Deep Space Telecommunications Systems Engineering

Joseph H. Yuen, Editor

Jet Propulsion Laboratory
California Institute of Technology

Deep Space Telecommunications Systems Engineering
(JPL Publication 82-76)

Library of Congress Catalog Card Number 82-084114

This book was prepared by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

First Printing July 1982
Second Printing April 1983

Foreword

The challenge of communication in planetary exploration has been unusual. The guidance and control of spacecraft depend on reliable communication. Scientific data returned to earth are irreplaceable, or replaceable only at the cost of another mission. In deep space, communications propagation is good, relative to terrestrial communications, and there is an opportunity to press toward the mathematical limit of microwave communication. Yet the limits must be approached warily, with reliability as well as channel capacity in mind. Further, the effects of small changes in the earth's atmosphere and the interplanetary plasma have small but important effects on propagation time and hence on the measurement of distance.

Advances are almost incredible. Communication capability measured in bits per second at a given range rose by a factor of 10^{18} in the 19 years from Explorer I of 1958 to Voyager of 1977. This improvement was attained through ingenious design based on the sort of penetrating analysis set forth in this book by engineers who took part in a highly detailed and amazingly successful program. Careful observation and analysis have told us much about limitations on the accurate measurement of distance.

It is not easy to get busy people to tell others clearly and in detail how they have solved important problems. Joseph H. Yuen and the other contributors to this book are to be commended for the time and care they have devoted to explicating one vital aspect of a great adventure of mankind.

John R. Pierce
July 1982

Preface

A vital, often predominant function in every space mission is that of communications. From the moment of launch, the only connection between spacecraft and earth is the communications system. This system is responsible for sending scientific data back to earth in the specified quality and quantity together with engineering data reporting the condition of the spacecraft. The communications system also provides the capability of tracking the spacecraft and commanding it to take certain actions. Without an effective communications system a successful mission would not be possible.

To appreciate the challenge that one faces in designing such systems for planetary exploration, one must consider the enormous distances that are involved. Voyager spacecraft, for example, are now more than one billion miles from earth, tens of thousands of times farther than the most distant communications satellite, and continue to transmit data and respond to commands.

The necessity of minimizing spacecraft weight presents a major problem to communications systems designers. The far-reaching implications of spacecraft weight become apparent as the designer considers the problems of providing power supply, antennas, and other necessary devices and supporting elements.

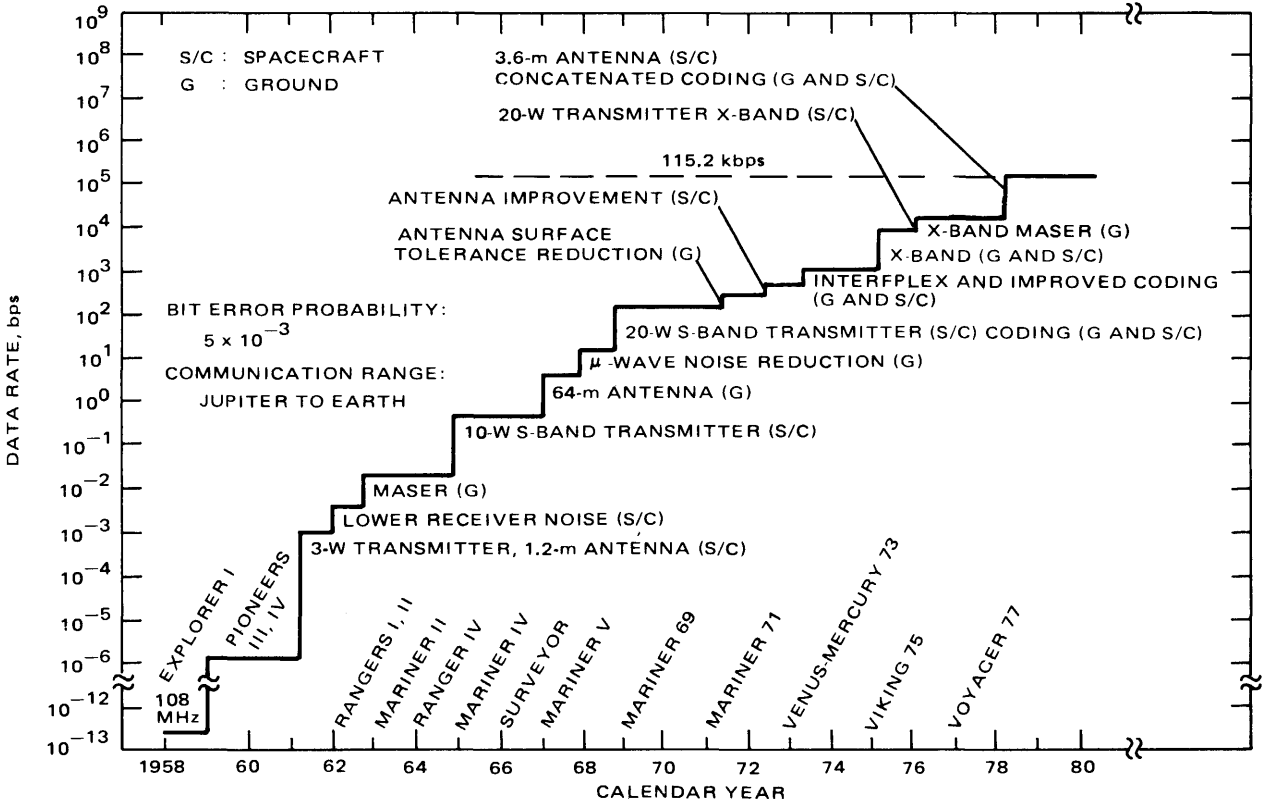
Another important challenge is the extreme reliability required of the communications system on the spacecraft. Once the spacecraft is launched, on-board failures can no longer be repaired except by use of redundant systems. System degradation due to aging, imperfect antenna pointing, or imperfect trajectories can be expected; and the designer must know how much degradation to expect from each case and must design the equipment, the operations, and the procedures of data analysis accordingly. The telecommunications engineer works with the most precise and advanced techniques of the engineering world.

Since the launch in 1958 of Explorer I, the first free-world satellite, there has been substantial progress in improving communications capability. The following figure illustrates the communications capability in terms of data bits per second sent back to earth. Each space mission has led to new telecommunications designs, and the major factors contributing to improvements are indicated in the figure. In order to give a fair comparison, performance has been normalized relative to the distance from earth to Jupiter, and at a bit error probability of 5×10^{-3}. This is, in effect, comparing the past 25 years of spacecraft transmission and receiving system capability with today's Voyager capability. The Voyager capability is 115.2 kbits per second at Jupiter range. Explorer I had a capability of about 10^{-13} bits per second at Jupiter range. Thus communications capability has been improved by a billion billion times, i.e., by 10^{18}. Had Explorer I been used to send pictures from Jupiter, it would have taken millions of years to send just one.

Even though substantial progress has been made in the last 25 years, space exploration is still in its infancy. There has been no exploration beyond the solar system. There are numerous galaxies and billions of stars to investigate. Bigger and tougher challenges are still ahead; more exciting times are yet to come. These challenges will undoubtedly call for more advanced telecommunications systems to transmit information to and from deep space. Telecommunications technology is still in its infancy.

Through the years, a number of telecommunications design techniques, procedures, and analyses contributing to the success of deep space exploration missions have been developed and applied. The purpose of this book is to provide descriptive and analytical information useful for the optimum design, specification, and performance evaluation of deep space telecommunications systems. The book emphasizes system performance information. Long, tedious derivations are not included. The book should serve to acquaint new telecommunications engineers with the techniques available to them and should summarize for the experienced engineers the analyses and information necessary for their work. It also provides a background for understanding the interface between the Deep Space Network and the spacecraft and is intended to facilitate the conceptual designs and analyses for the enhancement of telecommunications performance and assurance of compatibility between spacecraft and ground system capabilities.

The book, which was started more than one and a half years ago, was originally intended to be an update of the widely used *Telecommunications Systems Design Techniques Handbook*, edited by R. E. Edelson, and published on July 15, 1972, by the Jet Propulsion Laboratory. However, owing to advancing communications technology, it became apparent that a total rewrite was necessary.



Profile of deep space telemetry capability

Chapter 1 herein presents an overview of telecommunications systems. For design control, a statistical approach to predict telecommunications link performance and to assess its uncertainty is described. Chapter 2 reviews power spectral density of a class of digital modulations. The required transmission bandwidth is directly related to the power spectral density of the signaling scheme or modulation technique employed. The bandwidth required to transmit the chosen modulation and the ratio of energy-per-bit to noise spectral density are the two important factors in evaluating the efficiency of any communication system. Chapter 3 summarizes the basic fundamentals of phase-locked loop and Costas loop receivers used for synchronization, tracking, and demodulation of phase-coherent signals in residual carrier and suppressed carrier systems. Chapter 4 describes the radio tracking system used in deep space exploration. The radio tracking system provides radio reference carriers which are used for telemetry and command. It also performs radiometric functions in which information is obtained on spacecraft position and the radio propagation medium. Chapter 5 treats the telemetry system, whose function is to transmit information signals from spacecraft to ground receiver. Demodulation and decoding techniques are discussed. Practical system performances of various schemes are presented. Chapter 6 discusses the command system, whose function is to enable commands to be sent to the spacecraft. Chapter 7 describes the spacecraft radio frequency subsystem, which is a vital component for the three telecommunications functions of tracking, telemetry, and command. Chapter 8 examines spacecraft antennas. Chapter 9 outlines a model which combines the statistical description of telecommunication link performance with a statistical description of weather effects. Chapter 10 documents the computer software used for telecommunications performance analysis and monitoring. Software automation is crucial to deep space mission operation. Chapter 11 deals with the selection of radio signal frequencies to avoid interference from other spacecraft and other terrestrial sources. Chapter 12 presents the basic telecommunications interface properties of two new important interfaces: between spacecraft and the space shuttle and between spacecraft and the Tracking and Data Relay Satellite System.

Although this book is intended to be comprehensive, it cannot be. Improvements and innovations in deep space telecommunications are ongoing, but books have a stopping place. Also many important allied topics have been omitted to limit scope and size. Among the prominent omissions are spacecraft design, ground communication network, and mission control and communication facilities. As editor, I have tried to keep each chapter in perspective and to allocate to it the appropriate amount of space. However, individual authors with different specialties and backgrounds have different opinions and styles of presentation.

All the contributors to this book are with the Telecommunications Science and Engineering Division of the Jet Propulsion Laboratory, California Institute of Technology, with the exception of James C. Springett, who is a consultant

from Neocomm Systems, Inc. These individuals are active practitioners in their particular fields of endeavor. I am fortunate to have had their support and contributions. Some contributors have spent considerable amounts of their own time at night and on weekends to participate. Their professional attitudes have my highest respect and deepest appreciation. One individual who deserves special mention is Peter W. Kinman. Without his assistance, this book probably would not have been completed at this time.

Over the years, I have benefited from technical discussions and informal conversations with many friends, too numerous to list, for which I am grateful. I especially want to acknowledge Richard P. Mathison, William J. Weber, Lawrence L. Rauch, Edward C. Posner, James W. Layland, and Robert E. Edelson. I also want to thank Shirley A. Prather, who typed most of the manuscript of this book – some of the chapters several times.

Last, but definitely not least, I want to express my gratitude for the programmatic support and encouragement provided by Hugh S. Fosque of NASA Headquarters and Joel G. Smith and Catherine L. Thornton of the Jet Propulsion Laboratory, California Institute of Technology.

Joseph H. Yuen
Pasadena, California
July 1982

Contents

Chapter 1
Telecommunications Systems Design

Joseph H. Yuen

1.1 Introduction

Deep space telecommunications systems are used for the transmission of information between a spacecraft and a ground station. A typical deep space telecommunications system is illustrated in Fig. 1-1. It performs three basic functions: telemetry, command, and tracking.

The telemetry function involves the transmission of information from a spacecraft to earth. This information usually consists of science data, engineering data, and imaging data. Science data convey information gained from scientific experiments onboard the spacecraft. These data are moderate in volume but very valuable, with stringent quality and transmission accuracy requirements. Engineering data report the status of spacecraft instruments and systems. These data are low in volume and need to be of only moderate quality. Imaging data are of high volume. Because of the redundancy present in planetary scenes, imaging data need be of only moderate quality.

The command function involves the transmission of information from the ground to the spacecraft. Although spacecraft are designed with a high degree of automation and preprogrammed adaptabilities, command is needed to control spacecraft functions, to direct the spacecraft to take specified actions, with

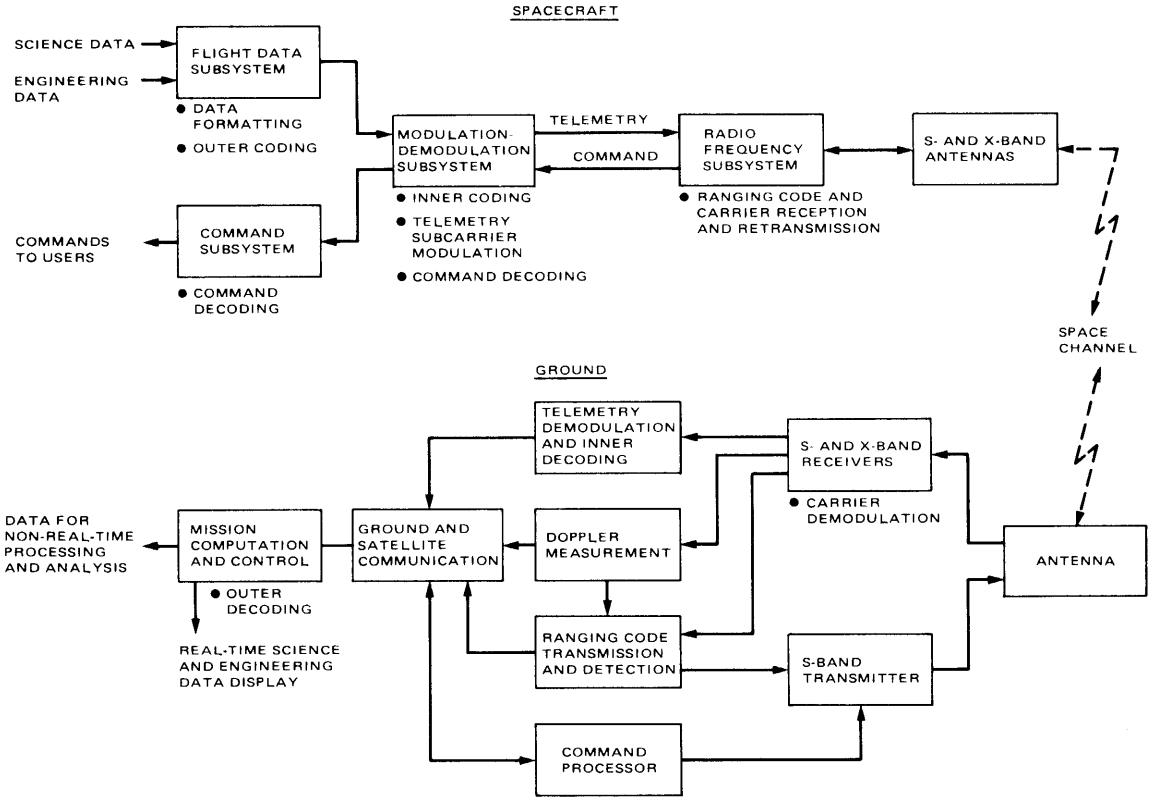


Fig. 1-1. A typical deep space telecommunications system

specified parameters, at a definite time—for example, to change flight path. The command link is characterized by a low data rate and a relatively low volume of data of extremely high quality to assure a minimum of errors of transmission to avoid misinterpretation of commands.

Tracking yields information on spacecraft position and velocity, the radio propagation medium, and the properties of the solar system, thus enabling trajectory monitoring and spacecraft navigation. In their passage through space, the signals transiting between spacecraft and ground station are altered by the medium through which they travel, by gravity, and by the relative velocity of the spacecraft and ground stations. Further, the signals are delayed substantially over these very long paths. By extremely accurate measurements of the radio signal characteristics, the position of the spacecraft, its velocity, and its acceleration can be determined. Simultaneously, important information is gained on the structure, composition, and temperature of the atmosphere of a body that may occult the signal, on the plasma state along the signal path, and on planetary and solar gravitational fields. Thus the process of transmitting radio signals to and from a cooperative spacecraft can allow us to navigate the spacecraft to its destination, measure properties of the gas and plasma in the signal path, and determine gravitational and relativistic effects. These tracking data are characterized by a very low rate, a need for long-term stability, extreme accuracy of measurement, and by the extent of data processing required to turn the signal into information.

The performance of these three functions—telemetry, command, and tracking—depends on the amount of signal that is present relative to the noisy environment in which the signal must be detected, i.e., a ratio of signal-to-noise strength, and the degree of efficiency with which this signal-to-noise ratio is used. Hence, to achieve best possible performance, we must design the telecommunication system which gives the highest signal power, lowest amount of noise, and most efficient use of signal-to-noise ratio, within constraints such as spacecraft weight, size, and cost.

Let us illustrate these points quantitatively by considering the telemetry link. To begin with, it is important to understand that interplanetary space is essentially the same as free space as far as the propagation of microwave signals is concerned. In the absence of weather effects, the earth's atmosphere is almost completely transparent at 2.3 GHz (S-band) and 8.4 GHz (X-band), which are the frequencies being used for deep space telecommunications. Therefore, the calculation of received signal strength is a geometrical problem. Let the spacecraft transmitted power be P_T at distance r from the receiver on earth. If the power were radiated isotropically, the power flux density at the receiver ρ would be the transmitted power divided by the area of a sphere of radius r. The actual flux density is increased by a factor equal to the gain of the transmitting antenna, G_T. Thus, [1-1],

$$\rho = \frac{P_T G_T}{4\pi r^2} \tag{1.1-1}$$

where $4\pi r^2$ is the area of a sphere of radius r. The received signal level P_R is obtained by multiplying the flux density at the receiver by the area of the receiving antenna A_r times the receiving antenna efficiency μ; i.e.,

$$P_R = \rho \mu A_r \tag{1.1-2}$$

The received power P_R can be increased by three actions. First, the transmitted power P_T can be increased. This, of course, has system repercussions. Even though the spacecraft-transmitted power is typically only 20 watts, which is less than a dim desk lamp, the input power required to generate it represents a substantial fraction of the total power available on the spacecraft. Second, the transmitting antenna gain G_t can be increased to focus the transmitted power more intensely on the receiving antenna. Better focusing requires a larger spacecraft antenna (within limits of size and weight constraints) and the ability to aim the resulting narrow beam. Third, the receiving antenna area can be increased to collect as much of the incoming signal power as possible.

Consider the Voyager spacecraft,[1] for the transmitting frequency of 8.4 GHz, where the transmitted power P_T is 21.3 W, the transmitting antenna gain G_T is 6.5×10^4 (unitless), and the distance from Jupiter to Earth is 6.8×10^{11} m. Thus, by using (1.1-1), we find the power flux density at the receiving antenna ρ to be

$$\rho = 2.383 \times 10^{-19} \text{ W/m}^2 \tag{1.1-3}$$

which is a very small number indeed because of the large distance involved. This is a measure of how much the transmitted power spreads in the 6.8×10^{11} meters. We want to recover some of this with as large a receiving antenna as possible. The largest antenna that NASA has is a 64-m-diameter dish (i.e., $A_R = 3.2 \times 10^3$ m^2) which has an efficiency of 42%. The received signal power is obtained by using (1.1-2) with (1.1-3),

$$P_R = 3.05 \times 10^{-16} \text{ W} \tag{1.1-4}$$

[1] Voyager is a NASA deep space mission with two spacecraft launched in 1977. Both reached Jupiter in 1979. Voyager 1 arrived at Saturn in late 1980, and Voyager 2 in the summer of 1981. Voyager 2 will encounter Uranus in 1986 and Neptune in 1989.

This is again a very small number. The ability to communicate depends not on the signal strength alone, but on the signal-to-noise ratio and, further, on how efficiently this signal-to-noise ratio is being used.

The reception process introduces most of the noise that corrupts the receiving signal. Every object radiates energy at radio frequencies. The omnipresent 2.7 kelvin thermal background of the universe produces a radio noise that is an important fraction of the signal power that we are able to supply to the ground receiving antenna. The deep space network (DSN) of NASA employs cryogenically cooled maser amplifiers, so that the total system noise power corrupting the received signal is very low. In typical operation condition, at 8.4 GHz, the DSN station receiving system equivalent noise temperature is 28.5 K. The noise spectral density is, [1-1],

$$N_0 = kT \tag{1.1-5}$$

where k is Boltzmann's constant = 1.38×10^{-20} mW/K Hz, and T is the system equivalent noise temperature = 28.5 K. Thus,

$$N_0 = 3.933 \times 10^{-22} \text{ W/Hz} \tag{1.1-6}$$

There are also other losses in the entire link such as circuit losses and antenna pointing losses, etc., other than that due to distance. For the Voyager telemetry link under consideration, the total loss is $L = 0.7$ (unitless). Therefore, the received signal-power-to-noise-spectral-density ratio, when this loss L is taken into account, is

$$\frac{P_R L}{N_0} = 5.428 \times 10^5 \text{ Hz} \tag{1.1-7}$$

How efficiently this signal-power-to-noise-spectral-density ratio is being used determines the ultimate telemetry capability. And it is the choice of modulation and coding methods that determines the efficiency with which this signal-to-noise ratio is used (see chapter 5). For Voyager imaging data, the acceptable telemetry data quality has a bit error probability of 5×10^{-3} [1-2]. Voyager uses PCM/PSK/PM modulation and convolutional code with Viterbi decoding, which requires a received signal-power-to-noise-spectral-density ratio of 2.05×10^5 Hz for a data rate of 115.2 kilobits per second. Comparing this with (1.1-7), we see that there is a small margin to cover uncertainties. This emphasizes the extreme attention which must be paid to link parameter tolerances throughout the system.

Before turning to chapters 2 through 12, which deal with various aspects of deep space telecommunications, we formalize the link design procedure and its performance uncertainty assessment technique.

1.2 Telecommunications Link Analysis

The performance of a telecommunications system depends on numerous link parameters. Advanced modulation techniques, coding schemes, modern antennas and transmitters, etc., all improve communication efficiency in their own ways. When it comes to designing the entire communications system, communications engineers put all the components or subsystems together and determine system performance capability. In this section the signal-to-noise-spectral-density ratios (SNRs) that characterize telecommunications performance are defined in terms of the link parameters.

1.2.1 Received Power

General equations used for performance computation are derived from the basic equations of communications in the medium between transmitting and receiving systems [1-1]. The first step in link analysis is to calculate the received signal power. Received power P_R is computed by the following equation:

$$P_R = P_T L_T G_T L_{TP} L_S L_A L_P L_{RP} G_R L_R \tag{1.2-1}$$

where P_R is the received signal power at the input to the receiver or preamplifier, P_T is the total transmitted power at antenna terminals, L_T is the transmitting circuit loss between transmitting antenna terminals and radio case due to cabling, G_T is the transmitting antenna gain, L_{TP} is the pointing loss of the transmitting antenna, L_S is the space loss, L_A is the atmospheric attenuation, L_P is the polarization loss between transmitting and receiving antennas due to mismatch in polarization patterns, L_{RP} is the pointing loss of the receiving antenna, G_R is the receiving antenna gain, and L_R is the receiving circuit loss between receiving antenna and receiver due to cabling. Equation (1.2-1) consists of a large number of parameters in product form. Different types of communications links have different components but the form of (1.2-1) remains unchanged.

The space loss, or numerical ratio of received power to transmitted power between two antennas, is given by

$$L_S = \left(\frac{\lambda}{4\pi r}\right)^2 \tag{1.2-2}$$

where λ is the wavelength of radio signal and r is the distance between spacecraft and ground antennas.

The transmitting antenna gain G_T can be related to the effective antenna aperture A_T as

$$G_T = \frac{4\pi A_T}{\lambda^2} \tag{1.2-3}$$

where λ is the wavelength of radio signal. The effective antenna aperture A_T is related to the actual antenna aperture A_t by the relations

$$A_T = \mu A_t \tag{1.2-4}$$

where μ is the antenna efficiency factor. The receiving antenna gain is similarly defined (see chapter 8 for more detailed discussions).

Some of the parameters in (1.2-1) are not defined in exactly the same way on all projects. For example, the transmitting circuit loss L_T is sometimes accounted for by decreasing the effective transmit antenna gain and/or by decreasing the effective transmitted power, obviating L_T. Also, the atmospheric attenuation (for clear, dry weather) is ordinarily accounted for in the ground antenna gain. No matter what the precise definitions are, it must account for the entire telecommunications link.

The received power is referenced to some point in the receiving circuit. Of course, the choice of reference point affects L_R. On the uplink (from ground to spacecraft), the point of reference is usually the input port of the spacecraft transponder. On the downlink (from spacecraft to ground), the point of reference is the input to the maser amplifier. Whatever the reference, the noise equivalent temperature of the receiving system must be referenced to that same point if signal-to-noise ratios are to be computed correctly.

1.2.2 Noise Spectral Density

The noise for an uplink is dominantly thermal. For the downlink, thermal noise is abetted by random radiation picked up by the ground antenna. This random radiation includes that from the atmosphere, hot bodies in the field of view of the antenna, the 2.7 K cosmic background, and that portion of the ground seen by antenna sidelobes.

It is assumed that the receiving system noise has uniform spectral density in the frequency band containing the signal. The one-sided noise spectral density N_0 (in units of watts/hertz) is defined as

$$N_0 = kT \tag{1.2-5}$$

where k is Boltzmann's constant = 1.380×10^{-23} J/K, $\widetilde{k} = 10 \log k = -198.6$ dBm/(Hz K), and T is the system equivalent noise temperature. Equation (1.2-5) is valid since microwave frequency signals are currently being used for deep space telecommunications. For signals in other frequency regions, such as in optical frequencies, different expressions should be used [1-1].

1.2.3 Carrier Performance Margin

Carrier performance margin or simply carrier margin M_c on either uplink or downlink is defined as

$$M_c = \frac{P_c}{2 B_{LO} N_0} \tag{1.2-6}$$

where P_c = portion of received power in the residual carrier, and B_{LO} = one-sided threshold loop noise bandwidth. Here, P_c is calculated from P_R using the modulation indices of the link and depends on the type of modulation used (see chapter 5).

The above definition of carrier margin was chosen because a phase-locked loop receiver loses lock when P_c drops below $2 B_{LO} N_0$ watts (see chapter 3). Thus, $P_c = 2 B_{LO} N_0$ defines carrier threshold. $\widetilde{M}_c$ is calculated as

$$\widetilde{M}_c = \widetilde{P}_c - \widetilde{2 B_{LO}} - \widetilde{N}_0 \tag{1.2-7}$$

and represents the number of decibels the received residual carrier is above carrier threshold. Another popular name for M_c is *carrier SNR in* $2 B_{LO}$. However, this is a misnomer since $B_{LO} N_0$, not $2 B_{LO} N_0$, is the noise power in a thresholding loop. So *carrier SNR in* $2 B_{LO}$ equals one-half the carrier signal-to-noise ratio in a thresholding loop.

The minimum acceptable carrier margin, in general, is not 0 dB. For swept acquisition of the uplink, the minimum useful P_c is in the neighborhood of $20 B_{LO} N_0$ watts. That is, the minimum useful carrier margin is about 10 dB. For the downlink, the DSN recommends that carrier margin be at least 10 dB. Furthermore, carrier margins for two-way doppler may need to be larger than 10 dB, depending on required radiometric accuracies.

1.2.4 Telemetry and Command Performance Margins

For both telemetry and command,

$$S T/N_0 \text{ to receiver} = \frac{S}{R N_0} \tag{1.2-8}$$

where S is the portion of received power in the data modulation sidebands, and R is the data bit rate. Here S is calculated from P_R using the modulation indices of the link. The parameter ST/N_0 *to receiver* is sometimes denoted by E_b/N_0, which is the signal energy per bit to noise spectral density ratio. And

$$ST/N_0 \text{ output} = [ST/N_0 \text{ to receiver}]\, L_{system} \tag{1.2-9}$$

where L_{system} is the system losses. Threshold ST/N_0 is defined by the bit error probability required of a link.

The bottom line of a telemetry or command link analysis is the performance margin. In dB,

$$\text{performance margin} = ST/N_0 \text{ output} - \text{threshold } ST/N_0 \tag{1.2-10}$$

1.2.5 Ranging Performance Margin

The ranging channel involves transmitting a ranging modulation or code from the DSN to the spacecraft, where it is demodulated and then, together with receiver noise, is used to modulate the downlink from the spacecraft to the DSN (see Chapter 4). The ranging signal-to-noise ratio at the spacecraft is

$$\text{ranging input SNR} = \frac{P_{R(u/l)}}{B_R N_{0(u/l)}} \tag{1.2-11}$$

where $P_{R(u/l)}$ is the portion of received uplink power in the ranging modulation sidebands, $N_{0(u/l)}$ is the uplink (i.e., one-sided noise spectral density of the spacecraft receiver), and B_R is the one-sided noise bandwidth of the transponder ranging channel. Here $P_{R(u/l)}$ is calculated from the uplink P_R using the modulation indices of the uplink. The ranging signal-to-noise-spectral-density ratio returned to the DSN is

$$\text{received SNR} = \frac{P_{R(d/l)}}{N_{0(d/l)}} \tag{1.2-12}$$

where $P_{R(d/l)}$ is the portion of received downlink power in the ranging modulation sidebands, and $N_{0(d/l)}$ is the downlink one-sided noise spectral density. Here, $P_{R(d/l)}$ is a function not only of the downlink P_R and the downlink modulation indices but also of *ranging input SNR*. This is because ranging is a turnaround channel. Some of the modulation sidebands on the downlink are turnaround noise sidebands.

$$output\ SNR = [received\ SNR]\ L_{radio} \quad (1.2\text{-}13)$$

where L_{radio} is the radio loss of ranging system. The value of the required SNR is specified by required radiometric accuracies and desired integration time (see chapter 4).

The bottom line of a ranging link analysis is the performance margin, in dB,

$$ranging\ performance\ margin = output\ SNR - required\ SNR \quad (1.2\text{-}14)$$

1.3 Telecommunications Design Control

We have seen in Section 1.1, using the telemetry link of Voyager as an example, that there is only a small margin to cover uncertainties. Indeed, a small number of decibels is usually all that separates an inadequate link design from a costly overdesign. This emphasizes the extreme attention which must be paid to performance prediction for deep space telecommunications systems.

If all link parameters were constant and precisely known to the telecommunications engineer, a simple accounting of the link parameters could predict performance. The real world is not so accommodating, however. Some link parameters vary with spacecraft environment, others with ground station parameters and the communications channel conditions. Some are associated with link components which have manufacturing tolerances.

In the early days of space exploration, engineers had little data and were relatively inexperienced in designing deep space telecommunications systems. Hence they tended to be very conservative. And a deterministic worst-case criterion [1-3] was used to assure sufficient link margins in guarding against uncertainties. Experience over many lunar and planetary flight projects has demonstrated that this approach is practical from the point of view of engineering and management [1-3, 1-4, 1-5]. The major disadvantage of this deterministic worst-case criterion is that it provides no information about the likelihood of achieving a particular design value. Hence cost tradeoff and risk assessment cannot be done quantitatively.

Over the years, more experience was gained in deep space telecommunications systems design. A statistical technique has evolved for treating telecommunications performance statistically [1-6] [1-7], removing the major disadvantages of the deterministic approach while preserving its advantages. Since 1975 this statistical technique has been used in the design of deep space telecommunications systems. It is described in this section.

1.3.1 Design Control Tables

The communication link margin is computed using an equation of the following form:

$$y = y_1 y_2 \cdots y_K \tag{1.3-1}$$

where y_i, $i = 1, 2, \cdots, K$ are parameters of the communication link such as in (1.2-1) and (1.2-6). This equation is presented in its general form, without its detail components. Different types of communications links have different components but the form of this equation remains unchanged.

The overall telecommunication system consists of a large number of parameters in product form. Hence, expressed in the dB domain, it becomes a sum of these parameters; i.e.,

$$x = x_1 + x_2 \cdots + x_K \tag{1.3-2}$$

where

$$x = 10 \log_{10} y \tag{1.3-3}$$

and

$$x_i = 10 \log_{10} y_i \ , \quad i = 1, 2, \cdots, K \tag{1.3-4}$$

In managing the system design, it is most convenient to put this in tabular form with these parameters as entries. This table is referred to as a design control table (DCT). All of the factors that contribute to system performance are listed in the order that one would find in tracing a signal through the system. As examples, sample DCTs of the telemetry, command and ranging links of Galileo[2] are shown in Tables 1-1 through 1-3.

To every parameter in a DCT a design value, along with its favorable and adverse tolerances, is assigned by designers. These tolerances are used not as a hidden safety margin of each parameter; rather, they reflect probable uncertainties, including measurement tolerance, manufacturing tolerance, environment tolerance, drift and aging of elements, parameter modeling errors, etc. The table readily indicates the parameters with the largest tolerances – hence the areas where more knowledge and hardware improvement might be most profitable.

[2]Galileo is a deep space mission, to be launched in 1985, sending a spacecraft to orbit Jupiter, with a probe descending into Jupiter's atmosphere.

Table 1-1. Sample telemetry link design control table of Galileo

	Design	Favorable tolerance	Adverse tolerance	Mean	Variance
Transmitter parameters					
(1) RF power to antenna, dBm				40.5	0.24
Transmitter power, dBm	40.60	1.00	−1.40	40.5	0.24
Transmitter circuit loss, dB	0.00	0.00	0.00	0.00	0.00
(2) Antenna circuit loss, dB	−0.20	0.10	−0.10	−0.2	0.00
(3) Antenna gain, dBi	50.10	0.50	−0.90	50.0	0.08
Cone angle = 0.48 deg					
(4) Pointing loss, dB	−0.81	0.71	−2.00	−1.2	0.33
Path parameters					
(5) Space loss, dB	−290.54			−290.5	0.00
Frequency = 8420.43 MHz					
Range = 9.529+08 km					
= 6.37 AU					
(6) Atmospheric attenuation, dB	0.00	0.00	0.00	0.00	0.00
Weather 50%					
Receiver parameters					
(7) Polarization loss, dB	−0.04	0.04	−0.04		
(8) Antenna gain, dBi	71.75	0.60	−0.60	71.7	0.12
(9) Pointing loss, dB	0.00	0.00	−0.10		
(10) Noise spectral density,					
dBm/Hz	−184.40	−0.53	0.47	−184.4	0.03
Total system noise					
temperature, K	26.30	−3.01	3.00		
T_{op} (zenith), K	20.00	−3.00	3.00		
Delta T_{op} (elev angle), K	6.30	−0.20	0.10		
Delta T_{op} (clouds), K	0.00	0.00	0.00		
Hot body noise, K	0.00	0.00	0.00		
Elev angle = 25.00 deg					
(11) 2-sided threshold loop					
noise bandwidth, dB-Hz	10.00	−0.46	0.41	10.0	0.03
Power summary					
(12) Received power, P_T, dBm				−129.8	0.78
(1+2+3+4+5+6+7+8+9)					
(13) Received P_T/N_0, dB-Hz				54.6	0.81
(12 − 10)					
(14) Carrier supp by ranging					
chan, dB	−0.20	0.10	−0.10	−0.2	0.00
(15) Carrier supp by telemetry					
mod, dB	−15.21	0.52	−0.56	−15.2	0.05
(16) Carrier power/total					
power, dB (14+15)				−15.4	0.05
(17) Received carrier power, dBm				−145.3	0.83
(12 + 16)					
(18) Carrier margin, dB				29.2	0.80
(17 − 10 − 11)				3σ =	2.81

Table 1-1 (contd)

	Design	Favorable tolerance	Adverse tolerance	Mean	Variance
Data channel performance					
(19) Data bit rate, dB	51.28	0.00	0.00	51.3	0.00
Bit rate = 134400.0 BPS					
(20) Data power/total power, dB	−0.33	0.12	−0.12	−0.3	0.00
TLM Mod Index = 80.0 deg					
Supp by rng chan = −0.2 dB					
(21) Data power to receiver, dBm				−130.2	0.78
(12 + 20)					
(22) ST/N_0 to receiver, dB				3.0	0.81
(21 − 19 − 10)					
(23) System losses, dB	−0.43	0.30	−0.30	−0.4	0.02
Radio loss, dB					
Demod, detection loss, dB					
Waveform distortion loss, dB					
(24) ST/N_0 output, dB				2.5	0.83
(22 + 23)					
(25) Threshold ST/N_0, dB	2.31	0.00	0.00	2.3	0.00
(26) Performance margin, dB				0.2	0.83
(24 − 25)					
(27) Array performance					
increase, dB				2.6	0.05
(28) Arrayed ST/N_0, dB				5.1	0.88
(24 + 27)					
(29) Arrayed performance					
margin, dB				2.8	0.88
(26 + 27)				3σ =	2.72

Table 1-2. Sample command link design control table of Galileo

	Design	Favorable tolerance	Adverse tolerance	Mean	Variance
Transmitter parameters					
(1) RF power, dBm	80.00	0.50	−0.50	80.0	0.04
Power output = 100.0 kw					
Transmitter circuit loss, dB	0.00	0.00	0.00	0.00	0.00
(2) Antenna gain, dBi	60.65	0.30	−0.70	60.4	0.08
Elevation angle = 25 deg					
(3) Pointing loss, dB	0.00	0.00	−0.10		
Path parameters					
(4) Space loss, dB	−278.54			−278.5	0.00
Frequency = 2114.68 MHz					
Range = 9.529+08 km					
= 6.37 AU					
(5) Atmospheric attenuation, dB	−0.03	0.00	0.00	−0.0	0.00
Weather 90%					

Table 1-2 (contd)

	Design	Favorable tolerance	Adverse tolerance	Mean	Variance
Receiver parameters					
(6) Polarization loss, dB	−0.26	0.21	−0.28		
(7) Antenna gain, dBi	6.59	2.08	−2.09	6.3	0.73
Cone angle = 0.48 deg					
(8) Pointing loss, dB	0.00	0.00	0.00	0.0	0.00
(9) Receiver circuit loss, dB	−2.00	0.40	−0.50	−2.0	0.07
(10) Noise spectral density, dBm/Hz	−167.68	−0.35	0.86	−167.4	0.04
Operating temp, K	1235.00	−95.00	270.00		
Hot body noise, K	0.00	0.00	0.00		
(11) 2-sided threshold loop, noise bandwidth, dB-Hz	12.00	−0.70	0.60	12.0	0.07
Power summary					
(12) Received power, P_T, dBm (1+2+3+4+5+6+7+8+9)				−133.8	0.93
(13) Received P_T/N_0, dB-Hz (12 − 10)				33.6	0.97
(14) Carrier suppression by ranging mod, dB	0.00	0.00	0.00	0.00	0.00
(15) Carrier suppression by command mod, dB	−4.02	0.27	−0.29	−4.0	0.01
(16) Carrier power/total power, dB (14 + 15)				−4.0	0.01
(17) Received carrier power, dBm (12 + 16)				−137.9	0.94
(18) Carrier margin, dB (17 − 10 − 11)				17.6 3σ =	1.05 3.08
Data channel performance					
(19) Data bit rate, dB	15.05	0.00	0.00	15.1	0.00
Bit rate = 32.0 bps					
(20) Data power/total power, dB	−2.19	0.18	−0.19	−2.2	0.01
Command mod index = 51.0 deg					
Suppression by ranging mod = 0.0 dB					
(21) Data power to receiver, dBm (12 + 20)				−136.0	0.93
(22) ST/N_0 to receiver, dB (21 − 19 − 10)				16.3	0.97
(23) System losses, dB	−2.40	0.50	−0.50	−2.4	0.4
Radio loss, dB					
Demod, detection loss, dB					
Waveform distortion, dB					
(24) ST/N_0 output, dB (22 + 23)				13.9	1.02
(25) Threshold ST/N_0, dB	9.59	0.00	0.00	9.6	0.00
(26) Performance margin, dB (24 − 25)				4.4 3σ =	1.02 3.02

Table 1-3. Sample ranging channel design control table of Galileo

	Design	Favorable tolerance	Adverse tolerance	Mean	Variance
Transmitter parameters					
(1) RF power, dBm	73.00	0.50	−0.50	73.0	0.04
Power output = 20.0 kw					
Transmitter circuit loss, dB	0.00	0.00	0.00	0.0	0.00
(2) Antenna gain, dBi	55.30	0.30	−0.70	55.1	0.08
Elevation angle = 25.00 deg					
(3) Pointing loss, dB	0.00	0.00	−0.10		
Path parameters					
(4) Space loss, dB	−278.54			−278.5	0.00
Frequency = 2114.68 MHz					
Range = 9.529+08 km					
= 6.37 AU					
(5) Atmospheric attenuation, dB	−0.03	0.00	0.00	−0.0	0.00
Weather 90%					
Receiver parameters					
(6) Polarization loss, dB	−4.25	1.22	−1.55		
(7) Antenna gain, dBi	36.40	0.60	−0.80	32.0	0.40
Cone angle = 0.48 deg					
(8) Pointing loss, dB	−0.18	0.18	−0.07	−0.1	0.00
(9) Receiver circuit loss, dB	−2.00	0.40	−0.50	−2.0	0.07
(10) Noise spectral density,					
dBm/Hz	−167.81	−0.22	0.78	−167.5	0.03
Operating temperature, K	1200.00	−60.00	235.00		
Hot body noise, K	0.00	0.00	0.00		
(11) 2-sided threshold loop					
Noise bandwidth, dB-Hz	12.20	−0.70	0.60	12.2	0.07
Power Summary					
(12) Received power, P_T, dBm				−120.6	0.60
(1+2+3+4+5+6+7+8+9)					
(13) Received P_T/N_0, dB-Hz				46.9	0.63
(12 − 10)					
(14) Carrier suppression by					
ranging mod, dB	−3.01	0.22	−0.23	−3.0	0.01
(15) Carrier suppression by					
command mod, dB	−0.54	0.04	−0.05	−0.5	0.00
(16) Carrier power/total					
power, dB (14+15)				−3.6	0.01
(17) Received carrier power, dBm				−124.2	0.61
(12 + 16)					
(18) Carrier margin, dB				31.2	0.71
(17−10−11)				2σ =	1.61

Table 1-3 (contd)

	Design	Favorable tolerance	Adverse tolerance	Mean	Variance
Transmitter parameters					
(1) RF power to antenna, dBm				40.5	0.24
Transmitter power, dBm	40.60	1.00	−1.40	40.5	0.24
Transmitter circuit loss, dB	0.00	0.00	0.00	0.00	0.00
(2) Antenna circuit loss, dB	−0.20	0.10	−0.10	−0.2	0.00
(3) Antenna gain, dBi	50.10	0.50	−0.90	50.0	0.08
Cone angle = 0.48 deg					
(4) Pointing loss, dB	−0.81	0.71	−2.00	−1.2	0.33
Path parameters					
(5) Space loss, dB	−290.54			−290.5	0.00
Frequency = 8420.43 MHz					
Range = 9.529+08 km					
= 6.37 AU					
(6) Atmospheric attenuation, dB	−0.35	0.00	0.00	−0.03	0.00
Weather 50%					
Receiver parameters					
(7) Polarization loss, dB	−0.02	0.02	−0.04		
(8) Antenna gain, dBi	66.20	0.60	−0.60	66.2	0.12
(9) Pointing loss, dB	0.00	0.00	−0.10		
(10) Noise spectral density,					
dBm/Hz	−181.39	−0.33	0.31	−181.4	0.01
Total system noise					
temperature, K	52.62	−3.91	3.91		
T_{op} (zenith), K	25.00	−2.50	2.50		
Delta T_{op} (elev angle), K	6.80	−3.00	3.00		
Delta T_{op} (clouds), K	20.82	0.00	0.00		
Hot body noise, K	0.00	0.00	0.00		
Elev angle = 25.00 deg					
(11) 2-sided threshold loop					
Noise bandwidth, dB-Hz	10.33	−0.46	0.46	10.3	0.04
Power summary					
(12) Received power, P_T, dBm				−135.7	0.78
(1+2+3+4+5+6+7+8+9)					
(13) Received P_T/N_0, dB-Hz				45.7	0.79
(12 − 10)					
(14) Carrier supp by ranging					
chan, dB	−0.22	0.05	−0.05	−0.2	0.00
(15) Carrier supp by telemetry					
mod, dB	−15.21	0.52	−0.56	−15.2	0.05
(16) Carrier power/total					
power, dB (14+15)				−15.4	0.05
(17) Receiver carrier power, dBm				−151.2	0.83
(12 + 16)					
(18) Carrier margin, dB				19.9	0.88
(17 − 10 − 11)				2σ =	1.87

Table 1-3 (contd)

	Design	Favorable tolerance	Adverse tolerance	Mean	Variance
Ranging channel performance					
(19) Uplink ranging power/ total power, dB Ranging mod index = 45.0 deg Supp by cmd mod = -0.5 dB	-3.55	0.26	-0.28	-3.6	0.01
(20) Ranging power to receiver, dBm (12U+19)				-124.2	0.61
(21) Uplink radio loss, dB	0.00	0.00	0.00	0.0	0.00
(22) Ranging power at receiver output, dBm (20+21)				-124.2	0.61
(23) Ranging noise bW, dB-Hz	61.76	-0.22	0.22	61.8	0.01
(24) Ranging noise power, dBm (10U+23)				-105.8	0.04
(25) Ranging input SNR, dB (22-24)				-18.4	0.64
(26) Downlink ranging power/ Total power, dB Downlink ranging MI = 12.9 deg rms Supp by telemetry mod = -14.2 dB	-47.77	3.76	-3.92	-47.8	1.86
(27) Received ranging power, dBm (12D+26)				-183.5	2.64
(28) Received P_R/N_0, dB (27-10D)				-2.1	2.66
(29) Downlink radio loss, dB	-0.50	0.00	0.00	-0.5	0.00
(30) P_R/N_0 output, dB (28+29)				-2.6	2.66
(31) Required P_R/N_0, dB	0.00	0.00	0.00	0.0	0.00
(32) Performance margin, dB (30-31)				-2.6 2σ =	2.66 3.26

The performance assessment and the design procedure are described in the following section.

1.3.2 Design Procedure and Criterion

Before we describe the statistical approach used for deep space telecommunications systems design, we examine the importance of weather effects on telecommunications performance.

1.3.2.1 Weather Effects. Weather requires special consideration. For carrier frequencies at or above X-band, the randomness that weather introduces to the link dominates all other sources of randomness. There are two techniques for

incorporating weather into telecommunications design control. The simple one of these, the percentile weather technique, is described in this section. It is a reasonable estimate of the weather effects on link performance. Often a reasonable estimate suffices for preliminary system design and performance assessment purposes. The percentile weather technique is attractive for its simplicity. On the other hand, for detailed design and link performance monitoring purposes, a more accurate estimate is required. A more elaborate technique is treated in Chapter 9 in this book.

The percentile technique for incorporating weather into telecommunications design control requires the preparation of two design control tables. In the first design control table, a dry atmosphere and clear sky over the DSS is assumed. In the second design control table, x-percentile inclement weather is assumed. By "x-percentile" weather is meant that with $x\%$ probability a pessimistic assumption is being made about weather effects; moreover, with $(100 - x)\%$ probability an optimistic assumption is being made. As an example, 95-percentile means that 95% of the time the degradation due to weather is less than predicted, while 5% of the time the weather degradation is worse.

1.3.2.2 Design Procedure. The design procedure is described here. The procedure unfolds as a sequence of six steps during which the philosophy of telecommunications design control reveals itself. The discussion below follows [1-6] and [1-7].

Step 1. Three values are assigned to most link parameters: design, favorable tolerance, and adverse tolerance. All three values are to be in decibel representation. Those parameters which are not assigned three values should receive only a design value (in decibels). Data bit rate, space loss, and threshold (or required) signal-to-noise ratios are regarded as deterministic and only have design values. The weather-dependent parameters – atmospheric attenuation and, on the downlink, incremental noise temperature due to clouds – should be assigned only design values (noise temperature is in units of kelvin, though, not decibels). In fact, the design values of the weather-dependent parameters should be based on the assumption of clear, dry weather. Later on, as explained in the previous paragraph, the design procedure is to be repeated with weather-dependent design values assigned on the basis of x-percentile inclement weather. The following definitions serve as a guide in the assignment of values to a link parameter:

Design value	=	the a priori estimate of a parameter,
Favorable tolerance	=	the best case of a parameter minus the design value,
Adverse tolerance	=	the worst case of a parameter (short of failure) minus the design value.

Noise temperatures, noise spectral densities, and noise bandwidths have favorable tolerances with negative values and adverse tolerances with positive values. The opposite is true of all other link parameters that get assigned tolerances. Tolerances reflect one or more of the following: a limit cycle, a manufacturing tolerance associated with a link hardware component, a dependence on spacecraft environment, and other uncertainties.

Step 2. Arrange the link parameters in a vertical listing – a design control table – and identify independent groups among them.

Step 3. Within each of the independent groups, add the design values and the favorable and adverse tolerances so that there is only one design value with its associated favorable and adverse tolerances for each group.

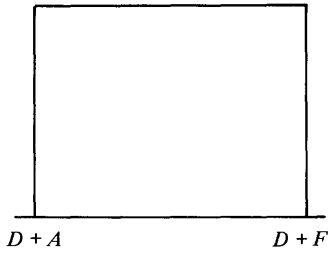
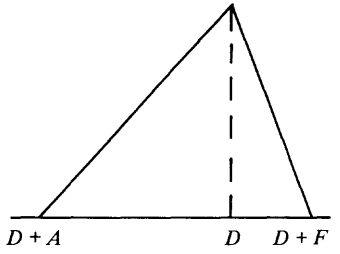
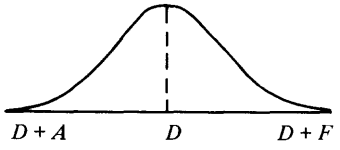
Step 4. Assign a probability density function (pdf) to each independent group. Typically, only uniform, triangular, Gaussian, and Dirac-delta (for those groups without tolerances) pdf's are used. The assignment made by the Telecommunications Prediction and Analysis Program (TPAP) are tabulated in Chapter 10. In case a probability density function is nonzero over the entire real line such as the Guassian density function, use the absolute sum of its favorable and adverse tolerances as its 6-sigma (6 standard deviations) measure.

Step 5. Compute for each independent group (random variable) its mean and variance. Table 1-4 illustrates the computation of mean and variance for the common pdf's. Having been computed from a design value and tolerances all expressed in decibels, the mean will, of course, be in decibels (and the variance in decibels squared).

Step 6. Compute the mean and variance of the desired performance or carrier margin by algebraically summing the means and adding the variances obtained in step 5. By algebraically summing is meant that some means – those corresponding to noise spectral density, noise bandwidth, data bit rate, and threshold (or required) signal-to-noise ratios – are subtracted rather than added.

It is certainly true that a precise probability density function of the overall link margin can be obtained by convolving the pdf's of the K independent random variables. However, the link margin tolerance distribution is well approximated by a Gaussian distribution by invoking the central limit theorem since the overall link consists of K independent random variables formed in step 2 above. This simplifies the computational complexity to the point that hand calculation is indeed practical. Moreover, the pdf's of the K independent random variables were only estimated. It seems difficult to justify using tedious convolution to achieve a precise solution based on imprecise information if an approximation is indeed satisfactory. A more worthwhile effort would be making a more accurate estimate of the pdf's of the K independent random variables.

Table 1-4. Computation of mean and variance of common probability density functions

	Probability density function
Uniform $M = D + (F + A)/2$ $V = (F - A)^2/12$	
Triangular $M = D + (F + A)/3$ $V = (F^2 + A^2 - AF)/18$	
Gaussian $M = D + (F + A)/2$ $V = (F - A)^2/36$	

D = design
M = mean
V = variance
A = adverse tolerance, which has a negative value for all link parameters except noise spectral density and noise bandwidth
F = favorable tolerance, which has a positive value for all link parameters except noise spectral density and noise bandwidth

For noise spectral density and noise bandwidth the favorable tolerance is a negative number while the adverse tolerance is a positive number.

The above procedure is repeated with the weather-dependent parameters being assigned design values based on x-percentile inclement weather. The performance or carrier margin is finally considered predicted with the specification of four members:

(1) Mean margin with clear, dry weather.

(2) Mean margin with x-percentile weather.

(3) n-sigma margin with clear, dry weather.

(4) n-sigma margin with x-percentile weather.

where "n-sigma margin" equals mean margin minus n standard deviations. The value n is typically 3 for command links and 2 otherwise.

1.3.2.3 Design Criterion. In order to assure successful operation and guard against adverse situations, we must provide sufficient link margins. Based on the design procedure described in the preceding section, it can be stated that, for example, link performance will not deviate from its mean margin by more than 3-sigma (three standard deviations) with probability 0.99. This 3-sigma value is used as an uncertainty measure for the link. Depending on how much risk we are willing to take, we can choose any number of sigma values. Hence a useful design criterion is: *the mean value of the link SNR must exceed the required SNR by an amount equal to or larger than n-sigma.* The value n is typically 3 for command links and 2 for telemetry and ranging links.

References

1-1. Pierce, J. R., and Posner, E. C., *Introduction to Communication Science and Systems*, Plenum Press, New York, 1980.

1-2 Edelson, R. E., Madsen, B. D., Davis, E. K., and Garrison, G. W., "Voyager Telecommunications: The Broadcast from Jupiter," *Science*, Vol. 204, No. 4396, pp. 913-921, June 1, 1979.

1-3 Edelson, R. E. (editor), *Telecommunications Systems Design Techniques Handbook*, Technical Memorandum 33-571, Jet Propulsion Laboratory, Pasadena, Calif., July 15, 1972.

1-4. Easterling, M. F., "From 8-1/3 Bits per Second to 100,000 Bits per Second in Ten Years," *Conference Proceedings, ITC '74*, Los Angeles, Calif., 1974.

1-5. Evanchuk, V. L., "117.6 kbps Telemetry from Mercury In-Flight System Analysis," *Conference Proceedings, ITC '74*, Los Angeles, Calif., 1974.

1-6. Yuen, J. H., *A Practical Statistical Model for Telecommunications Performance Uncertainty*, Technical Memorandum 33-732, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1975.

1-7. Yuen, J. H., "A Statistical Model for Telecommunication Link Design," *Proceedings, National Telecommunications Conference*, New Orleans, La., Dec. 1976.

Chapter 2
Spectral Characteristics of Digital Modulations

Marvin K. Simon

2.1 Introduction

Two factors that are important in evaluating the efficiency of any communication system are the bandwidth required to transmit the chosen signaling or modulation technique and the energy-per-bit-to-noise-spectral-density ratio. Whereas the latter factor is a well-defined quantity, the former can be defined in many ways [2-1]. Regardless of the definition, however, the required transmission bandwidth is directly related to the *power spectral density* (PSD) of the signaling scheme. Thus, efficient analytical methods for calculating this power spectral density are essential if one is to specify the spectral occupancy of the transmission channel.

One very important class of digital modulations is that corresponding to the simultaneous transmission of one or more *synchronous data pulse streams* where the underlying data sequences that generate these pulse streams have known statistical properties. Techniques for computing the PSD of such pulse streams have appeared in the literature and in some textbooks. In almost all cases, each sequence is assumed to be either wide sense stationary (WSS), with known

autocorrelation function or, more generally, an N-ary Markov source with known (fixed) stationary and transition probabilities. A special case of the latter is a purely random source, i.e., one that emits an elementary signal (pulse) in a given signaling interval independent of those emitted in previous signaling intervals. Such a source can be modeled as a degenerate case of an N-ary Markov source whose $N \times N$ transition matrix (the matrix of transition probabilities) has identical rows. When $N = 2$, and the two elementary signals are the same, we have the familiar case of antipodal binary signaling.

Occasionally, one finds situations where the generating data sequence is not WSS but, however, is *cyclostationary*, i.e., its first two moments are periodic. One such case occurs at the output of a convolutional encoder where the period (in code symbols) is equal to the reciprocal of the code rate. Thus, in order to characterize the spectral properties of the output of such an encoder, one must have an expression for the power spectral density of a synchronous data stream whose generating sequence is cyclostationary.

This section is devoted to reviewing the techniques described above along with illustrative examples of how they are used to compute the PSD of well-known modulation/coding schemes. We begin with a discussion of the PSD of a single synchronous data stream generated by a binary WSS sequence.

2.2 Power Spectral Density of a Synchronous Data Stream Generated by a Binary, Zero Mean, WSS Sequence

Consider the binary (±1) zero mean, WSS sequence $\{a_n\}$ for which it is known that

$$\left.\begin{aligned} \overline{a_n} &= 0 \\ \overline{a_n a_m} &= R(m-n) \end{aligned}\right\} \tag{2.2-1}$$

and the overbar denotes statistical expectation. From this sequence, we form the synchronous data pulse stream

$$m(t) = \sum_{n=-\infty}^{\infty} a_n p(t - nT) \tag{2.2-2}$$

where $p(t)$ is the elementary signal (pulse shape) and is not necessarily restricted to be time-limited to only a single signaling interval, i.e., the pulse train $m(t)$ can contain overlapping pulses. Irrespective of the properties of the generating sequence $\{a_n\}$, the data stream $m(t)$ is itself cyclostationary since the expected

value of the product $m(t)m(t+\tau)$ is, in addition to being a function of τ, a periodic function of t. Thus, to compute the power spectral density $S_m(f)$ of $m(t)$, we must first average

$$R(t;\tau) \stackrel{\Delta}{=} \overline{m(t)m(t+\tau)} \tag{2.2-3}$$

over t [the averaging is performed over the period of $R(t;\tau)$] and then take the Fourier transform of the result. Thus,

$$S_m(f) \stackrel{\Delta}{=} \mathcal{F}\{<R(t;\tau)>\} \tag{2.2-4}$$

where $<>$ denotes time average and $\mathcal{F}$ denotes Fourier transform. Substituting (2.2-2) into (2.2-3) and making use of (2.2-1), it can be shown [2-2] that

$$S_m(f) = S_p(f)S_a(f) \tag{2.2-5}$$

where, with $P(f)$ the Fourier transform of $p(t)$,

$$S_p(f) \stackrel{\Delta}{=} \frac{1}{T}|P(f)|^2 \tag{2.2-6}$$

is the power spectral density of the individual pulse $p(t)$ and

$$S_a(f) \stackrel{\Delta}{=} \sum_{l=-\infty}^{\infty} R(l)e^{-j2\pi flT} \tag{2.2-7}$$

is the spectral density of the sequence, i.e., the discrete Fourier transform of its correlation function. Note that if the data sequence is purely random, i.e.,

$$\overline{a_n a_m} = \begin{cases} 1\,; & m = n \\ 0\,; & \text{otherwise} \end{cases} \tag{2.2-8}$$

then equivalently from

$$R(l) = \begin{cases} 1\,; & = 0 \\ 0\,; & \text{otherwise} \end{cases} \tag{2.2-9}$$

and from (2.2-7)

$$S_a(f) = 1 \tag{2.2-10}$$

Hence,

$$S(f) = S_p(f) \tag{2.2-11}$$

which is a commonly used result.

2.3 Power Spectral Density of a Synchronous Data Stream Generated by a Binary, Zero Mean, Cyclostationary Sequence

Suppose now that we have a binary sequence $\{a_n\}$ which has the properties

$$\begin{aligned} \overline{a_n} &= 0 \\ \overline{a_n a_m} &= R(n; m-n) \end{aligned} \tag{2.3-1}$$

Furthermore,

$$R(n; m-n) = R(n+kN; m-n); \quad k = 0, \pm 1, \pm 2, \ldots \tag{2.3-2}$$

where N denotes the period of the correlation function $R(n; m-n)$. Then, it can be shown [2-2] that the analogous result to (2.2-5) is now

$$S_m(f) = S_p(f)\, S_{\bar{a}}(f) \tag{2.3-3}$$

where $S_p(f)$ is still given by (2.2-6) and $S_{\bar{a}}(f)$ is defined by

$$S_{\bar{a}}(f) = \sum_{l=-\infty}^{\infty} \left[\frac{1}{N} \sum_{n=1}^{N} R(n; l) \right] e^{-j2\pi flT} \tag{2.3-4}$$

Comparing (2.3-4) with (2.2-7) we notice, not surprisingly, that the only difference between the two is that the periodicity of the correlation function caused by the cyclostationary behavior of the sequence $\{a_n\}$ must be "averaged out" before taking the discrete Fourier transform.

2.4 Power Spectral Density of a Synchronous Data Stream Generated by a Binary, Nonzero Mean, Cyclostationary Sequence

When the generating sequence $\{a_n\}$ is not zero mean, then the spectrum of the corresponding synchronous data stream will have a discrete component in addition to the customary continuous component. An example of a situation where this might occur is at the output of a convolutional encoder whose input is random but not equiprobable binary data. The procedure for handling this case is as follows:

Define the zero mean cyclostationary sequence $\{A_n\}$ by

$$A_n = a_n - \bar{a}_n \tag{2.4-1}$$

which has the properties

$$\bar{A}_n = 0 \tag{2.4-2}$$

$$\overline{A_n A_m} = R_A(n; m-n) \tag{2.4-3}$$

Then, using the results of the previous section, the *continuous* component of the power spectrum $S_c(f)$ for a synchronous data stream generated by $\{a_n\}$ is given by

$$S_c(f) = S_p(f)\, S_{\bar{A}}(f) \tag{2.4-4}$$

where again $S_p(f)$ is defined in (2.2-6) and

$$S_{\bar{A}}(f) = \sum_{l=-\infty}^{\infty} \left[\frac{1}{N} \sum_{n=1}^{N} R_A(n; l) \right] e^{-j2\pi flT} \tag{2.4-5}$$

The discrete spectral component $S_d(f)$ is found from

$$S_d(f) = \mathcal{F} \left\{ \left\langle \sum_n \sum_m \overline{a_n a_m}\, p(t-nT)\, p(t+\tau-mT) \right\rangle \right\} \tag{2.4-6}$$

or, after much manipulation [2-2],

$$S_d(f) = \frac{1}{(NT)^2} \sum_k \left| P\left(\frac{k}{NT}\right) \right|^2 \left(\sum_{m=1}^{N} \overline{a_m}\, e^{-j2\pi mk/N} \right)$$

$$\times \left(\sum_{n=1}^{N} \overline{a_n}\, e^{j2\pi nk/N} \right) \delta\left(f - \frac{k}{NT}\right) \tag{2.4-7}$$

Note that when $\{a_n\}$ is a WSS sequence (i.e., $N = 1$), (2.4-7) reduces to the familiar result

$$S_d(f) = \frac{(\overline{a_1})^2}{T^2} \sum_k \left| P\left(\frac{k}{T}\right) \right|^2 \delta\left(f - \frac{k}{T}\right) \tag{2.4-8}$$

2.5 Examples and Applications

As a simple example of a WSS sequence consider a binary, zero mean, Markov source characterized by

$$\begin{aligned} Pr\{a_{n+1} \neq a_n\} &= p_t \\ Pr\{a_{n+1} = a_n\} &= 1 - p_t \end{aligned} \tag{2.5-1}$$

The correlation function for such a source is easily shown to be

$$R(l) = (1 - 2p_t)^{|l|} \tag{2.5-2}$$

Substituting (2.5-2) into (2.2-7) gives

$$S_a(f) = 1 + 2 \sum_{l=1}^{\infty} (1 - 2p_t)^l \cos 2\pi flT \tag{2.5-3}$$

Using the identity [2-3; p. 84, Eq. (454)]

$$\sum_{k=1}^{\infty} a^k \cos k\theta = \frac{a \cos \theta - a^2}{1 - 2a \cos \theta + a^2} \tag{2.5-4}$$

with $a = 1 - 2p_t$ and $\theta = 2\pi fT$, immediately gives the desired result

$$\left.\begin{aligned} S_a(f) &= \frac{1 - (1 - 2p_t)^2}{1 + (1 - 2p_t)^2 - 2(1 - 2p_t)\cos 2\pi fT} \\ &= \frac{4p_t(1 - p_t)}{2(1 - 2p_t)(1 - \cos 2\pi fT) + 4p_t^2} \end{aligned}\right\} \tag{2.5-5}$$

Furthermore, if the data stream generated by this sequence uses rectangular pulses, i.e.,

$$p(t) = \begin{cases} 1; & 0 \leqslant t \leqslant T \\ 0; & \text{otherwise} \end{cases} \tag{2.5-6}$$

then, using (2.2-6) and (2.5-5),

$$S(f) = T \frac{\sin^2 \pi fT}{(\pi fT)^2} \left[\frac{4p_t(1 - p_t)}{2(1 - 2p_t)(1 - \cos 2\pi fT) + 4p_t^2} \right] \tag{2.5-7}$$

As a second example, consider the sequence formed by interleaving N independent, first-order, Markov sources with respective transition probabilities p_{t_n}; $n = 1, 2, \ldots$. Then, the resulting sequence is cyclostationary with correlation function

$$R(n; l) = \begin{cases} (1 - 2p_{t_n})^{\frac{|l|}{N}} \\ \qquad ; \quad l = 0, \pm N, \pm 2N, \ldots . \\ 0; \quad \text{all other integer } l \end{cases} \tag{2.5-8}$$

The power spectral density $S_{\bar{a}}(f)$ of (2.3-4) is computed as (letting $l = kN$, $k = 0, \pm 1, \pm 2, \ldots$.)

$$\left.\begin{aligned} S_{\bar{a}}(f) &= \frac{1}{N}\sum_{n=1}^{N}\sum_{k}(1-2p_{t_n})^{|k|}\,e^{-j2\pi fkNT} \\ &= \frac{1}{N}\sum_{n=1}^{N}\left\{1+2\sum_{k=1}^{\infty}(1-2p_t)^k\cos 2\pi fkNT\right. \end{aligned}\right\} \tag{2.5-9}$$

Noticing the similarity between (2.5-3) and (2.5-9) (for fixed n), we can immediately write down the result

$$S_{\bar{a}}(f) = \frac{1}{N}\sum_{n=1}^{N}\left[\frac{4p_{t_n}(1-p_{t_n})}{2(1-2p_{t_n})(1-\cos 2\pi fNT)+4p_{t_n}^2}\right] \tag{2.5-10}$$

As a third example, consider a convolutional encoder with constraint length K and rate b/n. In the mth information interval, b information symbols a_{mb+j}; $j = 0, 1, 2, \ldots, b-1$ enter the encoder and n channel symbols X_{mn+p}; $p = 1, 2, \ldots, n$ exit the encoder. The structure of such a convolutional encoder can also be defined by a generator (connection) matrix, namely,

$$G = \begin{bmatrix} g_{1,1}g_{1,2}\cdots\cdots g_{1,Kb} \\ g_{2,1}g_{2,2}\cdots\cdots g_{2,Kb} \\ \vdots \\ g_{n,1}g_{n,2}\cdots\cdots g_{n,Kb} \end{bmatrix} \tag{2.5-11}$$

where $g_{i,j}$ is either one or zero depending, respectively, on whether the ith modulo summer is connected to the jth shift register stage. For mathematical convenience, we shall assume that both the input symbols $\{a_{mb+j}\}$ and the output symbols $\{X_{mn+p}\}$ take on values plus and minus one. This allows modulo 2 summation operations to be replaced by algebraic products. Thus, the encoder has the input/output relation

$$X_{mn+p} = \prod_{i=1}^{Kb}\left[a_{(m+1)b-i}\right]^{g_{p,i}} \tag{2.5-12}$$

and the output sequence $\{X_{mn+p}\}$ is clearly cyclostationary with period n.

If we now try to apply the results of Sections 2.3 and 2.4 to the encoder output sequence $\{X_i\}$ then, after considerable manipulation, it can be shown [2-2] that $S_A(f)$ [now appropriately denoted by $S_X(f)$] becomes

$$S_X(f) = \frac{1}{n} \sum_{p=1}^{n} \sum_{q=1}^{n} \sum_{l=0}^{\infty} \epsilon_l \operatorname{cov}(X_p, X_{nl+q}) \times \cos[2\pi(nl + q - p) fT] \tag{2.5-13}$$

where $\operatorname{cov}(X_i X_j)$ is the covariance function of the sequence $\{X_i\}$ defined by

$$\operatorname{cov}(X_i, X_j) = \overline{(X_i - \overline{X}_i)(X_j - \overline{X}_j)} \tag{2.5-14}$$

and ϵ_l is the Neumann factor defined by

$$\epsilon_l = \begin{cases} 1; & l = 0 \\ 2; & \text{otherwise} \end{cases} \tag{2.5-15}$$

Also, $S_d(f)$ of (2.4-7) becomes

$$S_d(f) = \frac{1}{(nT)^2} \sum_{k=-\infty}^{\infty} \left| P\left(\frac{k}{nT}\right) \right|^2 \left\{ \left[\sum_{m=1}^{n} \overline{X}_m \cos\left(\frac{2\pi mk}{n}\right) \right]^2 + \left[\sum_{m=1}^{n} \overline{X}_m \sin\left(\frac{2\pi mk}{n}\right) \right]^2 \right\} \delta\left(f - \frac{k}{nT}\right) \tag{2.5-16}$$

Here T denotes the encoder output symbol time.

Suppose now we assume the input to the encoder to be a sequence of independent binary symbols $\{a_j\}$ that take on values ± 1 with probabilities

$$Pr\{a_j = -1\} = 1 - Pr\{a_j = 1\} = p\star \tag{2.5-17}$$

and average symbol value

$$E\{a_j\} = 1 - 2p\star \triangleq \bar{a} \tag{2.5-18}$$

Consider first the special case of a random non-return-to-zero (NRZ) encoder input for which $p\star = 1/2$ and correspondingly, from (2.5-18), $\bar{a} = 0$. Then, if the rows of the generator matrix G of (2.5-11) are such that the constraints (which are illustrated in Fig. 2-1)

$$\left.\begin{aligned} g_{p,i} &= 0; \quad i = (K-l)\,b+1, (K-l)\,b+2, \ldots, Kb \\ g_{q,i} &= 0; \quad i = 1, 2, \ldots, lb \\ g_{p,i} &= g_{q,i+lb}; \quad i = 1, 2, \ldots, (K-l)\,b \end{aligned}\right\} \tag{2.5-19}$$

are *not* satisfied for any combination of p, q, and l (other than $p = q$ and $l = 0$), then (2.5-13) simplifies to [2-2]

$$S_X(f) = 1 \tag{2.5-20}$$

or

$$S_c(f) = T\left(\frac{\sin \pi fT}{\pi fT}\right)^2 \tag{2.5-21}$$

Furthermore, the discrete spectrum $S_d(f)$ of (2.5-16) vanishes. Thus, for all codes which do not satisfy (2.5-19), a random NRZ data input results in a

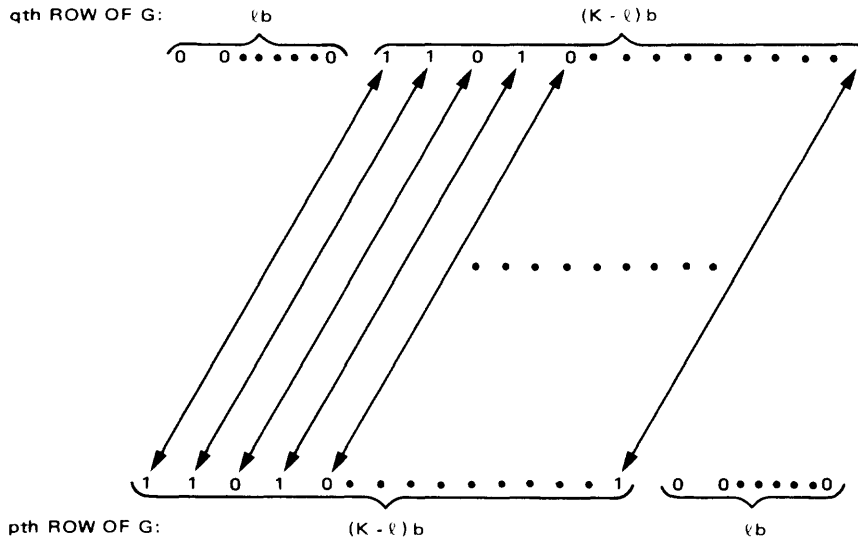


Fig. 2-1. An illustration of the code constraints of (2.5-19)

random NRZ data output scaled (expanded) in frequency by the reciprocal of the code rate, i.e., n/b. This class of codes was defined in [2-2] as "uncorrelated convolutional codes." Note, all rate 1/2, noncatastrophic, convolutional codes are uncorrelated codes.

On the other hand, if the rows of G are such that (2.5-19) can be satisfied for at least one combination of p, q, and l (other than $p = q$ and $l = 0$), then letting $\{p, q, l\}$ denote the set of combinations of p, q, l (other than $p = q$ and $l = 0$) for which (2.5-19) can be satisfied, (2.5-13) simplifies to

$$S_x(f) = 1 + \frac{1}{n} \sum\sum\sum_{\{p, q, l\}} \epsilon_l \cos [2\pi (nl + q - p)fT] \tag{2.5-22}$$

Also, the discrete spectrum $S_d(f)$ of (2.5-16) still vanishes. Since for such codes the encoder output spectrum differs (in form) from the input spectrum, this class of codes was defined in [2-2] as "correlated convolutional codes."

As an example of a "correlated convolutional code," we observe that any code whose generator matrix is such that two or more rows are identical satisfies (2.5-19) with $l = 0$. Short constraint length (up to and including $K = 14$) noncatastrophic codes with maximal free distance for rates 1/2, 1/3, and 1/4 were investigated and tabulated in [2-4]. Indeed, for rate 1/3 ($n = 3$), the tabulation includes only one correlated convolutional code which occurs for $K = 3$ ($b = 1$) and has the generator matrix

$$G = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{2.5-23}$$

This code was first found by Odenwalder [2-5]. Since rows two and three of G are identical, we must evaluate (2.5-22) for the two combinations $p = 2$, $q = 3$, $l = 0$, and $p = 3$, $q = 3$, $l = 0$. Doing so yields the simple result

$$S_X(f) = 1 + \frac{2}{3} \cos 2\pi fT \tag{2.5-24}$$

Multiplying $S_X(f)$ of (2.5-24) by $S_p(f)$ of (2.2-6) gives the encoder output spectrum $S_c(f)$ which is plotted in Fig. 2-2.

For rate 1/4 ($n = 4$), the tabulation in [2-4] reveals three correlated convolutional codes corresponding to constraint lengths $K = 3$, 4, and 7. For the $K = 3$

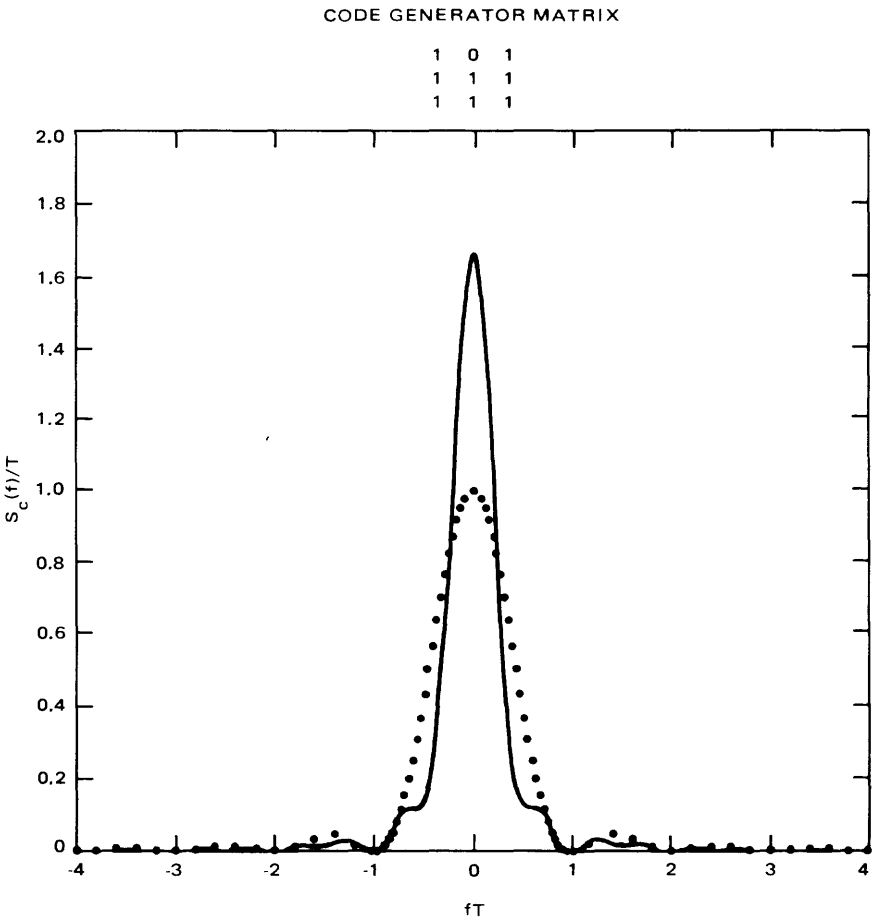


Fig. 2-2. Spectrum for best rate 1/3, constraint length 3, convolutional code. Dotted curve is spectrum of NRZ

code, for example, the free distance is ten (this is the maximum achievable), and the generator matrix is given by

$$G = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{2.5-25}$$

which is identical to (2.5-23) except for the addition of another identical row. Thus, following the same procedure as above, we find that

$$S_X(f) = 1 + \frac{1}{2}[2\cos 2\pi fT + \cos 4\pi fT] \qquad (2.5\text{-}26)$$

which, together with (2.2-6), gives the output encoder spectrum illustrated in Fig. 2-3 for a rate 1/4, constraint length 3 convolutional code and in Fig. 2-4 for a rate 1/3 code.

A second special case of a random NRZ encoder input is one for which $p\star \neq 1/2$ and hence $\bar{a} = 0$. Consider first the simple case of the optimum rate 1/2, constraint length 3 code whose generator matrix is given by

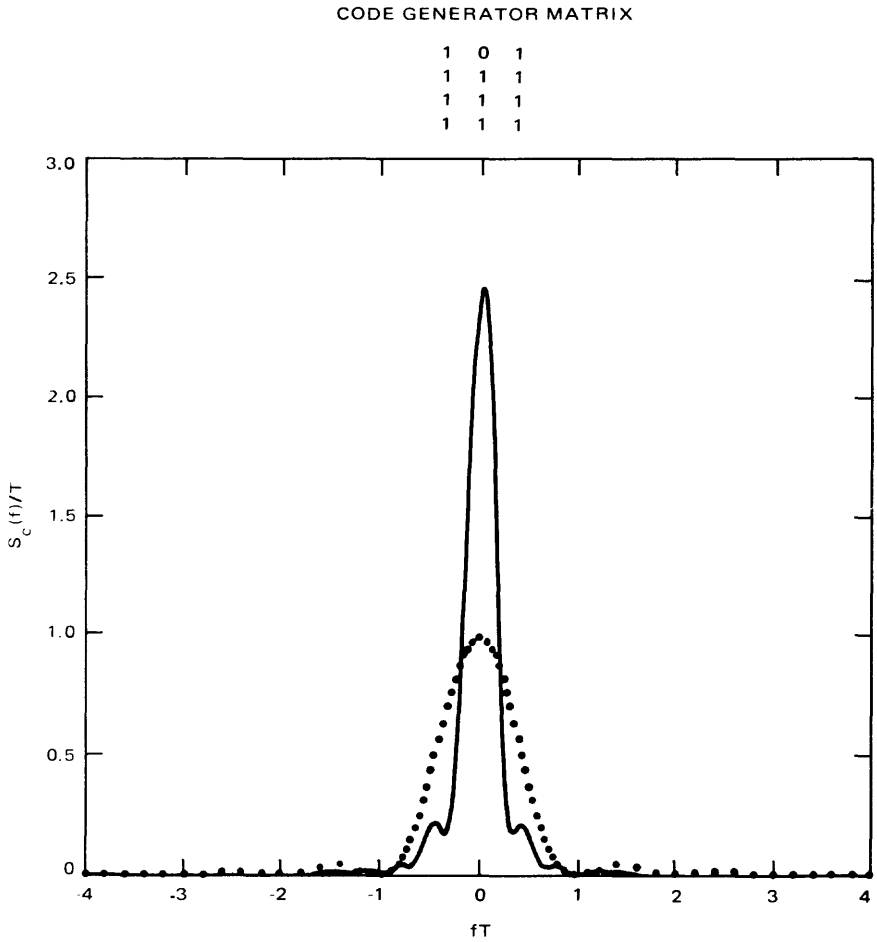


Fig. 2-3. Spectrum for best rate 1/4, constraint length 3, convolutional code. Dotted curve is spectrum of NRZ

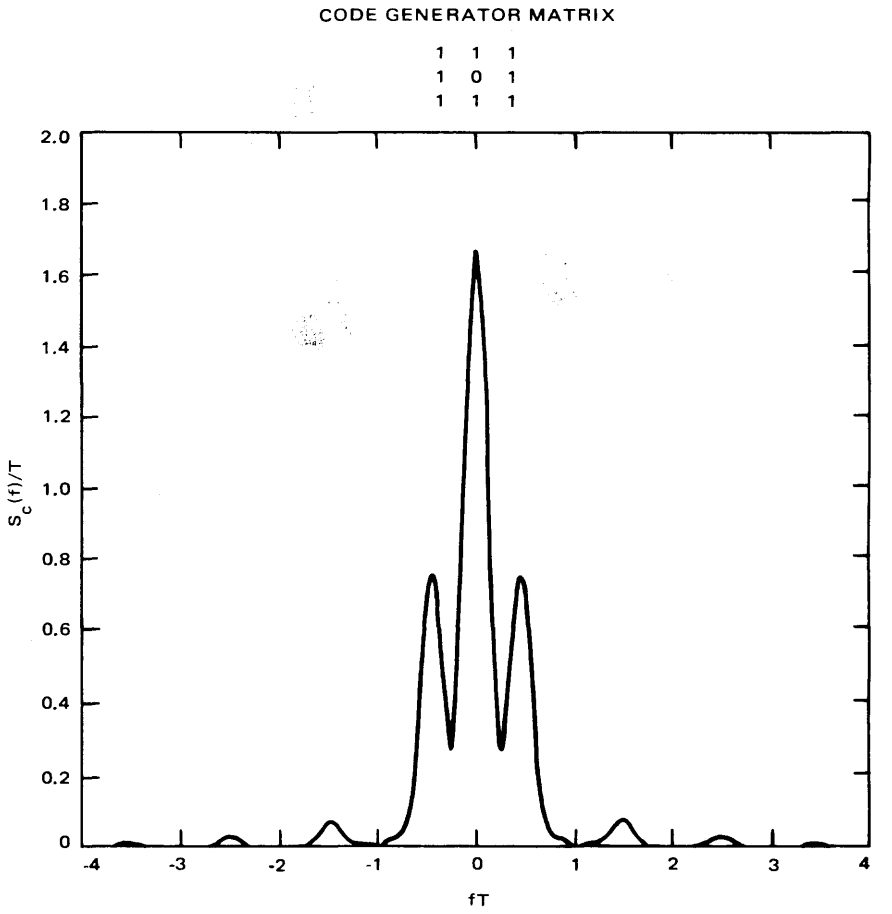


Fig. 2-4. Spectrum for best rate 1/3, constraint length 3, convolutional code

$$G = \begin{bmatrix} 1 & 0 & 1 \\ & & \\ 1 & 1 & 1 \end{bmatrix} \tag{2.5-27}$$

For this code (2.5-13) gives the closed form result [2-2]:

$$S_X(f) = 1 - \frac{1}{2}(\bar{a}^4 - \bar{a}^6) + (\bar{a} + \bar{a}^3 - 2\bar{a}^5)\cos 2\pi fT$$

$$+ (\bar{a}^2 - \bar{a}^6)\cos 4\pi fT + 2(\bar{a}^3 - \bar{a}^5)\cos 6\pi fT$$

$$+ (\bar{a}^2 - \bar{a}^6)\cos 8\pi fT + (\bar{a}^3 - \bar{a}^5)\cos 10\pi fT \tag{2.5-28}$$

Since this code is an uncorrelated convolutional code, (2.5-28) should reduce to $S_X(f) = 1$ for $\bar{a} = 0$, which indeed it does.

Figures 2-5 and 2-6 are illustrations of the continuous component of the encoder output spectrum $S_c(f)$ [obtained by multiplying $S_x(f)$ of 2.5-28) by $S_p(f)$ which is identical to $S_c(f)$ of (2.5-21)] for various values of $p\star = (1 - a)/2$. We observe that the spectrum becomes more and more concentrated as $p\star$ decreases. The continuous component of the input spectrum, on the other hand, remains unaltered in shape as $p\star$ is varied. In particular, for a unit power unbalanced NRZ input, the power spectral density is given by

$$S(f) = (1 - 2p\star)^2 \delta(f) + 4p\star (1 - p\star)\, T_b \left(\frac{\sin \pi f T_b}{\pi f T_b} \right)^2 \qquad (2.5\text{-}29)$$

where T_b is the input bit time.

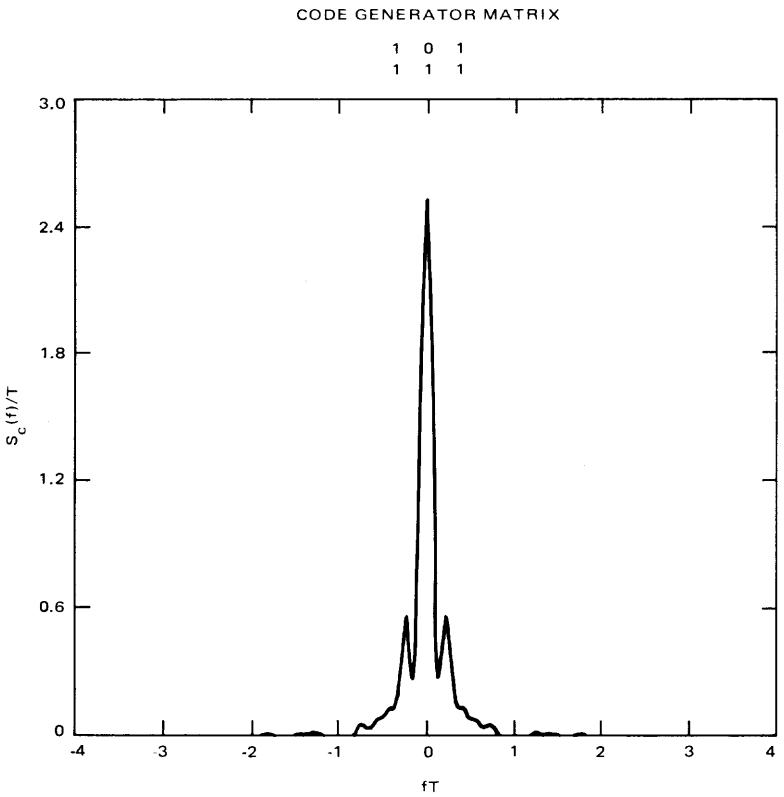


Fig. 2-5. Spectrum for best rate 1/2, constraint length 3, convolutional code; $p^\star = 0.1$

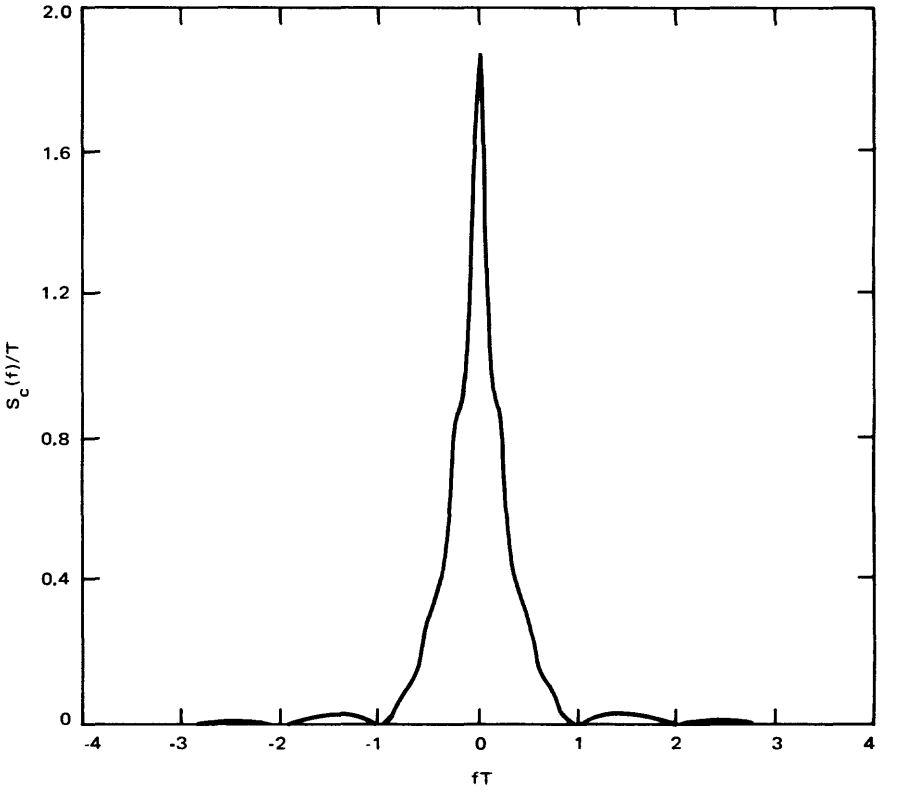


Fig. 2-6. Spectrum for best rate 1/2, constraint length 3, convolutional code; $p^* = 0.3$

The discrete component of the output spectrum $S_d(f)$ is easily found to be [2-2]

$$S_d(f) = (\bar{a}^3 - \bar{a}^2)^2 \sum_{k=1,3,5\ldots}^{\infty} \left(\frac{1}{\pi k}\right)^2 \delta\left(f - \frac{k}{2T}\right) \qquad (2.5\text{-}30)$$

As a second example, consider the optimum rate 1/2, constant length 7 code whose generator matrix is given by [2-4]:

$$G = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \qquad (2.5\text{-}31)$$

Because of the balance in the number of ones in both rows of the matrix, the encoder output will only contain a discrete spectrum at $f = 0$, i.e., $S_d(f) =$

$(\bar{a})^{10}\delta(f)$. The continuous component of the output spectrum can once again be obtained from the product of $S_X(f)$ and $S_p(f)$ where $S_X(f)$ is now given by [2-2]

$$
\begin{aligned}
S_X(f) = {} & 1 - \bar{a}^{10} + (\bar{a}^2 + \bar{a}^6 - 2\bar{a}^{10}) \cos 2\pi fT \\
& + (\bar{a}^4 + \bar{a}^6 - 2\bar{a}^{10}) \cos 4\pi fT + 2(\bar{a}^4 - \bar{a}^{10}) \cos 6\pi fT \\
& + 2(\bar{a}^6 - \bar{a}^{10}) [\cos 8\pi fT + \cos 14\pi fT] \\
& + (\bar{a}^4 + \bar{a}^8 - 2\bar{a}^{10}) \cos 10\pi fT + (\bar{a}^4 + \bar{a}^6 - 2\bar{a}^{10}) \cos 12\pi fT \\
& + 2(\bar{a}^8 - \bar{a}^{10}) [\cos 16\pi fT + \cos 20\pi fT + \frac{1}{2}\cos 22\pi fT \\
& + \cos 24\pi fT + \frac{1}{2}\cos 26\pi fT] + (\bar{a}^6 + \bar{a}^8 - 2\bar{a}^{10}) \cos 18\pi fT \quad (2.5\text{-}32)
\end{aligned}
$$

Figures 2-7 and 2-8 are illustrations of the corresponding continuous spectrum $S_c(f)$ for several values of $p\star = (1 - \bar{a})/2$. As before, we observe that the output encoder spectrum narrows as $p\star$ is decreased.

2.6 Power Spectral Density of a Synchronous Data Stream Generated by an *N*-ary Markov Source

Consider a random N-ary source which every T seconds emits an elementary signal from the set $\{s_i(t); i = 1, 2, \ldots, N\}$ with probability p_i. If the source is Markov, then the sequence of waveforms so generated is characterized by the set of probabilities $\{p_i; i = 1, 2, \ldots, N\}$ (often called the stationary probabilities) and the set of transition probabilities $\{p_{ik}; i, k = 1, 2, \ldots, N\}$. Transition probability p_{ik} is the probability that signal $s_k(t)$ is transmitted in any given transmission interval after the occurrence of the signal $s_i(t)$ in the previous transmission interval. These transition probabilities are conveniently arranged in a transition matrix P defined by

$$
P \triangleq \begin{pmatrix}
p_{11} & p_{12} & \cdots & p_{1N} \\
p_{21} & p_{22} & \cdots & p_{2N} \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
p_{N1} & p_{N2} & \cdots & p_{NN}
\end{pmatrix} \qquad (2.6\text{-}1)
$$

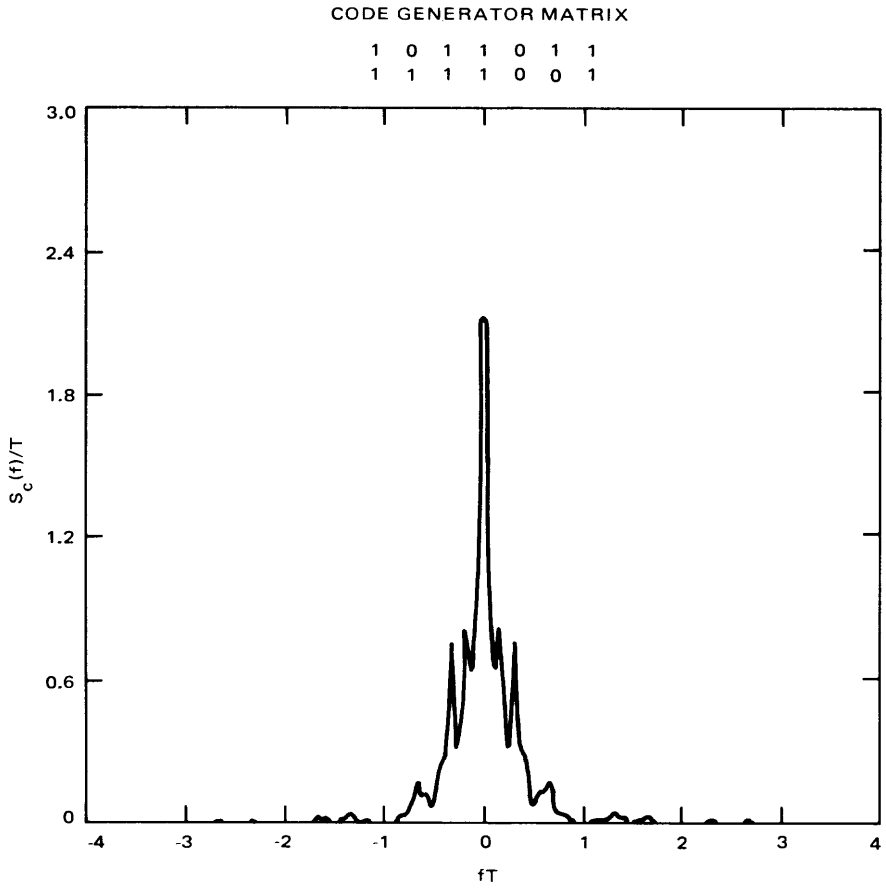


Fig. 2-7. Spectrum for best rate 1/2, constraint length 7, convolutional code; $p^* = 0.1$

From this statistical description of the source, the power spectral density of a data sequence generated by this source is given by [2-6, 2-7]

$$S_m(f) = \frac{1}{T^2} \sum_{n=-\infty}^{\infty} \left| \sum_{i=1}^{N} p_i S_i\left(\frac{n}{T}\right) \right|^2 \delta\left(f - \frac{n}{T}\right)$$

$$+ \frac{1}{T} \sum_{i=1}^{N} p_i |S_i'(f)|^2$$

$$+ \frac{2}{T} Re \sum_{i=1}^{N} \sum_{k=1}^{N} p_i S_i'^{*}(f) S_k'(f) p_{ik}(e^{-j2\pi fT}) \quad (2.6\text{-}2)$$

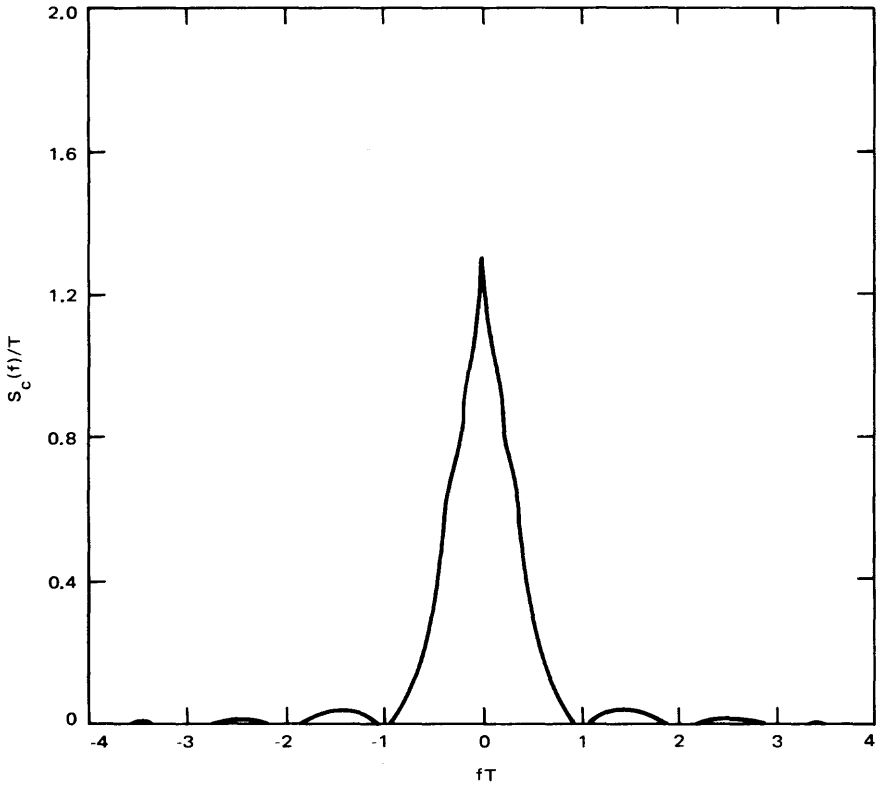


Fig. 2-8. Spectrum for best rate 1/2, constraint length 7, convolutional code; $p^* = 0.3$

where

$$S_i(f) \stackrel{\Delta}{=} \int_0^T s_i(t) e^{-j2\pi ft}\, dt; S_i'(f) \stackrel{\Delta}{=} \int_0^T s_i'(t) e^{-j2\pi ft}\, dt \tag{2.6-3}$$

and

$$p_{ik}(z) \stackrel{\Delta}{=} \sum_{n=1}^{\infty} p_{ik}^{(n)} z^n; s_i'(t) \stackrel{\Delta}{=} s_i(t) - \sum_{k=1}^{N} p_k s_k(t) \tag{2.6-4}$$

In (2.6-2), the asterisk denotes complex conjugate. The quantity $p_{ik}^{(n)}$ is defined as the probability that the elementary signal $s_k(t)$ is transmitted n signaling intervals after the occurrence of $s_i(t)$. Hence, from the properties of Markov sequences, $p_{ik}^{(n)}$ is the ikth element of the matrix P^n. Also by definition,

$$p_{ik}^{(1)} \triangleq p_{ik}.$$

Notice that the first term [i.e., the line (spike) spectrum] of (2.6-2) vanishes when

$$\sum_{i=1}^{N} p_i S_i\left(\frac{n}{T_s}\right) = 0 \tag{2.6-5}$$

which implies that a necessary and sufficient condition for the absence of a line spectrum is that

$$\sum_{i=1}^{N} p_i s_i(t) = 0 \tag{2.6-6}$$

Many special classes of signals exist for which the general power spectral density result of (2.6-2) can be simplified. One important class is that generated by a purely random source, i.e., one that emits an elementary signal in a given signaling interval independent of those emitted in previous signaling intervals. Such a source can be modeled as a degenerate case of a Markov source (i.e., P^{-1} does not exist) whose transition matrix is given by

$$P = \begin{pmatrix} p_1 & p_2 & \cdots & p_N \\ p_1 & p_2 & \cdots & p_N \\ \cdot & \cdot\ \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \cdot\ \cdot \\ p_1 & p_2 & \cdots & p_N \end{pmatrix} \tag{2.6-7}$$

and has the property that $P^n = P$ for all $n \geqslant 1$. In this case, the power spectral density of (2.6-2) simplifies to

$$S_m(f) = \frac{1}{T^2}\sum_{n=-\infty}^{\infty}\left|\sum_{i=1}^{N} p_i S_i\left(\frac{n}{T}\right)\right|^2 \delta\left(f-\frac{n}{T}\right)$$

$$+\frac{1}{T}\sum_{i=1}^{N} p_i(1-p_i)|S_i(f)|^2$$

$$-\frac{2}{T}\sum_{i=1}^{N}\sum_{\substack{k=1\\ i\neq k\\ i<k}}^{N} p_i p_k \, Re\,[S_i(f)S_k^*(f)] \qquad (2.6\text{-}8)$$

which is the generalization of (2.2-5) and (2.4-8) for an N-ary source.

As an example of the application of (2.6-2), consider the Miller coding scheme [2-8] which can be modeled as a 4-ary Markov source whose stationary probabilities are all equal to 1/4 and those transition matrix is given by

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix} \qquad (2.6\text{-}9)$$

Another property of the Miller code is that it satisfies the recursion relation

$$P^{4+l}S = -\frac{1}{4}P^l S \quad l \geqslant 0 \qquad (2.6\text{-}10)$$

where S is the signal correlation matrix whose ikth element is defined by

$$s_{ik} \triangleq \frac{1}{T}\int_0^T s_i(t)s_k(t)\,dt \quad i,k = 1,2,3,4 \qquad (2.6\text{-}11)$$

For the Miller code, the four elementary signals are defined by

$$\left.\begin{aligned} s_1(t) &= -s_4(t) = A \qquad \text{for } 0 \leqslant t \leqslant T \\ s_2(t) &= -s_3(t) = \begin{cases} A & \text{for } 0 \leqslant t < T/2 \\ -A & \text{for } T/2 \leqslant t \leqslant T \end{cases} \end{aligned}\right\} \qquad (2.6\text{-}12)$$

Substituting (2.6-12) into (2.6-11) and arranging the results in the form of a matrix, we get

$$S = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \tag{2.6-13}$$

Finally, using (2.6-9), (2.6-10) and (2.6-13) in the general power spectral density result of (2.6-2) yields the result for the Miller code:

$$\begin{aligned} \frac{S(f)}{E} = \frac{1}{2\theta^2(17 + 8\cos 8\theta)} (23 &- 2\cos\theta - 22\cos 2\theta \\ &- 12\cos 3\theta + 5\cos 4\theta \\ &+ 12\cos 5\theta + 2\cos 6\theta \\ &- 8\cos 7\theta + 2\cos 8\theta) \end{aligned} \tag{2.6-14}$$

where $\theta \overset{\Delta}{=} \pi f T$.

An example of the application of (2.6-7) is NRZ signaling where $N = 2$ and the elementary signal is a rectangular pulse of width T and amplitude A. Thus, using (2.6-3), we get

$$S_1(f) = -S_2(f) = AT \exp(-j\pi f T) \frac{\sin(\pi f T)}{\pi f T} \tag{2.6-15}$$

Finally, substituting (2.6-15) into (2.6-8) and letting $E = A^2 T$, gives

$$\frac{S(f)}{E} = \frac{1}{T}(1 - 2p)^2 \delta(f) + 4p(1 - p) \left[\frac{\sin^2(\pi f T)}{(\pi f T)^2} \right]$$

where $p = 1/2$, the dc spike at the origin disappears and

$$\frac{S(f)}{E} = \frac{\sin^2(\pi f T)}{(\pi f T)^2} \tag{2.6-16}$$

Another common example is bi-phase or Manchester code. Here $N = 2$ and

$$\left.\begin{aligned} s_1(t) &= A \quad \text{for } 0 \leqslant t \leqslant T/2 \\ s_1(t) &= -A \quad \text{for } T/2 \leqslant t \leqslant T \\ s_1(t) &= -s_2(t) \end{aligned}\right\} \tag{2.6-17}$$

Substituting the Fourier transform of (2.6-17) into (2.6-8) gives

$$\frac{S(f)}{E} = \frac{1}{T}(1-2p)^2 \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \left(\frac{2}{n\pi}\right)^2 \delta\left(f - \frac{n}{T}\right) + 4p(1-p)\left[\frac{\sin^4(\pi fT/2)}{(\pi fT/2)^2}\right] \tag{2.6-18}$$

For $p = 1/2$, the line spectrum disappears and

$$\frac{S(f)}{E} = \frac{\sin^4(\pi fT/2)}{(\pi fT/2)^2} \tag{2.6-19}$$

The power spectral densities of the NRZ, bi-phase, and delay modulation signaling schemes as given by (2.6-16), (2.6-19), and (2.6-14), respectively, are plotted in Fig. 2-9.

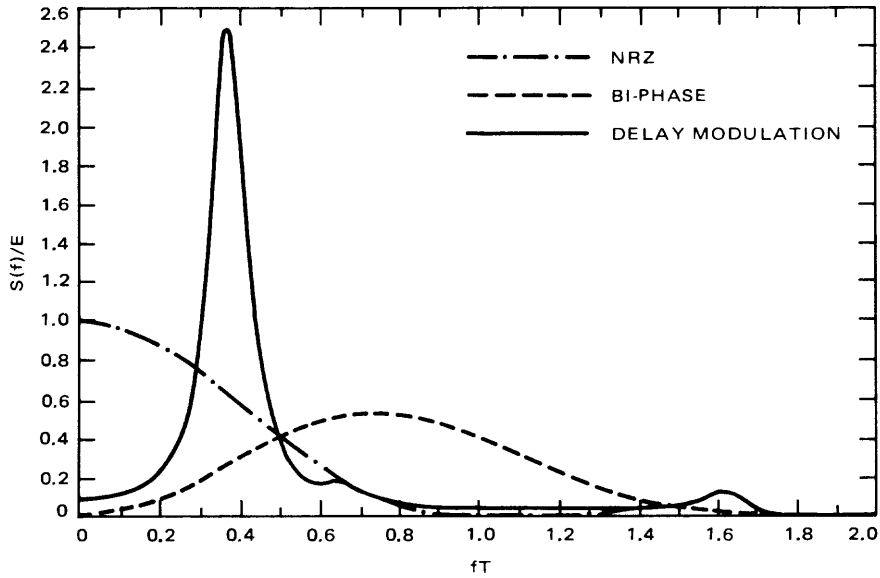


Fig. 2-9. Two-sided spectral densities of NRZ, bi-phase, and delay modulation waveforms

References

2-1. Scholtz, R. A., "How Do You Define Bandwidth?" *Proceedings of the International Telemetering Conference*, Los Angeles, Calif., Oct. 1972, pp. 281-288.

2-2. Divsalar, D., and Simon, M. K., "Spectral Characteristics of Convolutionally Coded Digital Signals," Publication 79-93, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 1, 1979.

2-3. Jolley, L. B. W., *Summation of Series*, Dover Publications, New York, N.Y., 1961.

2-4. Larsen, K. J., "Short Convolutional Codes with Maximal Free Distance for Rates 1/2, 1/3, and 1/4." *IEEE Transactions on Information Theory*, Vol. IT-19, No. 3, May 1973, pp. 371-372.

2-5. Odenwalder, J. P., "Optimal Decoding of Convolutional Codes," PhD Disseration, University of California, Los Angeles, Calif., 1970.

2-6. Titsworth, R. C., and Welch, L. R., *Power Spectra of Signals Modulated by Random and Pseudorandom Sequences*, Technical Report 32-140, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 1961.

2-7. Lindsey, W. C., and Simon, M. K., *Telecommunication Systems Engineering*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1973, Chapter 1.

2-8. Hecht, M., and Guida, A., "Delay Modulation," *Proceedings of the IEEE*, Vol. 57, No. 7, July 1969, pp. 1314-1316.

Chapter 3
Receiver Design and Performance Characteristics

Marvin K. Simon and Joseph H. Yuen

3.1 Introduction

This chapter presents an introductory discussion of the receivers used in deep space communications. The basic structures, principles of operations, and characteristics of these receivers are examined to provide an understanding of the important parameters in the design and performance evaluation.

There are two basic types of deep space communications receivers: the phase-locked loop (PLL) receiver and the Costas loop receiver—and variations of these. Phase-locked loop receivers are used in residual carrier systems—the traditional deep space communications systems. Costas loop receivers are used in suppressed carrier systems, which have been increasingly employed in recent years, in, for example, the Tracking and Data Relay Satellite System (TDRSS), the Space Shuttle, the Multimegabit Telemetry Demodulator/Detector (MTDD) System in the DSN, the Navigation Satellite Timing and Ranging Global Positioning System (NAVSTAR GPS), and the proposed Venus Orbiting Imaging Radar (VOIR) for the mapping of Venus's surface.

Phase-locked loop receivers are essential to deep space communications. They make the demodulation of phase-modulated carriers possible. They are integral to doppler tracking systems. Indeed, every spacecraft transponder has a PLL receiver, and the heart of a NASA Deep Space Station (DSS) closed-loop receiver is a PLL. The spacecraft PLL receiver demodulates command and ranging and provides a reference for generating a downlink carrier which is phase coherent with the uplink carrier (but shifted in frequency). On the ground, the DSS PLL receiver generates a reference signal which is phase coherent with the downlink carrier. This reference is used to demodulate telemetry and ranging and to obtain doppler data.

The Deep Space Network was developed at a time when weak signals and low data rates dominated. As a result, residual carrier phase modulation schemes with phase-locked loops tracking the residual carrier component were employed. However, over the years, technological advances in antennas, transmitters, and signal processing have caused a marked increase in the available signal power and hence the achievable data rates. In fact, essentially all of the current or near future deep space missions are already operating at or near the upper limit of the current DSN data rate capability. This leads to the choice of suppressed carrier systems employing Costas loop receivers. However, besides telemetry data rate considerations, for other uses of the deep space communication system, e.g., for ranging and radio science (see chapter 4), the residual carrier system may still be preferred.

3.2 Phase-Locked Loop Receivers

Phase-locked loop receivers designed for spacecraft and the DSN are invariably of the multiple-conversion heterodyne type. Three such examples are shown in Figs. 3-1 through 3-3. Figure 3-1 is the PLL receiver used on both Voyager and Galileo, but it is in many respects typical of spacecraft PLL receivers. Figures 3-2 and 3-3 are the Block III and Block IV DSS receivers, respectively. The noise temperature of a receiver is predominantly determined by the design of the first receiver stage because of its very high gain. For the DSS receivers, the first stage is the maser amplifier at the antenna. For spacecraft receivers, it is usually the first mixer. The heterodyne design is employed to translate the RF signal down to a frequency for which stable phase detectors can be built. Automatic gain control (AGC) is required to provide a signal whose amplitude is within the dynamic range of the intermediate frequency (IF) amplifier stages. The bandpass limiter (BPL) minimizes the total mean-square error of the loop over a wide range of input signal-to-noise ratios. This configuration has been shown in [3-1] to provide near-optimum PLL performance.

This section presents an introductory description of the design and performance of PLL receivers used in deep space communications. For a more complete treatment of PLLs, the reader is referred to [3-2] through [3-8].

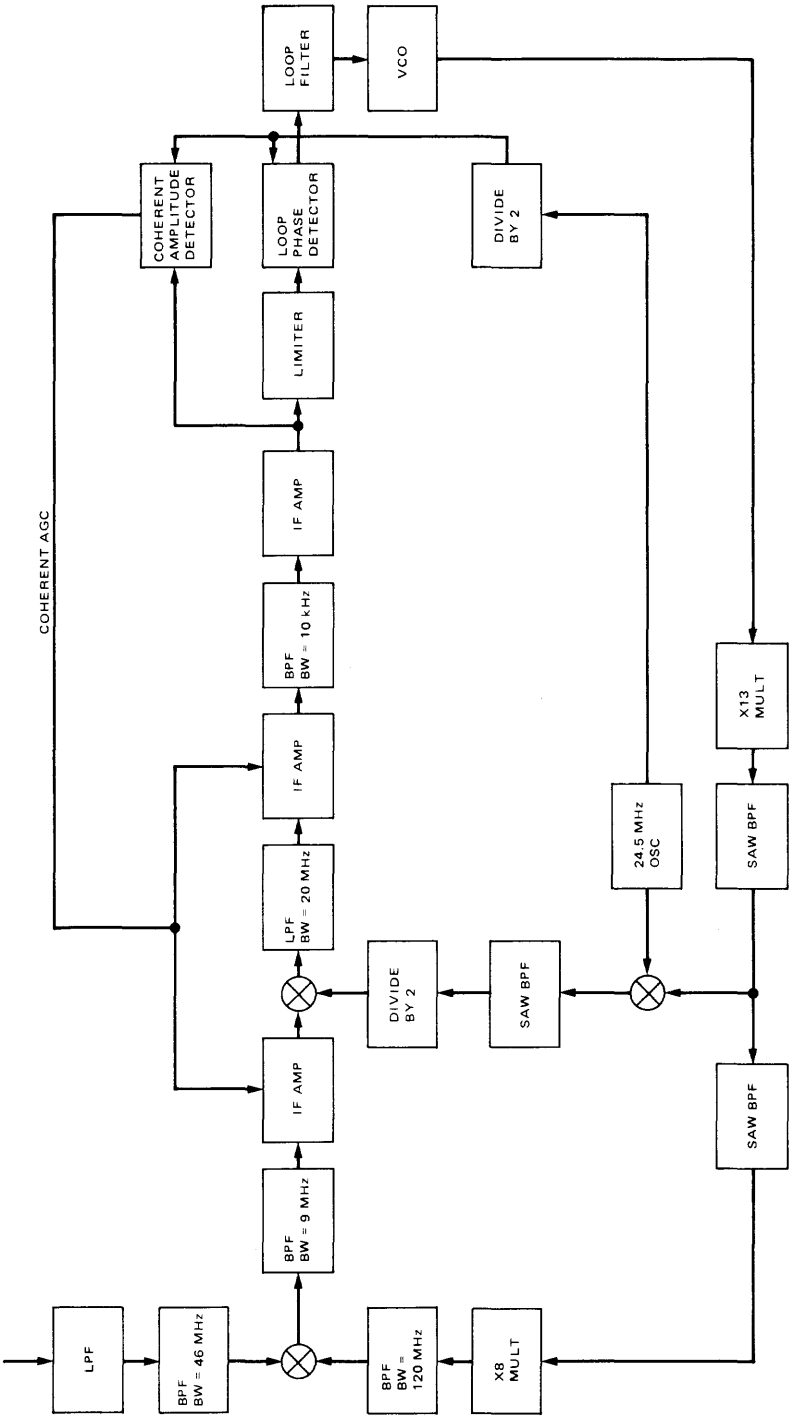


Fig. 3-1. Typical spacecraft PLL receiver

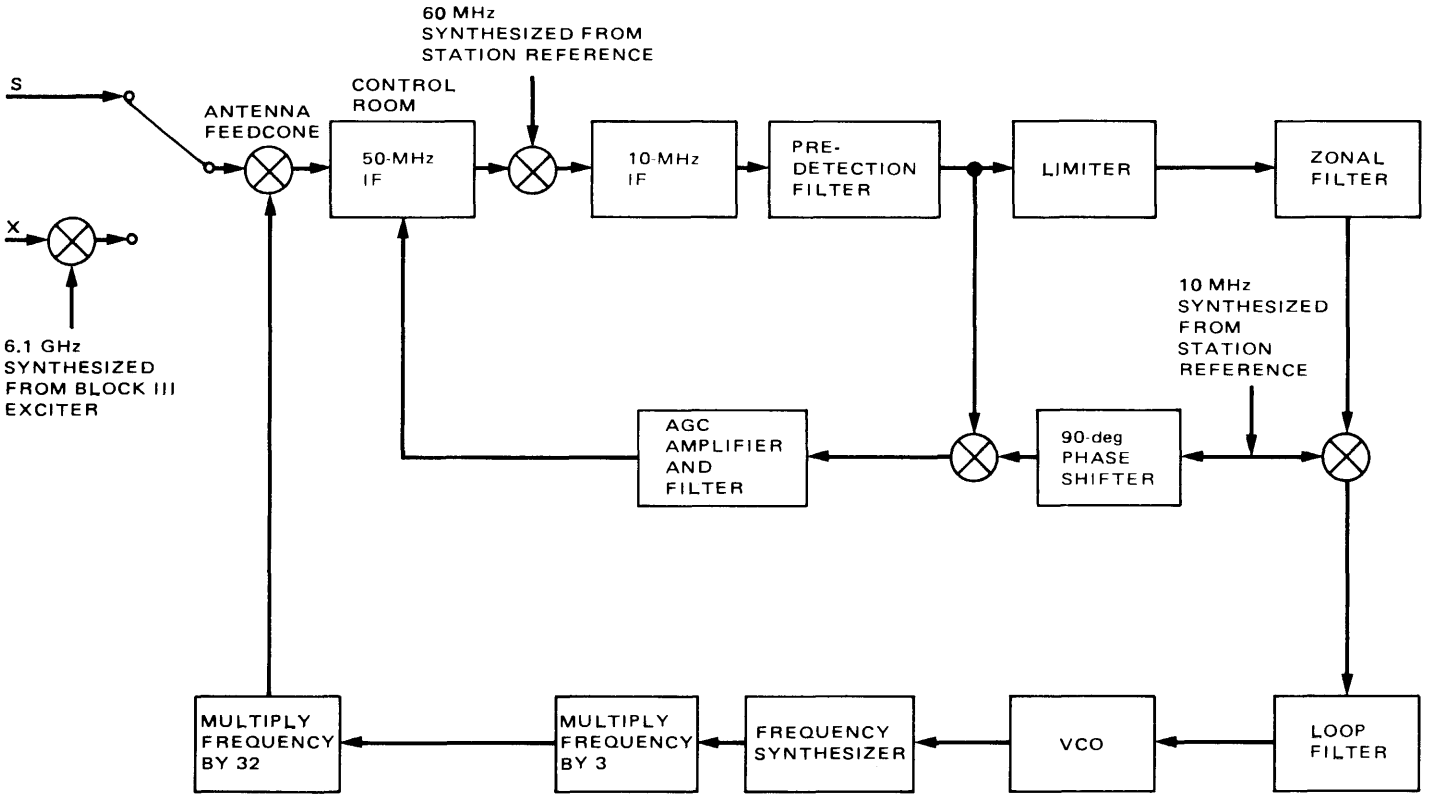


Fig. 3-2. Block III PLL receiver

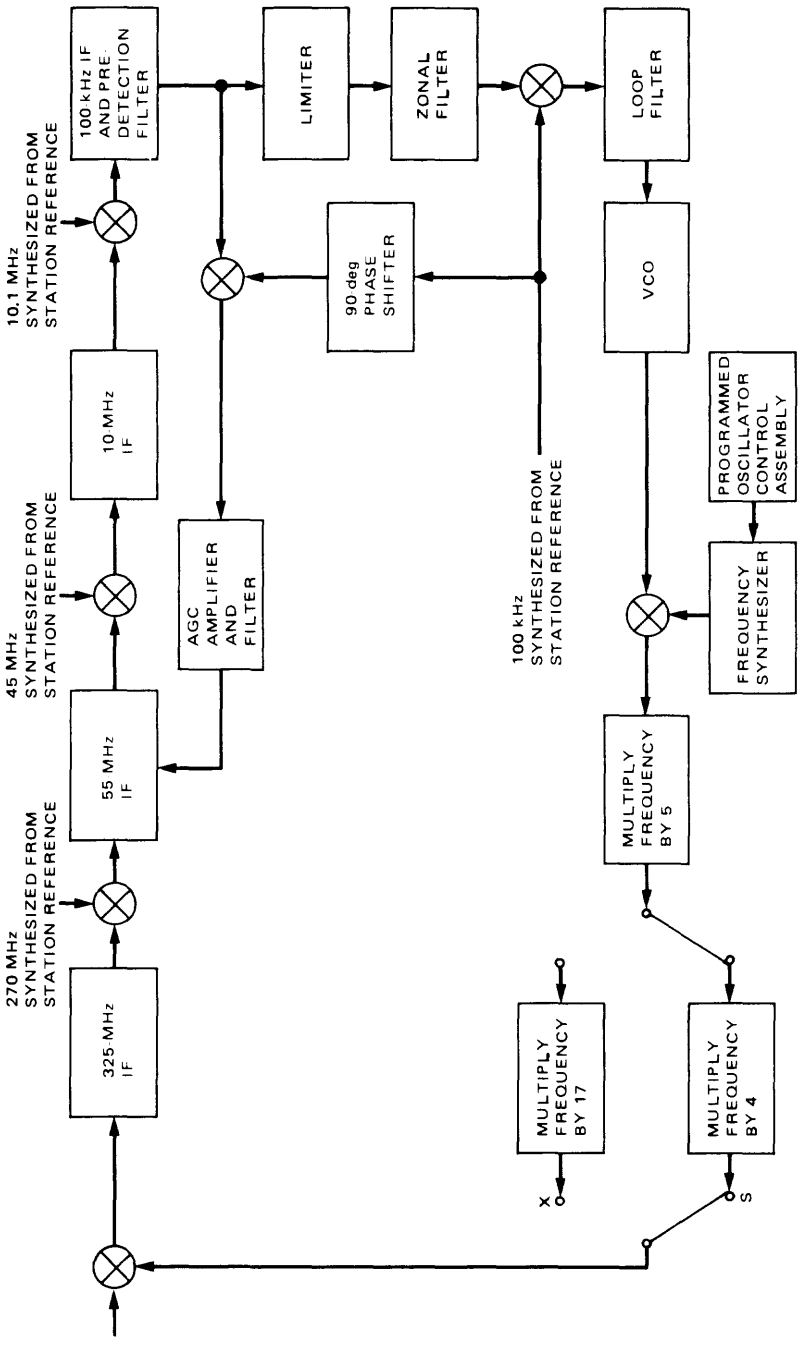


Fig. 3-3. Block IV PLL receiver

3.2.1 Phase-Locked Loop Operation

A phase-locked loop consists of three essential elements, a phase detector or multiplier, a loop filter, and a voltage controlled oscillator (VCO), as shown in Fig. 3-4. The input signal $y(t)$ is assumed to be the sum of a sinusoidal signal $s(t)$ and a channel additive noise $n_i(t)$,

$$y(t) = s(t) + n_i(t) = \sqrt{2P} \sin [\omega_0 t + \theta(t)] + n_i(t) \qquad (3.2\text{-}1)$$

where ω_0 is the signal frequency in radians, P is the input signal power, $\theta(t)$ is the input signal phase, and

$$n_i(t) = \sqrt{2} [n_c(t) \cos \omega_0 t - n_s(t) \sin \omega_0 t] \qquad (3.2\text{-}2)$$

is a narrowband process. Both $n_c(t)$ and $n_s(t)$ are assumed to be statistically independent, stationary, white Gaussian noise processes of single-sided spectral density N_0 W/Hz.

The reference signal at the VCO output is

$$r(t) = \sqrt{2} K_1 \cos [\omega_0 t + \hat{\theta}(t)] \qquad (3.2\text{-}3)$$

where K_1 is the rms output of the VCO, and $\hat{\theta}(t)$ is the phase estimate of $\theta(t)$ generated by the VCO. The signal $y(t)$ is multiplied by $r(t)$ to generate an error signal

$$\epsilon(t) = K_M y(t) r(t) \qquad (3.2\text{-}4)$$

where K_M is the multiplier gain. This error signal $\epsilon(t)$ is filtered by the loop filter to remove the high-frequency terms which are larger than and equal to $2\omega_0$, to produce the VCO control signal

$$z(t) = F(p) \epsilon(t) \qquad (3.2\text{-}5)$$

where $p = d/dt$ is the Heaviside operator. This filtered error signal $z(t)$ drives the VCO frequency and phase to match the input signal frequency and phase such that

$$\hat{\theta}(t) = K_V \int^t z(\xi)\, d\xi \qquad (3.2\text{-}6)$$

where K_V is the VCO gain in radians/sec/V.

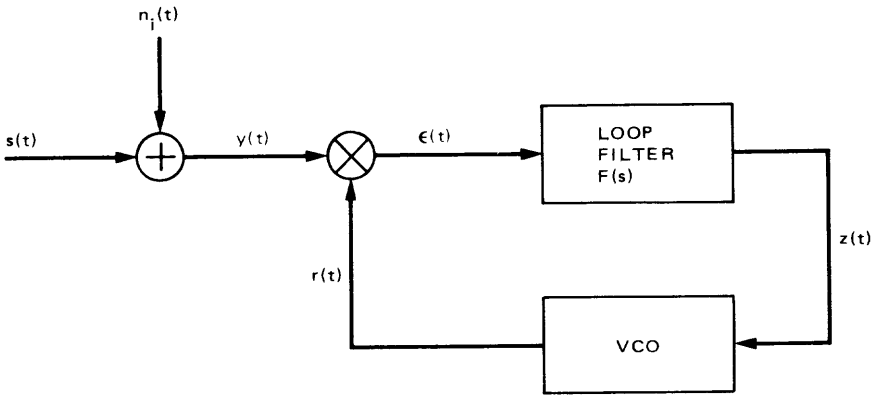


Fig. 3-4. A phase-locked loop model

Let the phase error be defined as

$$\varphi(t) = \theta(t) - \hat{\theta}(t) \tag{3.2-7}$$

Then, using (3.2-1) through (3.2-6), we obtain the characteristic equation which describes the dynamic behavior of the PLL,

$$\varphi(t) = \theta(t) - \frac{K\,F(p)}{p}\,[\sqrt{P}\,\sin\,\varphi(t) + n(t)] \tag{3.2-8}$$

where

$$K = K_1 K_M K_V \tag{3.2-9}$$

and the noise process $n(t)$ depends on $n_i(t)$, $\theta(t)$ and $\varphi(t)$. For deep space communication, it may be shown—[3-2] and [3-3]—that $n(t)$ is approximated by a low-pass white Gaussian noise process with the same spectral density as that of the original additive noise $n_i(t)$.

3.2.2 Noise-Free Phase-Locked Loop Characteristics

This section presents some of the important PLL characteristics in the absence of channel noise.

3.2.2.1 Closed-loop transfer function. The closed-loop transfer function $H(s)$ is defined by

$$H(s) = \frac{\hat{\theta}(s)}{\theta(s)} \tag{3.2-10}$$

where $\hat{\theta}(s)$ and $\theta(s)$ are the Laplace transforms of $\hat{\theta}(t)$ and $\theta(t)$, respectively.

When the phase error $\varphi(t)$ is small at all times,

$$\sin \varphi(t) \cong \varphi(t) \tag{3.2-11}$$

is a good linear approximation. This would be the case with a strong signal-to-noise power ratio. From (3.2-8), in the absence of noise, we have

$$H(s) = \frac{\hat{\theta}(s)}{\theta(s)} = \frac{\sqrt{P}\,K\,F(s)}{s + \sqrt{P}\,K\,F(s)} \tag{3.2-12}$$

The DSN and spacecraft receivers typically use passive second-order loop filters[1] of the form

$$F(s) = \frac{1 + \tau_2 s}{1 + \tau_1 s} \tag{3.2-13}$$

Substituting (3.2-13) into (3.2-12), we obtain

$$H(s) = \frac{1 + \tau_2 s}{1 + \left(\tau_2 + \dfrac{1}{\sqrt{P}\,K}\right)s + \dfrac{\tau_1}{\sqrt{P}\,K}\,s^2} \tag{3.2-14}$$

In deep space telecommunications, the time constants are usually chosen with $\tau_1 >> \tau_2$, so that $F(s)$ approaches the perfect integrator,

$$F(s) \cong \frac{1 + \tau_2 s}{\tau_1 s} \tag{3.2-15}$$

for $\tau_1 s >> 1$.

3.2.2.2 Loop bandwidth. The two-sided loop bandwidth W_L is defined to be

$$W_L = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} |H(s)|^2 \, ds \tag{3.2-16}$$

in Hz. The single-sided loop bandwidth is

$$B_L = \frac{1}{2} W_L \tag{3.2-17}$$

[1] The order of a PLL is $k + 1$ when k is the number of poles in the loop transfer function $F(s)$.

This loop bandwidth B_L (in Hz) so defined is the bandwidth of an ideal low-pass filter whose output variance is $N_0 B_L$ when the input is a white noise process of single-sided spectral density N_0 W/Hz. For the closed-loop transfer function in (3.2-14), the loop bandwidth is

$$W_L = \frac{1 + \dfrac{\sqrt{P} K \tau_2^2}{\tau_1}}{2\tau_2 \left(1 + \dfrac{1}{\sqrt{P} K \tau_2}\right)} \tag{3.2-18}$$

If we define the loop damping parameter r by

$$r = \frac{\sqrt{P} K \tau_2^2}{\tau_1} \tag{3.2-19}$$

and assume $r \tau_1 >> \tau_2$, then from (3.2-18) we obtain the approximation

$$W_L \cong \frac{1 + r}{2\tau_2} \tag{3.2-20}$$

3.2.2.3 **Steady-state phase error.** In deep space telecommunications, the input phase typically consists of modulation and of doppler due to the radial motion of the spacecraft relative to the tracking station of the DSN; i.e.,

$$\theta(t) = m(t) + d(t) \tag{3.2-21}$$

where $m(t)$ is the modulation and $d(t)$ is the doppler. This doppler signal can be expanded in a Taylor series so that

$$d(t) = \phi_0 + \Omega_0 t + \frac{\Lambda_0}{2} t^2 + \cdots + \frac{x_k t^k}{k!} + \cdots \tag{3.2-22}$$

where the first three terms are the most important for typical spacecraft trajectories, with the following physical interpretations:

ϕ_0 = initial phase offset of the incoming signal from the free-running VCO phase

Ω_0 = frequency offset of the incoming signal from the free-running VCO frequency

Λ_0 = rate of change of the incoming signal frequency

If $d(t)$ has n terms in the Taylor series, then a steady-state phase error

$$\varphi_{ss} = \lim_{t \to \infty} \varphi(t)$$

occurs if

$$n = k + 1$$

where k is the number of poles in the filter $F(s)$.

This steady-state phase error vanishes if $n < k + 1$. If $n > k + 1$, the phase error $\varphi(t)$ will become unbounded and the loop will eventually lose lock.

For the second-order loop with the filter of (3.2-13), and if $d(t)$ has only the first two terms in (3.2-22), i.e.,

$$d(t) = \phi_0 + \Omega_0 t \tag{3.2-23}$$

then the PLL tracks this Doppler with a steady-state phase error

$$\varphi_{ss} = \sin^{-1} \frac{\Omega_0}{\sqrt{P}K}$$

providing that the loop is initially in lock (i.e., $\varphi(t) = 0$) and no noise is present in the loop. The steady-state phase error is commonly referred to as static phase error.

3.2.2.4 Loop pull-in range. Given that the loop is out of lock, the PLL will be able to acquire lock in a finite period of time if $d(t)$ is given in (3.2-23) and Ω_0 is less than a certain magnitude called the loop pull-in range γ_m. That is, the loop will lock if

$$|\gamma_m| = \left|\frac{\Omega_0}{\sqrt{P}K}\right| < \frac{\dfrac{r W_L}{r+1}}{\dfrac{0.95\, r W_L}{r+1} + 0.04\sqrt{P}K} \tag{3.2-24}$$

When a small rate Ω_0 is present, the phase error will eventually reach a maximum value for which the loop will drop lock. The frequency offset corresponding to this condition is called the hold-in range of the loop, γ_{mh}, where

$$|\gamma_{mh}| = \left|\frac{\Omega_0}{\sqrt{P}K}\right| < \frac{\dfrac{r W_L}{r+1}}{\dfrac{0.85\, r\, W_L}{r+1} + 0.16\sqrt{P}K} \tag{3.2-25}$$

The formulas in (3.2-24) and (3.2-25) were experimentally verified, empirically determined [3-9], and appear to be valid for $0 < \sqrt{r}/\sqrt{P}K\tau_2 < 2$.

If the initial frequency offset Ω_0 is within the loop pull-in range, the time required to achieve frequency lock is given by [3-10]

$$t_{freq\ acq} = \pi \sqrt{\frac{2\tau_2}{\tau_1}} \left(\frac{r+1}{r}\right) \frac{\Omega_0^2}{W_L} \tag{3.2-26}$$

where W_L is the PLL loop bandwidth given in (3.2-18), r is the loop damping parameter given in (3.2-19), and the signal-to-noise power ratio (SNR) in the loop is greater than 10 dB. For SNR less than 10 dB, the equation consistently gives too small a value for the time to achieve lock.

3.2.2.5 Maximum sweep rate. Also of interest is the maximum Doppler rate for which the loop can acquire phase lock. This is termed the maximum sweep rate and is of interest because the DSN will sweep the uplink transmitter frequency over some range in order to acquire initial phase lock in the spacecraft receiver. The maximum sweep rate must be determined experimentally, but a good approximation may be found from [3-11]

$$\dot{f}_{acq} = \frac{(1 - \rho^{-1/2}) \left(\dfrac{\alpha_l}{\alpha_{lo}}\right)}{\pi \tau_2^2} \tag{3.2-27}$$

where

α_l = limiter signal amplitude suppression factor, discussed below

α_{lo} = limiter signal amplitude suppression factor, at loop threshold

ρ = signal-to-noise power ratio in the loop bandwidth

$\dot{f}_{acq}$ = the sweep rate that provides 90% probability of acquisition

According to Gardner [3-8], $\dot{f}_{acq}$ as predicted by (3.2-11) should be reduced by $\sqrt{2}$ to compensate for an error in the value of loop gain used in [3-11].

3.2.3 Phase-Locked Loop Characteristics in the Presence of Noise

Thermal noise is generated in the receiver, usually in the first front-end amplifier stage. This white Gaussian noise, which is added to the received signal as described in section 3.2.1, prevents the PLL receiver from making a perfect phase reference estimate even in the absence of Doppler.

The behavior of PLL in the presence of noise is important to the telecommunications system engineer because he must design the tracking, telemetry, and command systems to operate in the presence of noise to meet prescribed data quality and quantity.

3.2.3.1 **Phase error variance.** Under the linear approximation of (3.2-11), the variance of the phase error due to additive channel noise $n_i(t)$ is given by [3-3]

$$\sigma_\varphi^2 = \frac{N_0 W_L}{2P} \tag{3.2-28}$$

where W_L is the two-sided loop bandwidth defined in (3.2-16), and N_0 is again the single-sided noise spectral density of $n_i(t)$.

3.2.3.2 **Phase error probability density function.** Based on the Fokker–Planck method, the probability density function (pdf) of the nonstationary phase error is developed [3-3], which in the steady-state gives an unbounded variance due to the cycle clipping phenomenon associated with PLLs. The Fokker–Planck method was first successfully applied in [3-12], [3-13], and [3-14] to derive the pdf of the phase error of the first-order loop by recognizing that the phase error reduced modulo 2π is stationary and possesses a bounded variance. A first attempt to extend the Fokker–Planck method to a second-order PLL is given in [3-15], which suggests the possibility of approximating the second-order loop phase error pdf by the more trackable results obtained from analysis of the first-order loop.

The actual phase error process $\varphi(t)$ in a PLL undergoes diffusion much like a particle in Brownian motion; hence, the variance of the phase error becomes infinite in the steady-state. Previous work [3-12 through 3-14] in determining the pdf of the phase error in the steady-state of a first-order loop was accomplished by reducing the phase error modulo 2π to a process $\phi(t)$. For finding telemetry error probabilities, for example, this reduction gives sufficient information; whereas for estimating tracking accuracy, the statistical dynamics of the $\varphi(t)$ process itself must be studied.

To completely describe the $\varphi(t)$ process, one must account for the component of its variance that results from diffusion—that is, cycle slipping. The steady-state effect of cycle slipping must be combined with the variance of the phase error reduced modulo 2π to reflect the overall performance of the PLL.

3.2.3.2.1 *The second-order loop case.* The DSN and spacecraft receivers typically use second-order PLLs. For a second-order PLL with loop filter of (3.2-13), the steady-state pdf of the modulo 2π reduced phase error $\phi(t)$ is given to a good approximation by [3-3],

$$p(\phi) = \frac{\exp(\beta\phi + \alpha\cos\phi)}{4\pi^2 \exp(-\pi\beta)|I_{j\beta}(\alpha)|^2} \int_{\phi}^{\phi+2\pi} \exp(-\beta x - \alpha\cos x)\, dx \tag{3.2-29}$$

where $I_\nu(x)$ is the modified Bessel function of order ν and argument x. The domain of definition for ϕ in (3.2-29) is any interval of width centered about any lock point $2n\pi$, with n an arbitrary integer. The parameters α and β in (3.2-29) are related to the various system parameters by

$$\alpha = \left(\frac{r+1}{r}\right)\rho - \frac{1 - \left(\frac{\tau_2}{\tau_1}\right)}{r\,\sigma_G^2}$$

and

$$\beta = \left(\frac{r+1}{r}\right)^2 \frac{\rho}{2W_L}\left[\Omega_0 - \sqrt{P}K\left(1 - \frac{\tau_2}{\tau_1}\right)\overline{\sin\phi}\right] + \alpha\,\overline{\sin\phi} \tag{3.2-30}$$

where

$$\rho = \frac{2P}{N_0 W_L} \tag{3.2-31}$$

is the signal-power-to-noise-spectral-density ratio in the loop bandwidth, and

$$G = \sin\phi - \overline{\sin\phi} \tag{3.2-32}$$

where σ_G^2 is the variance of G, W_L is defined in (3.2-16), r is given by (3.2-19), Ω_0 is the frequency offset of the incoming signal from the free-running VCO frequency, τ_1 and τ_2 are parameters of the loop filter $F(s)$ of (3.2-13), P is the signal power, and K is given by (3.2-9).

The phase error variance can be obtained by using (3.2-29) in the following equation:

$$\sigma_\phi^2 = \int_{-\pi}^{\pi} \phi^2 p(\phi) d\phi - \left(\int_{-\pi}^{\pi} \phi p(\phi) d\phi \right)^2 \tag{3.2-33}$$

3.2.3.2.2 *The first-order loop case.* In this case, the analysis is much simplified. For a first-order PLL with loop filter

$$F(s) = 1 \tag{3.2-34}$$

the pdf of $\phi(t)$ was shown by [3-2] and [3-3] to be

$$p(\phi) = \frac{\exp(\rho \cos \phi)}{2\pi I_0(\rho)} \tag{3.2-35}$$

where $I_0(\rho)$ is the zeroth order imaginary Bessel function of argument ρ.

The phase variance of ϕ can be obtained by substituting (3.2-35) into (3.2-33).

3.2.3.2.3 *Comparison of results.* Analytical results for the variance of the phase error $\phi(t)$ have been obtained by (1) the linear model, given in (3.2-28) where $\sigma_\varphi^2 = \sigma_\phi^2$, (2) the second-order loop pdf, given in (3.2-29) with (3.2-33), and (3) the irst-order loop pdf, given in (3.2-35) with (3.2-33). These are plotted in Fig. 3-5 and compared with experimental results obtained in [3-15]. We observe that both the first-order loop model and the second-order loop model compare well with experimental results. Hence for phase error variance computation, the simpler first-order loop model can be used.

3.2.3.3 Cycle slipping. A PLL slips a cycle when the magnitude of its phase error $\varphi(t)$ exceeds 2π radians. The occurrence of a cycle slip is a random event caused by the noise in the PLL. Cycle slip introduces errors in Doppler tracking. The two parameters characterizing cycle slip are the mean time to first cycle slip τ and the average number of cycle slips per second $\bar{S}$.

The mean time to first cycle slip is defined as the average time the phase error $\varphi(t)$ takes to go from 0 to $\pm 2\pi$ radians, assuming the PLL is initially in phase lock; i.e., $\varphi(t) = 0$. Here τ can be computed by, [3-3],

$$\tau = \frac{1}{W_L} \left(\frac{r+1}{r} \right)^2 \frac{\rho}{2} \int_{-2\pi}^{2\pi} \int_{-2\pi}^{\varphi} [C - u(x)] \exp [U(x) - U(\phi)] \, dx d\varphi \tag{3.2-36}$$

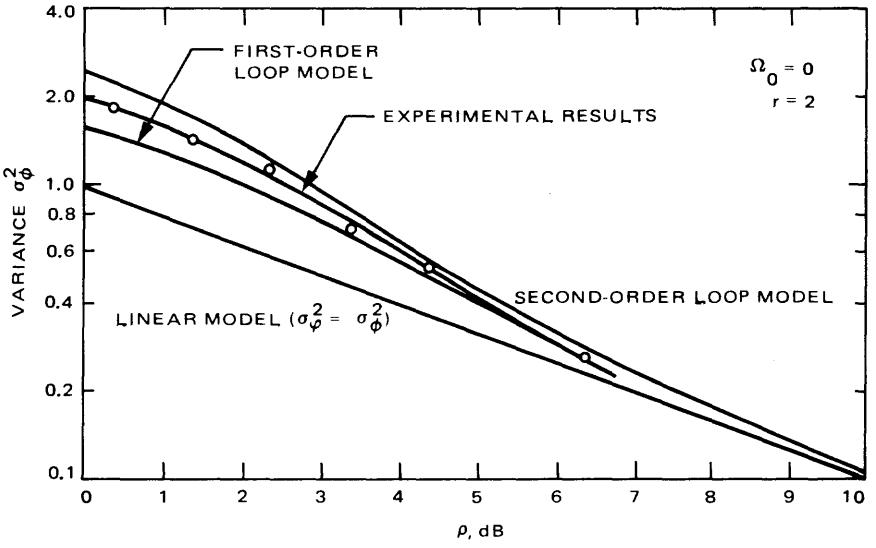


Fig. 3-5. A comparison of phase error variances based on various loop models

where

$$C = \frac{\int_0^{2\pi} \exp\,[U(x)]\,dx}{\int_{-2\pi}^{2\pi} \exp\,[U(x)]\,dx} \tag{3.2-37}$$

and

$$U(x) = -\left(\frac{r+1}{r}\right)\rho\cos x - \frac{\rho}{2r}x^2 - \rho\frac{\Omega_0}{\sqrt{P}K}x \tag{3.2-38}$$

In (3.2-36) and (3.2-37), $u(x)$ is the unit step function, r is given in (3.2-19), ρ is given in (3.2-31), and the loop bandwidth W_L is defined in (3.2-15).

Another way to examine cycle slipping is to assume that the phase error $\varphi(t)$ runs continuously and to compute the average number of cycle slips, both plus and minus, per second. A reasonable engineering approximation is:

$$\bar{S} \cong \frac{1}{\tau} \tag{3.2-39}$$

The parameter $\bar{S}$ may be used to compute the probability of loss of lock. Assume that the event $\{k$ slips in t sec$\}$ is Poisson; i.e.,

$$\Pr\{k \text{ cycle slips in } t \text{ sec}\} = \frac{(\bar{S}\,t)^k \exp(-\bar{S}\,t)}{k!} \tag{3.2-40}$$

This assumption is supported by experimental evidence [3-15 and 3-16].

Then the probability of losing phase lock in t seconds is

$$\begin{aligned} &\Pr\{\text{loss of phase lock in } t \text{ sec}\} \\ &= \Pr\{\text{one or more cycle slips in } t \text{ sec}\} \\ &= 1 - \exp(-\bar{S}\,t) \end{aligned} \tag{3.2-41}$$

Figure 3-6 shows a plot of $\bar{S}$ according to (3.2-39) and a simulation result of $\bar{S}$ as a function of SNR in the threshold loop bandwidth $2B_{L_o}$ (see section 3.2.4.2) for a typical second-order loop in deep space applications. Reference [3-17] gives cycle-slip simulation results for two PLLS in cascade. This is a model for the two-way system described in section 5.3.2.

For the case where $\Omega_0 \neq 0$, the average number of cycles slipped per second to the right, N_+, and to the left, N_-, is of interest. It has been shown [3-3] that

$$N_+ = \frac{J \exp(\gamma)}{2 \sinh \gamma} \tag{3.2-42}$$

$$N_- = \frac{J \exp(-\gamma)}{2 \sinh \gamma} \tag{3.2-43}$$

where

$$\gamma = \pi\rho \frac{\Omega_0}{\sqrt{P}\,K} \tag{3.2-44}$$

and J is the net average number of cycle slips per second,

$$J = N_+ - N_- \tag{3.2-45}$$

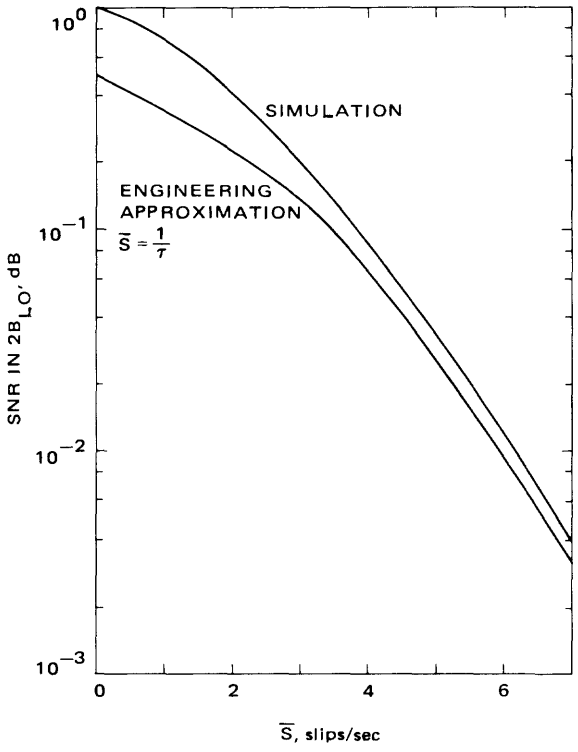


Fig. 3-6. Mean number of cycle slips per second vs SNR in $2B_{LO}$ ($2B_{LO}$ = 18 Hz)

The average number of cycle slips is then

$$\bar{S} = N_+ + N_- = \frac{J}{\tanh \gamma} \tag{3.2-46}$$

3.2.4 Phase-Locked Loop Preceded by Bandpass Limiter

The mechanization of a BPL is illustrated in Fig. 3-7; the limiter incorporated in the PLL system is shown in Fig. 3-8. BPLs are used in PLL receivers to maintain a constant total power at the input to the loop. This minimizes the total mean square error of the loop over a wide range of input signal-to-noise ratios. BPLs are used also to protect various loop components, the multiplier in particular, where signal and noise levels can vary over several orders of magnitude and exceed the dynamic range of these components. In this section, results needed to explain the behavior of a PLL when preceded by a BPL are summarized. The detailed theory and operation of the BPL are covered more completely in [3-3] and [3-18] through [3-20].

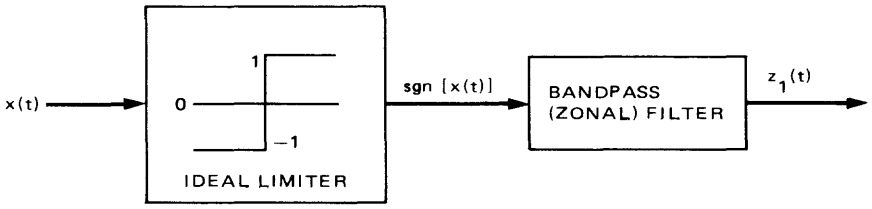


Fig. 3-7. A bandpass limiter model

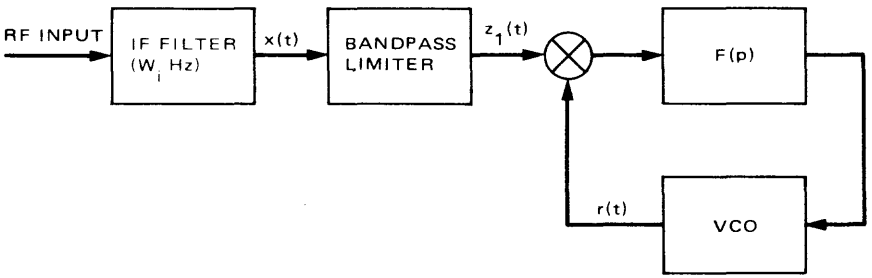


Fig. 3-8. A phase-locked loop preceded by a bandpass filter

3.2.4.1 **Bandpass limiting effects on phase-locked loop.** Assuming the output of the IF filter in Fig. 3-8 is given by (3.2-1), where the noise spectrum is flat over a bandwidth $W_i/2$ centered about the carrier frequency, then the first zone limiter output $z_1(t)$ can be expressed in the form,

$$z_1(t) = \alpha_1 \sqrt{2P_1} \sin [\omega_0 t + \theta(t)] + \sqrt{2}\{N_A(t) \cos [\omega_0 t + \theta(t)] - N_B(t) \sin [\omega_0 t + \theta(t)]\} \tag{3.2-47}$$

where $P_1 = 8/\pi^2$ represents the fraction of signal plus noise power that falls in the first zone, and the parameter α_1 is referred to as the signal amplitude suppression factor, and is given by [3-18],

$$\alpha_1 = \sqrt{\frac{\pi \rho_i}{4}} \exp\left(-\frac{\rho_i}{2}\right)\left[I_0\left(\frac{\rho_i}{2}\right) + I_1\left(\frac{\rho_i}{2}\right)\right] \tag{3.2-48}$$

where $\rho_i = 2P/N_0 W_i$, is the signal-to-noise ratio input to the limiter. Figure 3-9 illustrates the behavior of α_1 as a function of ρ_i. A rational function approximation to α_1, which is simpler to use in system design, is suggested in [3-5],

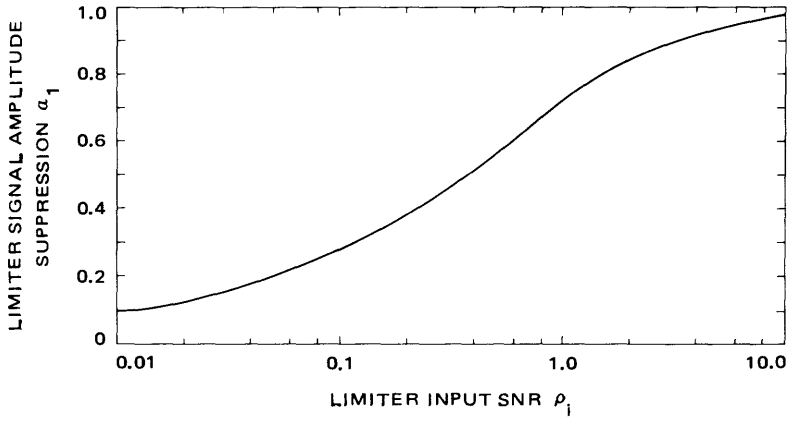


Fig. 3-9. Variation of limiter signal amplitude suppression vs limiter input SNR

$$\alpha_1 = \sqrt{\frac{0.785\,\rho_i + 0.4768\,\rho_i^2}{1 + 1.024\,\rho_i + 0.4768\,\rho_i^2}} \tag{3.2-49}$$

The noise processes $N_A(t)$ and $N_B(t)$ are zero mean and uncorrelated, but most important they are not, in general, Gaussian nor do they possess equal noise bandwidths or spectra [3-19].

The output of the phase detector is

$$\epsilon(t) = \alpha_1 \sqrt{P_1} \sin \varphi(t) + N_A(t) \cos \varphi(t) - N_B(t) \sin \varphi(t) \tag{3.2-50}$$

where double-frequency terms have been neglected. For steady-state operation in the linear region (small φ), we define an equivalent signal-to-noise ratio at the output of the phase detector by

$$\rho_0 = \frac{2P_e}{N_{0e} W_e} = \frac{\alpha_1^2 P_1}{\sigma_A^2} \tag{3.2-51}$$

where P_e, N_{0e} and W_e, are respectively the equivalent power, noise spectral density, and bandwidth at the phase detector output, and σ_A^2 denotes the variance of $N_A(t)$.

The parameter ρ_0 can be related to the input signal-to-noise ratio ρ_i by, [3-19]

$$\frac{\rho_0}{\rho_i} = \frac{2\alpha_1^2}{1 - \exp(-\rho_i)} = \frac{\pi\rho_i}{2[1 - \exp(-\rho_i)]} \exp(-\rho_i) \left[I_0\left(\frac{\rho_i}{2}\right) + I_1\left(\frac{\rho_i}{2}\right) \right]^2 \tag{3.2-52}$$

which is only a function of ρ_i.

We now define a factor Γ called the limiter performance factor by the ratio of the input signal-to-noise spectral density to that at the output:

$$\Gamma = \frac{P/N_0}{P_e/N_{0e}} = \frac{\rho_i W_i}{\rho_0 W_e} \tag{3.2-53}$$

Thus the reciprocal of Γ is the product of the input/output signal-to-noise ratio as given by (3.2-52) and the ratio of the equivalent noise bandwidth at the phase detector output to the IF bandwidth. For any value of ρ_i, it is clear that $W_e/W_i \geq 1/2$, since passing noise of known bandwidth through a zero-memory nonlinear device can only increase the equivalent noise bandwidth. The factor of 1/2 is due to the bandpass to low-pass bandwidth transformation. Hence, an upper bound on Γ is simply given by 2 times the reciprocal of (ρ_0/ρ_i). An exact analytical expression for Γ (or equivalently W_e/W_i) is quite difficult to develop. For our case of a rectangular noise spectrum, a good approximation to Γ is given by a rational function, which is simpler to use in system design, suggested in [3-5] (later corrected in [3-20]) and is given by

$$\Gamma = \frac{1 + \rho_i}{0.862 + \rho_i} \tag{3.2-54}$$

The function $1/\Gamma$ is plotted vs ρ_i in Fig. 3-10 along with its lower bound $\rho_0/2\rho_i$.

The variance of the phase error is an important parameter in specifying the loop response to a sine wave plus noise. Letting W_L also denote the loop bandwidth in the presence of the BPL (to be defined shortly), the phase error variance in the linear region is given by

$$\sigma_\varphi^2 = \frac{1}{\left(\frac{P_e}{N_{0e}}\right)\left(\frac{2}{W_L}\right)} = \frac{\Gamma}{\left(\frac{P}{N_0}\right)\left(\frac{2}{W_L}\right)} = \left(\frac{N_0 W_L}{2P}\right)\Gamma \tag{3.2-55}$$

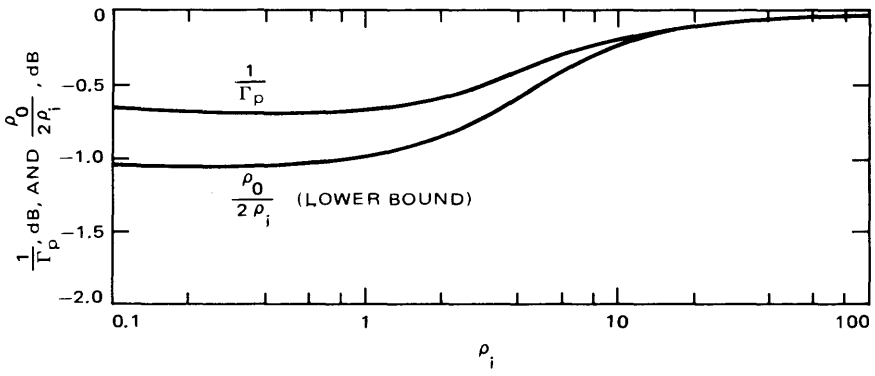


Fig. 3-10. Limiter performance factor as a function of input SNR

Since the effective signal power at the input to the PLL is now $\alpha_1^2 P_1$, the transfer function for a PLL preceded by a BPL is (from (3.2-14))

$$H(s) = \frac{1 + \tau_2 s}{1 + \left[\tau_2 + \frac{1}{\alpha_1 \sqrt{P_1}\, K}\right] s + \frac{\tau_1}{\alpha_1 \sqrt{P_1}\, K} s^2}$$

and the corresponding W_L becomes

$$W_L = \frac{1 + r}{2\tau_2 \left(1 + \frac{\tau_2}{r\tau_1}\right)} \tag{3.2-56}$$

where

$$r = \frac{\alpha_1 \sqrt{P_1}\, K\, \tau_2^2}{\tau_1} \tag{3.2-57}$$

Notice that W_L is now a function of the limiter suppression factor α_1, and hence the validity of the assumption $r\,\tau_1/\tau_2 >> 1$ depends on the value of the input signal-to-noise ratio ρ_i.

All PLL receivers used for deep space communications either on the spacecraft or on the ground are of the multiple-conversion heterodyne type preceded by a bandpass limiter, as shown in Figs. 3-1 through 3-3.

For multiple-conversion heterodyne PLL receivers,

$$r = \alpha_1 K_D K_{VCO} M K_{DC} 360 \tau_2^2/\tau_1$$

where

α_1 = limiter suppression factor

K_D = phase detector gain, V/deg

K_{VCO} = VCO gain, Hz/V

M = VCO multiplication factor

K_{DC} = gain of the loop filter.

3.2.4.2 Receiver threshold and design point. In practice, the nonlinear behavior of a PLL is of considerable interest in carrying out a particular design. It is useful in the design, specification, and performance analysis of PLL receivers to define a receiver threshold. By convention, the threshold condition occurs when the variance of the loop phase error is unity and the variance is obtained from a linearized PLL model. In actual fact, the variance of a linearized loop is different from that of an actual nonlinear loop when its variance is one. The fiction that the loop is linear in the definition of threshold is used to define the *threshold* condition

$$P_0 = N_0 W_{L0} \tag{3.2-58}$$

where the zero subscripts denote values at threshold.

Another definition used to specified PLL parameters at a *design point* has arbitrarily been taken to be the condition where

$$2 P_0 = N_0 W_{L0} \tag{3.2-59}$$

with

$$W_{L0} = \frac{1 + r_0}{2\tau_2 \left(1 + \dfrac{\tau_2}{r_0 \tau_1}\right)}$$

$$r_0 = \frac{\alpha_{10} \sqrt{P_1}\, K\, \tau_2^2}{\tau_1} \tag{3.2-60}$$

Again, the zero subscript on P, W_L, α_1 and r refers to their respective values at the design point. Let P_0 represent a signal power at which the linear PLL theory does not apply; hence, a suitable nonlinear model must be proposed from which one may predict system performance in the actual region of operation. When the loop bandwidth W_L is designed to be small relative to the equivalent noise bandwidth at the phase detector output W_e, it has been shown [3-6] that the effective signal-to-noise ratio in the loop bandwidth is given by

$$\rho = \left(\frac{2P}{N_0 W_L}\right)\frac{1}{\Gamma} = \frac{1}{\sigma_\varphi^2} \tag{3.2-61}$$

Keep in mind that both W_L and Γ are dependent on the input signal-to-noise ratio.

As we have already observed, the pdf of the phase error for a second-order PLL with zero detuning ($\Omega_0 = 0$) is characterized by the parameter α defined by (3.2-31) in terms of loop signal-to-noise ratio ρ and loop damping parameter r. It is sufficient, then, in applying this result and those derived from it to the case of a PLL preceded by a BPL, to use (3.2-61), together with (3.2-54) and (3.2-56) for ρ and (3.2-57) for r. We now consider the design of a PLL in terms of specifying system parameters at the design point.

If once again it is assumed that $r_0 \tau_1/\tau_2 >> 1$, then from (3.2-56) and (3.2-60) the actual operating condition loop bandwidth W_L can be related to the design point bandwidth W_{L0} by

$$W_L = \frac{1 + \dfrac{r_0}{\mu}}{1 + r_0} W_{L0} \tag{3.2-62}$$

where

$$\mu = \frac{\alpha_{10}}{\alpha_1} = \frac{r_0}{r} \tag{3.2-63}$$

and α_{10} is the value of α at the design point. Furthermore, the effective signal-to-noise ratio in the loop as defined in (3.2-61) can be written as

$$\rho = \left(\frac{2P}{N_0 W_{L0}}\right)\frac{1}{\Gamma}\left(\frac{1 + r_0}{1 + \dfrac{r_0}{\mu}}\right) \tag{3.2-64}$$

3.3 Suppressed Carrier Modulation and Tracking Loop Considerations

This section begins by examining the choice between suppressed and residual carrier systems [3-21]. The DSN emerged at a time when weak signals and low data rates dominated. As a result, residual carrier phase modulation schemes with phase-locked loops tracking the residual carrier component were employed. However, over the years, technological advances in antennas, transmitters, and signal processing have caused a marked increase in the available signal power and hence the achievable data rates. In fact, essentially all of the current or near-future deep space missions are already operating at or near the upper limit of the current DSN data rate capability.

When the DSN operated at 8 bps using residual carrier modulation, it was necessary to place the data modulation on subcarriers since direct modulation of the data on the carrier would cause most of the data power to fall within the bandwidth of, and to be tracked out by, the carrier phase-locked loop. The DSN is currently involved in new high rate telemetry system designs [3-22 through 3-24] which will push the data rate into the tens of megabits per second region. At these higher data rates the data signal spectrum is extremely broad so that even if subcarriers were not used, the part of the spectrum which would be contained in the loop bandwidth would be flat across this band and thus appear as white noise. Since the ratio of loop bandwidth to data rate is, in these situations, typically quite small, this additional white noise component degrades the tracking performance very little. This observation is further aided by the fact that we are already dealing with a greatly improved carrier loop SNR (relative to low data rate operation) by virtue of the fact that in order to support the higher data rates one must supply more signal (and hence carrier) power.

The above arguments suggest that subcarriers are not needed in high data rate systems. This leaves the question as to whether or not a residual carrier component should be provided for or is it preferable to fully suppress the carrier, thus allocating all of the transmitted power to the data modulation. Consider first a residual carrier signal of the form

$$y(t) = \sqrt{2P}\sin(\omega_0 t + \theta_m\, m(t)) + n_i(t)$$

$$= \sqrt{2P}\cos\theta_m \sin\omega_0 t + \sqrt{2P}\sin\theta_m m(t)\cos\omega_0 t + n_i(t) \qquad (3.3\text{-}1)$$

where P is the total signal power, ω_0 is the carrier radian frequency, θ_m is the modulation angle $(0 < \theta_m < \pi/2)$, $m(t)$ is a ± 1 binary data modulation with symbol rate $R_s = 1/T_s$, and as before $n_i(t)$ is a bandpass "white" Gaussian noise process. If the signal of (3.3-1) is demodulated by a reference signal

$$r(t) = \sqrt{2} \cos(\omega_0 t - \phi) \tag{3.3-2}$$

supplied by a phase-locked loop, then the loop phase detector output $\epsilon(t)$ is given by

$$\begin{aligned} \epsilon(t) &= \sqrt{P_c} \sin\phi + \sqrt{P_d}\, m(t) \cos\phi + n_1(t) \\ &\cong \sqrt{P_c}\, \phi + \sqrt{P_d}\, m(t) + n_1(t) \end{aligned} \tag{3.3-3}$$

where

$P_c = P \cos^2 \theta_m$ = carrier signal power

$P_d = P \sin^2 \theta_m$ = data modulation signal power

and

$$n_1(t) = \sqrt{2}\, n_i(t) \cos(\omega_0 t - \phi) \tag{3.3-4}$$

is an equivalent low-pass white Gaussian noise process with single-sided noise power spectral density N_0 W/Hz.

Assuming as above that the high rate data modulation (second term of (3.3-3)) has a flat spectrum over the phase-locked loop bandwidth with equivalent single-sided spectral density

$$N_d = \frac{P_d}{R_s/2} = 2\left(\frac{P_d T_s}{N_0}\right) N_0 = 2R_d N_0 \quad ; \quad R_d = \frac{P_d T_s}{N_0} \tag{3.3-5}$$

Then if B_L denotes the single-sided loop noise bandwidth, we have that the loop SNR ρ' is given by

$$\rho' = \frac{P_c}{(N_0 + N_d) B_L} = \left(\frac{P}{N_0 B_L}\right)\left(\frac{\cos^2 \theta_m}{1 + 2R_d}\right) \tag{3.3-6}$$

The first factor of (3.3-6) represents the loop SNR corresponding to the case where the total signal power is available for carrier tracking. The numerator of the second factor represents the degradation of this SNR due to the split of the total power between carrier and data modulation, while the denominator of this second factor represents the degradation produced by the modulation

self-noise in the loop. Clearly for large values of R_d (e.g., 9.6 dB for uncoded binary PSK at an error rate of 10^{-5}), the loop SNR is severely degraded by the need to track on a residual carrier component.

Aside from the increased noisy reference loss (additional signal power required to produce the same error probability performance as that corresponding to a perfect carrier sync reference) caused by the degraded loop SNR, there is an inherent loss in data power (relative to the total power available) just due to the split between carrier and data modulation as in (3.3-4).

Next we consider the situation in which the carrier is fully suppressed and the signal is of the form

$$y(t) = \sqrt{2P}\, m(t) \sin \omega_0 t + n_i(t) \tag{3.3-7}$$

Since $y(t)$ of (3.3-7) contains no residual carrier component, a PLL cannot be used to establish the desired carrier demodulation reference signal. Instead one must use a suppressed carrier type of tracking loop, e.g., a Costas loop (see Fig. 3-11). Since a detailed discussion of the behavior and performance of such a loop will be presented in the next section, we shall merely give the end result here for the purpose of comparison with (3.3-6), namely,

$$\rho'' = \left(\frac{P}{N_0 B_L}\right) S_L \tag{3.3-8}$$

where $S_L \leqslant 1$ is a loss factor (later on referred to as squaring loss) which depends on data SNR[2]

$$R_d = \frac{PT_s}{N_0}\,, \tag{3.3-9}$$

and the type of Costas loop arm filter, and the relation between the single-sided noise bandwidth B of these arm filters and the data rate R_s. As an example for ideal (rectangular frequency response) arm filters, S_L is given by the simple relation (also see (3.3-28))

[2]Note that in contrast with (3.3-5) since all of the total signal power P is available for the data modulation, the data SNR is defined accordingly.

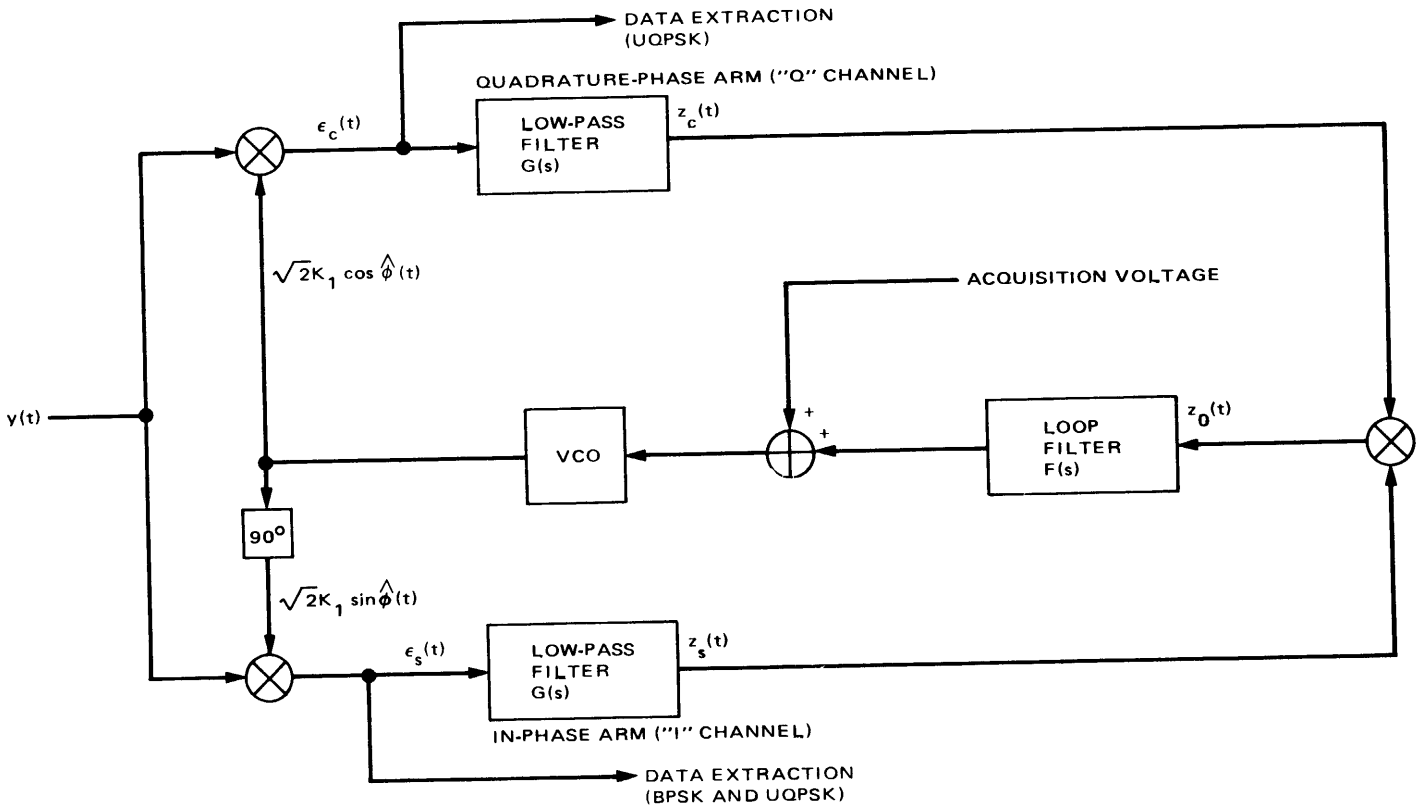


Fig. 3-11. Costas loop functional diagram with passive arm filters

$$S_L = \frac{K_2^2}{K_2 + \frac{BT_s}{R_d}} \tag{3.3-10}$$

where

$$K_2 = \frac{2}{\pi}\left[Si(2\pi BT_s) - \frac{\sin^2 \pi BT_s}{\pi BT_s}\right] \tag{3.3-11}$$

and $Si(x)$ is the *sine-integral function* defined by

$$Si(x) = \int_0^x \frac{\sin x}{x}\, dx \tag{3.3-12}$$

Table 3-1 presents an evaluation of the squaring loss S_L of (3.3-10) as a function of R_d for several values of the symbol arm filter time-bandwidth product BT_s [3-21]. Note that here the loss factor improves with increasing R_d. In fact for R_d on the order of 10 dB, a comparison of S_L with the loss factor for the residual carrier case (3.3-6) reveals that the suppressed carrier system enjoys about a 20 dB advantage in loop SNR (aside from the reduction in power due to the $\cos^2 \theta_m$ factor). Again we emphasize that the suppressed carrier system has the further advantage that all of the available power is simultaneously used for carrier tracking and data detection.

While the example we have pursued here strictly dealt with binary data modulation, i.e., BPSK, similar advantages of a suppressed carrier over a residual carrier format can be demonstrated for other forms of modulation, e.g., balanced quadriphase-shift-keying (QPSK) and unbalanced quadriphase-shift-keying (UQPSK). In the following sections, we present the various carrier synchronization options for these suppressed carrier modulation forms, beginning with a discussion of carrier recovery for BPSK modulation. Before proceeding to these detailed discussions, however, we wish to conclude this

Table 3-1. Evaluation of squaring loss S_L, in dB, for ideal arm filters

R_d, dB	$BT_s = 1$	$BT_s = 2$	$BT_s = 3$	$BT_s = 4$	$BT_s = 5$
0	−3.68	−5.14	−6.28	−7.19	−7.94
4	−2.03	−2.87	−3.65	−4.31	−4.90
10	−0.90	−1.05	−1.32	−1.60	−1.88

section by reemphasizing and stressing the key issues that distinguish these various carrier synchronization methods from the more traditional discrete carrier version, i.e., the PLL. It is hoped that the reader will understand the reasons for the amount of discussion that follows and its relative importance within the overall framework of this book.

It is clear from the word "suppressed" that a discrete spectral line does not appear in the input spectrum. Thus, the carrier synchronization loop must *reconstruct* the carrier from the data-bearing signal itself, which now contains all of the available signal power. The ability to accomplish this mission is thus dependent on the form of the signal modulation and hence a different synchronization structure must be discussed for each of these forms. (In some situations, it is possible to use a given structure for more than one modulation type.) Beyond this, the behavior of the synchronization loop and its performance depend on the data symbol format. Indeed, for a given modulation type, the design of a Costas-type receiver for optimum tracking performance requires that the bandwidth of its arm filters be selected in accordance with (1) the data symbol rate, (2) the arm filter type, e.g., 1- or 2-pole Butterworth, and (3) the data format (NRZ, Manchester, etc.). All of these degrees of design freedom require a multidimensional investigation over and above, but not separated from, the usual considerations of loop design associated with the loop filter (loop bandwidth, loop damping, etc.), the latter being the only design considerations required for a PLL. Indeed the PLL has no arm filters, and aside from the amount of power allocated to the residual carrier component, which might be a function of transmitted signal form, its design is in no way governed by the above modulation-dependent considerations.

3.3.1 Carrier Recovery for BPSK Modulation

As stated above, the form of carrier recovery loop is, in general, dependent on the form of modulation. For the three modulations mentioned above, namely, BPSK, QPSK and UQPSK, the carrier recovery loops are varietal versions of what is commonly referred to as a Costas loop [3-25]. Thus we begin this section with a description of the conventional Costas loop and its operational behavior for BPSK modulation.

3.3.1.1 The conventional Costas loop and its equation of operation. Consider the Costas loop illustrated in Fig. 3-11. An estimate of carrier phase is obtained by multiplying (using two phase detectors) the input suppressed carrier signal $s(t)$ plus noise $n_i(t)$ with the output of the VCO and a 90-deg phase shift of itself, respectively, filtering the results of these two multiplications, and using the product of the two filtered signals to control the phase and frequency of the loop's VCO output. When the filters in the in-phase and quadrature-phase arms of the Costas loop are mechanized with integrate-and-

dump circuits, then the loop will be referred to as a Costas loop with active arm filters.

The received signal $y(t)$ can be modeled as

$$y(t) = \sqrt{2P}\, m(t) \sin \Phi(t) + n_i(t) = s(t) + n_i(t) \tag{3.3-13}$$

where $\Phi(t) \triangleq \omega_i t + \theta(t)$, with ω_i the input radian carrier frequency and $\theta(t) = \Omega_0 t + \theta_0$ the input phase to be estimated, P is the carrier power, $m(t)$ is the data modulation (a ±1 digital waveform), and $n_i(t)$ is the additive channel noise which can be expressed in the form of a narrow-band process about the actual frequency of the input observed data; i.e.,

$$n_i(t) = \sqrt{2}\, [N_c(t) \cos \Phi(t) - N_s(t) \sin \Phi(t)] \tag{3.3-14}$$

where $N_c(t)$ and $N_s(t)$ are approximately statistically independent, stationary, white Gaussian noise processes with single-sided noise spectral density N_0 W/Hz [3-6] and single-sided bandwidth $B_H < \omega_0/2\pi$.

