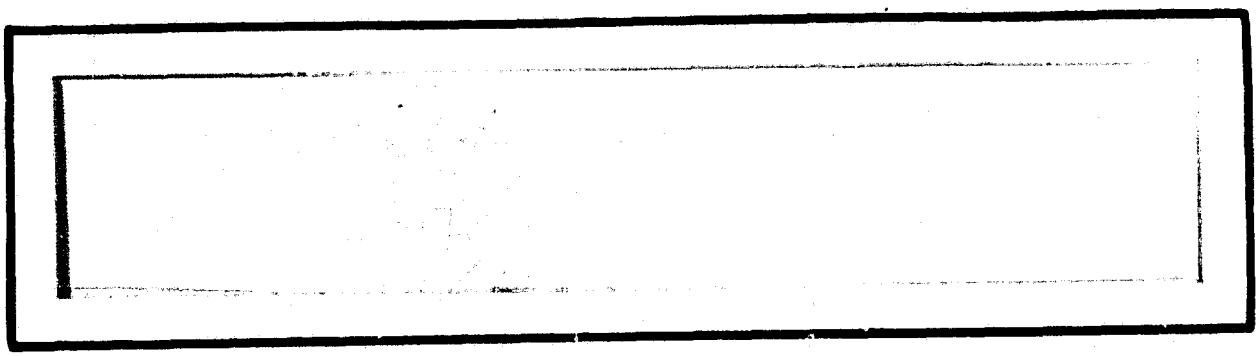
NASA-CR-167786



PAYLOAD/SHUTTLE-DATA-COMMUNICATION-LINK HANDBOOK

Contract No. NAS 9-16067

Exhibit C

Technical Monitor: William Teasdale

Prepared for

NASA Lyndon B. Johnson Space Center
Houston, Texas 77058

Prepared by

Axiomatix
9841 Airport Blvd., Suite 912
Los Angeles, California 90045

Axiomatix Report No. R8208-3
August 30, 1982

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

1.1 Purpose of Handbook

A user of the Shuttle Orbiter system will need to communicate with both the Orbiter and the ground. To implement such communication, the Orbiter contains a versatile set of payload-oriented avionic hardware and provides several communication links to various ground stations. A user wishing to communicate may make use of the Shuttle communication systems in either a standard or nonstandard manner. Standard accommodations will usually meet the majority of user requirements with maximum flexibility and reliability and with minimum concern and cost. Nonstandard capabilities, however, are provided so that 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.

It is the purpose of this handbook to provide the user with technical performance information that will allow the user to make his own assessment of which aspect of the Shuttle communication capability can best be employed to satisfy his requirements. Thus, the handbook outlines the various subsystem configurations and performance parameters and develops a rationale which the user may use to calculate the operation of his particular application. Based on this background and the current designs for the Shuttle communication equipment and links, a number of typical link performance tables are constructed as guidance examples.

1.2 Definition of a User Payload

A user payload may be 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 spacecraft are the payloads with which this handbook is primarily concerned.

Payloads may be divided into two distinct classes: (1) those which will be separated or become "detached" from the Orbiter and (2) those which will remain "attached" to the Shuttle in the associative surroundings of the payload bay. Many detached payloads will be transported into geosynchronous or other Earth orbits or placed on deep-space

trajectories by the Inertial Upper Stage (IUS). 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 RF channels.

1.3 Overview of the Payload/Orbiter Interface

In the user's planning for Shuttle operations, the key words are "standard" and "adaptable." Standard plans and equipment using standardized interfaces should be the rule. With the standard communication capabilities, the user can select from among several options in equipment, thereby tailoring a flight to his own needs. Users are therefore encouraged to design payloads that are compatible with the standard communication links.

The two largest user agencies of the Shuttle as a payload launcher will be NASA and DOD. Other users will be entities such as COMSAT, private industry, and foreign countries. NASA and DOD payload requirements and subsystem capabilities have been predominantly responsible for the design of the avionic subsystems, especially in terms of the detached payload communication links. Thus, "standard" capabilities have evolved to serve NASA and DOD. Nonstandard conditions have also been provided for, but generally with less operational capability, especially aboard the Orbiter.

The Orbiter communications and tracking subsystem provides links between the Orbiter and the payload, It also transfers payload telemetry and uplink data commands to and from the space networks. Most payload communications, tracking and data management can be accommodated by this Orbiter subsystem since provisions have been incorporated to maintain a large degree of flexibility.

The Orbiter can communicate with the ground stations directly or through the Tracking and Data Relay Satellite System (TDRSS). Payloads can communicate with the Orbiter through hardline cables for attached payloads or through the payload radio frequency (RF) link for detached payloads. Table 1.1 lists the major unmanned payload functions and the communication links over which they are handled and Figure 1.1 portrays the various communication links.

Table 1.1. Orbiter Avionics Services to Payloads

Function	Payload/Ground Direct or Through Tracking and Data Relay Satellite		Payload/Orbiter Hardline		Payload/Orbiter Radio Frequency 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 Attitude Control		X	X	X	X	
Caution & Warning	X			X		X
Master Timing			X			
Uplink Data		X	X			

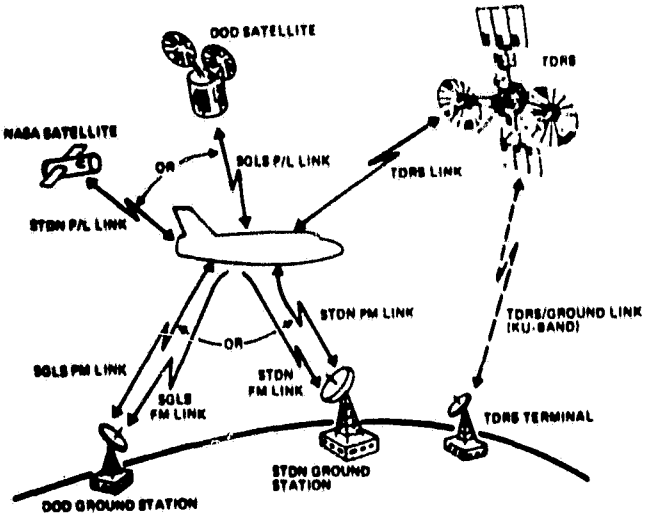


Figure 1.1. Space Shuttle Orbiter Communication Links

The data processing and software subsystem of the Orbiter furnishes the onboard digital computation necessary to support payload management and handling. The stations in the Orbiter aft flight deck for payload management and handling are equipped with data displays, cathode ray tubes (CRT) and keyboards for onboard monitoring and control of payload operations.

1.4 Scope of the Handbook

Subsequent sections of this handbook are designed to provide the user with (1) general information regarding the nature, use and restrictions of the various communication links and, (2) specific data on the characteristics and parameters which typify the links. Communication link models are developed, along with their link design budgets, and a number of representative design control tables are presented. This handbook concludes with some guidelines for devising and interpreting the overall communication link capability.

2.0 GENERAL DESCRIPTION OF THE AVAILABLE COMMUNICATION LINKS

2.1 Overview and Typical End-to-End Links

Figure 2.1 shows a pictorial representation of the Shuttle and the various RF channels which comprise the communication links between the Shuttle, payloads, and ground. The links between the Shuttle and payloads are at S-band, as are the Shuttle/ground direct links. The forward frequency from the Shuttle to the DOD/SGLS is L-band. Relay links through the TDRS are at both S-band and Ku-band.

Communication can generally be established with only one detached payload at any time. Similarly, only one coherent Shuttle/ground direct link is available. Since, however, the FM direct link utilizes separate equipment, it may be used simultaneously with the coherent direct link. The TDRSS relay can make use of both 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, but nearly continuous contact may be maintained using the TDRS.

2.1.1 Equipment Definitions

In the ensuing discussions, various pieces of functional Orbiter avionic equipment, called subsystems, will be denoted. These subsystems are listed in Table 2.1 by name (and acronym), and their principal internal functions are indicated. A narrative description and functional block diagram of most of these subsystems may be found in Appendix C.

2.1.2 Typical End-to-End Link Configurations

In this subsection, some typical end-to-end communication link configurations will be outlined. Although these configurations are by no means exhaustive, they do represent the majority of forms and generally portray the nature of the various forward and return data/signal links.

2.1.2.1 Detached payload standard telemetry S-band direct link

Figure 2.2 shows a block diagram of the detached payload standard telemetry S-band direct link. Standard telemetry for NASA and DOD payloads involves the transmission of digitally encoded data at specified bit rates. Within the payload itself, digital data must be modulated

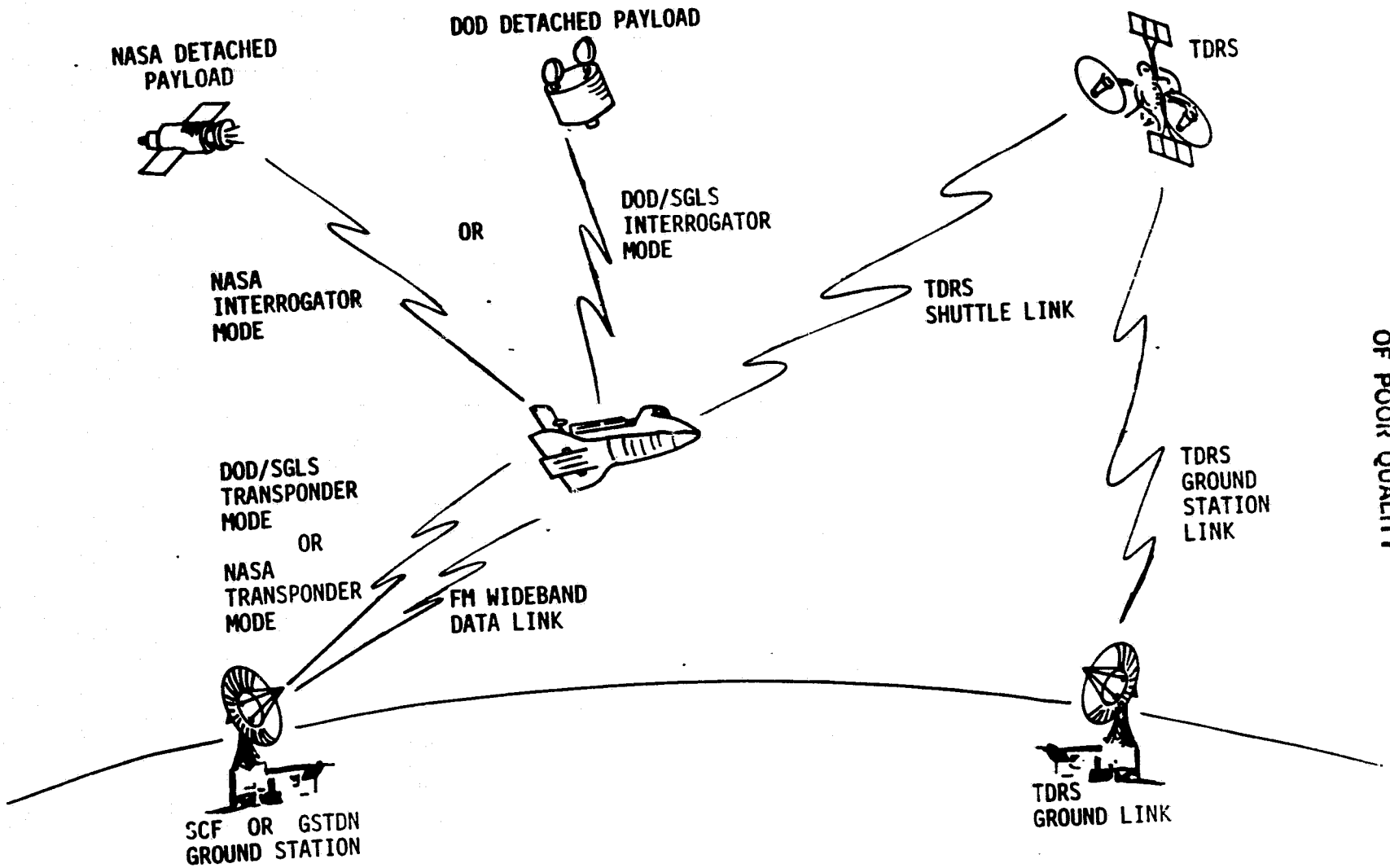


Figure 2.1. Payload Communication Links Overview

Table 2.1. Orbiter Avionics Subsystems and Functions

Subsystem Name	Acronym	INTERNAL FUNCTIONS									
		Carrier Modulation	Carrier Tracking & Demodulation	Subcarrier Demodulation	Data Synchronization & Detection	Command Modulation Generation	Data Stream Multiplexing	Data Stream Demultiplexing	Data Validity Check	Data Buffering & Rate Change	Data Decoding
Payload Interrogator	PI	X	X								
NASA Payload Signal Processor	PSP			X	X	X			X	X	X
DOD Communication Interface Unit	CIU			X	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 & Receiver	-	X	X								
Ku-Band Signal Processor	KuSP				X		X	X		X	X
Multiplexer/ Demultiplexer	MDM						X	X	X		X
General-Purpose Computer	GPC										

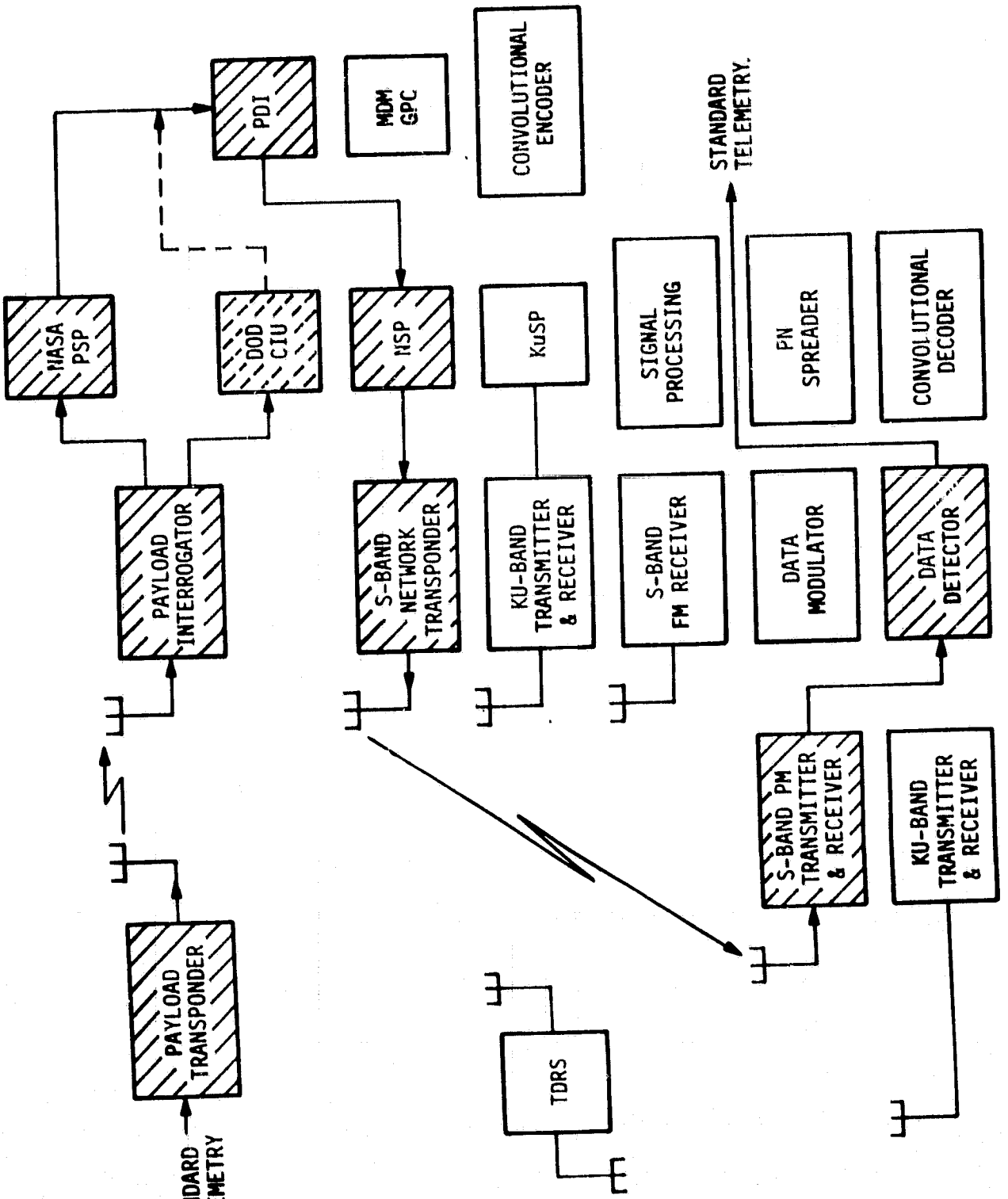


Figure 2.2. Detached Payload Standard Telemetry, S-Band Direct Link

onto subcarriers of specified frequency with an NRZ-type format. Table 2.2 summarizes the standard digital telemetry requirements. In addition, for DOD payloads, certain FM/FM analog telemetry on a subcarrier is allowed, as indicated in Table 2.2.

Table 2.2. Payload Standard Telemetry Requirements

Parameter	Parameter/Range	
	PSK Modulation	Frequency Modulation
Subcarrier Frequencies	1.024 MHz or 1.7* MHz	1.7* MHz
Bit Rates or Modulation Response	256,*† 128,*† 64,* 32,* 16, 10, 8, 4, 2, 1, 0.5,* 0.25* kbps	100 Hz to 200 kHz

*DOD only

†1.7 MHz subcarrier only

The standard telemetry is transmitted via the payload transponder and received and demodulated aboard the Orbiter in the Payload Interrogator (PI). For NASA payloads, the Payload Signal Processor (PSP) demodulates the subcarrier and detects the data while, for DOD payloads, the Communications Interface Unit (CIU) performs the comparable function. (Note that, in Figure 2.2, the DOD CIU is shown as an alternate path to the NASA PSP; these subsystems do not operate simultaneously on payload signals.)

The Payload Data Interleaver (PDI) and Network Signal Processor (NSP) function to multiplex the detected detached payload data from the PSP (or CIU) 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 (not shown functionally in Figure 2.2) so that the standard telemetry data stream, as it appeared at the input to the payload transponder, is delivered to the appropriate payload/user facility. Because of the noisy detection operations that take place in the PSP (or CIU) and in the ground data detector, some bits of information in the telemetry data stream are in error.

2.1.2.2 Detached payload standard telemetry S-band relay link

This configuration is nearly identical to the direct link (see subsection 2.1.2.1) except that the Tracking and Data Relay Satellite (TDRS) is used as an intermediate channel between the Orbiter and the ground. The link is shown in Figure 2.3.

(It should be noted that a different operating mode of the network transponder and an entirely different ground station and set of ground equipment is used with the TDRS link than with the direct link. Thus, the block diagrams employed in this handbook should be understood as generic to the end-to-end links and that specific flight hardware modes and ground equipment and configurations, apart from the indicated generalized functions, are not implied.)

2.1.2.3 Attached payload FM signal link

Figure 2.4 indicates the configuration for the attached payload S-band FM signal link. This system and link are provided for the transmission of payload wideband analog signals, television, and Shuttle main engine data.

The FM signal processor appropriately conditions the analog signal inputs, frequency modulates them as needed onto subcarriers, and forms the composite FM signal which is input to the FM transmitter. An S-band direct link to the ground is employed. The composite FM signal is received, carrier demodulated, and then signal processed to demodulate the subcarriers and filter the various analog telemetry signals. If digital data has been frequency-shift-keyed onto the carrier or a subcarrier, the signal processor must function to bit synchronize and detect the data after demodulation.

2.1.2.4 Detached payload nonstandard telemetry bent-pipe links

The standard telemetry data capability available for detached payloads (see subsection 2.1.2.1) provides for a reasonable degree of flexible operation. It may happen, however, that certain payloads are not able to avail themselves of the standard system. To accommodate payloads whose telemetry formats are not compatible with standard data rates and subcarrier frequencies, "bent-pipe" modes of operation are provided

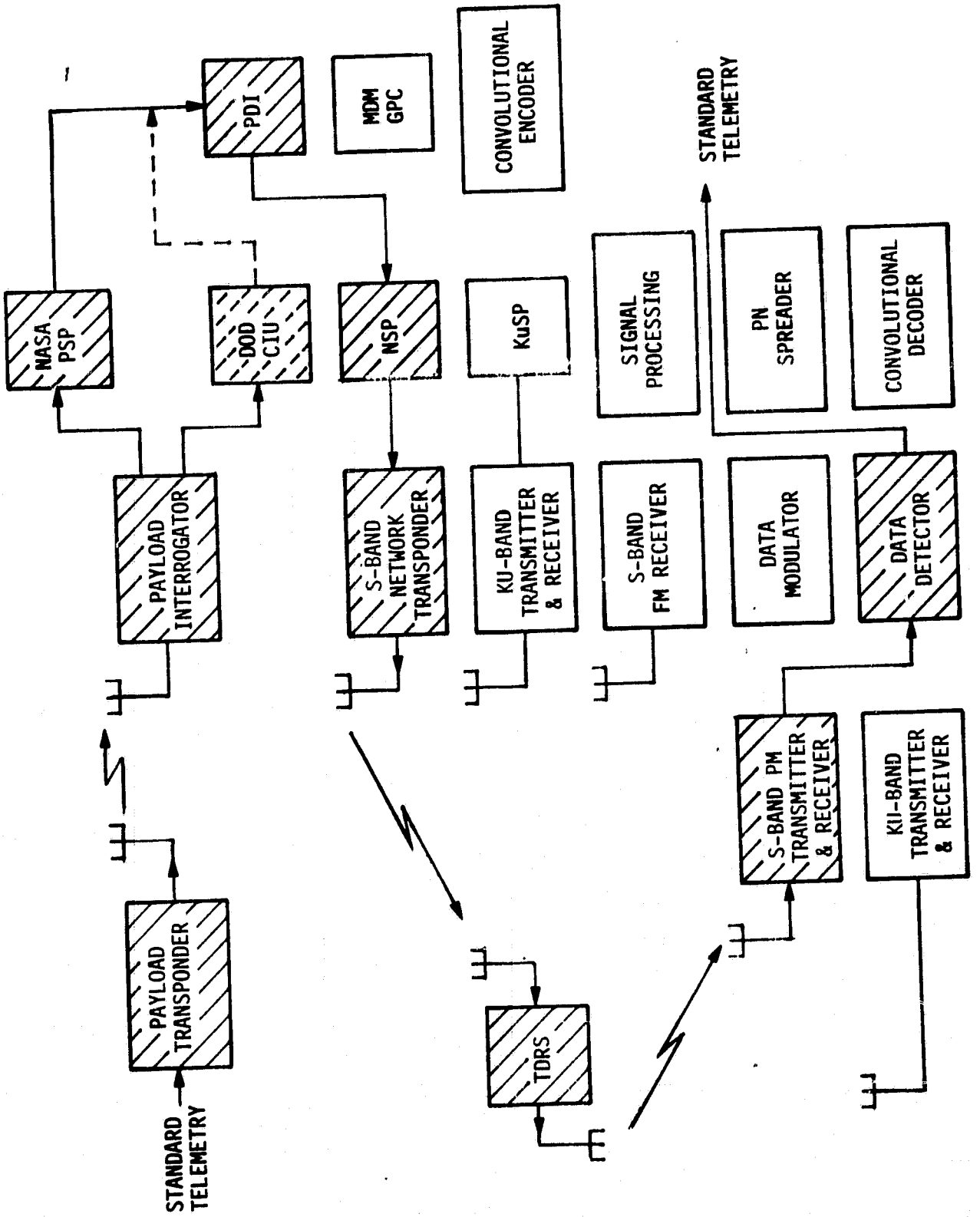


Figure 2.3. Detached Payload Standard Telemetry, S-Band Relay Link

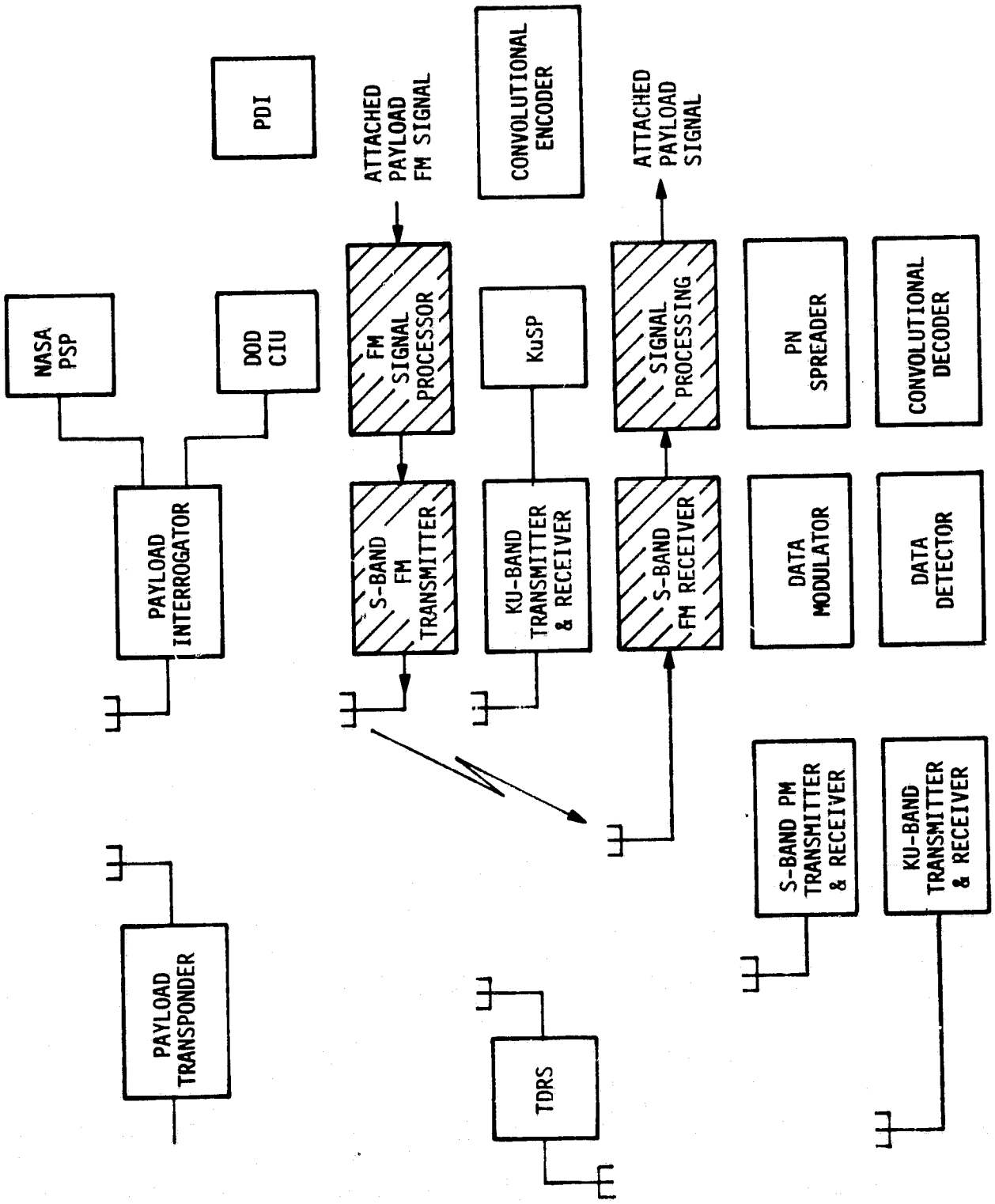


Figure 2.4. Attached Payload FM Signal Link

within the Shuttle's avionic equipment. Several signal paths acting as "transparent throughputs" are available for both digital and analog signals.

Digital data streams at rates higher than 64 kbps (which therefore cannot be handled by the PDI) may directly enter the Ku-band Signal Processor (KuSP), where they may be: (1) QPSK modulated onto an 8.5-MHz 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.

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 hard limited (i.e., a two-level or one-bit quantized waveform produced) and treated as "digital" signals by the 8.5 MHz subcarrier QPSK modulator. 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.

Figure 2.5 shows the subsystems that would be employed in an FM bent-pipe link.

2.1.2.5 Attached payload high-data-rate Ku-band relay link

Attached payloads may have digital data rate requirements that exceed the standard telemetry capacity. Thus, for data rates up to 50 Mbps, a high data rate link utilizing the Ku-band TDRS channel is available. The end-to-end configuration is shown in Figure 2.6.

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.

2.1.2.6 Detached payload command S-band relay link

Commands from the ground to detached payloads may be transmitted from the ground to the Orbiter by any one of three links: (1) S-band direct, (2) S-band TDRS relay or, (3) Ku-band TDRS relay. Regardless of which link is used, detected command data aboard the Orbiter is thoroughly checked for validity and errors before it is transmitted to the payload.

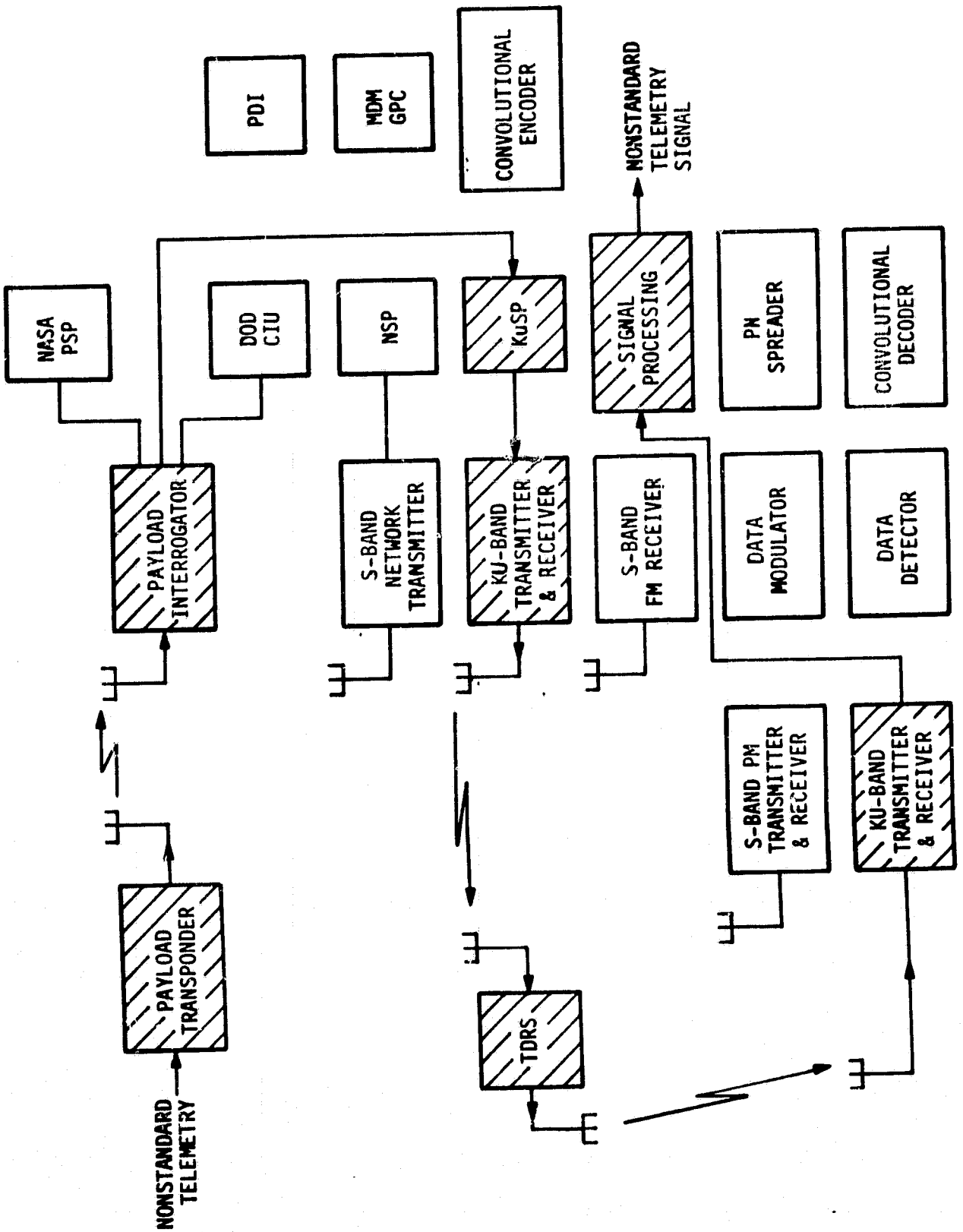


Figure 2.5. Detached Payload Nonstandard Telemetry Bent-Pipe Link

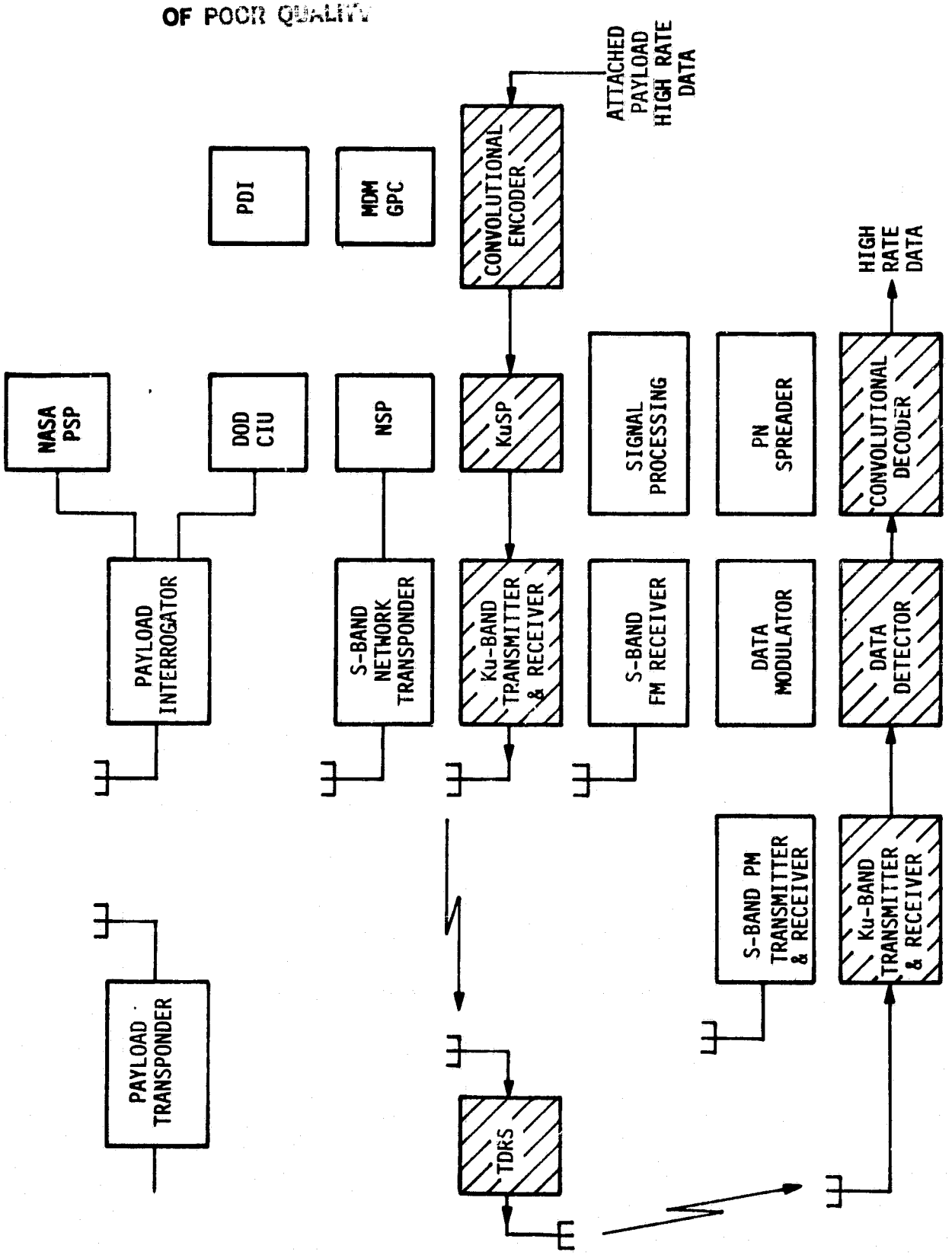


Figure 2.6. Attached Payload High Data Rate Ku-Band Link

Figure 2.7 shows the end-to-end subsystems employed in an S-band relay command link. Encoded (i.e., structured) payload command data bits at the ground station are multiplexed with Orbiter commands and data and PN code modulated in order to spread the data frequency spectrum. (This is a requirement of the TDRS 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 (or DOD CIU as the alternate path), where it is transformed into the proper payload subcarrier signal structure. Transmitted to the payload via the PI, then received, demodulated and detected by the payload's transponder, the command data is sent to the payload's decoder (not shown in Figure 2.7) for final decoding and disposition.

Commands to detached payloads are always in standard form; there is no nonstandard command equivalent to the nonstandard telemetry capability. Table 2.3 indicates the standard command conditions.

Table 2.3 Command System Parameters

NASA		
Subcarrier Frequency	-	16 kHz, sine wave
Bit Rates	-	$2000 \div 2^N$ bps (N = 0,1,2,...,8)
DOD		
Signal Tone Frequencies	-	65 kHz, 76 kHz, 95 kHz
Symbol Rates	-	1000 or 2000 symbols/second

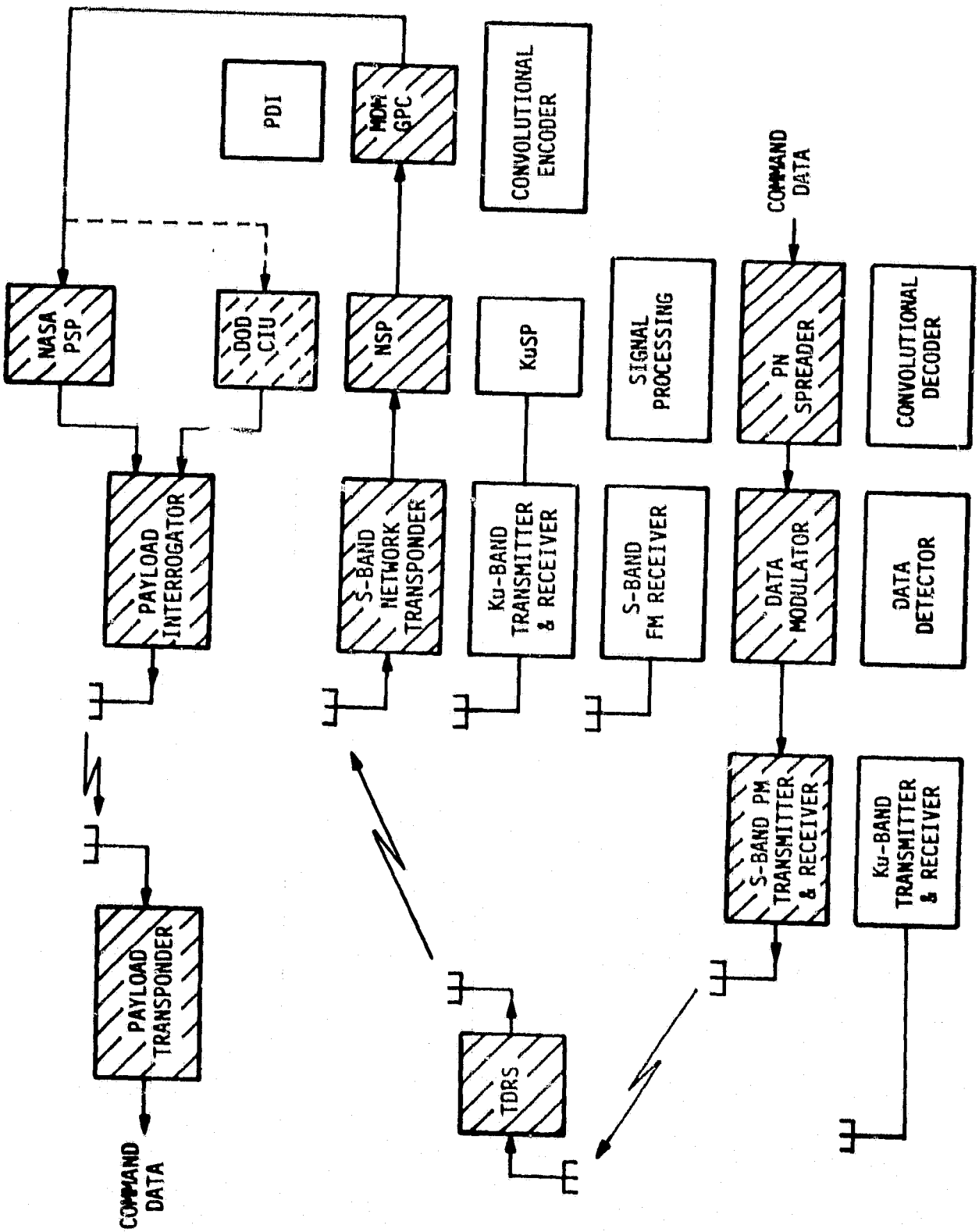


Figure 2.7. Detached Payload Command S-Band Relay Link

2.1.2.7 Attached payload high-data-rate Ku-band forward relay link

Certain attached payloads may require forward link (ground-to-payload) data rates in excess of those possible with the standard command capability. For such applications, the general link shown in Figure 2.8 is available to handle data rates up to 128 kbps.

The following subsections provide a technically oriented summary of the various links which comprise the end-to-end payload/ground communication system. Block diagrams show the principal signal-processing functions which are important to the structuring of the communication link design budget discussed in Section 4. Some of the link parameters are also introduced on the block diagrams; these and others are given substantive definition in Section 3.

2.2 Payload/Orbiter Link

2.2.1 Hardline Interfaces

2.2.1.1 Standard capabilities

Attached payloads communicate with the Orbiter over hardline umbilicals. Standard telemetry and command capability are identical to that for detached payloads (see Tables 2.2 and 2.3). Commands may be transferred from the Orbiter to the payload in either a direct data or subcarrier signal mode. Telemetry-modulated subcarriers may not be transmitted over the hardline.

Generally, the standard hardline interfaces by themselves are not governed by a link design budget, as all signals are very large compared with any ambient additive noise.

2.2.1.2 Nonstandard characteristics and restrictions

Nonstandard attached payload communications take place over hard lines. Nonstandard telemetry may involve a digital data stream up to 50 Mbps and/or analog signals having baseband bandwidths up to 4 MHz. Digital data rates above 5 Mbps are transmitted to the ground via the Ku-band relay link, while either digital data rates less than 5 Mbps or analog signals may be accommodated by the S-band FM direct link. Analog signals may also be handled by the Ku-band bent-pipe link. Digital data from the Orbiter to the payload may be transferred at a rate of 128 kbps over the hardline. No provision is made for Orbiter-to-payload analog signals.

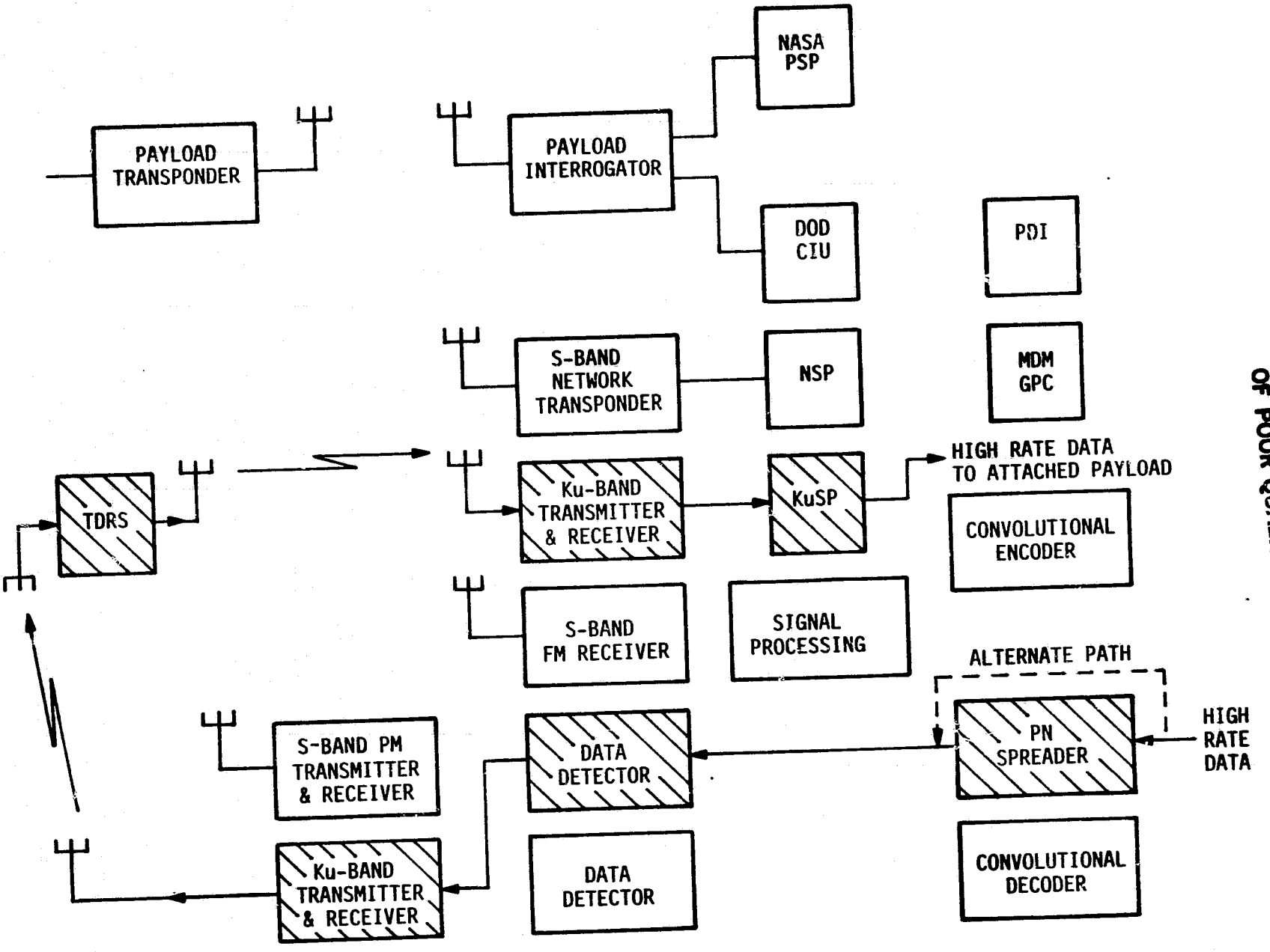


Figure 2.8. Attached Payload High-Data-Rate Ku-Band Forward Relay Link

Nonstandard hardline interfaces by themselves are not governed by a link design budget.

2.2.2 S-Band RF Link

2.2.2.1 Standard capabilities

Figure 2.9 shows the payload-to-Orbiter link and Figure 2.10 is the corresponding Orbiter-to-payload channel. These links are complete in the sense that data or signals input at one end are depicted as recovered at the opposite end. In the case of digital data streams, the output of the link is a two-state digital data sequence (with the possible exception of the bent-pipe signal) with bit errors.

The general symbology employed throughout these subsections is that P_X represents power, G_X antenna gain, L_S space loss, θ_X modulation index, (S+N) denotes a noisy signal, $(S/N)_X$ is a signal-to-noise ratio, and E_B/N_0 is a digital data detection bit-energy to noise-spectral-density ratio. Error probability is designated by the symbol P_E. Additionally, certain functions are internally characterized by internal noise sources (N_0), by two-sided tracking bandwidth $(2B_L)$ and by other bandwidth and signal component loss parameters not specifically shown on the block diagrams. Refer to Section 3 for a complete description of the link parameters and symbols and to Section 4 for their identification with the design control tables.

Both the forward and return payload/Orbiter links are phase coherent. Standard telemetry at rates shown in Table 2.2 are biphase modulated onto the appropriate subcarriers. The subcarriers, in turn, phase modulate the carrier at a NASA standard peak deviation of 1.0 rad or DOD standard peak deviations of 0.3 or 1.0 rad. In the PI receiver, the residual carrier component of the received signal is tracked by a phase-locked-loop (PLL) which establishes the coherent reference for demodulating the subcarriers from the carrier. The recovered subcarrier is in turn processed by the NASA PSP (or DOD CIU), where it is coherently demodulated and the coherent reference is established by tracking the subcarrier average phase in a Costas loop. Bit synchronization and detection then follow, with hard decisions being made as to bit polarity (state) on a bit-by-bit basis.

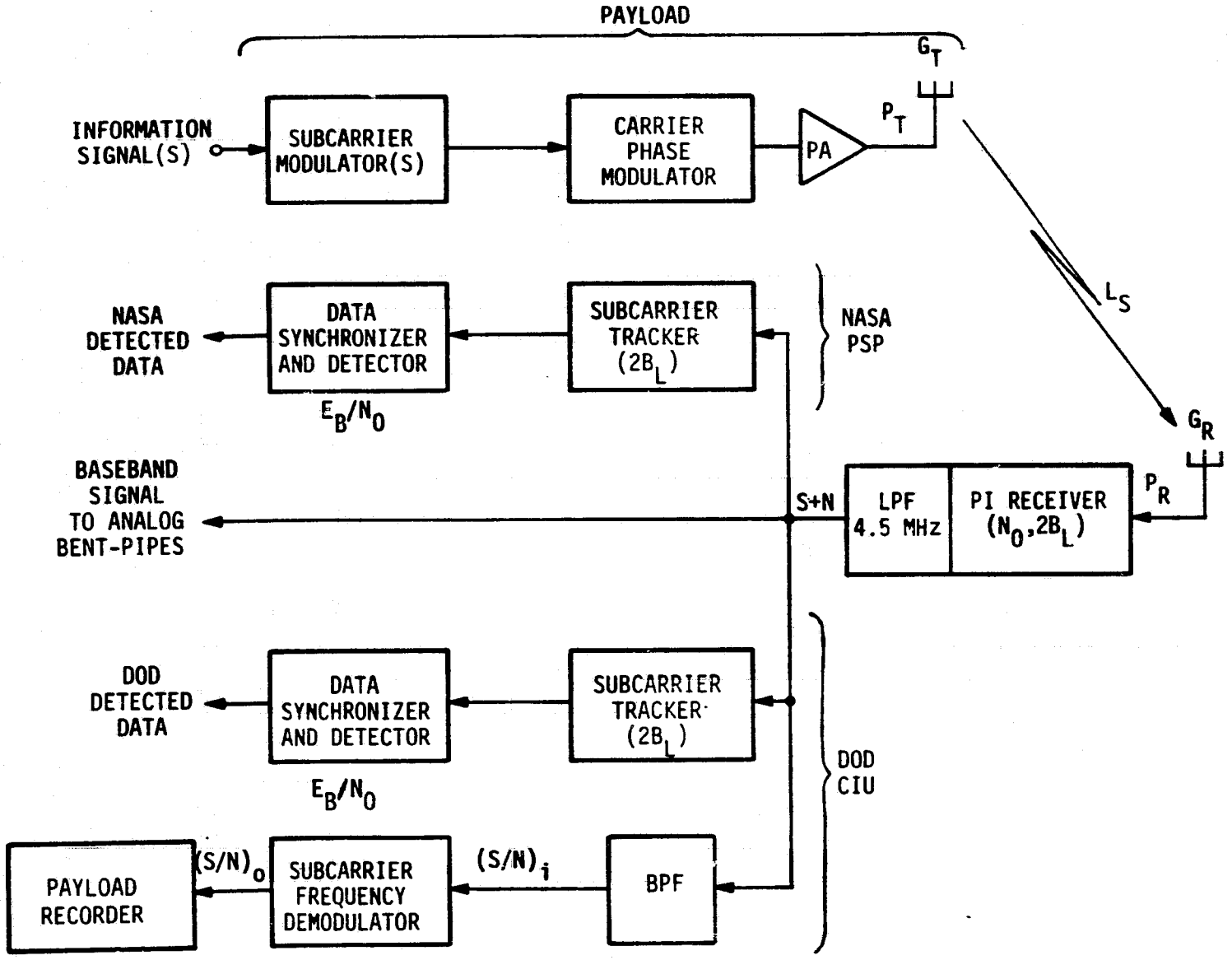


Figure 2.9. Payload-to-Orbiter S-Band Link

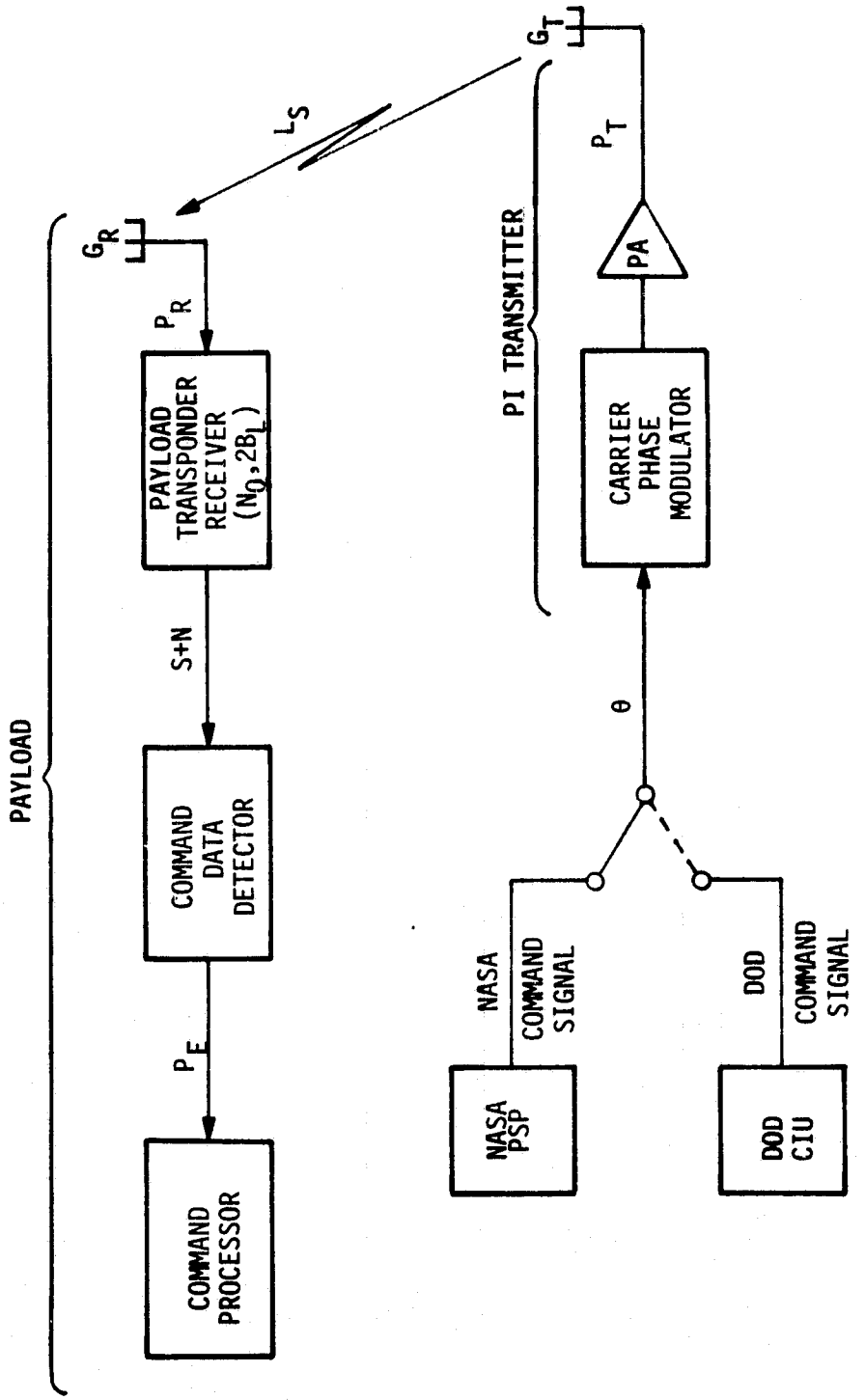


Figure 2.10. Orbiter-to-Payload S-Band or L-Band Link

In the DOD return link, analog telemetry may be transmitted and received using FM on the 1.7 MHz subcarrier. Within the CIU, the subcarrier is demodulated and filtered, and the composite analog signal is input to the Payload Recorder (PR).

The standard command forward link consists of phase modulating the PI transmitter by either the NASA or DOD command signal. Peak carrier phase deviation is 1.0 rad for the NASA command signal and either 0.3 or 1.0 rad for the DOD command signal. The NASA command signal consists of a 16-kHz subcarrier biphase-modulated by the command bits, and the DOD command signal is a set of three frequency-shift-keyed tones with the composite tone waveform amplitude modulated by a triangular command symbol timing function. Within the payload, the transponder receiver phase tracks the residual carrier component of the received signal and coherently demodulates the carrier. The NASA subcarrier or DOD tone set is then processed by the command data detector to demodulate and detect the command information. Detected commands are characterized by an error probability which is increasingly small. Decoding of the command follows, wherein the appropriate action aboard the payload takes place.

More detailed information regarding operation of the various forward and return link subsystems may be found in [1]. Detailed link parameter information is listed in Appendices A1, A2, A3, A8, and A9, and some typical standard link design control tables may be found in Appendices B1, B2 and B3.

2.2.2.2 Nonstandard characteristics and restrictions

Nonstandard telemetry forms (e.g., data rate, coding) and modulations (e.g., nonstandard subcarrier frequencies, direct carrier modulation by data) are allowed on a return link, provided they are compatible with the PI. Specifically, they must have a residual carrier component and sideband characteristics that do not promote PI receiver false lock or in any way compromise PI operation. Appendix D sets forth a complete set of nonstandard payload modulation restrictions. It must also be understood that, beyond acting as a bent pipe to nonstandard modulations, the Orbiter avionic subsystems cannot be used to demodulate, detect, decode, or otherwise process such signals. Design control for bent-pipe channels, beyond those portions of the link which are specified by the overall Orbiter/ground communication system, is the responsibility of the nonstandard user.

Nonstandard command signals to detached payloads are not possible because an external modulation port to the PI transmitter is not available. Only the PSP- or CIU-generated command signals may be transmitted to detached payloads.

2.3 Shuttle/Ground Links

2.3.1 S-Band Links

2.3.1.1 S-band PM direct link

Figure 2.11 shows the principal functions that concern data transfer from the Orbiter to the ground utilizing the S-band Ground Satellite and Tracking Data Network (GSTDN) PM direct link. Figure 2.12 depicts the corresponding ground-to-Orbiter channel.

These links are very similar in function to the payload/Orbiter links discussed in subsection 2.2.2.1 in that they employ a residual carrier for tracking and coherence and have established data formats and rates. (See subsection 2.2.2.1 and Section 3 for definitions of the symbols used.) Typical operating parameters are listed in Appendices A2, A3 and A4.

Two particular features of note to payloads are that (1) payload data (telemetry and commands) is multiplexed with other Orbiter data and (2) a turnaround ranging signal may be present. The presence of ranging is important only in that it receives a portion of the total transmitter power and must be taken into account when constructing the payload data portion of the link design control table. It is also noted that data subcarriers are not used by these links. A typical Orbiter-to-ground design control table appears in Appendix B4. The corresponding ground-to-Orbiter design control table may be constructed by analogy.

2.3.1.2 S-band PM TDRSS relay links

Shown in Figure 2.13 are the functional elements of the Orbiter-to-ground S-band TDRSS relay link. By comparison with the S-band direct link of subsection 2.3.1.1, this channel (1) employs convolutionally encoded data, (2) uses suppressed carrier PSK modulation with no residual carrier, and (3) is characterized by two RF links.

The two RF links complicate the design control procedure because of the presence of two independent noise sources (one in the TDRS receiver and the other in the ground station receiver) that are separated by intermediate signal-processing functions. Especially important is that the

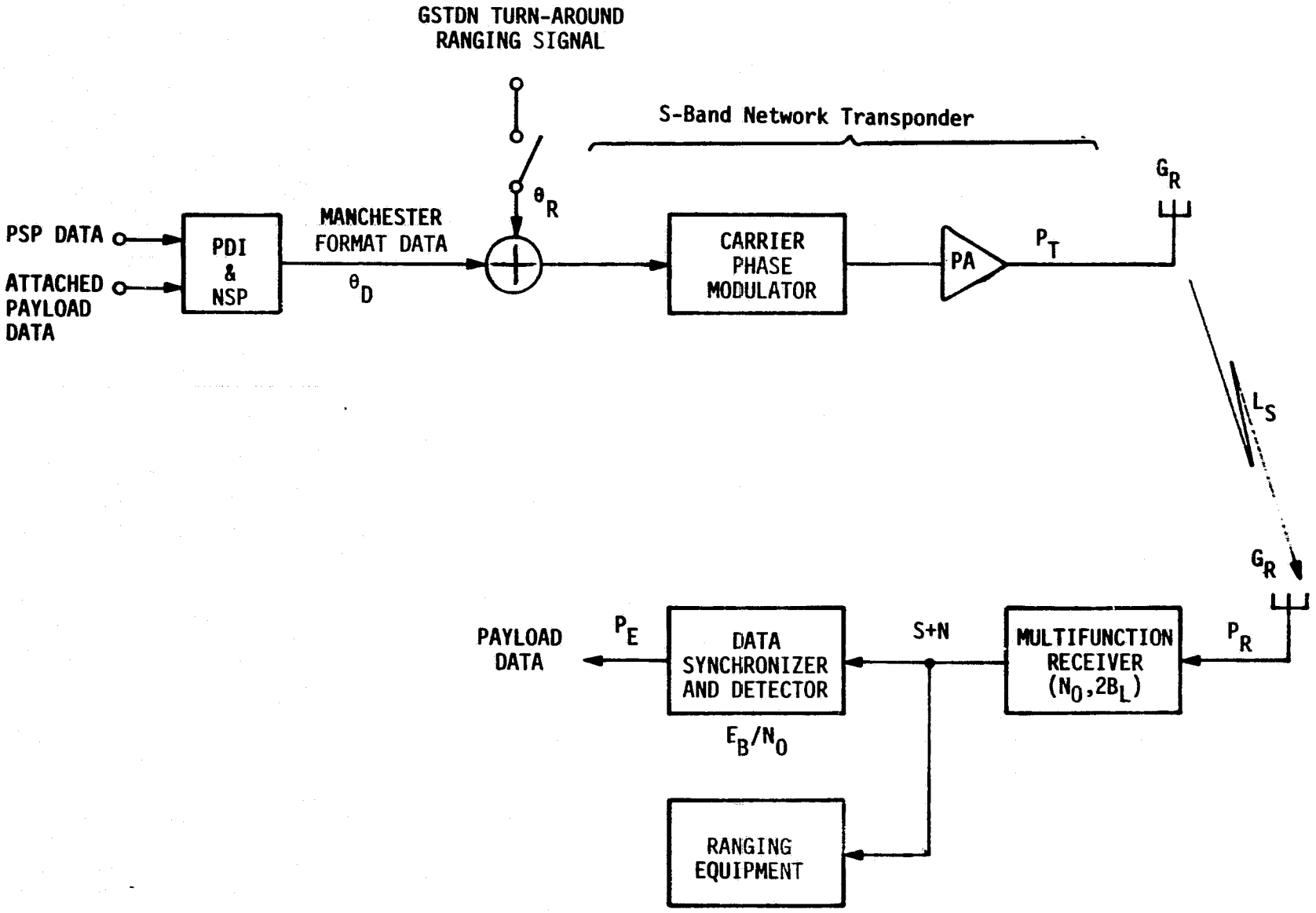


Figure 2.11. Orbiter-to-Ground S-Band GSTDN PM Link

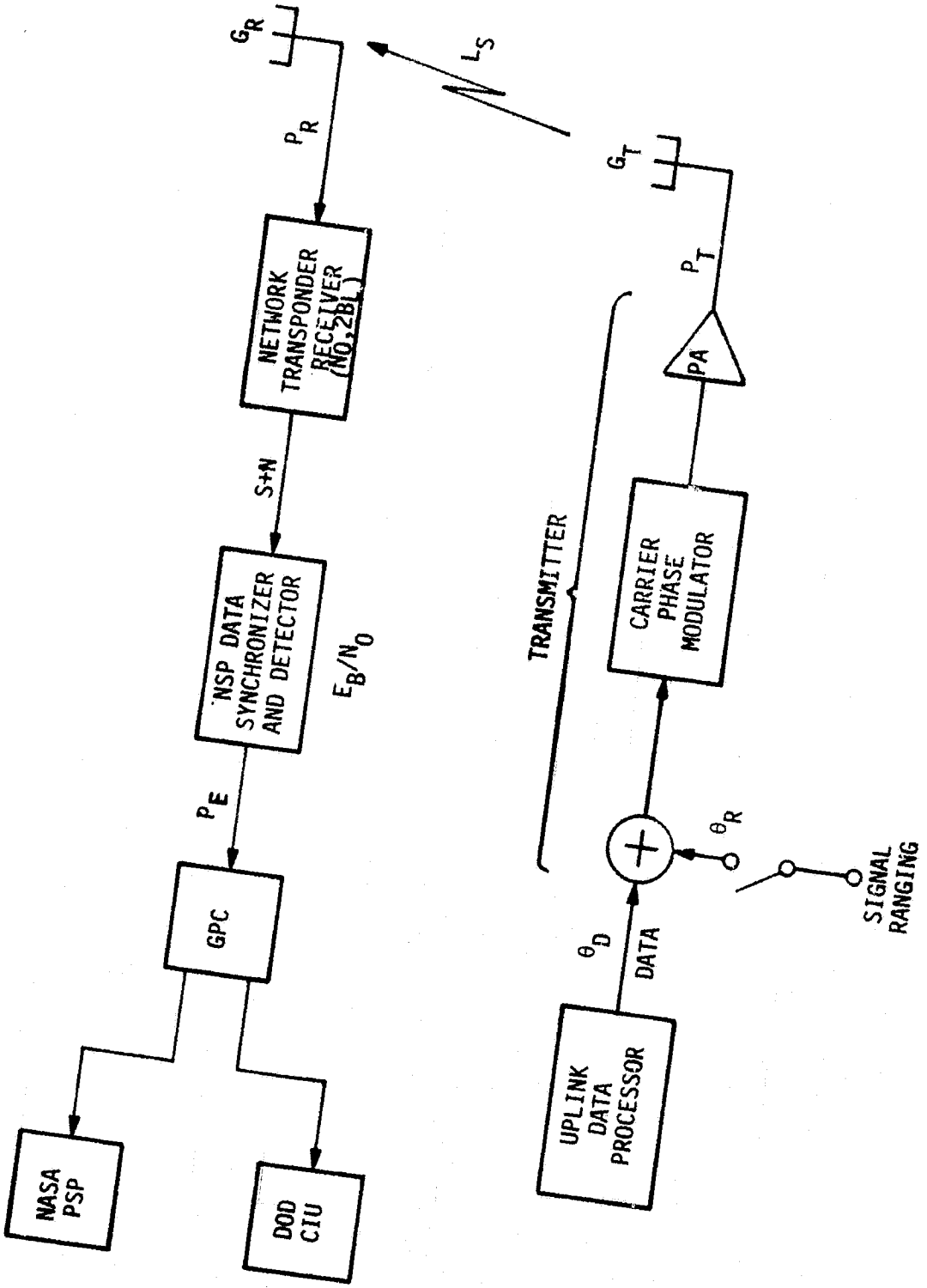


Figure 2.12. Ground-to-Orbiter S-Band GSTDN PM Link

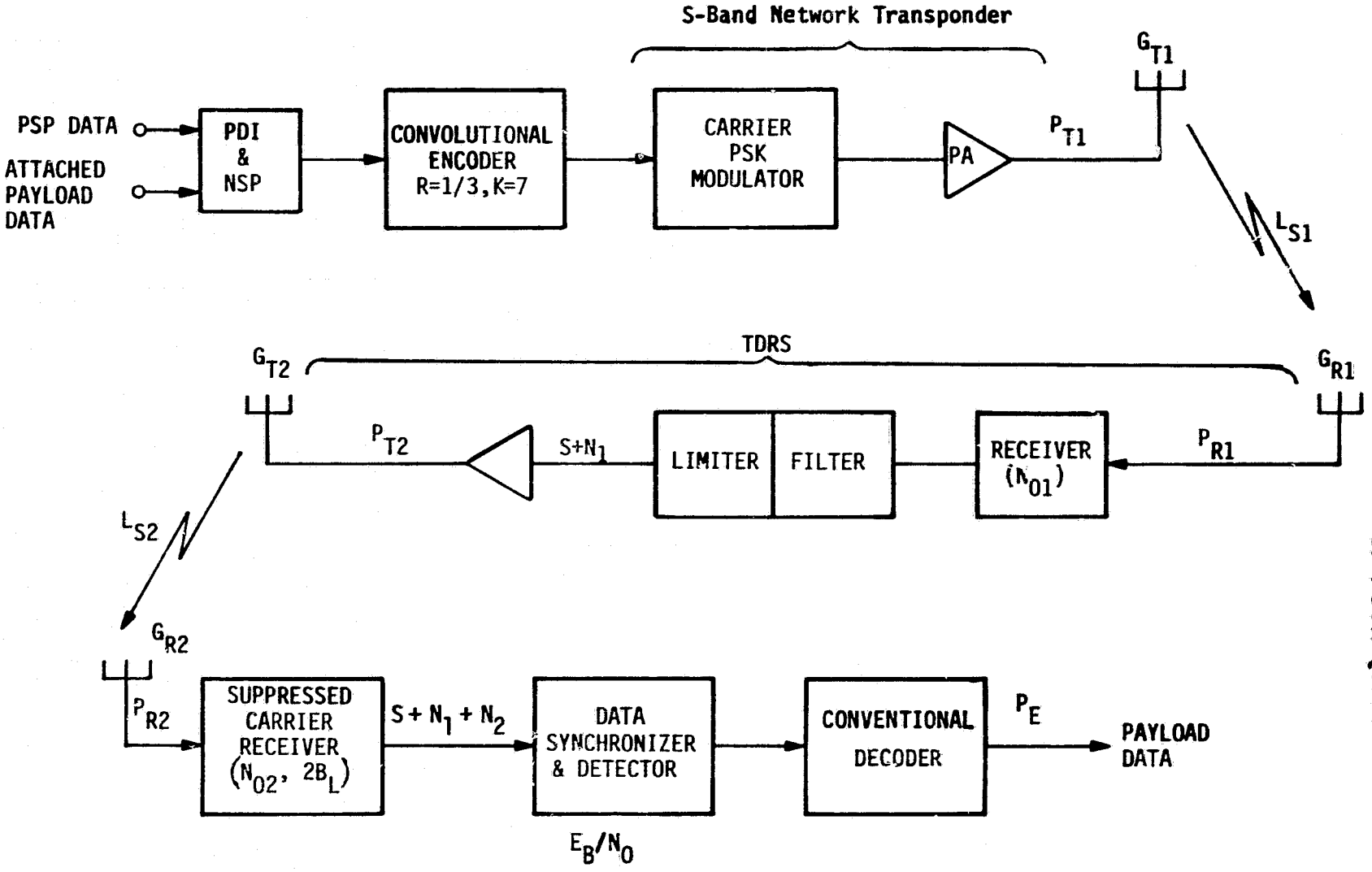


Figure 2.13. Orbiter-to-Relay-to-Ground S-Band TDRSS PSK Link

The two RF links complicate the design control procedure because of the presence of two independent noise sources (one in the TDRS receiver and the other in the ground station receiver) that are separated by intermediate signal processing functions. Especially important is that the TDRS signal processing involves an amplitude limiter which has a useful output signal component that is a function of the limiter input signal-to-noise ratio. (The parametric nature of this portion of the channel is discussed in subsection 3.2.15.)

Subsection 2.2.2.1 and Section 3 should be consulted concerning the definition of symbols. Appendix B6 contains a design control table example of the complete link.

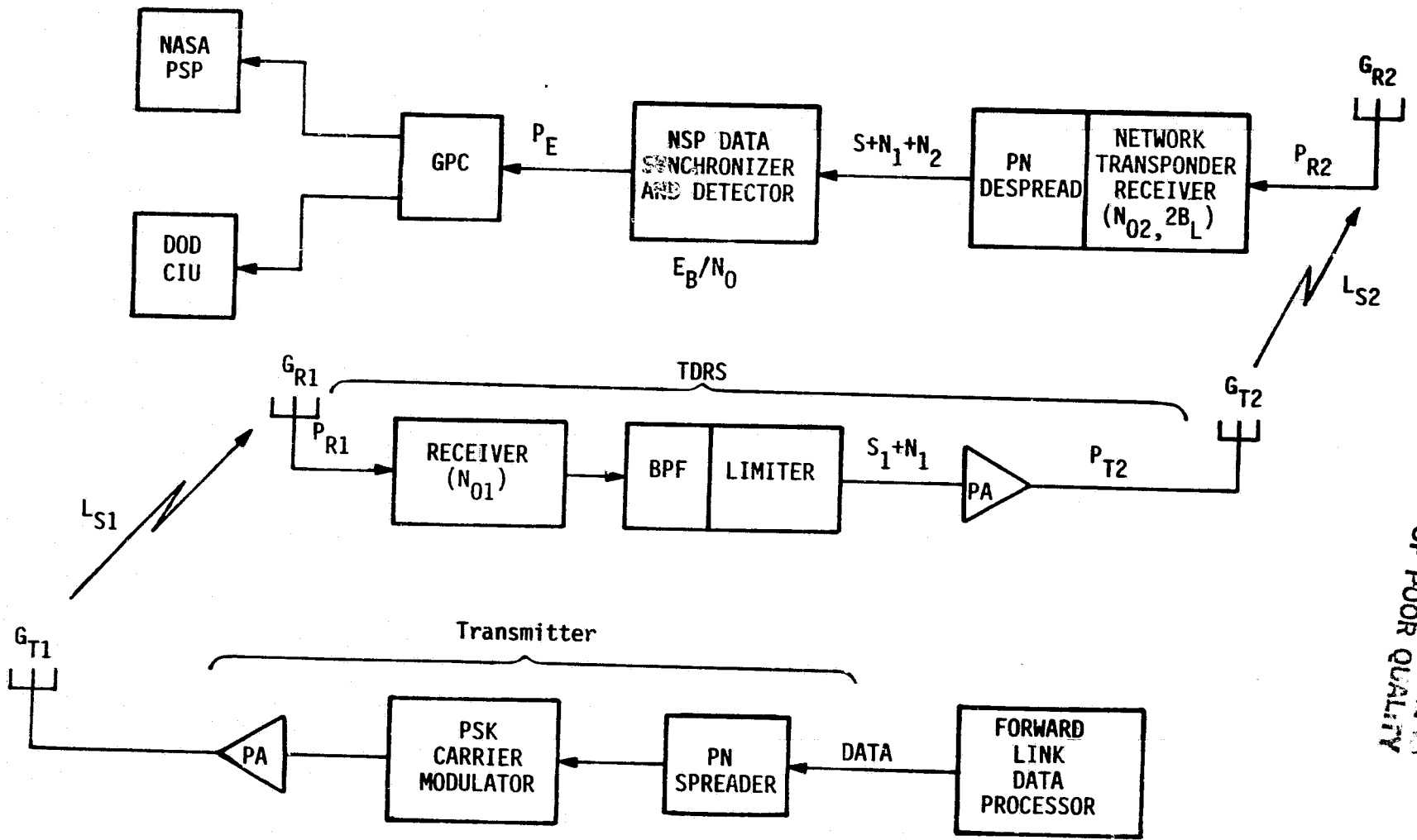
Figure 2.14 shows a block diagram of the ground-to-Orbiter S-band TDRSS relay link. Functionally, it differs from the Orbiter-to-ground TDRSS link in that it does not employ convolutional coding of the data, but is required to PN code spread the data. (See subsection 2.1.2.6.) PN code spreading and despreading does not improve the data detection probability of error as does convolutional coding/decoding; in fact, some fraction of the useful signal power is usually lost (see subsection 3.1.11.1.2).

2.3.1.3 S-band FM direct link

Figure 2.15 indicates the principal functions associated with the S-band FM Orbiter-to-ground direct link. (NOTE: There is no comparable ground-to-Orbiter link.)

The FM Signal Processor (FMSP) provides mode control of the input signals. Analog waveforms are simply throughput with some lowpass filtering effects above 4 MHz, while digital data is conditioned so that the FMSP output is bipolar at the correct drive level for the FM transmitter. It must be understood that the FM transmitter deviation by analog signals is directly proportional to the amplitude of the signal input to the FMSP.

Appendix B6 lists the system parameters for the FM link and Appendix B5 is a typical design control table for an analog signal.



NOTE: The ground-to-relay-to-Orbiter Ku-band link is functionally identical. The NSP is replaced by the KuSP, and DOD data rates do not flow through the GPC.

Figure 2.14. Ground-to-Relay-to-Orbiter S-Band TDRSS Link

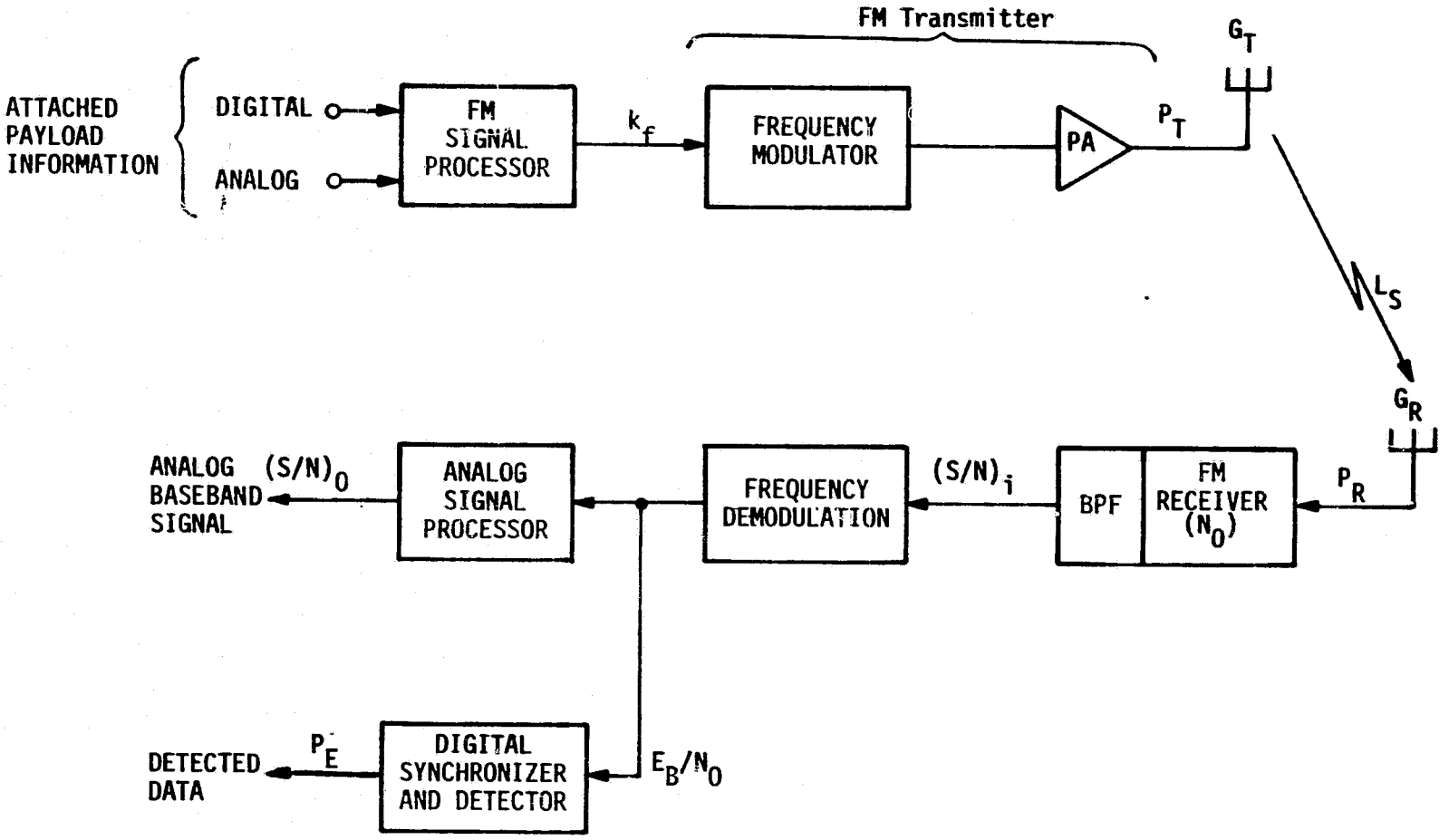


Figure 2.15. Orbiter-to-Ground S-Band GSTDN FM Link

2.3.2 Ku-Band TDRSS Relay Links

The Ku-band links are the most complex of all the Shuttle communication channels from the standpoint of modulation and coding. High data rate or wideband analog signals directly modulate the carrier, while lower data rate and narrowband analog signals are modulated onto a subcarrier. Particulars are given in the following subsections.

2.3.2.1 Ku-band digital channels

Shown in Figure 2.16 is a functional block diagram of the Orbiter-to-relay-to-ground Ku-band quadriphase-shift-keying (QPSK) link. QPSK is used at both the carrier and subcarrier levels. High rate payload data (up to 50 Mbps) is rate one-half convolutionally coded and receives 80% of the QPSK modulated carrier power. The remaining 20% resides in the square-wave 8.5-MHz subcarrier. A 80/20% subcarrier power split is also achieved by means of subcarrier QPSK modulation. Payload data receives the 80% allocation; however, since the square-wave subcarrier harmonics are lost in the ground subcarrier tracker/demodulator, 19% of this power is effectively lost.

Apart from the above modulation considerations, the remainder of the Ku-band digital channel is functionally similar to that of the S-band relay link discussed in subsection 2.3.1.2. Appendix A7 contains the pertinent Ku-band link parameters, and a typical design control table for subcarrier data may be found in Appendix B7.

2.3.2.2 Ku-band analog channels

Analog signals may be transferred from the payload to the ground via the Ku-band relay link operating in the FM mode. Nonstandard digital signals may also be accommodated in this mode. For nonstandard payload telemetry, the Ku-band relay channel is referred to as the "bent pipe" (see Figure 2.17).

Wideband (up to 4.5 MHz) analog signals are directly frequency modulated onto the carrier. Narrowband (up to 2 MHz) signals from payloads are amplitude limited to produce a two-level bipolar waveform which, in turn, modulates the 8.5-MHz subcarrier as if it were digital data. The subcarrier then frequency modulates the carrier. It is anticipated that the most common form of signal to be handled in the narrowband mode will

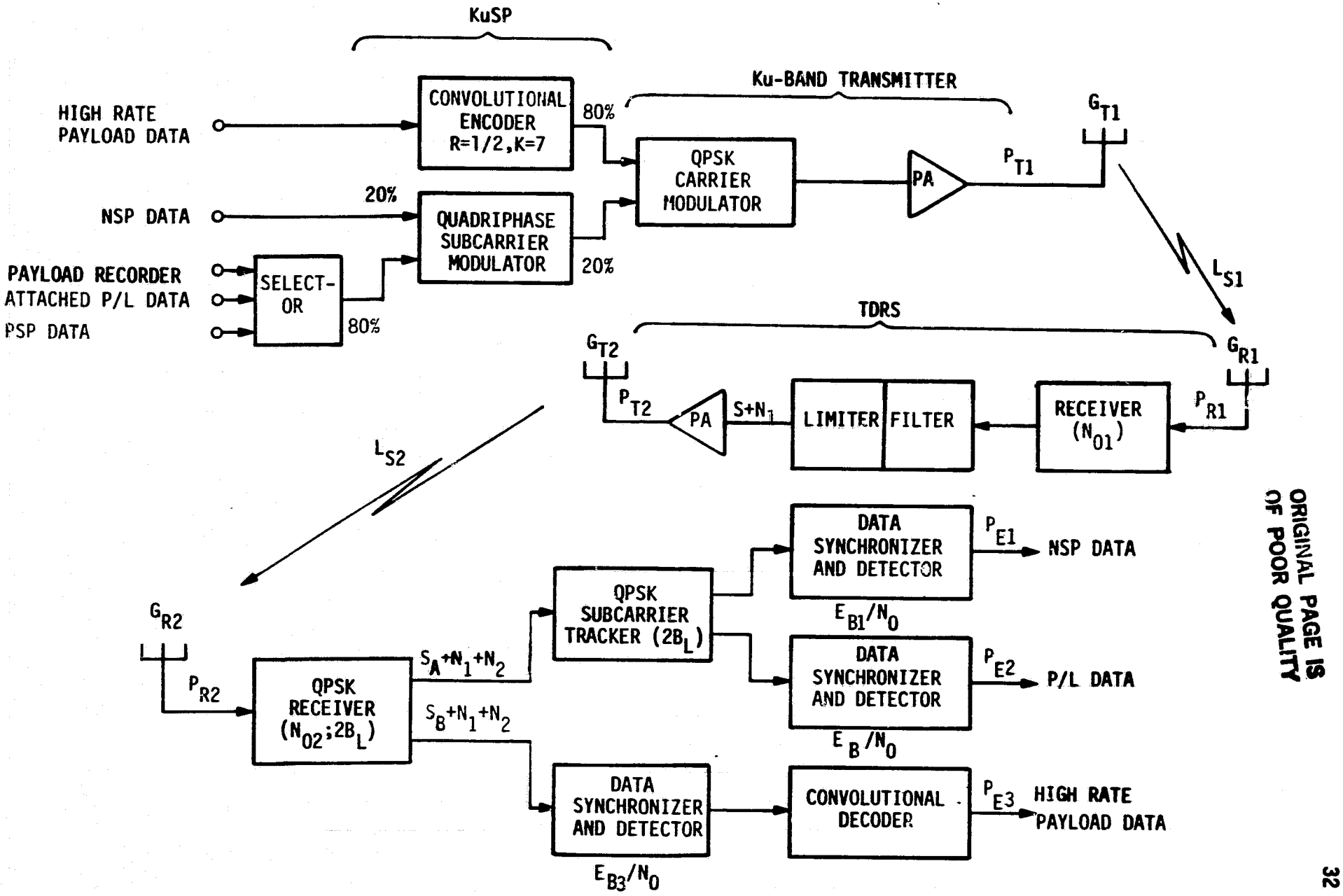


Figure 2.16. Orbiter-to-Relay-to-Ground Ku-Band QPSK Link

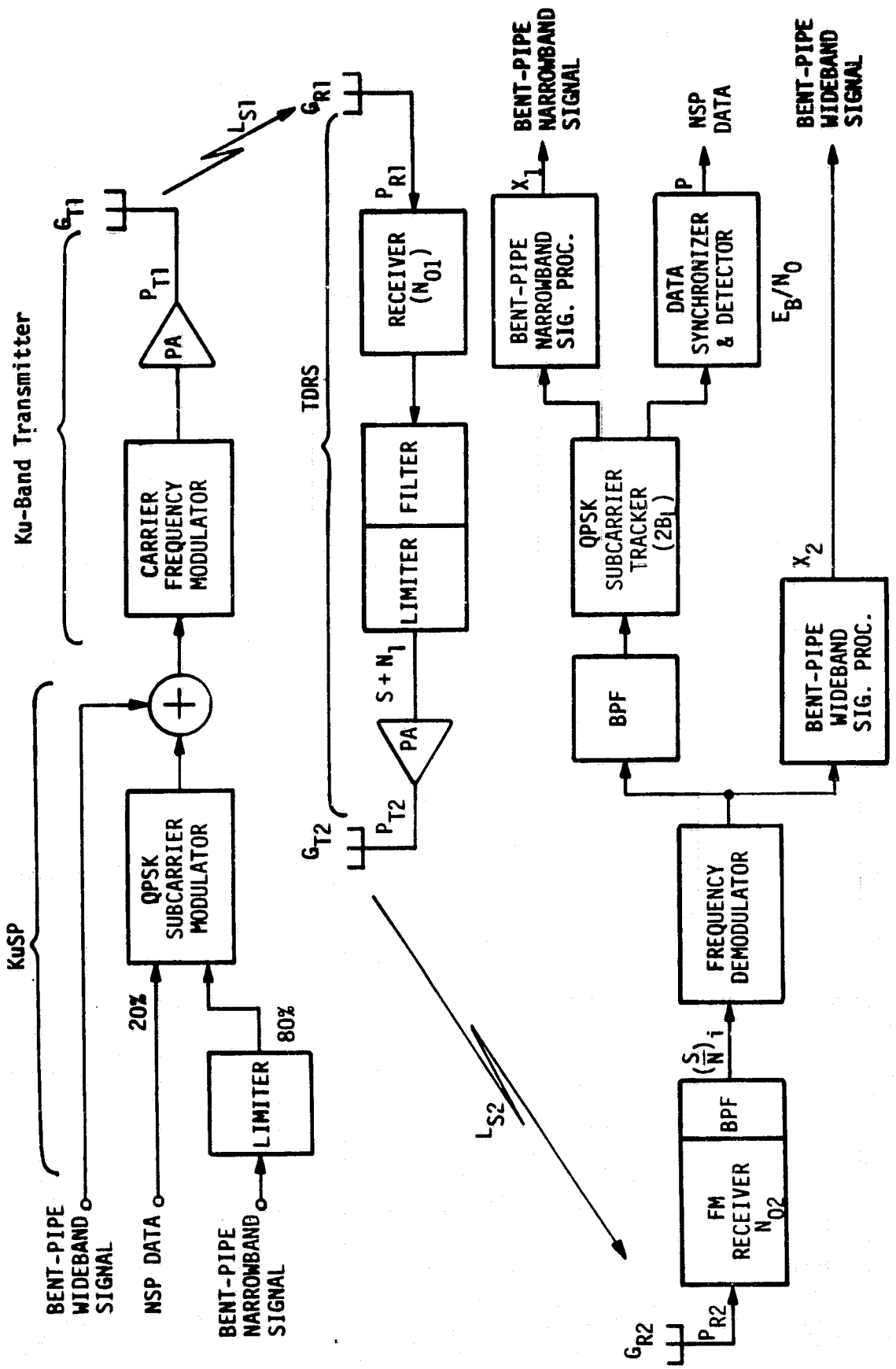


Figure 2.17. Orbiter-to-Relay-to-Ground Ku-Band Bent-Pipe FM Link

be a nonstandard modulated subcarrier from the payload, as derived from the PI receiver. Any narrowband or wideband nonstandard bent-pipe signal processing required on the ground is the responsibility of the payload user.

Ku-band bent-pipe parameters may be found in Appendix A7, and Appendix B8 is a design control table example for a wideband bent-pipe link.

3.0 COMMUNICATION LINK MODELS AND PARAMETERS

This section presents the typical communication links from a communication parameter perspective. The link model is comprised of an end-to-end series of "blocks," each of which has a set of pertinent parameters identified with it which must be specified for the link budget design. A short explanation is given of the theoretical/engineering nature of each parameter, along with the units for specification.

3.1 Typical Communication Link Models

A generalized communication link block diagram is shown in Figure 3.1. The links identified in the figure are typical of those which might be encountered in a given payload/Orbiter mission. For example, digital or analog data could be modulated directly onto the carrier or onto one or two subcarriers. Digital data could be encoded at rate R, with constraint (or block) length K for error correction at the receiver or left uncoded if the overall link bit error rate is satisfactory. Analog data could use preemphasis of H(f), depending on the spectrum of the analog waveform. Signal premodulation processing is presented in subsection 3.2.2.

When a signal modulated a carrier or subcarrier, partitioning of the total carrier or subcarrier power occurs in association with the modulation components. Since phase and frequency modulation are nonlinear processes, a certain amount of the total carrier or subcarrier power is also wasted or lost in unwanted harmonic-distortion-related terms.

Since the amount of available power residing in any desired or useful component is proportional to the modulating intensity of the particular signal giving rise to the component, this available power is expressed as a fraction of the total power. In general, if P_{mi} is the power of the ith modulation component and P_T is the total power available, the ratio P_{mi}/P_T represents the fraction of total power allocated to the ith modulating signal. Since this fraction is always less than unity, it is sometimes called "modulation loss."

Prior to modulation being applied to the transmitter, all the power resides in the discrete carrier (single frequency). With full modulation taking place, some fraction of the discrete carrier may remain; this is denoted by the ratio P_C/P_T.

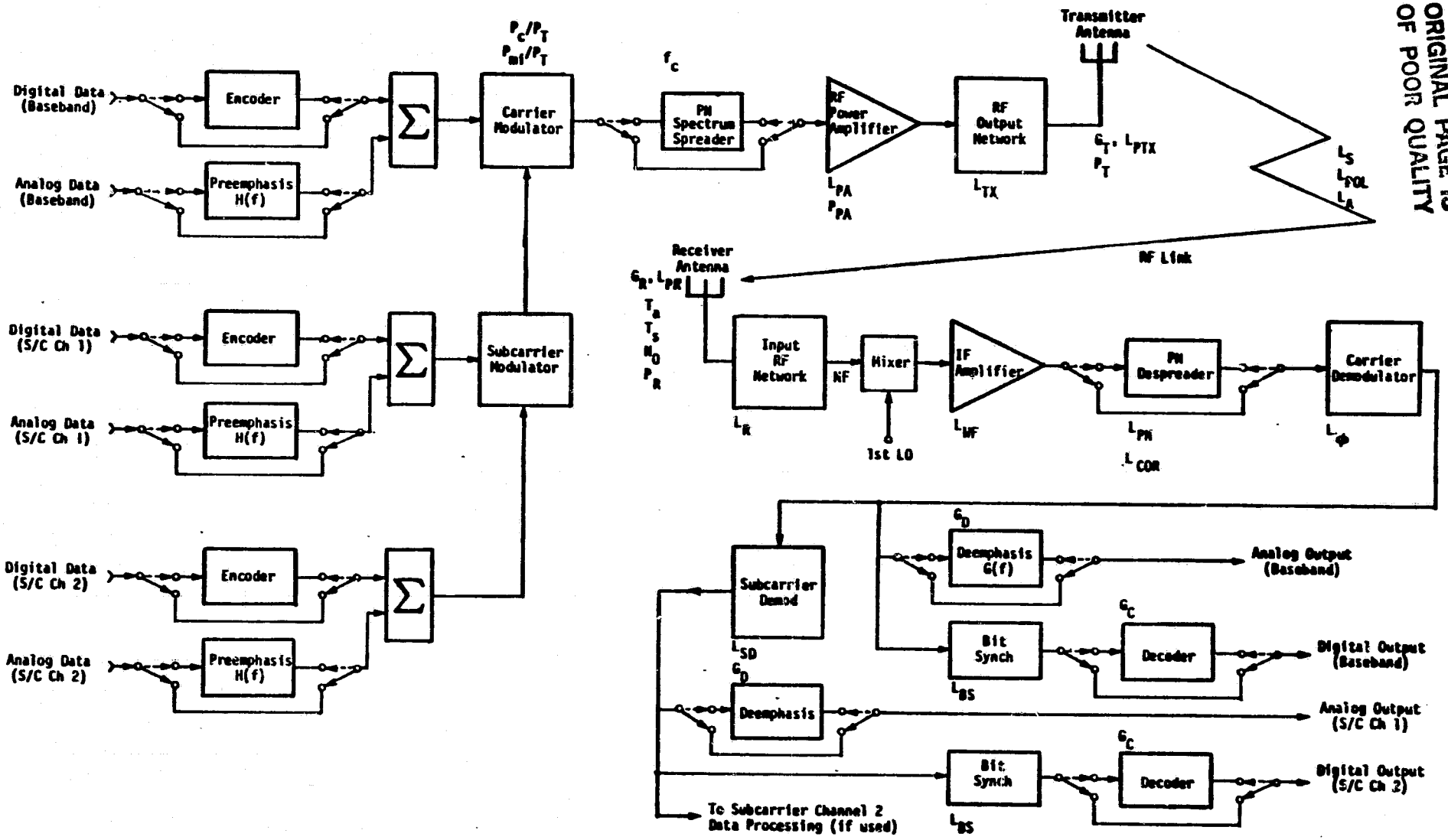


Figure 3.1. Generalized Link Block Diagram

Details concerning the relationships for calculating carrier and subcarrier power allocation as a function of the modulating signal's intensity are found in subsections 3.2.3, 3.2.4 and 3.2.5.

Following modulation, the communication signal could be spectrum spread by a PN sequence of chip rate f_c with period p chips. Whether or not the communication signal is spread, the signal is amplified in a power amplifier to power P_{PA}. Note that the power amplitude can cause signal distortion, and the associated performance loss as described in subsection 3.2.5 is denoted L_{PA}. Additional loss is caused by the RF output network that connects the amplified power with the antenna. The RF output network loss is denoted by L_{TX}. Finally, the communication signal is transmitted by the antenna which has gain G_T, pointing loss L_{PT}, and output power P_T, as discussed in subsection 3.2.6.

The RF channel over which the communication signal is transmitted exhibits several losses. These losses are discussed in detail in subsection 3.2.7 and include space loss (denoted L_S in Figure 3.1), polarization loss (L_{POL}), and ionospheric and atmospheric (weather) losses (L_A), as presented in detail in subsection 3.2.7.

The receiver antenna is also characterized by gain and received power, as was the transmitter antenna. The receiver antenna gain is denoted by G_R, pointing loss by L_{PR}, and the received power is denoted by P_R in Figure 3.1. Subsection 3.2.6 discusses the antenna gain in detail. In addition to the gain and received power at the receiver antenna, other parameters are necessary to describe the performance of the receiver RF front end. First, there are the contributions to system noise as presented in subsection 3.2.8. Figure 3.1 shows four of the contributions to the system noise: (1) atmospheric noise temperature T_A, (2) system noise temperature T_S, (3) noise spectral density N_0, and (4) noise figure of the RF front-end NF. Note that the RF network illustrated in Figure 3.1 also has a loss L_R associated with it, as discussed in subsection 3.2.6.

Following the mixing with the first local oscillator (LO) in the receiver, the communication signal and system noise are amplified by the IF amplifier shown in Figure 3.1. Since both signal and noise are amplified,

there is no increase in signal-to-noise ratio (SNR); however, the IF amplifier has a loss L_{WF} associated with it due to signal waveform distortion as far as the communication performance is concerned. When a spread spectrum signal is used, the receiver must remove the spread spectrum modulation from the communication signal. As shown in Figure 3.1, the PN despreader removes the spread spectrum modulation, but with some loss in performance, as denoted by L_{PN} and L_{COR} in subsection 3.2.10.

The carrier demodulator shown in Figure 3.1 reconstructs the carrier and separates the direct modulation on the carrier from the modulation on the subcarrier. The carrier-tracking loops and carrier demodulation are presented in subsections 3.2.10 and 3.2.11, respectively. The performance loss associated with carrier demodulation is denoted L_{ϕ}. Similarly, the subcarrier demodulator separates the data modulated on each of the subcarriers by reconstructing the subcarrier, as described in subsection 3.2.12. The performance loss incurred in the subcarrier demodulator is denoted L_{SD}.

The digital data modulated on either the carrier or subcarrier is detected by a bit synchronizer, as shown in Figure 3.1. The performance loss resulting from the bit synchronizer and data detection is denoted by L_{BS} and is discussed in detail in subsection 3.2.13. If the digital data was error-correction encoded at the transmitter, by using a decoder, a performance gain is then obtained which is denoted as coding gain G_C in Figure 3.1. The amount of coding gain depends on the error correction code and the type of decoder used. The trade-offs in choosing the error correction code and decoder type are presented in subsection 3.2.14.

The analog data is output to the user following the appropriate carrier or subcarrier demodulation. If preemphasis or compression was performed at the transmitter, then deemphasis or expansion must be performed at the receiver. These postdemodulation analog signal-processing techniques are presented in subsection 3.2.11. The resulting performance gain is denoted by G_D in Figure 3.1.

The communication link parameters shown in Figure 3.1 and presented in subsection 3.2 provide the necessary information to compute the circuit margin for each possible payload/Orbiter communication link. Section 4 provides the structure of the design control table in order to properly use the information derived in this section to determine the communication link budget and resulting circuit margin.

3.2 Communication Link Parameters

The communication link parameters presented in the last subsection describing Figure 3.1 are discussed in detail in this subsection.

3.2.1 Information Waveforms

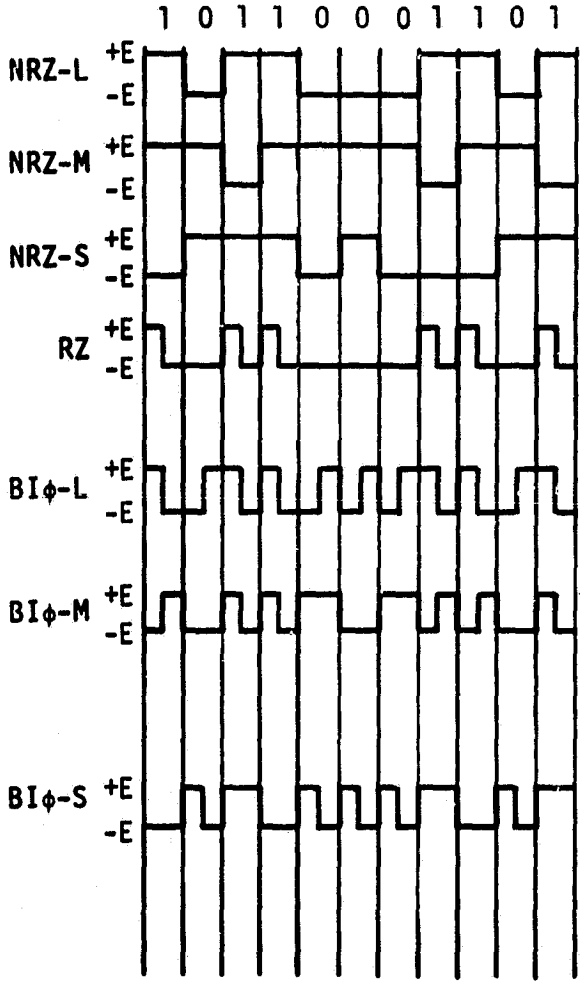
The data to be communicated may be either digital or analog in nature. The following two subsections describe the various forms which the digital and analog signals might take and present the pertinent parameters needed to be specified for each data form.

3.2.1.1 Digital data

Natural digital data are typically information such as commands, addresses, synchronization words, computer data, etc. Derived digital data is comprised of a series of binary digits (ones and zeros) to describe the voltage or current level of a sample taken from an analog waveform. This latter binary data derived from analog signals is called pulse code modulation (PCM).

To transmit digital data, some voltage waveform must exist to represent the "1: and "0" bits (binary digits). Several different waveforms have been used in the formulation of a bit stream. The Inter-Range Instrumentation Group (IRIG) of the Range Commanders' Council recognizes seven permissible digital formats [5]. Figure 3,2 illustrates these various formats. The "+E" and "-E" levels indicated in Figure 3.2 represent the actual voltage levels of the bit stream, while the "1" and "0" indicate binary logic levels. It should be noted that each of the formats illustrated in Figure 3.2 is actually a variant of one of the following three basic formats: (1) nonreturn-to-zero level (NRZ), (2) biphase level (biϕ-L), which is also referred to as split-phase and Manchester-coded data (the latter designation is used throughout this handbook), and (3) return-to-zero (RZ).

One important difference among the three basic formats is the transmission bandwidth required. The bandwidths associated with the various NRZ formats are the same, as are those of each of the biphase formats. The power density spectrums for the NRZ, Manchester, and RZ



NRZ-LEVEL (or NRZ-CHANGE); "1" is represented by one level, "0" is represented by the other level

NRZ-MARK; "1" is represented by a change in level, "0" is represented by no change in level

NRZ-SPACE; "1" is represented by no change in level, "0" is represented by a change in level

RZ; "1" is represented by a half-bit wide pulse, "0" is represented by no pulse condition

BIPHASE LEVEL (OR SPLIT PHASE, MANCHESTER CODE); "1" is represented by "10," "0" is represented by "01"

BIPHASE MARK; A transition occurs at the beginning of every bit period, "1" is represented by a second transition 1/2 bit period later, "0" is represented by no second transition

BIPHASE SPACE; A transition occurs at the beginning of every bit period, "1" is represented by no second transition, "0" is represented by a second transition 1/2 bit period later

Figure 3.2. PCM Code Formats

formats are shown in Figures 3.3 through 3.5, respectively. The values in parentheses correspond to frequencies which are multiples of the NRZ bit rate, R_b.

The actual bandwidth, BW, required for transmission of a bit stream depends on the desired signal fidelity. That is, a narrow transmission bandwidth results in a train of distorted pulses and a corresponding performance loss rather than the original square pulses corresponding to the ideal performance. The allowable transmission bandwidth is typically determined by specifications necessary to eliminate interference with the other communication channels onboard the Orbiter or payload. The relative bandwidth requirements for the various formats may be determined from Figure 3.6, which contains plots of the percentage of total code power contained within a given baseband (premodulation or postdetection) bandwidth. It is evident that the NRZ-L code requires the least bandwidth for a given percentage of total code power. The Manchester code requires the most bandwidth.

