements, and the equal path lengths of the corporate feed maintain the arrays broadside pointing when frequency shifts. If we scan a beam by setting phase shifters based on frequency and scan angle, the beam will scan toward broadside as frequency increases. The amount of angular beam shift is independent of the array size, as given by Eq. (3-14). For a planar surface array with its normal along the z-axis with beam scanned to θ_0, measured from the normal vector, changing frequency shifts the beam $\Delta\theta$:

$$\Delta\theta = \frac{f_2 - f_1}{f_2} \tan\left(\frac{\pi}{2} - \theta_0\right) \tag{12-8}$$

where f_1 and f_2 are the two frequencies. Figure 12-1 shows the beam scanning due to a 5% change in frequency. Figure 12-1 was drawn for a particular array, but the

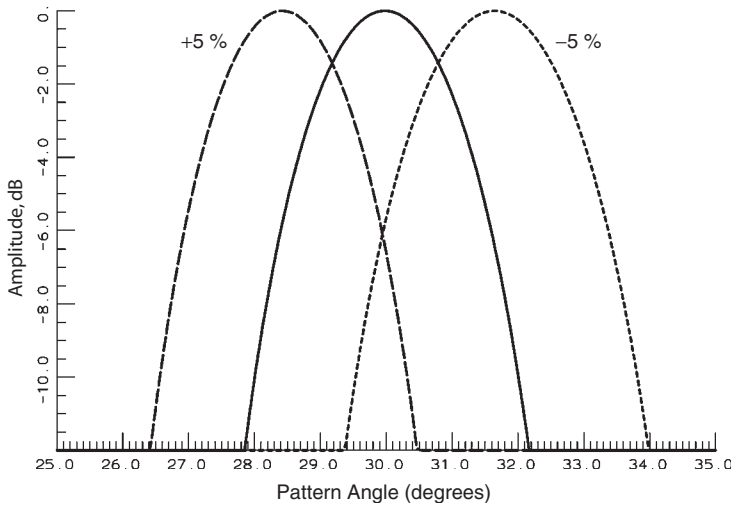


FIGURE 12-1 Scan of array with fixed phase shifters for 5% frequency shift scanned to 30°.

beam shift is independent of array size because array length affects only beamwidth. Larger arrays shift their beams the same amount as do short arrays. Only when we define bandwidth in terms of a factional part of the beamwidth do shorter arrays have greater bandwidth.

We can use frequency shift to scan a linear array beam direction. The array must be series-fed similar to waveguide slots, where a single line feeds the array by having each element coupling a portion of the remaining power traveling down the transmission line. Figure 12-2 shows the nature of the series feed in terms of slot radiators that can be replaced by any antenna element. We load the transmission line after the last element. We use this feeding arrangement for waveguide-fed slot arrays, except that we use straight sections of waveguide to reduce beam scanning as frequency changes. In this application we add a meander to increase the electrical length between elements. Given element-to-element spacing d and a length along the meander s, in general made from a waveguide with relative propagation constant P, we find the phasing equation of the radiated wave:

$$kd\cos\theta + 2\pi N = Pks + (-\pi) \quad \text{for} \quad P = \frac{\lambda}{\lambda_g} = \sqrt{1 - \left(\frac{\lambda}{\lambda_c}\right)^2} \tag{12-9}$$

The $-\pi$ term occurs when we have alternating phases on the elements, such as alternating slots. For coax line we use the dielectric constant for $P = \sqrt{\varepsilon_r}$. For a given N we solve for the meander length s:

$$\frac{s}{\lambda} = \frac{1}{P}\left[\frac{d}{\lambda}\cos\theta_M + \left(N + \frac{1}{2}\right)\right] \tag{12-10}$$

The $\frac{1}{2}$ term is included for alternating phases on the elements. Larger values of N increase the beam direction frequency slope:

$$f\frac{d\theta}{df} = \begin{cases} -\left[\dfrac{s}{d}\left(\dfrac{\lambda}{\lambda_c}\right)^2\dfrac{1}{P} + \dfrac{\left(N+\frac{1}{2}\right)\lambda}{d}\right]\dfrac{1}{\sin\theta}(\text{rad}) & \text{waveguide} \\ -\left[\dfrac{s}{d} + \dfrac{\left(N+\frac{1}{2}\right)\lambda}{d}\right]\dfrac{1}{\sin\theta}(\text{rad}) & \text{coax} \end{cases} \tag{12-11}$$

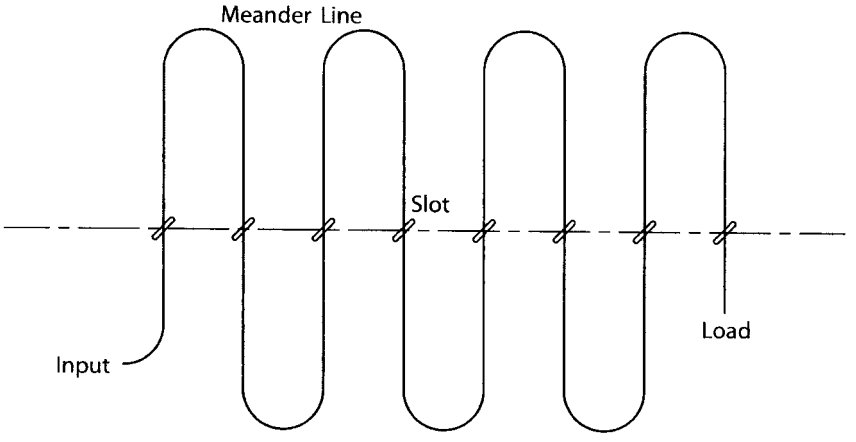


FIGURE 12-2 Frequency scanning array using the meander line between antennas.

Example For a coaxial feeder using $N = 8$ and a spacing $D/\lambda = 0.6$, find the phase slope when $\theta_M = 90^\circ$ and the elements do not have alternating phases.

We substitute these values into Eq. (12-11*b*):

$$\Delta\theta = -\frac{\Delta f}{f}\frac{8}{0.6} = -13.33\frac{\Delta f}{f}(\text{rad}) = -764\frac{\Delta f}{f}(\text{deg})$$

To shift the beam 30°, we change frequency by 3.93%.

Phasers have been designed using high-speed diode switches and lower-speed ferrite devices. The diode (or FET) switches enable rapid scan but have high loss compared to ferrite phasers. The switches redirect the RF path through different lengths of transmission lines and lead to digital phasers. A digital phase shifter has quantized values: 180°, 90°, 45°, 22.5°, and so on, at a particular frequency. Each phaser bit directs the signal through one of two paths, each containing two diode switches, whose nearly equal losses reduce amplitude modulation. But we must design carefully for impedance match in both states to prevent amplitude modulation. Each bit adds significant transmission loss due to the diode switches, and unfortunately, all bits are in series. Adding bits gives better control but increases loss.

The design of transmit/receive (T/R) modules eliminates phaser loss system impacts. If the antenna both transmits and receives signals, such as radar, we add two paths through the module and rely on the nonreciprocal nature of amplifiers to separate the channels. Of course, if the antenna does only one function, we eliminate its path. When transmitting, a final amplifier produces almost all the power when its output is connected to the antenna. The phaser operates on the drive signal and the amplifier overcomes the phaser loss. In the receive channel the signal first passes through a low-noise amplifier (LNA) connected to each antenna element. Section 1-15 shows that a LNA with sufficient gain overcomes the loss of networks with high loss (and noise figure) that follow in the receiver chain. The amplifier overcomes the phaser loss that follows it.

Ferrite phase shifters operate along a magnetic hysteresis loop and it is latched at a zero magnetic drive point for a given phase shift. This requires drive magnetic fields to cycle it along the loop to produce the desired value, but operate only when changing phase. For arbitrary phase shift the ferrite must be demagnetized before changing to a new state. An alternative design digitizes the ferrite cores by using shorter cores and pulse magnetic fields developed by current pulses through a wire that threads through the center of the core to latch it at the two opposite points along the hysteresis loop. We mount the ferrite toroidal cores in a waveguide and vary the lengths to alter the phase shift. The phase shift depends on the direction of the wave through the ferrite phaser. It is nonreciprocal, and in radar all phasers must be reset between transmittal and receipt of pulses. Reciprocal Faraday rotation phasers convert the input signal to circular polarization in the waveguide by using a quarter-wave plate tilted at 45° along the axis of a circular waveguide relative to the input linear polarization. The input wave divides into two signals at the quarter-wave plate, which retards or advances the two signals relative to a waveguide without the plate until their phase difference is 90°. The wave exits as circular polarized. The magnetized ferrite alters the phase of the CP wave independent of direction. On the output another quarter-wave plate converts the wave to a linearly polarized wave. Although ferrite phasers are analog devices

capable of any phase shift, the drive currents are controlled by digital commands and our designs end up having quantized values.

12-2 QUANTIZATION LOBES

Quantization of the phase shifters produces extra sidelobes for scans near broadside. We apply a linear-phase slope across the face to scan an array, and the distance between phase changes is large for scans near broadside because the phase slope is shallow. These widely spaced phase changes produce grating lobes. Given the scan angle θ_0 for an array with M bits in the phase shifters, the quantization lobes are evenly spaced in sin θ-space:

$$\theta_q = \sin^{-1}[\sin\theta_0(1 \pm N \cdot 2^M)] \tag{12-12}$$

The quantization lobes occur at every integer N until θ_q leaves visible space. Figure 12-3 gives the pattern of a 128-element linear array with amplitudes from sampling a 30-dB Taylor distribution scanned to 1° and illustrating quantization lobes. The inner two quantization lobes can be found from Eq. (12-12) for $N = 1$:

$$\theta_q(-1) = \sin^{-1}[\sin(1^\circ)(1-8)] = -7.02^\circ$$
$$\theta_q(1) = \sin^{-1}[\sin(1^\circ)(1+8)] = 9.04^\circ$$

Figure 12-4 shows how the quantization lobes appear at greater angles as the array scan increases, but the sidelobe levels are approximately the same in Figures 12-3 and 12-4. Figure 12-5 illustrates the effect of adding another bit to the phase shifters. The quantization lobes move out in angle and reduce in level for additional bits in the phase shifters. These sidelobes have predictable levels. We use the peak sawtooth phase error due to quantization $\beta = \pi/2^M$, and the lobe voltage amplitude is given by

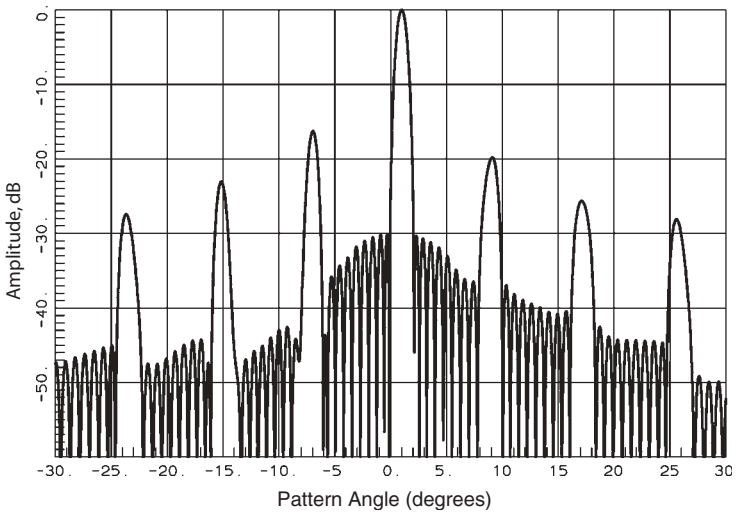


FIGURE 12-3 Quantization lobes of a 128-element line array (30-dB Taylor distribution) with 3-bit phase shifters scanned to 1°.

