COMMUNICATION ENGINEERING

COMMUNICATION SATELLITE ANTENNAS

SYSTEM ARCHITECTURE, TECHNOLOGY, AND EVALUATION

ROBERT DYBDAL

Communication Satellite Antennas: System Architecture, Technology, and Evaluation

Robert Dybdal

New York Chicago San Francisco Lisbon London Madrid Mexico City Milan New Delhi San Juan Seoul Singapore Sydney Toronto

The McGraw-Hill Companies

Library of Congress Cataloging-in-Publication Data
Dybdal, Robert.
Communication satellite antennas : system architecture, technology,
and evaluation / Robert Dybdal.
p. cm.
Includes bibliographical references and index.
ISBN-13:978-0-07-160918-0 (alk. paper)
ISBN-10:0-07-160918-0
1. Artificial satellites—Radio antennas. 2. Satellite dish antennas.
3. Artificial satellites in telecommunication. I. Title.
TL3035.D93 2009
621.382’54—dc22
2009017847

McGraw-Hill books are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. To contact a special sales representative, please visit the Contact Us page at www.mhprofessional.com.

Communication Satellite Antennas: System Architecture, Technology, and Evaluation

1 2 3 4 5 6 7 8 9 0 FGR FGR 0 1 9

ISBN 978-0-07-160918-0
MHID 0-07-160918-0

Sponsoring Editor
Wendy Rinaldi
Editorial Supervisor
Jody McKenzie
Project Manager
Madhu Bhardwaj, International Typesetting and Composition
Acquisitions Coordinator
Joya Anthony
Copy Editor
Michael McGee
Proofreader
Andy Saff
Indexer
Steve Ingle
Production Supervisor
George Anderson
Composition
International Typesetting and Composition
Illustration
International Typesetting and Composition
Art Director, Cover
Jeff Weeks
Cover Designer
12E Design

adequacy, or completeness of any information and is not responsible for any errors or omissions or the results obtained from the use of such information.

ABOUT THE AUTHOR

BOB DYBDAL has supported a broad base of military and commercial communication satellite programs and is affiliated with The Aerospace Corporation. His interest in antennas and RF systems developed at the ElectroScience Laboratory at Ohio State University, where he received a BSEE, MSc, and PhD in Electrical Engineering. He has been involved in a wide range of IEEE technical activities and is a past president of the Antenna Measurement Techniques Association. He holds patents in instrumentation, adaptive antennas, antenna tracking, satellite transponder designs, interferometry, and microwave components.

Contents at a Glance

Contents

Preface

Antenna systems are a fundamental part of communication satellite systems. Antenna technology has a long development history beginning with the fundamental experiments performed by Hertz in the 1880s, the development of broadcast antennas in the 1920s where fundamental concepts of antenna pattern shaping and array synthesis began, the microwave technology demonstrated during World War II, and today's technology and analysis capabilities. Antenna technology has had a significant impact not only on communication systems but also in radar, remote sensing, and other applications.

Antenna technology is extensively documented in IEEE publications and those of other organizations, including the Antenna Measurement Techniques Association. A number of excellent textbooks are available to educate future antenna developers, and a variety of books address specific antenna technologies. This book describes the way in which antenna technology is used in communication satellite systems. The book is motivated by a belief that practicing system designers and technology developers would benefit from a system view of antenna applications, a description of antenna technology, and guidance on methodologies needed in their evaluation. On an educational level, the material would be suitable for academic courses on applications of antenna technology to systems that have a major importance worldwide.

The material in this book has evolved from an innumerable collection of people spanning the development history and application of anten-nas. The technology heritage is very rich, spanning a variety of system applications, innovative designs, well-developed analysis capabilities, and instrumentation and measurement facilities. Future system development and application likewise depend on the efforts of a large number of people. Clearly, this publication is indebted to the efforts of many. On a personal level, the author is likewise indebted to many people, including peers, members of professional organizations, and the contractor and customer communities. One of life's riches is the opportunity to benefit from lively

technical debate, learn and teach, collaborate, and create and evolve in a technology area as vibrant as antennas and their system applications. The enthusiasm and encouragement of Wendy Rinaldi of McGraw-Hill and the careful editing of Madhu Bhardwaj and her colleagues are gratefully acknowledged.

Introduction

Satellite systems have had a profound effect on worldwide information dissemination. Early systems provided proof-of-concept demonstrations and established an initial operating capability. System capabilities have greatly extended beyond these early system designs in ways that were not foreseen at the inception of satellite systems. Early systems and technology available at that time provided limited service to large ground terminals and then dissemination by terrestrial means to system users. Today, a wide ranging number of services are available to individual system users having relatively modest user equipment requirements. Future system designs will continue to extend the services available to system users in ways that are not grasped today.

Existing satellite system maturity has been made possible by a wide range of enabling technologies. Today's launch vehicle, solar power arrays, and attitude stability technologies have resulted in satellite capabilities that could not have been imagined by early satellite developers. Today's satellite lifetimes greatly exceed those of the early satellites and often their own projected lifetimes. Electronic technologies likewise have made possible the development of capable systems for both the space and user segments comprising satellite systems. The development and demonstration of modulation formats and multiple access techniques that allow a collection of users to share satellite resources have had major roles in providing efficient and reliable communications for a multitude of system users and applications. Antenna systems have greatly increased in sophistication. Space segment antennas provide high gain capabilities to ease user requirements; can spatially isolate different portions of the field of view allowing the available spectra to be reused; and can mitigate interference. Of all the technologies used in the space segment, antenna systems are the most diverse as a result of different operating frequencies and system requirements. User segment antenna designs are also diverse, ranging from handheld designs for low data rate applications to very large ground terminals for high data rate transfer. The escalating

number of system users demand attention to cost-effective designs and economies of production to control system acquisition costs.

Future satellite systems will not only replenish existing capabilities but also provide capabilities that cannot be clearly envisioned today. While today's satellite system technologies are highly capable, future designs will benefit by development and further refinements and efficiencies. Technology evolution will continue to contribute to systems having additional capabilities and flexibilities, as well as reduced weight and power requirements and acquisition costs. This evolution will extend over all the diverse technologies used in satellite systems. In addition to component evolution, other developments in modulation, multiple access, and network techniques can also be envisioned. Utilization of software and digital technologies will also increase in future system designs. Like these other technologies, satellite antenna systems will continue to evolve to satisfy the objectives of future system designs.

Communication satellites have been developed for both commercial and military applications and the objectives of their applications differ. Commercial systems are configured to serve particular market segments and are intended to provide as much system capacity from the available frequency allocation as possible. These considerations result in system designs that have relatively fixed coverage requirements and techniques to expand system capacity by reusing the same frequency spectra. Serving the required coverage with multiple beams to isolate users in different portions of the coverage area and reusing the same frequency subband when sufficient spatial isolation is available is one technique. Another commonly used technique uses orthogonal polarizations to communicate independent data channels. Military systems, by contrast, require the capability to respond to capacity and coverage needs that change over the satellite's lifetime because of evolving geopolitical requirements. Additionally, military users have long had concerns regarding intentional interference or jamming. Techniques to protect systems from interference have been developed and used operationally.

While commercial and military systems have differing objectives, both share common development requirements. Independent of the application, SWaP, size, weight, and power, are of paramount importance for the space segment. Reliability is also essential and extensive system testing and redundant components are required to assure satisfying orbital lifetime objectives. Acquisition cost is another critical factor. As the number of system users continues to increase, providing sufficient performance to reduce user requirements and permitting the development of cost-effective user segment designs are the most important areas of system design and planning. Testing is an essential part of system development, and as the number of users continues to

increase, techniques to test on a production basis must be developed. These issues will have increased importance for future system designs as the level of complexity increases and the number of system users continues to grow.

System design is an iterative process, and the amount of iteration will grow as system complexity and the number of users continues to increase. The system design process illustrated in Fig. 1 indicates the iterative nature that must be addressed by system planners. At a top level, system-level objectives define the user data transfer and coverage requirements, the frequency allocations to be used, and preliminary assessments of *G*/*T* and ERP (effective radiated power) constraints for both the space and user segments. These top-level requirements are used to develop system design concepts based on preliminary assessments of performance capabilities for the space and user segments. A most important and fundamental part of system definition is questioning and understanding the impacts of system requirements. As the system definition proceeds, the requirements will evolve as necessary to configure viable system designs. The importance of questioning system requirements cannot be overstated. The system design concepts are compared with launch vehicle constraints for the space segment and compared with production costs for the user segment. Technology estimates play a major role in these preliminary system designs and development risk for implementation must be addressed. Other choices that are examined at this time are modulation formats to be used in user communications and multiple access techniques that allow users to share the space segment resources. A significant number of system tradeoffs exist and the process iterates multiple times in developing an

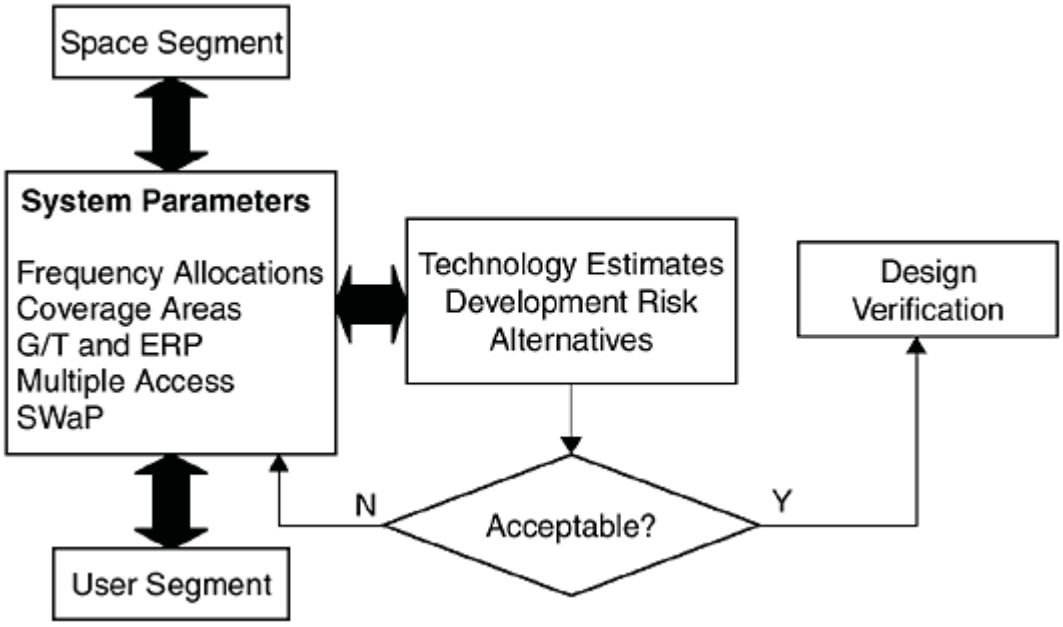


Figure 1 The system design process

acceptable system design. System design development and definition clearly must provide a balance between the space and user segment performance requirements in deriving system-viable implementations. As system capabilities increase and afford increased service requirements to service a greater number of system users, this iterative process becomes more complex and extensive.

While the system planning and development process is ongoing, the capabilities of many different technologies are also assessed in support of the system definition. The scope of this effort likewise becomes more extensive as system design complexity increases. Design implementation choices, such as the fabrication alternatives of MMIC (monolithic microwave integrated circuits) and ASIC (application-specific integrated circuits) implementations to support specialized needs of the system design and the use of digital technology, are addressed in selecting the system electronics. System design choices for space and user antenna requirements become extensive with the complexity of requirements and technology alternatives. Antenna systems in particular afford opportunities for creative solutions because the system requirements for each application differ and “standard” designs are nonexistent. In addition to the component selection, this preliminary system definition phase needs to address testing requirements and the associated facilities needed to evaluate not only components but integrated subsystems and systems. While many technology choices and technical issues must be addressed, acquisition costs must also be examined and tradeoffs in system design evaluated on a cost basis. System definition is a multifaceted undertaking that requires careful assessments of requirements, technology alternatives, the allocation of resources, and economic impacts.

Antenna technology to support system definition and development has a major role in devising viable system designs. System development, to date, has demonstrated a diverse antenna technology base to meet requirements for specific system applications. This antenna technology base has greatly contributed to existing system capabilities. Future system designs will continue to generate even more diverse antenna designs and extend component-level antennas to antennas integrated into system-level designs. Much opportunity exists here to develop creative solutions for future system needs. This book was prepared to provide guidance for future communication satellite antenna developments and endeavors to provide a system background to assist system planners and technology developers. Such development requires insight into system architectures, antenna technology alternatives, and methods to evaluate both their component- and system-level performance.

The organization of the book has the following format. An overview of the parameters that characterize antennas is presented to provide a basis to quantify antenna performance. Antenna technology required

in communication satellite systems is described in some detail. System architectures for both the space and user segment are reviewed so that antenna interfaces with system designs are understood. Practical system designs must assess propagation limitations and link analyses that determine the capabilities afforded by candidate system designs. The increased number of communication, radar, and navigation services and the substantial increase in user demands for these services result in potential interference between systems. Future system designs therefore will require increased design attention to interference susceptibility and include techniques to mitigate interference. Space and user segment antenna technologies are separately addressed, and technology applications to satisfy typical system requirements are discussed. Antenna performance evaluations must address facility alternatives and techniques to provide meaningful assessments of their performance. The processes used in the development and characterization of antenna systems are then reviewed.

Chapter

1

Fundamental Parameters

1.1 Overview

Communication satellite systems depend significantly on both space segment and user segment antenna designs. Space segment antennas must meet their performance requirements over their specified coverage areas with allowance for satellite attitude variations. User segment antennas likewise must meet their performance requirements while tracking the satellite in orbit. Antenna requirements depend on specific program needs, and a significant diversity of technology has been developed to accommodate the diverse objectives of individual programs. As a result, space segment antennas are the most diverse technology in the space segment, and specific designs for one application cannot be applied to other applications. User segment antenna hardware likewise exhibits a wide variety of antenna hardware ranging from small handheld technologies to much larger ground terminal antennas, which are often associated in the public's mind with communication satellite antenna systems. A review of the system parameters used to quantify antenna performance is presented as a basis for subsequent chapters.

1.2 Antenna Parameters

Antenna parameters must describe both the spatial characteristics and terminal interfaces with system electronics. The spatial characteristics specify the two-dimensional description of the antenna's sensitivity variations in a coordinate system embedded in the antenna. These spatial characteristics must also indicate the antenna's polarization properties that define the orientation of the electric field during one RF (radio frequency) cycle. The antenna's terminal impedance quantifies the interface relations

with system electronics. Satellite system antennas are commonly in the class of aperture antennas. The relationship between the aperture size and spatial characteristics is a most important issue in system sizing. This relationship dictates the antenna's gain levels and beamwidth requirements. Perhaps the most commonly asked question regarding antennas is the size required to meet system requirements. This question is typically followed by a request to explain why the size must be that large. Noise in receiving systems is an important system parameter and is characterized by the antenna noise temperature at the antenna's terminal. The antenna noise temperature added to the receiver noise temperature equals the total system noise temperature, an important factor in the performance of receiving antennas.

1.2.1 Spatial Characteristics

The spatial characteristics describe the spatial variation of the antenna's sensitivity. They also describe the vector nature of the antenna's field distribution in a coordinate system referenced to the antenna's structure. Commonly, satellite systems use aperture antennas that have a distribution of fields in the aperture and a corresponding distribution of fields in space. The coordinate system used for this specification generally places the aperture plane with the XY plane as indicated in Fig. 1-1. At a sufficient distance from the aperture (referred to as the antenna's far field), the variation of the fields becomes invariant with the range from the antenna's aperture. The electric field quantities, E_θ and E_φ, are orthogonal to one another as specified and vary with separation R from the aperture as 1/R. The power density in the far field is proportional to $(|E_\theta|^2 + |E_\varphi|^2)/Z_o$, where Z_o equals 120π and is the free space impedance.

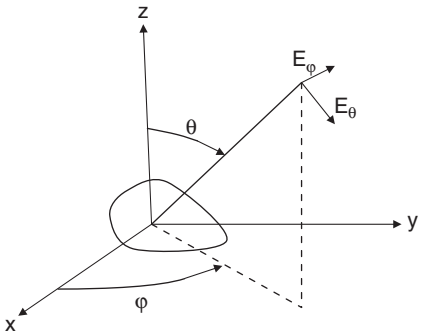


Figure 1-1 Coordinate system for aperture antennas

The relationship between the fields in an antenna's aperture to the spatial distribution is the radiation integral [1]

$$g(k_x, k_y) = \iint F(x, y) \exp(j\,(k_x x + k_y y)\ \mathrm{dx\ dy}$$

where $g(k_x, k_y)$ is the pattern (voltage), $F(x, y)$ is the field distribution in the aperture having coordinates x and y, the integration limits are the physical extent of the aperture, and

$$k_x = k\ \sin\ \theta \cos\ \varphi$$
$$k_y = k\ \sin\ \theta \sin\ \varphi$$

where k is the free space wavenumber equal to $2\pi/\lambda$, λ is the wavelength equal to c/f where c is the speed of light, and f is the RF frequency. The aperture fields are vector functions representing the polarization properties of the aperture fields. The variation of the antenna's sensitivity with direction is referred to as its pattern, and $g(k_x, k_y)$ is proportional to the electric field variation. This relation assumes the spatial fields are sufficiently separated from the aperture that the fields are independent of the range, a condition referred to as the far field. Commonly, the required far field separation for aperture antennas is taken as $2D^2/\lambda$, where D is the aperture width. It should also be noted that antennas generally satisfy reciprocity relations so that at the same frequency, the characteristics are identical independent of whether the antenna is transmitting or receiving. The exception is when the antenna incorporates nonreciprocal devices such as active amplifiers.

The relation between the aperture fields and the far field pattern is a two-dimensional Fourier transform. Similarly, the aperture field is the inverse two-dimensional Fourier transform of the far field pattern. The antenna size is thus related to the beamwidth in the far field, and likewise the beamwidth in the far field is related to the antenna size through the transform. The familiar properties of Fourier transforms are inherent in antenna design. If the aperture fields have an amplitude taper, the far field beamwidth broadens and the sidelobes surrounding the main beam are reduced. If the aperture fields are in phase over the extent of the aperture, the beam maximum is normal to the aperture plane. If the aperture fields have a linear phase gradient over the extent of the aperture, the beam maximum is normal to the phase gradient, a consequence of the familiar shifting theorem of Fourier transforms.

Antenna gain measures the antenna's ability to transfer or receive signals in a particular direction. It is referenced to an idealized lossless antenna having uniform sensitivity in all directions. In a sense, this reference for antenna gain follows the definition of electronics gain that is referenced to the transfer response of an idealized,

lossless "straight wire." The maximum value of antenna gain for aperture designs equals

$$G = \eta\,(4\pi \mathrm{A}/\lambda^2)$$

where η is the antenna efficiency (≤ 1), A is the physical area of the antenna's aperture, and λ is again the free space wavenumber. Ideally, an antenna having 100% efficiency is lossless and has an aperture distribution uniform in both amplitude and phase. Practically, this ideal antenna efficiency can only be approached, and the antenna efficiency of practical antenna designs falls short of the ideal value because of ohmic and impedance mismatch losses, the aperture amplitude and phase deviations from the ideal, and scattering and blockage from the antenna's structure. In determining the required antenna size or aperture area, an estimated value of the antenna's efficiency is required. The efficiency value depends on the specific antenna design.

Another term defined for receiving antennas is effective aperture, which equals

$$A_e = (\lambda^2/4\pi)\ \mathrm{G}$$

The received power equals the product of the incident power density and the effective aperture.

The far field parameters implicitly assume the antenna responds to an incident plane wave or a wave that approximates a plane wave. The far field criteria $2D^2/\lambda$ is derived based on the required range from the point of origin of a spherical wave such that the phase deviation over a planar surface of dimension D has a maximum value of 22.5° relative to an ideal in-phase plane wave.

Directivity or *directive gain* is another term that characterizes an antenna's directional properties. Directivity is a function of the antenna pattern or the variation of the antenna's sensitivity to different signal directions. Directivity differs from antenna gain because ohmic and mismatch losses are not included in directivity. Thus, antenna gain has a lower value than directivity. Directivity is defined by

$$D(\theta, \varphi) = 4\pi P(\theta, \varphi)/\iint(P(\theta, \varphi)\sin\theta\, d\theta d\varphi$$

The integral in the denominator is total power radiated or received from all directions. The fields of an antenna are vector quantities and (as will be discussed) have a principal polarized component with the design polarization state and, unavoidably, a cross-polarized component that is orthogonal to the principal polarization. Directivity is generally computed with the power pattern in the numerator limited to principal polarized fields and the total power in the denominator comprised of

both principal and cross-polarized terms. In this way, the directivity is determined relative to the design polarization of the antenna.

Antenna gain defines the signal power transfer and varies with angular coordinates. The antenna's beamwidth describes the angular width of the antenna's maximum response and is defined by the angular extent of the pattern within 3 dB of the peak antenna gain value or the HPBW, half-power beamwidth. The beamwidth of practical antennas can vary depending on which plane of the antenna pattern is used. Commonly, principal plane patterns display the patterns in the XZ and YZ planes in Fig. 1-1 when the beam maximum is coincident with the Z-axis. These patterns are great circle cuts through the sphere surrounding the antenna. When the beam is not coincident with the Z-axis, great circle cuts that intersect the peak gain level of the antenna are used. Depending on requirements, the patterns in other planes, also, great circle cuts are taken and sometimes referred to as φ cuts. When φ equals 45° or 135°, the patterns are referred to as diagonal cuts. Generally, multiple pattern cuts are used to judge the symmetry of the antenna's pattern. For aperture antennas, the beamwidth, θ_{hp}, equals

$$\theta_{hp} = K\lambda/D$$

where K is a constant that depends on the aperture distribution, λ is the wavelength, and D is the aperture width.

The parameters and their variation are illustrated by a simple analytic model. A circular aperture is assumed to have a uniform phase distribution and a rotationally symmetric amplitude having a $(1 - r^2)^p$ variation, where r is the aperture's radius. Example characteristics of this family of distributions are given in Table 1-1 where $J_{P+1}(x)$ is the Bessel function of order $p + 1$, and X equals $(\pi \mathrm{D}\lambda) \sin \theta$ with D equal to the aperture's diameter. When p equals 0, the amplitude distribution, like the phase distribution, is uniform over the aperture. The uniform aperture distribution has the maximum efficiency, a beamwidth factor of 58 in degrees, and a first sidelobe level that is 17.6 dB lower than the peak gain level. As the value of p increases, the efficiency decreases, the beamwidth broadens, and the sidelobe level decreases, all very familiar consequences of the Fourier transform relation between the

TABLE 1-1 Amplitude Taper Effects for Circular Apertures

P	Efficiency Loss, dB	Beamwidth Factor, K, degrees	First Sidelobe Level, dB	Pattern Variation
0	0	58	17.6	$J_1(X)/X$
1	1.2	73	24.6	$J_2(X)/X^2$
2	2.5	84	30.6	$J_3(X)/X^3$

aperture and the far field patterns. The pattern characteristics of reflector antennas are sometimes represented for p having a value of 1. These simple analytic forms lend themselves to simulation activities, and as the simulation is refined, characteristics of the actual antenna can be used to increase the simulation fidelity.

The antenna gain and the antenna beamwidth depend on the electrical size of the antenna, that is, the size in wavelengths. The antenna gain increases with the square of the electrical size while the beamwidth is inversely related to the electrical size. Both values clearly depend on the specifics of the antenna's design. For preliminary system sizing, an efficiency of 55% and a beamwidth factor of 70^{o} are often used. As the design evolves, such values are updated. Using these parameter values, the gain and beamwidth are plotted in Fig. 1-2 for various aperture sizes in wavelengths. Values of antenna gain and beamwidth for specific cases as a function of frequency are given in Figs. 1-3 and 1-4, respectively.

In practice, detailed computer codes are available to accurately project the performance of a wide variety of antenna technology used in communication satellite systems. Such analyses provide the means of refining the values of the nominal parameters used in preliminary system sizings, as indicated here. These nominal values can also be useful for "mental estimates" of antenna performance. Notice that the speed of light is approximately 1 ft/nsec and therefore the number of wavelengths per foot equals the frequency in GHz. For example, a 10-ft antenna at 10 GHz has a diameter of 100 wavelengths. Using a beamwidth factor of 70, the beamwidth equals about 0.7^{o}. For a circular aperture, the antenna gain

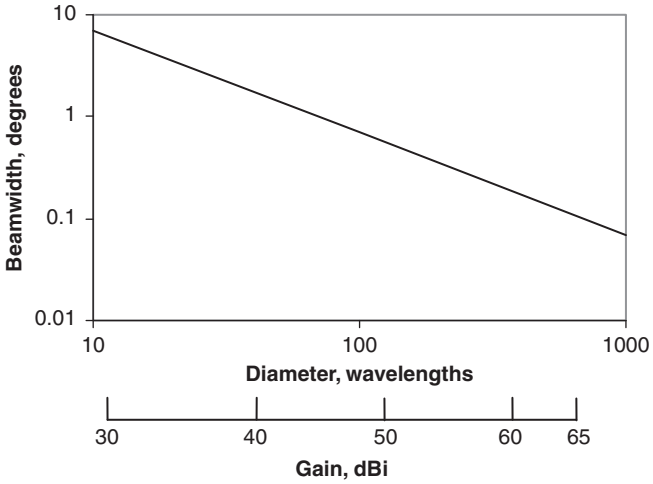


Figure 1-2 Nominal antenna gain and beamwidth values

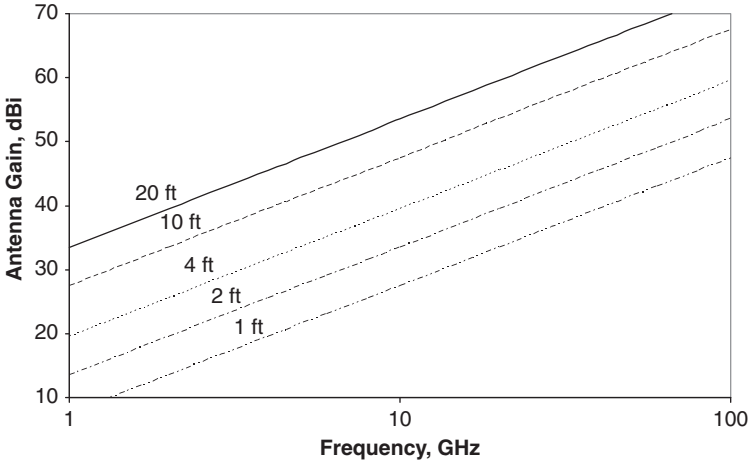


Figure 1-3 Antenna gain variation with frequency

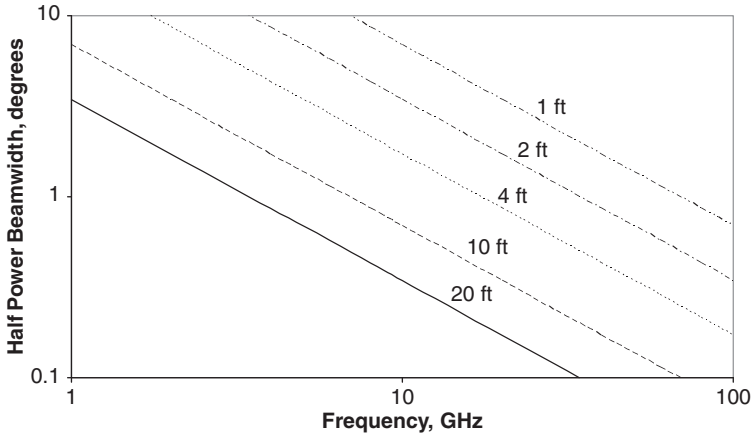


Figure 1-4 Antenna beamwidth variation with frequency

equals $\eta(\pi D/\lambda)^2$. The product $\eta\pi^2$ corresponds to about 7.3 dB, for a 55% antenna efficiency. The aperture diameter equals 100 wavelengths, and 20 times the log of 100 equals 40 dB. The antenna gain thus equals about 47.3 dBi (this indicates antenna gain relative to an isotropic gain level). This process may prove useful when rough estimates of antenna parameters are required and detailed calculation is unavailable.

1.2.2 Polarization

The vector nature of electromagnetic waves is specified by the polarization produced by an antenna, and propagating in free space. Polarization specifies the orientation of the electric field during one RF cycle. The most general polarization state is elliptical where the electric field traces out an ellipse. For every polarization, a unique orthogonal polarization exists, where orthogonal denotes ideal isolation between a receiving antenna and an incident field having orthogonally polarized states. Polarization is characterized by three parameters. One is an axial ratio equal to the ratio of the major and minor axes of the polarization ellipse. The second is the tilt angle specified by the alignment of the major axis of the ellipse in a reference frame. The third is the polarization sense specified by the familiar right- or left-hand rotation, as viewed in the direction of propagation.

The nominal orthogonal polarizations are linear and circular. Linear polarizations are typically indicated as vertical and horizontal, and ideal linear polarization confines the electric field to a plane. Linear vertically polarized antennas do not respond to horizontally polarized fields and thus the linear polarizations must be spatially aligned in use. Circular polarization is comprised of two orthogonal linear components having a 90^{o} phase difference. Over one RF period, the electric field traces out a circle. Circular polarization does not require the polarization alignment that linear polarization does, and for that reason circular polarization is widely used in satellite communication systems. Circular polarization components are orthogonal when their sense differs. Right-hand circular polarization sense is orthogonal to left-hand circular polarization sense. These polarization senses follow the familiar right- and left-hand rules when viewed in the direction of propagation.

Practical antennas are not ideally polarized and are mixtures of the two orthogonal components. The cross-polarized antenna response quantifies the degree to which the antenna deviates from the ideal polarization. At a system level, two issues result from the finite cross-polarized components:

1. What signal loss results from the cross-polarized components, a parameter referred to as polarization mismatch loss?
2. When orthogonally polarized signals are used to communicate independent data streams in polarization reuse designs, what is the isolation between orthogonal pairs?

The axial ratio, r, can be expressed [2] in terms of the circularly polarized components as

$$r = (E_R + E_L)/(E_R - E_L)$$

where E_R and E_L are the amplitudes of the right- and left-hand polarization components, respectively. Notice that the numeric value of axial ratio is positive for right-hand components and negative for left-hand components. Normally, axial ratio is given in a logarithmic value that involves the magnitude of the axial ratio. The level of the cross-polarized component relative to the principally polarized component can be calculated as presented in Fig. 1-5.

Normally, the axial ratio of incident fields and antenna systems are known, but the relative orientation of the tilt angles of their respective polarization ellipses are unknown. Both the polarization mismatch loss and polarization isolation depend on the relative orientation of the two polarization ellipses of the incident field and receiving antenna. A statistical approach [3] is presented as a means of understanding the variations resulting from unknown polarization ellipse alignment.

The polarization efficiency has been defined in terms of the axial ratios and orientation of the polarization ellipses of the incident field and receiving antenna. Polarization mismatch loss is determined from polarization efficiency when the incident field and receiving antenna have the same polarization sense. Polarization isolation is determined from polarization efficiency when the incident field and receiving antenna have opposite polarization senses. Polarization efficiency [2] equals

$$\eta_p = ½ + A + B\cos\Delta$$

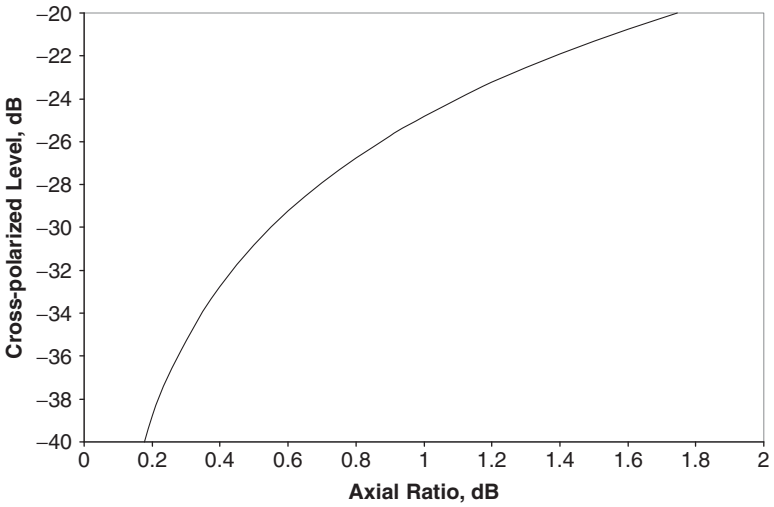


Figure 1-5 Cross-polarized level versus axial ratio

where

$$A = 2r_w r_r \Big/ \left[\left(1 + r_w^2\right)\left(1 + r_r^2\right) \right]$$

$$B = \left(1 - r_w^2\right)\left(1 - r_r^2\right) \Big/ \left[2\left(1 + r_w^2\right)\left(1 + r_r^2\right) \right]$$

where the subscripts "w" and "r" refer to the axial ratios of the incident wave and the receiving antenna, respectively. When the sense of incident field and the receiving antenna have the same polarization sense, A is positive because the product of the numeric value of axial ratios is positive when both senses are the same. When the senses of the incident field and receiving antenna have opposite polarization senses, A is negative. The angle, Δ, is the phase difference between the polarization components and equals twice the difference in the tilt angle orientations of the ellipses.

The commonly used bounds on polarization efficiency are $\frac{1}{2} + A \pm B$. The statistical variation of the polarization efficiency is derived by assuming the relative orientations of the tilt angle of the incident field and receiving antenna are equally likely and uniformly distributed over 0 to π, corresponding to Δ being equally likely and uniformly distributed over 0 to 2π. The first order (mean) statistics are determined from

$$E_p = (\tfrac{1}{2}\pi) \int \eta \, d\Delta$$

$$= \tfrac{1}{2} + A$$

where the integration extends over 0 to 2π. When both the incident field and receiving antenna have the same polarization sense, the mean efficiency is $\geq \frac{1}{2}$, and when their polarization senses are opposite, the efficiency is $\leq \frac{1}{2}$. Additionally, if either the incident field or the receiving antenna or both is ideally linear, the mean polarization efficiency is $\frac{1}{2}$, or the familiar 3 dB loss.

Similarly, the second order (variance) statistics are determined from

$$V_p = 1/2\pi \int (\eta - E_p)^2 \, d\Delta$$

$$= B^2/2$$

where again the integration extends over 0 to 2π. The standard deviation of the polarization efficiency, σ, about its mean value equals $B/2^{\frac{1}{2}}$. The polarization efficiency statistics have a non-zero mean value, so the second-order statistics are generally expressed as the $\pm 1\ \sigma$ spread about the mean value. Further, the polarization efficiency statistics are not Gaussian. The peak-to-peak bounds are $\pm B$ for these statistics, while the rms spread is ± 0.707 B. The bounds equal 1.41 times

the standard deviation. By contrast, peak-to-peak variations for Gaussian statistics are often taken as $\pm 3\sigma$, well in excess of the possible peak-to-peak excursion for the polarization efficiency.

Example values illustrate the statistical variations and two different cases, one for an incident field having a 0.5 dB axial ratio and the second with a 2 dB axial ratio. The polarization mismatch loss levels in Fig. 1-6

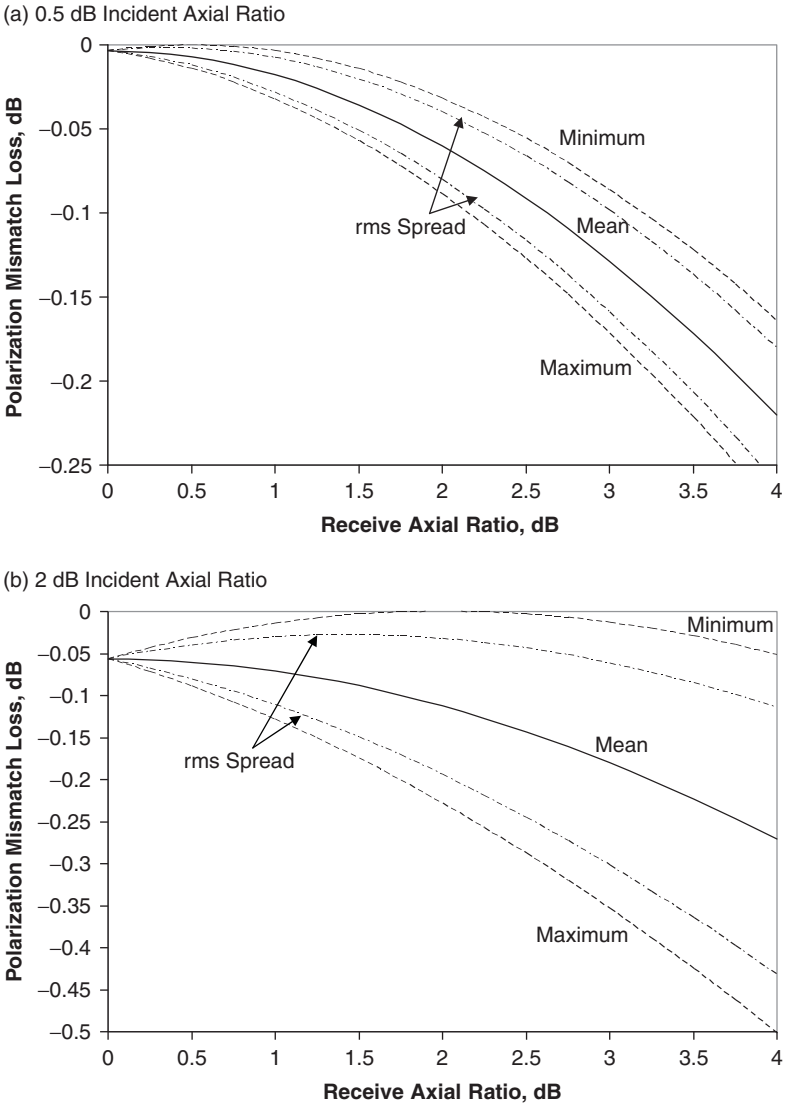


Figure 1-6 Polarization mismatch loss statistics

illustrate the mean, the minimum, the maximum, and the mean ± 1 σ values. The two examples illustrate the variation in the statistical values increase as the axial ratio of the incident field increases and as the axial ratio of the receiving antenna increases. Further, notice that the matched polarization condition when the incident field and receiving antenna have the same axial ratio value and their polarization tilt angles are coincident has a finite probability of having no polarization mismatch loss. System applications for polarization reuse require high polarization purity and incident fields having axial ratio values on the order of 0.5 dB. Other applications that seek reasonable polarization mismatch loss generally limit axial ratio values to about 2 dB. For example, if both the axial ratios of the incident field and receiving antenna are limited to 2 dB, the maximum possible polarization mismatch loss is less than 0.25 dB, and on average the polarization mismatch loss is about 0.1 dB, corresponding to the mean value.

The polarization isolation values in Fig. 1-7 similarly illustrate the statistical variations when the incident field has a 0.5 and 2 dB axial ratio. When orthogonally polarized components are used in frequency reuse designs to increase system capability, high levels of polarization purity are required of both the incident field and the receiving antenna. For example, if the incident field and the receiving antenna both have 0.5 dB axial ratios, the minimum polarization isolation is about 25 dB. When design attention is not paid to polarization purity, the isolation significantly degrades. Ideal polarization isolation requires the incident field and receiving antenna to have the same axial ratio and orthogonal polarization tilt angle orientations. The results indicate a finite probability of that condition being satisfied.

Often, circular polarization is produced by combining orthogonal linear components with a quadrature hybrid to produce circular polarization. The axial ratio resulting from amplitude and phase imbalance [4] in combining two orthogonal linearly polarized components is illustrated in Fig. 1-8. As the authors point out, the inherent cross polarization in the linear components is not included in this analysis and must be considered in practical designs.

1.2.3 Impedance

The interface between the antenna and the system electronics must also be specified. Maximum power transfer requires matched impedance characteristics, and deviations from ideal matched impedance result in reduced power transfer referred to as mismatch loss. Measured terminal parameters are generally performed using network analyzer instrumentation, and such measurements are expressed in scattering matrix parameters. The reflected components are expressed in the voltage

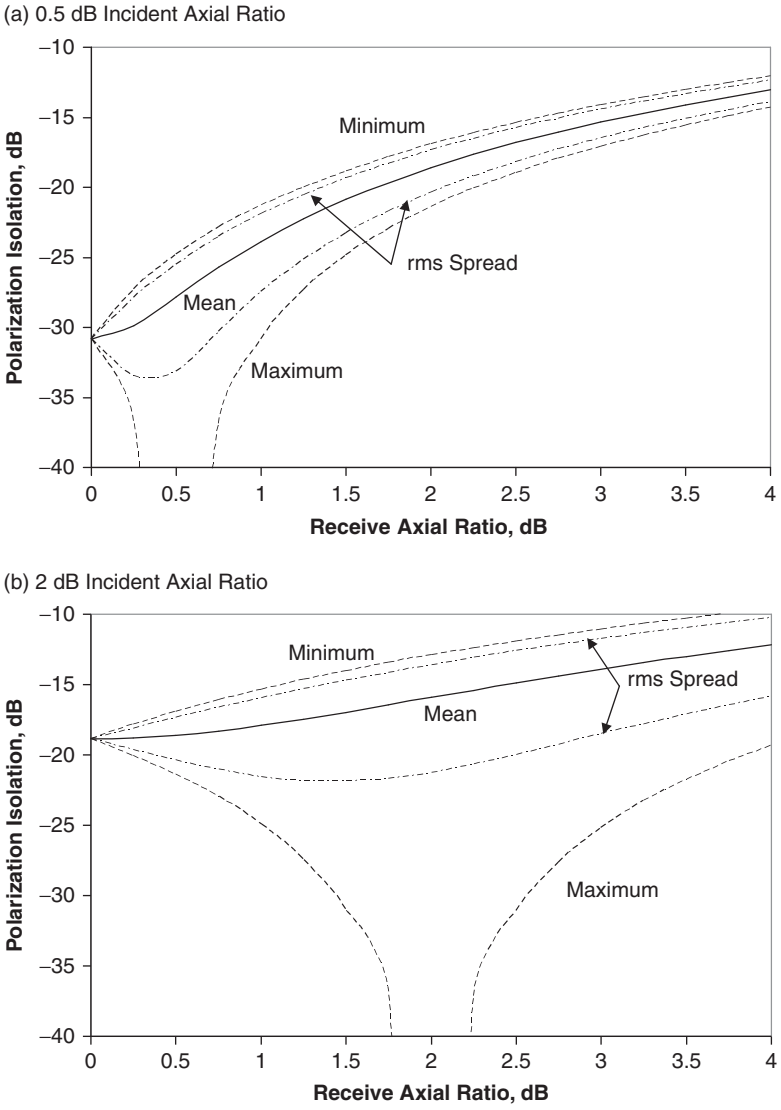


Figure 1-7 Polarization isolation statistics

reflection coefficient denoted by S_{11}, while the voltage transmission properties are denoted by S_{21}. The insertion loss for passive components equals $|S_{12}|^2$ between terminals 1 and 2 and for active electronics, $|S_{12}|^2$ is the insertion gain. The mismatch loss is $1 - |S_{11}|^2$. Antenna impedance

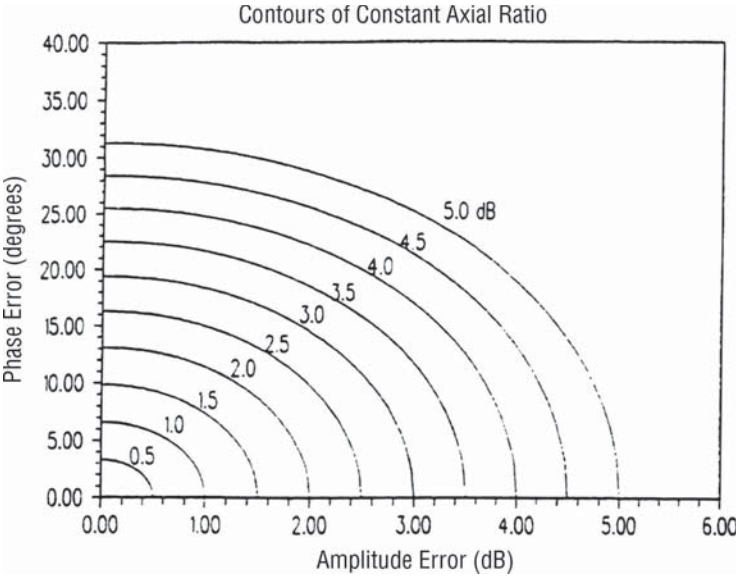


Figure 1-8 Hybrid imbalance impacts on axial ratio [4] (©1990 IEEE)

values are commonly specified by the voltage reflection coefficient S_{11}, the return loss RL, or the VSWR (voltage standing wave ratio), and are expressed as

$$\text{RL} = 20 \log (|S_{11}|)$$

and

$$\text{VSWR} = (1 + |S_{11}|)/(1 - |S_{11}|)$$

where physically VSWR is the ratio of the maximum and minimum values of the voltages along a transmission line. The return loss that equals $20 \log |S_{11}|$ is commonly used when network analyzer measurements are used. The mismatch loss is $1 - |S_{11}|^2$. Example values of impedance mismatch loss for the three common ways of expressing mismatch are given in Fig. 1-9. In specifications, VSWR and return loss are most commonly used, and the relationship between these parameters is given in Fig. 1-10.

In practical system designs, neither the antenna nor the interface electronics are ideally matched. As a result, multiple reflections between the antenna and electronics occur giving rise to amplitude and phase ripples over the system's operating bandwidth. Typically, the amplitudes of the reflection coefficients are known, but the phase values and their

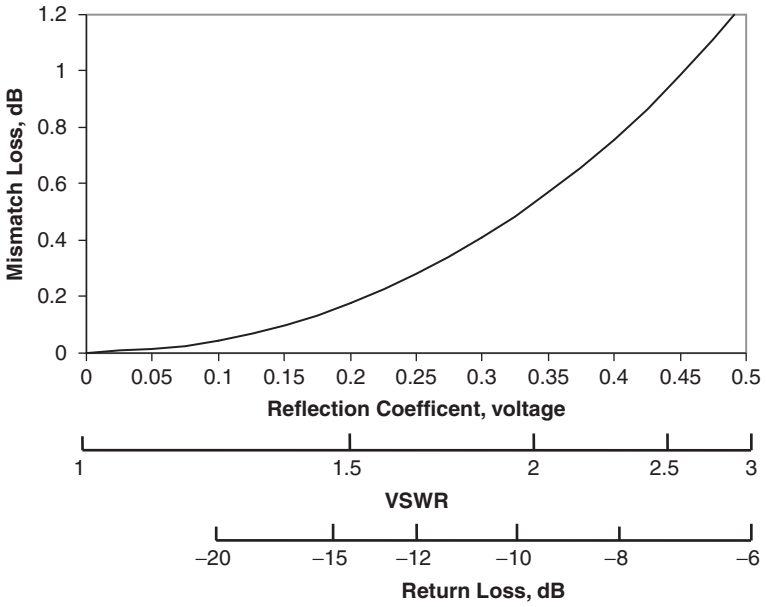


Figure 1-9 Impedance mismatch loss versus common impedance parameters

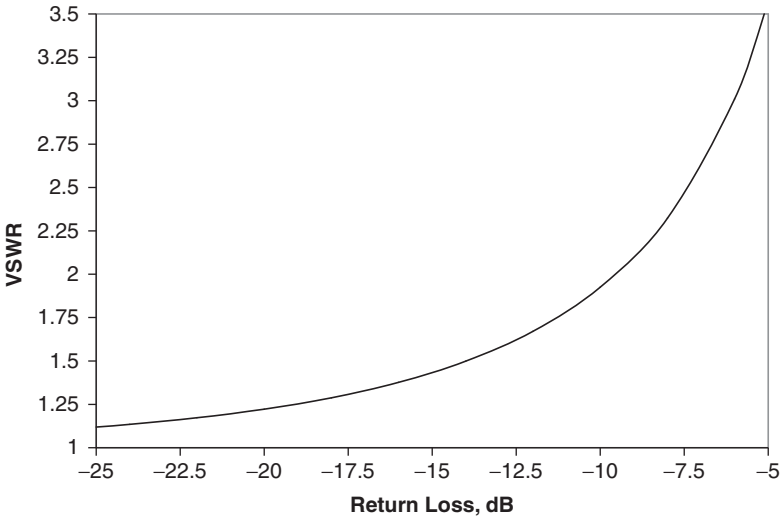


Figure 1-10 Relationship between return loss and VSWR

variation with frequency are not. Amplitude and phase ripple degrades signal detection performance and generally a specification is placed on the tolerable ripple. One approach to addressing the ripple uses coherent error statistics [5]. Other applications of these statistics,

besides assessing VSWR interactions, include assessing the effects of multipath facility reflection errors and antenna cross-polarization errors.

The coherent error statistics can be visualized by the phasor diagram in Fig. 1-11. The true value assumes a unit level, and the coherent error is represented by a phasor having an amplitude a and a phase α. Since the phase is assumed to be unknown, the error statistics are derived by assuming any phase value is equally likely and uniformly distributed between 0 and 2π. The resulting first- and second-order statistical values for power, voltage, and phase are given in Table 1-2. The statistics for power can be exactly integrated, while series expressions are derived for both voltage and phase statistics. The maximum errors for $a < 0.5$ (VSWR = 3, RL = –6 dB) are indicated and apply to most practical cases. The power and voltage statistics are non-zero mean and the peak-to-peak error bounds have finite values. The rms spread about the mean error and the peak-to-peak errors are presented in Fig. 1-12. The coherent error statistics are clearly not a Gaussian distribution and have finite bounds whose values are much less than that that would be anticipated from 3 σ confidence values for Gaussian statistics. In the case of limiting the amount of amplitude ripple resulting from VSWR interactions, the product of the return loss values for both impedance interfaces must be less than values indicated in Fig. 1-13. For example, if the ripple is to be less than 1 dB, the product of the return loss values must be less than about –25 dB. In some cases (e.g., a cable run from an LNA output to a downconverter input), the addition of loss can be advantageous in reducing the ripple and the product of the return loss

TABLE 1-2 Coherent Error Statistics

	Power	Voltage	Phase
Mean	$1 + a^2$	$\approx 1 + a^2/4$	0
Error a < 0.5	N/A	1.6%	N/A
Standard Deviation	$2^{½}a$	$\approx a/(2^{½})(1 - 3a^2/16)^{½}$	$\approx (a^2/2 + a^4/8)^{½}$
Error a < 0.5	N/A	0.1%	0.5%

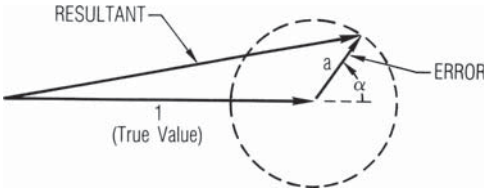


Figure 1-11 Phasor diagram for coherent errors [5] (©1989 IEEE)

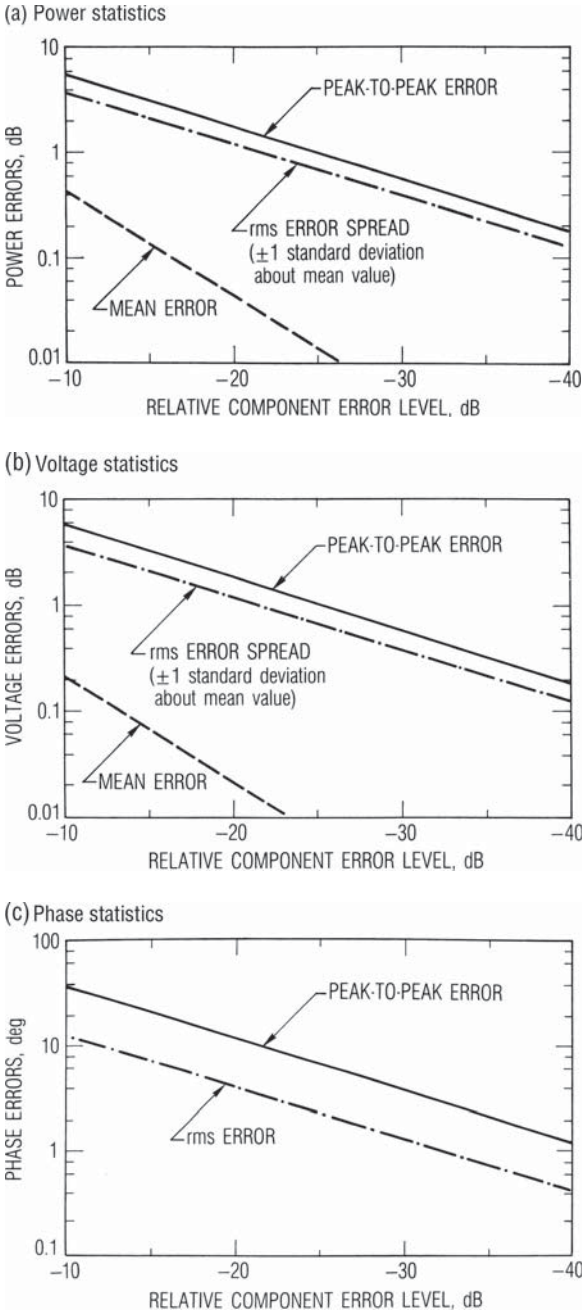


Figure 1-12 Coherent error statistical values [5] (©1989 IEEE)

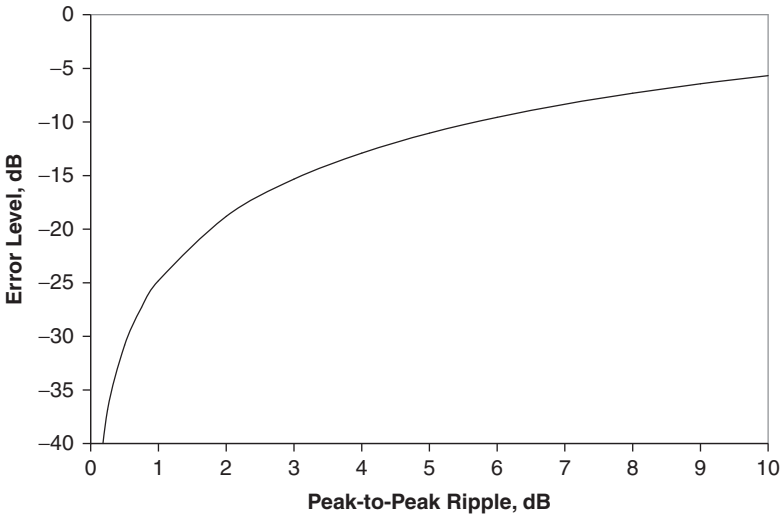


Figure 1-13 Error component level versus peak-to-peak ripple

values is increased by twice the attenuation value because the VSWR interaction component incurs a two-way path through the attenuation. The insertion of attenuators when measuring antennas that have significant mismatch is commonly done to reduce errors resulting from VSWR interactions. To be effective, the attenuators must have a good impedance match.

1.2.4 System Noise Temperature

The system figure of merit for receiving antennas is G/T, the antenna gain divided by the total system noise temperature T. The system noise temperature has two components. The antenna noise temperature includes noise components from the environment surrounding the antenna and the noise generated by losses within the antenna. A common reference terminal must be specified where antenna gain and system noise figures are both established. G/T is independent of the location of the reference terminal plane, but both antenna gain and system noise temperature values vary with the location of the terminal plane used for G/T determination. For example, the input terminal of an LNA is often convenient in measuring the receiver noise temperature. Generally, cabling and RF filtering follow the antenna terminals where the antenna gain values have been established. The losses in such components must be used to adjust the antenna gain value and antenna noise temperature so that the antenna gain level is referenced to the LNA's input terminal. Alternatively, the receiver noise temperature can

be measured, including the filter and cabling loss, so that the G/T is determined at the antenna terminal.

The antenna noise temperature can be measured as described in Chapter 8 or calculated in the following way. When calculating the antenna noise temperature, the antenna is initially assumed to be lossless, and the noise temperature is calculated from

$$T_{\text{ant}}' = \iint P(\theta - \theta_o, \varphi - \varphi_o)\, T_e(\theta, \varphi) \sin\theta\, d\theta\, d\varphi$$

where $P(\theta - \theta_o, \varphi - \varphi_o)$ is the power pattern of the antenna pointed with its beam maxima in the direction θ_o, φ_o and normalized to a unity value at the beam peak, $T_e(\theta, \varphi)$ is the emission background temperature that is described in further detail in Chapter 3. The power pattern includes both principal and cross-polarized components. An example of such a calculation for a reflector antenna is given in Fig. 1-14. The antenna is an 8-ft reflector fed with a low-loss conical horn illuminating a Cassegrain subreflector at a frequency of 11 GHz. The analysis was performed using NEC REF [7], a geometrical theory of diffraction code. The measurements and analytic results agree well.

The example in Fig. 1-14 has a low-loss feed that contributes little to the antenna noise temperature. However, practical antennas generally have losses in the antenna feed, the interconnecting cabling, and the filters are

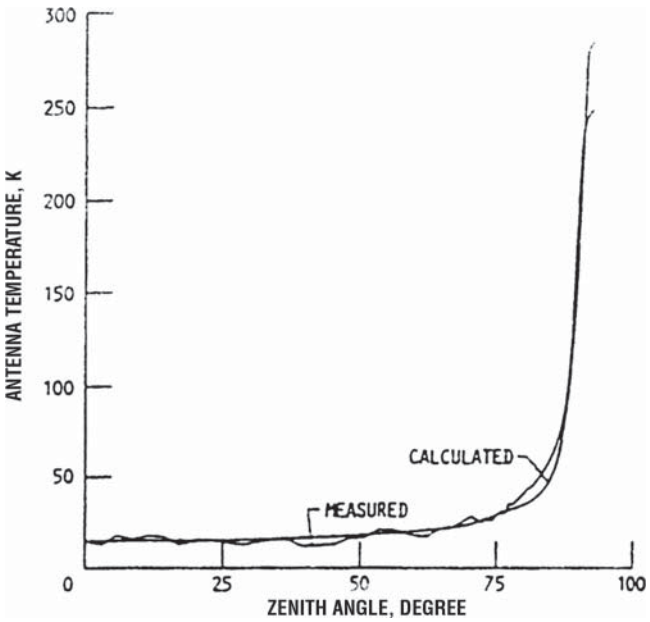


Figure 1-14 Example antenna temperature values [6]

required to limit the signal spectra to the operating bandwidth. The lossless antenna temperature [8] is then corrected for the noise contributed by these losses. The antenna noise temperature at the output of the filter and input to the receiver's LNA is given by

$$T_{\text{ant}} = (1 - \Gamma^2)\,[T_{\text{ant}}'L + 290(1 - L)]$$

where Γ is the magnitude of the reflection coefficient [for well-matched systems, $(1 - \Gamma^2)$ is very close to 1], T_{ant}' is the antenna noise temperature at the antenna terminals, and L is the ohmic loss. Physically, the noise power received by a lossless antenna is reduced by ohmic loss, but additional noise is generated by the loss. When impedance mismatch loss is significant, the antenna noise temperature is reduced by the impedance mismatch loss $(1 - \Gamma^2)$.

The receiver noise temperature at the LNA input terminal includes contributions from the LNA and other receiver components following the LNA. The receiver noise temperature at the LNA input, taking into account the LNA and receiver components following the LNA, is the cascade noise temperature, which equals

$$T_{\text{rec}} = T_{\text{lna}} + \Sigma T_i/G_{i-1}$$

where T_i is the noise temperature of the i^{th} component in the receiver and G_{i-1} is the insertion gain at the input to that component. System designs generally strive to make the LNA noise temperature dominate the receiver noise temperature, but contributions from the other receiver components can become significant when the receiver is required to have a wide linear dynamic range. The noise performance of receiver components is generally stated in terms of noise figure, NF, which is related to the noise temperature, T_n, by

$$T_n = 290(NF - 1)$$

Noise figure is defined as the input SNR (signal-to-noise ratio) divided by the output SNR. The total input noise can have many different values depending on how the component is used. Specifying the component's noise figure to a standard 290 K reference temperature allows the noise temperature to be calculated and used in applications where the input noise temperature differs from 290 K. In this application, the antenna noise temperature generally differs from 290 K, for example. Numerical values of the conversion between noise figure and noise temperature are given in Fig. 1-15.

The system noise temperature is the sum of the antenna and receiver noise temperature and equals

$$T = T_{\text{ant}} + T_{\text{rec}}$$

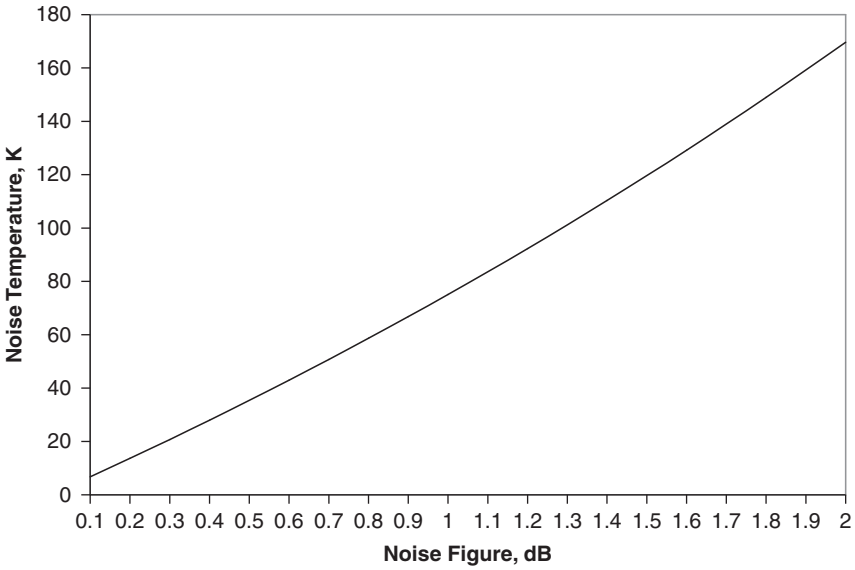


Figure 1-15 Noise temperature versus noise figure

The system noise temperature depends on the terminals in the system used for the specification and both the antenna and receiver noise temperature values must be specified for the same terminal.

In communication satellite applications, the space segment antennas for uplink receiving service view the earth's surface and generally have an ambient 290 K antenna temperature. User segment antennas look out towards space but incur atmospheric loss as discussed in Chapter 4, and as illustrated in Fig. 1-14, their antenna noise temperature varies with elevation angle and frequency of operation. Crosslink antennas that provide connectivity between satellites view a 3 K cosmic background temperature. Receiving system performance, however, depends on the system noise temperature. Today's LNA technology offers very low noise performance. However, the system noise performance can be dominated by the antenna noise temperature. This is illustrated in Fig. 1-16 where an ideal noiseless receiver is used as a reference value. The loss in receiving sensitivity or system temperature for various antenna noise temperature values is parametrically plotted as a function of receiver noise temperature. As the antenna noise temperature increases, the loss in sensitivity is less sensitive to the receiver noise temperature. Thus, low noise receiver temperatures can be effectively used if a low antenna noise temperature exists, whereas higher antenna noise temperature values derive less benefit from low noise receiver technology.

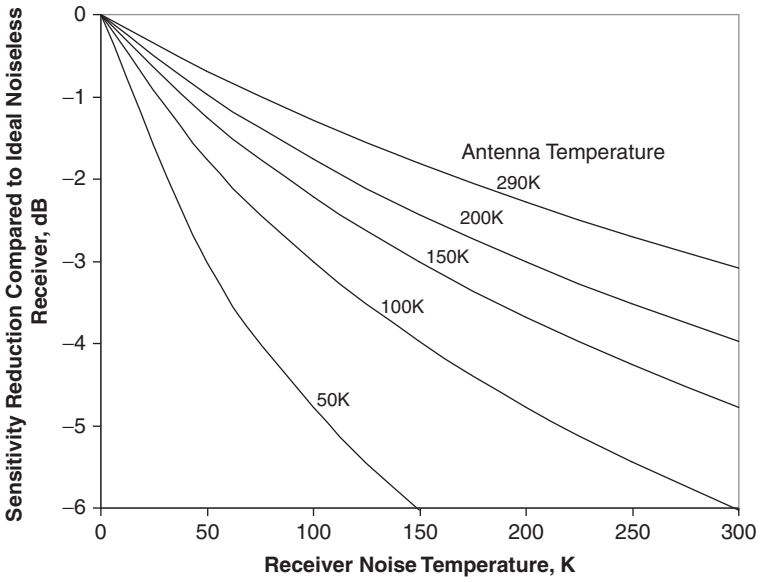


Figure 1-16 Sensitivity reduction from receiver noise temperature

1.2.5 System Parameters

The system parameters are *G*/*T* for receiving systems and ERP (Effective Radiated Power) for transmitting systems. The values of these parameters do not vary with the system terminals at which they are determined but their specification must use a consistent set of terminals. Examples will be used to illustrate the determination of these system parameters.

The *G*/*T* of an example system to operate at 10 GHz assumes a 4-ft antenna is used with a 0.5 dB noise figure receiver and a 0.5 dB loss exists between the antenna and receiver because of a bandpass filter and interconnection loss. The antenna is assumed to have a 55% efficiency, so that at 10 GHz the antenna has a 40.6 wavelength diameter, a gain value of 39.5 dBi, and a beamwidth of 1.7^{o} if a beamwidth factor value of 70 is assumed. The input terminals for *G*/*T* determination are the LNA input. The antenna gain level after the filter loss therefore equals 39.0 dBi. The antenna system specification references the *G*/*T* value to a 20^{o} elevation angle where the antenna noise temperature at the antenna terminals is assumed to be 30 K. The antenna noise temperature at the LNA input terminal after the 0.5 dB loss equals 58.2 K. The receiver noise temperature corresponding to the 0.5 dB noise figure equals 35.4 K so that the system noise temperature equals 93.6 K or 19.7 dBK (dB referenced to 1 K). The *G*/*T* value is obtained from

the antenna gain and system noise temperature referenced to the LNA input terminal. The G/T at the specified 20^{o} elevation angle thus equals 19.3 dBi/K.

In the preceding example, the loss not only reduces the antenna gain value but also increases the system noise temperature. In the design of user receiving systems where the antenna noise temperature values are generally significantly lower than 290 K, attention to reducing system loss is important in achieving good G/T performance. For space segment antennas where the antenna noise temperature is 290 K, the antenna noise temperature remains 290 K after loss and the G/T of uplink receiving antennas is reduced only by the loss. At VHF and UHF frequencies, the situation for user receiving antennas differs. As discussed in Chapter 4, the antenna noise temperature greatly exceeds 290 K and dominates the system noise temperature. The expression for G/T with loss can be written as

$$G/T = GL/(T_{\text{ant}}L + 290(1 - L) + T_{\text{rec}})$$

$$\approx G/T_{\text{ant}}$$

The G/T for high antenna noise temperature values becomes independent of the loss. Physically, because the antenna noise temperature greatly dominates the system noise temperature, the loss reduces signal and noise equally so that the system's G/T remains unaffected. When operating in the VHF and UHF frequencies where the antenna noise temperature can be several thousand K, the loss in filters needed at these frequencies does not reduce system sensitivity.

The determination of ERP is also illustrated by an example. Suppose the same 4-ft antenna used in the previous example is connected to a 10 W transmitter. The loss between the transmitter and antenna is assumed to be 1 dB. The 10 W transmitter power level is its saturated power output (discussed in Chapter 3) and must be backed off to satisfy linearity requirements. In this case, a 3 dB backoff is assumed to be adequate and so the transmitted power output at this design operating point is 5 W or 7 dBW (dB relative to a 1 W level). The ERP in this case equals 45.5 dBW.

References

1. S. Silver, *Microwave Antenna Theory and Design* (New York: McGraw-Hill, 1949).
2. ———, *IEEE Standard Test Procedures for Antennas* (New York: IEEE,1979).
3. R. B. Dybdal, "Polarization Efficiency Statistics," *1999 IEEE MILCOM Symposium Digest* (November 1999).
4. D. M. Pozar and S. Targonski, "Axial Ratio of Circularly Polarized Antennas with Amplitude and Phase Errors," *IEEE AP-S Magazine* (October 1990): 45–46.
5. R. B. Dybdal and R. H. Ott, "Coherent RF Error Statistics," *IEEE Trans Microwave Theory and Techniques*, vol. MTT-34 (December 1986): 1413–1419.

6. K. M. Lambert and R. C. Rudduck, "Calculation and Verification of Antenna Temperature for Earth-based Reflector Antennas," *Radio Science*, vol. 27, (January–February 1992): 23–30.
7. T. H. Lee, R. C. Rudduck, and K. M. Lambert, "Pattern Measurements of Reflector Antennas in the Compact Range and Validation with Computer Code Simulation," *IEEE Trans Antennas and Propagation*, vol. AP- 38 (June 1990): 889–895.
8. R. E. Collin in R. E. Collin and F. J. Zucker (eds), *Antenna Theory*, Chap 4 (New York: McGraw-Hill, 1969).

Chapter

2

Technology Survey

2.1 Overview

The antenna technology used for communication satellite systems is quite varied because of differing program requirements and frequencies of operation. As a result, antenna designs are arguably the most diverse technology used in satellite communication systems. Future antenna systems can be anticipated to become even more diverse as the complexity of system requirements continues to increase and as antennas become more integrated with RF (radio frequency) electronics. This diversity of antenna technology presents a problem in this discussion. A detailed comprehensive description of existing antenna designs would be of impractical length and would soon become outdated as technology is developed for future system designs. In many cases, the detailed designs are proprietary to the system developers, and in other cases the design descriptions are subject to ITAR (International Traffic in Arms Regulations) restrictions. Furthermore, the varied system requirements that result in the existing design diversity prevent applying designs for one application directly to other applications. Future antenna designs, while drawing on the existing technology base, will continue to evolve as requirements for future programs and applications are defined.

An overview of the generic antenna technology used in communication satellite applications is provided. The discussion describes technology as their beamwidth values decrease and aperture sizes increase. Wide coverage antennas are required in applications such as TT&C (Telemetry, Tracking, and Control) services that are used on every satellite. Other applications exist for wide coverage antennas, particularly for satellites in low earth orbits. Earth coverage antennas are also a common requirement and one that will have additional future application to allow users to access satellite resources irrespective of their

location within the available field of view. Aperture antennas providing higher gain levels are the backbone technology for both the space and user segments. Array antenna technology also has its applications for both space and ground segments. Finally, antenna tracking techniques that spatially align the antennas with the signal source are used in both space and user segment antenna designs. Further discussion of antenna technology applications is contained in Chapters 6 and 7 for space segment and user segment technologies, respectively.

2.2 Wide Coverage Antennas

Wide coverage space segment antennas are required in applications for TT&C subsystems and satellites having low altitude orbits. Wide coverage antennas are needed to satisfy TT&C capabilities required by every satellite. Other applications for wide coverage antennas arise particularly for low earth orbiting satellites such as the polar satellites used in meteorological remote sensing programs. Wide coverage antennas are also used in user segment designs for low data rate applications. The interactions of such antennas with their surrounding environment are likewise a well-known problem. A familiar example is GPS user antennas, where multipath components are one limitation on the accuracy of navigational solutions.

A principal design issue for wide coverage antennas is the ability to maintain their wide coverage characteristics in the presence of the surrounding environment. For example, the wide coverage achieved by space segment antennas in a free space environment unavoidably illuminates the spacecraft structure. The radiated scattering from the spacecraft structure degrades the wide coverage characteristics of the antenna. The challenge is to isolate the antenna from the spacecraft structure. The free space pattern of the antenna itself differs from the antenna's pattern when the antenna is located on the spacecraft. The scattering from the spacecraft structure results in pattern ripples and cross-polarization components that distort and degrade the pattern of the antenna in a free space environment.

A common application for wide coverage antennas results from the requirements imposed by TT&C applications that are required by every satellite. Often, these TT&C subsystems operate at lower microwave frequencies that have relatively long wavelengths. In other cases, such wide coverage antennas are used during launch operations, and on-orbit, the TT&C functions use the payload antennas that operate at higher frequencies. This operation using higher microwave frequencies can be anticipated in future designs as software becomes more heavily used in space segment designs. Because future space segment designs will need to upload software revisions and augmentations, these designs will require higher data

rate services than are necessary for commanding and health and status services. The general requirements for TT&C antennas during launch operations are to provide complete spherical coverage to allow access to the TT&C subsystem regardless of the satellite's orientation, to use circular polarization with sufficient polarization purity to avoid excessive signal loss and allow reliable antenna tracking with closed-loop tracking systems, and to have a compact size providing both uplink and downlink operation. In addition, the reliability of the TT&C subsystem is more critical than that of all the other subsystems in the satellite to ensure that the satellite can be controlled over its lifetime.

TT&C antennas have two coverage requirements. The launch phase of the program imposes a requirement to provide coverage over the complete sphere so that if the satellite starts to tumble or has other launch anomalies, commands for correction, or worse yet destruction, can be injected irrespective of the satellite's orientation. Adequate coverage over the complete sphere requires more than one antenna because of the inherent blockage of the spacecraft. Typically, two separate antennas are used that provide coverage over independent hemispheres, a "fore antenna" that is earth facing and an "aft antenna" that is facing away from the earth. A second coverage requirement exists when the spacecraft is established in its desired orbital position. In this case, the requirement is to provide coverage only over the earth's field of view. While such a coverage requirement can be satisfied by an earth coverage antenna having a higher gain level than a hemispheric coverage antenna, the cost and additional complexity and reliability issues of another antenna and associated switching circuitry are generally not warranted. Furthermore, the low data rates typically used for TT&C applications together with ground terminals configured for large signal margins result in a decision simply to use the earth-facing TT&C antenna.

During the launch trajectory, access to both hemispheric antennas is required. One approach assigns different frequency allocations during launch where one frequency is assigned to the fore antenna and the second to the aft antenna. In this way, the required spherical coverage is achieved by independent telemetry subsystems. This approach requires separate telemetry transponders for each operating frequency; however, such an approach advantageously provides a redundant transponder for reliability over the satellite's lifetime. In operation, the mission ground link uses both telemetry frequencies and bases the commanding on the system having the highest signal level. Another approach uses switching techniques to select the appropriate antenna depending on the vehicle's orientation. Received signal strength is indicated by the TT&C transponders' AGC (Automatic Gain Control) levels. If the AGC level of one transponder is lower than a predetermined threshold value,

the TT&C transponder is switched to the second antenna, where the received signal level should exceed the threshold signal level because higher antenna gain is available for that signal direction.

In many instances, proposals are made simply to combine the two antennas and route their combined signal to the TT&C transponder. However, the physical separation of the two antennas on the fore and aft sides of the satellite and their wide pattern characteristics results in a grating lobe structure where the two fore and aft antenna patterns overlap. The coherent combination of two widely separated antennas results in addition and subtraction of the two overlapping antenna patterns, creating a set of peaks and nulls as the signal direction changes. The two antennas, when coherently combined, form a two-element antenna array, and the separation between the two elements produces a lobe structure resulting from the addition and subtraction of the two antenna patterns. The resulting lobe structure is referred to as a grating lobe spectra. The different uplink and downlink frequencies result in grating lobe structures that generally do not overlap (i.e., the grating lobe peaks of the uplink frequency generally do not coincide with those of the downlink frequency). The resulting patterns significantly degrade the subsystem's antenna gain coverage. The uplink pattern can be aligned to a peak level providing good uplink performance while the downlink could be aligned to a pattern null, significantly degrading downlink signal reception. Similarly, the uplink could be in a pattern null while the downlink is in a pattern peak. Coherently combining the fore and aft antennas is specifically not recommended for this reason.

The principal development problem is isolating the TT&C antennas from the satellite to achieve the spherical coverage needed during the launch phase. The complex structure of a typical satellite results in blockage and diffraction mechanisms that degrade hemispheric coverage requirements. Consequently, a mounting location with a clear field of view over a hemisphere is typically difficult to find with the large, complex satellite structures used today. The interaction with the satellite can be examined either analytically with diffraction codes [1] or by measurement. Neither approach is a simple undertaking. The analytic challenge is modeling a complex structure and validating the results. In practice, satellites are covered with thermal blanket material that may not be smooth at the TT&C frequency of operation, which adds to the modeling challenge. Direct measurement of the antennas mounted on the spacecraft is generally impractical. Because the radiation arises from both the antenna and the spacecraft structure, excessive far field distances are involved. Such measurements would also need to be performed during spacecraft assembly and concerns of risk to flight hardware would likely preclude such measurements. Other features of the spacecraft that impact antenna coverage, such as solar arrays,

generally cannot be deployed for measurement. Finally, measurements of the vehicle mounted antennas over the spherical volume required in this application are generally impractical.

Clearly, a need exists to develop validated techniques to establish the TT&C antenna coverage characteristics. One approach might use scale model techniques and measurement at a scaled frequency. Scale model measurements are based on the theorem of similitude [2]. The dimensions of a scale model preserve the electrical size of actual design in wavelengths and the measurement is performed at the scaled frequency. In this application, a scale factor of at least 10 is required (i.e., a 1/10 scale model of the satellite is measured at ten times the operational frequency). The smaller scaled test article is required to result in a potentially practical measurement.

When the satellite reaches its on-orbit position, the antenna covering the earth-facing hemisphere can be used. In theory, somewhat higher gain performance can be obtained by switching to an antenna matched to the earth's field of view for the operational phase of the satellite. In practice, the additional complexities of two sets of antennas and the reliability of necessary switches limit the attractiveness of this approach. Moreover, TT&C systems are purposely sized to achieve large margins for reliable communications in the event of performance shortfalls in the space segment. Thus, the same antenna is generally used for both launch and on-orbit operations.

The existing TT&C frequency allocations at low microwave frequencies are often separated by a large percentage bandwidth, and consequently, simple frequency independent antennas [3] such as spirals are commonly used. While these designs are simple and relatively lightweight, pattern control beyond the required coverage area with a compact design is difficult to achieve for the existing low microwave frequencies and relatively high backlobe levels are common for these designs. These high backlobes illuminate not only the undesired coverage hemisphere reducing the antenna's gain in the desired hemisphere but also the satellite structure. The resulting reflected and diffracted components from the satellite structure in the desired coverage hemisphere degrade minimum gain values. The irregular shapes in the satellite structure and reflections from smooth surfaces also result in cross-polarized components. These cross-polarized components can result in regions of the desired hemispheric coverage where the sense of the signal's polarization is reversed. Closed-loop antenna tracking techniques can result in tracking shortfalls in situations where such polarization diversity exists, as will be discussed later in this chapter. Further, the polarization diversity in the desired coverage hemisphere requires diversity combining techniques to maintain signal reception.

A compact antenna having good polarization purity, sufficient bandwidth, and low backlobes is desired for TT&C applications. One design approach is the rolled edge cavity antenna [4] shown in Fig. 2-1. This design uses a cross dipole in a cavity whose edges are rolled in a semicircular shape. A variety of techniques can be used to increase the bandwidth of the dipoles such as bowtie configurations and the sleeve loading used in this design. The cavity extends to screwheads visible below the rolled edge terminal and the remaining portion of the cylindrical housing was used to enclose the electronics used in the antenna's application. The overall electrical diameter including the rolled edge is about 1.2 wavelengths. The calculated patterns in Fig. 2-2 were performed using a commercial HFSS (high frequency structural simulator) finite element modeling code and reveal low backlobe levels producing high gain levels in the desired forward hemisphere. The good pattern symmetry and low cross-polarized pattern levels are necessary to achieve good axial ratio performance in the forward hemisphere. The measured patterns [4] closely agree with these calculated results.

Figure 2-1 Rolled edge cavity antenna [4] (© 2004 IEEE)

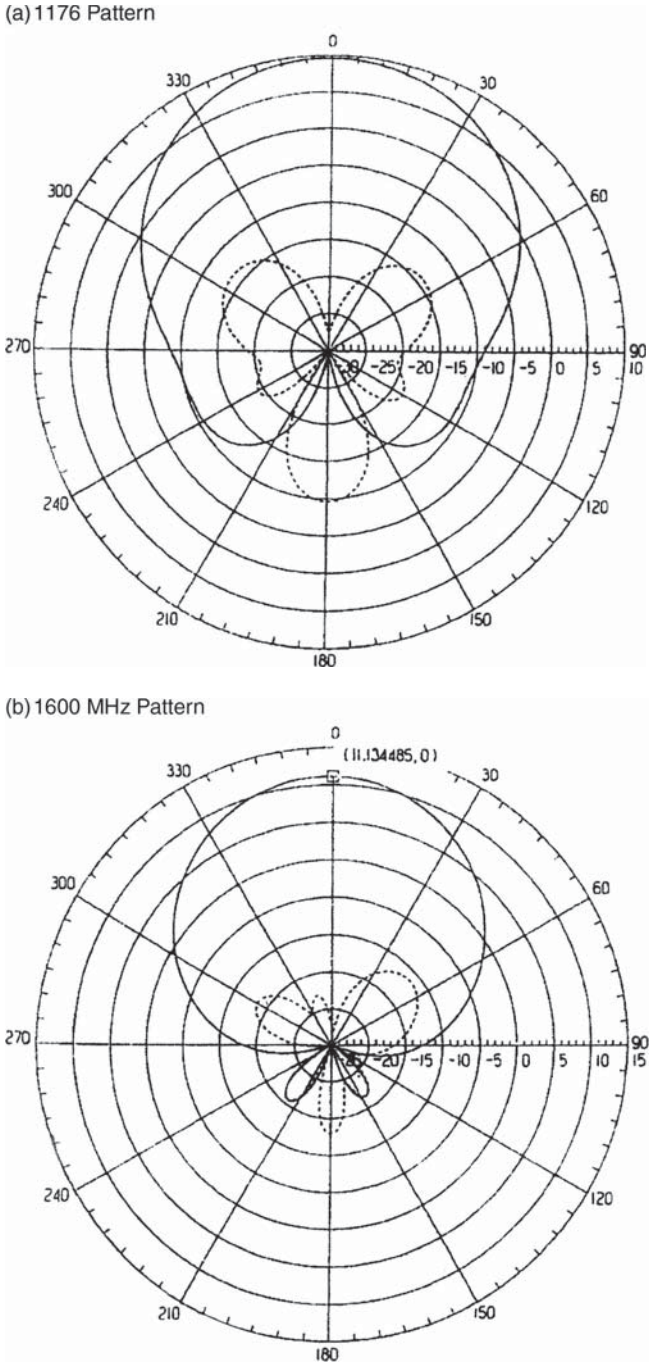


Figure 2-2 Calculated principal plane patterns of rolled edge cavity antenna [4] (© 2004 IEEE)

Two recommendations are made for future development of wide coverage antennas. The first recommendation is to further develop compact designs with reduced backlobe performance so that the illumination of satellite structural components is minimized both to maintain the antenna's free space coverage and to increase the antenna's directivity in the desired hemisphere. The second recommendation is further development and validation of techniques to project and validate the antenna's performance when mounted on the satellite. It is envisioned that a combination of measurement techniques and analyses could be developed and validated.

2.3 Earth Coverage Antennas

The earliest satellite antenna designs provided coverage over the entire field of view subtended by the earth together with the attitude uncertainty of the satellite. Requirements for full earth coverage still exist to communicate with users distributed throughout the field of view. Such designs offer access and broadcast to users within the earth's field of view. Applications for such coverage requirements continue to exist. For example, user access and satellite resource assignment broadcast require coverage to users throughout the satellite's field of view. Early satellites at low altitudes had primitive attitude control capabilities and used simple monopole antennas to provide very broad coverage. The low altitude and operating frequency used in early satellites has been supplanted by geosynchronous altitude satellites and greatly improved attitude stability. Geosynchronous satellites have a 16.9^{o} field of view subtended by the earth at their 20,000 nmi altitude above the earth. Today's satellites provide attitude stability on the order of 0.1^{o}. Simple horn antennas are commonly used for such earth coverage requirements and their beamwidths are selected to conform to the angle subtended by the earth with an additional margin to compensate for the satellite's attitude variations. Earth coverage antennas require a rotationally symmetric pattern to obtain not only efficient coverage but also to achieve low axial ratio performance over the earth's field of view. Low sidelobes beyond the earth's field of view increase the directivity of the antenna and thus the overall antenna efficiency. These general requirements are commonly satisfied by dual mode or corrugated horn designs [5, 6].

Since earth coverage antennas are used for both uplink and downlink requirements, a tradeoff exists between using a single antenna and diplexer to separate receive uplink and transmit downlink functions and using separate antennas for uplink and downlink frequencies. Physical separation between uplink and downlink antennas provides part of the required isolation between the transmitter and receiver that must be accommodated by a diplexer discussed in Chapter 3. Separate earth coverage antennas are attractive because the additional weight of

a second antenna is minimal and the transmit and receive filtering loss to isolate the receiver and transmitter is typically less than a diplexer for the single antenna design because of the isolation provided by the physical separation.

Horn antennas have a variety of other satellite communication applications such as antenna feeds for reflector antennas and an antenna standard for measurements. A useful compendium of horn technology [7] describes the diverse design approaches that have been used. The development of horn antennas continues and design parameters such as throat profiles [8] for wide bandwidth operation are explored to extend the horn technology capabilities. Another approach uses rolled edges [9, 10] to reduce the antenna's sidelobes and produce patterns with increased rotational symmetry.

An example of a rolled edge horn shown in Fig. 2-3 uses a horn antenna in a system application to receive low data rate meteorological satellite signals [11] from polar orbiting satellites. The overall system requirements for this application include a compact, efficient terminal design with low power consumption for transportable applications. The system includes an ASIC-based receiver located on top of the horn's input waveguide with the input bandpass filter and LNA located beneath the input waveguide. A simple elevation over azimuth positioner driven by DC motors positions the antenna. A GPS antenna mounted on the top of the elevation platform provides location and time-of-day information, and a compass and tilt angle sensor on the azimuth platform is provided to compensate motion if the system is mounted on a mobile platform providing an "image on the move" capability. In operation, the power consumption for the system is 18 W. A laptop computer provides system control, data processing, image display, and data storage.