Using standard analysis techniques analogous to those applied to phase-locked loops [3-6], it may be shown [3-25] that the stochastic integro-differential equation of operation of Fig. 3-11 becomes

$$2\frac{d\varphi(t)}{dt} = 2\Omega_0 - KF(p)\, \{P \overline{\langle \hat{m}^2(t) \rangle} \sin 2\varphi(t) + \nu_2 [t, 2\varphi(t)]\} \tag{3.3-15}$$

where $\varphi(t) \triangleq \Phi(t) - \hat{\Phi}(t)$ is the loop phase error ($\hat{\Phi}(t)$ is the VCO's reference phase), and K is the total open loop gain,

$$\begin{aligned} \nu_2 [t, 2\varphi(t)] \triangleq\ & [-\hat{N}_c^2(t) + \hat{N}_s^2(t) - 2\sqrt{P}\hat{m}(t)\hat{N}_s(t)] \sin 2\varphi(t) \\ & + [2\sqrt{P}\hat{m}(t)\hat{N}_c(t) - 2\hat{N}_c(t)\hat{N}_s(t)] \cos 2\varphi(t) \end{aligned} \tag{3.3-16}$$

and

$$\overline{\langle \hat{m}^2(t) \rangle} \triangleq K_2 = \int_{-\infty}^{\infty} S_m(f) |G(j2\pi f)|^2\, df \tag{3.3-17}$$

with $S_m(f)$ denoting the power spectral density of the data modulation $m(t)$ and $|G(j2\pi f)|^2$ the squared magnitude of the arm filter transfer function. Here the signal $\hat{m}(t)$ denotes the signal emerging from passage of the data modulation $m(t)$ through the low-pass arm filter $G(p)$, and likewise $\hat{N}_c(t)$ and $\hat{N}_s(t)$

are equivalently filtered versions of $N_c(t)$ and $N_s(t)$, respectively. Also, in (3.3-15) and (3.3-17) the overbar denotes statistical expectation and $\langle\ \rangle$ denotes time averaging, which is required since $m(t)$ is a cyclostationary process. In arriving at (3.3-15), we have also ignored the effect of the modulation self noise

$$n_\Delta\ [t,\ 2\ \varphi(t)] \triangleq P[\hat{m}^2(t) - \overline{\langle \hat{m}^2(t) \rangle}]\ \sin\ 2\ \varphi(t) \tag{3.3-18}$$

since for data SNRs on the order of 10 dB or less this omission has negligible effect on the loop performance.

3.3.1.2 Equivalent noise model. Since the bandwidth of the Costas loop is ordinarily designed to be narrow with respect to the equivalent noise bandwidth of $v_2(t, 2\varphi)$, we can further approximate $v_2(t, 2\varphi)$ as a delta correlated process with equivalent single-sided noise spectral density

$$N_{sq} \triangleq 2\int_{-\infty}^{\infty} R_{v_2}(\tau)d\tau \tag{3.3-19}$$

where

$$R_{v_2}(\tau) = \overline{v_2(t,\ 2\varphi)\ v_2(t+\tau,\ 2\varphi)} \tag{3.3-20}$$

Evaluation of N_{sq} results in [3-25]

$$N_{sq} = 4PN_0\left(K_2\ K_4 + \frac{N_0\ B}{P}\ K_L\right) \tag{3.3-21}$$

where K_L is a constant dependent only on the filter type and is defined by

$$K_L \triangleq \frac{\displaystyle\int_{-\infty}^{\infty} |G(j2\pi f)|^4\ df}{\displaystyle\int_{-\infty}^{\infty} |G(j2\pi f)|^2\ df} \tag{3.3-22}$$

K_4 is a constant dependent on both the baseband data power spectrum and the filter type, i.e.,

$$K_4 \triangleq \frac{\int_{-\infty}^{\infty} S_m(f)\,|G(j2\pi f)|^4\,df}{\int_{-\infty}^{\infty} S_m(f)\,|G(j2\pi f)|^2\,df} \tag{3.3-23}$$

and B denotes the single-sided noise bandwidth of the low-pass arm filter $G(j2\pi f)$, i.e.,

$$B \triangleq \int_0^{\infty} |G(j2\pi f)|^2\,df \tag{3.3-24}$$

3.3.1.3 Equivalent linear loop performance. Although the nonlinear differential equation of (3.3-15) can be solved in the steady state using Fokker-Planck techniques [3-6], typically the signal-to-noise ratio in the loop is sufficiently high that one can linearize the equation of operation and still obtain meaningful performance results. By linearize, we mean that sin 2φ is replaced by 2φ in (3.3-15); i.e.,

$$2\,\frac{d\varphi}{dt} = 2\,\Omega_0 - KF(p)\,\{P\,K_2\,2\varphi + v_2(t, 2\varphi)\} \tag{3.3-25}$$

Thus, under the above linear assumption, the solution for the modulo 2π reduced phase error 2ϕ can be shown to have a steady-state probability density function $p(2\phi)$ which is a Gaussian distribution with variance

$$\sigma_{2\phi}^2 = \frac{N_{sq}\,B_L}{P^2 D_m^2} \triangleq \frac{4}{\rho S_L} \tag{3.3-26a}$$

or

$$\sigma_{\phi}^2 = \frac{1}{4}\,\sigma_{2\phi}^2 = \frac{1}{\rho S_L} \tag{3.3-26b}$$

where again

$$\rho = \frac{P}{N_0 B_L} \tag{3.3-27}$$

is the signal-to-noise ratio in the loop bandwidth of an equivalent linear loop, B_L is the single-sided loop noise bandwidth, and the squaring loss S_L is the loop

signal-to-noise ratio penalty relative to a linear loop due to signal X signal, signal X noise, and noise X noise distortions.

Using (3.3-21) and (3.3-27) in (3-3-26b), we see that the loop squaring loss is given by the simple relation

$$S_L = \frac{K_2^2}{K_2 K_4 + K_L \frac{\zeta}{R_d}} \tag{3.3-28}$$

where $R_s = 1/T_s$ is the data symbol rate,

$$R_d = \frac{PT_s}{N_0} \tag{3.3-29a}$$

is the detection signal-to-noise ratio, and the parameter

$$\zeta = \frac{B}{R_s} \tag{3.3-29b}$$

3.3.1.4 Squaring loss evaluation. Expressing the data modulation $m(t)$ as a random pulse train, i.e.,

$$m(t) = \sum_{n=-\infty}^{\infty} a_n p(t - nT_s); E\{a_n a_m\} = \delta_{mn} \tag{3.3-30}$$

where $\{a_n\}$ is the data sequence taking on values ±1 with equal probability and $p(t)$ is the data pulse shape which for NRZ is

$$p(t) = \begin{cases} 1; & 0 \leq t \leq T_s \\ 0; & \text{otherwise} \end{cases} \tag{3.3-31}$$

while for Manchester coding

$$p(t) = \begin{cases} 1; & 0 \leq t \leq \frac{T_s}{2} \\ -1; & \frac{T_s}{2} \leq t \leq T_s \\ 0; & \text{otherwise,} \end{cases} \tag{3.3-32}$$

then the corresponding power spectral densities are

$$S_m(f) = T_s \frac{\sin^2(\pi f T_s)}{(\pi f T_s)^2} \quad \text{(NRZ)} \tag{3.3-33a}$$

and

$$S_m(f) = T_s \frac{\sin^4\left(\frac{\pi f T_s}{2}\right)}{\left(\frac{\pi f T_s}{2}\right)^2} \quad \text{(Manchester code)} \tag{3.3-33b}$$

Furthermore, assume that the arm filter is an n-pole Butterworth characterized by the transfer function

$$|G(j2\pi f)|^2 = \frac{1}{1 + \left(\frac{f}{f_c}\right)^{2n}} \tag{3.3-34}$$

where f_c, the 3-dB bandwidth, is related to the single-sided noise bandwidth B of the filter by

$$f_c = \left(\frac{2n}{\pi} \sin \frac{\pi}{2n}\right) B \tag{3.3-35}$$

As an example, consider the case of Manchester coding. Then, substituting (3.3-33b) and (3.3-34) into (3.3-17) and performing the required integration gives the modulation distortion factor K_2. Using K_2 in (3.3-28) and the fact that $K_L = (2n - 1)/2n$, Fig.3-12 illustrates the squaring loss S_L vs $\zeta = B/R_s$ with R_d as a parameter for single-pole Butterworth filters. We observe that for a fixed R_d there exists an optimum noise bandwidth B for the arm filters in the sense of minimizing the squaring loss. These values of optimum arm filter bandwidth occur in the vicinity of the Nyquist bandwidth and their sensitivity with signal-to-noise ratio diminishes as the number of filter poles increases.

Table 3-2 tabulates the closed form expressions for K_2 of (3.3-17) and K_4 of (3.3-23) corresponding to single-pole and two-pole Butterworth arm filters [$n = 1$ and $n = 2$ of (3.3-34)] and NRZ or Manchester data, thus allowing computation of squaring loss from (3.3-28). Closed form results for higher-order Butterworth arm filters ($n > 1$) or indeed any arm filter with rational transfer function can be obtained, if desired, from the results in [3-26]. In

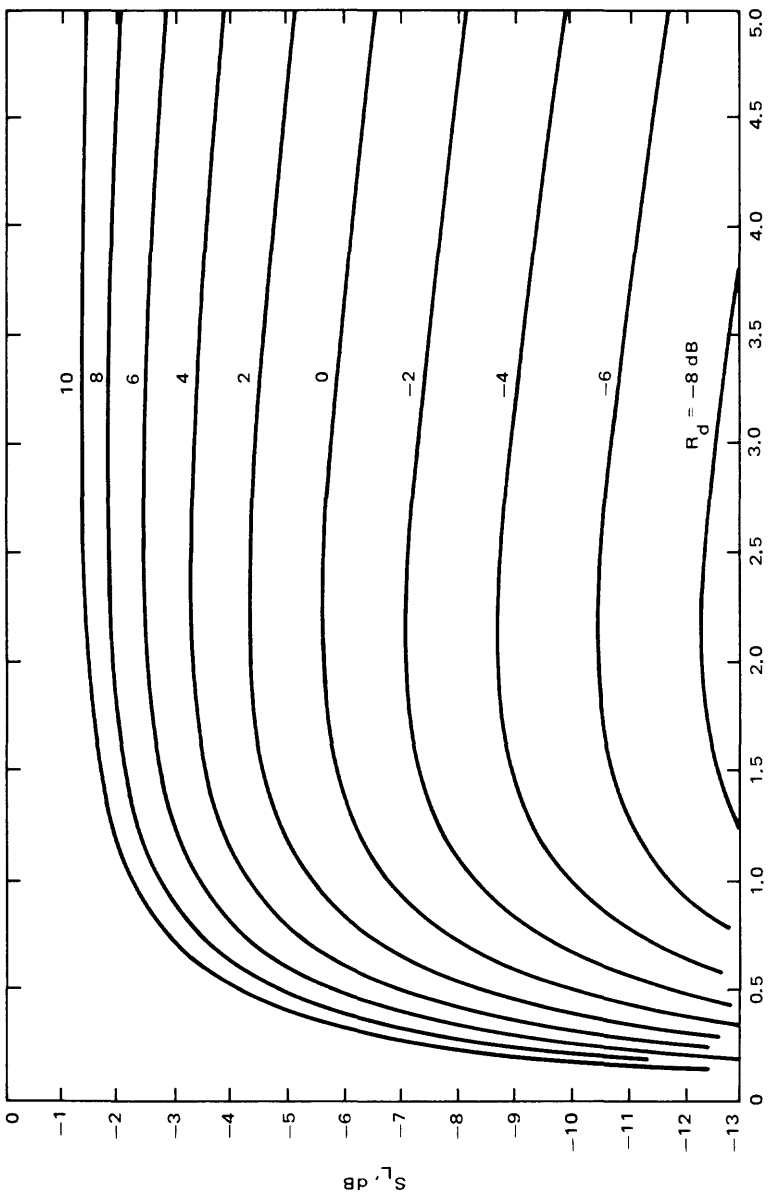


Fig. 3-12. Squaring loss variations vs B/R_s for various values of R_d; RC filter, Manchester coding

Table 3-2. Closed form expressions for the coefficients K_2 and K_4

1. NRZ data, single-pole Butterworth arm filters

$$K_2 = 1 - (1/\xi)\,[1 - \exp(-4\xi)]$$

$$K_2K_4 = 1 - (1/4\xi)\,[3 - (3 + 4\xi)\exp(-4\xi)]$$

2. NRZ data, two-pole Butterworth arm filters

$$K_2 = 1 - (1/8\xi)\,[1 - \exp(-4\xi)(\cos 4\xi - \sin 4\xi)]$$

$$K_2K_4 = 1 - (1/32\xi)\,\{5 - [8\xi\cos 4\xi + 5(\cos 4\xi - \sin 4\xi)]\exp(-4\xi)\}$$

3. Manchester data, single-pole Butterworth arm filters

$$K_2 = 1 - (1/4\xi)\,[3 - 4\exp(-2\xi) + \exp(-4\xi)]$$

$$K_2K_4 = 1 - (1/8\xi)\,[9 - 4(3 + 2\xi)\exp(-2\xi) + (3 + 4\xi)\exp(-4\xi)]$$

4. Manchester data, two-pole Butterworth arm filters

$$K_2 = 1 - (1/8\xi)\,\{3 - 4\exp(-2\xi)[\cos 2\xi - \sin 2\xi] + \exp(-4\xi)[\cos 4\xi - \sin 4\xi]\}$$

$$K_2K_4 = 1 - (1/32\xi)\,\{15 - [16\xi\cos 2\xi + 20(\cos 2\xi - \sin 2\xi)]\exp(-2\xi),$$

$$- (1/32\xi)\,[8\xi\cos 4\xi + 5(\cos 4\xi - \sin 4\xi)]\exp(-4\xi)\}$$

Note: ξ is defined in (3.2-29b).

particular, for $n = \infty$ (an ideal "brick wall" filter), we would have $K_4 = 1$ and for NRZ data, K_2 as previously given by (3.3-11).

In many applications, once symbol sync has been obtained, the passive arm filters of the Costas loop in Fig. 3-11 are replaced by active arm filters of the integrate-and-dump type. When the transmitted modulation is Manchester coding, it may be shown [3-25] that the variance of the phase error in the loop is still given by (3.3-26) where, however, the squaring loss is given by

$$S_L = \frac{1}{1 + \dfrac{1}{2R_d}} \tag{3.3-36}$$

This result is also valid for NRZ data.

Figure 3-13 compares the squaring losses which give minimum phase jitter achievable with passive arm filters to that obtained from the implementation with integrate-and-dump arm filters. Typically, the active matched filter

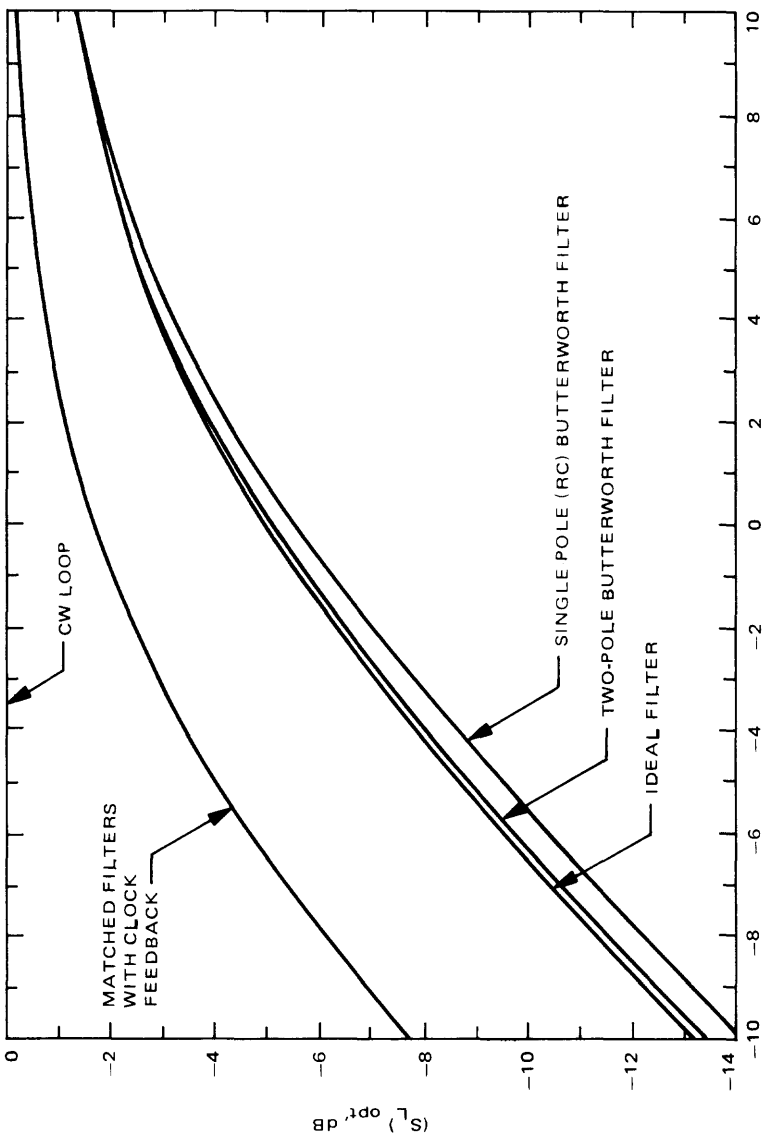


Fig. 3-13. Comparison of the effectiveness of various arm filters in reducing the squaring loss S_L for various values of R_d

(integrate-and-dump, sample-and-hold) gives approximately a 4 dB or greater advantage over the passive single-pole circuit for $R_d \leqslant 0$ dB. For example, at $R_d = -4$ dB, the advantage is 5.16 dB. When compared to an ideal filter, the improvement is typically only 4.4 dB at $R_d = -4$ dB. Notice also that for $R_d \geqslant 6$ dB, the Costas loop with active matched filters in the arms gives approximately the same performance (0.5 dB inferior) to that of a Costas loop with no modulation present (a CW loop).

3.3.1.5 The polarity-type Costas loop and its equation of operation. It is becoming increasingly popular (e.g., the carrier tracking loop in the Multi-megabit Telemetry Demodulator and Detector (MTDD) of the DSN) in the design of suppressed carrier receivers, which employ Costas loops for carrier reconstruction, to hard limit the output of the in-phase channel. Doing so allows replacement of the analog multiplier, which forms the loop error signal, with a chopper-type device which typically exhibits much less dc offset. Generally speaking, introduction of a limiter (hard or soft) into a system results in signal suppression the amount of which is a function of the signal-to-noise ratio at the limiter input. This signal suppression in turn reduces the total loop gain and, as a consequence, the loop bandwidth.

In this section, we address the tracking performance of such a hard-limited Costas loop (often referred to as a "polarity-type" Costas loop) and compare its performance with that of the conventional Costas loop previously discussed. Typically, for small detection signal-to-noise ratios, one will find that the inclusion of the hard limiter results in a tracking performance penalty, whereas for large detection signal-to-noise ratios, such inclusion indeed results in a performance improvement.

Realizing that the input signal and noise processes can still be modelled as in (3.3-13) and (3.3-14), respectively, then analogously to (3.3-15), the loop equation of operation for the polarity-type loop is given by [3-27]

$$2\frac{d\varphi}{dt} = 2\Omega_0 - KF(p)\left[\sqrt{P}\,\tilde{\alpha} f_1(2\varphi) + 2\hat{N}_e(t, 2\varphi)\right] \tag{3.3-37}$$

where

$$\hat{N}_e(t, 2\varphi) \triangleq \tilde{m}(t)\,\hat{N}(t, \varphi) \tag{3.3-38}$$

is the equivalent additive noise perturbing the loop. As before, the "hat" denotes filtering by the arm filters $G(s)$. Furthermore,

$$\tilde{m}(t) \triangleq \operatorname{sgn}\{\sqrt{P}\,\hat{m}(t)\cos\varphi(t) - [\hat{N}_s(t)\cos\varphi(t) + \hat{N}_c(t)\sin\varphi(t)]\} \tag{3.3-39}$$

where sgn $(x) = x/|x|$ denotes the signum function of the argument x,

$$\hat{N}[t, \varphi(t)] \triangleq \hat{N}_c(t) \cos \varphi(t) - \hat{N}_s(t) \sin \varphi(t) \tag{3.3-40}$$

and

$$\tilde{\alpha} f_1(2\varphi) = 2 \overline{< \hat{m}(t)\tilde{m}(t) >} \sin \varphi \tag{3.3-41}$$

where $f_1(x)$ is a nonlinearity which is periodic in x with period 2π and has *unit* slope at the origin; i.e., $x = 0$. Thus, $\tilde{\alpha}$ defined by

$$\tilde{\alpha} \triangleq \overline{< \hat{m}(t)\tilde{m}(t) >}\Big|_{\varphi=0} \tag{3.3-42}$$

is the *signal amplitude suppression factor* which results owing to the combined distortion effects on the input modulation $m(t)$ of the hard limiter and finite arm filter bandwidth. Finally, as was done for the conventional loop, in arriving at (3.3-37), we have ignored the modulation self-noise

$$n_\Delta[t, 2\varphi] \triangleq 2\sqrt{P}\,[\hat{m}(t)\tilde{m}(t) - \overline{<\hat{m}(t)\tilde{m}(t)>}] \sin \varphi \tag{3.3-43}$$

3.3.1.6 Calculation of the signal amplitude suppression factor. Expressing the input modulation $m(t)$ as in (3.3-30), then for a single-pole Butterworth arm filter and NRZ data, it can be shown that (3.3-46) evaluates to [3-27]

$$\tilde{\alpha} = \frac{1}{2}\int_0^1 \left[1 - 2\exp(-4\zeta x)\right] \operatorname{erf}\left\{\sqrt{\frac{R_d}{2\zeta}\left[1 - 2\exp(-4\zeta x)\right]}\right\} dx + \frac{1}{2}\operatorname{erf}\left(\sqrt{\frac{R_d}{2\zeta}}\right) \tag{3.3-44}$$

and is illustrated in Fig. 3-14. The curves are drawn for $\zeta = B/R_s \geqslant 0.75$ or equivalently $\tau_c/T \leqslant 1/3$, where $\tau_c \triangleq 1/\omega_c$ is the arm filter time constant. The validity of (3.3-44) in this region is somewhat justified by results obtained for a similar investigation [3-28].

3.3.1.7 Calculation of the equivalent noise spectral density. As was done for the conventional Costas loop in section 3.3.1.2, we can make the assumption that the bandwidth of the Costas loop is designed to be narrow with respect to the equivalent noise bandwidth of $\hat{N}_e(t, 2\varphi)$ defined in (3.3-38). Thus, we can further approximate $\hat{N}_e(t, 2\varphi)$ as a delta correlated process with equivalent single-sided noise spectral density

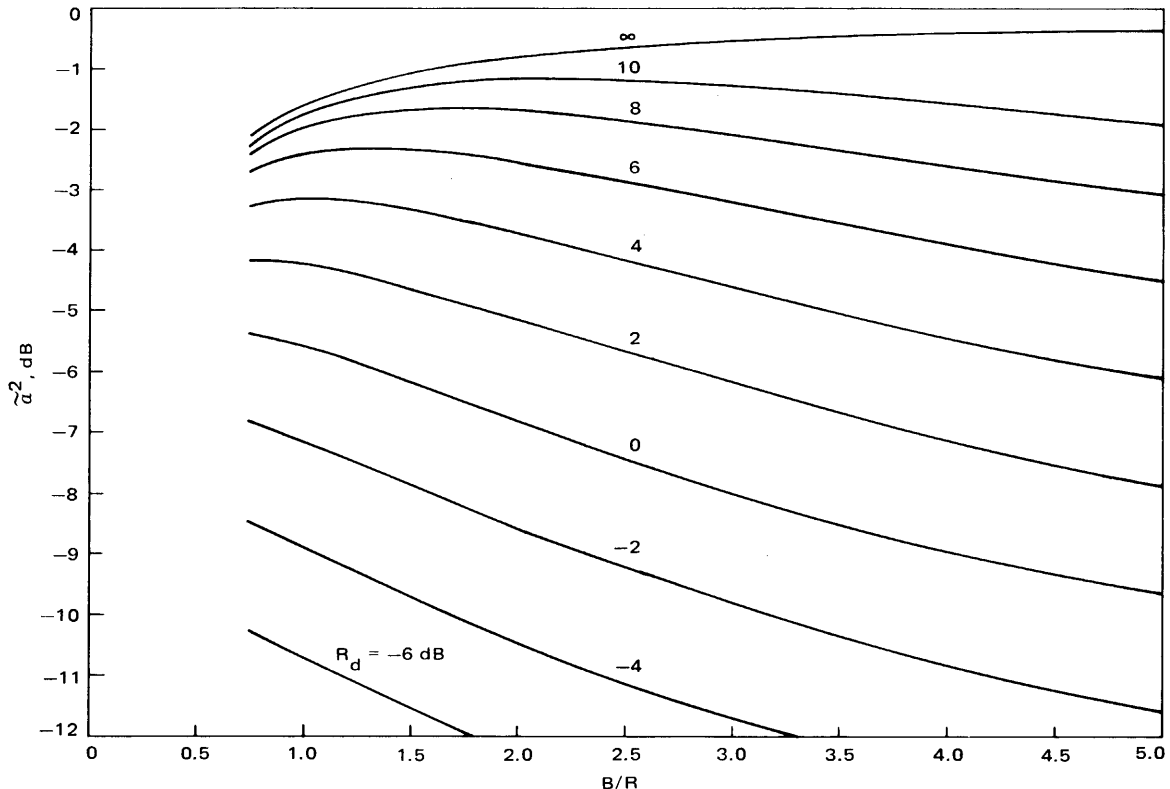


Fig. 3-14. Signal power suppression factor vs the ratio of two-sided arm filter bandwidth to data rate with data signal-to-noise ratio as a parameter

$$N_e \triangleq 2\int_{-\infty}^{\infty} R_{\hat{N}_e}(\tau)\, d\tau \tag{3.3-45}$$

where

$$R_{\hat{N}_e}(\tau) \triangleq \overline{<\hat{N}_e(t, 2\varphi)\, \hat{N}_e(t+\tau, 2\varphi)>} \tag{3.3-46}$$

Since, as previously mentioned, for good tracking performance the loop will operate in the neighborhood of $\phi = 0$, we need to evaluate (3.3-46) and hence (3.3-45) only at the point $\phi = 0$. Without elaborating the details (see [3-27]), the final results for NRZ data and single-pole arm filters are as follows.[3] Letting β_N denote the equivalent noise spectral density N_e of (3.3-45) normalized by the spectral density N_0 of the input noise, i.e.,

$$N_e \triangleq N_0 \beta_N \tag{3.3-47}$$

and defining

$$\rho_i = \frac{R_d}{\zeta} = \frac{P}{N_0 B} \tag{3.3-48}$$

with ζ given by (3.3-29b), then, in the case of small ρ_i:

$$\beta_N = \left(1 - \frac{2}{\pi}\right)(1 + \rho_i) - \frac{2\rho_i}{\pi}\left(\frac{1}{4\zeta}\right)\left[\frac{\pi}{2}\ln 2 - e^{-4\zeta}\left(\frac{3}{4} + \frac{1}{2}\ln 2 + 2\zeta\right)\right] \tag{3.3-49}$$

and in the case of large ρ_i:

$$\beta_N = 4\zeta \int_0^1 e^{-4\zeta y}\left[f_1(y) + e^{-4\zeta} f_3(y)\right] dy$$

$$+\frac{8\zeta}{\pi}\int_0^1 e^{-8\zeta y}\left[f_2(y) + e^{-8\zeta} f_4(y)\right] dy \tag{3.3-50}$$

[3] Closed form results can only be obtained in limiting cases of small and large values of signal-to-noise ratio. See [3-28] for further details.

where

$$f_1(y) = \left(\frac{2-y}{4}\right)\operatorname{erf}^2\left(\sqrt{\frac{\rho_i}{2}}\right) - \int_0^y g_1(x)\,dx + \int_{1-y}^1 g_1(x)\,dx$$

$$+\frac{1}{2}\int_0^{1-y} g_1(x)g_1(x+y)\,dx - \frac{1}{4}\int_0^y g_1(x)g_1(x+1-y)\,dx$$

$$f_2(y) = \left(\frac{2-y}{4}\right)\exp^2\left(-\frac{\rho_i}{2}\right) + \int_0^y g_2(x)\,dx + \int_{1-y}^1 g_2(x)\,dx$$

$$+\frac{1}{2}\int_0^{1-y} g_2(x)g_2(x+y)\,dx + \frac{1}{4}\int_0^y g_2(x)g_2(x+1-y)\,dx$$

$$f_3(y) = \left(\frac{1-y}{4}\right)\operatorname{erf}^2\left(\sqrt{\frac{\rho_i}{2}}\right) + \int_0^y g_1(x)\,dx - \int_{1-y}^1 g_1(x)\,dx$$

$$-\frac{1}{4}\int_0^y g_1(x)g_1(x+y)\,dx$$

$$f_4(y) = \frac{1}{4}\exp^2\left(-\frac{\rho_i}{2}\right) + \frac{1}{2}\int_0^1 g_2(x)\,dx + \frac{1}{4}\int_0^y g_2(x)g_2(x+1-y)\,dx$$

$$+\frac{1}{4}\int_0^{1-y} g_2(x)g_2(x+y)\,dx \qquad (3.3\text{-}51)$$

with

$$g_1(x) = \operatorname{erf}\left[\sqrt{\frac{\rho_i}{2}}(1 - 2e^{-4\zeta x})\right]$$

$$g_2(x) = \exp\left[-\frac{\rho_i}{2}(1 - 2e^{-4\zeta x})^2\right] \qquad (3.3\text{-}52)$$

3.3.1.8 Calculation of squaring loss and mean-square tracking jitter. As was done for the conventional Costas loop in arriving at (3.3-25), we linearize the

polarity-type Costas loop by replacing $f_1(2\varphi)$ of (3.3-41) by 2φ in the loop equation of operation, namely, (3.3-37). When this is done, the mean-square error of the modulo-2π reduced phase error process 2ϕ is given by a result analogous to (3.3-26), namely,

$$\sigma_{2\phi}^2 = \frac{4N_e B_L}{P\tilde{\alpha}^2} = 4\left(\frac{N_0 B_L}{P}\right)\left(\frac{\beta_N}{\tilde{\alpha}^2}\right)$$

$$\triangleq \frac{4}{\rho S_L} \tag{3.3-53}$$

where ρ is still given by (3.3-27). Thus the squaring loss of the polarity-type Costas loop becomes

$$S_L = \frac{\tilde{\alpha}^2}{\beta_N} \tag{3.3-54}$$

with $\tilde{\alpha}$ given by (3.3-44) and β_N given by (3.3-49) for the case of small ρ_i or (3.3-50) for the large ρ_i case. Additionally, using (3.3-48) in these defining equations for $\tilde{\alpha}$, β_N, and S_L, Figs. 3-15 and 3-16 illustrate S_L vs B_i/R_s corresponding to the calculation of β_N from (3.3-49) and (3.3-50), respectively. The curves in Fig. 3-15 were terminated whenever ρ_i exceeded 0.5, whereas the curves in Fig. 3-16 have a similar termination for $\rho_i < 2$. Superimposed on these curves (in dotted lines) are the corresponding results obtained from the previous section for the conventional Costas loop without the hard limiter. We observe from these results that, whereas at low signal-to-noise ratios the insertion of a hard limiter into the in-phase channel of a Costas loop results in a tracking performance degradation, at high signal-to-noise ratios, the hard-limited loop actually outperforms the conventional loop. These results are not surprising, at least from a qualitative standpoint, if one considers the Costas loop motivated by the maximum *a posteriori* (MAP) approach and approximates the integrate-and-dump filters which result from this approach by RC filters. In fact, using a similar technique, one can easily show that for the hard-limited Costas loop with integrate-and-dump filter in both arms, the squaring loss is given by

$$S_L = \text{erf}^2 \sqrt{R_d} \tag{3.3-55}$$

Figure 3-17 illustrates S_L of (3.3-55) vs R_d in dB and the corresponding result (3.3-36) in dotted lines for the conventional Costas loop with integrate-and-dump filters. The crossover in squaring loss or equivalently in tracking performance occurs at $R_d = -1.5$ dB.

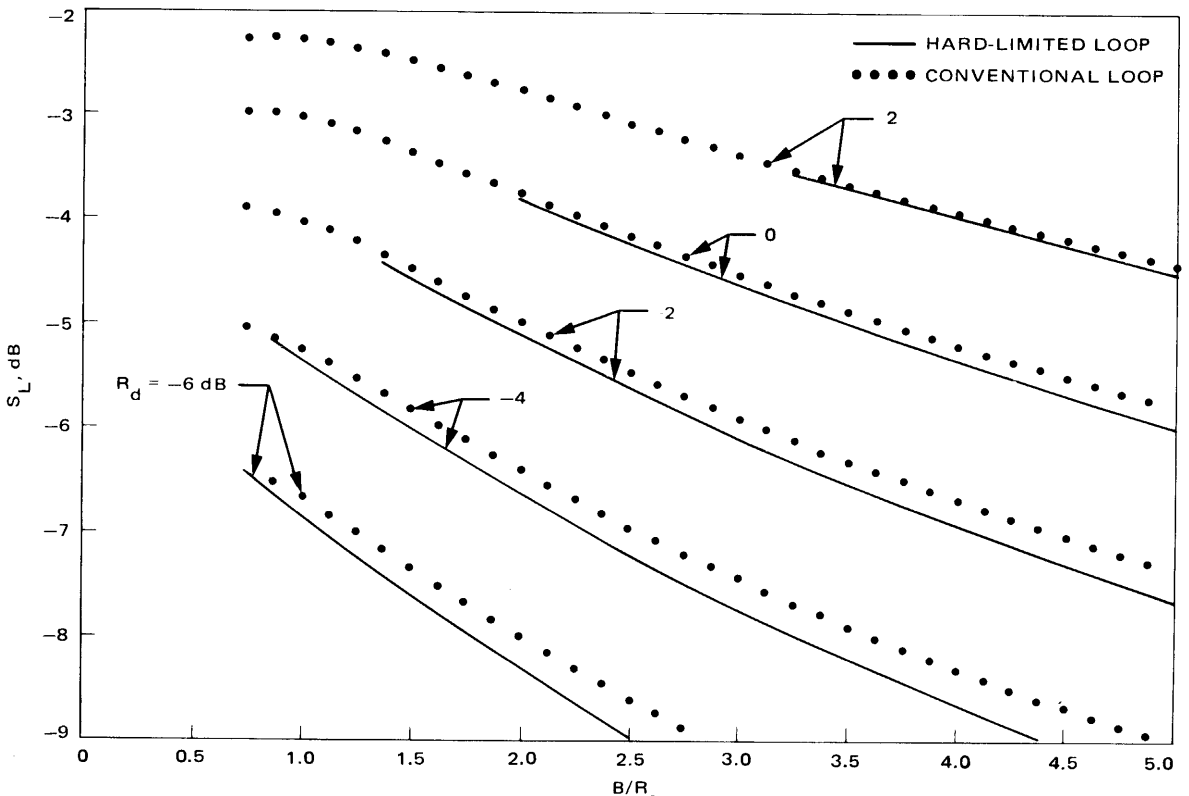


Fig. 3-15. Squaring loss variations vs B/R_s with R_d as a parameter; small ρ_I approximation; RC filter, NRZ data

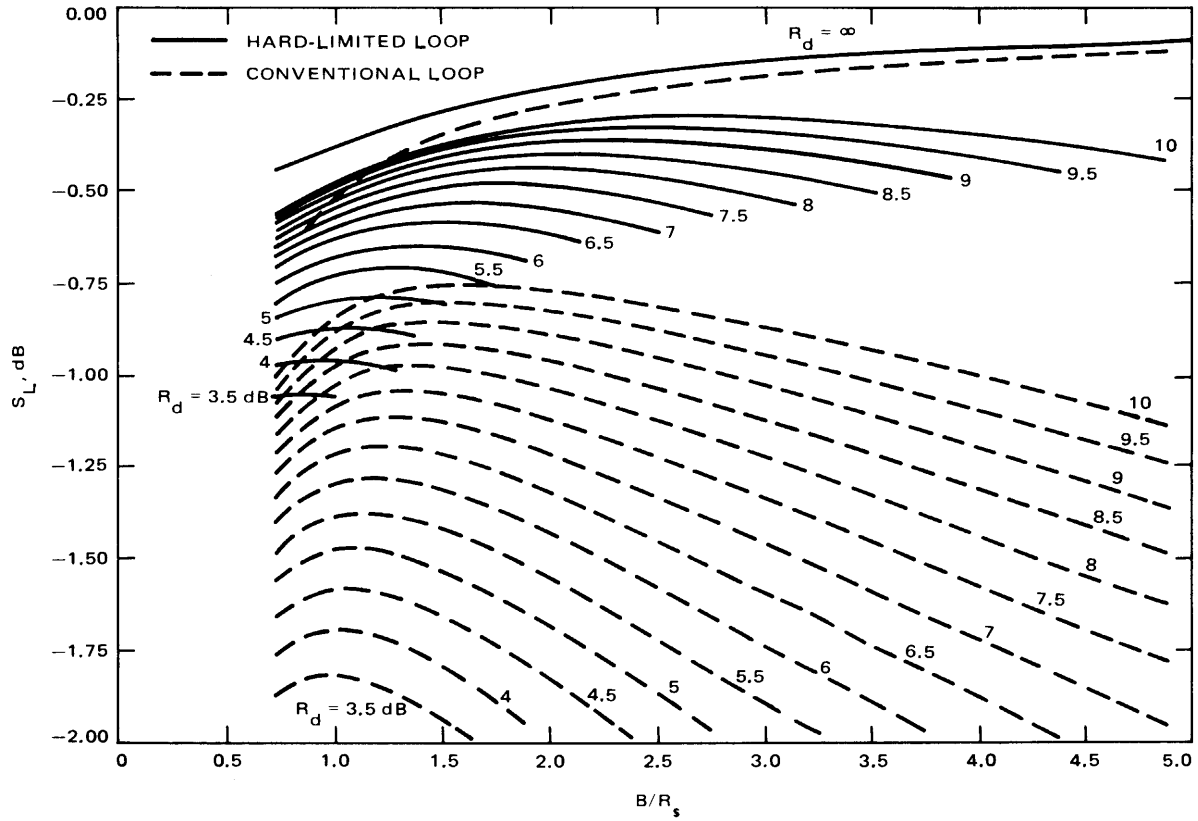


Fig. 3-16. Squaring loss variations vs B/R_s with R_d as a parameter; large ρ_I approximation; RC filter, NRZ data

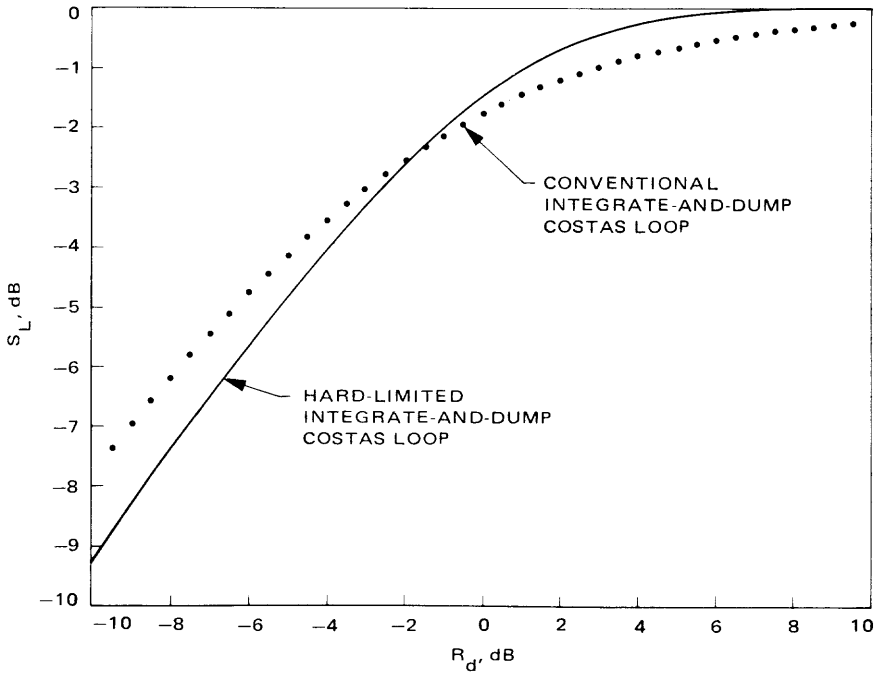


Fig. 3-17. A comparison of the squaring loss performances of conventional and hard-limited Costas loops with integrate-and-dump arm filters

3.3.2 Carrier Recovery for QPSK Modulation

A generalized version of the Costas loop of Fig. 3-11, which applies to BPSK, QPSK, and 8-PSK modulations, is illustrated in Fig. 3-18. Unfortunately, a complete analysis of this loop is not presently available. Thus we resort to considering the counterpart of this loop wherein the passive low-pass arm filters are again replaced by active arm filters of the integrate-and-dump type. This loop is generally referred to as a quadriphase polarity-type I-Q loop or quadriphase I-Q loop, depending respectively upon whether or not hard limiters are included in the in-phase arms. The former is illustrated in Fig. 3-19. Its analysis follows along the same lines as that of the equivalent biphase loop. Thus, we shall omit the details and merely present the final performance results.

The mean-square phase jitter tracking performance of the loop illustrated in Fig. 3-19 is given by a relation analogous to (3.3-26), namely,[4]

[4]Note that the loop now tracks the 4ϕ process.

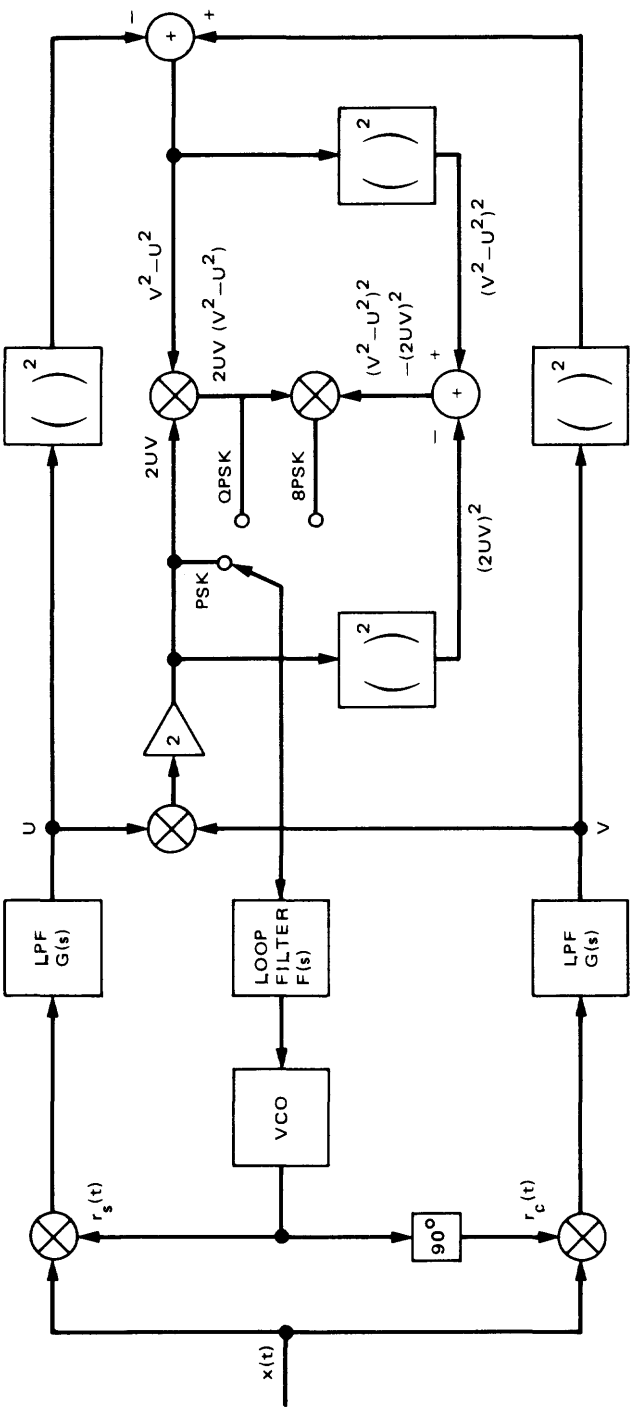


Fig. 3-18. Generalized Costas loop

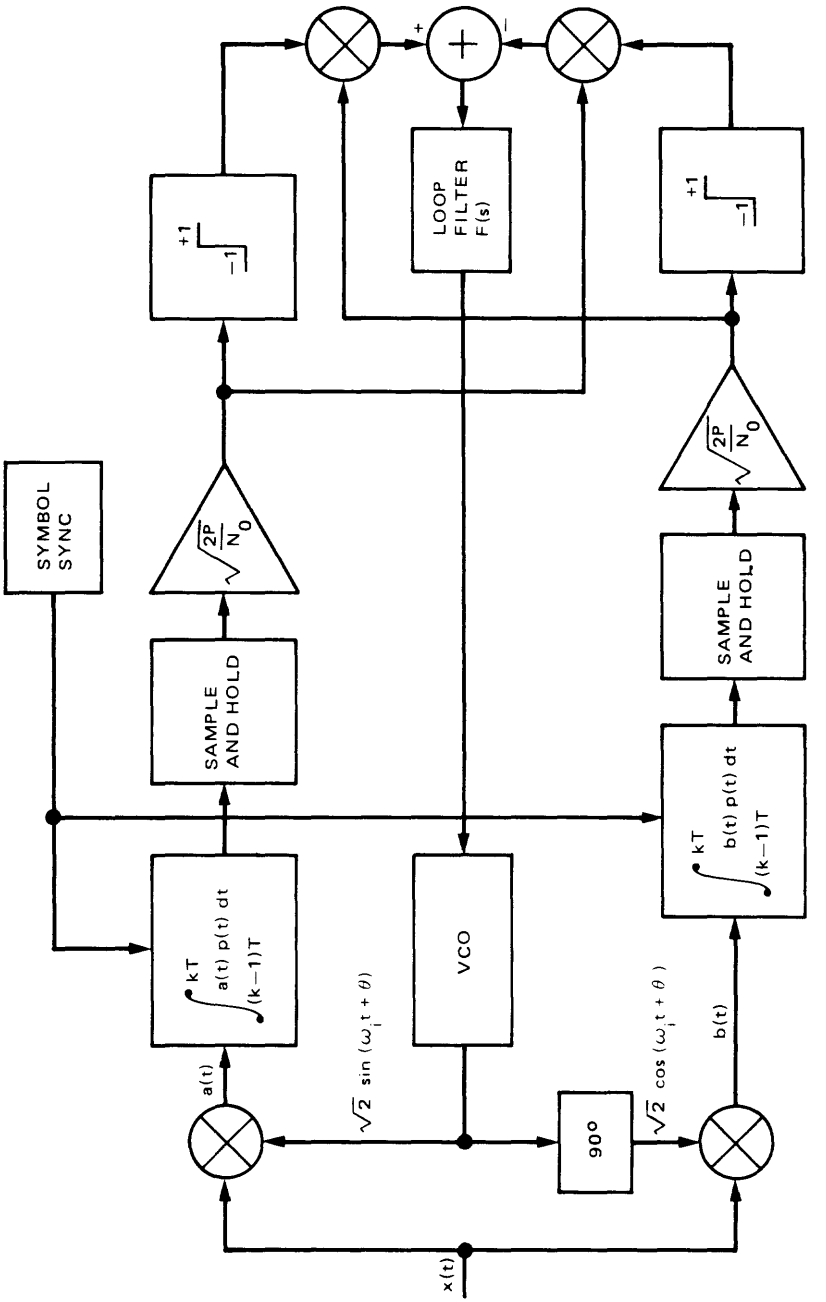


Fig. 3-19. The polarity-type QPSK carrier recovery loop

$$\sigma_{4\phi}^2 = \frac{16}{\rho S_L}; \sigma_{\phi}^2 = \frac{1}{16}\sigma_{4\phi}^2 = \frac{1}{\rho S_L} \tag{3.3-56}$$

where ρ is still given by (3.3-27) but the "squaring loss" S_L (actually a fourth power loss) is now given by the simple formula

$$S_L = \frac{\left[\operatorname{erf}\sqrt{\frac{R_d}{2}} - \sqrt{\frac{2R_d}{\pi}}\exp\left(-\frac{R_d}{2}\right)\right]^2}{1 + R_d - \left[\sqrt{\frac{2}{\pi}}\exp\left(-\frac{R_d}{2}\right) + \sqrt{R_d}\operatorname{erf}\sqrt{\frac{R_d}{2}}\right]^2} \tag{3.3-57}$$

where R_d is still given by (3.3-29).

In the absence of the hard limiters in Fig. 3-19, the squaring loss would be given by

$$S_L = \frac{1}{1 + \frac{9}{2R_d} + \frac{6}{R_d^2} + \frac{3}{2R_d^3}} \tag{3.3-58}$$

Figure 3-20 is a comparison of the squaring losses as given by (3.3-57) and (3.3-58) as a function of the *bit* signal-to-noise ratio $R_b = PT_b/N_0 = R_d/2$. Clearly, as already indicated, for high values of R_b, superior performance is obtained by including the hard limiters. Note that, as was true for the BPSK case, the above results hold for both NRZ and Manchester coded data.

3.3.3 Carrier Recovery for UQPSK Modulation

Unbalanced quadriphase-shift-keying (UQPSK) is a modulation format in which two quadrature components of a carrier are biphase modulated with independent binary data streams of unequal powers and different data rates. Such a scheme is most useful when it is desired to transmit the two different sources of data, either synchronously or asynchronously, in the same frequency band. This signal design will be used in the development of the Tracking and Data Relay Satellite System (TDRSS), the Navigation Satellite Timing and Ranging Global Positioning System (NAVSTAR GPS), the Space Shuttle S-band and Ku-band communication links, the Multimegabit Telemetry Demodulator/Detector (MTDD) system of NASA's DSN and is proposed for the projected Venus Orbiting Imaging Radar (VOIR) mission for the mapping of Venus's surface. In the case of the latter, the imaging data are to be transmitted on one channel at a multimegabit data rate and scientific data on the other channel at a relatively low kilobit data rate. Such a large ratio of data rates for the two channels or, equivalently, a large unbalanced channel

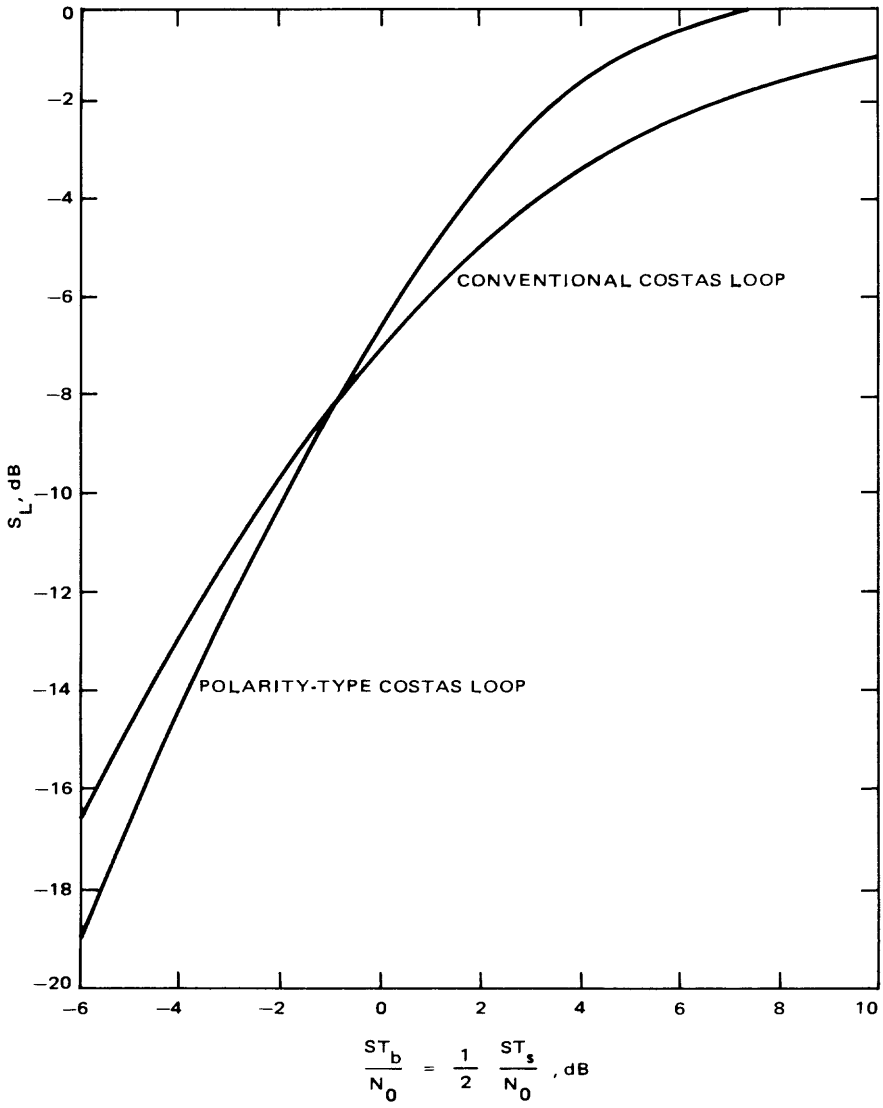


Fig. 3-20. A comparison of the squaring loss of the conventional and the polarity-type Costas loops

power ratio will provide the necessary motivation for the discussion which follows.

An unbalanced QPSK signal is defined analogous to (3.3-13) as

$$s(t) = \sqrt{2P_1}\, m_1(t) \cos(\omega_i t + \theta(t)) + \sqrt{2P_2}\, m_2(t) \sin(\omega_i t + \theta(t)) \tag{3.3-59}$$

where P_1 and P_2 represent the powers in the two quadrature components, $m_1(t)$ and $m_2(t)$ are two independent binary signal waveforms, and ω_i, $\theta(t)$ are previously defined for the BPSK case. Added to $s(t)$ of (3.3-59) is the noise process characterized by (3.3-14) which thus forms the received signal from which carrier synchronization is to be derived.

For a UQPSK modulation as characterized by (3.3-59), the system designer has two options with respect to selecting an appropriate carrier recovery loop. If the ratio of the channel powers P_1 and P_2 is sufficiently large (usually a ratio of 4 or more is sufficient), then a conventional Costas loop such as that in Fig. 3-11 can be used to reconstruct the suppressed carrier[5] [3-29, 3-30]. We shall restrict our attention here to the tracking performance of such a loop (see Fig. 3-11), particularly summarizing the results given in [3-29].

3.3.3.1 Loop tracking performance. We start by arbitrarily assuming that channel 2 corresponds to the high data rate channel. Next, we define the following parameters in terms of which the loop mean-square phase jitter and squaring loss can be expressed.

$$\begin{aligned} n &\triangleq \frac{T_{s1}}{T_{s2}} = \frac{R_{s2}}{R_{s1}} = \text{ratio of data symbol rates} \\ \lambda &\triangleq \frac{P_1}{P_2} = \text{ratio of channel powers} \\ P &= P_1 + P_2 = \text{total signal power} \\ R_{T2} &= \frac{P T_{s2}}{N_0} = \text{total power-to-noise ratio in high data rate bandwidth.} \end{aligned} \tag{3.3-60}$$

Then, based upon the results derived and presented in [3-29], the variance of the loop phase error 2ϕ is once again expressible in the form of (3.3-26a),

[5] The arm filter bandwidth is now chosen to be optimum for the high data rate channel.

where the linear loop signal-to-noise ratio ρ is now defined in terms of the total power-to-noise ratio; i.e.,

$$\rho \triangleq \frac{P}{N_0 B_L} \tag{3.3-61}$$

and the loop squaring loss is given by

$$S_L = \frac{(K_{22} - \lambda K_{21})^2}{(1+\lambda)\left[\lambda K_{21}K_{41} + K_{22}K_{42} + \frac{\left(\frac{B}{R_{s2}}\right)(1+\lambda)}{R_{T2}} K_L + \frac{2\lambda}{1+\lambda} R_{T2}\left(\frac{D_{12}}{T_{s2}}\right)\right]} \tag{3.3.62}$$

In (3.3-62), we have employed the following definitions analogous to (3.3-17) and (3.3-23), namely,

$$K_{2k} \triangleq \int_{-\infty}^{\infty} S_{m_k}(f)|G(j2\pi f)|^2\, df; \quad k = 1,2$$

$$K_{4k} \triangleq \frac{\int_{-\infty}^{\infty} S_{m_k}(f)|G(j2\pi f)|^4\, df}{\int_{-\infty}^{\infty} S_{m_k}(f)|G(j2\pi f)|^2\, df}; \quad k = 1,2 \tag{3.3-63}$$

and K_L is still defined as in (3.3-22). Furthermore, the crossmodulation distortion D_{12} is defined by

$$D_{12} \triangleq \int_{-\infty}^{\infty} S_{m_1}(f)\, S_{m_2}(f)\, |G(j2\pi f)|^4\, df \tag{3.3-64}$$

Since the demodulation reference signals of Fig. 3-11 are again at ω_0 rather than $2\omega_0$, the mean-square phase tracking jitter on the output data streams is given by (3.3-26b). We conclude then that, as in the BPSK case, characterization of the loop squaring loss in terms of the system parameters is sufficient for predicting the loop's tracking performance.

3.3.3.2 Squaring loss evaluation for practical filters and data modulation spectra. To graphically illustrate the theory previously developed, it would be

convenient to obtain closed-form expressions for the parameters defined in (3.3-63) and (3.3-64) for practical filters and data modulation formats. Indeed, such is possible for certain special cases of interest as we shall now illustrate. Let us consider again a simple RC filter with 3-dB cutoff frequency f_c for the Costas loop arm filters. Then, using (3.3-33) and (3.3-34) in (3.3-63), we arrive at results analogous to those in Table 3-2, where now R_s is replaced by R_{sk}; $k = 1, 2$.

All that remains is to present expressions for the cross-modulation coefficient D_{12} corresponding to the various combinations of modulation formats and relative data rates. These results are summarized in Table 3-3 for single-pole Butterworth arm filters, with

$$\zeta_1 \triangleq \frac{B}{R_{s1}},$$

$$\zeta_2 \triangleq \frac{B}{R_{s2}},$$

and

$$g(R_{s1}, R_{s2}) \triangleq \frac{D_{12}}{T_{s2}}, \text{ when both } m_1(t) \text{ and } m_2(t) \text{ are NRZ.}$$

The numerical evaluation of the tracking jitter, for a fixed ratio of arm filter noise bandwidth to loop noise bandwidth (B/B_L), is shown in Figs. 3-21 through 3-23. These figures illustrate the behavior of σ_ϕ, for a fixed channel power ratio λ as a function of the ratio of arm filter bandwidth to the higher data rate R_{s2} for all possible signal format combinations with R_{s2}/R_{s1} and PT_{s2}/N_0 as parameters. The minimum values of σ_ϕ for some ratios of R_{s1}/R_{s2} represents best design points when the combined effect of $N \times N$ distortion and cross-modulation noise or $S \times N$ power is minimal.

3.3.4 Automatic Gain Control Considerations

In practice, the implementation of a suppressed carrier reconstruction loop requires cross-coupled control loops and elements other than the ones which serve to track the phantom carrier. This is due to several facts of which the two most important are:

(1) The dynamic range of the received signal level may vary over a range of 100 dB or more.

(2) The received signal phase and frequency dynamics generally require a sync acquisition control loop to accommodate the uncertainties in the received carrier frequency.

Table 3-3. Normalized cross-modulation distortion D_{12}/T_{s2}

1. $m_1(t)$ is NRZ, $m_2(t)$ is NRZ

$$\frac{D_{12}}{T_{s2}} = g(R_{s1}, R_{s2})$$

$$= 1 - \frac{R_{s1}}{R_{s2}}\left(\frac{1}{3} + \frac{1}{4\zeta_2^2}\right) - \frac{R_{s1}}{64 R_{s2}\zeta_2^3}$$

$$\left\{5 - (5 + 4\zeta_2)\exp(-4\zeta_2) - (5 + 4\zeta_1)\exp(-4\zeta_1)\right.$$

$$+\frac{1}{2}\left[5 + 4\left(\frac{R_{s2}}{R_{s1}} - 1\right)\zeta_2\right]\exp\left[-4\left(\frac{R_{s2}}{R_{s1}} - 1\right)\zeta_2\right]$$

$$\left.+\frac{1}{2}\left[5 + 4\left(\frac{R_{s2}}{R_{s1}} + 1\right)\zeta_2\right]\exp\left[-4\left(\frac{R_{s2}}{R_{s1}} + 1\right)\zeta_2\right]\right\}, R_{s2} > R_{s1}$$

2. $m_1(t)$ is Manchester, $m_2(t)$ is NRZ

$$\frac{D_{12}}{T_{s2}} = \begin{cases} 2g(2R_{s1}, R_{s2}) - g(R_{s1}, R_{s2}), & R_{s2} \geq 2R_{s1} \\ 2g(R_{s2}, 2R_{s1}) - g(R_{s1}, R_{s2}), & R_{s1} \leq R_{s2} \leq 2R_{s1} \end{cases}$$

3. $m_1(t)$ is NRZ, $m_2(t)$ is Manchester

$$\frac{D_{12}}{T_{s2}} = 2g(R_{s1}, 2R_{s2}) - g(R_{s1}, R_{s2}), \quad R_{s2} \geq R_{s1}$$

4. $m_1(t)$ is Manchester, $m_2(t)$ is Manchester

$$\frac{D_{12}}{T_{s2}} = \begin{cases} 4g(2R_{s1}, 2R_{s2}) - 2g(2R_{s1}, R_{s2}) \\ -2g(R_{s1}, 2R_{s2}) + g(R_{s1}, R_{s2}), & R_{s2} \geq 2R_{s1} \\ 4g(2R_{s1}, 2R_{s2}) - 2g(R_{s1}, 2R_{s2}) \\ -2g(R_{s2}, 2R_{s1}) + g(R_{s1}, R_{s2}), & R_{s1} \leq R_{s2} \leq 2R_{s1} \end{cases}$$

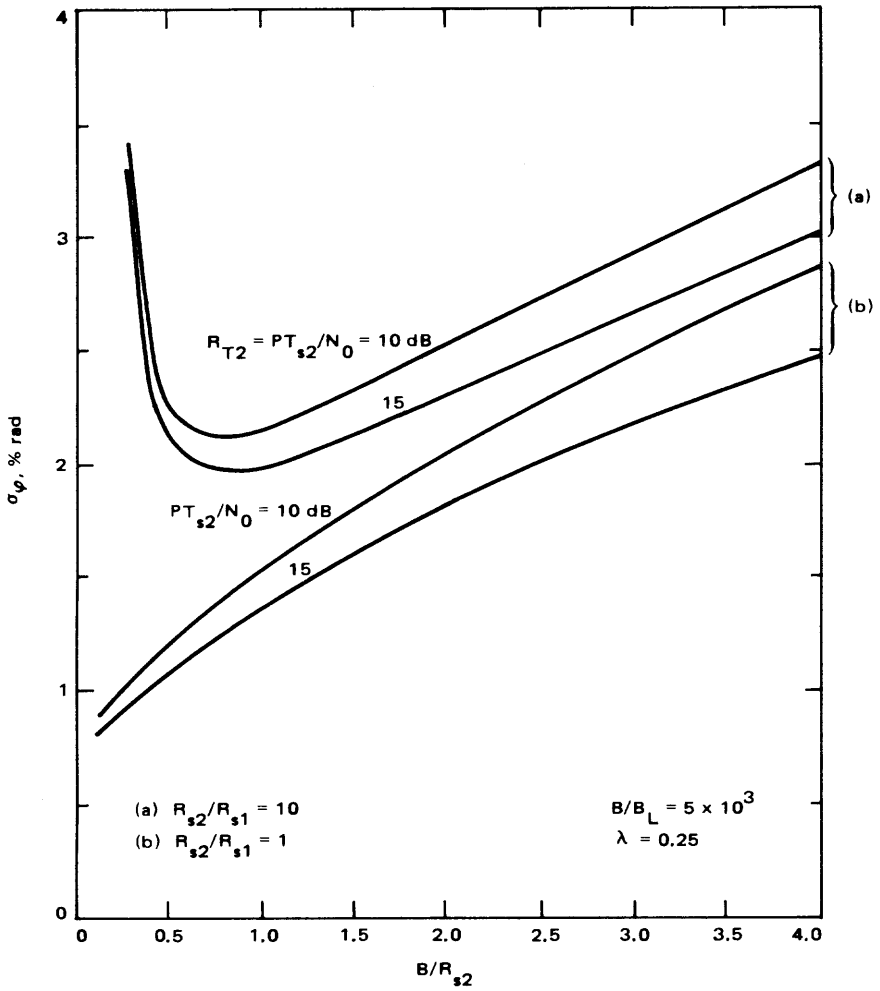


Fig. 3-21. Tracking jitter standard deviation vs ratio of arm filter bandwidth to high data rate; PT_{s2}/N_0 and R_{s2}/R_{s1} are parameters; $m_1(t)$ is NRZ, $m_2(t)$ is NRZ; $R_{s2} \geq R_{s1}$

Signal level variations are usually coped with by means of an automatic gain control (AGC) loop coupled with the suppressed carrier tracking loop. The major purpose of the AGC loop is to maintain a "constant" signal level into the receiver's front end so that linearity of receiver operation is maintained. As a by-product of the amplitude control loop, a lock detection signal for circuit supervision during sync acquisition to tracking handover is also generated.

The design of an AGC loop for a discrete carrier (PLL) receiver is discussed in [3-31] using a linearized model for predicting loop behavior corresponding

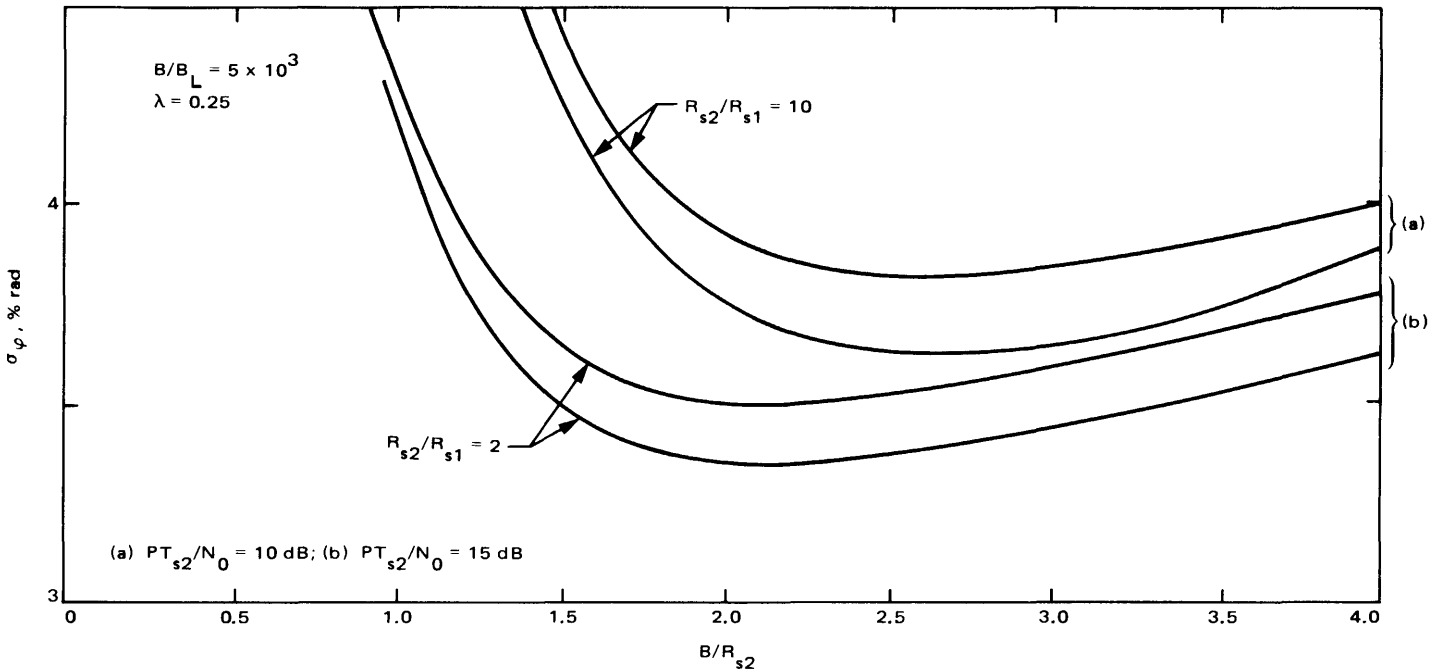


Fig. 3-22. Tracking jitter standard deviation vs ratio of arm filter bandwidth to high data rate; PT_{s2}/N_0 and R_{s2}/R_{s1} are parameters; $m_1(t)$ is Manchester code, $m_2(t)$ is Manchester code; $R_{s2} \geq R_{s1}$

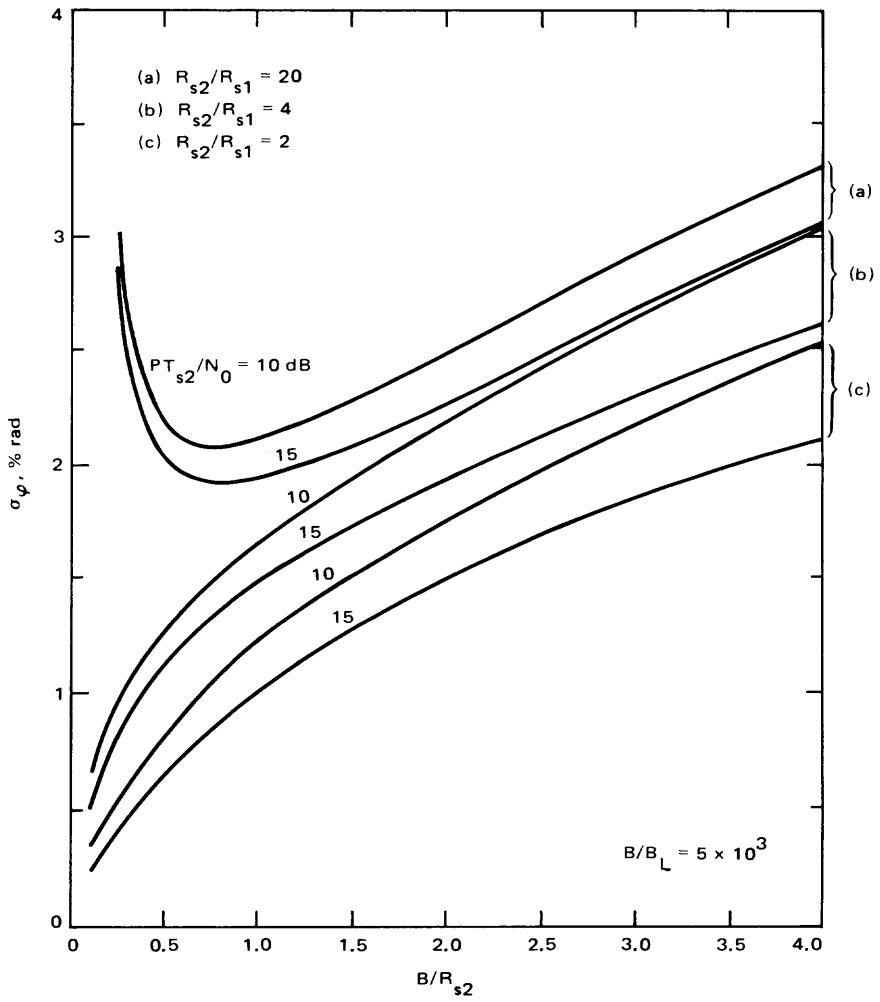


Fig. 3-23. Tracking jitter standard deviation vs ratio of arm filter bandwidth to high data rate; PT_{s2}/N_0 and R_{s2}/R_{s1} are parameters; $m_1(t)$ is Manchester code, $m_2(t)$ is NRZ; $R_{s2} \geq R_{s1}$

to small changes in input signal level. A subsequent large signal analysis [3-32] showed the linearized model to be valid under a much less restrictive condition. Briefly speaking, a simplified linear AGC system (as in [3-31]) can be illustrated as in Fig. 3-24. There, a coherent amplitude detector (CAD) translates the signal and noise at the IF to baseband. The AGC loop filter removes the noise and the clean dc output signal controls the IF amplifier gain.

On the contrary, very little information appears in the literature on AGC for suppressed carrier systems. The one exception to this statement is a paper

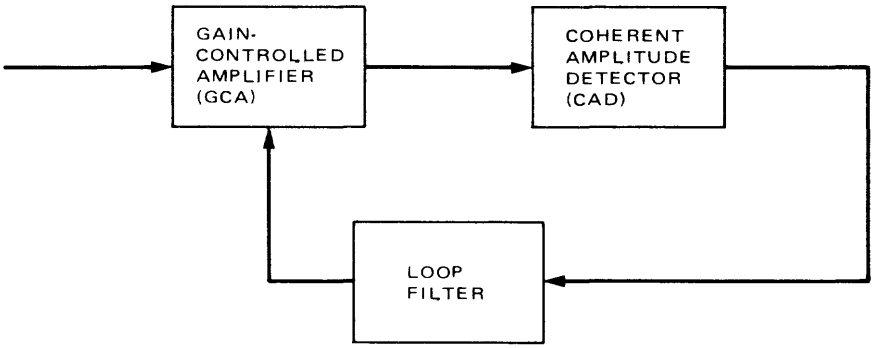


Fig. 3-24. A simple block diagram of an automatic gain control (AGC) loop

by Weber [3-33] which discusses a decision-directed AGC loop for use with M-ary amplitude- and phase-shift keying (MAPSK) systems, in particular M-ary amplitude-shift-keying (MASK) and quadrature amplitude-shift-keying (QASK).

A Costas loop receiver which incorporates both lock detector and AGC functions is illustrated in Fig. 3-25. The AGC input is derived from the square of the in-phase Costas loop output I^2, which is proportional to $P\cos^2\phi$. Thus this type of AGC is referred to as a partially coherent power AGC in that the AGC loop must be open during acquisition and closed by a switch upon detection and verification of lock. More often than not, a fully noncoherent type of power AGC is employed wherein the AGC input is derived from the sum of the squares of the in-phase and quadrature outputs ($I^2 + Q^2$), which is proportional to $P\ (\cos^2\phi + \sin^2\phi) = P$. Thus this type of AGC loop can remain closed during carrier acquisition. The primary disadvantage of this type of loop is that when the loop is in lock (nominally, $\phi = 0$), then the $I^2 + Q^2$ signal contributes twice the noise as compared to the I^2 signal of the partially coherent AGC.

It is well known [3-6] that the bandwidth, natural frequency and damping of a loop which employs the phase-lock principle all depend upon the loop gain. This gain is proportional to the received signal level. Thus, in addition to maintaining linearity in the receiver front end, the AGC serves to regulate the variation of the above performance parameters. We now examine this behavior in mathematical terms for the conventional Costas loop with and without a noncoherent AGC.

Figure 3-26 illustrates a conventional Costas loop with no AGC and a perfect integrating type of loop filter. For this case the threshold loop damping parameter (typically 0.707) at the threshold (minimum) signal-to-noise-spectral-density ratio P_0/N_0 is

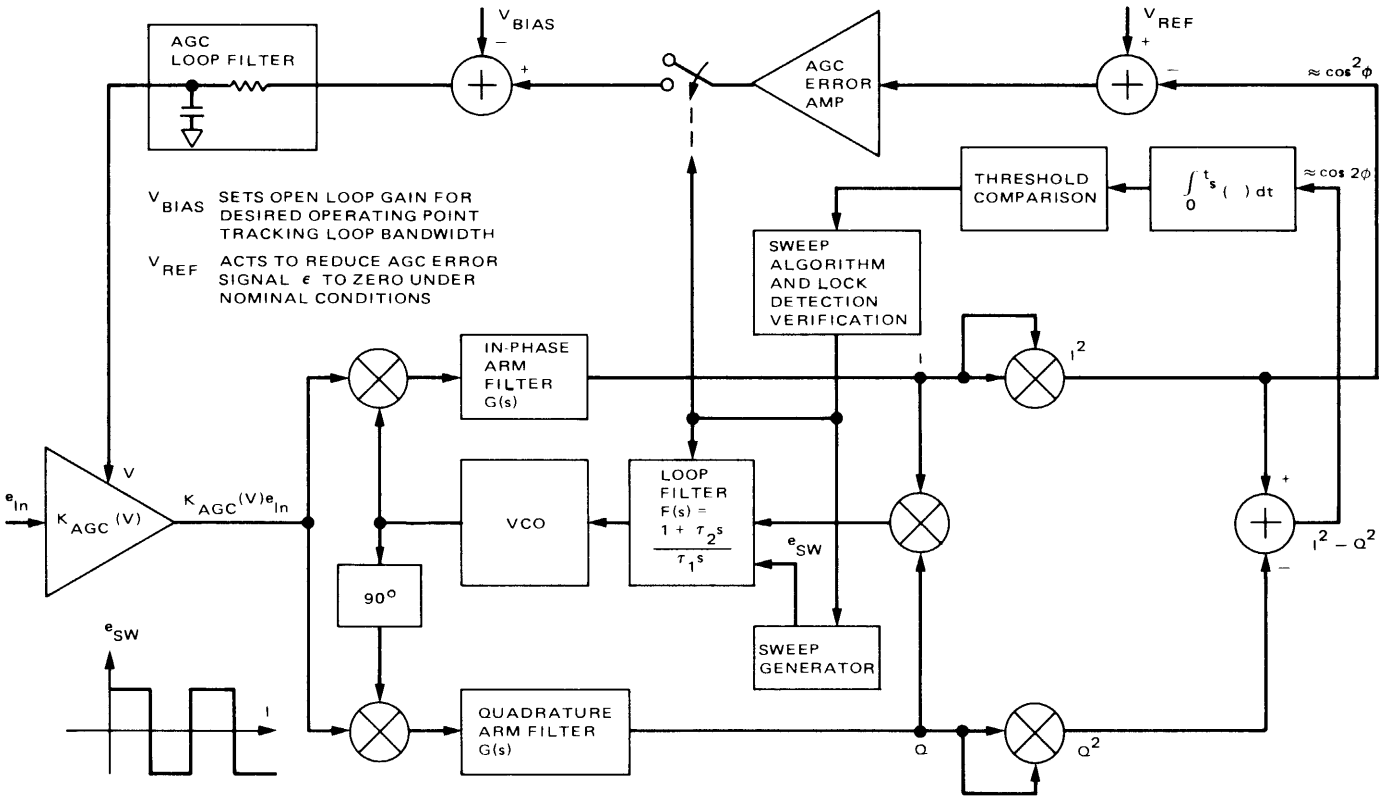
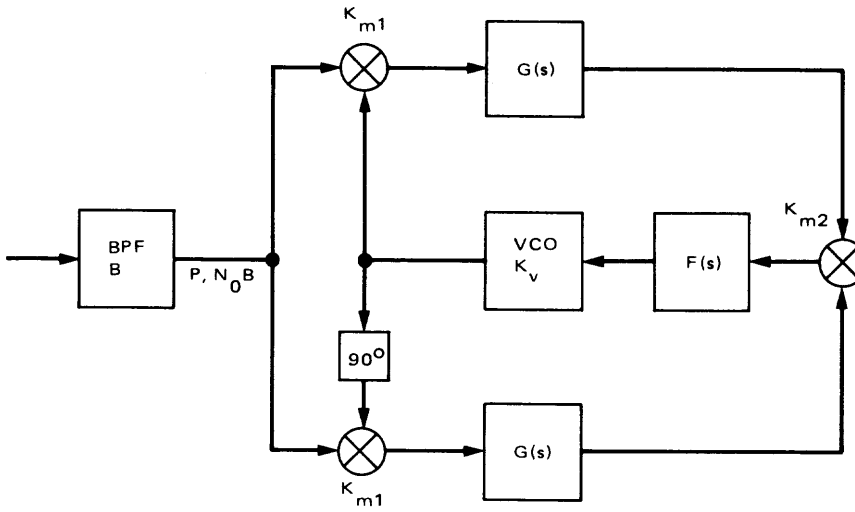


Fig. 3-25. Costas loop receiver with lock detector and AGC functions



OPEN-LOOP GAIN

$$K = K_1^2 K_{m1}^2 K_{m2} K_v$$

SECOND-ORDER LOOP FILTER

$$F(s) = \frac{1 + \tau_2 s}{\tau_1 s}, \quad \frac{\tau_2}{\tau_1} \ll 1$$

ASSUME G(0) = 1

Fig. 3-26. Costas loop design (no AGC)

$$\zeta_0 = \sqrt{\frac{P_0 K \tau_2^2}{4\tau_1}} \tag{3.3-65}$$

(Note: ζ_0^2 is proportional to signal power rather than signal voltage as would be the case in a CW loop, e.g., a PLL.) The threshold loop bandwidth is

$$B_{L0} = \frac{4\zeta_0^2 + 1}{4\tau_2} \tag{3.3-66}$$

The same operating point loop damping and operating point loop bandwidth parameters at the operating signal-to-noise-spectral-density ratio P/N_0 are

$$\zeta = \sqrt{\frac{P}{P_0}}\zeta_0 = \sqrt{\frac{\frac{P}{N_0}}{\frac{P_0}{N_0}}}\zeta_0 \tag{3.3-67}$$

and

$$B_L = \left(\frac{4\zeta^2 + 1}{4\zeta_0^2 + 1}\right) B_{L0} = \left[\frac{4\zeta_0^2 \frac{P}{P_0} + 1}{4\zeta_0^2 + 1}\right] B_{L0} \tag{3.3-68}$$

Figure 3-27 illustrates the same Costas loop design with a noncoherent AGC included. The threshold and operating point loop parameters are now given by

$$\zeta_0 = \sqrt{\frac{C_0 K \tau_2^2}{4\tau_1}}, B_{L0} = \frac{C_0 K \tau_2}{4\tau_1} \tag{3.3-69}$$

$$\zeta = \sqrt{\frac{C}{C_0}}\,\zeta_0 = \sqrt{\frac{1 + \frac{N_0 B}{P_0}}{1 + \frac{N_0 B}{P}}}\,\zeta_0 \tag{3.3-70}$$

$$B_L = \left[\frac{4\zeta_0^2 \frac{C}{C_0} + 1}{4\zeta_0^2 + 1}\right] B_{L0} = \left[\frac{4\zeta_0^2 \left(\frac{1 + \frac{N_0 B}{P_0}}{1 + \frac{N_0 B}{P}}\right) + 1}{4\zeta_0^2 + 1}\right] B_{L0} \tag{3.3-71}$$

$$\frac{C}{C_0} = \frac{1 + (SNR_0)^{-1}}{1 + (SNR)^{-1}} \tag{3.3-72}$$

where

$C = K_{AGC}^2 P$ = input signal power to Costas loop at operating point

$C_0 = K_{AGC}^2 P_0$ = input signal power to Costas loop at threshold

$SNR_0 \triangleq \frac{P_0}{N_0 B}$ = input SNR at threshold

$SNR \triangleq \frac{P}{N_0 B}$ = input SNR at operating point

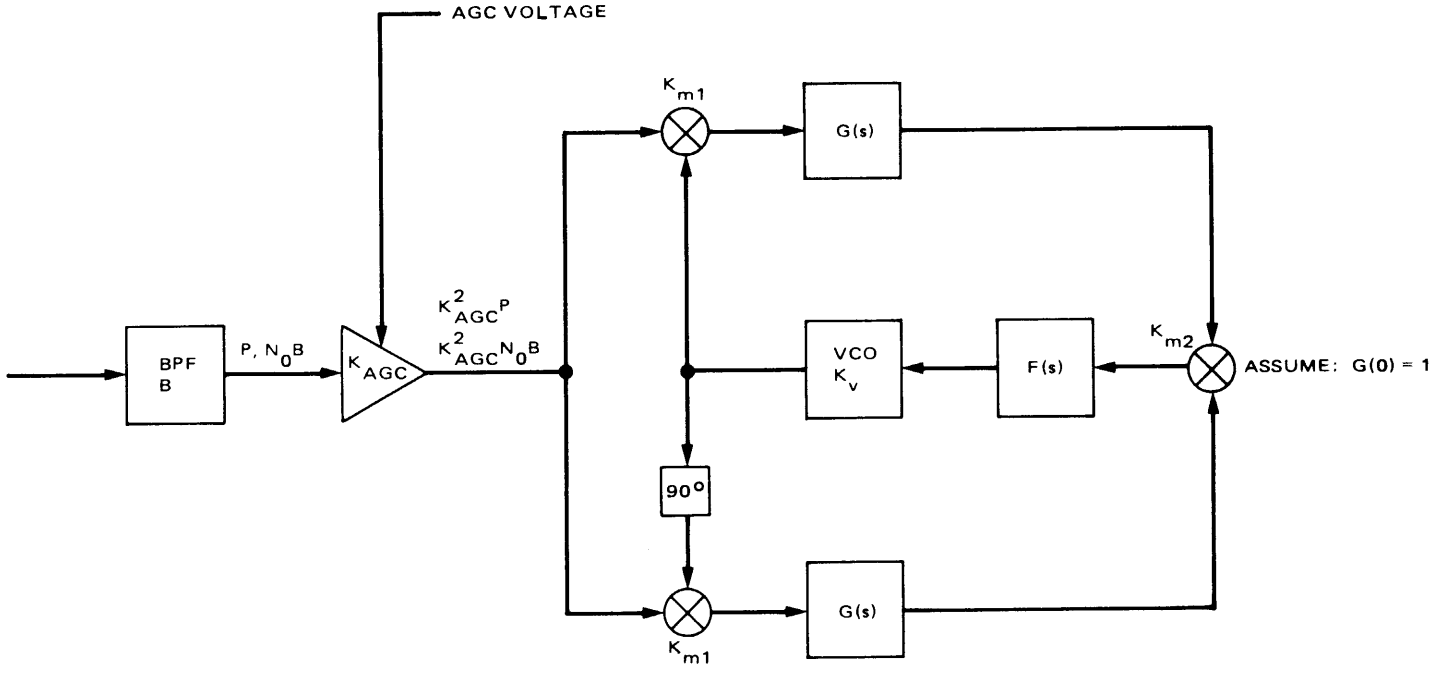


Fig. 3-27. Costas loop design with "ideal" fully noncoherent AGC

Figure 3-28 illustrates B_L/B_{L0} as a function of P/P_0 for various values of P/N_0B in dB. The case of no AGC, as computed from (3.3-68), is also shown for comparison.

3.3.5 Lock Detector Considerations

A lock detector generates a signal which indicates to the receiver when to switch from the sync acquisition to tracking mode of operation and vice versa. Such supervision of acquisition to tracking handover is ordinarily accomplished in two stages, the first being an indication of when to make the switch and the second being a verification (typically a specific algorithm) as to whether or not to proceed with the switch. The end result of this two-stage operation is a decision to declare in-lock and initiate carrier tracking if the loop was previously out-of-lock or to declare the loop out-of-lock and reinitiate carrier acquisition if the loop was previously in-lock.

A lock detector for suppressed carrier biphase reception is illustrated in Fig. 3-29. The $I^2 - Q^2$ signal derived from the I and Q signals produced by the Costas loop (see Fig. 3-25) is low-pass filtered to produce a voltage V which, when compared to a fixed threshold δ, is used to decide whether or

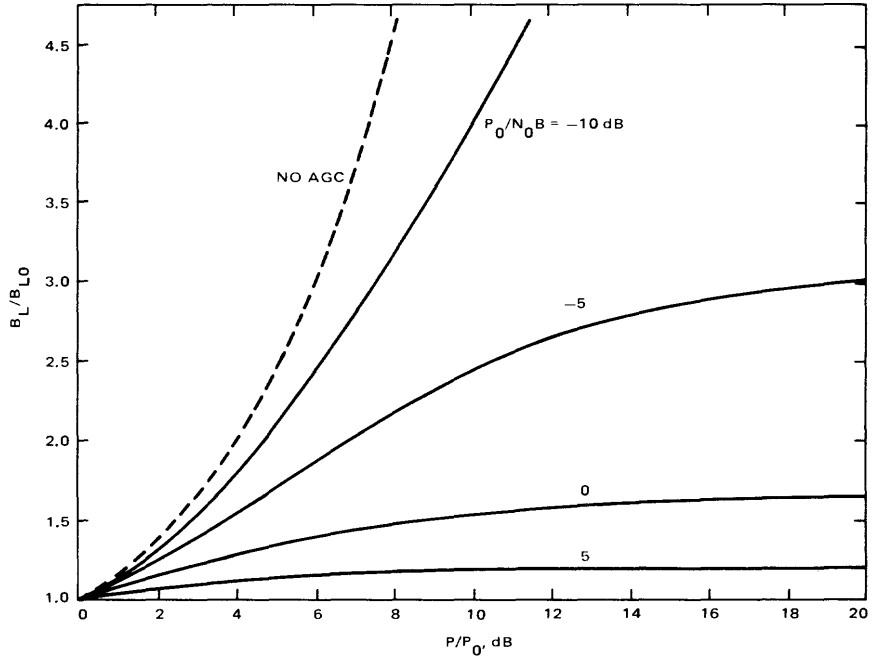


Fig. 3-28. Loop bandwidth expansion as a function of change in carrier power level with threshold signal-to-noise ratio in the IF bandwidth as a parameter

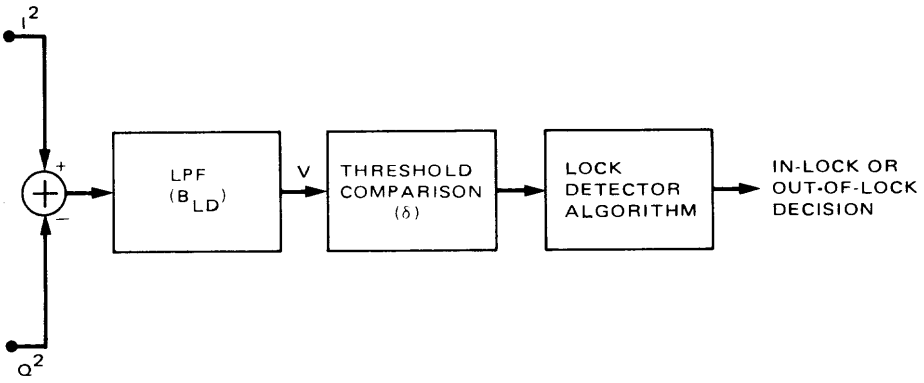


Fig. 3-29. A lock detector for suppressed carrier biphase signals

not to proceed with the lock verification algorithm. The mathematical analysis which describes the operation of the lock detector is now presented.

If, for simplicity, we assume all multiplier gains equal to unity in the Costas loop, then from Fig. 3-25 and (3.3-15), we have that the loop error voltage is given by

$$2(IQ) = P\hat{m}^2(t)\sin\overbrace{2\varphi(t)}^{\text{loop phase error}} + \underbrace{v_2\,[t, 2\varphi(t)]}_{\substack{\text{equivalent}\\ \text{additive}\\ \text{noise}}}\} \tag{3.3-73}$$

where $v_2\,[t, 2\varphi]$ is again given by (3.3-16). By similar analysis, it can be shown that the lock detector voltage has the identical form of (3.3-73) except that sin 2φ is replaced by cos 2φ and $v_2\,[t, 2\varphi]$ is replaced by $v_2\,[t, 2\varphi + \pi/2]$, namely,

$$I^2 - Q^2 = P\hat{m}^2(t)\cos 2\varphi(t) + v_2\left[t, 2\varphi(t) + \frac{\pi}{2}\right] \tag{3.3.74}$$

The consequence of replacing 2φ by $2\varphi + \pi/2$ in v_2 is merely of mathematical curiosity since both $v_2\,[t, 2\varphi]$ and $v_2\,[t, 2\varphi + \pi/2]$ have identical noise spectral densities, namely N_{sq} (see (3.3-19)) which is all that is important from a performance standpoint.

Typically, the lock detector is designed such that the noise bandwidth B_{LD} of the low-pass filter which acts on the $I^2 - Q^2$ signal is small compared to the data symbol rate R_s; i.e., $B_{LD}/R_s \ll 1$. Thus under this assumption the

filter output voltage V can be approximated as a Gaussian random variable with mean $\overline{V}$ (which depends on whether the loop is in-lock or out-of-lock) and variance σ_V^2.

When the loop is out-of-lock, it is reasonable to assume that the phase error 2φ is uniformly distributed on $(0, 2\pi)$, in which case

$$\begin{aligned} \overline{V}_{O.L.} &= 0 \\ \sigma_V^2 &= N_{sq} B_{LD} \end{aligned} \tag{3.3-75}$$

Alternately when the loop is in-lock, we can set $\varphi = 0$ and thus[6]

$$\begin{aligned} \overline{V}_{I.L.} &= P K_2 \\ \sigma_V^2 &= N_{sq} B_{LD} \end{aligned} \tag{3.3-76}$$

where K_2 is the modulation distortion factor defined in (3.3-17).

3.3.5.1 Detection probabilities. Since V is assumed to be Gaussian, then the probability of crossing the threshold is given by

$$p \triangleq \Pr\{V > \delta\} = \frac{1}{2}\operatorname{erfc}\left(\frac{\delta - \overline{V}}{\sqrt{2}\,\sigma_V}\right) = \frac{1}{2}\left[1 - \operatorname{erf}\left(\frac{\delta - \overline{V}}{\sqrt{2}\,\sigma_V}\right)\right] \tag{3.3-77}$$

For the in-lock condition, we use (3.3-76) in (3.3-77) to give

$$p = \frac{1}{2}\left\{1 + \operatorname{erf}\left[\sqrt{\frac{\rho_{LD}}{2}}\,(1 - \delta_0)\right]\right\} \tag{3.3-78}$$

where δ_0 is the normalized threshold, namely,

$$\delta_0 \triangleq \frac{\delta}{\overline{V}_{I.L.}} = \frac{\delta}{PK_2} \tag{3.3-79}$$

and ρ_{LD} is the effective lock detector signal-to-noise ratio defined by

[6]For the lock detector performance during in-lock conditions, we shall assume perfect carrier tracking.

$$\rho_{LD} \triangleq \frac{\overline{V}_{I.L.}}{\sigma_V^2} = \frac{(PK_2)^2}{N_{sq}B_{LD}} = \left(\frac{\rho}{4}S_L\right)\left(\frac{B_L}{B_{LD}}\right) \tag{3.3-80}$$

with ρ, B_L, and S_L as defined in (3.3-27). For the out-of-lock condition, the use of (3.3-75) in (3.3-77) produces

$$p = \frac{1}{2}\left\{1 - \text{erf}\left[\sqrt{\frac{\rho_{LD}}{2}}\,\delta_0\right]\right\} \tag{3.2-81}$$

3.3.5.2 Lock detection performance measures. Lock detection performance is typically measured by false alarm probability P_{FA} and mean time to false alarm $\overline{T}_{FA}$ when the loop is out-of-lock, and false dismissal probability P_{FD} and mean time to false dismissal $\overline{T}_{FD}$ when the loop is in-lock. These performance measures are defined below:

P_{FA} = probability that loop is declared in-lock (and thus sweep is terminated) when loop is indeed out-of-lock

$\overline{T}_{FA}$ = mean time for a false alarm event to occur

P_{FD} = probability that loop is declared out-of-lock (and thus sweep is reinitiated) when loop is indeed in-lock

$\overline{T}_{FD}$ = mean time for a false dismissal event to occur

In order to evaluate these performance measures and express them in terms of the detection probabilities of the previous subsections, or equivalently, the tracking loop and lock detector parameters, one must specify a specific lock detector verification algorithm.

3.3.5.3 Lock detector algorithm and false alarm performance. As an example, consider that an in-lock condition is declared only if the threshold is exceeded *twice* in succession. A Markov state diagram corresponding to this simple algorithm is illustrated in Fig. 3-30. Here, the transition probability p is the probability of V exceeding the threshold as given by (3.3-81) and $q = 1 - p$ is the probability that V does not exceed the threshold. Furthermore, T_i; $i = 0, 1$ is the dwell time for state i; $i = 0, 1$; e.g., T_0 corresponds to the time between successive samplings (tests) of the threshold output.

The analysis that produces P_{FA} and T_{FA} from Fig. 3-30 follows the well-known theory associated with finite Markov chains [3-34]. In particular, we have that

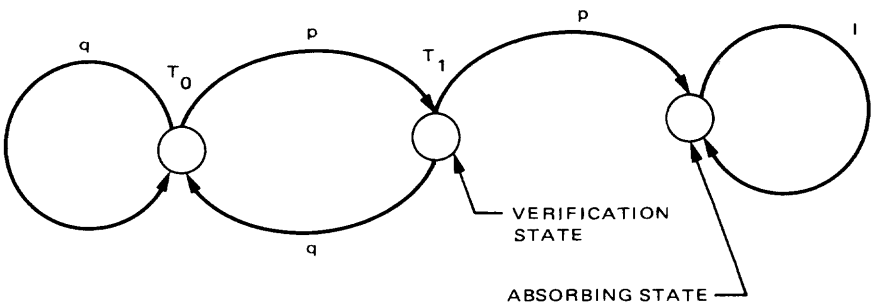


Fig. 3-30. A Markov state representation of a lock detector algorithm for declaring in-lock given an out-of-lock condition

$$P_{FA} = (p|_{O.L.})^2$$

$$\overline{T}_{FA} = \frac{T_0 + (p|_{O.L.})\,T_1}{(p|_{O.L.})^2} \tag{3.3-82}$$

3.3.5.4 Lock detection algorithm and false dismissal performance. Using a similar algorithm (see Fig. 3-31) to determine an out-of-lock condition when the loop is indeed in-lock, namely out-of-lock is declared only if the threshold is not exceeded *twice* in succession, we arrive at the following false dismissal performance measures:

$$P_{FD} = (q|_{I.L.})^2; \quad q|_{I.L.} = 1 - p|_{I.L.}$$

$$\overline{T}_{FD} = \frac{T_0 + (q|_{I.L.})\,T_1}{(q|_{I.L.})^2} \tag{3.3-83}$$

3.3.5.5 Lock detector performance. Combining (3.3-78) and (3.3-81) with (3.3-82) and (3.3-83), we can write the relations

$$p|_{O.L.} = \sqrt{P_{FA}} = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{\rho_{LD}}{2}}\,\delta_0\right) \tag{3.3-84a}$$

$$q|_{I.L.} = \sqrt{P_{FD}} = \frac{1}{2}\operatorname{erfc}\left[\sqrt{\frac{\rho_{LD}}{2}}\,(1 - \delta_0)\right] \tag{3.3-84b}$$

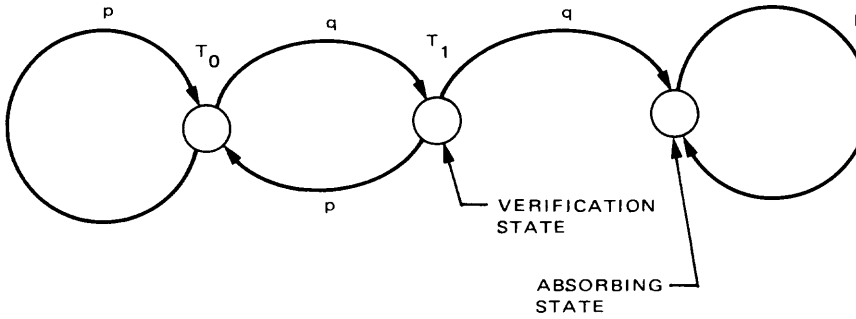


Fig. 3-31. A Markov state representation of a lock detector algorithm for declaring out-of-lock given an in-lock condition

Eliminating the normalized threshold δ_0 between the above two equations and rewriting the lock detector signal-to-noise ratio ρ_{LD} in the form

$$\rho_{LD} \triangleq \left(\frac{P}{4N_0B_L}\right) S_L\left(\frac{B_L}{B_{LD}}\right) = \frac{R_d}{4} S_L\left(\frac{R_s}{B_{LD}}\right); \tag{3.3-85}$$

where

$$R_d = \frac{P T_s}{N_0}$$

and

$$R_s = \frac{1}{T_s}$$

gives the lock detector performance equation

$$\text{erfc}^{-1}(2\sqrt{P_{FA}}) + \text{erfc}^{-1}(2\sqrt{P_{FD}}) = \sqrt{\frac{R_d}{8} S_L\left(\frac{R_s}{B_{LD}}\right)} \tag{3.3-86}$$

Figures 3-32 and 3-33 illustrate the behavior of (3.3-86) for NRZ data and two different values of false alarm probability. In each case the squaring loss has been minimized by proper selection of arm-filter-bandwidth-to-data-rate ratio as per the discussion previously given for Costas loop carrier tracking.

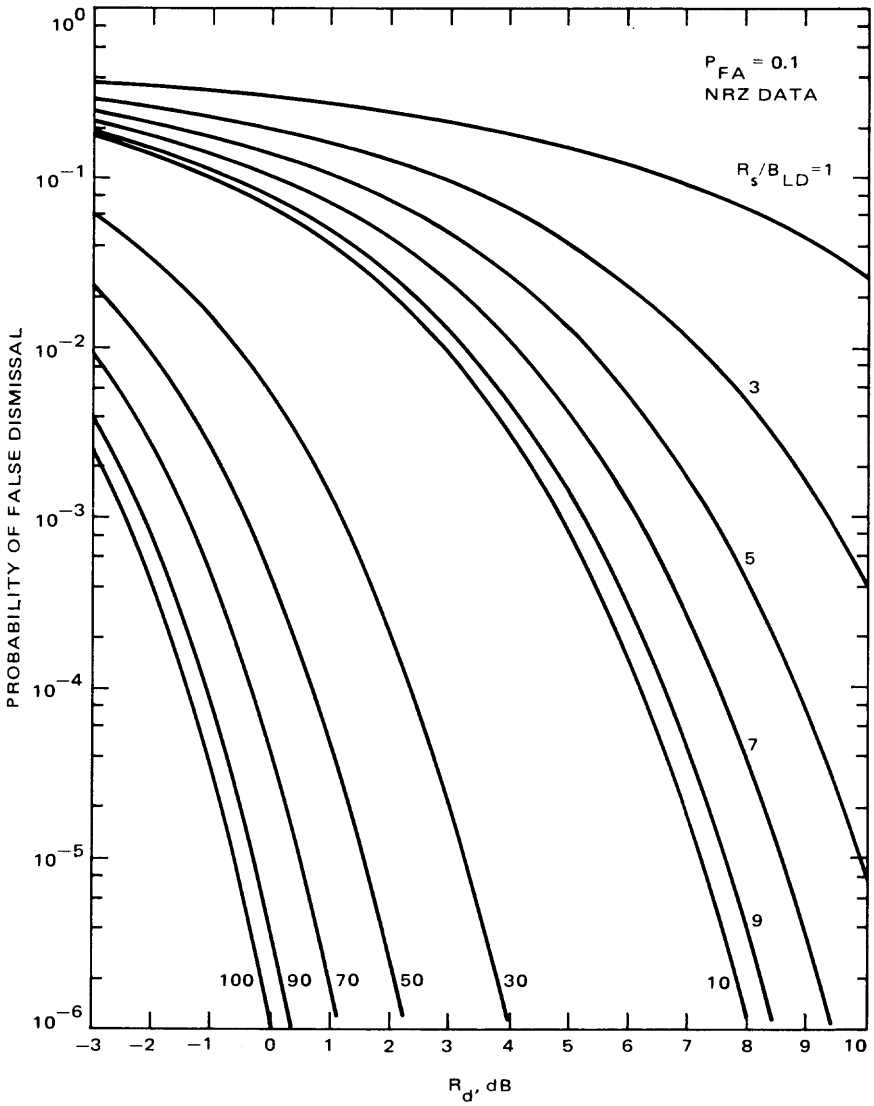


Fig. 3-32. False dismissal probability vs detection signal-to-noise ratio (P_{FA} = 0.1)

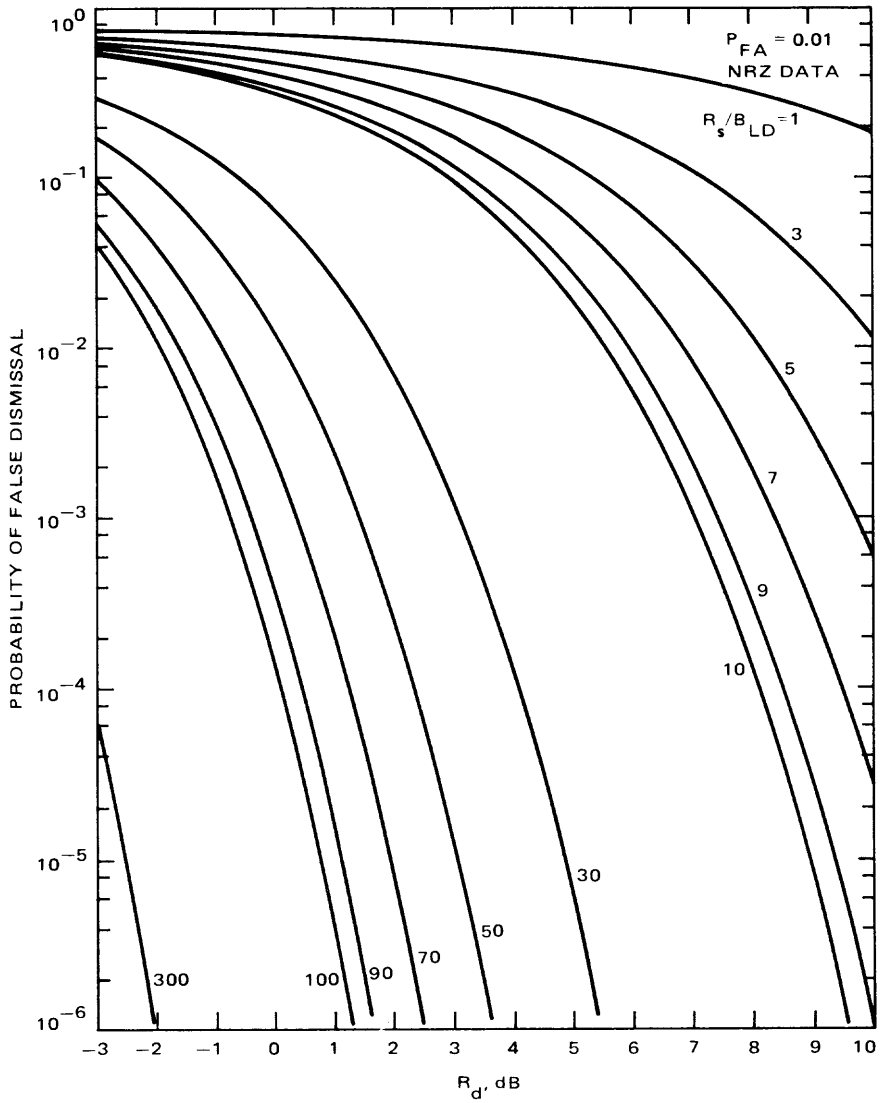


Fig. 3-33. False dismissal probability vs detection signal-to-noise ratio (P_{FA} = 0.01)

References

3-1. Jaffe, R. M., and Rechtin, E., "Design and Performance of Phase-Lock Circuits Capable of Near Optimum Performance Over a Wide Range of Input Signals and Noise Levels," *IRE Transactions on Information Theory*, Vol. IT-1, pp. 66–76, Mar. 1955.

3-2. Viterbi, A. J., *Principles of Coherent Communications*, McGraw-Hill, New York, 1966.

3-3. Lindsey, W. C., *Synchronization Systems in Communication and Control*, Prentice-Hall, Englewood Cliffs, N.J., 1972.

3-4. Blanchard, A., *Phase Locked Loops: Application to Coherent Receiver Design, Interscience*, New York, 1976.

3-5. Tausworth, R. C., *Theory and Practical Design of Phase-Locked Receivers*, Volume 1, Report No. 32-819, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 16, 1966.

3-6. Lindsey, W. C., and Simon, M. K., *Telecommunication Systems Engineering*, Prentice-Hall, Englewood Cliffs, N.J., 1973.

3-7. Lindsey, W. C., and Simon, M. K., editors, *Phase-Locked Loops and Their Application*, IEEE Press, New York, 1978.

3-8. Gardner, F. M., *Phaselock Techniques*, John Wiley and Sons, New York, 1979.

3-9. Holliday, D., "A Review of Some Phase-Locked Loop Theory," Proceedings, International Communication Conference, Montreal, Canada, pp. 3–2 to 3–8, June 1971.

3-10. Lindsey, W. C., "Coding and Synchronization Studies: Moments of the First Passage Time in Generalized Tracking Systems," *Space Program Summary 37–58*, Vol. III, pp. 63–66, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 31, 1969.

3-11. Frazier, J. P., and Page, J., "Phase-Lock Loop Frequency Acquisition Study," *Transactions of the IRE*, Vol. SET-8, Sept. 1962.

3-12. Tikhonov, V. I., "The Effect of Noise on Phase-Locked Oscillator Operation," *Automation and Remote Control*, Vol. 20, (1959) 1160–1168. Translated from *Automatika i Telemekhaniki*, Akademya Nauk SSSR, Vol. 20, Sept. 1959.

3-13. Tikhonov, V. I., "Phase-Lock Automatic Frequency Control Operation in the Presence of Noise," *Automation and Remote Control*, Vol. 21, (1960) 209–214. Translated from *Automatika i Telemekhaniki*, Akademya Nauk SSSR, Vol. 21, Mar. 1960.

3-14. Viterbi, A. J., "Phase-Locked Loop Dynamics in the Presence of Noise by Fokker-Planck Techniques," *Proceedings of the IEEE*, Vol. 51, No. 12, pp. 1737-1753, Dec. 1963.

3-15. Charles, F. J., and Lindsey, W. C., "Some Analytical and Experimental Phase-Locked Loop Results for Low Signal-to-Noise Ratios," *IEEE Proceedings*, Vol. 54, No. 9, pp. 1152-1166, Sept. 1966.

3-16. Tausworth, R., and Sanger, D., "Digital Communication and Tracking: Experimental Study of the First Slip Statistics of the Second Order Phase Locked Loop," *Space Programs Summary 37-43*, Vol. III, pp. 76-80, Jet Propulsion Laboratory, Pasadena, Calif., Jan. 31, 1967.

3-17 Yuen, J. H., "On First Cycle Slip Time of Phase Locked Loops in Cascade," *IEEE Transactions on Aerospace and Electronics Systems,* Vol. AES-10, No. 3, pp. 373-379, May 1974.

3-18. Davenport, W. B., "Signal-to-Noise Ratios in Band-Pass Limiters," *Journal of Applied Physics*, 24, pp. 720-727, June 1953.

3-19. Springett, J. C., and Simon, M. K., "An Analysis of the Phase Coherent-Incoherent Output of the Bandpass Limiter," *IEEE Transactions on Communication Technology*, Vol. COM-19, No. 1, pp. 42-49, Feb. 1971.

3-20. Tausworthe, R. C., "Information Processing: Limiters in Phase-Locked Loops: A Correction to Previous Theory," *Space Programs Summary 37-54*, Vol. III, pp. 201-204, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 31, 1968.

3-21. Lesh, J., "Tracking Loop and Modulation Format Considerations for High Rate Telemetry," *DSN Progress Report 42-44*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 117-124, Apr. 15, 1978.

3-22. Reasoner, R., Stevens, G., and Woo, K. T., "Costas Loop Demodulation of Suppressed Carrier BPSK in the DSN Environment-Experimental Results Obtained at TDL," *DSN Progress Report 42-51*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 94-104, June 15, 1979.

3-23. Woo, K. T., "Effects of Asymmetric Passband Filtering on the Phase of the Costas Loop's Reconstructed Carrier," *DSN Progress Report 42-51*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 105-112, June 15, 1979.

3-24. Stevens, G., and Woo, K. T., "Design of Costas Loop to Operate with the Block III Receiver and Its Predicted Performance," *DSN Progress Report 42-51*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 113-123, June 15, 1979.

3-25. Simon, M. K., and Lindsey, W. C., "Optimum Performance of Suppressed Carrier Receivers with Costas Loop Tracking," *IEEE Transactions on Communications*, Vol. COM-25, No. 2, pp. 215-227, Feb. 1977.

3-26. Simon, M. K., "On the Calculation of Squaring Loss in Costas Loops with Arbitrary Arm Filters," *IEEE Transactions on Communications*, Vol. COM-26, No. 1, pp. 179–184, Jan. 1978.

3-27. Simon, M. K., "Tracking Performance of Costas Loops with Hard-Limited In-Phase Channel," *IEEE Transactions on Communications*, Vol. COM-26, No. 4, pp. 420–432, Apr. 1978.

3-28. Brockman, M. H., "MMTS Subcarrier Demodulator," *Space Programs Summary 37–46*, Vol. III, Jet Propulsion Laboratory, Pasadena, Calif., pp. 189–204, July 31, 1967.

3-29. Simon, M. K., and Alem, W. K., "Tracking Performance of Unbalanced QPSK Demodulators; Part I–Biphase Costas Loop with Passive Arm Filters," *IEEE Transactions on Communications*, Vol. COM-26, No. 8, pp. 1147–1156, Aug. 1978.

3-30. Simon, M. K., "Tracking Performance of Unbalanced QPSK Demodulators; Part II–Biphase Costas Loop with Active Arm Filters," *IEEE Transactions on Communications*, Vol. COM-26, No. 8, pp. 1157–1166, Aug. 1978.

3-31. Victor, W. K., and Brockman, M. H., "The Application of Linear Servo Theory to the Design of AGC Loops," *Proceedings of the IRE*, Vol. 48, pp. 234–238, Feb. 1960.

3-32. Ohlson, J. E., "Exact Dynamics of Automatic Gain Control," *IEEE Transactions on Communications*, Vol. COM-22, pp. 72–74, Jan. 1974.

3-33. Weber, W. J., III, "Decision-Directed Automatic Gain Control for MAPSK Systems," *IEEE Transactions on Communications*, Vol. COM-23, No. 5, pp. 510–517, May 1975.

3-34. Kemeny, J., and Snell, J., *Finite Markov Chains*, Van Nostrand, N.Y., 1960.

Chapter 4
Radio Tracking System

Julian C. Breidenthal and Tomas A. Komarek

4.1 Introduction

This chapter describes the principles and techniques of deep space radio tracking. The emphasis here is on the measurement functions of radio tracking, since the problems of establishing and maintaining radio contact are well covered elsewhere in this book.

The organization of this chapter is deductive, that is, it proceeds from a general description to increasing levels of detail. An effort has been made, however, to stop at the level of detail that might be found in the "Theory of Operation" section in a particular instrument's operation manual. This section briefly describes the tracking system and how it is used. Section 4.2 describes the uses of tracking data, which motivate the specific measurements made. Section 4.3 focuses on measurement techniques, while Section 4.4 examines the errors inherent in these measurements.

4.1.1 Purpose

The radio tracking system has a dual purpose for deep space exploration. First, the radio tracking system performs radiometric functions in which information is obtained on spacecraft position, the radio propagation medium, and

the properties of the solar system. This information makes important contributions both to the navigation of spacecraft and to the scientific value of deep space missions. Second, the tracking system provides radio frequency carriers and additional reference signals that are used for telemetry and command functions. The latter function is discussed elsewhere in this book.

4.1.2 System Description

The tracking system consists of all equipment that is used to generate radio signals to and from the spacecraft and to receive, track, and record the properties of these signals. The actual elements of the system differ depending on whether two-way (round-trip) or one-way measurements are being made.

The two-way system begins with a Deep Space Station frequency standard, exciter system, transmitter, and microwave system, which are used to generate an S-band (2-GHz) radio frequency carrier signal.[1] This signal can be modulated by command or ranging signals and is then amplified to a power level of 20 kW or more (Fig. 4-1). The resulting radio wave is then collimated and directed toward the spacecraft by a 34- or 64-m parabolic antenna and an associated pointing system. The radio wave from the earth is intercepted and focused by an antenna on the spacecraft.

The spacecraft receiver uses a phase-locked loop to lock on to and track the uplink carrier, producing a reference signal whose phase is coherent with the uplink carrier (Fig. 4-2). This reference signal is used to demodulate ranging and command signals from the uplink carrier.

The ranging signal demodulated by the receiver is passed through a bandpass filter with a typical bandwidth of 1.5 MHz. This baseband ranging signal is controlled in amplitude to provide a constant power of the signal and uplink noise combined. Additionally, sine wave signals in the range 0.5 to 20 MHz are generated for differential one-way ranging (DOR). These DOR signals are obtained by frequency division from either the received carrier or from another oscillator on board the spacecraft. Both baseband ranging signals, the ranging signal from the uplink and the DOR signal, can be used to modulate the downlink carrier.

The downlink exciters coherently multiply the receiver reference frequency to obtain S- and X-band carrier signals that are slightly higher in frequency than the received carrier signals, by the ratio 240/221 for S-band and 880/221 for X-band. Alternatively, fixed frequencies unrelated to the uplink carrier, from either an auxiliary crystal oscillator or an ultrastable crystal oscillator, can

[1] Or possibly also X-band (8 GHz) on some spacecraft.

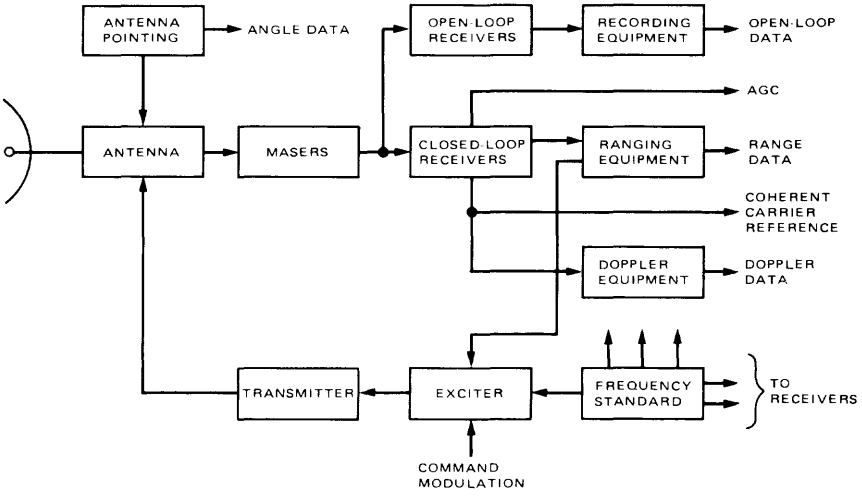


Fig. 4-1. Ground tracking system block diagram

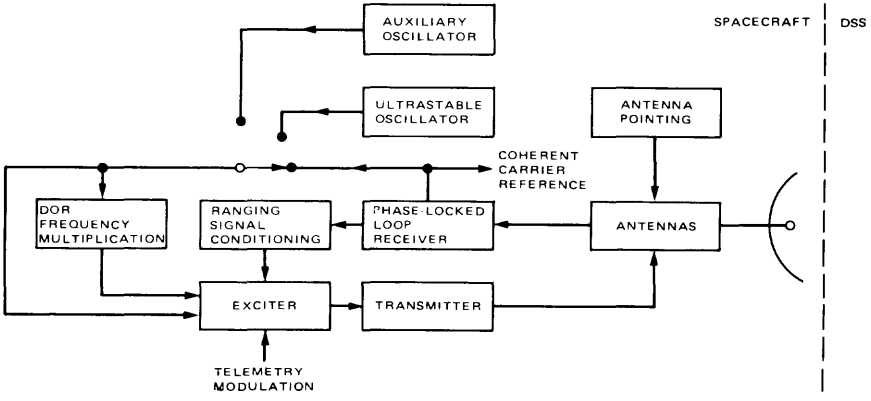


Fig. 4-2. Spacecraft tracking system block diagram

be used to generate the downlink carriers. The exciters also phase modulate the downlink carrier signals with ranging and DOR modulation. The phase-modulated signals are then amplified by a transmitter and collimated by an antenna for transmission back to earth.

The radio waves from the spacecraft are intercepted and focused by a 34- or 64-m antenna on the earth and are immediately amplified by a low-noise maser amplifier. The amplified signal can then be received in one of two distinct manners: either open loop, in which a programmed local oscillator is used to translate the received signal to an intermediate frequency for subsequent recording, or closed loop, in which a phase-locked loop is used to produce a reference signal coherent with the received signal. In the latter case, the reference signal is used to demodulate telemetry and ranging signals and estimate the difference in frequency between the reference signal and the received signal (doppler frequency). Separate equipment can be used to estimate the time delay between the transmitted and received ranging modulation (ranging) or the relative time delay between two stations (DOR).

The one-way tracking system begins with the spacecraft oscillator and proceeds through the ground system in the same manner as the two-way system, with the exception that round-trip ranging cannot be performed.

The tracking system produces the following outputs:

(1) Coherent carrier reference signals.

(2) Frequency of received signal (doppler).

(3) Round-trip time delay to spacecraft (range).

(4) Direction of received signal (angles).

(5) Intensity of received signal (AGC).

(6) Recordings of waveform and spectrum of received signal (open-loop recording).

4.1.3 Tracking Technique

The following paragraphs describe the various ways in which the telecommunication link can be configured for radio tracking. An often used equivalent description of tracking mode specifies the type of measurement being made. For instance, guided by the list of outputs above, the terms are "coherent carrier tracking," "doppler tracking," "range tracking," "angle tracking," etc.

(1) *One-way tracking.* The spacecraft generates a radio signal (a downlink) from an onboard oscillator that is then received and tracked by an earth-based receiver. No transmission is made to the spacecraft.

(2) *Two-way tracking.* An earth-based transmitter generates a radio signal (an uplink) that is tracked by the spacecraft receiver. The spacecraft transmits a downlink that is coherent with the uplink, which is then tracked by an earth-based receiver using the same frequency standard and antenna as the original earth-based transmitter (Fig. 4-3).

(3) *Three-way tracking.* Two-way tracking is performed by one earth-based station, while another station using either a different frequency standard or a different antenna tracks the downlink.

(4) *Two-way noncoherent tracking.* Identical to two-way tracking, except that the spacecraft transmits a downlink that is not coherent with the uplink, but is derived instead from a free-running onboard oscillator, usually an ultrastable crystal oscillator.

4.2 Uses of Tracking Data

4.2.1 Navigation

4.2.1.1 Doppler frequency. Consider the distance between a spacecraft and an earth-based tracking station. Under the assumption that the distance to the spacecraft is large compared to the radius of the earth, Hamilton and Melbourne have shown that this distance obeys the formula [4-1, 4-2]

$$\dot{\rho}(t) = \dot{r}(t) + \omega r_s \cos \delta_0 \sin (\omega t + \phi + \lambda + \alpha_0) \qquad (4.2\text{-}1)$$

where

$\dot{\rho}(t)$ = time derivative of distance between spacecraft and tracking station

$\dot{r}(t)$ = time derivative of distance between spacecraft and center of earth

ω = angular rotation rate of earth, rad/sec

r_s = distance of tracking station from earth rotation axis

δ_0 = declination of spacecraft

t = time past epoch

ϕ = phase angle that depends on epoch

λ = longitude of tracking station

α_0 = right ascension of spacecraft

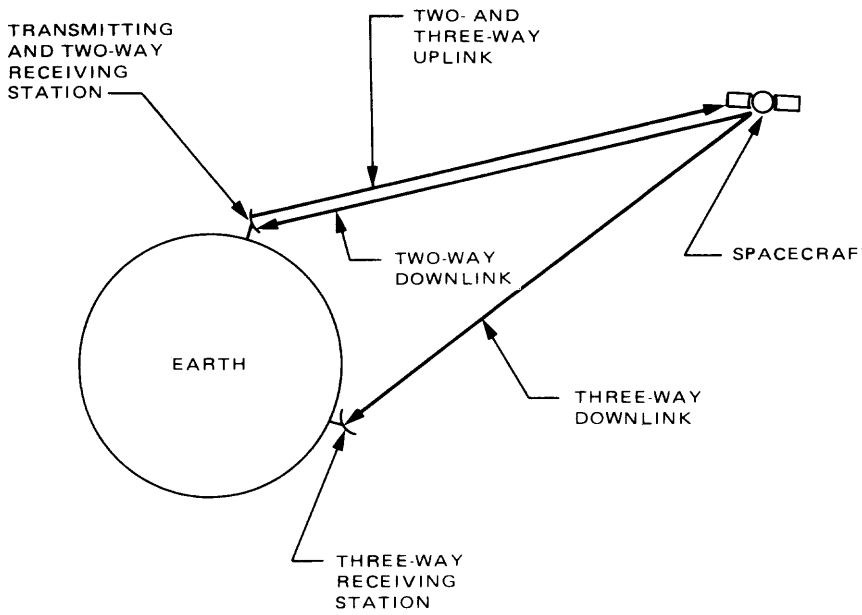


Fig. 4-3. Transmitting and receiving configuration for two- and three-way data

The coordinates to which the Hamilton-Melbourne equation applies are described in Fig. 4-4.

It is possible to measure $\dot{\rho}(t)$ by observing the change in frequency introduced on a radio link between the spacecraft and the tracking station. This change in frequency arises as a consequence of the Lorentz transformation in the special theory of relativity. In the simplest possible case the spacecraft would transmit a constant known frequency f_T, and the tracking station would receive a frequency [4-3]

$$f_R = \gamma f_T (1 - \beta \cos \theta) \tag{4.2-2}$$

where

f_R = received frequency

f_T = transmitted frequency

v = magnitude of spacecraft velocity

c = speed of light

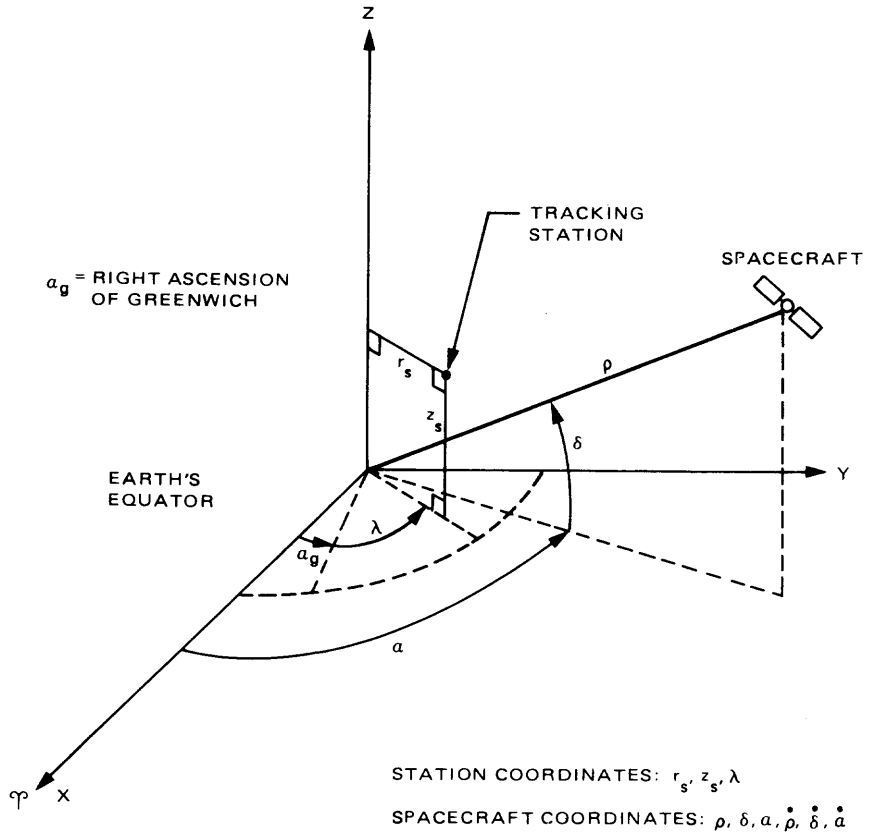


Fig. 4-4. Spacecraft and station coordinates

β = v/c

γ = $1/\sqrt{1-\beta^2}$

θ = angle between spacecraft velocity vector and vector from receiver to spacecraft

If the spacecraft motion is purely along the line of sight, then the received frequency is

$$f_R = \sqrt{\frac{c-\dot{\rho}}{c+\dot{\rho}}}\, f_T \cong (1-\frac{\dot{\rho}}{c}) f_T \qquad (4.2\text{-}3)$$

if $\dot{\rho} \ll c$. The one-way doppler frequency is then

$$f_{D1} \equiv f_R - f_T \cong -\frac{\dot{\rho}}{c} f_T \tag{4.2-4}$$

Substitution of (4.2-1) shows that the doppler frequency will undergo a sinusoidal variation due to the rotation of the earth, about a mean which varies slowly due to the relative motion between the spacecraft and the earth (Fig. 4-5). Whenever the station location and the transmitted frequency are known, the declination of the spacecraft may be found from the amplitude of the sinusoidal variation, and the right ascension from the phase of the sinusoid. The geocentric velocity may be found from the mean value, and the geocentric acceleration from changes in the mean value of the sinusoid [4-4, 4-5].

Practical measurements of doppler frequency almost always make use of two-way or three-way techniques, in which an earth-based reference oscillator transmits a signal to the spacecraft, where it is coherently multiplied in frequency, retransmitted, and subsequently received at the original transmitter (two way) or another earth station (three way). These two round-trip techniques are preferred because the most stable reference oscillators, e.g., a hydrogen maser, are too massive for spacecraft use. Then (4.2-2) must be modified for practical

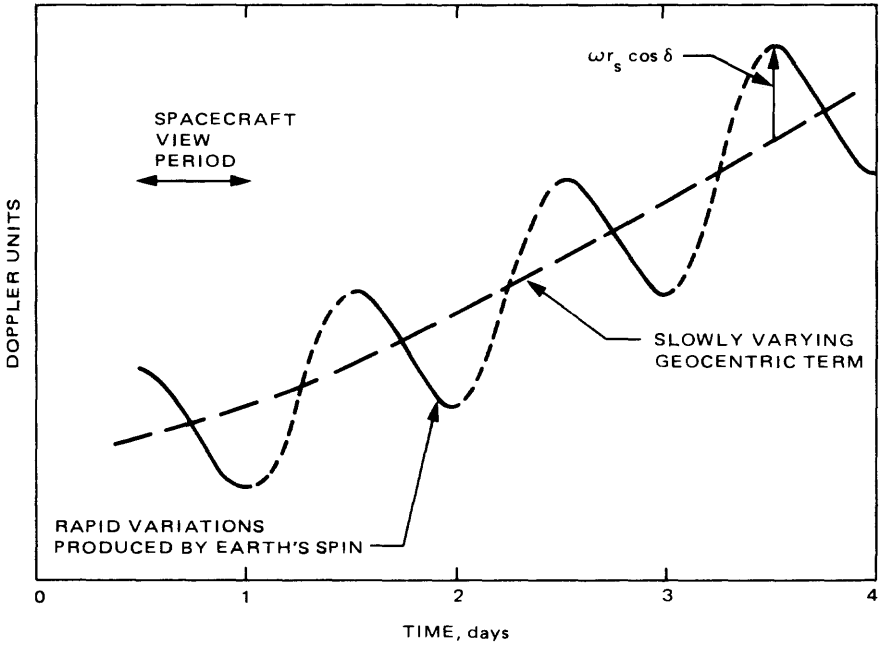


Fig. 4-5. Schematic illustration of doppler observable

measurements to take into account several other effects. These are the change in radial velocity between the tracking station and spacecraft during the time of flight of the radio signal, the spacecraft coherent frequency multiplication, general relativistic effects, and errors in frequency introduced by hardware and refractive media. The first three items are discussed in detail in [4-6]; the observation errors are discussed in Section 4.4.

It is possible to determine the orbit of a spacecraft inside the solar system using measurements of doppler frequency alone. Three parameters are needed to describe the position, and three parameters are needed to describe the velocity of the spacecraft, giving a total of six independent parameters that must be obtained from observations of the doppler frequency. A single day's observation can provide the right ascension, declination, and geocentric radial velocity; differencing the results with a second day's observation indirectly gives the velocity in right ascension and declination and the geocentric radial acceleration. Thus, we have found within a short time six parameters describing the location and motion of the spacecraft. However, only five of these parameters relate directly to the parameters describing position and velocity. The radial distance to the spacecraft has not been measured directly, but remains hidden in a combination of the radial acceleration with the other observed parameters.

In this case a model of the dynamics of the solar system is constructed, and the equations of motion for the spacecraft are integrated forward from some instant when the spacecraft position and velocity are known (e.g., immediately after launch) to obtain a predicted trajectory. Then the actual measurements are compared to the measurements that would have been obtained had the spacecraft followed the predicted trajectory, and the parameters of the original model are adjusted to minimize the weighted squares of the differences between the actual and predicted measurements. The radial distance to the spacecraft can then be determined from the model trajectory. This process is termed orbit determination and is usually repeated frequently for any one spacecraft (see Fig. 4-6). The accuracy of orbit determination depends on the accuracy of the measurements, the geometric characteristics of the trajectory, the time span over which the observations are made, and the accuracy of earlier knowledge of the model parameters.

4.2.1.2 Ranging. As described previously, it is possible to determine the position of a spacecraft inside the solar system using only measurements of doppler frequency and the equations of motion of the spacecraft. However, in the presence of noise and limited spans of data, and in some geometries, the position can be determined much more accurately if an independent measurement of the distance to the spacecraft is made [4-2, Table 1].

Furthermore, the velocity given in (4.2-1) has approximately the same sinusoidal component for all declinations with cosines near unity, making it difficult to measure declinations less than 5 deg [4-7]. In this case it is desirable to

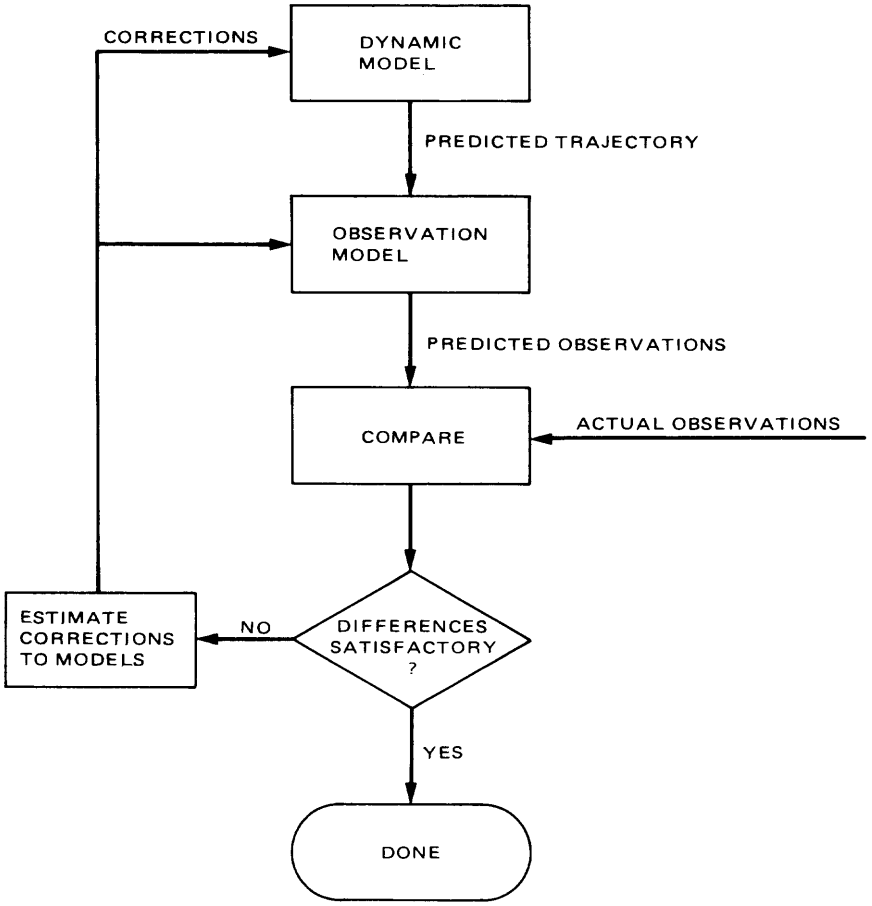


Fig. 4-6. Orbit determination process

measure the difference in range from two stations simultaneously or near-simultaneously (Figs. 4-7 and 4-8). For ideally located stations this would give

$$\rho_2 - \rho_1 = D \sin \delta \tag{4.2-5}$$

where

ρ_i = distance between station i and spacecraft

D = distance between stations perpendicular to the earth's equator

δ = declination of spacecraft

which is much less sensitive to errors near zero declination.

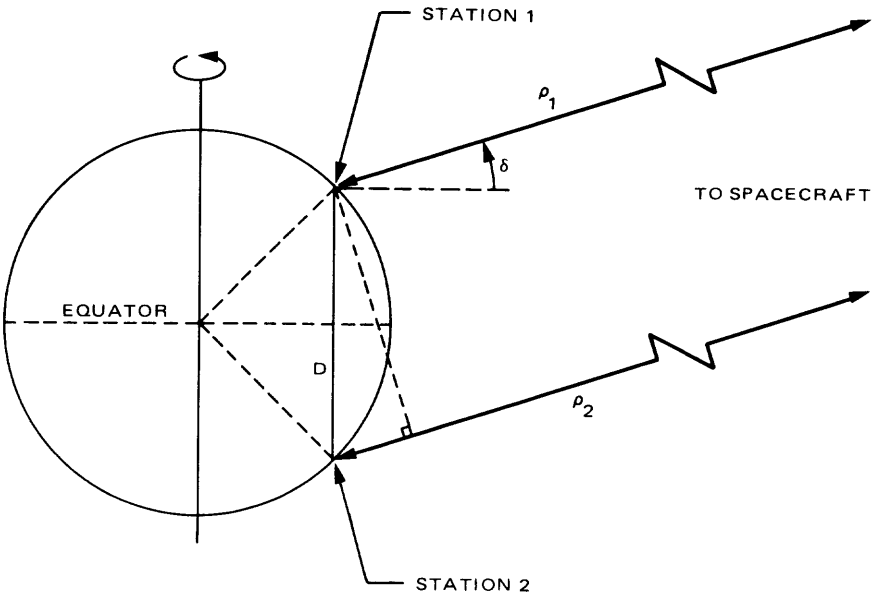


Fig. 4-7. Simultaneous ranging from two stations

The distance between a spacecraft and a tracking station may be measured by impressing modulation on the radio signal from the station to the spacecraft. If the modulation is retransmitted by the spacecraft and received on the ground, it will have been delayed by a time equal to the time of flight of the radio signal. For instance, suppose that a continuous sinusoidal modulation $A \cos \omega t$ is placed on the carrier to the spacecraft. The modulation subsequently received at time t will be $B \cos \omega \, [t - (\rho_u + \rho_d)/v_g]$, and the cross-correlation between the transmitted and received modulation will be

$$\eta \equiv \frac{1}{T} \int_{t_0}^{t_0+T} A \cos \omega t \, B \cos \omega \left(t - \frac{\rho_u + \rho_d}{v_g} \right) dt \qquad (4.2\text{-}6)$$

$$\cong \frac{AB}{2} \cos \frac{\omega}{v_g} (\rho_u + \rho_d), \text{ if } T \gg \frac{1}{\omega} \qquad (4.2\text{-}7)$$

where

η = cross-correlation between transmitted and received waveforms

A,B = amplitude of transmitted and detected modulation, respectively

ω = angular frequency of modulation

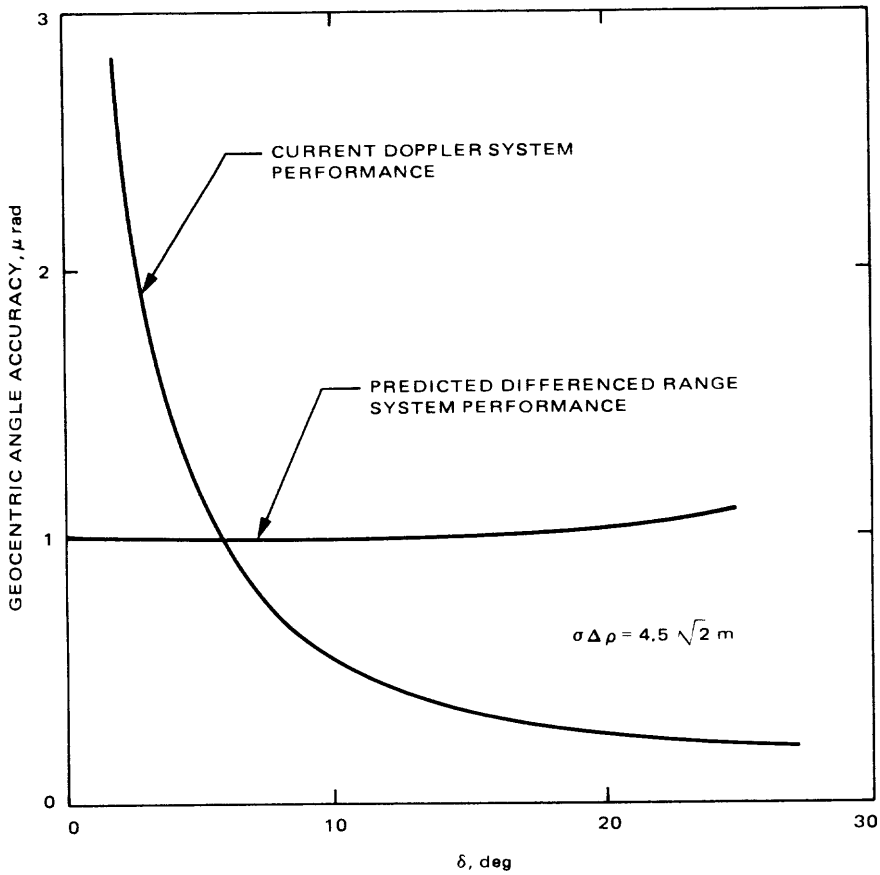


Fig. 4-8. Accuracy performance of doppler and ranging systems versus nominal probe declination

T = correlation interval

t_0 = start of correlation interval

ρ_u = distance to spacecraft on uplink

ρ_d = distance to spacecraft on downlink

v_g = velocity of propagation of modulation

The round-trip distance may then be found from

$$\rho_u + \rho_d = \frac{v_g}{\omega}\left[\cos^{-1}\frac{2\eta}{AB} + 2\pi N\right] \tag{4.2-8}$$

The ambiguity factor N, which arises in any modulation of finite period, must in general be determined by independent means, e.g., by extrapolation from previous knowledge of the spacecraft trajectory. There is an obvious tradeoff between how accurately N can be predicted and the precision with which the round-trip distance can be measured; higher modulation frequencies can yield more precise determination of distance, but place more stringent demands on the fractional error of N. This tradeoff is usually circumvented by using a succession of frequencies, in which the highest frequency determines the precision of the measurement, the lowest frequency determines the required accuracy of prediction of N, and the intervening frequencies provide enough information to determine the precise ambiguity of the highest frequency [4-8].

It should be noted that the time delays introduced by hardware must be removed so that the measured time delay corresponds to a distance, and that certainly some assumption must be made about the speed of propagation of the ranging modulation. This calibration process is described in detail in Section 4.4.

4.2.1.3 Very long baseline interferometry. This subsection, as edited, is adapted extensively from an article by D. W. Curkendall [4-7].

In very long baseline interferometry (VLBI), each of two widely spaced antennas observes a single (broad band) radio source, e.g., a quasar, simultaneously recording the received signal over a specified frequency interval. The recordings of the received voltage (Fig. 4-9) are digitized at the one-bit level, and timing information is added so that the recordings may be cross-correlated later when brought to a central site. The (expected) cross-correlation function can be shown to be approximately [4-9]

$$E(R(t, \Delta\tau)) = \frac{\sin \pi W \Delta\tau}{\pi W \Delta\tau} \cos \phi_1(t) \qquad (4.2\text{-}9)$$

where

$\Delta\tau = \tau_g - \tau_m$

τ_g = geometric delay as shown in Fig. 4-8

τ_m = *a priori* delay estimate inserted to bring the correlation function to near its maximum during data processing

W = bandwidth of recorded signal

$\phi_1 = \omega_1 \Delta\tau$

ω_1 = frequency at the center of the bandpass

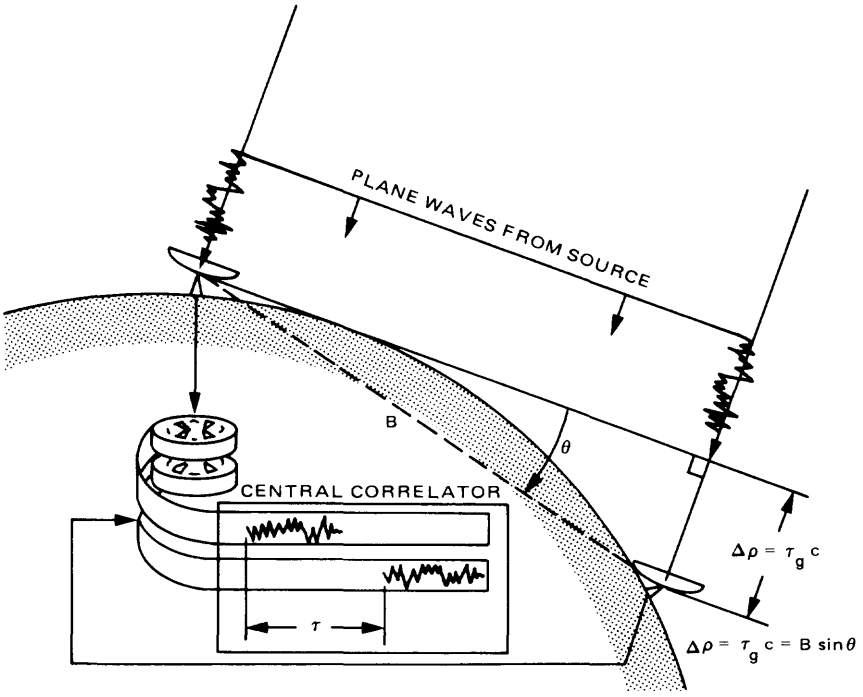


Fig. 4-9. Measuring differenced time delay with VLBI

Using typical values of 2300 MHz and 2 MHz as the RF and recorded bandwidth, respectively, this correlation function goes through one complete cycle for every change in $\Delta\tau$ equal to the period of the RF frequency (<5.0 nsec). In addition it also manifests a $(\sin x)/x$ characteristic envelope reaching its first null at 0.5 μsec delay. These two components are called, respectively, "fringes" and the "delay resolution function." For these same typical values, τ_g (or equivalently $\Delta\rho/c$, Fig. 4-8) can be measured directly, by adjusting τ_m so as to maximize the delay resolution function, to an order of 10-nsec precision (3 m in one-way distance). More powerful measurements can be obtained, however, in each of two conceptually different ways:

(1) Observation of the source continuously over the common visibility period of the two stations (approximately 4 hr on the baseline afforded by the DSN for sources near the ecliptic) produces a continuous record of the phase of the fringes versus time. The record thus obtained will contain a diurnal sinusoidal modulation term due to the earth's rotation, whose phase and amplitude are parametric in the source location and baseline parameters. This is exactly analogous to the single-station coherent doppler tracking, except that the equatorial baseline projection and longitude play the roles of the distance off the spin axis, r_s, and

station longitude, λ, respectively. The differential frequency of the two clocks replaces the geocentric velocity term observed by the coherent data.

(2) Observation of the source at a second center frequency, ω_2, produces a second measurement of the fringe phase, ϕ_2, at a single instant of time. Because

$$\frac{\partial \phi}{\partial \omega} = \tau_g \cong \frac{\phi_2 - \phi_1}{\omega_2 - \omega_1}$$

a direct measurement of τ_g can be obtained in the short time required to achieve a high S/N for the ϕ measurements (typically 10 min). This is the "bandwidth synthesis" technique, so-called because large effective bandwidths can be obtained without the need for commensurate high recording rates [4-9, 4-10]. The geometry is exactly as for the differenced range measurement already discussed (cf Figs. 4-6 and 4-8); the τ_g measurement obtained can be used directly to estimate either the baseline projection or source location.

Equivalently τ_g may be determined by cross-correlating a modulated spacecraft carrier received at the two VLBI antennas, and applying (4.2-7) and (4.2-8) with the quantity $(\rho_u + \rho_d)/v_g$ replaced by τ_g. Repetition of this technique, alternating between a spacecraft and a quasar, reduces several error sources (notably clock and media effects), and is known as differential one-way ranging (DOR). Again, the geometry is the same as for differenced range, but DOR is usually only used to estimate the spacecraft location; quasar location and baseline projection are assumed known.

With spanned bandwidths, $\omega_2 - \omega_1$, on the order of 40 MHz, the precision of the measurement can easily be brought to the centimeter level; its accuracy is dominated by other effects such as clock performance and systematic calibration errors.

When estimating source locations with this method, a second baseline is usually employed for the second component of position. An effective combination for the two baselines is to have a large polar component associated with the first and a large equatorial projection associated with the second so that they can produce largely uncorrelated estimates of right ascension and declination.

Thus, these two methods, often referred to as narrow- and wide-band VLBI, respectively, have a one-to-one correspondence with the two coherent modes,

doppler and differenced ranging, normally used for spacecraft tracking. Their normal applications complement each other in a bootstrap process. VLBI is usually employed to estimate the station baselines; precise source coordinates are needed to enable this. Coherent tracking is normally used to estimate the spacecraft coordinates; precision station locations are needed for this task.

4.2.2 Radio Science

Each of the parameters of an electromagnetic wave that may be measured, namely, its frequency, phase and group delay, intensity, and polarization, is affected by the properties of the media traversed by the wave. Because the radio waves to and from the spacecraft traverse so many media, these signals have proven to be a rich and complex source of information about the solar system. This is especially true when measurements of several electromagnetic parameters are combined together or combined with some other complementary observation.

The following sections describe the most important of the uses to which tracking data have been put so far. New uses are constantly being devised; therefore, the reader is also encouraged to take advantage of the scientific literature. Comprehensive descriptions of the uses of spacecraft radio signals are contained in [4-11], [4-12], and [4-13]. A more extensive guide to references appears in the appendix.

4.2.2.1 Mass and gravity fields of planets and satellites. The motion of a spacecraft over almost all of its journey through the solar system is governed by the gravitational fields it encounters. The motion can, of course, be discerned through doppler, ranging, and VLBI measurements, so the spacecraft becomes a test probe for the gravity fields through which it moves. The acceleration induced during planetary encounters or orbits is usually large enough so that the total mass of a planet can be determined, as well as several moments of the mass distribution. This information can be used to limit the class of plausible interior models of the planet, including its chemical composition. Due to their smaller effects, it has only been possible to measure the total mass of satellites, although some higher moments might be determined for the largest satellites in the future.

4.2.2.2 Pressure, temperature, and composition of atmosphere. Measurements of the received frequency of the radio signals from a spacecraft during planetary occultation can provide information on the angle of refraction in a planetary atmosphere. Knowledge of this angle and the spacecraft trajectory makes it possible to estimate the index of refraction as a function of height, which in turn can determine the relative temperature and pressure as a function of height. Knowledge of atmospheric composition can make these profiles absolute.

The intensity of radio signals passing through an atmosphere provides an independent measurement of the index of refraction as a function of height, in

this case because differential refraction leads to defocusing of the signal collimated by the spacecraft antenna. In addition, the loss due to microwave absorption or scattering by cloud condensates, vapors, or primary atmospheric constituents at low altitudes can provide information on the location, density, and other characteristics of the absorbing or scattering material.

4.2.2.3 Electron density of ionospheres, solar corona, and the interplanetary medium. The velocity of propagation of radio waves in a tenuous plasma depends on both the average density of electrons in the plasma and, inversely, on the frequency of the wave squared (4.4-6, 4.4-7). Knowing the distance to the spacecraft, and measuring the difference in arrival time of signals at two frequencies, one can infer the average electron density along the ray path. If this difference is observed over a period of time, it is possible to infer the density distribution of plasma about the solar system or within the solar corona. Similarly, the difference in doppler frequency during a planetary occultation can be used to infer the distribution of plasma in an ionosphere.

4.2.2.4 Experiments in general relativity. The gravitational field of the sun induces an extra time delay on radio signals passing near the sun. The photons comprising the radio waves to and from the spacecraft are deflected in a manner similar to the deflection of massive particles in a hyperbolic trajectory around the sun, although the effect has a different magnitude. The effect is largest for ray paths that graze the solar limb, and measurements of the round-trip signal delay to a spacecraft have provided the most sensitive test so far of ray bending by the sun's mass.

An important experiment in general relativity that has not yet been performed, but which is in development at the time of this writing, concerns the detection of gravitational radiation. Hypothetical bursts of gravitational radiation, arising possibly from the collapse of galactic nuclei into extremely massive black holes, would induce a characteristic three-pulse response in the measurements of two-way doppler frequency. This response is due to small changes in ground and spacecraft oscillator frequency caused by a temporary general relativistic red shift within the passing gravitational wave. If sufficient frequency stability of spacecraft radio signals can be obtained, it is possible that the characteristic signature of gravitational radiation may be found in two-way doppler frequency.

Finally, it is also possible to measure the general relativistic red shift of spacecraft oscillators within the gravity field of Jupiter and Saturn to an accuracy of approximately 1%. While this is not competitive with experiments performed on earth-orbiting spacecraft, measurements at Jupiter and Saturn constitute independent and simpler tests.

4.3 Measurement Techniques

Measurements with the tracking system are made with several hardware assemblies which are described below: the Metric Data Assembly, the Ranging Demodulator Assemblies, the Planetary Ranging Assembly, and the Radio Science Subsystem (Fig. 4-10).

4.3.1 Metric Data Assembly

The Metric Data Assembly (MDA) measures the frequency of the doppler signal obtained from the doppler extractor of the Receiver-Exciter Subsystem (Fig. 4-11). Additionally, the Metric Data Assembly controls many of the ground elements of the tracking system and provides radiometric data transfer to and from a ground station, although these functions are not discussed here.

The measurement of doppler frequency is performed by doppler counters and resolvers within the MDA. The phase $\theta(t)$ of the doppler signal is continuously measured by counting positive zero crossings of $\sin \theta(t)$, and the count is sampled at fixed time intervals T_i ranging from 0.1 to 60 sec (Fig. 4-12). The resolver further refines the phase reading by measuring the elapsed time between the latest sample point and the next following positive zero crossing. The time interval is measured by a clock of a relatively high frequency[2] (100 MHz) so that only a small quantization error results.

The doppler signal obtained from the Receiver-Exciter Subsystem is a *biased* doppler signal. That is, instead of from (4.2-4), a one-way doppler signal is obtained from

$$f'_D = f_R - f_T + f_B = f_D + f_B \tag{4.3-1}$$

where

f'_D = biased doppler frequency

f_D = unbiased doppler frequency

f_R = earth-received carrier frequency

f_T = spacecraft transmitted carrier frequency

f_B = doppler bias frequency (±1 to ±5 MHz)

The bias frequency is introduced to distinguish between the positive and negative doppler frequencies caused by approach or recession of a spacecraft.

[2] 500 MHz in the future.

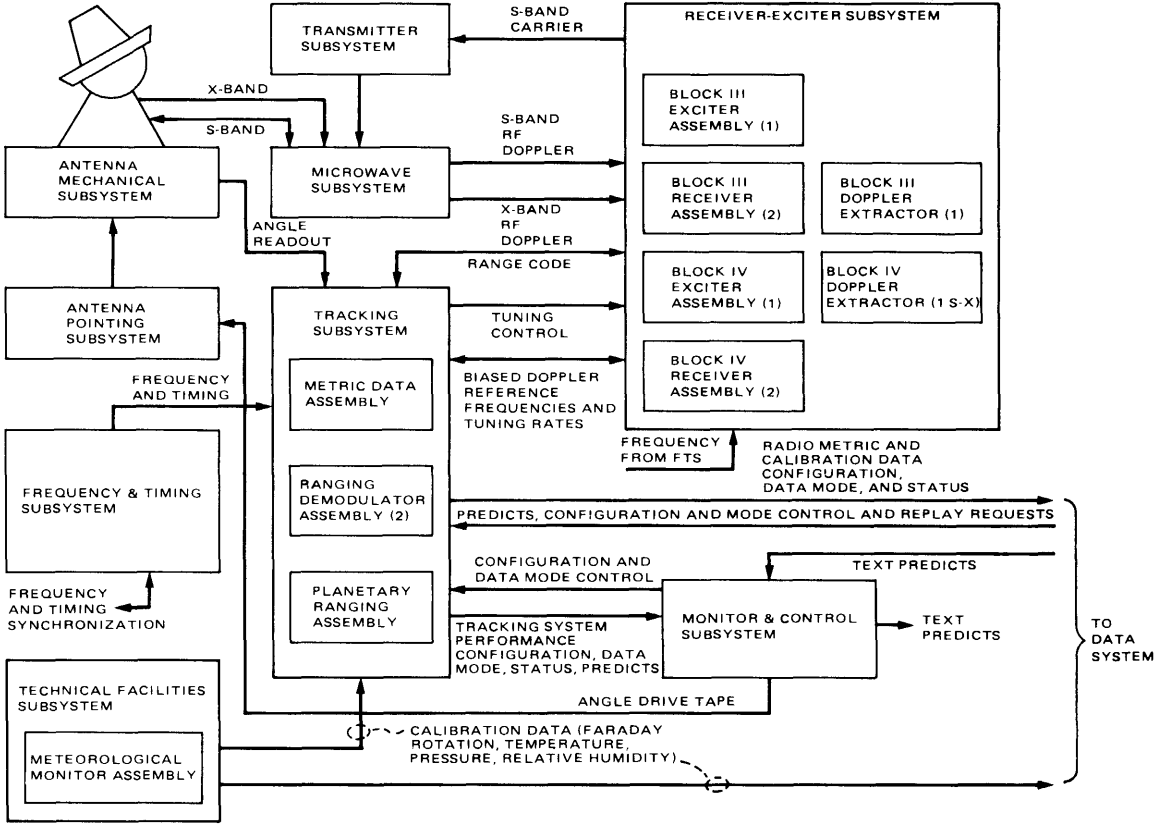


Fig. 4-10. Ground tracking system for navigation

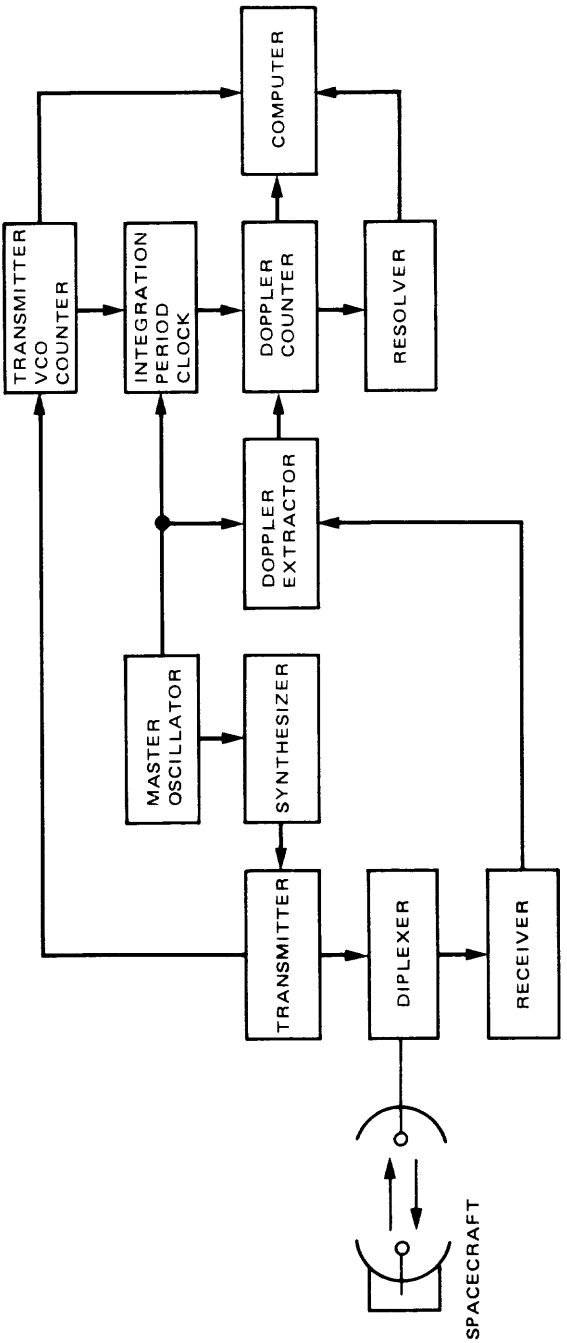


Fig. 4-11. Doppler measurement technique

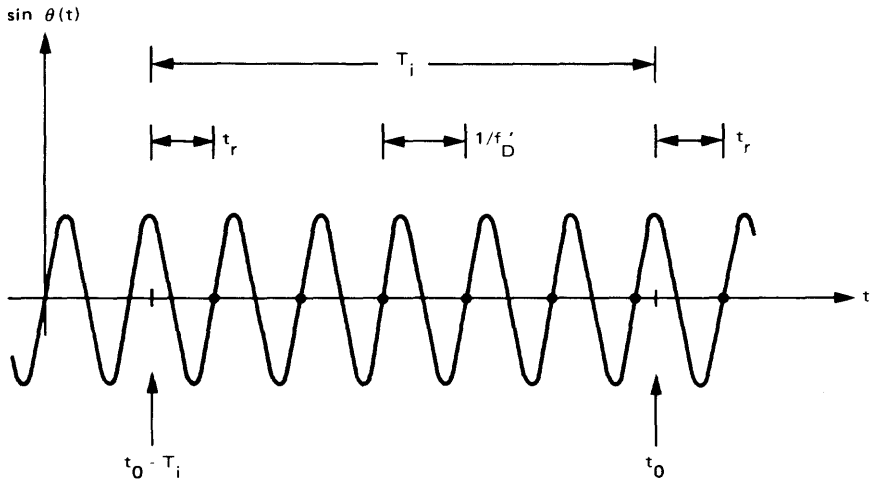


Fig. 4-12. Doppler counter and resolver parameters

The count of zero crossings performed by the doppler counters and the time measured by the resolvers translate directly into an increment of phase during the sampling interval given by

$$\Delta\theta = 2\pi \left[N(t_0) - N(t_0 - T_i) - 10^{-8} f'_D \left(n(t) - n(t_0 - T_i)\right)\right] \qquad (4.3\text{-}2)$$

where

$\Delta\theta$ = phase increment

t_0 = time at end of sampling interval

T_i = length of sampling interval

$N(t)$ = doppler count at time t

$n(t)$ = $t_r(t) \times 10^8$ = doppler resolver count at time t

$t_r(t)$ = time measured by the resolver, sec

Since frequency is the time derivative of phase, the average biased doppler frequency can be obtained from

$$f'_D = \frac{\Delta\theta}{2\pi T_i} \qquad (4.3\text{-}3)$$

and the average unbiased doppler frequency from

$$f_D = f'_D - f_B \tag{4.3-4}$$

4.3.2 Ranging Demodulation Assemblies

Two Ranging Demodulator Assemblies (RDAs), one for S-band and one for X-band, cross-correlate the received ranging signals with a ranging signal from the Planetary Ranging Assembly (Figs. 4-13, 4-14, and 4-15). The ranging signal from the Planetary Ranging Assembly is a square-wave model of the expected received ranging signal, and is provided in two versions differing in phase by 90 deg. An intermediate frequency consisting of a 10-MHz carrier phase modulated by the received ranging signal is provided by the Receiver-Exciter Subsystem. This intermediate frequency is passed through a 1-MHz single-sided bandwidth band-pass filter centered at 10 MHz and mixed with the model ranging signals. The output of the mixers is a 10-MHz sine wave whose phase depends on the relative phase between the model and received ranging signal. This phase difference is then detected and averaged by mixing with a 10-MHz reference signal and passing the signal through a low-pass filter to yield the final correlation voltages [4-14].

4.3.3 Planetary Ranging Assembly

The Planetary Ranging Assembly (PRA) measures round-trip group delay from the correlation voltages provided by the Ranging Demodulation Assemblies (Fig. 4-13). The PRA generates a square-wave ranging signal, which is phase-modulated on the uplink carrier by the Receiver-Exciter Subsystem, and also generates a square-wave model of the ranging signal for use by the Ranging Demodulation Assemblies. The PRA then estimates the time delay between the transmitted and received ranging signal, the signal-to-noise ratio of the received ranging signal, and the differenced range versus integrated doppler [4-14].

The transmitted ranging signal is generated by coherent frequency division from the frequency reference used to generate the uplink carrier. Usually a code consisting of a succession of frequencies is generated, starting with approximately 1 MHz and decreasing by factors of two to as low as approximately 1 Hz. This is done in order to provide a precise determination of time delay up to an unavoidable ambiguity of 1 μsec, and subsequently increase the size of the ambiguity to a value which could easily be determined by other means.

The generation of the model ranging signal is similar to the generation of the transmitted signal, although two factors complicate this process. First of all, the received ranging signal has been shifted in frequency by coherent multiplication in the spacecraft, by the doppler effect, and to a much smaller extent by intervening refractive media. Obviously, it would be useless to attempt to measure

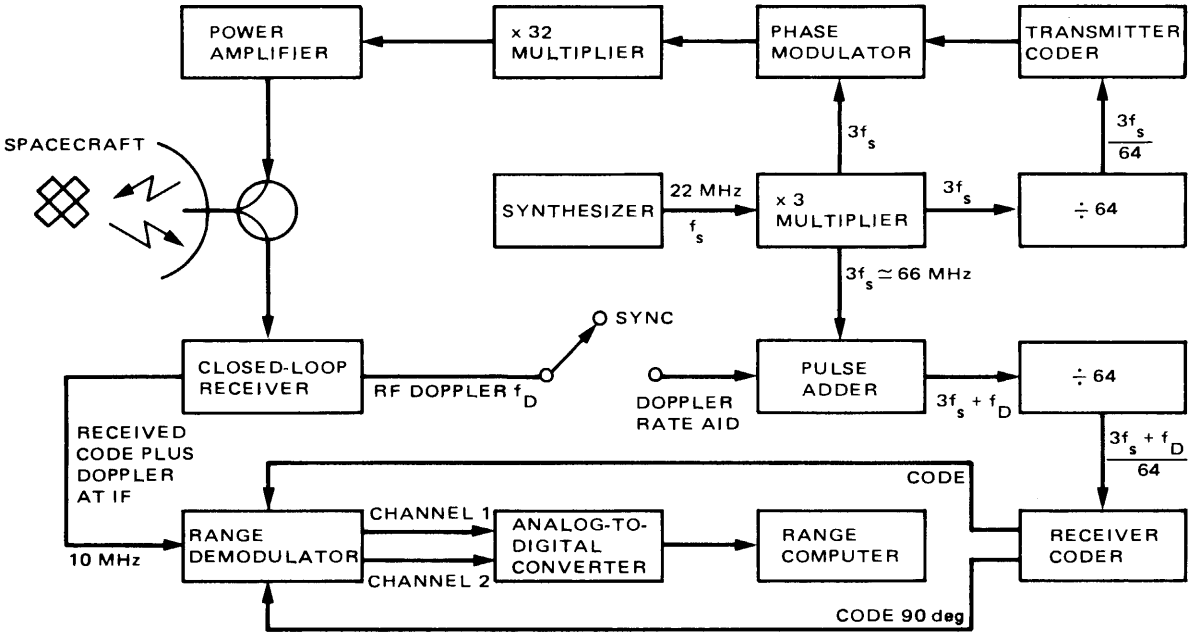


Fig. 4-13. Range measurement technique

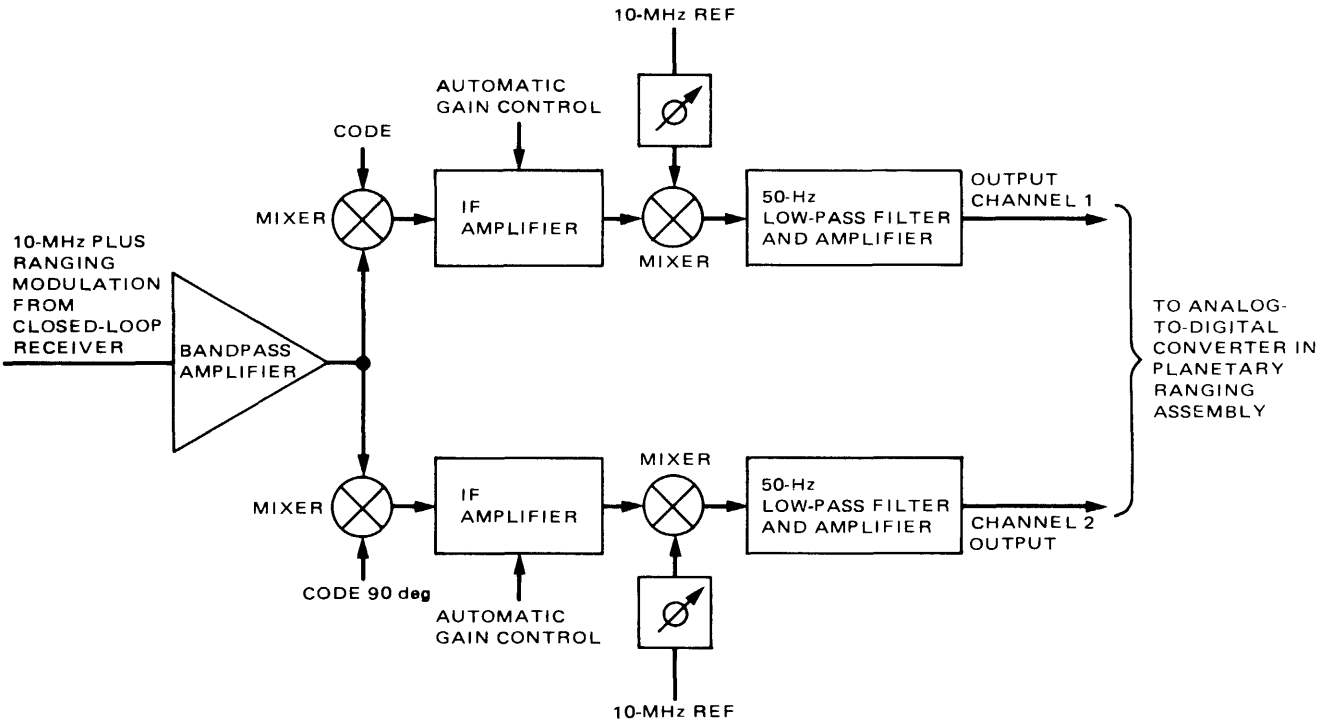


Fig. 4-14. Ranging Demodulator Assembly

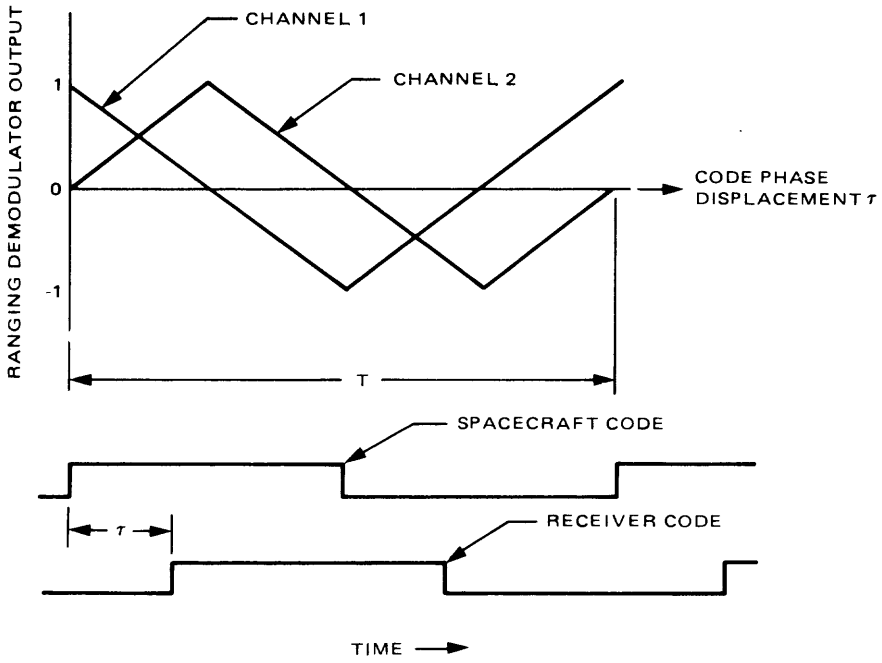


Fig. 4-15. Ranging Demodulator output

range by cross-correlating two signals of different frequencies; therefore, a process known as "doppler rate aiding" is used. In doppler rate aiding the reference frequency used to generate the model code is modified to take account of the downlink frequency shift. The reference frequency is identical to that used for the transmitted code until a designated time T_0, when the doppler signal from the Receiver-Exciter Subsystem is mixed with the reference signal to obtain a frequency equal to the reference plus the doppler frequency. Thereafter the model code is coherent with the received code and cross-correlation becomes meaningful.

Secondly, the received ranging signal necessarily is available for correlation only after travelling to and from the spacecraft, possibly as long as several hours after transmission. Thus, it is necessary to record the time of transmission of a range code and arrange the model code to coincide with the predicted arrival of the received code.

The time delay between the transmitted and received ranging signals and the noise on the received signal are estimated by sampling the outputs of the Ranging Demodulator Assemblies every 1/4 sec for a period of time and applying the following formulas (see also Fig. 4-16):

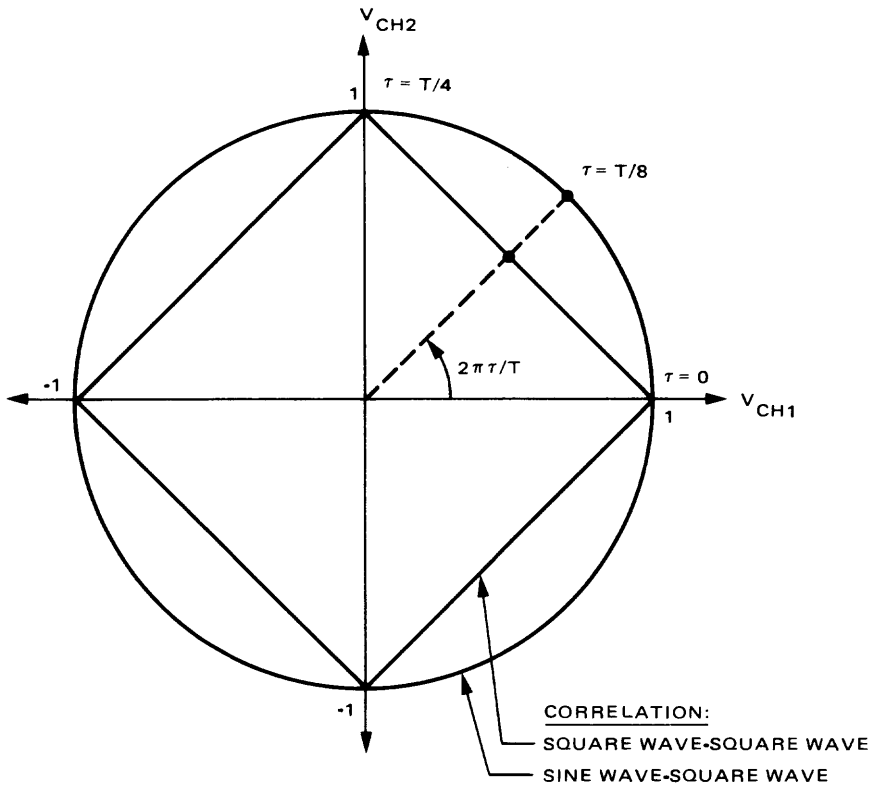


Fig. 4-16. Ranging Demodulator outputs with time delay as a parameter

For square-wave-square-wave correlation [4-14]:

$$\tau = \left(\frac{\sum_{t=0}^{t_i} V_{CH2}}{\left| \sum_{t=0}^{t_i} V_{CH1} \right| + \left| \sum_{t=0}^{t_i} V_{CH2} \right|} \right) \frac{T}{4}, \quad 0 < \tau < \frac{T}{4} \qquad (4.3\text{-}5)$$

$$\tau = \left(2 - \frac{\sum_{t=0}^{t_i} V_{CH2}}{\left| \sum_{t=0}^{t_i} V_{CH1} \right| + \left| \sum_{t=0}^{t_i} V_{CH2} \right|} \right) \frac{T}{4}, \quad \frac{T}{4} < \tau < \frac{3T}{4} \qquad (4.3\text{-}6)$$

$$\tau = \left(4 + \frac{\sum_{t=0}^{t_i} V_{CH2}}{\left| \sum_{t=0}^{t_i} V_{CH1} \right| + \left| \sum_{t=0}^{t_i} V_{CH2} \right|} \right) \frac{T}{4}, \quad \frac{3T}{4} < \tau < T \tag{4.3-7}$$

$$\frac{P_r}{N_0} = \text{Bandwidth} \cdot \frac{\left(\left| \sum_{t=0}^{t_i} V_{CH1} \right| + \left| \sum_{t=0}^{t_i} V_{CH2} \right| \right)^2}{\sigma_{CH1}^2 + \sigma_{CH2}^2} \tag{4.3-8}$$

For sine-wave-square-wave correlation [4-16]:

$$\tau = \frac{T}{2\pi} \tan^{-1} \frac{\sum_{t=0}^{t_i} V_{CH1}}{\sum_{t=0}^{t_i} V_{CH2}} \tag{4.3-9}$$

where t_i = correlation interval, and T = period of ranging signal.