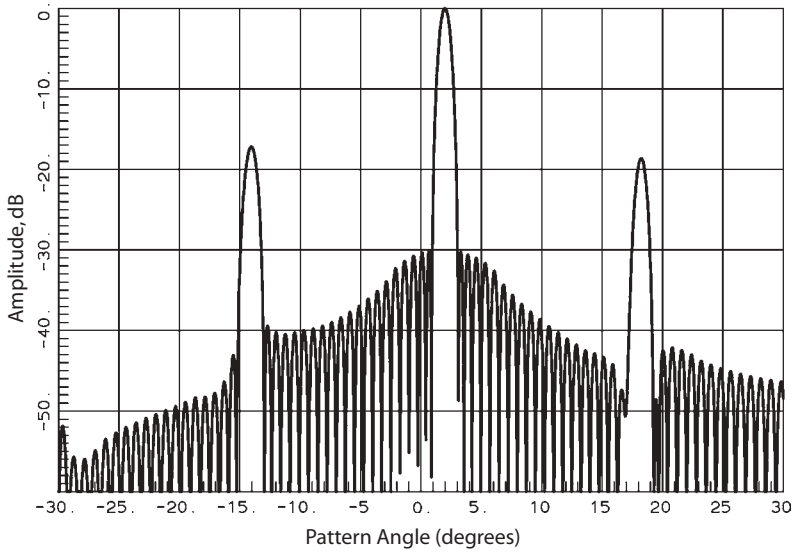


FIGURE 12-4 Quantization lobes of a 128-element line array (30-dB Taylor distribution) with 3-bit phase shifters scanned to 2°.

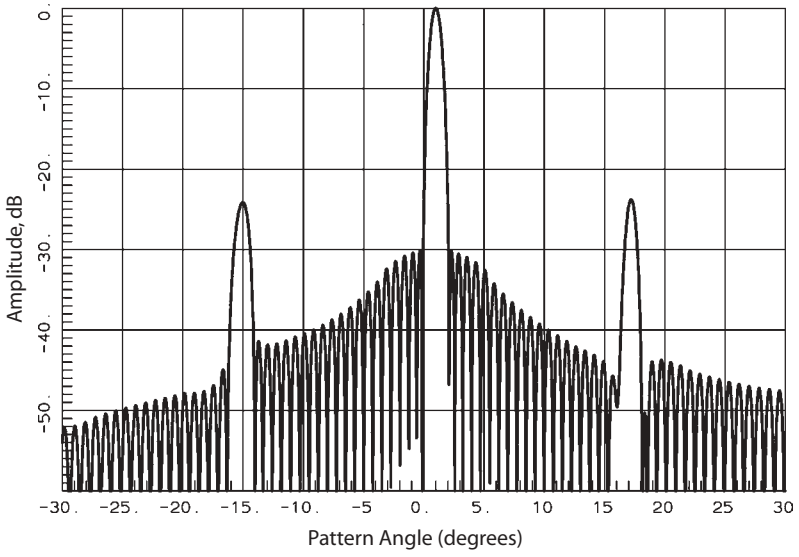


FIGURE 12-5 Quantization lobes of a 128-element line array (30-dB Taylor distribution) with 4-bit phase shifters scanned to 1°.

$\text{sinc}(\beta \pm N\pi)$ for a uniform amplitude array. $N = 0$ is the main beam, which gives the quantization loss due to phase shifters. Table 12-1 lists the main-beam loss and quantization lobe (QL) amplitudes versus the number of bits M in the phase shifters for a positive scan. The negative QL appear on the opposite of the scan. There are no QL values for a scan of zero. We can solve Eq. (12-12) for the scan angle that places

TABLE 12-1 Quantization Sidelobe Levels and Scan of Last QL for a Phased Array

Phase Bits, M	Main-Beam Loss (dB)	−2 QL (dB)	−1 QL (dB)	+1 QL (dB)	+2 QL (dB)	Scan Angle of Last QL	Scan Increment Ratio of BW
2	0.91	17.8	10.4	14.9	20.0	19.47	0.243
3	0.22	23.7	17.1	19.3	24.8	8.21	0.121
4	0.056	29.9	23.6	24.7	30.4	3.82	0.0607
5	0.014	36.0	29.8	30.4	36.3	1.85	0.0304
6	0.003	48.1	42.1	42.2	42.2	0.91	0.0152

the first quantization lobe at -90° and eliminates them for all higher scan angles:

$$\theta_0 = \sin^{-1} \frac{1}{2^M - 1} \tag{12-13}$$

Table 12-1 lists the scan angle that places the first quantization lobe at -90°. The number of phase bits limits the scan increments given as a ratio of the 3-dB beamwidth θ_3 [3]. Table 12-1 lists this ratio:

$$\frac{\theta_{\text{incr}}}{\theta_3} = \frac{1}{2^M(1.029)} \tag{12-14}$$

12-3 ARRAY ERRORS [4; 5; 6, pp. 393–399]

Phase arrays contain random errors that lower gain and increase sidelobe levels. We assume a Gaussian distribution for random amplitude and phase errors with zero mean and variances of $\overline{\Delta}^2$ (volts)^2 for amplitude and $\overline{\delta}^2$(radians)^2 for phase. Although quantization of phase shifters is a systematic error, we can assign a variance to it. The peak error of a quantized phase shifter is $\pi/2^M$ (radians). We calculate the equivalent Gaussian distribution variance of this triangular error distribution:

$$\delta_Q^2 = \frac{\pi^2}{3(2^{2M})} \qquad \text{rad}^2 \tag{12-15}$$

We add this variance to the phase term. Array element failure lowers gain and raises sidelobes. P_e is the probability of survival for each element in the array. The advantage of an array is the graceful degradation of the pattern as elements fail and no single failure shuts down the system. The gain reduction is given by

$$\frac{G}{G_0} = \frac{P_e}{1 + \overline{\Delta}^2 + \overline{\delta}^2} \tag{12-16}$$

We find the efficiency of an array in the same manner as amplitude taper efficiency of a continuous distribution (Section 4-1) by replacing the integrals with summations and using constant phase for the voltage (current) amplitudes a_i of the elements:

$$\eta_a = \frac{\left(\sum_{i=1}^{N} |a_i|\right)^2}{N \sum_{i=1}^{N} |a_i|^2} \tag{12-17}$$

The average sidelobe level σ_s^2 (volts)^2 is related to the beam peak provided that we consider the far sidelobes:

$$\sigma_s^2 = \frac{(1 - P_e) + \overline{\Delta}^2 + \overline{\delta}^2}{P_e N \eta_a} \tag{12-18}$$

Equation (12-18) shows that to achieve low sidelobes, we must control the amplitude and phase errors of elements in an array and that increasing the number of elements decreases the requirements on feeding errors. Sidelobes follow a Rayleigh distribution, and the probability that a sidelobe exceeds a level v_0^2 (volts)^2 is given by

$$P(v > v_0) = e^{-v_0^2/\sigma_s^2} \tag{12-19}$$

We solve Eq. (12-18) for requirements on arrays to achieve average sidelobes:

$$\text{phase error variance } \overline{\delta}^2 = \sigma_s^2 N \eta_a - (1 - P_e + \overline{\Delta}^2)$$

$$\text{amplitude error variance } \overline{\Delta}^2 = \sigma_s^2 P_e N \eta_a - \overline{\delta}^2 - 1 + P_e$$

$$\text{element survivability } P_e = \frac{1 + \overline{\Delta}^2 + \overline{\delta}^2}{\sigma_s^2 N \eta_a + 1}$$

$$\text{number of elements } N = \frac{1 - P_e + \overline{\Delta}^2 + \overline{\delta}^2}{P_e \sigma_s^2 \eta_a}$$

If we substitute Eq. (12-14) into Eq. (12-18), we find the average sidelobe level due to phase quantization:

$$\text{SL(dB)} = 10 \log \left(\frac{\pi^2}{3(2^{2M})} \right) - 10 \log(N) - 10 \log(\eta_a) \tag{12-20}$$

Table 12-2 gives the first term of Eq. (12-20).

Example Given a 64-element array with amplitudes found by sampling a 30-dB Taylor distribution, find the RMS sidelobes versus phase quantization.

TABLE 12-2 Average Sidelobe Factor Due to Phase Quantization Normalized to the Number of Elements and Array Efficiency

Number of Bits	Sidelobe Factor (dB)
1	−0.9
2	−6.9
3	−12.9
4	−18.9
5	−24.9
6	−31.0
7	−37.0

The efficiency of the Taylor distribution is 0.66 dB, from Table 4-5. The element number factor 10 log(N) is 18.06 dB.

$$\text{SL(dB)} = \begin{cases} -12.9 - 18.06 + 0.66 = -30.3 & \text{for 3 bits} \\ -18.9 - 18.06 + 0.66 = -36.3 & \text{for 4 bits} \end{cases}$$

Of course, the distribution will determine the sidelobe levels in the pattern. With only 3 bits, we can expect the peak sidelobes to rise above 30 dB. Quantization lobes exceed these levels for scan angles close to broadside. This example and Eq. (12-20) show that the sidelobe level is proportional to 10 log(N).

12-4 NONUNIFORM AND RANDOM ELEMENT EXISTENCE ARRAYS [4]

12-4.1 Linear Space Tapered Array

Array synthesis discussed in Chapter 4 uses uniformly spaced elements. For arrays along a line, we can analyze them by the Schelkunoff unit circle. These arrays control the sidelobe level by tapering element amplitudes: high in the center and tapered at both ends. We can design planar arrays by sampling aperture distributions, such as the circular Taylor distribution for arrays confined to circles or hexagons. These taper element amplitudes as well. An alternative approach is to taper the spacing between elements fed with uniform amplitude so that the average follows an aperture distribution. When we sample an aperture distribution, we have each element sample a length along the distribution equal to the element spacing. We either find the element amplitude by point sampling the distribution at the element location or we integrate the aperture distribution plus and minus the half-element spacing. Integrating along the distribution produces a better match between the pattern of array and the aperture distribution, but for large arrays we notice little difference between methods. For a nonuniformly spaced array we assign half spacing on each side of the element for integration along the sampled aperture distribution. Because we feed all elements with the same amplitude, we space elements so that the integral of the aperture distribution over each spacing interval is the same for each element.