Figure 2-3 Rolled edge horn used in transportable readout terminal [11]

Applications such as this one often use a small reflector antenna. However, because of feed blockage, such electrically small reflector antennas generally have low antenna efficiency values. Horn antennas provide a more efficient and compact design for this application. The addition of the rolled edges to the aperture results in low sidelobe, rotational symmetric patterns, as illustrated in Fig. 2-4. The measured

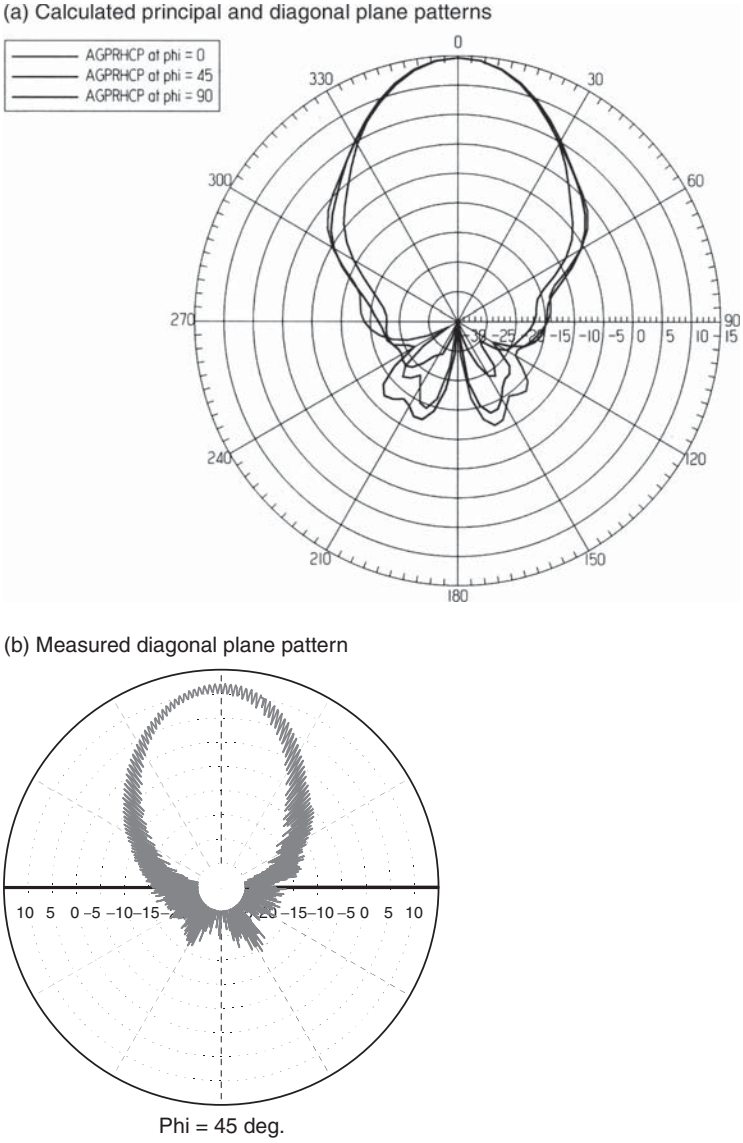


Figure 2-4 Rolled edge horn patterns [10] (© 1982 IEEE)

patterns were performed by illuminating the horn with a test source that mechanically rotates a linearly polarized test signal. The resulting "ripple" on the pattern indicates the extreme variations of the polarization ellipse as the rotating polarization of the illumination source varies during the angular changes as the pattern is measured. In this way, the pattern measurements display the antenna's axial ratio variation.

The measured patterns in Fig. 2-4 agree well with calculated results from HFSS modeling that was used to optimize the horn's design parameters. The antenna also uses a septum polarizer to produce the required circular polarized patterns. As discussed in Chapter 1, attention to RF loss is necessary to maintain high G/T performance. The ohmic loss of the horn and polarizer was measured to be 0.1 dB using zenith radio source techniques. This measured loss value was confirmed by using the HFSS model to determine the antenna's gain difference for a perfectly conducting horn and a horn employing the conductivity of the aluminum used to fabricate the horn.

2.4 Narrow Coverage Antennas

Antennas for narrow coverage requirements for space segment designs and high gain requirements for user segment antennas predominately use reflector antenna technology. Reflector antennas are commonly associated with communication satellite systems both from photographs of satellites and from user experience with applications such as direct broadcast services where offset reflector "dish" antennas are widely familiar to system users. Reflector antennas have a variety of advantages. The cost of reflector antennas is minimal and a variety of fabrication techniques are available to produce high-quality surfaces. The reflector antennas are based on geometric optics principles. The parabolic shape has the property that any path from the reflector's focal point that is reflected by the parabolic surface has the same length to a projected planar aperture surface. This property of parabolic shapes results in a projected aperture surface that has equal time delays between the antenna feed located at the focal point and the planar focal surface. As a consequence of this true time delay variation, the antenna feed rather than the reflector itself limits the antenna's bandwidth. Operation over broad bandwidths is possible, and operation at multiple frequency bands by physically locating or optically imaging the feed to the focal point is commonly done to satisfy communication satellite requirements. The bandwidth of practical reflector antennas is limited by the bandwidth of the antenna's feed system. The most common reflector antenna feed uses horn antenna technology. The reflector's shape is defined by the f/D value, which is defined by the focal length f and the reflector diameter D. Increasing the f/D value results in

a flatter reflector surface, but one that requires a more directive feed pattern to illuminate the reflector surface properly.

A variety of reflector designs are used as described in a useful compendium [12]. The most popular reflector design is a prime focus configuration where the reflector's antenna feed is placed at the focal point. One limitation is the reduction of the antenna's efficiency because the feed blocks a portion of the aperture field. A first-order estimate of feed blockage loss assumes the fields obscured by the feed result in an aperture distribution having zero fields in the area blocked by the feed. The blockage loss then becomes the ratio of the aperture area reduced by the area of the blockage divided the aperture area. For circular blockage having a diameter d_b, the blockage loss L_b equals

$$L_b = 1 - (d_b/D)^2$$

where D is the reflector diameter. The resulting pattern can be viewed as the superposition of two patterns, one being that of the unblocked aperture and the second being a pattern corresponding to the blocked portion of the aperture that is out phase with the first pattern so that the aperture fields are zero in the blocked area. The gain of resulting pattern is reduced by the out-of-phase second pattern. The second pattern being derived from a smaller aperture representing the blockage is much broader than the unblocked pattern that has a larger aperture than the blocked area. The superposition of the two patterns results in variations of the antenna's sidelobe levels. The phase of the sidelobes relative to the main beam's phase alternate (e.g., the first sidelobe from the main beam is out of phase from the main beam, the second sidelobe from the main beam is in phase with the main beam, and so forth). When combined with the broad pattern of the blocked area, the sidelobes of the overall antenna pattern are perturbed as the pattern of the unblocked aperture adds or subtracts from the pattern representing the block area. Blockage not only reduces the antenna's gain but also perturbs the antenna's sidelobe levels near the main beam.

Feed blockage effects can be avoided by using an offset reflector design that only uses a portion of the reflector surface. In this way, the antenna's feed system located at the focal point is positioned so that the aperture fields are not blocked by the antenna's feed. As a result, the antenna efficiency of offset reflector designs is not reduced by blockage loss, but since only a portion of the parabola is illuminated by the antenna feed, a larger, more directive feed design is required in comparison to the prime focus geometry. A very familiar offset reflector application is the user terminals for direct broadcast services. Another example is illustrated in Fig. 8-3, where the near field of an offset reflector is used to produce a uniform test field for antenna measurements.

Larger, higher gain antennas often use dual reflector antenna designs based on optical telescope designs. Such designs use a subreflector to image the physical feed into the focal point. The subreflector must have a diameter greater than at least ten wavelengths to image the reflector's feed into the focal point effectively. Smaller subreflectors result in excessive spillover loss [13] and the feed contributions do not effectively illuminate the radiating reflector. The blockage from the subreflector would be excessive unless the reflector gain exceeds 40 dBi. Consequently, dual reflector antennas [14] are used in high-gain applications. The most common dual reflector design is the Cassegrain configuration where the subreflector having a hyperbolic shape is located between the reflector's focus and the radiating reflector. A Gregorian design is used less often and the subreflector has an elliptical shape and is located beyond the parabola's focal point. One advantage of these designs results in transmitting antennas. The feed system location allows a relatively short waveguide run between the transmitter and antenna feed. Prime focus designs use a transmission line that generally runs along a support strut for the feed system and along the antenna's radius on the back side of the reflector. This relatively long transmission line becomes increasingly lossy as the frequency increases. The shorter transmission line of dual reflector designs reduces transmission line loss, improving the overall transmission efficiency. Receiving electronics are compact and require little prime power so that they can be located with the antenna's feed. Transmitters, however, are larger and consume much more prime power than receiver electronics and generally require access for servicing and dummy loads for diagnostics purposes. As a consequence, transmitters for applications requiring high-gain antennas and high transmitter levels are generally located away from the antenna's structure. The interconnection between the antenna and transmitter generally incorporates rotary joints to allow antenna positioning or, at high frequencies where their reflector sizes are compact, beam waveguide techniques that transmit the RF signals through mirrors to provide a low loss connection between the transmitter and antenna.

The efficiency of high-gain dual reflector antennas can be increased by shaping techniques [15]. As discussed in Chapter 1, aperture amplitude tapering reduces the antenna efficiency while broadening the beamwidth and lowering sidelobe levels near the main beam. The basic idea for antenna shaping techniques is to distort both the radiating reflector and subreflector in such a way that the resulting aperture amplitude distribution becomes more uniform while preserving a flat phase distribution in the aperture plane. In this way, the aperture's amplitude taper loss is reduced, increasing the antenna's efficiency. Typically, the reflector edge illumination for nonshaped designs has a value about 10 dB

below the feed's peak gain level on its axis. This illumination level [16] results in a compromise between amplitude taper loss and spillover loss that depends on the portion of the antenna feed's pattern that does not illuminate the antenna. Shaping techniques applied to both the subreflector and radiating reflector surfaces result in a more nearly uniform amplitude aperture distribution that increases the antenna's efficiency. In practice, shaped reflector antennas with aperture efficiency values in excess of 75% have been realized.

The polarization properties of reflector antennas are generally dominated by the polarization purity of the antenna's feeds. Other factors result from the depolarization produced by antenna blockage of the feed and supporting struts. Offset reflector designs avoid aperture blockage effects, but their optics [17] can result in cross-polarization generation and slightly different boresight axis locations for opposite sense circular polarization components. These effects are reduced as the reflector's f/D value increases, resulting in a flatter reflector surface. When high polarization purity is required, design attention to the antenna feed's polarization and reflector optics is needed.

In some applications, reflector antenna designs require beams that are displaced from the reflector's axis. For example, reflector antennas providing multiple beam capabilities use a cluster of feeds in the reflector's focal region. Each feed produces an independent beam whose direction depends on the feed's off-axis position. Commonly, the off-axis response is referred to as "beam scanning." The ability to produce an antenna beam from the axis requires an aperture distribution having a linear phase gradient normal to the off-axis direction. When antenna feeds are moved off-axis, the phase distribution has a linear gradient to first order. The parabola focuses ideally on-axis, but as the feed moves off-axis, the focusing is no longer ideal and the desired linear phase gradient has perturbations known as aberrations. These aberrations increase the further the beam is moved from an axial position [18], limiting the beam scanning performance of the antenna. At off-axis positions, the phase aberrations result in reduced gain, beam broadening, and increased sidelobe levels compared to axial values. Multiple beam antennas for space segment applications typically required off-axis beams with both low sidelobes and low cross-polarization levels while maintaining the on-axis efficiency performance. A combination of dual reflector technology and reflector shaping technique [19, 20] provides design techniques to achieve such performance goals. Available computer codes typically have the provision to calculate the off-axis performance of such reflector antenna designs.

Applications that require operation at multiple separated frequencies often arise. For example, commercial systems commonly use both C- and Ku-band frequencies, and accommodating both frequency bands

in a single antenna is attractive for system users. Both frequency ranges must locate their feeds in the focal region to use a single reflector. One approach uses a C-band feed in a prime focus location and images the Ku-band feed to the focus by using a subreflector. In this case, the subreflector is constructed from a frequency selective surface [21] that is transparent at the C-band frequency to allow prime focus operation and is reflective at the higher Ku-band frequency to allow operation as a Cassegrain configuration. In such applications, the higher frequency is used in a Cassegrain configuration to minimize the subreflector size and associated blockage. Other applications for multifrequency designs [22] are described where frequency selective surfaces, frequency selective reflectors, and beam combining for contour coverage are combined in novel ways to provide multifrequency satellite capabilities.

Reflector surfaces have limitations on the mechanical precision of their manufacture. Deviations from the design reflector surface degrade antenna performance since the surface imperfections result in errors in the desired aperture phase distribution. The gain loss resulting from random mechanical tolerance deviations [23] is generally determined from

$$L_{\text{Tol}} = \exp[-(4\pi\varepsilon_T/\lambda)^2]$$

where ε_T is the rms surface error and λ is the wavelength. The surface deviations are assumed to have a random variation in this analysis. Systematic deviations result in performance loss that can be addressed by examining the effect of systematic phase errors in the aperture distribution on the antenna's radiation performance. Example values for loss resulting from random phase errors in Fig. 2-5 illustrate the frequency variation of tolerance loss for a range of values. If the tolerance loss is limited to 0.1 dB, the rms error must be less than about 1/82 of

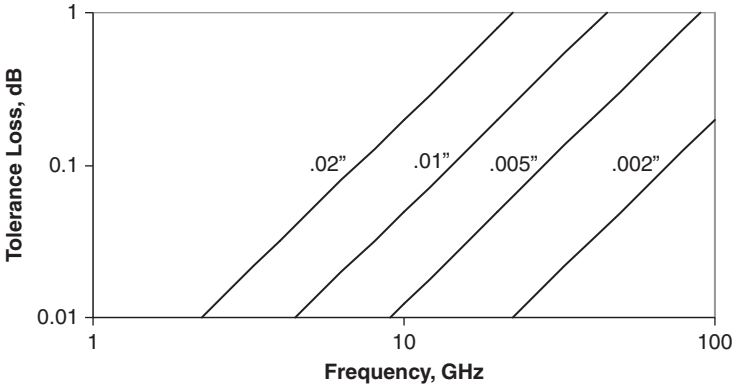


Figure 2-5 Reflector tolerance loss

a wavelength or if the tolerance loss is limited to 0.5 dB, the required tolerance is relaxed to a value of about 1/37 of a wavelength.

Space segment antenna requirements can result in aperture diameters whose value exceeds the launch vehicle's fairing dimensional limitations. In such cases, reflector antenna designs having deployable surfaces are used. These surfaces are comprised of a woven metallic mesh and a deployable structure. An early example of deployable reflector antenna technology [24] was flown on the ATS-6 satellite. A 30-ft diameter prime focus reflector was demonstrated for operation up to C-band frequencies. Development of deployable designs continues [25] and application of deployable antennas can be anticipated in future systems as narrow beamwidth space segment antennas are required to further increase link performance. Deployable antenna technology has three design issues. The first issue is the reflectivity of the mesh surface. Generally at least three openings per wavelength are required. The second issue arises in multicarrier communication applications where PIM (Passive Intermodulation) products discussed in Chapter 9 degrade communication performance. Contamination between mesh elements can produce nonlinear rectification, producing the intermodulation products. The third issue is the mechanical accuracy of the surface. The rms tolerance errors discussed earlier must be met. However, such analyses are based on random surface errors. Design attention must be paid to assure that systematic phase errors do not result from producing the surface from the mesh material.

While reflector antennas are commonly used to satisfy high gain requirements, lens antennas offer an alternative technology. Lens antennas have the advantage that the feed does not produce blockage. The inner and outer lens surfaces provide design freedom to optimize the beam scanning performance of the design. Dielectric lens designs are limited by the weight of the dielectric and the practical problem of obtaining a sufficient dielectric size with uniform dielectric properties. Dielectric lens applications are limited to designs at higher frequencies that require relatively small lens diameters. The dielectric weight can be reduced by zoning techniques. In this case, the dielectric aperture is divided into radial zones and the dielectric thickness is reduced by satisfying the aperture phase distribution on a modulo 2π sense rather than the true time delay of an unzoned lens. Zoning techniques result in bandwidth limitations since the aperture phase distributions satisfy the required uniform phase characteristics at the design frequency chosen for zone step size.

Another alternative is a waveguide lens [26, 27] where the phase velocity of waveguide functions like a dielectric. Dielectric slows the phase velocity so that the thicker surface at the center of the lens slows the wave to compensate for the longer path lengths to the edges of

the lens surface. By contrast, the phase velocity of waveguide increases the phase velocity so the lens is thicker at the edges than the center to compensate for the longer path length at the edges of the lens. Waveguide is a dispersive media since the phase velocity varies with frequency. In this case, zoning techniques can actually increase the bandwidth because shorter lengths of waveguide reduce the aperture phase variation resulting from dispersion.

2.5 Array Antennas

Array antennas [28] may be thought of as digital representations of analog aperture distributions. Their overall electrical characteristics such as beamwidth are dictated by the same considerations. The antenna efficiency of array designs depends on the specific implementation. The array gain performance equals the gain of the array element times the number of elements to first order. Array antennas, however, have fundamental limitations. Like any digital representation, the analog aperture distribution has minimum sampling requirements. To avoid aliasing, the sampling of the analog aperture distribution must be less than ½ wavelength. Aliasing in this case results in additional lobes in the array's patterns, referred to as grating lobes. The array elements combine together to produce a main beam when the array elements produce an in-phase condition in a plane normal to the main beam's direction. If the element spacing exceeds ½ wavelengths, this same phasing condition results at angles other than the main beam direction and grating lobes are produced. Practical array designs endeavor to minimize the required number of array elements to reduce design complexity, and the spacing of the array elements is often selected so that grating lobes are removed from the earth's surface. Such designs are sometimes referred to as "thinned" arrays since the number of elements is less than that of an array whose element spacing is selected to avoid grating lobes. Grating lobes reduce the directivity of the array, compromising the antenna efficiency.

The gain of an array antenna to first order equals NG_e, where N is the number of array elements and G_e is the antenna gain of an array element. The elements are combined together with amplitude coefficients that dictate the sidelobe characteristics and phase coefficients that dictate the array's beam direction. The array elements are generally identical, so in the pattern analyses the array element is separable from the array excitation coefficients that combine together to form an "array factor." This array factor is the array pattern if the array elements had isotropic gain characteristics. The overall electrical size of the array and the array element amplitude coefficients determine the array's beamwidth. The array phase coefficients are set to produce

the array beam's direction. However, the element phase coefficients are adjusted by phase shifters. These phase adjustments are adjusted on a modulo 2π basis at a design frequency rather than an analog linear phase gradient, as produced by other technologies. As a consequence, the array's beam steering has a frequency dependence, so the farther the array beam is scanned from the normal to the array surface, the faster the beam scans with frequency changes.

Array designs generally connect active elements directly to the array elements to avoid the effects of loss in phase shifters and element combining circuitry. Preamplifiers are generally connected directly to array elements to establish the system noise temperature. Array transmit modules are directly attached to the array elements so that phase shifter and combining circuitry losses do not reduce ERP levels. The ERP of an active transmit array equals $N^2 G_e P_e$, where P_e is the power output of the array's transmit modules. The factor of N^2 results because the array's antenna gain is N times the number of array elements and the transmitted power equals N times the power of a single array transmit module. The active transmit arrays can be thought of as a spatial power combiner.

The array elements are combined by a corporate feed structure whose topology resembles an organization chart. If the array is required to generate multiple beams to service different coverage areas, the corporate feed structure must be replicated to control the individual beams. Operation at separated frequency bands is difficult for array designs since the element characteristics and spacing are selected for operation at a single frequency band. Array element phase settings required for beam steering capabilities are generally produced by phase shifters. In addition to the design complexity, power consumption for array designs is much greater than reflector antenna systems. The individual active elements in the array must be powered and the phase shifters also require power for both their setting and a control system that commands their phase shift values. In addition, special thermal control designs are needed to maintain the amplitude and phase tracking requirements of the active elements in the array design. The inherent design complexity increases testing schedules. The combination of design complexity, power consumption, and thermal control limits the application attractiveness of array technology and increases the cost compared to other technologies. Consequently, array designs are generally used in those applications where requirements cannot be easily satisfied by other technologies. For example, conformal arrays in aircraft antenna applications are used to satisfy aerodynamic requirements.

Space segment applications of array antenna technology are subject to the usual SWaP (size, weight, and power) limitations of space systems.

Geosynchronous satellite antennas generally require narrow beamwidths, and an excessive number of array elements are required to achieve acceptable performance. As a consequence, high-gain space segment antenna applications are commonly satisfied by reflector antenna technology. Geosynchronous satellites have a limited field of view equal to 16.9^{o} and this field of view is compatible with the limited beam scan capabilities of reflector antenna designs. Array designs for low orbiting satellites [29, 30] are capable of providing beams over the much wider field of view that exists at low orbital altitudes than the narrower field of view for geosynchronous orbits, as discussed in Chapter 3. The wider angular field of view for low orbiting satellites results in antenna requirements for broader beamwidths and corresponding smaller antenna apertures that result in the array design having a reasonable number of elements. This is an example where array designs can be effectively used because the required field of view is incompatible with reflector antenna technology.

An advantage of array technology that is commonly touted is graceful degradation. Within reason, element failures result in a small reduction in the antenna gain level. However, element failures also degrade sidelobe performance, an important factor when isolation is to be maintained between separated coverage areas. Sidelobe degradation depends on the number of failed elements and their distribution within the array. Sidelobe degradation when element failures have a random distribution in the array is less pronounced than when adjacent array elements fail. The pattern when adjacent elements fail can be viewed as the subtraction of the array pattern of the failed elements from the array pattern with all elements functional. Since the failed elements in this case have a smaller, lower gain pattern that has a directive pattern, its subtraction from the pattern without failures results in addition and subtraction with the sidelobe structure of the array pattern without element failures. When the distribution of element failures is random, the corresponding pattern of failed elements is less directive than the pattern of failed adjacent elements and generally has grating lobes that reduce directivity. Consequently, the pattern when elements fail has a more random distribution has a less pronounced effect on the array pattern without element failures.

If the location of the failed array elements can be determined, a possibility [31] exists to rephrase the array to attempt to maintain the required sidelobe performance. The design sidelobe performance can be maintained if a small number of elements fail, but as the number of failed elements increases, the design sidelobe cannot be recovered. If this technique is used, a means of identifying failed elements on-orbit is required, a capability that can use additional development attention.

A well-known application of array technology is the antenna used in the GPS satellites [32]. The GPS satellites provide earth coverage

from 10,900 nmi orbits and use arrays of helical elements to shape the antenna pattern [33] so that users distributed over the earth's surface will receive comparable incident power density values independent of the user's elevation angle to the satellite. The GPS array design is comprised of two circumferential arrays that are combined by subtracting a small portion of the more directive outer array from the less directive inner array. The pattern level on-axis is reduced by the subtraction, while the pattern reduction at wider angles from the subsatellite direction is smaller because of the more directive pattern of the outer circumferential array. The result is a pattern variation whose gain increases towards the edge of the earth. This gain increase offsets the longer range towards the edge of the earth so that the power density users receive has comparable levels within the earth's field of view. The array technology provides a pattern shaping capability and the simple array design does not require phase shifters for beam steering and has sufficiently low loss that active elements are not required.

2.5.1 Arrays of High-Gain Antennas

Normally, array antennas are configured with relatively small antenna elements to reduce the effects of grating lobes. Because the small array elements have relatively low gain performance, a large number of elements must be combined to achieve high gain levels since the array gain to first order equals the antenna element gain multiplied by the number of elements. High gain levels require so many elements that the array design becomes impractical. An alternative approach is to combine a number of high-gain antenna elements coherently to achieve higher gain levels. One application of this approach is being explored as a replacement for 70 m reflector antennas used in JPL's Deep Space Network [34]. Another potential application is combining high-gain array elements on aircraft to achieve higher gain levels needed for high data rate applications while controlling the physical size of the array elements to satisfy aerodynamic requirements.

The coherent combination of high-gain array elements imposes several requirements. The elements must be sited with sufficient separation that array elements do not block other array elements. Like any array design, the element electronics must have matched amplitude and phase responses so that the coherent combining response is maintained over the required bandwidth. The array element separations result in signal arrival time differences at the individual array elements that require not only phase compensation but time delay compensation to maintain coherent combining over a bandwidth. The high-gain antenna elements have narrow beamwidths so that antenna tracking must be used to align the antenna elements with the signal's direction.

The required separation of the array depends on the minimum elevation angle anticipated in system operation. A simple expression can be derived if it is assumed the signals can arrive at any angle above the minimum elevation angle and that the antennas have a common diameter D and are located on a planar surface. The minimum antenna element separation in Fig. 2-6 equals $D/\sin\varepsilon$, where D is the antenna element diameter and ε is the required minimum elevation angle. For example, the required antenna separation must be at least 11.5 antenna diameters for a 5° minimum elevation angle.

The coherent combination of array elements in this case must compensate for the signal arrival time differences at each antenna element. Very narrow bandwidth signals require phase compensation, but wider bandwidth signals require both phase and time delay compensation. The tolerances for these compensation requirements can be derived by considering a two-element antenna array [35, 36] that has an ideal coherent combination efficiency that increases the peak antenna gain level by a factor of 2 (3 dB) greater than the gain of a single antenna element. The combining efficiency of two antenna elements is

$$C(\theta,\omega) = 2\ [\cos\ \{[(\omega((S/c)\ \sin\ \theta - \tau) - \alpha]/2\}]^2$$

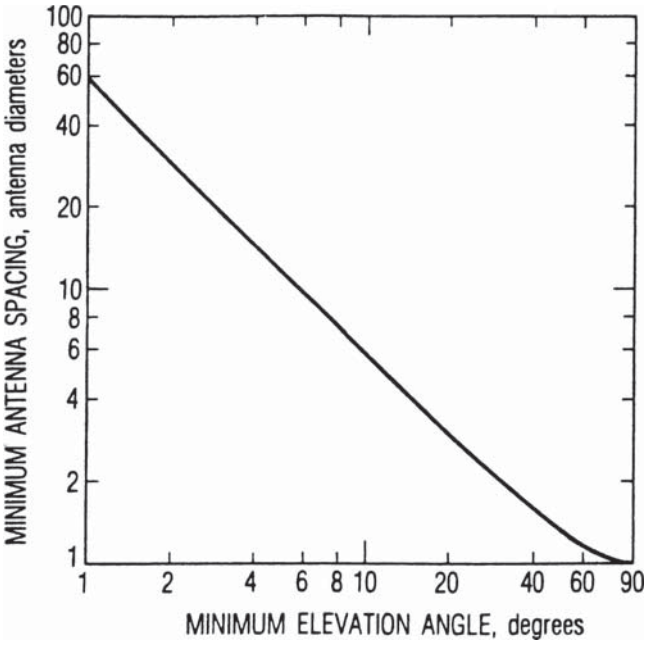


Figure 2-6 Array element separation requirements [38] (© 2008 IEEE)

where θ is the signal direction, S is the separation (baseline) between antenna elements, ω is the radian frequency, τ is the time delay adjustment in the array combining circuitry, and α is the insertion phase difference between the antenna elements at a center frequency. The adjustment tolerances can be expressed as an uncompensated time delay, $\Delta\tau = (S/c) \sin \theta - \tau$, and an uncompensated phase $\delta\phi = \omega_0\Delta\tau - \alpha$ at the center frequency. These tolerances are the deviations from ideal compensation. The combining efficiency becomes

$$C = 1 + \cos(\delta\omega\,\Delta\tau + \delta\phi)$$

where the radian frequency has been expanded about the center frequency as $\omega = \omega_0 + \delta\omega$. Ideal compensation is achieved when $\Delta\tau$ equals 0, corresponding to adjusting τ to equal $(S/c) \sin \theta$ and adjusting the insertion phase at the center frequency so that $\alpha = 0^o$. Notice that ideal compensation is independent of frequency and thus has unlimited bandwidth.

Practical systems communicate finite bandwidth signals that have less stringent tolerance requirements for compensation than the ideal values that are frequency independent. These tolerances are derived by examining the average combining efficiency for the required bandwidth, BW. The average combining efficiency is computed by integrating the combining efficiency over the finite bandwidth and dividing the result by the bandwidth that yields

$$C_{\text{ave}} = 1 + (\sin X/X) \cos\, \delta\phi$$

where $X = \pi \text{BW}\, \Delta\tau$. Compensation tolerance loss values depend on allowable loss, and in this case a 0.1 dB loss is used for both the uncompensated time delay and the phase tolerances. The time delay tolerance depends on X that is directly proportional to the signal bandwidth. A 0.1 dB combining loss limits X to a value of 0.5265. Time delay tolerance values plotted in Fig. 2-7 illustrate that the required time delay tolerance becomes increasingly stringent as the signal bandwidth increases. The time delay tolerance for a 1 MHz bandwidth signal equals 167 nsec, but a 100 MHz bandwidth signal imposes a 1.67 nsec tolerance. The tolerance precision becomes more challenging for large signal bandwidths. Since the speed of light is about 1 ft/nec, the combination of the array geometry and the signal's elevation angle can be adequate to set time delay values for narrow bandwidth signals. A more precise means of determining and maintaining the time delay compensation is required for wide bandwidth coherent combining.

The tolerance for array element phase compensation at the center frequency does not depend on the signal bandwidth. If the phase compensation tolerance loss is assumed to be 0.1 dB loss, the phase

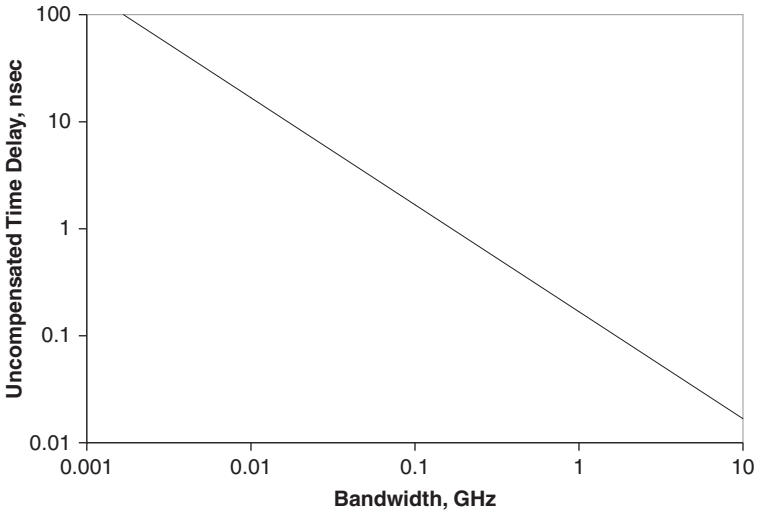


Figure 2-7 Time delay compensation tolerance versus bandwidth [38] (© 2008 IEEE)

compensation tolerance equals a 17.4° value. Other values of coherent combining efficiency loss due to phase error compensation accuracy are illustrated in Fig. 2-8. When the uncompensated phase error exceeds 90°, the combining efficiency is less than 1 and the combined gain of the array is lower than that of a single array element.

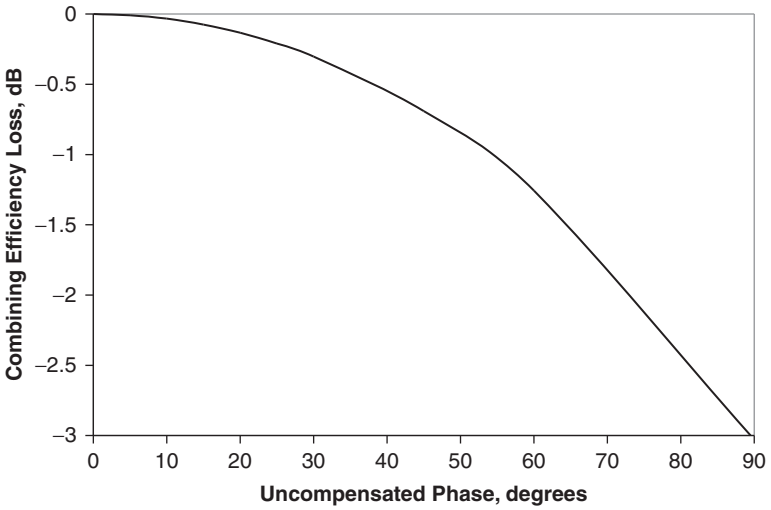


Figure 2-8 Combining efficiency sensitivity to phase compensation errors [38] (© 2008 IEEE)

Physically, the overall array angular response contains a large number of grating lobes. As frequency varies, the grating lobes scan in angle. The time delay compensation limits the amount of frequency scan for grating lobes close to the signal's direction. Good array performance over a bandwidth requires aligning a grating lobe with sufficient time delay compensation with the signal's direction. The alignment of the grating lobe peak with the signal direction is accomplished by the phase compensation.

In operation, coherently combining high-gain antenna elements requires a quick, reliable method to align the antenna with the signal's direction. This alignment requires the array antenna elements to track the signal direction and a means to determine, implement, and maintain the time delay and phase compensation. The time delay compensation for wide bandwidth signal is a particular challenge. A proposed means of implementing this array alignment [37, 38] uses a beacon signal for array alignment purposes that is transmitted along with the data signal. A wide bandwidth pseudorandom code would be used for the beacon signal. The coded signal and its processing gain provide sufficient signal strength to allow antenna element tracking using pseudo-monopulse tracking techniques described later in this chapter. Likewise, an adequate level of the pseudorandom beacon signal is needed for array element time delay and phase alignment. This beacon signal level, because of processing gain, would not interfere with the data signal and would consume only a fraction of the transmitted signal power. Time delay compensation values at each array element are determined by correlation with a replica of the coded signal. The array output signals are also correlated with a replica of the beacon signal to verify appropriate compensation values. The beacon carrier frequency would also be used to determine the required phase compensation, and if related to the data signal's carrier frequency, would advantageously assist in data signal acquisition, but with a signal level that does not interfere with the desired signal reception.

An implementation of the beacon alignment technique illustrated in Fig. 2-9 has an architecture comprised of the antenna elements with

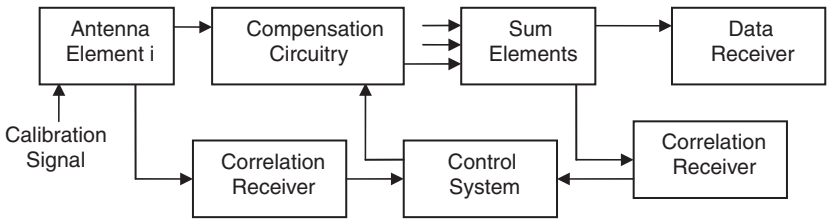


Figure 2-9 Array architecture for high-gain antenna elements [38] (© 2008 IEEE)

calibration test signals; correlation receivers to determine the element time delay values; element compensation circuitry with time delay, amplitude, and phase adjustment capabilities; a summer for the individual array element signal components; a correlation receiver to determine and maintain the phase alignment between array elements; and a data receiver. The individual array elements use a calibration signal to determine their insertion amplitude and phase values, and the compensation circuitry is adjusted so that the array electronic responses match element to element. The interconnecting cabling also is measured and any variation in their response is adjusted with the compensation circuitry to provide matched responses at the array element summing circuitry. The goal of this calibration is to have matched responses for the element electronics over the bandwidth. Design attention should be paid to component selection and thermal stability requirements during system development so that the desired active component's transfer functions match and that this match is maintained for reasonable time periods. The element calibration signals might use one signal spectra for initial alignment and out-of-band tones for monitoring during array operation to avoid interference with the desired data signals. Such calibration signals also provide a built-in test and diagnostic capability for array electronics.

The overall objective for array time delay compensation is to achieve a $\sin X/X$ value in the average combining efficiency that is close to 1. If the data signal is correlated, its time delay resolution is on the order of 1/BW. The time delay of the pseudorandom code is on the order of 1/(10 BW) assuming the code bandwidth equals the data signal's bandwidth. If the delay tolerances are equal to the time resolution of the signals, the signal correlation results in an X value of π so that $\sin X/X$ equals 0. The average combining efficiency in this case equals 1 and coherent combining affords the same performance as a single array element. By contrast, the correlation of the beacon signal results in an X value $\pi/10$, assuming the time delay compensation error is again the time delay resolution of the pseudorandom beacon signal. The value of $\sin X/X$ in this case is very close to 1 and results in a combining efficiency that is within 0.04 dB of the ideal value. The time resolution performance of the coded beacon signal allows accurate determination of the time delay compensation requirements. Combining efficiency performance close to ideal values should be achievable with proper phase compensation at the center frequency.

2.6 Antenna Tracking

The main beam of the receiving antenna must be aligned with the incident signal direction to avoid reduced signal levels. This reduction is commonly referred to as antenna pointing loss. Antenna alignment

requirements depend on the angular beamwidth of the antenna and the uncertainty of the knowledge of the signal's direction. Very broad beamwidth antennas require only a general orientation in the signal direction, whereas narrow antenna beamwidths must be aligned with precision. The process of commanding an antenna to a given angular location is referred to as antenna pointing while dynamically maintaining alignment with the signal direction is referred to as antenna tracking. A variety of user segment needs exist that range from simple orientation to closed-loop tracking systems capable of high angular precision. The space segment can also impose antenna tracking requirements. For example, crosslink systems generally use very narrow beamwidths, and closed-loop antenna tracking techniques are needed to compensate for the satellite's attitude variations. Antenna tracking techniques [39] are based on open-loop or closed-loop designs depending on the required accuracy and dynamics of the signal's direction.

Antenna pointing information for user segment antennas is derived from knowledge of the user's geographic location and the satellite's ephemeris. The satellite's orbital location position is defined by the ephemeris data, which specifies the orbit altitude, eccentricity, inclination angle, and time-of-epoch parameters in inertial space. The user's location and the satellite's ephemeris are used to define the required time variation of the antenna's azimuth and elevation variation. Uncertainties and inaccuracies in this information dictate the requirements for antenna tracking that can range from open-loop positioning to closed-loop tracking designs.

Analyses of antenna tracking performance generally are concerned with the antenna's main beam alignment. A convenient representation of the high-level portions of the antenna's main beam is a Gaussian function that is a good fit for practical antenna designs and is given by

$$f(\theta) = \exp[-K(\theta/\theta_{hp})^2] \text{ (voltage)}$$

where θ is the angle measured from the main beam boresight, θ_{hp} is the antenna's half-power beamwidth, and $K = 1.3816$, as can be determined by evaluating the expression at the half-power point, $\theta_{hp}/2$.

The angular accuracy of antenna tracking analyses is generally normalized [40] to the antenna's half-power beamwidth. This normalization results in expressions that are independent of the antenna's beamwidth and are suitable for use with error budget assessments of the factors that limit antenna tracking uncertainty. The antenna pointing loss that results from the antenna tracking uncertainty can be derived from the Gaussian pattern representation given in Fig. 2-10. The required tracking accuracy for communication systems is generally accepted as 1/10 of the antenna's beamwidth. This angular uncertainty limits the antenna pointing loss to about 0.1 dB. Better angular accuracy for communication

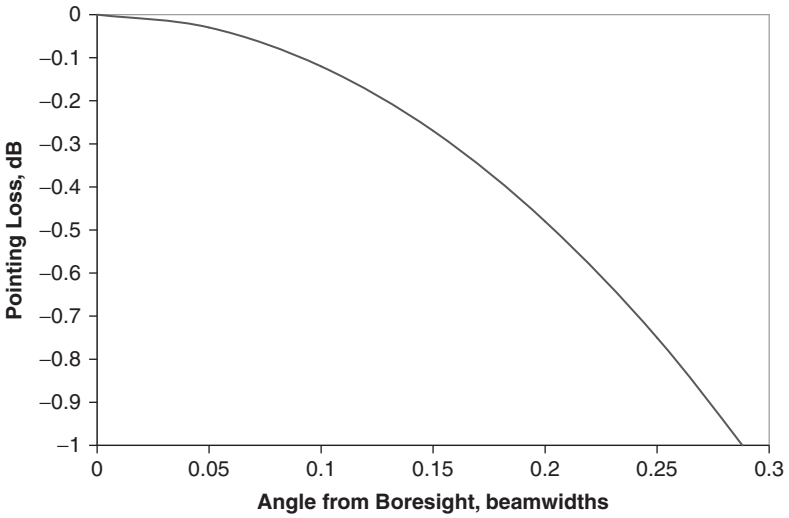


Figure 2-10 Antenna pointing loss versus misalignment [39]

applications would have an insignificant impact on link performance. Other applications, such as tracking radars, require as much tracking accuracy as possible in order to locate targets with precision, requiring more complex implementations.

Antenna tracking techniques can be separated into open- and closed-loop system designs. Open-loop systems principally require software control of the antenna pointing, while closed-loop systems require a more complex antenna design, additional receiver electronics, a control system, and supporting software.

2.6.1 Open-Loop Antenna Tracking

When the antenna's beamwidth significantly exceeds the uncertainty in the signal's direction, either fixed pointing or program track techniques are used. A typical example of the fixed pointing is user satellite TV antennas, whose beamwidths are sufficiently broad to permit fixed mounting after alignment with the desired satellite. Program track for user applications uses the satellite's ephemeris values and the geographic location of the terminal to calculate the time variation of the antenna's azimuth and elevation positions. In most program track applications, an ACU (antenna control unit) is programmed with the satellite's ephemeris and user location to command open-loop antenna pointing to the signal's direction. Geosynchronous altitude satellites have relatively little motion, and deviations in the satellite's inclination result in a small figure-eight

variation over a 24-hour period. The extent of this motion is controlled by occasional thruster adjustments to the orbit, a process referred to as stationkeeping. Polar orbits are used by meteorological satellites, for example, and have much more dynamic antenna tracking requirements. This orbital choice allows remote sensing with high-resolution performance because of the short range to the earth's surface and because global sampling of meteorological conditions results from the inclined orbit. User antenna tracking for such orbits is much more dynamic than geostationary antenna tracking.

An example system in Fig. 2-11 that uses program track can receive data from both polar and geosynchronous meteorological satellites. This capability is attractive to meteorologists because the resolution of low-altitude satellites is available when they are in view, and when polar satellites are not in view, "rapid refresh" data can be collected from geosynchronous satellites that scan the visible earth's surface continuously, providing information on weather dynamics. Data from meteorological satellites in both orbit are therefore complementary. This prototype design for meteorological satellite readout terminals provides both L- and S-band receive capabilities by using a scaled version of the rolled edge cavity previously described in Fig. 2-2. A 6-ft diameter reflector is used for high-resolution geostationary satellite data and has beamwidth values of 6.7° and 5.2° at L- and S-band, respectively. The terminal's location and the satellite's ephemeris permit determination of the required time history of the necessary azimuth and elevation angle variations. The antenna beamwidths are sufficiently broad in this case

Figure 2-11 Antenna that uses program track techniques

to allow reliable tracking performance using program track. Using the 1/10 beamwidth accuracy requirement that limits the pointing loss to 0.1 dB, the required positioning accuracy is on the order of 0.5°, which is achieved through antenna leveling, positioner alignment to true north, and the readout accuracy of absolute encoders in the azimuth and elevation axes of the antenna's positioner.

2.6.2 Step Track

If uncertainty exists regarding the accuracy of program track, another open-loop tracking technique, referred to as step track [41], can be used to verify correct antenna tracking. The antenna beam is positioned based on program track information, and is displaced from this position by equal and opposite angular increments, as indicated in Fig. 2-12. The received power levels are measured at both angular offsets. If the antenna's boresight axis is aligned with the signal's direction, the power levels at both angular offset positions will be identical. Differences in the power levels at each angular offset indicate misalignment with the signal direction. If the main beam's pattern is represented by a Gaussian function, the required angular correction θ_ε can be shown to equal

$$\theta_\varepsilon = -(\theta_{\text{hp}}^2/4K\theta_o)\ln R$$

where θ_{hp} is the antenna's beamwidth, θ_o is the angular offset used in the measurement, and R is the amplitude ratio of the signals at

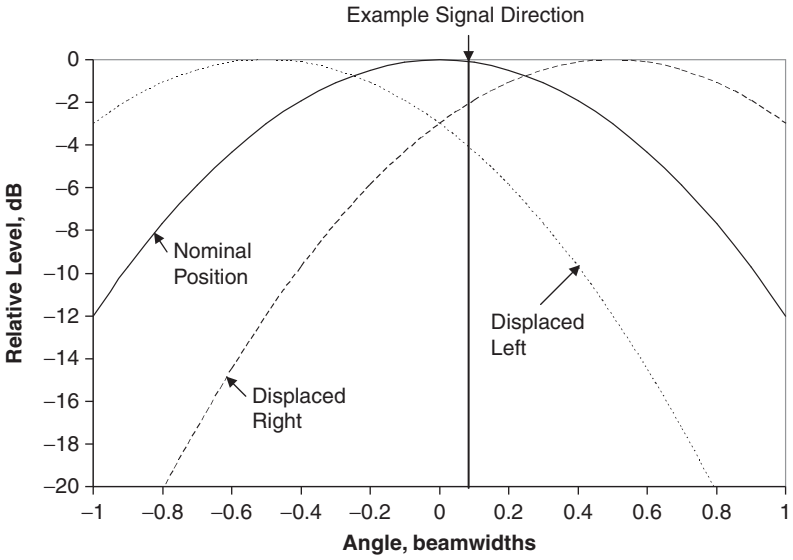


Figure 2-12 Step track operation [39]

the angular offset positions. Thus, the antenna's misalignment with the signal direction θ_ε can be determined from the measured signal level at equal and opposite angular offsets.

The angular accuracy for step track has been derived using the Gaussian representation of the antenna pattern. Antenna tracking accuracy [40] is defined by the rms angular error normalized by the antenna's half-power beamwidth. This normalization is directly related to the antenna pointing loss (e.g., a 1/10 beamwidth tracking accuracy results in a 0.1 dB antenna point loss). The angular accuracy can be shown to equal

$$\sigma_\theta/\theta_{\text{hp}} = (\theta_{\text{hp}}/2K\theta_{\text{o}})\sigma_A$$

where σ_θ is the rms angular error and σ_A is the rms amplitude measurement accuracy. Amplitude measurement accuracy is limited by SNR (signal-to-noise ratio) and the rms amplitude measurement accuracy equals $1/(2\text{SNR})^{1/2}$. The antenna tracking accuracy can be expressed as $k_{\text{st}}/(\text{SNR})^{1/2}$, where k_{st} is the step track accuracy coefficient. Numerical values for this coefficient in Fig. 2-13 are shown as a function of the pattern levels used in the angular offset values.

The selection of an angular offset value for step track measurements is based on having a small value of the step track accuracy coefficient. The behavior of k_{st} results from the following factors. Small angular offsets have limited measurement sensitivity because the antenna pattern

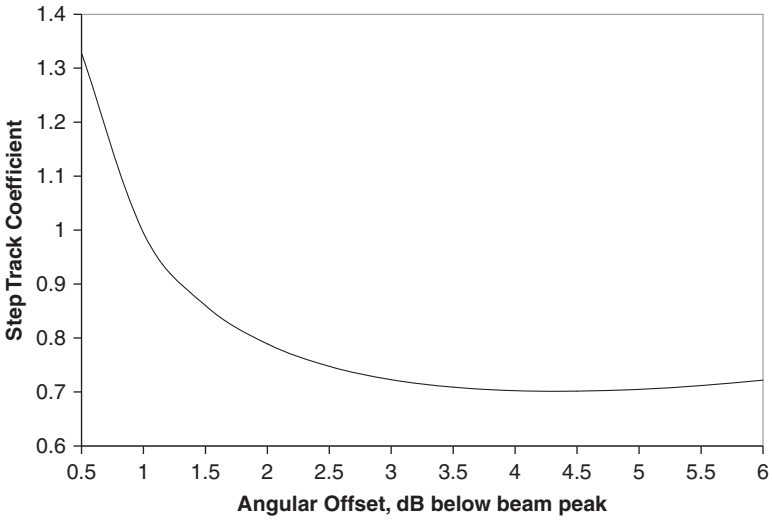


Figure 2-13 Step track coefficient [39]

near the main beam peak is relatively flat and high amplitude measurement accuracy is required to discern amplitude differences at the two offset angles. Large angular offsets benefit from the larger antenna pattern slope at the large angular offset angles; however, this increased measurement sensitivity is also accompanied by reduced SNR values because the offset angles are sampled at lower antenna gain levels. These factors are indicated in a large angular displacement; the ability to measure the amplitude at both angular displacements is degraded by the loss in SNR from the antenna pattern reduction at the large angular displacement. These factors result in the values shown in Fig. 2-13, where the desired small values of k_{st} have a broad minima region around the minima at a pattern level of 4.3 dB lower than the peak gain on the antenna's boresight axis. In practice, systems normally have a system margin value relative to their threshold sensitivity. One strategy selects the angular offset value based on the available system margin. In this way, the received data are not impacted during the step track's angular sampling by antenna pattern levels that would be lower than the gain level required to receive the threshold signal power.

For many communication systems, antenna tracking is based on program track and verified and refined by step track means. Geosynchronous satellites have relatively little motion in orbit, and stationkeeping using occasional spacecraft thrusting limits the amount of orbital movement. This movement follows a north–south figure-eight path. For narrow beamwidths, some angular tracking may be required, and step track verification might be performed four times a day to follow the figure-eight trace. For other satellite orbits, the user antenna must dynamically follow the satellite's orbital trajectory. In this case, the satellite's ephemeris and user's location can be used to predict the time history of the required azimuth and elevation variation, as is done in program track. Conventional step track techniques assume the signal direction remains constant over the time required to offset the antenna to perform the received power measurements necessary in the step track procedure. The conventional step track procedure has been extended to apply to situations where the signal direction is varying during the measurement process, a technique referred to as rate-corrected step track [42].

Rate-corrected step track initially requires acquiring the signal as the satellite clears the local horizon and then verifies the predicted time history variation of azimuth and elevation. Prior to the estimated satellite arrival time, the precalculated azimuth angle is used to position the antenna. Typically, ephemeris errors are offsets in time due to satellite drag; perturbations of the satellite in its orbital plane require momentum changes. At low elevation angles, the receiving antenna is typically limited by multipath errors so it is recommended that

the antenna boresight be positioned somewhat above the horizon and that the antenna be swept back and forth in azimuth over an anticipated uncertainty value maintaining a constant elevation angle until the receiver acquisition occurs. By maintaining a constant elevation angle, the effect of multipath changes on the received signal level with elevation angle increases is minimized. If the receiver has not acquired by the anticipated time period, a wider azimuth search is required. The azimuth and elevation rates of the orbital trajectory are modest as the satellite clears the horizon and allows time to search for the satellite.

The antenna's spatial alignment proceeds after the receiver has acquired the signal. The alignment in azimuth is performed first so that multipath variations in the elevation direction do not perturb the received signal level. Commanded angular offsets are processed in the same way as conventional step track procedures since only azimuth variations are sampled. Alternatively, if the received signal levels are recorded during the azimuth sweep, a fit between the known antenna pattern and the measured signal power variations can be made to determine the correct azimuth position. In this way, azimuth multipath variations can be averaged in aligning the antenna's azimuth axis.

Throughout this period, the elevation angle of the antenna remains fixed. As the satellite increases in elevation, the received signal varies in accordance with the antenna pattern. Since this portion of the trajectory is at low elevation angles where multipath is an issue, measuring the received signal during the elevation rise and fitting the data to the antenna pattern is an appropriate averaging technique. From such data, the antenna boresight can be realigned with the satellite's elevation angle. At this point, the receiver has acquired the signal and the antenna is correctly aligned in both azimuth and elevation coordinates.

Following this initial acquisition of the signal and antenna boresight axis alignment, the estimated trajectory is compared with the measured values for this time period of the satellite's trajectory. Discrepancies in this comparison are generally time offsets, used with validation to follow the satellite. The accuracy of this trajectory derived from the ephemeris is indicated in part by the correspondence between the anticipated and actual acquisition parameters (i.e., the azimuth angle and time of acquisition comparison). Initially, the antenna is positioned in accordance with the estimated trajectory, and the correctness of this positioning is validated by the rate-corrected step track procedure.

While the conventional step track technique provides angular displacements in azimuth and elevation directions, since the signal source is moving in both coordinates, such conventional step track's azimuth and elevation sampling results in coupled measurements between the angular coordinates. The motion for the sampling for the rate-corrected step track is in the cross track and in-track directions, and the sampling

times also vary in accordance with the azimuth and elevation rates. At low elevation angles after initial acquisition, the angular rates are modest but increase at higher elevation angles. More frequent sampling is performed at higher elevation angles where the angular rates are more dynamic.

The cross track sampling is made by commanding equal and opposite angular offsets in the cross track direction. The motion of the antenna is an "s-shaped" curve about the trajectory, and based on the power measurement samples during the excursion of the "s," the antenna can be realigned from the estimated trajectory to the actual trajectory. Two alternatives exist for the in-track alignment. One alternative is to command the antenna by equal and opposite amounts in the in-track direction. The realignment is then based on the signal differences at the commanded positions. The second alternative accelerates the antenna in the in-track direction and then slows the antenna in the in-track direction to allow the signal to drift through the main beam peak. Again, like the procedure recommended for the elevation alignment in initial acquisition, the measured signal variation can be fit to the antenna pattern and realignment to the beam peak can be made based on this pattern fit.

2.6.3 Closed-Loop Tracking

When the uncertainty in open-loop tracking performance exceeds the required angular uncertainty, closed-loop antenna tracking systems are used to dynamically track the signal's direction. Closed-loop techniques track the dynamics of the signal's direction variations, compensate for platform motion, and/or offset wind loading disturbances to maintain antenna tracking accuracy. The closed-loop antenna tracking is referred to as monopulse tracking, which was developed for radar applications where variations in the received signal level limited other tracking techniques. The term derives from the ability to determine the signal's direction from a single radar pulse. In practical radar system designs, a tracking capability that is invariant to signal level fluctuations is required. Communication applications have well-behaved received signal levels in comparison to radar systems, allowing design simplification.

Monopulse systems operate by forming two antenna beams: a sum beam that is the normal antenna main beam that receives the signal with highest sensitivity, and a difference beam used by the tracking system that has a pattern null aligned with the antenna's boresight axis. In operation, a closed-loop system is configured to align the difference null with the signal direction. The ratio of the difference and sum beam output levels is referred to as the error response, which is the

closed-loop system's input signal. As will be shown, the error response is zero on the antenna's boresight axis and has a linear variation for angular deviations from the boresight axis that are positive on one side of the axis and negative on the other side of the axis. Thus, the error signal response error follows the classic variation used in control systems. Minimizing the error response aligns the antenna with the signal direction; the magnitude of the error response is proportional to the separation of the signal direction from the antenna's boresight axis, and the sign of the error response indicates which side of the axis corresponds to the signal direction.

Radar systems commonly use three receivers, one for the sum beam and two more that separately detect the azimuth and elevation difference beams. The receiver outputs are simultaneously processed so that antenna tracking is unaffected by changes in signal level. Communication signal levels do not have the same dynamic variations as radar systems' signals so that azimuth and elevation axes are sequentially sampled by a single receiver and averaged. The tracking for communication applications is referred to as pseudo-monopulse since both azimuth and elevation channels are not simultaneously and continually received. The difference pattern channels are sequentially coupled onto the data channel by a switching circuit, and synchronous detection of the resulting AM (amplitude modulation) on the data channel provides the tracking information. The AM level varies linearly with the signal's displacement from the antenna boresight, commonly referred to as the "monopulse error slope." The control system functions to minimize the AM value as an "error signal" to align the antenna's boresight axis with the signal direction.

The tracking accuracy for the monopulse technique can be derived [39, 43] using Gaussian functions such as the step track analysis. Two Gaussian beams are displaced from the boresight axis and overlap at their half-power points. When these two beams are added together, a sum beam results, and when the two beams are subtracted, a difference beam results, as illustrated in Fig. 2-14. The ratio of these difference and sum beam patterns gives the error response shown in Fig. 2-15. The angular accuracy of the monopulse design using these beam representations becomes

$$\sigma_\theta/\theta_{\text{hp}} = 0.361/(\text{SNR})^{1/2}$$

The monopulse accuracy coefficient is smaller than the step track accuracy coefficient and, generally, monopulse designs have better angular accuracy than step track designs. The errors in this discussion consider the RF tracking outputs. Practical designs have other errors resulting from mechanical imperfections such as backlash, wind gusts,

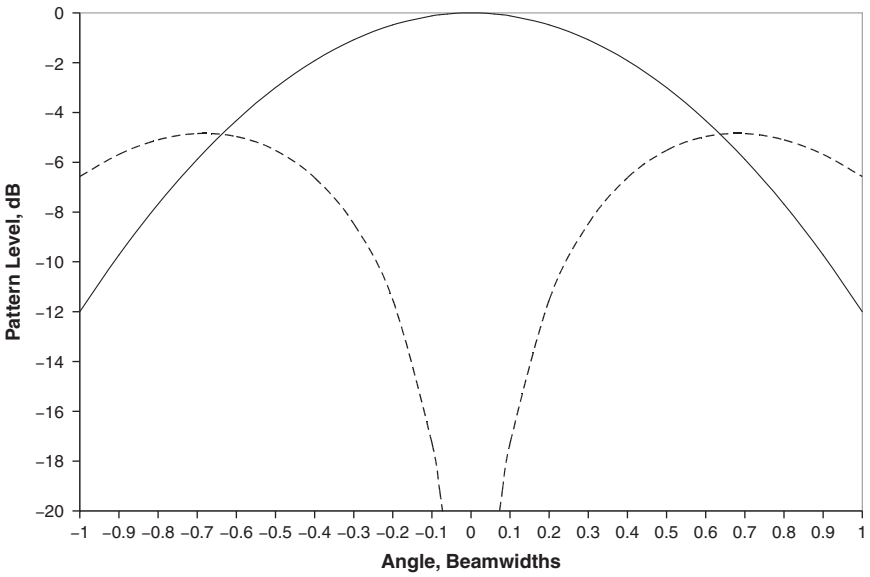


Figure 2-14 Sum (solid) and difference (dashed) patterns [39]

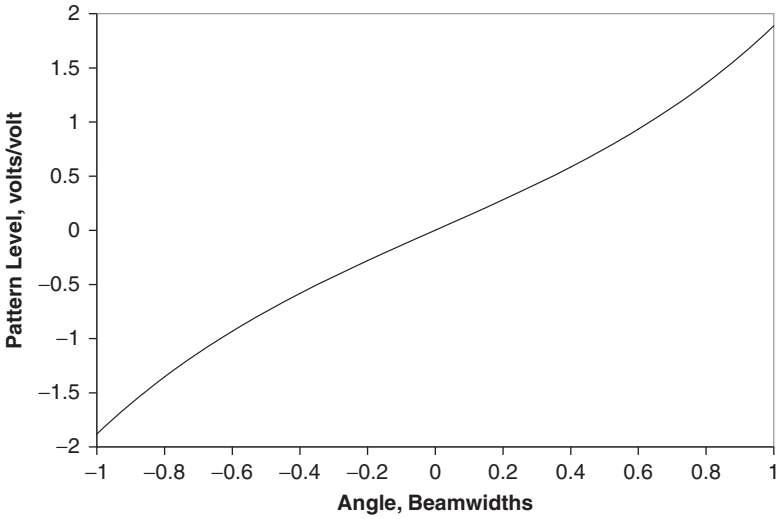


Figure 2-15 Error response corresponding to the sum and difference beams in Figure 2-14 [39]

control system offsets, and RF null offsets for monopulse designs, and positional encoder errors for open-loop tracking techniques. The overall tracking accuracy requires examination of these other error sources, and from their component errors an error budget is constructed from

the rss sum of the individual rms values of the component errors. For antennas that are not enclosed in radomes to protect the antenna from wind gusts, the error budget generally considers tracking performance in calm conditions and under wind loading conditions separately. The analyses of antenna tracking variations with wind loading [44] follow commonly established procedures that assess both steady state and gusty conditions.

2.6.4 Monopulse Feed Designs

While monopulse designs provide accurate tracking and can dynamically follow changes in signal direction, their implementation cost is higher because of the control system and additional receiver requirements, but in large part because of a more complex antenna feed design. Early monopulse systems used a feed design comprised of multiple horns centered on the boresight axis. More recently, multimode feed designs have been developed. The sum beam uses the usual waveguide mode excitation, while the difference beams are formed from higher-order waveguide mode excitation. These two design approaches are discussed in turn.

The multiple horn tracking feeds form the sum beam by adding the horns together to produce an antenna pattern that has a peak main beam level coincident with the antenna's boresight axis. The difference beams that provide antenna tracking capabilities are produced by subtracting the horns on opposite sides of the boresight axis. The subtracted antenna pattern has a null on-axis and is positive on one side of the boresight axis and negative on the other side. This process was used to determine the sum and difference patterns shown in Fig. 2-14 and the error response shown in Fig. 2-15, where the off-axis beams that are formed by the feed cluster in the focal region use Gaussian functions for their representation. The error response is the input to the antenna's position control system and is the ratio of the difference and sum patterns. When the antenna is ideally tracking, the error response equals zero. Tracking deviations from the antenna's boresight axis have a linear deviation that is positive on one side of the axis and negative on the other. Thus the monopulse circuitry results in a classic control system response, minimizing the error response and a well-behaved linear response when the error is non-zero. This linear response persists for a significant portion of the antenna's main beam angular extent but deviates for larger values. Thus, the initial antenna pointing must have sufficient accuracy to align the antenna with the signal direction so the control system functions properly. Further discussion follows regarding the acquisition issues.

One problem with this simple combination of off-axis antenna beams is that the sum and difference patterns cannot be separately optimized [45]. If the off-axis beams are summed to form a desired aperture

amplitude taper, the individual beams that form the difference beam would have high crossover levels and the resulting difference patterns would have a shallow linear slope. If the off-axis beams are selected to provide difference patterns with a higher linear slope by increasing the angular separation of the component beams from the boresight axis, the beamwidth of the feed cluster sum would become narrower, increasing the amplitude taper of the reflector illumination and thus reducing antenna efficiency. An alternative approach is to use a normal feed design for the sum beam and integrate smaller antenna elements into the throat or perimeter of the sum pattern feed. One example [46] used four polyrod radiators located within the throat of a sum beam feed horn. These elements are arranged in four quadrants surrounding the central horn, and opposing pairs of elements can be subtracted to form vertical and horizontal difference beams.

Hybrid networks are generally used to combine the antenna elements in multiple aperture monopulse feed designs. The amplitude and phase accuracy requirements [47] have been examined for such hybrid combining circuitry. Two distinct monopulse combining circuitry networks are addressed using Gaussian beam representations of the individual beams being combined to produce the monopulse tracking feed patterns. The first combining circuitry, referred to as "hybrid combined," produces the sum and difference beam patterns by combining multiple horns. The second, referred to as "separate sum beam," uses a single central feed to form the sum beam and hybrid networks to combine four antenna elements to form difference beams. The pattern characteristics and error responses for two hybrid combining circuit designs are presented in Table 2-1.

TABLE 2-1 Patterns and Error Response for Hybrid Monopulse Circuits

	Hybrid Combined	**Separate Sum Beam**
Sum Beam	$[\exp(-X^2/K)][2\cosh(X) + \delta]/2$	$\exp(-X^2/K)$
Difference Beam	$[\exp(-X^2/K)][2\sinh(X) - \delta]/2$	$[\exp(-X^2/K)][2\sinh(X) - \delta]/2$
Error Response	$(2\sinh(X) - \delta)/(2\cosh(X) + \delta)$	$[2\sinh(X) - \delta]/2$

where $X = K\theta/\theta_{\text{hp}}$ and $\delta = (A - 1)\exp(-X)$ when $\theta_o = \theta_{\text{hp}}/2$. The ideal hybrid circuitry results when the parameter A equals 1. Amplitude and phase deviations from this ideal value of 1 result in the hybrid circuit imperfections expressed by the parameter δ and the combining tolerance, where A is the amplitude and/or phase tolerance deviation from ideal combining where A equals 1. The error response is the ratio of the difference and sum beams and is the control system input.

Amplitude errors in monopulse combining circuitry result in angular shifts in the boresight axes of the difference and sum beams. The difference pattern null shift ε_d equals

$$\varepsilon_d = [\ln A/2\text{K}]\ \theta_{\text{hp}}$$

This error shifts the difference pattern null toward the component beam having the lower value. Amplitude errors in the hybrid combined design also shift the sum beam position, but in a direction toward the component beam with the higher level. The sum beam shift is derived by setting the derivative of the sum beam to zero, which yields a sum error ε_s that equals

$$\varepsilon_s = \{-(A + 1 - K) + [(A + 1 - K)^2 - 4(A - 1)K]^{1/2}\}\ \theta_{hp}/4K$$

The boresight shifts for the difference beam and the sum beam and total boresight shift for the hybrid combined case are presented in Fig. 2-16. Amplitude errors also result in a minor variation of the error response slope values.

Phase errors in the combining circuitry result in filling in the difference pattern null because the quadrature error component cannot be cancelled. The difference pattern null depth, ND, equals

$$\text{ND} = [1 - \cos\ \varphi]^{1/2}/2^{1/2}$$

where φ is the phase imbalance relative to the zero ideal value, as illustrated in Fig. 2-17. The difference pattern null filling also results in a non-zero error response at boresight as indicated in Fig. 2-18 for the SSB case with a 15° phase error. The ideal error response makes a 180° phase transition at the boresight axis, but this phase transition is rounded when phase error is present.

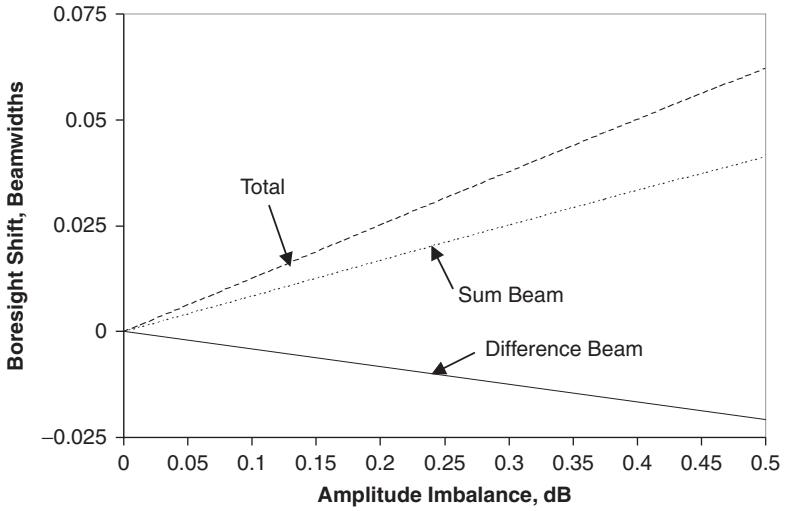


Figure 2-16 Boresight shifts with amplitude imbalance [47] (© 2007 IEEE)

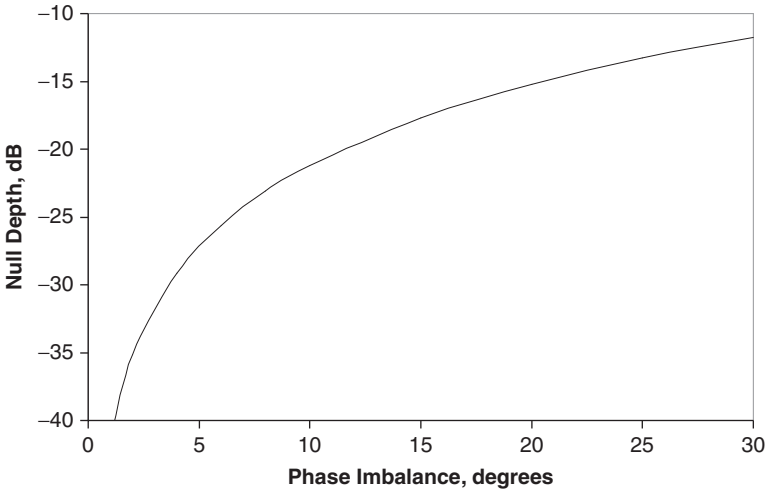


Figure 2-17 Null depth versus phase imbalance [47] (© 2007 IEEE)

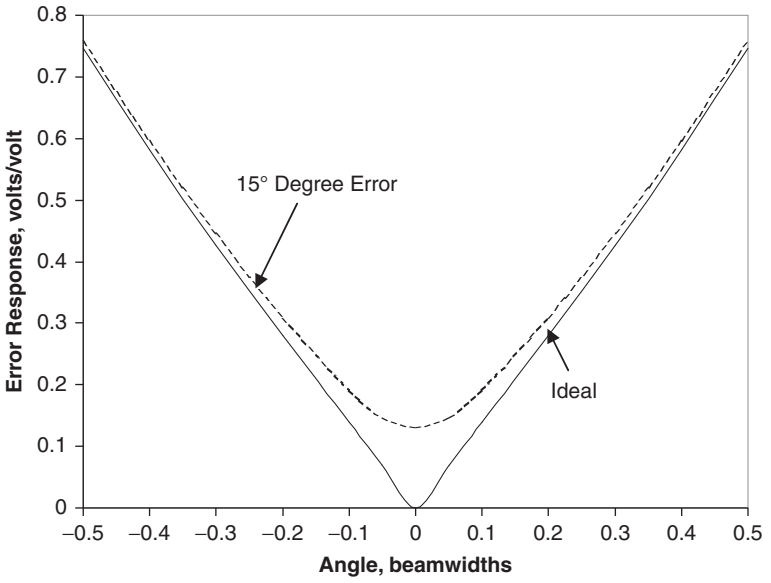


Figure 2-18 Null filling with 15° phase imbalance [47] (© 2007 IEEE)

The allowable amplitude and phase imbalances in practical designs depends on the specific application. Typically, the overall tracking accuracy of one-tenth of the antenna's beamwidth limits the antenna pointing loss to 0.1 dB, but this performance includes other error sources

besides hybrid circuitry imbalances. A representative error component allocation for combining circuitry imbalances might limit the values to 0.05 beamwidths. The amplitude imbalance values for a 0.05 beamwidth error are about 1.2 dB for the separate sum beam design and about 0.4 dB for the hybrid combined design, as illustrated in Fig. 2-16. The tighter hybrid combined tolerance results from the oppositely directed and roughly doubled magnitude errors for the sum beam in the hybrid combined case. The null depth limitation can be used to derive a phase tolerance value for a 0.05 beamwidth deviation. Without phase error, an ideal pattern has a –23 dB null depth at a 0.05 beamwidth deviation from the boresight axis. A 25 dB null depth limitation from phase imbalance is assumed adequate and results in a null depth that is 2 dB lower than the ideal null depth of 23 dB at a 0.05 beamwidth offset. The 25 dB null depth value corresponds to a phase imbalance of about 6^{o}, as illustrated in Fig. 2-17.

Another monopulse antenna feed technology [48, 49] uses a multimode aperture design. The dominant mode produces the sum beam, and higher-order waveguide modes produce the difference patterns. The dominant mode has an in-phase aperture distribution, while the difference mode pattern has a 360^{o} phase progression that results in a pattern null on the boresight axis. The same 360^{o} phase progression exists in the far field antenna pattern. In this case, the antenna's angular misalignment is expressed in a polar ρ, φ form rather than the x, y form previously discussed. The magnitude of the error response, ρ, indicates the radial distance from the antenna's boresight axis and the phase φ that is measured relative to the sum beam pattern indicates the azimuth position of the antenna's angular misalignment. These feed designs require higher-order mode couplers to separate dominant and higher-order mode components. Such feeds generally sequentially sample the higher-order mode components in angular quadrants and add the sampled signal components to the received sum beam signal. The sequential sampling results in an amplitude modulation pattern onto the received sum beam signal level. The height of the amplitude modulation indicates the angular displacement from the antenna's boresight axis. The high and low values of the amplitude modulation where the coupled higher-order mode adds and subtracts with the sum signal respectively indicate the azimuth plane where the antenna misalignment occurs.

In such designs, inaccuracies in the phase of the sum beam and difference beams result in tracking errors that are referred to as "crosstalk." For example, if the antenna is displaced a horizontal distance from the antenna's boresight axis, the control system should respond with motion within the horizontal plane. If the antenna follows both a horizontal and vertical trajectory in returning to the boresight position,

the vertical motion referred to as crosstalk results from phase offsets between the sum and difference channels that need to be corrected for proper tracking operation. For narrow bandwidth applications, the phase between the sum and difference beams is adjusted with a phase shifter to minimize antenna crosstalk errors. When wide bandwidth requirements must be satisfied, the group delay differences rather than the phase shift between the sum and difference beams must be equalized to minimize crosstalk. The evaluation of antenna tracking systems is discussed in Chapter 8.

2.6.5 Signal Acquisition Issues

When antenna tracking techniques are used, the antenna is initially pointed under command to a nominal azimuth and elevation position based on a priori knowledge such as ephemeris data. If the initial pointing knowledge is sufficiently accurate, the receiver should acquire the signal. The accuracy required of the initial pointing direction clearly depends on the antenna's beamwidth. Generally, antenna tracking performance is limited to a portion of the angular width of the antenna's main beam that is referred to as the signal acquisition angular width of the antenna's tracking design. If the receiver does not acquire the signal, a search for the signal's direction is performed by varying the antenna's pointing about the nominal pointing direction until the receiver acquires the signal. Both raster and spiral scan patterns are used in practice in seeking signal acquisition.

Narrow beamwidth antennas are required to satisfy high data rate applications and require accurate, a priori knowledge of the signal's location to minimize search time. Two techniques can be used to extend the acquisition field of view for narrow beamwidth antennas. Both of these techniques capitalize on the sensitivity of the tracking receiver. Generally, tracking receivers use carrier tracking techniques where high sensitivity is achieved by a narrow bandwidth receiver response and averaging techniques to reduce the variance of the receiver signal level. By contrast, a data receiver requires higher signal levels to demodulate the signal with adequate fidelity.

One technique for narrow beamwidth signal acquisition uses a smaller antenna having a broader beamwidth than the antenna used for data reception to acquire the signal and measure its direction. The gain of the smaller antenna must be sufficient to allow reliable antenna tracking performance. Both step track or monopulse tracking techniques can be used by the broader beamwidth antenna to measure the signal's direction with sufficient accuracy to allow pointing alignment within the data antenna's tracking acquisition width. This technique results in the additional complexity of a separate smaller antenna.

Since high-gain antennas commonly used dual reflector designs, the smaller signal acquisition antenna can be located behind the subreflector so that boresight coincidence of the smaller and data antennas is established by design.

The second technique [50] extends the angular acquisition while using the narrow beamwidth antenna's reflector. In addition to a central feed system for normal tracking and data reception, additional feed elements are placed in the focal region and each of these feed elements are connected to a tracking receiver. During signal acquisition, the signal arrival direction is identified by the tracking receiver having the highest signal level. Since the boresight axes for each antenna beam is established by design, knowledge of the required pointing correction is obtained by the angular separation of the feed with the highest signal level and the central feed. This acquisition technique requires additional antenna feeds and tracking receivers but avoids the required additional smaller antenna needed by the first technique. Tradeoffs exist regarding the number of antenna feeds and their separation from the central feed, and these tradeoffs depend on the sensitivity of the tracking receivers.

The high sensitivity of the tracking receiver can result in situations where the antenna is tracking on an antenna sidelobe rather than the main beam. Some indication of sidelobe alignment results when the data receiver either does not acquire the signal or has excessive noise. However, main beam alignment verification [51] is desired in many applications and alignment verification before initiating closed-loop tracking operation is also desired. Three techniques for main beam alignment verification are described.

One technique uses a separate smaller antenna in addition to the larger main antenna required for signal reception. The smaller antenna, generally referred to as a guard antenna, is designed so that its main beam gain level is comparable to the sidelobe antenna gain values of the main antenna. Tracking signal levels are compared for both antennas. If the tracking signal level in the main antenna is significantly larger than that for the guard antenna, the signal is aligned with the antenna's main beam. If the signal levels in each tracking receiver are comparable, a sidelobe of the main antenna is possibly aligned with the signal direction. Open-loop commands are then made to reposition the antenna sequentially in equal and opposite angular offsets in azimuth and elevation. The angular offset sizes are on the order of one-half of the antenna's beamwidth. A comparison of the tracking receiver outputs is made at each offset position. If the signal level of the main antenna's tracking receiver at one of the four commanded offsets is significantly higher than the tracking receiver's output level of the guard antenna, the required main beam alignment is in the direction of that offset direction. Further open-loop commands using smaller angular offset values provide additional confidence of main beam alignment and thus normal tracking operation can

be initiated. If none of the four angular offsets results in tracking receiver signal levels in the main antenna that exceed the guard antenna's tracking receiver output, the antenna boresight axis is more than one sidelobe removed from the signal direction. Further open-loop commanding using step sizes based on the antenna's sidelobe response are required to align the antenna's main beam with the signal direction.

A second technique applies to antenna systems that use open-loop tracking techniques. An examination of antenna pattern characteristics reveals the angular width of the main beam is considerably greater than the angular width of a sidelobe. In this case, step track commands are made and the measured values at the angular offsets are fit to the known main beam's angular response. If the measured values do not fit the anticipated response, an antenna sidelobe is aligned with the signal direction. Again open-loop commanding using the known antenna pattern to select pointing offsets as described earlier is used to locate the signal direction. If the commanded angular offset results in a significant increase in the tracking receiver's output level, the antenna's main beam alignment is in the vicinity of the signal direction, and smaller angular offsets are used to determine the signal direction. Finally, the normal step track operation is initiated. The open-loop commanding used in this process can also be used as a means of aligning closed-loop antenna tracking designs with the signal direction.

The third technique is based on the closed-loop antenna response. Closed-loop antenna tracking is based on a linear variation and predetermined slope of the monopulse error slope. This error response behavior, however, persists only over a limited portion of the antenna's main beam, and departs significantly from that linear response beyond the main beam region. Error response variations for wider angles are illustrated in Fig. 2-19. Since the angles in Fig. 2-19 extend into the

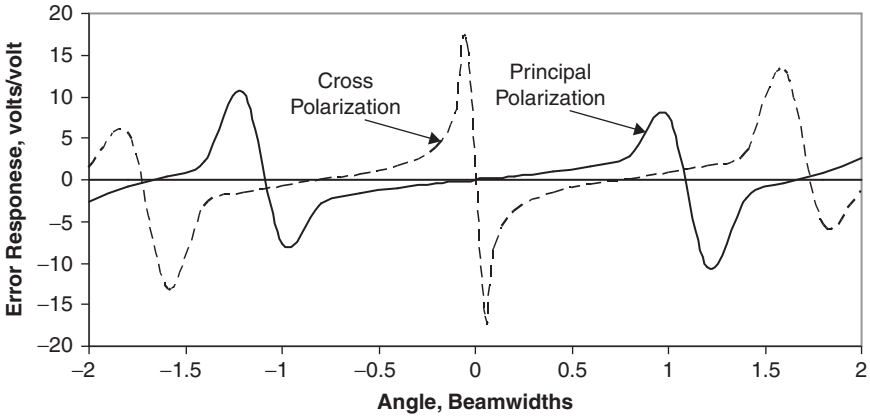


Figure 2-19 Monopulse error slope variation

antenna's sidelobe region, the error response was computed from composite beams having a sinX/X function to represent the sidelobe response. If available, measurements of the error response values for the actual antenna can be used. If sidelobe alignment with the signal is a program concern, measurements of the error response well beyond the design linear region are recommended. The error functions are shown for both a principally polarized response and a cross-polarized response (to be discussed later).

Main beam alignment verification for closed-loop antenna tracking designs can be based on two factors concerning the monopulse error response. The first factor is the determination of excessive amplitude modulation results at the antenna's initial pointing direction. Most tracking receivers can measure a limited amount of amplitude modulation, and excessive amplitude modulation or receiver modulation saturation indicates antenna sidelobe alignment with the signal direction. As illustrated in Fig. 2-14, at the edges of the main beam the level of the difference pattern exceeds that of the sum pattern, resulting in overmodulation and deviation from the error response's linear slope. A second factor concerns the slope of the monopulse error response. As illustrated in Fig. 2-19, the error response slope in the linear region is significantly steeper in the sidelobe region than the design error slope in the main beam. Open-loop commanded angular offsets can be used to measure the error slope to determine if the antenna's main beam or a sidelobe is aligned with the signal's direction. Again open-loop commanded angular offsets and tracking receiver output measurements are used to locate the main beam. Closed-loop tracking is initiated once main beam alignment is verified.

The error response in Fig. 2-19 also illustrates the behavior when the received signal has the design polarization sense and when the received signal is cross-polarized. Within the design acquisition limits, the error response for the correct polarization sense has the desired behavior, namely zero error on the boresight axis and a linear gradient for small angular displacements from the boresight axis. However, the error response for the cross-polarized signals has a decidedly different behavior, the region near the boresight axis has the wrong slope with an exceedingly high value, and zero error points are located away from the boresight axis at angles corresponding to about 10 dB below the sum beam peak level for the design polarization. The error response for cross-polarized signals can result in unstable tracking, as will be illustrated with measured responses to cross-polarized signals in Chapter 8.

Antenna tracking designs [52] can be implemented with difference antenna elements having either linear or circular polarization. When the difference elements are linearly polarized, the antenna properly tracks signals having the design linear polarization and for both senses of

circular polarization. However, the antenna does not track signals having the orthogonal linear polarization, simply because no signal component is received. When the difference elements are circularly polarized, the antenna properly tracks signals having design sense circularly polarized signals and linearly polarized signals having any orientation. However, antenna tracking can become unstable when circularly polarized signals having the opposite sense to the design sense are received.

When antenna tracking systems are operated in applications where the received signal has varying polarization properties, two requirements must be satisfied by the system design. Such polarization diverse applications require a means to maintain the received signal level without incurring significant polarization mismatch loss and maintain antenna tracking performance for varying polarization conditions. The first requirement of controlling polarization mismatch loss is satisfied by an antenna design whose sum channel has orthogonally polarized outputs and by the combining of the two orthogonally polarized sum channel outputs in a polarization diversity combiner that dynamically tracks the received signal's polarization. One implementation for polarization diversity combiner designs combines the orthogonally polarized antenna outputs in a vector modulator that can combine the two signals in any combination of amplitude and phase. One of the vector modulator's outputs is connected to the data receiver. The power level in the second vector modulator output is used as an input to a control system that maintains the adjustment of the vector modulator by minimizing the power at that second vector modulator output used by the controller. This second vector modulator output is the error signal to the control system, and by minimizing that error, the polarization of the received signal is dynamically matched by the receiving antenna to minimize polarization mismatch loss.

The second requirement is to maintain antenna tracking when the received signal's polarization is varying. Several alternatives exist in addressing diverse polarization issues in antenna tracking. As with the sum channel for data reception, tracking elements with orthogonal polarization output are required. One means to maintain antenna tracking in polarization diverse environments is to command the polarization of the tracking elements from the polarization weighting of the sum channel based on the polarization diversity combiner. This approach requires an additional vector modulator combiner. An alternative approach [53, 54] is to sequentially detect the orthogonally polarized tracking channel outputs and sum them. This sum can be shown to be dominated by the tracking channel output that is most closely matched to the polarization of the incident signal. In comparison to a design whose tracking channels dynamically match the received signal's polarization, the implementation of this alternative approach is less complex and costly but

results in a somewhat reduced tracking sensitivity since the tracking response of the channel with the dominant polarization response is measured only one-half of the time and incurs some polarization mismatch loss. However, as discussed, the sensitivity of the tracking channels exceeds the sensitivity of the data channels so that the loss of tracking sensitivity in this alternative approach can be accommodated.

References

1. W. D. Burnside and R. J. Marhefka, "Antennas on Aircraft, Ships, or Any Other Large, Complex Environment," in Y. T. Lo and S. W. Lee (eds), *Antenna Handbook* (New York: Chapman and Hall, 1993).
2. J. A. Stratton, *Electromagnetic Theory* (New York: McGraw-Hill, 1941).
3. D. S. Filipovic and T. Cencich, "Frequency Independent Antennas," J. L. Volakis (ed), *Antenna Engineering Handbook* (New York: McGraw-Hill, 2007).
4. D. E. Ping, J. T. Shaffer, L. U. Brown, and R. B. Dybdal, "A Broadband Rolled Edged Cavity Antenna," *2004 IEEE AP-S Symposium Digest*, Monterey CA, June 2004.
5. P. D. Potter, "A New Horn Antenna with Suppressed Sidelobes and Equal Beamwidths," *Microwave Journal*, vol. VI (June 1963): 71–78.
6. R. E. Lawrie and L. Peters, "Modifications of Horn Antennas for Low Sidelobe Levels," *IEEE Trans Antennas and Propagation*, vol. AP-14 (September 1966): 605–610.
7. A. W. Love (ed), *Horn Antennas* (New York: IEEE Press, 1976).
8. C. Granet, G. L. James, R. Bolton, and G. Moorey, "A Smooth-Walled Spline Profile Horn as an Alternative to the Corrugated Horn for Wide Band Millimeter Wave Applications," *IEEE Trans Antennas and Propagation*, vol. AP-52 (March 2004): 848–854.
9. W. D. Burnside and C. W. Chuang, "Aperture Matched Horn Design," *IEEE Trans Antennas and Propagation*, vol. AP-30 (July 1982): 790–796.
10. J. T. Shaffer, D. E. Ping, and R. B. Dybdal, "A Simple High Performance Horn Design," *2002 IEEE AP-S Symposium Digest* (June 2002) San Antonio TX.
11. J. D. Michaelson and R. B. Dybdal, "Development/Demonstration of a Tactical RDS Terminal," *MAXI 2002 Digest* (October 2002) Silver Spring MD.
12. A. W. Love (ed), *Reflector Antennas* (New York: IEEE Press, 1978).
13. W. V. T. Rusch, Scattering from a Hyperboloidal Reflector in a Cassegrain Feed System, *IEEE Trans Antennas and Propagation*, vol. AP-11 (July 1963): 414–421.
14. C. Dragone and D. C. Hogg, "The Radiation Pattern and Impedance of Offset and Symmetric Near Field Cassegrainian and Gregorian Antennas," *IEEE Trans Antennas and Propagation*, vol. AP-22 (May 1974): 472–475.
15. G. W. Collins, "Shaping of Subreflectors in Cassegrain Antennas for Maximum Aperture Efficiency," *IEEE Trans Antennas and Propagation*, vol. AP-21 (May 1973): 309–313.
16. S. Silver, *Microwave Antenna Theory and Design* (New York: McGraw-Hill, 1949).
17. T. S. Chu and R. H. Turrin, "Depolarization Properties of Offset Reflector Antennas," *IEEE Trans Antennas and Propagation*, vol. AP-21 (May 1973): 339–345.
18. W. A. Imbriale, P. G. Ingerson, and W. C. Wong, "Large Lateral Feed Displacements in a Parabolic Reflector," *IEEE Trans Antennas and Propagation*, vol. AP-22 (November 1974): 742–745.
19. R. Jorgensen, P. Balling, and W. J. English, "Dual Offset Reflector Multibeam Antenna for International Communications Satellite Applications," *IEEE Trans Antennas and Propagation*, vol. AP-33 (December 1985): 1304–1312.
20. C. M. Rappaport and W. P. Craig, "High Aperture Efficiency Symmetric Reflector Antennas with up to 60° Field of View," *IEEE Trans Antennas and Propagation*, vol. 39 (March 1991): 336–344.
21. B. A. Munk, *Frequency Selective Surfaces* (New York: Wiley, 2000).