In the presence of time varying dispersive media, for instance, when the ranging signal passes through the interplanetary plasma, the doppler rate aiding scheme described above will not be exact. This is because the phase delay measured by doppler will evolve differently from the group delay measured by range, causing the received code to drift from the model code. This drift between successive estimates of range is called differenced range versus integrated doppler (DRVID) and is measured continuously by the Planetary Ranging Assembly.

4.3.4 Radio Science Subsystem

Open-loop recording performed by the Radio Science Subsystem is usually considered to be part of the Radio Science System rather than the tracking system, but is included here because of its importance in the use of tracking data, for both navigation and radio science.

The open-loop recorders in the Radio Science Subsystem are used to record on tape spacecraft waveforms that have been translated to baseband by the open-loop receivers (Fig. 4-17). Open-loop recordings are used to determine the spectrum of a spacecraft signal in radio science experiments, or to determine the group delay of quasar or spacecraft radio signals in VLBI or DOR measurements. Additionally, these recordings are made for two orthogonal polarizations for each frequency band recorded, e.g., in right-circular and left-circular polarization. This makes it possible to determine the orientation of a received linear polarization (from the relative phase between the circular components) or the degree of depolarization of any initially known polarization (from the relative amplitude, or noise on the relative phase between the circular components).

Two types of recording devices are employed, which may be used independently or in combination. The Occultation Data Assembly (ODA) performs an 8- or 12-bit analog-to-digital conversion of the voltages applied to four input channels and records the results on a computer-compatible tape, up to a bandwidth of 10 kHz per channel (narrow bandwidth recording). The ODA also controls the local oscillator for open-loop receivers and records frequency control and measurement information along with the digitized voltages.

The Digital Recording Assembly (DRA) performs an 8-bit analog-to-digital conversion of one channel, or 1-bit conversion of up to 28 channels, up to a bandwidth of 2 MHz per channel. The DRA can be used to record a single wideband signal from an open-loop receiver (wide bandwidth recording), or the signals of four receivers which have been stacked in frequency by the ODA to form a single composite signal (medium bandwidth recording).

4.4 System Error Control

The tracking system capability allows measurement of the combined effects of spacecraft velocity and propagation media variation with an accuracy of better than 0.1 mm/sec rms for frequency observations integrated for 60 sec. Range measurements of the combined effects of spacecraft position and media time delay can be measured to an accuracy of better than 2.0 m rms, for range measurements integrated over 60 sec [4-17]. It is expected that VLBI measurements will yield angular accuracies of a few tens of milliarcseconds. Each of these numbers assumes signal-to-noise ratios typical of deep space missions.

Actual measurements of spacecraft position and velocity do not achieve these accuracies because of errors introduced in calibration for the media effects discussed in Subsections 4.4.1.1 and 4.4.1.2. Final determinations of velocity and range achieve accuracies of approximately 1 mm/sec and 30 m, respectively.

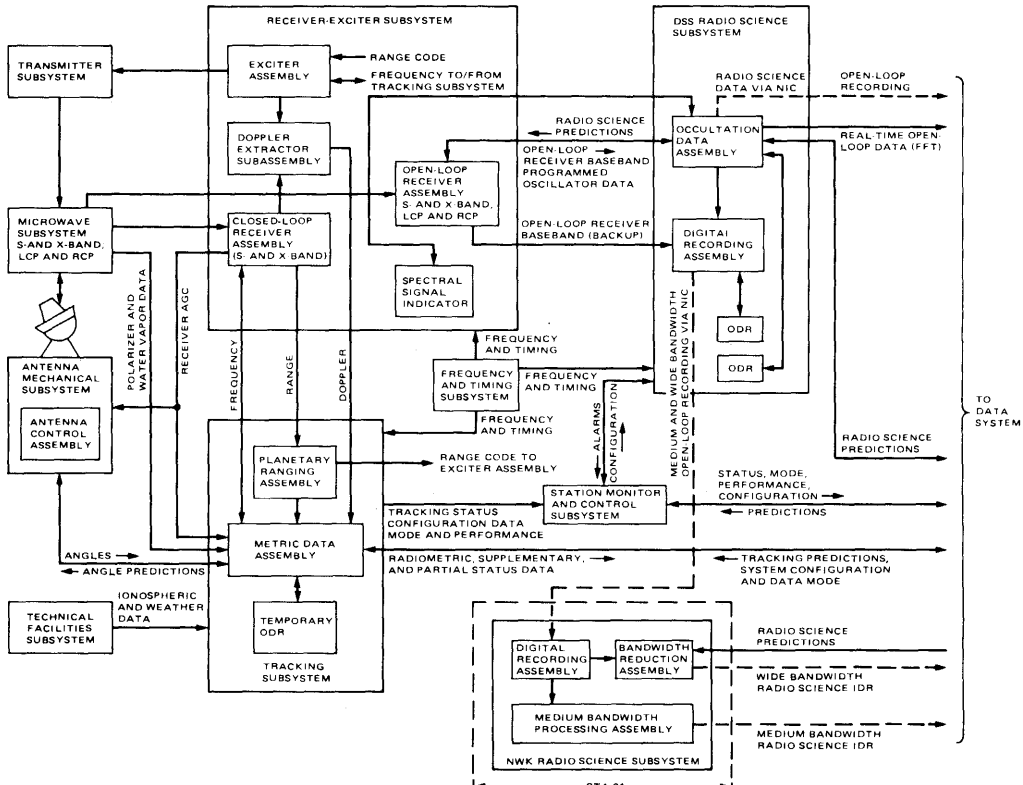


Fig. 4-17. Ground tracking system for radio science

4.4.1 Sources of Error

4.4.1.1 Troposphere. The effective path length of an electromagnetic wave in a refractive medium is

$$s = \int_{path} n \, dl \tag{4.4-1}$$

where

s = effective path length

n = index of refraction = $c/v < 1$

dl = element of length along the ray path

In general, the index of refraction is a function of frequency and location, but the frequency dependence may be neglected for spacecraft radio signals in the earth's troposphere. The dependence on location cannot be neglected, and through Snell's law leads to bending of the ray path. The increase in path length due to the troposphere ranges from 2 m at the zenith to 100 m at the horizon, and the angular deflection ranges from zero at the zenith to 1/2 deg at the horizon [4-18]. The difference in path length between S-band and X-band is less than 1 cm.

The refractivity of the troposphere at a given location was approximated by Smith and Weintraub [4-19] as

$$\eta = 77.6 \times 10^{-6}\left(\frac{P}{T} + 4810\,\frac{U}{T^2}\right) \tag{4.4-2}$$

where

η = refractivity

n = index of refraction = $1 + \eta$

P = pressure, mbar

T = temperature, K

U = water vapor partial pressure, mbar

The first term gives the refractivity of a dry atmosphere, and within 7 km of the earth accounts for 90% of the total refractivity. The second term gives the refractivity due to water vapor.

The total refractivity at the zenith can be related to the surface weather conditions, provided that the troposphere is assumed to be (1) in hydrostatic equilibrium, (2) an ideal gas, (3) spherically symmetric, and (4) subject to constant gravitational acceleration. Under these assumptions, the one-way range corrections at the zenith for the dry and wet components of the atmosphere are, respectively [4-18],

$$\Delta\rho_{\text{dry}}\,(90^\circ) \cong 2.27\,P_0 \tag{4.4-3}$$

$$\Delta\rho_{\text{wet}}\,(90^\circ) \cong D\left(3.733 \times 10^{-3}\,\frac{U_0}{T_0}\right) \tag{4.4-4}$$

where

$\Delta\rho(90^\circ)$ = increase in one-way range at zenith, mm

P_0 = surface pressure, mbar

T_0 = surface temperature, K

U_0 = surface water vapor partial pressure, mbar

D = troposphere vertical dimension, mm

The one-way range correction for an arbitrary elevation angle can be obtained by scaling $\Delta\rho(90^\circ)$ with the function $R(\theta)$, which is approximately $1/\sin\theta$. An empirical function for $R(\theta)$, which is accurate to within a few percent[3], is

$$R(\theta) \cong [\sin\theta + A/(\tan\theta + B)]^{-1} \tag{4.4-5}$$

where

θ = elevation angle

A = 1.43×10^{-3} (dry), 3.5×10^{-4} (wet)

B = 4.45×10^{-2} (dry), 1.7×10^{-2} (wet)

Systematic errors in phase and group delay measurements calibrated with the above equations arise to the extent that the assumptions mentioned previously fail to describe the troposphere. Systematic errors in one-way range at zenith of at least 2 cm for the dry component or 2 to 4 cm for the wet component have been noted [4-18]. The zenith errors are correspondingly increased by $R(\theta)$ at smaller elevation angles.

[3]Numerical tables of greater accuracy can be obtained.

4.4.1.2 Charged particles. It can be shown that the phase and group velocity of an electromagnetic wave within a tenuous plasma have the following form [4-20]:

$$v_p = c\left[1 + \frac{A}{f^2}\ N_e(x, t)\right] \tag{4.4-6}$$

$$v_g = c\left[1 - \frac{A}{f^2}\ N_e(x, t)\right] \tag{4.4-7}$$

where

v_p = phase velocity

v_g = group velocity

c = speed of light in vacuum

f = frequency of wave, Hz

$N_e(x,t)$ = number density of electrons, m^{-3}

A = $e^2/8\,\pi^2\,m_e = 40.3$ (mks)

e = charge on an electron, coulomb

m_e = mass of an electron, kg

Hence, the phase velocity is increased, and the group velocity is decreased by the same amount, in the presence of a tenuous plasma.

4.4.1.2.1 *Phase delay.* The signal transmitted to a spacecraft is

$$g(t) = \sin(2\pi f_u t) = \sin\theta_T(t) \tag{4.4-8}$$

The phase of the signal received at the spacecraft at time t is the same as that transmitted at time $t - T_u$:

$$\theta(t) = 2\pi f_u\,(t - T_u) \tag{4.4-9}$$

where

$\theta(t)$ = spacecraft received signal phase

f_u = uplink frequency

T_u = uplink phase transit time

The transit time from a station ($x = 0$) to a spacecraft ($x = \rho$) may be found from

$$T_u(t) = \int_0^{\rho(t)} \frac{dx}{v_p} \cong \int_0^{\rho(t)} \frac{dx}{c} \left[1 - \frac{A}{f_u^2} N_e\left(x, t - \frac{\rho - x}{c}\right)\right]$$

$$= \frac{1}{c}\left(\rho(t) - \frac{A}{f_u^2} I_u(t)\right) \qquad (4.4\text{-}10)$$

where

t = time of arrival at the spacecraft

$I_u(t)$ = integral of electron density along uplink ray path, or columnar electron density, m^{-2}

$$= \int_0^{\rho(t)} N_e\left(x, t - \frac{\rho - x}{c}\right) dx$$

Substituting T_u into (4.5-9), one obtains for the received phase

$$\theta_{SR} = 2\pi f_u \left[t - \frac{\rho(t)}{c} + \frac{A}{cf_u^2} I_u(t)\right] \qquad (4.4\text{-}11)$$

corresponding to an instantaneous frequency

$$f_{SR} = \frac{1}{2\pi}\frac{d\theta_{SR}}{dt} = f_u\left[1 - \frac{\dot{\rho}(t)}{c} + \frac{A}{cf_u^2}\dot{I}_u(t)\right] \qquad (4.4\text{-}12)$$

This is identical to (4.2-4) except for the last term, which is due to the time rate of change of the columnar electron density.

The spacecraft received frequency is coherently multiplied by the spacecraft and, neglecting the very small change in the downlink charged-particle effect due to uplink radial motion and charged particles, yields the following equation for the earth-received phase:

$$\theta_R(t) = 2\pi K f_u \left[t - \frac{2\rho(t - T_d)}{c} + \frac{A}{cf_u^2}\left(I_u(t - T_d) + \frac{1}{K^2} I_d(t)\right)\right] \qquad (4.4\text{-}13)$$

where

$\theta_R(t)$ = earth-received phase

$\rho(t - T_d)$ = spacecraft range at $(t - T_d)$

t = time of arrival at earth station

T_d = downlink phase transit time

$$= \frac{\rho(t - T_d)}{c} - \frac{A}{cf_{ST}^2 (t - t_d)} I_d(t)$$

$I_d(t)$ = downlink columnar electron content

K = transponder turnaround ratio
= 240/221 (S-band); 880/221 (X-band)

The instantaneous received frequency is, therefore,

$$f_R = \frac{1}{2\pi} \frac{d\theta_R}{dt} = Kf_u \left[1 - \frac{2\dot{\rho}}{c} + \frac{A}{cf_u^2} \left(\dot{I}_u + \frac{1}{K^2} \dot{I}_d\right)\right] \quad (4.4\text{-}14)$$

Thus, the frequency of the received signal will be altered by any changes in the columnar electron content. This change in frequency typically reaches values at S-band corresponding to a velocity of about one millimeter per second [4-18] due to the diurnal variation of the earth's ionosphere, but can be calibrated by several means. For instance, since the effect of charged particles is dependent on frequency, the difference in frequency between the S- and X-band downlinks on some spacecraft yields a measure of the variation of the columnar electron content on the downlink.

4.4.1.2.2 *Group delay*. Proceeding in the same manner as for phase delay, and using (4.4-7), it is possible to show that the uplink group transit time is

$$T_u = \int_0^{\rho(t)} \frac{dx}{v_g} = \frac{\rho(t)}{c} + \frac{A}{cf_u^2} I_u(t + T_u) \quad (4.4\text{-}15)$$

where

T_u = uplink group transit time

$\rho(t)$ = true spacecraft range

A = 40.3 (mks)

c = speed of light in vacuum

I_u = uplink columnar electron density

t = time of transmission from earth

$t + T_u$ = time of reception at spacecraft

f_u = uplink frequency

The round-trip group transit time is

$$T_{u+d} = \frac{\rho(t) + \rho(t + T_u)}{c} + \frac{A}{cf_u^2}\left(I_u(t + T_u) + \frac{1}{K^2} I_d(t + T_u + T_d)\right) \tag{4.4-16}$$

where

K = transponder turnaround ratio

= 240/221 (S-band) or 880/221 (X-band)

As in the phase delay measurement, it is possible to measure the downlink columnar electron content on spacecraft which have two downlink frequencies

$$\Delta T_{u+d} = \frac{A}{c}\left(\frac{1}{f_1^2} - \frac{1}{f_2^2}\right) I_d(t + T_u + T_d) \tag{4.4-17}$$

The magnitude of the S-band group delay due to charged particles in the earth's ionosphere ranges from about 0.5 to 15 m [4-18], while the group delay due to interplanetary plasma ranges from about 10 m for spacecraft at opposition at a distance of a few AU to 1 km or more for spacecraft at superior conjunction.

4.4.1.2.3 *Differenced range versus integrated doppler (DRVID).* As the name suggests, DRVID is defined as the difference between changes in the apparent range as measured by phase and group delay [4-21]. The apparent phase delay rate $\dot{\rho}_p$ is the output observable of doppler measurements. Integrating the phase change between t_0 and t_1, and subtracting it from the difference between two

measurements of the group (ranging) delay, we get DRVID:

$$\text{DRVID}(t) = \rho_g(t_1) - \rho_g(t_0) - \int_{t_0}^{t_1} \dot{\rho}_p(t)\, dt \qquad (4.4\text{-}18)$$

where

$\dot{\rho}_p$ = time derivative of range measured with phase delay

ρ_g = range measured with group delay

It is possible to show that, like doppler frequency, DRVID depends on the time rate of change of columnar electron content. However, DRVID measurements of plasma change can be made with a single carrier frequency and also measure uplink as well as downlink effects:

$$\text{DRVID}(t) = \frac{A}{f_u^2} \int_{t_0}^{t_1} \left[\dot{I}_u(\tau - T_d) + \frac{1}{K^2} I_d(\tau) \right] d\tau \qquad (4.4\text{-}19)$$

where

A = 40.3 (mks)

f_u = uplink frequency, Hz

c = speed of light, m/s

t_1 = end of integration period

t_0 = start of integration

$I_{u/d}$ = time derivative of uplink/downlink columnar electron content

T_d = downlink time delay

K = transponder turnaround ratio

= 240/221 (S-band); 880/221 (X-band)

4.4.1.3 Thermal noise

4.4.1.3.1 *Doppler.* It is shown in Chapter 3 that the phase-locked-loop receivers used in spacecraft and in the Deep Space Network develop a random phase error when the received signal is corrupted by white gaussian noise. The variance of

the doppler phase difference at the ground receiver in two-way tracking, and under strong signal conditions [4-22], is

$$\sigma^2_{\phi_{d2}} = \frac{1}{d_1} + \frac{1}{\rho_2} \tag{4.4-20}$$

where

ϕ_{d2} = doppler phase difference = $\hat{\theta}_2 - \theta_0$

$\hat{\theta}_2$ = DSS receiver estimate of the received carrier phase

θ_0 = transmitted uplink carrier phase

$\sigma^2_{\phi_{d2}}$ = variance of doppler phase difference

$$d_1 = \frac{\rho_1}{G_f^2 K_D}$$

G_f = downlink carrier frequency to uplink carrier frequency ratio

$$K_D = \frac{1}{r_1 + 1}\left[\frac{r_2(r_1+1) + r_1(r_1 + r_2 + r_1 r_2)(\xi + \xi^2) + r_1(r_1 r_2 + r_1)\,\xi^3/r_2}{r_2 + r_1 r_2\,\xi + r_1(r_1 + r_2 - 2)\,\xi^2 + r_1^2\,\xi^3 + r_1^2\,\xi^4/r_2}\right]$$

r_i = loop damping parameters (defined in Chapter 3), $i = 1, 2$

$$\xi = \frac{W_{L1}(r_2 + 1)}{W_{L2}(r_1 + 1)}$$

W_{Li} = receiver loop bandwidths (defined in Chapter 3), $i = 1, 2$

ρ_1 = uplink input signal-to-noise ratio in bandwidth of the linearized spacecraft receiver loop

ρ_2 = downlink input signal-to-noise ratio in bandwidth of the linearized ground receiver loop

and $i = 1$ indicates uplink and $i = 2$ indicates downlink.

The doppler frequency estimated by the tracking system is

$$f_d = \frac{1}{2\pi}\frac{\phi_{d2}(t_1) - \phi_{d2}(t_0)}{t_1 - t_0} = \frac{1}{2\pi}\frac{\Delta\phi_{d2}}{\Delta t} \tag{4.4-21}$$

with the corresponding variance

$$\sigma^2_{f_D} = \left(\frac{\sqrt{2}}{2\pi\,\Delta t}\right)^2 \sigma^2_{\phi_{d2}} \tag{4.4-22}$$

where

f_D = estimated doppler frequency

t_1 = end of doppler sampling interval

t_0 = start of doppler sampling interval

Δt = length of doppler sampling interval

4.4.1.3.2 *Ranging*. The variance of the time delay estimated by the Planetary Ranging Assembly can be shown to be [4-15]

$$\sigma^2_{\hat{\tau}} = \frac{T^2}{32T_i} f^2(\tau)\frac{N_0}{P_r} \tag{4.4-23}$$

for square-wave ranging (strong signal case) or

$$\sigma^2_{\hat{\tau}} = \frac{T^2}{64T_i}\,\frac{N_0}{P_r} \tag{4.4-24}$$

for sine-wave[4] ranging [4-16], where

$\hat{\tau}$ = time delay estimate

$\sigma^2_{\hat{\tau}}$ = variance of time delay estimate

T = period of ranging signal

T_i = correlation interval

P_r = power of the received ranging signal, W

N_0 = one-sided spectral density of input noise, W/Hz

[4] Actually square-wave versus sine-wave correlation, also known as 1-MHz filtered ranging. This is the current DSN technique.

$f^2(\tau)$ = function that depends on measured time delay

$= (32\,\tau_m^2/T^2) - (8\,\tau_m/T) + 1$

τ_m = τ modulo $T/4$ $(0 \leqslant \tau_m \leqslant T/4)$

4.4.1.4 Instrumentation errors

4.4.1.4.1 *Doppler sampling interval.* The timing signals generated by the Deep Space Station frequency and timing subsystem contain an rms error of 5 nsec. For biased doppler frequencies near 1 MHz, a 5-nsec error in sampling time corresponds to an error in phase of 1.8 deg.

4.4.1.4.2 *Doppler resolver quantization error.* The doppler resolver uses a 100-MHz (500 MHz in the future) clock to measure the phase of the biased doppler frequency, leading to a 10-nsec (2 nsec in the future) uniformly distributed error in the measured phase. Assuming a biased doppler frequency near 1 MHz, this results in a uniformly distributed error of a 3.6-deg width, or an rms error of $3.6/\sqrt{12} = 1.0$ deg.

4.4.1.4.3 *Reference oscillator instability.* Reference oscillator instability has two effects: a drift in frequency that occurs during the time of flight of the signal to the spacecraft (described in Fig. 4-18), and an accumulated phase error during a doppler sampling interval (described in Figs. 4-19 through 4-22).

4.4.1.4.4 *Path delay instability.* Figure 4-23 shows the error requirement for path delay for the Receiver–Exciter Subsystem. Since the actual distribution of the path delay variation is not known, it is assumed to be uniform within the given bounds. The corresponding doppler frequency variance is

$$\sigma_{fD}^2 = (360\,P(T)/\sqrt{12}\,\lambda)^2 \qquad (4.4\text{-}25)$$

where

σ_{fD}^2 = doppler frequency variance, deg^2

$P(T)$ = peak delay variation, m

T = sample time

λ = wavelength of received signal, m

4.4.1.4.5 *Total path delay.* The two-way group delay measured by the Tracking Subsystem contains elements of delay due to spacecraft and ground hardware (Fig. 4-24). This systematic error is removed by the following technique [4-23]:

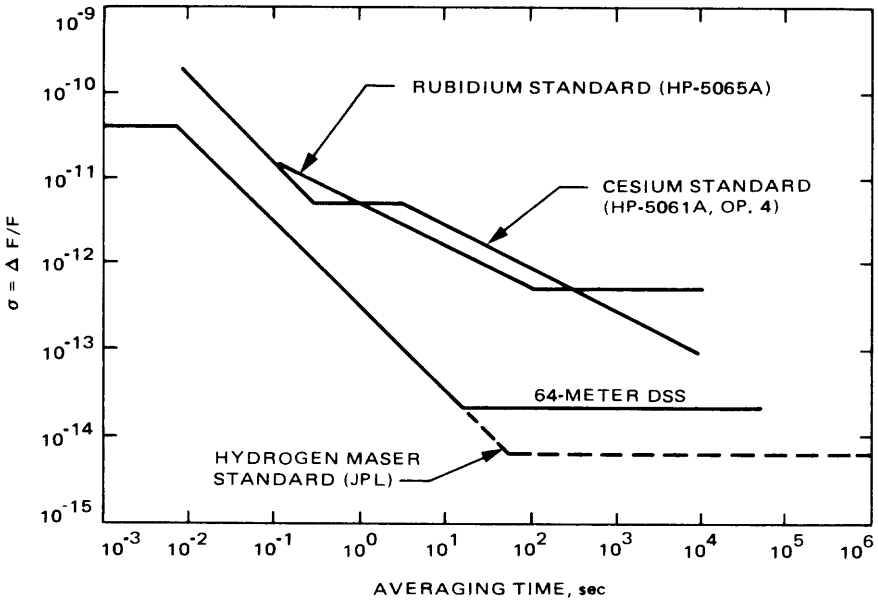


Fig. 4-18. DSS frequency stability

$$RTPT = D - BIAS_{SC} - BIAS_{DSS} - Z_{CORRECTION} \qquad (4.4\text{-}26)$$

where

$RTPT$ = round-trip propagation time from DSS reference location to spacecraft reference location back to DSS reference location

D = time delay measured by tracking system

$BIAS_{SC}$ = spacecraft turnaround delay

$BIAS_{DSS}$ = two-way time delay to a DSS turnaround device in Antenna Microwave Subsystem

$Z_{CORRECTION}$ = time delay correction which takes account of the distance from the turnaround device to the DSS reference location

This technique leaves several residual errors:

(1) Random error due to thermal noise on measurement of $BIAS_{DSS}$ and $BIAS_{SC}$ with magnitude given by (4.4-24).

(2) Systematic error due to path delay uncertainty during ranging measurements, magnitude given in Fig. 4-23.

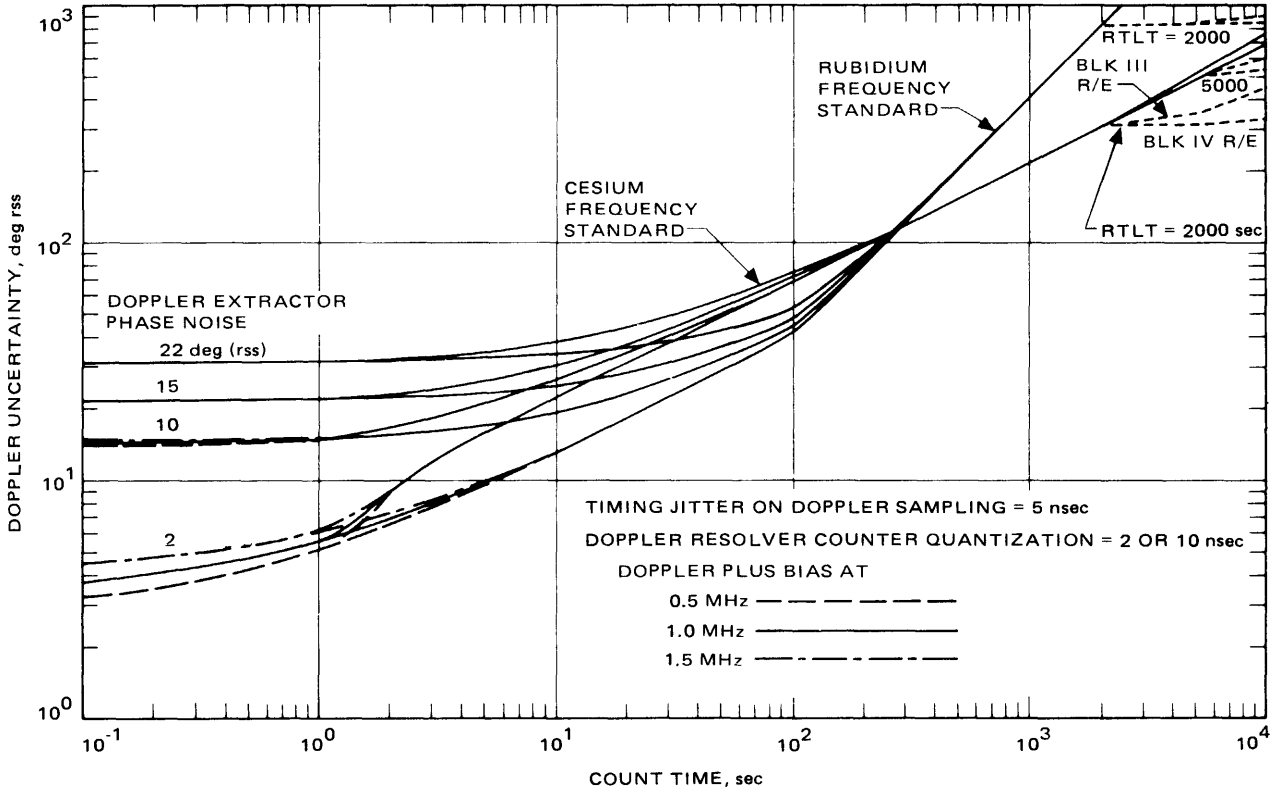


Fig. 4-19. S-band doppler system phase uncertainty versus count time: rubidium and cesium frequency standards

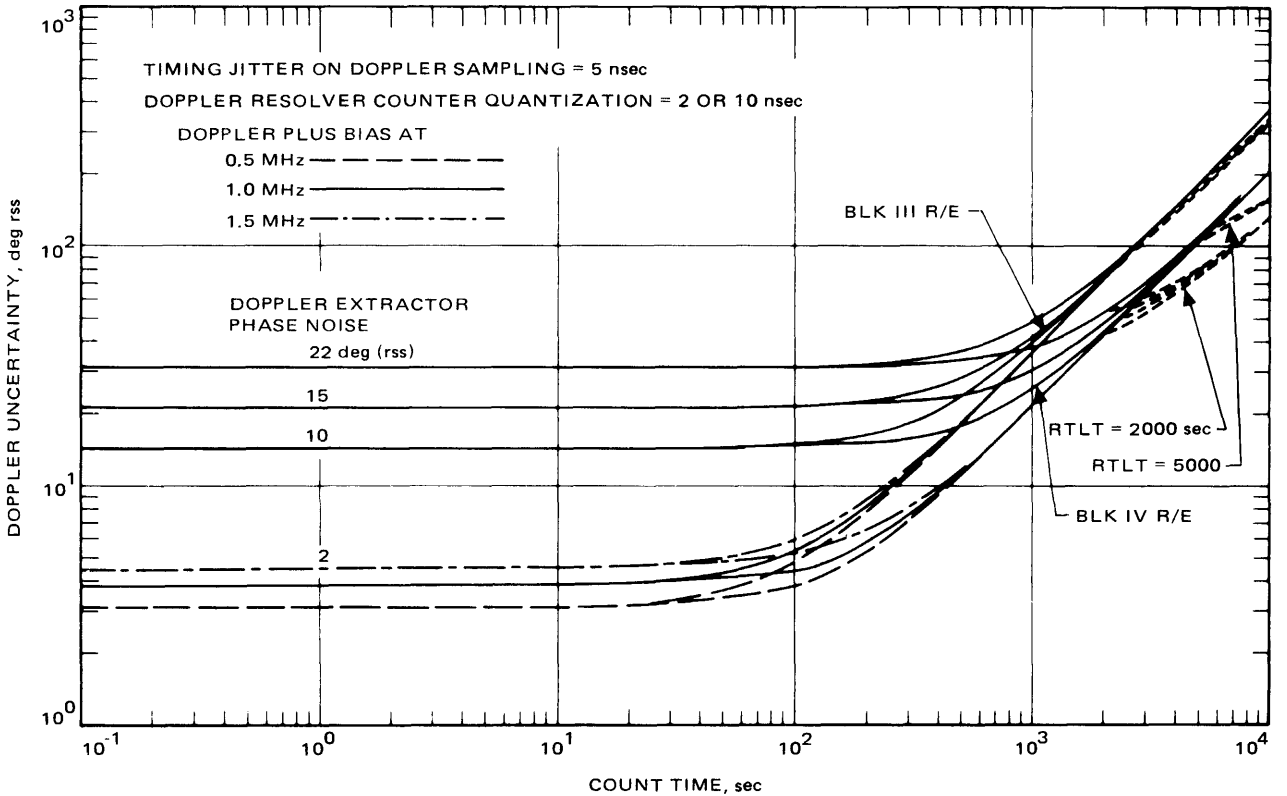


Fig. 4-20. S-band doppler system phase uncertainty versus count time: hydrogen maser frequency standard

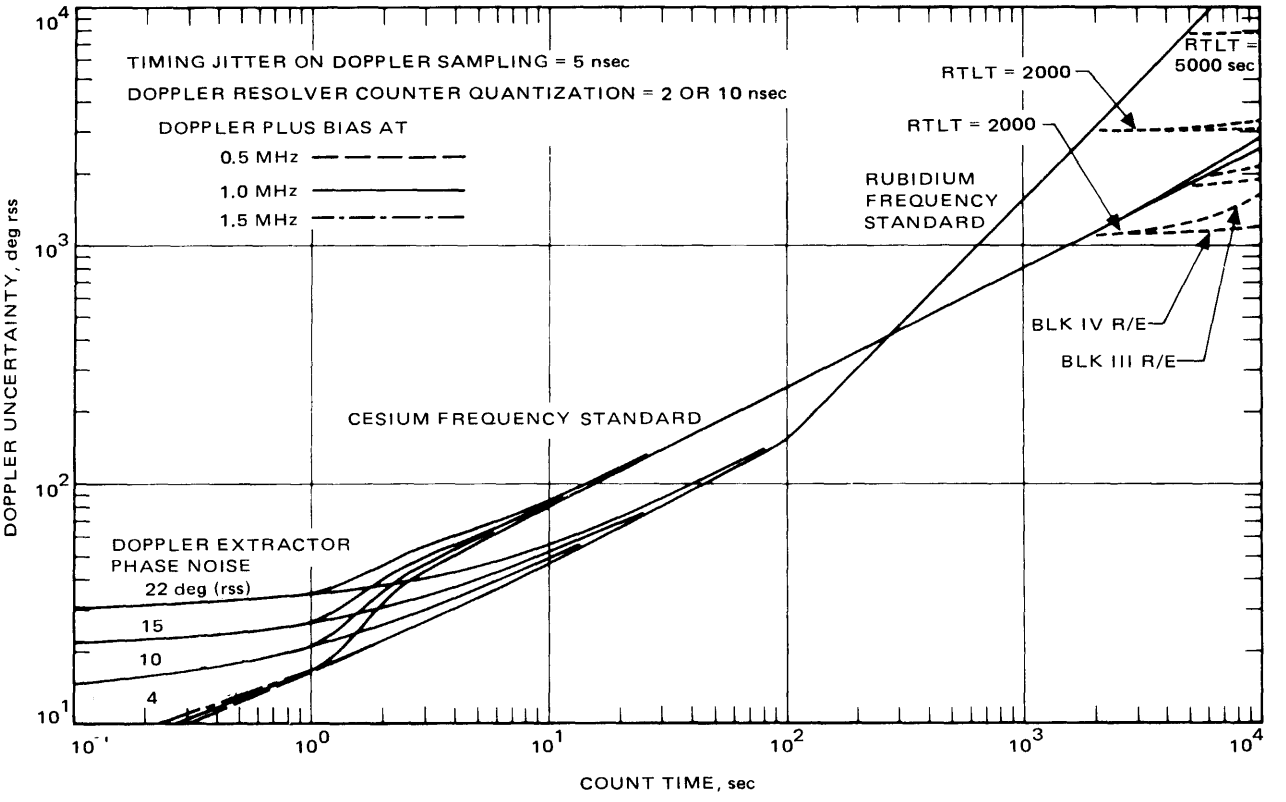


Fig. 4-21. X-band doppler system phase uncertainty versus count time: rubidium and cesium frequency standards

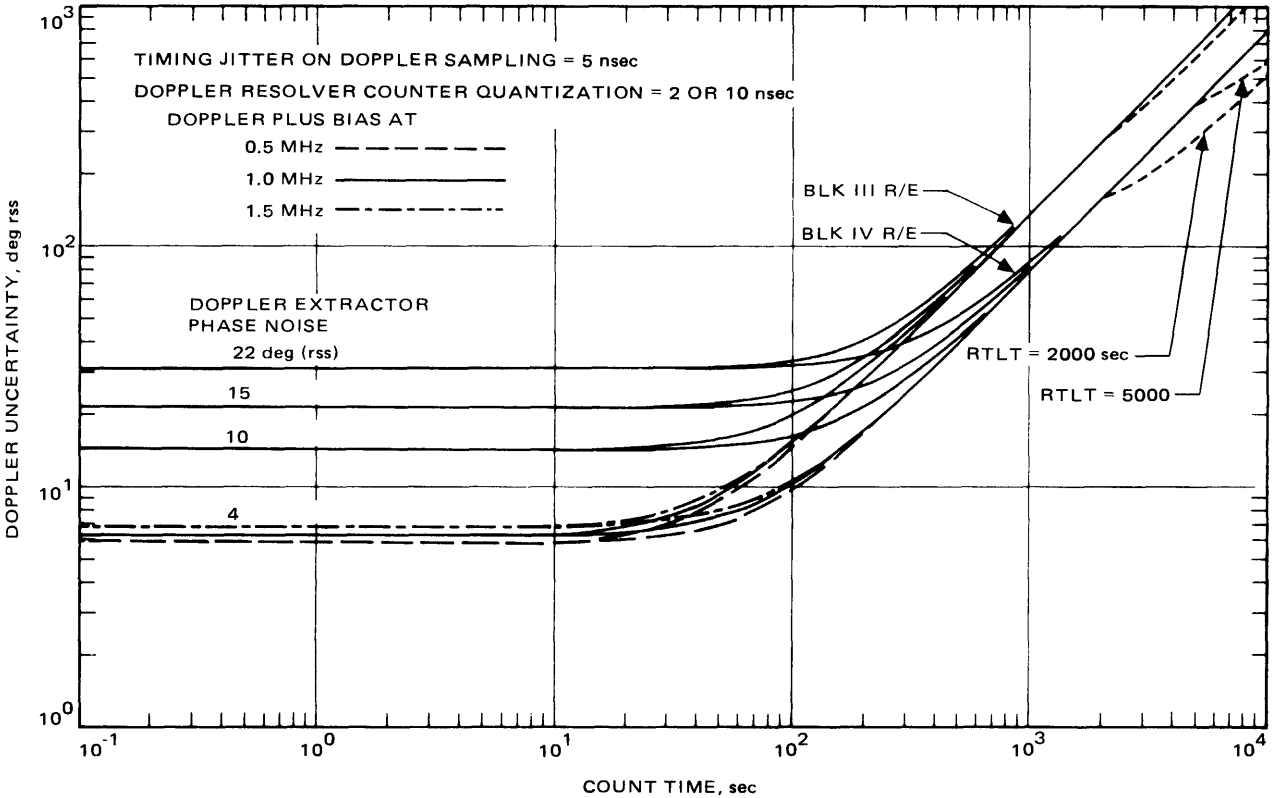


Fig. 4-22. X-band doppler system phase uncertainty versus count time: hydrogen maser frequency standard

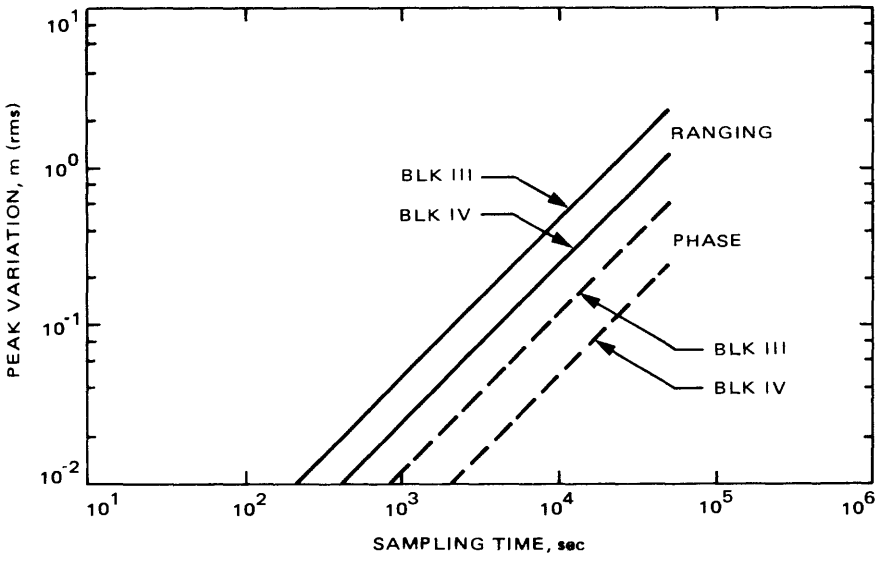


Fig. 4-23. Instrumental path delay error requirement

(3) Systematic error due to aging of the spacecraft transponder with magnitude unknown, but believed to be less than 6 nsec.

(4) Systematic error due to calculation of $Z_{CORRECTION}$ from measurements on individual components, magnitude unknown, but believed to be less than 1 nsec [4-24].

4.4.1.4.6 *Waveform distortion.* Filtering of a square-wave ranging signal by the 1.5-MHz spacecraft transponder bandwidth and nonlinearities in modulators and power amplifiers lead to a systematic error when time delay is determined by correlation between assumed perfect square waves. This error depends on time delay and has a peak value of about 7.4 nsec for 500-kHz unfiltered ranging, but less than 0.6 nsec for 1-MHz filtered ranging [4-25]. The current mechanization of the Planetary Ranging Assembly measures waveform distortion and plasma variation simultaneously with the DRVID mode, so this error can be removed when there is small plasma activity, with an accuracy limited by thermal noise.

4.4.1.4.7 *Multipath.* Reflection of received carrier signals from antenna supports and microwave interfaces and subsequent superposition with the main signal can result in systematic error in group delay measurements. Several studies have shown that these effects are less than 5 nsec for 64-m antennas, and less than 0.5 nsec for 26-m antennas when the signal source is in the far field [4-26 to 4-28] but can be as large as 10 to 15 m for near-field measurements [4-29].

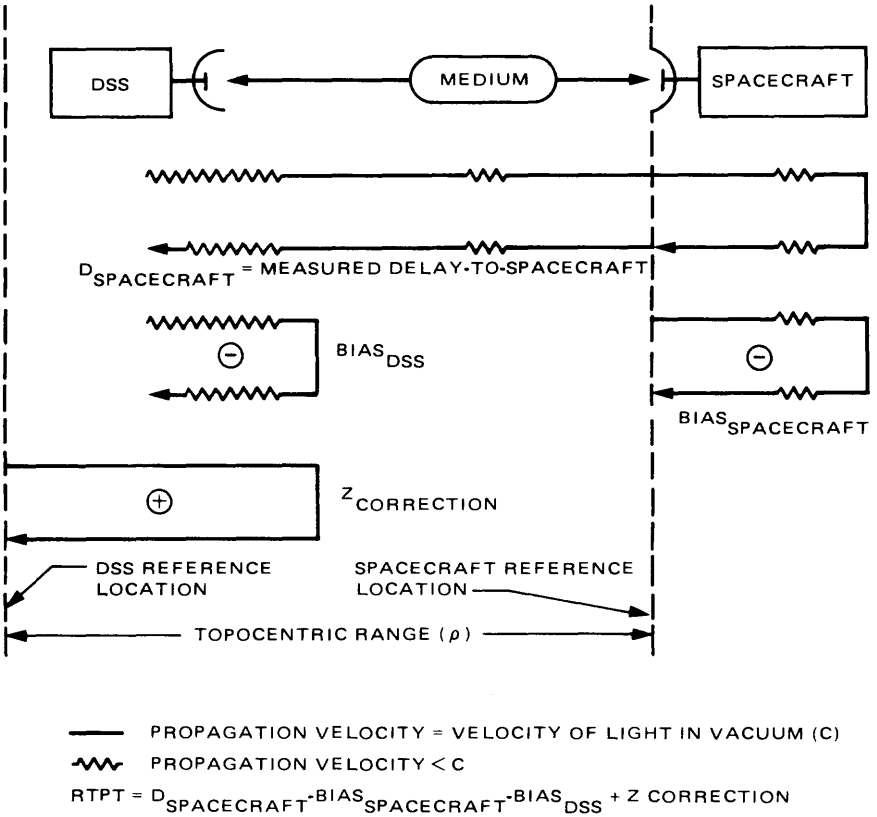


Fig. 4-24. Removal of hardware delay from range measurement

4.4.2 The Radiometric Design Control Table

The performance of the tracking system is controlled in a manner similar to that used for control of the performance of the telemetry and command systems, namely, through a design control table (DCT) [4-30]. In principle, all three types of DCTs tabulate the properties of a number of random variables that are summed to obtain the random variable of interest. In the case of telemetry and command DCTs, the variable of interest is a signal-to-noise ratio or performance margin; for radiometric DCTs the variable of interest is the value of a measurement of frequency, range, or angle, etc.

A slight difference in technique arises because radiometric system performance cannot be measured in terms of the mean or "true" value of a radiometric measurement in the way that communication system performance can be measured in terms of the mean value of a performance margin. In radiometric systems the performance is measured in terms of the standard deviation of a mea-

surement, and hence the standard deviation is the only value tabulated in a radiometric DCT.

Ideally, the probability distributions of all sources of random error would be known, and calibration schemes leaving at most a random error of known distribution would exist for all forms of systematic error. Then the expected measurement error could be determined by convolution of all the known probability distributions. This is far from the case in the radio tracking system, and, although the situation has been consistently improved, we are forced to make several simplifying assumptions.

Assumption 1. All random errors are stationary, Gaussian, and zero mean. This is true for random error due to thermal noise, but is justified for other random errors by the fact that the combination of a number of independent probability distributions tends toward a Gaussian probability distribution, in the limit of large numbers of probability distributions (Central Limit Theorem). This assumption leads to the root sum of square (RSS) technique for determination of the expected standard deviation of a measurement. Specifically,

$$\sigma(q) = \sqrt{\sum_i \left(\frac{\partial q}{\partial x_i} \sigma(x_i) \right)^2} \qquad (4.4\text{-}27)$$

where

$\sigma(q)$ = standard deviation of the quantity being measured, q

$\sigma(x_i)$ = standard deviation of a random variable x_i that generates error in q

$\frac{\partial q}{\partial x_i}$ = rate of change of q with respect to x_i

The quantity tabulated in radiometric DCTs is always $(\partial q/\partial x_i)\, \sigma(x_i)$, except for totals, which are always $\sigma(q)$.

Assumption 2. Calibrated systematic errors leave a residual random error which is assumed to be stationary, Gaussian, and zero mean. This is motivated by the experience that over many days or years systematic error appears to be randomly distributed. In some cases, such as calibration of plasma effects through dual-frequency measurements, or calibration of ground station group delay, a significant fraction of the remaining systematic error arises from thermal noise during the calibrating measurement, and hence may be considered to be random over a large number of measurements.

Assumption 3. Uncalibrated systematic error is assumed to be uniformly distributed within a certain region. The standard deviation of the uniform probability distribution is then used as an entry to the table:

$$\sigma(x_i) = \frac{a}{\sqrt{12}} \tag{4.4-28}$$

where a is the width of uniform distribution.

One final comment is required concerning the radiometric DCT. Generally, the DCT applies only to a single, specific combination of link, spacecraft, ground, and processing parameters. However, several combinations of parameters may appear in a single table under the guise of adverse and favorable tolerances. For instance, in the representative[5] DCT in Table 4-1, the adverse column uses the assumption of a heavily modulated carrier (telemetry at 8.5 dB and ranging at 6 dB carrier suppression) and relaxed spacecraft temperature control (4.0°C peak to peak). These factors lead to increased thermal noise and decreased spacecraft frequency stability, respectively. The favorable and adverse tolerances shown do not represent uncertainty in the assignment of standard deviations, but rather they represent the range of conditions that might occur during mission operations.

[5] A note of caution: This table does not apply to an existing system.

Table 4-1. International Solar Polar Mission doppler frequency stability design control table (adapted from [4-31])

Error source	Sigma (Delta F/F) 10^{-15}			
	Nominal	Favorable	Adverse	Note
Spacecraft RSS	5.03	0.53	20.0	Calibrated
Cables, microwave	0.09	0.09	0.2	
VCO	0.50	0.50	1.0	
Transponder path	5.0	0.5	20.0	
Power amplifier	0.13	0.13	0.22	
Link RSS	1.5	1.1	5.4	Gaussian
Polarization	–	–	–	
Thermo noise	1.5	1.1	5.4	
Spacecraft	–	–	–	
Ground station	–	–	–	
Propagation media RSS	17.7	6.1	51.4	
Troposphere				
Uplink	11.0	3.9	32.0	Uncalibrated
Downlink	11.0	3.9	32.0	Uncalibrated
Ionosphere				
Uplink	0.83	0.26	2.6	
Downlink	0.83	0.26	2.6	Calibrated
Solar plasma				
Uplink	8.3	2.5	17.0	
Downlink	0.8	0.2	17.0	Calibrated
Ground station RSS	4.6	3.6	43.0	Calibrated
Frequency standard	1.1	1.1	–	
Frequency distribution	0.1	0.1	–	
Receiver path	–	–		
Maser, microwave	1.8	1.8	–	
System cables	0.1	0.1	–	
Receiver (w/o VCO)	0.9	0.9	–	
VCO	–	–		
Doppler extractor	0.8	0.8	–	
Transmitter path				Calibrated
Exciter	2.5	2.5	–	
Power amp., buffer	0.3	0.3	–	
System cables	–	–		
Microwave	0.1	0.1	–	Uncalibrated
Antenna mechanical				
Gravity	3.0	1.0	–	
Conscan	–	–		
Wind loading	–	–		
Vibration	–	–		
Data acquisition				
Timing	0.1	0.1	–	Uniform
Quantization	0.3	0.3	–	distribution

Table 4-1 (contd)

Error source	Sigma (Delta F/F) 10^{-15}			
	Nominal	Favorable	Adverse	Note
Unmodeled motion RSS	0	0	0	
Spacecraft	–	–	–	
Platform	–	–	–	
Total	19.0	7.1	70.1	
Requirement	3.0	3.0	3.0	
Performance margin	0.16	0.42	0.043	

Conditions:

Spacecraft: ISPM	DSS: 34 m
Mission phase: 2nd opposition	Link: HGA, 20 kW
Range: 4.2 AU	Elevation: 45 deg
Sampling interval: 1000 sec	Cmd modulation: off
Frequency: X-up, X-down	Tlm modulation: off (Nom & Fav), 8.5 dB (Adv)
Tropospheric cals: none	Rng modulation: off (Nom & Fav), 6 dB (Adv)
Plasma cals: X-down	S/C temp deviation: 1.0/0.1/4.0°C
Freq source: H_2 maser	(Nom/Fav/Adv)

Appendix

Uses of Tracking Data for Radio Science

Frequency or Phase Delay

Doppler frequency

(1) Mass and mass distribution of planets and mass of satellites [4-32, 4-33, 4-34, 4-35].

(2) Planet spin axis orientation, motion, spin rate [4-36].

(3) Planet size, shape, and gravity field [4-32, 4-37].

(4) Gravitational waves [4-38, 4-39].

(5) Pressure, temperature, and density of atmospheres [4-40].

(6) Electron concentration of ionospheres and the interplanetary plasma [4-11].

Spectrum

(1) Turbulence of atmospheres [4-11].

(2) Turbulence of the solar wind [4-41].

(3) Distribution of planet surface slopes [4-42].

(4) Size and velocity distributions of planetary ring particles [4-11].

Group Delay

(1) Electron concentration of ionospheres, and the interplanetary plasma [4-11].

(2) Solar corona shape, density [4-43], large scale flow patterns in the solar wind [4-44].

(3) Parameters of general relativistic theory [4-45, 4-46].

Intensity

(1) Pressure, temperature, and density profiles of atmospheres [4-11].

(2) Orientation of gravitational equipotential surfaces in an atmosphere [4-11].

(3) Location, density, and composition of clouds [4-11].

(4) Sizes of planetary ring particles [4-11].

Polarization

(1) Shape of planetary ring particles [4-11].

(2) Electron density and magnetic field in ionospheres and the solar corona [4-47].

(3) Dielectric properties of planetary surfaces [4-42, 4-48].

(4) Test of the weak equivalence principle [4-49].

References

4-1. Ondrasik, V. J., and Rourke, K. H., "Application of New Radio Tracking Data Types to Critical Spacecraft Navigation Problems," *Quarterly Technical Review*, Vol. 1, No. 4, Jet Propulsion Laboratory, Pasadena, Calif., Jan. 1972.

4-2. Curkendall, D. W., and McReynolds, S. R., "A Simplified Approach for Determining the Information Content of Radio Tracking Data," *Journal of Spacecraft and Rockets*, Vol. 6, No. 5, pp. 520–525, May 1969.

4-3. Kong, J. A., *Theory of Electromagnetic Waves*, Wiley, New York, N.Y., p. 68, 1975.

4-4. Melbourne, W. G., "Navigation Between the Planets," *Scientific American*, Vol. 234, No. 6, pp. 58–74, June 1976.

4-5. Curkendall, D. W., and Stephenson, R. R., "Earthbased Tracking and Orbit Determination," *Astronautics and Aeronautics*, pp. 30-36, May 1970.

4-6. Moyer, T.D., *Mathematical Formulation of the Double-Precision Orbit Determination Program (DPODP)*, Technical Report 32-1527, Chapters VIII and IX, Jet Propulsion Laboratory, Pasadena, Calif., May 15, 1971.

4-7. Curkendall, D. W., "Radiometric Technology for Deep Space Navigation: A Development Overview," paper presented at AIAA/AAS Astrodynamics Conference, Palo Alto, Calif., Aug 7–9, 1978, American Institute of Aeronautics and Astronautics, New York, N. Y.

4-8. Martin, W. L., and Zygielbaum, A. I., *MU-II Ranging*, Technical Memorandum 33-768, Jet Propulsion Laboratory, Pasadena, Calif., pp. 3–6, May 15, 1977.

4-9. Rogers, A.E.E., "Very Long Baseline Interferomery with Large Effective Bandwidth for Phase-Delay Measurements," *Radio Science*, Vol. 5, No. 10, pp. 1239–1247, Oct. 1970.

4-10. Thomas, J. B., "An Analysis of Long Baseline Radio Interferometry," Technical Report 32-1526, Vol. VII, pp. 37–50, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1972.

4-11. Eshleman, V. R., et al., "Radio Science Investigations with Voyager," *Space Science Reviews*, Vol. 21, pp. 207–232, Nov. 1977.

4-12. Michael, W. H., et al., "The Viking Radio Science Investigations," *Journal of Geophysical Research*, Vol. 82, pp. 4293–4295, Sept. 30, 1977.

4-13. Howard, H. T., *Planetary Data System Requirements: Multi-Mission Radio Science Requirements for the 1978 to 1988 Era*, JPL Publication 79-74, Jet Propulsion Laboratory, Pasadena, Calif., pp. 5–45, Oct. 1, 1979.

4-14. Martin, W. L., "A Binary-Coded Sequential Acquisition Ranging System," *The Deep Space Network*, Space Programs Summary 37-57, Vol. II, pp. 72-81, Jet Propulsion Laboratory, Pasadena, Calif., May 31, 1969.

4-15. Komarek, T. A., "Ranging Jitter and Mean Estimate Error," IOM 3395-75-276, Jet Propulsion Laboratory, Pasadena, Calif., Sept. 26, 1975 (an internal document).

4-16. Dolinsky, S., "Ranging Jitter and Variance for Squarewave-Sinewave Correlation," IOM Voyager-NAV-80-148, Jet Propulsion Laboratory, Pasadena, Calif., Sept. 12, 1980 (an internal document).

4-17. Renzetti, N. A., *DSN/Flight Project Interface Design Book*, Document 810-5, Rev. D, Chapters TRK20 and TRK30, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1975 (an internal document).

4-18. Lesh, J. R., and Winn, F. B., *Navigation Network Study, Phase A Study Report*, Document 890-112, pp. 9-1 to 9-18, Jet Propulsion Laboratory, Pasadena, Calif., May 12, 1979 (an internal document).

4-19. Smith, E. K., and Weintraub, S., "The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies," *Proceedings of IRE*, Vol. 41, pp. 1035-1037, Aug. 1953.

4-20. Komarek, T. A., "Radiometric Measurements – A Simplified Analysis," IOM 3395-74-173, Jet Propulsion Laboratory, Pasadena, Calif., May 3, 1974 (an internal document).

4-21. MacDoran, P. F., "A First-Principles Derivation of the Differenced Range Versus Integrated Doppler (DRVID) Charged Particle Calibration Method," *The Deep Space Network*, Space Program Summary 37-62, Vol. II, pp. 28-34, Jet Propulsion Laboratory, Pasadena, Calif., Mar. 31, 1970.

4-22. Lindsey, W. C., and Simon, M. K., *Telecommunication Systems Engineering*, Prentice-Hall, Englewood Cliffs, N.J., 1973.

4-23. Komarek, T., and Otoshi, T., "Terminology of Ranging Measurements and DSS Calibrations," *DSN Progress Report 42-36*, pp. 35-40, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1976.

4-24. Hartop, R., "Microwave Time Delays in the DSN 34- and 64-m Antennas," *DSN Progress Report 42-51*, pp. 183-185, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1977.

4-25. Layland, J. W., Zygielbaum, A. I., and Hubbard, W. P., "On Improved Ranging," *DSN Progress Report 42-46*, pp. 40-45, Jet Propulsion Laboratory, Pasadena, Calif., May 1978.

4-26. Otoshi, T. Y., and Brunn, D. L., "Multipath Tests on 64-m Antennas Using the Viking Orbiter-1 and -2 Spacecraft as Far-Field Illuminators," *DSN*

Progress Report 42-31, pp. 41-49, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1976.

4-27. Otoshi, T. Y., and Taylor, T. H., "An Experimental Investigation of the Effects of Antenna Pointing Errors on Range Delays," *DSN Progress Report 42-53*, pp. 141-147, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 15, 1979.

4-28. Otoshi, T. Y., "Experimental Investigation of the Effects of Antenna Pointing Errors on Range Delays (Part II)," *DSN Progress Report 42-56*, pp. 143-146, Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1980.

4-29. Otoshi, T. Y., "S-band Zero-Delay Device Multipath Tests on the 64-m Antenna at DSS 43, DSS 63, and DSS 14," *DSN Progress Report 42-29*, pp. 20-30, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 15, 1975.

4-30. Yuen, J. H., *A Practical Statistical Model for Telecommunications Performance Uncertainty*, Technical Memorandum 33-732, p. 6, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1974.

4-31 Komarek, T. A., and Meeker, J., *X-Band Uplink Technology Demonstration*, Experiment Requirements Document 900-944, pp. B-38 to B-43, Jet Propulsion Laboratory, Pasadena, Calif., Mar. 1, 1981 (an internal document).

4-32. Null, G. W., et al., "Gravity Field of Jupiter from Pioneer II Tracking Data," *Science*, Vol. 188, pp. 476-477, May 2, 1975.

4-33. Howard, H. T., et al., "Results on Mass, Radius, Ionosphere and Atmosphere from Mariner 10 Dual Frequency Radio Signals," *Science*, Vol. 185, pp. 179-180, July 12, 1974.

4-34. Sjogren, W. J., et al., "Gravity Anomalies on Venus," *Journal of Geophysical Research*, Vol. 85, No. A13, pp. 8295-8302, Dec. 30, 1980.

4-35. Anderson, J. D., Hubbard, W. B., and Slattery, N. L., "Structure of the Jovian Envelope from Pioneer 10 Gravity Data," *Astrophysical Journal*, Vol. 193, pp. L149-L150, Nov. 1974.

4-36. Michael, W. H., et al., "Viking Lander Location and Spin Axis at Mars – Determination from Tracking Data," *Science*, Vol. 193, p. 803, Aug. 27, 1976.

4-37. Sjogren, W. L., et al., "Mars Gravity Field Based on a Short-Arc Technique," *Journal of Geophysical Research*, Vol. 80, No. 20, pp. 2899-2908, July 10, 1975.

4-38. Thorne, K. S., and Braginsky, V. B., "Gravitational-Wave Bursts from Nuclei of Distant Galaxies and Quasars: Proposal for Detection Using Doppler Tracking of Interplanetary Spacecraft," *Astrophysical Journal*, Vol. 204, pp. L1-L6, Feb. 15, 1981.

4-39. Estabrook, F. B., and Wahlquist, H. D., "Response of Doppler Spacecraft Tracking to Gravitational Radiation," *General Relativity and Gravitation*, Vol. 6, No. 5, pp. 439-477, 1975.

4-40. Kliore, A. J., and Patel, I. R., "Vertical Structure of the Atmosphere of Venus from Pioneer Venus Orbiter Radio Occultations," *Journal of Geophysical Research*, Vol. 85, No. A13, pp. 7957-7962, Dec. 30, 1980.

4-41. Woo, R., "Radial Dependence of Solar Wind and Properties Deduced from Helios 1/2 and Pioneer 10/11 Radio Scattering Observations," *Astrophysical Journal*, Part I, Vol. 219, pp. 727-739, Jan. 15, 1978.

4-42. Simpson, R. A., et al., "Viking Bistatic Radar Observations of the Hellas Basin on Mars–Preliminary Results," *Science*, Vol. 203, pp. 45-46, Jan. 5, 1979.

4-43. Tyler, G. L., et al., "The Viking Solar Corona Experiment," *Journal of Geophysical Research*, Vol. 82, No. 28, pp. 4335-4340, Sept. 30, 1977.

4-44. Winn, F. B., et al., "A Solar Plasma Stream Measured by DRVID and Dual-Frequency Range and Doppler Radio Metric Data," *DSN Progress Report 42-37*, pp. 43-54, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1977.

4-45. Shapiro, I. I., et al., "The Viking Relativity Experiment," *Journal of Geophysical Research*, Vol. 82, pp. 4329-4334, Sept. 30, 1977.

4-46. Anderson, J. D., et al., "Experimental Test of General Relativity Using Time-Delay Data for Mariner 6 and Mariner 7," *Astrophysical Journal*, Vol. 200, pp. 221-223, Aug. 15, 1975.

4-47. Volland, H., et al., "Helios-1 Faraday Rotation Experiment: Results and Interpretations of the Solar Occultations in 1975," *Journal of Geophysics*, Vol. 42, No. 6, pp. 659-672, 1977.

4-48. Tang, C. H., Boak, T. I. S., III, and Grossi, M. D., "Bistatic Radar Measurements of the Electrical Properties of the Martian Surface," *Journal of Geophysical Research*, Vol. 82, No. 28, pp. 4305-4315, Sept. 30, 1977.

4-49. Dennison, B., et al., "Deflection of Polarized Radiation: Relative Phase Delay Technique," *Nature*, Vol. 273, No. 5657, pp. 33-35, May 4, 1978.

Chapter 5
Telemetry System

Joseph H. Yuen, Dariush Divsalar, Peter W. Kinman, and Marvin K. Simon

5.1 Introduction

A telemetry system conveys data collected by scientific instruments, as well as engineering information on the status of spacecraft, to the Deep Space Network (DSN). These days, telemetry systems are entirely digital. Not since the Rangers and the Lunar Orbiters of the 1960s has an analog telemetry system been designed into a deep-space mission. Of course, an analog information source can be made compatible with a digital telemetry system through the processes of sampling and quantization. Accordingly, this chapter addresses the means of telemetering digital data from a spacecraft to the DSN.

5.2 Description of the Telemetry System

A typical deep-space telemetry system is illustrated in Fig. 5-1. The data are encoded for energy efficiency, then used to phase-modulate a square-wave subcarrier, which in turn phase-modulates a microwave sinusoidal carrier. Upon reception by the DSN, the signal is routed from antenna to receiver. The receiver tracks the carrier and translates the signal down to an intermediate frequency (IF) coherent with the station reference. The Subcarrier Demodulator Assembly (SDA) translates the signal down to baseband and demodulates the subcarrier.

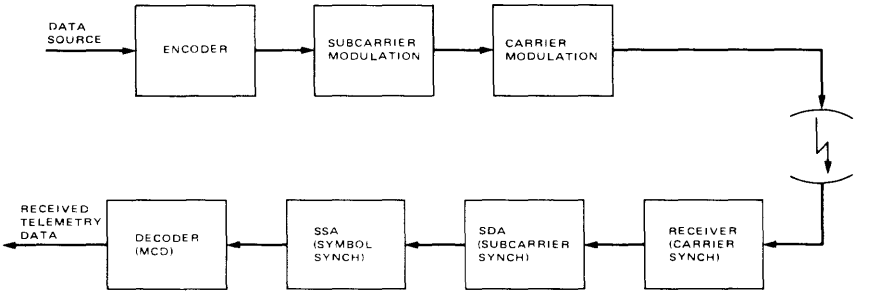


Fig. 5-1. Typical deep space telemetry system

The output of the SDA is synchronized by the Symbol Synchronizer Assembly (SSA), then decoded by a Maximum Likelihood Convolutional Decoder (MCD) if the spacecraft uses a convolutional code.

5.2.1 Deep Space Telemetry Phase-Shift-Keying Modulation

A class of modulation schemes has proven best for the deep-space channel. The most typical telemetry modulation for deep space is to phase-shift-key the data onto a square-wave subcarrier and then phase-modulate onto a sinusoidal carrier (PCM/PSK/PM). This is the scheme depicted in Fig. 5-1. However, an alternate scheme phase-modulates the data directly onto the carrier. In this latter case, the data being binary phase-shift-keyed is called PCM/PM. In the abbreviations PCM/PSK/PM and PCM/PM, the acronym PCM means "Pulse Code Modulation" and serves as a reminder that the information is digital. PCM/PSK/PM and PCM/ PM are well suited to the deep-space channel because: (1) phase-shift-keying is the most efficient binary signaling scheme [5-1 and 5-2] and (2) a phase-modulated carrier has a constant envelope.

The constant envelope property is important because the final downlink amplification stage on the typical spacecraft is a traveling-wave tube operated in a nonlinear (saturation) mode. If the modulated downlink carrier does not have a constant envelope, distortion will result.

Before proceeding to the mathematical descriptions of PCM/PSK/PM and PCM/PM, the issue of data formats needs to be addressed. Various data formats are illustrated in Fig. 5-2. With a Non-Return to Zero-Level (NRZ-L) data format, a logical "one" is represented by one level and a logical "zero" by the other. With NRZ-M (Mark), a logical "one" is represented by a change in level and a logical "zero" by no change. NRZ-S (Space) is the same as NRZ-M but with the roles of "one" and "zero" reversed. NRZ-M and NRZ-S are often referred to as differential encoding. It is important to note that differential

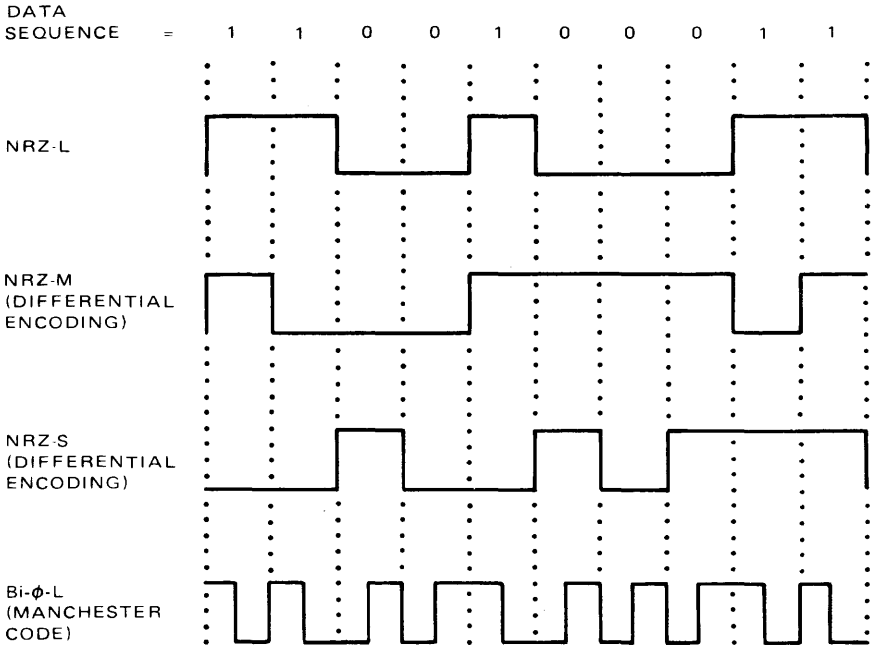


Fig. 5-2. Data formats

encoding is *not* an example of error-correction encoding. With Bi-Phase-Level (Bi-ϕ-L), a logical "one" is represented by one cycle of a square wave and a logical "zero" by one cycle of a square wave with reversed polarity. Bi-ϕ-L is often called a Manchester code; but, again, it is *not* an example of an error-correcting code. Bi-ϕ-L is identical to NRZ-L when the latter's data are phase-reversal-keyed (i.e. phase-shift-keyed with a signaling level of $\pm 90^\circ$) onto a square-wave subcarrier whose cycle period equals the bit period, so that a transition occurs at mid-bit. Thus, PCM/PM with Bi-ϕ-L data formatting is a special case of PCM/PSK/PM with NRZ-L data formatting.

A modulated carrier, in the cases of both PCM/PSK/PM and PCM/PM, may be (assuming nothing else modulates the carrier) represented mathematically by

$$s_T(t) = \sqrt{2P} \sin \left[\omega_0 t + \sum_{i=1}^{M} \theta_i s_i(t) \right] \qquad (5.2\text{-}1)$$

where P = total signal power; ω_0 = carrier angular frequency; θ_i = modulation index, $i = 1, 2, \cdots, M$; M = total number of subcarriers or data channels; and

$s_i(t)$ represents a normalized data sequence (in the case of PCM/PM) or a normalized modulated square-wave subcarrier (in the case of PCM/PSK/PM). In the latter case, the telemetry is invariably phase-*reversal*-keyed (i.e., phase-shift-keyed with signaling levels of $\pm 90^\circ$) onto the subcarrier. Furthermore, phase-reversal-keying a square wave with a data sequence is the same as multiplying the square wave by the data sequence. So

$$s_i(t) = \begin{cases} d_i(t), & \text{PCM/PM} \\ d_i(t)\ \mathrm{Sin}\ \omega_{sci}t, & \text{PCM/PSK/PM} \end{cases} \tag{5.2-2}$$

and

$$s_i(t) = \pm 1, \quad i = 1, 2, \cdots, M \tag{5.2-3}$$

where $d_i(t) = \pm 1$ is the normalized data sequence (one of the waveforms of Fig. 5-2), and $\mathrm{Sin}\ \omega_{sci}t = \pm 1$ is the square-wave subcarrier. Using trigonometric identities with the fact that $s_i(t) = \pm 1$, $i = 1, 2, \cdots, M$, (5.2-1) may be expanded for $M = 1$ as

$$s_T(t) = \sqrt{2P}\cos\theta_1 \sin\omega_0 t + \sqrt{2P}\, s_1(t) \sin\theta_1 \cos\omega_0 t \tag{5.2-4}$$

If $0^\circ < \theta < 90^\circ$, this phase-modulated carrier comprises a pilot tone (residual carrier) and a Double-Sideband (DSB) modulated carrier. A system with $\theta < 90^\circ$ is called a *residual* carrier system. A system with $\theta = 90^\circ$ is called a *suppressed* carrier system. A residual carrier receiver employs a phase-locked loop to track the pilot tone and provide a coherent reference for demodulating the DSB-modulated carrier [5-3]. A Costas loop may be used as a suppressed carrier receiver [5-3]. Up until now, all deep-space-probe–DSN digital telemetry systems have been of the residual carrier type.

For $M = 2$, (5.2-1) may be expanded as

$$\begin{aligned} s_T(t) = {} & \sqrt{2P}\cos\theta_1 \cos\theta_2 \sin\omega_0 t \\ & + \sqrt{2P}\, s_1(t) \sin\theta_1 \cos\theta_2 \cos\omega_0 t \\ & + \sqrt{2P}\, s_2(t) \cos\theta_1 \sin\theta_2 \cos\omega_0 t \\ & - \sqrt{2P}\, s_1(t)\, s_2(t) \sin\theta_1 \sin\theta_2 \sin\omega_0 t \end{aligned} \tag{5.2-5}$$

Similar expansions may be carried out for the cases $M \geqslant 3$.

When we have a single data channel, as in (5.2-4), the first term is the carrier, while the second term is the data channel. Hence the modulation index θ_1 has allocated the total power P in the transmitted signal $s_T(t)$ to the carrier and to the data channel, where the carrier power and the data power are, respectively,

$$\begin{aligned} P_c &= P\cos^2\theta_1 \\ P_D &= P\sin^2\theta_1 \end{aligned} \tag{5.2-6}$$

When we have two data channels, as in (5.2-5), the first term is the carrier component, the second term is the modulated subcarrier $s_1(t)$ component, the third term is the modulated subcarrier $s_2(t)$ component, and the fourth term is the cross-modulation loss component. Hence, the corresponding powers in these four components are allocated by the modulation indices θ_1 and θ_2, with

$$\begin{aligned} P_c &= P\cos^2\theta_1\cos^2\theta_2 \\ P_{D1} &= P\sin^2\theta_1\cos^2\theta_2 \\ P_{D2} &= P\cos^2\theta_1\sin^2\theta_2 \\ P_M &= P\sin^2\theta_1\sin^2\theta_2 \end{aligned} \tag{5.2-7}$$

Similar interpretations can be given to the cases when $M \geqslant 3$. Table 5-1 lists the power allocations for phase modulation using square-wave subcarriers.

5.2.2 Uncoded Quadriphase-Shift-Keying (QPSK) Modulations

5.2.2.1 QPSK and staggered QPSK. The QPSK and SQPSK modulations are basically dual binary systems. An input binary bit sequence $\{d_k\}$ arrives at a rate of $1/T$ baud (bits/sec) and is separated into two bit streams consisting, respectively, of the even and odd bits in the original stream. An example of this procedure is illustrated in Fig. 5-3, where $d_k(t)$, $d_I(t)$, and $d_Q(t)$ are the binary-valued (± 1) waveforms corresponding to the three bit sequences. The inphase and quadrature pulse trains modulate two carriers which are in phase quadrature and the sum $s(t)$, the modulated QPSK signal, can be represented as

$$s(t) = \sqrt{P}\,d_I(t)\cos\omega_0 t + \sqrt{P}\,d_Q(t)\sin\omega_0 t \tag{5.2-8}$$

The two terms in (5.2-8) represent two BPSK signals and can be demodulated and detected independently due to the orthogonality of the carriers $\cos\omega_0 t$ and $\sin\omega_0 t$.

Table 5-1. Power allocations for phase modulation using square-wave subcarriers

Type of channel	Power allocation for each component
Single data channel	$P_c = P \cos^2 \theta_1$
	$P_D = P \sin^2 \theta_1$
Two data channels	$P_c = P \cos^2 \theta_1 \cos^2 \theta_2$
	$P_{D1} = P \sin^2 \theta_1 \cos^2 \theta_2$
	$P_{D2} = P \cos^2 \theta_1 \sin^2 \theta_2$
Three data channels	$P_c = P \cos^2 \theta_1 \cos^2 \theta_2 \cos^2 \theta_3$
	$P_{D1} = P \sin^2 \theta_1 \cos^2 \theta_2 \cos^2 \theta_3$
	$P_{D2} = P \cos^2 \theta_1 \sin^2 \theta_2 \cos^2 \theta_3$
	$P_{D3} = P \cos^2 \theta_1 \cos^2 \theta_2 \sin^2 \theta_3$

The SQPSK signal can be represented similarly to (5.2-8), the only difference being in the alignment of the two bit streams. The odd and even bit streams, transmitted at the rate of $1/2T$ baud, are synchronously aligned in QPSK (Fig. 5-3), such that their transitions coincide. SQPSK modulation is obtained by a shift or offset in the relative alignments of $d_I(t)$ and $d_Q(t)$ by an amount equal to a bit time T. Figure 5-4 shows the offset, and the following equation defines the SQPSK signal.

$$s(t) = \sqrt{P}\, d_I(t + T) \cos \omega_0 t + \sqrt{P}\, d_Q(t) \sin \omega_0 t \qquad (5.2\text{-}9)$$

Using a well-known trigonometric identity (5.2-8) can also be written as

$$s(t) = \sqrt{2P} \cos (\omega_0 t + \theta(t)) \qquad (5.2\text{-}10)$$

where, as shown in the state space diagram of Fig. 5-5, $\theta(t)$ can take on the values 45°, 135°, 225°, and 315°, corresponding to the four combinations of $d_I(t)$ and $d_Q(t)$.

In QPSK, due to the coincident alignment of $d_I(t)$ and $d_Q(t)$, the carrier phase can change only once every $2T$. The carrier phase over any $2T$ interval is any one of the four phases shown in Fig. 5-5, depending on the pair of values for $\{d_I(t), d_Q(t)\}$. In the next $2T$ interval, if neither bit stream changes sign, then the carrier phase remains the same. If one component ($d_I(t)$ or

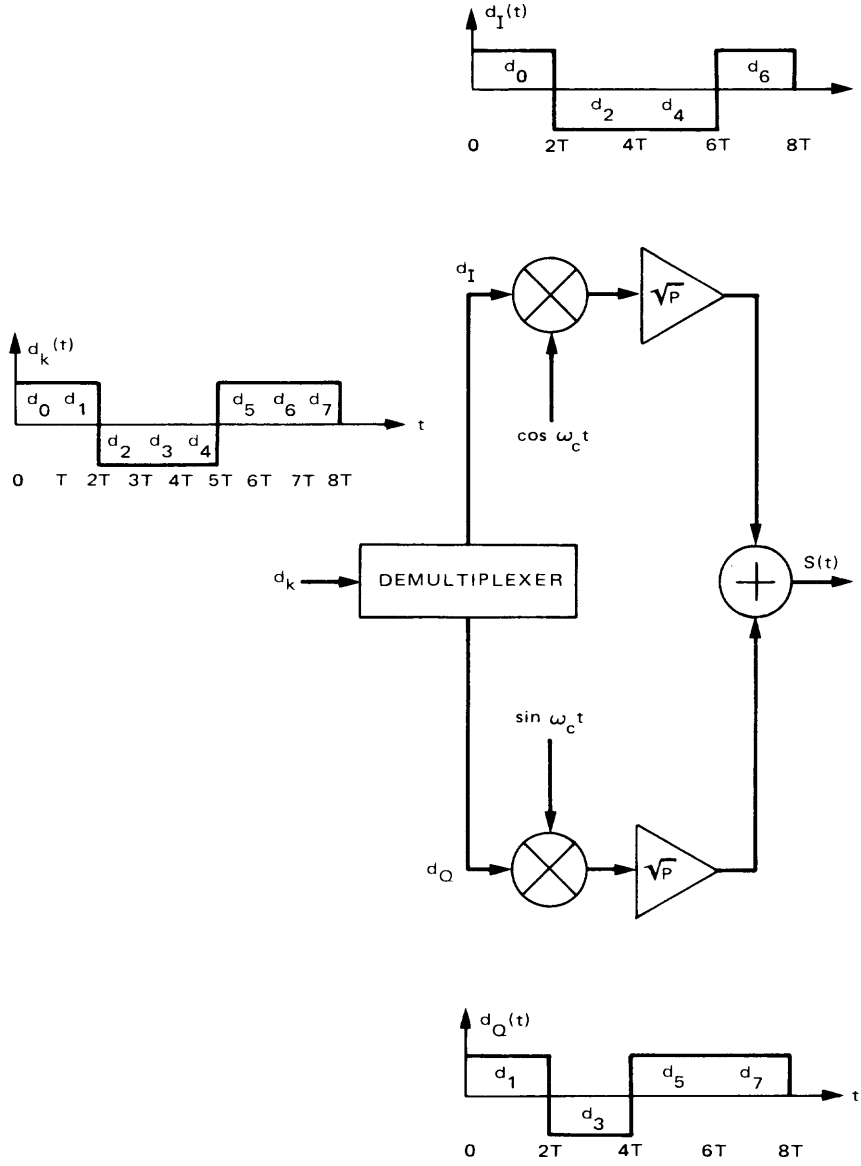


Fig. 5-3. QPSK modulator

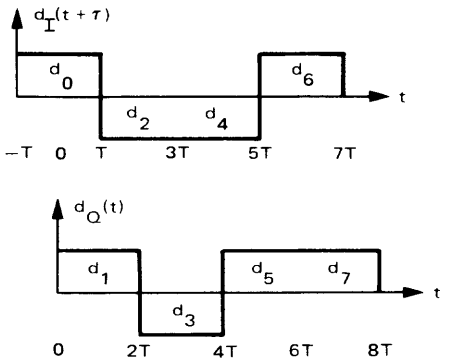


Fig. 5-4. Staggering of data streams in SQPSK

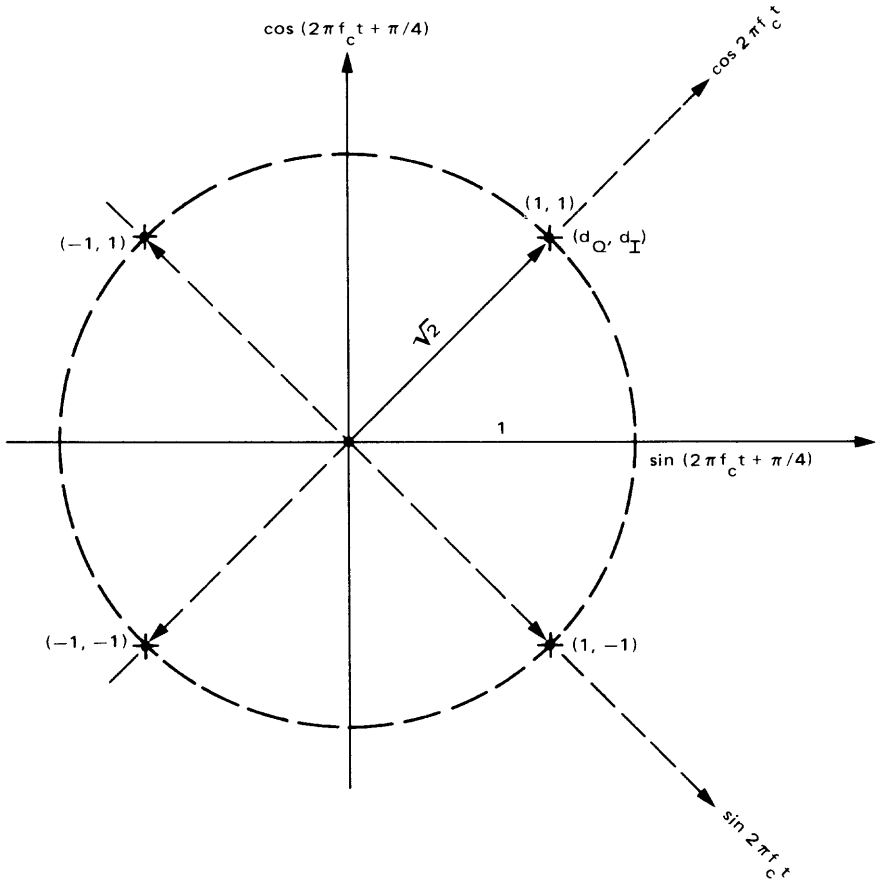


Fig. 5-5. Signal space diagram for QPSK and SQPSK

$d_Q(t)$) changes sign, a phase shift of ±90° occurs. A change in both components results in a phase shift of 180°. Figure 5-6(a) shows a typical QPSK signal waveform for the sample bit streams shown in Fig. 5-3.

In SQPSK, the binary components cannot change states simultaneously. One component has transitions in the middle of the other symbol and hence only one component can switch at a time. This eliminates the possibility of 180° phase changes. Furthermore, these changes are limited to 0°, ±90°, every T seconds. Figure 5-6(b) shows a typical SQPSK waveform for the example bit streams in Fig. 5-4. When an SQPSK signal undergoes band-limiting, the resulting intersymbol interference causes the envelope to droop slightly in the region of 90° phase transitions. Since 180° phase transitions have been avoided, the envelope does not go to zero as it does in the band-limited QPSK case.

The advantages of QPSK or SQPSK over BPSK is that, with both data rate and power being fixed, QPSK or SQPSK achieve the same power efficiency as BPSK using only half the bandwidth.

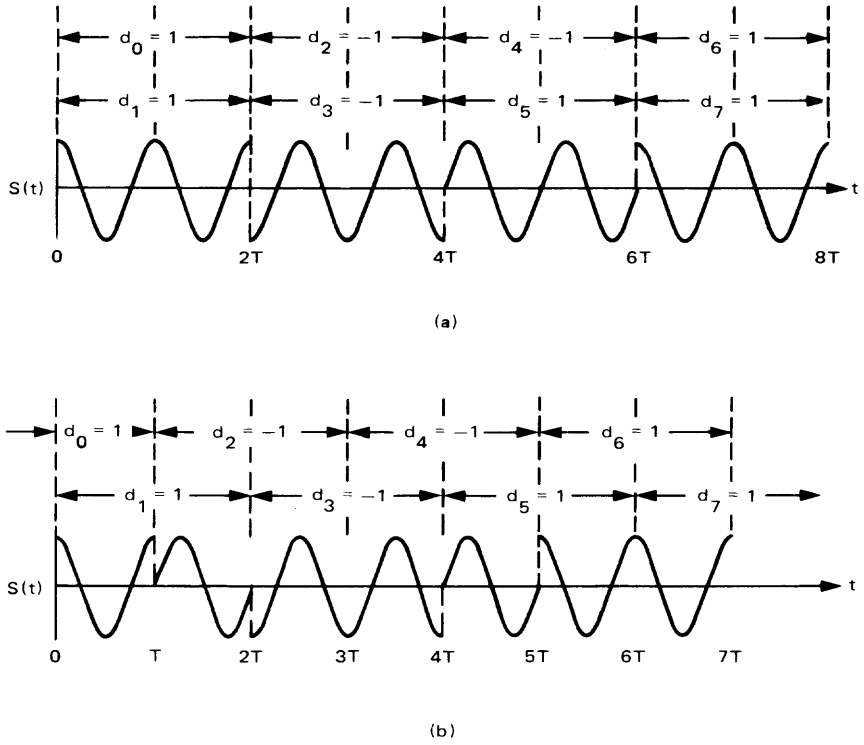


Fig. 5-6. (a) QPSK waveform; (b) SQPSK waveform

With symmetrical bandpass Gaussian noise, the inphase and quadrature noise terms are independent and the bit error probability performance of QPSK is the same as that of BPSK.

Since the phase changes in SQPSK occur in smaller step sizes than in QPSK, SQPSK would tend to suffer less degradation due to bandpass filtering than QPSK.

5.2.2.2 Unbalanced quadriphase-shift-keying (UQPSK). The UQPSK modulation is similar to QPSK modulation. Here two independent binary data streams, $d_1(t)$ and $d_2(t)$, with different rates and powers, modulate the two quadrature components of the carrier. The UQPSK modulator (suppressed carrier) is shown in Fig. 5-7.

The output of the modulator can be represented as

$$s(t) = \sqrt{2P_1}\, d_1(t) \sin(\omega_0 t + \theta_0) + \sqrt{2P_2}\, d_2(t) \cos(\omega_0 t + \theta_0) \tag{5.2-11}$$

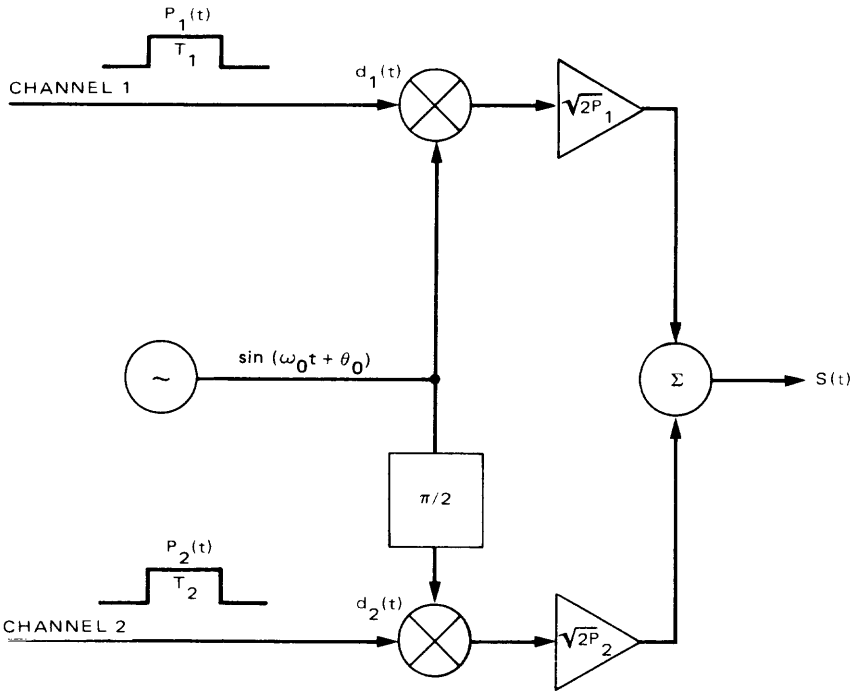


Fig. 5-7. UQPSK modulator

where P_i, $i = 1, 2$ is the channel power, θ_0 is the unknown carrier phase, and ω_0 is the carrier frequency in rad/sec. The data waveforms can be expressed as

$$d_i(t) = \sum_{n=-\infty}^{\infty} a_{ni} p_i (t - (n - 1) T_i - \epsilon_i), \quad i = 1, 2 \qquad (5.2\text{-}12)$$

where a_{ni}, $i = 1, 2$ are independent binary sequences of ±1's; $p_i(t)$, $i = 1, 2$ is the unit-power, symbol pulse in the ith channel and is defined to be nonzero on the interval $(0, T_i)$, $i = 1, 2$, where T_i is channel symbol time. Since, in general, the two channels are asynchronous (not aligned), we assume that the pulse epoch ϵ_i, $i = 1, 2$ is uniformly distributed in the interval $(0, T_i)$, $i = 1, 2$, and that ϵ_1 is independent of ϵ_2.

5.2.3 Description of the Subcarrier Demodulation Assembly (SDA)

This section follows the treatment in [5-4]. Another useful reference is [5-5].

The Subcarrier Demodulation Assembly (SDA) is an integral part of the multiple mission telemetry system now incorporated in the DSN. The primary purpose of the SDA is to provide a demodulated noisy data stream to the SSA for detection. Figure 5-1 shows where the SDA fits into the overall communications system. Its inputs are the received radio signal, at an intermediate frequency, and a phase-coherent reference, both of which are furnished by the receiver/exciter subsystem. The following is a brief simplified mathematical description of how the SDA works.

With reference to Fig. 5-1, let the input to the subcarrier modulator be $d(t)$, the binary waveform representing the data bits; then the output of the modulator can be written as:

$$d(t) \, \mathrm{Sin} \, \omega_{sc} t$$

where ω_{sc} is 2π times the subcarrier frequency in rad/sec and Sin represents a square wave having the same zero crossings as $\sin \omega_{sc} t$. Let both $d(t)$ and Sin have values of +1 or −1 (see (5.2-3)). The output of the carrier phase modulator can then be written as:

$$\sqrt{2P} \sin [\omega_0 t + \Delta d(t) \, \mathrm{Sin} \, \omega_{sc} t]$$

where ω_0 is the carrier angular frequency, Δ is the modulation index, which determines the phase deviation of the modulation (compare with (5.2-1); we

have set $\theta_1 = \Delta$). This signal is then transmitted through space and received by one of the DSN stations. The received signal is a replica of the transmitted signal with the addition of noise and can be written as:

$$\sqrt{2P} \sin [\omega_0 t + \Delta d(t) \operatorname{Sin} \omega_{sc} t + \theta] + n''(t)$$

where θ is an arbitrary phase angle and $n''(t)$ is the additive noise.

The receiver now translates this signal in frequency to an intermediate frequency (IF) so that the IF signal passing to the SDA is just:

$$\sqrt{2P} \sin [\omega_{IF} t + \Delta d(t) \operatorname{Sin} \omega_{sc} t + \theta] + n'(t)$$

The receiver has a phase-locked loop (PLL) that tracks the frequency and phase of this signal and provides a reference signal to the SDA, which is phase-coherent with the received signal (see Chapter 3). Its form is:

$$\sin [\omega_{IF} t + \hat{\theta}]$$

where $\hat{\theta}$ is the PLL estimate of θ. The cosine version of this signal is sent to the SDA. For the purpose of discussing the SDA, the assumption will be made that $\theta \cong \hat{\theta} = 0$. Also the amplitude $\sqrt{2P}$ is not important in the discussion, so it can be taken as unity. Thus, the two signals entering the SDA from the receiver are:

IF Signal:

$$s(t) = \sin [\omega_{IF} t + \Delta d(t) \operatorname{Sin} \omega_{sc} t] + n'(t)$$

Coherent Reference:

$$r_{IF}(t) = \cos \omega_{IF} t$$

It is the goal of the SDA to take these two signals and furnish the following signal to the SSA:

$$d(t) + n(t)$$

The signal $s(t)$ can be rewritten in a more convenient form. First, expanding by a trigonometric identity, $s(t)$ is (ignoring noise):

$$s(t) = \sin \omega_{IF} t \cos [\Delta d(t) \operatorname{Sin} \omega_{sc} t]$$

$$+ \cos \omega_{IF} t \sin [\Delta d(t) \operatorname{Sin} \omega_{sc} t]$$

But $d(t)$ Sin $(\omega_{sc}t)$ takes on only the values of +1 or -1 so that

$$\cos\ [\Delta d(t)\ \text{Sin}\ \omega_{sc}t] = \cos\ \Delta$$

$$\sin\ [\Delta d(t)\ \text{Sin}\ \omega_{sc}t] = d(t)\ \text{Sin}\ \omega_{sc}t \sin\ \Delta$$

Thus $s(t)$ is just

$$s(t) = \cos\ \Delta \sin\ \omega_{IF}t + d(t)\ \text{Sin}\ \omega_{sc}t \sin\ \Delta \cos\ \omega_{IF}t$$

Since Δ is only a constant, the first term is just the carrier component. The second term contains all the sideband modulation components displaced spectrally on either side of the carrier frequency.

A basic principle in coherent demodulation is the idea of translating a frequency by "multiplying" it by a coherent signal. This multiplication is variously referred to as heterodyning, conversion, mixing, or phase detection, depending on the application, but mathematically it can be thought of as multiplying two signals of the same or different frequencies. Only three trigonometric identities will be needed for the rest of the discussion:

$$\cos x \cos y = 1/2\ [\cos\ (x - y) + \cos\ (x + y)]$$

$$\sin x \sin y = 1/2\ [\cos\ (x - y) - \cos\ (x + y)]$$

$$\sin x \cos y = 1/2\ [\sin\ (x - y) + \sin\ (x + y)]$$

As an example, suppose a radio signal is received having the form:

$$d(t) \cos\ \omega_0 t$$

and it is desired to recover the modulation, $d(t)$. If a coherent reference, $\cos\ \omega_0 t$, can be generated in the receiver, then the signal can be coherently demodulated by "multiplying" the signal and the reference together, namely,

$$d(t) \cos\ \omega_0 t \cos\ \omega_0 t$$

From the first trigonometric identity this is

$$1/2\ d(t)\ [\cos\ 0 + \cos\ 2\omega_0 t] = \frac{d(t)}{2} + \frac{d(t)}{2} \cos\ 2\omega_0 t$$

The second term is a high-frequency term that can easily be eliminated by a low-pass filter, thus leaving simply modulation $d(t)$ times a scale factor.

With this basic principle, the operation of the SDA is more easily understood. Figure 5-8 shows the present functional block diagram of the SDA in a simplified form. The Mark IV configuration is somewhat different, but functionally the same. The configuration shown accomplishes two things simultaneously:

(1) The data, $d(t)$, is coherently demodulated and sent (with noise) to the SSA

(2) A tracking loop provides the coherent reference needed for demodulation

Note that item (1) depends on item (2). Moreover, in the type of tracking loop used in the SDA, an estimate of the data, $d(t)$, is needed. Thus, item (2) depends on item (1).

In explaining the loop operation, the assumption will be made that the loop is in lock, that is (see Fig. 5-8),

$$r_{sc}(t) = \text{Sin } \omega_{sc} t \tag{5.2-13}$$

and also that the rest frequency of the voltage-controlled oscillator (VCO) is equal to the subcarrier frequency, ω_{sc}.

The upper branch of the loop provides the data demodulation. Ignoring noise, the output of the first multiplier is

$$s(t)\, r_{sc}(t) = \cos \Delta \text{ Sin } \omega_{sc} t \sin \omega_{IF} t$$

$$+ d(t) \text{ Sin } \omega_{sc} t \text{ Sin } \omega_{sc} t \sin \Delta \cos \omega_{IF} t \tag{5.2-14}$$

The first term has frequency components greater than $\omega_{sc} + \omega_{IF}$ and less than $\omega_{sc} - \omega_{IF}$, as can be seen from the trigonometric identity. These components will be filtered by the bandpass filter (BPF). Also, in the second term, Sin $\times$ Sin equals unity for all time so that the output of the top BPF is

$$d(t) \sin \Delta \cos \omega_{IF} t$$

As in the above example, this signal is coherently demodulated by multiplying by $r_{IF}(t) = \cos \omega_{IF}(t)$. The output of the coherent demodulator is thus

$$\frac{\sin \Delta}{2}\, d(t)\, [1 + \cos 2\omega_0 t]$$

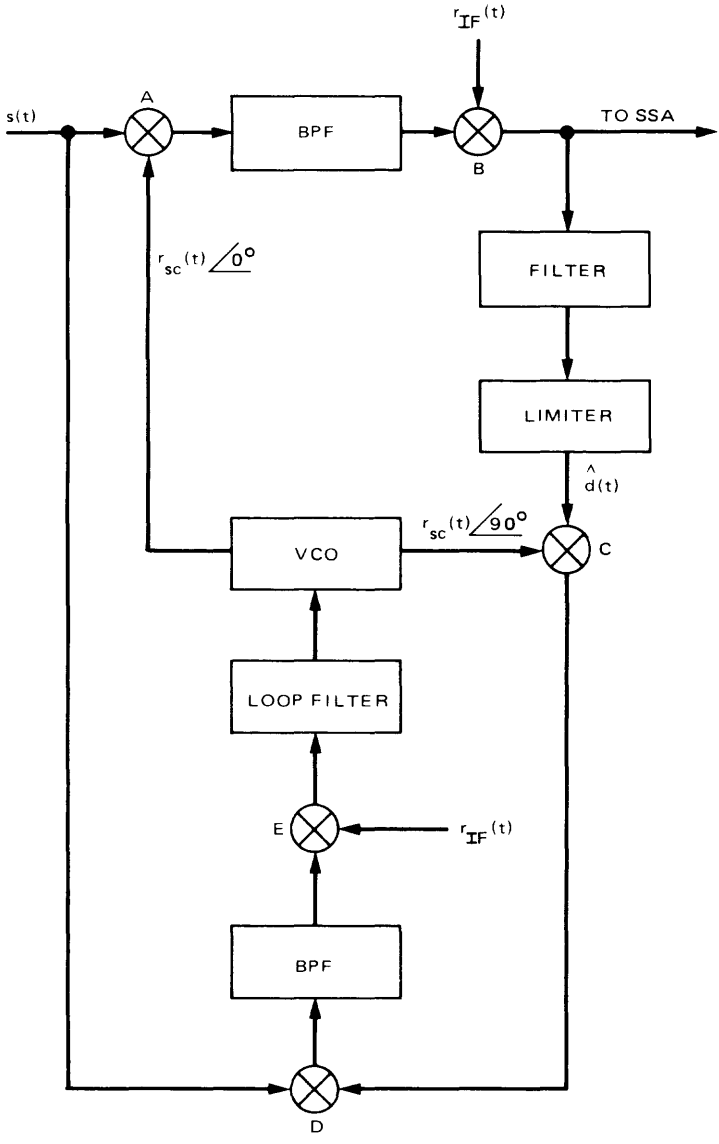


Fig. 5-8. Block diagram of SDA

This signal is low-pass filtered to obtain $d(t)$ and is sent to the SSA for detection and synchronization.

Proceeding along in the loop, the above signal is filtered to optimize the signal-to-noise ratio and then hard-limited[1] to produce a replica of the data stream, $\hat{d}(t)$. This concludes the data-detection branch of the loop. The rest of the loop provides the coherent subcarrier reference.

At multiplier C the output is just

$$\hat{d}(t) \text{ Cos } \omega_{sc} t$$

where Cos $\omega_{sc} t$ is a square wave with the same zero crossings as cos $\omega_{sc} t$. This signal goes to multiplier D, where it is combined with the input signal to produce

$$s(t)\,\hat{d}(t) \text{ Cos } \omega_{sc} t = \hat{d}(t) \cos \Delta \sin \omega_{IF} t \text{ Cos } \omega_{sc} t + d(t)\,\hat{d}(t) \sin \Delta \text{Sin } \omega_{sc} t \text{ Cos } \omega_{sc} t \cos \omega_{IF} t \tag{5.2-15}$$

Again the first term has components outside the bandpass of the BPF. Also, we note that

$$d(t)\,\hat{d}(t) \cong d^2(t) = 1 \quad \text{for all } t \tag{5.2-16}$$

Thus, the important output of multiplier D is

$$\sin \Delta \text{ Sin } \omega_{sc} t \text{ Cos } \omega_{sc} t \cos \omega_{IF} t$$

The term Sin X Cos is just a square wave having a frequency of $2\omega_{sc}$. Thus, the above term has frequency components greater than $\omega_{IF} + 2\omega_{sc}$ and less than $\omega_{IF} - 2\omega_{sc}$. Accordingly, this whole term is filtered by the BPF leaving the output of multiplier E as zero and the input to the VCO as zero. But this is in agreement with our first assumption that the loop was in lock which implies a zero error signal going to the VCO.

In actuality, however, some phase error ϕ will exist between the two subcarrier terms above due to noise. The product of the Sin X Cos will now be a

[1] A hard limiter puts out +V volts for any input signal greater than 0 volts and –V volts for any signal less than 0.

$2\omega_{sc}$ term with a duty cycle not equal to 50%. This gives the product a dc bias which can be called $e(t)$. Thus, the important output of multiplier D is

$$\sin \Delta \, \text{Sin}\, \omega_{sc} t \, \text{Cos}(\omega_{sc} t - \phi) \cos \omega_{IF} t = \sin \Delta \, [e(t) + 2\omega_{sc} \text{ terms}] \cos \omega_{IF} t \tag{5.2-17}$$

Now the output of the BPF is

$$\sin \Delta \, e(t) \cos \omega_{IF} t$$

Multiplying at E by $r_{IF}(t)$ we get

$$\frac{\sin \Delta}{2} \, e(t) \, (1 + \cos 2\omega_{IF} t)$$

The second term is low-pass filtered by the loop filter, leaving only the error signal $e(t)$ going to the VCO. The VCO output frequency is proportional to the input error voltage, and thus any phase error ϕ will produce an error signal $e(t)$, which will tend to correct the phase error and bring it back to 0. This completes the closed loop, which provides a synchronous reference for the coherent demodulation.

5.2.4 Description of the Symbol Synchronizer Assembly (SSA)

This section follows the treatment in [5-6]. Another useful reference is [5-7].

The Symbol Synchronizer Assembly (SSA), which is part of the Telemetry and Command Data Handling Subsystem, has two primary functions: (1) to detect uncoded data or, alternatively, to provide symbol integrations of coded data and (2) to provide the timing or synchronization needed to perform function (1). While the SSA performs a number of other functions, this section will be concerned with a description of how these two functions specifically are accomplished.

Binary data (1's and 0's) can be transmitted over an uncoded telemetry channel as the same sequence of 1's and 0's, which in this case are called bits. If the data are coded by forming a different sequence, the 1's and 0's are called symbols. The bits (or symbols) are represented by a modulating waveform consisting of two levels, $+A$ or $-A$, as shown in Fig. 5-9(a). This signal waveform (referred to as non-return to zero (NRZ) modulation) is transmitted on a radio frequency (RF) carrier and is demodulated (recovered) by the DSN receiver. The received waveform has been corrupted, however, with the addition of noise, as shown in Fig. 5-9(b). The function of the SSA is the extraction of the original signal waveform from this noisy waveform. It is to be

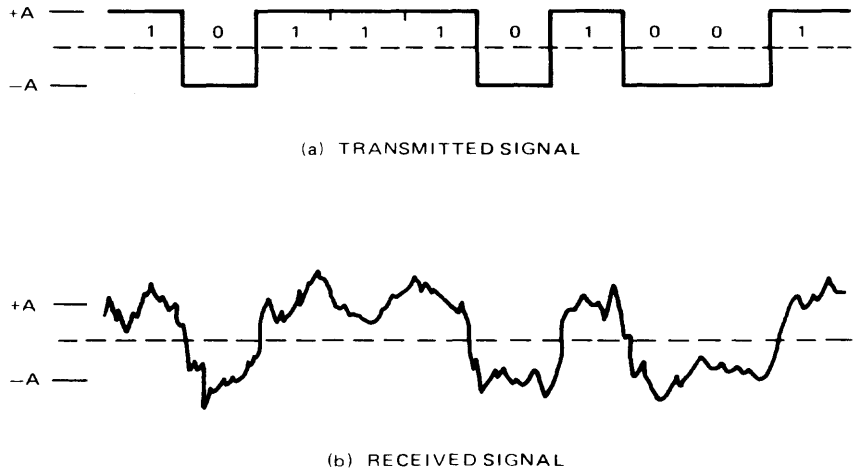


Fig. 5-9. Telemetry signal waveforms

emphasized that the SSA is configured to operate only on an NRZ data format. Thus the use of a Manchester code requires modification to the basic structure.

To understand how the SSA recovers the desired bits, it is necessary to understand the method for detecting a single bit. This is accomplished by integrating the signal over the bit time and choosing a one if the integral is positive or a zero if the integral is negative. This circuit is referred to as an "integrate and dump" circuit because it integrates for a bit time and then resets to zero at the end of each T second interval. Figure 5-10(a) shows the bit and its integral value at the end of the T second integration. Figure 5-10(b) shows the same decision made on a noisy signal. Figure 5-10(c) shows the noise being such that an error is made.

Naturally, a communications link is designed such that the probability of noise-caused errors is low. This is accomplished by maintaining the signal level at a relatively high value compared to the noise level. One assumption that is made, however, is that the beginning of the bit must be known to start the integration properly. An example of integration over a bit time beginning before the actual start of the bit is shown in Fig. 5-11, where the "1" bit is preceded by a "0" bit. The correct decision is made with no noise, but with a small amount of noise the probability is great that the wrong decision will be made, as shown in Fig. 5-11(b). Thus, it can easily be seen that accurate time information as to the start of each bit is necessary to keep the probability of making an error as low as possible. This is the second function of the SSA; i.e., to provide bit (or symbol) synchronization to the data detector circuits.

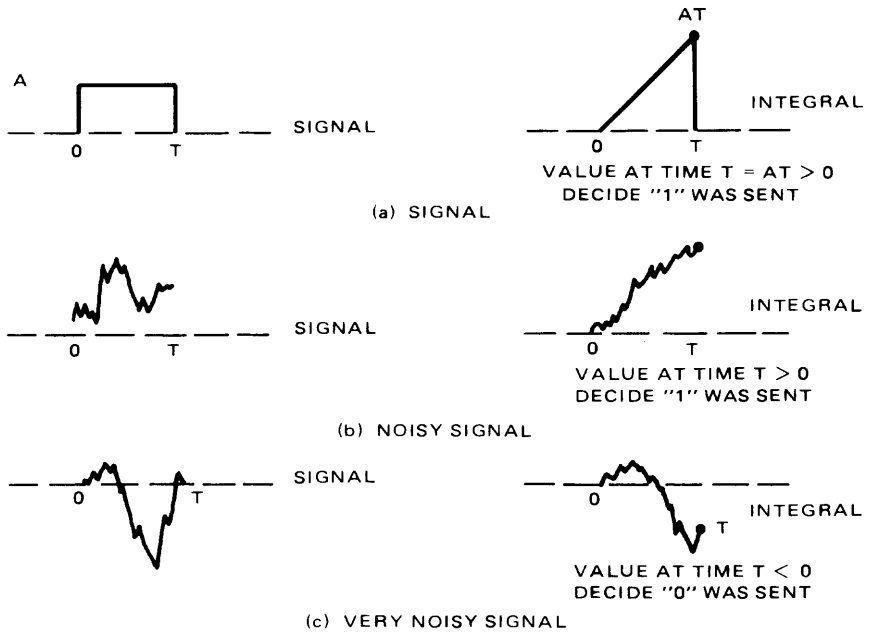


Fig. 5-10. Bit detection

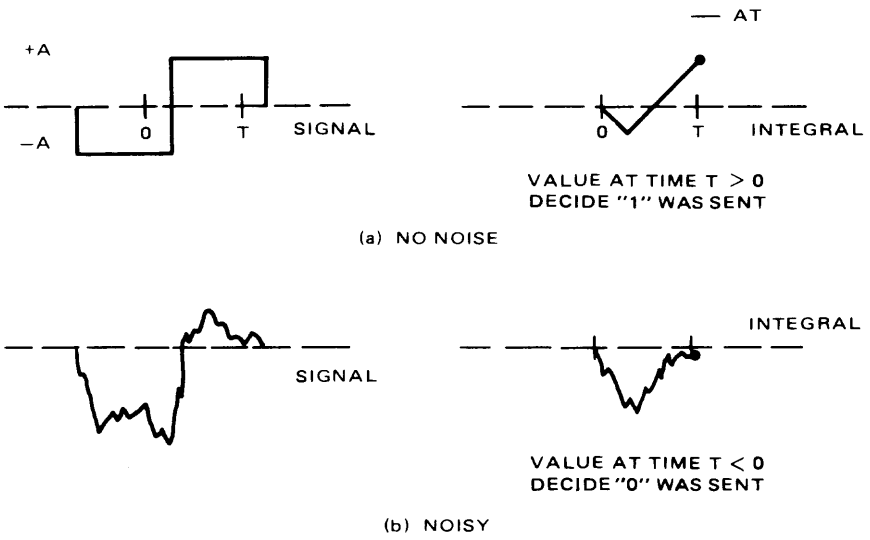


Fig. 5-11. Nonsynchronized bit detection

The functional block diagram of the synchronizing circuitry of the SSA is shown in Fig. 5-12. The in-phase integrate and dump circuit followed by the decision device compose the bit detector described above. This is followed by a transition detector, which outputs a plus or minus one if a transition occurs in the data (e.g., a "1" followed by a "0") and a zero if no transition occurs. The lower mid-phase integrate and dump circuit integrates symmetrically around the estimated data transition point for an integrating period equal to 1/4 the symbol period. The selection of a 1/4-symbol-period integration interval for the lower (quadrature) integrate-and-dump represents a good compromise between tracking and loss of lock performance. This output is delayed so that the outputs of the upper and lower branches appear at the multiplier simultaneously. The correct timing is shown in Fig. 5-13. The purpose of the loop filter is to smooth out the error signal from the multiplier, thus leaving only the slowly varying component of the error signal. The timing generator is essentially a voltage-controlled oscillator (VCO) whose frequency is controlled by the input voltage. The timing generator controls the timing of the integrate and dump circuits shown in Fig. 5-13.

For the purposes of explaining the operation of the loop, the loop can be thought of as being in lock with the timing exactly correct and the timing generator having the same frequency as the incoming data. If the incoming signal has no noise then the mid-phase integrator will integrate to zero over a

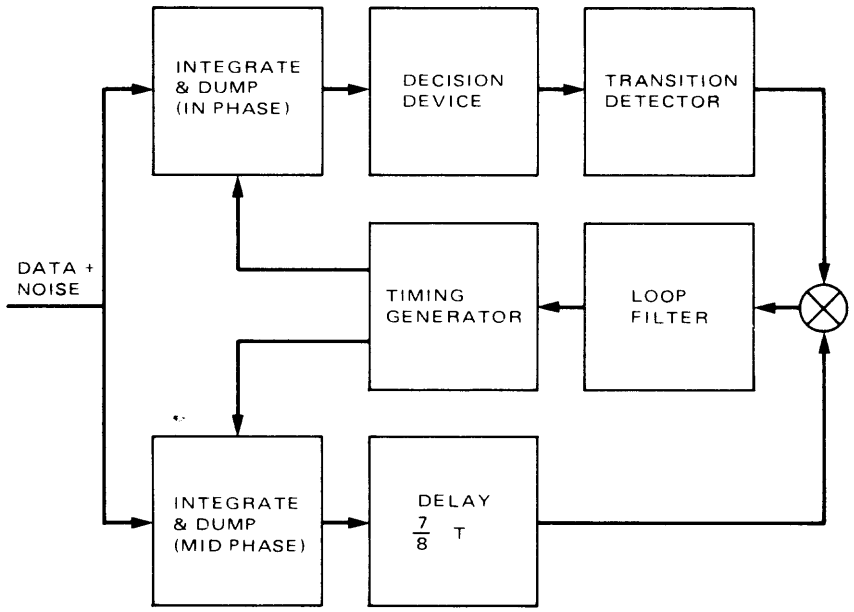


Fig. 5-12. Block diagram of the SSA bit synchronizer

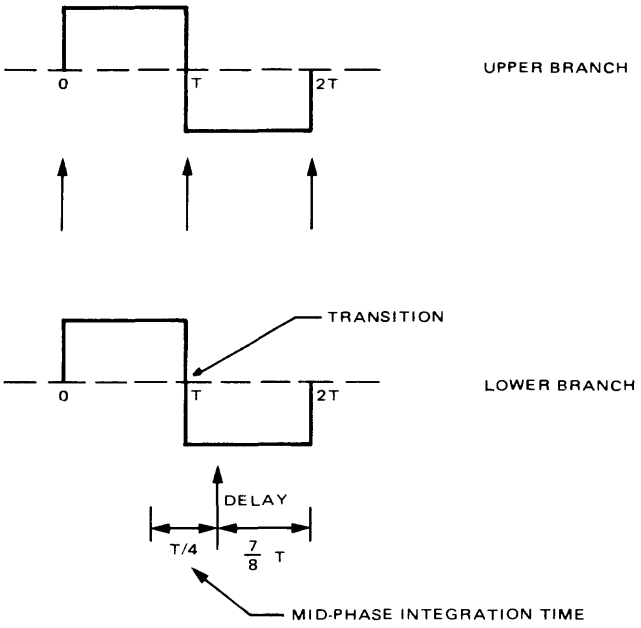


Fig. 5-13. Bit synchronization

transition. Thus, for an occurring transition the multiplier output will be zero, i.e., no error signal to the VCO. If no transition occurs, the output of the transition detector will be zero and again the multiplier output or error signal will be zero. If the timing had originally been slightly off, the mid-phase integrator would not integrate to zero and hence would produce an error signal at the input to the VCO. The sign of this error signal would be such that the error voltage would change the frequency slightly so that the timing signals would be aligned properly and the error voltage would again be zero.

The key point in the operation of the synchronizer is that the error signal is always zero when no transitions occur; i.e., no correction to the timing can occur if no transitions occur. The error voltage is only produced when a transition occurs and the mid-phase integrator does not integrate to zero. For this reason the following situation could occur. The incoming signal could have a doppler rate; i.e., its frequency could be changing with time. If a long string of no transitions occurs in the data, the VCO frequency will remain constant while the incoming signal will be changing frequency. Thus, the loop cannot track the incoming signal during this period. When transitions do start occurring again, many errors in detection will be made while the loop tries to reacquire the new frequency and timing. This condition puts restrictions on the data sequence as to the transition density of the data, i.e., how often transitions occur on the average, and what the longest string of no transitions

is that can occur in the data stream. Naturally, the greater the transition density, the greater the ability of the loop to track doppler rates.

A final note is that this section gives a functional description of the SSA; the actual mechanization of the SSA functions described here is not as shown in Fig. 5-12. The synchronizing loop is in fact a hybrid design of both digital and analog components and functions. Also, the SSA performs a number of other functions not described here.

5.3 Uncoded Phase-Shift-Keying Telemetry System Performance

The performance of a digital telemetry system is measured as a bit error probability, P_b, or bit error rate (BER). The BER, in turn, depends on the bit energy-to-noise spectral density ratio E_b/N_0 and the inefficiencies of the demodulation/detection process. The bit energy-to-noise spectral density ratio is defined as

$$E_b/N_0 = \frac{S}{N_0 R} \tag{5.3-1}$$

where

S = received data power

N_0 = one-sided noise spectral density

R = information bit rate

The demodulation/detection inefficiencies are dominated by the imperfection with which the receiver achieves carrier, subcarrier, and symbol synchronization.

If one assumes no demodulation/detection inefficiencies (implying perfect carrier, subcarrier, and symbol synchronization at the receiver), the resulting functional dependence of the bit error probability, P_b, upon E_b/N_0 is termed the baseline performance. In this chapter the baseline performance is frequently denoted by $f(\cdot)$:

$$P_b = f(E_b/N_0) \tag{5.3-2}$$

Of course, $f(\cdot)$ is different for different coding/modulation schemes. In some cases (see, for example, subsection 5.3.1), $f(\cdot)$ can be determined theoretically. Otherwise, $f(\cdot)$ must be determined by simulation or a combination of

simulation and theory. In all cases, the channel is modeled as additive white Gaussian—an excellent model for the deep-space channel.

5.3.1 One-Way Telemetry Performance

In this section, the baseline performance $f(\cdot)$ is found for the case of no error-correcting coding and no differential encoding, i.e., the uncoded case. The performance of a receiver for a binary digital signal depends on the phase-coherence of the receiver and the type of modulation employed at the transmitter. A transmitted signal is expressed as

$$s_k(t) = A_k(t) \sin [\omega_0 t + m_k(t) + \theta(t)], \; 0 < t < T_s, \; k = 1, 2 \tag{5.3-3}$$

where T_s is the symbol period, ω_0 is the signal frequency, and $A_k(t)$ and $m_k(t)$ are determined by the modulation scheme, whereas the ability of the receiver to track $\theta(t)$ is a measure of its coherence.

The received waveform at the receiver, with additive noise $n_i(t)$, is

$$y(t) = s_k(t) + n_i(t)$$

The optimum detector with minimum error probability is a matched filter, or correlation detector. It is this type of detector that is employed by the Deep Space Network (DSN) to detect transmitted data. (For discussions of matched filter detection, see [5-8] and [5-9]). This correlation operation can be performed by a multiplier-integrator combination with readout at time $t = T_s$ where T_s is the duration of the kth transmitted symbol, or by a matched filter with readout at $t = T_s$. A functional diagram of a coherent correlation detector for binary signals that uses a carrier reference for detection is illustrated in Fig. 5-14, where both the perfect reference and the noisy reference cases are included. This detector model is predicated on the assumption that each transmitter symbol is equally likely, and that each symbol has equal energy.

The energy of each symbol out of the detector is E_s. A measure of the correlation of the two signals is the normalized inner product

$$\lambda = \frac{1}{E_s} \int_0^{T_s} s_1(t) \, s_2(t) \, dt \tag{5.3-4}$$

Let the received signal be represented by

$$y(t) = \sqrt{2P} \, m_k(t) \cos [\omega_0 t + \theta(t)] + n_i(t) \tag{5.3-5}$$

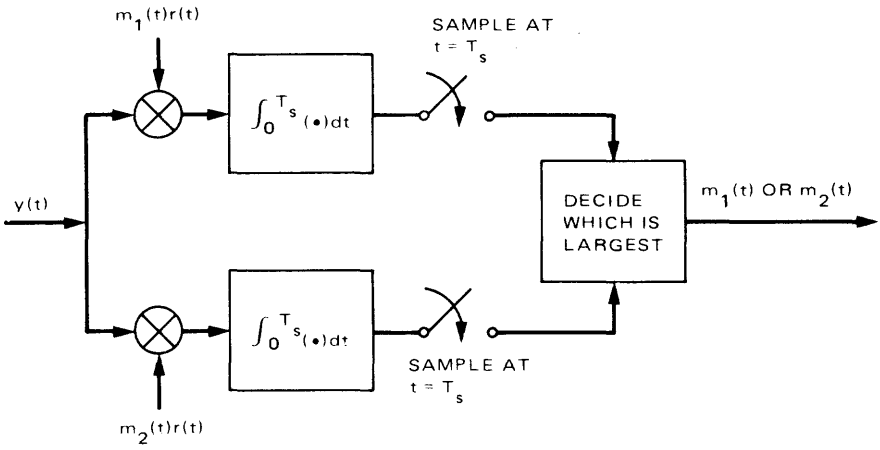


Fig. 5-14. Coherent correlation detector; perfect reference case $\theta(t) = \hat{\theta}(t)$, noisy reference case $\theta(t) \neq \hat{\theta}(t)$

where $\theta(t)$ is the random phase shift, a slowly varying quantity, introduced by the channel; $n_i(t)$ is the additive white Gaussian channel noise with one-sided spectral density N_0 watts/Hz; and $m_k(t)$ represents the ±1 modulation corresponding to the kth signal in the transmitter signaling alphabet. The signal $y(t)$ is demodulated at the receiver.

For a perfectly matched filter, the receiver obtains perfect knowledge of the transmitted phase reference $\theta(t)$, and of transmitted symbol transition time. In this case, the carrier reference signal $r(t)$ is exactly in phase with the signal component of the received waveform, i.e.,

$$r(t) = \sqrt{2}\cos\left[\omega_0 t + \hat{\theta}(t)\right] \tag{5.3-6}$$

with an estimate of the phase-shift

$$\hat{\theta}(t) = \theta(t) \tag{5.3-7}$$

and the readout instances are known a priori at the receiver. Then, as shown in [5-1] and [5-3] among numerous other references, the symbol error probability is

$$P_s = 1 - \int_{-\infty}^{z} \frac{\exp(-u^2/2)}{\sqrt{2\pi}}\,du = \frac{1}{2}\,\mathrm{erfc}\left(\frac{z}{\sqrt{2}}\right) \tag{5.3-8}$$

where

$$z = \sqrt{\frac{E_s(1 - \lambda)}{N_0}}$$

and

$$\text{erfc}\,(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt \tag{5.3-9}$$

If the transmitted symbols are 1 and -1, they are antipodal ($\lambda = -1$). Antipodal signals are ordinarily transmitted by spacecraft.

In an uncoded PSK telemetry system, the binary symbols are the data bits. If the bits have a signal power S watts and last for $T_s = T_b$ seconds,

$$E_s = E_b = S\,T_b \tag{5.3-10}$$

Thus, using (5.3-10) and $\lambda = -1$, the bit error probability for uncoded PSK transmission over a Gaussian channel with white additive noise of one-sided spectral density N_0 watts/Hz, is

$$P_b = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{5.3-11}$$

In practice, we rarely have perfect knowledge of the transmitted phase reference (i.e., a perfectly matched filter), since the channel continuously changes the phase of the transmitted signal relative to that of the received signal. Thus the receiver usually derives the reference signal, $r(t)$, by means of a phase-locked loop (see Chapter 3), and uses this $r(t)$ in the detection process as though it were noise-free. The reference signal $r(t)$ in (5.3-6) no longer has the condition in (5.3-7). Instead, $r(t)$ is a noisy reference, and there is a phase error process

$$\phi(t) = \theta(t) - \hat{\theta}(t) \tag{5.3-12}$$

The accuracy of the estimate $\hat{\theta}(t)$ is dependent on the power allocated to the residual carrier component. The conditional bit error probability (conditioned on a given ϕ) is [5-3],

$$P_b(\phi) = \frac{1}{2}\text{erfc}\left[\sqrt{\frac{E_b}{N_0}}\, Y(\phi)\right] \tag{5.3-13}$$

where

$$Y(\phi) = \frac{1}{T_b} \int_0^{T_b} \cos \phi(t)\, dt \tag{5.3-14}$$

Then the bit error probability can be obtained by averaging $P_b(\phi)$ over ϕ, i.e.,

$$P_b = \int_{-\pi}^{\pi} P_b(\phi)\, p(\phi)\, d\phi \tag{5.3-15}$$

where $p(\phi)$ is the pdf of ϕ given in Chapter 3.

Equation (5.3-13) can be computed if the pdf of the random variable, $Y(\phi)$, is known. However, computation of the pdf of $Y(\phi)$ appears to be formidable, unless some simplifying assumptions are made. We consider the following three special cases of practical interest.

5.3.1.1 High-data-rate case. When the data rate is high with respect to the receiver tracking loop bandwidth, the phase process $\phi(t)$ varies slowly and is essentially constant over a bit interval T_b.[2] Hence, the phase process $\phi(t)$ can be approximated as a constant within the bit interval T_b. Thus, the conditional bit error probability as given in (5.3-13) becomes

$$P_b(\phi) = \frac{1}{2} \operatorname{erfc}\left[\sqrt{\frac{E_b}{N_0}} \cos \phi\right] \tag{5.3-16}$$

while the average bit error probability as obtained by using (5.3-16) in (5.3-15) is

$$P_b = \frac{1}{2} \int_{-\pi}^{\pi} \operatorname{erfc}\left[\sqrt{\frac{E_b}{N_0}} \cos \phi\right] p(\phi)\, d\phi \tag{5.3-17}$$

where the pdf $p(\phi)$ is as given in Chapter 3 and is repeated here for convenience. In the absence of loop stress or detuning, $p(\phi)$ may be approximated by the Tikhonov pdf

[2] For a coded system, we should compare the rate of the variation of the phase process $\phi(t)$ with the time it takes to decode one bit in order to decide if the high data rate case is applicable or if the following two cases (namely, low data rate and medium data rate) are more appropriate.

$$p(\phi) = \begin{cases} \dfrac{\exp(\rho \cos \phi)}{2\pi I_0(\rho)}; & |\phi| \leq \pi \\ 0; & \text{otherwise} \end{cases} \tag{5.3-18}$$

where

$$\rho = \frac{P_c}{N_0 B_L}$$

is the tracking loop SNR defined in Chapter 3.

Using (5.3-18), the bit error probability P_b given in (5.3-17) is evaluated as a function of bit energy-to-noise spectral density ratio E_b/N_0 for various values of ρ. The results are plotted in Fig. 5-15, illustrating this noisy reference effect on system performance. System losses including noisy reference loss (or radio loss) will be treated in Section 5.5 of this chapter. Note that the presentation in Fig. 5-15 is not valid for suppressed carrier systems, which will be discussed in Section 5.6 of this chapter.

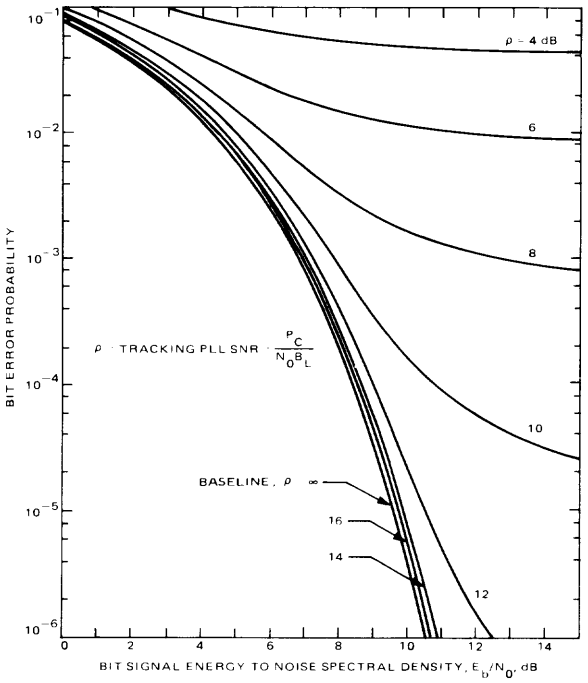


Fig. 5-15. Bit error probability performance of an uncoded PSK system

5.3.1.2 Low-data-rate case. When the data rate is low relative to the receiver tracking loop bandwidth, the phase process $\phi(t)$ varies rapidly over the bit interval T_b. Hence the random variable $Y(\phi)$, as given above in (5.3-14), is a good approximation of the true time average of the function $\cos \phi(t)$. If $\phi(t)$ is an ergodic process, the time average may be replaced by the statistical mean. Thus

$$Y(\phi) \cong \overline{\cos \phi} \tag{5.3-19}$$

and the bit error probability is approximated by

$$P_b \cong \frac{1}{2}\text{erfc}\left[\sqrt{\frac{E_b}{N_0}}\,\overline{\cos \phi}\right] \tag{5.3-20}$$

5.3.1.3 Medium-data-rate case. When neither of the above two extreme cases is valid, we have a medium-data-rate case. Here the data rate is comparable with the receiver tracking loop bandwidth. We must seek an alternate approximation for the pdf of $Y(\phi)$. One technique which is valid in many practical situations is to employ the assumption that the phase error $\phi(t)$ is small enough so that the linear PLL theory can be used. Then $\phi(t)$ is a Gaussian process. This is a good approximation in high loop SNR situations. Hence

$$Y(\phi) \cong 1 - \frac{1}{2T_b}\int_0^{T_b} \phi^2(t)\,dt \tag{5.3-21}$$

and the computation of its pdf reduces to computing the pdf of the integral of the square of a Gaussian process. Then the bit error probability can be obtained by using (5.3-21) in (5.3-13), (5.3-15), and (5.3-18). Another technique is the interpolation scheme presented in subsection 5.5.1.3.

5.3.2 Two-Way Telemetry Performance

The phase coherent two-way telemetry link can be modeled as in Fig. 5-16. The ground or reference system transmits an uplink signal

$$\sqrt{2P_{c1}}\,\sin(\omega_1 t + \theta_0)$$

where P_{c1} is the uplink carrier power, ω_1 is the carrier frequency, and θ_0 is a constant phase that can be assumed zero without loss of generality. The channel introduces an arbitrary and unknown phase shift $\theta_1(t)$ in the transmitted signal.

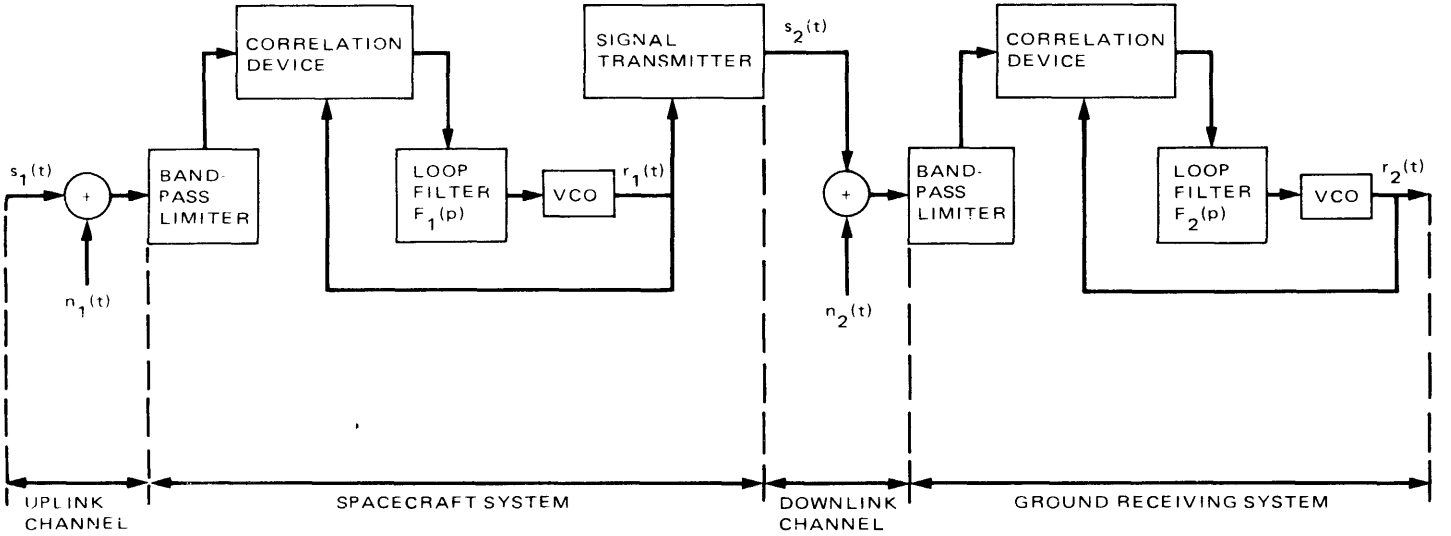


Fig. 5-16. A phase-coherent two-way link

At the spacecraft, the receiver input is

$$\sqrt{2P_{c1}} \sin [\omega_1 t + \theta_1(t) + \theta_0] + n_1(t)$$

where $n_1(t)$ is the additive channel noise, which is assumed to be a stationary white Gaussian noise process with one-sided spectral density N_{01} watts/Hz. The spacecraft receiver phase-locked loop coherently tracks the uplink carrier and produces a local reference signal,

$$\sqrt{2} \cos [\omega_1 t + \hat{\theta}_1(t)]$$

where $\hat{\theta}_1(t)$ is the PLL estimate of $\theta_1(t) + \theta_0$. The phase error of this one-way link is

$$\phi_1(t) = \hat{\theta}_1(t) - \theta_1(t) - \theta_0 \tag{5.3-22}$$

The spacecraft then transmits a downlink signal that is coherently derived from the uplink signal—more precisely, from the uplink signal recovered by the spacecraft receiver phase-locked loop. Because the transponder has a frequency translation ratio G_f, of downlink frequency to uplink frequency, the downlink signal is

$$\sqrt{2P_{c2}} \sin G_f[\omega_1 t + \hat{\theta}_1(t)] = \sqrt{2P_{c2}} \sin [\omega_2 t + G_f \hat{\theta}_1(t)] \tag{5.3-23}$$

where P_{c2} is the downlink carrier power and ω_2 is the carrier frequency. The channel introduces an arbitrary phase shift, $\theta_2(t)$, in the downlink signal.

At the reference system, the receiver input is

$$\sqrt{2P_{c2}} \sin [\omega_2 t + G_f \hat{\theta}_1(t) + \theta_2(t)] + n_2(t)$$

where $n_2(t)$ is the additive channel noise, which is assumed to be a stationary white Gaussian process with one-sided spectral density N_{02} watts/Hz.

The ground receiver phase-locked loop (PLL) coherently tracks the downlink carrier and produces a local reference signal,

$$\sqrt{2} \cos [\omega_2 t + \hat{\theta}_2(t)]$$

where $\hat{\theta}_2(t)$ is the PLL estimate of $G_f \hat{\theta}_1(t) + \theta_2(t)$

There are two important error estimates in this two-way link that are essential to system design. The first is the two-way phase error, defined as

$$\phi_2(t) = \hat{\theta}_2(t) - [G_f\hat{\theta}_1(t) + \theta_2(t)] \tag{5.3-24}$$

i.e., the difference between the ground receiver estimate and the received signal phase. This is important for telemetry analysis. The second is the two-way doppler phase, defined as

$$\phi_{d2}(t) = \hat{\theta}_2(t) - \theta_0$$

i.e., the difference between the ground receiver estimated phase of the turn-around signal and the original phase transmitted by the ground transmitter. This is important for doppler considerations (see Chapter 4). Without loss of generality, one can assume that if $\theta_0 = 0$, then

$$\phi_{d2}(t) = \hat{\theta}_2(t) \tag{5.3-25}$$

The present practical planetary two-way system employs second-order PLLs both in the spacecraft receiver and the ground receiver. The transfer function of the loop filter in the spacecraft PLL is given by

$$F_1(p) = \frac{1 + \tau_{21}p}{1 + \tau_{11}p} \tag{5.3-26}$$

and the transfer function of the loop filter in the ground PLL is given by

$$F_2(p) = \frac{1 + \tau_{22}p}{1 + \tau_{12}p} \tag{5.3-27}$$

Note that if $\tau_{2i} = \tau_{1i}$, $i = 1, 2$, then $F_i(p)$ is a first-order loop.

From the linear theory given in [5-3] and [5-10], we know that the carrier phase variance is

$$\sigma_2^2 = \frac{G_f^2 K_R}{\rho_1} + \frac{1}{\rho_2} \tag{5.3-28}$$

where

$$\rho_1 = \text{carrier loop SNR in spacecraft PLL} = \frac{P_{c1}}{N_{01}B_{L1}}$$

$$\rho_2 = \text{carrier loop SNR in ground PLL} = \frac{P_{c2}}{N_0 B_{L2}}$$

$$K_R = \frac{r_1 \xi}{r_2(r_1+1)} \left[\frac{r_2 + r_1 r_2 \xi (1+\xi) + (1+\xi)\xi^2 + r_1 \xi^3}{\frac{r_2}{r_1} + r_2 \xi + (r_1 + r_2 - 2)\xi^2 + r_1 \xi^3 + \frac{r_1 \xi^4}{r_2}} \right]$$

$$\xi = \frac{B_{L1}(r_2+1)}{B_{L2}(r_1+1)}$$

$$B_{L1} = \text{one-sided loop bandwidth of spacecraft PLL} = \frac{1+r_1}{4\tau_{21}}$$

$$B_{L2} = \text{one-sided loop bandwidth of ground PLL} = \frac{1+r_2}{4\tau_{22}}$$

$$r_i = \frac{\sqrt{P_{ci}} K_i \tau_{2i}^2}{\tau_{1i}}, \quad i = 1, 2$$

$$K_i = \text{loop gain in spacecraft } (i = 1) \text{ or ground } (i = 2) \text{ PLL}$$

Of course, if bandpass limiters precede the spacecraft and/or ground PLLs, then the above set of equations defining σ_2^2 needs to be adjusted to account for the limiter suppression factor and the limiter signal-to-noise spectral density ratio degradation, as discussed in Chapter 3.

The two-way phase error mode is completed by specifying its probability density function, $p(\phi_2)$.

$$p(\phi_2) = \begin{cases} \dfrac{\exp\left(\dfrac{\cos\phi_2}{\sigma_2^2}\right)}{2\pi I_0\left(\dfrac{1}{\sigma_2^2}\right)}; & -\pi \leqslant \phi \leqslant \pi \\ 0; & \text{otherwise} \end{cases} \tag{5.3-29}$$

If $f(\cdot)$ represents the baseline probability of error versus E_b/N_0 curve, then the bit error probability, P_b, for a two-way coherent-system high-data-rate case is given by

$$P_b = \int_{-\pi}^{\pi} f\left(\frac{E_b}{N_0}\cos^2\phi_2\right) p(\phi_2)\, d\phi_2 \qquad (5.3\text{-}30)$$

where $p(\phi_2)$ is given in (5.3-29).

5.4 Channel Coding and Coded Telemetry System Performance

The goal of the channel encoder and decoder is to map the input digital sequence into a channel input sequence, and, conversely, the channel output sequence into an output digital sequence, such that the effect of the channel noise is minimized—that is, such that the number of discrepancies (errors) between the output and input digital sequences is minimized. The approach is to introduce redundancy with the channel encoder and to use this redundancy at the decoder to reconstitute the input sequence as accurately as possible. Therefore, if an efficient code is selected and a good decoding scheme is used, we may require for a given bit error rate and a given data rate less ST_b/N_0 than if the data is transmitted uncoded.

For a treatment of channel coding that is more nearly complete than that given in this section, the reader should consult [5-11] through [5-19].

5.4.1 Types of Codes

The encoder of Fig. 5-17 accepts at its input a continuous sequence of information digits. At its output, it produces another sequence with somewhat more digits, which is fed to the modulator. The decoder accepts a sequence of channel symbols from the demodulator and translates it into a somewhat shorter sequence of information digits. The rules under which the encoder and decoder operate are specified by the particular code that is employed.

There are two fundamentally different types of codes. The encoder for a *block code* breaks the continuous sequence of information digits into K-symbol sections or blocks. It then operates on these blocks independently according to the particular code to be employed. With each possible information block is associated an N-tuple of channel symbols, where $N > K$. The result, now called a *code word*, is transmitted, corrupted by noise, and decoded independently of all other code words. The quantity N is referred to as the code length or *block length*.

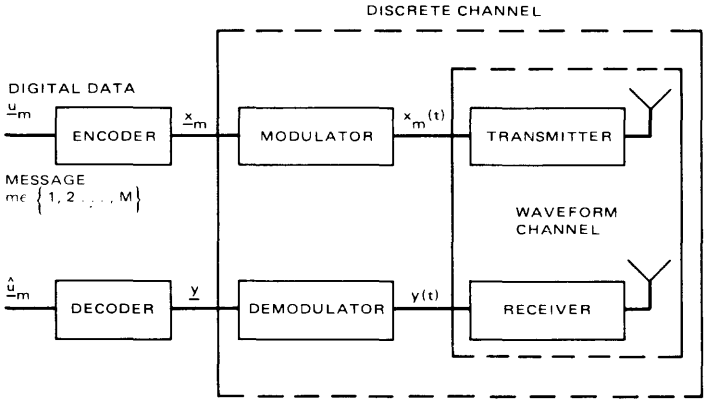


Fig. 5-17. Block-coded digital communication system

The other type of code, called a *tree code*, operates on the information sequence without breaking it up into independent blocks. Rather, the encoder for a tree code processes the information continuously and associates each long (perhaps semi-infinite) information sequence with a code sequence containing somewhat more digits. The encoder breaks its input sequence into b-symbol blocks, where b is usually a small number. Then, on the basis of this b-tuple and the preceding information symbols, it emits an n-symbol section of the code sequence. The name "tree code" stems from the fact that the encoding rules for this type of code are most conveniently described by means of a tree graph.

The class of *convolutional codes* forms a subset of the class of tree codes. These codes are important since they are simpler to implement than other types of tree codes. In this book the convolutional codes are the only tree codes considered.

Block codes and convolutional codes have similar error-correcting capabilities and the same fundamental limitations. In particular, Shannon's fundamental theorem for the discrete noisy channel holds for both types of codes. This result states that a channel has a well-defined capacity and that by using suitable codes it is possible to transmit information at any rate less than channel capacity with an arbitrarily small probability of decoding erroneously.

Since the appearance of Shannon's classic papers[3] in 1948 and 1949, a great deal of research has been devoted to the problem of designing efficient schemes by which information can be coded for reliable transmission across channels that are corrupted by noise.

[3] For a reprint of these papers, see [5-20].

From a practical standpoint, the essential limitation of all coding and decoding schemes proposed to date has not been Shannon's *capacity* but the complexity (and *cost*) of the decoder. For this reason, efforts have been directed toward the design of coding and decoding schemes that could be easily implemented.

For transmission from a spacecraft, for example, where the amount of equipment at the transmitter is much more important than the total amount of equipment for the communication system, the encoding equipment should be as simple as possible, lightweight, and more reliable with respect to equipment failure. The complete decoding procedures can be performed on the earth, where equipment limitations are not severe.

5.4.2 Linear Block Codes

Block encoding can also be regarded as a table-look-up operation, where each of the $M = q^K$ (q is the alphabet size of information symbols) code signal vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M$ is stored in an N-stage register of a memory bank and, whenever message $\mathbf{u}_m$ is to be transmitted, the corresponding signal vector $\mathbf{x}_m$ is read into the modulator. Alternatively, we may label each of the $M = q^K$ messages, as in Fig. 5-17, by a K-vector over a q-ary alphabet. Then the encoding becomes a one-to-one mapping from the set of message vectors $\{\mathbf{u}_m = (u_{m1}, \cdots, u_{mK})\}$ into the set of code signal vectors $\{\mathbf{x}_m = (x_{m1}, \cdots, x_{mN})\}$. We shall concern ourselves primarily with binary alphabets, i.e., $q = 2$; thus, initially, we take $u_{mn} \in \{0, 1\}$ for all m, n.

For binary-input data, a linear code consists simply of a set of modulo-2 linear combinations of the data symbols, which may be implemented as shown in Fig. 5-18. The K-stage register stores the K incoming information bits that correspond to data block $\mathbf{u}_m = (u_{m1}, \cdots, u_{mK})$.

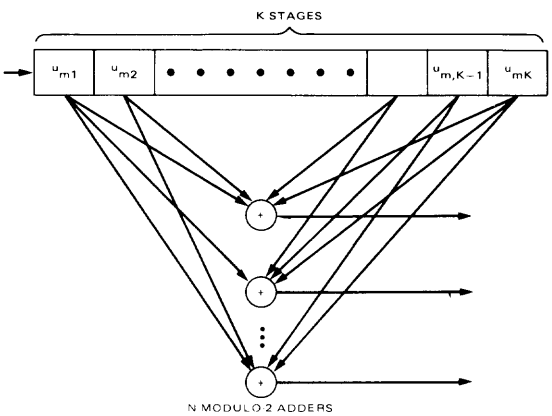


Fig. 5-18. Linear block encoder

The encoder then consists of N modulo-2 adders, each of which adds together a subset of the data symbols $u_{m1}, u_{m2}, \cdots, u_{mK}$ to generate one *code symbol* x_{mn} where $n = 1, 2, \cdots, N$ as shown in Fig. 5-18. We shall refer to the vector $\mathbf{x}_m = (x_{m1}, x_{m2}, \cdots, x_{mN})$ as the *code vector* or *codeword* with *block length* N. There are 2^K codewords. K is called the *dimension* of the code. We call the code an (N, K) *block code*. This code uses N symbols to send K message symbols, so it is said to have *code rate* $r = K/N$.

Thus the linear coding operation for binary data can be represented by

$$\begin{aligned} x_{m1} &= u_{m1}g_{11} \oplus u_{m2}g_{21} \oplus \cdots \oplus u_{mK}g_{K1} \\ x_{m2} &= u_{m1}g_{12} \oplus u_{m2}g_{22} \oplus \cdots \oplus u_{mK}g_{K2} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_{mN} &= u_{m1}g_{1N} \oplus u_{m2}g_{2N} \oplus \cdots \oplus u_{mK}g_{KN} \end{aligned} \tag{5.4-1}$$

where $\oplus$ denotes modulo-2 addition, and $g_{kn} \in \{0, 1\}$ for all k, n. The term $u_{mk}g_{kn}$ is an ordinary multiplication, so that u_{mk} enters into the particular combination for x_{mn} if and only if $g_{kn} = 1$. The matrix

$$G = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1N} \\ g_{21} & g_{22} & \cdots & g_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ g_{K1} & g_{K2} & \cdots & g_{KN} \end{bmatrix} = \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_K \end{bmatrix} \tag{5.4-2}$$

is called the *generator matrix* of the linear code and $\{\mathbf{g}_i\}$ are its row vectors. Thus (5.4-1) can be expressed in vector form as

$$\mathbf{x}_m = \mathbf{u}_m G = \sum_{k=1}^{K} u_{mk}\mathbf{g}_k \tag{5.4-3}$$

where $\sum$ means modulo-2 addition and both $\mathbf{u}_m$ and $\mathbf{x}_m$ are binary row vectors.

An important property of linear codes is known as *closure*: namely, the property that the modulo-2 termwise sum of two code vectors $\mathbf{x}_m$ and $\mathbf{x}_k$

$$\mathbf{x}_m \oplus \mathbf{x}_k = (x_{m1} \oplus x_{k1}, x_{m2} \oplus x_{k2}, \cdots, x_{mN} \oplus x_{kN}) \tag{5.4-4}$$

is also a code vector. An interesting consequence of the closure property is that the set of Hamming distances from a given code vector to the $(M-1)$ other code vectors is the same for all code vectors. Define the *Hamming weight* of a binary vector as the number of ones in the vector. The *Hamming distance* between two vectors $\mathbf{x}_m$ and $\mathbf{x}_k$ is then just the Hamming weight of their modulo-2 termwise sum, denoted $w(\mathbf{x}_m \oplus \mathbf{x}_k)$. The *minimum distance* of a code is the minimum (Hamming) distance between its codewords. Minimum distance is an important parameter of a linear code.

5.4.2.1 Systematic linear codes. Consider a linear code whose generator matrix has the form of (5.4-5).

$$G = \begin{bmatrix} 1\ 0\ 0\ \cdots\ 0\ g_{1,K+1}\ \cdots g_{1N} \\ 0\ 1\ 0\ \cdots\ 0\ g_{2,K+1}\ \cdots g_{2N} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 0\ 0\ 0\ \cdots\ 1\ g_{K,K+1}\ \cdots g_{KN} \end{bmatrix} \tag{5.4-5}$$

We note that a linear code (5.4-3) generated by a matrix (5.4-5) has its first K code symbols identical to the data symbols. A codeword with this property is

$$\mathbf{x} = \underbrace{x_1 \cdots\cdots x_K}_{\text{message symbols}}\ \underbrace{x_{K+1} \cdots\cdots x_n}_{\text{check symbols}}$$

Such a code, which transmits the K data symbols unchanged together with $N-K$ "parity check" symbols is called a *systematic code*. It can be shown that any one-to-one linear code [5-11, 5-12] is equivalent in performance to a systematic code.

5.4.2.2 Decoding linear block codes. Suppose the message $\mathbf{u} = u_1 \cdots u_K$ is encoded into the codeword $\mathbf{x} = x_1 \cdots x_N$, which is then sent through the channel. Because of channel noise, the received vector $\mathbf{y} = y_1 \cdots y_N$, $y_n \in \{0, 1\}$ may be different from $\mathbf{x}$. Let's define the error vector

$$\mathbf{e} = \mathbf{y} \oplus \mathbf{x} = (e_1 \cdots e_n) \tag{5.4-6}$$

The decoder (Fig. 5-17) must decide from $\mathbf{y}$ which message $\mathbf{u}$ or which codeword $\mathbf{x}$ was transmitted. Since $\mathbf{x} = \mathbf{y} \oplus \mathbf{e}$, it's enough if the decoder finds $\mathbf{e}$. The decoder can never be certain what $\mathbf{e}$ was. Its strategy therefore will be to choose the most likely error vector $\mathbf{e}$, given that $\mathbf{y}$ was received. Provided the codewords are all equally likely, this strategy is optimum in the sense that it minimizes the probability of the decoder making a mistake, and is called

maximum likelihood decoding. For many channels it can be shown that an equivalent decoding scheme is simply to compare y with all 2^K codewords and pick the closest in (Hamming) distance. It can be shown that a code with minimum distance d can correct $(d-1)/2$ errors. If d is even, the code can simultaneously correct $(d-1)/2$ errors and detect $d/2$ errors. If more than $d/2$ errors occur, the received vector may or may not be closer to some other codeword than to the correct one. If it is close to some other codeword, the decoder will be fooled and will output the wrong codeword. This is called a *decoding error*. When the receiver makes no attempt to correct errors, but just tests the received vector to see if it is a codeword, this is called *error detection*.

It can be shown [5-14] that the maximum likelihood decoding of any binary code transmitted over the Binary Symmetric Channel (BSC) is equivalent to minimum distance decoding. That is

$$\hat{\mathbf{u}}_m = \mathbf{u}_m \text{ if } w(\mathbf{x}_m \oplus \mathbf{y}) \leqslant w(\mathbf{x}_{m'} \oplus \mathbf{y}) \text{ for all } m' \neq m \tag{5.4-7}$$

We now demonstrate a simpler table look-up technique for decoding systematic linear codes on the BSC. Define the $N \times (N-K)$ matrix H^T by

$$H^{\mathrm{T}} = \begin{bmatrix} g_{1,K+1} & \cdots & g_{1,N} \\ g_{2,K+1} & \cdots & g_{2,N} \\ \cdots & \cdots & \cdots \\ g_{K,K+1} & \cdots & g_{K,N} \\ 1 & 0 \cdots \cdots & 0 \\ 0 & 1 \cdots \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 \cdots \cdots & 1 \end{bmatrix} \tag{5.4-8}$$

The matrix H is called the *parity-check matrix*. Any code vector multiplied by H^T yields the null $\mathbf{0}$ vector

$$\mathbf{x}_m H^T = \mathbf{0} \tag{5.4-9}$$

Now consider post-multiplying any received vector y by H^T. The resulting $(N-K)$-dimensional binary vector is called the *syndrome* of the received vector and is given by

$$\mathbf{s} = \mathbf{y} H^T \tag{5.4-10}$$

The optimum decoding procedure is outlined below.

Step 1 – Initially prior to decoding, for each of the 2^{N-K} possible syndromes s, store the corresponding minimum weight vector e satisfying (5.4-11), below. A table of 2^{N-K} N-bit entries will result.

$$\mathbf{e}H^T = \mathbf{s} \tag{5.4-11}$$

Step 2 – From the N-dimensional received vector y, generate the $(N-K)$-dimensional syndrome s by the linear operation (5.4-10); this requires an N-stage register and $N-K$ modulo-2 adders.

Step 3 – Do a table look-up in the table of Step 0 to obtain $\hat{\mathbf{e}} = \mathbf{e}$ corresponding to the syndrome s found from Step 2.

Step 4 – Obtain the most likely code vector by the operation $\mathbf{x}_m = \mathbf{y} \oplus \hat{\mathbf{e}}$. The first symbols are the data symbols.

Denote the minimum Hamming distance among all pairs of codewords by d_{min}. For the maximum likelihood decoder on the BSC outlined above, any number of errors in a given block that is less than $d_{min}/2$ will be corrected.

5.4.2.3 Examples of linear block codes. Many linear block codes have been devised. In what follows we consider two particular block codes that have recently been employed for the deep-space channel—Golay and Reed-Solomon codes.

The reader interested in block codes and the algebraic theory behind block codes may wish to consult [5-21] through [5-29] as a supplement to the general coding references cited earlier.

5.4.2.3.1 *Golay codes.* Two linear block codes, the Golay (23, 12) and (24, 12) codes, have minimum distances equal to 7 and 8 respectively. The Golay (23, 12) is called a *perfect code* and Golay (24, 12) a quasi-perfect code. The Golay (24, 12) code is used more often than (23, 12) for various reasons including its slightly better performance on the AWGN channel. The Golay (24, 12) code is capable of correcting all patterns of three or fewer bit errors. It is known that the parity-check matrix of the (24, 12) Golay code may be written as

$$H = [I|A] \tag{5.4-12}$$

where I is the 12 X 12 identity matrix and

$$A = \left[\begin{array}{cccccccccccc} 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{array}\right] \qquad (5.4\text{-}13)$$

The Golay codes are among the few linear codes for which all the code vector weights are known. These are summarized in Table 5-2.

On the BSC, for the (24, 12) code, minimum distance decoding always corrects 3 or fewer errors and corrects one-sixth of the weight 4 error vectors. On the other hand, error vectors of weight 5 or more are never corrected. Similarly for the (23, 12) code, all error vectors of weight 3 or less, and only these, are corrected.

Table 5-2. Weight of code vectors in Golay codes

Weight, w	Number of code vectors of weight w, N_w	
	(23, 12) code	(24, 12) code
0	1	1
7	253	0
8	506	759
11	1288	0
12	1288	2576
15	506	0
16	253	759
23	1	0
24	0	1
Total	4096	4096

The performance of the Golay (24, 12) code on the AWGN channel is shown in Fig. 5-19.

5.4.2.3.2 *Reed-Solomon codes.* The Reed-Solomon (RS) codes are a particularly interesting class of linear block codes. The block length N of a Reed-Solomon code is $q - 1$, with q being the alphabet size of the symbols. It can be seen that these codes are useful only for large alphabet sizes. RS codes with K information symbols and block length N have a minimum distance, $d_{min} = N - K + 1$. This code has been used effectively in a concatenation code scheme (see Section (5.4.4)), where the symbols in the Reed-Solomon code are codewords in a smaller, embedded code. The error probability is an exponentially decreasing function of the block length, and the decoding complexity is proportional to a small power of the block length. Reed-Solomon codes can be used directly on a channel with a small input alphabet by representing each letter in a codeword by a sequence of channel letters. Such a technique is useful on channels where the errors are clustered, since the decoder operation depends only on the number of sequences of channel outputs that contain errors.

Using symbols with $q = 2^J$ for some J, the block length is $N = 2^J - 1$. For an arbitrarily chosen odd minimum distance d, the number of information

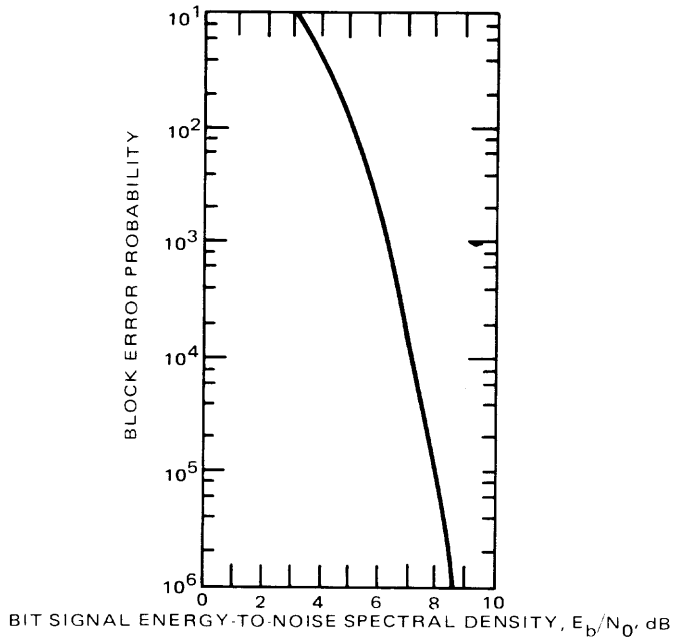


Fig. 5-19. Block error probability for Golay (24, 12) coded signals on the AWGN channel

symbols is $K = N - d + 1$ and any combination of $(d - 1)/2 = (N - K)/2$ errors can be corrected. If we represent each letter in a codeword by J binary digits, then we can obtain a binary code with KJ information bits and block length NJ bits. Any noise sequence that alters at most $(N - K)/2$ of these J length sequences can be corrected, and thus the code has a burst correcting capability of $J[(N - K)/2 - 1] + 1$, along with the capability of correcting many combinations of multiple shorter bursts. Therefore Reed-Solomon codes are very appropriate on burst noisy channels such as a channel consisting of a convolutional encoder-AWGN channel-Viterbi decoder. Reed-Solomon codes will be discussed further in Section 5.4.4.

5.4.3 Convolutional Codes

A constraint length K, code rate $r = b/n$ (bits/symbol) convolutional encoder is a linear finite-state machine consisting of a bK-stage shift register and n linear algebraic function generators (Fig. 5-20). The input data, which is usually binary, is shifted at a bit rate of $1/T_b$ along the register b bits at a time. The output symbol rate is equal to $1/T_s = (rT_b)^{-1}$. In comparison to block codes, convolutional codes encode the input data bits continuously rather than in blocks. The convolutional encoder implementation is simpler and thus less costly and more reliable for spacecraft operation than a block encoder. A (b, n) linear block code can be regarded as a special case of a convolutional code with $K = 1$.

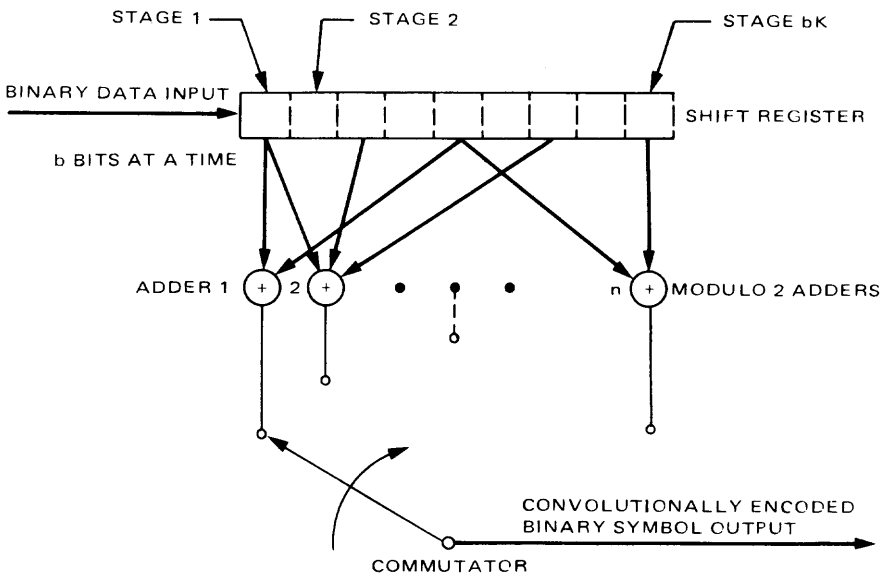


Fig. 5-20. Constraint length *K*, code rate *b/n* convolutional encoder

The particular code structure depends on the manner in which the adders are connected to the shift register. These connections are denoted by a set of vectors

$$\mathbf{g}_i = (g_{i1}, g_{i2}, \cdots, g_{ibK}), \quad i = 1, 2, \cdots, n \tag{5.4-14}$$

where $g_{ij} = 1$ denotes a connection between the jth stage of the shift register and the ith adder, and $g_{ij} = 0$ denotes the absence of a connection. The complete set of $\mathbf{g}_i$ and the parameter b define the code. Figure 5-21 demonstrates the meaning of $\mathbf{g}_i$.

If the data bits are sent as part of the output, the code is termed "systematic." Otherwise, it is "nonsystematic." Thus, a systematic code is one which for some $1 \leqslant i \leqslant n$ and $0 \leqslant m \leqslant K - b$ (usually $i = 1$ and $m = 0$)

$$g_{i,m+j} = \begin{cases} 1; j = 1, 2, \cdots, b \\ 0; \text{otherwise} \end{cases} \tag{5.4-15}$$

Systematic codes have the advantage of the simple quick-look property. That is, since the data is present explicitly in the symbol stream, it can be examined without decoding.

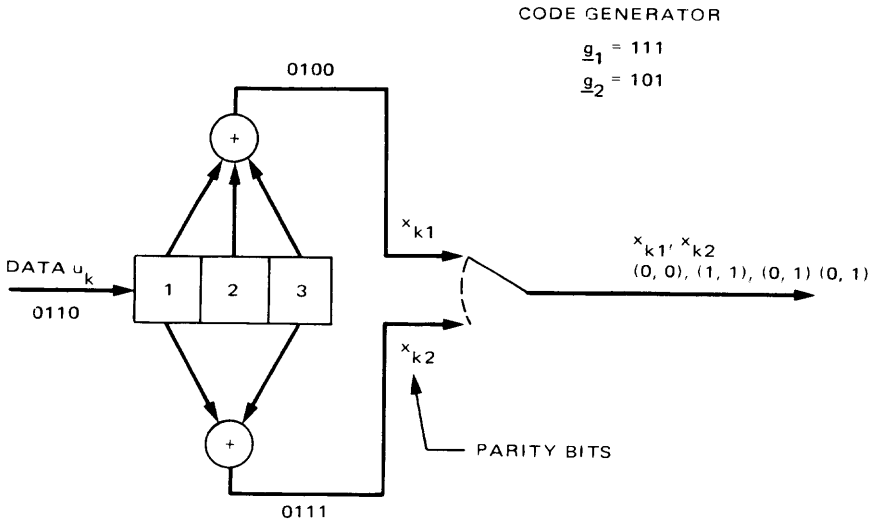


Fig. 5-21. Nonsystematic convolutional encoder with $K = 3, r = 1/2$; the code generator denotes the top positions

Some of the code structures are transparent to code inversion—that is, if the signs of the input bits are reversed, the coded output bit sequence is simply inverted. Thus, if the numbers of "ones" or weights of all g_i's are odd, then the code is transparent to a sign inversion. That is, the decoded output bit stream has the same sign ambiguity as the input. This transparency is valuable if biphase-modulated PSK is used with its ensuing sign ambiguity, for it permits decoding prior to ambiguity removal.

In deep-space telecommunication, data-derived symbol synchronization, rather than a separate synchronization channel, is used. Inverting alternate symbols of a convolutional encoder output provides sufficient density of symbol transitions to guarantee adequate symbol synchronizer performance, a guarantee otherwise lacking. Although alternate symbol inversion may increase or decrease the average transition density, depending on the data source model, it does limit the number of contiguous symbols without transition for a particular class of convolutional codes, independent of the data source model. Further, this limit is sufficiently small to guarantee acceptable symbol synchronizer performance for typical applications. Subsequent inversion of alternate detected symbols permits proper decoding. In Fig. 5-22, two NASA Planetary Standard constraint length 7, convolutional codes with code rates 1/2 and 1/3 and symbol inverters are shown. The maximum number of contiguous symbols without transition for the 1/2 and 1/3 code-rates of the convolutional codes of Fig. 5-23 are 14 and 12, respectively.

Before turning to the subject of decoding convolutional codes, we call [5-2], [5-3], and [5-30] through [5-32] to the reader's attention. These references, along with the general coding references cited earlier, are a representative cross section of the literature on the subject of convolutional codes.

5.4.3.1 Viterbi decoding of convolutional codes. In this subsection we consider the performance of Viterbi decoders, including the effect of noisy carrier reference. For discussions on the design of Viterbi decoders, the reader should consult Appendix A and [5-33] and [5-34].

The Viterbi Algorithm decoder of the Standard NASA $K = 7$, $r = 1/2$ convolutional codes with 8-level (3-bit) quantization is called the Maximum Likelihood Convolutional Decoder (MCD). For the MCD the probability of bit error baseline (that is, with perfect carrier reference) has been determined by simulation. If $f(\cdot)$ denotes the probability of bit error, then, according to [5-52] and [5-53]

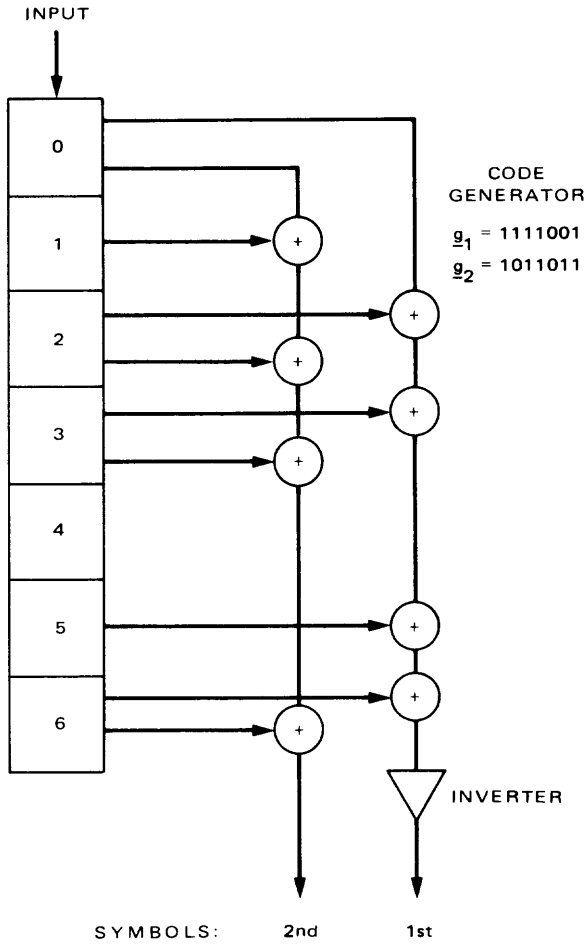


Fig. 5-22. Constraint length $K = 7$, code rate 1/2, NASA standard convolutional code

$$f(x^2) = \begin{cases} \exp(\alpha_0 - \alpha_1 x^2); & x \geqslant \sqrt{\dfrac{\ln 2 + \alpha_0}{\alpha_1}} \\ \dfrac{1}{2}; & |x| = \sqrt{\dfrac{\ln 2 + \alpha_0}{\alpha_1}} \\ 1 - \exp(\alpha_0 - \alpha_1 x^2); & x \leqslant -\sqrt{\dfrac{\ln 2 + \alpha_0}{\alpha_1}} \end{cases} \tag{5.4-16}$$

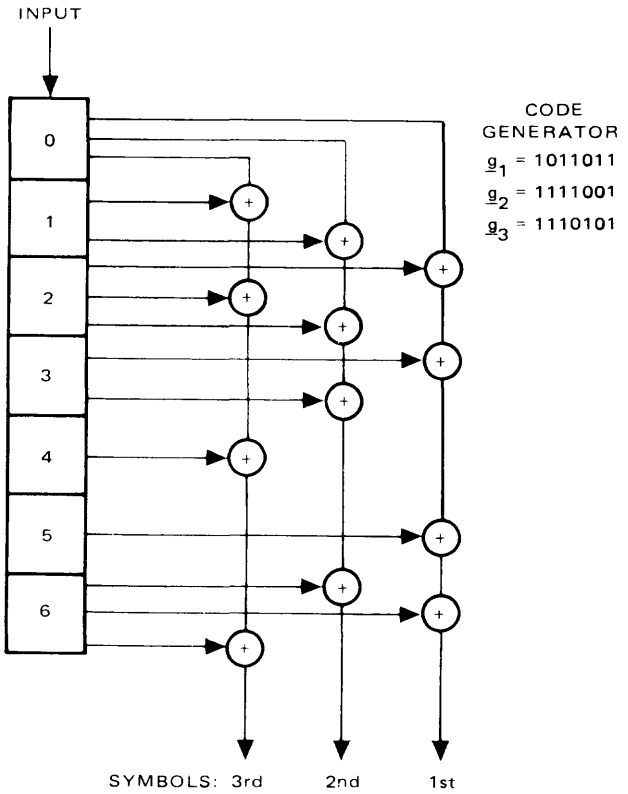


Fig. 5-23. Constraint length K = 7, code rate 1/3, NASA standard convolutional code

where α_0 = 4.4514 and α_1 = 5.7230. For perfect carrier reference, i.e., the baseline case, we have $x^2 = E_b/N_0$. This model has been compared with two sets of experimental data and shown to yield excellent agreement, as illustrated in Fig. 5-24.

A modification to the use of this model for computing bit error probability performance over channels with carrier tracking was suggested in [5-53]. Basically this modification accounts for the effect of symbol stream inversion when the absolute value of the carrier tracking phase-locked loop phase error is between $\pi/2$ and π. Since in a biphase suppressed carrier system with perfectly resolved ambiguity this cannot occur [see (5.6-23)], then the above modification is unnecessary.

The effect of carrier phase jitter on the Viterbi decoding of convolutionally encoded data can be found [5-34] as

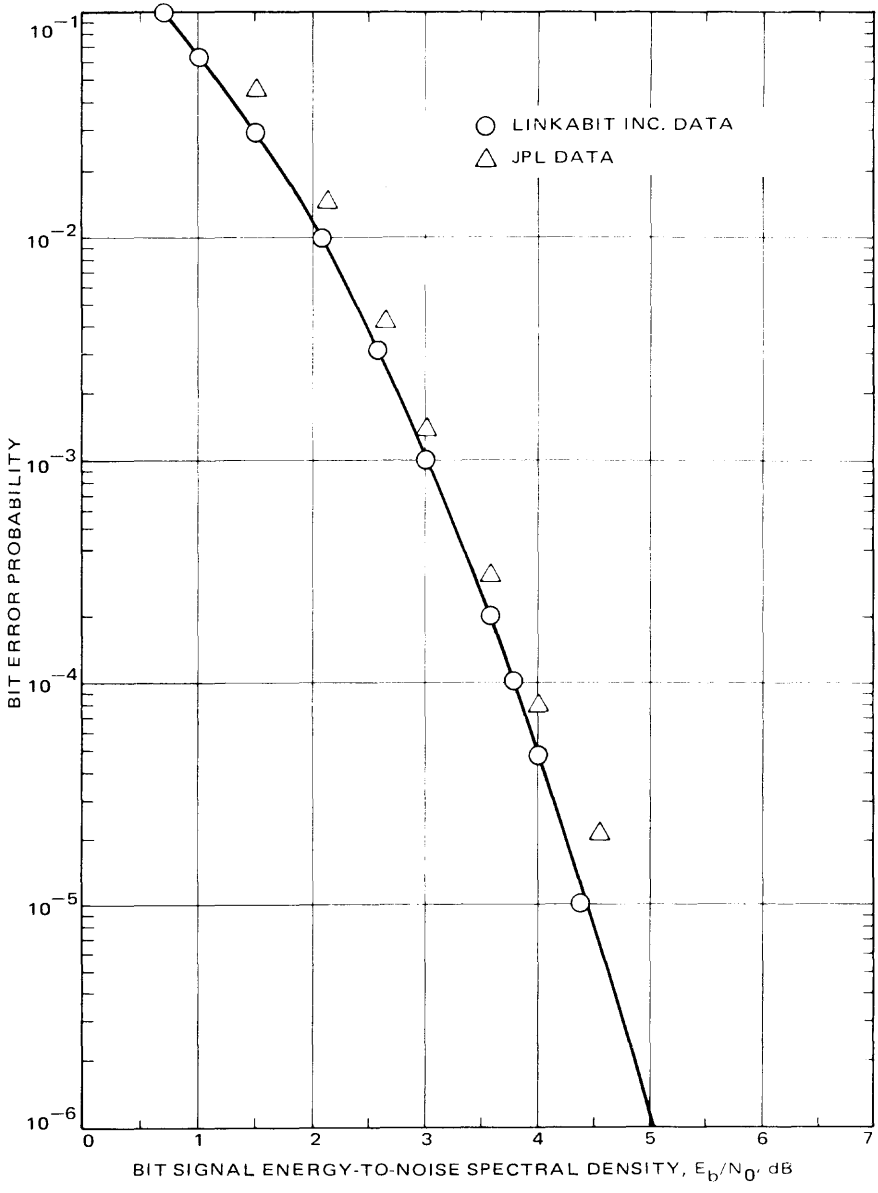


Fig. 5-24. Baseline performance Viterbi decoder for $K = 7, r = 1/2$, convolutional code; soft quantized, $Q = 3$

$$P_b = 2\int_0^{\pi/2} f\left(\frac{E_b}{N_0}\cos^2\phi\right) p(\phi)\, d\phi + 2\int_{\pi/2}^{\pi}\left[1 - f\left(\frac{E_b}{N_0}\cos^2\phi\right)\right] p(\phi)\, d\phi \tag{5.4-17}$$

where $p(\phi)$ is a one-way or two-way carrier phase error density function. For one-way tracking we can use a model for the pdf of phase error in (5.3-18), with loop SNR

$$\rho = \frac{P_c}{N_0 B_L \Gamma} \tag{5.4-18}$$

for PLL preceded by a bandpass limiter (see Chapter 3). In (5.4-18), P_c is the carrier power, B_L is the noise loop bandwidth and Γ is the performance factor of a bandpass limiter (see Chapter 3). The loop SNR can easily be related to carrier margin. This relation depends on loop noise bandwidth at threshold B_{L0}. The relation between ρ and the carrier margin is shown in Fig. 5-25. The effect of a noisy reference on P_b is shown in Fig. 5-26.

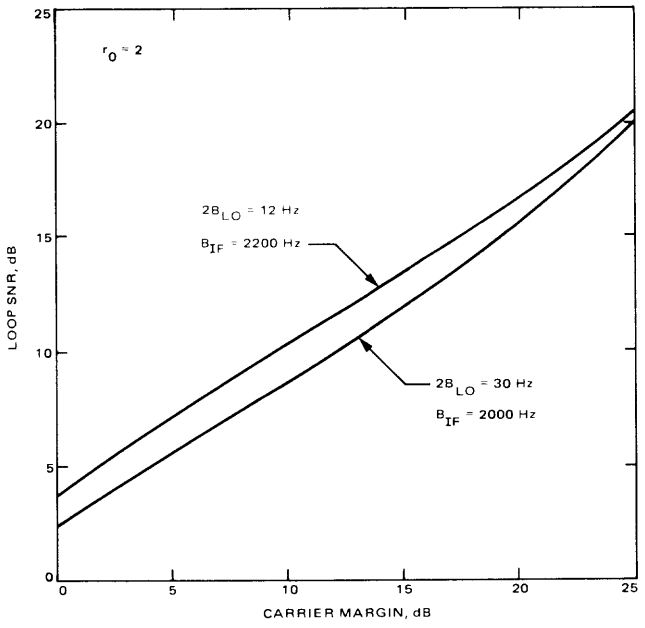


Fig. 5-25. Relation between loop SNR and carrier margin

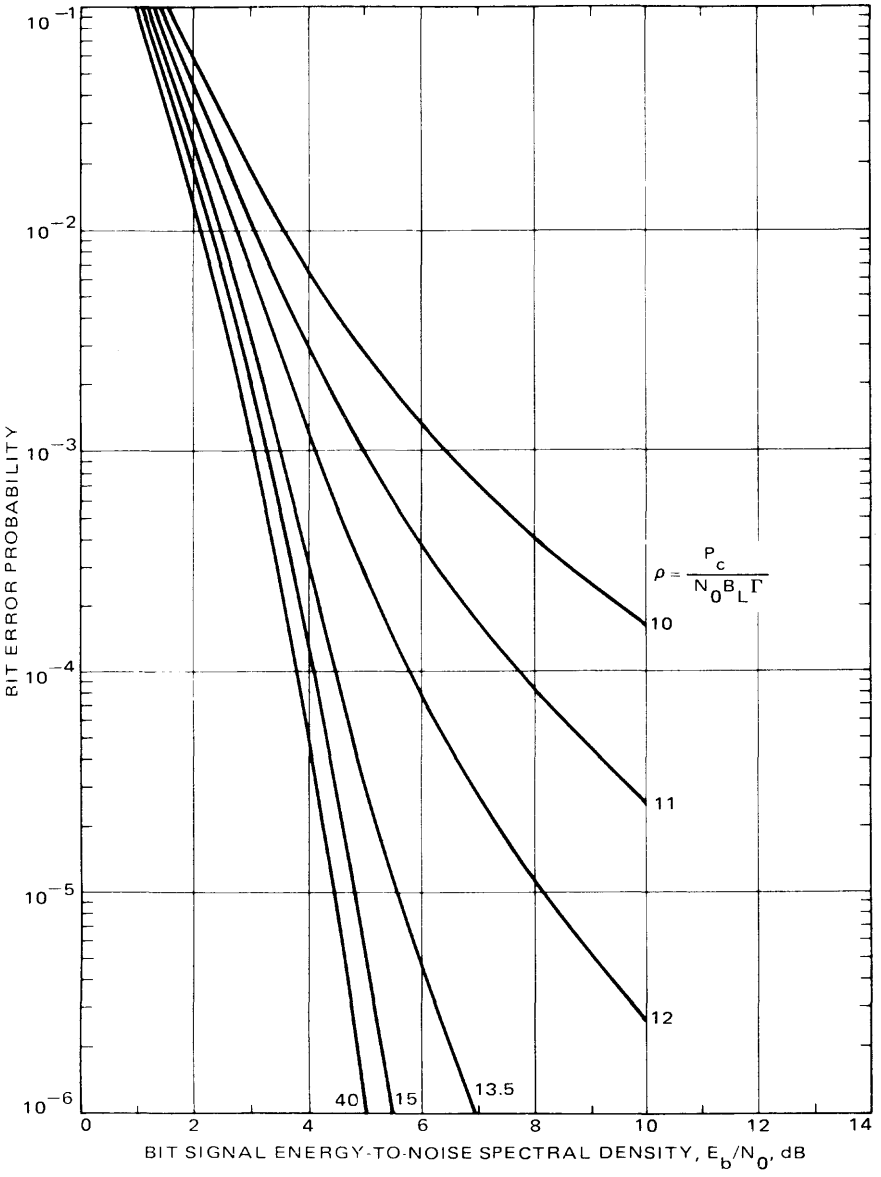


Fig. 5-26. Bit error probability performance of Viterbi decoder (high rate model) as a function of receiver input SNR for varying carrier loop SNR (dB)

Radio loss for a given bit error rate P_b is defined as the excess of E_b/N_0 required under noisy reference to achieve the same P_b as under perfect reference; i.e., given P_b, let $(E_b/N_0)_1$ satisfy

$$P_b = f\left[\left(\frac{E_b}{N_0}\right)_1\right] \tag{5.4-19}$$

and $(E_b/N_0)_2$ satisfy

$$P_b = \overline{f\left[\left(\frac{E_b}{N_0}\right)_2 \cos^2 \phi\right]} \tag{5.4-20}$$

where the average is over ϕ and is given by (5.4-17). Then

$$\text{radio loss in dB} = \left(\frac{E_b}{N_0}\right)_2 - \left(\frac{E_b}{N_0}\right)_1 \tag{5.4-21}$$

For a given P_b, the radio loss curves are shown in Fig. 5-27. Radio loss will be further discussed in Section 5.5.1.