3.2.1.2 Analog data

Analog data may be voice, television, sensor responses, etc. In order to transmit analog data, several processes must be performed, as shown in Figure 3.1. The goal in the transmission of analog data is to produce an undistorted and fairly noise-free output at the receiver. Distortionless transmission does not necessarily imply that the output is identical to the input. Certain differences can be tolerated and not classified as distortion. More precisely, given an input signal x(t), the output is undistorted if it differs from the input only by a multiplying constant K and a finite time delay t_0. Analytically, if

$$y(t) = K\,x(t - t_0) \qquad (3\text{-}1)$$

then y(t) is considered to be a distortionless signal. (NOTE: t_0 must be positive or zero in a realizable system.) In order for a transmission network to produce distortionless transmission, the network must have constant amplitude response and a negative linear phase shift; that is

$$|H(f)| = \text{constant}, \quad \theta(f) = -\omega t_0 \pm m\pi, \qquad (3\text{-}2)$$

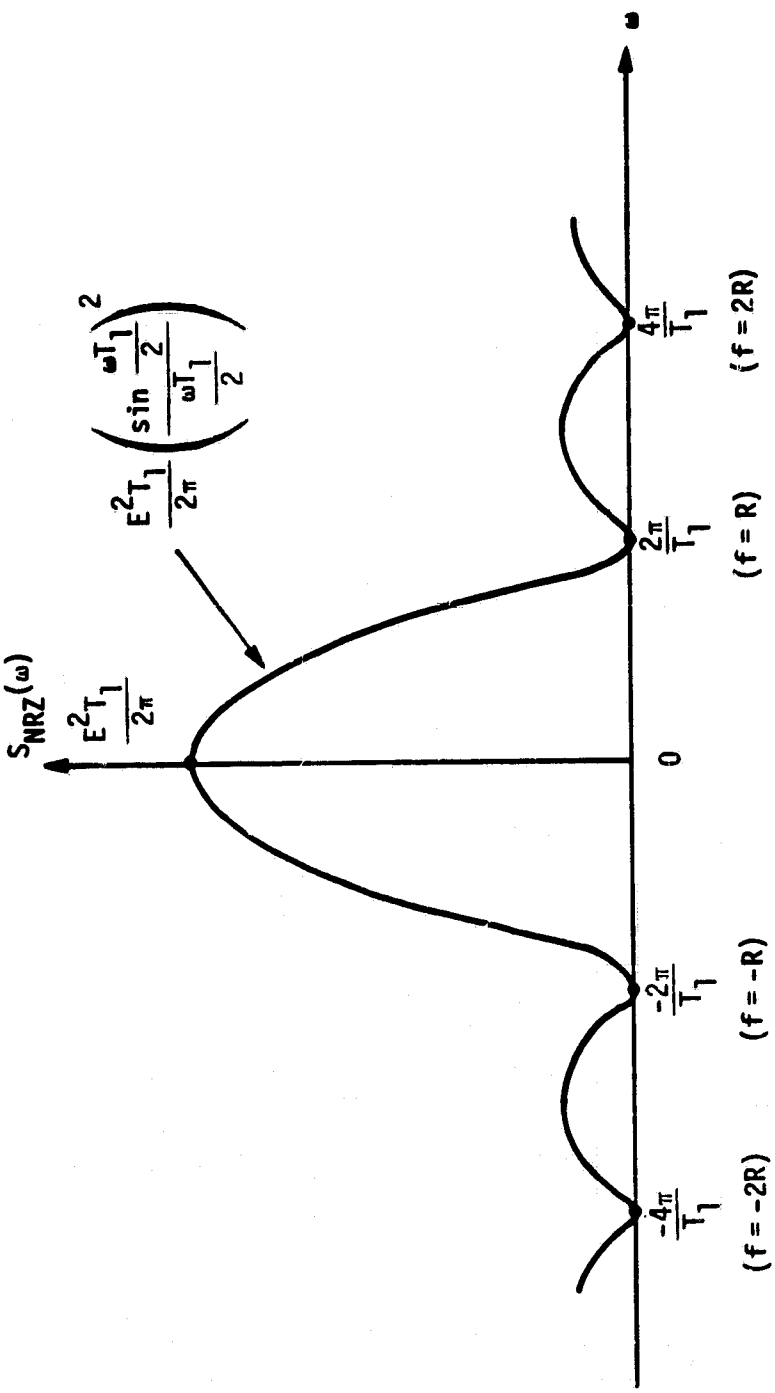


Figure 3.3. Power Density Spectrum of an NRZ Code (Random Bit Pattern)

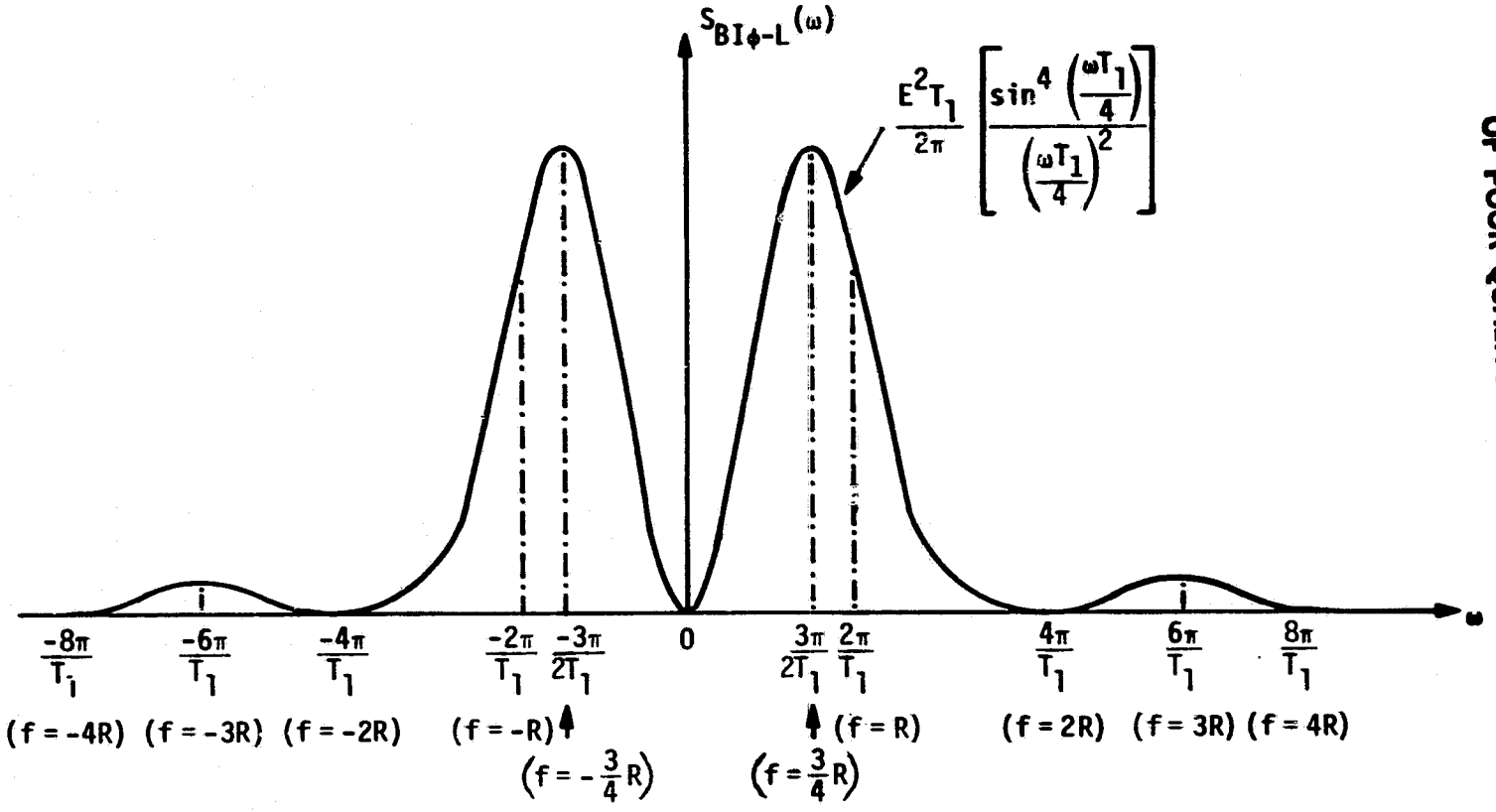


Figure 3.4. Power Density Spectrum of a Manchester Code (Random Bit Pattern)

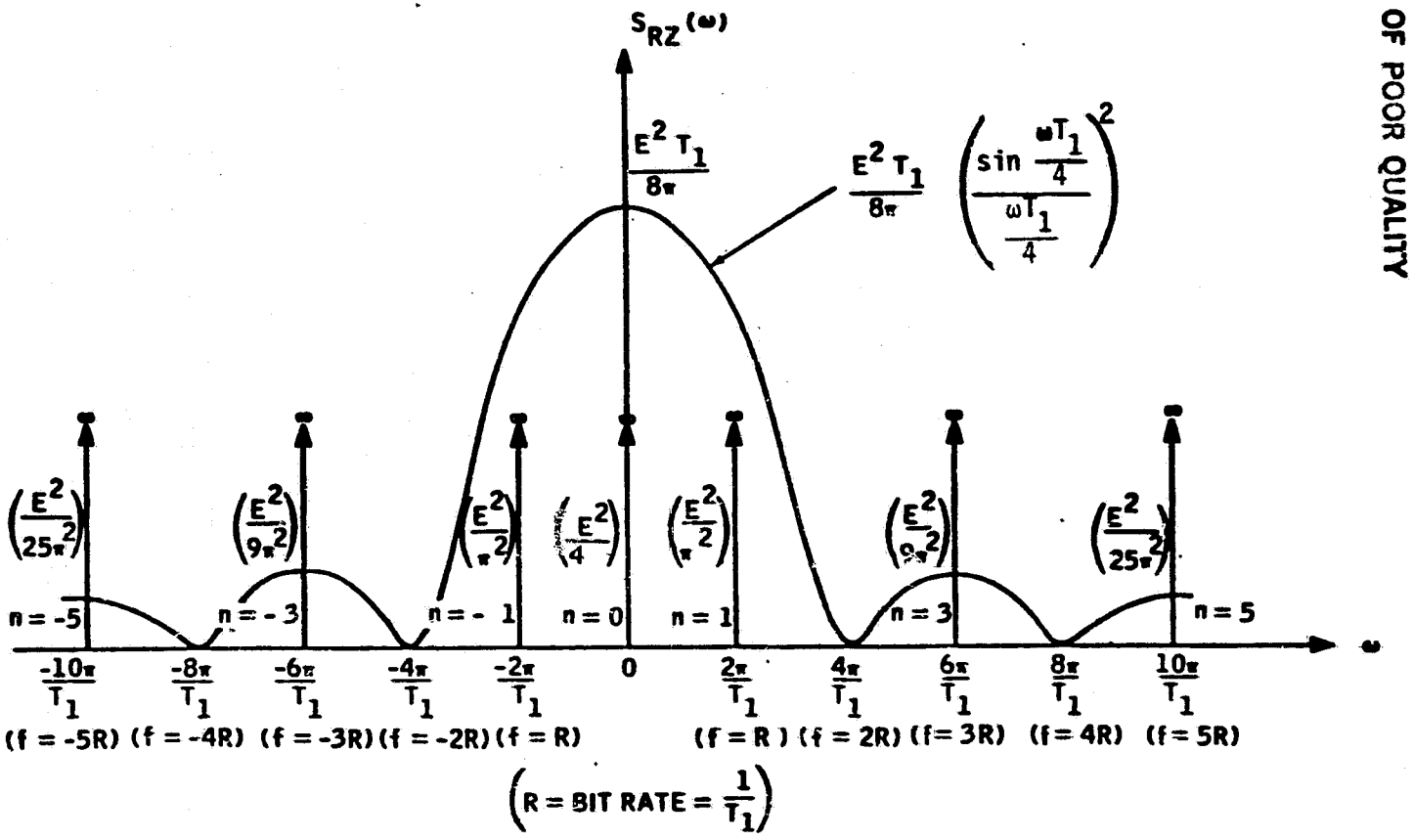


Figure 3.5. Power Spectrum of an RZ Code (Random Bit Pattern)

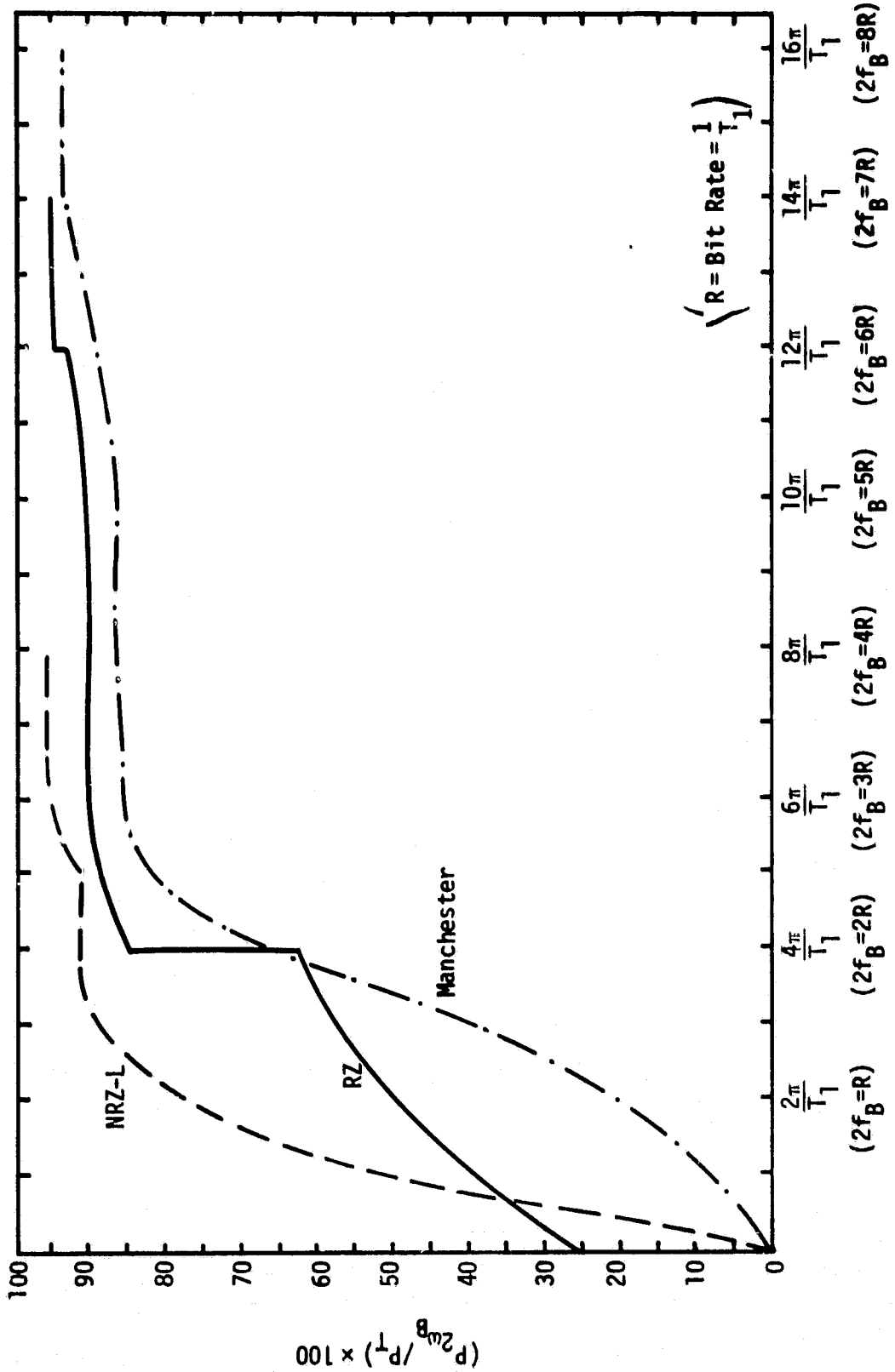


Figure 3.6. Percentage of Total Power of a PCM Code Contained in the Frequency Band Extending From $-\omega_B$ to $+\omega_B$

where $|H(f)|$ is the amplitude transfer response and $\theta(f)$ is the phase transfer response of the network. The term $\pm m\pi$ allows for the constant K to be positive or negative. One qualification that can be added to (3-2) is that these conditions are required only over those frequencies for which the input signal has a nonzero spectrum. Thus, if $x(t)$ is bandlimited to W, then (3-2) need be satisfied only for $|f| < W$.

In practice, (3-2) is a stringent condition which can be only approximately satisfied, at best. The three major types of distortion are: (1) amplitude distortion, where $|H(f)| \neq$ constant, (2) phase or delay distortion, where $\theta(f) \neq -\omega t_0 \pm m\pi$, and (3) nonlinear distortion. In the third case, the transmission network includes nonlinear elements, and its transfer function may not be defined in a linear manner.

Amplitude distortion is easily described in the frequency domain; it means simply that the output frequency components are not in correct proportion. Since this is caused by the $|H(f)|$ not being constant with frequency, amplitude distortion is sometimes called frequency distortion. The most common forms of amplitude distortion are excess attenuation and enhancement of extremely high and low frequencies in the signal spectrum. The results of amplitude distortion are difficult to quantify without an experimental study of specific signal types. Even the experimental studies of these signal types are usually couched in terms of the required frequency response; that is, the range of frequencies over which $|H(f)|$ must be constant to within a certain tolerance (e.g., ±1 dB) so that the amplitude distortion is sufficiently small. Table 3.1 lists typical frequency response requirements for signals commonly encountered in communication systems.

Table 3.1. Typical Frequency Response Requirements

Signal Type	Frequency Range (Hz)
Voice	
High fidelity	20 - 20,000
Average broadcast quality	100 - 5,000
Average telephone quality	200 - 3,200
Barely intelligible speech	500 - 2,000
Television video	60 - 4,200,000

A linear phase shift yields a constant time delay for all frequency components in the signal. This, coupled with constant amplitude response, yields an undistorted output. If the phase shift is not linear, the various frequency components suffer different amount of time delay, and the resulting distortion is termed phase or delay distortion. Delay distortion can be critical for television video transmission. On the other hand, the human ear is fairly insensitive to delay distortion. Thus, delay distortion is seldom of concern in voice transmission.

Nonlinear distortion results from nonlinear or amplitude saturation in the transmission network. The input and output of these nonlinear elements are related by a curve or function, commonly called the transfer characteristic. The most familiar nonlinear characteristic is the flattening out of the output for large input excursions typical of the saturation and cut-off effects of amplifiers. A quantitative measure of nonlinear distortion is provided by taking a simple cosine wave, $x(t) = \cos \omega_0 t$, as the input. The nonlinear distortions appear as harmonics of the input wave. The amount of second-harmonic distortion is the ratio of the amplitude of the second harmonic to that of the fundamental (i.e., the input wave). Higher order harmonics are treated similarly. If the input is a sum of two cosine waves (say, $\cos \omega_1 t + \cos \omega_2 t$), the output will include all the harmonics of f_1 and f_2 plus cross-product terms which yield $f_1 - f_2$, $f_1 + f_2$, $f_2 - 2f_1$, etc. These sum and difference frequencies are designated as intermodulation distortion. Generalizing the intermodulation effect, if $x(t) = x_1(t) + x_2(t)$, then $y(t)$ contains the cross product $x_1(t)x_2(t)$. In the frequency domain, $x_1(t)x_2(t)$ becomes the convolution of their power density spectra [i.e., $S_{x_1}(\omega) * S_{x_2}(\omega)$] and, even though $S_{x_1}(\omega)$ and $S_{x_2}(\omega)$ may be separated in frequency, $S_{x_1}(\omega)$ $S_{x_2}(\omega)$ can overlap both of them, thus producing one form of crosstalk.

3.2.2 Premodulation Processing

Premodulation processing is performed on the data to compensate for some of the deleterious effects of the communication signal transmission system. In terms of digital data, some form of error-correcting code may be employed to decrease the probability of error due to noise at the receiver. Similarly, for analog data, preemphasis might be used to enhance the SNR by performing deemphasis at the receiver.

3.2.2.1 Digital coding

Coding is the process of mapping k data bits into n code symbols to expand the transmission bandwidth in a particular way. Decoding is the inverse process performed by mapping the detected n symbols estimated into k data bits. If an efficient code is selected and a good decoding scheme is used, the k data bits will have a lower average bit probability for the same energy per bit to single-sided noise spectral density (E_b/N_0) than if they were transmitted uncoded.

When a block of data bits, taken k at a time, are mapped into a sequence of n symbols, taken from a finite set of possible symbol sequences, the process is known as block coding and the code is called an (n,k) block code. Each of the 2^k possible data words is uniquely mapped into one of 2^n possible codewords. If each bit in the data word is independent of the other bits and is equally likely to be a "1" or a "0," then all the possible codewords are equally probable. Typical block codes used for space communication links are the Golay (24,12) code, the quadratic residue (48,24) code, the (16,5) biorthogonal code, the (32,6) biorthogonal code, various Bose-Chaudhuri-Hocquenghem (BCH) codes, and various Reed-Solomon (RS) codes. For details of these codes and their implementation, see Peterson and Weldon [6].

Another class of error-correcting codes used for space communication links is convolutional codes. Convolutional codes have produced performance that is uniformly better than that of block coding for given symbol-rate-to-data-rate ratios. In addition, the encoder implementation is simpler and thus less costly and more reliable for spacecraft operation than that of a block encoder.

Binary convolutional codes may be generated in the manner shown in Figure 3.7. The binary input data is shifted bit by bit into a K-stage shift register. With every shift, the commutator samples n output lines from adder 1 through adder n and forms a binary symbol output of the serial samplings, called a code branch. Thus, if the input bit rate is $1/T_b$, the output symbol rate $1/T_s$ is equal to n/T_b. Such an encoder generates a constraint length K, rate $R = 1/n$ (bits/symbol) convolutional code. K is called the constraint length because each input bit affects, or constrains, the subsequent Kn output symbols derived from K input bit shifts.

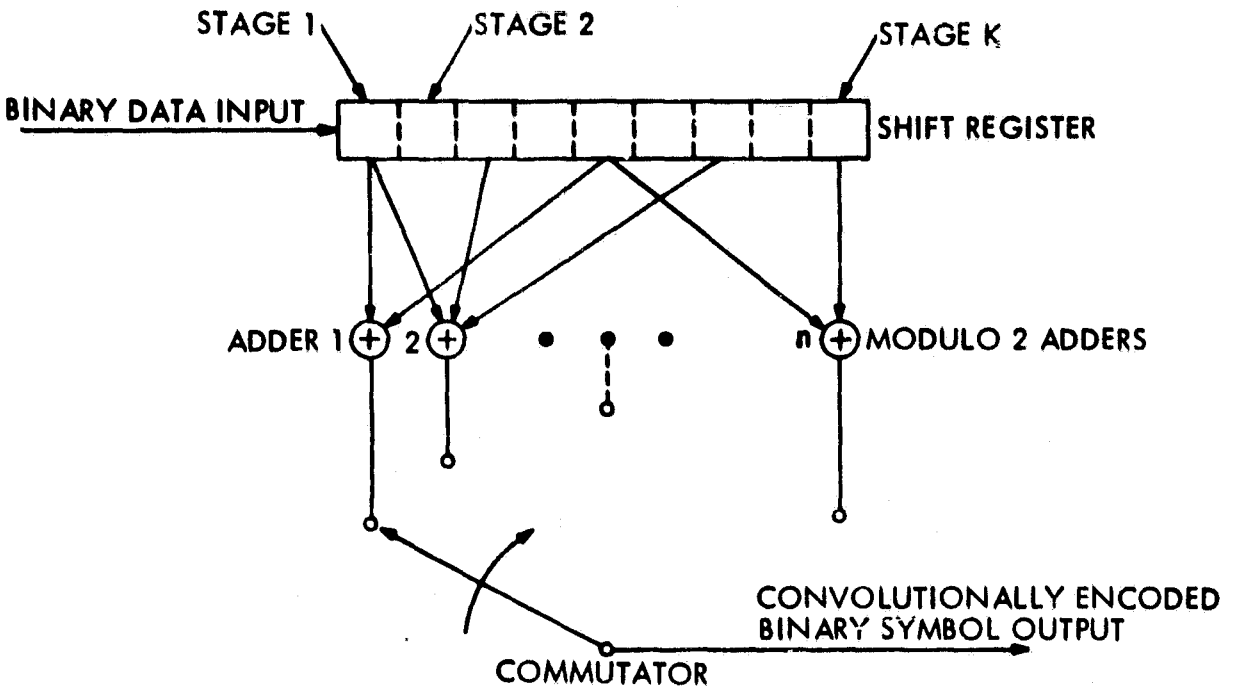


Figure 3.7. Length K, Rate R = 1/n Convolutional Encoder

In order to send a block of l data bits, the shift register is initialized to the all-zero state. The bits are then shifted in until all L have been shifted out of the register and it is again filled with zeros. Thus, L input bits give (L + K)n output symbols, where the Kn output symbols are due to the "tail" necessary to flush the shift register with input K zeros following L data bits.

3.2.2.2 Analog processing

Suppose that an analog baseband signal is to be transmitted using frequency modulation which requires the best possible SNR for a given carrier power, noise spectral density, R_b bandwidth and baseband bandwidth. Clearly then, the baseband signal to the frequency should produce as large a carrier deviation as possible which is consistent with the linearity of the frequency modulator and receiver bandwidth. Distortion eventually occurs because the frequency deviation exceeds the linearity and/or bandwidth capabilties. Thus, the modulating signal level may be raised only to the point where the distortion exceeds a specified value.

However, for some baseband signals, we find that something further can be done. Such a baseband signal is one which has the characteristic that its power spectral density (PSD) is relatively high in the low-frequency range and falls off rapidly at higher frequencies. For example, speech has little PSD above about 3 kHz. As a consequence, when the spectrum of the sidebands associated with a carrier which is frequency modulated by this type of baseband signal is examined, it is found that the PSD of the sidebands is greatest near the carrier and relatively small near the limits of the allowable frequency band allocated to the transmission. The method which takes advantage of these spectral features in order to improve the performance of an FM system is shown in Figure 3.8.

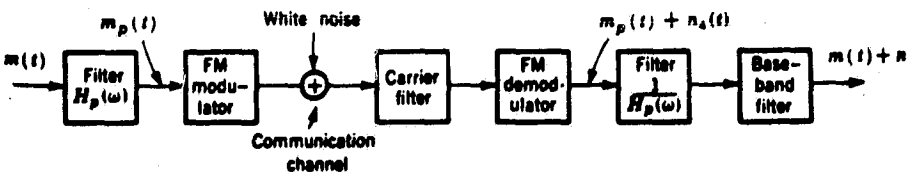


Figure 3.8. Preemphasis and Deemphasis in an FM System

In Figure 3.9 observe that, at the transmitting end, the baseband signal m(t) is not applied directly to the FM modulator, but is first passed through a filter of transfer characteristic $H_p(\omega)$, so that the modulating signal is $m_p(t)$. The modulated carrier is transmitted across a communication channel; during which process, noise is added to the signal. The receiver is a conventional discriminator except that a filter has been introduced before the baseband filter. The transfer characteristic of this filter is the reciprocal of the characteristic of the transmitter filter. This receiver filter of transfer characteristic $1/H_p(\omega)$ may equally well be placed after the baseband filter since both filters are linear. Observe that any modification introduced into the baseband signal by the first filter, prior to modulation, is presently undone by the second filter which follows the discriminator. Hence, the output signal

at the receiver is exactly the same as it would be if the filters had been omitted entirely. However, the noise passes through the receiver filter only, and this filter may then, to some extent, be used to suppress the noise.

Selection of the transfer characteristic $H_p(\omega)$ is based on the following consideration. At the output of the demodulator, the spectral density of the noise increases with the square of the frequency. Hence, the receiver filter will be most effective in suppressing noise if the response of the filter falls off with increasing frequency; that is, if the filter transmission is lowest where the spectral density of the noise is highest. In such a case, the transmitter filter must exhibit a rising response with increasing frequency. Initially, assume that the transmitter filter is designed so that it serves only to increase the spectral density of the higher frequency components of the signal m(t). Such a filter must necessarily increase the power in the modulating signal, thereby increasing the distortion above its specified maximum value. However, as noted above, the spectral density of the modulated carrier is relatively small near the edges of the allowed frequency band. Thus, such a filter may possibly raise the signal spectral density only near the edges of the allowed frequency band and cause only a small increase in distortion. In this case, if the modulating signal power is lowered to decrease the distortion to the allowed value, there is a net advantage; i.e., the improvement due to raising the spectral density near the edges of the allowable band outweighs the disadvantage due to the need to lower the level of the modulating signal. The premodulation filtering in the transmitter used to raise the PSD of the baseband signal in its upper frequency range is called preemphasis (or predistortion). The filtering at the receiver used to undo the signal preemphasis and suppress noise is called deemphasis. The spectral density of the noise at the output of an FM demodulator increases with the square of the frequency. Hence, a deemphasis network at the receiver will be most effective in suppressing noise if its response falls with increasing frequency. In commercial FM, the deemphasis is performed by the simple lowpass resistance capacitance (RC) network. This network has a transfer function $H_d(f)$ given by

$$H_d(f) = \frac{1}{1 + j(f/f_1)}, \qquad (3\text{-}3)$$

where $f_1 = 1/2\pi rC$ is the lower frequency "break-point." An inverse network is required at the transmitter.

The preemphasis network transfer function $H_p(f)$ is given by

$$H_p(f) = \frac{R}{r}(1 + j\omega Cr) = \frac{R}{r}\left(1 + j\left(\frac{f}{f_1}\right)\right); \quad r > R, \qquad (3\text{-}4)$$

where, as before, $f_1 = 1/2\pi rC$, and $f_2 = 1/2\pi RC$ is the higher frequency break-point. Hence, $H_p(f)$ has a frequency dependence inverse to $H_d(f)$ as required in order that no net distortion be introduced into the signal; thus, $H_p(f)H_d(f) = R/r = \text{constant}$.

Another premodulation-processing technique typically employed by analog signals is companding. The word "compandor" is derived from two words which describe its functions: compressor/expandor, to compress and to expand. The simplified functional diagram and its analogy are shown in Figure 3.9. The compressor compresses the intensity range of the baseband signal at the input circuit of a communication channel by imparting more gain to weak signals than to strong signals. At the far-end output of the communication circuit, the expandor performs the reverse function. It restores the intensity of the signal to its original dynamic range.

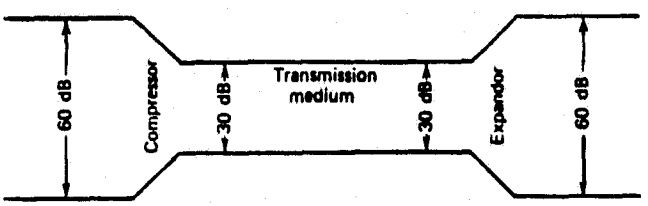


Figure 3.9. Functional Analogy of a Compandor

The three advantages of compandors are that they:

(1) Tend to improve the SNR on noisy circuits

(2) Limit the dynamic power range of signals, reducing the chances of overload of carrier systems

(3) Reduce the possibility of crosstalk.

An important parameter of a compandor is the compression/expansion ratio, which is the degree to which speech energy is compressed and expanded. It is expressed by the ratio of the input power to the output power (dB) in the compressor and expandor, respectively. Compression ratios are always greater than 1 and expansion ratios are less than 1. The most common compression ratio is 2. The corresponding expansion ratio is thus 1/2. The meaning of a compression ratio of 2 is that the dynamic range of the speech volume has been cut in half from the input of the compressor to its output.

Another important criterion for a compandor is its companding range. This is the range of intensity levels a compressor can handle at its input. Usually, 50-60 dB is sufficient to provide the expected SNR and reduce the possibility of distortion. High-level signals appearing outside this range are thereby limited without markedly affecting the intelligibility.

3.2.3 Subcarrier Modulation

Modulation of digital data on a subcarrier is usually by means of phase-shift-keying (PSK) or frequency-shift-keying (FSK). Actually, amplitude-shift-keying (ASK) could also be used but, generally, there are nonlinearities in the communication system that significantly degrade the performance of ASK. Analog data is typically modulated by frequency modulation (FM). Again, amplitude modulation (AM) could be used, but is not typically employed for subcarrier modulation.

3.2.3.1 PSK modulation

To PSK modulate a signal that takes on the values $x(t) = \pm 1$, the modulated waveform is

$$s(t) = \sqrt{2P}\,\sin\left(\omega_0 t + \frac{\pi}{2}\,x(t)\right), \qquad (3\text{-}5)$$

where P is the transmitted power and ω_0 is the subcarrier frequency. Equation (3-5) can be written in an alternate form by expanding the sine as

$$s(t) = \sqrt{2P}\,\sin\left(\frac{\pi}{2}\,x\,(t)\right)\cdot\cos\,(\omega_0 t) - \cos\left(\frac{\pi}{2}\,x\,(t)\right)\sin(\omega_0 t)$$

$$= \sqrt{2P}\,x(t)\,\cos\,\omega_0 t\ , \qquad (3\text{-}6)$$

since $\sin(\pm\pi/2) = \pm 1$ and $\cos(\pm\pi/2) = 0$. The waveform of (3-6) may be generated by applying the waveform $x(t)$ and the subcarrier $\cos\,\omega_0 t$ to a balanced modulator. A balanced modulator yields an output waveform which, aside from a constant factor, is the product of its input waveforms.

The PSK modulation can be generalized to multiple PSK (MPSK) where the signals $s_i(t)$ are given by

$$s_i(t) = \sqrt{2P}\,\cos(\omega_0 t + \theta_i)\ ; \qquad 0 \leq t \leq T\ , \qquad (3\text{-}7)$$

with $\theta_i = 2(i-1)\pi/M$ for $i = 1,2,3,\ldots,M$. Note that, for $M = 2$, $s_1(t)$ and $s_2(t)$, the antipodal (PSK) signals are given in (3-5). The case $M = 4$ produces quadriphase-shift-keying (QPSK), and $M = 8$ yields octaphase-shift-keying. These three signal sets are depicted in Figure 3.10, along with the decision thresholds that are implemented in the optimum receiver.

The spectrum of a PSK signal is the same as shown in Figures 3.3 through 3.5, depending on the format of the binary signal $x(t)$. The only difference is that the spectra are controlled about the frequency ω_0. The spectrum of an MPSK signal with $N > 4$ is most closely approximated by the NRZ code spectrum centered at ω_0, and the nulls of the spectrum are at $\omega_0 \pm (2\pi k)/T$, where $T = T_1 \log_2 M$ and k is any integer.

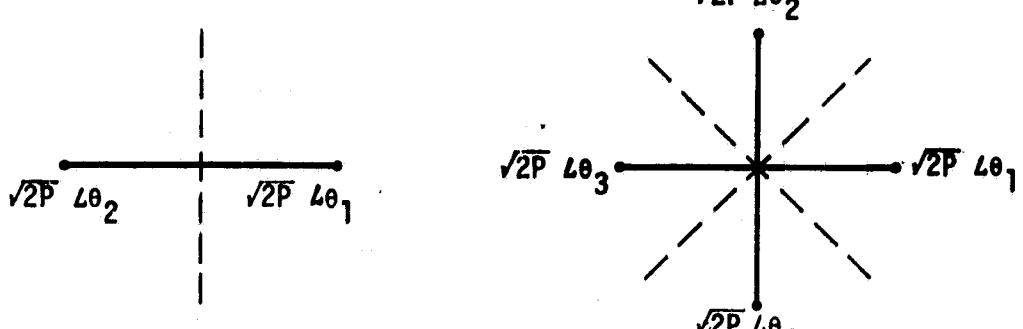
3.2.3.2 FSK modulation

In FSK, the binary signal $x(t)$ is used to generate a waveform

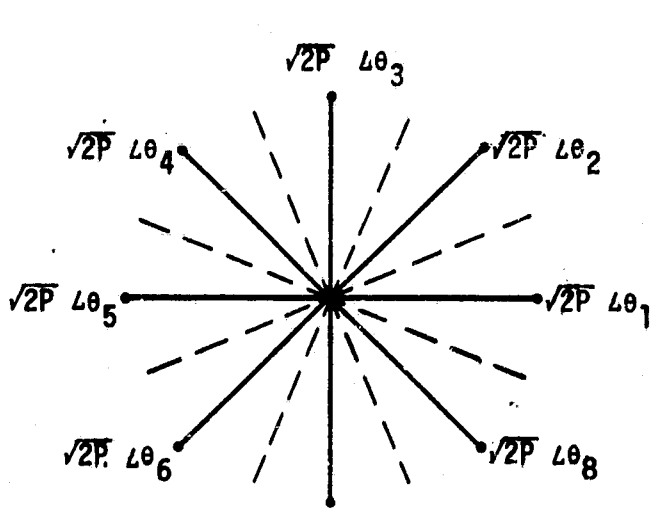
$$s(t) = \sqrt{2P}\,\cos(\omega_0 - \Omega_i)t\ ; \qquad 0 \leq t \leq T\ , \qquad (3\text{-}8)$$

in which the plus sign or minus sign applies, depending on whether $x(t)$

Dotted lines represent decision thresholds in the receiver.



(a) Antipodal Signals (M=2) (b) Quadriphase Signals (M=4)



(c) Octaphase Signals (M = 8)

Figure 3.10. Signal Vectors for Various Polyphase Signal Sets

is +1 or -1. The transmitted signal is then of amplitude $\sqrt{2P}$ and has an angular frequency $\omega_0 + \Omega$ or $\omega_0 - \Omega$, with Ω a constant angular frequency offset from the center frequency ω_0. For the FSK signals s(t) to be orthogonal over the bit duration T, Ω must equal $(\pi k)/T$, where k is any integer. The spectrum of FSK beyond the first null is the sum of two NRZ spectra, with one centered at $\omega_0 + \Omega$ and the other at $\omega_0 - \Omega$. To minimize the required bandwidth to transmit an FSK signal, Ω should be chosen to equal π/T.

The generalization of FSK is multiple FSK (MFSK), where the signal set is given by

$$s_i(t) = \sqrt{2P} \cos(\omega_0 - \Omega_i)t \; ; \quad 0 \leq t \leq T , \tag{3-9}$$

in which $\Omega_i = (M/2 - i)\pi/T$ for M even. The spectrum of MFSK is the sum of M NRZ spectra centered at $\omega_0 - \Omega_i$, where $T = T_1 \log_2 M$, with T_1 shown in Figure 3.3.

3.2.3.3 Subcarrier PM and FM

The basic signal form of PM and FM is given by

$$s(t) = \sqrt{2P} \cos\left(\omega_0 t + \theta(t)\right) \tag{3-10}$$

where, for an input signal x(t), the instantaneous phase angle for the PM signal is

$$\theta(t) = m_p x(t) , \tag{3-11}$$

with m_p defined as the phase modulation index. For the FM signal,

$$\theta(t) = m_f \int_0^t x(t)\, dt, \tag{3-12}$$

where m_f is defined as the modulation index for FM. Note that, in PM, the instantaneous phase of the modulated signal depends linearly upon the modulating signal x(t) while, in FM, the instantaneous frequency varies linearly with x(t), namely,

$$\omega_i(t) = \omega_0 + m_f x(t) , \tag{3-13}$$

which is equivalent to the statement that the phase depends linearly upon the integral of x(t). As an example, when the modulating signal is a signal-tone sinusoid so that $x(t) = \cos \omega_m t$, the PM signal is

$$s_p(t) = \sqrt{2P} \cos(\omega_0 t + m_p \cos \omega_m t) \tag{3-14}$$

and the FM signal is

$$s_f(t) = \sqrt{2P} \cos\left(\omega_0 t + \frac{m_f}{\omega_m} \sin \omega_m t\right). \tag{3-15}$$

This is illustrated in Figure 3.11, where the modulating signal is a sinusoid; amplitude modulation of the carrier is also included for comparison. It should be noted here that a distinction can be made between the phase- and frequency-modulated carrier only when compared with the modulating signal, as illustrated in Figure 3.11. Also note that, for phase modulation, the peak phase deviation or modulation index m_p is a constant independent of the frequency of the modulating signal and, for frequency modulation, the peak phase deviation is given by m_f/ω_m, which is inversely proportional to the frequency of the modulating signal.

The spectral analysis of PM and FM can best be performed by considering specific examples of $\theta(t)$. As the first example, let $\theta(t) = \beta \sin \omega_m t$. Therefore,

$$s(t) = \sqrt{2P} \cos(\omega_0 t + \beta \sin \omega_m t) , \tag{3-16}$$

where β is the peak phase deviation of the modulated signal of angular frequency ω_m. Hence, the instantaneous frequency $\omega_i(t)$ is equal to

$$\omega_i(t) = \omega_0 + \omega_m \beta \cos \omega_m t = \omega_0 + \Delta\Omega \cos \omega_m t , \tag{3-17}$$

where $\Delta\Omega = \omega_m \beta$ is the maximum frequency deviation of the modulated frequency ω_0. If $\beta \ll 1$, then s(t) is a narrowband PM signal and

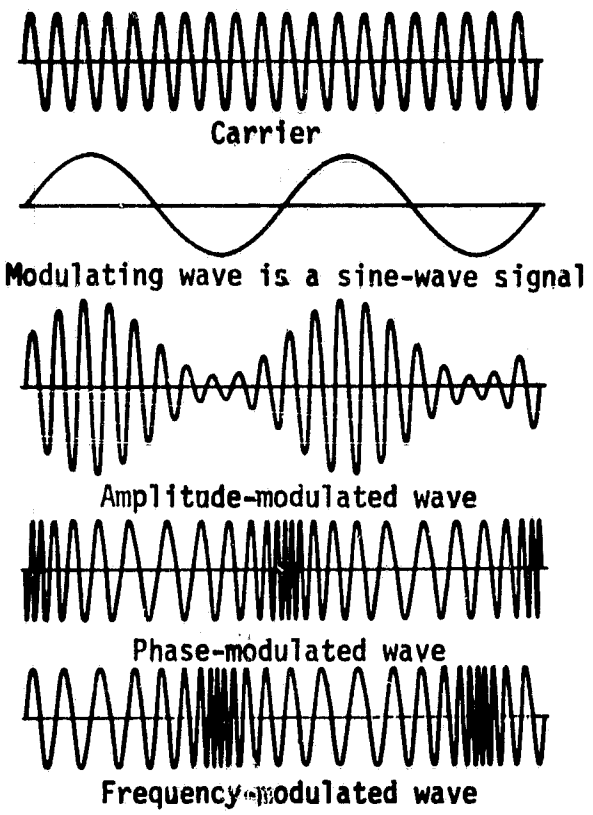


Figure 3.11. Amplitude, Phase and Frequency Modulation of a Sine-Wave Carrier by a Sine-Wave Signal

$$s(t) = \sqrt{2P}\left[\cos\omega_0 t \cos(\beta \sin\omega_m t) - \sin\omega_0 t \sin(\beta \sin\omega_m t)\right]$$

$$= \sqrt{2P}\left[\cos\omega_0 t - \beta \sin\omega_m t \sin\omega_0 t\right] \qquad \beta << 1$$

$$= \sqrt{2P}\left[\cos\omega_0 t - \frac{\beta}{2}\cos(\omega_0 - \omega_m)t + \frac{\beta}{2}\cos(\omega_0 + \omega_m)t\right]. \qquad (3\text{-}18)$$

From (3-18), note that the spectrum of narrowband PM consists of a carrier and two sidebands which are 180° out of phase.

If β is not small, then $s(t)$ is a wideband PM signal and

$$s(t) = \sqrt{2P}\left[\cos\omega_0 t \cos(\beta \sin\omega_m t) - \sin\omega_0 t \sin(\beta \sin\omega_m t)\right]$$

$$= \sqrt{2P}\left[\cos\omega_0 t \sum_{n=-\infty}^{\infty} J_n(\beta)\cos(n\omega_m t) - \sin\omega_0 t \sum_{n=-\infty}^{\infty} J_n(\beta)\sin(n\omega_m t)\right]$$

$$= \sqrt{2P}\Big\{J_0(\beta)\cos\omega_0 t - J_1(\beta)\left[\cos(\omega_0 - \omega_m)t - \cos(\omega_0 + \omega_m)t\right]$$

$$+ J_2(\beta)\left[\cos(\omega_0 - 2\omega_m)t + \cos(\omega_0 + 2\omega_m)t\right]$$

$$- J_3(\beta)\left[\cos(\omega_0 - 3\omega_m)t + \cos(\omega_0 + 3\omega_m)t\right]$$

$$+ \cdots\Big\} \qquad (3\text{-}19)$$

Thus, the wideband PM signal consists of a carrier and an infinite number of sidebands whose amplitudes are proportional to $J_n(\beta)$ spaced at frequencies $\pm\omega_m$, $\pm 2\omega_m$,...$(\pm n\omega_m)$ away from the carrier, as shown in Figure 3.12.

In practice, the spectra of PM and FM signals are not infinite because, beyond a certain frequency range from the carrier, depending on the magnitude of β, the sideband amplitudes which are proportional to $J_n(\beta)$ become negligibly small. Thus, the bandwidth of the transmitter and receiver can be restricted to encompass only the significant bands without introducing an excessive amount of harmonic distortion.

In order to compare FM to the spectral analysis of PM, note that, assuming a sinusoidal modulating signal $\cos\omega_m t$ with maximum frequency deviation $\Delta\Omega$, the instantaneous frequency is given by

$$\omega_i(t) = \omega_0 + \Delta\Omega \cos\omega_m t\,; \qquad \Delta\Omega << \omega_0\,, \qquad (3\text{-}20)$$

where the frequency deviation $\Delta\Omega$ is independent of the modulating frequency and is proportional to the amplitude of the modulating signal.

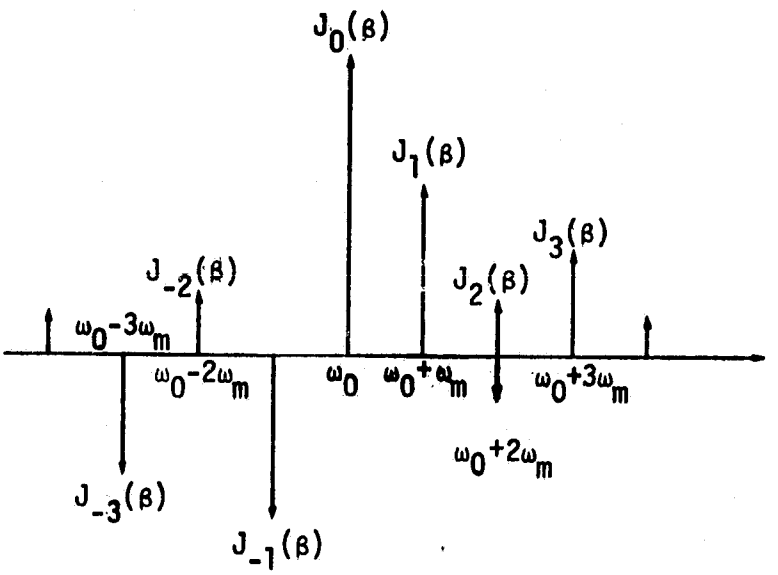


Figure 3.12. Composition of FM Wave into Sidebands

The instantaneous phase angle $\theta_i(t)$ for this special case is given by

$$\theta_i(t) = \int_0^t \omega_i(t)\, dt = \omega_0 t + \frac{\Delta\Omega}{\omega_m} \sin \omega_m t + \theta_0 , \qquad (3\text{-}21)$$

where θ_0 may be taken as zero by referring to an appropriate phase reference so that the FM signal is given by

$$s(t) = \sqrt{2P} \cos(\omega_0 t + \beta \sin \omega_m t) , \qquad (3\text{-}22)$$

where β is inversely proportional to the modulating angular frequency ω_m; namely, $\beta = \Delta\Omega/\omega_m$, which is of the same form as (3-16) for PM. However, it should be emphasized that, while (3-22) and (3-16) are of the same form, the respective modulation indices β assume different significance. In PM, the maximum phase deviation β is constant, independent of the modulating angular frequency ω_m, but, in FM, the maximum frequency deviation $\Delta\Omega$ is constant, independent of ω_m.

Since the expressions representing phase- and frequency-modulated signals are identical, as given by (3-22) and (3-16), it follows that, for a single modulating frequency, the spectral representations are identical. However, it is of interest to examine the spectral behavior as ω_m is varied, $\Delta\Omega$ being constant. As shown in Figure 3.13, as $\beta \to \infty$, the number of sidebands increases and the spectral components become more and more confined to the band between $\omega_0 \pm \Delta\Omega$.

The bandwidth is equal to $2\Delta f_0 (\Delta\Omega = 2\pi\Delta f_0)$ only for a very large modulation index. For smaller values of β, we determine the bandwidth by counting the significant number of sidebands. The word "significant" is usually taken to mean those sidebands which have a magnitude at least 1% of the magnitude of the unmodulated carrier.

It has been shown that, for $\beta << 1$, the spectrum of an FM signal is essentially a carrier and one pair of significant sidebands. As β increases, the number of significant sidebands increases, while the total average power of the sidebands plus the carrier remains constant, equal to P. Since the amplitudes of the carrier and sidebands are proportional to the Bessel function $J_n(\beta)$, where $n = 0, \pm 1, \pm 2, \ldots$, the variation of the amplitudes as a function of the modulation index β can be obtained from the plot of the Bessel functions of the first kind, as shown in Figure 3.14.

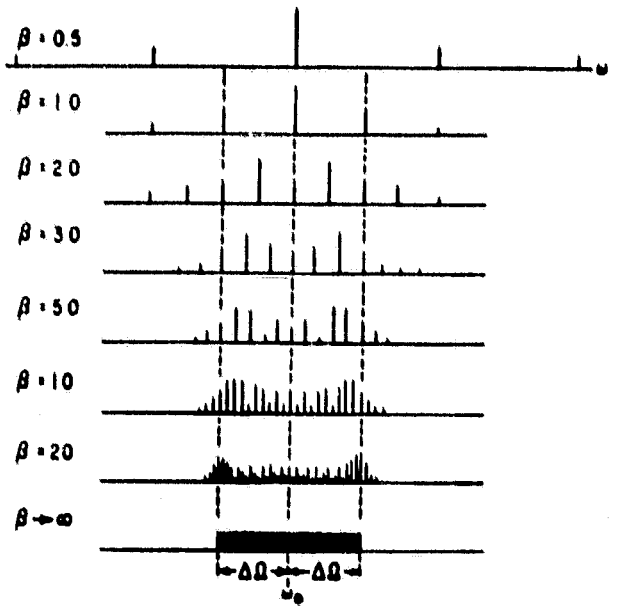


Figure 3.13. FM Spectrum of Single-Tone Modulation Bandwidth Versus Modulation Index β

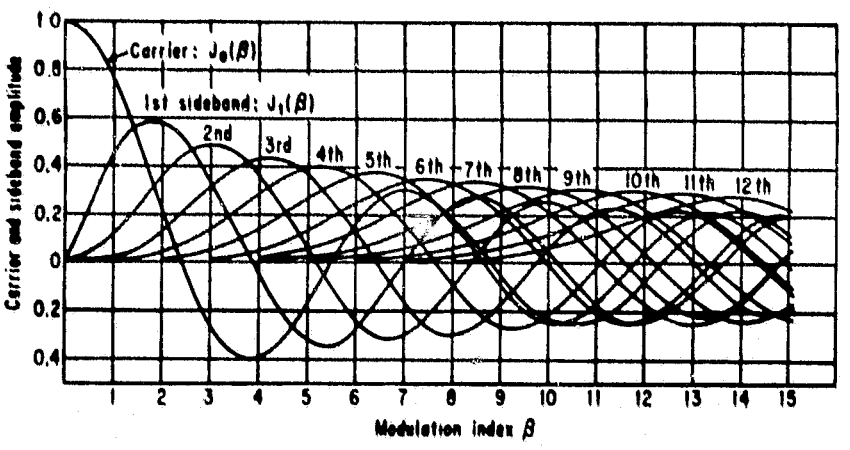


Figure 3.14. Plot of Bessel Function of the First Kind as a Function of Argument β

The behavior of the carrier amplitude is determined from the zero-order Bessel function $J_0(\beta)$. As β is increased, the value of the Bessel function drops off rapidly until, as $\beta = 2.404$, the amplitude is zero. As seen from the plot, the zero-order Bessel function is oscillatory with decreasing peak amplitude, the spacing between zeros asymptotically approaching the constant value π. This follows from the approximation

$$J_0(\beta) = \frac{\cos(\beta - \pi/4)}{\sqrt{\beta\pi/2}} . \qquad (3\text{-}23)$$

An important modulating signal to be considered for FM is a binary data stream, which is another way of generating FSK. However, when FSK is generated using FM, the phase is continuous at each bit transition, and the resulting spectrum is different than the FSK discussed in subsection 3.2.3.2. This form of FSK is referred to as continuous phase FSK (CPFSK). As an example of the spectrum of CPFSK generated by FM, consider a square-wave FM signal, shown in Figure 3.15.

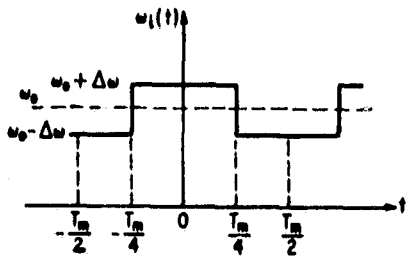


Figure 3.15. Frequency Modulation by a Square Wave

The instantaneous frequency can be described as follows:

$$\omega_i(t) = \omega_0 - \Delta\omega, \quad -\frac{T_m}{2} < t < -\frac{T_m}{4}$$

$$\omega_i(t) = \omega_0 + \Delta\omega, \quad -\frac{T_m}{4} < t < \frac{T_m}{4}$$

$$\omega_i(t) = \omega_0 - \Delta\omega, \quad \frac{T_m}{4} < t < \frac{T_m}{2} . \qquad (3\text{-}24)$$

The corresponding instantaneous phase variation may be described as

$$\theta(t) = \begin{cases} -\frac{\Delta\omega}{\omega_m}(\pi + \omega_m t), & -\frac{T_m}{2} < t < -\frac{T_m}{4} \\ \Delta\omega t, & -\frac{T_m}{4} < t < \frac{T_m}{4} \\ \frac{\Delta\omega}{\omega_m}(\pi - \omega_m t), & \frac{T_m}{4} < t < \frac{T_m}{2} \end{cases} \qquad (3\text{-}25)$$

Figure 3.16 illustrates the spectrum of the square-wave FM signal where $\Delta\omega$ is kept constant and ω_m is varied to show the square-wave spectrum amplitudes. For low $\beta(\omega_m > \Delta\omega)$, the energy is now distributed among the carrier and the first pair of sidebands. As β is increased, more and more energy is contributed by the sidebands in the vicinity of $(\omega_0 \pm \Delta\omega)$. As $\beta \to \infty$, which corresponds to the wave frequency having an infinitely long half-period at both $(\omega_0 + \Delta\omega)$ and $(\omega_0 - \Delta\omega)$, the energy and amplitude spectra will consist of only two spectral lines, indicating that all the energy is at $(\omega_0 \pm \Delta\omega)$.

In working with the complicated mathematical expressions often involved in complex-wave and multitone spectrum analysis, it is extremely important to note that the spectral contributions of an FM wave are functions of the rate of change of frequency during the entire cycle; the longer the frequency of an FM wave remains in a certain range of frequencies, the greater will be the spectral contribution in that frequency range. This is clearly shown in Figure 3.17, where several complex-wave FM waveforms are shown with typical corresponding spectra resulting from large modulation indices.

3.2.3.4 Subcarrier modulation losses

When multiple data streams are to be modulated on the same subcarrier, there is a performance loss. This loss is the result of two effects. First, if more than one data stream is modulated onto a subcarrier, then the available power must be divided between the data streams. The second contribution to the subcarrier performance loss is the spectral restrictions placed on the data streams when modulated onto the subcarrier. Two digital data streams can easily be modulated onto the same subcarrier by PSK modulating the data streams onto the subcarrier separately and

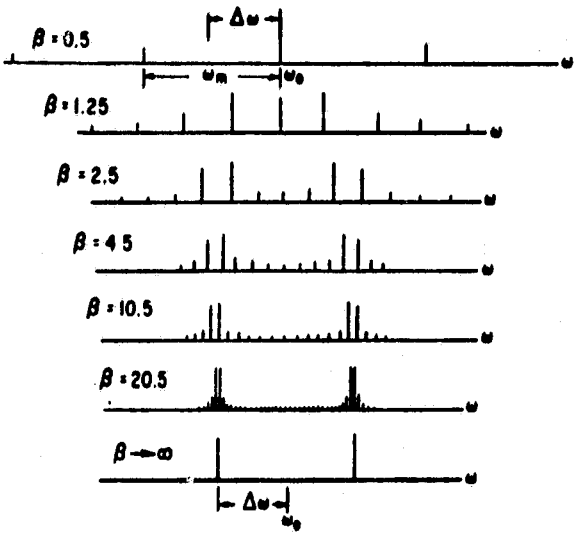


Figure 3.16. Spectrum Behavior of a Square-Wave FM Wave as β is Increased with Δω Constant

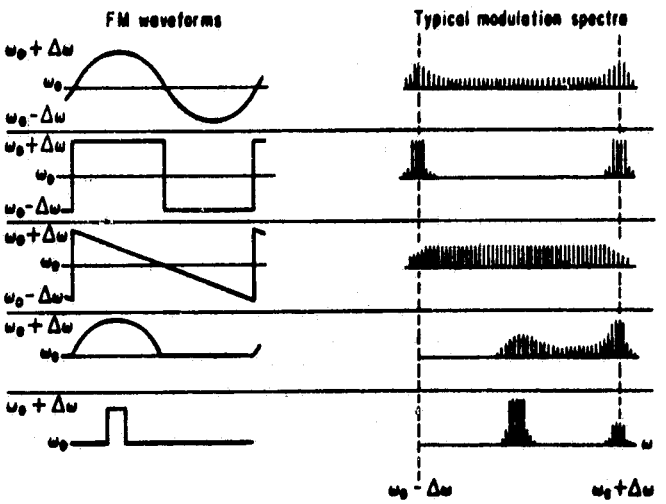


Figure 3.17. FM Waveforms and Corresponding Typical Modulation Spectra

then summing the two modulated signals in quadrature, as shown in Figure 3.18. Similarly, using narrowband PM (or FM), two analog data streams or one analog and one digital data stream can be modulated onto a single subcarrier. When narrowband PM (or FM) or PSK for digital data is used to combine two data streams, there is no other subcarrier performance loss than that due to the division of power. However, if more than two data streams are combined on a single subcarrier or wideband PM (or FM) is used to modulate the data, then the data streams must not have spectra that overlap. Therefore, if any filtering is performed on the data streams in order to avoid spectral overlap, the amplitude and phase distortion introduced by the filtering, as described in subsection 3.2.1.2 must be included as a contribution to the subcarrier performance loss.

3.2.4 Carrier Modulation

The modulation of the carrier is similar to the modulation of the subcarrier by analog and digital data. Analog data and data-modulated subcarriers can be modulated onto the carrier using PM or FM. If all the baseband data is digital and the subcarriers are square waves that have been multiplied by the digital data, then the data and modulated subcarriers can be modulated onto the carrier using PSK or QPSK (or, in general, MPSK). Finally, the modulated carrier can be spread using spread spectrum techniques to achieve low detectable waveforms, multiple-access capability or interference protection.

3.2.4.1 Carrier phase modulation

This section is concerned with communication systems that phase modulate an RF carrier with multiple modulated data subcarriers. As an example, first consider two modulated subcarriers. The frequency spectra of these two subcarriers must be appropriately selected since the modulation technique requires that they be nonoverlapping (i.e., lie in separate portions of the RF channel) and situated outside the passband of the RF carrier-tracking loop. This particular concept is extremely important in assuring successful demodulation of telemetry and/or command data while simultaneously performing the tracking function. Figure 3.19 illustrates a typical frequency spectrum of the transmitted signal in a two-channel system.

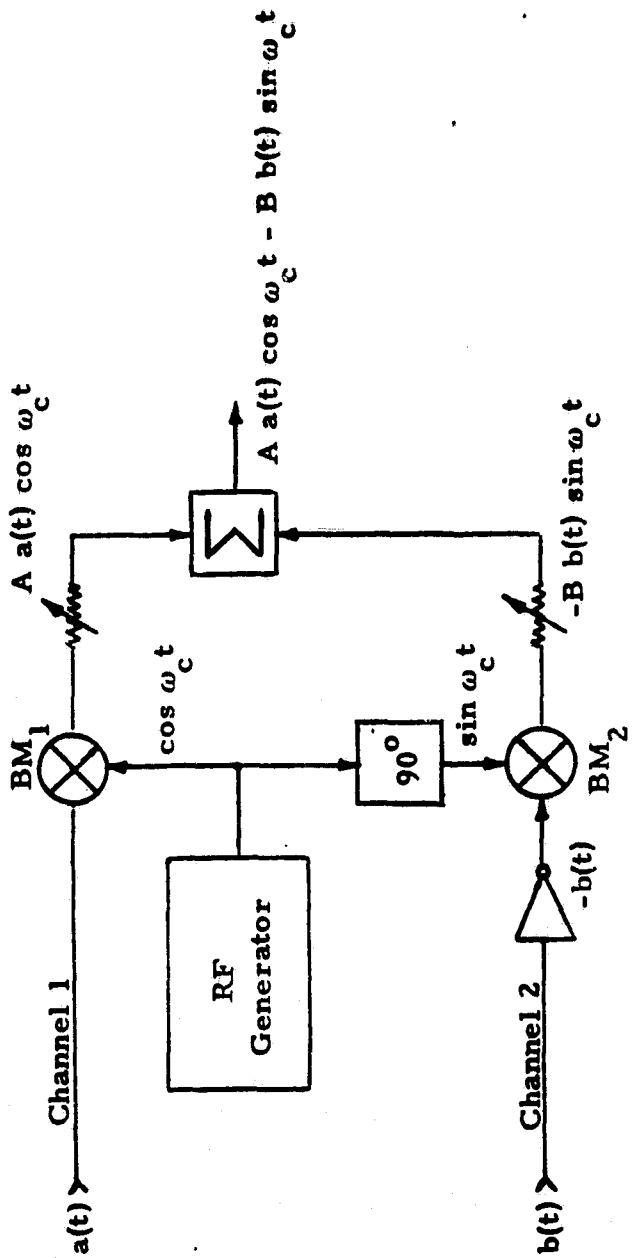


Figure 3.18. Two-Channel Quadriphase Modulator

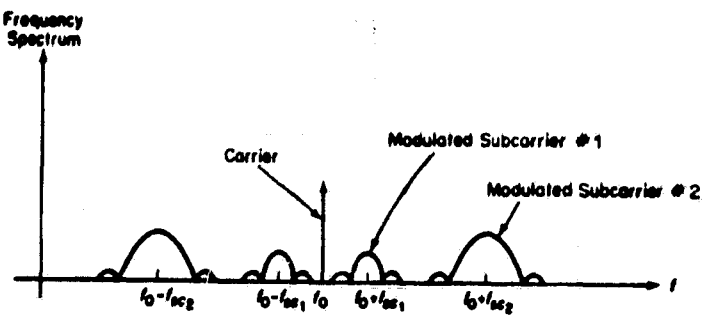


Figure 3.19. Frequency Spectrum of an RF Carrier Which is Phase Modulated by Two Modulated Subcarriers

The carrier modulated with two subcarriers is given by:

$$s(t) = \sqrt{2P}\cos\left[\omega_0 t + \beta_1 y_1(t) + \beta_2 y_2(t)\right], \qquad (3\text{-}26)$$

where $y_1(t)$ is data-modulated subcarrier 1 and $y_2(t)$ is data-modulated subcarrier 2, β_1 and β_2 are the peak phase deviations in radians for subcarriers 1 and 2, respectively, ω_0 is the carrier frequency, and P is the transmitted power.

As an example of the resulting spectral properties of a phase-modulated carrier, consider the two sinusoidal modulations as subcarriers to phase modulate the carrier. That is,

$$y_1(t) = \sin \omega_1 t$$

$$y_2(t) = \sin \omega_2 t \qquad (3\text{-}27)$$

Substituting (3-27) into (3-26),

$$s(t) = \sqrt{2P}\cos\left[\omega_0 t + \beta_1 \sin \omega_1 t + \beta_2 \sin \omega_2 t\right] \qquad (3\text{-}28)$$

Expanding the cosine term,

$$s(t) = \sqrt{2P}\Big\{\cos\omega_0 t\Big[\cos(\beta_1 \sin\omega_1 t)\cos(\beta_2 \sin\omega_2 t) - \sin(\beta_1 \sin\omega_1 t)\sin(\beta_2 \sin\omega_2 t)\Big] - \sin\omega_0 t\Big[\sin(\beta_1 \sin\omega_1 t)\cos(\beta_2 \sin\omega_2 t) + \sin(\beta_2 \sin\omega_2 t)\cos(\beta_1 \sin\omega_1 t)\Big]\Big\} \qquad (3\text{-}29)$$

and using Bessel function series expansion for the terms within the brackets,

$$s(t) = \sqrt{2P}\Bigg\{\cos\omega_0 t\Bigg\{\Big[J_0(\beta_1) + 2\sum_{\substack{n=2\\ \text{even}}}^{\infty} J_n(\beta_1)\cos n\omega_1 t\Big] \times \Big[J_0(\beta_2) + 2\sum_{\substack{n=2\\ \text{even}}}^{\infty} J_n(\beta_2)\cos n\omega_2 t\Big] - \Big[2\sum_{\substack{n=1\\ \text{odd}}}^{\infty} J_n(\beta_1)\sin n\omega_1 t\Big]\Big[2\sum_{\substack{n=1\\ \text{odd}}}^{\infty} J_n(\beta_2)\sin n\omega_2 t\Big]\Bigg\} - \sin\omega_0 t\Bigg\{\Big[2\sum_{\substack{n=1\\ \text{odd}}}^{\infty} J_n(\beta_1)\sin n\omega_1 t\Big]\Big[J_0(\beta_2) + 2\sum_{\substack{n=2\\ \text{even}}}^{\infty} J_n(\beta_2)\cos n\omega_2 t\Big] + \Big[2\sum_{\substack{n=1\\ \text{odd}}}^{\infty} J_n(\beta_2)\sin n\omega_2 t\Big] \times \Big[J_0(\beta_1) + 2\sum_{\substack{n=2\\ \text{even}}}^{\infty} J_n(\beta_1)\cos n\omega_1 t\Big]\Bigg\}\Bigg\} \qquad (3\text{-}30)$$

Note that, when $y_2(t)$ is zero (i.e., a single subcarrier to phase modulate the carrier), (3-30) reduces to the same form as (3-19), which was for a subcarrier phase modulated by a sinusoid.

Another important example of carrier PM is for a sinusoidal signal representing analog data and a square-wave signal representing digital data to be modulating the carrier. That is,

$$y_1(t) = \sin \omega_1 t$$

$$y_2(t) = \text{sq}\ \omega_2 t\,. \qquad (3\text{-}31)$$

Therefore,

$$s(t) = \sqrt{2P}\cos\left[\omega_0 t + \beta_s \sin \omega_1 t + \beta_d\ \text{sq}\ \omega_2 t\right], \qquad (3\text{-}32)$$

where β_s is the peak phase deviation for the sinusoidal signal and β_d is the peak phase deviation for the digital signal. Expanding (3-32) using Bessel functions gives

$$\begin{aligned} s(t) = \sqrt{2P}\Bigg\{ & \cos \omega_0 t \Bigg\{ \left[J_0(\beta_s) + 2 \sum_{\substack{n=2 \\ \text{even}}}^{\infty} J_n(\beta_s) \cos n\omega_1 t \right] \cos \beta_d \\ & - 2 \sum_{\substack{n=1 \\ \text{odd}}}^{\infty} J_n(\beta_s) \sin n\omega_1 t \sin \beta_d\ \text{sq}\ \omega_2 t \Bigg\} \\ & - \sin \omega_0 t \Bigg\{ 2 \sum_{\substack{n=1 \\ \text{odd}}}^{\infty} J_n(\beta_s) \sin \omega_1 t \cos \beta_d \\ & + \text{sq}\ \omega_2 t \sin \beta_d \left[J_0(\beta_s) + 2 \sum_{\substack{n=2 \\ \text{even}}}^{\infty} J_n(\beta_s) \cos n\omega_1 t \right] \Bigg\} \Bigg\}. \end{aligned} \qquad (3\text{-}33)$$

Equation (3-33) presents the spectral properties of carrier PM for two communication modes on the Orbiter. The first is the forward-link signal when data plus ranging is modulated on the carrier. The second case is the narrowband "bent-pipe" mode for nonstandard payload telemetry signals from the payload to the Orbiter to be transmitted to the ground.

3.2.4.2 Carrier PSK and QPSK

The general form of carrier PM with PSK modulation on all the subcarrier data streams is

$$s(t) = \sqrt{2P}\sin\left[\omega_c t + \theta(t)\right], \qquad (3\text{-}34)$$

where ω_c is the carrier frequency and the combined PM signal resulting from N channels is defined by

$$\theta(t) = \sum_{n=1}^{N} \beta_n s_n(t) , \qquad (3\text{-}35)$$

where β_n is the modulation angle associated with the nth channel and

$$s_n(t) = d_n(t)\ sq\ (\omega_n t) = \pm 1 , \qquad (3\text{-}36)$$

which defines the nth digital data stream $d_n(t)$ which is PSK modulated on a square-wave subcarrier of frequency ω_n.

As an example of a PSK/PM system, Figure 3.20 presents the modulator for a three-channel system. For the case of three channels, the signal is

$$s(t) = \sqrt{2P}\ \sin\left[\omega_c t + \beta_1 s_1(t) + \beta_2 s_2(t) + \beta_3 s_3(t)\right] . \qquad (3\text{-}37)$$

Equation (3-37) can be rewritten in the form which shows the orthogonal components of the RF signal:

$$\begin{aligned} s(t) &= \sqrt{2P}\ \sin \omega_c t\ \cos\left[\beta_1 s_1(t) + \beta_2 s_2(t) + \beta_3 s_3(t)\right] \\ &\quad + \sqrt{2P}\ \cos \omega_c t\ \sin\left[\beta_1 s_1(t) + \beta_2 s_2(t) + \beta_3 s_3(t)\right] \\ &= \sqrt{2P}\ \sin \omega_c t\ \cos \theta(t) + \sqrt{2P}\ \cos \omega_c t\ \sin \theta(t) . \end{aligned} \qquad (3\text{-}38)$$

The trigonometric manipulations of the cos θ(t) and sin θ(t) terms results in detailed expansion of (3-38) into its components:

$$\begin{aligned} s(t) = \sqrt{2}\ \cos \omega_c t \Big[& \sqrt{P_1}\ s_1(t) + \sqrt{P_2}\ s_2(t) + \sqrt{P_3}\ s_3(t) - \sqrt{P_{cm}(1,2,3)} \\ & \times s_1(t)s_2(t)s_3(t)\Big] \\ + \sqrt{2}\ \sin \omega_c t \Big[& \sqrt{P_c} - \sqrt{P_{cm}(2,3)}\ s_2(t)s_3(t) - \sqrt{P_{cm}(1,2)} \\ & \times s_1(t)s_2(t) - \sqrt{P_{cm}(1,3)}\ s_1(t)s_3(t)\Big] \end{aligned} \qquad (3\text{-}39)$$

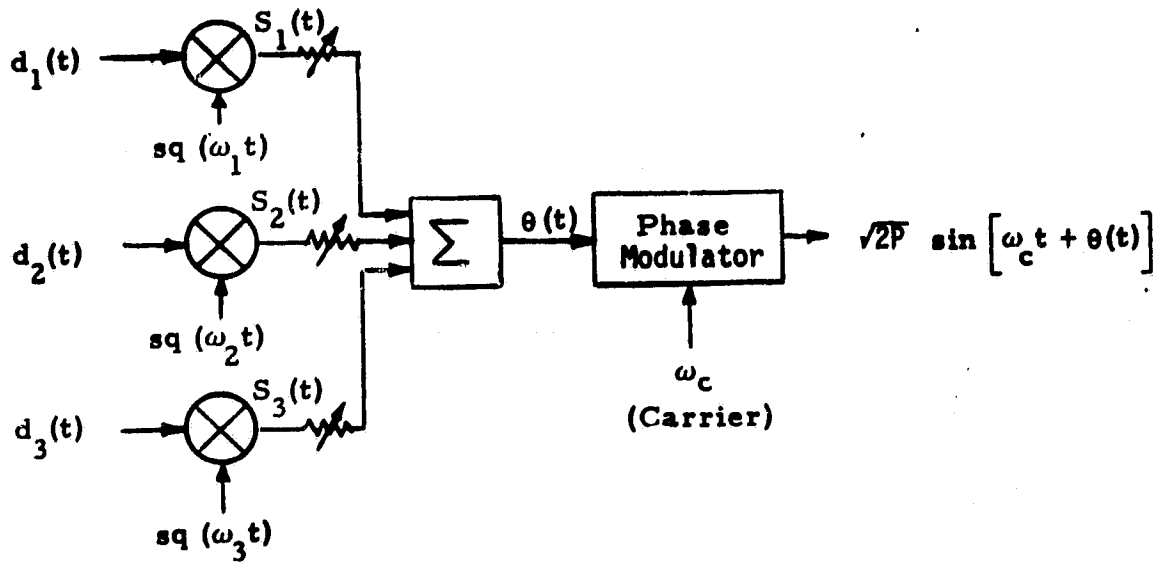


Figure 3.20. Three-Channel PSK/PM Modulator

where

$$P_1 = P \sin^2 \beta_1 \cos^2 \beta_2 \cos^2 \beta_3 = \text{Power in channel 1}$$

$$P_2 = P \cos^2 \beta_1 \sin^2 \beta_2 \cos^2 \beta_3 = \text{Power in channel 2}$$

$$P_3 = P \cos^2 \beta_1 \sin^2 \beta_2 \sin^2 \beta_3 = \text{Power in channel 3}$$

$$P_c = P \cos^2 \beta_1 \cos^2 \beta_2 \cos^2 \beta_3 = \text{Unmodulated carrier power}$$

$$P_{cm}(1,2,3) = P \sin^2 \beta_1 \sin^2 \beta_2 \sin^2 \beta_3 = \text{Third-order cross modulation}$$

$$\left.\begin{aligned} P_{cm}(1,2) &= P \sin^2 \beta_1 \sin^2 \beta_2 \cos^2 \beta_3 \\ P_{cm}(2,3) &= P \cos^2 \beta_1 \sin^2 \beta_2 \sin^2 \beta_3 \\ P_{cm}(1,3) &= P \sin^2 \beta_1 \cos^2 \beta_2 \sin^2 \beta_3 \end{aligned}\right\} = \text{Second-order cross-modulation powers} \qquad (3\text{-}40)$$

Note the obvious waste of RF power due to the cross-modulation terms. This loss must be included as part of the modulation loss for each channel.

With an interplex implementation of a PSK/PM system [7], the major cross products of the conventional system are converted into a different set of modulation components by multiplying all of the secondary channels by the signal generated by the primary channel. The primary channel is generally the channel which carries the highest data rate and thus has the major portion of the RF power assigned to it. Thus, for an N-channel interplex system, the phase modulation term $\theta(t)$ becomes

$$\theta(t) = \beta_1 s_1(t) + \sum_{n=2}^{N} \beta_n s_1(t) s_n(t) . \qquad (3\text{-}41)$$

Figure 3.21 shows the implementation of a three-channel interplex modulator. The corresponding carrier PSK/PM signal is

$$s(t) = \sqrt{2P} \sin\left[\omega_c t + \beta_1 s_1(t) + \beta_2 s_1(t) s_2(t) + \beta_3 s_1(t) s_3(t)\right] , \qquad (3\text{-}42)$$

where all of the terms are as defined earlier.

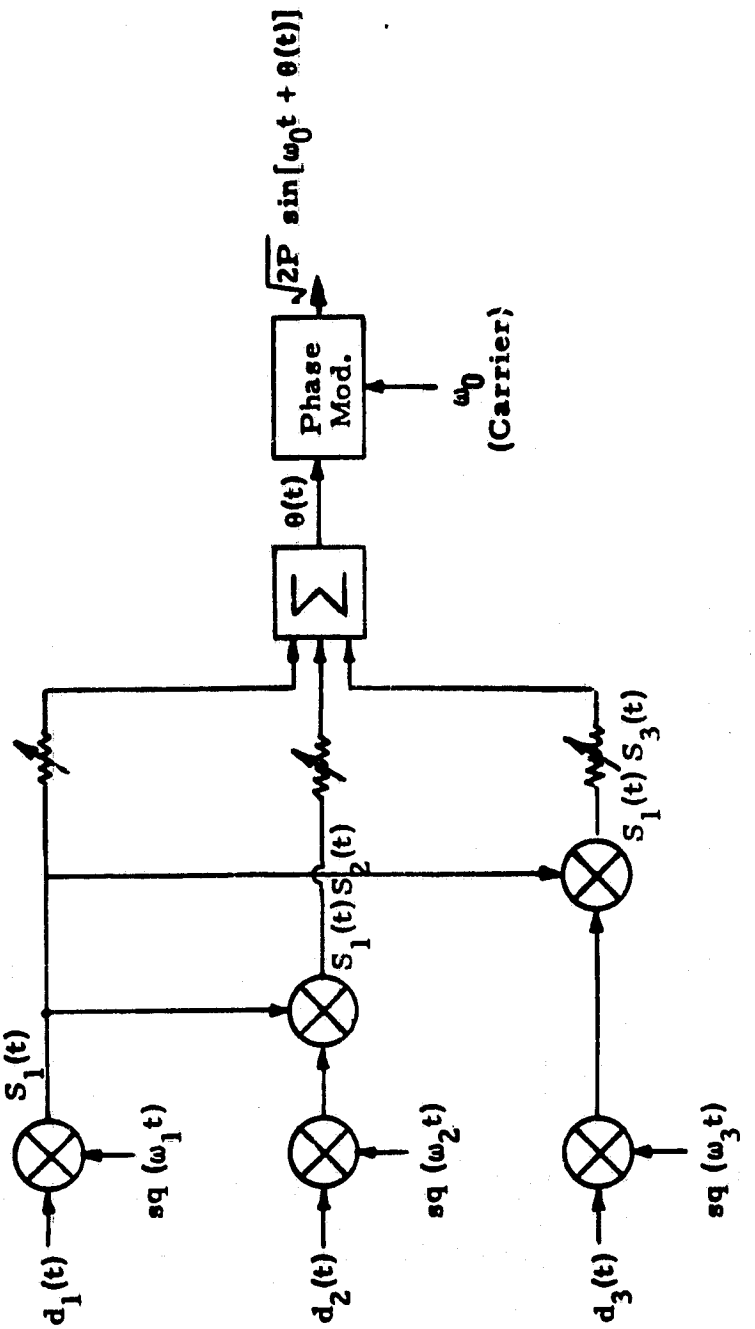


Figure 3.21. Three-Channel Interplex Modulator

Rewriting (3-42) in terms of its orthogonal RF components, one obtains

$$s(t) = \sqrt{2P}\sin\omega_c t\cos\left[\beta_1 S_1(t) + \beta_2 S_1(t) S_2(t) + \beta_3 S_1(t) S_3(t)\right]$$
$$+ \sqrt{2P}\cos\omega_c t\sin\left[\beta_1 S_1(t) + \beta_2 S_1(t) S_2(t) + \beta_3 S_1(t) S_3(t)\right]$$
$$= \sqrt{2P}\sin\omega_c t\cos\theta^*(t) + \sqrt{2P}\cos\omega_c t\sin\theta^*(t), \qquad (3\text{-}43)$$

where $\theta^*(t)$ indicates interplex time-varying modulation angle.

Following the same procedure as for conventional PSK/PM, (3-43) is expanded into its power term components. This results in

$$s(t) = \sqrt{2}\cos\omega_c t\left[\sqrt{P_1}\, S_1(t) - \sqrt{P_{cm}(1,2,3)}\, S_1(t) S_2(t) S_3(t)\right.$$
$$\left. + \sqrt{P_{cm}(1,2)}\, S_1(t) S_2(t) + \sqrt{P_{cm}(1,3)}\, S_1(t) S_3(t)\right]$$
$$+ \sqrt{2}\sin\omega_c t\left[\sqrt{P_c} - \sqrt{P_{cm}(2,3)}\, S_2(t) S_3(t) - \sqrt{P_2}\, S_2(t) - \sqrt{P_3}\, S_3(t)\right] \qquad (3\text{-}44)$$

where

$$P_1 = P\sin^2\beta_1\cos^2\beta_2\cos^2\beta_3 = \text{Power in channel 1}$$

$$P_2 = P\sin^2\beta_1\sin^2\beta_2\cos^2\beta_3 = \text{Power in channel 2}$$

$$P_3 = P\sin^2\beta_1\cos^2\beta_2\sin^2\beta_3 = \text{Power in channel 3}$$

$$P_c = P\cos^2\beta_1\cos^2\beta_2\cos^2\beta_3 = \text{Unmodulated carrier power}$$

$$P_{cm}(1,2,3) = P\sin^2\beta_1\sin^2\beta_2\sin^2\beta_3 = \text{Third-order cross-modulation power}$$

$$\left.\begin{aligned} P_{cm}(1,2) &= P\cos^2\beta_1\sin^2\beta_2\cos^2\beta_3 \\ P_{cm}(2,3) &= P\cos^2\beta_1\sin^2\beta_2\sin^2\beta_3 \\ P_{cm}(1,3) &= P\cos^2\beta_1\cos^2\beta_2\sin^2\beta_3 \end{aligned}\right\} = \text{Second-order cross-modulation powers} \qquad (3\text{-}45)$$

The obvious significance of (3-45) is that letting $\beta_1 = \pi/2$ provides the following advantages:

(a) Power in all three channels is maximized

(b) Unmodulated carrier is completely suppressed

(c) All second-order cross-modulation terms are suppressed.

For the case of $\beta_1 = \pi/2$, the signal portion of (3-44) can be written as

$$s(t) = \sqrt{2P}\cos\omega_c t\left[S_1(t)\cos\beta_2\cos\beta_3 - S_1(t)S_2(t)S_3(t)\sin\beta_2\sin\beta_3\right] - \sqrt{2P}\sin\omega_c t\left[S_2(t)\sin\beta_2\cos\beta_3 + S_3(t)\cos\beta_2\sin\beta_3\right]. \quad (3\text{-}46)$$

Equation (3-46) is useful for providing a geometrical representation of the three-channel interplex PSK/PM signal.

The interference to channel 1 caused by the cross-modulation term depends on the relationship between the signal formats and the data rates of the three channels. If the bandwidth of $S_1(t)$ is such that it extends into the frequency regions occupied by $S_2(t)$ and $S_3(t)$, interference will show up in channel 1 but it will not exceed the bounds set by power allocations indicated by (3-45).

The two-channel interplex system power allocations can be derived from (3-45) by setting $\beta_3 = 0$. This results in

$$\begin{aligned} P_1 &= P\sin^2\beta_1\cos^2\beta_2 = \text{Power in channel 1} \\ P_2 &= P\sin^2\beta_1\sin^2\beta_2 = \text{Power in channel 2} \\ P_c &= P\cos^2\beta_1\cos^2\beta_2 = \text{Unmodulated carrier power} \\ P_{cm} &= P\cos^2\beta_1\sin^2\beta_2 = \text{Cross-modulation power.} \end{aligned} \quad (3\text{-}47)$$

Setting $\beta_1 = \pi/2$ for maximum efficiency reduces the two-channel interplex system to a QPSK system with unequal powers in the two orthogonal components. Thus, (3-46) becomes simply

$$s_u(t) = \sqrt{2P}\cos\omega_c t\left[S_1(t)\cos\beta_2\right] - \sqrt{2P}\sin\omega_c t\left[S_2(t)\sin\beta_2\right] \quad (3\text{-}48)$$

where the subscript u indicates an unequal-power-division QPSK modulation. The corresponding polar form of this equation is

$$s_u(t) = \sqrt{2P} \cos\left[\omega_c t + S_1(t)\, S_2(t)\, \beta_2\right]. \qquad (3\text{-}49)$$

Now consider a signal which results from the addition of two quadrature carriers, each of which is amplitude modulated (DSB-SC) by a separate bipolar data sequence. The block diagram for the generation of such a composite signal is shown in Figure 3.22. The resultant QPSK signal is given by

$$f(t) = A\,a(t) \cos \omega_c t - B\,b(t) \sin \omega_c t, \qquad (3\text{-}50)$$

where

$a(t)$ = binary sequence (± 1) of channel 1
A = amplitude of channel 1
$b(t)$ = binary sequence (± 1) of channel 2
B = amplitude of channel 2.

(The inversion of $b(t)$ polarity has been introduced only for sign consistency with interplex.) It is customary to reduce (3-50) to an equivalent polar form:

$$f(t) = R \cos\left(\omega_c t + \theta\right). \qquad (3\text{-}51)$$

For this, using the trigonometric identity:

$$\cos(\omega_c t + \theta) = \cos \omega_c t \cos\theta - \sin \omega_c t \sin\theta\,. \qquad (3\text{-}52)$$

Thus,

$$\begin{aligned} f(t) &= A\,a(t) \cos \omega_c t - B\,b(t) \sin \omega_c t \\ &= R \cos \omega_c t\,[a(t) \cos\theta] - R \sin \omega_c t\,[b(t) \sin\theta]. \end{aligned} \qquad (3\text{-}53)$$

Note that (3-52) is of the same form as (3-48). Furthermore, for (3-52) to be true for all values of t, the coefficients of $\cos \omega_c t$ and $\sin \omega_c t$ must be such that

$$\begin{aligned} A\,a(t) &= R\,a(t) \cos\theta \\ B\,b(t) &= R\,b(t) \sin\theta \\ R &= \sqrt{A^2 + B^2} \\ \theta &= \tan^{-1}\left[\frac{B\,b(t)}{A\,a(t)}\right]. \end{aligned} \qquad (3\text{-}54)$$

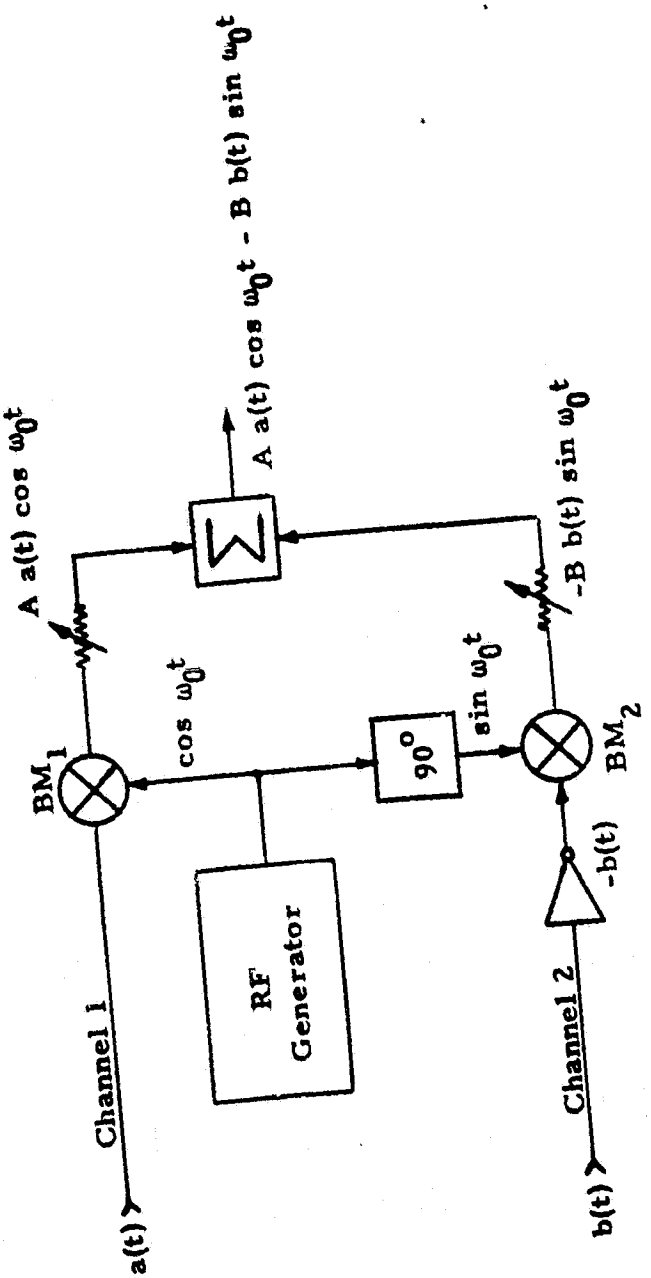


Figure 3.22. Two-Channel Quadriphase Modulator

Hence, (3-50) can be written as

$$f(t) = \sqrt{A^2+B^2}\cos\left\{\omega_c t + \tan^{-1}\left[\frac{B\,b(t)}{A\,a(t)}\right]\right\}. \qquad (3\text{-}55)$$

The equivalence with two-channel interplex ($\beta_1 = \pi/2$) is established by conversion coefficients

$$A = R\cos\beta_2 = \sqrt{2P}\cos\beta_2$$

$$B = R\sin\beta_2 = \sqrt{2P}\sin\beta_2$$

$$A/B = \tan\beta_2$$

$$a(t) = S_1(t)$$

$$b(t) = S_2(t). \qquad (3\text{-}56)$$

Thus, for the two-channel case and antipodal modulation of the primary (high power) channel, the results of interplex and quadriphase modulation are identical.

From the standpoint of implementation, there are differences which must be taken into account. The interplex transmitter performs all the data combining at baseband and thus requires a phase modulator whose linearity and gain must be carefully controlled. This requirement is particularly important for suppressing the cross-modulation and carrier terms. Quadriphase transmitter implementation, however, is based on combining individually modulated orthogonal components of the same carrier, a task which is relatively easy to implement.

A three-channel multiplexer/modulator which does not require a linear phase modulator, such as required by an interplex transmitter, can be synthesized from the basic configuration of an unbalanced two-channel quadriphase modulator. Such a three-channel multiplexer/modulator is shown in Figure 3.23. This configuration can be considered as a hybrid between the interplex and quadrature carrier multiplexing. The similarity with interplex arises from the fact that angular modulation of the two secondary channels is proportional to the fractional power allocated to each of these channels. On the

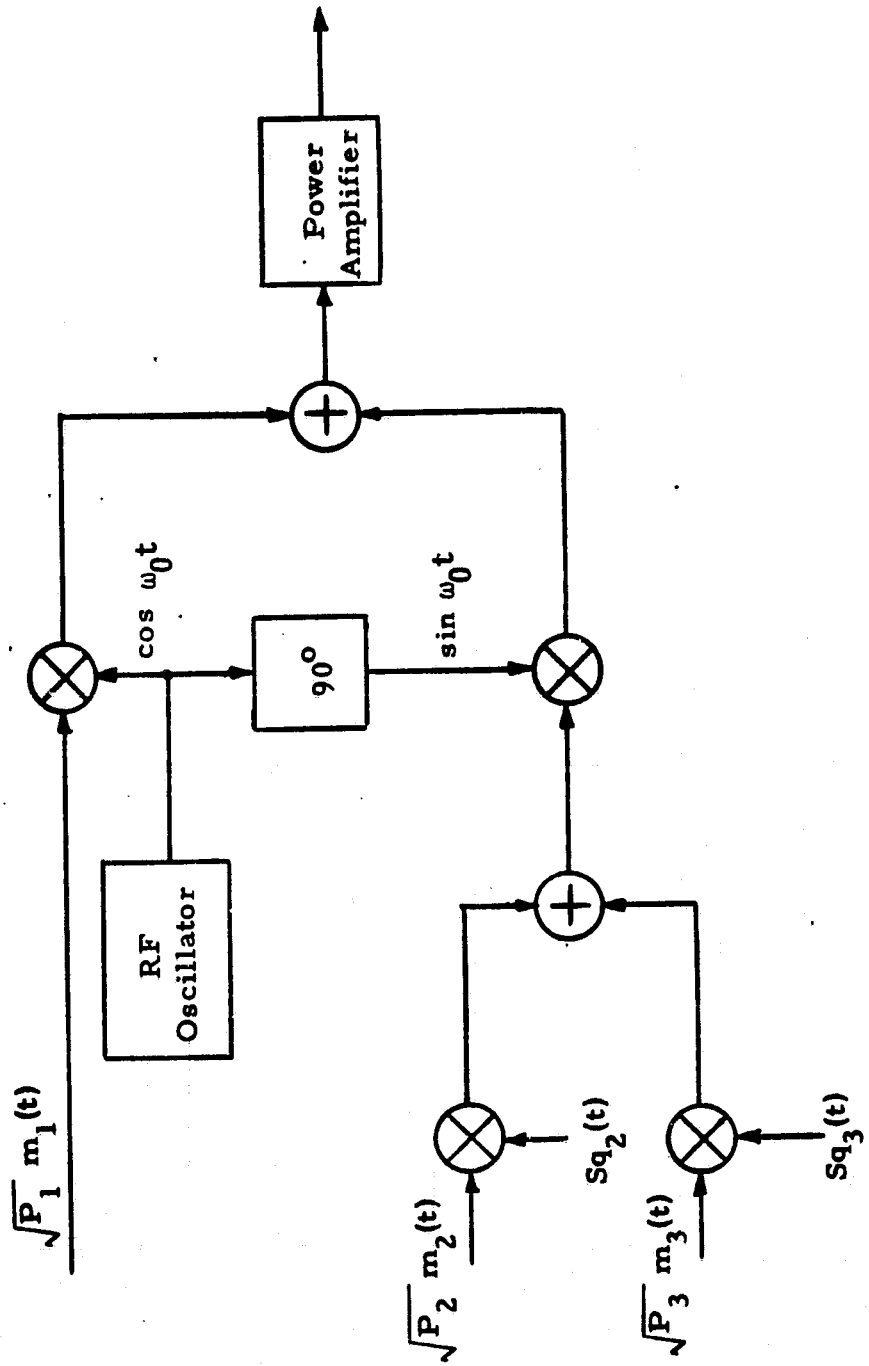


Figure 3.23. Three-Channel Quadrature Multiplex Modulator

other hand, the similarity with quadriphase modulation is because the two quadrature components of the same RF carrier are used to synthesize the composite signal.

The structure of the modulator shown in Figure 3.23 is such as to form an unbalanced QPSK signal wherein the high-rate data signal, $\sqrt{P_1}\, m_1(t)$, is biphase modulated on the inphase carrier $\sqrt{2}\sin\omega_0 t$, and the sum of the two lower rate signals, $\sqrt{P_2}\, m_2(t)$ and $\sqrt{P_3}\, m_3(t)$, after being modulated onto separate square-wave subcarriers, $Sq_2(t)$ and $Sq_3(t)$, are added and biphase modulated onto the quadrature carrier, $\sqrt{2}\cos\omega_0 t$. The sum of the inphase and quadrature-modulated carriers [i.e., the unbalanced QPSK signal s(t)] is then bandpass hard-limited and power amplified.

The modulator block diagram illustrated in Figure 3.23 is used to generate a signal for simultaneous transmission of three channels of information on the Orbiter's return Ku-band link. Two of the channels represent independent data channels (one having a data rate up to 50 Mbps while the other has a rate up to 2 Mbps) and the third channel consists of operational data at a rate of 192 kbps.

The signal generated by the modulator of Figure 3.23 is

$$s(t) = \sqrt{2}\left[C(t)\cos\omega_0 t + S(t)\sin\omega_0 t\right], \tag{3-57}$$

where

$$C(t) = \sqrt{P_2}\, Sq_2(t)\, m_2 t + \sqrt{P_3}\, Sq_3(t)\, m_3(t)$$

$$\triangleq \sqrt{P_2}\, s_2(t) + \sqrt{P_3}\, s_3(t)$$

$$S(t) = \sqrt{P_1}\, m_1(t) \triangleq \sqrt{P_1}\, s_1(t) \tag{3-58}$$

and $s_1(t)$, $s_2(t)$, $s_3(t)$ are ± 1 binary waveforms. Alternately, in polar coordinates (amplitude and phase), (3-57) can be rewritten as

$$s(t) = \sqrt{2}\, V(t)\sin(\omega_0 + \phi(t)), \tag{3-59}$$

where

$$V(t) = \sqrt{C^2(t) + S^2(t)}$$

$$\phi(t) = \tan^{-1}\frac{C(t)}{S(t)}. \tag{3-60}$$

Passing s(t) of (3-59) through a bandpass hard-limiter preserves the phase $\phi(t)$. Thus, the resultant first zone output is given by

$$z_1(t) = \sqrt{2P}\,\sin\left(\omega_0 t + \phi(t)\right), \tag{3-61}$$

where P is the total power in the first zone after amplification. In terms of inphase and quadrature components, (3-61) can be rewritten as

$$z_1(t) = \sqrt{2P}\left[\frac{S(t)}{V(t)}\sin\omega_0 t + \frac{C(t)}{V(t)}\cos\omega_0 t\right]. \tag{3-62}$$

From (3-58) and (3-60),

$$V(t) = \sqrt{P_1 s_1^2(t) + P_2 s_2^2(t) + P_3 s_3^2(t) + 2\sqrt{P_2P_3}\,s_2(t)\,s_3(t)}$$

$$= \sqrt{P_1 + P_2 + P_3 + 2\sqrt{P_2P_3}\,s_2(t)\,s_3(t)}\,. \tag{3-63}$$

Let $P_T \triangleq P_1 + P_2 + P_3$ denote the total input power; then,

$$V(t) = \sqrt{P_T\left(1 + 2\sqrt{P_2/P_T}\sqrt{P_3/P_T}\,s_2(t)\,s_3(t)\right)} \tag{3-64}$$

or, equivalently,

$$\frac{1}{V(t)} = \frac{1}{\sqrt{P_T}}\,\frac{1}{2}\underbrace{\left[\frac{1}{\sqrt{1 + 2\sqrt{P_2/P_T}\sqrt{P_3/P_T}}} + \frac{1}{\sqrt{1 - 2\sqrt{P_2/P_T}\sqrt{P_3/P_T}}}\right]}_{C_1}$$

$$- s_2(t)\,s_3(t)\,\frac{1}{2}\underbrace{\left[\frac{1}{\sqrt{1 - 2\sqrt{P_2/P_T}\sqrt{P_3/P_T}}} - \frac{1}{\sqrt{1 + 2\sqrt{P_2/P_T}\sqrt{P_3/P_T}}}\right]}_{C_2} \tag{3-65}$$

Substituting (3-58) and (3-65) into (3-62) gives

$$s_1(t) = \sqrt{2P}\left\{\left[C_1\sqrt{P_2/P_T} - C_2\sqrt{P_3/P_T}\right] s_2(t) + \left[C_1\sqrt{P_3/P_T} - C_2\sqrt{P_2/P_T}\right] s_3(t)\right\}\cos\omega_0 t$$

$$+ \sqrt{2P}\left\{C_1\sqrt{P_1/P_T}\, s_1(t) - C_2\sqrt{P_1/P_T}\, s_1(t)\, s_2(t)\, s_3(t)\right\}\sin\omega_0 t$$

$$= \sqrt{2}\left[\sqrt{\tilde{P}_1}\, s_1(t) - \sqrt{\tilde{P}_d}\, s_1(t)\, s_2(t)\, s_3(t)\right]\sin\omega_0 t$$

$$+ \sqrt{2}\left[\sqrt{\tilde{P}_2}\, s_2(t) + \sqrt{\tilde{P}_3}\, s_3(t)\right]\cos\omega_0 t\,, \qquad (3\text{-}66)$$

where

$$\tilde{P}_1 = (P/P_T) P_1 C_1^2$$

$$\tilde{P}_2 = (P/P_T)\left[C_1\sqrt{P_2} - C_2\sqrt{P_3}\right]^2$$

$$\tilde{P}_3 = (P/P_T)\left[C_1\sqrt{P_3} - C_2\sqrt{P_2}\right]^2$$

$$\tilde{P}_d = (P/P_T) P_1 C_2^2\,. \qquad (3\text{-}67)$$

If the output level of the power amplifier is arbitrarily set equal to the total input power level, i.e., $P = P_T$, for fixed values of P_1/P_T, P_2/P_T and P_3/P_T, one can then compute from (3-67) the corresponding values of $\tilde{P}_1/P_T$, $\tilde{P}_2/P_T$, $\tilde{P}_3/P_T$ and $\tilde{P}_d/P_T$ for the transmitted unbalanced quadriphase signal of (3-66). For example, suppose that the high-rate channel contains 80% of the total power while the remaining 20% is split, with 80% going to the next highest rate channel and 20% to the lowest rate channel. Equivalently, $P_1/P_T = 0.8$, $P_2/P_T = 0.16$ and $P_3/P_T = 0.04$. Then, from (3-67), the transmitted power ratios are $\tilde{P}_1/P_T = 0.8157$, $\tilde{P}_2/P_T = 0.1503$, $\tilde{P}_3/P_T = 0.0287$ and $\tilde{P}_d/P_T = 0.00529$.

3.2.4.3 Carrier frequency modulation

Carrier frequency modulation is very similar to subcarrier FM, as described in subsection 3.2.3.3. However, an additional concept that is frequently employed in telemetry systems is carrier FM with the output

of one or more subcarrier oscillators which, in turn, have been frequency modulated by data signals. As an example, consider an FM/FM signal for one subcarrier.

Let ω_c be the carrier angular frequency, ω_s the subcarrier frequency, and ω_m the modulating angular frequency. The subcarrier signal is given by

$$s(t) = \sqrt{2P_s} \cos\left[\omega_s t + \phi_s + \frac{\Delta\omega_s}{\omega_m} \sin(\omega_m t + \phi_m)\right]$$

$$= \sqrt{2P_s} \cos\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right], \qquad (3\text{-}68)$$

where $\Delta\omega_s$ is the peak frequency deviation of the subcarrier and β_s is the peak phase deviation of the subcarrier. The instantaneous frequency of the carrier wave is given by

$$\omega_i(t) = \omega_c + \Delta\omega \cos\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right], \qquad (3\text{-}69)$$

where $\Delta\omega$ is the maximum frequency deviation of the carrier.

Let the modulated carrier wave be represented by

$$s(t) = \sqrt{2P} \cos\theta(t), \qquad (3\text{-}70)$$

where

$$\theta(t) = \int_0^t \omega_i(t)\, dt$$

$$= \int_0^t \left\{\omega_c + \Delta\omega \cos\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right]\right\} dt$$

$$= \omega_c t + \phi_c + \Delta\omega \int_0^t \cos\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right] dt. \qquad (3\text{-}71)$$

Evaluating the integral,

$$I = \int_0^t \cos\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right] dt \qquad (3\text{-}72a)$$

$$I = \frac{1}{\omega_s} \sin\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right] - \frac{\Delta\omega_s}{\omega_s}$$

$$\times \int_0^t \cos(\omega_m t + \phi_m) \cos\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right] dt . \qquad (3\text{-}72b)$$

In practice, $\omega_s >> \omega_m$, and the term $\cos(\omega_m + \phi_m)$, which is a slowly varying function compared to $\cos \omega_s t$, may be treated as a constant in front of $\cos(\omega_s t + \phi_s + \beta_s \sin \omega_m t + \phi_m)$ during the integration period. Therefore,

$$I = \frac{1}{\omega_s} \sin\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right] - \frac{\Delta\omega_s}{\omega_s} \cos(\omega_m t + \phi_m) I$$

or

$$I\left[1 + \frac{\Delta\omega_s}{\omega_s} \cos(\omega_m t + \phi_m)\right] = \frac{1}{\omega_s} \sin\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right] .$$

Since $\Delta\omega_s/\omega_s << 1$, the term $(\Delta\omega_s/\omega_s) \cos(\omega_m t + \phi_m)$ can be neglected and

$$I = \frac{1}{\omega_s} \sin\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right] . \qquad (3\text{-}73)$$

by substituting this result in (3-71),

$$\theta(t) = \omega_c t + \phi_c + \frac{\Delta\omega}{\omega_s} \sin\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right] . \qquad (3\text{-}74)$$

The modulated carrier wave of (3-70) is given by

$$s(t) = \sqrt{2P} \cos\theta(t)$$

$$= \sqrt{2P} \cos\left\{\omega_c t + \phi_c + \beta \sin\left[\omega_s t + \phi_s + \beta_s \sin(\omega_m t + \phi_m)\right]\right\}, \qquad (3\text{-}75)$$

where $\beta = \Delta\omega/\omega_s$. This can be written in the complex form

$$s(t) = \sqrt{2P}\ \mathrm{Re}\left(e^{j(\omega_c t+\phi_c)}\ e^{j\beta \sin\left[\omega_s t+\phi_s+\beta_s \sin(\omega_m t+\phi_m)\right]}\right)$$

$$= \sqrt{2P}\ \mathrm{Re}\ e^{j(\omega_c t+\phi_c)} \sum_{p=-\infty}^{\infty} J_p(\beta)\ e^{jp(\omega_s t+\phi_s)} \sum_{q=-\infty}^{\infty} J_q(p\beta_s)\ e^{jq(\omega_m t+\phi_m)}. \tag{3-76}$$

Finally,

$$s(t) = \sqrt{2P} \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} J_p(\beta)\, J_q(p\beta_s)$$

$$\times \cos\left[(\omega_c + p\omega_s + q\omega_m)t + \phi_c + p\phi_s + q\phi_m\right]. \tag{3-77}$$

In the case of n subcarriers, it can be shown that (3-77) assumes the form

$$s(t) = \sqrt{2P} \prod_{k=1}^{n} \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} J_p(\beta_k)\, J_q(p\beta_{sk})$$

$$\times \cos\left[\left(\omega_c + p\sum_{k=1}^{n} \omega_{sk} + q\sum_{k=1}^{n} \omega_{mk}\right)t + \phi_c + p\sum_{k=1}^{n} \phi_{sk} + q\sum_{k=1}^{n} \phi_{mk}\right], \tag{3-78}$$

where n subcarriers of frequencies $\omega_{s1}, \omega_{s2}, \ldots, \omega_{sn}$ are respectively modulated by the data input signals of frequencies $\omega_{m1}, \omega_{m2}, \ldots, \omega_{mn}$.

3.2.4.4 Spread spectrum

Spread spectrum and modulation systems are designed to permit communication of message information under the condition of very low SNR, such as for low detectability, or to combat interfering transmissions, as in military antijam or multiple-access environments.

One of the requirements for NASA communication systems to use spread spectrum is determined by the guidelines established by the CCIR (International Radio Consultative Committee) of the ITU (International Telecommunications Union) on the maximum-flux density produced at the earth's surface by emissions from space communication systems. These guidelines are intended to prevent interference with earth-based communication systems and theoretically can be exceeded

only after it has been established that there is no possibility that interference exists at the frequency under consideration. At S-band frequencies between 2200 MHz and 2300 MHz, it is not possible to guarantee that there will be no interference with other communication systems; therefore, the CCIR guidelines must be followed. Table 3.2 presents the CCIR guidelines for S-band transmission. The worst case is for a tangential angle of incidence $\theta = 0°$ since the difference between the path length for $\theta = 0°$ and $\theta = 90°$ is great enough to cause a 10 dB change in signal energy due to the path length.

Table 3.2. CCIR Guidelines for Maximum Flux Density for S-Band Transmission Between 2200 MHz and 2300 MHz

Angle of Arrival Above the Horizontal Plane θ (Angle of Incidence)		
$\theta = 0°$ to $5°$	$\theta = 5°$ to $25°$	$\theta = 25°$ to $90°$
-154 dBW/m^2 per 4 kHz	$\left(-154 + \frac{\theta-5}{2}\right)$ dBW/m^2 per 4 kHz	-144 dBW/m^2 per 4 kHz

The most straightforward way to widen the spectrum of a message signal is to multiply (modulate) it by a wideband signal. Such a spreading signal must have correlation properties that aid in acquisition and tracking. The best signal that fits these specifications is the binary pseudonoise (PN) signal. The basic modulator is shown in Figure 3.24. The data is modulated onto subcarriers and then onto a carrier using any of the modulation techniques described in this section. The balanced modulator multiplies by +1 or -1 (i.e., inverts the signal) according to the PN code.

The power spectrum of a PN waveform is given by

$$S(\omega) = \frac{p+1}{p^2}\left(\frac{\sin\frac{\omega t_c}{2}}{\frac{\omega t_c}{2}}\right)^2 \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \delta\left(\omega - \frac{2\pi n}{pt_c}\right) + \frac{1}{p^2}\delta(\omega), \qquad (3\text{-}79)$$

where p is the period of the sequence, $t_c = 1/f_c$ is the chip period (or signaling period) of the PN waveform, and $\delta(\omega)$ is the Dirac delta

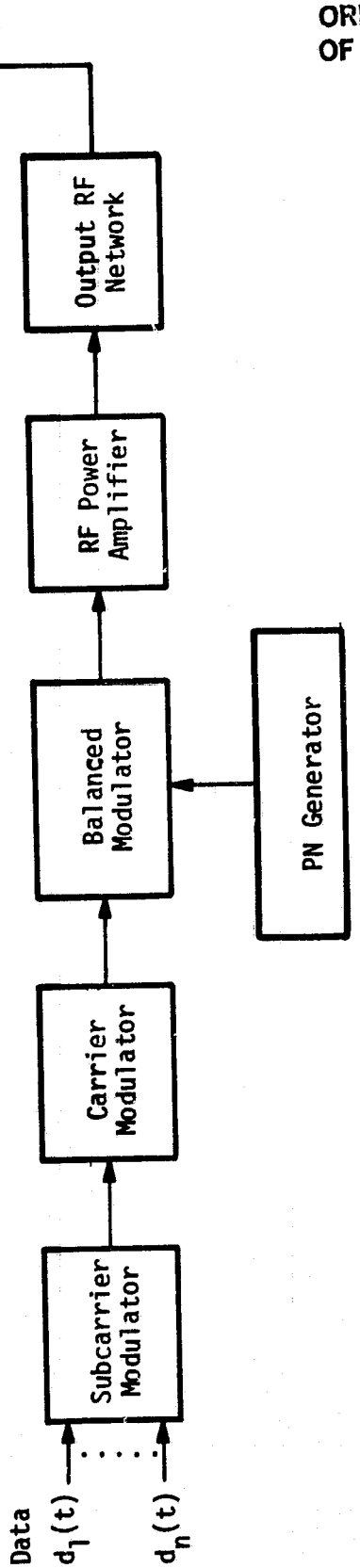


Figure 3.24. Biphase PN Modulator/Transmitter

function. Note that this spectrum is a line spectrum, with frequencies at multiples of the fundamental frequency. Since the binary waveform is a constant-amplitude square wave having constant power, there is a scale factor inversely proportional to the period of the sequence. Thus, if the period of the sequence is doubled, the lines in the spectrum become twice as dense and the power of each is reduced by a factor of 2. The envelope of the spectrum is defined completely by the chip period. This means that the possible spectrum spreading is determined independent of the period of the sequence and solely on the chip period. For very long periods, the spectrum is essentially continuous, having the form

$$S(\omega) = t_c \left(\frac{\sin \frac{\omega t_c}{2}}{\frac{\omega t_c}{2}} \right)^2 . \qquad (3\text{-}80)$$

which is the same as the NRZ spectrum shown in Figure 3.3. Note that multiplying the normal communication signal by the PN signal has the effect in the frequency domain of convolving the normal communication signal spectrum with the PN spectrum. The convolution of the two spectra tends to make the PN line spectrum in (3-79) merge into a continuous spectrum, so even short-period PN codes can be used to spread a normal communication signal.