Figure 12-6*a* illustrates sampling the aperture distribution with an array so that the area under the region assigned to each element is the same. To find element locations, we start by generating a table of the cumulative distribution. We calculate this by integrating the normalized length distribution from the start at -0.5 to a given point x:

$$I(x) = \int_{-0.5}^{x} E(x)\,dx \tag{12-21}$$

We divide this cumulative distribution into equal-length intervals along the ordinate as shown in Figure 12-6*c*, equal to the number of elements in the array. We locate the position of the element by projecting the center of each interval to the abscissa. Instead of using a graphical technique, we calculate a cubic spline with $I(x)$ as the independent variable normalized to 1 and x as the dependent variable. We evaluate the spline at $(i - 0.5)/N$ to find element location x_i. So far we have only found the relative element location normalized to an array length of 1. We have two choices. We can choose the spacing of the closest elements because our planned elements have

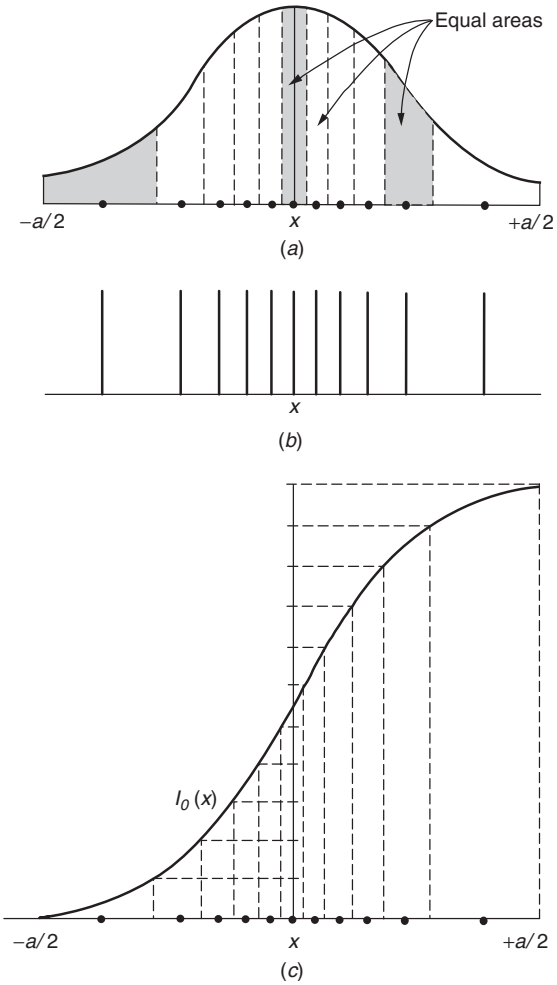


FIGURE 12-6 Deterministic density taper for a line array: (*a*) amplitude taper divided in equal areas; (*b*) location of elements; (*c*) cumulative distribution equally divided along ordinate. (From [4], Fig. 6.2, © 1969 McGraw-Hill.)

a given size and we need to limit mutual coupling. This scales the location of all elements and determines the effective array length. In the second choice we design to a given beamwidth that determines the total array length by considering the half-power beamwidth factor of the distribution sampled. In this case we scale element locations by this length.

A 32-element space tapered array was designed for a 30-dB Taylor linear distribution. Figure 12-7 gives the relative element locations along the line, and Table 12-3 lists element locations when the center element spacing is 0.45λ. Figure 12-8 shows the calculated pattern of this array and another designed with 24 elements, both using elements with 90° beamwidths. The 24-element array has a larger beamwidth because it is shorter. These two patterns show that increasing the number of elements improves the ability of the array to realize the aperture distribution pattern. With so

FIGURE 12-7 Locations of a 32-element space tapered line array for 30-dB Taylor distribution.

TABLE 12-3 Space-Tapered 32-Element Array for 30-dB Taylor Linear Distribution

Element	Position	Element	Position
32, 1	±10.288	24, 9	±3.552
31, 2	±8.798	23, 10	±3.036
30, 3	±7.649	22, 11	±2.539
29, 4	±6.739	21, 12	±2.059
28, 5	±5.971	20, 13	±1.590
27, 6	±5.291	19, 14	±1.130
26, 7	±4.671	18, 15	±0.676
25, 8	±4.095	17, 16	±0.225

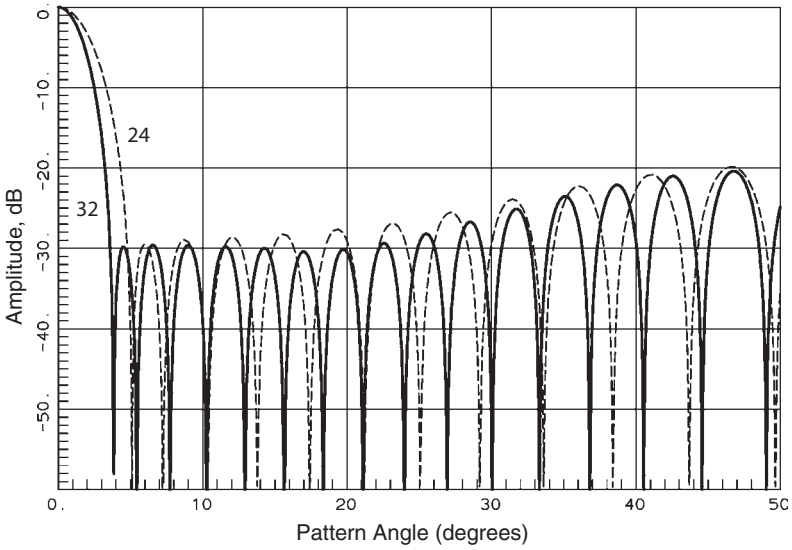


FIGURE 12-8 Pattern of space tapered arrays using 90° beamwidth elements: 32 elements (solid); 24 elements (dashed).

few elements the far sidelobes rise because we narrow the beamwidth relative to a uniformly spaced array that does not radiate grating lobes. The nonuniformly spaced array will not produce lobes at the same amplitude of the main beam, grating lobes of the uniformly spaced array, but spreads the power in these lobes among all far sidelobes.

12-4.2 Circular Space Tapered Array

Space tapered arrays can be designed for circular aperture distributions by evenly spacing array elements in rings [7]. We start by designing a circular aperture distribution,

such as the Taylor distribution (Sections 4-18 and 4-19). We choose a distribution for the spacing of elements in the rings as a function of radius, $d_e(r)$: generally wider near the edge, to match the circular aperture distribution. A table of the cumulative distribution is found by integrating the voltage distribution times the element spacing function $d_e(r)$ from zero to a given radius r:

$$I(r) = \int_0^r d_e(\rho) E(\rho)\, d\rho \tag{12-22}$$

Figure 12-9 shows the division of the cumulative distribution into 10 rings for a 30-dB circular Taylor distribution where the element spacing $d_e(r)$ is a linear function from 0.66λ to 1.4λ from the center to the edge. A ring is located in the center of each region along the radius where the maximum aperture radius is 8.5λ. We space the elements evenly around each ring of radius r_i:

$$N_i = \frac{2\pi r_i}{d_e(r_i)} \tag{12-23}$$

Because N_i must be integers, the radial element spacing only approximately satisfies the $d_e(r)$ distribution when we design small arrays. Table 12-4 lists the parameters of a 10-ring design.

The diagram of the element positions (Figure 12-10) shows the increasing element spacing as radius increases. The combination of Figures 12-9 and 12-10 illustrates the nearly even ring spacing, due to a design with increasing element spacing with radius. Figure 12-11 plots the pattern response and shows the nearly equal sidelobe response for the first few sidelobes. Adding rings improves the match between the aperture distribution and the space tapered array radiation patterns.

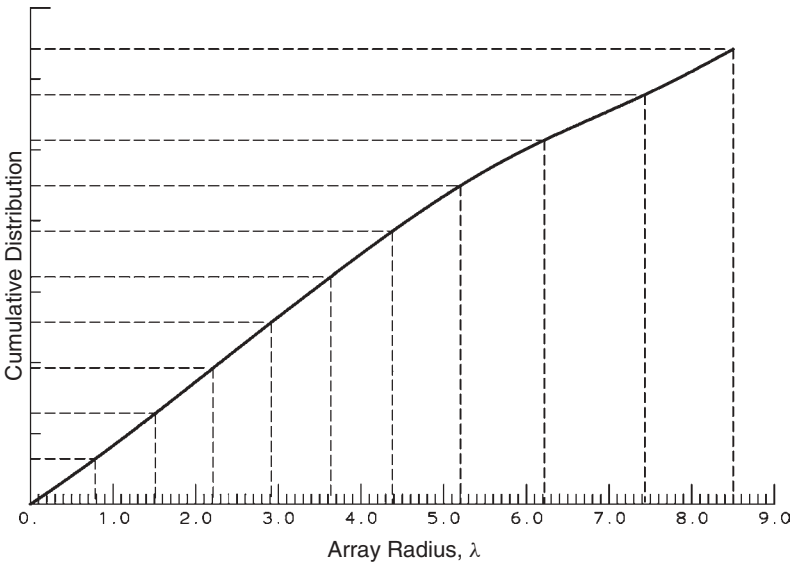


FIGURE 12-9 Cumulative aperture distribution of a 10-ring space tapered (30-dB Taylor distribution) ring array with linear taper on the element spacing from 0.66λ to 1.4λ.

TABLE 12-4 Design of a 10-Ring (223 Elements) Space Tapered Array for 30-dB Circular Taylor Distribution with Linearly Tapered Radial Spacing from 0.66λ to 1.4λ

Ring	Radius	N_i	d_e	Ring	Radius	N_i	d_e
1	0.403	4	0.695	6	4.004	25	1.009
2	1.155	10	0.761	7	4.779	28	1.076
3	1.864	14	0.822	8	5.677	31	1.154
4	2.562	18	0.883	9	6.826	34	1.254
5	3.272	22	0.945	10	7.989	37	1.356

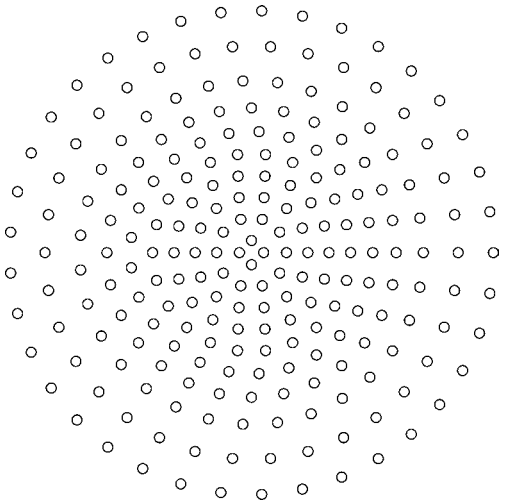


FIGURE 12-10 Layout of a 223-element ring array for 30-dB circular Taylor distribution with 10 rings.

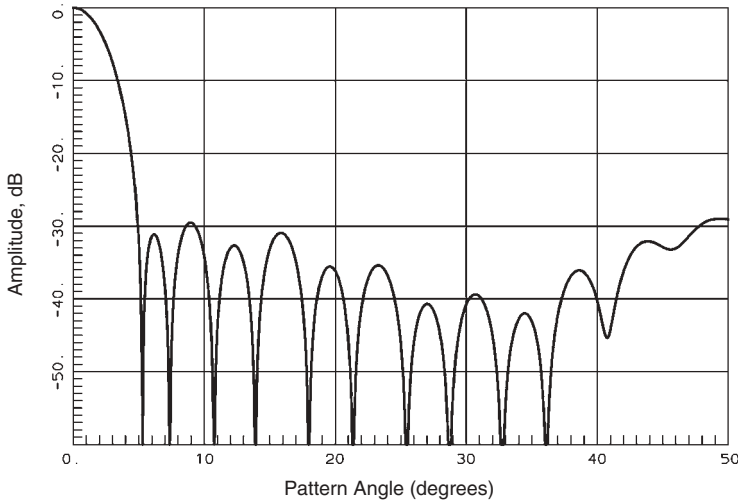


FIGURE 12-11 Pattern of a 223-element space tapered ring array with 90° beamwidth elements.

12-4.3 Statistically Thinned Array [8]

Low sidelobe patterns can be obtained from uniform-amplitude fed arrays by thinning a uniformly spaced array statistically. Each element has a random chance of existence, determined by using the aperture distribution as the probability density function. In this method we lay out a uniformly spaced array and use a random number generator output times the desired aperture distribution to determine whether to include each element. We multiply the product of the random number [0,1] by the aperture distribution [0,1] and by a scale factor k to scale to the level of thinning. We call $k = 1$ natural thinning because the process removes about half the elements. When the random number is higher than the product of k times the aperture distribution, the element is removed, and therefore a lower k produces greater thinning.