For two-way tracking we can use a model for phase error in (5.3-29), with

$$\rho_i = \frac{P_{ci}}{N_{0_i} B_{L_i} \Gamma_i}; \quad i = 1, 2 \tag{5.4-22}$$

The index $i = 1$ indicates the uplink and $i = 2$ indicates the downlink. Using $p(\phi)$ given by (5.3-29) and (5.4-22) in (5.4-17) we can get the effect of two-way carrier tracking on P_b. The performance of the system when two-way carrier tracking is used for various link conditions is given in Figs. 5-28 through 5-37.

5.4.3.2 Sequential decoding of convolutional codes. In this subsection we consider the performance of sequential decoders for convolutional codes, including the effect of a noisy carrier reference. For more details on the performance of sequential decoders, the reader should consult [5-35] through [5-38]. For discussions on the design of sequential decoders, the reader should consult Appendix B, [5-2], and [5-39].

From a performance point of view there are two types of decoding failure. The first, called an *undetected error*, occurs when the decoder accepts a

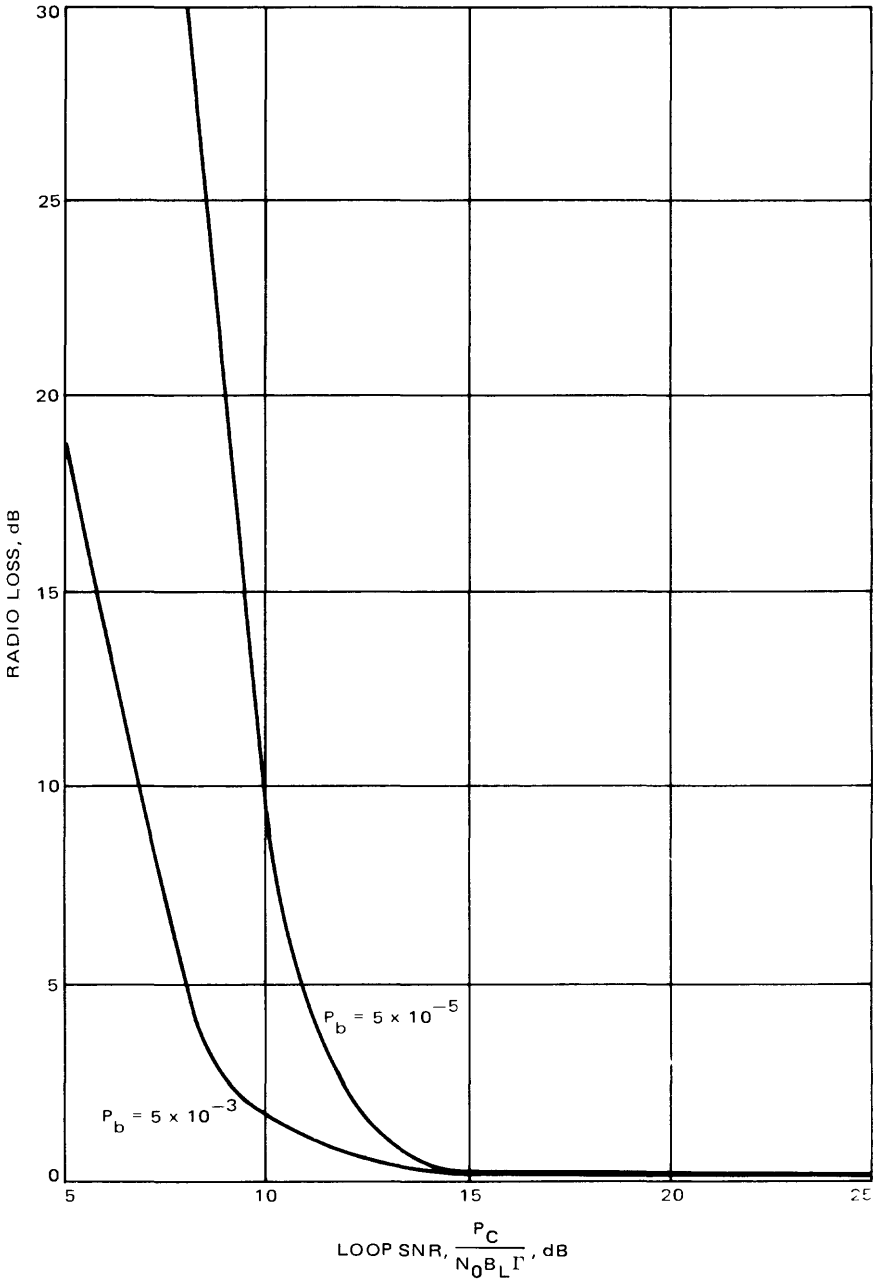


Fig. 5-27. One-way radio loss vs carrier loop SNR for Viterbi decoder

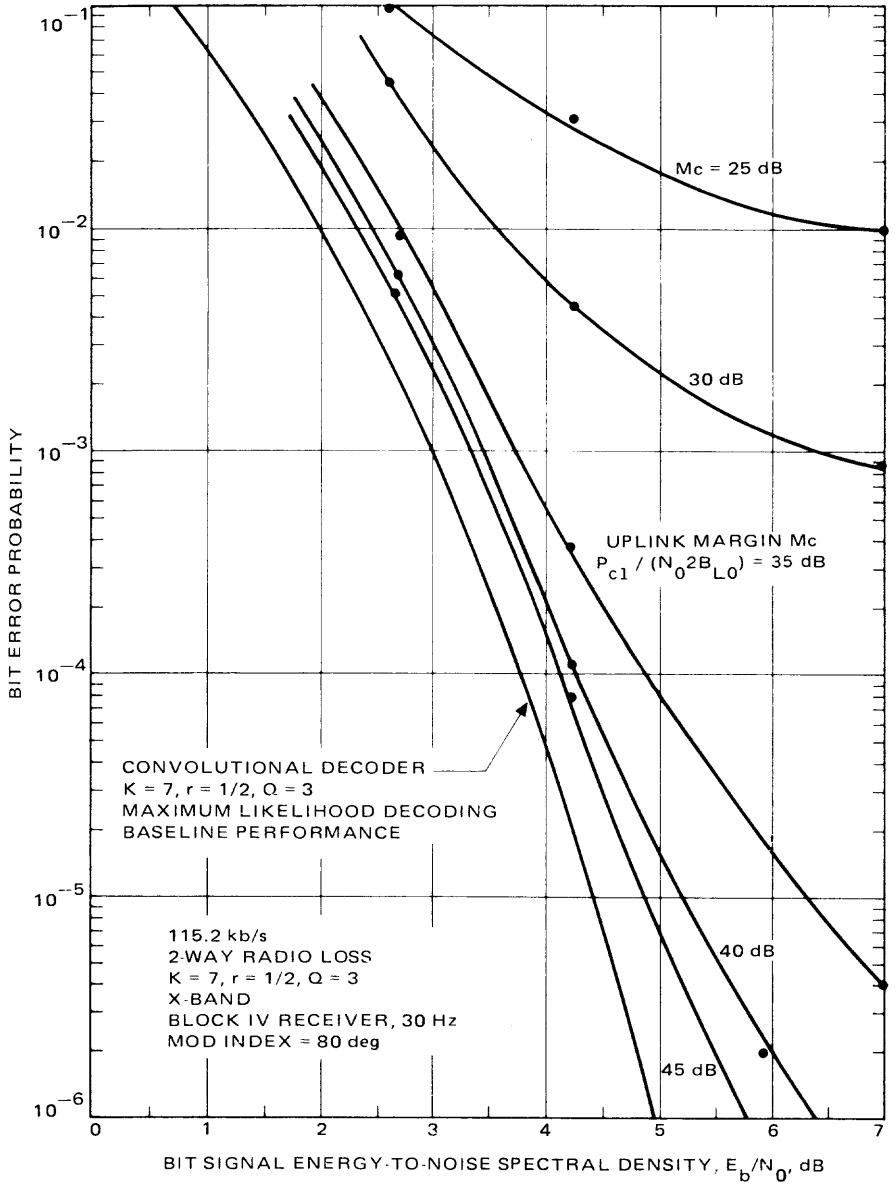


Fig. 5-28. Two-way link bit error probability performance of Viterbi decoder for $K = 7, r = 1/2$, convolutional code

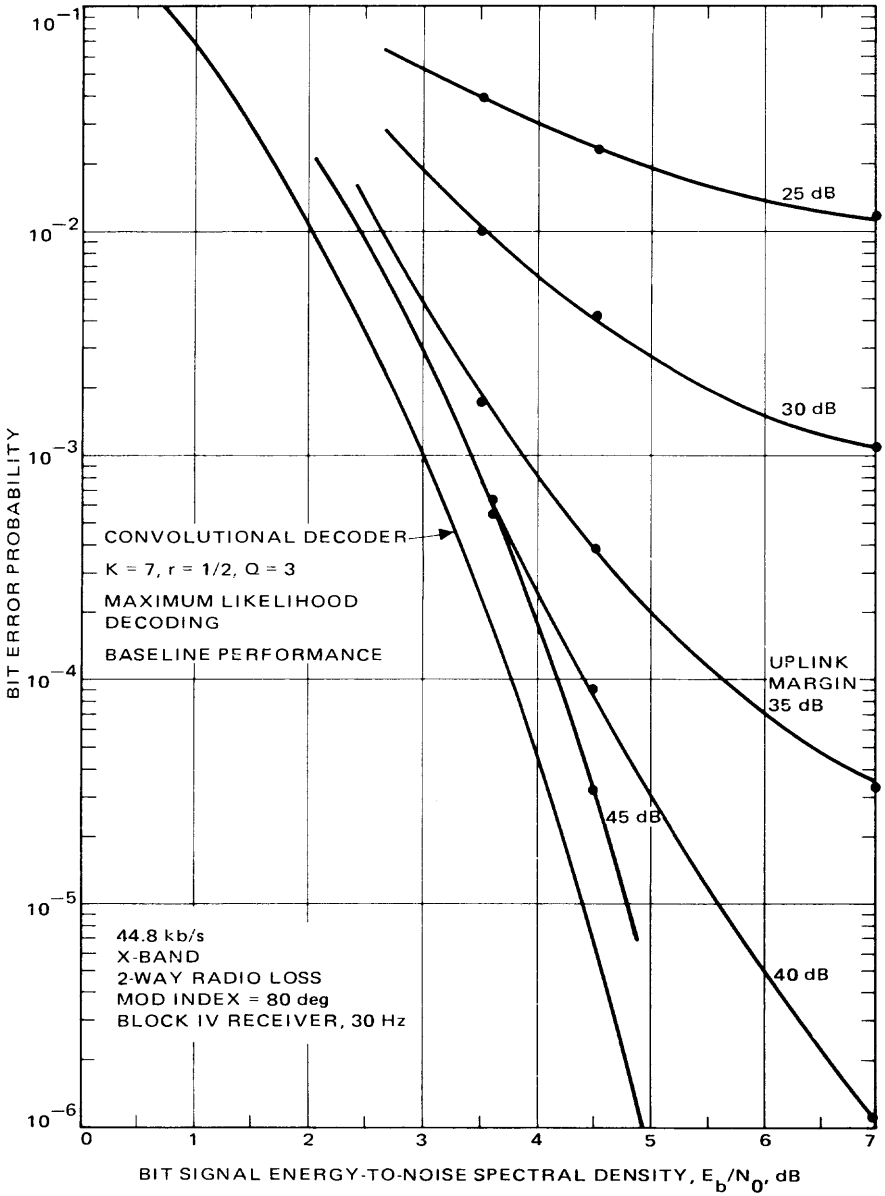


Fig. 5-29. Two-way link bit error probability performance of Viterbi decoder for $K = 7, r = 1/2$, convolutional code

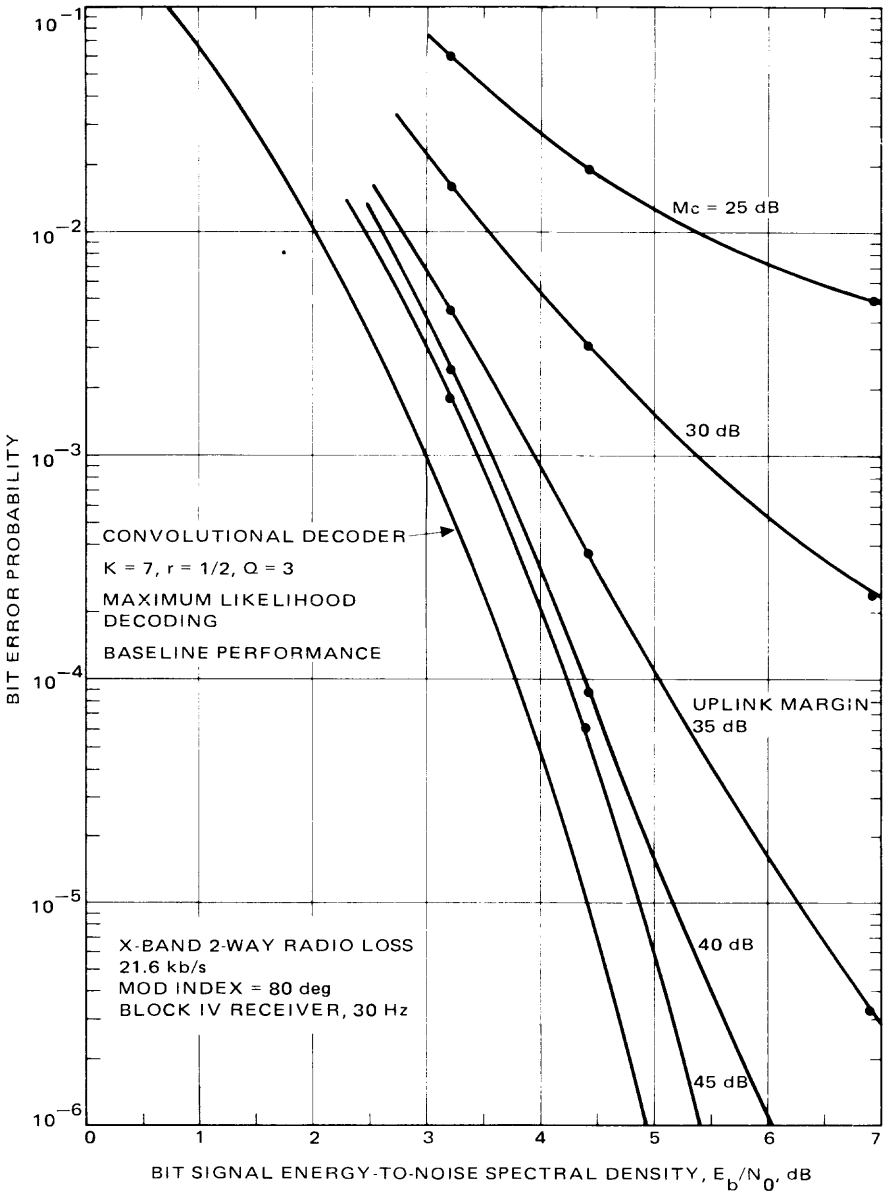


Fig. 5-30. Two-way link bit error probability performance of Viterbi decoder for $K = 7, r = 1/2$, convolutional code

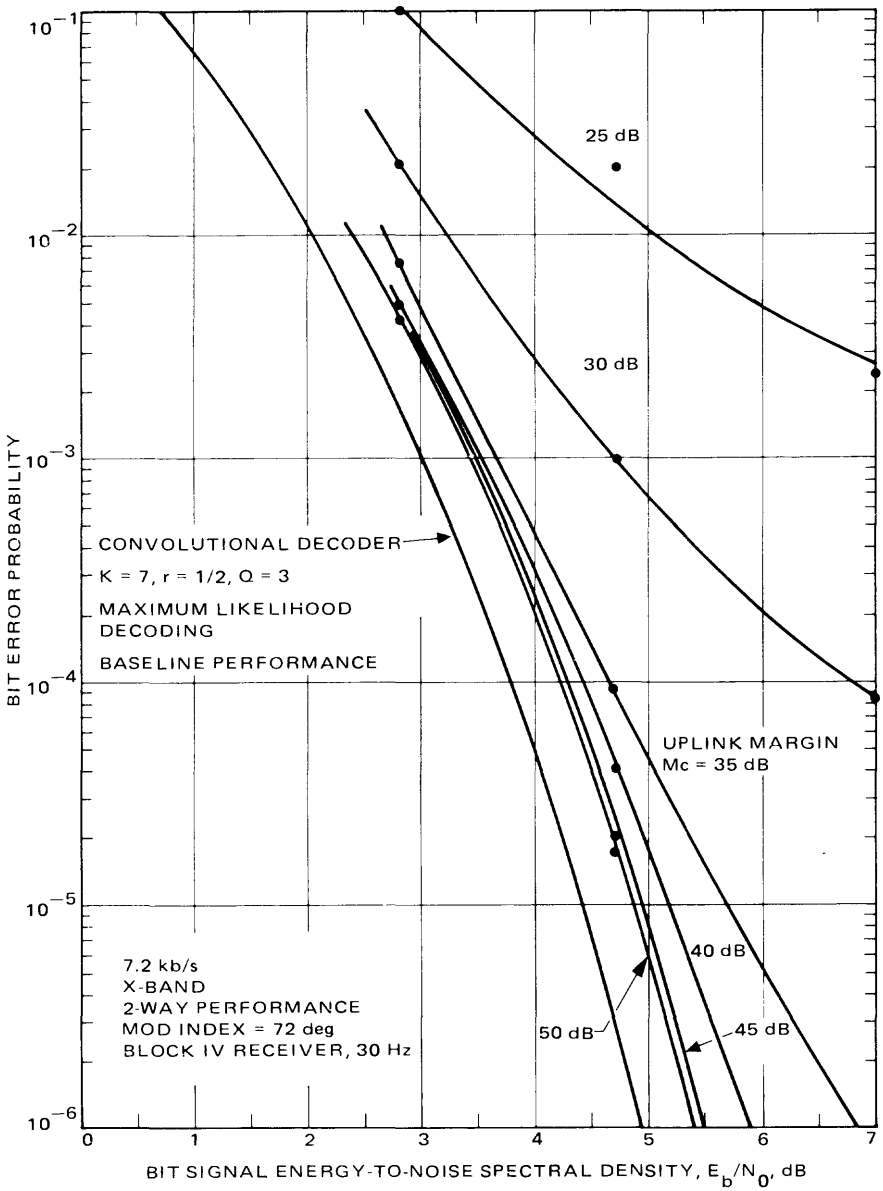


Fig. 5-31. Two-way link bit error probability performance of Viterbi decoder for $K = 7$, $r = 1/2$, convolutional code

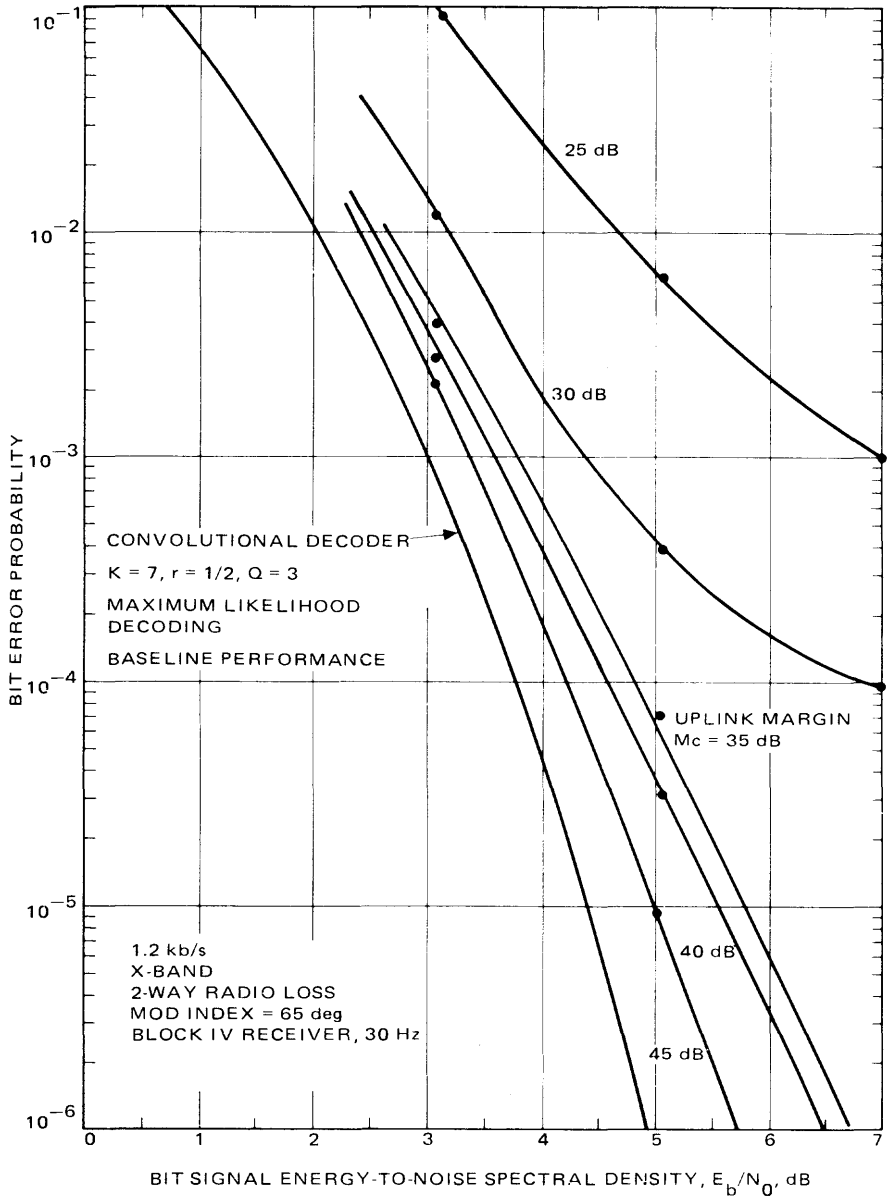


Fig. 5-32. Two-way link bit error probability performance of Viterbi decoder for $K = 7, r = 1/2$, convolutional code

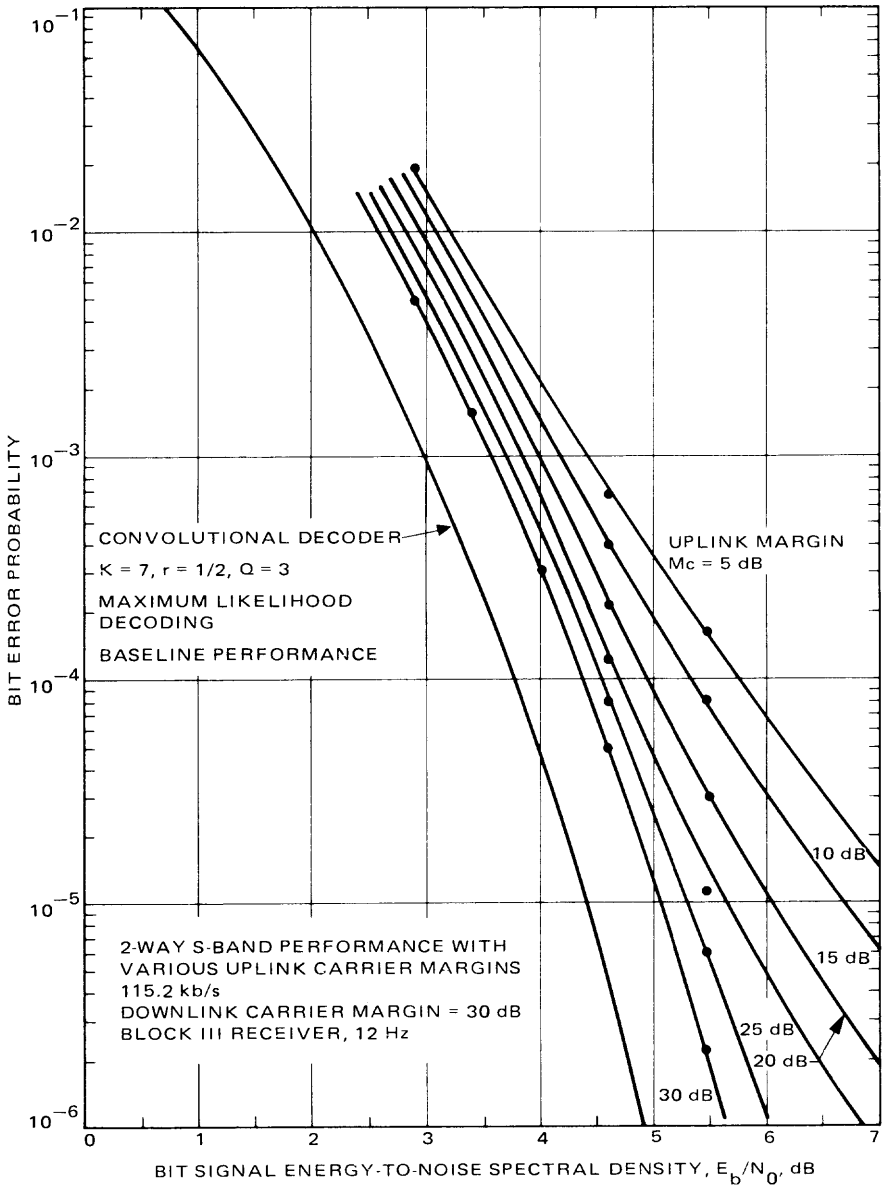


Fig. 5-33. Two-way link bit error probability performance of Viterbi decoder for $K = 7$, $r = 1/2$, convolutional code ($Q = 3$, 115.2 kbps, Block III receiver)

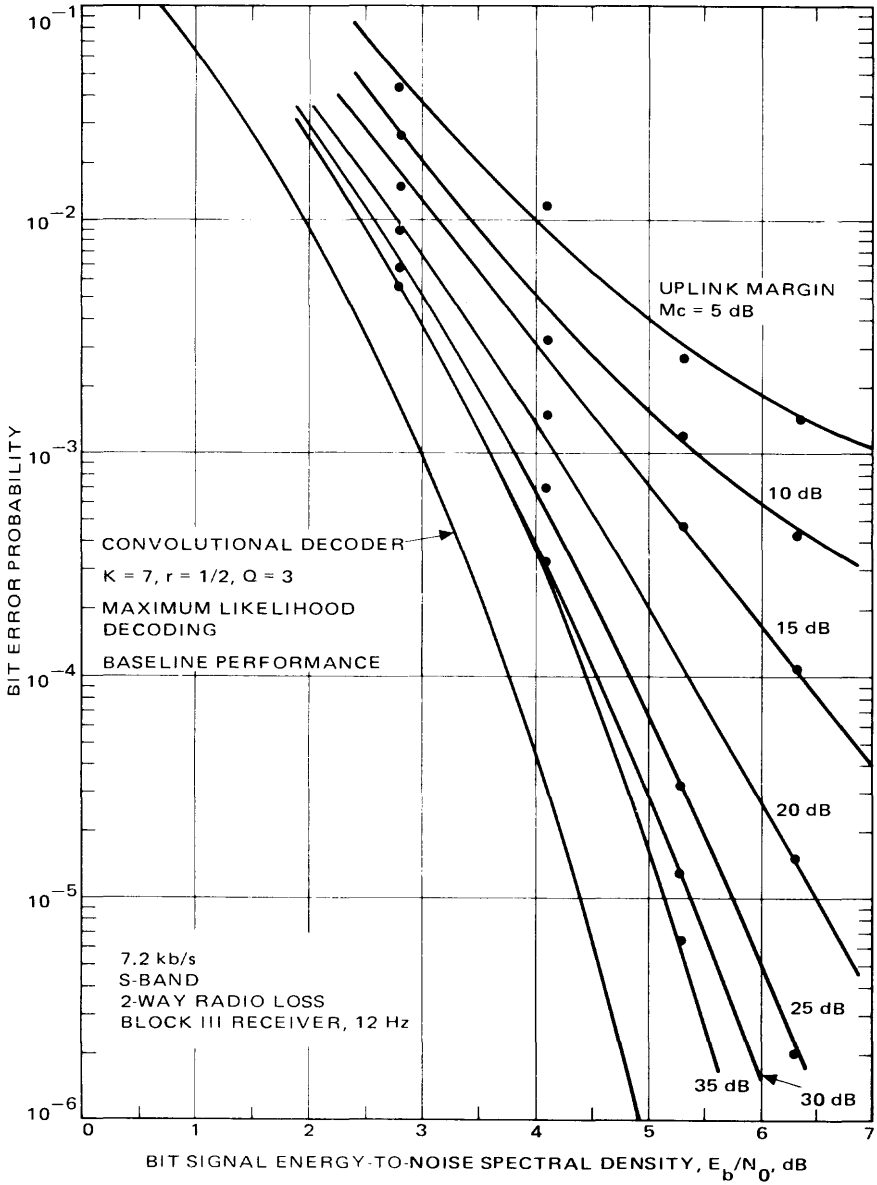


Fig. 5-34. Two-way link bit error probability performance of Viterbi decoder for $K = 7$, $r = 1/2$, convolutional code ($Q = 3$, 7.2 kbps, Block III receiver)

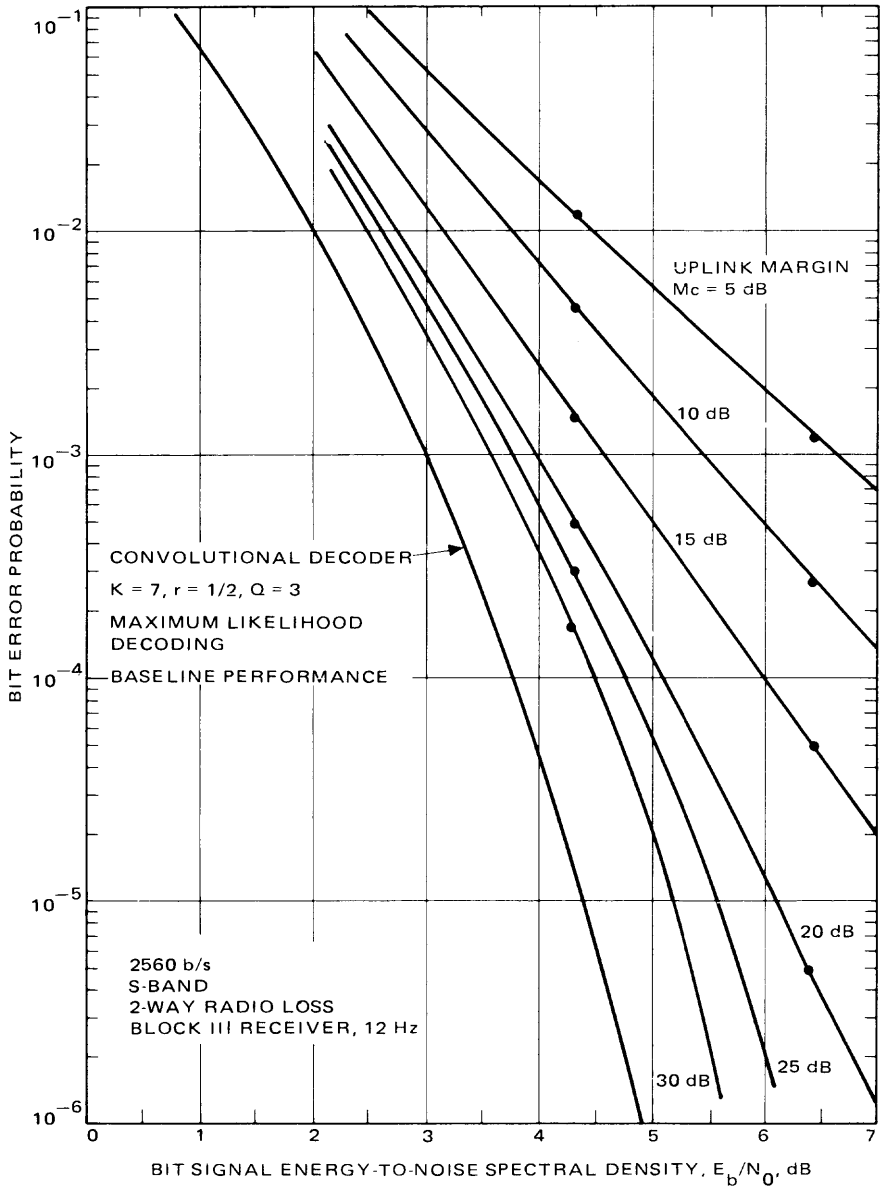


Fig. 5-35. Two-way link bit error probability performance of Viterbi decoder for $K = 7$, $r = 1/2$, convolutional code ($Q = 3$, 2560 bps, Block III receiver)

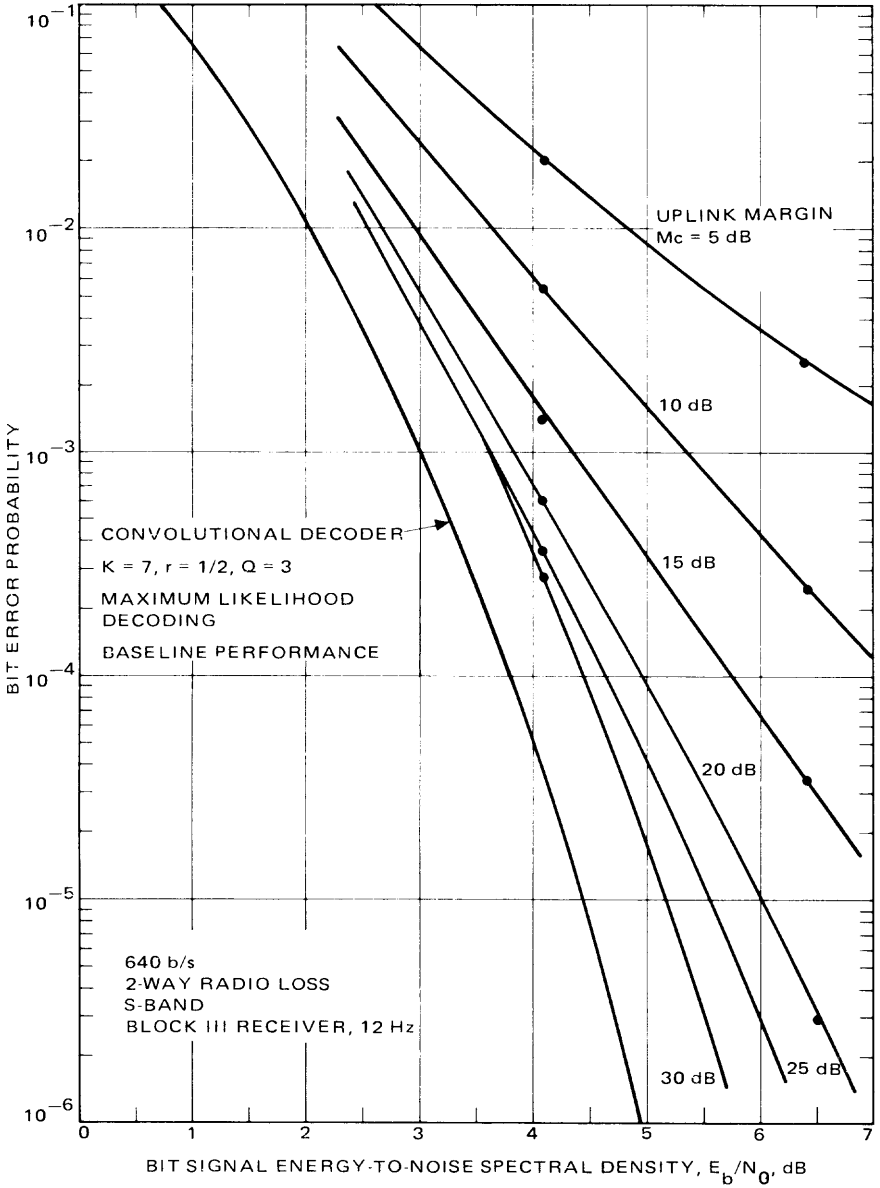


Fig. 5-36. Two-way link bit error probability performance of Viterbi decoder for $K = 7$, $r = 1/2$, convolutional code ($Q = 3$, 640 bps, Block III receiver)

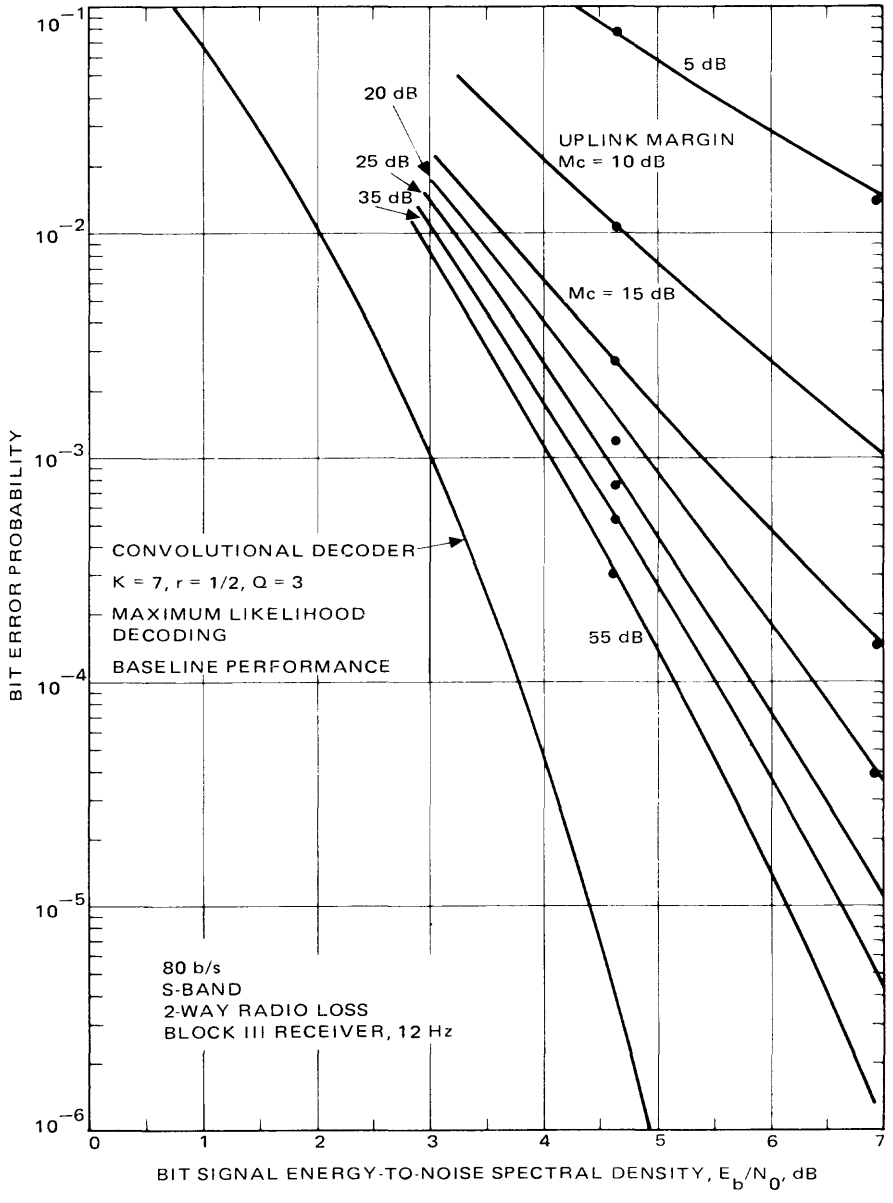


Fig. 5-37. Two-way link bit error probability performance of Viterbi decoder for $K = 7$, $r = 1/2$, convolutional code ($Q = 3$, 80 bps, Block III receiver)

number of wrong hypotheses and moves ahead anyway. The probability associated with this event is called the *undetected error probability* or the bit error probability. The second, called a *buffer overflow*, occurs when the number of computations permitted per frame[4] is exceeded. In this case, the frame cannot be decoded, and is considered an *erased* or *deleted* frame. The frequency with which this event happens is called the *frame deletion probability* or *frame erasure probability*. At this point, the decoder is forced into a resynchronization mode, and until the decoder is again decoding normally, uncorrected data are delivered to the user, containing the original channel errors.

Currently, for the Deep Space Network (DSN), decoding is possible for either of the two encoding implementation configurations shown in Figs. 5-38 and 5-39.

The convolutional codes, which are sequentially decoded, typically have a large enough constraint length so that the undetected error probability of the decoder is negligible compared to the probability that a block cannot be successfully decoded in the time allowed. Thus the limiting factor for sequential decoding is the probability that large amounts of computation are required to decode a frame of the code, rather than the probability of error. Thus, in this part we only consider the effects of noisy carrier reference on deletion probability, although the results can easily be extended to undetected error probability with the same procedure that is going to be used for determining deletion probability.

5.4.3.2.1 *Sequential decoding with noisy carrier reference: high-rate model.* Before going to the derivation of the effect of a noisy carrier on deletion (erasure) probability, we should note two facts. The first fact is that whenever the number of computations needed to decode a code frame is large, its distribution is dominated by single long computations, representing decoder penetration from node level $j - l$ to j, for some j, l. The second fact is that whenever the number of computations is small, the distribution function represents the sum of many small computations.

Now, for high-data-rate systems, if the ratio $\delta = R/2\,B_L$ is very large, where B_L is the phase-locked loop bandwidth and R is the data symbol rate, then it is reasonable to assume that the phase error ϕ is essentially constant over a frame of data. Thus, in this situation we can average the deletion probability curves conditioned on bit SNR (which depends on phase error ϕ) over the distribution of the phase error, to derive an estimate of decoding performance

[4]The encoder is set to a known state at the beginning of each L bits of data. Each L-bit block of data constitutes a frame.

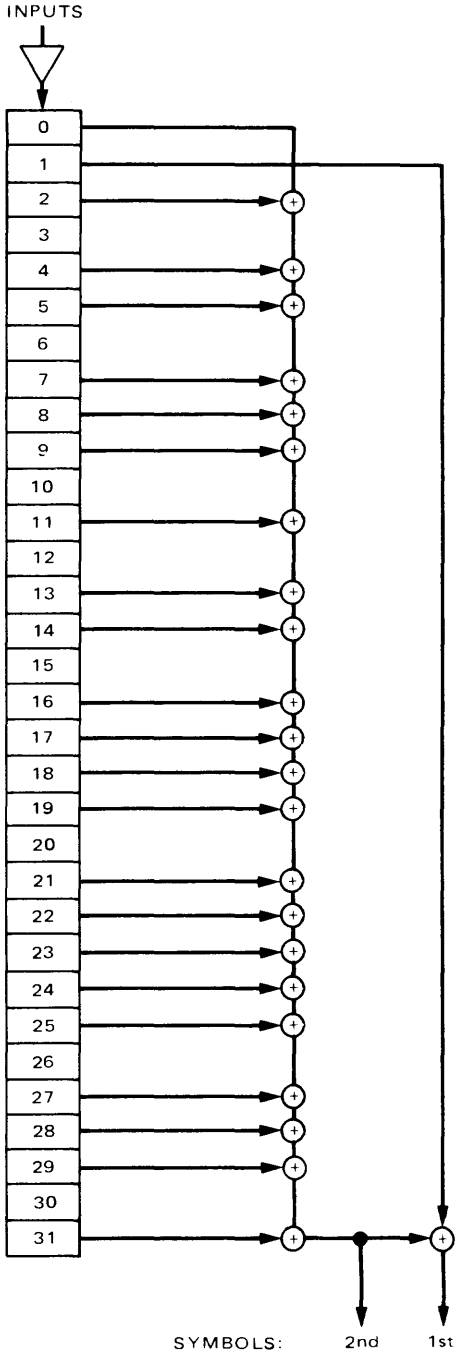


Fig. 5-38. NASA Standard $K = 32$, $r = 1/2$ convolutional encoder without interleaving

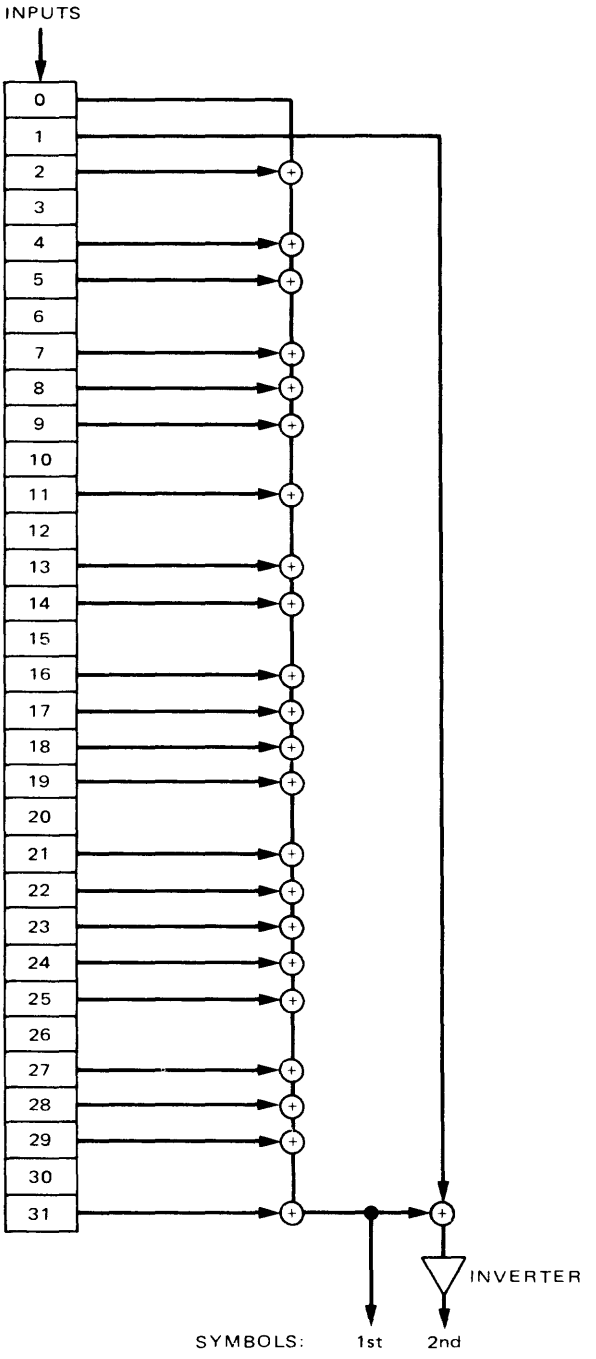


Fig. 5-39. NASA Standard $K = 32$, $r = 1/2$ convolutional encoder with interleaving

with noisy reference. The pseudo-theoretical distribution of decoding computations, which results from treating phase error as constant over an entire frame by averaging the perfect-reference computation distribution over the phase error distribution, has thus two levels of validity: it is an accurate estimate of the low probability, long computation events, which correspond to erasures in a typical system; and it represents an upper bound to the degradation due to phase error for the sums of several short searches, which occur with higher probability and which seldom represent erasures. To obtain numerical results, the baseline deletion probability curves should be approximated by a function of bit $SNR = E_b/N_0$ and an average number of computations per bit N (if we fix the decoder speed μ, which is the number of computations per branch, and the buffer size B branches, we should get $NL = \mu B$). The chosen approximating function [5-35 and 5-36] is

$$P\{deletion \mid SNR\} = \exp\left[\sum_{i=-1}^{1}\sum_{j=0}^{2} A_{ij}(SNR)^j (\ln N)^i\right] \quad (5.4\text{-}23)$$

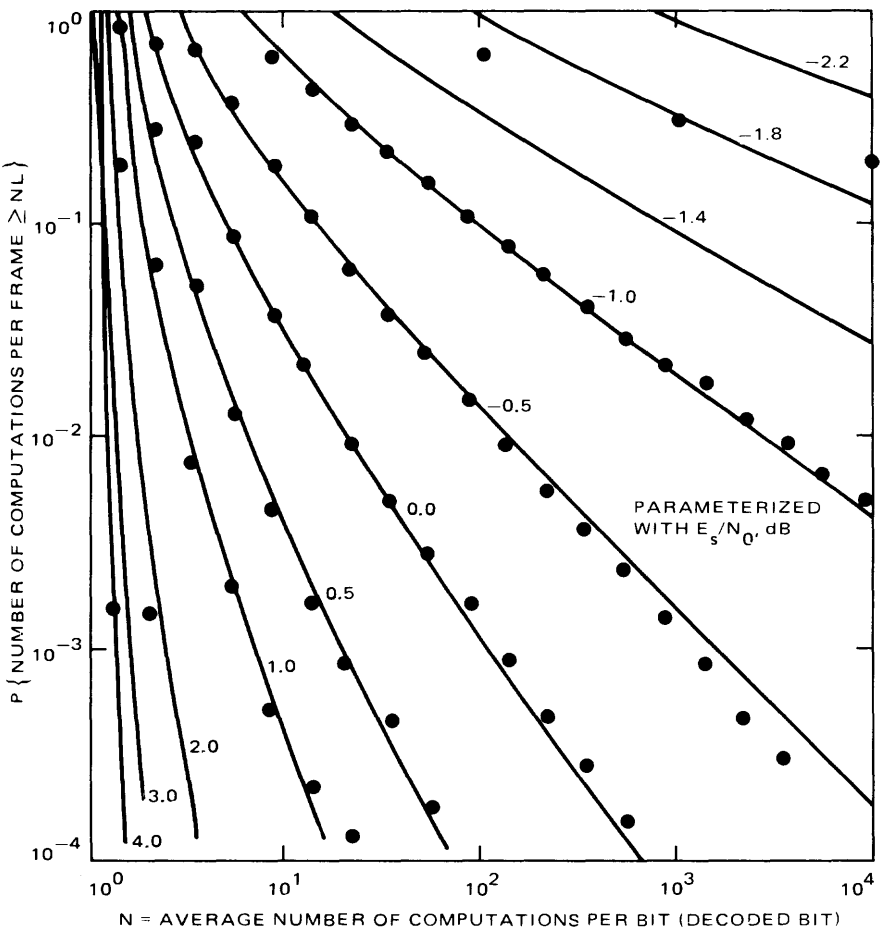
The coefficients A_{ij} were determined by a two-dimensional, least-square polynomial fit. The coefficients depend on frame length L. The coefficients A_{ij} are given in Table 5-3 for frames where L = 192 bits, used in Pioneer spacecraft, and in Table 5-4 for frames where L = 1152 bits, used in Project Helios. This function for frame length L = 1152 bits is shown in Fig. 5-40. Therefore, using

Table 5-3. Coefficients A_{ij} for L = 192

	i		
j	−1	0	1
0	−4.48	5.81	−0.329
1	4.91	−5.01	0.806
2	−1.25	0.995	−0.685

Table 5-4. Coefficients A_{ij} for L = 1152

	i		
j	−1	0	1
0	2.397	8.824	−0.9887
1	−0.5331	−6.788	1.569
2	0.02303	0.8848	−0.8543



$E_s = rE_b$

r = CODE RATE

NL = μB

μ = THE NUMBER OF COMPUTATIONS PER BRANCH

B = BUFFER SIZE IN TERMS OF THE NUMBER OF BRANCHES

E_b = BIT ENERGY

L = FRAME LENGTH = 1152

Fig. 5-40. Distribution of computations for sequential decoding of Helios frame (dots indicate the measurement results)

an approximated function (5.4-23), we can find the probability deletion (given l, μ and B) as

$$P\{deletion\} = \int_{-\pi}^{\pi} P\left\{deletion \left| \frac{E_b}{N_0} \cos^2 \phi \right.\right\} p(\phi)\, d\phi \qquad (5.4\text{-}24)$$

where $P\{deletion|(E_b/N_0) \cos^2 \phi\}$ is given by (5.4-23), replacing $SNR = E_b/N_0$ with $SNR = (E_b/N_0) \cos^2 \phi$, and $p(\phi)$ is the probability density function of the phase error ϕ in a phase-locked loop [5-3] (see also Chapter 3), as given in (5.4-18). Finally, note that evaluation of decoding performance from the high-data-rate models always can serve as an upper bound on decoding performance even if the phase error ϕ is not constant during the entire frame.

5.4.3.2.2 *Sequential decoding with noisy carrier reference: low-rate model.* If the ratio

$$\delta = \frac{R}{2 B_L} = \frac{1}{2 B_L T_b} \qquad (5.4\text{-}25)$$

is small, we can assume that the phase error is independent from bit to bit and can compute the expected loss in signal amplitude into the decoder by averaging over the distribution of ϕ. Therefore [5-37], approximating

$$\frac{1}{T_m} \int_0^{T_m} \cos \phi(t)\, dt \cong 1 - \frac{x}{2} \qquad (5.4\text{-}26)$$

where T_m is the integration time and x is defined as

$$x \triangleq \frac{1}{T_m} \int_0^{T_m} \phi^2(t)\, dt \qquad (5.4\text{-}27)$$

we get effective bit SNR $(E_b/N_0)_{eff}$ as

$$\left(\frac{E_b}{N_0}\right)_{eff} \cong \frac{E_b}{N_0} \int_0^2 \left(1 - \frac{x}{2}\right)^2 p(x)\, dx \qquad (5.4\text{-}28)$$

where $p(x)$ is the probability density function of x. Noting that the autocorrelation function of phase error $\phi(t)$ is approximately

$$R_\phi(\tau) \cong \sigma_\phi^2 \exp(-4 B_L |\tau|) \qquad (5.4\text{-}29)$$

it can then be shown [5-36] and [5-38] that $p(x)$ is approximately

$$p(x) = \sqrt{\frac{a}{\pi}} \exp\left(\sqrt{2ab}\right) \sqrt{\rho x} \exp\left(-a\rho x - \frac{b}{\rho x}\right); \; 0 \leqslant x \qquad (5.4\text{-}30)$$

where

$$\begin{aligned} a &= \frac{B(\delta)}{4}\left[1 + \sqrt{1 + \frac{4}{B(\delta)}}\right] \\ b &= a - 1 + \frac{1}{4a} \\ \delta &= \frac{1}{2 B_L T_m} \end{aligned} \qquad (5.4\text{-}31)$$

$$B(\delta) = \frac{1}{\delta - \frac{\delta^2}{4}\left[1 - \exp\left(-\frac{4}{\delta}\right)\right]}$$

for the low-rate model. Where ϕ is independent from bit to bit, clearly the integration time is equal to bit time T_b, i.e.,

$$T_m = T_b \qquad (5.4\text{-}32)$$

Now, if we replace $(E_b/N_0)_{eff}$ of (5.4-28) with the bit SNR in (5.4-23) we get the deletion probability of sequential decoding for the low-rate model.

5.4.3.2.3 *Sequential decoding with noisy carrier reference: medium-rate model.* For medium data rates between 10 to 1000 bps an approach has been suggested for obtaining the effect of noisy carrier reference on sequential decoding [5-36] as

$$P\{deletion\} = \int_0^2 P\left\{deletion|SNR = \frac{E_b}{N_0}\left(1 - \frac{x}{2}\right)^2\right\} p(x)\, dx + \int_2^\infty p(x)\, dx \qquad (5.4\text{-}33)$$

where $P\{deletion|SNR\}$ is given by (5.4-23) and $p(x)$ is given by (5.4-30). Here, integration time T_m is not T_b and should be chosen appropriately. It

has been argued that T_m for rate $1/n$ convolutional codes should be approximated by

$$T_m \cong 2\left[1 - \frac{1}{\mu B}\log_2\left(1 + \frac{\mu B}{2}\right)\right]T_b \tag{5.4-34}$$

Effects of noisy carrier reference on deletion probability as a function of the total power-to-noise spectral density (P/N_0) are shown in Figs. 5-41 through 5-44 using (5.4-24) and appropriate data-rate models. We note that

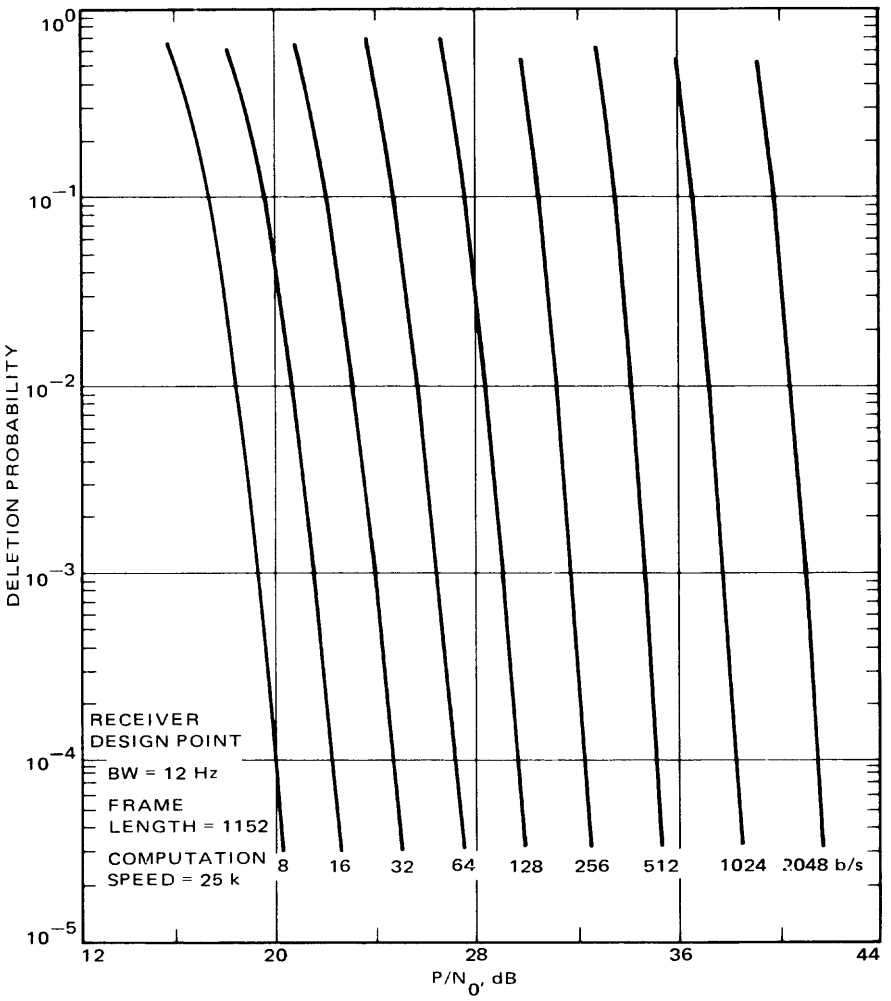


Fig. 5-41. Modeled deletion probability for Helios, *MI* = 42 deg

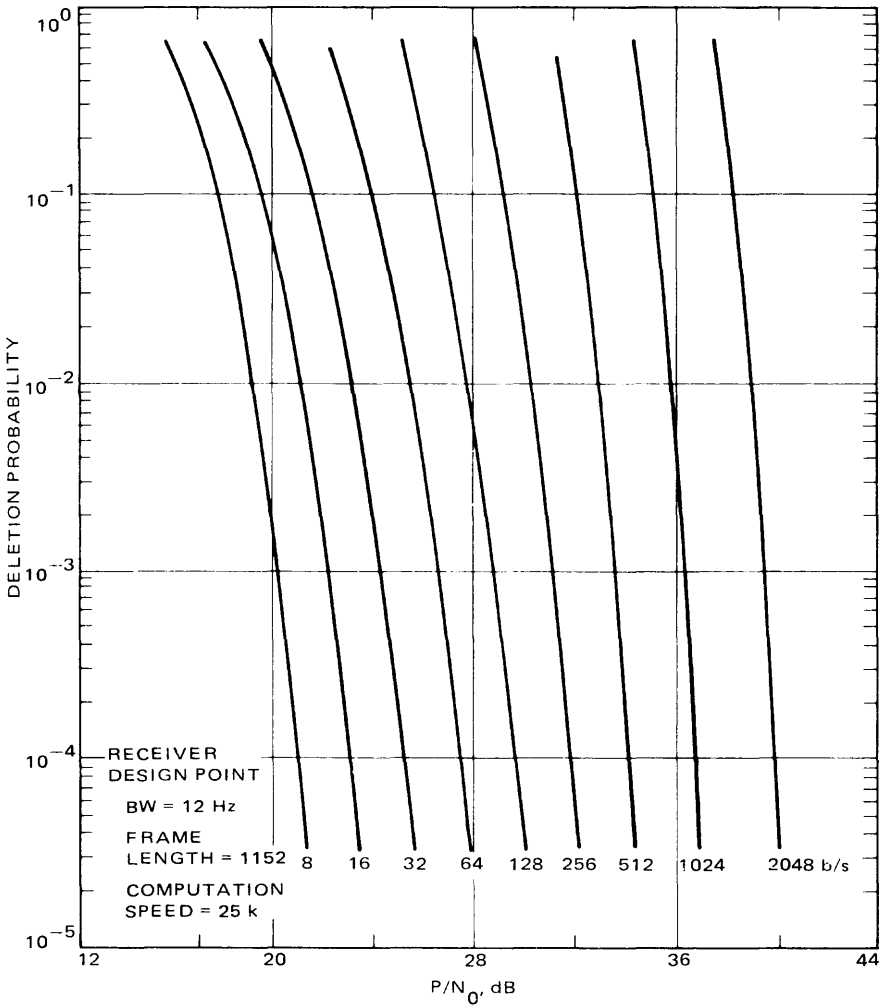


Fig. 5-42. Modeled deletion probability for Helios, *MI* = 55 deg

this method for the medium-rate model can also be used for Viterbi decoding if the integration time T_m is properly redefined.

5.4.4 Reed-Solomon/Viterbi Concatenated Coding

This subsection discusses a particular concatenated coding scheme of importance for deep-space communication. It is the Reed-Solomon/Viterbi concatenated code, consisting of a Reed-Solomon outer code and a convolutional inner code (which is Viterbi decoded). The reader may wish to consult [5-40] for the theory of concatenated coding and [5-41] through [5-46] for more information on the Reed-Solomon/Viterbi concatenated code.

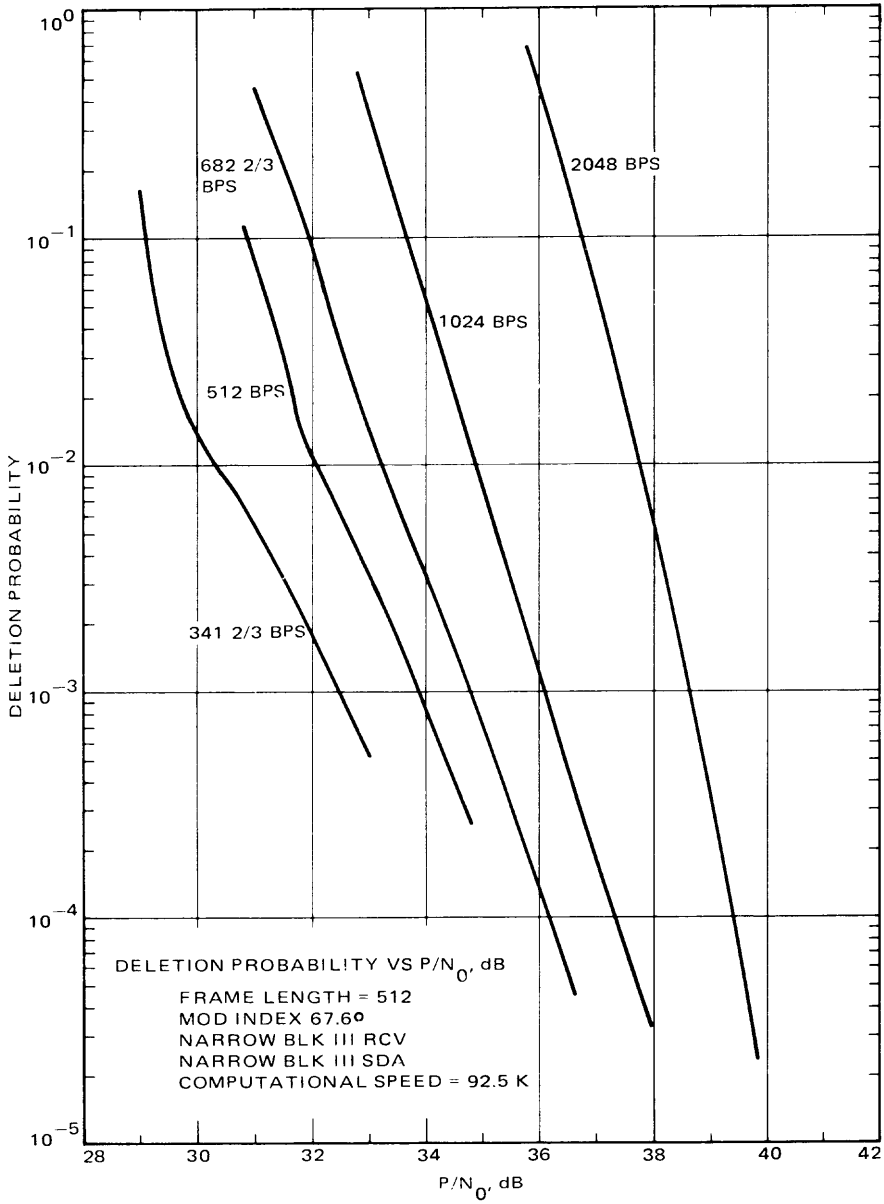


Fig. 5-43. Modeled deletion probability for Pioneer Venus '78, *MI* = 67.6 deg

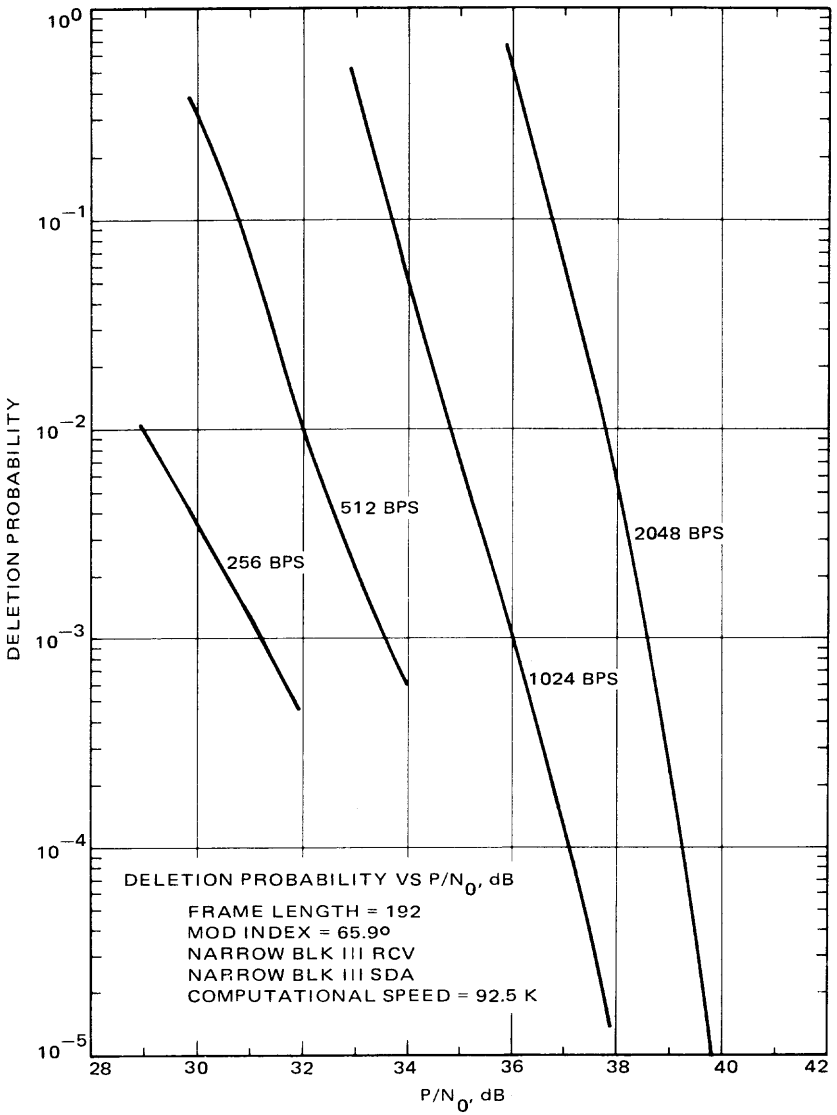


Fig. 5-44. Modeled deletion probability for Pioneer 10, 11, *MI* = 65.9 deg

A block diagram of the concatenated coding system is given in Fig. 5-45. The binary input data sequence is divided into J bit sequences to form symbols over a 2^J-ary alphabet. The Reed-Solomon (RS) code then encodes the symbols such that any combination of E or fewer symbol errors per RS word ($2^J - 1$ symbols per word) can be corrected. A very simple block diagram of a Reed-Solomon (RS) block coder is shown in Fig. 5-46.

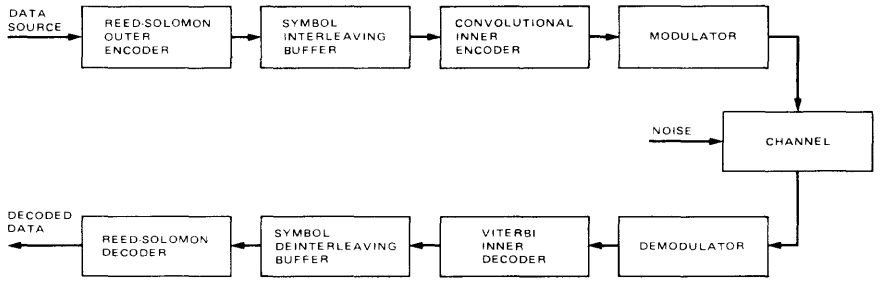


Fig. 5-45. Concatenated coding system block diagram

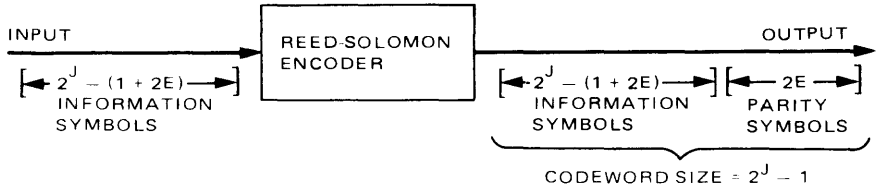


Fig. 5-46. Reed-Solomon encoder input–output relations

The RS code is nonbinary. An RS symbol consists of a sequence of J bits so that there are 2^J possible RS symbols. All coding and decoding operations involve RS symbols, not individual bits.

$$K = 2^J - (1 + 2E) \tag{5.4-35}$$

information symbols (or $J[2^J - (1 + 2E)]$ information bits) from some data source enter the RS coder to the left. The result of coding operations is a codeword of length

$$N = 2^J - 1 \tag{5.4-36}$$

symbols, of which the first $2^J - (1 + 2E)$ are the same symbols as those entering to the left. This makes the code *systematic*. The remainder of the codeword is filled in with $2E$ *parity symbols*. An RS symbol is in error if any of the J bits making up the symbol are in error. E represents the number of correctable RS symbol errors in an RS codeword. That is, if E or less RS symbols are in error in any way, the decoder will be capable of correcting them.

We put attention on the specific system with $J = 8$, $E = 16$. The basic codeword structure of this specific codeword with $J = 8$, $E = 16$, is given in Fig. 5-47. Note that the overhead associated with the parity symbols is only around 15 percent. If desired, a "quick look" at the data (information bits) would still be possible since the code is systematic. Typically, the inner convolutional code corrects enough errors so that a high-code-rate outer code can reduce the error probability to the desired level.

The interleaving buffers are required because the inner Viterbi decoder errors tend to occur in bursts, which occasionally are as long as several constraint lengths. Without interleaving, Viterbi decoder burst error events would tend to occur within one RS codeword. That one codeword would have to correct all of these errors. Thus over a period of time there would be a *tendency* for some codewords to have "too many" errors to correct (i.e., greater than 16).

Figure 5-48 [5-41] illustrates two interleaver methods with an encoder code array of I RS codewords (for I levels of interleaving). The figure gives the sequence for reading and writing symbols from interleavers represented by the code array. Interleaver B has the advantage that all of the data symbols of the code array do not have to be stored before the symbols can be transmitted to the convolutional encoder. However, when an RS word error is made, it will, in general, cause more source block errors than interleaver A. The purpose of interleaving and de-interleaving is to make the RS symbol errors, at the input of the RS decoder, independent of each other and to distribute the RS symbol errors uniformly; in other words, to distribute the

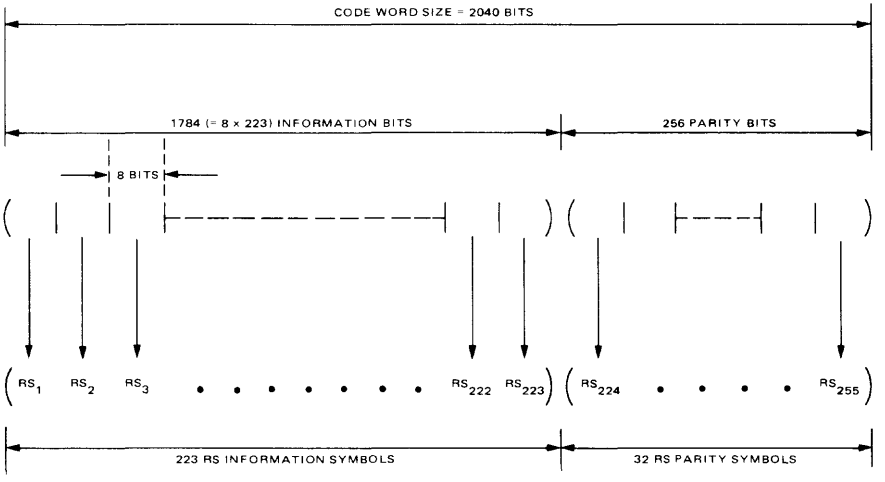
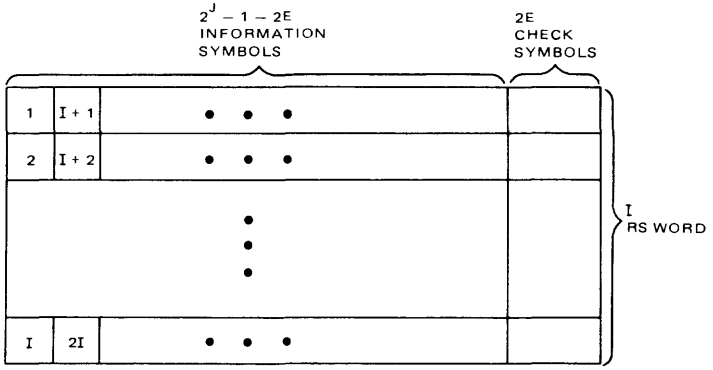


Fig. 5-47. Reed-Solomon codeword structure, $J = 8, E = 16$



INTERLEAVER A

WRITE SYMBOLS INTO LOCATIONS	$1, I+1, 2I+1, \ldots,$ $2, I+2, \ldots, I(2^J-1)$
READ SYMBOLS FROM LOCATIONS	$1, 2, 3, \ldots, I(2^J-1)$

INTERLEAVER B

WRITE SYMBOLS INTO LOCATIONS	$1, 2, 3, \ldots, I(2^J-1)$
READ SYMBOLS FROM LOCATIONS	$1, I+1, 2I+1, \cdots, 2, I+2, \cdots, I(2^J-1)$

Fig. 5-48. Encoder code array representation of interleaving methods

burst errors out of the Viterbi decoder among several codewords. The performance of the RS decoder is severely degraded by highly correlated errors among several successive symbols.

The level of interleaving I corresponds to the number of RS codewords involved in the interleaving and de-interleaving operation. Interleaving and de-interleaving operations over a channel can be explained simply by considering two $I \times 2^J - 1$ matrices, one at the input of the channel and one at the output (see Fig. 5-48). For interleaving, put the codeword with length $2^J - 1$ in rows $1, 2, \cdots, I$ of the matrix, then transmit the symbols of columns $1, 2, \cdots, 2^J - 1$ through the channel. For de-interleaving, do the reverse operation. For the Reed-Solomon code (255, 223), simulation results have shown that the interleaving level of $I = 16$ is adequate for making the RS symbol errors independent of each other. Thus, henceforth we can assume that the interleaving, convolutional code, AWGN channel, Viterbi decoder, and de-interleaving in tandem create an equivalent 2^J-ary Discrete Memoryless Channel (DMC) with transition error probability π.

In decoding the RS codewords, essentially three events may happen. The first event happens if there are E or less RS symbol errors in a codeword. In this case the decoder successfully corrects the errors and outputs the correct information block. The second event happens if the number of RS symbol

errors in a codeword are more than E but the corrupted codeword is not close to any other codeword within the distance of E symbols. In this case the RS decoder *fails to decode* and outputs the first $2^J - 1 - 2E$ undecoded information symbols that may contain symbol errors. The third event happens if the number of RS symbol errors in a codeword is more than E, and the corrupted codeword is closer to some other codeword within the distance of E symbols. In this case the decoder is fooled, *decodes incorrectly*, and outputs a wrong information block. In other words, it claims the decoded block as a correct one and by doing this it may add up to E more symbol errors to the decoded output. Fortunately for most of the RS codes of interest with large alphabet size, in particular for the (255, 223) RS code, the probability that the third event happens is very small [5-47]. This probability has very little effect on the error probability performance of an RS code in the range of interest. In [5-47] it has been shown that the probability of the third event, i.e., an incorrect decoding event, is less than 10^{-13}. Therefore, for the practical range of interest in error probability performance, it almost surely can be assumed that only the first and second events happen.

5.4.4.1 Error performance.

(1) RS codeword error probability:

Denote the word error probability for an (N, K) RS code by $P_w(N, E)$. An RS codeword is in error when there are more than E channel symbol errors in a received codeword $\mathbf{y}$. Therefore, we should consider all combinations of j symbol errors within the $2^J - 1$ symbol, for all $j > E$.

The RS codeword error probability is (see also [5-41])

$$P_w(N, E) = P(\hat{\mathbf{x}}_m \neq \mathbf{x}_m) = \sum_{j=E+1}^{N} \binom{N}{j} \pi^j (1 - \pi)^{N-j} \qquad (5.4\text{-}37)$$

where π is the RS decoder input symbol error probability; i.e., π is the probability for one or more bit errors occurring in a string of J consecutive bits. For the experimentally determined statistics of π with a non-noisy carrier reference, the resulting RS word error probability is shown in Fig. 5-49.

(2) RS information-block error probability:

Note that if a codeword is in error it is not necessary that the corresponding information block be in error. This error probability is important for source coding (data compression). Denote the information-block error probability by P_I. Based on our assumptions, we should subtract the probability of all possible patterns of symbol errors that

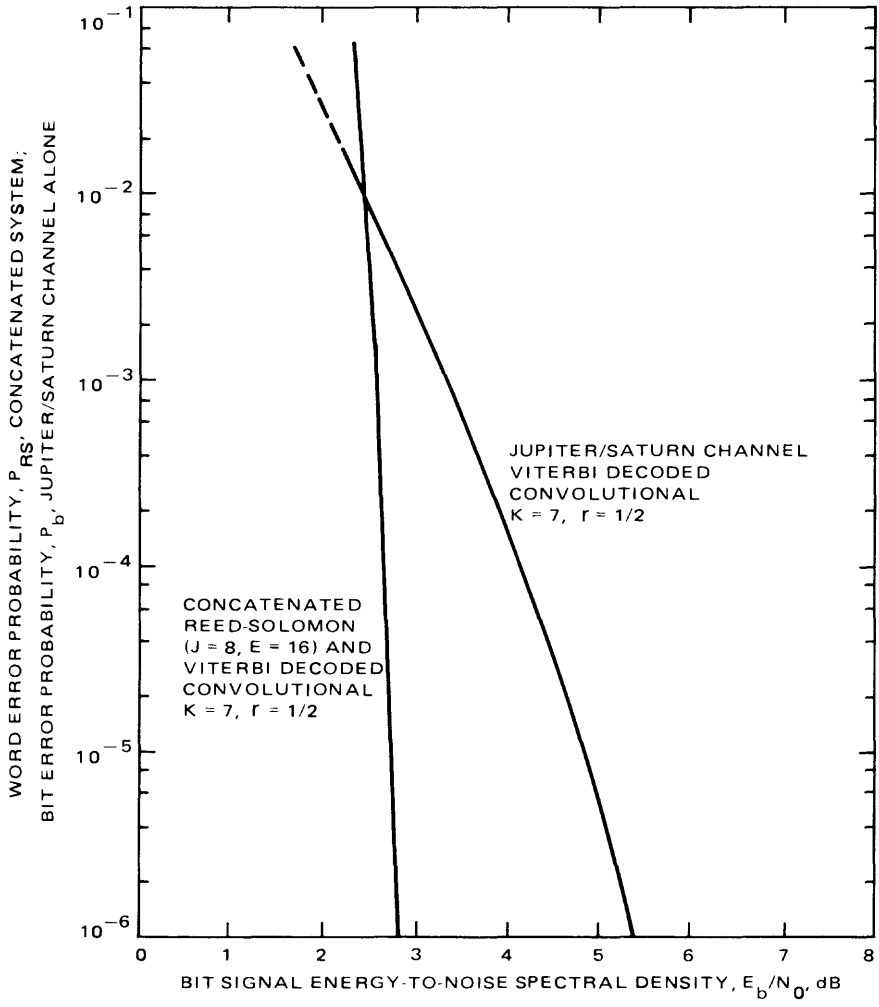


Fig. 5-49. Concatenated Reed-Solomon/Viterbi decoder performance curves

happen only in parity check symbols, from P_w. This implies that the information block is still correct. Thus, we have

$$P_I = P(\hat{\mathbf{u}}_m \neq \mathbf{u}_m) = \sum_{j=E+1}^{N} \binom{N}{j} \pi^j (1-\pi)^{N-j} - \sum_{j=E+1}^{2E} \binom{2E}{j} \pi^j (1-\pi)^{N-j} \tag{5.4-38}$$

Obviously, in the practical range of error probability, the last term can be ignored and we have approximately,

$$P_I \cong P_w(N, E) \tag{5.4-39}$$

(3) RS symbol error probability (out of the RS decoder):

Denote the RS symbol error probability by P_s. When the RS decoder fails to decode, the symbol error pattern is exactly the same at the input and output of the RS decoder. Now, in order that an output RS symbol at, say, position k, $k = 1, 2, \cdots, 2^J - 1 - 2E$ be in error, the input RS symbol at the same position k should be in error. In addition, there should be E or greater input RS symbol errors in the remaining $N - 1$ symbols of the received codewords out of the Viterbi decoder, i.e., other than at the position k. The probability that the input RS symbol at position k be in error is π. The probability that there will be E or more input RS symbol errors in the remaining positions, i.e., $N - 1$ places of received codeword, is

$$P_w(N - 1, E - 1) = \sum_{j=E}^{N-1} \binom{N-1}{j} \pi^j (1 - \pi)^{N-1-j} \tag{5.4-40}$$

Thus

$$P_s = \pi P_w(N - 1, E - 1) \tag{5.4-41}$$

(4) RS bit error probability:

Denote the RS bit error rate by $P_b(RS)$. In order that a bit at, say, position m, $m = 1, 2, \cdots, J$ within the RS output symbol at, say, position k, $k = 1, 2, \cdots, 2^J - 1 - 2E$ be in error, the bit at the position m within the input RS symbol at position k of an uncoded codeword should be in error and there should be E or more RS symbol errors in the received codeword out of the Viterbi decoder in m other positions than k. But the probability that a bit would be in error in an undecoded RS codeword is the bit error rate of the Viterbi decoder denoted by $P_b(Viterbi)$. The probability that there are E or more RS symbol errors in an undecoded codeword in positions other than k is given by (5.4-40). Thus, the RS bit error probability is [5-47]

$$P_b(RS) = P_b(Viterbi)\, P_w(N - 1, E - 1) \tag{5.4-42}$$

5.4.4.2 A functional model for the RS decoder input-symbol error probability. Based on simulation results of Linkabit for π given in [5-42] and using a least-squares curve fit, we proposed a simple model for π as

$$\pi\left(\frac{E_b}{N_0}\right) = \begin{cases} \exp\left(\beta_0 - \beta_1 \dfrac{E_b}{N_0}\right); & \dfrac{E_b}{N_0} \geqslant \dfrac{\beta_0}{\beta_1} \\ 1 - 2^{-J}; & 0 \leqslant \dfrac{E_b}{N_0} \leqslant \dfrac{\beta_0}{\beta_1} \end{cases} \tag{5.4-43}$$

where $\beta_0 = 4.9551$, $\beta_1 = 5.2275$ and E_b/N_0 is the bit SNR of the Viterbi channel (inner code). Figure 5-50 shows the baseline symbol error probability π for the (255, 223) RS code vs E_b/N_0 of the Viterbi channel.

Based on our functional model for π, $P_b(RS)$, P_s, and P_w are shown in Fig. 5-51 for concatenated coding with the (255, 223) RS code and the (7, 1/2) convolutional code, versus the E_b/N_0 of the concatenated channel. Note that the code rate of the RS code is

$$RS\ code\ rate = \frac{2^J - 1 - 2E}{2^J - 1} = \frac{K}{N} \tag{5.4-44}$$

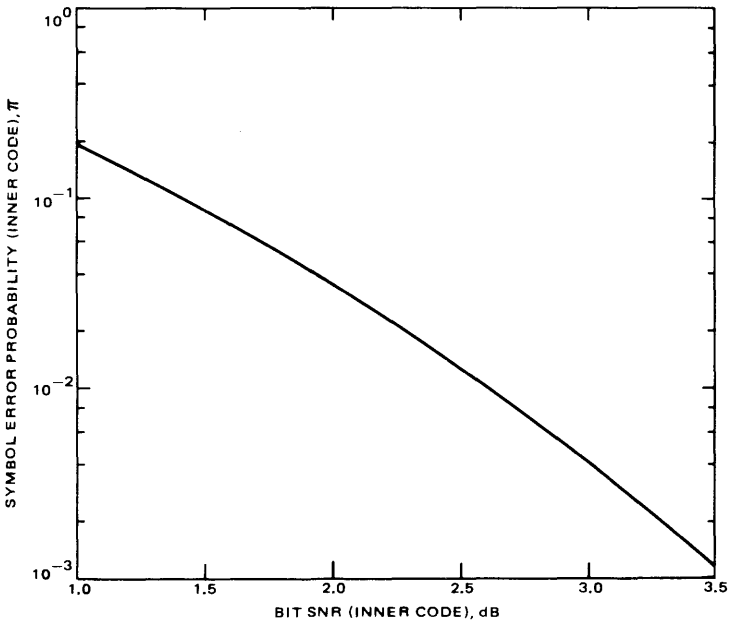


Fig. 5-50. Symbol error probability π for (255, 223) Reed-Solomon code vs bit SNR of the Viterbi inner code

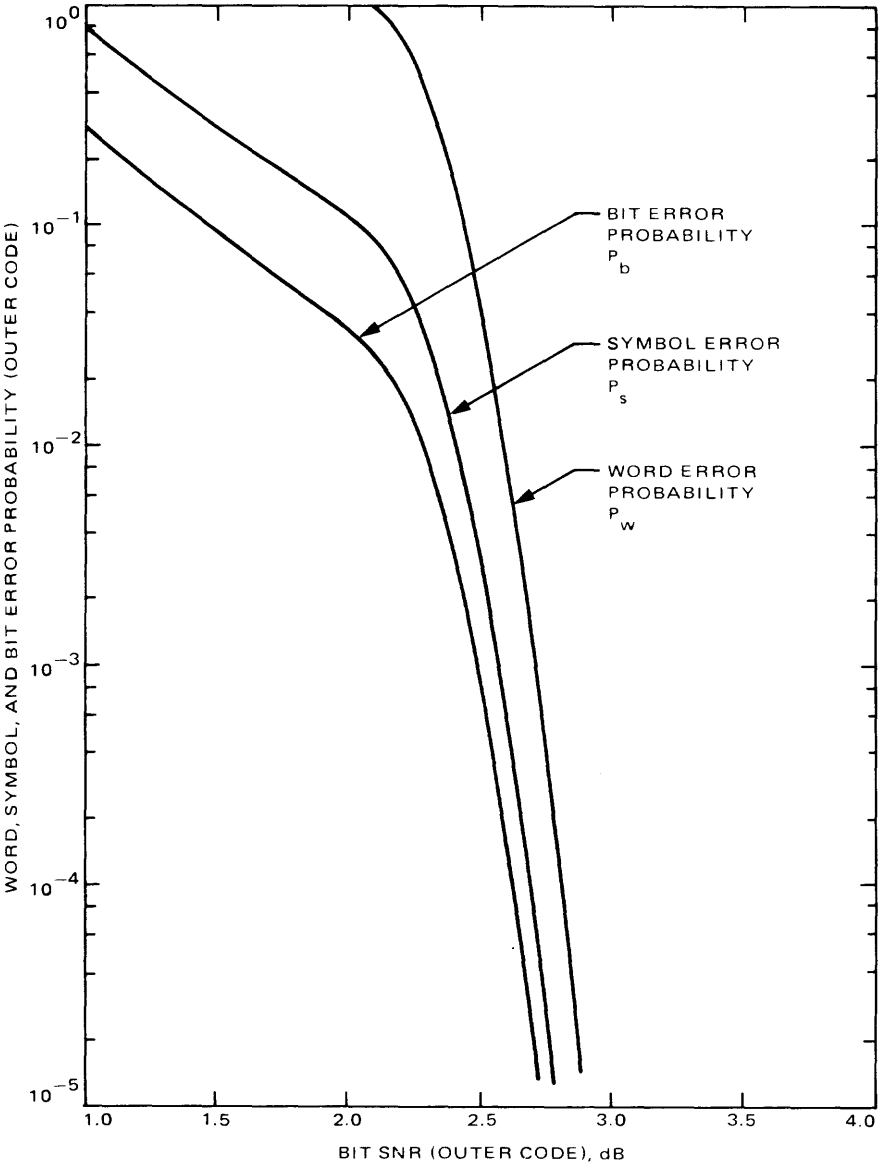


Fig. 5-51. P_b, P_s, and P_w performance curves

5.4.4.3 Effects of phase jitter. If the data rate is large enough compared with the PLL loop bandwidth so that the phase error does not vary significantly during the Viterbi decoder error bursts, then the phase error can be assumed to be constant during one RS symbol. In addition, since we have $I = 16$ level of interleaving, the phase errors affect RS input symbols approximately independently. Noting the approach we have taken for derivation of P_w, P_s, and $P_b(RS)$, we have

$$P_w = E\{P_w(\phi_1, \phi_2, \cdots, \phi_N)\} = \sum_{j=E+1}^{N} \binom{N}{j} \overline{\pi(\phi)}^j (1 - \overline{\pi(\phi)})^{N-j} \quad (5.4\text{-}45)$$

and

$$P_b = \overline{P_b(Viterbi|\phi)} \sum_{j=E}^{N-1} \binom{N-1}{j} \overline{\pi(\phi)}^j (1 - \overline{\pi(\phi)})^{N-1-j} \quad (5.4\text{-}46)$$

where

$$\overline{\pi(\phi)} = \int_{-\pi}^{\pi} \pi\left(\frac{E_b}{N_0} \cos^2 \phi\right) p(\phi)\, d\phi \quad (5.4\text{-}47)$$

$p(\phi)$ is the probability density function of the phase jitter given by (5.4-18), and $\pi(\cdot)$ is given by (5.4-43), replacing E_b/N_0 with $(E_b/N_0)\cos^2\phi$. $\overline{\pi(\phi)}$ is shown in Fig. 5-52 for various values of loop SNR ρ. The effect of a noisy carrier reference on the Reed-Solomon/Viterbi bit error rate performance given by (5.4-46) is shown in Fig. 5-53. Note that in (5.4-46) $\overline{P_b(Viterbi|\phi)}$ can be computed from (5.4-17). For the effect of a noisy carrier reference on RS/Viterbi, we also recommend [5-46] to the interested reader. The effect of node synchronization loss of the Viterbi decoder on RS bit error probability is discussed in [5-47].

5.4.4.4 Effects of slow fading on the RS/Viterbi decoder performance. The RF telemetry signal can be expressed as

$$s(t) = \sqrt{2} A \sin(\omega_0 t + m(t)) \quad (5.4\text{-}48)$$

where $m(t)$ contains the telemetry information that can be a subcarrier, phase-modulated with a binary data stream. Now if this signal is passed through a fading channel, at the output we have

$$r(t) = \sqrt{2}\, a(t) \sin(\omega_0 t + m(t) + \theta(t)) + n(t) \quad (5.4\text{-}49)$$

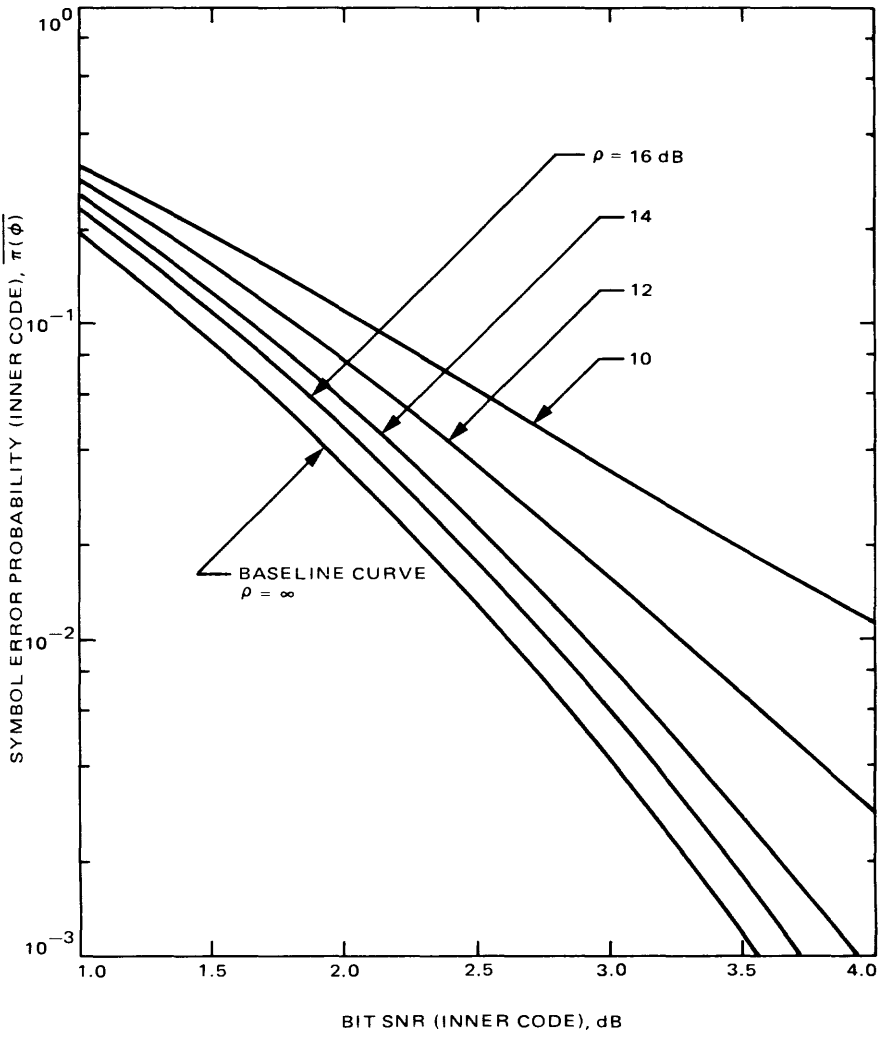


Fig. 5-52. Effects of noisy carrier reference on symbol error probability π

where $a(t)$ is a random amplitude process and $\theta(t)$ is a random phase process. If $a(t)$ and $\theta(t)$ change slowly with time, and the spectral bandwidths of $a(t)$ and $\theta(t)$ are narrow with respect to the average loop bandwidth of the receiver PLL, then $\theta(t)$, together with the carrier phase, will be tracked.

Here we consider a slowly varying fading channel with perfect tracking. We can have Rayleigh, Rician, or Lognormal channels. Consider first the Rician

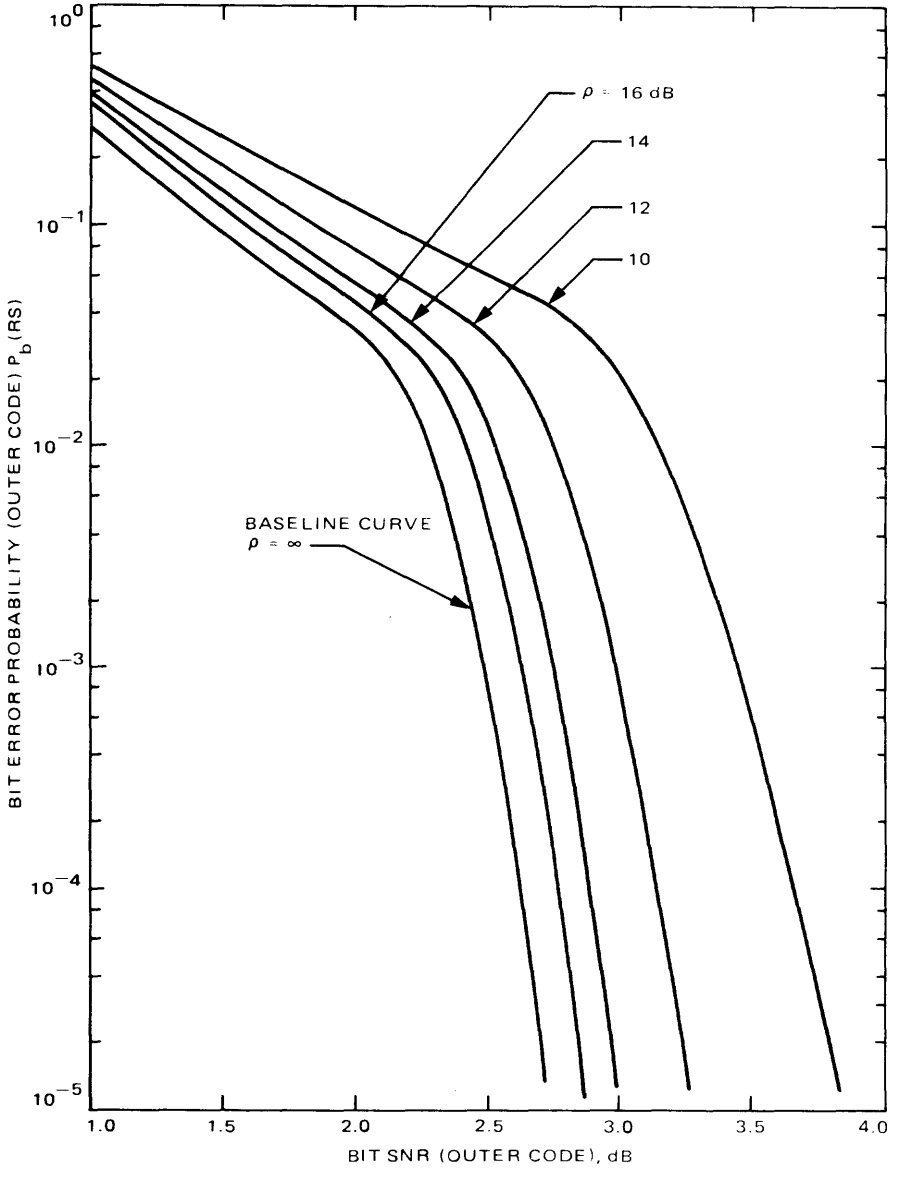


Fig. 5-53. Effects of noisy carrier reference on concatenated Reed-Solomon/Viterbi bit error rate P_b(RS)

channel, since it is Rayleigh with an added specular component. If we pass a transmitted signal (5.4-48) through a Rician channel we get

$$r(t) = \sqrt{2}A \sin(\omega_0 t + m(t)) + \sqrt{2}n_c(t) \cos(\omega_0 t + m(t))$$

$$+ \sqrt{2}n_s(t) \sin(\omega_0 t + m(t)) + n(t) \tag{5.4-50}$$

where n_c and n_s are zero mean Gaussian random processes, each with variance σ^2. Then the received signal amplitude $a(t)$ has the Rician probability density function.

$$p(a) = \begin{cases} \dfrac{a}{\sigma^2} \exp\left(-\dfrac{(a^2 + A^2)}{2\sigma^2}\right) I_0\left(\dfrac{aA}{\sigma^2}\right); & a \geq 0 \\ 0; & \text{otherwise} \end{cases} \tag{5.4-51}$$

Define

$$\gamma^2 \triangleq \frac{A^2}{2\sigma^2} \tag{5.4-52}$$

as the ratio of the specular power to the fading power. Suppose the received power is P where

$$P = A^2 + 2\sigma^2 \tag{5.4-53}$$

Let's normalize $a(t)$ as

$$y(t) = \frac{a(t)}{\sqrt{P}} \tag{5.4-54}$$

Then

$$p(y) = \begin{cases} 2y(1 + \gamma^2) \exp\left[-(1 + \gamma^2)y^2 - \gamma^2\right] I_0\left[2y\sqrt{\gamma^2(1 + \gamma^2)}\right]; & y \geq 0 \\ 0; & \text{otherwise} \end{cases} \tag{5.4-55}$$

with

$$\overline{y^2} = 1 \tag{5.4-56}$$

For a Rayleigh channel, we don't have a specular component, which means $A = 0$ or $\gamma^2 = 0$; then the pdf for y is

$$p(y) = \begin{cases} 2y \exp(-y^2); & y \geqslant 0 \\ 0; & \text{otherwise} \end{cases} \tag{5.4-57}$$

Figure 5-54 shows $p(y)$, given in (5.4-55) and (5.4-57). For a Lognormal channel the received signal amplitude is of the form

$$a(t) = A\, e^{x(t)} \tag{5.4-58}$$

where $x(t)$ is a Gaussian random process with variance σ_x^2 and the pdf of normalized $a(t)$ as in (5.4-54) is

$$p(y) = \begin{cases} \dfrac{1}{y\sqrt{2\pi}\,\sigma_x} \exp\left[-\dfrac{(\ln y + \sigma_x^2)^2}{2\,\sigma_x^2}\right]; & y \geqslant 0 \\ 0; & \text{otherwise} \end{cases} \tag{5.4-59}$$

Therefore in all cases we can assume the received signal is

$$r(t) = \sqrt{2P}\, y \cos\left[\omega_0 t + m(t) + \theta(t)\right] + n(t) \tag{5.4-60}$$

where the pdf of y is given by (5.4-55), (5.4-57), or (5.4-59).

If the signal amplitude changes very slowly in comparison to the bit rate, y will remain constant over a large number of bits. On the other hand, E levels of interleaving make y affect RS symbols independently. Therefore, we have

$$\pi = \overline{\pi(y)} = \int_0^\infty \pi\left(\frac{E_b}{N_0} y^2\right) p(y)\, dy \tag{5.4-61}$$

and

$$P_b(Viterbi) = \overline{P_b(Viterbi|y^2)} = \int_0^\infty f\left(\frac{E_b}{N_0}\, y^2\right) p(y)\, dy \tag{5.4-62}$$

where $f(\cdot)$ is given by (5.4-16).

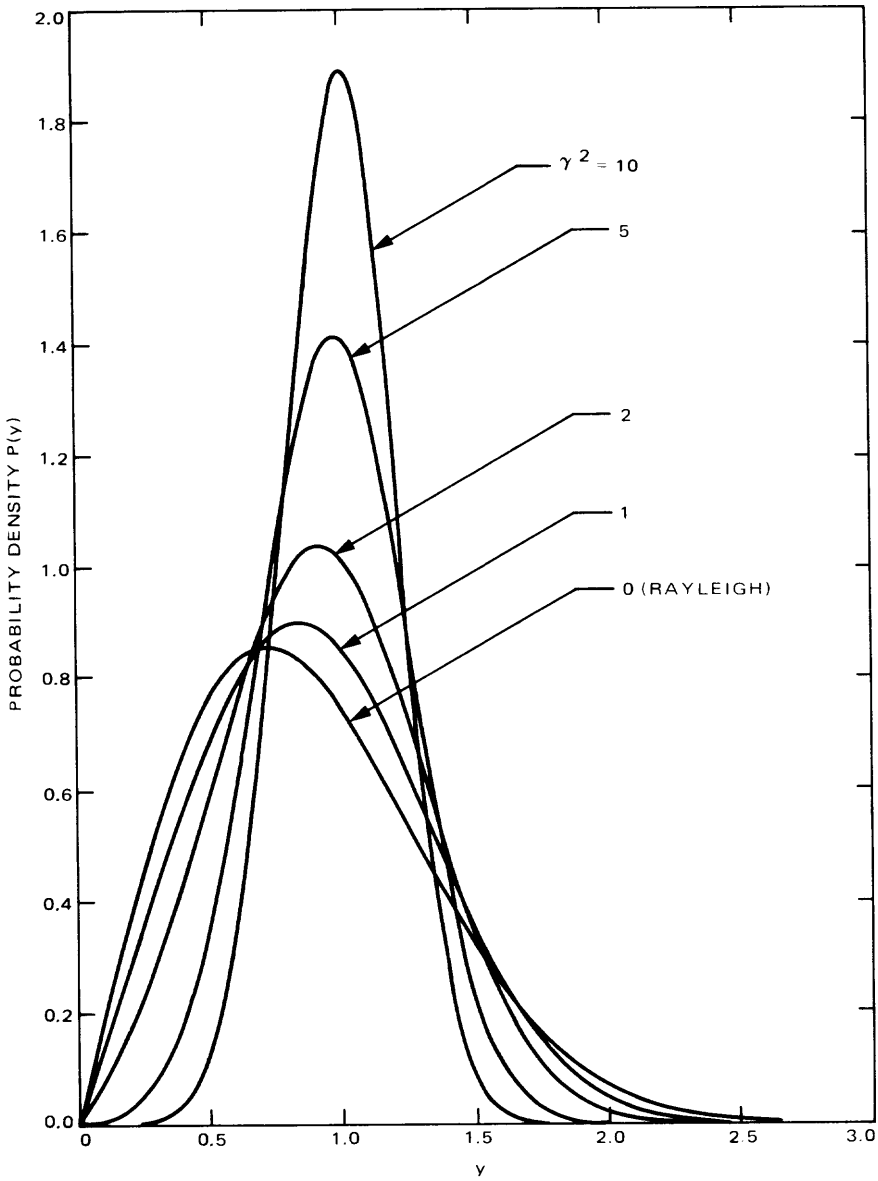


Fig. 5-54. Rician ($\gamma^2 > 0$) and Rayleigh ($\gamma^2 = 0$) probability density functions

Using these averages in (5.4-37), (5.4-4), and (5.4-42) we can get the performance of a concatenated coded system in the presence of a slow fading channel. The corresponding performance curves are shown in Figs. 5-55 and 5-56.

5.5 Telemetry System Losses

When the carrier phase, the subcarrier phase, and the symbol and word timing references are all derived perfectly at the Deep Space Station (DSS), the bit error probability, P_b, is a well-defined function of E_b/N_0. It is convenient to denote that function as $f(\cdot)$.

$$P_b = f\left(\frac{E_b}{N_0}\right) \text{ no system losses} \tag{5.5-1}$$

where $f(\cdot)$ is the system's bit error probability as a function of bit SNR in the case of perfect carrier tracking. Of course, $f(\cdot)$ will depend on the particular modulation, coding, and decoding schemes used.

In general, the carrier phase, the subcarrier phase, and the symbol and word timing references are not derived perfectly at the DSS. Under these circumstances, it is necessary to define telemetry system loss η_S as

$$P_b = f\left(\frac{E_b}{N_0}\ \eta_S\right) \tag{5.5-2}$$

Or, equivalently,

$$\eta_S = \frac{f^{-1}(P_b)}{\dfrac{E_b}{N_0}} \tag{5.5-3}$$

In the analysis of telemetry system loss, the assumption is generally made that the total system loss is the product of four independent subsystem losses [5-48]. This assumption allows the analysis to proceed with four separate, tractable problems. Thus, with the subcarrier phase and the symbol and word timing references assumed to be perfectly derived at the DSS, the telemetry performance degradation resulting from imperfect carrier phase recovery determines radio loss or noisy reference loss η_{RL}. Subcarrier demodulation loss η_{SDL} and symbol (or bit) synchronization and detection loss η_{BSDL} are similarly defined. Waveform distortion loss η_{WDL} is a transmitter/channel

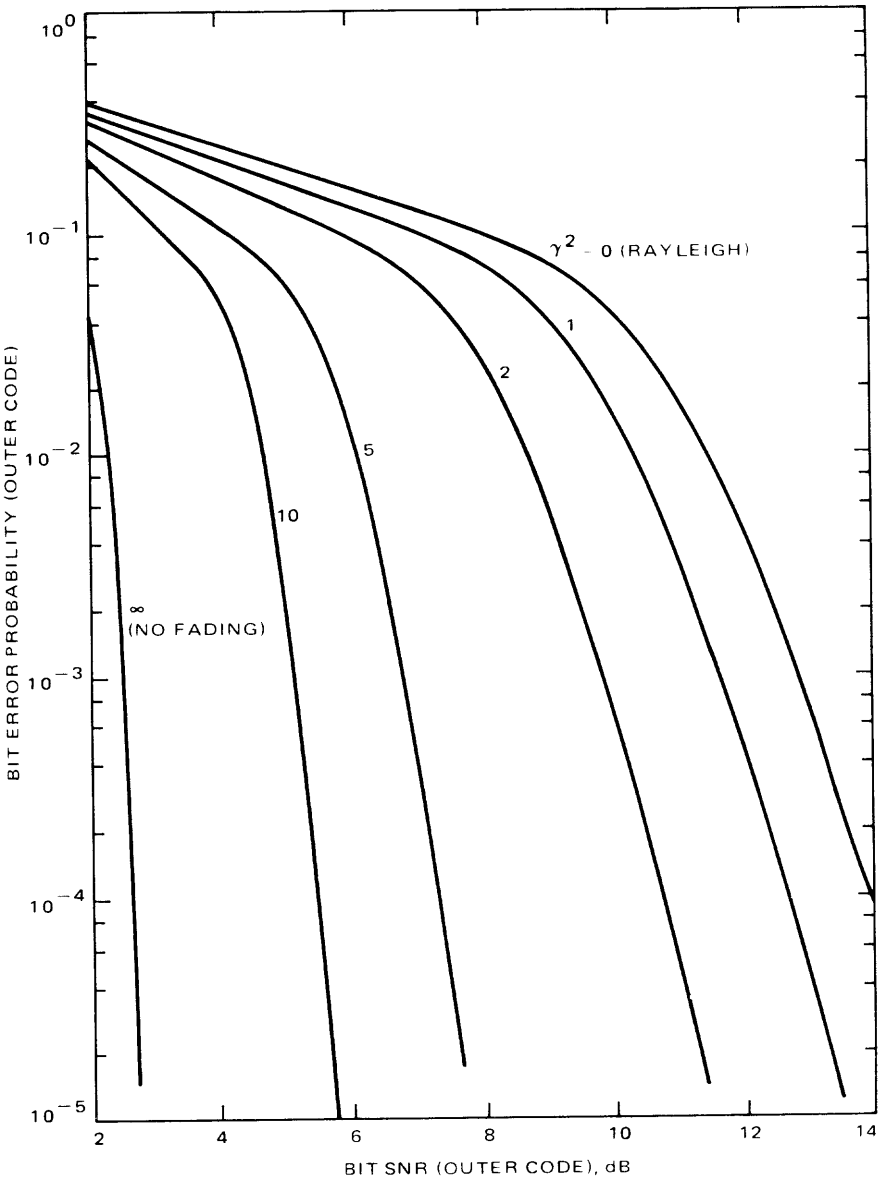


Fig. 5-55. P_b (RS) vs bit SNR for Rician ($\gamma^2 > 0$) and Rayleigh ($\gamma^2 = 0$) fading channels

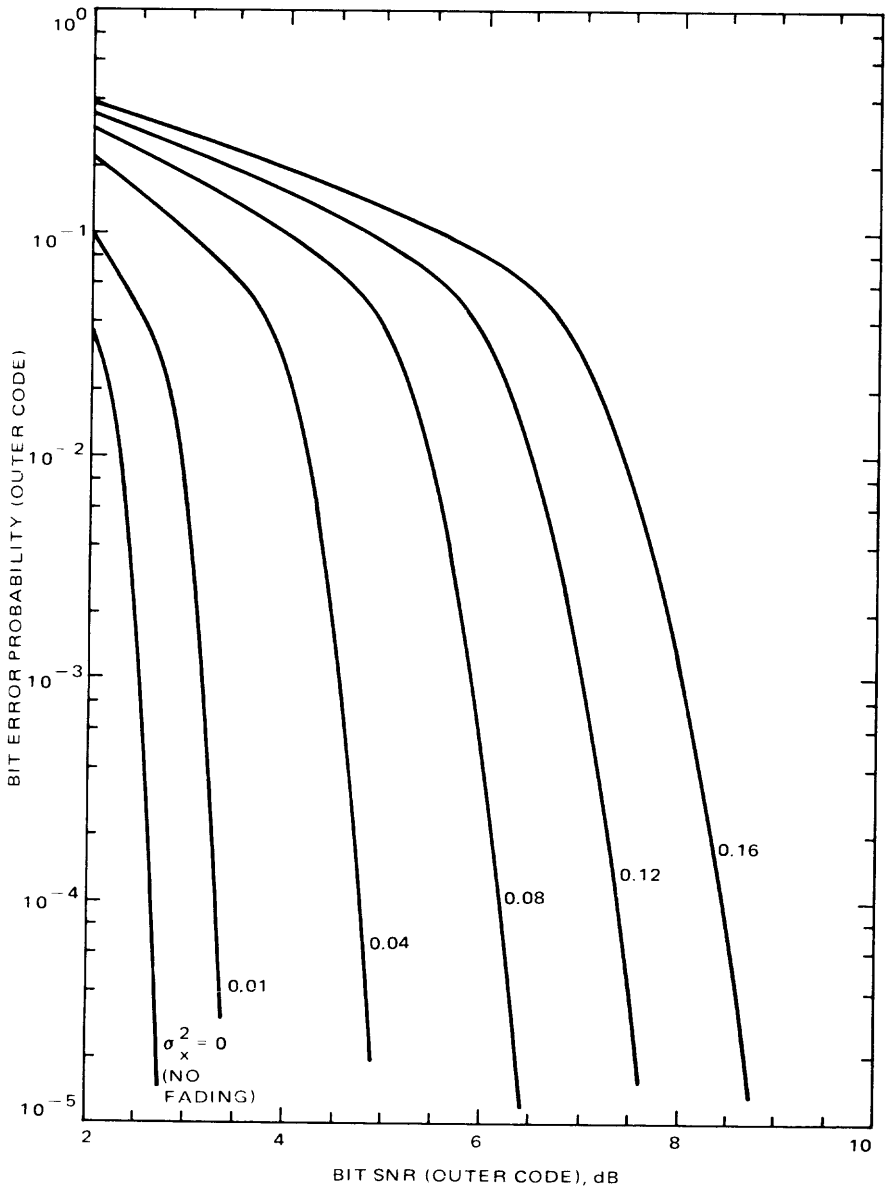


Fig. 5-56. P_b(RS) vs bit SNR for log-normal fading channel

effect whereby the subcarrier is not a symmetric square-wave. Telemetry system loss is conceptually defined as in (5.5-2) and (5.5-3), and these are useful definitions if the loss is to be determined experimentally. However, for analytical determinations of η_S, the operational definition shall be

$$\eta_S = \eta_{RL}\, \eta_{SDL}\, \eta_{BSDL}\, \eta_{WDL} \tag{5.5-4}$$

General mathematical models for η_{RL}, η_{SDL}, η_{BSDL}, and η_{WDL} are given in the subsections to follow. An approach discussed in section 5.6.6 in connection with suppressed carrier systems can also be applied here to obtain good approximate results for these losses. The carrier and subcarrier synchronization (sync) are the principal source of performance degradation in deep space telecommunications. The further deterioration of performance, due to the lack of perfect sync information such as symbol, word, and frame sync, is negligible relative to the carrier and subcarrier sync in most applications.

5.5.1 Radio Loss

When only the carrier phase is imperfectly derived at the DSS, the telemetry bit error probability $P_b(\phi)$ conditioned on the carrier phase tracking error ϕ is

$$P_b(\phi) = f\left\{\frac{E_b}{N_0}\left[\frac{1}{T}\int_0^T \cos\phi\, dt\ \right]^2\right\} \tag{5.5-5}$$

where T = detector integration time. $\phi(t)$ is a stochastic process whose probability density [5-10] (see also Chapter 3) is given in (5.3-18).

At this point, three cases must be distinguished: low data rate, high data rate, and medium data rate. For coded communication systems, in order to determine whether a high-, medium-, or low-data-rate model should be used, one should consider the phase error dynamics (how fast it is changing) and the time it takes to decode one bit. For our discussion here, the use of a data-rate-to-carrier tracking-loop bandwidth ratio is sufficient to illustrate these three cases.

5.5.1.1 Radio loss–low-data-rate case. If the data-rate-to-carrier tracking-loop bandwidth is low enough that $\phi(t)$ varies rapidly during time T, then the following low-rate approximation is appropriate.

$$\frac{1}{T}\int_0^T \cos\phi\, dt = \int_{-\pi}^{+\pi} \cos\phi\, p(\phi)\, d\phi \tag{5.5-6}$$

From which it follows that

$$\eta_{RL(LR)} = \left[\frac{I_1(\rho)}{I_0(\rho)}\right]^2 \tag{5.5-7}$$

5.5.1.2 Radio loss–high-data-rate case. If the data-rate-to-carrier tracking-loop bandwidth ratio is high enough so that any resulting carrier phase error, $\phi(t)$, changes slowly over the interval $(0, T)$, then $\phi(t)$ can be assumed constant over this interval. Hence,

$$\frac{1}{T}\int_0^T \cos\phi\, dt = \cos\phi \tag{5.5-8}$$

The high-rate radio loss then becomes

$$\eta_{RL(HR)} = \frac{f^{-1}\left\{\int_{-\pi}^{\pi} f\left[\frac{E_b}{N_0}\cos^2\phi\right] p(\phi)\, d\phi\right\}}{\frac{E_b}{N_0}} \tag{5.5-9}$$

5.5.1.3 Radio loss–medium-data-rate case. If neither the high-rate nor the low-rate assumptions are appropriate, then the following interpolation scheme is employed.

$$\eta_{RL(MR)} = \eta_{RL(HR)}\, a_c + \eta_{RL(LR)}\,(1 - a_c) \tag{5.5-10}$$

where [5-48]

$$a_c = \frac{\delta_c}{4}\left\{1 - \frac{\delta_c}{8}\left[1 - \exp\left(\frac{-8}{\delta_c}\right)\right]\right\} \tag{5.5-11}$$

and

$$\delta_c = \frac{1}{B_L T} \tag{5.5-12}$$

Examples of high-, medium-, and low-data-rate radio losses as a function of carrier loop SNR are plotted in Fig. 5-57 for the case of uncoded telemetry.

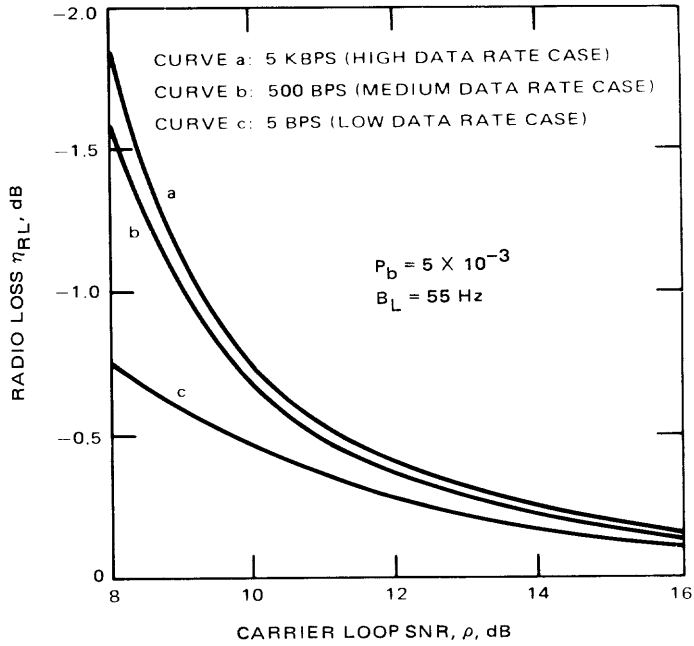


Fig. 5-57. One-way radio loss versus carrier loop SNR for uncoded telemetry

The radio loss as a function of carrier loop SNR for the convolutionally coded high-rate telemetry using a Viterbi decoder was plotted in Fig. 5-27.

5.5.2 Subcarrier Demodulation Loss

When only the subcarrier phase is imperfectly derived at the DSS the telemetry bit error probability $P_b(\phi_{sc})$ conditioned on the subcarrier phase tracking error ϕ_{sc} is

$$P_b(\phi_{sc}) = f\left\{\frac{E_b}{N_0}\left[\frac{1}{T}\int_0^T \left(1 - \frac{2}{\pi}|\phi_{sc}|\right) dt\right]^2\right\} \qquad (5.5\text{-}13)$$

where T = detector integration time.

$\phi_{sc}(t)$ is a stochastic process whose probability density is [5-5]

$$p(\phi_{sc}) = \frac{1}{\sqrt{2\pi\sigma_\phi^2}} \exp\left(-\frac{\phi_{sc}^2}{2\sigma_\phi^2}\right) \qquad (5.5\text{-}14)$$

where

$$\sigma_\phi^2 = \frac{\Gamma_{SL}\left(\frac{\pi}{2\alpha'}\right)^2}{\rho_{sc}} \tag{5.5-15}$$

$$\alpha' = 0.769\left[\frac{0.877 + 0.2\left(\frac{E_b}{N_0}\right)^{1.2}}{1.0 + 0.2\left(\frac{E_b}{N_0}\right)^{1.2}}\right]\text{erf}\left[\sqrt{\frac{2}{3}\left(\frac{E_b}{N_0}\right)}\right] \tag{5.5-16}$$

$$\Gamma_{SL} \cong \frac{1 + 0.345\,\rho' + 50(\rho')^5}{0.862 + 0.69\,\rho' + 50(\rho')^5} \tag{5.5-17}$$

$$\rho' = (\alpha')^2\,\rho_{sc}\frac{B_L}{B_{IF}} \tag{5.5-18}$$

ρ_{sc} = SNR in the subcarrier loop bandwidth

In the above, the suppression factor α' is for the assumed 50% symbol transition density. Again, three cases must be distinguished: low data rate, high data rate, and medium data rate.

5.5.2.1 Subcarrier demodulation loss–low-data-rate case. If the data-rate-to-subcarrier-tracking-loop bandwidth ratio is low enough that $\phi_{sc}(t)$ varies rapidly during time T, then the following low-rate approximation is appropriate.

$$\frac{1}{T}\int_0^T \left(1 - \frac{2}{\pi}|\phi_{sc}|\right) dt = \int_{-\infty}^{+\infty}\left(1 - \frac{2}{\pi}|\phi_{sc}|\right)p(\phi_{sc})\,d\phi_{sc} \tag{5.5-19}$$

From which it follows that

$$\eta_{SDL(LR)} = \left(1 - \frac{8\sigma_\phi^2}{\pi^3}\right)^2 \tag{5.5-20}$$

5.5.2.2 Subcarrier demodulation loss–high-data-rate case. If the data-rate-to-subcarrier-tracking-loop bandwidth ratio is high enough that any resulting subcarrier phase error is constant over the time interval $(0, T)$, then

$$\frac{1}{T}\int_0^T \left(1 - \frac{2}{\pi}|\phi_{sc}|\right) dt = 1 - \frac{2}{\pi}|\phi_{sc}| \tag{5.5-21}$$

The high-rate subcarrier demodulation loss then becomes

$$\eta_{SDL(HR)} = \frac{f^{-1}\left\{\int_{-\infty}^{+\infty} f\left[\frac{E_b}{N_0}\left(1 - \frac{2}{\pi}|\phi_{sc}|\right)^2\right] p(\phi_{sc})\, d\phi_{sc}\right\}}{\frac{E_b}{N_0}} \tag{5.5-22}$$

5.5.2.3 Subcarrier demodulation loss–medium-data-rate case. If neither the high-rate nor the low-rate assumptions are appropriate, then the following interpolation scheme is employed.

$$\eta_{SDL(MR)} = \eta_{SDL(HR)}\, a_{sc} + \eta_{SDL(LR)}\,(1 - a_{sc}) \tag{5.5-23}$$

where

$$a_{sc} = \frac{0.09135\,\delta_{sc} + \delta_{sc}^2}{1 + 3.3718\,\delta_{sc} + \delta_{sc}^2} \tag{5.5-24}$$

and

$$\delta_{sc} = \frac{1}{B_{Lsc}T} \tag{5.5-25}$$

B_{Lsc} is the SDA loop bandwidth (one-sided). Subcarrier demodulation loss for uncoded telemetry is plotted in Fig. 5-58 for several data rates.

5.5.3 Symbol or Bit Synchronization and Detection Loss

The symbol (for coded system) or bit (for uncoded) synchronization error is defined in Fig. 5-59. Symbol synchronization and detection loss is the result

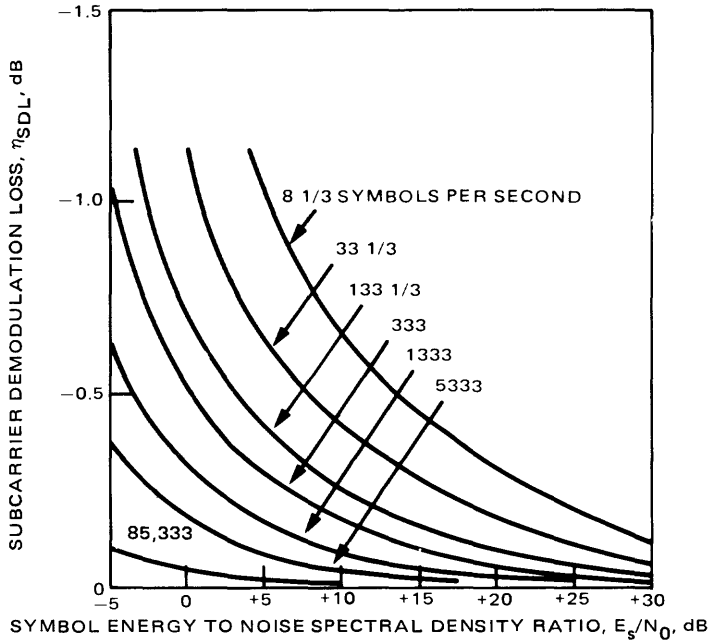


Fig. 5-58. Subcarrier demodulation loss versus E_s/N_0 for the Block III SDA in the medium noise bandwidth mode (for uncoded telemetry, $E_s = E_b$)

of the SSA operating imperfectly. The problem of symbol sync deals with the estimation of the time of arrival of the received data symbols. Symbol sync is important, for it is necessary to maintain a high efficiency in the data detection process; inaccurate symbol sync directly reduces the probability of making correct decisions. The symbol sync and detection loss is

$$\eta_{BSDL} = \frac{f^{-1}\left\{\int_{-T}^{+T} f\left[\frac{E_b}{N_0}\left(1-\left|\frac{\tau}{T}\right|\right)^2\right] q(\tau)\, d\tau\right\}}{\frac{E_b}{N_0}} \qquad (5.5\text{-}26)$$

Assuming a 50% transition density, the function $q(\tau)$ in (5.5-26) can be represented as

$$q(\tau) = \frac{1}{2}\,\delta(\tau) + \frac{1}{2}\,p(\tau); \quad -T \leqslant \tau \leqslant +T \qquad (5.5\text{-}27)$$

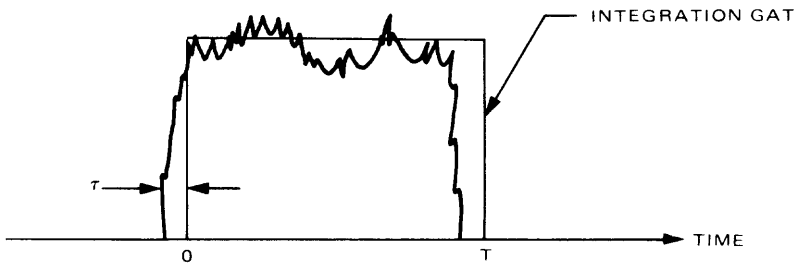


Fig. 5-59. Bit synchronization and detection loss

where $\delta(\cdot)$ is the Dirac-delta function and $p(\tau)$ is the pdf of sync error and approximately Gaussian; [5-49] gives standard deviations for $p(\tau)$ for several bit synchronizers.

The most common form of data-derived symbol sync is where the receiver includes a device that extracts the sync directly from the information-bearing signal. The principal advantage of extracting sync from the data-bearing signal is that no additional power expenditure and frequency spectrum are required. Thus, all the transmitter sideband power is available to the data and sync signals simultaneously. Hence symbol sync and detection loss is generally very small. In deep space telecommunications systems, it is usually less than 0.1 dB.

5.5.4 Waveform Distortion Loss

The predominant source of subcarrier waveform distortion loss is asymmetry, illustrated in Fig. 5-60. The resulting loss is given by

$$\eta_{WDL} = \left(1 - \frac{2\Delta T_{sc}}{T_{sc}}\right)^2 \tag{5.5-28}$$

where T_{sc} is the subcarrier period and ΔT_{sc} is as defined in Fig. 5-60.

5.6 Noisy Reference Performance for Suppressed Carrier Receivers

In Section 3.2 of Chapter 3, we considered the phase tracking performance of various suppressed carrier synchronization loops corresponding to BPSK, QPSK, and UQPSK modulations. The measure of performance used in each case was the mean-square phase jitter caused by the presence of thermal noise at the receiver front end. The purpose here is to tie together these mean-square phase jitter results with their degrading effect on system error probabil-

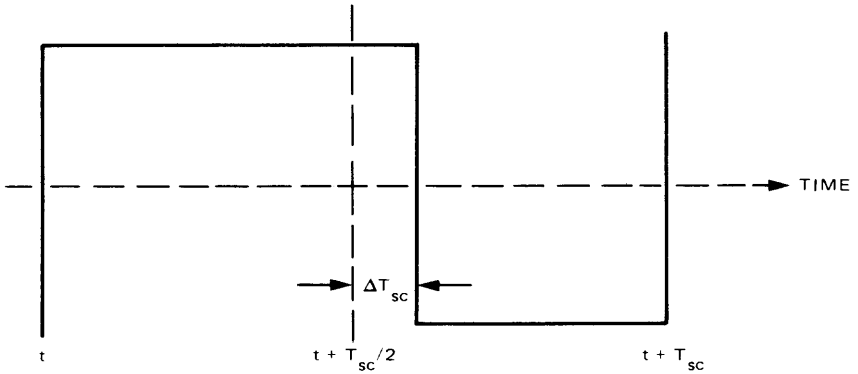


Fig. 5-60. Waveform distortion loss arising from subcarrier asymmetry (T_{sc} = subcarrier period)

ity performance. For small values of mean-square phase jitter, simple-to-use closed form formulas will be presented that allow quick evaluation of the noisy reference loss (or radio loss) corresponding to each of the above modulation techniques. Numerical evaluations corresponding to both uncoded and convolutionally coded systems will be given as illustrative examples of the application of the theory.

5.6.1 Bit Error Probability Performance of BPSK in the Presence of Carrier Sync Error (Phase Jitter)

In the absence of carrier sync error, the average bit error probability P_b is well-known [5-3] to be given by

$$P_b = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{5.6-1}$$

where $E_b/N_0 \triangleq CT_b/N_0$ is the bit energy-to-noise ratio. Assuming that the ratio of data-rate-to-carrier sync-loop bandwidth is large (the usual high-rate assumption), then in the presence of carrier sync error we have

$$P_b = \int_{-\pi/2}^{\pi/2} P_b(\phi)\, p(\phi)\, d\phi = \int_{-\pi/2}^{\pi/2} \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{E_b}{N_0}}\cos\phi\right) p(\phi)\, d\phi \tag{5.6-2}$$

where

$$p(\phi) = \begin{cases} \dfrac{2 \exp\left[(2\sigma_\phi)^{-2} \cos 2\phi\right]}{2\pi I_0\left[(2\sigma_\phi)^{-2}\right]}; & |\phi| \leqslant \dfrac{\pi}{2} \\ 0 \qquad\qquad\qquad\qquad ; & \text{otherwise} \end{cases} \tag{5.6-3}$$

is the Tikhonov probability density function for the phase error process ϕ. Note that since the synchronization loop indeed tracks the 2ϕ process defined on $(-\pi, \pi)$, the standard Tikhonov density has been appropriately scaled so as to apply to the ϕ process. Also assumed is that the 180° phase ambiguity caused by this division of the 2ϕ process by a factor of two has been perfectly resolved either by a "magic genie" or more appropriately by differential encoding of the input data.[5] Finally, in (5.6-3) we have substituted for the equivalent loop signal-to-noise ratio ρ_l (which normally appears in the Tikhonov density) an evaluation of that ratio based on *linear* theory, namely,

$$\rho_l = (2\sigma_\phi)^{-2} = \frac{1}{4\sigma_\phi^2} \tag{5.6-4}$$

where σ_ϕ^2 is obtained from (3.4-26b), i.e., $\sigma_\phi^2 = (\rho S_L)^{-1}$.

It is important to note, at this point, that σ_ϕ^2 is a function of E_b/N_0 through (3.4-28) and (3.4-20) (here $T_s = T_b$ for uncoded BPSK). Thus, a plot of P_b versus E_b/N_0 for *constant* σ_ϕ (as is commonly done for discrete carrier loops) is unjustified in the suppressed carrier case. More will be said about this later on.

For small σ_ϕ, we can make the following simplifications:

$$p(\phi) \cong \frac{1}{2\pi\sigma_\phi^2} \exp\left(-\frac{\phi^2}{2\sigma_\phi^2}\right)$$

$$P_b \cong \underbrace{\frac{1}{2}\operatorname{erfc}\sqrt{\frac{E_b}{N_0}}}_{\substack{\text{Ideal performance} \\ \text{(perfect sync)}}} + \underbrace{\frac{1}{2}\sqrt{\frac{E_b}{\pi N_0}} \exp\left(-\frac{E_b}{N_0}\right)\sigma_\phi^2}_{\substack{\text{Degradation due to} \\ \text{imperfect carrier sync} \\ \text{(phase jitter)}}} \tag{5.6-5}$$

[5] The result in (5.6-2) must be modified when differential encoding of the BPSK modulation is employed, as will be seen shortly.

The result for P_b in (5.6-5) was obtained by expanding $P_b(\phi)$ of (5.6-2) in a Maclaurin (power) series in ϕ, keeping only terms of second order or lower and using (5.6-5) rather than (5.6-3) to perform the required averaging over ϕ. Equation (5.6-5) is in a highly desirable form from the standpoint of computation in that it avoids the necessity of numerical integration and, in addition, clearly isolates the degradation due to the imperfect carrier synchronization reference.

An important point to note is that even if σ_ϕ were to be held constant as E_b/N_0 varies, evaluation of (5.6-2) or (5.6-5) does *not* result in an irreducible bit error probability as occurs in discrete carrier receivers. The principal reason for this is that here ϕ is restricted to lie in the region $(-\pi/2, \pi/2)$ insofar as the evaluation of (5.6-2) is concerned. Thus cos ϕ is always positive, and hence, for any ϕ in this interval, erfc $(\sqrt{E_b/N_0}\cos\phi)$ will approach zero as E_b/N_0 approaches infinity. Certainly since, as previously mentioned, σ_ϕ is not constant as E_b/N_0, the plot of P_b versus E_b/N_0 will be even steeper yet.

If differential encoding is employed to resolve the 180° phase ambiguity associated with the carrier synchronization loop (Costas loop), then $P_b(\phi)$ of (5.6-2) must be replaced by

$$P_b(\phi) = \text{erfc}\left[\sqrt{\frac{E_b}{N_0}}\cos\phi\right]\left\{1 - \frac{1}{2}\,\text{erfc}\left[\sqrt{\frac{E_b}{N_0}}\cos\phi\right]\right\} \quad (5.6\text{-}6)$$

For small ϕ, one obtains a result analogous to (5.6-5) except that the degradation term due to the imperfect carrier sync reference is now twice as large and the perfect sync term is obtained by setting $\phi = 0$ in (5.6-6).

Figure 5-61 illustrates P_b as computed from (5.6-2) vs E_b/N_0 with $\delta_b = 1/B_L T_b$ as a parameter (note that from (3.3-27), we can write $p = (E_b/N_0)\delta_b$). The Costas loop which provides the noisy synchronization reference is assumed to have integrate-and-dump arm filters so that the squaring loss is related to $R_d = E_b/N_0$ by the simple formula in (3.3-36) Superimposed on these curves in dashed lines are the results corresponding to evaluation of P_b from (5.6-5). Figure 5-62 illustrates the same results assuming differential encoding is employed for resolution of the 180° phase ambiguity associated with the Costas loop. In both of these cases, we note the presence of extremely small noisy reference losses even for a value of δ as low as 10. Typical systems have values of δ on the order of hundreds or thousands, in which case the noisy reference loss becomes vanishingly small. It is to be emphasized that the reason this occurs is that, contrary to discrete carrier systems, ρ now varies directly with E_b/N_0 for fixed B_L and T_b.

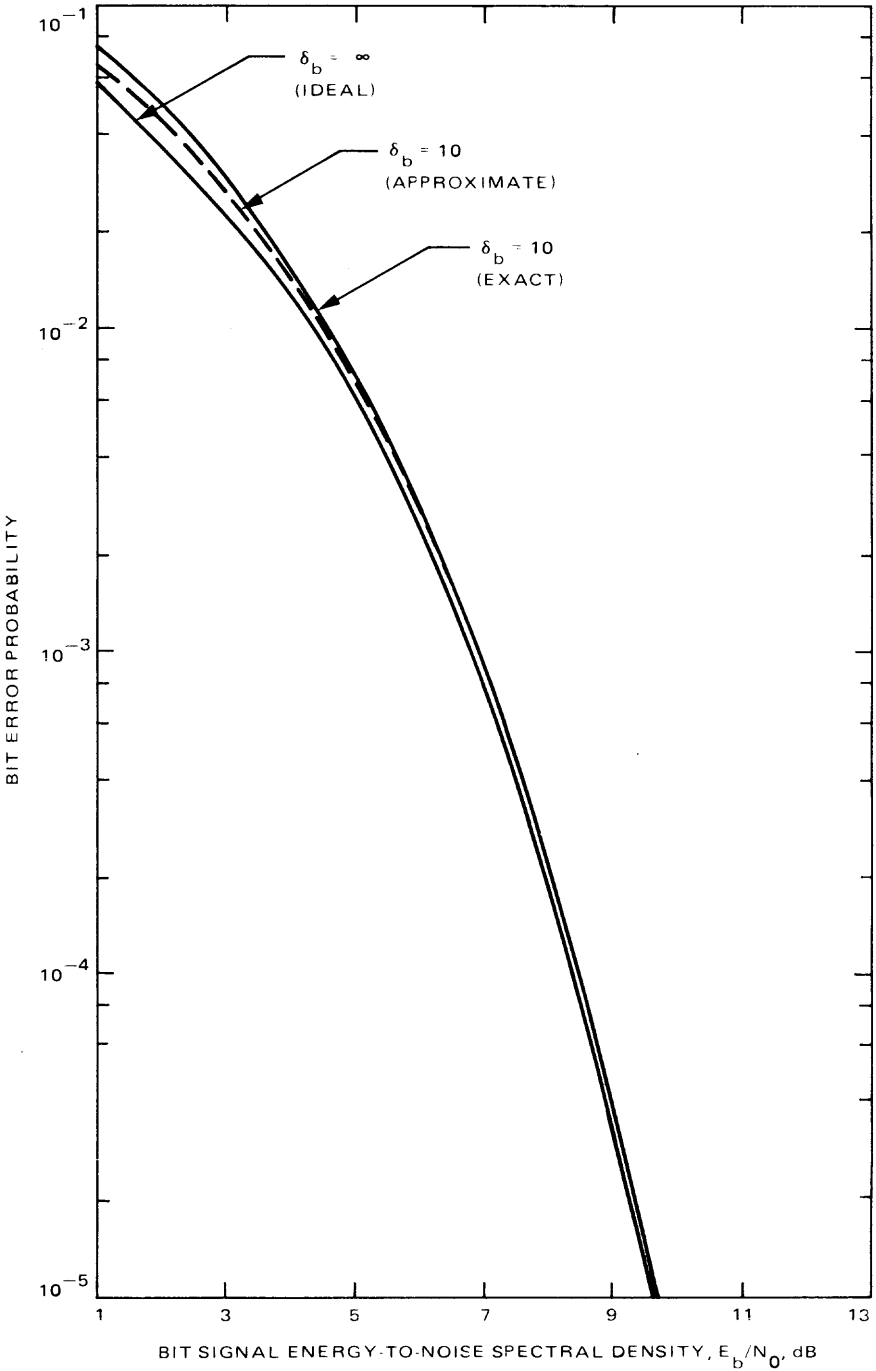


Fig. 5-61. Noisy reference loss performance of BPSK with Costas-loop tracking; integrate-and-dump arm filters; perfect ambiguity resolution

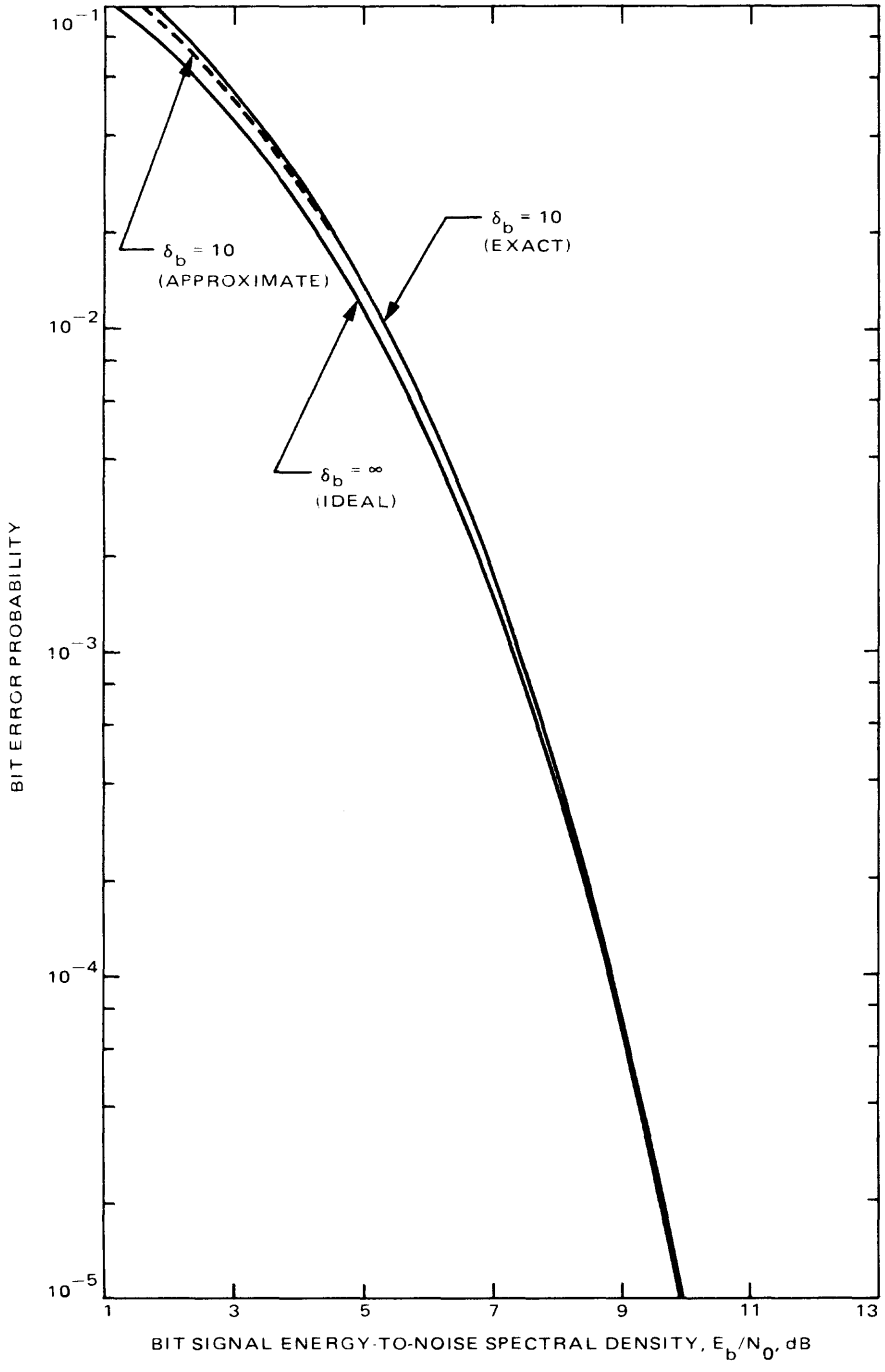


Fig. 5-62. Noisy reference loss performance of BPSK with Costas-loop tracking; integrate-and-dump arm filters; differentially encoded data for ambiguity resolution

5.6.2 Bit Error Probability Performance of QPSK in the Presence of Carrier Sync Error (Phase Jitter)

In the absence of carrier sync error, the average bit error probability performance of QPSK is *identical* to that of BPSK, namely, (5.6-1). In the presence of carrier sync error and the high-data-rate assumption made previously, we have

$$P_b = \int_{-\pi/4}^{\pi/4} P_b(\phi)\, p(\phi)\, d\phi$$

$$P_b(\phi) = \frac{1}{4}\,\text{erfc}\left[\sqrt{\frac{E_b}{N_0}}\,(\cos\phi - \sin\phi)\right] + \frac{1}{4}\,\text{erfc}\left[\sqrt{\frac{E_b}{N_0}}\,(\cos\phi + \sin\phi)\right]$$

$$p(\phi) = \begin{cases} \dfrac{4\exp\left[(4\sigma_\phi)^{-2}\cos 4\phi\right]}{2\pi I_0\left[(4\sigma_\phi)^{-2}\right]}; & |\phi| \leqslant \dfrac{\pi}{4} \\ 0; & \text{otherwise} \end{cases} \tag{5.6-7}$$

where σ_ϕ is now determined from (3.4-72). For small σ_ϕ, we again get the simplified results:

$$p(\phi) \cong \frac{1}{\sqrt{2\pi\sigma_\phi^2}}\exp\left(-\frac{\phi^2}{2\sigma_\phi^2}\right)$$

$$P_b \cong \underbrace{\frac{1}{2}\,\text{erfc}\sqrt{R_b}}_{\substack{\text{Ideal performance} \\ \text{(perfect sync)}}} + \underbrace{\frac{1}{2}\sqrt{\frac{R_b}{\pi}}\,(1 + 2R_b)\exp(-R_b)\,\sigma_\phi^2}_{\substack{\text{Degradation due to} \\ \text{imperfect carrier sync} \\ \text{(phase jitter)}}} \tag{5.6-8}$$

where

$$R_b \triangleq \frac{E_b}{N_0} = \text{Bit energy-to-noise ratio} \tag{5.6-9}$$

Figure 5-63 illustrates P_b as computed from (5.6-7) vs E_b/N_0 with $\delta_b = 1/B_L T_b$ as a parameter. The quadriphase Costas loop which provides the noisy synchronization reference is again assumed to have integrate-and-dump arm filters so that the squaring loss is related to $R_d = 2E_b/N_0$ by the simple formula in (3.3-58). Also superimposed in dashed lines is the approximate result from (5.6-8). Although for a fixed value of δ the noisy reference loss for QPSK is larger than that for BPSK, once again one observes that practical values of δ on the order of hundreds or thousands will produce negligible degradation. Furthermore, since for $|\phi| < \pi/4$ we have $\cos\phi - \sin\phi > 0$, then the argument of the complementary error function in the first term of (5.6-7) will never go negative; hence, an irreducible error probability cannot occur.

5.6.3 Bit Error Probability Performance of UQPSK in the Presence of Carrier Sync Error (Phase Jitter)

In the absence of carrier sync error, the per channel average bit error probability performance is given by

$$P_{bi} = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{P_i T_{si}}{N_0}}\right);\quad i = 1, 2 \tag{5.6-10}$$

where

$$\begin{aligned} P_i &= \text{average signal power in channel } i;\ i = 1, 2 \\ T_{si} &= \text{baud interval on channel } i;\ i = 1, 2 \\ P = P_1 + P_2 &= \text{total transmitted power} \end{aligned} \tag{5.6-11}$$

In the presence of carrier sync error, we now get

$$P_{bi} = \int_{-\pi/2}^{\pi/2} P_{bi}(\phi)\, p(\phi)\, d\phi;\quad i = 1, 2$$

$$\begin{aligned} P_{bi}(\phi) = {} & \frac{1}{4}\,\mathrm{erfc}\left[\sqrt{\frac{P_i T_{si}}{N_0}}\cos\phi + \sqrt{\frac{P_i T_{si}}{N_0}}\, d_{\bar{i}\bar{i}}\sin\phi\right] \\ & + \frac{1}{4}\,\mathrm{erfc}\left[\sqrt{\frac{P_i T_{si}}{N_0}}\cos\phi - \sqrt{\frac{P_i T_{si}}{N_0}}\, d_{\bar{i}\bar{i}}\sin\phi\right];\quad i = 1, 2 \end{aligned} \tag{5.6-12}$$

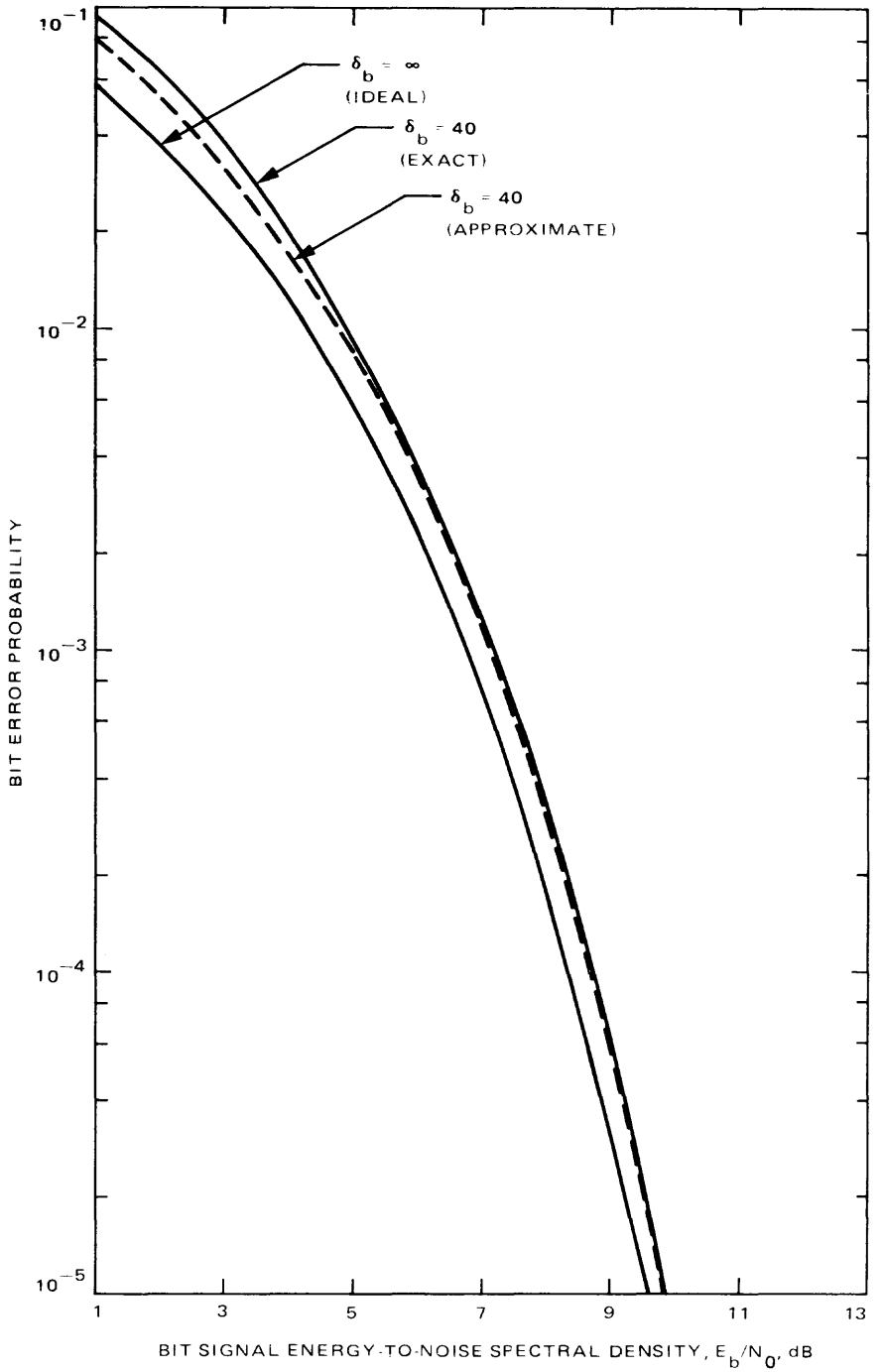


Fig. 5-63. Noisy reference loss performance of QPSK with Costas-loop tracking; integrate-and-dump arm filters; perfect ambiguity resolution

where the overbar denotes average over $d_{\bar{i}i}$, and $\bar{i}$ is the "complement" of i, i.e.,

$$\bar{i} = \begin{cases} 1; & i = 2 \\ 2; & i = 1 \end{cases} \tag{5.6-13}$$

$d_{\bar{i}i}$ = crosstalk of channel $\bar{i}$ on channel i

$$= \frac{1}{T_{si}} \int_0^{T_{si}} \underbrace{m_{\bar{i}}(t)}_{\text{Data modulation on channel } \bar{i}} \underbrace{p_i(t)}_{\text{Pulse shape of modulation on channel } i} dt \tag{5.6-14}$$

Finally, consider the following two cases.

(1) The case when σ_ϕ is small.

In this case, the phase error probability is

$$p(\phi) = \frac{1}{2\pi\sigma_\phi^2} \exp\left(-\frac{\phi^2}{2\sigma_\phi^2}\right)$$

$$P_{bi} \cong \frac{1}{2} \operatorname{erfc} \sqrt{\frac{P_i T_{si}}{N_0}}$$

$$+ \frac{1}{2} \sqrt{\frac{P_i T_{si}}{\pi N_0}} \left[1 + 2\,\overline{d_{ii}^2}\left(\frac{P_i T_{si}}{N_0}\right)\right] \exp\left(-\frac{P_i T_{si}}{N_0}\right) \sigma_\phi^2; \quad i = 1, 2 \tag{5.6-15}$$

and

$$\overline{d_{\bar{i}i}^2} = \text{mean-square crosstalk of channel } \bar{i} \text{ on channel } i \tag{5.6-16}$$

Depending on the data formats on the two channels, $\overline{d_{12}^2}$ is evaluated in Tables 5-5 and 5-6. Furthermore,

$$\overline{d_{21}^2} = \frac{1}{n}\,\overline{d_{12}^2} \tag{5.6-17}$$

Table 5-5. Evaluation of $\overline{d_{12}^2(t)}$; synchronous symbol clocks; (n is integer)

$m_1(t)$	$m_2(t)$ Manchester	$m_2(t)$ NRZ
Manchester	0; n even $\frac{1}{n}$; n odd	1 ; n even $\frac{n-1}{n}$; n odd
NRZ	0	1

Table 5-6. Evaluation of $\overline{d_{12}^2(t)}$; asynchronous symbol clocks; (n need not be an integer)

$m_1(t)$	$m_2(t)$ Manchester	$m_2(t)$ NRZ
Manchester	$\frac{1}{6}[n^2 - \frac{5}{n} + 12 - 6n]$; $R_{s1} \leqslant R_{s2} \leqslant 2R_{s1}$ $\frac{1}{2n}$; $R_{s2} \geqslant 2R_{s1}$	$n - \frac{1}{6}n^2 - 1 + \frac{1}{3n}$; $R_{s1} \leqslant R_{s2} \leqslant 2R_{s1}$ $1 - \frac{1}{n}$; $R_{s2} \geqslant 2R_{s1}$
NRZ	$\frac{1}{6n}$	$1 - \frac{1}{3n}$

where

$$n \triangleq R_{s2}/R_{s1} \tag{5.6-18}$$

is the ratio of data rates on the two channels.

(2) The case when n is large.

For deep-space applications, the ratio of data rates of the high-rate channel to the low-rate channel is usually very large, $n \geqslant 20$. Then, for an uncoded UQPSK, the conditional bit error probability performance [5-50] of the low-data-rate channel (Channel 1) is

$$P_{b1}(\phi) \cong \frac{1}{2}\operatorname{erfc}\left\{\frac{\sqrt{R_{s1}}\cos\phi}{\sqrt{1+2R_{s2}}\sin^2\phi}\right\} \tag{5.6-19}$$

and the high-data-rate channel is

$$P_{b2}(\phi) \cong \frac{1}{2}\text{erfc}\,(\sqrt{R_{s2}}\cos\phi) \tag{5.6-20}$$

We get the unconditional bit error probability by averaging $P_{bi}(\phi)$,

$$P_{bi} = \int_{-\pi/2}^{\pi/2} P_{bi}(\phi)\,\frac{\exp(\alpha\cos 2\phi)}{\pi I_0(\alpha)}\,d\phi;\quad i = 1, 2 \tag{5.6-21}$$

where α is the Costas loop parameter related to the signal-to-noise ratio in the loop ρ and to the squaring loss S_L [5-51], with

$$\alpha = \frac{\rho S_L}{4} \tag{5.6-22}$$

$$\rho = \frac{P_1 + P_2}{N_0 B_L} = \frac{P}{N_0 B_L} \tag{5.6-23}$$

B_L is the single-sided Costas loop bandwidth.

Figures 5-64 and 5-65 illustrate the bit error probabilities vs bit SNRs of the low-data-rate channel and the high-data-rate channel respectively. Figures 5-66 and 5-67 plot the bit error probabilities of both channels as a function of n, and bit SNR of the higher-data-rate channel, respectively.

When n, given in (5.6-18), does not satisfy the condition $n \geqslant 20$, the above approximation does not hold. A performance analysis of a UQPSK system, which does not depend on the approximation $n \geqslant 20$, has been documented in [5-50].

5.6.4 Bit Error Probability Performance of Convolutionally Coded BPSK in the Presence of Carrier Sync Error

In this section, we compute the noisy reference performance of convolutionally coded BPSK using a Costas loop for carrier tracking. The procedure for accomplishing this is as follows. Suppose that $f(E_b/N_0)$ is the bit error probability for a convolutionally coded channel with a perfect carrier reference. The particular form of the function $f(\cdot)$ depends on the specific code in use. Then, again assuming that the bit rate for the channel is very much greater than the bandwidth of the receiver suppressed carrier tracking loop,

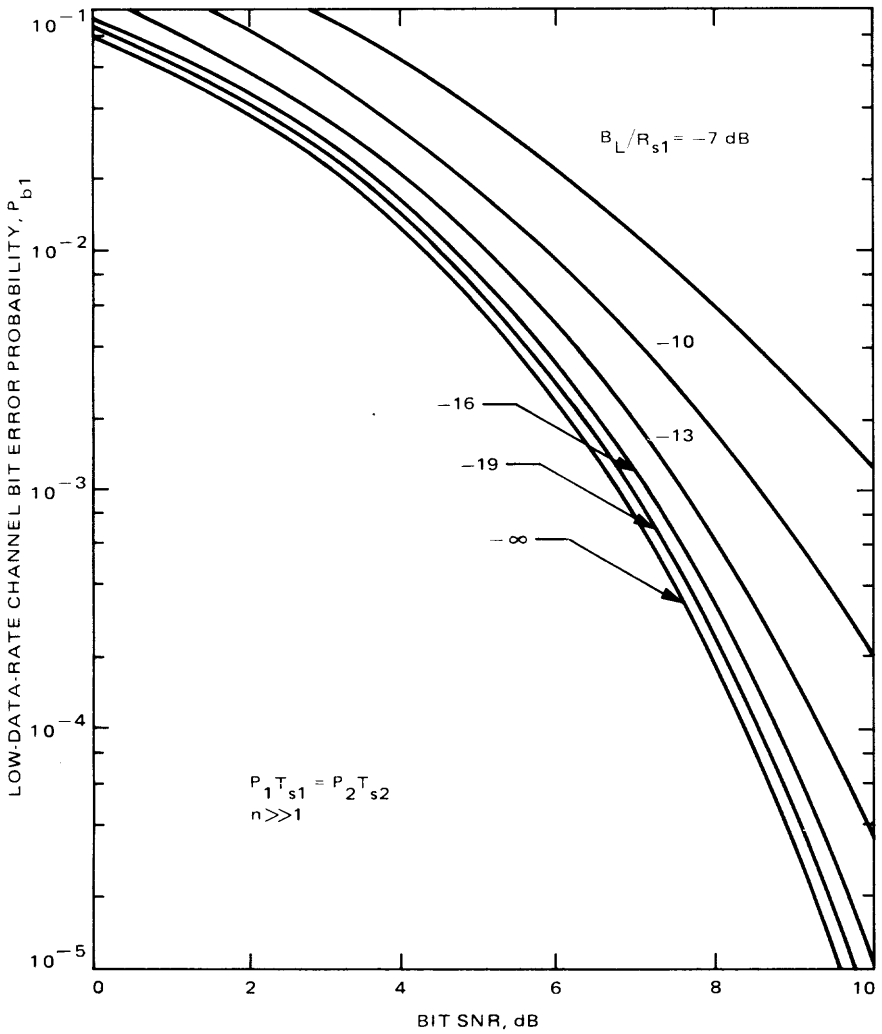


Fig. 5-64. Bit error probability of the low-data-rate channel vs bit SNR

the bit error probability for the channel with an imperfect (noisy) carrier reference is, analogously to (5.6-2),

$$P_b = \int_{-\pi/2}^{\pi/2} f\left(\frac{E_b}{N_0}\cos^2\phi\right) p(\phi)\, d\phi \qquad (5.6\text{-}24)$$

where $p(\phi)$ is again given by (5.6-3). Similar methods have been used by others to compute the bit error probability performance of convolutionally

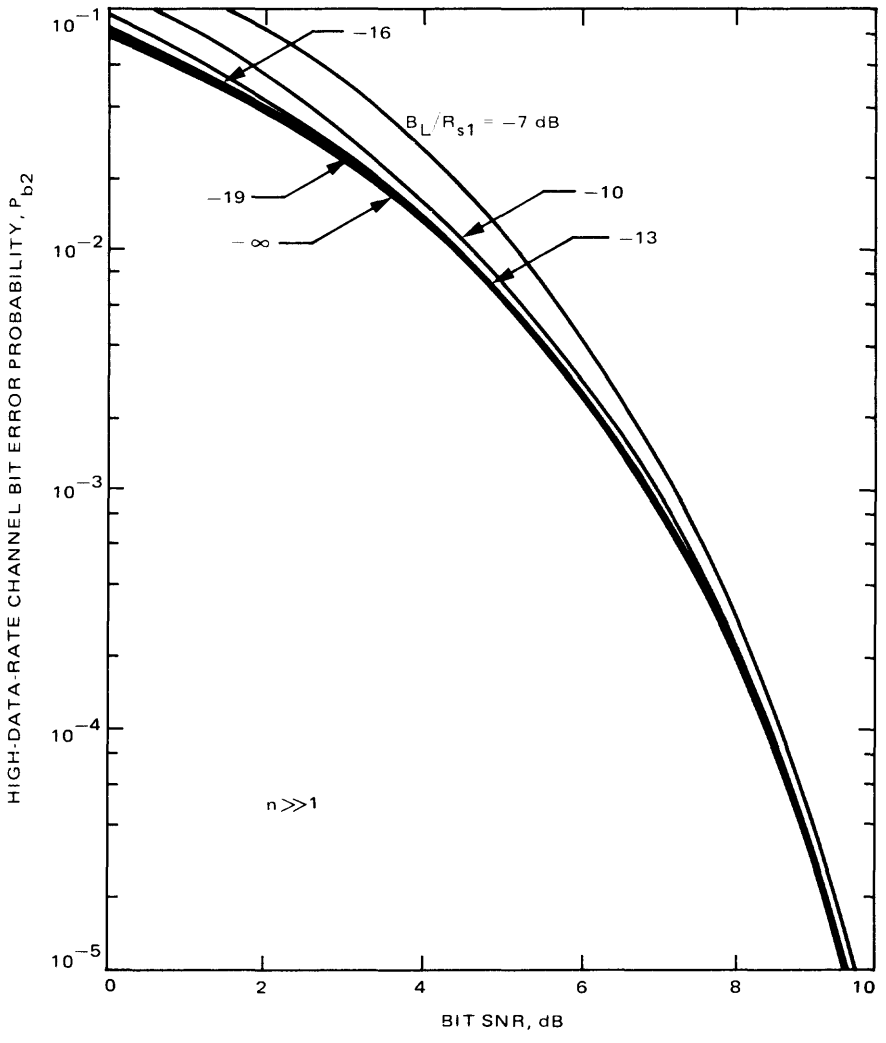


Fig. 5-65. Bit error probability of the high-data-rate channel vs bit SNR

coded channels with a *discrete* carrier tracking loop (i.e., a phase-locked loop) acting as the source of noisy reference. In particular, such performance of a rate 1/2, constraint length 7 convolutionally encoded/Viterbi-decoded channel during one-way operation is given in [5-52] and [5-53], and has been discussed in Section 5.4.3.1. For the perfect carrier, case $f(E_b/N_0)$ is given in (5.4-16).

Figure 5-68 illustrates P_b as computed from (5.6-24) vs E_b/N_0 in dB with $\delta_s = 1/B_L T_s$ as a parameter with T_s denoting the code symbol duration. Here,

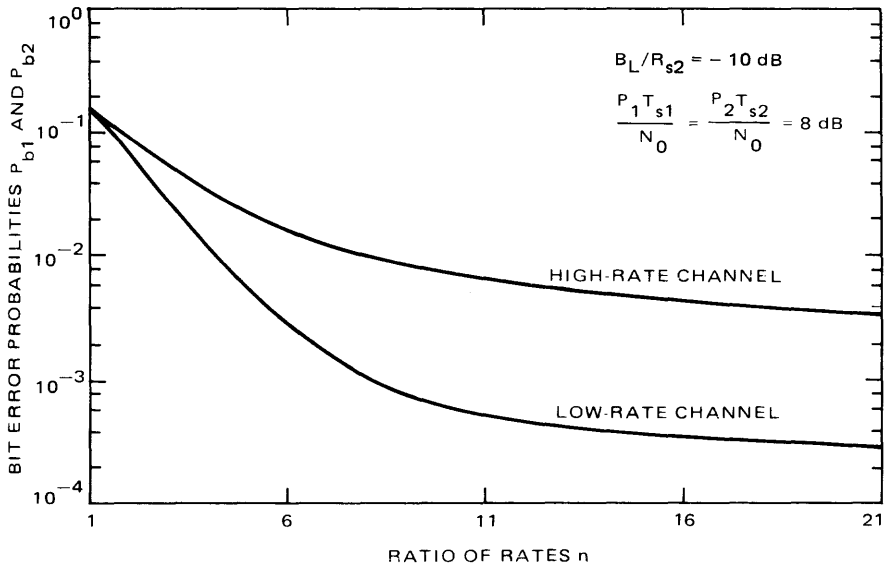


Fig. 5-66. Bit error probabilities of both the high- and low-data-rate channels vs the ratio of the data rates

since the convolutional code is rate 1/2, we have $T_s = T_b/2$, and likewise $R_d = E_b/2N_0$ is used in the squaring loss formula of (3.4-40). Once again we observe that reasonably large values of δ_s will result in very small noisy reference losses.

5.6.5 Bit Error Probability Performance of Convolutionally Coded UQPSK in the Presence of Carrier Sync Error

In this section we compute the performance of convolutionally coded UQPSK in the presence of a noisy carrier reference, using a Costas loop for carrier tracking. A block diagram of this coded UQPSK system is depicted in Fig. 5-69.

Let $f_i(E_b/N_0)$, $i = 1, 2$, be the bit error probability of a convolutionally coded/Viterbi-decoded channel with a perfect carrier reference. For deep-space communication, n (5.6-18), the ratio of the high-data-rate channel (Channel 2) to the low-data-rate channel (Channel 1) is usually large, say $n \geqslant 20$. And the data rates of both Channels 1 and 2 are much greater than the suppressed carrier tracking loop bandwidth. Thus, the bit error probabilities of the two channels [5-54] are

$$P_{bi} = \int_{-\pi/2}^{\pi/2} P_{bi}(\phi)\, p(\phi)\, d\phi \qquad (5.6\text{-}25)$$

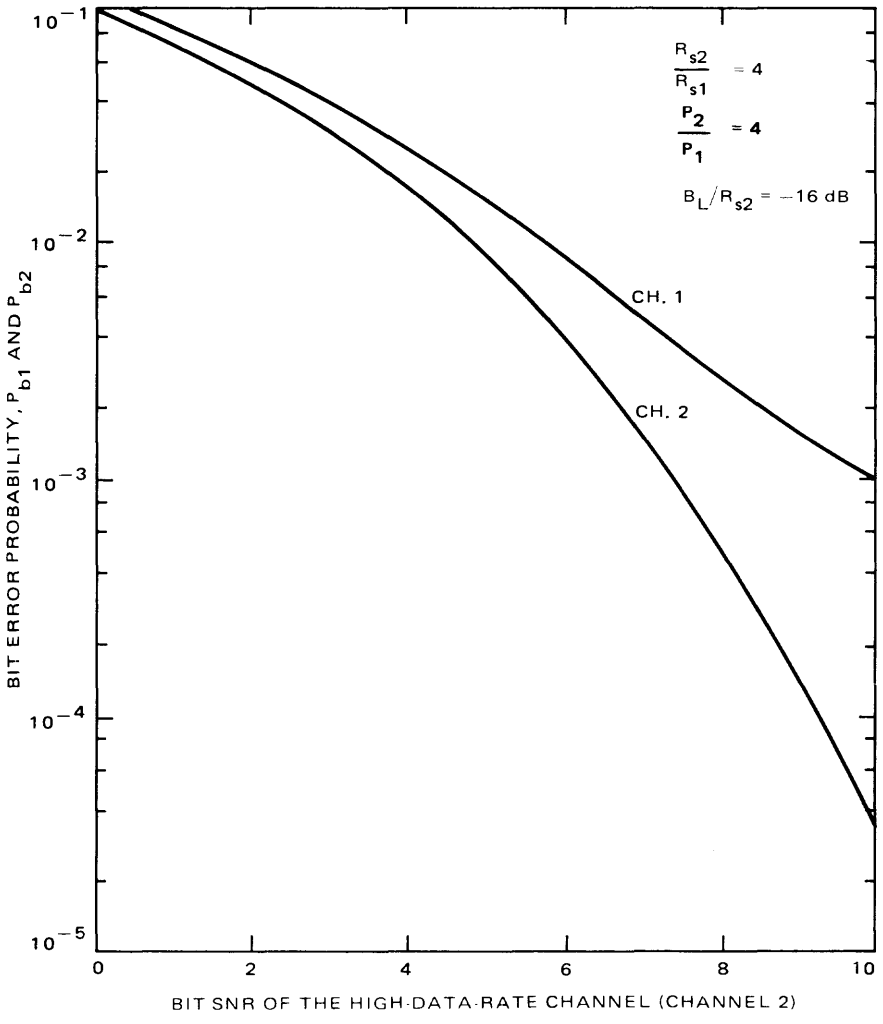


Fig. 5-67. Bit error probability of both the high- and low-data-rate channels vs the ratio of the high-data-rate-channel bit SNR

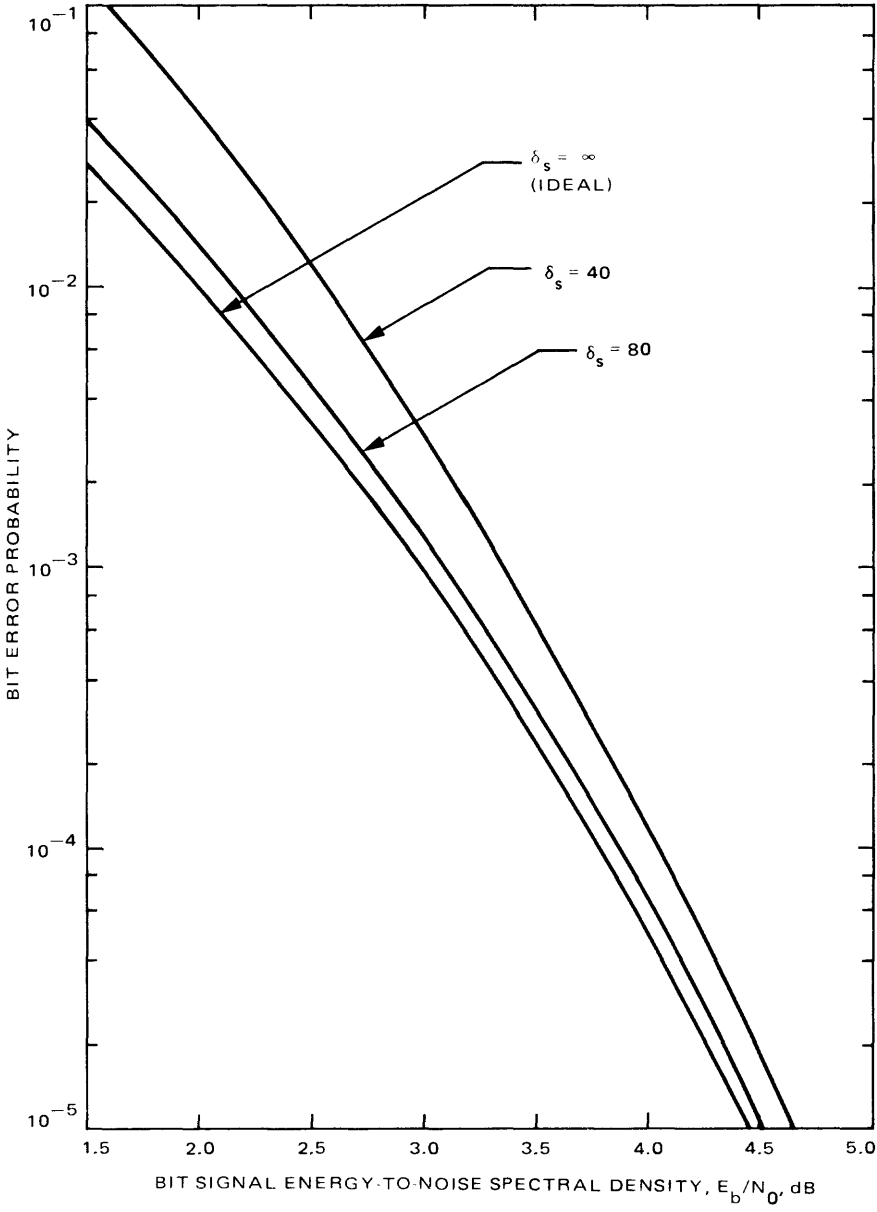


Fig. 5-68. Noisy reference loss performance of $K = 7$, rate 1/2 convolutionally encoded BPSK with Costas-loop tracking; integrate-and-dump arm filters

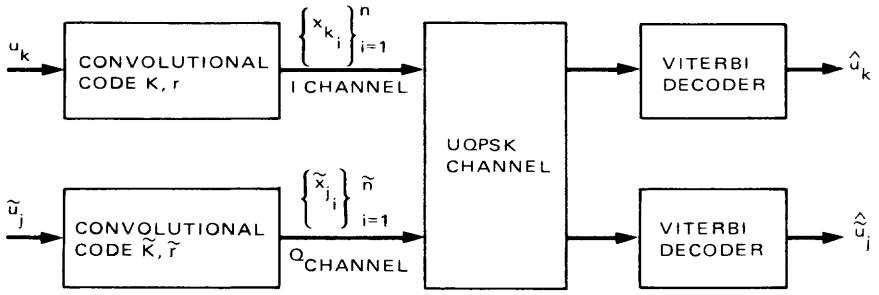


Fig. 5-69. Block diagram of the system

where the Costas loop phase error density $p(\phi)$ is given by (5.6-7). For the high-data-rate channel, we have

$$P_{b2}(\phi) \leqslant \frac{1}{2} f\left[\frac{E_{b2}}{N_0}\left(\cos\phi - \sqrt{\frac{E_{b1}\sin^2\phi}{nE_{b2}}}\right)^2\right] + \frac{1}{2} f\left[\frac{E_{b2}}{N_0}\left(\cos\phi + \sqrt{\frac{E_{b1}\sin^2\phi}{nE_{b2}}}\right)^2\right] \quad (5.6\text{-}26)$$

where n is defined in (5.6-18). For the low-data-rate channel, we have

$$P_{b1} = f\left[\left(\frac{E_{b1}}{N_0}\cos^2\phi\right) \Big/ \left(1 + \frac{E_{b2}}{N_0}\sin^2\phi\right)\right] \quad (5.6\text{-}27)$$

Note that for large n, (5.6-18), the bit error probabilities in (5.6-25), (5.6-26), and (5.6-27) are valid whether the data streams in the two channels are synchronized or not.