3.2.5 Transmitter Power Allocation

The efficiency of a communication system design is measured by how much power is required to achieve a given data rate with a given maximum-bit-error rate or equivalent analog data criterion. Therefore, one of the important system parameters to be determined is the allocation of the total transmitter power to the data streams, to the residual carrier, and to undesired cross-modulation terms (modulation loss terms). The following sections present the transmitter power allocation results for phase modulation (PM), phase-shift-keying (PSK) and frequency modulation (FM).

3.2.5.1 Phase modulation power allocation

Subsection 3.2.4.1 presented the spectral properties of two subcarriers phase modulated onto a carrier. The first example was two sinusoidal subcarriers with peak phase deviations β_1 and β_2 (in radians) for subcarriers 1 and 2, respectively. From (3-30), it can easily be shown that the power in the carrier P_c is given by

$$P_c = J_0^2(\beta_1) J_0^2(\beta_2) P_T , \tag{3-81}$$

where P_T is the total transmitted power. Similarly, the power in the subcarriers is given by

$$P_{s1} = 2 J_1^2(\beta_1) J_0^2(\beta_2) P_T$$

$$P_{s2} = 2 J_0^2(\beta_1) J_1^2(\beta_2) P_T . \tag{3-82}$$

The power division (depending on whether the data is digital or analog) shown in Figure 3.1 is then P_{s1}/P_T or P_{s2}/P_T, given in (3-82) for subcarriers 1 and 2, respectively. Figures 3.25 and 3.26 plot P_c and P_{s1} as a function of β_1 and β_2. Since the expressions for P_{s1} and P_{s2} are symmetric with respect to β_1 and β_2, only one plot is shown.

If only one subcarrier is modulated onto the carrier, then the power in the carrier P_c is given by

$$P_c = J_0^2(\beta) P_T \tag{3-83}$$

and the power in the subcarrier is given by

$$P_s = 2 J_1^2(\beta) P_T . \tag{3-84}$$

Figures 3.27 and 3.28 plot P_c and P_s as a function of β.

The second example presented in subsection 3.2.4.1 was for a sinusoidal subcarrier and a square-wave digital data stream modulated onto the carrier. In this case, it can be shown from (3-33) that the power in the carrier P_c is given by

$$P_c = J_0^2(\beta_s) \cos^2 \beta_d \, P_T . \tag{3-85}$$

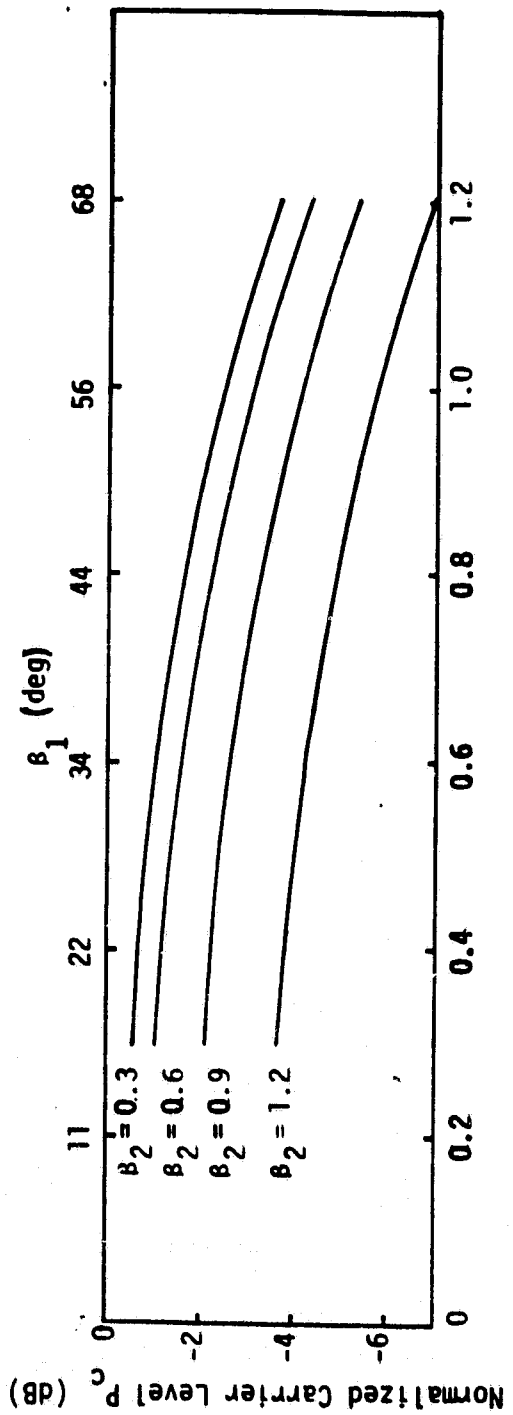


Figure 3.25. Carrier Suppression Factor as a Function of Modulation Indexes

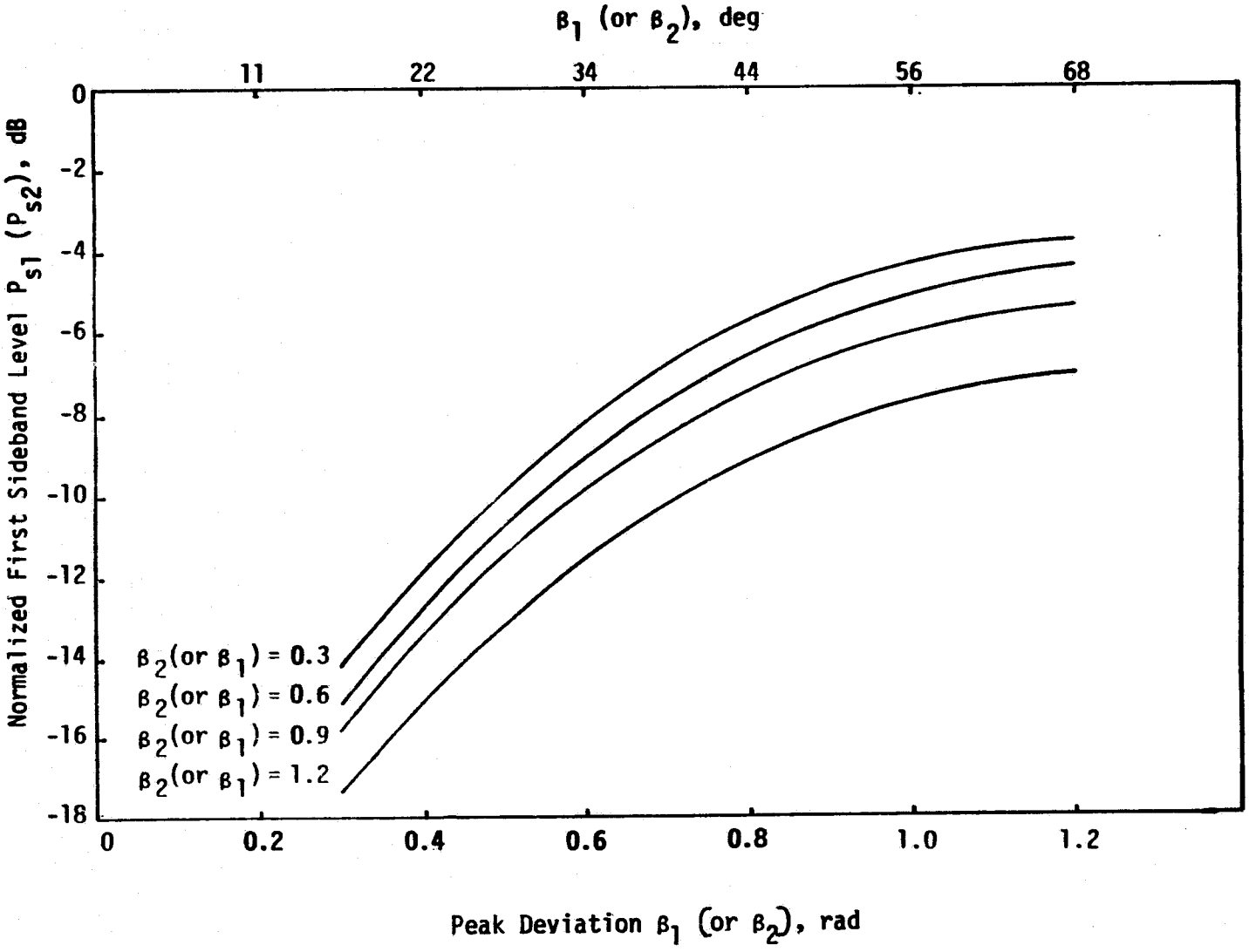


Figure 3.26. First Sideband Relative to Power Level as a Function of Modulation Indexes

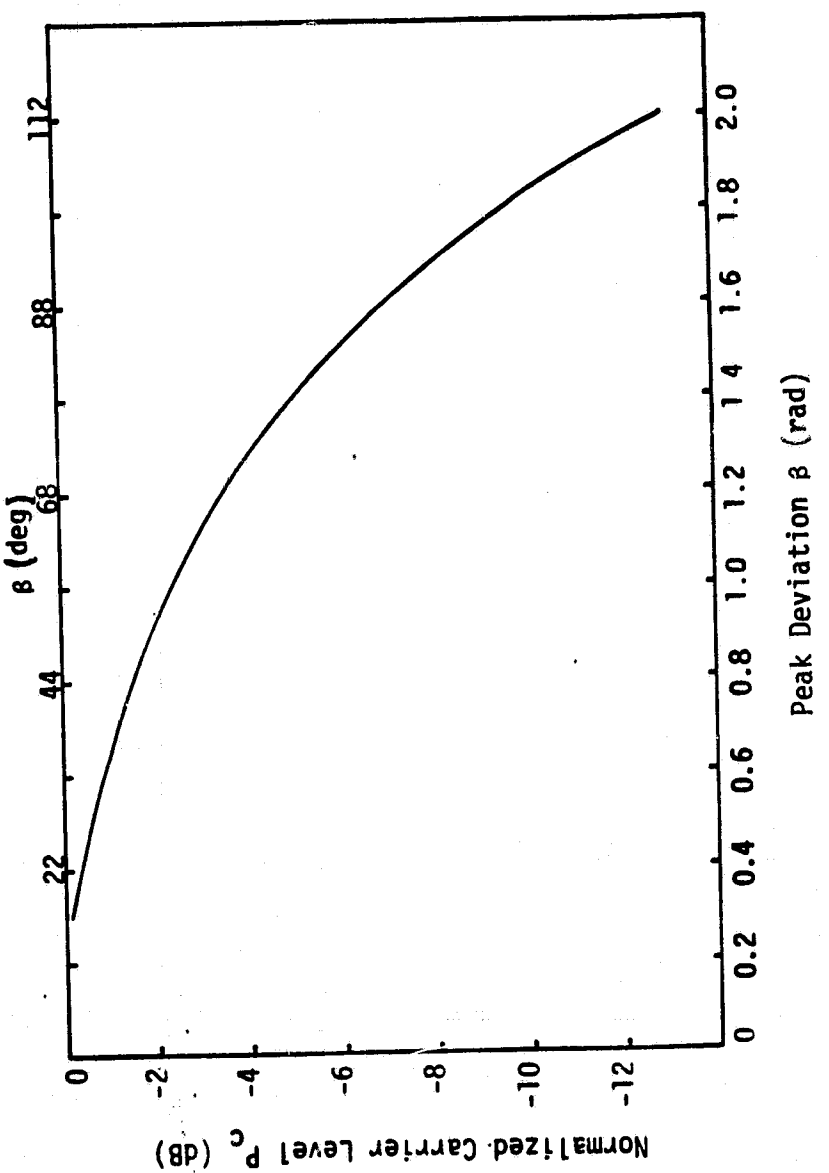


Figure 3.27. Carrier Suppression Factor as a Function of Modulation Index

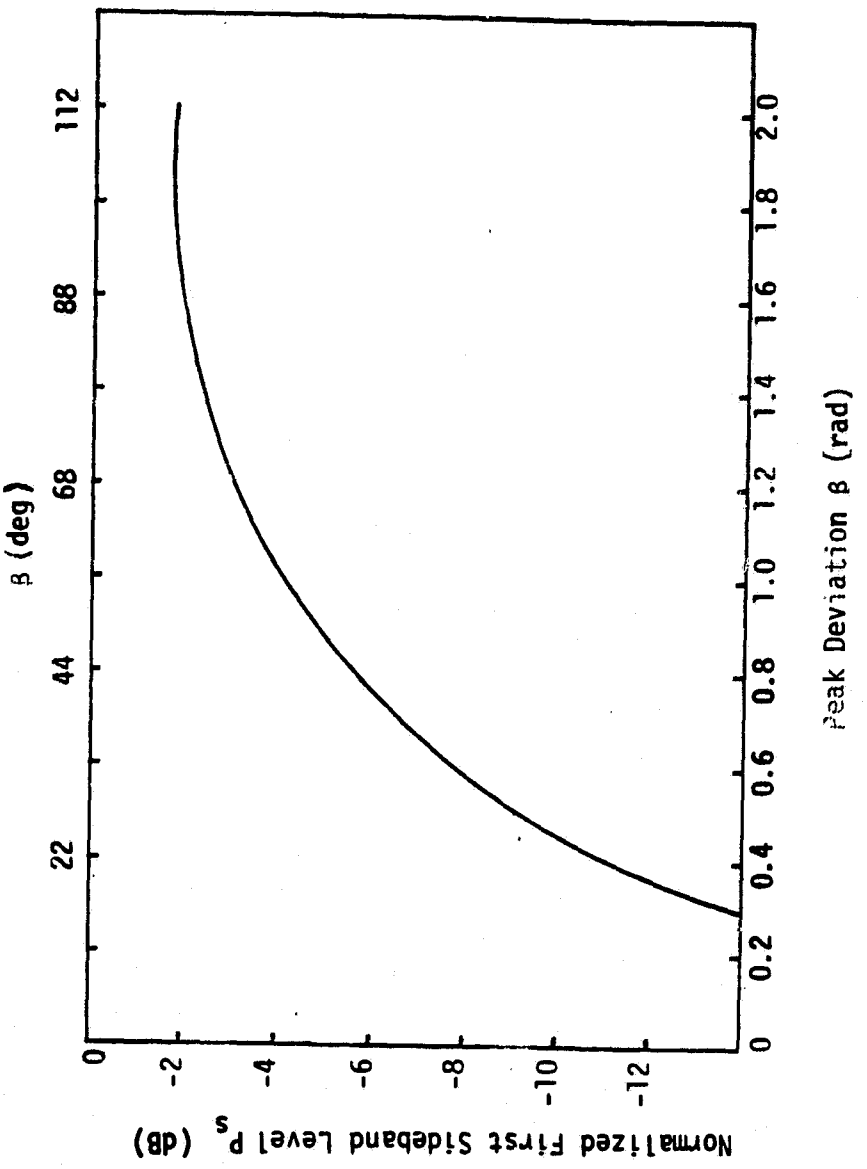


Figure 3.28. First Sideband Power Suppression Factor as a Function of Modulation Index

The power in the sinusoidal subcarrier is given by

$$P_s = 2J_1^2(\beta_s)\cos^2\beta_d P_T , \tag{3-86}$$

and the power in the square-wave digital data stream P_d is given by

$$P_d = \frac{1}{2}J_0^2(\beta_s)\sin^2\beta_d P_T . \tag{3-87}$$

The power divisions shown in Figure 3.1 are P_d/P_T and P_s/P_T. Figures 3.29 through 3.31 plot P_c, P_s, and P_d, respectively.

3.2.5.2 PSK and QPSK power allocation

The transmitter power allocations for PSK/PM and QPSK were presented in subsection 3.2.4.2. For reference, Table 3.3 summarizes the transmitter power allocation for each of the modulation techniques presented in subsection 3.2.4.2. Note that the data power is maximized in the single-channel case or the interplex and QPSK cases if $\beta_1 = \pi/2$, which results in completely suppressed carrier operation.

3.2.5.3 FM modulation index versus bandwidth

The problem of bandwidth design is of paramount practical importance in the design of communication systems. The bandwidth of the bandpass filters for the transmission and reception of the modulated carrier should be adequate enough to encompass the significant sidebands only.

As noted previously, the theoretically infinitely wide frequency band of an FM signal is in practice limited to a finite number of significant sidebands compatible with the distortion and intermodulation specifications, the number of significant sidebands being a function of the modulation index β and the peak deviation $\Delta\omega$. For $\beta < 1$, the spectrum is limited to the carrier and one pair of significant sidebands only of amplitudes A_c and $\beta A_c/2$, respectively, because $J_0(\beta) = 1$ and $J_1(\beta) = \beta/2$. The second and subsequent sidebands can be neglected because, for small β we have $J_n(\beta) = \beta^n/2^n n!$. for $\beta >> 1$ or $\Delta\omega >> \omega_m$, the required angular bandwidth is given by $2\Delta\omega$, as shown in Figure 3.13.

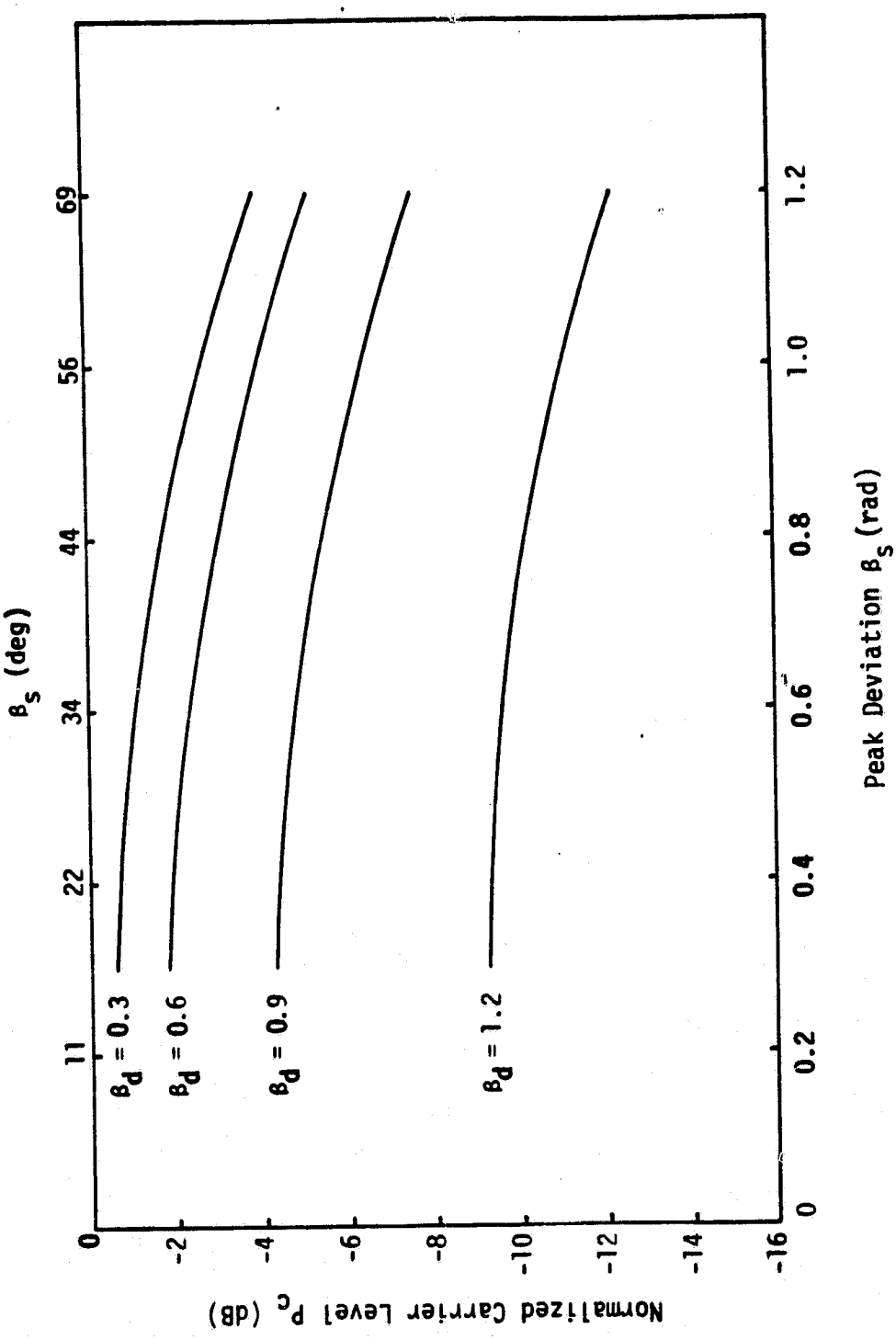


Figure 3.29. Carrier Suppression Factor as a Function of Modulation Indexes

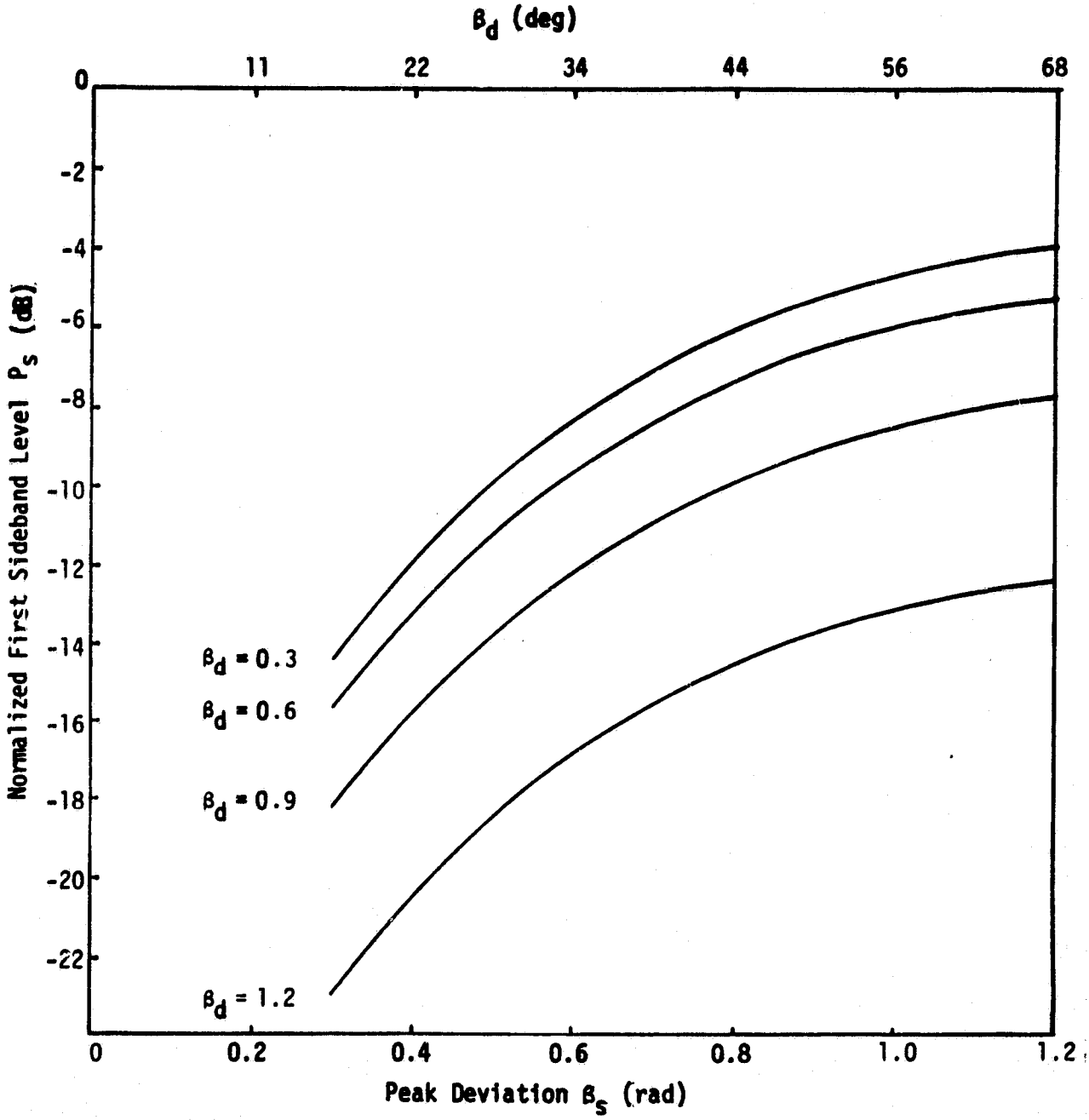


Figure 3.30. First Sideband Power Level for Sinusoidal Modulation as a Function of Modulation Indexes

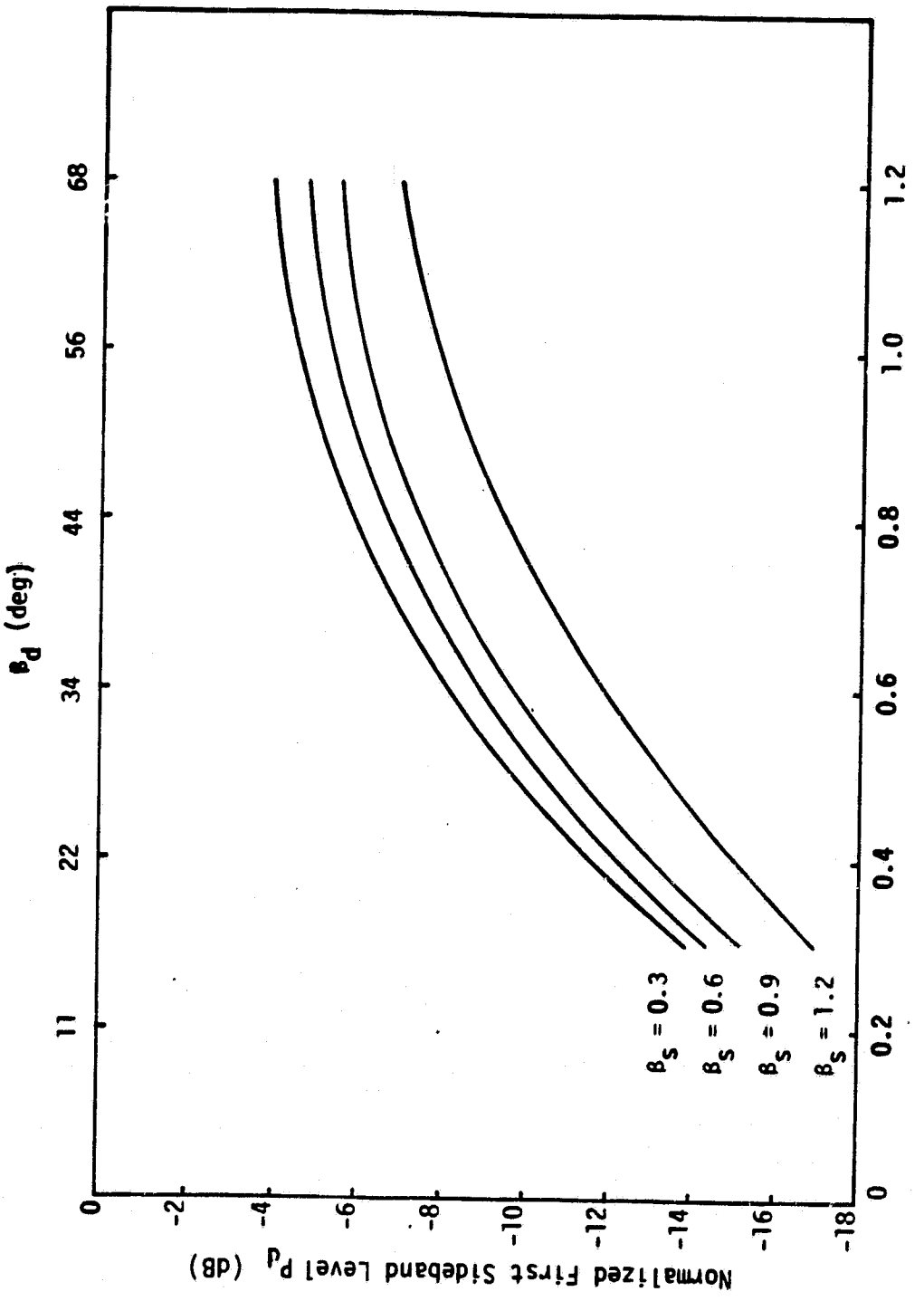


Figure 3.31. First Sideband Power Level for Square Modulation as a Function of Modulation Indexes

Table 3.3. Transmitter Power Allocations for PSK/PM and QPSK

Type of Modulation	Channel Power Ratios	Power Allocation for Each Component
Single-Channel PSK/PM	P_c/P_T	$\cos^2\beta$
	P_d/P_T	$\sin^2\beta$
Two-Channel PSK/PM	P_c/P_T	$\cos^2\beta_1\ \cos^2\beta_2$
	P_{d1}/P_T	$\sin^2\beta_1\ \cos^2\beta_2$
	P_{d2}/P_T	$\cos^2\beta_1\ \sin^2\beta_2$
Two-Channel Interplex or QPSK	P_c/P_T	$\cos^2\beta_1\ \cos^2\beta_2$
	P_{d1}/P_T	$\sin^2\beta_1\ \cos^2\beta_2$
	P_{d2}/P_T	$\sin^2\beta_1\ \sin^2\beta_2$
Three-Channel PSK/PM	P_c/P_T	$\cos^2\beta_1\ \cos^2\beta_2\ \cos^2\beta_3$
	P_{d1}/P_T	$\sin^2\beta_1\ \cos^2\beta_2\ \cos^2\beta_3$
	P_{d2}/P_T	$\cos^2\beta_1\ \sin^2\beta_2\ \cos^2\beta_3$
	P_{d3}/P_T	$\cos^2\beta_1\ \cos^2\beta_2\ \sin^2\beta_3$
Three-Channel Interplex or QPSK	P_c/P_T	$\cos^2\beta_1\ \cos^2\beta_2\ \cos^2\beta_3$
	P_{d1}/P_T	$\sin^2\beta_1\ \cos^2\beta_2\ \cos^2\beta_3$
	P_{d2}/P_T	$\sin^2\beta_1\ \sin^2\beta_2\ \cos^2\beta_3$
	P_{d3}/P_T	$\sin^2\beta_1\ \cos^2\beta_2\ \sin^2\beta_3$

P_T = Total transmitted power

P_c = Carrier power

P_{di} = Data power in channel i

β = Peak phase deviation

In case of multitone modulation, as the number of modulating signals is increased, the amplitudes of the sidebands decrease rapidly, as expected, since the total energy of the modulated wave is constant and independent of the form of the modulating signal. Since the sidebands extend from zero to infinite frequencies, it is impossible to assign a definite bandwidth for the frequency-modulated wave, but it is obvious that, in general, beyond a certain frequency range from the carrier, the amplitudes of the sidebands are negligibly small and consequently may be neglected. In fact, if each value of ω_m is fairly large compared to unity, the sidebands which lie very far outside the band $\omega_c \pm \sum_{m=1}^{N} \beta_m \omega_m$ can be neglected.

If the component waves of the complex signal have a low modulation index, the spectrum will consist of a pair of sidebands for each component wave of frequencies $(\omega_c \pm \omega_m)$. The bandwidth is thus twice the highest modulation frequency. In some practical applications, the complex signal will have a low-frequency component of high modulation index and a high-frequency wave of low modulation index. In this case, the IF bandwidth is approximately given by

$$B_{IF} = 2\Delta f + 2f_m , \tag{3-88}$$

where Δf is the peak frequency deviation of the carrier due to the low-frequency component and f_m is the frequency of the highest frequency component of the complex modulation signal. Thus, the bandwidth equals twice the maximum deviation of the low-frequency wave plus twice the maximum modulating frequency.

For intermediate values of β, there are several approximate expressions available for use in the design of multiplex communication systems. An approximate formula which is commonly used for the IF bandwidth required for such systems is

$$B_{IF} = 2(\Delta f + 2f_m) = 2\Delta f\left(1 + \frac{2}{\beta}\right) , \tag{3-89}$$

where Δf is the peak frequency deviation of the system and f_m is the highest baseband frequency. A plot of (3-89) is shown in Figure 3.32, where the bandwidth occupied by the significant sideband is plotted

(normalized for Δf) versus the modulation index β. Note that, for $\beta \gg 1$, $B_{IF} \rightarrow 2\Delta f$, as expected.

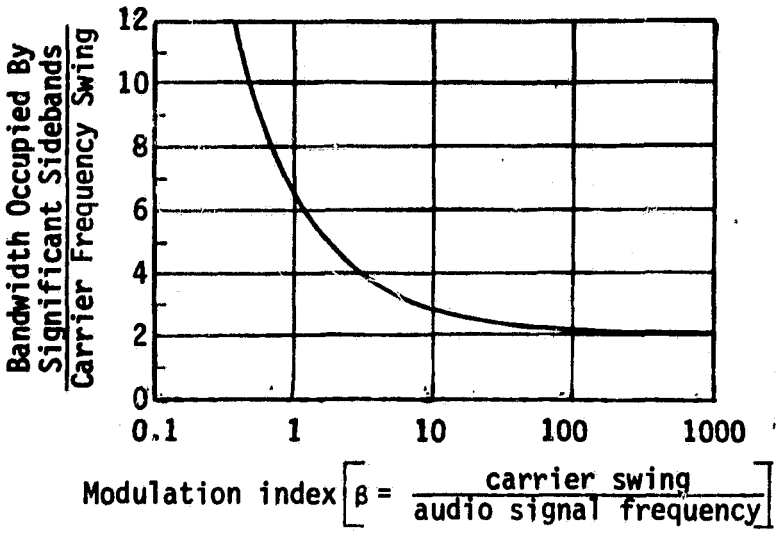


Figure 3.32. Significant Bandwidth (Normalized) versus Modulation Index

3.2.6 Antennas

In order to permit the transfer of information between the transmitter and receiver, the radio signal is radiated into space by an antenna. The simplest conceptual antenna is the isotropic radiator, which distributes the radiated power uniformly into space. Thus, for a total radiated power P, the radiation intensity is $P/4\pi$ per unit solid angle, independent of angular position with respect to the antenna.

With a preferred direction or range of directions, it becomes desirable to concentrate the radiated power in the useful areas. For this purpose, a directional antenna may be designed so that the radiation intensity is a function of the direction, Ω.

Defining $U(\Omega)$ as the radiation intensity for a total radiated power P and $U_0 = P/4\pi$, the antenna gain function is then

$$G(\Omega) \triangleq \frac{U(\Omega)}{U_0} \qquad (3\text{-}90)$$

Plots of $G(\Omega)$ = constant are known as antenna pattern contours. $G_m \triangleq \max [G(\Omega)]$ is called peak antenna gain or, frequently, antenna gain. In general,

$$\int_{\text{all } \Omega} G(\Omega)\, d\Omega = 4\pi P_t , \tag{3-91}$$

where P_t, the power transmission coefficient, is less than 1 for realizable antennas because of resistive losses in the antenna structure.

3.2.6.1 Antenna gain function

The antenna gain function $G(\Omega)$ is generally a complicated relationship which depends on the type of antenna, the type of feed, the structure on which the antenna and the feed are mounted, and the perturbations due to fabrication. It can be closely approximated by numerical techniques during design.

It is important to note that the gain functions for receive and transmit are identical for a given frequency by reciprocity. However, when referring to a receiving antenna, the concept of capture area is useful. The antenna capture area, A_C, is defined as the ratio of the power P delivered by the feed to a matched load and the power density P_I (watts per unit area) normally incident on the antenna aperture by an infinite plane wave:

$$A_C \triangleq \frac{P}{P_I} . \tag{3-92}$$

The aperture efficiency η is defined as

$$\eta = A_C/A , \tag{3-93}$$

where A is the physical aperture area. In general, η ranges between 0 and 1 but, theoretically, may be greater than 1. The power transmission coefficient is absorbed into η. A_C may be found by the relationship

$$A_C = \frac{G_m \lambda^2}{4\pi} , \tag{3-94}$$

where λ is the wavelength of the signal.

Simple representations of the antenna gain function may be used to understand parametric variations but are not useful for detailed analysis because of the invalidity of the simplifying assumptions.

Finding $G(\omega)$ for a uniform feed is a considerable undertaking and gets too deeply involved in antenna design rather than system design. Thus, the best course of action is to use measured patterns scaled approximately. Table 3.4 presents some scaling parameters.

For a quick estimation of the antenna gain pattern and nulls of an axisymmetric antenna, a useful approximation formula is:

$$G(\theta) = k_1\left(\frac{\sin k_2\theta}{\theta}\right)^2, \tag{3-95}$$

where $k_1 = G_m/k_2^2$, $k_2 = 2.78/\theta_0$, and $G_m = 9.42\theta_0^{-1.96}$. G_m is the on-axis gain (the antenna gain) expressed as a ratio, θ_0 is the half-power beamwidth [3 dB beamwidth, i.e., $G(\theta_0/2) = G_m/2$], and θ_0 and θ are measured in radians.

In (3-95), the parameter θ_0, the 3-dB beamwidth of the antenna, is used; θ_0 is defined as the central planar angle over which the antenna gain function remains within 3 dB of the peak antenna gain. Depending on the type of antenna, the 3-dB beamwidth may differ in different directions and in different planes; however, it is commonly used as a measure of the directivity of the antenna, a small beamwidth implying a highly directive antenna and vice versa. Thus, for increasing gain, the beamwidth decreases and proper antenna pointing becomes more important. The concept of beamwidth is usually used only for antennas with a single primary lobe, but can be useful in other cases if care is used in specification.

3.2.6.2 Pointing loss

The maximum communication system performance is obtained when both the transmitter and receiver antennas are lined up along their maximum gain points. Deviation from the maximum gain point of either antenna results in a pointing loss. The transmitter pointing loss L_{PT} is given by

Table 3.4. Properties of Typical Antennas

Configuration	Gain Above Isotropic Radiator	3-dB Beamwidth (deg)
Isotropic radiator	1	360
Infinitesimal dipole or loop	1.5	90 (toroidal pattern)
Half-wave dipole	1.64	78 (toroidal pattern)
Paraboloid (area A, diameter D)	$(6.3 \text{ to } 8.8)A/\lambda^2$	$(60 \text{ to } 70)\lambda/D$
Open-mouth waveguide (area A, E-plane dimension D_E, H-plane D_H)	$10A/\lambda^2$	$\theta_E = 56\ \lambda/D_E$ $\theta_H = 67\ \lambda/D_H$
Optimum horn (mouth area A, E-plane dimension D_E, H-plane D_H)	$7.5A/\lambda^2$	$\theta_E = 56\ \lambda/D_E$ $\theta_H = 67\ \lambda/D_H$
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)

$$L_{PT} = \frac{G_T(\Omega)}{G_{Tm}} , \tag{3-96}$$

where $G_T(\Omega)$ is the transmitter antenna gain along the axis determined by the autotrack error or open-loop pointing error and G_{Tm} is the maximum gain of the transmitter antenna. With autotracking, the pointing error is due to receive noise in the antenna servo-tracking loop. The normalized RMS tracking error for a monopulse antenna is

$$\frac{\sigma_\theta}{\theta_0} = \frac{1}{K\sqrt{(SNR)_L}} , \tag{3-97}$$

where σ_θ is the RMS tracking error, $(SNR)_L$ is the SNR in the servo loop bandwidth, and K is the gain slope constant. Therefore, the narrower the servo loop bandwidth, the greater the $(SNR)_L$ and the smaller the normalized error.

Open-loop pointing error occurs because the antenna maximum gain is not along the assumed line-of-sight. There are several contributions to the open-loop pointing error. First, there are errors due to the inaccurate knowledge of the receiver location. Second, there are errors due to the alignment of the mechanical boresight and the electrical boresight with the direction of maximum gain.

The receiver pointing loss L_{PR} is similarly given by

$$L_{PR} = \frac{G_R(\Omega)}{G_{Rm}} , \tag{3-98}$$

where $G_R(\Omega)$ is receiver antenna gain function along the axis determined by the autotrack error or open-loop pointing error, and G_{Rm} is the maximum gain of the receiver antenna.

3.2.6.3 Polarization loss

The gain function of an antenna does not completely characterize its performance. $G(\Omega)$ 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 electric field (E) 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 theory 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, and the ratio of the power delivered to the antenna terminals, P_R, to the available power, P_A, which would be received by an antenna matched to the incident polarization is called the polarization efficiency, ν;

$$\nu = \frac{P_R}{P_A} . \tag{3-99}$$

This is 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° 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,

$$p = \pm j \frac{R_1}{R_2} \triangleq \pm jR ,$$

where $j = \sqrt{-1}$ and R 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 decibels, where

$$\text{ellipticity} = 20 \log_{10} R . \tag{3-100}$$

Generally, the antenna is characterized by a gain pattern and an ellipticity pattern.

For R = 1, the E vector traces a circle corresponding to 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-hand circularly polarized (RHCP). For p = j, the wave is left-hand circularly polarized (LHCP).

All antennas have some ellipticity. It may be shown [8] that the polarization loss, L_{POL}, between two stations with RHCP generally can be written:

$$L_{POL} = \left[\frac{\left(1+R_R^2\right)\left(1+R_T^2\right) + 4R_TR_R + \left(1-R_R^2\right)\left(1-R_T^2\right)\cos 2\phi}{2\left(1+R_R^2\right)\left(1+R_T^2\right)}\right], \qquad (3\text{-}101)$$

where R_R is the receiving antenna axial ratio, R_T is the transmitting antenna axial ratio, and ϕ is the angle between the major axes of the polarization ellipses.

In general, the relative orientation of the polarization ellipses is not known. Therefore, ϕ is assumed random between 0° and 90° (for greater than 90°, the pattern repeats since the angular factor is 2ϕ). If the polarization loss is denoted by $L_{POL}(\phi)$, then $L_{POL}(0°)$ is a maximum (favorable), $L_{POL}(90°)$ is a minimum (adverse), and $L_{POL}(45°)$ is the average value. Figure 3.33 presents a plot of $L_{POL}(45°)$ versus transmitter and receiver antenna ellipticities.

3.2.6.4 Antenna feeds and cables

The antenna system looking outward from the transmitter (or from the receiver) consists of a transmission line (waveguide or coaxial cable), the antenna reflector (or device), and a feedhorn (or feeding device). The antenna feeds and transmission lines, as well as any included circulators, directional couplers, filters, and switches, attenuate the transmitter or receiver signal. This power loss due to attenuation from the transmitter to the antenna is defined as the transmitter circuit loss, L_T. Similarly, the loss due to signal attenuation between the antenna and the preamplifier in the receiver is defined as the receiver circuit loss, L_R.

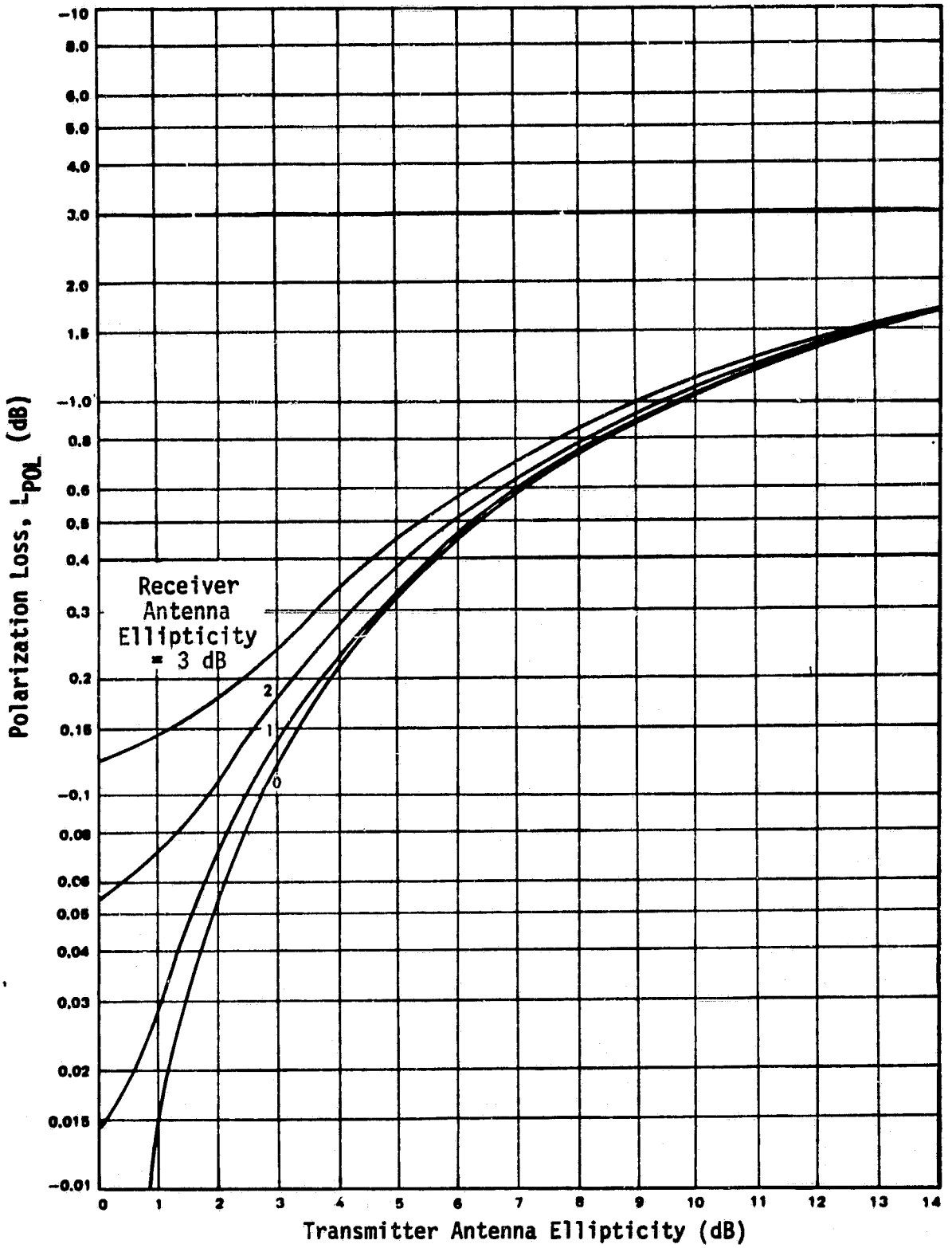


Figure 3.33. Nominal Polarization Loss Versus Transmitter and Receiver Antenna Ellipticity

These circuit losses must be measured to find the actual radiated and received powers. The individual component losses are a complex function of both dissipative and mismatch (reflective) losses, which are commonly lumped together as insertion losses used to characterize the net attenuation of L_T and L_R. However, the mismatch loss concept is somewhat misleading in that the reflected waves from each component in the system combine such that the total system circuit loss is not truly multiplicative. The concept of impedance transformations is required to account for the mismatch losses of the final circuit, and it is therefore essential to actually measure the final total attenuation of the entire circuit. The individual insertion and mismatch losses of each component are specified to be as low as possible to ensure that the final circuit losses are reasonable.

The mismatch loss, L_{MM} (in dB) is equal to

$$L_{MM} = 10 \log (1 - |\rho|^2), \tag{3-102}$$

where the reflection coefficient ρ is obtained by the ratio $|\rho| = |\text{reflected voltage/incident voltage}|$, which is obtained from voltage standing-wave ratio (VSWR) measurements:

$$\text{VSWR} \triangleq \frac{1 + |\rho|}{1 - |\rho|} = \frac{\text{maximum peak voltage}}{\text{minimum peak voltage}}. \tag{3-103}$$

Figure 3.34 plots L_{MM} versus VSWR.

The final total system mismatch loss then can vary depending on the circuit configuration because it relies on the phase relationships of reflected waves (i.e., VSWR) of each component in the system. The insertion loss consists of both the attenuation and the mismatch losses so that it characterizes the decrease in available signal output. Due to the complex nature of the cascaded mismatch losses, the total system insertion loss can be obtained only by measurement of the final system configuration.

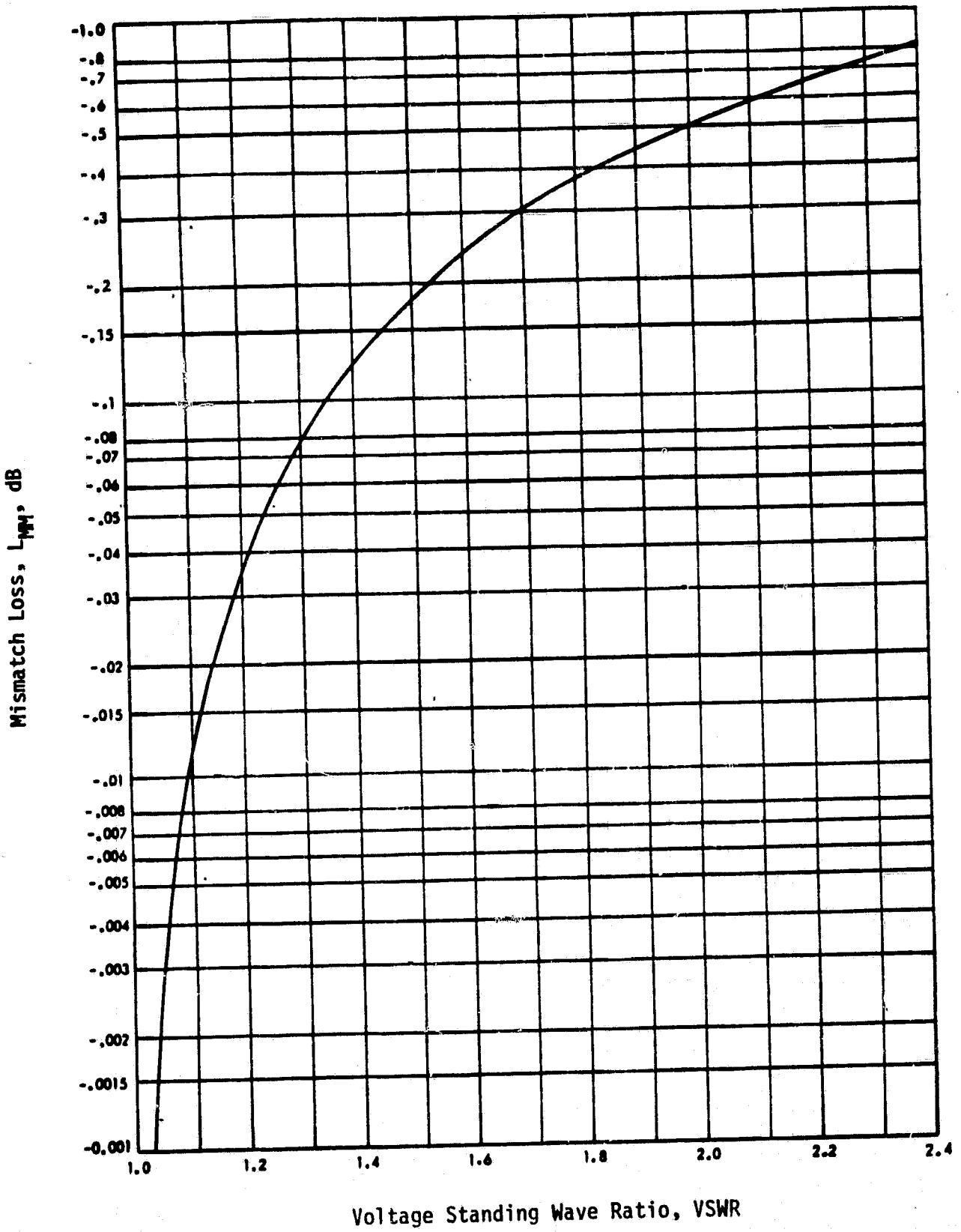


Figure 3.34. Mismatch Loss Versus Voltage Standing-Wave Ratio

3.2.7 The RF Channel

The basic communication transmission path for space communication consists of a propagation medium which is a vacuum (i.e., free space). However, significant effects may be observed if the RF signal is transmitted through the atmosphere and ionosphere to the ground. The atmosphere and ionosphere may alter the signal power, the frequency spectrum, the signal phase relationships, and the ray path because of their dissipative and dispersive nature.

3.2.7.1 Space loss

The power P_R available at the aperture of a receiving antenna from an isotropic transmitting antenna is equal to the ratio of the area of the receiving antenna aperture A_R to the area of a sphere whose radius is the distance d between the two antennas times the power output of the transmitting antenna P_T. Thus,

$$P_R = \left(\frac{A_R}{4\pi d^2}\right) P_T . \tag{3-104}$$

But, substituting for A_R from (3-94),

$$\frac{P_R}{P_T} = G_R\left(\frac{\lambda}{4\pi d}\right)^2 . \tag{3-105}$$

The so-called space loss L_S is defined as

$$L_S = \left(\frac{\lambda}{4\pi d}\right)^2 . \tag{3-106}$$

For d in kilometers and the carrier frequency f in MHz, the space loss can be expressed in dB as

$$L_S\ (\text{dB}) = 32.446 + 20 \log_{10} d + 20 \log_{10} f \tag{3-107}$$

or, if d is in nautical miles,

$$L_S\ (\text{dB}) = 37.796 + 20 \log_{10} d + 20 \log_{10} f . \tag{3-108}$$

3.2.7.2 Atmospheric and ionospheric effects

When an RF signal passes through an atmosphere, it is altered in a manner generally inimical to communication systems performance. The communication system designer must be aware of atmospheric effects since they can have a large effect on signal power, system noise temperature, system frequency choice, and even antenna pointing. The major effects on signals at S-band and higher frequencies are:

(1) Attenuation of the signal caused by weather conditions such as rain, clouds, wind, snow, etc.

(2) Increase in the system noise temperature due to weather

(3) Bending and defocusing of the signal due to the differential index of refraction between the atmosphere and the ionosphere.

Weather conditions at the ground station can cause severe attenuation of the electromagnetic signal. The effect is frequency dependent so that, while X-band and higher frequency signals are severely attenuated, the effect on an S-band signal may be negligible. The atmosphere contains oxygen, water vapor, and water droplets which absorb and reradiate electromagnetic waves. Since the oxygen content is fairly constant in time, it produces a constant background noise. However, the water vapor and water droplet content of the atmosphere fluctuate radically in time and in geographical location as the weather changes.

The atmospheric water content causes two major effects: attenuation and increase in noise temperature of the ground system. At centimeter wavelengths, attenuation is caused primarily by water droplets of rain and fog, with rain the major offender. For rain, the attenuation depends on such parameters as drop size, geometry and distribution, rain rate, signal propagation distance, and the signal frequency. For clouds, the attenuation depends on the water content and cloud thickness.

The attenuation due to rain increases with frequency and increasing precipitation rate. Figure 3.35 shows the frequency dependence of attenuation due to rain and fog. The contribution to the atmospheric loss, L_A, due to rain is a function of equivalent path

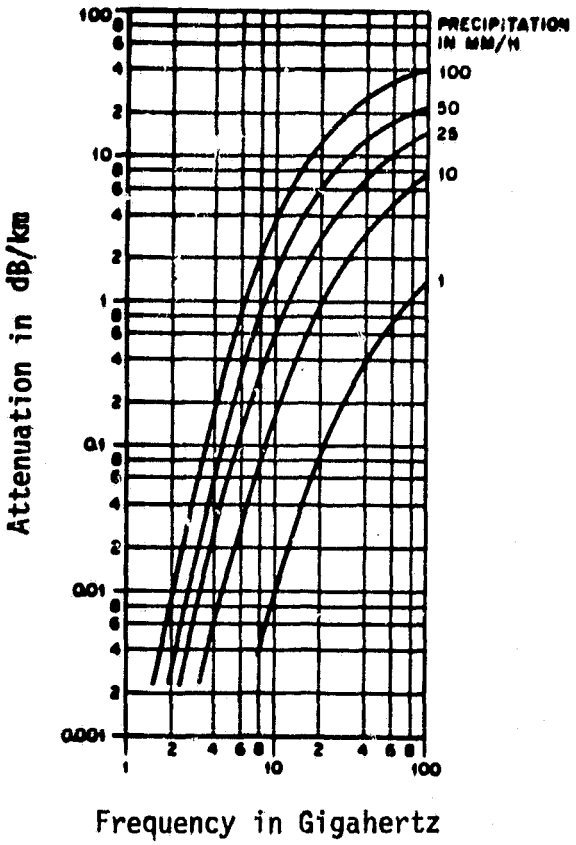


Figure 3.35. Attenuation Due to Precipitation [9]

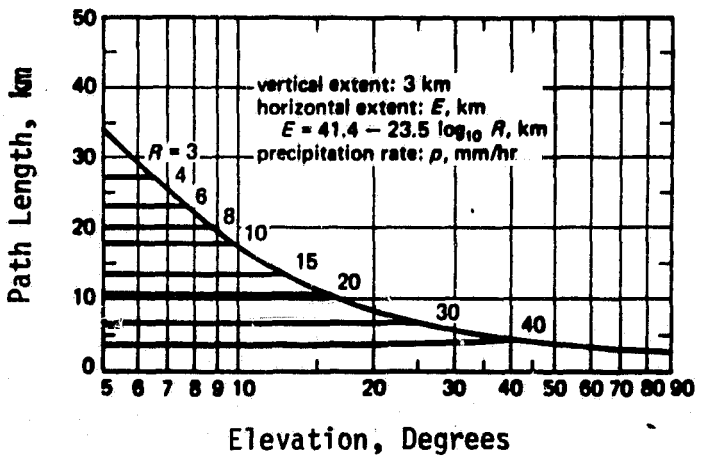


Figure 3.36. Effective Path Length Through Rain Versus Elevation Angle and Rain Rate [10]

length through the rain. Figure 3.36 shows the equivalent path length versus elevation angle for various rainfall rates. At very heavy rain rate , R > 40 mm/hr, the effective path length is equal to 4 km and is nearly independent of elevation angle.

At frequencies above 10 GHz, the atmospheric attenuation due to water vapor and oxygen can have significant effects on the communication link. These effects are even greater at some frequencies than the effects of rainfall. Curves for both 0% and 100% humidity are shown in Figure 3.37. Water vapor causes the peaks at 22.2 and 183.3 GHz, and oxygen causes a family of absorption lines at 60 GHz (56-65 GHz) and an isolated line at 118.8 GHz, as shown in this figure. The curves show the minimum attenuation values that occur between the oxygen absorption lines. Hence, the individual lines are not visible at 60 MHz in Figure 3.37. The very high attenuations at the oxygen absorption frequencies, which are over 100 dB more than the background, render these frequencies unusable for earth/satellite links.

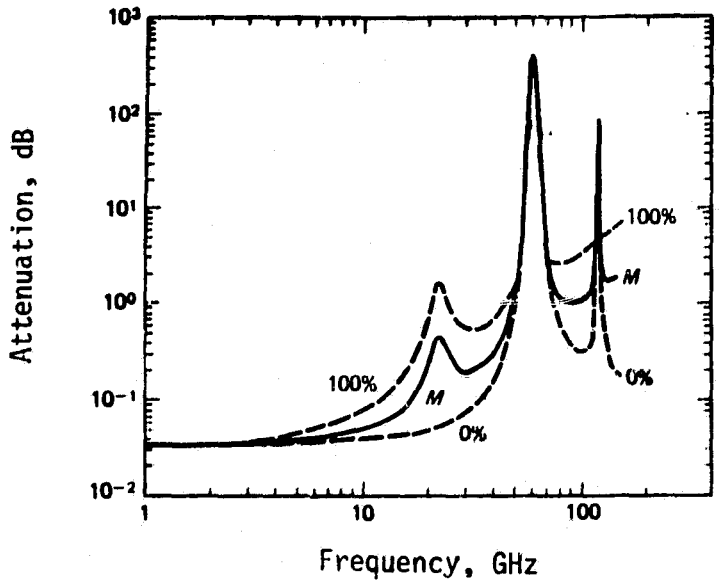


Figure 3.37. Zenith Attenuation Versus Frequency for Various Humidity Levels [11]

Attenuation for other elevation angles θ_e can be computed using the effective path length r as related to the vertical extent of the troposphere $r_v = 10$ km and the effective earth radius r_0:

$$r = r_0^2 \sin^2 \theta_e + r_v(2r_0 + r_v)^{1/2} - r_0 \sin \theta_e . \qquad (3\text{-}109)$$

The total attenuation A (dB) for a given effective path length r relative to the vertical attenuation A_v (dB) is given by $A/A_v = r/r_v$. Oxygen and water vapor attenuation, as a function of elevation angle, is plotted in Figure 3.38, using data for standard atmospheric conditions and the value of r calculated above.

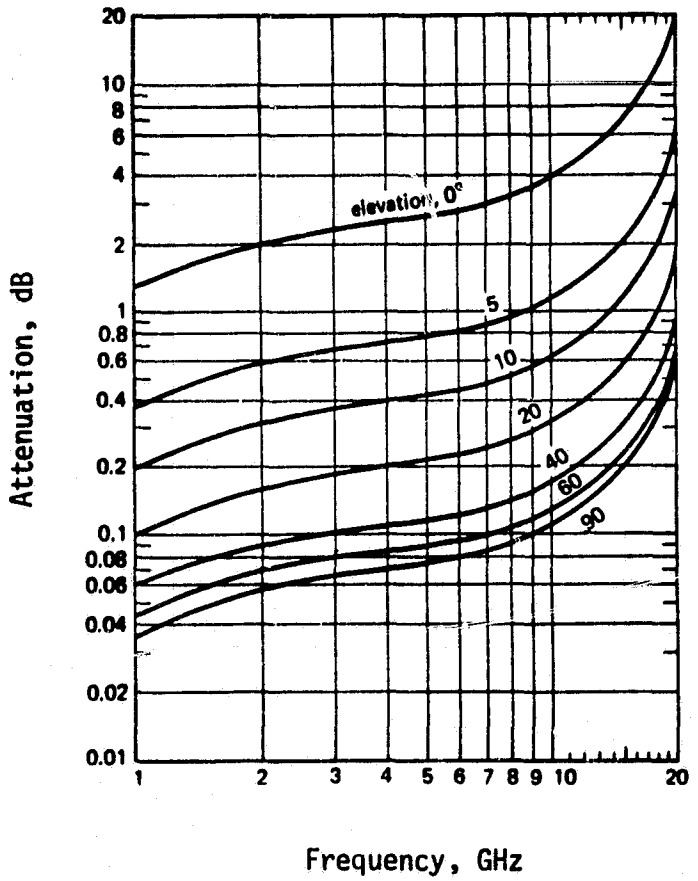


Figure 3.38. Atmospheric Attenuation Due to Oxygen and Water Vapor [10]

Another frequency-dependent attenuation effect is ionospheric scintillation. Ionospheric effects are of primary importance at frequencies below 1 GHz. However, even at the microwave frequencies above 1 GHz, the effects can still be of significance.

Ionospheric scintillation is caused by irregularities in the the night-time F-layer ranging from 200 to 600 km in altitude. The irregularities appear to be elongated regions with the major axis parallel to the earth's magnetic field lines. Axial ratios greater than 60 to 1 have been measured. The effect of these irregularities is to alternatively produce signal enhancement and negative fades. The refractive index of the ionosphere is a function of radio frequency, and irregularities in the ionosphere have progressively less effect as the frequency increases. The exact nature of the frequency dependence appears to depend somewhat on the ionospheric conditions, but the absolute value of the scintillation attenuation seems to vary approximately as the square of the wavelength.

Measured data for a 95% confidence of fading less than a given value yield results generally consistent with the following:

f	250 MHz	2.3 GHz	7.3 GHz
Attenuation, dB	22 dB	2 dB	<0.5 dB

For the most part, scintillation effects above 7.3 GHz are practically absent.

For wideband signals, the phase distortion of the ionosphere becomes an important factor. For propagation through a uniform ionosphere, the phase shift can be written as [12]:

$$\beta = \sqrt{\left(\omega^2 - \omega_p^2\right)/c^2} \quad , \tag{3-110}$$

where

$$\omega_p^2 = \frac{Nq^2}{\varepsilon_0 m} = 324\,\pi^2 N \tag{3-111}$$

and

c = velocity of light in free space (m/sec)

$\omega = 2\pi f$, radian frequency (rad/sec)

N = free electron density (number/m^3)

q = electron charge (coulombs)

ε_0 = dielectric constant of free space .

For propagation through distance r, the total phase shift is $\Theta = \beta r$. To determine the phase distortion, β can be expanded in a power series about the center frequency ω_c as follows.

$$\Theta = \beta r = r\left\{\beta(\omega_c) + (\omega - \omega_c)\left.\frac{d\beta}{d\omega}\right|_{\omega_c} + \frac{1}{2}(\omega - \omega_c)^2\left.\frac{d^2\omega}{d\omega^2}\right|_{\omega_c} + \cdots\right\}$$

$$= \frac{r}{c}\left\{\left(\omega_c^2 - \omega_p^2\right)^{1/2} + \frac{\omega_c(\omega - \omega_c)}{\left(\omega_c^2 - \omega_p^2\right)^{1/2}} - \frac{\omega_p^2(\omega - \omega_c)^2}{2\left(\omega_c^2 - \omega_p^2\right)^{3/2}} + \cdots\right\} \qquad (3\text{-}112)$$

The first term of the expansion is a fixed carrier phase shift which does not affect the correlation amplitude. The second term represents a fixed delay which the receiver tracks. The third term is the first actual distortion term Θ_d. Thus,

$$\Theta_d(\omega - \omega_c) = -\frac{r\omega_p^2(\omega - \omega_c)^2}{2c\left(\omega_c^2 - \omega_p^2\right)^{3/2}} \simeq \frac{r\omega_p^2(\omega - \omega_c)^2}{2c\,\omega_c^3}, \quad \text{for } \omega_p < \omega_c . \qquad (3\text{-}113)$$

For notation convenience, let

$$\omega_0 = \omega - \omega_c \qquad (3\text{-}114)$$

and

$$D = \frac{r\omega_p^2}{2c\,\omega_c^3} . \qquad (3\text{-}115)$$

To determine the performance loss due to phase distortion, consider that the transmitted spectrum is ideal rectangular bandpass filtered to bandwidth 2B. The ionsophere then distorts the phase.

The correlation output of a wideband PN spread spectrum signal can be found by multiplying the filter characteristic times the filter response, then taking the Fourier transform. Thus, the peak correlation R(0) is given by

$$|R(0)| = \left| \frac{T_c}{2\pi} \int_{-\pi B}^{\pi B} \left[\frac{\sin(\omega_0 T_c/2)}{\omega_0 T_c/2} \right]^2 e^{jD\omega_0^2} \, d\omega_0 \right| , \qquad (3\text{-}116)$$

where T_c is the PN chip duration. Equation (3-116) is numerically evaluated (step size $\delta\omega = 2\pi/100\ T_c$) in Figure 3.39 for correlation loss (performance loss). Note that the loss can increase as the bandwidth becomes wider due to phase cancellation when the distortion is present. The normalization in Figure 3.39 is the phase distortion at the first null. That is,

$$\Theta_d(2\pi/T_c) = 4\pi^2 D/T_c^2 = \text{Phase distortion at first null (in radians)} . \qquad (3\text{-}117)$$

In Figure 3.39, the performance becomes significant when the phase distortion becomes greater than 1 radian. Therefore, the maximum-allowed distortion for a PN system might be set at 2 radians. Then,

$$\Theta_d(\omega_0) = 2 = \frac{r\omega_p^2}{2c\omega_c^3} \left(\frac{2\pi}{T_c}\right)^2 \qquad (3\text{-}118)$$

or the maximum allowed chip rate $R_c = 1/T_c$ is

$$R_c < \sqrt{\frac{2c f_c^3}{\pi r f_p^2}} . \qquad (3\text{-}119)$$

It may be noted [13] that a worst-case value of Nr is roughly 2×10^{18} electrons/m^2. Assuming $r = 10^5$ m for this worst case, then $N = 2^{13}$ electrons/m^3 and, from (3-111), $f_p = 40$ MHz. For S-band at $f_c = 2200$ MHz, the maximum-allowable PN chip rate is 113 MHz.

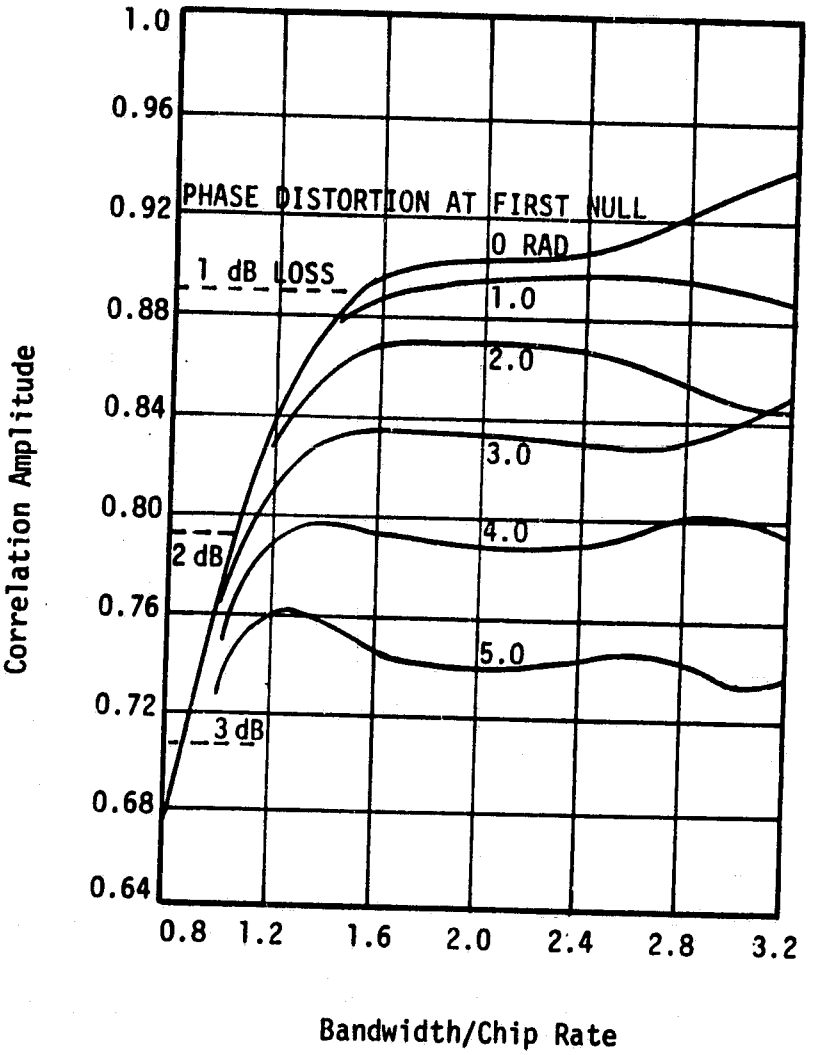


Figure 3.39. Performance Loss Due to Phase Distortion and Finite Bandwidth

3.2.8 System Noise

Noise and interference from other communication systems are two factors limiting the useful operation of all radio equipment. There are a number of different sources of radio noise, the most important being atmospheric noise, galactic noise, man-made noise, and receiver noise. The levels of noise may be expressed in various ways; perhaps the most convenient way is to refer the received noise power to the thermal noise power at a reference ambient temperature of 290°K.

In estimating the noise level at the receiver due to external sources, the gain and orientation of the receiving antenna must be considered. Since, in general, the available noise power is proportional to bandwidth, it may be expressed in terms of an effective antenna noise factor f_a, which is defined by

$$f_a = \frac{N_a}{kT_0B} = \frac{T_a}{T_0} , \qquad (3\text{-}120)$$

where N_a = noise power available from an equivalent loss-free antenna (watts)

k = Boltzmann's constant = 1.38×10^{-23} watts per °K per Hz

T_0 = reference ambient temperature, taken as 290 K

B = effective receiver noise bandwidth (Hz)

T_a = effective antenna temperature in the presence of external noise (°K); that is,

$$T_a = \frac{N_a}{kB} . \qquad (3\text{-}121)$$

3.2.8.1 Sources of noise

Figure 3.40 summarizes the sources of noise for a ground terminal. The noise level f_a (i.e., f_a in dB) above kT_0B is shown in Figure 3.40, as well as T_a (in °K), as a function of frequency for the more important sources of radio noise. Atmospheric noise curves [14] are shown for New York City for the summer nighttime and winter daytime. These two curves represent the extremes in expected

levels of atmospheric radio noise at this location. This atmospheric noise is mostly produced by lightning discharges in thunderstorms. The noise level thus depends on frequency, time of day, weather, season of the year, and geographical location. Subject to variations due to local stormy areas, noise generally decreases with increasing latitude on the surface of the earth. Noise is particularly severe during the rainy seasons in areas such as the Caribbean, East Indies, equatorial Africa, northern India, etc. However, note from Figure 3.40 that this atmospheric noise is not a significant contribution to the total system noise for UHF and higher frequencies.

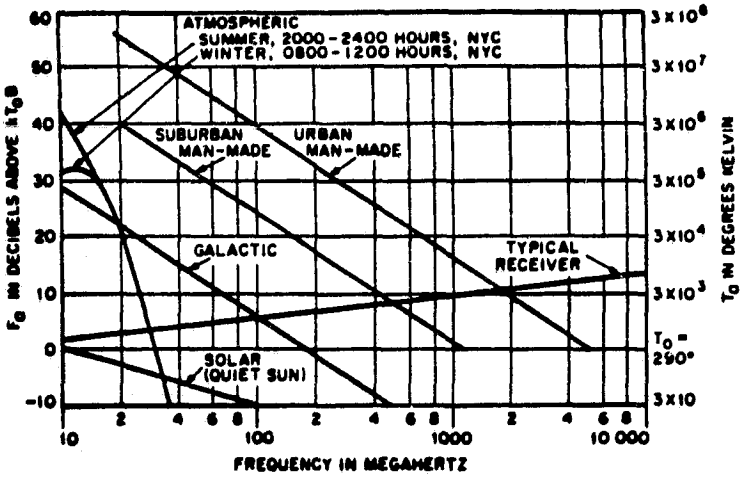


Figure 3.40. Median Values of Average Noise Power Expected from Various Sources (Omnidirectional Antenna Near Surface)

Noise due to atmospheric absorption can be of importance above 1 GHz when low-noise amplifiers are employed at the receiver. The sky temperature for an infinitely sharp beam is given in Figure 3.41. For For an elevation angle of 0.1 radian, the typical effective sky temperature is 13 K for 1 GHz and 25 K at 10 GHz, increasing to 80 K at 20 GHz.

Galactic noise may be defined as the noise at RF caused by disturbances that originate outside the earth or its atmosphere. The chief causes of such radio noise are the sun and a large number of

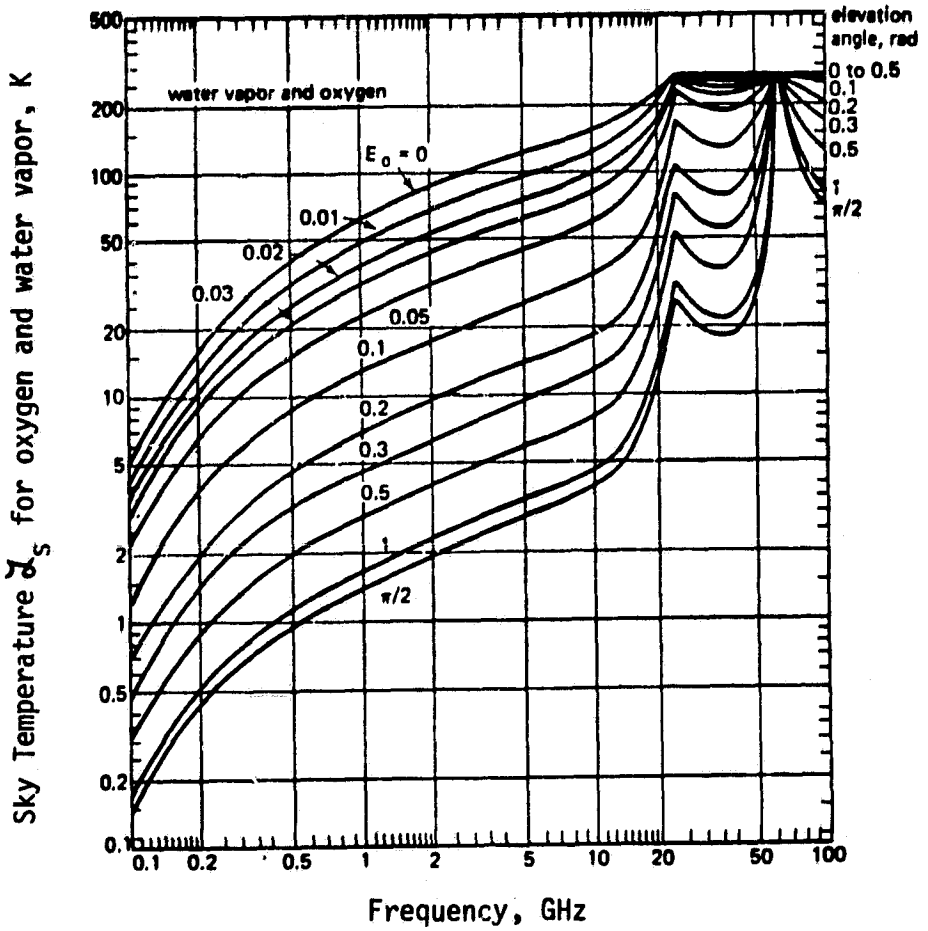


Figure 3.41. Sky-Noise Temperature Due To Reradiation by Oxygen and Water Vapor

discrete radio sources distributed chiefly along the galactic plane. Galactic noise reaching the surface of the earth ranges from about 15 MHz to 100 GHz, being limited by ionospheric absorption at the low end of the spectrum and by atmospheric absorption at the high end. In practice, the importance of galactic noise is restricted to frequencies no lower than about 18 MHz by atmospheric noise, and to frequencies not higher than 500 MHz by receiver noise and antenna gain, as shown in Figure 3.40. However, with a high-gain receiving antenna directed at the sun, the antenna noise temperature may exceed 290 K at frequencies as high as 10 GHz.

Figure 3.42 shows the level of galactic noise in decibels relative to a noise temperature of 290 K when receiving on a half-wave dipole. The noise levels shown in this figure assume no atmospheric absorption typical of spacecraft and refer to the following sources of galactic noise.

Galactic Plane: Galactic noise from the galactic plane in the direction of the center of the galaxy. The noise levels from other parts of the galactic plane can be as much as 20 dB below the levels given in Figure 3.42.

Quiet Sun: Noise from the "quiet" sun; that is, solar noise at times when there is little or no sunspot activity.

Disturbed Sun: Noise from the "disturbed" sun. The term disturbed refers to times of sunspot and solar flare activity.

Cassiopeia: Noise from a high-intensity discrete source of cosmic noise known as Cassiopeia. This is one of more than a hundred known discrete sources, each of which subtends an angle at the earth's surface of less than half a degree.

The level of cosmic noise received by an antenna directed at a noise source may be estimated by correcting the relative noise levels with a half-wave dipole (from Figure 3.42) for the receiving antenna gain realized on the noise source. Since the galactic plane is an extended nonuniform noise source, free-space antenna gains cannot be realized, and 10 to 15 dB is approximately the maximum antenna gain that can be realized here. However, on the sun and other discrete sources of galactic noise, antenna gains of 50 dB or more can be obtained.

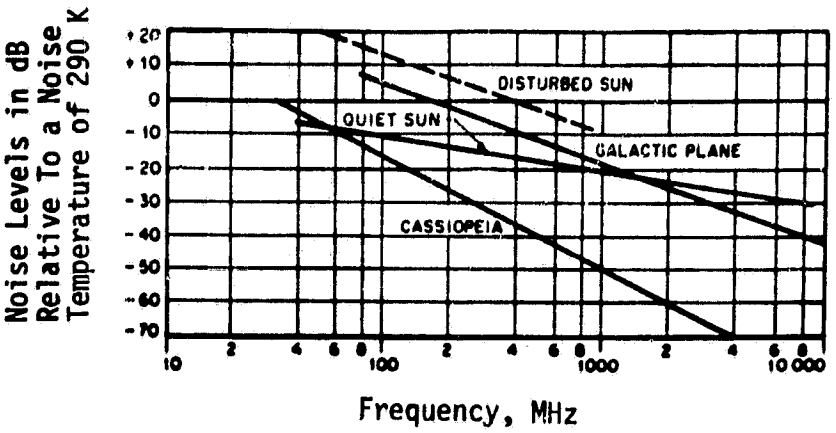


Figure 3.42. Galactic Noise Levels for a Half-Wave-Dipole Receiving Antenna

The effective antenna noise temperature is defined by

$$T_{ae} = \frac{\int_{4\pi} T(\Omega)G(\Omega)d\Omega}{\int_{4\pi} G(\Omega)d\Omega} \qquad (3\text{-}122)$$

where $T(\Omega)$ 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

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} = 43,000 K, entire beam intercepted by the sun.

The effective antenna noise temperature (brightness temperature) decreases significantly as the RF frequency is increased above 2 GHz, as is shown in Figure 3.43.

The remaining external noise source shown in Figure 3.40 is man-made noise. The amplitude of man-made noise decreases with increasing frequency and varies considerably with location. It is

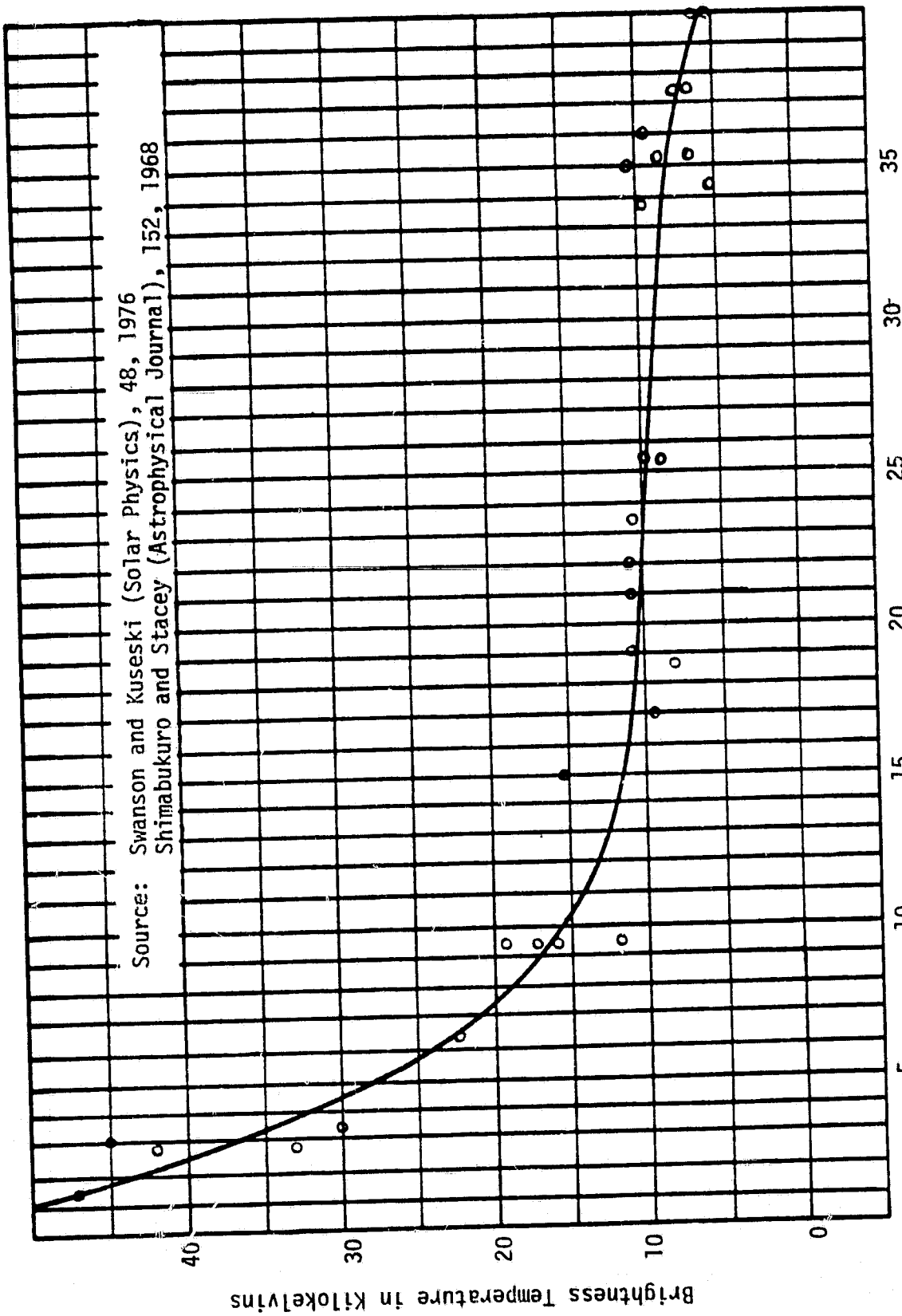


Figure 3.43. Solar Brightness Temperature of the Quiet Sun Versus Frequency (1 to 40 GHz)

due chiefly to electric motors, neon signs, power lines, and ignition systems located within a few hundred yards of the receiving antenna; certain high-frequency medical appliances and high-voltage transmission lines may, however, cause interference at much greater distances. The average level of man-made noise power can be 16 dB or more higher in urban than in suburban areas of the United States; in remote rural locations, the level may be 15 dB below that experienced in a typical surburban site. In quiet remote locations, the noise level from man-made sources will usually be below galactic noise in the frequency range above 10 MHz.

Propagation of man-made noise is chiefly by transmission over power lines and by ground wave; however, it may also be by ionospheric reflection at frequencies below about 20 MHz.

Measurements indicate that the peak level of man-made noise is not always proportional to bandwidth for bandwidths greater than about 10 kHz. According to the best available information, the peak field strengths of man-made noise (except diathermy and other narrow-band noise) increase as the receiver bandwidth is substantially increased for bandwidths greater than about 10 kHz, as shown in Figure 3.44.

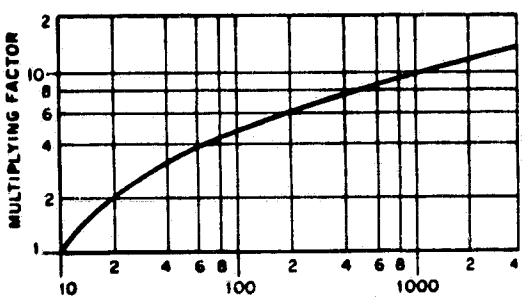


Figure 3.44. Bandwidth Factor for Man-Made Noise

Internal noise sources are referred to as receiver noise, and the RF and IF stages, mixers, etc., often add more noise to the system than the external noise. In the next subsection, receiver noise is computed and added to the external noise in the proper way to compute the overall system noise temperature.

3.2.8.2 Noise temperature, figure and spectral density

In the system considered here, the noise contribution to the signal can be approximated by an additive white Gaussian process. Thus, the noise power into the detector is equal to the system noise spectral density, N_0, times the system noise bandwidth, B_n.

The two-sided noise bandwidth, B_{nj}, of a system element j is defined by:

$$B_{nj} = \frac{\frac{1}{2\pi}\int_{-\infty}^{\infty} G_j(\omega)d\omega}{[G_j(\omega)]_{max}}, \qquad (3\text{-}123)$$

where $G_j(\omega) = |H_j(i\omega)|^2$ is the gain function of the element j, $H_j(i\omega)$ is its frequency characteristic, and $i = \sqrt{-1}$.

B_{n_j} can be viewed as the bandwidth of an ideal bandpass filter whose maximum response is the same as that of $H_j(i\omega)$ and whose output power in the presence of white noise is equal to that of $H_j(i\omega)$.

The system noise bandwidth is then:

$$B_n = \frac{\int_{-\infty}^{\infty} G_s(f)df}{[G_s(f)]_{max}} \qquad (3\text{-}124)$$

where

$$G_s(f) = \prod_{j=1}^{M} G_j(f) \qquad (3\text{-}125)$$

is the system gain function of a system of M elements.

The contribution to the noise spectral density of a system element j, N_{0j}, can be represented by

$$N_{0j} = kT_jG_j, \qquad (3\text{-}126)$$

where

k = Boltzmann's constant

G_j = gain of element $j \triangleq [G_j(f)]_{max}$

$T_j \triangleq$ effective noise temperature of element j.

The effective noise temperature T_j is that temperature at the input to a noiseless element which would result in a noise spectral density N_{0j} at its output.

The effective system noise temperature can be found by considering each element separately and combining them, as shown in the following paragraphs.

System Noise Temperature, T_{es}. A receiving system is a cascade of electronic elements, each of which contributes noise power to the system. The total noise power depends on the gains and bandwidths of the various components. Since the ratio of signal power to noise power into the detector determines the system performance, it is vital to know the noise power contribution to the system.

It is convenient to replace the receiver by an equivalent noise source and define the effective noise temperature, T_n, of the receiver by

$$N_0 = \frac{N}{B_n} = k(T_n + T_i)G_s \triangleq kT_{es}G_s , \qquad (3\text{-}127)$$

where

N_0 = system noise spectral density

N = total noise power

B_n = receiver noise bandwidth

$T_i = N_{0i}/k$, where N_{0i} is the input noise spectral density.

G_s is the total system gain from the antenna terminals to the measurement point, $kT_iG_sB_n$ is the input noise power in the receiver bandwidth multiplied by the system gain, and T_i is equal to T_{ae}, the effective antenna noise temperature.

$T_n + T_i$ is called the effective system temperature T_{es}, referenced to the antenna terminals. Note that T_{es} differs from T_n in that different

systems are involved. The theory is identical for both, but it is convenient to use T_n since T_1 is an exogenous variable.

Noise Temperature of an Element. The effective noise temperature referenced to the input of an element can be determined from the ratio of input to output SNR. For a physical temperature T_{ij} at the input to element j, we define the noise factor or noise figure NF of element j by

$$NF_j(T_{ij}) = \frac{S_{inj}/N_{inj}}{S_{outj}/N_{outj}}, \tag{3-128}$$

where

S_{inj} = input signal power

N_{inj} = input noise power $\triangleq kT_{ij} B_{nj}$

S_{outj} = output signal power

N_{outj} = output noise power

Remembering that the output noise power is the sum of the input noise power and the noise power generated by an element, it can be shown that

$$NF_j(T_{ij}) = 1 + \frac{T_j}{T_{ij}}, \tag{3-129}$$

where T_j, the effective noise temperature of element j, is that temperature at the input of a noiseless element which would result in N_{outj} at the output.

For convenience in comparing different elements, the noise figure is usually measured at an input noise temperature:

$$T_{ij} = T_0 = 290 \text{ K}, \tag{3-130}$$

giving

$$NF_j = 1 + T_j/T_0 \tag{3-131}$$

or

$$T_j = (NF_j - 1)T_0. \tag{3-132}$$

The unmodified term "noise figure" in the literature refers to this $NF_j = NF_j(T_0)$.

Note that

$$NF_j(T_{ij}) = (NF_j - 1)\frac{T_0}{T_{ij}} + 1 , \tag{3-133}$$

so that the effective noise temperature measured at T_{ij} is

$$T_j = \left(NF_j(T_{ij}) - 1\right)T_{ij} \equiv (NF_j - 1)T_0 . \tag{3-134}$$

For an active element, such as an amplifier, NF_j determines T_j independent of the input noise temperature and the physical temperature of the input. However, for a passive element (such as a cable) of gain G_j (loss $L_j = 1/G_j$) in thermal equilibrium with its environment,

$$T_{ij} = T_{outj} = \text{ambient temperature} \triangleq T \tag{3-135}$$

and

$$NF_j(T_{ij}) = \frac{1}{G_j} = L_j , \tag{3-136}$$

so that the effective noise temperature of a loss is

$$T_{LJ} = (L_j - 1)T , \tag{3-137}$$

where T is the ambient temperature. Note that for a loss, the effective noise temperature is dependent on the physical temperature of the component. The noise factor NF_j could be defined as

$$NF_j = 1 + (L_j - 1)\frac{T}{T_0} . \tag{3-138}$$

Since NF_j for a loss depends on the physical temperature of the component, it is not an especially useful concept and, when speaking of the noise temperature of a passive element, (3-137) is used instead.

Receiver System Temperature. For a cascaded system of M elements, passive or active, it can be shown that the system effective noise temperature can be written as

$$T_n = T_1\frac{B_{n_1}}{B_n} + \frac{T_2}{G_1}\frac{B_{n_2}}{B_n} + \cdots + \frac{T_M}{G_1G_2\cdots G_{M-1}}\frac{B_{n_M}}{B_n} \tag{3-139}$$

where

$$T_n = \frac{N_s}{k B_n G_s}$$

N_s = noise power contributed by the system = $N_{out} - N_i G_s$

N_{out} = total output noise power of the system due to N_s and the input noise

The temperatures T_j are referenced as shown in Figure 3.45.

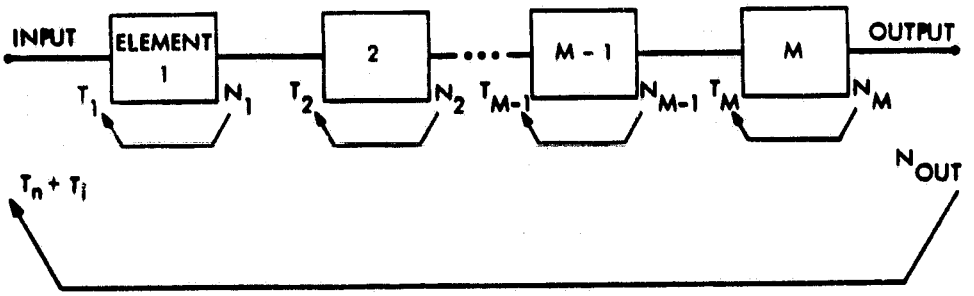


Figure 3.45. Noise Powers and Their Associated Noise Temperature Reference Points

For $B_{nj} = B_n$, we have the more common form of the system effective noise temperature:

$$T_n = T_1 + \frac{T_2}{G_1} + \cdots + \frac{T_M}{G_1 G_2 \cdots G_{M-1}} . \qquad (3\text{-}140)$$

We shall make this assumption in what follows.

Note that, if the gain of one element in the cascade is very high and the noise figures of succeeding elements are not large, the expression for effective noise temperature may be truncated after the noise temperature term for that element. For example, if $G_3 \gg 1$,

$$T_n \simeq T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 G_2} \qquad (3\text{-}141)$$

and $N_s \simeq kT_n B_n G_s$, since the noise power contributions from elements 4 to M are small.

The effective system noise temperature as defined here is the temperature at the input to the system which would result in a noise spectral density $N_0 = kT_{es}G_s$ at the system output. It is also called the system noise temperature referenced to the system input.

Frequently, the system noise temperature is noted as referenced to a convenient point within the system. Thus, one may speak of the system noise temperature referenced to the receiver input. This noise temperature is defined as that temperature at the reference point which would result in the system noise spectral density at the system output. What this change of reference point amounts to is a lumping of the appropriate subsystem gains into the definition of noise temperature.

The system noise temperature referenced to the input of element j would be:

$$T_{ej} = G_1 G_2 \cdots G_{j-1} T_{es} \qquad (3\text{-}142)$$

and

$$N_0 = kT_{ej}\left(G_j G_{j+1} \cdots G_M\right) \equiv kT_{es}G_s \,. \qquad (3\text{-}143)$$

The term in parentheses is called G_s' and is the system gain which follows the reference point.

3.2.8.3 Cumulative noise (tandem links)

A tandem link contains one or more transponders as relays. Two examples of tandem links are (1) the Orbiter using the TDRSS as a relay and (2) a payload using the Orbiter "bent-pipe" mode. The bent-pipe mode is a two-relay tandem link since the Orbiter is used as a relay by the payload and the Orbiter uses the TDRSS as a relay. Figure 3.46 illustrates a simplified two-relay tandem link.

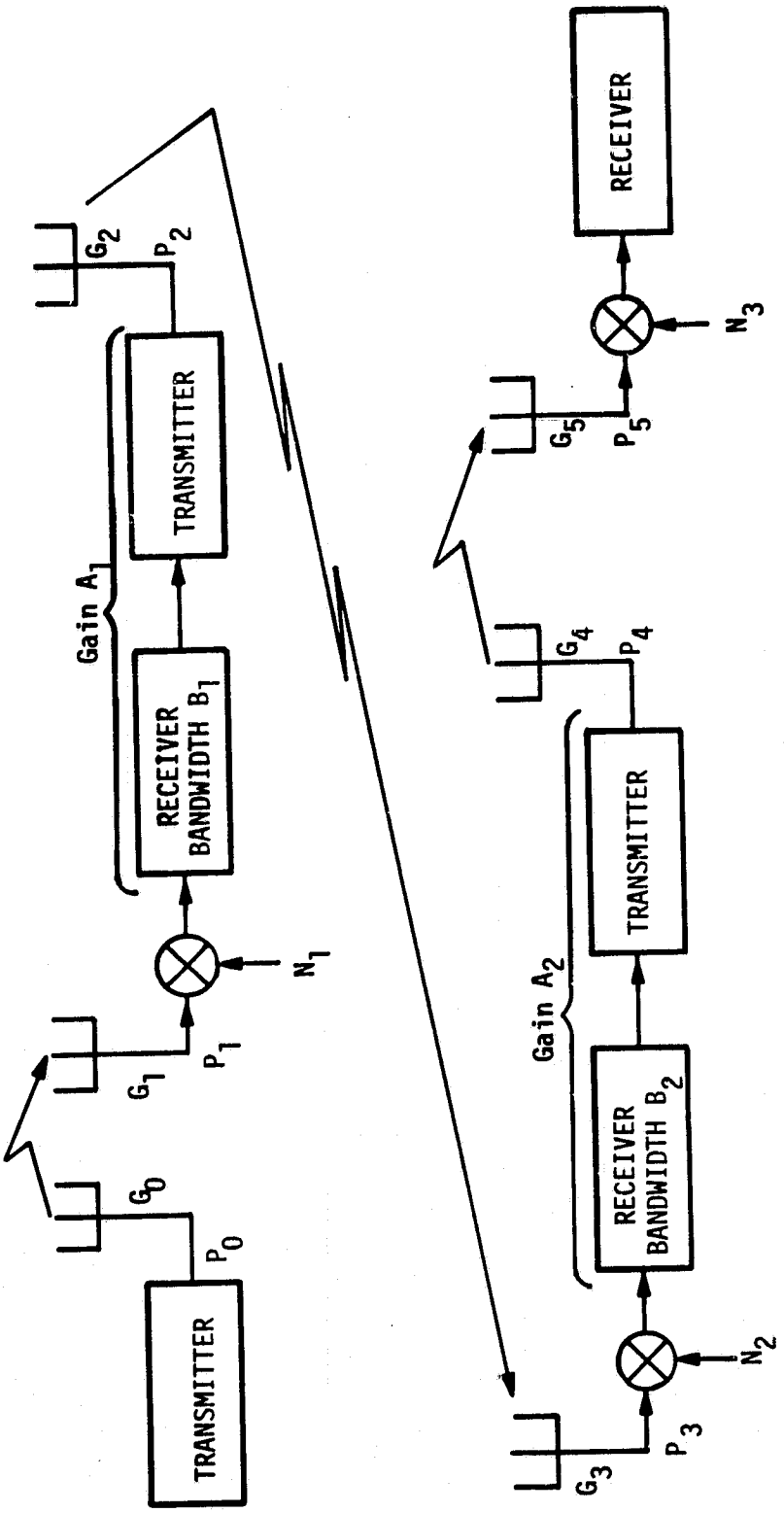


Figure 3.46. Simplified Tandem Link Configuration

In a tandem link, the desired signal, as well as noise, is amplified through the relays. Therefore, the noise is cumulative. Consider the example shown in Figure 3.46. The total received signal power in the first relay is P_1, which is received along with a noise power of N_1B_1, where N_1 is received noise spectral density and B_1 is the transponder bandwidth for a linear transponder. The total power to be transmitted is P_2, which is shared proportionately by the signal and noise according to their respective powers. For a nonlinear channel, the signal is suppressed or enhanced by α_1 with respect to the noise, depending on the ratio $\rho_1 = P_1/N_1B_1$. The output SNR is related to the input SNR by

$$(SNR)_{OUT} = \alpha(SNR)_{IN} \tag{3-144}$$

where, for a hard limiter, the limiter suppression factor α is given by [16]

$$\alpha = \frac{\frac{\pi}{4}\exp(-\rho)\left[I_0(\rho/2) + I_1(\rho/2)\right]^2}{1 - \frac{\pi}{4}\rho\exp(-\rho)\left[I_0(\rho/2) + I_1(\rho/2)\right]^2}, \tag{3-145}$$

where $I_n(\ )$ is the modified Bessel function of order n.

For the linear transponder the transmitted relay power P_2 depends on the transponder gain,

$$A_1 = \frac{P_2}{P_1 + N_1B_1} \tag{3-146}$$

and the total input power level $P_1 + N_1B_1$. Thus, the transmitted power is

$$P_2 = A_1(P_1 + N_1B_1) = P_2\gamma_1\left[P_1 + N_1B_1\right], \tag{3-147}$$

where $\gamma_1 = 1/(P_1 + N_1B_1)$ is the denominator of the power-sharing factor between signal and noise. For a hard-limited transponder,

$$P_2 = P_2\gamma_1'\left[\alpha_1\rho_1 + 1\right], \tag{3-148}$$

where $\gamma_1' = 1/(\alpha_1\rho_1 + 1)$.

The signal plus noise is received by the next transponder, along with its noise power N_2B_2, for a total received power-plus-noise of

$$P_3 + N_2B_2 = P_2L_2\gamma_1\left[P_1 + N_1B_1\right] + N_2B_2 \tag{3-149}$$

or

$$P_3 + N_2B_2 = P_2L_2\gamma_1'\left[\alpha_1\rho_1 + 1\right] + N_2B_2, \tag{3-150}$$

where L_2 represents the system losses and gains over the second link. By similar arguments, the transmitted power at the second transponder output is

$$P_4 = P_3\gamma_2\left(P_2L_2\gamma_1\left[P_1 + N_1B_1\right] + N_2B_2\right), \tag{3-151}$$

where the power-sharing factor

$$\gamma_2 = \left(P_2L_2\gamma_1\left[P_1 + N_1B_1\right] + N_2B_2\right)^{-1} \tag{3-152}$$

For a hard-limited transponder,

$$P_4 = \left[P_3\gamma_2'\,\alpha_2\rho_2 + 1\right],$$

where

$$\rho_2 = \frac{P_2L_2\gamma_1'\alpha_1\rho_1}{P_2L_2\gamma_1 + N_2B_2}, \tag{3-153}$$

$$\gamma_2' = 1/(\alpha_2\rho_2 + 1), \tag{3-154}$$

and the hard-limiter suppression factor is computed from (3-145) using P_2.

The signal is finally received along with ground-station receiver noise N_3B_3 for a total received power for the linear transponders of

$$P_5 + N_3B_3 = P_4L_3\gamma_2\left(P_2L_2\gamma_1\left[P_1 + N_1B_1\right] + N_2B_2\right) + N_3B_3 , \qquad (3\text{-}155)$$

where L_3 represents the system losses and gains over the third link. For the hard-limited transponders, the received power-plus-noise is

$$P_5 + N_3B_3 = P_4L_3\gamma_2\left[\alpha_2\rho_2 + 1\right] + N_3B_3 . \qquad (3\text{-}156)$$

It is seen from this example that noise is added for each relay link. If the SNR for any of the links to the relays is small, the relay transponder noise dominates the output power and is wasted transmitted noise. Thus, if $P_1/N_1B_1 \gg 1$, the output $(SNR)_g$ at the ground station for linear transponders then becomes

$$(SNR)_g = \frac{P_4L_3\gamma_2\ P_2L_2\gamma_1P_1}{P_4L_3\gamma_2(P_2L_2\gamma_1N_1B_1 + N_2B_2) + N_3B_3} . \qquad (3\text{-}157)$$

But, $\gamma_1 \simeq 1/N_1B_1$. Let $P_2L_2 \gg N_2B_2$; then $\gamma_2 \simeq 1/P_2L_2$ and

$$(SNR)_g = \frac{P_4L_3/N_1B_1}{P_4L_3 + N_3B_3} . \qquad (3\text{-}158)$$

If $P_4L_3 \gg N_3B_3$, then

$$(SNR)_g = \frac{P_1}{N_1B_1} , \qquad (3\text{-}159)$$

which is just the received SNR at the first relay transponder. In the cases when the assumptions $P_2L_2 \gg N_2B_2$ and $P_4L_3 \gg N_3B_3$ are not valid, then these links add to the noise as well. In fact, if $P_2L_2 \ll N_2B_2$ and $P_4L_3 \ll N_3B_3$, then in the example,

$$(SNR)_g = \left(\frac{P_1}{N_1B_1}\right)\left(\frac{P_2L_2}{N_2B_2}\right)\left(\frac{P_4L_4}{N_3B_3}\right) , \qquad (3\text{-}160)$$

and the SNRs for each link are multiplied together to give a very small SNR on the ground.

3.2.9 Received Signal-to-Noise Spectral Density

In the previous subsection, SNR was used to illustrate the contribution of signal and noise. To establish the overall system performance, however, the parameters of most interest are: (1) the received signal power to single-sided noise spectral density (P_R/N_0) measured in dB-Hz, (2) the energy per bit to single-sided noise spectral density (E_b/N_0) measured in dB for digital data, and (3) demodulator output $(SNR)_0$ for analog data.

The value of P_R is found from the transmitter power P_T by multiplying by the system losses and the transmitter and receiver antenna gains presented in the previous subsections. The value of N_0 is computed as described in subsection 3.2.8. The resulting P_R/N_0 is independent of the reference point in the system and becomes a fundamental system parameter. By dividing P_R/N_0 by different bandwidths, the SNR at a given point can be computed. For example, in the previous subsection, the SNR at the demodulator (or receiver) input was used. In this case, the demodulator (or receiver) input bandwidth B_{IN} was used to divide P_R/N_0 to compute the $(SNR)_{IN}$. That is,

$$(SNR)_{IN} = \frac{P_R}{N_0 B_{IN}} . \tag{3-161}$$

3.2.9.1 Carrier power-to-noise spectral density

For coherent demodulators, the carrier power-to-noise spectral density (C/N_0) is an important parameter. For residual carrier PM links,

$$C/N_0 = \left(\frac{P_R}{N_0}\right)\left(\frac{P_C}{P_T}\right)\left(\frac{P_T}{P_R}\right) , \tag{3-162}$$

where P_T/P_R is equal to the product of the system losses and antenna gains, and P_C/P_T is the transmitter power allocations to the residual carrier. Subsection 3.2.5 presents the relationships of P_C/P_T for each type of residual carrier PM. Note that another important parameter used in connection with coherent demodulators is the net carrier signal power to noise spectral density (P_{NC}/N_0), which is

$$P_{NC}/N_0 = (C/N_0)L_{IF} \ , \tag{3-163}$$

where L_{IF} is the receiver IF losses which include, as shown in Figure 3.1, noise oscillator loss (L_{OSC}), PN filtering loss (L_{PN}), and PN correlation loss (L_{COR}) for the spread spectrum case, and limiter loss (L_{LIM}) if a bandpass limiter is used before the phase-locked loop (PLL). These receiver IF losses are discussed in detail in subsequent subsections.

For suppressed carrier modulation links, C/N_0 is defined as being equal to P_R/N_0, and

$$P_{NC}/N_0 = (C/N_0)L_{IF} = \left(\frac{P_R}{N_0}\right)L_{IF} \ . \tag{3-164}$$

But there is additional loss included in L_{IF}. As shown in subsection 3.2.10.2, suppressed carrier-tracking loops exhibit a squaring loss L_{SQ}. Included in L_{SQ} is the miscellaneous waveform loss L_{WF} resulting from filtering and nonlinear distortion present in the IF amplifier chain.