22. K. Uneo, T. Itanami, H. Kumazawa, and I. Ohtomo, "Design and Characteristics of a Multi-band Communication Satellite Antenna System," *IEEE Trans Aerospace and Electronic Systems*, vol. 31 (April 1995): 600–606.
23. J. Ruze, "Antenna Tolerance Theory—A Review," *Proc IEEE*, vol. 54 (April 1966): 633–640.
24. J. P. Corrigan, "ATS-6 Experiment Summary," *IEEE Trans Aerospace and Electronic Systems*, vol. AES-11 (November 1975): 1004–1013.
25. M. W. Thomson, "The Astromesh Deployable Reflector," *1999 IEEE AP-S Symposium Digest* (June 1999) Orlando FL.
26. A. R. Dion and L. J. Ricardi, "A Variable Coverage Satellite Antenna System," *Proc IEEE*, vol. 59 (February 1971): 252–262.
27. H. E. King, J. L. Wong, R. B. Dybdal, and M. E. Schwartz, "Experimental Evaluation of a Circularly Polarized Metallic Lens Antenna," *IEEE Trans Antennas and Propagation*, vol. AP-18 (May 1970): 412–414.
28. R. J. Mailloux, *Phased Array Antenna Handbook* (Norwood MA: Artech, 2005).
29. F. J. Dietrich, P. Metzen, and P. Monte, "The Globalstar Cellular Satellite System," *IEEE Trans Antennas and Propagation*, vol. AP-46 (June 1998): 933–942.
30. J. J. Schuss, J. Upton, B. Meyers, T. Sikina, A. Rohwer, P. Makridakas, R. Francois, L. Wardle, and R. Smith, "The IRIDIUM Main Mission Antenna Concept," *IEEE Trans Antennas and Propagation*, vol. AP-47 (March 1999): 416–424.
31. T. J. Peters, "A Conjugate Gradient-based Algorithm to Minimize the Sidelobe Level of Planar Arrays with Failed Elements," *IEEE Trans Antennas and Propagation*, vol. AP-39 (October 1991): 1497–1504.
32. ———, Special Issue on: GPS The Global Positioning System, *Proc IEEE*, vol. 87 (January 1999).
33. C. T. Brumbaugh, A. W. Love, G. M. Randall, D. K. Waineo, and S. H. Wong, "Shaped Beam Antenna for the Global Positioning Satellite," *1976 IEEE AP-S Symposium Digest* (October 1976).
34. D. S. Bagri, J. I. Statman, and M. S. Gatti, "Proposed Array-based Deep Space Network for NASA," *Proc IEEE*, vol. 95 (October 2007): 1916–1922.
35. K. M. Soo Hoo and R. B. Dybdal, "Tolerances for Combining High Gain Antennas," *1994 IEEE AP-S Symposium Digest* (June 1994).
36. R. B. Dybdal and K. M. Soo Hoo, "Arraying High Gain Antennas," *2000 IEEE SP-S Symposium Digest* (July 2000).
37. R. B. Dybdal and S. J. Curry, "Arraying Large Antennas," *2007 IEEE AP-S Symposium Digest* (June 2007).
38. R. B. Dybdal, "Coherently Combining High Gain Antennas," *2008 IEEE MILCOM Symposium Digest* (November 2008).
39. R. B. Dybdal, "Antenna Tracking," Chapter 42 in J. L. Volakis (ed), *Antenna Engineering Handbook* (New York, McGraw-Hill, 2007).
40. D. K. Barton and H. R. Ward, *Handbook of Radar Measurement* (Englewood Cliffs NJ: Prentice Hall, 1969).
41. R. B. Dybdal, "Step Tracking Performance and Issues," *1998 IEEE AP-S Symposium Digest* (June 1998).
42. R. B. Dybdal and D. D. Pidhayny, "Rate Corrected Step Track," *2002 IEEE AP-S Symposium Digest* (June 2002), see also R. B. Dybdal and D. D. Pidhayny, "Method of Tracking a Signal for a Moving Signal Source" (May 4, 2004): U.S. Patent 6,731,240.
43. R. B. Dybdal, "Monopulse Resolution of Interferometric Ambiguities," *IEEE Trans Aerospace and Electronic Systems*, vol. AES-22 (March 1986): 177–183.
44. ———, *Electrical and Mechanical Characteristics of Earth Station Antennas for Satellite Communications*, Electronic Industries Alliance/Telecommunications Industry Rept TIA/EIA-411 Revison A (September 1986).
45. P. W. Hannan, "Optimum Feeds for All Three Modes of a Monopulse Antenna," *IRE Trans Antennas and Propagation*, vol. AP-9 (September 1961): 444–461.
46. J. DiTullio, "A High Performance Feed for Earth Station Satellite Communications," *19th Annual USAF Antenna R&D Symposium Digest* (October 1968) Robert Allerton Park, IL.

47. R. B. Dybdal, "Monopulse Antenna Tolerance Errors," *2007 IEEE AP-S Symposium Digest*, June 2007.
48. Y. H. Choung, K. R. Goudey, and L. O. Bryans, "Theory and Design of a Ku-band TE_{21} Mode Coupler," *IEEE Trans Microwave Theory and Techniques*, vol. MTT-30 (November 1982): 1862–1866.
49. Y. H. Choung, "Wideband TM_{01}-mode Traveling Wave Coupler," *Proc IEE*, vol. 144 (October 1997).
50. R. B. Dybdal, "Extended Spatial Acquisition for Tracking Antennas," *1997 IEEE AP-S Symposium Digest* (July 1997); see also R. B. Dybdal, "Extended Spatial Acquisition for Tracking Antennas," (September 14, 1999): U.S. Patent 5,952,962.
51. R. B. Dybdal and D. D. Pidhayny, "Main Beam Alignment Verification," *2004 IEEE AP-S Symposium Digest* (June 2004); see also, R. B. Dybdal and D. D. Pidhayny, "Main Beam Alignment Verification for Tracking Antennas" (August 30, 2005): U.S. Patent 6,937,186.
52. R. B. Dybdal and D. D. Pidhayny, "Polarization Optimization for Tracking Antennas," *2003 IEEE AP-S Symposium Digest* (June 2003).
53. R. B. Dybdal, "Polarization Limitations in Antenna Tracking," *2004 IEEE AP-S Symposium Digest* (June 2004).
54. R. B. Dybdal and D. D. Pidhayny, "Methods and Systems for Tracking Signals with Diverse Polarization" (July 3, 2007): U.S. Patent 7,239,275.

Chapter

3

Communication Satellite System Architectures

3.1 Overview

The overall communication satellite system is comprised of the orbiting space segment payload and the user segment terminals. The system designs for specific applications vary greatly because of differing system objectives and requirements. This trend will continue to expand in future systems as demand increases for additional satellite services and as further extensions are made to provide personal services.

The space segment has several alternative architectures. The most common communication satellite system simply relays information between users using the satellite payload. The space segment is referred to as a transponder, which receives the uplink information from the users and then broadcasts the information to users on a frequency translated downlink. In other cases, the same information is broadcast from the satellites to numerous users, an architecture referred to as a direct broadcast design. Satellite television broadcast designs are a particularly popular example of this architecture. Other services are point-to-point designs, such as satellite crosslinks that provide communications between adjacent satellites without any relay through ground assets. Crosslinks thus provide connectivity to users who do not have mutual visibility to a common satellite. Finally, all satellites require a TT&C (Tracking, Telemetry, and Control) subsystem, another communication system that assists in verifying the satellite's on-orbit position, provides data regarding the satellite's on-orbit health and status using telemetry multiplexed on the TT&C downlink, and controls the satellite's operation through commanding and authentication.

The user segment likewise has varied architectures. Early satellite systems with very limited spaceborne resources needed large ground terminals to achieve adequate performance and disseminated data to individual users via terrestrial links in a "hub and spoke" arrangement. With more capable satellite designs, direct distribution of more information to a wider class of users became possible, a trend that continues to extend communications between individual user terminals. Present system designs include the proliferation of user terminals for direct broadcast services that require cost-effective terminal designs having receive-only capabilities and communication networks using modest-sized terminal referred to as VSAT (very small aperture terminals) to link terrestrial users. These system designs are made possible by increased satellite capabilities and improved user terminal designs. Thus, linking individual users to extend personal communications by using satellite links is accomplished and will be extended in future developments to additional services available to users with handheld equipment. Connectivity with other assets such as terrestrial communication networks, a feature required by the early satellites, continues to be an objective of future systems, particularly in expanding Internet services. This evolution from limited networks and terrestrial relay of low data rate services to individual users found in early systems to an exploding number of user terminals with increased service capabilities is a key reason for existing and future satellite system growth. This same expansion to individual users requires increased emphasis of cost-effective user terminals to control overall system acquisition costs.

The overall satellite communication architectures have many variations for specific program applications, but the number of generic architectural designs is limited. Practical designs vary significantly as required to meet their program objectives; however, the generic designs are repeated to provide channelized services to serve needs over the satellite's field of view. Likewise, a wide variety of user designs exist to support diverse services. Space and user segment architectures are described at a top level and orbital alternatives and their impact on link requirements are discussed.

3.2 Space Segment Architectures

Several architectures are commonly used in the space segment. Three different transponder architectures can be distinguished for space segment applications that provide connectivity between individual users. The first and most widely used design is a linear frequency translating transponder, where the received signals on the uplink are translated in frequency and amplified for retransmission on the downlink, a transponder architecture commonly referred to as a bent pipe repeater. The second type

of transponder design is referred to as the regenerative repeater, where the uplink signals are partially or completely demodulated, reformatted and remodulated, and retransmitted. Some system architectures perform the demodulation, remodulation, and routing on the ground by transmitting the received uplink signals to the ground-based gateway terminals, processing them, and retransmitting the processed signals to the satellite for broadcast to the users on the downlink. The ground segment where the processing is performed is commonly referred to as a gateway terminal.

The advantages of the more complex regenerative repeater transponder are the capability to reduce the downlink broadcast of uplink interference power, thereby increasing the effectiveness of the downlink power; added flexibility in routing information to a variety of destinations; and maintaining user downlink power control. With advances in digital technology, a third type of transponder design is distinguished and future designs will capitalize on digital technology to far greater extents. Digital technology has the potential flexibility to tailor satellite assets to changing demands for satellite resources that evolve over the satellite's lifetime. Memory technology can be used to average out the peak capacity demands by storing data that are not time critical when capacity demands are at their peak levels and broadcasting stored data as capacity becomes available. Digital technology can also be used in frequency translating repeaters where the transponder uses a digital channelizer to separate users received in a given frequency subband, enabling the routing of individual users to different downlink subbands that serve different coverage areas. This user signal separation capability provides added flexibility in routing user traffic.

Other space segment architectures provide services for direct broadcast, point-to-point links for crosslink and earth link capabilities, and TT&C requirements. Direct broadcast services can either broadcast signals received from an uplink transmission, as is the case for satellite TV broadcast, or broadcast information derived onboard the satellite. Point-to-point architectures provide connectivity between satellites in crosslink designs or in earth link services such as satellite-to-gateway communications. Crosslink capabilities allow global communications between users that do not have mutual visibility to a common satellite and in this way additional relays through ground relay stations are not required. Every satellite requires a TT&C subsystem that assists in determining the satellite's orbital location, reports the health and status of the satellite, and provides a commanding capability for satellite operation.

3.2.1 Frequency Translating Transponder

The frequency translating transponder architecture described in Fig. 3-1 is a simple architecture and the most widely used. The uplink signals collected by the receive antennas are preamplified to establish

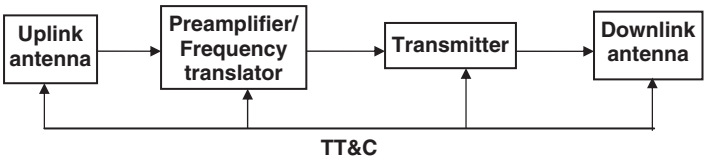


Figure 3-1 Frequency translating transponder

the uplink system noise level. The uplink signal collection is translated to the downlink frequency assignment by a frequency converter(s). In practice, the frequency converter is simply a mixer, a stable local oscillator, a filter to reject the image frequency, and an amplifier. In some cases, the frequency conversion is performed in more than one step. In this way, the uplink frequencies are mapped into the downlink frequencies by a fixed frequency offset following this transponder's name. Suitable amplification is then provided to meet the transmitted power levels. The amplification can be varied by ground command through the TT&C subsystem to optimize the power level of the transmitter.

Satellite transmitters are typically operated near their saturated output level to obtain the most power-efficient transmitter operation. However, the nonlinear behavior of the transmitter not only results in intermodulation products from multiple users sharing the transponder but it also distorts the signal modulation, thus degrading communication performance. The output power of the transmitter must therefore be controlled to achieve the maximum levels consistent with an acceptable degradation in user signal reception. The transponder's power output is controlled by the transponder's analog gain not only to compensate for on-orbit transponder gain variations but also to maintain the desired transmitter operating point under diverse user loading conditions. The TT&C subsystem not only commands the transponder gain setting but also selects antenna beams and pointing directions, commands redundant elements to replace failures, and monitors the health and status via telemetry outputs.

Simplicity is the principal advantage of this transponder and, at the same time, its principal disadvantage. These transponders are easily constructed and thus minimize the very important SWaP in satellite systems. Because of this, the frequency translating design is the most common transponder architecture. A generic operational problem is maintaining power control among the individual users so each user has its equable share of the satellite downlink transmitter. The goal of operating the transponder transmitter near its saturated output level to maximize power efficiency requires power control among the individual users. A frequency translating transponder does not distinguish between

desired signals and interference and its operation can be degraded by strong interference. Finally, since individual users are not separated in a frequency translating transponder, user connectivity in different coverage areas requires dividing the frequency into subbands so connectivity to different coverage areas can be provided.

Transmitters [1] have a limited power output. If the input power levels to the transmitter are allowed to increase, eventually the power output of the transmitter will not increase above a maximum level, commonly referred to as the saturated output power. The amplification at saturation is nonlinear, which degrades communication reception as a result of intermodulation products between multicarrier signals, signal suppression, and signal distortion that includes spectral regrowth of the signals' modulation sidelobes, thus increasing co-channel interference levels.

Two alternative specifications are used by vendors to characterize the signal levels that result in a nonlinear response from their devices. The first alternative specification is a 1 dB compression point defined by a 1 dB deviation from the device's linear gain response at small signal input powers. For example, a typical LNA has a 1 dB compression point when its output level is about 10 to 13 dBm. A second alternative specification is the third-order intercept point. The third-order intermodulation product is determined by exciting the device with two equal amplitude CW (continuous wave) tones and increasing the input power levels of the tones while observing the growth of the intermodulation product level of the tones. The small signal level gain response is linearly extrapolated to regions where the device's power is nonlinear. This linear gain response has a one-to-one correspondence between the output and input signal levels. The intermodulation product has a cubic variation with input power levels. The third-order intercept level is the signal input level where the extrapolated linear response of the small signal gain and the extrapolated cubic response of the third-order intermodulation response intercept and are equal.

In operation, satellite transponders generally support a multicarrier signal collection, and the intermodulation product spectra is more varied than that for two CW tones so that multiple carrier distortion sidebands are not fully addressed. An alternative measurement approach replaces the two CW tones by two separated noise spectra having sufficient filtering to allow observing the linear noise power levels between the two spectra for linear small signal operation. The power levels of the two noise spectra are increased to values resulting in nonlinear operation and the intermodulation spectra between noise components result in increased power levels in the bandwidth occupied by noise power when the system response is linear. The NPR (noise power ratio) is the power difference between the amplified noise level and the power level within the bandwidth between the original two noise spectra. The NPR

characterization may be thought of as a generalization of the standard two CW tone characterization used to obtain the third-order intercept parameter, but a characterization that is more realistic in representing the intermodulation spectra for multicarrier operation. Further evaluation of NPR techniques [2] expresses concern regarding signal distortion characterizations by such techniques and proposes other measures to characterize co-channel distortion effects.

While these alternative specifications of nonlinear device response are commonly used to quantify device nonlinear performance, the impact on communication signals requires a representation of the device's nonlinear response. Nonlinear amplifier characteristics are specified by AM/AM distortion that describes the amplifier's gain compression and the AM/PM distortion that describes the amplitude-phase distortion. In practice, satellite transmitters generally operate at a power output level "backed off" from their nonlinear saturated power output level to maintain operation that is sufficiently linear to control the degradation to user signal reception. The required amount of backoff depends on the amplifier's nonlinear response and the sensitivity of the user's modulation formats to distortion.

The nonlinear response of the amplifier is characterized by measurement techniques that are used to develop nonlinear models to assess the impacts on communication system performance. Vector network analyzer measurements provide one means of measuring the AM/AM and AM/PM parameters as a function of the input signal level. More accurate measurement techniques [3, 4] have been developed and demonstrated to dynamically characterize both solid state and TWT amplifiers. The resulting measurements were used to develop a "3 box" modeling representation of a nonlinear amplifier. The impacts of amplifier nonlinearities on signal reception performance have been addressed [5] and the tolerances for nonlinear amplifier responses to different modulation formats are quantified. Finally, recent development attention has been paid to linearizers [1, 6] that modify the input signals using feed forward, predistortion, or adaptive techniques to allow operation closer to the amplifier's saturated output level. In this way, somewhat increased transmitter power output levels can be obtained while maintaining sufficient linearity to satisfy signal fidelity requirements.

The preceding considerations and design approaches form the basis of establishing a design transmitter operating point that determines the amount of transmitter backoff to be maintained during system operation. Typically, the transponder contains an analog gain control that can be commanded to maintain the desired operating point. This gain control is generally in the drive electronics and establishes the input power to the final transmit amplifier. For the frequency translating transponder, individual user signals are not separated and thus the

input signal power is subject to variations in the uplink user signals as well as the number of system users.

Discipline among the individual users is required to meet the goals of operating the transmitter near saturated conditions and maintaining an equable distribution of the downlink power. Users with higher signal powers within the transponder have a greater share of the limited downlink transmitter power. If an individual user increases its uplink power level to enhance its own link performance, a greater share of the downlink power is available to that user to the detriment of the other users. If the entire user collection increases the uplink signal power or if additional users access the satellite, the uplink input power levels increase. Since the electronics gain through the transponder has a fixed but selectable level, the increased uplink input power level results in transmitter operation more closely located to the nonlinear region. The transponder gain in this case must be reduced by command to maintain the design operating point of the final amplifier in the transmitter. Thus, the transmitted levels among the individual users and variations in the total uplink signal power must be controlled to achieve the proper power balance so each user has its equable share of satellite resources and the transmitter design operating point is maintained to provide as much downlink power as possible while satisfying the linearity required for user operation. The responsibility to maintain the downlink transmitter design operating point rests with the mission control segment's selection of the transponder gain setting.

While discipline among users is required, in some applications, link impairments further confound the problem. As discussed in Chapter 4, EHF (extremely high frequency) systems are impacted by weather, particularly rainfall that attenuates uplink and downlink signals. Similarly, a variety of link impairments impact UHF (ultrahigh frequency) systems. In such cases, user power control techniques are addressed to maintain user link performance. Experiments using the NASA ACTS program [7] provide an experimental demonstration of open-loop techniques. In these experiments, power control over an 18 dB dynamic range was exercised and the power control accuracy during rain events was believed to be about $\pm$2.5 dB. Another approach integrates a radiometric sensor into a frequency hopped EHF system [8] to provide a cost-effective means of power control. Finally, a satellite beacon system for UHF communications [9] and a user receiver using this beacon signal have been proposed to provide a real-time indication of the severity of several link impairments, to indicate communication capabilities, and to provide the basis of power control techniques.

Strong interference is an even more severe problem for frequency translating transponders and degrades their downlink performance in several ways. Interference obscures desired user signals that occupy

the same spectra as the interference degrading user reception. Since the transmitter is operated close to its saturated level under interference-free conditions, strong interference drives the transmitter into its nonlinear region, suppressing the desired signals as well as distorting them and generating undesired intermodulation products. Operation at the transmitter's design operating point can be restored by selecting an appropriate transponder gain level to allow user reception in spectral regions not obscured by interference. However, downlink resources are wasted because interference power is being transmitted, reducing the downlink power available to user signals. Thus, a fundamental limitation of the frequency translating transponder design is its susceptibility to interference signals.

Aside from the inherent problems of user power control and interference susceptibility, an advantage of frequency translating transponders in addition to their design simplicity is that it is transparent to its use. User frequency assignments within individual transponder bandwidths can be varied, multiple access techniques among the system users can be changed, and different modulation formats can be used. Thus, the frequency translating transponder has a significant amount of flexibility in the way in which its resources are used during the satellite's lifetime.

The basic transponder design has been applied to a variety of program applications and differing frequency plans that divide the allocated frequency band into different channels and coverage areas that subdivide the satellite's field of view into different regions served by the antenna system design. Small antenna beamwidths, orthogonal polarizations, and division of the bandwidth into subbands isolate separated coverage areas within the field of view. The isolation between coverage areas allows reusing the same frequency subbands for independent data streams, thus increasing the throughput capacity and efficiently using the frequency allocations. These techniques are referred to as frequency and polarization reuse, respectively and will be discussed in more detail later.

Practical satellite systems provide coverage to different portions of the available field of view, and use the frequency and polarization reuse techniques to isolate communications between coverage areas. Generally, communication connectivity with users is required not only within the same coverage area but also between coverage areas. As a result the available frequency allocations are divided into subbands to provide the connectivity flexibility. A simple example in Fig. 3-2 illustrates these techniques where two coverage areas, A and B, are located within the field of view. The allocated bandwidth in this example is divided into four subbands as illustrated in Fig. 3-3, where A1 and B1 provide connectivity within their respective coverage areas and A2 and B2 provide connectivity between the two coverage areas. A functional description of the transponder for this case is illustrated in Fig. 3-4 describing the subband channelization.

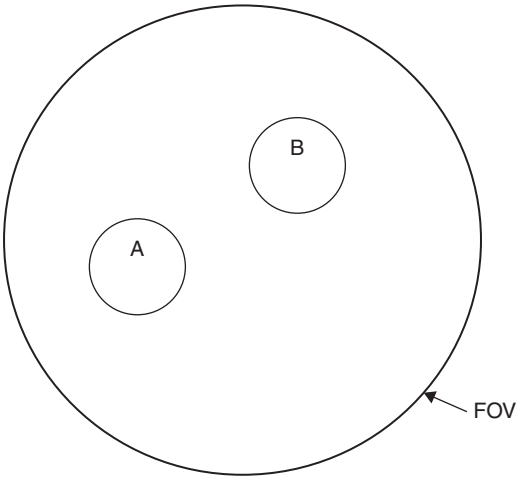


Figure 3-2 Example coverage areas

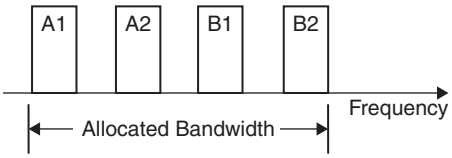


Figure 3-3 Example Frequency plan

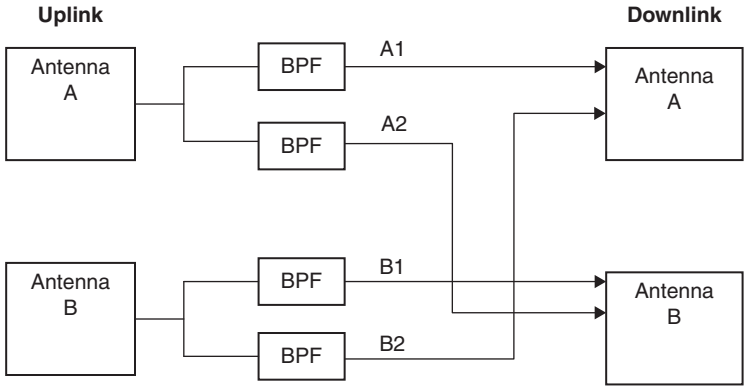


Figure 3-4 Channelized transponder function diagram

The INTELSAT VII/VIIA transponder architecture [10] describes a representative commercial design. This transponder operates in two frequency bands, C- and Ku-band. Both frequency reuse and polarization reuse techniques are used to expand system capacity. A total of

38 C-band transponders and six Ku-band transponders are used in this design that illustrate the proliferation of the frequency translating transponder architectures within a single satellite to provide communication capabilities to the various coverage areas. The antenna system design provides global, hemispheric, zone and spot beams. The antenna design is based on reflector technology and design attention has been paid to polarization purity to maintain isolation in the polarization reuse plan. The C-band antennas provide hemispheric and global coverage areas by combining 88 feed horns, while the Ku-band antennas provide steerable spot coverage areas. Both frequency and polarization reuse techniques (described more fully in Chapter 6) are used by this transponder architecture to use the available frequency allocation efficiency by using the same frequency simultaneously in different coverage areas and in orthogonal polarizations. The ability to service different coverage areas independently and simultaneously and to operate using orthogonal polarizations illustrates the significant role that payload antennas play in modern satellite systems. The transponder architecture described in this reference illustrates the use of subbands and polarization and frequency reuse techniques that increase the flexibility to provide communication networks linking different coverage areas.

3.2.2 Regenerative Repeater Transponders

The second type of transponder design, the regenerative repeater illustrated in Fig. 3-5, addresses some of the limitations of the frequency translating transponder architecture at the expense of increased implementation complexity. Like the frequency translating transponder, the system noise temperature is established by the preamplifier and frequency conversion is used. This type of transponder demodulates the uplink signal collection to separate individual users. The demodulated signals within this collection are individually routed and multiplexed with other uplink signals to form downlink channels that are remodulated, amplified, and retransmitted. Thus, the uplink information is "regenerated" for downlink transmission. In some cases, the demodulation within the transponder may not be completely done. For example, in systems using spread spectrum modulation [11], the transponder might

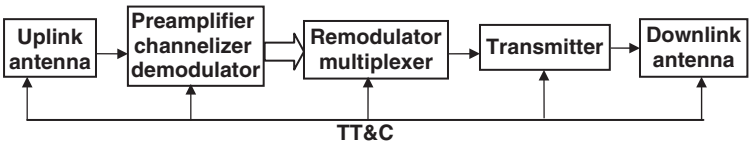


Figure 3-5 A regenerative repeater transponder

only despread the uplink signals rather than fully demodulate them. These despread signals can be remodulated, respread, and retransmitted. Any uplink interference is despread and thereby reduced by the spread spectrum processing gain to reduce the radiated downlink interference radiated power. The reduced downlink interference power results in an increased amount of downlink power being available to the desired users. Variations of this transponder architecture and their benefits have been known for a long time; however, implementation of these architectures awaited the technology development of suitable payload capabilities and manufacturing techniques for the required processing equipment within reasonable weight and power restrictions. The capabilities rely on ASIC (Application Specific Integrated Circuit) and MMIC (Monolithic Microwave Integrated Circuit) technologies to achieve practical designs with reasonable SWaP.

The regenerative repeater transponder architecture addresses the principal disadvantages of the frequency translating transponder and provides routing flexibility for individual users and reduced susceptibility to interference. Since the individual users are separated in the transponder, flexibility to route individual users to desired coverage areas and power leveling for each user can also be performed within the transponder so each user obtains a fair share of the satellite downlink. The interference power is reduced by the demodulation process prior to transmission so that the downlink resources are not wasted by transmitting interference power. Thus, the individual users enjoy reduced interference through processing and receiving more of the downlink power since interference power is not being transmitted. In future system designs for personal communications, some of the potential channels may not be occupied continuously, and by demodulating and retransmitting only the occupied channels, the downlink transmitter resources are not wasted by transmitting the noise power in unoccupied channels.

While the regenerative repeater transponder addresses the inherent limitations of the frequency translating transponder design, the regenerative repeater is inherently more complex. This inherent complexity translates into additional weight and power requirements for the satellite transponder, which is not desirable. In addition, the inherent flexibility of the frequency translating transponder designs in using the channel bandwidths with different multiple access techniques, modulation formats, and number of users is limited by the regenerative repeater design. When the demodulation and remodulation circuitry of the regenerative repeater is implemented using analog technology, options to change modulation, multiple access formats, and so on (discussed in Chapter 4) during the satellite's lifetime are unavailable. Digital implementations provide potential flexibility to alter modulation and multiple access schemes on-orbit. The additional complexity of

the regenerative repeater transponder is therefore justified when the system requirements demand protection from interference or when the flexibility is required in routing individual users to a variety of coverage areas.

The transponder design for the regenerative repeater illustrated in Fig. 3-6 indicates the manner in which connectivity is provided for users within the same coverage area and for users in different coverage areas. Again the same example coverage areas illustrated in Fig. 3-2 will be used and isolation between coverage areas will again be provided by frequency and polarization reuse plans. Each antenna separates its received signal collection into individual user data streams and demodulates those data streams. A switch network routes the demodulated data streams to the appropriate downlink beam where the collection of multiplexed data streams is remodulated for downlink transmissions. In comparison to the frequency translating transponder architecture discussed earlier, the regenerative repeater transponder requires dividing the allocated frequency into subbands for frequency reuse purposes. Unlike the frequency translating transponder, subbands are not required to provide user connectivity between different coverage areas. The reduced number of subbands provides some simplification of the transponder's analog circuitry at the expense of channelizers, demodulators, switch matrices, multiplexers, and remodulators. The advances in digital technology anticipated in future years will make the regenerative repeater architecture more attractive.

The ITALSAT multiple beam transponder [12] is an example of the regenerative repeater architecture and services six different coverage areas covering the Italian nation. The frequency plan provides six subbands to isolate the beams and each beam has a 110 MHz bandwidth. The six coverage areas are produced by using two reflector antennas and each antenna provides three independent beams. The system operates at Ka-band frequencies, and the antenna beamwidths for the service areas are sufficiently small that active tracking of ground beacons is required to maintain coverage alignment in the presence of satellite attitude variations. The uplink data is demodulated onboard and

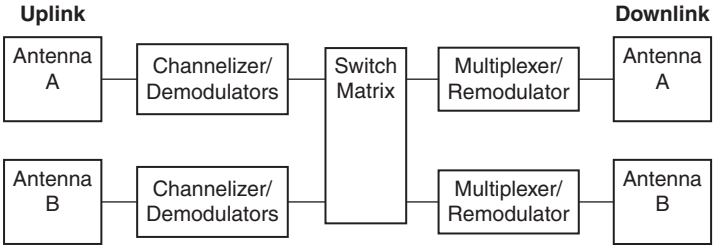


Figure 3-6 Regenerative repeater architecture for multiple beams

switched and routed to the appropriate service area. This system uses a spread spectrum time division multiple access waveform and user assignment to the satellite resources so that high capacity and flexibility is achieved. Other examples of regenerative architecture are used in military applications to achieve high levels of interference protection.

3.2.3 Digital Transponder Designs

Digital technology has progressed to the point that sufficient technology is available to present serious alternatives to traditional analog implementations. The preceding transponder architectures are based on analog design concepts. While digital systems can be configured as replacements for analog designs, the additional flexibility afforded by digital circuitry, the differences in design philosophies between digital and analog circuitry, and interfaces with other digital technologies such as memory devices result in transponder architectures that fundamentally differ from their analog counterparts. These differences justify a separate discussion of transponders that are implemented with digital technology.

The functional block diagram of a digital transponder in Fig. 3-7 assumes array antennas are used with digital beamforming techniques. Digital beamforming techniques [13, 14] have long been known. Practical applications in the past have been limited by digital technology capabilities, and other system sensors such as sonar that have sufficiently small bandwidth requirements that could be accommodated by earlier digital technology have capitalized on digital beamforming. The substantial increase in digital technology capabilities results in communication satellite system designs based on digital technology. Additional application of digital technology in future system designs is apparent.

The transponder design in Fig. 3-7 processes the analog outputs of the receive array elements using analog technology to provide a suitable input to the A/Ds (analog-to-digital converters) that convert

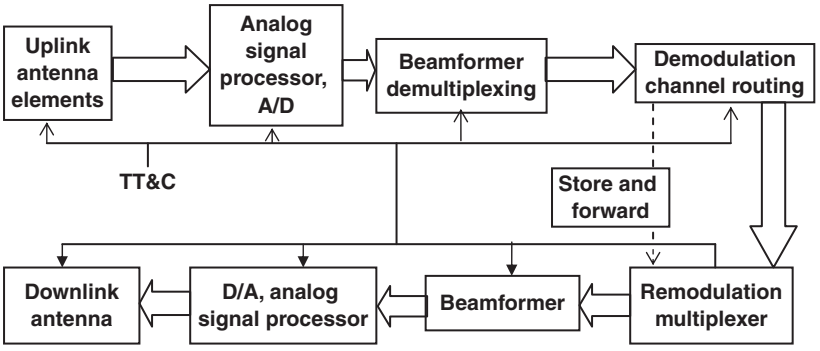


Figure 3-7 A digital transponder

the analog array element input signals to the digital domain. The analog processing includes preamplification to establish the system noise temperature and frequency conversion to the spectra where the A/D transforms the analog data received by each array element into a digital data stream for each array element. The collection of digitized baseband signals from each receive array element is processed by two-dimensional Fourier transforms to produce the antenna array patterns. The individual users in these beams can be separated by digital demultiplexers and, if desired, demodulated. Digital demultiplexing and multiplexing [15] may be accomplished by capitalizing on hardware normally used to implement FFT (fast Fourier transform) operations. The individual data streams are remodulated if required and routed to multiplexers. A two-dimensional Fourier transform is then used to produce the digital excitations for the downlink array elements. Analog signal processing is applied and is comprised of the D/As' analog upconversion to the RF transmitter frequency, and amplification to the required array element power level and then radiation by the array. For narrow bandwidth applications, complex matrix multiplication of multiplexed samples of the same information stream [16] can be performed in place of the two-dimensional Fourier transform to perform the beam steering in the same manner as analog phase shifters.

In some cases, the digital processing functions would be performed within the satellite transponder, while in other cases, the uplink signal collection is relayed to a gateway ground terminal where the processing is performed and the processed data are transmitted to the satellite for rebroadcast to users. Onboard processing within weight and power restrictions of practical space segment designs is limited to applications with limited bandwidth and user channels. Further digital exploitation in space segment transponder designs can be anticipated as digital technology continues to mature.

Another digital application applies digital memory technology to store selected uplink signals for later rebroadcast to the user terminals. A network description of this technique [17] illustrates potential implementations. Two distinct applications exist for such a capability. The first application stores uplink signals when downlink capacity is not available and rebroadcasts the signals at a later time when the capacity is available. In this way, the peak demands for space segment capacity are averaged to allow efficient satellite operation. A second application applies to satellites in a low-altitude constellation [18]. When users of such systems do not have mutual visibility and crosslink resources are not available to relay their communications between satellites, memory technology can be used to rebroadcast the information when the satellite is visible to the receiving user in a subsequent part of the orbital trajectory.

One example of RF digital beamforming designs [19] has been reported that illustrates the hardware implementation and achieved performance. In this case, multiple beams produced by digital beamforming have been used with a mobile user platform to maintain alignment with an L-band satellite. The array is comprised of 16 elements arranged in a 4×4 geometry. In this experiment, FPGA (Field Programmable Gate Array) digital devices were used for the digital processing. An adaptive processing technique was employed to implement the beamsteering and adaptive co-channel interference cancellation as proposed in this design. In operation, the satellite signal alignment was performed by summing the in-phase and quadrature power levels and using the beam with the highest level. The array was mounted on a van and driven in an urban area. In this way, several beam positions were used to maintain alignment with the satellite as the van maneuvered through an urban area and the effects of shadowing by overpasses and buildings were observed in the received signal levels. Additional development and technology experience with digital beamforming can be anticipated.

Digital transponder designs and architectures contain inherent implementation tradeoffs that address critical system parameters such as weight and power requirements. For example, the architecture in Fig. 3-7 performs beamforming over the operating bandwidth and then demultiplexing (or multiplexing). This approach advantageously results in potential power savings in systems that permit turning off beams when users are not present. Another alternative is to demultiplex (or multiplex) individual frequency channels at each array element and perform narrow bandwidth beamforming on the multiplexed channels. Tradeoffs in power consumption in these two approaches depend on the number of channels and the user traffic models.

Two implementation alternatives exist for the digital hardware technology. One alternative uses FPGA technology that provides flexibility in configuring the processing. The second alternative uses custom ASIC technology, whose design is tailored to meet the system's design requirements. This second alternative is more costly than the first but results in lower power consumption. The tradeoffs between these two approaches depend on the specific requirements of the application, cost differences, and required schedule impacts. When ASIC technology is developed, sufficient time should be allowed to produce two ASIC versions so that any required updates to the first version can be produced in the second version for flight application.

A surprising variety of digital technology for transponder applications is available and progress in developing more capable digital technology will continue in the future. Both the sampling rates and the quantization levels of digital hardware can be expected to increase in future

years and decreased power consumption can be anticipated. Typically, the sampling rates are selected to provide 2.5 oversampling. The digital quantization is selected to accommodate the uplink signal collection dynamic range. Each bit of quantization provides 6 dB of dynamic range so that 10-bit quantization provides a 60 dB dynamic range. The effective number of bits is somewhat lower because the least significant bits need to correspond to signal input levels somewhat below the thermal noise level so quantization noise does not impact the system noise temperature. The dynamic range must account for not only the power variations anticipated in the signals received from the users but also the possibility of interfering signals. The effects of interference on A/D converters used in digital beamformers [20] illustrate antenna pattern distortion that results in high-level interference. The system design needs to address the analog gain distribution so the appropriate signal levels are present at the A/D inputs and so sufficient analog anti-aliasing filtering is provided, insuring that out-of-band interference does not degrade the desired signal reception.

Digital implementations need to be compared with analog alternatives to develop systems with acceptable SWaP. For example within the generic transponder design in Fig. 3-7, the alternatives of analog beamforming and digital beamforming need to be addressed particularly in regard to SWaP. Digital beamforming techniques have manageable array complexity levels for the wide field of view requirements for low-altitude satellite designs. In such applications, the overall aperture size is modest and an array design can be implemented with a reasonable number of elements, complexity, and power consumption. For geostationary satellites, however, the array complexity becomes excessive, and analog antenna designs are more appropriate. Digital processing in the form of channelizers to separate and route individual users is a reasonable alternative to pursue. Transponder applications of digital technology are rapidly developing at this writing, and far greater use of these technologies and design implementations can be anticipated in the future.

The flexibility of future transponder designs imposes greater burdens for managing the space segment resources that greatly exceed the requirements of routine commanding as used by earlier satellites. An adaptive control technique [21] in Fig. 3-8 has been proposed as a means of managing the resources. The uplink and downlink antenna designs provide user communication capabilities and the uplink collection of individual user signals is demultiplexed and routed to downlink destinations where the user signals are multiplexed for downlink transmission. The digital technology provides routing complexity unavailable in fixed analog subbands to respond to varying communication needs and traffic variations. Digital technology provides the flexibility to vary

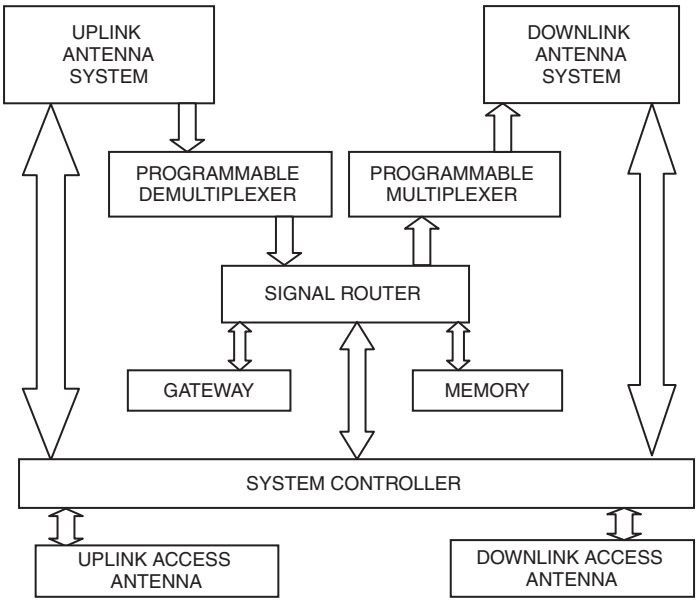


Figure 3-8 Adaptive transponder control [21]

the transponder resource used during the satellite's lifetime to accommodate changing program objectives. It is anticipated the opportunity to upload software upgrades to capitalize on the design flexibility would be exercised. Gateway use and the allocation of onboard memory are also part of the payload management. Finally, user access throughout the satellite's field of view is provided for user requests for satellite availability and notification of resource allocations to the users. The system controller manages the onboard hardware assets and system resource allocations in both an autonomous and commandable fashion. Future designs can be anticipated to depend on such adaptive control techniques to manage system operations.

3.2.4 Direct Broadcast

A variety of direct broadcast architectures exist that use satellites to relay the same information to a multitude of users. At present, direct broadcast is commonly thought of as relaying television coverage over wide areas. The program information is relayed to the satellite from a single ground terminal and the downlink transmission is received by a variety of users. Other forms of direct broadcast services result in navigation and remote sensing satellites. The information for navigation satellites is derived from an onboard clock and pseudorandom codes that allow users to determine the time difference of arrival values needed to

calculate a navigational solution [22]. The onboard sensors in remote sensing satellites gather data that are broadcast to user terminals to be processed into meteorological data and other remote sensing parameters. The common feature in all cases is that information is delivered to a multitude of users by direct satellite broadcast without any return link from the users themselves.

The direct broadcast transponder design illustrated in Fig. 3-9 is less complex than other transponder architectures. The signal source is well controlled so that power control among users is not an issue. Relatively simple space segment antennas are required to meet the requirements for user broadcast service. User equipment for direct broadcast services requires only a receiving capability simplifying their design requirements and providing opportunities for cost-effective user equipment.

In recent years, the size of user satellite TV terminals has shrunk as satellites with increased performance have come into service. More program alternatives continue to be available to system users. The production volume and development of the commercial user equipment result in a wide variety of cost-effective alternatives for users. A similar situation exists with navigation satellites such as GPS, where low-cost terminals have become widely available. User applications for navigational signals have expanded well beyond the applications envisioned originally. Commercial equipment for precision surveying and airport traffic control are representative example applications that extend far beyond the objectives of the original navigational function. Remote sensing satellites are another class of direct broadcast satellites that have also enjoyed commercial application and user system development for weather data received by inexpensive terminals for low-resolution services to more capable terminals for higher-resolution services. A wide variety of commercial alternatives for user reception are available. These examples will enjoy greater user interest in future years. System planners for direct broadcast services need to provide satellites with adequate ERP performance to allow the development of cost-effective user terminals. As the number of users and available applications continues to increase, the total system acquisition cost results in a greater

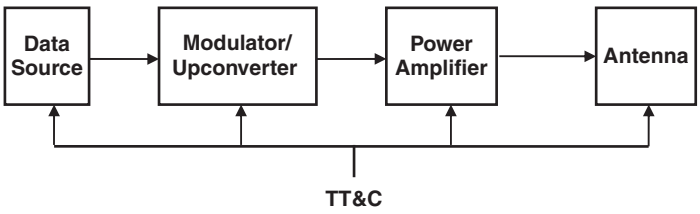


Figure 3-9 A direct broadcast transponder

proportion of the cost in the user segment. Increased space segment performance reduces user performance requirements and cost, as well as provides additional applications for the user segment.

3.2.5 Crosslinks and Earth Links

Crosslink subsystems are point-to-point architectures that provide communication services between two satellites. These systems achieve global user connectivity without the need to relay communications through intermediary ground terminals. This capability provides communications between users that do not have mutual visibility to the same satellite. Generally, satellites have two crosslink subsystems, so connectivity between adjacent satellites is achieved. Crosslink systems commonly use 60 GHz because the oxygen absorption spectra (described in Chapter 4) provide protection from ground-based interference [23]. In addition, the required antenna size for point-to-point services is reduced by increasing the RF frequency, and 60 GHz crosslink designs have compact high-gain antennas, transponder hardware with a well-established technology, and a reasonable weight and size—factors that have resulted in the popularity of 60 GHz crosslink operation. High data rate transfer in crosslink designs generally requires closed-loop antenna tracking capabilities (described in Chapter 2) to compensate for satellite attitude variations.

This subsystem (described in Fig. 3-10) has a relatively straightforward architecture. The communication interfaces are with the satellite transponder and are well defined by the system design. The frequency plan for crosslink operation with geosynchronous satellites requires four independent frequencies to avoid mutual interference when an odd number of satellites are in the constellation and/or when crosslinks are configured between satellites that are across the constellation rather than to adjacent satellites. These frequency plans are based on not allowing a given satellite to have to both transmit and receive operation on the same frequency subband. An example frequency plan in Fig. 3-11 illustrates the connectivity between a constellation of four satellites where four frequencies are required so that connectivity to either adjacent or cross-constellation satellites can be made. Two frequency

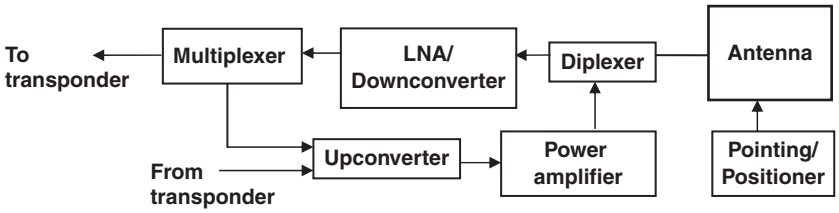


Figure 3-10 A crosslink transponder

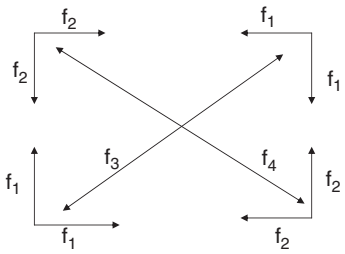


Figure 3-11 An example of a crosslink frequency plan

subbands can provide connectivity between adjacent satellites when an even number of satellites exists in the constellation. Three frequency subbands can provide connectivity between adjacent satellites when an odd number of satellites exist in the constellation. Providing connectivity across the constellation requires a fourth frequency.

Crosslink subsystem antennas for geosynchronous satellites must be capable of principally scanning in the azimuth coordinate and provide minor elevation scanning to offset stationkeeping and satellite attitude variations. An examination of the link geometry in Fig. 3-12 describes the range separation and angle scanning requirements as a function of the angular separation between the satellites in a geostationary constellation. The widest angular separation in this example is 160°, and for this separation, the line of sight between the separated satellites misses the earth's surface by 630 nmi. This azimuth scanning requirement is large and is a ± value because the crosslink

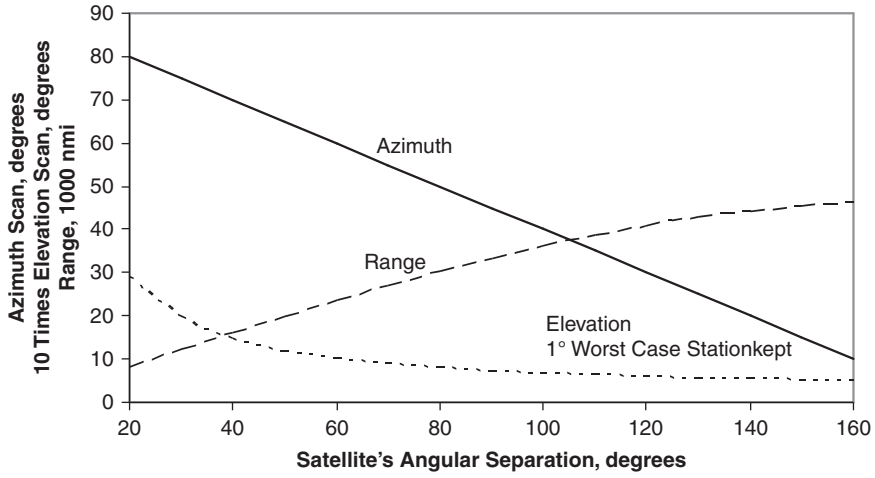


Figure 3-12 Geometric parameters for crosslinks

antenna must be able to "look" both left and right from the spacecraft. If the angular separation between satellites is as close as 20^{o}, then the antenna's scan range is 160^{o} to cover both left and right satellite locations. The elevation scan requirements are much smaller. The elevation scanning requirements (assuming the stationkeeping maintains the satellites within 1^{o}) is only 2.9^{o} for a 20^{o} angular satellite separation and the worst-case orbital phasing for the two satellites. Satellite separation values greater than 20^{o} have smaller elevation angle scanning requirements, as indicated in this figure. The range between adjacent satellites as a function of the satellites' angular separation is also indicated in this figure. These beam scanning requirements form the basis of a crosslink antenna design described in Chapter 6.

Other point-to-point antenna requirements can exist for gateway services and dedicated earth links [24] for high data rate communications. Ground coverage requirements do not apply to these designs since only a single ground terminal is used. Like the crosslink antennas, such point-to-point services require more compact antennas as the operating frequency increases. However, rain attenuation at EHFs (described in Chapter 4) limit the attractiveness of increased RF operation.

3.2.6 TT&C Systems

The TT&C subsystem plays a vital role in every satellite. This subsystem assists in determining the precise location of the satellite in orbit, reports the health and status of the satellite by gathering data from subsystem sensors to form telemetry messages, and commands changes in the satellite operation. The TT&C subsystem is operated in both the satellite launch phase and during its operational life in orbit. During the launch phase, telemetry data are used to determine the correctness of the orbital trajectory, to provide commanding in the orbital insertion as required, and to command the operation of satellite subsystem deployments. During this phase of the program, the TT&C antennas must provide coverage over a complete sphere so that if the satellite begins to tumble during launch ascent, commands can be injected to control the satellite or order its destruction. On-orbit, the TT&C subsystem is the only means of controlling the satellite operation during its lifetime. Thus, the reliability and availability of both space segment and ground segment TT&C system elements' reliability has paramount importance. The TT&C system design purposely provides generous link margins and attention to redundancy.

The space segment TT&C subsystem communicates with satellite mission control terminals that monitor the satellite operation and issue commands to change its operation. With the exception of remote sensing satellites that multiplex their mission data onto the TT&C downlink,

TT&C systems have no direct involvement with the user segment. A generic block diagram of a TT&C satellite transponder is presented in Fig. 3-13. Space segment transponders for TT&C applications [25] now rely heavily on digital technology and thus have the flexibility to serve diverse TT&C applications.

The overall determination of the satellite's orbital position results from complying data from a variety of sources, tracking radars, beacon systems, carrier Doppler variations, and so on, and the TT&C subsystem, which can assist in this process by providing a ranging code. Pseudorandom codes are used to provide accurate range values. The range data and data from the remaining collection of sources are processed using Kalman filtering techniques [26] to obtain a best estimate of the satellite's orbital position defined by the satellite's ephemeris. The ranging code is typically broadcast by the satellite control terminal and converted to the downlink frequency onboard the satellite using dividing and multiplying techniques to maintain code coherence. The ranging information along with angular tracking data from the satellite control terminal antennas provides information used to project the satellite's orbital position.

The telemetry gathered by the satellite's subsystems provides data to monitor the satellite's health and status. This information typically includes satellite prime power data, attitude control data, temperature information gathered from various parts of the satellite, subsystem voltages and current draws, and others. These data, together with any onboard diagnostics, are the only means of identifying causes of operational problems. Therefore, it is critically important early in the development of a satellite to assure that an adequate amount of telemetry points are available, that their response can be clearly interpreted, and that the telemetry link has adequate data capability so that telemetry can be received in a timely manner.

Satellite commanding is the third function of the TT&C subsystem. This commanding has a variety of forms, including deploying subsystems during the early on-orbit launch phase, controlling attitude thrusters, switching to redundant components to replace failed items, changing

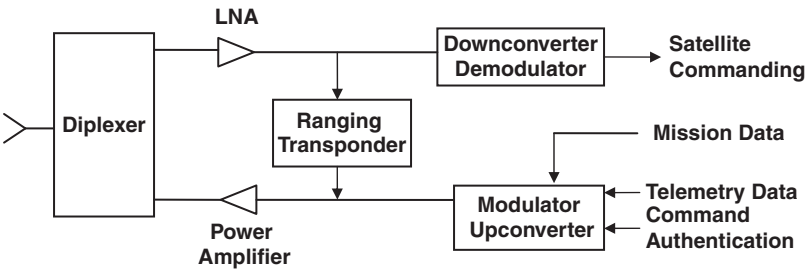


Figure 3-13 A TT&C transponder

antenna coverage by commanding the operation of beamforming networks or repositioning spot beams, and so on. If performance anomalies arise operationally, the TT&C system can command additional telemetry data, including data that are not routinely reported, to provide additional diagnostic data. The commanding must be extremely reliable both in terms of its operation and performance. Command authentication techniques are often used in critical commands to avoid mistaken or misinterpreted commands to the extent practical.

3.3 User Segment Architectures

User segment terminals have had much development and are available in a wide variety of configurations depending on the application and data rates. User terminal designs range from handheld terminals that evolve from technology used in terrestrial applications to very large terminals for high data rate services. Terminal designs are required to satisfy user needs in ground, naval, and airborne applications. Terminals for direct broadcast services do not require a transmitting capability and their design and cost are more modest than terminals requiring a transmission capability. Rather than dwell on terminal designs for different applications, a description of a generic terminal and its operating principles and features is provided.

The user terminal architecture illustrated in Fig. 3-14 describes the most general user terminal design. The terminal is comprised of the antenna, an ACU (antenna control unit), a diplexer that connects the receiver and transmitter to the antenna, a receiver, and a transmitter. User terminals generally use a single antenna for both receive and transmit operations and a diplexer isolates the terminals receiver and

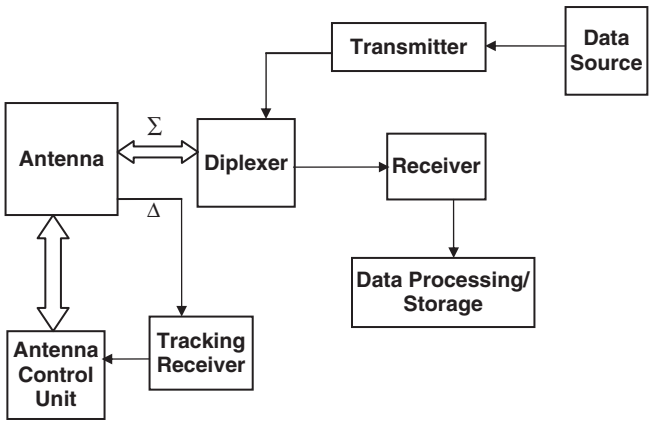


Figure 3-14 A user terminal functional diagram

transmitter. The antenna is positioned by the ACU, which takes a variety of forms depending on the antenna's beamwidth. Personal systems have broad antenna coverage with no pointing requirements. At the other extreme, large ground terminals for high data rate systems use closed-loop tracking systems to maintain antenna tracking, as described in Chapter 2. The ACU subsystem uses the terminals' geographic locations and the satellite's ephemeris information that describes the satellite's orbit to derive nominal antenna pointing information. Signals from the antenna feed are used to align the antenna with the satellite, as described in Chapter 2. Commonly, the ACU also contains star tracking information that is used in radio source antenna calibration techniques (described in Chapter 8).

A diplexer is a device that allows connections between the antenna and the receiver and transmitter while providing adequate isolation so that the transmitter operation does not interfere with the receiver. While such isolation can be achieved by physical separation between the receiver and transmitter used with separate antennas, generally the cost and real estate required for separate receive and transmit antennas are prohibitive. The required diplexer isolation is achieved by providing adequate filtering in the receive and transmit paths. The transmit filter must pass the in-band transmitted signal with minimum loss while suppressing the out-of-band transmitter noise over the receive bandwidth. The receive filter must likewise pass the in-band received bandwidth with minimum loss and suppress the out-of-band transmit spectra sufficiently to maintain linear receiver operation.

The diplexer isolation requirements are derived from the factors illustrated in Fig. 3-15. The filtering requirement in the transmitter path must suppress the transmitter spectra within the receive bandwidth sufficiently to avoid perturbing the receiver noise temperature. The out-of-band transmitter spectra contain noise, harmonics, and spurious components. Commonly, the out-of-band transmitter spectra are required to be suppressed to a level 10 dB lower than the system noise temperature. Such suppression limits the increase in the system noise temperature to about 0.1 dB. Increased receiver noise levels with transmitter operation provide a basis of test requirements to determine the adequacy of the diplexer filter. The receiver noise levels with and without transmitter operation are compared to determine compliance with the diplexer requirements. Similarly, the filtering in the receive path must be sufficient to suppress the transmitter output to avoid exceeding the linear dynamic range of the receiver. Typically, the filtering in the receiver downconversion isolates that portion of the receiver from the transmitted spectra, and attention must be typically focused on maintaining linearity of the LNA and possibly the first downconverter. Generally, sufficient receiver filtering to reduce the

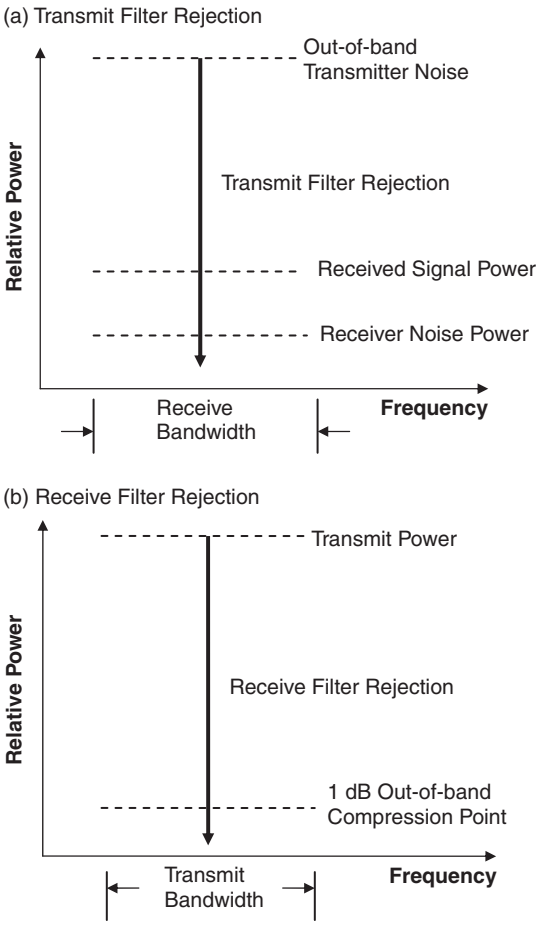


Figure 3-15 Diplexer filter requirements

transmitted spectra to a level 6 to 10 dB below the 1 dB compression point is used in establishing the rejection requirements. The receive filter requirements also must consider anticipated signal levels from other external interference in determining the filter's requirements. Verification of adequate isolation compares the receiver noise level with and without the transmitter operating; typically, the noise level increase with transmitter operation is limited to 0.1 dB.

The receiver provides preamplification, downconversion, synchronization, and demodulation of the received signal. Depending on the system, this process can involve varying levels of complexity. When satellites have multiple access schemes to allow many users to share

the satellite downlink, the means to separate the user of interest is required. Because of the relative motion between the satellite and the user, Doppler variations must be compensated. The larger antennas often use a separate tracking receiver to measure the downlink power for antenna tracking purposes. Digital signal formats and spread spectrum modulation formats have a variety of synchronization requirements that the receivers must satisfy. Digital receiver processing techniques can provide bit and frame synchronization, decoding error correction coding and deinterleaving, and demodulation in software radio implementations.

Similarly, the transmitter requirements depend on the program. The input data must be modulated, amplified, and upconverted to the transmit frequency and amplified to the required output power. The details of this process also depend on the specific system. Synchronization as appropriate to the program must be obtained to avoid interfering with other system users. Modulation formats differ from one program to the next. The transmitter provides the final amplification and its output level is typically varied by adjusting the level of the input signal power to the transmitter so that the transmitter output level does not exceed a threshold value dictated by linearity requirements. Transmitter protection circuitry and power monitoring test ports are typically provided in the terminal design.

In addition to these functional requirements, large ground terminals generally have further requirements to maintain operation when components fail. Larger terminals are often required to operate remotely to the extent practical while being subject to incentives to maintain availability. Such requirements demand attention to effective BITE (built-in test equipment) capabilities to identify component shortfalls rapidly and to command redundant replacements to maintain operation. Significant design attention must be paid to develop systems capable of remote operation and high availability.

3.4 Orbital Alternatives

Most satellite communication programs configure a constellation of geostationary satellites [27] so that the satellite remains stationary with respect to ground users. The ability to remain fixed with respect to the ground users is advantageous in many respects. Systems that service only one geographical area benefit from geosynchronous orbits because fixed antenna coverage is provided to user communities, and the antenna beams are not constantly moving to maintain coverage to fixed regions of the earth. With the exception of the polar regions, three satellites provide almost global coverage. For these reasons, geosynchronous altitude satellites are a traditional choice for communication satellites.

This choice of geosynchronous satellites has its limitations, one of which is that the required satellite altitude be 20,000 nmi. This altitude requires significant performance levels for space and user segment designs. Voice communications using geostationary satellites suffer a roundtrip time delay equal to about a quarter second that some users find objectionable. These limitations have prompted investigations of other orbital alternatives.

Recently, other orbital geometries [19] have been widely considered for personal communication systems. Two orbital classes, LEO (low earth orbit) for altitudes up to about 800 nmi and MEO (medium earth orbit) that are generally above 6000 nmi, are distinguished. The two orbital classes straddle the Van Allen radiation belts where the high ionization results in challenges in protecting electronic components.

Lower satellite altitudes likewise have their own limitations. Since these satellites are not stationary with respect to the ground, the individual satellites are in view of a particular geographic location for a limited time, and the lower the altitude, the less time the satellite is in view of a particular location. If continuous coverage is required, orbital constellations with a large number of satellites are required. Developing orbital constellations to satisfy particular system requirements [28] requires strategies in selecting the orbital inclination values and the number of satellites per orbital plane to minimize the required number of satellites. The economic differences continue to be debated between a large number of small LEO satellites versus fewer large MEO satellites versus three larger satellites in geosynchronous orbits.

An examination of orbital geometries illustrates some of the system tradeoffs. The angular field of view varies with the satellite's altitude, as shown in Fig. 3-16, which indicates significant increases in required

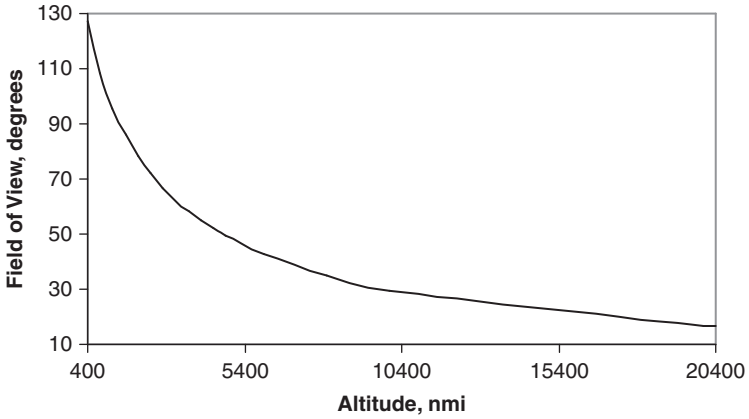


Figure 3-16 Instantaneous field of view versus satellite altitude

angular coverage as the altitude decreases. The angular field of view subtended by the earth, FOV, can be calculated from

$$\text{FOV} = 2 \sin^{-1} (3440/(3440 + h))$$

where the earth's radius equals 3440 nmi and h is the satellite's altitude above the earth measured in nmi. The increased angular field of view for LEO orbits significantly impacts the appropriate antenna technology. If the angular field of view is divided into multiple coverage areas, the beamwidth required to service a given coverage area also increases as the satellite's altitude decreases. The increased beamwidth for coverage areas as the satellite altitude decreases also results in reduced antenna gain levels. The wide field of view requirements are inappropriate for limited field-of-view antenna designs such as reflector technology. Further, as the field of view is divided into multiple coverage regions, the users transverse the coverage areas more rapidly, resulting in increased challenges in transferring users from one beam position to the next as the satellite proceeds along its orbital trajectory. While the field of view increases with decreasing satellite altitude, the instantaneous coverage from a given satellite also decreases. The instantaneous coverage of a satellite normalized to the total surface area of the earth, as shown in Fig. 3-17, illustrates how little of the earth's surface is covered at a given time for low-altitude satellites.

Another parameter illustrated in Fig. 3-18 is the difference in the spatial attenuation between users located at the nadir position beneath the satellite's location and users located at the edges of the instantaneous

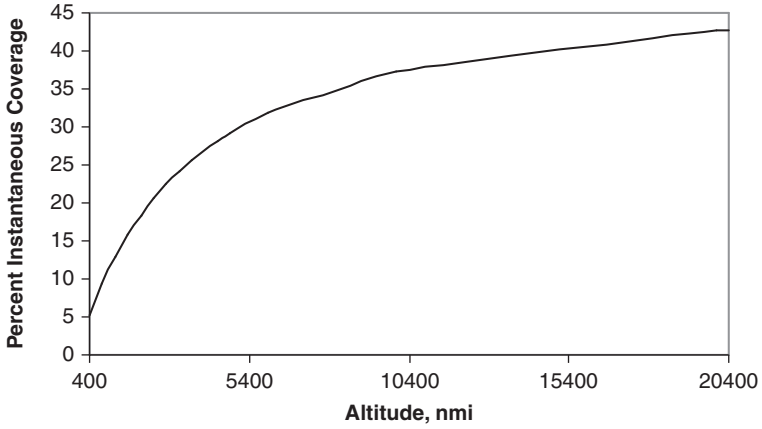


Figure 3-17 Altitude variation of the percentage of the earth's surface area with the field of view

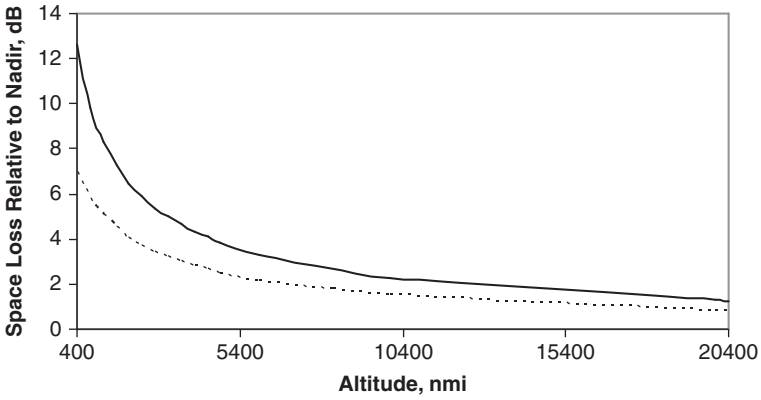


Figure 3-18 Space loss relative to the nadir for the horizon (solid line) and a 20^{o} elevation angle (dashed line)

field of view. The range from the satellite to a user at an elevation angle ε to the satellite is given by

$$R = (3440 + h) \cos (\varepsilon + \alpha)/\cos \varepsilon$$

where α is measured from the satellite's nadir axis and equals

$$\alpha = \sin^{-1} (3440 \cos \varepsilon/(3440 + h))$$

As the satellite altitude decreases, the space loss at the edges of coverage becomes significantly greater than that of the nadir position. A significant difference exists in the spatial attenuation values for users at a 20^{o} elevation angle compared to the spatial attenuation for users at the edge of coverage. Thus, part of the tradeoffs in developing an orbital constellation for low-altitude satellites is defining a constellation based on restricting users to a minimum elevation angle. Low-altitude system designs typically provide services for low data rate requirements. In such cases, user terminal antennas provide broad coverage to meet low data rate link closure requirements. As the user elevation angles to the satellite become lower, additional problems exist in combating multipath and terrain obscuration issues. Thus, two reasons exist to develop orbital constellations where the user locations exceed a minimum elevation angle value. The first reason concerns excessive spatial attenuation values, and the second reason concerns the limitations resulting from multipath and terrain obscuration.

The debate regarding orbital requirements can be anticipated to continue. It is anticipated that low-altitude satellite systems will have a role in low data rate applications where simple cost-effective user equipment

and satellites can be configured for low data rate applications. Higher data rate applications for users located outside of the polar regions can be anticipated to use higher altitude satellite orbital configurations.

References

1. F. H. Raab, P. Asbeck, S. Cripps, P. B. Kenington, S. B. Popovic, N. Potecary, J. F. Sevic, and N. O. Sokal, "Power Amplifiers and Transmitters for RF and Microwave," *IEEE Trans Microwave Theory and Techniques*, vol. 50 (March 2002): 814–826.
2. J. C. Pedro and N. B. de Carvalho, "Characterizing Nonlinear RF Circuits for Their In-band Signal Distortion," *IEEE Trans Instrumentation and Measurement*, vol. 51 (June 2002): 420–426.
3. C. J. Clark, C. P. Siva, A. A. Moulthroup, and M. S. Muha, "Power Amplifier Characterization Using a Two Tone Measurement Technique," *IEEE Trans Microwave Theory and Techniques*, vol. 50 (June 2002); 1590–1602.
4. C. J. Clark, G. Chrisikos, M. S. Muha, A. A. Moulthrop, and C. P. Silva, "Time Domain Envelope Measurement Technique with Application to Wideband Power Amplifier Modeling," *IEEE Trans Microwave Theory and Techniques*, vol. 46 (December 1998): 2531–2540.
5. M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems* (New York: Kluwer/Plenum, 2000).
6. A. Katz, "Linearization: Reducing Distortion in Power Amplifiers," *IEEE Microwave Magazine*, vol. 2, no 4 (December 2001): 37–49.
7. A. W. Dissanayake, "Application of Open Loop Uplink Power Control in Ka-band Satellite Links," *Proc IEEE*, vol. 85 (June 1997): 959–969.
8. R. B. Dybdal, "Radiometer Integrated with Communications Terminal," *1984 IEEE Symposium Digest* (June 1984): 645–647.
9. R. B. Dybdal and M. A. Rolenz, "Satellite Beacon for UHF System Users," *2003 IEEE MILCOM Symposium Digest* (October 2003); see also R. B. Dybdal and M. A. Rolenz, "Method of Determining Communication Link Quality Employing Beacon Signals" (May 13, 2008): U.S. Patent 7,373,105.
10. D. J. Cornelius, A. J. Herridge, R. Silk, and P. T. Thompson, "The INTELSAT VIII/VIIIA Generation of Global Communication Satellites," *Inter Jour Satellite Communications*, vol. 13 (January/February 1995): 39–48.
11. R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications* (Prentice Hall: 1995).
12. F. Carducci and M. Francesi, "The ITALSAT Satellite System," *International Jour Satellite Communications*, vol. 13 (January–February 1995): 49–81.
13. D. E. Dudgeon, "Fundamentals of Digital Array Processing," *Proc IEEE*, vol. 65 (June 1977): 898–904.
14. R. J. Mailloux, *Phased Array Antenna Handbook*, Second Edition (Norwood MA: Artech, 2005).
15. M. G. Bellanger and J. L. Daguet, "TDM-FDM Transmultiplexer: Digital Polyphase and FFT," *IEEE Trans on Communications*, vol. COM-22 (September 1974): 1199–1204.
16. T. Gebauer and H. G. Goekler, "Channel Individual Adaptive Beamforming for Mobile Satellite Communications," *IEEE Trans Selected Areas on Comm*, vol. SAC-13 (February 1995): 439–447.
17. J. P. Havlicek, J. C. McKeeman, and P. W. Remaklus, Jr., "Networks of Low Earth Orbit Store and Forward Satellites," *IEEE Trans Aerospace and Electronic Systems*, vol. AES-31 (April 1995): 543–553.
18. R. Miura, T. Tanaka, I. Chiba, A. Horie, and Y. Karasawa, "Beamforming Experiment with a DBF Multibeam Antenna in a Mobile Satellite Environment," *IEEE Trans Antennas and Propagation*, vol. AES-45 (April 1997): 707–713.
19. F. Ananasso and F. D. Priscoli, "The Role of Satellites in Personal Communication Services," *IEEE Journal Selected Areas on Communications*, vol. SAC-13 (February 1995): 180–195.

20. D. R. Ucci and R. G. Petroit, "The Effects of ADC Nonlinearities on Digital Beamformers," *1989 IEEE MILCOM Symposium Digest* (October 1989).
21. R. B. Dybdal, "Adaptive Control of Multiple Beam Satellite Transponders," *1997 IEEE MILCOM Symposium Digest* (November 1997); see also R. B. Dybdal, "Adaptive Control of Multiple Beam Communication Transponders" (April 25, 2000): U.S. Patent.
22. ———, "Special Issue on GPS The Global Positioning System," *Proc IEEE*, vol. 87 (January 1999).
23. R. B. Dybdal and F. I. Shimabukuro, "Electronic Vulnerability of 60 GHz Crosslinks," *1984 IEEE MILCOM '84 Symposium Digest* (October 1984).
24. R. B. Dybdal and H. J. Wintroub, "An Antenna Based High Data Rate Concept," *1995 IEEE MILCOM Symposium Digest* (November 1995).
25. M. C. Comparini, F. De Tiberis, R. Novello, V. Piloni, L. Simone, D. Gelfusa, S. Cocchi, F. Marchetti, A. Delfino, A Rapposelli, and G. Basile, "Advances in Deep Space Transponder Technology," *Proc IEEE*, vol. 95, (October 2007): 1994–2006.
26. H. W. Sorenson, "Least Squares Estimation: From Gauss to Kalman," IEEE Spectrum, vol. 7 (July 1970): 63–68.
27. J. R. Pierce and R. Kompfner, "Transoceanic Communication by Means of Satellites," *Proc IRE*, vol. 47 (March 1959): 372–380.
28. V. A. Chobotov (ed), *Orbital Mechanics*, Second Edition American Institute of Aeronautics and Astronautics (1996).

Chapter

4

Propagation Limitations and Link Performance

4.1 Overview

Communication system performance depends on the receiver and transmitter capabilities, the propagation medium, and the modulation and multiple access techniques. The propagation of satellite links at microwave frequencies is relatively close to ideal vacuum conditions, but the UHF and EHF extremes are impacted by propagation impairments. A useful compendium of propagation data [1] covering both UHF and EHF characteristics describes predictive techniques to address these impairments. UHF systems are subject to ionospheric effects and the relatively high galactic and manmade noise levels that have a significant impact on system performance. In addition, the relative long wavelengths and reasonable size limitations constrain the benefits that antennas can provide at higher frequencies where a given physical size corresponds to a greater electrical size. Both the normal water vapor and oxygen absorption characteristics within the atmosphere result in moderate attenuation away from the EHF resonant frequencies, while hydrometeors, clouds, ice crystals, hail, snow, and particularly rain have more pronounced effects on EHF propagation. The size of atmospheric hydrometeors becomes a significant fraction of the EHF wavelengths, resulting in atmospheric attenuation and depolarization that can significantly constrain system performance. Lower microwave frequencies are relatively unaffected because the hydrometeors are smaller compared to microwave wavelengths. In configuring system designs, signal modulation formats, coding and interleaving techniques, and multiple access techniques that allow multiple users to share satellite resources must be addressed and their

selection also dictates the capacity of systems. This chapter discusses the propagation issues and the process of link analyses.

4.2 Propagation Limitations

The ionospheric propagation is but one of the link impairments that impact the operation of UHF systems [2, 3]. The long UHF wavelength and reasonable constraints on physical size limits the electrical size of the UHF antennas. For users, the resulting broad pattern characteristics create further impairments arising from multipath and blockage by manmade and terrestrial features in the terrain. In addition, spectral crowding and demands for other services result in interference issues. Additionally, the background noise level at UHF results from not only the natural galactic background (the radiation noise created by the Milky Way) but also manmade noise that continues to increase with levels that exceed the natural galactic background. Thus, in addition to propagation limitations, other factors result in a variety of link impairments that impact UHF systems. For this reason, a beacon has been proposed [4] to provide system users with a real-time indication of the values of individual UHF link impairments and a means to determine the available communication capabilities.

Propagation at EHF frequencies is sensitive to weather. The principal concern is rainfall, but clouds and ice crystals also impair propagation. Three major effects must be addressed. The first is attenuation that both reduces signal levels and produces additional noise. The second is depolarization that results since hydrometeors are not ideally spherical. Cross-polarized levels increase as rain rates increase, and for systems that operate using orthogonal polarization and have high rain margins, performance is degraded by not only signal loss but also cochannel interference resulting from increased cross-polarization. The third factor is the "wet antenna" problem, which is often overlooked in link analyses. A significant effort to understand these limitations and develop predictive techniques has evolved over the last 40 years.

4.2.1 Ionospheric Limitations

The ionosphere extends from roughly 60 to 600 miles above the earth's surface and is characterized by electron density profiles that vary with altitude, time of day, and solar activity. The electrons interact with the earth's geomagnetic field and result in an anisotropic media causing propagation to depend on orientation with magnetic field lines. At very low frequencies (e.g., HF), the ionosphere is completely reflective, thus precluding propagation between satellites and earth-based users. At higher UHF frequencies, the ionosphere is no longer completely reflective,

but ionospheric effects still limit propagation [5, 6]. The electron density increases in daylight hours because of the ionizing effect of the sun.

Irregularities in the electron density profiles produce scintillations that most often occur at night in equatorial regions, less often in polar regions, and seldom in midlatitude regions. The irregularities tend to be aligned with the geomagnetic field, and their scale size is about 1 km transverse to the field and is elongated along the field. The scintillations resulting from the irregularities depend on the frequency of operation and the magnetic field activity. Their frequency variation tends to follow $\alpha\ \lambda^n$ dependence, with n varying between 1.5 and 2 so that ionospheric scintillation impacts decrease with increasing frequency.

The ionosphere is an anisotropic plasma resulting from the electron density and the earth's geomagnetic field. As a consequence, propagation has two circularly polarized modes, each having different phase velocities. If linear polarization is used, the signal is coupled into both of the circularly polarized modes and because of the different phase velocity of the modes, the polarization of the linearly polarized field rotates as it transverses the ionosphere. This rotation is referred to as Faraday rotation and the number of rotations depends on the frequency, increases at low elevation angles because of the longer path through the ionosphere, and varies with the day and night ionization conditions. Consequently, circular polarization is generally used to avoid the uncertain orientation of linear polarization caused by Faraday rotation that requires continuously tracking the orientation of the linear polarization.

Other factors limit operation at UHF. The frequency allocations necessarily have a limited bandwidth compared to other alternatives, thus UHF is constrained to supporting users with low data rate requirements. The long wavelengths in comparison to microwave frequencies result in the inability to construct highly directive antennas within practical size constraints. Spot beam designs for the space segment result in inordinately large antennas. Similarly, directive user antennas are constrained by physical size, and the resulting broad coverage of practical antenna designs has an inherent susceptibility to multipath that further degrades communication performance. Increasing the user elevation angle reduces multipath and terrain blockage effects; therefore, configuring satellite constellations based on a minimum elevation angle requirement for users is prudent. Another limiting factor is the high noise background. The noise levels over a wide frequency range, as shown in Fig. 4-1 [1], illustrate this factor. Two distinct noise contributions exist: one from the natural galactic background and the second from manmade sources. The galactic background levels vary with solar activity and the increased manmade noise level increase in urban environments substantially exceeds the natural background. For example, at 137 MHz, a frequency used by low-resolution meteorological services, manmade noise levels of several

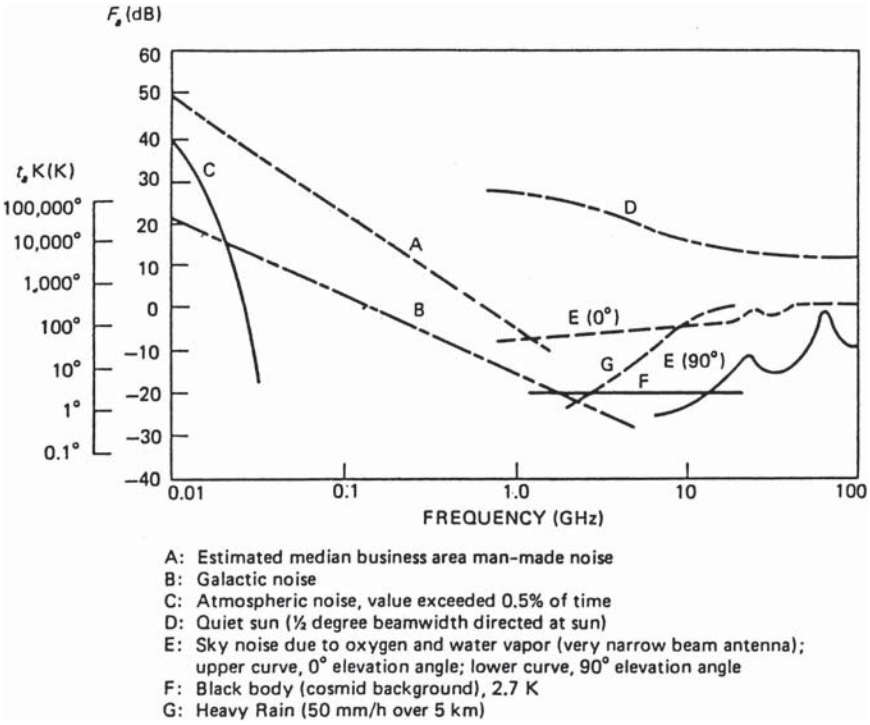


Figure 4-1 Representative noise levels [1]

thousand Kelvin are commonly encountered, thus the system noise temperature is dominated by external noise. As discussed in Chapter 1, VHF and UHF receiving systems can tolerate significant filtering loss since, within reason, both signal and noise are equally attenuated. Such filtering is required because the heavy usage of these frequencies for other services imposes significant interference problems. While these link impairments constrain operation and pose reliability concerns, the cost of equipment is relatively modest. Thus, some applications such as the transfer of low-resolution meteorological data find this lower frequency range useful.

4.2.2 EHF Limitations

The EHF frequency range has limitations resulting from molecular absorption and scattering from hydrometeors. The "clear" background is impacted by molecular absorption. Electric dipole resonances of water vapor and resonances of the magnetic dipole spin states of oxygen are present within the EHF range. Away from the resonant frequencies, their response persists over a broad frequency range and results in

atmospheric propagation loss. Liquid water in the form of rain, clouds, and fog has a major impact on EHF propagation, and frozen water in the form of ice crystals, snow, and hail has lesser effects. The liquid and frozen water particles are commonly referred to as hydrometeors. The size of atmospheric hydrometeors becomes a significant fraction of the EHF wavelength, resulting in atmospheric attenuation and depolarization that significantly constrain system performance. The longer microwave wavelengths are much larger than hydrometeors and propagate with relatively little effect. A further factor at EHF frequencies that is sometimes overlooked is the "wet antenna" problem. The clear propagation, rain attenuation, depolarization, and wet antenna issues are discussed in turn.

Molecular absorption impacts propagation in both clear and inclement conditions. An overview in Fig. 4-2 [7] presents the molecular absorption characteristics resulting from electric dipole resonances of water vapor and the magnetic dipole spin state resonances. These data present the specific attenuation (e.g., dB/km) and indicate the resonant responses of water vapor and oxygen, and the attenuation at frequencies has effects well away from their resonant responses. The resonances of particular interest for satellite communications are the water vapor line at 22 GHz, which is comparatively weak; the oxygen absorption at 60 GHz and the single oxygen line at 118 GHz, and the higher water vapor absorption line at 183 GHz. Between these absorption lines, the atmosphere is relatively transparent, and communication systems using 44, 30, and 20 GHz not only have little degradation from molecular absorption but also have allocated frequencies. At frequencies below the resonant values, relatively little attenuation exists from the molecular absorption. However, this attenuation results in the noise contributions to user uplink antenna noise temperature as will be discussed.

The specific attenuation in Fig 4-2 together with atmospheric models allows determination of the attenuation for satellite links. The zenith attenuation for various altitudes [7] is presented in Fig. 4-3. Attenuation values for other elevation angles are obtained by dividing the zenith attenuation by sin ε, where ε is the elevation angle of the path. This is a good approximation for elevation angles greater than 10^{o}. Further modeling of these molecular absorption characteristics [8] has been done for frequencies up to 1000 GHz. This model was used to examine the zenith (90^{o} elevation angle) attenuation of the oxygen absorption spectra as a function of altitude [9] in Fig. 4-4. At low altitudes where oxygen is abundant, a broad spectra exist and the resonances are filled because of pressure broadening. At high altitudes where oxygen is tenuous and atmospheric pressure does not broaden the lines, the individual magnetic spin state resonances are evident. The attenuation values clearly preclude reasonable communications through the atmosphere. However, the attenuation afforded by the oxygen absorption spectra

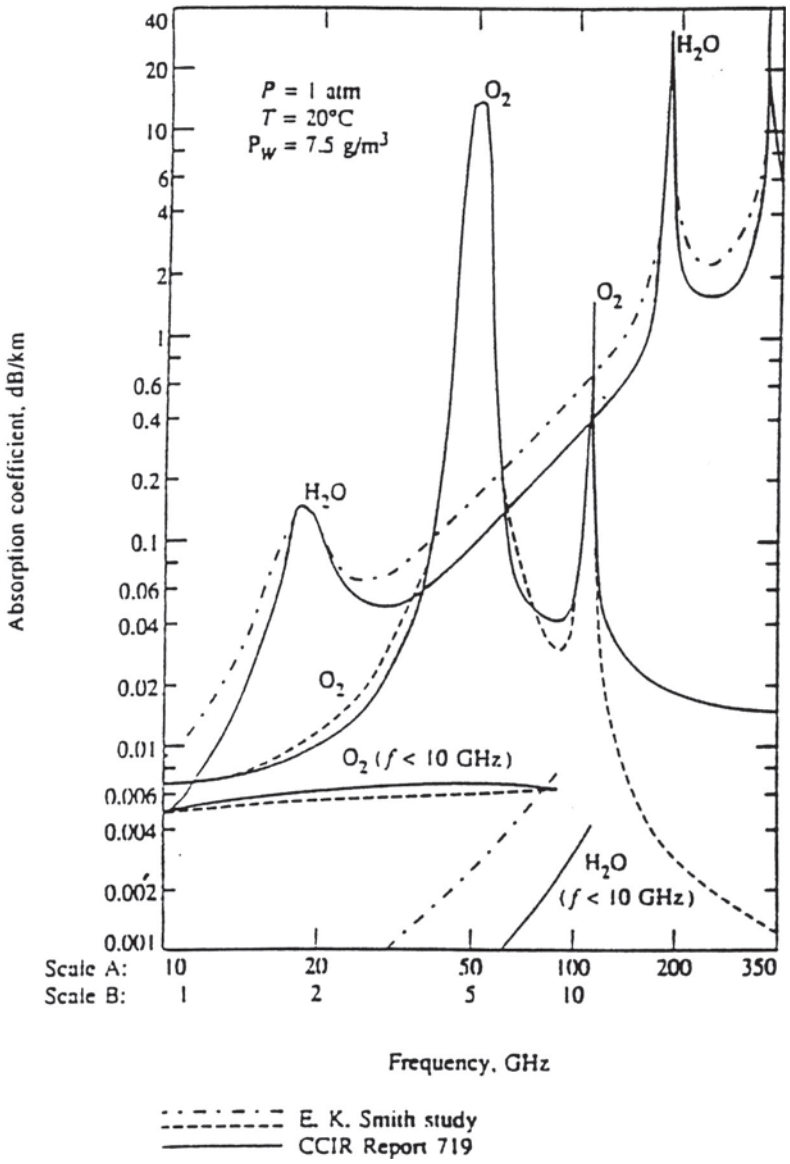


Figure 4-2 Specific attenuation values [7]

also isolates satellites from ground-based interference, one reason for the popularity of 60 GHz satellite crosslinks.