Using the results in [5-10], it can be shown that for moderate to high loop SNRs and small static phase error $\phi_{ss} < 20°$, the Costas loop phase error probability density function is well approximated by

$$p(\phi) = \begin{cases} \dfrac{1}{\pi I_0(\alpha)} \exp\left(-2\alpha\phi_{ss}^2 + 2\alpha\phi_{ss}\phi + \alpha\cos 2\phi\right); & -\dfrac{\pi}{2} \leqslant \phi \leqslant \dfrac{\pi}{2} \\ 0; & \text{otherwise} \end{cases} \quad (5.6\text{-}28)$$

where

$$\alpha = \left(\frac{P_1 + P_2}{N_0 B_L}\right)\left(\frac{S_L}{4}\right)$$

$$\phi_{ss} = \frac{\Omega_0}{K_T} \tag{5.6-29}$$

$\Omega_0 = 2\pi f_0$ is the doppler frequency offset in rad/sec, K_T is the total loop gain (including AGC and suppression factors), B_L is the one-sided loop bandwidth, and S_L is the Costas loop squaring loss (see Chapter 3). Here we are able to use the squaring loss derived for the uncoded case because we deal with uncorrelated convolutional codes [5-55] and the performance of the Costas loop depends only on the second order statistics of the signal stream. Hence, uncorrelated code symbols mean independent code symbols.

For a constraint length 7, code rate 1/2 convolutionally coded/Viterbi-decoded UQPSK channel, the function $f(\cdot)$ is given by (5.6-16). Using (5.6-26) through (5.6-29) in (5.6-25), we obtain the bit error probabilities of the high-rate and low-rate channels; these are illustrated in Figs. 5-70 and 5-71, respectively. Figure 5-72 plots the bit error probabilities of both channels vs the total received power-to-noise spectral density ratio for various values of doppler frequency offset.

5.6.6 Direct Evaluation of Noisy Reference Loss

Often one is interested in directly computing the additional bit energy-to-noise ratio required in the presence of a noisy synchronization reference to produce the same error probability as the perfectly synchronized system. This additional bit energy-to-noise ratio is referred to as the *noisy reference loss* for the specified error probability and as such represents a *horizontal* displacement of a given point on the ideal performance error probability vs the bit energy-to-noise ratio curve (see Fig. 5-73). In mathematical terms, letting η denote the dB equivalent of E_b/N_0, then at the particular bit error probability P_{b_0}, the noisy reference loss (in dB) denoted by L, would be given by

$$L = \eta_0 - \eta_{I_0} \tag{5.6-30}$$

Unfortunately, the approximate relations of (5.6-5), (5.6-8), and (5.6-15) give us the *additional error probability at a given bit energy-to-noise ratio* caused by the synchronization reference being imperfect; i.e., a *vertical* dis-

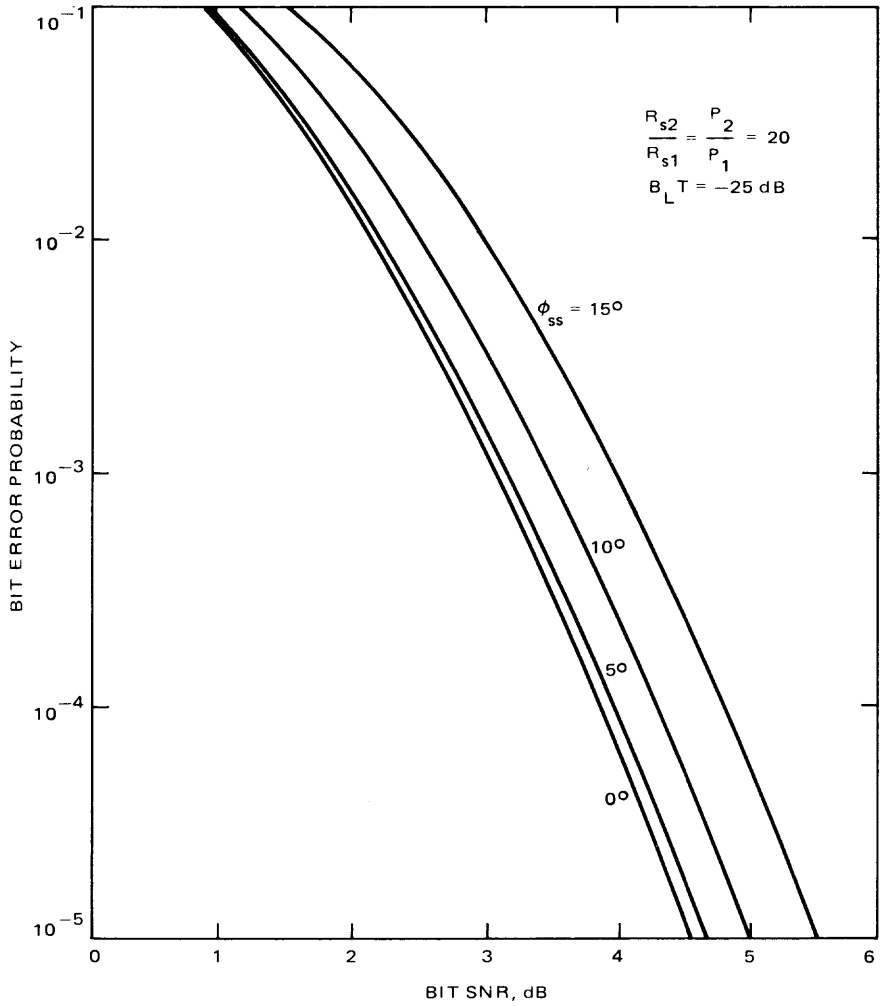


Fig. 5-70. Bit error probability of high-rate channel vs bit SNR with ϕ_{ss} as parameter

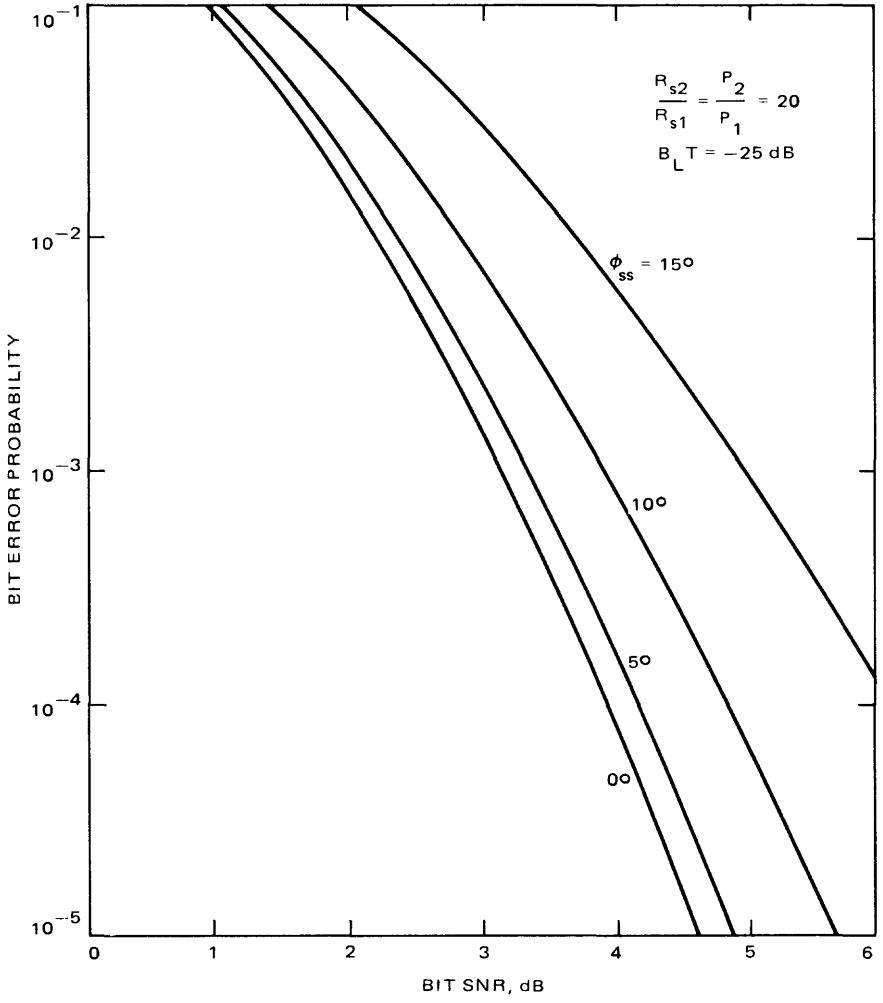


Fig. 5-71. Bit error probability of low-rate channel vs bit SNR with ϕ_{ss} as parameter

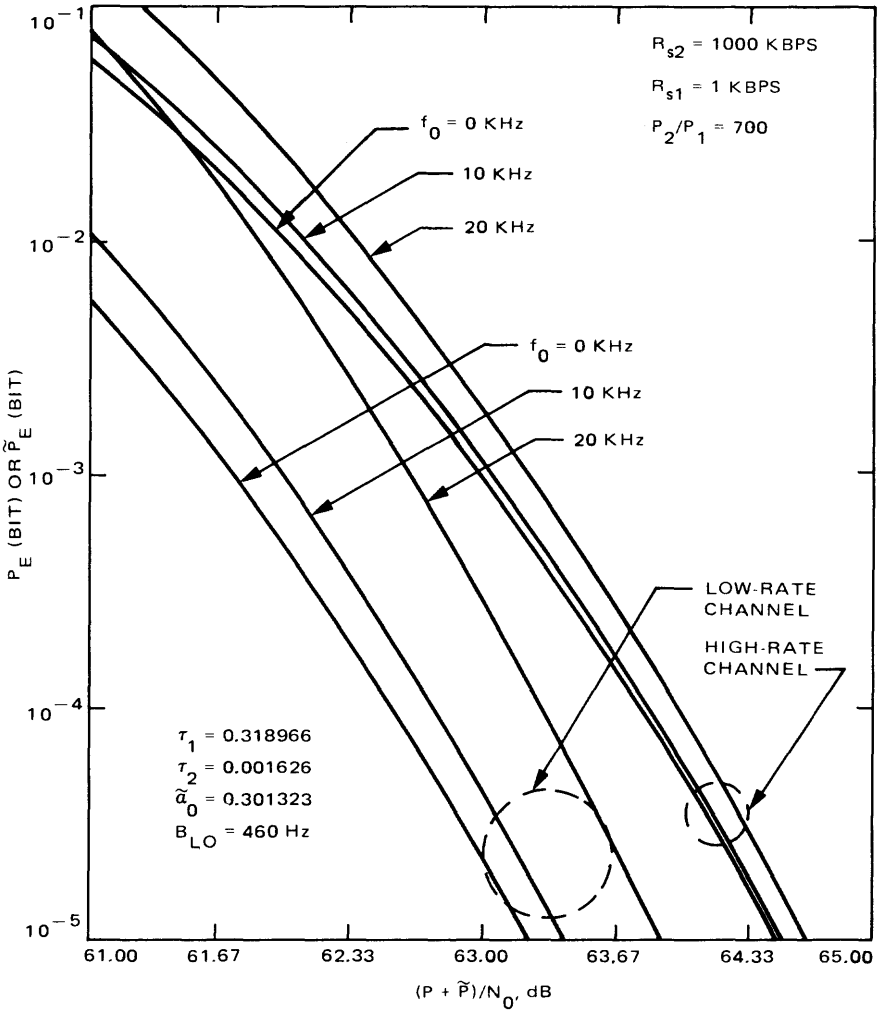


Fig. 5-72. High- and low-data-rate channel's bit error rates vs $(P_1 + P_2)/N_0$ with doppler frequency offset as parameter

placement of a given point on the ideal performance curve. Thus, we present here a simple technique that allows transformation of this vertical displacement into the horizontal displacement required for directly evaluating noisy reference loss.

When the noisy reference is small and likewise the system error probability, then a curve of $\log_{10} P_b$ vs η will be essentially parallel to the corresponding ideal performance curve, namely, $\log_{10} P_{b_I}$ vs η. This behavior can easily be observed by looking at Figs. 5-61, 5-62, and 5-63 as examples. Thus, in view

of the above observation, one can approximate the slope of the ideal performance curve as follows:

$$\frac{d \log_{10} P_{b_I}}{d\eta} = \text{slope of ideal performance curve at point of interest}$$

$$\cong \frac{\log_{10} P_b - \log_{10} P_{b_I}}{\eta - \eta_I} = \frac{\log_{10}\left(\frac{P_b}{P_{b_I}}\right)}{L} \tag{5.6-31}$$

or

$$L = \frac{\log_{10}\left(\frac{P_b}{P_{b_I}}\right)}{\frac{d \log_{10} P_{b_I}}{d\eta}} \tag{5.6-32}$$

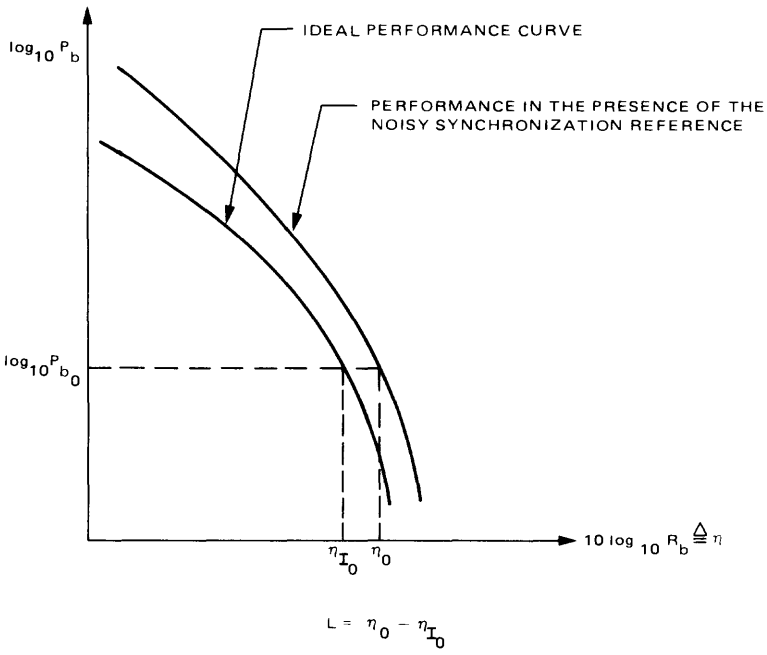


Fig. 5-73. Illustration for the evaluation of noisy reference loss

For BPSK, the slope of the ideal performance curve is obtained by differentiating (5.6-1); thus,

$$\log_{10} P_{E_I} = \log_{10}\left[\frac{1}{2}\operatorname{erfc}\sqrt{R_b}\right] = \log_{10}\left[\frac{1}{2}\operatorname{erfc}\sqrt{10^{\eta_b/10}}\right]$$

$$\frac{d\log_{10} P_{E_I}}{d\eta_b} = -\frac{1}{10\sqrt{\pi}}\,\frac{\sqrt{R_b}\exp(-R_b)}{\operatorname{erfc}\sqrt{R_b}} \tag{5.6-33}$$

Since for large arguments,

$$\operatorname{erfc} x \cong \frac{\exp(-x^2)}{\sqrt{\pi}\,x} \tag{5.6-34}$$

then (5.6-33) simplifies to

$$\frac{d\log_{10} P_{b_I}}{d\eta_b} \cong \frac{R_b}{10} \tag{5.6-35}$$

Furthermore, applying (5.6-34) to (5.6-5) results in

$$\frac{P_b}{P_{b_I}} = 1 + \frac{\frac{1}{2}\sqrt{\frac{R_b}{\pi}}\exp(-R_b)\,\sigma_\phi^2}{\frac{1}{2}\operatorname{erfc}\sqrt{R_b}} \cong 1 + R_b\,\sigma_\phi^2 \tag{5.6-36}$$

Finally, substituting (5.6-35) and (5.6-36) in (5.6-32) gives the desired simple result for noisy reference loss, namely,

$$L = \frac{10\log_{10}(1 + R_b\,\sigma_\phi^2)}{R_b}\;;\quad R_b = [\operatorname{erfc}^{-1} 2P_{E_I}]^2 \tag{5.6-37}$$

By a similar procedure, we can obtain expressions for noisy reference loss for QPSK and UQPSK modulations. The results are as follows:

- QPSK

$$L_1 = L_2 = \frac{10\log_{10}[1 + R_b(1 + 2R_b)\,\sigma_\phi^2]}{R_b} \tag{5.6-38}$$

where L_1 and L_2 are the noisy reference loss in Channels 1 and 2 respectively.

- UQPSK

$$L_2 = \frac{10 \log_{10} \left[1 + R_{s2} \left(1 + 2R_{s1} \frac{T_2}{T_1} \overline{d_{12}^2}\right) \sigma_\phi^2\right]}{R_2}$$

$$L_1 = \frac{10 \log_{10} \left[1 + R_{s1} (1 + 1R_{s2} \overline{d_{12}^2}) \sigma_\phi^2\right]}{R_{s1}} \tag{5.6-39}$$

where

$$R_{si} = \frac{P_i T_{si}}{N_0}, \; i = 1, 2 \tag{5.6-40}$$

5.7 Effect of Doppler on Telemetry Performance

The uncompensated doppler shift of the downlink produces a doppler stress in the receiver. Stated mathematically, the phase error has a nonzero mean [5-10]. Of course, this means a performance degradation to the telemetry channel. In this section, the performances of an uncoded (subsection 5.7.1) and a convolutionally coded (subsection 5.7.2) telemetry channel operating with doppler-stressed receivers are considered. Both theoretical and simulation approaches are taken, and the results from distinct approaches compared. This section follows the treatment in [5-56]. Only the high-data-rate case, which is the one of most practical interest under this situation, is discussed.

5.7.1 Uncoded PSK Telemetry Performance in the Presence of Uncompensated Doppler

Uncoded PSK telemetry performance is specified in terms of bit error probability P_b, which is given by (5.3-15). For the high-data-rate case under discussion, the conditional probability function $P_b(\phi)$ is given by (5.3-16). In the presence of uncompensated doppler, the pdf $p(\phi)$ of the receiver PLL steady-state phase error may be approximated [5-10, p. 489] by

$$p(\phi) \cong \begin{cases} \dfrac{\exp(\beta\phi + \rho \cos \phi)}{2\pi I_0(\rho)}; & |\phi| \leqslant \pi \\ 0; & \text{otherwise} \end{cases} \tag{5.7-1}$$

where $\beta = \rho \sin \phi_{ss}$, ρ is the SNR in the PLL noise bandwidth B_L, ϕ_{ss} is the static phase error in the PLL. A more accurate normalization constant for the pdf is obtained by replacing $I_0(\rho)$ in (5.7-1) with

$$C_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left[\rho\left(\phi \sin \phi_{ss} + \cos \phi\right)\right] d\phi$$

Hence

$$p(\phi) = \begin{cases} \dfrac{1}{C_0} \exp\left[\rho\left(\phi \sin \phi_{ss} + \cos \phi\right)\right]; & |\phi| \leq \pi \\ 0; & \text{otherwise} \end{cases} \tag{5.7-2}$$

The resulting expression for the uncoded PSK bit error probability is

$$P_b = \int_{-\pi}^{\pi} P_b(\phi) \frac{\exp\left[\rho\left(\phi \sin \phi_{ss} + \cos \phi\right)\right]}{C_0} d\phi \tag{5.7-3}$$

where $P_b(\phi)$ is given by (5.3-16) and C_0 is given by (5.7-2).

This expression is evaluated as a function of E_b/N_0 for constant values of receiver tracking loop SNR, ρ and static phase error (θ). The results are plotted in Figs. 5-74 through 5-76. For each figure a constant value for the loop static phase error was selected (ϕ_{ss} = 5, 10, and 20 degrees corresponding to Figs. 5-74 through 5-76, respectively) and seven curves of constant loop SNR were plotted (ρ = 4, 6, 8, 10, 12, 14, and 16 dB). For reference, baseband bit error probability versus E_b/N_0 was also plotted in each figure. Baseband bit error probability for an uncoded channel is given by (5.3-11). Radio loss at a selected probability of bit error may be determined from these plots by taking the difference (in E_b/N_0) between the baseband curve and a selected constant ρ curve at the desired P_b.

5.7.2 Convolutionally Encoded/Viterbi-Decoded PSK Telemetry Channel Performance (High-Data-Rate Model)

The convolutionally encoded (K = 7, r = 1/2) Viterbi-decoded telemetry channel bit error probability, P_b, is of the same form as for the uncoded channel case and is given by (5.3-15).

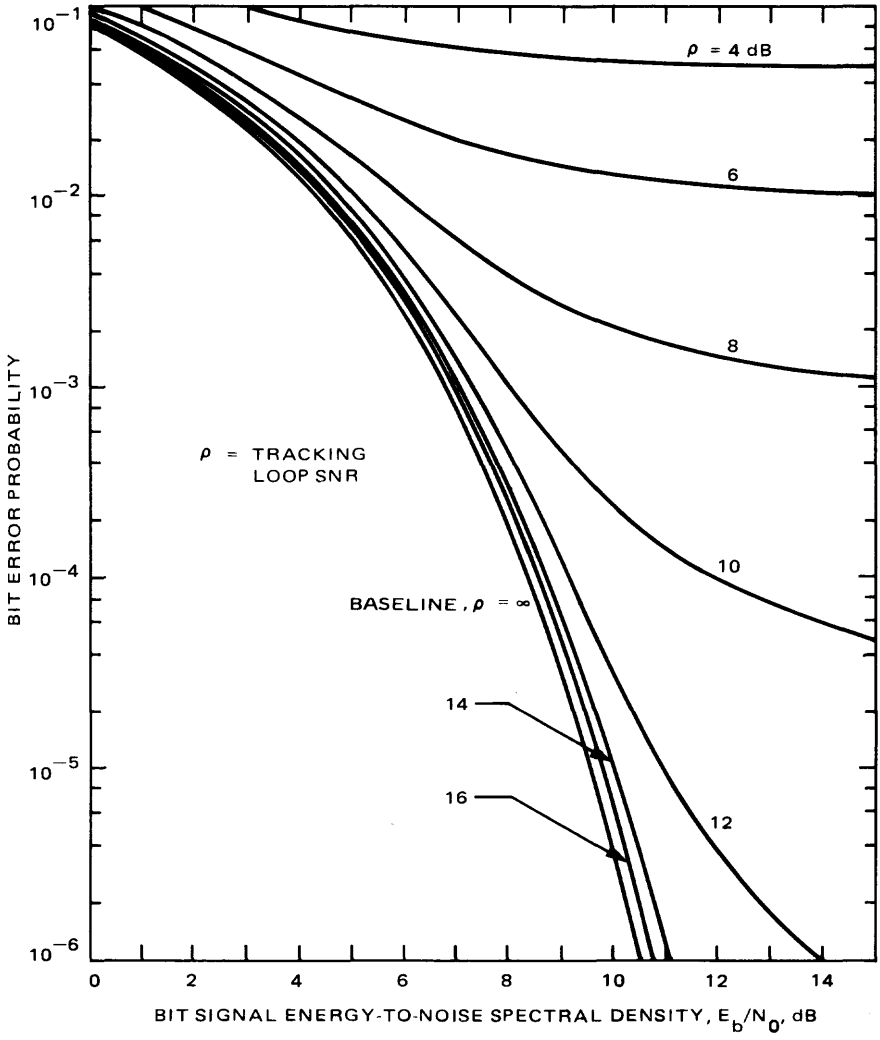


Fig. 5-74. Bit error probability performance (PSK, $\phi_{ss} = 5°$)

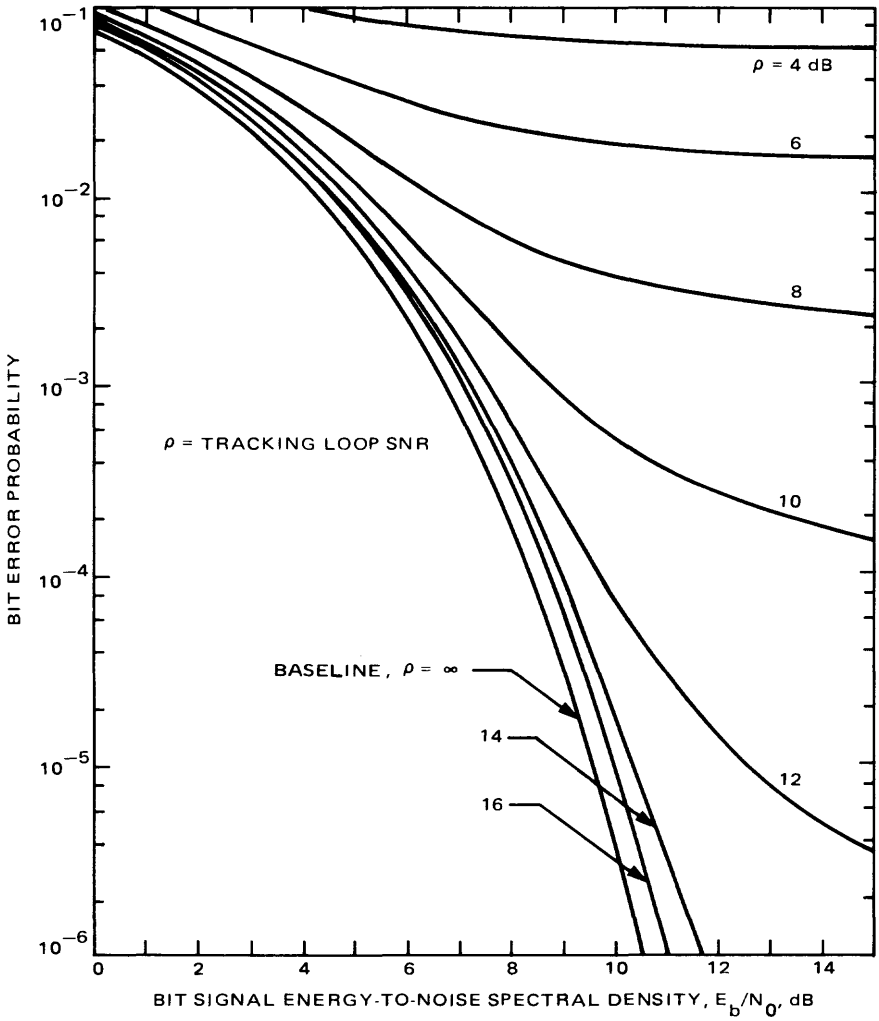


Fig. 5-75. Bit error probability performance (PSK, $\phi_{ss} = 10°$)

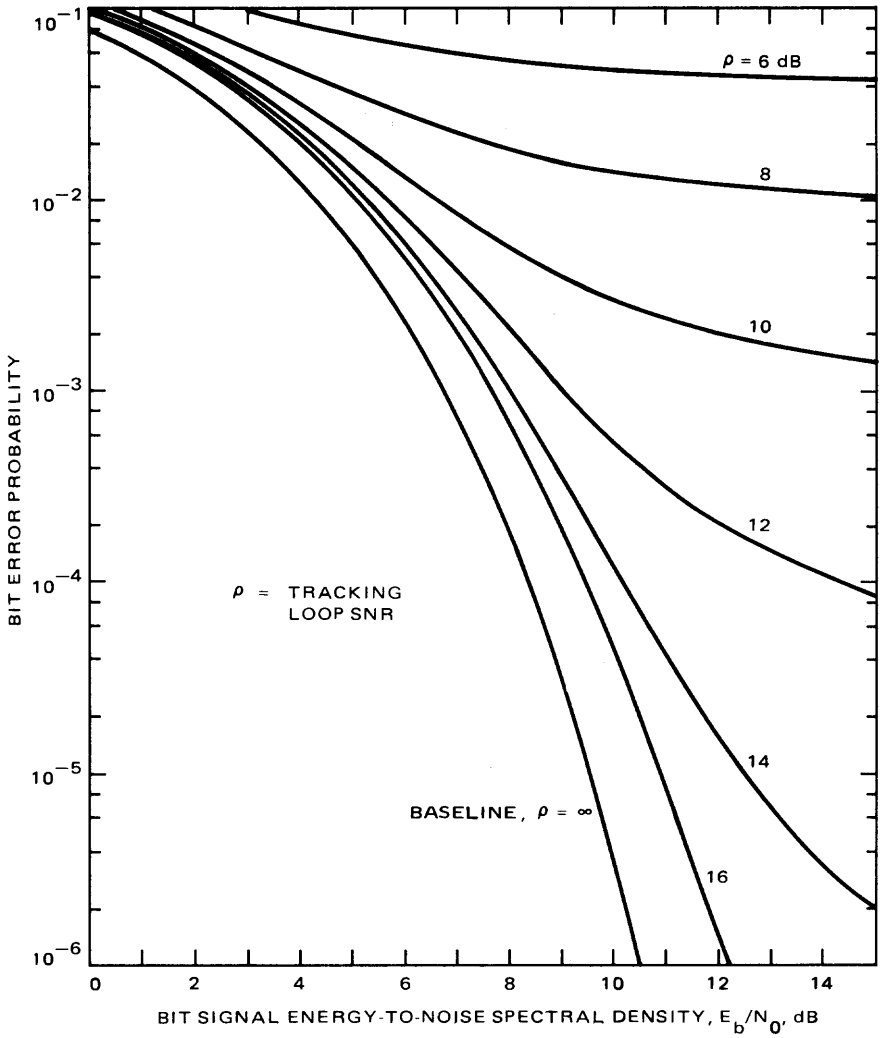


Fig. 5-76. Bit error probability performance (PSK, $\phi_{ss} = 20°$)

Again, in this expression, $P_b(\phi)$ is the conditional probability of a bit error, given a time-varying receiver tracking-loop phase error, ϕ, that is constant over several code-symbol times, and $p(\phi)$ is the probability density function (pdf) for the receiver PLL steady state phase error. The $P_b(\phi)$ for a convolutionally encoded/Viterbi-decoded PSK channel is approximated by (5.4-16) with $x^2 = (E_b/N_0)\cos\phi$; i.e.,

$$P_b(\phi) \cong \begin{cases} \exp\left(\alpha_0 - \alpha_1 \dfrac{E_b}{N_0}\cos\phi\right); & 0 \leqslant |\phi| \leqslant \phi_0 \\ 1 - \exp\left(\alpha_0 - \alpha_1 \dfrac{E_b}{N_0}\cos\phi\right); & \pi - \phi_0 < |\phi| \leqslant \pi \\ 0.5; & \phi_0 < |\phi| \leqslant \pi - \phi_0 \end{cases} \quad (5.7\text{-}4)^6$$

where ϕ_0 is determined by

$$\phi_0 = \cos^{-1}\left(\frac{\ln 2 + \alpha_0}{\alpha_1 \dfrac{E_b}{N_0}}\right)^{1/2} \quad (5.7\text{-}5)$$

The expression for the pdf, $p(\phi)$, is the same as for the uncoded case, and, again, may be approximated by (5.7-3).

Substituting (5.7-1), (5.7-2), and (5.7-4) in (5.3-15), we obtain the bit error probability P_b, which is evaluated as a function of E_b/N_0 for constant values of receiver tracking-loop SNR and static phase error ϕ_{ss}. The results are plotted in Figs. 5-77 through 5-79. For each figure, a constant value for the loop static phase error was selected (ϕ_{ss} = 5, 10, and 20 degrees corresponding to Figs. 5-77 through 5-79, respectively) and seven curves of constant loop SNR were plotted (ρ = 4, 6, 8, 10, 12, 14, and 16 dB). For reference, baseband bit error probability versus E_b/N_0 was also plotted in each figure. Baseband bit error probability for a convolutionally encoded (K = 7, r = 1/2) Viterbi-decoded channel is approximated by

$$P_b \cong \exp\left(\alpha_0 - \alpha_1 \frac{E_b}{N_0}\right) \quad (5.7\text{-}6)$$

[6] Note: [5-53] discusses the need to account for the effect of symbol-stream inversion for $\pi/2 \leqslant |\phi| \leqslant \pi$.

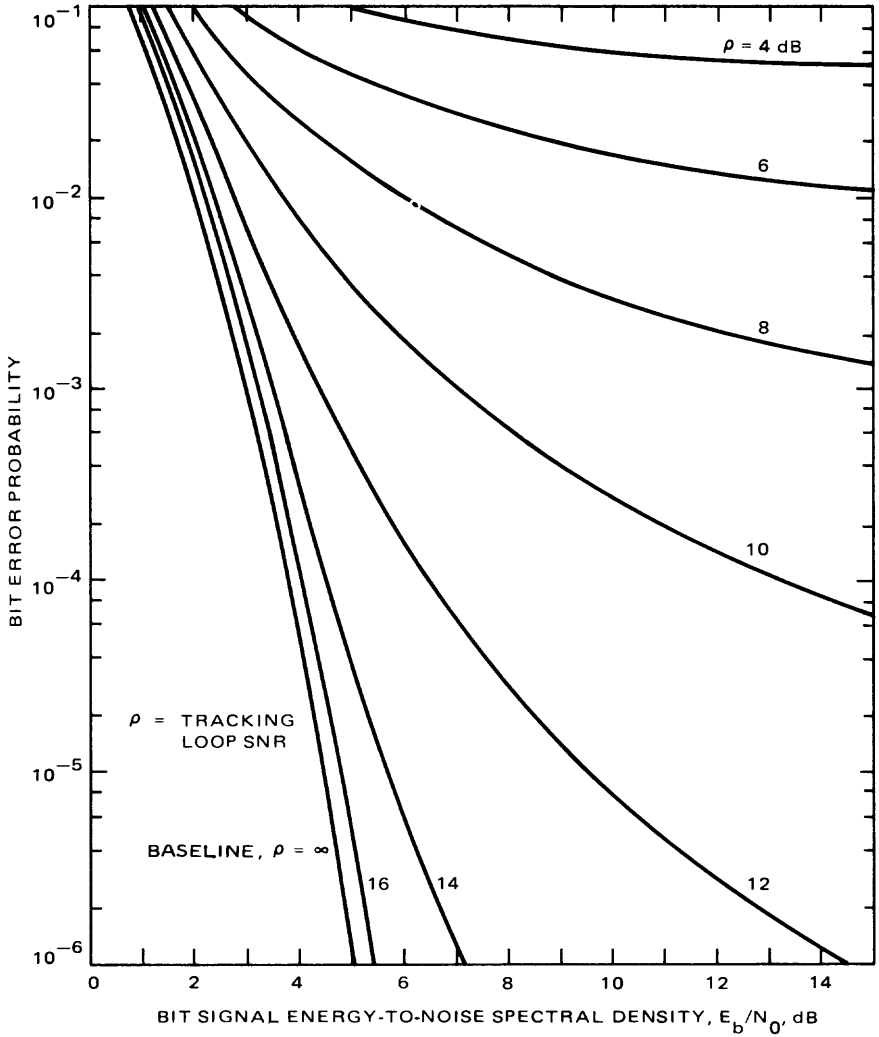


Fig. 5-77. Bit error probability performance (convolutionally coded PSK, $r = 1/2, K = 7, \phi_{ss} = 5°$)

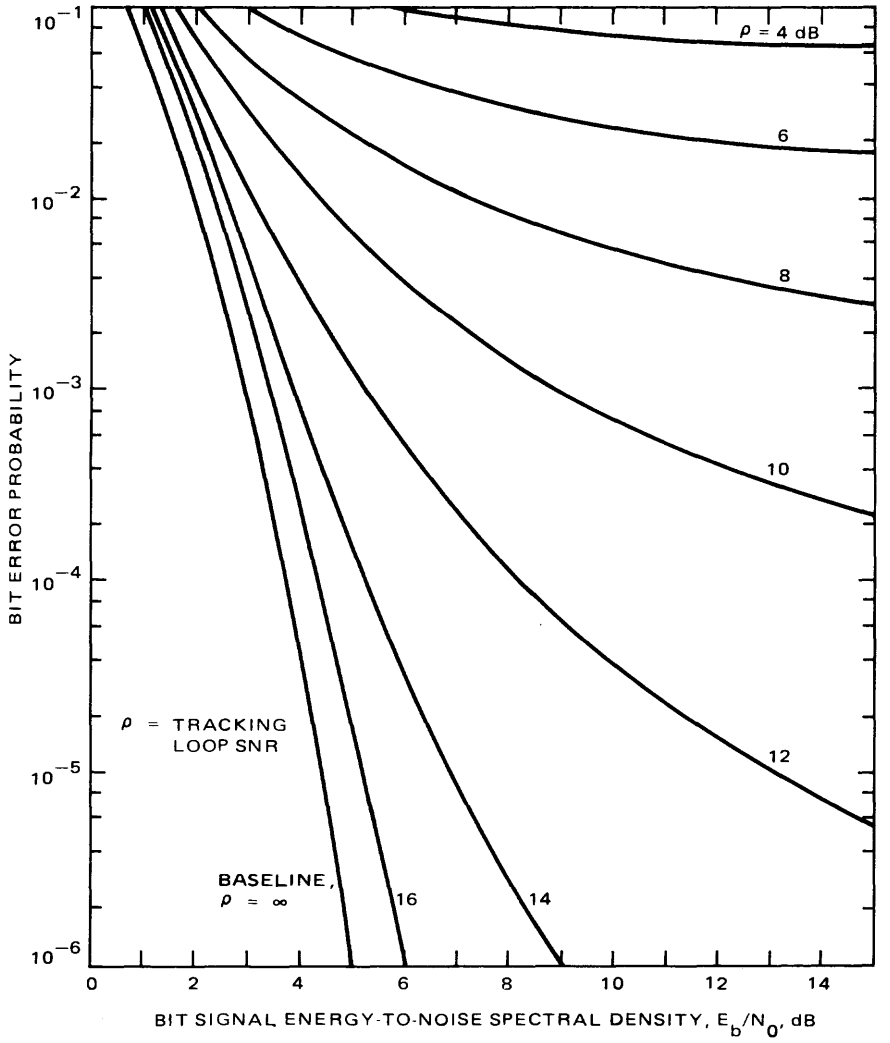


Fig. 5-78. Bit error probability performance (convolutionally coded PSK, $r = 1/2, K = 7, \phi_{ss} = 10°$)

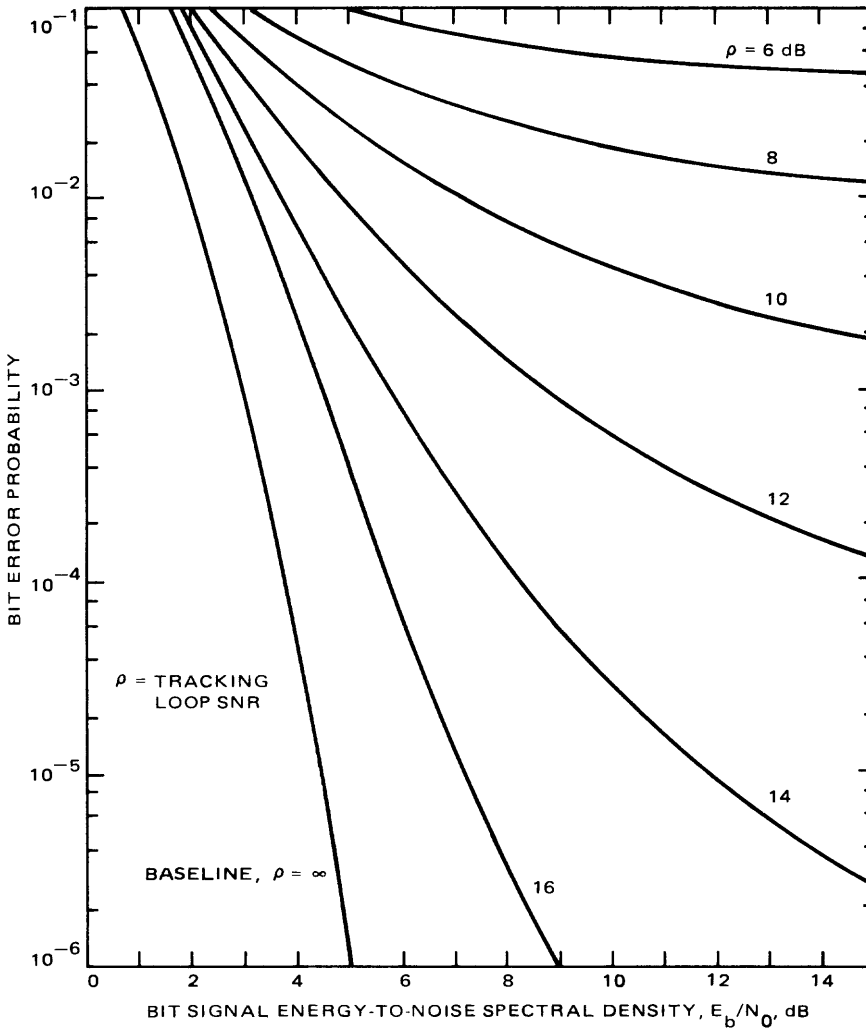


Fig. 5-79. Bit error probability performance (convolutionally coded PSK, $r = 1/2, K = 7, \phi_{ss} = 20°$)

where α_0, α_1 and E_b/N_0 are as defined above for (5.7-4). Radio loss at a selected bit error probability may be determined from these plots by taking the difference (in E_b/N_0) between the baseband curve and a selected constant ρ curve at the desired P_b.

5.7.3 Simulation Results

The simulation results [5-56] shown in Figs. 5-80 through 5-82 were obtained by replacing the theoretical probability density function $p(\phi)$ of the loop phase error ϕ with one derived from computer simulation of the RF receiver phase-locked loop. As the "bin" size for forming the histogram during the computer simulation was 0.1 degree, the maximum quantization error was ±0.05 degree, and the probability density function derived from the computer simulation will have the form

$$p(\phi) = \sum_{k=-1800}^{1800} p_k \, \delta\left(\phi - k\frac{\pi}{1800}\right) \tag{5.7-7}$$

For $-1799 \leqslant k \leqslant 1799$, p_k is the fraction of the time during the computer simulation that the phase-locked loop phase error is between $k/10 - 0.05$ and $k/10 + 0.05$ degree. P_{1800} is the fraction of the time that the phase error is between 179.95° and 180°, and P_{-1800} is the fraction of the time that the phase error is between -180° and -179.95°. Then, using (5.7-12) in (5.7-7), the bit error probability calculated using the phase error probability density function from the computer simulation would be

$$P_b = \sum_{k=-1800}^{1800} p_k P_b\left(\frac{k\pi}{1800}\right) \tag{5.7-8}$$

where $P_b(\phi)$ is specified in (5.3-16).

The RF receiver phase-locked loop considered herein is assumed to have finite gain K and loop filter transfer function given in (3.2-13).

The inputs to the computer program are the one-sided threshold loop bandwidth B_{L0}, the threshold damping factor r_{L0}, the threshold loop gain K_{L0}, the noise bandwidth B_{LI} of the filter and receiver stages preceding the limiter in the bandpass filter, and the threshold signal-to-noise ratio γ_{L0} in B_{L0} at the receiver input. For all the data presented in this section, γ_{L0} is equal to two. Note that this is equivalent to 0 dB in the two-sided threshold loop bandwidth. Given these parameters and the carrier margin M_L (above threshold) at which the RF receiver phase-locked loop is operating, the computer

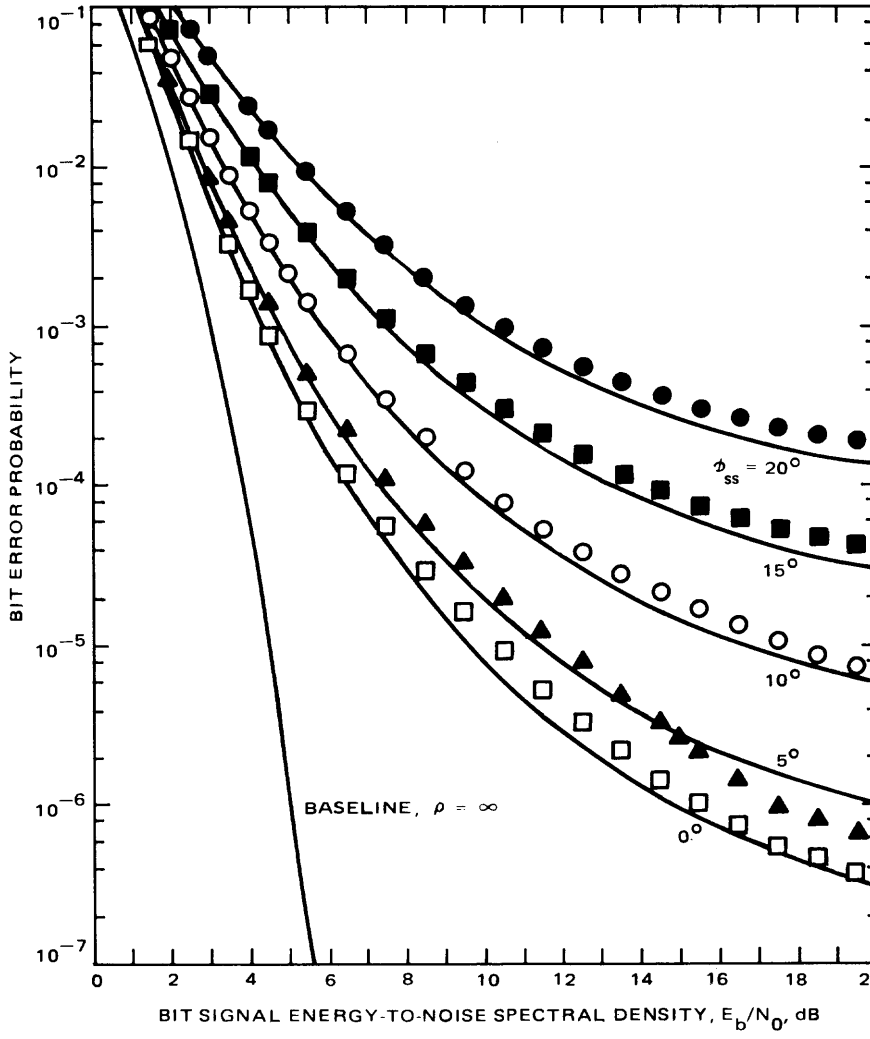


Fig. 5-80. Viterbi decoder performance: theoretical vs simulation results (ρ = 11.5 dB)

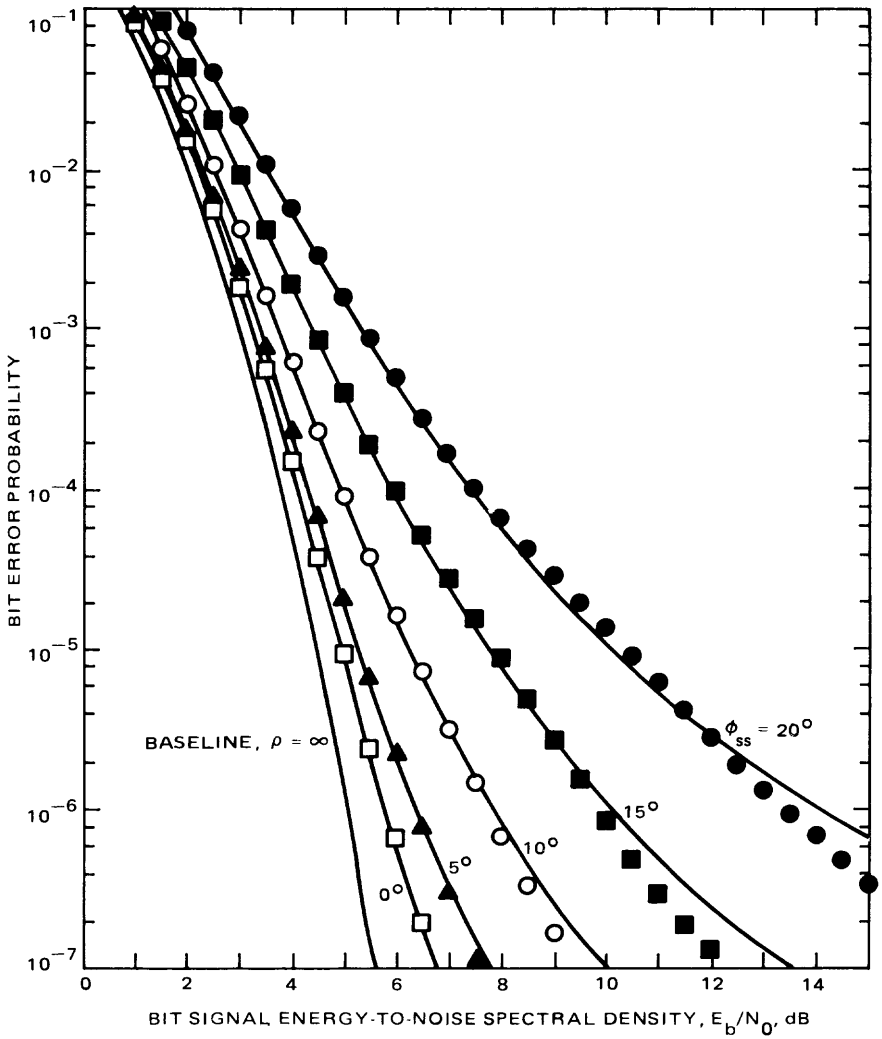


Fig. 5-81. Viterbi decoder performance: theoretical vs simulation results (ρ = 14.5 dB)

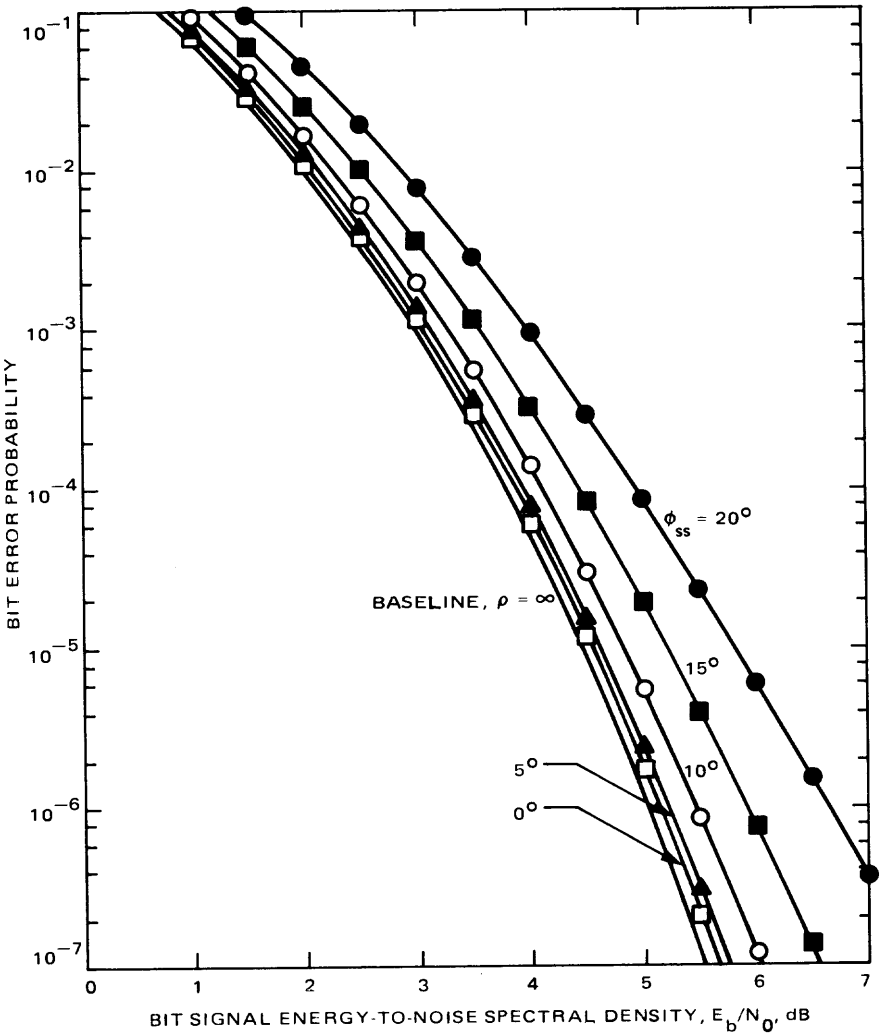


Fig. 5-82. Viterbi decoder performance: theoretical vs simulation results (ρ = 18 dB)

program, using the linear theory for a phase-locked loop preceded by a bandpass limiter (see Chapter 3), calculates τ_1, τ_2, K, and the one-sided noise bandwidth B_L and loop signal-to-noise ratio ρ.

The computer simulation of the RF receiver numerically solves the differential equations for the loop phase error $\phi(t)$ for a phase-locked loop driven by the normalized noise process $n(t)$. For convenience, the differential equations are expressed in state variable form. Thus, according to [5-10],

$$\dot{\phi} = -\Omega_0 - \sqrt{PK}\, F_0 [\sin \phi + n] + \dot{y} \tag{5.7-9}$$

where y is the state variable,

$$\dot{y} = -\frac{1}{\tau_1} y - (1 - F_0) \frac{\sqrt{PK}}{\tau_1} [\sin \phi + n] \tag{5.7-10}$$

and F_0 is the ratio τ_2/τ_1 of the loop filter time constants. The time interval between successive iterations was

$$\Delta t = \frac{0.01}{B_L} \tag{5.7-11}$$

and 16,220,160 iterations were used for each simulation.

The variance of the Guassian process $n(t)$ was adjusted to be

$$\sigma_n^2 = \frac{1}{2\rho B_L \Delta t} \tag{5.7-12}$$

In (5.7-9), Ω_0 is the angular frequency (radians/sec) offset between the received RF frequency and the RF receiver "best lock" frequency. For the simulation data presented herein, Ω_0 was selected so as to make

$$\sin^{-1}\left(\frac{\Omega_0}{\sqrt{PK}}\right) = \phi_{ss} \tag{5.7-13}$$

where ϕ_{ss} is the loop static phase error that would be observed with loop gain $\sqrt{PK}$ in the absence of noise. Simulation data are presented for ϕ_{ss} equal to 0, 5, 10, 15 and 20 degrees.

The simulation results are plotted in Figs. 5-80, 5-81, and 5-82. For each figure a constant value of receiver tracking loop SNR (ρ) was selected (ρ = 11.5 dB, 14.5 dB, and 18.0 dB for Figs. 5-80, 5-81, and 5-82, respectively)

and five curves of constant loop static phase error were plotted (θ = 0, 5, 10, 15, and 20 degrees). The curves were plotted using (5.3-15) (with appropriate substitutions for $P_b(\phi)$ from (5.7-3) and for $p(\phi)$ from (5.7-2)), and provide bit error probability as a function of E_b/N_0. For reference, baseband bit error probability versus E_b/N_0 was also plotted in each figure, using (5.7-6). The simulation data are indicated on these plots by symbols. A unique symbol is used to identify each of the five loop static phase errors considered. Note that all three figures show good agreement between the theoretical curves and the simulation results. Agreement improves with increased loop signal-to-noise ratio (ρ), and with decreased E_b/N_0. Perturbations in the simulation results for bit error rates less than 10^{-5} are probably due to too few trials per simulation point.

5.8 Arrayed System

There are two types of arraying pertinent to the DSN: baseband combining and carrier arraying. Baseband combining is the combining at baseband of received telemetry streams from several geographically separated DSSs. The purpose of baseband combining is to improve overall bit error rate performance. Carrier arraying is the combining at RF of received carrier signals from several geographically separated DSSs. The purpose of carrier arraying is to improve overall carrier loop SNR.

5.8.1 Baseband Combining

When baseband combining is the only arraying technique being used, the arrayed system configuration is as diagramed in Fig. 5-83.

As indicated in Fig. 5-83, the baseband signal from each DSS must be properly delayed and weighted, to ensure optimum combining. If E_{bi}/N_{0i} denotes the energy-per-bit-to-noise spectral density ratio for the ith station in the array, then the optimum weighting factor β_i for the ith station is given by

$$\beta_i = \sqrt{\frac{E_{bi}}{E_{b1}} \frac{N_{01}}{N_{0i}}} \tag{5.8-1}$$

Under the conditions of strong carrier loop SNR at each contributing station and optimum combining, the effective bit energy-to-noise spectral density ratio E_b/N_0 after the combiner is

$$\frac{E_b}{N_0} = \sum_{i=1}^{N} \frac{E_{bi}}{N_{0i}} \tag{5.8-2}$$

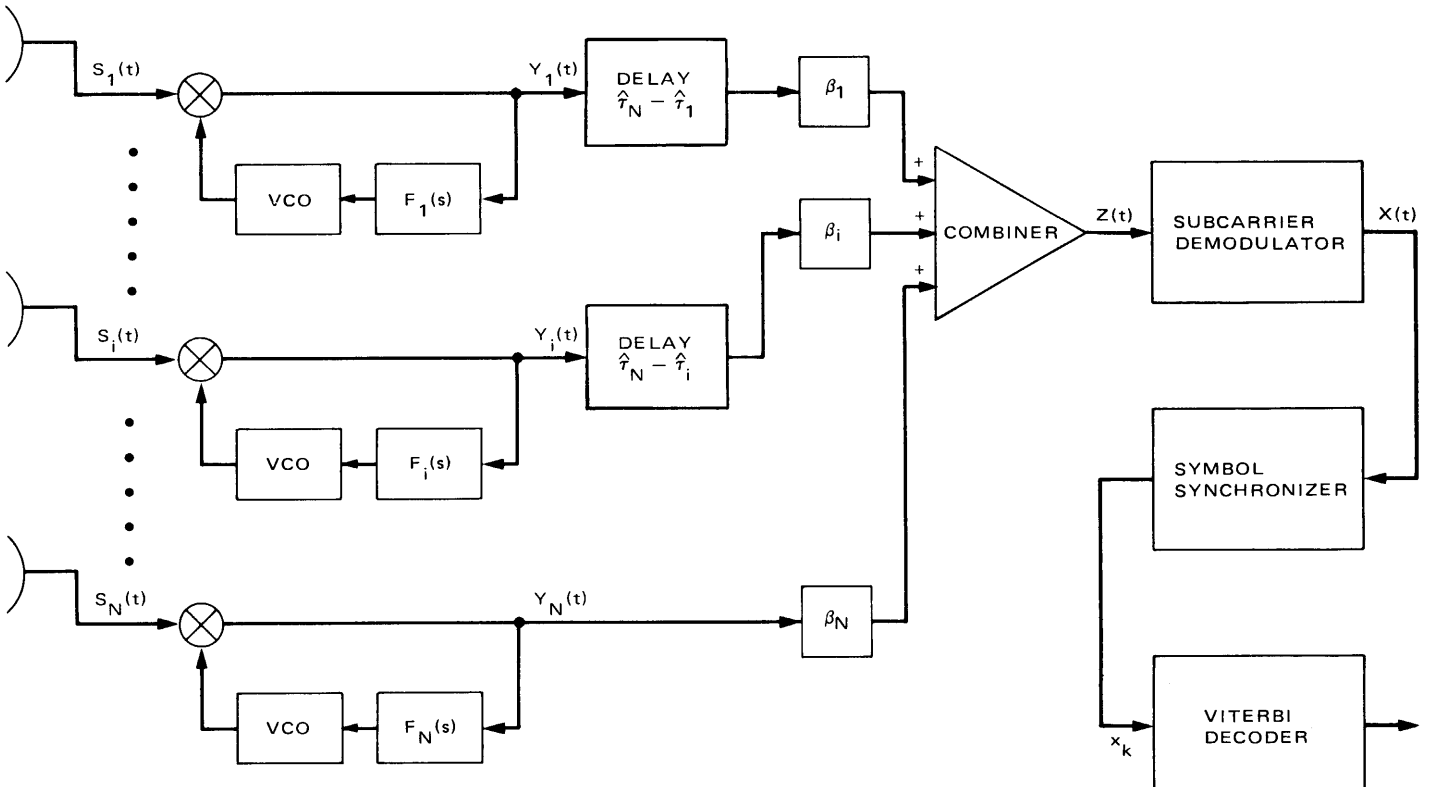


Fig. 5-83. Configuration for arrayed network with baseband combining

If the carrier loop SNR at each contributing station is *not* strong, then the bit error probability P_b for the arrayed system is given by

$$P_b = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi} f\left[\frac{\left(\sum_{i=1}^{N}\frac{E_{bi}}{N_{0i}}\cos\phi_i\right)^2}{\sum_{i=1}^{N}\frac{E_{bi}}{N_{0i}}}\right] p(\phi_1)p(\phi_2)\cdots p(\phi_N)d\phi_1 d\phi_2\cdots d\phi_N \tag{5.8-3}$$

where

$f(\cdot)$ = the baseline bit error rate performance

ϕ_i = the carrier tracking phase error of the ith station

and the ϕ_i are distributed according to the Tikhonov pdf (replacing ϕ by ϕ_i, and ρ by ρ_i in (5.3-18)). The parameter ρ_i denotes the carrier loop SNR for the ith station.

Figure 5-84 plots P_b, as given in (5.8-3), versus E_b/N_0 (the effective bit energy-to-noise spectral density ratio after combining) for an arrayed network, with the following assumptions applying:

(1) convolutional code $K = 7$, $r = 1/2$, with Viterbi decoding, $Q = 3, f(x)$ is given by (5.6-30)

(2) the optimum weighting factors are

$$\beta_1 = 1$$

$$\beta_2 = 0.513$$

$$\beta_3 = \beta_4 = 0.589 \tag{5.8-4}$$

The intent of assumption 2 is to model a four-station array composed of a 64-m, and three 34-m antennas, with Station 1 being the station with the 64-m antenna.

For a derivation of the results of this subsection, the reader should consult [5-57].

5.8.2 Carrier Arraying

When an arrayed system has the capability for both baseband combining and carrier arraying, its configuration is as diagramed in Fig. 5-85.

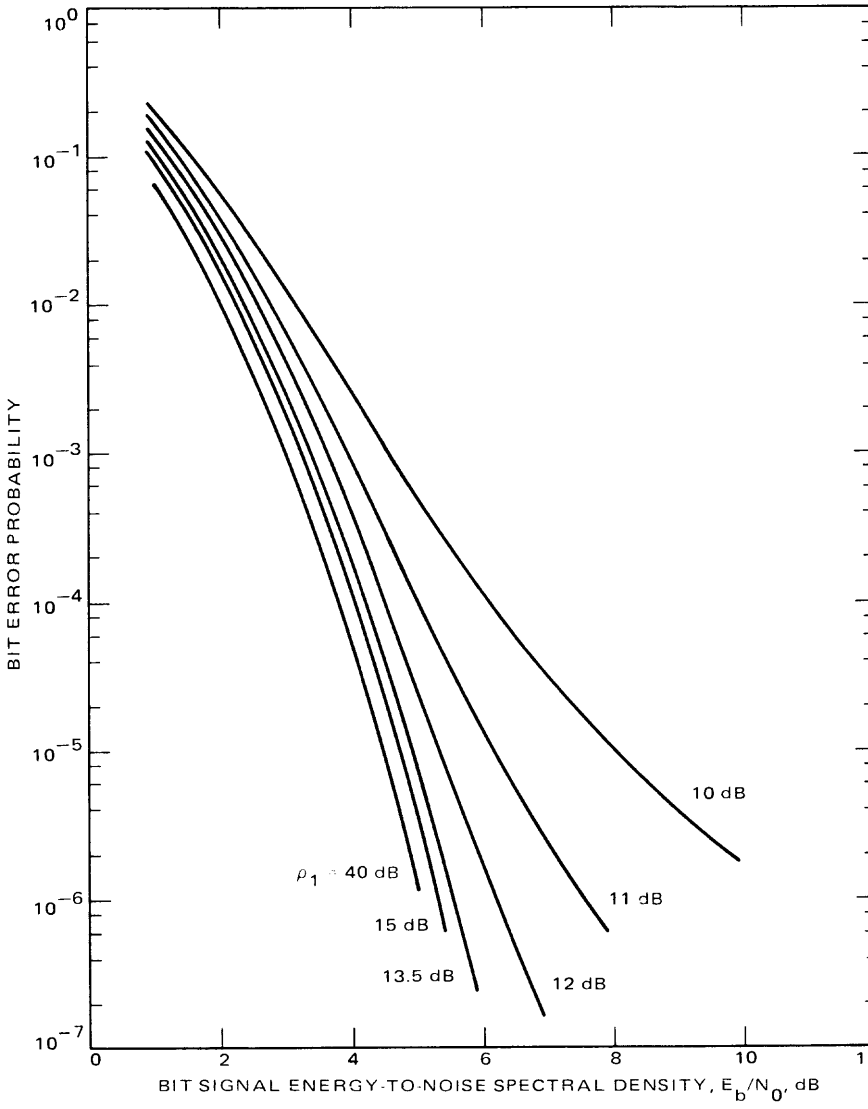


Fig. 5-84. Performance curves for a four-element array using a baseband combiner

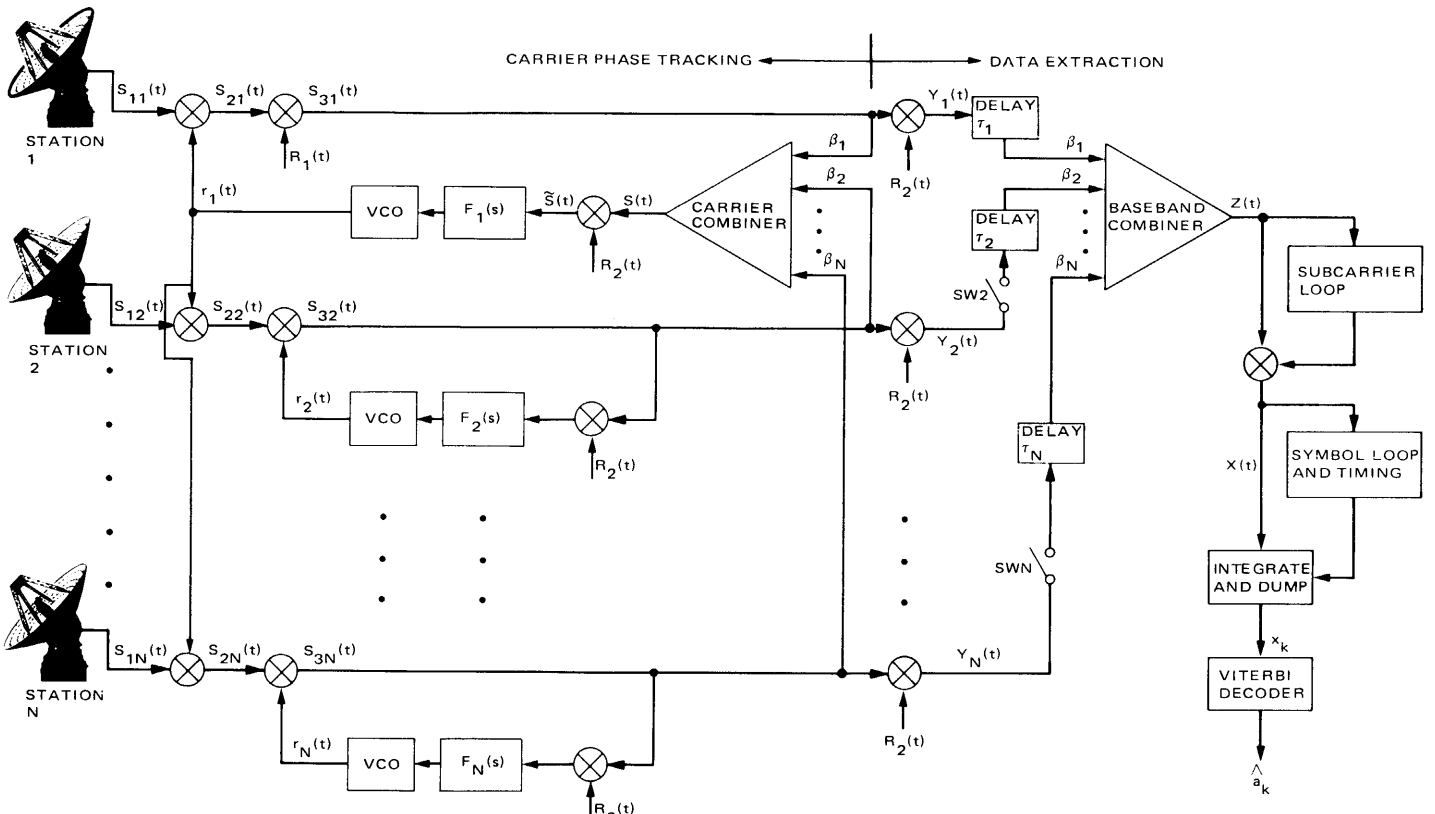


Fig. 5-85. Configuration for arrayed network with carrier and baseband arraying

Telemetry performance with simultaneous baseband combining and carrier arraying will be deferred until the next subsection. In this subsection, carrier arraying alone is considered. So, for the purposes of this subsection, the switches SW2 through SWN of Fig. 5-85 are open.

The loop-SNR improvement factor of carrier arraying (relative to the carrier loop SNR of Station 1 unaided) [5-58] is

$$\text{Improvement in loop SNR} = \frac{\rho}{\rho_1} = \frac{G(\tilde{\alpha}_0 + 2\tilde{\alpha}_1)\,\Gamma_1}{(\tilde{\alpha}_0 + 2\tilde{\alpha})\,\Gamma} \tag{5.8-5}$$

where

ρ_1 = the loop SNR of Station 1 unaided

$\tilde{\alpha}_1$ = the limiter suppression factor of Station 1 unaided

Γ_1 = the limiter performance factor of Station 1 unaided

$$\tilde{\alpha} = \sqrt{\frac{0.7854\rho_{in} + 0.4768\rho_{in}^2}{1 + 1.024\rho_{in} + 0.4768\rho_{in}^2}} \tag{5.8-6}$$

$$\Gamma \cong \frac{1 + \rho_{in}}{0.862 + \rho_{in}} \tag{5.8-7}$$

$\tilde{\alpha}_0$ = $\tilde{\alpha}$ evaluated at carrier threshold

$$\rho_{in} = \frac{P_1 G^2}{\sum_{i=1}^{N} \beta_i^2 N_{0i} B_{IF}} \tag{5.8-8}$$

$$G = \sum_{i=1}^{N} \beta_i \gamma_i \tag{5.8-9}$$

$$\gamma_i = \sqrt{\frac{P_i}{P_1}} \tag{5.8-10}$$

B_{IF} = the one-sided bandwidth of the IF filter preceding the limiter

P_i = the carrier power at station i

and the optimum weighting factors for carrier arraying are

$$\beta_i = \gamma_i \frac{N_{01}}{N_{0i}} \tag{5.8-11}$$

With a 64-m station considered as Station 1, the loop-SNR improvement factor is plotted in Fig. 5-86. For more details and a derivation of the carrier loop SNR improvement factor, the reader should consult [5-58].

5.8.3 Simultaneous Baseband Combining and Carrier Arraying

This subsection considers telemetry performance with simultaneous baseband combining and carrier arraying. In Fig. 5-85, the switches SW2 through SWN are to be considered closed. With optimum weighting factors employed, the bit error probability performance is

$$P_b = 2\int_0^{\pi/2} P_b(\phi_1)p(\phi_1)d\phi_1 + 2\int_{\pi/2}^{\pi} [1 - P_b(\phi_1)]\, p(\phi_1)d\phi_1 \tag{5.8-12}$$

The bit error probability conditioned on ϕ_1, $P_b(\phi_1)$, is

$$P_b(\phi_1) = f\left(\sum_{i=1}^{N} \frac{E_{bi}}{N_{0i}} \cos^2 \phi_1\right) \tag{5.8-13}$$

The carrier arrayed tracking phase error ϕ_1 has the Tikhonov pdf (see (5.3-18), replacing ϕ by ϕ_1), where

$$\rho = \sum_{i=1}^{N} \frac{P_i \cos^2 \theta_m}{N_0 \widetilde{B}_L \Gamma} \tag{5.8-14}$$

θ_m = telemetry modulation index (5.8-15)

$$\widetilde{B}_L = \frac{1 + \widetilde{r}}{4\, \tau_{21}} \tag{5.8-16}$$

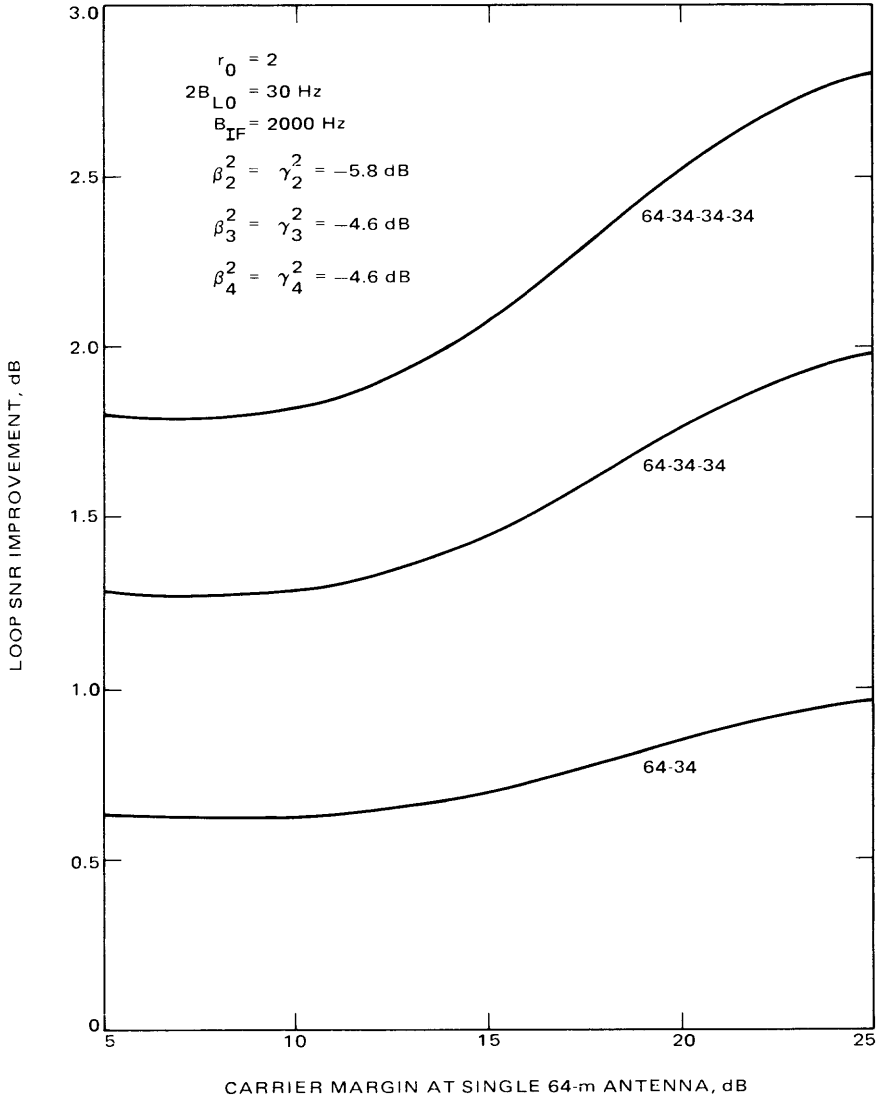


Fig. 5-86. Loop SNR improvement for carrier arraying with coupled PLLs with respect to a single 64-m antenna station

$$\tilde{r} = \frac{\sqrt{8}\tilde{\alpha} r_1}{\pi \sqrt{P_1} \cos \theta_m} \tag{5.8-17}$$

$$r_1 = \left(\sqrt{P_1} K_1 \frac{\tau_{21}^2}{\tau_{11}}\right) \cos \theta_m \tag{5.8-18}$$

$$K_1 = \text{loop gain at Station 1} \tag{5.8-19}$$

τ_{21} and τ_{11} are the time constants of the Station 1 loop whose loop filter has the transfer function

$$F_1(s) = \frac{1 + \tau_{21} s}{1 + \tau_{11} s} \tag{5.8-20}$$

For the assumptions (1) and (2) of subsection 5.8.1 and with simultaneous baseband combining and carrier arraying, telemetry bit error probability performance has been calculated. The results are shown in Fig. 5-87. The horizontal axis is E_{b1}/N_0, which is the bit energy-to-noise spectral density ratio at Station 1 (the 64-m station). This is unlike Fig. 5-84 whose horizontal axis is the E_b/N_0 of (5.8-2). For derivation of the equations of this subsection, see [5-58].

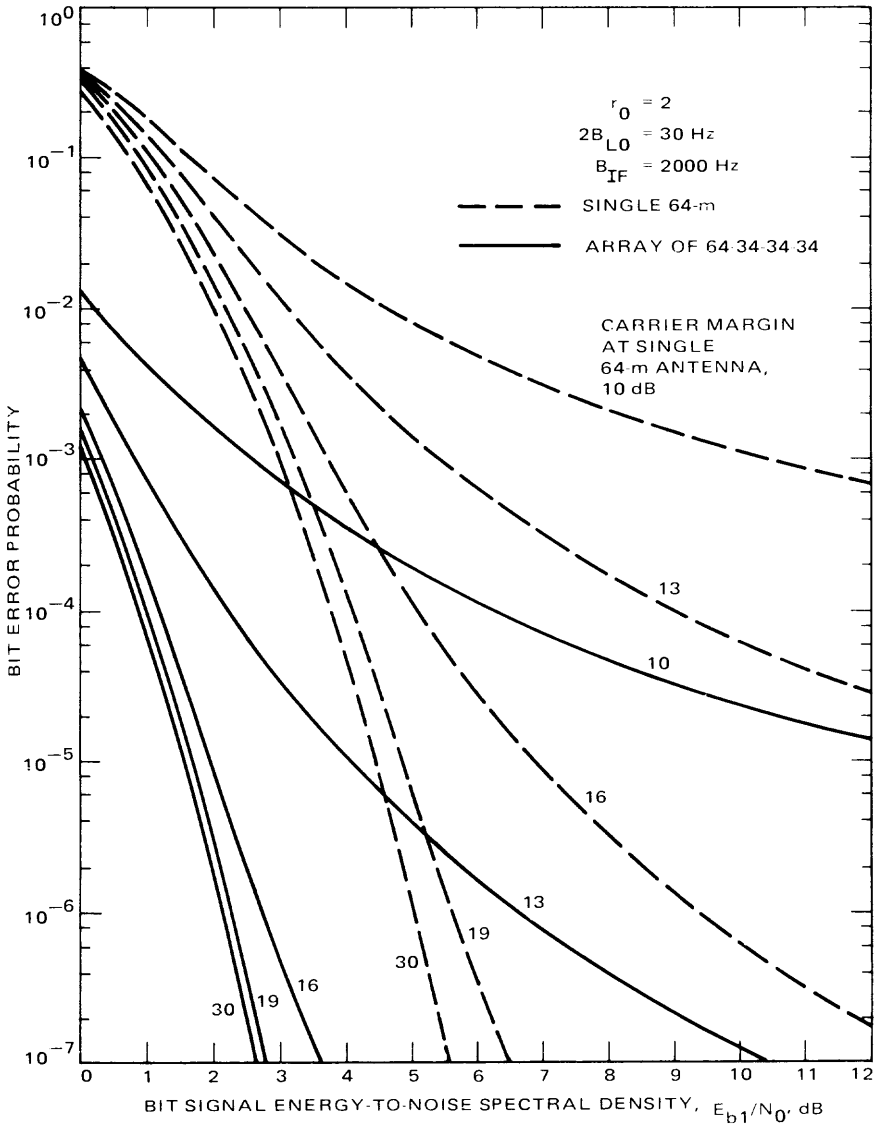


Fig. 5-87. Comparison of bit error probabilities of combined carrier and baseband arraying (64-34-34-34) with a single 64-m, for various carrier margins at a single 64-m antenna

References

5-1. Viterbi, A. J., *Principles of Coherent Communication*, McGraw-Hill, N.Y., 1966.

5-2. Wozencraft, J. M., and Jacobs, I. M., *Principles of Communication Engineering*, Wiley, N.Y., 1965.

5-3. Lindsey, W. C., and Simon, M. K., *Telecommunication Systems Engineering*, Prentice-Hall, Englewood Cliffs, N.J., 1973.

5-4. Weber, W. J., "Description of the SDA," unpublished notes.

5-5. Brockman, M. H., "MMTS: Performance of Subcarrier Demodulation," *Space Programs Summary 37-52*, Vol. II, pp. 127-141, Jet Propulsion Laboratory, Pasadena, Calif., July 31, 1968.

5-6. Weber, W. J., "Description of SSA," unpublished notes.

5-7. Layland, J. W., "Telemetry Bit Synchronization Loop," *Space Programs Summary 37-46*, Vol. III, pp. 204-215, Jet Propulsion Laboratory, Pasadena, Calif., July 31, 1967.

5-8. Schwartz, M., Bennett, W. R., and Stein, S., *Communication Systems and Techniques*, pp. 63-68, 310-313, McGraw-Hill, N.Y., 1966.

5-9. Turin, G. L., "An Introduction to Matched Filters," *IRE Transactions on Information Theory*, June 1960.

5-10. Lindsey, W. C., *Synchronization Systems in Communication and Control*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1972.

5-11. Peterson, W. W., and Weldon, E. J., *Error-Correcting Codes*, 2nd edition, MIT Press, Cambridge, Mass., 1972.

5-12. Lin, S., and Costello, D., *An Introduction to Error-Correcting Codes*, Prentice-Hall, Englewood Cliffs, N.J., 1982.

5-13. Forney, G. D., "Coding and Its Application in Space Communications," *IEEE Spectrum*, pp. 47-58, June 1970.

5-14. Viterbi, A. J., and Omura, J. K., *Principles of Digital Communication and Coding*, McGraw-Hill, N.Y., 1979.

5-15. Clark, G. C., and Cain, J. B., *Error-Correction Coding for Digital Communication*, Plenum Press, N.Y., 1981.

5-16. McEliece, R. J., *The Theory of Information and Coding*, Addison-Wesley, Reading, Mass., 1977.

5-17. Hamming, R. W., *Coding and Information Theory*, Prentice-Hall, Englewood Cliffs, N.J., 1980.

5-18. Abramson, N., *Information Theory and Coding*, McGraw-Hill, N.Y., 1963.

5-19. Pless, V., *Introduction to the Theory of Error-Correcting Codes*, Wiley, N.Y., 1981.

5-20. Shannon, C. E., and Weaver, W., *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, Ill., 1963.

5-21. Berlekamp, E. R., *Algebraic Coding Theory*, McGraw-Hill, N.Y., 1968.

5-22. Berlekamp, E. R., editor, *Key Papers in the Development of Coding Theory*, IEEE Press, N.Y., 1974.

5-23. Berlekamp, E. R., "The Technology of Error-Correcting Codes," *Proceedings of the IEEE*, Vol. 68, No. 5, pp. 564-593, May 1980.

5-24. MacWilliams, F. J., and Sloane, N. J. A., *The Theory of Error-Correcting Codes*, Elsevier/North Holland, N.Y., 1977.

5-25. Reed, I. S., and Solomon, G., "Polynomial Codes over Certain Finite Fields," *J. Soc. Industrial and Applied Mathematics*, 8, pp. 300-304, 1960.

5-26. Blake, I. F., and Mullin, R. C., *The Mathematical Theory of Coding*, Academic Press, N.Y., 1975.

5-27. Blake, I. F., and Mullin, R. C., *Introduction to Algebraic and Combinatorial Coding Theory*, Academic Press, N.Y., 1976.

5-28. Hartnett, W. E., *Foundations of Coding Theory*, Kluwer, Boston, Mass., 1975.

5-29. Van Lint, J. H., *Coding Theory*, Springer-Verlag, Berlin, 1971.

5-30. Massey, J. L., *Error Bounds for Tree Codes, Trellis Codes, and Convolutional Codes with Encoding and Decoding Procedures*, Springer-Verlag, Berlin, 1977.

5-31. Viterbi, A. J., "Convolutional Codes and Their Performance in Communication Systems," *IEEE Transactions on Communications Technology*, Vol. COM-19, No. 5, Pt. 2, pp. 751-772, Oct. 1971.

5-32. Simon, M. K., and Smith, J. G., "Alternate Symbol Inversion for Improved Symbol Synchronization in Convolutionally Coded Systems," *IEEE Transactions on Communications*, Vol. COM-28, No. 2, pp. 228-237, Feb. 1980.

5-33. Forney, G. D., "The Viterbi Algorithm," *Proceedings of the IEEE*, Vol. 61, pp. 268-278, 1973.

5-34. Heller, J. A., and Jacobs, I. M., "Viterbi Decoding for Satellite and

Space Communication," *IEEE Transactions on Communications Technology*, Vol. COM-19, No. 5, Pt. 2, pp. 835-848, Oct. 1971.

5-35. Layland, J. W., *Sequential Decoding with a Noisy Carrier Reference*, Technical Report 32-1526, Vol. XII, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1972.

5-36. Layland, J. W., *A Sequential Decoding Medium Rate Performance Model*, Technical Report 32-1526, Vol. XVIII, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1973.

5-37. Layland, J. W., *A Note on Noisy Reference Detection*, Technical Report 32-1526, Vol. XVII, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 15, 1973.

5-38. Lindsey, W. C., and Blake, I. F., "Effect of Phase Locked Loop Dynamics on Phase Coherent Communication," *Space Programs Summary 37-54*, Vol. III, Jet Propulsion Laboratory, Pasadena, Calif., pp. 192-195, 1968.

5-39. Jacobs, I. M., "Sequential Decoding for Efficient Communication from Deep Space," *IEEE Transactions on Communications Technology*, Vol. COM-15, No. 4, pp. 492-501, Aug. 1967.

5-40. Forney, G. D., *Concatenated Codes*, MIT Press, Cambridge, Mass., 1967.

5-41. Rice, R. F., *Channel Coding and Data Compression System Considerations for Efficient Communication of Planetary Imaging Data*, Technical Memorandum 33-695, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1974.

5-42. Odenwalder, J. P., "Concatenated Reed-Solomon/Viterbi Channel Coding for Advanced Planetary Missions: Analysis, Simulations, and Tests," submitted to JPL by Linkabit Corp., San Diego, Calif., Contract No. 953866, 1974.

5-43. Odenwalder, J. P., et al., "Hybrid Coding Systems Study Final Report," prepared for NASA Ames Research Center, Moffett Field Calif., by Linkabit Corp., San Diego, Calif., Contract No. NAS2-6722.

5-44. Miller, R. L., Deutsch, L. J., and Butman, S. A., *On the Error Statistics of Viterbi Decoding and the Performance of Concatenated Codes*, Publication 81-9, Jet Propulsion Laboratory, Pasadena, Calif., Sept. 1, 1981.

5-45. Liu, K. Y., and Lee, J. J., *An Experimental Study of the Concatenated Reed-Solomon/Viterbi Channel Coding System Performance and Its Impact on Space Communications*, Publication 81-58, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1981.

5-46. Liu, K. Y., *The Effects of Receiver Tracking Phase Error on the Perfor-*

mance of the Concatenated Reed-Solomon/Viterbi Channel Coding System, Publication 81-62, Jet Propulsion Laboratory, Pasadena, Calif., Sept. 1, 1981.

5-47. Divsalar, D., and Yuen, J. H., "Performance of Concatenated Reed-Solomon Viterbi Channel Coding," *TDA Progress Report 42-71*, pp. 81-94, Jet Propulsion Laboratory, Pasadena, Calif., Nov. 15, 1982.

5-48. Edelson, R. E., (editor), *Telecommunications Systems Design Techniques Handbook*, Technical Memorandum 33-571, Jet Propulsion Laboratory, Pasadena, Calif., July 15, 1972.

5-49. Simon, M. K., "Nonlinear Analysis of an Absolute Value Type of an Early-Late Gate Bit Synchronizer," *IEEE Transactions on Communications Technology*, Vol. COM-18, No. 5, pp. 589-596, Oct. 1970.

5-50. Divsalar, D., and Yuen, J. H., "Performance of Unbalanced QPSK in the Presence of Noisy Reference and Crosstalk," *National Telecommunication Conference Proceedings*, Washington, D.C., 1979.

5-51. Simon, M. K., and Alem, W. K., "Tracking Performance of Unbalanced QPSK Demodulators, Part I," *IEEE Transactions on Communications*, Vol. COM-26, No. 8, pp. 1147-1156, Aug. 1978.

5-52. Brown, D. S., "A High Rate One-Way Radio Loss Model for Maximum Likelihood Convolutional Decoding," JPL IOM 3396-76-021, Feb. 16, 1976 (an internal document).

5-53. Koerner, M. A., "A Modified Model for the Performance of a Convolutionally Encoded ($K = 7$, $R = 1/2$) X-band Telemetry Link Using Maximum-Likelihood Viterbi Decoding ($Q = 3$) During Two-Way Operation with an S-band Uplink," JPL IOM 3392-80-031, Feb. 29, 1980 (an internal document).

5-54. Divsalar, D., and Yuen, J. H., "Performance of Convolutionally Coded Unbalanced QPSK Systems," *National Telecommunication Conference Proceedings*, Houston, Texas, 1980.

5-55. Divsalar, D., and Simon, M. K., "Spectral Characteristics of Convolutionally Coded Digital Signals," *IEEE Transactions on Communications*, Vol. COM-28, No. 2, pp. 173-186, Feb. 1980.

5-56. Burow, N. A., and Koerner, M. A., "Effect of RF Receiver Phase-Locked Loop Static Phase Error on Telemetry Performance," JPL IOM 3392-82-024, March 29, 1982 (an internal document).

5-57. Divsalar, D., et al., "The Effect of Noisy Carrier Reference On Telemetry with Baseband Arraying," *TDA Progress Report 42-63*, pp. 128-135, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1981.

5-58. Divsalar, D., and Yuen, J. H., "Improved Carrier Tracking Performance with Coupled Phase-Locked Loops," *TDA Progress Report 42-66*, pp. 148-171, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1981.

5-59. Zigangirov, K. Sh., "Some Sequential Decoding Procedures," Problemy Peredachi Informatsii, Vol. 2, pp. 13-25, 1966.

5-60. Jelinek, F., "A Fast Sequential Decoding Algorithm Using a Stack," *IBM J. Research and Development*, Vol. 13, pp. 675-685, 1969.

5-61. Fano, R. M., "A Heuristic Discussion on Probabilistic Decoding," *IEEE Transactions on Information Theory*, Vol. IT-9, pp. 69-74, 1963.

5-62. Gallager, R. G., *Information Theory and Reliable Communication*, Wiley, N.Y., 1968.

Appendix A
The Viterbi Algorithm

1. Preliminary

Before proceeding to the Viterbi algorithm, a discussion of the trellis-code representation of the convolutional encoder is desirable.

For a constraint length K, code rate $r = b/n$ convolutional encoder, the state is defined by the $b(K-1)$ most recent bits in the shift register. The output bits and transitions between states can be recorded by the trellis diagram of Fig. A-1. The diagram starts in the all-zero state, node a, and makes transitions corresponding to the next data bit. These transitions are denoted by a solid line (branch) for an "0" and dotted line (branch) for a "1". Thus, node a proceeds to node a or b with outputs bits 00 or 11. A branch weight is the number of "1"s in the n code symbols in the branch.

2. The Algorithm

The maximum-likelihood or Viterbi decoding algorithm is an efficient and practical decoding technique for short-constraint-length convolutional codes [5-31]. It has been shown [5-33] that the Viterbi algorithm is, in fact, maximum-likelihood decoding. A brute-force maximum-likelihood decoder would calculate the likelihood of the received data for code symbol sequences on all paths through the trellis. The path with the largest likelihood would then be selected, and the information bits corresponding to that path would form the decoder output. Unfortunately, the number of paths for an L bit information sequence is 2^L; thus, this brute force decoding quickly becomes impractical as L increases.

With Viterbi decoding, it is possible to greatly reduce the effort required for maximum-likelihood decoding by taking advantage of the special structure of the code trellis. Referring to Fig. A-1, it is clear that the trellis assumes a fixed periodic structure after trellis depth K is reached.

The paths are said to have diverged at some state, and some depth j, if at depth $j+1$, its b information bits disagree. Later, paths can remerge after $b(K-1)$ consecutive identical information bits.

The maximum-likelihood sequence estimation problem is formally identical to the problem of finding the shortest route through a certain graph. The Viterbi algorithm then arises as a natural recursive solution. Consider a rate b/n convolutional code. Let $\mathbf{u}_0 \cdots \mathbf{u}_{k-1}\ \mathbf{u}_k\ \mathbf{u}_{k+1} \cdots$ denote the information

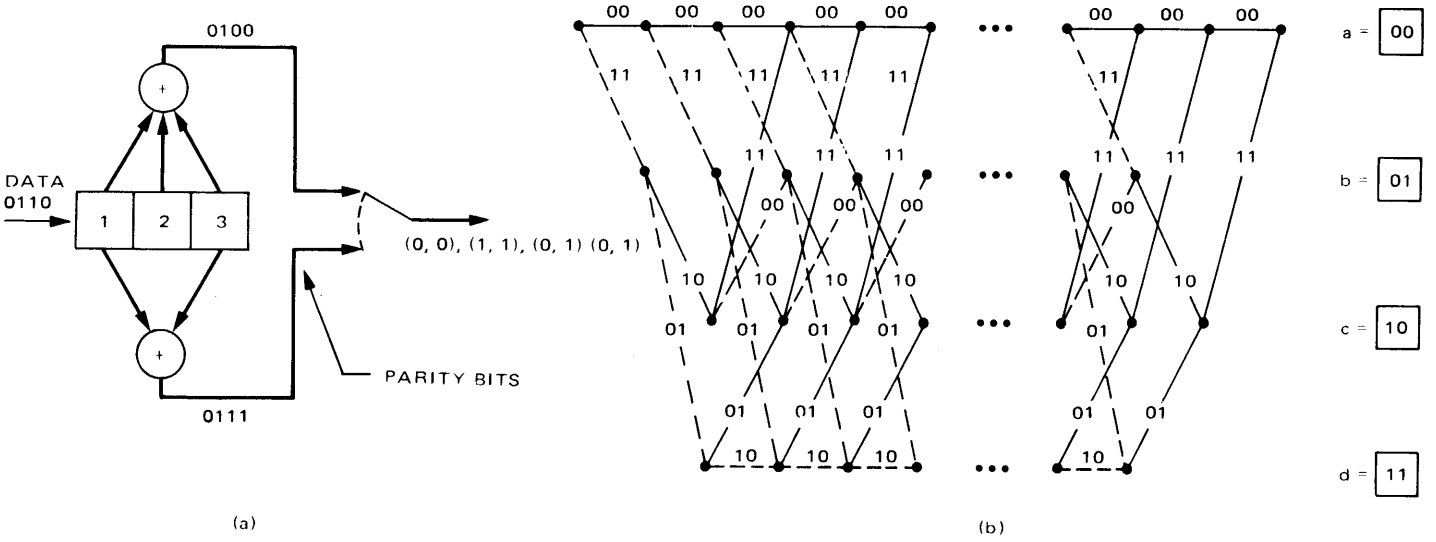


Fig. A-1. (a) Convolutional encoder; (b) trellis-code representation for encoder of (a)

bits to the encoder where $\mathbf{u}_k = (\mathbf{u}_{k1} \cdots \mathbf{u}_{kb})$. At time k define the encoder state as

$$s_k = \mathbf{u}_k \cdots \mathbf{u}_{k-k+1} \tag{A-1}$$

Given a sequence of observations, $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_L$; $\mathbf{y}_i = (\mathbf{y}_{i1} \cdots \mathbf{y}_{in})$, every path may be assigned a "length" proportional to metric $-\ln p(\mathbf{y}|\mathbf{s})$, where $p(\mathbf{y}|\mathbf{s})$ is the likelihood function and $\mathbf{s} = (s_0, \cdots s_L)$ is the state sequence associated with that path.

This will allow us to solve the problem of finding the state sequence for which $p(\mathbf{y}|\mathbf{s})$ is maximum, or equivalently, by finding the path whose length $-\ln p(\mathbf{y}|\mathbf{s})$ is minimum. Note that to every possible state sequence $\mathbf{s}$ there corresponds a unique path through the trellis, and vice versa. If the channel is memoryless, then

$$-\ln p(\mathbf{y}|\mathbf{s}) = \sum_{k=1}^{L} \lambda(s_k, s_{k-1}) \tag{A-2}$$

where

$$\lambda(s_k, s_{k-1}) = -\ln p(\mathbf{y}_k | s_k, s_{k-1}) = -\ln p(\mathbf{y}_k | \mathbf{x}_k)$$

is the branch "length" or metric. $\zeta_k(s_k, s_{k-1})$ denotes the transition from state s_{k-1} to s_k associated with branch symbols $\mathbf{x}_k = (x_{k1} \cdots x_{kn})$, which correspond to the information sequence

$$\mathbf{u}_k \cdots \mathbf{u}_{k-K} \tag{A-3}$$

Therefore, we can define the state transition as $\zeta_k(s_k, s_{k-1}) = \mathbf{u}_k \cdots \mathbf{u}_{k-K}$. We denote by $\mathbf{s}(s_k)$ a segment $(s_0, s_1, \cdots, s_k)$ consisting of the states up to time k of the state sequence $\mathbf{s}$. In the trellis, $\mathbf{s}(s_k)$ corresponds to a path segment starting at the state s_0 and terminating at state s_k. For any particular time k and state s_k, there will in general be several such path segments, each with some length

$$\lambda(\mathbf{s}(s_k)) = \sum_{i=1}^{k} \lambda(s_i, s_{i-1}) \tag{A-4}$$

The shortest such path segment is called the survivor, corresponding [5-31] to the state s_k, and is denoted $\hat{\mathbf{s}}(s_k)$. For any time $k > 0$, there are $2^{b(K-1)}$ survivors in all, one for each s_k.

Thus at any time k we need remember only the $2^{b(K-1)}$ survivors $\hat{s}(s_k)$ and their lengths $\Gamma(s_k) \triangleq \lambda(\hat{s}(s_k))$. To get to time $k+1$, we need only extend all time k survivors by one time unit, compute the lengths of the extended path segments, and for each state s_{k+1} select the shortest extended path segment terminating in s_{k+1} as the corresponding time $(k+1)$ survivor. Recursion proceeds indefinitely without the number of survivors ever exceeding $2^{b(K-1)}$.

We illustrate the algorithm for a simple four-state trellis covering 5 time units in Fig. A-2 [5-33]. Figure A-2(a) shows the complete trellis, with each branch labeled with a length for some received data. (In a real application, the lengths would be functions of the received data.) Figure A-2(b) shows the 5 recursive steps by which the algorithm determines the shortest path from the initial to the final state. At each stage only the 4 (or fewer) survivors are shown, along with their lengths. A formal statement of the Viterbi algorithm [5-31 and 5-33] follows.

Storage:

k (time index)

$\hat{s}(s_k)$; $1 \leqslant s_k \leqslant 2^{b(K-1)}$ (survivor terminating in s_k)

$\Gamma(s_k)$; $1 \leqslant s_k \leqslant 2^{b(K-1)}$ (survivor length)

Initialization:

$k = 0$

$\hat{s}(s_0) = s_0$

$\hat{s}(m)$ arbitrary; $m \neq s_0$

$\Gamma(s_0) = 0$

$\Gamma(m) = \infty$; $m \neq s_0$

Recursion:

Compute

$$\lambda(s_{k+1}, \hat{s}(s_k)) \triangleq \Gamma(s_k) + \lambda(s_{k+1}, s_k)$$

for all possible transitions $\zeta_{k+1}(s_{k+1}, s_k)$.

Find

$$\Gamma(s_{k+1}) = \min_{s_k} \lambda(s_{k+1}, \hat{s}(s_k)$$

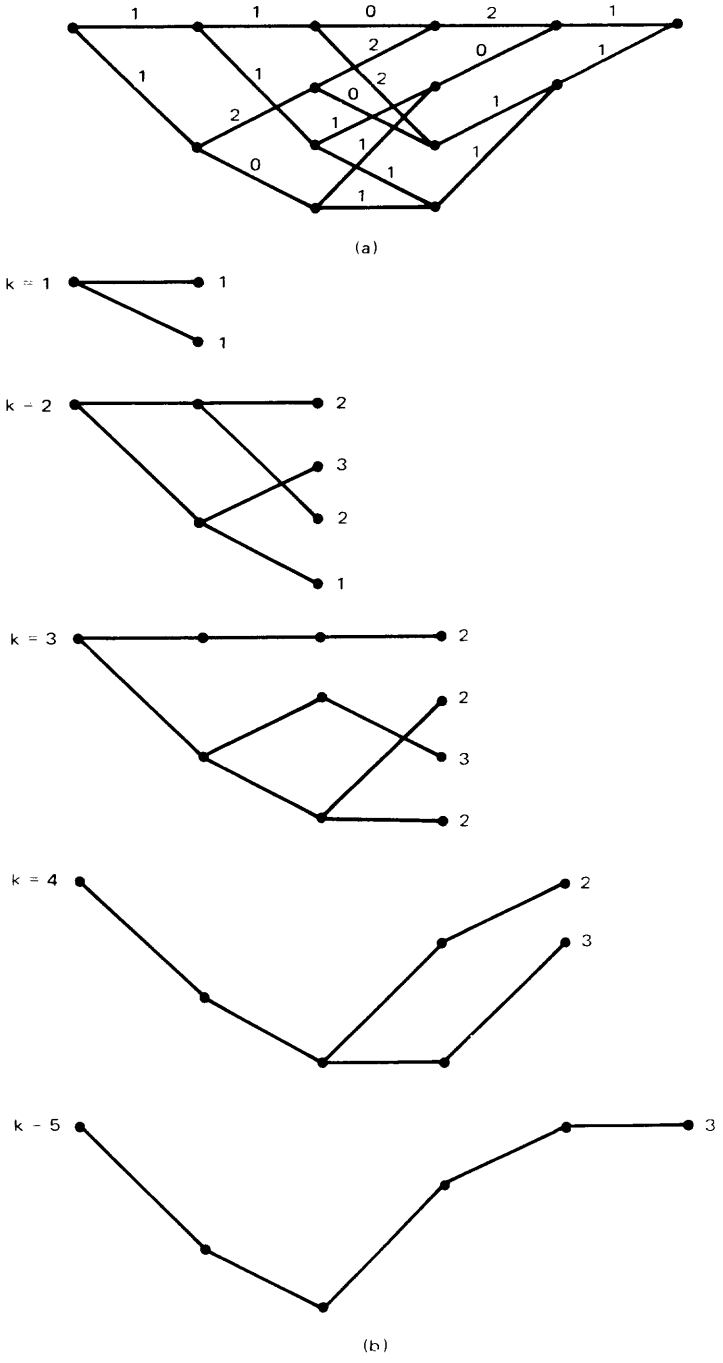


Fig. A-2. (a) Trellis labeled with branch lengths 1 for code of Fig. A-1 ($L = 5$); (b) recursive determination of the shortest path via the Viterbi algorithm

For each s_{k+1}; store $\Gamma(s_{k+1})$ and the corresponding survivor $\hat{\mathbf{s}}(s_{k+1})$; set k to $k+1$ and repeat until $k = L$.

With finite state sequence **s**, the algorithm terminates at time L with the shortest complete path stored as the survivor $\hat{\mathbf{s}}(s_L)$.

In summary, a Viterbi decoder calculates the length (likelihood or metric) of each of the 2^b paths entering a given state and, except for the most likely (shortest length) path, eliminates from further consideration all paths that lead to that state. This is done for each of $2^{b(K-1)}$ states at a given trellis depth; after each decoding operation only one path remains leading to each state. The decoder then moves one level deeper into the trellis and repeats the process.

The great advantage of the Viterbi maximum-likelihood decoder is that the number of decoder operations performed in decoding L bits is only $L2^{b(K-1)}$, which is linear in L. Of course, Viterbi decoding as a practical technique is limited to relatively short-constraint-length codes due to the exponential dependence of decoder operations, per decoded bit, on K.

Appendix B

Sequential Decoding

The complexity of the Viterbi algorithm is an exponential function of the code's constraint length K. A decoding algorithm that will work on convolutional codes with very large values of K is needed. There is a class of such algorithms, called *sequential decoding algorithms.* They are not quite as good as maximum-likelihood algorithms for a fixed code, but that defect is largely compensated for by the fact that they can be used to decode some codes with very large K.

The key to understanding sequential decoding algorithms is the *tree diagram.* In a shift register encoder for a b/n convolutional code, b information bits entering the decoder cause n encoded bits to leave. Thus it is possible to view the encoding process conceptually as a walk through a tree, as depicted in Fig. B-1 ($b = 1$, $n = 2$).

The encoder begins at the INITIAL node and after j inputs it will be at some node in the tree at depth j. If the next input is "0", it moves to depth $j + 1$ along the branch in the upward direction; if it is "1", in the downward direction. For example, if the input is $01100\cdots$, the encoder will follow the path (solid line) indicated in Fig. B-1. In general, each branch of the tree will be labeled with the n bits the encoder will output as it travels the branch. The tree in Fig. B-1 apparently extends to infinite depth, but for any concrete realization of a code the tree will be finite. If the Lth truncation of the code is being used, the tree will terminate at depth $L + K - 1$; and beyond depth $L - 1$ there will be no bifurcation of the paths, since the last $K - 1$ inputs to the encoder will always be "0". Figure B-1 is the completely labeled tree corresponding to the code of Fig. A-1. A code sequence is a path through the tree. The encoding operation can be viewed as a process in which the encoder traces a particular path through the code tree according to the instructions of the input message blocks. Conversely, the decoding operation may be regarded as a process in which the decoder, based on the received sequence and the channel statistics, attempts to retrace the transmitted path through the code tree.

From this viewpoint, each transmitted sequence corresponds to a path through the tree, beginning at the INITIAL node and ending at depth $L + K - 1$. The received sequence will be a noisy version of this path, and the decoder's job will be to guess which path the encoder actually took. Since there will be a total of 2^L possible paths, if L is large it will not in general be possible to compare the received message to each of these paths. The approach taken to this problem by sequential decoding algorithms is to

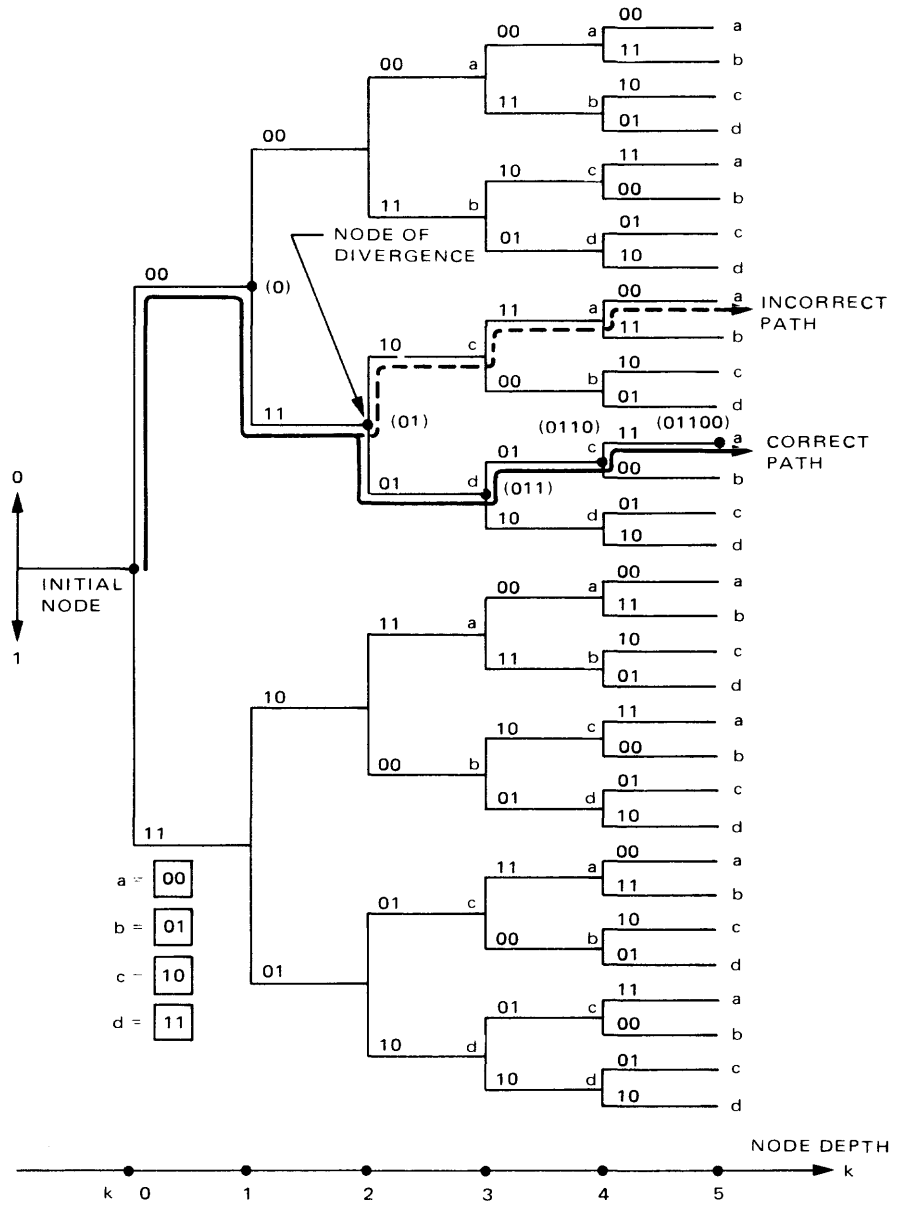


Fig. B-1. A binary code tree for the convolutional code of Fig. A-1

explore a very small subset of the possible paths. The idea of sequential decoding is to program the decoder to act much as a driver who occasionally makes a wrong choice at a fork in the road, but quickly discovers his error, goes back, and tries the other.

Sequential decoding is a procedure for systematically searching through a code tree and using received information as a guide, with the objective of eventually tracing out the path representing the actually transmitted information sequence. The two best-known sequential decoding algorithms are the stack algorithm [5-59 and 5-60] and the Fano algorithm [5-61].

Most sequential decoder implementations to date have used some modification of the Fano algorithm. Briefly, the Fano algorithm operates as follows. Starting at the INITIAL node in the code tree, a path is traced through the tree by moving ahead one node at a time. At each node encountered, the decoder evaluates a branch metric for each branch stemming from that node. The branch metric is a function of the transition probabilities between the received symbols and the transmitted symbols along the hypothesized branch. The decoder will initially choose the branch with the largest metric value (corresponding to the closest fit to the received symbols). The metric is then added to a path metric, which is the running sum of branch metrics along the path presently being followed. This path metric is called the Fano metric [5-30] and is given by

$$\Gamma(\mathbf{x}(k), \mathbf{y}(k)) = \sum_{i=1}^{k} \lambda_i(\mathbf{x}_i, \mathbf{y}_i) \tag{B-1}$$

where $\lambda_i(\mathbf{x}_i, \mathbf{y}_i)$ are branch metrics and are given by

$$\lambda_i(\mathbf{x}_i, \mathbf{y}_i) = \sum_{j=1}^{n} \left[\log_2 \frac{p(y_{ij}|x_{ij})}{p(y_{ij})} - r \right] \tag{B-2}$$

$\mathbf{x}(k) \triangleq (\mathbf{x}_1, \cdots, \mathbf{x}_k)$ is a k branch path stemming from the initial node of the code tree, and $\mathbf{y}(k) \triangleq (\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_k)$ is the first k block segment of the received sequence $\mathbf{y}$. The $p(y_{ij}|x_{ij})$ are the channel transition probabilities. The $p(y_{ij})$ are the probability distributions on the channel output alphabet, and r is the code rate. The x_{ij}'s and y_{ij}'s are the channel input and output symbols, respectively, corresponding to the ith branch.

Along with the path metric, the decoder keeps track of the running threshold T. As long as the path metric keeps increasing, the decoder assumes it is on the right track and keeps moving forward, raising T to lie within a fixed constant Δ, below the path metric. If, on the other hand, the path metric

decreases at a particular node, such that it becomes less than T, the decoder assumes it may have made a mistake and backs up. It will then systematically search nodes at which the path metric is greater than T until it finds a path that starts increasing again, or until it exhausts all nodes lying above T. At this point it is forced to lower T, by Δ, and search again. Eventually the decoder will find a path that appears to have an increasing path metric, and will penetrate sufficiently deeply into the tree that there is a high probability the first few branches followed will be correct. These branches will not be returned to by the decoder in a backward search. At this point, the information bits corresponding to these branches can be considered decoded, and the decoder may erase received data pertaining to these branches.

In summary, at every stage the decoder is located at some node in the tree. From this node the decoder looks forward (i.e., deeper) into the tree. If it sees a node it likes, it moves forward to that node. If not, it moves backward and then tries to move forward along another branch. It decides whether or not it likes a given node by comparing the metric of that node to a running threshold T. The decoder stops as soon as it reaches a terminal node (say, one at depth $L + m$) and outputs the information sequence corresponding to the branch ending at that node. If the channel is quiet, the decoder can trace the correct path through the termination node. However, when the channel gets noisy, it may happen that the decoder is fooled and traces an incorrect path. In this case we have undetected errors made by the decoder. Correct and incorrect paths and nodes of divergence of two paths are shown in Fig. B-1. For example, if $0110\cdots$ is the transmitted sequence and $0100\cdots$ is the decoded sequence (see Fig. B-1) then we have an undetected error at position 3.

The decoder consists of a buffer to hold quantized received samples, a copy of the encoder, and equipment to measure the merit of the encoder's output relative to the received symbol sequence. For any given sequence of information digits $\mathbf{u}$, the decoder can run $\mathbf{u}$ through the *replica encoder*, find the associated $\mathbf{x}$, and calculate $\Gamma(\mathbf{x}, \mathbf{y})$, which is the value of the node associated with $\mathbf{u}$. Now we reexamine the three kinds of moves from one node to another: *forward*, *sideways*, and *backward*. On a *forward move*, the decoder goes one branch to the right in the received value tree from the previously hypothesized node. Instrumentally, this corresponds to shifting the replica encoder shift register to the right and inserting "b" new hypothesized information bits at the left. The new value for the metric can be found as

$$\Gamma_{k+1} \triangleq \Gamma(\mathbf{x}(k+1), \mathbf{y}(k+1)) = \Gamma(\mathbf{x}(k), \mathbf{y}(k)) + \lambda_{k+1}(\mathbf{x}_{k+1}, \mathbf{y}_{k+1}) \tag{B-3}$$

The digits involved in this calculation are simply the n channel input digits coming out of the replica encoder. A *sideways move* is a move from one node to another node differing only in the last branch of the tree. Instrumentally, this corresponds to changing the "b" left-most bits in the replica encoder shift register. Again, the change in value from one node to the other is determined by the change in the last n channel input digits. A *backward move* is a move one branch to the left in the received value tree. Instrumentally, this corresponds to shifting the encoder shift register to the left and reinserting the last "b" bits to pass out of the shift register on the right. The new value is calculated from the old value by subtracting off the last term in the summation over i. Thus, for each possible move, the change in value from the old node to the new is a function only of the last n hypothesized channel inputs. The typical behavior of Γ_k for the correct path and incorrect paths is shown in Fig. B-2. This behavior of Γ_k will be used to indicate whether the decoder is tracing the correct path or an incorrect path. As the decoder traces a path $\mathbf{x}(k)$ through the code tree one branch at a time, it computes the Fano metric Γ_k of $\mathbf{x}(k)$. When Γ_k increases as k increases, $\mathbf{x}(k)$ is presumably the correct path. However, when Γ_k starts to decrease with increasing k, the decoder is presumably tracing an incorrect path. The node of divergence is where an incorrect path branches off from the correct path.

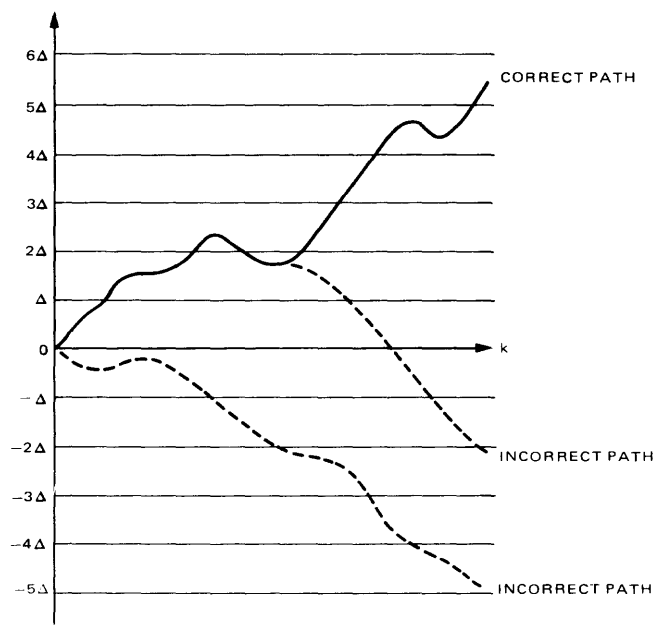


Fig. B-2. Typical behavior of Γ_k for correct path and incorrect paths

The Fano decoding algorithm is then a set of rules by which the decoder is able both to recognize whether it is tracing the correct path or an incorrect path and to initiate a systematic search for the node of divergence when the decoder senses the possibility that it is tracing an incorrect path. For each node of the code tree at level k we assign a metric value that is equal to the Fano metric Γ_k of the path leading to the node. Obviously, the metric value Γ_0 of the initial node is zero.

The essential features of the Fano decoding algorithm [5-2, 5-12, and 5-62] are described below:

(1) The decoder begins searching for the correct path from the initial node of the code tree. It moves one branch at a time.

(2) The decoder must first order the 2^b branches (for rate $1/n$ convolutional codes we have 2 branches) stemming from node at depth k in accordance with their branch metrics λ_k corresponding to the kth block of the received sequence. The branch with largest metric is first, the branch with the next largest metric is second and so on. Note that for large b or for a saving in computation we may choose not to order the branch metrics. In this case the Fano algorithm will be slightly changed from its original version.

(3) Suppose that the decoder has traced a path $\mathbf{x}(k)$ up to a node N_1 at depth k, as shown in Fig. B-3. From node N_1, the decoder either moves along a branch forward to a node N_2 at depth $(k+1)$ or steps back to the node N_3 at depth $(k-1)$, which was accepted previously.

(4) As the decoder moves from one node to another, it maintains a running threshold $T = l\Delta$, where l is an integer and Δ is the threshold spacing, which is a design parameter. The running threshold at the initial node is set to zero.

(5) Suppose that the decoder is at node N_1 at depth k and that the current running threshold is T_k. First, the decoder attempts to make a forward movement. It *looks forward* (computes the next node metric and compares it with the threshold) and considers the node N_2 at depth $(k+1)$, which is connected to node N_1 by the first branch stemming from node N_1. The metric value of node N_2, denoted $\Gamma_{k+1}(N_2)$, is computed and compared with the current running threshold T_k. If $\Gamma_{k+1}(N_2) \geq T_k$, the decoder *moves forward* (increasing k) and accepts node N_2. If this is the first time that the decoder has ever reached node N_2 (this is true if $\Gamma_k(N_1) < T_k + \Delta$), the running threshold is raised to $T_{k+1} = T_k + j\Delta$, where j is an integer such that $T_k + j\Delta \leqslant \Gamma_{k+1} < T_k + (j+1)\Delta$. We say that the running threshold is *tightened*. In the same manner, the decoder attempts to move forward to a node at depth $(k+2)$ from node N_2. If $\Gamma_{k+1}(N_2) < T_k$, the

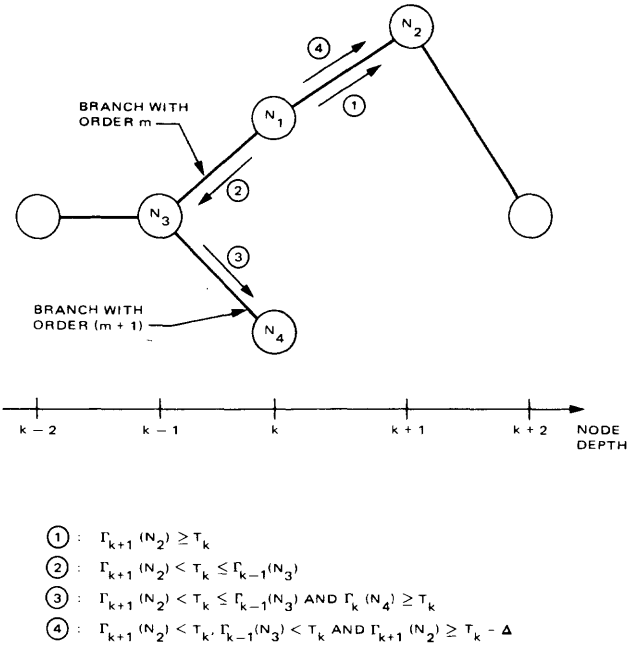


Fig. B-3. The motion of a decoder from the Node N_1

decoder *looks back* (computes the previous node metric and compares it with the threshold) from node N_1 to the node N_3 at depth $k-1$, from which the decoder moved to N_1 in a previous forward movement. The metric value of node N_3, $\Gamma_{k-1}(N_3)$, is compared to T_k. If $\Gamma_{k-1}(N_3) > T_k$, the decoder *moves back* (decreasing k) to node N_3. From node N_3, the decoder looks forward to the node N_4 at depth k (*looks sideways*), which is connected to N_3 by a branch whose order is one higher than the order of the branch that connects node N_1 to node N_3. If $\Gamma_k(N_4) > T_k$, the decoder accepts node N_4 (*moves sideways*). From node N_4, the decoder attempts to move forward again.

(6) If both $\Gamma_{k+1}(N_2)$ and $\Gamma_{k-1}(N_3)$ are less than the current running threshold T_k at node N_1, the decoder can neither move forward nor *backward*. In this event, the running threshold T_k is reduced by Δ. Thus, the running threshold at node N_1 becomes $T_k - \Delta$. We say that the running threshold is *loosened*. With this reduced running threshold, the decoder looks forward to node N_2 again.

The running threshold is tightened when and only when the decoder moves forward to a node that has never been examined before. The

running threshold is never tightened during backward motion. The running threshold is reduced when and only when the decoder is trapped at a node that can move neither forward nor backward. Before tightening the threshold after accepting N_2, we have to check that the metric of the previous node satisfies $\Gamma_k(N_1) < T_k + \Delta$. This corresponds to checking that we are visiting the current node N_2 for the first time.

The purpose of the restriction $\Gamma_k(N_1) < T_k + \Delta$ is to prevent the threshold from being raised again on one of these nodes that has been previously hypothesized. In fact, if this restriction were not imposed, the decoder would quickly find itself in a loop, hypothesizing the same node again with the same threshold.

The detailed movement of the decoder is illustrated by the flow chart shown in Fig. B-4.

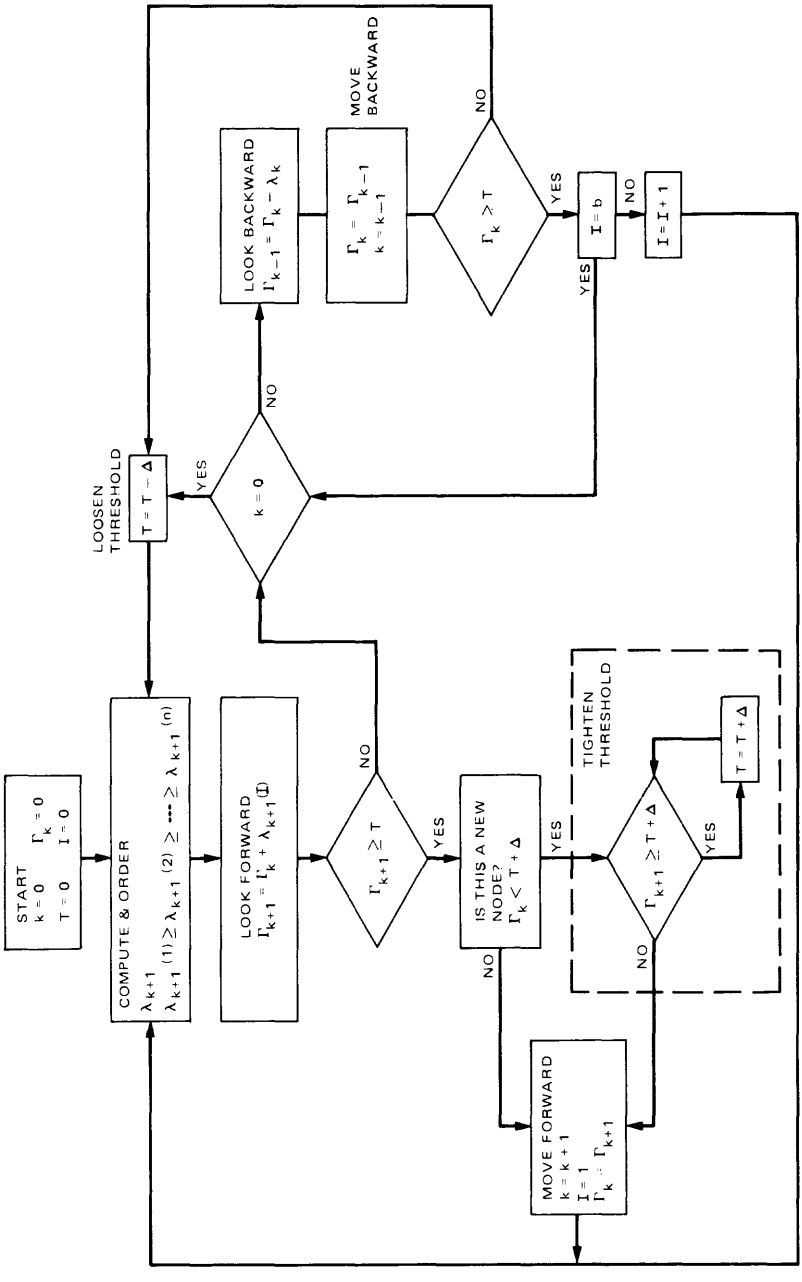


Fig. B-4. Flow diagram of Fano sequential decoding algorithm

Chapter 6
Command System

Nathan A. Burow and Man K. Tam

6.1 Introduction

The Command System (CMD) provides the means by which a project controls the activities of its spacecraft from the earth. This chapter presents an overview of the Multimission Command (MMC) System. The major components within the MMC System are discussed, with the emphasis on the telecommunication-related implementations. Two versions of the spacecraft command detection system – the Viking heritage command detector and the NASA standard command detector – are discussed in detail. The former prevails in the existing flight projects and the latter will likely be adopted by the missions of the near future. The preparation of Design Control Tables for the control of command link performance between Deep Space Stations (DSSs) and the spacecraft is also discussed.

6.2 Multimission Command System

The Multimission Command System [6-1] extends from the point at which validated commands are entered at the Mission Operations Center (MOC) to the point on the spacecraft where commands, after error detection and correction, are distributed to spacecraft subsystems. Thus, the MMC System includes the

command functions performed by the MOC, the Deep Space Network (DSN), and spacecraft command detector and command decoder. The MMC System does not include the project-dependent Mission Sequence System, where commands are formulated and validated, nor does it include the spacecraft subsystems which intepret and respond to the commands. The Command System, thus delimited, performs a multi-mission command delivery and accounting function. A block diagram showing the boundaries of the MMC System is given in Fig. 6-1.

The MCC System employs a "store-and-forward" method for delivery of command sequences to a spacecraft. The ability to store commands at various points in the command delivery system eliminates time criticality from the transport of commands. Large files of spacecraft commands may be generated by a project, sent by the MOC to the DSS, and correct receipt verified before a command is to be "forwarded" to the spacecraft. This allows a project to prepare large files of spacecraft commands in advance and then to forward several files to the DSS at the beginning of a spacecraft track. The same store-and-forward concept is employed between the DSS and the spacecraft. Commands may be radiated from the DSS to the spacecraft, correct receipt verified, and the commands stored onboard the spacecraft for later execution. It should be noted that this store-and-forward mode of operation does not preclude near-real-time commanding of the spacecraft. The MCC store-and-forward operations are represented functionally in Fig. 6-2.

For discussion purposes, the Command Systems may be divided into a ground segment and a spacecraft segment. The functional operations for these two portions are briefly described below.

6.2.1 Ground Portion of Command System Operations

On the ground side of the Command System, command sequences generated by the projects for the different spacecraft are stored at the MOC. Commands for a particular spacecraft are selected from the command files, formatted into

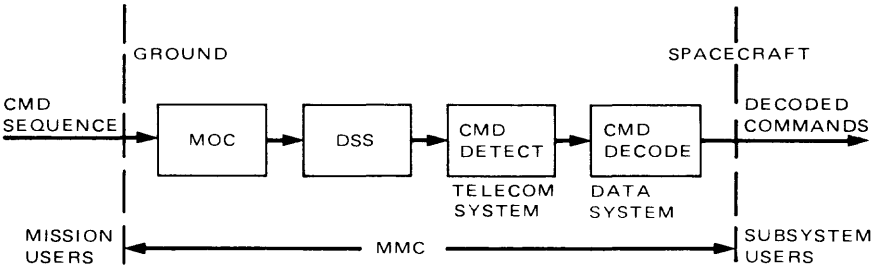


Fig. 6-1. MMC system uplink block diagram

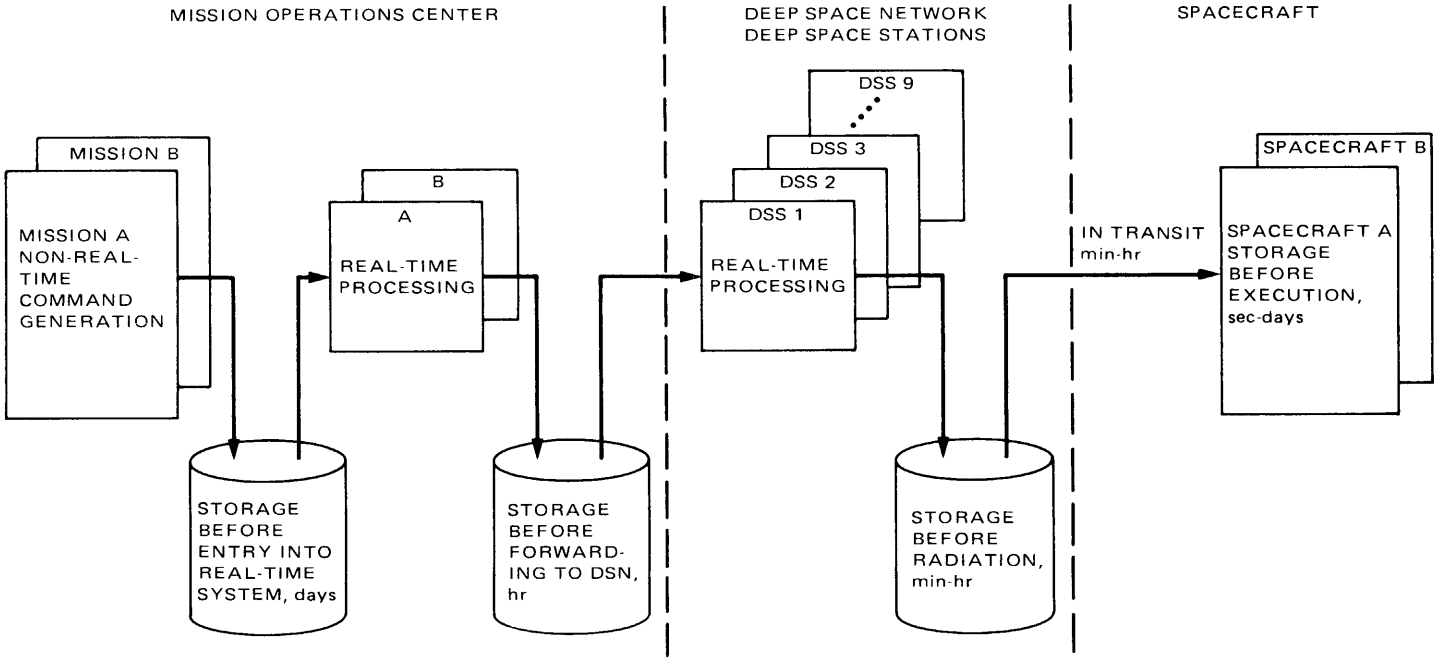


Fig. 6-2. MMC system "store-and-forward" operations (with typical storage times)

messages, and stored for transmittal, via the Ground Communication Facility (GCF), to a specified DSS. The command messages contain the commands to be radiated to the spacecraft, timing instructions to the DSS for the radiation of the commands, and error detection/correction coding. The DSS verifies correct receipt of the command files or requests retransmission of files containing errors. The spacecraft commands are extracted from the messages by the DSS Command Processor Assembly (CPA) and stored in the CPA until the time specified for their radiation to the spacecraft. When a command is to be radiated from the DSS to the spacecraft it is placed in a queue in the CPA and sent to the DSS Command Modulator Assembly (CMA). The CMA modulates the command bits onto a subcarrier which then modulates the DSS transmitter for radiation of the command to the spacecraft.

6.2.2 Spacecraft Portion of Command System Operations

The spacecraft portion of the Command System consists of two subassemblies: the command detector and the command decoder. Upon receipt of the composite command signal from the spacecraft receiver, the command detector removes the modulated subcarrier, detects the binary command bit stream, and presents the bit stream to the command decoder. The decoder then determines the validity of its input and, if it is a valid command, decodes it and passes the decoded command to the addressed spacecraft subsystem. The decoder, therefore, depends only on command format and tends to be oriented toward a specific mission. The detector, on the other hand, is dependent on the uplink modulation scheme and tends to be less mission-dependent.

6.3 DSN Command System Capabilities

The capabilities of the DSN Command System that affect spacecraft designs are described in this section. More detailed information is contained in [6-2] through [6-4]. The DSN Command System provides the capability for commanding one spacecraft from each Deep Space Station. The functions performed by the DSN Command System include the following:

(1) Establishing the DSN configuration for the specified spacecraft.

(2) Receiving the command data from the MOC and storing it at the DSS.

(3) Queuing command data to be radiated to the spacecraft.

(4) Radiating the command data to the spacecraft.

(5) Monitoring and reporting DSN Command System status and events.

A detailed diagram of the DSN Command System (Mark III-80) is presented in Fig. 6-3.

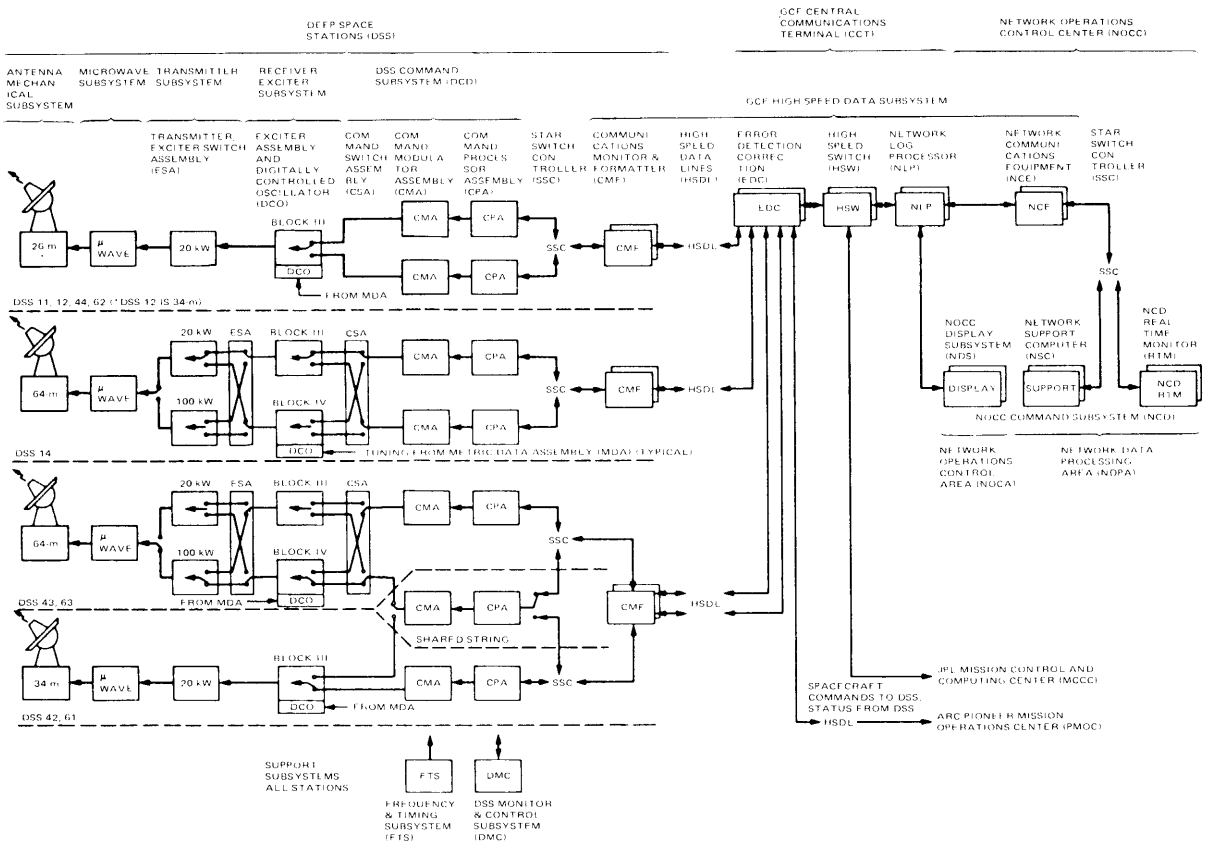


Fig. 6-3. DSN Command System Mark III-80

The DSN Command System at each DSS consists of a Command Processor Assembly (CPA) for software functions, a Command Modulator Assembly (CMA) for generating the command waveform, and the exciters, transmitters and antennas for RF carrier modulation and transmission. Note the redundancy of these units at each DSS in Fig. 6-3. The DSN Command System produces a pulse code modulation (PCM) non-return-to-zero (NRZ-L) data waveform. This waveform is then used to bi-phase modulate a subcarrier in the phase-shift-keyed (PSK) mode, or to switch between two subcarriers in the frequency-shift-keyed (FSK) mode. Data rates are variable from 1 to 128 bits per second. An upgrade to 2000 bps is planned [6-5]. PSK data rates can be either noncoherent or coherent with the subcarrier frequency. PSK and FSK subcarrier frequencies may be selected in the range of 100 Hz to 16 kHz, with a resolution of 0.1 Hz. PSK subcarriers may be either sinewave or squarewave. For FSK modulation, only sinewave subcarriers are available. The RF carrier may be modulated at modulation index angles from 0.1 to 1.5 radians peak, for both sinewave and squarewave subcarriers. Complete performance specification for the CMA is presented in [6-6].

Command carrier power suppression and data power suppression as functions of modulation index angle are:

Sinewave subcarrier:

$$\frac{P_C}{P_T}\ (\text{dB}) = 10 \log [J_0^2 (\theta_D)] \tag{6.3-1}$$

$$\frac{P_D}{P_T}\ (\text{dB}) = 10 \log [2J_1^2 (\theta_D)] \quad \text{(upper and lower sidebands)} \tag{6.3-2}$$

Squarewave subcarrier:

$$\frac{P_C}{P_T}\ (\text{dB}) = 10 \log [\cos^2 (\theta_D)] \tag{6.3-3}$$

$$\frac{P_D}{P_T}\ (\text{dB}) = 10 \log [\sin^2 (\theta_D)] \quad \text{(all sidebands)} \tag{6.3-4}$$

where

θ_D = data modulation index, radians peak

P_T = total power

P_C = carrier power

P_D = data power

J_0 = zero-order Bessel function

J_1 = first-order Bessel function

The nominal total RF power output available for single-channel command modulation is 20 kW at the output of the transmitter. At the 64-meter antenna stations (DSS 14, 43, and 63), an additional transmitter is available for emergency operation having a nominal carrier output power of 100 kW. A more complete definition of existing and planned DSN command capabilities is contained in [6-4] and [6-5], respectively.

6.4 Spacecraft Command System Capabilities and Performance

The capabilities and performance of the spacecraft command system are, in general, characterized by the choice of the command detector configuration and the link design. The following specifications are important factors for the hardware design of the command detector:

(1) Maximum bit rate.

(2) Minimum preamble to acquire synchronization.

(3) Requirement for external data ambiguity resolution.

(4) Modulation characteristics.

(5) Subcarrier characteristics.

(6) Data transition dependency.

The performance of the command link is specified by the following parameters:

(1) Required per bit energy-to-noise spectral density ratio (ST_B/N_0) to achieve a given bit error rate (BER).

(2) Required duration (bit-times) to achieve bit synchronization, with an associated false acquisition probability at a given ST_B/N_0.

(3) Error statistics for in-lock and out-of-lock indications.

Past spacecraft missions have employed the following types of command detectors: a single-channel analog command detector, a two-channel PSK command detector and a single-channel digital command detector (Viking Heritage). A new single-channel digital command detector (NASA standard) has been developed and will probably be used on new missions of the near future. Table 6-1 shows the general capabilities and performance of these four command detectors and indicates their usage in terms of past and expected future space missions. At present, only the two digital command detectors are in use. The operation of these two command detectors is described in detail below.

6.4.1 Viking Heritage Command Detector Operation

The Viking Heritage command detector was originally developed for the Viking Mission, and was later adopted by the Voyager and Galileo Projects. It accepts a 512-Hz squarewave subcarrier modulated by data bits from the RF receiver output. The bit rate is defined by selecting the number of subcarrier cycles per bit (denoted by M), where M is constrained to be an even integer. Command data bit rates in the range of 1 bps through 256 bps may be selected (hardwired prior to launch) based on use of the 512 Hz squarewave subcarrier and a requirement to have at least two subcarrier cycles per bit.

The Viking Heritage command detector first establishes two levels of synchronization, i.e., subcarrier sync and bit sync, before data bits are detected. To accomplish this synchronization, transmitted commands are prefixed by a length of unmodulated subcarrier, and then by a length of subcarrier modulated by bit sync only. The Viking Heritage command detector first establishes subcarrier sync by comparing the unmodulated subcarrier with a reference signal. Then a replica of the bit-sync modulated subcarrier is used as a reference to establish bit sync. After establishing these references the system can detect data. All of the functions of synchronization and data detection are performed by the process of cross-correlation. One unique design feature of this system is the use of a method of discrete-step phase tracking which compensates for small frequency offsets due to Doppler shifts and/or oscillator instabilities. This method of phase tracking is performed by accumulating, for an interval of one data bit period, a sum of samples taken near the transitions of the signal. The polarity of the sum determines the direction of the adjustment made to the demodulation reference. The detailed performance analysis for the subcarrier tracking system is presented in [6-7].

The block diagram for the Viking Heritage command detector is presented in Fig. 6-4. The system requires a lock detector to ensure the presence of the

Table 6-1. Capabilities and performance summary of four spacecraft command detectors

Parameter	2-channel PN detector	Single-channel analog detector	Single-channel digital detector (Viking Heritage)	Single-channel digital detector (NASA standard)
Modulation format	PSK data PCM sync	PSK	PCM	PCM
Data rate	1 bps	4 bps	Selectable up to 256 bps	Variable in binary steps up to 2 kbps
Subcarrier waveform	Sinewave	Sinewave	512-Hz squarewave	16-kHz sinewave
Data transition dependency	None	Dependent	None	During acquisition
Multiple data rate capability	None	None	1 to 256 bps	7.8 to 2000 bps
External ambiguity resolving requirement	No	No	No	Yes
Command ST_B/N_0 at BER = 10^{-5}	11.7 dB ($P_{sync} = 2P_{data}$)	10.9 dB	10.5 dB	10.5 dB
Maximum acquisition time at threshold ST_B/N_0	511 T_B	360 T_B	90 T_B	132 T_B
Lock detector probabilities:				
P (in/noise)	$<1.5 \times 10^{-2}$	$<10^{-3}$	$<10^{-5}$	$<10^{-4}$
P (out/signal)	$<1.6 \times 10^{-3}$	$<2 \times 10^{-6}$	$<10^{-5}$	$<2.5 \times 10^{-9}$
Past and expected missions	Mariner Venus '62, Mars '69, Mars '71, MVM '73	Helios	Viking Orbiter (VO'75), Voyager, Galileo	All missions using NASA standard deep space transponder

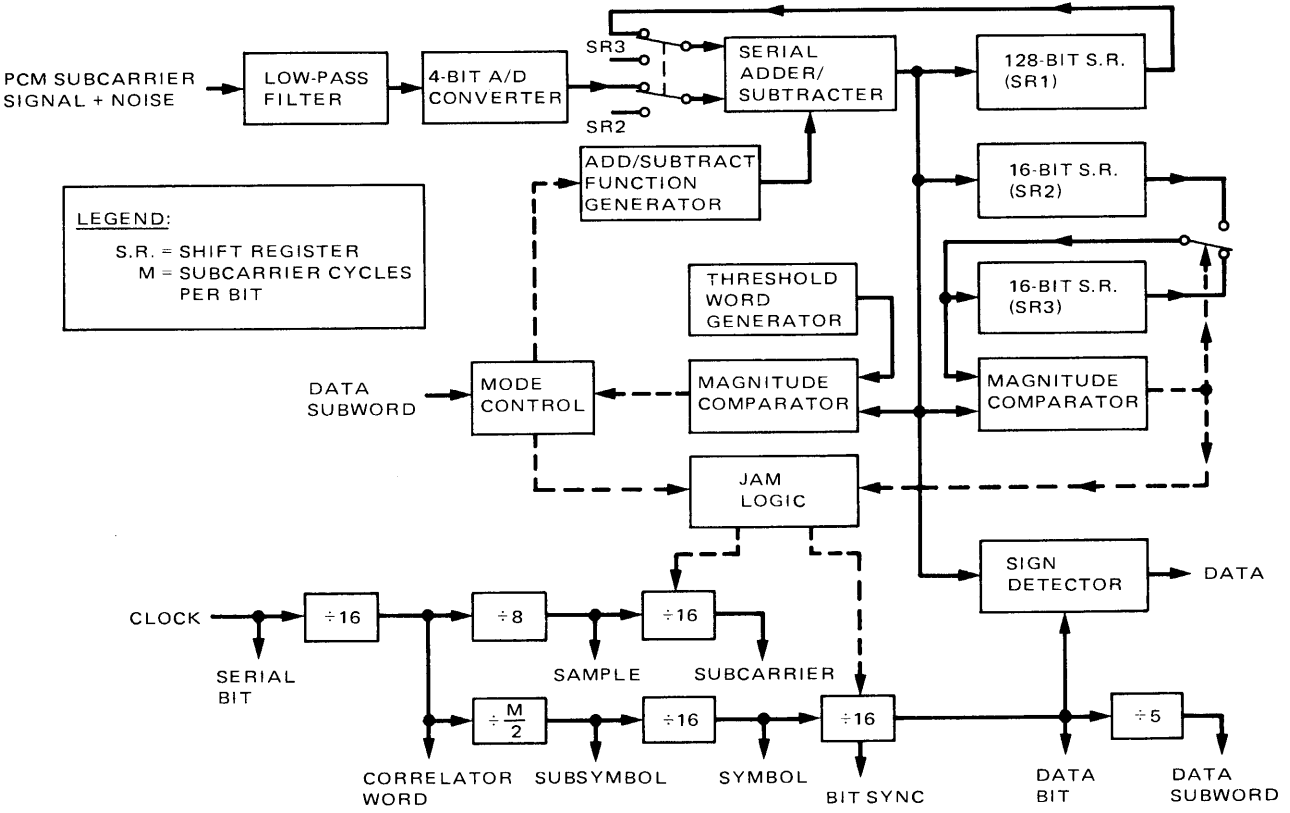


Fig. 6-4. Viking Heritage command detector block diagram

proper signal. The in-lock condition is determined by comparing the sum of the absolute values of five consecutive bits against a preset threshold value.

6.4.1.1 Detection algorithm. The Viking Heritage command detector is basically a correlation device. During subcarrier acquisition the correlator estimates the peak point on the correlation curve by cross-correlating the input signal with internally generated, phase-shifted estimates of the subcarrier. The estimate whose phase most closely agrees with that of the input signal is chosen as the subcarrier demodulation reference. The command input waveform is sampled 16 times per subcarrier cycle and correlated with 16 reference waveforms which are offset with respect to each other by increments of one sample period. Thus the digital correlation curve seen by the system is the staircase function shown in Fig. 6-5b, which approximates from the continuous correlation curve of Fig. 6-5a.

Subcarrier and bit sync acquisitions are performed in sequence according to the basic system algorithm shown in Fig. 6-6. In its initial state, the system simultaneously correlates 16 phase-shifted replicas of the subcarrier with the incoming signal for 5 bit periods. A correlation period of 5 bit times is required to make each of the threshold tests reliable. At the end of the 5 bit times, if the largest correlation value exceeds a preset threshold value, the system advances to mode 2. Othewise, the system stays in mode 1 and repeats the subcarrier correlations.

In mode 2 the correlation against the subcarrier reference is repeated, and threshold is again checked. If the check is confirmed the system enters mode 3, where the replica having the largest correlation value is chosen as the subcarrier estimate and the subcarrier clock is "jammed" to correspond to this phase. The system then leaves mode 3 and automatically enters mode 4, which is a wait period of 50 bits. During mode 4 the system tracks subcarrier phase offsets between the transmitted signal and its internally generated replica caused by Doppler shifts and/or oscillator instabilities. The 50 bits are needed to accomplish this tracking with a small steady-state average phase error. Entry into mode 5 is automatic after completion of the 50 bits.

In mode 5 the system attempts to determine if bit sync is present. During this mode the system continues the subcarrier threshold tests. As long as threshold tests are successful, the system assumes that bit sync is not present on the incoming signal and continues tracking the subcarrier. If a test fails, the system assumes bit sync is on and enters mode 6.

In mode 6 the system attempts to establish bit synchronization between the local bit sync estimates and the received command signal. During this mode the incoming signal is correlated for 5 bit periods against the 16 phase-shifted sequences of the bit sync squarewave, which is synchronous in a preset way with

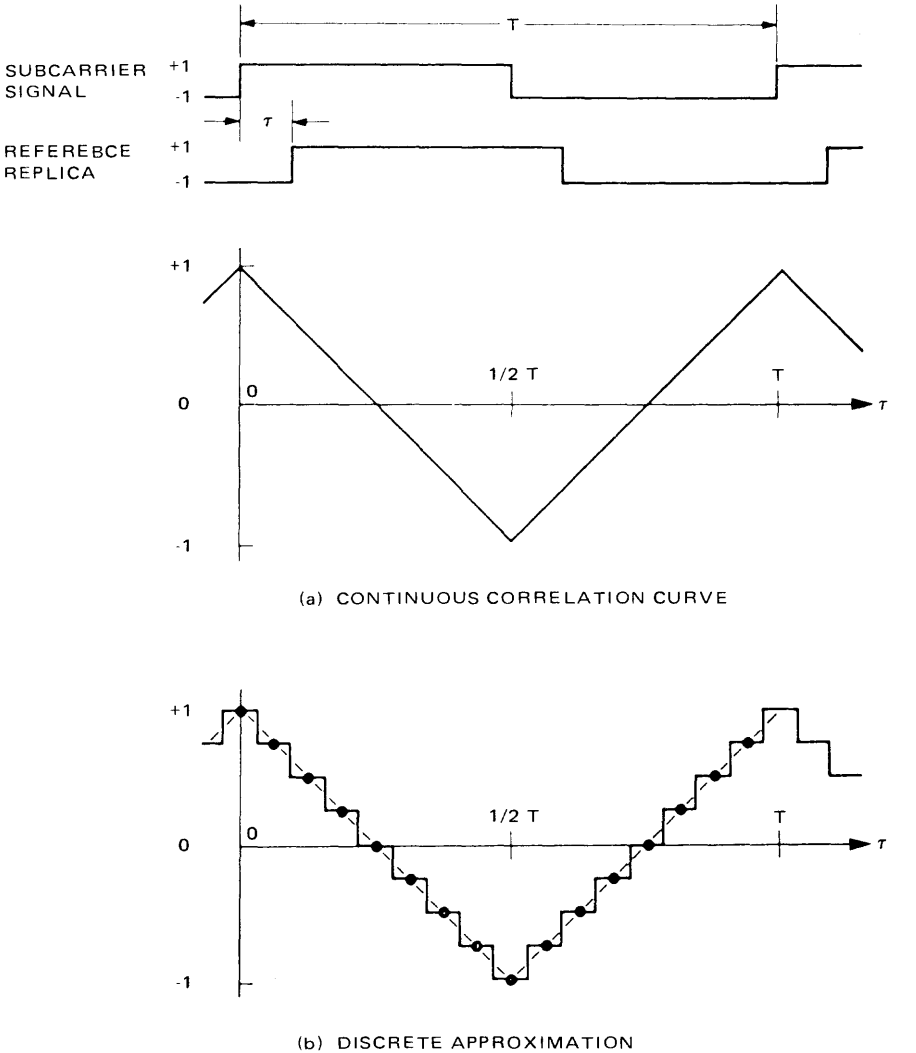


Fig. 6-5. Correlation properties of the subcarrier signal for the Viking Heritage command detector

the data bits. If threshold is not exceeded, the system goes to mode 10; otherwise it goes to mode 7, where the bit sync clock is jammed to its proper phase in the same manner as for subcarrier phase selection. The system then enters mode 8, where it waits for the start of the next data bit and then proceeds to mode 9. In mode 9 the system tracks both subcarrier and bit sync phase offsets for 10 bits. After 10 bit times all synchronization is complete and the system begins to detect data bits. During mode 9 the system tracks both subcarrier and

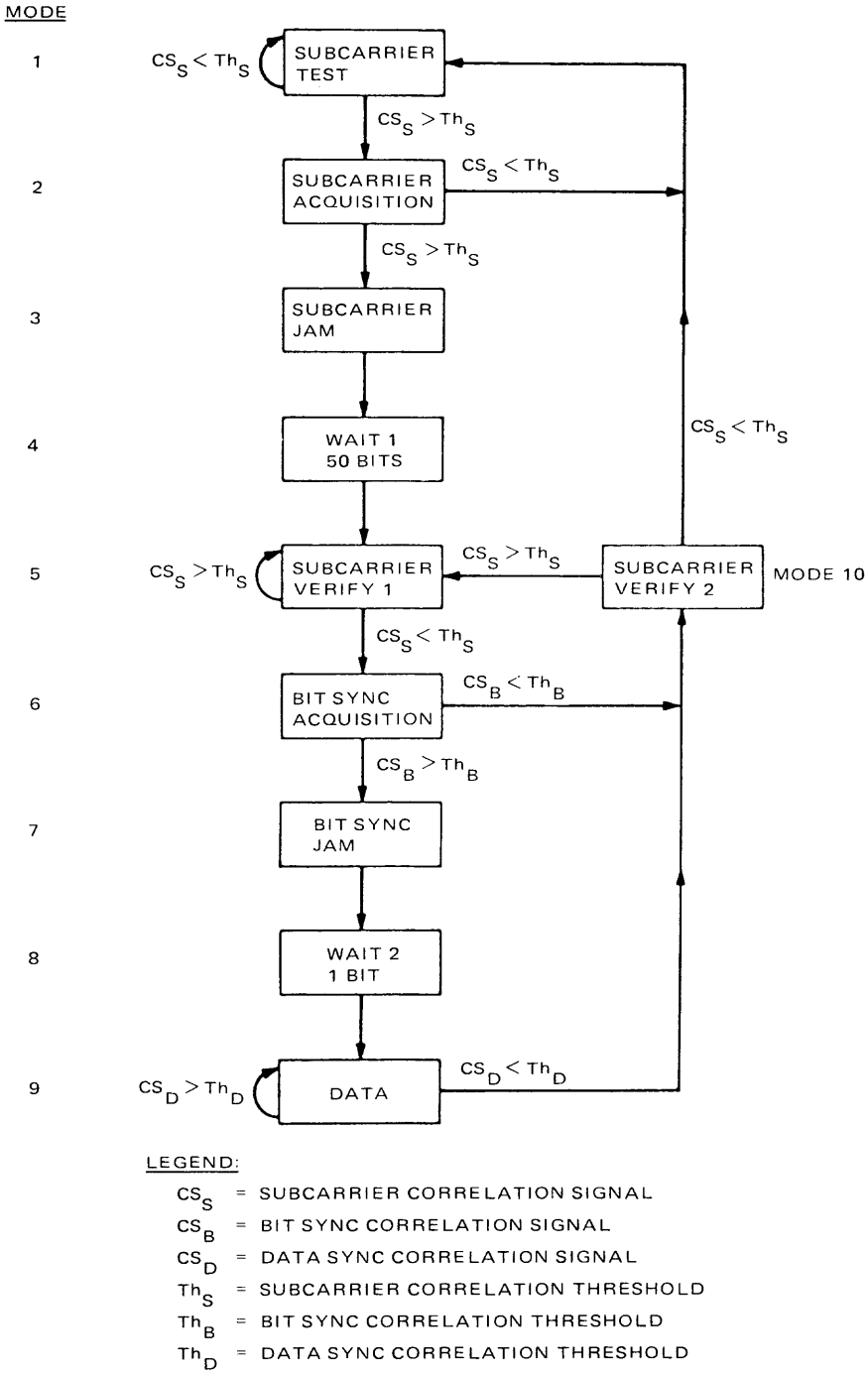


Fig. 6-6. Viking Heritage command detector operation flow diagram

bit sync phase offsets. Tracking is accomplished by forming a sum of samples taken near the known transitions of signal. The sum is compared to zero at the end of 1 bit of accumulation time, and the sign of the sum determines the direction of the clock bump. The clock is bumped a fixed discrete amount in either direction (advance or retard).

Detection of data is accomplished simply by performing the correlation operations described earlier. With all phases now determined, however, only the in-phase correlation is of concern. This in-phase correlation is an approximation to the output of a matched filter, which is known to be the optimum detector.

In mode 9 the system also monitors the presence of command signal by the process of lock detection. The lock detector acquires the sum of the magnitudes of 5 data bit correlations and compares this sum to a threshold. If threshold is not exceeded, the system declares "out-of-lock" and goes to mode 10. In mode 10 a subcarrier threshold test is performed. If threshold is exceeded, the system enters mode 5; if not, the system returns to mode 1.

6.4.1.2 Command prefix. The acquisition procedures described in Section 6.4.1.1 cannot succeed unless the proper signal component is present during each of the correlation modes. Specifically, during modes 1 and 2 unmodulated subcarrier must be present, and during mode 6 subcarrier modulated by the bit sync must be present. To meet these requirements, the prefix shown in Fig. 6-7 is sent before each series of commands.

In designing this prefix signal, all the worst possible conditions are considered to assure that incorrect interpretation of the signal is minimized. When the system is initially in mode 1, it has no sync information and may begin to correlate at an arbitrary time with respect to the transmitted prefix. Consequently, there may exist a mode 1 correlation for which the prefix is only present for part of the correlation time. If this first correlation does not exceed threshold, mode 1 must be repeated and a situation such as that shown in Fig. 6-8a will exist. To allow for successful completion of mode 1 and mode 2 in this case, 10 more command bit times ($10 \times M$ subcarrier cycles) of unmodulated subcarrier will be required after the unsuccessful mode 1 correlation is completed. Therefore, 15 command bit times of the prefix are allowed for unmodulated subcarrier.

A different situation occurs if the mode 1 correlation is successful, even through the prefix is only partially present. This condition is shown in Fig. 6-8b. (Note that the "jam" modes consume a negligible amount of time and therefore are not shown.) In such a case, mode 1 and mode 2 have more than ample time for completion, but the timing of the mode 5 bit sync correlation becomes critical. If the mode 1 correlation were completed very early in the prefix, it might be possible for the system to enter mode 5 before the bit sync portion of the prefix arrived. This situation is prevented by the mode 4 waiting period. As

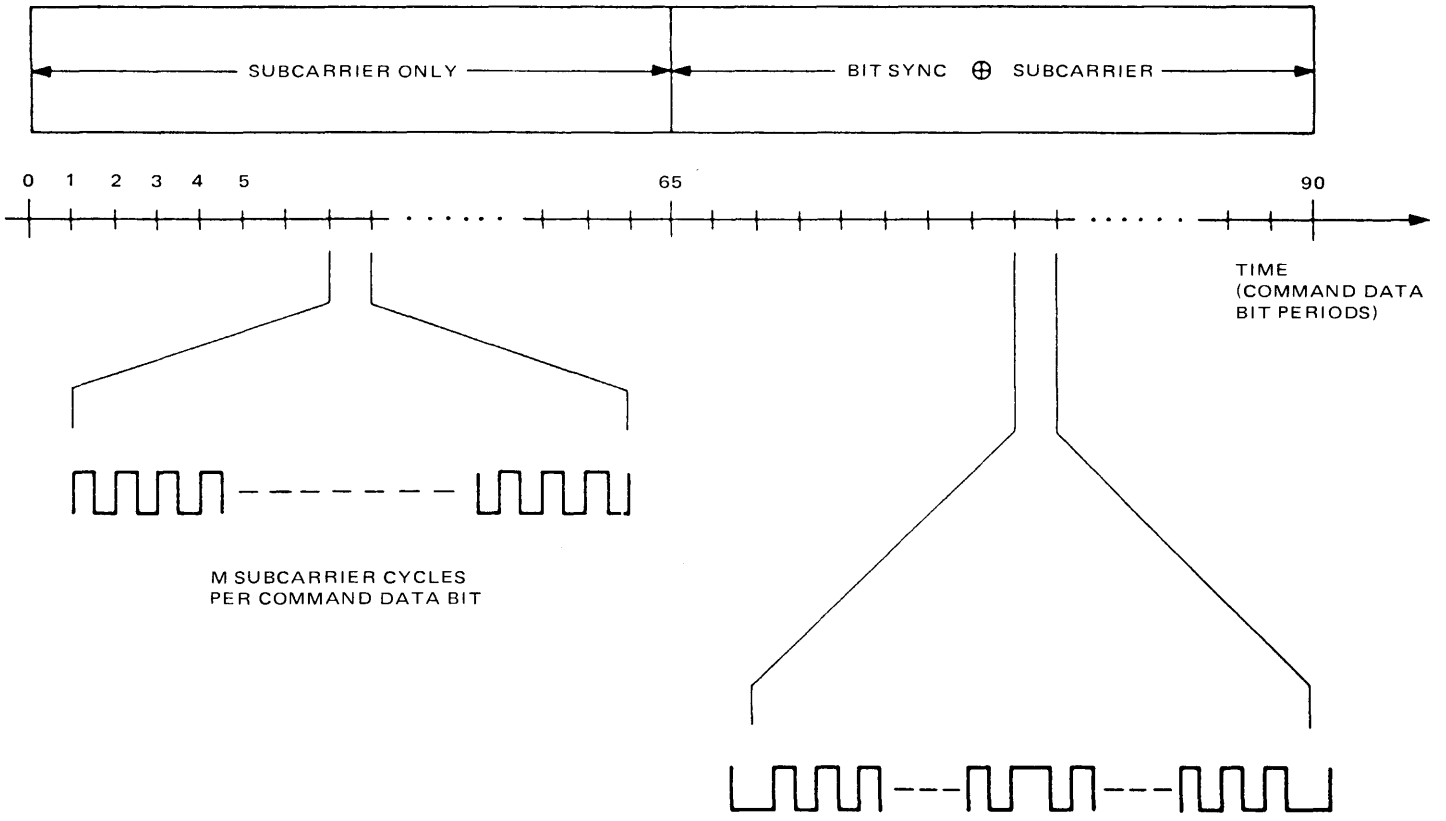


Fig. 6-7. Viking Heritage command detector acquisition signal format

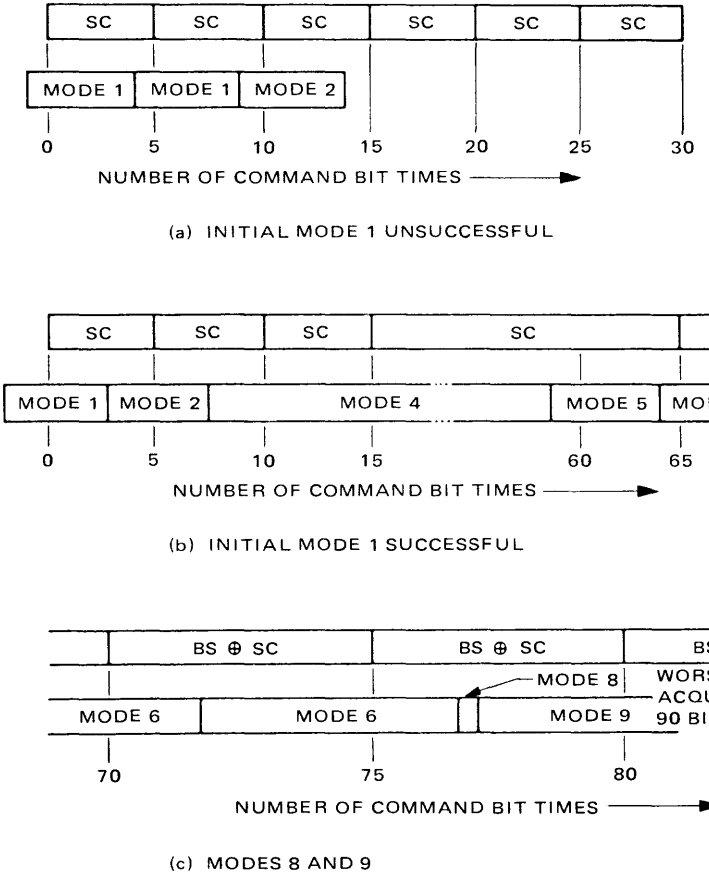


Fig. 6-8. Viking Heritage command detector prefix timing diagram

a very conservative worst-case design, it may be assumed that the mode 1 correlation is completed at the very beginning of the prefix. In this case the required waiting period would have to be 50 bits, which is the time allotted for mode 4. Figure 6-8b also points up the reason for the existence of mode 2; if the mode 1 correlation is successful with only partial correlation with the prefix, the resultant correlation values are likely to be very bad estimates of the subcarrier phase. The existence of mode 2 guarantees that prefix will be present over the entire subcarrier phase-determining correlation.

Bit sync acquisition occurs in modes 6 through 8. From Fig. 6-8b, we see that if worst-case conditions prevail, the mode 4 waiting period will consume 50 command bit times of unmodulated subcarrier. In this case 20 command bit times of bit-sync-modulated subcarrier would be needed to successfully complete modes 5 and 6. To allow for this eventuality, 20 command bit times of

the prefix are allotted to bit sync modulated subcarrier. Thus, with the prefix signal shown in Fig. 6-7, the system should be able to successfully acquire subcarrier sync and bit sync. However, it should be noted that modes 5 and 6 (including the small amount of time for mode 7) may be completed at any point during an actual transmitted data bit period, since the initial phase offset is random. Even though sync has been established, it will be necessary for the system to wait until the start of the next bit before it begins to detect data. This is the purpose of mode 8, as shown in Fig. 6-8c. After mode 8 the system tracks the phase offsets of subcarrier and bit sync for 10 command bit times. This 10 bits of tracking ensures that all synchronization is complete and that the system is ready to output data.

6.4.1.3 Command detector power allocation. The Viking Heritage command detector must detect and acquire both subcarrier and bit sync references in sequence before detecting data. To accomplish this synchronization in sequence, commands transmitted from the Deep Space Stations are prefixed by a short burst of unmodulated subcarrier and of the subcarrier bi-phase modulated by the bit sync. During the data detection mode the command information bits are bi-phase modulated by the subcarrier and the bit sync. The signal transmitted by the DSS for each mode described above can be written as

$$\sqrt{2P_{tr}} \sin [\omega_c t + \theta_D \mathrm{Sq}(\omega_s t)] \quad \text{for subcarrier detection and acquisition modes} \quad (6.4\text{-}1)$$

$$\sqrt{2P_{tr}} \sin [\omega_c t + \theta_D BS \oplus \mathrm{Sq}(\omega_s t)] \quad \text{for bit sync detection and acquisiton modes} \quad (6.4\text{-}2)$$

and

$$\sqrt{2P_{tr}} \sin [\omega_c t + \theta_D D(t) \oplus BS \oplus \mathrm{Sq}(\omega_s t)] \quad \text{for data detection mode} \quad (6.4\text{-}3)$$

where

ω_c = RF carrier frequency

ω_s = subcarrier frequency

θ_D = data modulation index

$D(t)$ = data bits

BS = bit sync squarewave of frequency ω_s/M

P_{tr} = transmitted power

$\mathrm{Sq}(\omega_s t)$ = squarewave subcarrier of frequency ω_s

The ratio of data power P_D to total receiver power P_T is determined by the modulation index as

$$\frac{P_D}{P_T} = \sin^2 \theta_D \tag{6.4-4}$$

The ratio of carrier power P_C to total power with command modulation on is given by

$$\frac{P_C}{P_T} = \cos^2 \theta_D \tag{6.4-5}$$

where θ_D, P_D, P_C and P_T have been defined in Section 6.3.

6.4.1.4 Command detector channel performance. The performance of the command channel is measured in terms of its overall acquisition time, its bit error rate, and its ability to detect the presence of the transmitted signal. Since subcarrier and bit sync acquisition are performed in sequence according to the basic system algorithm, the overall probability as well as the probability associated with each mode must be specified. For phase-coherent, uncoded demodulation, the bit error probability P_{BE}, is given by

$$P_{BE} = \frac{1}{2} \operatorname{erfc} \left(\sqrt{ST_B/N_0}\right) \tag{6.4-6}$$

The received signal at the input to the command detector is degraded by the system losses. Therefore,

$$\frac{ST_B}{N_0} \triangleq \left(\frac{P_T}{N_0}\right)\left(\frac{P_D}{P_T}\right) T_B \eta_s \tag{6.4-7}$$

where

η_s = system efficiency

The system loss is primarily composed of losses due to a noisy carrier reference η_{RL}, low-pass filtering loss η_{LPF}, analog-to-digital quantization loss $\eta_{A/D}$, subcarrier demodulation loss η_{SDL}, and bit synchronization and detection loss η_{BSDL}. Thus,

$$\eta_s = \eta_{RL}\,\eta_{LPF}\,\eta_{A/D}\,\eta_{SDL}\,\eta_{BSDL} \tag{6.4-8}$$

6.4.1.4.1 *Low-pass filter loss and choice of sampling time.* The subcarrier demodulation process of the Viking Heritage Command Detector can be modeled as a second-order all-digital phase-locked loop as shown in Fig. 6-9. The low-pass filtering loss can be estimated by assuming a squarewave input to a flat, sharp cutoff, low-pass filter model. In this design, the low-pass filter bandwidth W is chosen to be

$$W = 16/T_{sc} \tag{6.4-9}$$

where T_{sc} (= T_B/M) is the subcarrier period. With this choice, the first seven harmonics of a squarewave input pass the filter, and the low-pass filter loss, η_{LPF}, is 0.2 dB.