3.2.9.2 Information energy-to-noise spectral density

The E_b/N_0 is an important fundamental parameter for digital data links. In general, E_b/N_0 is defined by

$$E_b/N_0 = \frac{ST_b}{N_0} \ , \tag{3-165}$$

where S is the signal power for the particular digital data channel and T_b is the bit duration, or

$$T_b = 1/R_b \ , \tag{3-166}$$

where R_b is the bit rate. When error-correcting coding is used, it is important to distinguish between energy per coded symbol E_s and E_b. An error-correcting code introduces parity bits into the channel and increases the total number of symbols (i.e., information bit plus parity bits) that must be transmitted. Thus, an error-correcting

code is described by a code rate R (symbols/information bit) which is equal to k/n for block codes and 1/n for simple convolutional codes, as presented in subsection 3.2.3.1. In order to transmit the bits at rate R_b, the symbol rate must be

$$R_S = \frac{R_b}{R} . \tag{3-167}$$

Hence, the symbol duration T_s is

$$T_s = \frac{1}{R_s} = \frac{R}{R_b} = RT_b , \tag{3-168}$$

and the energy-per-symbol-to-single-sided noise spectral density E_s/N_0 is given by

$$\frac{E_S}{N_0} = \frac{ST_S}{N_0} = \frac{SRT_b}{N_0} = R\frac{E_b}{N_0} . \tag{3-169}$$

Thus, even though the carrier tracking loop or subcarrier-tracking loop and the bit synchronizer operate at the symbol rate R_s, the output of the decoder is at the information bit rate R_b. Therefore, the fundamental parameter at the information output is E_b/N_0, whether error-correcting coding is used or not.

The received E_b/N_0 computed from P_R/N_0 is given by

$$E_b/N_0 = \left(\frac{P_R}{N_0}\right)\left(\frac{P_D}{P_T}\right)\left(\frac{P_T}{P_R}\right) L_{IF} L_D \frac{1}{R_b} , \tag{3-170}$$

where P_D/P_T represents the transmitter power allocations to the data and L_D is the demodulation loss. Subsection 3.2.5 presents the relationships of P_D/P_T for each type of modulation. The demodulation loss L_D includes the carrier phase noise loss L_ϕ and the miscellaneous waveform losses L_{WF} presented in subsection 3.2.11, the subcarrier demodulation loss L_{SD} presented in subsection 3.2.12, and the bit synchronization loss L_{BS} presented in subsection 3.2.13.

The demodulator output $(SNR)_0$ for analog data cannot be expressed as straightforwardly as C/N_0 or E_b/N_0. The expression for $(SNR)_0$ under various conditions is presented in subsection 3.2.11.2.

The value of $(SNR)_0$ is a function of the predemodulation $(S/N)_I$, the peak carrier deviation ΔF, and the postdemodulation bandwidth B_0. The predemodulation SNR is given by

$$(S/N)_I = \left(\frac{P_R}{N_0}\right) L_{SF} \left(\frac{1}{B_i}\right), \tag{3-171}$$

where L_{SF} is signal filtering loss presented in subsection 3.2.11.2 and B_i is the predemodulation bandwidth.

3.2.10 Carrier-Tracking Loop

Successful transmission of information through a phase-coherent communication system requires, by definition, a receiver capable of determining or estimating the phase and frequency of the received signal with as little error as possible. Quite often the data-bearing signal is modulated on an RF carrier in such a way that a residual component at RF exists in the overall signal power spectrum. This component could be tracked with a narrowband phase-locked-loop (PLL) and used to provide the desired reference signal. On the other hand, the power contained in this residual component does not convey any information other than the phase and frequency of the carrier. Thus, it represents power not available for the transmission of data and in practice, techniques that conserve energy are always of interest. More specifically, the data is often angle modulated on a secondary carrier (i.e., a subcarrier) which, in the absence of a DC component in the data signal spectrum, requires some type of suppressed carrier tracking loop for establishing the coherent subcarrier reference.

3.2.10.1 Discrete carrier-tracking loops

The essential elements of a phase-locked-loop which govern the behavior of a PLL receiver are shown in Figure 3.47. The input signal x(t) is multiplied by the output signal of the VCO in the phase detector. An error signal is generated which is filtered to remove the high-frequency terms of the multiplication process. The filtered error signal drives the VCO frequency in such a manner that the VCO

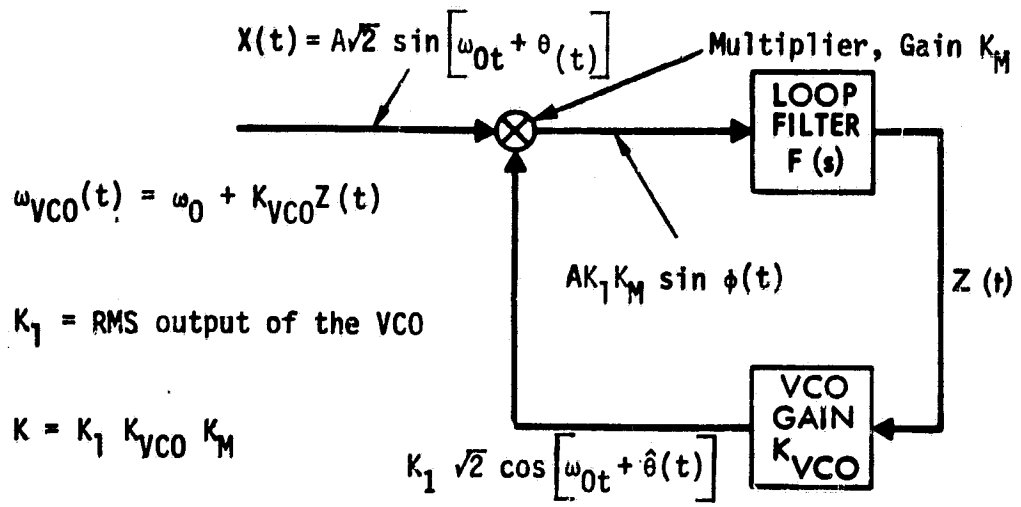


Figure 3.47. Simple Model of a Phase-Lock Loop

frequency and phase approximate the frequency and phase of the input signal. The characteristic equation which describes the dynamic behavior of the loop is

$$\theta(t) = \phi(t) + \sqrt{2P}\, K \frac{F(p)}{p} \sin \phi(t)$$

or

$$\hat{\theta}(t) = \left[\frac{\sqrt{2P}\, KF(p)}{p}\right] \sin \phi(t) \quad , \tag{3-172}$$

where

$\theta(t)$ = input signal phase
$\phi(t)$ = error signal = $\theta(t) - \hat{\theta}(t)$
$\hat{\theta}(t)$ = phase estimate generated by the VCO
P = input signal power
K = loop gain
$F(p)$ = transfer function of the loop filter
p = differential operator, d/dt

The loop order L is $L = n + 1$, where n is the number of poles in the loop filter transfer function

Typically, the input phase consists of modulation and doppler due to the radial motion of the spacecraft relative to the transmitter. Thus,

$$\theta(t) = x(t) + d(t) , \tag{3-173}$$

where $x(t)$ is the modulation and $d(t)$ is the doppler shift.

The doppler shift can be expanded in a Taylor series to obtain

$$d(t) = \phi_0 + \Omega_0 t + \frac{\Lambda_0}{2} t^2 + \cdots + \frac{x_n t^n}{n!} , \tag{3-174}$$

where

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

x_n = nth time derivative of the incoming carrier phase.

The first three terms of (3-174) are the most important for typical spacecraft trajectories.

A steady-state phase error ϕ_{ss} will occur when $d(t)$ has a finite number of Taylor series coefficients such that

$$n = \ell + 1 , \tag{3-175}$$

where ℓ is the number of poles at the origin of $F(s)$ and n is defined in (3-174). If $n > \ell + 1$, $\phi(t)$ is unbounded and the loop will eventually lose lock.

Most typical spacecraft receivers use passive second-order loop filters of the form:

$$F(s) = \frac{1 + \tau_2 s}{1 + \tau_1 s} . \tag{3-176}$$

The time constants are usually chosen with $\tau_1 \gg \tau_2$, so that $F(s)$ approaches the perfect integrator:

$$F(s) = \frac{1 + \tau_2 s}{\tau_1 s} \quad \text{for } \tau_1 s \gg 1 . \tag{3-177}$$

For the filter of (3-175), the PLL can track the doppler signal

$$d(t) = \phi_0 + \Omega_0 t \tag{3-178}$$

with a steady-state phase error [17]:

$$\phi_{ss} \triangleq \sin^{-1}\left(\frac{\dot{d}(t)}{\sqrt{2P}\,K} + \frac{\ddot{d}(t)\tau_1}{\sqrt{2P}\,K}\right) = \sin^{-1}\left(\frac{\Omega_0}{\sqrt{2P}\,K}\right), \tag{3-179}$$

given that the loop is initially in lock and that no noise is present in the loop. The phase error is commonly referred to as static phase error and Ω_0 is referred to as the doppler frequency shift.

Given that the loop is out of lock, the PLL will be able to acquire lock in a finite period of time if $d(t) = \phi_0 + \Omega_0 t$ and Ω_0 is less than a certain magnitude, called the loop pull-in range. The loop pull-in range must be determined experimentally since current models are of insufficient accuracy.

If the initial frequency offset, Ω_0, is within the loop pull-in range, the time required to achieve lock is given by [18]:

$$t_{\text{freq acq}} = \sqrt{2\pi^2 \frac{\tau_2}{\tau_1}}\left(\frac{r+1}{r}\right)\frac{\Omega_0^2}{\omega_L}, \tag{3-180}$$

where ω_L is the PLL bandwidth, $r = \sqrt{2P}\,K\tau_2^2/\tau_1$, and the signal-to-noise power ratio in the loop is greater than 10 dB. For SNR less than 10 dB, the equation consistently gives the time a smaller value.

If a small rate (Λ_0) is present in d(t), then the ϕ_{ss} will vary slowly according to

$$\phi_{ss} = \sin^{-1}\left(\frac{\Omega_0}{\sqrt{2P}\,K} + \frac{\Lambda_0 t}{\sqrt{2P}\,K} + \frac{\Lambda_0\tau_1}{\sqrt{2P}\,K}\right). \tag{3-181}$$

Eventually, the loop will reach a maximum value of ϕ for which it will lose lock. The frequency offset corresponding to this condition is called the hold-in range of the loop.

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 many receivers sweep the 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 [19]:

$$\dot{f}_{acq} = \frac{\left[1 - (\eta)^{-1/2}\right]\left(\frac{\alpha_\ell}{\alpha_{\ell 0}}\right)}{\pi \tau_2^2} , \qquad (3\text{-}182)$$

where

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

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

α_ℓ = limiter signal amplitude suppression factor

$\alpha_{\ell 0}$ = limiter signal amplitude suppression factor, at loop threshold.

According to Gardner [20], $\dot{f}_{acq}$ predicted by (3-182) should be reduced by $\sqrt{2}$ to compensate for an error in the value of loop gain used in [19].

Two additional characteristics of PLL receivers are their transfer function H(s) and two-sided loop bandwidth $\omega_L (= 2B_L)$*. The transfer function is defined by

$$\hat{\theta}(s) = H(s)\,\theta(s) , \qquad (3\text{-}183)$$

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

When the phase error ϕ is small, $\phi = \sin\phi$ and H(s) is approximately:

$$H(s) = \frac{F(s)\sqrt{2P}\,K}{s + \sqrt{2P}\,KF(s)} . \qquad (3\text{-}184)$$

---

*ω_L is the two-sided bandwidth of the PLL. B_L is the single-sided loop bandwidth. Both notations are conventionally used in the literature.

When the loop filter is of the form

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

then,

$$H(s) = \frac{1+\tau_2 s}{1+\left(\tau_2 + \frac{1}{\sqrt{2P}\,K}\right)s + \left(\frac{\tau_1}{\sqrt{2P}\,K}\right)s^2}. \qquad (3\text{-}185)$$

The loop bandwidth, ω_L, is defined to be

$$\omega_L \triangleq \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} |H(s)|^2\, ds. \qquad (3\text{-}186)$$

Using the linearized loop,

$$\omega_L \simeq \frac{r+1}{2\tau_2}, \qquad (3\text{-}187)$$

where the assumption $r\tau_1 >> \tau_2$ has been used.

Receivers used in most spacecraft are of the double-conversion superheterodyne type preceded by a bandpass limiter, as shown in Figure 3.48. 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 is required to provide a signal whose amplitude is within the dynamic range of the amplifier stages. The bandpass limiter minimizes the total mean-square error of the loop over a wide range of input SNR's. This configuration has been shown in [21] to provide near-optimum PLL performance.

The noise-free characteristics described in the preceding paragraphs apply to these receivers, with the exception that the amplitude gain product, $\sqrt{2P}\,K$, is computed using the relation:

$$\sqrt{2P}\,K = 360\,\alpha_\ell\, K_D\, K_{VCO}\, M\, K_{DC} \qquad \left[\text{sec}^{-1}\right], \qquad (3\text{-}188)$$

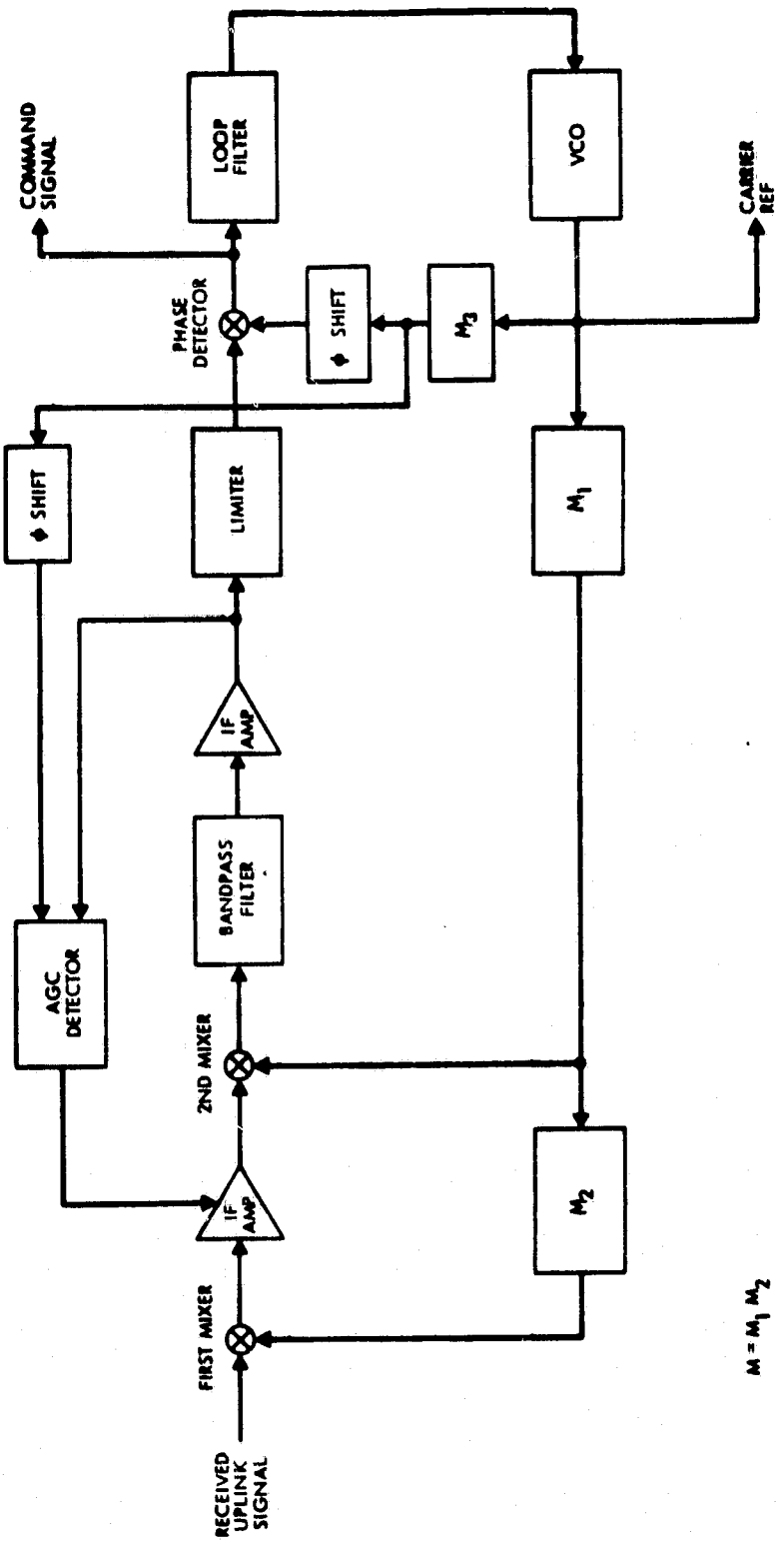


Figure 3.48. Typical Spacecraft Receiver

where

α_ℓ = limiter suppression factor (to be discussed)

K_D = phase detector gain (volts/deg)

K_{VCO} = VCO gain (Hz/volt)

M = VCO multiplication factor

K_{DC} = gain of the loop filter

In its first stage, the spacecraft receiver will generate white Gaussian noise which is added to the received signal. This noise prevents the receiver from making a perfect phase reference estimate even in the absence of doppler shift or phase modulation.

The probability density function of the PLL phase error for a first-order loop when the signal is corrupted by white Gaussian noise was shown by Viterbi [22] to be:

$$p(\phi) = \frac{\exp(\eta \cos \phi)}{2\pi I_0(\eta)}, \quad |\phi| \leq \pi, \tag{3-189}$$

where

$p(\phi)$ = probability density function of the loop phase error at the phase detector output

$\eta = \frac{2P}{N_0 L}$ = SNR in the bandwidth of the linearized loop, i.e., SNR in B_L

ω_L = PLL two-sided loop bandwidth

$L_0(\ )$ = 0th order imaginary Bessel function

N_0 = input noise spectral density, W/Hz.

Bandpass limiters (a bandpass filter followed by an amplitude hard limiter) are used in phase-locked-loop 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 SNR.

The bandpass limiter may be characterized by the filter bandwidth, ω_ℓ. If η_ℓ is the SNR in the filter bandwidth, then the power in the signal component of the limiter output spectrum, α_ℓ^2, is determined from

$$\alpha_\ell = \sqrt{\frac{\pi}{2}} \left(\frac{\eta_\ell}{2}\right)^{1/2} \exp\left(-\frac{\eta_\ell}{2}\right) \left[I_0\left(\frac{\eta_\ell}{2}\right) + I_1\left(\frac{\eta_\ell}{2}\right)\right], \tag{3-190}$$

where

$$\eta_\ell = \frac{2P}{N_0\omega_\ell} = 2\eta_0\left(\frac{\omega_{LO}}{\omega_\ell}\right) = \eta_0\frac{2B_{LO}}{B_\ell} \tag{3-191}$$

and η_0, defined as $P/(N_0\omega_{LO})$, or SNR in the threshold loop bandwidth $(2B_{LO})$, is not the SNR in $2B_L$ at threshold in the bandpass limiter case. The quantity α_ℓ is commonly referred to as the limiter signal amplitude suppression.

Equation (3-190) may be approximated by the rational polynomial [17]:

$$\alpha_\ell^2 = \frac{0.7854\,\eta_\ell + 0.4768\,\eta_\ell^2}{1 + 1.024\,\eta_\ell + 0.4768\,\eta_\ell^2}\ , \tag{3-192}$$

which is plotted in Figure 3.49. For the linear model of a PLL, the variance of the phase error is given by

$$\sigma_\phi^2 = \frac{N_0\omega_L}{2P}\ . \tag{3-193}$$

For the linear model of a PLL preceded by a bandpass limiter,

$$\sigma_\phi^2 = \left(\frac{N_0\omega_L}{2P}\right)\Gamma = \frac{1}{\eta}\ , \tag{3-194}$$

where Γ is the limiter performance factor or limiter loss [17], and η is the SNR in B_L. Γ can be approximated by [23]:

$$\Gamma \simeq \frac{1+\eta_\ell}{0.862+\eta_\ell}\ , \quad \text{for } \omega_\ell > 10\,\omega_L \tag{3-195}$$

(refer to Figure 3.50). The loop bandwidth is now

$$\omega_L = \frac{1+r}{2\tau_2\left(1+\dfrac{\tau_2}{r\tau_1}\right)}\ , \tag{3-196}$$

where r becomes

$$r = \frac{360\,\alpha_\ell\,K_D\,K_{VCO}\,M\,K_{DC}\,\tau_2^2}{\tau_1}\ . \tag{3-197}$$

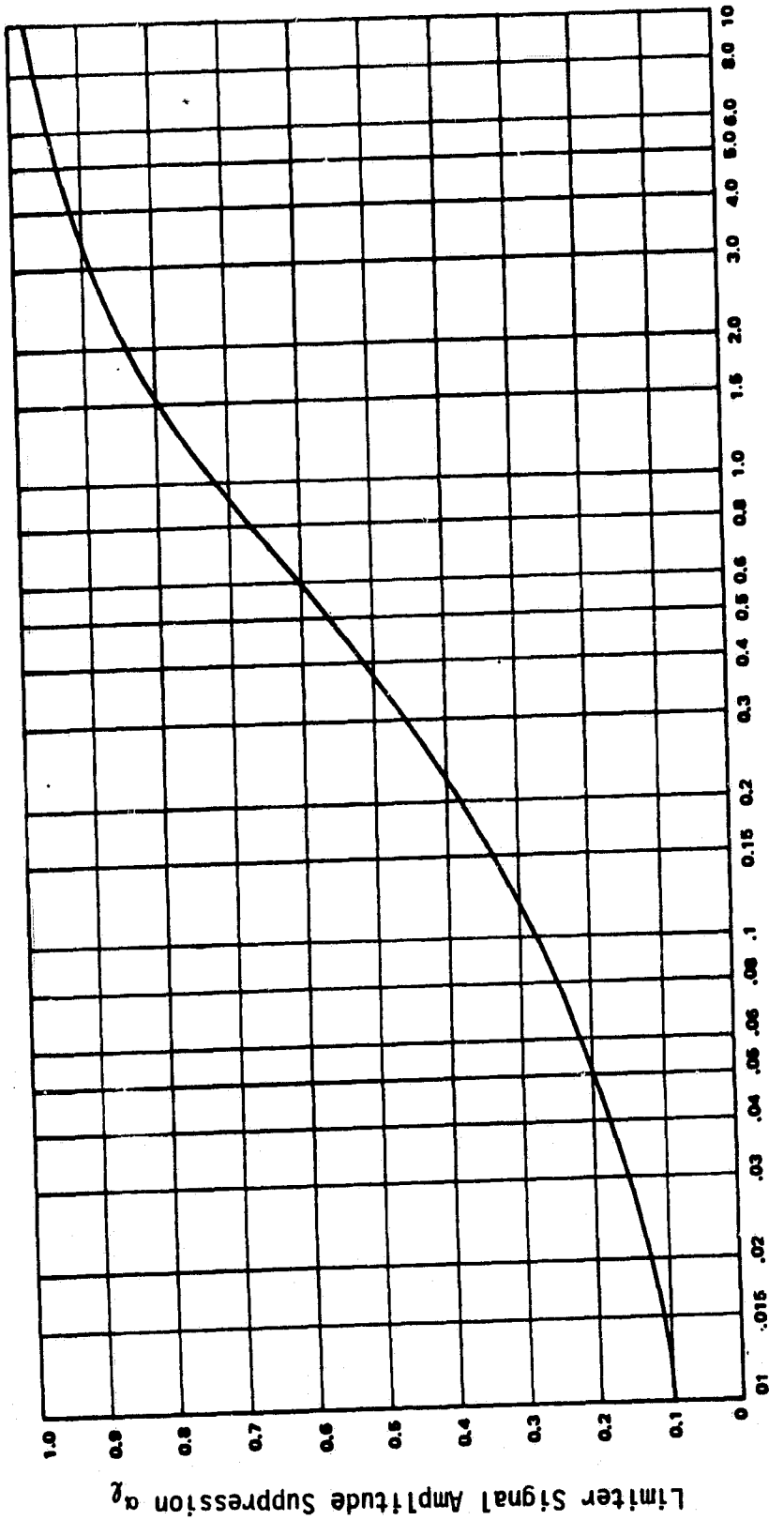


Figure 3.49. Variation of Limiter Signal Amplitude Suppression with Limiter Input SNR

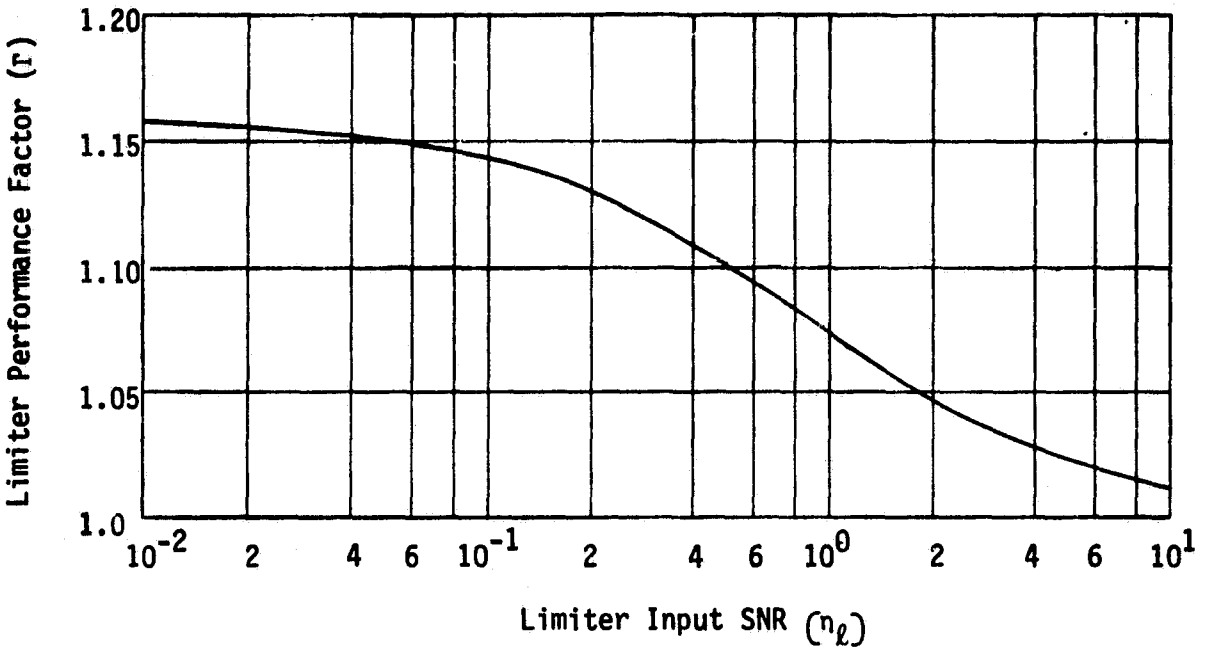


Figure 3.50. Variation of the Limiter Performance Factor With Limiter Input SNR

It is useful in the design, specification and performance analysis of PLL receivers to define a receiver threshold or design-point condition. By convention, the design point condition occurs when the variance of the loop phase error is unity and the variance is represented by a linearized PLL. In actual fact, the variance of a linearized loop is different from that of an actual nonlinear loop when its variance is 1. The fiction that the loop is linear in the definition of threshold is used so that the mathematically tractable expression:

$$P_0 \triangleq N_0 \omega_{LO} \tag{3-198}$$

at threshold* can be used to <u>define</u> the threshold condition.

In practice, PLL receivers are designed by specifying the threshold loop bandwidth ω_{LO} and the receiver noise spectral density N_0. The receiver sensitivity P_0 is then determined by definition from (3-198). The limiter suppression factor at threshold may be calculated by substituting $\eta_{\ell 0} = 2B_{LO}/B_\ell$ into (3-190) or (3-192) to obtain $\alpha_{\ell 0}$.

Now the PLL receiver characteristics may be determined as a function of a signal level referenced to the threshold loop bandwidth. For instance, the actual loop bandwidth is

$$\omega_L = \omega_{LO}\left(\frac{1+\frac{r_0}{\mu}}{1+r_0}\right)\left(\frac{1+\frac{\tau_2}{r_0\tau_1}}{1+\frac{\mu\tau_2}{r_0\tau_1}}\right), \tag{3-199}$$

where

$$\mu = \frac{\alpha_{\ell 0}}{\alpha_\ell}$$

$$r_0 = \frac{360\,\alpha_{\ell 0} K_D K_{VCO} M K_{DC} \tau_2^{\,2}}{\tau_1} = \mu r\,. \tag{3-200}$$

---

*By convention, the subscript $(_0)$ represents a threshold condition.

Typically, $r_0\tau_1 >> \tau_2$, so that

$$\omega_L \approx \omega_{LO}\left(\frac{1+\frac{r_0}{\mu}}{1+r_0}\right) . \tag{3-201}$$

With no limiter in the system, $B_L/B_{LO} = 1$. However, with a bandpass limiter, B_L/B_{LO} is a function of η_0, r_0, and $2B_{LO}/B_\ell$. For $2B_{LO}/B_\ell$ very small (<0.002), there is no significant dependence on this parameter. For this case and $r_0\tau_1 >> \tau_2$, Figure 3.51 plots B_L/B_{LO} versus η_0, Figure 3.52 plots B_L/B_{LO} versus η, and Figure 3.53 plots η versus η_0 for various values of r_0. Note that η is the SNR in B_L, whereas η_0 is defined as the SNR in $2B_{LO}$, and that B_L/B_{LO} evaluated at $\eta = 3$ dB is not equal to 1. This is because η_0 is a definition and does not include the limiter performance factor Γ; thus, the SNR in $2B_{LO}$ is not the actual SNR in the phase-locked-loop bandwidth. These curves were derived through equations (3-191), (3-192), (3-194), (3-195), and (3-201), and the definition

$$\eta = \eta_0\left(\frac{2B_{LO}}{B_L}\right)\frac{0.862 + N_0\frac{2B_{LO}}{B_\ell}}{1+\eta_0\frac{2B_{LO}}{B_\ell}} . \tag{3-202}$$

When the signal level at the input to a PLL preceded by a bandpass limiter is so low that the loop is in its nonlinear region of operation, phase error variance σ^2 is [24]:

$$\sigma^2 = \frac{\pi^2}{3} + \frac{4}{I_0(\eta)}\sum_{k=1}^{\infty}\frac{(-1)^k I_k(\eta)}{k^2} , \tag{3-203}$$

where

$$\eta = \left(\frac{2P}{N_0\omega_{LO}}\right)\left(\frac{1}{\Gamma}\right)\left(\frac{1+r_0}{1+\frac{r_0}{\mu}}\right) . \tag{3-204}$$

Above a value of η of 7 or 8 dB, the linear theory results of (3-194) hold quite accurately.

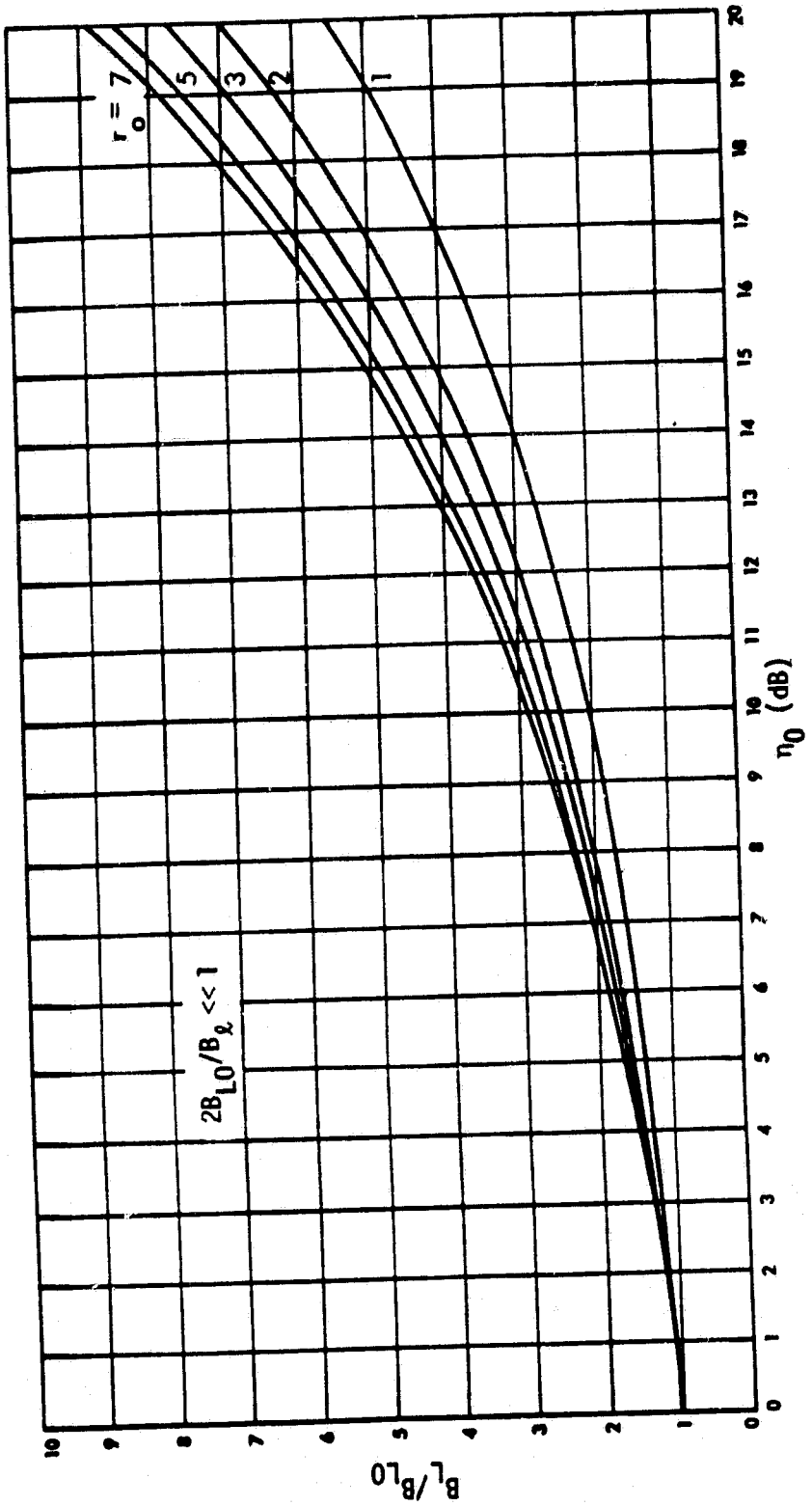


Figure 3.51. B_L/B_{LO} Versus SNR in $2B_{LO}$ with r_0 as a Parameter

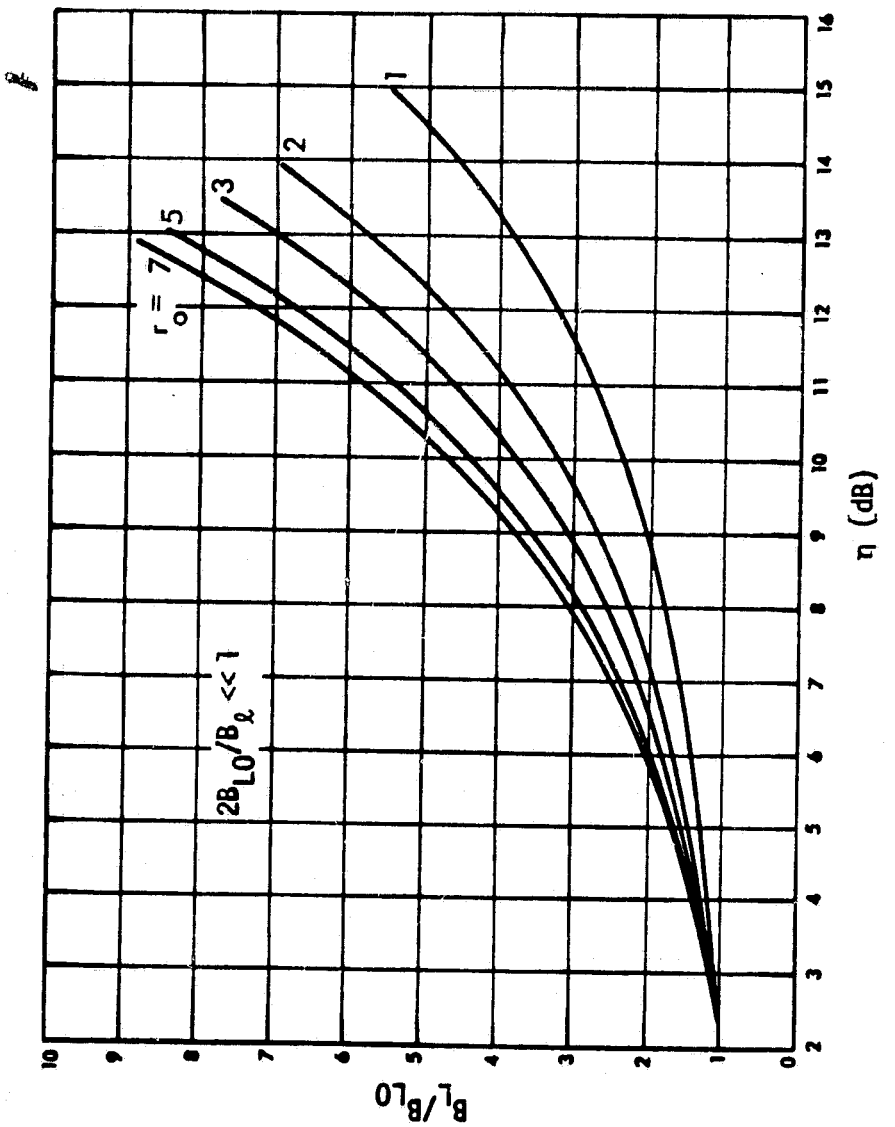


Figure 3.52. B_L/B_{LO} Versus SNR in B_L with r_0 as a Parameter

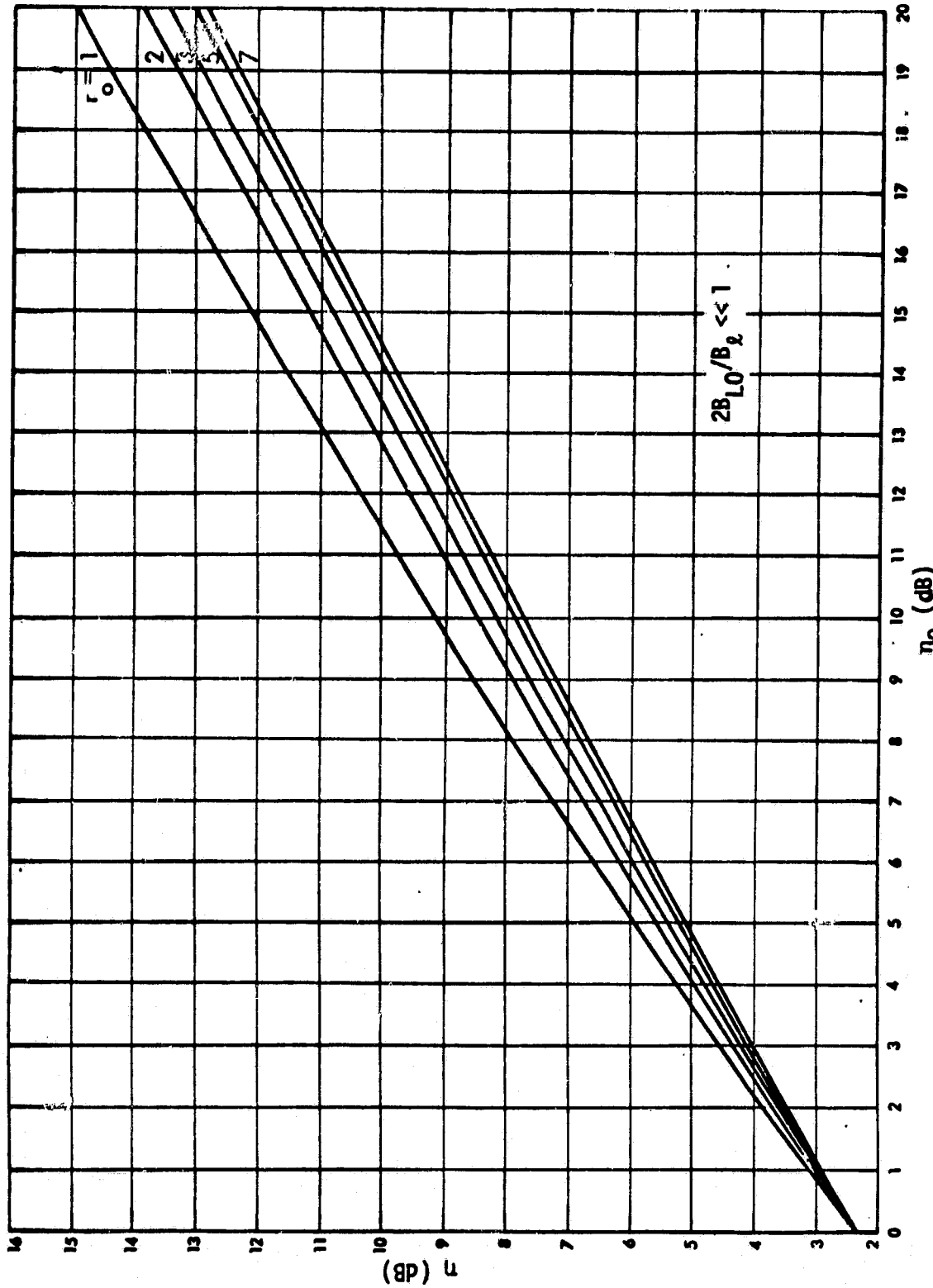


Figure 3.53. SNR in B_L Versus SNR in $2B_{LO}$ with r_0 as a Parameter

In order to accurately demodulate the data, the SNR in the loop η must be large enough that the probability of slipping a cycle is very low. The loop will lose lock momentarily if the instantaneous phase angle, ϕ, exceeds the range $\pm\pi/2$ radians. The probability of this occurrence can be chosen to lie at some multiple of σ. Then (3-193) gives the corresponding η, and use of (3-199) and (3-194) with the proper parameters gives the minimum SNR in $2B_{LO}$ to provide for low probability of loss of lock. Figure 3.53 also plots the relation between η and η_0 for $2B_{LO}/B_{\ell} << 1$.

Empirical results which include limiter suppression and system efficiency suggest a minimum recommended operating SNR in $2B_{LO}$ of 10 dB.

3.2.10.2 Suppressed carrier-tracking loops

Suppressed carrier modulation is an efficient technique for transmission of digital data over space and terrestial links. Because of carrier suppression, demodulators which use information contained in data sidebands must be employed to provide for coherent carrier tracking and data recovery. Such demodulators consist of either a squaring loop, a Costas loop, or a data-aided loop.

For the purpose of the following discussion, the received signal that is input to the receiver will have the form

$$x(t) = \sqrt{2P}\, d(t) \cos \omega_c t + n_i(t) , \qquad (3\text{-}205)$$

where the input noise $n_i(t)$ is bandpass with one-sided noise spectral density N_0 [W/Hz].

The basic squaring loop is shown in Figure 3.54. The signal $x(t)$ is bandpass-filtered and squared (multiplied by itself) to produce $z(t)$, which has a discrete frequency component at $2\omega_c$. The loop locks to this component, and the output of the VCO is divided in frequency by 2 and phase-shifted 90° to provide the reference $r_2(t)$ which coherently demodulates the carrier. The output of the data detector is the maximum-likelihood estimate $\hat{d}(t)$ of the transmitted symbols.

The performance of the squaring loop, as measured by the variance of the phase error is related to the PLL by

$$\sigma_{2\phi}^2 = \frac{4}{\eta L_{SQ}} , \qquad (3\text{-}206)$$

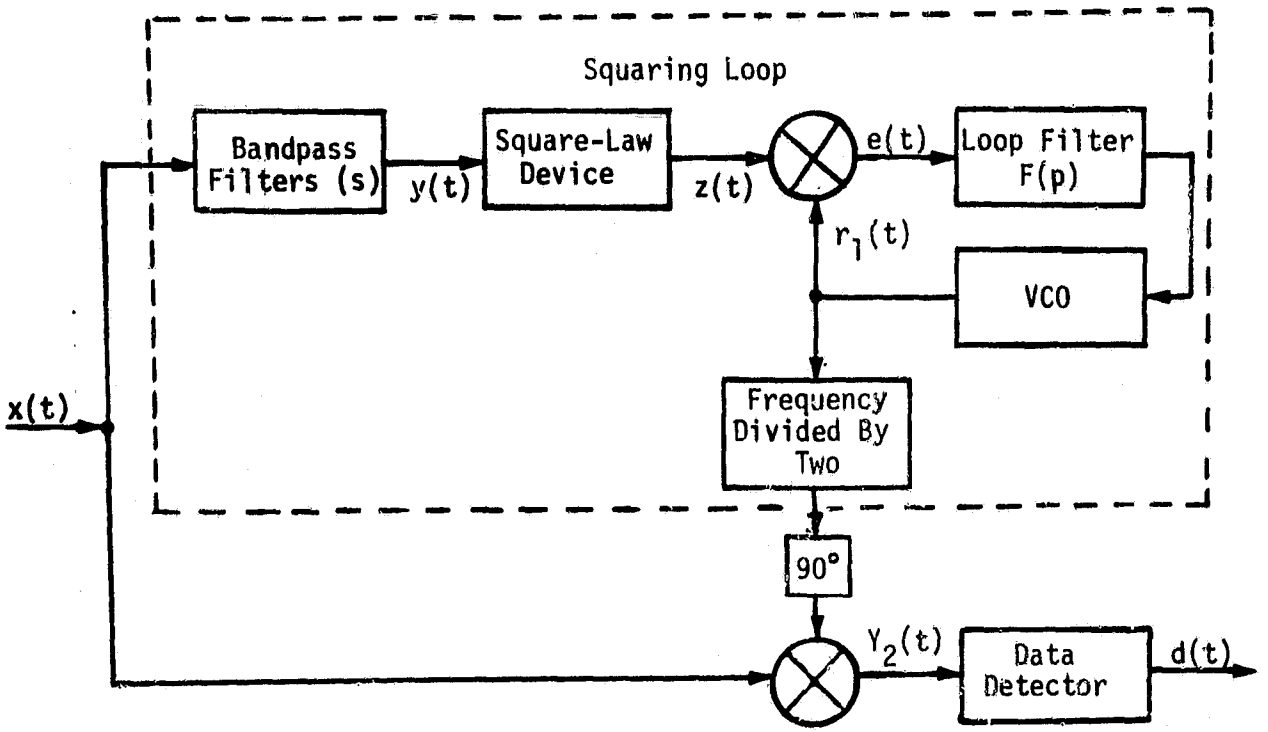


Figure 3.54. Squaring-Loop Receiver

where the variance is measured as twice the phase error ϕ (i.e., 2ϕ) since the PLL in the squaring loop is tracking $2\omega_c$ and η is the SNR in B_L. The squaring loss L_{SQ} is largely a function of the SNR into the square-law device. This loss L_{SQ} is given by [25]:

$$L_{SQ} = \frac{D_m}{K_s + K_L \frac{B_i/R_s}{2 D_m R_d}} \tag{3-207}$$

where

D_m = modulation distortion factor due to the input bandpass filter

B_i = two-sided noise bandwidth of the lowpass equivalent of the input bandpass filter

K_s = constant depending on the data modulation spectrum and the filter type

K_L = constant depending on the filter type

R_s = data rate (symbols/second)

$R_d = E_s/N_0$.

The two-sided noise bandwidth B_i corresponding to the input bandpass filter is defined as

$$B_i = \int_{-\infty}^{\infty} |H_\ell(j2\pi f)|^2 \, df \, , \tag{3-208}$$

where $H_\ell(s)$ is the lowpass equivalent of the bandpass filter transfer function H(s). For an n-pole Butterworth filter,

$$|H_\ell(j2\pi f)|^2 = \frac{1}{1 + (f/f_c)^{2n}} \, . \tag{3-209}$$

Evaluating (3-208) using (3-209),

$$B_i = \frac{f_c}{\frac{n}{\pi} \sin \frac{\pi}{2n}} \, . \tag{3-210}$$

The modulation distortion factor due to the input bandpass filter is defined as

$$D_m = \int_{-\infty}^{\infty} S(f) \, |H_\ell(j2\pi f)|^2 \, df \, , \tag{3-211}$$

with S(f) denoting the power spectral density (PSD) of the data modulation. For NRZ coding of equiprobable independent transmitted symbols, the PSD is

$$S(f) = \frac{1}{R_s} \frac{\sin^2 (\pi f/2R_s)}{(\pi f/2R_s)^2} \, . \tag{3-212}$$

Alternately, for Manchester-coded data, the PSD is

$$S(f) = \frac{1}{R_s} \frac{\sin^4(\pi f/2R_s)}{(\pi f/2R_s)^2} . \qquad (3\text{-}213)$$

Figure 3.55 presents the results of numerically integrating (3-211) for Manchester-coded data and various filters.

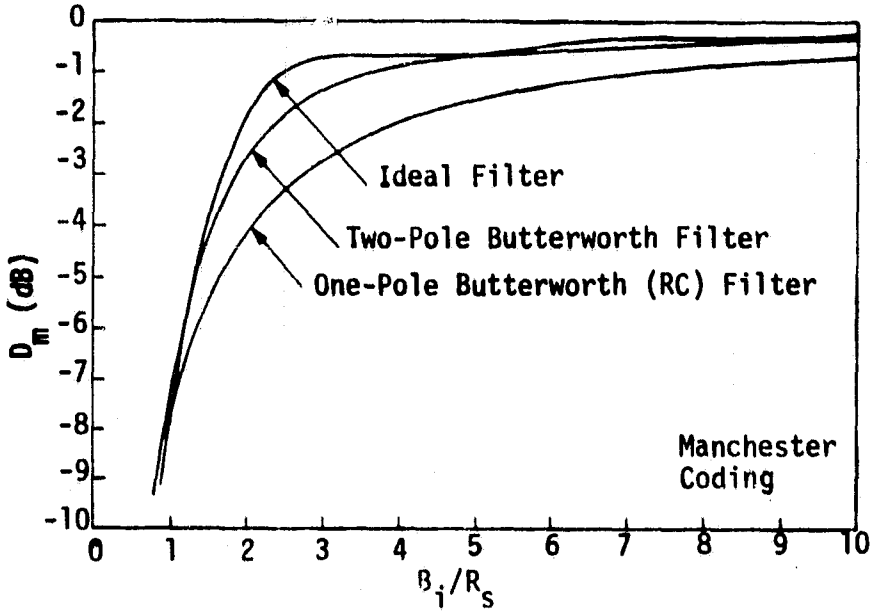


Figure 3.55. Modulation Distortion Factor Versus B_i/R_s for Various Input Bandpass Filters

The two constants K_s and K_L in (3-207) are defined as

$$K_s = \frac{\int_{-\infty}^{\infty} S(f)\,|H_\ell(j2\pi f)|^4\,df}{\int_{-\infty}^{\infty} S(f)\,|H_\ell(j2\pi f)|^2\,df} \qquad (3\text{-}214)$$

and

$$K_L = \frac{\int_{-\infty}^{\infty} |H_\ell(j2\pi f)|^4 \, df}{\int_{-\infty}^{\infty} |H_\ell(j2\pi f)|^2 \, df} , \qquad (3\text{-}215)$$

where K_s is dependent on both the baseband data power spectrum density and the filter type and, like D_m, must be numerically integrated. Note that K_L is dependent only on the filter type, with $K_L = 1$ for an ideal lowpass filter and $K_L = (2n-1)/2n$ for an n-pole Butterworth filter.

Using the results for D_m from Figure 3.55 in (3-207), Figures 3.56 through 3.58 illustrate the squaring loss L_{SQ} versus B_i/R_s with R_d as a parameter for one-pole, two-pole, and infinite-pole Butterworth filters. In each case, it is observed that, for a fixed R_d, there exists an optimum noise bandwidth B_i for the input bandpass filters in the sense of minimizing the squaring loss. These values of optimum input bandpass filter bandwidth occur in the vicinity of the Nyquist bandwidth. The optimum bandwidth (which minimizes the effects of the squaring loss on loop-tracking performance) choice for the input bandpass filters as a function of R_d is illustrated in Figure 3.59. These results indicate that, as the number of filter poles increases, the sensitivity of the optimum input bandpass filter bandwidth with SNR ratio diminishes.

Although the numerical results given in Figures 3.55 through 3.59 pertain to the case of Manchester-coded data, similar results for NRZ data (or any other data format) can be easily obtained from (3-207). In fact, the only terms in this equation which are data-format dependent are D_m and K_s which, for NRZ data, can be calculated using (3-212) rather than (3-213) in (3-211) and (3-214). Figure 3.60 plots D_m in dB versus the ratio B_i/R_s for single-pole, two-pole, and infinite-pole Butterworth filters, the latter being equivalent to an ideal lowpass filter. For a single-pole filter, D_m can be evaluated in closed form with the result [26]:

$$D_m = 1 - \frac{1}{2B_i/R_s}\left[1 - \exp\left(-2B_i/R_s\right)\right] . \qquad (3\text{-}216)$$

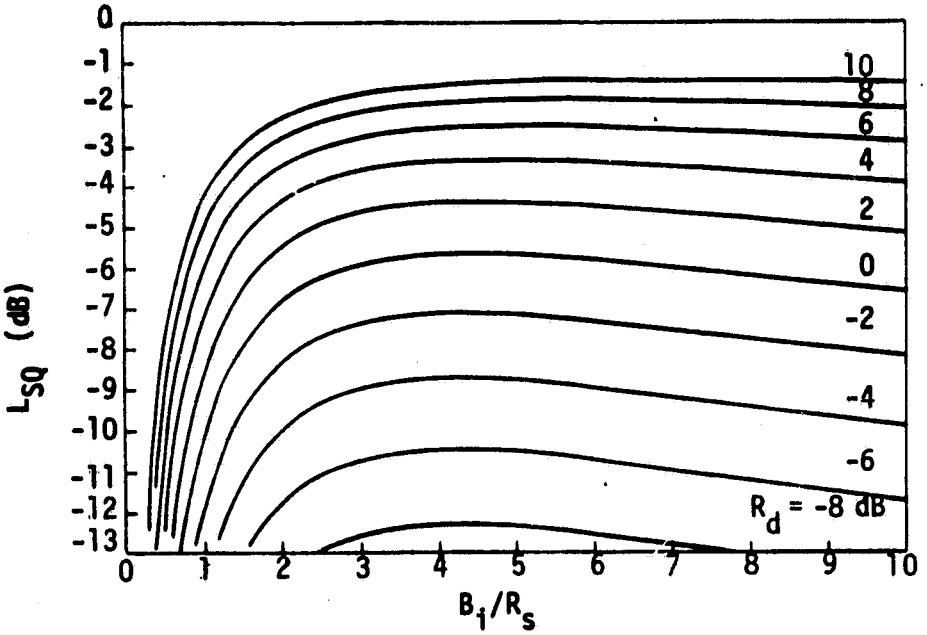


Figure 3.56. Squaring Loss Variations Versus B_i/R_s for Various Values of R_d; RC Filter, Manchester Coding

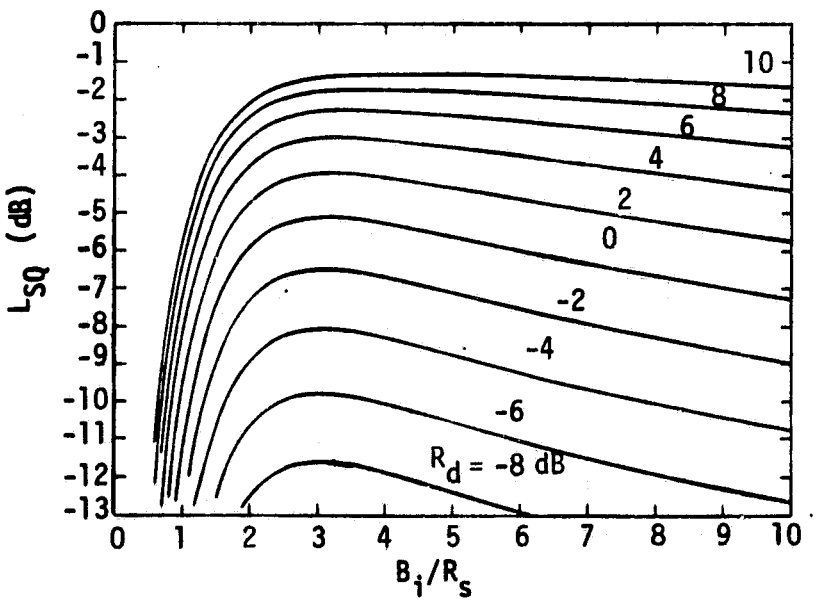


Figure 3.57. Squaring Loss Variations Versus B_i/R_s for Various Values of R_d; Two-Pole Butterworth Filter, Manchester Coding

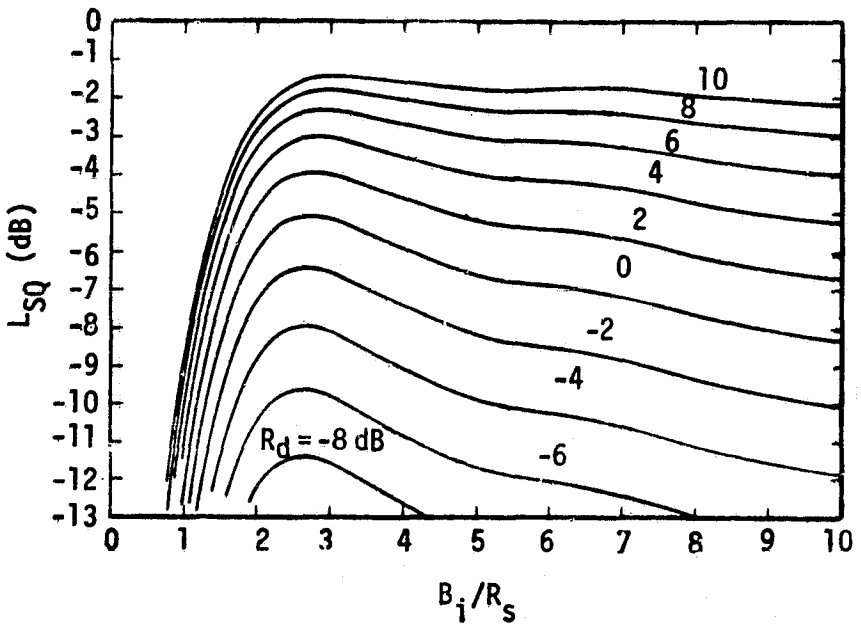


Figure 3.58. Squaring-Loss Variations versus B_i/R_s for Various Values of R_d; Ideal Filter, Manchester Coding

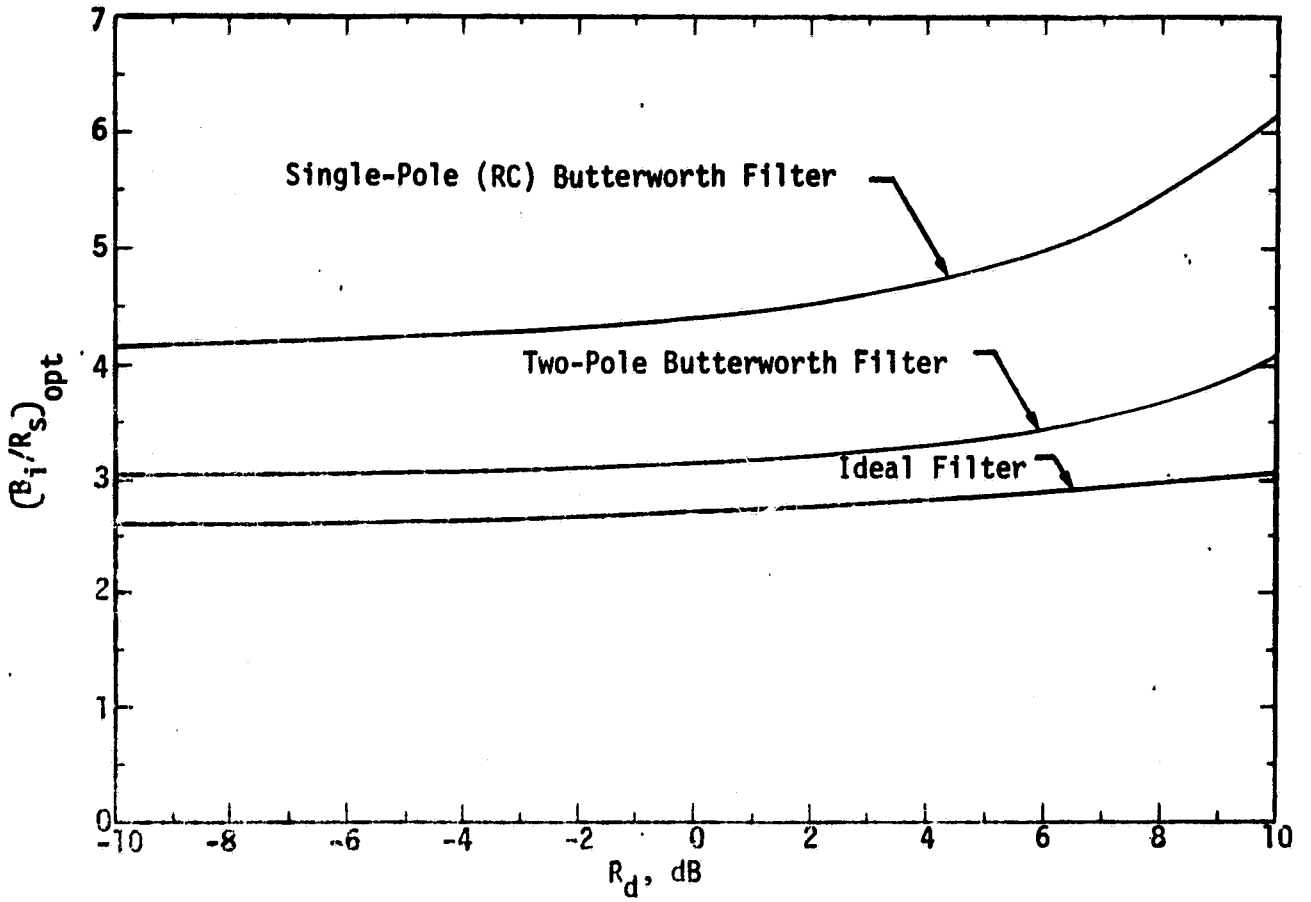


Figure 3.59. Ratio of Optimum Input Bandpass Filter Bandwidth to Symbol Rate Versus Symbol Energy-to-Noise Ratio

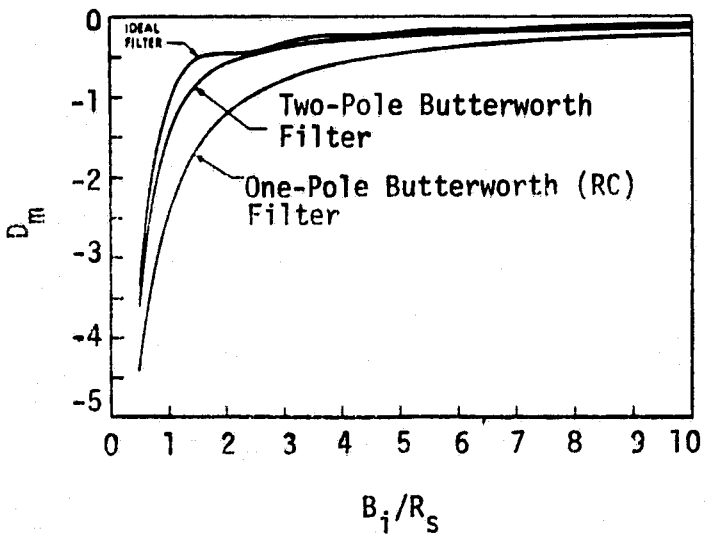


Figure 3.60. Modulation Distortion Factor Versus B_i/R_s for Various Input Bandpass Filter Types

The constant K_s can also be evaluated in closed form for a single-pole filter as:

$$K_s = \frac{1 - \dfrac{3 - (3 + 2B_i/R_s)\exp(-2B_i/R_s)}{4B_i/R_s}}{1 - \dfrac{1}{2B_i/R_s}\left[1 - \exp(-2B_i/R_s)\right]} \quad . \tag{3-217}$$

Also, for an ideal filter, $K_s = 1$. Further NRZ numerical results are included as a special case of a more general problem concerning bandpass limiters preceding the suppressed carrier-tracking loops discussed in subsequent paragraphs.

The Costas loop, shown in Figure 3.61, does not require that the input BPF have a specified bandwidth; rather, its performance depends upon the specification of the two lowpass filters, G(s). Actually, the Costas loop is nothing more than a lowpass version of the squaring loop and has the identical stochastic differential equation as that of the squaring loop, provided the noise spectrum formed by the filters for each configuration is the same. Thus, the performance of the Costas loop and the squaring loop have the identical performance if G(s) is substituted for $H_\ell(s)$ in (3-208), (3-209), (3-211), (3-214), and (3-215). Figures 3.55 through 3.59 present the results for the Costas loop if B_i refers to the arm filter bandwidth rather than the input bandpass filter bandwidth.

It should be noted that, for both the squaring and Costas loops, no attempt is made to optimize the receiver performance with respect to the modulating signal d(t). That is, the nature of d(t), other than its power spectrum, plays no real part in the formation of the loop error signal. An optimum receiver would, however, make proper use of the data estimate $\hat{d}(t)$. A receiver configuration which accomplishes this is the data-aided loop (DAL); the basic configuration is shown in Figure 3.62 for Manchester-coded data. For NRZ data, the integration limit on the integrate-and-dump circuits would be 0 to T_s rather than the divided integration limits required for Manchester-coded data. The topology of the DAL is very similar to that of the Costas loop; however, there are major differences. First, the data detector is now an integral part of the loop. Secondly, the lowpass filters

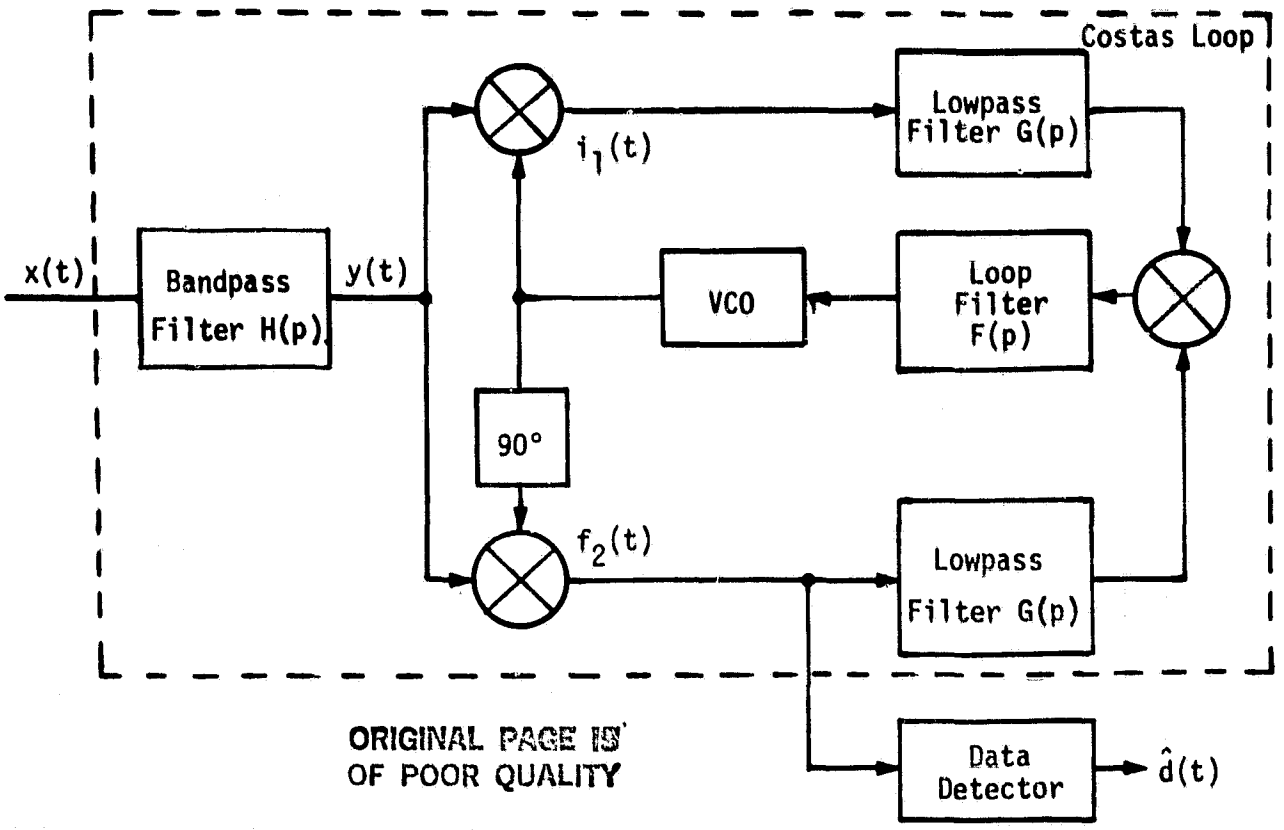


Figure 3.61. Costas Loop Receiver

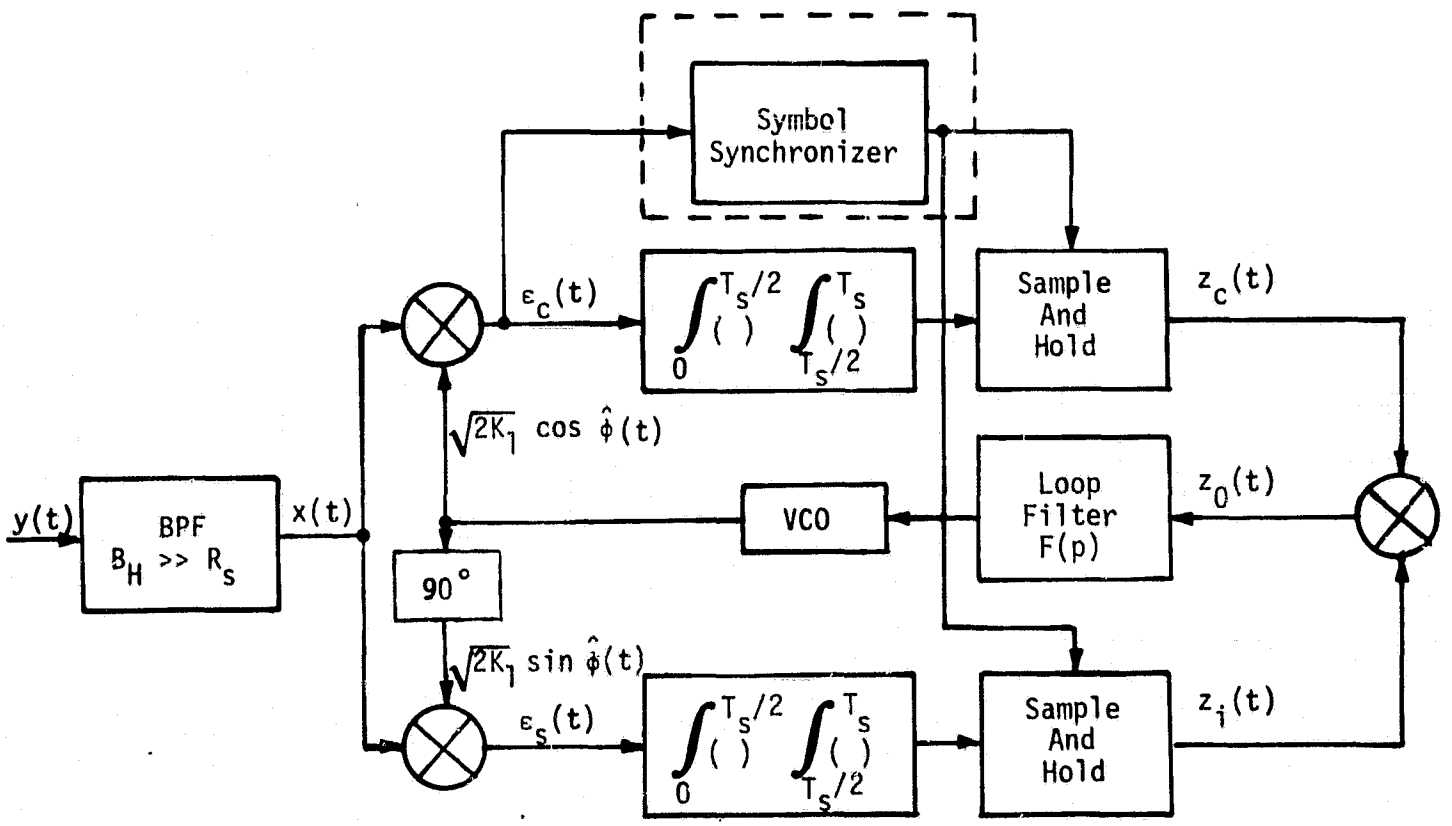


Figure 3.62. Costas Loop With Active Arm Filters

G(s) are not required, and the input BPF serves only to limit the total input noise voltage.

Operation of the Costas loop with matched filters in the arms required accurate estimates of the instants in time at which the symbols may change states; however, symbol synchronization acquisition with a Costas-type loop that requires symbol synchronization would seem to involve one of those unstable situations in which the carrier cannot be acquired unless symbol synchronization has been obtained and vice versa.

Fortunately, this is not the case as, under certain conditions, it is possible to obtain a good estimate of symbol synchronization even though the carrier loop is not locked. This is qualitatively explained in what follows.

The effect of the carrier loop being out of lock is to amplitude modulate (100%) the data symbol stream by the quadrature error signal $\cos \phi(t)$, $\phi(t)$ being the instantaneous phase error. If the symbol rate is less than the beat-note frequency out of the coherent amplitude detector, then the symbol synchronization loop sees $\pm\cos \phi(t)$ and the symbol transitions which carry the clock information are not affected by carrier lock. The fact that amplitude modulation exists at the beat-note frequency can be accounted for in the design of the symbol synchronizer by considering the average SNR reduction which it would see during symbol synchronization acquisition. This is easily argued (worst-case) to be 3 dB. Thus, for a given frequency uncertainty region assumed to be less than the symbol rate, the symbol synchronizer can be designed such that, as the carrier VCO is swept through the zero-beat frequency region, it acquired rapidly, thus giving the carrier loop a coarse estimate of the clock. Assuming that this takes place quite rapidly, the carrier VCO will acquire the swept signal as it passes through the zero-beat region. The lock detector recognizes this state and correspondingly kills the sweep. This coupled synchronization process does affect the receiver/bit synchronizer interface, which should create no major problems in practice.

The equivalent squaring-loss for a Costas loop which uses integrate-and-dump circuits as arm filters is given by the simple expression [25]:

$$L_{SQ} = \frac{1}{1 + \frac{1}{2R_d}} \tag{3-218}$$

for either Manchester-coded data or NRZ coding.

At large values of $R_d = E_s/N_0$, better performance can be obtained from a DAL with a hard-limiter in the upper arm of the Costas loop in Figure 3.61. In this case, the squaring-loss is given by [27]:

$$L_{SQ} = (1 - 2P_s)^2 , \tag{3-219}$$

where P_s is the average symbol probability of error. For P_s less than 10^{-2}, L_{SQ} is less than -0.2 dB.

Figure 3.63 illustrates the comparison of the squaring losses which give the minimum phase jitter achievable with passive arm filters to that implemented with active (integrate-and-dump sample-and-hold) arm filters. In addition, the squaring-loss performance is compared with that of a CW phase-locked-loop operating with the same S/N_0 and no data modulation. Typically, the active matched filter (integrate-and-dump, sample-and-hold) gives approximately a 4 dB or greater advantage over the passive single-pole circuit for $R_d \leq 0$ dB. For example, at $R_d = -4$ dB, advantage is 5.16 dB. When compared to an ideal filter, the improvement is typically only 4.4 dB at $R_d = -4$ dB. This improvement directly translates into an effective improvement in S/N_0. Also notice that, for $R_d \geq 6$ dB, the Costas loop with active matched filters in the arms gives approximately the same performance (0.5 dB inferior) to that of a CW PLL with no modulation present. With RC filters in the arms, it is 1.4 dB inferior at $R_d = 10$ dB, and 2.48 dB inferior at $R_d = 6$ dB. On the other hand, for two-pole Butterworth filters and ideal lowpass filters, the performance is, respectively, 2.33 dB and 2.31 dB inferior to the CW loop for $R_d = 6$ dB. In conclusion, as the SNR increases, the value of the optimum squaring loss becomes less sensitive to the order of the filter.

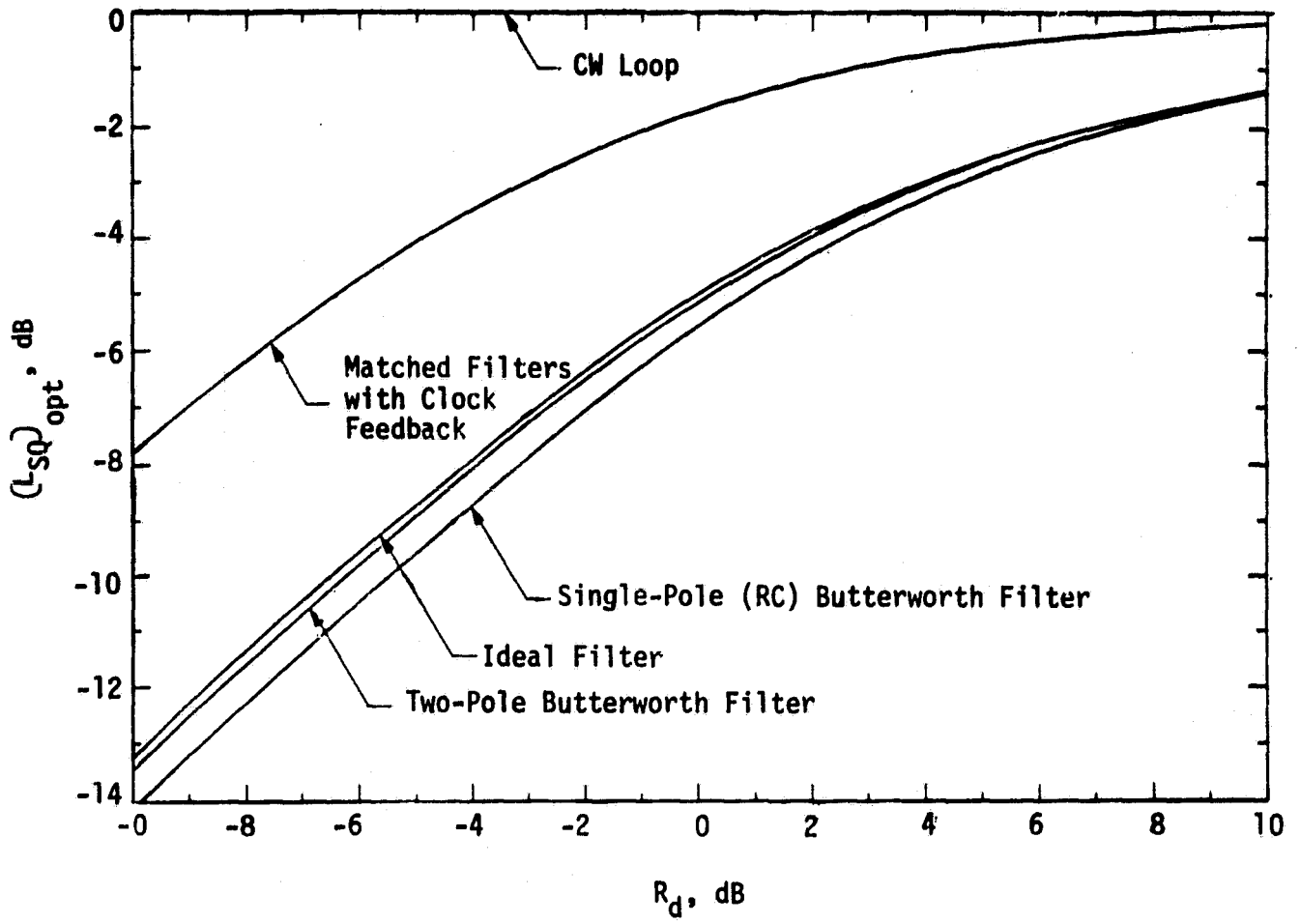


Figure 3.63. Comparison of the Effectiveness of Various Arm Filters in Reducing the Squaring Loss for Various Values of R_d

Bandpass limiters (BPL's) are used in suppressed-carrier receivers for the same reasons they are used with the phase-locked-loop receivers described in subsection 3.2.10.1. Ideally, one might like to implement a soft-limiter characteristic, which serves to control loop gain during synchronization acquisition and where the signal level is larger than the tracking loop can tolerate. Also, the bandpass limiter serves the additional function of protecting various loop components (e.g., the loop phase detector, where signal and noise levels may vary over several orders of magnitude, possibly exceeding the dynamic range of its constituent components). Once the coherent AGC loop sets the signal level to the desired design point value, i.e., after phase lock, the soft-limiter characteristic would manifest itself in the receiver as a linear amplifier. In this case, the limiter degradation-to-loop performance would be effective only during synchronization acquisition.

In what follows, the tracking performance of a Costas loop is discussed for the case where the bandpass limiter is characterized by a soft limiter which has the transfer characteristic [27]:

$$y(x) = L \operatorname{erf}\left(\frac{K\sqrt{\pi}}{2L} x\right), \tag{3-220}$$

where L is the peak limiter output, K is the slope of the transfer characteristic at $x = 0$, and

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz . \tag{3-221}$$

Obviously, if K approaches infinity, then

$$\lim_{K\to\infty} y(x) = L \operatorname{sgn}[x] \tag{3-222}$$

In other words, the soft limiter approaches a hard limiter. Furthermore, as K approaches zero, then

$$\lim_{K\to 0} y(x) = Kx , \tag{3-223}$$

which is the result for a linear amplifier of gain K (i.e, no limiter). Thus, the parameter K represents a measure of the "softness" of the limiter, as well as a design parameter selectable by the design engineer. Furthermore, by applying the above limiting arguments on K to all the results which follow, the hard-limiter and no-limiter applications fall out as special cases of the general theory.

Consider the block diagram of the BPL illustrated in Figure 3.64. The first block is a bandpass filter (usually the IF filter) with center radian frequency ω_0 and equivalent single-sided noise bandwidth $B_H << \omega_0/2\pi$ defined by

$$B_H \triangleq \frac{1}{2\pi}\int_0^\infty \frac{|H(j\omega)|^2}{|H(j\omega_0)|^2}\, d\omega \,, \tag{3-224}$$

where $H(j\omega)$ is the transfer function of the filter. The second block in Figure 3.64 is a soft limiter with input/output transfer characteristic given in (3-220). Finally, the third block of the BPL is a zonal filter which passes only the first harmonic (the term at ω_0) of the limiter output. This first zone output, denoted by $y_1(t)$, serves as the input to the Costas loop. Herein it is assumed that the input IF filter H(s) completely passes the input signal modulation but filters the input noise, thereby setting the input SNR.

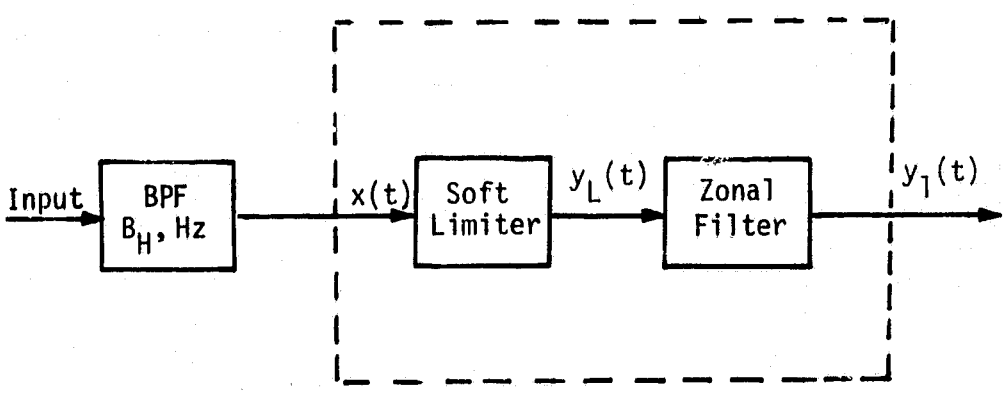


Figure 3.64. Bandpass Limiter Model

The $\tilde{\alpha}_1$ is the limiter amplitude suppression factor defined in (3-190). However, for a soft limiter,

$$\eta_\ell \triangleq \frac{\rho_H}{1+D}, \tag{3-225}$$

where $\rho_H \triangleq P/N_0B_H$ is the SNR input to the limiter and D is a "softness" parameter defined by

$$D \triangleq \frac{4\rho_H}{\pi}\left(\frac{1}{\sqrt{2}\,K\sqrt{P}}\right)^2. \tag{3-226}$$

Note that, as $K\to\infty$, D goes to 0, η_ℓ goes to ρ_H, and the suppression factor is for a hard limiter. The parameter $P_1 = 8/\pi^2$ represents the fraction of the signal-plus-noise power that falls in the first zone output.

The equivalent limiter suppression factor α_s for a Costas loop depends on the noise variances at the phase detector, as well as the filtered data. It is of interest to compare the signal suppression for the Costas loop α_s^2 to that of the CW loop preceded by a soft BPL. Figure 3.65 gives this comparison as a function of IF SNR and for a single-pole RC filter in the arms of the loop. In this figure, the optimum bandwidth-to-data-rate ratio is chosen so as to minimize the combined limiter-squaring loss. Notice that α_s^2 changes by approximately two orders of magnitude as ρ_H varies between -20 dB and -5 dB; as one would expect, the loop bandwidth and the square of the loop damping would also vary by this amount over the dynamic range of the SNR ρ_H. Figure 3.66 illustrates this variation in the loop bandwidth versus ρ_H for those values of B_i/R_s which minimize the combined limiter-squaring loss. Although the curves in Figures 3.65 and 3.66 have been drawn for the single-pole RC filter, these results are virtually insensitive to the choice of arm filter type.

Now consider the evaluation of the modulation distortion factor, D_m, when the input modulation m(t) is an NRZ coding of equiprobable independent transmitted symbols. Since it is assumed that the input IF filter bandwidth is wide enough so as not to distort the modulation, the evaluation of D_m (due to the arm filters) is independent of the presence

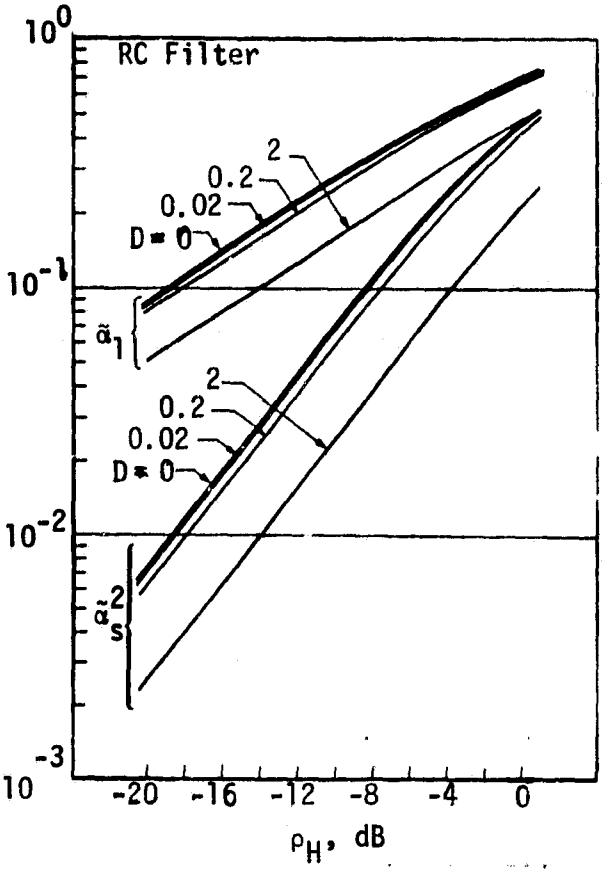


Figure 3.65. Signal Amplitude Suppression Factor Versus IF SNR for the CW and Costas Loop

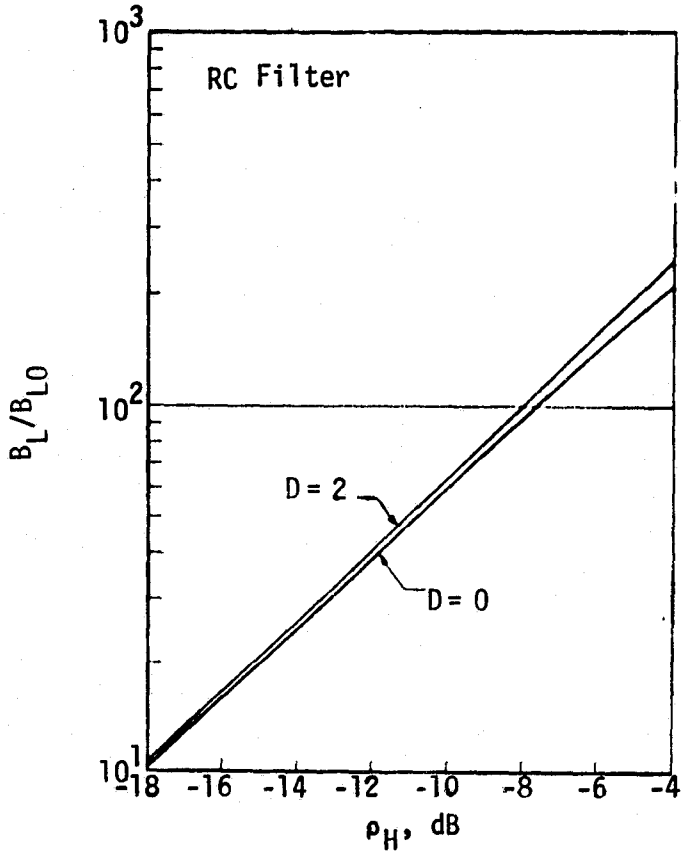


Figure 3.66. Variation of Loop Bandwidth with Signal Level for Optimum Values of B_i/R_s

of a BPL. Figures 3.67a-d illustrate the squaring loss L_{SQ} versus B_i/R_s and D, the "softness" of the BPL, for two (low and high) values of R_d as a parameter for single-pole and infinite-pole Butterworth filters. The IF filter is assumed to be a Gaussian filter. The dotted curve on each figure corresponds to the hard-limiter case, whereas the curves for D = 200 essentially represent the results for the no-limiter case, i.e., $D = \infty$. In each case, it is observed that, for fixed values of D and R_d, there is an optimum bandwidth B_i 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. As an example, for the ideal lowpass filter, the optimum ratio B_i/R_s varies from 1.0 to 1.6 as R_s varies between -10 dB and +10 dB and D varies between 0 and 200.

Figure 3.68 plots the ratio L (in dB) of L_{SQ} computed for the Costas loop preceded by a BPL with the L_{SQ} for the same Costas loop in the absence of the BPL at the optimum values of B_i/R_s as a function of R_d.

3.2.10.3 Unbalanced QPSK carrier-tracking loops

Unbalanced quadriphase-shift-keying (UQPSK) used on the Ku-band return link is an attractive means for transmitting two digital data streams which have different average powers. The two data streams are not constrained to have identical data rates nor must they have the same data format, e.g., one might be an NRZ sequence and the other a Manchester code. In fact, it is the difference in data rates which causes the imbalance of power when it is desired to have symbol energies and therefore error rates on the two channels within the same order of magnitude.

When the unbalanced power ratio is large (e.g., approximately 4:1 or greater), a biphase Costas loop, as presented in subsection 3.2.10.2, is more efficient for carrier tracking than the normal QPSK fourth-power tracking loop. To characterize the Costas loop for carrier tracking of UQPSK, L_{SQ} must be computed. The analysis of the Costas loop for tracking UQPSK is given in [29-31]. Using these results, the evaluation of L_{SQ} can be obtained for each case of interest.

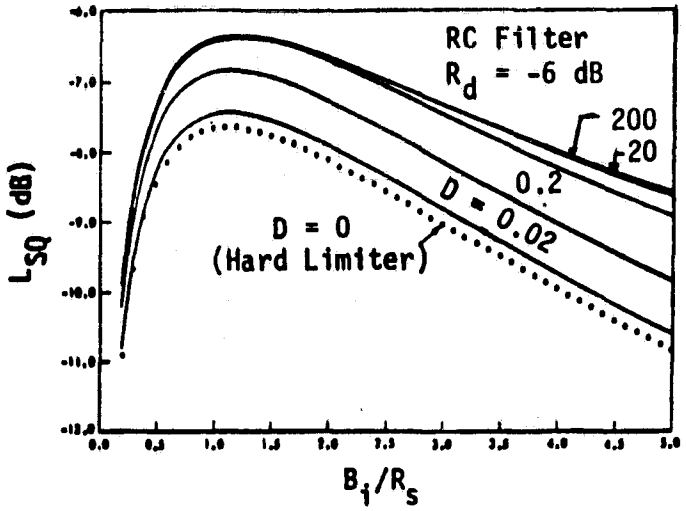


Figure 3.67a. Combined Limiter Squaring Loss Versus B_i/R_S for Various Values of D;R_d = -6 dB; RC Filter

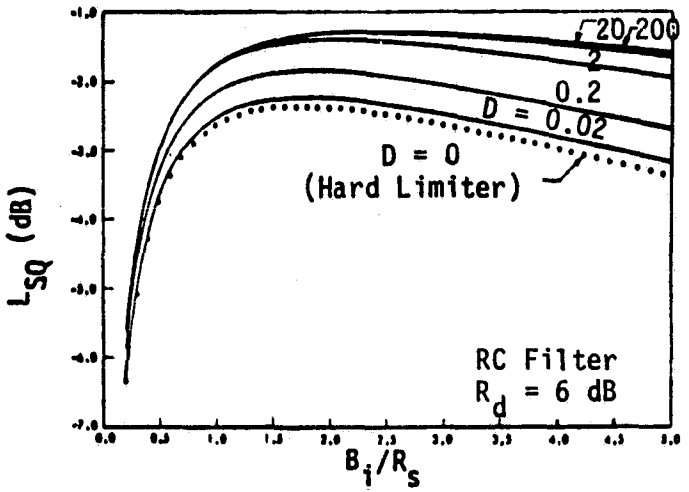


Figure 3.67b. Combined Limiter Squaring Loss Versus B_i/R_S for Various Values of D;R_d = 6 dB; RC Filter

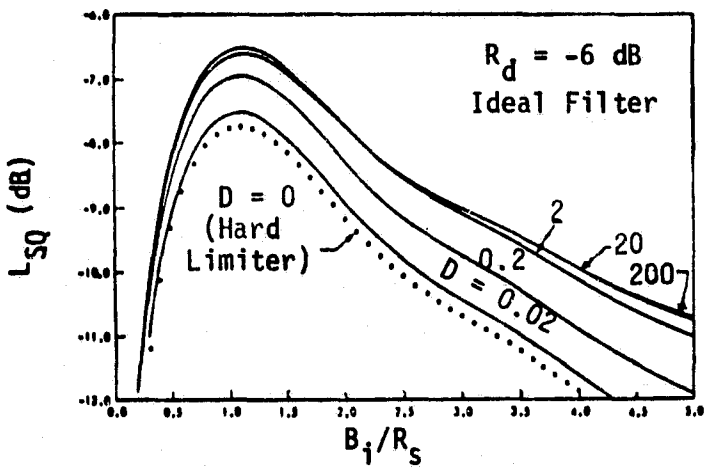


Figure 3.67c. Combined Limiter Squaring Loss Versus B_i/R_S for Various Values of D;R_d = -6 dB; Ideal Filter

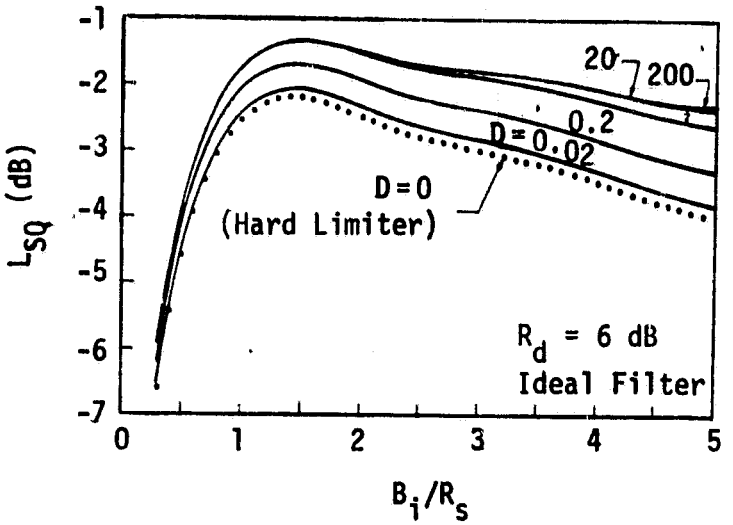


Figure 3.67d. Combined Limiter Squaring Loss Versus B_i/R_s for Various Values of D; R_d = 6 dB; Ideal Filter

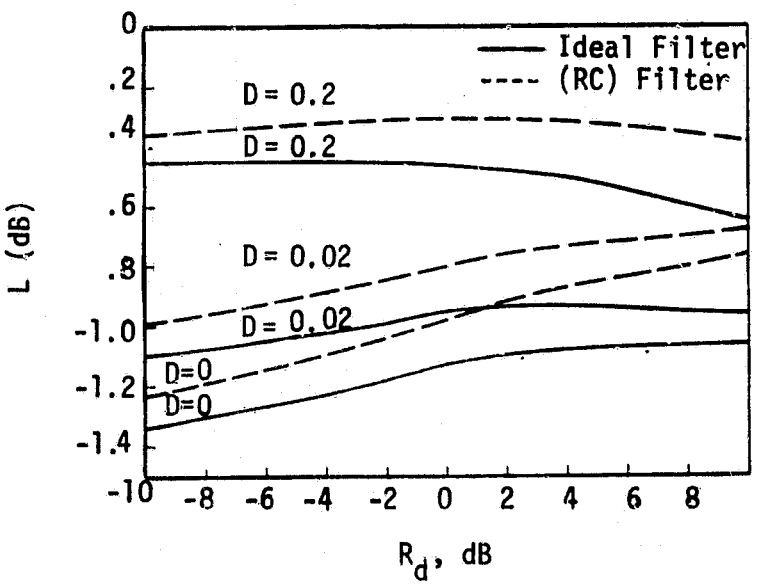


Figure 3.68. Loss in Loop SNR Due to the Presence of a Soft Limiter Versus R_d for Various Values of D

As a numerical illustration, consider the case where the low-rate modulation $m_1(t)$ is Manchester-coded data, the high-rate modulation $m_2(t)$ is NRZ data, and the arm filters are single-pole Butterworth (RC) filters. Then Figure 3.69 illustrates (for fixed fractional channel powers η_1 and η_2 corresponding to a 4:1 power ratio on the two channels) the behavior of L_{SQ} as a function of the ratio of two-sided arm filter noise bandwidth B_i to the higher of the two data rates $R_2 = 1/T_2$, with the ratio of data rates R_2/R_1 and $P_T T_2/N_0$ as parameters, where P_T/N_0 is the total power-to-noise spectral density. Assuming P_T/N_0 to be fixed, the variation of squaring loss with $P_T T_2/N_0$ directly reflects the effect of changing the high data rate R_2. Furthermore, at low values of B_i/R_2, we observe from Figure 3.69 that additional interesting peaks and valleys of the squaring loss characteristic occur. These extremes represent trade-offs between SxS distortion and cross-modulation noise or SxN power, depending on which of the latter dominates the total noise.

The corresponding numerical evaluation of the tracking jitter, for a fixed ratio of arm filter noise bandwidth to the loop noise bandwidth (B_i/B_L) is shown in Figure 3.70. The minimum values of σ_ϕ for some ratios of R_2/R_1 represent best design points when the combined effects of NxN distortion and cross-modulation noise or SxN power is minimal. Assuming P_T/N_0 to be fixed, the variation of σ_ϕ with $P_T T_2/N_0$ is shown in Figure 3.71. As is intuitively true, the tracking jitter performance improves with the increase of $P_T T_2/N_0$.

In a Costas loop study for biphase modulation [25], it was demonstrated that considerable improvement in tracking performance could be obtained by employing active arm filters of the integrate-and-dump type as opposed to passive arm filters. An investigation to determine if a similar performance improvement can be obtained for an unbalanced OPSK modulation is presented in detail in [29].

It is observed from the previous discussion that, for given values of the data rate ratio R_2/R_1, power ratio P_1/P_2, and total SNR in the high data rate bandwidth $P_T T_2/N_0$, an optimum arm filter bandwidth (or equivalently, B_i/R_2) exists in the sense of maximizing L_{SQ} (i.e., minimizing the squaring loss). Using that value of B_i/R_2,

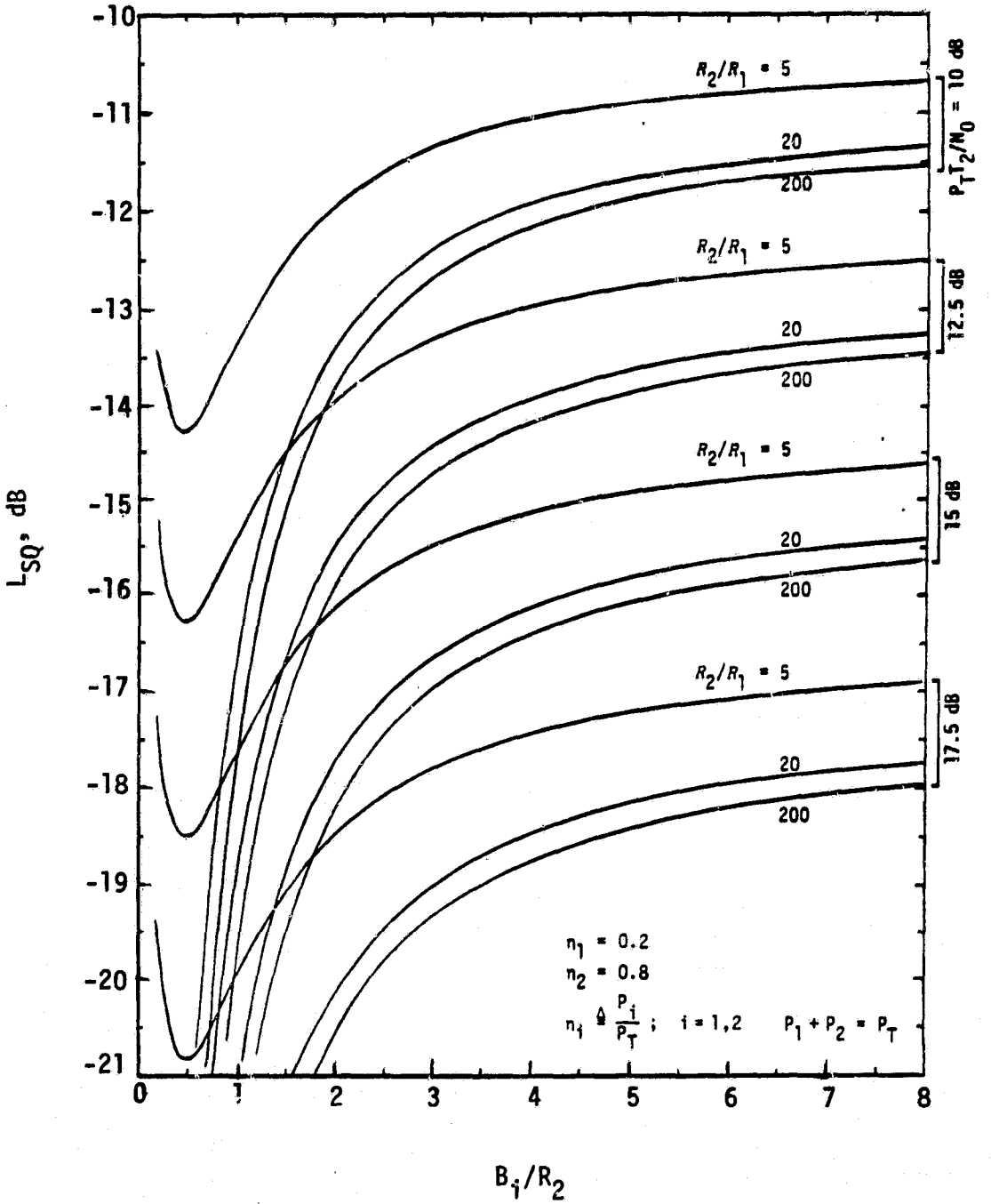


Figure 3.69. Squaring Loss Versus Ratio of Arm Filter Bandwidth to High-Data Rate; P_TT_2/N_0 and R_2/R_1 are Parameters; $m_1(t)$ is Manchester Code, $m_2(t)$ is NRZ; $R_2 \geq 2R_1$

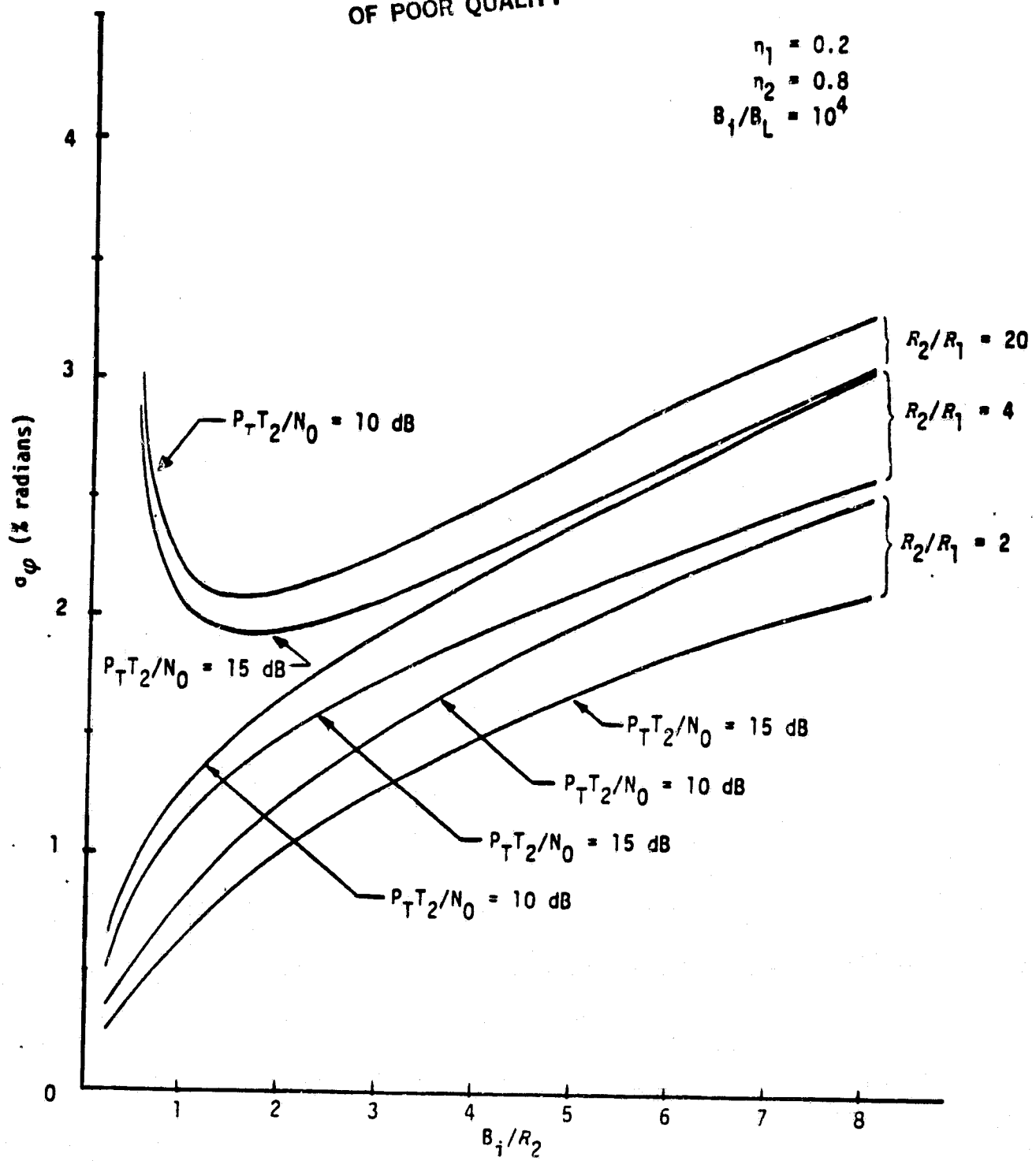


Figure 3.70. Tracking Jitter Standard Deviation Versus Ratio of Arm Filter Bandwidth to High-Data Rate; P_TT_2/N_0 and R_2/R_1 are Parameters; $m_1(t)$ is Manchester Code, $m_2(t)$ is NRZ; $R_2 \geq R_1$

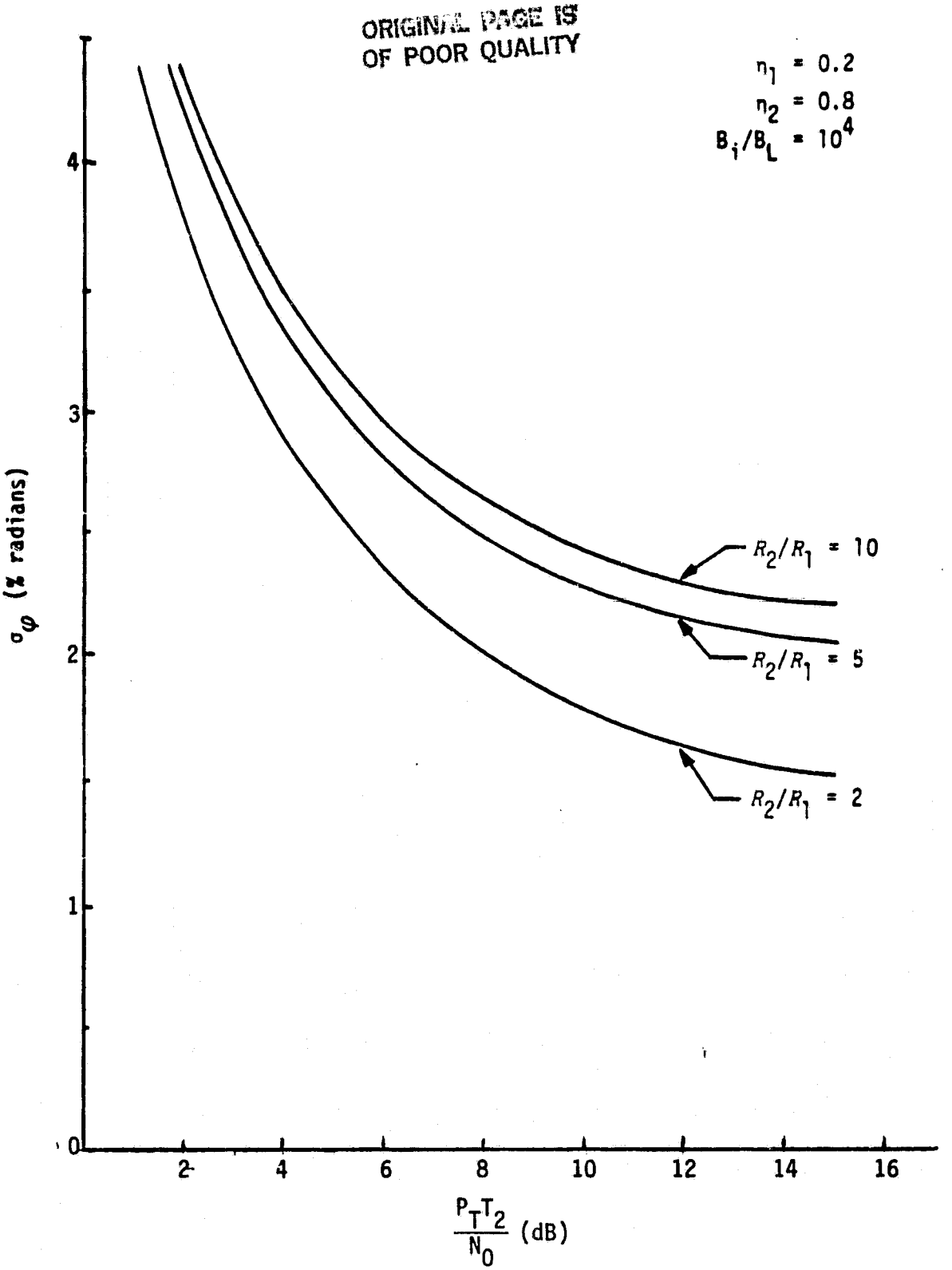


Figure 3.71. Tracking Jitter Standard Deviation versus $P_T T_2/N_0$; B_i/R_2=4; R_2/R_1 is a Parameter; $m_1(t)$ is Manchester Code, $m_2(t)$ is NRZ; $R_2 \geq R_1$

$(B_1/R_2)_{opt}$, and defining the corresponding value of L_{SQ} by $(L_{SQ})_{opt}$, the minimum improvement in tracking performance (or, equivalently, in squaring-loss performance) obtained by employing active integrate-and-dump arm filters as opposed to active arm filters is given by $I = L_{SQ}/(L_{SQ})_{opt}$, where L_{SQ} is used to denote the squaring loss of the Costas loop with integrate-and dump arm filters.