Thinning reduces gain by the ratio of the remaining elements to the initial total, but array size determines beamwidth. We obtain patterns determined by aperture extent but with reduced gain. The number of elements determines the average sidelobe level at 10 $\log(N)$ dB. Provided that we use sufficient elements, the near-in sidelobes will follow the aperture distribution, but the far-out sidelobes rise to the average level. For initially large arrays, 90% thinning is reasonable. A convenient method in a computer program is to mark elements for removal and report on the remaining element number so that the algorithm can be repeated after adjusting k until the desired number is achieved before actual removal. We obtain a different number of elements with the same k because of the variability of the random number generator.

To show the method we pick a hexagon array with 0.66λ element spacing first confined to a radius $= 8.7\lambda$. A 30-dB circular Taylor distribution used as the probability density and $k = 1$ produced an array with 337 elements starting from an array with 637 elements inside the circle. Figure 12-12 shows the layout of the remaining elements. The low aperture distribution near the rim removed many of the outer elements while

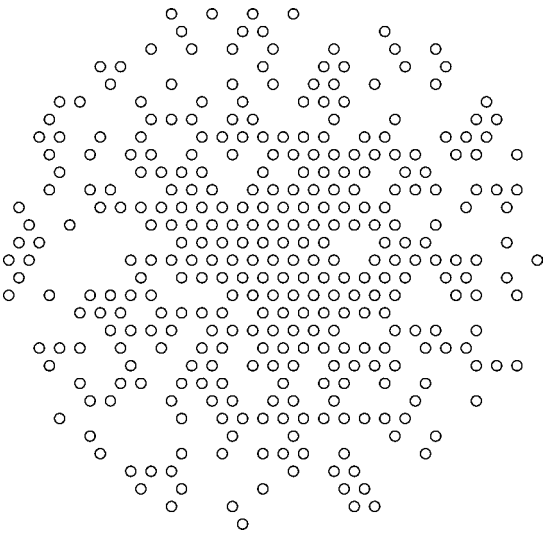


FIGURE 12-12 Layout of a 337-element statistically tapered array for 30-dB circular Taylor distribution.

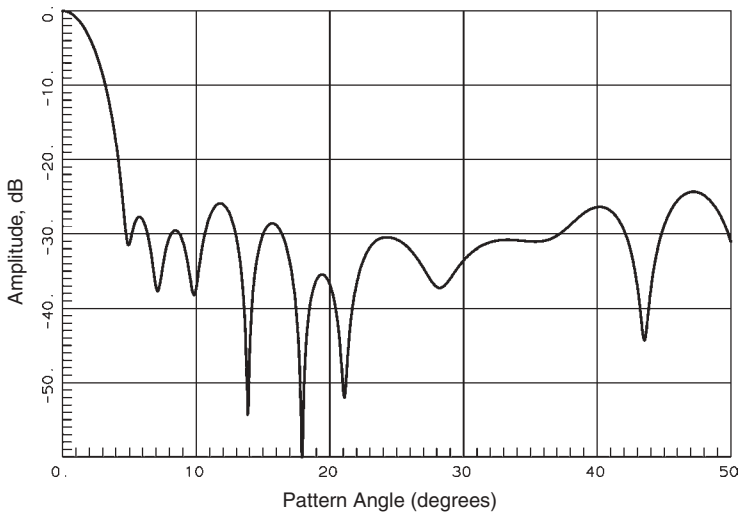


FIGURE 12-13 Pattern of a 337-element statistically tapered array for 30-dB circular Taylor distribution using 90° beamwidth elements.

most of the inner elements remain. Figure 12-13 plots the pattern response of this array when 90° beamwidth elements are used in the array.

12-5 ARRAY ELEMENT PATTERN

When we insert an antenna element into an array, its effective pattern changes. Throughout the book we have shown that the pattern of isolated elements change when we mount them on finite ground planes. An antenna mounted over a ground plane excites currents on it that radiate and contribute to the pattern. We place many antennas near each other in a phased array. Radiation from one antenna excites currents on its neighbors, and the sum of these currents radiates a pattern different from that of the isolated element. If we want to apply pattern multiplication between the array factor and the element pattern, we need to modify the element pattern to include the radiation from the current excited on its neighbors. This is a simplifying assumption because in the final configuration the array will be mounted on a finite ground plane, and not all elements will be surrounded by the same configuration of elements. Each element will have a different pattern, but to first order we use the array element pattern.

We could assume current basis functions on the elements and on the nearby structure and use an integral equation method to solve for the coefficients of these basis functions (moment method). The final pattern calculated would be found from a weighted sum of currents. By increasing the number of basis functions, the calculation converges to the measured result (within measurement error). We simplify the problem by assuming an infinite mounting structure to eliminate edge effects that eliminate the basis functions on the ground plane, and we can use image elements. To find mutual impedance we calculate the reaction integral over the surface of one element, with an assumed current distribution, times the field radiated from the first antenna. The first antenna and its image radiate to the second antenna. Since the reaction integral is over the surface of the second antenna, its image is not included in the integral. The elements along the

edge can have a different pattern than the central elements, and a full moment method solution accounts for differing current distributions. The next approximation assumes that the antenna elements radiate in fundamental modes with fixed current distributions. For example, these could be sinusoidal current distributions on dipoles that reduce the basis functions to one on each antenna.

Because we have a linear problem, we can calculate the coupling (or mutual impedance) between elements using pairs. Assuming only one mode on each antenna reduces the matrix problem to one with rank equal to the number of elements. We multiply the inverse impedance matrix by the excitation voltage vector to find the current on each element (one mode on each). This method becomes intractable for a large array, because we must invert the impedance matrix.

The next level of approximation assumes that all antennas are surrounded by the same element configuration. We assume that the edge elements have the same effective pattern as the inner elements. If we supply power to one element, its radiation is received by all other elements in the array that excites currents on the element simply as a scaled version of the fundamental mode. For this analysis we feed only a single element and calculate the current excited on nearby elements loaded by their characteristic impedance. This suggests an experimental approach of feeding a single element surrounded by other loaded elements. We call this the *array element pattern* or *scan element pattern* (formally the *active element pattern*). Analytically, we calculate the mutual impedance matrix and add twice the element resistance to the diagonal elements to account for the loop of source and load resistance on the antennas. We invert this matrix and find the current on the fed element and on its loaded nearby neighbors. This current sum radiates a pattern that we associate with the fed element and call it the array element pattern. We assume that all elements radiate the same pattern and multiply by the array factor to calculate the array pattern.

Consider an array of V-dipoles mounted on a ground plane. To equalize the E- and H-plane beamwidths, we tilt the elements 30° toward the ground plane. This broadens the E-plane beamwidth by removing the pattern null along the horizontal plane due to the dipole current. We reduce the mutual coupling when we locate a horizontal dipole over a ground plane because its image radiation decreases the radiated field of antenna along the ground plane. Figure 12-14 plots the E-plane pattern of the V-dipole tilted 30° when located in a hexagon array of loaded elements for 2 (19) and 3 (37) rings of elements. The radiation from currents excited on the extra dipoles widens the pattern beamwidth and lowers gain. The H-plane pattern (Figure 12-15) has lower gain than the isolated element, but narrower beamwidth. The increased E-plane beamwidth has decreased gain and the H-plane narrowed beamwidth fails to offset the gain loss significantly. The analysis shows the increased pattern ripple as more elements are added. These plots illustrate results expected from calculations or measurements on an antenna element where the ground plane reduced mutual coupling.

Figure 12-16 gives the pattern of an edge element in the V-dipole array in the E-plane. The analysis placed the elements over an infinite ground plane to eliminate ground-plane effects. We see significant asymmetry in the pattern due to its placement. Of course, we expect a large variation from this small array for the edge element. As we increase array size, these effects diminish. It is reasonable to start with an idealized element pattern for the initial design since array element pattern effects are small for good elements. As the design progresses an array element pattern should be measured (or calculated) and used for analysis at the next level.

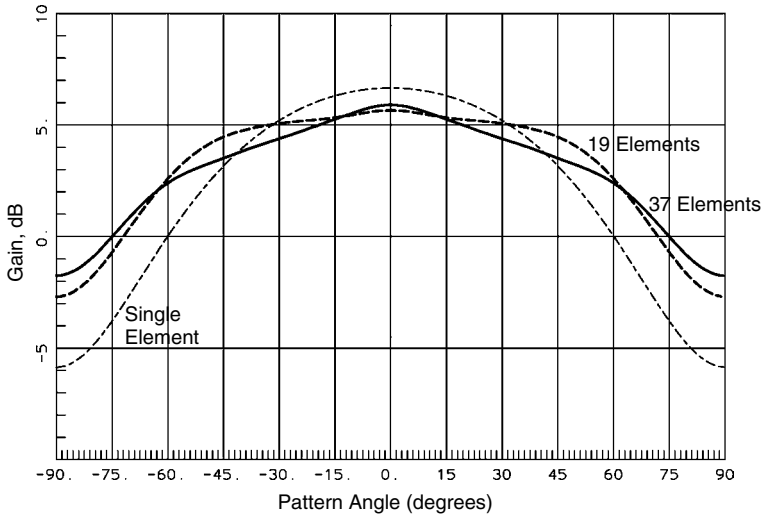


FIGURE 12-14 *E*-plane pattern of a V-dipole tilted 30° located in the center of a hexagon array.

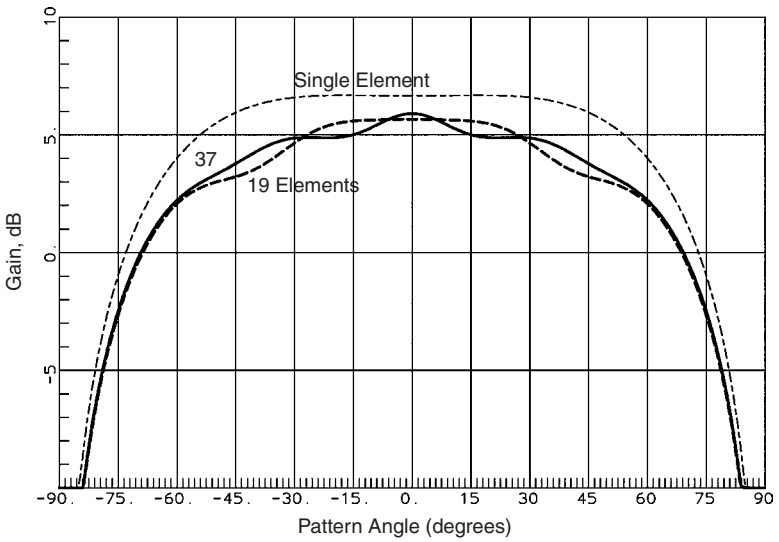


FIGURE 12-15 *H*-plane pattern of a V-dipole tilted 30° located in the center of a hexagon array.