The propagation loss not only attenuates signal levels but also generates noise like any other loss. The molecular absorption loss results in one component of the antenna noise temperature; other components

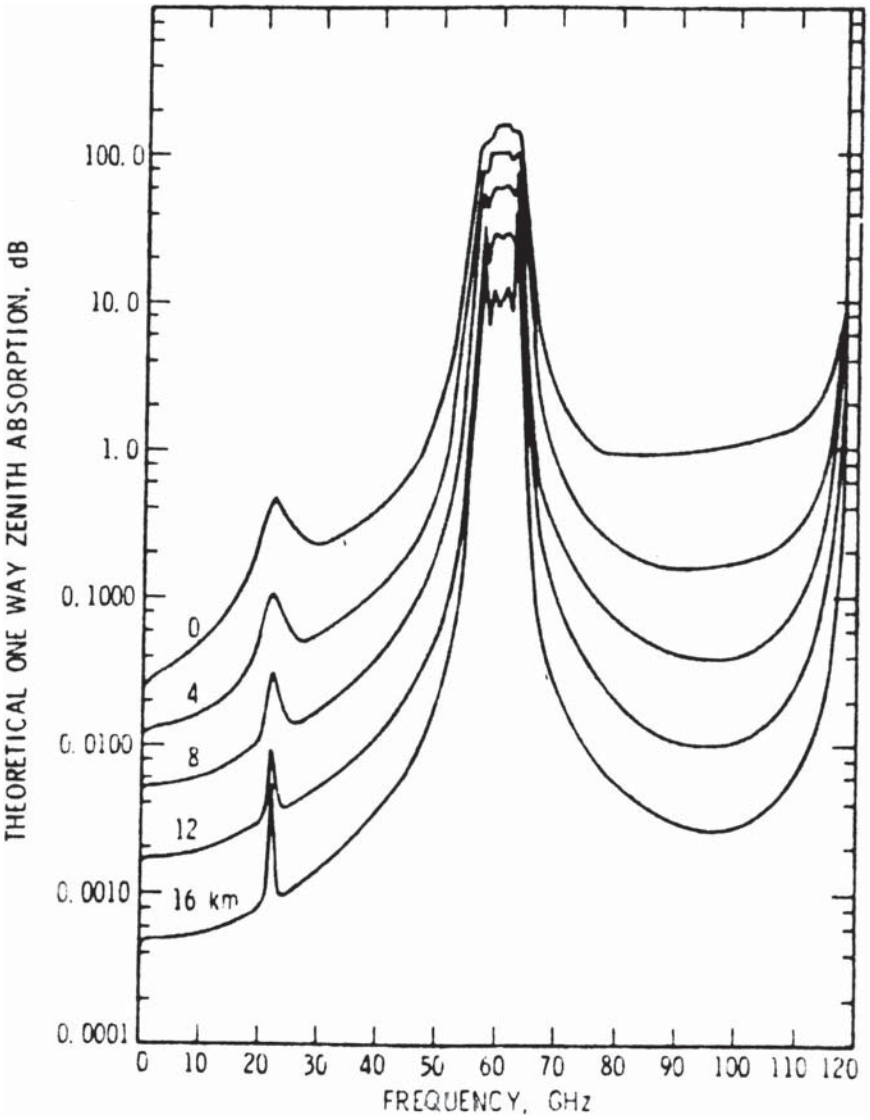


Figure 4-3 Zenith attenuation characteristics [7]

include ground emission and ohmic loss, as described in Chapter 1. The noise contribution from the molecular absorption is expressed as sky temperature values (described in Fig. 4-5 [7]) for various elevation angles. The sky temperature is the noise temperature an ideal lossless antenna with an infinitely small beamwidth would receive due to molecular absorption. These values, shown in Fig. 4-5, are exclusive of galactic and manmade emissions at the lower frequencies (illustrated in

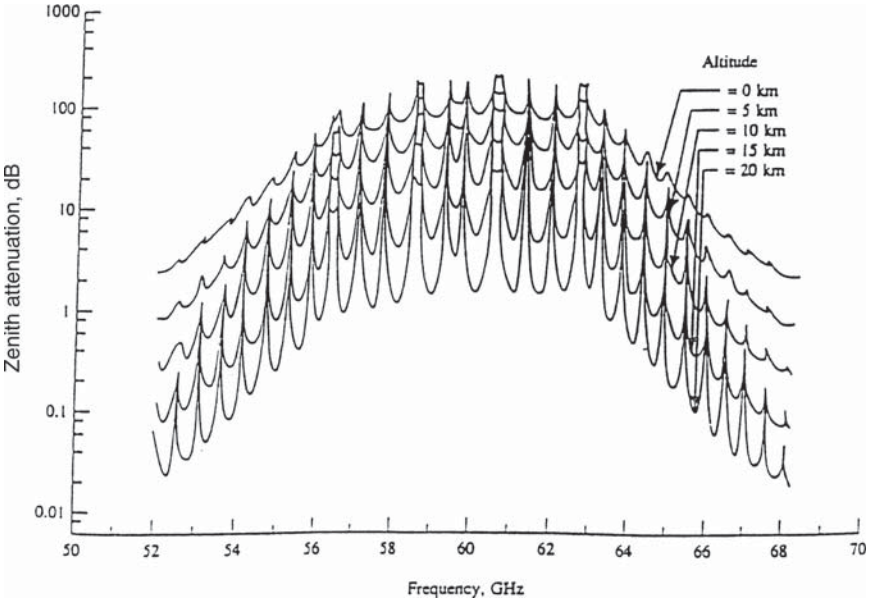


Figure 4-4 60 GHz zenith attenuation for various altitudes [9]

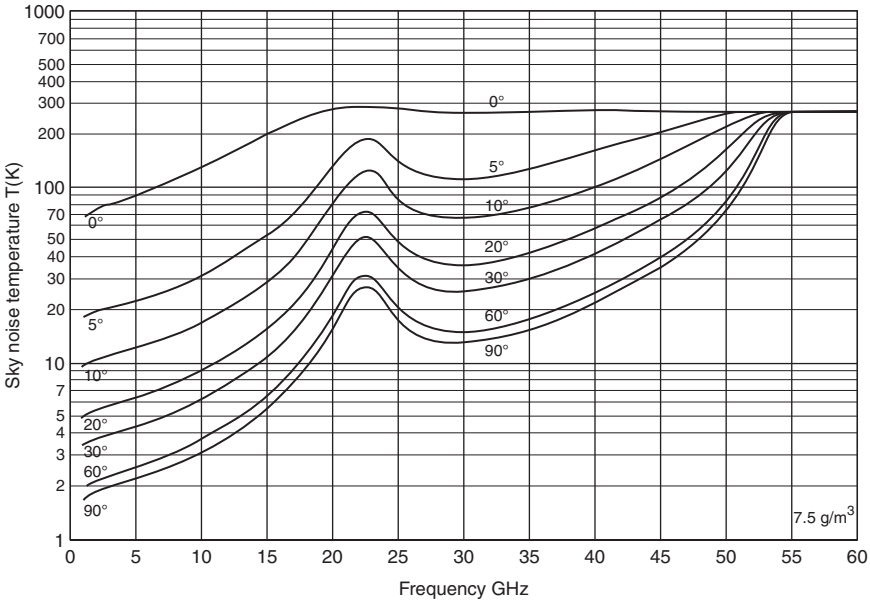


Figure 4-5 Sky temperature values [7]

Fig. 4-1) and exclusive of hydrometeor loss at the higher EHF frequencies. If molecular absorption and other losses were not present, the sky temperature would have the 3 K value for the cosmic background. As discussed in Chapter 1, user terminal performance is reduced by not only the propagation loss but also the increased system noise temperature. By contrast, uplink space segment antenna performance is reduced by only the propagation loss because the earth emission temperature is assumed to be 290 K and the additional attenuation from propagation loss does not change the antenna noise temperature from its 290 K value.

Generally, EHF link analyses address system performance under clear and inclement conditions separately. User segment antennas incur signal attenuation from the loss and noise generated by the loss (as described in Chapter 1). The antenna noise temperature of user antennas under clear conditions is computed taking into account the sky temperature values (indicated in Fig. 4-5) and other noise components (illustrated in Fig. 1-14). The increased antenna noise resulting from rain attenuation loss is addressed separately, as will be described later. The antenna noise temperature values are typically much less than 290 K. The additional hydrometeor loss both attenuates the signal and increases the antenna noise temperature. Space segment antennas, by contrast, unavoidably view the ambient earth background temperature, and generally, the antenna noise temperature of space segment antennas is assumed to be the ambient 290 K. Additional propagation loss does not increase the 290 K background temperature so that downlink paths are degraded only by the signal attenuation. In reality, the earth background temperature [10] differs from 290 K depending on the emission from the earth's surface as viewed by the antenna. Ocean areas are typically much less than 290 K, for example. The differences in the earth's emission characteristics are the basis of microwave radiometric sensors used in remote sensing applications to derive estimates of earth surface conditions. While assuming the antenna noise temperature of earth viewing antennas is not rigorously correct, the 290 K value is contractually convenient, avoids making a detailed assessment of the earth's emission for various orbital locations and antenna footprints, and provides some additional link margin.

The second EHF propagation factor is scattering from hydrometeors, upon which much has been written [11, 12]. Hydrometeors are water in nonvaporous forms having a characteristic particle size. Hydrometeors include rain, fog, clouds, hail, sleet, ice crystals, and snow. Liquid water has a higher index of refraction than frozen water, and consequently a greater impact on propagation. Similarly, the particle size compared to the system's wavelength has a significant impact. Consequently, rain has the most significant effect on propagation, and the small particle size in clouds and fog has a lesser effect. At microwave frequencies,

even raindrops are much smaller than the wavelength, so little effect is noted at those frequencies. However, the smaller EHF wavelengths are effected significantly by hydrometeors.

Early investigators during World War II [13] noted rain attenuation and observed measured data could be fit to a power law. Since that time, additional measured data, a rigorous derivation of scattering, and drop-size distributions have produced a detailed understanding of hydrometeor scattering [14]. The end result is that the specific attenuation (dB/km) for rain can be expressed as aR^b, where R is the rain rate in mm/hr. The reference gives coefficient values for a and b over a wide frequency range and provides a discussion of the scattering models and the raindrop size distributions commonly used. The specific attenuation needs to be integrated along the propagation path; consequently, the distribution of rain becomes important. Much work has been performed to model these effects.

The nature of rain and our inability to predict its distribution and occurrence present several difficulties. As may be appreciated from personal experience, much variability exists. The problem faced by communications system developers is the ability to determine the reliability of service that is generally expressed as availability. At lower microwave frequencies, straightforward answers can be provided since these frequencies are relatively insensitive to weather conditions. However, the variability of rain and the uncertainty of weather conditions introduce significant complications in projecting the reliability of EHF communication services. Because of the inherent variability, a statistical approach is required. However, such statistics are stable only in a multi-year sense (i.e., some years are wetter or drier than others). Similarly, if link availability on, say, next Thursday afternoon is to be determined, one is at the mercy of meteorologists. Confidence in weather predictions, particularly several days in advance, is limited.

These inherent limitations notwithstanding, the statistical approach is commonly used and has value in many applications. The objective of this statistical approach is to determine a weather margin for the system that can be used in link analyses. The long-term statistics are useful in applications such as television broadcasts, where outage over the satellite's lifetime is required. Rain attenuation values are also useful in system studies where comparative evaluations of different frequencies are performed.

With these reservations, the problem of assessing weather effects proceeds as follows. Data on rainfall rates at given sites are needed. Commonly, global climate models [15], such as that shown in Fig. 4-6, are used. These climate models are characterized by rain rate distributions in a cumulative statistical sense, as in Fig. 4-7. Specifying an availability of service (e.g., 99%) results in a corresponding rain rate for

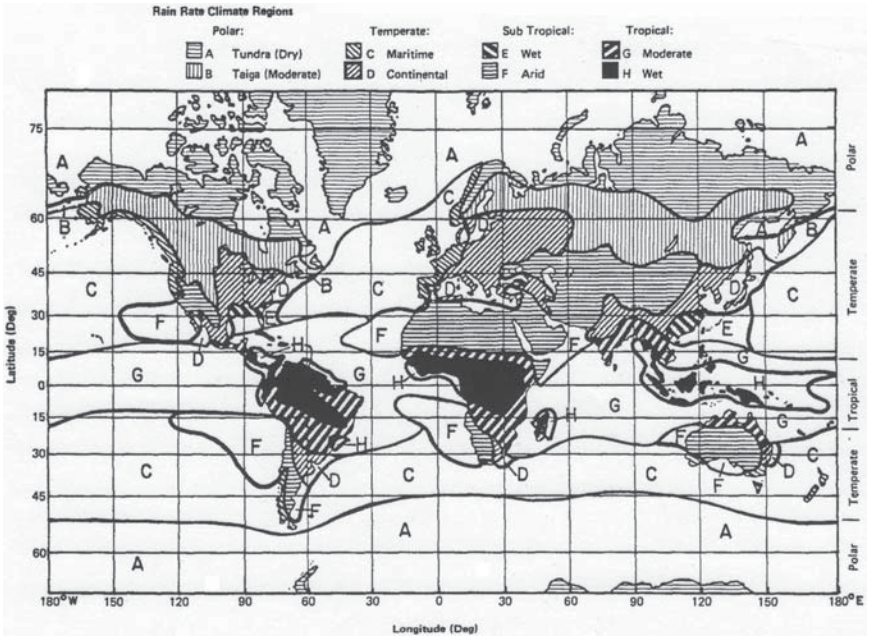


Figure 4-6 Global climate models [15] (© 1980 IEEE)

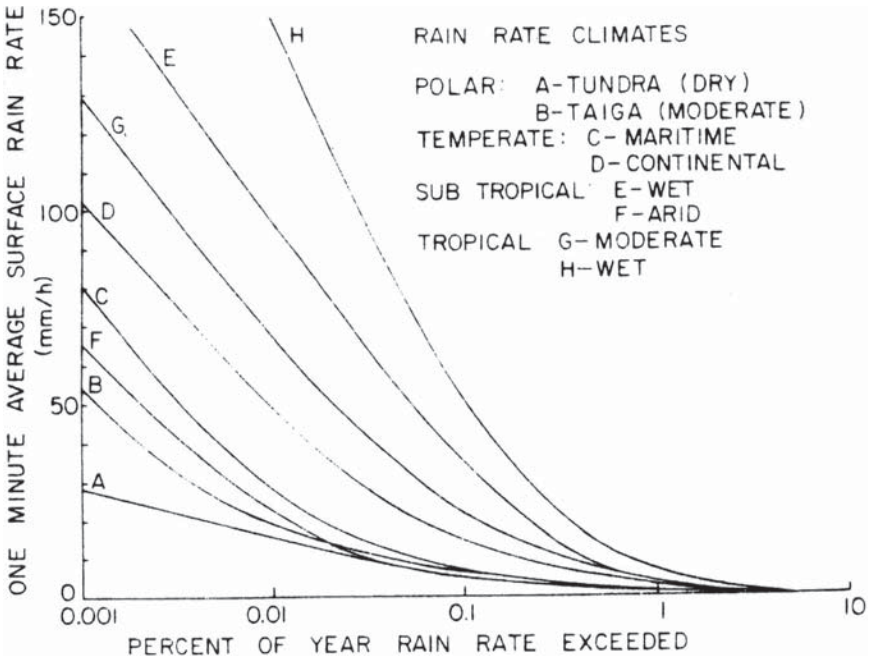


Figure 4-7 Rain rate distributions [15] (© 1980 IEEE)

a particular climate. This rain rate, the specific attenuation for the frequency, the elevation angle, and the isotherm height (shown in Fig. 4-8) provide the means to integrate along the path to obtain the attenuation for the path. The isotherm height is the point at which liquid water freezes. Above this height, relatively little attenuation occurs because of the low index of refraction of ice. This process of calculating the attenuation of a satellite link is conveniently performed with readily available computer programs. The attenuation for the rain rate is used in link analyses, and in user uplinks the increase in system temperature must also be addressed. Several refinements to this process have been made, particularly the statistical description of rain rates. Efforts to integrate the effects of weather-related link impairments have also been carried out [16].

The prediction capabilities for hydrometeor attenuation have been supported by a variety of experiments. A particularly comprehensive study [17] has been performed by NASA's ACTS (Advanced Communication Technology Satellite) program. This satellite-provided beacon downlinks at 20.2 and 27.5 GHz, and terminals at a large number of locations in five different rain regions were used to gather data. The results of these experiments provide an opportunity to examine the fidelity of the predictive techniques.

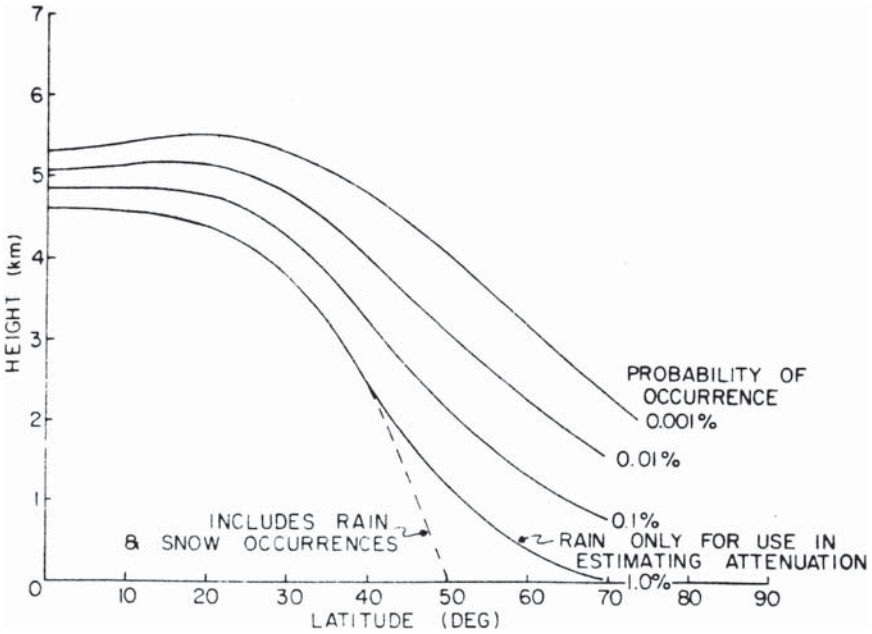


Figure 4-8 Isotherm height [15] (© 1980 IEEE)

The extensive ACTS data base provided a means to assess the fidelity of the predictive techniques. The comparison of predictive techniques [18] revealed that the Dissanayake, Allnutt, and Haidara model [16] appeared to provide "acceptable results." Predictive techniques also require rain rate distribution statistics and a comparison of the measurements with predictions [19] revealed "not one of the model combinations provided good predictions." Further refinements of rain rate prediction techniques [20] have been recently proposed. This work extends the two component modeling techniques that distinguished between the intense rain cells referred to as volume cells and the rain rates in regions beyond the volume cells referred to as debris.

Weather effects on communication links have been extensively measured. Two distinct types of measurement techniques are used. These techniques complement one another and both were used in the ACTS program. One technique uses a satellite beacon whose power level is carefully controlled since changes in the transmitted power can be misinterpreted as link attenuation. The satellite beacon has two strong advantages. A narrow bandwidth tone can be used and a narrow bandwidth ground receiver response produces high sensitivity, providing the ability to measure attenuation over a wide dynamic range. A second advantage is that two orthogonally polarized beacon signals can be transmitted so that the depolarization effects of weather events can be measured. The principal disadvantages of this technique are the cost of the satellite beacon and the inability to explore the effects of elevation angle dependence at a particular site since geosynchronous satellites with a fixed elevation angle are generally used to provide continuous monitoring at a given location.

The second technique uses a ground-based radiometer that avoids the cost of a satellite package and can point to any elevation angle or operating frequency desired. The radiometer detects the thermal emissions from the attenuating rain cell. A radiometer is a simple receiver design (described in Fig. 4-9) that generally is comprised of an LNA, downconverter, a square law detector, and an integrator. In practice, the radiometric output also varies with insertion gain variations of the electronics so that the radiometric output is calibrated by an established reference temperature, generally either a termination at a controlled ambient temperature or a noise source. An example radiometric receiver design [21] has been described.

Figure 4-9 A radiometric receiver

The output of the radiometer is proportional to the input noise level that depends, in part, on the antenna's noise temperature, as illustrated in Fig. 4-10. The antenna's noise temperature depends on the propagation loss, thereby providing a means of determining the attenuation. Recall that the rain attenuation generates noise like any other loss component and, as explained in Chapter 1, has an impact on the system noise temperature of the radiometer. The change in the emission temperature when rainfall is present, indicated by the radiometer, is related to the attenuation. The emission temperature increase, shown in Fig. 4-11, illustrates this effect. Three different empirical temperature values are indicated in this figure: 290 K is an ambient value commonly used for hardware loss; 270 K is commonly used because of the temperature decrease at high altitudes containing raindrops; and 250 K is an average value for high frequencies (e.g., 90 GHz), where at high rain rates, signal energy is scattered out of the beam. The figure indicates that the measurement's dynamic range is limited by the knowledge of the empirical temperature, thus a radiometer is most useful for measuring rain attenuation at lower rain rates.

A novel application of this technique [22] integrates a radiometer with the terminal receiver that provides a real-time measurement of the downlink attenuation. This design concept was developed for frequency hopped systems and can be implemented inexpensively. Other receiver modulations can offset the radiometric measurement channel from the operating bandwidth to achieve this attenuation measurement capability. The radiometer and the system receiver share the expensive system

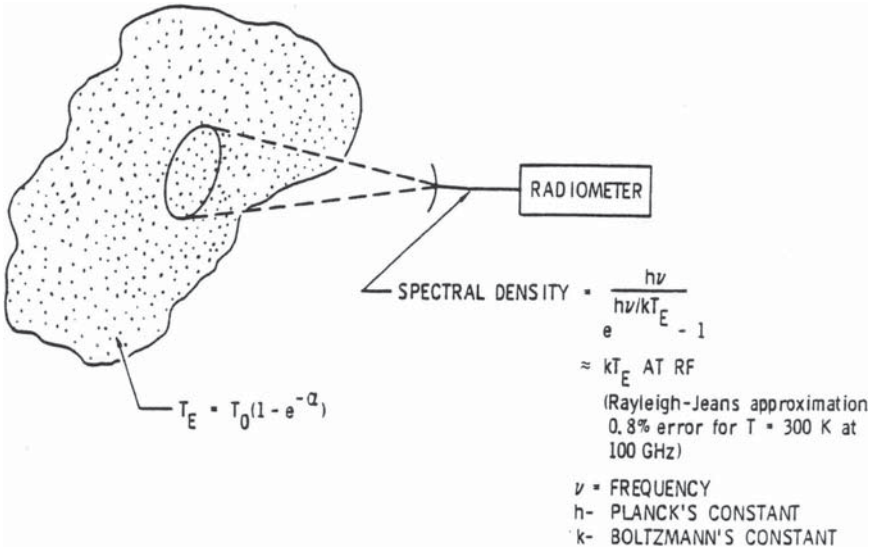


Figure 4-10 Radiometric emission [22] (© 1984 IEEE)

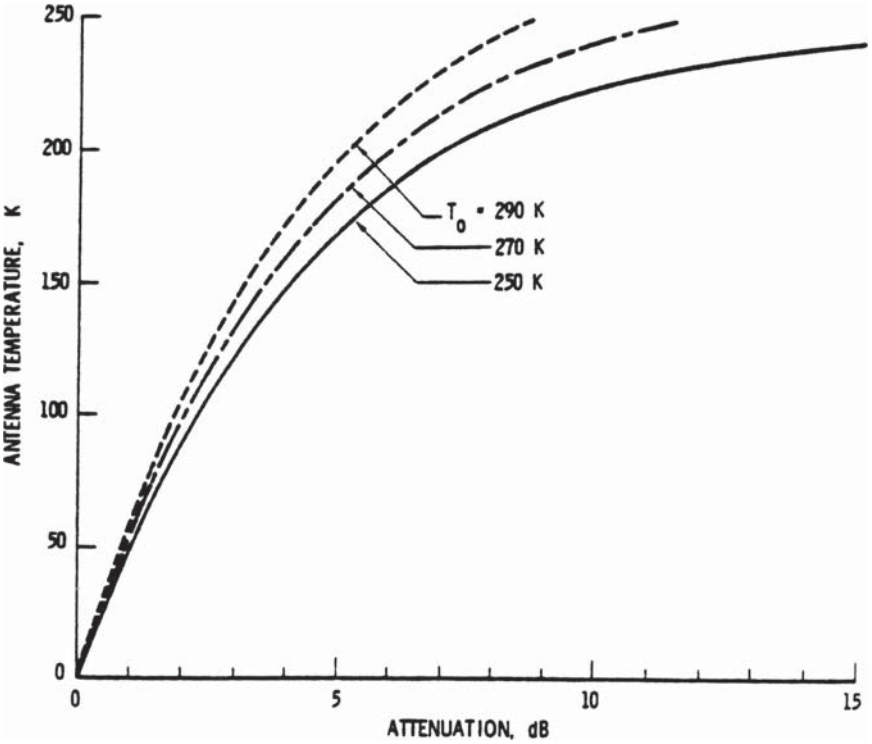


Figure 4-11 Emission temperature versus attenuation [22] (© 1984 IEEE)

components, the antenna and the receiver front-end. Such a capability provides a real-time measure of rain attenuation so that the potential for link closure is indicated, BITE diagnostics for the terminal's receiver front-end are also provided, and the presence of interference is indicated.

Beacon and radiometric techniques for rain attenuation are complementary capabilities. Beacon measurements can provide high dynamic range data capable of characterizing high attenuation values. Beacon measurements can also characterize depolarization resulting from rain. Radiometric measurements do not require a satellite signal and can be inexpensively constructed. Because a satellite signal is not used, radiometers have the flexibility to measure characteristics at selected elevation angles and frequencies. Radiometers do not have the capability to measure depolarization or accurately measure high attenuation values. This complementary capability was recognized in the ACTS program [23], and both types of techniques were used. Both the beacon and radiometric techniques require design attention to the wet antenna problem (as will be discussed) so that the loss in the path rather than the combined loss in the path and the antenna is measured.

Systems that simultaneously use orthogonal polarizations in polarization reuse designs to double the data rate to a given coverage region have further limitations. Such systems must maintain adequate isolation between the polarizations to avoid degradation from co-channel interference. If a system has high signal margins to maintain availability during rain events, the system may be limited by loss of isolation from depolarization rather than being limited by propagation loss. Three distinct mechanisms contribute to depolarization. The first is the polarization purity of the transmit and receive antennas, as has been discussed in Chapter 1. The second is the depolarization resulting from rain in the propagation path. The third is the depolarization resulting from the wet antenna. The development of dual polarized systems must address each of the three factors. Antenna developers must configure antennas with a high level of polarization purity, and the requirements for such purity in achieving isolation have been previously discussed in Chapter 1. The depolarization resulting from rain and wet antennas will be discussed.

Raindrops are not ideally spherical and consequently rain causes depolarization, degrading isolation between orthogonally polarized signal components. The shape of the small raindrops in light rain is close to spherical because of the surface tension. Larger raindrops in heavy rain are distorted by wind forces and the longer axis tends to be in a horizontal direction. Thus, depolarization increases with increasing rain rate. Systems configured with high rain margins must be concerned not only with signal loss but also increased depolarization that results in co-channel interference. Several references [24, 25, 26, 27, 28] provide data on depolarization. System designs where orthogonal polarizations are used, as well as high rain margins, can be limited by co-channel interference resulting from the depolarization before they are limited by rain attenuation. Adaptive cross-polarization cancellation systems that dynamically maintain isolation [29, 30] have also been described.

In addition to weather effects on the propagation path, the wet antenna problem further reduces link performance. Two cases need to be distinguished. The first is when the antenna is enclosed by a radome, and the second is when the antenna is not enclosed by a radome.

Antennas that use protective radomes incur loss when wet because the additional unintended dielectric resulting from moisture degrades radome transmission efficiency. The loss in transmission efficiency depends on the moisture distribution and the operating frequency. Much work has been done to develop radome coatings that cause moisture beading in order to avoid uniform water layers that degrade radome transmission. These coatings are described as hydrophobic and result in rain beading on the surface rather than sheeting. While such beading reduces the increased radome loss compared to a uniform

water layer, the beaded water surface results in scattering that increases the antenna's sidelobes and produces additional cross-polarization, a concern for polarization reuse systems. The hydrophobic coatings have a limited lifetime, much like a coat of wax on a car's finish. The coatings require reapplication to maintain their effectiveness.

Measured results have been reported on typical radome materials. Data were taken on new, used, cleaned, and recoated radomes [31] over a 13 to 21 GHz frequency range. A typical result in Fig. 4-12 applies to rain at a 30° incidence angle, which indicates a loss of about 2 dB at the higher rain rates. Data taken over a range of rain incidence angles show minimal effects for incidence angles greater than 30° and increasing attenuation for smaller incidence angles, probably due to puddling. In operation, only a limited amount of the antenna's aperture would intercept portions of the radome having small rain incidence angles. Other measurements [32] were made at 20 GHz with comparable results and include cross-polarized data. Two different hydrophobic coatings were tested. The more effective coating results in somewhat lower attenuation but higher cross-polarized levels during rainfall.

Measurements have also been performed on antennas unprotected by radomes. Two distinct issues impact performance. The first is the problem of water accumulation on the feed's radome, and the second is water accumulation on the reflector surface. The presence of water on the feed's radome has a significant effect, much like the radome enclosing the entire antenna. Protecting the radome from water accumulation remains a design issue—shrouding and providing a forced

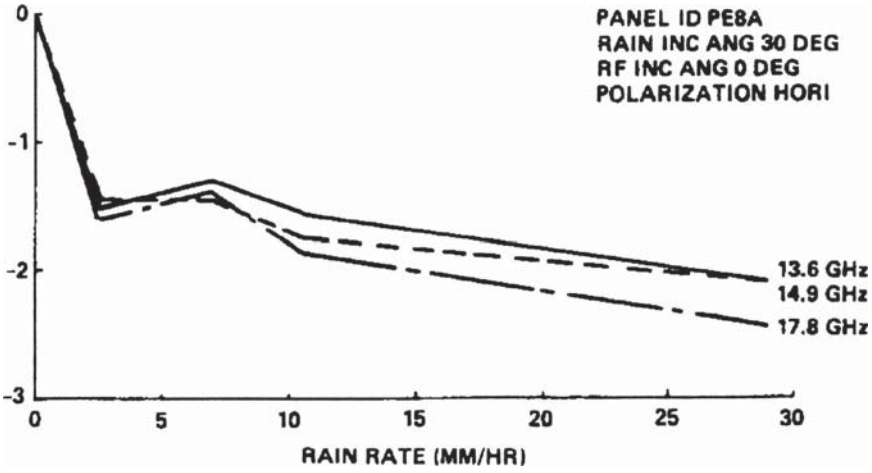


Figure 4-12 Wet radome loss [31] (© 1988 IEEE)

air flow over the feed radome are potential solutions. The antennas used in the ACTS experiment have been investigated [33] and are a limiting factor in obtaining correlation with the path attenuation projections predicted. In this case, the reflectors had a crinkled plastic surface that accumulated water more significantly than a smooth surface. These analyses were performed with the objective of providing a correction factor so the path attenuation could be separated from the total measured attenuation. Measurements using sprayers and analysis techniques were used in these studies. The attenuation for the wet feed, crinkled and smooth plastic surfaces, and a metal surface are indicated in Fig. 4-13 as a function of rain rate. It was also observed that wet snow accumulating on the crinkled reflector surface resulted in significant attenuation.

Another approach to minimizing the effects of wet antennas has been applied to radiometric measurements [35, 36]. The antenna is an offset reflector that is enclosed by a shroud to keep the antenna dry. A flat plate at a 45° angle images the aperture and rotation about an axis, while maintaining the 45° angle permits operation at different elevation angles. The flat reflector is the only component exposed to rain. Measurements and analyses of this flat plate were performed to determine the contributions to the antenna noise temperature.

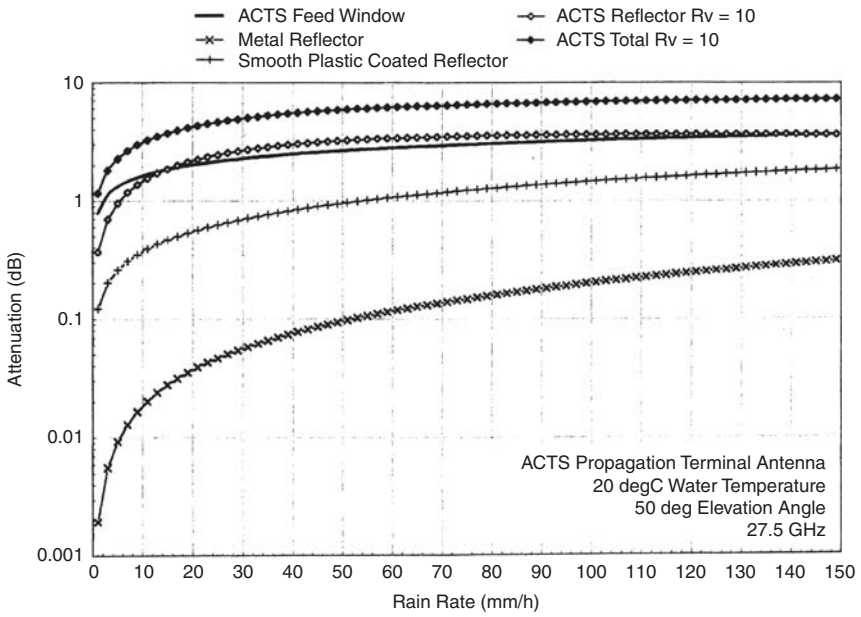


Figure 4-13 Wet antenna loss [33] (© 2002 IEEE)

The ability to configure practical EHF systems depends on coping with weather limitations. As the saying goes, "everyone talks about the weather, but nobody does anything about it." Clearly, coping with weather limitations depends on the specific system requirements for availability, the local climate, and the required elevation angle. System requirements for high availability impose demanding and generally costly impacts on terminal performance. Four distinct approaches exist for dealing with weather effects. The first is to provide adequate rain margin to achieve the required system availability. However, at higher EHF frequencies and for systems requiring high availability at low elevation angles, the required margin may preclude practical or affordable system designs. The second alternative is to select locations with little rainfall and relatively high elevation angles to minimize the rain margin requirements so that a practical design can be achieved. For example, locating terminals with high availability and low elevation angle requirements in the tropics is probably ill-advised. The third alternative is to reduce the data rate during inclement weather so that link closure can be maintained; many system designs, however, do not have the flexibility to vary data rates. The fourth alternative is to use a second terminal separated from the first. This technique, referred to as spatial diversity, uses adequate separation between the two terminals [37] so that the probability of simultaneous intense rain at both sites is low.

Studies and measurements have investigated the separation requirements for site diversity. Very intense rain having high attenuation is relatively localized; typical rain cells having very high rain rates are limited to 1 to 2 km in extent. As the distance from the intense rain cell increases, the rain rate also decreases. Site diversity separation requirements are thus based on the probability that one or more separated sites have a sufficient rain margin to meet the availability requirement. In practice, a system is configured for a particular rain margin, and in site diversity, two or more terminals are configured based on that rain margin. The required separation is based on the probability that link closure is possible with one of the terminals for a given level of availability. For example, suppose the overall availability is to be 99% and the individual terminals in the diversity plan have a margin that provides a 98% availability corresponding to a rain margin and corresponding rain rate. The minimum separation from an intense rain cell is thus based on a sufficient distance where the link will close, corresponding to a 98% availability. The minimum separation for a given availability clearly is reduced as the design rain margin increases.

Two parameters are defined in connection with site diversity. The first is diversity gain that is defined for a specified availability value as the

difference between the margin required for a single site and the margin required for multiple sites. This parameter is always bounded by the design rain margin. The second parameter is the diversity improvement that is defined as the difference in the availability of a single site and the availability of multiple sites for a given design margin value. An example of measurements made in the ACTS program [38] is given in Fig. 4-14. Three sites in Virginia and Maryland having separations of 30 to 40 km were addressed. As an example, if the design rain margin is 5 dB, the single site availabilities are on the order of 98%, whereas the three site availability is about 99.99%. These data are an example of the diversity improvement and illustrate the benefit of increasing the design rain margin value. Likewise, the diversity gain is indicated in Fig. 4-15 for the three sites. If the system is sized for a single site design rain margin of 5 dB, the diversity gain is about 3 dB. Thus, if the availability of the single site is to be maintained for three sites, the design rain margin required at each of the three sites is 2 dB rather than 5 dB. It should be noted that these site diversity experiments are limited to one geographic location and elevation angle and subject to rain conditions in that locale. Nevertheless, the data illustrate site diversity benefits and provides a basis to predict performance for other cases.

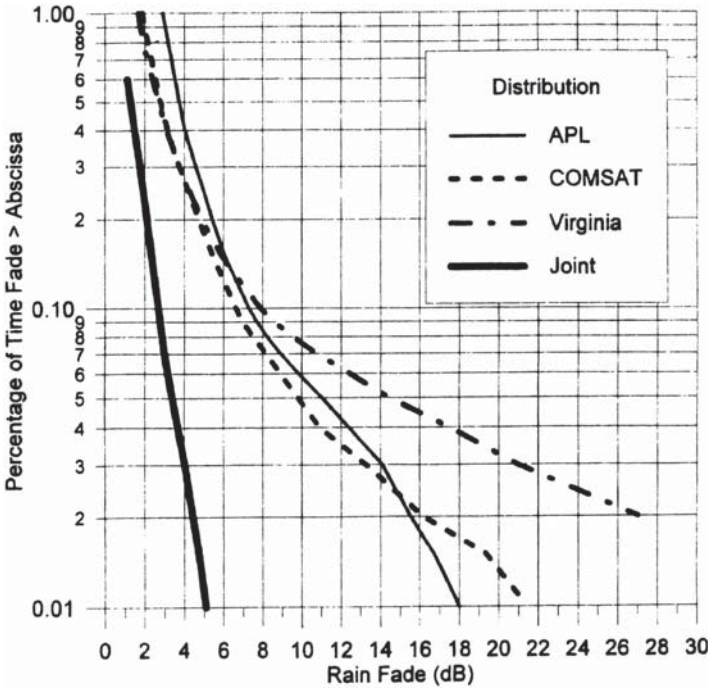


Figure 4-14 Single and multiple site availability [38] (© 1997 IEEE)

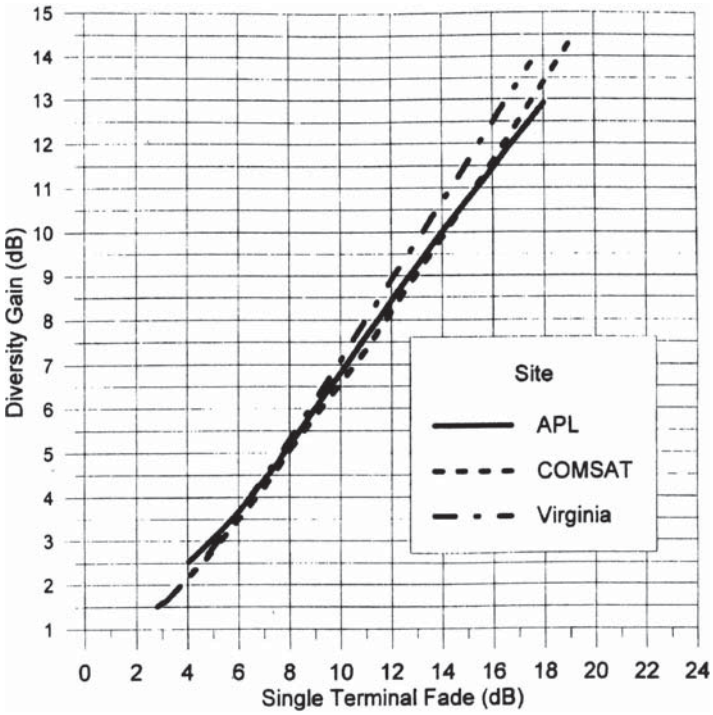


Figure 4-15 Diversity gain [38] (© 1997 IEEE)

4.3 Modulation and Multiple Access

Information is conveyed by modulating the RF signal, and a variety of modulation formats have been used. Digital modulation techniques have almost entirely supplanted analog modulation used by early satellites. Space segment transponders generally simultaneously service many individual users that must be isolated from one another while sharing the satellite's resources. Multiple access methods are used to isolate system users from one another to avoid mutual interference. These topics are discussed at a top level.

A variety of digital modulation techniques exist to satisfy differing program objectives. As in many situations, no one modulation format has universal appeal. Independent of the specific modulation format, system performance is measured by the received data fidelity. The measure of signal fidelity for digital modulation techniques is BER (bit error rate), which quantifies the number of errors statistically occurring in a specified number of digital bits. System performance is specified by requirements to achieve a particular BER performance (e.g., a BER of 10^{-6}), which means that statistically one wrong bit has been received out of a million bits.

Signal detection performance depends on the available signal energy in comparison to the receiver's noise level and for digital modulation techniques, the performance measure is E_b/N_o, the energy per bit divided by the noise spectral density. Signal fidelity i.e., BER depends on E_b/N_o, and reduced BER levels require increased E_b/N_o. The relation between these two parameters depends on the specifics of the modulation formats, and the notional values shown in Fig. 4-16 illustrate typical behavior that is sometimes referred to as "waterfall" curves. These values are derived from theoretical analyses and represent the performance that could be achieved by an ideal receiver. Practical receivers do not achieve ideal performance and the difference between the ideal performance and the performance of a practical receiver is the implementation loss that is a component of practical link analyses. Specifically, the implementation loss is the difference between the theoretical E_b/N_o value required to meet the specified BER performance and the required E_b/N_o to achieve the specified BER value. Receiver implementation loss performance is measured using specialized test set instrumentation or hardware configurations using operational system modulation modems that are developed to evaluate the performance of operational systems.

The required bandwidth and the frequency spectra of the signals also depend on the modulation formats as illustrated in Fig. 4-17. Like antenna systems that have a finite aperture, digital modulation has a finite bandwidth so that modulation sidelobes exist extending beyond the main response. Also like an antenna, phase errors increase the sidelobe levels and the phase distortion from nonlinear transmitter operation (described in Chapter 2) result in increased modulation sidelobe levels

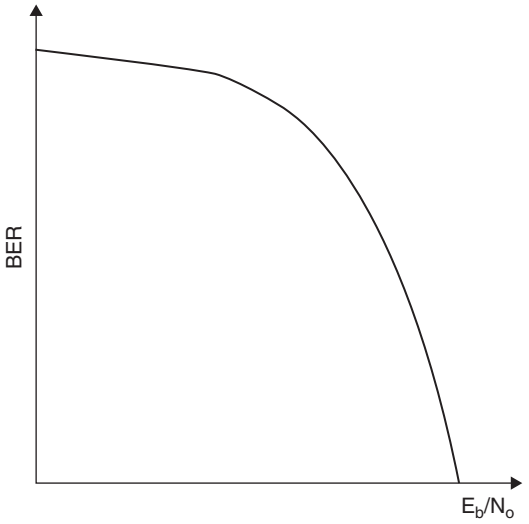


Figure 4-16 BER values

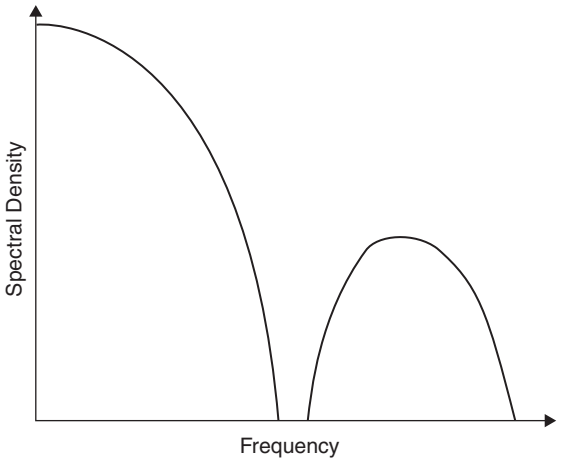


Figure 4-17 Modulation baseband spectral characteristics

or spectral regrowth. Filtering techniques are needed to control these modulation sidelobes to avoid interference to other system users and/or other systems. The ability to share the system with other users and the potential of interference to other systems will require additional design attention in future systems.

The choice of digital modulation formats depends on the system objectives. One objective is to communicate as much information as possible through available bandwidth. Modulation formats for this objective measure their effectiveness by the number of bits per hertz that can be communicated. For example, BPSK modulation provides 0.5 bits per hertz, QPSK modulation provides 1 bit per hertz, and 8PSK provides 2 bits per hertz. These higher-order modulation formats that increase the number of bits per hertz are referred to as "bandwidth efficient modulation." Detection performance depends on the spacing between the bits in the modulation constellation and as the number of bits per hertz increases, additional E_b/N_o levels are required to maintain BER performance.

Another system objective is to minimize the system resources by using error-correcting coding and interleaving techniques to reduce the required received signal level. Additional bits are added to the original information with the objective of enhancing detection performance at low signal levels. These error-correcting codes assume that incorrectly detected bits are randomly distributed. Interleaving techniques that redistribute the order of the original information bits are used to randomize the bit errors. The progress in error-correction encoding [39] has greatly reduced the required signal levels to achieve a specified BER value. While error correction coding reduces the required signal levels, the steepness of the "waterfall" increases and greater attention

must be given to assure user-received signal levels will exceed the required threshold value. The reduced signal level requirements result in reduced user resources, allowing the use of more compact user terminal designs. Consequently, error-correction coding techniques are commonly used.

A third system objective is to protect the signal from interference, and a variety of spread spectrum modulation techniques have been developed to protect system users from interference. Spread spectrum modulation uses a much greater bandwidth than that normally required to transfer the information. Two distinct techniques are used in spread spectrum modulation. The first is frequency hopping, in which the signal's carrier frequency is hopped in a pseudorandom pattern. The second technique is referred to as direct sequence and modulates the signal with a wide bandwidth pseudorandom code to spread the information over a much wider bandwidth. The frequency hopped pattern and the wide bandwidth pseudorandom code are known only to the users so that the desired signal can be separated from the received signal spectrum and adversaries cannot exploit the codes to degrade system performance. Spread spectrum modulation distributes the user's information over a wide bandwidth to dilute the effects of interference power. The spread spectrum effectiveness in reducing the impacts of interference is generally measured by its processing gain. To first order, the processing gain is the ratio of the spread bandwidth to the bandwidth normally required for signal communication.

The space segment provides communication services to multiple system users, and the individual users must be isolated to avoid mutual interference. Signal isolation techniques are limited to spatial location, polarization, frequency, time, and orthogonal coding. Antenna systems provide spatial and polarization isolation and separate coverage areas by frequency and polarization reuse plans to avoid mutual interference between coverage areas. Each coverage area is assigned a frequency subband. Individual users sharing the same frequency subband within a given coverage area use signal modulation methods to isolate individual system users through the use of multiple access methods.

Three multiple access methods are commonly used, and in some cases, hybrid combinations of multiple access methods are employed. The most widely used modulation method is FDMA (frequency division multiple access), which divides the satellite transponder subbands into frequency assignments that are allocated to individual users. Each user is required to maintain its modulation within frequency assignment, and since the total power within the bandwidth is the composite of the individual powers, power control among users in frequency translating transponder architectures is necessary so each user has an equable share of the satellite downlink, as discussed in Chapter 3.

The second modulation method is TDMA (time division multiple access), where each user can use the entire available spectrum for an allotted time period. A synchronized time code and notification of the time allocation are needed in this technique to prevent users from intruding on each other. The third technique is CDMA (code division multiple access), where each user is assigned one member of an orthogonal code set. The individual user processes his information by correlation with code assigned to the user.

Again, no one multiple access technique or hybrid of these techniques is universally advantageous. A variety of implementation alternatives exist, and the efficiency of these techniques depends on the number of users, their duty cycles, the traffic model, the channel bandwidth, the dynamic range of user ERP levels, and the tolerance to co-channel interference. Clearly, the choices of multiple access techniques are specific to the application and debate of the merits of alternative multiple access techniques continues. For example, the projected capacity of CDMA systems may be reduced [40] when time-delayed multipath components are present, reducing the orthogonality of the user's codes. Such issues are currently being explored by performing propagation measurements and simulations.

4.4 Link Analyses

Communication system performance is projected by using link analyses. The link equation uses the characteristics of the transmitting terminal, the receiving terminal, the system losses, the propagation conditions, and system noise levels to determine the received signal level and contrast that value with the level needed to produce the desired data fidelity. The process of link analyses is described in Fig. 4-18. This equation is derived by starting with the power density P_d incident on the receiving terminal that is obtained from

$$P_d = P_t G_t L_p / 4\pi R^2$$

$$= ERP\ L_p / 4\pi R^2$$

where P_t is the transmit power, G_t is the transmit antenna gain, L_p is the propagation loss, and R is the separation between the transmitting and receiving terminals (at orbital distances, spherical wave propagation exists).

The power density multiplied by the receiving antenna's effective aperture yields the power received, P_r, by the antenna as

$$P_r = ERP\ G_r L_p / (4\pi R/\lambda)^2$$

This is the familiar Friis transmission formula [41].

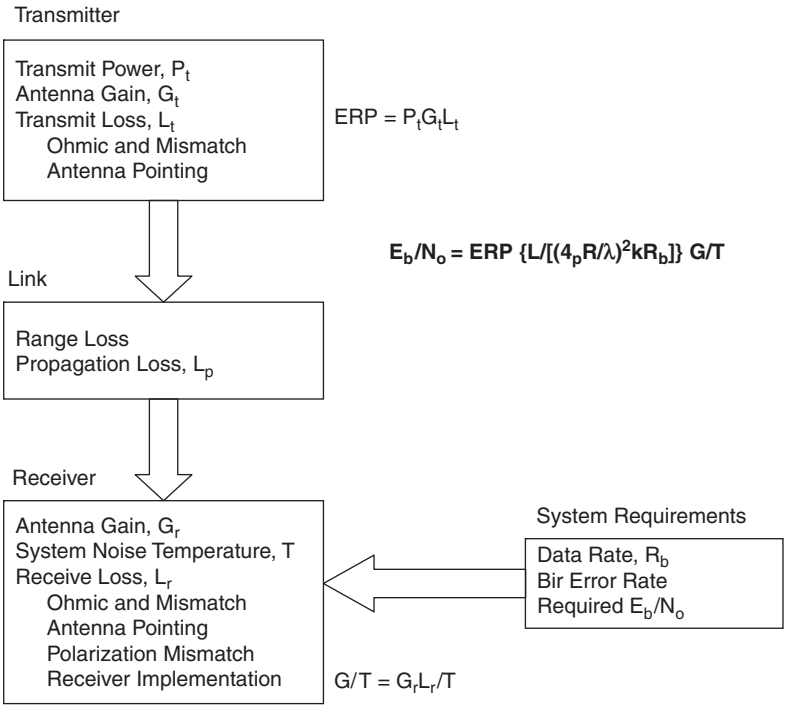


Figure 4-18 Link analyses process

For digital communication systems, the E_b/N_o, the energy per bit divided by the noise spectral density, is required. The energy per bit equals the power multiplied by the bit time, which is the inverse of the bit rate, R_b. The noise spectral density is Boltzmann's constant k multiplied by the system temperature T. Thus,

$$E_b/N_o = ERP\ G_r/T\{L/[(4\pi R/\lambda)^2 kR_b]\}$$

where L is the total system loss that incorporates the propagation loss L_p, the receiver's implementation loss, any pointing loss and polarization loss, and others (ohmic and mismatch losses are included within the antenna gain values). It is apparent in this equation why ERP and G/T are important values in characterizing communication system performance.

The frequency dependence of the link equation is also worth examining. Two distinct situations exist. In the first situation, space segment antennas are generally configured to provide service over a specified coverage area. As such, the gain of a space segment antenna over the coverage area can be written as $K_cEOC/(\theta_c)^2$, where K_c is a constant depending on antenna efficiency, EOC is the edge of coverage gain, and

θ_c is the angular extent of a coverage area that is assumed to be circular. The antenna gain to maintain a specified coverage area is independent of frequency to first order. Similarly, antennas for the space segment are generally constrained to a given physical size and (as discussed in Chapter 1) the gain for a circular aperture equals $\eta(\pi D/\lambda)^2$. The link equation then becomes

$$E_b/N_o = P_t K_c EOC/(\theta_c)^2 \eta(D)^2 \{L/[16R^2 kTR_b]\}$$

Thus, to first order, when a specified coverage area is to be served having users with a common antenna size, the link equation is independent of frequency. However, other factors such as allocated frequency and its bandwidth dictate the system design along with variations of the parameters assumed to be frequency independent. For example, the rain margin required to maintain EHF availability results in significant increases in system loss values and, on the downlink, increased system noise temperature values from the emission from propagation loss.

The second situation arises for a point-to-point service where it is assumed both the space segment and user segment antenna diameters are size constrained. Such point-to-point services arise in crosslink and gateway services. The antennas are again assumed to be circular, with respective antenna efficiencies of η_r and η_t. In this case, the link equation becomes

$$E_b/N_o = P_t \eta_t (D_t)^2 \eta_r (D_r)^2 [L\pi^2/(16R^2 kR_b]/\lambda^2$$

Again assuming the parameters are frequency independent to first order, operation of point-to-point services favor higher frequencies because of the $1/\lambda^2$ dependence of the link equation. This is one reason crosslink systems are configured for EHF operation.

A simple example was contrived to illustrate link analyses procedures. Service is to be provided to a number of small users located within a 2^o spot size at a 0.5 Mbps rate within a 40 MHz subband where the users are isolated by an FDMA multiple access method. The modulation is assumed to be QPSK having a one bit per Hertz rate. Error correction coding will be used at a rate ½ that doubles the required bandwidth since each information bit has a corresponding coded bit. The bandwidth required for a 0.5 Mbps service is therefore 1 MHz. The required E_b/N_o is assumed to be 5 dB, including implementation loss. An uplink frequency of 10 GHz and a downlink frequency of 9 GHz are assumed; these frequencies are used for illustration and are not authorized for communication satellite operation. The link budget details are presented in Table 4-1.

The details of the table entries and the assumptions for their values are described. The space segment antenna parameters assume a 55% efficiency and a beamwidth factor of 70, as discussed in Chapter 1.

TABLE 4-1 Example Link Analyses

	Uplink 2-ft Antenna	Uplink 1-ft Antenna	Downlink 2-ft Antenna	Downlink 1-ft Antenna
ERP, dBm				
Antenna Gain, dBi	33.5	27.5	34.3	34.3
Transmit Power, dBm	30	36	43	43
Insertion Loss, dB	0.5	0.5	0.5	0.5
Propagation Loss				
Space Loss, dB	203.8	203.8		
Atmospheric Loss, dB	0.3	0.3	0.3	0.3
Receive Antenna Gain, dBi	35.2	35.2	32.6	26.6
Insertion Loss, dB	0.5	0.5	0.5	0.5
Design Margin, dB	3	3	3	3
Received Signal Power, dBm	–109.4	–109.4	–97.3	–103.3
Noise Spectral Density, dBm/Hz	–173.0	–173.0	–176.8	–176.8
Eb/No, dB	6.6	6.6	22.5	16.5
Required Eb/No, dB	5	5	5	5
Margin (Single User)	1.5	1.5	17.5	11.5

The 2^{o} spot size with these assumptions requires an antenna diameter of about 41" and results in an uplink peak gain level of 38.2 dBi and a downlink gain level of 37.3 dBi, where its beamwidth equals 2.2^{o}. The edge of coverage gain for the space segment antenna is 3 dB lower than the peak gain levels. Two values of user antenna diameters, 1 ft and 2 ft, are used in this example, and the antenna efficiency assumptions for the space segment antenna will also be used for the user segment. The 2 ft antenna has an uplink antenna gain of 33.5 dBi and a downlink antenna gain of 32.6; the corresponding values for the 1-ft antenna are 6 dB lower since antenna gain is proportional to the square of the antenna diameter. The user antenna noise temperature is assumed to be 50 K, and with a 0.5 dB insertion loss between the antenna and LNA, the antenna noise temperature referenced to the LNA input equals 76.1 K (as discussed in Chapter 1). A receiver noise figure of 1 dB corresponding to a 75 K noise temperature is assumed for both the space and user segment. The system noise temperature for the user segment is 151.1 K so that the noise spectral density, kT_s (k, Boltzmann's constant, equals –198.6 dBm/Hz/K), equals –176.8 dBm/Hz. The space segment noise temperature assumes a 290 K earth background resulting in a 365 K system noise temperature and a –173 dBm/Hz noise spectral density. The E_b/N_o values are obtained by dividing the received power levels by the product of the noise spectral density and the bit rate (57 dB).

The link budget assumes the uplink performance with the smaller antenna is maintained by increasing the 1 W transmit power level

assumed for the 2-ft antenna to 4 W for the 1-ft antenna to offset the smaller antenna's 6 dB lower antenna gain level. The downlink signal margin allows 40 users to use the system while achieving a 1.5 dB link margin when a 2-ft terminal design is used, while the smaller user terminal could support ten users with a 1.5 dB link margin. However, a problem exists with the system design that uses the 2-ft antenna, since 40 users would occupy 40 MHz of bandwidth without any guard band between user frequency assignments, and ten users would sparsely occupy the 40 MHz bandwidth. Satellite systems are generally limited by the satellite's downlink resources and can be bandwidth limited, as is the case for the 2-ft user terminal, or power limited, as is the case for the 1-ft user terminal. Typically, satellite systems are downlink power limited. If the number of system users is to be increased for the 1-ft terminal design, the 20 W downlink transmitter could be replaced by a 40 W design, and the additional 3 dB downlink performance could support 20 users in place of ten. Clearly, a variety of tradeoffs exist in the system design (as was discussed in the Introduction) and trade-off exist between the space and user segments in evolving candidate system designs.

References

1. L. J. Ippolito, "Propagation Effects Handbook for Satellite Systems Design," Fifth Edition, ITT Industries Doc AF01-0006, September 2000.
2. R. B. Dybdal, "Antenna Tradeoffs for UHF SATCOMS," *2000 IEEE MILCOM Symposium Digest* (October 2000).
3. R. B. Dybdal, "UHF SATCOM Antenna Architectures," *2000 IEEE MILCOM Symposium Digest* (October 2000).
4. R. B. Dybdal and M. A. Rolenz, "Satellite Beacon for UHF System Users," *2003 IEEE MILCOM Symposium Digest* (October 2003); see also R. B. Dybdal and M. A. Rolenz, "Method of Determining Communication Link Quality Employing Beacon Signals" (May 13, 2008): U.S. Patent 7,373,105.
5. R. K. Crane, "Ionospheric Scintillations," *Proc IEEE*, vol. 65 (February 1977): 180–199.
6. J. Aarons, H. E, Whitney, and R. S. Allen, "Global Morphology of Ionospheric Scintillations," *Proc IEEE*, vol. 59 (February 1971): 159–172.
7. E. K. Smith, "Centimeter and Millimeter Wave Attenuation and Brightness Temperature Due to Atmospheric Oxygen and Water Vapor," *Radio Science*, vol. 17, (November–December 1982): 1455–1464.
8. H. J. Liebe, "An Updated Model for Millimeter Wave Propagation in Moist Air," *Radio Science*, vol. 20 (September–October 1985).
9. R. B. Dybdal and F. I. Shimabukuro, "Electronic Vulnerability of 60 GHz Crosslinks," *1984 IEEE MILCOM Symposium Digest* (October 1984).
10. E. G. Njoku and E. K. Smith, "Microwave Antenna Temperature of the Earth from Geostationary Orbit," *Radio Science*, vol. 20 (May–June 1985): 591–599.
11. D. C. Hogg and T. S. Chu, "The Role of Rain in Satellite Communications," *Proc IEEE*, vol. 63 (September 1975): 1308–1331.
12. R. K. Crane, *Electromagnetic Wave Propagation Through Rain* (New York: Wiley, 1996).
13. D. E. Kerr, *Propagation of Short Radio Waves* (New York: McGraw-Hill, 1951).
14. R. L. Olsen, D. V. Rogers, and D. B. Hodge, "The aR^b Relation in the Calculation of Rain Attenuation," *IEEE Trans Antennas and Propagation*, vol. AP-26 (March 1978): 318–329.

15. R. K. Crane, Prediction of Attenuation by Rain," *IEEE Trans on Communications*, vol. COMM-28 (September 1980): 1717–1733.
16. A. Dissanayake, J. Allnutt, and F. Haidara, "A Predictive Model that Combines Rain Attenuation and Other Propagation Impairments Along Earth-Satellite Paths," *IEEE Trans Antennas and Propagation*, vol. AP-45 (October 1997): 1546–1557.
17. ———, Special Issue on Ka-band Propagation Effects on Earth-Satellite Links, *Proc IEEE*, vol. 85 (June 1997).
18. R. K. Crane and A. W. Dissanayake, "ACTS Propagation Experiment: Attenuation Distribution Observations and Prediction Model Comparisons," *Proc IEEE*, vol. 85 (June 1997): 879–891.
19. R. K. Crane and P. C. Robinson, "ACTS Propagation Experiment: Rain Rate Distribution Observations and Prediction Model Comparisons," *Proc IEEE*, vol. 85 (June 1997): 946–958.
20. R. K. Crane, "A Local Model for the Prediction for Rain Rate Statistics for Rain Attenuation Models," *IEEE Trans Antennas and Propagation*, vol. 51 (September 2003): 2260–2272.
21. F. O. Guiraud, J. Howard, and D. C. Hogg, "A Dual Channel Microwave Radiometer for Measurements of Precipitable Water Vapor and Liquid," *IEEE Trans Geoscience Electronics*, vol. GE17 (October 1979): 129–136.
22. R. B. Dybdal, "Radiometer Integrated with Communication Terminal," *1984 IEEE AP-S Symposium Digest* (June 1984).
23. R. K. Crane, X. Wang, D. B. Westenhaver, and W. J. Vogel, "ACTS Propagation Experiment: Experiment Design, Calibration, and Data Preparation and Archival," *Proc IEEE*, vol. 85 (June 1997): 863–877.
24. D. C. Cox, "Depolarization of Radio Waves by Atmospheric Hydrometeors in Earth-Space Paths, A Review," *Radio Science*, vol. 16 (September–October 1981): 1566–1570.
25. T. S. Chu, "A Semi-Empirical Formula for Microwave Depolarization Versus Rain Attenuation on Earth-Space Paths," *IEEE Trans on Communications*, vol. COM-30 (December 1982): 2550–2554.
26. C. E. Hendrix, G. Kulon, C. S. Anderson, and M. A. Heinze, "Multigigabit Transmission Through Rain in a Dual Polarization Frequency Reuse System: An Experimental Study," *IEEE Trans on Communications*, vol. COM-41 (December 1993): 1830–1937.
27. W. L. Stutzman, C. W. Bostian, A. Tsolakis, and T. Pratt, "The Impact of Ice Along Satellite-to-Earth Paths on 11 GHz Depolarization Statistics," *Radio Science*, vol. 18 (May–June 1983): 946–954.
28. W. L. Stutzman and D. L. Runyon, "The Relationship of Rain-Induced Cross Polarization to Attenuation for 10 to 30 GHz Earth-Space Radio Links," *IEEE Trans Antennas and Propagation*, vol. AP-32 (July 1984): 705–711.
29. T. S. Chu, "Restoring the Orthogonality of Two Polarizations in Radio Communication Systems I," *BSTJ*, vol. 50 (November 1971): 3063–3069.
30. T. S. Chu, "Restoring the Orthogonality of Two Polarizations in Radio Communication Systems II," *BSTJ*, vol. 52 (March 1973): 319–327.
31. F. J. Dietrich and D. B. West, "An Experimental Radome Panel Evaluation," *IEEE Trans Antennas and Propagation*, vol. AP-36 (November 1988): 1566–1570.
32. C. E. Hendrix, J. E. McNalley, and R. A. Monzingo, "Depolarization and Attenuation Effects of Radomes at 20 GHz," *IEEE Trans Antennas and Propagation*, vol. AP-37 (March 1988): 320–328.
33. R. K. Crane, "Analysis of the Effects of Water on the ACTS Propagation Terminal Antenna," *IEEE Trans Antennas and Propagation*, vol. AP-50 (July 2002): 954–965.
34. J. Y. C. Cheah, "Antenna Effect on VSAT Rain Margin," *IEEE Trans on Communications*, vol. 41 (August 1993): 1238–1244.
35. M. D. Jacobson, D. C. Hogg, and J. B. Snider, "Reflectors in Millimeter Wave Radiometry-Experiment and Theory," *IEEE Trans Geoscience and Remote Sensing*, vol. GE-24 (September 1986): 784–790.
36. D. C. Hogg, F. O. Guiraud, J. Howard, A. C. Newell, and A. G. Repjar, "An Antenna for Dual Wavelength Radiometery at 21 and 32 GHz," *IEEE Trans Antennas and Propagation*, vol. AP-27 (November 1979): 764–770.

37. D. B. Hodge, "An Empirical Relationship for Path Diversity Gain," *IEEE Trans Antennas and Propagation*, vol. AP-24 (March 1976): 250–251.
38. J. Goldhirsh, B. H. Musiani, A. W. Dissanayake, and K. T. Lin, "Three Site Space Diversity Experiment at 20 GHz Using ACTS in the Eastern United States," *Proc IEEE*, vol. 85 (June 1997): 970–979.
39. ———, Special Issue on: Turbo-Information Processing: Algorithms, Implementations, and Applications, *Proc IEEE*, vol. 95 (June 2007).
40. P. Monsen, "Multiple Access Capability in Mobile User Satellite Systems," *IEEE Jour Selected Areas on Communications*, vol. SAC-13 (February 1995): 222–231.
41. H. T. Friis, "A Note on a Simple Transmission Formula," *Proc IRE*, vol. 34 (May 1946): 254–256.

Chapter

5

Interference Susceptibility and Mitigation

5.1 Overview

The popularity of microwave systems and their extensive use in both satellite and terrestrial services for many applications continues to rapidly increase. This trend results in the limited available spectrum becoming still more crowded, particularly at the lower microwave frequencies, with signals for communications, navigation, radar, and remote sensing services. While each microwave service has its own frequency allocation, this increased microwave activity raises the potential for interference between systems. Thus, enhanced system immunity to interference has increased importance in the development of future systems.

Interference concerns differ between space and user segments. The space segment is separated from terrestrial interference sources by the orbital distances but is visible from significant portions of the earth's surface. The principal space segment concerns arise from interference between the payload's subsystems, which is addressed in EMI/EMC (electromagnetic interference, electromagnetic compatibility) assessments, unintentional interference to and from system users resulting from inadequate isolation between users, and possible terrestrial interference because the satellite is accessible from large portions of the earth's surface visible to the satellite. The requirements for increased system capacity have led to satellites that reuse frequency and polarization in multiple coverage areas to increase system capacity. Adequate isolation between system users must be assured to avoid co-channel interference. Isolation between users in separated coverage areas is achieved through antenna polarization purity and sidelobe control. Intentional interference or jamming has long been a concern to military system developers.

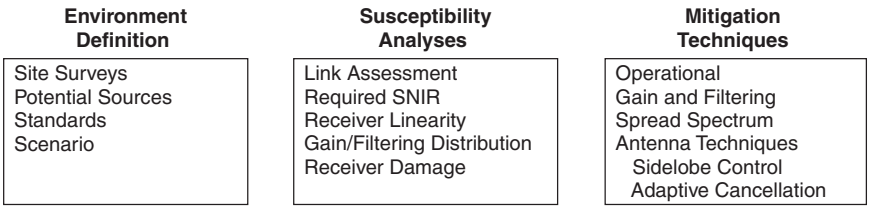


Figure 5-1 Interference assessment overview

This concern has resulted in a well-developed technology base to negate interference effects. However, both military and commercial users face the prospects of increased unintentional interference resulting from future microwave activity.

The user segment, by contrast, is much closer to other terrestrial systems that can potentially interfere with the user's receiver but is shielded from more distant systems by terrain and manmade features. Interference from terrestrial services close to communication satellite users, albeit outside of the user's bandwidth, is troublesome because the power received from the terrestrial services is often much higher than the power received from the user's space segment. For example, nearby high-power pulsed radars can result in receiver nonlinearity at out-of-band frequencies because of high received power levels and user segment designs having insufficient filtering. Unlike the space segment that must accommodate multiple simultaneous users, the user segment is impacted only by assuring its spectral characteristics conform to the space segment's user isolation requirements. Opportunities to augment and update interference protection during the satellite program's lifetime are more available to the user segment than the space segment and can be enhanced by design upgrades.

The process of addressing a system's susceptibility interference (described in Fig. 5-1) is comprised of (1) defining the interference environment, (2) addressing the receiver's interference susceptibility, and (3) determining a cost-effective combination of techniques that provide reliable interference protection as appropriate. This process will be described in some detail, and antenna interference mitigation techniques, passive sidelobe control, and adaptive interference cancellation will be discussed, with specific applications described in further detail in Chapters 6 and 7 for the space and user segment, respectively.

5.2 Interference Environment Definition

Defining the interference environment is a non-trivial task. Space segments must address potential co-channel interference between users sharing the same frequency subbands, potential EMI/EMC shortfalls,

and susceptibility to terrestrial sources. Intentional interference defined by scenarios is mitigated through signal waveform design, antenna coverage selection, and antenna design techniques such as sidelobe control and adaptive interference cancellation. The effectiveness of such techniques and a determination of an appropriate combination of them must be evaluated in the system design process.

User segment systems that have fixed locations generally perform site surveys to determine interference levels present at the site prior to the user equipment's installation. User equipment design must assess other services that use nearby frequency allocations to define out-of-band filtering requirements. Standards are another source that defines both the allowable emissions from the user's equipment as well as protection requirements to reduce susceptibility from other services. The limits for radiation and susceptibility levels are provided in EMI/EMC standards for a wide variety of applications and specify interference protection requirements that need to be incorporated into system designs. Intentional interference is generally addressed by defining scenarios that specify the number of interference sources, their location(s), power and spectral characteristics, and strategies for their deployment.

5.2.1 Site Surveys

Site surveys for the user segment are commonly performed to measure interference levels prior to the installation of fixed sites or in typical operating environments for mobile users. These site surveys must address two questions:

1. What high-level signals either in-band or out-of-band can damage or saturate the receiver, cause loss of acquisition, or degrade antenna tracking?
2. What in-band interference and its spectrum characteristics are present that can degrade signal reception, and what is its frequency of occurrence?

These surveys are conducted to determine filtering and linearity requirements for the receiver design and to identify additional requirements for interference protection needed at specific locations.

Such surveys are also used to diagnose existing systems experiencing interference. Surveys conducted prior to the installation use a separate survey antenna. Surveys to diagnose existing system shortfalls can use the actual operational antenna to measure interference power levels, and in such cases, the antenna's spatial position needs to be varied over the range of angles used operationally to determine the received interference power's dynamic range. The operational receiver's bandpass filtering and LNA also serve to determine out-of-band interference levels and spectral

characteristics that the receiver will experience. The interference levels can also be observed at different receiver interfaces to demonstrate filtering effectiveness. Many receiver system designs permit observing the effectiveness of the filtering at the output of the LNA, the IF levels, and other portions of the receiver design.

Terrestrial interference generally arrives through the user antenna's sidelobes because the antenna's main beam is pointing at the satellite. Consequently, an antenna gain reference level for the sidelobes is required to assess received power levels. The sidelobes away from the antenna's main beam region are no longer dominated by the antenna's aperture distribution and result from second-order radiation mechanisms. Additionally, the polarization differences between the interference and the antenna's sidelobe response are typically uncertain. When measurements are performed using a separate survey antenna such as a log periodic design, interference levels should be measured for orthogonal polarization orientations to quantify the total incident interference power. The broad beamwidth of log period antennas typically requires sampling four azimuth quadrants. The user's antenna response depends on the antenna technology, the direction of the interference, and the angular range over which the antenna is operationally used.

In many cases, interference assessments assume the user antenna's sidelobes are bounded by an isotropic gain level; thus measurements using a separate survey antenna are referenced to that gain value. The isotropic sidelobe level typically represents the worst-case level in regions where the sidelobe response, for example, results from spillover in reflector antenna designs. However, the actual pattern characteristics of the antenna and the range of elevation angles used operationally should be assessed to establish a reference level for specific applications. The received interference power can be adjusted by the sidelobe antenna gain level relative to the power received by an isotropic antenna.

Site surveys using existing site antennas observe the actual interference power levels received by the site's antenna. The sidelobe levels of the site antenna can be determined, if desired, by comparing the site antenna response with the response of a separate survey antenna. Repositioning the antenna over the range of operational pointing directions indicates the variation of received interference levels. Using a site antenna to measure the interference power is advantageous in removing uncertainty in the antenna's response and the operational input RF filter so that the input power to the receiver's LNA is directly measured along with the variation of the received interference levels over the anticipated operational pointing directions. These data, in some cases, can be used as a basis of arguments for "sector blanking" limitations.

Sector blanking restricts the operation of offending systems in the direction of the terminal. For example, high-power densities produced by radar systems can be controlled by limiting the radar's azimuth and elevation coverage when radar information in those directions may not be of particular interest. The sector blanking constraints limit radar illumination over portions of the radar's scan to avoid interference with other systems, an example of an operational solution to interference problems.

Site surveys use spectrum analyzer instrumentation to measure interference power. The system illustrated in Fig. 5-2 [1] uses a log periodic antenna to sample the incident interference, performs the measurements with a portable spectrum analyzer, covers a 500 MHz to 18 GHz frequency range, and has preamplifiers with protective filters to offset the inherently high noise figure of the spectrum analyzer. The photograph illustrates the log periodic antenna, the preamplifiers on the shelf of the tripod, and the spectrum analyzer located on top of the system's shipping cases. This particular system provides high measurement sensitivity over a very broad bandwidth to accommodate a variety of survey tasks. Generally, site survey equipment requires high

Figure 5-2 RF site survey equipment [1]

sensitivity over the bandwidth used by the system, and a filter whose response emulates the actual system's input bandpass filter is used to determine interference levels at the operational system's LNA input terminals. Outside of the operating bandwidth of the system, the user's system filtering is effective in reducing interference levels. Thus, the required sensitivity of the survey equipment is reduced compared to its in-band requirements, and the preamplification is not required.

Several issues must be addressed in calibrating the survey equipment. The antenna gain of the survey antenna must be determined so that the received signals can be related to an isotropic gain level for reference purposes. For example, the log periodic antenna has a boresight gain value of 7 dBi for the system shown in Fig. 5-2. When preamplifiers are used, their gain values are also required to relate the spectrum analyzer's indicated level to the power at the terminals of the survey antenna. This system uses two preamplifiers at each of two lower frequency bands that cover the 500 MHz to 2 GHz and 2 to 6 GHz bandwidths, respectively, and a single preamplifier that covers the 6 to 18 GHz bandwidth. The system's sensitivity expressed in a *G/T* (gain over temperature) value (shown in Fig. 5-3) illustrates the sensitivity benefits of the preamplifiers and the boresight antenna gain of the log periodic antenna. For reference purposes, the *G/T* of a system having a 200 K system noise temperature and an isotropic antenna gain in the sidelobe region is also indicated.

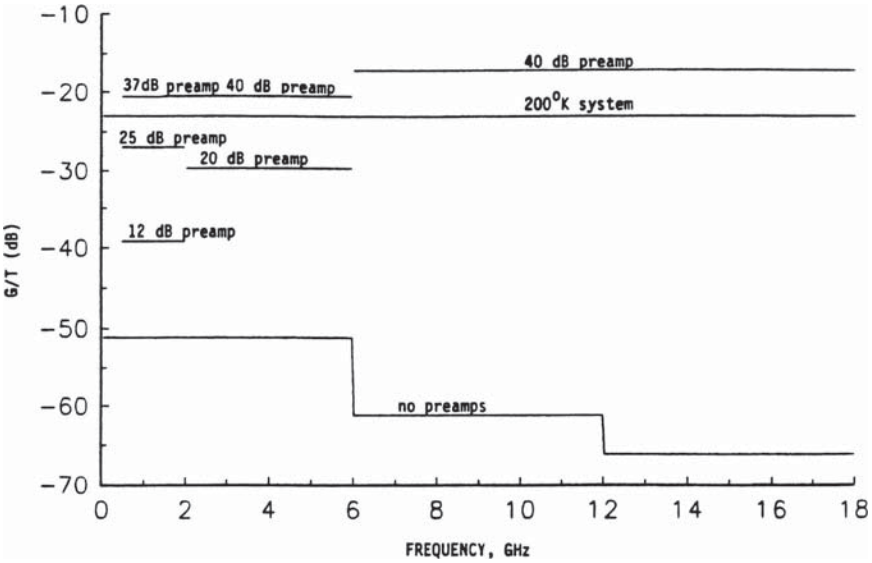


Figure 5-3 *G/T* sensitivity of survey equipment [1]

Like the operational system, the survey equipment requires linear operation to obtain valid results. For this system, the 1 dB compression points of the preamplifiers and spectrum analyzer were used to indicate the linearity limitations, and the power levels indicated at the spectrum analyzer referenced to the 1 dB compression point are indicated in Fig. 5-4. The preamplifiers in this case have 1 dB compression levels of about 10 dBm at their output, while the spectrum analyzer in this case has a 1 dB compression level of –5 dBm at its input terminal without attenuation. The filter selectivity of the spectrum analyzer and the use of the spectrum analyzer's attenuators provide linear operation when the spectrum analyzer measures spectral characteristics well removed from high-level sources. A combination of fixed filters and variable YIG band reject filters were also used to avoid spectrum analyzer and LNA (low-noise amplifier) compression to maintain system linearity in measurement situations where high-level interference was present. In addition, the preamplifiers could also be removed to increase the linear dynamic range for high-level signals at the expense of sensitivity for low-level signals, as indicated in Fig. 5-3.

Site surveys should allow sufficient measurement time to observe all potentially interfering systems because some systems are not operated continuously. The impacts of future nearby systems must be addressed

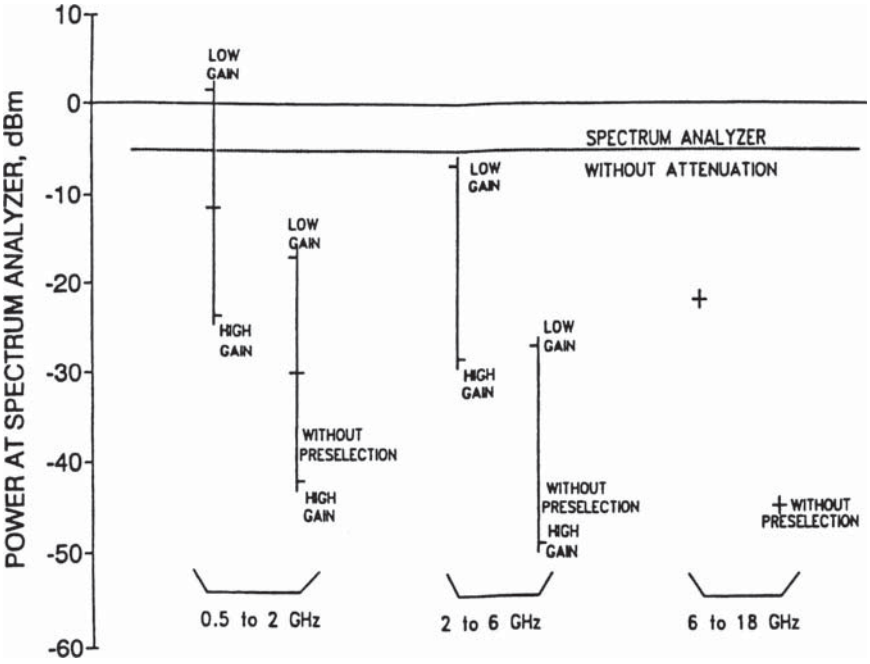


Figure 5-4 Linearity limitations of survey equipment [1]

by separate analyses. In some cases, site frequency managers are available to provide data on existing systems and the installations of future systems. These data are used in analyses that account for the separation between systems, any terrain blockage in the path between systems, transmit power levels, and the gains of the potentially interfering transmit antenna and the receiver's antenna in the directions between the two sites and the interference spectra.

5.2.2 Potential Interference Sources

A second means of defining the interference environment examines potential interference sources that operate near the receiver's design bandwidth and/or have possible harmonics or out-of-band noise that falls within the receiver's bandwidth. This approach is often used for mobile receivers that have to operate in arbitrary locations. The mobile equipment does not experience every possible interference environment, but representative interference sources can be addressed. Often, the goal of these analyses is to determine the minimum required separation between the potential interference and the receiver needed to achieve acceptable receiver operation.

Such analyses proceed by determining the power density variations of the interfering system as a function of separation from that system. When the systems are separated by far field distances, the power density has a $1/R^2$ variation. The interference power received by the system being analyzed equals the incident power density multiplied by the effective aperture of the system being analyzed. The effective aperture equals $\lambda^2/4\pi$ multiplied by the antenna gain, by definition. Generally, a worst-case antenna gain level is used initially to determine the potential of a problem. If a potential problem is identified, the assessment can be refined using updated values. For the space segment, the worst-case levels are the peak gain of the antenna for main beam illumination, and as discussed, for the user segment, an isotropic gain level is commonly used when the interference power is received through the antenna's sidelobes that are widely separated from the main beam. The effects of any terrain blockage between the interference source and the user's antenna must also be addressed to obtain realistic results.

In some cases, the interfering and victim systems are not separated by far field distances. The mutual coupling between payload antennas on a spacecraft is a common example. User systems operating close to large ground terminals can also be within the interference source's near field. While the near field antenna characteristics can be analyzed, an understanding of near field behavior provides sufficient insight to obtain first-order power density estimates. Such estimates can quickly assess the potential of a problem. If this potential exists, more refined analyses and/or measurements can be performed.

Two distinct near field regions exist. The near field of the aperture fields propagates as a collimated distribution located around the antenna's boresight. When transmitting, the near field power density in this region is roughly equal to the transmit power divided by the aperture area. This near field region is seldom of interest because other antennas within this region would block the antenna's main beam and consequently reduce the antenna's gain level. For example, search radars commonly scan in azimuth but strongly illuminating objects close to the main beam would shadow more distant targets or result in multipath errors, thus reducing detection performance. Most search radars have reduced antenna gain levels at low elevation angles to minimize terrestrial illumination so the radar's operation is not limited by clutter returns from the surrounding terrain. The reduced antenna gain level also results in reduced interference power received by user terminals.

Away from this main beam region, the fields are produced by the same mechanisms that produce the wide angle far field sidelobes. For example, the wide angle sidelobes of reflector antennas are produced by the complex sum of direct feed illumination and spillover, blockage from the feed supports, and edge diffraction. Not surprisingly, the highest power densities in this region are typically in the spillover regions and in regions where feed support struts scatter the signal. These same mechanisms produce the far field wide angle sidelobes. The levels of these sidelobes are typically bounded by an isotropic gain level, because if their level were higher, the antenna's directivity would be reduced. These same sidelobe mechanisms combine in a complex sum in the near field, but their phasing, finite range separations, and different angular relationships produce a range dependence unlike the far field behavior. The peak power densities in the near field are still bounded by an isotropic gain value. These assertions have been verified by analytic projections of near field power densities and further validated by measurement.

Measurements can be performed to establish the near field power densities from a transmitting antenna that is a potential interference source. The spectrum analyzer instrumentation described for site surveys can be used for such measurements. The power levels received by the survey antenna can be performed at various locations and distances to measure the near field antenna transfer function that could be experienced by a user terminal operating in the interfering antenna's vicinity. The antenna's near field distributions can also be predicted by using analytic codes; an example of such analyses is described in Fig. 8-7, where a portion of the collimated near field region provides field uniformity sufficient for antenna testing. Such analytic techniques allow identification of near field regions having high-power density values to save measurement time.

The analysis and/or measurement for transmit antennas can also address a related problem. High-power transmit antennas can produce power density levels that raise biological concerns, the radiation hazard problem. A commonly used power density value for radiation hazard levels [2] is 10 mW/cm^2. Measured near field power density levels provide data needed to assess radiation hazard concerns. The power density values of concern are very high and allow measurements to be conducted at reduced transmit power levels. A signal generator can be used in place of the operational high-power transmitter to protect personnel while performing the measurements. In addition, narrow bandwidth waveforms from the signal generator and the spectrum analyzer sensitivity afford ample dynamic range in the measurements. The end objective is to determine locations that personnel should avoid when the system is transmitting at full power.

5.2.3 Standards

A third source of data for interference environment levels results from requirements set forth in standards imposed on system designs. These standards are derived from experience with a variety of systems and operating environments. The levels specified in the standards require system operation without malfunction or performance degradation. The standards address both the allowable radiation from the equipment in stipulating radiated emission requirements and the illumination of the equipment in terms of radiated susceptibility. Both the radiated emission and radiated susceptibility levels cover very broad bandwidths, encompassing both in-band and out-of-band frequencies used by the system. The standards apply to both component testing and integrated system testing. For spacecraft systems, for example, the integrated satellite is evaluated in specialized EMI/EMC facilities as a part of the qualification testing.

EMI/EMC susceptibility standards stipulate radiation levels that couple into the system though the antennas that are commonly referred to as “front door” illumination, as well as interference power that can couple into the receiver through leakage paths, or at IFs (intermediate frequencies) that are commonly referred to as “backdoor” illumination. Generally, EMI/EMC susceptibility measurements illuminate receiver components over wide frequency ranges that can couple into receiver components at a specified incident field level. At microwave frequencies, the power density can be determined from the incident electric field strength because the electric and magnetic fields are related by the free space impedance equal to $120\pi\,\Omega$. However, at low frequencies, such as HF (high frequencies), the free space impedance based on plane wave propagation no longer applies and so levels of both electric and

magnetic field strengths are required to demonstrate EMI/EMC compliance at those frequency ranges. Such low frequencies are a backdoor consideration and principally concern system electronic components. EMI/EMC compliance and attention to shielding receiver components in the design result in the principal interference concern being front door illumination through the antenna system.

A typical level for the susceptibility testing at microwave frequencies is 20 volts/meter. This field strength corresponds to a power density of 1.06 W/m^2 or 0.26 dBW/m^2, required for test purposes. As an aside, the radiation hazard level previously discussed equals 10 mW/cm^2, which corresponds to 20 dBW/m^2. The 20 volt/meter illumination of EMI/EMC measurements is roughly 20 dB lower than the radiation hazard level. The radiation hazard level was derived from solar flux density values so that thermal heating from EMI/EMC testing is not a concern. The susceptibility levels are used in evaluation payload components, but for antenna systems, other limits specify the in-band susceptibility levels as well as the out-of-band rejection requirements near the operating bandwidth. Such limits specify the linear dynamic range required of the receiver and furnish design guidance on filtering requirements. EMI/EMC compliance requires demonstration that the component operates without malfunction or performance degradation. An examination of the values used in the standards as they evolve with time indicate that recent standards have more stringent requirements than earlier standards, reflecting increased awareness of the growing problem of interference.

The testing to establish EMI/EMC compliance necessarily covers the broad bandwidth stipulated in the standards. Calibrated antennas for such measurements are available from specialty vendors. In many cases, small probe antennas are used in "sniff tests" to both receive equipment emissions and to illuminate component pieces. Measurements of spacecraft antennas may use payload antenna models and mockups of portions of the spacecraft in initial studies as a means of reducing the risk of complying during qualification tests. Such measurements are generally within the antennas' near fields and consideration of radiated harmonic levels must be addressed. Generally, the payload antennas are isolated from the payload electronics within the vehicle's bus structure by thermal blankets, panel structure, and Faraday shielding. The principal EMI/EMC attention for spacecraft antennas is therefore directed towards the interaction between payload antennas. For the antenna systems, additional measurements at out-of-band frequencies may be required to evaluate potential susceptibilities. Performing such measurements during development testing for the in-band antenna performance may be schedule- and cost-effective.

These standards also provide design guidance during system development to avoid internally generated interference susceptibilities and

to reduce emissions from the systems that could interfere with other systems. System design generally includes extensive EMI/EMC analyses for the various subsystems comprising the design to avoid interference between subsystems. This process starts with frequency planning within the subsystems to minimize the potential interference (e.g., the selection of local oscillator frequencies, IFs, and so forth). The potential intermodulation products, harmonics, spurious outputs, and noise levels from various system components are included in these design development activities. Such analyses also address frequency components conducted on the prime power. The potential interfering paths in a typical spacecraft can number in the thousands. At the design level, the plausibility of the potential paths should be assessed, and those paths that appear troublesome or have little margin should be specifically identified for further study and evaluation. In many cases, EMI/EMC shortfalls are the result of workmanship errors. Part of the EMI/EMC analyses should address system elements where such errors could lead to noncompliance. Careful attention to these detailed EMI/EMC analyses benefits system testing and identifies factors that merit special attention during such tests.