Closed-form expressions for L_{SQ} are derived in [29] for all combinations of NRZ and Manchester data formats for the two channels and both synchronized and unsynchronized symbol clocks. As an example, assuming single-pole (RC) arm filters as the basis of comparison for the Costas loop with passive arm filters and assuming unsynchronized symbol clocks for the Costas loop with integrate-and-dump filters, Figures 3.72 and 3.73 illustrate I (in decibels) versus the channel power ratio P_1/P_2 with the data rate ratio R_2/R_1 as a parameter and values of total power-to-noise ratio $P_T T_2/N_0$ typical of coded and uncoded systems. It is observed from these figures that the improvement in squaring-loss performance of using integrate-and-dump arm filters as opposed to single-pole arm filters is an increasing function of P_1/P_2 and depends heavily on the choice of data formats for $m_1(t)$ and $m_2(t)$.

Similarly, using many numerical illustrations, it is shown [29] that, for a fixed ratio of data rates and total power-to-noise ratio in the higher data bandwidth, the squaring loss itself increases with the ratio of powers in the two channels, and the rate at which this loss increases (i.e., tracking performance deteriorates) also depends heavily upon the data formats in each channel. Thus, it is concluded that, when the ratio of data rates is of the same order of magnitude as the inverse of the power ratio, i.e., approximately equal signal energies in the two channels, the biphase Costas loop can be used as an efficient demodulator of QPSK. On the other hand, if the energy in the two channels is very unbalanced, e.g., one channel is coded and one is uncoded, then it is still possible to efficiently use a biphase Costas loop for demodulation of unbalanced QPSK, provided the higher data rate channel is Manchester coded. It is understood that the foregoing conclusions are quite general and are not intended

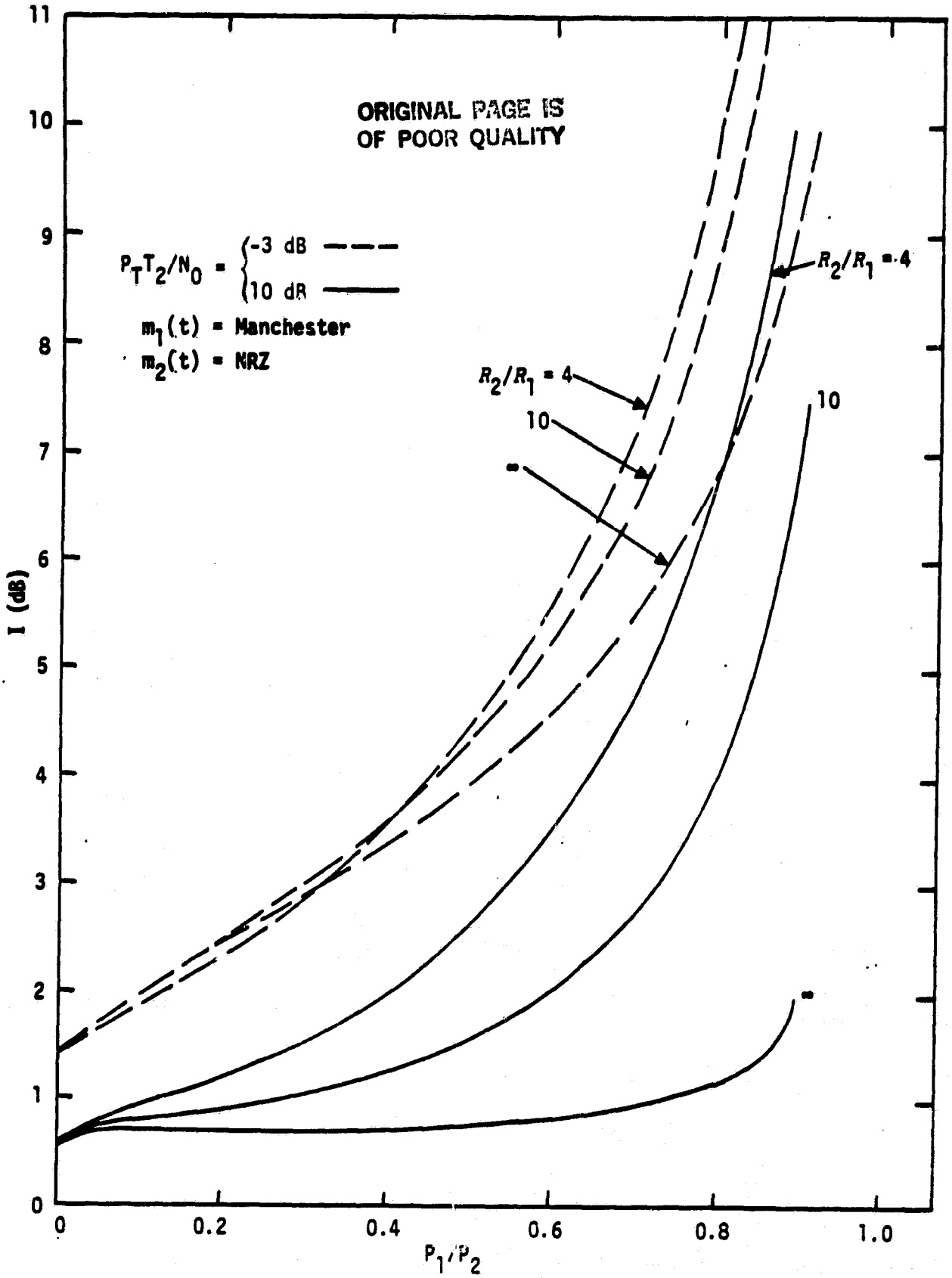


Figure 3.72. Squaring-Loss Improvement (in dB) Using Integrate-and-Dump Arm Filters as Opposed to Single-Pole Passive Arm Filters

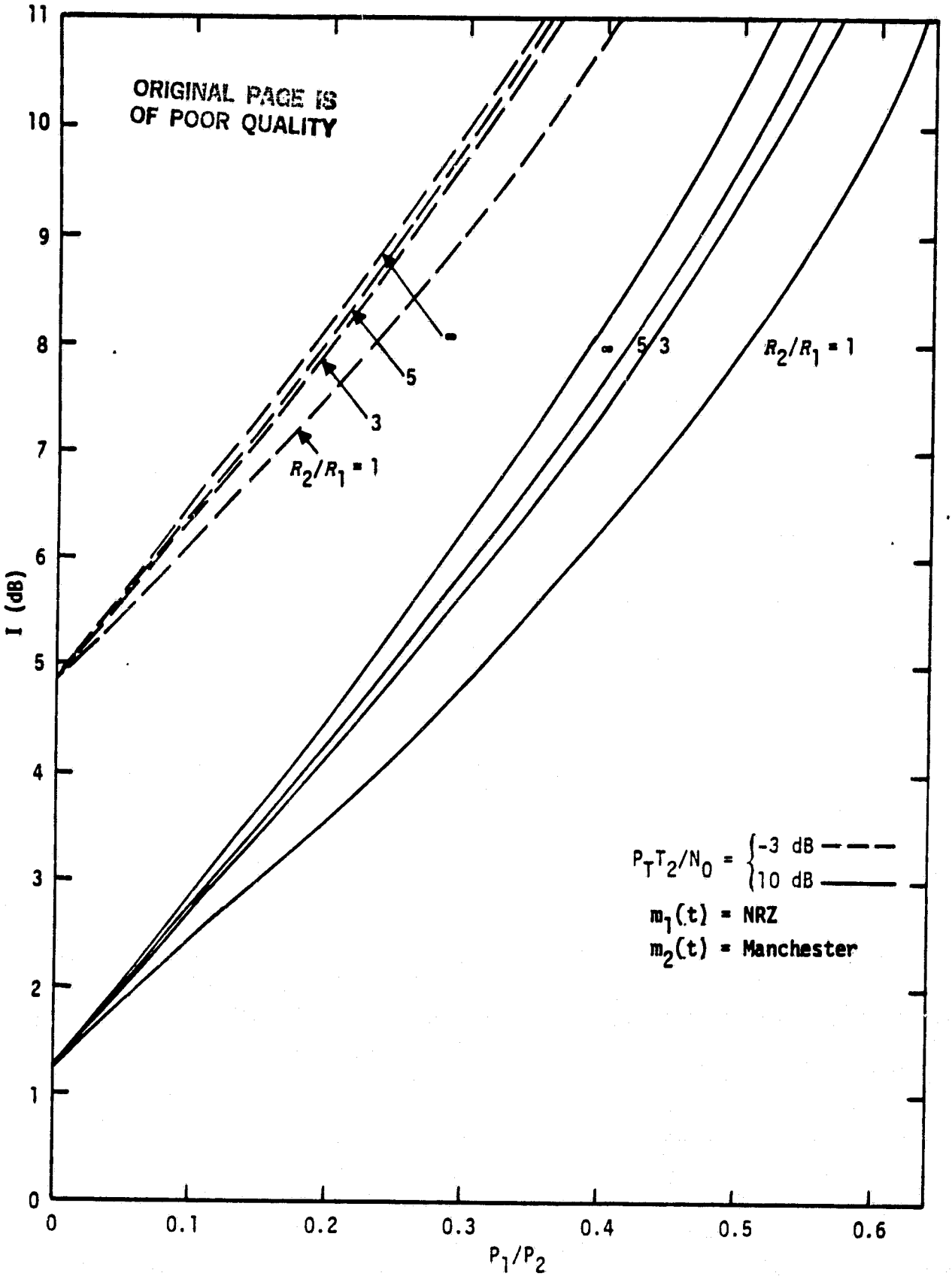


Figure 3.73. Squaring-Loss Improvement (in dB) Using Integrate-and-Dump Arm Filters as Opposed to Single-Pole Passive Arm Filters

to rule out specific design situations in which sufficient total power-to-noise ratio is available to tolerate large squaring losses. In an individual situation, one must resort to the specific numerical results given in the illustrations to determine the suitability of employing a biphase Costas loop for demodulation of unbalanced QPSK.

So far, the problem of tracking an unbalanced QPSK signal with a conventional biphase Costas loop with analog input phase detectors and an analog third multiplier (the one that forms the loop error signal) has been addressed. Because of DC offsets associated with analog multipliers, it is common practice to hard-limit the inphase* channel arm filter output and replace the analog third multiplier with a chopper-type device (switched multiplier) which typically exhibits much less offset (see Figure 3.74). While it is also possible to replace the input inphase and quadrature analog phase detectors with switched multipliers, the impact on the resulting tracking performance will be minimal since the arm filters will pass only the first harmonic of these phase detector outputs. Thus, aside from the $8/\pi^2$ power loss associated with the first harmonic of a square wave, the performance of the loop would be identical to that given in the previous paragraphs for an analog third multiplier or that to be presented here for a switched third multiplier. For ease of terminology, a conventional biphase Costas loop with a switched third multiplier is referred to as a "biphase polarity-type Costas loop" or, even simpler, a "polarity-type Costas loop."

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 SNR at the limiter input. This signal suppression in turn reduces the total loop gain and, as a consequence, the loop bandwidth. Another potential problem with the limiter under strong signal conditions is that it may increase the tendency of the loop to false lock.

---

*For unbalanced quadriphase, we shall arbitrarily refer to the in-phase channel as that corresponding to the point of data extraction for the higher power signal

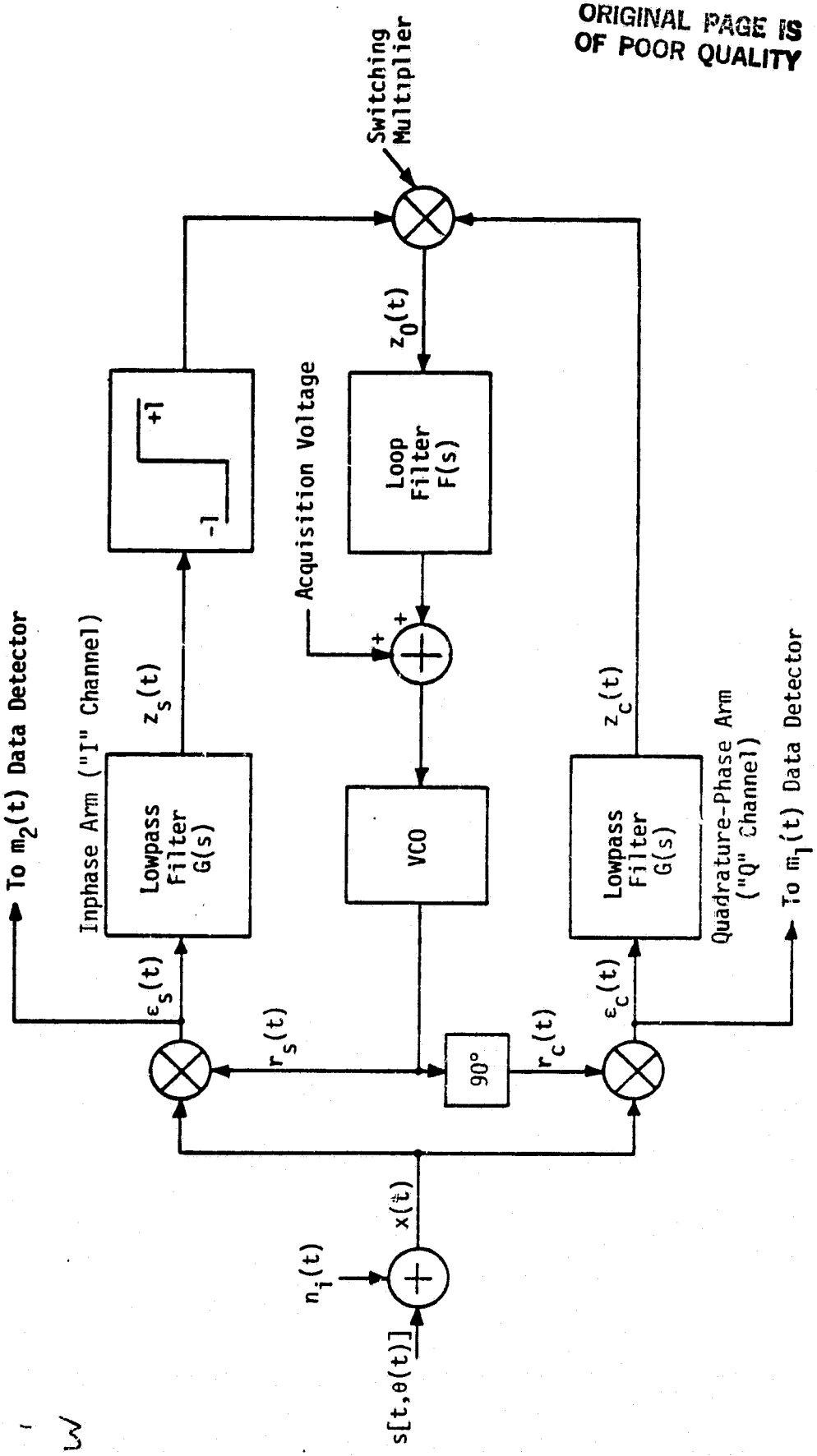


Figure 3.74. Costas Loop with Hard-Limited Inphase Channel

The tracking behavior of the polarity-type Costas loop with unbalanced QPSK input is discussed in detail in [29] and its performance compared to that of the conventional Costas loop. In particular, for the case of single-pole Butterworth (RC) arm filters and a particular combination of NRZ and Manchester-coded data on the two channels, the squaring loss (tracking jitter penalty relative to a linear loop) is evaluated and illustrated as a function of the ratio of arm filter bandwidth to higher data rate and total signal power-to-noise ratio in this higher data rate bandwidth. Also numerically illustrated is the corresponding mean-squared tracking-jitter performance as a function of these same receiver parameters. A summary of the results is given as follows. Figure 3.75 illustrates the variations of L_{SQ} versus B_i/R_2 with $P_T T_2/N_0$ as a parameter. Superimposed on these curves (in dashed lines) are the corresponding results obtained from Figure 3.69 for the biphase Costas loop with passive arm filters. It is observed from these numerical results that, for high SNR, the hard-limited loop actually outperforms the conventional loop and, depending on the data rate ratio, the improvement (in terms of squaring loss) might be as high as 2.8 dB. Also, for a given SNR and an arm filter bandwidth to high-data rate ratio, the squaring loss does not change significantly with data rates when the ratio of the data rates is high. This is particularly true for small values of arm filter bandwidth to high-data rate ratio. A comparison with the dotted curves of Figure 3.75 reveals that the same is not true for the conventional Costas loop with passive arm filters. However, the fact that the polarity-type Costas loop produces an improvement in tracking performance at high SNR over the biphase Costas loop with passive arm filters is not surprising in view of similar results recently demonstrated for biphase modulation [22].

3.2.10.4 PN spread-spectrum-tracking loops

There are two PN tracking-loop configurations which are used extensively in PN spread spectrum communication systems These loops are the delay-lock-loop (DLL) shown in Figure 3.76 and the tau-dither loop (TDL) or time-shared loop shown in Figure 3.77. Consider the noncoherent

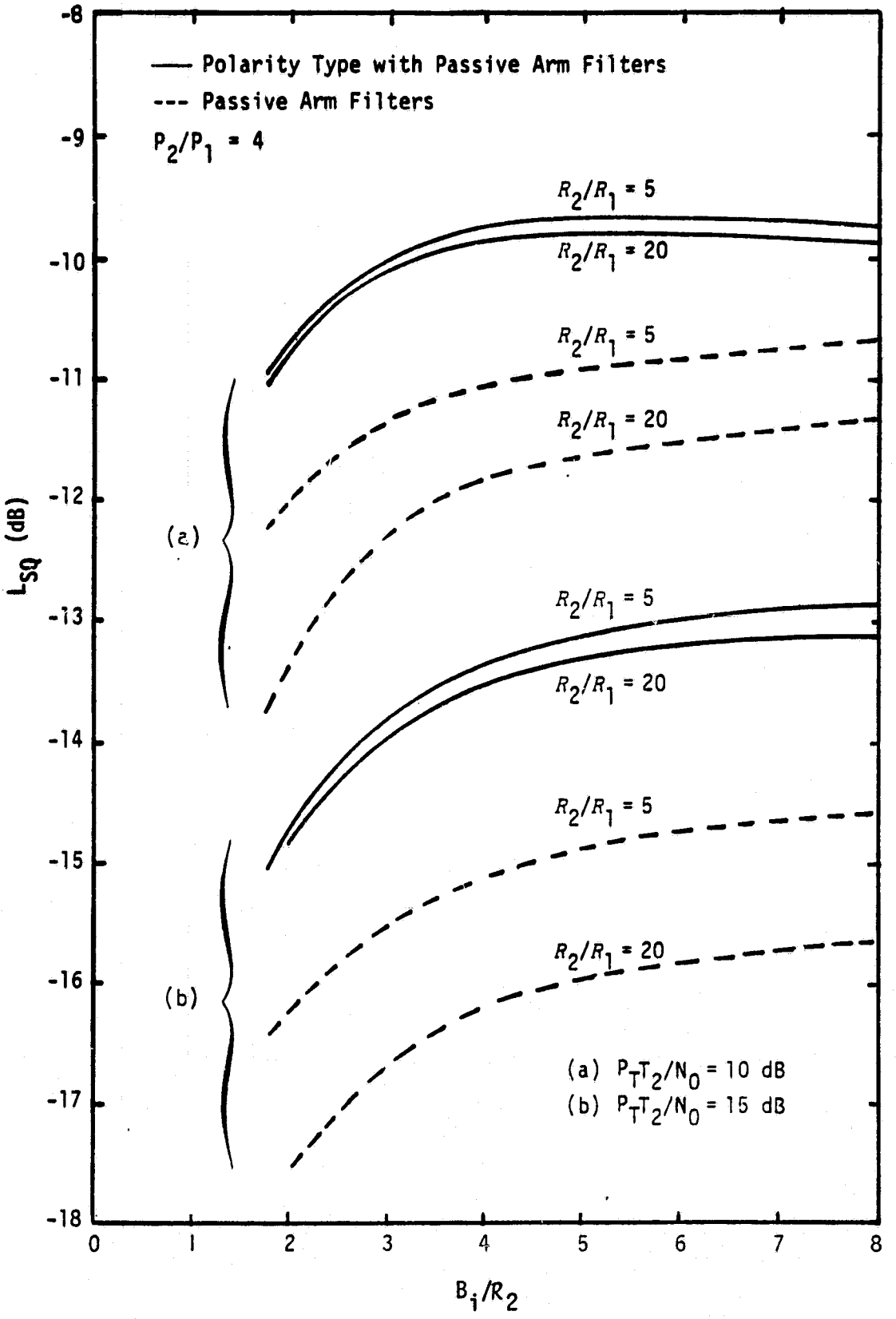


Figure 3.75 Squaring-Loss Variations Versus B_i/R_2 with R_2/R_1 and $P_T T_2/N_0$ as Parameters; $m_1(t)$ is Manchester Code, $m_2(t)$ is NRZ

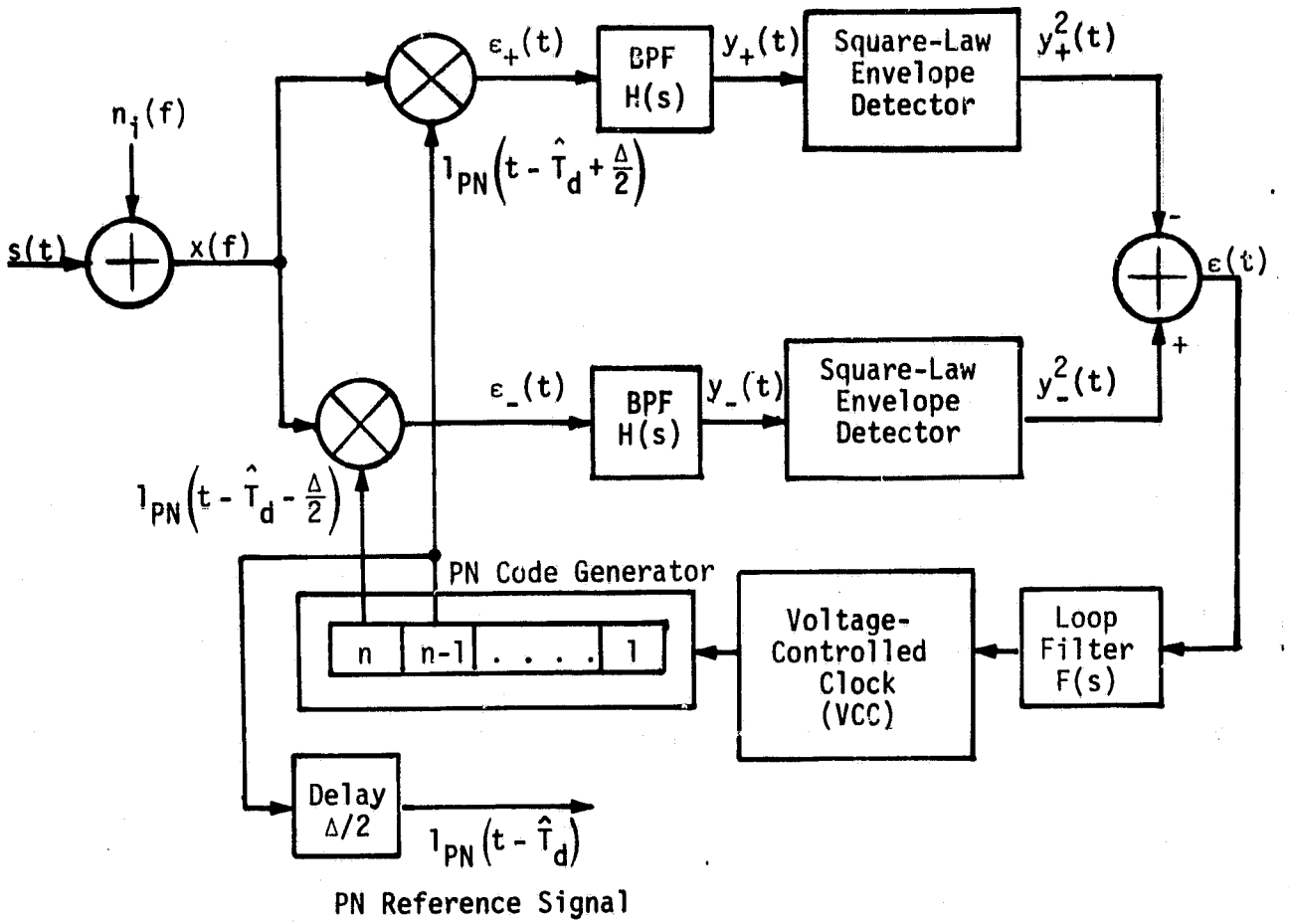


Figure 3.76. A Noncoherent "One-Δ" Delay-Locked Loop

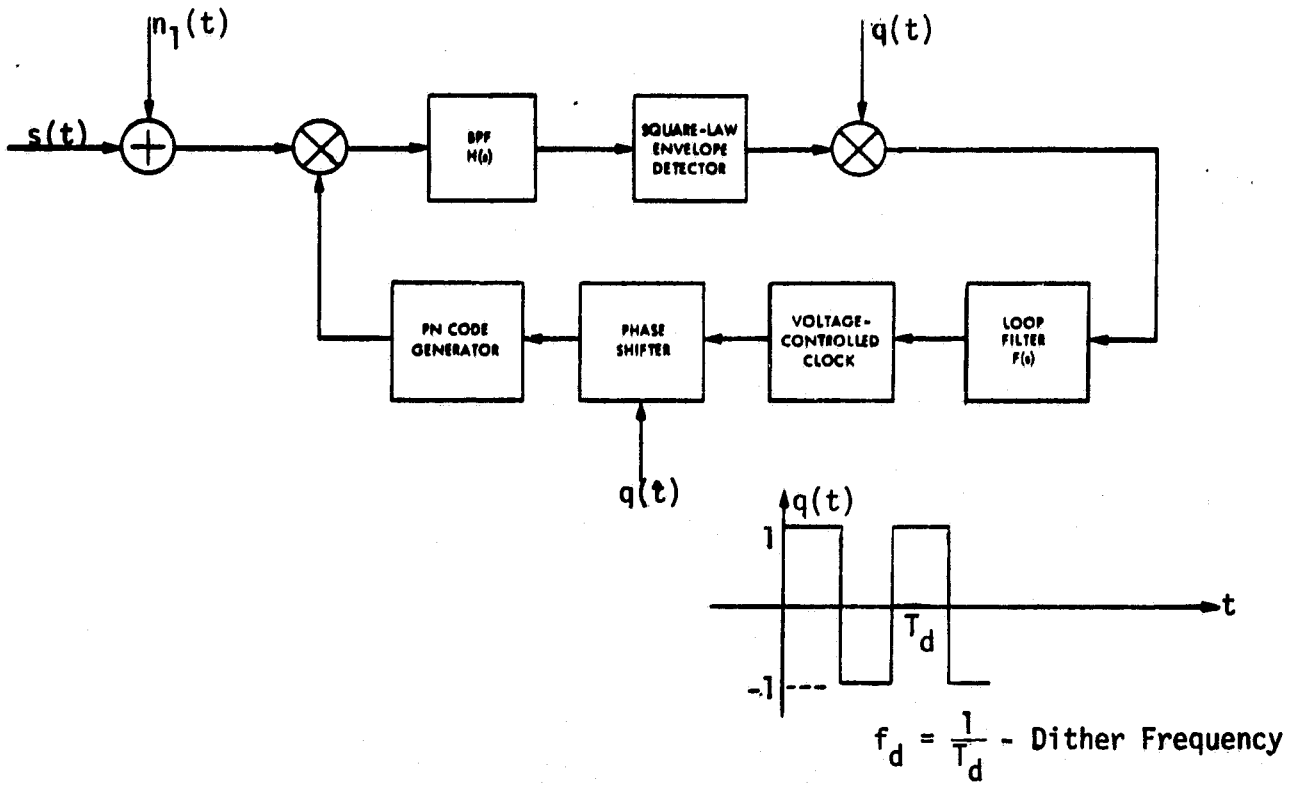


Figure 3.77. A Noncoherent "One Δ" Tau-Dither Loop

noncoherent delay-locked-loop illustrated in Figure 3.76. The input signal x(t) is cross correlated with advanced and retarded versions of the local PN code-generator sequence. The results of these cross-correlation operations are then bandpass-filtered, squared and differenced to produce an error (discriminator) characteristic. The loop is closed by applying this differenced output to a loop filter and voltage-controlled clock (VCC) that drives the PN code generator from which the PN reference sequence is obtained. When the advance (and retard) interval is equal to one-half of a PN code chip, the loop is commonly referred to as a "one-Δ" loop, where Δ denotes the length (in seconds) of a PN code chip.

The tracking performance of the DLL is measured by the variance of the timing jitter. For large η, to a first approximation,

$$\left(\frac{\sigma_t}{\Delta}\right)^2 = \frac{\overline{\tau_e^2}}{\Delta^2} = \frac{1}{2\eta L_{SQ}} , \tag{3-227}$$

where τ_e is the delay error equal to $\tau_d - \hat{\tau}_d$. Note that (3-227) is the identical form as (3-206) for a squaring loop except for a factor of 8. The L_{SQ} for a DLL is given by

$$L_{SQ} = \frac{D_m}{K_s + K_L \dfrac{2B_i/R_s}{R_d D_m}} , \tag{3-228}$$

where the terms in the expression for L_{SQ} are identical to the L_{SQ} for the squaring loop in (3-207) except for a factor of 4 in the second term of the denominator of L_{SQ}. This factor of 4 has the effect of decreasing the effective R_d by 6 dB. Therefore, the curves used for the squaring loop can be applied to the DLL, and vice versa, by adjusting R_d by 6 dB. Thus, if the R_d in Figures 3.57 and 3.58 is increased by 6 dB, the L_{SQ} for the DLL is obtained.

A more exact value of $\sigma_\gamma = \sigma_t/\Delta$ can be computed by performing a nonlinear tracking performance analysis of the DLL [33]. Figures 3.78 and 3.79 illustrate $(\sigma_\gamma)_{min}$ versus R_d with $\delta = R_s/B_L$ as a parameter for a two-pole Butterworth and ideal equivalent lowpass filter,

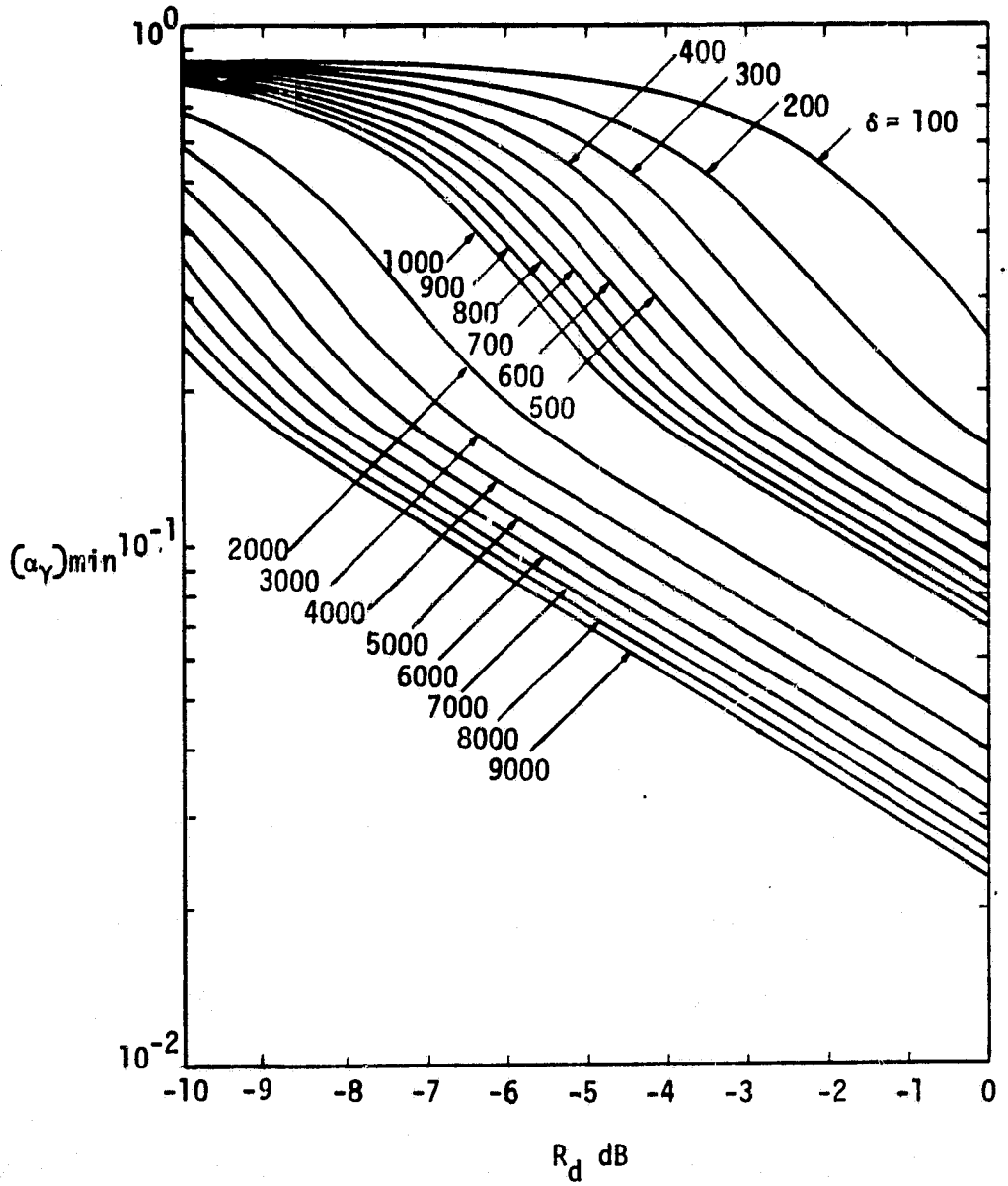


Figure 3.78. Nonlinear Tracking Jitter Performance of Noncoherent Delay-Lock-Loop; Two-Pole Butterworth Filter

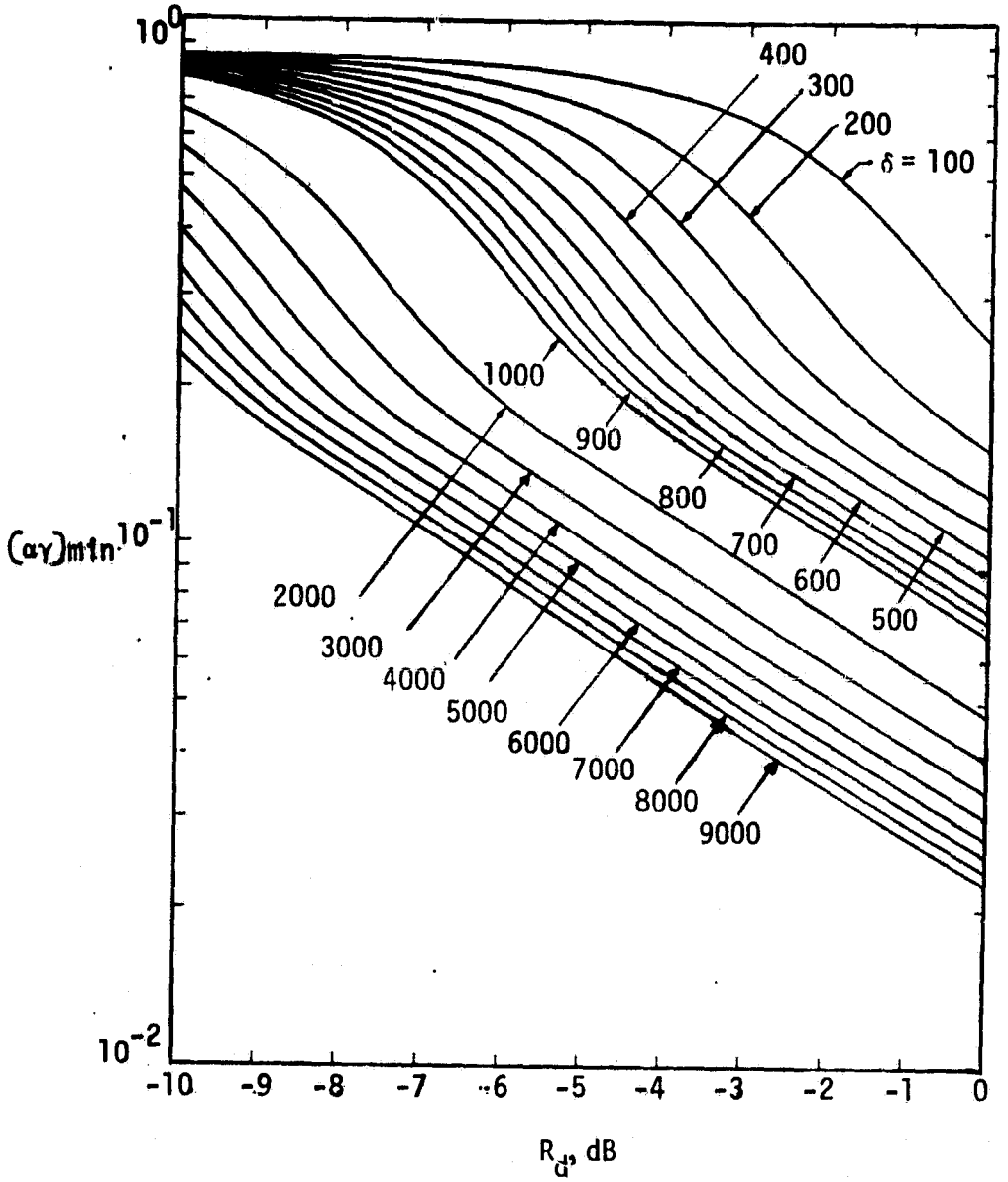


Figure 3.79. Nonlinear Tracking Jitter Performance of Noncoherent Delay-Lock-Loop; Ideal Filter

respectively. Note that $(\sigma_\gamma)_{min}'$ is obtained by minimizing σ_γ with respect to B_i/R_s. It may be observed from Figures 3.78 and 3.79 that the results are relatively insensitive to the filter type.

Now consider the noncoherent TDL illustrated in Figure 3.77 whose operation is described as follows. The received signal-plus-noise is alternately (as opposed to simultaneously) correlated with the advanced and retarded versions of the locally generated PN code (thus, the name "dither" loop) to produce an error signal which, when bandpass filtered, envelope detected and alternately inverted in synchronism by the binary signal q(t), drives the VCC through the loop filter F(s). One obvious advantage of the TDL over the DLL, which was alluded to previously, is the fact that only a single input correlator/filter is required, thus eliminating the problems of gain imbalance and other mismatches which are present in a two-channel loop such as the DLL.

The tracking performance of the TDL is measured by (σ_t'/Δ) which is, for large η, to a first approximation

$$\left(\frac{\sigma_t'}{\Delta}\right)^2 = \frac{1}{2\eta L_{SQ}'}, \tag{3-229}$$

L_{SQ}' is the squaring loss of the TDL. Thus, a comparison of the linear tracking-jitter performances of the DLL and the TDL depends simply on the ratio of L_{SQ}' to L_{SQ}, namely,

$$\frac{L_{SQ}'}{L_{SQ}} = \frac{K_s + K_L \dfrac{2B_i/R_s}{R_d D_m}}{K_s' + K_L' \dfrac{2B_i/R_s}{R_d D_m}}, \tag{3-230}$$

where

$$K_s' = K_s + 2 \sum_{\substack{n=1 \\ \text{odd}}}^{\infty} \left(\frac{2}{n\pi}\right)^2 \frac{\displaystyle\int_{-\infty}^{\infty} S(f)|H_\ell(j2\pi f)|^2 \left|H_\ell\left[j2\pi\left(\frac{n}{T_d} - f\right)\right]\right|^2 df}{\displaystyle\int_{-\infty}^{\infty} S(f)\,|H_\ell(j2\pi f)|^2\, df} \tag{3-231}$$

and

$$K_L' = K_L + 2 \sum_{\substack{n=1 \\ \text{odd}}}^{\infty} \left(\frac{2}{n\pi}\right)^2 \frac{\int_{-\infty}^{\infty} |H_\ell(j2\pi f)|^2 \left|H_\ell\left[j2\pi\left(\frac{n}{T_d} - f\right)\right]\right|^2 df}{\int_{-\infty}^{\infty} |H_\ell(j2\pi f)|^2 \, df} \tag{3-232}$$

Since the ratio of the integrals in (3-231) and (3-232) are bounded by K_s and K_L, respectively, then

$$\begin{aligned} K_s' &< 2K_s \\ K_L' &< 2K_L \end{aligned} \tag{3-233}$$

and

$$\frac{L_{SQ}'}{L_{SQ}} > \frac{1}{2} \tag{3-234}$$

or, equivalently, the linear theory mean-squared timing error for the TDL is less than 3 dB worse than that of the DLL.

Although the integrals in (3-231) and (3-232) are difficult to evaluate, Figures 3.80 and 3.81 plot K_L' and K_s', respectively, as a function of B_iT_d for an ideal lowpass filter and Manchester-coded data with various values of filter bandwidth to data rate, B_i/R_s.

At this point, it is reasonable to expect that, if one were to plot L_{SQ}' versus B_i/R_s with R_d as a parameter, then there would be an optimum filter bandwidth at each value of R_d in the sense of minimizing the loop's squaring loss. Figure 3.82 illustrates the validity of this statement for the case of an ideal filter. In Figure 3.82, B_iT_d is chosen equal to 4. As B_iT_d is increased (typically, by lowering the dither frequency relative to the arm filter bandwidth), the performance penalty in comparison to the DLL also increases, approaching 1.5 dB in the limit as B_iT_d approaches infinity. Clearly, this situation is never reached in practice or theory since the assumption that the dither frequency is large relative to the loop bandwidth breaks down in the analysis.

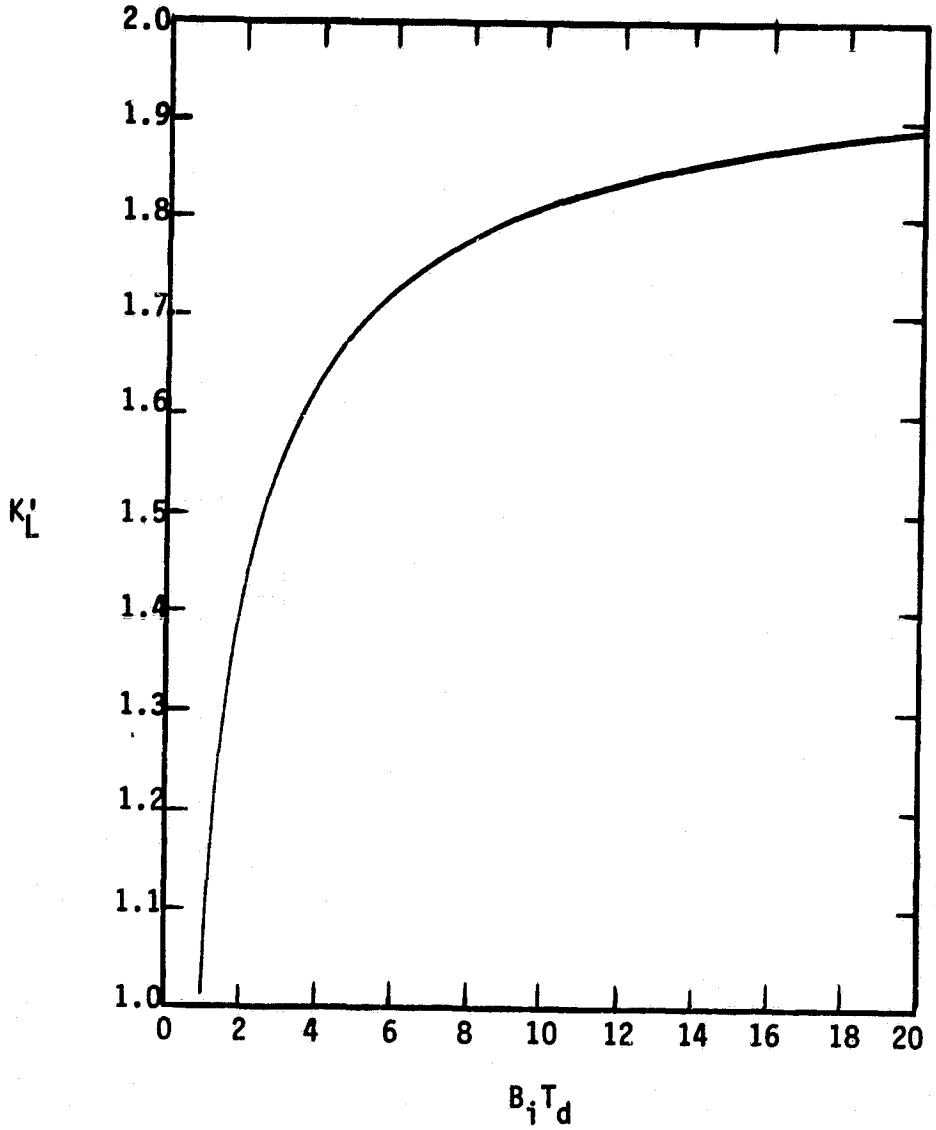


Figure 3.80. Plot of K_L' Versus B_iT_d; Ideal Filter

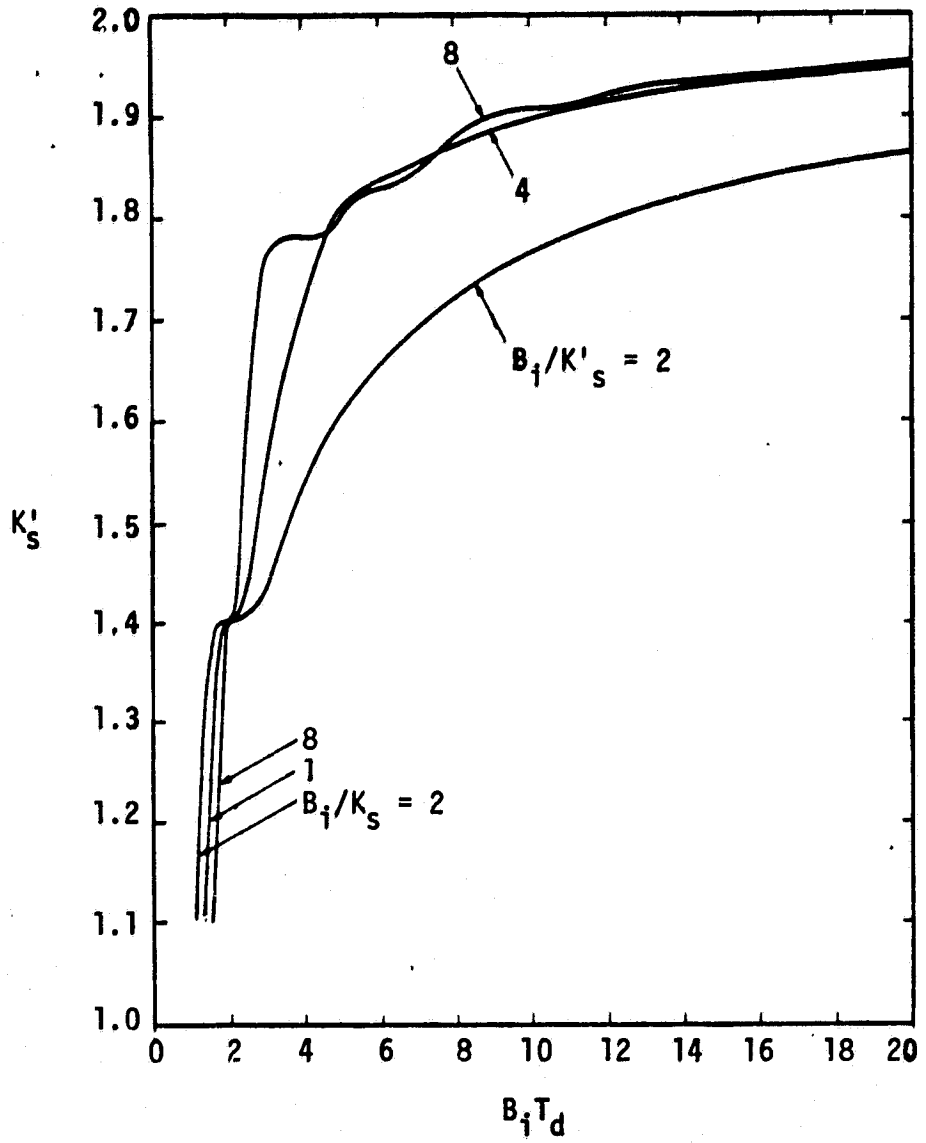


Figure 3.81. Plots of K'_s Versus B_iT_d With B_i/R_s as a Parameter; Ideal Filter; Manchester-Coded Data

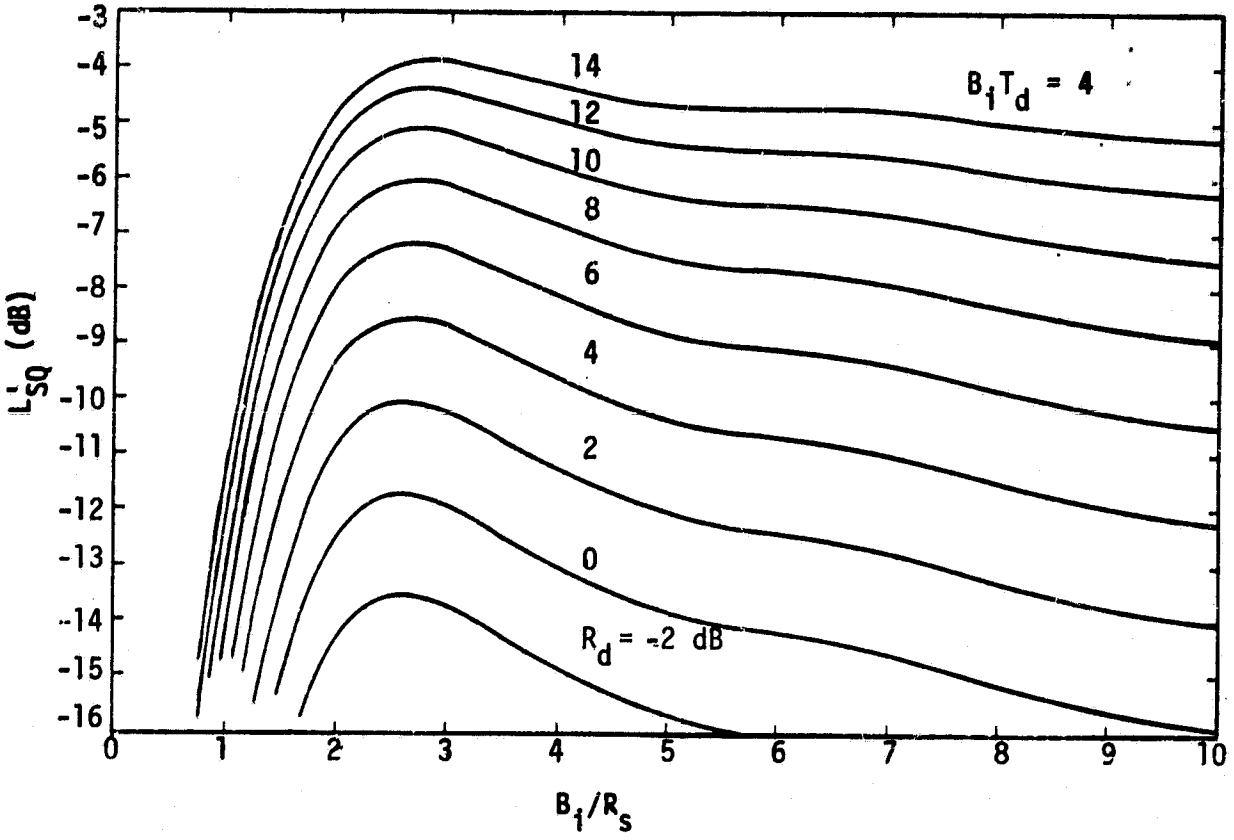


Figure 3.82. Squaring-Loss Variations Versus B_i/R_s for Various Values of R_d; Ideal Filter

Again, a more exact value of $\sigma_\gamma' = \sigma_t'/\Delta$ can be computed by performing a nonlinear tracking-performance analysis of TDL [33]. Figure 3.83 illustrates $(\sigma_\gamma')_{min}$ versus R_d with δ as a parameter for $B_i T_d = 4$ and an ideal arm filter. Note that $(\sigma_\gamma')_{min}$ is obtained by minimizing σ_γ with respect to B_i/R_s.

3.2.11 Carrier Demodulation

This subsection identifies the various losses associated with carrier demodulation. Both coherent demodulation losses for carrier demodulation using the carrier-tracking loops presented in subsection 3.2.10 and FM demodulation losses are presented.

3.2.11.1 Coherent demodulation losses

The additional coherent demodulation losses that were not discussed in subsection 3.2.10 are noisy oscillator loss L_{OSC}, PN despreading losses (including the PN filter loss L_{PN} and the PN correlation loss L_{COR}), the waveform distortion loss L_{WF}, and the carrier phase noise loss L_ϕ. In order to define losses due to demodulation, it is helpful to establish the ideal performance of coherent demodulation. The most commonly used performance measure for digital data is the probability of symbol error P_s versus E_s/N_0 or, equivalently, the bit error probability P_b versus E_b/N_0 for no error-correction coding. For antipodal signals (binary PSK) and ideal coherent demodulation, P_s is given by [35]

$$P_s = Q\left(\sqrt{2E_s/N_0}\right) \tag{3-235}$$

where the function Q( ) is defined by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-y^2/2\right) dy \ . \tag{3-236}$$

For orthogonal signals (binary FSK) and ideal coherent demodulation, P_s is given by [35]

$$P_s = Q\left(\sqrt{E_s/N_0}\right) . \tag{3-237}$$

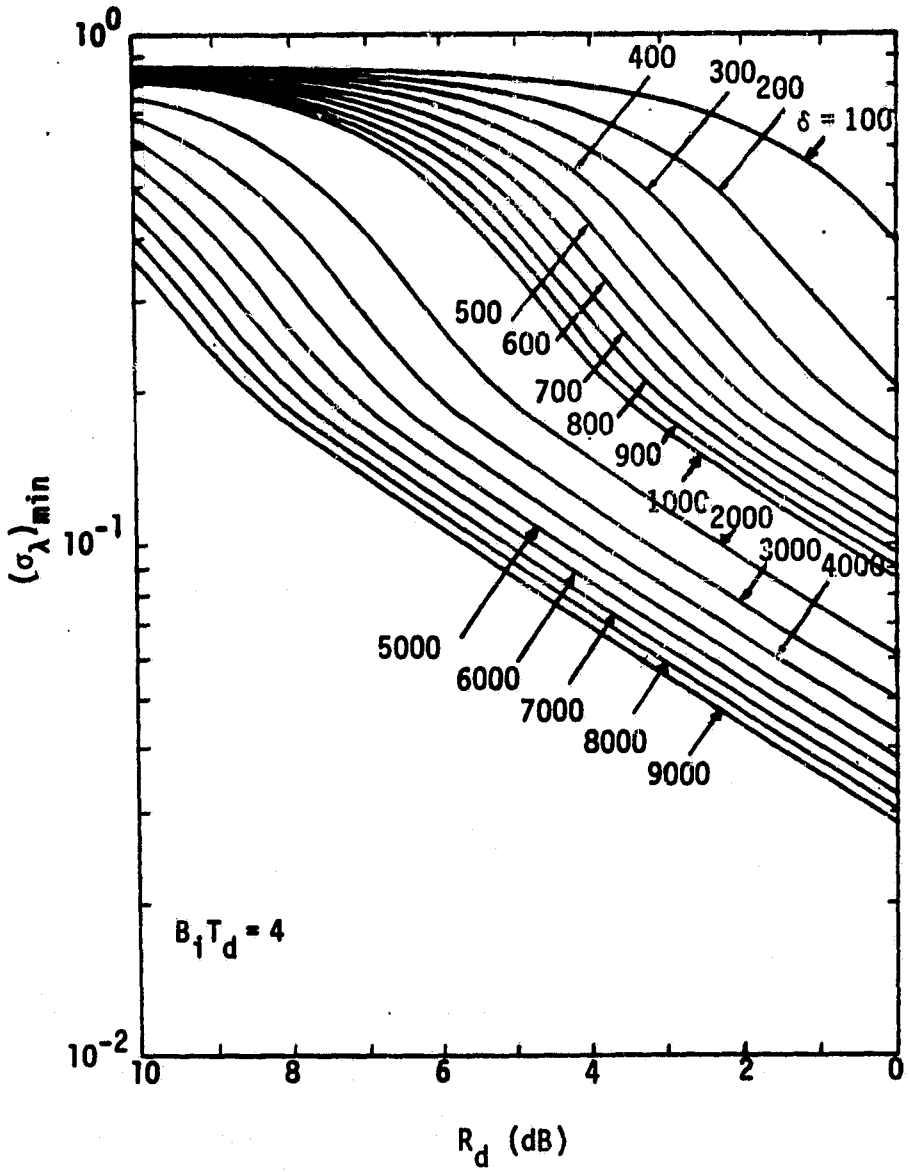


Figure 3.83. Nonlinear Tracking Jitter Performance of a Noncoherent Tau-Dither Loop; Ideal Filter

Figure 3.84 presents P_s versus E_s/N_0 for antipodal and orthogonal signals. The demodulation loss due to a particular contribution is defined as the increase in E_s/N_0 that is required to achieve the same average P_s as an ideal coherent demodulator.

To determine the demodulation loss due to oscillator instabilities, the oscillator phase statistics must be modeled and related to commonly measured frequency stability measures. The carrier, neglecting modulation at the input to the carrier-tracking loop, is assumed to be of the form

$$s(t) = \sqrt{2P}\sin\left[\omega_0 t + \phi(t) + \frac{dt^2}{2}\right] \tag{3-238}$$

where d represents the aging effect (long-term drift) of an oscillator and $\phi(t)$ represents the phase jitter. Several oscillators in the communication link may have contributed to this phase noise. All the local oscillators or frequency synthesizers in the transmitter upconverter, the relay frequency translator and the receiver downconverter contribute to the phase noise.

The phase noise $\phi(t)$ is a superposition of causally generated signals and nondeterministic random noise. The causally generated effects are created by changes in the oscillator temperature, supply voltage or power-line fluctuations, magnetic field, humidity, physical vibration, or output load impedance.

The PSD representation of phase noise relies on an assumption of a wide-sense stationary model for the process that produced the collected oscillator phase noise statistics. However, one need assume only that the measured data is consistent with a stationary model over the interval of interest, not that the oscillator physical noise sources are completely stationary since, in nature, very few sources are.

A spectral model that has been found very useful represents the phase noise PSD (one-sided) as:

$$G_\phi(f) = \underbrace{k_{-4}\frac{f_0^2}{f^4}}_{\substack{\text{random}\\ \text{frequency}\\ \text{walk}}} + \underbrace{k_{-3}\frac{f_0^2}{f^3}}_{\substack{\text{flicker}\\ \text{frequency}\\ \text{noise}}} + \underbrace{k_{-2}\frac{f_0^2}{f^2}}_{\substack{\text{random}\\ \text{phase}\\ \text{walk or}\\ \text{white}\\ \text{frequency}\\ \text{noise}}} + \underbrace{k_{-1}\frac{f_0^2}{f}}_{\substack{\text{flicker}\\ \text{phase}\\ \text{noise}}} + \underbrace{k_0 f_0^2}_{\substack{\text{white}\\ \text{phase}\\ \text{noise}}} \quad f_l \le f \le f_h \tag{3-239}$$

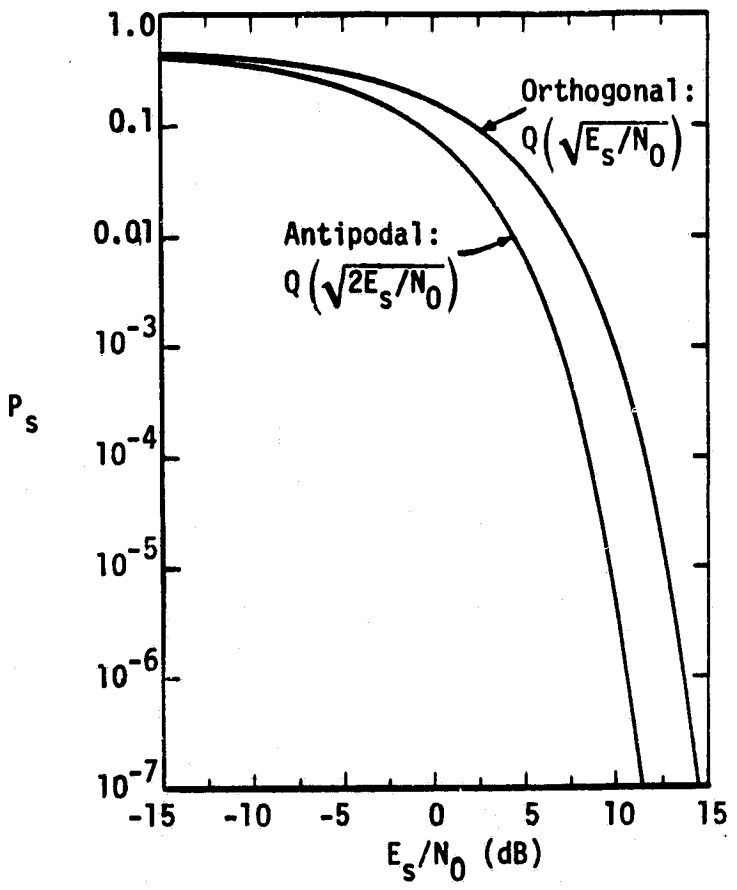


Figure 3.84. Probability of Error for Binary Antipodal and Orthogonal Signaling with Equally Likely Messages

The phase noise spectral density is zero outside of this frequency range. The frequency cutoff f_ℓ may be selected as a frequency below the fundamental frequency of the longest observation period. Notice that the low-frequency cutoff is necessary if $G_\phi(f)$ is to have finite power. However, much of the following discussion is meaningful in the limiting case as $f_\ell \to 0$ and, indeed, some noise spectral effects have been found at as low a frequency as one cycle per year.

The variance of $\phi(t)$ at the output of the carrier-tracking loop is computed as

$$\sigma_\phi^2 = \frac{1}{2}\int_0^\infty |1-H(j2\pi f)|^2 G_\phi(f)\, df \,, \tag{3-240}$$

where $H(j2\pi f)$ is the two-sided closed-loop transfer function of the carrier-tracking loop. Assuming that $\phi(t)$ is a zero-mean Gaussian process, then

$$p(\phi) = \frac{1}{\sqrt{2\pi}\,\sigma_\phi} \exp\left(-\phi^2/2\sigma_\phi^2\right). \tag{3-241}$$

For PSK modulation with a phase error ϕ, the energy per symbol is $P_s \cos^2 \phi$. Therefore, from (3-235),

$$P_s(\phi) = Q\left(\sqrt{2E_s \cos^2 \phi/N_0}\right) \tag{3-242}$$

and the average probability of error is computed by averaging by the probability density of ϕ; that is,

$$P_s = \int_{-\infty}^{\infty} Q\left(\sqrt{2E_s \cos^2 \phi/N_0}\right) \frac{1}{\sqrt{2\pi}\,\sigma_\phi} \exp\left(-\phi^2/2\sigma_\phi^2\right) d\phi \,. \tag{3-243}$$

Figure 3.85 presents a set of probability of symbol error curves as a function of E_s/N_0 for various values of σ_ϕ in degrees. At $P_s = 10^{-4}$ and σ_ϕ less than 10°, L_{OSC} is less than 0.2 dB. The value of L_{OSC} can be computed for other modulation types by computing

$$E_S = P_D/R_S \,, \tag{3-244}$$

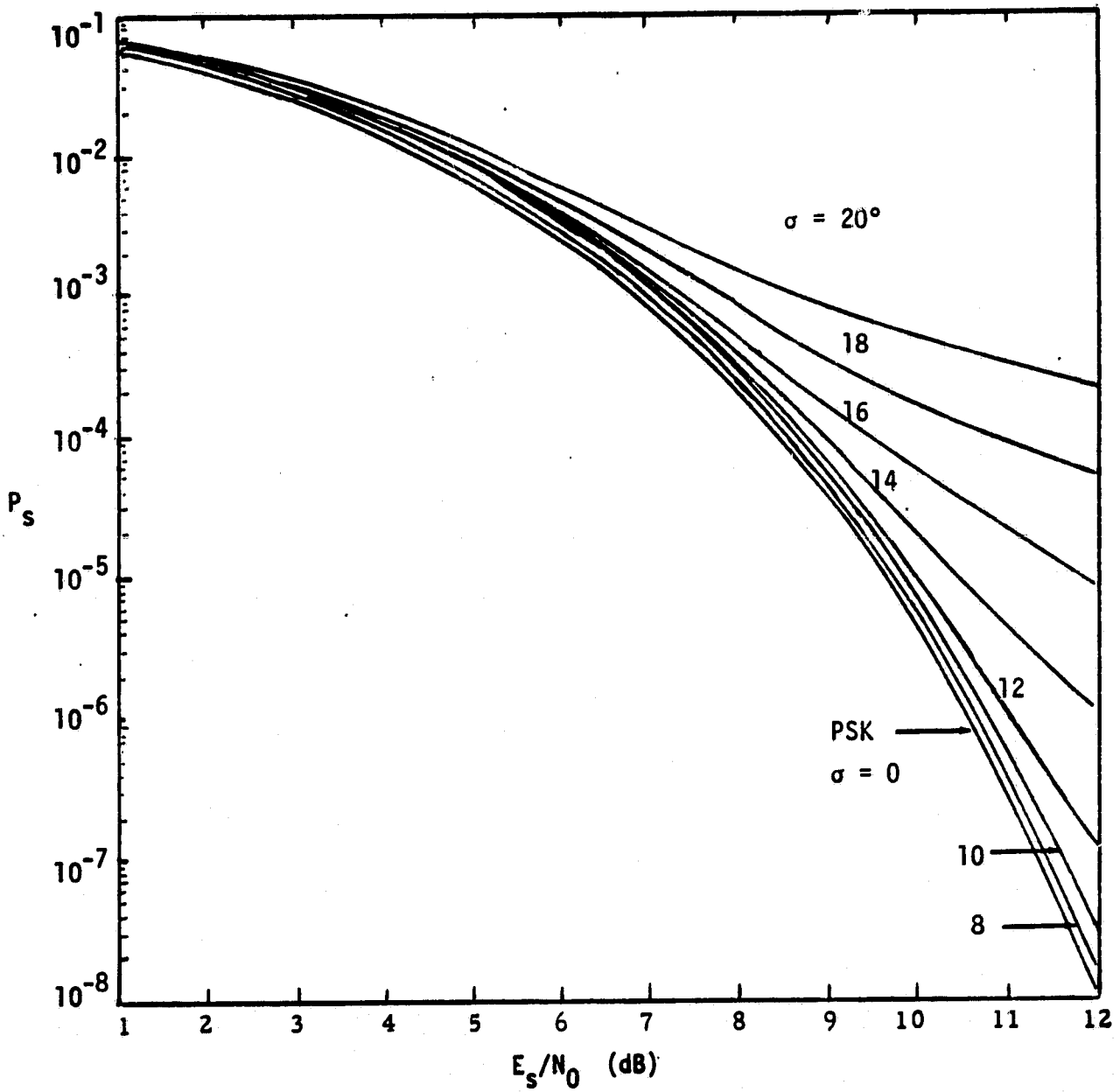


Figure 3.85. Symbol Error Probability due to Phase Noise

where P_D is the data power in the baseband channel, as given in Table 3.3 and subsection 3.2.5.

The PN despreading losses are caused by filtering the PN signal and imperfect tracking of the PN signal. The first contribution is denoted L_{PN} and the second contribution is denoted L_{COR}.

Almost any communication channel required bandpass filtering in order to reduce received interference and noise. This filtering causes a certain amount of signal distortion and intersymbol interference, thereby degrading the system noise performance. For a filtered signal with parabolic phase distortion, L_{PN} can be computed by

$$L_{PN} = \left| \frac{1}{2\pi f_c} \int_{-\infty}^{\infty} \left[\frac{\sin(\omega/2f_c)}{\omega/2f_c} \right] |H(j\omega)|^2 \exp\left(jD\omega^2\right) d\omega \right| , \qquad (3\text{-}245)$$

where f_c is the PN chip rate, $H(j\omega)$ is the filter transfer function, and D is the phase distortion in radians at the first null of the PN spectrum (i.e., $\omega = 2\pi f_c$). Figure 3.86 plots L_{PN} for an ideal rectangular bandpass filter of bandwidth 2B. The loss L_{PN} can be seen to become significant when the phase distortion is greater than one radian. Note that L_{PN} can increase as the bandwidth becomes wider due to phase cancellation when the phase distortion is present.

The PN despreader imperfect tracking causes an effective signal loss at the punctual correlator (i.e., the communication signal channel). This signal loss is defined as L_{COR} and is given by [35]:

$$L_{COR} = 10 \log_{10}\left(1 - 1.6\, \sigma_\gamma\right) , \qquad (3\text{-}246)$$

where σ_γ is the normalized RMS tracking error given in Figure 3.77 and 3.78 for the DLL and in Figure 3.82 for the TDL.

The waveform distortion loss L_{WP} is typically due to filtering. The filtering may be at the transmitter or the receiver, or both. The transmission filter is usually employed to limit the interference with adjacent communication channels. The receiver filter is used to reject interference from adjacent channels and keep RF front-end noise from overloading the demodulator circuits. The filtering of the signal

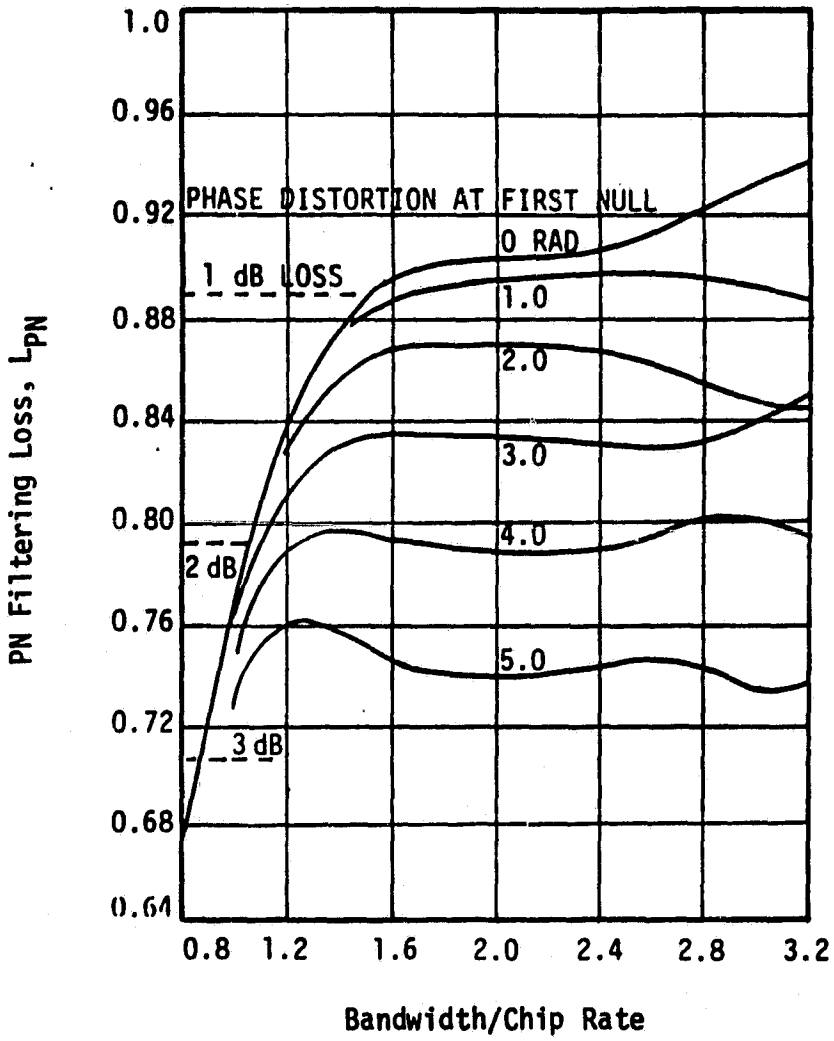


Figure 3.86. PN Filtering Loss Due to Phase Distortion and Finite Bandwidth

causes filter distortion and intersymbol interference. A filter at the transmitter causes the filter distortion, but it occurs in the link before the introduction of the noise. If the same filter were placed in the link after the addition of noise to the signal, it would serve to filter out some of the noise prior to detection and thereby tend to compensate somewhat for distorting the signal. Figure 3.87 compares the use of a symmetrical bandwidth-limiting filter acting as either a transmission filter or a predetection filter. The filter chosen for this comparison is a five-pole Chebyshev type with 0.1 dB ripple in the passband. The data filter is of the integrate-and-dump type.

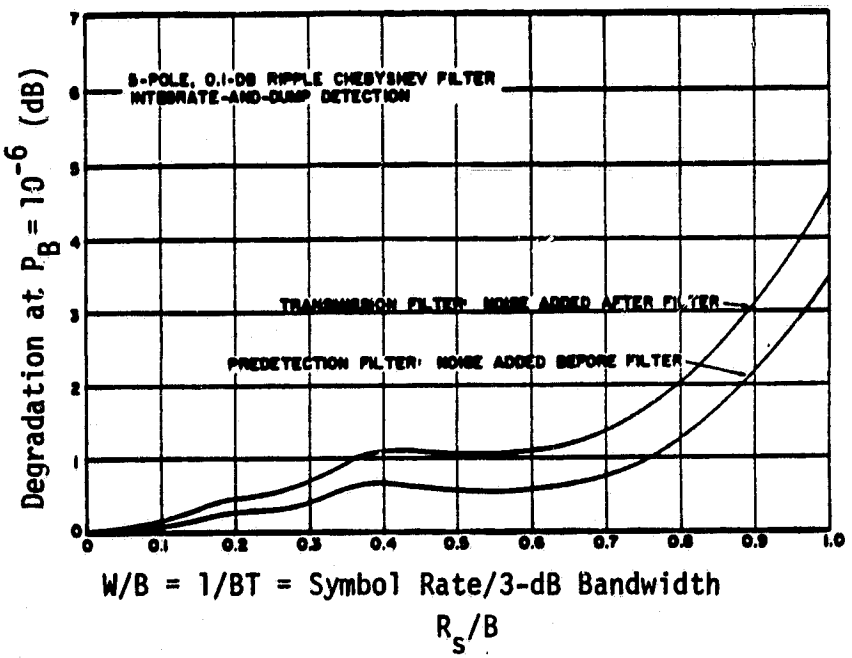


Figure 3-87. Comparison of Transmission and Predetection Filtering

Curves of the resulting degradation for both applications are compared in Figure 3.87 as a function of R_S/B, where B is the 3-dB RF bandwidth of the filter [36]. As Figure 3.87 indicates, the bandwidth-limiting degradation is less for a filter used as a predetection filter than for one used as a transmission filter.

The lower loss is due strictly to noise rejection of the predetection filter. With the use of other data filters (e.g., the two-pole Butterworth) instead of the integrate-and-dump, the improvements would be less significant.

Using an integrate-and-dump filter, L_{WF} increases slowly at first with R_s/B to about the vicinity of $R_s/B = 0.5$, beyond which the loss grows rapidly as the filter bandwidth is decreased further. At the point $R_s/B = 0.5$, the filter bandwidth is about twice the symbol rate, which corresponds to passing only the major lobe of the signal of the PSK signal spectrum. As the filter bandwidth shrinks inside the width of the major spectral lobe, degradation rapidly increases. Clearly then, limiting the channel bandwidth to the symbol rate is out of the question. To avoid excessive degradation for this class of filters, the 3 dB RF filter bandwidth should be at least twice the symbol rate ($R_s/B \leq 0.5$). At the point $R_s/B = 0$, the bandwidth-limiting degradation is zero with the integrate-and-dump filter. Figure 3.87 shows L_{WF} for $P_s = 10^{-6}$ (P_B if there is no error correction coding). The degradation is typically smaller for larger P_s, but Figure 3.87 can be used as a worst-case L_{WF}.

The carrier phase noise loss L_ϕ is computed very similarly to L_{OSC}. Using $P_s(\phi)$ in (3-242) and $P(\phi)$ in (3-189), then

$$P_s = \int_{-\pi}^{\pi} Q\left(\sqrt{2E_s \cos^2 \phi/N_0}\right) \frac{\exp(\eta \cos \phi)}{2\pi I_0(\eta)} d\phi , \qquad (3\text{-}247)$$

where η is the SNR in B_L and, for a PLL, is related to the variance of the phase error σ_ϕ by

$$\eta = \frac{1}{\sigma_\phi^2} \qquad (3\text{-}248)$$

and, for a suppressed carrier-tracking loop, by

$$\eta = \frac{4}{\sigma_{2\phi}^2 L_{SQ}} \qquad (3\text{-}249)$$

The loss L_ϕ is then found by solving

$$Q\left(\sqrt{2L_\phi E_s/N_0}\right) = \int_{-\pi}^{\pi} Q\left(\sqrt{2E_s \cos^2\phi/N_0}\right) \frac{\exp(\eta \cos\phi)}{2\pi I_0(\eta)} d\phi . \quad (3\text{-}250)$$

Figure 3.88 plots L_ϕ versus η as a function of E_s/N_0. Note that L_ϕ is greater for the same value of η if error-correction coding is employed. Figure 3.89 plots L_ϕ versus E_b/N_0 as a function of η for convolutionally encoded/Viterbi-decoded data with rate 1/2, constraint length 7, and eight-level quantization (see subsections 3.2.2 and 3.2.14).

3.2.11.2 FM demodulation

An idealized FM receiver is shown in Figure 3.92. A frequency-modulated signal of transmission bandwidth B_T (Hz) is first passed through an ideal limiter which removes all amplitude variations. The limiter output, after filtering, goes to the discriminator which is assumed to give an output directly proportional to the instantaneous frequency of the signal, then to an ideal lowpass filter of bandwidth B Hz ($B < B_T/2$). B is the maximum bandwidth of the actual information signal f(t) being transmitted. The unmodulated carrier plus noise is of the form

$$\begin{aligned} v(t) &= A_c \cos\omega_0 t + n(t) \\ &= (A_c + x)\cos\omega_0 t - y(t)\sin\omega_0 t \\ &= r(t)\cos\omega_0 t + \phi(t) . \end{aligned} \quad (3\text{-}251)$$

In this unmodulated carrier case, it is apparent that $\phi(t)$ represents the noise at the discriminator output. In particular, since

$$\phi = \tan^{-1}\left(\frac{y}{x + A_c}\right) , \quad (3\text{-}252)$$

where x and y are the inphase and quadrature noise components, respectively, the discriminator output is given by

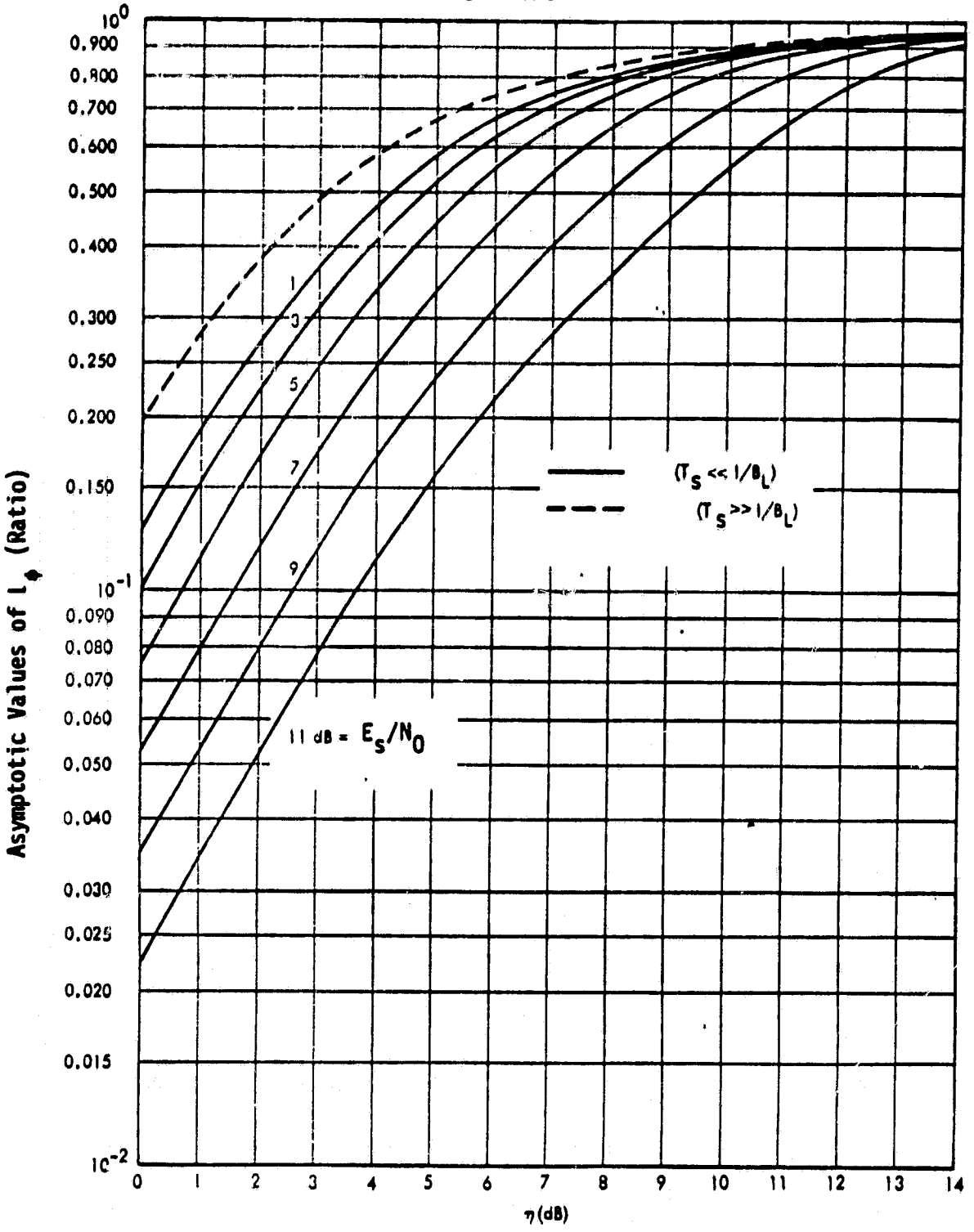


Figure 3.88. Asymptotic Values of L_ϕ Versus SNR in B_L for Uncoded Data

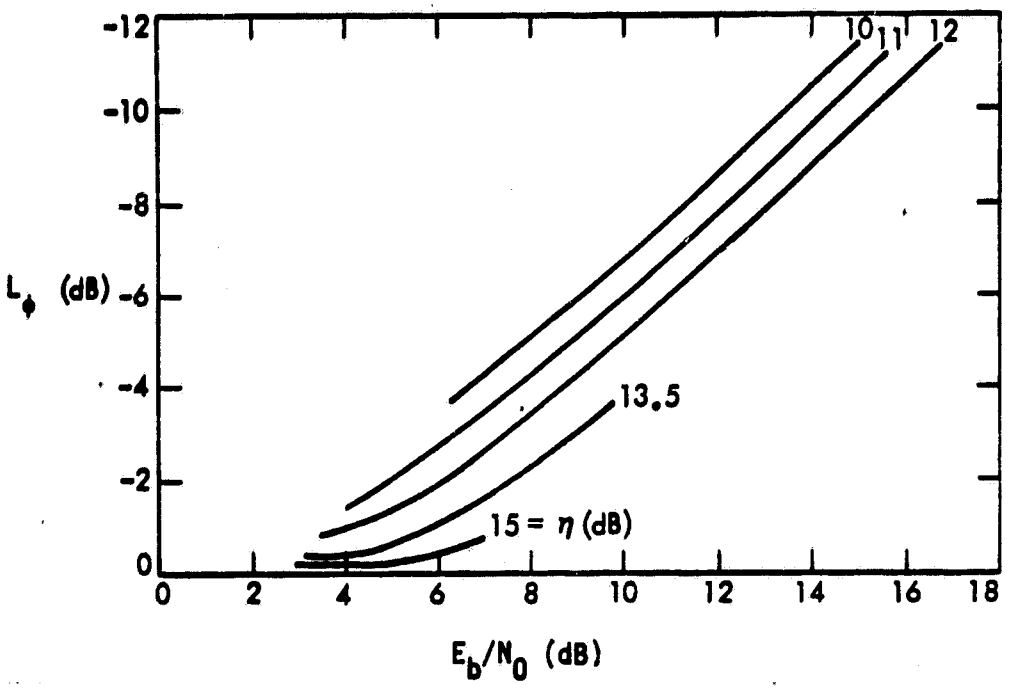


Figure 3.89. Value of L_ϕ Versus E_b/N_0 With SNR in B_L as a Parameter for Rate 1/2, k = 7, Q = 8 Convolutionally Encoded/Viterbi-Decoded Data

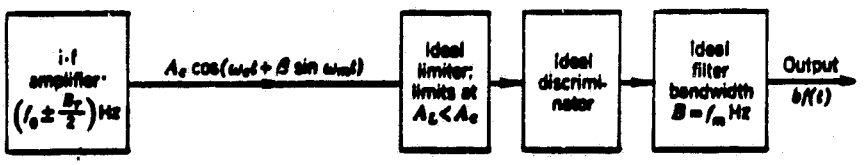


Figure 3.90. Idealized FM Receiver Block Diagram

$$\dot{\Phi} = \frac{(x+A_c)\dot{y} - y\dot{x}}{y^2 + (x+A_c)^2} \quad . \tag{3-253}$$

For large C/N_0, this expression simplifies to

$$\dot{\Phi} = \frac{1}{A_c}\dot{y} \quad . \tag{3-254}$$

The discriminator output noise is thus proportional to the time derivative of the quadrature noise term y(t).

Since the discriminator is followed by a lowpass filter, the spectral density of the discriminator output noise $\dot{\Phi}$ must be found in order to take the effect of the filter into account. To do this, it is simply noted that differentiation is a linear operation. Hence, (3-254) indicates that $\dot{\Phi}$ may be considered the response at the output of a (linear) differentiator H( ) with y applied at the input. Thus,

$$S_{\dot{\Phi}_n}(f) = |H(\omega)|^2 S_y(f) \quad . \tag{3-255}$$

But differentiation of a time function corresponds to multiplication of its Fourier transform by $j\omega$. Then $H(\omega) = j\omega/A_c$, $|H(\omega)|^2 = \omega^2/A_c^2$, and

$$S_{\dot{\Phi}_n}(f) = \frac{\omega^2}{A_c^2} S_y(f) \quad . \tag{3-256}$$

As an example, first consider a rectangular IF spectrum as shown in Figure 3.91(a) where $S'(f) = S_y(f) = N_0$, a constant. The two-sided spectral density of the detected output noise is shown in Figure 3.91(b). The detected output noise can be expressed in terms of the input carrier-to-noise ratio $(C/N)_i = P_R/N_0B_i$ as

$$S_{\dot{\Phi}_n}(f) = \frac{2\pi^2}{B_i} \frac{f^2}{(C/N)_i} ; \qquad -B_i/2 < f < B_i/2 \quad . \tag{3-257}$$

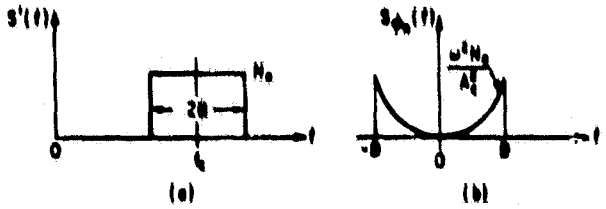


Figure 3.91. FM Noise Spectrum Input, Rectangular IF Spectrum:
(a) Spectrum of Input IF Noise
(b) Spectrum of Detected Output Noise

Assuming an ideal lowpass filter of cut-off frequency B_0 at the FM detector output, the total output noise power is

$$(N)_0 = k_d^2 \int_{-B_0}^{B_0} S_{\phi_n}(f)\, df = k_d^2 \frac{4}{3} \frac{\pi^2}{B_1} \frac{B_0^3}{(C/N)_1} = \left(\frac{k_d}{A_c}\right)^2 (2\pi)^2 \frac{4N_0}{3} B_0^3 \,. \tag{3-258}$$

For a Gaussian IF spectrum with a 3-dB bandwidth equal to 2B, the one-sided spectral density is given by

$$S'(f) = \frac{N e^{-(f-f_c)^2/2f_0^2}}{\sqrt{2\pi}\, f_0} \,, \tag{3-259}$$

where the parameter $f_0 = 1.7B_1$ and N equals the mean noise power at the IF output. The two-sided noise spectral density at the output of the FM detectors is then

$$S_{\phi_n}(f) = \frac{\omega^2}{A_c^2} \frac{N e^{-f^2/2f_0^2}}{\sqrt{2\pi}\, f_0} \frac{2\pi^2}{\sqrt{2\pi}\, f_0} \frac{f^2 e^{-f^2/2f_0^2}}{(C/N)_1} \tag{3-260a}$$

and

$$S_{\dot{\phi}_n}(f) = \frac{0.47\,\pi^2}{B_i}\,\frac{f^2\,e^{-f^2/0.7B_i}}{(C/N)_i} \quad . \tag{3-260b}$$

It should be noted here that, in the region where $f \ll B_i$, the spectral distribution for the Gaussian IF case is very nearly equal to that of the rectangular IF case, as illustrated in Figure 3.92.

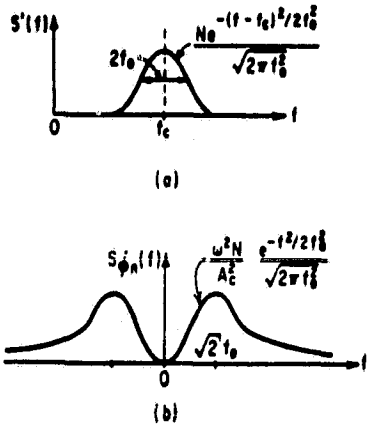


Figure 3.92. FM Noise Spectrum Input, Gaussian IF Spectrum:
(a) Spectrum of Input IF Noise
(b) Spectrum of Detected Output Noise

Assuming again an ideal lowpass filter at the detector output, the total output noise power for $f \ll B_i$ is

$$(N)_0 = k_d^2 \int_{-B_0}^{B_0} S_{\dot{\phi}_n}(f)\,df = k_d^2\,\frac{4}{3}(0.94)\,\frac{\pi^2}{B_i}\,\frac{B_0^3}{(C/N)_i} \quad . \tag{3-261}$$

The signal-to-noise improvement ratio for the high $(C/N)_i$ and rectangular IF spectrum is found as follows. As before, the $(S/N)_0$ is defined as the ratio of the mean-squared output signal voltage (in the absence of noise) to the mean-squared noise voltage at the output of the ideal FM detector (in the absence of signal):

$$S_0 = \frac{k_d^2(\Delta\Omega)^2}{2}, \tag{3-262}$$

where k_d is the constant of the discriminator and $\Delta\Omega$ is the peak angular frequency deviation. Hence the output SNR for high $(C/N)_i$ and rectangular IF spectrum is

$$\left(\frac{S}{N}\right)_0 = \frac{3}{2}\beta^2\left(\frac{B_i}{N_0}\right)\left(\frac{C}{N}\right)_i = \frac{3}{2}\beta^2\left(\frac{P_R}{N_0 B_0}\right). \tag{3-263}$$

For large β, the transmission bandwidth $B_i = 2\Delta f$ and (3-263) reduces to

$$\left(\frac{S}{N}\right)_0 = 3\left(\frac{B_i}{2B_0}\right)^3. \tag{3-264}$$

A similar equation can be derived for the $(S/N)_0$ for a particular data-modulated subcarrier. Consider the noise contribution due to a particular channel in the baseband $(0 - f_m)$ occupying the frequency range from f_1 to f_2. Thus,

$$N_0(f_1 - f_2) = \left(\frac{k_d}{A_c}\right)^2 2\int_{f_1}^{f_2} f^2 S'(f_c + f)\, df. \tag{3-265}$$

For white noise, this becomes

$$N_0(f_1 - f_2) = \left(\frac{k_d}{A_c}\right)^2 2N_0 \frac{f_2^3 - f_1^3}{3}. \tag{3-266}$$

Let the channel bandwidth $B_c = f_2 - f_1$ and the subcarrier frequency $f_s = (f_1 + f_2)/2$. In practice, the channel bandwidth $B_c \ll f_{mj}$; therefore,

$$\frac{f_2^3 - f_1^3}{3} = B_c f_s^2 \tag{3-267}$$

and the channel noise power is

$$N_0(f_1 - f_2) = \left(\frac{k_d}{A_c}\right)^2 2N_0B_c f_s^2 . \qquad (3\text{-}268)$$

It should be noted that the noise contribution of a particular channel varies with the position of the channel in the baseband, being greater for the high-frequency channels. The problem of noise equalization in a multichannel system is therefore of great practical importance in the design of communication systems.

The $(S/N)_0$ in a particular channel is readily found to be

$$\left(\frac{S}{N}\right)_0 = \frac{1}{2}\left(\frac{C}{N_i}\right)\left(\frac{B_i}{B_c}\right)\left(\frac{\Delta f_m}{f_s}\right)^2 = \frac{1}{2}\left(\frac{P_R}{N_0B_c}\right)\left(\frac{\Delta f_m}{f_s}\right)^2 , \qquad (3\text{-}269)$$

where Δf_m is the peak channel-frequency deviation. It should be noted that the IF bandwidth B_i does not affect $(S/N)_0$. However, the IF bandwidth does play a very important role in setting the threshold level. Figure 3.93 shows a comparison of AM and FM receivers for $\beta = 4$ and $\beta = 1$. Note that, for $\beta = 4$, the FM output $(S/N)_0$ deteriorates rapidly for P_R/N_0B_i less than 13 dB. For $P_R/N_0B_i > 15$ dB, however, the FM system shows an improvement of 14 dB. For $\beta = 4$, the theoretical improvement should be $3\beta^2 = 48$, or 17 dB. For $\beta = 1$, the threshold level is experimentally found to occur at 2 dB. Above this value of AM carrier-to-noise ratio, the FM improvement over AM is 3 dB. The theoretical improvement would be expected to be $3\beta^2 = 3$, or 4.8 dB. It should be emphasized here that the above equations for $(S/N)_0$ hold provided $(C/N)_i$ is greater than a specified threshold level; thus, the IF bandwidth determines the service range of the communication systems, and consequently, it should be designed as narrow as possible in keeping with the allowable distortion requirements.

The maximum operating range of any communication system is determined by the location at which the carrier falls below the improvement threshold. When this occurs, even in ordinary double-sideband amplitude modulation, there is a relatively sudden large

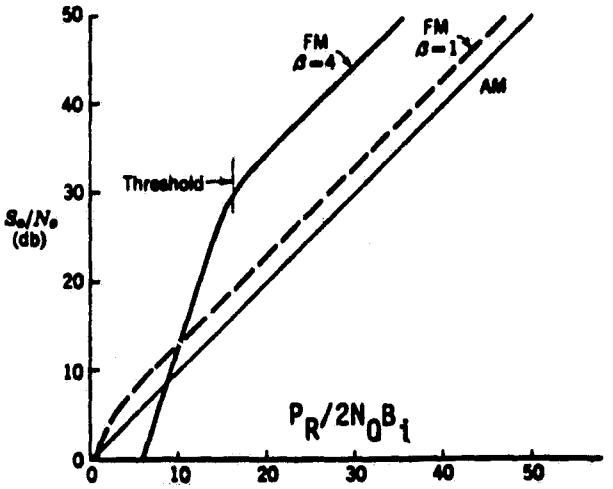


Figure 3.93. Measured Characteristics, FM and AM Receivers

rise of the noise level which effectively blankets the signal. This fact should be considered in the design of communication systems. If the $(C/N)_i$ ratio is much higher than the improvement threshold, the bandwidth of the IF prior to the limiter can be increased many-fold, and as long as the bandwidth is again narrowed to its optimum value by the audio (or video) amplifier, there will be no loss in $(S/N)_0$ ratio in the final output. However, in case of the $(C/N)_i$ ratio near the improvement threshold, widening the predetection (IF) bandwidth to the extent that the $(C/N)_i$ ratio falls into the range of the improvement threshold will cause a rise in noise which cannot be eliminated by narrowing the bandwidth of the audio amplifier. This threshold level therefore determines the point at which the communication system fails or determines the maximum operating range.

The FM performance presented above is degraded for narrow IF filter and postdetection bandwidths by causing signal distortion. Thus, there is a limit as to how narrow the IF filter bandwidth can be made and, hence, how low the threshold can be made. The loss due to signal distortion of the IF and postdetection filters is denoted

L_{SF}. While there are analytical techniques that aid in the calculation of L_{SF}, typically, the communication system is tested to determine L_{SF} since the effect of signal distortion depends a great deal on the type of baseband signal that was modulated.

In subsection 3.2.2.2, preemphasis (or predistortion) by a filter at the transmitter was presented as a way of improving the $(S/N)_0$ at the receiver by using deemphasis. The deemphasis network discussed in subsection 3.2.2.2 was

$$H_d(f) = \frac{1}{1 + j2\pi f\tau} \quad , \tag{3-270}$$

where τ is the time constant of the deemphasis filter. With deemphasis, the spectral distribution of the noise power is modified by $|H(j2\pi f)|^2$, and the total mean noise power at the output of the ideal lowpass filter of bandwidth B_0 is given by

$$(N)_{0,d} = \left(\frac{k_d}{A_c}\right)^2 N_0 \int_{-B_0}^{B_0} \frac{f^2\, df}{1 + 4\pi^2 f^2 \tau^2} = \left(\frac{k_d}{A_c}\right)^2 2N_0 \left(B_0 - \frac{1}{2\pi\tau} \tan^{-1} \left(2\pi B_0 \tau\right)\right) . \tag{3-271}$$

The improvement factor σ_{FM} is given by

$$\sigma_{FM} = \frac{(N)_0}{(N)_{0,d}} = \frac{\left(2\pi B_0 \tau\right)^3}{3\left(2\pi B_0 \tau - \tan^{-1} \left(2\pi B_0 \tau\right)\right)} . \tag{3-272}$$

For narrowband FM, $\sigma_{FM} \to 1$. For wideband, B_0 is large, and

$$\sigma_{FM} \to \frac{\left(2\pi B_0 \tau\right)^2}{3} . \tag{3-273}$$

Thus, for small B_0, the improvement factor due to deemphasis reduces to unity. For large B_0, the improvement factor is proportional to B_0^2. Finally, the mean S/N output power ratio for FM with deemphasis is given by

$$\left(\frac{S}{N}\right)_{0,d} = \sigma_{FM}\left(\frac{S}{N}\right)_0 = \frac{3}{2}\sigma_{FM}\beta^2\left(\frac{P_R}{N_0 B_i}\right). \quad (3\text{-}274)$$

With $\tau = 75$ μs and $B_0 = 15$ kHz, $\sigma_{FM} = 20$, or the S/N improvement due to deemphasis is 13 dB. For $B_0 = 21$ kHz, the improvement due to deemphasis is 16 dB.

Demodulation of FSK signals can also be performed by an FM demodulator. In subsection 3.2.3.3, FSK signals generated by FM were presented. The frequency discriminator demodulation approach is of general interest because the same modulation/demodulation equipment used for transmission of binary FSK data can then be used for transmission of analog data.

The analysis of systems employing discriminator detection of FSK is complicated by (1) the fact that it is very difficut to account for the effects of signal distortion due to bandpass filtering and (2) the presence of non-Gaussian noise at the discriminator output and the resulting difficulties associated with computation of error probabilities.

Many studies of error probabilities in noncoherent FSK systems have been performed. However, these studies assumed a sufficiently broad bandpass filter in the system for negligible distortion of the FSK signal. In fact, it is possible to make a favorable trade-off between signal distortion and input noise reduction, so these results do not indicate error rate performance of the "optimum" FSK system employing discriminator detection.

Bennett and Salz [37] determined error rates for a binary FSK system, taking into account the effects of distortion due to a bandpass filter. However, their receiver model did not include a data-matched filter after the discriminator.

Tjhung and Wittke [38] evaluated error probabilities for a binary FSK system (utilizing discriminator detection), taking into account the effects of both a bandpass filter and a data-matched filter. In order to account for the FM signal distortion due to bandpass filtering, a periodic modulating signal (a 30-bit pseudorandom sequence) was used. The particular sequence used was 11000 00101 10111 00111 11010 01000, and it was determined that the FM spectum for this signal

was a good approximation to the spectrum for FM by a random binary signal. The predetection bandpass filter was assumed to have a symmetrical passband and a linear phase characteristic. Results were obtained for two filter models--rectangular passband and Gaussian passband. Using Rice's click theory of FM noise, Tjhung and Wittke computed overall error probabilities by taking the average of the error probabilities for the individual bits. A number of error rate curves were calculated as functions of E_b/N_0 (for the unfiltered FM signal), with $2\Delta f$ and BT (the product of the filter bandwidth and the bit period or, alternately, the ratio of the filter bandwidth to the bit rate) as parameters. The simulation results indicate that, for a given filter type and bit rate, there is a bandwidth B and a frequency deviation Δf that minimize the probability of error. Tables 3.5 and 3.6 were provided by Tjhung and Wittke to allow some degree of precision in determining the optimum values of these parameters for an error probability of 10^{-4}. It can be seen from these tables that, for both the Gaussian and rectangular bandpass filters, a value of $2\Delta f = 0.7R$ is best in that it requires the smallest value of E_b/N_0 to achieve a 10^{-4} bit error probability. The optimum IF bandwidth for $P_e = 10^{-4}$ is seen to be 1.2 times the bit rate for the rectangular bandpass filter and 1.0 times the bit rate for the Gaussian filter.

It is very significant that (from Table 3.5) using discriminator detection of binary FSK, it is apparently possible to achieve an error probability of 10^{-4} for $E_b/N_0 = 10.65$ dB. This is only 2.25 dB more than is required for coherent PSK and is within 0.1 dB of the best performance achievable using coherent detection of FSK. Thus, the results of Tjhung and Wittke indicate that the performance bound represented by coherent FSK is almost achievable using discriminator detection, given that some discretion is exercised in the choice of frequency deviation and IF filter bandwidth.