12-6 FEED NETWORKS

12-6.1 Corporate Feed

Figures 12-17 and 12-18 illustrate schematically the two types of constrained feeds. In these networks we route signals through transmission lines and power dividers to deliver power to each element for transmittal and, by reciprocity, route the receive signals to single outputs. A phase shifter is placed between the power division network and every antenna element to scan the beam. The corporate feeds in Figure 12-17

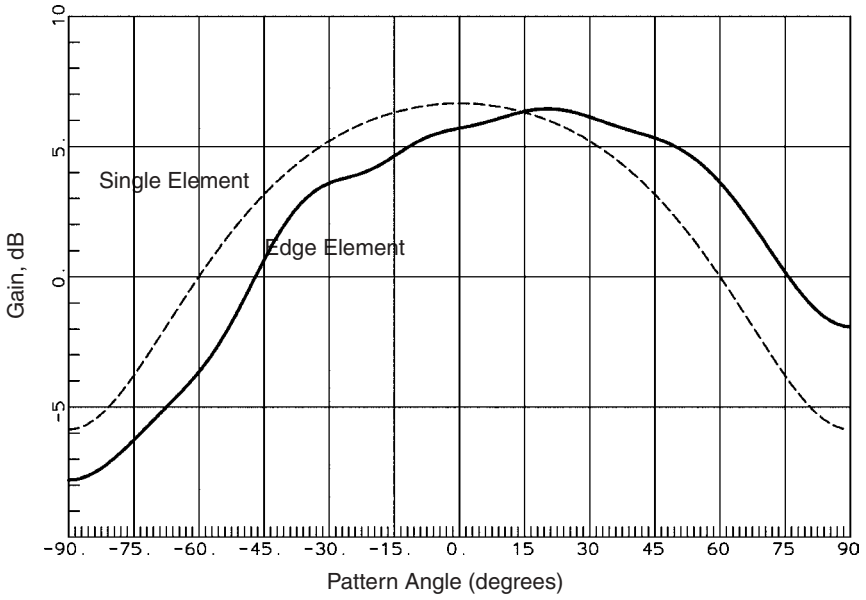


FIGURE 12-16 *E*-plane pattern of a V-dipole tilted 30° located at the edge of a hexagon array.

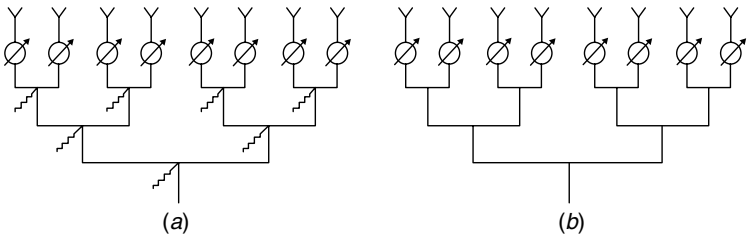


FIGURE 12-17 Corporate feeds: (*a*) matched with isolation resistors; (*b*) reactive. (From T. C. Cheston and J. Frank, Array antennas, Chapter 11 in M. I. Skolnik, ed., *Radar Handbook,* Fig. 38, © 1990 McGraw-Hill.)

usually divide to 2^N elements because two-way power dividers have the simplest designs. The matched corporate feed contains isolation resistors on the power dividers. We fabricate power dividers from four-port microwave circuits where one port is isolated, which means that in normal matched-impedance operation it receives no signal. When the output ports do not have matched loads, a portion of the reflected signal will be dissipated in the load on the isolated port. The impedance of phased array elements change as we scan the beam and it is impossible to match the elements at all scan angles. The isolation resistors prevent propagation of these reflected signals beyond the first network, whereas reactive power divider networks cannot prevent the accumulation of errors as the phase array scans. These networks work best in fixed-beam operation, not in phased arrays. Isolation resistors dissipate no power under impedance-matched conditions. In Section 11-7 we discuss how to analyze the connection between the feed network and the array with its mutual coupling, to calculate the resulting feed coefficients of the array.

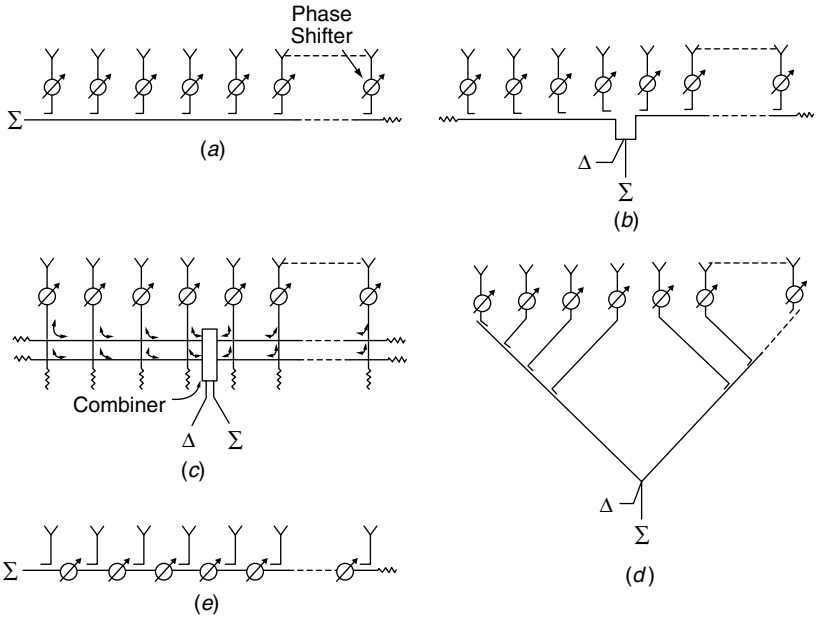


FIGURE 12-18 Series feed networks: (*a*) end feed; (*b*) center feed; (*c*) center feed with separately optimized sum and difference channels; (*d*) equal-path-length feed; (*e*) series phase shifters. (From T. C. Cheston and J. Frank, Array antennas, Chapter 11 in M. I. Skolnik, ed., *Radar Handbook,* Fig. 37, © 1990 McGraw-Hill.)

12-6.2 Series Feed

Figure 12-18 shows types of series feeds. We use series feeds for frequency scanning, but these contain phase shifters that can scan the beam in other directions not determined by frequency scanning. Of course, a series feed has limited bandwidth because the long transmission lines cause frequency scanning. The end feed in Figure 12-18*a* is the same as the waveguide-fed slot array, which has a bandwidth $50\%/N$, where N is the number of elements along the line. Feeding in the center, Figure 12-18*b* doubles the bandwidth and allows both sum and difference feeding used in monopulse. The network of Figure 12-18*c* has separate lines for sum and difference distributions so that each can be optimized. This double line configuration does suffer from coupling between the two rows of couplers because of the limited isolation in the couplers and the changing impedance as the array scans. If we place the phase shifters along the feeder line as shown in Figure 12-18*e*, all can be set to the same value to scan the beam. The problem is that each phase shifter has loss and these accumulate as the signal travels down the feed line. The network in Figure 12-18*a* has the loss of only one phase shifter. By equalizing the path lengths in each arm (Figure 12-18*d*), we increase the frequency bandwidth to that of the phase shifters and array size, because series feed frequency scanning is eliminated.

12-6.3 Variable Power Divider and Phase Shifter

Figure 12-19 gives the schematic of two transmission-line couplers connected in tandem with lines that contain phase shifters. We build 3-dB power divider couplers in stripline made with a centerboard with totally overlapped lines a quarter-wavelength

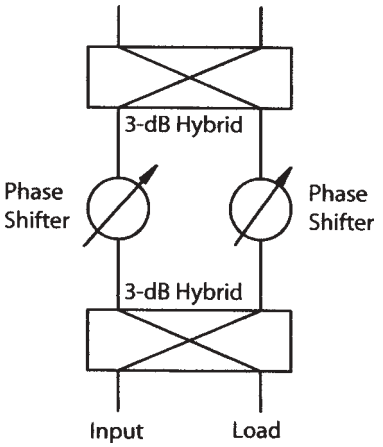


FIGURE 12-19 Variable power divider and phase shifter combination.

long in the dielectric. In a dielectric of 2.21 the centerboard is 0.013 in. (0.33 mm) thick and the outer spacers to the two ground planes are 0.062 in. (1.60 mm) thick. The totally overlapped lines are 0.065 in. (1.65 mm) wide to produce a 3-dB coupler. By crossing over the two coupled lines, the outputs occur on the side opposite the feed.

If the phase shifters have the same value, the network produces a 0-dB coupler where signals at the left input flow to the right output. Of course, the network has symmetry, and a signal at the right input couples totally to the left output. Through the two couplers the signal switches from one side of the centerboard to the other side. This dual coupler can be used to cross signals in a feed network without using the vertical vias that are useful for the construction of Butler matrices and other feed networks where crossovers are needed.

With two phase shifters in the lines between the two couplers, as shown in Figure 12-19, we can build a combination variable power divider and phase shifter. The voltage coupling c_p is determined by the phase difference in the two paths:

$$\phi_2 - \phi_1 = \cos^{-1}(1 - 2c_p^2) \tag{12-24}$$

The output phase of the two arms is given by

$$\text{phase} = \phi_1 - \tan^{-1}\frac{\sin(\phi_2 - \phi_1)}{1 - \cos(\phi_2 - \phi_1)} \tag{12-25}$$

Figure 12-20 shows a feed network using variable power dividers so that the output can be any amplitude and phase distribution. We discussed the formation of multiple beams from a single input in Section 3-6. We can form any combination of beams when we have control of both amplitude and phase. Of course, the array size limits our ability to form distinct beams, but the beamwidth of each beam is determined by the full size of the array. Each power divider is variable. Since the output from a variable power divider with one phase shifter can be any amplitude, we only need the extra phase shifter, either in the last variable power divider or on the output as shown. Because the network provides one input, the gain drops as we form multiple beams. No power is lost in the loads for impedance-matched outputs. Depending on

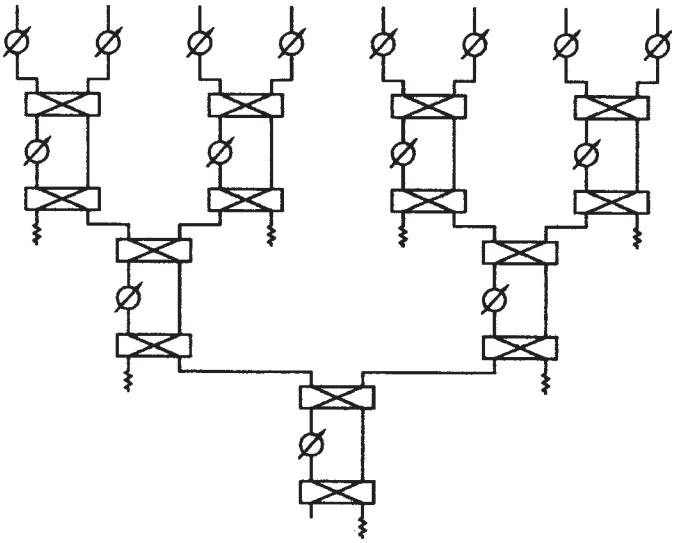


FIGURE 12-20 Eight-element feed network using variable power dividers with final phase shifters at the elements.

the phase of the outputs, the reflected signals from the antennas are either dissipated in the isolation loads or appear at the input.

12-6.4 Butler Matrix [9]

A Butler matrix is a feed network with 2^N inputs and 2^N outputs used to feed a uniformly spaced line array that produces a beam for every input port. The network consists of 3-dB hybrid couplers and fixed phase shifters. Figure 12-21 shows the schematic of the microwave feed network for eight ports. Each input port produces a uniform-amplitude distribution on the outputs with a uniform phase slope across the ports. We can use 0-dB couplers for the internal line crossings because adding couplers to an etching for the centerboard of a stripline has no cost. We make the output line crossings by using equal-length coaxial lines connecting the feed network to the antennas.