In terms of susceptibility analyses, the EMC/EMI requirement levels provide guidance in interference studies to determine potential problems. Compliance with the EMI/EMC requirements indicates a level of incident power density that can be tolerated without system malfunction. Thus, compliance with standards provides a "safe" level to address potential system degradation and definition of interference power densities that must be exceeded to produce potential system shortfalls. In some cases, a capability is required to monitor the field strength of incident signals and identify time periods where threshold field strength levels are exceeded. A simple broad bandwidth receiver for this purpose is described in Chapter 7.

5.2.4 Interference Scenarios

Interference analyses often use scenarios to define anticipated numbers and power levels of the interference sources, their locations, spectral characteristics, and operation. Such scenarios are commonly used to assess the effects of intentional interference but are also useful in defining a representative environment to determine the susceptibility of mobile systems. System performance is then evaluated using a defined scenario to determine the degradation caused by interference.

Scenarios unavoidably involve a large number of possible cases that result in variations of the various parameters describing the interference. The large number of potential cases is typically addressed by simulations. The simulation provides answers projecting the system's

performance at a statistical level. The simulation is exercised on a Monte Carlo basis to span the number of simulation variables. The results of this Monte Carlo approach generate the statistical measures of system performance. Such simulations are practical to exercise in software, but would involve an inordinate test time. The simulation is validated by measuring a representative number of test cases. In this way, a practical measurement time is achieved. The agreement between these measurements and the simulation results are used to validate the simulation. Once validated, the simulation is exercised on a Monte Carlo basis to define system performance.

While such a procedure is commonly used, the basis of the scenario in many cases has a subjective nature. Moreover, over the satellite's lifetime, the basis of the scenario and its values evolves and generally becomes more stringent. Further, while a simulation provides a means to quantify system performance, it should be recognized that the operational performance can differ from that projected on the basis of the simulation. Therefore, the examination of the sensitivity of the results to the scenario parameters is important and recommended. An important part of evaluating the performance of a system against a specified scenario is determining not only the system limitations imposed in satisfying the scenario requirements but also how performance is degraded by more stringent scenario parameters.

5.3 Susceptibility Analyses

A wide range of issues must be addressed to evaluate a system's susceptibility to interference. The range of susceptibility outcomes covers damage to the receiver to only minimally degraded operation. The susceptibility analyses must address both the antenna and receiver responses to interference. The antenna response depends on its gain level, G_I, relative to the interference source in both direction and frequency. The interference power received at the output of the antenna's terminal is the product of the interference source's incident power density and effective aperture of the antenna. Receiver susceptibilities to interference include degradation to the desired signal reception; loss of signal acquisition; damage for extreme interference levels; and, in the case of closed-loop antenna tracking, errors in the antenna's tracking performance. The receiver operation for acceptable performance is generally measured by an SNIR (signal to noise + interference) threshold value that can be tolerated in achieving minimum data fidelity and limitations on linear performance and signal acquisition. When closed-loop antenna tracking is implemented by a separate tracking receiver, the assessment of interference susceptibility must be extended to the

tracking receiver. Two questions must be addressed in susceptibility analyses:

1. What is the SNIR relative to its threshold value for acceptable communication service?
2. What is the received interference power level relative to the limits of linear receiver operation?

Interference susceptibility analyses are based on a defined interference environment, the antenna response to interference sources within that environment, and the receiver's response to the interference.

5.3.1 Antenna Response to Interference

The antenna response for the space segment is relatively straightforward. The in-band gain and coverage characteristics are well established in order to verify the system requirements, and reasonable estimates can be made for sidelobe levels beyond the coverage area and within the earth's field of view. In some cases, multiple beams are combined through beamforming networks. The gain levels within the coverage area are roughly the gain of one of the beams reduced by the number of beams combined. The sidelobe levels beyond the coverage area in this case follow the sidelobe variation of an individual beam. Generally, such beamforming networks are passive and the beamformer output is routed to an input bandpass filter and LNA. If the beamformer contains active devices, the response is the same as long as the beamformer's active elements are in a linear operating range. However, if the interference is sufficiently strong to saturate active devices in one or more of the multiple beams, then the performance at those beam positions is degraded by not only the interference but also by the nonlinear response of the electronics. Other beams in the combined outputs can also be degraded by intermodulation products from the beams having a nonlinear response. Active uplink array designs must also address nonlinear responses in addition to interference degradation when the array has a linear response. Active uplink arrays are used so that the uplink G/T is established without being impacted by phase shifter and combiner losses. As long as the active array elements remain linear, the array beam patterns are unaffected. The array element patterns cover the entire earth field of view and the interference must be sufficiently strong to saturate the active electronics of the element. However, sufficiently strong interference would saturate all of the array elements, and so all of the beam patterns formed by the array would be affected. Since the array element pattern has very broad coverage characteristics, the interference source can be located anywhere in the satellite's field of view.

The in-band antenna response for space segment antennas is relatively straightforward to determine. The out-of-band performance is more difficult to address. Space segment uplink antennas have input bandpass filters before the active devices. At a minimum, the received out-of-band power is bounded by the in-band antenna gain levels reduced by the rejection characteristics of the input filtering. Depending on the antenna technology and the interference source's frequency relative to the antenna's design bandwidth, the out-of-band antenna gain values can differ from the in-band performance. An examination of the specific antenna technology is required in this case. In some cases where waveguide is used in the antenna's input circuitry, interference at frequencies below the waveguide's cutoff frequency experience significant attenuation.

The situation for user segment antennas is more complex. In many cases, the interference source does not have a clear line-of-sight path to the receiving antenna and estimates must be made of the terrain blockage losses between the interference source and the receiving antenna. Such losses depend on the specific environments surrounding the ground segment user. Terrestrial interference is received through the antenna's sidelobes because the antenna's main beam is pointed at the satellite. Generally, the antenna gain levels close to the main beam are determined to specify the system's performance, and less attention is given to the antenna's sidelobe response. Often, both the antenna and receiver responses are required at out-of-band frequencies when high-level out-of-band interference can impact the system's performance. Thus, additional analyses and measurement may be needed to quantify the user antenna's response. Further, the exact angular location of the interference is generally unknown; therefore, the antenna's sidelobe levels to be used in link analyses span a large dynamic range. Representative levels in the sidelobe regions must be determined in defining the receiver's susceptibility. In some instances, sidelobe envelope values discussed in Chapter 7 are a convenient means to establish the wide angle sidelobe values. Thus, attention is required to make reasonable estimates of the bounding antenna gain levels to assess the susceptibility to interference for specific applications.

5.3.2 Receiver Response to Interference

The end objective of defining the interference susceptibility is specifying the interference power and spectrum at the receiver's input terminal and determining the impact of received interference power [3] on receiver operation. The two distinct problems, the SNIR for acceptable communication performance and the received power levels relative to the receiver's linearity limitations, must be addressed. The receiver's input terminal is a convenient reference plane to separate the antenna

from the receiver electronics and is also convenient in evaluating system hardware by injecting both the desired signal and interfering signals into the receiver to evaluate receiver operation. Some user segment designs provide couplers for test signal injection as a part of the BITE (built-in test equipment) capabilities to maintain and diagnose system performance. Such couplers also provide a means of injecting signals representing the interference power level and spectral characteristics to evaluate susceptibility. In most cases, however, bench tests are used to evaluate receiver performance.

Receiver responses to a range of signal levels are illustrated in Fig. 5-5 for the receiver's design bandwidth. The receiver operation for interference-free conditions is established by a threshold signal level. This threshold value is generally derived by measuring the receiver's BER (bit error rate) performance. The system requirements stipulate the acceptable BER value, and the relation between the threshold signal level and the system noise level is established by this process. The receiver's implementation loss is the difference between the measured and ideal E_b/N_o values that can be converted into S/N values. Increased input power levels can result in saturation of the analog circuitry, producing desired signal suppression, distortion, and intermodulation products that degrade receiver performance. Further input power increases result in loss of the receiver acquisition, and still more increases in input power damage the receiver.

Digital circuitry in the receiver design must also be considered. Generally, sufficient digital quantization must be provided in the

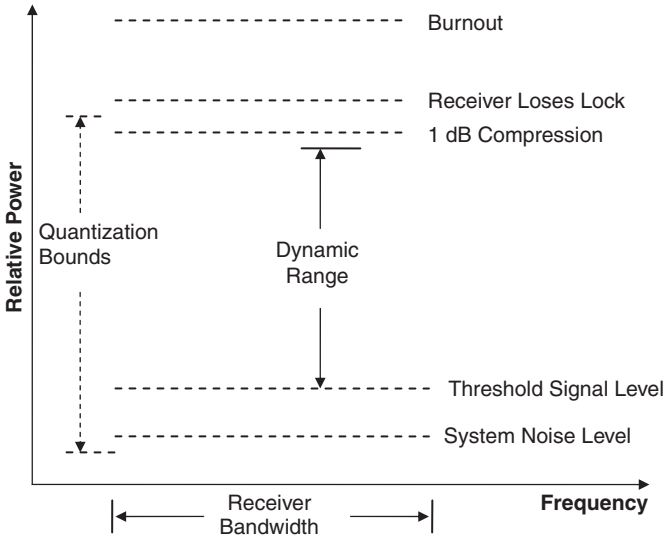


Figure 5-5 Receiver responses versus input power

design to satisfy two objectives. The drive level for the A/D and least significant bit is selected so that the digital quantization noise is lower than the thermal noise floor to avoid degrading the system noise temperature. The required number of bits of quantization is then determined by the objective of having sufficient linearity to extend from the least significant bit to a level somewhat higher than the 1 dB compression point of the analog circuitry. The quantization provides a dynamic range of 6 dB/bit so that 12-bit quantization spans a 72 dB dynamic range, for example. Additional attention must also be given to analog anti-aliasing filtering so that out-of-band interference does not fold into the desired signal spectra. The linear dynamic range for user signals extends from the threshold signal power to a level somewhat short of the 1 dB compression point that has sufficient linearity to avoid signal distortion. In systems that require very high dynamic range, AGC (automatic gain control) circuitry can vary the receiver gain levels in response to signal input power. In interference environments, however, care must be exercised so that interference signals do not capture the AGC circuitry and suppress the desired signal components into the thermal noise.

The operation of carrier tracking loops and/or frame or bit synchronization can also be disrupted by interference. Typically, such disruption results from a high-level pulse, but loss of lock or synchronization can also occur when another signal with similar characteristics to the desired signal and a slightly high-power level "captures" the loop and/or synchronization. The receiver must reacquire the desired signal after the interference is past, and the receiver is disrupted for the interference period, the time for the interference to pass through the receiver's filters, and the reacquisition time for the desired signal.

Like the antenna response, the receiver response to out-of-band interference must also be addressed. In comparison to the in-band signal levels depicted in Fig. 5-5, the out-of-band receiver's response to interference power is impacted by the receiver's filtering in the analog circuitry. The filtering before the LNA dictates the spectrum that can enter the receiver circuitry. The level of out-of-band interference is reduced by the filter selectivity performance. Similarly, filtering at the IF level further rejects out-of-band interference. Such filtering is even more effective than RF filtering because of increased selectivity and more selective filter shapes because IF insertion loss is not as significant an issue as RF insertion loss. Additionally, adequate filtering at the IF level must be provided to avoid aliasing out-of-band interference into the receiver's design bandwidth. Interference somewhat outside of the receiver's operating bandwidth is rejected by the receiver's processing and does not degrade the receiver's performance as long as the receiver maintains a linear response.

The receiver design has two conflicting requirements. The first requirement is to minimize the receiver noise temperature. The second requirement is to maximize the linear dynamic range. The first requirement implies minimal RF filtering and a high gain, low noise LNA so that the cascaded noise contributions of the remainder of the receiver electronics are minimized. High linear dynamic range requires minimum analog gain to increase the input signal power to result in saturation and significant filtering to reject out-of-band interference. These conflicts are resolved in a process known as "gain and filtering partitioning" that examines the interference environment for a specific application and tries to satisfy the linear dynamic range requirements while maintaining a low receiver noise temperature.

Gain partitioning examines the noise figure and saturation characteristics of candidate analog components for a receiver. A typical specification for receiver saturation is the 1 dB compression point at the receiver input. At this input level, the receiver output differs from its linear response by 1 dB. This receiver saturation can occur at frequencies well removed from the design bandwidth so that the distribution of the receiver filtering requires examination. The selection of the filter shape for the RF filter and the gain of the LNA requires examination of the anticipated levels of the out-of-band interference. Increased filter selectivity is accompanied by increased insertion loss. Reducing the gain of the LNA increases the power level at the receiver input needed to saturate the LNA, but the noise contributions for the components following the LNA are also increased. The filtering at the IF levels provides additional selectivity and rejection of image components. Selectivity tradeoffs and the gain selection at the IF levels require examination of an anticipated interference environment close to the design bandwidth and analyses to assure sufficient anti-aliasing filtering is provided for the digital circuitry. While the goal is to assure the receiver noise temperature is dominated by the initial preamplification, requirements for linear dynamic range can result in some increases in the receiver noise temperature.

The evaluation of receivers begins with quantifying their performance in interference-free conditions. Testing generally starts by using network analyzer instrumentation to verify the gain and filtering distribution in the analog circuitry. The linearity is also measured using signal generator measurements. The receiver noise temperature is measured by using Y factor techniques or noise figure instrumentation (described in Chapter 8). The analog circuitry after its evaluation is connected with the digital circuitry and the demodulation evaluation is started. Receiver development testing generally includes BER tests to evaluate receiver performance and to quantify receiver implementation losses relative to an ideal receiver for purposes of determining link budget losses. Implementation loss is defined as the difference between the

ideal E_b/N_o and the value required of the actual receiver to meet the BER specified for the system. These measurements are made using standard BER test sets or the modems to be used operationally. The receiver evaluation when interference is present is then started.

The required SNIR must be established and depends not only on the received interference power but also the spectral characteristics of the interference. Generally, measuring interference effects using the specific receiver hardware is the most satisfactory means of evaluating interference effects. Both desired signals and interfering signals are required. The BER test set or operational modems are used for desired signal representations since SNIR levels are also evaluated using BER measured. Additional signal generation is required to represent interference signals. Additional testing to quantify interference limitations should be made since such testing can be easily performed in a cost-effective manner at this time. The interference can be injected into the receiver along with the bit stream used for BER evaluations and different types of interference (e.g., noise, CW tones, pulses, and so on can be readily obtained from signal generators). The goals of such measurements are to determine a threshold SNIR that achieves the specified data fidelity expressed by a required BER value and to evaluate interference levels needed to result in loss of lock and/or synchronization along with the required receiver acquisition time. The user receiver should have indicators for loss of lock and/or synchronization and for interference presence, and the operation of these indicators can also be evaluated. While these measurements are recommended during receiver development, similar measurements are often performed on existing receivers to evaluate the effects of interference experienced in operational conditions. The effects of specific interference sources can be evaluated by operating the receiver in that environment and measuring interference values with a spectrum analyzer.

The end goal for the receiver examination is to determine limits of the existing receiver design as well as possible impacts to receiver performance of possible changes to further reduce receiver vulnerability to interference. For example, more stringent filtering at the receiver input increases the filtering loss and degrades the receiver noise temperature and consequently sensitivity. The receiver dynamic range may be increased by reduced preamplifier gain and more selective filtering in the downconverter and IF amplifiers at the expense of increased receiver noise. Additional quantization in digital receiver designs increases the dynamic range but may be limited by component technology. Automatic gain control techniques might be employed, but practical time constants in such circuitry may limit effectiveness for pulse interference. Such tradeoffs in the receiver design need to be understood to address techniques for possible reductions to receiver interference vulnerability.

This discussion concerns a general receiver architecture that would be used in user segment applications. The space segment designs using linear frequency translators contain the analog portions of the receiver but not the demodulation circuitry. Additional analog circuitry exists in frequency translation and transmission circuitry. The filtering and dynamic range issues of the analog transponder follow the previous discussion. Regenerative repeater architectures do have demodulation and signal separation circuitry that follows the previous discussion. Closed-loop antenna tracking designs often have separate tracking receivers. If interference degrades the tracking receiver performance, antenna point loss increases. In some cases, the presence of interference is indicated by AGC levels in the tracking receiver or separate thresholding circuitry, such as used in the receiver described in Chapter 7, providing identification of excessive interference power. One approach to interference mitigation for tracking receivers simply commands the antenna to program track when the presence of interference is identified.

5.3.3 Receiver Damage

Low-noise receivers have gone through much development, and LNAs with very low noise temperatures are widely available over a broad frequency range. The low noise temperatures, however, require devices having very small (submicron) dimensions. While these dimensions are necessary to achieve the low-noise performance, the end result is a device that is limited by its thermal dissipation. Consequently, sufficiently high-level RF inputs [4, 5, 6, 7, 8, 9] can exceed the thermal dissipation limits, cause the device's substrate to melt, and irreversibly damage the device. This "burnout" results in attenuation rather than the amplification the device is intended to provide. Typical power input values for device burnout, shown in Fig. 5-6, illustrate these points. For very short exposure times, the required burnout power varies as $1/(t)^{½}$, a behavior typical of thermal dissipation limits. Very shortly, roughly 10 nanoseconds after pulse initiation, the required input power level for device burnout reaches a steady state limit. When the power level is less than this threshold value, the device survives no matter how long the power is applied. Thus, the power level rather than the amount of energy is the critical factor in burnout. The RF power needed for burnout exceeds the dc bias so burnout occurs whether or not the device is powered. Since low-noise devices have only 6 or 7 dB of gain, typical LNAs have several devices in series to achieve the overall amplifier gain performance. The devices typically saturate at a 10 to 13 dBm level at their output. Sufficiently high power level inputs will damage the first device of the amplifier but the devices following the first device

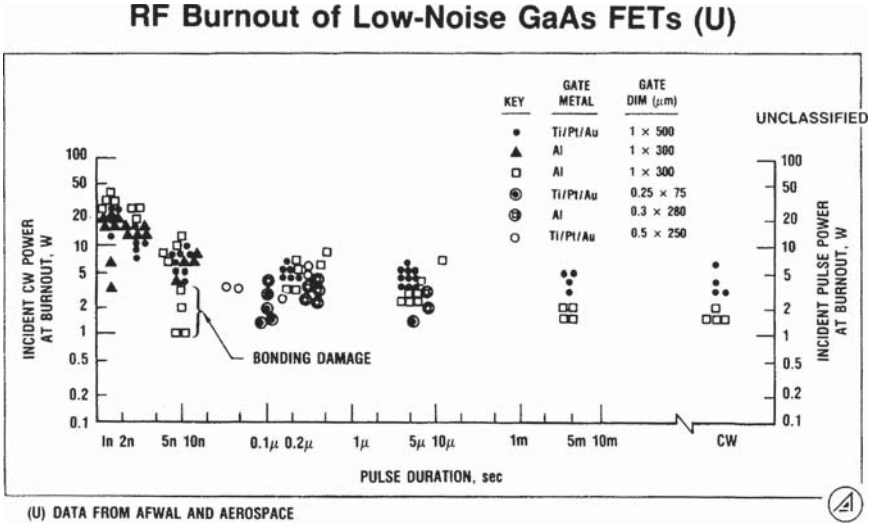


Figure 5-6 LNA burnout characteristics [4] (© 1984 IEEE)

are protected because of the limited saturated output levels of the devices.

Low-noise devices have a wide bandwidth response and when they are integrated into the low-noise amplifier, device matching techniques are used to meet the specified noise figure and gain flatness requirements over their design bandwidth. The broad bandwidth of the device persists after matching is performed. While not compliant with noise figure and gain flatness performance in the operating bandwidth, the amplifier has a high gain response over a considerable bandwidth. For this reason, adequate filtering must be provided to maintain receiver linearity when high-level out-of-band signals are present.

If systems are required to operate in environments where burnout is a potential concern, diode limiters can be placed before the LNA to protect the LNA from high power input levels. The input RF filter is required to precede the limiter so that high-level out-of-band interference does not activate the limiter needlessly. The insertion loss of highly selective filters and protective limiters that are necessary in adverse environments degrades the receiver noise temperature. Practical limiter designs for these applications have a few tenths of a dB insertion loss. The limiter requires series and shunt diodes to achieve response times short enough in comparison to the short times that devices can be damaged. Other parameters specify limiter performance besides insertion loss and response time. The activation level describes the input power level that initiates limiter operation. Spike leakage is the peak power level leaked

through the limiter for a pulse having a rise time much shorter than that required for burnout. The steady state rejection is the power level passed by the limiter after the transient response to an input RF pulse. Finally, the power handling capability is the maximum input power level that can be continuously applied without damaging the limiter. If limiters are used, integrating the limiter with the LNA is generally considered to avoid the losses and reliability concerns of connectors.

5.3.4 Link Analyses

The susceptibility of receiving systems to interference is analyzed by link analyses and an examination of the receiver response to the interference source's spectrum. The link analyses generally examine received interference levels relative to the required SNIR value and the received power level relative to the linearity limitations of the receiver's design. These analyses are examined in turn.

The SNIR can be written in the following manner

$$\text{SNIR} = M/[N/S + I/S]$$

where M is the system's margin relative to the threshold BER performance in interference-free conditions and I/S is the interference-to-signal ratio referenced to the signal level satisfying the same threshold BER performance. In this way, the interference level is referenced to the specified threshold signal level and the margin achieved under operational conditions directly increases the achieve SNIR. The threshold signal-to-noise ratio, S/N, in interference-free conditions is related to the required E_b/N_o to meet BER requirements as

$$S/N = E_b/N_o\, R_b/B$$

where R_b is the bit rate and B is the signal's bandwidth. The interference-to-signal ratio, I/S, can be expressed from the respective link equations and equals

$$I/S = (\text{ERP}_I/\text{ERP}_S)(G_I/G_S)(L_I/L_S)(R_S/R_I)^2(\lambda_I/\lambda_S)^2$$

where the subscript "I" denotes the interference parameters and "S" denotes the signal parameters.

The parameters take on different values depending on the system's interference analysis. The ratio of the ERP values clearly depends on the signal and interference values being assessed. The antenna gain G generally differs between the antenna gain level in the signal and interference directions. The interference received by user systems typically arrives in the antenna's sidelobe region so that the G_I/G_S ratio has a small value. For space segment uplink antennas, the angular separation between the coverage area and the interference source dictates the

antenna gain ratio value. System loss, L, can have a significant variation in values for interference and signal paths. For example, interference to user systems typically does not have a free space path and terrain blockage loss can result in a low value of the loss ratio. The loss values also include *RF* filter rejection so that the interference power levels at the receiver input terminals are obtained. The range values R can be comparable for interference paths for the space segment and vastly different for the user segment. The wavelength, λ, should be comparable for the interference and signal components in most cases so that their ratio is close to unity.

The received interference levels are a straightforward application of link analysis and equal the incident power density multiplied by the effective aperture

$$I = \mathrm{ERP}_I \, G_I \, L_I / (4\pi R_I / \lambda_I)^2$$

As is the case for the SNIR analyses, a wide range of parameter values exist, and depend on space or user applications.

5.4 Interference Mitigation Techniques

A variety of techniques exist to mitigate interference. Interference mitigation has both applications to specific problems that arise in operation and system planning and design. Operational solutions commonly arise in user segment applications. Such solutions may include additional filtering, moving the system away from interference sources, and limiting the operation of the interference source. Filtering can be added to the user's system to reduce interference or additional filtering can be added to the interference source. For example, the addition of filters to an offending system to reduce harmonic radiation that interferes with user operation might be a good operational solution. Operation in a location separated from the interference source can be carried out, or relocating the user antenna to provide shielding from buildings is another example. Using spectrum analyzer instrumentation to guide such relocations is helpful in selecting relocation alternatives. Another form of operational solutions is to limit the operation of the offending system. For example, high-power radar systems may not need full coverage, and sector blanking techniques can be employed. Sector blanking refers to turning off the radar transmitters in some sectors where radar information does not have high interest and where the user terminal is located. In this way, the user terminal is illuminated by the lower gain sidelobe levels of the radar's antenna rather than the higher level illumination from the radar antenna's main beam. Such operational solutions are particularly useful for specific problems in user system applications.

The user segment clearly has more options to incorporate upgrade designs affording increased interference protection than the space segment. If interference proves to be troublesome at specific locations, quantifying the interference levels and spectra as described earlier is the first step. Operational solutions should be examined for specific applications since such solutions are generally the most cost-effective. In other cases, additional interference protection can be obtained by minor system upgrades such as increased filter selectivity or antenna modifications to reduce sidelobe levels that can be applied to specific applications. Generally, system upgrades prove to be costly and, therefore, the specification of interference mitigation techniques in system requirements should have a conservative bias to minimize the necessity of costly upgrades and retrofits.

Interference mitigation capabilities for specific programs start with the system definition of requirements. The space segment's requirements for coverage often have some inherent requirements for interference mitigation. For example, polarization and frequency reuse techniques require bounds on antenna polarization purity and sidelobe levels so that users sharing the same frequency band in different coverage areas have sufficient isolation to avoid co-channel interference. Providing multiple narrow beams to satisfy coverage requirements is effective at isolating interference to a small portion of the coverage area, but comes at the expense of a larger and more complex antenna. The choice of modulation waveforms is a further example. Spread spectrum modulation techniques can be specified in space segment requirements as a design feature to provide interference protection.

Three distinct design techniques exist to provide additional protection from interference: (1) spread spectrum modulation, (2) low sidelobe antenna designs, and (3) adaptive interference cancellation. Spread spectrum modulation techniques are well developed and applied in many applications. Antenna techniques for interference mitigation [10] have been widely demonstrated. Devising a combination of these techniques to provide adequate cost-effective protection from interference for a given application is the system challenge. Each of these techniques will be briefly described in turn with additional discussion on the application of the techniques in Chapters 6 and 7.

5.4.1 Spread Spectrum Modulation

Spread spectrum modulation [11] distributes the desired information over a much wider bandwidth than normally used and consequently dilutes the effects of interference. Two classes of spread spectrum modulation exist. One class is frequency hopping, where the carrier is hopped in a random sequence over a wide bandwidth to distribute the information.

The second class is direct sequence, which modulates the desired information with a wide bandwidth pseudorandom sequence. In both cases, the transmitter and receiver understand how the information is spread (i.e., the frequency hop sequence or the random code). This information must remain unknown to the interference source.

Irrespective of the details of the spread spectrum modulation, the interference is effectively reduced by the spread spectrum processing gain. The spread spectrum processing gain to first order equals the ratio of the spread bandwidth to the bandwidth normally used to transmit the data. The spread spectrum processing gain is therefore ultimately limited by the available bandwidth and the data rate of the signal. For example, if a QPSK signal that provides 1bit/Hz capability is used for a 1 Mbps rate signal and the signal is spread to a 100 MHz bandwidth, the spread spectrum processing gain is a factor of 100, or 20 dB.

Spread spectrum modulation is generally used with error correcting coding and interleaving. Coding techniques are effective for random distributions of errors and interleaving endeavors to randomize the distribution of errors so that coding is effective. These modulation techniques provide a powerful means of protecting receivers from interference. Generally, the modulation techniques must be addressed during the design definition of the program.

5.4.2 Low Sidelobe Antennas

Interference received through the antenna sidelobes can be reduced by lowering the antenna's sidelobe levels. Antenna sidelobe reduction techniques provide a twofold advantage for systems whose antennas are required to both transmit and receive. The interference levels received from other sources and potential interference transmitted to other receivers are reduced by lowering the antenna's sidelobe levels. Techniques for antenna sidelobe control have been developed for a variety of applications and technologies.

Techniques for antenna sidelobe control can be divided into two angular regions. The first region applies to those sidelobes near the main beam. In this region, the antenna pattern is principally dictated by the aperture distribution. As illustrated earlier in Table 1-1, antenna sidelobe control near the antenna's main beam requires maintaining a linear phase distribution and tapering the amplitude distribution. Reducing the sidelobes near the antenna's main beam unavoidably reduces antenna efficiency and broadens the antenna's beamwidth. The sidelobe levels near the antenna's main beam also depend on aperture blockage from feeds, subreflectors, and support struts in reflector antennas. Offset reflector designs are advantageous because aperture blockage is avoided. The space segment antennas are excellent examples in

which the control of sidelobes near the beam is an important design issue. The achievable sidelobe rejection (as will be discussed in Chapter 6) depends on the interference location relative to the design coverage area and the sidelobe rejection for interference sources located beyond the design coverage area. Sidelobe reduction is particularly important in providing isolation between coverage areas in space segment antenna designs so that frequency reuse can be used without incurring co-channel interference between users in different coverage areas that share the same frequency subband.

The second angular region is the sidelobes removed from the main beam region. The wide angle sidelobe levels result from radiation mechanisms other than the aperture distribution. Antenna feed spillover and direct radiation, edge diffraction, and blockage scattering and diffraction in reflector antenna technology are examples of mechanisms that produce wide angle sidelobes. Since these sidelobes result from mechanisms other than the aperture distribution, their control has little effect on the antenna's main beam gain and efficiency. The need for wide angle sidelobe control commonly arises for user antennas because their main beams are directed towards the satellite and, consequently, terrestrial interference arrives through the wide angle sidelobes.

The sidelobe control requirements differ between the space and user segments. The space segment is required to service its design coverage area. Ideally, the sidelobes removed from this design coverage would be reduced to avoid interference from the remaining field of view. The use of narrow antenna beamwidths increases the interference protection for users located outside of the design coverage area, and existing and future designs, particularly at higher frequencies where compact antenna systems can be configured, will benefit from the sidelobe rejection of narrow beamwidth antennas. The user segment generally requires interference protection from terrestrial interference that arrives through the antenna's wide angle sidelobes. Chapter 7 describes a variety of wide angle sidelobe control techniques that have little effect on the antenna's main beam gain level. As interference mitigation becomes more important in future years, commercial development and availability of user antenna designs with much lower wide angle sidelobe levels than existing commercial designs can be anticipated.

5.4.3 Adaptive Interference Cancellation

Adaptive interference cancellation techniques have developed greatly in recent years [12, 13]. Adaptive systems dynamically combine antenna elements in such a way that a null in the overall antenna pattern is aligned with the direction(s) of the interference. These pattern nulls are dynamically formed in response to interference and "cancel" the interference.

Changes in the interference environment result in changes in the pattern null positions. Adaptive cancellation systems require:

1. A means of distinguishing desired signals from interference
2. An ability to sense the presence and initiation of interference
3. A control processor capable of determining the appropriate adaptive weight values that combine a set of antenna elements to satisfy an optimization criterion
4. The necessary antenna elements and the circuitry to adjust the weighting determined by the control processor
5. The capability to monitor the performance measured by the optimization criterion to identify changes in the interference that may require readjustment and appropriately respond

Clearly, such steps can be satisfied in a variety of ways and an appropriate choice depends significantly on the application. Example designs may be found in Chapter 6 and 7 for the space and user segments, respectively.

Adaptive cancellation is inherently a subtraction process. A pattern null requires the adaptive antenna to produce two antenna responses having equal amplitude and opposite phasing so their sum produces a pattern null. The antenna elements are combined by adaptive weighting values to produce a null(s) in the direction of interference sources. Subtraction to form significant nulls is accompanied by tight tolerances on amplitude and phase matching. The tolerances illustrated in Fig. 5-7

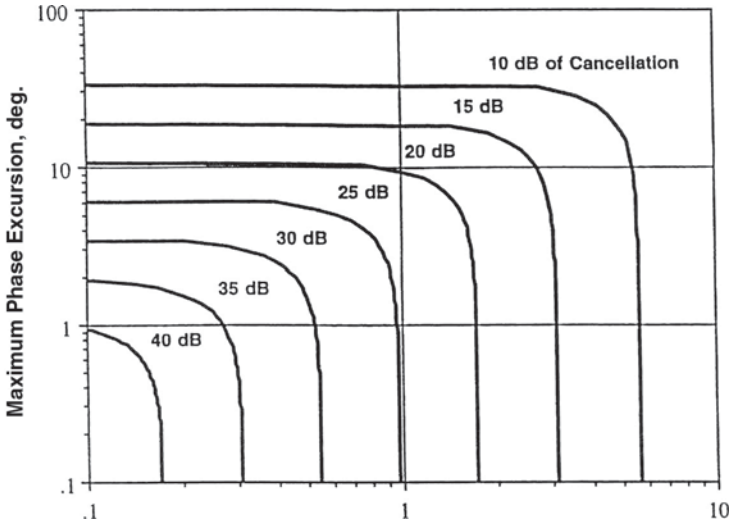


Figure 5-7 Amplitude and phase-matching requirements [10] (© 1995 IEEE)

indicate that tolerances become increasingly stringent as the required cancellation level increases. This amplitude and phase-matching performance must be maintained over the system's operating bandwidth to achieve effective cancellation. Cancellation performance is degraded by many factors, including dispersion contributions from the antenna, amplitude and phase tracking imperfections in the system electronics up to the point at which adaptive combination occurs, and the precision of the determination and adjustment of the adaptive weight values used to combine the antenna elements. The design complexity also increases with the required number of pattern nulls (e.g., interference sources and the cancellation bandwidth).

The adaptive weighting values are derived to satisfy an optimization criterion [13] and are conventionally determined in three alternative ways. The most commonly used technique is an LMS (least mean square) algorithm. The weighting values are determined by the adaptive processor that functions as a control system. The error in this case is the interference, and it can be shown the weighting solution is a quadratic process so that a simple recursive algorithm can be used. Such a solution rapidly provides a solution that maximizes the SNIR. Another technique is an iterative weight adjustment process that varies the weighting while monitoring the interference power levels. This process is relatively slow and subject to dynamic variations in the interference. A third technique is referred to as a sample matrix inversion, where the processing forms an interference correlation matrix that is inverted to yield a solution. This technique is limited by the accuracy in determining the correlation matrix values.

A fourth technique has been developed [14, 15, 16] and is described as DEADEN (DEterministic ADaptive Environmental Nuller). This technique departs from conventional algorithms by measuring the interference sources' directions and using a priori antenna design information to command pattern nulls. A simple example using array antenna technology assumes the interference is received through the antenna sidelobes. The array forms another antenna beam that sequentially samples the antenna sidelobes to identify those sidelobes that contain interference power. In this way, the direction of the interference source can be determined. Once the array sidelobe containing interference power is identified, another array beam is steered to that sidelobe's location, commanded to match the known level of the sidelobe, and subtracted from the sidelobe to produce a null canceling the interference. The extension of DEADEN techniques to adaptive uplink multiple beams has been explored.

Adaptive system development is generally guided by simulation efforts to identify the performance capabilities of the design. The simulation program can address the parameters describing the interference

sources on a Monte Carlo basis and can make a statistical projection of adaptive system performance. As the system development proceeds, measured hardware performance of the system can be incorporated into the simulation to improve its fidelity. Testing the adaptive system on a Monte Carlo basis would require an impractical amount of test time. In practice, the testing uses the simulation program to develop a limited number of test cases. This limited number of test cases is measured to validate the simulation. Having validated the simulation, Monte Carlo results from the simulation program are used to establish system compliance of the design. Thus, the development of a detailed simulation is a very important part of adaptive system development. Generally, adaptive system designs must address two system-level questions.

1. What is the steady state performance when interference is present?
2. After the initiation of interference, what length of time is required to achieve steady state performance?

For the space segment, the steady state performance is measured by the percent of the design coverage area remaining after the adaptive weights converge, where communications are possible when interference is present. For the user segment, the steady state performance is judged on the basis of the SNIR achieved under steady state conditions. The time after the initiation of the interference to reach steady state performance is generally measured by the time for convergence of the adaptive weight values to their steady state levels. The user segment evaluation must address both the time for the adaptive system response and the time for receiver reacquisition if interference causes the receiver to break lock.

References

1. R. B. Dybdal, G. M. Shaw, and T. T. Mori, "A RFI Measurement System for Field Sites," *1995 AMTA Symposium Digest* (November 1995).
2. ———, *IEEE Standard Test Procedures for Antennas* (New York: Wiley-Interscience, 1979).
3. R. B. Dybdal, "Analyzing Receiver Vulnerability to Interference," *1995 IEEE MILCOM Symposium Digest* (November 1995).
4. R. B. Dybdal, J. F. Heney, M. McColl, and H. J. Wintroub, "The Technology of the Electromagnetic Threat to Space System Receivers," *1984 IEEE MILCOM Symposium Digest* (October 1984).
5. R. B. Dybdal, K. M. SooHoo, and H. J. Wintroub, "Limitations of Multiple Source Combining Against Survivable Communication Systems," *1989 IEEE MILCOM Symposium Digest* (October 1989).
6. A. A. Moulthroup, M. H. Muha, R. B. Dybdal, and H. J. Wintroub, "HPM Damage Thresholds of Low Noise GaAs FETs and HEMPTs," *Fifth National HPM Conference* (June 1990).
7. M. A. Johnson, R. B. Dybdal, S. Sokolsky, and M. Hopkins, "High Power Radio Frequency Weapon Threat Analysis for Geosynchronous Satellites," *2007 IEEE MILCOM Symposium Digest* (October 2007).

8. J. J. Whalen, M. C. Calcatera, and M. L. Thorn, "Microwave Nanosecond Pulse Burnout Properties of GaAs MESFETS," *IEEE Trans Microwave Theory and Techniques*, vol. MTT-27 (December 1979): 1026–1031.
9. J. J. Whalen, R. T. Kemberley, and E. Rastefano, "X-Band Burnout Characteristics of GaAs MESFETS," *IEEE Trans Microwave Theory and Techniques*, vol. MTT-30 (December 1982): 2206–2211.
10. R. B. Dybdal, "Assessment of Antenna Interference Reduction Techniques for Communication Satellite Systems," *1995 IEEE MILCOM Symposium Digest* (November 1995).
11. R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications* (New York: Prentice Hall, 1995).
12. L. H. Sibul, *Adaptive Signal Processing* (New York: IEEE Press, 1987).
13. R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays* (New York, Wiley, 1980).
14. R. B. Dybdal and R. H. Ott, "Apparatus and Method for Employing Adaptive Interference Cancellation over a Wide Bandwidth" (August 8, 1995): U.S. Patent 5,440,306.
15. D. J. Hinshilwood, R. B. Dybdal, and K. M. SooHoo, "The Cancellation of Interference by Sidelobe Annihilation," *1991 IEEE AP-S Symposium Digest* (June 1991).
16. R. B. Dybdal and D. J. Hinshilwood, "DEADEN: A New Adaptive Cancellation Technique," *1995 IEEE MILCOM Symposium Digest* (November 1995).

Chapter

6

Space Segment Antenna Technology

6.1 Overview

Antennas greatly contribute to the performance and capabilities of the space segment [1] and are visually prominent features of satellite designs. The differing requirements for satellite systems result in a wide range of satellite antenna technologies, and an even more diverse technology range has been carried through development. The space segment antenna technology has more diversity than any other payload subsystem because of the differences in program requirements.

Space segment antennas can be broadly separated into technologies that support different coverage areas. Early satellite antennas were necessarily constrained by limited payload capabilities, satellite attitude stability, and low frequency operation, thus simple antenna designs provided very broad coverage. As described in Chapter 2, broad coverage antennas and earth coverage antennas continue to have applications, but the narrow coverage antennas described in this chapter provide system designs that are largely responsible for present-day communication capabilities. These narrow coverage antenna systems became practical as satellite payload capabilities and attitude stability increased and as launch capabilities to geosynchronous altitudes became available. Today's systems benefit from operation at higher frequencies where broad bandwidth frequency allocations are available and compact antenna systems can be implemented. At the same time, increased demand for satellite services, particularly in commercial designs, is made possible by frequency and polarization reuse techniques that achieve greatly increased communication capacity. A compendium

that describes the variety of satellite programs [2] and the top-level characteristics of the technologies used in their implementation illustrates the diversity of technology that has been developed to satisfy a variety of system requirements and applications.

Spot beam antenna technologies that communicate the same information to a single coverage area, and multiple-beam antenna technologies that provide independent beams that individually service portions of a composite coverage area, have been widely used in space segment antenna designs. Spot coverage antennas and multiple-beam antennas are widely used for both commercial and military space applications. In recent years, the trend towards integrating antennas with system electronics has increased, resulting in antenna systems that provide performance benefits while imposing additional challenges in development and testing. Adaptive antenna systems to reduce uplink interference and active aperture downlink antennas to combine the outputs of many low-level solid state transmitters illustrate antenna systems whose performance strongly depends on system electronics. Space segment antennas that provide point-to-point connectivity are used in crosslink subsystems and for earth links having high data rate requirements.

6.2 Spot Beam Antennas

Spot coverage antennas service only a portion of the earth's field of view and have several distinct benefits. Spot antennas provide higher gain performance than earth coverage antennas that service the entire earth field of view. This increased antenna gain performance for spot antennas' smaller coverage requirement enhances users' system communication capacities and/or reduces users' performance requirements. The antenna beamwidth necessary to cover a portion of the earth is generally defined by the antenna's footprint, which outlines the antenna coverage projected on the earth's surface. Representative values of spot coverage for geosynchronous satellites (shown in Fig. 6-1) illustrate the minimum footprint values at the subsatellite point and the longest footprint dimension for users at a 30^{o} elevation angle where the footprint spreads out over the earth's surface. Separating different portions of the earth field of view into different coverage areas affords the possibility of reusing the frequency and/or polarization to obtain greater capacity from the allocated frequency spectrum. These benefits are accompanied by larger antenna dimensions compared with earth coverage designs, more numerous antennas to service the satellite's field of view, and the need for routing signals between antenna coverage areas within the transponder. Spot beam coverage is widely exploited in both commercial and military space segment designs to increase system capacity,

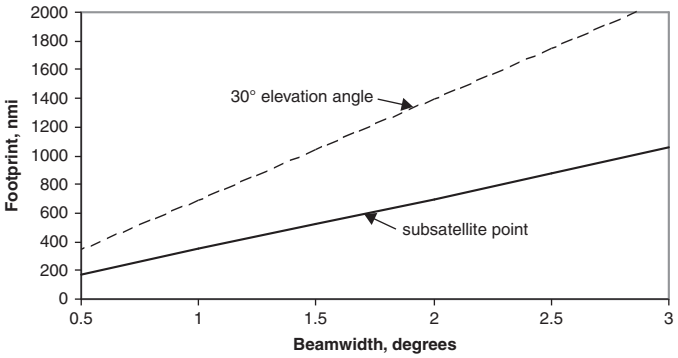


Figure 6-1 Spot coverage versus antenna beamwidth

and these space segment antennas for spot coverage requirements are prominent features of present and future satellite designs.

A variety of different antenna techniques are used for spot coverage requirements. A very simple design for spot coverage uses a reflector antenna to generate a narrow beamwidth, which is selected to conform to the specified coverage area and is mechanically positioned to point at the centroid of the desired coverage area. As communication needs change during the satellite lifetime, the antenna can be repositioned to service other coverage areas. Typically, satellite designs use more than one spot beam antenna to provide communications in different locations. If the spot beam antenna's coverage areas are sufficiently separated, the same frequency band and polarizations can be reused in both coverage areas to increase communication capacity because of the spatial isolation between coverage areas afforded by the antennas' sidelobe responses.

In some communication applications, geopolitical boundaries must be serviced for the entire satellite's lifetime. Often, these applications also have regulatory requirements concerning the allowable radiated power densities outside of the geopolitical boundaries. In such situations, the coverage provided by a single spot beam may be inadequate because the coverage of irregular geopolitical areas is not uniform or the regulatory requirements on power density levels outside the coverage area are not satisfied. Antenna techniques to conform more closely to such irregular contours have been developed.

One means of satisfying irregular coverage requirements uses a larger aperture having a cluster of feeds within the focal region in place of a smaller antenna having a single spot beam. The independent beams produced by the feed elements in the focal region are then combined to produce pattern contours that conform more closely to the irregular

coverage boundary. The antenna design in this case typically uses offset reflector technology, and the arrangement of the individual feed horns within the cluster mimics the boundaries of the desired coverage area shape. The fit to the irregular coverage boundary improves as a larger number of narrower antenna beams are combined to produce the coverage contour. In addition, the antenna sidelobe reduction beyond the coverage area increases since the sidelobe rolloff follows the pattern of a narrower independent antenna beam pattern. However, this improved fit to the coverage area boundaries is achieved at the expense of an increased number of feeds in the cluster and a larger overall aperture size. An example of generating coverage areas by beam combining [3] and multiple-beam coverage is illustrated in multifrequency designs for the Japanese nation. The coverage characteristics can be varied during the satellite's lifetime if provision is made in the design to change the combination of antenna beams that produce contour coverage.

A second means of satisfying coverage requirements for irregular areas also uses a larger aperture compared to a smaller antenna, providing coverage by a single beam. In this design approach [4], the larger aperture is fed by a single feed and the reflector surface is deformed to produce the required coverage area. The design of the deformed reflector surface can be developed using a synthesis approach [5]. The coverage contours are fixed by the particular deformation configuration used in the reflector surface and cannot be changed on-orbit.

While both of these approaches to producing irregular coverage areas require larger apertures than that for a single antenna that produces a spot beam, the larger aperture provides four benefits:

1. The sidelobe levels beyond the design coverage area are lower than those in a smaller spot beam antenna using a single coverage beam.
2. These reduced sidelobe levels beyond the design coverage area reduce interference both to and from sources located outside of the design coverage area.
3. These reduced sidelobe levels provide increased isolation between other coverage areas in frequency reuse applications, allowing either an increased number of coverage areas or closer angular spacing between coverage areas.
4. The reduced sidelobe levels and beam combining result in a more directive pattern having more uniform coverage, thus increasing communication performance.

While payload designs generally endeavor to reduce weight and complexity, the performance advantages of these spot antenna designs for

irregular coverage areas and the system requirements for frequency reuse to expand system capacity dictate their use.

6.3 Multiple-Beam Designs

In contrast to spot beam antennas that service a coverage area with a beam pattern that communicates a single collection of signals, a multiple-beam antenna services its coverage area with multiple beams having independent data streams routed to different portions of the composite coverage area. In this way, increased capabilities are provided to the composite coverage area. Since each beam services only a portion of the coverage area, the space segment's antenna gain level in each portion is greater than it would be if the composite coverage area were serviced by a single beam. This increased space segment antenna gain level results in higher data rate services to users and/or reduced user performance requirements.

Generally, multiple beams are produced from a single aperture capable of simultaneously generating more than one beam. The ITALSAT design [6] discussed in Chapter 3 is an exception that uses more than one aperture to achieve the total multiple-beam coverage. Two apertures each providing three multiple beams produce the six beam coverage of the Italian nation. When a single aperture is used, each of the multiple beams has a common aperture phase center located at the aperture's center. The benefit of this feature is that beams can be coherently combined to serve a larger coverage area without suffering grating lobes that would degrade performance. Another significant benefit of multiple-beam designs is that they are more compact and lighter than multiple, discrete antennas. The available space on a satellite, particularly on its earth-facing side, is in high demand; therefore, simultaneously generating multiple beams from a single aperture is an attractive SWaP feature. Multiple-beam designs can be configured to cover the satellite's available field of view or only a limited portion of it.

The nominal parameters for multiple-beam antennas are derived as a means to examine design complexity versus RF performance. Generally, multiple-beam antennas use reflector antenna designs, and offset reflector configurations are typically used to avoid efficiency loss from feed cluster blockage. Multiple-beam antennas are configured by surrounding a central beam by a number of beams arranged with beam centers located on an equilateral triangular pattern. The minimum antenna gain level in the field of view for multiple-beam designs occurs at the point where three adjacent beam patterns overlap and is referred to as the "triple point." The equilateral triangular beam arrangement [7] results in a maximum value of this minimum gain level at the adjacent beam overlapping point compared to other beam arrangement configurations. The adjacent beams in this triangular arrangement,

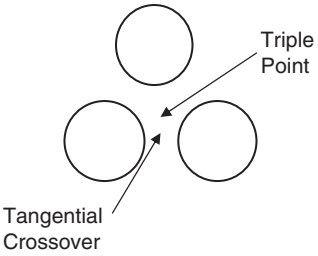


Figure 6-2 Tangential crossover and triple point locations

shown in Fig. 6-2, illustrate the minimum gain level between adjacent beams indicated by the triple point. The tangential crossover point between two adjacent beams is also indicated in this figure and the tangential crossover level corresponds to the pattern level where adjacent beam contours intersect halfway between the beam centers.

The multiple-beam pattern arrangement starts with a central beam aligned with the antenna's boresight axis. The central beam is surrounded by a ring of six additional beams; this in turn is surrounded by another ring having 12 beams, which in turn is surrounded by another ring of 18 beams, and so forth. The total number of beams, n, for a multiple-beam antenna having N rings of beams surrounding the central beam, equals

$$n = 1 + 3(N + 1)N$$

A typical antenna beam arrangement for a multiple-beam antenna has beams spaced so that the edge of coverage is tangential to the half-power points and the tangential crossover level between adjacent beams is 4.3 dB below the beam peak. The pattern level 4.3 dB lower than the beam peak corresponds to an angular separation from the beam peak equal to 0.5986 beamwidths using a Gaussian pattern representation for the antenna beams. With the equilateral separation between beam centers, the triple point where the three beam patterns intersect is 5.6 dB lower than the peak gain value and for the individual beams corresponds to the minimum gain level of the individual beams within the field of view. As an example, the beam pattern geometry for a 91-beam antenna is illustrated in Fig. 6-3.

Nominal characteristics for multiple-beam antennas can be derived by defining the required field of view and determining the number of beams needed to service that coverage area. The antenna beamwidth, θ_b, to service a FOV (field of view) having a specified angular extent can be determined from

$$\text{FOV} = [1 + 2N(2 \times 0.5986)]\,\theta_b$$

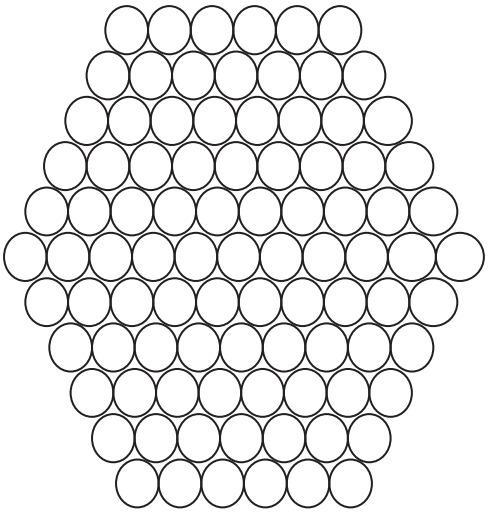


Figure 6-3 Beam arrangement for 91 beams

The required aperture diameter D can then be determined from the beamwidth of the individual beams.

As an example that provides full earth coverage from a geostationary satellite, the beamwidth of an individual beam is assumed to equal $65\lambda/D$. The peak gain level of an individual beam depends on the antenna efficiency, and this example assumes a 55% antenna efficiency. The minimum antenna gain level within the field of view occurs at the triple point between adjacent beams and is obtained by subtracting the 5.6 dB value from the peak antenna gain level. These nominal parameter assumptions were used to construct Table 6-1, which illustrates the gain levels and

TABLE 6-1 Nominal Multiple-Beam Antenna Parameters

Number of Rings	Number of Beams	Beamwidth, degrees	Diameter, Wavelengths	Peak Gain, dBi	Minimum Gain, dBi	Incremental Gain, dB
1	7	4.97	13.1	29.7	24.1	
2	19	2.92	22.3	34.3	28.7	4.6
3	37	2.06	31.5	37.3	31.8	3.0
4	61	1.60	40.7	39.5	34.0	2.2
5	91	1.30	50.0	41.3	35.8	1.8
6	127	1.10	59.2	42.8	37.2	1.5
7	169	0.95	68.4	44.0	38.5	1.3
8	217	0.84	77.6	45.1	39.6	1.1
9	271	0.75	86.8	46.1	40.6	1.0
10	331	0.68	96.1	47.0	41.4	0.9

beamwidth of a single beam within the collection. This table assumes the satellite is in a geostationary orbit that has a 16.9° earth field of view. The last column of the table indicates the incremental gain level relative to the values achieved when one less ring of feeds is used. The beams become more densely packed as the number of rings increases; however, the gain advantage afforded by an increased number of beams becomes less significant as the number of rings increases. Practical designs using multiple-beam antenna technology are limited by the required aperture size at low frequencies and by the tolerable system complexity of the supporting electronics as the number of beams increases.

The ability to independently use the multiple beams requires achieving isolation between beam positions to avoid mutual interference between independent data streams communicated on adjacent beams. This isolation is achieved by the antenna sidelobe response and by polarization. In practice, frequency and polarization reuse techniques are developed to provide the required isolation. The required isolation depends on the signal modulation formats used by the system and the dynamic range of the user signal levels. The tolerable co-channel interference level for a given modulation format is generally determined through measurements using the operational signal modems, one of which represents the desired signal and the other representing the interfering co-channel interference. The level of the interfering signal representation is increased until the SNIR (signal-to-noise plus interference) for a specified BER level is determined. The dynamic range of the uplink user signals also must be addressed in determining the required isolation between beams assigned to the same frequency subband. In a simple example of a frequency reuse plan [8], illustrated in Fig. 6-4, the available spectrum is divided into three subbands and the

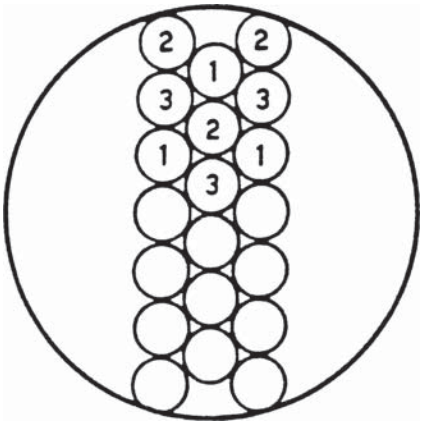


Figure 6-4 Example frequency reuse plan [8]

subbands are assigned to beam positions as indicated so that adjacent beam positions do not use the same subband. More commonly, a frequency reuse based on seven frequency subbands is used to increase the angular spacing between coverage areas using the same frequency subband. The increased angular spacing results in lower sidelobe levels in adjacent beam positions sharing the same frequency subband and increased isolation between beams as a result.

A principal design issue for multiple-beam antennas [7] is how to achieve individual beams with low sidelobe performance while maintaining high tangential beam crossover levels so the minimum antenna gain levels in the field of view are not reduced. Low sidelobe levels (as discussed in Chapter 1) require an aperture having a tapered amplitude distribution produced by the feed illumination. The tapered amplitude distribution requires more directive feed patterns, which in turn require a larger feed aperture. At the same time, the minimum gain within the field of view is also an important system parameter. The amplitude taper needed for low sidelobe beams requires relatively large feed dimensions to produce the desired illumination taper. However, the dimensions of these relatively large feeds also impose correspondingly large angular separations between adjacent beam positions. These large separations between adjacent beam positions result in tangential beam crossover levels that (relative to the antenna beam's peak gain level) are at a lower level than desired.

Two ways have been explored to address the conflict between high tangential beam crossover levels and low antenna sidelobe levels. One way [9] is to combine a given feed with small amounts of the adjacent six feeds with weighting circuitry to achieve low sidelobes and high crossover levels. The result of combining small values of adjacent beams with a central beam is to increase the feed aperture size electrically, producing an aperture amplitude taper. A second way [10] to achieve high tangential beam crossover levels and low antenna sidelobe levels uses a cluster of feeds that is displaced from the focal region to allow larger feed apertures and underilluminates the aperture to achieve low sidelobes. This approach requires a somewhat larger aperture than is normally used, but its potential in achieving high crossover levels and low sidelobes warrants further development.

Another different approach to achieve both low sidelobe antenna beams and high values of the minimum antenna gain level in the field of view configures the antenna design to have low sidelobe beams and a relatively low crossover level. The problem of obtaining a high value of the minimum antenna gain level in the field of view is addressed by combining the three adjacent beams to produce another beam whose axis is coincident with the triple point locations. A final approach, mentioned earlier and used by the ITALSAT design [6], employs two separate

apertures to form six independent beams, with each aperture forming three independent beams. The beams formed by each aperture have sufficient angular separation to provide the necessary aperture amplitude taper. The antenna pointing by each aperture allows alternative beams from each aperture to form an overall multiple-beam arrangement with closely spaced adjacent beams. This approach does not permit combining beams from the separated apertures because the phase center separation between the apertures would result in grating lobes. The separate aperture approach also increases the number of apertures needed to produce the composite multiple-beam arrangement but can be an effective solution for applications like the ITALSAT design.

Multiple-beam antenna technology must produce beams with high fidelity over the required field of view. Design attention must be paid to minimizing the scan loss over the field of view and maintaining sidelobe performance at the edges of coverage that is comparable with the central beams. Multiple-beam antennas for geosynchronous satellites that cover the entire available field of view must produce multiple beams over about an 18° field of view. Reflector antenna technology is generally considered a limited scan design and, with proper design attention, can maintain good pattern fidelity and minimal scan loss over the required field of view. Offset reflector technology is commonly used because its geometry avoids blockage effects of relatively large feed clusters used to produce the multiple beams. Long f/D values, dual reflector configurations, and shaping techniques [11, 12] are commonly used, and analysis codes are available to develop detailed designs to satisfy requirements for specific applications.

When multiple beams are used for lower altitude satellites, the required field of view expands greatly (as shown in Fig. 3-17) and exceeds the capability of limited scan reflector designs. At these lower altitudes, the required beamwidth to achieve footprint dimensions comparable to geosynchronous satellites is much larger than the geosynchronous case. These factors result in array antenna systems that are small enough to have a practical number of elements and can achieve the required field of view with acceptable scan degradation. Multiple-beam generation is accomplished by using multiple corporate feed structures, where each feed structure contains the means to steer the beam independently to the desired location. The Globalstar design [13] uses separate uplink and downlink arrays that produce 16 beams. The uplink array operates at L-band and forms 16 beams from 61 array elements and the downlink array operates at S-band and forms 16 beams from 91 array elements. The IRIDIUM array design [14] is configured in three array panels having active transmit/receive modules for each of the more than 100 array elements. Each array covers a 120° angular sector and the array panels are tilted away from the satellite's nadir axis to improve

coverage at wide angles towards the earth's horizon. Each array panel produces 16 beams so that the field of view for each satellite is serviced by 48 beam positions.

An example of multiple-beam operation within a spot size [15] has been applied to theater coverage applications. When multiple beams service a smaller spot size than the entire field of view, electrically large apertures are required. Such aperture sizes become practical at EHF frequencies where a reasonably compact antenna design results. This design is proposed for uplink operation at 44 GHz to provide communications for military applications. Very high antenna gain performance is achieved by the narrow beamwidth in this design, and active tracking of ground beacons (as has been demonstrated by the ITALSAT program [6]) is required to compensate for satellite attitude variations. In comparison with the gain levels for antenna system designs that form a single antenna beam, the gain levels are increased by the number of antenna beams servicing the spot size. For military applications, the narrow antenna beamwidths isolate interference sources to a very small portion of the coverage area since the sidelobes of other narrow beams removed from the beam position containing interference reduce received interference power levels. The system design also uses spread spectrum modulation for interference protection and the presence of interference that impacts performance is indicated by excessive received power levels in the beam positions receiving interference power. In such cases, the interference power that impacts desired signal communication performance must exceed the signal power level by roughly the spread spectrum processing gain. Adaptive interference cancellation can also be implemented in these designs. Since spread spectrum is also used, adaptive interference cancellation is required only when the spread spectrum interference cancellation protection is inadequate. In such cases, interference is identified by excessive power levels in uplink beam positions. As discussed in the next section, direction finding techniques using the multiple beams and adaptive interference cancellation based on measured interference signal directions can be used.

Example values for a 2^{o} coverage area [15] in Table 6-2 illustrate the beamwidth, gain levels, size, and interference suppression performance achieved as the number of multiple beams in the spot coverage increases. The interference performance assumes a pattern null is located on the interference source, and communication can be performed when users are within 10 dB of the beam peak. The angular separation between the pattern null and a pattern level 10 dB below the beam peak is 0.38 beamwidths. EHF systems require link margin for rain attenuation, and it is assumed that this signal margin is available when interference is present. The required separation between interference and

TABLE 6-2 Nominal Multiple-Beam Antenna Parameters for a 2° Coverage Area [15]

Number of Beams	Beamwidth, deg	Gain, dBi	Diameter, ft	Delta R, km	Delta R, km- 30°	Delta R, km- 20°
7	0.835	45.8	1.88	204.2	408.4	597
19	0.418	51.8	3.75	102.1	204.2	298.5
37	0.278	55.4	5.63	68.1	136.1	199
61	0.209	57.9	7.51	51	102.1	149.3
91	0.167	59.8	9.39	40.8	81.7	119.4
127	0.139	61.4	11.26	34	68.1	99.5
169	0.119	62.7	13.14	29.2	58.3	85.3
217	0.104	63.9	15.02	25.5	51	74.6
271	0.093	64.9	16.89	22.7	45.4	66.3
331	0.084	65.8	18.77	20.4	40.8	59.7

a desired user has been computed in the table as the "delta R" parameter. Separation values are indicated for the subsatellite point, and user elevation angles of 20^{o} and 30^{o} where the worst-case values are presented corresponding to the satellite, the center of the earth, the desired user, and the interference source are located on a common plane.

Multiple-beam designs, particularly when very narrow beamwidths are used, have issues in dealing with capacity allocations—and for military systems, dealing with interference. Capacity demands for communication services are not uniformly distributed over the collection of beams and can also be time-varying. This suggests a frequency reuse plan [16] that provides a basic capability with fixed subband allocations and another subband that can provide increased communication capabilities in beam positions with high-capacity demands. This additional subband can also be used in beam positions where interference is present. The change of frequency between the basic fixed subband to the additional subband may be adequate to reduce interference, and if not, adaptive interference techniques can be applied without impacting the basic frequency reuse plan.

A typical multiple-beam transponder design [9] (described in Fig. 6-5) uses multiple-beam antennas on the uplink and downlink frequencies. This system maps K uplink beams into N downlink beams, which is controlled by what is called a "message router." The message router can have a variety of forms. The simplest form is simply a hardwired fixed mapping of uplink-to-downlink beams. While such a design is very simple, connectivity between users in different coverage areas requires incorporating switching networks and frequency subbands into the message router design. Such techniques follow the linear frequency translating transponder architecture and have minimal weight and power impacts on the space segment designs.

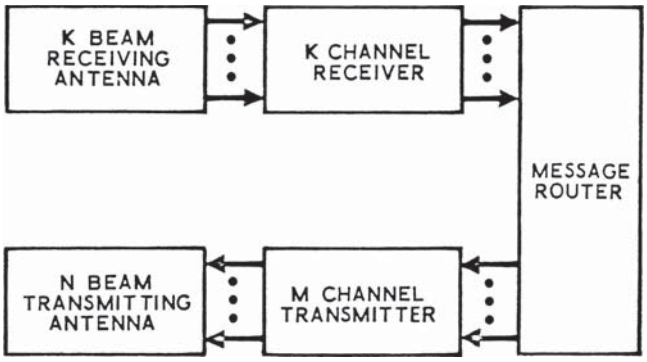


Figure 6-5 Multiple-beam transponder architecture [8] (© 1974 IEEE)

The message router has other forms that follow the regenerative repeater transponder architecture. In this case, the message router includes demodulation and remodulation capabilities that separate individual system users. After such uplink signal separation, the individual uplink signals user can be routed to the appropriate downlink beam positions. Regenerative repeater designs provide flexibility in user routing as well as reducing the interference power transmitted on the satellite downlink. The multiple-beam system can contain digital demultiplexing and multiplexing circuitry to separate the uplink user signals within the individual beam channels and combine and route them to form the downlink signal collection. This signal separation and combining is referred to as a channelizer architecture (discussed earlier in Chapter 3). Individual beams can also be assigned to subbands whose signal collections can be routed to downlink beams as a group, thus simplifying the channelizer design. The individual signal or frequency subbands can be combined from different beam positions to form larger downlink coverage areas than a single beam's coverage for broadcast purposes. In some cases, the message routing is performed on the satellite while other system designs locate the message router on the ground, where the separation, switching, and routing operations are performed. The uplink signal collection is transmitted to a gateway ground terminal where the appropriate processing is performed to form the downlink signal collection. The information is then transmitted to the satellite by the gateway terminal, and the satellite rebroadcasts the information to the user segments. Gateway architectures are practical for low data rate applications where the multiple-beam signal collections can be accommodated through the gateway links with reasonable data transfer resources.

The specific architecture for the multiple-beam systems depends largely on the application. Commercial systems are commonly configured to

provide fixed coverage in orbit. Thus, these systems tend to combine feed elements in a fixed manner to follow geopolitical boundaries. Channelization for different user segments is commonly performed and provides linking not only within a given coverage area but also between coverage areas. The need to use the existing spectral allocation to the fullest extent possible results in high demands on frequency and polarization reuse techniques. These factors are apparent in the existing designs for geosynchronous satellites. Military systems place different demands on their systems than commercial systems. Political situations can vary greatly during the satellite's lifetime, so that flexibility to reconfigure coverage areas is essential. Variations in coverage are also needed to cover naval fleet motions and service to tactical operations. Military users are concerned with the possibility of intentional interference, or jamming, as well as privacy, and spread spectrum modulation used with regenerative repeater transponders and adaptive interference cancellation provide a powerful means of reducing interference and minimizing the amount of downlink power wasted by transmitting interference.

As the examples in this discussion indicate, multiple-beam antenna systems can be used in a large variety of ways, and their design flexibility is the reason for the importance of multiple-beam antenna systems. Uplink receiving antennas use preamplifiers between the antenna ports and the beamformer circuity to establish the system noise temperature and avoid degrading system performance with beamforming loss. Beamforming can also be accomplished at IF frequencies with potential savings in weight and cost. When active components are used between the antenna ports and beamforming networks, design attention must be paid to the unit-to-unit amplitude and phase tracking characteristics of the active devices to maintain ideal beamforming performance. The low weight and insensitivity to loss for receive uplink beamformers allow significant flexibility in their operation. The downlink beamforming has less flexibility because of system loss in beamformers. Present practice is to distribute a group of transmitters to the downlink multiple-beam collection. Signal routing loss from the transmitters to the beam ports of the multiple-beam antennas degrades ERP performance. Ideally, each antenna port would have its own transmitter that can be powered as required to service particular coverage areas. In this way, the transmit beamforming can be accomplished prior to the final transmit amplifiers, and beamforming loss would not reduce downlink ERF performance.

6.4 Adaptive Uplink Antennas

Adaptive uplink antenna designs have been developed to reduce interference received by the satellite uplink antenna. System degradation from interference has long been a concern of military users. More recently,

adaptive techniques [17] to locate user signals, reduce unintentional interference in multiple-beam designs, and increase isolation between beams in frequency reuse techniques have been addressed for the commercial communications sector. Adaptive uplink antenna development [18] involves many disciplines, the antenna hardware, algorithm selection, software implementation, and system design simulation. Adaptive antennas are an example of an antenna system whose performance strongly depends on the system electronics that are integrated with the antenna. Test techniques for adaptive antenna designs are described in Chapter 8.

Adaptive multiple-beam designs [19] have been used to protect users in spot coverage areas, as illustrated in Fig. 6-6. In such applications, adaptive cancellation of interference within the coverage area protects users,

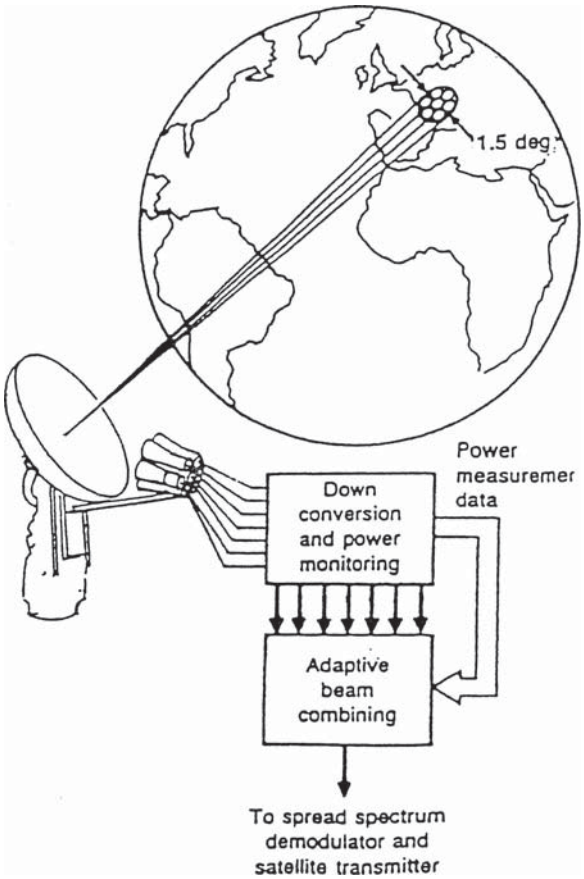


Figure 6-6 An adaptive uplink antenna [19] (© 1989 IEEE)

and sidelobe protection is provided from interference sources located beyond the coverage area. Narrow beamwidths produced by the multiple-beam antenna provide both resolution between interference and desired users and reduce the amount of the design coverage affected when a null is formed to cancel interference. The system design generally uses spread spectrum modulation formats so that adaptive interference cancellation operation is required only when adequate interference protection from spread spectrum processing gain is not provided. Interference is thus indicated when the power level in the beam exceeds a threshold level as measured by power monitoring and when the received signal components do not have the spread spectrum coding. Since such a design provides spot coverage, mechanical repositioning is used to move the coverage area as communication needs change. Beyond the coverage area, the system is protected from interference by the sidelobe response of the antenna, and unless very strong interference is very close to the design coverage, adaptive resources are not needed to reduce interference.

Adaptive antenna designs for uplink satellite applications cancel interference sources, but at the same time must provide communication services to users located within the coverage area. The system design must address four performance measures. The first measure is the steady state performance when interference is present. Adaptive cancellation produces nulls within the coverage area. One performance measure concerns the effectiveness of the interference cancellation. The second performance measure is the percent coverage area that is the portion of the original coverage area where existing and potential users can communicate after adaptive cancellation. The third performance measure is the minimum angular separation between interference sources and system users. The fourth performance measure is the time required for the adaptive weighting to converge to steady state values.

Multiple-beam antennas produced by reflector antenna technology are commonly used in such applications. Multiple-beam designs produce individual beams from a common aperture that has the important benefit of phase center coincidence for all beams. The phase center coincidence permits antenna beam combining without dispersion, which limits adaptive cancellation performance. Thinned array designs received consideration in the past [19, 20], but have several inherent disadvantages. In principle, the array antenna elements can be more widely separated than the diameter of the multiple-beam apertures. The wider element spacing has the potential of better resolution of interference sources. However, the element spacing results in dispersion as a consequence of the frequency scanning properties of arrays. Such dispersion limits the cancellation bandwidth of the design. When adaptive cancellation is implemented with array elements, the periodicity also produces additional nulls (referred to as grating nulls)

within the field of view, thus reducing the available coverage area. The thinned array design has less G/T than the multiple-beam approach and the array sidelobe levels outside the coverage area are higher than the multiple-beam design so that protection from interference outside of the coverage area is not provided.

A simulation program for adaptive cancellation operation is developed in conjunction with the system's hardware and software developments to demonstrate design compliance with requirements. The simulation program is initially used to project the system performance in terms of the four performance measures previously discussed. As the program proceeds, the simulation is updated by measured component performance and used to demonstrate design compliance with requirements. The simulation addresses the steady state and transient responses of the design and uses Monte Carlo performance evaluations of the interference scenario parameters to obtain statistical measures of communication performance. Since testing on a Monte Carlo basis is impractical, a limited number of cases are selected from the simulation results, and measurements from these cases are used to validate the simulation.

The minimum separation between interference sources and system users depends on the angular resolution of adaptive multiple-beam antennas. The angular resolution is a function of the slope of the antenna pattern between the desired signal and the interference that in turn depends on the antenna's electrical size. Consequently, the resolution between desired signals and interference increases as the beamwidth of the individual beams comprising the antenna decreases. The required separation between interference and system users (illustrated in Fig. 6-7) considers two different user elevation angles, 90° where the best angular resolution occurs, and 20° where the beam is spread by the earth curvature, and the worst-case value corresponding to the longest footprint dimension is indicated. The angular resolution also depends on the amount of gain performance that can be sacrificed in achieving the threshold SNIR. Two different values of signal loss, 3 and 10 dB, illustrate the differences in resolution performance that are achieved, and illustrate the importance of designing the antenna to achieve the maximum gain within the design coverage area. The angular resolution is calculated by aligning the angular separation of the pattern level and the pattern null between the main beam and first sidelobe. This simple analysis has been verified by more detailed simulations that have the additional capability of quantifying the percent coverage area where desired users can communicate. The important factors for practical designs are illustrated in this figure. A narrow beamwidth is selected to provide good resolution of interference subject to a system design with a practical number of beams. Antenna designs with the highest practical G/T are needed so that minimizing system losses,

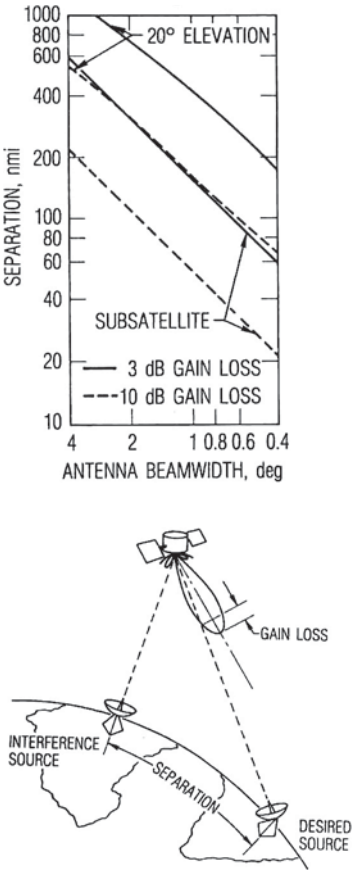


Figure 6-7 Multiple-beam adaptive antenna resolution [19] (© 1989 IEEE)

using low-noise preamplifiers, and controlling the noise contributions from other system electronics up to the point of beam combining are required to provide a system margin that can be sacrificed during adaptive operation. These system electronics must have adequate amplitude and phase tracking performance over the design bandwidth to meet the tolerance requirements (previously given in Fig. 5-7) to obtain effective cancellation.

A novel adaptive antenna design [21, 22] has been devised that does not require the usual adaptive weighting circuitry. The design uses a narrow antenna beamwidth that is repositioned to a lower gain level of the main beam in the direction of the interference. The basic idea of this design (illustrated in Fig. 6-8) is that the antenna beam is shifted away from the interference source. Beam repositioning reduces the desired

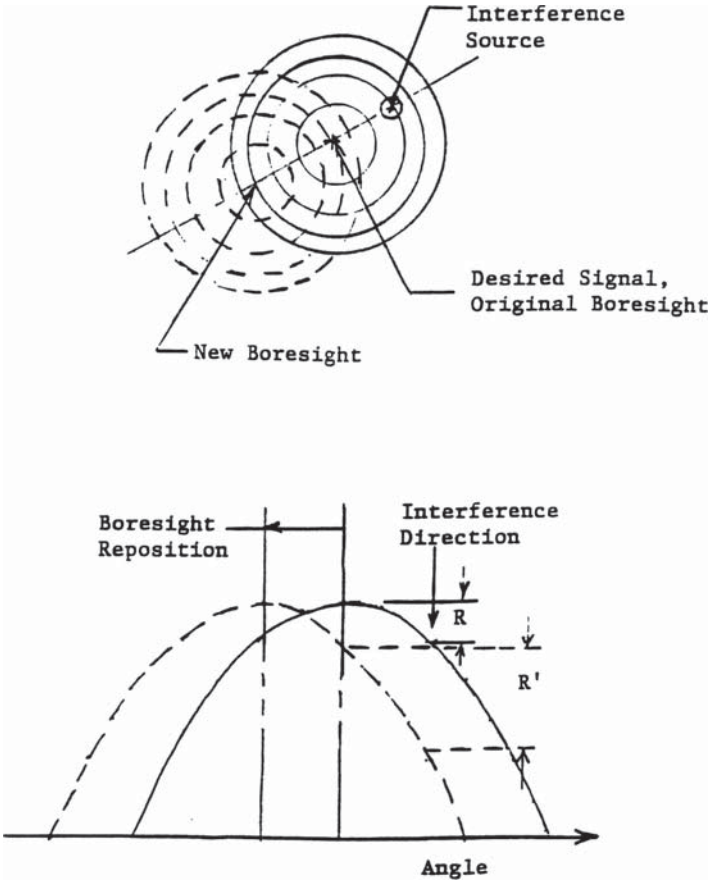


Figure 6-8 Contour and pattern description of adaptive beam respositioning [22] (© 1996 IEEE)

signal level, but because of the slope of the main beam pattern, the interference power is reduced by a greater amount. The design assumes sidelobe levels away from the main beam provide sufficient interference protection. Ideally, the beam is shifted sufficiently that the interference source is aligned with the pattern null between the main beam and the first sidelobe. The loss of the desired signal power as a function of the angular separation between the desired and interference signals (illustrated in Fig. 6-9) assumes the pattern null is aligned with the interference source.

The beam repositioning is implemented by correlation techniques using the circuitry shown in Fig. 6-10. The narrow beamwidth antenna is assumed to use closed-loop antenna tracking to maintain alignment with the desired signal by compensating for satellite attitude variations

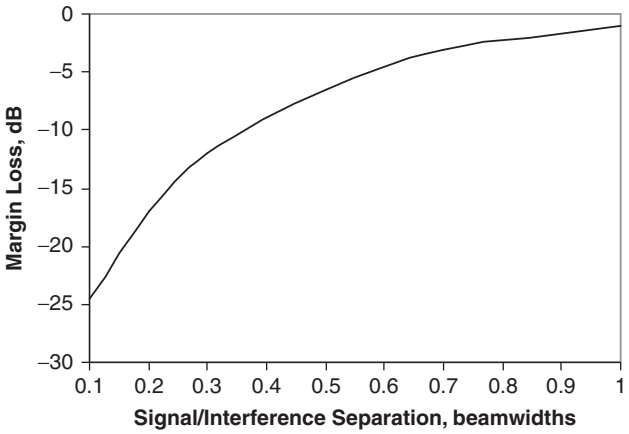


Figure 6-9 Margin loss versus signal/interference separation [1, 22] (© 1996 IEEE)

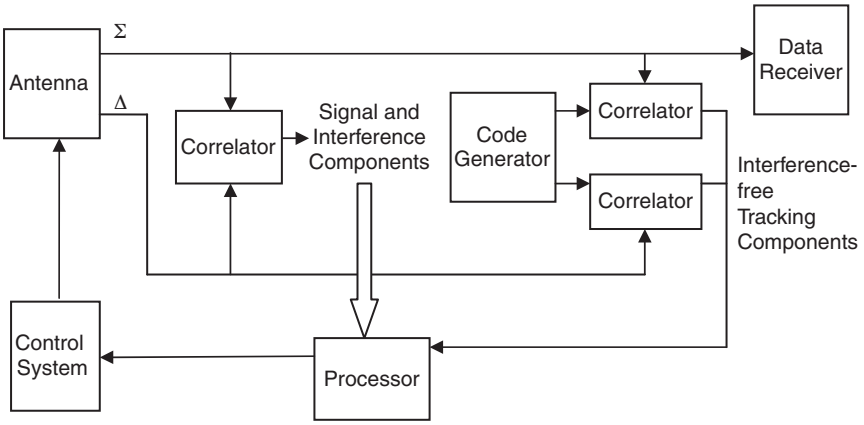


Figure 6-10 Circuitry for adaptive beam respositioning [1, 22] (© 1996 IEEE)

in interference-free conditions. The uplink signal is also assumed to contain a low-level pseudorandom code for operation of the tracking system. Correlation of the sum and difference with the pseudorandom code is used to track the desired signal. When interference is present, the normal error response indicates a response comprised of both desired and interference signal components. This error response is used to steer the antenna away from the interference source. During the beam steering, the error response increases because of the movement of the desired signal away from the antenna's boresight axis and

decreases as the interference is aligned with lower gain portions of the main beam. Since the desired signal's error response is measured by the error signal correlation with the pseudorandom code, the minimization of the interference can be determined by compensating the uncorrelated error response for the increased error response of the desired signal components.

This beam repositioning technique has also been extended [23] to operation with multiple-beam antennas. In this case, power monitors are used to identify beams having excessive received power levels that result from interference power. The three beams having the highest levels of excessive power are processed to identify the location of the interference source. Since the beams are arranged on an equilateral triangular basis, the direction finding techniques are based on subtracting pairs of the three beams to identify the interfering signal's location. Once the location is identified, a null can be commanded to that direction using a priori information on the antenna's beam response. Like the beam repositioning technique, the user signals contain a pseudorandom code and correlation techniques are used to monitor the reception of desired signal components.

Another simple adaptive system design [24] applies to relatively wide spot beam antenna designs and is based on "punching a hole" in the spot beam coverage. In operation, the spot coverage antenna beam and a much narrower antenna beam are amplitude weighted and subtracted, as indicated in Fig. 6-11. The spot coverage antenna is aligned with the interference source using tracking techniques. The design endeavors to locate the phase centers of both antennas close to one another and time delay compensation is used to minimize the location differences. Alternatives exist in selecting the subtracted level of the narrow beam antenna. If the gain levels of the two antennas are identical when subtracted, a relatively broad pattern null results, as indicated in Fig. 6-12. If the subtracted gain level of the narrower beamwidth antenna has a

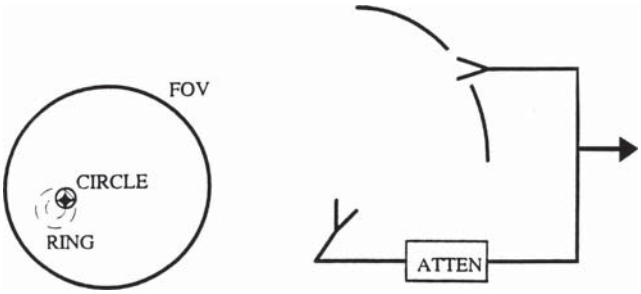


Figure 6-11 Adaptive nulling within spot beams [24] (© 1996 IEEE)

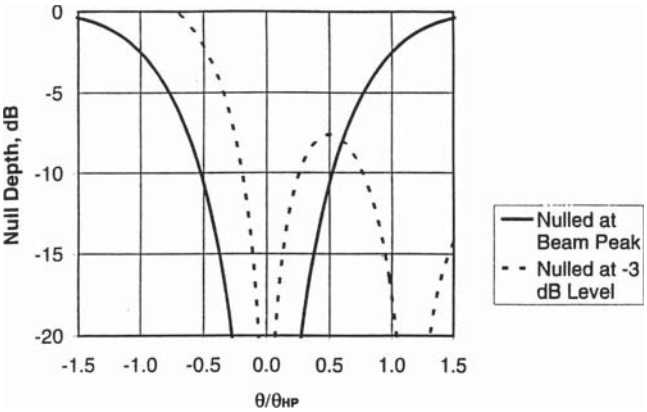


Figure 6-12 Resolution performance of spot null types [24] (© 1996 IEEE)

higher value than the gain of the spot coverage antenna in the direction of the interference, a ring null is formed, as shown in Fig. 6-12. When the subtracted gain level of the narrower beamwidth equals the gain level of the spot beam, the null is relatively broad. The pattern level of the narrower beamwidth antenna's main beam has little gain variation near the antenna's boresight axis, so the subtraction is effective over a range of angles. By contrast, when the subtracted gain value of the narrower beamwidth antenna is higher than the spot beam antenna's gain level in the direction of the interference, the subtraction is effective over a smaller angular width so that resolution of the ring null is better than broader pattern null when the subtracted gain levels of the two antennas are equal. This resolution is beneficial when desired signals are near the interference source. However, the ring null extends to other portions of the coverage area. The choice of the subtraction level depends on the specific interference situation. If the interference source is near the edge of coverage, for example, the alignment of the narrow beamwidth antenna and a subtraction level higher than the spot coverage antenna might be used to form a ring null that extends beyond the coverage area.

6.5 Active Aperture Antennas

While multiple-beam array designs have been previously discussed, another array design for downlink service is the active aperture design illustrated in Fig. 6-13. In operation, active aperture design would sequentially service low data rate users that are sparsely but widely distributed over the earth field of view. The beam can be rapidly

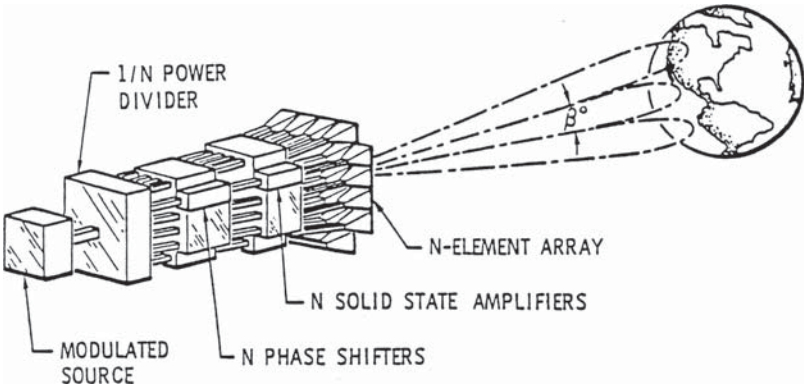


Figure 6-13 Active aperture antenna

repositioned to different user locations with the speed of electronically changing the phase shifters used to steer the beam. The individual array elements are waveguide radiators that are contiguous to achieve the maximum possible array element gain levels. Since only one beam is radiated at any given time, the sidelobe requirements are minimal, and thus a uniform amplitude distribution can be used to achieve the minimum beamwidth for the aperture size. Uniform amplitude also permits each amplifier to be driven to the same near-saturation operating conditions for maximum power efficiency, thus achieving the highest ERP level.

The first-order design proceeds by selecting a beamwidth for coverage purposes, dividing the aperture into a sufficient number of elements so that with the selected element transmitter power output, the desired ERP is achieved. The selection of the element size also needs to avoid grating lobes within the earth field of view and to achieve the ERP requirements at the edges of the field of view. A maximum element spacing of 3.4 wavelengths is required to keep the grating lobes away from the earth's field of view for geosynchronous orbits. The peak ERP level to first order is given by

$$\text{ERP} = \eta N^2 P_e G_e$$

where η is the antenna efficiency, N is the number of elements, P_e is the transmitted power per element, and G_e is the element gain in the array environment. The antenna efficiency, as used here, is really the array efficiency, and includes the effect of amplitude and phase tracking and the directivity loss caused by the grating lobes. Other efficiency loss includes ohmic and mismatch loss that is included in

the antenna element gain. Similarly, the element pattern rolloff has not been included; a 1.8 wavelength element spacing yields a 1 dB pattern rolloff at the edge of the earth. Design attention must be paid to the system's gain distribution. The solid state devices have both limited power efficiency and gain. The gain distribution through the antenna needs to be optimized to minimize the prime power consumption. Attention must also be given to the thermal control design to maintain the design operating temperature and provide sufficient temperature uniformity to maintain amplitude and phase tracking performance of the active devices. A means of calibration must also be devised to determine the element amplitude and phase tracking performance in orbit, so that compensation can be provided by adjusting the phase shifters.