The choice of W also determines the optimum sampling time of the A/D converter. The mean and variance of the signal amplitude leaving the A/D are given by:

$$\mu = K\sqrt{S} \tag{6.4-10}$$

and

$$\sigma^2 = KW\,\frac{N_0}{2} \tag{6.4-11}$$

where

$K = T_B/t_s$ = number of samples per bit time

S = signal power

$N_0/2$ = two-sided noise spectral density of the Gaussian noise process entering the A/D

t_s = sampling time and is an integer fraction of T_{sc} when the loop is in lock

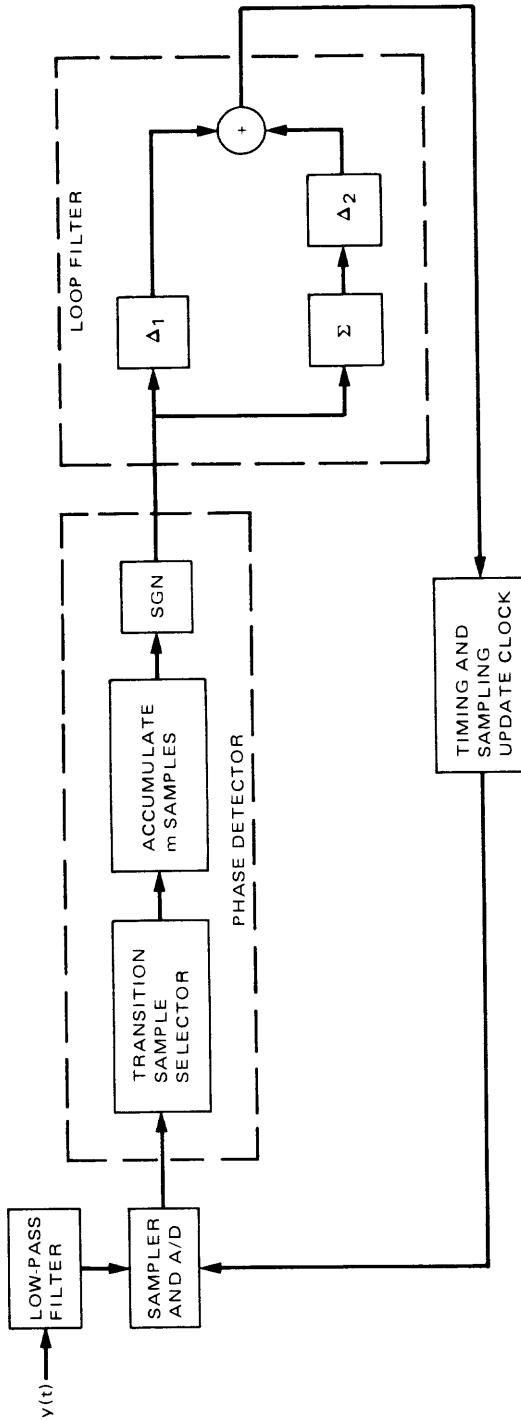


Fig. 6-9. Block diagram for all-digital PLL Viking Heritage command detector

and we have assumed that the loop is synchronized and that the samples are independent. The signal-to-noise ratio (SNR) into the PLL is then given by μ^2/σ^2, or,

$$SNR = \frac{KS}{\frac{N_0}{2}W} = \left(\frac{2ST_B}{N_0}\right)\left(\frac{K}{T_B W}\right) = \left(\frac{2ST_B}{N_0}\right)\left(\frac{t_s}{W}\right) \qquad (6.4\text{-}12)$$

The autocorrelation function of the noise samples from an ideal low-pass filter is

$$R_n(t) = \sigma^2 \frac{\sin(\pi W t)}{\pi W t} \qquad (6.4\text{-}13)$$

Sampling with period $1/W$, $2/W$, $3/W$, etc., ensures that the samples are uncorrelated, which is equivalent to independence for a Gaussian process. Clearly, the SNR is maximized if sampling is done at the fastest rate, or

$$t_s = \frac{1}{W} = \frac{T_{sc}}{16} \qquad (6.4\text{-}14)$$

This yields a SNR of $2ST_B/N_0$, which is the theoretical maximum.

6.4.1.4.2 *Analog-to-digital quantization loss, $\eta_{A/D}$.* By converting an analog signal to a digital form, some loss in effective signal energy to noise spectral density ratio occurs.

Reference [6-8] shows that, for a quantizer (A/D converter) of infinite range, the quantizer may be considered to add independent random noise with variance

$$\sigma^2_{\text{A/D}} = \frac{q^2}{12} \qquad (6.4\text{-}15)$$

when the width of a quantization level q is a small fraction of the dynamic range of the quantized variable. If the quantized variable is Gaussian, this criterion is equivalent to requiring that $q \leqslant \sigma$, where σ is the standard deviation of the noise process into the A/D converter, $\sigma = \sqrt{N_0 W/2}$.

When the range of the quantizer is adjusted so that saturation occurs only for noise of amplitude greater than 3σ, the quantizer may be modeled as an infinite quantizer with the results quoted above.

In the system described, a 4-bit A/D is used, with 16 quantization levels, and the gain of the A/D is adjusted to conform to the 3σ criterion, and the loss, $\eta_{A/D}$, is 0.1 dB.

6.4.2 NASA Standard Command Detector Operation

The NASA standard command detector was developed as an integral part of the NASA standard transponders. It was designed to be in compliance with the NASA Planetary Program Flight/Ground Data System Standards [6-9]. The unit is operable in two modes—the NRZ mode without a subcarrier for TDRS application and the PSK mode with a 16-kHz sinewave subcarrier for near-earth and deep-space applications. This chapter is concerned with the PSK mode only. A total of nine command data bit rates are in-flight selectable. The allowed bit rates are specified as $2000/2^N$, where $N = 0, 1, 2, \ldots 8$.

Unlike the Viking Heritage system, the NASA standard command detector establishes subcarrier and bit synchronizations simultaneously, thus eliminating the cumbersome multimode acquisition process required by the former. For the NASA standard command detector, detections of subcarrier, bit sync and data are performed simultaneously with a single correlation process. The subcarrier and bit sync phases are coherently related and are controlled by a quadrature subcarrier tracking loop and a data-transition bit synchronization loop, respectively.

A block diagram for the NASA standard command detector is shown in Fig. 6-10. Functionally, this command detector consists of a coherent automatic gain control (AGC) unit, a sample-and-hold (SH) circuit, an analog-to-digital converter (ADC), a second-order data-aided subcarrier tracking loop, a data transition bit synchronization loop, and a lock detector. Structurally, it consists of the signal conditioning assemblies, i.e., AGC, SH, and ADC, read only memory (ROM), random access memory (RAM), and a digital processing assembly employing a custom-LSI approach. Figure 6-11 illustrates the functional architecture of this command detector.

6.4.2.1 Subcarrier tracking. The NASA standard command detector utilizes a coherent sampling scheme to implement a second-order suppressed subcarrier data-aided loop. The theory and analysis of the data-aided loop configuration have been documented [6-10]. The subcarrier tracking loop utilizes a perfect integrator and a quasicontinuously variable phase correction (resolution 1/64 of one subcarrier cycle). The maximum subcarrier phase correction step is 45 degrees. The entire tracking algorithm is stored in ROM, with ROM-resident loop coefficients selected during the acquisition and the tracking phases so as to meet the acquisition time requirements and minimize steady-state phase jitter.

6.4.2.1.1 *Coherent sampling.* The sampling process of the NASA standard command detector takes place at the subcarrier level as illustrated in Fig. 6-12. Two

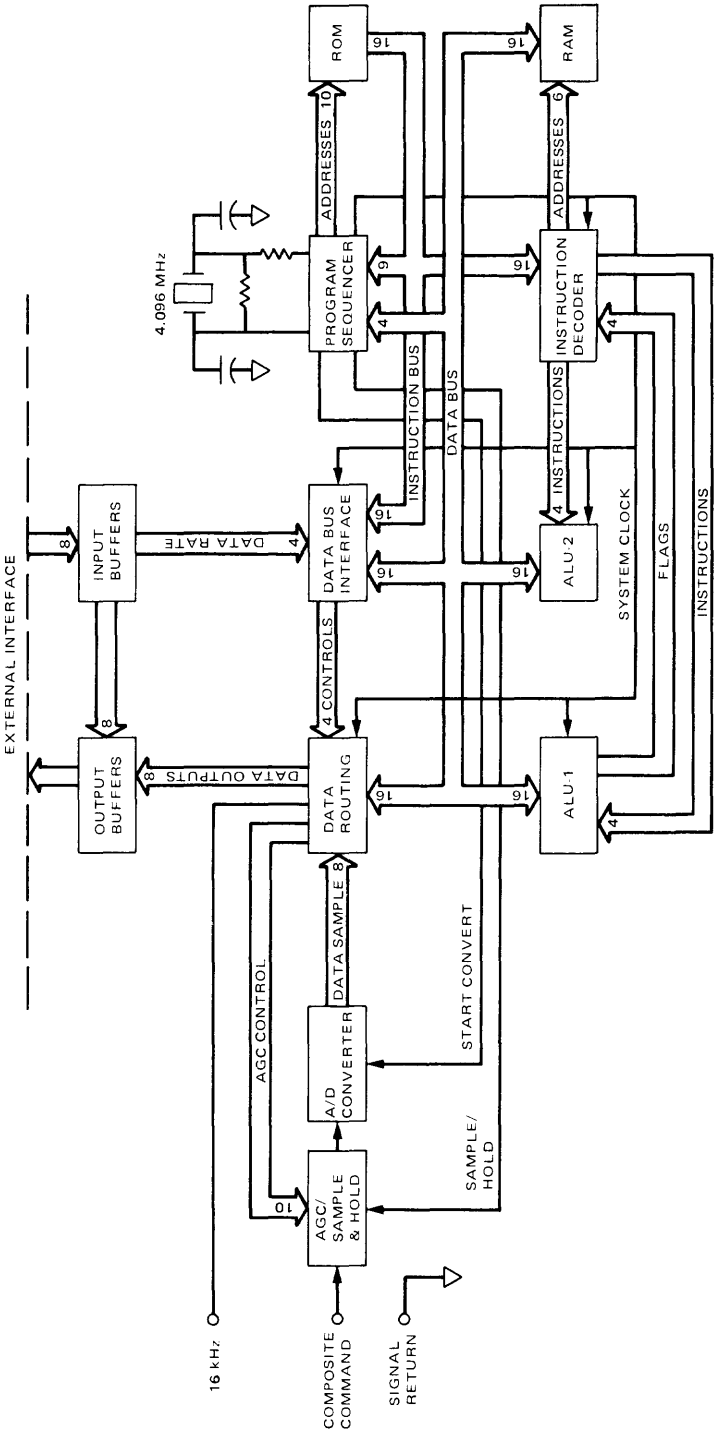


Fig. 6-10. NASA standard command detector block diagram

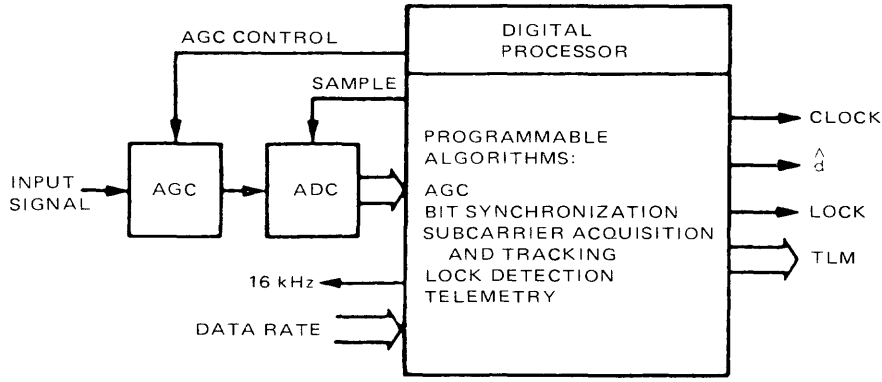


Fig. 6-11. NASA standard command detector functional architecture

samples 90 deg apart are taken in each subcarrier cycle. This process effectively removes the subcarrier and translates the input sequences to baseband. By sampling in quadrature, one channel represents the in-phase or data channel and the other the quadrature phase or error channel. The integrated sampling results of the error channel are used to maintain subcarrier synchronization, and those of the data channel are used to maintain bit synchronization and detect the data bits. Subcarrier phase corrections are executed at the end of each detected bit. The amounts of correction or bump sizes for various conditions are predetermined and are stored in ROM.

With this demodulation scheme, the subcarrier reference has two possible stable lock points. As a result, the polarity of the detected bits may be in phase or 180 deg out of phase. This ambiguity of data polarity must be resolved in the decoding process. This function is usually performed outside of the command detector domain (see section 6.4.3.1 below).

6.4.2.1.2 *Tracking loop model.* The command detector subcarrier tracking loop may be modeled as shown in Fig. 6-13. Sampled values of the error channel are used to estimate the instantaneous phase error of the subcarrier zero-crossing. It is readily seen that if the zero-crossing phase error is ϕ, the sampled value at that instant is $A_r \sin \phi$, where A_r is the voltage amplitude of the input sinusoid. These samples are converted into eight-bit digital numbers which are accumulated over one bit period. A total of N samples are taken for each accumulation, where N is the ratio of the bit and subcarrier periods. At the end of the accumulation, the integrated value is scaled according to the bit rate and dumped into the loop filter. The scaling is used to maintain a constant subcarrier loop dynamic range at all permissible data rates. The subcarrier loop filter is a first-order digital filter. The output of this filter is the summation of two branches. One branch is simply the scaled input value and the other is a scaled running accumulation of the input. The loop filter output, together with the data estimator from the bit

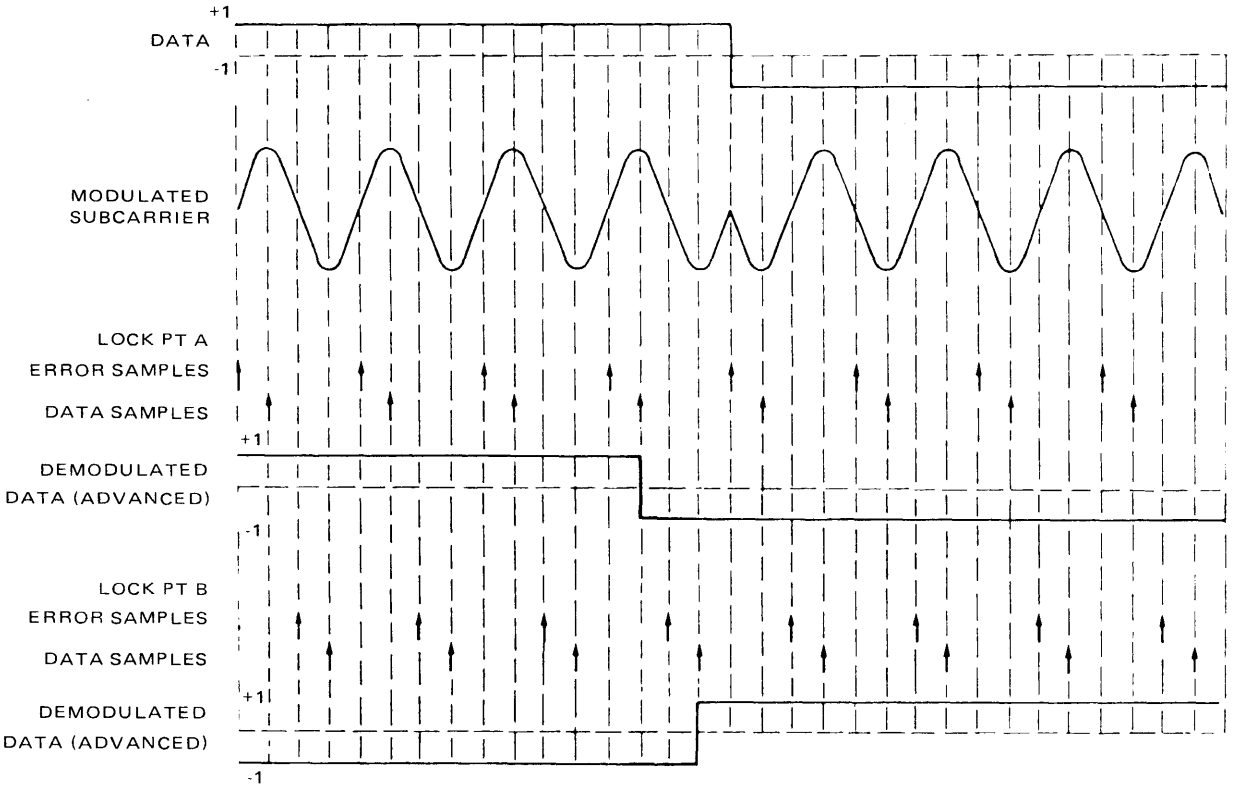


Fig. 6-12. NASA standard command detector coherent sampling demodulation

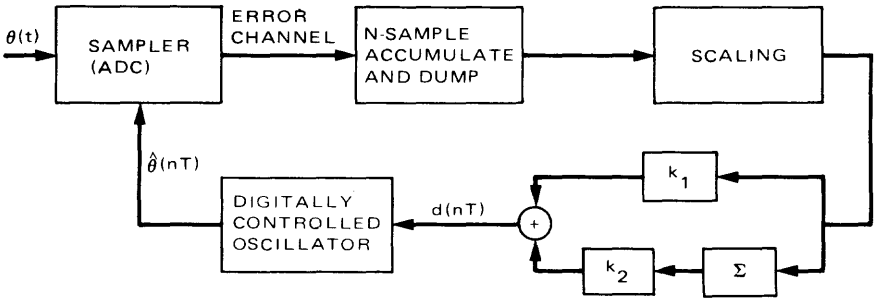


Fig. 6-13. NASA standard command detector subcarrier loop model

sync loop, provides a quantitative value to the digitally controlled oscillator to either advance or retard its phase so as to reduce the zero-crossing phase error of the subcarrier. An analysis of this tracking loop has been documented [6-11].

6.4.2.2 Bit sync tracking. For bit sync tracking and data detection, the sampled values of the data channel are accumulated over a bit interval. A measure of the normalized bit sync phase error is obtained by integrating over a bit period centered at the expected data transition point. This process is called the mid-phase accumulation, as distinguished from the in-phase accumulation normally performed for data detection.

6.4.2.2.1 *Bit sync algorithm.* Proper detection of the demodulated data stream requires the use of a matched filter. Implementation of this matched filter, in turn, requires the generation of accurate end-of-bit estimates.

The bit sync configuration of the NASA standard command detector utilizes a digital data transition tracking loop which consists of two parallel branches. The in-phase branch, from which the data estimates are produced, monitors the polarity of any data transition, while the mid-phase branch obtains a measure of the lack of bit synchronization.

The theories of the digital data transition tracking loop have been extensively analyzed [6-12, 13, 14]. The implementation of the NASA standard command detector bit sync loop was largely based on the results of these analyses.

6.4.2.2.2 *Bit sync loop model.* Figure 6-14 illustrates a model for the bit sync loop. The data samples are taken from the data channel which is in quadrature to the subcarrier tracking error channel. If the subcarrier zero-crossing phase error is ϕ, then the instantaneous input value of the data sample is $A_r \cos \phi$. These samples are taken once per subcarrier cycle and are converted into digital numbers in the same manner as the subcarrier tracking process and are accumulated into two separate accumulators—one in-phase and one mid-phase. At the

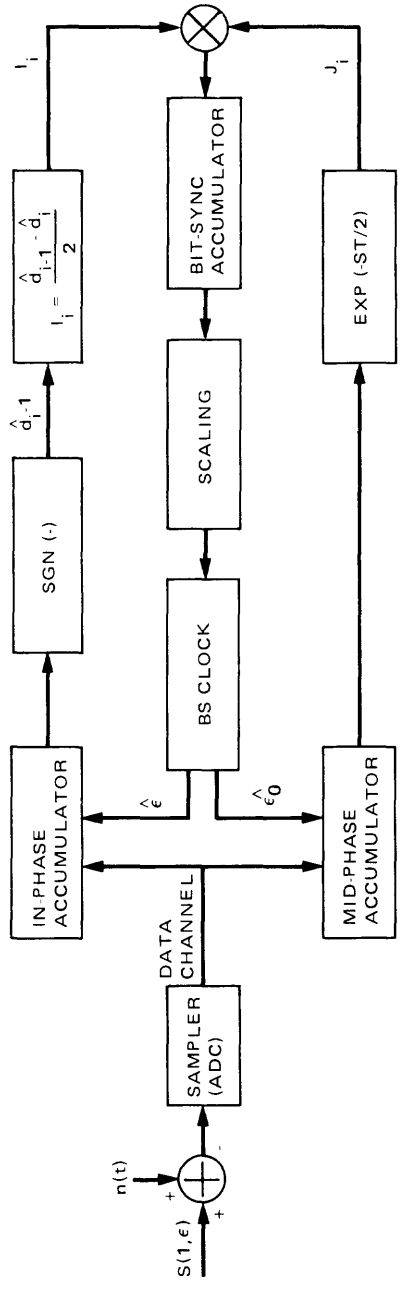


Fig. 6-14. NASA standard command detector bit sync loop model

end of each in-phase bit period, the sign of the in-phase accumulator represents the bit estimate. If there is a data transition between two adjacent bits of data, the contents of the mid-phase accumulator will be retained and further integrated into a third accumulator. Here, K transitions are required to determine the relative phase error of the bit sync estimate, where K is a constant of the bit sync loop residing in the ROM, and an advance or retart correction is made proportionately. (K equals 4 is currently being used.) Correction can be made to the nearest subcarrier cycle. This resolution is augmented by the subcarrier tracking loop performance to provide adequate overall bit sync resolution.

6.4.2.3 Automatic gain control. Automatic gain control establishes a constant, predetermined input signal amplitude to the command detector over a wide dynamic range of received carrier signal power. The NASA standard command detector employs a coherent AGC system which, designed with a log-linear characteristic, provides a dynamic range in excess of 40 dB with loop performance independent of signal amplitude and scaled with data rate. A block diagram for the AGC loop model is shown in Fig. 6-15.

The central element of the AGC system is the gain controllable amplifier which precedes the signal sampler. The data for AGC processing are taken from the output of the in-phase accumulator as described in the bit sync tracking loop. The magnitude of M consecutive values from the in-phase accumulator is integrated in a second accumulator and the result, after appropriate scaling, is compared against a predetermined reference value. The value M is a loop constant programmable in the resident ROM (M equals 9 is currently being used.) The difference between the actual and the reference magnitude is continuously integrated in another digital integrator, the output of which operates on an AGC function circuit which produces the desired gain setting of the AGC amplifier.

6.4.2.4 Lock detector. The lock detector of the NASA standard command detector determines whether or not a command signal is present and provides this decision to the command decoder. In the NASA standard command detector, the lock detection algorithm and threshold coefficients are ROM-resident, providing complete flexibility in tailoring the detector to meet mission-specific acquisition and deacquisition performance requirements.

6.4.2.4.1 *Lock detector specifications.* The general performance requirements for the NASA standard command detector in-lock indication for all data rates are given in Table 6-2.

6.4.2.4.2 *Lock detector model.* The lock detection algorithm is closely related to the AGC algorithm discussed in the preceding section. The model used in describing the lock detector operation and in determining its lock thresholds is shown in Fig. 6-16. As can be seen from this model, the same information used for the AGC function (i.e., output values of the bit magnitude accumulator) is

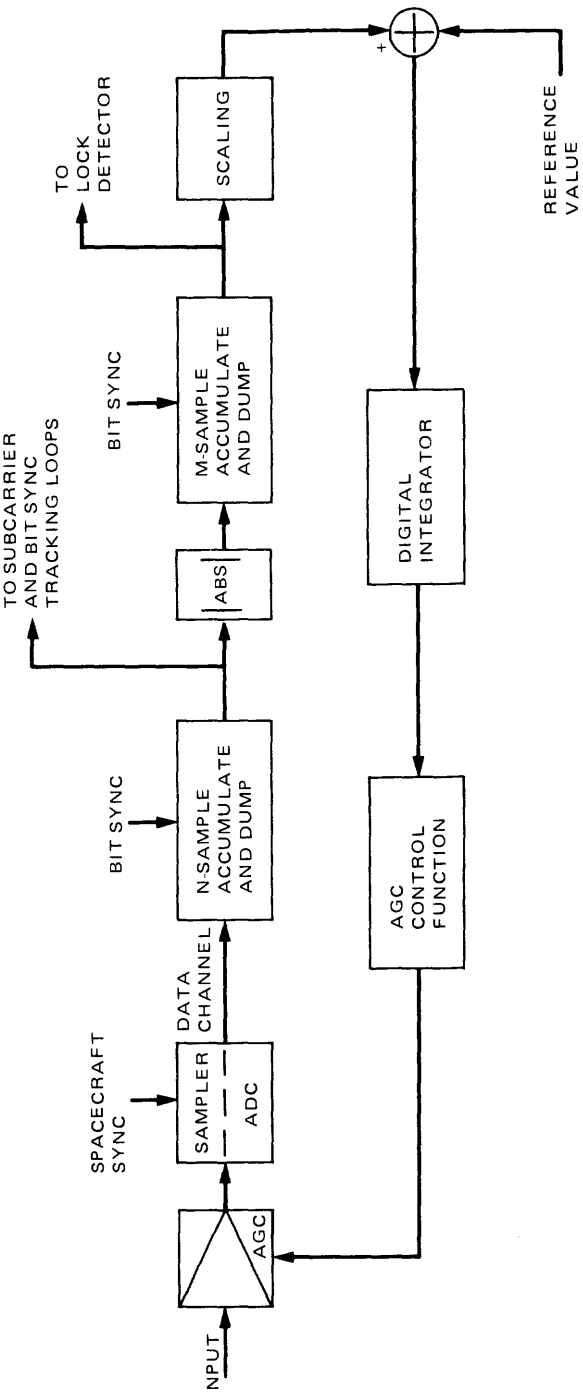


Fig. 6-15. NASA standard command detector AGC loop model

Table 6-2. NASA standard command detector lock detector specifications

Parameter	Condition	Probability
Probability of not acquiring	In the presence of 132-bit acquisition sequence	$\leqslant 1 \times 10^{-4}$
Probability of false acquisition	In any 18-bit interval of noise	$\leqslant 1 \times 10^{-4}$
Probability of dropping lock	Within 27 bits of signal removal	$\leqslant 2 \times 10^{-2}$
Probability of false drop lock	In any 18-bit interval of signal	$\leqslant 2.5 \times 10^{-9}$

also used for the in-lock condition evaluation. To determine the presence or absence of command signal, the magnitude of the matched filter for each bit time is taken and is integrated over M bit times. Then, after appropriate scaling, the integrated value is compared with a lock threshold ("LKTH") if the detector has been out of lock, or with an unlock threshold ("ULKTH") if it has been in-lock. The possible outcomes of these tests are summarized in Table 6-3. Notice that a single test result is not sufficient to change the lock indication.

A change in lock indication, whether from out of lock to in lock or vice versa, requires at least two consecutive threshold tests indicating the same result. The choice of the number of consecutive tests to effect a lock indication change dictates the probabilities for various parameters regarding acquisition and dropping lock, and determines the response time of the lock detector loop.

When the command detector is out of lock, the AGC amplifier is driven to its maximum gain setting and held fixed. Only after lock has been achieved is the AGC permitted to seek its quiescent point of operation.

6.4.2.5 Command prefix. The NASA standard command detector does not require a fixed transmission procedure to acquire lock. However, to insure that no valid data bits are lost before in-lock indication is declared, a sequence of alternate 1's and 0's is transmitted prior to the data bit stream to assist rapid bit sync acquisition. To meet the lock statistics specified in Table 6-2, a 132-bit preamble is required.

6.4.2.6 Power allocation. Since the NASA standard command detector assumes a sinewave subcarrier, the ratios of data power and carrier power to the total received power as a function of the modulation index are

$$\frac{P_C}{P_T} = J_0^2(\theta_D) \tag{6.4-16}$$

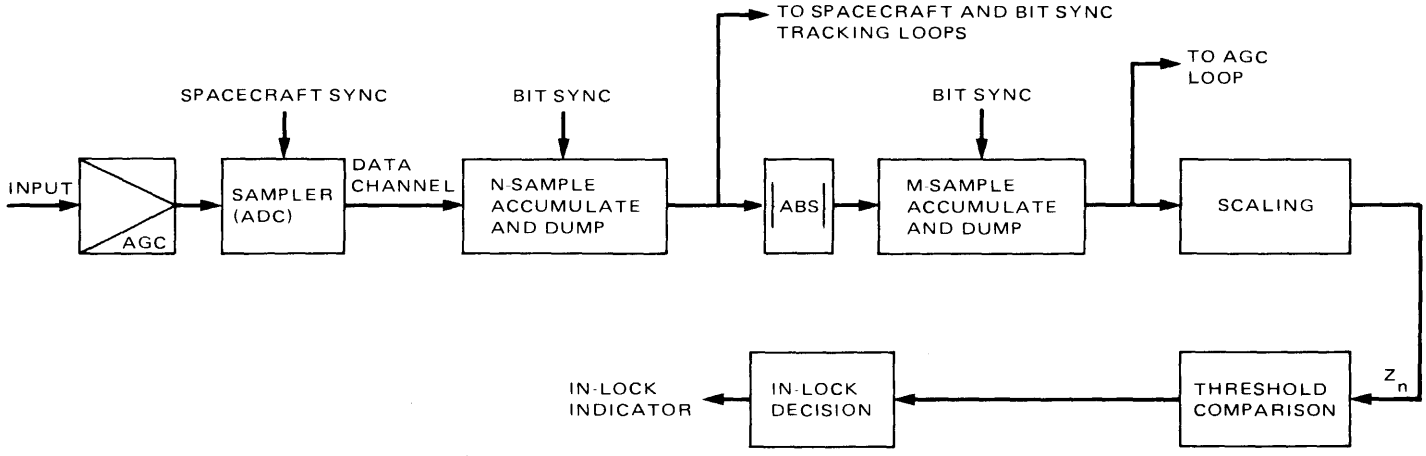


Fig. 6-16. NASA standard command detector, lock detector model

Table 6-3. Lock indication threshold test summary

In-lock condition	Test	Result	Implication
Previously out of lock	Compare with LKTH	$Z_n >$ LKTH	Indicate in lock
		$Z_n <$ LKTH	Remain out of lock
Previously in lock	Compare with ULKTH	$Z_n >$ ULKTH	Remain in lock
		$Z_n <$ ULKTH	Indicate out lock

and

$$\frac{P_D}{P_T} = 2J_1^2(\theta_D) \tag{6.4-17}$$

where θ_D, P_D, P_C, P_T, J_0 and J_1 have been previously defined in Section 6.3.

6.4.2.7 **Channel performance.** As mentioned previously, the design concepts are based on theoretical results of related works [6-8], [6-10], [6-11], [6-13], and [6-14]. No rigorous analyses were attempted to identify the degradation due to the various individual command detector components. However, the aggregate detector degradation was obtained through simulation and tests. Table 6-4 shows the observable performance and degradation of the NASA command detector under various bit rates and frequency offset conditions.

6.4.3 Command Decoding

The functions of the spacecraft command detector are limited to demodulation and detection of the composite command signal. The contents of the command sequence are transparent to the command detector. In order to process the command sequence and deliver the individual commands to the users, the detected command bit stream must be decoded and distributed. These functions are the responsibility of the spacecraft command decoder.

The proper decoding of the command sequence requires a complete understanding of the command structure, word format, and contents. These elements are mission dependent. Though the command decoding process is an important discipline of spacecraft system design, it is beyond the scope and purposes of this chapter. The only detector-related decoding problem is the data ambiguity inherent in the NASA Standard Command Detector.

6.4.3.1 **Real-time decoder ambiguity resolver.** Since the NASA standard command detector may deliver either normal or inverted data bits, depending upon

Table 6-4. NASA standard command detector performance summary (with PN data pattern)

Bit rate, bps	Frequency offset, PPM	BER = 10^{-4}		BER = 10^{-5}	
		ST_B/N_0, dB	CDU degradation, dB	ST_B/N_0, dB	CDU degradation, dB
2000	0	8.8	0.4	10.2	0.6
2000	3906	8.8	0.4	9.8	0.2
125	0	8.9	0.5	9.9	0.3
125	244	9.0	0.6	9.9	0.3
7.8	15.3	9.4	1.0	10.5	0.9

the axis at which the subcarrier is locked, care must be exercised to assure that the right commands are executed. For an elaborate command decoding system utilizing a sophisticated computer, the data ambiguity may be simultaneously resolved when the command frame sync is detected. However, with a decoder of less sophistication, a real-time ambiguity resolver may be implemented.

To resolve the data ambiguity on a real-time basis, a sync word must precede the command bit stream following the alternate 1 and 0 preamble. Careful selection of the sync word will minimize the probability of error at a given length of the sync word. When the detector indicates an in-lock condition, the ambiguity resolver will first search for this sync word and determine whether it is in its normal or inverted position. The subsequent command bits will then be accepted in the polarity as indicated by the sync word polarity decision. A seven-bit Barker code has been employed for such a sync word for near-earth applications.

6.5 Command Channel Design Control

As a major function of the spacecraft telecommunications system, the command channel is subject to the provisions of the telecommunications design control policy [6-15]. This section identifies the parameters relating to the command function and presents an example of the command channel design control table.

6.5.1 Carrier Channel

This section presents the definitions of all the parameters which affect the performance of spacecraft receiver carrier tracking.

6.5.1.1 Total transmitting power at antenna. This parameter is the total power into a DSS antenna. The nominal ground transmitted power is 20 kW. However, 100 kW (and possibly 400 kW) transmitted power may be used for emergency and other occasions.

6.5.1.2 Transmitting (ground) antenna gain. This is the peak gain of a DSS antenna including pointing and circuit loss, and wind and gravity deformation loss. Two antenna nets are normally in use: 34- and 64-m antennas.

6.5.1.3 Space loss. This is the loss due to the transmission medium. In a normal mission, space loss is a function of carrier frequency and spacecraft-earth range. The absorption loss due to atmospheric and plasma effects represents additional losses and should be taken into consideration.

6.5.1.4 Polarization loss. This is the loss due to mismatch in polarization between the ground and the spacecraft antennas. A measure of ellipticity is used to determine the polarization loss.

6.5.1.5 Receiving (spacecraft) antenna gain. This parameter represents the gain of the spacecraft antenna in the direction of the incoming signal. Traditionally, the value of this parameter is derived by factoring the expected loss due to antenna pointing error into the boresight antenna gain.

6.5.1.6 Spacecraft circuit loss. This is the loss due to transmission cables or waveguides from the spacecraft antenna terminal to the receiver input port.

6.5.1.7 Total received power. This is the realizable power at the receiver input port.

6.5.1.8 System noise spectral density. This is the Gaussian white noise power per unit frequency (N_0) introduced by the environment, cabling, switches, antenna and receiver. This parameter is proportional to the spacecraft circuit loss and the effective system noise temperature. The latter is a function of the receiver noise figure, circuit loss, cable physical temperature, and antenna noise temperature.

6.5.1.9 Ratio of total receiver power to noise spectral density. This is the ratio of total received power to noise spectral density (P_T/N_0) at the input of the receiver.

6.5.1.10 Ratio of carrier power to total received power. This parameter reflects the allocation of power to the carrier as a fraction of the total power. It is determined by the modulation indices for the command and ranging (if present) signals.

6.5.1.11 Carrier threshold tracking bandwidth, $2B_{LO}$. This is the two-sided noise bandwidth of the carrier tracking loop at threshold P_C/N_0.

6.5.1.12 Signal-to-noise ratio in $2B_{LO}$. This is the signal-to-noise power ratio into the spacecraft carrier tracking phase-locked loop for a bandwidth of $2B_{LO}$.

6.5.1.13 Threshold signal-to-noise ratio in $2B_{LO}$. This is the minimum value of the signal-to-noise power ratio defined in 6.5.12 acceptable for the required accuracy of carrier tracking.

6.5.1.14 Carrier margin. This is the amount of carrier power in excess of the minimum acceptable value for proper carrier tracking.

6.5.2 Command Channel

This paragraph presents the definitions of all the parameters relevant to command detection.

6.5.2.1 Bit rate. This parameter refers to the number of command bits transmitted over a period of one second.

6.5.2.2 Ratio of command power to total received power. This power ratio reflects the allocation of power to the command channel as a fraction of the total power. It is determined by the modulation index for the command signal. The power available to the command detector is readily computable with this allocation.

6.5.2.3 Command detector losses. This is the aggregate loss within the command detector. Factors contributing to the command detector losses include radio loss, waveform distortion, subcarrier demodulation loss, bit sync detection loss and circuit loss.

6.5.2.4 Ratio of command detector per bit energy to noise spectral density. This ratio reflects the net energy over one bit period at the output of the command detector to the noise spectral density (ST_B/N_0).

6.5.2.5 Threshold ST_B/N_0. This is the minimum acceptable value of the ST_B/N_0 defined in 6.5.2.4 to achieve a given bit error rate.

6.5.2.6 Performance margin. This is the amount of command power in excess of the minimum acceptable value required to achieve the specified bit error rate.

6.5.3 Command Function Design Control Table

Table 6-5 is a typical design control table for the uplink command function. It consists of two parts – the carrier channel and the command channel. The signal-to-noise ratio in relation to the bandwidth for both channels must exceed their respective thresholds for the command link to be established.

Table 6-5. Typical command design control table

	Uplink carrier				
	Design	Favorable tolerance	Adverse tolerance	Mean	Variance
Transmitter parameters					
(1) RF power, dBm Power output = 20.0 kW	73.00	0.00	0.00	73.0	0.00
(2) Circuit loss, dB	0.00	0.00	0.00	0.0	0.00
(3) Antenna gain, dBi Elevation angle = 10.00 deg	60.60	0.30	-0.70	60.3	0.08
(4) Pointing loss, dB	-0.10	0.05	-0.05		
Path parameters					
(5) Space loss, dB Frequency = 2115.00 MHz Range = 2.581 + 08 KM = 1.73 AU	-267.19			-267.2	0.00
(6) Atmospheric attenuation, dB	-0.02	0.00	-0.00	-0.0	0.00
Receiver parameters					
(7) Polarization loss, dB	-0.10	0.05	-0.05		
(8) Antenna gain, dBi	34.50	0.40	-0.40	34.4	0.03
(9) Pointing error, dB	-0.06	0.03	-0.03	-0.1	0.01
Limit cycle, deg	0.20	-0.20	0.00		
Angular errors, deg	0.00	0.00	0.00		
(10) Receiver circuit loss, dB	-1.50	0.50	-0.50	-1.5	0.08
(11) Noise spectral density, dBm/Hz	-169.26	-0.21	0.87	-168.9	0.03
Operating temperature, K	860.00	-40.00	190.00		
Hot body noise, K	0.00	0.00	0.00		
(12) Carrier threshold noise BW, dB-Hz	12.55	-0.97	0.79	12.5	0.13
Total power summary					
(13) Received power, P_T, dBm (1+2+3+4+5+6+7+8+9+10)				-101.1	0.21
(14) Received P_T/N_0, dB-Hz (13-11)				67.9	0.24
(15) Ranging suppression, dB	0.00	0.00	0.00	0.0	0.00
(16) Command suppression, dB	-3.06	0.10	-0.10	-3.1	0.00
(17) Carrier power/total power, dB (15+16)				-3.1	0.00
(18) Received carrier power, dBm, (13+17)				-104.1	0.21
(19) Carrier SNR in $2B_{LO}$, dB (18-11-12)				52.3	0.37
				3S = 1.8	

Table 6-5. (contd)

	Command channel				
	Design	Favorable tolerance	Adverse tolerance	Mean	Variance
Data channel performance					
(20) Data bit rate, dB Bit rate = 31.3 bps	14.95	0.00	0.00	14.9	0.00
(21) Data power/total power, dB	-3.35	0.10	-0.10	-3.3	0.00
(22) Data power to receiver, dBm (13+15+21)				-104.4	0.21
(23) ST_B/N_0 to receiver, dB (22-20-11)				49.6	0.25
(24) System losses, dB	-3.00	0.50	-0.50	-3.0	0.04
Radio loss, dB	0.00	0.00	0.00		
Demodulator, detect loss, dB	0.00	0.00	0.00		
Waveform dist loss, dB	0.00	0.00	0.00		
(25) ST_B/N_0 output, dB (23+24)				46.6	0.29
(26) Threshold ST_B/N_0, dB	9.60	0.00	0.00	9.6	0.00
Threshold bit error rate	1.00-05				
(27) Performance margin, dB (25-26)				37.0	0.29
					3S = 1.6

References

6-1. *Multimission Command System Requirements for the 1977-1987 Era*, Document 663-2, Jet Propulsion Laboratory, Pasadena, Calif., June 3, 1977 (an internal document).

6-2. Thorman, H. C., "DSN Command System Mark III-80," *The Telecommunications and Data Acquisition Progress Report, March and April 1980*, Publication 42-57, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1980.

6-3. Thorman, H. C., "DSN Command System," *The Telecommunications and Data Acquisition Progress Report, May and June 1981*, Publication 42-64, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1981.

6-4. *Deep Space Network/Flight Project Interface Design Book*, Document 810-5, Rev. D., Vol. I, "Existing DSN Capabilities," Section CMD-10, "Command System," Jet Propulsion Laboratory, Pasadena, Calif. (an internal document).

6-5. *Deep Space Network/Flight Project Interface Design Book*, Document 810-5, Rev. D., Vol. II, "Proposed DSN Capabilities," CMD-10 (PC) "Command System," Jet Propulsion Laboratory, Pasadena, Calif. (an internal document).

6-6. *Command Modulator Assembly Operations and Maintenance Technical Manual*, Publication TM011-5A, Jet Propulsion Laboratory, Pasadena, Calif., July 15, 1977 (an internal document).

6-7. Holmes, J. K., and Tegnelia, C. R., *Digital Command System Second-Order Subcarrier Tracking Performance*, Technical Report 32-1540, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 1971.

6-8. Widrow, B., "A Study of Rough Amplitude Quantization by Means of Nyquist Sampling Theory," *IRE Transactions on Circuit Theory*, Vol. CT-3, Dec. 1956.

6-9. *NASA Planetary Program Flight/Ground Data System Standards*, NASA Document (unnumbered), Section 3, "Command Standard," June 1977.

6-10. Simon, M. K., and Springett, J. C., "The Theory, Design, and Operation of the Suppressed Carrier Data-Aided Tracking Receiver," Technical Report 32-1583, Jet Propulsion Laboratory, Pasadena, Calif., June 1973.

6-11. *NASA Standard Command Detector Unit Engineering Report*, Motorola, February 1977.

6-12. Lindsey, W. C., and Tausworthe, R. C., "Digital Data-Transition Tracking Loops," *Space Programs Summary 37-50*, Vol. III, Jet Propulsion Laboratory, Pasadena, Calif., April 1968.

6-13. Simon, M. K., "An Analysis of the Steady-State Phase Noise Performance of a Digital Data-Transition Tracking Loop," *Space Programs Summary 37-55*, Vol. III, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 1969.

6-14. Simon, M. K., "Optimization of the Performance of a Digital-Data-Transition Tracking Loop," Jet Propulsion Laboratory, Pasadena, Calif., *IEEE Transactions on Communication Technology*, Oct. 1970.

6-15. "Policy for the Design of Deep Space Telecommunication System." Document 601-12, Jet Propulsion Laboratory, Pasadena, Calif., Sept. 1975 (an internal document).

Chapter 7

Spacecraft Radio Frequency Subsystem

Man K. Tam and Carroll F. Winn

7.1 Introduction

The flight Radio Frequency Subsystem (RFS) is a vital component for the three spacecraft telecommunications functions of tracking, command, and telemetry. It is the radio and the signal processing equipment residing in the spacecraft that interfaces with the Control & Data Subsystem and performs two-way communications with the earth-based Deep Space Network. The RFS consists of all the elements for RF reception, demodulation, modulation, and transmission, including those for command detection and telemetry modulation.

This chapter reviews briefly the historical background for the development of the planetary RFS, and describes the performance capabilities of the various functional subassemblies.

7.2 Historical Background

Since the initiation of the nation's space exploration program in the late fifties, the spacecraft RFS design has experienced several stages of evolution,

both in parametric improvement and in component technology breakthrough. While no dramatic changes in microwave hardware have been evidenced during the brief history of the spacecraft radio development, the progress of the flight RFS has been dominant in the areas of the transponder and power amplifier design. The continual increase in the carrier frequency, increase in RF power amplifier output, and improved modulation techniques have all contributed to the improved performance of the telecommunications links.

During the infant period of the space program exemplified by the Pioneer and Ranger missions, the RFS consisted of an L-band radio and a three-watt cavity amplifier. The Mariner-class spacecraft of the sixties and early seventies witnessed the implementation of the S-band link, and traveling-wave tube amplifiers (TWTA) prevailed in obtaining RF power amplification. It was not until the Mariner 73 and the Viking 75 missions that an X-band experimental transmitter was introduced. Prior to the Viking era, the flight RFS was constructed predominantly with discrete component parts.

A major advancement in RFS fabrication was attained shortly after the Viking launch when a miniature transponder replaced its discrete component counterpart for the Voyager 77 mission. The miniature transponder was the forerunner of the NASA standard transponder which was developed subsequently. The Voyager RFS also marked the beginning of the full-up S/X-band dual frequency downlink capabilities, and solid-state RF amplifiers (SSA) had reached the maturity to compete seriously with the TWTAs for planetary applications. This generation of RFS is expected to service the Galileo mission scheduled to launch in 1985.

Up to the miniature transponder era, the command detector unit (CDU) and the telemetry modulation unit (TMU) had been implemented separately as the Modulation Demodulation Subsystem (MDS). However, due to the technological breakthrough in large-scale integration of digital circuitry, it was possible to incorporate the CDU in the transponder proper when the NASA standard planetary transponder was developed late in the seventies. The standard transponder had been chosen by the International Solar Polar Mission and the Halley Intercept Mission, and will likely be employed by new projects started during this decade.

While the standard transponder does not include the TMU, it is conceivably a simple task (as well as a cost-effective one) to combine the TMU into the RFS implementation. This arrangement will likely be favored for future flight projects. Therefore, this document includes the CDU and TMU as parts of the RFS description.

The next generation of RFS transponder is now being investigated. Features of the new transponders will probably include an X-band receiver and an adaptive phase-locked loop with digital control.

Aside from the transponder and power-amplifier domains, the RFS has also seen significant changes in the modulations of the information channels. The ranging channel has evolved from its primitive tone-ranging scheme to PN sequence and to the current sequential-ranging implementation. The command and telemetry channels have also progressed from the simple FM system to the advanced pulse-code-modulation schemes. Data rates on both links are much higher than they used to be especially on the telemetry link where megabit capability is now forthcoming. Sophisticated coding techniques are also implemented for the telemetry channels.

7.3 RFS Description

The RFS comprises the following principal subassemblies:

(1) S/X-band transponder

(2) S-band power amplifier

(3) X-band power amplifier

(4) Command detector unit (CDU)

(5) Telemetry modulation unit (TMU)

(6) Control and interface unit

Complementary microwave hardware such as switches, diplexers, filters, and hybrids provide the architecture for the desired RFS configuration.

7.3.1 RFS Block Diagram

A typical RFS block diagram is shown in Fig. 7-1. This configuration is applicable to the dual-antenna spacecraft (usually a low-gain and a high-gain antenna) prevalent in past projects. For other antenna subsystem arrangements, RF switches and diplexers may be changed to provide the proper signal paths. Redundant subassemblies are indicated in the diagram to reflect compliance of the traditional planetary mission's requirement that no single failure shall cause the loss of any required spacecraft function. Failure detection is provided to allow switching to the redundant units.

7.3.2 RFS Functional Operation

As shown in Fig. 7-1, the S-band uplink signal is directed to the transponder receivers from the antenna subsystem via the appropriate diplexers and the RF

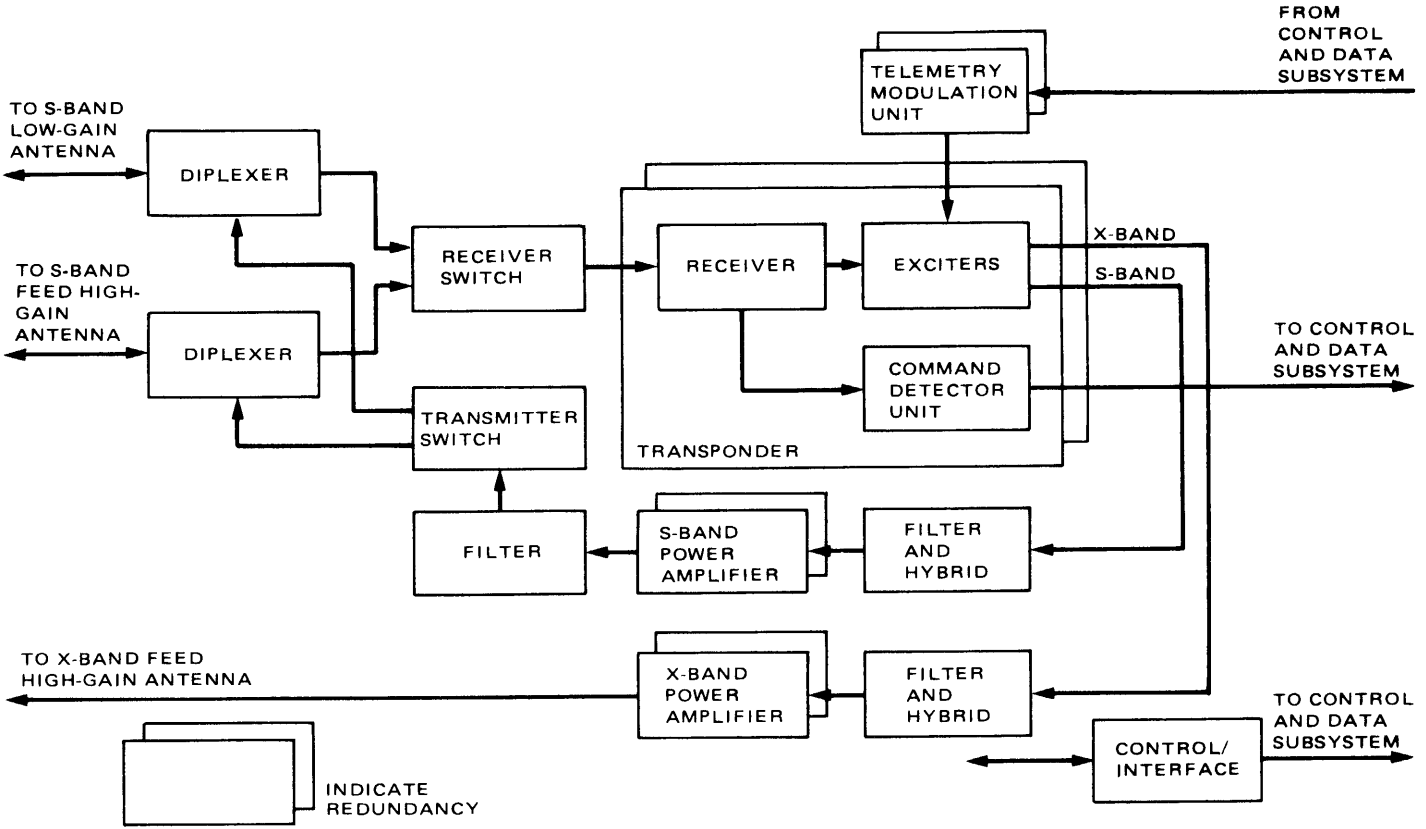


Fig. 7-1. Typical RFS block diagram

switch. The receiver acquires and tracks the uplink carrier by means of a phase-locked loop and produces a reference signal whose phase is coherent with the uplink carrier. With the aid of the reference signal, the ranging and command components of the composite uplink signal are demodulated. The ranging component is delivered to a turnaround ranging channel for downlink modulation, and the command component is delivered to the CDU for subcarrier and baseband detections.

When a coherent downlink transmission is required, the receiver reference signal frequency is multiplied in the exciters to obtain coherent S-band and X-band carriers which are modulated by the turnaround ranging signal and the telemetry signal from the TMU. When coherency is not required, the downlink carrier is derived from an internal frequency source provided by the spacecraft.

Because of the distance involved for planetary missions, the power levels of the exciters are generally inadequate. Sufficient power levels are attained by employing RF power amplifiers such as TWTAs or SSAs. The outputs of the power amplifiers are filtered and directed to the antenna subsystem for radiation via the appropriate RF switches and diplexers.

The RFS housekeeping chores are performed by the control/interface unit. This subassembly sets the configuration of the RFS as commanded by the spacecraft Control and Data Subsystem (CDS), and provides to the CDS the RFS engineering telemetry to form a part of the downlink data package.

7.3.3 RFS Subassembly Description

The following subsection presents detailed descriptions for the major RFS subassemblies.

7.3.3.1 S/X-band transponder. There are two versions of S/X-band transponder currently in use for planetary missions. They are the Voyager-class miniature transponder and the NASA standard transponder. Functionally, the two transponder versions are essentially the same. Both versions accept only S-band uplink and both provide a coherent frequency translation with the ratio 240/221 for S-band downlink to S-band uplink and 880/221 for X-band downlink to S-band uplink. Uplink and downlink data, ranging signal, and carrier turnaround are processed in the transponder. A block diagram showing the typical modules in a transponder and the frequency generation scheme are shown in Fig. 7-2 and Fig. 7-3, respectively. The receiver is implemented as a double-conversion superheterodyne radio. The two local oscillators are generated from a voltage-controlled oscillator (VCO) which is phase locked to the uplink carrier. The downlink carriers are generated by frequency multiplication from the VCO for the coherent mode or from a spacecraft frequency standard for the noncoherent mode.

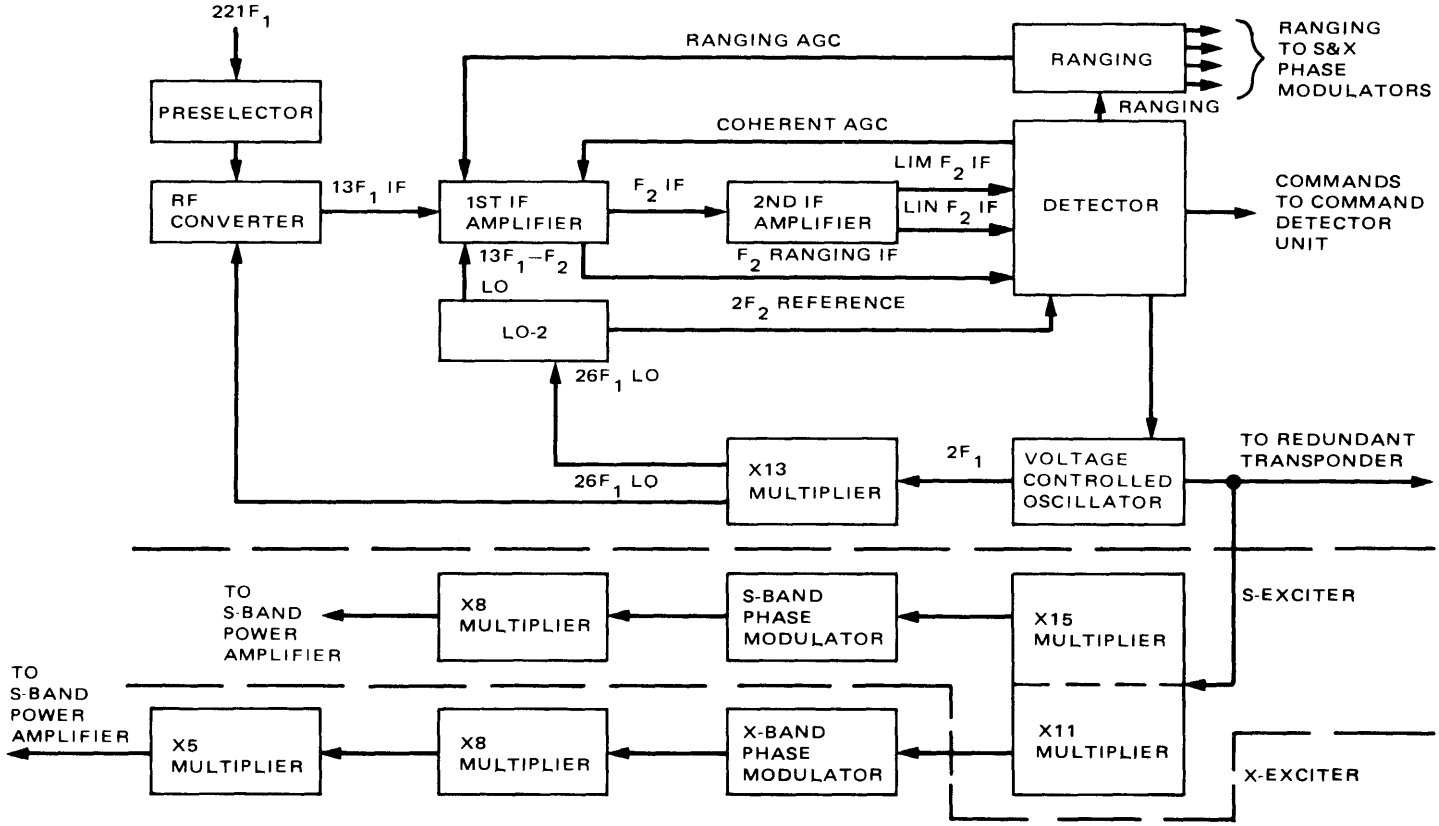


Fig. 7-2. Transponder block diagram

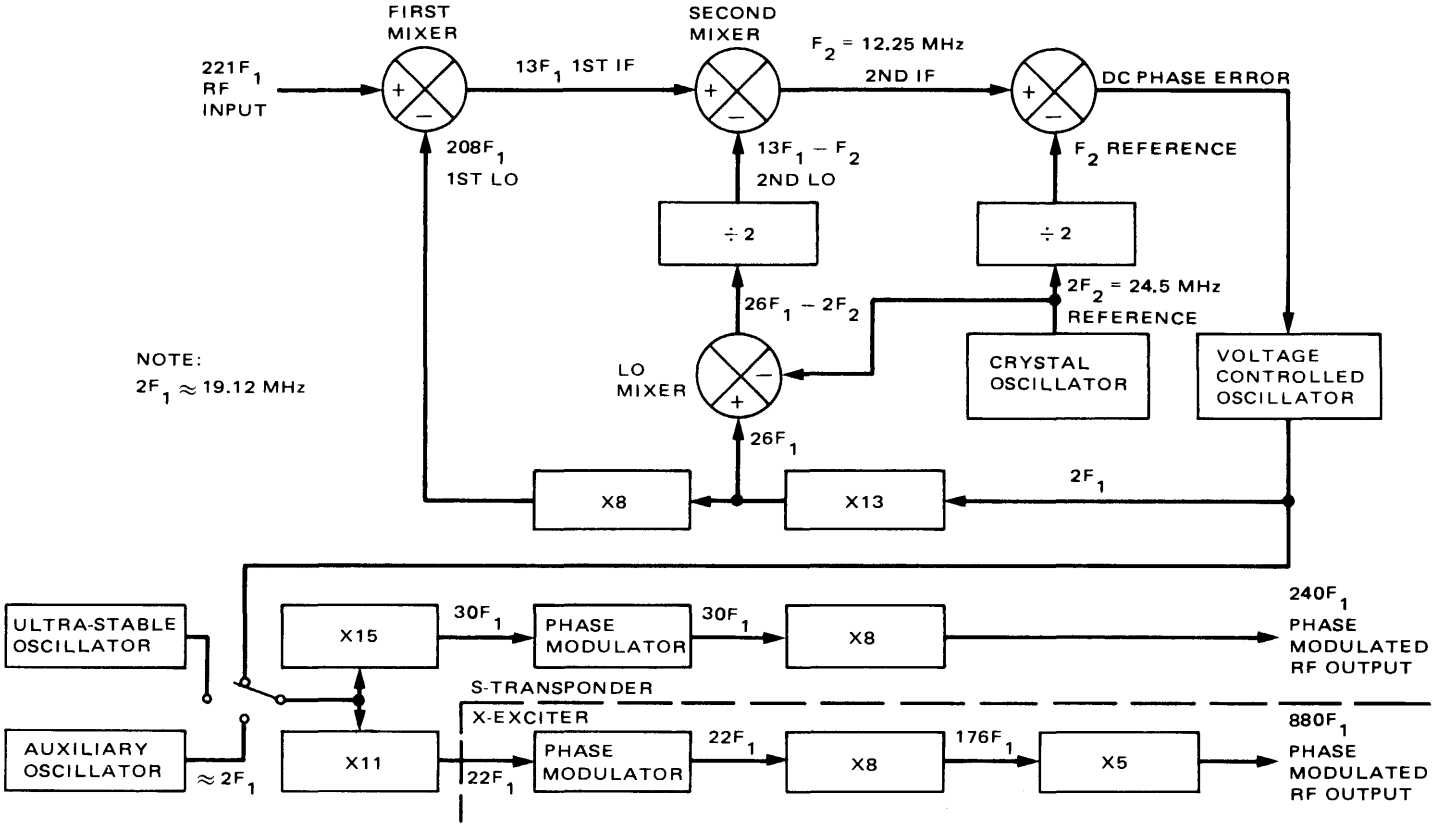


Fig. 7-3. Transponder frequency scheme

7.3.3.1.1 *Transponder receiver.* The transponder's receiver is mechanized as a second-order phase-locked loop (PLL). The loop is locked to the incoming carrier component. The transfer function of the receiver's automatic phase-control loop is designed to be approximately equivalent to the low-pass filter form of

$$H(s)_{\text{design}} = \frac{1 + \dfrac{4\zeta_O^2 + 1}{4B_{LO}} S}{1 + \dfrac{4\zeta_O^2 + 1}{4B_{LO}} S + \dfrac{\alpha_O}{\alpha}\left(\dfrac{4\zeta_O^2 + 1}{8\zeta_O B_{LO}}\right)^2 S^2}$$

where

$2B_{LO}$ = double-sided loop noise bandwidth at carrier threshold

α = limiter suppression factor

α_O = limiter suppression factor at carrier threshold

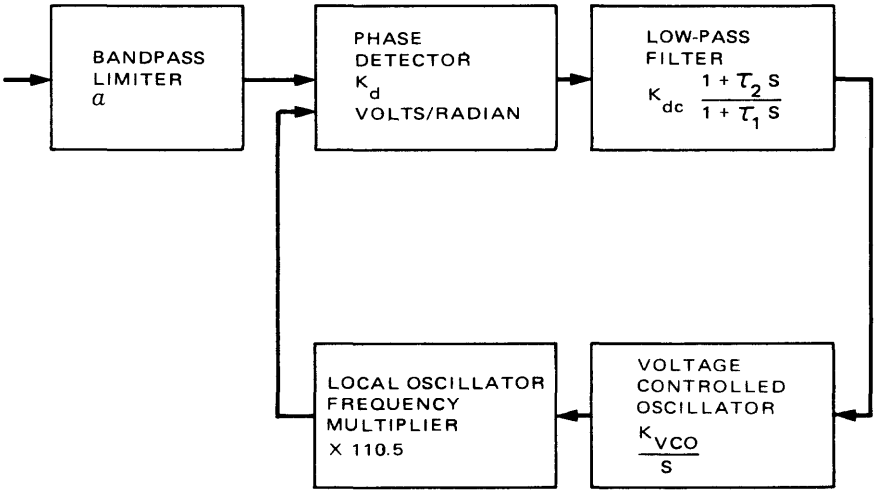
ζ_O = loop damping factor at carrier threshold

The mechanization of this design has a closed-loop transfer function of:

$$H(s) = \frac{1 + \tau_2 S}{1 + \left(\dfrac{1}{K_v} + \tau_2\right) S + \dfrac{\tau_1}{K_v} S^2}$$

where K_v is the total loop gain and τ_1 and τ_2 are the time constants of the loop filter as shown in the simplified loop of Fig. 7-4. The mechanized filter achieves the design if $1/K_v \ll \tau_2$, which is the case in practice.

This mechanization is for maximally flat dampling (or $\zeta_0 = 0.707$) at carrier threshold with $2B_{LO}$ = 18 Hz. (Carrier threshold is defined as zero signal-to-noise ratio in the carrier loop.) A second mechanization that is used in deep space receivers is to set the point for maximally flat damping at a threshold defined by the data error rate in the uplink command data channel. The usual threshold is a command-channel bit-error rate of 1 in 10^5 bits. The transfer function of the loop in terms of ζ and W_n (the natural frequency) is



FOR NASA STANDARD TRANSPONDER

$110.5\ K_d\ K_{dc}\ K_{VCO} = 5.23 \times 10^7$

$\tau_1 = 2910$ sec

$\tau_2 = 0.0833$ sec

NOTE:

τ_2, τ_1 } ARE USED IN TEXT

Fig. 7-4. Simplified receiver phase-locked loop

$$H(S) = \frac{1 + \tau_2 S}{1 + \frac{2\zeta}{W_n} S + \frac{1}{W_n^2} S^2}$$

where

$$W_n = \frac{K_v}{\tau_1} \quad \text{and} \quad \zeta \cong \frac{1}{2} W_n \tau_2$$

The phase-locked loop receiver has been documented in a number of books and papers [7-1 through 7-7].

The limiter suppression factor was used in the above equations without definition. The effect of this variable is to reduce the total loop gain of the receiver's phase-locked loop as the signal-to-noise ratio (SNR) in the bandpass limiter becomes less than one. The closed-loop bandwidth of the receiver is

reduced as the signal-to-noise ratio decreases. Thus, we have an adaptive-loop bandwidth with signal level which, in typical deep space designs, gives more than a 10-to-1 bandwidth reduction from strong signal to threshold. The signal-to-noise ratio and the signal reduction in a bandpass limiter have been analyzed by a number of authors. One of the first was Davenport [7-8] who studied the variation in an output signal as a function of input signal-to-noise ratio. He found that the loss in output SNR is small for any input SNR less than one. The limiting value of this loss is about 1.05 dB.

Tausworthe [7-5] developed a convenient and very close approximation for signal suppression, which is given by:

$$\alpha = \left(\frac{0.7854P_i + 0.4768P_i^2}{1 + 1.024P_i + 0.4678P_i^2} \right)^2$$

where P_i = limiter input signal-to-noise ratio. This works well over the region of interest in most transponder problems. For a complete analysis of the bandpass limiter in a PLL, see the work done by Lindsey [7-2].

The receiver's IF amplification is gain-controlled in the early stages and limited, as discussed above, in the stages before the carrier is detected. Input signal powers may vary from -154 to -70 dBm, giving an 84-dB range. To assure that limiting occurs only in stages designed as limiters, and to provide a linear ranging channel, automatic gain control (AGC) of the carrier signal is used. The AGC error signal is produced by a phase detector whose reference is shifted 90 deg from the carrier PPL phase-detector's reference. The signal used for AGC is picked off before the limiters. The AGC loop acts as a first-order loop with a bandwidth in the order of 1 Hz. For an analysis of the AGC loop see [7-1], [7-2], and [7-4].

7.3.3.1.2 *Ranging channel.* In addition to maintaining a carrier lock, the transponder is also required to handle the ranging signal. The ranging signal is detected by a separate phase detector in phase with the carrier phase-locked loop in a wide-band channel. Due to changes in modulation indexes on the uplink carrier, changes in distance to the spacecraft, and the wide channel bandwidth, the ranging signal is gain-controlled with a limiter or a second AGC loop after it is separated from the carrier channel. The channel separation takes place just after the last stage that is controlled by the carrier AGC.

When AGC is used to control the ranging signal, the total power (signal plus noise) in the ranging channel is detected using a square-law detector. The current transponders use ranging channels with about 1.5×10^6 Hz noise bandwidths, which results in SNRs less than one for most ranging-signal levels. Since

the ranging code is not recovered in the transponder and signal-plus-noise is retransmitted, the ranging signal strength varies as the range to the fourth power when detected back in the DSN.

7.3.3.1.3 *Command channel.* Two command-modulation methods are currently employed which differ in the choice of subcarrier frequency. The Viking-vintage 512-Hz subcarrier version is used with the Voyager-class miniature transponder. The command subcarrier is detected along with the carrier in the phase-locked loop where the command signal is split off before the loop filter. A disadvantage of this method is that the command signal cannot be completely eliminated from the carrier loop and an undesired feedthrough is experienced on the downlink as the reference VCO frequency is being multiplied. The degradation on the X-band downlink due to command feedthrough is quite pronounced. Another disadvantage is that the command bit rate is limited to 256 bps because of the low subcarrier frequency.

The second version is the 16-kHz subcarrier which is used with the NASA standard transponder. A command rate of 2 kbps is possible and the command feedthrough problem is eliminated. The command signal, however, must be detected in a separate channel in the same manner as ranging. The detected composite command signal is routed to a command detector for data recovery.

7.3.3.1.4 *Exciters.* There are two exciters (an S-band and an X-band) used in the transponder and they employ the same design concept. An exciter is a series of frequency multipliers which bring the transponder reference frequency to the desired downlink frequency. Within the multiplier chain, a phase modulator is implemented to permit modulation of the ranging and/or the telemetry data. Figure 7-2 shows the construction (schematically) of the two exciters. The modulating signals are first combined by a summing amplifier before being applied to the phase modulator. In the S-band and the X-band exciters, the phase modulators are operating in the low "hundred-MHz" region which allows low indexes to be modulated. The frequency multiplication that follows increases the indexes by the factor of the multiplication ratio. In this fashion, the transponder is able to provide modulation of square wave in excess of 90 deg.

7.3.3.2 Power amplifiers. As pointed out previously, two types of RF power amplifiers have been and continue to be used for planetary missions. They are the TWTA and the SSA. Despite the tedious and sometimes difficult procedures in trimming the amplifiers to the required specifications, an RF amplifier is in fact, configuratively speaking, a relatively simple device. For the TWTA, it consists of a power supply to provide the various voltages required for TWT operations and a traveling-wave tube which performs the RF power amplification. The SSA consists of a power converter and a solid-state amplifier. Some supplementary circuits are provided for power switching and other housekeeping purposes.

The RF amplifiers are used to raise the exciter outputs to a power level sufficient for RF transmission. Different amplifiers must be designed for different frequency bands. At the present time, S-band TWTAs up to 28 watts, X-band TWTAs up to 20 watts, and S-band SSAs up to 20 watts have all been flown.

7.3.3.3 Command detector unit. As with the case of the transponder, there are also two command detector versions currently in use for planetary missions. They are the Viking Heritage CDU which is compatible with the Voyager-class transponder and the NASA standard CDU which was developed with and became an integral part of the NASA standard transponder. Since both transponder-CDU combinations will be used in future missions, both CDUs are described.

7.3.3.3.1 *Viking Heritage CUD.* The Viking CDU is the first generation of all-digital bit synchronizers used for spacecraft command-detection purposes. The only analog elements in the detector are the front-end signal-conditioning circuits and the analog-to-digital converter. The command baseband signal for the Viking CDU consists of a 512-Hz subcarrier modulated by the bit sync which in turn is modulated by the data bits. The composite command signal may be represented by

$$\textit{composite command signal} = SC \oplus BS \oplus D$$

where

SC is the 512-Hz subcarrier

BS is the $512/_n$-Hz bit sync or clock, $n = 2$ to 512 in binary steps

D is data bits

In order to detect the command bits, two levels of synchronization, subcarrier sync and clock sync, must first be established. To accomplish these synchronizations, the transmitted commands are prefixed first by a short burst of unmodulated subcarrier, followed by a period of subcarrier modulated by bit sync. Hence, the detections of subcarrier, bit sync, and data are performed sequentially with a multimode correlation process. In other words, the CDU first establishes subcarrier lock when the unmodulated subcarrier is present; then it uses the detected subcarrier to extract the bit sync during the period when no data ($SC \oplus BS$ only) are present. Finally, the knowledge of SC and BS enables the CDU to detect data bits when they subsequently appear. The three-level (SC, BS, and data) detection process is accomplished by means of digital correlation. A block diagram for the CDU and a detection flow diagram are given in Fig. 7-5 and Fig. 7-6, respectively.

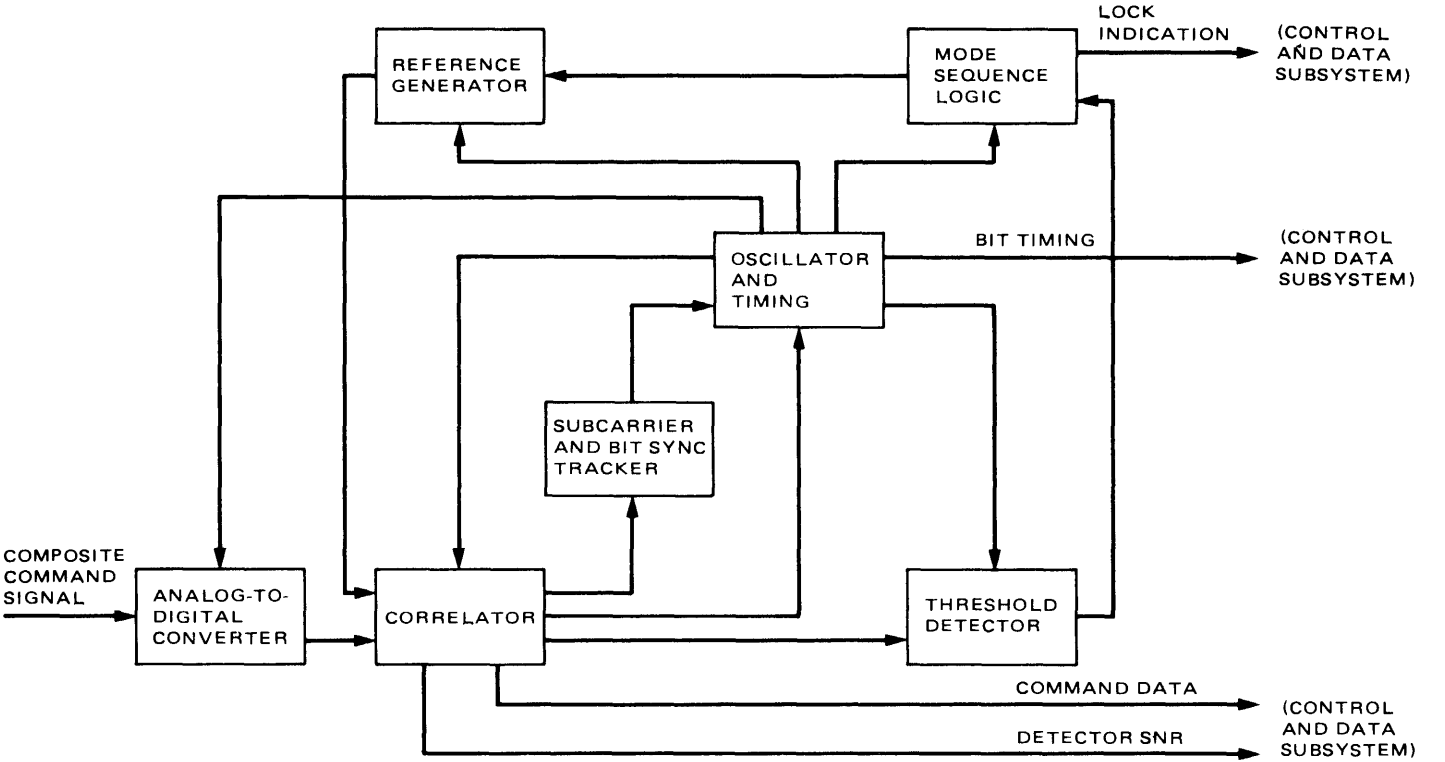


Fig. 7-5. Viking Heritage CDU functional block diagram

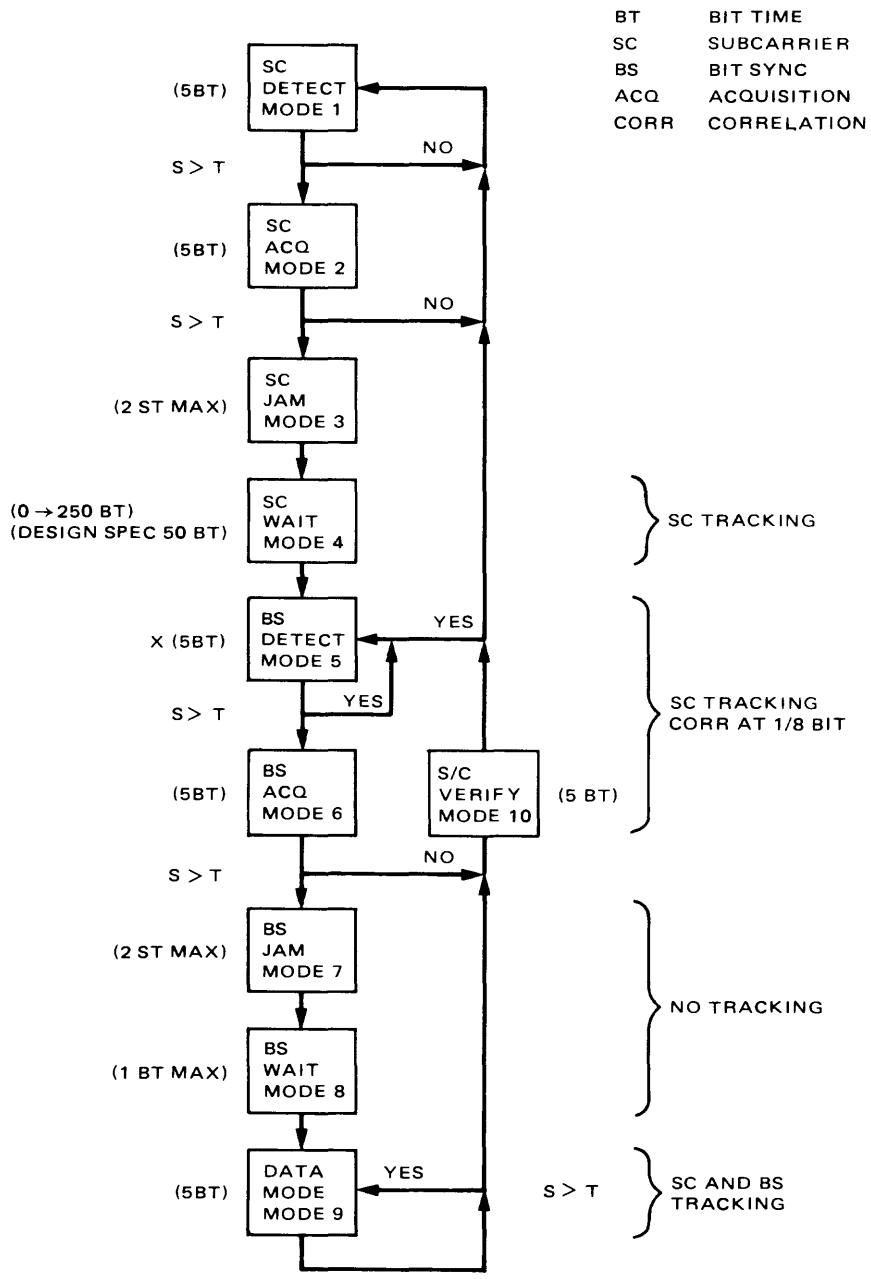


Fig. 7-6. Viking Heritage CDU mode flow diagram

7.3.3.3.2 *NASA standard CDU.* Like its Viking counterpart, the NASA standard CDU is an all-digital bit synchronizer. Unlike the Viking CDU, the NASA standard composite command signal consists of a 16-kHz subcarrier modulated by data only. Bit sync estimates are derived based on the knowledge of data transitions. Detections of subcarrier, bit sync, and data are performed simultaneously with a single correlation process. To facilitate acquisition, the transmitted commands are prefixed by a string of alternate 1s and 0s. Because of the higher subcarrier frequency, the maximum bit rate for the NASA standard CDU is 2 kbps.

The NASA standard CDU is constructed with dedicated MOS-LSI chips organized in the form of a microprocessor. Quadrature samplings are performed directly at the composite command signal to detect the zero crossings of the subcarrier from which the subcarrier phase lock is established. Correlations of subcarrier, bit sync, data, and in-lock indication are performed by the microprocessor in a manner similar to that of the Viking CDU. A block diagram for the NASA standard CDU is shown in Fig. 7-7.

7.3.3.4 Telemetry modulation unit. The current TMU design owes its origin to the Viking Modulation Demodulation Subsystem. Perhaps because of its relative simplicity, there have been no attempts to standardize the TMU or to reconstruct the unit with LSI circuitry. Therefore, within its basic structure, the TMU may be tailored to suit a specific project's needs.

In general, the TMU consists of a convolutional coder which may or may not be used, a subcarrier generator, modulators, and modulation-index select circuits. Two telemetry channels are normally provided to offer data mode flexibility. A typical TMU block diagram is given in Fig. 7-8.

In terms of the receiving station, the TMU provides two telemetry modes, Tracking and Data Relay Satellite (TDRS) and Deep Space Network (DSN). The fundamental difference between the two modes is simply that the subcarrier is not used in the TDRS mode.

7.3.3.5 Microwave hardware. A set of microwave components is normally required in an RFS. These microwave components include diplexers, filters, hybrids, and switches. The diplexer is used to combine the uplink and downlink signals into a single signal. It is a set of filters that prevents the transmitter power from saturating the receiver input and the wideband noise out of the power amplifiers from degrading the receiver SNR. Most diplexers use a bandpass filter at the uplink frequency in the receiver arm that rejects the transmitter power. A band-reject filter at the uplink frequency in the transmit arm prevents power amplifier noise from degrading the receiver SNR.

The hybrid is a power divider which couples energy from multiple sources to multiple loads. It is constructed with the conventional transmission-line

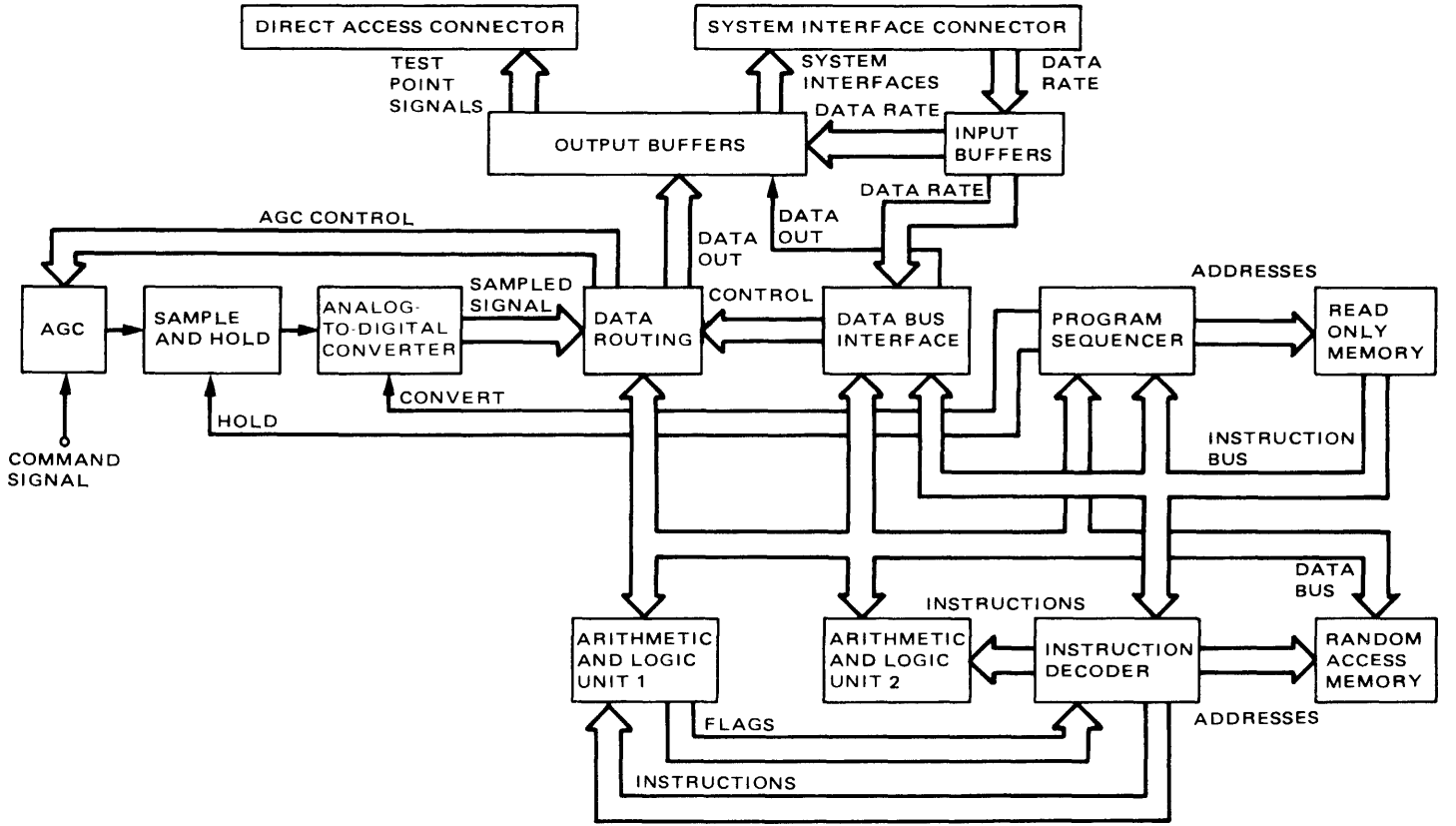


Fig. 7-7. NASA standard CDU block diagram

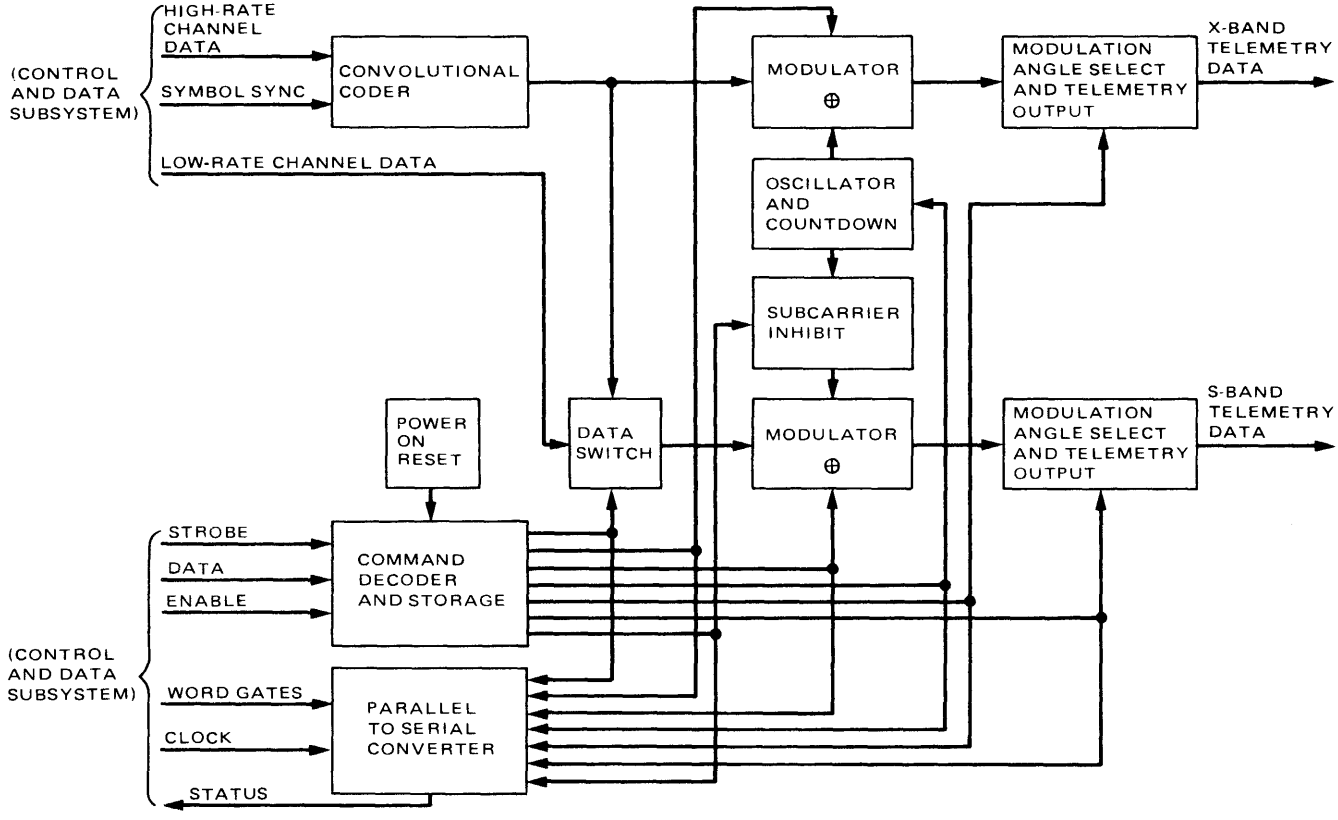


Fig. 7-8. Typical TMU functional block diagram

principles and is traditionally used to cross-strap the redundent exciter outputs and the redundent power amplifier inputs. The switch is a signal-directing device. Ferrite circulating switches are commonly used in the RFS. In general, there is no standard configuration for the microwave components, as they are mission-dependent.

7.4 RFS Subassemblies Performance Specifications

The following section compiles the performance specifications for the major RFS subassemblies. The data are provided from [7-9] through [7-15].

7.4.1 RFS Transponder Specifications

The specifications for elements of the RFS transponder are listed in the tables which follow.

7.4.1.1 Carrier channel. The parameters and specifications for the carrier channel are tabulated below.

Parameter	Specification
Uplink frequencies	Any NASA-designated S-band channel
Coherent transponding	S-band: 240/221 times uplink frequency X-band: 880/221 times uplink frequency
Noncoherent operation (one-way downlink)	S-band: 120 times internal oscillator frequency
Internal oscillator frequency	19.1 MHz depending on channel assignment
Receiver frequency setting accuracy	±2 ppm at 24 ± 2°C
Receiver frequency stability	
Short-term (15 hr)	±1.5 ppm at any constant temp ±0.5°C in range of 10 to 40°C
Aging (1 yr)	±3 ppm at any constant temp ±0.5°C in range of 10 to 40°C
Temperature	±0.4 ppm per °C, 10 to 40°C ±14 ppm per °C, -10 to 55°C (flight acceptance range)

Parameter	Specification
	±20 ppm per °C, -20 to 65°C (type approval range)
Receiver signal level	-154 to -70 dBm
Antenna terminal input impedance	50 ohm, VSWR ≤ 1.5 to 1
Receiver preselector input impedance	50 ohm, VSWR ≤ 1.2 to 1
Receiver noise figure	< 5.5 dB with 50 ohm VSWR ≤ 1.1 to 1
PLL noise bandwidth ($2B_{LO}$)	18 Hz ±10%, 10 to 40°C 18 Hz ±20%, type approval range
PLL filter damping factor (ζ)	0.707 SNR = 1 in $2B_{LO}$
Predetector bandwidth	2-stage, single-pole filters 250 kHz ±20% cascaded with 3 kHz ±10%
Tracking threshold	-154 dBm min at best lock frequency
Tracking capture range	±1.3 kHz, ≥ -90 dBm carrier
Tracking range	±150 kHz min, ≥ -90 dBm carrier
Acquisition and tracking rate	600 Hz/sec min, ≥ -90 dBm carrier
Tracking error (SPE)	< 1 deg per 30 kHz, ≥ -90 dBm carrier
Carrier AGC loop bandwidth (2-sided)	1.8 ± 0.9 Hz
Carrier AGC linearity	Within 10% from -140 to -70 dBm
VCO frequency	19.1 to 19.2 MHz determined by uplink frequency
Internal oscillator frequency setting accuracy (for one-way)	±2 ppm at 24 ±2°C

Parameter	Specification
Internal oscillator frequency stability	
Short-term	2×10^{-9} rms fractional frequency deviation of 10 min, with 10 sec integration time at any constant temp ±0.5°C in range of 10 to 40°C
Aging (1 yr)	±3 ppm at any constant temp ±0.5°C in range of 10 to 40°C
Temperature	±0.3 ppm per °C, 10 to 40°C ±10 ppm per °C, flight acceptance range ±15 ppm per °C, type approval range
Exciter residue phase noise	3° rms max at S-band 11° rms max at X-band
S-band exciter output power	+21 dBm ± 1 dB flight acceptance +21 dBm ± 2 dB type approval at 50 ohm resistive load
S-band exciter output impedance	50 ohm nominal VSWR ⩽ 1.5 to 1
X-band exciter ouput power	+12 dBm ±1 dB flight acceptance +12 dBm ±2 dB type approval at 50 ohm resistive load
X-band exciter output impedance	50 ohm nominal VSWR ⩽ 1.5 to 1

7.4.1.2 Ranging channel. The parameters and specifications for the ranging channel are tabulated below.[1]

Parameter	Specification
Linearity	5%
Channel nonsymmetry	2%, 10 to 40°C 5%, flight acceptance range 10%, type approval range

[1] All specifications are under specific test conditions

Parameter	Specification
Downlink RG modulation index	0.2 to 1 rad peak
Downlink RG modulation index	
Stability	±5%, 10 to 40°C ±10%, flight acceptance range ±13%, type approval range
RG channel frequency response (-3 dB)	1 kHz to 1.2 MHz
RG delay variation	
S-band up, S- or X-band down	< 20 nsec peak-to-peak, 10 to 40°C < 30 nsec peak-to-peak, flight acceptance range < 40 nsec peak-to-peak, type approval range
Differential S- and X-band down	< 8 nsec peak-to-peak, 10° to 40°C < 12 nsec peak-to-peak, flight acceptance range < 16 nsec peak-to-peak, type approval range
Phase delay variation	
S-band up, S- or X-band down	< 1.5 nsec peak-to-peak, 10 to 40°C < 3 nsec peak-to-peak, flight acceptance range < 3.5 nsec peak-to-peak, type approval range
Differential S- and X-band down	< 0.75 nsec peak-to-peak, 10 to 40°C < 1.5 nsec peak-to-peak, flight acceptance range < 1.8 nsec peak-to-peak, type approval range

7.4.1.3 Telemetry channel. The parameters and specifications for the telemetry channel are tabulated below.

Parameter	Specification
TLM channel frequency response	100 Hz to 3 MHz (-3 dB) 1.5 to 500 kHz (flat within 1 dB)

Parameter	Specification
Input signal level	1.8 V rms
Input impedance	50 ± 5 ohm, 100 pf shunt max
Asymmetry	1% max up to 2 rad mod index
Amplitude nonlinearity	2% max to 1 rad mod index 5% max to 2 rad mod index
Modulation sensitivity	2 rad peak/volt peak
Modulation index stability	±5% flight acceptance range ±7% type approval range

7.4.1.4 **Command channel.** The parameters and specifications for the command channel are tabulated below.

Parameter	Specification
Implementation	Voyager-class external to transponder NASA standard internal to transponder
All other parameters and characteristics listed under CDU.	

7.4.2 Power Amplifiers Specifications

The specifications for elements of the power amplifiers are listed in the tables which follow.

7.4.2.1 **20-watt S-band SSA.** The parameters and specifications for the 20-watt S-band SSA are tabulated below.

Parameter	Specification
Primary DC voltage source	30 ± 3 V
Primary voltage ripple	25 mV peak-to-peak max
Primary voltage source over-voltage	15% max

Parameter	Specification
Primary DC power	33.5 W max low power 87.5 W max high power
Power converter efficiency	Not specified
RF amplifier efficiency	24% min low power 27% min high power
Input frequency	2,290 to 2,302 MHz
Input carrier modulation	Phase modulated square wave
Modulation bandwidth (one-sided)	3 MHz
Input carrier suppression	19 dB from unmodulated level
Input power level	3 to 7 dBm
Input power level variation	±1.5 dB
Input impedance	50 ohms VSWR < 1.3 to 1
Output impedance	50 ohm VSWR < 1.3 to 1
Output power level	6 W min low power 20 W min high power
Output power variation	
– temperature	0.2 dB per 10°C
– frequency	±0.1 dB (input frequency range)
– short term	±0.1 dB (30-min interval)
– aging	Degraded < 0.4 dB (5 yr)
Output amplitude modulation	0.5% mod index = 83°
Output spectral purity	2nd harmonic 30 dB down 3rd harmonic 40 dB down > 3rd harmonic 40 dB down
Output noise power spectral density	-35 dBm/MHz (2,013 to 2,213 MHz) -30 dBm/MHz (2,245 to 2,345 MHz)

Parameter	Specification
Phase stability	
– input level	14.4 psec per dB
– primary voltage	3 psec with 1 V change
– temperature	12 psec per 10°C
– frequency	0.0003 psec per 1 kHz (input frequency range)
– short term	0.5 psec over 12 hr
Group delay stability	
– input level	0.1 nsec per dB
– primary voltage	0.1 nsec with 1 V change
– temperature	0.1 nsec per 1°C
– frequency	0.00001 nsec per 1 kHz (input frequency range)
– long term	0.1 nsec over 5 yr

7.4.2.2 **28-watt S-band TWTA**. The parameters and specifications for the 28-watt S-band TWTA are tabulated below.

Parameter	Specification
Primary dc voltage source	30 ± 3 V
Primary voltage ripple	25 mV peak-to-peak maximum
Primary voltage source over-voltage	15% max (20 msec or less)
Primary dc power	33 W ±6% low power 86 W ±6% high power
Power converter efficiency	88% min low power 90% min high power
RF amplifier efficiency	38% min low power 41% min high power
Input frequency	2,290 to 2,302 MHz
Input carrier modulation	Phase modulated square wave

Parameter	Specification
Modulation bandwidth (one-sided)	3 MHz
Input carrier suppression	19 dB from unmodulated level
Input power level	3 to 7 dBm
Input power level variation	±1 dB
Input impedance	50 ohm VSWR < 1.3 to 1
Output impedance	50 ohms VSWR < 1.3 to 1
Output power level (worst case)	9.5 W low power 26.7 W high power
Output power variation	
– temperature	0.1 dB per 10°C (0 to 55°C)
– frequency	±0.1 dB (input frequency range)
– short term	±0.1 dB (30-min interval)
– aging	Degraded < 0.4 dB (5 years)
Output amplitude modulation	0.5% mod index = 83°
Output spectral purity	2nd harmonic 10 dB below fundamental 3rd harmonic 20 dB below fundamental >3rd harmonic 30 dB below fundamental Total 9.5 dB below fundamental
Output noise power spectral density	-35 dBm/MHz (2,013 - 2,213 MHz) -30 dBm/MHz (2,245 - 2,345 MHz)
Phase stability	
– input level	14.4 psec per dB
– primary voltage	3 psec with 1 V change
– temperature	12 psec per 10°C
– frequency	0.0003 psec per 1 kHz (input frequency range)
– short term	0.5 psec over 12 hr

Parameter	Specification
Group delay stability	
– input level	0.1 nsec per dB
– primary voltage	0.1 nsec with 1 V change
– temperature	0.1 nsec per 1°C
– frequency	0.00001 nsec per 1 kHz (input frequency range)
– long term	0.1 nsec over 5 yr
Warmup time	180 sec max.

7.4.2.3 20-watt X-band TWTA. The parameters and specifications for the 20-watt X-band TWTA are tabulated below.

With the following exception, the 20-W X-band TWTA has the same specifications as indicated in 7.4.2.2.

Parameter	Specification
Primary dc power	47 W ±6% low power 65 W ±6% high power
Power converter efficiency	89% min low power 90% min high power
RF amplifier efficiency	38% min low power 45% min high power
Input frequency	8,400 - 8,442 MHz
Input power level	0 to 4 dBm
Output power level (worst case)	12.7 W low power 21.6 W high power
Output power variation – aging	Degraded < 0.4 dB (10 years)
Output noise power spectral density	-30 dBm/MHz (8,315 - 8,515 MHz)
Phase stability – input level	3.3 psec per dB

7.4.3 Command Detector Unit Specifications

The specifications for elements of the command detector unit are listed in the tables which follow.

7.4.3.1 **NASA standard CDU.** The parameters and specifications for the NASA standard CDU are tabulated below.

Parameter	Specification
Modulation	Non-return-to-zero bi-phase-modulated
Subcarrier	16 kHz sine wave
Bit-rates	$2000/2^n$, $n = 0, 1, 2, \ldots\ldots, 8$
Bit error rate	$\leqslant 1 \times 10^{-5}$ at command threshold
Command threshold (E_b/N_0)	9.6 dB, -1.0 dB + 1.1 dB
Lock detection	
– probability of false acquisition	$\leqslant 1 \times 10^{-4}$
– probability of deacquisition	> 0.98 within 27 bits of command removal
– probability of out-of-lock indicate	$\leqslant 2.5 \times 10^{-9}$ at command threshold
– probability of failure to acquire	$\leqslant 1 \times 10^{-4}$ at command threshold
Maximum frequency offset for acquisition and tracking	1.95 times bit rate at command threshold
Maximum frequency offset rate for acquisition and tracking	0.396 times (bit rate)^2 at command threshold
Acquisition sequence	132 bits of alternate 1s and 0s
Probability of acquisition	0.9999 at bit rate = 31 bps
Command data output	Non-return-to-zero data (polarity unresolved)
Command bit sync	Square wave at data rate
Lock indicator	Bilevel signal

7.4.3.2 Viking Heritage CDU. The parameters and specifications for the Viking Heritage CDU are tabulated below.

Parameter	Specification
Modulation	Manchester coded bi-phase-modulated
Subcarrier	512-Hz square wave
Bit rates	1 to 256 bps in binary steps
Bit error rate	$< 1 \times 10^{-5}$ at command threshold
Command threshold (E_b/N_0)	9.6 dB, -1.0 dB + 1.1 dB
Lock detection	
– probability of false acquisition	$< 1 \times 10^{-8}$
– probability of out-of-lock indicate	$\leqslant 1 \times 10^{-5}$ at command threshold
– probability of failure to acquire	$\leqslant 1 \times 10^{-4}$ at command threshold
Frequency offset	±0.04%
Acquisition sequence	Subcarrier only (65 bit times), Subcarrier $\oplus BS$ (15 bit times)
Probability of acquisition failure	5.9×10^{-6}
Command data output	Non-return-to-zero data sequence
Command bit sync	Square wave at data rate
Lock indicator	Bilevel signal

7.4.4 Telemetry Modulation Unit Specifications

The specifications and parameters for the telemetry modulating unit are listed in the table which follows.

Parameter	Specification
Modulation	Non-return-to-zero or convolutional coded Non-return-to-zero, biphase modulation
Subcarrier	Low rate 1.5 times data rate min square wave High rate 1.5 times data symbol rate min
Data rates	Low rate 40 bps uncoded for Voyager High rate up to 134 kbps
Subcarrier frequency tolerance	Initial frequency set $\leqslant 0.005\%$
Subcarrier frequency stability	0.002% over flight acceptance 0.0005%, 12 hr, $\pm 5^{\circ}$C
Subcarrier phase jitter	Low rate, 0.2 rad rms in 0.03-Hz BW High rate, 0.2 rad rms in 2-Hz BW

References

7-1. Gardner, F. M., *Phase Lock Techniques*, 2nd ed., Wiley, New York, 1979.

7-2. Lindsey, W. C., *Synchronization Systems in Communication and Control*, Prentice-Hall, Englewood Cliffs, N.J., 1972.

7-3. Blanchard, A., *Phase-Locked Loops*, Wiley, New York, 1976.

7-4. Lindsey, W. C., and Simon, M. K., *Phase-Locked Loops and their Applications*, IEEE Press, New York, 1978.

7-5. Tausworthe, R. C., *Theory and Practical Design of Phase-Locked Receivers*, Technical Report 32-819, Vol. 1, Jet Propulsion Laboratory, Pasadena, Calif., 1966.

7-6. Gupta, S. G., "Phase-Locked Loops," *Proceedings of the IEEE*, Vol. 63, pp. 291-306, Feb. 1975.

7-7. Lindsey, W. C., and Chie, C. M., "A Survey of Digital Phase-Locked Loops," *Proceedings of the IEEE*, Vol. 69, pp. 410-431, Apr 1981.

7-8. Davenport, W. B., Jr., "Signal-to-Noise Ratios in Band-Pass Limiters," Journal of Applied Physics, 24, No. 6, pp. 720-727, June 1953.

7-9. NASA Standard Deep Space Transponder Design Requirement (STD-336-MOI-DS), Jet Propulsion Laboratory, Pasadena, Calif., 1978 (an internal document).

7-10. VO'75 Modulation Demodulation Subsystem Engineering Report, Motorola, Sept. 1972.

7-11. Galileo Orbiter Radio Frequency Subsystem Design Requirement (GLL-2002-1), Jet Propulsion Laboratory, Pasadena, Calif., June 1979 (an internal document).

7-12. Galileo Orbiter Modulation Demodulation Subsystem Functional Requirement (GLL-4-2003), Jet Propulsion Laboratory, Pasadena, Calif., 1979 (an internal document).

7-13. Dual Power S-band Solid-State Amplifier Equipment Specification (ES 511745), Jet Propulsion Laboratory, Pasadena, Calif., Aug. 1976 (an internal document).

7-14. Dual Power S-Band Traveling Wave Tube Amplifier Equipment Specification (ES 509347), Jet Propulsion Laboratory, Pasadena, Calif., Apr. 1974 (an internal document).

7-15. Dual Power X-Band Traveling Wave Tube Amplifier Equipment Specification (ES 509348), Jet Propulsion Laboratory, Pasadena, Calif. Apr. 1974 (an internal document).

Chapter 8
Spacecraft Antennas

Yahya Rahmat-Samii and Gary K. Noreen

8.1 Introduction

The theoretical foundations of antennas rest on Maxwell's equations presented before the Royal Society of London in 1864 by James Clerk Maxwell (1831-1879). These equations have proven to be one of mankind's greatest discoveries for predicting and describing nature's electromagnetic phenomena. In 1886 Heinrich Hertz (1857–1894) was able to verify experimentally the existence of electromagnetic phenomena and their propagation characteristics. This added a new chapter to human history since the electromagnetic spectrum, outside the visible region, could be employed for communication purposes through the use of radio. Figure 8-1 presents the names of some renowned scientists who contributed to the advancements and applications of the theory of electromagnetism in the era spanned from Gilbert to Einstein [8-1].

A radio antenna is a device for radiating or receiving electromagnetic waves which provides a transition between a free-space wave and a guided wave (receiving) or vice versa (transmitting). The electromagnetic characteristics of antennas can be accurately predicted using Maxwell's equations. Since the science of antennas is mature, there are numerous publications devoted to this topic and the interested reader is referred to some of the most recent ones

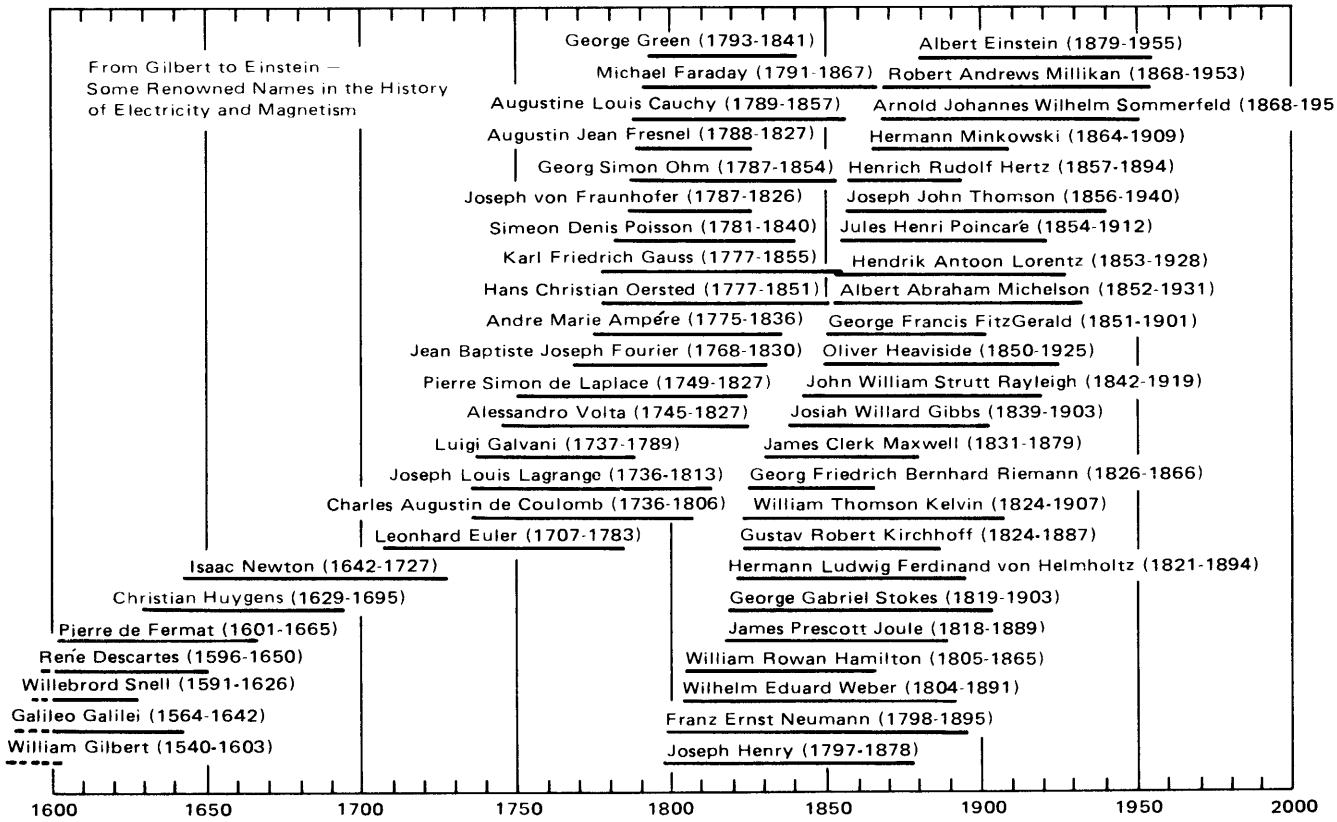


Fig. 8-1. From Gilbert to Einstein—some renowned names in the history of electricity and magnetism [8-1]

[8-2 through 8-8]. As technological needs have demanded more advanced antennas, many newly developed analytical, numerical and experimental methods have emerged, and a variety of antenna configurations have been built and examined. Among them one may refer to linear antennas, log periodic antennas, horn antennas, reflector antennas, arrays, lenses, microstrip antennas, etc. Obviously, not all of these can be examined here, and it is the responsibility of the design engineer to keep himself updated to make sure that the choice of antenna is optimum for the task.

In this section, a tutorial description is given for spacecraft antennas used for deep-space-to-earth communication. For this reason, this discussion is focused primarily on reflector antennas as they are used most frequently in deep-space spacecraft. The objective is to present some fundamental definitions, tables, and graphs to enable the system engineer to originate a preliminary design for a specified task. The discussion is by no means exhaustive, and the user must consult more detailed sources and references to refine a design.

8.2 Radiation Pattern Parameters

8.2.1 Radiation Pattern, Beamwidths and Sidelobes

To permit the transfer of information between a spacecraft and earth, the radio signal is radiated into space by an antenna. The radiation characteristics of an antenna as a function of direction are given by the antenna pattern. By the principle of reciprocity, this pattern is the same for both receiving and transmitting conditions in most circumstances. The simplest antenna, hypothetically speaking, is the isotropic radiator, which distributes the radiated power uniformly into space. Thus, for a total radiated power of W_r, one observes a radiation intensity of

$$U_0 = \frac{W_r}{4\pi} \tag{8.2-1}$$

per unit solid angle independent of the angular position with respect to this hypothetical antenna. However, with a known target direction or directions, it becomes desirable to concentrate the radiated power into specific areas. For this purpose, a directional antenna may be designed so that the radiation intensity is a function of the direction.

If the radiation intensity pattern is measured at a sufficient distance from the antenna (typically larger than $2D^2/\lambda$, where D = antenna diameter and λ = wavelength) so that an increase in distance causes no change in the relative intensity, the pattern will be the far-field pattern. Closer measurements yield Fresnel and near-field patterns which are functions of both angle and distance.

For instance, Fig. 8-2 shows how the radiation pattern can change as a function of the distance from an antenna [8-9]. Notice that after a large distance ($2D^2/\lambda$), the pattern does not change appreciably, which indicates a far-field pattern. This is an important observation since the use of increasingly large antennas (in terms of wavelength) makes it difficult to measure their true far-field patterns using standard far-field ranges. In the discussion of this section, only the far-field patterns are considered.

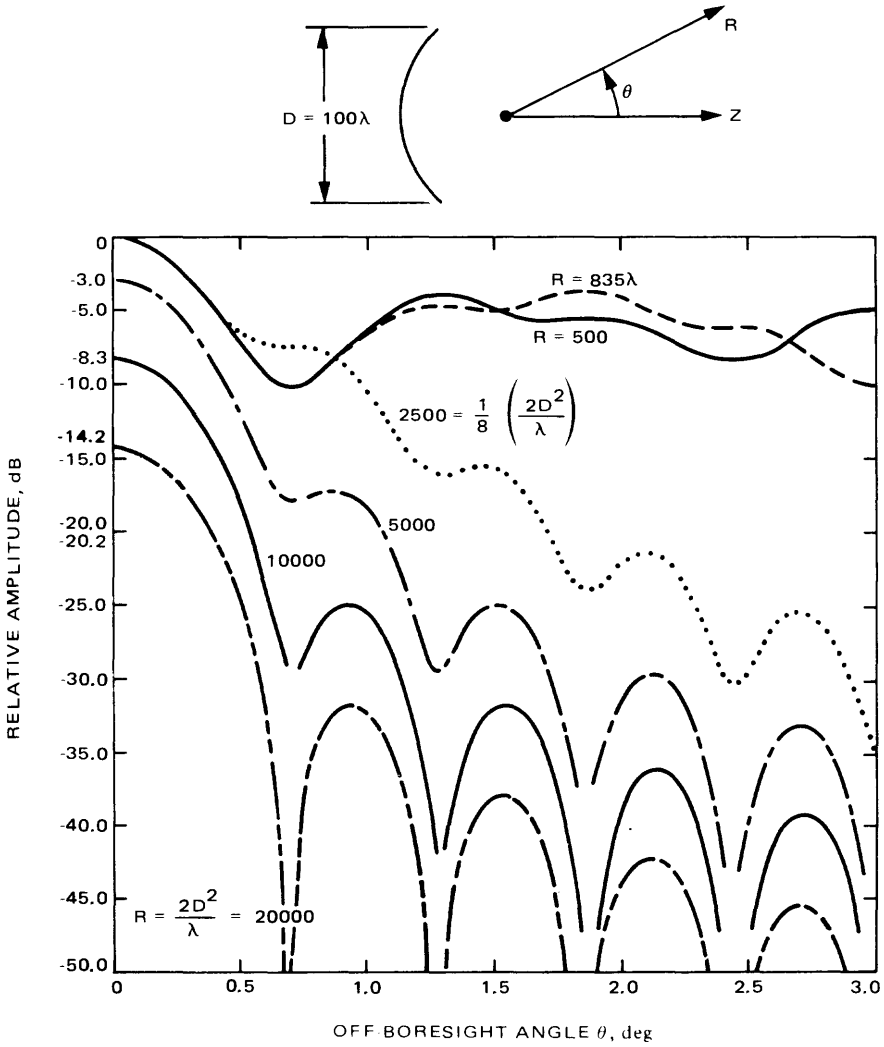


Fig. 8-2. Evolution of reflector antenna pattern as a function of distance *R*. At a distance of $R \geq 2D^2/\lambda$, the far-field pattern is formed

The pattern may be expressed in terms of the field intensity (field pattern) or in terms of the Poynting vector or radiation intensity (power patterns). The latter is used most often. A typical antenna power pattern is shown in Fig. 8-3 as a polar plot in linear units, and is shown in Fig. 8-4 in rectangular coordinates on a dB scale. The latter presentation shows the minor-lobe structure in more detail. In most cases, the main lobe is the desired lobe and, thus, the sidelobes are the minor lobes. Typically, the sidelobes have alternately positive and negative values. In fact, the far field in its most general form may be complex-valued and its magnitude and phase will then be used to define the power and phase patterns. A single pattern (single cut) would be sufficient to completely specify the angular variation in the radiation pattern provided the pattern is symmetric. If the pattern is not symmetric, a three-dimensional diagram or contour plot will be required to show the pattern in its entirety. However, in practice, two patterns (one through the narrowest part of the lobe; the other, perpendicular to it through the widest part of the lobe) may suffice. These mutually perpendicular patterns through the main-lobe axis are called the principal-plane patterns and, based on their polarizations, sometimes are referred to as the E- and H-plane patterns. Since one encounters vector fields when working with antennas, there are occasions for referring to the copolar and cross-polar patterns in order to demonstrate two orthogonally polarized components of the far field.

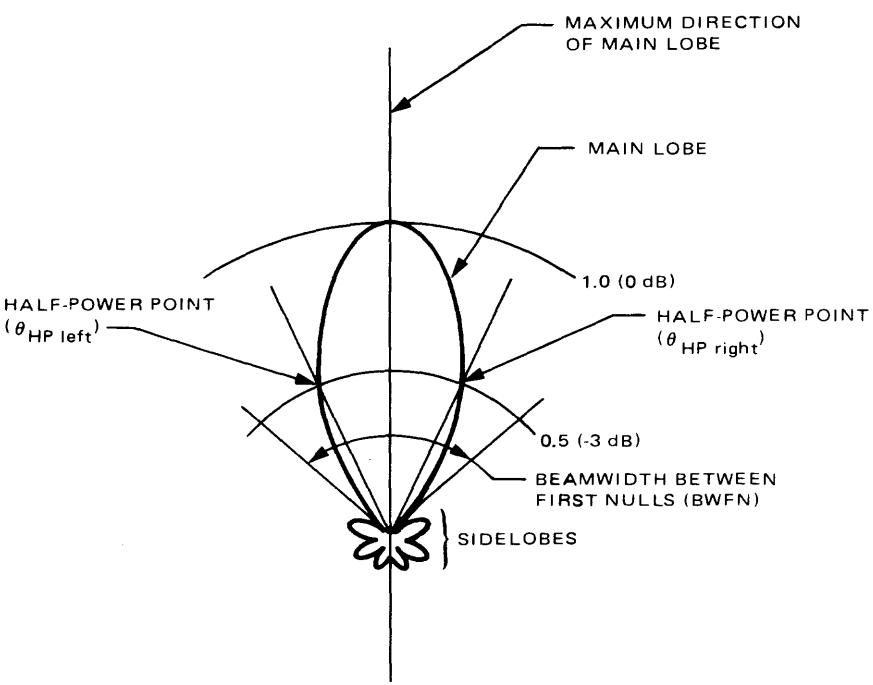


Fig. 8-3. Normalized antenna pattern in polar coordinates and linear power scale

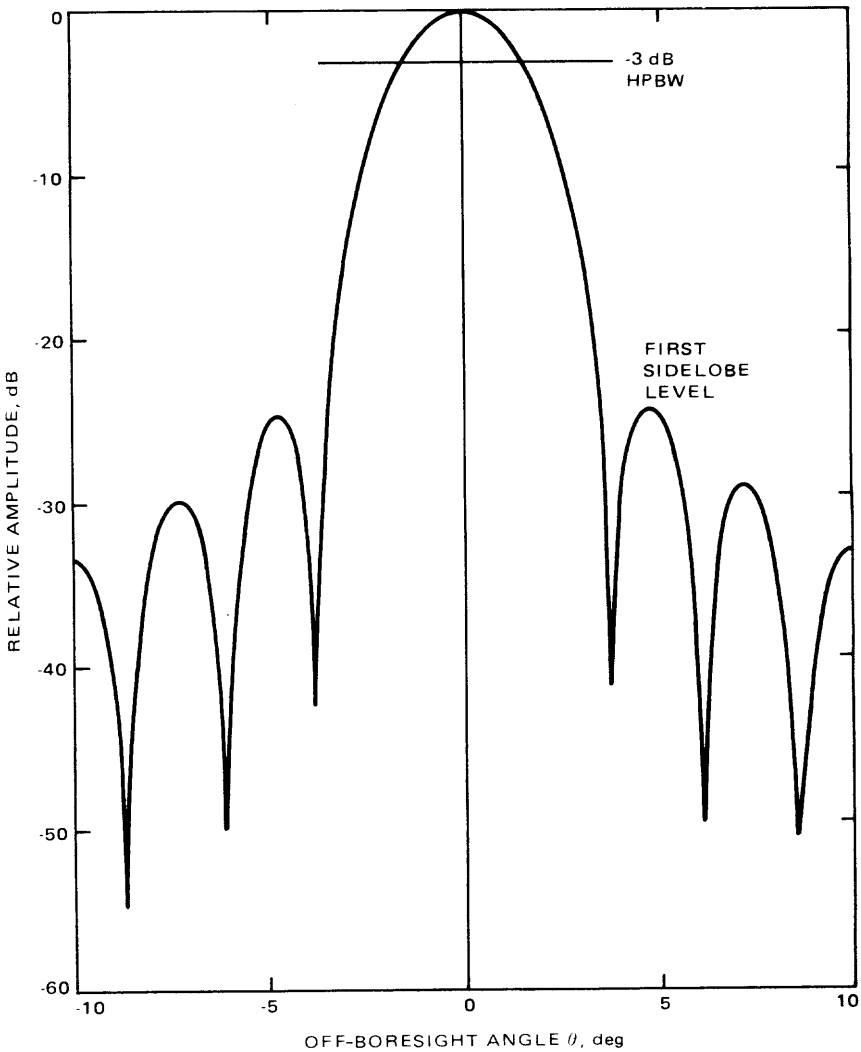


Fig. 8-4. Normalized antenna pattern in rectangular coordinates in dB power scale

An important numerical specification of the pattern is the angular width of the main lobe at a particular level. The angle at the half-power level or half-power beamwidth (HPBW) is most commonly used and defined as

$$\text{HPBW} = |\theta_{HP\ left}| + |\theta_{HP\ right}|$$

where $\theta_{HP\ left}$ and $\theta_{HP\ right}$ are angles at half-power levels down from the left and right of the peak, respectively, as shown in Fig. 8-3. The antenna HPBW can

be different in different pattern cuts for nonsymmetric patterns. Also defined are the beamwidth between first nulls (BWFN) or the beamwidth at -10 or -20 dB below the pattern maximum. Another important parameter is sidelobe level, which is the ratio of the pattern value of a sidelobe peak to the pattern value of the main lobe peak and typically is expressed in dB. For example, in Fig. 8-4 the first sidelobe level is at -24 dB.

8.2.2 Directivity, Gain, and Aperture Efficiency

An important characteristic of an antenna is how well it concentrates energy in one direction with reference to radiation in other directions. This characteristic is most easily described using the Poynting vector concept defined as

$$\mathbf{P} = \mathbf{E} \times \mathbf{H}^* \tag{8.2-2}$$

where **E** and **H** are the electric and magnetic fields and * designates the conjugate operator. The total average power radiated by an antenna is

$$W_r = \frac{1}{2} Re \int\int (\mathbf{E} \times \mathbf{H}^*) \cdot \hat{r}\, d\Omega \tag{8.2-3}$$

in which $\hat{r}$ is the unit vector in the propagation direction, $d\Omega$ defines the differential of the solid angle, and Re is the operator which takes the real part of a complex number. Furthermore, the radiation intensity can be defined as

$$U(\Omega) = \frac{1}{2} Re\,(\mathbf{E} \times \mathbf{H}^*) \cdot r^2 \hat{r} \tag{8.2-4}$$

where r^2 is introduced to make $U(\Omega)$ independent of distance. Furthermore, using (8.2-3) in (8.2-1) one can obtain the average radiated power. The directivity pattern is now defined as

$$D(\Omega) = \frac{4\pi\, U(\Omega)}{W_r} \tag{8.2-5}$$

and its maximum value, called directivity, is given by

$$D = \max\,[D(\Omega)] \tag{8.2-6}$$

which is typically in the main beam direction.

The directivity is solely determined by the radiation pattern of an antenna. However, when an antenna is used in a system (say, transmitting), one is actually

interested in knowing how efficiently the antenna transforms available power at its input terminals to radiated power in a particular direction. Obviously, because of resistive losses, the radiated power is always less than the total input power. To this end, power gain (or simply gain) is defined as 4π times the ratio of the radiation intensity in a given direction to the net power accepted by the antenna from the connected transmitter; i.e.,

$$G(\Omega) = \frac{4\pi U(\Omega)}{W_{in}} \tag{8.2-7}$$

where W_{in} is the input power (this definition does not include losses due to mismatches of impedance). The maximum value of power gain, called gain, is

$$G = \max\,[G(\Omega)] \tag{8.2-8}$$

which is typically in the main beam direction.

Comparison of (8.2-5), (8.2-1), and (8.2-7) prompts one to define radiation efficiency, e, as

$$e = \frac{W_r}{W_{in}} \tag{8.2-9}$$

which results in

$$G = eD \tag{8.2-10}$$

where

$$0 < e \leqslant 1 \tag{8.2-11}$$

From (8.2-10) one concludes that the maximum power gain of an antenna is equal to its purely directional characteristic of directivity reduced by its radiation efficiency. Furthermore, one can easily establish that

$$\iint G(\Omega)\, d\Omega = 4\pi e \tag{8.2-12}$$

Since gain is a power ratio, it can be calculated in decibels as follows:

$$G_{dB} = 10 \log_{10} G \tag{8.2-13}$$

Unfortunately, the terminology found in the literature is inconsistent on the topics of directivity and gain, and they are often confused and sometimes used interchangeably. Thus, the reader should be careful to understand the proper usage.

For aperture-type antennas, the gain can be related to the geometrical area of the aperture by the following well-known expression:

$$G = \eta_{ap} \frac{4\pi}{\lambda^2} A_p \tag{8.2-14}$$

where η_{ap} and A_p are the aperture efficiency and aperture area of the reflector, respectively. The aperture efficiency η_{ap} may further be expressed by its constituent components as

$$\eta_{ap} = e\eta_t\eta_1\eta_2\eta_3 \cdots \tag{8.2-15}$$

where

e = radiation efficiency due to ohmic losses, which are usually very small, except for lossy devices included as part of the feed system.

η_t = aperture taper efficiency due to the gain losses based on an aperture illumination that is tapered relative to uniform illumination, which produces maximum gain. For example, Fig. 8-5 (from [8-7]) shows the variation of η_t as a function of illumination taper.

η_1 = spillover (or feed) efficiency due to the portion of the feed radiation which is not intercepted by the reflector. As the aperture taper increases, the spillover decreases while the aperture taper efficiency η_t decreases. The tradeoff between η_t and η_1 has an optimum solution as indicated by the combined curve in Fig. 8-5. Typically, this optimum solution occurs at about -11 dB edge taper [8-7], and the value of -10 dB edge taper is frequently quoted.

η_2 = random surface error efficiency due to surface rms errors. This efficiency can be obtained from

$$\eta_2 = e^{-(4\pi\kappa\delta/\lambda)^2} \tag{8.2-16}$$

where δ is the surface rms error and κ is a factor which depends on the f/D ratio as shown in Fig. 8-6.

η_3 = aperture blockage efficiency due to the presence of a feed system or a subreflector. Values of η_3 are given in Fig. 8-7 as a function of the

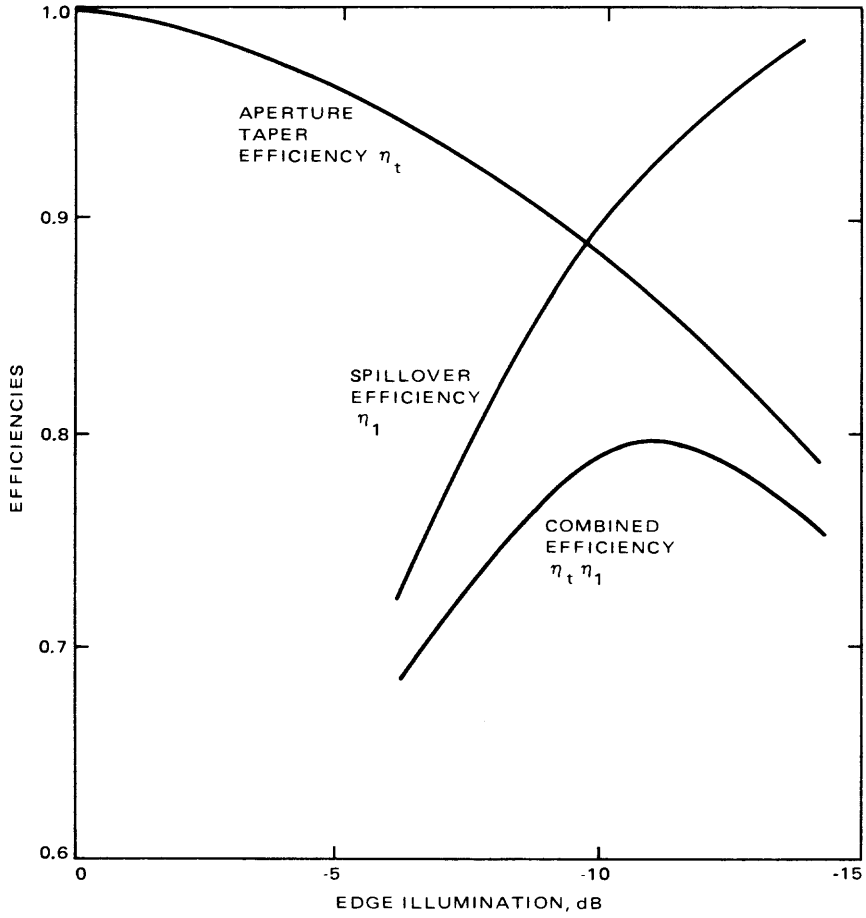


Fig. 8-5. Aperture taper and spillover efficiencies as a function of edge illumination

ratio of the diameter of the blocking structure near the focal plane to the main reflector diameter. Also shown are the effects of the aperture blockage on the sidelobe levels.

η_4 = strut blockage efficiency due to the presence of struts. Values are given in Table 8-1 for various main reflector sizes and a number of half-wavelength-thick struts [8-7].

η_5 = squint efficiency due to the lateral displacement of the feed. If the beam shifts off-axis by one half-power beamwidth, $\eta_5 = 0.98$.

η_6 = astigmatism efficiency due to axial displacement of the feed. This is a function of frequency and the f/D ratio and may take the values of 0.996, 0.98, and 0.93 for a 0.1λ axial displacement and for f/D's 1/2, 1/3 and 1/4, respectively [8-7].

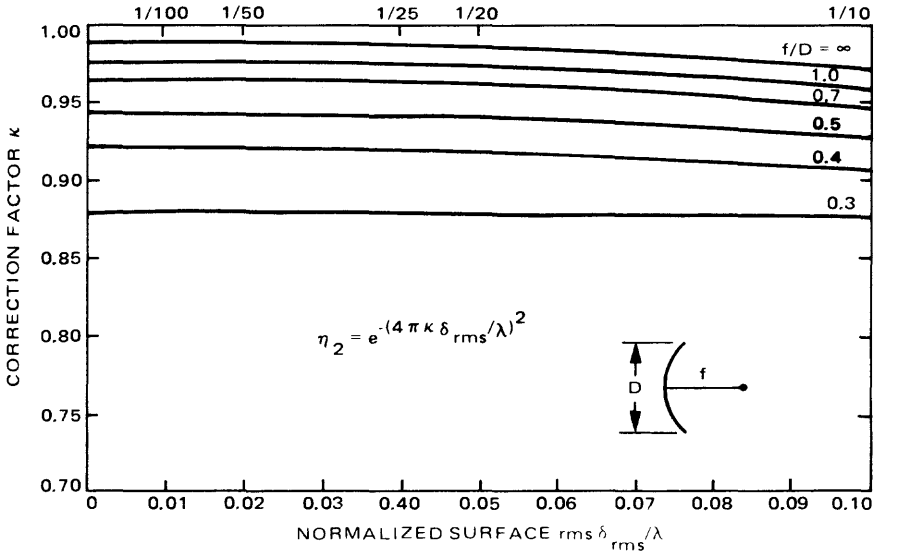


Fig. 8-6. Correction factor κ as a function of the surface rms (in the normal to the surface direction) δ_{rms}/λ for different f/D values

η_7 = surface leakage efficiency due to mesh-type surfaces. It is typically 0.99 for a mesh with several grid wires per wavelength.

η_8 = depolarization efficiency due to the power generated in the polarization state orthogonal to that desired. It typically takes a value greater than 0.98.

For antennas with high efficiency feeds, the overall efficiency η_{ap} is typically quoted to vary between 0.45 and 0.70.

8.2.3 Simple Models for Pattern Functions

The majority of well-focused circular aperture antennas, such as reflector antennas, can be approximated as circular apertures with aperture field amplitude distributions which are tapered from the center of the aperture toward the edge. Furthermore, for many applications (with feed at the focal point) the circular aperture distributions are close to being radially symmetric. With these assumptions, one can obtain a simplified mathematical model which describes the antenna pattern function. To this end, an aperture amplitude distribution of the type

$$E(\rho) = C + (1 - C)\left[1 - \left(\frac{\rho}{a}\right)^2\right]^n \qquad (8.2\text{-}17)$$

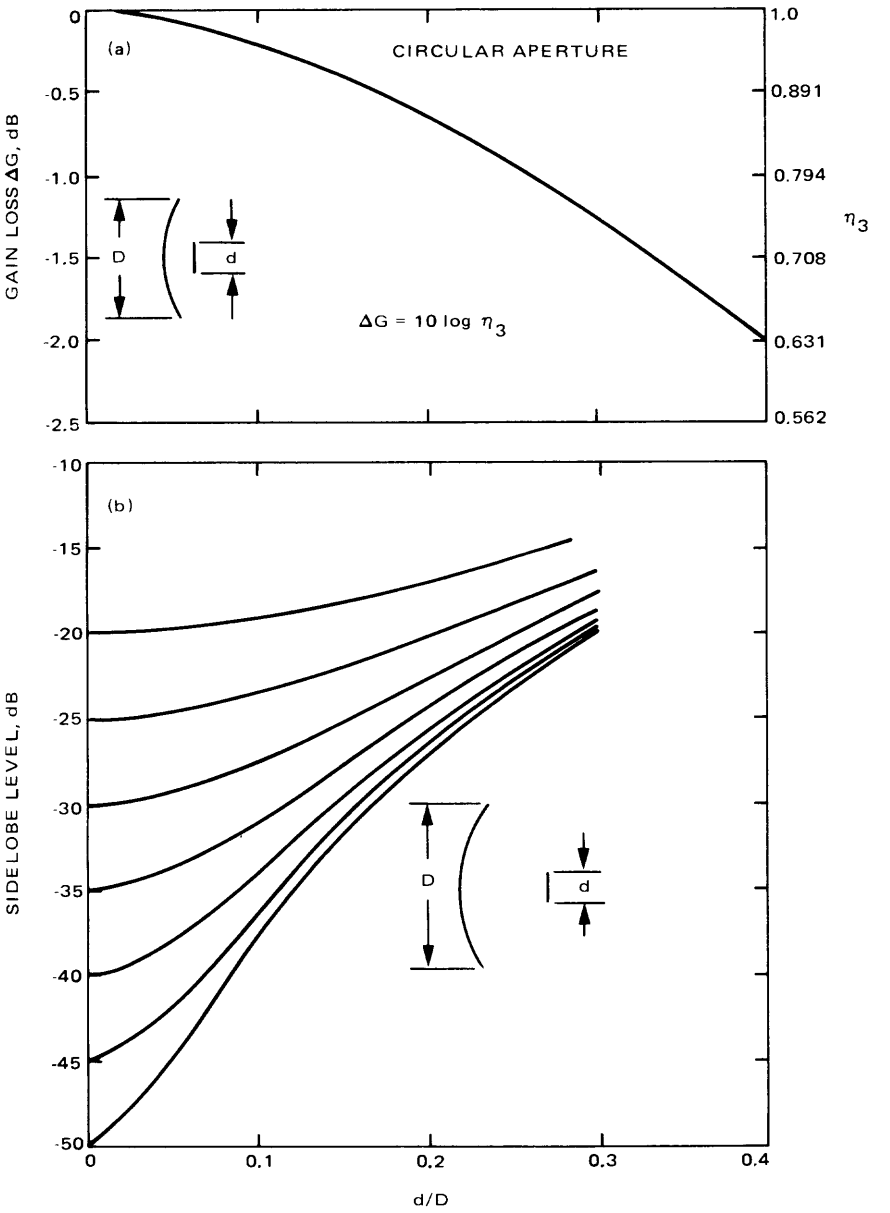


Fig. 8-7. Blockage effects on the gain and sidelobe levels: (a) gain loss due to blockage, and (b) sidelobe level vs aperture blockage ratio for a circular aperture with various illuminations

Table 8-1. Blockage efficiency η_4 due to the struts

	D		
N	10λ	100λ	200λ
3	0.946	0.995	0.999
4	0.935	0.994	0.998

D = reflector diameter; N = number of support struts which are $\lambda/2$ thick.

is used in which C is a parameter which controls the reflector aperture edge taper (i.e., $ET = 20 \log C$), a is the radius of the reflector aperture, n is a parameter which controls the amplitude shape, and ρ is the radial distance such that $0 \leqslant \rho \leqslant a$. Using the above distribution, one can determine the normalized far-field patterns as [8.7]

$$P(\theta, n, C) = \frac{Cf(\theta, n = 0) + \dfrac{1 - C}{n + 1} f(\theta, n)}{C + \dfrac{1 - C}{n + 1}} \tag{8.2-18}$$

where

$$f(\theta, n) = \frac{2^{n+1}(n + 1)!\, J_{n+1}\,[2\pi(a/\lambda) \sin \theta]}{[2\pi(a/\lambda) \sin \theta]^{n+1}} \tag{8.2-19}$$

and where J_{n+1} is the $(n + 1)$th-order Bessel function. From the foregoing expressions one can also determine the aperture taper efficiency η_t as

$$\eta_t = \frac{\left[C + \left(\dfrac{1 - C}{n + 1}\right)\right]^2}{C^2 + 2C\left(\dfrac{1 - C}{n + 1}\right) + \dfrac{(1 - C)^2}{2n + 1}} \tag{8.2-20}$$

The plots of E from (8.2-17) for various values of C and n are depicted in Figs. 8-8 and 8-9. For these cases, the far-field patterns, normalized to the 0 dB taper case, are constructed using (8.2-17) with final results as shown in Figs. 8-10 and 8-11. These plots clearly demonstrate how the relative gain and sidelobe levels can vary as functions of the illumination tapers and shapes. Based on these results, one can also tabulate the important characteristic

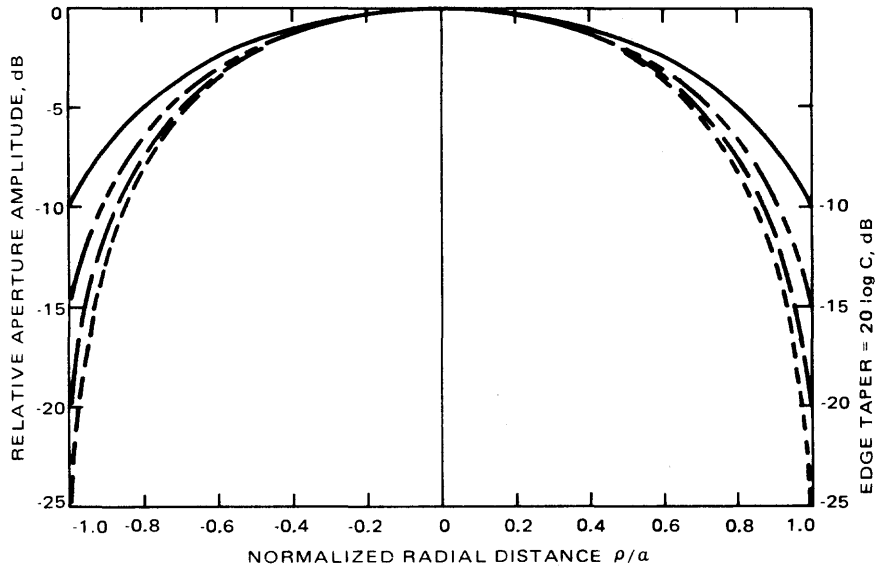


Fig. 8-8. Aperture distribution based on (8.2-17) for different edge tapers (−10, −15, −20, −25 dB) and n=1

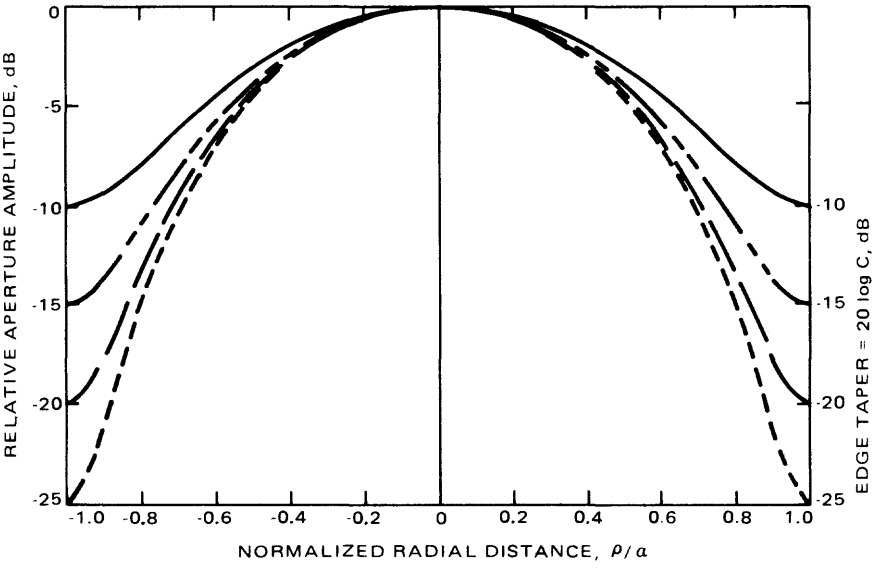


Fig. 8-9. Aperture distribution based on (8.2-17) for different edge tapers (−10, −15, −20, −25 dB) and n=2

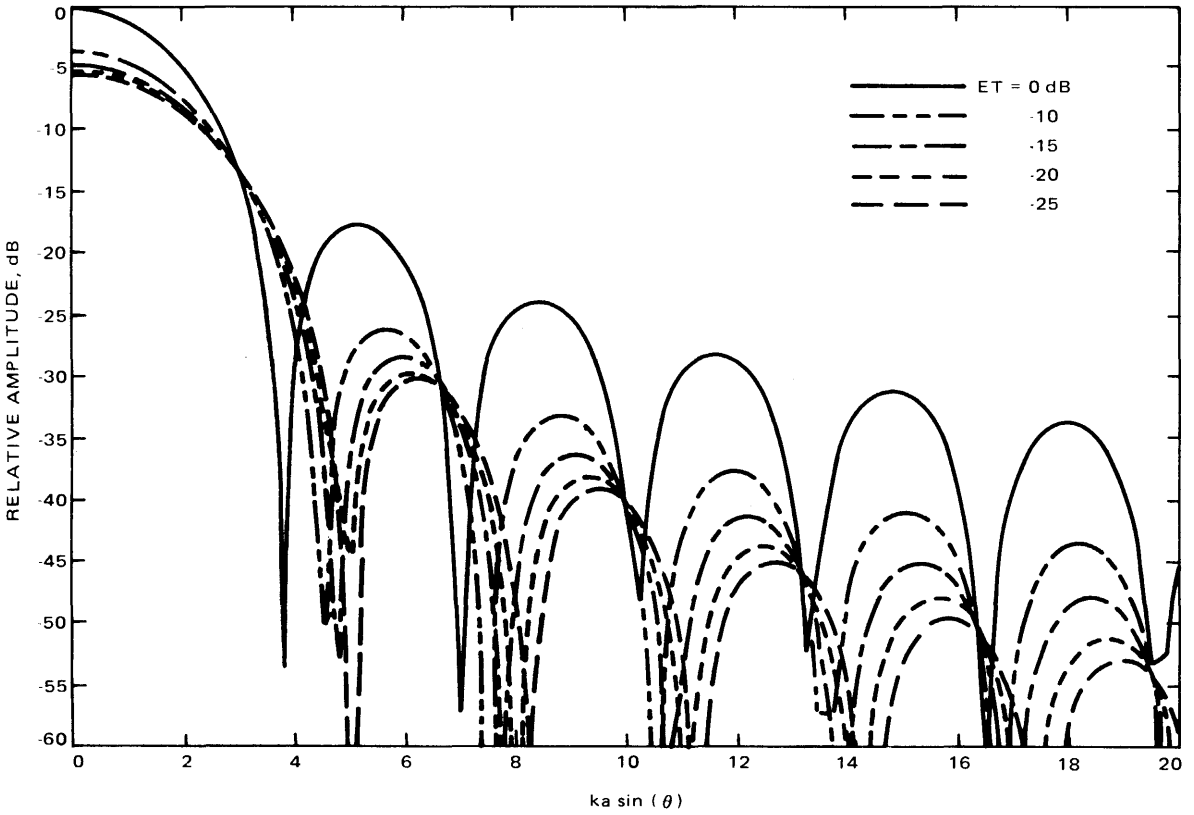


Fig. 8-10. Far-field patterns based on the aperture distribution (8.2-17) for different edge tapers and $n = 1$; $k = 2\pi/\lambda$, a = antenna radius, θ = off-boresight angle. (Patterns are normalized to the peak of the 0 dB edge taper case)

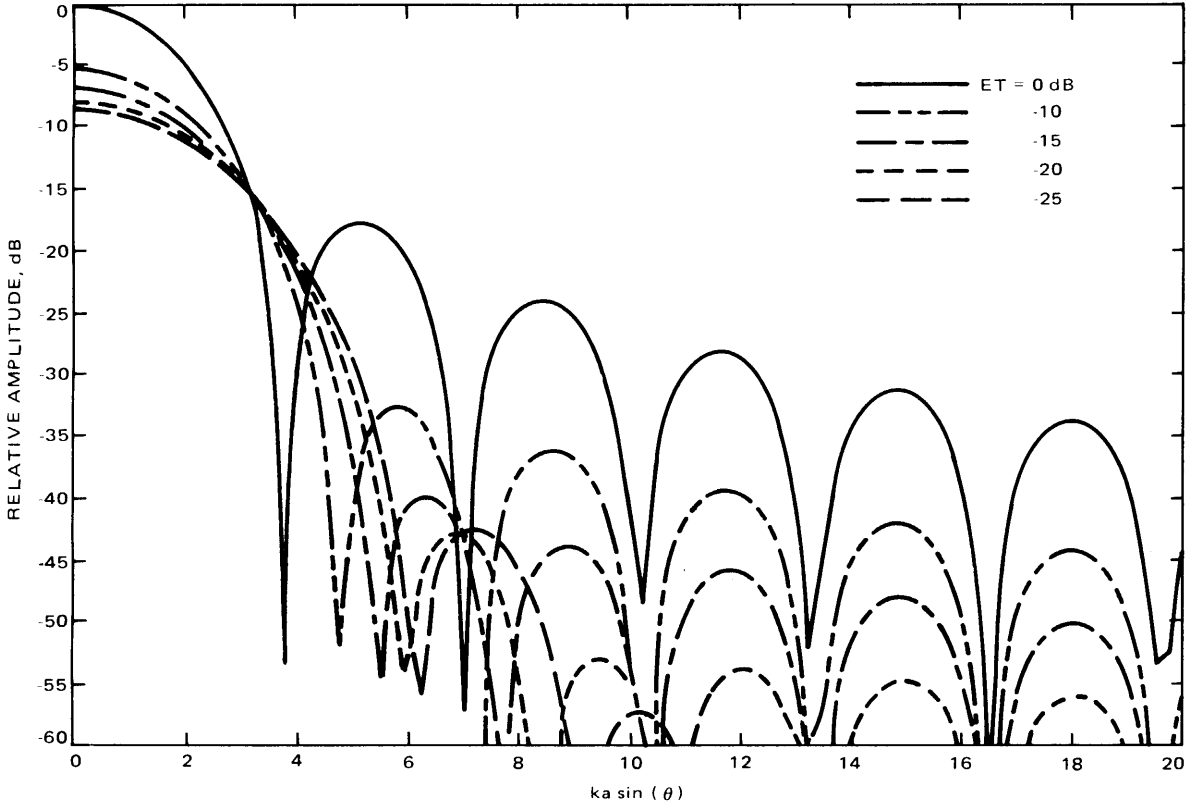


Fig. 8-11. Far-field patterns based on the aperture distribution (8.2-17) for different edge tapers and $n = 2$; $k = 2\pi/\lambda$, a = antenna radius, θ = off-boresight angle. (Patterns are normalized to the peak of the 0 edge taper case)

parameters of the patterns as shown in Table 8-2. These patterns and tabulated data can be used as a good approximation for many reflector antenna patterns in the angular regions close to the main beam.

In order to demonstrate the applicability of the foregoing simple expressions, a realistic case of the Viking Orbiter high-gain antenna is discussed. The geometry of this antenna is shown in Fig. 8-12. The antenna diameter is $2a$ = 58 inches and the operating frequency is set at X-band (f = 8.415 GHz). The far-field patterns of this reflector as measured in two different cuts are shown in Fig. 8-13. Clearly, owing to the effects of the struts and other components, the patterns in two planes are not the same; however, they demonstrate similar features, particularly about the main beam. These measurements are obtained by both measuring the antenna in a standard outdoor far-field range, and also in the near-field range, using a new technique developed by Rahmat-Samii, et al. [8-10]. This technique is finding an ever-increasing application in measuring the far-field patterns and gain of large directive (high-gain) antennas. Figure 8-14 demonstrates the basic experimental setup of this near-field procedure. The prime advantages of the near-field technique are: (1) the antenna can be measured in the controlled environment, and (2) true far-field patterns can be obtained independent of the size of the antenna. Briefly explained, the near-field amplitude and phase data of the antenna are measured in a distance typically less than a diameter. These data are then used via some computer programs [8-10] to generate the true far-field patterns of the antenna. As shown in Fig. 8-13, the comparison between the far-field patterns using the near-field technique and the direct far-field measurement is very good. The gain of this antenna is 38.8 dB with an aperture efficiency of η_{ap} = 44.91%, which includes

Table 8-2. Characteristics of tapered circular aperture distribution (8.2-17 through 8.2-20)

Edge illumination		n = 1			n = 2		
ET (dB)	C	HPBW, rad	Sidelobe level, dB	η_t	HPBW, rad	Sidelobe level, dB	η_t
0	1	1.01 $\lambda/2a$	−17.6	1	1.01 $\lambda/2a$	−17.6	1
−8	0.398	1.12 $\lambda/2a$	−21.5	0.942	1.14 $\lambda/2a$	−24.7	0.918
−10	0.316	1.14 $\lambda/2a$	−22.3	0.917	1.17 $\lambda/2a$	−27.0	0.877
−12	0.251	1.16 $\lambda/2a$	−22.9	0.893	1.20 $\lambda/2a$	−29.5	0.834
−14	0.200	1.17 $\lambda/2a$	−23.4	0.871	1.23 $\lambda/2a$	−31.7	0.792
−16	0.158	1.19 $\lambda/2a$	−23.8	0.850	1.26 $\lambda/2a$	−33.5	0.754
−18	0.126	1.20 $\lambda/2a$	−24.1	0.833	1.29 $\lambda/2a$	−34.5	0.719
−20	0.100	1.21 $\lambda/2a$	−24.3	0.817	1.32 $\lambda/2a$	−34.7	0.690

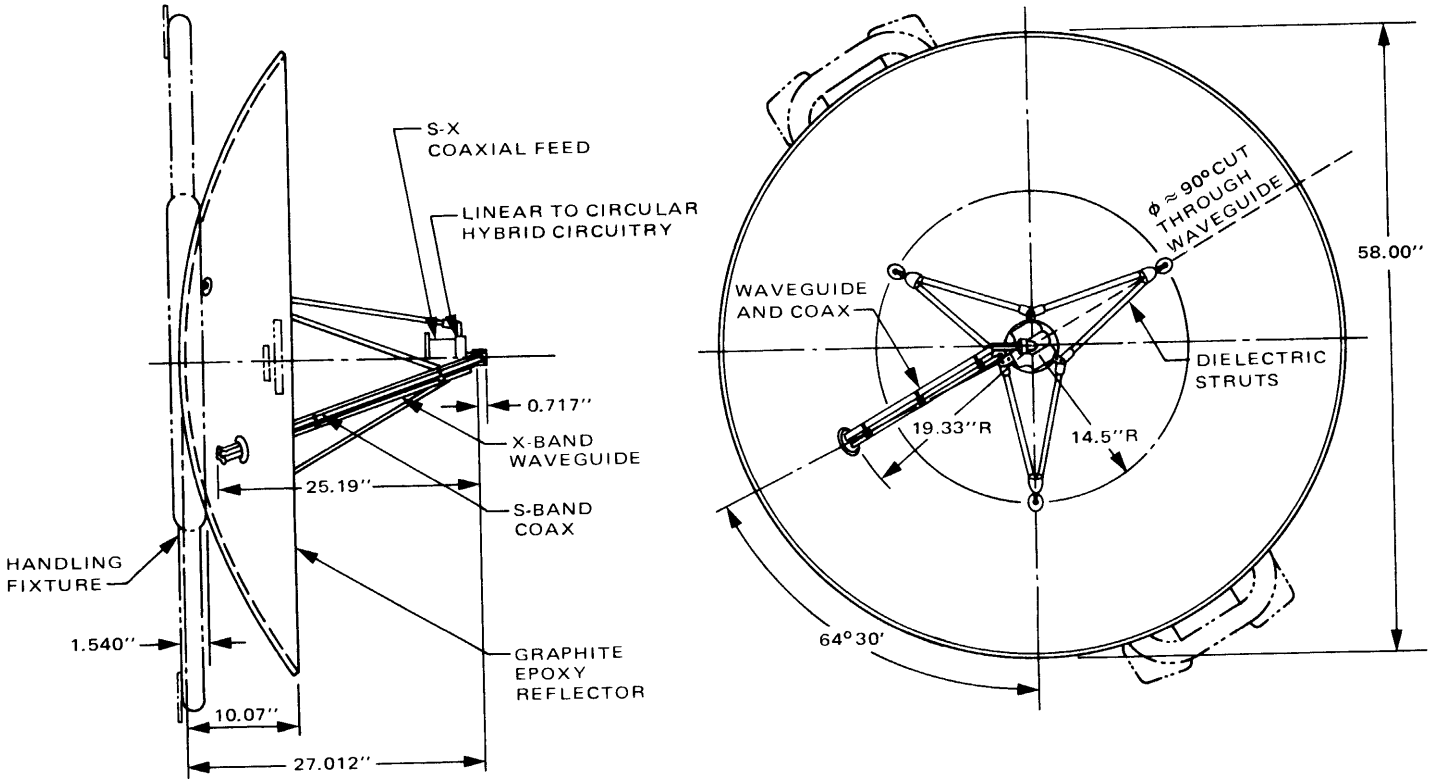


Fig. 8-12. Geometrical layout of the Viking Orbiter high-gain antenna

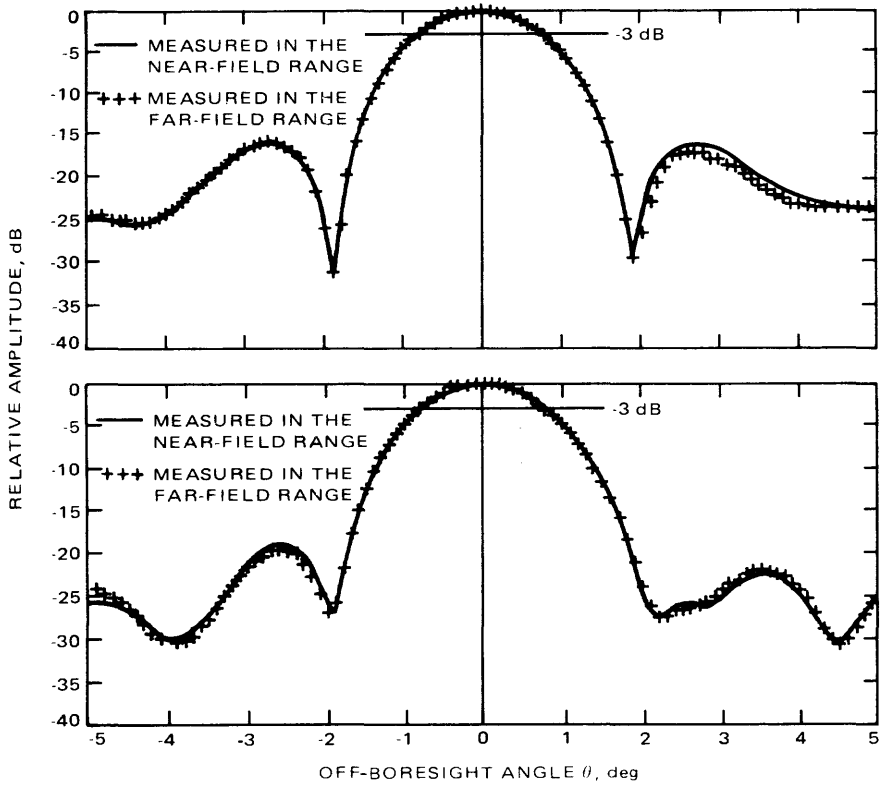


Fig. 8-13. Far-field patterns of the Viking Orbiter high-gain antenna measured in two different cuts using both the direct far-field and near-field measurement techniques.

the waveguide losses. For this reflector the feed is designed to provide an edge taper of about -10 dB. This measured pattern can be approximated with good accuracy using the simple formula of (8.2-18). In this case, $C = 0.316$ (-10 dB taper) and $n = 1$. The normalized far-field pattern based on these values can be constructed from (8.2-18) with the final result plotted in Fig. 8-15. As is evident, a good approximation for the actual pattern is obtained in the main beam region. Note that due to the blockage effects, the sidelobes cannot be accurately predicted by the simple model. It is worthwhile to mention that the simple model presented by (8.2-18) is based on a simplified scalar description of circular reflector antennas. There are vector-diffraction analysis techniques available which accurately predict both the copolar and cross-polar patterns of reflector antennas. The reader may refer to [8-2] or to the more recent works of Rahmat-Samii, et al. [8-11 and 8-12].

Although equation (8.2-18) provides a first-cut estimate of the pattern, on many occasions it will be useful to have a quick estimate of the reflector's gain and beamwidth. Table 8-3 provides such information for a variety of antennas.

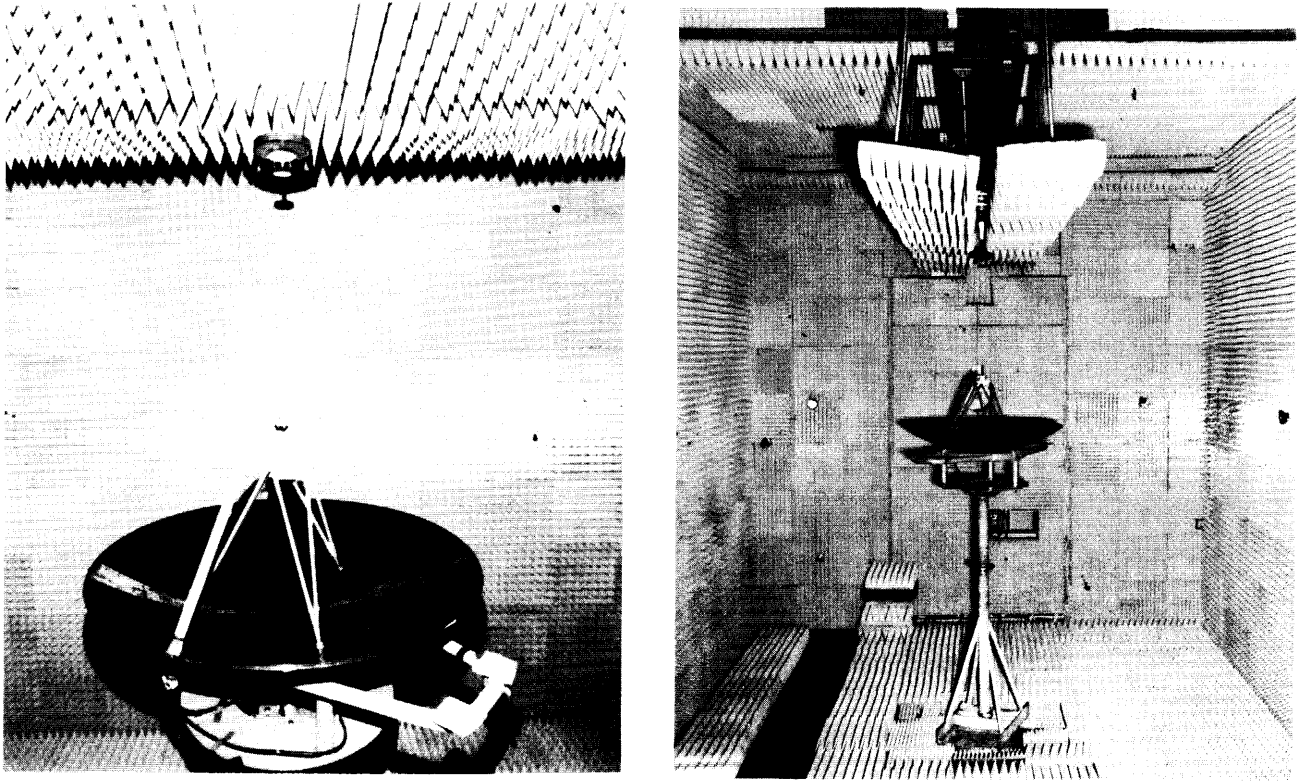


Fig. 8-14. Experimental setup: Near-field plane–polar measurements for far-field pattern determination of the Viking high-gain reflector antenna

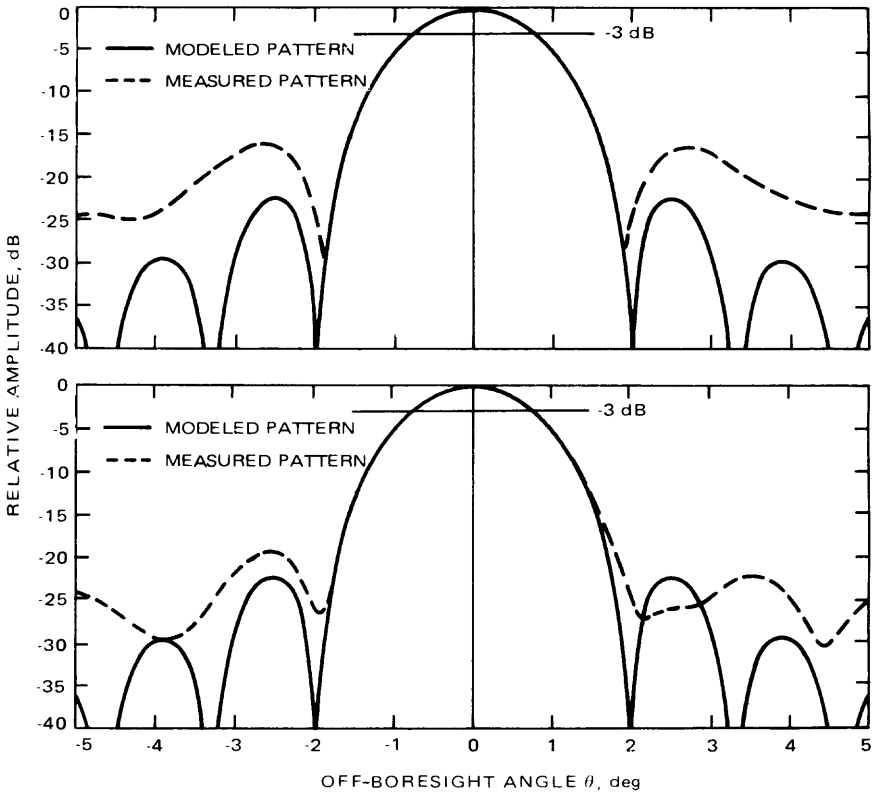


Fig. 8-15. Comparison between the results of the simple model (8.2-18) and the experimental data of Fig. 8-13. In (8.2-18) the values of $a = 58''/2 = 20.68\lambda$, $n = 1$ and $C = 0.316$ (−10 dB taper) are used

8.2.4 Choice of Antenna

For a telecommunications system, one of the least expensive methods of obtaining increased performance per dB is to vary the antenna gain, provided proven antenna types are used. The only other method that rivals antenna choice is bit rate, but this is generally set by other system requirements. For a given power-gain product there is a particular combination of antenna gain and transmitter power at which the spacecraft weight is minimized. Thus, in weight-critical vehicles, it may be necessary to increase transmitter power rather than antenna gain, despite a possibly higher cost.

The telecommunications designer must obtain an idea of the minimum antenna gains necessary to accomplish the mission for each mode by using state-of-the-art values for spacecraft parameters and the committed DSN parameters for the mission data which provide a starting point for antenna choice. Typical

Table 8-3. Approximate antenna gain and 3-dB beamwidths for different antennas

Configuration	Gain above isotropic radiator	3-dB beamwidth, deg
Isotropic radiator	1	360
Infinitesimal dipole or loop	1.5	89.9 (toroidal pattern)
Half-wave dipole	1.64	78 (toroidal pattern)
Paraboloid (area A, diameter D)	(6.3 to 8.8) A/λ^2	(60 to 70) λ/D
Open mouth waveguide (area A, E-plane dimension d_E, H-plane, d_H)	$10\,A/\lambda^2$	$\theta_E = 50\,\lambda/d_E$ $\theta_H = 77\,\lambda/d_H$
Optimum horn (mouth area A, E-plane dimension d_E, H-plane d_H)	$7.5\,A/\lambda^2$	$\theta_E = 54\,\lambda/d_E$ $\theta_H = 78\,\lambda/d_H$
Optimum biconical horn (height h)		
Vertical polarization	$1.2\,h/\lambda$	$56\,\lambda/h$ (toroidal pattern)
Horizontal polarization	$1.6\,h/\lambda$	$67\,\lambda/h$ (toroidal pattern)

antenna gain capabilities for several types of antennas are shown in Fig. 8-16. In order to maximize the reflector antenna gain, antenna designers have used dual-shaped reflector antennas. The surface of these shaped reflector antennas departs from the classical shapes such as the paraboloid (for main reflector) and the hyperboloid (for subreflectors). For example, the 1977 Voyager spacecraft used a 3.7-m shaped-dual-reflector system for high data-rate telemetry transmission [8-13]. The system was optically shaped for high efficiency at X-band. The dichroic subreflector was transparent to radiation from an S-band prime-focus horn nestled behind it. The geometrical layout of this dual-shaped reflector is shown in Fig. 8-17. A photograph of the model is shown in Fig. 8-18. The gains of this reflector antenna at X-band and S-band are 48.2 dB and 35.3 dB, respectively.

In order to further improve the pattern characteristics of reflectors and eliminate the effects of unwanted struts and subreflector blockages, the antenna designers have started using the concept of offset (clear aperture) reflector antennas [8-14]. This concept is particularly useful when there is a need for a low sidelobe reflector design such as satellite communication antennas [8-2 and 8-12].

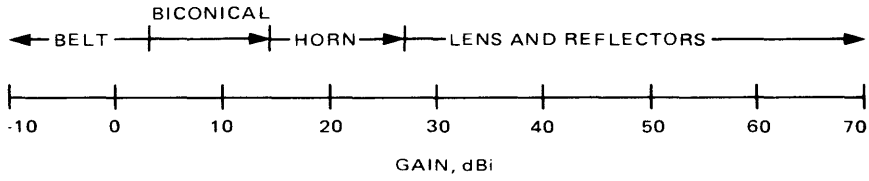


Fig. 8-16. Gain capabilities of several types of antennas

8.3 Pointing Errors

The maximum spacecraft telecommunications performance is obtained when the target is aligned with the maximum-gain point of the spacecraft antenna. It is difficult, if not impossible, to maintain exact target tracking. Therefore, an acceptable pointing error is usually specified. This determines the positioning requirements for a given antenna. Total pointing error is a function of mission time and is usually specified by degrees away from boresight.

Discussions of pointing error sources and statistics are presented here, followed by a discussion on pointing loss. Total pointing error and pointing loss are analyzed (see appendix) for a spinning spacecraft. The analysis was specifically developed for the Galileo spacecraft, but adaptation to similar pointing dynamics is simple.

8.3.1 Pointing Control Errors

A change in target positions over the antenna gain pattern occurs with mission time due to the basic limitations of the antenna pointing system. For example:

(1) For a fixed antenna position, the spacecraft-to-earth vector is aligned with the peak gain for only short periods of time.

(2) For a single-axis pointing system, the antenna does not track earth exactly.

(3) In a discrete pointing system, pointing error will occur because of non-continuous position updating.

(4) Optimizing antenna position for a range of trajectories may result in pointing errors for a particular flight vehicle trajectory.

Spinning antennas also suffer from:

(5) Misalignment between the mechanical boresight and spacecraft spin axis.

(6) Misalignment between the antenna electrical and mechanical boresight.

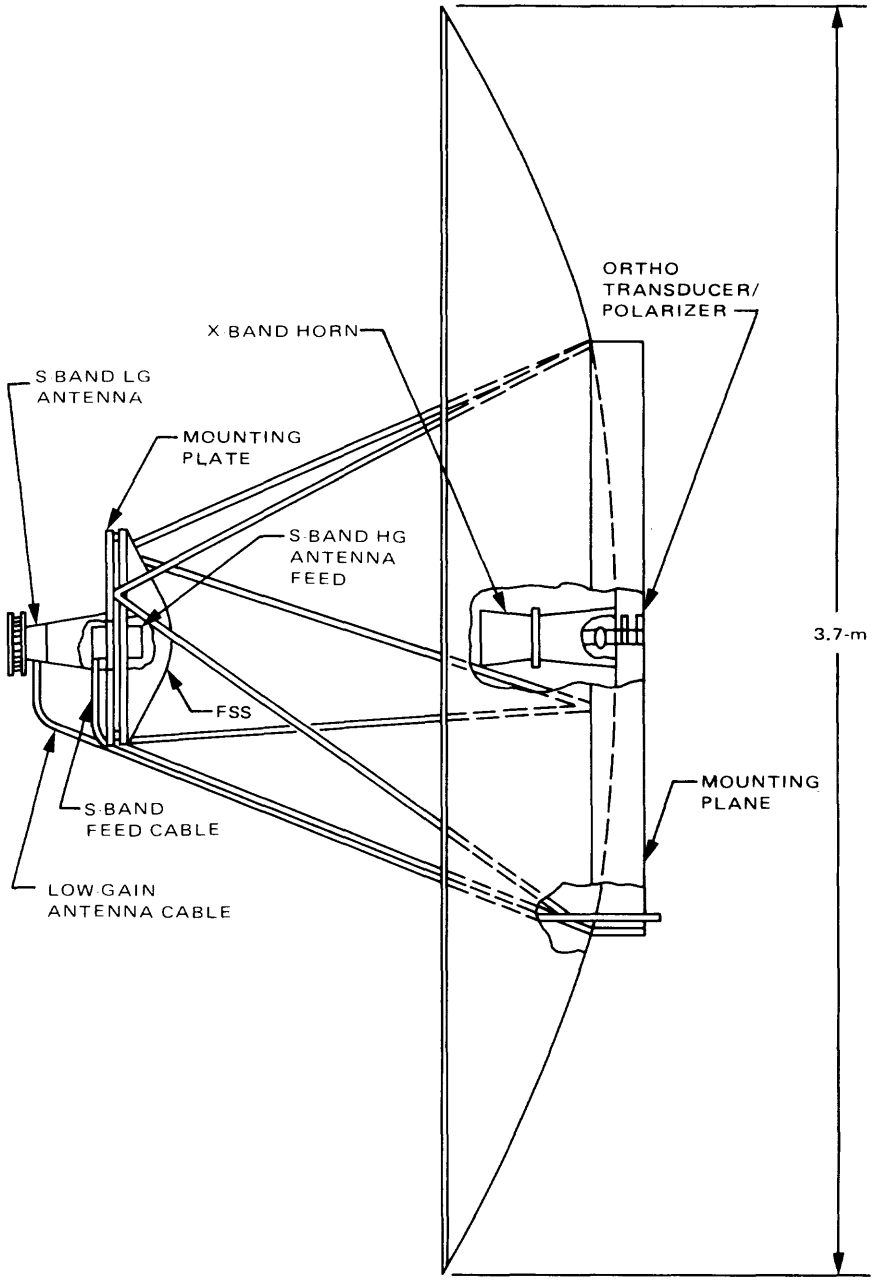


Fig. 8-17. Voyager high-gain antenna overall layout

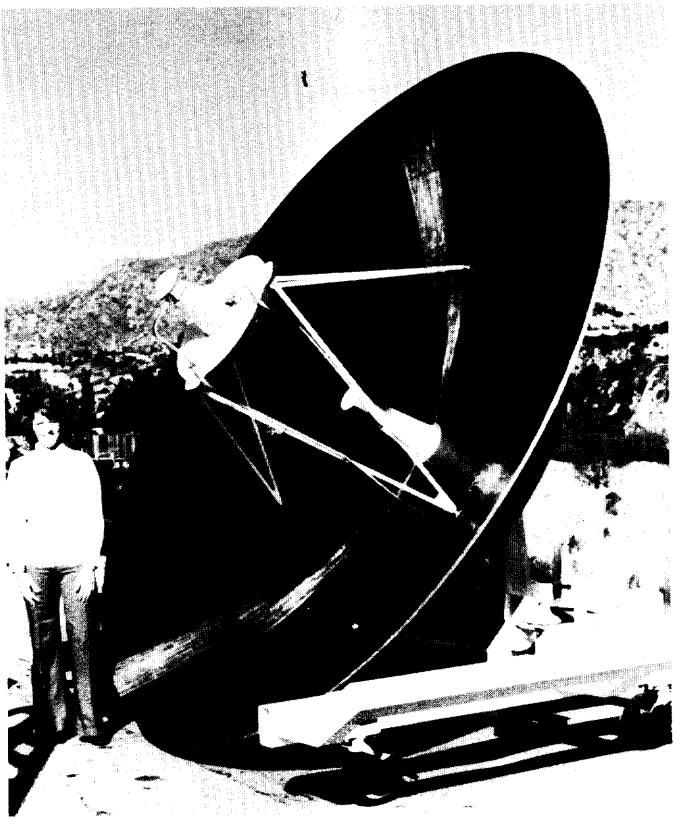


Fig. 8-18. Photograph of the Voyager high-gain antenna

(7) Wobble arising from center-of-gravity offsets and coupling between axes of inertia.

(8) Nutation arising from external torques.

An unacceptable aiming error results if the antenna position is not changed to compensate for spacecraft-earth motion. A change in antenna position usually requires the use of an attitude control system equipped with gas jets if the antenna is body-fixed. In order to keep gas consumption to realistic levels, the attitude control system typically operates in a "bang-bang" mode whereby a restoring torque is exerted whenever the pointing error grows to a specified level. The spacecraft is free, then, to point only within a designated cone.

The Galileo spacecraft, for example, controls its pointing offsets within a 1.64-mr deadband. Whenever the angular momentum vector estimator – used

for attitude determination – drifts beyond the deadband, the spacecraft reorients itself toward earth.

These pointing errors are not errors in the usual sense. They are planned deviations from peak antenna gain to lower costs and increase the reliability of the antenna subsystem. Since they are due to limitations in the ability to control the antenna pointing direction, they are referred to as pointing control errors.

8.3.2 Pointing Knowledge Errors

There are several types of errors that contribute to uncertainty in pointing direction and hence to antenna pointing error. These are referred to as pointing knowledge errors.

Some pointing knowledge errors occur because the positions of the high-gain antenna and the attitude sensing devices that detect the correct spacecraft orientation are not precisely known. Other pointing knowledge errors are due to limitations in ability to resolve sensor outputs. These errors result in an antenna attitude that differs slightly from the nominal position in an unpredictable way. For a typical vehicle, these errors include:

(1) Alignment uncertainty between the mechanical boresight and the spacecraft principal axis.

(2) Alignment uncertainty between the antenna electrical and the mechanical boresight.

(3) Sun sensor alignment uncertainty.

(4) Star tracker alignment uncertainty.

(5) Sun sensor digital resolution.

(6) Star tracker digital resolution.

8.3.3 Pointing Error Statistics

Each contributor to pointing error must have a probability distribution associated with it which translates to antenna clock and cone coordinate distributions. The sum of these angular errors forms the uncertainty component of pointing error. By combining the probability distributions of all contributions to pointing error appropriately, the probability density function $p_e(\theta)$ of the total off-boresight pointing error θ can be computed.

8.4 Pointing Loss

Pointing loss is found by locating the spacecraft-earth vector on the antenna pattern. The magnitude of this vector is referred to as the pointing error, which

is discussed in Section 8.3. Nominal pointing loss, usually called pointing loss, is determined by

$$P_L = G(\theta, \phi) - G_m \tag{8.4-1}$$

where $G(\theta, \phi)$ is the nominal antenna gain function in dB, θ is the antenna cone angle and ϕ the antenna clock angle of the spacecraft-earth vector, and G_m is the nominal peak antenna gain in dB.

The mean pointing loss $\overline{P}_L$ is defined as the expected value of pointing loss

$$\overline{P}_L = \int\int [G(\theta, \phi) - G_m]\, p_e(\theta, \phi)\, d\theta\, d\phi \tag{8.4-2}$$

where $p_e(\theta, \phi)$ is the joint probability density function of θ and ϕ.