Each input port of the Butler matrix has the full gain of the array. On transmit we provide 2^N inputs and insert 2^N times the power of a single input into the feed network that provides full power to each beam. If we scan an array to an angle different from the angle of the input wave, the array will either dissipate the signal in internal loads or scatter it, but it does not collect at the input port. A Butler matrix provides an input port for the sequence of waves incident from 2^N directions. Signals from these directions are not dissipated or scattered, but the array collects them with the full gain of the array.

Figure 12-22 plots the patterns for all eight ports of a Butler matrix feeding an array with $\lambda/2$-element spacing. The array elements have 90° beamwidths and the further scanned patterns show the decrease due to element pattern. The uniform-amplitude arrays radiate 13.2-dB sidelobes, which limits the separation of the signals into the various input ports. The pattern peaks occur at evenly spaced values of u-space, where

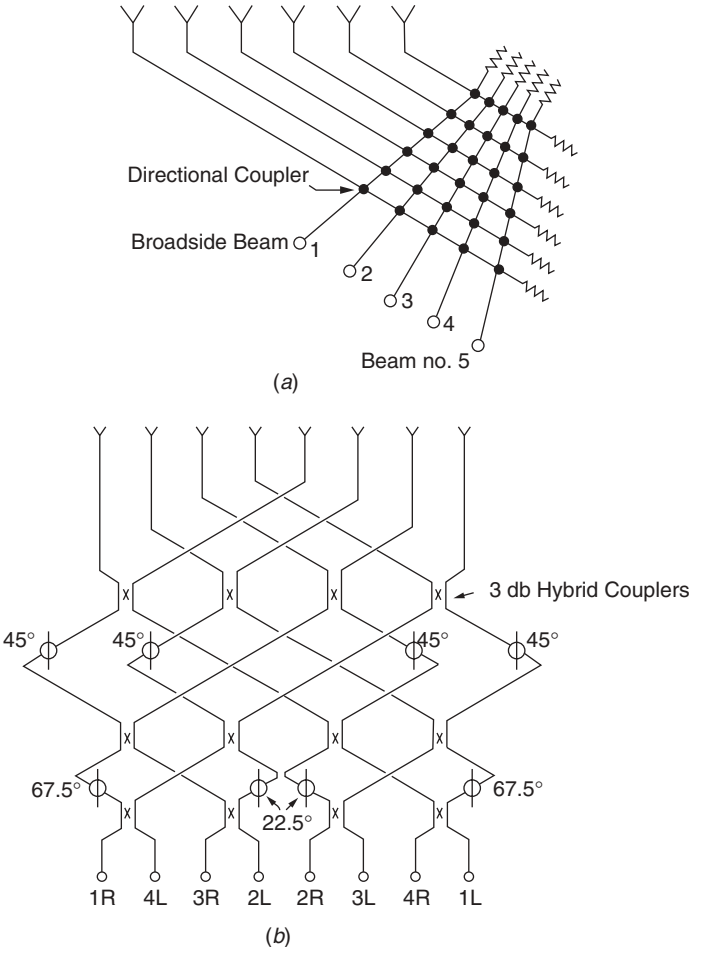


FIGURE 12-21 Eight-element Butler matrix network: (*a*) equal-path-length beam-forming matrix; (*b*) eight-element, eight-beam matrix. (From T. C. Cheston and J. Frank, Array antennas, Chapter 11 in M. I. Skolnik, ed., *Radar Handbook,* Fig. 57, © 1990 McGraw-Hill.)

$u_i = (d/\lambda)(\sin\theta - \sin\theta_i)$ for an element spacing of d:

$$\sin\theta_i = \pm\frac{i\lambda}{2Nd} \qquad i = 1, 3, 5, \ldots, N-1 \tag{12-26}$$

The uniform amplitude line array has a pattern:

$$\frac{E_e(\theta)\sin N\pi u_i}{N\sin\pi u_i}$$

Since the minimum value of the index i is 1, the first beams scan from broadside. The crossover level between beams occurs at -3.92 dB for an isotropic element pattern.

Equation (12-26) shows that the beams move closer together as we increase the electrical distance between elements. Figure 12-23 gives the pattern response of a

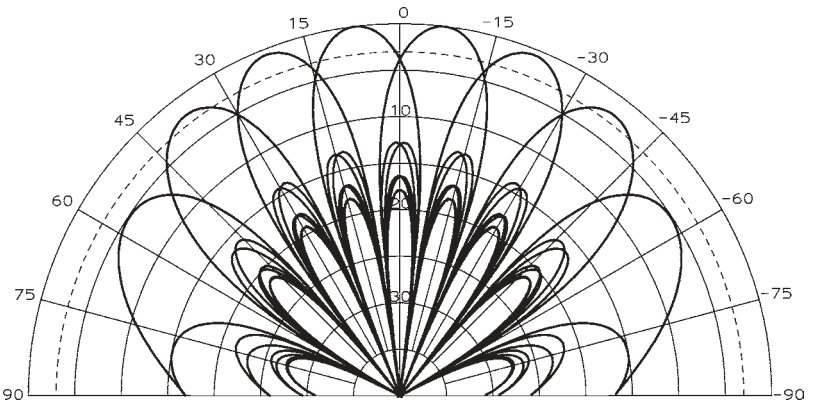


FIGURE 12-22 Patterns of an eight-element line array spaced λ/2 with Butler matrix feed.

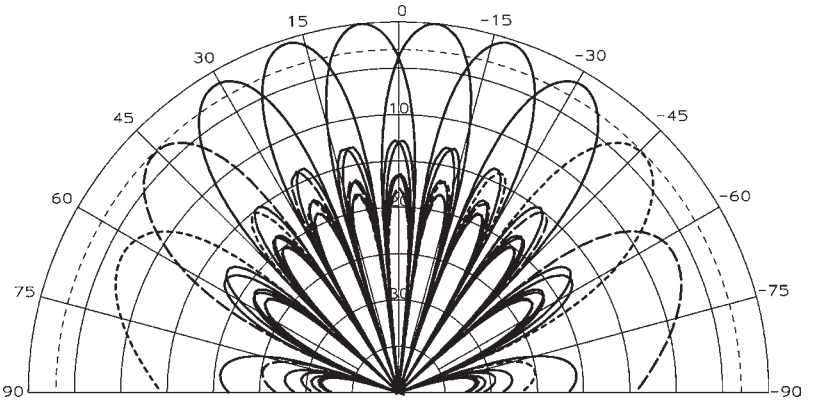


FIGURE 12-23 Patterns of an eight-element line array spaced 0.6λ with Butler matrix feed.

Butler matrix feeding an array with 0.6λ element spacing. All beams shift toward broadside. The crossover level is still at −3.92 dB and the first sidelobe level at 13.2 dB. The two dashed patterns radiate grating lobes reduced in amplitude only by the element pattern. The response seems to have 10 beams. Although the Butler matrix can be designed to cover an octave bandwidth by using quarter-wavelength hybrid couplers and compensated line Schiffman phase shifters [10], an array will not operate over this bandwidth without forming grating lobes. An octave-bandwidth Butler matrix is useful as the beamformer for a multiple-mode spiral antenna (Section 11-5.4).

12-6.5 Space Feeding

Instead of using a feed network with the signals constrained in transmission lines, we can use a feed antenna and an array of receiving antennas to distribute the signal. Figure 12-24 illustrates space feeding where a single feed antenna radiates to the backside of an array, where pickup antennas collect the signal and deliver it to a phase shifter and amplifier chain that feed the radiating antennas of the array. Of course, this feeding method can be used for reception as well. Although Figure 12-24 shows a line array, we actually have a planar array.

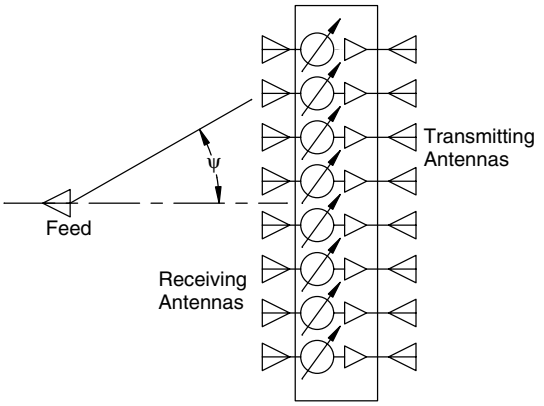


FIGURE 12-24 Space-fed array.

We analyze this feed network as a lens. The phase shifters compensate for the phase shift due to the extra transmission distance and eliminate the quadratic phase error, but the network cannot compensate for the amplitude taper. The feed-antenna-narrowed beamwidth, which matches the angular extent of the plane to reduce spillover loss, reduces amplitude to the edge elements. Similarly, the receiving antennas have a pattern that adds to the amplitude taper. We equate feed power in differential area to aperture power in a differential area [Eq. (9-11)]:

$$\frac{A(r,\phi)}{F(\psi,\phi)} = \frac{\sin\psi}{r}\frac{d\psi}{dr}$$

We add the geometry of the flat-plate receiving surface and carry out the indicated operations to find the distribution on the array:

$$A(r,\phi) = \frac{F(\psi,\phi)R(\psi,\phi)\cos^3\psi}{f^2} \tag{12-27}$$

Equation (12-27) includes a receive antenna power pattern $R(\psi,\phi)$. Of course, the feed arrangement can be offset instead of centrally fed as shown, with phase shifters compensating for the different feed path lengths.

12-6.6 Tapered Feed Network with Uniform-Amplitude Subarrays

One approach to produce a wider bandwidth array is to combine a feed network containing time-delay networks that feed subarrays. We construct the subarrays with identical networks containing phase shifters and a uniform-amplitude distribution. We apply the amplitude taper to the smaller feed array that feeds the subarrays to lower sidelobes. The beam scanned with time-delay networks does not shift when frequency changes, but the phase shifters cause phase stepping across the array face. This causes quantization lobes.

Initially, we consider the array at broadside, where it does not matter whether the input feed network contains time-delay networks or phase shifters. We consider the subarray to be the element pattern of the smaller input feed network. Figure 12-25 shows the pattern at broadside using an eight-element subarray and a 16-element input array that zero-sampled a linear 30-dB Taylor distribution ($\overline{n} = 6$). Because the

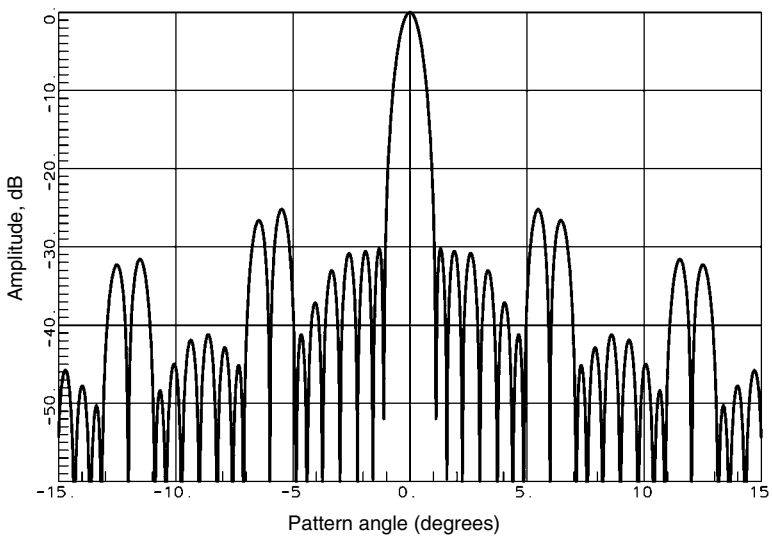


FIGURE 12-25 Paired sidelobes of a 128-element array with a 16-element subarray with an eight-element 30-dB Taylor distribution feed network.

sidelobes of the uniform array do not occur at the same angles as the Taylor distribution, Figure 12-25 contains paired high sidelobes. Decreasing the number of elements in the feed array to eight causes the paired sidelobes to rise above 30 dB because the distribution controls only seven nulls instead of the 15 shown in Figure 12-25.