This concept has many attractive advantages. The ERP level can be increased by increasing the number of elements and/or the power per element that may be accomplished by combining active devices. In this way, the ERP to a given size coverage area can be increased without the development of a higher-power transmitter needed in antenna designs with a single transmitter. The beam can be steered to an arbitrary location so that the peak ERP can be centered on the desired community of users. The patterns in Fig. 6-14 use an array of 64 elements arranged in an 8×8 configuration. The first pattern illustrates a diagonal plane response when the beam is scanned 5° off-axis. In addition, the size of the coverage area can be increased by rephasing the aperture distribution. This capability is also illustrated in Fig. 6-14, where again the beam is scanned 5° off-axis in a diagonal plane and the beam is broadened by a factor of three. A beam broadening technique is often accomplished by adding quadrature phase error to the aperture distribution. In this case, the quadratic phase error simply raised the sidelobe without appreciable beam broadening. The pattern results were achieved by taking every other element to form three subarrays. The central subarray was steered to the desired scan angle and the other subarrays were phased to steer their beams to opposite sides of the central beam, and the patterns were added together to produce the broadened pattern. Since the subarray elements are more widely spaced than the adjacent array elements, a grating lobe is apparent in Fig. 6-14. Finally, the active aperture design can be viewed as a spatial combining technique. Combining circuitry that forms a single power output port is well known [26, 27], but the spatial combining technique avoids the loss in the output combiner, and the high isolation between array elements results in little effect on the combining efficiency of the remaining devices when an active array device fails.

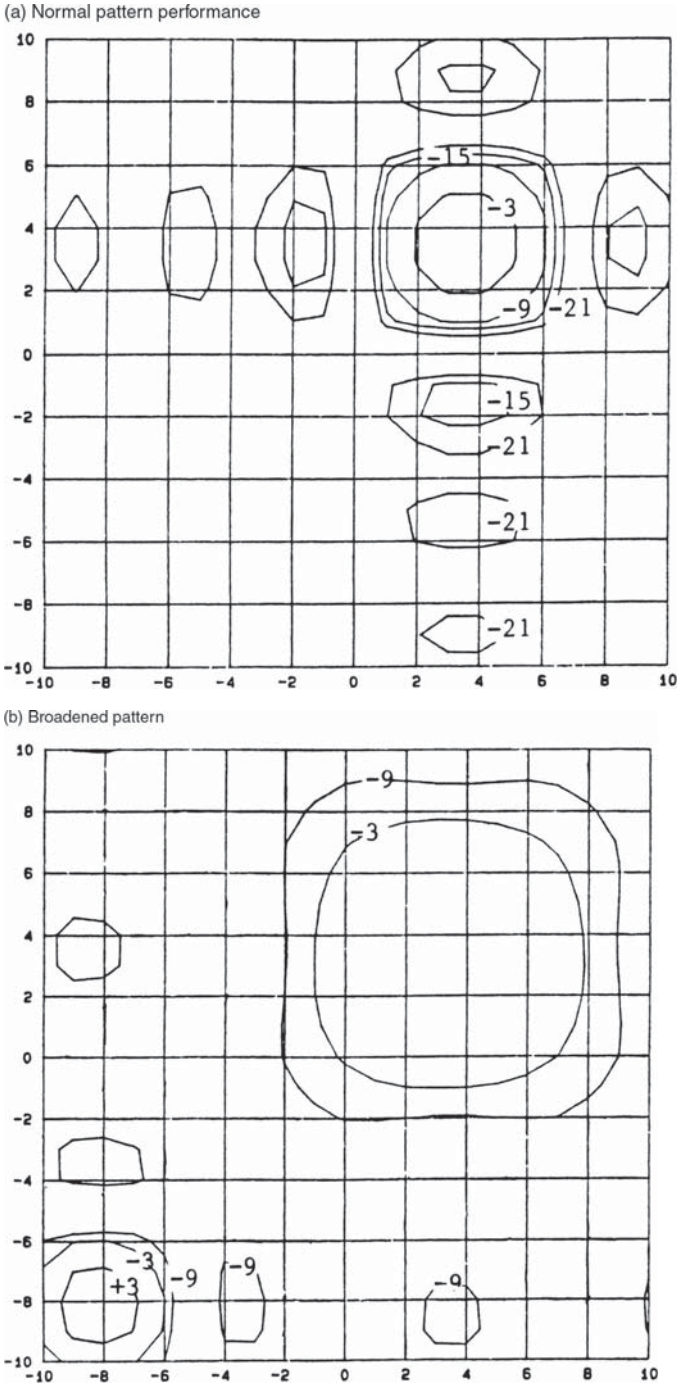


Figure 6-14 Normal and broadened diagonal plane patterns [25] (© 1989 IEEE)

6.6 Point-to-Point Antennas

Antennas for point-to-point communication links are required for crosslink subsystems and for gateway applications having high data rate requirements. An example antenna system design for a crosslink subsystem for geostationary satellites is described. The overall transponder architecture, antenna beam steering requirements, and the need for four frequency subbands were explained previously in Chapter 3. The beam steering requirements are principally in the azimuth plane, with modest requirements for elevation scanning to accommodate satellite stationkeeping variations. This antenna design evolved to satisfy several development objectives:

1. Provide a low loss means of antenna beam steering while avoiding beam waveguide and rotary joint complexities.
2. Minimize the mass needed to be moved for beam steering.
3. Allow a fixed location for the feed and RF electronics to provide isolation from the external environment, access to the spacecraft's thermal control system, and low loss to the RF electronics.
4. Achieve low insertion loss in the receive path to capitalize on the low noise background temperature to achieve a low system noise temperature.
5. Provide very high polarization purity to allow the use of orthogonal polarizations to limit required bandwidth for high data rate applications.
6. Satisfy the requirements to select two of the four frequency bands required in the crosslink operation, one for transmitting and the second for receiving.

The antenna system concept shown in Fig. 6-15 is an offset reflector design having a feed and geometry that requires moving only the main reflector to meet the azimuth and elevation scanning requirements previously described in Fig. 3-13. The wide azimuth plane scanning requirement is satisfied by rotating the offset reflector about the axis containing the reflector center and the reflector focus imaged by the subreflector. The optics and scan motion function like a periscope to achieve the wide azimuth scanning requirement. The elevation scan is provided in the axis orthogonal to the azimuth scan axis to compensate for satellite attitude variations and satellite stationkeeping variations. The long effective focal length from the reflector further magnified by the subreflector permits the required elevation scan capability. This approach permits the required beam steering by only positioning the main reflector and the rest of the system components, including the

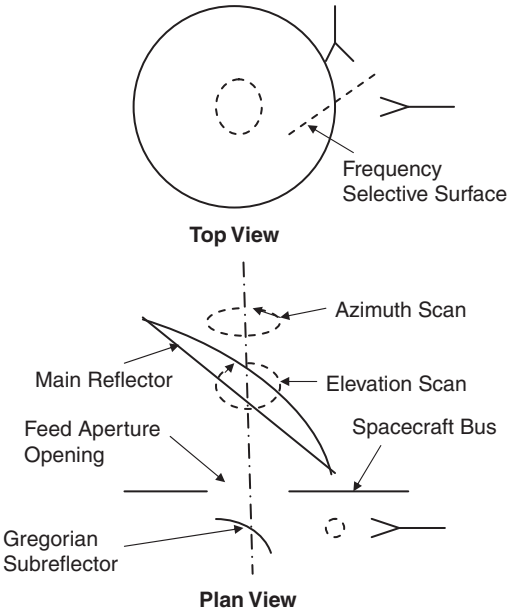


Figure 6-15 The crosslink antenna concept

Gregorian subreflector, and the feed assembly are fixed to, and enclosed by, the spacecraft. This approach reduces the required mass to be positioned for beamsteering and allows the electronics to be integrated with the spacecraft thermal control system and to be shielded from the external space environment.

The feed system and the subreflector use an opening in the spacecraft panel to illuminate the main reflector. This optics arrangement minimizes the mass that is repositioned to provide the necessary beamsteering. Further, the feed system is fixed within the spacecraft to minimize the loss to the RF electronics, isolate the components from the external environment, and provide an interface to the bus thermal control. The feed system in Fig. 6-15 is comprised of two feed elements separated by a frequency selective surface that reflects two of the feeds and transmits the other two feeds. Each feed services two of the four frequency bands required in this application. This arrangement permits feed optimization for operation over the bandwidth encompassed by two of the bands rather than four bands. The frequency selective surface also provides additional isolation between the frequency bands. The polarization purity is provided by the Gregorian subreflector, but design attention in the frequency selective surface development is required to match both the amplitude and phase of the reflection and transmission properties of the surface for orthogonal polarization components over the range

of incident angles of the feed illumination. The matched polarization responses of the surface are necessary to avoid degrading the axial ratio performance of the signals reflected from, or transmitted through, the frequency selective surface. The feed system is anticipated to use either dual mode or corrugated horn technology, and a higher-order waveguide mode needs to be generated for the system's autotrack function. Both feed technologies are capable of providing low sidelobe rotationally symmetric illumination patterns that are required to achieve high polarization purity. This design attention to the feed system, the long reflector focal length, and the polarization correction of the Gregorian subreflector design results in the capability to achieve the axial ratio performance needed to isolate the orthogonal polarizations. The selection of a large f/D ratio further magnified by the subreflector minimizes the beamscanning loss.

References

1. R. B. Dybdal, "Satellite Antennas," Chap. 44 in J. L. Volakis (ed), *Antenna Engineering Handbook* (New York: McGraw-Hill, 2007).
2. D. H. Martin, P. R. Anderson, and L. Bartamian, *Communication Satellites*, Fifth Edition (El Segundo CA: The Aerospace Press, 2007).
3. K. Ueno, T. Itanami, I. Kumazama, and I. Ohtomo, "Design and Characteristics of a Multiband Communication Satellite Antenna System," *IEEE Trans Aerospace and Electronic Systems*, vol. AES-31 (April 1995): 600–607.
4. A. R. Cherrett, S. W. Lee, and R. J. Acosta, "A Method for Producing a Shaped Contour Radiation Pattern Using a Single Shaped Reflector and a Single Feed," *IEEE Trans Antennas and Propagation*, vol. 37 (June 1989): 698–706.
5. O. M. Bucci, G. D'Elia, G. Mazzarello, and G. Panariello, "Antenna Pattern Synthesis: A New General Approach," *Proc IEEE*, vol. 82 (March 1994): 358–371.
6. F. Carducci and M. Francesi, "The ITALSAT Satellite System," *International Jour Satellite Communications*, vol. 13 (January–February 1995): 49–81.
7. R. J. Mailloux, *Phased Array Antenna Handbook* (Norwood MA: Artech, 2005).
8. R. B. Dybdal, "Multiple Beam Communication Satellite Antenna Systems," *IEEE ICC Conference Digest* (June 1974).
9. K. S. Rao, G. A. Morin, M. Q. Tang, S. Richard, and K. K. Chan, "Development of a 45 GHz Multiple Beam Antenna for Military Satellite Communications," *IEEE Trans Antennas and Propagation*, vol. AP-43 (October 1995): 1036–1047.
10. P. Ingerson and C. A. Chen, "The Use of Non-focusing Aperture for Multibeam Antennas," *1983 IEEE AP-S Symposium Digest* (May 1983).
11. R. Jorgensen, P. Balling, and W. J. English, "Dual Offset Reflector Multibeam Antenna for International Communications Satellite Applications," *IEEE Trans Antennas and Propagation*, vol. AP-33 (December 1985): 1304–1312.
12. C. M. Rappaport and W. P. Craig, "High Aperture Efficiency Symmetric Reflector Antennas with up to 60° Field of View," *IEEE Trans Antennas and Propagation*, vol. 39 (March 1991): 336–344.
13. F. J. Dietrich, P. Metzen, and P. Monte, "The Globalstar Cellular Satellite System," *IEEE Trans Antennas and Propagation*, vol. AP-46 (June 1998): 933–942.
14. J. J. Schuss, J. Upton, B. Meyers, T. Sikina, A. Rohwer, P. Makridakas, R. Francois, L. Wardle, and R. Smith, "The IRIDIUM Main Mission Antenna Concept," *IEEE Trans Antennas and Propagation*, vol. AP-47 (March 1999): 416–424.
15. R. B. Dybdal, S. J. Curry, and M. A. King, "An Uplink Multiple Beam Concept for Theater Coverage," *2002 IEEE MILCOM Symposium Digest* (October 2002).

16. R. B. Dybdal, "Adaptive Control of Multiple Beam Transponders," *1997 IEEE MILCOM Symposium Digest* (November 1997); see also R. B. Dybdal, "Adaptive Control of Multiple Beam Communication Transponders" (April 25, 2000): U.S. Patent 6,055,431.
17. T. Gebauer and H. G. Gockler, "Channel Individual Adaptive Beamforming for Mobile Satellite Communications," *IEEE Jour Selected Areas on Communication*, vol. SAC-11 (February 1995): 439–448.
18. R. B. Dybdal, D. J. Hinshilwood, and K. M. SooHoo, "Development Considerations in the Design and Simulation of Adaptive MBAs for Satellite Communications," *1993 IEEE MILCOM Symposium Digest* (October 1993).
19. K. M. SooHoo and R. B. Dybdal, "Resolution Performance of an Adaptive Multiple Beam Antenna," *IEEE MILCOM '89 Symposium Digest* (October 1989).
20. J. T. Mayhan, "Area Coverage Adaptive Nulling from Geosynchronous Satellites: Phased Arrays Versus Multiple Beam Antennas," *IEEE Trans Antennas and Propagation*, vol. AP-27 (March 1986): 410–419.
21. R. B. Dybdal and H. J. Wintroub, "An Antenna Based High Data Rate Concept," *1995 IEEE Symposium Digest* (November 1995).
22. R. B. Dybdal and S. J. Curry, "Adaptive Beam Pointing," *1996 IEEE MILCOM Symposium Digest* (October 1996); see also R. B. Dybdal and S. J. Curry, "Adaptive Receiving Antenna for Beam Repositioning" (April 14, 1998), U.S. Patent 5,739,788.
23. R. B. Dybdal and S. J. Curry, "An Uplink Antenna for Electronic Beam Steering and Interference Reduction," *2002 IEEE AP-S Symposium Digest* (June 2002).
24. G. M. Shaw and R. B. Dybdal, "A Simple Nulling Antenna for Satellite Systems," *1996 IEEE AP-S Symposium Digest* (July 1996).
25. G. M. Shaw and R. B. Dybdal, "Beam Broadening for Active Aperture Antennas," *1989 IEEE AP-S Symposium Digest* (June 1989).
26. K. Chang and C. Sun, "Millimeter Wave Power Combining Techniques," *IEEE Trans Microwave Theory and Techniques*, vol. MTT-31 (February 1983): 91–107.
27. Z. Galani, J. L. Lampen, and S. J. Temple, "Single Frequency Analysis of Radial and Planar Amplifier Combiner Circuits," *IEEE Trans Microwave Theory and Techniques*, vol. MTT-29 (July 1981): 642–654.

Chapter

7

User Segment Antennas

7.1 Overview

User segment antennas, like space segment antennas, have unique requirements for communication satellite applications. While space segment antennas maintain coverage over a specified field of view, user segment antennas must point and track the satellite in orbit. The space segment antennas must service the same coverage areas for both the uplink and downlink frequencies. The differences in the uplink and downlink frequencies often result in using different space segment antennas for the uplink and downlink services. The uplink and downlink space segment antennas have different sizes as required to achieve the same beamwidth values at both frequencies needed to preserve the same angular coverage characteristics. User segment antennas do not have coverage requirements at both uplink and downlink frequencies, and consequently, the different antenna beamwidth values at the two frequencies do not present a problem. Since the same aperture is used for both uplink and downlink frequencies, design complexity and cost are reduced. Space segment antennas point at a warm earth background, whereas the user segment antennas point at the cold sky background. Consequently, user antennas require particular design attention to RF losses to capitalize on low noise receiver technology to achieve a low system noise temperature to enhance the user system's G/T performance (as discussed in Chapter 1). User segment antennas are susceptible to interference from other nearby terrestrial services, particularly at the lower microwave frequencies, and their sidelobe envelope requirements and control techniques will have increased importance in future years. User systems for multiple frequency operation are commonplace, allowing the same antenna to operate with

different satellite programs, a trend that will increase in the future. The number of user terminals continues to increase with demands for satellite service, a trend that will continue in the future as communications continue to be extended to individual users.

Future development of user segment antennas [1] can logically be anticipated in several areas. Low sidelobe antenna designs will be required to provide increased protection from interference. Packaging techniques to satisfy user mobility requirements are needed to satisfy user mission requirements. User terminal designs capable of sequentially communicating with different satellite services will require further development of multifrequency capabilities. As satellite services extend to individual users in cellular networks, much of the terrestrial cellular user antenna development can be anticipated to apply to satellite communication services. Cost-effective designs for user terminals will be necessary to control overall system costs and, as discussed in Chapter 9, efficient techniques for production testing will be required as well.

User segment designs initially required large dedicated ground terminals for link closure purposes. Such terminals served as a hub, and communications to individual users were distributed by terrestrial means. Today, large terminals are still required for high data rate transfer in applications such as gateways and satellite uploads. Typically, such terminal designs use shaped reflector techniques (described in Chapter 2) to increase antenna efficiency. In addition, many systems are designed for use with VSAT (very small aperture terminals) applications that have much more modest performance requirements than earlier designs because of increased satellite performance capabilities. Direct broadcast television services are an excellent example of where initial services requiring about an 8-ft user antenna have transitioned to much more compact antennas. Further, these terminals are surprisingly affordable as a result of their production volume. Future satellite designs are being developed for handheld personal equipment. The trend for more diverse collections of terminal technology will continue in the future, and system sizing must accommodate a multitude of small individual users and their cost constraints.

7.2 User Antenna Technology

The requirements for user antennas favor reflector antenna technology, and indeed, in the public's mind, large reflector antennas are commonly associated with satellite communications. Reflector antennas are cost-effective, have a broad bandwidth capability, and impose minimal power consumption and design complexity. Other antenna technologies are selected only when compelling system requirements exist (e.g., conformal

array designs needed to satisfy aerodynamic constraints for high-performance aircraft). Such array designs have a significant cost, are limited in bandwidth, inherently have high complexity, and require prime power for active elements and beamsteering requirements.

User antenna size is constrained for reasons that differ from the space segment. A compact antenna size is less expensive than a larger antenna. Large antennas are susceptible to antenna pointing errors from wind loading, both in terms of their narrow beamwidth and the torque resulting from wind loading for a large aperture. Antenna pointing stability under wind conditions requires a stiff reflector design and positioner drive power to offset the wind-induced torques. Methodologies to address wind loading effects on antenna pointing performance [2] commonly assess steady state and gusting wind conditions. Alternatively, a protective radome is commonly used to isolate the antenna from wind perturbations. Operation with or without a radome involves economic tradeoffs involving not only satisfying operation during high wind levels but also providing protection from weather deterioration of the equipment enclosed within the radome.

User segment antenna designs endeavor to achieve high antenna efficiency to minimize the required antenna size, and for high gain antennas using dual reflector configurations, reflector shaping techniques have been quite successful. Reflector shaping techniques and high-efficiency feeds (such as corrugated horns having low sidelobe rotationally symmetric patterns) are key reasons that such designs achieve efficiency values in excess of 75%, as discussed in Chapter 2. The shaping of the main reflector and the subreflectors results in a more uniform aperture distribution that offsets the amplitude taper loss to achieve increased efficiency. Smaller antennas for VSAT applications commonly use offset reflector technology that increases antenna efficiency by eliminating feed blockage.

User antenna designs for polarization reuse satellite systems must achieve the necessary polarization purity to avoid co-channel interference from, and to, other users. A typical requirement is to reduce the cross-polarized fields to 27 dB below the principal polarization over the field of view subtended by the tracking uncertainty. This design attention includes feed designs with low axial ratio capabilities for circular polarization. Such designs require rotationally symmetric feed patterns to maintain polarization purity over the radiating aperture plane. When offset reflector antenna technology is used, correction techniques (discussed in Chapter 2) provide the means to produce designs with high polarization purity.

User antennas for commercial C- and Ku-band frequencies often require operation at both bands so that multiple frequency feed designs are required. Frequency selective surfaces that allow transmission

through the surface at one frequency band and reflection from the surface at another frequency band permit the separation of frequency bands. The feeds are physically separated, giving the design enough freedom to optimize antenna feed performance for each band. Commonly, the lower frequency is in a prime focus configuration and a frequency selective subreflector is used to provide a Cassegrain configuration at the high frequency. The overall subreflector size is minimized by this approach, and for transmitting applications, the high frequency band has a short waveguide run, whereas the longer waveguide run to the prime focus location has reduced loss/length compared with the high frequency. The overall aperture size must be sufficient to satisfy G/T and ERP requirements at both bands.

User antennas for future satellite-based cellular communications have relatively low data rate capabilities and present future design challenges and development needs. Of necessity, such antennas need to be compact for practical designs. In addition, their pattern coverage must be broad to avoid requirements for precise pointing, which would represent an additional burden to the users. The broad pattern coverage resulting from these requirements together with typical operating environments results in situations where multipath levels generated by reflections from the surrounding environment significantly degrade system performance. These concerns are expressed [3] for CDMA systems where multipath can reduce the isolation between users and consequently degrade the capacity of the system. The reference also compares the capacity differences between CDMA and FDMA/TDMA multiple access techniques in multipath environments. Further research is required and suitable antenna designs must be developed for these applications that evolve from terrestrial cellular designs. Additional understanding and quantifications of the propagation limitations of such systems are required in the development and operation of these systems.

Adaptive techniques are commonly proposed as a means to cope with the multipath environments when broad coverage user antennas are required. Rake receivers that provide equalization and diversity combining [4] can protect receivers from multipath fading. The time-delayed multipath signals are equalized by an "adaptive transversal filter," which is a series of tapped delay lines with adaptive amplitude and phase weighting. Such circuitry aligns the direct and delayed multipath signals in time. The delay time spread is bounded by the autocorrelation function of the signal. Diversity combining is also incorporated into the Rake design to maximize signal power. More recently, adaptive concepts have been applied [5] to transmitting antennas. In this case, reflected components of the transmitted signals from terrain features are sampled. These reflected components are also multipath sources. By transmitting the signal through adaptive

transversal equalizers into the antenna elements and adjusting the circuitry to minimize the reflected components, the multipath can be controlled in the transmit case. An LMS algorithm constrained by the pointing direction to the satellite is envisioned. A compact antenna design is required in both cases; a multimode aperture adaptively combined is one potential technology.

7.3 Antenna Sidelobe Control

Antenna sidelobe control techniques are advantageous in reducing interference both to and from other terrestrial systems. In operation, the main beam of the user systems is pointing at the satellite and because of the range separation between the user and space segments, the received signal levels are low and the transmitted signal levels are high from the user terminals. As a consequence, interference to and from a user terminal for satellite communications from terrestrial sources becomes an issue, and spectral control for the adjacent frequency allocations requires attention. These concerns are particularly apparent at the lower microwave frequencies, and this problem will become more troublesome as the use of microwave systems continues to increase.

User antennas, particularly in the congested C- and Ku-band frequencies, have requirements for sidelobe control to avoid interference with adjacent satellites in the crowded geosynchronous orbit. These sidelobe control requirements specify an envelope for the sidelobe levels, and as years have passed, the envelope requirements have become more stringent, indicating greater awareness of potential interference. Sidelobe envelope requirements have spread to designs for user antennas over a broad range of frequencies and communication applications. Compliance with such requirements is generally a feature of vendor offerings. A typical example [6] is the CCIR recommendations (465-3) for sidelobe envelopes that apply to ground terminal antennas operating from 2 to 30 GHz and follow

$$\begin{aligned} G &= 32 - 25 \log\theta, \text{ dBi}, \ 1^\circ < \theta < 48^\circ \\ &= -10 \text{ dBi}, \ 48^\circ < \theta < 180^\circ \\ &\text{for } D/\lambda > 100 \end{aligned}$$

and

$$\begin{aligned} G &= 52 - 10 \log(D/\lambda) - 25 \log\theta, \text{ dBi}, \ (100\ \lambda/D)^\circ < \theta < 48^\circ \\ &= 10 - 10 \log(D/\lambda), \text{ dBi for } 48^\circ < \theta < 180^\circ \\ &\text{for } D/\lambda \leq 100 \end{aligned}$$

In addition, the CCIR recommendation 524-2 addresses the ERP spectral density, and the more detailed report 1001 provides greater data and explanation. Compliance with these sidelobe envelope requirements is commonly specified during user antenna procurements.

While vendor designs are commonly available that are compliant with such antenna sidelobe envelope requirements, design techniques exist to develop antennas having much lower wide angle sidelobe levels. The control of sidelobe levels near the antenna's beam requires providing an aperture distribution with significant amplitude taper and a planar aperture distribution, as discussed in Chapter 1. The wide angle sidelobe response of typical reflector antennas is comprised of the coherent sum of the aperture distribution, the feed response exclusive of that portion illuminating the reflector, blockage effects from the feed and struts, reflector edge diffraction, and spillover. Three examples of design approaches will be used to illustrate how these radiation mechanisms can be controlled to reduce sidelobe levels well below the levels of conventional reflector antenna technology.

An early effort to control antenna sidelobes for satellite communications [7, 8] uses an offset reflector antenna, as illustrated in Fig. 7-1. The antenna feed in this design is enclosed by shrouds to reduce feed radiation contributions to the antenna pattern while also avoiding feed blockage impact to antenna sidelobes. This geometry did not address feed spillover lobes that are present in the pattern, but later development [9] added serrated "blinders" to the design to reduce such contributions. The original rationale for this early development was to reduce antenna noise temperature levels by suppressing wide angle sidelobe levels that would couple ground emission components into the antenna, thereby increasing the performance of user segment antenna designs. This design instead has found terrestrial communication applications in municipal telephones exchanges, where they are a familiar sight. Municipal telephone exchanges have several links converging to a common location, and the reduced sidelobe levels provide isolation between such links.

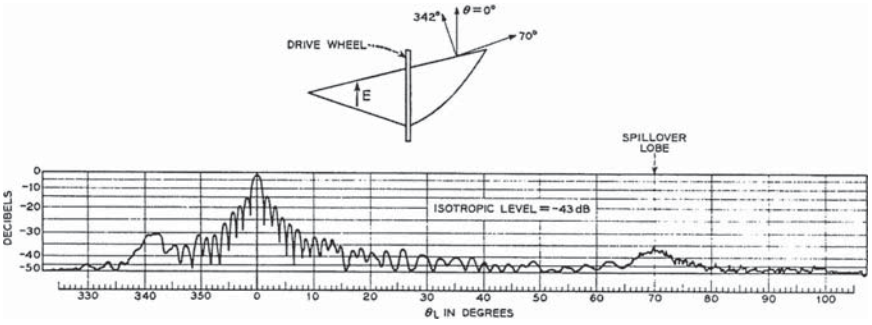
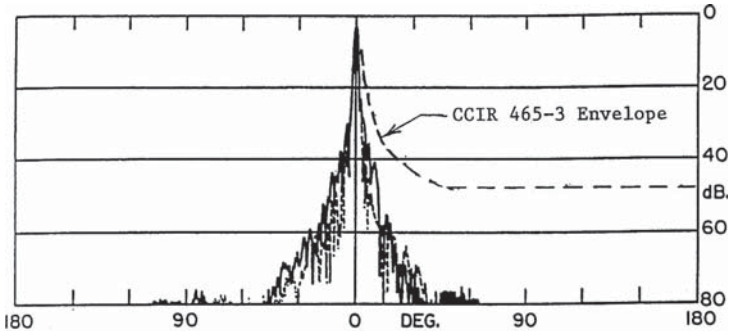


Figure 7-1 Horn reflector antenna [8]

A second example of low sidelobe antenna technology [10] further reduces the wide angle sidelobes by adding an absorber-lined tunnel to the aperture and uses amplitude tapering to reduce the near-in sidelobe levels as well. The patterns in Fig. 7-2 illustrate extremely low sidelobe levels at wide angles that are well below isotropic antenna gain levels. The absorber-lined tunnel is effective in eliminating the spillover lobes that are apparent in Fig. 7-1. In contrast to the previous low sidelobe antenna, the offset reflector is fed with a relatively large corrugated horn that provides a rotationally symmetric illumination of the reflector. The corrugated horn produces a low illumination of the offset reflector surface edges, whereas the antenna in Fig. 7-1 extends the feed aperture and illuminates the reflector surface with basically a waveguide mode distribution. The low edge illumination produces a significant amplitude taper in the aperture distribution so that the near-in sidelobe levels are also reduced. The pattern response of this design is largely limited to the antenna's aperture distribution. The CCIR sidelobe envelope values have been added to the pattern in this figure and serve to illustrate that sidelobe levels well below the envelope values can be achieved.

The third low sidelobe antenna example [11] also uses absorber-lined tunnels, and the aperture distribution has a significant amplitude taper.



Measured Pattern [10] (© 1975 IEEE)

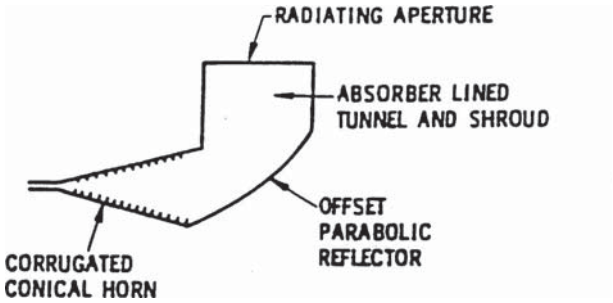


Figure 7-2 Low sidelobe offset reflector

This design applies an absorber-lined tunnel to a 6" diameter reflector. Pattern measurements at 92 GHz, as illustrated in Fig. 7-3, were made with and without the absorber-lined tunnel in place. A diagonal horn feed in a prime focus configuration provides about a 25 dB edge illumination taper, and thus low sidelobe levels near the antenna's main beam result. The feed and input waveguide blockage limits the effects of the amplitude taper and, particularly in the H-plane, the levels of the near-in sidelobes flatten rather than monotonically decrease, as would be anticipated if the sidelobes were totally dictated by the aperture distribution.

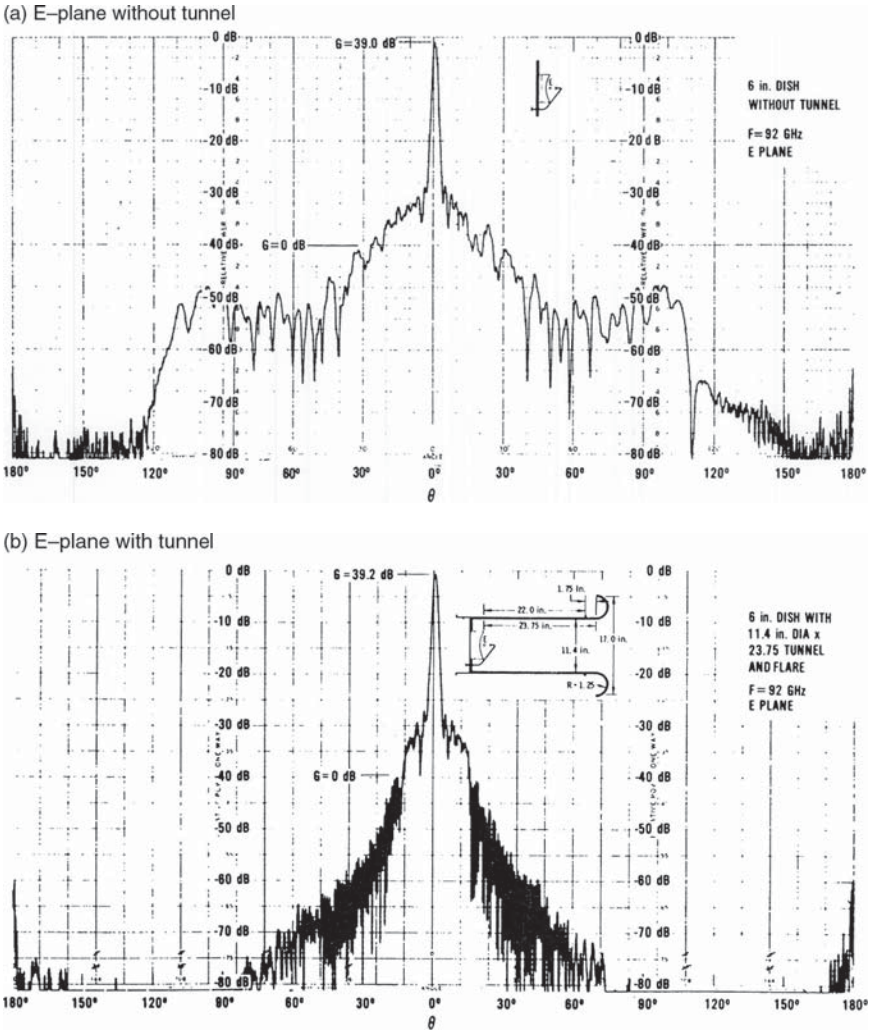


Figure 7-3 Absorber-lined tunnel reflector [11] (© 1983 IEEE)

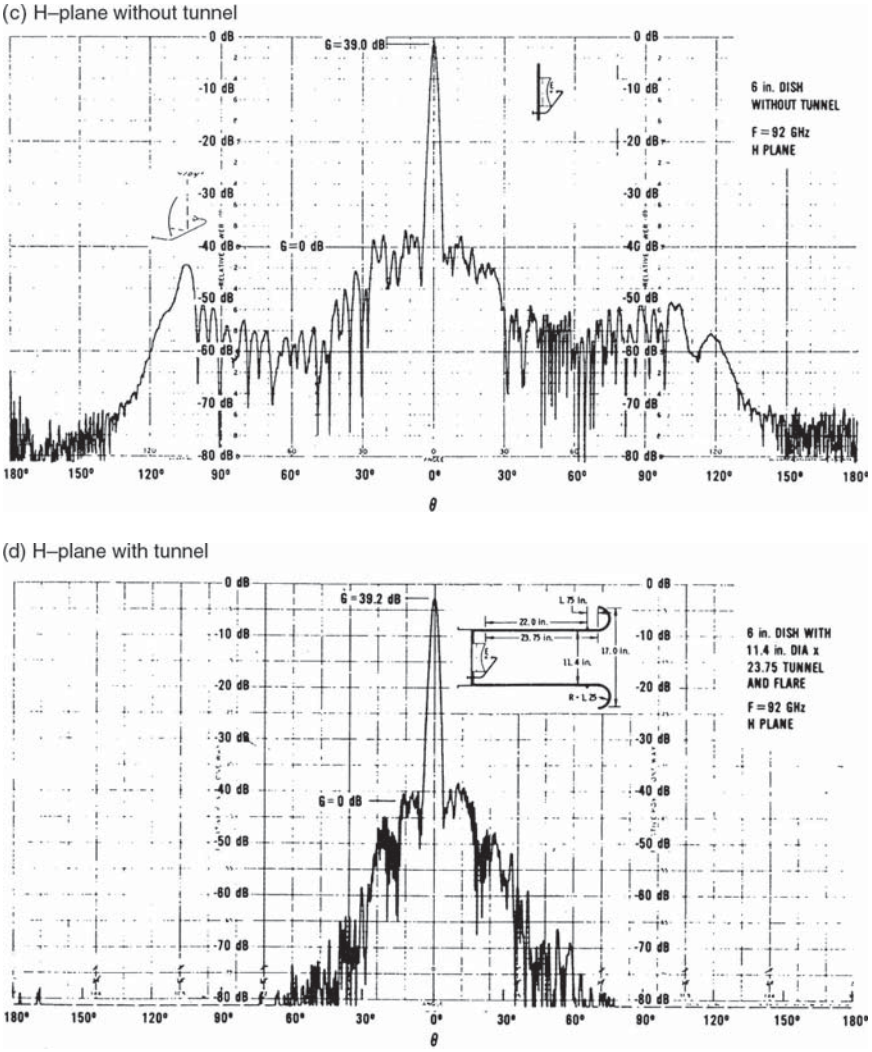


Figure 7-3 Absorber-lined tunnel reflector [11] (© 1983 IEEE) (*Continued*)

The spillover lobes from the reflector surface and mounting plate occur at about 100° from the main beam. The peak pattern levels of these spillover lobes are about –10 dBi and are lower than typical levels as a result of the large illumination taper. The patterns with the absorber-lined tunnel further reduce the wide angle sidelobes, and beyond 75° from the main beam, the levels are below –40 dBi. The measurements illustrate that the tunnel becomes effective when it blocks the radiation from the antenna reflector. Thus, the angular regions where sidelobe

reduction can be achieved can be determined from simple geometric optics arguments. The tunnel's edges are also rounded to reduce wide angle sidelobes further. The measured antenna gain with and without the absorber tunnel is identical within measurement error, illustrating that wide angle sidelobe reduction does not impact main beam gain performance since the radiation mechanisms responsible for the wide angle sidelobes do not contribute to the main beam. The control of antenna sidelobes for terrestrial applications is required principally in the azimuth plane, and partial tunnels blocking azimuth directions where interference can access sidelobes have been shown to be effective in that case.

Each of these low sidelobe antenna examples illustrates different aspects of low sidelobe design. The horn reflector antenna design in Fig. 7-1 illustrates the effectiveness of offset reflector antennas and reduced sidelobe levels resulting from direct feed radiation. Both the low sidelobe offset reflector in Fig. 7-2 and the absorber-lined tunnel prime focus reflector in Fig. 7-3 illustrate the benefits of reducing the edge illumination of the reflector that include lower sidelobe levels near the main beam, reduced spillover levels, and wide angle sidelobe levels that are significantly lower than an isotropic gain value. The effects of feed blockage are also shown to limit the sidelobes near the antenna's main beam in the absorber-lined prime focus antenna in Fig. 7-3. Absorber-lined tunnels are effective in blocking wide angle sidelobes and feed spillover. The angles at which tunnels become effective can be anticipated from geometric optics considerations at angles where the tunnel begins to block the reflector antenna's wide angle radiation. Antenna gain measurements with and without the absorber-lined tunnel illustrated in Fig. 7-3 indicate that the reduction of the wide angle sidelobes does not have an appreciable impact on the main beam gain. While the principles of reducing antenna sidelobes are well understood, satisfying system needs for low sidelobe levels will require commercial development. An examination of these factors reveals that such sidelobe control techniques result in antenna pattern responses that principally result from only the antenna's aperture distribution.

While some of these low sidelobe level design techniques are well known and can be applied to an existing antenna, several precautions must be observed. Large reflector antennas are typically constructed from a collection of panels. Leakage through the panels contributes to the sidelobe levels behind the reflector. Similarly, some reflector surfaces contain holes or cutouts for the spars supporting the subreflector or prime focus feed. Radiation from these openings also contributes to the sidelobe levels behind the reflector. Thus, an examination of the existing antenna is needed to identify additional contributors to the sidelobes. Sidelobe levels near the main beam can be controlled by replacing the

existing antenna feed with a larger and more directive feed horn to increase the aperture amplitude taper. Such approaches inherently reduce antenna efficiency and when aperture blockage is present, the reduction of the sidelobes near the antenna's main beam is not as effective as would result from an unblocked antenna aperture. Sidelobe control using tunnels surrounding the aperture have length constraints when radomes enclose the antenna. If the radome has a space frame construction, scattering from the space frame members that support the individual radome panels also contributes to the sidelobe levels. If the sidelobes are to be reduced in the azimuth plane, only a portion of the tunnel is required, in a manner similar to blinders on a horse.

Finally, whether an existing antenna is modified or a new antenna is developed, attention must be paid to leakage from microwave components and transmission line interfaces. Such leakage can exceed the radiation from the antenna itself. Electroforming techniques are an effective means of reducing leakage components by minimizing the number of component junctions. Terminating the feed aperture and measuring its pattern and gain level when terminated is a means of quantifying the leakage levels in comparison to the sidelobe response of the antenna itself. The tuning screws of microwave filters, waveguide junctions (particularly at higher frequencies), and mixer ports are typical sources of leakage radiation. Leakage components are also a result of workmanship errors. Developing effective enclosures and seals for such enclosures should be addressed to control component leakage.

7.4 Adaptive User Antennas

Adaptive interference cancellation systems for user antennas must provide effective cancellation over a much wider field of view than space segment. In principle, interference can arrive over a hemispheric volume if both ground and airborne interference sources are considered. Pointing variations either resulting from satellite motion or using the same terminal design in different locations can expand this hemispheric angular volume. The limited field of view for space segment antennas requires adaptive cancellation in those angular regions where the antenna response is principally dictated by the aperture distribution. By contrast, user segment antennas are pointed towards the satellite and terrestrial interference is received through a wider angular range and separation from the antenna's main beam. Thus, adaptive designs for user antennas must be effective for sidelobes dictated by principally second-order mechanisms, such as feed radiation and blockage, edge diffraction, and so on.

These differences in angular coverage requirements and in the mechanisms that produce the sidelobes have a significant impact on the

design requirements for adaptive interference cancellation. The sidelobe response of the space segment antennas is principally dictated by the aperture distribution having a well-defined phase center. By contrast, the response of user antennas at wide angles from the main beam results from the coherent sum of many different radiation mechanisms that individually have widely separated phase centers. The antenna sidelobe response is a summation of these individual components and each radiation component has a corresponding time delay in its arrival time. Antenna analyses for the user segment antenna form the phasor sum of these radiation mechanisms with appropriate phase delays to a common phase reference. These analyses produce the familiar patterns describing antenna performance at a single frequency. If adaptive cancellation is required at a single frequency, a phasor value containing an interference sample could cancel the interference. However, practical cancellation techniques must respond to both the interference and the desired signal's bandwidth values.

Adaptive cancellation requires the antenna response over the operating bandwidth at aspect angles corresponding to interference directions. The time delay spread or dispersion in the user antenna response at a given direction must be matched to the response of the antenna elements used in the adaptive cancellation. The frequency responses of the main antenna and the antenna elements used in the cancellation process must be identical within the tolerances previously shown in Fig. 5-7. These antenna and cancellation antenna element responses have different frequency responses, not only on a component level but also as a result of the physical separation and their combining. The adaptive weighting must provide a means of compensating for the frequency response differences between the user antenna and the antenna element combinations used in the adaptive cancellation. The adaptive weighting circuitry accordingly requires frequency-dependent weighting values to achieve effective cancellation over the bandwidth, resulting in adaptive equalization requirements. The adaptive weighing circuitry generally takes the form of adaptive transversal filters comprised of time delay components and amplitude and phase weighting for each time delay tap. Finally, interference is much closer to user antennas than space segment antennas, and protection for much higher power levels is required. The linearity of the user receiver becomes an issue for nearby high-level interference sources, and the possibility of saturating the receiver must be addressed. Thus, design attention to the system's linear dynamic range is also required. If the receiver response remains nonlinear after adaptive interference cancellation, the user receiver performance is still degraded.

Adaptive user antenna designs follow a sidelobe cancellation configuration [12], one of the original adaptive antenna designs. This design

configures a main antenna to operate as a user terminal antenna in interference-free conditions. The interference is assumed to arrive through the main antenna's sidelobe structure because the antenna's main beam is directed at the satellite. Auxiliary antennas, typically small horn antennas, are used during adaptive cancellation and sample the received interference. The gain level of these auxiliary antennas are selected so that their gain levels meet or exceed the sidelobe gain of the main antenna in the angular region covered by each auxiliary antenna element in the collection. Higher gain antenna elements cover the sidelobes near the main beam, and less directive auxiliary antenna elements are used to cover the lower gain values of the wide angle sidelobes. Thus, the main antenna receives the desired signal level through the main beam, the interference signals are received through the main antenna's sidelobes, and the main antenna has a thermal noise component. The auxiliary antennas predominately receive interference power, an independent thermal noise component, and a much smaller desired signal power than the main antenna since the auxiliary antenna does not have a high gain main beam. An example system design and measured results for representative interference signals [13] have been described.

In operation, the adaptive sidelobe cancellation design measures the correlation between the main antenna and the auxiliary antenna elements. If interference power is not present, the correlation is primarily that of independent thermal noise samples, which equals zero. If interference is present, the correlation of the interference levels received by the main antenna's sidelobes and the interference levels received by the auxiliary antenna element is non-zero because the interference is coherent upon itself. Since the interference samples are coherent, the output correlation of the main antenna and auxiliary antenna elements are non-zero. While the desired signal samples received by the main antenna and the auxiliary antenna elements are also coherent, their correlation values are quite low because the desired signal levels received by the antenna's main beam are so much higher than the desired signal levels in the auxiliary antenna elements. The existence of a correlation product indicates the presence of interference. The closed-loop adaptive system operation typically uses a least means square algorithm to determine adaptive weight values iteratively. The auxiliary antenna elements, adaptively weighted, are combined with the main antenna output. The adaptive weight values are updated until the correlation values become zero. Zero correlation is equivalent to canceling the interference at the antenna system output.

If the pattern of the antenna system comprised of the main antenna combined with the weighted auxiliary antenna outputs is plotted, antenna pattern nulls in the interference direction(s) would be observed

in the sidelobes of the main antenna. The main beam of the antenna system is not affected by the cancellation process since the auxiliary antenna element gain is much lower. The adaptive modifications to the antenna sidelobes have a minimal effect on the antenna gain. The system noise level can increase somewhat because of noise contributions from the auxiliary antennas. This noise level depends on the gain difference between the main antenna's sidelobe levels and the auxiliary antenna's gain. The earlier discussion indicated the auxiliary antennas' gain levels should exceed the main antenna's sidelobe gain values. The more the auxiliary antenna gain exceeds the main antenna's sidelobe gain levels, the less noise is added by the auxiliary antenna elements. Finally, the adaptive weight values are constrained so that the desired signal is not attacked—in other words, the adaptive weight values are bounded so that desired signal correlations cannot reduce the main beam gain. As appropriate, if the system attempts to cancel the desired signal, threshold levels in the correlation products can be used to inhibit the cancellation circuitry.

Performance measures for adaptive user antennas are the achieved SNIR and the time required to reach steady state performance. The SNIR principally depends on the residual interference levels after interference cancellation, changes in the system noise temperature resulting from noise contributions from the adaptively weighted auxiliary antennas, and changes in the overall antenna gain with adaptive weight contributions. The residual interference power depends on the effectiveness of the adaptive interference cancellation that in turn depends on the interference power levels. The system noise temperature contributions are generally small if the gain of the auxiliary antenna elements exceeds the main antenna's sidelobe levels. Changes in the antenna gain are small because pattern nulls in the antenna's sidelobe structure have a minimal impact on the antenna's directivity. The time to reach steady state performance depends on the algorithm used and its transient performance. If the interference upsets the receiver's acquisition of the desired signal, the time to recover communication includes the time for the receiver to reacquire the desired signal.

As discussed, the antenna dispersion is addressed by using an adaptive transversal equalizer to generate frequency-dependent weighting values. The total delay in the equalizer corresponds to the delay spread in the main antenna sidelobe response. One way of determining the required delay spread is to measure the antenna response over a bandwidth with a vector network analyzer. The resulting frequency domain data are transformed into the time domain using internal network analyzer capabilities. The time domain data display the antenna's time delay spread. A second way of determining the required delay spread analytically estimates the time delay spread using diffraction analyses.

The analytically derived data of the equalizer requirements for reflector antennas [14] are presented in Fig. 7-4 [15]. The desired cancellation performance is plotted as a function of a time bandwidth product with the number of taps in the transversal equalizer as parameters. The time bandwidth product equals the time delay spread in the antenna multiplied by the cancellation bandwidth. Typically, the time delay spread in reflector antennas equals two to three times the transit time across the aperture. Thus, achieving good cancellation performance in large reflector antennas requires a large number of taps in the transversal filter, and for large operating bandwidths, the required complexity of the equalizer becomes impractical.

Techniques to increase the cancellation bandwidth have had limited investigation. One proposed design [16] recommended using additional feed elements in the focal region of the main antenna as auxiliary antenna elements. The reasoning in this design concept was that such auxiliary antenna elements would contain similar dispersion characteristics as the main antenna, providing the necessary equalization to obtain broad bandwidth cancellation. A more detailed examination [17] reveals that the dispersion differences between the main and auxiliary feeds were sufficient enough that broad bandwidth cancellation would not be achieved.

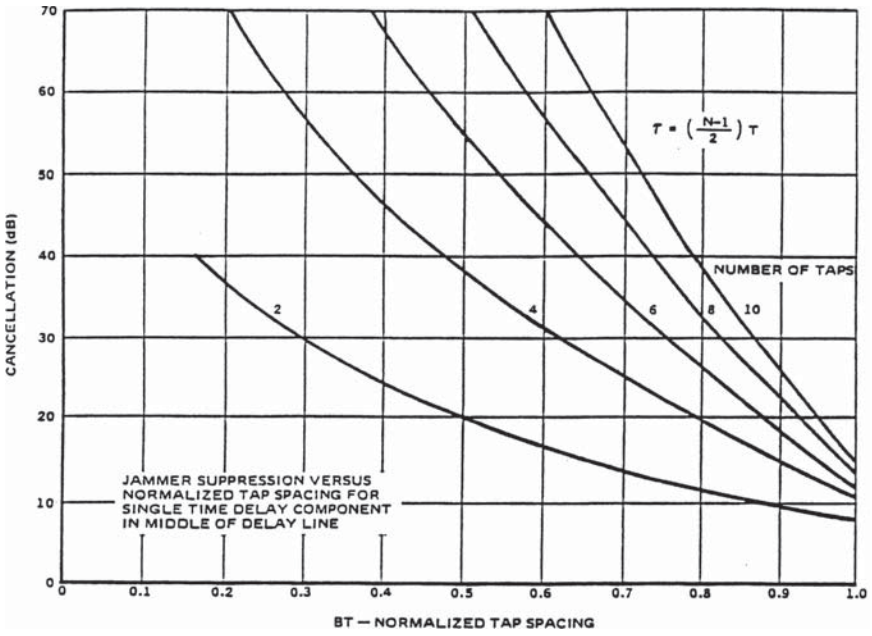


Figure 7-4 Equalization requirements for reflector antennas [14, 15] (© 1995 IEEE)

Antenna design has also been investigated in terms of adaptive system performance. The differences between Cassegrain and offset reflector antenna designs were examined [18]. Offset reflector designs have a less complex antenna response than Cassegrain designs that have the feed and subreflector blockage contributions. The less complex antenna response of the offset reflector design results in better cancellation bandwidth performance than the Cassegrain design. Additional development efforts should be devoted to antenna designs whose time domain response simplifies adaptive processing requirements.

More to the point, system designs are configured to provide a specified level of interference mitigation. Antenna interference mitigation techniques include both passive sidelobe control techniques and adaptive interference cancellation. Further development attention to achieving specified levels of interference protection by a combination of both antenna techniques is believed to be profitable. Design attention to antenna sidelobe control for user antennas addresses many of the same second-order mechanisms that result in the dispersion that limits adaptive cancellation performance. Indeed, such antenna sidelobe control techniques result in antenna pattern responses whose sidelobe levels are reduced by controlling the second-order radiation mechanisms. The resulting antenna pattern is dominated by the aperture distribution that also is not dispersive since a well-defined phase center exists. This same reasoning also results in antenna responses having less dispersion, simplifying the adaptive equalization requirements. The challenge then is to configure antenna system designs capable of providing the specified interference protection with minimal design complexity. This design approach achieves much of the interference protection at wide angles from the antenna's main beam by passive sidelobe control techniques. Adaptive cancellation protects the system from interference closer to the antenna's main beam, and since the antenna dispersion is reduced by the sidelobe control, cancellation circuitry becomes less complex because of reduced equalization requirements. Further study of this approach is recommended to meet interference requirements with passive antenna design techniques and an adaptive design with a practical level of complexity for wide bandwidth applications.

7.5 Mission Control Assets

Satellite operations depend on a variety of mission control assets that support launch operations, TT&C services, gateway stations, and monitoring functions specific to program needs. While these assets depend on the needs of a particular program, typical requirements and examples can be discussed. The mission control assets have increased emphasis on their calibration so that the on-orbit performance can be established and

monitored with minimal measurement uncertainty. In some instances, specialized systems are needed to measure performance or establish RF parameters and two examples illustrate such systems. The first example describes a terminal to monitor the on-orbit performance of GPS satellites, and the second example describes a wide bandwidth monitor receiver to identify time periods when incident signal values exceed high field strengths that could potentially damage systems.

7.5.1 Mission Control Stations

Mission control assets are an important part of the ground segment and are used by satellite operators to communicate with the satellite. The mission control station must satisfy several objectives in its operation. A primary objective is to monitor the satellite's health that is indicated by the telemetry data. These data routinely monitor such parameters as the prime power voltage and current consumption, thermal measurements, attitude variations, and other parameters used to determine the on-orbit status of the satellite. A second objective is to provide a commanding function to control the satellite's operation throughout its lifetime. Such commanding includes requesting additional data that are not routinely reported to obtain further insight if shortfalls are experienced, repositioning the space segment antennas to accommodate desired changes in their coverage requirements, substituting redundant components for failed items, and instituting and monitoring thrusters to maintain or change the satellite's orbital position. Commanding is done with extreme care, and so command authentication techniques are used to verify that the proper commands have been received prior to their execution. A third objective is to provide information used in determining the satellite's ephemeris data. The angular location provided by the mission control antenna tracking provides a limited amount of information. A second source of information is a capability to transmit a ranging code that is rebroadcast by the satellite. This information together with data from tracking radars and other sources and orbital monitoring techniques is used in a Kalman filtering algorithm to define the satellite's orbital position. A fourth objective of the mission control station is a measurement capability of the satellite's on-orbit performance. When the satellite reaches its on-orbit position and is configured for operation, measurements are made to assess compliance with the satellite's specification. Often, such measurements involve financial incentives, and calibration accuracy for the measurements is carefully assessed.

Mission control requirements are specific to program needs and space segment designs. Future mission control requirements will include capabilities to upload space segment software whose use will become more extensive in future designs. Links with a sufficiently high data

rate capability to accomplish the software transfer in a timely manner are required. Another requirement concerns the amount of link margin necessary for low data rate commanding in the event of space segment TT&C transponder shortfalls. Clearly, such margin requirements involve subjective decisions, but most mission control segments are heavily margined so that a commanding capability remains if space segment shortfalls need to be overcome. Other mission control requirements are more extensive. One example is NASA's mission control involving the TDRS satellites that support a multitude of other satellites for data relay services. The requirements to interface with TDRS satellites are described in detail [19].

7.5.2 Monitor for On-Orbit GPS Performance

A capability to monitor GPS on-orbit satellite performance is one example of ground assets that utilize equipment developed to support mission operations. The general requirements for the monitoring capability include a directive antenna to isolate a single satellite within the GPS constellation and avoid multipath errors, a capability to monitor received signal levels and their polarization accurately, a means to separate individual navigation codes and determine their power ratio, and a way to measure other signal parameters used in GPS operation. A prototype design [20, 21], illustrated in Fig. 7-5, was developed to demonstrate calibration techniques used in its operation. The achievable measurement uncertainty in received power levels is of particular interest, and an error budget projection based on measured performance was developed to assess measurement uncertainty.

This prototype design uses a 4-ft reflector antenna with the rolled edge cavity dipole feed, shown in Fig. 7-5, to cover the L5 to L1 L-band

Figure 7-5 Antenna for on-orbit GPS monitoring [20, 21]

frequencies used by GPS. The feed design has been previously described in Figs. 2- 1 and 2-2. The rotational pattern symmetry provided by this feed design results in low antenna axial ratio performance, so the axial ratio of the incident GPS signals and signal power measurements can be performed with minimum uncertainty. The receiver uses correlation techniques to isolate the code components; this prototype design examined only the L1 C/A GPS code component. The receiver implementation is based on FPGA technology that was also used to generate code components for both correlation processing and a calibration signal to determine the receiver's electronic response.

The GPS system is specified in part by the power levels received by an idealized reference user antenna that has a 3 dBi linearly polarized gain level. The monitor capability must address two questions in its calibration for received power:

1. What is the relationship between the monitoring antenna's gain and the reference antenna used in the GPS specification?
2. How do the correlation receiver output levels relate to the signal power received by the monitoring antenna?

The monitoring antenna's gain was measured by two independent methods. A far field range measurement of the antenna yielded a gain level of 22.4 dBi. The axial ratio that is required to address uncertainty in polarization parameters was measured to be 0.6 dB. A second antenna gain measurement was performed using solar radio source techniques (described in Chapter 8). The G/T value from the radio source measurement and the system noise temperature derived from the receiver noise temperature and measured antenna noise temperature yields an antenna gain level of 22.6 dBi.

The measured antenna performance must be related to the specified reference antenna. The total power of the incident signal received from the satellite is the sum of the detected outputs for the two orthogonally polarized antenna terminals. The differences in the received power output levels at these two terminals are used to calculate the incident signal's axial ratio. The total power received by an isotropic antenna matched to the incident signal's polarization equals the summed output power levels reduced by the antenna gain relative to an isotropic circularly polarized antenna. The reference antenna is a 3 dBi linearly polarized antenna. If the incident signal had ideal circular polarization, the isotropic circularly polarized antenna and the 3 dBi linearly polarized antenna would have the same signal levels. However, when the incident signal has a finite axial ratio, the received signal level for linear polarization depends on the orientation of the polarization ellipses of the incident signal and receiving antenna. For example,

if the orientation of the linearly polarized antenna were allowed to rotate, the variation in the received power would measure the axial ratio. The interpretation of the specification is the minimum power received by the linear antenna that occurs when the linear antenna is aligned with the minor axis of the incident signal's polarization ellipse. The polarization statistics (described in Chapter 1) are used to determine the worst-case polarization mismatch loss so that the total received power can be corrected for the worst-case polarization alignment. The values in Fig. 7-6 account for the 3 dB polarization mismatch loss of the linearly polarized antenna and indicate the correction for the minimum signal level for the worst-case polarization as well as the mean loss, the rms spread equal to the mean value $\pm 1\ \sigma$, the maximum value corresponding to alignment with the major axis of the polarization ellipse, and the peak-to-peak value that equals the incident field's axial ratio. The values in this figure assume the receiving antenna has a 0.6 dB axial ratio and span the incident field's axial ratio values to the 2 dB specified maximum axial ratio for GPS satellites.

The receiver's electronics were calibrated by measuring the transfer characteristics between the antenna's output terminal and the correlator. The analog components and cabling were initially measured using network analyzer techniques to establish nominal system values. The calibration uses the injection of a C/A signal waveform through a coupler

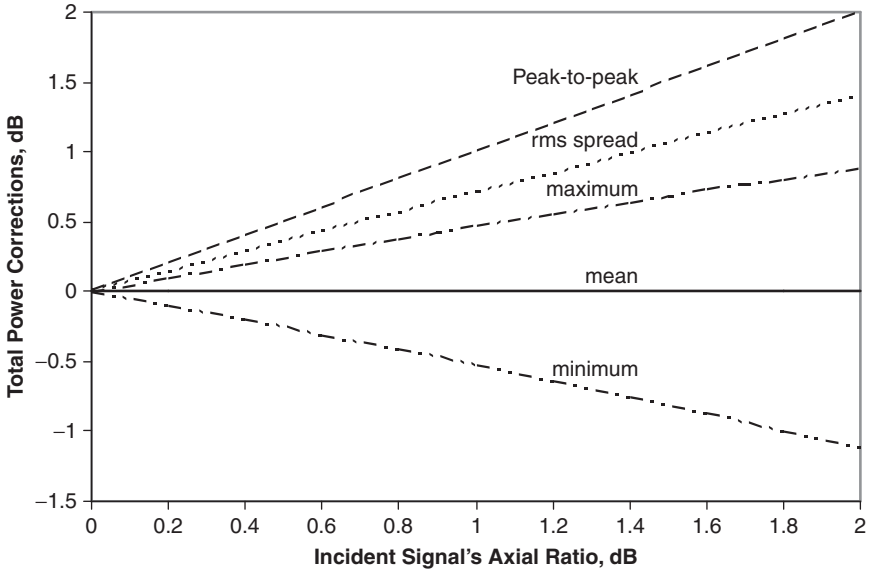


Figure 7-6 Polarization correction for linear orientation

as a calibration and measurements of its response in the receiver's correlation output. The injected code component's power level is measured using a power meter to establish a reference level accurately. The attenuators, cabling, and coupler's coupling coefficient that reduce the test signal's power level to values that are comparable to signal levels received from GPS satellites were separately measured to establish the signal level injected into the system's correlation receiver. The electronic transfer function was then determined from the output correlation levels and the signal level of the injected calibration signal. Separate measurement of the analog circuitry using swept frequency generators were used as a means of independently verifying the calibration. The test generator was also directly injected into the system's correlation receiver to evaluate the response of the digital circuitry. The use of a C/A code for calibration advantageously includes the correlation receiver's implementation loss in the calibration process. A comparison of the two independent means of measuring the receiver's transfer function provided confidence in the calibration values. Finally, since the analog circuitry's gain level can vary during the time required for the satellite signal's measurement, the time stability of the overall transfer function was also measured and included in the error budget projection.

The error budget projection of measurement uncertainty in Table 7-1 reflects the elements of the calibration process discussed earlier. This error budget separately addresses the antenna's calibration, the test signal injection used to calibrate the receiver's transfer function, and the calibration of the system electronics. The projected measurement uncertainty is the rss sum of the error components and has a value of 0.79 dB.

TABLE 7-1 Error Budget for Power Measurement Uncertainty [20, 21]

Item	1 Sigma Uncertainty, dB
Antenna	
Gain Uncertainty, dB	0.5
Polarization Mismatch, dB	0.025
Pointing Loss, dB	0.3
Test Signal Injection	
Level Reading, dB	0.2
Attenuator Setting, dB	0.2
Coupling Uncertainty, dB	0.1
Electronics	
Correlation Response	0.35
Carrier Tracking	0.2
Measurement Accuracy, dB	0.3
Total	0.79 dB

7.5.3 Incident Signal Level Monitor

A second example of a specialized system is a monitor used during satellite launch processing. Launch sites typically contain other support systems such as tracking radars that are located near launch processing facilities. Such support systems can produce incident power densities at launch processing facilities that exceed levels used in testing satellite subsystems and concerns regarding damage to satellite components prior to the launch. Recognition of this situation resulted in imposing sector blanking requirements on the support systems, but in addition, a means to monitor continuously the incident power densities was desired. This application imposes several requirements on such a monitor. A broad instantaneous bandwidth that spans the frequency range of nearby support systems is required. While spectrum analyzer instrumentation is capable of providing a detailed examination of the spectral characteristics of individual emitters, the detection of such emitters is possible only for those time periods when the spectrum analyzer covers the emitter's bandwidth. Accordingly, a large instantaneous bandwidth is required in this application. Commercially available monitors are available with a broad instantaneous bandwidth, but such monitors respond to average power levels rather than peak power levels. As discussed in Chapter 5, vulnerable devices such as LNAs are susceptible to peak power levels, not average power. A further requirement is to provide broad angular coverage to limit the number of required sensors. A cost-effective design capable of operating in an outdoor environment with minimal power consumption is needed.

A simple system design was developed to satisfy these requirements. The monitor was designed to provide an indication based on when threshold levels in incident field strengths are exceeded. The incident power density in terms of field strength E_i equals

$$P_d = |E_i|^2/Z_o$$

where the magnitude of the incident field is the scalar sum of the orthogonally polarized components of the incident electric field, and Z_o is the free space impedance equal to $120\pi\,\Omega$. The total incident electric field requires an antenna system that responds to orthogonally polarized components, and the power received by each of the elements is summed after detection so the total power from the incident field is received. The received power equals the incident power density multiplied by the effective aperture of the antenna, equaling

$$P_r = \lambda^2|E_i|^2G_r/(4\pi Z_o)$$

where λ is the wavelength and G_r is the gain of the receiving antenna. The response of the power detectors is independent of frequency, thus

the frequency response of the antenna gain must vary as $1/\lambda^2$ so that incident field levels independent of their operating frequency are received by the detectors. While a variety of networks can be used to produce such an antenna gain response, a simple means was used in this case. The orthogonal antenna elements used in this design were simple linear patch antenna elements that are orthogonally oriented to respond to orthogonal linearly polarized incident field components. The mismatch loss of such an antenna to first order has a frequency response below its resonant frequency that follows the desired $1/\lambda^2$ dependence. This behavior was used to provide the desired frequency response of the antenna system.

The overall receiver in Fig. 7-7 illustrates the two linearly polarized microstrip patch antenna elements that are printed on a duroid board having an 8-inch-square size. The system is required to cover a 1 to 6 GHz frequency range for this application to detect the power density from potential high-power emitters. The antenna pattern responses in the principal planes in Fig. 7-8 illustrate the broad coverage characteristics of the design and the desired reduced gain levels at the lower frequencies. The H-plane patterns are roughly cosine-dependent, as anticipated. The E-plane patterns can be visualized as two equivalent slot radiators located at the ends of the linear elements. At low frequencies, the equivalent slots have field components that are oppositely directed, and their excitation tends toward an equal phasing that produces a pattern behavior that tends towards a null on-axis, as illustrated by the 1 GHz pattern. As the frequency increases towards the resonance value, the phase shift from one end of the patch to the other increases towards 180^{o}. The field components

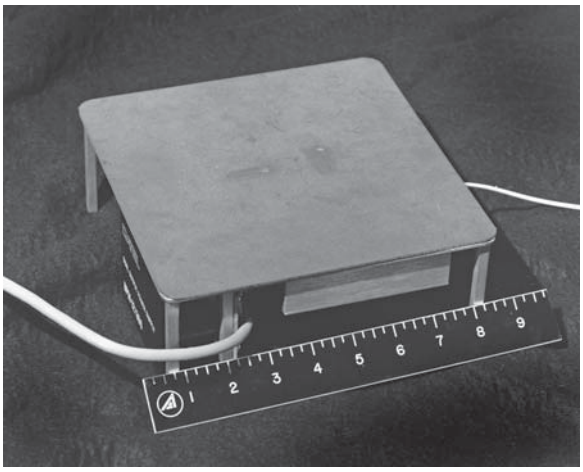


Figure 7-7 Microstrip antenna elements for incident signal level monitor [22] (© 1992 IEEE)

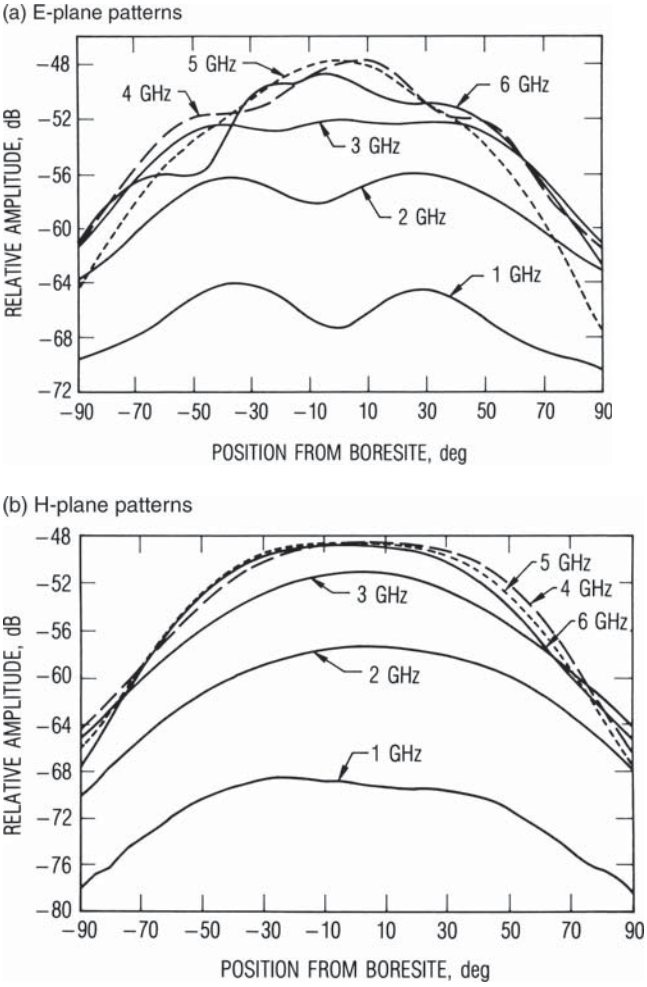


Figure 7-8 Principal plane patterns for microstrip antenna [22] (© 1992 IEEE)

in the equivalent slots are oppositely directed so the radiation from the two slots becomes more directive on-axis. The frequency coverage of the antenna design is limited by the axial pattern nulls at low frequencies and loss of coverage at the high frequency by a directive pattern. The frequency coverage is dominated by the impedance mismatch loss and this is one example where antenna mismatch loss can be usefully employed. The antenna gain variation over a quadrant of coverage is indicated in Fig. 7-9.

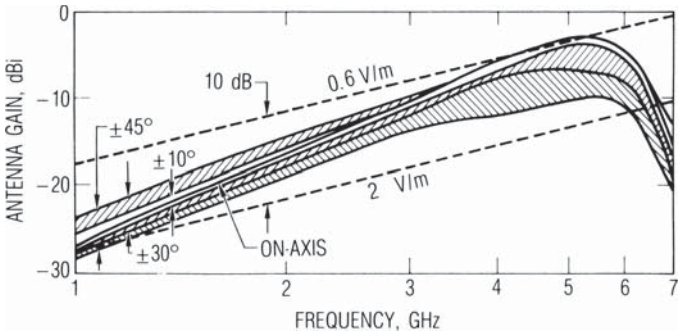


Figure 7-9 Antenna coverage characteristics [22] (© 1992 IEEE)

The receiver electronics follow Fig. 7-10 and detect three threshold field strength values, 2, 10, and 50 volts/meter. Diode detectors follow the microstrip patch elements and their outputs are summed. Such detectors, as illustrated in Fig. 7-11, have a very broad bandwidth, flat frequency response. A comparator circuit forms the threshold function and a one-shot circuit stretches received pulses sufficiently to allow recording on a stripchart recorder. The video amplifiers following the detectors have a 10 MHz bandwidth so that received pulses with a width as small as 250 nsec can be detected with little energy loss.

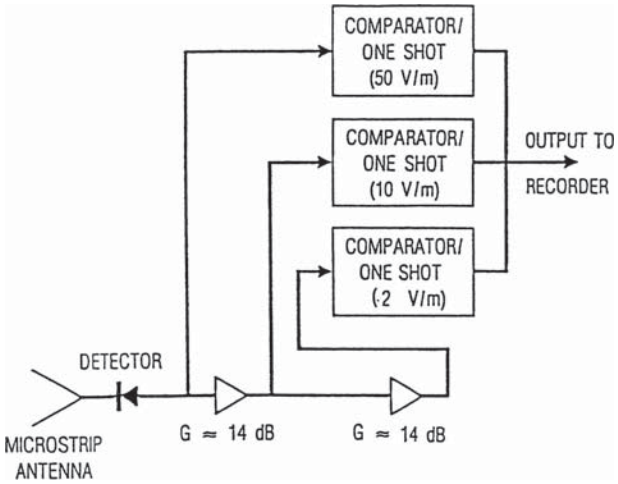


Figure 7-10 Receiver for monitor [22] (© 1992 IEEE)

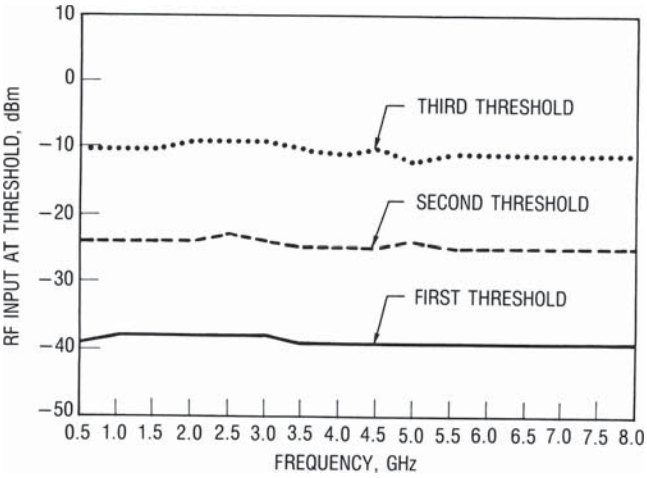


Figure 7-11 Electronics response of receiver [22] (© 1992 IEEE)

An example response of the receiver in Fig. 7-12 illustrates the receiver's detection capability. In this case, a nearby radar that scans in azimuth was detected. When the radar's main beam illuminates the receiver, the second threshold level (10 volts/meter) is exceeded, while the radar's sidelobe illumination can exceed the first threshold (2 volts/meter).

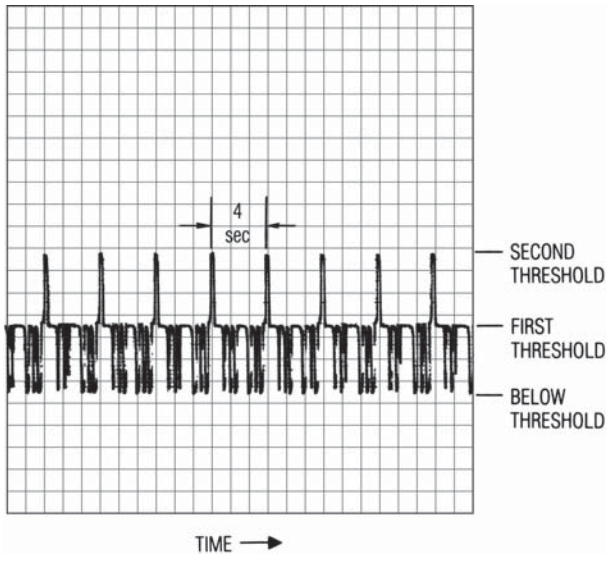


Figure 7-12 Measured response to radar signal [22] (© 1992 IEEE)

References

1. R. B. Dybdal, "User Segment Antenna Development Issues," *1998 IEEE MILCOM Symposium Digest* (October 1998).
2. ———, *Electrical and Mechanical Characteristics of Earth Station Antennas for Satellite Communications*, Electronic Industries Alliance/Telecommunications Industry Rept TIA/EIA-411, Revision A (September 1986).
3. P. Monsen, "Multiple Access Capability in Mobile User Satellite Systems," *IEEE Trans Selected Areas on Communications*, vol. SAC-13 (February 1995): 222–231.
4. R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications* (New York:, Prentice -Hall, 1995), Chapter 8.
5. R. B. Dybdal and S. J. Curry, "Adaptive Transmit Antenna," *1997 IEEE AP-S Symposium Digest* (July 1997): 2410–2413; see also R. B. Dybdal and S. J. Curry, "Adaptive Transmitting Antenna" (July 14, 1998): U.S. Patent 5781845.
6. ———, CCIR Recommendation 465-3.
7. A. B. Crawford, D. C. Hogg, and L. E. Hunt, "A Horn Reflector Antenna for Space Communications," *BSTJ*, vol. 40 (July 1961): 1095–1116.
8. J. N. Hines, T. Li, and R. H. Turrin, "The Electrical Characteristics of the Conical Horn-Reflector Antenna," *BSTJ*, vol. 42 (July 1963): 1185–1211.
9. D. T. Thomas, "Design of Multiple-Edge Blinders for Large Horn Reflector Antennas," *IEEE Trans Antennas and Propagation*, vol. AP-21 (March 1973): 153–158.
10. H. P. Coleman, R. M. Brown, and B. D. Wright, "Parabolic Reflector Offset Fed with a Corrugated Horn," *IEEE Trans Antennas and Propagation,* vol. AP-23 (November 1975): 817–819.
11. R. B. Dybdal, "Millimeter Wave Antenna Technology," *IEEE Trans Selected Areas on Communications*, vol. SAC-1 (September 1983): 633–644.
12. P. W. Howells, "Intermediate Frequency Sidelobe Canceller" (August 24, 1965): U.S. Patent 3,202,990.
13. K. M. SooHoo and W. Masenten, "Adaptive Sidelobe Canceller Designs for Large Earth Terminals," *1984 IEEE MILCOM Symposium Digest* (October 1984).
14. R. A. Dell Imagine and K. M. SooHoo, "Adaptive Sidelobe Canceller Designs for Large Earth Terminals," *Proc 1980 Adaptive Antenna Symposium*, RADC Document TR-80-378.
15. R. B. Dybdal, "Assessment of Antenna Interference Reduction Techniques for Communication Satellite Systems," *1995 IEEE MILCOM Symposium Digest* (November 1995).
16. G. G. Chadwick, J. C. Charitat, W. Gee, C. C. Hung, and J. L. McFarland, "Adaptive Antenna/Receiver Processor System," *1981 IEEE AP-S Symposium Digest* (June 1981).
17. R. B. Dybdal, "Limitations of a Sidelobe Canceller Concept," *1991 IEEE AP-S Symposium Digest* (June 1991).
18. R. H. Ott and R. B. Dybdal, "The Effects of Reflector Antenna Diffraction on the Interference Cancellation Performance of Coherent Sidelobe Cancellers," *IEEE Trans Antennas and Propagation*, vol. AP-34 (March 1986): 432–439.
19. ———, *Space Network Users Guide*, NASA Rept 450-SNUG (June 2002).
20. M. D. Partridge and R. B. Dybdal, "Design/Calibration of GPS On-Orbit Monitoring Capability," *Institute of Navigation National Technical Meeting Digest* (January 2005).
21. R. B. Dybdal and M. D. Partridge, "Calibration of GPS On-Orbit Monitor," *2005 AMTA Symposium Digest* (October/November 2005).
22. R. B. Dybdal, T. T. Mori, and A. M. Castaneda, "A Low Cost High Power Microwave Receiver," *IEEE Trans Instrumentation and Measurement*, vol. IM-41 (June 1992): 349–352.

Chapter

8

Antenna Test Facilities and Methodologies

8.1 Overview

Antenna testing is a very important part of communication satellite development and qualification. Antenna test techniques, instrumentation, and test facilities are well developed for a wide variety of applications [1, 2, 3] and form the basis to evaluate satellite antennas. The AMTA (Antenna Measurement Techniques Association) has long been a forum to discuss measurement facilities, methodologies, evaluation techniques, and measurement uncertainties as documented in symposium records. Satellite antenna evaluations, however, pose unique requirements that are not normally part of antenna system evaluation. Accordingly, test techniques and methodologies, instrumentation, and facility requirements must be tailored to meet the specific requirements for satellite antenna designs.

Antenna testing characterizes both their spatial and terminal properties, and clearly such testing and evaluation criteria depend on the application and program requirements. At a system level, the system's G/T and ERP levels dictate the design communication performance, but these parameters in turn depend on the antenna's characteristics. The spatial properties of antennas include gain, pattern, and polarization parameters and their variation over their operating bandwidth requirements. Space segment antennas provide service to earth-located coverage areas, thus the system-level parameters are the minimum values within the coverage area burdened by antenna pointing errors and satellite attitude uncertainty. User segment antennas must point to the satellite and their projected tracking errors likewise burden the user

segment parameters. The antenna's polarization must also be established. The antenna's polarization characteristics and the polarization of the incident field together define polarization mismatch and polarization isolation values for operational system designs, as described in Chapter 1. The terminal characteristics of an antenna define the interface impedance with the system's RF electronics and the resultant signal transfer characteristics.

The combination of increased design complexity, the evolution from the evaluation of passive RF parameters to system-level performance measures, and increased integration of electronics with the antennas provide challenges for both space and user segment antenna designs. These challenges place increased emphasis on measurement facility requirements and techniques to quantify measurement uncertainty. RF evaluation facilities are well developed but the specialized needs of future satellite antenna design require further extensions of existing capabilities. Test techniques for adaptive antenna technology, antennas integrated with system electronics, and antenna tracking techniques provide examples of such extended test requirements. The system-level evaluations for these examples illustrate the need for increased instrumentation, and facility requirements and performance measures in their evaluation expand from the conventional parameters used for antenna testing to system performance parameters. Increased use of computer techniques for instrumentation control, data processing, and archiving will be required.

8.2 General-Purpose Test Facilities

Antenna parameters, as they are defined, assume that the antenna responds to a uniform plane wave. Test facilities therefore must provide uniform test fields to evaluate the antenna's response to a plane wave. The extent of the facility's test region that has sufficient uniformity is commonly referred to as the quiet zone. Three generic test facilities have been developed for this purpose: (1) far field ranges, (2) compact ranges, and (3) near field sampling techniques. As is the case in many instances, no one type of facility is universally advantageous, so that the selection of facilities to evaluate test articles must address the technology to be evaluated, the application's requirements, and facility availability. The instrumentation to support the facility is equally important and the calibration of antenna standards to establish absolute gain values and polarization properties of test articles is also required. Present general-purpose instrumentation and supporting software is well developed. Future antenna designs that integrate the antenna with system electronics will need to interface at IF or digital levels. The evaluation of adaptive systems further expand test requirements to represent both

desired and interfering signals having differing arrival directions and spectral characteristics, as well as additional instrumentation to evaluate adaptive system performance. Software control and data processing will have increased application as test requirements and data volume continue to expand to evaluate the more demanding requirements of future antenna systems. The validation of test software will also have increased emphasis.

8.2.1 Far Field Ranges

The earliest and most common antenna test facility is the far field range. This facility uses a separated test source to illuminate the antenna under test, as shown in Fig. 8-1. The antenna under test is illuminated by a source separated sufficiently from the test article to provide uniform illumination with minimal phase curvature over the antenna under test. A positioning system, such as the one that will be illustrated in Fig. 8-6, supports the antenna under test and orients it to measure the antenna response over the required angular volume. An important issue in far field measurements is reflections from the range facility that degrade measurement fidelity. The multipath component indicated in the figure is one example of facility reflections that differ from ideal free space conditions.

The antenna under test and the illuminating test source must satisfy far field conditions. At very low frequencies, the far field requirement is a separation of 8 to 10 wavelengths to avoid inductive coupling between the test source and the antenna under test. However, for most satellite antennas, the far field requirement is based on sufficient separation from the source antenna that the antenna under test is uniformly illuminated. Generally, the antenna under test is larger than the test source antenna so that the far field requirements of the antenna under test dictate the required separation. The far field conditions limit the quadratic phase error of the test source's spherical wave to less than 22.5°. This condition requires a $2D^2/\lambda$ separation value, where D is the overall antenna size and λ is the operating wavelength.

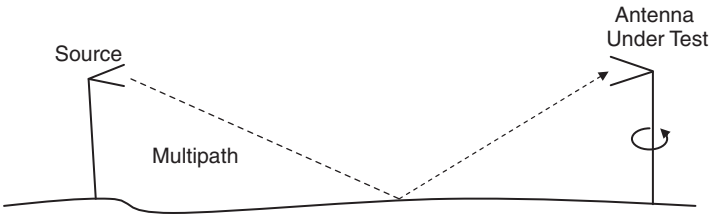


Figure 8-1 Far field range

The dimension D encompasses the geometry that radiates. For high gain antennas, D is the diameter of the radiating aperture. However, if a wide field of view antenna illuminates the satellite structure, the dimension, D, encompasses the antenna and the portions of the satellite's structure that radiate. The far field requirement in such cases greatly exceeds that of the antenna by itself. Example far field values in Fig. 8-2 illustrate that large antennas at high frequencies have excessive far field requirements. Such far field requirements typically are unavailable and also have increased potential for multipath degradation (indicated in Fig. 8-1) so that alternative techniques must be devised. For more modest range separations, existing instrumentation is capable of time domain processing to reject these facility components. At large distances and shallow illumination angles, the time delays between the direct and multipath signals become small enough that time-delay rejection of multipath becomes limited.

Far field ranges are commonly thought of as outdoor facilities. Outdoor measurements are usually precluded for flight hardware evaluations because of the risk in moving flight hardware and the potential of damage from wind and contamination. Indoor far field facilities have been developed and are referred to as anechoic chambers because their surfaces are covered with absorber to reduce facility reflections. Absorber is most effective in reducing reflections when the incident RF energy is normal to the absorber, and the absorber effectiveness is reduced as the incidence angle approaches grazing angles. Such issues and design techniques to improve reflectivity at grazing incidence [3] have been investigated based on resistance card fences to reduce reflections at grazing incidence. Anechoic chambers have been developed that permit operation over very wide frequency ranges [4]. Such facilities are used in evaluating relatively small flight antennas such as earth coverage horns. Measurement for larger higher-gain antennas is often limited by their far field requirements that exceed the facility's dimensions.

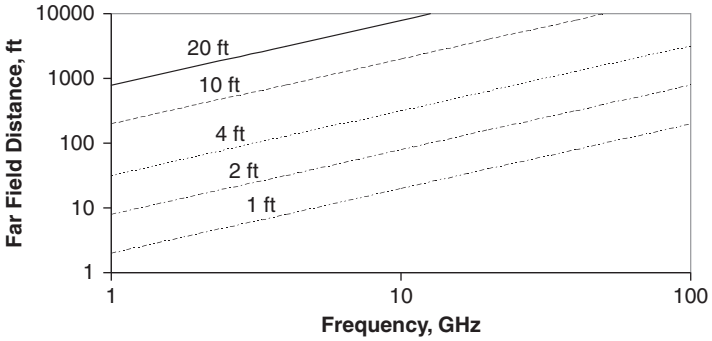


Figure 8-2 Far field distance

In operation, the antenna under test is mounted on the positioner to allow measurements of the antenna's angular variations, as will be illustrated in Fig. 8-6. Positioners are typically elevation over azimuth design, where a positioning arm that allows elevation variation is placed over an azimuth turntable. The antenna's mounting on a positioner must be referenced to the antenna coordinates. Typically, the antenna's boresight axis where the maximum antenna gain level occurs is established by measuring the antenna's main beam on both sides of the boresight axis at different pattern levels, as will be described. Once the antenna's RF boresight axis is established, the reference to the mechanical axes of the antenna structure is known and must be documented. For larger antennas, this documentation may use tooling balls or optical alignment cubes. In the coordinate system previously defined in Fig. 1-1, the boresight axis is aligned with the z-axis of that coordinate system. Patterns are then taken in the principal planes that are great circle cuts. A convenient feature is a polarization head mounted on the elevation arm that provides a third axis of rotation about the z-axis. This arrangement permits the measurement of the antenna pattern variation in different azimuth angles. For example, patterns measured at a 45° azimuth angle to the principal planes result in diagonal pattern cuts. An example of a measured diagonal plane pattern has been previously given in Fig. 2-4.

The polarization characterization of the antenna under test is established by measuring its response to the polarization of the incident field produced by the test facility. Commonly, the illuminators provide the capability to provide orthogonally polarized illumination for test purposes. In this way, the principally polarized and the cross-polarized response of the antenna under test can be determined. As a practical matter, linear-polarized incident fields with high polarization purity are often easier to achieve than circular polarization with comparable polarization purity. The polarization characteristics of antennas under test that are designed for high polarization require attention to the test facility's polarization purity to obtain accurate results. The statistics for polarization efficiency described in Chapter 1 provide a means to assess the uncertainty in polarization measurements. If the antenna is characterized by its response to two orthogonal polarizations and the relative phase between the two polarization responses, the response to any polarization state may be determined. This process yields a full polarization characterization of the antenna under test and requires careful attention to the polarization purity of the incident field and the accuracy with which the relative phase between polarization components can be measured.