The variance σ_L^2 of the pointing loss is then

$$\sigma_L^2 = \int\int [G(\theta, \phi) - G_m]^2\, p_e(\theta, \phi)\, d\theta\, d\phi - \overline{P}_L^2 \tag{8.4-3}$$

P_L and σ_L^2 are inserted into the telecommunication link design control table (DCT) to account for the pointing loss degradations of the link. As an example, the pointing loss statistics of the Galileo spacecraft high-gain antenna are derived in the appendix.

8.5 Polarization Loss

The gain function of an antenna does not completely characterize its performance. $G(\theta, \phi)$ includes information on the magnitude of the electric field vectors of the antenna radiation, but discards information as to their relative phase. An antenna transmits an E field in a preferred way which can be described by the polarization factor p of the wave at each point in space, where $p = E_y/E_x$ in a known coordinate system. By the principle of reciprocity, a receiving antenna will select incoming waves of the polarization of its transmit pattern and discard all others (the antenna is said to have a polarization pattern). The ratio of the power delivered to the antenna terminals P_R to the power P_{Avail} received by an antenna matched to the incident polarization is called the polarization efficiency, ν, viz

$$\nu = \frac{P_R}{P_{Avail}} \tag{8.5-1}$$

This is also referred to as polarization loss, measured in dB, and it must be added to the transmission equation to obtain the true received power.

In general, p is a function of the coordinate system chosen, and is a complex number. In practice, the antenna is usually designed so that orthogonal E-vector components are 90 deg out of phase with one another in time. In this case, the E vector describes an ellipse with major axis R_1 and minor axis R_2. For this case,

$$p = \pm\frac{jR_1}{R_2} \triangleq \pm jAR$$

where $j = \sqrt{-1}$ and AR is called the axial ratio. The sign denotes the direction of E-vector rotation. The term ellipticity is used to express the axial ratio in dB, where

$$\text{ellipticity} = 20 \log_{10} AR \tag{8.5-2}$$

Generally, the antenna is characterized by a gain pattern and an ellipticity pattern [8-15]. For example, Fig. 8-19 shows the polarization characteristics of an antenna as a function of the ratio parameter p. For $AR = 1$, the E vector traces a circle, resulting in a circular polarization. If the E vector rotates counterclockwise to an observer looking from the receiver to the source, $p = -j$, and the wave is right-circularly polarized (RCP). For $p = j$, the wave is left-circularly polarized (LCP). Circular polarization is usually desirable for deep-space communication because, ideally, a change in orientation of the spacecraft antenna with respect to the target antenna causes no polarization loss. However, for application to landers, or for certain atmospheric experiments, other considerations may dominate. For example, under certain conditions the multipath from a lander transmitting RCP will be almost entirely LCP and may be rejected by an RCP antenna.

All antennas have some ellipticity. It may be shown [8-4] that the polarization loss in dB between two antennas with gains specified relative to radiators with perfectly matched polarization can be written as

$$L_p(\Psi) = 10 \log_{10} \frac{(1 + R_R^2)(1 + R_T^2) + 4R_T R_R + (1 - R_R^2)(1 - R_T^2)\cos 2\Psi}{2(1 + R_R^2)(1 + R_T^2)} \tag{8.5-3}$$

where

R_R = receiving antenna axial ratio

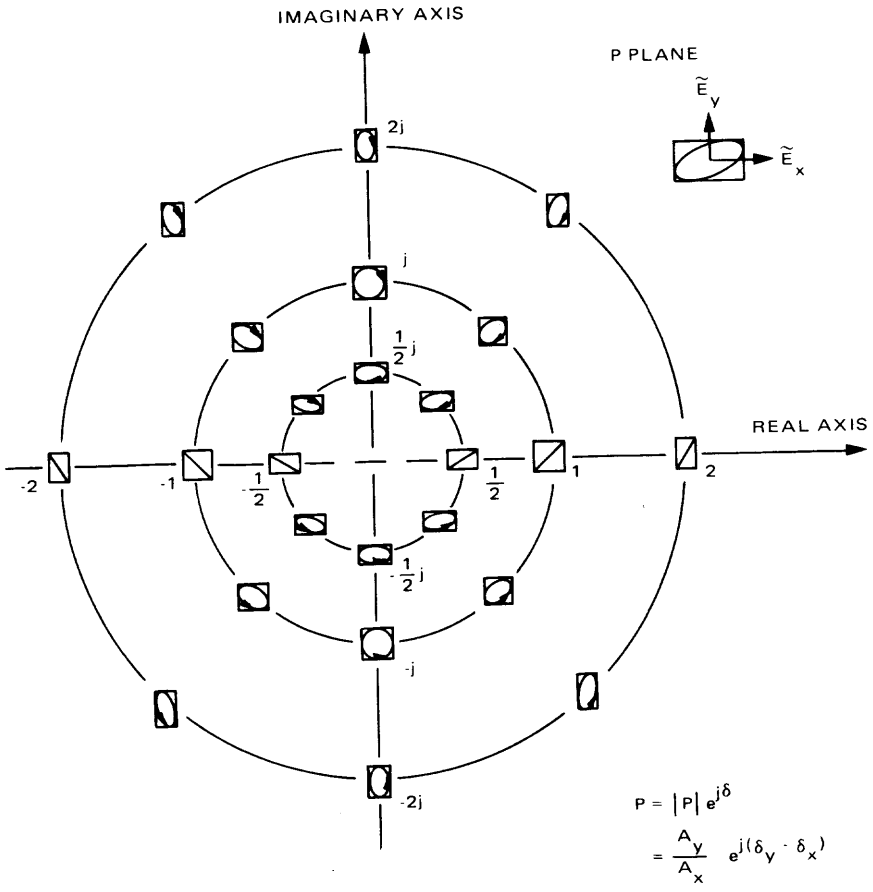


Fig. 8-19. Complex plane with polarization ellipses

R_T = transmitting antenna axial ratio

Ψ = angle between the major axes of the polarization ellipses

If the gain of the transmitting antenna is specified relative to a right-hand circular isotropic radiator instead, the polarization loss in dB will be

$$L_p(\Psi) = 10 \log_{10} \frac{[(1 + R_R^2)(1 + R_T^2) + 4 R_T R_R + (1 - R_R^2)(1 - R_T^2) \cos 2\Psi]}{(1 + R_R^2)\ (1 + R_T)^2} \tag{8.5-4}$$

In general, one does not know the relative orientation of the polarization ellipses. Thus, Ψ is random between 0 and 90 deg (for greater than 90 deg, the pattern repeats since the angular factor is 2Ψ). The mean and variance can be computed from (8.5-3) or (8.5-4) by assuming a uniform density function for Ψ.

8.6 Antenna Noise Temperature and Noise Spectral Density

The noise power component seen by an antenna is due to electromagnetic radiation generated in the antenna bandwidth by

(1) Celestial bodies within the antenna beam.

(2) Atmospheric absorption and reradiation.

(3) Absorption and reradiation by physical bodies surrounding the antenna.

The effective antenna noise temperature is defined by [8-4]

$$T_{ae} = \frac{1}{A_G} \int_{4\pi} T(\Omega)\, G(\Omega)\, d\Omega \tag{8.6-1}$$

where $A_G = \int_{4\pi} G(\Omega)\, d\Omega$, and $T(\Omega)$ is the blackbody temperature of the environment in the direction Ω. At 2 GHz, typical spacecraft values of T_{ae}, including noise contributions due to surface finish errors and other antenna generated sources, are [8-16]

T_{ae} = 50 K, no bright noise sources occupying a significant fraction of the beam

T_{ae} = 200 K, entire beam intercepted by the moon

T_{ae} = 290 K, entire beam intercepted by the earth

$T_{ae} \sim$ 10,000 K, entire beam intercepted by the sun

DSN antennas may have noise temperatures at S-band of 15 K or less when aimed at a cold sky (no major celestial source) at high elevation angles. At X-band, T_{ae} is greatly increased by the presence of water vapor. Reference [8-17] should be consulted for typical values.

Frequently, the antenna noise temperature is divided into the following two components:

$$T_g = \frac{1}{A_G} \int_{main\ lobe} T(\Omega)\, G(\Omega)\, d\Omega \tag{8.6-2}$$

and

$$T_a = \frac{1}{A_G} \int_{side\ lobes} T(\Omega)\, G(\Omega)\, d\Omega \tag{8.6-3}$$

Then

$$T_{ae} = T_g + T_a \tag{8.6-4}$$

This division is particularly useful for ground-based antennas at S-band, since for a cold sky and elevation angles greater than 10 degrees T_g is constant and T_a is a function of azimuth and elevation of the antenna. At X-band, T_g may be modeled as the sum of an input temperature dependent on celestial sources and an effective noise temperature $T_{Ae} = (L_A - 1)\, T_A$ due to atmospheric attenuation, where T_{Ae} is the effective temperature of the atmosphere, L_A is the total loss due to energy absorption, and T_A is a density-averaged temperature of the atmosphere equal to about 280 to 290 K. This is equivalent to regarding the atmosphere as a passive circuit loss.

For spacecraft antennas, the distinction between T_g and T_a is not usually made, although it might prove useful for steerable antennas if the spacecraft effects are large and variable.

The noise spectral density N_0 at the antenna terminal is

$$N_0 = kT_{ae} \tag{8.6-5}$$

where k is Boltzmann's constant, 1.38×10^{-20} mW-sec/K, or -198.6 dBm/Hz K. By definition (8.6-1), antenna gain is included in the computation of T_{ae}. Frequently, noise contributors such as surface finish errors are lumped into T_{ae}, since there is no effective way to separately measure these sources when measuring T_{ae}.

8.7 Flight Environment

Mounting antennas on a spacecraft changes their patterns and introduces interference effects which can be estimated and measured. A typical spacecraft mounts at least two antennas of the same design frequencies. These are connected by circulator switches to the same receiver. Because of line mismatches

and imperfect switches, some fraction of power into the receiver comes from the antenna that is not in use. Similarly, in transmission a fraction of the power transmitted comes from the nonoperating antenna. Since the path lengths traveled from the antenna to the receiver are different, the signals will not be in phase when they are summed. Thus, an interferometer effect will occur.

Figure 8-20 gives a model for interferometer equations, where

a_1 = unknown electrical path length from antenna 1 phase center to summing point

a_2 = unknown electrical path length from antenna 2 phase center to summing point

ℓ = free space distance between antenna phase centers

θ = angle between the spacecraft-to-earth vector and the line joining the antenna phase centers

$d = \ell \cos \theta$ = extra distance the wave travels from antenna 1 to antenna 2

k_1 = voltage gain (loss) for antenna 1 signal between antenna 1 and summing point

k_2 = voltage gain (loss) for antenna 2 signal between antenna 2 and summing point

The relative voltages received at antenna 1 and antenna 2, respectively, are

$$V_1 = g_1(\Omega_1) \tag{8.7-1}$$

and

$$V_2 = g_2(\Omega_2)\, e^{j\beta d} \tag{8.7-2}$$

where

g_1 = the voltage gain function of antenna 1

g_2 = the voltage gain function of antenna 2

Ω_1 = vector angle from the axis of antenna 1 to the spacecraft-to-earth vector

Ω_2 = vector angle from the axis of antenna 2 to the spacecraft-to-earth vector

$\beta = 2\pi/\lambda$

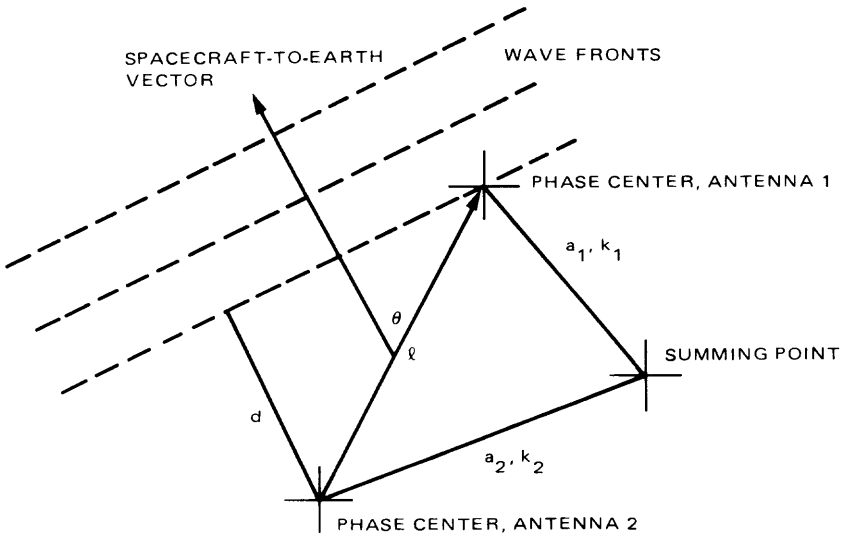


Fig. 8-20. Model for interferometer effect

Thus, at the summing point the received voltage is

$$V_R = k_1 g_1(\Omega_1)\, e^{j\beta a_1} + K_2 g_2(\Omega_2)\, e^{j\beta(d+a_2)} \tag{8.7-3}$$

with average power

$$P_R = |V_R|^2 = k_1^2 g_1^2(\Omega_1) + k_2^2 g_2^2(\Omega_2)$$

$$+ 2\, k_1 k_2 g_1(\Omega_1) g_2(\Omega_2) \cos(\beta(d + a_2 - a_1)) \tag{8.7-4}$$

If the operating antenna is chosen as antenna 1, the received power desired is

$$P_{R1} = k_1^2 g_1(\Omega_1)^2 = k_1^2 P_0 G_1(\Omega_1) \tag{8.7-5}$$

where

$G_1(\Omega_1)$ = power gain of antenna 1

P_0 = power received at the antenna terminals of an isotropic antenna located at the phase center of antenna 1 (or at the phase center of antenna 2 since the range difference is negligible)

which results in

$$\frac{P_R}{P_{R1}} = 1 + \left(\frac{k_2}{k_1}\right)^2 \frac{G_2(\Omega_2)}{G_1(\Omega_1)}$$

$$+ 2\left(\frac{k_2}{k_1}\right)\sqrt{\frac{G_2(\Omega_2)}{G_1(\Omega_1)}} \cos\left[\beta(d + a_2 - a_1)\right] \qquad (8.7\text{-}6)$$

where

$G_2(\Omega_2)$ = power gain of antenna 2 and P_R/P_{R1} is the power gain (loss) due to the interferometer effect.

If the spacecraft attitude is changing, the fading should be examined as a function of time to ensure that no fades deep enough to cause loss of lock develop. In addition, for a large effect, the assumption of Gaussian noise breaks down, and the statistics of data reception must be reexamined.

The above analysis can be easily extended to more than two antennas by adding the appropriate received voltages.

References

8-1. Tai, C. T., "On the Presentation of Maxwell's Theory," *Proceedings of the IEEE*, Vol. 60, No. 8, pp. 936–945, Aug. 1972.

8-2. Love, A. W., *Reflector Antennas*, IEEE, New York, 1978.

8-3. Clarricoats, P. J. B., and Poulton, G. T., "High Efficiency Microwave Reflector Antennas – A Review," *Proceedings of the IEEE*, Vol. 65, No. 10, pp. 1470–1504, Oct. 1977.

8-4. Rusch, W. V. T., and Potter, P. D., *Analysis of Reflector Antennas*, Academic Press, New York, 1970.

8-5. Jasik, H., *Antenna Engineering Handbook*, McGraw-Hill, New York, 1961.

8-6. Wood, P. J., *Reflector Antenna Analysis and Design*, Peter Peregrinus Ltd., London, 1980.

8-7. Stutzman, W. L., and Thiele, G. A., *Antenna Theory and Design*, John Wiley & Sons, New York, 1981.

8-8. Elliot, R. S., *Antenna Theory and Design*, Prentice-Hall, New Jersey, 1981.

8-9. Galindo-Israel, V., and Rahmat-Samii, Y., "A New Look at Fresnel Field Computation Using the Jacobi-Bessel Series," *IEEE Transactions on Antennas and Propagation*, Vol. 29, No. 6, pp. 885–898, Nov. 1981.

8-10. Rahmat-Samii, Y., Galindo-Israel, V., and Mittra, R., "A Plane Polar Approach for Far-Field Construction from Near-Field Measurements," *IEEE Transactions on Antennas and Propagation*, Vol. 28, No. 2, pp. 216–230, Mar. 1980.

8-11. Rahmat-Samii, Y., and Galindo-Israel, V., "Shaped Reflector Antenna Analysis Using the Jacobi-Bessel Series," *IEEE Transactions on Antennas and Propagation*, Vol. 28, No. 4, pp. 425–435, July 1980.

8-12. Rahmat-Samii, Y., and Galindo-Israel, V., "Scan Performance of Dual Offset Reflector Antennas for Satellite Communications," *Radio Science*, Vol. 16, No. 6, pp. 1093–1099, Nov.-Dec. 1981.

8-13. Brejcha, A. G., and Smith, C. A., "Telemetry Antennas for Deep Space Probes," International Telemetry Conference, Los Angeles, Calif., Oct. 18–20, 1977.

8-14. Smith, C. A., "A Review of the State-of-the-Art in Large Spaceborne Antenna Technology," Publication 78–88, Jet Propulsion Laboratory, Pasadena, Calif., Nov. 1978.

8-15. Balanis, C. A., *Antenna Theory*, Harper and Row, New York, 1982.

8-16. *TRW Space Data*, 3rd Edition, J. B. Kenduck, ed., TRW Systems Group, 1967.

8-17. *Deep Space Network/Flight Project Interface Design Handbook*, DSN Standard Practice, 810–5, Jet Propulsion Laboratory, Pasadena, Calif., (an internal document).

Appendix

Antenna Pointing Analysis for Spinning Spacecraft

Shlomo Dolinsky

In this appendix the calculation of the total pointing error for a spinning spacecraft is summarized. The analysis is developed for the special case of the Galileo Orbiter X-band high-gain antenna. However, it can be adapted to spacecraft with similar pointing dynamics.

1. Galileo Pointing Errors

The Galileo Orbiter is a dual–spun spacecraft with a spinning high-gain antenna and as such has pointing dynamics different from a three-axis stabilized spacecraft. Table A-1 lists the errors common to spacecraft in dual- and all-spin mode. These values are mission-specific and were introduced for computational purposes. Additional errors unique to other missions can be accommodated in a like manner.

An explanation of each error follows:

(1) Angular momentum (**H**). The spacecraft angular momentum vector **H** should ideally be aligned with the spacecraft–earth line. In practice, it is impossible to prevent some drift of the spin axis with respect to earth.

Table A-1. Galileo Orbiter X-band pointing errors (3 sigma) [a]

Error source	Dual spin, mrad
Δ = control deadband size	1.64
C = attitude determination error	0.5
N = steady-state nutation torque exerted by scan platform slewing	0.5
W = wobble	0.14
Misalignment: (a) M = spin axis/mechanical boresight	0.5
(b) E = mechanical boresight/ electrical boresight	0.58

[a] In the context of a spinning spacecraft, "3 sigma" refers to that value which yields 99.73% confidence level for a Rayleigh distribution. All errors are assumed to be Rayleigh, with the exception of the deadband, which is deterministic.

Therefore, a certain amount of drift is permitted. The spacecraft attitude control system generates an estimate $\hat{\mathbf{H}}$ of the angular momentum vector to determine when this vector has drifted outside a predefined range, the deadband.

(2) Deadband (Δ). The HGA deadband controls the maximum earth-antenna drift offset. As time passes, the angular momentum–earth angle widens. When an estimate $\hat{\mathbf{H}}$ of the angular momentum vector **H** is larger than the deadband, the Galileo Orbiter motors are fired to reorient the spacecraft back toward the earth.

(3) Attitude determination error (**C**). This is the knowledge error associated with estimating the attitude of the angular momentum vector **H**. It consists of errors due to computation, drift, resolution, etc., in the gyros, start-scanner, scan platform, and accelerometers. These units are periodically calibrated in flight (every 100 to 200 days), and the updates are incorporated into the onboard control algorithm. In this way built-up biases are eliminated.

(4) Nutation (**N**). Nutation is the name given to the circular motion traced out by the principal axis of inertia under torquing actions (see Fig. A-1). It rotates at a rate approximately 30 percent higher than the spacecraft spin rate and is dependent on torques from sources such as scan platform slewing or motor firing. Dampers attenuate the nutation, reducing it to a steady-state value. Once the torquing action stops, the dampers will completely remove any residual nutation.

(5) Wobble (**W**). This is an error introduced by center-of-gravity offsets and non-zero cross products of inertia. It rotates at the spacecraft spin rate (see Fig. A-1) and cannot be reduced without a reorientation of spacecraft mass.

(6) Mechanical misalignments (**M**). These errors are a result of structural misalignments in the antenna and the antenna mounting.

(7) Electrical misalignment (**E**). This error is due to misalignment of the mechanical boresight and the electrical boresight.

Figure A-2 illustrates the relative positions of these errors. In Fig. A-3 the magnitude and direction of these errors is shown projected onto the XY plane tangent to the unit sphere centered at point O as illustrated in Fig. A-2. The total pointing error is

$$\mathbf{Z} = e_{\hat{H}\text{-}EARTH} + \mathbf{C} + \mathbf{N} + \mathbf{W} + \mathbf{M} + \mathbf{E} \qquad \text{(A-1)}$$

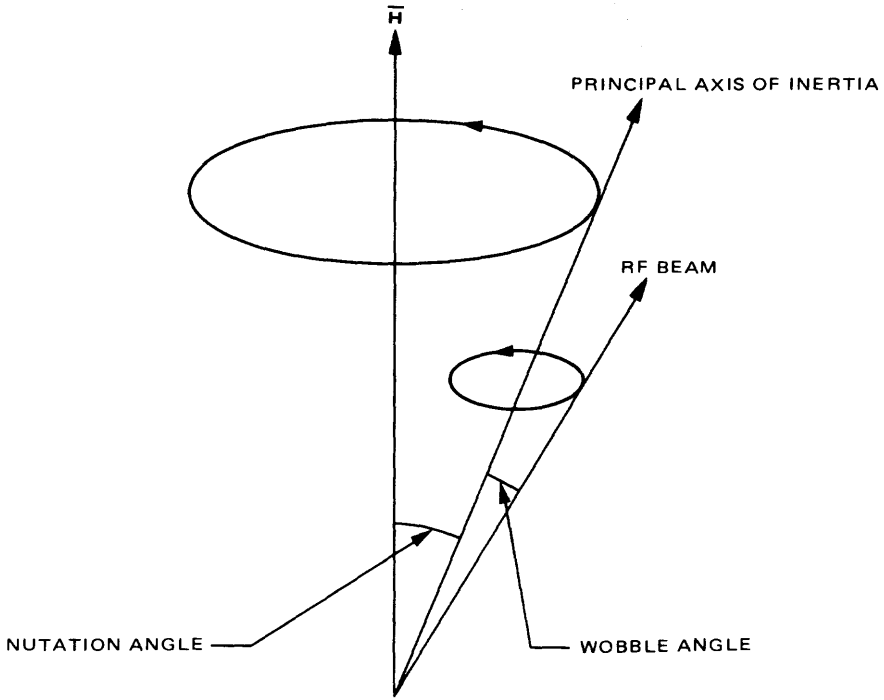


Fig. A-1. Nutation and wobble

2. Pointing Error Statistics

It is reasonable to assume that

(1) The five pointing errors defined in Section A-1 are statistically independent.

(2) Each error e has e_X and e_Y components which are independent and identically distributed zero mean Gaussian random variables.

It can be proven that

(1) The magnitude of each pointing error $|e| = \sqrt{e_X^2 + e_Y^2}$ is Rayleigh and the orientation of the error, $\tan^{-1} e_Y/e_X$, is uniform on $[0, 2\pi]$. These are natural consequences of assumption 2.

(2) The relationship between the per-axis standard deviation σ_{e_x} of the Gaussian distribution and the mean and standard deviation of $|e|$ is

$$\text{mean}(|e|) = \sqrt{\frac{\pi}{2}}\sigma_{e_x}, \sigma_{|e|} = \sqrt{\left(\frac{4-\pi}{2}\right)}\sigma_{e_x} \qquad \text{(A-2)}$$

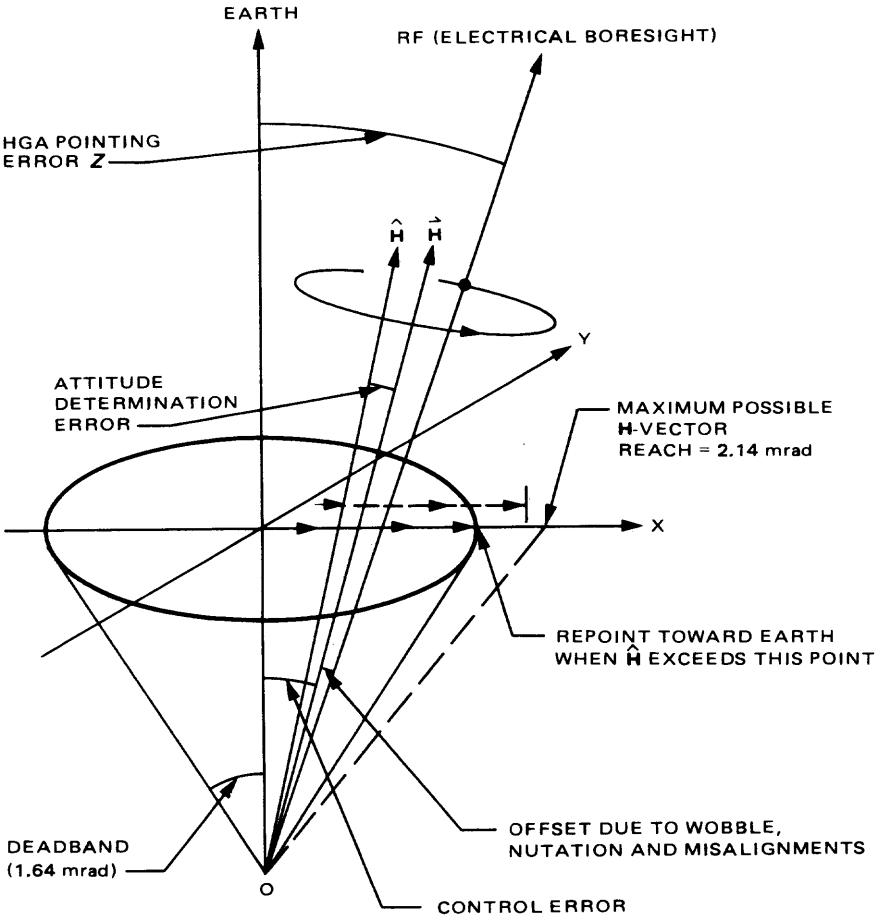


Fig. A-2. Galileo Orbiter HGA pointing errors

(3) The probability that $|e| \leqslant 1.15\ (3\ \sigma_{e_X})$ is 99.73%. Thus, the 3-sigma value for each e_X or e_Y is derived by dividing the corresponding error in Table A-1 by 1.15.

(4) The maximum pointing error Z before spacecraft tests is

$$Z = \Delta + |\mathbf{C}| + |\mathbf{N}| + |\mathbf{W}| + |\mathbf{M}| + |\mathbf{E}| \qquad \text{(A-3)}$$

which is approximated by a Gaussian distribution with

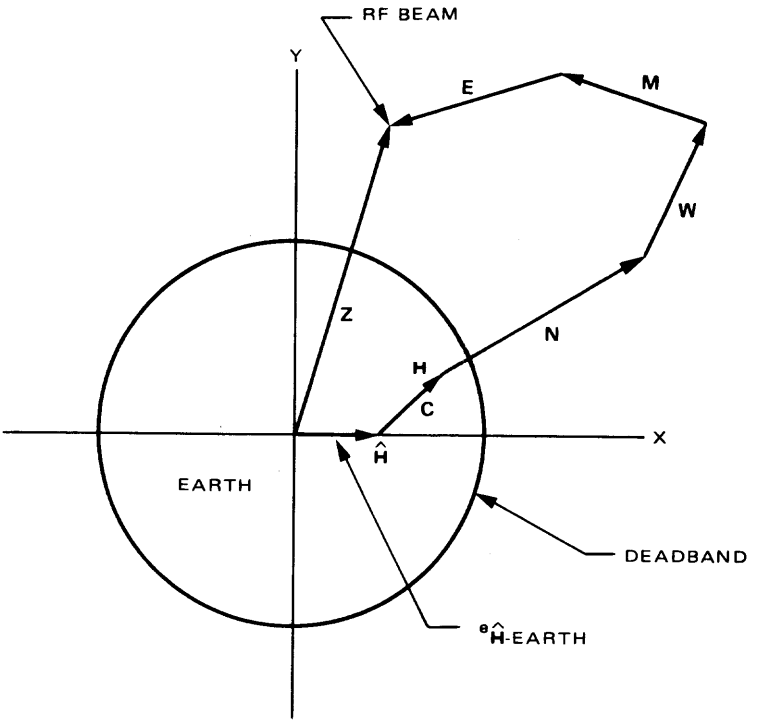


Fig. A-3. Pointing error configuration for Galileo

$$\text{mean}(Z) = \Delta + E|\mathbf{C}| + E|\mathbf{N}| + E|\mathbf{W}| + E|\mathbf{M}| + E|\mathbf{E}|$$

$$= \Delta + \sqrt{\frac{\pi}{2}}\,(\sigma_C + \sigma_N + \sigma_W + \sigma_M + \sigma_E) \qquad \text{(A-4a)}$$

$$= 2.45 \text{ mrad}$$

and

$$\text{VAR}(Z) = \text{VAR}|\mathbf{C}| + \text{VAR}|\mathbf{N}| + \text{VAR}|\mathbf{W}| + \text{VAR}|\mathbf{M}| + \text{VAR}|\mathbf{E}|$$

$$= \left(\frac{4-\pi}{2}\right)(\sigma_C^2 + \sigma_N^2 + \sigma_W^2 + \sigma_M^2 + \sigma_E^2) \qquad \text{(A-4b)}$$

$$= 0.040 \text{ mrad}$$

Here, σ_e is understood to be σ_{e_x} or the standard deviation of the X component of the error. The X has been suppressed in the notation.

(5) A total pointing error defined by

$$Z = \Delta + |\mathbf{C}| + |\mathbf{N}| + \sqrt{(W_X + M_X + E_X)^2 + (W_Y + M_Y + E_Y)^2}$$

can be approximated by the deadband size plus a Gaussian distribution with

$$\begin{aligned} \text{Mean}(Z) &= \Delta + \sqrt{\frac{\pi}{2}}(\sigma_C + \sigma_N + \sigma_{W_X+M_X+E_X}) \\ &= \Delta + \sqrt{\frac{\pi}{2}}(\sigma_C + \sigma_N + \sqrt{\sigma_W^2 + \sigma_M^2 + \sigma_E^2}) \quad \text{(A-5a)} \\ &= 2.29 \text{ mrad} \end{aligned}$$

and

$$\begin{aligned} \text{VAR}(Z) &= \left(\frac{4-\pi}{2}\right)(\sigma_C^2 + \sigma_N^2 + \sigma_{W_X+M_X+E_X}^2) \\ &= \left(\frac{4-\pi}{2}\right)(\sigma_C^2 + \sigma_N^2 + \sigma_W^2 + \sigma_M^2 + \sigma_E^2) \quad \text{(A-5b)} \\ &= 0.040 \text{ mrad} \end{aligned}$$

The magnitude $|\mathbf{W} + \mathbf{M} + \mathbf{E}|$ is also Rayleigh, thus (A-5a) and (A-5b) follow.

3. Choosing Design, Adverse, and Favorable Pointing Errors

Selecting appropriate design values and tolerances for the total pointing loss means that all possible pointing configurations must be considered. The philosophy applied to Galileo has three different configurations for design, adverse, and favorable pointing. These, in turn, define the design and tolerances for the pointing loss.

3.1 Design Value

If the Galileo pointing mechanism was error free, the maximum pointing offset would equal the deadband. Thus, the design value for pointing error is at least as big as 1.64 mrad.

Furthermore, wobble and mechanical and electrical misalignments are actualized during spacecraft construction. At the conclusion of the spacecraft tests, an exact measurement of their combined pointing offset will be available. This deterministic error will then be added to the deadband. Before manufacturing, these errors must be treated statistically. Together they can be expected to contribute to the design value a total error equal to the mean of the magnitude of their vector sum.

Finally, the attitude determination error and the nutation occur in random directions. When aligned with the aforementioned errors, their contribution is greatest. Thus they can each be expected to add the mean of their magnitude to the design value.

Stated mathematically, the potentially worst-case design value pointing error is

$$Z_{DESIGN} = \Delta + E\,(|\mathbf{C}| + |\mathbf{N}| + |\mathbf{W} + \mathbf{M} + \mathbf{E}|) = 2.29 \text{ mrad} \qquad \text{(A-6)}$$

That is, Z_{DESIGN} is given by (A-5a).

3.2 Adverse Pointing

The maximum pointing offset on Galileo happens in the unlikely event that all errors are aligned. From (A-3)

$$Z_{MAX} = \Delta + |\mathbf{C}| + |\mathbf{N}| + |\mathbf{W}| + |\mathbf{M}| + |\mathbf{E}|$$

The mean plus three-sigma value of Z_{MAX} is defined as $Z_{ADVERSE}$. Hence,

$$Z_{ADVERSE} = \Delta + \sqrt{\frac{\pi}{2}}\,(\sigma_C + \sigma_N + \sigma_W + \sigma_M + \sigma_E) + 3\sqrt{\left(\frac{4-\pi}{2}\right)(\sigma_C^2 + \sigma_N^2 + \sigma_W^2 + \sigma_M^2 + \sigma_E^2)} \qquad \text{(A-7)}$$

$$= 3.05 \text{ mrad}$$

3.3 Favorable Pointing

The worst-case minimum pointing offset on Galileo happens in the unlikely event that all the errors are aligned away from the deadband toward the earth, i.e., toward the center of the deadband. Thus,

$$Z_{MIN(worst\text{-}case)} = \Delta - (|\mathbf{C}| + |\mathbf{N}| + |\mathbf{W}| + |\mathbf{M}| + |\mathbf{E}|) \qquad \text{(A-8)}$$

The mean minus three-sigma value of $Z_{MIN(worst\text{-}case)}$ is defined as $Z_{FAVORABLE}$. Stated mathematically,

$$Z_{FAVORABLE} = \Delta - \left[\sqrt{\frac{\pi}{2}}\,(\sigma_C + \sigma_N + \sigma_W + \sigma_M + \sigma_E) + 3\sqrt{\left(\frac{4-\pi}{2}\right)(\sigma_C^2 + \sigma_N^2 + \sigma_W^2 + \sigma_M^2 + \sigma_E^2)}\right] \quad \text{(A-9)}$$

$$= 0.23 \text{ mrad}$$

4. Galileo High-Gain Antenna Pointing Loss

The Galileo high-gain antenna is modelled by

$$\text{Relative gain} = 10 \log \cos^{6.26}\left(\frac{D}{\lambda}\,\theta\right) \text{(dB)} \quad \text{(A-10)}$$

where

D = 4.8-m-diameter antenna

λ = 3.57-cm X-band wavelength

θ = angle offset from peak

Thus, the pointing losses resulting from pointing errors defined in Section 3 in this appendix are shown in Table A-2 below.

5. Galileo Pointing Loss Simulation

Over time the earth-RF beam pointing error on Galileo traces out a complex ever-widening path (Figs. A-2 and A-3). The path begins with $\hat{\mathbf{H}}$ aligned with the earth. **H** is oriented randomly within an area with a 0.5-mrad radius. As time passes, the **H** vector slowly drifts away from its initial position. (In reality, it

Table A-2. Design, adverse, and favorable pointing errors and losses

	Design	Adverse	Favorable
Pointing error, mrad	2.29	3.05	0.23
Pointing loss, dB	−1.31	−2.36	−0.01

is the earth that drifts away from **H**.) $\hat{\mathbf{H}}$ follows suit in a locus parallel to that of **H** as the onboard computer updates it every 7 seconds. Every 500 msec the HGA pointing algorithm checks to make sure $\hat{\mathbf{H}}$ has not drifted outside of the assigned deadband. When this happens, the onboard control system reorients the spacecraft toward earth and the pointing process repeats.

The wobble and misalignments cause the Galileo high-gain antenna to spin about its angular momentum vector **H** at the spacecraft spin rate. Nutation can appear suddenly and in any direction. It induces an additional spin about the **H** vector at a rate 30% above the spacecraft spin rate. With continuous torquing applied to the spacecraft, the nutation will settle down to a steady-state value. The net effect on the earth-RF beam offset is a circular looping action with a variable diameter.

A simulation has been written to quantify the pointing losses for Galileo. The simulation calculates the total pointing error:

$$Z = \sqrt{A^2 + B^2} \tag{A-11}$$

where

$$\begin{aligned} A &= x \text{ component of pointing error} \\ &= C_x + rt + R\cos 2\pi ft + |\mathbf{N}|\cos 2\pi f_N t \\ B &= y \text{ component of pointing error} \\ &= C_y + R\sin 2\pi ft + |\mathbf{N}|\sin 2\pi f_N t \end{aligned} \tag{A-12}$$

and

R = sum or RSS of ($|\mathbf{W}|$, $|\mathbf{M}|$, $|\mathbf{E}|$)

r = drift rate of earth relative to **H** = 4.36 mrad/day (maximum rate at Jupiter distance)

f = frequency of wobble and misalignments = 3.15 per/min (spacecraft spin rate)

f_N = $1.3f$

C_x, C_y = x and y components of **C**

$|\mathbf{N}|$ = steady-state nutation

The pointing loss is calculated from (A-10), by setting $\theta = Z$,

The HGA pointing simulation begins with the choice of an initial start point for **H**, which sits at a Rayleigh 99.73% confidence distance of 0.5 mrad from $\hat{\mathbf{H}}$ (see Fig. A-4). C_x and C_y are the initial **H** vector coordinates used and may be chosen deterministically or according to a per-axis Gaussian distribution with 3 sigma = 0.5/1.15 mrad. As time passes, **H** drifts to the right. In practice, when $\hat{\mathbf{H}}$ reaches the deadband, a repointing maneuver is executed. Equivalently, the program will repoint the antenna and start over when **H** reaches $D = DEADBAND + C_x$.

The first simulation is representative of the adverse case. The initial position of **H** is chosen to be along the X axis, and R is the sum of wobble plus misalignments. Values chosen for the pointing error simulations are given in Table A-3 below. They are chosen to yield a mean plus 3 sigma error of 3.05 mrad (A-7). In the second simulation R is the RSS of (mean wobble, mean misalignments) and the values are chosen to satisfy (A-6).

Galileo pointing loss over time is shown in Figs. A-5 and A-6. The former is the loss under the unusual condition that all adverse pointing errors are periodically aligned. The second is the expected loss over time. Note that the worst-case design and worst-case adverse pointing losses of Table A-2 can occur just before an antenna repointing maneuver. Since these losses can occur, the philosophy behind choosing the worst-case design and tolerance values is sound. It will not lead to unnecessarily optimistic telemetry predictions. A simulation with favorable pointing was not run since there is little loss with a 0.23-mrad offset.

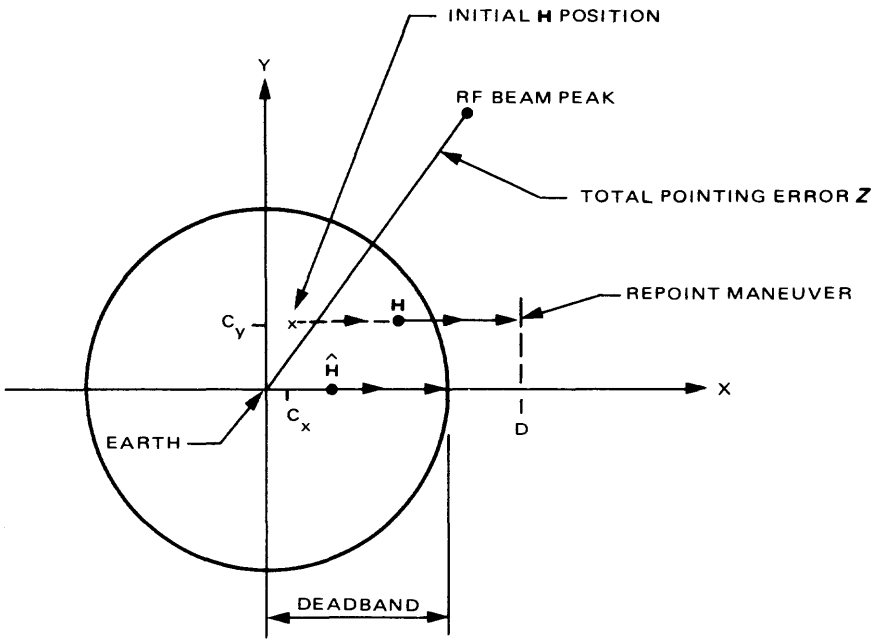


Fig. A-4. Galileo HGA pointing

Table A-3. Pointing error values chosen for the simulations

Error parameter	Values chosen for simulations, mrad	Rayleigh confidence level of each parameter, %	Maximum Pointing errors	Gaussian confidence level for sum of errors, %	Maximum pointing loss, dB
N, W, M *E*, C_x, C_y	0.31, 0.087, 0.31, 0.36, 0.31, 0.	90.91	$Z_{ADVERSE}$ = 3.05 mrad	99.87	−2.36
N, W, M *E*, C_x, C_y	0.18, 0.05, 0.18, 0.21, 0.18, 0	54 (Rayleigh mean)	Z_{DESIGN} = 2.29 mrad	*50*	−1.31

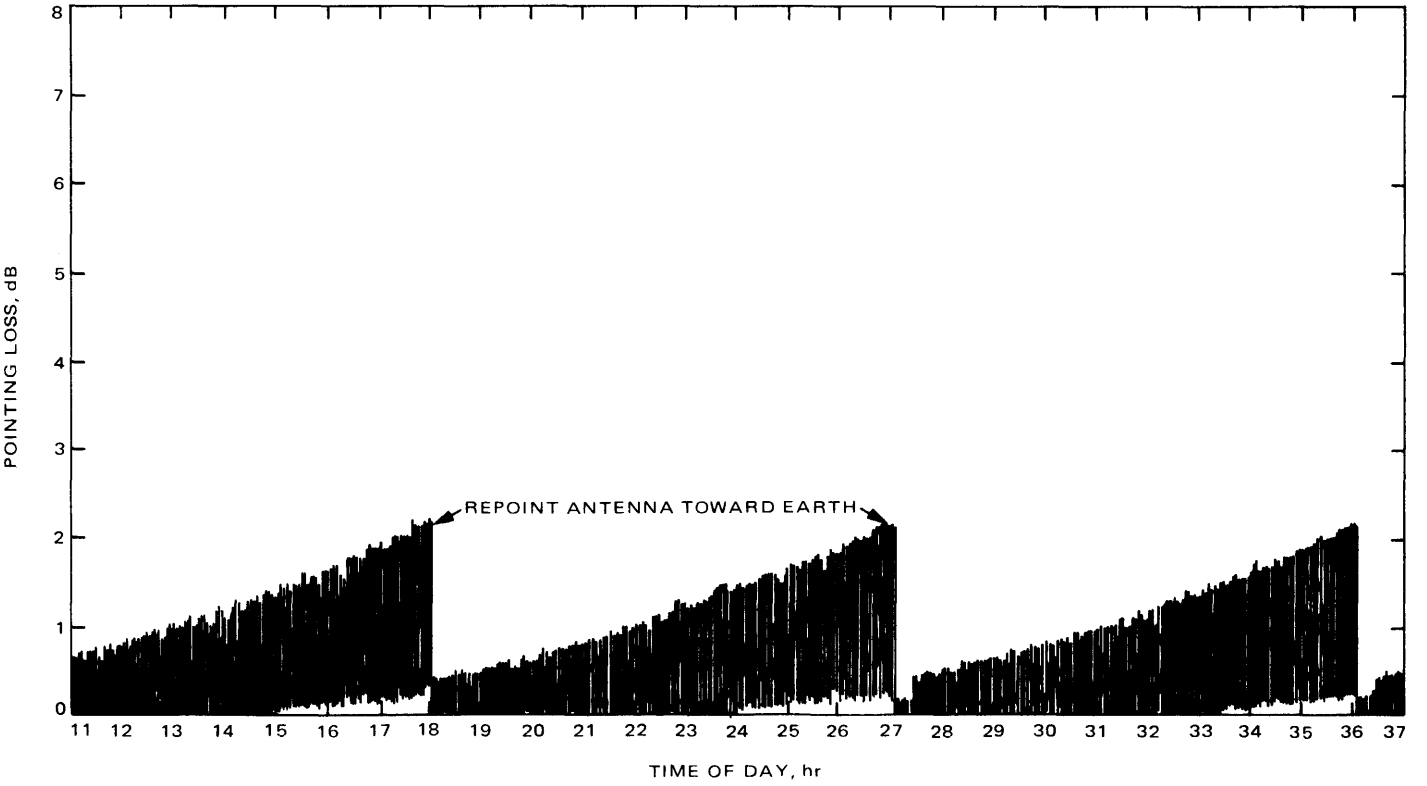


Fig. A-5. Galileo Orbiter X-band HGA pointing loss characteristics; maximum sum of errors at mean plus 3-sigma confidence level (A-7)

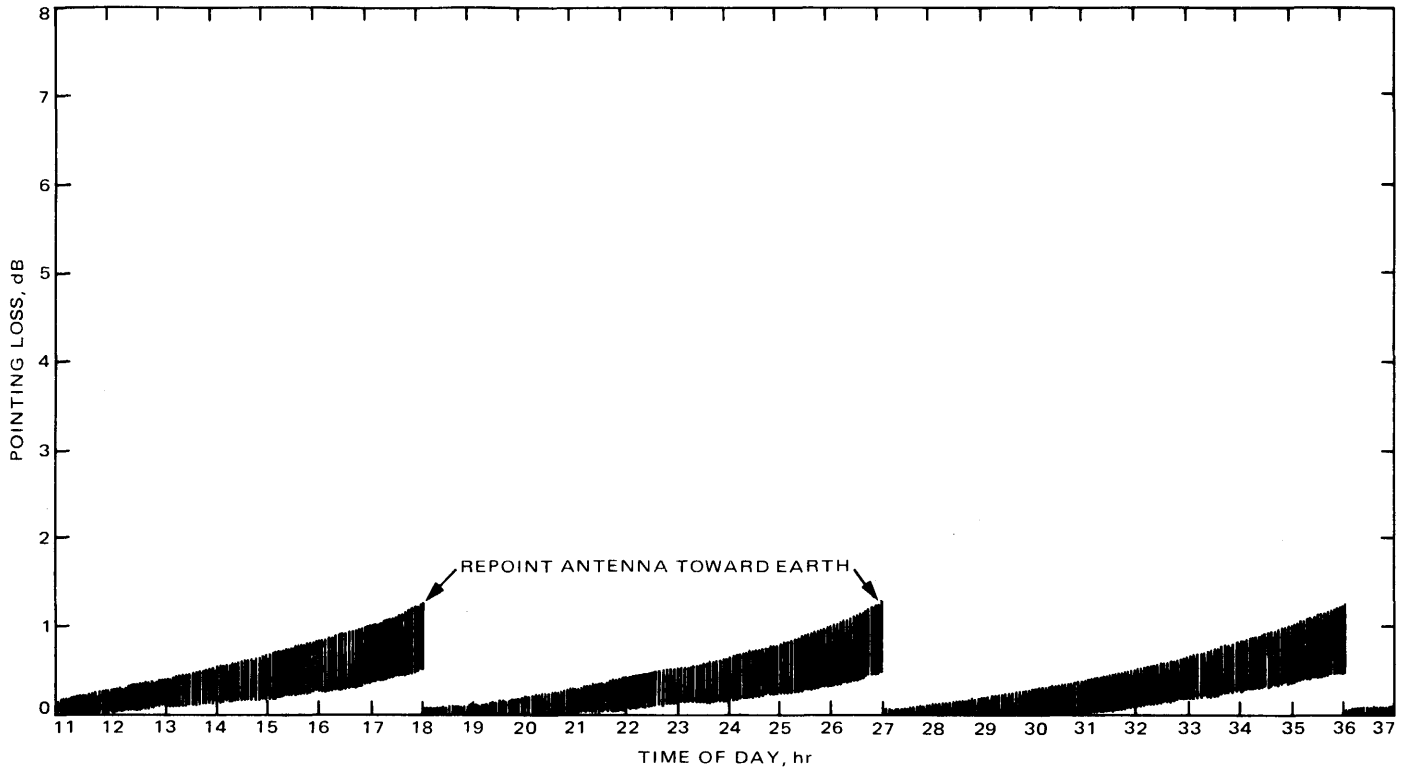


Fig. A-6. Galileo Orbiter X-band HGA pointing loss characteristics; maximum sum of errors at deadband plus attitude determination mean plus nutation mean plus RSS (wobble mean, misalignment means) (A-6)

Chapter 9
The Prediction of Weather Effects on Telecommunications Link Performance

Richard L. Horttor and Shlomo Dolinsky

9.1 Introduction

This chapter presents a theoretical model which combines the statistical description of communications link performance [9-1] with a statistical description of weather effects. The measured effects of weather, namely atmospheric attenuation and increase in ground receiver system noise temperature, are additive (in decibels). The combined probability density function is therefore the convolution of the separate density functions.

The density function for deep space communications link performance (without weather effects) derived in Chapter 1 of this book has been shown to be Gaussian by application of the central limit theorem. The next step is to derive the density of the weather effects. Both types of weather effects have been related to each other and to atmospheric cloud cover [9-2]. Since cloud cover is described by cumulative statistics and both weather effects are related to cloud cover, weather effects can also be modeled from derived cumulative statistics.

The statistical descriptions of attenuation and noise temperature derived from cloud cover are given in Tables A-1, A-2, and A-3 and Figs. A-2 through A-8 of the appendix. These same figures show data measured by a radiometer [9-3 through 9-6]. These data are useful for two reasons. They represent actual measurement of effects (instead of extrapolations of cloud models) and are less pessimistic than the estimates derived from the cloud cover model.

At this time, only Goldstone, Calif., has a statistically meaningful collection of radiometer data. However, the model is flexible enough to accommodate future measured statistics at the overseas complexes.

As a practical example of the theory developed in this chapter, it is applied in detail to the Project Galileo telecommunications link performance predictions (Section 9.4). The details of applying the model are rather tedious. However, the performance improvements, 0.6 dB and 1.8 dB depending on ground station, are significant and justify the labor.

9.2 Weather Effects at X-Band

Weather effects described in this chapter relate to downlink communications performance in the 8.4-GHz region of the spectrum. There are two related effects. One is attenuation of the received signal; the other is an increase in the ground receiver system noise temperature. The relations of the two effects to each other and to cloud cover are discussed more completely in [9-2].

9.2.1 Atmospheric Attenuation of Received Signal Power

The attenuation on received signal power P_T at station elevation angle θ above the horizon is denoted by $A(\theta)$. The value varies with ray path length through the atmosphere. Consequently, the value $A(\theta)$ is given approximately by

$$A(\theta) = \frac{A(90^\circ)}{\sin\theta} \quad \text{(dB)} \quad 5^\circ < \theta < 90^\circ \tag{9.2-1}$$

where $A(90^\circ)$ is the value at station zenith. The value $A(90^\circ)$ is characterized statistically in the appendix, as will be described later.

The power received at elevation angle θ is given by

$$P_T(\theta) = \frac{P_T}{A(\theta)} \tag{9.2-2}$$

9.2.2 Increase in Ground Receiver System Noise Temperature

The increase in system noise temperature at elevation angle θ is denoted by $\Delta SNT(\theta)$. From [9-2] $\Delta SNT(\theta)$ is related to attenuation at angle θ by

$$\Delta SNT(\theta) = 260 \cdot 10^{-0.043/10\sin\theta} \left[1 - 10^{-A(\theta)(\text{dB})/10}\right] (K) \tag{9.2-3}$$

where $\Delta SNT(90^\circ)$ is the increase at zenith. The value $\Delta SNT(90^\circ)$ is characterized statistically in the appendix, as will be described later.

Receiver noise temperature at elevation angle θ is given by

$$T_{SYS}(\theta) = T_0 + T_H + T_{EL}(\theta) + \Delta SNT(\theta) \tag{9.2-4}$$

where T_0 is the RF system temperature, T_H is the hot body noise (sun or planet in the antenna beam), $T_{EL}(\theta)$ is the deterministic increase caused by ground terrain in the antenna side lobes [9-7] and $\Delta SNT(\theta)$ is the weather component.

Receiver noise spectral density is given by

$$\begin{aligned} N_0(\theta) &= k\, T_{SYS} \\ &= k\, [T_0 + T_H + T_{EL}(\theta) + \Delta SNT(\theta)] \end{aligned} \tag{9.2-5}$$

where k is Boltzmann's constant.

The weather-induced change is normalized to the deterministic value at elevation angle θ by

$$N_0(\theta) = k\, [T_0 + T_H + T_{EL}(\theta)] \left[1 + \frac{\Delta SNT(\theta)}{T_0 + T_H + T_{EL}(\theta)}\right] \tag{9.2-6}$$

Thus, the weather-induced noise factor can be denoted separately by $Y_{\Delta SNT}$ in

$$Y_{\Delta SNT}(\theta) = \left[1 + \frac{\Delta SNT(\theta)}{T_0 + T_H + T_{EL}(\theta)}\right] \tag{9.2-7}$$

9.2.3 Temperature and Attenuation Relationships

Reference [9-2] gives the relationships between $A(\theta)$ and $\Delta SNT(\theta)$. These are repeated in (9-2.8) and 9-2.9).

$$\Delta SNT(\theta) = 260 \cdot 10^{-0.043/10 \sin\theta}\, [1 - 10^{-A(\theta)(\mathrm{dB})/10}]\ (K) \tag{9.2-8}$$

Or solving for $A(\theta)$,

$$A(\theta) = -10 \log \left[1 - \frac{\Delta SNT(\theta)}{260} \cdot 10^{0.043/10 \sin\theta}\right] (\mathrm{dB}) \tag{9.2-9}$$

The factor -0.043 dB is the X-band attenuation at zenith on a clear, dry day. The S-band factor is -0.035 dB.

9.2.4 Weather Effects on Link Performance

The principal measure of downlink performance is the ratio of received power P_T to noise spectral density N_0. The complete dependence on elevation angle and weather effects is given by

$$\frac{P_T}{N_0}(\theta) = \frac{P_T}{N_0} \frac{1}{A(\theta)\left[1 + \dfrac{\Delta SNT(\theta)}{T_0 + T_H + T_{EL}(\theta)}\right]} \tag{9.2-10}$$

where P_T is the weather-free received power and N_o is the weather-free system noise spectral density $k\,(T_0 + T_H + T_{EL}(\theta))$.

The weather degradation factor $Y(\theta)$ is given by

$$Y(\theta) = A(\theta)\left[1 + \frac{\Delta SNT(\theta)}{T_0 + T_H + T_{EL}(\theta)}\right] \tag{9.2-11}$$

or, if the translation is made to decibels

$$Y_\theta = 10 \log_{10} Y(\theta) \quad \text{(dB)} \tag{9.2-12}$$

or

$$Y_\theta = A_\theta + Y_{\Delta SNT(\theta)}\ \text{(dB)}$$

The composite weather loss factor Y_θ, will be modeled statistically and combined with the link performance distribution in later sections.

9.3 The Stochastic Weather Model

The weather degradation factor Y_θ was defined by (9.2-12). It depends on noise temperature and attenuation, which are related by (9.2-8) and (9.2-9). Since both are random variables (the inherent randomness of weather has never been seriously questioned), the cumulative distribution of Y_θ can be constructed from the cumulative distributions of either system noise temperature increase or attenuation. An example follows.

9.3.1 Stochastic Weather Data

The Deep Space Network published the cumulative statistics of cloud attenuation and system noise temperature increase [9-2]. The necessary data are contained in Tables A-1, A-2, and A-3 and Figs. A-2 through A-8 of the appendix.

Using the Goldstone June-August quarter contained in Table A-2, (9.2-12) yields values of Y_θ at cumulative probabilities 0.50, 0.70, 0.80, 0.85, 0.90, and 0.95. These are plotted in Fig. 9-1. Actually, the plot shows $1 - P(Y_\theta)$ instead of $P(Y_\theta)$ to take advantage of the logarithmic scale.

The shape of the curve suggests an exponential distribution would be a convenient fit. Such a fit is drawn also, with parameters explained below.

The exponential density is defined by (9.3-1):

$$p_w[Y(\theta)] = \begin{cases} (1-p_0)\,\delta[Y(\theta)-Y_0(\theta)] + \dfrac{p_0}{W_0(\theta)} \exp\left[-\dfrac{Y(\theta)-Y_0(\theta)}{W_0(\theta)}\right]; & Y(\theta) \geq Y_0(\theta) \\ 0\ ; & Y(\theta) < Y_0(\theta) \end{cases} \tag{9.3-1}$$

The cumulative distribution is defined by (9.3-2):

$$P_w[Y(\theta)] = \begin{cases} (1-p_0)\,u[Y(\theta)-Y_0(\theta)] + p_0\left\{1-\exp\left[-\dfrac{Y(\theta)-Y_0(\theta)}{W_0(\theta)}\right]\right\}; & Y(\theta) \geq Y_0(\theta) \\ 0\ ; & Y(\theta) < Y_0(\theta) \end{cases} \tag{9.3-2}$$

where $u(\cdot)$ and $\delta(\cdot)$ are the unit step and delta functions, respectively. The step allows flexibility in fitting the data nearing zero kelvin. The model parameters are interpreted as follows: $(1 - p_0)$ is the probability of weather losses less than or equal to $Y_0(\theta)$, and $W_0(\theta)$ is proportional to the reciprocal of the slope of the plotted cumulative function. In other words, $W_0(\theta)$ is the characteristic weather loss.

$$W_0(\theta) = -\log_{10} e/\text{slope}$$

For this example, $Y_0(90°) = 0.12$ dB, $W_0(90°) = 0.73$ dB and the probability $(1 - p_0)$ of degradation less than or equal to $Y_0(90°)$ is 0.8.

9.3.1 Stochastic Weather Data

The Deep Space Network published the cumulative statistics of cloud attenuation and system noise temperature increase [9-2]. The necessary data are contained in Tables A-1, A-2, and A-3 and Figs. A-2 through A-8 of the appendix.

Using the Goldstone June-August quarter contained in Table A-2, (9.2-12) yields values of Y_θ at cumulative probabilities 0.50, 0.70, 0.80, 0.85, 0.90, and 0.95. These are plotted in Fig. 9-1. Actually, the plot shows $1 - P(Y_\theta)$ instead of $P(Y_\theta)$ to take advantage of the logarithmic scale.

The shape of the curve suggests an exponential distribution would be a convenient fit. Such a fit is drawn also, with parameters explained below.

The exponential density is defined by (9.3-1):

$$p_w[Y(\theta)] = \begin{cases} (1-p_0)\,\delta[Y(\theta)-Y_0(\theta)] + \dfrac{p_0}{W_0(\theta)} \exp\left[-\dfrac{Y(\theta)-Y_0(\theta)}{W_0(\theta)}\right]; & Y(\theta) \geq Y_0(\theta) \\ 0\ ; & Y(\theta) < Y_0(\theta) \end{cases} \tag{9.3-1}$$

The cumulative distribution is defined by (9.3-2):

$$P_w[Y(\theta)] = \begin{cases} (1-p_0)\,u[Y(\theta)-Y_0(\theta)] + p_0\left\{1-\exp\left[-\dfrac{Y(\theta)-Y_0(\theta)}{W_0(\theta)}\right]\right\}; & Y(\theta) \geq Y_0(\theta) \\ 0\ ; & Y(\theta) < Y_0(\theta) \end{cases} \tag{9.3-2}$$

where $u(\cdot)$ and $\delta(\cdot)$ are the unit step and delta functions, respectively. The step allows flexibility in fitting the data nearing zero kelvin. The model parameters are interpreted as follows: $(1 - p_0)$ is the probability of weather losses less than or equal to $Y_0(\theta)$, and $W_0(\theta)$ is proportional to the reciprocal of the slope of the plotted cumulative function. In other words, $W_0(\theta)$ is the characteristic weather loss.

$$W_0(\theta) = -\log_{10} e/\text{slope}$$

For this example, $Y_0(90°) = 0.12$ dB, $W_0(90°) = 0.73$ dB and the probability $(1 - p_0)$ of degradation less than or equal to $Y_0(90°)$ is 0.8.

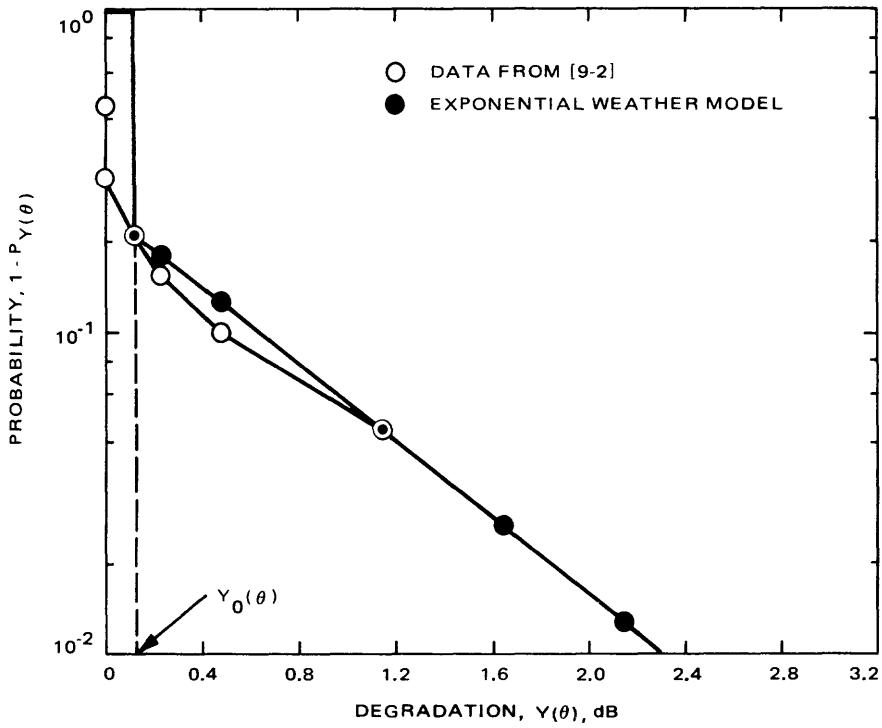


Fig. 9-1. Cumulative distribution for weather degradation $Y(\theta)$, Goldstone, June–August quarter, elevation angle $\theta = 90$ deg

The above exercise must be performed at each elevation angle because the weather losses are elevation dependent. The resulting values of the parameters at several elevation angles are listed in Table 9-1. The table includes values at each complex (Goldstone, Madrid and Canberra) for annual quarters.

9.3.2 Radiometer Data

Results thus far have been based on theoretical models relating visually observed cloud cover to system temperature increases and attenuation. Two objections arise. One challenges the theoretical cloud model because direct physical measurements would be preferable. The other is practical, namely, that accurate cloud cover estimates are possible only during daylight hours. Consequently, half the time little data is available.

The water vapor radiometer answers both objections. It directly measures system noise temperature (increases) from which attenuation can be computed, and it can operate continuously. At this writing, only Goldstone, Calif., has a reasonably complete statistical base.

Table 9-1. Cloud-cover-based weather model parameter values

		Elevation deg										
	θ	90	60	45	35	30	25	20	15	10	6	
Goldstone												
Dec.-Feb.	$Y_0(\theta)$	0.386	0.424	0.483	0.555	0.599	0.666	0.761	0.898	1.101	1.437	
	$W_0(\theta)$	1.089	1.176	1.308	1.463	1.553	1.687	1.868	2.116	2.473	3.031	
Mar-May	$Y_0(\theta)$	0.190	0.209	0.239	0.276	0.298	0.333	0.382	0.454	0.561	0.742	
	$W_0(\theta)$	0.598	0.651	0.734	0.832	0.890	0.978	1.100	1.269	1.515	1.903	$p_0 = 0.2$
June-Aug.	$Y_0(\theta)$	0.123	0.136	0.155	0.179	0.194	0.217	0.249	0.296	0.368	0.489	
	$W_0(\theta)$	0.734	0.799	0.900	1.020	1.091	1.200	1.346	1.553	1.853	2.327	
Sept.-Nov.	$Y_0(\theta)$	0.137	0.150	0.172	0.199	0.215	0.240	0.276	0.328	0.407	0.540	
	$W_0(\theta)$	0.333	0.364	0.413	0.472	0.507	0.561	0.636	0.742	0.898	1.149	
Canberra												
Dec.-Feb.	$Y_0(\theta)$	0.123	0.136	0.155	0.179	0.194	0.217	0.249	0.296	0.368	0.489	
	$W_0(\theta)$	1.439	1.550	1.716	1.910	2.024	2.192	2.419	2.733	3.190	3.922	
Mar.-May	$Y_0(\theta)$	0.083	0.091	0.104	0.120	0.130	0.146	0.167	0.200	0.248	0.331	
	$W_0(\theta)$	1.053	1.140	1.272	1.428	1.519	1.656	1.841	2.099	2.472	3.062	$p_0 = 0.3$
June-Aug.	$Y_0(\theta)$	0.190	0.209	0.239	0.276	0.298	0.333	0.382	0.454	0.561	0.742	
	$W_0(\theta)$	1.228	1.325	1.470	1.640	1.739	1.886	2.085	2.358	2.755	3.385	
Sept.-Nov.	$Y_0(\theta)$	0.137	0.151	0.172	0.199	0.215	0.240	0.276	0.328	0.407	0.540	
	$W_0(\theta)$	1.615	1.736	1.917	2.127	2.250	2.431	2.675	3.013	3.508	4.308	
Madrid												
Dec.-Feb.	$Y_0(\theta)$	0.096	0.106	0.121	0.140	0.152	0.169	0.195	0.232	0.288	0.384	
	$W_0(\theta)$	1.356	1.457	1.607	1.781	1.883	2.034	2.237	2.518	2.932	3.600	
March-May	$Y_0(\theta)$	0.437	0.479	0.546	0.627	0.676	0.751	0.857	1.010	1.236	1.609	
	$W_0(\theta)$	2.039	2.174	2.376	2.606	2.741	2.937	3.200	3.564	4.115	5.039	$p_0 = 0.3$
June-Aug.	$Y_0(\theta)$	0.014	0.015	0.018	0.020	0.022	0.025	0.028	0.034	0.042	0.057	
	$W_0(\theta)$	0.922	1.001	1.122	1.265	1.350	1.477	1.650	1.893	2.244	2.803	
Sept.-Nov.	$Y_0(\theta)$	0.204	0.224	0.256	0.295	0.319	0.356	0.408	0.484	0.599	0.791	
	$W_0(\theta)$	1.412	1.519	1.680	1.867	1.977	2.139	2.356	2.656	3.094	3.793	

For comparison purposes, four years of measured data are drawn along with cloud cover data (Table A-2) on Fig. 9-2. The curves labeled model A and model B represent two possible exponential fits. The succeeding example chooses model B. The corresponding weather loss cumulative function is drawn in Fig. 9-3.

Thus, for the Goldstone June-August quarter, the model parameters are $Y_0(90°) = 0.36$ dB, $(1 - p_0) = 0.88$ and $W_0(90°) = 0.37$ dB. Again, each elevation angle has different statistics. Table 9-2 shows the parameter values for several elevation angles for each quarter at Goldstone.

9.4 The Combined Model for Link and Weather Effects

Downlink telecommunications performance indicators, such as (P_T/N_0), ground received SNR, and the telemetry performance margin, are aggregates of 10 to 15 independent parameters, each characterized with favorable and adverse tolerances and a suitable density function. Chapter 1 has shown how to describe these aggregates statistically by assuming that they are normal with zero mean

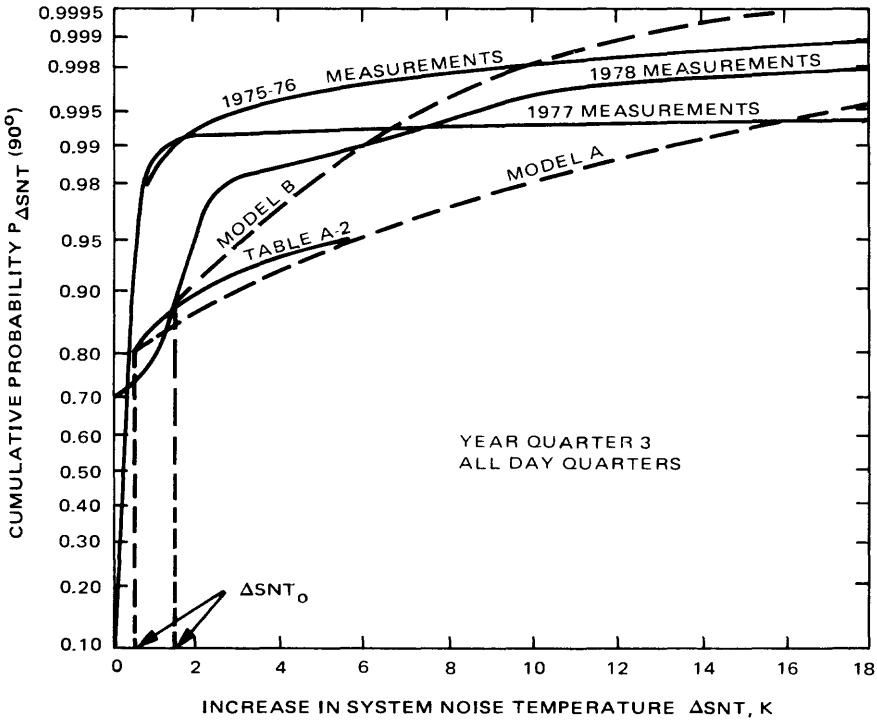


Fig. 9-2. Cumulative distribution function for zenith noise temperature increase, Goldstone, June–August quarter

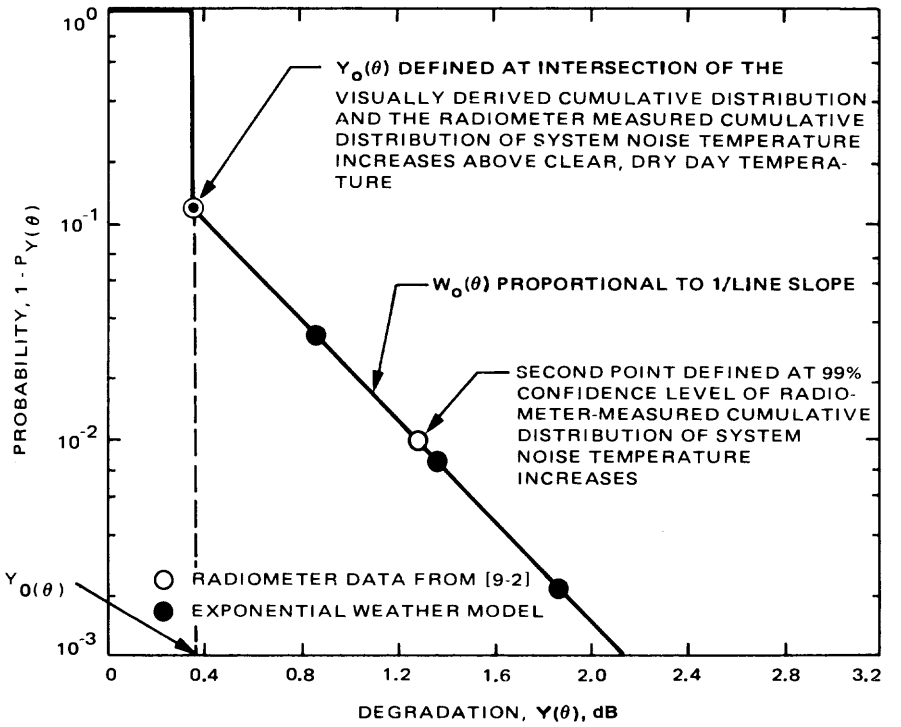


Fig. 9-3. Cumulative distribution for weather degradation Y_θ, Goldstone, June–August quarter, elevation angle $\theta = 90$ deg

and have a standard deviation equal to the rss of the individual parameter standard deviations. The random variable Z_θ in the link performance can then be described by

$$Z_\theta(\text{dB}) = X(\text{dB}) - Y_\theta(\text{dB}) \tag{9.4-1}$$

where X is the link uncertainty and Y_θ is the weather-induced degradation.

The density function for the total link plus weather degradation Z_θ is the convolution of (9.3-1) and a Gaussian density $N(0, \sigma^2)$. The resulting cumulative function of Z is given in (9.4-2).

$$P_T[Z(\theta)] = \Phi\left[\frac{Z(\theta) + Y_0(\theta)}{\sigma}\right] + p_0 \exp\left[\frac{\sigma^2}{2\,W_0^2(\theta)} + \frac{Z(\theta) + Y_0(\theta)}{W_0(\theta)}\right]$$

$$\times\ \Phi\left\{-\left[\frac{Z(\theta) + Y_0(\theta)}{\sigma} + \frac{\sigma}{W_0(\theta)}\right]\right\} \tag{9.4-2}$$

Table 9-2. Radiometer-based weather model parameter values, Goldstone

		Elevation, deg										
	θ	90	60	45	35	30	25	20	15	10	6	
Dec.-Feb.	$Y_0(\theta)$	0.399	0.438	0.499	0.573	0.618	0.687	0.785	0.926	1.135	1.481	$p_0 = 0.2$
	$W_0(\theta)$	0.617	0.664	0.737	0.820	0.869	0.941	1.038	1.171	1.363	1.665	
Mar.-May	$Y_0(\theta)$	0.360	0.396	0.451	0.519	0.560	0.623	0.712	0.841	1.032	1.350	$p_0 = 0.15$
	$W_0(\theta)$	0.482	0.522	0.582	0.653	0.694	0.756	0.839	0.954	1.119	1.377	
June-Aug.	$Y_0(\theta)$	0.360	0.396	0.451	0.519	0.560	0.623	0.712	0.841	1.032	1.350	$p_0 = 0.12$
	$W_0(\theta)$	0.370	0.402	0.450	0.507	0.541	0.591	0.660	0.754	0.890	1.103	
Sept.-Nov.	$Y_0(\theta)$	0.190	0.209	0.239	0.276	0.298	0.333	0.382	0.454	0.561	0.742	$p_0 = 0.17$
	$W_0(\theta)$	0.981	1.055	1.165	1.292	1.366	1.476	1.624	1.827	2.126	2.606	

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy \tag{9.4-3}$$

The mean value of Z is given by

$$E(Z(\theta)) = - [Y_0(\theta) + p_0 W_0(\theta)] \tag{9.4-4}$$

The density function form of $p_T(Z(\theta))$ is illustrated in Fig. 9-4 for Goldstone, June to August quarter at several elevation angles. Figure 9-5 does the same for $P_T(Z(\theta))$. Weather model parameters are listed in Table 9-2.

The function $Z(\theta)$ is the real product of the model. That is, it shows how much margin in dB the design value of a communications link must have at an elevation angle θ in order to maintain a given level of confidence. Clearly, margin changes with elevation angle, though not significantly at angles above 30 deg.

By numerically iterating through (9.4-2), $Z(\theta)$ may be found in terms of a desired confidence level, defined as $1 - P_T(Z)$, and the specified model parameters. For the last example, $Z(\theta)$ is plotted in Fig. 9-6. It applies to Goldstone in the June-August quarter, with weather model parameters specified by Table 9-2.

9.5 Time-Dependent Statistics and Performance

Since the weather data base is given only for discrete daily or yearly quarters, the model parameters can be evaluated only for these periods. To find the total degradation $Z_{TOTAL}(\theta(t))$ for an intermediate time t, the total degradation is evaluated at two quarters, designated I and II below, adjacent to the intermediate time t, and then weighted accordingly. Thus,

$$Z_{TOTAL}(\theta(t)) = a_1(t)\, Z(\theta; \text{quarter I}) + a_2(t)\, Z(\theta; \text{quarter II})$$

$$\text{quarter I} \leqslant t \leqslant \text{quarter II} \tag{9.5-1}$$

where

$$a_1(t) + a_2(t) = 1 \tag{9.5-2}$$

for all t, and $a_1(t)$ and $a_2(t)$ are unitary weighting functions.

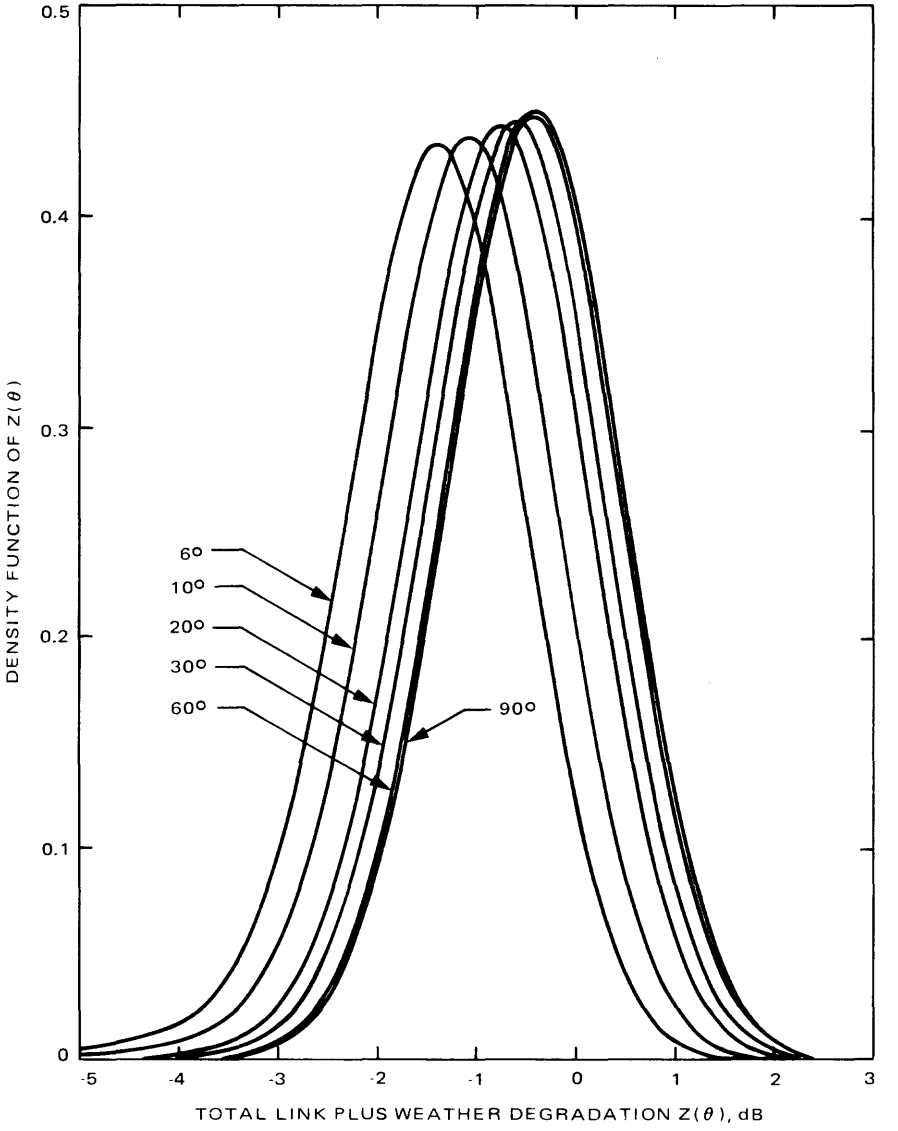


Fig. 9-4. Density function of uncertainties in end-to-end deep space link, Goldstone, June–August quarter, includes radiometer data

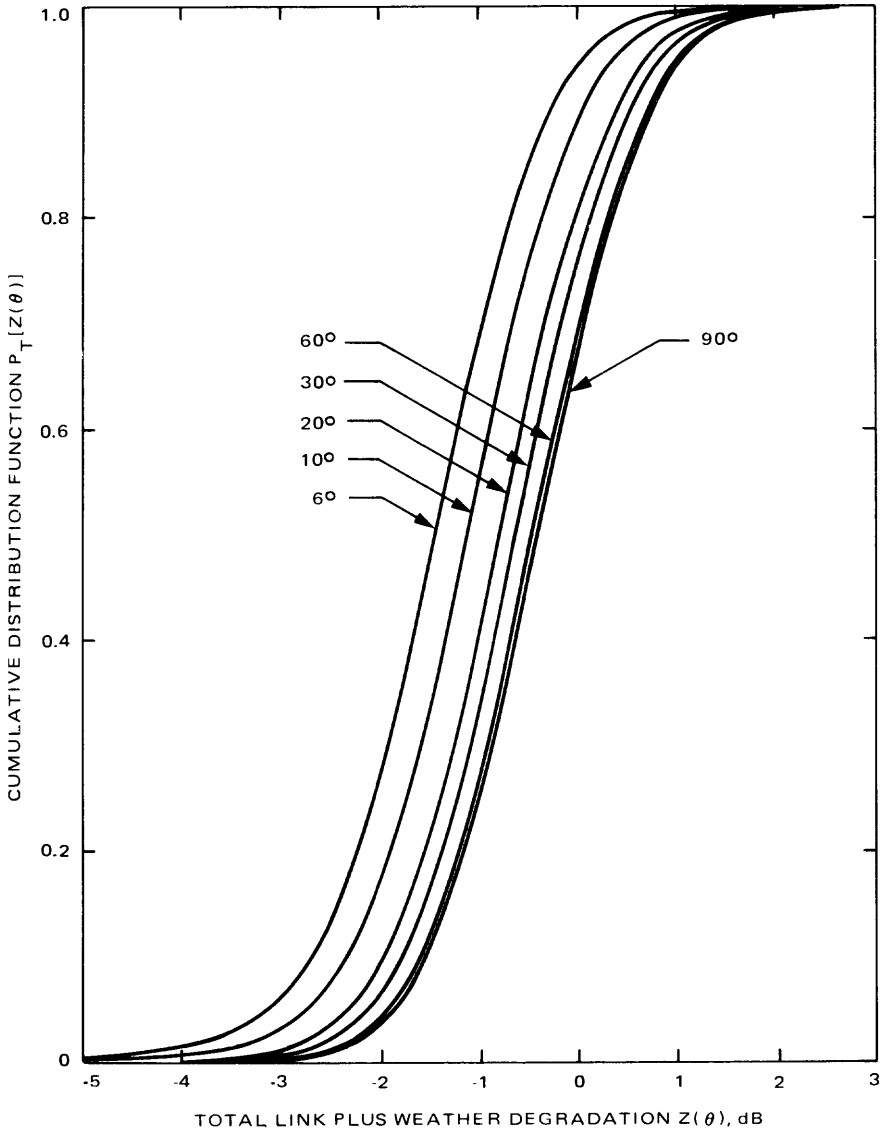


Fig. 9-5. Distribution function of uncertainties in end-to-end deep space link (Goldstone, June–August quarter, includes radiometer data)

All other factors being equal, there can be no continuity in degradation values calculated by the hour or day, since there is no continuity in the data base from hour-to-hour or quarter-to-quarter. Data appearing in Tables A-1 and A-2 are only averages over these time periods: 2100-0300 hours, 0300-0900 hours, 0900-1500 hours and 1500-2100 hours; December to February, March to May, June to August and September to November. Adapting the combined link plus weather model to intermediate times means that the degradations computed at two adjacent time periods must be weighted. Linear weighting has been the choice, though nothing in the model precludes another selection, such as cubic or $\cos^2$ functions.

If hourly resolution is desired, the expression for $Z_{TOTAL}(\theta(t))$ is given by

$$Z_{TOTAL}(\theta(t)) = \sum_j b_j[\theta(t)] \sum_i a_i[\theta(t)]\, Z_{ij}[\theta(t)],$$

$$\sum_i a_i(t) = \sum_j b_j(t) = 1 \qquad (9.5\text{-}3)^1$$

where

$Z_{ij}[\theta(t)]$ = total link plus weather degradation for day quarter i, year quarter j, at time t

$a_i(t)$, $b_j(t)$ = unitary weighting functions of Figs. 9-7 and 9-8

In both cases, Z_{TOTAL} is evaluated at each of several elevation angles. By linearly interpolating between the $Z_{TOTAL}(\theta)$ values the total degradation $Z(t)$ corresponding to the elevation profile is computed. Higher-order interpolation techniques can also be implemented.

9.6 An Example: Galileo Communications Link Design

A telecommunications link prediction program was written for the Galileo Project utilizing the results described in this chapter. Given a confidence level, the weather model parameters applicable to the date of interest, a standard deviation for the link and the ground station elevation profile, a computer search algorithm computes the total weather plus link degradation as a function of elevation angle. Since elevation angle is related to time by spacecraft trajectory, degradation is therefore related to time. However, the weather statistics are

[1] Note that the elevation dependence of Z in (9.5-1) and (9.5-2) has been replaced by time dependence in (9.5-3). Elevation and time are interchangeable since, for a pass, specifying one also specifies the other.

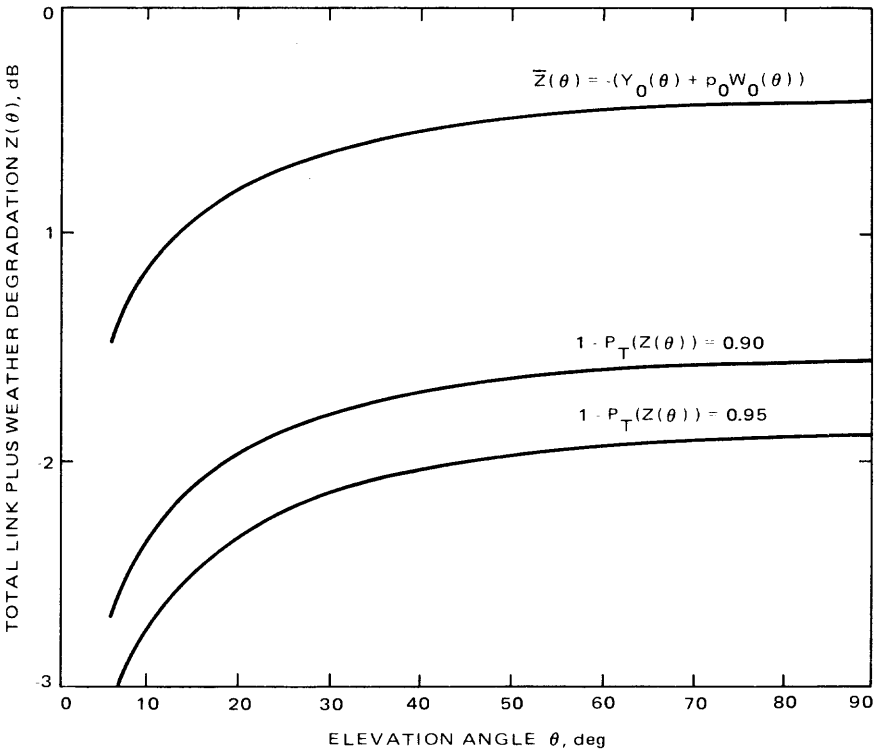


Fig. 9-6. Degradation in the deep space end-to-end link uncertainties at mean, 90, 95% confidence for $6° \leq \theta \leq 90°$, Goldstone, June–August quarter, including radiometer

available only for day quarters, year quarters, or all year. The total degradation at a given hour on a given day must be calculated from a weighted sum of computed degradations on two adjacent day quarters and year quarters. A performance prediction at the given confidence level is obtained by subtracting the total weather plus link degradation from weather-free estimates of P_T/N_0, ground carrier SNR, symbol telemetry, or other link observables.

Galileo had been using an ad hoc procedure wherein adverse link performance predictions were computed by subtracting a degradation corresponding to 90% yearly weather and 2-sigma link uncertainty from weather-free expected value link performance predictions. Accounting for link tolerances and weather effects in this manner resulted in optimistic estimates of the mean link performance and unnecessarily conservative estimates of the adverse link performance. Moreover, little flexibility was available for choice of confidence levels or time-varying weather statistics. As it turns out, the margin between the previous and current refined link performance predictions is significant. In the example developed below, the difference between the two measures of adverse link performance ranges from 0.6 dB at Goldstone to 1.8 dB at Madrid.

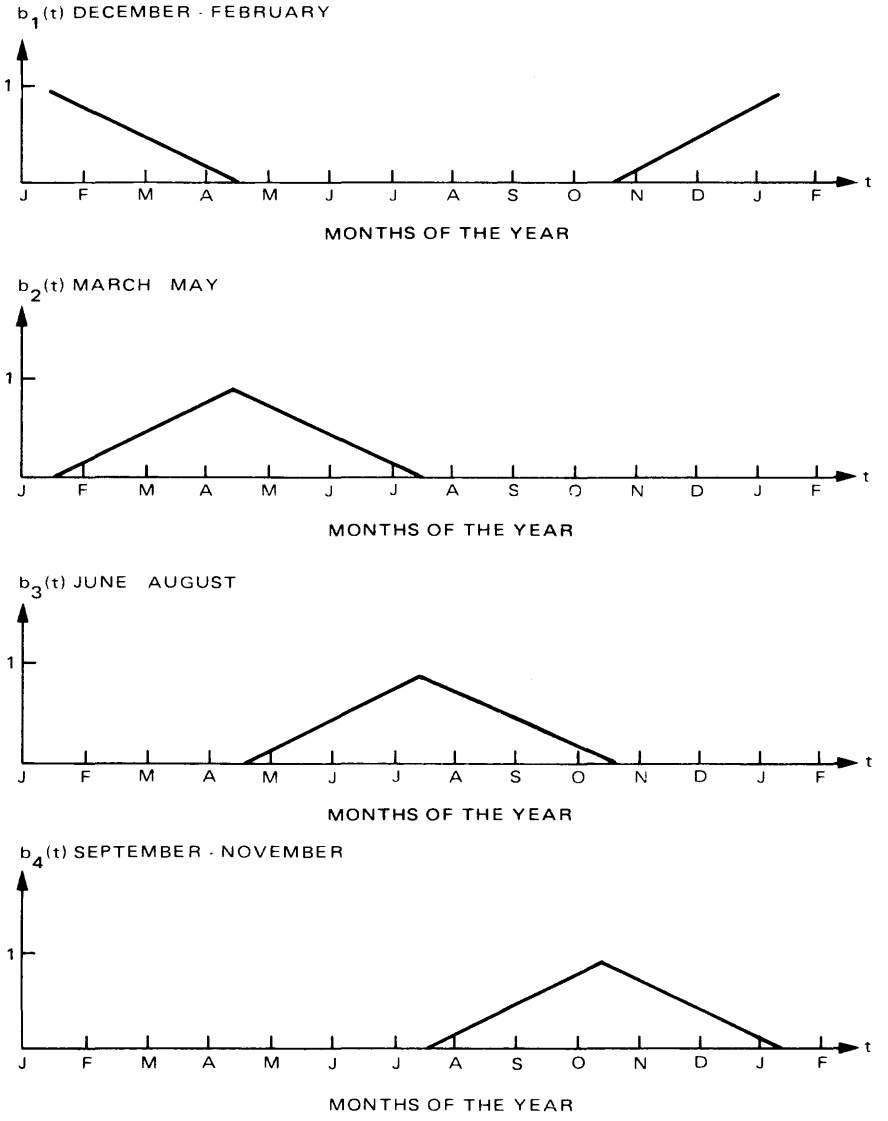


Fig. 9-7. The year quarters weighting function

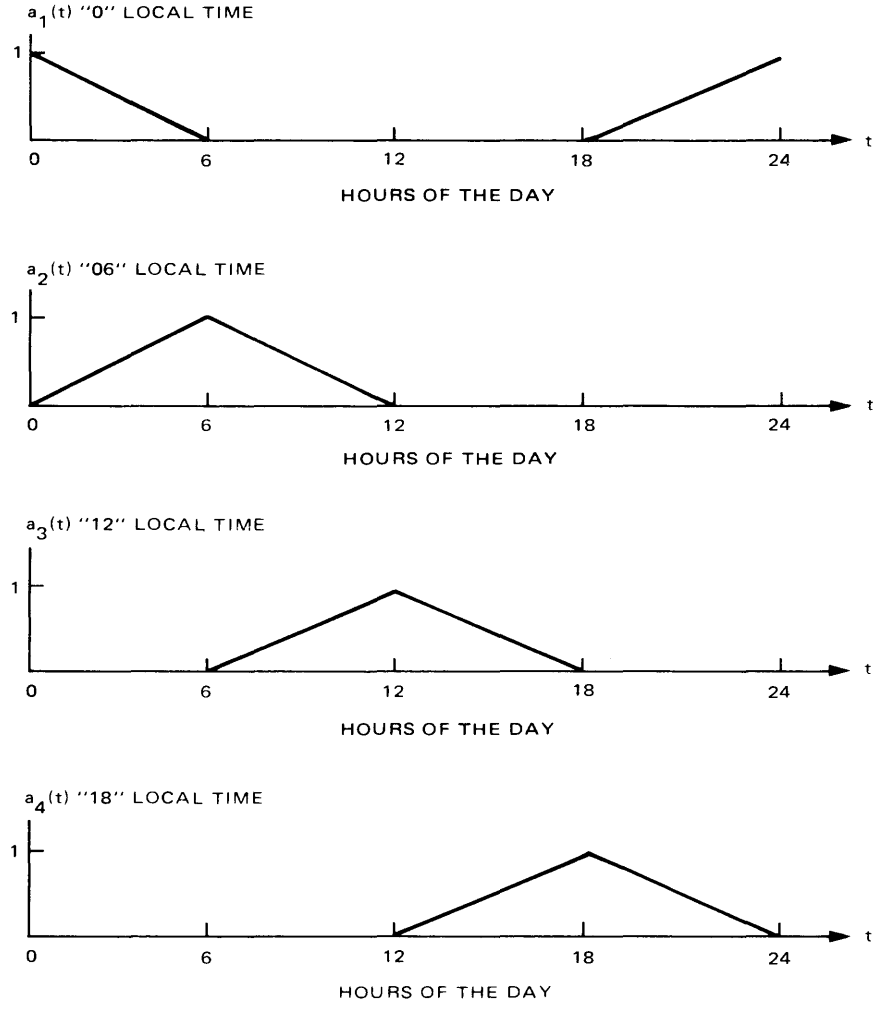


Fig. 9-8. The day quarters weighting function

9.7 The Galileo Link Prediction Algorithm

The Galileo link prediction program calculates the total link and weather degradation $Z(\theta(t))$ vs time. To run it, the user specifies the Deep Space Station (DSS), day of year, and desired confidence levels. He must also decide whether to use day quarters (Table A-1) and/or yearly quarters (Table A-2) data.

After the user specifies which data to use, the program selects the appropriate weather parameter values. If year quarterly data is desired, Tables 9-1 and 9-2 yield the model parameters. With the chosen confidence levels at hand the pro-

gram searches for the corresponding $Z(\theta)$ by iterating through (9.4-2). The $Z(\theta)$ is evaluated at several elevation angles ranging from 6 to 90 deg, for each of two adjacent quarters and then weighted. Sufficient values of $Z(\theta)$ are now computed to allow a smooth curve presentation of results.

If the confidence level specified is the mean of $Z(\theta)$, the program computes $E(Z(\theta))$ directly from (9.4-4). A similar algorithm is followed when $Z(\theta)$ is calculated for day quarters, only it is evaluated at several elevation angles at two adjacent day quarters and two adjacent year quarters.

A note of caution is advised. Daily variations, while useful, suffer two disadvantages. Six-hour periods are probably too short to average statistical inconsistencies. Also the basic data are derived from cloud cover, hence are pessimistic. Table A-2 averages the daily statistics given in Table A-1 and is therefore more reliable. Choosing yearly quarter data (Table A-2) has the added advantage that model parameters have incorporated radiometer-measured data for a more accurate evaluation of weather effects. For the time being these results are available only at Goldstone.

Once $Z(t)$, the total link plus weather degradation, has been computed for a given hour and day at the desired confidence level, the Galileo Link Prediction Program subtracts it from the weather-free estimate of P_T/N_0. The resulting profile is the link performance at the desired confidence level.

The Galileo project specifies the link design value to be the expected value (approximately the 50% confidence level) of the overall link performance and the link adverse value to be the 90% confidence level of the overall link performance. The telecommunications predictions for clear, dry and (2-sigma, 90% weather) appear as profiles A and D, respectively, in Fig. 9-9, while the expected value and 90% confidence level of the link (including weather) are plotted as curves B and C, respectively.

Figure 9-10 shows the same conditions, except for curve C, which was computed for 95% confidence. To summarize both figures, profile A is the sum of the (weather-free) link parameter design values; profile B adds the expected value weather loss; profile C is the 90 or 95% confidence level, respectively; profile D is for −2 sigma link deviation plus 90% (yearly) weather loss (Table A-3) relative to clear, dry, or curve A.

Values used in the link-plus-weather model were taken from Tables 9-1 and 9-2. Goldstone incorporated a combination of the most pessimistic radiometer-measured data and cloud cover data while Canberra and Madrid used attenuation data only. Since the day for this pass is August 14, $Z(\theta)$ was first computed for the June-August and September-November quarters over the specified elevation angles. The weather-free link performance standard deviation is

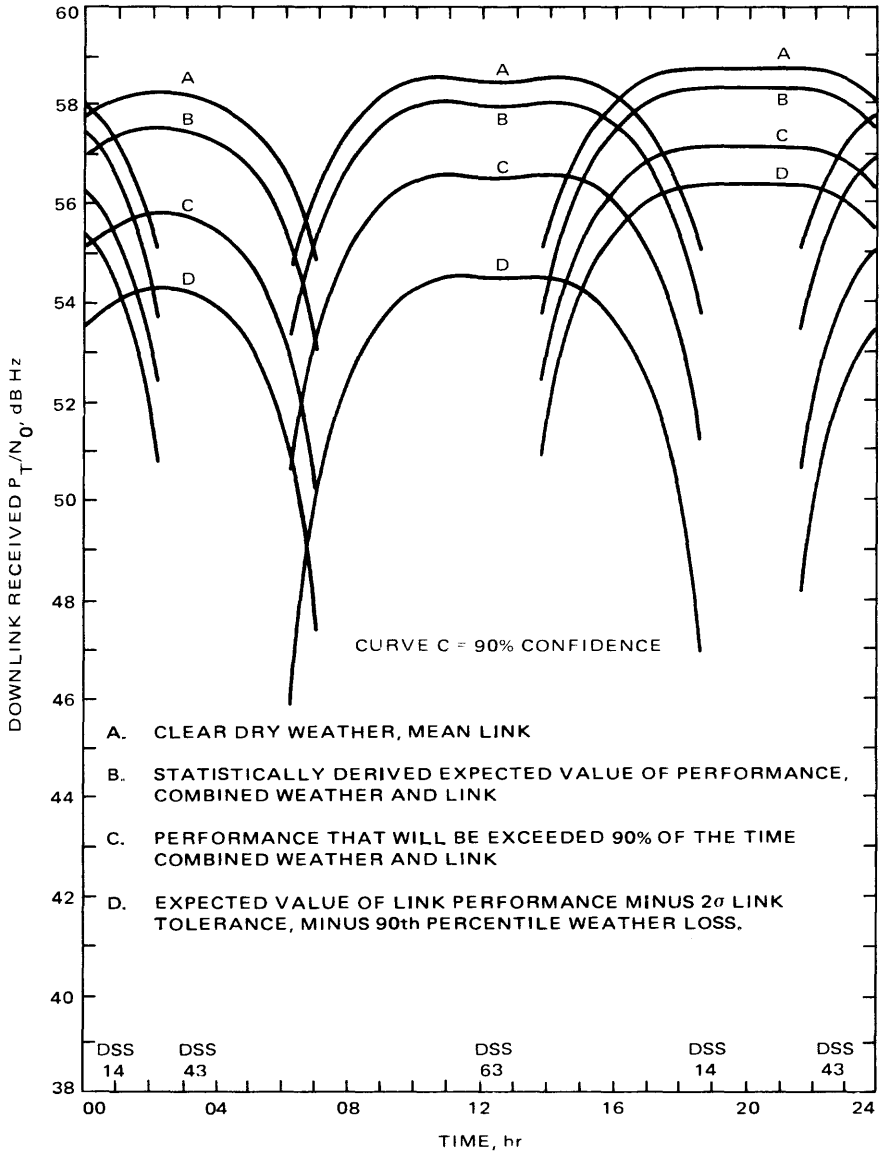


Fig. 9-9. Downlink received P_T/N_o vs time on Aug. 14, 1991; curve C = 90% confidence

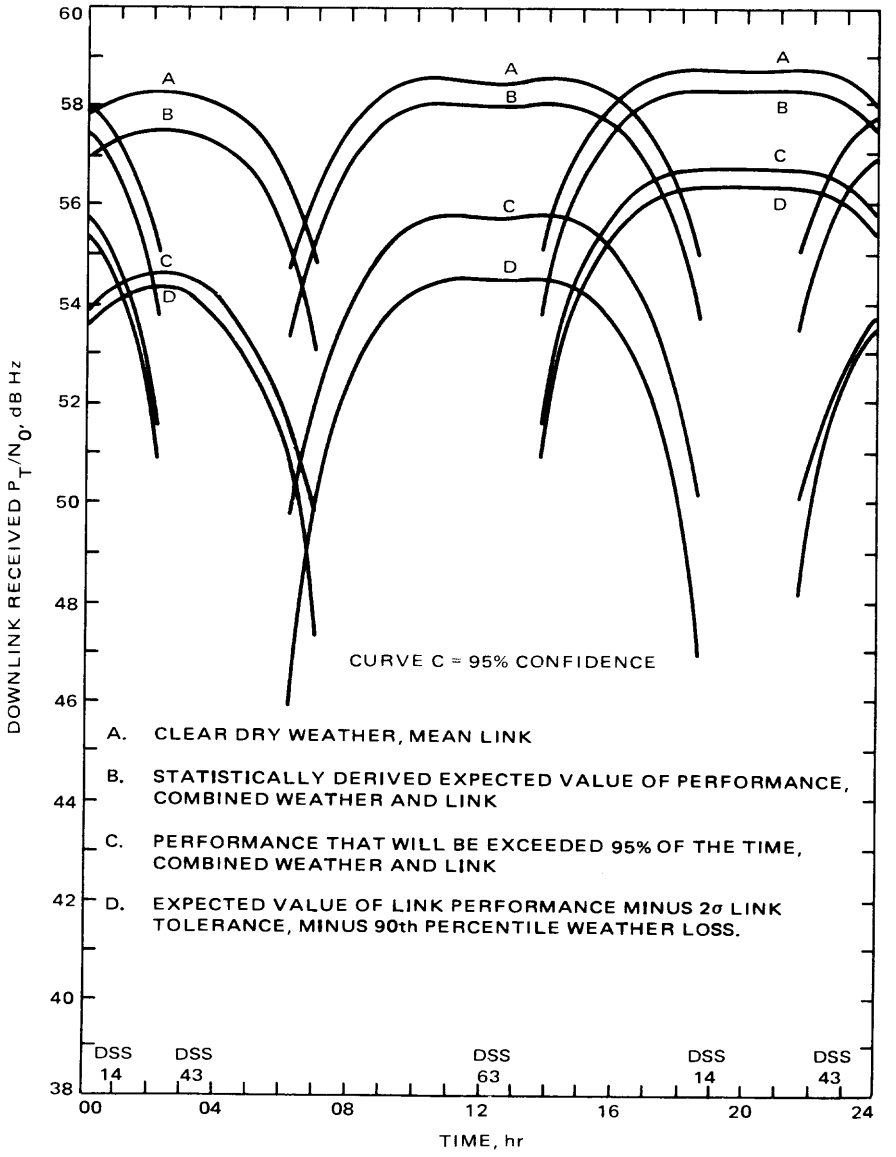


Fig. 9-10. Downlink received P_T/N_0 vs time on Aug. 14, 1991; curve C = 95% confidence

0.87 dB. Using a weighted average corresponding to the date, $Z_{TOTAL}(\theta(t))$ was derived from (9.5-1). The total degradation Z(t) corresponding to the elevation profile was computed by linear interpolation of the $Z_{TOTAL}(\theta(t))$ values at each angle of θ. These steps were repeated for the mean, 90% and 95% confidence levels at each of the three DSS stations. Defining profiles B and C followed subtraction of the quarterly weighted degradations $Z(t)$ from profile A.

Profile D is the conventional way that Galileo specified adverse link conditions. It is computed by subtracting from profile A the sum of a 90% yearly weather contribution (Table A-3) and a tolerance contribution equivalent to a 99.72% (two sigma) confidence level. The resulting difference between C and D at DSS 43 (Fig. 9-9) is attributed to inherent conservatism of the previous estimates of adverse link performance. Specifically, it is a consequence of the following:

(1) 0.1 dB more degradation in C due to the incorporation of quarterly rather than yearly data.

(2) 1.0 dB less degradation in C attributed to the difference between the 90% degradation predicted by the link-plus-weather model and the degradation arising from adding the individual 90% degradations of the link and weather.

(3) 0.6 dB less degradation in C due to the difference between a 90% (1.28 sigma) and a 99.7% (2 sigma) link performance level.

References

9-1 *A Practical Statistical Model for Telecommunications Performance Uncertainty*, Technical Memorandum 33-372, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1975.

9-2 Brockman, M. H., "DSN Telecommunications Interfaces, Atmospheric and Environmental Effects," 810-5, Rev. D, TCI-40, Rev. A, Jet Propulsion Laboratory, Pasadena, Calif., Sept. 15, 1980, (an internal document).

9-3 Slobin, S. D., et al., "X-band Atmospheric Noise Temperature Statistics at Goldstone, DSS 13, 1979 and 1980, and Clean Air Noise Temperature Models for Goldstone," *TDA Progress Report 42-64*, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1981.

9-4 Slobin, S. D., et al., "X-band Atmospheric Noise Temperature Data and Statistics at Goldstone DSS 13, 1977-79," *DSN Progress Report 42-52*, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1979.

9-5 Slobin, S. D., et al., "X-band Atmospheric Noise Temperature Statistics at Goldstone DSS 13, 1975-76," *DSN Progress Report 42-38*, Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1977.

9-6 Greenhall, C. A., "Examination of the DSN X-band Weather Specifications," *DSN Progress Report 42-45*, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1978.

9-7 Mulhall, B. D. L., "DSN Telecommunications Interfaces, 64-Meter Antenna Subnet," 810-5, Rev. D, TCI-10, Rev. A, Jet Propulsion Laboratory, Pasadena, Calif., Jan. 15, 1980 (an internal document).

Appendix

Visually Measured Atmospheric Attenuation and Radiometer-Measured System Noise Temperature Increase Statistics

Table A-1. Distributions of X-band zenith cloud attenuation, dB (from [9-2])

Complex	Year-Quarter[a]	Local time[a]	50[b]	70[b]	80[b]	85[b]	90[b]	95[b]
Goldstone, California	DF	00	0.000	0.000	0.012	0.043	0.083	0.145
	DF	06	0.000	0.000	0.020	0.036	0.078	0.166
	DF	12	0.000	0.005	0.048	0.078	0.122	0.189
	DF	18	0.000	0.013	0.036	0.057	0.088	0.164
	MM	00	0.000	0.000	0.000	0.007	0.024	0.059
	MM	06	0.000	0.000	0.012	0.024	0.041	0.072
	MM	12	0.000	0.000	0.022	0.038	0.059	0.093
	MM	18	0.000	0.001	0.027	0.043	0.062	0.089
	JA	00	0.000	0.000	0.000	0.004	0.015	0.057
	JA	06	0.000	0.000	0.000	0.011	0.019	0.071
	JA	12	0.000	0.000	0.014	0.023	0.044	0.086
	JA	18	0.000	0.002	0.019	0.041	0.074	0.124
	SN	00	0.000	0.000	0.000	0.010	0.020	0.029
	SN	06	0.000	0.000	0.014	0.023	0.037	0.067
	SN	12	0.000	0.000	0.011	0.019	0.028	0.061
	SN	18	0.000	0.000	0.012	0.017	0.024	0.042
Canberra, Australia	DF	00	0.000	0.014	0.096	0.163	0.232	0.307
	DF	06	0.000	0.014	0.096	0.163	0.232	0.307
	DF	12	0.000	0.014	0.051	0.089	0.142	0.222
	DF	18	0.000	0.000	0.010	0.020	0.043	0.094
	MM	00	0.000	0.013	0.065	0.104	0.145	0.199
	MM	06	0.000	0.013	0.065	0.104	0.145	0.199
	MM	12	0.000	0.008	0.042	0.076	0.114	0.161
	MM	18	0.000	0.000	0.009	0.022	0.057	0.116
	JA	00	0.000	0.017	0.059	0.098	0.151	0.240
	JA	06	0.000	0.017	0.059	0.098	0.151	0.240
	JA	12	0.000	0.014	0.047	0.082	0.133	0.211
	JA	18	0.000	0.010	0.037	0.073	0.124	0.198
	SN	00	0.000	0.005	0.063	0.140	0.234	0.351
	SN	06	0.000	0.005	0.063	0.140	0.234	0.351
	SN	12	0.000	0.023	0.070	0.114	0.180	0.288
	SN	18	0.000	0.000	0.021	0.046	0.096	0.197

Table A-1 (contd)

Complex	Year-Quarter[a]	Local time[a]	50[b]	70[b]	80[b]	85[b]	90[b]	95[b]
Madrid,	DF	00	0.003	0.019	0.097	0.140	0.189	0.267
Spain	DF	06	0.005	0.042	0.139	0.218	0.321	0.479
	DF	12	0.010	0.064	0.139	0.190	0.265	0.378
	DF	18	0.016	0.053	0.103	0.132	0.166	0.227
	MM	00	0.000	0.025	0.079	0.132	0.217	0.320
	MM	06	0.000	0.021	0.083	0.216	0.380	0.646
	MM	12	0.001	0.079	0.180	0.258	0.357	0.520
	MM	18	0.000	0.031	0.114	0.174	0.261	0.394
	JA	00	0.000	0.000	0.010	0.021	0.055	0.118
	JA	06	0.000	0.001	0.016	0.032	0.070	0.138
	JA	12	0.000	0.001	0.025	0.043	0.088	0.192
	JA	18	0.000	0.006	0.018	0.025	0.039	0.145
	SN	00	0.000	0.002	0.037	0.113	0.181	0.273
	SN	06	0.000	0.014	0.105	0.169	0.249	0.364
	SN	12	0.000	0.022	0.088	0.123	0.166	0.240
	SN	18	0.000	0.021	0.070	0.100	0.138	0.191

[a]DF = Dec.-Feb.
MM = Mar.-May
JA = June-Aug
SN = Sept.-Nov.

00 = 2100-0300 hours
06 = 0300-0900 hours
12 = 0900-1500 hours
18 = 1500-2100 hours

} Local standard time

[b]Probability (percent) that attenuation does not exceed tabulated value.

Table A-2. Averaged year-quarter distributions of zenith cloud attenuation for X-band, dB (from [9-2])

Complex	Year Quarter[a]	50[b]	70[b]	80[b]	85[b]	90[b]	95[b]
Goldstone, California	DF	0.000	0.000	0.029	0.055	0.094	0.168
	MM	0.000	0.000	0.014	0.028	0.049	0.082
	JA	0.000	0.000	0.009	0.017	0.036	0.093
	SN	0.000	0.000	0.010	0.017	0.026	0.046
Canberra, Australia	DF	0.000	0.009	0.043	0.093	0.065	0.262
	MM	0.000	0.006	0.037	0.077	0.120	0.176
	JA	0.000	0.014	0.050	0.088	0.140	0.224
	SN	0.000	0.010	0.050	0.102	0.186	0.305
Madrid, Spain	DF	0.007	0.045	0.116	0.159	0.224	0.331
	MM	0.000	0.033	0.117	0.191	0.292	0.464
	JA	0.000	0.001	0.017	0.029	0.063	0.144
	SN	0.000	0.015	0.077	0.123	0.177	0.266

[a]DF = Dec.-Feb.
MM = Mar.-May
JA = June-Aug.
SN = Sept.-Nov.

[b]Probability (percent) that attenuation does not exceed tabulated value.

Table A-3. Averaged whole-year distribution of zenith cloud attenuation for S- and X-band, dB (from [9-2])

Complex	50[a]	70[a]	80[a]	85[a]	90[a]	95[a]
			X-Band			
Goldstone	0.000	0.000	0.013	0.025	0.048	0.097
Canberra	0.000	0.010	0.045	0.088	0.146	0.241
Madrid	0.000	0.018	0.072	0.122	0.188	0.300
			S-Band			
Goldstone	0.000	0.000	0.000	0.001	0.003	0.006
Canberra	0.000	0.000	0.003	0.007	0.011	0.018
Madrid	0.000	0.001	0.004	0.008	0.014	0.023

[a]Probability (percent) that attenuation does not exceed tabulated value

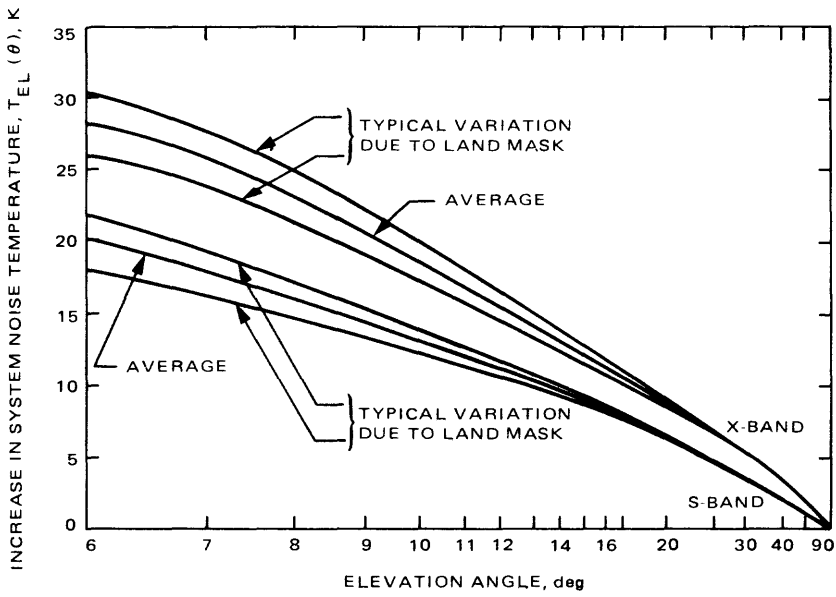


Fig. A-1. Increase in system noise temperature vs elevation angle for 64-meter antennas, S- and X-band (no weather effects) (from [9-2])

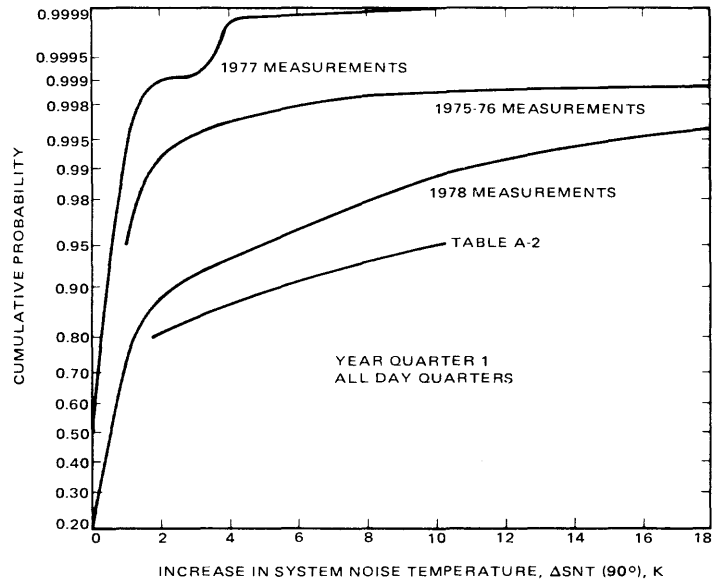


Fig. A-2. Increase in system noise temperature vs elevation angle for 64-meter antennas, S- and X-band (no weather effects) (from [9-2])

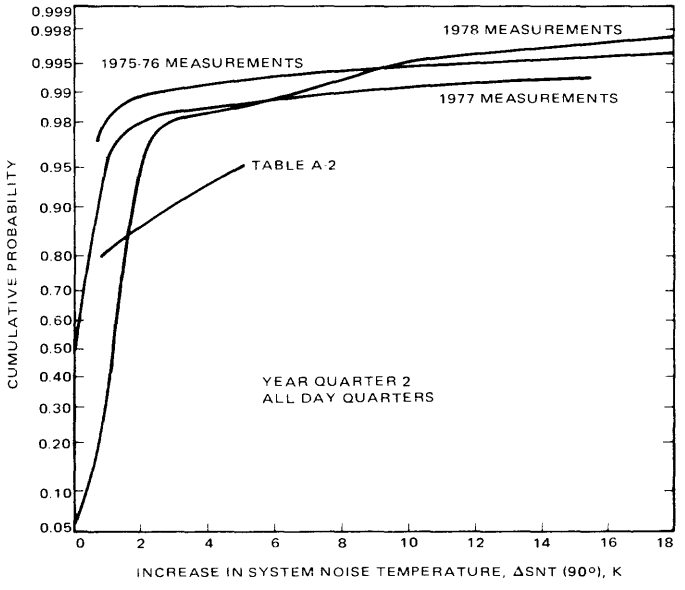


Fig. A-3. Cumulative distribution function for zenith noise temperature increase, Goldstone X-band, March–May (from [9-2])

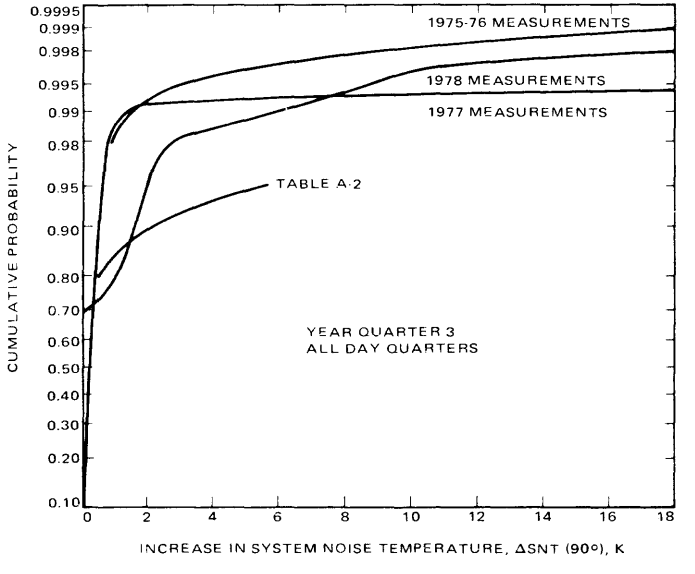


Fig. A-4. Cumulative distribution function for zenith noise temperature increase, Goldstone X-band, June–August (from [9-2])

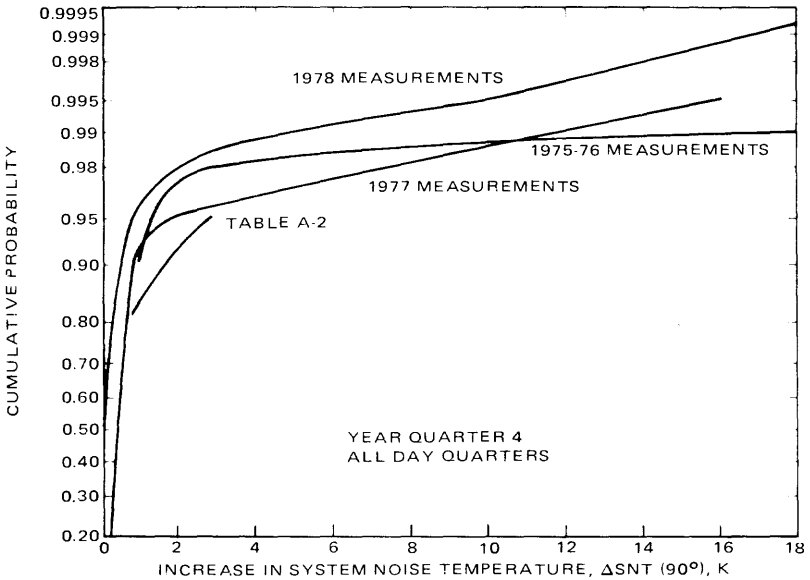


Fig. A-5. Cumulative distribution function for zenith noise temperature increase, Goldstone X-band, September–November (from [9-2])

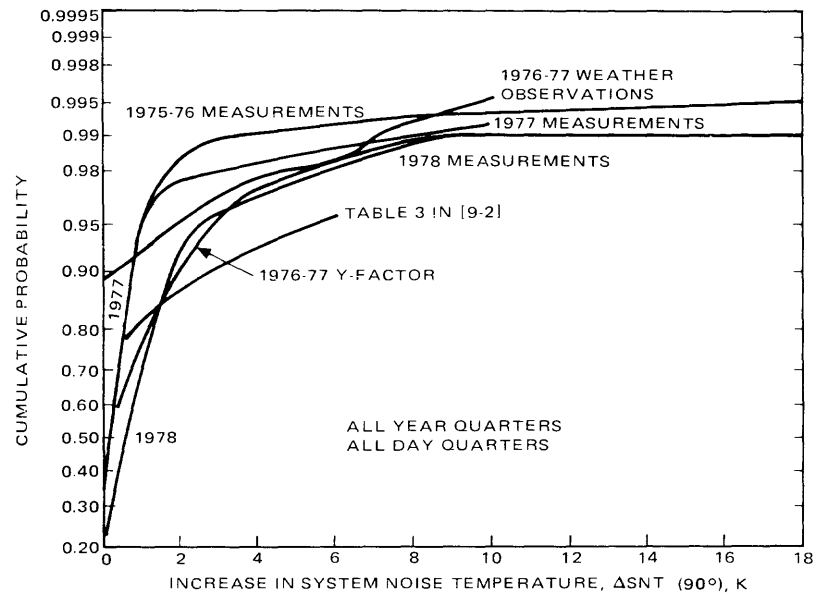


Fig. A-6. Cumulative distribution function for zenith noise temperature increase, Goldstone X-band, all year quarters (from [9-2])

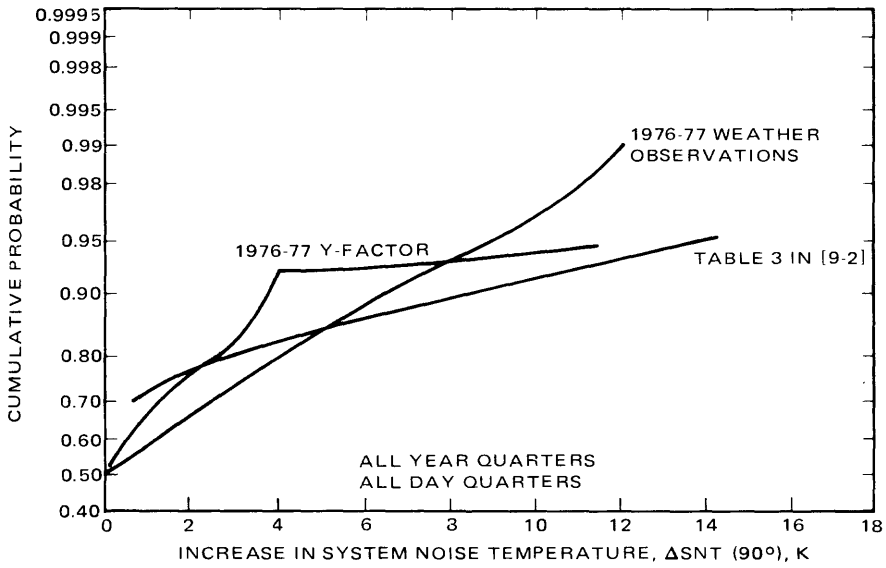


Fig. A-7. Cumulative distribution function for zenith noise temperature increase, Canberra X-band, all year quarters (from [9-2])

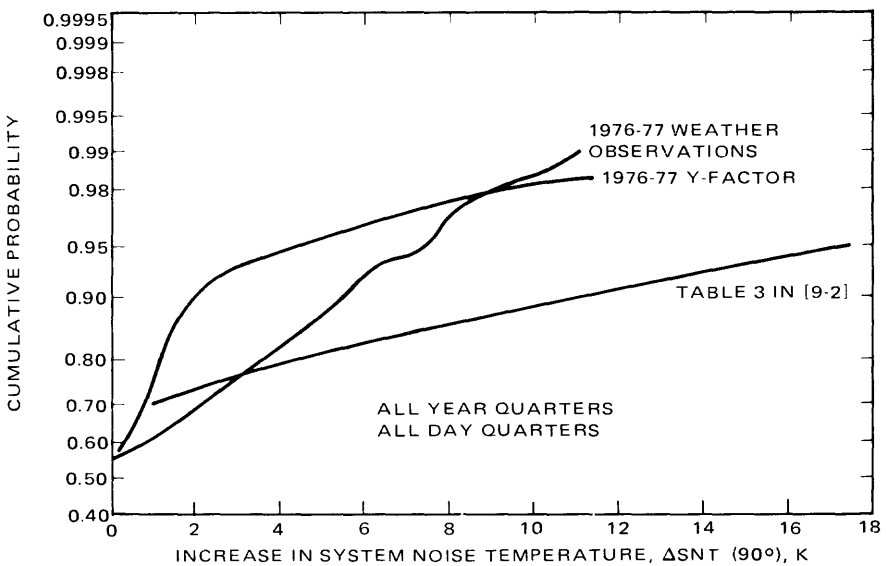


Fig. A-8. Cumulative distribution function for zenith noise temperature increase, Madrid X-band all year quarters (from [9-2])

Chapter 10
Telecommunications Performance Analysis and Monitoring Software

Jim F. H. Taylor

10.1 Introduction

Jet Propulsion Laboratory flight projects and the Deep Space Network require accurate prediction of telecommunications link capability. Link capability determines spacecraft command message error rate, science data telemetry error rate, and radiometric angle, velocity, and position errors during a mission. To establish link performance capability, the analyst commonly translates these error rates into corresponding signal-to-noise ratios or received signal power levels. To validate predicted capability, the analyst compares actual (measured) signal-to-noise ratios or signal levels with previously predicted values for the same link configuration and time.

Prediction of link capability begins in the earliest stages of spacecraft communications system design and spacecraft-ground system tradeoffs. The comparison of actual and predicted data begins with the first use of the radio links after launch and continues until the end of the mission. Computations for both prediction and comparison are repetitive in nature over mission durations that may exceed a decade. Link analysis software provides the needed accuracy, consistent computations, and outputs in standard formats. The software accesses link parameter data bases having traceable updates over the mission duration.

The JPL Telecommunications Systems Section developed the Telecommunications Prediction and Analysis Program for use with Mariner spacecraft in the 1960s. The program is known as TPAP for short, with the word "analysis" meaning the comparison of predicted with actual performance. The program has evolved from project to project to meet specific needs and to accommodate changing data input and output interfaces. Flight projects first provided access to the actual data via computer files in the 1970s. The resulting TPAP link comparison occurred days or weeks after data receipt; that is, in non-real-time. In the same era, the Telecommunications Systems Section added the capability to profile the maximum telemetry data rate available for Viking and to predict the effects of weather on X-band links for Voyager. As link prediction and comparison software requirements have evolved, TPAP has become the Telecommunications Performance Analysis System (TPAS) in the 1980s. TPAS is a joint development between JPL flight projects and the DSN, responding to the needs of both organizations. TPAS capabilities will include access to actual data via an electrical interface for real-time comparison, as well as link prediction, data rate profiling, and non-real-time comparison.

The utilization of TPAP for the Viking 1975 project and the Voyager 1977 project is described in [10-1] through [10-4]. The functional and design requirements for the successor system, TPAS, are given in [10-5] through [10-8]. (The initial version of the real-time comparison portion of TPAS is called "Voyager Telecommunications On-Line Processing Subsystem," or VTOPS.)

10.2 Prediction of Link Capability Using TPAP

This section describes the prediction functions, data flow, and outputs of TPAP as the program exists at the end of 1982. The version described produces predictions for Voyager mission operations and for Galileo preflight spacecraft and mission design. The TPAP comparison and profiling functions are inactive on Voyager. Section 10.4 describes comparison and profiling as they will be implemented in TPAS for post-1984 missions.

The TPAP user may generate any of the following predictions.

(1) Telemetry: downlink carrier and data channel.

(2) Command: uplink carrier and data channel.

(3) Turnaround ranging: uplink carrier, downlink carrier, and data channel.

(4) Trajectory: range, angles, and derived quantities.

10.2.1 Functional Description and Data Flow

TPAP operates in the Univac 1100 computer of the JPL Information Processing Center (IPC). Figure 10-1 is a functional flow diagram of the prediction por-

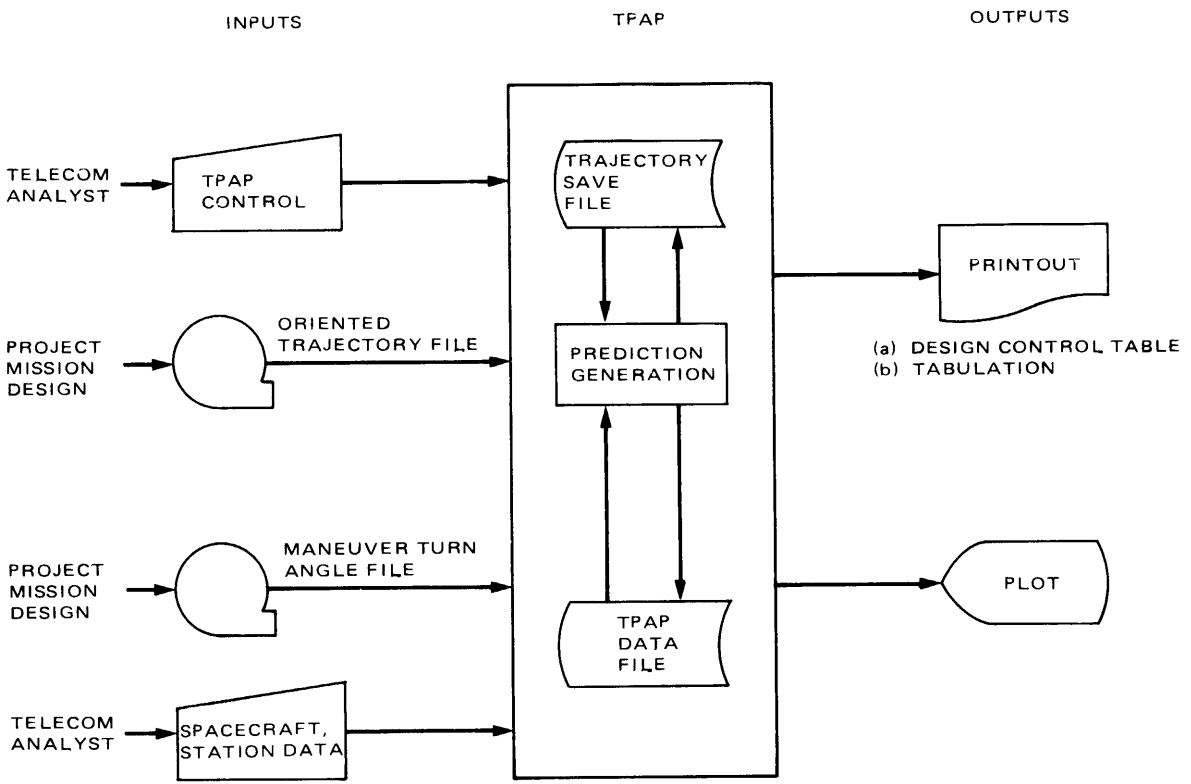


Fig. 10-1. TPAP functional flow diagram for predictions

tion of TPAP, showing its major interfaces with external data and its major internal components. Table 10-1 summarizes the information required to make a prediction run. Card images are commonly input from a demand terminal. Save tapes are in the project-controlled tape library. TPAP data files are permanently catalogued files or tapes; they are project-controlled during mission operations.

Tables 10-2 and 10-3 define the TPAP data file elements used in uplink predictions and downlink predictions, respectively. As suggested by their names, the elements are grouped according to the portion of the link each specifies: uplink carrier, command data channel, downlink carrier, telemetry data channel, or ranging data channel. The tables also show the probability density function type associated with each data element type. Depending on the number of spacecraft and ground station configurations available (for example, S-band transmitter high-power or low-power), TPAP may have one or several individual data elements of each type.

The telecom analyst controls TPAP by a runstream. The runstream consists of a series of card images that specify the type of prediction desired, the trajectory tape identification, the interval over which predictions are desired, the particular set of data elements that define the link configuration and mode, the times that DCTs are output, the particular link quantities that are plotted or tabulated, and formats and scales for the outputs.

10.2.2 Prediction Outputs

Types of link prediction outputs are design control tables (DCTs), plots, and tabulations. The DCTs display configuration and performance at one given time.

Table 10-1. Inputs to prediction portion of TPAP

Input	Typical information included	Source of input
TPAP control	Type of prediction	Card images
	Output quantities and formats	
	Prediction start and stop times	
Trajectory	Range, range rate	Trajectory save tape, maneuver save tape
	Spacecraft cone, clock angles (station-centered or geocentric)	
Spacecraft and station data	Power levels and system losses	TPAP data file or card images
	Spacecraft antenna patterns	
	Antenna gains, pointing losses, and ellipticities	
	Data rates and threshold levels	

Table 10-2. Uplink data element types

Element name	Description	Probability density function
ULCARR (1,1)	Station transmitter power	Triangular
ULCARR (1,2)	Station transmit circuit loss (if not included in ULCARR (1,1))	(not currently used)
ULCARR (1,3)	Station uplink antenna gain	Uniform
ULCARR (1,4)	Station uplink antenna pointing loss	Uniform
ULCARR (1,5)	Uplink polarization loss	Triangular
ULCARR (1,6)	Spacecraft uplink antenna gain	Triangular
ULCARR (1,7)	Spacecraft uplink antenna pointing loss	Triangular
ULCARR (1,8)	Spacecraft receiving circuit loss	Uniform
ULCARR (1,9)	Spacecraft system noise temperature	Gaussian
ULCARR (1,10)	Spacecraft carrier noise bandwidth at threshold	Triangular
ULCARR (1,11)	Uplink ranging carrier suppression	Triangular
ULCARR (1,12)	Uplink command carrier suppression	Triangular
CMDCHN (1,1)	Command channel modulation loss	Triangular
CMDCHN (1,2)	Total uplink system losses (in lieu of CMDCHN (1,3), CMDCHN (1,4))	Triangular
CMDCHN (1,3)	Uplink radio loss	(not currently used)
CMDCHN (1,4)	Uplink demodulator/detector loss	(not currently used)
CMDCHN (1,5)	Uplink threshold S_{TB}/N_0	Triangular
RNGCHN (1,1)	Uplink ranging power/total power	Triangular
RNGCHN (1,2)	Uplink ranging control radio loss	Triangular
RNGCHN (1,3)	Ranging channel noise bandwidth	Triangular

A plot or tabulation shows the variation in performance of one or more link parameters as a function of time.

The TPAP DCT is an orderly listing of the design and tolerance values for all the parameters defining a link, showing how each individual parameter affects the link. For the DCT, TPAP calculates the design value link performance as a linear sum of the design values of the individual parameters, each expressed in decibels. The design value is the expected link performance. The program also calculates an adverse value link performance as a specified multiple (n-sigma) of the statistical standard deviation of the link. The data element probability density functions determine the mean and variance for each parameter from the design value and favorable and adverse tolerances. The link variance is the sum of the individual variances. The DCT threshold parameter values come from

Table 10-3. Downlink data element types

Element name	Description	Probability density function
DLCARR (1,1)	Spacecraft transmitter power	Triangular
DLCARR (1,2)	Spacecraft transmit circuit loss	Uniform
DLCARR (1,3)	Spacecraft downlink antenna circuit loss	Uniform
DLCARR (1,4)	Spacecraft downlink antenna gain	Triangular
DLCARR (1,5)	Spacecraft downlink antenna pointing loss	Triangular
DLCARR (1,6)	Downlink polarization loss	Uniform
DLCARR (1,7)	Station downlink antenna gain	Uniform
DLCARR (1,8)	Station downlink antenna pointing loss	Uniform
DLCARR (1,9)	Station zenith system noise temperature	Gaussian
DLCARR (1,10)	Clear dry noise temperature increase with elevation	Gaussian
DLCARR (1,11)	Station carrier noise bandwidth at threshold	Triangular
DLCARR (1,12)	Downlink carrier suppression due to ranging channel	Triangular
DLCARR (1,13)	Downlink carrier suppression due to telemetry	Triangular
DLCARR (1,14)	Hot body system noise temperature increase	Gaussian
TLMCHN (1,1)	Telemetry channel modulation loss	Triangular
TLMCHN (1,2)	Total downlink system losses (in lieu of TLMCHN (1,3), TLMCHN (1,4), TLMCHN (1,5))	Triangular
TLMCHN (1,3)	Downlink radio loss	Triangular
TLMCHN (1,4)	Downlink demodulator/detector loss	Triangular
TLMCHN (1,5)	Downlink waveform distortion loss	Triangular
TLMCHN (1,6)	Downlink threshold S_{TB}/N_0	Triangular
RNGCHN (1,4)	Downlink ranging channel modulation loss	Triangular
RNGCHN (1,5)	Downlink ranging radio loss	Triangular
RNGCHN (1,6)	Required ranging P_R/N_0	Triangular

theoretical criteria that relate required signal levels to the individual project telecommunication system mission requirements.

For examples of TPAP-produced DCTs for Galileo command, telemetry, and ranging performance, see Tables 1-2, 1-3, and 1-4 in chapter 1.

The TPAP generates the values for plots and tabulations from DCTs generated (but not necessarily output) at the interval specified between plotted or tabulated points. Figures 10-2, 10-3, and 10-4 demonstrate TPAP plot capability.

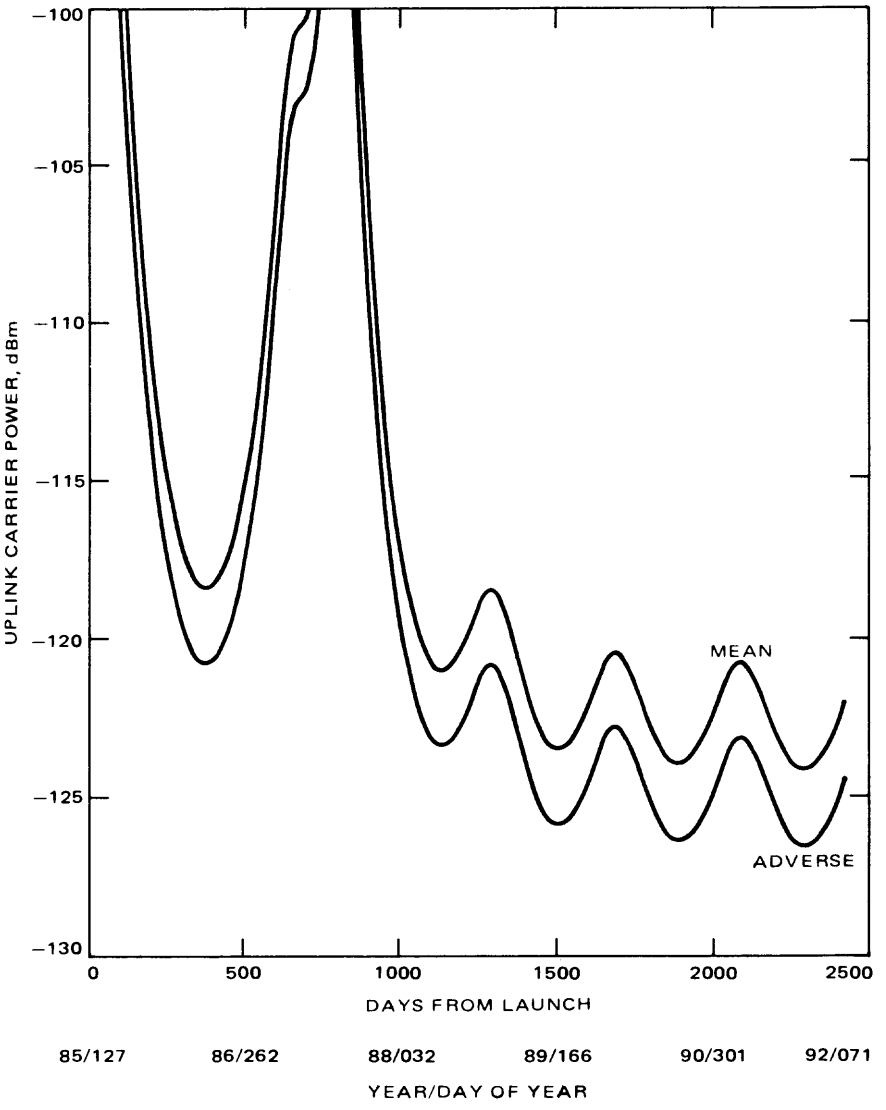


Fig. 10-2. Long-period TPAP prediction output

Figures 10-2 and 10-3 show, respectively, the long-term and short-term prediction of a selected link parameter. The top curve in each plot is the parameter's expected value; the bottom curve is the minus two-sigma value. The long-term plot is for a single spacecraft and ground station configuration, for one station and fixed elevation angle of the station antenna. The short-term plot covers 24 hours for a single configuration, but it shows performance at each station as the elevation angle of the tracked spacecraft changes.

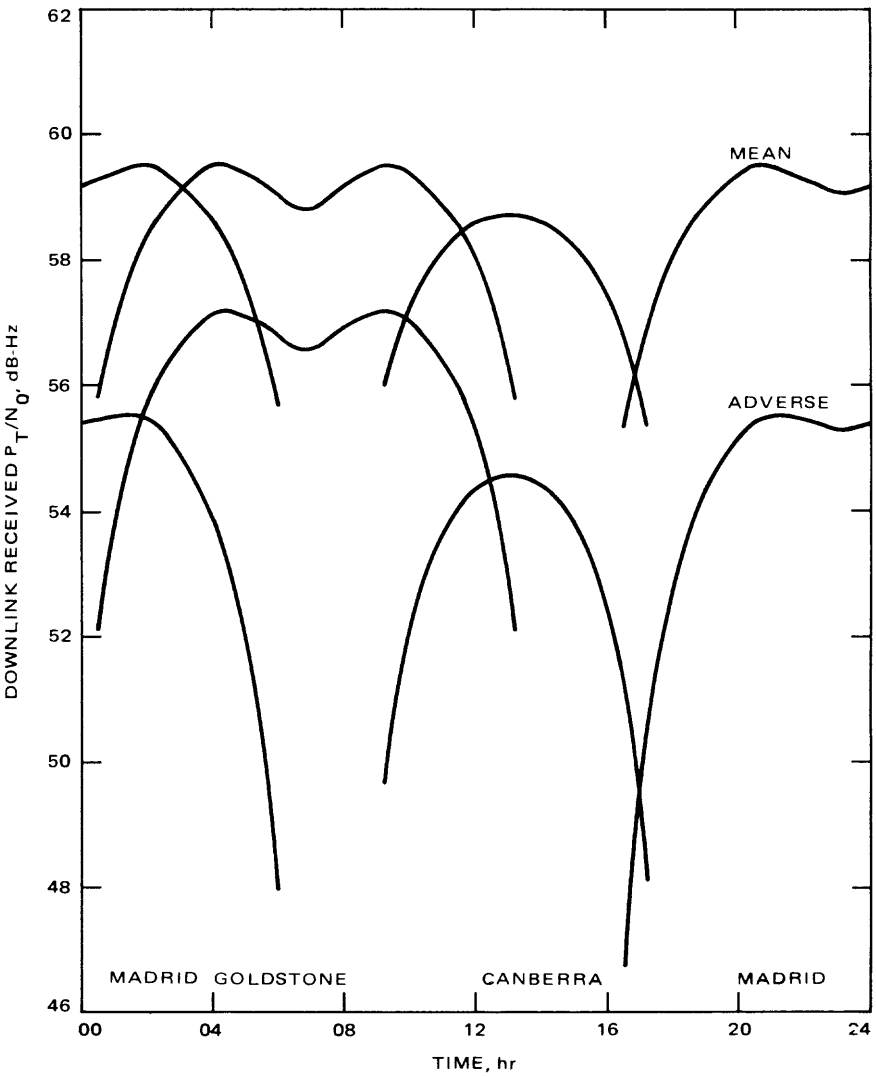


Fig. 10-3. Short-period TPAP prediction output

Figure 10-4 plots a predicted trajectory quantity. "Predicted" means displayed from the trajectory input file. Like other such quantities, the earth-to-spacecraft range can be predicted relative to the center of the earth (geocentric) or relative to a station location. Unlike link parameters, trajectory quantities do not have tolerances associated with them.

Table 10-4 is a TPAP tabulation, showing a selected set of telemetry link parameters and their tolerances over a mission interval.

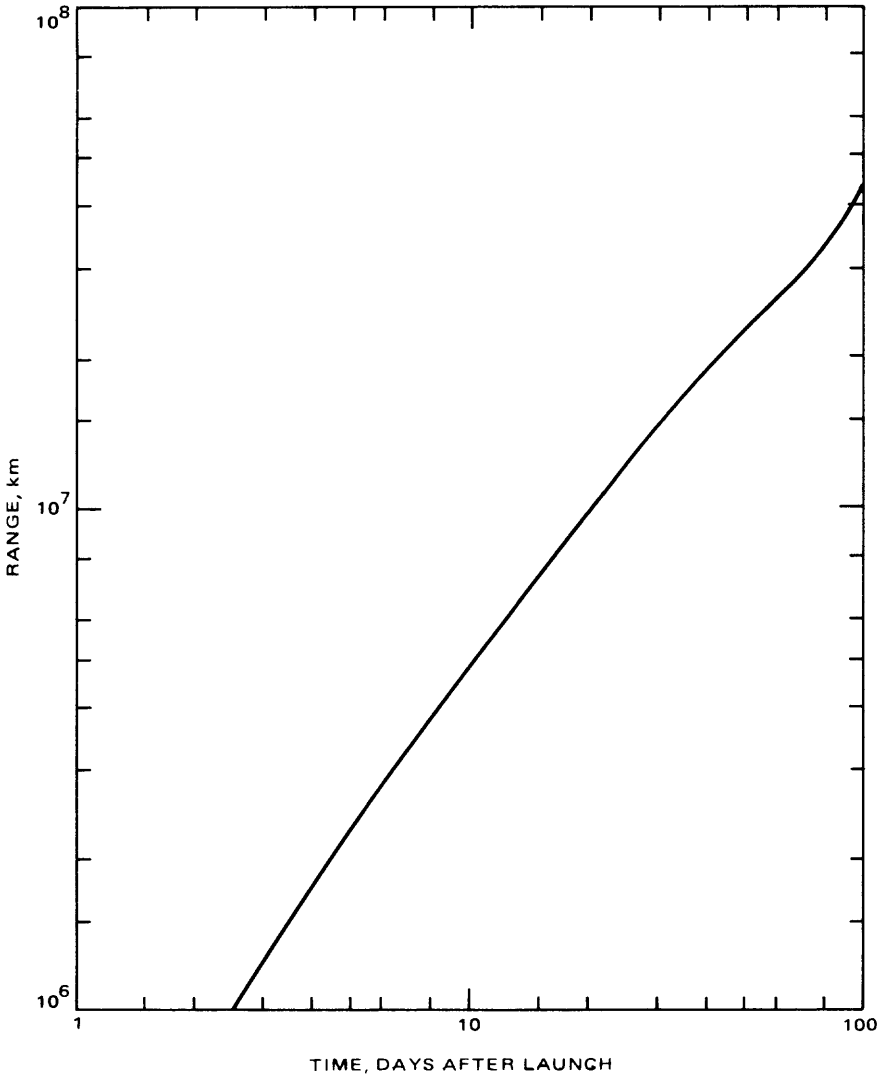


Fig. 10-4. TPAP trajectory prediction output

10.3 Mini-TPAP

Mini-TPAP serves the special needs of early spacecraft and mission design. It facilitates input of trajectory data via card images and the rapid buildup of data base elements. Its outputs are design control tables, plots, and tabulations that look similar to TPAP outputs. Mini-TPAP accommodates few-of-a-kind runs for preliminary design and tradeoffs. TPAP is best for repetitive production runs

Table 10-4. Typical TPAP telemetry performance prediction tabulation

Year-month-day	Days from launch	P_T/N_0		Downlink carrier margin	
85-05-08	1.0	63.96	4.81	45.59	4.83
85-05-15	8.0	37.46	4.55	19.10	4.57
85-05-22	15.0	35.36	3.91	17.00	3.93
85-05-29	22.0	35.24	3.28	16.87	3.31
85-06-05	29.0	35.85	2.64	17.48	2.67
85-06-12	36.0	36.53	2.16	18.17	2.19
85-06-19	43.0	37.09	1.98	18.73	2.02
85-06-26	50.0	38.04	1.92	19.68	1.96
85-07-03	57.0	39.21	1.89	20.85	1.94
85-07-10	64.0	40.18	1.98	21.81	2.02
85-07-17	71.0	41.00	2.15	22.64	2.18
85-07-24	78.0	41.71	2.23	23.34	2.26
85-07-31	85.0	42.14	2.25	23.77	2.29
85-08-07	92.0	42.26	2.24	23.89	2.28
85-08-14	99.0	41.96	2.12	23.59	2.16
85-08-21	106.0	41.08	2.08	22.71	2.12
85-08-28	113.0	39.65	2.16	21.28	2.20
85-09-04	120.0	37.96	2.25	19.59	2.29
85-09-11	127.0	36.32	2.28	17.95	2.32
85-09-18	134.0	34.74	2.25	16.38	2.29
85-09-25	141.0	33.29	2.23	14.93	2.26
85-10-02	148.0	31.94	2.23	13.57	2.26
85-10-09	155.0	30.71	2.23	12.35	2.26
85-10-16	162.0	29.60	2.23	11.23	2.26
85-10-23	169.0	28.59	2.23	10.23	2.26
85-10-30	176.0	27.68	2.23	9.316	2.26
85-11-06	183.0	26.85	2.23	8.488	2.26
85-11-13	190.0	26.11	2.22	7.742	2.26
85-11-20	197.0	25.45	2.23	7.079	2.26
85-11-27	204.0	24.85	2.23	6.488	2.26
85-12-04	211.0	24.32	2.23	5.954	2.26
85-12-11	218.0	23.84	2.23	5.478	2.26
85-12-18	225.0	23.42	2.23	5.057	2.26
85-12-25	232.0	23.05	2.23	4.687	2.26
86-01-01	239.0	22.73	2.23	4.363	2.26
86-01-08	246.0	22.45	2.23	4.084	2.26
86-01-15	253.0	22.21	2.23	3.846	2.26
86-01-22	260.0	22.01	2.23	3.646	2.26
86-01-29	267.0	21.84	2.23	3.479	2.27
86-02-05	274.0	21.70	2.24	3.333	2.28
86-02-12	281.0	21.59	2.25	3.220	2.29
86-02-19	288.0	21.50	2.27	3.137	2.30
86-02-26	295.0	21.45	2.28	3.082	2.31
86-03-05	302.0	21.42	2.29	3.054	2.33
86-03-12	309.0	21.40	2.28	3.029	2.32

using established link configurations specified by runstreams. With a few exceptions, the data base element probability density function types and link equations common to TPAP and mini-TPAP are identical.

In the TPAS era, TPAS will provide capability and ease of operation similar to the present mini-TPAP.

10.4 Prediction, Profiling, and Comparison With TPAS

TPAS provides project and DSN telecom analysts involved with Galileo and subsequent projects the capability to predict, profile, and compare the performance of spacecraft-to-DSN telecommunications links. TPAS prediction capabilities are functionally similar to TPAP capabilities (Section 10.2), with similar outputs. Profiling is the calculation of the maximum telemetry data rate that can be supported as a function of time, based on a planned sequence of link configurations, on a stated degree of link conservatism, and on the underlying predictions. Comparison is the calculation of the difference between a measured (actual) link quantity and its predicted value at a given time, the result being a residual, and the calculation of statistics of the residuals.

TPAS-profile functional requirements originate from the Viking Orbiter TPAP Playback Opportunities Profile (TPBOP) program. TPAS-compare functional requirements originate from the Mariner 1971 TPAP-compare program and the Voyager Telecom On-line Processing System (VTOPS). This section includes typical outputs from Viking TPBOP and Voyager TOPS that show expected TPAS capability. TPAS-profile and TPAS-compare will be available in 1984.

Figure 10-5 is a functional flow diagram for TPAS. The Deep Space Network and the JPL flight projects jointly support the implementation and operation of TPAS. TPAS software resides in the DSN Network Operations Control Center (NOCC) Support Controller (NSC) computer, in the JPL Information Processing Center (IPC) Univac 1100 computer, and in a microprocessor called Automated Office Data Center (AODC). TPAS performs prediction, profiling, off-line comparison and on-line comparison functions.

The DSN uses the TPAS-predict element called NSC-TPAP to make its predictions for long-range planning and for each individual station tracking pass. Flight projects use the TPAS element called U1100-TPAP for prediction and profiling of available telemetry data rates, and for off-line comparison of spacecraft and station data accumulated in a file against the predictions for the same pass. On-line comparison for flight projects takes place in the TPAS element, containing both hardware and software, called the Telecommunications On-line Processing System (TOPS). TOPS contains link verification (LINVER) software

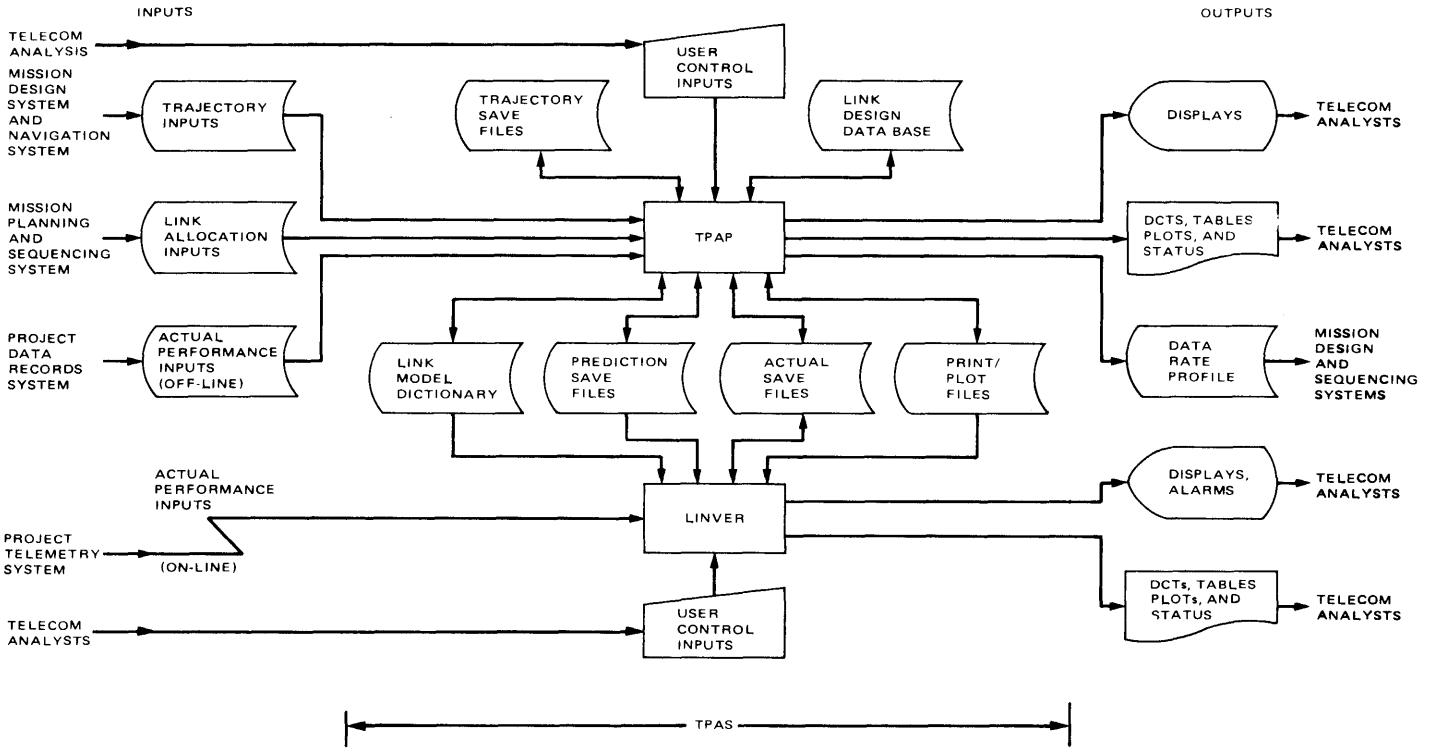


Fig. 10-5. Functional flow diagram for telecommunications performance analysis system (TPAS)

that operates in a particular AODC configuration called the Telecommunications On-line Computer Subsystem (TOCS). On-line comparison requires a direct electrical connection with the project computer supplying the data for comparison in real-time.

TPAS will incorporate a data-base system, residing in the IPC 1100, with the program having a shared link model dictionary interacting with separate link design data bases for each project. All data bases will have a common structure and access method. The link model dictionary contains information describing data entries, processing, and output formats. The link design control document for each project is from the Deep Space Network/Flight Project Interface Design Handbook (a JPL internal document). Each time the U1100 TPAP data base is updated, a subset is transferred to the NSC as part of NSC-TPAP.

10.4.1 Functional Description and Data Flow

Referring to Fig. 10-5, which omits reference to the different computers containing TPAS elements, all of the inputs from external systems go to the TPAP portion except for the on-line actual data. The TPAP inputs include trajectory files, link allocation files, and off-line actual data files. TPAP requires trajectory and link allocation data to generate a profile. The link allocation data defines the Deep Space Stations scheduled for the project, the handover time from one to the next, the spacecraft transmitter power mode, the on/off status of the spacecraft ranging channel, and other information required to define link configuration completely. TPAP requires off-line telecom data, and LINVER requires on-line telecom data to perform a comparison. The telecom data is the actual spacecraft telemetry and the station performance and status monitor data received when the spacecraft is being tracked by the station during mission operations.

The internal TPAS interfaces include the trajectory save file, the prediction save file, and the actual save file. TPAP creates a trajectory save file from the trajectory input, and this trajectory save file is used for subsequent predictions over that time interval. TPAP-predict creates a prediction save file for each of a series of standard configurations, and these prediction save files are used for the generation of subsequent profiles. Either TPAP-compare or LINVER can create an actual save file; it contains the link data accumulated during the comparison, the resulting residuals (a residual is equal to the actual value minus the predicted value of a link parameter), the residual statistics, and "alarm" messages flagging occurrences of the configuration not meeting its standard or the performance exceeding its limit.

Both U1100-TPAP and LINVER produce outputs for use by the project telecom analyst, and NSC-TPAP produces predictions for operations controllers and analysts.

For the project, U1100 TPAP creates print/plot files that LINVER accesses to generate DCTs, tabulations, and plots on the TOCS line printer, digital plotter, or cathode ray tube display. The same outputs may be produced through U1100 devices. TPAP-profile produces a telemetry data rate profile as an U1100 file. This file contains a header with input identification and link configuration data for profile verification. It also contains data files having spacecraft times and the maximum telemetry data rate achievable during each interval. TPAP-compare and LINVER produce printed tabulations and plots displaying the predicted, actual, and residual values of selected link parameters. The program also generates histograms to show the statistical distribution of residuals of selected link parameters over a specified time span. LINVER will enable the implementation of link alarms. The standards and limits triggering the alarms come from the prediction save file. Link parameters in alarm appear on the TOCS cathode ray tube display and go into the actual save file.

NSC-TPAP produces a prediction data file for each pass; the file contains quantities to be transmitted to the Deep Space Stations, used by other software in the NSC, or provided to DSN cathode ray tube displays.

10.4.2 Profile Outputs

TPAP-profile will produce a downlink events and telemetry signal-to-noise ratio plot similar to Fig. 10-6. Key station and spacecraft configuration events from the link allocation file appear at the top as a function of time. The signal-to-noise ratio that results from the standard predictions and these configurations appear on the plot. (Figure 10-6, for Viking Orbiter, comes from TPBOP.) TPAP-profile also will produce a printed profile, corresponding to the data rate profile. Table 10-5 is an example; it was produced by Viking TPBOP. The TPAS printed profile will display events in both spacecraft and station time, using one-way light-time from the prediction save file. It will list events from the link allocation file and their times. Not all such events may cause a data rate change. Available data rates may differ for different types of downlink data, according to project criteria.

10.4.3 Compare Output Plot Formats

TPAS-compare plot formats will show the variation of one or several link parameters or their residuals as a function of time. This section includes some figures as examples of plot formats. These formats, which were useful on pre-TPAS projects, will be the basis for the TPAS formats.

Figure 10-7, from Voyager TOPS, is the fundamental plot for on-line link assessment. This particular plot shows the values of the residuals of the X-band received carrier power as a function of time. A similar format is available for the uplink or the downlink received carrier power and for the telemetry symbol signal-to-noise ratio. Formats are also available that show residuals of two or

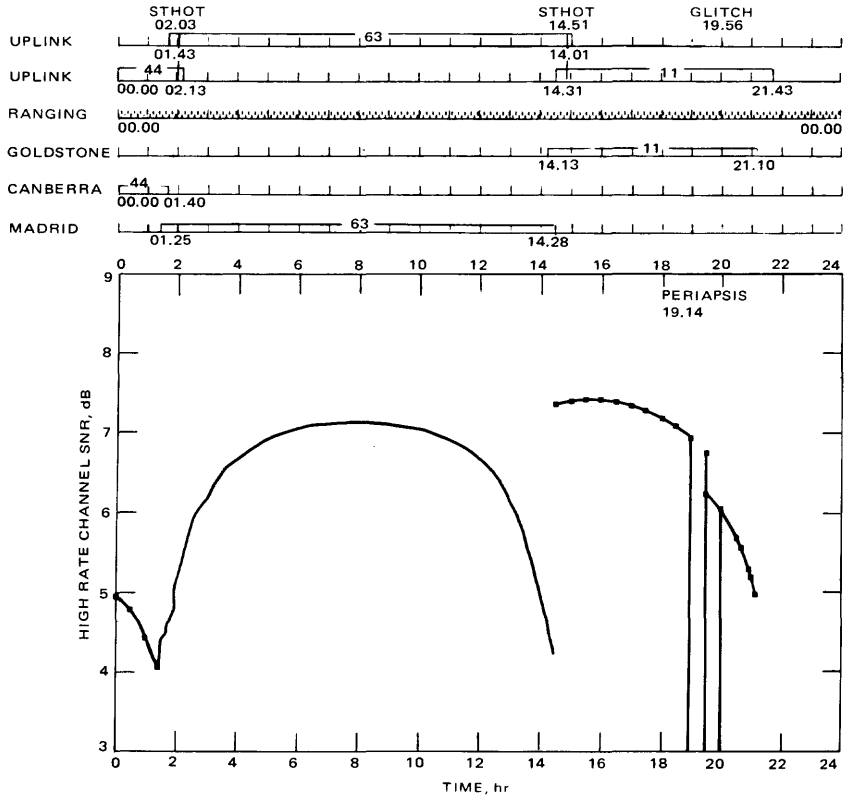


Fig. 10-6. Typical Viking TPBOP downlink events and telemetry SNR plot

three parameters, all on a common time axis, or that show the actual parameters as a function of time.

Figure 10-8, from TPAP-compare for Mariner 1971, is for long-term trend analysis. The plot has the predicted value and the actual data on a common time axis, in this case spanning one month. A similar format, but with the residual as a single curve rather than the predicted and actual as two curves, appears in Fig. 10-9. Figure 10-9, from the Viking TPRESD manual non-real-time compare program, also displays the mean and the standard deviation of the parameter over the plotted interval. Figure 10-10, from TPAP-compare for Mariner 1971, is a histogram for the residual of a link parameter over an interval.

10.5 Real-Time Comparison for Voyager Using TOPS

By late 1981, Voyager telecom analysts began making real-time link comparisons using the Voyager telecommunications on-line processing system (VTOPS).

Table 10-5. Typical Viking TPBOP telemetry data rate printed profile

GMT day/time	Maximum data rates				Comments
	Type 1	Type 2	Type 3	Type 4	
213/00:00					Start of sequence, downlink, DSS 44
					Ranging channel is on
	0K	0K	0K	0K	SNR at 8K = -4.1 dB
01:20					Relay playback end
01:25					Downlink, DSS 44 to DSS 63
	8K	4K	2K		SNR at 8K = 4.1 dB
01:30				4K	Resume relay playback
01:58				0K	Relay playback end
02:03					Uplink handover, DSS 44 to DSS 63
02:06		8K			SNR at 8K = 5.2 dB
02:08				4K	Resume relay playback
02:16	16K				SNR at 8K = 5.5 dB
02:25				8K	SNR at 8K = 5.7 dB
13:31				4K	SNR at 8K = 5.7 dB
13:40	0K				SNR at 8K = 5.5 dB
13:52		4K			SNR at 8K = 5.2 dB
14:23				0K	Relay playback end
14:28					Downlink DSS 63 to DSS 11
	2K	0K	0K		SNR at 8K = -2.1 dB
14:30					SNR shifted 0.5 dB
					SNR fudged by 4.0 dB
14:31			2K		SNR at 8K = -1.6 dB
14:33					Resume relay playback
14:46					Relay playback end
14:51					Uplink handover, DSS 63 to DSS 11
14:56					Resume relay playback
17:53			0K		SNR at 8K = -1.8 dB
18:54					Playback end
					Enter earth occultation (750 km)
	0K				SNR at 8K = 0.0 dB
19:14					Begin periapsis suppression
					Periapsis, orbit number 410
					End periapsis suppression
19:28					Exit earth occultation (750 km)
					Resume playback
	2K				SNR at 8K = -2.3 dB
19:30					SNR shifted 0 dB
19:56					Playback end
	0K				SNR at 8K = 0.0 dB
20:01					Resume playback
	2K				SNR at 8K = -3.0 dB
	0K				SNR at 8K = -3.0 dB
21:10					Playback end
					0 dB, downlink DSS 11 end, no downlink coverage
21:43					Uplink DSS 11 end, one-way start

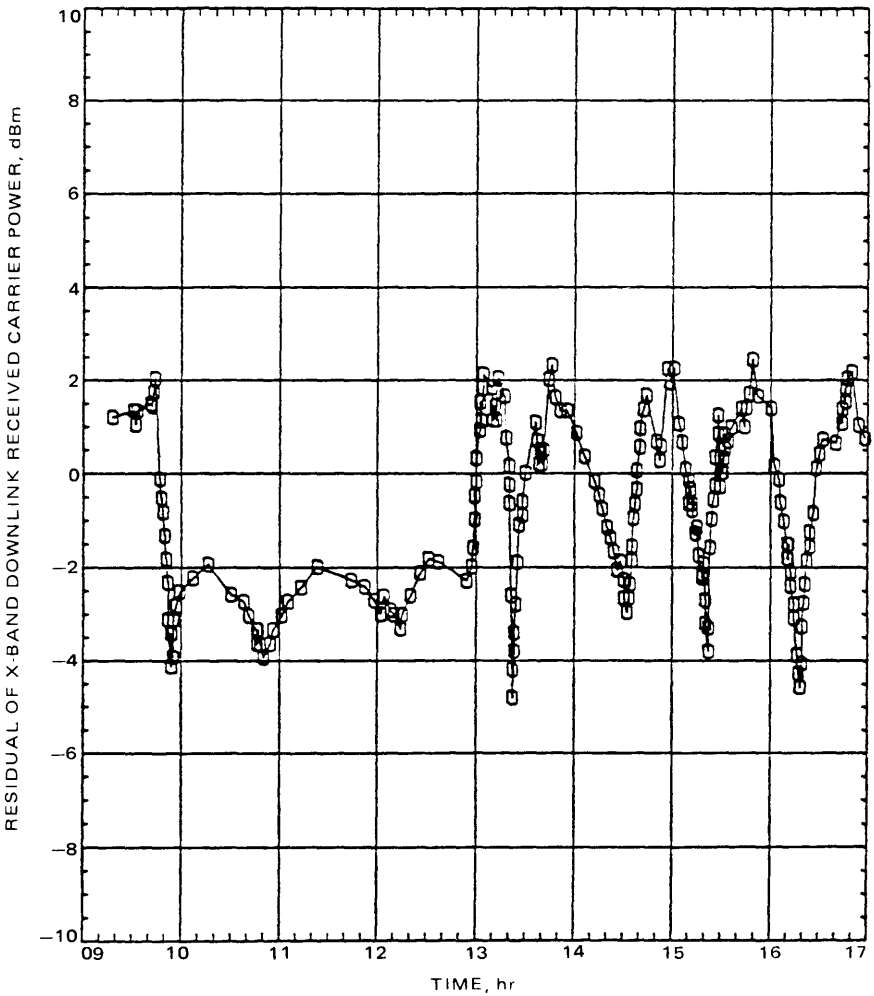


Fig. 10-7. Typical Voyager TOPS real-time residual plot

The VTOPS hardware is a dedicated microcomputer of the JPL automated office data center (AODC) type. The software, VLINVER, will serve as a prototype for TPAS LINVER.

VTOPS, as well as the successor TOPS, is a semiautomatic means for calculating, displaying, plotting, and storing link residuals. Figure 10-11 is a functional flow diagram for TOPS that also shows its interfaces for data and predictions. The predictions come from Voyager TPAP (Section 10.2). Referring also to the project sequence of events (SOE), the telecom analyst enters predictions into VLINVER via the VTOCS keyboard prior to each tracking station pass or

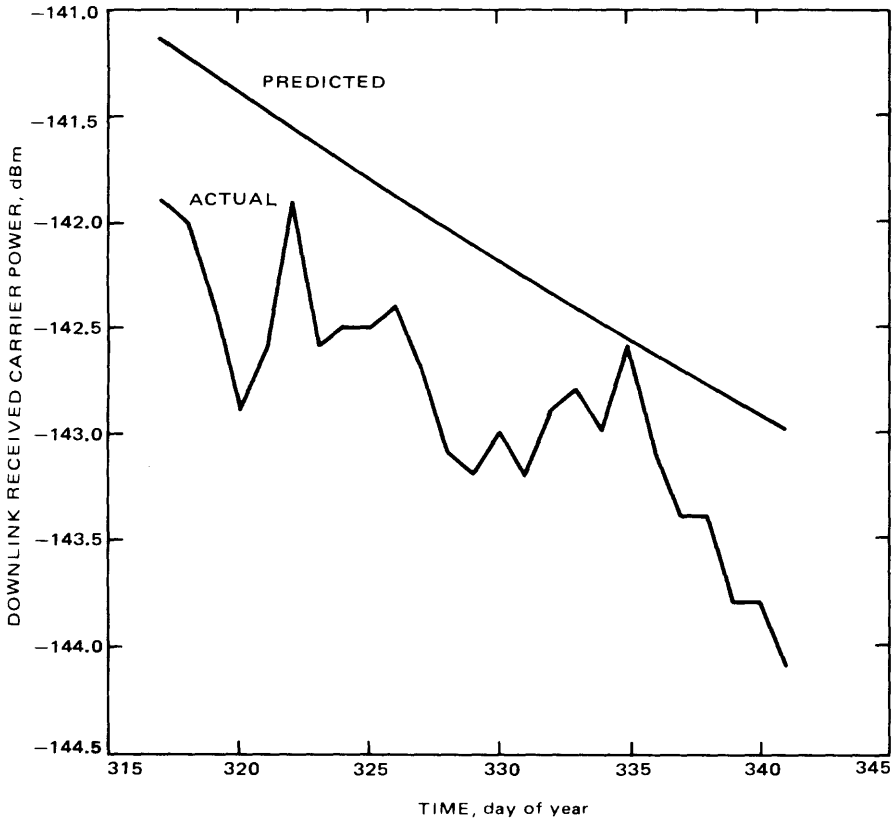


Fig. 10-8. TPAP-compare long-term trend plot, Mariner-Mars 1971 telecommunications prediction analysis

series of passes. The actual data comes from the same on-line electrical interface that drives the Voyager telecom character printer. The actual data, a subset of that used by TPAS-compare, includes uplink carrier signal level, S-band and X-band downlink received carrier power, symbol signal-to-noise ratio for the telemetry data channels, and some spacecraft and station status and identification words.

Voyager LINVER contains a prediction save file from which it can generate predictions as a function of time over a station pass. As each actual data channel updates, LINVER calculates the residuals of the link quantity, using spacecraft and station configuration data from status words entered by the analyst or included in the actual data input. At the end of each pass, it extracts for each quantity the largest positive and largest negative residual value, and it calculates the mean and variance. LINVER outputs a pass summary on the line printer. The summary contains the statistics for each link quantity as well as status and

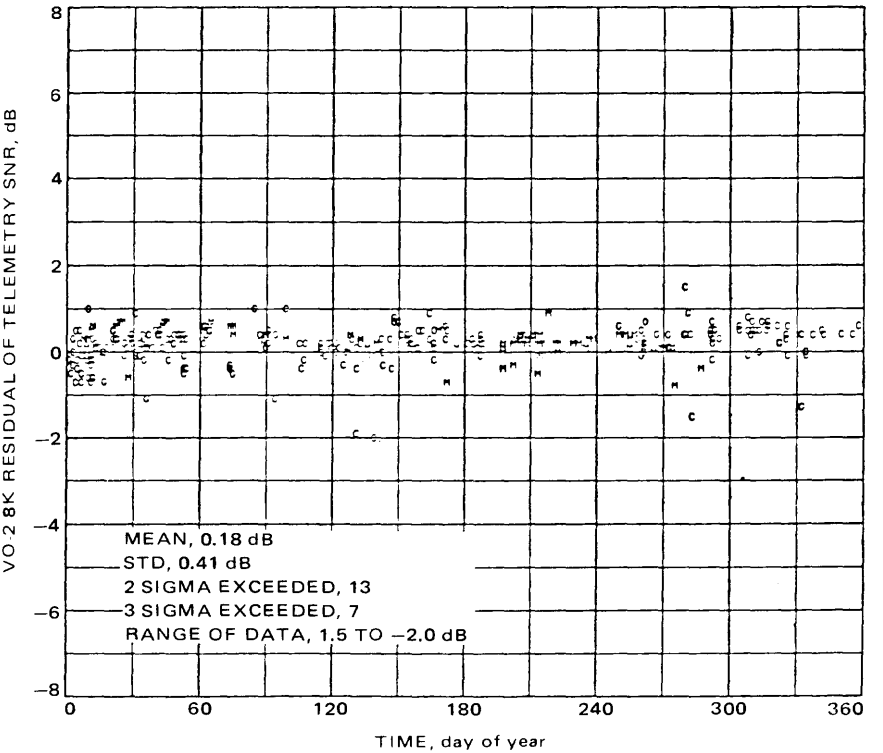


Fig. 10-9. Viking TPRESD residual plot with statistics (actual computer plot)

error messages accumulated during the pass. LINVER outputs the link residuals nearly in real-time as plots versus time on the cathode ray tube terminal. The residual data can be input again to produce tabulations of selected residuals.

10.6 Use of Telecommunications Software (Operations Scenario)

The JPL Telecommunications Systems Section is charged with planning and evaluating telecommunications link performance of JPL flight spacecraft. The Deep Space Network has the responsibility for the performance of the Deep Space Stations and for the transport of the data from the stations to JPL.