When we scan the beam, we apply phase shift to both feed and subarray so that the beams remain aligned. This presents no pattern problem if we use continuous phase shifters, but when we quantize the phase shifters, misalignment occurs unless we space elements on $\lambda/2$ centers. Figure 12-26 gives the pattern responses a for 128-element

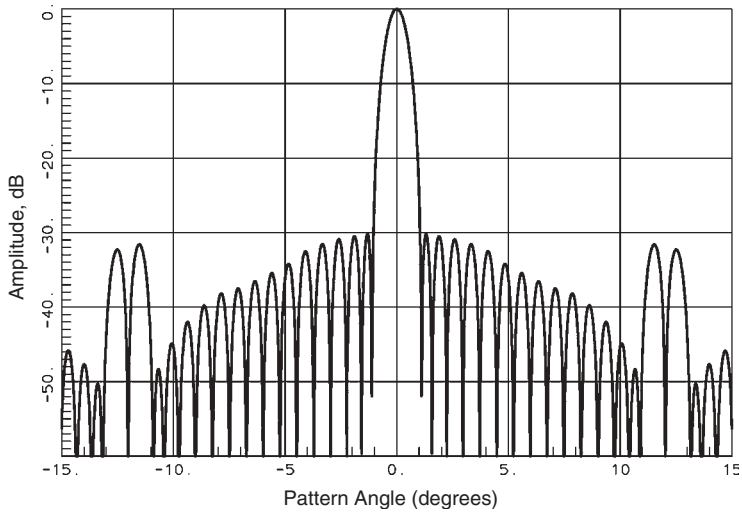


FIGURE 12-26 Paired sidelobes of a 128-element array with an eight-element subarray with a 16-element 30-dB Taylor distribution feed network.

array with an eight-element uniform-amplitude subarray and a 16-element feed array with a 30-dB Taylor ($\overline{n} = 6$) sampled distribution for 0.6λ element spacing. The curve occurs when the feed array has 4-bit phase shifters and the feed array has 3-bit phase shifters. Decreasing the feed array to 3 bits or reducing the number of elements degrades the sidelobes further. Adding time-delay networks does not eliminate these problems because the subarray and feed network sidelobes increase their misalignment when frequency shifts even for elements initially on $\lambda/2$ centers.

The solution to these problems lies in using an overlapped subarray scheme [11]. In this case we need multiple inputs to the feed array, such as a Butler matrix, whose outputs feed multiple subarrays. One method uses space feeding so that all Butler matrix output ports feed every subarray. We phase the multiple input ports in a Woodward synthesis (Section 4-8) to form a flat-top pattern output for the subarray. The flat-top pattern eliminates the misalignment between the subarray and feed patterns and allows the use of time-delay networks. The overlapped subarray method greatly complicates the feed network.

12-7 PATTERN NULL FORMATION IN ARBITRARY ARRAY

In Chapter 4 we discussed the control of an array pattern by manipulation of pattern nulls. In Section 4-9 we applied the Schelkunoff unit-circle method to line arrays. In Sections 4-21 and 4-22 we combined nulls found for rhombic and linear arrays to synthesize planar arrays by convolution. We can extend these ideas to arbitrarily positioned and oriented arrays. We start with the equation for the pattern for any array with feed coefficients w_n:

$$\mathbf{E}(\theta, \phi) = \sum_{n=1}^{N} [E_{\theta n}(\theta, \phi)\hat{\theta} + E_{\phi n}(\theta, \phi)\hat{\phi}] w_n e^{j\mathbf{k}\cdot\mathbf{r_n}} \tag{12-28}$$

The array elements can be oriented arbitrarily and located at $\mathbf{r}_n = (x_n, y_n, z_n)$ with the vector propagation constant $\mathbf{k} = k(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. We calculate the pattern for one element of the array in a particular direction $\mathbf{k}_m = k(\sin\theta_m\cos\phi_m, \sin\theta_m \sin\phi_m, \cos\theta_m)$:

$$\mathbf{S}_{m,n} = [E_{\theta n}(\theta_m, \phi_m)\hat{\theta} + E_{\phi n}(\theta_m, \phi_m)\hat{\phi}] e^{j\mathbf{k}_m\cdot\mathbf{r_n}} \tag{12-29}$$

The array pattern in direction m is found by combining Eqs. (12-28) and (12-29):

$$\mathbf{E}_m(\theta_m, \phi_m) = \sum_{n=1}^{N} \mathbf{S}_{m,n} w_n \tag{12-30}$$

Since we only have one set of coefficients, w_n, only a single linear combination of the theta and phi components can form a null in a given direction. We project the polarization components of each element onto a particular polarization that reduces Eq. (12-29) to a single scalar term. For a given direction we define the vector $[\mathbf{S}_{i,n}]$ using patterns of the array elements in direction $\mathbf{k}_i$ evaluated projected on a given polarization:

$$[\mathbf{S}_{i,n}] = [S_{i,1}, S_{i,2}, \ldots, S_{i,N}]^{\mathrm{T}} \tag{12-31}$$

The operation $[\cdot]^{\mathrm{T}}$ is the matrix transpose and $[\mathbf{S}_{i,n}]$ is a column vector.

We start with an initial vector $[w_0]$ of array feeding coefficients that includes amplitudes and phases. We modify the coefficients to form a pattern null $\mathbf{k}_1 = k(\sin\theta_1\cos\phi_1, \sin\theta_1\sin\phi_1, \cos\theta_1)$ by multiplying the coefficient vector by a matrix found from the pattern of each element in that direction:

$$[w_p] = \left\{\mathbf{I} - \frac{[S_{1,n}]^*[S_{1,n}]^{\mathrm{T}}}{[S_{1,n}]^{\mathrm{T}}[S_{1,n}]^*}\right\}[w_0] = \left\{\mathbf{I} - \frac{[S_{1,n}]^*[S_{1,n}]^{\mathrm{T}}}{A}\right\}[w_0] = [P_1][w_0] \quad (12\text{-}32)$$

The product of the column vector and its transform generates an $n \times n$ matrix. The coefficient A is a summation: $A = \sum_{n=1}^{N} S_{1,n}S_{1,n}^*$. Equation (12-32) is derived from Gram–Schmidt vector orthogonalization [12, pp. 41–46]. The vector $[w_0]$ is the initial state of the array that includes any amplitude distribution to lower sidelobes and the phasing to scan the beam. We form nulls after scanning the beam.

This process can be extended to multiple nulls, but numerical precision and the number of elements limit the process and it is exact only for the last null. Equation (12-32) extends to m ($< N - 1$) nulls by matrix multiplication:

$$[w_p] = [P_m]\cdots[P_2][P_1][w_0] \quad (12\text{-}33)$$

The method does not exactly match the locations of initial nulls, but it improves when we have large arrays. We do not need to store the matrix $[P_i]$ because its elements are easily formed from the vector $[S_{i,n}]$ when needed, and by starting with $[w_0]$ we work Eq. (12-33) from right to left and store the new excitation vector at each step.

To design an array we specify $N - 1$ nulls and one desired signal direction. We extend Eq. (12-30) to a matrix for a given beam direction $E_1(\theta_1, \phi_1)$ and the direction of $N - 1$ nulls:

$$\begin{bmatrix} E_1(\theta_1,\phi_1) \\ E_2(\theta_2,\phi_2) \\ \vdots \\ E_N(\theta_N,\phi_N) \end{bmatrix} = \begin{bmatrix} S_{1,1} & S_{1,2} & \cdots & S_{1,N} \\ S_{2,1} & S_{2,2} & \cdots & S_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ S_{N,1} & S_{N,2} & \cdots & S_{N,N} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} \quad (12\text{-}34)$$

We invert the matrix to find the feeding coefficients:

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} S_{1,1} & S_{1,2} & \cdots & S_{1,N} \\ S_{2,1} & S_{2,2} & \cdots & S_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ S_{N,1} & S_{N,2} & \cdots & S_{N,N} \end{bmatrix}^{-1} \begin{bmatrix} E_1(\theta_1,\phi_1) \\ E_2(\theta_2,\phi_2) \\ \vdots \\ E_N(\theta_N,\phi_N) \end{bmatrix} \quad (12\text{-}35)$$

All N coefficients of the vector E are not necessarily zero, we could solve for multiple beams with individually specified levels. The array has N degrees of freedom. This analysis ignores the effects of mutual coupling, whereas scattering from nearby objects can be included in the element patterns because this development has not assumed identical array elements. We can use the results of Section 3-11 to compensate the network for mutual coupling, but this development leads to an adaptive array which changes its feeding coefficients based on incident fields.

12-8 PHASED ARRAY APPLICATION TO COMMUNICATION SYSTEMS

Radar has been the chief application of phased arrays, but small arrays find application in communication systems. Terrestrial line-of-sight (LOS) microwave systems use multiple antennas to reduce downtime due to fading, but these systems operate at very high connectivity. Cellular telephone operates with high multipath that produces a Rayleigh probability distribution because most of the time there is no direct signal path (Section 1-19). Diversity combining improves connectivity in the same manner as the LOS system by providing an alternative path when the signal fades in the first path. The LOS system has slowly varying fades due to changing atmospheric conditions, while cellular telephone has rapid fades as the user moves. Diversity combining reduces signal null depths but does little to increase the average signal.

The second cellular telephone system problem is channel capacity. The number of users continues to grow and their demand for service soon saturates existing systems. Adding phase arrays to base stations provides multiple beams to subdivide the cells that allow improvements without building new sites. Beyond using fixed multiple beams formed by networks such as Butler matrices, adaptive arrays (*smart antennas*) increase capacity. Antennas are the same pieces of metal; adaptive electronic beamforming is the smart part. Since adaptive array analysis and design are mostly signal processing, they will not be discussed in detail.

When signals fade in one channel because the user or a scattering object moves, a second signal uncorrelated to the cause of the fade may provide a better communication path. If we have two signals, it may be better to combine them to reduce fading. We discuss two signals, and the extension to more is straightforward. Imagine two antennas and a network that connects them. The best system uses *maximal ratio combining* (MRC), which adjusts both amplitude and phase (complex phasers) of the two signals before adding them. MRC requires the most complicated equipment, because we need a variable power divider combined with a phase shifter. *Equal gain combining* (EGC) is a phased array with only a phase shifter that brings both signals to a common phase before combining them. *Selective combining* (SEC) detects the strongest signal and switches to it and ignores the second signal. Finally, we have *switch combining* (SWC), which switches when the signal drops below a threshold. Of course, SWC can switch back and forth rapidly when the signal drops in both paths. SWC is the lowest-cost system because it needs only one receiver front end.