Figure 3.96 compares the measured suboptimum FSK system performance (for best-case parameters) with the theoretical performance for coherent PSK and coherent FSK systems [39]. It should be noted that

Table 3.5. E_b/N_0 in dB Required to Achieve a 10^{-4} Bit Error Probability in Binary FSK Systems Employing Discriminator Detection (Rectangular Bandpass Filter)

	BT or B/R				
$2\Delta f$	1.0	1.2	1.6	2.0	3.0
0.5R	12.27	10.95	11.7	12.63	
0.7R	11.28	10.65	11.7	12.23	
1.0R	13.8		13.25	12.8	

Table 3.6. E_b/N_0 in dB Required to Achieve a 10^{-4} Bit Error Probability in Binary FSK Systems Employing Discriminator Detection (Gaussian Bandpass Filter)

	BT or B/R					
$2\Delta f$	0.8	1.0	1.2	1.6	2.0	3.0
0.5R	13.2	12.26	12.08	12.42	13.0	
0.7R	11.09	10.74	11.0	11.73	12.45	14.06
1.0R		12.38	12.23	12.53		

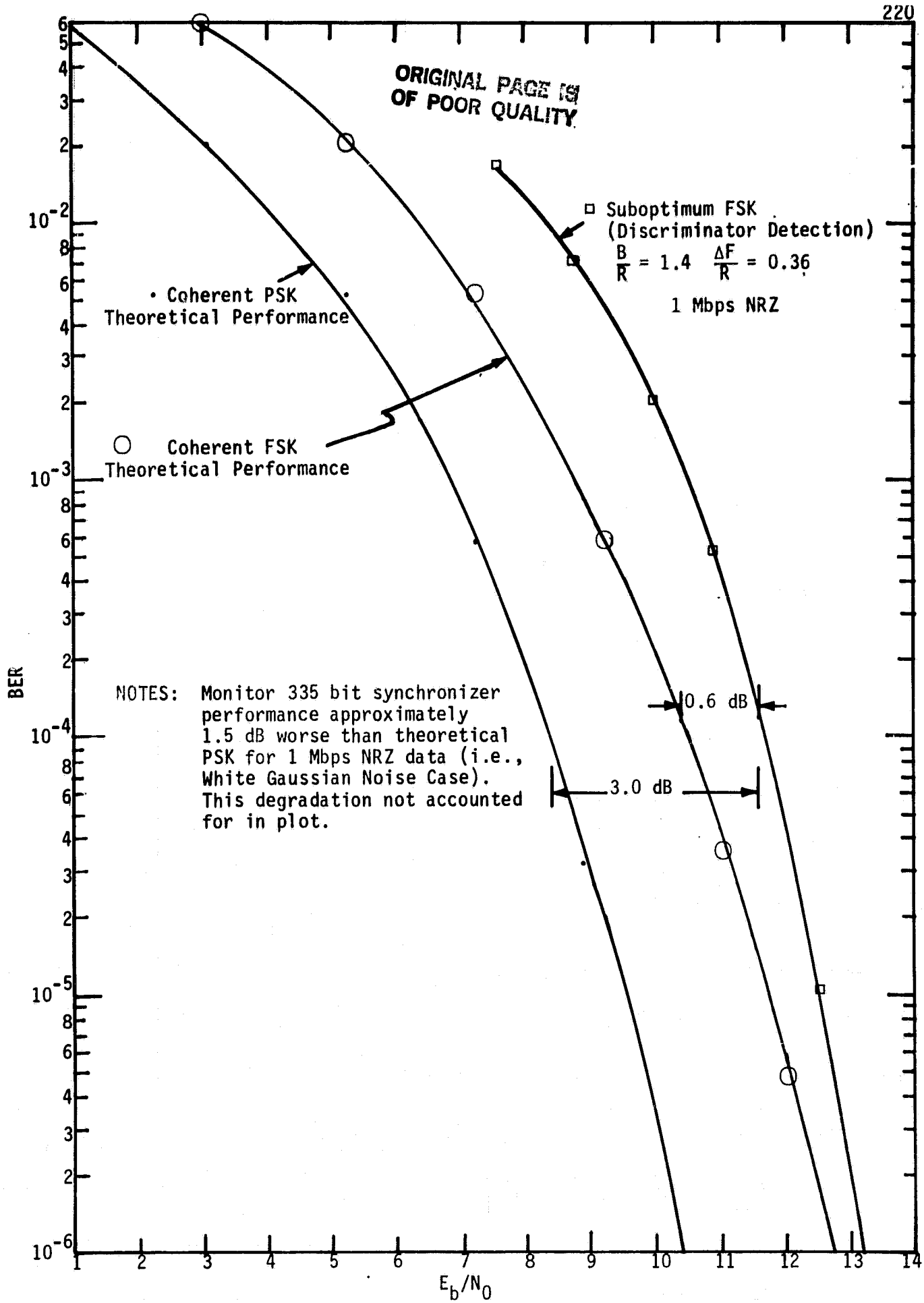


Figure 3.94. Comparison of Measured (Best-Case) Suboptimum FSK System Performance with Coherent PSK and FSK Theoretical Performance (NRZ-L Data Format)

the bit synchronizer degradation is included as part of the measured suboptimum FSK system performance. No attempt was made to adjust the measured data or the theoretical predictions to account for bit synchronizer degradation in the presence of the non-Gaussian, white noise encountered with discriminator detection.

As seen from Figure 3.94, the suboptimum system performance was surprisingly good even without accounting for the degradations in the measured performance due to the monitor bit synchronizer (1.5 dB for white Gaussian noise condition) and in fact is approaching the bound for the theoretical coherent FSK system.

These results agree quite favorably with the simulation results of Tjhung and Wittke discussed previously, and are presented in tabular form in Table 3.7 for comparison with Tables 3.5 and 3.6.

Table 3.7. Measured E_b/N_0 in dB Required to Achieve a 10^{-4} BER in Binary FSK Systems Employing Discriminator Detection (NRZ Data Format, Integrate-and-Dump Bit Detection*)

	BT or B/R			
$2\Delta f$	0.98	1.4	2.7	3.5
0.5R	15.2	13.0	15.7	15.9
0.72R	13.6	11.4	12.8	13.0
1.0R	14.8	12.5	13.2	13.6

*Bit detector performance degradation is 2.5 dB for white Gaussian noise operation.

There are no theoretical or simulation results available for FSK Manchester-coded data system performance under the conditions of suboptimum (discriminator) detection. Therefore, the test results described herein provided the basis for the Shuttle Orbiter FSK system parameter optimization and performance margin calculations. The data is felt to be general enough in nature to be utilized for any FSK system employing discriminator detection and has been tacitly verified by the NRZ system test results discussed previously which were shown to agree very well with earlier theoretical and/or simulation results.

Figure 3.95 shows the comparison of the suboptimum FSK system (discriminator detection) best-case performance with coherent PSK and coherent FSK theoretical performance. It should be noted again that the measured performance data includes the monitor bit synchronizer performance degradation (2.0 dB for Manchester-coded data in the presence of white Gaussian noise). There has been no attempt to adjust either the measured or theoretical curves to account for bit synchronizer degradation when operating in the presence of the nonwhite, non-Gaussian noise encountered at the discriminator output of the FSK system. For completeness, Table 3.8 provides data for the biphase-level FSK system performance similar to that for the NRZ-L system performance provided in Table 3.7.

Table 3.8. Measured E_b/N_0 in dB Required to Achieve a 10^{-4} BER in Binary FSK Systems Employing Discriminator Detection (Manchester Data Format, Integrate-and-Dump Bit Detection*)

	BT or B/R				
$2\Delta f$	1.4	2.7	5.0	10.5	21
1.0R	16.4	14.2	15.4		
1.24R	15.3	13.4	15.0		
2.0R	24.6	15.2	16.3		
2.42R		15.7	16.3	16.4	
4.84R					16.9

*Bit synchronizer performance degradation is 2.0 dB for white Gaussian noise operation.

3.2.12 Subcarrier Demodulation

The losses due to subcarrier demodulation for coherent demodulation are, for the most part, the same as those presented for carrier tracking in subsections 3.2.10 and 3.2.11. The subcarrier demodulation losses for FM and FSK using a limiter/discriminator demodulator are also

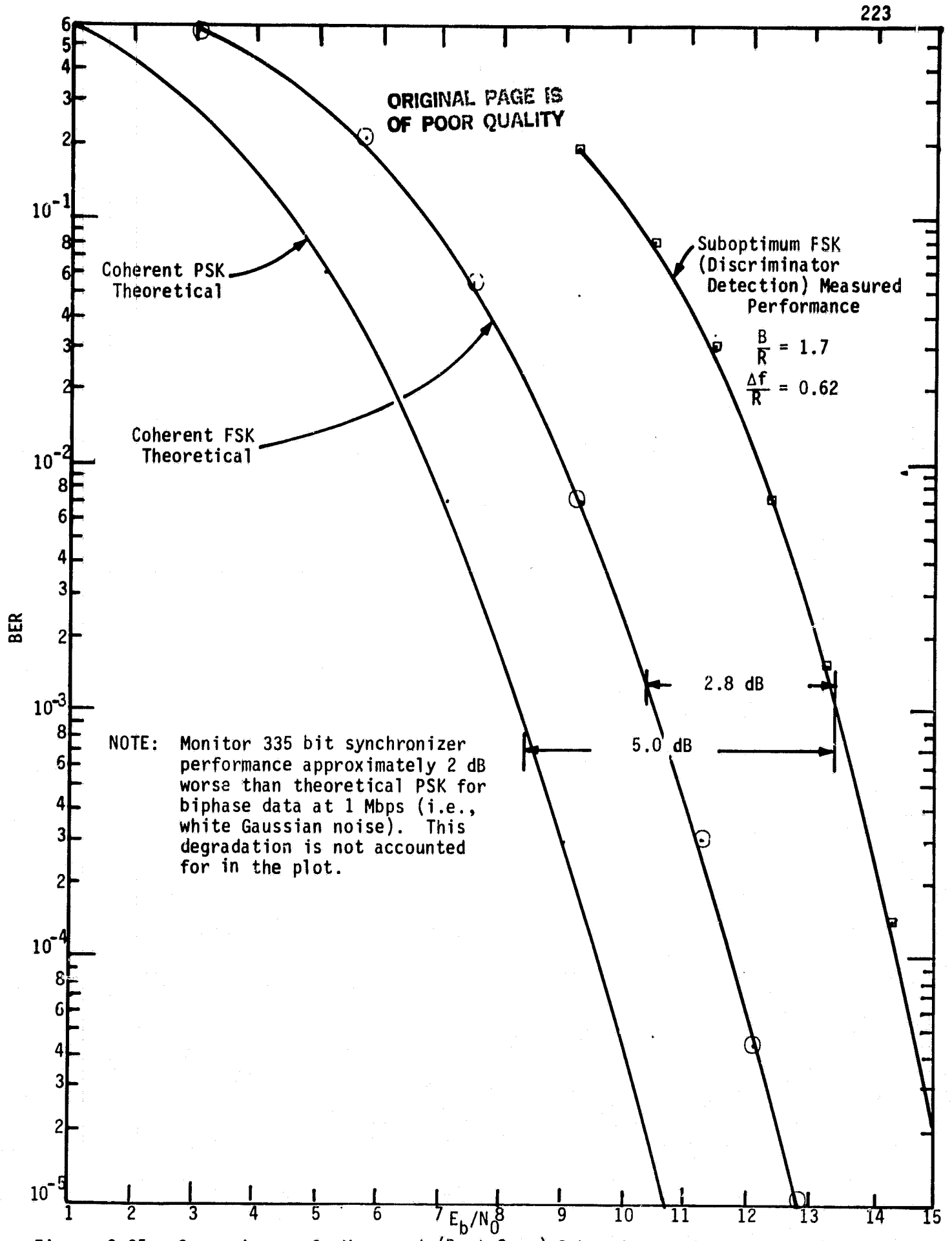


Figure 3.95. Comparison of Measured (Best-Case) Suboptimum FSK System Performance with Coherent PSK and FSK Theoretical Performance (Manchester-Data Format)

the same as the results presented in subsection 3.2.11. Therefore, this subsection relies heavily on the two preceding subsections. In addition, however, demodulation of FSK using a bandpass filter and an envelope detector is treated in this subsection.

3.2.12.1 PSK subcarrier demodulation

When digital data is modulated on a subcarrier, the subcarrier is typically suppressed by using PSK or QPSK modulation. Therefore, the results presented in subsections 3.2.10 and 3.2.11 for suppressed-carrier tracking using a squaring loop, Costas loop, or a DAL are equally valid for subcarrier tracking. However, an important subcarrier tracking topic is the demodulation of the three-channel Orbiter Ku-band return link.

Having examined in subsection 3.2.10 the many ways in which a biphase Costas loop can track an unbalanced QPSK signal, the tracking performance of the subcarrier loop for the three-channel Orbiter Ku-band return link can be analyzed. The primary difference between the subcarrier loop tracking behavior and that of the loops discussed in subsection 3.2.10 is the fact that the inphase and quadrature data modulations which are input to the subcarrier loop are biphase modulated onto quadrature square-wave subcarriers as opposed to sine-wave subcarriers. A secondary difference is the fact that, since the output of the quadrature phase detector of the carrier-tracking loop serves as the input to the subcarrier loop, the performance of the latter depends on the carrier-tracking loop phase error. An analysis which takes both of these differences into account is given in [29]. For simplicity of numerical evaluation, it was convenient to assume perfect carrier tracking. The additional degradation due to the phase-tracking jitter of the carrier loop itself can easily be assessed from the results in [40], wherein the performance of a Costas loop for recovering the carrier from the three-channel quadrature multiplex signal was studied.

Some specific results [29] are summarized herein. In addition to perfect carrier tracking which was already mentioned, the following assumptions were made. The subcarrier frequency is 8.5 MHz, the high

rate modulation $m_1(t)$ in Mode 1 is rate 1/2, constraint length 7, convolutional code (NRZ format) with data rate R_1 = 50 Mbps, and the arm filters in the subcarrier loop are one-pole Butterworth (RC). Under these assumptions, Figures 3.96 and 3.97 illustrate the subcarrier tracking-jitter behavior for a fixed ratio of arm filter noise bandwidth to loop noise bandwidth (B_i/B_L). In particular, Figure 3.96 illustrates the case where $m_2(t)$ is NRZ data at 500 kbps and $m_3(t)$ is Manchester-coded data at 192 kbps, while Figure 3.97 corresponds to the same parameter values as Figure 3.96 with the exception that now $m_2(t)$ is also Manchester coded. In both figures, the RMS tracking jitter, σ_ϕ (in % radians), is plotted versus the ratio of two-sided arm filter noise bandwidth (B_i) to the higher data rate (R_2), with total power-to-noise ratio in the higher data rate bandwidth $(P_T T_2/N_0)$ as a parameter. It is observed that the changes in the subcarrier tracking jitter as a function of B_i/R_2 are more obvious when $m_2(t)$ and $m_3(t)$ are both Manchester codes than when $m_2(t)$ is NRZ and $m_3(t)$ is Manchester. Upon establishing a subcarrier reference signal, the two lower rate modulations, $m_2(t)$ and $m_3(t)$, can then be demodulated.

Evaluation of error probability performance of BPSK, QPSK, and offset (staggered) QPSK receivers has been extensively covered in the literature. While the techniques used there are certainly applicable to demodulation of unbalanced QPSK as on the two lower rate channels of the three-channel signal, the complexity of the evaluation when the ratio of data rates in the two channels is large prompts one to look for a simpler calculation procedure. Indeed, such an approach is possible when the noisy reference loss* is small or, equivalently, the effective SNR in the tracking loop bandwidth is large, i.e., the loop operates in its so-called linear region. Making such an assumption for purposes of error probability performance evaluation is quite reasonable when one realizes that this very same assumption has already been implied in assessing the tracking performance of the subcarrier

*"Noisy reference loss" is defined here as the equivalent increase in signal power required to produce the same error probability as obtainable in a perfectly synchronized system.

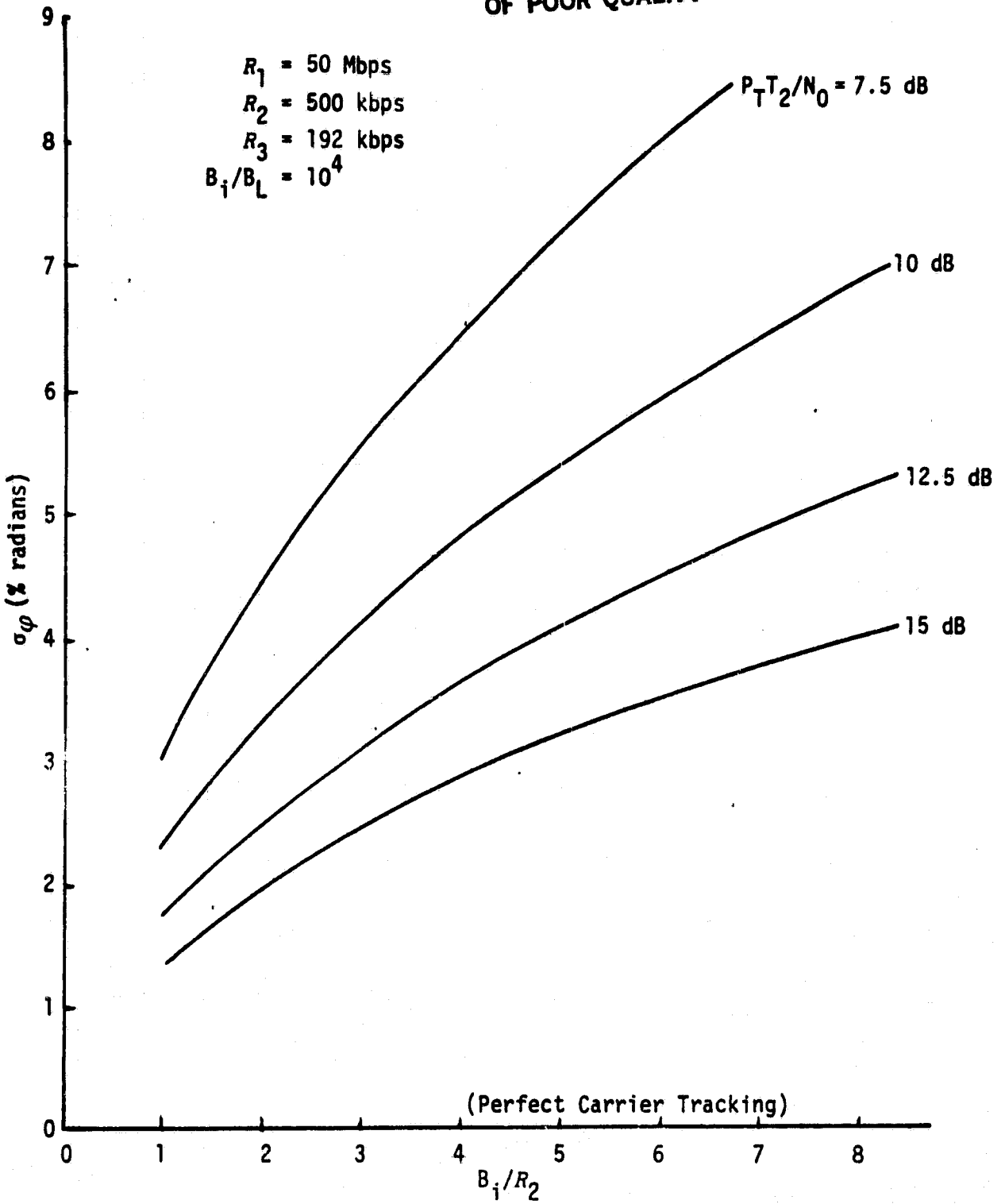


Figure 3.96. Subcarrier-Tracking Jitter Versus Ratio of Arm Filter Bandwidth to High Subcarrier Data Rate R_2; P_TT_2/N_0 is a Parameter; $m_1(t)$ and $m_2(t)$ are NRZ, $m_3(t)$ is Manchester; $R_1 > R_2 > R_3$.

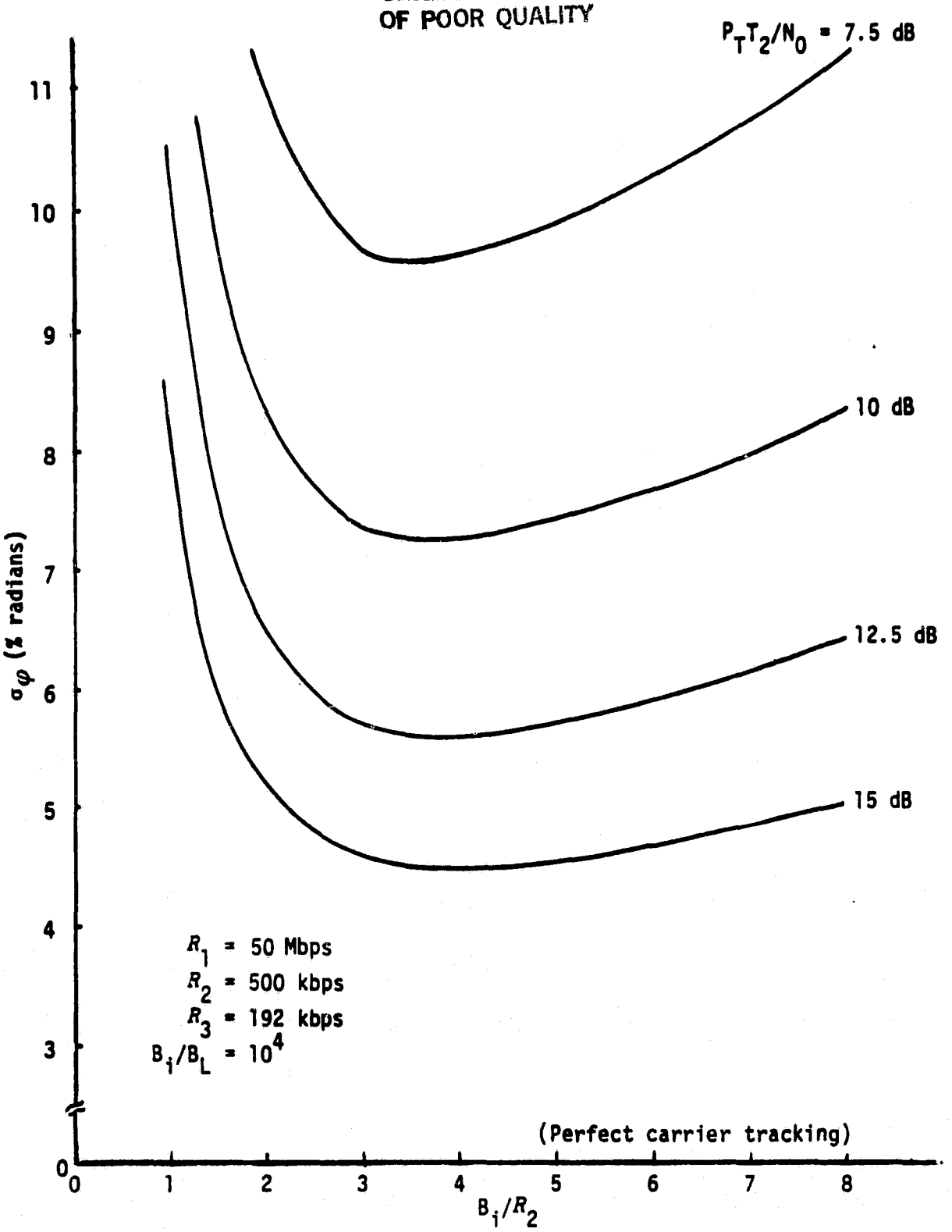


Figure 3.97. Subcarrier-Tracking Jitter Versus Ratio of Arm Filter Bandwidth to High Subcarrier Data Rate R_2; $P_T T_2/N_0$ is a Parameter; $m_1(t)$ is NRZ, $m_2(t)$ and $m_3(t)$ are Manchester codes; $R_1 > R_2 > R_3$

loop. The approach taken is to expand the error probability conditioned on the subcarrier loop phase error ϕ in a power (Maclaurin) series in ϕ and then, keeping only the first few terms of this series, average this conditional error probability over the probability density function (PDF) of ϕ. By doing this, we obtain the additional error probability due to a noisy subcarrier reference as an additive term directly proportional to the mean-squared phase jitter σ_ϕ^2 directly associated with the receiver's subcarrier-tracking loop. In this regard, the results [29] play an important role in assessing this error probability performance. Finally, similar arguments can be advanced to give closed-form results for the noisy reference loss itself.

A detailed discussion of the general problem of assessing the error probability performance of unbalanced QPSK receivers under the abovementioned assumptions is presented in [29]. The key results from this discussion are summarized as follows. In terms of the total SNR R_{T2} ($R_{T2} = P_T T_2/N_0$) in the higher data bandwidth and the transmitted modulation indices η_2 and η_3 defined by

$$R_{T2} = \frac{(P_2 + P_3)T_2}{N_0}; \qquad \eta_i = \frac{P_i}{P_2 + P_3}, \quad i = 2,3 . \tag{3-275}$$

the error probability performance of channels 2 and 3 become

$$P_{E_2} = \frac{1}{2}\operatorname{erfc}\sqrt{R_{T2}\eta_2} + \frac{1}{2}\sqrt{\frac{R_{T2}\eta_2}{\pi}}\exp(-R_{T2}\eta_2)\left[1 + 2R_{T2}\eta_3\overline{\tilde{m}_{32}^2}\right]\sigma_\phi^2$$

$$P_{E_3} = \frac{1}{2}\operatorname{erfc}\sqrt{R_{T2}\gamma_T\eta_3} + \frac{1}{2}\sqrt{\frac{R_{T2}\gamma_T\eta_3}{\pi}}\exp(-R_{T2}\gamma_T\eta_3)\left[1 + 2R_{T2}\eta_2\overline{\tilde{m}_{32}^2}\right]\sigma_\phi^2 , \tag{3-276}$$

where $\gamma_T \triangleq R_2/R_3$ is the ratio of data rates and the normalized mean-squared crosstalk $\overline{\tilde{m}_{32}^2}$ is tabulated in [29] for various combinations of data formats in the two channels. The corresponding expressions

for the noisy reference loss or subcarrier demodulation loss L_{SD} (in decibels) itself are

$$L_{SD2} = \frac{10 \log_{10}\left\{1 + R_2\left[1 + 2R_3\left(\overline{\tilde{m}_{32}^2}/\gamma_T\right)\right]\sigma_\phi^2\right\}}{R_2}$$

$$L_{SD3} = \frac{10 \log_{10}\left\{1 + R_3\left[1 + 2R_2\overline{\tilde{m}_{32}^2}\right]\sigma_\phi^2\right\}}{R_3} \qquad (3\text{-}277)$$

In applying the results of (3-276) and (3-277) to Channels 2 and 3 of the three-channel Orbiter Ku-band return link signal, it is assumed that $m_2(t)$ is NRZ data with a maximum rate $R_2 = 2$ Mbps and $m_3(t)$ is a Manchester-coded data stream at $R_3 = 192$ kbps. Further, it is assumed that the power allocation is chosen so that, for the given data rate ratio $\gamma_T = R_2/R_3 = 10.42$, the SNR's R_2 and R_3 in the two channels are made equal, i.e., both channels operate at the same error rate. Thus, $R_2 = R_3$ implies $P_2T_2 = P_3T_3$, and the modulation indices η_2, η_3 become

$$\eta_2 \triangleq \frac{P_2}{P_2 + P_3} = \frac{\gamma_T}{\gamma_T + 1} = 0.9124$$

$$\eta_3 \triangleq \frac{P_3}{P_2 + P_3} = \frac{1}{\gamma_T + 1} = 0.0876 \; . \qquad (3\text{-}278)$$

Further, it is typical to design the Costas loop bandwidth on the order of $R_2/100$ (or less) since most of the power is in the high rate channel which controls the performance of the tracking loop. Thus, assuming $B_L/R_2 = 0.01$*, Figures 3.98 and 3.99 illustrate B_i/R_2 versus L_{SD2} and L_{SD3} of (3-277) for error probabilities of 10^{-4}, 10^{-5}, and 10^{-6}, corresponding respectively to $R_2 = R_3 = 8.4$, 9.6, and 10.5 dB. Several conclusions may be drawn from these figures. First, the

*Smaller values of B_L/R_2 as would be typical in practical receiver design would yield insignificant losses in L_{SD2} and L_{SD3}.

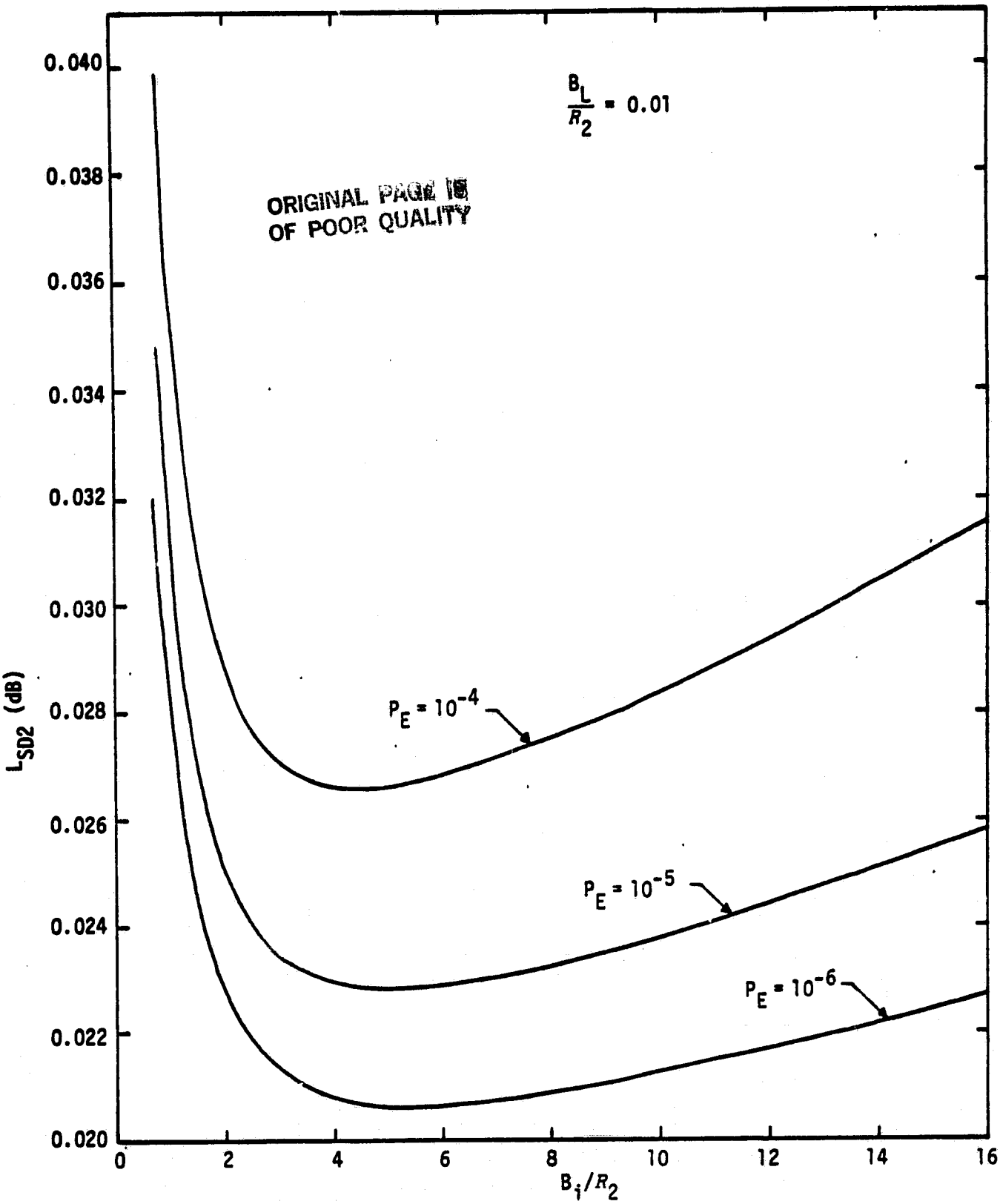


Figure 3.98. Channel 2 Noisy Reference Loss (in dB) Versus the Ratio of Two-Sided Costas Loop Arm Filter Bandwidth to Channel 2 Data Rate; $m_3(t)$ is 192-kbps Manchester-Coded Data, $m_2(t)$ is 2-Mbps NRZ Data

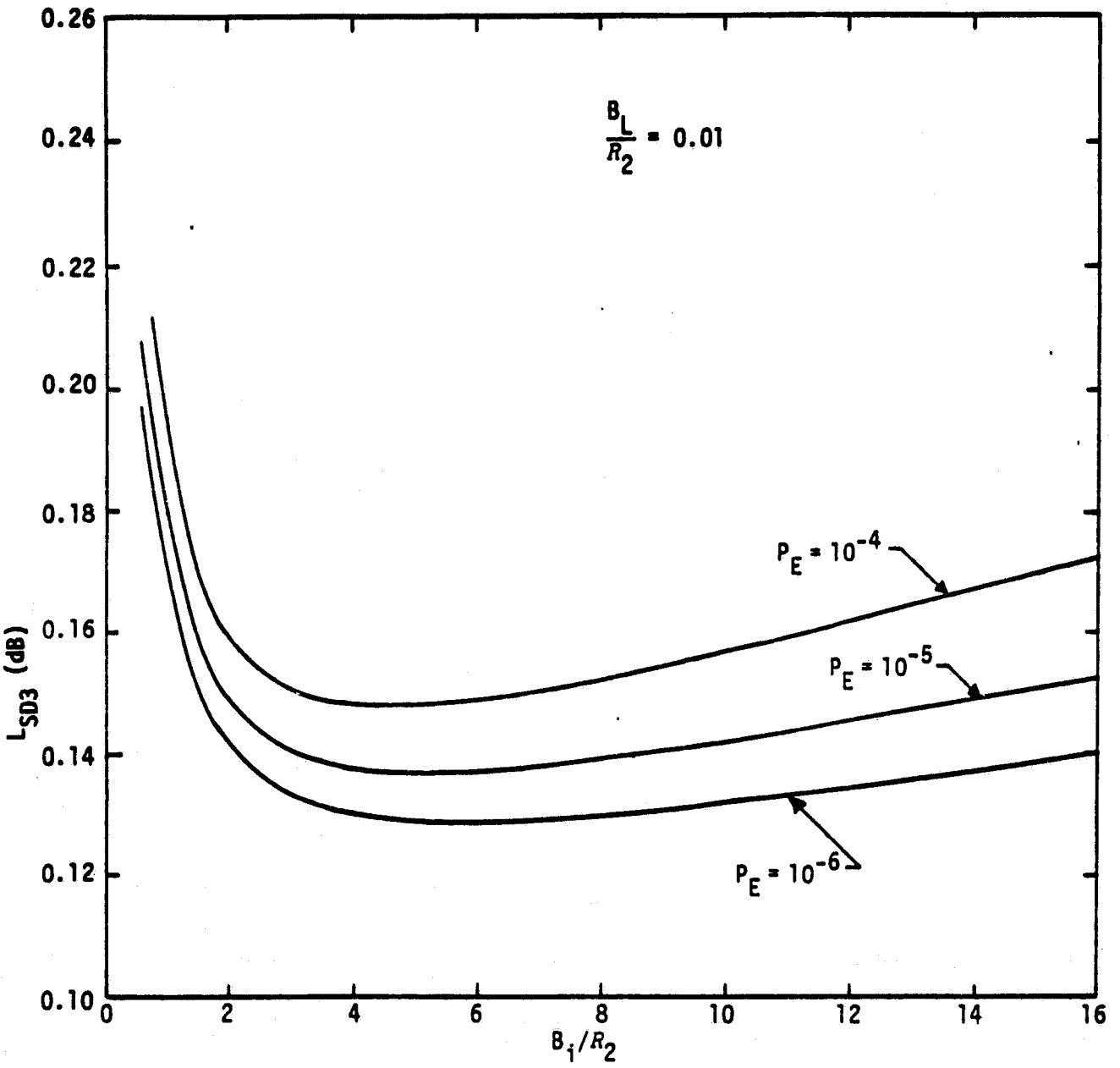


Figure 3.99. Channel 3 Noisy Reference Loss (in dB) versus the Ratio of Two-Sided Costas Loop Arm Filter Bandwidth to Channel 2 Data Rate; $m_3(t)$ is 192-kbps Manchester Coded Data, $m_2(t)$ is 2-Mbps NRZ Data

the noisy reference loss on Channel 2 is considerably smaller than that of Channel 3. The principal reason for this can be easily explained in terms of the result in (3-277), where it is observed that the effective cross-modulation loss on Channel 2, $\tilde{m}_{32}^2$, is divided by γ_T which, in this case, has a value equal to 10.42. Secondly, for either channel, the noisy reference loss decreases with increasing error probability. This is intuitively satisfying when one realizes that the slope of the error probability versus SNR curve becomes steeper as P_E becomes smaller and, thus, for a given σ_ϕ, the parallel ideal and noisy synchronization error probability curves become closer together. Finally, observe that there is an optimum arm filter bandwidth (for fixed R_2) in the sense of minimizing L_{SDi} $(i = 2,3)$. Since only σ_ϕ^2 depends upon this bandwidth, it is clear that this bandwidth choice is identical to that which minimizes σ_ϕ^2 or, equivalently, the loop squaring loss. Note that, if B_L/R_2 is decreased, then the noisy reference loss will also decrease since the equivalent loop SNR ρ increases.

In summary, it is concluded that the crosstalk degradation due to noisy subcarrier demodulation references is quite small (on the order of tenths of a decibel or less, depending on the particular channel and the ratio of loop bandwidth to data rate in that channel. When the higher data rate channel is 1-Mbps Manchester-coded data, since both channels are now Manchester coded, the crosstalk loss would then be even smaller.

While the results of [29] have been directed principally toward the demodulation of unbalanced QPSK by a conventional (single-channel) Costas loop, the expressions for average error probability [see (3-276)] and noisy reference loss [see (3-277)] apply in a much broader sense. In particular, the two-channel-type Costas loops have a mean-squared phase jitter given by (3-206) where, however, L_{SQ} is a much more complex function of the various system parameters such as data rates and channel power ratios. Nevertheless, once L_{SQ} $\left(\text{and thus } \sigma_\phi^2\right)$ is determined, (3-276) and (3-277) apply directly toward evaluation of the noisy synchronization reference effects of these loops on error probability performance.

3.2.12.2 FSK subcarrier demodulation

In subsection 3.2.11.2, FSK demodulation was performed by a limiter discriminator. Another technique for FSK demodulation is to use envelope detection. Figure 3.100 presents a binary FSK demodulator where one channel is tuned to frequency f_1 and the other channel to f_2.

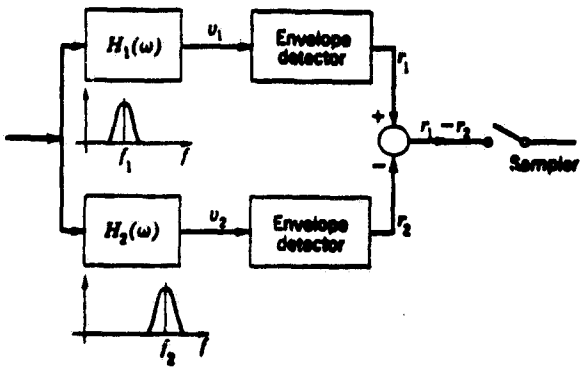


Figure 3.100. Noncoherent FSK Envelope Detection

The outputs of the two envelope detectors are compared to determine whether one binary symbol or the other was transmitted. The envelope detector in each channel is really just measuring the energy in each channel, but the random statistics are different depending on whether there is noise or signal (tone) plus noise in the filter. When only noise is present in the filter, the random variable r at the output of the envelope detector is Rayleigh distributed:

$$f_n(r) = \frac{r}{N} \exp\left(-r^2/2N\right); \quad r \geq 0, \tag{3-279}$$

where

$$N = N_0 \int_{-\infty}^{\infty} |H_\ell(j2\pi f)|^2 \, df = N_0 B_i \tag{3-280}$$

with $H_\ell(j2\pi f)$ equal to the lowpass equivalent of the bandpass filter transfer function $H(j\omega)$ and B_i is the two-sided noise bandwidth. If the signal is a sine wave for amplitude $\sqrt{2P}$ for all time, the random variable r at the output of the envelope detector is Rician distributed:

$$f_s(r) = \frac{r}{n} \exp - \frac{r^2}{2N} \exp\left(-\frac{P}{N}\right) I_0\left(\frac{r\sqrt{2P}}{N}\right), \qquad (3\text{-}281)$$

where P is the average signal power and P/N is the power SNR.

For correct FSK demodulation, the sampled value of $(r_1 - r_2)$ should be positive for a "0" transmitted (i.e., f_1) and negative for a "1" transmitted (i.e., f_2). Therefore, the probability of symbol error, P_s, assuming a "0" transmitted is

$$P_s = \int_0^\infty \frac{r_1}{N}\left(-\frac{r_1^2}{2N}\right)\exp\left(-\frac{P}{N}\right) I_0\left(\frac{r_1\sqrt{2P}}{N}\right)\int_1^\infty \frac{r_2}{N}\exp\left(-\frac{r_2^2}{2N}\right) dr_2 dr_1$$

$$P_S = \frac{1}{2}\exp\left(-\frac{P}{2N}\right) \qquad (3\text{-}282)$$

For communication signals, the signal tone is present for only T_s seconds then switches to the other tone or stays the same, with equal probability. Therefore, since the bandpass filter bandwidth is chosen to maximize the output SNR, not all of the signal power P may be passed. This loss of signal power due to filtering corresponds to the signal distortion factor D_m presented in subsection 3.2.10.2 where

$$D_m = \int_{-\infty}^{\infty} S(f)\,|H_\ell(j2\pi f)|^2\, df\,, \qquad (3\text{-}283)$$

where S(f) is the power spectral density of the data modulation. For data (symbol) rate R_s, Figures 3.54 and 3.59 present D_m versus B_i/R_s for Manchester and NRZ data formats, respectively. Including D_m in the calculation of probability of error,

$$P_s = \frac{1}{2}\exp\left(-\frac{P\,D_m}{2N_0B_i}\right)$$

$$= \frac{1}{2}\exp\left(-\frac{E_s}{2N_0}L_{BE}\right), \qquad (3\text{-}284)$$

where the FSK demodulation bandpass-envelope detector loss, L_{BE}, is given by

$$L_{BE} = \frac{D_m}{B_T/R_s} \quad . \tag{3-285}$$

3.2.13 Digital Data Detection

After the carrier and subcarrier demodulation have been performed, the data is at baseband. This subsection presents the losses from ideal detection that result in practical systems. In order to establish performance losses, the ideal detection performance must be established. The ideal detector is referred to as a matched filter.

3.2.13.1 Matched filter detection

The digital data detection problem is to minimize the probability of error for a given E_b/N_0. Consider the received signal S(t) representing one digital pulse arriving at t = 0 which is contaminated by white noise with single-sided noise spectral density N_0. Presumably a filter should be inserted ahead of the demodulator to limit the noise. The optimum receiver filter has transfer function H(s) which maximizes the SNR at the filter output or maximizes the expression

$$SNR = \frac{\left| \int_{-\infty}^{\infty} H(j2\pi f)\, S(2\pi f)\, \exp\,(j2\pi f\tau)\, df \right|^2}{N_0 \int_{-\infty}^{\infty} |H(j2\pi f)|^2\, df} \quad , \tag{3-286}$$

where τ is some time at which the SNR is maximized and $S(2\pi f)$ is the signal power spectral density. The only variable to maximize SNR in (3-286) is $H(j\omega)$. The optimum filter is

$$H(j\omega) = S^*(j\omega)\, \exp\,(-j\omega\tau) \quad , \tag{3-287}$$

which has impulse response

$$h(t) = s(\tau - t) \quad . \tag{3-288}$$

Filters having the property of (3-288) are said to be <u>matched filters</u>, matched to s(t), in that the impulse response has the shape of s(t) delayed by τ and reversed in time. Note that matched filtering is of the form of correlation detection. That is, in the absence of noise, the filter output is s(t) correlated with itself, since

$$y(t) = s(t)*h(t) = \int_{-\infty}^{\infty} h(t')\, s(t - t')\, dt'$$

$$= \int_{-\infty}^{\infty} s(\tau - t')\, s(t - t')\, dt' . \qquad (3\text{-}289)$$

At $t = \tau$, the peak output is then

$$y(\tau) = \int_{-\infty}^{\infty} s^2(\tau - t')\, dt' = \int_{-\infty}^{\infty} s^2(t)\, dt . \qquad (3\text{-}290)$$

There are two practical difficulties with matched filters: (1) usually they cannot be physically realized and, (2) the output y(t) has roughly twice the duration of s(t), leading to problems of intersymbol interference. At the cost of increased transmission bandwidth, both of these problems can be overcome by using rectangular pulses (or the modulated equivalent). Then the matched filter's impulse response is also rectangular and, in the absence of noise, the output waveform is triangular with twice the duration of the input pulse. To prevent overlapping outputs, the filter is discharged (shorted) just after the output peaks. Since this process is equivalent to integrating over the input pulse duration, such devices are known as integrate-and-dump filters.

As an alternative to the integrate-and-dump method, one can design the filter in reverse order, with y(t) being a specified output pulse shape chosen to minimize intersymbol interference. The time delay t_0 is then relatively unimportant and can be as large as required, within reason, to facilitate filter approximation.

For binary data transmission in the presence of white Gaussian noise, one of two equiprobable pulses, $s_0(t)$ and $s_1(t)$, is transmitted.

The energies of the pulses are given as

$$E_0 = \int_{-\infty}^{\infty} s_0^2(t)\, dt$$

$$E_1 = \int_{-\infty}^{\infty} s_1^2(t)\, dt \qquad (3\text{-}291)$$

and the average energy is

$$E = \frac{1}{2}\left(E_0 + E_1\right) . \qquad (3\text{-}292)$$

The correlation coefficient of $s_0(t)$ and $s_1(t)$ is defined as

$$\rho = \frac{1}{E}\int_{-\infty}^{\infty} s_0(t)\, s_1(t)\, dt , \qquad (3\text{-}293)$$

where $-1 \leq \rho \leq 1$.

Because there are two pulses in question, optimum detection involves a pair of matched filters

$$h_0(t) = s_0(\tau - t) \; ; \quad h_1(t) = s_1(\tau - t) . \qquad (3\text{-}294)$$

In the absence of noise, each filter produces a maximum output only for its matched input. Therefore, the filter outputs can be subtracted and an appropriate threshold can be used to determine which pulse actually arrived. Figure 3-101a diagrams the complete detector. Alternately, since the filters are linear, the same effect is achieved more conveniently by a single filter with

$$h(t) = h_1(t) - h_0(t) = s_1(\tau - t) - s_0(\tau - t) , \qquad (3\text{-}295)$$

giving the simplified system of Figure 3.101b.

For either arrangement, the detection is properly classified as being coherent in the sense that the receiver has available stored copies of the uncontaminated pulse shapes. Emphasizing this point,

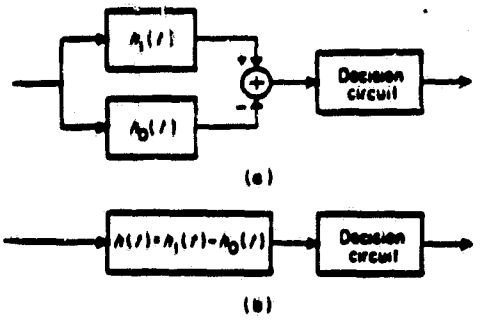


Figure 3.101. Coherent Binary Detection. (a) Two Matched Filters; (b) Equivalent Single Matched Filter

Figure 3.102 shows the equivalent product-type coherent detector with the reference signal $r(t) = s_1(t) - s_0(t)$.

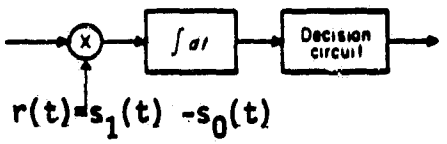


Figure 3.102. Coherent Detection With Reference Signal

The performance in terms of probability of symbol error, P_s, for signals of equal symbol energy E_s is equal to

$$P_s = Q\left(\sqrt{\frac{E_s(1-\rho)}{N_0}}\right). \tag{3-296}$$

The two cases of most interest in terms of the correlation coefficient are:

$$\begin{aligned} \rho &= -1 \quad \text{antipodal} \\ \rho &= 0 \quad \text{orthogonal} . \end{aligned} \tag{3-297}$$

Binary PSK is an example of $\rho = -1$ and binary FSK is an example of $\rho = 0$.

When the received carrier phase is not known, the received signal is of the form

$$y(t) = \sqrt{2s}\cos\left(w_k t + \theta\right) + n(t);\ 0 \leq t \leq T, \qquad (3\text{-}298)$$

for a pulse of duration T where θ is the random phase shift, introduced by the channel, which is assumed to be a uniformly distributed random variable between $-\pi$ and π and independent of frequency. It is assumed for the matched-filter performance that the receiver is operating with perfect time and frequency synchronization.

For PSK, the information is lost unless θ can be estimated in some way or eliminated. Differentially coherent reception (DPSK) assumes that θ does not change over two symbol times (i.e., 2T). Therefore, if an NRZ-M format is used (or Manchester M format), then, for $s_0(t)$ transmitted, followed by an input data stream of 1101, the next four transmitter signals are $s_1(t)$, $s_1(t)$, $s_0(t)$, $s_1(t)$. Also, for PSK, $s_1(t) = -s_0(t)$. Since the phase is constant over 2T, then

$$s_i(t)\, s_i t+T) = \begin{cases} 1 \text{ for input symbol } 0 \\ -1 \text{ for input symbol } 1\ . \end{cases} \qquad (3\text{-}299)$$

Figure 3-103 presents the matched filter DPSK demodulator. Note that the decision output does not depend on the value of θ as long as it is constant. The symbol probability of error for DPSK can be shown to be [41]:

$$P = \frac{1}{2}\exp\left(-E/N_0\right). \qquad (3\text{-}300)$$

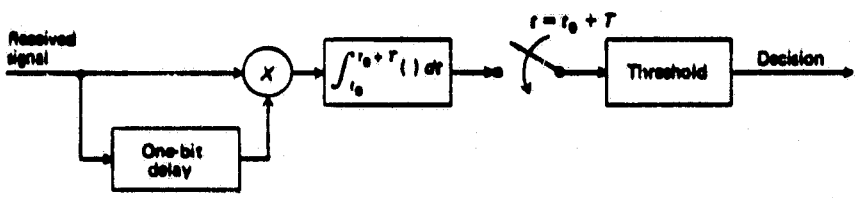


Figure 3.103. Optimum DPSK Demodulation

For FSK, k = 1,2...N for the carrier frequency ω_k in (3-298). The optimum noncoherent receiver for FSK ($\rho = 0$) is to eliminate the influence of θ by computing the following set of samples:

$$M_k^2 = X_k^2 + Y_k^2 ; \quad k = 1,2,\ldots,N \tag{3-301}$$

where

$$X_k = \sqrt{2S}\int_0^T y(t) \cos \omega_k t \, dt$$

$$Y_k = \sqrt{2S}\int_0^T y(t) \sin \omega_k t \, dt \tag{3-302}$$

and select, on the basis of these samples, the signal that gives rise to the largest M_k^2. Note that (3-301) and (3-302) depend on the envelope of the cross-correlation function of the received signal y(t) and the kth stored message waveform. By "stored," it is meant that the waveforms are stored in time and in frequency (assumed perfect). Alternately, the probability computer contains a set of N matched filters (one matched to each signaling waveform), each followed by an envelope detector (usually, a linear rectifier and a lowpass filter are sufficient) and a device which samples the output envelopes. Figure 3.104 illustrates the cross-correlation version of the receiverrand its alternate matched-filter realization for one particular element.

The symbol probability of error for binary FSK (N=2) with noncoherent reception is given by [41]:

$$P_s = \frac{1}{2} \exp\left(-E_s/2N_0\right) . \tag{3-303}$$

Figure 3.105 compares the performance of binary PSK and FSK for optimum coherent, differentially coherent, and noncoherent reception. Note from (3-296), (3-300), and (3-303) that coherent reception of FSK (i.e., $\rho = 0$, orthogonal) is a factor of 2 (3 dB) worse than PSK (i.e., $\rho = -1$, antipodal) and that noncoherent reception of FSK is 3 dB worse than differentially coherent reception of

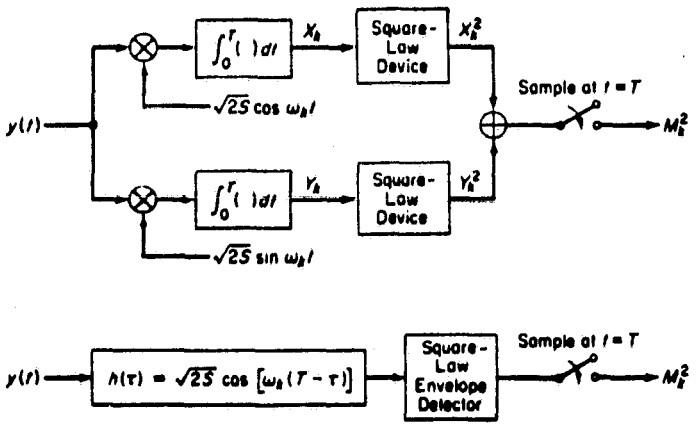


Figure 3.104. Alternate Ways to Mechanize an Element of the Optimum Noncoherent Receiver

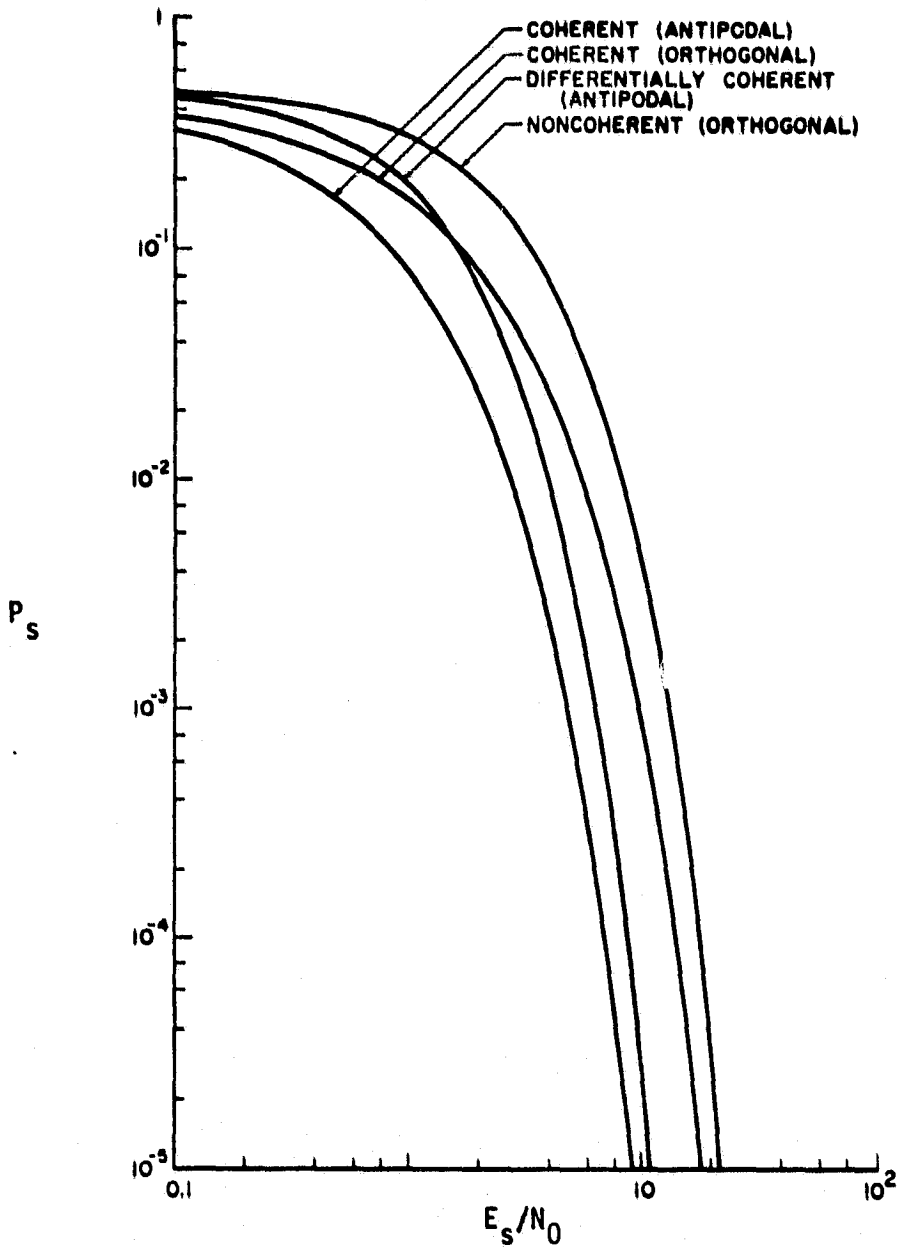


Figure 3.105. Comparison of Error Probabilities for Coherent, Noncoherent, and Differentially Coherent Reception

PSK. As E_s/N_0 becomes large, Figure 3.105 shows that the performance of coherent and differentially coherent reception of PSK become essentially equal, which is also true of coherent and noncoherent reception of FSK.

Binary signaling can readily be extended to the N-ary case by considering one of N known signals to be transmitted. The mechanization of the optimum receiver can be realized with a bank of N multipliers and finite-time integrators of the type shown in Figure 3.106, with the input to the multiplier now being $S_1(t)$. These devices are again referred to as correlators and the process as correlation detection. The quantities q_i are referred to as the correlator outputs.

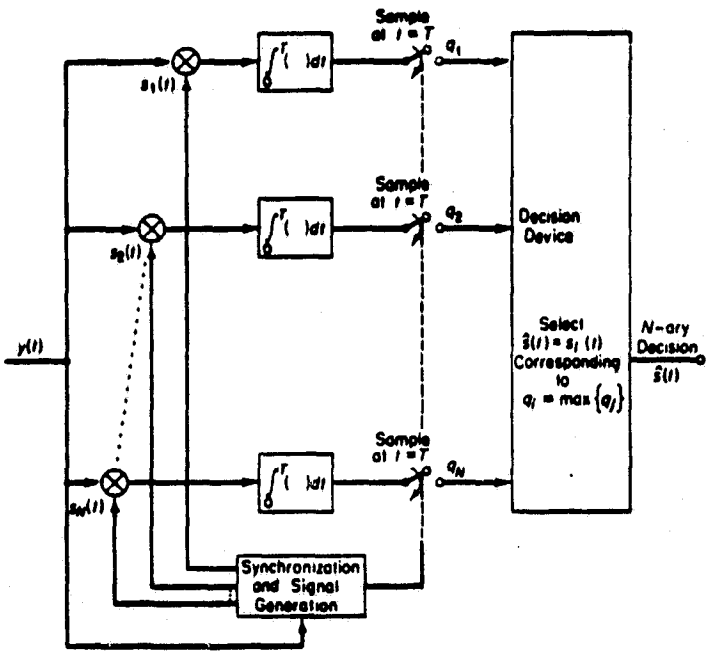


Figure 3.106. Correlation Receiver For N-ary Decision Problem

At the conclusion of the signaling interval, a decision mechanism examines the outputs, determines the greatest, and announces that signal for which the output is greatest. To describe the N signal properties, it is convenient to denote the set of normalized signal inner products λ_{ij} of the signal set $\{s_i(t)\}$ for all i and j as

$$\lambda_{ij} \triangleq \frac{1}{E}\int_0^T s_i(t)\, s_j(t)\, dt \ , \qquad (3\text{-}304)$$

where it is assumed that all N signals have equal energy:

$$E = \int_0^T s_i^2(t)\, dt \ ; \quad i = 1,2,\ldots,N \ . \qquad (3\text{-}305)$$

The set of signals $\{ s_i(t), i = 1,2,\ldots,N \}$ is said to be orthogonal only if

$$\lambda_{ij} = \begin{cases} 0 \ ; & i \neq j \\ 1 \ ; & i = j \end{cases} \qquad (3\text{-}306)$$

Multiple FSK (MFSK) is an example of orthogonal signal sets.

The biorthogonal signal set can be obtained from an orthogonal set of N/2 signals by augmenting it with the negative of each signal. Here,

$$\lambda_{ij} = \begin{cases} 0 \ ; & i \neq j, \ i - j \neq N/2 \\ -1 \ ; & i \neq j, \ i - j = N/2 \\ 1 \ ; & i = j \ . \end{cases} \qquad (3\text{-}307)$$

PSK and QPSK correspond to biorthogonal signals with N = 2 and 4, respectively.

Figures 3.107 and 3.108 present the binary symbol probability of error $P_b(N)$ versus $R_b = ST_b/N_0$ for orthogonal and biorthogonal signals, respectively. Note that T_b is the input bit duration at the modulator. Figure 3.109 presents $P_b(M)$ versus R_b for MFSK with noncoherent demodulation.

3.2.13.2 Bit/symbol synchronization

Thus far, two aspects of the synchronization problem (carrier and subcarrier synchronization) have been discussed. The problem of symbol synchronization deals with the estimation of the time-of-arrival or epoch of the received data symbols. Data-derived symbol synchronization is where the receiver includes a device that extracts the

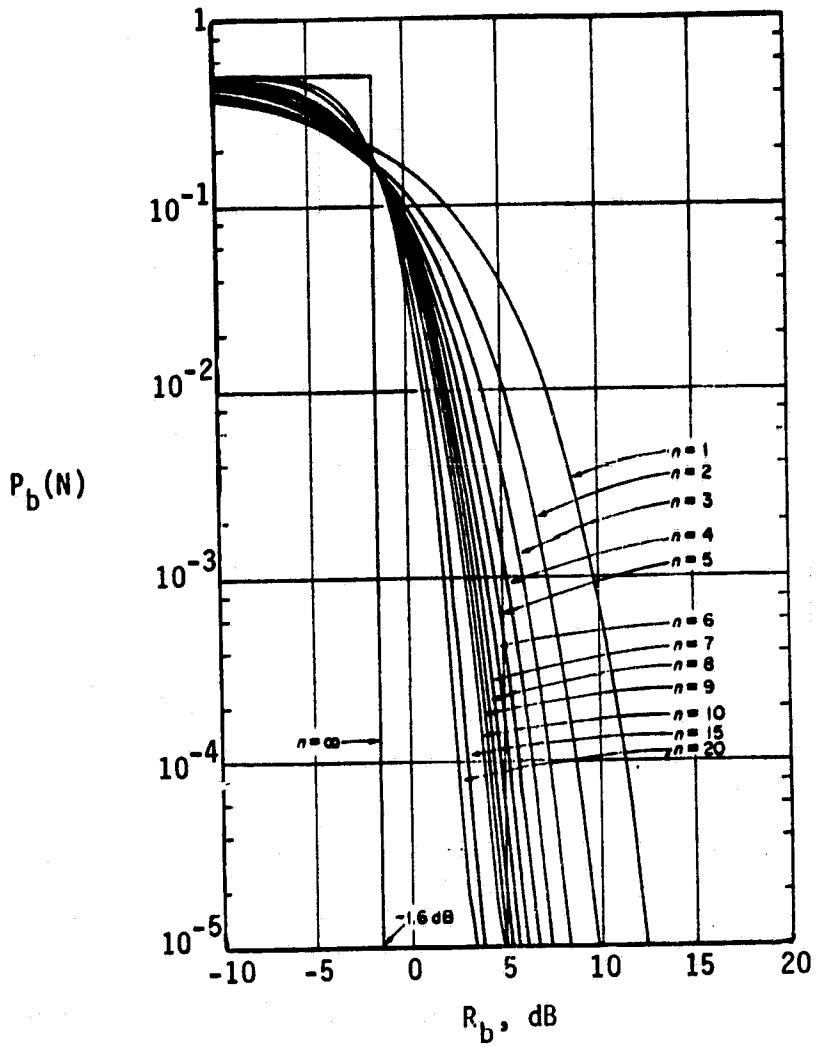


Figure 3.107. Binary Symbol Error Probability Performance of a System Transmitting an Orthogonal Signal Set with Coherent Demodulation

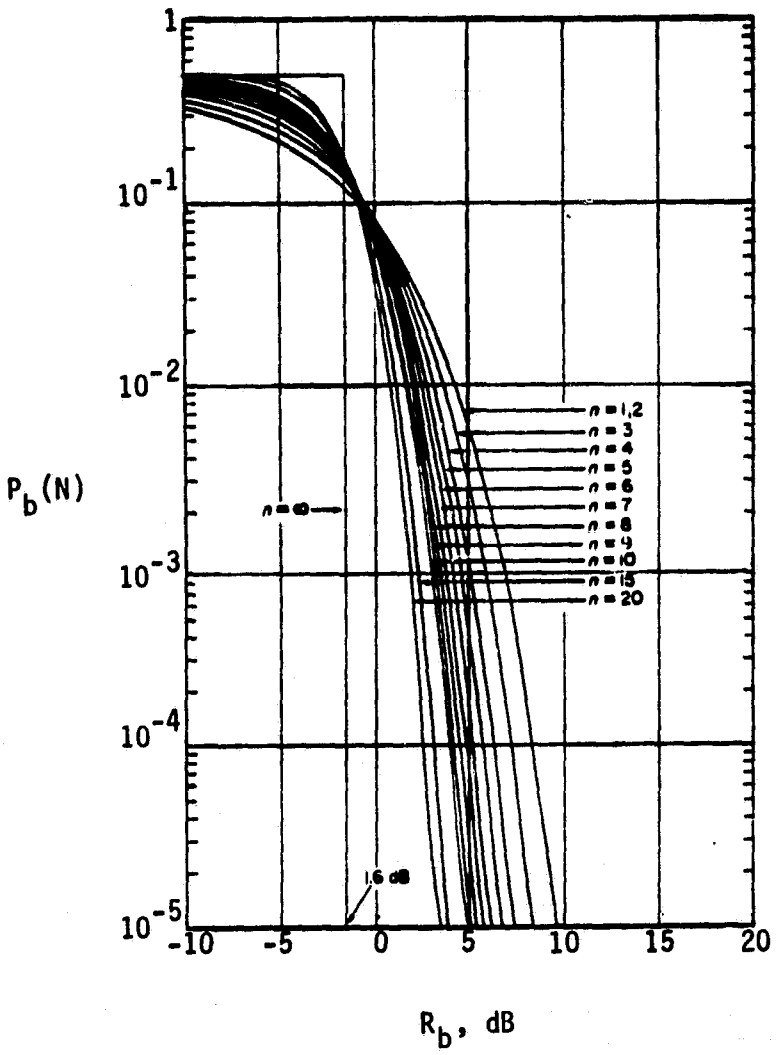


Figure 3.108. Binary Symbol Error Probability Performance of a System Transmitting a Biorthogonal Signal Set With Coherent Demodulation

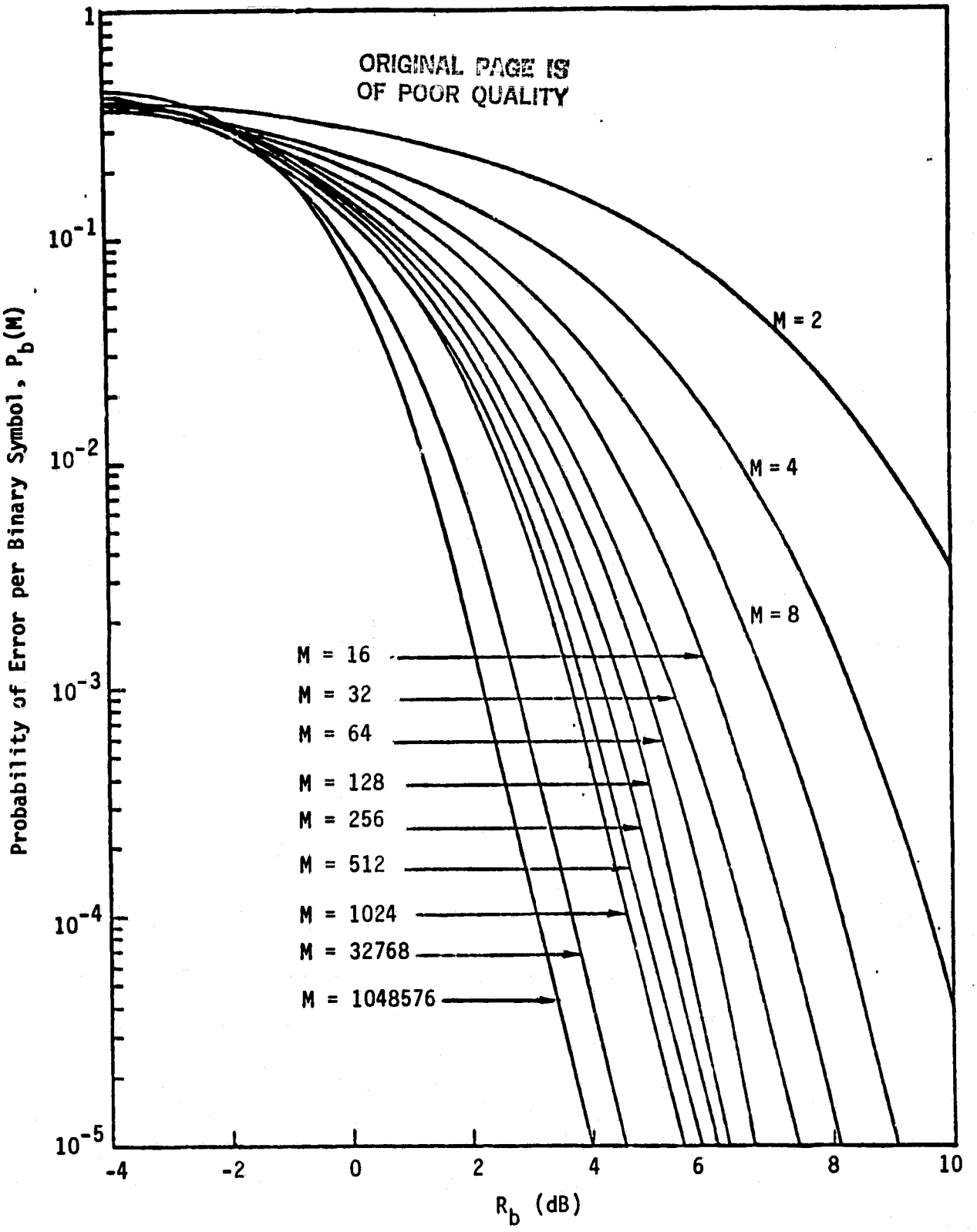


Figure 3.109. Performance of Noncoherent MFSK Modulation

synchronization directly from the information-bearing signal. This subsection investigates several data-derived symbol synchronizers but is in no way a complete catalog of possible configurations.

For digital implementations of the symbol synchronizer, the data-transition tracking loop (DTTL) shown in Figure 3.110 is one of the most promising configurations. The input noise-free signal $s(t,\varepsilon)$ is a random pulse train as characterized by

$$s(t,\varepsilon) = \sum_{k=0}^{K} a_k \, p_s[t-(k-1)T-\varepsilon] \, , \qquad (3\text{-}308)$$

where a_k is the polarity (± 1) of the kth transmitted symbol with random epoch ε assumed to be constant for KT seconds. The basic pulse shape $p_s(t)$ is defined to be nonzero only over the interval $(0,T)$. For rectangular pulses, $p_s(t)$ is equal to $\sqrt{S}$ for $0 \leq t \leq T$. The sum of signal-plus-noise $y(t)$ is passed through two parallel branches which are triggered by a timing pulse generator according to a digitally filtered

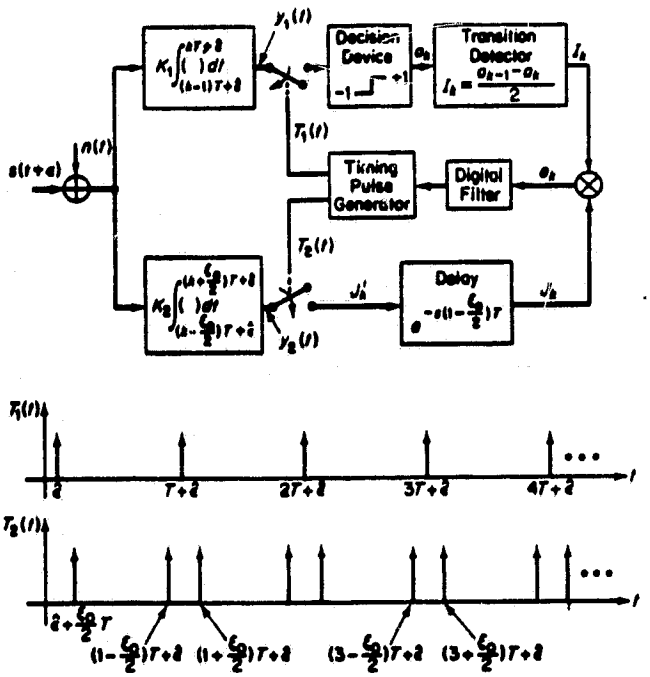


Figure 3.110. Digital-Data-Transition Tracking Loop (DTTL)

version of an error signal formed from the product of the branch outputs. Furthermore, the two branches are held at a fixed phase relationship with one another by the timing generator. Basically, the inphase branch monitors the polarity of the actual transitions of input data and the quadrature (midphase) branch obtains a measure of the lack of synchronization. The particular way in which these two pieces of information are derived and combined to synchronize the loop is described below.

The input signal is passed through inphase and quadrature integrate-and-dump circuits. The output of the inphase integrate-and-dump is sampled at intervals of T and a plus/minus decision is made corresponding to each input symbol. The decision device is simply a hard limiter with a signum function input-versus-output characteristic. The transition detector then examines two adjacent decisions a_{k-1}, a_k and records an output I_k according to the following rules:

$$\begin{aligned} &\text{If } a_k = a_{k-1}, && \text{then } I_k = 0; \\ &\text{if } a_k = -1,\ a_{k-1} = +1, && \text{then } I_k = +1; \\ &\text{if } a_k = +1,\ a_{k-1} = -1, && \text{then } I_k = -1. \end{aligned} \tag{3-309}$$

The output J_k' of the quadrature integrate-and-dump is also sampled at intervals of T and must be delayed before multiplication with the appropriate I_k. An improvement in steady-state mean-squared synchronization error and mean-time-to-first-slip performance can be obtained by integrating in the quadrature branch only over a portion of the symbol interval (e.g., $\xi_0 T$; $0 \leq \xi_0 \leq 1$). Then, for proper loop operation, the delay in the quadrature branch must be chosen equal to $(1 - \xi_0/2)T$. As an example, for full symbol integration in the quadrature branch corresponding to true midphase sampling in that branch, the required delay for correct formation of the error signal $e_k \triangleq I_k J_k$ is T/2. Actually, for many practical applications, the loop bandwidth-symbol time product is much less than unity; hence, the delay factor $\exp\left[-sT(1 - \xi_0/2)\right]$ has a negligible effect on the analysis of the loop operation. The error signal e_k is digitally filtered, with the resulting output being used to control the instantaneous frequency of the

timing pulse generator. Thus, an estimate $\hat{\epsilon}$ of the input random epoch is formed.

Two other implementations of the symbol synchronizer of Figure 3.110 have been considered. In one case, each branch is split into two subbranches, each having an integrator. Now, the integrators need not be dumped, since one can be resetting while the other is integrating. The outputs of the two subbranches are then multiplexed to give the resultant output waveform for that branch. In fact, when $\xi_0 \leq 1/2$, one subbranch from each branch can be eliminated since the waveform will no longer overlap in time.

In another implementation, the output of the integrator is processed before decision by removing the previously integrated value, a procedure equivalent to dumping the integrator after each sampling instant.

The normalized symbol synchronization error is

$$\lambda = (\epsilon - \hat{\epsilon})/T \ . \tag{3-310}$$

Figure 3.111 presents the variance of the normalized symbol synchronizer error σ_λ^2 versus $R_d = ST/N_0$ for values of $\delta_{s0} = 20$ and 100, when

$$\delta_{s0} = \frac{K_2 K_g}{B_L T} \ , \tag{3-311}$$

where K_2 is equal to the gain of the quadrature arm (e.g., see Figure 3.110) and K_g is phase detector gain given by [42]:

$$K_g = \text{erf}\left(\sqrt{R_d}\right) - \frac{\xi_0}{2}\sqrt{\frac{R_d}{\pi}}\ \exp\left(-R_d\right) . \tag{3-312}$$

Also included in Figure 3.111 (solid line) are the corresponding values obtained by optimizing the value of ξ_0. Several interesting results can be observed in Figure 3.111, namely:

(1) For small $R_d\delta_{s0}$, the PDF $p(\lambda)$ becomes uniform in the interval (-1/2, 1/2) and thus is independent of ξ_0. In fact,

$$\lim_{R_d \to 0} \sigma_\lambda^2 = \frac{1}{12} \qquad \text{for all } \xi_0 . \tag{3-313}$$

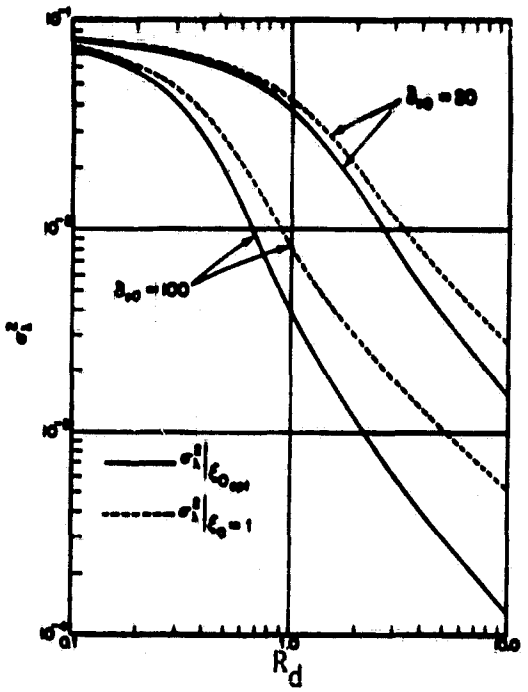


Figure 3.111. Variance of the Normalized Symbol Synchronization Error Versus Signal-to-Noise Ratio for Various Values of δ_{s0}; $\xi_0 = 1$, and $\xi_0 = \xi_{opt}$

(2) For large $R_d\delta_{s0}$ and $\xi_0 = 1$, $p(\lambda)$ has the form

$$p(\lambda) = C_1\left[1 + 4R_d\lambda^2\right]^{-\left(1+\delta_{s0}/4\right)}; \quad |\lambda| \leq \frac{1}{2}, \tag{3-314}$$

with zero mean and variance

$$\sigma_\lambda^2 = \frac{1}{2R_d\delta_{s0}}\left[\frac{1}{1 - 2/\delta_{s0}}\right]. \tag{3-315}$$

The factor $1 - 2/\delta_{s0}$ represents the increase in mean-squared synchronization error due to considering the variation of the noise spectrum with λ which, for large δ_{s0}, becomes negligible.

It is desired to optimize the performance of the symbol synchronizer by varying the integration interval $\xi_0 T$ in the quadrature branch along with the input SNR R_d. In terms of the above model, a

system constraint is imposed whereby the average power of the equivalent feedback reference signal that is cross-correlated with the input signal-plus-noise will be maintained constant as ξ_0 is varied at fixed R_d. Curves of σ_λ^2 versus ξ_0 with R_d and δ_{s0} as parameters are illustrated in Figures 3.112 and 3.113. It may be observed that, for each value of R_d, an optimum $\xi_0 \left(\xi_{0\,opt}\right)$ exists, along with a corresponding value of σ_λ^2. Figure 3.114 illustrates the variation of $\xi_{0\,opt}$ versus R_d for $\delta_{s0} = 20$ and 100.

It is also possible to optimize the mean time to first half slip performance of the symbol synchronizer by varying ξ_0 at fixed R_d. Figure 3.115 plots $T_{1/2}$ versus ξ_0 as a function of R_d. In this figure, the normalized time-bandwidth product δ_{s0} is held constant. For each SNR, there is an optimum (maximum) mean time to first half slip. The value of ξ_0 at which the maximum occurs $\left(\xi_{0\,opt}\right)$ is a function of the SNR R_d and δ_{s0}. This functional relationship is illustrated in Figure 3.116. The optimum mean time to first half slip $T_{1/2_{opt}}$ is plotted against R_d in Figure 3.117 for $\delta_{s0} = 100$ along with the values of $T_{1/2_{|\xi=1}}$, the latter corresponding to the case of no optimization. It is interesting to compare the values of ξ_0 that minimize σ_λ^2 with those that maximize $T_{1/2}$. In the actual design of a system, ξ_0 is chosen as some compromise between these two optimum values, depending on the relative importance of σ_λ^2 and $T_{1/2}$. Alternately, one can constrain $T_{1/2}$ to be above a certain threshold value and choose ξ_0 on the basis of σ_λ^2.

In practice, there are several other classes of symbol synchronizer configurations that exploit the idea of shaping the equivalent loop nonlinearity. One such class contains synchronizers of the early/late gate integration type (Figure 3.118). Several possible circuit topologies (Figure 3.119) have been suggested for implementing the associated phase detector characteristic. Of these, the one of most interest is the absolute-value type (Figure 3.119a). Also, for comparison purposes, the results for the phase detector topology of Figure 3.119b, which synthesizes a difference-of-squares (DSL) type of loop error signal, is presented. The limiter approximation topology of Figure 3.119c is merely a circuit simplification of Figure 3.119b that allows one to use a chopper rather than an analog multiplier to form the loop error signal.

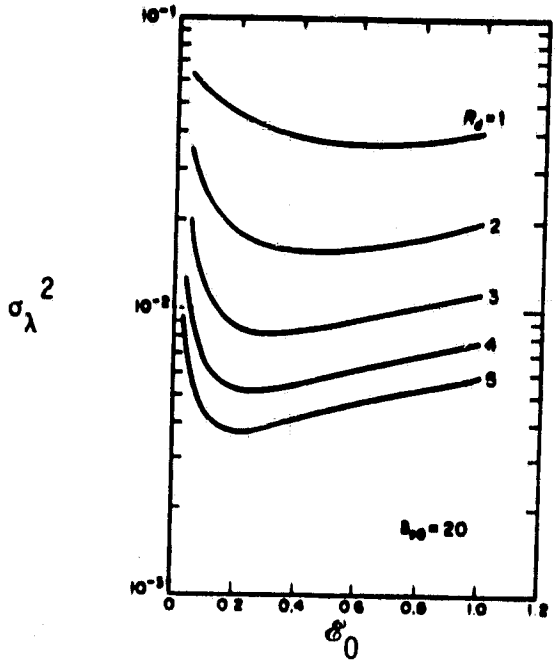


Figure 3.112. Variance of the Normalized Symbol Synchronization Error Versus Normalized Integration Interval for the Quadrature Branch (δ_{s0} = 20; SNR is a Parameter)

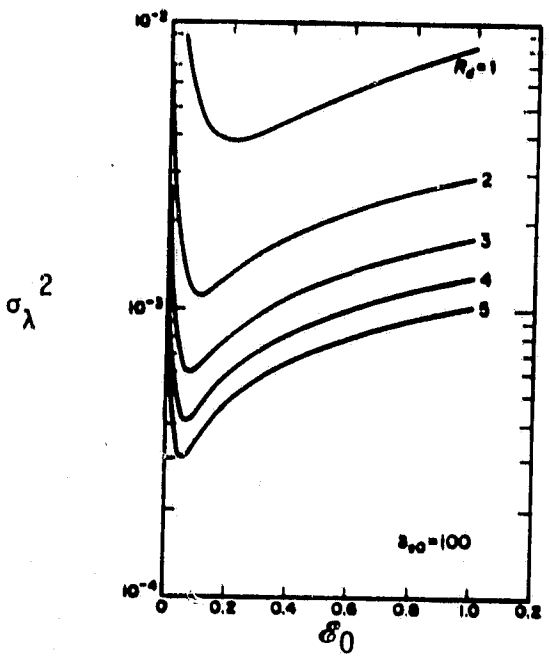


Figure 3.113. Variance of the Normalized Symbol Synchronization Error Versus Normalized Integration Interval in the Quadrature Branch (δ_{s0} = 100; SNR is a Parameter)

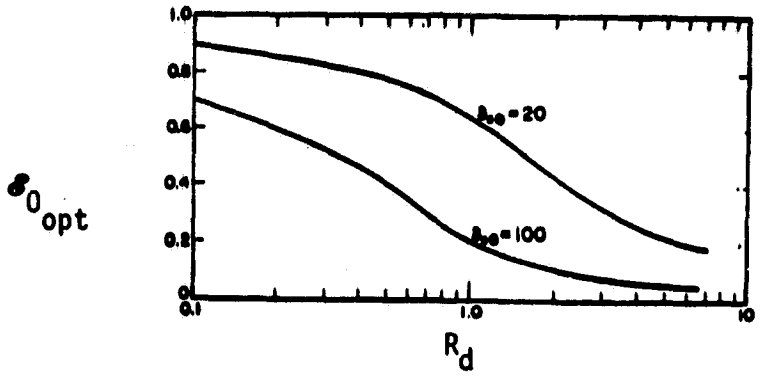


Figure 3.114. Optimum Normalized Integration Interval for the Quadrature Branch Versus SNR with δ_{s0} as a Parameter

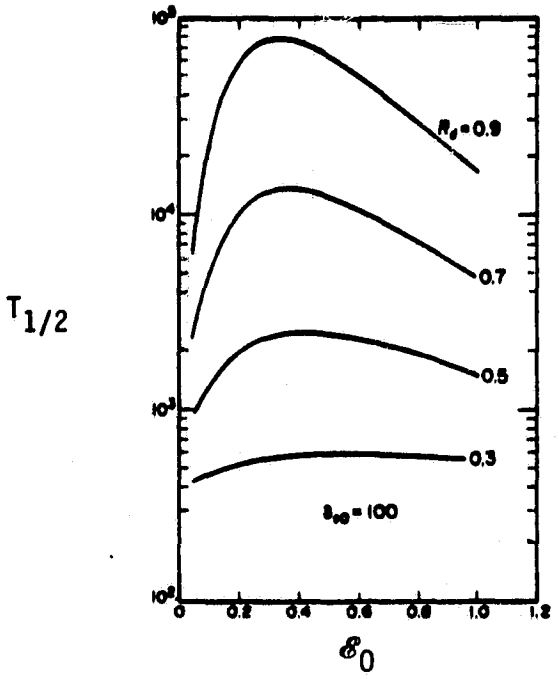


Figure 3.115. Mean-Time to First-Half Cycle-Slip Versus Normalized Integration Interval for the Quadrature Branch (δ_{s0} = 100; SNR is a Parameter)

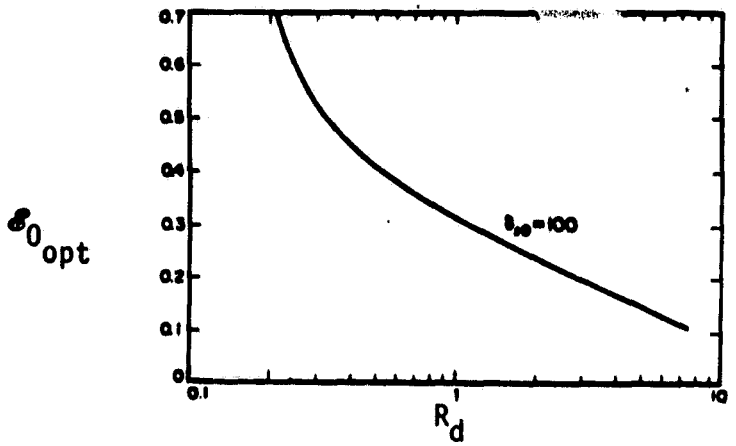


Figure 3.116. Optimum Normalized Integration Interval for the Quadrature Branch Versus SNR (δ_{s0} = 100)

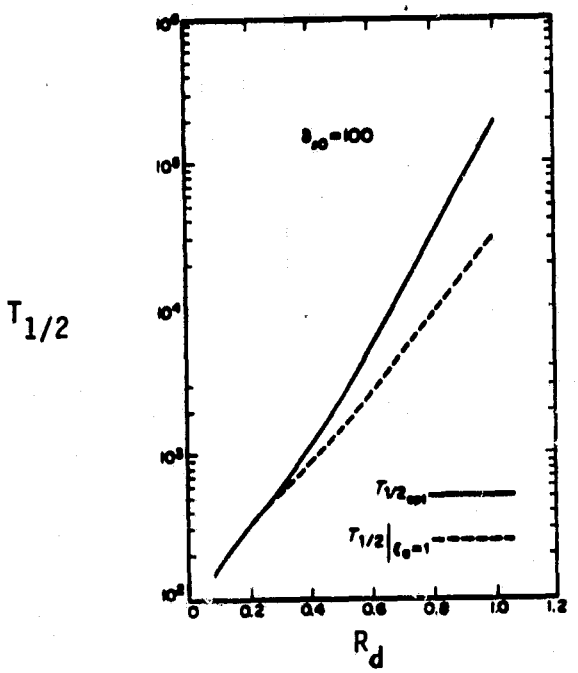


Figure 3.117. Mean-Time to First-Half Cycle-Slip Versus SNR (δ_{s0} = 100; ξ = 1 and $\xi_0 = \xi_{0\,opt}$)

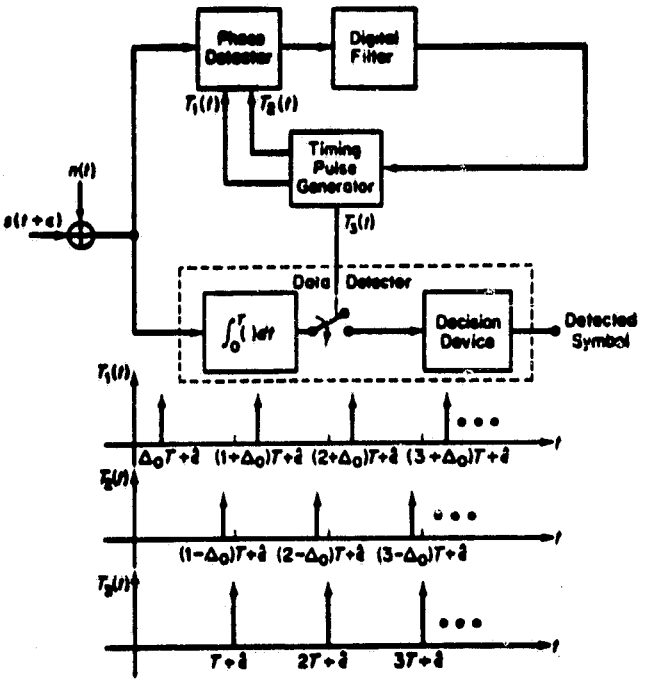


Figure 3.118. Early/Late Gate Symbol Synchronizer and Associated Data Detector

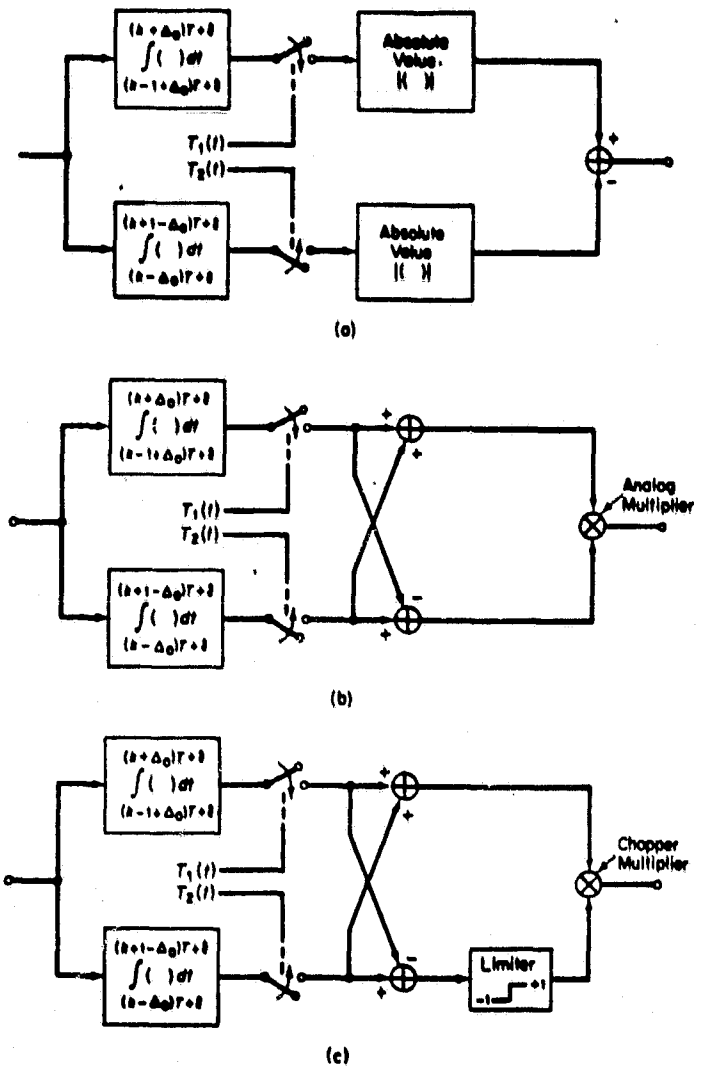


Figure 3.119. Three Different Phase Detector Topologies

Figure 3.120 is a functional diagram of the absolute value type (AVTS) of the early/late gate symbol synchronizer. An error signal e_k is generated by differencing the absolute values of two integrals of the input signal-plus-noise, each taken over a full symbol interval T, starting $\Delta_0 T$ early and $\Delta_0 T$ late, respectively. In the absence of noise the ideal operation of the loop is as follows. For a given phase offset λT between the actual transition time $\{t_k\}$ and their local estimates $\{\hat{t}_k\}$, the error signal e_k is (1) zero when no transition occurs at t_k and (2) linearly proportional to λ when a transition occurs at t_k, independent of its polarity. The polarity independence in (2) is a direct consequence of taking absolute values before differencing. The filtered error signal is used to drive a timing pulse generator, which controls the charging and discharging instants of the matched filters.

When comparing the performance of several different symbol synchronizer configurations, one must choose a fixed operating condition that is common to all. One possible basis for comparison is to require equal loop bandwidths for all configurations at every R_d. This comparison is meaningful when considering the actual design of a loop having a specified bandwidth operating over a given range of input SNR's.

Figure 3.121 plots σ_λ^2 versus R_d for $\delta_s = 20$ and 100. Included in this plot are the corresponding curves for the DTTL with full integration in the quadrature branch and DSL. The asymptotic behaviors of σ_λ^2 for large R_d are described as

$$\text{AVTS:}\quad \sigma_\lambda^2 = \frac{1}{4 R_d \delta_s}\,; \qquad \text{DTTL:}\quad \sigma_\lambda^2 = \frac{1}{2 R_d \delta_s}$$

$$\text{DSL:}\quad \sigma_\lambda^2 = \frac{1}{16 R_d \delta_s}\,. \tag{3-316}$$

Hence, in the linear region, the AVTS synchronizer offers a 3-dB advantage over the DTTL and a 0.97-dB advantage over the DSL. Furthermore, the AVTS appears to be uniformly better than the other two configurations at all values of SNR. In the limit, as $R_d \to 0$, all three configurations give the variance of a uniform distribution, that is, $\sigma_\lambda^2 = 1/12$.

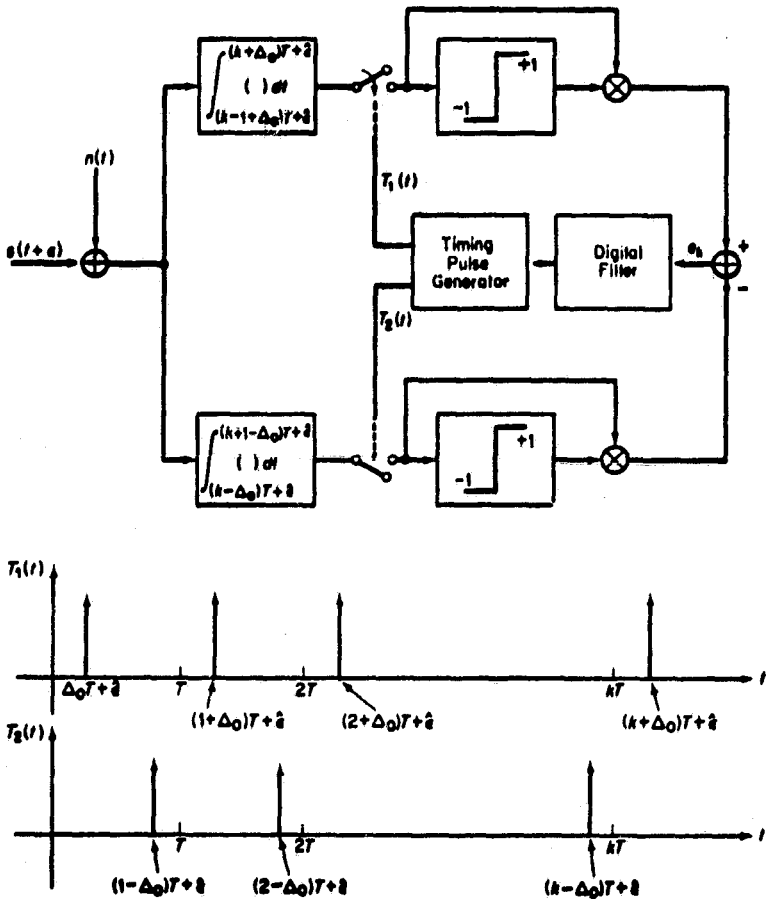


Figure 3.120. An Absolute-Value Type of Early/Late Gate Symbol Synchronizer (AVTS)

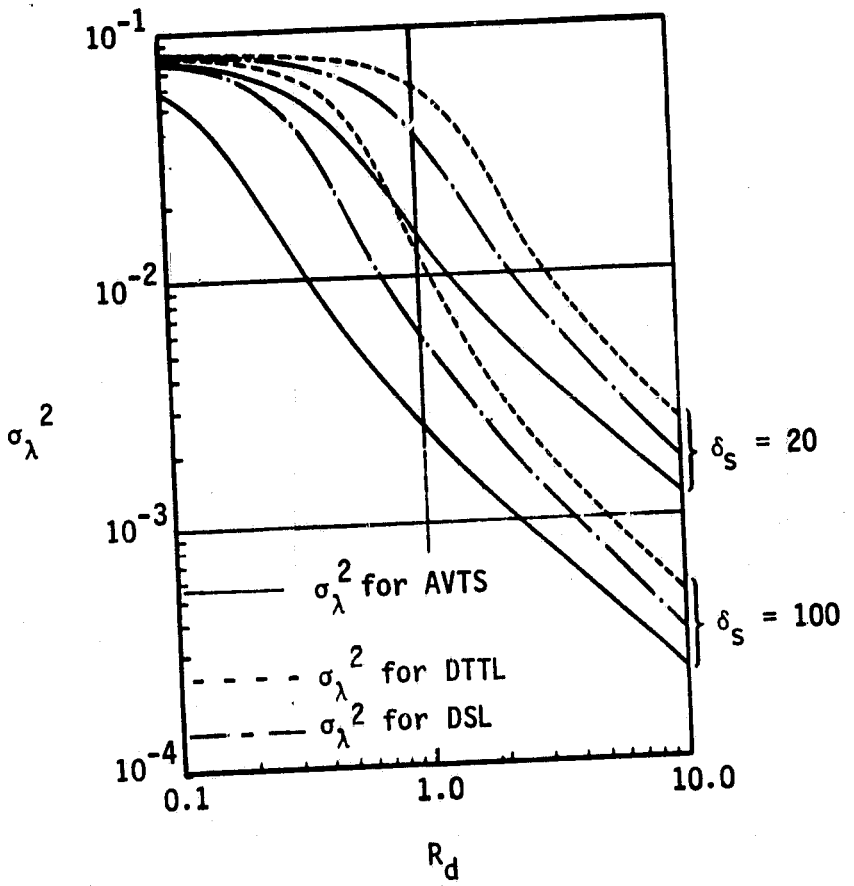


Figure 3.121. Performance Comparison of Symbol Synchronizer Configurations in Terms of the Variance of Normalized Symbol Synchronization Error Versus SNR (δ_s = 20 and δ_s = 100)

A comparison of performance on the basis of mean time to first slip can also be made for the three symbol synchronizer topologies. As one might expect, the relative rankings of the three configurations remain the same, as in the case where σ_λ^2 is compared at each value of R_d and δ_s.

To determine the bit synchronizer loss L_{BS}, the effect of symbol synchronization error on symbol error probability needs to be computed. Since the optimum detector for known signals is a cross-correlator, the interest is primarily in examining the performance degradation of this device when the symbol synchronization reference is noisy. The first step in the computation is to determine the error probability of the correlation detector conditioned on a symbol synchronization error. This conditional error probability is then averaged over the PDF $p(\lambda)$ of the synchronization error to yield the average error probability of the receiver. When compared with a source of perfect symbol synchronization, the performance degradation due to symbol synchronization error can be assessed.

Figures 3.122 through 3.124 present the average probability of symbol error P_s versus E_s/N_0, with σ_λ as a parameter for NRZ, Manchester and Rz data formats, respectively. The value of L_{BS} is equal to the amount that E_s/N_0 (in decibels) must be increased for a given σ_λ to achieve the same error probability as the perfectly synchronized ($\sigma_\lambda = 0$) system.

3.2.13.3 Detection signal losses

In addition to the loss due to the symbol synchronizer tracking-loop timing error, there are waveform losses. Detection signal losses due to filter distortion effects are presented in subsection 3.2.11.1. A loss that has not been considered thus far is data asymmetry. To quantitatively determine the degrading effect of NRZ symbol asymmetry on error rate performance, one must develop a suitable asymmetry model which accurately describes the physical source from which the asymmetry originates. The data symmetry model assumes that +1 NRZ symbols are elongated at $\Delta T/2$ (relative to their nominal value of T seconds) when a

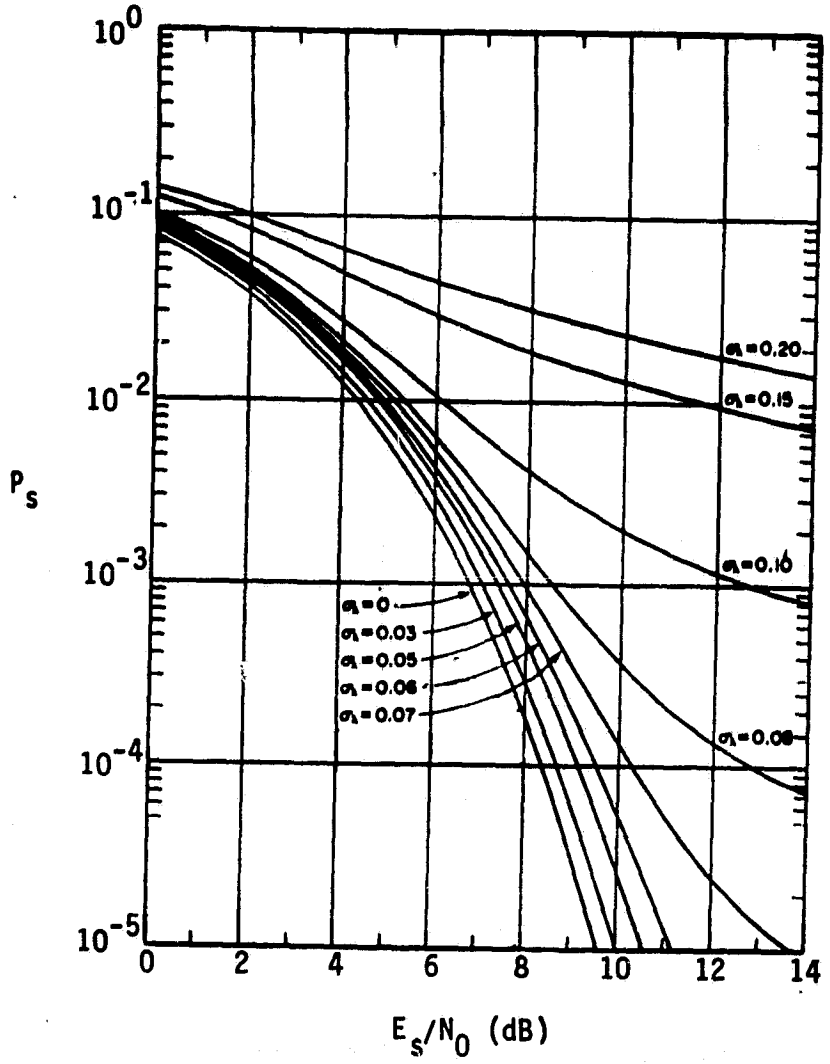


Figure 3.122. Average Probability of Error Versus SNR with Standard Deviation of Symbol Synchronization Error as a Parameter (NRZ)

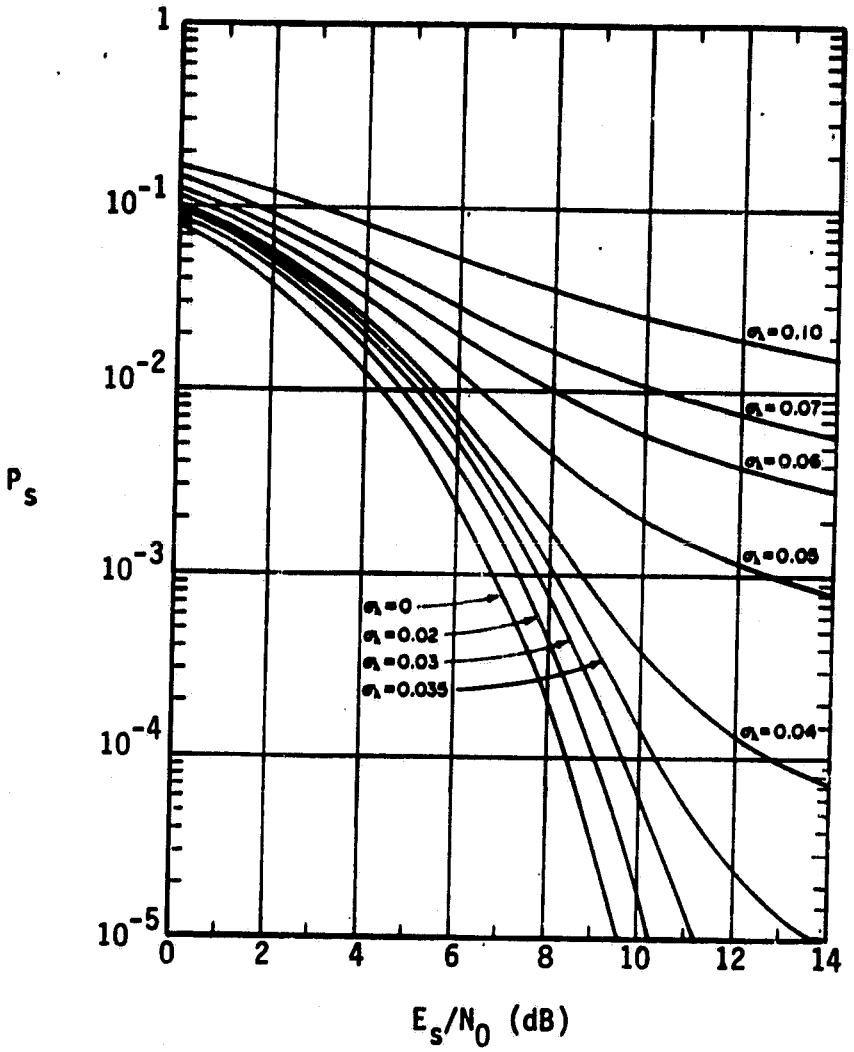


Figure 3.123. Average Probability of Error Versus SNR with Standard Deviation of Symbol Synchronization Error as a Parameter (Manchester Code)

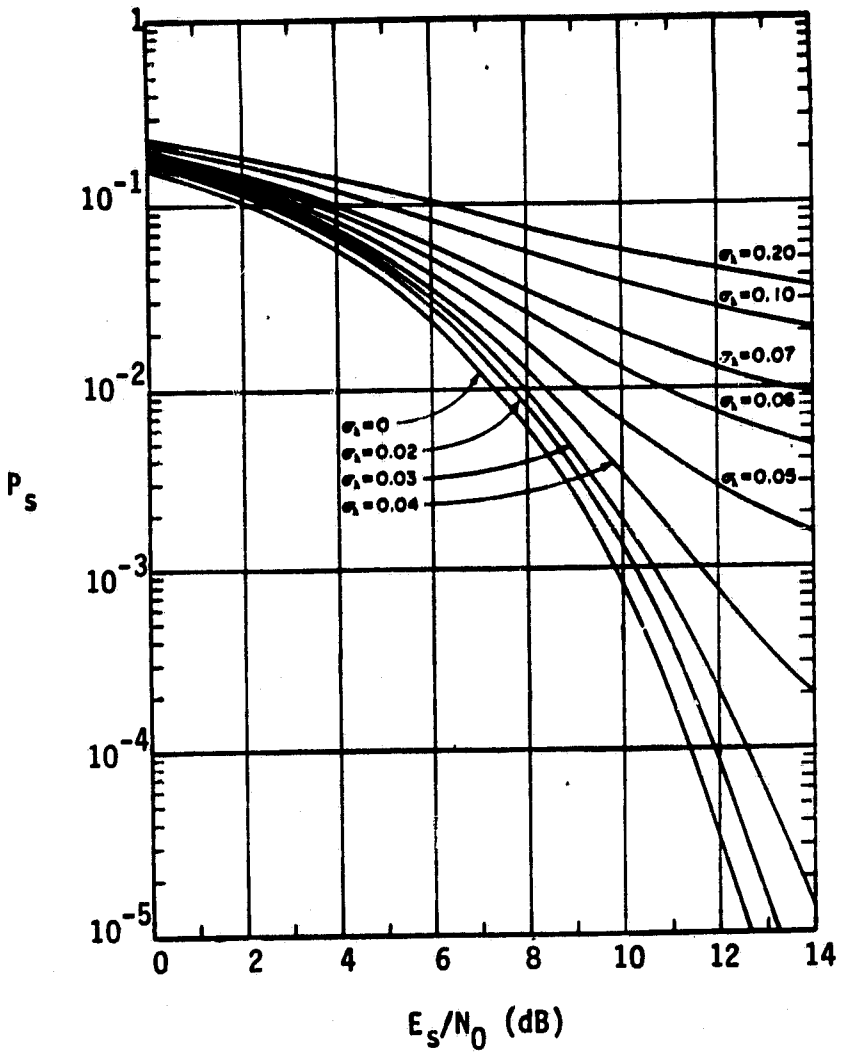


Figure 3.124. Average Probability of Error Versus SNR with Standard Deviation of Symbol Synchronization Error as a Parameter (RZ)

negative-going data transition occurs and -1 symbols are shortened by the same amount when a positive-going data transition occurs.* Otherwise (when no transition occurs), the symbols maintain their normal T-second width. Thus, ΔT represents the relative difference in length between the elongated +1 and shortened -1 symbols. An example demonstrating this model is illustrated in Figure 3.125. Data asymmetry is defined as the difference in length between the shortest and longest pulses in the sequence divided by their sum. This definition gives

$$\text{Asymmetry} \triangleq \eta = \frac{T\left(1+\frac{\Delta}{2}\right) - T\left(1-\frac{\Delta}{2}\right)}{T\left(1+\frac{\Delta}{2}\right) + T\left(1-\frac{\Delta}{2}\right)} = \frac{\Delta}{2} . \qquad (3\text{-}317)$$

In the absence of noise, the timing instants for the inphase integrate-and-dump occur at $t = T(n + \Delta/4)$; $n = 0, \pm 1, \pm 2, \ldots$. Based on the foregoing definition of asymmetry and the accompanying clock misalignment, for random NRZ data, the average symbol error probability associated with hard decisions made on the inphase integrate-and-dump output of the symbol synchronizer is given by [29]:

$$P_s = \frac{5}{8} Q\left(\sqrt{2E_s/N_0}\right) + \frac{1}{4} Q\left(\sqrt{2E_s/N_0}\,(1-\eta)\right) + \frac{1}{8} Q\left(\sqrt{2E_s/N_0}\,(1-2\eta)\right) . \qquad (3\text{-}318)$$

The asymmetry performance loss, L_{DA}, is the additional E_s/N_0 required due to asymmetry to produce the same value of symbol error probability when $n = 0$ $\left(\text{i.e., } P_s^0\right)$, where

$$P_s^0 = Q\left(\sqrt{2E_s/N_0}\right) \qquad (3\text{-}319)$$

It is assumed in (3-318) that the symbol synchronizer produces a perfect clock which is locked up with a misalignment equal to half the asymmetry. The losses due to asymmetry and symbol synchronizer timing jitter do not add directly.