Activities carried out by the Telecommunications Systems Section and the DSN in meeting their responsibilities appear in Fig. 10-12. The zero reference time in the 20-year span of the figure is the acquisition of data at a station from an in-flight spacecraft. The TPAS software supports the 14 project activities in the top half of Fig. 10-12. Each activity appears as a time-bar, indicating duration of the activity. Prediction begins years in advance of mission operations;

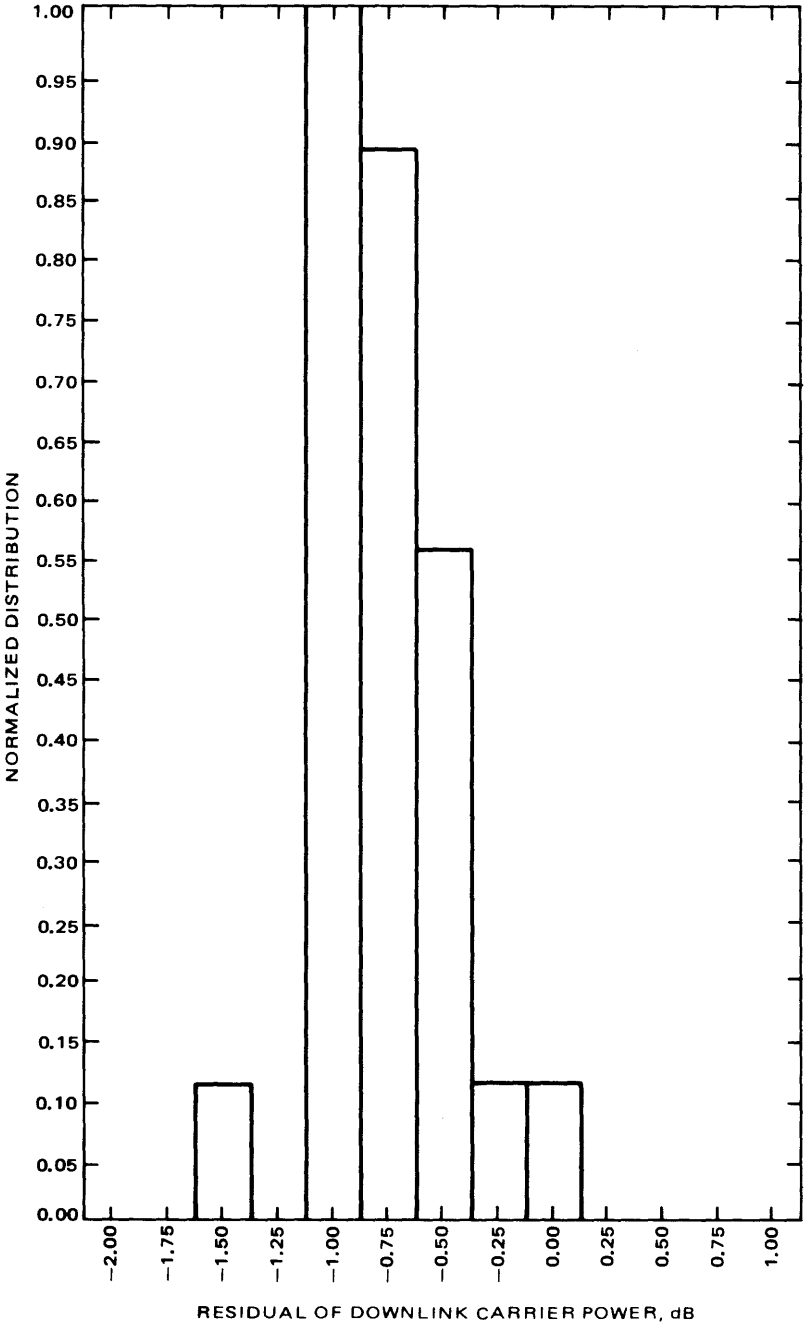


Fig. 10-10. TPAP-compare histogram with statistics

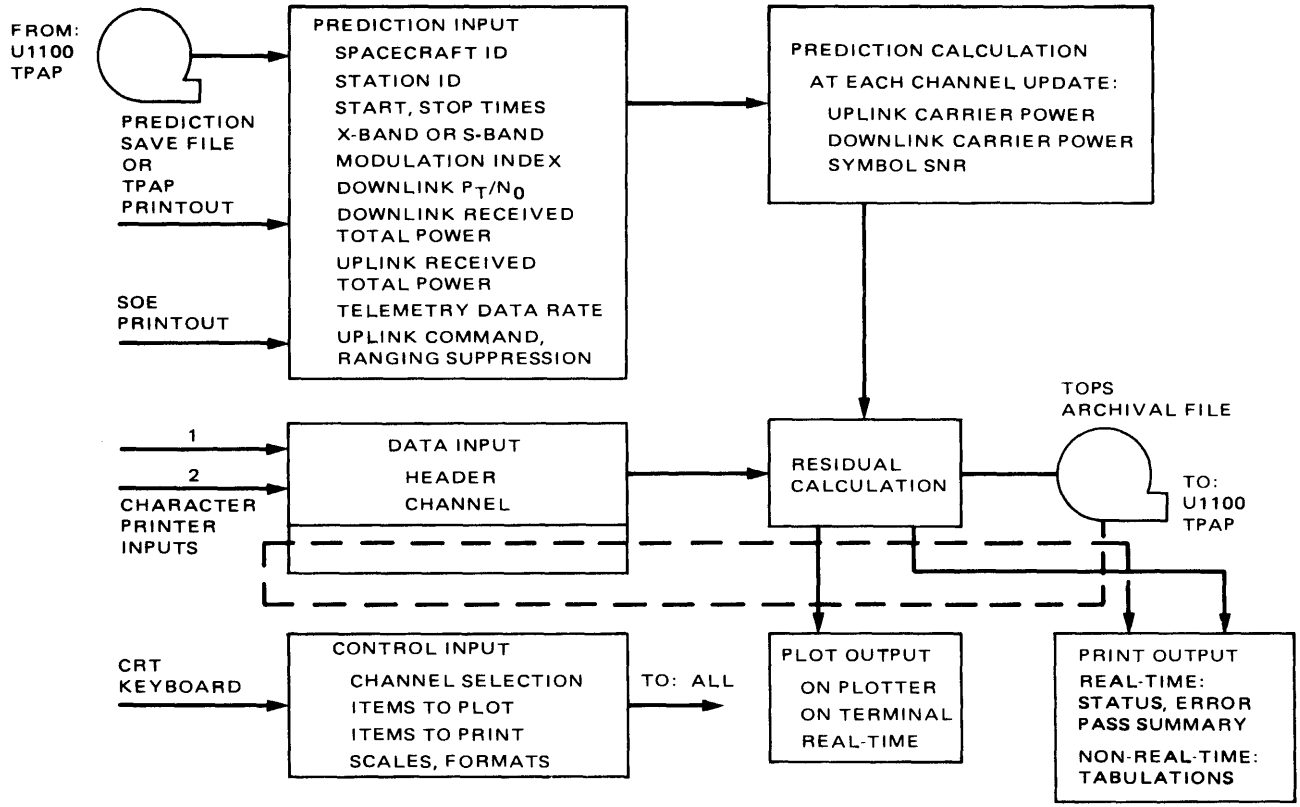


Fig. 10-11. TOPS functional data flow diagram

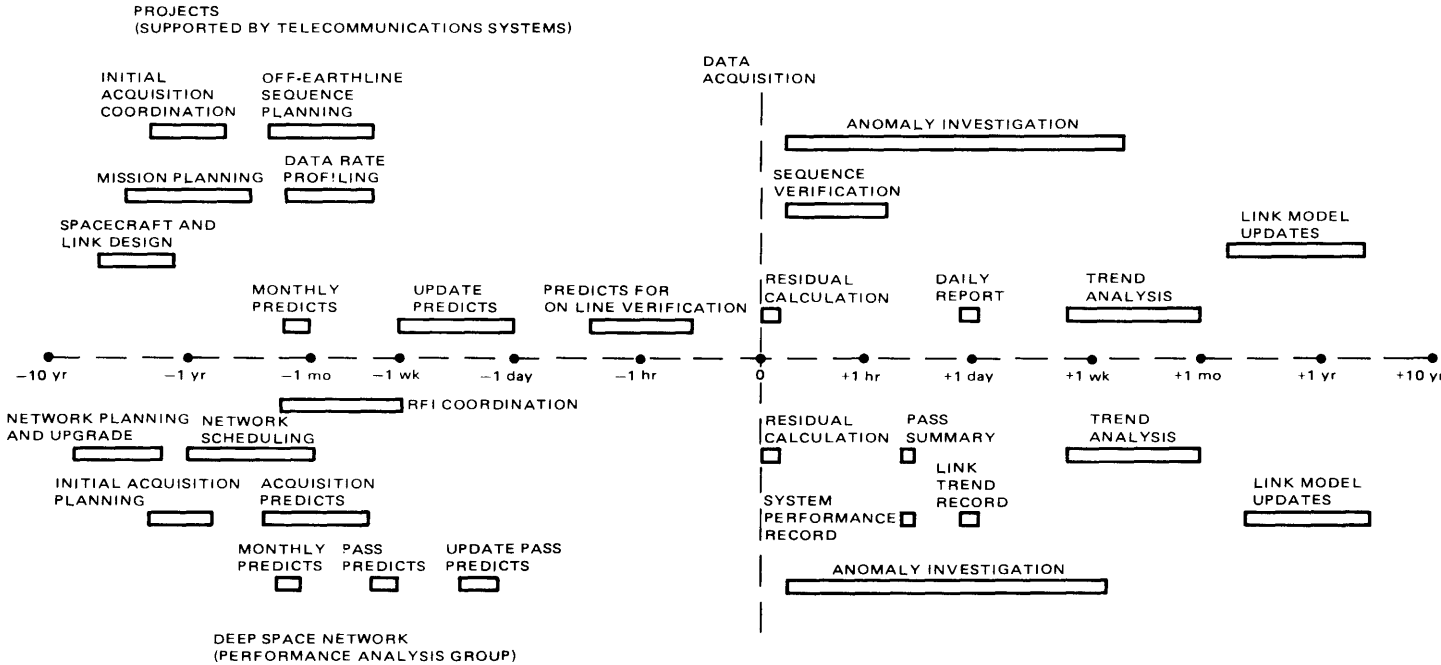


Fig. 10-12. Activity timeline for link prediction, profiling, and comparison

profiling occurs, for most projects, a week to a month before acquisition of data. On-line link comparison takes place within a minute of data acquisition, while off-line comparison for anomaly resolution and trend analysis is largely complete within a month after data acquisition. Each of the 14 project telecom activities involves the use of one or more of the TPAS capabilities. Table 10-6 shows the activities, the portion of TPAS involved, and a summary of the important attributes of each activity.

At a given time on the project side, three or four people are actively using or supporting TPAS. This number assumes a level of flight support similar to 1977, when JPL mission operations teams supported two projects and telecom analysts were working on the preproject definition or prelaunch development of other projects. Of the 14 activities in Table 10-6, most occur during flight operations. In terms of the people using or supporting TPAS, most are involved with flight operations. In units of work months per month during mission operations, TPAS-predict and TPAS-profile require about 0.5, TPAS on-line compare about 1.3, TPAS-predict and TPAS-profile about 0.5, TPAS off-line compare about 1.0, and software and TOCS hardware maintenance about 0.4. TPAS-predict for preproject and prelaunch use is at the level of 0.5 work month per month.

The NSC-TPAP functions appear in the lower left quarter of Fig. 10-12. For the DSN, the initial plan is to operate NSC-TPAP in the batch mode. The DSN functions 24 hours a day, seven days a week. NSC TPAP requires the equivalent of 2 people to maintain the data base (using the U1100), prepare input data files for the program, run the predictions, and verify the outputs for correctness. The input files include ephemeris and spacecraft trajectory data and sequence of event data for each tracking pass. In the late 1980s, the DSN plans to include NSC-TPAP in an end-to-end system that supports all planning functions for the tracking passes, including scheduling, sequence of events generation, ephemeris and trajectory generation, telecommunications link prediction, and generation of standards for station equipment settings and limits for equipment operating parameters. The staffing level for NSC-TPAP itself in that era would reduce to the equivalent of about one-half person, mainly for data base maintenance and output verification.

Table 10-6. Fourteen project telecommunications activities supported by software

1. Spacecraft and link design (TPAP-predict)
   - Create and update data bases
   - Generate standard case predictions
2. Mission planning (TPAP-predict)
   - Predicts for spacecraft design
   - Predicts for MOS design
   - Predicts for mission design
3. Initial acquisition coordination (TPAP-predict)
   - Predict doppler rates, coverage
   - Predict signal level profiles
4. Off-earthline sequence planning (TPAP-predict)
   - Use cone/clock trajectory data
   - Generate signal level predicts with rapid turnaround
5. Data rate profiling (TPAP-profile)
   - Predicts for mission plan
   - Predicts for sequence of events file (SOE) output to project software with rapid turnaround
6. Monthly predicts (TPAP-predict)
   - Predicts for link verification
   - Predicts for hypothesized future situations, project questions
   - Maintained as Predict Save File
7. Update predicts (TPAP-predict, profile)
   - Respond to requests, SOE updates
   - Plot, tab, or file output may be required with rapid turnaround
8. Predicts for on-line verification (TPAP-predict)
   - Generate from Predict Save Files, SOEs, for next several passes
   - Provide as plots, tabulations
9. Residual calculation (LINVER)
   - Highly automated process, *not* labor-intensive, occurs in real-time
   - Produce Actual Save Files
10. Sequence Verification (LINVER)
    - Utilize TPAS (TOPS) alarms
    - Human operations, using TOPS outputs (occurs whenever data are available at input)
11. Anomaly investigation (TPAP-compare)
    - Performed in non-real-time only
    - Reformat, replot capability emphasized
12. Daily report (LINVER)
    - Standardized pageprint format
    - Minimum operation input, rapid turnaround
13. Trend analysis (TPAP-compare)
    - Standard plots for monthly reports
    - Actual Save File, Prediction Save File used
14. Link model update (TPAP-data base)
    - Result of in-flight performance and tests
    - Data transfer to preflight, study projects
    - Refine in-flight prediction accuracy

References

10-1 "Program Description Document for the Viking Telecommunications Prediction and Analysis Program (TPAP)," 620-320, Jet Propulsion Laboratory, Pasadena, Calif. (an internal document).

10-2 "Program Description Document for the Applications Viking Telecommunications Prediction and Analysis Program (AVTPAP)," 620-322, Jet Propulsion Laboratory, Pasadena, Calif. (an internal document).

10-3 Timpe, C., "Software Requirements for the Voyager Telecom Analysis Subsystem Category II Software," memo 3396-76-72, Jet Propulsion Laboratory, Pasadena, Calif., June 25, 1976 (an internal document).

10-4 Timpe, C., "Revision to Software Requirements for Telecom Analysis Subsystem Category II Software," memo 3396-78-19, Jet Propulsion Laboratory, Pasadena, Calif., Mar. 7, 1978 (an internal document).

10-5 "Functional Requirement: Telecommunications Performance Analysis System (TPAS)," 900-952, Jet Propulsion Laboratory, Pasadena, Calif. (an internal document).

10-6 "Design Requirement: Project Voyager Telecommunications On-Line Processing Subsystem," DM-512796, Jet Propulsion Laboratory, Pasadena, Calif. (an internal document).

10-7 "Software Requirements for the Telecommunications Prediction and Analysis System (TPAP)," 901-30, Jet Propulsion Laboratory, Pasadena, Calif. (an internal document).

10-8 *Galileo Mission Operations System Requirements*, module MOS-GLL-5-212.2, Telecommunications Analysis Subsystem, 625-500, Jet Propulsion Laboratory, Pasadena, Calif. (an internal document).

Chapter 11

Radio Frequency Selection and Interference Prevention

Norman F. de Groot

11.1 Introduction

Successful radio communication between spacecraft and earth stations depends upon the use of appropriate radio frequencies and upon having sufficient freedom from interference. Radio frequencies for deep-space research may be chosen from bands that are allocated for that purpose. The choice of particular frequencies within those bands is guided mostly by the need to avoid interference between deep-space missions. Interference may also result from terrestrial sources making transmissions not related to space research.

In this chapter we discuss the bands available for deep-space communications, and the choice of particular mission frequencies. The more general susceptibility of deep-space earth stations to various kinds of interference is then presented. An associated topic is the development of protection criteria that specify maximum allowable levels of interference.

Next, the prediction of interference from near-earth satellites is described, with particular emphasis on the problems and uncertainties of such predictions. The chapter concludes with a brief description of other activities aimed at the prevention or avoidance of interference to deep-space radio communications.

11.2 International Allocation of Radio-Frequency Bands

Use of the radio frequency spectrum is regulated by international treaty. The International Telecommunications Union (ITU), a special agency of the United Nations, is the body responsible for determining the regulations [11-1]. Over 150 nations are members of the ITU, which has its headquarters in Geneva, Switzerland.

The radio regulatory action of the ITU is based on work conducted by its International Radio Consultative Committee (CCIR). As a result of the formal process of adopting technical papers presented at international meetings, the CCIR quadrennially publishes a set of official reports and recommendations. These deal with technical issues concerning the use and regulation of the radio frequency spectrum by all radio services.

CCIR includes 11 study groups that treat the various radio services. Study Group 2 is concerned with space research and radio astronomy. The technical basis for the allocation and use of bands for deep-space research is a part of the work of CCIR Study Group 2. Reports and recommendations concerning technical requirements, preferred frequencies, band sharing, and interference protection criteria relating to deep-space communications may be found in [11-2].

11.3 Bands Allocated to Deep-Space Research

Allocation of radio frequency bands is defined in terms of radio services. One of these services is space research. Deep-space communications must take place in bands allocated for space research. In recognition of the special characteristics of deep-space communications, e.g., very weak received signals and very high-power earth stations, all or part of some space research bands have been restricted to deep space only. In all cases, the bands allocated to deep-space research are shared bands. This means that the bands are also used by other specified services. The characteristics of these other services have been shown to be generally compatible with deep-space links from the standpoint of interference. Continuing analysis of potential interference in connection with increased usage of the shared bands is necessary for the protection of deep-space communications. Case-by-case analysis and resolution of actual interference problems is also required.

The bands that may be used for and that are restricted to deep-space communications in the space research service are listed in Table 11-1. Excerpts of the more complete language contained in the international Radio Regulations can be found in Appendix A.

Table 11-1. Band allocations for deep-space research

Band	Direction	Class[a]
2110–2120 MHz	Earth to space	Primary
2290–2300 MHz	Space to earth	Primary
7145–7190 MHz	Earth to space	Primary
8400–8450 MHz	Space to earth	Primary
12.75–13.25 GHz	Space to earth	Secondary
16.6–17.1 GHz	Earth to space	Secondary
31.8–32.3 GHz	Space to earth	Primary
34.2–34.7 GHz	Earth to space	Primary

[a]Primary means co-equal rights with other primary users.

Secondary means permission to operate with co-equal rights with respect to other secondary users, but without protection from interference *from* primary users, and without co-equal rights with respect to interference caused *to* primary users.

11.4 Band Selection

The selection of a band or bands to be used for a particular mission is based on several considerations:

(1) Communication reliability

(2) Communication performance

(3) Propagation effects on radio navigation and radio science experiments

(4) Existing or planned DSN capability

(5) Equipment technology

In practice, the choice of bands is limited at present (circa 1982) to the 2-GHz uplink-downlink pair and the 8-GHz downlink. This limitation is imposed by existing capability of the DSN and the NASA standard transponder for deep-space applications. A 7-GHz uplink capability is being developed. Application and technology studies may be expected to lead to future use of the newer 12-, 17-, 32-, and 34-GHz allocations.

11.5 Channel Plans and Coherence

To provide for orderly selection and assignment of frequencies for deep-space missions, channel plans have been developed within the United States. The plans were based on bandwidth, hardware implementation, and frequency ratio considerations. All U.S. deep-space missions utilize frequencies included in the channel plans.

For spacecraft navigation and some types of scientific measurement it is necessary that uplink (earth-to space) and downlink (space-to-earth) transmissions be phase coherent. This means that the uplink frequency and phase received by the spacecraft transponder must be translated by a fixed ratio and used to control the downlink frequency and phase from the spacecraft. The requirement for coherence applies between the uplink and downlink frequencies of a band pair. The requirement can also apply to simultaneous transmissions in more than one band pair.

Of the four band pairs listed in Table 11-1, channel plans have been developed for the first two [11-3][1]. Table 11-2 shows the associated frequency ratios.

In the list of channels, Table 11-3, Column 3, we see that there are 27 channel center frequencies within the 2290-2300-MHz downlink band. Channel frequencies in the other bands are a necessary result of the frequency ratios. A particular channel number – for example, channel 17 – specifies frequencies in all four bands.

Because of the spacing between allocations and the frequency ratios embodied in the channel plans, some channels in each band may not be usable from the standpoint of coherence. When complete coherence between the two uplink and two downlink bands is required, the frequencies selected can only be chosen from channels 5-27.

Channel planning for the 12-, 17-, 32-, and 34-GHz allocation has not yet been accomplished.

Table 11-2. Channel frequency ratios

Band pair	Channel frequency ratio
2110–2120 MHz uplink	221
2290–2300 MHz downlink	240
7145–7190 MHz uplink	749
8400–8450 MHz downlink	880
2290–2300 MHz downlink	3
8400–8450 MHz downlink	11

[1] The 7-GHz uplink channels have been planned but are not shown in TRK-20, Rev. B of [11-3].

Table 11-3. Channel center frequencies

Channel	2110–2120-MHz uplink channel center frequency, MHz	2290–2300-MHz downlink channel center frequency, MHz	7145–7190-MHz uplink channel center frequency, MHz	8400–8450-MHz downlink channel center frequency, MHz	Remarks
1		2290.185185	7147.286265		
2		2290.555556	7148.442131		
3		2290.925926	7149.597994		
4		2291.296296	7150.753857		
5	2110.143056	2291.666667	7151.909724	8402.777780	
6	2110.584105	2292.037037	7153.065587	8404.135803	
7	2110.925154	2292.407407	7154.221450	8405.493826	
8	2111.266204	2292.777778	7155.377316	8406.851853	Channels
9	2111.607253	2293.148148	7156.533179	8408.209876	
10	2111.948303	2293.518519	7157.689045	8409.567903	5–27
11	2112.289352	2293.888889	7158.844908	8410.925927	
12	2112.630401	2294.259259	7160.000771	8412.283950	are
13	2112.971451	2294.629630	7161.156637	8413.641977	
14	2113.312500	2295.000000	7162.312500	8415.000000	fully
15	2113.653549	2295.370370	7163.468363	8416.358023	
16	2113.994599	2295.740741	7164.624229	8417.716050	coherent
17	2114.335648	2296.111111	7165.780092	8419.074073	
18	2114.676697	2296.481481	7166.935955	8420.432097	in
19	2115.017747	2296.851852	7168.091821	8421.790124	
20	2115.358796	2297.222222	7169.247684	8423.148147	all
21	2114.699846	2297.592593	7170.403550	8424.506174	
22	2116.040895	2297.962963	7171.559413	8425.864197	four
23	2116.381944	2298.333333	7172.715276	8427.222220	
24	2116.722994	2298.703704	7173.871143	8428.580248	bands
25	2117.064043	2299.074074	7175.027006	8429.938271	
26	2117.405092	2299.444444	7176.182868	8431.296294	
27	2117.746142	2299.814815	7177.338735	8432.654321	
28	2118.087171		7178.494597	8434.012344	
29	2118.428241		7179.650464	8435.370371	
30	2118.769290		7180.814838	8436.738395	
31	2119.110339		7181.962190	8438.086418	
32	2119.451389		7183.118056	8439.444445	
33	2119.792438		7184.273919	8440.802468	
34			7185.429783	8442.160493	
35			7186.585617	8443.518517	
36			7187.741511	8444.876542	
37			7188.897375	8446.234566	
38				8447.592591	
39				8448.950616	

11.6 Channel Selection

For each new deep-space mission it is necessary to select the radio frequencies that will be used for uplink and downlink communication. The selection has the objective of avoiding or minimizing the possibility of radio interference between existing and planned missions. The selection of mission channels within a band is not influenced by link performance. This is because the frequency-dependent parameters of link performance, e.g., antenna gain, space loss, etc., do not vary significantly over the relatively narrow width of band allocations.

11.6.1 Channel Selection Process

The process of channel selection is based on calculation and analysis of interference-to-signal power ratios (ISR) as a function of time for each possible pair of missions [11-4]. The initial calculation assumes that both spacecraft in each pair are using the same channel.

Co-channel operation is often possible because of the very narrow beams of earth-station antennas and the diverse position and motions of spacecraft engaged in deep-space missions.

The worst case ISR is compared to a criterion of acceptable interference. If the ISR meets this criterion for all spacecraft, any channel may be selected for the new spacecraft. This is true because the co-channel condition was used for the calculation.

If the criterion of acceptable interference is not met for the co-channel case, alternatives must be examined. A separate, unused channel may be required. Another possibility is that interference may be acceptable at certain times during each mission.

The ISR calculation is made for each of the two spacecraft in a pair. First, one is assumed to be the desired spacecraft and the other is considered a potential source of interference. The calculation is then repeated for the opposite situation.

11.6.2 Modes of Interference

The radio frequency signal for deep-space communications normally includes a carrier and one or more sets of data sidebands. Examples of two sets of data sidebands are combined telemetry and ranging signals on a downlink, or combined command and ranging signals on the uplink. When separate data streams are carried on individual subcarriers, there are additional intermodulation products, compared to the single-data-stream case.

Given the necessary frequency relationships, the signal components to or from two spacecraft may interfere in one or more of the following ways:

(1) Carrier-to-carrier

(2) Data-to-carrier

(3) Intermodulation product-to-carrier

(4) Carrier-to-data

(5) Data-to-data

(6) Intermodulation product-to-data

These interference modes can occur between uplinks or between downlinks. For the uplink case, there is an additional interference mode. Consider the situation where one uplink signal is being transmitted and intended for a particular spacecraft. If this signal is received by another spacecraft and has sufficient strength and the necessary frequency components, it is possible for the receiver in the unintended spacecraft to lock to the uplink signal. This must be avoided if independent operation of the two spacecraft is to be maintained. This interference mode is called one-way uplink interference.

In practice, the carrier and the telemetry sidebands are usually the most susceptible to interference between downlinks. The predominant mode for uplinks is the one-way uplink interference.

11.6.3 Interference Protection Ratio

The acceptable ratio of interference to signal power is called the interference protection ratio. The required ratio is considered to be -20 dB during critical mission phases and -15 dB at other times. The ratio expresses the maximum interference power with respect to the power of the desired signal. The powers in the ratio are those of the spectral components being considered for a particular interference mode.

A -15 dB ratio will produce a usually negligible effect on carrier tracking performance [11-5, 11-6], a 0.4 dB degradation of telemetry performance, and a 1.0 dB degradation of command performance, assuming that both the telemetry and command are operating at a symbol error rate of 10^{-5}. The -20 dB test provides greater protection for critical times and events.

These protection ratios may seem particularly conservative. They are justified by the fact that the frequency selection study is usually performed in a very early phase of a mission when uncertainties exist about many parameters that

can affect the interference situation. Additionally, it is a goal of channel selection to provide the safest environment so that mission planners will have maximum flexibility.

The maximum acceptable uplink one-way interference is called the uplink one-way threshold. It is considered to be equal to the spacecraft receiver sensitivity. This sensitivity is usually in the range from -155 dBm to -165 dBm.

11.6.4 Interference Analysis

There are four steps in the interference analysis and channel selection process:

(1) Determine power levels of the interference and the desired signal

(2) Determine the likely modes of interference, e.g., carrier-to-carrier interference, carrier-to-data interference, etc.

(3) Determine the time and duration of potential interference

(4) Select one or more channels to avoid potential interference during critical mission events and to minimize potential interference at other times

To achieve these steps, it is necessary to have the following information for all the missions involved:

(1) Characteristics of the telecommunication system

(2) Expected modes of operation as a function of time

(3) Dates of important mission events

(4) Orbital elements that specify the mission trajectory

A complete list and description of the above items are provided in Appendix B.

A computer program has been developed to assist in the calculations needed for interference analysis. This program examines two missions at a time. It accepts as its input a set of orbital elements that completely specify the trajectories of the missions being examined. Based on these orbital elements, it computes for both missions the spacecraft-to-earth range, the doppler rate, and the angular separation between the two spacecraft. From these data and an assumed fixed e.i.r.p., it then computes the ratio of the total received signal power to the total received interference power (TSIR)[2], and also the uplink interference power level as received by the spacecraft. The calculations are made for selected

[2]The program calculates TSIR; the protection ratio is usually specified inversely, that is, interference-to-signal, ISR.

intervals of time during the period of operation that is common for a particular pair of missions.

11.6.4.1 Downlink interference analysis. The earth-station receiving antenna is assumed to be pointing at the desired spacecraft and to be receiving data from it (Fig. 11-1). The signal from the other spacecraft is treated as potential interference. The downlink TSIR is calculated by the computer program for both missions, assuming an equal e.i.r.p. The equations for the computation are:

$$\mathrm{TSIR}_1 = G_{MAX} - G(\theta) - 20 \log_{10} (R_1/R_2)$$

$$\mathrm{TSIR}_2 = G_{MAX} - G(\theta) + 20 \log_{10} (R_1/R_2)$$

where G_{MAX} is the gain of the receiving antenna in dBi, R is the spacecraft-to-earth range, $G(\theta)$ is the gain of the receiving antenna in the direction of θ, and the subscripts 1 and 2 refer to the two missions, with 1 being arbitrarily assigned to one mission and 2 to the other. The off-axis antenna gain, $G(\theta)$, is modeled by the following expressions:

$$G(\theta) = G_{MAX} - (G_{MAX} - 32) \cdot \theta \quad \text{dB} \qquad \text{for } 0^\circ \leqslant \theta < 1^\circ,$$

$$= 32 - 25 \log_{10} (\theta) \ \text{dB} \qquad \text{for } 1^\circ \leqslant \theta < 48^\circ$$

$$= -10 \ \text{dB} \qquad \text{for } 48^\circ \leqslant \theta.$$

The next step is to determine the potential modes of interference based on the spectra of the desired signal and the interference.

The interference modes can be determined by finding out which part of the unwanted signal can spectrally interfere with which part of the wanted signal, i.e., carrier-to-carrier interference, data-to-carrier interference, etc. An unwanted signal can spectrally interfere with a wanted signal if they are close in frequency. Since Doppler-shift can move the two spectra closer to or farther apart from each other, it may be necessary to consider the Doppler effects in determining the modes of interference, particularly when the Doppler rate is significant.

The computer program calculates the total interference power. The analyst must consider the amount of interference power that applies to a particular interference mode. The effective interference power of an in-band CW interference is simply equal to the power of the interference. The effective interference power of an interference having a dense spectrum is equal to the power of the interference reduced by a factor equal to the bandwidth conversion factor.

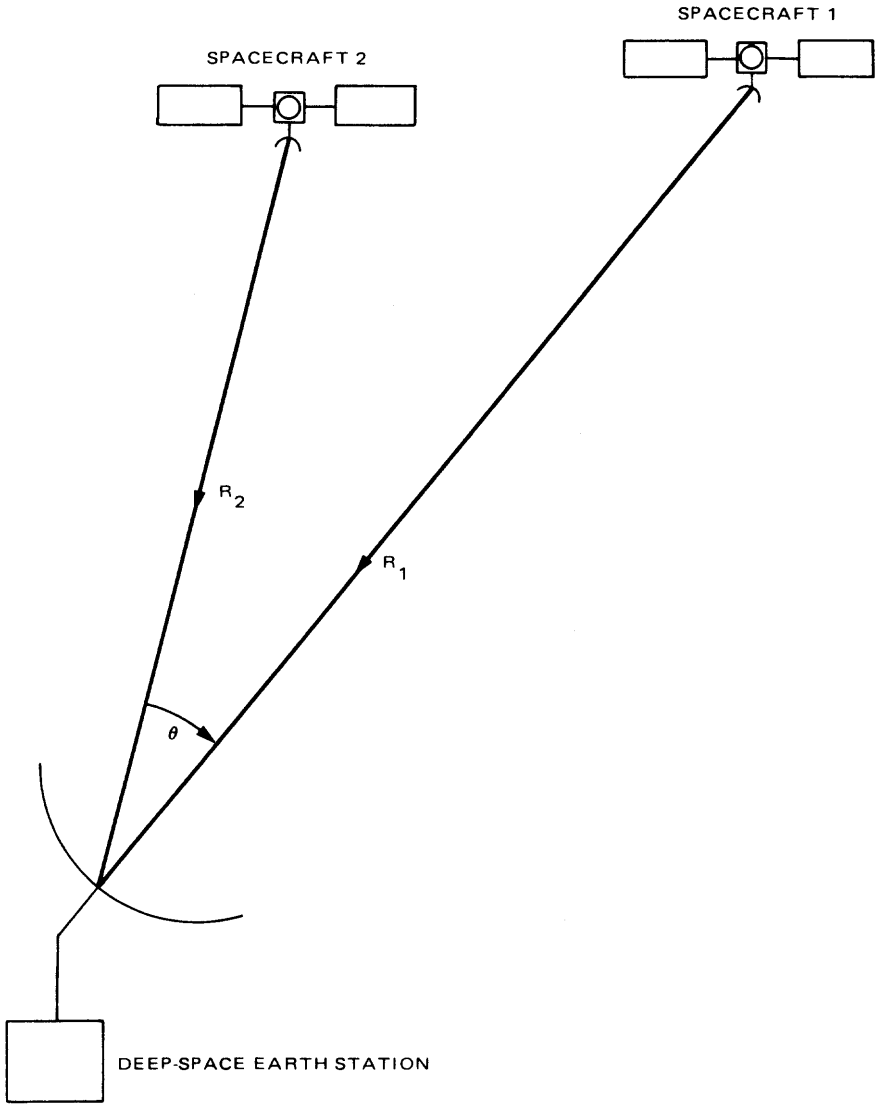


Fig. 11-1. Downlink interference situation with spacecraft 2 as the interferer

The bandwidth conversion factor is defined as the ratio of the bandwidth of the interfered channel to the bandwidth of the interference. The maximum value of the bandwidth conversion factor is unity.

Having determined the interference modes, it is then necessary to calculate the effective interference-to-signal ratio (EISR). This ratio can be derived from the total signal-to-interference ratio as follows:

$$\text{EISR} = P_W - P_I - \text{TSIR}$$

where P_W denotes the power of the wanted signal being interfered with and P_I the effective power of the interference. For example, if the interference mode is carrier-to-data, then the effective interference power is the carrier power of the unwanted signal and the signal power is the power in the data sidebands of the wanted signal.

The effective interference-to-signal ratio can then be compared to the protection ratio. If the protection ratio is exceeded, interference exists. The total number of days for which interference exists is calculated and used as a measure of the amount of potential interference between a given mission pair. This process is repeated for all mission pairs of interest.

A detailed step-by-step analysis of the downlink potential interference for a hypothetical system is provided in Appendix C.

11.6.4.2 Uplink interference analysis. There are two situations in which an unwanted uplink signal constitutes an interference to a spacecraft. The first one is when the effective interference-to-signal ratio exceeds the protection ratio. The second one is when the level of an uplink signal as received by a spacecraft exceeds the receiver threshold of that spacecraft for which the uplink signal is not intended. For the first situation, it is assumed that the two earth stations have the same antenna gain and pattern, and the same transmitter power. For spacecraft in deep space, this assumption yields

$$\text{TSIR} = G_{MAX} - G(\theta)$$

for both missions.

To evaluate the amount of potential interference for the second situation, the power of an unintended uplink as received by a spacecraft located in a direction θ degrees from the main axis of the DSN transmitting antenna is calculated:

$$PU_1 = PT + G(\theta) + 20 \log_{10} \left(\frac{\lambda}{4\pi} \right) - 20 \log_{10} (R_1),$$

and

$$PU_2 = PT + G(\theta) + 20 \log_{10} \left(\frac{\lambda}{4\pi} \right) - 20 \log_{10} (R_2)$$

where PU is the power of the unintended signal, PT is the DSN transmitter output power, $G(\theta)$ is the gain of the transmitting antenna in the direction of θ, λ is the wavelength of the uplink signal, and R and the subscripts 1 and 2 are as defined before. The interference power as calculated by the above expressions assumes an isotropic receiving antenna. It is therefore necessary to increase the interference power by an amount equal to the receiving antenna gain before it can be compared to the uplink one-way threshold to determine the amount of potential interference.

11.6.5 Additional Selection Considerations

Deep-space missions usually have periods of intense activity separated by longer periods of relatively low activity. Interference during the quiet periods will often be acceptable. The potential interference may be strong and virtually destroy successful telecommunications, or it may only cause a slight degradation of performance. The severity of interference therefore affects the choice of channel.

Often, more than one frequency band is used on a mission. The potential interference is usually different for different frequency bands. It may be necessary to examine the potential interference for all frequency bands involved before a selection can be made. Since these frequencies are related to each other by a fixed translation ratio, the resulting channels may be optimal in one band, but not the other.

11.6.6 Channel Availability

It has been possible, so far, to make channel selections such that significant interference between deep-space missions does not occur. This is because of the narrow beamwidth of the DSN earth-station antennas and the characteristics of mission trajectories. Common usage of a single channel is often possible. As other countries and agencies begin to make use of deep-space bands, there may be increasing difficulty in finding channels free of interference for the total duration of each mission. It will then be necessary to accept some interference during noncritical times.

11.7 Frequency Assignment

Following the recommendation of channels for a new U.S. mission, formal application for frequency assignment is made by the JPL Frequency Manager.

In the United States, use of the radio frequency spectrum is managed by two separate agencies. The Federal Communications Commission (FCC) manages usage for all non-Federal civil purposes, including those of state and local governments. The National Telecommunications and Information Administration (NTIA) of the Department of Commerce manages spectrum usage by the Federal Government [11-7]. NASA deep-space missions are considered Federal users.

In order to assure electromagnetic compatibility with existing and planned telecommunications systems, NTIA requires a four-stage review of proposed new systems. The various stages are related to the cycle of project development and procurement. These activities may not begin without favorable action at the corresponding part of the review process. In practice, JPL prepares two submissions for review. NASA completes the needed interaction with NTIA.

For the first submission, the JPL Frequency Manager prepares a system review package. The package includes a brief mission description, technical characteristics of the planned communication links and equipment, and the results of the channel selection study. The review package is sent to NASA and then on to the Spectrum Planning Subcommittee (SPS) of NTIA.

Assuming favorable review, SPS will recommend that NTIA grant spectrum support and that the Frequency Assignment Subcommittee (FAS) of NTIA act favorably on an application for assignment when such is received. Spectrum support means that NTIA agrees with the proposed usage of the requested band for the purposes of the new mission, and will eventually act to assign specific frequencies.

The initial approval by SPS will include a recommendation that Advance Publication Information on Planned Satellite Systems be sent to the International Frequency Registration Board (IFRB), a part of the ITU in Geneva. A draft of this Advance Publication Information document is included in the first system review package. Upon receipt of the approved document by the IFRB, it is translated into the three working languages (French, English, and Spanish) and mailed to member countries.

The purpose of the Advance Publication Information on Planned Satellite Systems is to inform these countries of the proposed frequency assignment, and to invite comments. The comment period of ninety days allows any country to initiate discussions of compatibility.

For the second submission by JPL, the Frequency Manager will prepare a new review package, including up-to-date technical data, response to questions raised by SPS, and a draft formal Notification of Frequency Assignment. This new package is sent to NASA and on to SPS.

With favorable review, SPS will recommend that NASA apply for a frequency assignment, that FAS grant the requested assignment, and that a Notification of Frequency Assignment be sent to IFRB.

Ultimately, the mission and frequencies are published in the Master International Frequency Register. The Register is a large set of catalogs and supplements that include several hundred thousand frequency assignments by countries around the world. Inclusion of new assignments in the Register is the result of IFRB procedures aimed at ensuring compliance with applicable regulations and technical standards. To enjoy the protection offered by the regulations, a new station or system must be registered.

11.8 Susceptibility to Ratio Frequency Interference

In connection with channel selection, Section 11.6 of this chapter described the analysis of potential interference between missions. The ISR criteria used to identify the possibility of downlink interference are based upon the characteristics of the receiving systems employed at deep-space earth stations. The study of their interference susceptibility also applies to the more general problem of RFI from other terrestrial radio services. These problems arise because the bands used for deep space communication are shared with other users. This shared use creates the potential for interference that must be analyzed and managed.

The susceptibility of deep-space earth stations and spacecraft to various kinds of interference is the subject of continuing analysis and test. Mathematical models of receiver circuits and processes are developed. These models are then verified by laboratory testing in which the effects of various kinds of interference are measured. An overview of the development of susceptibility models can be found in [11-8].

The verified models may then be used to predict interference that could arise from known existing and potential sources. An additional use is to provide information to other agencies that may be concerned with or affected by the DSN. For example, under an agreement between NASA and the Department of Defense, JPL provides susceptibility data to the Electromagnetic Compatibility Analysis Center. That organization is involved in frequency and interference management for the military services.

11.8.1 Interference Types

The spectral characteristic of a radio signal depends on the kind of modulation. For assessment of potential interference it is convenient to consider three classes of spectra: CW, noise-like, and pulse. CW (continuous wave) interference is characterized by a single spectral line such as an unmodulated carrier. Noise-like interference refers to complex spectra whose width is greater than the bandwidth of the victim receiver – that is, spectra that affect the victim receiver in a manner similar to noise. Pulse interference generally has a spectrum containing a number of relatively distinct spectral lines. Many systems have spectral characteristics that do not fall neatly within one of the three classes. A frequency-hopping pulse radar is an example.

11.8.2 Interference Effects

Radio frequency interference (RFI) can affect deep-space communications in several different ways, depending upon the strength and nature of the interference. Interference that falls within the tracking bandwidth of a phase-locked loop can induce phase jitter and degraded performance. Still stronger interference can cause the loop to drop lock and, in some cases, track the interfering signal. These effects may occur in one or more of the several functional loops typical of receivers used for deep-space communications: carrier tracking, subcarrier demodulation, and symbol synchronization.

Another class of interference is that where the interfering signal is outside the various loop bandwidths but is of sufficient strength to result in saturation of one or more receiver stages. The effect of saturation is to cause gain compression and corresponding nonlinearity in the amplification and detection of a desired signal to which the receiver is tuned. Nonlinearity also leads to the generation of harmonics, spurious signals, and intermodulation products that may cause interference.

Several aspects of these interference effects have been studied and reported in [11-2], [11-5], [11-6], and [11-9 through 11-12]. The following sections of this chapter present information taken from those reports. This is done to acquaint the reader with a general idea of the CW interference susceptibility of the Block IV receiving system used in the DSN.

11.8.3 Susceptibility to Saturation

CW signals of sufficient strength can saturate the MASER preamplifier or subsequent stages of the earth-station receiver. The particular stage or stages that may be affected depend upon the frequency difference between the interfering signal and the frequency to which the receiver is tuned. The amount of interference power that causes a particular degree of saturation depends upon the strength of the desired signal, since that signal determines the gain of several

receiver stages via the automatic gain control loop. For the purpose of developing a curve of maximum allowable interference power, a threshold (weakest) level is assumed for the desired signal.

Figure 11-2 gives a curve that meets the criterion of not more than 1 dB of gain compression due to saturation. In a somewhat arbitrary manner, 1 dB of compression has come to be considered an acceptable value.

11.8.4 Carrier Tracking Susceptibility to CW Interference

The effect of CW interference on the carrier tracking function of a DSN Block IV receiver has been analyzed and modeled [11-6]. Figure 11-3 illustrates the analytic relationship between the interference-to-signal ratio that results in carrier-loop drop lock as a function of the frequency difference between the desired carrier signal and the interfering CW signal. Curves are given for three values of margins, i.e., the ratio of carrier power to noise power in the tracking-loop bandwidth. Figure 11-4 shows tracking-loop phase jitter as a function of interference-to-signal ratio for several values of frequency offset.

Combining these and similar analytic relationships with laboratory test data results in a curve of maximum allowable CW interference ratios and levels. An

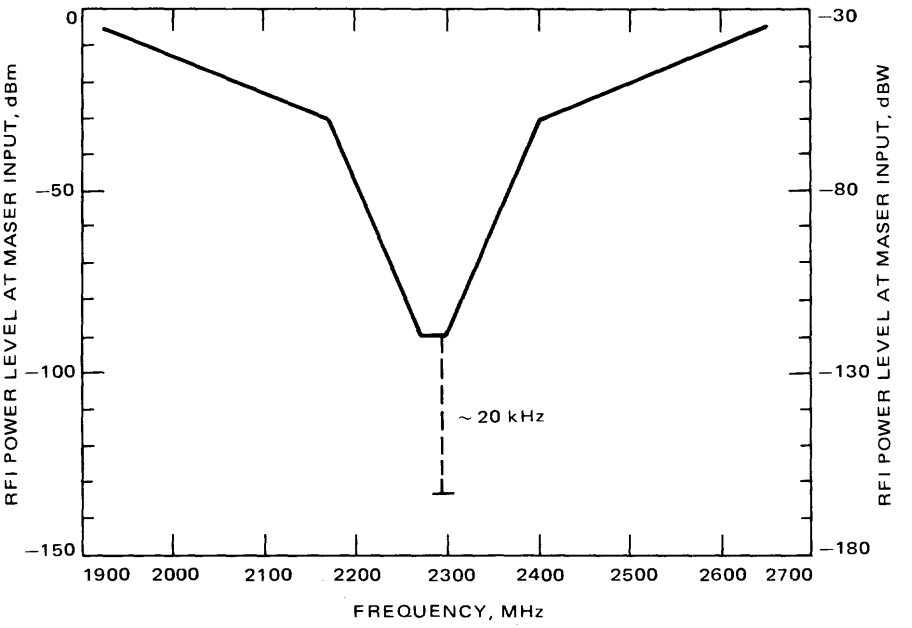


Fig. 11-2. Maximum allowable CW RFI power into S-band maser for 1 dB or less gain compression of maser or Block IV receiver operating at channel 14

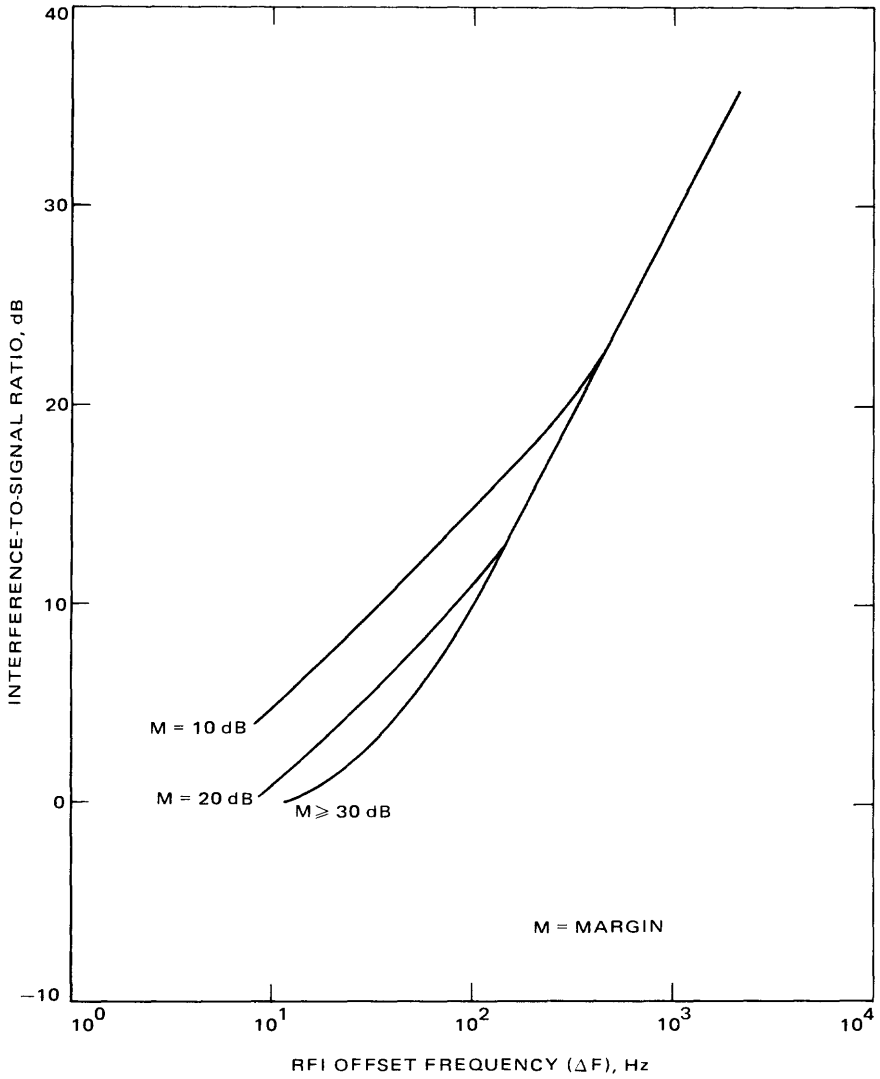


Fig. 11-3. Receiver drop lock levels vs RFI offset frequency with margin as a parameter (Block IV receiver settings: wide/2 kHz/10 Hz)

example is shown in Fig. 11-5. The criterion satisfied by this curve is a phase jitter of not more than 10 deg, a value generally considered suitable for protection of the carrier tracking function needed for navigation and other Doppler dependent measurements.

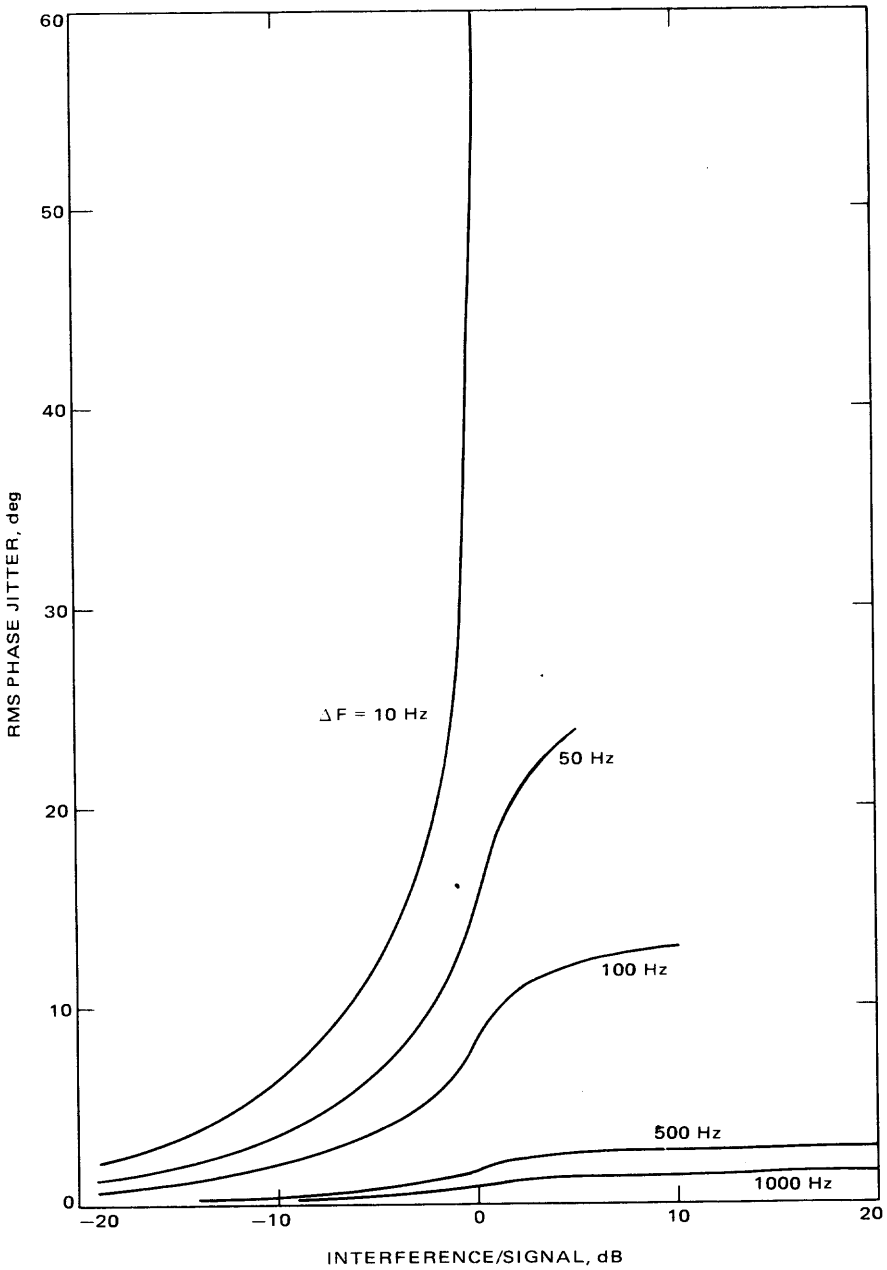


Fig. 11-4. Rms phase jitter vs ISR with ΔF as a parameter (Block IV receiver, strong signal, 2 BL_0 = 10 Hz, wide mode)

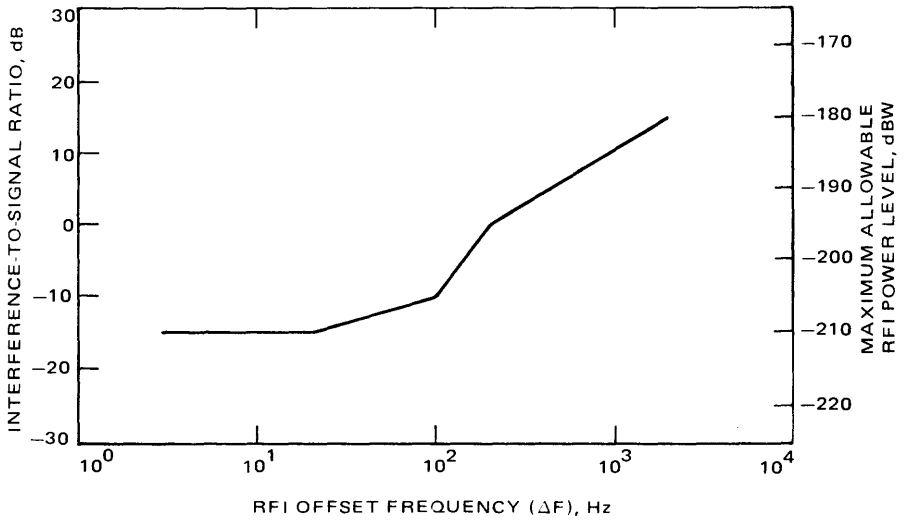


Fig. 11-5. Example of maximum CW ISR and interference power as a function of frequency offset, for the case of 10 dB carrier margin and −205 dBW noise power in 10 Hz bandwidth.

11.8.5 Telemetry Susceptibility to CW Interference

CW interference that falls in or near the telemetry predetection bandwidth can result in degraded performance as expressed in terms of error rate. Perhaps less obvious is the degradation that can be caused by other CW RFI that is translated in frequency by the squarewave demodulation of the telemetry subcarrier. This translation process can move the interfering signal into the telemetry passband, often with sufficient amplitude to result in degraded performance. This section presents data resulting from models that describe the degradation of coded and uncoded data streams as caused by in-band and translated CW RFI. The performance described by the models has been verified by laboratory testing and found to be generally representative of actual interference effects. Keep in mind that the information presented is for the Block IV receiving system and for CW interference. Other receiver or interference types are not included.

The probability of symbol error as a function of the ratio of symbol energy to noise spectral density, with ISR as a parameter, is shown in Fig. 11-6. The curves apply to the error probability for an uncoded data stream. The equivalent loss in symbol energy-to-noise spectral density ratio as a function of ISR is another way to express the effect of CW RFI and this is shown in Fig. 11-7.

The effect of in-band CW RFI on the performance of a telemetry system using convolutional codes is difficult to analyze mathematically. Nevertheless,

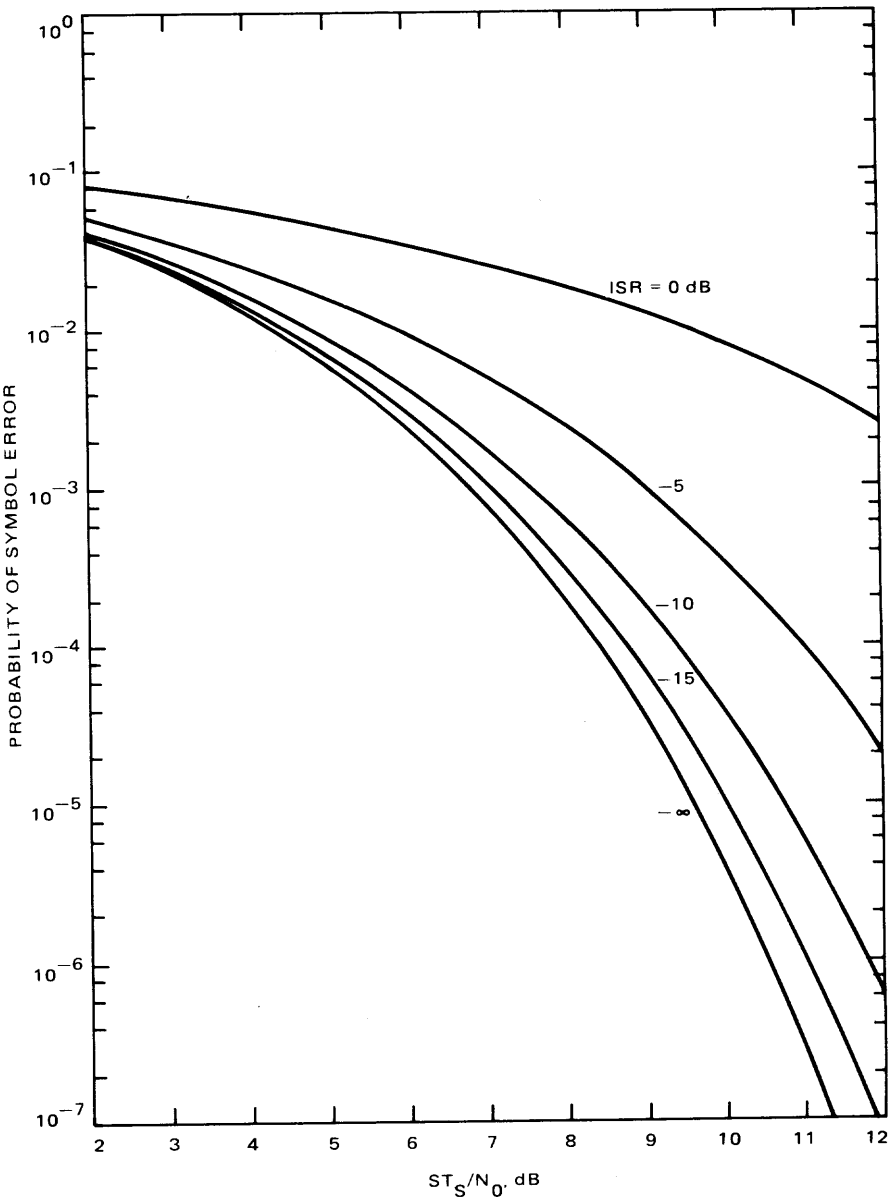


Fig. 11-6. Probability of symbol error vs ST_s/N_0

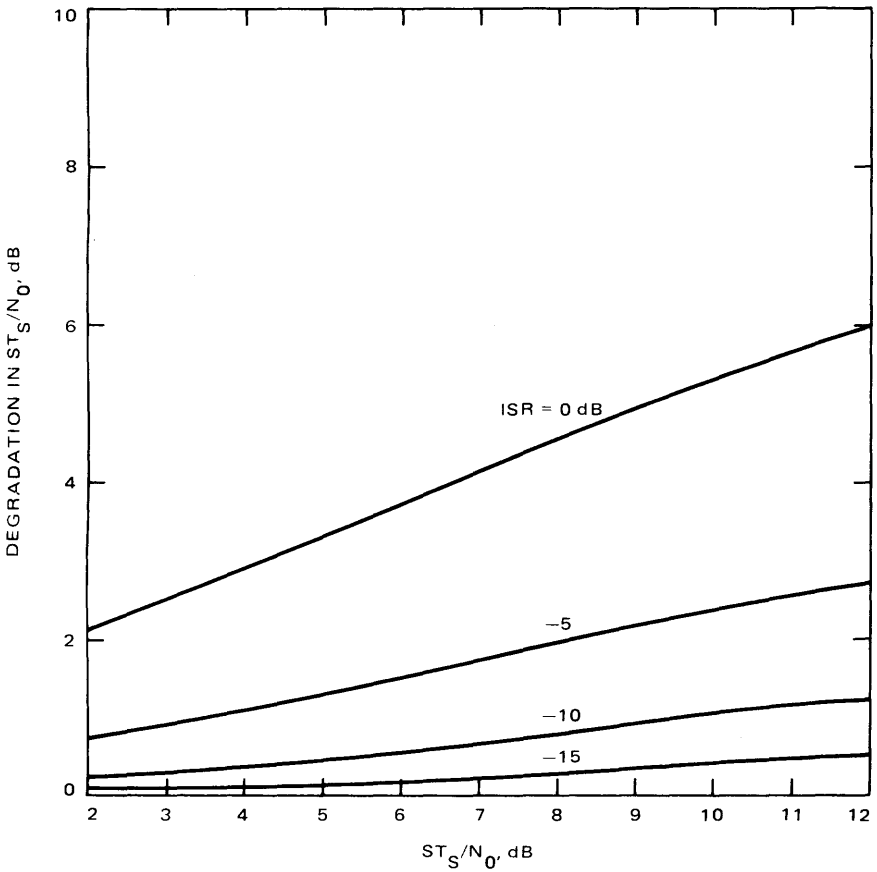


Fig. 11-7. Degradation in ST_s/N_0 vs ST_s/N_0 − dB

a combination of analysis and experimental data results in a model that gives the curves shown in Fig. 11-8.

Through the process of subcarrier demodulation, a CW RFI outside of the telemetry passband is decomposed into a set of components separated by twice the subcarrier frequency. The power of each component is related to its harmonic order. If one of the components falls within the telemetry passband, it will result in performance degradation. Table 11-4 gives the additional power required for an out-of-band CW interference to have the same effect as an in-band interference. For example, in order to produce the same degradation, a CW signal separated from the carrier of the desired signal by 11 times the subcarrier frequency would have to be 22.8 dB stronger than an interfering CW signal that fell within the telemetry passband around the subcarrier frequency. The key point to observe is that a strong enough signal far removed from the desired carrier

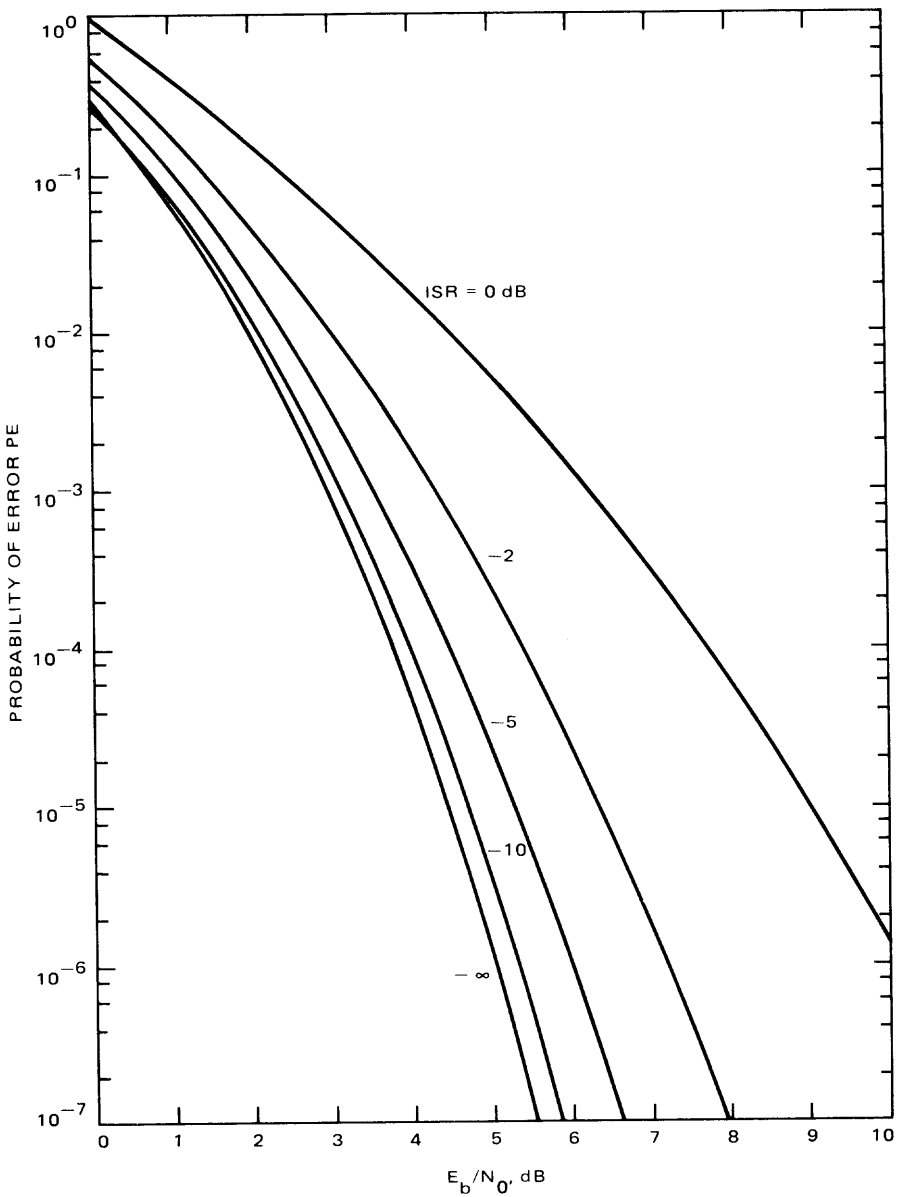


Fig. 11-8. Probability of bit error vs E_B/N_0 for selected values of ISR

Table 11-4. Additional power required for out-of-band CW interference to produce the same effect as an in-band interference

Frequency separation between interference and desired carrier (multiple of sub-carrier frequency, f_{sc}	Additional power, dB
$1 \times f_{sc}$	0
$3 \times f_{sc}$	9.5
$5 \times f_{sc}$	14.0
$7 \times f_{sc}$	17.0
$9 \times f_{sc}$	19.1
$11 \times f_{sc}$	22.8
$13 \times f_{sc}$	25.3
$15 \times f_{sc}$	29.5
$17 \times f_{sc}$	44.6
$19 \times f_{sc}$	67.6

can nevertheless cause performance degradation as a result of the translation effect.

11.8.6 Interference Protection

The degree of susceptibility of deep-space communication systems to various kinds of interference determines the maximum allowable level of interference that can be tolerated without harm. Based upon acceptable degrees of performance degradation, interference protection criteria have been developed at JPL. These criteria are shown in Tables 11-5 and 11-6, and have been adopted by the CCIR. Note that these criteria are single valued for each allocated band, as opposed to the susceptibility curves presented earlier. The purpose of the single value is to provide protection anywhere within each band, thus giving the maximum freedom to assign frequencies to deep-space missions.

A principal use of the protection criteria is in the calculation of potential interference from new systems that may share the band allocations. The procedure, use, and consequences of these calculations are described in the Radio Regulations [11-1].

11.9 Satellite Interference Prediction

This section presents a discussion of the prediction of radio frequency interference from satellites to deep-space downlinks. A description of the information needed for prediction is given, along with an example of one approach to

Table 11-5. Interference protection criteria for DSN earth station receivers

Band GHz	Maximum power spectral density of interference, dBW/Hz	Maximum power spectral flux density of interference, dBW/m^{2}Hz
2.3	−222.5	−255.5
8.4	−220.9	−253.2
13.0	−220.5	−251.7
32.0	−217.3	−239.1

Table 11-6. Interference protection criteria for receivers in deep space

Band, GHz	Maximum power spectral density of interference, dBW/20 Hz
2.1	−190.8
7.2	−188.6
17.0	−186.1
34.5	−183.8

the computation of an interference effect. The difficulty of accurate prediction is described and the consequences of a worst case calculation are pointed out. It is intended that the reader gain an appreciation of the prediction problem, rather than the specific ability or information needed to make predictions.

11.9.1 Satellite Interference

Some earth-orbiting satellites operate in or near deep-space band allocations. Because these satellites are relatively near the earth, they possess a significant potential for interference with the weaker signals from deep space. Since the satellites repeatedly and predictably come into view of deep-space earth stations, it is desirable and, in principle, possible to forecast the possibility of interference.

11.9.2 Information Needed for Prediction

In order to successfully predict the time and effect of satellite interference it is necessary to know:

(1) The characteristics of the satellite-transmitted signal spectrum

(2) The gain of the satellite antenna in the direction of the deep-space earth station

(3) The location and radial velocity of the satellite as a function of time

(4) The gain of the earth-station antenna in the direction of the satellite

(5) The susceptibility of the earth-station receiving system to the Doppler-shifted interfering satellite signal

It is also necessary to have similar information about the spacecraft being tracked, since the effect of interference depends in part upon the strength and spectral characteristics of the desired signal.

11.9.3 Prediction Models and Calculations

To compute the time and effect of interference, mathematical models of susceptibility are employed. One approach is to model the effect of a CW signal and then to test each interfering spectral line against the CW model. For example, the possibility of carrier-loop drop lock may be tested as follows. Drop lock is predicted if:

$$I > P_C \qquad \text{when } \Delta F < B$$

$$\left(I - 20 \log \frac{\Delta F}{B}\right) \geqslant P_C \qquad \text{when } \Delta F \geqslant B$$

where

I is the power of an interfering spectral line.

P_C is the power of the carrier being tracked.

ΔF is the frequency difference between the spectral line and the carrier.

B is the bandwidth of the carrier tracking loop.

Given the proper models, similar computation can be used to predict telemetry drop lock, and degradation of carrier tracking or telemetry performance as expressed by signal-to-noise ratio.

The success of this method of prediction depends upon the accuracy of the models, the ability to accurately compute the amplitude and frequency of each of several signal components, and the validity of the assumption that the interfering signal may be characterized by a set of spectral lines. In practice, it is difficult to make accurate predictions because of uncertainties, particularly with regard to the spectral line amplitude and frequency.

11.9.4 Missed Detection and False Warning

A desirable prediction is one that forecasts an event that later takes place. The objectives for interference prediction are to forecast interference that does occur, and to minimize forecasts of interference that does not occur. Interference that is unpredicted is called a missed detection, and predicted but unrealized interference is called a false warning. For the protection of deep-space communication it is vitally important that missed detections do not occur, at least during critical mission phases. False warnings are less serious, but can damage the credibility that allows effective interference management with other agencies and satellite operations.

The probabilities of missed detection and false warning are not independent and are governed by the information and computation that leads to prediction. Figure 11-9 illustrates how amplitude and frequency uncertainties influence prediction. Carrier drop lock is expected whenever the combination of ΔF and ISR of an interfering line falls between the two curves. Assume that the calculated ΔF and ISR for a spectral line yield the point shown, for a particular instant of time.

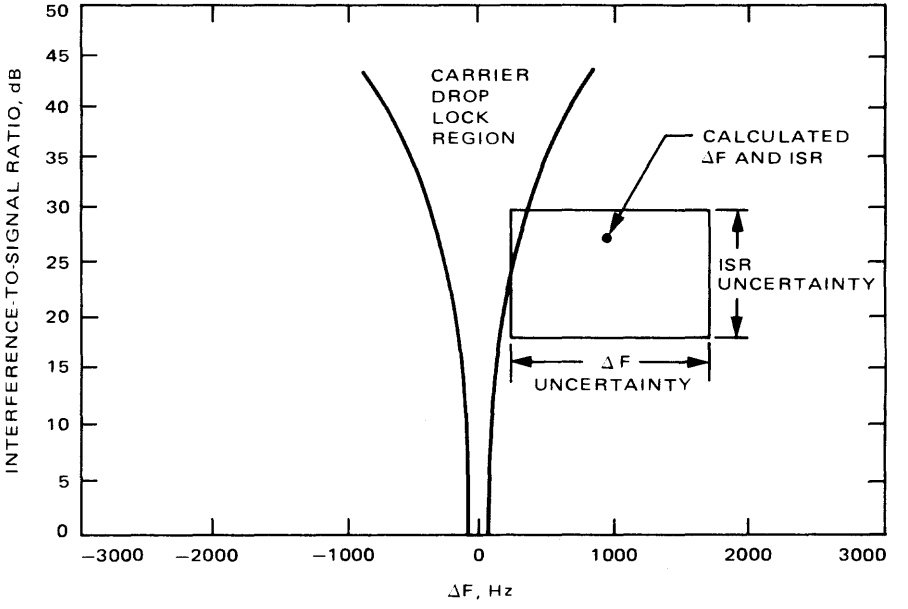


Fig. 11-9. RFI prediction uncertainty for carrier drop lock at a particular instant of time (note: the figure is representative only and may not accurately show the drop lock characteristic)

There is an uncertainty in the value of ΔF that results from error in the calculation of satellite radial velocity and the corresponding Doppler shift, and from imprecise knowledge of the satellite transmitted frequency. Similarly, the ISR is uncertain to a degree determined by error in modeling the off-axis gain of the earth-station antenna in the direction of the true satellite position. This error is a function of time because the differences between the modeled and true antenna gains are angle, and hence, time dependent. In addition, the time dependent antenna gain of the satelliite antenna in the direction of the earth station is imperfectly known.

In the example of Fig. 11-9, one corner of the area of uncertainty surrounding the calculated point falls within the drop lock region. If a missed detection is to be avoided, a drop lock prediction must be issued. Necessarily attendant to this prediction is a high probability that drop lock will not occur, as indicated by the uncertainty area outside of the drop lock region.

It will be seen that a worst case prediction of interference possibility can be made, accompanied by a high rate of false warnings. Accurate prediction, within the usual meaning of accuracy, is difficult. The prediction of satellite interference possibility is nevertheless useful, for it provides the opportunity of corrective or evasive action.

Several courses of action are available. The timing of some spacecraft events may be adjusted to avoid periods of possible interference. Spacecraft or satellite telemetry or tracking modes can sometimes be changed to alter spectral relationships in a way to reduce interference. To protect critical mission phases, it may be possible to negotiate satellite silence during those times. This has been done successfully, even with satellites of other nations.

11.10 RFI Prevention and Avoidance

In foregoing sections of this chapter we have discussed a number of regulatory, cooperative, and technical procedures, methods, and techniques for managing and protecting the radio frequencies used for deep-space communication. There are several other JPL activities that are aimed at protecting the DSN, flight projects, and radio science activities from harmful radio interference. These activities include:

(1) Monitoring the radio frequency environment in the Goldstone area by means of automatic spectrum analysis that examines the deep-space bands and reports signals that exceed a threshold. Although the monitoring equipment is much less sensitive than the DSN earth stations with their large antennas and maser amplifiers, the detection of strong interfering signals provides information that is useful in managing the electromagnetic environment. Signals from nearby fixed, mobile, and airborne

sources, as well as those from earth-orbiting satellites, can be detected by the monitor.

(2) Participation in organizations and procedures for control of the radio emissions in the Goldstone area. Originally a quiet site for reception of weak signals, Goldstone is now surrounded by extensive radio activity, some of which may contaminate the signals from deep space. To control the environment for the benefit of all, there is a coordination group that meets regularly to discuss RFI problems and seek operational or technical solutions to them.

(3) Development of techniques and equipment for improved immunity to interference. Deep-space communication links, and the DSN earth stations in particular, have been designed to give maximum performance in a quiet radio frequency environment. In some instances the design characteristics increase the susceptibility to interference. For example, the maser bandwidth is, from an RFI viewpoint, too wide. One approach to minimize the effect of RFI is to employ filters between the antenna and maser. These filters could have a fixed bandpass just wide enough to accept a single mission spectrum, or could be tunable to reject particular interference. In either case, the filters must be cryogenically cooled to a few kelvins if system performance is to be maintained. Another possibility for improved immunity to RFI is the use of spread spectrum modulation of the downlink spacecraft signal. This creates a bandwidth mismatch between the desired and interfering signals, to the advantage of the desired signal.

The complete scope of acquisition, management, and protection of frequencies for deep-space research is thus seen to be unexpectedly broad and an essential element of mission success.

References

11-1. Radio Regulations, International Telecommunication Union, Geneva, 1981.

11-2. *Recommendations and Reports of the CCIR, 1978, Volume II*, Space Research and Radio Astronomy, International Telecommunications Union, Geneva, 1978 and subsequent issues.

11-3 *Deep Space Network/Flight Project Interface Design Handbook*, 810-5 Rev. D., TRK-20, Rev. B, Jet Propulsion Laboratory, Pasadena, Calif., January 1981 and subsequent revisions (an internal document).

11-4. de Groot, N. F., and Sue, M. K., "Selection of Frequencies for Deep-Space Telecommunications," *TDA Progress Report 42-69*, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1982.

11-5. Hersey, D. R. and Sue, M. K., "Maximum CW RFI Power Levels for Linear Operation of the DSN Block IV Receiver at S-Band Frequencies," *TDA Progress Report 42-56*, Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1980.

11-6. Sue, M. K., "Block IV Receiver Tracking Loop Performance in the Presence of a CW RFI," *TDA Progress Report 42-60*, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1980.

11-7. *Manual of Regulations and Procedures for Federal Radio Frequency Management*, National Telecommunications and Information Administration, United States Department of Commerce, Washington, D.C. 1979 and subsequent revisions.

11-8. Sue, M. K., "DSN RFI Susceptibility Model Development: Program Overview," *TDA Progress Report 42-68*, Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1982.

11-9. Low, P. W., "Radio Frequency Interference Effects of Continuous Sinewave Signals on Telemetry Data," *DSN Progress Report 42-40*, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1977.

11-10. Low, Pauldit, "Radio Frequency Interference Effects of CW Signals on Telemetry Data: Part II," *DSN Progress Report 42-51*, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1979.

11-11. Sue, M. K., "Telemetry Degradation Due to a CW RFI Induced Tracking Error for the Block IV Receiving System with Maximum Likelihood Convolutional Decoding," *TDA Progress Report 42-61*, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1981.

11-12. Sue, M. K., "Performance Degradation of the Block IV Telemetry System Due to the Presence of a CW Interference," *TDA Progress Report 42-69*, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1982.

Appendix A

Band Allocations for Deep-Space Research

Radio frequency band allocations for deep-space telecommunications were adopted by the General World Administrative Radio Conference held in Geneva in late 1979 and became effective in January 1982. The allocations shown below apply to Region 2 (North and South America). Allocations in other parts of the world are similar but not always identical. Variations are confined to the status of other services that share the bands. Footnotes listed are those most relevant to the deep-space allocations. There are other footnotes that apply to other services in these bands.

For complete allocation information, including technical limitations such as maximum power flux density, coordination distance, and minimum elevation angles, consult the Radio Regulations, published by the International Telecommunications Union, Geneva.

NOTE: Capital letters indicate primary allocation, e.g., SPACE RESEARCH. Lower case letters indicate secondary allocation, e.g., Space Research.

1710-2290 MHz

FIXED
MOBILE
Footnote 3707B, 3707D

Footnote 3707B: Subject to agreement obtained under the procedure set forth in Article N13A, the band 2110-2120 MHz may also be used for earth-to-space transmissions in the space research (deep-space) service.

Footnote 3707D: Subject to agreement obtained under the procedure set forth in article N13A, the band 2110-2120 MHz may also be used in Japan for the space research (earth-to-space) and space operation (earth-to-space) services until 31 December 1990.

2290-2300 MHz

FIXED
MOBILE, except aeronautical mobile
SPACE RESEARCH (space-to-earth) (deep-space)

7075-7250 MHz

FIXED
MOBILE
Footnote 3763, 3762A

Footnote 3763: Subject to agreement under the procedure set forth in Article N13A, the band 7145-7235 MHz may be used for earth-to-space transmissions in the space research service. The use of the band 7145-7190 MHz is restricted to deep-space; no emissions to deep-space shall be effected in the band 7190-7235 MHz.

Footnote 3762A: Subject to agreement obtained under the procedure set forth in Article N13A, in Region 2, the band 7125-7155 MHz may be used for earth-to-space transmissions in the space operation service.

8400-8500 MHz FIXED
MOBILE, except aeronautical mobile
SPACE RESEARCH (space-to-earth)
Footnote 3771A

Footnote 3771A: In the space research service, the use of the band 8400-8450 MHz is limited to deep-space.

12.75-13.25 GHz FIXED
FIXED-SATELLITE (earth-to-space)
MOBILE
Space Research (deep-space) (space-to-earth)

16.6-17.1 GHz RADIO LOCATION
Space Research (deep-space) (earth-to-space)

31.8-32.0 GHz RADIO NAVIGATION
Space Research
Footnote 3807, 3807E, 3807D

32.0-32.3 GHz INTER-SATELLITE
RADIO NAVIGATION
Space Research
Footnote 3807, 3807E, 3807D

Footnote 3807: Different category of service: in Bulgaria, Cuba, Hungary, Mongolia, Poland, the German Democratic Republic, Czechoslavakia, and the U.S.S.R., the allocation of the band 31.8-32.3 GHz to the space research service is on a primary basis.

Footnote 3807E: Different category of service: In Australia, Spain, and the United States, the allocation of the band 31.8-32.3 GHz to the space research (deep-space) service in the space-to-earth direction is on a primary basis. This use shall not impose power flux density constraints on the intersatellite service in the bands 32-32.3 GHz.

Footnote 3807D: Subject to agreement obtained under the procedure set forth in Article N13A, the band 31.8-33.8 may also be used in Japan for space-to-earth transmissions in the fixed satellite service up to 31 December 1990.

34.2-35.2 GHz RADIO LOCATION
Space Research
Footnote 3808A

Footnote 3808A: Different category of service: in Australia, Spain, and the United States, the allocations of the band 34.2-34.7 GHz to the space research (deep-space) (earth-to-space) service is on a primary basis.

Article N13A:

This article is a supplementary procedure to be applied in cases where a footnote in the Table of Frequency Allocations requires an agreement with an administration. See Article N13A in the Radio Regulations [11-1], for the full text of procedures to be followed.

Appendix B

Parameters Needed for Interference Analysis and Channel Selection

1. Information Needed for Detailed Study

The information needed to perform an interference analysis and subsequent channel selection is detailed in the following paragraphs.

1.1 Information About The Characteristics of the Telecommunication Systems

(1) Spacecraft antenna gain (for all antennas at their operating frequencies)

(2) Spacecraft transmitter power output (for all possible transmitter modes)

(3) Modulation schemes (including carrier modulation, subcarrier modulation, subcarrier frequencies, number of subcarriers, type of subcarriers (i.e., squarewaves or sinusoids), modulation indices, coding schemes, etc.)

(4) Transmitted data rate or symbol rate

(5) Number of uplink carrier frequencies, downlink carrier frequencies, and their interrelationship

(6) Frequencies preferred for reasons independent of interference considerations

(7) Earth receiving-station antenna gain

(8) Spacecraft receiver sensitivity

1.2 Expected Modes of Operation as a Function of Mission Phase

These include, for example, different combinations of antennas, transmitter power, frequency bands, etc.

1.3 Dates of Important Mission Phases

These include encounters, maneuvers, landings, etc., where interference is less tolerable.

1.4 Mission Trajectory Data

Orbital elements that specify the trajectory of a mission are needed. It is preferred that the orbital elements be provided in the form of classical orbital

elements using "Sun-Centered Earth Equator and Equinox of 1950" as a reference frame. The following parameters are needed:

(1) Six classical orbital elements

semi-major axis (a)
eccentricity (e)
inclination (i)
node angle (Ω)
argument of periapsis (ω)
mean anomaly (M)

(2) Epoch for the above set of elements

(3) Reference frame used for the above orbital elements. (It is preferred to use "Sun-Centered Earth Equator and Equinox of 1950" as a reference frame.)

(4) Time period for which the above set of elements are applicable.

(5) Launch date

(6) Arrival date

(7) Destination

(8) Expected date for end of mission (EOM)

2. Information Needed for Rough Estimation

Parameters listed in preceding paragraphs constitute a complete set of information needed for frequency selection study. Some of these data may not be available at the early phase of mission design. It is possible, even with an incomplete set of data, to perform a frequency selection study by using typical parameter values or worst case values, whichever is appropriate. This, however, may place unnecessary restrictions on channel selection. As a minimum, the following parameters are needed for a crude estimation:

2.1 Characteristics of the Telecommunication Systems

(1) Spacecraft antenna gain

(2) Spacecraft transmitter power output

(3) Earth-station antenna gain

(4) Number of uplink and downlink carrier frequencies and their interrelationship

(5) Frequency preferred for reasons independent of interference consideration

2.2 Mission Trajectory Data

In general, all trajectory parameters listed in the previous paragraph are needed. In some special cases where the trajectory of a spacecraft does not consist of any breaks, it is possible to estimate the trajectory by specifying the following parameters:

(1) Launch date

(2) Arrival date

(3) Destination

(4) End of mission date

Appendix C

Determination of Potential Downlink Interference

1. Introduction

This appendix illustrates the steps necessary to determine if potential interference exists between two deep space downlinks. The two spacecraft examined are designated as 1 and 2 and spacecraft 1 is arbitrarily chosen as the wanted spacecraft and the other as the source of interference. Thus, this appendix examines only interference to spacecraft 1 from spacecraft 2. The interference from spacecraft 1 to spacecraft 2 can be obtained by following the same procedures.

2. Assumptions

The hypothetical system used in this appendix has the following characteristics:

	Spacecraft 1	Spacecraft 2
Downlink e.i.r.p. (dBW)	35	30
Antenna gain (dBi)	30	25
Telemetry symbol rate (BPS)	10K	20K
Telemetry subcarrier frequency (kHz)	300	300
Telemetry modulation angle (deg)	80	40
Number of subcarrier channels	1	1

The Doppler frequency and the total signal-to-interference ratio computed by the program are assumed to have the following values:

	Spacecraft 1	Spacecraft 2
Downlink TSIR (dB)	0	40
Doppler frequency (kHz)	2	1

Both spacecraft are assumed to occupy the same frequency channel.

3. Interference Modes

To determine the interference modes, it is necessary to examine the spectra of the signal and the interference. A sketch of the spectra of these two signals is shown in Figure C-1.

From the sketch, it can be seen that there are two interference modes: carrier-to-carrier and data-to-data. The potential interference for both modes is discussed in the following sections.

4. Carrier-to-Carrier Interference

To determine if potential interference exists for this mode, it is necessary to calculate the power of the wanted signal and the power of the unwanted signal in the carrier channel. The power of the wanted signal in the carrier channel is simply the carrier power of spacecraft 1. The unwanted power is the carrier power of spacecraft 2. The wanted and unwanted power can be calculated as follows:

Spacecraft 2 downlink e.i.r.p.	35.0 dBW
Modulation loss (20 log (cos (80°)))	-15.2 dB
Wanted carrier power level	19.8 dBW

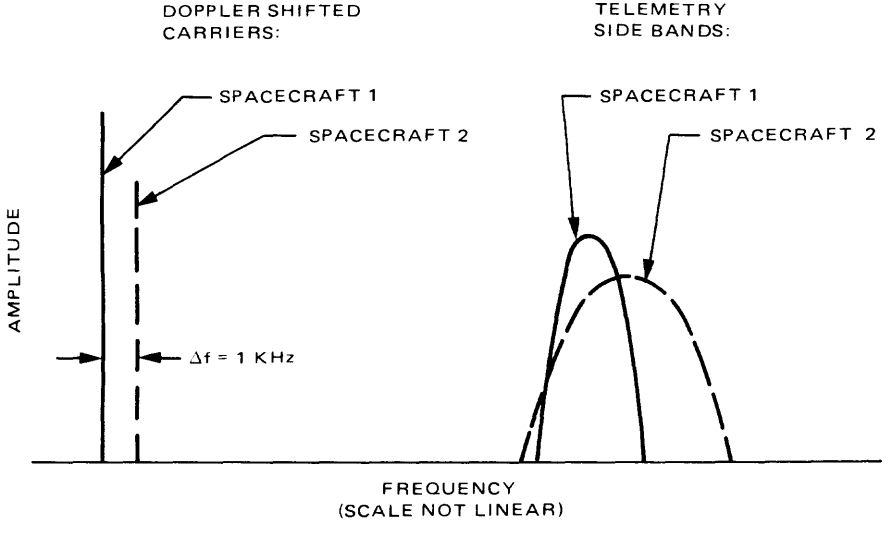


Fig. C-1. Signal spectra for desired spacecraft 1 and interfering spacecraft 2 (carrier-to-carrier and data-to-data interference, co-channel transmission)

Spacecraft 2 downlink e.i.r.p.	30.0 dBW
Modulation loss (20 log (cos (40°)))	-2.3 dB
Unwanted carrier power level	27.7 dBW

The effective interference-to-signal power is thus equal to:

$$EISR = P_W - P_I - TSIR$$

$$= 19.8 - 27.7 - 0 = -7.9 \text{ dB}$$

which exceeds the -15 dB protection ratio by about 7 dB. Potential interference to the carrier of spacecraft 1 thus exists.

5. Data-To-Data Interference

To determine if this mode of interference exists, it is necessary to calculate the data power of spacecraft 1 and spacecraft 2. These power levels can be calculated as follows:

Spacecraft 1 downlink e.i.r.p.	35.0 dBW
Modulation loss (20 log (sin (80°)))	-0.1 dB
Wanted data power level	34.9 dBW

Spacecraft 2 downlink e.i.r.p.	30.0 dBW
Modulation loss (20 log (sin (40°)))	-3.8 dB
Unwanted data power level	26.2 dBW

Since the unwanted signal has a dense spectrum, it is necessary to adjust the unwanted signal power according to the bandwidth conversion factor. The bandwidth conversion factor, B_F, is defined as follows:

$$B_F = \begin{cases} B_w/B_I & B_w < B_I \\ 1 & B_w > B_I \end{cases}$$

where B_w is the bandwidth of the wanted signal and B_I is the bandwidth of the interference. Since the bandwidth of a data channel is proportional to the data rate, the bandwidth conversion factor can thus be calculated from the following equation.

$$B_F = \frac{(\text{Data Rate})_1}{(\text{Data Rate})_2} = \frac{10\text{K}}{20\text{K}} = 1/2 \text{ or } -3 \text{ dB}$$

where the subscripts 1 and 2 refer to spacecraft 1 and spacecraft 2 respectively.

The effective interference-to-signal power is:

$$EISR = P_W - P_I - TSIR$$

$$= 34.9 - (26.2 - 3) - 0 = 11.7 \text{ dB}$$

which exceeds the -15 dB protection ratio. Potential interference to the spacecraft 1 telemetry thus exists.

6. Conclusion

The foregoing analysis of a simple hypothetical case serves as an illustration of the techniques used for interference analysis. Real spacecraft often employ more than one subcarrier channel on the downlink and their spectra are thus more complicated. As a result, interference analysis of an actual system may be more involved.

Chapter 12
Space Shuttle and TDRSS Telecommunications System Interfaces

James C. Springett and Peter W. Kinman

12.1 Introduction

There are two important telecommunications system interfaces: (1) between the spacecraft and the space shuttle, and (2) between the spacecraft and the Tracking and Data Relay Satellite System (TDRSS). Since the shuttle and TDRSS are new, an introductory discussion of their basic telecommunication interface properties is provided. However, this chapter is far from a comprehensive treatment of the subject. For a more complete description of shuttle and TDRSS telecommunications, the reader should consult [12-1], [12-2], and [12-3].

12.2 Space Shuttle Telecommunications Interfaces

Among the various missions that can be carried out by the shuttle orbiter, the primary function is the use of the shuttle as a launch vehicle for satellites to be placed into earth orbit and for spacecraft to be sent into deep space. A user

of the shuttle orbiter system for launch purposes will have need to communicate with the orbiter and with the ground. To provide such communications, the orbiter contains a versatile set of payload-oriented avionic hardware and furnishes a variety of communication links. A user that requires communication links may make use of the shuttle communication systems in either a standard or nonstandard manner. Standard accommodations will meet the majority of user requirements with maximum flexibility and reliability, and with minimum concern and cost. Nonstandard capabilities, however, are also provided so that special/unique user needs may be met. In the nonstandard situation, the user bears a much greater responsibility for the design, implementation, and operation of the communication link.

12.2.1 User/Payload

A user/payload is defined as any system which is carried by the shuttle into orbit, but which is not in any way a functional part of the orbiter itself. More specifically, unmanned satellites and spacecraft are the payloads with which this document is primarily concerned.

Payloads are divided into two distinct classes: (1) those which separate or become "detached" from the orbiter, and (2) those which remain "attached" to the shuttle in the associative surroundings of the cargo bay. Many detached payloads will be transported into geosynchronous or other earth orbits or placed on deep space trajectories using upper stage launchers. Certain detached payloads (known as free-flyers) will simply operate away from the orbiter in co-orbit, and some of these will be subsequently recovered by the shuttle for return to the ground. Usually, attached payloads will be serviced via hardwire links, while communications with detached payloads must use radio frequency (RF) channels.

12.2.2 Payload/Shuttle/Ground Communications Network

The shuttle communications and tracking system provides links between the orbiter and ground and between the orbiter and attached/detached payloads. The orbiter communicates with ground stations directly or through the TDRSS. Payloads communicate with the orbiter through hardline cables (attached payloads) or the payload RF link (detached payloads). Table 12-1 lists the major unmanned payload communication functions and the links over which they are handled.

Figure 12-1 is a pictorial representation of the shuttle and the principal RF channels that comprise the communication links between the shuttle, payloads, and the ground. The links between the shuttle and payloads are at S-band; the shuttle/ground direct links are also S-band. Relay links through the Tracking and Data Relay Satellite (TDRS) are at S-band and Ku-band. Only one detached payload may be communicated with at any time. Similarly, only one coherent

Table 12-1. Orbiter avionics services to payloads

Communication function	Payload/ground direct or through TDRS		Payload/orbiter hardline		Payload/orbiter RF link	
	Payload to ground via orbiter	Ground to payload via orbiter	Orbiter to attached payload	Attached payload to orbiter	Orbiter to detached payload	Detached payload to orbiter
Scientific data	X			X		
Engineering data	X	X		X		X
Command		X	X		X	
Guidance, navigation, and control		X	X	X	X	
Caution and warning	X			X		X
Master timing			X			
Uplink data		X	X			

shuttle/ground direct link is available. Since, however, the FM direct link utilizes separate equipment, it may be worked simultaneously with the coherent direct link. TDRSS relay can make use of the S-band and Ku-band capabilities concurrently. Since the shuttle operates in low orbit (100 to 500 nmi), the time that it may communicate with any direct ground station is limited, while nearly continuous contact can be maintained using the TDRSS.

12.2.3 Shuttle Avionic Equipment

Table 12-2 lists by subsystem name (and acronym) the principal orbiter avionic equipment that have a role in the transfer of payload communication signals through the orbiter.

12.2.4 Principal End-to-End Link Configurations

About a dozen different end-to-end configurations are possible for payload telemetry and command signal transfer. In the following, the five most important end-to-end links are outlined with respect to the orbiter avionic subsystems and ground equipment functions involved.

Figure 12-2 is a block diagram of the detached payload standard telemetry S-band direct link. Standard telemetry for NASA payloads involves the transmission of digitally encoded data at specified bit rates. Within the payload itself,

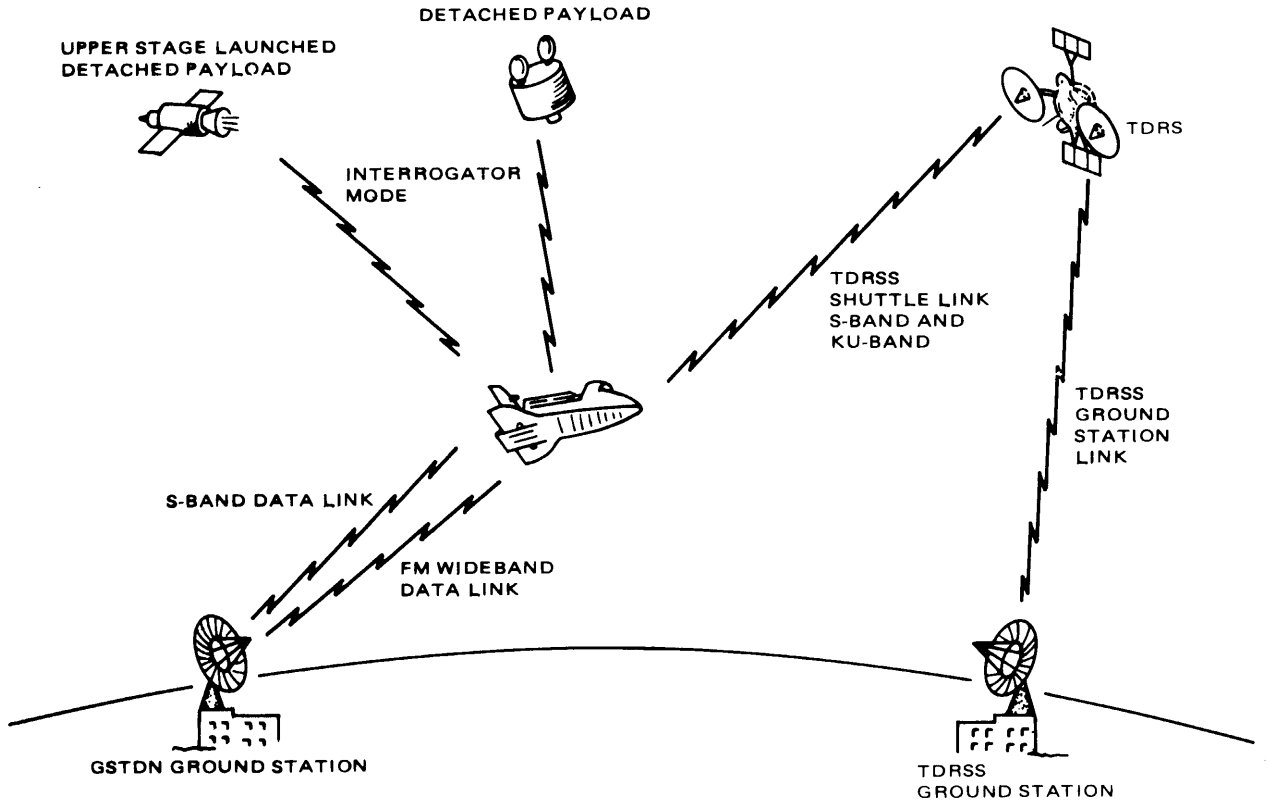


Fig. 12-1. Payload communication network

Table 12-2. Orbiter avionic subsystems and functions

Subsystem name	Acronym	Internal functions									
		Carrier modulation	Carrier tracking and demodulation	Subcarrier demodulation	Data synchronization and detection	Command modulation generation	Data stream multiplexing	Data stream demultiplexing	Data validity check	Data buffering and rate change	Data decoding
Payload interrogator	PI	X	X								
Payload signal processor	PSP			X	X	X			X	X	X
S-band network transponder	–	X	X								
Payload data interleaver	PDI						X				
Network signal processor	NSP				X		X	X		X	X
Ku-band transmitter and receiver	–	X	X								
Ku-band signal processor	KuSP				X		X	X		X	X
Multiplexer/demultiplexer	MDM						X	X	X		X
General-purpose computer	GPC						X	X	X		X

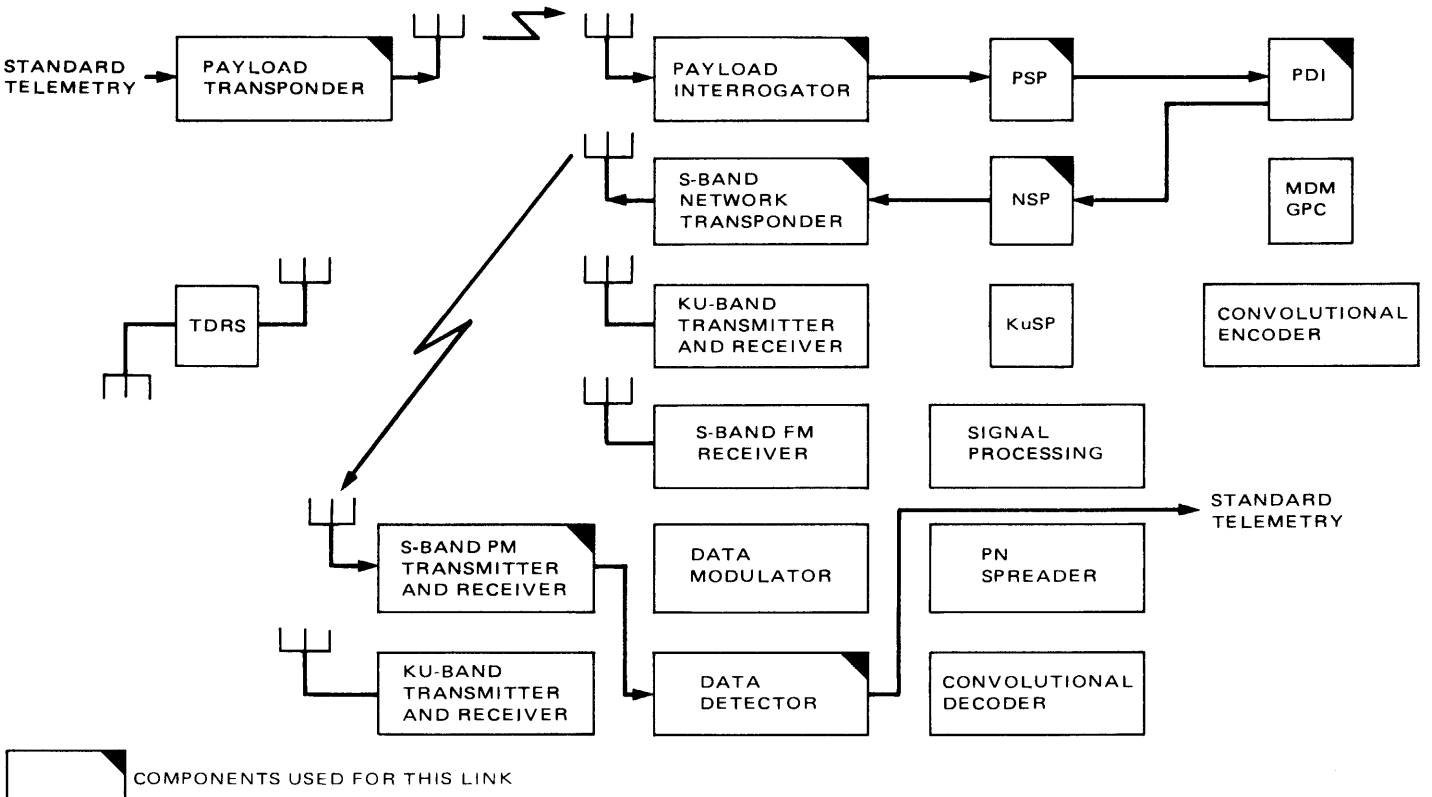


Fig. 12-2. Detached payload standard telemetry S-band direct link

the digital data must be modulated onto a subcarrier of specified frequency and with a non-return-to-zero-(NRZ)-type format.

Standard telemetry is transmitted via the payload transponder and received, tracked, and demodulated aboard the orbiter by the payload interrogator (PI). The payload signal processor (PSP) demodulates the subcarrier and detects the data, following which the payload data interleaver (PDI) and the network signal processor (NSP) function to multiplex the detected detached payload data from the PSP with other attached payload data and orbiter data. A composite digital data stream is then transmitted directly to the ground station via the S-band network transponder.

At the ground station, the telemetry signal is received, demodulated, and detected. It is also demultiplexed so that the standard telemetry data stream, as it appeared at the input to the payload transponder, can be delivered to the appropriate payload/user facility. Because of noisy detection operations that take place in the PSP and in the ground data detector, some bits of information in the telemetry data stream received at the user's facility are in error.

A diagram of the signal flow for the detached payload standard telemetry S-band relay link appears in Fig. 12-3. This configuration is nearly identical to the direct link just described except that the TDRS is used as an intermediate channel between the orbiter and ground. (It should be noted that a separate operating mode of the network transponder and an entirely different ground station and set of ground equipment are used with the TDRS link than with the direct link. Thus, the block diagrams that are employed to depict the end-to-end links should be understood as generic, and specific flight hardware modes and ground equipment and configurations, apart from the indicated generalized functions, are implied.)

Standard telemetry data capability for detached payloads provides a reasonable degree of flexible operation. Certain payloads, however, may not be able to avail themselves of the standard system. To accommodate payloads whose telemetry formats are not compatible with the standard data rates and subcarrier frequency, "bent-pipe" modes of operation are provided within the shuttle's avionic equipment. Several signal paths acting as "transparent throughputs" are available for both digital and analog signals.

Figure 12-4 is a diagram of the signal path for nonstandard telemetry, capable of supporting detached payloads with very high data rates and/or having "analog" telemetry waveforms.

Digital data streams at nonstandard rates or at rates higher than 64 kbps, which cannot be handled by the PSP and PDI, may directly enter the Ku-band signal processor (KuSP), where they are either (1) QPSK-modulated onto a

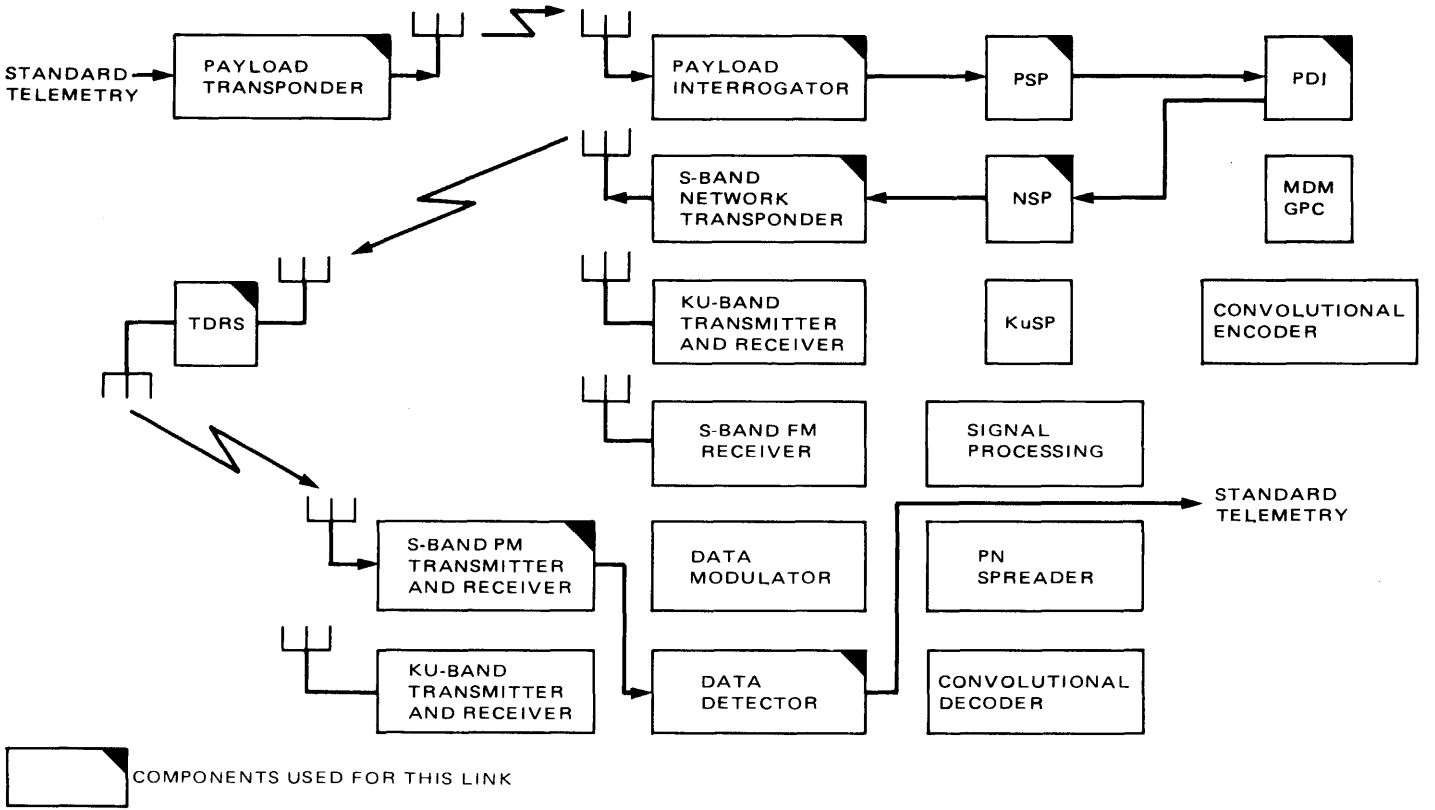


Fig. 12-3. Detached payload standard telemetry S-band relay link

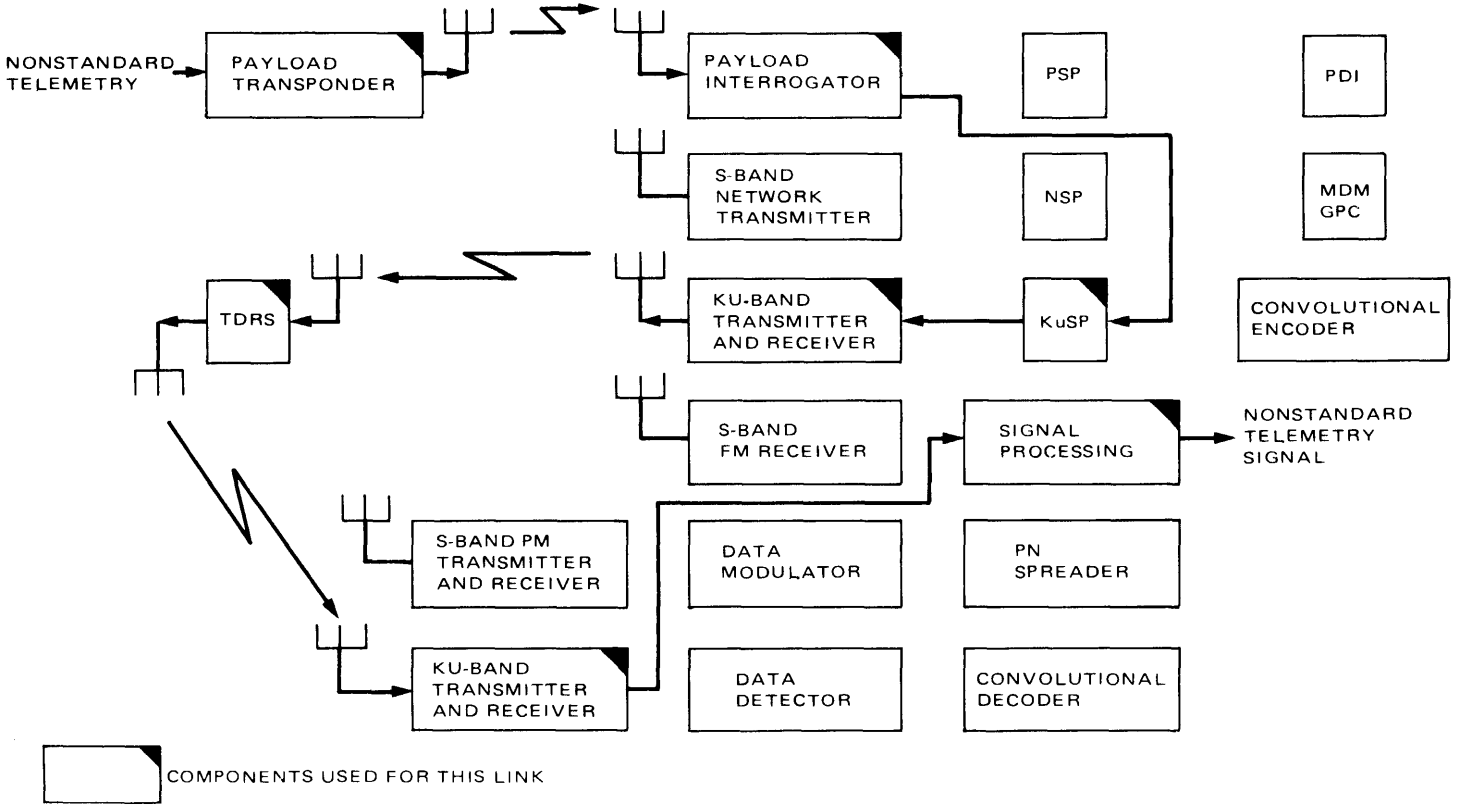


Fig. 12-4. Detached payload nonstandard telemetry bent-pipe link

subcarrier, (2) QPSK-modulated onto the Ku-band carrier, or (3) frequency-modulated onto the Ku-band carrier. Detection and processing of all such data occur at the ground stations employing user-supplied equipment.

Analog signals may take one of two paths. If they are in the form of a modulated subcarrier and do not have significant frequency components above 2 MHz, they may be hardlimited (i.e., one-bit-quantized) and treated as "digital" signals by the KuSP. On the other hand, if the analog signal is baseband in nature on a frequency range up to 4.5 MHz, it may be transmitted via the Ku-band link utilizing FM. Again, all processing is accomplished on the ground by the user's equipment.

Attached payloads may have digital data rate requirements that exceed the standard telemetry capacity. Thus, for data rates up to 50 Mbps, a high-rate data link utilizing the Ku-band TDRS channel is available. The end-to-end configuration is shown in Fig. 12-5.

All data streams transmitted in this mode are convolutionally encoded aboard the orbiter and convolutionally decoded on the ground, in order to provide low error rate communication up to the highest possible data rates.

The final link considered is one that provides ground command capability to detached payloads. Commands may be transmitted from the ground to the orbiter by any one of three links: (1) S-band direct link, (2) S-band TDRS relay link, and (3) Ku-band TDRS relay link. Irrespective of which link is used, they all have similar performance, and detected command data aboard the orbiter is thoroughly checked for validity and errors before it is transmitted to the payload.

Figure 12-6 shows the end-to-end subsystems employed in an S-band relay command link. Encoded (i.e., structured) payload command bits at the ground station are multiplexed with orbiter commands and other data, and PN code-modulated in order to spread the carrier frequency spectrum. (This is a requirement of the TDRSS forward link in order to satisfy transmitted power versus frequency flux density limitations.) The resultant signal is then carrier-modulated and transmitted to the orbiter through the TDRS.

At the orbiter, the S-band network transponder acquires, tracks, despreads (removes the PN code), and demodulates the composite command data stream. In turn, the NSP bit synchronizes and detects the command bits, while the MDM/GPC performs demultiplexing and validation.

The payload command bit stream is input to the NASA PSP, where it is transformed into the proper payload subcarrier signal structure. Transmitted to the payload via the PI, and received, demodulated, and detected by the payload's

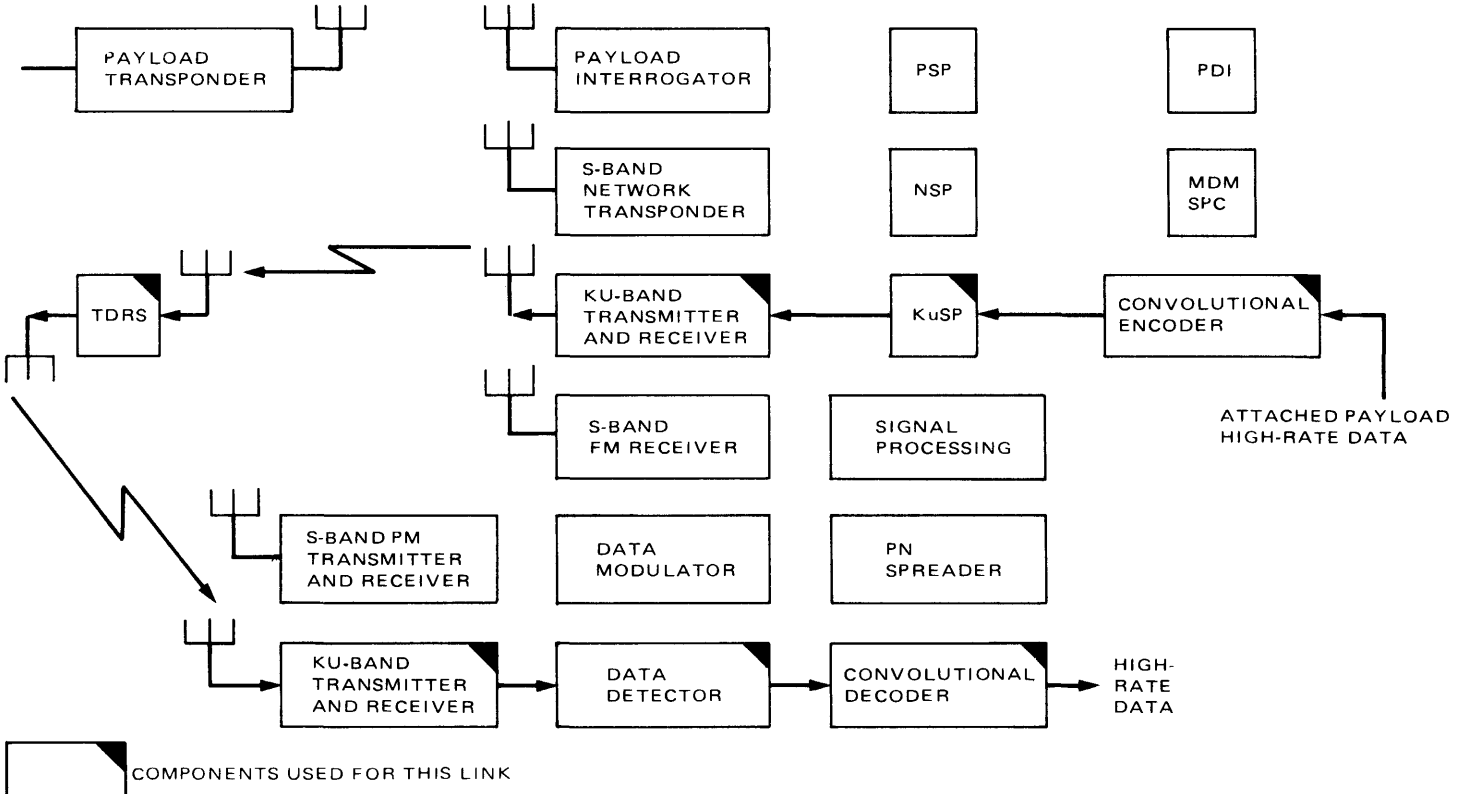


Fig. 12-5. Attached payload high-rate data Ku-band link

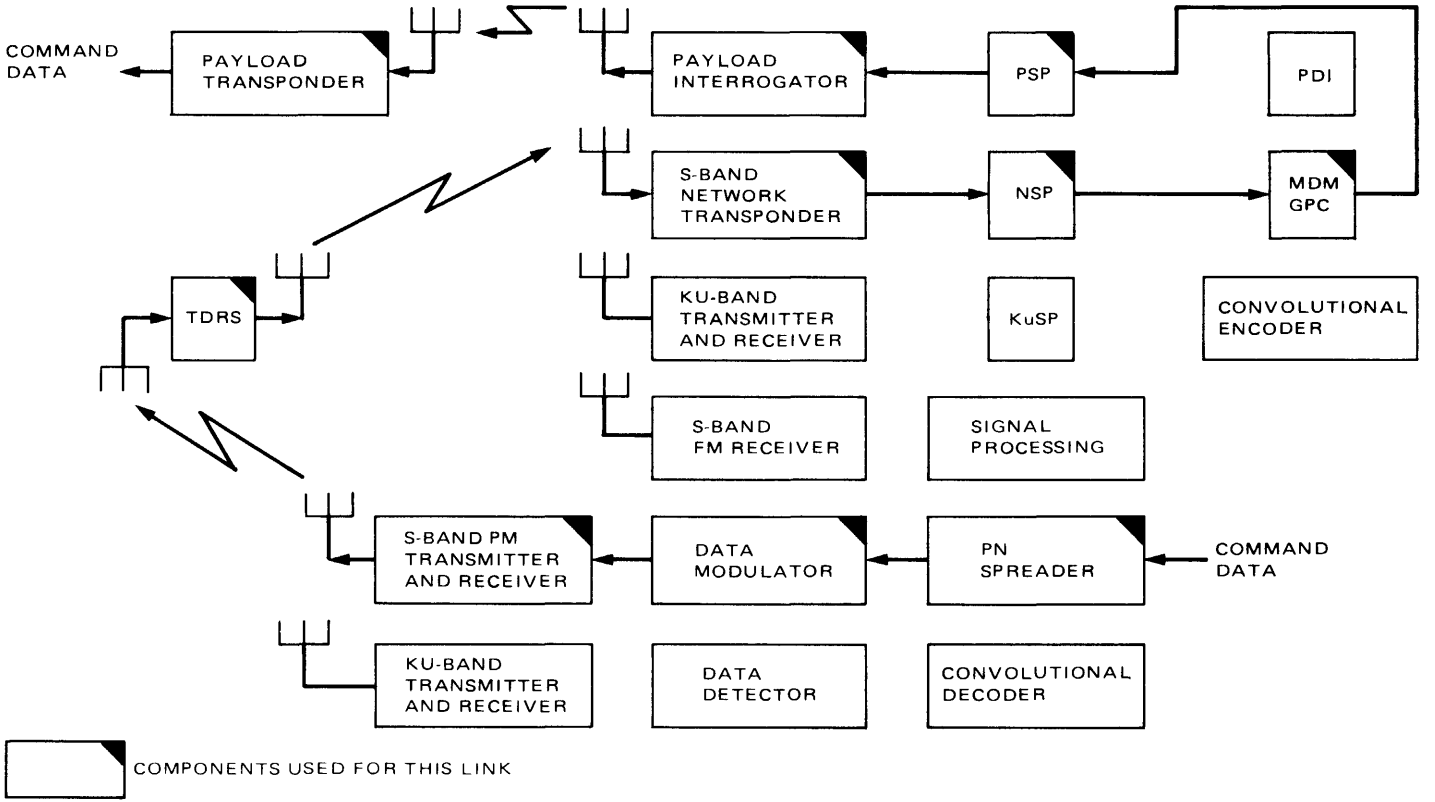


Fig. 12-6. Detached payload command S-band relay link

transponder, the command data is sent to the payload's decoder for final decoding and disposition.

Commands to detached payloads must always be in a standard form; there is no nonstandard command equivalent to the nonstandard telemetry capability.

12.2.5 Attached Payload Electrical Interface Constraints

Probably the most difficult requirement for the attached payload interfaces is that of meeting the interface signal parameters, which are anything but standard between the various avionic subsystems. Table 12-3 summarizes the digital interface parameters, and Table 12-4 tabulates the analog interface parameters. Some specifications are incomplete. It has been recognized that all interface parameters should be standardized, but this will not be accomplished until some future time.

12.2.6 Detached Payload Electrical Constraints and Requirements

The important characteristics of the payload interrogator are listed in Tables 12-5 and 12-6. For standard telemetry modulation, Tables 12-7 and 12-8 are relevant. Table 12-7 contains a breakdown of loss components in the PI and PSP. Table 12-8 lists minimum payload EIRP requirements.

12.2.7 Detached Payload RF Signal Equations

12.2.7.1 Standard signals. The orbiter to payload RF command signal equation is given by:

$$s_c(t) = \sqrt{2P_R} \cos [\omega_R t + \theta_c c(t) \cos \omega_{sc} t] \qquad (12.2\text{-}1)$$

where

P_R = signal total received power

ω_R = payload transponder receive frequency = $2\pi f_R$

θ_c = command subcarrier phase modulation index = 1 rad

$c(t)$ = command data bit stream

ω_{sc} = command subcarrier frequency = $2\pi f_{sc} = 2\pi \times 16 \times 10^4$ Hz

Equation (12.2-1) may be decomposed into its two principal components, viz.,

$$\text{Carrier} = \sqrt{2P_R} J_0(\theta_c) \cos \omega_R t \qquad (12.2\text{-}2)$$

$$\text{Command} = \sqrt{2P_R} J_1(\theta_c) c(t) \{\sin (\omega_R - \omega_{sc}) t + \sin (\omega_R + \omega_{sc}) t\} \qquad (12.2\text{-}3)$$

Table 12-3. Attached payload digital interface parameters

Parameter	Payload recorder	Payload data interleaver	FM signal processor	Ku-band signal processor
Signal level	3.7–9 V p-p, line-to-line	2–12 V p-p, line-to-line (2.5–9.0 V open PIRN)	5.0 ± 0.5 V p-p, line-to-line (wideband) 1.0 ± 0.5 V p-p, line-to-line (DOD)	2.3–5.0 V p-p, line-to-line
Rms SNR	–	–	45 dB rms to rms noise dc to 2 MHz	35 dB minimum
Data rates	25.5 kbps - 1.024 Mbps	10 bps - 64 kbps	200 bps - 5 Mbps, NRZ-L 200 bps - 2 Mbps, Biphase-L 250 bps - 256 kbps (DOD)	16 kbps - 2 Mbps NRZ-L, M, S 16 kbps - 1.024 Mbps Biphase-L, M, S 2–50 Mbps NRZ-L, M, S 16 kbps–4 Mbps, NRZ-L, M, S
Signal code	Biphase-L	NRZ-L, M, S Biphase-L, M, S	NRZ-L Biphase-L	NRZ-L, M, S Biphase-L, M, S
Rise and fall times	10% of bit duration	Maximum: 5 μsec or 10% of bit duration, whichever is less Minimum: 100 nsec	Maximum: 50 nsec (wideband) 100 nsec (DOD)	5% or 50 nsec, whichever is less (KuSP specification) ICD: 40 nsec to 1 Mbps 20 nsec 1–2 Mbps 5 nsec 2–4 Mbps 10% of bit duration 2–50 Mbps
DATA/$\overline{\text{DATA}}$ skew		±200 nsec	–	–

Table 12-3 (contd)

Parameter	Payload recorder	Payload data interleaver	FM signal processor	Ku-band signal processor
Data asymmetry	Included in bit jitter specification	–	(±3% TDRSS user constraint; JSC recommends ±10%)	–
Waveform distortion	–	Overshoot/undershoot less than 20% of signal level	–	–
Clock frequency jitter	–	–	–	±0.1% rms of data rate
Bit jitter	<2% of bit duration (p-p) (includes data asymmetry)	–	–	(<0.1% TDRSS user constraint)
Bit rate accuracy	–	±2%	–	–
Bit rate stability	–	1 part in 10^5 (60-sec average)	–	0.01% (long term)
Clock/data skew	–	5% of clock period or 10 μsec, whichever is less	–	–
Clock duty cycle	–	50.0 ± 5%	–	–
Load impedance	75 ± 10% ohms	Min: 74 ohms Max: 91 ohms	75 ± 10% ohms	75 ± 5 ohms 50 ± 10% ohms (2–50 Mbps)

Table 12-3 (contd)

Parameter	Payload recorder	Payload data interleaver	FM signal processor	Ku-band signal processor
Cable impedance at 1 MHz	75 ± 5 ohms,	75 ± 5 ohms, TSP 2900 pF capacitance[a]	75 ± 5 ohms, TSP	75 ± 5 ohms, TSP 50 ohms (RF 142BU) (2-50 Mbps)
Source impedance	TTL-compatible (TI 5514 IR equivalent)	–	75 ± 10% ohms	–
Noise immunity	–	150 mV p-p[b] dc to 100 kHz	–	–
Common mode voltage (line-to-signal ground)	±15 V	±3 V	–	±10 V

[a] 18–23 pF per foot of cable.

[b] 100 mV p-p maximum EMI, 50 mV p-p from payload.

Table 12-4. Attached payload analog interface parameters

Parameter	Payload recorder	FM signal processor	Ku-band signal processor
Signal level	1 V rms ± 6 dB	1 V ± 10% peak-to-peak line-to-line	2 V rms peak-to-peak line-to-line 7 V maximum
Bandwidth	1.9 kHz - 1.6 MHz	300 Hz - 4 MHz	Dc to 4.5 MHz
Signal/noise			45 dB rms to rms noise (minimum)
Coupling type	Differential	Balanced differential	Balanced differential Direct coupled
Load impedance	75 ± 5% ohms	75 ± 10% ohms	75 ± 5 ohms
Cable impedance	75 ± 5 ohms, TSP	75 ± 5 ohms, TSP	75 ± 5 ohms, TSP
Source impedance	Capable of driving 150 ft of cable	75 ± 10% ohms	75 ± 5 ohms
Common mode rejection	±15 V	±1 Vdc to 2 MHz	±15 Vdc to 10 kHz

Table 12-5. Principal PI receiver characteristics

Parameter	Value	Units
Input frequency range	2200–2300	MHz
Input signal level operating range	−124 to +10	dBm
AGC range	−124 to −20	dBm
Noise figure	7.0 max	dB
Carrier thresholds:		
Acquisition	−120.0	dBm
Tracking	−124.0	dBm
Acquisition sweep range	±125	kHz
Acquisition sweep rate	330	kHz/sec
Frequency rate tracking	17	kHz/sec
Input frequency offset range	±87	kHz
Tracking bandwidth	2320	Hz
Maximum phase noise	15	deg-rms
Maximum SPE	10	deg
Throughput bandwidth	< 5.5	MHz
Throughput SNR losses	2.1 max	dB

Table 12-6. Principal PI transmitter characteristics

Parameter	Value	Units
L-band frequency range	1763–1840	MHz
S-band frequency range	2025–2120	MHz
Carrier frequency tolerance	±0.001	%
Carrier phase noise	10 max	deg-rms
Output spurs	> 67	dBc
Phase modulator	0.2–2.5	rad
Frequency sweep ranges	±75	kHz
	±55	kHz
	±33	kHz
Frequency sweep rates	10	kHz/sec
	250	Hz/sec
Power level:		
High	39	dBm
Medium	30	dBm
Low	7	dBm

Table 12-7. PI/PSP SNR maximum loss components for standard modulations

Component	Loss, dB
PI interference degradation	1.0
PI phase noise loss	0.2
PI demodulation phase offset loss	0.5
PI filtering loss	0.2
PI nonlinear loss	0.2
PSP subcarrier demodulator and bit synchronizer loss	0.8

Table 12-8. NASA payload minimum EIRP requirements for standard telemetry signals

Bit rate, kbps	EIRP, dBm
16	28.2
8	25.1
4	22.1
2	19.1
1	16.2*

*Set by carrier minimum acquisition level requirement.

where $J_0(\theta_c)$ and $J_1(\theta_c)$ are Bessel functions of the first kind. The carrier and command component powers are given, respectively, by:

$$\text{Carrier power} = P_R J_0^2(\theta_c)$$

$$= 0.5855 P_R$$

$$= P_R \text{ (dBm)} - 2.3 \text{ (dB)} \qquad (12.2\text{-}4)$$

$$\text{Command power} = 2P_R J_1^2(\theta_c)$$

$$= 0.3874 P_R$$

$$= P_R \text{ (dBm)} - 4.1 \text{ (dB)} \qquad (12.2\text{-}5)$$

The payload to orbiter RF telemetry signal equation is given by:

$$s_T(t) = \sqrt{2P_T} \cos [\omega_T t + \theta_T m(t) \cos \omega_{ST} t] \qquad (12.2\text{-}6)$$

where

P_T = signal total transmit power

ω_T = payload transponder transmit frequency = $2\pi f_T = (221/240)\ \omega_R$

θ_T = telemetry subcarrier phase modulation index = 1 rad

$m(t)$ = telemetry data bit or symbol stream

ω_{ST} = telemetry subcarrier frequency = $2\pi f_{ST} = 2\pi \times 1.024 \times 10^6$ Hz

Equation (12.2-6) may be decomposed into its two principal components, viz.,

$$\text{Carrier} = \sqrt{2P_T} J_0 (\theta_T) \cos \omega_T t \qquad (12.2\text{-}7)$$

$$\text{Telemetry} = \sqrt{2P_T} J_1 (\theta_T)\, m(t) \{\sin (\omega_T - \omega_{ST})\, t + \sin (\omega_T + \omega_{ST})\, t\} \qquad (12.2\text{-}8)$$

The carrier and telemetry component powers are given, respectively, by:

$$\begin{aligned} \text{Carrier power} &= P_T J_0^2 (\theta_T) \\ &= 0.5855 P_T \\ &= P_T\,(\text{dBm}) - 2.3\ (\text{dB}) \end{aligned} \qquad (12.2\text{-}9)$$

$$\begin{aligned} \text{Telemetry power} &= 2P_T J_1^2 (\theta_T) \\ &= 0.3874 P_T \\ &= P_T\,(\text{dBm}) - 4.1\ (\text{dB}) \end{aligned} \qquad (12.2\text{-}10)$$

12.2.7.2 **Nonstandard telemetry signals.** The following equations represent typical acceptable forms of nonstandard telemetry signals. (Nonstandard command signals are not allowed.)

(1) Two subcarrier data modulations

$$s(t) = \sqrt{2P_T} \cos [\omega_T t + \theta_1 m_1(t) \cos \omega_1 t + \theta_2 m_2(t) \cos \omega_2 t] \quad (12.2\text{-}11)$$

where

θ_1, θ_2 = modulation indicies

$m_1(t), m_2(t)$ = data streams

ω_1, ω_2 = subcarrier frequencies

(2) Single square-wave subcarrier

$$s(t) = \sqrt{2P_T} \cos [\omega_T t + \beta m(t) \, \mathrm{Sq}(\omega_S t)] \quad (12.2\text{-}12)$$

where

β = modulation index

$m(t)$ = data stream

ω_S = square-wave subcarrier frequency

$Sq(\omega_S t)$ = square-wave subcarrier

$= Sgn \cos(\omega_S t)$

(3) Direct carrier data modulation

$$s(t) = \sqrt{2P_T} \cos [\omega_T t + \beta m(t)]$$

$$= \sqrt{2P_T} \cos(\beta) \cos \omega_T t - \sqrt{2P_T} \sin(\beta) \, m(t) \sin \omega_T t \quad (12.2\text{-}13)$$

(4) Frequency-shift-keyed (FSK) subcarrier

$$s(t) = \sqrt{2P_T} \cos \left[\omega_T t + \beta \cos \left\{ \omega_S t + \Delta\omega \int_{-\infty}^{t} m(\lambda) \, d\lambda \right\} \right] \quad (12.2\text{-}14)$$

where $\Delta\omega$ is the FSK deviation ($\Delta\omega = 2\pi \Delta f$). Note that the instantaneous subcarrier frequency is $\omega_S + \Delta\omega m(t)$.

(5) Analog FM subcarrier

$$s(t) = \sqrt{2P_T} \cos \left[\omega_T t + \beta \cos \left\{ \omega_S t + 2\pi K_f \int_{-\infty}^{t} x(\lambda) d\lambda \right\} \right] \quad (12.2\text{-}15)$$

where

K_f = subcarrier FM deviation sensitivity

$x(t)$ = analog modulating signal

12.3 TDRSS Telecommunications Interfaces

TDRSS is a satellite-based data transfer and tracking facility. Using two operational geostationary satellites and a single ground facility, the TDRSS is able to service user satellites in earth orbit between altitudes of 200 and 12,000 km.

12.3.1 General System Configuration and Services

The TDRSS consists of two operational satellites in geostationary orbit separated 130 deg in longitude. Additionally, an in-orbit spare satellite is available to replace either of the operational satellites in the event of a failure. The principal ground facility that communicates with both operational satellites is located within the White Sands Test Facility (WSTF).

A user satellite is defined as an earth satellite operating generally between the altitude limits of 200 and 12,000 km. (Specific coverage limits are given in Subsection 12.3.2.) Because the TDRSS satellites are in geostationary orbit (approximately 35,800-km altitude), the user satellite "looks up" to establish communication. A single TDRSS satellite is able to "see" somewhat more than 50% of all the user satellites at any given time.

The RF link from the WSTF through the TDRSS satellite to the user satellite is known as the forward link. The reverse situation is the return link. Since the TDRSS is a communication service, it is not explicitly concerned with the contents of the data that is transferred on either the forward or return link. No processing of data in either direction, other than reception, frequency translation, and amplification, is performed by the satellite. Thus, the satellite functions as a "bent-pipe" repeater, with all signal processing being accomplished aboard the user satellite or at the WSTF.

Four principal services are available: (1) doppler tracking, (2) ranging, (3) forward link data (command), and (4) return link data (telemetry). Single access (SA) links (one user per assigned channel) are available at both S-band and Ku-band, while multiple access (MA) links (several users per assigned channel by means of spread spectrum modulation) are provided at S-band only. Multiple access is designed to serve relatively low data rate users ($\leqslant$1.5 Mbps[1]), while single access will accommodate data rates as high as 300 Mbps[1] at the Ku-band frequency.

Finally, the interface between the WSTF and the user's ground facility/center is provided by the NASA Ground Communication System (NASCOM). NASCOM's capabilities are not discussed in this document; interested readers are referred to [12-4].

[1] Return link limits.

Before proceeding to the detail sections that follow, it should be noted that it is not the purpose of this document to provide an exhaustive treatment of the TDRSS. Rather, the intent is to orient prospective users and interested readers to the most important features of the system so that a working understanding of the principal functions and how they are accomplished may be attained. Performance parameters as provided are therefore somewhat abridged, and full particulars may be obtained from [12-4] and [12-5].

12.3.2 User Spacecraft Orbital Coverage

Figure 12-7 shows the TDRSS satellite placement and orbital coverage for user satellites [12-6]. With a longitudinal spacing of 130 deg between the two satellites, there exists a small shadow zone, or zone of exclusion (ZOE), centered on the Indian Ocean.

The ZOE establishes the lower-altitude coverage limits for TDRSS users, and the amount of coverage that can be provided to user spacecraft is a function of the user's altitude and inclination. Users at low altitudes and low inclinations will pass through the ZOE each orbit and receive the least coverage. Users at high altitudes and high inclinations will pass through the ZOE only periodically (e.g., a user at 1000 km and 99-deg inclination will pass through the ZOE once per day or less, although the duration of this passage will be greater than for a lower inclination satellite).

For orbital altitudes greater than 1200 km, 100% coverage can be provided up to 2000 km for the MA service and 12,000 km for the SA service. In summary, the following general coverage may be expected:

(1) Minimum coverage of 85% at 200 km.

(2) Coverage at 100% between 1200 and 2000 km for the MA service and between 1200 and 12,000 km for the SA service.

(3) Coverage decreases above 2000 and 12,000 km for the MA and SA services, respectively.

Figure 12-8 shows the ZOE for several user satellite altitudes.

12.3.3 Frequency Plan

Figure 12-9 depicts the MA and SA forward and return link frequency plan, and Table 12-9 summarizes the number of active channels available for each type of data service. MA designates multiple access (available at S-band only), while SSA and KSA are, respectively, S-band single access and Ku-band single access.

Figure 12-9 indicates that a single S-band frequency pair is assigned to the forward and return MA service. For the case of the MA return link, Table 12-9

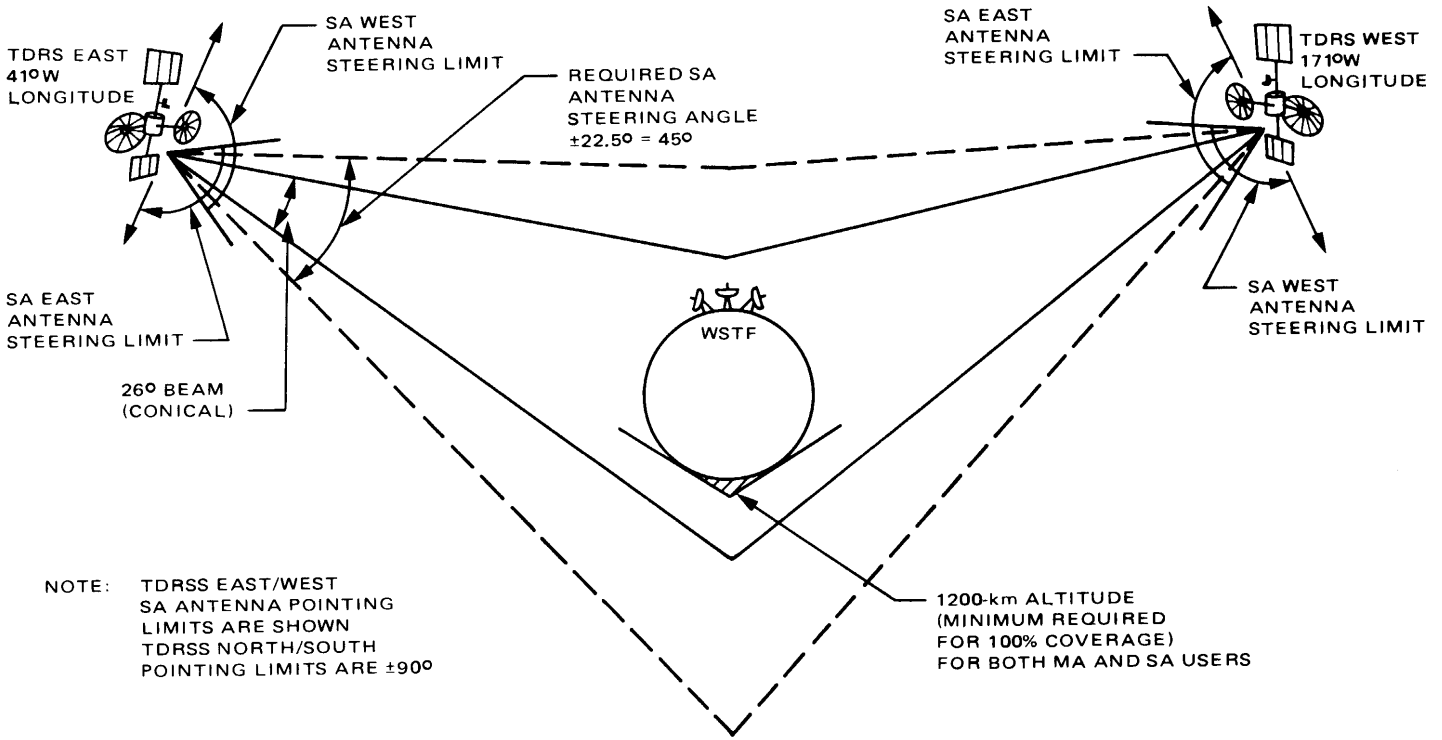


Fig. 12-7. TDRSS satellite orbit coverage

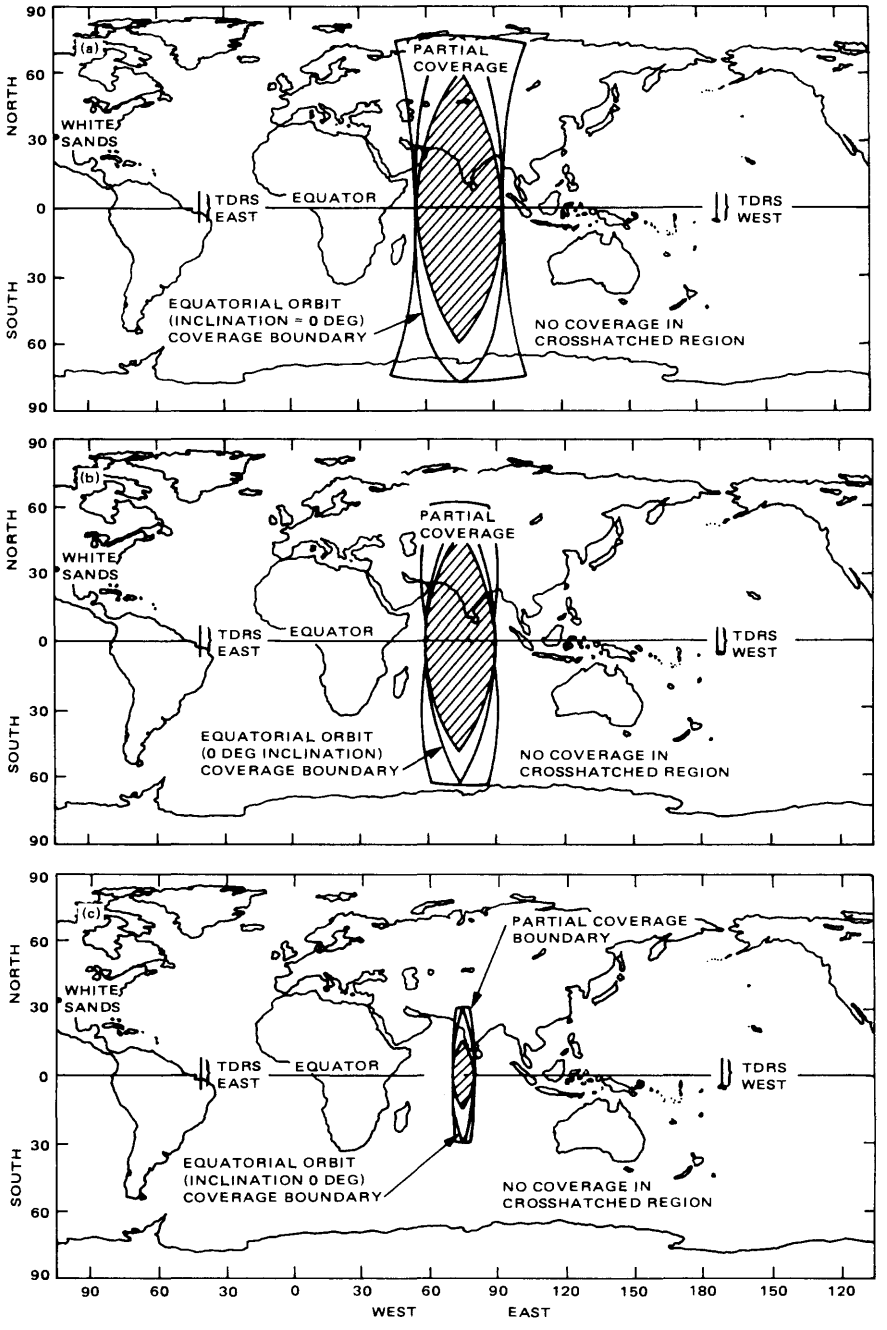


Fig. 12-8. Coverage for (a) 200, (b) 400, (c) 1000 km user satellite altitude 7-deg TDRS inclined orbits with 180-deg phasing

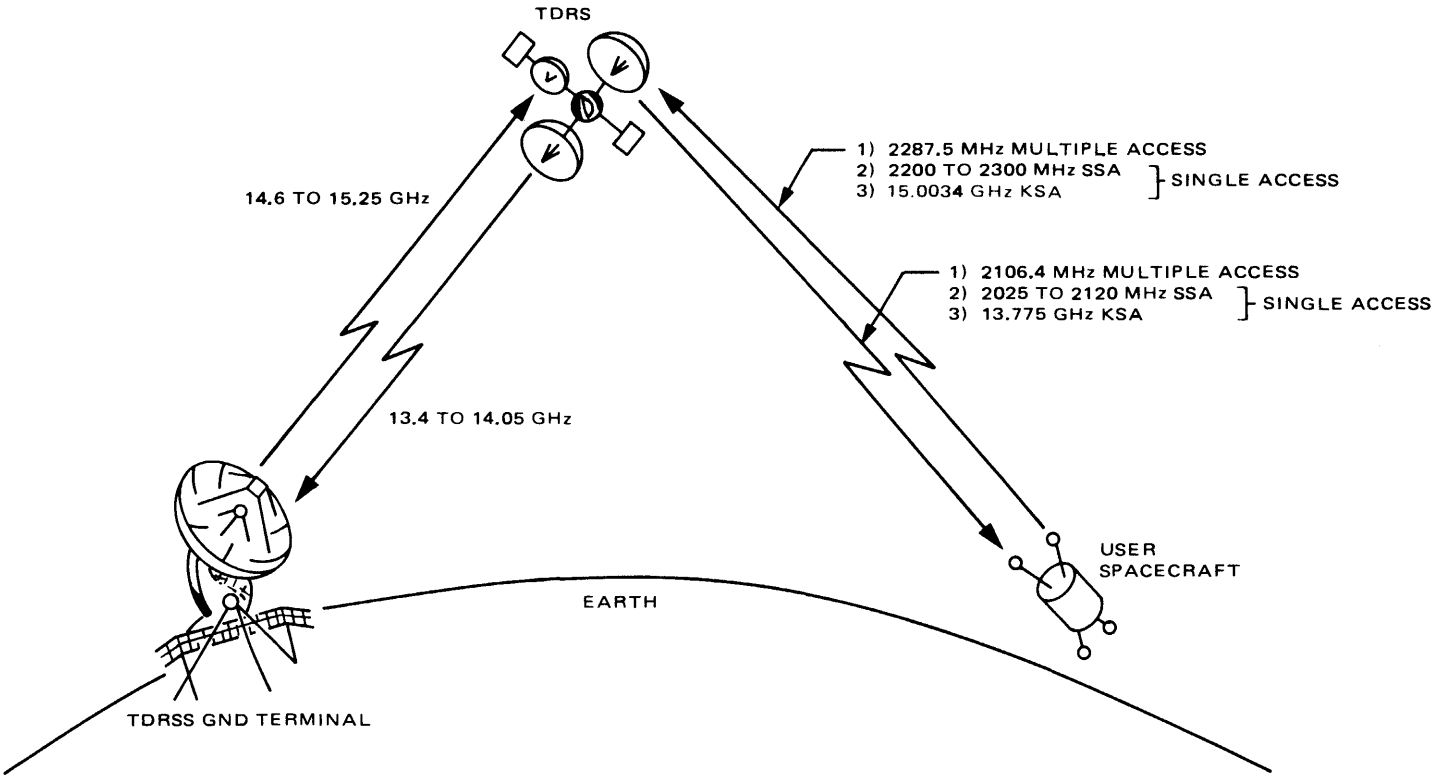


Fig. 12-9. Forward and return link frequency plan

Table 12-9. Active channels available

Service	MA	SSA	KSA
Forward link services			
Quantity of links per satellite	1	2	2
Total links for the TDRSS	3*	6*	6*
Return link services			
Quantity of links per satellite	20	2	2
Total links for the TDRSS	20	6*	6*

*Totals include the spare satellite.

shows that 20 links are possible per satellite, but that the system can accommodate a maximum of 20 MA channels at any given time irrespective of whether 1, 2, or 3 satellites are involved. The reason for this limitation is that the MA satellite antenna system uses an array of 30 helix antennas to form up to 20 independent beams directed toward user satellites. A limit of 20 beams is set by the ground station equipment which functions to actively command the satellite's beam forming network in order to place and maintain each beam on its designated user satellite.

SSA services are provided for a range of assignable frequencies, while the KSA services operate on a single Ku-band frequency pair. Each satellite can provide a maximum or two SSA links and two KSA links because there are two independent SA reflector antennas per satellite, each having dual-frequency feeds. However, since SA antenna pointing is accomplished by mechanical steering of the reflector, two user satellites may be serviced simultaneously by the antenna (one at S-band and the other at Ku-band), provided they are both spatially within the beam.

12.3.4 Definitions of User Services

12.3.4.1 Multiple access forward link data. The MA forward link service through each satellite can support only one user at a time. The TDRSS (including the in-orbit spare satellite) provides the capability for three MA forward link services. Only the two MA forward link services through the two operational satellites will be continually available for user support. These services will be time-shared by all MA users. All MA users will operate at the same frequency and polarization, and will be discriminated by unique PN codes and antenna beam pointing. MA forward link data is usually used for spacecraft command purposes, and is limited to a maximum bit rate of 10 kbps.

12.3.4.2 Single access forward link data. Each satellite can provided two SSA forward link services. The TDRSS (including the in-orbit spare satellite) provides

the capability for six SSA forward link services. Only the four SSA forward link services through the two operational satellites will be continually available for user support. These services will be time-shared by all SSA users. SSA users will be discriminated by frequency, polarization, unique PN codes, and antenna beam pointing. An SSA forward link can also be used to provide forward link service with increased signal EIRP to an MA user. Maximum SSA data rate is 300 kbps.

Each satellite has the capability to provide two KSA forward link services. The TDRSS (including the in-orbit spare satellite) provides the capability for six KSA forward link services. Only the four KSA forward link services through the two operational satellites will be continually available for user support. These services will be time-shared by all KSA users. KSA users will be discriminated by polarization, unique PN codes, and antenna beam pointing. The KSA forward link maximum data rate is 25 Mbps.

12.3.4.3 Multiple access return link data. The MA return link data services provide simultaneous real-time and dedicated operations to low earth-orbiting user spacecraft with real-time data rates up to 50 kbps. Return link support can be provided to each of 20 users during the entire portion of their orbit visible to a satellite (a minimum of 85% of the orbital period). Based upon current mission model projections, no scheduling restrictions should be encountered for return link service. All MA users operate at the same frequency and polarization, and will be discriminated by unique PN codes and antenna beam pointing.

12.3.4.4 Single access return link data. In general, the SA return link data services provide a number of data rate options depending on whether S-band or Ku-band links are employed. Return link service at S-band will provide data rates up to 12 Mbps, and a Ku-band will provide data rates up to 300 Mbps, although maximum rate services will only be provided on a priority scheduled basis, and will not normally be used for dedicated support to any user satellite. SA return link data services are time-shared by all users. Discrimination is based on a combination of frequency, unique PN codes, antenna beam pointing, and polarization.

12.3.4.5 Cross-support data services. The MA services may be thought of as a subset of SSA services since the MA radio frequency falls within the SSA frequency band, and the MA data rates and formats are a subset of the SSA capabilities. Thus, any user satellite designed to operate as MA could be supported in the TDRSS SSA mode. The principal advantage of such cross support is the higher TDRSS satellite antenna gain (approximately 8 dB). Thus, a given user satellite could be designed to provide continuous real-time return link data operating in the MA mode, while periodically receiving SSA support of, say, a special high data rate experiment.

12.3.4.6 Tracking services. All of the above data services (12.3.4.1 through 12.3.4.5) simultaneously provide range and range-rate (R&RR) tracking data for each user satellite. This information is derived from a combination of the PN code displacement (delay) and doppler processing. Accuracies are comparable to that from the ground-based STDN.

12.3.4.7 Simulation and verification services. The TDRSS includes a user simulator for calibration and verification of forward, return, and tracking services, and also for simulation of user/TDRSS forward and return links. The user simulator makes it possible to verify or simulate one MA, one SSA, and two KSA services simultaneously. TDRSS users are also provided a flexible set of testing functions prior to and as part of operational activity with the TDRSS. This includes user data and RF compatibility, network data flow for interface testing, and a full range of simulation services for a network operations and procedures familiarization.

12.3.5 Forward Link General Signal Form

All forward link services have a common signal form so that the user's satellite transponder receiver may be functionally identical for each class of service (MA, SSA, KSA). The main difference involves the frequency conversions from RF (S-band or Ku-band) to a common IF. In fact, an S-band transponder needs simply an external converter to enable it to operate with the Ku-band signal.

Each forward link service signal consists of a PN spread spectrum command channel in carrier phase quadrature with a PN code modulated ranging channel. The mathematical form of the signal is given by:

$$s(t) = \underbrace{\sqrt{2(0.91)P_T}\,PN_I(t)\,c(t)\cos\omega t}_{\text{I channel}} + \underbrace{\sqrt{2(0.09)P_T}\,PN_Q(t)\sin\omega t}_{\text{Q channel}} \qquad (12.3\text{-}1)$$

where

$\omega = 2\pi F$ = carrier frequency

P_T = signal total power

$c(t)$ = command bit stream

PN_I = command channel PN code

PN_Q = ranging channel PN code

The PN_I code length is $2^{10} - 1 = 1023$ chips, while the PN_Q code length is $2^{18} - 2^8 = 261{,}888$ chips. The ratio of the Q to I code lengths is 256, and both

codes run from the same clock (approximately 3.1 MHz). Both codes are epoch-synchronized so that when the shift-register generator which produces the PN_Q code passes through the all "1" state, the PN_I code shift-register generator is also passing through its all "1" state. Such epoch synchronization facilitates PN code acquisition in the user's transponder.

It is also noted that the power ratio between the I and Q channels is 10:1. Thus, 91% of the power is devoted to the command data. No data modulates the Q channel.

The forward link signal parameters for all classes of service are given in Table 12-10. A user transponder receiver's frequency is nominally specified to be f. The nominal carrier frequency (f) transmitted by the TDRSS satellite is compensated for doppler so that the received carrier, f_R, arrives at the user spacecraft within a predictable tolerance (e) of f, i.e., $f_R = f \pm e$. This feature minimizes the doppler resolution requirements of the user transponder and the corresponding time to acquire. Doppler compensation is available continuously to facilitate reacquisition by the user spacecraft in the event of temporary loss of signal. Doppler compensation is inhibited during periods of coherent two-way doppler measurement.

The use of a short-cycled PN code for the range channel allows optimization of the command channel PN code. This feature permits use of Gold codes for the command channel providing a code library with good cross-correlation properties. The complete PN code library, documented in [12-7], is sufficiently large to allow a unique code set assignment for each user satellite.

The PN chip rate is coherently related to the transmitted carrier frequency in all cases. This feature permits the user transponder to use the receiver PN clock (acquired before the carrier) to predict received carrier frequency, thereby minimizing transponder complexity and acquisition time.

Forward link data is directly modulo-2 added to the command channel PN sequence. Since the PN clock and carrier are doppler-compensated, the forward link data will generally be asynchronous with the PN code. Command data greater than 300 kbps will binary-phase-shift-key (BPSK) modulate the KSA forward link carrier, and the range channel will not be transmitted for this condition.

An example of a forward link MA service link budget is presented in Table 12-11.

Table 12-10. Forward link signal parameters

Parameter	Definition
Transmit carrier frequency, Hz	f
Carrier frequency arriving at user spacecraft, Hz	f_R
$\frac{\text{Command channel radiated power}}{\text{Range channel radiated power}} = \frac{P_I}{P_Q}$	10 dB
Range channel	
Carrier frequency	Command channel carrier frequency delayed $\pi/2$ rad
PN modulation	PSK, $\pm \pi/2$ rad
Carrier suppression	30 dB minimum
PN chip rate	Synchronized to command channel PN chip rate
PN code length, chips	$(2^{10} - 1) \times 256$
PN code epoch reference	All "1" condition synchronized to the command channel PN code
PN code family	Truncated 18 stage shift register sequences
Command channel	
Carrier frequency, Hz	Transmit carrier frequency (f)
PN modulation	PSK, $\pm \pi/2$ rad
Carrier suppression	30 dB minimum
PN code length, chips	$2^{10} - 1$
PN code family	Gold codes

	MA	SSA	KSA
PN chip rate, chips/sec	$\frac{31}{221 \times 96} \times f$	$\frac{31}{221 \times 96} \times f$	$\frac{31}{1469 \times 96} \times f$
Data format	NRZ	NRZ	NRZ
Data rate restrictions	0.1 - 10 kbps	0.1 - 300 kbps	1 kbps - 25 Mbps
Data modulation	Modulo-2 added asynchronously to PN code		

Table 12-11. Typical forward link MA service link budget

BER	10^{-5}
TDRS antenna gain, dB	23.0
TDRS transmit power, dBW	13.0
RF transmit loss, dB	-1.0
Transmitted EIRP, dBW, peak ($S + N$)	35.0
TDRS transponder loss, dB	-1.0
Peak signal EIRP, dBW	34.0
Antenna pointing loss, dB	0.0
Signal EIRP, dBW	34.0
Space loss, dB	-191.6
Polarization loss, dB	-0.5
User gain/noise temperature, dB/K	G/T
Power/noise spectral density, dB-Hz	$70.5 + G/T$
UQPSK loss, dB	-0.5
Demodulation loss, dB	-1.5
PN loss, dB	-1.0
Required E_b/N_0, dB (ΔPSK)	-9.9
User margin, dB	-3.0
Achievable data rate, dB	$54.6 + G/T$

12.3.6 Return Link Data Groups and Modes

The return link data services are divided into two data groups: DG1 and DG2. DG1 signal parameters are further subdivided into three operating modes, distinguished as follows:

(1) Mode 1 is used for two-way doppler and range measurements. The return link PN code length is identical to, and epoch-synchronized with, the forward link PN code received from the TDRSS. Acquisition of the Mode 1 return link signal by the WSTF may take place only after the forward link has been acquired by the user's transponder receiver, and the transponder is functioning in the coherent turn-around mode. When a user transponder functions only in Mode 1, it is not necessary to maintain forward link acquisition once the return link has been established. However, loss of forward link tracking by the user transponder must not introduce "transients" that can cause loss of return link tracking.

(2) Mode 2 is used when return link acquisition is desired without the requirement for prior forward link acquisition. This mode of operation uses short, easily acquired PN codes.

(3) Mode 3 is used when two-way range and doppler measurements are required simultaneously with high-rate telemetry data. Restrictions on Mode 3 acquisition are identical to that for Mode 1. In Mode 3, the Q channel contains only data and is not restricted by the PN rate. For MA users, Mode 3 is available only when supported by the SSA service.

DG2 parameters are used when the data rate requirement exceeds the capability of DG1. DG2 operation cannot provide range tracking since return link PN modulation is not used. The DG2 carrier can be either coherently related to or independent of the forward link carrier frequency. Two-way doppler tracking will be provided when the DG2 carrier is coherently related to the forward link carrier frequency. DG2 operates in a single mode.

12.3.7 Return Link General Signal Form

Like the forward link, a degree of commonality exists between the various return link signal forms.

The mathematical form of the signal for DG1, Modes 1 and 2 ,is:

$$s(t) = \sqrt{2P_I}PN_I(t)d_I(t)\cos\omega_1 t + \sqrt{2P_Q}PN_Q(t - T_c/2)d_Q(t)\sin\omega_1 t \tag{12.3-2}$$

where

$\omega_1 = 2\pi f_1$ = carrier frequency

P_I, P_Q = component powers, with three possibilities: $P_Q/P_I = 1, 2, 4$

PN_I, PN_Q = different, equal length, PN sequences

T_c = PN chip period

d_I, d_Q = independent data bit streams that may have identical or different bit rates

Note that the PN_Q sequence is time displaced by 1/2 chip relative to the PN_I sequence. The distinct difference between the signal for each mode is that the PN code lengths for Mode 1 are $2^{18} - 2^8$ chips, while for Mode 2 they are $2^{11} - 1$ chips in length.

The mathematical form of the signal for Mode 3 is:

$$s(t) = \sqrt{2P_I}PN_I(t)d_I(t)\cos\omega_1 t + \sqrt{2P_Q}d_Q(t)\sin\omega_1 t \qquad (12.3\text{-}3)$$

All components are as defined for (12.3-2). No PN is employed on the Q channel; otherwise, the qualifications of the components are identical to Mode 1.

The DG2 signal mathematical form is:

$$s(t) = \sqrt{2P_I}d_I(t)\cos\omega_2 t + \sqrt{2P_Q}d_Q(t)\sin\omega_2 t \qquad (12.3\text{-}4)$$

where $\omega_2 = 2\pi f_2$ = carrier frequency. Other components are as defined for (12.3-2) with the exception that $P_I/P_Q = 1$ or 4 for QPSK modulation. BPSK modulation is also allowed, in which case $P_I = 0$ or $P_Q = 0$.

Table 12-12 summarizes the salient return link signal parameters. Qualifying information is given in the following paragraphs.

The user transmit frequency for DG1 Mode 2 or DG2 service must be defined by the user to an accuracy of ±700 Hz for MA or SSA and ±5 kHz for KSA when requesting return link service.

Staggered quadriphase PN (SQPN) modulation is used for DG1 Modes 1 and 2. With this feature, the spectral characteristics out of a saturated power amplifier will to a great degree retain the spectral characteristics of the band-limited input signal. This results in better control of out-of-band emissions, which provides more efficient communications and less interference to spacecraft using adjacent frequency channels.

For Mode 1 operations, the I and Q channel PN codes are generated from a single linear shift register. The I and Q channel PN codes are identical but offset by at least 20,000 chips. This separation is adquate for TDRSS to uniquely identify each data channel without requiring a unique linear shift register for each channel. User-unique PN code assignments (refer to [12-7]) include shift register tap connections for generating the assigned codes.

The return link data is modulo-2 added asynchronously to the PN code (DG1, Modes 1, 2, and 3). In the case of Mode 1 or Mode 3 operation, this eliminates the doppler on the PN code from affecting the telemetry data rate. For Mode 2 it eliminates the need for synchronizing the spacecraft data block with the user transponder PN clock.

The PN clock is coherently related to the transmitted carrier frequency in all cases. This feature permits the user transponder to use a common source for generating carrier and clock, and also permits TDRSS to use the PN clock to aid carrier acquisition.

Table 12-12. Return link signal parameters

Parameter	Definition
Transmit carrier frequency (Hz)	
Data Group 1	f_1
Data Group 2	f_2
Data Group 1	
PN modulation	
Modes 1 and 2	SQPN
Mode 3, I channel	PSK, $\pm \pi/2$ rad
PN code length, chips	
Modes 1 and 3	$(2^{10} - 1) \times 256$
Mode 2	$2^{11} - 1$
PN code epoch reference	
Mode 1	
I channel	All "1" condition synchronized to all "1" condition of received forward link range channel
Q channel	All "1" condition delayed $x + 1/2$ PN chips relative to I channel epoch
Mode 2	
I channel	Spacecraft oscillator
Q channel	Delayed 1/2 PN chip period relative to I channel epoch
Mode 3, I channel	Same as Mode 1, I channel
PN code family	
Modes 1 and 3	Truncated 18-stage shift register sequences
Mode 2	Gold codes
Data format	
Without convolutional coding	NRZ-L, NRZ-M, NRZ-S, Biϕ-L, Biϕ-M, Biϕ-S
With convolutional coding	NRZ-L, NRZ-M, NRZ-S (see text)
Data modulation	
Modes 1 and 2	Modulo-2 added asynchronously to PN code
Mode 3	
I channel	Modulo-2 added asynchronously to PN code
Q channel	PSK, $\pm \pi/2$ rad

Table 12-12 (contd)

Parameter	Definition		
	MA	SSA	KSA
PN chip rate, chips/sec	$\frac{31}{240 \times 96} \times f_1$	$\frac{31}{240 \times 96} \times f_1$	$\frac{31}{1600 \times 96} \times f_1$
Mode 1 data rate restrictions			
Total	0.1 - 50 kbps	0.1 - 600 kbps	1 - 600 kbps
I channel	0.1 - 50 kbps	0.1 - 300 kbps	1 - 300 kbps
Q channel	0.1 - 50 kbps	0.1 - 300 kbps	1 - 300 kbps
Mode 2 data rate restrictions			
Total	1 - 50 kbps	1 - 600 kbps	1 - 600 kbps
I channel	1 - 50 kbps	1 - 300 kbps	1 - 300 kbps
Q channel	1 - 50 kbps	1 - 300 kbps	1 - 300 kbps
Mode 3 data rate restrictions			
Total	I + Q	I + Q	I + Q
I channel	0.1 - 159 kbps	0.1 - 300 kbps	1 - 300 kbps
Q channel	1 kbps - 1.5 Mbps	1 kbps - 6 Mbps	1 kbps - 150 Mbps
Carrier (f_1) reference, Hz			
Modes 1 and 3	$\frac{240}{221} \times f_R$	$\frac{240}{221} \times f_R$	$\frac{1600}{1469} \times f_R$
Mode 2	Spacecraft Oscillator		
	MA	SSA	KSA
Data Group 2			
Carrier (f_2) reference, Hz	NA	$\frac{240}{221} \times f_R$ or Spacecraft oscillator	$\frac{1600}{1469} \times f_R$ or Spacecraft oscillator
Data format	NA	See text	See text
Data rate restrictions			
Total	NA	1 kbps - 12 Mbps	1 kbps - 300 Mbps
I channel	NA	1 kbps - 6 Mbps	1 kbps - 150 Mbps
Q channel	NA	1 kbps - 6 Mbps	1 kbps - 150 Mbps
Data modulation	NA	Quadriphase PSK or BPSK	Quadriphase PSK or BPSK

For both Mode 1 and 2 operations, the return link can have either a single telemetry data signal or two independent data signals. For a single data signal, the identical data must appear simultaneously on the I and Q channels. For two

independent data signals, one data signal will appear on the I channel, the other on the Q channel. The I and Q channel power division in the user spacecraft transmitter can be weighted up to a maximum weight of 1:4 for either a single data signal or two independent data signals.

For Mode 3, the return link can have either a single telemetry data signal or two independent data signals. For a single data signal, the data will appear only on the Q channel while the I channel will be used for range tracking only. For two independent data signals, one data signal will appear on the I channel and the other on the Q channel. The I and Q channel power division in the user spacecraft transmitter can be weighted up to a maximum weight of 1:4 for either a single data signal or two independent data signals.

The DG2 return link can have either a single telemetry data signal or two independent data signals. The I and Q channel power division in the user spacecraft transmitter can be either 1:1 or 1:4.

The use of convolutional data coding is mandatory for all MA configurations, and the data rate restrictions given for the MA modes in Table 12-12 are for the convolutionally encoded (rate 1/2) data symbols. Whenever Biϕ formats are specified, all data rate restrictions for all services must be reduced by a factor of 2.

The choice of which data format should be used depends upon how the user wishes the recovered data to be made available or channelized out of the TDRSS WSTF. For the options available, [12-4, Section 3.3.3] should be consulted.

TDRSS S-band return link data modes are summarized in Table 12-13. An example return link MA service link budget is shown in Table 12-14.

12.3.8 Ranging and Doppler

The TDRSS provides one-way doppler measurements and two-way range and doppler measurements. One-way doppler measurements are available on return channels for all mode and/or data group configurations. Two-way doppler measurements are available to all user spacecraft for which the coherent transmit/receive carrier frequency turnaround ratio is 240/221 for MA and SSA and 1600/1469 for KSA. Two-way range measurements are available to user spacecraft operating DG1 Modes 1 and 3 when the user transmitted PN code is time-synchronized to the PN code received from the TDRSS.

The TDRSS will provide the following simultaneous tracking services on a continuous basis:

(1) One-way doppler measurement for any 10 return links.

(2) Two-way range and doppler measurement for three MA links (includes the in-orbit spare satellite).

Table 12-13. Summary of TDRSS S-band return link data modes

	Data Group 1 (DG1)			Comments
	Mode 1	Mode 2	Mode 3	
Type of service	Multiple access (MA)	Multiple access (MA)	MA requires SSA service	SSA service is ground interface
Return link carrier	Coherent: $250/221 \times f_R$	Noncoherent spacecraft oscillator	Coherent: $240/221 \times f_R$	f_R = Forward link, receiver phase-locked
Code source:			1 data channel / 2 data channels	
I channel	Stored PN	Stored PN	Stored PN / Data	
Q channel	Stored PN	Stored PN	Data / Data	
Ranging				Code epoch sync.: All "1" state on return link synchronized with all "1" state on forward link
I channel	PN code epoch sync	Spacecraft oscillator	PN code epoch sync	
Q channel	Epoch synch + (x + 1/2 PN chip)	Epoch sync + (1/2 PN chip)	(Used for high-speed data)	
Data modulation	Modulo-2 added asynchronously	Module-2 added asynchronously	Modulo-2 added asynchronously	SQPN = staggered quad-riphase pseudo-noise (PN)
I channel	SQPN	SQPN	SQPN	
Q channel	SQPN	SQPN	PSK $\pm \pi/2$ rad	
Data rate restrictions (MA)	Convolutional coding	Convolutional coding	Convolutional coding	Data on I and Q channels can be independent and asynchronous; reduce rates by 2 for biphase format
I channel	0.1 - 50 kbps	1.0 - 50 kbps	0.1 - 150 kbps	
Q channel	0.1 - 300 kbps	1.0 - 50 kbps	1.0 kbps - 1.5 Mbps	
Total	0.1 - 600 kbps	1.0 - 50 kbps	I + Q	
Data rate restrictions (SSA)				For SSA service, the minimum spread bandwidth shall produce a maximum power flux density at the earth's surface which does not exceed -151 dBw/m^2 in any 4-kHz bandwidth for angles from 0 to 5 deg above horizontal
I channel	0.1 - 50 kbps	0.1 - 300 kbps	0.1 - 300 kbps	
Q channel	0.1 - 50 kbps	0.1 - 300 kbps	1.0 kbps - 6 Mbps	
Total	0.1 - 300 kbps	0.1 - 600 kbps	I + Q	
PN code chip rate (MA and SSA)	$\frac{31}{(240 \times 96)} \times f_1$	$\frac{31}{(240 \times 96)} \times f_1$	$\frac{31}{(240 \times 96)} \times f_1$	f_1 = transmit (return link) frequency
PN code length	$(2^{10} - 1) \times 256$	$2^{11} - 1$	$(2^{10} - 1) \times 256$	All PN codes stored in transponder

Table 12-14. Typical return link MA service link budget

BER	10^{-5}
User EIRP, dBW	EIRP
Space loss, dB	-192.2
Polarization loss, dB	-1.0
TDRS antenna gain at ±13°, dB	28.0
P_s at output of antenna, dBW	-165.2 + EIRP
T_s (antenna output terminals), K	824
T_i (due to direct other user interference), K	255
$K(T_s + T_i)$, dBW/Hz	-198.3
$P_s/K(T_s + T_i)$, dB-Hz	33.1 + EIRP
Transponder loss, dB	-2.0
Demodulation loss, dB	-1.5
PN loss, dB	-1.0
Antenna beam forming loss, dB	-0.5
User margin, dB	-3.0
Required E_b/N_0, dB (ΔPSK)	-9.9
FEC gain, R = 1/2, K = 7, dB	5.2
*Achievable data rate, dB	20.4 + EIRP

*This achievable data rate is the user's information rate. It should not be confused with the channel symbol rate, which is twice the information rate.

(3) Two-way range and doppler measurement for six SA links.

Figure 12-10 shows the general geometry involved between a user satellite and the two TDRSS operational satellites and the WSTF. Range and range rate are determined by measuring the time required for signals to propagate through the closed loop from the WSTF, through a TDRSS satellite, turned around by the user satellite's transponder, and then back through the TDRSS satellite to the WSTF. The TDRSS can also measure range and range rate using a loop from the WSTF through one TDRSS satellite to the user and back through the other TDRSS satellite to the WSTF. The latter mode requires the user satellite to see both TDRSS satellites (a user satellite antenna coverage problem), but will generally provide improved position accuracy over the one TDRSS satellite mode.

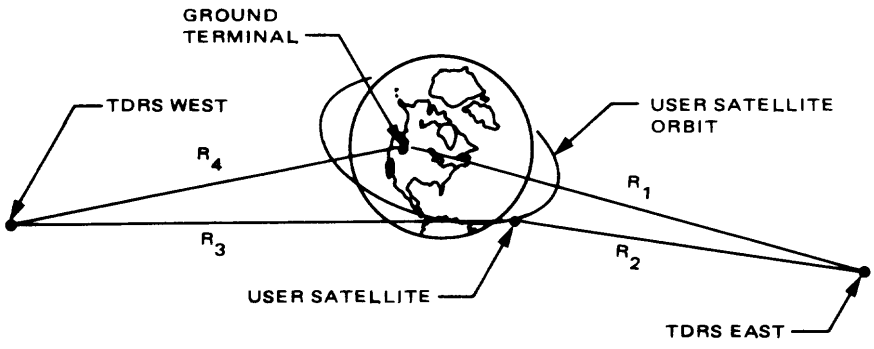


Fig. 12-10. General tracking service geometry

References

12-1. Springett, J. C., *Space Shuttle Telecommunication Interfaces*, Document 890-123, Jet Propulsion Laboratory, Pasadena, Calif., Mar. 15, 1981 (an internal document).

12-2. "Special Issue on Space Shuttle Communications and Tracking," *IEEE Transactions on Communications*, Vol. COM-26, No. 11, Nov. 1978.

12-3. Springett, J. C., *Tracking and Data Relay Satellite System Interfaces*, Document 890-124, Jet Propulsion Laboratory, Pasadena, Calif., Apr. 15, 1981 (an internal document).

12-4. *Tracking and Data Relay Satellite System (TDRSS) User's Guide*, Document STDN No. 101.2, Goddard Space Flight Center, Greenbelt, Md., Jan. 1980.

12-5. *Performance Specification for Services via the Tracking and Data Relay Satellite System*, Document S-805-1, Change 1, Goddard Space Flight Center, Greenbelt, Md., May 1979.

12-6. Holmes, W. M., Jr., "NASA's Tracking and Data Relay Satellite System," *IEEE Communications Society Magazine*, Sept. 1978.

12-7. *PN Codes for Use with the Tracking and Data Relay Satellite System (TDRSS)*, Document STDN No. 108, Goddard Space Flight Center, Greenbelt, Md., Dec. 1976.

Index

☆ U. S. GOVERNMENT PRINTING OFFICE: 1983—684-338