When the axial ratio of the antenna under test is required, another means of measurement can be performed that is referred to as a rotating

linear response. An illuminator having linear polarization with high purity is rotated while the antenna under test is being measured. The resulting variations in the measured response are the minimum and maximum values of the polarization ellipse of the antenna under test. These values define the axial ratio, and the variation of the axial ratio is measured as the antenna under test is rotated by the positioner. Example patterns measured with a rotating linearly polarized source have been previously illustrated in Fig. 2-4. While the axial ratio is directly measured by this method, the tilt of the polarization ellipse of the antenna under test requires the relative phase between orthogonal orientations. Thus, while the axial ratio is determined by rotating linear measurements, the full polarization characterization is not made and a means to transform the measured response to any arbitrary polarization is not possible. The full polarization characterization entails a significant amount of additional measurement attention, and knowledge of the polarization ellipse of the incident field used operationally is required. Such information is generally not available, thus the specification of the axial ratio is generally used and the uncertainty of the polarization efficiency can be determined (as described in Chapter 1).

In addition to the characterization of the antenna's angular variation described by relative pattern levels, the absolute gain level of the antenna must also be established. Commonly, this gain level is established by comparing the response of the antenna under test at its peak level on the boresight axis with the response of a standard antenna having a known calibration at the frequencies of interest. The absolute gain of the antenna under test is therefore a comparative measurement in this case. In principle, absolute gain levels can be determined from a link analysis and calibration of the illuminating antenna; however, the measurement uncertainties in establishing link performance typically exceed the measurement uncertainties of a comparative approach.

Measurement uncertainty for far field facilities must address several error sources [5]. These error sources include illumination, facility interactions, instrumentation, and knowledge and measurement of reference antennas for comparative gain measurements. Antenna gain is defined based on illumination of a uniform plane wave, thus the capabilities of the measurement facility to provide uniform illumination of the antenna under test are a key requirement. The measurement of the uniformity of the test fields over the volume to be occupied by the antenna under test is a key requirement. Such measurements are generally performed by moving a probe antenna over the test volume and comparing the measured amplitude and phase variations with those of an ideal plane wave. The portion of this test volume that has adequate amplitude and phase flatness is commonly referred to as the "quiet zone." The quiet zone criteria are generally a phase variation less than

22.5° and an amplitude variation of less than a few tenths of a dB. Field probe measurements of the quiet zone often use linear actuator technology to position the probe within the quiet zone for measurements.

Measurement uncertainty from illumination errors can then be assessed by exploring the effects of the illumination imperfections with existing computer analysis codes to quantify the measurement uncertainty. For example, perturbing the design aperture distribution with the measured amplitude and phase imperfections and comparing the gain and sidelobe levels that result with an ideal plane wave response is a common way to quantify illumination errors close to the main beam. Facility reflections are coherent with respect to the direct fields in the quiet zone. The coherent interaction between the direct and reflected components can be addressed by using the coherent RF error statistics (discussed in Chapter 1). For example, a 0.5 dB peak-to-peak variation can result from a reflected component that is 30 dB lower than the direct component (see Fig. 1-15). Thus, careful attention must be paid to facility reflections. The variations clearly depend upon the dynamic range between direct and reflected components. For example, in measuring a high-gain antenna's sidelobe levels, the antenna's main beam may also coincide with a source of the facility's reflection, such as a side wall, and the instrumentation receives both the desired direct signal path and the reflected path coupled through the antenna's main beam. An assessment of the difference between the antenna's main beam reflected by the facility and the sidelobe level to be measured yields a means to estimate the uncertainty in measuring the sidelobe level.

A more detailed assessment of illumination errors describes other processing techniques to assess measurement uncertainty. Field probe measurements in the quiet zone can be processed [6, 7] to obtain a plane wave spectrum of the illumination errors. The resulting plane wave spectrum that identifies the spatial distribution of illumination can be used together with existing analyses codes to project the pattern characteristics of the antenna that include the effects of facility imperfections. A comparison of this pattern that includes facility imperfections and the pattern with ideal plane wave illumination provides estimates for the measurement uncertainty of a specific antenna over a range of angular coordinates. In this way, the effects of illumination errors on the main beam gain and various regions of the sidelobe structure [3] can be established.

The preceding factors limit the accuracy of relative pattern measurements. Absolute gain levels are commonly established by using a calibrated antenna in a "gain by comparison" procedure. The uncertainty in this process results from two factors. The first is the accuracy with which the gain of the reference calibrated antenna can be established. The second is the accuracy with which the reference calibrated antenna

can be measured in the facility. Further error sources include polarization mismatch loss uncertainty (discussed in Chapter 1) and antenna pointing errors. These latter error components typically have small values.

The overall measurement uncertainty is assessed through an error budget projection [8]. Typical component errors, indicated in Table 8-1, assess illumination errors, polarization uncertainty, positioning alignment, and antenna gain uncertainty for comparative standard for both their calibration and measurement within the facility. The overall measurement uncertainty is the algebraic sum of the mean errors and the rss (root sum square) of the rms errors given by the standard deviation of the error components. The statistical distribution of some of the error components is not Gaussian, as has been shown for the coherent RF error and polarization mismatch statistics (as discussed in Chapter 1). In other cases, the distributions are unknown. From a statistical standpoint, the rms errors have meaning but the confidence values of the errors cannot be validly projected because the distribution of the component errors is not known. The central limit theorem is often invoked, and with a sufficient number of independent errors, the distribution of the values often tends towards a "bell-shaped" distribution. The overall error distribution, however, is not Gaussian, particularly for distribution values beyond the standard deviation. Accordingly, confidence values in measurement uncertainty cannot be made, as has been discussed [9].

8.2.2 Compact Ranges

The compact range was developed to permit antenna measurements at much shorter separations than required by far field facilities. The plane wave that illuminates the antenna under test is generated by the near field of an optical antenna whose size is at least twice as large as the test antenna. The near field of an optical antenna near its boresight axis is a collimated field, and a test volume results over a region of this near field

TABLE 8-1 Typical Antenna Measurement Errors

Illumination Errors
Incident field uniformity
Facility multipath
Polarization Uncertainties
Incident field impurities
Facility multipath
Positioning alignment
Antenna gain standard
Calibration uncertainty
Measurement uncertainty

that is sufficiently uniform for test purposes. While the original compact range [10] used a lens antenna, present-day designs use reflector technology. The principal design issues in compact ranges are the control of the edge diffraction that distorts the aperture's plane wave fields in the test region and leakage from the feed illuminator of the reflector. Reflector edge treatments that include rolled surfaces and serrated edges [11] are commonly used. Pattern control and shrouding techniques are used to limit the contributions from the illuminator's sidelobes and backlobes in the quiet zone direction. The goal is to produce the quiet zone fields wholly from the collimated near field of the compact range reflector.

Compact ranges offer two distinct advantages over far field facilities. Test signals from a far field illuminator spread spherically, whereas the collimated near field propagates as a plane wave without spherical spreading. Thus, for a given illuminator power, the power density in the test region of a compact range is higher than that of a far field range. The increased power density increases measurement sensitivity. The second advantage is that the collimated near field provides reduced illumination of the side walls, floor, and ceiling of the measurement facility, reducing their errors and absorber reflectivity requirements. In both anechoic chambers and compact ranges, the back wall of the facility is illuminated, but this illumination is at the normal incidence, where absorber performance is most effective.

One example of a compact range, illustrated in Fig. 8-3, has a particularly simple design [12]. This design uses a commercial offset reflector. Rather than incurring the expense and complexity of an edge treatment,

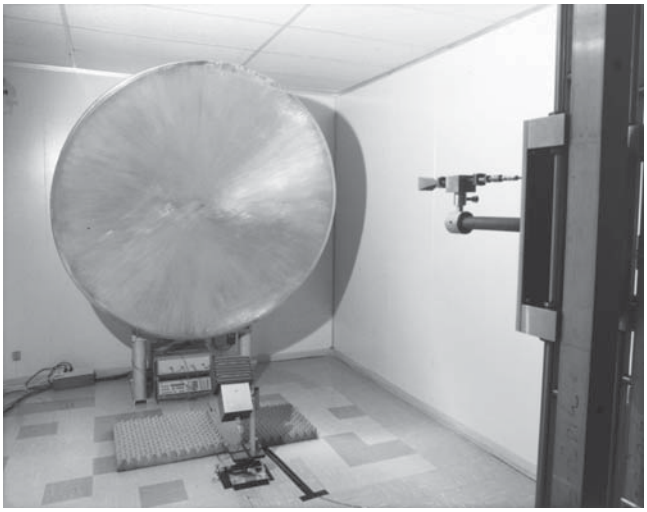


Figure 8-3 Simple compact range [12]

the illuminator was designed to strongly under-illuminate the reflector. The diffraction from the reflector edges is reduced by this under-illumination; the amplitude taper at the reflector edge is about 25 dB. The compact range illuminator at the focal point also has a simple design [13, 14] and is comprised of a standard gain horn surrounded by an absorber-lined tunnel that extends 4" beyond the horn aperture. The absorber tunnel reduces the normal high E-plane sidelobes of the standard gain horn. While this low sidelobe horn antenna is a very simple design in comparison to corrugated or dual mode horns, its outdoor applications are limited by the problem of a waterproofing absorber, and its space applications are limited by outgassing and ultraviolet deterioration of the absorber. The measured E- and H-plane patterns in Fig. 8-4 of this horn illustrate comparable beamwidths and a low sidelobe response of the horn in both planes; the measurements agree well with an iterative absorbing screen analysis [14] of the absorber-lined tunnel. The absorber-lined tunnel has very low wide angle sidelobes to minimize disturbances of the test zone fields. The field uniformity was measured by the probe antenna mounted on an xy scanner, seen in the foreground of Fig. 8-3. Measured and calculated amplitude contours (3 dB increments between contours) of the compact range's quiet zone fields in Fig. 8-5 illustrate good agreement. The measured phase contours have

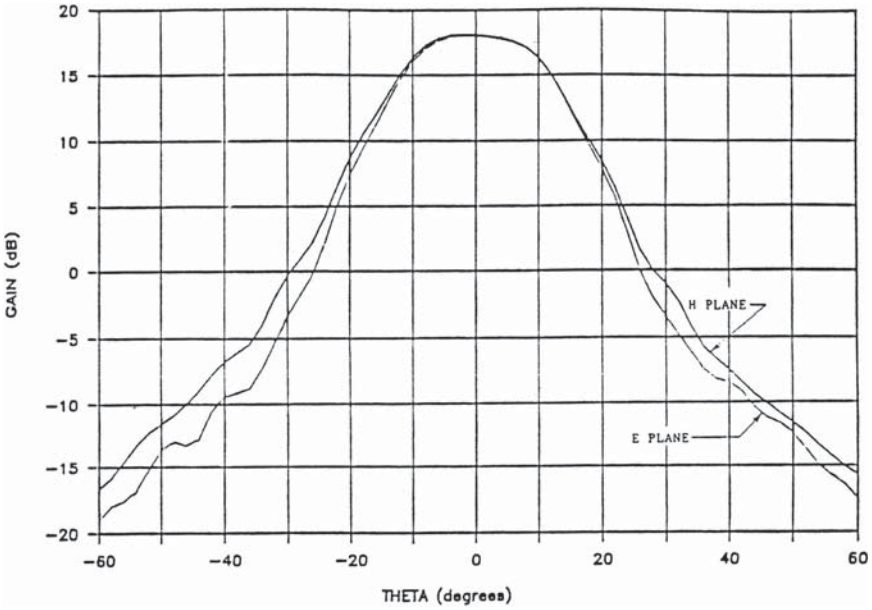


Figure 8-4 Compact range illuminator patterns [14]

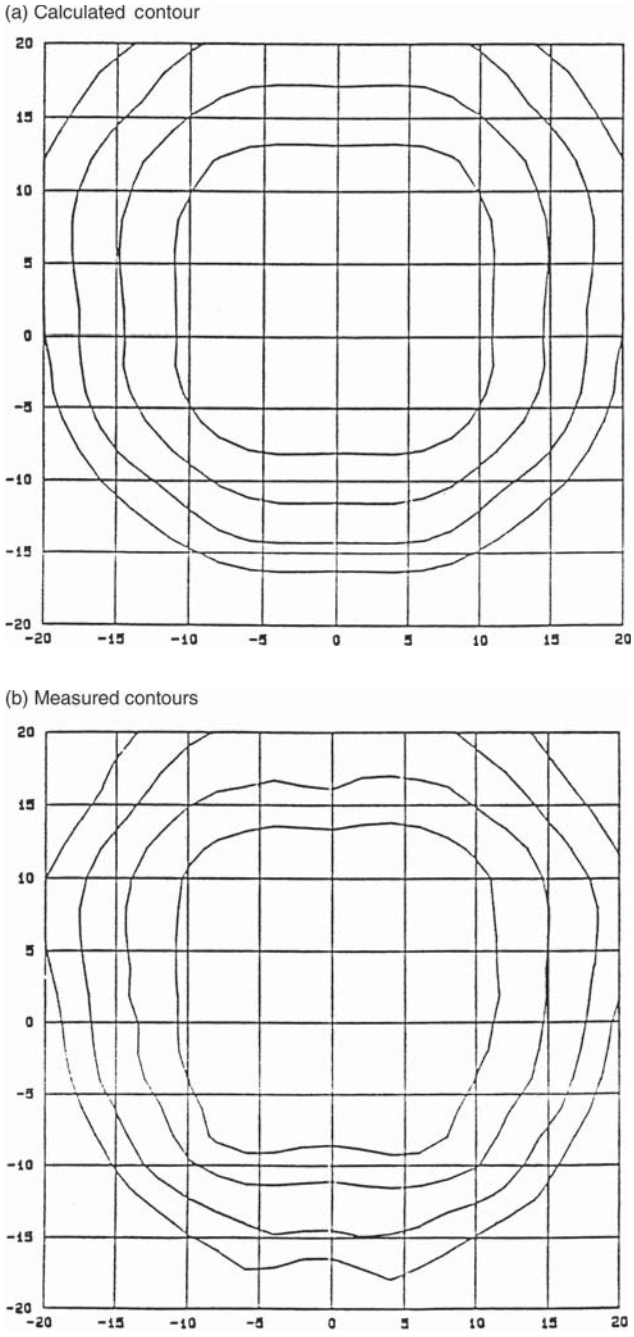


Figure 8-5 Quiet zone amplitude contours [12]

the same character as the calculated phase contours, and differences in their values are attributed to alignment uncertainty of the scanner's probe. The quiet zone characteristics are representative of results over the 21 to 25 GHz frequency range used in the testing. The quiet zone measurements did not discern any ripple in the test region from edge diffraction or illuminator sidelobes.

Another novel example of a simple compact range [15], illustrated in Fig. 8-6, uses a conventional prime focus reflector antenna mounted on the tower in this photograph and provides illumination of a test article mounted on the positioner. This compact range design was used to perform 18 GHz measurements. The test region was determined by examining the reflector's near field distribution to select a range separation from the reflector (47 ft) that conforms to the required amplitude and phase flatness. The near field contours illustrated in Fig. 8-7 were calculated

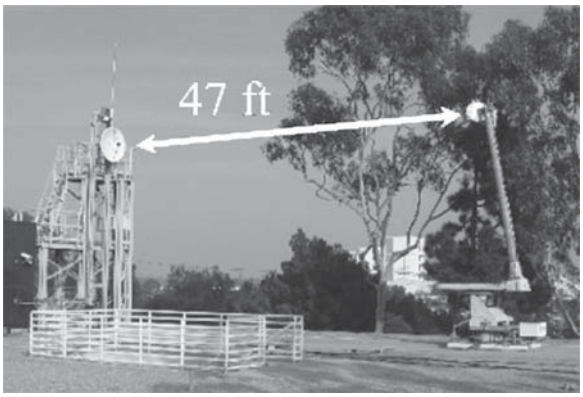


Figure 8-6 Quasi-compact range [15]

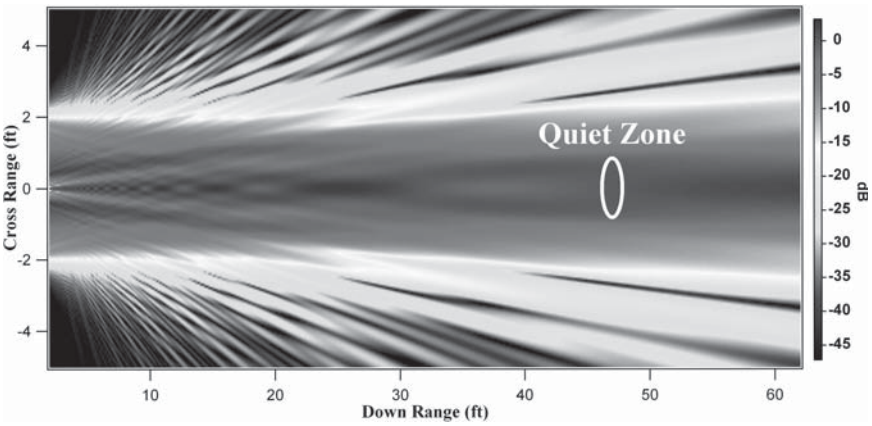


Figure 8-7 Near field power contours [15]

using a GRASP computer code. The distance from the reflector aperture is required so that the scattering from the feed blockage and supporting struts that propagate as a spherical wave is much lower than the scattering from the aperture's near field that propagates as a plane wave. For this reason, the design is referred to as a "quasi-compact range." As in other compact range designs, the 4-ft diameter reflector antenna is larger than the 20" quiet zone. The frequency variation of this design depends on the frequency dependence of the antenna's near field uniformity. Example field distributions in Fig. 8-8 illustrate the amplitude variation at 18 GHz. Further examination of the quiet zone fields over a 17 to 19 GHz frequency range reveal minor variations of this performance. In this example, measurements were performed at a range (47 ft) that is less than half the conventional far field distance.

Compact ranges are generally indoor facilities that have less expensive absorber requirements than far field anechoic chambers because the side walls, floor, and ceiling of the facility are far more weakly illuminated by the collimated field than the spherical wave used by far field ranges. These indoor facilities can be extensively used for development testing. The development of lightweight portable compact ranges would provide the capability to perform a more extensive evaluation of flight hardware in payload assembly areas.

Another novel compact range application is the evaluation of adaptive uplink antenna designs [16]. Adaptive antenna evaluations for

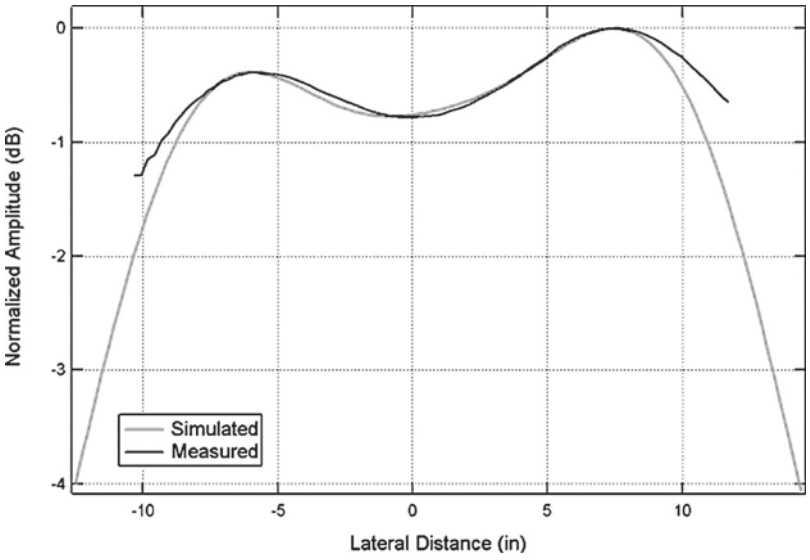


Figure 8-8 Measured and calculated transverse quiet zone fields [15]

applications such as uplink antennas require the capability to generate incident signals produced by independent signal sources from different directions to represent desired and interfering signals over the limited field of view from a geosynchronous altitude. Likewise, with design attention, a reflector antenna used in a compact range is also capable of producing scanned beams over a limited field of view. Thus, a compact range using a Gregorian subreflector and multiple illuminators with independent signal generators can generate high-fidelity plane waves arriving from different directions that are needed to produce desired and interfering signals for adaptive antenna measurements.

Compact ranges are generally indoor facilities that are commonly used in both development and qualification test phases for space segment antennas. The designs can be used over a wide frequency range. At low frequencies, the design is limited by physical size and the associated electrical size. The high frequency limitation is the mechanical surface tolerance like other reflector antennas. The development of portable designs would benefit testing in qualification in launch processing facilities.

Measurement uncertainty for compact ranges follows a similar error budget process described for far field ranges in Table 8-1. The illumination errors in this case can be assessed through measurements of the amplitude and phase perturbations in the quiet zone. Two kinds of perturbations result. Amplitude taper over the quiet zone results from the near field characteristics of the reflector, and amplitude and phase ripple result from residual edge diffraction and illuminator leakage components. As discussed, the facility reflections from the side walls, floor, and ceiling benefit from reduced illumination. Signal processing techniques provided by instrumentation and the effectiveness of the absorber for normally incident fields also contribute to reduced facility reflections. As discussed, the plane wave propagation of the collimated near field in contrast to the spherical wave radiation of the far field range typically results in higher power density levels in the quiet zone. Thus, noise limitations in the instrumentation occur at lower pattern levels in the measurement, reducing the measurement uncertainty in sidelobe regions so that the measurement's dynamic range is increased. Like far field ranges, absolute gain levels are commonly established using comparisons with a calibrated antenna, and errors for the standard include both calibration and measurement within the facility.

8.2.3 Near Field Sampling

Another approach to antenna testing samples the amplitude and phase characteristics over a surface surrounding the antenna under test and mathematically transforms these data to obtain the far field antenna

characteristics [17]. This process directly follows the Fourier transform relationship between the aperture fields and far field patterns (discussed in Chapter 1). The amplitude and phase values are measured with a probe antenna that is mechanically positioned by a scanner, as illustrated in Fig. 8-9. The amplitude and phase values sampled at discrete points are related to far field characteristics by a transform. Three types of sampling surfaces have been used. The most common surface is planar, and FFT techniques are used to transform the amplitude and phase samples to far field characteristics. Cylindrical sampling surfaces are also used, and cylindrical Bessel functions are employed in the transform, when spherical sampling surfaces are utilized, and spherical Bessel functions are used in the transform. In this case, the probe might remain fixed as the antenna is moved over a spherical surface by an antenna positioner.

In operation, the characteristics of the measurement probe [18] need to be established, because in the transform process the probe functions as a spatial filter. Commonly, an open-ended waveguide is used as a probe that has a linearly polarized response. In operation, when a linearly polarized probe is used, the surface is scanned twice with an orthogonal rotation to the probe between scans to obtain the total vector field over the scan surface. Absolute gain measurements [19] can be obtained through the probe calibration, both by comparison with a standard gain antenna and by using the three-antenna method. This reference also describes experience in measuring the ERP and transmitter saturation characteristics of the INTELSAT VI satellite.

This measurement technique has another important advantage, particularly for array antenna evaluations. The measured amplitude and phase data can be transformed not only to the far field but also back to the aperture plane of the antenna [20]. This is a significant benefit for array diagnostics because array excitation errors can be identified. The untransformed amplitude and phase data themselves have useful diagnostics. Aperture phase errors contained within these data, for example, can be used to refocus reflector antennas by identifying the quadratic phase error. For payload antennas requiring characterization over the earth's field of view, planar scanning techniques are often used. The probe

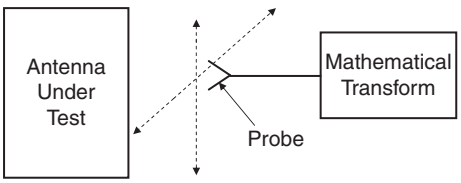


Figure 8-9 Near field sampling technique

cannot sample a plane of infinite extent. Typically, the sampled portion of the infinite plane for pattern measurements out to an angle θ_s from the aperture normal [21] equals the projected aperture area at the probe plane, increased by the projections of the aperture plane, subtended by the field of view, as shown in Fig. 8-10. Scanning over very wide fields of views generally requires spherical scanning techniques.

This measurement technique is commonly used in both development and qualification test phases, and is generally used to characterize antenna systems indoors. The usefulness of this technique could be increased for development testing if portable scanners were developed for use within payload assembly areas. This capability will increase in importance in future years as RF electronics continue to be integrated with antenna systems. A suitably portable scanner design would be particularly useful in launch processing areas.

A methodology to assess uncertainty of planar near field sampling measurements [21] has been developed and is commonly employed by many organizations. The terms in the error budget are listed in Fig. 8-11, and methods and examples of determining the individual error component values are described. The error sources include the probe calibration and errors incurred in the measurement process. Commonly, the facility errors are indicated by repeating the scan measurements with the probe and antenna under test moved by one quarter

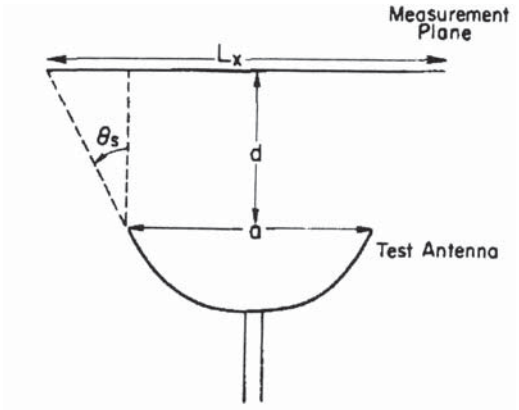


Figure 8-10 Planar near field sampling area [21] (© 1988 IEEE)

ERROR SOURCES IN PLANAR NEAR-FIELD MEASUREMENTS

Source of Error	Primary Methods of evaluating: Computer Simulation	Test on Measurement System	Error Equations
1) Probe relative pattern			x
2) Probe polarization ratio			x
3) Probe gain measurement			x
4) Probe alignment error			x
5) Normalization constant			x
6) Impedance mismatch factor			x
7) AUT alignment error		x	x
8) Data point spacing (aliasing)		x	x
9) Measurement area truncation	x		x
10) Probe x, y-position errors	x		x
11) Probe z-position errors	x		x
12) Multiple reflections (probe/AUT)		x	
13) Receiver amplitude nonlinearity	x	x	x
14) System phase error due to:		x	x
Receiver phase errors			
Flexing cables/rotary joints			
Temperature effects			
15) Receiver dynamic range		x	
16) Room scattering		x	
17) Leakage and crosstalk		x	x
18) Random errors in amplitude/phase	x	x	x

Figure 8-11 Error budget terms of near field scanning measurements [21] (© 1988 IEEE)

of a wavelength. Similarly, repeating the scan measurements with the separation between the antenna and the probe moved by one quarter of a wavelength provides a means to evaluate interactions between the probe and the antenna under test. Further details on random errors in near field measurements [22] are also described. The analyses have been recently extended [23] to address measurement uncertainty in spherical near field measurements.

8.2.4 Instrumentation and Antenna Gain Standards

Instrumentation and supporting software for general-purpose antenna measurements are well developed. Typically, network analyzer instrumentation is used for antenna pattern and terminal impedance measurements. Today's network analyzers provide a highly linear response over a wide frequency range and dynamic range, as is required for antenna measurement. Two types of network analyzers can be distinguished. The first type is a scalar network analyzer that permits amplitude measurements. The second type is a vector network analyzer capability of performing amplitude and phase measurements over a much wider dynamic range. The dynamic range is limited at the high signal levels by compression and at low signal levels by the received SNR (signal-to-noise ratio). Generally, compression is a problem only when additional amplification is added to the network analyzer to increase the dynamic range.

At low signal levels, the error results from noise and has an rms (root mean square) value equal to $1/(2SNR)^{1/2}$. Commonly, averaging techniques can also be applied to reduce the thermal noise error variations. The variance of the error is reduced by n, the number of samples averaged, so that the rms value is reduced by $1/n^{1/2}$.

Vector network analyzers that provide amplitude and phase measurement capabilities can gather information over a bandwidth. Fundamental mixing techniques [23] are used to maintain sensitivity over very wide bandwidths. A capability internal to vector network analyzers exists to transform the measured frequency domain data into the time domain. Once in the time domain, responses distributed in time can be associated with the antenna response and the facility response. Windowing techniques can be applied to isolate the antenna response from the facility response. The windowed time domain response can be transformed back into the frequency domain. In this way, the antenna's frequency response in the absence of those facility contributions that can be windowed out is obtained, improving measurement accuracy. Another alternative is time gating with low-reflection diode switches.

Vector network analyzers can also be used to measure the antenna's impedance values. In this case, impedance values are obtained by calibrating the network analyzer interface with short and open circuit reference loads. Typically, antenna impedance measurements are also windowed in the time domain to reduce facility reflection contributions. Network analyzer measurements of the mutual coupling between antenna systems or array elements [24] can also be performed, and benefit from windowing techniques.

The technology trends of future antenna designs will require more general instrumentation capabilities. One example is active receive and/or transmit antennas using monolithic fabrication incorporating frequency conversion. In such cases, RF antenna terminals are unavailable and the instrumentation interface is an IF frequency. Thus, one desirable extension for future instrumentation is the ability to sample the oscillators used in the design's frequency conversion to obtain a frequency reference. This frequency reference could then be used with separate frequency converters to obtain a coherent reference for the network analyzer. Such a capability would capitalize on the existing benefits of network analyzer instrumentation. A more challenging example is future receive and/or transmit antennas that use digital beamforming techniques. In this case, more generalized instrumentation and testing using specialized waveforms appropriate to the system application are needed. One approach is to use D/A technology for receive antennas and A/D technology for transmit antennas to provide an analog interface for measurement purposes.

System development testing also requires other instrumentation such as spectrum analyzers and BER test sets. For example, spectrum analyzer instrumentation is required to address the presence of intermodulation products or to evaluate the adequacy of diplexing filtering. BER (bit error rate) test sets are mandatory in evaluating adaptive antenna designs to establish threshold SNIR levels. Satellite qualification testing greatly expands the instrumentation requirements. These requirements result in an elaborate test set that is collectively known as AGE (aerospace ground equipment). The test set commonly incorporates modulation modems used by the system users and performs end-to-end tests, typically using BER measurements. The extent of the AGE equipment requires computer control of the instrumentation and additional computer resources to gather and store the data.

When absolute gain measurements are performed using comparative techniques, the calibration of the antenna standard becomes an important issue. Commonly, horn antennas are used as a gain standard because their design can be analyzed by well-established techniques, and their calibration methodologies are well known. The antenna gain values are generally established by using a three-antenna method [1]. In this method, the power transfer is measured for alternative pairs of three antennas, one of which is the antenna under test. The separation between the phase centers of each pair being measured is also determined. The three measured received power levels and the spatial separation distance result in three equations and three unknowns, namely the gain levels of the three antennas. In this way, the gain levels of each antenna are established. Often, two of the antennas used in these measurements have previous calibration histories, and agreement of the measured results with previous results provides confidence in the measurement. This technique can also be extended to determine the polarization characterization [24, 25] of antennas. Experience with the calibration of standard gain horns [26] indicates that uncertainties on the order of 0.3 dB can be achieved. Swept frequency measurements are recommended because the antenna's gain value has a slight ripple, resulting from the interactions between the horn's input waveguide and the aperture reflections. This reference also describes the calibration of higher gain antennas required in some applications.

Care must be taken in using standard gain horn antennas because of their broad beamwidths and high sidelobe levels in the E-plane. Such antennas [27] can interact with reflecting surfaces located close to them. These reflections result in pattern ripple and changes in the gain value compared to the free space environment in which the antenna was calibrated. Rolled edges (as described in Chapter 2) can be added to the E-plane aperture edges to reduce those sidelobe levels. Measurements of a horn with the rolled edges in the same reflection environment as the

horn without the rolled edges illustrate the reduced sensitivity to the surrounding environment when the E-plane sidelobes are reduced. However, the horn with the rolled edges is more directive than a standard gain horn, and thus its absolute gain level must be recalibrated.

8.3 Radio Source Techniques

Antenna measurement techniques based on using radio sources provide a means of evaluating a wide class of user segment antenna designs. The basic radio source measurement yields the G/T of the antenna under test. These techniques were developed to measure large ground terminals [28, 29, 30, 31] because of their large far field distances and practical logistical limitations of measuring large antennas in a conventional test facility. The far field requirements for larger antennas at higher frequencies become excessive, as shown in Fig. 8-2. Radio source measurements technique can also be used to determine the antenna's gain in the transmit band by substituting a preamplifier for the transmitter and measuring the G/T and the system noise temperature, T, to obtain the transmit antenna gain, G. The wide availability of cost-effective low-noise amplifiers makes radio source measurements practical.

Radio source measurements are based on the known flux density values of radio stars. The largest radio source is the sun, whose flux density depends on frequency and varies with solar activity. The solar flux density is measured daily by NOAA. The differences between the quiet sun and active sun flux density values are illustrated in Fig. 8-12. The sun, however, subtends a 0.5° width, and when the antenna's beamwidth is smaller than that value, the sun no longer approximates a point source. The finite extent of the sun has been treated by what is referred to as a "beamfilling factor," which assumes the flux density is uniformly distributed over the solar disk. However, because of solar flare activity, the flux density distribution over the solar disk is decidedly non-uniform, particularly when the sun is in an active period and consistent measurements are hard to achieve when the antenna's beamwidth is less than 0.5°. Fortunately, other radio sources are available, and commonly Cassiopeia A, Cygnus, and Taurus are used, albeit these sources have a lower flux density than the sun.

The basic G/T measurement proceeds by aligning the antenna with the source and measuring the received noise power. This noise power measurement consists of the incoherent sum of the radio source's flux density, the cold sky background, and the system noise power. The antenna is then rotated away from the source, maintaining a constant elevation angle, and a second noise power measurement is made, having only the cold sky noise power and the system noise power. The ratio of these two noise powers—"hot" to "cold"—is referred to as a Y factor.

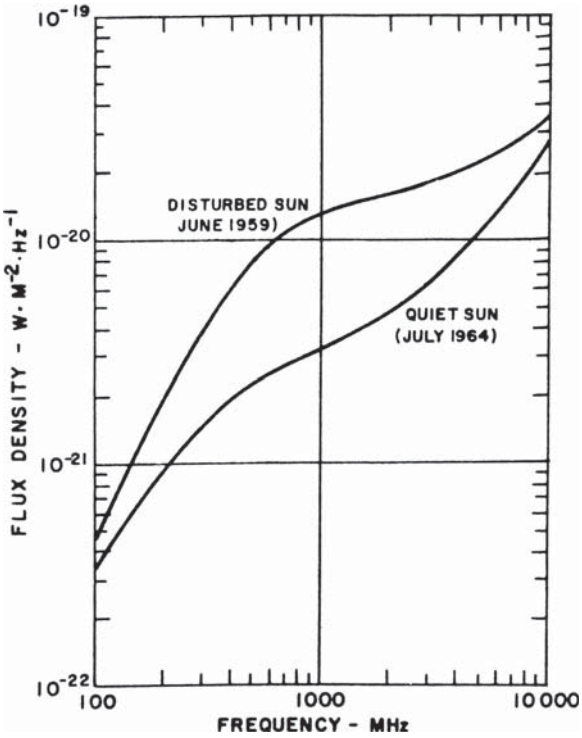


Figure 8-12 Solar flux values [30] (© 1971 IEEE)

These noise measurements are made at an elevation angle in the radio source's direction and allow the calculation of the G/T at the elevation angle of the radio source. A consistent elevation between hot and cold measurements is required so that both measurements have comparable path loss values and, hence, antenna temperature values.

The antenna noise temperature is a function of elevation angle (as discussed in Chapter 1). The specification of a G/T value requires stipulating a reference elevation angle value so that the antenna temperature can be properly defined. A typical reference elevation angle is 20°, a value that is high enough so that local obscura, or terrain blockage, does not contribute to the antenna temperature. This elevation angle, as opposed to a lower value, results in a specification that is not strongly dependent on the specific site's local terrain. The G/T can be determined for the reference elevation angle by performing a third cold sky background measurement at that reference elevation angle value. The difference in the two background noise values directly measures the effect of the antenna noise temperature's elevation angle dependence on

the system noise temperature T. The ratio of the two noise values thus corrects the G/T at the measurement elevation angle to obtain the G/T at the specified elevation angle.

The G/T can be determined from these measured noise powers and the source's flux density. The noise power referenced to the LNA input terminal when pointed at the source equals

$$P_{n1} = k(T_{\text{rs}} + T)B$$

where k is Boltzmann's constant (–198.6 dBm/Hz/K), T_{rs} is the equivalent noise temperature of the radio source, T is the total system noise temperature at the elevation angle of the source, and B is the RF bandwidth. The noise is the incoherent sum of the cold sky background noise and the noise from the radio source. The noise power received from a radio source whose flux density is S equals one-half times the flux density, multiplied by the effective aperture of the antenna, and multiplied by the noise bandwidth, or

$$P_{\text{rs}} = \tfrac{1}{2}S(\lambda^2/4\pi)GB = kT_{\text{rs}}B$$

where λ is the wavelength and G is the antenna's gain value. The reported flux density values equal the total flux density of the source that is randomly polarized. The antenna receives one-half of the total flux density because the antenna has a single polarization state. If an orthogonally polarized antenna could also receive the flux, the remaining half of the flux density of the radio source would be received.

The noise power referenced to the LNA input when pointed at the cold sky at the elevation angle of the radio source equals

$$P_{n2} = k(T)B$$

The Y factor for G/T measurements is defined as the ratio of the noise power when the antenna is pointed at the source and the noise power when the antenna is pointed at the cold sky background (i.e., P_{n1}/P_{n2}). Combining these equations yields

$$G/T = 8\pi k(Y - 1)/(S\lambda^2)$$

At higher EHF frequencies, the propagation loss, and hence the antenna noise temperature, has a significant dependence on weather conditions. Care must be exercised to perform the measurements on a "clear" day, but sometimes, clear is a subjective measure. Obviously, EHF radio source measurements should not be performed during rain, but cloud cover differences impact the path attenuation and hence the antenna temperature. An independent radiometric measurement is one approach

to defining the path loss. The system specification in such cases must stipulate not only a reference elevation angle but also path loss conditions if the specification is stated by G/T values. The alternative is to specify the antenna gain value and a measurement of the antenna temperature at the same time the G/T is measured. Subsequent measurements of the antenna noise temperature under varying weather conditions can yield a better understanding of the weather-induced G/T impacts. These considerations will become increasingly important as the receiver noise temperature performance continues to decrease in future designs.

The radio source technique is widely used for large ground antennas. What is not commonly appreciated is that the technique can be used with much smaller antenna sizes, a consequence of the low-noise receiver technology used in today's systems. Such antennas use the sun as a radio source and their beamwidth generally exceeds the 0.5° width of the solar disk. An example [32] assumes the antenna efficiency is 55% and the system temperature is 200K. The flux density of the sun under quiet conditions is assumed, as well as a measurement Y factor of 0.75 dB. These assumptions are used to illustrate in Fig. 8-13 that small antenna sizes can be successfully measured using radio sources.

The G/T of a wide class of user antenna designs can be measured using radio source techniques. Small antennas can be measured using solar radio source techniques. A variety of other radio sources can be used for much larger antennas. Further discussion of measuring antennas between these two extremes may be found in Chapter 9. Because radio sources are generally randomly polarized (some sources are partially polarized), the polarization of the antenna under test cannot be characterized. Radio source techniques will have increased future application as the trend of integrating user antennas with electronics continues. Such integration results in a measurement problem because the antenna and electronics cannot be separated and a terminal to measure the antenna by itself is not available. Further discussion of this problem is provided in Section 8.5.

Radio source measurements depend on measuring the noise power differences. The presence of any interfering signal invalidates the measurement, a factor that is particularly significant at lower microwave frequencies because of spectral congestion. Interference power results in inconsistent noise power measurements. For user antennas, interference generally arrives through the antenna's sidelobes and thus the received interference power varies as the antenna is repositioned to make the hot and cold noise power measurements. Repositioning the antenna at the cold noise power positioning and observing if the values are stable is one means of indicating the presence of interference. Small changes in position (e.g., 1°) have little effect on the antenna noise temperature.

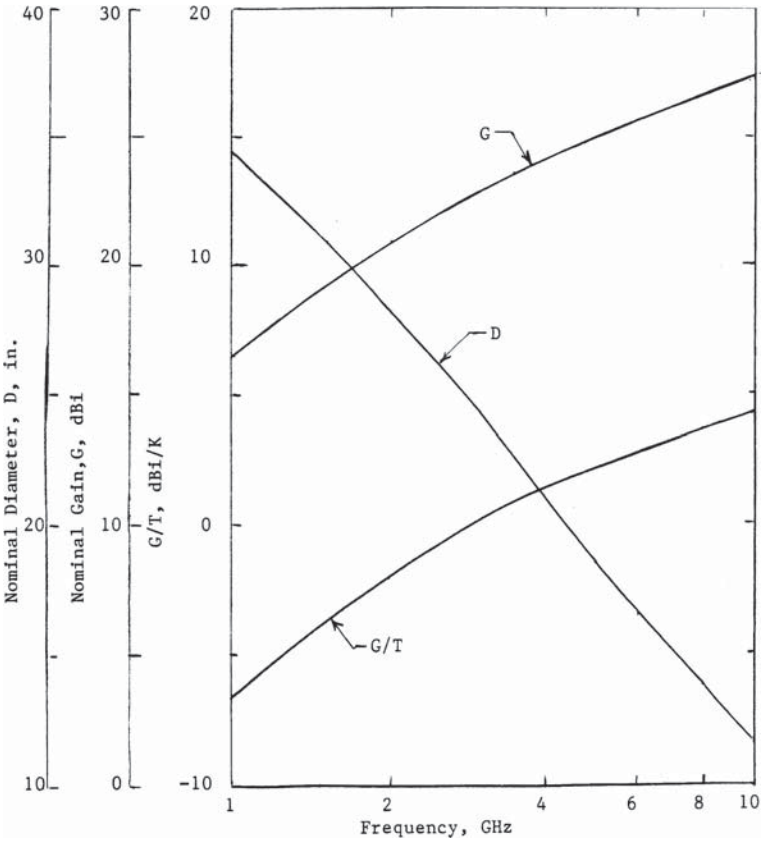


Figure 8-13 G/T measurement of small antennas [34]

If the measurement is free of interference, the cold noise temperatures should not vary. If interference is present, the received interference power contributions to the power measurements vary because of different sidelobe antenna gain level variations at the two angular positions. This same process is followed to determine if the antenna has been rotated adequately away for the radio source so that the source's flux density received by the antenna's sidelobes does not contribute to the cold noise power measurement. Radio sources move as time passes. If interference is present in the hot noise power measurement, the noise power values will again change as the antenna's position varies. Similarly, the measured noise power values should remain stable for small changes in the measurement frequency. This process examining the stability of the noise power measurements when the antenna's angular position and the measurement frequency are varied provides

confidence in establishing the presence of interfering signals that invalidate the measurement.

The recommended process for radio source measurements is to repeat the measurements three times in succession and examine the noise power values. Variations in the values indicate interference is present in the measurements. A second cause of variations in the hot noise measurements is antenna misalignment with the radio source. Thus, repeating the measurements three times in succession and achieving consistent results is a necessary condition for correct measurements.

The equipment required for this measurement is relatively modest. The antenna must have the means to point at the positions required for the measurements and since the sources move, tracking and verification of the boresight alignment is necessary during the course of the measurements. Ground terminal antennas typically have star tracking capabilities in their antenna control units to automate the process of radio source measurements. The noise power can be measured in several ways. A very simple low-cost measurement can be made if the receiver's AGC voltage is available. The AGC linearity over the different noise levels should be verified. The AGC output also responds to interference and measures an average G/T over the receiver's bandwidth. For wide bandwidth systems, G/T variations at discrete frequencies may be required. Power meters and narrow bandpass filters at an IF level are an alternative means to measure the noise power values, and again their linearity needs to be evaluated. More commonly, radio source G/T measurements are performed by a spectrum analyzer. The high noise level of the spectrum analyzer requires care to assure the spectrum analyzer noise does not increase the system temperature. Additional preamplification may be needed before the spectrum analyzer to avoid degrading the system noise temperature. The spectrum analyzer has two distinct advantages. The presence of interfering signals is clearly indicated by frequency variations of the noise power displayed by the spectrum analyzer and can also be observed by the stability of the noise measurements for small perturbations of the antenna's angular position. In some cases, the spectrum analyzer is useful in identifying a portion of the received spectrum that appears to be interference-free. The second advantage is that the resolution bandwidth of the spectrum analyzer provides G/T measurements at narrow portions of the operating bandwidth (e.g., the center and end frequencies of the operating bandwidth). This capability allows evaluation of the G/T variations over the operating bandwidth. The dynamic range and averaging features in spectrum analyzers are also useful in performing the measurements.

A simple way exists to determine the impact of the spectrum analyzer noise on the total system temperature. The system noise temperature

with the spectrum analyzer inserted, T', follows the cascaded noise temperature discussion in Chapter 1 and equals

$$T' = T_{ant} + T_{rec} + T_{sa}/G'$$

where T_{ant} is the antenna temperature, T_{rec} is the noise temperature of the receiver up to the point at which the spectrum analyzer is used, T_{sa} is the noise temperature of the spectrum analyzer, and G' is the electronics gain of the receiver between the LNA input and the spectrum analyzer input terminal. The values can be calculated from the system parameters and the spectrum analyzer's noise power to estimate if preamplification is needed. The noise power of spectrum analyzers is typically specified by the average noise power in a 1 KHz bandwidth when terminated in a matched load, which can be converted to the value of the spectrum analyzer temperature T_{sa}. The measurement to determine the spectrum analyzer's contribution to the system noise temperature is performed by measuring two noise levels. One noise level is the indicated spectrum analyzer noise level when the analyzer is connected to the system under test and has an indicated level of $G'T' B$, where B is the noise bandwidth since the electronics gain G' amplifies the front-end noise. The second noise level is the indicated spectrum analyzer noise when the spectrum analyzer is terminated in a matched load and equals $T_{sa}B$. The ratio of these noise measurements, Y_{sa}, equals

$$Y_{sa} = G'T'/T_{sa}$$

This equation, solved to yield the spectrum analyzer's contribution to the system temperature, yields

$$T' = (T_{ant} + T_{rec})(1 + 1/(Y_{sa} - 1))$$

The additional term in the system temperature T' is the noise contributed by the spectrum analyzer and an appropriate correction can be applied to the G/T. The values in Fig. 8-14 indicate that a 10 dB Y_{sa} ratio results in about a 0.5 dB reduction of G/T, and if Y_{sa} exceeds about 17 dB, the G/T reduction is less than 0.1 dB.

8.3.1 Noise Temperature Measurements

If G/T measurements are performed and the antenna's gain value is required, the system noise temperature must be measured. Similarly, if the antenna's gain value is determined by other means and the G/T is required, the system noise temperature must be measured. The system noise temperature is comprised of two components, the antenna temperature and the receiver temperature. These two noise components

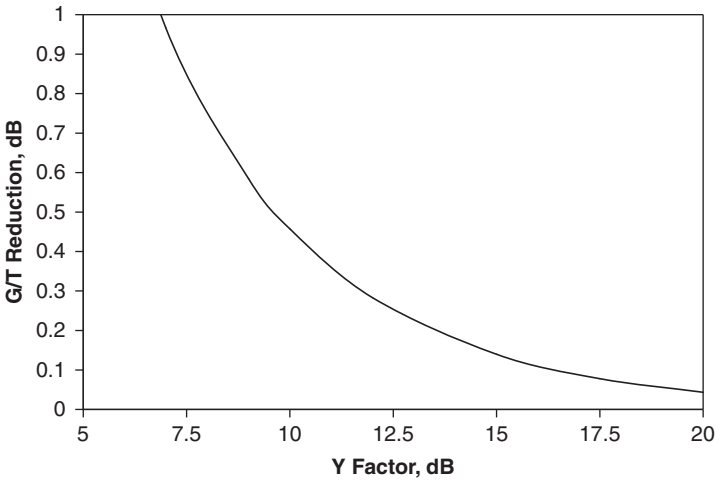


Figure 8-14 Correction for spectrum analyzer noise [33]

must be referenced to the same terminal plane along with the antenna gain. Often, a convenient location for this terminal plane is the input to the low-noise amplifier for ease of measurement. Two ways of proceeding with measuring the system noise temperature are described.

One way is to measure the receiver and antenna temperature values separately at the input to the low-noise amplifier. Conventional noise measurements are used to measure the receiver temperature. Two alternatives exist to measure the receiver noise temperature. One alternative is to use commercial noise figure measurement instrumentation, where a cascaded arrangement is used to measure the noise figure and gain of the amplifier. The noise figure is then converted to a noise temperature value, as described in Chapter 1. A second alternative is to use a hot/cold load measurement. The LNA is terminated with a matched load and the output noise power is measured. This measurement terminates the LNA at an ambient (≈ 290 K) temperature. A second measurement is made with the terminating load immersed in liquid nitrogen. This measurement terminates the LNA at a 77 K temperature. The ratio of these two noise power measurements, the terminating load at an ambient temperature, T_a, and the terminating load immersed in liquid nitrogen, T_c, forms a Y_R factor. The receiver temperature T_{rec} is calculated from

$$T_{\text{rec}} = (T_a - T_c Y_R)/(Y_R - 1)$$

The antenna noise temperature is also measured at the same terminals that the receiver noise temperature is measured. Since the antenna

temperature at microwave frequencies is typically lower than the ambient temperature, a simple measurement can be performed to obtain the antenna temperature. The antenna is pointed at the elevation angle where the G/T is measured (or at the specified elevation angle for the system when the antenna temperature is to be measured separately) and the noise power at the receiver output is again measured. The antenna is then disconnected and replaced by a matched load. The noise power is again measured to obtain Y_{ant}. The antenna temperature T_{ant} is obtained from

$$T_{\text{ant}} = T_a/Y_{\text{ant}} + T_{\text{rec}}(1 - Y_{\text{ant}})/Y_{\text{ant}}$$

If the antenna noise temperature variation with elevation angle is desired, this process can be repeated with the antenna pointed at different elevation angles. As shown in Fig. 1-14, the greatest variation in antenna noise temperature occurs at low elevation angles, thus the elevation angle sampling density should be increased at low elevation angles.

A second way to measure the system temperature uses a noise source to inject a hot temperature reference into the receiver by a coupler. In some designs, this coupler is also used to inject test signals for BITE capabilities, and this same coupler can be used to inject a noise reference from a noise diode. Noise diodes are commonly used as a noise source and specified by their ENR (excess noise ratio). The coupling coefficient of the coupler used to inject the noise source must be carefully measured. Noise diodes produce noise over extremely broad bandwidths, but are not well matched at all frequencies. Typically, several noise diodes are available, and selection of one that is well matched at the calibration frequencies is recommended. Commercial services are also available to calibrate noise diodes at specific frequencies, and calibration of the noise diode and the coupler can typically be done to reduce measurement uncertainty.

The antenna noise temperature is measured by pointing the antenna at a given elevation angle and measuring the noise power at the receiver's output with and without an active noise source. The antenna elevation angle where the G/T was measured can be used if the antenna gain value is required. Alternatively, the antenna can be pointed at the specified elevation angle for the system when the system temperature is to be measured separately. A range of elevation angles can also be used to determine the system noise temperature variation with the elevation angle. Another Y factor, Y_s, is obtained from the ratio of these two noise power measurements. The system temperature, T_s, equals

$$T_s = \Delta T/(Y_s - 1)$$

where ΔT is the equivalent noise temperature injected into the receiver at the output of the coupler. This technique is commonly referred to as

the noise added method. It requires careful calibration of the coupler and compensation for the effects of coupler loss on the system temperatures. This technique is particularly useful when the antenna temperature is close to the ambient 290 K, where Y_{ant} is close to 1 (0 dB).

8.3.2 Radio Source Measurement Uncertainty

Radio source measurements rely on measuring only the radio source's and terminal's noise power levels. Other interfering signal sources invalidate the measurements. Extreme care is required to insure interference is not in the measurement. An initial examination of the received spectrum with a spectrum analyzer should be made to insure interfering signals are not present within the bandwidth to be used for the radio source measurement and that high-level sources are not present, possibly resulting in nonlinear receiver operation that suppresses the noise power levels. The stability of the noise power levels at the cold sky and hot sky positions should be observed for small changes in frequency and positioner angles as described earlier. Once assured the measurements are not corrupted by interference signals, the measurement's error sources can be examined.

Three distinct error sources degrade radio source measurement accuracy. The first factor is the accuracy with which the antenna is aligned with the source. The antenna pointing loss in terms of misalignment normalized to the antenna's beamwidth was previously given in Fig. 2-10. Offsetting the antenna's angular position, the following step track procedures can be applied to radio source measurements to verify the antenna's alignment if sufficiently high Y factor values exist. In this way, antenna tracking alignment with the radio source has an uncertainty of about 1/10 of a beamwidth, corresponding to a 0.1 dB antenna pointing uncertainty.

A second error budget component is uncertainty in the radio source's flux density value. Because G/T is inversely proportional to the flux density value, the G/T error is inversely proportional to the flux density error. The G/T error resulting from the flux density error equals

$$\varepsilon_S = S/S'$$

where S is the correct flux density value and S' is the flux density value in error.

The G/T uncertainty also depends on the uncertainty in the Y factor measurement. The G/T value is proportional to $Y - 1$, so that the G/T error resulting from Y factor measurement errors equals

$$\varepsilon_Y = (Y' - 1)/(Y - 1)$$

where Y' is the Y factor in error and Y is the correct Y factor value. The variation of the error with Y factor values is illustrated in Fig. 8-15 when

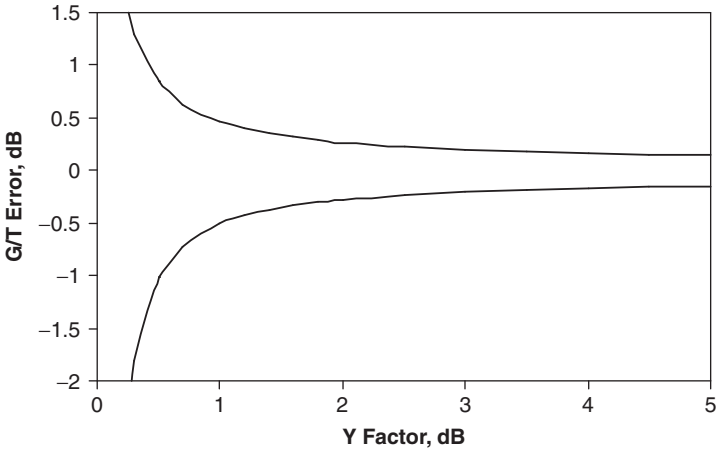


Figure 8-15 *G*/*T* errors for a 0.1 dB *Y* factor error

the *Y* factor error is ±0.1 dB. These errors asymptotically approach the *Y* factor error value for large *Y* factors but are minimal for reasonable *Y* factor values. For example, the *G*/*T* error for a 2 dB *Y* factor value is about 0.27 dB, which drops to a 0.12 dB value when the *Y* factor is about 7 dB.

If the antenna's gain value is to be determined from *G*/*T* measurements, the system noise temperature must also be measured to determine the antenna's gain value. Likewise, if the antenna's gain value is determined by other means and *G*/*T* is required, the system noise temperature must also be determined. The errors in the system noise temperature measurements need to be addressed in both cases. The receiver noise temperature is separately determined from either commercial noise figure meter measurements or from hot/cold load measurements. Another *Y* factor measurement is used to determine the antenna noise temperature. The receiver and antenna noise temperature values are then summed to obtain the total system noise temperature. A consistent set of terminals is required to quantify the noise temperature measurements, and the antenna gain value burdened by any system loss between its measurement terminal and the reference terminal used for the noise measurements must be referenced to that same terminal. Two error components must be addressed in determining the uncertainty in the system noise temperature: the receiver noise temperature measurement error and the *Y* factor measurement error for the antenna noise temperature.

The error in the receiver noise temperature measurement result in an error system noise temperature can be determined by noting the

system noise temperature. The system noise temperature error value equals

$$\varepsilon_{\text{tr}} = (T_A + T_R')/(T_A + T_R)$$

where T_A is the antenna noise temperature, T_R' is the receiver noise temperature value in error, and T_R is the correct receiver noise temperature. Numerical values of the system noise temperature measurement error in Fig. 8-16 assume a 50 K antenna noise temperature value and a 0.1 dB error in the receiver noise figure value. The system noise temperature error increases as the noise figure decreases.

The second error source in the system noise temperature measurement is a Y factor measurement error in determining the antenna noise temperature. The Y factor in this case is denoted by Y_A to distinguish it from the Y factor used in the G/T measurement. The measurement of Y_A is formed by the ratio of the noise power when the receiver is terminated in a matched load and the noise power when the receiver is connected to the antenna that is pointed at the elevation angle of the radio source where the G/T is measured. The Y_A factor equals

$$Y_A = (290 + T_R)/(T_A + T_R)$$

The antenna noise temperature is computed from

$$T_A = (290 + (1 - Y_A)\, T_R)/Y_A$$

When the Y_A factor is in error, the antenna noise temperature value calculated from the preceding relationship is also in error. An example

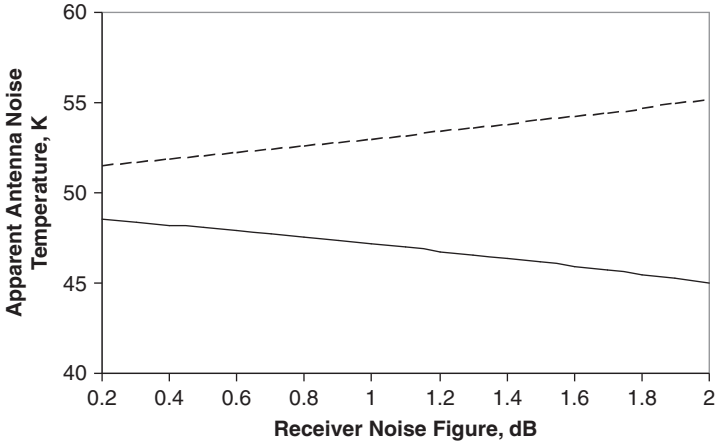


Figure 8-16 Antenna noise temperature measurement error for a 0.1 dB receiver noise temperature error

of the antenna noise temperature error in Fig. 8-17 assumes the true value of the antenna noise temperature is 50 K and the Y_A factor error is 0.1 K where the solid curve corresponds to a positive error and the dashed curve corresponds to a negative error.

The system noise temperature error resulting from an error in Y_A can be derived from the preceding expressions. The antenna noise temperature computed by using the Y_A' factor that is in error equals

$$T_A' = (290 + (1 - Y_A')\, T_R)/Y_A'$$

The system noise temperature error that results equals

$$\begin{aligned} \varepsilon_{Ya} &= (T_A' + T_R)/(T_A + T_R) \\ &= (1/Y_A')(290 + T_R)/(T_A + T_R) \\ &= (Y_A/Y_A') \end{aligned}$$

The system noise temperature error resulting from an error in the Y_A value is inversely proportional to that error. Example system noise temperature error values in Fig. 8-17 assume a 50 K antenna noise temperature and a 0.1 dB error in the Y_A factor measurement. Positive errors correspond to the solid curve while negative errors correspond to the dashed curve.

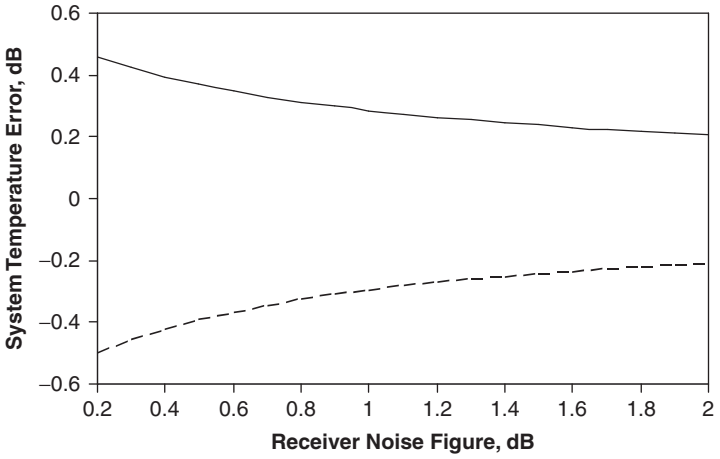


Figure 8-17 System noise temperature error resulting from a 0.1 dB error in Y_A factor measurement

8.4 Adaptive Antenna Evaluation

The evaluation of adaptive antenna systems [16, 35, 36] extends the scope of conventional antenna testing to system-level testing to quantify the effectiveness of adaptive interference rejection. Such testing evaluates the design integration with system electronics, the control and weighting circuitry, and system performance measures. The quiescent performance of the antenna in interference-free conditions is initially established using conventional antenna test techniques to characterize the coverage, polarization, bandwidth, and G/T performance. Additional bench testing is performed to evaluate the performance of the adaptive electronic components of the design. The evaluation of the adaptive operation of the system is then initiated and quantifies the steady state performance when interference is present and the transient performance that measures the time required for the system to adapt to the interference and quantifies the disruption of user communications when disrupted by interference. Adaptive antenna evaluations therefore address two system-level issues:

1. What is the steady state loss in performance when interference is present?
2. How long does it take to reach steady state conditions after interference initiation?

The requirements of the adaptive antenna design are stated in terms of scenarios that describe the interference environment. These requirements form the basis of both the design and testing. Adaptive antenna development proceeds by constructing a detailed simulation of the proposed design and exercising the simulation to evaluate design compliance. As the development proceeds, the simulation is augmented by measured component performance of the adaptive design elements to improve the simulation's fidelity. The simulation is exercised in a Monte Carlo sense following the possible variations defined by the scenario to obtain the required statistical answers to adaptive system performance. While the Monte Carlo approach is necessary to provide the statistical answers to adaptive system performance, a Monte Carlo approach to adaptive system testing has impractical schedule requirements. The simulation is used to define a limited number of test cases, and adaptive antenna testing is conducted as defined by these test cases. The test results are then compared with the simulation results, and their agreement validates the simulation. Once the simulation has been validated through the test case agreement, the simulation is then exercised in a Monte Carlo manner to address system compliance with requirements. Adaptive antenna testing is therefore a significant departure from the

normal antenna testing and compliance determination used in conventional antenna testing.

Adaptive antenna testing is contrasted with conventional antenna testing in Fig. 8-18. Conventional antenna testing is well supported by standards, established facilities and instrumentation, and analytic computer codes that allow independent verification of measured results. The test results for conventional testing address the component-level parameters that define the antenna's performance. By contrast, adaptive antenna evaluations are guided by an interference scenario that is specific to the program. The detailed simulation of the adaptive system design and the Monte Carlo treatment of simulation results are used to select test cases for adaptive system evaluations. Comparison of the measured and simulated system performance is used to validate the simulation. By contrast, conventional antenna testing typically compares measurements with the calculated results from analysis codes to provide confidence in the measured results.

Adaptive facility requirements have more general requirements than conventional antenna testing since the desired signal collection must be generated along with interference signals, representing the collection of interference sources. The collection of desired and interference signals also is required to have different arrival directions. The signals must replicate operational and interference waveforms and

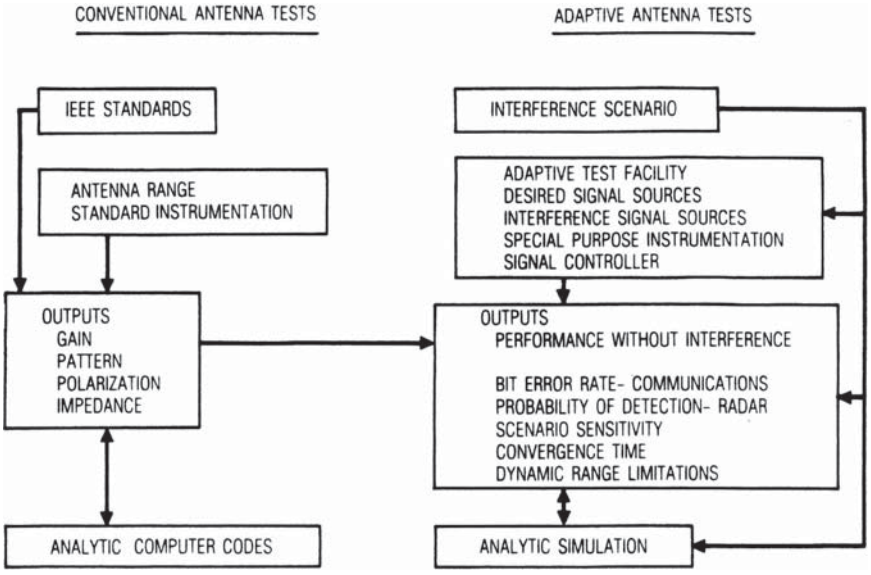


Figure 8-18 Comparison of conventional and adaptive antenna testing [35]

use specialized instrumentation such as BER test sets so that parameters such as *SNIR* can be evaluated. A variety of interference signal spectra can be postulated and performance measures relative to the interference spectra must be evaluated. Thus, the test signals used in adaptive system evaluation are clearly more diverse and general than the swept frequency measurements used with network analyzers that are employed in conventional antenna testing. The levels of these test signals must also be varied so that the performance for different signal and interference levels can be evaluated. Clearly, computer control is needed for the signals and instrumentation, and storage and display of the data are required. Finally, the outputs of the adaptive system evaluations are system-level measures of performance rather than the component-level parameters used in conventional antenna testing.

Uplink adaptive antennas provide interference protection over a limited field of view subtended by the earth. The performance of these designs [16, 35] must be established for both quiescent and adaptive cancellation conditions. The quiescent performance of the antenna provides service over the design coverage area when interference is not present and the testing follows conventional antenna measurements. When interference is initiated, the adaptive cancellation responds by dynamically forming pattern nulls in the direction of interfering sources. The resulting pattern nulls impact the performance available to users within the design coverage area. A threshold value of the acceptable *SNIR* value must be established to define user requirements for communication. The performance measure for adaptive system effectiveness of uplink antennas is the percent coverage area. The percent area coverage is defined by the amount of the design coverage area where the threshold *SNIR* is exceeded after adaptive cancellation, divided by the design coverage area where communication service is provided under quiescent conditions. The percent area coverage varies with the precise location of interference sources. Cumulative statistics gathered as scenario parameters and treated on a Monte Carlo basis are used to measure adaptive cancellation effectiveness. The transient performance is established by measuring the time required after interference initiation for the adaptive weights to converge to their steady state values.

Adaptive antennas for user applications generally follow the sidelobe canceller architecture (described in Chapter 7). Interference arrives through the antenna sidelobes while the main beam is directed towards the satellite. The sidelobe canceller design purposely constrains the auxiliary antenna collection to avoid canceling the main antenna's main beam. The auxiliary antennas are configured so that their antenna gain exceeds the sidelobe gain of the main antenna so that noise contributions from the auxiliary antennas when adaptive cancellation is exercised have a minimal impact on the system noise temperature.

The primary objectives of adaptive evaluations of these designs are to determine residual interference levels and changes in the system noise temperature for steady state operation and the transient performance of the adaptive circuitry. Again, a threshold *SNIR* value is used as a means of evaluating the effectiveness of sidelobe canceller designs.

Facility requirements for adaptive testing differ from conventional testing. General-purpose antenna test facilities are based on a single illumination source that produces fields of sufficient fidelity in their quiet zones to permit antenna measurements. By contrast, adaptive system evaluations require both desired signal and interference signal components from differing directions to illuminate the adaptive antenna under test. Space segment adaptive antennas for uplink antenna applications receive signals and interference from the limited field of view subtended by the earth. Extensions of compact range technology described earlier have been found to be an attractive way to generate the desired and interfering signals over a limited field of view.

User antennas with adaptive cancellation capabilities are generally impacted by terrestrial interference. Testing adaptive user antennas requires setting up a collection of interference sources for test purposes surrounding the antenna. As discussed in Chapter 7, sidelobe cancellation performance is impacted by multipath from terrain features surrounding the antenna. The selected configuration of interference sources to evaluate sidelobe canceller adaptive designs must be sensitive to siting so that the system responds to the interference illuminators and measured results are not obscured by multipath. A simple example illustrated the measurement sensitivity to multipath. A two-element adaptive array, illustrated in Fig. 8-19, is perturbed by a multipath component indicated by the dashed element. The multipath component is assumed to be 30 dB lower than the direct signal components received by the two array elements. The effect of this multipath error on the array's antenna pattern measurement in interference-free conditions is minor, 0.56 dB in amplitude and 3.5^{o} in phase using the coherent error statistics described in Chapter 1. The effect of this multipath component on the adaptive cancellation performance is illustrated in Fig. 8-20, where the multipath component has in-phase and out-of-phase conditions.

Separate measurements of the illuminators representing interference sources can be made to determine the presence of multipath. If sufficient bandwidth is available, network analyzer measurements of the signal received by the main antenna from the interference illuminator can be processed in the time domain to identify multipath components and their level. When the bandwidth is too narrow, the illuminator's location can be varied, and differences in the power transfer between the illuminator and the antenna can be used to identify the presence of multipath. Generally, locating the illuminators relatively close to the antenna being

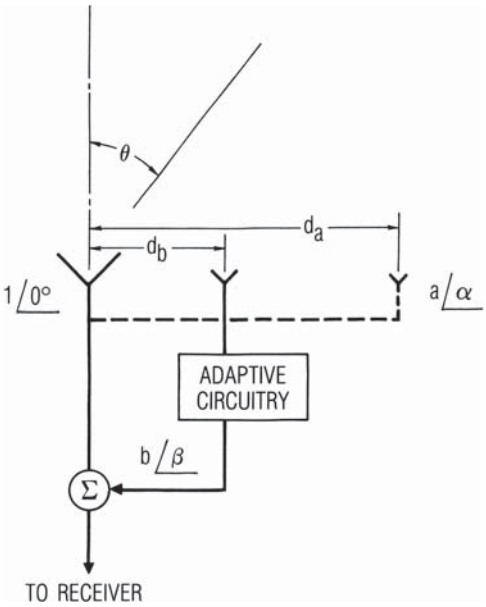


Figure 8-19 Adaptive antenna model [35]

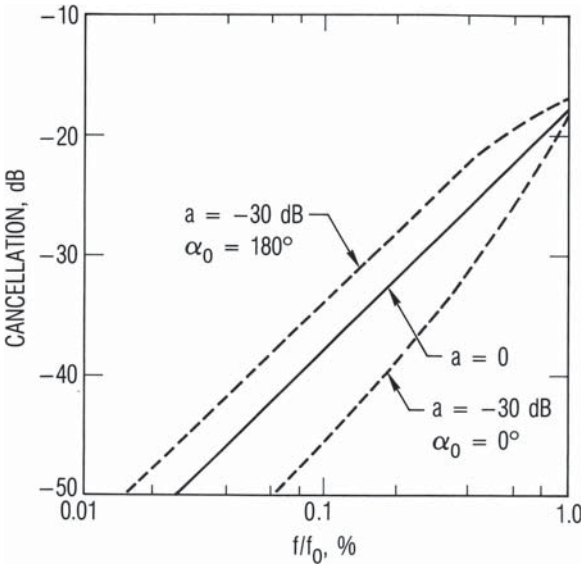


Figure 8-20 Multipath perturbation of adaptive cancellation [35]

evaluated is desired. In this case, far field requirements are considered. The sidelobe response of the main antenna arises from the collection of second-order radiation components that have a different far field requirement than that dictated by the main antenna's aperture size. The phasing between the secondary radiation components close to the main antenna may be somewhat different than their far field values, but their time delay differences are comparable to far field values. The dispersion resulting from the time delay differences is the limiting factor in adaptive cancellation performance so that an illuminator location closer than the antenna's far field does not impact measurement results. The principal requirement for the source illuminator is a separation that is sufficient enough that the main antenna and auxiliary elements are uniformly illuminated by the interference illuminator.

8.5 Evaluation of Antennas Having Integrated Electronics

The evaluation of antennas having integrated electronics and active array antennas typically poses problems because a terminal that separates the antenna and the electronics is not available for test purposes. Two examples of integrated antenna technology will be discussed. The first example is antenna designs where the antenna feed, LNA, and downconverter are integrated into a single package. Such designs are commonly found in reflector antenna systems for user segment applications. The second example is the evaluation of active array designs; both receive and transmit arrays will be discussed. In both cases, the active electronics should receive an adequate burn-in time to screen out infant mortality failures prior to assembling the integrated antenna.

When the antenna, LNA, and downconverter are an integrated assembly, the relative pattern levels and polarization properties can be measured by using the IF interface. If a sample of the local oscillator output can be obtained, mixing techniques can be used so that network analyzer instrumentation can be used. If the performance of a "first article" model for high production designs is being tested to determine design compliance, modifications of the first article to sample the output of the system's LNA can be used to obtain an RF sample. In this way, design compliance can be established, and other test techniques can be used in production testing to assure proper operation on a qualitative basis. Production units can be tested by signal injection techniques where test signals are injected into the antenna feed aperture much like hat couplers are used in thermal vacuum testing.

The G/T of the users' integrated antenna can be established by radio source techniques. This assumes the antenna aperture is large enough to allow solar radio source measurements. If this is not the case,

comparative G/T measurements can be made using an antenna that can be measured by radio source techniques and by separately comparing the SNR values when both antennas are illuminated by the same signal source. In some cases, the measured G/T is adequate to evaluate performance. In other cases, the antenna gain and the receiver noise figure are desired measurement parameters. Such measurements can be performed using system noise temperature measurement techniques previously discussed. Specifically, the receiver noise temperature can be measured by terminating the feed aperture with an absorber and measuring the output noise power levels when the absorber is at an ambient temperature and when the absorber is immersed in liquid nitrogen, providing a 77 K reference temperature. This measurement constitutes a hot/cold load measurement of the receiver with the reference terminal being the aperture plane of the feed horn. The resulting Y factor measurement allows determination of the receiver noise temperature. The feed can be assembled into the reflector and the antenna noise temperature then measured. The noise power levels are measured when the antenna feed is enclosed by the absorber at an ambient temperature and when the antenna is pointed at the elevation angle where the G/T measurements were performed.

The measurement of active antenna arrays is a considerably more involved and costly endeavor. The individual active elements comprising the array design must satisfy amplitude and phase tracking tolerances to maintain array performance. The elements for operational use are typically selected from a larger lot of elements based on their amplitude and phase tracking tolerances. The required number of elements in this selection must exceed the number in the array to account for replacements of elements that develop shortfalls in the subsequent assembly and testing phases. The selected active electronic devices are subjected to the required burn-in period and testing is performed to assure an adequate number of devices exist at that point, not only to populate the array but also to provide replacements for failed units uncovered in subsequent testing.

A similar process is used to select the phase shifter elements used in the design. The tolerance requirements at each bit position are examined along with insertion loss to select an adequate number of units for both the array operation and an allowance for replacement of units subsequently determined to have shortcomings. The control circuitry needed for phase shifter commanding should be measured and verified at this time.

The assembly of the array begins by integrating the selected units into subassemblies that will comprise the overall array design. Testing of the array subsystems includes evaluating not only their RF performance but also testing environmental suitability. Operation after vibration and

thermal cycling can be used to identify shortcomings at the subsystem level. The objective is to determine potential shortfalls in the subsystems at the lowest possible level to reduce the possibility of later replacements that delay the schedule as a result of component replacement and penalty testing needed to assure that the integrated assembly after component replacement complies with requirements. A surplus of subsystems that have completed the required testing is needed so that unit replacements are not delayed by additional testing at the subsystem level. Further integration of the subsystem units is then performed, and testing at each point of the integration is required to reduce the possible need for replacement elements. Again, environmental testing as the array assembly integration proceeds is recommended to identify potential shortfalls. Independent of their application, cables and connectors are the least reliable elements of satellite systems. Array antennas require a large number of interfaces between subassemblies, and thus the integrity of their interconnections must be verified at each stage of the assembly process. Mechanical and thermal testing provides a means of identifying interconnection shortfalls and reduces the risk of subsequent replacement and required retesting. Performing some environmental testing during the process of assembling the array is prudent because the replacement of failed array components at the assembly level is generally easier and more efficient compared to disassembling a fully completed array to replace failed elements in lower array assemblies.

Array designs depend on the amplitude and phase tracking performance of the electronics, and potential variations must be evaluated in their testing. The amplitude and phase tracking of the array electronics depend in part on their thermal environment, which for the space segment depends on the performance of the thermal control system for the array. Receive and transmit arrays also have different test issues.

Receive arrays require attention to establishing the system noise temperature. The receive arrays are operated in the linear region of the devices. Thermal control variations and power supply fluctuations all affect the amplitude and phase tracking performance, and the array sensitivity to these variations must be evaluated. For satellite uplink designs, the G/T must be established. A comparative measurement, for example, can be made by measuring the difference in the SNR of the array and the SNR of a standard system, and the antenna's G/T is derived from the SNR differences. The reference standard system typically is configured from a standard gain horn and a preamplifier.

Generally, measurements of uplink array antennas are performed in indoor facilities to protect the flight hardware. Indoor facilities envelope the antenna in an ambient noise temperature environment, whereas on-orbit, the antenna experiences an ambient noise temperature over the angular region subtended by the earth and the remaining angular

region has a 3 K cosmic background noise temperature. If comparative G/T measurements are made in indoor facilities, the antenna noise temperature must be corrected to account for the portions of the on-orbit field of view having the cosmic background temperature. One way to determine the array antenna's system noise temperature is by comparing the array's receiver noise temperature referenced to the array's aperture plane. The system noise temperature, comprised of the antenna and receiver noise temperatures, must be established. The receiver noise temperature is determined by measuring the output noise power when the array aperture is terminated by an absorber-lined foam container at an ambient temperature and when the absorber is immersed in liquid nitrogen at a 77 K temperature. The receiver noise temperature is obtained from these two noise measurements using standard Y factor techniques. This measurement includes the impact of array losses. The receiver noise temperature determined in this way is referenced to the array's aperture plane. Within the measurement facility, the antenna noise temperature referenced to the aperture plane is the ambient temperature. The array's antenna noise temperature when on-orbit can then be calculated using the emission background temperature and array patterns following the procedures described in Chapter 1. The on-orbit emission background includes an ambient temperature over the earth's field of view and a 3 K cosmic background for the remaining field of view. This emission background is particularly important because array designs endeavor to minimize the number of array elements by allowing grating lobes to exist beyond the earth's field of view. The emission contribution to the antenna noise temperature for the field of view beyond the earth is coupled by the array grating lobes, with the result that the array's antenna noise temperature is lower than the ambient temperature of the earth's background. The on-orbit G/T is determined from the measured G/T in the facility multiplied by the ratio of the system temperature when the calculated on-orbit antenna noise temperature is used and the system noise temperature when the ambient antenna noise temperature is used.

Transmit arrays have additional challenges. The array devices must be driven to their design operating point, generally close to saturation. Thus the operating point of the array must be maintained. The low power efficiency and gain of these devices result in significant thermal dissipation. The evaluation of the thermal design is a critical issue. Thermal control must be furnished during evaluation, but the on-orbit performance of the thermal control must be determined. The presence of intermodulation products, spurious responses, harmonics, and PIMs (passive intermodulation products) must be identified in the measurements.

Because of the long lifetime required by satellite operation, the sensitivity to element failures in the array must be addressed in the measurement.

The gain loss resulting from uplink array element failures equals the ratio of the number of elements still operating divided by the number of elements in the array design. The ERP loss resulting from downlink array element failures is the square of that ratio because both the downlink array gain and the total downlink transmit power are reduced. Element failures also impact the array sidelobe performance. As discussed in Chapter 2, analysis techniques are generally used to analyze the increased sidelobe levels statistically. Validating these analyses in the array testing efforts is recommended and can be accomplished by measuring array patterns with selected elements unpowered.

Finally, a means of calibrating array designs on-orbit is generally required to provide diagnostics and a means to correct the system's phase shifters to recover lost performance. The means of determining the effectiveness of planned calibration techniques must be evaluated as a part of the array testing.

8.6 Antenna Tracking Evaluation

The measurement requirements for antenna tracking designs become progressively more involved as the tracking designs proceed from open-loop program tracks to closed-loop monopulse tracking techniques. Independent of the tracking technique, the boresight axis of the design must be initially established so that the antenna's position can be referenced in a defined coordinate system. The positioner must be evaluated for its accuracy and repeatability in positioning the antenna, as well as the ability to cover the specified range of motion at the specified angular rates. The RF evaluation of monopulse designs requires measuring the difference pattern characteristics as well as the sum channel characteristics. Closed-loop designs require evaluating the control system for its transient response and stability, and the impact of the antenna and positioner mechanical properties typically measured by the design's locked rotor frequency.

8.6.1 Antenna Boresight Measurements

Antenna pattern measurements of the antenna's main beam are used to measure the location of the antenna's boresight axis. The positioner is commanded to various angular offsets from a nominal value of the boresight axis within the antenna's main beam and the power levels at each of these angular offsets are measured. These angular offsets are commanded in two orthogonal planes. The antenna's main beam is generally symmetric about the boresight axis for the higher-level portions of the main beam used in these measurements. The measured relative pattern levels at the offsets and their corresponding angular

offset values are plotted to form a pattern of the antenna's main beam behavior. Pattern levels on the pattern fitted to the measured values are then selected on both sides of the nominal boresight axis. Since the main beam at the higher pattern levels is assumed to be symmetric, the angular position of the boresight axis is located halfway between those two measured values. Several pattern-level values at positions on both sides of the nominal boresight axis should be measured to provide independent estimates of the boresight axis location that can be averaged to refine the results.

The selection of angular offset values should be made based on the pattern slope at that offset. Since the antenna pattern at its beam peak has zero slope, sampling the main beam response where the pattern has some slope provides higher accuracy than attempting to measure close to the main beam's peak level. Like the step track angular offset selection described in Fig. 2-13, the optimum measurement sensitivity is achieved for pattern levels that are 3 to 4 dB lower than the main beam's peak level. For example, select values that are 2.5 dB down from the estimated peak and define the boresight axis as halfway between the two angular positions; repeat for 2.75 dB, 3.0 dB, 3.25 dB, and so on from the peak, and then average the individual boresight locations at the sampled levels to locate the boresight position. The pattern derived from the measured angular offsets can be compared with an existing measured pattern or its fit to a simple analytic function, such as Gaussian, based on a measured beamwidth value. The measured pattern levels at the angular offsets should be repeated for antenna motion in two angular directions (e.g., left to right and right to left) to observe any backlash in the antenna positioning.

A similar technique is applied to evaluating the boresight axes for monopulse difference patterns. The boresight axes for both the sum and difference patterns need to be established and ideally, both sum and difference patterns have a common boresight axis. Differences in their boresight locations indicate imbalance in the difference pattern circuitry. Like the sum beam, measurement of the difference pattern near the boresight axis poses a problem since the difference null has a finite value. Sampled points away from the boresight location are used to identify the boresight axis's location. Similarly, antennas for multiple frequency operation require verifying boresight coincidence at all of the frequencies. In this case, the antenna boresight axes are measured at the individual frequencies used in the design. The highest frequency of operation provides the most precise measurement since the narrowest antenna beamwidth results at the highest frequency.

The boresight measurements are generally performed using the system's positioner. Verification of the positioner's operation during the boresight measurement is performed at the same time. The range of

motion and angular rates are measured and compared with specified values. The antenna drive system can be evaluated by monitoring the drive current over the full azimuth and elevation travel. Typically, a sampling resistor is furnished to monitor the drive current values, and unanticipated variation of these values over the range of travel indicates a problem with the positioner's mechanical drive system. The velocity and acceleration capabilities of the system can be measured by timing intervals on the positional readout. The ability to track signals at high elevation angles having the maximum azimuth rates is a final evaluation. As appropriate, the ability to track a low-altitude satellite at high elevation angles can be assessed. The time variation of the positioner's motion as indicated by angle encoders can be measured and compared to corresponding values predicted from the positioner's location and the satellite's ephemeris values. The latter predicted values are taken as "truth" since the calculated values are based on well-established ephemeris values.

The alignment of the positioner in the system's reference coordinate system can be verified by using signals at known locations. Such signals may be a source antenna on a pattern range for small antennas, while satellite signals or radio sources are used for larger antennas. Several signal source directions are used to determine any biases in antenna pointing. For fixed installations, the objective is to determine the antenna's alignment with respect to true north so the antenna can be properly commanded to point in desired directions. The antenna is commanded to point at the source locations, and step track techniques provide a means to determine if the commanded positions correspond to boresight alignment with the signal. Repeating such measurements from different directions evaluates any backlash in the positioner drive. Most large antennas use dual motor drives that virtually eliminate backlash errors, but such pointing measurements perceive any backlash when the antenna is commanded from different directions.

8.6.2 Closed-Loop Antenna Tracking Evaluation

Closed-loop tracking systems [39, 40] expand the testing to evaluate not only the RF aspects of the antenna design but also the control system response to tracking requirements. The sum and difference patterns must be measured to verify the RF performance. The error response, the difference divided by the sum beam responses, should also be measured at the control system input. The error response is evaluated by measuring the control voltage inputs for commanded angular offsets from a test signal. The control system response depends on both the RF receiver drive loop and structural properties of the antenna. Both the transient and steady state performance require evaluation.

The sum and difference patterns for the monopulse design can be evaluated using conventional antenna pattern measurements. The boresight coincidence of the two pattern types must be verified as previously described. Depending on the design and system requirements, these patterns can also vary with polarization. The null depth of the difference pattern can also limit the minimum angular resolution; this depth depends on the monopulse circuitry balance and the antenna's cross-polarization response.

As discussed in Chapter 2, the sampled difference channel is coupled onto the sum beam for pseudomonopulse designs. This coupling can result in insertion phase offsets for narrow bandwidth designs or time-delay offsets for wide bandwidth designs that result in tracking errors. These offsets or biases result in coupling between the positional axes referred to as "crosstalk." Crosstalk measurements require an illumination source at a known position, and care must be used to avoid multipath that results in an angular displacement from the position of the free space source response. In some cases, a satellite beacon or downlink signal fulfills these requirements. Crosstalk errors are evaluated by commanding an angular offset, and allowing the control system to respond to return the antenna to its boresight position. If the angular offset is in the azimuth direction, for example, the return to the boresight location should result in only azimuth positioning of the antenna. Variation of the antenna's position in the elevation direction indicates the presence of crosstalk. The antenna's motion during repositioning to the boresight alignment can be monitored by the antenna's encoders. Such tests are commonly referred to as "snap-on" tests.

If crosstalk is indicated by the system's response, the insertion phase or the group delay differences must be adjusted to reduce crosstalk levels. For narrow bandwidth designs, the phase differences between the sum and difference channels can be adjusted by properly setting a phase shifter that is typically inserted into the difference channel. The required phase adjustment is obtained by observing whether the crosstalk values have negative or positive values and the measured deviation from the desired tracking trajectory. If the system is required to operate at different frequencies, the tests are repeated to obtain compensating phase values at each required frequency. The sum and difference channels generally use waveguide, and the combination of waveguide frequency dispersion and different lengths of waveguide make it necessary to use different phase shifter values at each frequency. The alternative is to compensate group delay differences in the sum and difference channels. The group delay adjustment becomes more tedious. A sequence of narrow band or swept measurements can be made to determine the phase slope, and the required group delay compensation can be estimated from those values. The group delay is adjusted by adding additional waveguide into generally the sum channel path so that both the difference and sum channels

have the same group delay values. Group delay compensation for the wide bandwidth requirement is a key design issue, and attention should be paid to the proper group delay compensation during the feed development and calibration prior to integrating the feed with the antenna.

The control system response must also be quantified. This response depends on not only the RF error response but also the mechanical design of the antenna and positioning system. The mechanical aspects of the design are commonly expressed by the locked rotor resonant frequency. This value depends on the structural and antenna mounting stiffness, and bounds for its values [41] derived from measured antenna responses are given in Fig. 8-21.