---

*Due to symmetry in the data itself, it is immaterial whether the elongated pulse is of positive or negative polarity, and vice versa for the shortened pulse.

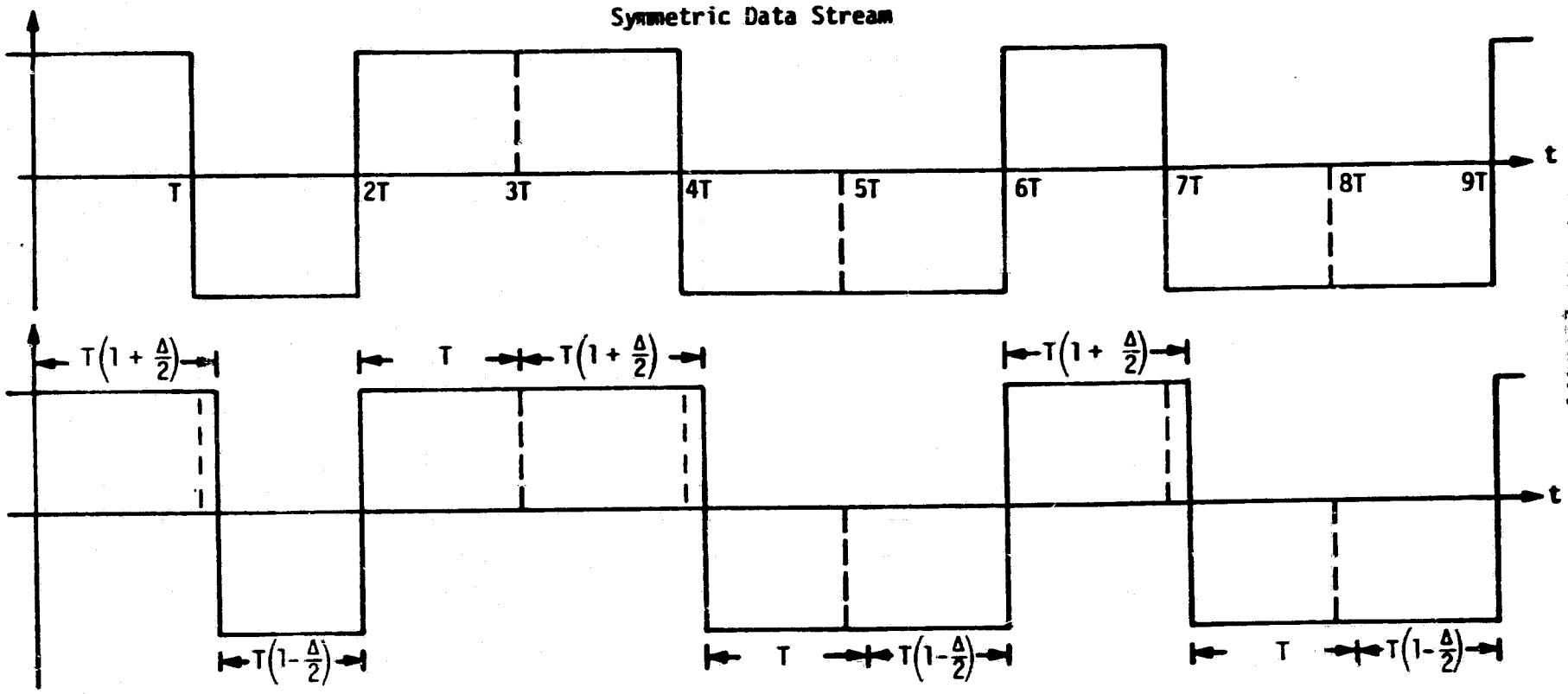


Figure 3.125. Asymmetric Data Stream Definition

An approach to assessing SNR degradation due to both asymmetry and symbol synchronization misalignament is to assume a model probability density function $p(\lambda)$ (PDF) for λ and average the condition error probability $P_s(\lambda)$ over this PDF to obtain the average error probability performance. In this regard, a Tikhonov PDF for $p(\lambda)$ which is entirely characterized in terms of the variance σ_λ^2 of the synchronization error is used. Thus, for NRZ data,

$$p(\lambda) = \frac{\exp\left[\cos 2\pi\lambda/(2\pi\sigma_\lambda)^2\right]}{I_0\left[(1/2\pi\sigma_\lambda)^2\right]} \; ; \quad |\lambda| \leq \frac{1}{2} \; . \tag{3-320}$$

For random NRZ data with equiprobable symbols, the average probability of error conditioned on the misalignment (λT) of the symbol synchronization clock relative to its nominal position is given by

$$P_s(\lambda) = \begin{cases} \frac{5}{8} Q\sqrt{2E_s/N_0} + \frac{1}{8} Q\left[\sqrt{2E_s/N_0}\,(1-\eta-2|\lambda|)\right] \\ + \frac{1}{8} Q\left[\sqrt{2E_s/N_0}\,(1-\eta+2|\lambda|)\right] + \frac{1}{8} Q\left[\sqrt{2E_s/N_0}\,(1-2\eta)\right] ; \\ \qquad\qquad 0 \leq |\lambda| \leq \frac{\eta}{2} \\ \frac{1}{2} Q\sqrt{2E_s/N_0} + \frac{1}{4} Q\left[\sqrt{2E_s/N_0}\,(1-\eta-2|\lambda|)\right] \\ + \frac{1}{4} Q\left[\sqrt{2E_s/N_0}\,(1-\eta+2|\lambda|)\right] ; \quad \frac{\eta}{2} \leq |\lambda| \leq \frac{1}{2} \end{cases} \tag{3-321}$$

Thus, using (3-320) and (3-321), the average error probability can be computed from

$$P_s = \int_{-1/2}^{1/2} P_s(\lambda)\, p(\lambda)\, d\lambda \; . \tag{3-322}$$

Figures 3.126 and 3.127 illustrate P_s [as computed from (3-322)] versus E_s/N_0 in dB with σ_λ as a parameter for asymmetry values of 3% and 6%, respectively.

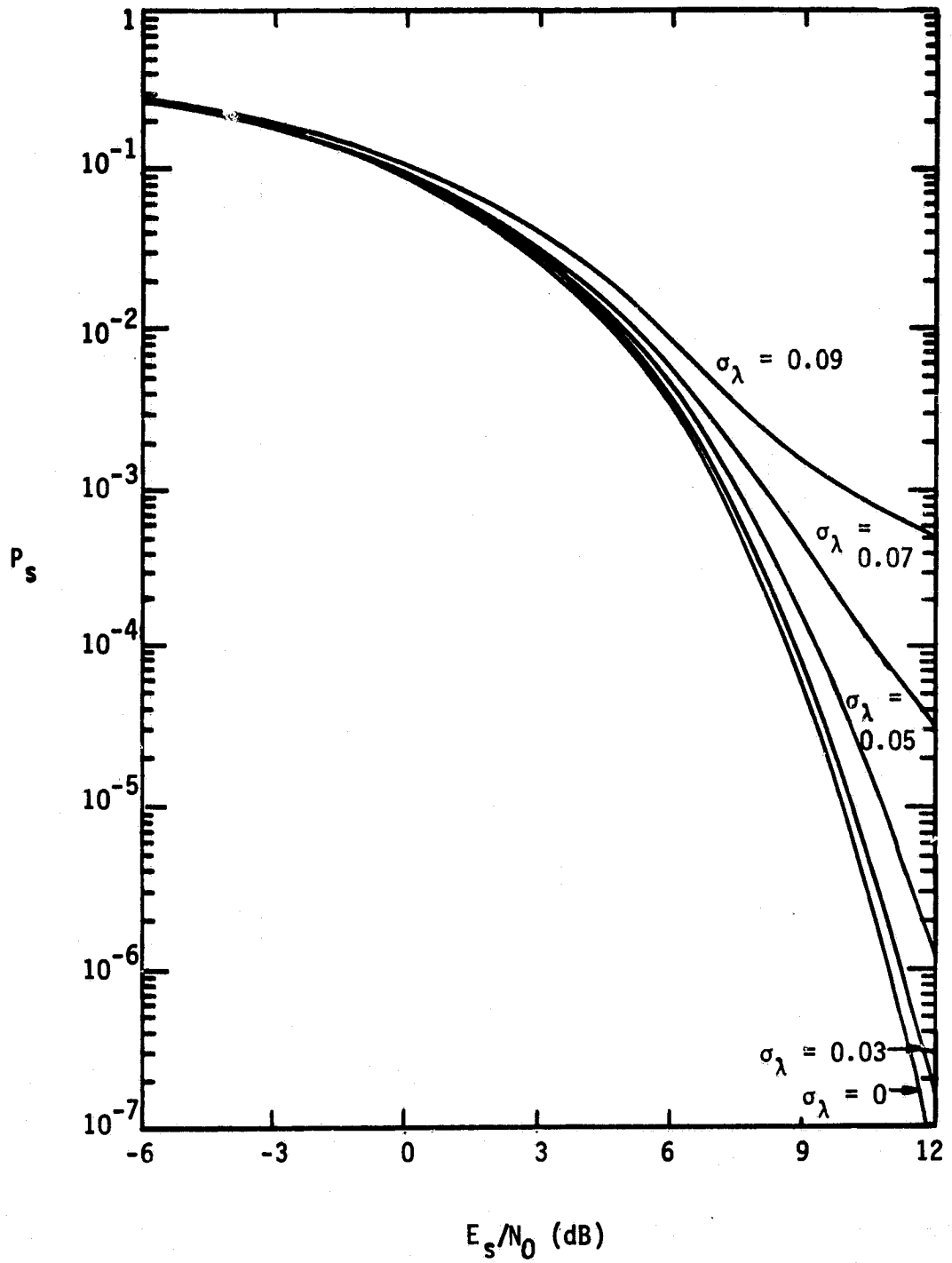


Figure 3.12b. Average Error Probability with Symbol Synchronization Error as a Parameter

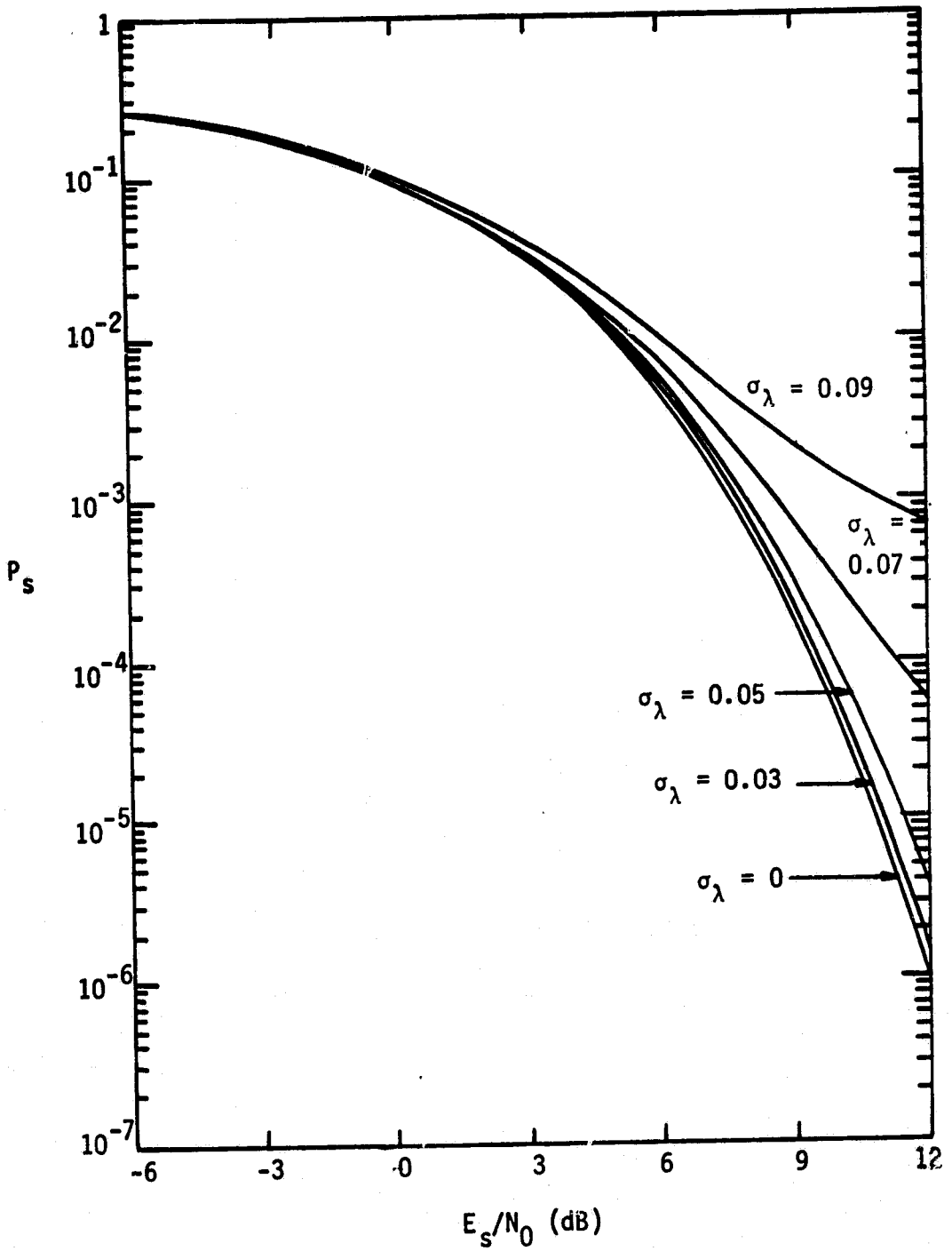


Figure 3.127. Average Error Probability With Symbol Synchronization Error as a Parameter (6% Asymmetry), NRZ Data

Data detectors that employ DC restoration* tend to reduce the degrading effects of data asymmetry. The effect of DC restoration on data detection is most easily accounted for by artificially shifting the decision threshold (nominally at zero) against which the matched-filter output is compared. The amount of this artificial shift in threshold depends upon the specific way in which DC restoration comes about.

The simplest method of achieving DC restoration is to capacitively couple the input signal to the symbol synchronizer.. In this case, the artificial threshold shift equals the DC component of the asymmetric data waveform in front of the capacitor which, for random data with transition density D, is

$$\Delta_t = \eta D\sqrt{E_s/T} \ , \tag{3-323}$$

where $\sqrt{E_s/T}$ is the data pulse amplitude in Figure 3.125. Computing the matched-filter output for the eight possible three-symbol sequences made up of the present, preceding and succeeding symbols, and shifting these outputs by Δ_t gives the result for the error probability performance of asymmetric NRZ data with DC restoration by capacitive coupling, namely,

$$P_s = \frac{1}{2} Q\left[\sqrt{2E_s/N_0}\,(1 - \eta D)\right] + \frac{1}{8} Q\left[\sqrt{2E_s/N_0}\,(1 - 2\eta + \eta D)\right]$$
$$+ \frac{1}{4} Q\left[\sqrt{2E_s/N_0}\,(1 - \eta + \eta D)\right] + \frac{1}{8} Q\left[\sqrt{2E_s/N_0}\,(1 + \eta D)\right] . \tag{3-324}$$

For equiprobable data symbols (D = 0.5), (3-324) simplifies to

$$P_s = \frac{3}{4} Q\left[\sqrt{2E_s/N_0}\left(1 - \frac{\eta}{2}\right)\right] + \frac{1}{8} Q\left[\sqrt{2E_s/N_0}\left(1 - \frac{3\eta}{2}\right)\right] + \frac{1}{8} Q\left[\sqrt{2E_s/N_0}\left(1 + \frac{\eta}{2}\right)\right] . \tag{3-325}$$

*DC restoration refers to the process by which the DC value of the asymmetric data waveform is forced to zero.

Comparing (3-325) with (3-318), note that the effect of DC restoration is to compensate for the data asymmetry by shifting the effective decision threshold away from the shortened symbols.

For a given value of asymmetry, the value of E_s/N_0 required to obtain P_s [as computed from (3-325)] equal to 10^{-5} can be calculated. Comparing this value of E_s/N_0 with that obtained from (3-319) for the same P_s gives the SNR degradation for asymmetric NRZ data with DC restoration by capacitive coupling. Figure 3.128 illustrates this NRZ degradation versus asymmetry along with the comparable results obtained from (3-318) corresponding to no DC restoration (direct coupling).

Another method of achieving DC restoration, which depends specifically on the symbol synchronizer implementation itself, is to require the matched-filter output to have zero crossings at the center of each symbol period that starts with a data transition. In this case, the effective shift in decision threshold relative to its nominal (zero) value is

$$\Delta_t = \eta\sqrt{E_s/T} \quad . \tag{3-326}$$

Comparing (3-326) with (3-323), it can be concluded that the error probability for this method of DC restoration is given by (3-324) with D = 1, i.e.,

$$P_s = \frac{5}{8} Q\left[\sqrt{2E_s/N_0}\,(1-\eta)\right] + \frac{1}{4} Q\left[\sqrt{2E_s/N_0}\right] + \frac{1}{8} Q\left[\sqrt{2E_s/N_0}\,(1+\eta)\right] . \tag{3-327}$$

Again by determining those values of E_s/N_0 required to obtain $P_s = 10^{-5}$ for various values of asymmetry, one can compute the SNR degradation for DC restoration based on symbol timing. The results of these calculations are illustrated in Figure 3.129, along with experimental test results taken in the Electronic Systems Test Laboratory (ESTL) at JSC for the sake of comparison. It is to be noted that the experimental results include the effect of bandlimiting, whereas the theory as predicted by (3-327) in no way accounts for this effect. Furthermore, the data detector used in the experimental setup is not a true

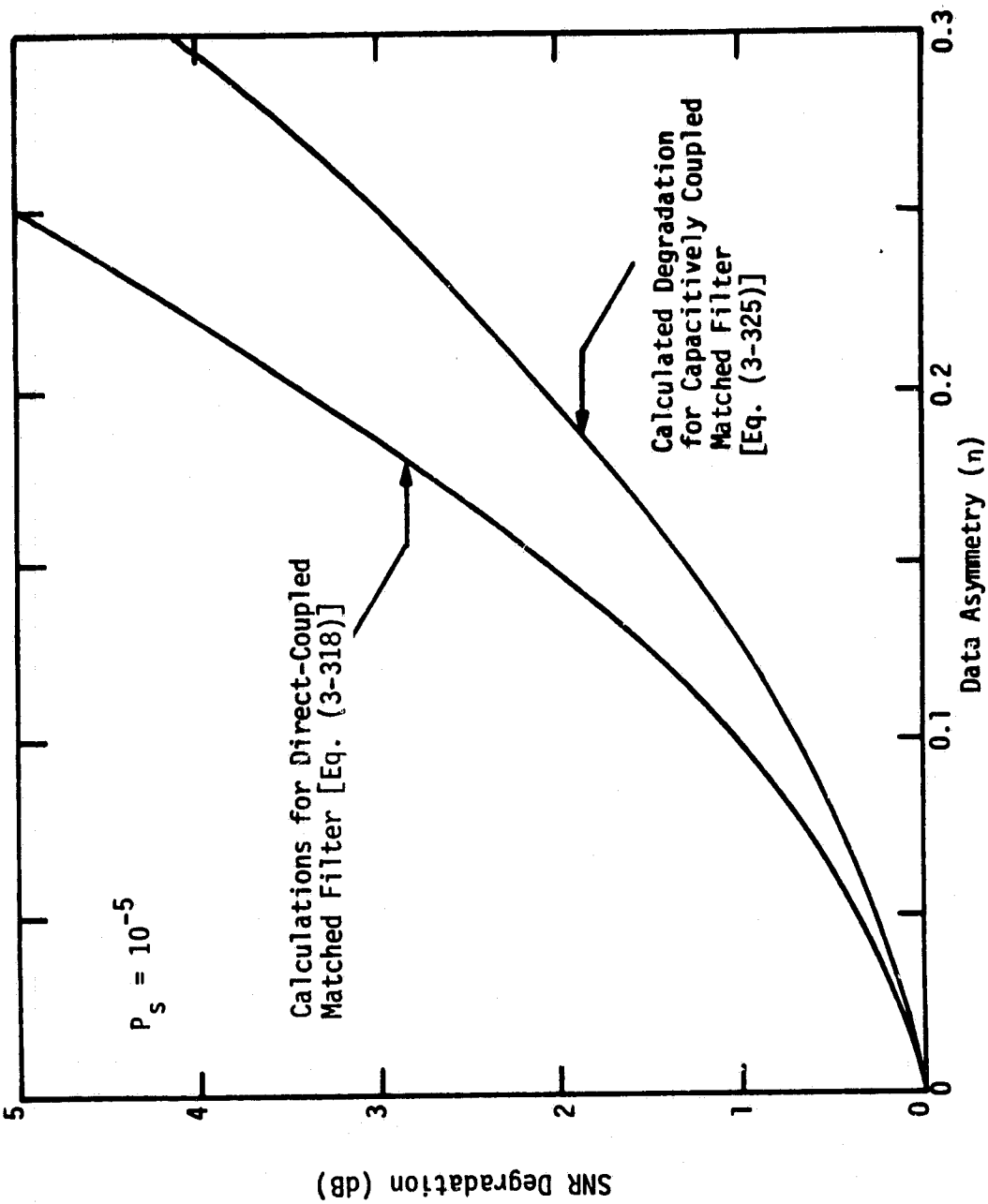


Figure 3.128. Performance Degradation for Capacitively Coupled Matched Filters; Random Data (D = 0.5)

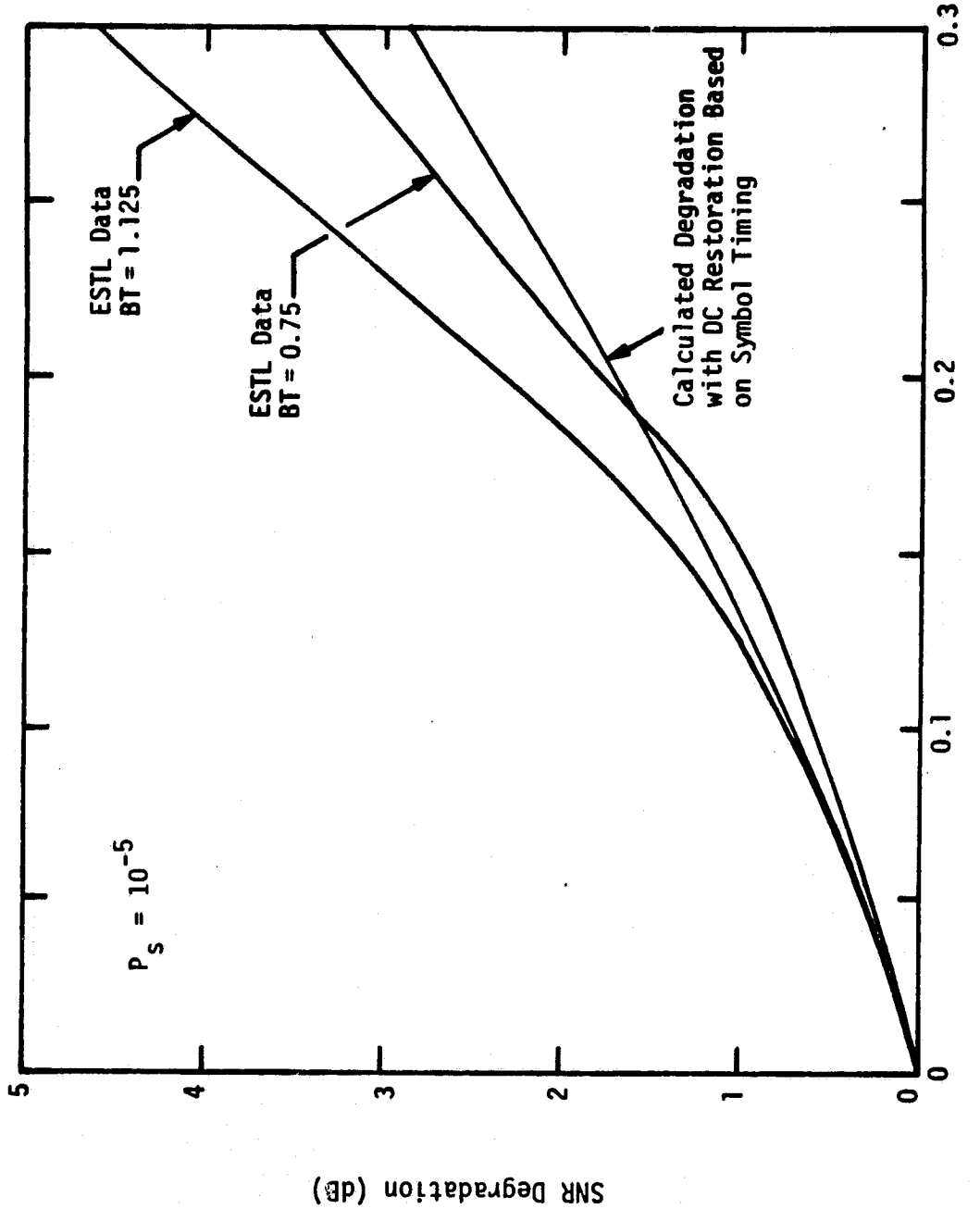


Figure 3.129. Performance Degradation for DC Restoration Based on Symbol Timing

matched filter as is assumed for the analytical model. Surprisingly enough, however, the analytic and experimental results show reasonably good agreement.

3.2.14 Digital Data Decoding

Digital coding was introduced in subsection 3.2.2.1. The coding gain G_C is the decrease in E_b/N_0 required to achieve a given probability of bit error, P_b, from the E_b/N_0 required for the uncoded channel. This subsection presents G_C and the performance curves (i.e., P_b versus E_b/N_0) for various codes and decoders. The two basic classes of error-correcting codes considered are block codes and convolutional codes. Against random channel errors, convolutional codes have consistently been shown [43] to outperform block codes with coding and decoding equipment of equal complexity. For instance, Figure 3.130 presents a comparison of rate 1/2 codes for PSK modulation. The block codes shown are the best available in terms of decoding simplicity and coding gain, yet the convolutional codes still outperform the block codes. For comparison purposes, the hard decision decoder for the (24,12) Golay code is about as complex as the K = 5 hard decision Viterbi decoder, while the (48,24) quadratic residue decoder is about equivalent to the K = 7 hard decision Viterbi. The new (23,12) soft-decision decoder for the Golay code is more complex than the K = 7 soft-decision decoder. The most complex mechanization considered for a desired output BER of 10^{-3} is the soft-decision sequential decoder. The complexity of this mechanization is due to the large input buffer and fast processor needed to prevent buffer overflow and loss of message.

From Figure 3.130, it is clear the convolutional codes do indeed outperform block codes of equivalent mechanization complexity. There are several reasons. The most basic reason is that Viterbi decoding actually performs maximum-likelihood decoding, and large buffer sequential decoding very nearly approximates it. This means that no matter what channel response is received, the decoder will find the codeword which is the closest match to the received sequence. Practical block decoders, on the other hand, are limited in the sense that the received sequence must not be different than the true codeword

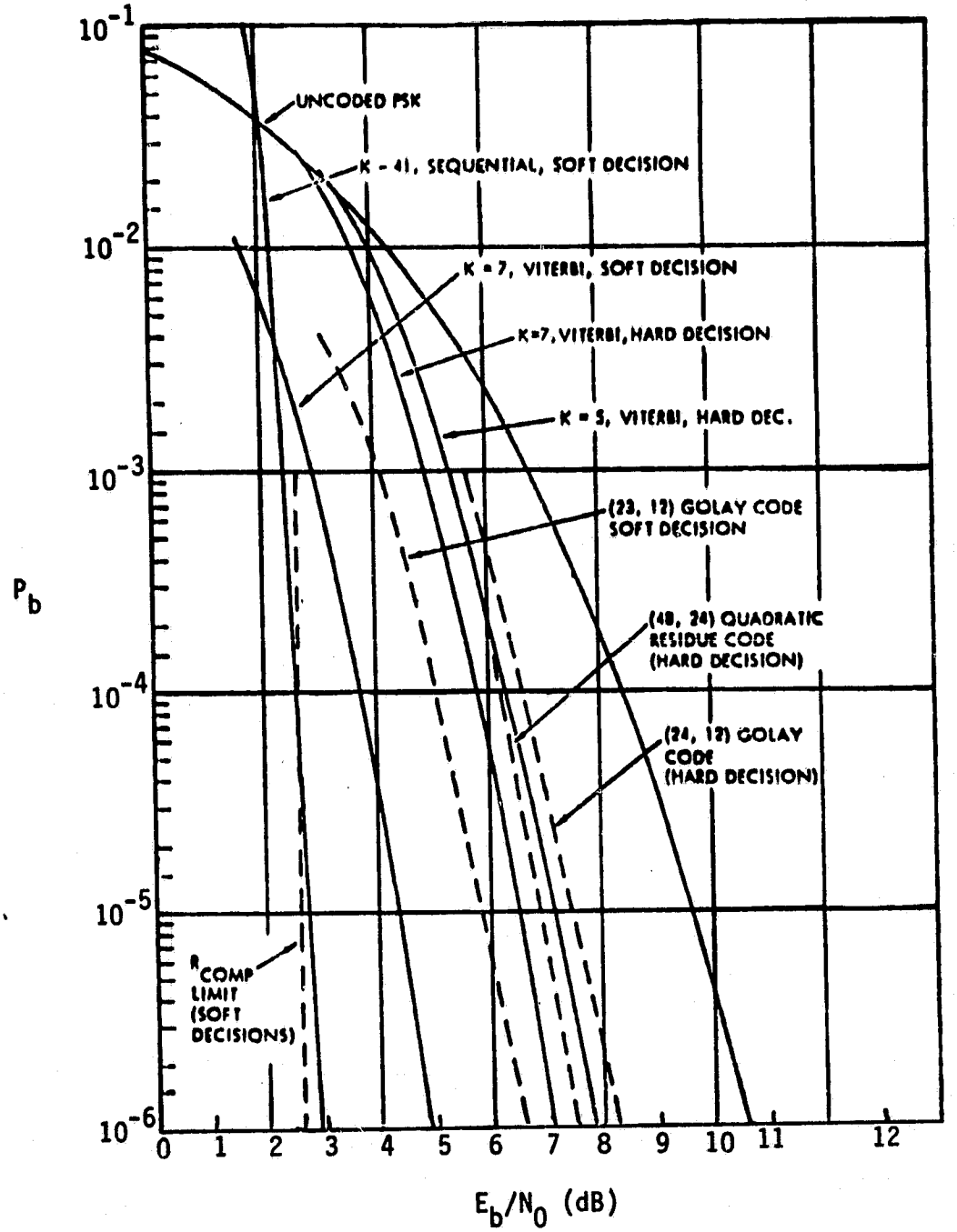


Figure 3.130. Rate 1/2 Code Choices for PSK Modulation

in more than t symbols before decoding correctly. This type of decoding is, in general, limited to half the number of identical symbols between the two closest codewords in the codebook. So if the block code is a three-error-correcting code and a received sequence occurs which has four symbols different from the correct codeword, the decoder makes an error, even if the correct codeword is the closest match to the received sequence.

The second advantage of convolutional over block codes is the ability to perform simple soft-decision decoding. Until very recently, block codes were constrained to hard-decision decoding in which the received sequence is hard quantized and the decoder had to work with this binary message stream. On the other hand, convolutional codes, using either Viterbi or sequential decoding, can operate on "soft" decisions from the channel. For binary channels, this means that a measure of the probability of the transmitted bit being a one or zero is used directly in the decoding. Simulations of the Viterbi algorithm [44] and the sequential algorithm [45] have shown that an incremental 2 dB in coding gain may be realized by using soft decisions as opposed to hard decisions. Recent theoretical work on block codes [46] has made possible soft-decision decoding of these codes. The performance of soft decoding the (23,12) Golay codes is also shown in Figure 3.130. Note that the use of soft decoding in this case only gains about 1.5 dB (not maximum-likelihood decoding) and that the resulting decoder is slightly more complex than the full K = 7 soft Viterbi decoder which does 1.5 dB better.

Block decoders also have a disadvantage in the area of synchronization. Block codes must be synchronized to the edge of each codeword because the encoding/decoding is always performed on a fixed-length basis. Although this does not affect performance, it does represent additional complexity in the decoder.

Three classes of convolutional decoding algorithms are in general use: the Viterbi algorithm, the Fano sequential algorithm, and the feedback decoder. The Viterbi decoding algorithm performs optimum decoding and generates the maximum-likelihood codeword based on the whole received sequence. Practical Viterbi decoders operate

on a moving window of five constraint lengths of data and perform within 0.05 dB of the theoretical optimum limit. The algorithm decodes by postulating, at each clock time, all 2^K possible encoder states for a "constraint length" K encoder. Each postulated state is examined for all possible ways to reach that state, and the likelihood of each way is evaluated. This evaluation consists of summing an incremental metric (log likelihood) of the received message conditioned on the old state to new state hypothesis plus the stored minimum metric of getting to the hypothesized old state. This total metric is compared for all ways of getting into the postulated state, and the minimum metric entry is determined. This becomes the stored "minimum metric" of getting into the postulated state to be used during the next clock period (iteration of the algorithm). Since all possible states must be examined at each time, the decoder complexity grows exponentially in constraint length K. For this reason, practical Viterbi decoders are limited to codes with constraint lengths on the order of 7. Figures 3.131 and 3.132 present the performance of convolutional encoding/Viterbi decoding in terms of P_b versus E_b/N_0. Figure 3.131 shows how the required E_b/N_0 to achieve a given P_b can be reduced by increasing the constraint length K. Figure 3.132 illustrates the decrease in E_b/N_0 to achieve a given P_b by decreasing the code rate R. Figures 3.133 through 3.135 compare the performance of biorthogonal (codes with rate $R = 1/3$) convolutional encoding/Viterbi decoding for various constraint lengths. Note in Figure 3.133 that a $K = 8$ biorthogonal block code is needed to give approximately the P_b as a $K = 4$ Viterbi-decoded convolutional code.

Sequential decoders are suboptimum "valley-seeking" decoders which attempt to follow along the locally best symbols of a codeword as long as good agreement is observed between the local symbols and the received sequence. Where poor agreement is noted, back searches are initiated to find codewords with better agreement. For fast processors and large input buffers (during back searches, the inputs continue), very deep back searches can be tolerated, and the undetected error performance of the suboptimum sequential algorithm approaches optimum. However, the time to decode each bit is still random,

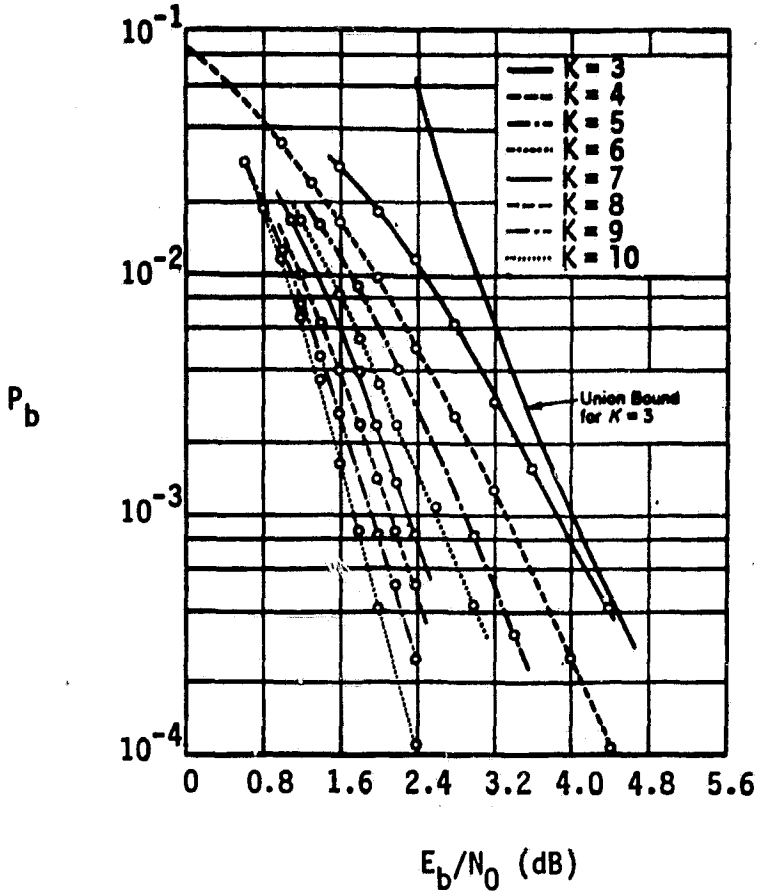


Figure 3.131. Bit Error Probabilities of R = 1/3 Convolutional Codes

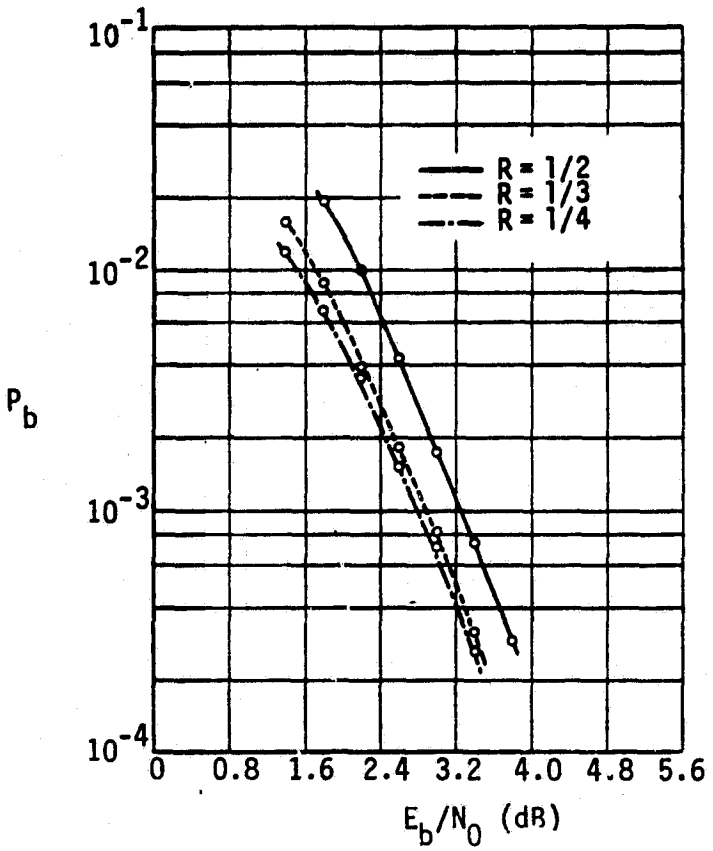


Figure 3.132. Bit Error Probabilities of K = 5 Codes

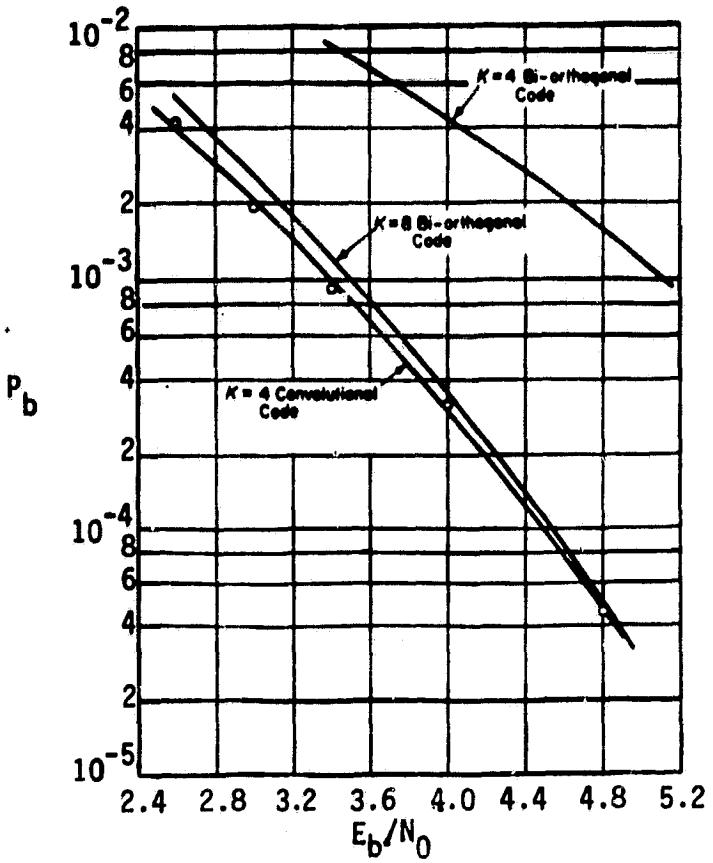


Figure 3.133. Bit Error Probabilities for K = 4 Convolutional and Biorthogonal Codes and Comparison with K = 8 Biorthogonal Code

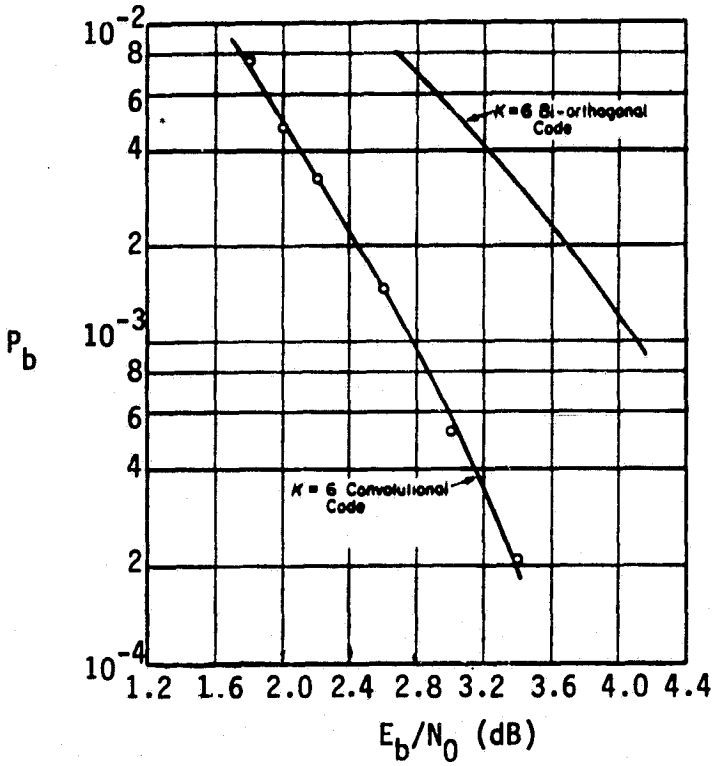


Figure 3.134. Bit Error Probabilities for K = 6 Convolutional and Biorthogonal Codes

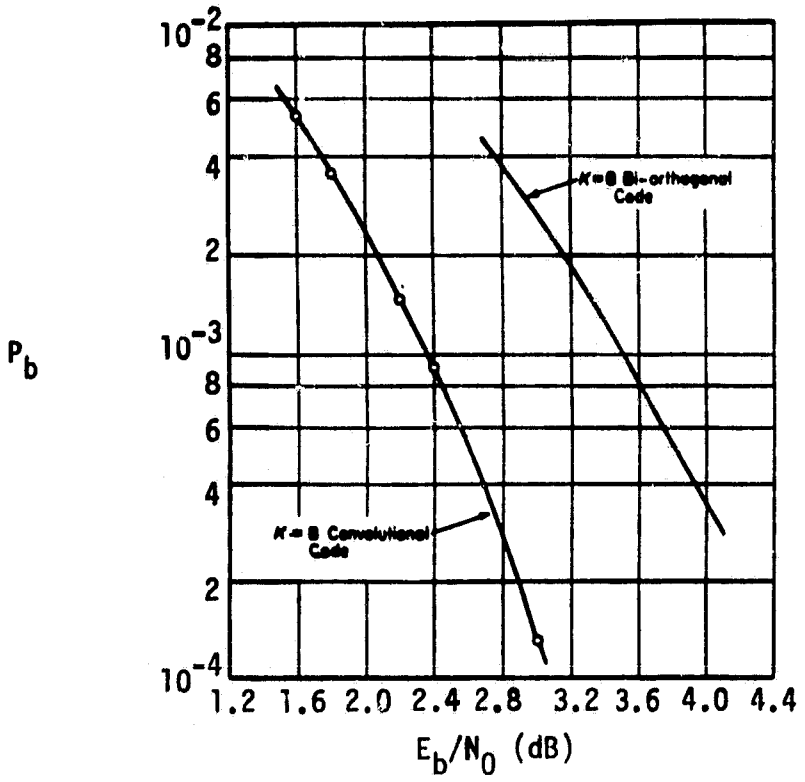


Figure 3.135. Bit Error Probabilities for K = 8

depending on the number of back searches needed, and the inputs must be buffered to allow for this delay. If the input buffer capacity is exceeded, then a portion of data is lost, causing a gap in the output stream. For $P_b = 10^{-3}$ voice, this may be unacceptable, or at least less desirable than sending through a message, even if it is noisy. The problem can be circumvented by using a systematic code in which one of the channel bits is the uncoded data stream, and this stream may be directly output (still noise corrupted) when buffer overflow occurs. However, the fact still remains that the sequential decoder has large complexity in the fast processor and large backup buffer.

Sequential decoding complexity is independent of constraint length, so very long codes (K = 40 or above) are used for extremely good error-correcting capability. The source of errors is thereby restricted to the overflow problem mentioned above. For fixed processor rate and

buffer size, the likelihood of buffer overflow can be found by calculating the probability that the number of computations will exceed the product of computation rate times the time to fill the buffer. This quantity has been evaluated and found to exhibit a Pareto distribution. For a desired ouput bit error rate, the sequential decoder speed and buffer are sized so that the probability of overflow times the number of bits lost due to overflow meets the requirements for P_b.

There is a fundamental limitation on the amount of improvement available, however. Above a certain channel bit error rate, the expected number of computations needed to decode goes to infinity. By evaluating the expected number of computations as a function of error rate, the information rate at which the expected value becomes infinite may be determined. This value is known as R_{comp}.

Note that the difference in performance between the rate 1/2, K= 7, soft Viterbi decoder and the rate 1/2, K= 41, soft sequential decoder is less than 0.5 dB at $P_b = 10^{-3}$. Similar analysis at rate 1/3 shows that K= 7 soft Viterbi decoding is within 0.3 dB of the R_{comp} limit at $P_b = 10^{-3}$.

Another type of suboptimal decoding is feedback decoding, a special case of which is threshold decoding [47]. Although the performance of Viterbi and sequential decoding is generally superior to that of feedback decoding, the latter has the distinct advantage of being easy to implement. The operation of a feedback decoder is described as follows. The first $L_t n$ symbols (L_t is a parameter of the decoder) are examined, and the path yielding the largest metric with respect to the first $L_t n$ symbols of the received sequence is selected. At this point, the first information bit corresponding to the first n symbols on the maximum metric path chosen is decoded. This decision uniquely moves the decoder to a specific encoder state. Starting there, all paths of length $L_t n$ symbols emanating from that node are considered and, again, a maximum metric selection is made--now with respect to the sequence of $L_t n$ received code digits whose first is the n+1st. Then the second information bit is decoded. This process continues in the manner just described. The technique is called feedback decoding because each decoding decision on an information bit is fed back to affect future decisions. Hence, decoding schemes of this type possess the undesirable

property of error propagation caused by incorrect decisions. For a large class of convolutional codes, however, the error propagation is very short.

It should be clear that the decoding accuracy of a feedback decoding scheme improves as the parameter L_t increases. Unfortunately, however, the complexity of the decision device also grows as L_t gets large. Threshold decoding provides a partial solution to this paradox, but it is only useful for a specially constructed class of convolutional codes and only for moderate values of L_t. Further details of this specific type of feedback decoder are well documented [47].

While the previous discussion presented the performance for PSK modulation, it is not necessary to restrict the specific form of modulation. The results presented apply to a general class of memoryless channels, which include additive white Gaussian noise channels, using polyphase and MFSK modulation. However, when soft decisions are employed, the performance depends on the PDF of the symbol detector. In many applications only hard decisions are available. In this case, the decoding performance is specified in terms of input symbol error probability, P_s, to output bit error probability, P_b. Figure 3,136 presents the hard decision coding performance of several codes. The convolutional encoding/Viterbi decoding performance is for $K = 7$ and rates $R = 1/2$ and $1/3$. The convolutional encoding/binary feedback decoding is $R = 1/2$ and $K = 10$. The convolutional encoding/sequential decoding is $R = 1/2$ and $K = 41$. It is important to note that the results presented in Figure 3.136 are valid for any memoryless channel, regardless of the type of modulation. The coding gain G_C or E_b/N_0 required to achieve a given P_b can be determined by finding the P_s that corresponds to the given P_b for a particular decoder. Let E_s/N_0 be required to achieve P_s for the type of modulation used. Then, the required (E_b/N_0) to achieve P_b for the particular decoder is

$$(E_b/N_0)_C = \frac{1}{R} E_s/N_0 \quad . \tag{3-328}$$

Let $(E_b/N_0)_U$ be required to achieve P_b with no coding; then

$$G_C = (E_b/N_0)_U - (E_b/N_0)_C \quad . \tag{3-329}$$

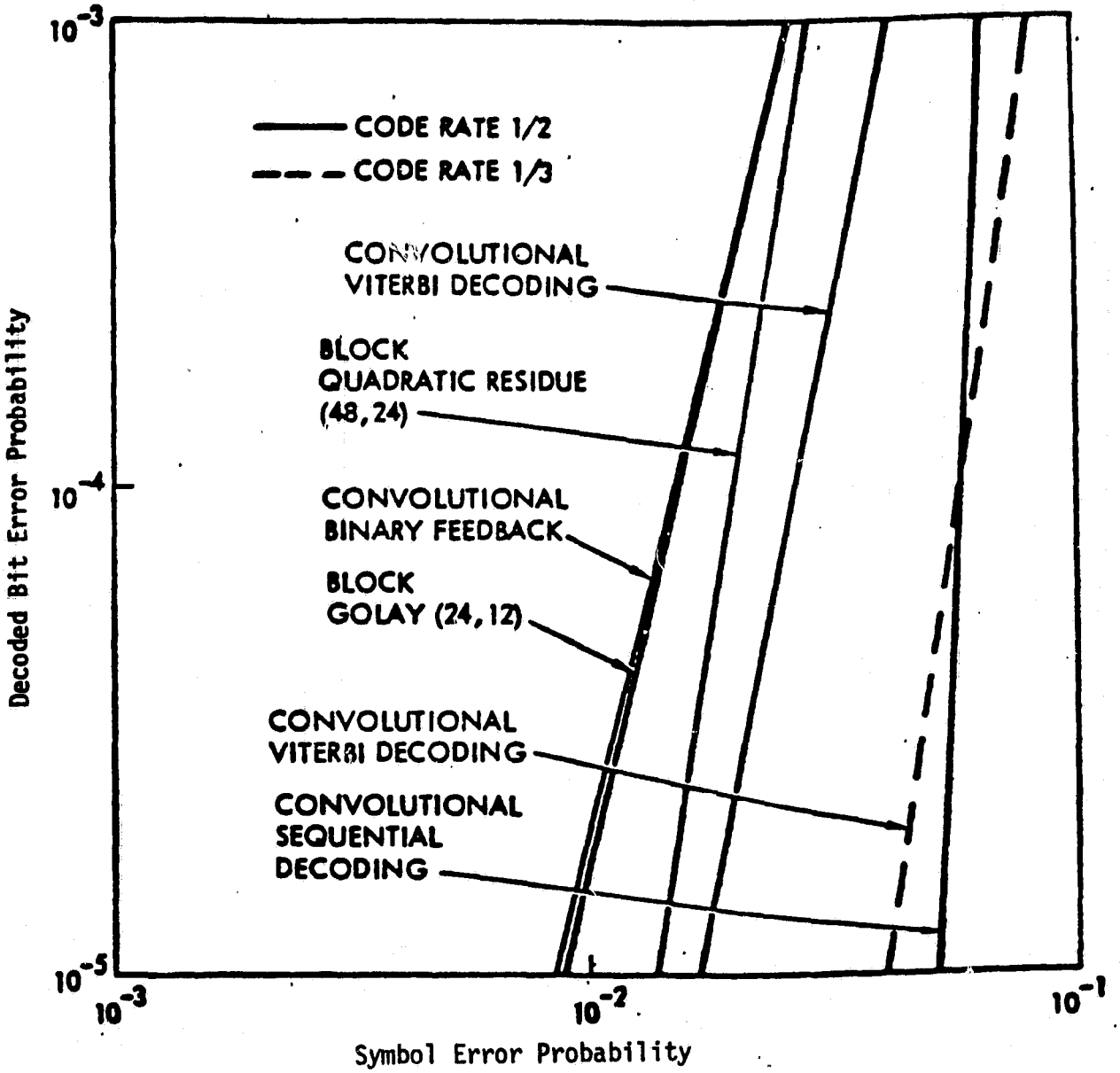


Figure 3.136. Hard-Decision-Coding Performance

3.2.15 Tandem Link Considerations

A tandem link is defined in this handbook as one that contains two or more independent and accountable noise sources. Figure 3.137 shows the general model for a tandem link with three additive noises. Each noise source is introduced between a transmitter and receiver pair. The noise at the output of receiver #1 depends only on noise #1, while the noise output of receiver #2 is a function of both noise #1 and noise #2. Finally, the end output noise depends upon all three noise sources.

An example of the generalized tandem link depicted in Figure 3.137 would be the payload-to-Orbiter-to-TDRS-to-ground link shown in Figure 2.3. Functionally, this link is comprised of Figures 2.9 and 2.13 combined. An identification of the Figure 3.137 blocks with subsystems would be:

Transmitter #1 = Payload transponder transmitter
Receiver #1 = Payload interrogator
Transmitter #2 = Network transponder transmitter
Receiver #2 = TDRS S-band receiver
Transmitter #3 = TDRS S-band transmitter
Receiver #3 = Ground station receiver.

The next two subsections discuss two particular tandem link models that are useful to the design control procedure.

3.2.15.1 The TDRS model

A functional block diagram of the TDRS return link channel may be found in Figure 2.13. The analytical model appears in Figure 3.138.

Two channels comprise the TDRS link. One of two, the TDRS-to-ground channel is relatively strong by virtue of the fact that high-gain antennas are employed on the satellite and at the ground station. On the other hand, the Orbiter-to-TDRS link tends to be SNR limited because the Orbiter's antenna has a comparatively much lower gain. Thus, when considering the composite TDRS channels, the effects of noise from the TDRS to the ground may be neglected.

What remains to be accounted for is the signal-to-noise spectral density (S/N_0) loss associated with the TDRS itself. Mathematically,

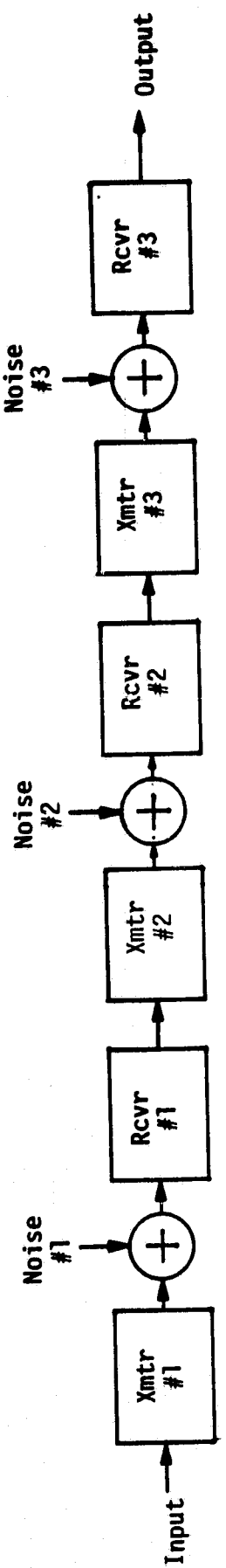


Figure 3.137. Generalized Tandem Link

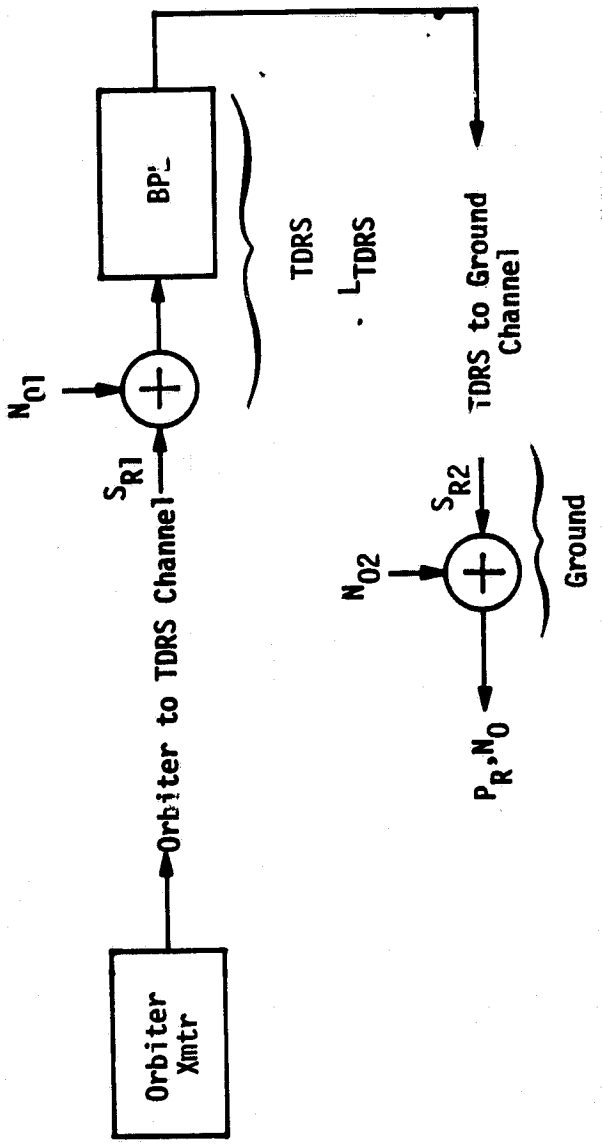


Figure 3.138. TDRS Link Analytical Model

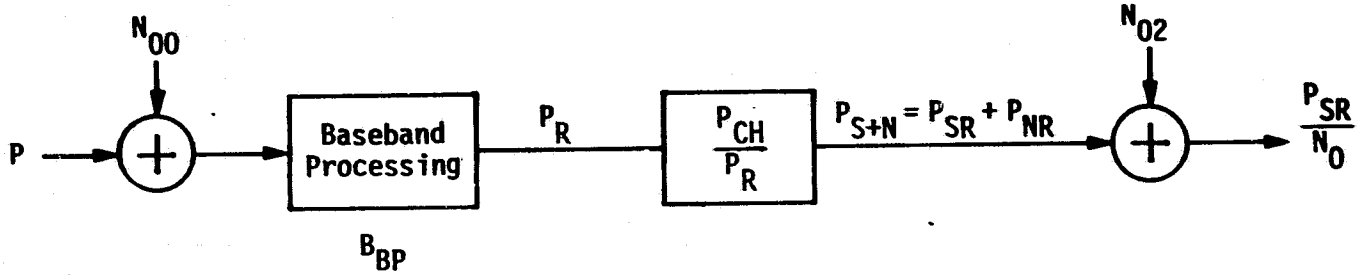
the satellite may be modeled as a bandpass limiter which transforms the input signal-plus-noise into a constant-envelope bandpass waveform. The effect of this operation insofar as the throughput is concerned is to decrease the input S/N_0 by a factor L_{TDRS}. L_{TDRS} is a function of the SNR at the limiter input, and analyses [48] have shown that the loss is fairly constant for low SNR. Therefore, when dealing with the TDRS return link, L_{TDRS} is assigned a value of -2.0 dB, indicative of a low SNR condition.

The TDRS forward link may be analytically modeled akin to the return link. In this case, however, the SNR into the limiter is relatively large and, as a result, the value assigned to L_{TDRS} is -1.0 dB.

3.2.15.2 Accounting for multiple noise sources

The bent-pipe link (see subsections 2.1.2.4 and 2.3.2.2) is a prime example of a tandem link where two independent noise sources must be accounted for. An analytical model is shown in Figure 3.139. The left-hand noise source, N_{00}, is to be associated with the PI receiver, while the other noise, N_{02}, is the TDRS receiver. A block labeled Baseband Processing is a combination of the PI receiver and KuSP filtering and amplitude regulation circuits. The block P_{CH}/P_R accounts for the Ku-band transmitter power allocation to the particular bent-pipe channel under consideration (see subsection 2.3.2).

The model may be explained as follows. A signal with power P and noise with spectral density N_{00} are added at the input to the baseband processor. The processor functions to filter and amplitude scale the signal-plus-noise. Amplitude scaling may be comprised of (1) a linear RMS or peak-to-peak regulator or (2) a hard-limiter (clipper). Whatever the nature of the amplitude scaling, the output of the processor may be viewed as having a signal component plus noise components in a noise bandwidth of B_{BP}. Total output power of the processor as manifest in the received total power at the TDRS is designated as P_R. This power is decreased by the factor P_{CH}/P_R to account for the power allocation to the particular bent-pipe channel under consideration, the effective signal-plus-noise power being given the symbol P_{S+N}. The individual signal and noise power components are then obtained respectively by



$$P_{SR} = (P_{S+N})\left(\frac{P_{SR}}{P_{S+N}}\right)$$

$$P_{NR} = (P_{S+N})\left(\frac{P_{NR}}{P_{S+N}}\right)$$

$$N_0 = \frac{P_{NR}}{B_{BP}} + N_{02} = N_{01} + N_{02}$$

Figure 3.139. Bent-Pipe Dual-Noise-Source Analytical Model

multiplication with the factors P_{SR}/P_{S+N} and P_{NR}/P_{S+N}. These factors have values that are dependent upon the nature of the baseband processing. (NOTE: As of the first edition of the handbook, analytical and graphical data which relate these factors to the physical subsystems involved was not available. This information is planned to be supplied in subsequent editions of the handbook.) Finally, the tandem equivalent noise spectral density, N_0, is obtained by adding the bandwidth-normalized received noise power component, $P_{NR}/B_{BP} = N_{01}$ (this creates the noise spectral density form of the received noise power), with the TDRS receiver noise spectral density, N_{02}. (Note that this addition is accomplished with absolute forms of the noise spectral densities, not their dBW/Hz equivalents.) Table 4.6 in section 4.2 shows the design control procedure for handling the case of two noise sources.

4.0 THE COMMUNICATION LINK DESIGN BUDGET

4.1 The Nature and Structure of the Link Equation

Section 2.0 discusses the functional composition of the various end-to-end links which comprise the overall payload/Orbiter/ground communication system. The concept of a generalized link block diagram is then introduced in Section 3.0, and the various types of performance parameters which characterize the end-to-end link operations are defined and explained. In this section, the practical means for calculating the expected end-to-end link performance are developed, and guidelines are given which allow a link design budget to be generated for any channel of interest to a payload user.

To begin with, it may be simply stated that calculation of a particular desired performance quantity is obtained through multiplication and division of the separate link parameters, viz.,

$$Z = \frac{X_1 X_2 \cdots X_N}{Y_1 Y_2 \cdots Y_M}, \tag{4-1}$$

where Z is the performance quantity to be calculated and X_n and Y_m are the link parameters. A simple example would be the calculation of the effective power radiated at a transmitting directional antenna. Given that a power amplifier supplies power P_T to a cable having transmission loss L_C which feeds the antenna having on-axis gain G_T, the effective radiated power calculation is

$$P_{radiated} = P_T L_C G_T .$$

A second example would be the determination of receiver SNR given the received signal power P_R, a lossy cable L_C, a receiver noise figure NF, the ambient noise spectral density N_0, and a bandwidth B. The receiver SNR is then calculated by

$$SNR = \frac{P_R L_C}{(NF) N_0 B} .$$

It is much easier to add and subtract numbers than to multiply and divide them. Thus, it is common practice to deal with (4-1) in its logarithmic form. Taking $10 \log_{10}$ of both sides of (4-1) gives

$$10 \log_{10} (Z) = 10 \log_{10} \left[\frac{X_1 X_2 \cdots X_N}{Y_1 Y_2 \cdots Y_M} \right]$$

$$= 10 \log_{10} (X_1) + 10 \log_{10} (X_2) + \cdots + 10 \log_{10} (X_N)$$

$$- 10 \log_{10} (Y_1) - 10 \log_{10} (Y_2) - \cdots - 10 \log_{10} (Y_M) .$$

In order to write this result in more compact form, it is desirable to introduce new symbols for each of the variables $10 \log_{10} ( )$. However, since there are so many parameters for a typical link, symbology is already a problem, and a dual set of symbols to denote absolute and logarithmic quantities would be confusing. Therefore, the same symbols are employed, the above equation being written as

$$Z = X_1 + X_2 + \cdots + X_N - Y_1 - Y_2 - \cdots - Y_M . \qquad (4\text{-}2)$$

If the parameters X_n and Y_m are given in absolute form, then (4-1) is used while, if they appear in logarithmic form (as denoted by their dB or dBx units), (4-2) applies.

At this point, a word of caution is invoked: absolute and logarithmic units must not be mixed in the calculation. This is a common error. When parameters are initially specified or determined from the physical subsystems involved, some will often appear in absolute units and others in logarithmic form. Since the usual procedure is to construct the link design budget using parameters in logarithmic form, all quantities given with absolute units must be converted before the link performance calculation is made.

An end-to-end link equation, for which (4-2) is the archetype, involves many parameters. Thus, a tabular form of the equation, known as a design control table, is used. Typical link equations will therefore not be explicitly written out but will be structured according to the tabular organization presented in the subsequent subsection.

4.2 Typical Design Control Table Configuration

A design control table (DCT) is a tabulation of all pertinent link parameters for a given channel in an order akin to tracing through the end-to-end system from input to output. Additionally, the entries are arranged in such a way as to expedite the logical order of calculation, along with instructions for adding and subtracting the numbers.

Since RF channels may be signal-wise subdivided into their various component parts (e.g., a typical downlink may be partitioned into its carrier component, telemetry component and ranging component), the DCT is sectionally organized according to such components. Consider the simplest of links, which involves only one transmitter and one receiver. The first section would then involve the calculation of the total SNR spectral density ratio at the receiver. Table 4.1 shows the DCT total channel section. Note that this table has a provision for entering the nominal parameter value and parameter tolerance information. The use of tolerances is discussed in section 4.3. The grouping brackets under the parameter column are also associated with these tolerance calculations. Each parameter has an entry number, symbol (for keying to the link block diagram), and a logarithmic unit. The entry number is prefixed with an alphabetic code to indicate the DCT section or table type. In the case of Table 4.1, the "T" denotes "total signal." Other common prefixes would be as follows:

CT = carrier tracking
DC = data channel
DCX = data channel "X"
A = analog channel
CM = command channel
R = ranging channel

The prefixes allow for ease of cross-table parameter reference when the entire DCT is comprised of several sections, as will be discussed subsequently. Calculation procedure/notes appear in the parameter column enclosed in parentheses (or brackets), where unprefixed numbers refer to entries within the immediate table and prefixed numbers are entries found in the corresponding tables.

Table 4.1. Design Control Table, Total Signal for a Single Channel

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
T1	Total Transmit Power	P_T	dBW					
T2	Transmit Circuit Losses	L_{TX}	dB					
T3	Transmit Antenna Gain	G_T	dB					
T4	Transmit Pointing Loss	L_{PTX}	dB					
T5	Space Loss, f = _____ MHz	L_S	dB					
T6	Atmospheric Loss	L_A	dB					
T7	Receive Antenna Gain	G_R	dB					
T8	Polarization Loss	L_{POL}	dB					
T9	Receive Pointing Loss	L_{PR}	dB					
T10	Receive Circuit Losses	L_R	dB					
T11	Total Received Power (Σ1 through 10)	P_R	dBW					
T12	Receive Noise Density NF = _____ dB	N_0	dBW/Hz					
T13	Receiver Signal/Noise Density (11 - 12)	P_R/N_0	dB-Hz					

Coherent communication systems require a carrier phase-tracking function regardless of whether the carrier is spectrally discrete or suppressed. Some noncoherent FM receivers employ AFC loops which, in effect, track the mean carrier frequency. For these reasons, a DCT is needed for the carrier-tracking function; an example is shown in Table 4.2. This table is universal in that it may be used with any carrier-tracking function found in any of the payload links. The difference concerns which of the loss terms (entry numbers 3 through 7) are considered applicable. The nature and relevance of each of these losses is discussed under Section 3.0. (Note that the first line entry is obtained from line 13 of Table 4.1 or line 14 of Table 4.5 or line 12 of Table 4.6.) Examples for the various types of tracking functions may be found in Appendix B.

Table 4.2 also introduces the use of the measure "margin" which is found in all of the end-point performance DCT's. "Nominal value margin" is simply the difference between some DCT-calculated nominal performance parameter and the specified/desired value of the same performance parameter. This latter value is supplied by the user as a necessary minimum level of operation. Interpretation of margin, taking into consideration the parameter tolerances, is discussed in the following subsection.

A digital data channel DCT appears in Table 4.3. Again, this table is universal to all digital configurations, with and without subcarriers, coding and PN spreading. Entries are defined in Section 3.0, and several examples of its use are found in Appendix B.

Table 4.4 is for an FM analog channel. Computation of the demodulator output SNR (line A7) requires the use of an appropriate formula relative to the input SNR conditions and whether preemphasis/deemphasis is employed. For a discussion of such formulas, see subsection 3.2.11.2. Appendix B5 provides an example of this table. If the modulation is FSK and a postdemodulation synchronizer/detector is used, it is usually necessary to compute the equivalent noise spectral density (see 3.2.11.2.4). Following this, Table 4.3 may be employed to complete the FSK performance calculation.

Two tandem-link DCT's are now presented. Table 4.5 is a total-signal DCT for use with the TDRS, where the Orbiter/TDRS links are SNR limited. It may therefore be used for the Orbiter-to-TDRS-to-ground link where the transmit and receive parameters apply respectively to the

Table 4.2. Carrier-Tracking Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
CT1	Receiver Signal/Noise Density (T12 or T13 or T14)	P_R/N_0	dB-Hz					
CT2	Carrier/Total Power	P_C/P_R	dB					
CT3	Noisy Oscillator Loss	L_{OSC}	dB					
CT4	PN Filtering Loss	L_{PN}	dB					
CT5	PN Correlation Loss	L_{COR}	dB					
CT6	Squaring Loss	L_{SQ}	dB					
CT7	Limiter Loss	L_{LIM}	dB					
CT8	Net Carrier-Signal/Noise-Density (Σ 1 through 7)	P_{NC}/N_0	dB-Hz					
CT9	Carrier Loop Noise Bandwidth, Two-Sided	$2B_L$	dB-Hz					
CT10	Carrier Loop SNR (8 - 9)	$(\frac{S}{N})_{2B_L}$	dB-Hz					
CT11	Required Carrier Loop Minimum SNR		dB					
CT12	Carrier Loop SNR Margin (10 - 11)		dB					
CT13	Adverse Margin [12 + (12 Adv. Tol.)]		dB					
CT14	Mean Tolerance Margin [12 + (12 Mean Tol.)]		dB					
CT15	3σ Tolerance Margin [14 + 3×(12 Std. Dev.)]		dB					

Table 4.3. Digital Data Channel Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
DC1	Receiver Signal/Noise Density (T12 or T13 or T14)	P_R/N_0	dB-Hz					
DC2	Data/Total Power	P_D/P_R	dB					
DC3	PN Filtering Loss (CT4)	L_{PN}	dB					
DC4	PN Correlation Loss (CT5)	L_{COR}	dB					
DC5	Carrier Phase Noise Loss	L_ϕ	dB					
DC6	Subcarrier Demodulation Loss	L_{SD}	dB					
DC7	Bit Synchronization Loss	L_{BS}	dB					
DC8	Miscellaneous Waveform Loss	L_{WF}	dB					
DC9	Data Bit Rate	R_b	dB-bps					
DC10	Net Bit Energy/Noise Density [(Σ1 through 8) - 9]	E_b/N_0	dB					
DC11	Required E_B/N_0 P_E = _____ Coded ___ Uncoded ___		dB					
DC12	Data Channel Margin (10 - 11)		dB					
DC13	Adverse Margin [12 + (12 Adv. Tol.)]		dB					
DC14	Mean Tolerance Margin [12 + (12 Mean Tol.)]		dB					
DC15	3σ Tolerance Margin [14 + 3×(12 Std. Dev.)]		dB					

Table 4.4. FM Analog Channel Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
A1	Receiver Signal/Noise Density (T12 or T13 or T14)	P_R/N_0	dB-Hz					
A2	Predemodulation Bandwidth	B_i	dB-Hz					
A3	Signal Filtering Loss	L_{SF}	dB					
A4	Predemodulation SNR (1 - 2 + 3)	$(\frac{S}{N})_{B_i}$	dB					
A5	Peak Carrier Deviation	Δf	Hz					
A6	Postdemodulation Bandwidth	B_o	Hz					
A7	Demodulator Output SNR (Formula)	SNR_0	dB					
A8	Minimum Required SNR_0		dB					
A9	SNR_0 Margin (7 - 8)		dB					
A10	Adverse Margin [9+(9 Adv. Tol.)]		dB					
A11	Mean Tolerance Margin [9+(9 Mean Tol.)]		dB					
A12	3σ Tolerance Margin [11+3×(9 Std. Dev.)]		dB					

Table 4.5. Design Control Table, Total Signal for the TDRS Link

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
T1	Total Transmit Power	P_T	dBW					
T2	Transmit Circuit Losses	L_{TX}	dB					
T3	Transmit Antenna Gain	G_T	dB					
T4	Transmit Pointing Loss	L_{PTX}	dB					
T5	Space Loss f = _____ MHz	L_S	dB					
T6	Receive Antenna Gain	G_R	dB					
T7	Polarization Loss	L_{POL}	dB					
T8	Receive Pointing Loss	L_{PR}	dB					
T9	Receive Circuit Losses	L_R	dB					
T10	Atmospheric Loss	L_A	dB					
T11	TDRS Loss	L_{TDRS}	dB					
T12	Total Received Power (Σ1 through 11)	P_R	dBW					
T13	Receiver Noise Density NF = _____ dB	N_0	dBW/Hz					
T14	Receiver Signal/Noise Density	P_R/N_0	dB-Hz					

Orbiter and the TDRS or for the ground-to-TDRS-to-Orbiter link, where the transmit parameters are those of the TDRS and receive parameters are associated with the Orbiter In this latter link, care must be exercised when dealing with the TDRS multiple-access mode to make sure the correct amount of "per-channel" power is inserted on line 1; for the Ku-band return link, the proper transmitter power components must be accounted for (see subsection 2.3.2). TDRS/ground system performance is accounted for by a single entry, the "TDRS loss" on line 11. Note that the atmospheric loss, now on line 10, has been statistically associated with the TDRS loss. A typical DCT for the Ku-band link is found in Appendix B6.

A second type of tandem link is the "bent-pipe." The nature of the noise equivalent calculation involved for feedthrough modulation noise is discussed in subsection 3.2.1.5.2, and Table 4.6 shows the accounting procedure. Appendices B7 and B8 provide sample calculations.

In summary, this subsection has introduced six typical DCT's that cover the majority of possible payload/ground configurations. These DCT's have been made as universal as possible; however, some revision/customizing by the user to exactly suit the end-to-end link being considered is likely to be required.

4.3 Tolerances

4.3.1 The Source and Nature of Tolerances

Tolerances are assigned to each nominal parameter by cognizant individuals, and are based upon both objective and subjective considerations. Some of the major factors which give rise to tolerances are as follows:

(a) Inability to manufacture electrical and mechanical components exactly to specification

(b) Nonconstant environmental conditions (e.g., temperature, pressure)

(c) Component aging

(d) Electrical operating-point drifts, especially power supplies

(e) Inability to measure performance beyond a certain level of accuracy

(f) Imperfect theoretical performance models

(g) Past experience with identical or similar equipment

(h) Risk conditions.

Table 4.6. Bent-Pipe Effective Signal and Noise Levels

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
BP1	Total Received Power (T12)	P_R	dBW					
BP2	Ku-Band Channel/Total Power	P_{CH}/P_R	dB					
BP3	Channel S+N Power (1 + 2)	P_{S+N}	dBW					
BP4	S/(S+N) Power	P_{SR}/P_{S+N}	dB					
BP5	Effective Received Signal Power (3 + 4)	P_{SR}	dBW					
BP6	N/(S+N) Power	P_{NR}/P_{S+N}	dB					
BP7	Effective Received Noise Power (3 + 6)	P_{NR}	dBW					
BP8	Channel Noise Bandwidth	B_{BP}	dB-Hz					
BP9	Effective Received Noise Density (7 - 8)	N_{01}	dBW/Hz					
BP10	Receiver Noise Density (T13)	N_{02}	dBW/Hz					
BP11	Total Noise Density (*)	N_0	dBW/Hz					
BP12†	Channel Signal/Noise Density (5 - 11)	P_{SR}/N_0	dB-Hz					

*$N_0 = 10 \log \left[\log^{-1} \left(\frac{N_{01}}{10} \right) + \log^{-1} \left(\frac{N_{02}}{10} \right) \right]$

†NOTE: When supplying this entry to line DC1 of a Data Channel DCT, also enter values of zero on line DC2.

Ku-Band Mode _____
Ku-Band Channel _____

When tolerances are specified, they should be realistic. Under no circumstances should they be "padded" to embody "safety margins." Generally, they should be small for communication links--on the order of several tenths of a decibel. Tolerances that exceed 1 dB or more should be suspect and investigated as to their validity. Meaningful tolerances are specified only after careful consideration of all factors, including previous knowledge and experience.

Electrical circuits that must maintain linearity or some prescribed level of constant performance usually have the larger tolerances. Microwave components (especially power amplifiers) and subsystems are found to have the largest tolerances of all. Microwave connectors and joints, when used in sufficient quantities, give rise to undue cumulative tolerances. Digital circuits and subsystems, on the other hand, have relatively low tolerance margins. Digital functions, rather than being graded by tolerances, are usually rated by their probability of failure (i.e., either they work or they don't).

Some parameters, such as atmospheric loss, are in reality random variables with a mean value which is called the parameter nominal value. Other parameters, such as data bit rate, have such small tolerances that it can be argued that no tolerances need be assigned. In addition, there are parameters such as required thresholds or desired operating-point values which have no tolerances because they are specified operational conditions which are inviolate with respect to equipment performance.

4.3.2 Tolerance Accounting

The usual approach to tolerance specification is to supply the most adverse and most favorable values that a particular parameter may attain. From a statistical point of view, such values represent the extremes of the distribution of possible values for the parameter; as a result, their likelihood of occurrence is rather small. Yet, there has been a prevailing feeling that, unless the communication link design margin is sufficient to overcome the most adverse tolerance possibilities, the link cannot be relied upon. Although this approach had much merit in the earlier days of the U.S space programs, many years of experience,

together with significant refinements in equipment and techniques, make its use in the Shuttle era overly conservative. Therefore, in addition to the adverse or worst-case tolerance method, a statistical procedure* is also outlined in the following paragraphs.

Reference to any of the DCT's in section 4.2 shows that columns have been provided for listing the favorable and adverse tolerances of each parameter. The calculation information listed in the parameter column of the DCT used to calculate nominal values is also used to calculate tolerance. Specifically, with reference to Table 4.1, all the parameter nominal values and favorable and adverse tolerances are supplied on the DCT on lines 1 through 10 inclusive. The parameter nominal value and tolerances for the entry on line 11 are then calculated by summing the values of lines 1 through 10 as indicated by the (Σ1 through 10) found in the parameter column on line 11.

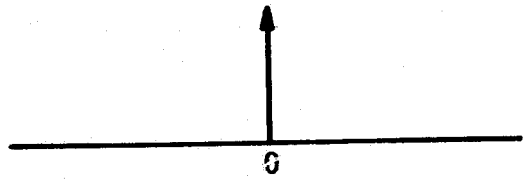
The statistical or probabilistic method of tolerance accounting is now reviewed. In the DCT's, this involves the two tolerance columns designated "mean" and "std. dev." (standard deviation). Use of the statistical method requires that a PDF be assigned to each of the parameters. In order to limit the number of PDF's employed, only four different forms are considered: (1) discrete, (2) rectangular, (3) triangular, and (4) Gaussian. Figure 4.1 shows these PDF's and the relationship of their mean and standard deviation to the extreme points. When applying them to the DCT, the extreme points of the PDF are set exactly equal to the specified adverse and favorable tolerances. Thus, in Figure 4.1, "a" is the adverse number and "f" is the favorable number, while μ and σ are, respectively, the mean and standard deviation. (Note that, since the adverse tolerance is expressed in decibels, "f" is generally a positive number and "a" is a negative number.)

The assignment of a PDF form to any parameter is dependent upon experience and measurements. The example given in Table 4.7 has its PDF selections based upon practice and may therefore be used as a guideline. Once the PDF's have been assigned along with the favorable and adverse tolerances, the parameter means and standard deviations are easily

---

*This method is based upon the approach developed in [3].

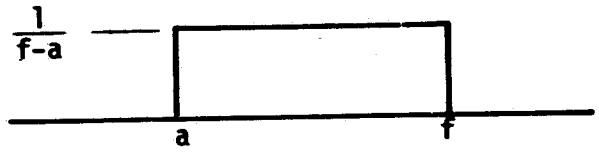
(1) Discrete PDF



$\mu = 0$

$\sigma = 0$

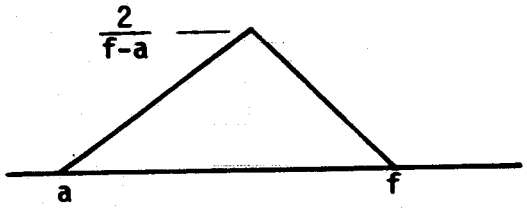
(2) Rectangular PDF



$$\mu = \frac{1}{2}(f + a)$$

$$\sigma = \left[\frac{1}{2}(f^2 + af + a^2) - \frac{1}{4}(f + a)^2\right]^{1/2}$$

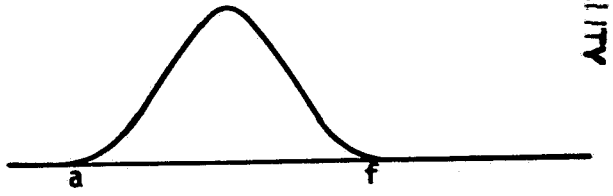
(3) Triangular PDF



$$\mu = \frac{1}{3}(f + a)$$

$$\sigma = \left[\frac{1}{6}(f^2 + af + a^2) - \frac{1}{9}(f + a)^2\right]^{1/2}$$

(4) Gaussian PDF



$$\mu = f + a$$

$$\sigma = \frac{1}{6}(f - a)$$

Figure 4.1. Parameter Tolerance Probability Density Functions

calculated from the formulas given in Figure 4.1. In the DCT's previously introduced, some parameters are grouped because of a statistical interrelationship between them. Where this occurs, the favorable tolerances are summed together, as are the adverse tolerances, to form a combined parameter tolerance pair used to compute the mean and standard deviation of the grouping.

Once all of the means and standard deviations have been determined for the specified parameters, the calculated parameters have their means and standard deviations calculated according to the relationships

$$\mu_{Total} = \sum_{n=1}^{N} \mu_n \qquad (4\text{-}3)$$

$$\sigma_{Total} = \left[\sum_{n=1}^{N} \sigma_n^2\right]^{1/2} \qquad (4\text{-}4)$$

in accordance with the calculation instructions found in the parameter column of the DCT, which determine N.

An example at this point will best serve to dispel any confusion. Table 4.7a shows a total signal DCT for a hypothetical link; thus, the numbers, although realistic, do not represent any particular channel. Note for the sake of this example that an additional column has been added to indicate the PDF form used (D = discrete, R = rectangular, T = triangular, G = Gaussian). On line 11, the "mean" and "std. dev." values are calculated using (4-3) and (4-4), respectively, with N = 7. The nominal value on line 11 is taken to have a Gaussian PDF by virtue of applying the central-limit theorem when combining all of the specified parameters.

Table 4.7b continues the calculation begun in Table 4.7a through the carrier-tracking DCT. The results of Table 4.7b are discussed in the next subsection.

Table 4.7a. Design Control Table, Total Signal for a Single Channel (Hypothetical Example)

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance				PDF
					Favorable	Adverse	Mean	Std. Dev.	
T1	Total Transmit Power	P_T	dBW	-8.5	1.2	-1.5	-0.10	0.55	T
T2	Transmit Circuit Losses	L_{TX}	dB	-0.4	0.2	-0.3	-0.05	0.14	R
T3	Transmit Antenna Gain	G_T	dB	12.5	0.6	-0.6			
T4	Transmit Pointing Loss	L_{PTX}	dB	-0.3	0.2	-0.2	-0.03	0.35	T
T5	Space Loss f = 15,000 MHz	L_S	dB	-206.0	0.0	0.0	0.00	0.00	D
T6	Atmospheric Loss	L_A	dB	-0.8	0.8	-0.8	0.00	0.27	G
T7	Receive Antenna Gain	G_R	dB	36.2	0.7	-0.9			
T8	Polarization Loss	L_{POL}	dB	-0.3	0.1	-0.2			
T9	Receive Pointing Loss	L_{PR}	dB	-0.1	0.1	-0.1	-0.10	0.43	T
T10	Receive Circuit Losses	L_R	dB	-0.2	0.1	-0.1	0.00	0.04	R
T11	Total Received Power (Σ1 through 10)	P_R	dBW	-167.9	4.0	-4.7	-0.28	0.84	G
T12	Receiver Noise Density NF = 4.0 dB	N_0	dBW/Hz	-200.0	-0.3	0.3	0.00	0.10	G
T13	Receiver Signal/Noise Density (11 - 12)	P_R/N_0	dB-Hz	32.1	4.3	-5.0	-0.28	0.85	G

Table 4.7b. Carrier-Tracking Design Control Table (Hypothetical Example)

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance				PDF
					Favorable	Adverse	Mean	Std. Dev.	
CT1	Receiver Signal/Noise Density (T13)	P_R/N_0	dB-Hz	32.1	4.3	-5.0	-0.28	0.85	G
CT2	Carrier/Total Power	P_C/P_R	dB	0.0	0.0	0.0	0.00	0.00	D
CT3	Noisy Oscillator Loss	L_{OSC}	dB	-0.3	0.0	0.0	0.00	0.00	D
CT4	PN Filtering Loss	L_{PN}	dB	-0.5	0.1	-0.1			
CT5	PN Correlation Loss	L_{COR}	dB	-0.8	0.3	-0.1	0.07	0.12	T
CT6	Squaring Loss	L_{SQ}	dB	-0.6	0.1	-0.2	-0.05	0.11	R
CT7	Limiter Loss	L_{LIM}	dB	0.0	0.0	0.0	0.00	0.00	D
CT8	Net Carrier Signal/Noise Density (Σ1 through 7)	P_{NC}/N_0	dB-Hz	29.9	4.8	-5.4	-0.26	0.87	G
CT9	Carrier Loop Noise Bandwidth, Two-Sided	$2B_L$	dB-Hz	17.8	0.0	0.0	0.00	0.00	D
CT10	Carrier Loop SNR (8 - 9)	$(\frac{S}{N})_{2B_L}$	dB	12.1	4.8	-5.4	-0.26	0.87	G
CT11	Required Carrier Loop Minimum SNR		dB	9.0					
CT12	Carrier Loop SNR Margin (10 - 11)		dB	3.1	4.8	-5.4	-0.26	0.87	G
CT13	Adverse Margin [12 + (12 Adv. Tol.)]		dB	-2.3					
CT14	Mean Tolerance Margin [12 + (12 Mean Tol.)]		dB	2.84					
CT15	3σ Tolerance Margin [14 + 3×(12 Std. Dev.)]		dB	0.23					

4.4 Interpreting Link Performance

4.4.1 Specified Versus Calculated Performance-Link Margin

As previously stated in section 4.2, "margin" is an important end-point performance measure. Any nominal-value margin is defined as the difference between a given DCT-calculated nominal performance parameter and its specified or desired minimum value. When, therefore, the margin is positive, the designer knows that nominally an adequate link design exists. On the other hand, when it is negative, the link design is inadequate and some redesign or tighter specifications are required.

Table 4.7b continues the hypothetical example introduced in the previous subsection through the carrier-tracking function. Line 12 of Table 4.7b therefore shows a nominal carrier loop SNR margin of +3.1 dB. On the surface, this might appear quite adequate; however, when tolerances are considered, it will be seen to be only barely adequate, as discussed in the next subsection.

4.4.2 Tolerance Conditions

Line 12 of Table 4.7b illustrates a carrier loop SNR margin of +3.1 dB, along with the margin tolerances. If the favorable tolerance is considered, the margin could be as large as 7.9 dB. However, this is very likely and, even if it did occur, it is of little merit as the improved link performance is, by definition, not required.

On the other hand, the adverse margin considering the adverse tolerance is -2.3 dB. This has been entered on line 13 of Table 4.7b. Since the adverse margin is negative, the link, taking the most conservative point of view, could be judged inadequate. However, consider the probabilistic margins based upon the PDF parameters. The mean-tolerance margin (line 14) is almost as large as the nominal margin. Taking into account the 3σ adverse tolerance, the 3σ tolerance margin (line 15) is still positive, although quite small. The 3σ tolerance margin for this example is interpreted to mean that the carrier loop SNR margin will not fall below +0.23 dB with probability 0.99. From a risk point of view, this would appear to be satisfactory performance and should be accepted in lieu of the -2.3 dB adverse margin which would call for link redesign.

In summary, the exclusive use of adverse margin results in an overly conservative approach to acceptable link performance. The 3σ tolerance margin should generally be applied when dealing with channels for which there is considerable experience and corroborating data.

5.0 DESIGN CONTROL TABLE PREPARATION

5.1 Organizing The Effort

The DCT procedure should begin by identifying the necessary payload functions and requirements. Standard telemetry and command capability should be reviewed next to determine whether it will satisfy the needs. If it will, then the remaining DCT procedure is straightforward, and it is likely that the link configurations and DCT's presented in this handbook will suffice with little, if any, modification.

Given that the standard capabilities are inappropriate, the next two steps should be: (1) establish the most appropriate end-to-end link configuration that will meet the requirement (refer to subsection 2.1.2) and (2) construct a functional block diagram showing those flight and ground subsystems that will be involved (see subsection 2.1.3 for guidance). Major system/subsystem performance parameters should be identified on the functional diagram. Next, DCT forms should be prepared for all sections of the link that have separate end-point measures. The DCT forms presented in section 4.2 should be used as the basis for this aspect of the effort.

5.2 User Variable Parameters

Once the DCT forms have been prepared, parameter nominal values and tolerances must be specified. Actually, the payload user has only a few independent options insofar as the parameters are concerned. All of the Orbiter and much of the ground parameters are inviolate. They comprise the bulk of the DCT entries and may be readily obtained by reference to Appendices A and B of this handbook.

Parameters over which the user does have control are obviously those associated with the payload itself and any special-purpose ground equipment. Certain payload parameters, however, are constrained (see Appendix D for such restrictions).

When selecting parameters over which the user has a choice, sound engineering judgment and potential cost consciousness must be exercised. The state-of-the-art relative to hardware capability must be appreciated such that desired parametric values may be easily realized. Further, "inventive" approaches should be avoided, as they will probably increase

costs considerably in development. All in all, the degree of flexibility available to the user is limited, and proven techniques should therefore be implemented.

5.3 Maximizing Desired Performance

When specifying parameters over which the user has control, trade-off studies should be conducted to maximize the needed transmitter power/antenna gain product relative to cost. Improvements on the order of a few decibels may be most economically obtained through the use of coding or highly efficient modulation methods. Clearly, as outlined in subsection 4.4.2, small but positive 3σ tolerance margins should be accepted as indicative of valid link performance, whenever possible.

5.4 How To Obtain Additional Information

It may be necessary for those users who do not have extensive experience with DCT procedures or the Shuttle communication system to obtain guidance beyond that which is provided in this handbook. Technical questions and inquiries relative to parameter values and tolerances should be addressed to:

Mr. William Teasdale, EE8
Systems Analysis Branch
NASA Johnson Space Center
Houston, Texas 77058
(713)483-4647.

REFERENCES

1. "Special Issue on Space Shuttle Communications and Tracking," IEEE Transactions on Communications, Vol. COM-26, No. 11, November 1978.

2. "Telecommunications Systems Design Techniques Handbook," Technical Memorandum No. 33-571, Jet Propulsion Laboratory, July 14, 1972.

3. "A Statistical Model for Telecommunication Link Design," Paper No. 25-17, National Telecommunications Conference, New Orleans, December 1975.

4. "Tracking and Data Relay Satellite System (TDRSS) Users' Guide," STDN No. 101.2, Rev. 3, Goddard Space Flight Center, January 1978.

5. "Telemetry Standards," IRIG Document No. 106-66, March 1966.

6. W. W. Peterson and E. J. Weldon, Jr., Error-Correcting Codes, Second Edition, The MIT Press, Cambridge, Mass., 1972.

7. S. Butman and U. Timon, "Interplex--An Efficient PSK/PM Telemetry System," IEEE Transactions on Communication, Vol. COM-20, No. 3, June 1972, pp. 415-419.

8. W. V. T. Rusch and P. D. Potter, Analysis of Reflector Antennas, Academic Press, New York, 1970.

9. "CCIR 11th Plenary Assembly," Vol. II, Report No. 234, Oslo, 1966.

10. A. Benoit, "Signal Attenuation Due to Neutral Oxygen and Water Vapour, Rain and Clouds," Microwave Journal, November 1968, pp. 73-80.

11. R. K. Crane, "Propagation Phenomena Affecting Satellite Communication Systems," Proceedings of the IEEE, February 1971, p. 178.

12. H. Staras, "The Propagation of Wideband Signals Through the Ionosphere," IRE Proceedings, Vol. 49, July 1961, p. 1211.

13. "System 621B--Satellite System for Precise Navigation," Hughes Aircraft Company Final Report (Part 2), SAMSO TR69-4, January 1969, Appendix III.

14. "World Distribution and Characteristics of Atmospheric Radio Waves," CCIR Report No. 322, 10th Plenary Assembly, Geneva, 1963.

15. D. L. Rice, et al., "Transmission Loss Prediction for Tropospheric Loss Communication Circuits," Tech. Note 101, National Bureau of Standards, U.S. Government Printing Office, January 1967.

16. W. B. Davenport, "Signal-to-Noise Ratios in Bandpass Limiters," Journal of Applied Physics, Vol. 24, No. 6, June 1953.

17. R. C. Tausworthe, "Theory and Practical Design of Phase-Locked Receivers," Vol. 1, Jet Propulsion Laboratory Technical Report No. 32-819, February 15, 1966.

18. W. C. Lindsey, "Coding and Synchronization Studies: Moments of the First Passage Time in Generalized Tracking Systems," Jet Propulsion Laboratory, Space Programs Summary No. 37-58, Vol. III, pp. 63-66.

19. J. P. Frazier and J. Page, "Phase-Lock-Loop Frequency Acquisition Study," Transactions of the IRE, Vol. SET-8, September 1962.

20. F. M. Gardner, Phaselock Techniques, John Wiley & Sons, New York, 1966.

21. R. M. Jaffe and E. Rechtin, "Design and Performance of Phase-Lock-Loop Circuits Capable of Near-Optimum Performance Over a Wide Range of Input Signals and Noise Levels," IRE Trans. on Info. Theory, Vol. IT-1, March 1955, pp. 66-70.

22. A. J. Viterbi, Principles of Coherent Communication, McGraw-Hill, New York, 1966.

23. R. C. Tausworthe, "Information Processing: Limiters in Phase-Locked Loops, A Correction to Previous Theory," Jet Propulsion Laboratory, Space Program Summary No. 37-54, Vol. III, December 31, 1968, pp. 201-204.

24. W. C. Lindsey, "A Theory for the Design of One-Way and Two-Way Phase-Coherent Communication Systems, Phase-Coherent Tracking Systems," Jet Propulsion Laboratory Technical Report No. 32-986, July 15, 1969.

25. M. K. Simon and W. C. Lindsey, "Optimum Performance of Suppressed-Carrier Receivers with Costas Loop Tracking," IEEE Trans. on Comm., Vol. COM-25, No. 2, February 1977, pp. 215-227.

26. W. C. Lindsey and M. K. Simon, "Optimum Design and Performance of Costas Receivers Containing Soft Bandpass Limiters," IEEE Trans. on Communication, Vol. COM-25, No. 8, August 1977, pp. 822-831.

27. M. K. Simon and J. C. Springett, "The Theory, Design and Operation of the Suppressed-Carrier Data-Aided Tracking Receiver," Jet Propulsion Laboratory Technical Report No. 32-1583, June 15, 1973.

28. J. R. Lesh, "Signal-to-Noise Ratios in Coherent Soft Limiters," IEEE Trans. on Communication, Vol. COM-22, No. 6, June 1974, pp. 803-811.

29. W. K. Alem, G. K. Huth, and M. K. Simon, "Integrated Source and Channel-Encoded Digital Communication System Design Study," Final Report on NASA Contract No. 9-15240A, Axiomatix Report No. R7803-7, March 31, 1978.

30. M. K. Simon and W. K. Alem, "Tracking Performance of Unbalanced QPSK Demodulators: Part I - Biphase Costas Loop with Passive Arm Filters," IEEE Trans. on Comm., Vol. COM-26, No. 8, August 1978, pp. 1147-1156.

31. M. K. Simon, "Tracking Performance of Unbalanced QPSK Demodulators, Part II - Biphase Costas Loop with Active Arm Filters," IEEE Trans. on Communication, Vol. COM-26, No. 8, August 1978, pp. 1157-1166.

32. M. K. Simon, "Tracking Performance of Costas Loops with Hard-Limited In-Phase Channel," IEEE Trans. on Comm., Vol. COM-26, No. 4, April 1978, pp. 420-432.

33. M. K. Simon, "Noncoherent Pseudonoise Code-Tracking Performance of Spread Spectrum Receivers," IEEE Trans. on Comm., Vol. COM-25, No. 3, March 1977, pp. 327-345.

34. J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering, John Wiley & Sons, Inc., New York, 1965.

35. G. A. DeCouvreur, "Effects of Random Synchronization Errors in PN and PSK Systems," IEEE Trans. on AES, January 1970, pp. 98-100.

36. J. J. Jones, "Filter Distortion and Intersymbol Interference Effects on PSK Symbols," IEEE Trans. on Comm. Tech., Vol. COM-19, No. 2, April 1971, pp. 120-132.

37. W. R. Bennett and J. Salz, "Binary Data Transmission by FM Over a Real Channel," Bell Sys. Tech. Jour., Vol. 42, September 1963, pp. 2387-2426.

38. T. T. Tjhung and P. H. Wittke, "Carrier Transmission of Binary Data in a Restricted Band," IEEE Trans. on Comm. Tech., Vol. COM-18, August 1970, pp. 295-304.

39. B. H. Batson, J. W. Seyl and B. G. Smith," Experimental Results for FSK Data Transmission Systems Using Discriminator Detection," presented at 1976 National Telecommunications Conference, Dallas, Texas, November 29, 1976, pp. 5.5.1 to 5.5-9.

40. G. K. Huth, et al., "Integrated Source and Channel-Encoded Digital Communication System Design Study - Final Report," Axiomatix Report No. R7607-3 (under NASA Contract 9-13467), July 31, 1976.

41. S. Stein, "Unified Analysis of Certain Coherent and Noncoherent Binary Communication Systems," IEEE Trans. on Info. Theory, Vol. IT-10, January 1964, pp. 43-51.

42. W. C. Lindsey and M. K. Simon, Telecommunication Systems Engineering, Prentice Hall, Inc., Englewood Cliffs, N.J., 1973.

43. A. J. Viterbi, "Convolutional Codes and Their Performance in Communication Systems," IEEE Trans. on Comm. Tech., Vol. COM-19. No. 5, October 1971, pp. 751-772.

44. J. A. Heller and I. M. Jacobs, "Viterbi Decoding for Satellite and Space Communication," IEEE Trans. on Comm. Tech., Vol. COM-19, No. 5, October 1971, pp. 849-855.

45. J. W. Layland and W. A. Lushbaugh, "A Flexible High-Speed Sequential Decoder for Deep-Space Channels," IEEE Trans. on Comm. Tech., Vol. COM-19, No. 5, October 1971, pp. 813-820.

46. D. Chase, "A Class of Algorithms for Decoding Block Codes with Channel Measurement Information," IEEE Trans. on Info. Theory, Vol. IT-18, No. 1, January 1972, pp. 170-182.

47. J. L. Massey, Threshold Decoding, MIT Press, Cambridge, Mass., 1963.

48. J. C. Springett and M. K. Simon, "An Analysis of Phase Coherent-Incoherent Output of the Bandpass Limiter," IEEE Trans. on Comm. Tech., Vol. COM-19, No. 1, February 1971, pp. 42-49.

APPENDIX A

PERFORMANCE PARAMETERS

A1 PAYLOAD INTERROGATOR MAJOR PERFORMANCE PARAMETERS

The function of the Payload Interrogator (PI) is to provide the RF communication link between the Orbiter and detached payloads. For communication with the NASA payloads, the PI operates in conjunction with the Payload Signal Processor (PSP). During DOD missions, the PI is interfaced with the Communication Interface Unit (CIU). Nonstandard (bent-pipe) data received by the PI from either NASA or DOD payloads is delivered to the Ku-band Signal Processor (KuSP), where it is encoded for transmission to the ground via the Shuttle/TDRS link.

The tables on the following pages summarize the major performance parameters of the Payload Interrogator equipment.