A communication system provides connectivity at a given probability and we accept dropped calls and lost data. Fortunately, voice encoding can tolerate high BER (bit error rate), whereas LOS systems use error detection and retransmit to lower errors to near zero. Diversity gain measures the increased signal level at a given probability at the average signal level. Propagation models give us the log-normal (median) signal level, about which the level varies rapidly due to multipath. Diversity gain is usually given relative to the 90% signal reliability level, the improvement that allows a lower median signal level. We relate the improvement to the difference in signal level in the two channels Δ (dB) and the branch correlation ρ. Branch correlation measures the independence between the two signals. For example, polarization separation is one method to generate two channels, and cross-polarization in the two channels produces signal correlation. In other implementations two antennas are separated either vertically or horizontally, so a fade in reception at one antenna has a minor effect on the other. Two channels that receive the same signal have $\rho = 1$, whereas the ideal two-channel

system has $\rho = 0$. The following diversity gain (dB) at the 90% signal reliability level has been found empirically for the various combining techniques [13, p. 48]:

$$\text{Gain (dB)} = \begin{cases} 7.14e^{-0.59\rho}e^{-0.11\Delta} & \text{MRC} \\ -8.98 + 15.22e^{-0.2\rho}e^{-0.04\Delta} & \text{EGC} \\ 5.71e^{-0.87\rho}e^{-0.16\Delta} & \text{SEC} \end{cases} \tag{12-36}$$

MRC always produces a positive diversity gain. Selecting the larger signal has less benefit than MRC, and the EGC system can actually give a negative diversity gain under unfavorable situations. If the signal drops in one channel, adding two signals reduces overall performance because in essence the high signal level channel is reduced by 3 dB by the power divider.

Adaptive arrays (smart antennas) improve the system by directing a main beam at a particular user while producing nulls in directions of interfering signals. These signals could be from nearby cells using the same frequencies. Adaptive receiving arrays enable closer frequency reuse between cells and increases system capacity while improving quality of transmission. We either adaptively control element excitation at RF or detect and combine them at IF. The weak link in the cellular telephone system is the transmission from the mobile to the base station, and the base station contains the array used to separate signals, so we concentrate on the receiving path.

The array either uses the least mean squares (LMS) algorithm or an eigen-decomposition technique to set the element weights. LMS requires a known reference signal on which the system minimizes the errors. This system adjusts the weights to maximize the S/N ratio. The known signal can be a portion of the coded waveform transmitted by the mobile unit, which increases overhead and reduces transmit rate slightly when each unit sends a unique reference signal. The LMS algorithm can adjust weights to optimize the constant amplitude of receive signal, since a phased modulated wave (PSK, QPSK) has constant amplitude. Systems that optimize on a known reference signal include recursive least squares and Wiener filter methods.

Eigen-decomposition techniques operate without a known reference signal. A cross-correlation matrix is formed from the input signals by considering their locations using the terms of Eq. (12-29). A cross-correlation matrix is the product of the pairs of these terms using the complex conjugate of one factor. By solving the eigenvalue problem, the algorithm detects directions of incoming signals and the beamforming is based on the angular information of these beams. Two commonly used algorithms are MUSIC and ESPRIT.

To increase system performance significantly, we use digital beamforming, where we connect a receiver to every antenna. We can make as many copies of the signals after detection and form the beams digitally without loss of S/N. We had to detect the signals anyway to apply the adaptive array algorithms that set the element weights at RF. Digital beamforming eliminates the phase shifters and variable power dividers on the antennas and replaces them with receiver components. Element location and quantization values of the receiver components still affect the array patterns, but by using adaptive algorithms, enable production of identical arrays. The array adjusts to its location and the input signals by using signal processing—a smart antenna.

12-9 NEAR-FIELD MEASUREMENTS ON PHASED ARRAYS

Near-field measurements produce excellent diagnostics of phased arrays. We exercise the phase shifters and measure element outputs directly. Initially we built small loop

probes and moved them across the array face and recorded measurements. The small probes have little effect on the array elements, due to mutual coupling, because they are poor antennas. We improve precision by fabricating scanners to hold and position the probes at the array elements. Unfortunately, if we move the probe away from the array surface to reduce the probe effect on the elements, increased signal is received from more than one array element and it is difficult to separate element responses. The answer lies in using the transforms of near-field measurements to separate element responses.

Figure 12-27 shows the simulated response of radiation to a 90° beamwidth probe positioned directly over a 64-element linear array made with 90° beamwidth elements. The array sampled a 30-dB Taylor linear aperture distribution to set its amplitudes, and the twentieth element amplitude was reduced by 6 dB. The two dashed curves give the responses at $\lambda/2$ and 3λ above the surface. The probe located at $\lambda/2$ produces a response with about a 3-dB dip at the twentieth element. The probe at 3λ above the array face gives a vague dip over a number of elements. Both probe heights measure responses with significant ripples. Many arrays have been measured and corrected using this approach.

A planar near-field measurement combined with simple mathematical transforms produces better results. We scan the near-field probe sufficiently above the array to reduce mutual coupling effects that change what we are measuring. Planar near-field measurements take readings at evenly spaced intervals and use the fast Fourier transform (FFT) to calculate the far-field response at points evenly spaced in $\sin\theta$-space, where θ is measured from the planar normal. The planar measurements use $k_x = \sin\theta_x = \sin\theta\cos\phi$ and $k_y = \sin\theta_y = \sin\theta\sin\phi$ found from an FFT along each axis. For example, we transform each row along the x-axis using the FFT algorithm and transform resulting columns of FFT on the rows along the y-axis. The FFT requires 2^N samples along the array face and produces the same number of pattern directions

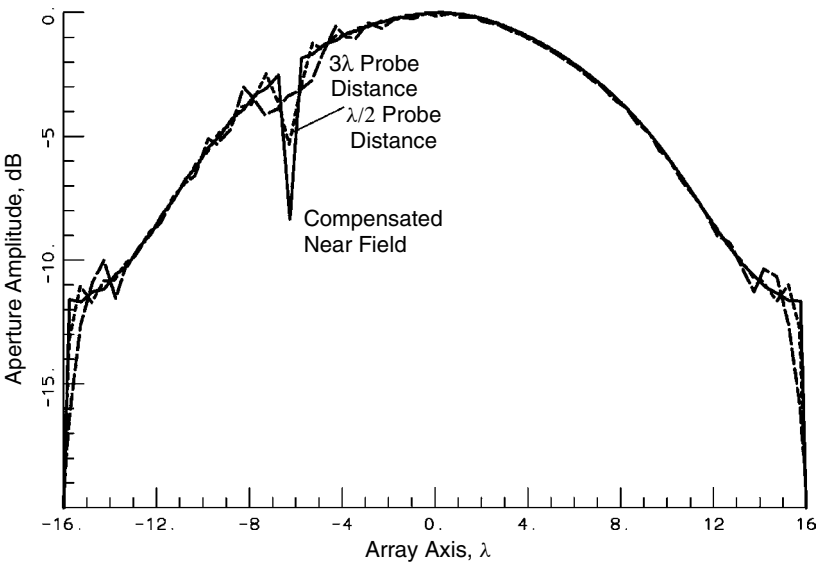


FIGURE 12-27 Near field measurements of a 64-element line array with 30-dB Taylor distribution at probe distances of 3λ, $\lambda/2$, and compensated to face.

TABLE 12-5 Planar Near-Field Sampling to a Far-Field Pattern Using FFT

Aperture Spacing, d (λ)	Sin θ_{max}	θ_{max} (deg)
0.25	2	Invisible space
0.5	1	90
0.75	2/3	41.8
1.0	0.5	30
2.0	0.25	14.48

evenly spaced in $\sin\theta$-space. We find the range of k_x-space from the aperture sampling spacing d:

$$\sin\theta_{max} = \pm\frac{\lambda}{2d} \tag{12-37}$$

Table 12-5 shows that sampling closer than $\lambda/2$ produces pattern points in invisible space, and for most arrays, larger sampling spacing fails to produce a point over each element. Since we want to capture the radiation from the edge elements, it is necessary to scan the near-field probe beyond the edge of the array face. If we space the probe a distance a above the array space, we should sample out to the maximum transformed angle. Given an array of length D and probe length L, the maximum angle for accurate measurements is given by $\theta_c = \tan^{-1}[(L - D)/(2a)]$. If we use $\lambda/2$ sampling distance, L extends to ∞ and we compromise by also assuming that the array element beamwidth captures most of the power radiated. After probing the field, we transform to the far field by using FFT^{-1}. This pattern has been multiplied by the probe antenna pattern and the pattern of a Huygens source (Section 2-2) of an aperture. The Huygens source has a pattern $E_H(\theta) = \cos^2(\theta/2)$. The far-field pattern also has a phase distribution relative to the array face, due to the distance between the face and the probe plane a distance a above it. We apply probe compensation and shift the far-field phase so that when we transform the k-space pattern back to the aperture, the scan plane lies on the array face. Given the desired position z for the new aperture plane, we first adjust the far-field pattern:

$$E_{new}(\theta_x, \theta_y) = \frac{E(\theta_x, \theta_y)e^{-jk\cos\theta_x(z-a)}e^{-jk\cos\theta_y(z-a)}}{E_{probe}(\theta_x, \theta_y)\cos^2(\theta_x/2)\cos^2(\theta_y/2)} \tag{12-38}$$

The solid curve on Figure 12-29 shows the results of these calculations on the array with altered element amplitude when $z = 0$. If the array contained more feeding errors, all of them would appear in the FFT of the far field to the near-field aperture plane. We use the same method to determine errors in paraboloidal reflectors, but we apply multiple steps to the z-axis plane position in the back transform to produce planar cross sections through the reflector surface.

REFERENCES

1. H. A. Wheeler, The radiation resistance of an antenna in an infinite array or waveguide, *Proceedings of IRE*, vol. 36, April 1948, pp. 478–488.

2. H. A. Wheeler, Simple relations derived from a phased-array antenna made of an infinite current sheet, *IEEE Transactions on Antennas and Propagation*, vol. AP-13, no. 4, July 1965, pp. 506–514.
3. R. C. Hansen, Linear arrays, Chapter 9 in A. W. Rudge et al., eds., *Handbook of Antenna Design*, IEE/Peter Peregrinus, London, 1983.
4. M. I. Skolnik, Nonuniform arrays, Chapter 6 in R. E. Collin and F. J. Zucker, eds., *Antenna Theory*, Part 2, McGraw-Hill, New York, 1969.
5. R. J. Mailloux, Periodic arrays, Chapter 13 in Y. T. Lo and S. W. Lee, eds., *Antenna Handbook*, Van Nostrand Reinhold, New York, 1992.
6. R. J. Mailloux, *Phase Array Antenna Handbook*, Artech House, Boston, 1994, pp. 393–399.
7. T. A. Milligan, Space tapered circular (ring) arrays, *IEEE Antennas and Propagation Magazine*, vol. 46, no. 3, June 2004.
8. M. I. Skolnik, J. W. Sherman III, and F. C. Ogg, Jr., Statistically designed density tapered arrays, *IEEE Transactions on Antennas and Propagation*, vol. AP-12, no. 4, July 1964, pp. 408–417.
9. J. L. Butler, Digital, matrix, and intermediate-frequency scanning, in R. C. Hansen, ed., *Microwave Scanning Antennas*, Vol. III, Academic Press, New York, 1966.
10. B. M. Schiffman, A new class of broadband microwave 90° phase shifters, *IRE Transactions on Microwave Theory and Techniques*, Vol. MTT-6, no. 4, April 1958, pp. 232–237.
11. R. Tang, Survey of time-delay beam steering techniques, *1970 Phased Array Conference*, Artech House, Boston, 1972.
12. J. E. Hudson, *Adaptive Array Principles*, IEE/Peter Peregrinus, Stevenage, Hertfordshire, England, 1981.
13. A. Paulraj et al., Space-time processing in wireless communications, *Proceedings of the 3rd Workshop on Smart Antennas in Wireless Mobile Communications*, Stanford University, Stanford, CA, 1996.

INDEX

Modern Antenna Design, Second Edition, By Thomas A. Milligan