An illumination signal source is required to provide a test signal for the tracking system. The control system response is typically evaluated by measuring its step response. The step response measures the system's response to transient inputs, and displays the classic control system response and stability. The step response is evaluated by using a function generator to inject a square wave into the control system's input. Since the control system responds to produce a minimum value of the control system's input, the injected square wave produces step responses so that the control system's transient behavior can be observed. A ramp response is obtained by injecting a triangular wave from the function generator into the control system's input. The ramp response evaluates the monopulse error channel response. The triangular wave drives the antenna up and down the linear cone that describes the antenna's difference-over-sum ratio. This measurement is repeated for different angular locations and polarizations to quantify the "antenna portion" of the closed-loop system.

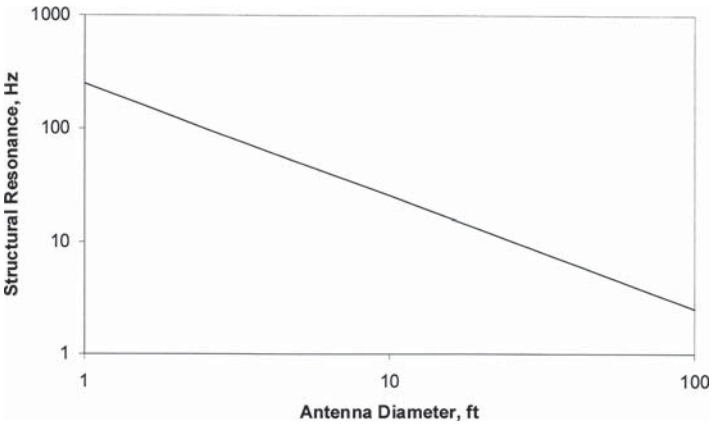


Figure 8-21 Antenna mechanical resonance bounds [41]

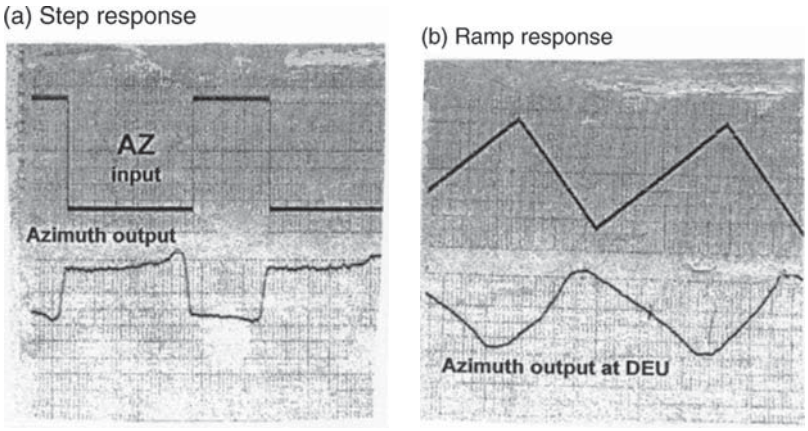


Figure 8-22 Example control channel responses [39]

Example measurements of the control system response [39] in Fig. 8-22 indicate a well-behaved step response for the antenna being measured. This figure is reproduced from a strip chart recorder record that is commonly used for such measurements. The measured ramp response is also well behaved and indicates the system follows a trajectory about the tip of the difference-over-sum ratio cone. These data use a 1/10 Hz frequency for the square wave and triangular wave inputs to the control system's error port obtained from a function generator. The amplitude of these injected waveforms is selected by the design's scale factor to span the nominal tracking range of the system.

The measured responses in Fig. 8-22 were performed when the incident illuminating signal had the design circular polarization sense. As discussed in Chapter 2, incident signals with the cross-polarized opposite sense can result in unstable antenna tracking. The measured responses in Fig. 8-23 illustrate the unstable operation when the antenna receives incident signals having the cross-polarized polarization sense. The antenna behavior in this case was quite dynamic following a coning motion that drove the antenna away from the illuminating signal's direction. The measured step response illustrates the unstable behavior and the antenna fails to track the signal. The ramp response further reveals the underlying reason. The ramp response does not have the linear variation about the boresight axis that is predicated in the control system design. Thus, an important part of evaluating closed-loop antenna tracking system limitations addresses the sensitivity to the incident polarization sense.

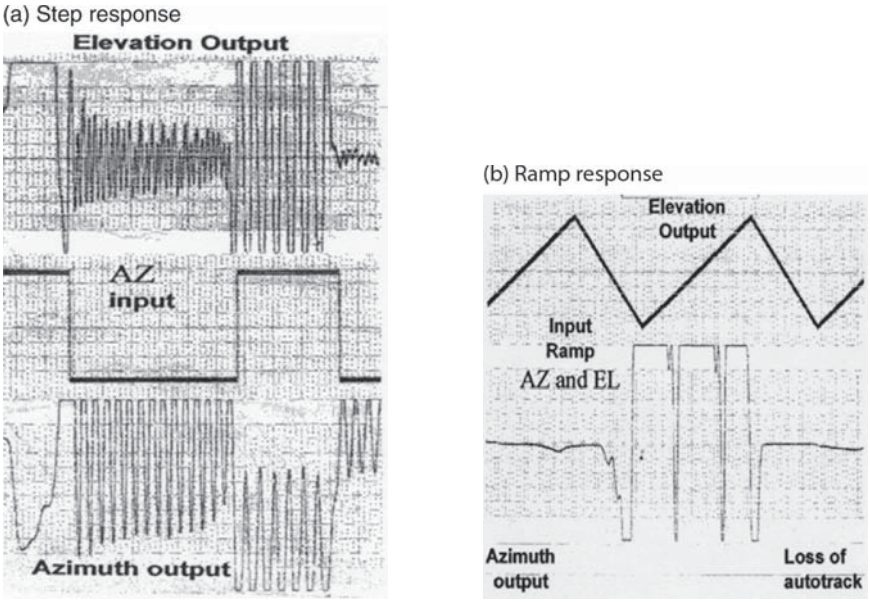


Figure 8-23 Unstable responses for wrong sense polarization [39]

8.7 System Evaluation

System performance depends on the G/T and ERP levels produced by antenna systems that must be quantified. The level of complexity in evaluating the system parameters depends on both the antenna design and application. Some antenna designs have a relatively straightforward means of evaluating the antenna parameters. For example, user antenna designs for VSAT applications require evaluating the antenna gain parameters and determining either the system noise temperature to obtain the G/T value or the transmitter power levels to determine the ERP value. Straightforward techniques exist to evaluate the antenna gain, the receiver noise figure and the antenna noise temperature can be measured, and the transmitter power output can be measured using standard techniques. In production testing of such designs, the antenna gain is evaluated using a design example. The feed system can be addressed to determine workmanship errors that would result in noncompliance, and random checks can be made to assure those workmanship errors do not exist in production items. The electronics can be separately evaluated to insure their compliance, integration of the antenna and electronics can be demonstrated on a design example, and a detailed evaluation can be conducted to assure design compliance.

Volume production items can be evaluated in this manner to insure the specified performance levels are maintained.

The system evaluation for limited production antennas for the user segment is generally performed for each item and more rigorous testing is conducted. A variety of antenna testing techniques can be used depending on the specific measurement requirements for the application. General-purpose facilities and instrumentation can be used for smaller user antennas. Larger antennas pose test challenges that are described in further detail in Chapter 9. Larger antennas are generally tested in two phases. The first phase is conducted at the vendor's location and establishes design compliance with system requirements. The second test phase is performed at the antenna's operational site to verify correct assembly and pointing alignment. Individual components such as the antenna feed are subject to a detailed evaluation that can be conducted in general-purpose facilities. The integrated antenna system is then evaluated at the vendor facility and tested to determine compliance with the elements of the requirements verification matrix for the design, as described in Chapter 9. The vendor measurements provide the most detailed measurements and qualify the design performance. Once the vendor tests have demonstrated design compliance, separate installation testing is performed at the site. The installation testing has the objective of demonstrating that the antenna system has been correctly assembled and its performance complies with the levels determined in the vendor tests.

Measurements of the antenna and system electronics should duplicate methodologies that were demonstrated at the vendor location and should be used in the future to demonstrate that the system performance is maintained over the system's lifetime. The vendor testing should endeavor to use alternative measurement methodologies and supporting analyses to provide confidence in the measured performance. The installation tests are therefore a subset of the vendor tests and provide baseline performance data that can be used in subsequent maintenance testing to determine if the system performance has been maintained during the lifetime. Separate evaluations and methodologies of the system electronics are also performed during the installation tests, and like the antenna, performed during the system's lifetime. In many cases, systems using antennas in this category are operated remotely, and an important part of the installation testing is demonstration that system shortfalls can be identified by commanding BITE (built-in test equipment) resources and commanding the substitution of redundant equipment to replace failed units to maintain system operation. Chapter 9 provides further discussion of these issues.

While a variety of methodologies and techniques can be used to evaluate user segment designs on a system level, the measurement of space

segment designs pose further challenges. Chapter 9 discusses the measurement processes used in development, qualification, and on-orbit test phases in detail.

On-orbit test methodologies for *G*/*T* and ERP performance [42] merit further discussion. Such testing is performed by a mission control station using link analyses. The calibration uncertainty of the mission control station must be carefully examined. Generally, such testing is performed shortly after the satellite's launch to assess compliance, and the satellite's transponder is exercised by test signals.

The downlink ERP can be measured by uploading a test signal into the transponder and measuring this signal with the mission control station. CW (continuous wave) tones are commonly used for frequency translating transponders; regenerative transponders use test signals that are compatible with their modulation formats and BER values are measured. The signal received by the mission control station can be measured in two ways [43]. The first method measures the received SNR and uses the calibrated *G*/*T* value of the mission control terminal. The second method measures the received signal power that requires knowledge of the mission control terminal's antenna gain value and the insertion gain of the electronics between the terminal where the antenna gain value is referenced and the indicated receiver output. When CW tones are used for testing, a means of measuring the insertion gain is to inject another CW tone offset somewhat in frequency from the downlink CW test tone through a calibrated coupler at the antenna's terminal plane. By adjusting the level of the CW tone to equal the level of the downlink CW tone, and measuring the injected CW tone's power level and the coupler's coupling coefficient, the insertion gain of the electronics can be established. The uplink *G*/*T* is measured by adjusting the transponder gain so that the downlink test signal is not noise limited. The uplink test signal level transmitted to the satellite is then reduced until the uplink noise becomes apparent in the downlink. Both ERP and *G*/*T* values are derived using link equations. In addition to the calibration attention needed for the mission control terminal, atmospheric propagation loss must be determined and the space segment's antenna pattern level in the direction of the mission control terminal must be obtained from antenna data taken prior to the satellite's launch.

References

1. ———, *IEEE Standard Test Procedures for Antennas* (New York: Wiley-Interscience, 1979).
2. J. S. Hollis, T. J. Lyon, and L. Clayton (eds), *Microwave Antenna Measurements* (Atlanta GA: Scientific-Atlanta, 1970).
3. W. D. Burnside, I. J. Gupta, and T-H. Lee, "Indoor Antenna Measurements," Chapter 50, in J. L. Volakis (ed), *Antenna Engineering Handbook* (New York: McGraw-Hill: 2007).

4. R. B. Dybdal and C. O. Yowell, "VHF to EHF Performance of a 90-ft Quasi-Tapered Anechoic Chamber," *IEEE Trans Antennas and Propagation*, vol. AP-21 (July 1973): 579–581.
5. R. B. Dybdal, "Methodology to Project Antenna Measurement Accuracy," *1989 AMTA Symposium Digest* (October 1989).
6. I. J. Gupta, W. D. Burnside, and E. K. Walton, "Time and Direction of Arrival Estimation of Stray Signals in a RCS/Antenna Range," *1996 AMTA Symposium Digest* (September/October 1996).
7. D. A. Leatherwood and E. B. Joy, "Plane Wave, Pattern Subtraction, Range Compensation," *IEEE Trans Antennas and Propagation*, vol.AP-49 (December 2001): 1843–1851.
8. R. B. Dybdal, "Error Statistics in RF Measurements," *1999 AMTA Symposium Digest* (October 1999).
9. W. B. Davenport and W. L. Root, *An Introduction to the Theory of Random Signals and Noise* (New York, McGraw-Hill, 1958).
10. J. R. Mentzer, "The Use of Dielectric Lenses in Reflection Measurements," *Proc IRE*, vol. 41 (February 1953): 252–256.
11. T. H. Lee and W. D. Burnside, "Performance Trade-off Between Serrated Edge and Blended Rolled Edge Compact Range Reflectors," *IEEE Trans Antennas and Propagation*, vol. AP-44 (January 1996): 87–96.
12. R. B. Dybdal and G. E. Stewart, "Development of a Small Compact Range Facility," *1988 AMTA Symposium Digest* (September 1998).
13. R. B. Dybdal, "Horn Antenna Sidelobe Reduction Using Absorber Tunnels," *1977 IEEE AP-S Symposium Digest* (June 1977).
14. G. E. Stewart and R. B. Dybdal, "Analysis and Measurement of Horns in Absorber Lined Tunnels," *1988 AMTA Symposium Digest* (September 1988).
15. D. A. Thompson, R. B. Dybdal, and F. A. Pisano, "Quasi-Compact Range," *2007 AMTA Symposium Digest* (November 2007).
16. R. B. Dybdal and K. M. SooHoo, "Evaluation of Adaptive Multiple Beam Antennas," *1990 AMTA Symposium Digest* (October 1990).
17. D. Slater, *Near Field Antenna Measurements* (Norwood MA, Artech, 1991).
18. A. G. Repjar, A. C. Newell, and M. H. Francis, "Accurate Determination of Planar Near Field Correction Parameters for Linearly Polarized Probes," *IEEE Trans Antennas and Propagation*, vol. AP-36 (June 1988): 855–868.
19. A. C. Newell, R. D. Ward, and E. J. McFarlane, "Gain and Power Parameter Measurements Using Planar Near Field Techniques," *IEEE Trans Antennas and Propagation*, vol. AP-36 (June 1988): 792–803.
20. J. J. Lee, E. M. Ferren, D. P. Woollen, and K. M. Lee, "Near Field Probe Used as a Diagnostic Tool to Locate Defective Elements in an Array Antenna," *IEEE Trans Antennas and Propagation*, vol. AP-36 (June 1988): 884–889.
21. A. C. Newell, "Error Analysis Techniques for Planar Near Field Measurements," *IEEE Trans Antennas and Propagation*, vol. AP-36 (June 1988): 754–768.
22. A. C. Newell and C. F Stubenrauch, "Effect of Random Errors in Planar Near Field Measurement," *IEEE Trans Antennas and Propagation*, vol. AP-36 (June 1988): 769–773.
23. G. Hindman and A. C Newell, "Simplified Spherical Near Field Accuracy Assessment," *IEEE Antennas and Propagation Magazine*, vol. 49, (February 2007): 233–240.
24. R. B. Dybdal, T. T. Mori, and H. E. King, "High Sensitivity Millimeter-Wave Instrumentation," *1981 IEEE AP-S Symposium Digest* (June 1981); see also, T. T. Mori, R. B. Dybdal, and R. D. Etcheverry, "High Sensitivity Millimeter-Wave Measurement Apparatus," (April 9, 1985) U.S. Patent 4,510,622.
25. E. B. Joy and D. T. Paris, "A Practical Method for Measuring the Complex Polarization Ratio of Arbitrary Antennas," *IEEE Trans Antennas and Propagation*, vol. AP-21 (July 1973): 432–435.
26. G.M. Shaw and R. B. Dybdal, "Time Domain Processing of Mutual Coupling Measurements," *1990 IEEE AP-S Symposium Digest* (May 1990).
27. A. C. Newell, "Improved Polarization Measurements Using a Modified Three Antenna Technique," *IEEE Trans Antennas and Propagation*, vol. AP-36 (June 1988): 852–854.

28. A. C. Newell, C. F. Stubenrauch, and R. C. Baird, "Calibration of Microwave Antenna Gain Standards," *Proc IEEE*, vol. 74 (January 1986): 129–132.
29. J. T. Shaffer and R. B. Dybdal, "Using Standard Gain Horns," *2000 AMTA Symposium Digest* (October 2000).
30. D. A. Guidice and J. P. Castelli, "The Use of Extra Terrestrial Radio Sources in the Measurement of Antenna Parameters," *IEEE Trans Aerospace and Electronic Systems*, vol. AES-7 (March 1971): 226–234.
31. J. W. M. Baars, "The Measurement of Large Antennas with Cosmic Radio Sources," *IEEE Trans Antennas and Propagation*, vol. AP-21 (July 1973): 461–474.
32. D. F. Wait, "Precision Measurement of Antenna System Noise Using Radio Stars," *IEEE Trans Instrumentation and Measurement*, vol. IM-32 (March 1983): 110–116.
33. R. B. Dybdal, "On G/T Radio Source Measurements," *2000 AMTA Symposium Digest* (October 2000).
34. R. B. Dybdal, "G/T Measurement of Small Antennas," *1997 AMTA Symposium Digest* (November 1997).
35. R. B. Dybdal, "Test Methodology for Adaptive Antennas," *1987 AMTA Symposium Digest* (September/October 1987).
36. R. B. Dybdal, "Adaptive Uplink Antenna Testing," *1994 IEEE MILCOM Symposium Digest* (October 1994).
37. R. B. Dybdal, "Measurement if Antennas with Integrated Electronics," *1992 AMTA Symposium Digest* (September/October 1992).
38. R. B. Dybdal, "Qualification Measurements for Satellite Uplink Arrays," *2003 AMTA Symposium Digest* (October 2003).
39. R. B. Dybdal and D. D. Pidhayny, "Evaluation of Antenna Tracking Systems," *2001 AMTA Symposium Digest* (October 2001).
40. T. Inoue and T Kaitsuka, "K-band Tracking System for Domestic Satellite Communication System," *IEEE Trans Aerospace and Electronic Systems,* vol. AES-17 (July 1981): 561–570.
41. D. D. Pidhayny, private communication
42. R. B. Dybdal, "On-Orbit Communication Satellite Measurements," *2002 AMTA Symposium Digest* (November 2002).
43. R. B. Dybdal, "ERP Measurement Issues," *1998 AMTA Symposium Digest* (October 1998).

Chapter

9

Satellite Antenna System Evaluation

9.1 Overview

Antenna systems for satellite applications pose interesting test and evaluation challenges compared with other applications. The space segment antennas encompass a diverse range of technologies and the requirements of their applications similarly pose equally diverse test and evaluation requirements. Testing requirements vary significantly with the antenna design complexity. The reliability required of space hardware and the inability to service the satellite on-orbit present the challenge of adequately testing the hardware. For this reason, space segment antenna testing is rigorously performed. Space segment testing must address not only the antenna design's RF performance but also its capability to withstand the launch and on-orbit environments. Such testing typically requires support from specialists in a variety of disciplines. Test and evaluation for space segment antennas include three distinct phases spanning the development, qualification, and on-orbit phases. Each phase has differing requirements and objectives. Extensive testing of space segment antennas is essential to satisfy on-orbit reliability objectives and ensure survival during launch and on-orbit operations.

The user segment likewise poses test and evaluation challenges as their requirements become increasingly stringent. High-production user antennas require test techniques that provide high confidence in achieving performance compliance, while also being capable of efficient execution to minimize production costs. Production testing

for user segment antennas results in the trend towards establishing the compliance of a design, identifying those aspects of the design that potentially could produce non-compliance, and evaluating the test item to insure the non-complying elements are not present in the production items. This test philosophy differs markedly from past testing where individual items were thoroughly evaluated within a reasonable schedule. Larger terminals are produced in a limited number and the combination of requirements specific for their application and the complexity of their design dictates testing of individual systems. General-purpose antenna measurement facilities are often unsuitable because of the antenna's physical size and excessive far field requirements. Such antenna designs require specialized test techniques for their evaluation. The evaluation of polarization purity when orthogonal polarizations are used in polarization reuse techniques and the evaluation of sidelobe envelope compliance are two examples of current requirements that require increased attention to testing. Test and evaluation of user segment antennas include four distinct stages spanning the development, qualification, integration/acceptance, and sustainment phases. Test issues for the space segment and user segment are discussed in turn.

Both space segment and user segment developments start with their respective system-level requirements, derive design requirements, and develop design candidates. The test process is initiated by developing a requirements verification matrix that indicates the test methods to be used in assuring all system requirements are satisfied. The determination of the adequacy of test capabilities and the development of a testable payload designs are critical activities at the program's initiation. It is therefore essential to address test plans and methodologies early in the program, identify test terminals to be incorporated into the antenna's design, and determine the necessary facilities and techniques that require development and validation prior to the start of the testing. Failure to comprehensively address the test methodologies and facility requirements at the initiation of a program is often accompanied by significant expense and schedule penalties.

9.2 Space Segment Antenna Testing

Antenna test techniques, instrumentation, and test facilities, as described in Chapter 8, are well developed for a wide variety of applications and form the basis to evaluate satellite antennas. Space segment antenna test and evaluation, however, pose unique requirements that are not normally part of antenna system evaluation. Space qualifying antenna systems [1] is multi-disciplinary, not only assuring RF compliance but also compatibility with launch and on-orbit environments. Accordingly, test techniques and methodologies, instrumentation, and

facility requirements must be tailored to meet the specific requirements for satellite antenna designs. An overview of the test and evaluation process for space segment antennas is provided to guide application to specific programs.

9.2.1 The Space Segment Antenna Testing Process

The process for space segment testing described in Fig. 9-1 begins by examining system requirements. These system-level requirements are then used to develop design candidates capable of complying with the system requirements and to devise a requirements verification matrix that identifies how the compliance with design requirements is to be verified. Typically, four distinct means of verification are addressed: (1) inspection, (2) analysis, (3) demonstration, and (4) test. Inspection typically concerns reviewing vendor component characteristics. Analysis often applies well-known procedures to characterize component performance. For example, a variety of analyses codes exist to project the RF performance of antennas. The results of such analyses are used to assess the RF performance compliance and to provide a basis of comparison with the measured performance. Demonstration generally

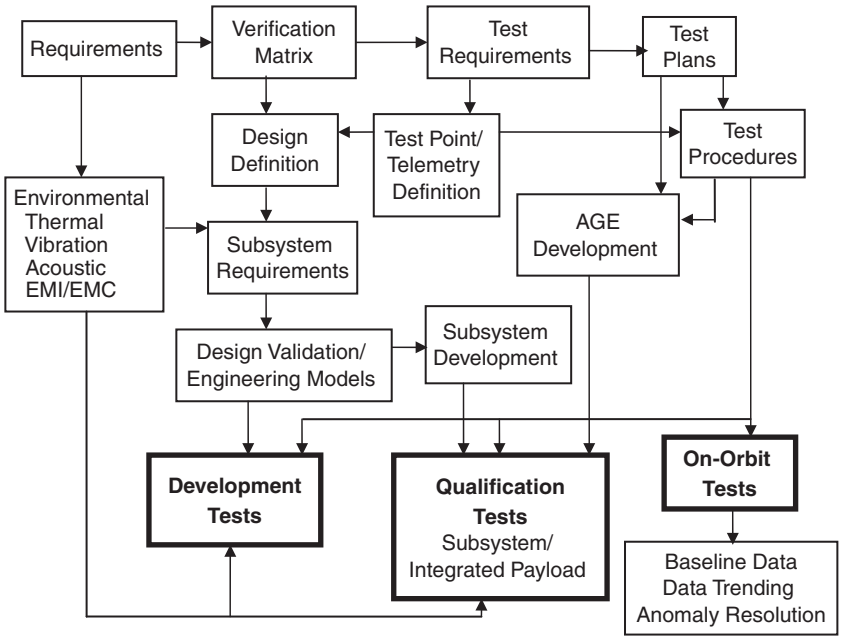


Figure 9-1 The space segment testing process

evaluates the item's function through observation. Examples of demonstration are positioning an antenna over its required range of motion or operating a deployment mechanism. Testing demonstrates compliance on either a component or integrated assembly basis and typically involves the use of other instrumentation and detailed data collection, as well as analyses to establish compliance with requirements. In many instances, a combination of verification means (e.g., analysis and test) is used to verify compliance. The type of testing identified in the verification matrix then allows determination of the test requirements. The verification matrix also guides the system design definition.

The test requirements identified in the verification matrix are then examined in further detail to identify facility and instrumentation requirements. The test requirements also impact the system design. Hardware design must include the appropriate test points to demonstrate compliance and parameters, and interface requirements for telemetry readout requirements must be identified for on-orbit operations. Careful attention to defining and developing test requirements further identifies facility and software needs that will be used in development and qualification testing. In some cases, the on-orbit verification of compliance also imposes additional test requirements. Test requirements also define facility needs and development requirements for development and qualification testing. The associated requirements for software control of system testing and the software data processing must also be established. The examination of test requirements is essential shortly after the program's initiation to identify resource requirements and needed development to avoid schedule and cost penalties later in the program.

The test point and telemetry requirements are then factored into the design definition. Typically, at this point of the program, alternative design implementations exist, and the test point and telemetry requirements for the alternative designs should be factored into the design definition. The design definition and the environmental requirements are then considered in defining more precise subsystem requirements leading to a candidate design. Depending on the design and existing heritage, decisions can be made on the design validation and the extent to which engineering model development must be pursued to reduce development risk. Once the design selection is established, subsystem development can proceed.

In parallel with the design definition, test requirements are used to develop test plans that identify facility, instrumentation, and resource requirements; the need to develop any specialized facility and/or test fixtures; and the software requirements for test instrumentation control, data processing, and data storage at a top level. The test plans provide a definition of the detailed test procedures needed in the program. The test plan addresses test objectives and a description of the specific test

parameters related to the requirements verification matrix compliance. The qualification tests of the integrated satellite are performed by a comprehensive test set referred to as AGE (aerospace ground equipment). This collection of instrumentation is assembled to provide end-to-end satellite payload testing during the qualification testing and emulates the overall operation of the payload in its on-orbit configuration. The interface of the AGE with the payload antennas must be addressed during test planning and the identification of specialized test equipment to evaluate antenna performance in the qualification testing phase. The overall test plan must address testing needs in both performance demonstrations and also the environmental testing at both a subsystem and an integrated spacecraft level.

The test procedures describe the test facility, instrumentation, and calibration needs; the detailed test methodologies to be used, the data collection, processing, and storage requirements; and the pass/fail limits of the data. The complex antenna technology in today's programs is sufficiently involved that software for instrumentation control, data accumulation, and data analyses is generally developed to manage the testing; the validation of such software is required.

Three distinct test phases, development, qualification, and on-orbit testing, exist. Their top-level objectives are:

1. The development phase validates compliance with the system design and environmental requirements.
2. The qualification phase has three objectives: to demonstrate the satellite antenna conforms with the performance specifications that were previously established in the development phase, to address any workmanship issues that could limit on-orbit reliability, and to establish the flightworthiness of the satellite.
3. The on-orbit phase establishes the on-orbit performance compliance and provides a trending and diagnostic capability over the satellite's lifetime.

Each test phase has differing objectives, parameter requirements, test methodologies, and instrumentation and test facility needs. An overview of the test phases in Fig. 9-2 illustrates the principal differences that will be discussed in further detail. Notice that the RF test parameters evolve from the component level to the system level as the test progresses from development to on-orbit phases. Environmental testing is limited to development and qualification testing. Development testing on components limits risks during qualification testing. Qualification testing is performed at both the subsystem and integrated payload levels. Similarly, facilities vary from general-purpose RF facilities and instrumentation during development testing, to RF test facilities to

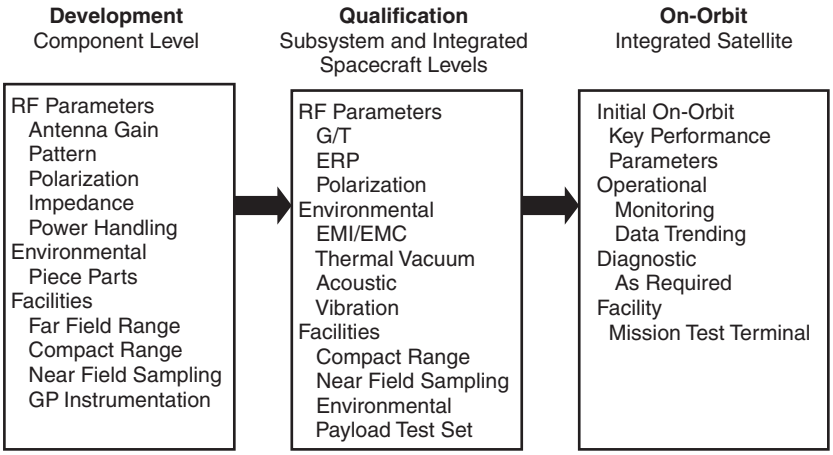


Figure 9-2 An overview of space segment testing [1]

reduce the risk to flight hardware, to comprehensive payload test sets and specialized test configurations to support environmental testing that supports qualification testing, to mission control test terminals that support on-orbit testing.

This test process is accompanied by a review process. An SRR (System Requirements Review) is typically held shortly after the verification matrix is established to insure the system requirements are properly and mutually understood and to approve the verification matrix and its associated verification methods. A PDR (Preliminary Design Review) is held at the completion of the design definition to review the proposed implementation and test plans. A CDR (Critical Design Review) is held at the completion of development testing to assure the system design is sufficiently mature to warrant proceeding to qualification testing. The SRR, PDR, and CDR activities are held at both the subsystem and system levels. Following CDR, a series of MRR (manufacturing readiness review) are conducted prior to the fabrication of each of the subsystems. All testing is proceeded by a series of TRR (test readiness review) at the component, subsystem, and integrated payload levels to insure test objectives, test methodologies, facility and instrumentation capabilities, test documentation, and elements of the requirements verification matrix to be verified are understood.

Qualification testing conducted at both the subsystem and integrated satellite levels establishes both performance compliance and environmental compatibility. Additional testing to insure the payload health is conducted during launch processing. The additional integration of system electronics and growth of design complexity require more comprehensive testing prior to launch. The trend is to transition from "light bulb" testing

that indicates the electronics function at a qualitative level to more quantitative performance testing to assure previously measured performance has been maintained. Such end-to-end verification can be performed using test couplers to evaluate payload electronics, but as electronics are integrated into the antenna systems, end-to-end testing cannot be performed without including the integrated electronics within the antenna system. These more quantitative test requirements are the reason for the previously recommended development of portable compact range and near field sampling capabilities for launch processing evaluations.

When the satellite is in its orbital position, the on-orbit testing commences to verify the on-orbit performance complies with system specifications. This testing is generally accomplished by the mission control station. The mission control station therefore requires attention to calibration and diagnostic capabilities. In some programs, performance and financial incentives are applied to system margin values so the measurement uncertainty becomes a major issue. When the initial on-orbit testing is completed, sell-off and customer acceptance activities commence. After this initial performance testing that establishes a baseline data set, routine testing is conducted for trending assessments, and testing to verify performance after redundant components are commanded into operation is required to assure on-orbit performance is maintained. Additional test capabilities and commanding to access telemetry data is not routinely used are also required by mission control assets to provide diagnostic capabilities over the satellite's lifetime. These requirements place demands on the mission control stations to maintain calibration, provide its internal BITE capability so indicated anomalies can distinguish between satellite and mission control station malfunction, and devote attention to data collection and retention. The same requirements also place demands on the satellite to provide an adequate number of telemetry points so that perceptive diagnostics can be obtained over the satellite's lifetime.

The development, qualification, and on-orbit test phases are contrasted in Fig. 9-2. The testing evolves from component-level testing during development testing to system-level testing during on-orbit testing. This evolution is reflected in the test measures for each phase. The facility requirements similarly evolve from general-purpose test facilities used in development testing to special-purpose capabilities found in mission control stations.

9.2.2 Space Segment Development Testing

The development phase is the initial and perhaps most fundamental phase in configuring the test and evaluation of satellite antennas. Several elements are included in this phase. At the inception of a program,

alternative design approaches for the antenna system can exist. In such cases, a plan and criteria must be devised to select the flight design. This plan requires addressing differences in the performance capabilities of the design and the relative development risks of alternative design approaches. Basic parameters, gain, pattern coverage, vehicle effects, polarization purity, and so on are measured to establish the design's RF compliance. RF testing in this phase typically capitalizes on general-purpose facilities and instrumentation and is conducted on breadboard or prototype antenna hardware since the principal objective is to establish the RF performance. Depending on the development status of the design, the need for engineering model design(s) and their required fidelity must be devised to demonstrate RF performance compliance and to address issues for flight suitability. The evolution of past heritage designs is an important factor in determining the extent of required engineering model development. Any design aspects that are unique or fail to have previous flight heritage are tested to reduce program risk and to allow time in the schedule to address aspects where development testing has identified shortfalls in performance. Assessments of development risk accompany this selection process and such assessments are generally a part of the selection of the flight design.

The development testing provides experience with the test facilities, instrumentation, methodologies, data processing, measurement uncertainty assessments, adequacy of test points and telemetry data, and operations to test software. Antenna control requirements and associated supporting software must be demonstrated for both test and operations. Interface requirements with other payload electronics must also be established at this time. Experience in development testing also identifies the need for additional augmentation of test methodologies, data processing, and test control software requirements needed to support future qualification testing. As the complexity of antenna designs increases, software increasingly becomes more important in conducting the required testing, and verification of software during development testing reduces risk during subsequent qualification testing. Development testing not only establishes the antenna design's compliance with system requirements but also provides data to support other RF antenna testing requirements. Such RF requirements include EMI/EMC (electromagnetic interference/electromagnetic compatibility), ESD (electrostatic discharge), multipaction, corona, PIM (passive intermodulation), and other issues. Environmental testing also must be addressed during development testing to establish design suitability. Such testing can impose additional requirements to monitor RF performance during the environmental testing, and the development and demonstration of suitable monitoring techniques during the development phase reduce risk during qualification testing. For example, RF measurement in thermal

vacuum chambers requires specialized attention since RF measurement facilities generally cannot be integrated into the thermal vacuum chamber. One approach is a "hat coupler" design to allow measurement in qualification facilities. Design verification and calibration of such couplers are required in development tests so that the capability is available during qualification testing. Finally, efforts to identify potential workmanship issues in the proposed design also need attention early in the program so subsequent measurement of flight hardware can identify potential shortfalls in the qualification test phases.

It is essential to address these test and evaluation issues at the initiation of the program to avoid schedule and cost impacts later in the program. The development of the necessary test capabilities must be undertaken so that qualification testing is not delayed. A detailed examination of these requirements is also necessary to provide a basis of estimate for resource requirements. Further, such examination identifies the need to expand the range of measurement parameters during development testing. For example, test and evaluation requirements at out-of-band frequencies may be necessary to support EMI/EMC compliance studies. The isolation between payload antennas is required in these studies, and often the antennas are not only out-of-band from their design bandwidth but also within the near field. In such cases, measurements using prototype antennas and a mockup of the satellite structure may be the most expedient means of determining the desired isolation data and should be performed as a part of the development phase rather than as an afterthought during the qualification phase.

Development testing is typically conducted using general-purpose instrumentation and test facilities and techniques that are well established. The emphasis in this program phase is establishing compliance with contractually specified performance. This compliance verification benefits from a well-defined requirements verification matrix. RF performance evaluations require a detailed examination of the components comprising the antenna system design to ensure performance expectations are fulfilled. Often, the expected RF performance of a candidate antenna design is based on projections obtained by exercising available computer codes. Agreement between measurements and these analytic projections add confidence to the performance values. It is not unusual to evaluate design tradeoffs and alternatives at the development phase of the program to finalize the flight hardware configuration.

Other evaluations are required to reduce risk in meeting the design's environmental requirements. Basic evaluations include vibration, thermal vacuum operation, and acoustic survivability (as indicated in Fig. 9-3) for the components comprising the antenna designs. Antenna designs that have deployment and gimbal mechanisms require careful attention in their reliability assessment, operational range of movement,

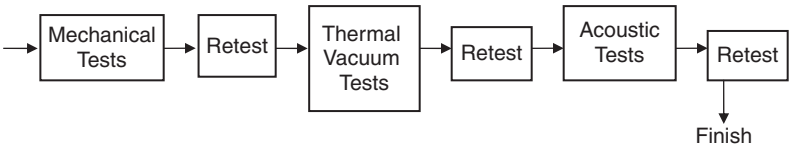


Figure 9-3 The environmental test process [1]

and critical clearance compliance. Design attention is needed to assess potential ESD issues within the antenna, as well as ESD coupling into the payload antennas from external features of the spacecraft. Transmit antennas need to address multipaction susceptibility and PIM generation. Still other evaluations are required to assure compliance with mass and power requirements and to reduce risk in producing flight hardware.

Often, these evaluation requirements result in producing an engineering model of the flight hardware to reduce qualification risk. The engineering model provides not only a means to verify RF performance compliance but also offers a means to test the launch and on-orbit environmental compliance. The engineering model tests for environmental factors can be used to determine design margins without risk to flight hardware. The engineering model is also useful in developing/demonstrating the interfaces with system hardware, software, and control capabilities. Development testing also provides opportunities to evaluate test methodologies and gain experience with test facilities and instrumentation along with their accuracy limitations. The development testing affords opportunities to assess test procedures and test data processing that benefits subsequent qualification testing. The present trend is to develop software scripts to implement such testing, and experience gained through exercising software in development testing will benefit subsequent qualification testing. Finally, it is important to identify those features of the design that could pose potential workmanship issues in flight hardware in order to anticipate testing needs in the qualification phase.

9.2.3 Space Segment Qualification Testing

The qualification phase of the testing addresses two aspects: (1) the compliance of the flight hardware with the performance demonstrated in development testing, and (2) the demonstration that the flight hardware is capable of surviving the launch and orbital environments. The end goal of this testing is assuring that the antennas are flightworthy and comply with system performance requirements. The antenna hardware at this point is generally the flight items. In some cases, where potential risks are identified, a separate engineering development item

is produced that is constructed in the flight configuration for further demonstration as a means of reducing program risk. The RF testing at the qualification phase is more specialized since the goal of the testing is to demonstrate performance compliance and survival in launch and operational environments. The hardware is subjected to environmental testing comprising thermal, acoustic, and shock and vibration loads that emulate the launch conditions and on-orbit environments. Testing is performed over the thermal extremes in a vacuum to establish survival of the on-orbit conditions.

The "hat coupler" mentioned earlier is used to evaluate antenna performance or inject RF signals while subsystems or integrated payloads are subject to environmental testing. The hat coupler encloses the antenna to provide a stable test environment for the antenna isolated from the test facility and incorporates test probes to inject and receive RF tests while environmental testing is ongoing. Hat coupler design is specific to the antenna systems being evaluated. Design and calibration of such couplers are recommended during development testing to avoid qualification testing delays. Attention needs to be applied to the calibration and repeatability of hat couplers so the meaningful data can be obtained in environmental testing. Hat coupler techniques have been used for many years, but the additional complexity of today's antenna designs and the desire for more comprehensive test data provide challenges in hat coupler development. Evaluation of hat coupler designs using engineering model hardware is strongly recommended if reliance on test results is required.

Materials in the antenna that can be subject to failure (e.g., solder connections that may work-harden with thermal variations) should be subject to a sufficient number of mechanical vibration cycles, thermal cycles, and acoustic levels to establish on-orbit reliability. Any active devices in the antenna design are evaluated for lifetime limitations. Generally, all active devices are subject to burn-in requirements that stipulate a required number of hours of operation to screen out infant mortality failures.

Qualification testing differs in detail from development testing both in terms of constraints and requirements. Further, qualification measurements also focus on potential workmanship errors in both the RF and environmental testing. Identification of potential workmanship errors during the development testing thus benefits qualification testing. Like development testing, detailed component-level testing is required for a portion of the components comprising the antenna (e.g., the feed(s) for reflector antennas). However, testing of the overall antenna design must be sensitive to the flight hardware status. Such testing is generally performed in indoor facilities, like compact range or near field sampling facilities, to protect the flight hardware. The RF test objectives focus on

the key performance features of the antenna design in the performance evaluation and demonstrate that the flight hardware performance matches that established during development. The qualification testing puts increased emphasis on system-level parameters compared to development testing, rather than the component-level requirements used in development testing that demonstrate design compliance.

Recent trends in satellite antennas include designs that integrate RF electronics with the antenna assembly. When active electronics are integrated into the antenna system, it is essential to identify the infant mortality in active elements as soon as possible to avoid cost and schedule penalties resulting from replacement of an integrated assembly. In addition to the normal burn-in requirements for such electronics, evaluation of thermal control designs for protecting such electronics is required. Additionally, the test parameters when active electronics are integrated into the antenna system are typically G/T for receive antennas and ERP for transmit antennas rather than gain values for passive antenna designs. The dynamic range of receiving antennas and the linearity of transmit antennas are key performance requirements for such designs, and characterizations to demonstrate design compliance place demands on the required test instrumentation to fully characterize the design compliance.

Additional testing attention depending on the specific implementation is imposed to verify the reliability of antenna deployment mechanisms and gimbals used to reposition the antenna's coverage areas. Such testing is performed at the subsystem and integrated spacecraft levels. Antenna designs that use mechanical positioning in their operation require testing to ensure that the required range of mechanical travel when integrated with the spacecraft can be achieved and that the encoders that measure that positioning accurately report the antenna's boresight positioning. Range of motion testing additionally verifies not only that the motion can be performed but also verifies critical clearance requirements that specify the minimal clearance between the antenna and spacecraft components. Critical clearance values are selected to provide margin to account for thermal changes in dimensions that could result in restricting motion of the antenna. The antenna positioning on the spacecraft is verified by demonstrating the coincidence of the mechanical and electrical alignments. The mechanical and electrical alignments are established during the antenna's RF evaluation, and optical alignment cubes are often attached to the antenna structure so the established alignment can be used during spacecraft assembly and integration.

Environmental requirements are a major part of the qualification testing. The environmental testing objectives are twofold: assure the flight hardware can withstand the launch and orbital environments;

and identify any workmanship errors in the flight hardware. These tests are performed at both the subsystem and integrated payload levels. Environmental testing has three distinct categories, which are identified in Fig. 9-3. The first category of testing evaluates the mechanical integrity of the test article. The test limits and spectral shape of the environment are derived from the launch vehicle characteristics and the mechanical load transfer to the test item. Finite element modeling is used to define these tests, and margins are placed on the test values to accommodate analysis uncertainties. The mechanical resonances of the structure are evaluated and often "tap" tests are conducted, where the structure is struck and accelerometers are used to measure the resonant frequency values. Specially developed shake tables are used to simulate launch loading, and spectral characteristics in three axes provide further verification of the mechanical limitations of the design. The second category of environmental tests is thermal vacuum tests, where test articles are installed in a vacuum chamber and the temperature is cycled multiple times between the projected thermal extremes. The thermal extreme values used in the testing are derived from heat transfer models and, typically, additional margin is added to the analytically projected bounds. The third type of testing is acoustic, where again modeling is used to specify the incident acoustic pressure values and their spectral characteristics. Acoustic chambers capable of providing the test levels of acoustic pressure and spectral characteristics are used to demonstrate compliance. Each of these tests requires specialized assessments to establish the required test parameters. Such assessment heavily relies on analyses techniques, and additional margin is attached to the values projected from analysis.

A typical test flow for qualification tests starts with mechanical tests, proceeds to thermal vacuum tests, and finishes with acoustic tests. The mechanical tests examine the test articles for compliance with requirements, and the presence of workmanship errors such as loose connectors. Thermal vacuum tests cycle the test articles over the thermal bounds several times while monitoring the payload's performance to the extent practical. Operation over the thermal extremes is also a good screen for workmanship shortfalls. Acoustic testing is then performed. Retest is indicated between these three test categories to assure performance is maintained after such testing. The required retest depends on the specifics of the antenna design. For example, VSWR (voltage-standing-wave-ratio) tests for earth coverage horn antennas generally suffice, while more extensive testing is required for more complex antenna designs. The alignment of mechanically positioned antennas, for example, is typically reverified after testing. Such testing typically uses optical cubes attached to the flight hardware. The payload operation during the integrated spacecraft qualification tests is often monitored, and signals are

injected and extracted through test couplers integrated into the hardware or through probes integrated into the hat couplers.

Testing in the qualification phase involves both antenna evaluation on a subsystem basis and payload testing when the antenna is integrated into the payload. Generally, the integrated payload portion of the qualification testing is performed at a system level and the evaluation typically measures the overall system performance (e.g., BER characteristics). The testing at the payload level evaluates the interfaces of the antenna with the payload and the commanding and control of the antennas as appropriate. This testing must include all operational variations in the antenna configuration that are used on-orbit. For example, antenna systems capable of varying their on-orbit coverage characteristics require qualification testing to verify their capabilities. Likewise, commanding to redundant elements and the performance with these redundant elements require verification in the qualification testing. Such testing, or a limited subset of the same, is generally conducted at the payload level during the environmental testing to assure performance is maintained over the thermal extremes. Depending on the antenna system complexity, testing may be required after shipment to the launch site to verify the prelaunch health of the antenna system. Such testing would require the recommended development of portable antenna test facilities that are capable of operation in a launch processing environment.

9.2.4 Space Segment On-Orbit Testing

The on-orbit antenna testing [2, 3] has three objectives: to initially establish compliance of the specified performance; to monitor performance during the satellite's lifetime to assure satellite capabilities are maintained; and to provide the capability to perform diagnostics in the event of performance shortfalls during the satellite's lifetime. The initial on-orbit testing establishes the system-level performance generally identified as KPPs (key performance parameters) addressing G/T, ERP, coverage, and polarization capabilities of the orbited hardware. This database serves not only to establish system compliance but also forms a baseline performance measure to allow data trending as a means to identify performance degradation during the satellite's lifetime.

Diagnostic capabilities are also required to identify and resolve potential performance shortfalls that arise during the satellite's orbital lifetime. These capabilities, together with telemetry data, provide the means to evaluate the satellite's performance state. Clearly, one important task during the satellite development is configuring the satellite in a way that provides an adequate means of determining its on-orbit performance. The planning for such diagnostic capabilities encompasses both the satellite's telemetry readout capabilities and the requirements

for on-orbit diagnostic capabilities reported to the mission control terminals.

Specialized ground terminals are used to perform the on-orbit measurements. These terminals are generally a part of the program's mission control segment. The terminals are configured to exercise the on-orbit payload through a variety of modes within the payload capability. Careful attention is paid to the terminal's calibration so that accurate measurements can be made. In some cases (e.g., EHF), the test terminal requires auxiliary capabilities, such as the ability to measure EHF propagation loss during measurement intervals. Likewise, diagnostic capabilities for the terminal are carefully examined so that observed anomalies can easily be resolved between the test terminal and satellite shortcomings. Future trends in satellite antenna technology include array designs. A feature of array antennas that is commonly touted is a graceful degradation in performance with array element failures. While this graceful degradation applies the antenna gain reduction with element failure, other system parameters such as sidelobe response do not have the same graceful degradation. In cases where the antenna sidelobe performance is an important system parameter (e.g., the isolation required of multiple beam antennas), the sensitivity requirements of the ground terminal are increased to be able to accurately measure the dynamic range imposed by the sidelobe response. The accuracy requirements and the corresponding requirements for received signal power result in test ground terminal requirements that exceed those of typical ground terminals.

9.3 Space Segment Test Issues

While test requirements are specific to each program, generic issues arise for space segment antenna testing. Specialized testing for the space segment is described because such testing is not normally a part of antenna evaluations. Space segment antennas that have broad coverage requirements impose test issues to determine the impact of antenna interactions with the spacecraft structure. Several specialized tests are conducted to evaluate space segment transmitters and include PIM (passive intermodulation) evaluations, and multipactor and corona discharges. Other RF tests address ESD (electrostatic discharge) and EMI/EMC evaluations.

9.3.1 Vehicle Interactions

An issue that commonly arises concerns the interactions of a broad coverage antenna with the surrounding spacecraft structure. A typical example is the TT&C antennas used on every satellite. These antennas typically are required to provide hemispheric coverage, with one antenna

providing forward coverage and a second antenna providing rearward coverage. These systems normally use low microwave frequencies, and the corresponding long wavelength results in challenges in isolating the antenna from its surrounding spacecraft structure. The issue is to distinguish the antenna coverage in the presence of the satellite from the antenna coverage in free space. In the case of the TT&C antennas, a typical specification is to exceed a minimum antenna gain level over a specified (e.g., 95%) portion of their respective hemispheres. The broad coverage from simple antenna designs often results in high backlobes that interact with the spacecraft's structure, and these interactions distort the pattern coverage by combining in and out of phase with the direct illumination of the antenna. This sensitivity to satellite interaction often results in the TT&C antenna being placed on a mast in front of the spacecraft structure.

Design and analysis codes are widely available for a diverse class of antenna designs. These analysis capabilities, however, address the RF performance of the antenna itself in free space, and further development is necessary to extend the capabilities to include the antenna interactions with the spacecraft structure. Measurements, likewise, can be readily performed on the antenna itself, but measurements of the antenna mounted on the spacecraft present challenges. Full-scale measurements require a large test facility, and because the antenna excites the spacecraft structure, large far field distances result, increasing measurement facility requirements. In most cases, deployment of solar arrays and their positioning over the range of sun tracking angles is not possible, so antenna interactions with solar arrays cannot be included in full-scale measurements. Some relief from these problems can be obtained by performing scale-model measurements, but attention is required to obtain valid scaled measurements. Further development and application of analyses techniques and measurement methods are recommended to quantify vehicle interaction effects and facilitate the selection of antenna placement on the vehicle. Similarly, developing broad coverage antenna designs with reduced backlobes and a reasonably compact size (as suggested in Chapter 2) compared to existing designs also increases the ability to isolate the antenna from the spacecraft.

9.3.2 Transmitter Issues

Space segment transmitters receive extensive evaluations to establish compliance with their RF and environmental requirements. Part of the system evaluation examines the ability to maintain the desired transmitter operating points. When active transmit array designs are used, their evaluation is generally performed in specialized facilities where the RF performance can be established and the thermal control designs

can be evaluated. Other specialized transmitter issues concern PIM products, multipaction, and corona that require specialized measurement techniques.

Space segment antennas generally simultaneously communicate multiple user signals, and system linearity is required to avoid degradation resulting from intermodulation between signals. The design operating point selected for the transmitter controls the intermodulation product generated by the transmitter. However, nonlinearities in the connection between the transmitter and antenna and the antenna itself can also produce intermodulation products—in this case, PIM products. Normally, antennas are thought of as passive components, but when sufficient power densities illuminate junctions, PIM products can result. Dissimilar metals and/or contamination in these junctions form weak diodes that generate the intermodulation products between multiple carrier components. Mitigation techniques to reduce PIM levels include avoiding contacting joints and using electroforming techniques to fabricate RF components, such as diplexers, to minimize the number of component junctions and thus reduce the generation of PIM products. Addressing potential PIM issues early in the development phase on prototype hardware is strongly recommended to assess the design susceptibility to PIM generation. Further evaluation of PIM levels is required in qualification testing on the integrated flight hardware because, in most cases, PIM products result from workmanship issues.

PIM testing [4, 5] uses the test transmitters to provide two tones whose frequencies are selected to provide known intermodulation frequencies within the transmit band and within the bandwidths of on-board receivers. These tones are injected into the passive circuitry between the transmitter location and the radiating element that is the transmit path containing mechanical junctions having the potential to generate PIMs [6, 7, 8, 9]. The test transmitters deliver power levels that exceed those used operationally by 6 dB, and spectrum analyzer instrumentation is used to determine the potential presence of PIM products. The PIM levels can be evaluated in two ways: one method references the levels to the transmitter power (e.g., dBc), and the second method measures the PIM levels received by a probe antenna. Care must be taken to filter the transmitted tones to avoid exceeding the dynamic range limitations of the spectrum analyzer. A small test facility capable of withstanding the test power levels is required, and the facility itself must be measured to assure that PIM products generated by the facility do not obscure the PIM levels to be measured in the test.

Multipaction [10, 11, 12] and corona [13, 14, 15] are two phenomena that can damage transmit antennas and RF components between the transmitter and antenna. Multipaction results when a sufficiently high electric field bridges a gap and strips electrons from the surfaces, creating

an avalanche. Multipaction can result in component damage because the component surfaces are eroded. The multipaction discharge depends on the RF frequency, gap spacing, secondary surface electron emission, and the field strength. Corona is sometimes referred to as microwave breakdown and also results in an avalanche condition when the free electrons accelerated by the RF fields ionize gas molecules producing a plasma. The avalanche results when the ionization exceeds the electron diffusion into regions of lower density. Corona depends on the field density and gas pressure, and for space applications, attention to component venting is necessary to avoid corona issues during the early phases of on-orbit operation. Systems operated during launch assent (e.g., TT&C subsystems) need to pay particular attention to venting and gas pressures to assure corona does not occur. The vacuum conditions in space, assuming proper venting, are commonly used to assess corona susceptibility of satellite transmission components. Typically, both multipaction and corona occur at high field densities at the minimum separation between surfaces such as the minimum inner dimensions of filters.

The vulnerability of systems to multipaction and corona is initially assessed through susceptibility analyses and depends on the outcome of these analyses; testing may be required to establish the design's suitability. Susceptibility analyses proceed by determining the minimum device dimensions and the field strengths across these surfaces. If the susceptibility analyses indicate transmitter power levels at least 6 dB higher than those used operationally, verification testing is not required. When the margin of the transmitter power level is 6 dB or less, testing is necessary to verify the actual susceptible transmitter power levels. Such testing requires test transmitters whose power output exceeds the operational transmitter power output by at least 6 dB. The components under test must be located within a controlled vacuum, and bell jar configurations are typically used in these measurements.

9.3.3 Other Electromagnetic Measurements

ESD [16, 17] results when on-orbit charge accumulation on dielectric materials discharges. Such charging is most pronounced during geomagnetic storms. Common culprits for charge build-up include dielectric materials commonly used in protective thermal covers both for antennas and for other spacecraft components. If the materials are allowed to accumulate charges, a discharge eventually occurs. The spectrum of discharges has a peak value at relatively low frequencies, and reduces away from the peak spectral levels. Such discharges have significant energy even at frequencies high enough to couple into payload antennas. Both interference and damage to RF electronics can result.

A conducting path is required to preclude charge accumulation that would result in ESD. This conducting path is provided by adding

conductivity to the dielectric material and providing a conducting path from the dielectric material to the spacecraft ground. For antenna thermal covers, however, the conductivity must be adequate to avoid charge accumulation while not having sufficient insertion loss that would degrade the antenna's RF performance. In addition to conductivity within the material, specially developed ESD paint has been used to reduce surface charging. In addition to the surface treatments, the integrity of a ground path between the dielectric material and the spacecraft ground must be maintained to bleed any accumulated charges.

Testing for ESD susceptibility requires a means to determine the surface charging properties of the material and to evaluate the effectiveness of the grounding plan to mitigate ESD. Specialized facilities capable of illuminating the material with an electron beam are used to determine whether or not sufficient material charging exists to produce a discharge. A probe antenna and spectrum analyzer instrumentation are used to measure the levels and spectrum of discharges. The probe is located at a specified distance from the material discharge, and the probe location is selected by the location of spacecraft receivers and the surfaces that can accumulate the static charge. These measurements are often performed on samples of the dielectric materials. In addition to the dielectric material issues, installation in the spacecraft also requires addressing the integrity of ground strap designs that connect the dielectric material to the spacecraft ground. When the material is used to protect antennas from the thermal environment, a further examination should be made to determine if the antenna performance can be degraded if the thermal cover is drawn close to the antenna by differential electrostatic forces.

Additional electromagnetic measurements are required in evaluating EMI/EMC issues. EMI/EMC is a spacecraft-wide activity to assure that the operation of satellite subsystems is not limited by interference to or from other subsystems. At the initiation of the satellite's development, the frequency plan for the entire spacecraft should be examined to identify the fundamental and harmonics of the spacecraft's frequency components and their coupling to other spacecraft subsystems. Such coupling involves a large number of potential paths because both fundamental and harmonic frequency components must be included in the assessments. The number of potential paths can easily exceed 10,000 in today's spacecraft designs. Estimates of the coupled levels and the subsystem susceptibility to interference are then compared. The number of potential paths with possible concerns dramatically drops, typically to less than 100. Those paths with more than a remote susceptibility should also be examined to identify possible workmanship issues that could produce a concern. The efforts put forth in conducting these analyses benefit the subsequent qualification testing that is conducted on component, subsystem, and integrated spacecraft levels.

While the internal payload electronics assessment is a significant activity, the EMI/EMC issues for the spacecraft antennas provide design and test concerns that result in a significant activity as well. The antennas are generally isolated from the internal EMI/EMC sources by the satellite bus shielding. The operating frequencies and other frequency sources used in frequency conversion are the principal concerns in addressing antenna system EMI/EMC issues. Particular attention should be given to those potential paths where workmanship can compromise the EMI/EMC susceptibility. In many cases, the coupling between payload antennas occurs in the near field of these antennas because of the limited physical separation of the antennas mounted on the spacecraft. Estimated values of antenna near field coupling provide a preliminary assessment of EMI/EMC susceptibilities. If potential concerns result from these preliminary estimates, measurements of the actual antenna and mockups of the surrounding spacecraft structure can be used to refine the preliminary estimate. These measurements are required to determine the out-of-band response of the antennas as well as address near field coupling at out-of-band frequencies. Initial measurements during development can be performed to address the potential of EMI/EMC susceptibility.

The integrated flight hardware is measured in specialized EMI facilities. The ambient RF environment of these facilities is established with the test sets activated to identify facility and equipment RF levels that would degrade or obscure required measurement sensitivities. The spacecraft hardware is then activated and tested for all operational functions to assure EMI/EMC compliance. Particular attention is focused on the previously identified paths with potential shortfalls and those areas having potential workmanship issues. Often such areas are examined with probes and spectrum analyzer instrumentation to identify sources of leakage. A general survey of RF emissions is also required to assure unanticipated emissions are not present.

9.4 User Segment Antenna Testing

The evaluation of user antennas is generally straightforward. User antennas are not subject to the space qualification and flight hardware protection requirements of space segment antenna testing. User antenna evaluations can generally capitalize on general-purpose facilities and instrumentation capabilities. Unlike the space segment antennas that cannot be serviced and maintained on-orbit, reliability does not have the same critical importance. Nonetheless, user antennas provide some test challenges and requirements peculiar to specific applications. For example, shipboard installations impose environmental requirements and a means of platform motion compensation typically sensed by

gyroscopes must be implemented. Thus, user terminal testing, like space segment testing, can also be a multidimensional problem depending on application requirements.

The basic antenna system parameters for user antenna evaluations are the G/T and ERP levels and the ability to track the satellites as appropriate. In many cases, the antenna and the system electronics are measured separately, but testing of the integrated terminal may be required. Antenna gain, pattern, polarization, and impedance properties are generally established during development measurements. The system testing is directed towards establishing the performance of the integrated design. The system's G/T depends on both the antenna and receiver noise contributions. Such measurements can be performed using radio source techniques or separate measurements of the antenna temperature and receiver noise temperatures. Additional receiver measurements are made to establish its implementation loss, a parameter needed in link budget assessments. Such measurements are performed by BER test set instrumentation or special test sets incorporating operational modulation modems. The ability of the receiver to acquire the satellite signal at a specified signal level is also established. Receiver operation over the required range and rates of Doppler frequencies must also be demonstrated.

The system's G/T performance depends on the antenna's gain and the system noise temperature, which is comprised of the antenna noise temperature and the receiver noise temperature. The antenna's gain and the system noise temperature must be determined at the same terminal plane as discussed earlier. Antenna noise temperature varies with elevation angle and so a reference elevation angle for G/T evaluation must be provided in the specification. The G/T specification is problematic for EHF systems where the antenna noise temperature varies with weather conditions because of the path loss variation.

The system's ERP performance depends on the antenna's gain and transmitter power output. Typically, the antenna and transmitter are evaluated separately. The rated output of the transmitter is verified and the design operating point to maintain adequate linearity is established. Any intermodulation products, spurious responses and harmonics in the transmitter output power are assessed. When integrated with the antenna, the system's power handling capability must be verified, particularly when high-power transmitters are used. The power handling capability addresses thermal rises of components and the possibility of arcing. Some transmitters are protected from reflected power levels by isolators, while others are protected by measuring the reflected power in a coupler. The transmitter protection circuitry when a coupler is used is based on operation to a maximum VSWR level that, if exceeded, turns

the transmitter off to avoid transmitter damage. When the transmitter is routed to the antenna through rotary joints, the VSWR at all possible positions must be established and compared to the VSWR value that turns off the transmitter. Finally, the out-of-band noise characteristics of the transmitter must be established and the radiated ERP levels for these out-of-band components must be determined to assure the system does not interfere with other systems. When integrated with the antenna, further testing is required to demonstrate proper operation of the diplexer. A simple test (along with a means of specification) is to measure the change in the receiver's noise level when the transmitter is turned on and off. Further testing is needed to insure the transmitter does not saturate the receiver at out-of-band frequencies.

Larger antennas require evaluation of the antenna acquisition and tracking performance. As will be discussed, these antennas often have excessive far field requirements, but may have insufficient performance to allow evaluation using radio stars to determine antenna gain characteristics. Measurements are required to evaluate antenna pointing accuracy and dynamic tracking accuracy to maintain antenna alignment with the satellite. The actual satellite, if available, provides an opportunity for such testing. For antenna system applications for reception from low-altitude satellites, the ability of the receiver and antenna to acquire the signal in a specified time period or by a specified elevation angle requires demonstration. The acquisition and tracking evaluation also needs to verify that the antenna has acquired the satellite on its main beam rather than a sidelobe. The coincidence of the receive and transmit boresight axes also must be demonstrated.

Two different classes of user antennas exist and the testing methodologies differ in the two cases. One class of user antennas is produced in a high volume, and a test methodology capable of providing high confidence in the results while being conducted in an efficient manner is required. The second class of user antennas has a limited production volume. These user antennas are typically large fixed designs that have specialized requirements to meet specific applications. The antenna testing for this class of antennas is generally performed for each antenna system.

The first class of user antennas is produced in large volumes and typically uses relatively small reflector antennas for cost and performance reasons. A key challenge for this class of terminal designs is accomplishing testing on a production basis with high confidence, while at the same time controlling costs. Detailed testing of individual production items for this class of user antenna would impose impractical time and resource requirements. The cost limitations of such terminals also depend on implementations based on design simplicity using established techniques with previous operating experience. Testing for this

class of user antennas proceeds by demonstrating that a "first article" complies with system requirements, and then devising test techniques to assure production items maintain the performance achieved by the first article by testing a limited number of randomly selected production units.

For example, antennas for VSAT applications are generally simple reflector designs that can be manufactured inexpensively, satisfy established manufacturing tolerance and sidelobe envelope requirements, and have previous experience of operation in a variety of environments. In many cases, the antenna's reflector and positioner are available as COTS items. The relatively broad beamwidths of these designs generally preclude antenna tracking requirements, and since the users typically communicate with geosynchronous satellites, fixed antenna pointing suffices. The feed systems are again straightforward designs for commonly using frequency ranges such as C- and Ku-band. The principal challenges in testing for such designs is assuring that the feed manufacturing maintains the required polarization purity since satellites commonly used for VSAT applications use orthogonal polarization to obtain the capacity benefits of polarization reuse. The testing for such designs is principally in the development stages, with a limited amount of spot testing during production to assure quality control. For example, polarization purity imposes mechanical tolerances on the feed, and depending on the manufacturing methods—tooling wear, for example—this can result in increased cross-polarization levels, so that mechanical measurements on a spot testing basis can be used to assure compliance at a production level.

The second class of user antennas includes larger ground terminals that have a limited production volume and must satisfy requirements specific for these applications. Such designs generally require testing each antenna. These designs have more complex testing requirements than the first class of high-production antennas as a consequence of their performance requirements. Program applications for these larger ground terminals include TT&C service, mission control, user segments for communication satellites, gateway terminals for communication satellites, very high data rate communications, terminals to upload data for direct broadcast services, and readout (receive only) terminals for mission data from direct broadcast satellites. While each application has specific requirements, generic test elements exist. For example, mobile users have requirements to maintain connectivity with the satellite that are not an issue with fixed terminals. Large ground terminals support satellite programs as they evolve over a long lifetime. Reliability and obsolescence, in addition to updated requirements and capabilities, dictate component replacements, which in turn require interface definition and system testing. Further, today's requirements include remote

operation in many cases, and so additional attention must be focused on BITE (built-in test equipment) capabilities, redundant components, and sparing strategies.

9.4.1 User Segment Measurement Processes

The process for user segment antenna testing in Fig. 9-4 differs from the space segment antenna testing described in Fig. 9-1. Relative to the two classes of user antennas, the test process for high-production volume designs applies to the compliance demonstration of first article design, while the test process for the second class of user antennas having a limited production follows the test process for each item. Space segment testing has three distinct test phases: development, qualification, and on-orbit. User segment testing has four distinct test phases: demonstration, qualification, integration/acceptance, and sustainment. Each phase has differing objectives and requirements, as shown in the following list.

1. Demonstration testing is comprised of evaluating alternative vendor COTS products that largely comprise the system design.
2. Qualification testing addresses the performance of components that are specifically developed for the system's application.
3. Integration/acceptance testing provides an evaluation of the overall system's compliance with system requirements.

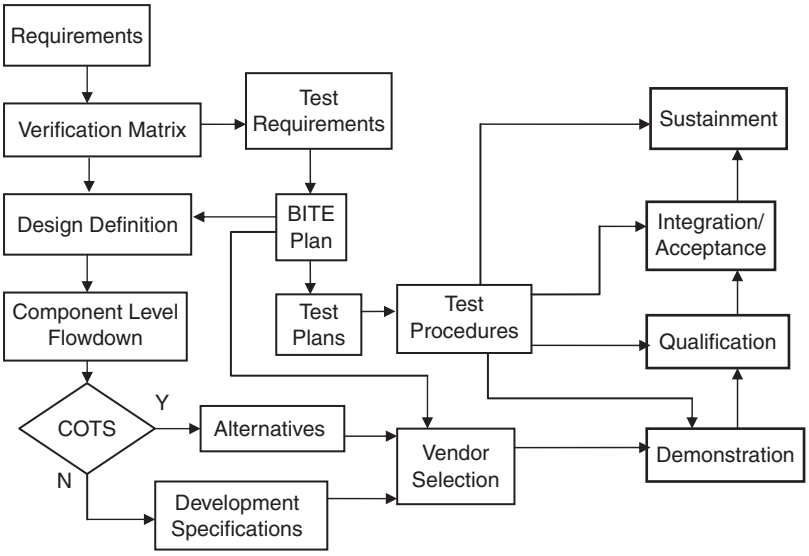


Figure 9-4 The overall user segment test process

4. Sustainment testing provides assurance that performance levels determined in integration/acceptance testing are maintained and that system upgrades to enhance performance and replace obsolete elements of the design comply with updated requirements.

As with the space segment testing, the process for user segment testing starts with the system requirements, and the elements of the requirements verification matrix are derived from the system requirements. Clearly, system requirements need to be stated in a straightforward way that allows the requirements to be tested. Like the space segment testing, the requirements verification matrix has four verification categories: (1) inspection, (2) analysis, (3) demonstration, and (4) test. Testing is used on either a component or an integrated assembly basis and typically involves the use of other instrumentation and detailed data collection, as well as analyses to establish compliance with requirements.

The design definition evolves directly from the requirements and verification matrix. The overall design must be developed to satisfy the system-level requirements and the candidate components of the overall system design are examined. The requirements form a basis to flowdown the performance needed by the system's components to yield a compliant design. Many of the individual components and subsystems comprising user segment systems are available as COTS products from a variety of vendors, and alternative offerings must be assessed. These assessments must carefully address the appropriate items in the requirements verification matrix. Generally, the testing takes the form of inspection of the COTS products and selection from offerings of alternative vendors. In other cases, development is needed that is specific to the system's requirements and implementation. For example, the filters are typically specific to an individual system, and specifications for their required development must be prepared. In this case, the testing is performed in the qualification phase to assure the requirements are satisfied.

Equally important, the requirements verification matrix also establishes the test requirements for the system. These test requirements must identify the methodologies used in the testing. As appropriate, candidate methodologies are compared to assess measurement uncertainties, and in some instances, more than one methodology might be used to increase confidence in the results. Attention should be paid to develop error budget projections to assess measurement uncertainties and to address the necessary measurements to quantify the uncertainties as required. Another very important part of test requirement development is the identification of system test points and their adequacy in all test phases of the program.

Two factors have a major impact on the required emphasis for testing. Independent of the application, the system design must operationally minimize downtime and often, minimal downtime has financial incentives. Accordingly, significant attention to BITE (built-in test equipment) capabilities is essential to have the capability to resolve potential shortfalls rapidly in system operation. Attention to test techniques that may be conducted on a non-interference basis during system operation also minimizes system downtime. Clearly, sufficient instrumentation must be available to diagnose system shortfalls. Additionally, significant attention must also be paid to redundancy, and strategies for onsite sparing must be developed with the objective of minimizing system outage. The second factor is the desire in many cases to operate the terminal remotely as a means to minimize operating expense. The ability to diagnose system operation and shortfalls by remote command also requires significant attention and planning to include an appropriate amount of redundancy.

Addressing test requirements at the initiation of system development is exceedingly important. Early resolution of test requirements provides the necessary definition in the design definition, establishes the necessary test points for both evaluation and sustainment, develops an effective BITE capability, provides redundancy to maintain operation with equipment shortfalls under remote command, and makes spares available to minimize downtime for other failures. Examination of test requirements also provides the necessary information for subsequent development of test plans and procedures.

The test plans and procedures address the details of the required testing. The plans state in general terms the detailed objectives of the tests, their scope, schedules and interrelationships, the elements of the requirements verification matrix that are satisfied, and a summary of the methodology to be used. The test procedures define the equipment required, a detailed description of the methodology or methodologies, the steps to be followed in the test, an error budget projection along with a description of the supporting measurements to determine measurement uncertainties, formats to record the data, data processing requirements, the capability to record any discrepancies discovered during the testing, and the elements of the requirements verification that are covered by the test procedures. Procedures for each of the four system development test phases are required.

An overview of the four test phases, shown in Fig. 9-5, summarizes their differences. Development testing is performed at a component level to verify the development objectives, including performance and environmental issues. Interface verification with the system also needs to be addressed. Generally, development testing is performed using general-purpose instrumentation and facilities. Qualification testing is performed

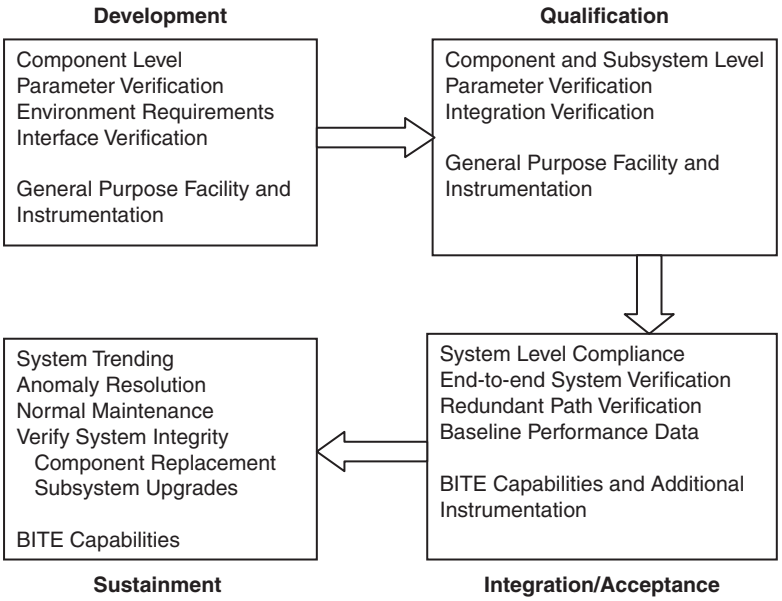


Figure 9-5 An overview of user segment test phases

at component and subsystem levels and differs from development testing in the emphasis on integration with other system components and on operational compatibilities. Generally, qualification testing is also performed using general-purpose instrumentation and facilities. Integration and acceptance testing is carried out at the system level and is more formal and detailed than development and qualification testing. The end objective of this testing is to demonstrate compliance with the requirements verification matrix. Integration and acceptance testing must exercise all redundant paths and operational modes and relies heavily on BITE capabilities augmented with other test instrumentation as appropriate to validate BITE results. Demonstration of remote operation is an important objective if it is a system requirement. Data from integration and acceptance testing form a baseline data set to be used throughout the system's lifetime and thus documentation of test results in an archival form is required. Finally, sustainment testing is used operationally and has the key objectives of data trending, the resolution of system anomalies that may occur, supporting maintenance, and system integrity verification after component replacement. Sustainment testing relies on BITE capabilities, a factor that further increases the importance of an effective BITE capability. Large ground terminals have relatively long lifetimes because of their high cost, and as time passes, existing equipment needs upgraded capabilities and replacement of obsolete subsystems for

maintainability and reliability reasons. The evolution of ground terminal equipment poses additional sustainment test requirements that are subsets of the overall testing described here.

Like space segment testing, system review processes are ongoing with the hardware testing process. The review process is comprised of an SRR (system requirements review) held soon after the system development is initiated, a PDR (preliminary design review) held at the completion of the design definition, and a CDR (critical design review) held at the completion of demonstration and qualification testing to assure the system design is sufficiently mature to warrant proceeding to integration and acceptance testing. Such reviews are applied to the overall system design reviews and to development components supplied by vendors. Formal system sell-off occurs at the completion and documentation of the test activities. All system testing is proceeded by a TRR (test readiness review) to ensure that test objectives, test methodologies, test documentation, and the elements of the requirements verification matrix to be verified by the testing are understood by all parties.

9.5 User Segment Test Issues

The testing of user segment antennas depends on their performance capabilities and requirements, which are specific to their application. Specific test requirements consequently depend on the antenna design. The most general test requirements pertain to larger antenna terminals, and other terminals are subsets of those test requirements. One of the generic test issues concerns the evaluation of the larger user antennas that have excessive far field requirements. Recommendations are furnished on test parameter requirements and test elements, as well as the problems involved in measuring larger user antennas.

9.5.1 Test Requirements

Both component-level and system-level testing is required for user antennas. General test requirements can be grouped into the three categories (illustrated in Table 9-1). The aspects encompassed in the

TABLE 9-1 Typical Aspects of Antenna Testing

RF	Mechanical	Safety
Antenna Gain	Positioning	Interlocks
Antenna Pattern	Range of Travel	Transmit Operation
Polarization	Accuracy	Aircraft Warning
Impedance	Response Time	
Tracking	Wind Loading	
Power Handling	Tracking Response	
RF Hazard	Thermal	
EMI/EMC		

testing include the normal evaluation of the antenna's RF parameters, addressing mechanical requirements, and insuring safety features perform properly.

The RF parameters for user segment testing include antenna gain, pattern, and polarization variations over the operating receive and transmit bandwidths. These component values are used to establish the system's ERP performance for transmit operation as well as the G/T performance for receive operation. Ground terminal antennas must also comply with the sidelobe envelope requirements (described in Chapter 7) that are intended to reduce interference between systems. In applications where orthogonal polarizations are used to communicate independent data streams to increase the system's data throughput, particular attention must be paid to evaluating polarization purity to assure that adequate isolation exists to avoid co-channel interference for such operations. Antennas must be integrated with other system electronics and the interface impedances must be measured. Antenna tracking designs must be evaluated to assure accurate alignment with the signal's direction. The antenna system must be capable of operating with the required transmit power levels. The transmit operation imposes power handling requirements that must be demonstrated for both the antenna feed and the connecting transmission lines between the antenna field and transmitter output terminal. The transmission lines typically include waveguide and often have rotary joints to allow for the necessary mechanical motions when the transmitter is not located on the antenna. In most cases, flexible waveguide or coaxial cable sections are provided for strain relief to compensate for thermal variations in dimensions. High-power transmitters used in ground terminal designs can result in radiation hazard concerns that also need to be quantified. The antenna and its electronics also must satisfy EMI/EMC requirements, and particular attention must be paid to the radiated emission of signals other than the desired transmitted spectra.

In addition to RF evaluations, mechanical requirements must also be addressed. The antenna must be able to point to the satellites accurately, to follow the dynamics of the signal's direction changes over the required range of motion, and to move from one satellite to another in a specified time. When the antenna is not enclosed in a radome, the antenna must maintain tracking alignment with the signal while perturbed by the wind. Procedures to address antenna wind loading effects [18] are commonly followed. At extreme wind velocities, the antenna survival becomes an important issue; generally, manufacturers stipulate a wind velocity that encourages operators to position the antenna to a "stow" position (i.e., a 90^{o} elevation angle) to minimize wind loading. A still higher wind velocity is stipulated for survival. Finally, the antenna

must be able to maintain operation over the thermal extremes and other environmental requirements specific to user applications.

Other requirements pertain to safety features that are specific to an antenna's requirements and design. The safety requirements include system interlocks for the antenna positioning and transmitter operation. Larger antennas provide the capability to be positioned by commands from the antenna control system for normal operation. A provision is also made for "local control" that allows the antenna to be positioned so it can be serviced or so access can be gained to RF equipment mounted on the antenna. Under local control, safety interlocks are provided so the antenna cannot be commanded by personnel outside of the immediate vicinity of the antenna, in order to avoid accidental injury to personnel working on the antenna. In addition, "kill" switches are placed on the antenna in locations where personnel can logically be located. These "kill" switches freeze the antenna positions. Transmitter operation imposes other safety requirements. In designs where radiation hazard limitations are even a remote possibility, indicators are required to report when the transmitter is operating. In many designs with high-power transmitters, waveguide switches are used to select connection to the antenna or connection to a dummy load employed in transmitter diagnostic measurements. Additional indicators are provided at the dummy load locations when the transmitter is terminated by dummy loads. Very large antennas may also require aircraft warning lights.

9.5.2 Measurement of Large Ground Terminal Antennas

Large user antennas impose special requirements for their measurement. The antenna size is often larger than can be accommodated by conventional general-purpose antenna measurement facilities, and the far field distances often exceed available real estate limits. The required far field distance can extend to thousands of feet for the larger ground terminal antennas.

The measurement process for such antennas generally proceeds in two phases. One phase is a detailed evaluation of an assembled antenna at a vendor location. The process is to perform a detailed measurement of the antenna feed that can be accomplished in a general-purpose measurement facility (because of the feed's small size), an analytic computation of the overall antenna performance using existing validated computer codes, an examination of the reflector surface to ensure its precision does not limit the RF performance, and the performance characterization of the overall antenna. The objective of this first phase is to establish the compliance of the antenna's RF design with the system requirements. Once the antenna's design compliance

is established, the second test phase is installation tests performed at the antenna's operational site. The objectives of the installation tests are to demonstrate that the antenna has been correctly assembled at the site and that its position is properly indexed in inertial space to minimize pointing errors.

The objective of installation testing is to demonstrate that the antenna system has been correctly assembled and that its performance complies with the levels determined in the vendor tests. Measurements of the antenna and system electronics should duplicate at least one of the methodologies demonstrated at the vendor location that will also be used in the future to demonstrate that system performance is maintained over the system's lifetime. The vendor testing should endeavor to use alternative measurement methodologies and supporting analyses to provide confidence in the measured performance. The installation tests are therefore a subset of the vendor tests. The installation tests also provide baseline performance data that can be used in subsequent maintenance testing to determine if the system performance has been maintained during the antenna's lifetime. Separate evaluations and methodologies of the system electronics are also performed during the installation tests, and like the antenna, performed during the system's lifetime.

In many cases, systems using antennas in this category are operated remotely, and an important part of the installation testing is demonstration that system shortfalls can be identified by commanding both BITE resources and the substitution of redundant equipment to replace failed units in order to maintain system operation. The RF testing is then typically performed using radio source techniques, and comparison of the site results with those previously performed in the vendor tests is used for acceptance.

Very large user antennas can be evaluated by radio source techniques using radio stars such as Cassiopeia A, Cygnus, and Taurus. At the other extreme, small user antennas can be measured by using the sun as a source. However, the sun subtends a 0.5^{o} angular width, and because of solar flare activity, the flux density is not uniformly distributed over the solar disk. Antennas between these two extremes present test challenges. One alternative radio source is the moon [19], which like the sun is an extended source, but which unlike the sun does not have active regions. In some cases, specifying the antenna's gain, or rather its G/T value, is more straightforward because the weather-dependent antenna noise temperature is no longer an issue. The moon's EHF emission varies with the lunar cycle. Another alternative [20] uses satellite beacon signals employed in many programs. Such beacons are commonly used on satellites that utilize polarization reuse to increase system capacity and which emit CW tones with high polarization purity,

an attractive feature if the antenna being measured also has polarization purity requirements that can be evaluated by measuring the beacon response.

A comparison of three measurement techniques [21] was made, which included radio source, satellite signal, and boresight tower measurements. Boresight towers are a far field measurement technique where a test illuminator is located on a distant tower. Field probing measurements and adjustment of the illuminator height provide the means to establish the uniformity of the test field. Such field probing is commonly done at vendor locations, but trying to establish and maintain a boresight tower facility at an operational site has often proven to be problematic. A variety of satellite signals have been used and have significant utility for measuring the downlink receiving performance of the user antenna, but which do not provide a signal source to evaluate the uplink transmitting performance of the antenna. No one technique is universally advantageous.

Another technique [22, 23] for measuring user segment G/T values is based on comparative measurements. The measurement proceeds by calibrating a smaller antenna and comparing its response with a larger antenna that has excessive far field distances, has a beamwidth smaller than the 0.5° width of the solar disk, and has insufficient performance to use other radio sources such as Cassiopeia A. Both radio source measurements and coherent signal source measurements were considered in the comparative measurements. As an example, assume a 16-ft antenna is to be measured at Ku-band frequencies—its beamwidth is about 0.3° and its far field distance is about 1.4 miles. If a 4-ft antenna is used as a comparative measurement, its far field distance is about 460 ft, but other measurement facilities, such as near field sampling and compact ranges, are available for its calibration. Two types of measurements can be performed: antenna gain comparison and antenna G/T comparison. The antenna gain is a gain by comparison technique that has been discussed in Chapter 8. A spectrum analyzer is assumed to be used as a receiver for these measurements. The antenna gain is based on comparative signal power measurements, and the insertion gain between the antenna terminals and the indicated spectrum analyzer signal level must be established by signal injection techniques for both antennas. The antenna noise temperature and receiver noise temperature must be separately measured to obtain the G/T value. The second type of measurement is a "G/T by comparison" where the SNR values of both antennas are compared to establish the G/T differences. If the antenna gain differences are required, the system noise temperature differences for each antenna must be measured. As a practical matter, performing both types of measurements involves little extra effort and

provides more confidence in the results, as well as additional insight into the performance of the antenna system being measured.

References

1. R. B. Dybdal, "Space Qualifying Antennas," *2007 AMTA Symposium Digest* (November 2007).
2. R. B. Dybdal, "On-Orbit Communication Satellite Measurements," *2002 AMTA Symposium Digest* (November 2002).
3. R. B. Dybdal, "ERP Measurement Issues," *1998 AMTA Symposium Digest* (October 1998).
4. J. W. Boyhan, H. F. Lenzing, and C. Koduru, "Satellite Passive Intermodulation: System Considerations," *IEEE Trans Aerospace and Electronic Systems*, vol. 32 (July 1996): 1058–1063.
5. J. W. Boyhan, "Ratio of Gaussian PIM to Two Carrier PIM," *IEEE Trans Aerospace and Electronic Systems*, vol. 36 (October 2000): 1336–1342.
6. J. Henrie, A. Christianson, and W. J. Chappell, "Prediction fo Passive Inermodulation From Coaxial Connectors in Microwave Networks," *IEEE Trans Microwave Theory and Techniques*, vol. 56 (January 2008): 209–216.
7. C. Vicente and H. L. Hartnagel, "Passive Intermodulation Analysis Between Rough Rectangular Waveguide Flanges," *IEEE Trans Microwave Theory and Techniques*, vol. 53 (August 2005): 2515–2525.
8. C. Vicente, D. Wolk, H. L. Hartnagel, B. Gimeno, V. E. Boria, and D. Raboso, "Experimental Analysis of Passive Intermodulation at Waveguide Flange Bolted Connections," *IEEE Trans Microwave Theory and Techniques*, vol. 55 (May 2007): 1018–1028.
9. P. L. Aspden and A. P. Anderson, "Identification of Passive Intermodulation Product Generation on Microwave Reflecting Surfaces," *Proc IEE*, vol. 139 (August 1992): 337–342.
10. A. D. Woode and J. Petit, "Design Data for the Control of Multipactor Discharge in Spacecraft Microwave and RF Systems, *Microwave Journal* (January 1992): 142–155.
11. ..., "Space Engineering: Multipacting Design and Test," vol. ECSS-20-01A, ESA Publication Division, May 2003.
12. C. Kudsia, R. Cameron, and W. C. Tang, "Innovations in Microwave Filters and Multiplexing Networks for Communication Satellite Systems," *IEEE Trans Microwave Theory and Techniques*, vol. 40, (June 1992): 1133–1149.
13. D. Anderson, U. Jordan, M. Lisak, T. Olsson, and M. Ahlander, "Microwave Breakdown in Resonators and Filters," *IEEE Trans Microwave Theory and Techniques*, vol. 47, (December 1999): 2547–2556.
14. U. Jordan, D. S. Dorozhkina, V. E. Semenov, T. Olsson, D. Anderson, M. Lisak, J. Puech, I. M. Nefedov, and I. A. Shereshevskii, "Microwave Corona Breakdown Around Metal Corners and Wedges," *IEEE Trans Plasma Science*, vol. 35, (June 2007): 542–549.
15. U. Jordan, D. Anderson, L. Lapierre, M. Lisak, T. Olsson, J. Puech, V. E. Semenov, J. Sombrin, and R. Tomala, "On the Effective Diffusion Length for Microwave Breakdown," *IEEE Trans Plasma Science*, vol. 34, (April 2006): 421–429.
16. C. Bowman, A. Bogorad, G. Brucker, S. Seehra, and T. Lloyd, "ITO-Coated RF Transparent Materials for Antenna Sunshields—Space Environment Effects," *IEEE Trans Nuclear Science*, vol. 37 (December 1990): 2134–2137.
17. H. C. Koons and T. S. Chin, "Broadband RF Spectrum for Electrostatic Discharges on Spacecraft," *Aerospace Corp Tech Rept TR-93(3940)-6* (May 1, 1993).
18. ———, "Electrical and Mechanical Characteristics of Earth Station Antennas for Satellite Communications," *Electronic Industries Alliance/Telecommunications Industry Rept TIA/EIA-411*, Revision A (September 1986).
19. D. A. Guidice and J. P. Castelli, "The Use of Extraterrestrial Radio Sources in the Measurement of Antenna Parameters," *IEEE Trans Aerospace and Electronic Systems*, vol. AES-7 (March 1971): 226–234.

20. R. B. Dybdal, "Antenna Measurements Using Satellite Beacons," *2005 AMTA Symposium Digest* (October/November 2005).
21. R. B. Dybdal, "Measurement Techniques for Large Terminal Antennas," *2006 AMTA Symposium Digest* (October 2006).
22. R. B. Dybdal, "Measuring 'Not So Big' Antennas," *2008 IEEE AP-S Symposium Digest* (July 2008).
23. R. B. Dybdal, "*G*/*T* Comparative Measurements," *2008 AMTA Symposium Digest* (November 2008).

Index

C

D

E

G

Q