Table A1.1. Payload Interrogator Operating Channels

	Frequency Range (MHz)	Number of Channels	Channel Spacing (kHz)
STDN			
Transmit	2025.833 - 2118.7	888	115
Receive	2200 - 2300.875	808	125
DSN			
Transmit	2110.243 - 2119.792	29	341.049
Receive	2290.185 - 2299.814	27	370.37
SGLS			
Transmit	1763.721 - 1839.795	20	4004
	2202.5 - 2297.5	20	5000

Table A1.2. Payload Interrogator Antenna Characteristics

Parameter	Value
Operating frequency	1740 to 1850 MHz (VSWR < 1.5:1) 2000 to 2300 MHz (VSWR ≤ 2.0:1)
Gain	2.5 dB
Beamwidth (3 dB)	±50° (100° cone)
Polarization	Right-hand circular (RHCP)
RF cable losses	9.8 dB (max)

Table A1.3. Receiver and Transmitter Performance Characteristics

RECEIVER			
Noise figure	7 dB max		
Doppler accommodated	±75 kHz		
Doppler rate accommodated	44 kHz/s		
Static phase-tracking error	±3°		
Mean phase-tracking error	≤15°		
Acquisition threshold (carrier component)	-122.2 dBm		
Acquisition time (probability 0.9)	≤3 s		
Tracking threshold (carrier component)	-129.2 dBm		
Mean time to lose lock	>10 s		
Dynamic range	110 dB		
	Low	Medium	High
Receiver sensitivity (maximum no-damage input)	36 dBm	30 dBm	20 dBm
Minimum signal level (modulated)	-87 dBm	-107 dBm	-120 dBm
TRANSMITTER			
RF power output (2:1 load) (minimum)	4 dBm	27 dBm	37 dBm
Transmit frequency stability	≤ ±0.001%		
Sweep capability			
Wide (STDN or SGLS)	±50 kHz @ 35 kHz/s		
Narrow (DSN)	±30 kHz @ 0.6 kHz/s		
Phase noise (steady state)	4° RMS		

Table A1.4. Baseband Receive and Transmit Characteristics

RECEIVE PM CHARACTERISTICS (TELEMETRY DATA)	
NASA (STDN/DSN)	
Data rate	1, 2, 4, 8 or 16 kbps
Subcarrier waveform	1.024 MHz PSK
Modulation index	1.0 rad
DOD (SGLS)	
Data rate	0.25 to 256 kbps in steps of 2
Subcarrier waveform	1.024 MHz PSK and/or 1.7 MHz FM/FM
Modulation index	0.3 or 1.0 rad each subcarrier
Bent-Pipe	
Any of the above NASA or DOD forms or baseband analog signals up to 4.5 MHz or 3 Mbps digital data (biphase-L or NRZ-L) phase modulated onto the carrier	
TRANSMIT PM CHARACTERISTICS (COMMAND DATA)	
NASA (STDN/DSN)	
Data rate	125/16, 125/8, 125/4, 125/2, 125, 250, 500, 1000, 2000 bps
Subcarrier waveform	16 kHz PSK
Modulation index	1.0 rad
DOD (SGLS)	
Data rate	1 ksps or 2 ksps
Subcarrier waveform	Ternary FSK/AM (65, 76 or 95 kHz tones)
Modulation index	0.3 or 1.0 rad

A2 NASA PAYLOAD SIGNAL PROCESSOR MAJOR PERFORMANCE PARAMETERS

The Payload Signal Processor (PSP) performs the following functions: (1) it modulates NASA payload commands onto a 16-kHz sinusoidal subcarrier and delivers the resultant signal to the PI and the attached payload umbilical, (2) it demodulates the payload telemetry data from the 1.024 MHz subcarrier signal provided by the PI, and (3) it performs bit and frame synchronization of demodulated telemetry data and delivers this data and its clock to the Payload Data Interleaver (PDI).

Presented in the following tables are the salient parameters of the NASA Payload Signal Processor equipment.

Table A2.1. PSP Command Processing Section Parameters

PSP/MDM COMMAND INTERFACE	
Mode	Half duplex
Data rate	1 Mbps
PCM format	Manchester
COMMAND DATA OUTPUT	
Bit rate	125/16, 125/8, 125/4, 125/2, 125, 250, 500, 1000 or 2000 bps
PCM format	NRZ-L, NRZ-M or NRZ-S
Modulation	PSK on 16-kHz sine-wave subcarrier
Data asymmetry	$\leq 2\%$
Phase jitter	$\leq 3\%$

Table A2.2. PSP Telemetry Processing Section Parameters

PCM format	NRZ or Manchester L, M, or S
Data word length	8 bits
Data frame length	1, 2, ..., 1023 or 1024 words
Frame sync word	
Pattern	Programmable
Length	8, 16, 24 or 32 bits
Subcarrier Demodulator/Bit Synchronizer	
Subcarrier frequency	1.024 MHz
Subcarrier stability	0.01%
Bit rate	1, 2, 4, 8 or 16 kbps
Bit rate stability	0.1%
Transition density	$\leq$ 64 transitions in 512 bits $\leq$ 64 consecutive bits without transition
Acquisition threshold E_b/N_0	4 dB
Tracking threshold E_b/N_0	2 dB
Hard decision data detection (2 dB $\leq E_b/N_0 \leq$ 11 dB)	<1.5 dB from theory
Acquisition time (0.9 probability)	2 seconds maximum
Mean time to lose lock (subcarrier demodulation)	$\geq$ 10 seconds minimum

A3 DOD COMMUNICATION INTERFACE UNIT SALIENT PARAMETERS

The Communication Interface Unit (CIU) is the DOD equivalent of the NASA PSP. The major differences are that the CIU (1) handles ternary commands in both baseband and FSK tone formats and (2) accepts Orbiter crew-generated commands. The salient parameters of the CIU are summarized in Table A3.1 on the following page.

Table A3.1. CIU Parameters Summary

Item	Parameter and Range
Outputs	
FSK/AM Commands	
Destination	Payload Interrogator or attached payload
Modulation	FSK/AM tones
Data rate	1 or 2 ksps
Ternary Commands	
Destination	Attached payload
Signal form	Two-level four-line signaling
Telemetry	
Destination	Payload Data Interleaver or Ku-Band Signal Processor
Signal form	NRZ
Data rate	Up to 64 kbps to PDI; up to 256 kbps to Ku-Band Signal Processor
Inputs	
Detached Payload Telemetry	
Source	Payload Interrogator
Subcarrier	1.024 MHz and/or 1.7 MHz
Modulation	PSK (1.024 and 1.7 MHz subcarriers) FM/FM (1.7 MHz subcarrier)
Data rate	0.25 to 256 kbps in steps of two
Attached Payload Telemetry	
Source	Payload umbilical
Subcarrier	1.024 MHz and/or 1.7 MHz
Modulation	PSK (1.024 MHz subcarrier) FM/FM (1.7 MHz subcarrier)
Onboard Commands	
Source	Crew inputs
Ground-Generated Commands	
Source 1 Input data rate	Ku-Band Signal Processor 128 kbps
Source 2 Input data rate	Network Signal Processor 1 Mbps bursts

A4 S-BAND PHASE-MODULATED (PM) DIRECT LINK

The S-band PM direct link consists of the forward (ground-to-Orbiter) and return (Orbiter-to-ground) segments. The onboard equipment servicing this link consists of the network transponder and the Network Signal Processor (NSP). The tables presented below summarize the salient parameters of the direct S-band PM link.

Table A4.1. Data Rates for Network PM Links in Direct Mode

Link	Rate	Channels Available	Composite Link Rate
Forward (Ground to Orbiter)	High	Two 32 kbps Voice One 8 kbps Command	72 kbps
	Low	One 24 kbps Voice One 8 kbps Command	32 kbps
Return (Orbiter to ground)	High	Two 32 kbps Voice One 128 kbps Telemetry	192 kbps
	Low	One 32 kbps Voice One 64 kbps Telemetry	96 kbps

Table A4.2. S-Band Direct Forward Link Network Subsystem Bit Error Rate Performance

Network Transponder/Network Signal Processor Bit Error Rate							
				Required Input Levels			
				32 kbps		72 kbps	
Mode	Receiver Input Port	Modulation	Maximum BER	P_T/N_0 (dB-Hz)*	Power (dBm)	P_T/N_0 (dB-Hz)*	Power (dBm)
SGLS/STDN-Lo	COMM	PSK	10^{-4}	58.8	-107.2	62.3	-103.7
SGLS/STDN-Lo	COMM	PSK	10^{-6}	60.9	-105.1	64.4	-101.6
SGLS/STDN-Lo	COMM	PM Composite	10^{-4}	62.1	-103.9	65.6	-100.4
SGLS/STDN-Lo	COMM	PM Composite	10^{-6}	64.3	-101.8	67.7	-98.3
STDN-Hi	TDRS	PSK	10^{-4}	58.8	-92.2	62.3	-88.7
STDN-Hi	TDRS	PSK	10^{-6}	60.9	-90.1	64.4	-86.6
STDN-Hi	TDRS	PM Composite	10^{-4}	62.1	-88.9	65.6	-85.4
STDN-Hi	TDRS	PM Composite	10^{-6}	64.2	-86.8	67.7	-83.3

*COMM Port: Noise density corresponding to 8 dB noise figure.
TDRS Input Port: Noise density of -151 dBm/Hz added to input signal (see note in Appendix A5).

Table A4.3. Network Transponder Characteristics Pertinent to S-Band Direct-Link Operation

Receiver	
Noise figure	8 dB max
Doppler accommodated	±60 kHz
Doppler rate accommodated	5 kHz/s max
Tone-ranging response (±1 dB)	1.5 to 1.9 MHz
Tone-ranging delay	1.0 μs max
Tone-ranging delay variation	39.0 ns max
Dynamic range	-30 to -122 dBm min
Receiver	
RF power SGLS/STDN 2:1 load	2 to 6 W
Output TDRS matched load	1 W ± .8 dB
Phase noise (during vibration)	7° RMS max
Probability of phase lock in 6 seconds > 0.9:	

Mode	Modulation
STDN/SGLS	Biphase-L, PSK (no ranging)

Minimum Signal Level into Transponder

32 kbps	72 kbps
-121.0 dBm	-120.0 dBm

Table A4.4. Network Signal Processor Salient Performance Characteristics

Bit Synchronizer	
Acquisition and tracking threshold E/N_0	-5 dB
Hard decision uncoded data detection (-5 dB $\leq E/N_0 \leq 10.6$ dB)	<0.6 dB from theory
Acquisition time	
Uncoded, $E/N_0 > 0$ dB	1 second maximum
Coded, $E/N_0 > -3$ dB	2 seconds maximum
Viterbi Decoder	
For BER 10^{-1} to 10^{-6}	<1 dB from theory
Bose-Chaudhuri-Hocquenghem (BCH) Command Decoder	
Probability of undetected error	10^{-18} maximum

A5 S-BAND PM TDRSS RELAY LINK

The function of the S band TDRSS relay link is the same as that of the direct S-band STDN link with the exception of the requirement to operate via the TDRS. The latter requirement necessitates the addition of the following functions and equipment to the direct S-band link onboard communication equipment:

(1) The forward link signal is spread in frequency by means of a PN code

(2) A low-noise preamplifier is used with the network transponder for the reception of the TDRS signal*

(3) A traveling wave tube (TWT) amplifier is added to boost the output power of the network transponder

(4) Both the forward and return signals are rate 1/3 encoded to provide an additional margin for both links.

The tables which follow describe the performance of the equipment required for operating the S-band link via the TDRS.

---

*Since receiver noise figure is usually referenced to the transponder receive port which follows the preamplifier, the received signals and noise levels at the receive port in the TDRSS configuration are nominally 20 dB greater than in the non-TDRSS configuration.

Table A5.1. Data Rates for Network PM Links in the Relay Mode

Link	Rate	Channels Available	Composite Link Symbol Rate (Rate 1/3 Coded)
Forward (ground-to-Orbiter)	High	Two 32 kbps Voice One 8 kbps Command	216 ksps (72 × 3)
	Low	One 24 kbps Voice One 8 kbps Command	96 ksps (32 × 3)
Return (Orbiter-to-ground)	High	Two 32 kbps Voice One 128 kbps Telemetry	576 ksps (192 × 3)
	Low	One 32 kbps Voice One 64 kbps Telemetry	288 ksps (96 × 3)

Table A5.2. Network Transponder Forward-Link Performance in the TDRS Relay Mode

Probability of Phase Lock in 6 sec > 0.9			
Mode	Modulation	Minimum Signal Level into Transponder	
		32 kbps	72 kbps
STDN High Power	Biphase-L, PSK* (no tone ranging)	-105.3 dBm	-104.3 dBm
TDRS*	Biphase-L, PSK	-103.5 dBm	-99.8 dBm
	Carrier Only	-105.1 dBm	-101.3 dBm

P_T/N_0 (dB-Hz)	Average Acquisition Time (sec)		Average Time to Unlock (sec)
	96 ksps	216 ksps	
48	≤ 48**	NA	≥ 3600
51	≤ 15	≤ 22.5	≥ 3600
54	≤ 7.5	≤ 7.5	≥ 3600

*Input noise power spectral density of -151 dBm/Hz, following preamp.

**Code doppler 100 chips/sec.

Table A5.3. S-Band TDRS Relay Forward-Link Bit Error Rate Performance

Network Transponder/Network Signal Processor Bit Error Rate							
				Required Input Levels			
				32 kbps		72 kbps	
Mode	Receiver Input Port	Modulation	BER	P_D/N_0 (dB-Hz)	Power (dBm)	P_D/N_0 (dB-Hz)	Power (dBm)
TDRS	TDRS	PSK	10^{-2}	49.8	-101.2	53.3	-97.5
TDRS	TDRS	PSK	10^{-4}	51.6	-99.4	55.1	-95.7
TDRS	TDRS	PSK	10^{-6}	52.7	-98.3	56.2	-94.6

NOTES: TDRS Port: Noise density of -151 dBm/Hz added to input signal, measured at the receive port following the preamplifier.

BER in TDRS mode is for a convolutionally-coded link at the output of the Viterbi decoder. Zero doppler rate assumed for all modes.

Table A5.4 TDRS/High-Power STDN Mode Orbiter
RF Equipment Characteristics

Power Amplifier (Each Channel)

RF output power (into 2:1 load mismatch) (input between 26.0 and 31.5 dBm)	100 W min 175 W max
Gain ripple (0.5 dB p-p)	10 MHz
Phase fluctuation - added	20 mrad max
AM/PM conversion	14°/dB max

Output maximums:

Harmonic	2	3	4	5	6	7	8	9	10+
Level (dBm)	+42	+32	+27	+22	+17	+12	+7	+2	-3

Preamplifier (Each Channel)

		Receive	Transmit
Passband	(MHz)	24 (3 dB)	5 (1 dB)
Gain ripple (1 dB)	(MHz)	16 (1 dB)	5 (1 dB)
Self-interference rejection	(dB)	103 at f_{tr}	115 at f_{rec}
Noise figure	(dB)	2.6	
Gain	(dB)	20-26	
Input dynamic range	(dBm)	-130 to -55	

Transmit channel attenuation	1.3 dB max
Maximum input power	200.0 W max

A6 S-BAND FM LINK

The S-band FM downlink function provides one-way communication between the Orbiter and either the NASA-STDN or the DOD-SCF ground stations. The avionics equipment used on board the Orbiter to provide the FM downlink consists of the following three major subunits: (1) FM signal processor, (2) FM transmitter, and (3) antenna. The salient parameters of these subunits are summarized in the following tables.

Table A6.1. FM Signal Processor Characteristics

TV channel input	EIA TV Standard RS 170
TV channel gain	19 dB (+0.8 to -0.25 dB)
TV channel dynamic range	51 dB ±0.25 dB
Frequency response ±0.25 dB and phase ripple ±1.0°	DC to 4.5 MHz
CCIR K factor	<2%
Main Engine	
Data in three channels	60 kbps Manchester
Subcarrier frequencies	576 kHz, 768 kHz, 1024 kHz
Subcarrier modulation	±180° at ±15°
Analog data bandwidth	300 Hz to 4 MHz
Wideband digital data rate	200 bps to 5 Mbps NRZ or 200 bps to 2 Mbps Manchester
Recorder data, two-channel data rate	25.5 to 1024 kbps
Narrowband DOD digital data rate	250 bps to 256 kbps
Input common mode voltage (DC to 2 MHz)	TV max

Table A6 2. FM Transmitter Characteristics

Frequency	2250.0 MHz ±0.003%
Output power	10 W min, 15 W max
Deviation sensitivity for deviation up to ±4.5 MHz pk)	1 MHz/V pk ±10%
Frequency response ±1 dB	DC to 5.0 MHz
Incidental AM	5% max over input range
Incidental PM	<5 kHz RMS over modulation BW
Intermodulation distortion (2-tone equal amplitude)	≥40 dB with frequency deviation ±1 MHz

Table A6.3. FM Downlink Antenna Parameters

Parameter	Value
Operating frequency	2250 MHz (VSWR≤1.5:1)
Gain	1 dB*
Beamwidth (3 dB)	±72.5° (145° cone)*
Polarization	Right-hand circular
RF cable loss	6 dB

*These are nominal design goal values. Actual values are expected to be as follows: upper hemi gain of 1 dB with 140° cone (±70°) coverage and lower hemi gain of 1.5 dB with 120° × 100° coverage.

A7 KU-BAND TDRSS RELAY LINK

The Ku-band TDRSS relay link provides for two-way communication between the ground and the Orbiter. Similar to the S-band PM TDRSS link, the Ku-band link has the forward and return link segments.

The forward link carries the data from the ground via the TDRS to the Orbiter. The salient parameters of this link are summarized in Table A7.1. The primary operating mode is Mode 1, which employs a 216-kbps biphase-L-encoded data stream composed of (1) 72-kbps operational data, (2) 128-kbps scientific and instrument data, and (3) 16-kbps overhead. Because these data streams are time-division-multiplexed (TDM), a demultiplexer is employed at the Orbiter. The demultiplexer separates the 72-kbps data stream and applies it to the Network Signal Processor (NSP), which is not a part of the Ku-band equipment. The recovered 128-kbps stream may be routed to either an attached payload or other equipment on board the Orbiter.

In Mode 2, a selectable data configuration is available. Because no TDM is used, the demultiplexer is bypassed and any one of the data streams (32, 72, 96 or 216 kbps) is applied directly to the NSP.

In both modes, the forward link, like the S-band relay signal, is spread by superimposing a PN code on the data-modulated carrier. The Ku-band code rate, however, is 3.03 Mcps and the code length is 1023 chips.

The forward-link budget is given in Table A7.2. The acquisition and track entries refer to open-loop TDRSS pointing (acquisition) and closed-loop TDRSS pointing (track). No bit error rate requirements are imposed while TDRSS is acquiring angle track of the Shuttle, which is to be accomplished within 10 seconds of detection of the return-link signal from the Shuttle. Table A7.3 defines the performance characteristics of the Ku-band forward-link subsystem the PN despreader, carrier recovery loop and bit synchronizer. The cumulative implementation loss of these subsystems is 3.5 dB.

The return-link carries data via TDRSS from the Orbiter to the ground. This link has two modes: Mode 1 is a digital PM link, and Mode 2 is a composite digital/analog link employing FM. The return-link channelization is given in Table A7.4. Return-link power sharing is shown in Table A7.5. In both Modes 1 and 2, the relative power split between the two low data rate channels is determined by the phase shift on the 8.5-MHz subcarrier. The subcarrier to Channel 3 power split is determined by the relative amplitude of the 8.5-MHz subcarrier

and Channel 3 convolutional encoder output in Mode 1, or the Channel 3 data in Mode 2 The power sharing shown for Mode 2 corresponds to peak input amplitude.

Return-link throughput fidelity is determined not only by the Orbiter Ku-band system, but also by the total TDRSS performance. However, if the Orbiter Ku-band system performs within specified limits, the total Ku-band distortion will be within specified limits, as defined in [1]. The relevant Orbiter Ku-band constraints are given in Table A7.6. Parameters specific to the transmitter are given in Table A7.7.

[1] TDRSS "Telecommunication Performance Interface Document (TPID)," TRW Document No. 29000-200-016, dated February 7, 1979.

Table A7.1. Forward-Link Data Parameters

Channel Number	Mode 1		Mode 2	
	Rates	Modulation Type	Rates	Modulation Type
Single Channel Only	216 kbps (Composite, see Comments)	Biphase on carrier Data format biϕ-L	32 kbps or 72 kbps or 96 kbps or 216 kbps (See Comment)	Biphase on carrier
Comments	1. 216 kbps is composed of: 72 kbps operational data 128 kbps scientific and instrument data 16 kbps overhead All three are time-division-multiplexed (TDM) 2. PN code superimposed on carrier to reduce spectral density		PN code superimposed on carrier to reduce spectral density.	

NOTES: 1. Orbiter receives forward link data on carrier frequency of f_{RX} = 13.775 GHz.

2. PN code clock = 3.028 Mcps; code length = 1023 chips.

Table A7.2. Forward-Link Budget

Item	Nominal Specifications	Worst Case	
		Acquisition	Track
1. TDRS EIRP, dBW	49.5	40.0	46.5
2. Path Loss, dB/m^2	-163.5	-163.5	-163.5
3. Incident Flux Density, dBW/m^2	-114.0	-123.5	-117.0
4. Antenna Area, dBW	-5.7	-5.7	-5.7
5. Pointing Loss, dB	-0.3	-0.3	-0.3
6. Received Power, dBW	-120.0	-129.5	-123.0
7. System Temperature, dB°K	32.0	32.0	32.0
8. Boltzmann's Constant, dBW/°K Hz	-228.6	-228.6	-228.6
9. N_0, dBW/Hz	-196.6	-196.6	-196.6
10. C/N_0, dBHz	76.6	67.1	73.6
11. Data Rate, dB-Hz (216 kbps)	53.3	53.3	53.3
12. Received E_b/N_0, dB	23.3	13.8	20.3
13. Implementation Loss, dB	-3.5	-3.5	-3.5
14. E_b/N_0, dB	19.8	10.3	16.8
15. Required E_b/N_0, dB (10^{-6} BER)	10.5	DNA	10.5
16. Margin, dB	9.3	DNA	6.3
17. Bit Error Rate	$\ll 10^{-6}$	DNA	$\ll 10^{-6}$

Table A7.3. Forward-Link Performance Characteristics

PN Code Chip Rate	3.028031 MHz ±1 Hz
PN Code Acquisition Threshold (P_{ACQ} = 0.99)	64 dB-Hz
Minimum Available Acquisition C/N_0	67.1 dB-Hz
Maximum Time to Acquire PN Code at Threshold	10 seconds
False-Alarm Probability (10-second interval)	$\leq 10^{-6}$
IF Center Frequency	647.025 MHz ± 141 kHz
Carrier Acquisition Threshold (P_{ACQ} = 0.99)	64 dB-Hz
Minimum Available Acquisition C/N_0, TDRS Open-Loop Pointing	67.1 dB-Hz
Maximum Time to Acquire Carrier at Threshold	0.6 second
Implementation Loss in PN and Carrier-Tracking Loops	2.0 dB
Data Rate (Biphase-L)	216,000 bps
Prebit Synchronizer Filter One-Sided Noise Bandwidth	2.16 MHz
Bit Synchronizer Acquisition Threshold (P_{ACQ} = 0.99)	10.5 dB
Minimum Available E_b/N_0, TDRS Open-Loop Pointing	10.3 dB
Minimum Available E_b/N_0, TDRS Angle Tracking	16.8 dB
Maximum Time to Bit and Frame Synchronization at $E_b/N_0 \geq 0$ dB	10 seconds
Implementation Loss in Bit Synchronization	1.5 dB
Total Implementation Loss	3.5 dB

Table A7.4. Return-Link Source Data--Channelization Summary

Source	Data Rate	Channel	Mode	Data Type
OPS NSP 1	192 kbps	1	1 or 2	Biphase-L
OPS NPS 2	192 kbps	1	1 or 2	Biphase-L
LDR Payload Digital	16 kbps - 2 Mbps/ 16 kbps - 1.024 Mbps	2	1 or 2	NRZ-L-M-S/ Biϕ-L-M-S
Payload Recorder Digital	25.5 kbps-1.024 Mbps	2	1 or 2	Biphase-L
OPS Recorder System	25.5 kbps-1.024 Mbps	2	1 or 2	Biphase-L
PI 1 (LDR)	Channel 2 Compatible	2	1 or 2	Arbitrary
PI 1 (HDR)	0 - 4.5 MHz	3	2	Arbitrary
PI 2 (LDR)	Channel 2 Compatible	2	1 or 2	Arbitrary
PI 2 (HDR)	0 - 4.5 MHz	3	2	Arbitrary
HDR Payload Digital	16 kbps - 4 Mbps	3	2	NRZ-L-M-S
HDR Payload Analog	0 - 4.5 MHz	3	2	Analog
CCTV	0 - 4.5 MHz	3	2	Analog TV
HDR Digital Max	2 - 50 Mbps	3	1	NRZ-L-M-S

Table A7.5. Return-Link Power Summary

Mode	Channel	Nominal Performance	Tolerance
1	1*	4.4%	3% Minimum
	2*	11.8%	10.4% Minimum
	3	80%	75% Minimum
2	1+2	± 6 MHz Deviation	+30%/-20%
	3	±11 MHz Deviation	±5%

*Includes subcarrier bandpass filter loss

Table A7.6. Return-Link Distortion Requirements

Parameter	Requirement
Transmitter Noise	
• Mode 1	≤ 17° RMS (1 Hz to 1 kHz) ≤ 3° RMS (1 kHz to 50 MHz)
• Mode 2	≤ 75 kHz RMS (1-second average)
Spurious	
• Radiated Emissions	≤ -40 dBc (inside data band) ≤ -60 dBc (outside data band)
• Phase Modulation (Mode 1)	≤ 2° RMS (100 Hz to 100 MHz)
Incidental AM	≤ 5%
I/Q Orthogonality	
• Carrier (Mode 1)	90° ± 2.5°
• Subcarrier	90° ± 2.5°
Mode 2 Distortion	
• Differential Phase	±2°/1.5°/2° at 10/50/90 APL
• Differential Gain	±0.25 dB/0.15 dB/0.25 dB at 10/50/90 APL
Subcarrier	
• Frequency	8.5 MHz ± 600 Hz
• Phase Noise	≤ 1° RMS (1 kHz to 4 MHz)

Table A7.7. Return-Link Transmitter Parameters

Parameter	Requirement
Transmit Carrier Frequency	15.0034 GHz ± 0.001%
Narrow-Beam Antenna	
• Polarization	RHCP
• Axial Ratio	1.5 dB, maximum on axis 3 dB, maximum over 3-dB beamwidth
Wide-Beam Antenna	
• Alignment	Within ±1° of electrical axis of high-gain antenna
• Polarization	Linear
EIRP	
• High-Gain Antenna	52 dBW minimum*; 56 dBW maximum
• Wide-Beam Antenna	30 dBW minimum over minimum of ±15°
Transmitter Frequency Stability	
• Mode 1	$\pm 1 \times 10^{-9}$ 1-second $\pm 1 \times 10^{-5}$ 5-hour $\pm 1 \times 10^{-5}$ 48-hour
• Mode 2	± 1.5 MHz

*Includes pointing loss and polarization loss.

A8 TYPICAL NASA AND DOD TRANSPONDERS

Table A8.1 lists the salient parameters associated with typical NASA and DOD payload transponders.

Table A8.1. Typical Payload Transponder Characteristics

Item	Parameter and Range
Receive frequency range	
L-band frequency (DOD)	1760-1840 MHz
S-band frequency (NASA)	2025-2120 MHz
Transmitter frequency range	2200-2300 MHz
Tracking-loop bandwidth	18, 60, 200, or 2000 Hz
Tracking-loop order	Second
AGC dynamic range	100 dB
Command channel frequency response	1 kHz to 130 kHz
Ranging-channel frequency response	1 kHz to 1.2 MHz
Noise figure	5 dB to 8 dB
Transmitter phase deviation	Up to 2.5 rad
Transmitter output power	200 mW to 5 W*

*Up to 200 W with external power amplifiers

APPENDIX B

TYPICAL DESIGN CONTROL TABLES

B1 TYPICAL DESIGN CONTROL TABLES

The tables in this appendix have been assembled using preliminary and unofficial data pertaining to the parameter values. Many of the numbers are specified or expected values and do not reflect actual hardware capability. They are reasonable, however, and are based upon sound engineering judgment and past experience. Tolerances have not been supplied as it is too early for rational values to be assigned to many parameters.

Table B1.1. Orbiter-to-Payload S-Band Total Channel Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
T1	Total Transmit Power	P_T	dBW	7.0				
T2	Transmit Circuit Losses	L_{TX}	dB	-6.0				
T3	Transmit Antenna Gain	G_T	dB	2.5				
T4	Transmit Pointing Loss	L_{PTX}	dB	0.0				
T5	Space Loss 10 nmi f = 2100 MHz	L_S	dB	-124.2				
T6	Atmospheric Loss	L_A	dB	0.0				
T7	Receive Antenna Gain	G_R	dB	1.5				
T8	Polarization Loss	L_{POL}	dB	-0.3				
T9	Receive Pointing Loss	L_{PR}	dB	0.0				
T10	Receive Circuit Losses	L_R	dB	-2.0				
T11	Total Received Power (Σ 1 through 10)	P_R	dBW	-121.5				
T12	Receiver Noise Density NF = 7.0 dB	N_0	dBW/Hz	-197.0				
T13	Receiver Signal/Noise Density (11 - 12)	P_R/N_0	dB-Hz	75.5				

Table B1.1. Orbiter-to-Payload S-Band Carrier Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
CT1	Receiver Signal/Noise Density (T13)	P_R/N_0	dB-Hz	75.5				
CT2	Carrier/Total Power (1.0 rad)	P_C/P_R	dB	-2.3				
CT3	Noisy Oscillator Loss	L_{OSC}	dB	-0.1				
CT4	PN Filtering Loss	L_{PN}	dB	0.0				
CT5	PN Correlation Loss	L_{COR}	dB	0.0				
CT6	Squaring Loss	L_{SQ}	dB	0.0				
CT7	Limiter Loss	L_{LIM}	dB	-0.3				
CT8	Net Carrier Signal/Noise Density (Σ1 through 7)	P_{NC}/N_0	dB-Hz	72.8				
CT9	Carrier Loop Noise Bandwidth, Two-Sided (2000 Hz)	$2B_L$	dB-Hz	33.0				
CT10	Carrier Loop SNR (8 - 9)	$(\frac{S}{N})_{2B_L}$	dB-Hz	39.8				
CT11	Required Carrier Loop Minimum SNR		dB	15.0				
CT12	Carrier Loop SNR Margin (10 - 11)		dB	24.8				
CT13	Adverse Margin [12+(12 Adv. Tol.)]		dB					
CT14	Mean Tolerance Margin [12+(12 Mean Tol.)]		dB					
CT15	3σ Tolerance Margin [14+3×(12 Std. Dev.)]		dB					

Table B1.3. **Orbiter-to-Payload S-Band NASA Command Design Control Table**

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
DC1	Receiver Signal/Noise Density (T13)	P_R/N_0	dB-Hz	75.5				
DC2	Data/Total Power (1.0 rad)	P_D/P_R	dB	-4.1				
DC3	PN Filtering Loss (CT4)	L_{PN}	dB	0.0				
DC4	PN Correlation Loss (CT5)	L_{COR}	dB	0.0				
DC5	Carrier Phase Noise Loss	L_ϕ	dB	-0.1				
DC6	Subcarrier Demodulation Loss	L_{SD}	dB	-0.4				
DC7	Bit Synchronization Loss	L_{BS}	dB	-0.3				
DC8	Miscellaneous Waveform Loss	L_{WF}	dB	-0.2				
DC9	Data Bit Rate (2000 bps)	R_b	dB-bps	33.0				
DC10	Net Bit-Energy/Noise-Density [(Σ1 through 8) - 9]	E_b/N_0	dB	37.4				
DC11	Required E_b/N_0 $P_e = 10^{-5}$ Coded ___ Uncoded x		dB	9.6				
DC12	Data Channel Margin (10 - 11)		dB	27.8				
DC13	Adverse Margin [12 + (12 Adv. Tol.)]		dB					
DC14	Mean Tolerance Margin [12 + (12 Mean Tol.)]		dB					
DC15	3σ Tolerance Margin [14 + 3×(12 Std. Dev.)]		dB					

Table B2.1. Payload-to-Orbiter S-Band Total Channel Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
T1	Total Transmit Power	P_T	dBW	3.0				
T2	Transmit Circuit Losses	L_{TX}	dB	-2.0				
T3	Transmit Antenna Gain	G_T	dB	0.0				
T4	Transmit Pointing Loss	L_{PTX}	dB	0.0				
T5	Space Loss 10 nmi f = 2300 MHz	L_S	dB	-125.0				
T6	Atmospheric Loss	L_A	dB	0.0				
T7	Receive Antenna Gain	G_R	dB	2.5				
T8	Polarization Loss	L_{POL}	dB	-0.5				
T9	Receive Pointing Loss	L_{PR}	dB	0.0				
T10	Receive Circuit Losses	L_R	dB	-6.0				
T11	Total Received Power (Σ1 through 10)	P_R	dBW	-128.0				
T12	Receiver Noise Density NF = 7.0 dB	N_0	dBW/Hz	-197.0				
T13	Receiver Signal/Noise Density (11 - 12)	P_R/N_0	dB-Hz	69.0				

Table B2.2. Payload-to-Orbiter S-Band Carrier Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
CT1	Receiver Signal/Noise Density (T13)	P_R/N_0	dB-Hz	69.0				
CT2	Carrier/Total Power (1.0 rad)	P_C/P_R	dB	-2.3				
CT3	Noisy Oscillator Loss	L_{OSC}	dB	-0.2				
CT4	PN Filtering Loss	L_{PN}	dB	0.0				
CT5	PN Correlation Loss	L_{COR}	dB	0.0				
CT6	Squaring Loss	L_{SQ}	dB	0.0				
CT7	Limiter Loss	L_{LIM}	dB	0.0				
CT8	Net Carrier Signal/Noise Density (Σ1 through 7)	P_{NC}/N_0	dB-Hz	67 4				
CT9	Carrier Loop Noise Bandwidth, Two-Sided (1260 Hz)	$2B_L$	dB-Hz	31.0				
CT10	Carrier Loop SNR (8 - 9)	$(\frac{S}{N})_{2B_L}$	dB-Hz	36.4				
CT11	Required Carrier Loop Minimum SNR		dB	15.0				
CT12	Carrier Loop SNR Margin (10 - 11)		dB	21.4				
CT13	Adverse Margin [12+(12 Adv. Tol.)]		dB					
CT14	Mean Tolerance Margin [12+(12 Mean Tol.)]		dB					
CT15	3σ Tolerance Margin [12+3×(12 Std. Dev.)]		dB					

Table B2.3. Payload-to-Orbiter S-Band Telemetry Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
DC1	Receiver Signal/Noise Density (T13)	P_R/N_0	dB-Hz	69.0				
DC2	Data/Total Power (1.0 rad)	P_D/P_R	dB	-4.1				
DC3	PN Filtering Loss (CT4)	L_{PN}	dB	0.0				
DC4	PN Correlation Loss (CT5)	L_{COR}	dB	0.0				
DC5	Carrier Phase Noise Loss	L_ϕ	dB	-0.1				
DC6	Subcarrier Demodulation Loss	L_{SD}	dB	-0.5				
DC7	Bit Synchronization Loss	L_{BS}	dB	-1.0				
DC8	Miscellaneous Waveform Loss	L_{WF}	dB	0.0				
DC9	Data Bit Rate (16 kbps)	R_b	dB-bps	42.0				
DC10	Net Bit-Energy/Noise-Density [(Σ1 through 8) - 9]	E_b/N_0	dB	21.3				
DC11	Required E_b/N_0 P_e = 10^{-5} Coded ___ Uncoded x		dB	9.6				
DC12	Data Channel Margin (10 - 11)		dB	11.7				
DC13	Adverse Margin [12+(12 Adv. Tol.)]		dB					
DC14	Mean Tolerance Margin [12+(12 Mean Tol.)]		dB					
DC15	3σ Tolerance Margin [14+3×(12 Std. Dev.)]		dB					

Table B3.1. Orbiter-to-Payload L-Band Total Channel Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
T1	Total Transmit Power	P_T	dBW	7.0				
T2	Transmit Circuit Losses	L_{TX}	dB	-6.0				
T3	Transmit Antenna Gain	G_T	dB	2.5				
T4	Transmit Pointing Loss	L_{PTX}	dB	0.0				
T5	Space Loss 10 nmi f = 1800 MHz	L_S	dB	-122.9				
T6	Atmospheric Loss	L_A	dB	0.0				
T7	Receive Antenna Gain	G_R	dB	0.0				
T8	Polarization Loss	L_{POL}	dB	-0.3				
T9	Receive Pointing Loss	L_{PR}	dB	0.0				
T10	Receive Circuit Losses	L_R	dB	-1.0				
T11	Total Received Power (Σ1 through 10)	P_R	dBW	-120.7				
T12	Receiver Noise Density NF = 5.0 dB	N_0	dBW/Hz	-199.0				
T13	Receiver Signal/Noise Density (11 - 12)	P_R/N_0	dB-Hz	78.3				

Table B3.2. Orbiter-to-Payload L-Band Carrier Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
CT1	Receiver Signal/Noise Density (T13)	P_R/N_0	dB-Hz	78.3				
CT2	Carrier/Total Power (0.3 rad)	P_C/P_R	dB	-0.2				
CT3	Noisy Oscillator Loss	L_{OSC}	dB	-0.1				
CT4	PN Filtering Loss	L_{PN}	dB	0.0				
CT5	PN Correlation Loss	L_{COR}	dB	0.0				
CT6	Squaring Loss	L_{SQ}	dB	0.0				
CT7	Limiter Loss	L_{LIM}	dB	0.0				
CT8	Net Carrier Signal/Noise-Density (Σ1 through 7)	P_{NC}/N_0	dB-Hz	78.0				
CT9	Carrier Loop Noise Bandwidth, Two-Sided (4000 Hz)	$2B_L$	dB-Hz	36.0				
CT10	Carrier Loop SNR (8 - 9)	$(\frac{S}{N})_{2B_L}$	dB-Hz	42.0				
CT11	Required Carrier Loop Minimum SNR		dB	20.0				
CT12	Carrier Loop SNR Margin (10 - 11)		dB	22.0				
CT13	Adverse Margin [12 + (12 Adv. Tol.)]		dB					
CT14	Mean Tolerance Margin [12 + (12 Mean Tol.)]		dB					
CT15	3σ Tolerance Margin [14 + 3×(12 Std. Dev.)]		dB					

Table B3.3. Orbiter-to-Payload L-Band DOD Command Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
DC1	Receiver Signal/Noise Density (T13)	P_R/N_0	dB-Hz	78.3				
DC2	Data/Total Power (0.3 rad)	P_D/P_R	dB	-13.6				
DC3	PN Filtering Loss (CT4)	L_{PN}	dB	0.0				
DC4	PN Correlation Loss (CT5)	L_{COR}	dB	0.0				
DC5	Carrier Phase Noise Loss	L_ϕ	dB	-0.1				
DC6	Subcarrier Demodulation Loss	L_{SD}	dB	-3.2				
DC7	Bit Synchronization Loss	L_{BS}	dB					
DC8	Miscellaneous Waveform Loss	L_{WF}	dB					
DC9	Data Bit Rate (2000 sps)	R_b	dB-bps	33.0				
DC10	Net Bit-Energy/Noise-Density [(Σ1 through 8) - 9]	E_b/N_0	dB	28.4				
DC11	Required E_b/N_0 Coded ___ P_e = 10^{-5} Uncoded x		dB	13.3				
DC12	Data Channel Margin (10 - 11)		dB	15.1				
DC13	Adverse Margin [12 + (12 Adv. Tol.)]		dB					
DC14	Mean Tolerance Margin [12 + (12 Mean Tol.)]		dB					
DC15	3σ Tolerance Margin [14 + 3×(12 Std. Dev.)]		dB					

Table B4.1. S-Band PM Direct-Link Total Channel Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
T1	Total Transmit Power	P_T	dBW	3.0				
T2	Transmit Circuit Losses	L_{TX}	dB	-6.6				
T3	Transmit Antenna Gain	G_T	dB	3.0				
T4	Transmit Pointing Loss	L_{PTX}	dB	-0.5				
T5	Space Loss 1122 nmi f = 2287.5 MHz	L_S	dB	-166.0				
T6	Atmospheric Loss	L_A	dB	0.0				
T7	Receive Antenna Gain	G_R	dB	43.5				
T8	Polarization Loss	L_{POL}	dB	-0.5				
T9	Receive Pointing Loss	L_{PR}	dB	-0.5				
T10	Receive Circuit Losses	L_R	dB	0.0				
T11	Total Received Power (Σ1 through 10)	P_R	dBW	-124.6				
T12	Receiver Noise Density NF = -3.2 dB	N_0	dBW/Hz	-207.2				
T13	Receiver Signal/Noise Density (11 - 12)	P_R/N_0	dB-Hz	82.6				

Table B4.2. S-Band PM Direct-Link Data Channel Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
DC1	Receiver Signal/Noise Density (T13)	P_R/N_0	dB-Hz	82.6				
DC2	Data/Total Power (1.0 rad)	P_D/P_R	dB	-1.5				
DC3	PN Filtering Loss (CT4)	L_{PN}	dB	0.0				
DC4	PN Correlation Loss (CT5)	L_{COR}	dB	0.0				
DC5	Carrier Phase Noise Loss	L	dB	-0.1				
DC6	Subcarrier Demodulation Loss	L_{SD}	dB	0.0				
DC7	Bit Synchronization Loss	L_{BS}	dB	-1.5				
DC8	Miscellaneous Waveform Loss	L_{WF}	dB	-0.2				
DC9	Data Bit Rate (192 kbps)	R_b	dB-bps	52.8				
DC10	Net Bit-Energy/Noise Density [(Σ1 through 8) - 9]	E_b/N_0	dB	26.5				
DC11	Required E_b/N_0 $P_e = 10^{-4}$ Coded ___ Uncoded x		dB	8.4				
DC12	Data Channel Margin (10 - 11)		dB	18.1				
DC13	Adverse Margin [12 + (12 Adv. Tol.)]		dB					
DC14	Mean Tolerance Margin [12 + (12 Mean Tol.)]		dB					
DC15	3σ Tolerance Margin [14 + 3×(12 Std. Dev.)]		dB					

Table B5.1. S-Band FM Direct-Link Total Channel Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
T1	Total Transmit Power	P_T	dBW	10.0				
T2	Transmit Circuit Losses	L_{TX}	dB	-7.2				
T3	Transmit Antenna Gain	G_T	dB	1.0				
T4	Transmit Pointing Loss	L_{PTX}	dB	0.0				
T5	Space Loss 1122 nmi f = 2250 MHz	L_S	dB	-165.8				
T6	Atmospheric Loss	L_A	dB	0.0				
T7	Receive Antenna Gain	G_R	dB	43.5				
T8	Polarization Loss	L_{POL}	dB	-0.5				
T9	Receive Pointing Loss	L_{PR}	dB	-0.5				
T10	Receive Circuit Losses	L_R	dB	0.0				
T11	Total Received Power (Σ1 through 10)	P_R	dBW	-119.5				
T12	Receiver Noise Density NF = -3.2 dB	N_0	dBW/Hz	-207.2				
T13	Receiver Signal/Noise Density (11 - 12)	P_R/N_0	dB-Hz	87.7				

Table B5.2. FM Direct-Link Analog Signal Design Control Table

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
A1	Receiver Signal/Noise Density (T13)	P_R/N_0	dB-Hz	87.7				
A2	Predemodulation Bandwidth (13.2 MHz)	B_i	dB-Hz	71.2				
A3	Signal Filtering Loss	L_{SF}	dB	-1.0				
A4	Predemodulation SNR (1 - 2 + 3)	$(\frac{S}{N})_{B_i}$	dB	15.5				
A5	Peak Carrier Deviation	Δf	MHz	4.0				
A6	Postdemodulation Bandwidth	B_o	MHz	2.0				
A7	Demodulation Output SNR (Formula)	$(SNR)_0$	dB	34.5				
A8	Minimum Required (SNR)		dB	15.5				
A9	$(SNR)_0$ Margin (7 - 8)		dB	19.5				
A10	Adverse Margin [9 + (9 Adv. Tol.)]		dB					
A11	Mean Tolerance Margin [9 + (9 Mean Tol.)]		dB					
A12	3σ Tolerance Margin [11 + 3×(9 Std. Dev.)]		dB					

Table B6.1. Design Control Table, Total Signal for Ku-Band TDRSS Return Link

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
T1	Total Transmit Power	P_T	dBW	17.0				
T2	Transmit Circuit Losses	L_{TX}	dB	-3.6				
T3	Transmit Antenna Gain	G_T	dB	35.4				
T4	Transmit Pointing Loss	L_{PTX}	dB	-0.7				
T5	Space Loss 22,786 nmi f = 15,000 MHz	L_S	dB	-208.5				
T6	Receive Antenna Gain	G_R	dB	52.6				
T7	Polarization Loss	L_{POL}	dB	-0.3				
T8	Receive Pointing Loss	L_{PR}	dB	-0.5				
T9	Receive Circuit Losses	L_R	dB	-0.3				
T10	Atmospheric Loss	L_A	dB	-1.0				
T11	TDRS Loss	L_{TDRS}	dB	-2.0				
T12	Total Received Power (Σ1 through 11)	P_R	dBW	-111.9				
T13	Receiver Noise Density NF = 4.9 dB	N_0	dBW/Hz	-199.1				
T14	Receiver Signal/Noise Density (12 - 13)	P_R/N_0	dB-Hz	87.2				

Table B6.2. Design Control Table, Data Channel for Ku-Band TDRSS Return Link (Mode 1, Channel 3)

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
DC1	Receiver Signal/Noise Density (T14)	P_R/N_0	dB-Hz	87.2				
DC2	Data/Total Power	P_D/P_R	dB	-1.0				
DC3	PN Filtering Loss (CT4)	L_{PN}	dB	0.0				
DC4	PN Correlation Loss (CT5)	L_{COR}	dB	0.0				
DC5	Carrier Phase Noise Loss	L_ϕ	dB	0.0				
DC6	Subcarrier Demodulation Loss	L_{SD}	dB	0.0				
DC7	Bit Synchronization Loss	L_{BS}	dB	-1.0				
DC8	Miscellaneous Waveform Loss	L_{WF}	dB	-0.5				
DC9	Data Bit Rate (50 Mbps)	R_b	dB-bps	77.0				
DC10	Net Bit-Energy/Noise-Density [(Σ 1 through 8) - 9]	E_b/N_0	dB	7.7				
DC11	Required E_b/N_0 Coded x / Uncoded ___ P_e = 10^{-4}		dB	3.2				
DC12	Data Channel Margin (10 - 11)		dB	4.5				
DC13	Adverse Margin [12 + (12 Adv. Tol.)]		dB					
DC14	Mean Tolerance Margin [12 + (12 Mean Tol.)]		dB					
DC15	3σ Tolerance Margin [T4 + 3×(12 Std. Dev.)]		dB					

Table B7.1. Design Control Table, Total Signal for a Bent-Pipe Digital Channel

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
T1	Total Transmit Power	P_T	dBW	17.0				
T2	Transmit Circuit Losses	L_{TX}	dB	-3.6				
T3	Transmit Antenna Gain	G_T	dB	35.4				
T4	Transmit Pointing Loss	L_{PTX}	dB	-0.7				
T5	Space Loss 22,786 nmi f = 15,000 MHz	L_S	dB	-208.5				
T6	Receive Antenna Gain	G_R	dB	52.6				
T7	Polarization Loss	L_{POL}	dB	-0.3				
T8	Receive Pointing Loss	L_{PR}	dB	-0.5				
T9	Receive Circuit Losses	L_R	dB	-0.3				
T10	Atmospheric Loss	L_A	dB	-1.0				
T11	TDRS Loss	L_{TDRS}	dB	-2.0				
T12	Total Received Power (Σ 1 through 11)	P_R	dBW	-111.9				
T13	Receiver Noise Density NF = 4.9 dB	N_0	dBW/Hz	-199.1				

Table B7.2. Effective Signal and Noise Levels for a Bent-Pipe Digital Channel

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
BP1	Total Received Power (T12)	P_R	dBW	-111.9				
BP2	Ku-Band Channel/Total Power	P_{CH}/P_R	dB	-1.9				
BP3	Channel S+N Power (1 + 2)	P_{S+N}	dBW	-113.8				
BP4	S/(S+N) Power	P_{SR}/P_{S+N}	dB	-0.1 (hypothetical)				
BP5	Effective Received Signal Power (3 + 4)	P_{SR}	dBW	-113.9				
BP6	N/(S+N) Power	P_{NR}/P_{S+N}	dB	-16.4 (hypothetical)				
BP7	Effective Received Noise Power (3 + 6)	P_{NR}	dBW	-130.2				
BP8	Channel Noise Bandwidth (2.5 MHz)	B_{BP}	dB-Hz	64.0 (hypothetical)				
BP9	Effective Received Noise Density (7 - 8)	N_{01}	dBW/Hz	-194.2				
BP10	Receiver Noise Density (T13)	N_{02}	dBW/Hz	-199.1				
BP11	Total Noise Density (*)	N_0	dBW/Hz	-193.0				
BP12†	Channel Signal/Noise Density (5 - 11)	P_{SR}/N_0	dB-Hz	79.1				

*$N_0 = 10 \log \left[\log^{-1}\left(\frac{N_{01}}{10}\right) + \log^{-1}\left(\frac{N_{02}}{10}\right) \right]$

†NOTE: When supplying this entry to line DC1 of a Data Channel DCT, also enter values of zero on line DC2.

Ku-Band Mode 1
Ku-Band Channel 2

Table B7.3. Digital Data Channel for the TDRS Bent-Pipe Link

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
DC1	Receiver Signal/Noise Density (BP12)	P_R/N_0	dB-Hz	79.1				
DC2	Data/Total Power	P_D/P_R	dB	0.0				
DC3	PN Filtering Loss (CT4)	L_{PN}	dB	0.0				
DC4	PN Correlation Loss (CT5)	L_{COR}	dB	0.0				
DC5	Carrier Phase Noise Loss	L_ϕ	dB	0.0				
DC6	Subcarrier Demodulation Loss	L_{SD}	dB	-0.5				
DC7	Bit Sycnronization Loss	L_{BS}	dB	-1.0				
DC8	Miscellaneous Waveform Loss	L_{WF}	dB	-0.5				
DC9	Data Bit Rate (325 kbps)	R_b	dB-bps	55.1				
DC10	Net Bit-Energy/Noise-Density [(Σ 1 through 8) - 9]	E_b/N_0	dB	22.0				
DC11	Required E_b/N_0 P_e = 10^{-5} Coded ___ Uncoded x		dB	9.6				
DC12	Data Channel Margin (10 - 11)		dB	12.4				
DC13	Adverse Margin [12 + (12 Adv. Tol.)]		dB					
DC14	Mean Tolerance Margin [12 + (12 Mean Tol.)]		dB					
DC15	3σ Tolerance Margin [14 + 3×(12 Std. Dev.)]		dB					

Table B8.1. Design Control Table, Total Signal for a Bent-Pipe Analog Channel

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
T1	Total Transmit Power	P_T	dBW	17.0				
T2	Transmit Circuit Losses	L_{TX}	dB	-3.6				
T3	Transmit Antenna Gain	G_T	dB	35.4				
T4	Transmit Pointing Loss	L_{PTX}	dB	-0.7				
T5	Space Loss 22,786 nmi f = 15,000 MHz	L_S	dB	-208.5				
T6	Receive Antenna Gain	G_R	dB	52.6				
T7	Polarization Loss	L_{POL}	dB	-0.3				
T8	Receive Pointing Loss	L_{PR}	dB	-0.5				
T9	Receive Circuit Losses	L_R	dB	-0.3				
T10	Atmospheric Loss	L_A	dB	-1.0				
T11	TDRS Loss	L_{TDRS}	dB	-2.0				
T12	Total Received Power (Σ 1 through 11)	P_R	dBW	-111.9				
T13	Receiver Noise Density NF = 4.9 dB	N_0	dBW/Hz	-199.1				

Table B8.2. Effective Signal and Noise Levels for a Bent-Pipe Analog Channel

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
BP1	Total Received Power (T12)	P_R	dBW	-111.9				
BP2	Ku-Band Channel/Total Power	P_{CH}/P_R	dB	-1.0				
BP3	Channel S+N Power (1 + 2)	P_{S+N}	dBW	-112.9				
BP4	S/(S+N) Power	P_{SR}/P_{S+N}	dB	-1.1 (hypothetical)				
BP5	Effective Received Signal Power (3 + 4)	P_{SR}	dBW	-114.0				
BP6	N/(S+N) Power	P_{NR}/P_{S+N}	dB	-6.5 (hypothetical)				
BP7	Effective Received Noise Power (3 + 6)	P_{NR}	dBW	-119.4				
BP8	Channel Noise Bandwidth (5 MHz)	B_{BP}	dB-Hz	67.0 (hypothetical)				
BP9	Effective Received Noise Density (7 - 8)	N_{01}	dBW/Hz	-186.4				
BP10	Receiver Noise Density (T13)	N_{02}	dBW/Hz	-199.1				
BP11	Total Noise Density (*)	N_0	dBW/Hz	-186.2				
BP12†	Channel Signal/Noise Density (5 - 11)	P_{SR}/N_0	dB-Hz	72.2				

*$N_0 = 10 \log \left[\log^{-1}\left(\frac{N_{01}}{10}\right) + \log^{-1}\left(\frac{N_{02}}{10}\right) \right]$

†NOTE: When supplying this entry to line DC1 of a Data Channel DCT, also enter values of zero on line DC2.

Ku-Band Mode 2
Ku-Band Channel 3

Table B8.3. FM Analog Channel Design Control Table for the TDRS Bent-Pipe Link

Entry No.	Parameter (Calculation)	Symbol	Units	Nominal Value	Tolerance			
					Favorable	Adverse	Mean	Std. Dev.
A1	Receiver Signal/Noise Density (BP12)	P_R/N_0	dB-Hz	72.2				
A2	Predemodulation Bandwidth (10 MHz)	B_i	dB-Hz	70.0 (hypothetical)				
A3	Signal Filtering Loss	L_{SF}	dB	-1.0 (hypothetical)				
A4	Predemodulation SNR (1 - 2 + 3)	$(\frac{S}{N})_{B_i}$	dB	1.2				
A5	Peak Carrier Deviation	Δf	MHz	2.0 (hypothetical)				
A6	Postdemodulation Bandwidth	B_0	MHz	3.0 (hypothetical)				
A7	Demodulator Output SNR (Formula)	$(SNR)_0$	dB	?	Because the predemodulation SNR (A4) is below 10 dB, the FM demodulator is below threshold			
A8	Minimum Required $(SNR)_0$		dB	15.0				
A9	$(SNR)_0$ Margin (7 - 8)		dB					
A10	Adverse Margin [9 + (9 Adv. Tol.)]		dB					
A11	Mean Tolerance Margin [9 + (9 Mean Tol.)]		dB					
A12	3σ Tolerance Margin [11 + 3×(9 Std. Dev.)]		dB					

APPENDIX C

FUNCTIONAL PARAMETERS

C1 PAYLOAD INTERROGATOR

The function of the Payload Interrogator (PI) is to provide the RF communication link between the Orbiter and detached payloads. For communication with the NASA payloads, the PI operates in conjunction with the Payload Signal Processor (PSP). During DOD missions, the PI is interfaced with the Communication Interface Unit (CIU). Nonstandard (bent-pipe) data received by the PI from either NASA or DOD payloads is delivered to the Ku-band Signal Processor (KuSP), where it is processed for transmission to the ground via the Shuttle/TDRSS link.

Simultaneous RF transmission and reception is the primary mode of PI operation with both NASA and DOD payloads. The Orbiter-to-payload link carries the commands, while the payload-to-Orbiter link communicates the telemetry data. In addition to this duplex operation, the PI provides for "transmit only" and "receive only" modes of communication with some payloads.

Figure C.1 shows the functional block diagram for the Payload Interrogator. The antenna connects to an input/output RF port which is common to both the receiver and transmitter of the PI unit. Because of a requirement to operate the PI simultaneously with the Shuttle/ground S-band network transponder which radiates and receives on the same frequency bands, a dual triplexer is employed. The S-band network transponder emits a signal at either 2217.5 MHz or 2287.5 MHz; both frequencies thus fall directly into the PI receive band of 2200 MHz to 2300 MHz. Conversely, the payload transmitter, operating in either the 2025-2120 MHz (NASA) or 1764-1840 MHz (DOD) bands, can interfere with uplink signal reception by the S-band network transponder receiver. Therefore, by use of the triplexer and by simultaneously operating the PI and network transponder in the mutually exclusive subbands, the interference problem is effectively eliminated.

When detached payloads are in the immediate vicinity of the Orbiter, excessive RF power levels may impinge on the Payload Interrogator antenna. Thus, the RF preamplifier of the receiver is protected by a combination of sensitivity control attenuators and a diode breakdown limiter. The output of the preamplifier is applied to the first mixer where it is converted to the first IF for amplification and level control. The first

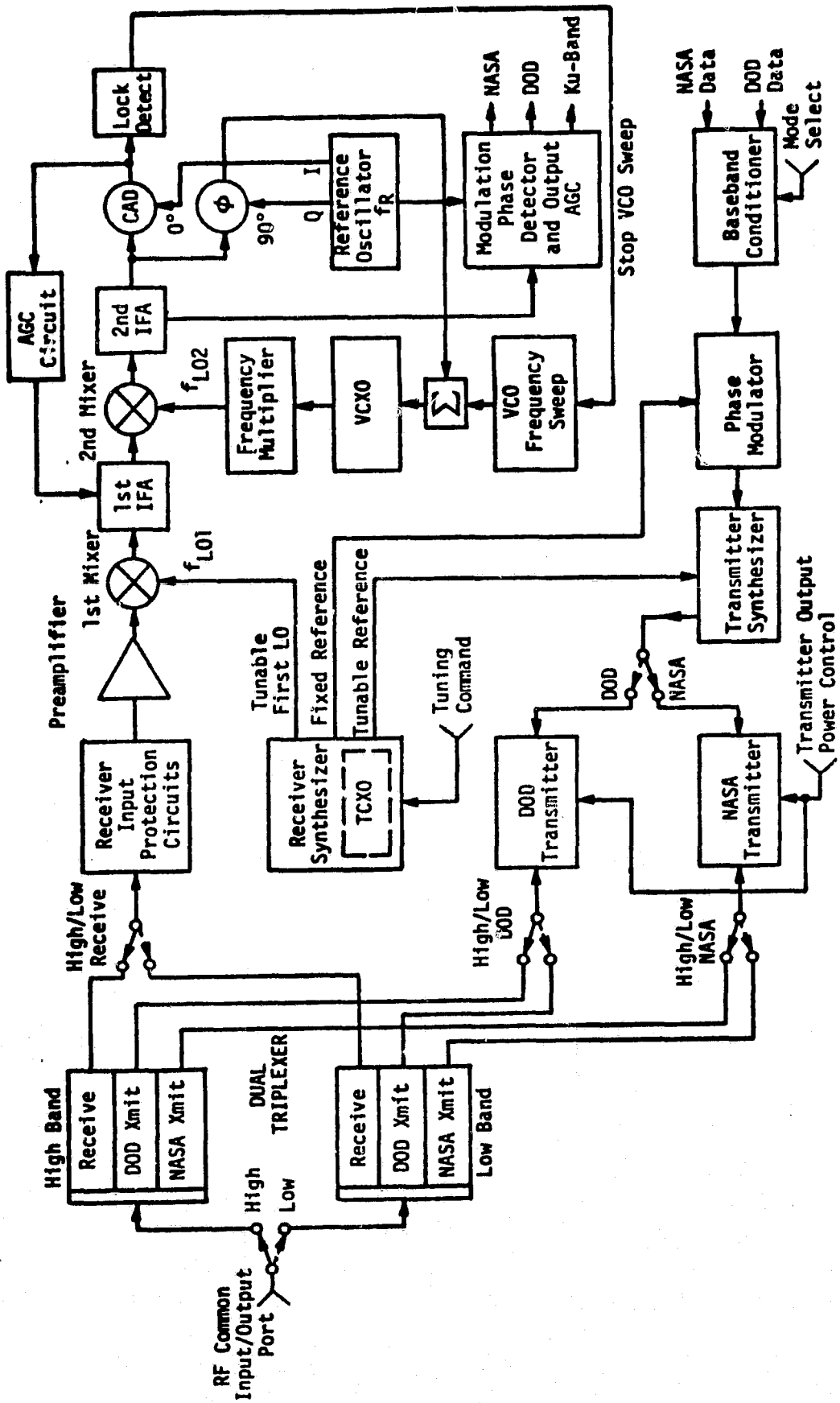


Figure C.1. Payload Interrogator Functional Block Diagram

local oscillator frequency, f_{LO1}, is tunable and its frequency corresponds with the desired PI receive channel frequency. Except for channel selection, however, f_{LO1} is fixed. Consequently, any unspecified frequency difference between the received payload signal and f_{LO1} will appear within the first IF amplifier and at the input to the second mixer.

The receiver frequency and phase tracking loop begin at the second mixer. As shown in Figure C.1, the output of the first IF amplifier is downconverted to the second IF as a result of mixing with a variable second LO frequency, f_{LO2}. The portion of the second IF which involves only the carrier-tracking function is narrowband, passing the received signal residual carrier component and excluding the bulk of the sideband frequencies. Demodulation to baseband of the second IF signal is accomplished by mixing with a reference frequency, f_R. The output of the tracking-phase detector, after proper filtering, is applied to the control terminals of a VCO which provides the second local oscillator signal, thereby closing the tracking loop. Thus, when phase track is established, f_{LO2} follows frequency changes of the received payload signal.

For the purpose of frequency acquisition, the f_{LO2} may be swept over a ±75 kHz uncertainty region. Sweep is terminated when the output of the coherent amplitude detector (CAD) exceeds a preset threshold, indicating that the carrier-tracking loop has attained lock. The output of the CAD also provides the AGC to the first IF amplifier. To accommodate payload-to-Orbiter received signal level changes due to range variation from about a few feet to 10 nmi, 110 dB of AGC is provided in the first IF amplifier.

A wideband phase detector is used to demodulate the telemetry signals from the carrier. The output of this detector is filtered, envelope-level controlled, and buffered for delivery to the PSP, CIU, and Ku-band Signal Processor units.

The PI receiver frequency synthesizer provides the tunable first LO frequency and the corresponding exciter frequency to the transmitter synthesizer. It also delivers a reference signal to the transmitter phase modulator. Baseband NASA or DOD command signals modulate the phase of this reference signal, which is in turn supplied to the transmitter synthesizer, where it is upconverted to either the NASA or DOD transmit frequency and applied to the power amplifiers.

For transmitter efficiency optimization, separate NASA and DOD RF power amplifier units are used. Depending on the operating band selected, transmitter output is applied to either the high- or low-band triplexer. To compensate for varying distances to payloads, each transmitter has three selectable output power levels.

C2 PAYLOAD SIGNAL PROCESSOR

The Payload Signal Processor (PSP) performs the following functions: (1) it modulates NASA payload commands onto a 16-kHz sinusoidal subcarrier and delivers the resultant signal to the PI and the attached payload umbilical, (2) it demodulates the payload telemetry data from the 1.024-MHz subcarrier signal provided by the PI, and (3) it performs bit and frame synchronization of demodulated telemetry data and delivers this data and its clock to the Payload Data Interleaver (PDI).

The PSP also transmits status messages to the Orbiter's general-purpose computer (GPC); the status messages allow the GPC to control and configure the PSP and validate command messages prior to transmission.

The functional block diagram for the PSP is shown in Figure C.2. The PSP configuration and payload command data are input to the PSP via a bidirectional serial interface. Transfer of data in either direction is initiated by discrete control signals. Data words 20 bits in length (16 information, 1 parity, 3 synchronization) are transferred across the bidirectional interface at a burst rate of 1-Mbps, and the serial words received by the PSP are applied to word validation logic which examines their structure. Failure to the incoming message to pass a validation test results in a request for a repeat of the message from the GPC.

Command data is further processed and validated as to content and the number of command words. The function of the command buffers is to perform data rate conversion from the 1-Mbps bursts to one of the selected standard command rates. Command rate and format are specified through the configuration message control subunit.

From the message buffers, the command bits are fed via the idle pattern selector and generator to the 16-kHz subcarrier biphase modulator. The idle pattern (which in many cases consists of alternating "ones" and "zeros") precedes the actual command word and is usually also transmitted in lieu of command messages. Subcarrier modulation is biphase NRZ only.

The 1.024-MHz telemetry subcarrier from the PI is applied to the PSK subcarrier demodulator. Since the subcarrier is biphase modulated, a Costas-type loop is used to lock onto and track the subcarrier. The resulting demodulated bit stream is input to the bit synchronizer subunit, where a DTTL bit synchronization loop provides timing to an

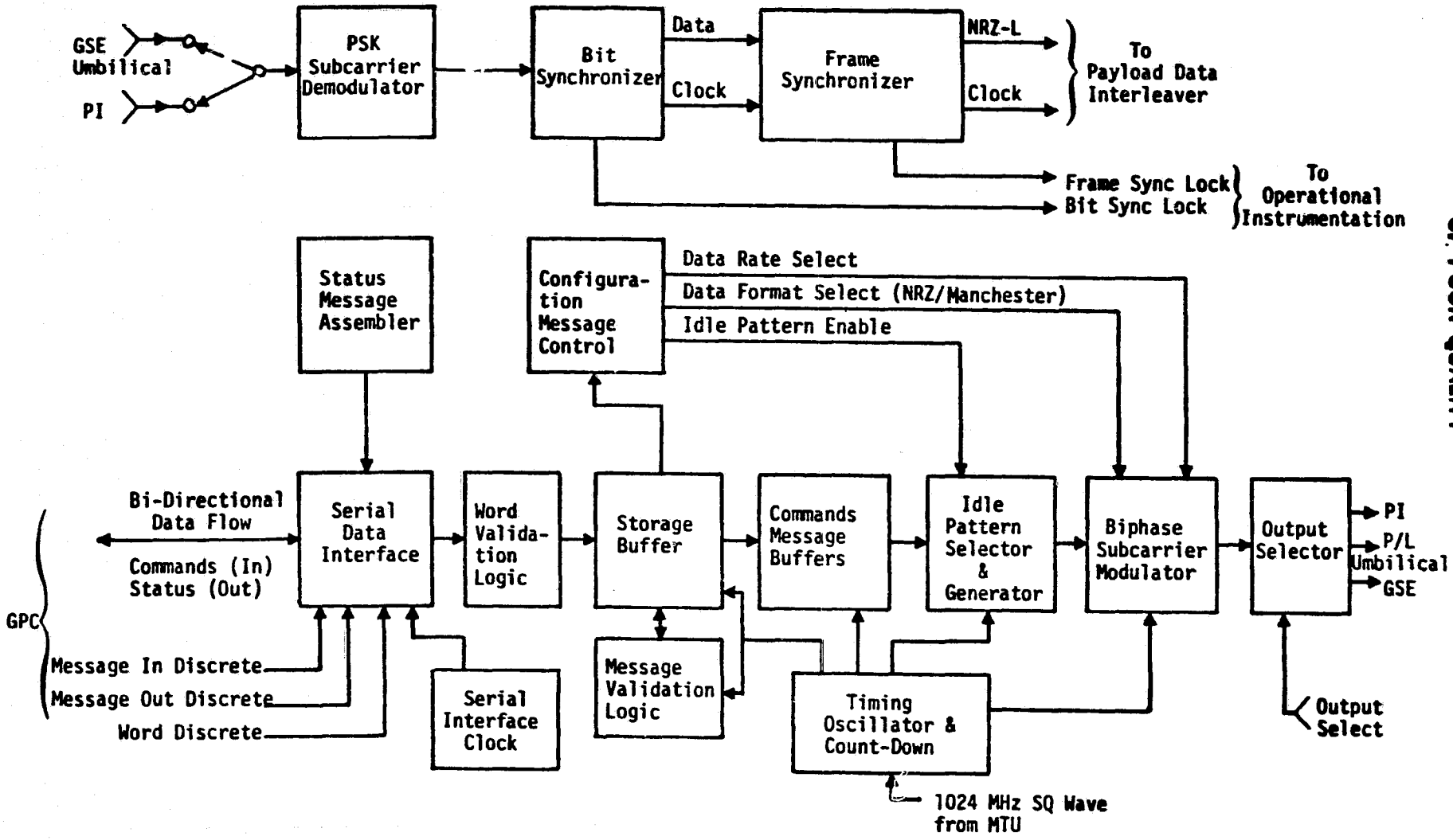


Figure C.2. NASA Payload Signal Processor Functional Block Diagram

integrate-and-dump matched filter which optimally detects and reclocks the telemetry data.

Detected telemetry bits, together with clock, are input to the frame synchronizer where frame synchronization is obtained for any one of the four NASA standard synchronization words. The frame synchronizer also detects and corrects the data polarity ambiguity caused by the PSK demodulator Costas loop.

From the frame synchronizer, the telemetry data with corrected frame synchronization words and clock are fed to the PI. The telemetry detection units also supply appropriate lock signals to the Orbiter's operational instrumentation equipment, thus acting to indicate the presence of valid telemetry.

C3 PAYLOAD DATA INTERLEAVER

A block diagram of the Payload Data Interleaver (PDI) is shown in Figure C.3. The PDI is basically a multiplexer capable of combining various asynchronous data streams into a single serial data stream. The PDI provides for the reception of up to six asynchronous payload pulse code modulation (PCM) streams, five from attached payloads and one from the PSP that is active (detached payload). An input switch matrix selects four of the inputs for the bit synchronizers. The "chain" functions of bit synchronization, decommutation, and word selection are provided for up to four simultaneous PCM streams in two possible modes.

Mode 1

In this mode of operation, a chain bit synchronizes, master frame synchronizes, minor frame synchronizes, and word synchronizes to the incoming data stream. The word selector blocks data into the proper words for storage in the data random access memory (RAM), and/or toggle buffer (TB). PCM code type, bit rate, PCM format, synchronization codes, and word selection are programmable under control of the decommutator format memories. Two word selection capabilities for this mode of operation are as follows:

Type I: The first type selects all of, or a subset of, the words in a payload PCM format minor frame (or master frame for formats without minor frames) for storage in the toggle buffer.

Type II: The second type of word selection is by parameter. The specification of a parameter consists of its word location within a minor frame, the first minor frame in which it appears, and its sample rate. The specification is provided as part of the decommutator control memory format load.

Mode 2

In this mode of operation a chain bit synchronizes to the incoming data, blocks it into eight-bit words, blocks the eight-bit words into frames, supplies synchronization pattern at the start of each frame, and includes the status register as the last three 16-bit words of

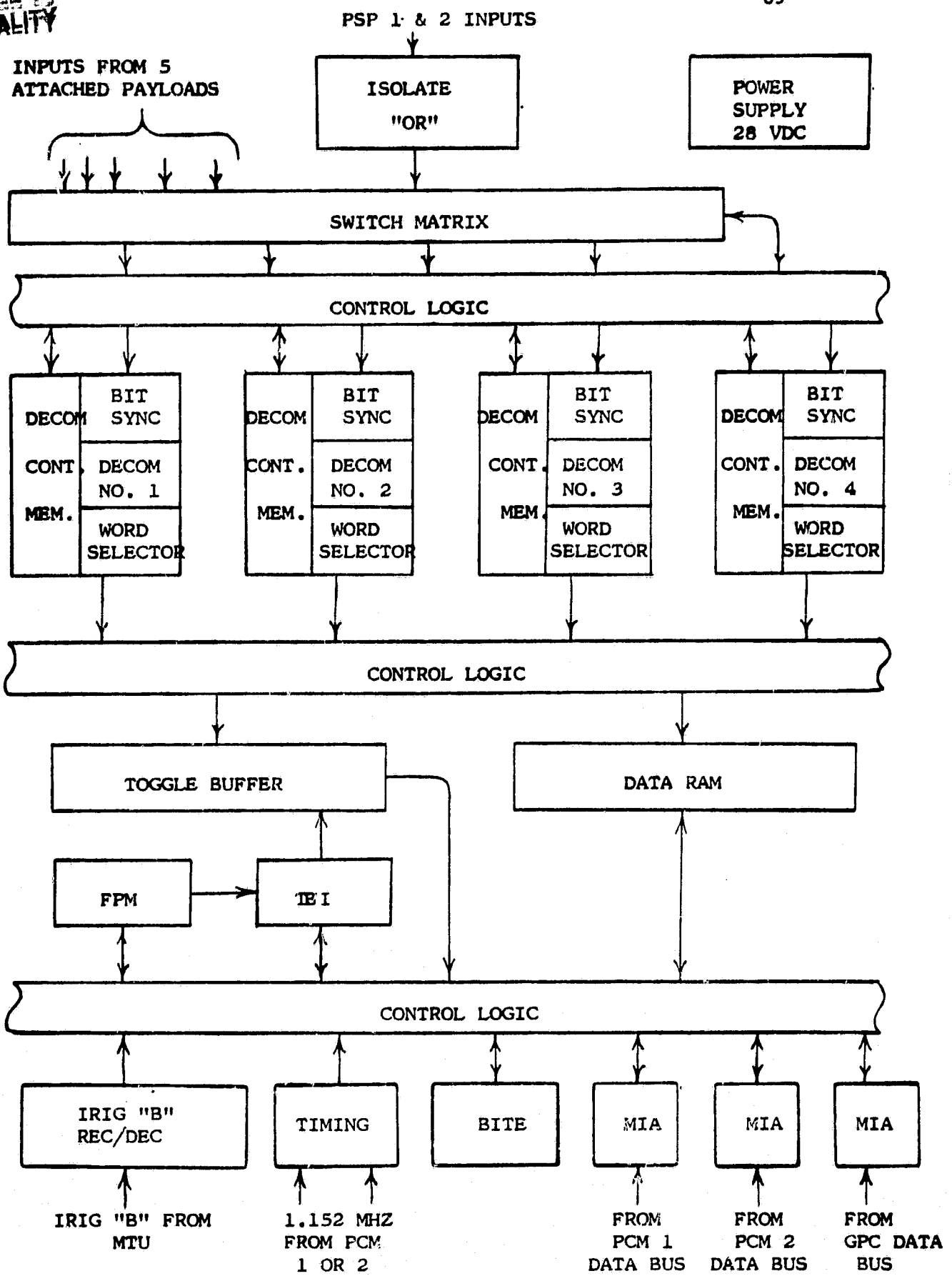


Figure C.3. PDI Block Diagram

each frame. A homogeneous data set for this mode of operation is defined as all information within this PDI-created frame. Code type, bit rate, frame length, and synchronization pattern are programmable under control of the decommutator format memories. The frames are supplied to the toggle buffer for storage in homogeneous data sets. No data is supplied to the data RAM in this mode of operation.

A status register containing the status and time for a given chain operation is provided by the word selector to the toggle buffer control logic. This logic regulates access to and from the half buffers by the word selectors and the data buses. All requests for TB data by the data bus ports are processed through the fetch pointer memory (FPM) and the toggle buffer identifier (TBI). The TB control logic also partitions data from the word selector into homogeneous data sets for access by the data bus ports.

The FPM is used to identify which TB is to be accessed by a data bus port. It also allows access to any location in the data RAM by any of the PDI data bus ports at any time. FPM control logic routes all requests for TB data to the location in the FPM identified by the data bus command word. It further provides for loading and reading of formats to and from the FPM at any time by the data bus ports.

A data RAM for the storage of data from the word selector by parameter is provided. The data RAM control logic steers data provided by the word selector into addresses in the data RAM as specified by the decommutator control memory.

There are three data bus ports for interface with the Orbiter's GPC that have read and write access into the switch matrix, the decommutator control memory, the FPM, the PDI, and the data RAM.

An IRIG "B" receiver/decoder accepts an IRIG "B" code from an external source, decodes time, and supplies it to the four status registers.

C4 KU-BAND SIGNAL PROCESSOR (SPA)

The signal processor (shown in Figure C.4) performs the functions of data and signal processing for the Ku-band forward and return links. For the forward link, two modes are available, as follows:

(1) A special mode from amplification and impedance matching of data from the Ku-band receiver and communication processor assemblies for delivery to the NSP.

(2) A nominal mode which performs the operations of bit synchronization, clock generation, ambiguity resolution (data and clock), bit detection, frame synchronization and data decommutation of Ku-band received data.

Return-link signals are handled in the KuSP by modulating the data in one of two modes before upconversion to Ku-band frequencies. The two selectable modes multiplex three channels carrying a wide variety of data. In mode 1, the PM mode, the high-rate data channel is convolutionally encoded before modulation onto the carrier. The lower rate data channels 1 and 2 are QPSK modulated onto a square-wave subcarrier which is, in turn, PSK modulated in quadrature with channel 3 onto the carrier.

In mode 2, the FM mode, the two lower rate channels are QPSK modulated onto a square-wave subcarrier, as in mode 1. The resulting signal is summed with the third wideband channel, and the composite signal is then frequency modulated (FM) onto the carrier.

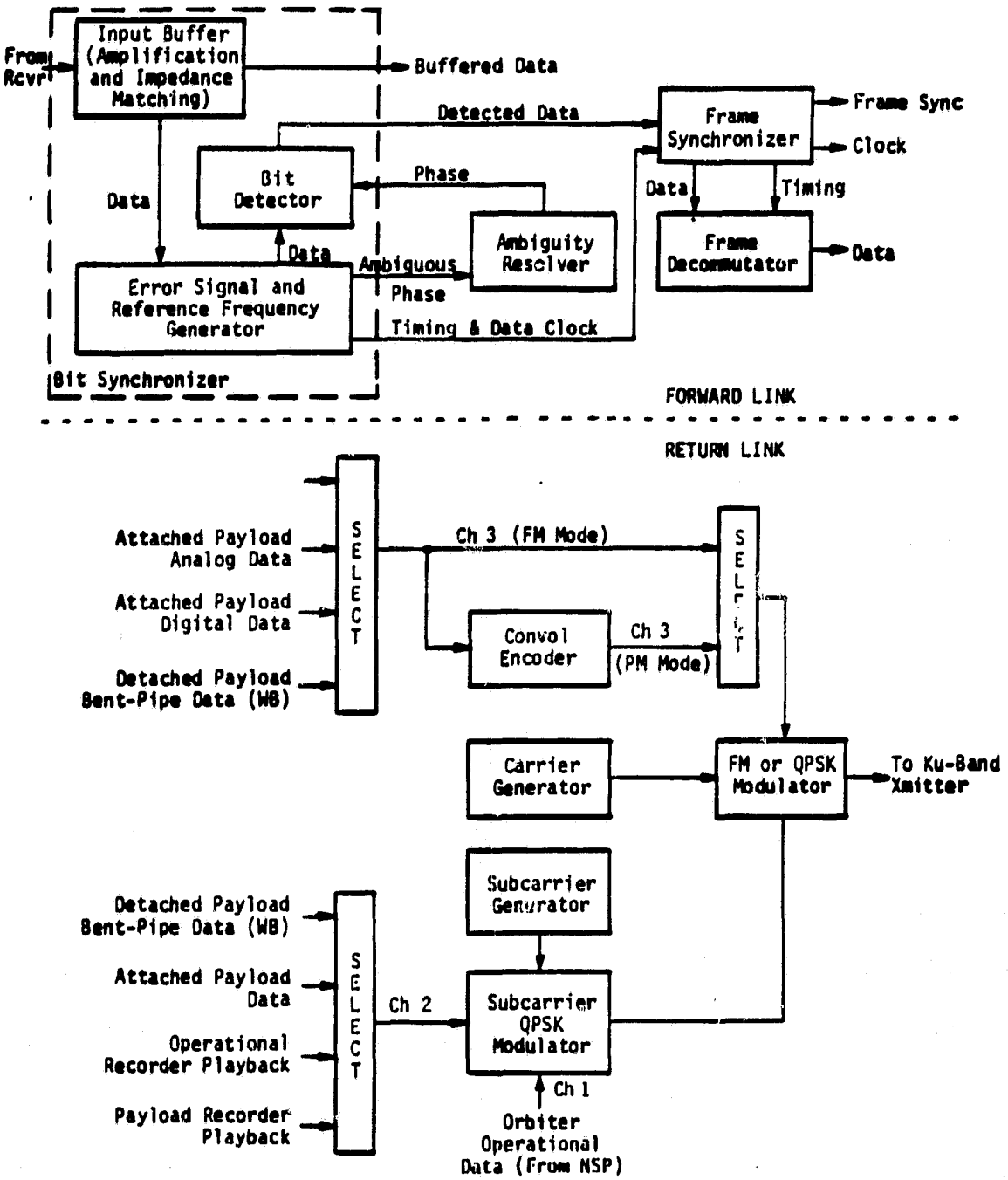


Figure C.4. Ku-Band Signal Processor Block Diagram

C5 FM SIGNAL PROCESSOR AND TRANSMITTER

Figure C.5 shows the diagram of the FM signal processor and transmitter. The functions of baseband modulation, mixing, routing, impedance matching and operational switching are accomplished by the signal processor. Payload signals, whether wideband analog, high-rate digital or low-rate digital, are buffered in a matching network and passed through the mode selector and wideband amplifier to the FM transmitter.

The FM transmitter provides the functions of carrier frequency modulation and RF power amplification. No preemphasis or other special processing is employed.

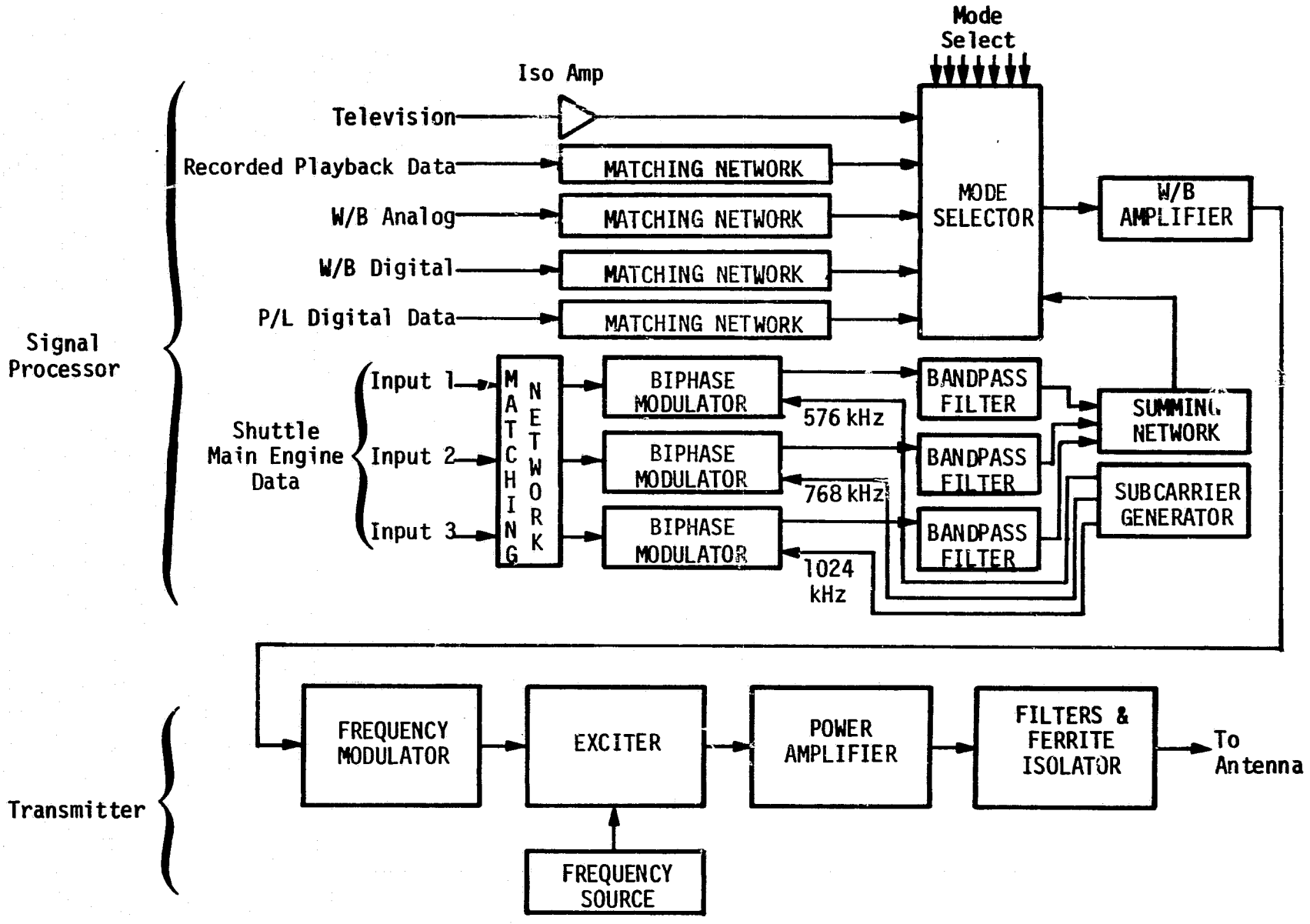


Figure C.5. FM Signal Processor and Transmitter

C6 NETWORK SIGNAL PROCESSOR

The block diagram of the Network Signal Processor (NSP) is shown in Figure C.6. The NSP consists of individual forward link, return link, and record mode processing circuits. The three processes operate concurrently, thus providing, in addition to the record mode processing, full duplex operation of the forward and return links.

Mode controls define the particular data rates, the nature of the data, the need for convolutional encoding and decoding, and the need for voice delta modulation or demodulation. Interface controls define the input data source and the PCM telemetry source.

All input data is introduced through the bit synchronizer, with four input controls identifying the data source, one input control identifying the data rate, and another input control identifying the hard or soft decision. When bit synchronization is achieved, a status bit is provided to the multiplexer/demultiplexer (MDM) via the summary built-in test equipment (BITE) circuit.

The bit synchronization data output and the derived clock are delivered to the convolutional decoder and to the selector A via a data invert control logic. Selector A is where the mode control determines if the convolutional decoder is to be employed. In the coded mode, the convolutional decoder provides its own data-inversion capability. At selector B, if the data is identified as DOD data by the mode control, it is output to the COMSEC unit and clocked back into the NSP after decryption.

Following detection (and decoding), the data is presented to the frame synchronization logic for frame pattern recognition. Once frame synchronization lock has been achieved, a lock signal is sent to the MDM via the summary BITE circuit. Finally, the forward link function of demultiplexing and rate buffering is performed.

Command data is checked for errors in the BCH decoder, modified appropriately and stored in a buffer. A message-valid pulse is sent to the MDM via the summary BITE circuit for every command word that passes the BCH check and vehicle address check. After 10 commands have been received, a signal is sent to the MDM indicating a data-present status. Upon request, 32 16-bit words are sent to an associated subsystem. The first word contains the BITE status of the NSP, the second through the

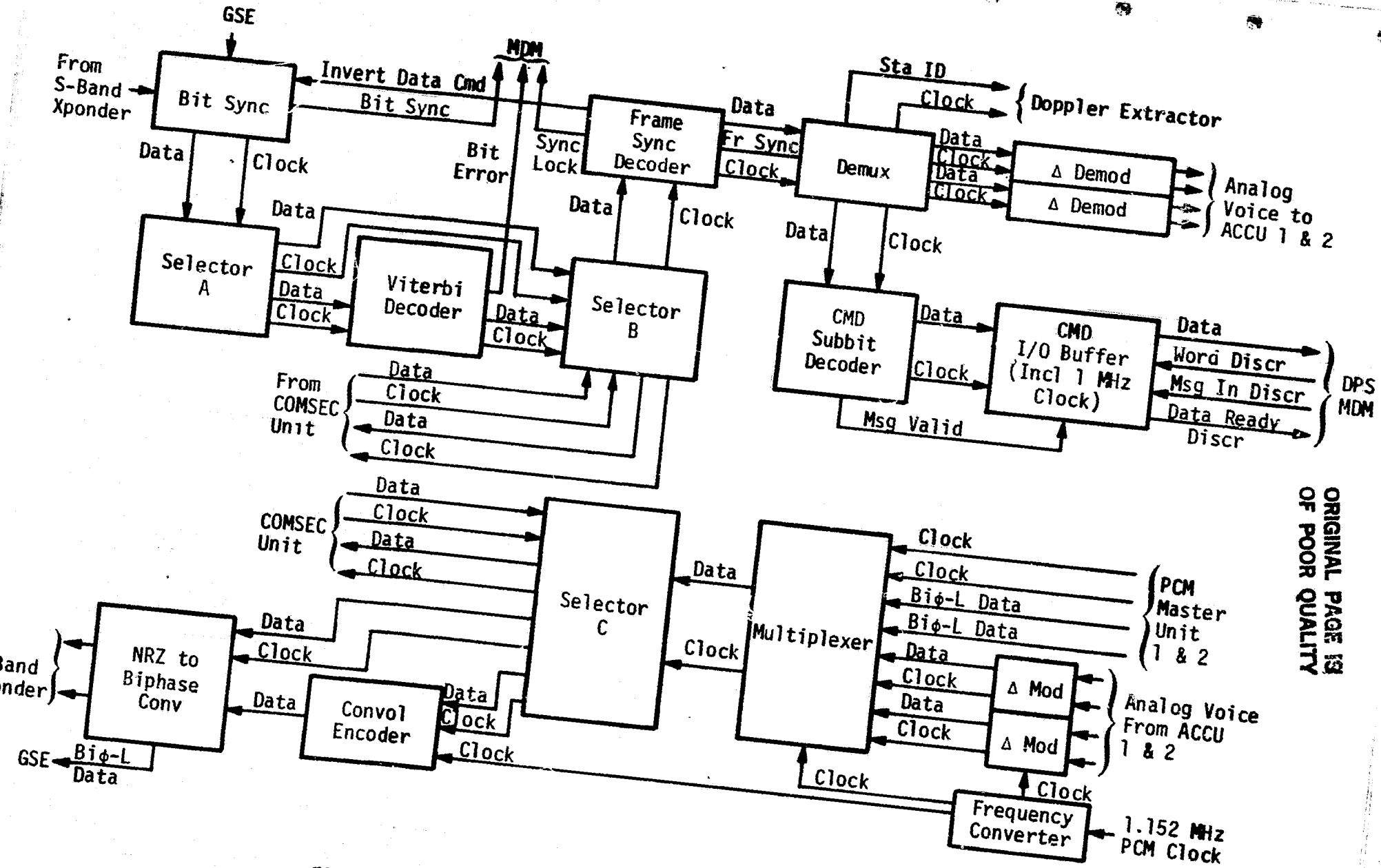


Figure C.6. Network Signal Processor Functional Diagram

thirty-first words contain commands, and the thirty-second word contains a bit for each command transmitted, representing that command's validity.

The return link consists of multiplexing telemetry and voice data. The multiplexing function is keyed to the frame synchronization pattern included with the telemetry data. For DOD data, once the multiplexing function has been performance, the data is routed to the COMSEC equipment for encryption. All data (NASA or DOD) may also be convolutionally encoded, as desired. Finally, the coded or uncoded data is NRZ-to-Manchester converted prior to transmission. The return link is provided simultaneously to the S-band and Ku-band network.

The record mode multiplexes the voice data only with the selected 128-kbps PCM data. In NASA submode 1, the 128-kbps telemetry is multiplexed with the two dedicated voice channels. In NASA submode 2, the 128-kbps telemetry is simply routed to the drivers for transmission to the recorders. In the DOD mode, the recorder data is taken from the return-link COMSEC encrypter (effectively bypassing the entire record mode processing logic).

C7 S-BAND NETWORK TRANSPONDER

A functional network transponder block diagram is shown in Figure C.7. The received signal, processed through the preamplifier in the TDRS mode or through the transponder triplexer receive filter (high or low) in the SGLS or STDN direct link modes, is amplified by a low-noise S-band input amplifier prior to downconversion to approximately 240 MHz. A second coherent downconversion brings the signal to 31 MHz where, in the TDRS mode, despreading is accomplished by the spread spectrum processor which uses a noncoherent code search loop. The TDRS despread signal is routed to the carrier Costas loop used to derive phase tracking information. In the SGLS and STDN modes, the Costas loop configuration is also used to track the residual carrier. Demodulation of command and ranging signals is accomplished using an offline wideband phase detector so that the Costas loop detector predetection bandwidth is optimized for tracking performance. Both tone ranging and data outputs from the receiver are noncoherently AGC'd to maintain a constant RMS signal plus noise level to the associated subsystems.

All frequencies are derived from two switchable voltage-controlled crystal oscillator (VCXO) subassemblies and one reference crystal oscillator. The reference oscillator operates at 31 MHz and thus places the second IF at 31 MHz. This is sufficiently high in frequency to provide good first IF image rejection and still allow the use of narrowband second IF filters. Channel selection is provided by changing the VCXO frequency. Each VCXO subassembly contains four VCXOs for two-channel operation in either the SGLS or STDN/TDRS modes.

A simple unique multiplier configuration is used employing phase-locked oscillators to accomplish the X25 (second LO), X14 or X15 (first LO), and X15 or X16 (transmitter drive) multiplication. By simply changing the divider feedback ratios, the multiplication factor can be changed. This technique provides the wide percentage bandwidth multiplication required for multimode operation while yielding very low spurious products. The final first LO multiplication ratio (X6 or X7) is selected as a function of the mode.

The third mixer in the second LO chain offsets the second LO frequency using a 62-MHz reference signal, so that the second IF is fixed and does not vary as a function of received frequency. Therefore, the

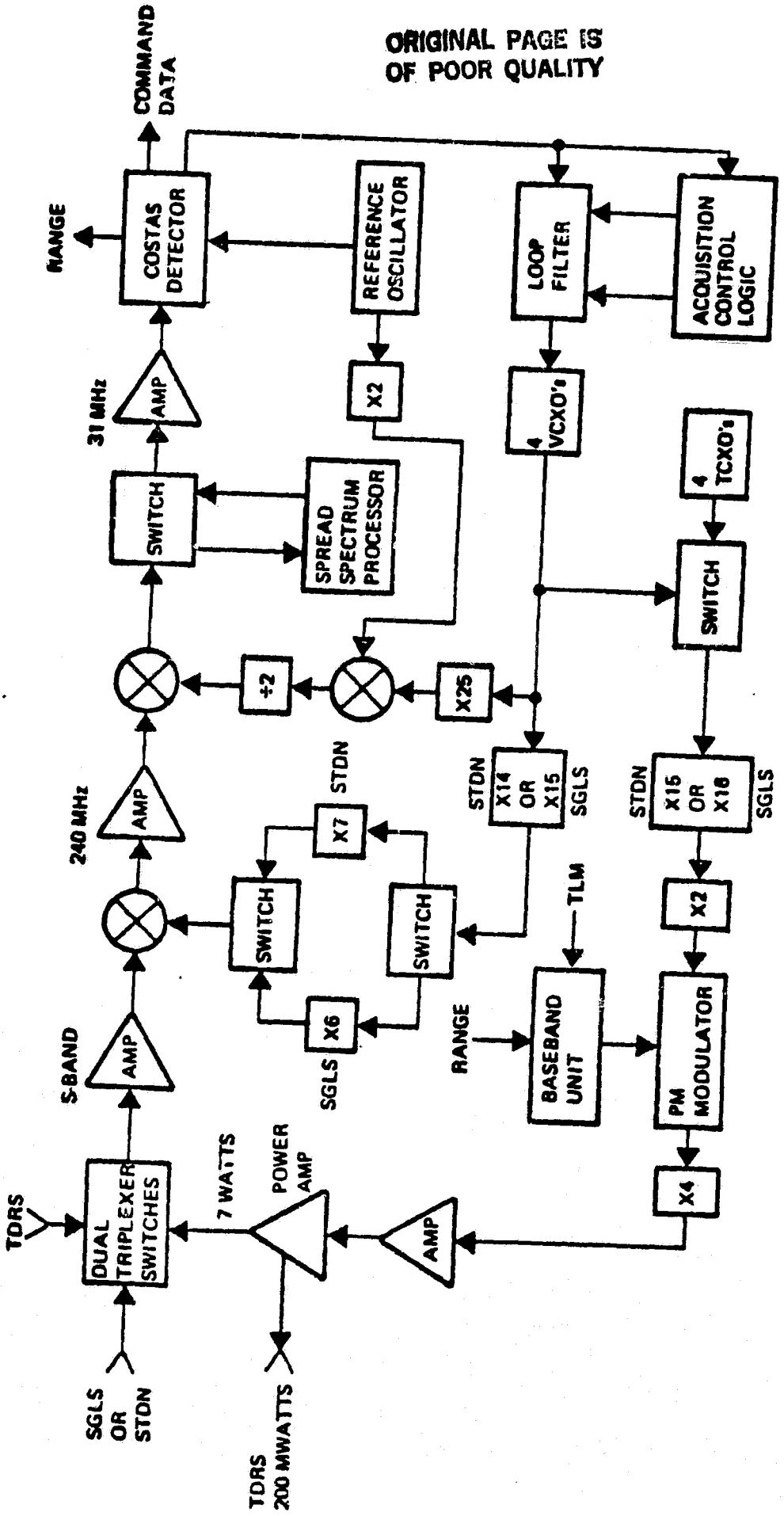


Figure C.7. S-Band Network Transponder Block Diagram

spread spectrum processor and the Costas loop preselection filters can be operated at very low SNR at the same center frequency regardless of input channel selection. The drive frequencies to the third mixer are at twice the first IF and twice the reference oscillator frequency. This eliminates the potential problem of generating a high-level signal at the third mixer exactly equal to the first IF frequency, which could result in a self-lock condition.

Downlink STDN or SGLS linear modulation is accomplished at about 560 MHz and then multiplied by 4 up to S-band. An S-band solid-state power amplifier provides a low-level (TDRS) or high-level (STDN/TDRS) output, depending on mode selection.

C8 KU-BAND RECEIVING AND TRANSMITTING SUBSYSTEM

The Ku-band receiving and transmitting subsystem frequency translates and amplifies the Orbiter Ku-band forward and return link signals. A functional diagram of this subsystem is given in Figure C.8.

The received signal from the TDRS is split into two channels: (1) the sum channel (Σ) which carries the received signal, and (2) the difference channel (Δ) which carries the angle information necessary to point the receiving antenna.

These two channels are received in the forward link, amplified in low-noise FET amplifiers, and downconverted before being delivered to the communication subunit. This subunit has three major functions: (1) IF processing, (2) spectrum despreading, and (3) acquisition and data demodulation.

The sum channel is frequency translated and bandpass filtered to drive the PSK demodulator and the PN despreader. The PSK demodulator consists of a Costas tracking loop and a lock detector, and provides the demodulated communication data and in-lock indication.

When the received signal is spread spectrum, it is despread in the PN despreader which contains a tau-dither acquisition and tracking circuit and a PN lock detector.

After the first downconversion, the sum and difference channels are combined into a sum-plus-difference channel; this signal is filtered, gain controlled, and envelope detected to provide the communication track information.

The transmitting system takes either the QPSK or FM modulator output from the SPA and processes it for transmission. The SPA signal is mixed with a signal generated by a Gunn VCO and the mixer output is filtered and amplified by a four-stage GaAs FET limiting amplifier. The low AM/PM conversion of this amplifier permits a significant relaxation of the TWT AM/PM conversion specifications. The TWT in the transmitter amplifies the upconverted signal and sends it to the antenna through a circulator.

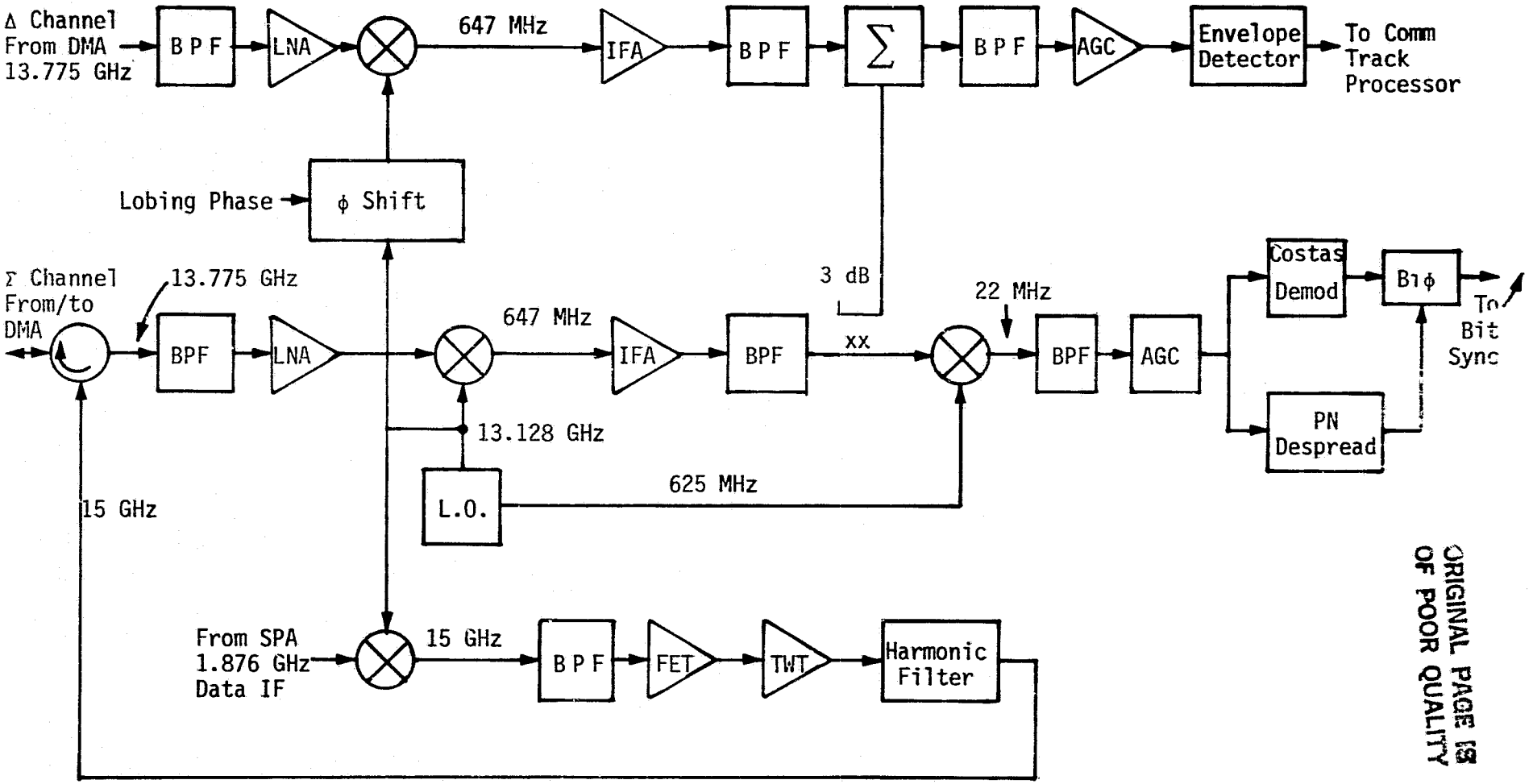


Figure C.8. Ku-Band Receiving/Transmitting Subsystem

APPENDIX D

USER'S GUIDELINE FOR NONSTANDARD MODULATION FORMATS INPUT TO THE SHUTTLE PAYLOAD INTERROGATOR RECEIVER

APPENDIX D

USER'S GUIDELINE FOR NONSTANDARD MODULATION FORMATS INPUT TO THE SHUTTLE PAYLOAD INTERROGATOR RECEIVER

D1 PAYLOAD INTERROGATOR RECEIVER CAPABILITIES AND PARAMETERS

1.1 Type of Receiver

The Payload Interrogator (PI) utilizes a discrete carrier phase-locked tracking loop and a single quadrature-phase (relative to the phase of the discrete carrier component) demodulator for modulation recovery.

1.2 Carrier-Tracking Loop Bandwidth

The minimum tracking-loop bandwidth of 1216 Hz (two-sided) is designed to occur at a discrete carrier signal level (in the high sensitivity mode) of -124 dBm. As the received signal level increases to its maximum-allowable value, the tracking-loop bandwidth increases to 1364 Hz (carrier loop in-lock and AGC functioning). When the carrier loop is out of lock and prior to proper AGC regulation of receiver gain for maximum signal levels, the tracking-loop bandwidth is on the order of 3800 Hz. (This maximum-signal-level bandwidth is a function of the IF amplifier amplitude saturation characteristics.)

1.3 Post-Demodulation Lowpass Bandwidth

The post-demodulation 3-dB lowpass bandwidth is 4.5 MHz.

1.4 Demodulation Phase Detector Characteristics

A product-type (double-balanced mixer) phase detector is employed for modulation recovery. (The amplitude-versus-phase characteristic is sinusoidal in form.

1.5 Carrier-Swept Frequency Acquisition

The PI receiver tracking-loop VCO is linearly frequency swept at a rate of 10 kHz/s, so that the receiver searches ±75 kHz about its nominal frequency in order to obtain carrier lock. Modulation sidebands within this frequency range must be such that receiver false-lock to sidebands is precluded.

D2 GENERAL PAYLOAD TRANSMITTER MODULATION CRITERIA

2.1 Allowable Modulations

Phase modulation (PM) of the carrier is the only allowable type of modulation. Frequency modulation (FM) and amplitude modulation (AM) of the carrier are not permitted. Quadriphase modulation is also not allowed*.

2.2 Maximum Carrier Suppression

The maximum allowable carrier suppression due to the composite of all phase-modulating sources shall not exceed 10 dB.

2.3 Subcarrier Modulation

When subcarriers are employed, they may be either phase or frequency modulated. Amplitude-modulated subcarriers are not permitted. Restrictions on the use of subcarriers are given under sections 3.2 and 3.3.

2.4 Direct-Carrier Modulation by Baseband Signals

Direct-carrier modulation by analog-type baseband signals is not allowed. Direct-carrier modulation by digital-type baseband signals is allowed, subject to the restrictions given under section 3.4

D3 SPECIFIC NONSTANDARD MODULATION RESTRICTIONS

3.1 Discrete Frequency Component Sideband Levels

Carrier phase modulation by periodic signals (sinusoids, square waves, etc.) is not permitted. No incidental and/or spurious discrete frequency component sideband levels shall be greater than -48 dBc in a frequency range of ±100 kHz about the carrier frequency.

A possible exception to the above restriction may be allowed for narrow-pulse, low-duty-cycle modulations, where the maximum sideband level of the largest discrete frequency component shall not be greater than -48 dB below the unmodulated carrier level. The allowance of such modulation shall be subject to the review and approval of the cognizant NASA authority.

---

*Although certain forms of quadriphase with a discrete carrier could be acquired by the PI receiver, no provision is made in the receiver for demodulating and outputting the in-phase component.

3.2 Frequency-Modulated Subcarriers

3.2.1 Analog Modulations

No analog signal frequency-modulated subcarrier, in a frequency range of ±100 kHz about the carrier frequency, shall be allowed to phase modulate the carrier if the inequality

$$f_m \Delta f > 1.1 \times 10^4 \tag{1}$$

is violated, where f_m is the bandwidth or maximum frequency of the baseband analog signal, and Δf is the peak frequency deviation of the subcarrier. Provided that the inequality (1) is satisfied, the maximum-allowable carrier phase modulation index, β, by the frequency-modulated sinusoidal subcarrier is given by

$$J_1(\beta)/J_0(\beta) \leq 4.9 \times 10^{-7} \; f_m \Delta f. \tag{2}$$

The value of β may be determined with the aid of Figure D.1.

3.2.2 Digital Modulations

No frequency-shift-keyed (FSK) subcarrier, in a frequency range of ±100 kHz about the carrier frequency, shall be allowed to phase modulate the carrier if the inequality

$$R_b > 6.3 \times 10^2 \tag{3}$$

is violated, where R_b is the data bit rate (bps). Provided that the inequality (3) is satisfied, the maximum-allowable carrier phase modulation index, β, by the FSK-modulated sinusoidal subcarrier is given by

$$J_1(\beta)/J_0(\beta) \leq 1.4 \times 10^{-8} \; R_b^2. \tag{4}$$

The value of β may be determined with the aid of Figure D.1.

3.3 Phase-Modulated Subcarriers

3.3.1 Analog Modulations

Phase modulations of subcarriers by analog baseband signals is not recommended due to inefficiency. As a result, no such modulations are expected, and no guidelines have been developed.

3.3.2 Digital Modulations

No phase-shift-keyed (PSK) subcarrier, in a frequency range of ±100 kHz about the carrier frequency, shall be allowed to phase modulate the carrier if the inequality

$$R_b > 6.3 \times 10^2 \tag{5}$$

is violated, where R_b is the data bit rate in bits per second. Provided that the inequality (5) is satisfied, the maximum-allowable carrier phase modulation index, β, by the PSK-modulated sinusoidal subcarrier is given by

$$J_1(\beta)/J_0(\beta) \leq 1 \times 10^{-8} R_b^2 . \tag{6}$$

The value of β may be determined with the aid of Figure D.1.

3.4 Direct-Carrier Modulations

3.4.1 Analog Modulations

Direct phase modulation of the carrier by an analog baseband signal is not recommended due to inefficiency. As no such modulations are expected, no guidelines have been developed.

3.4.2 Digital Modulations

The criterion for the minimum-allowable bit rate is based upon a carrier-tracking loop RMS phase noise component, due to modulation sidebands tracking, of 10° or less. The allowable NRZ bit rate must therefore satisfy the following inequality

$$R_b > 2.24 \times 10^4 \tan^2 (\beta) , \tag{7}$$

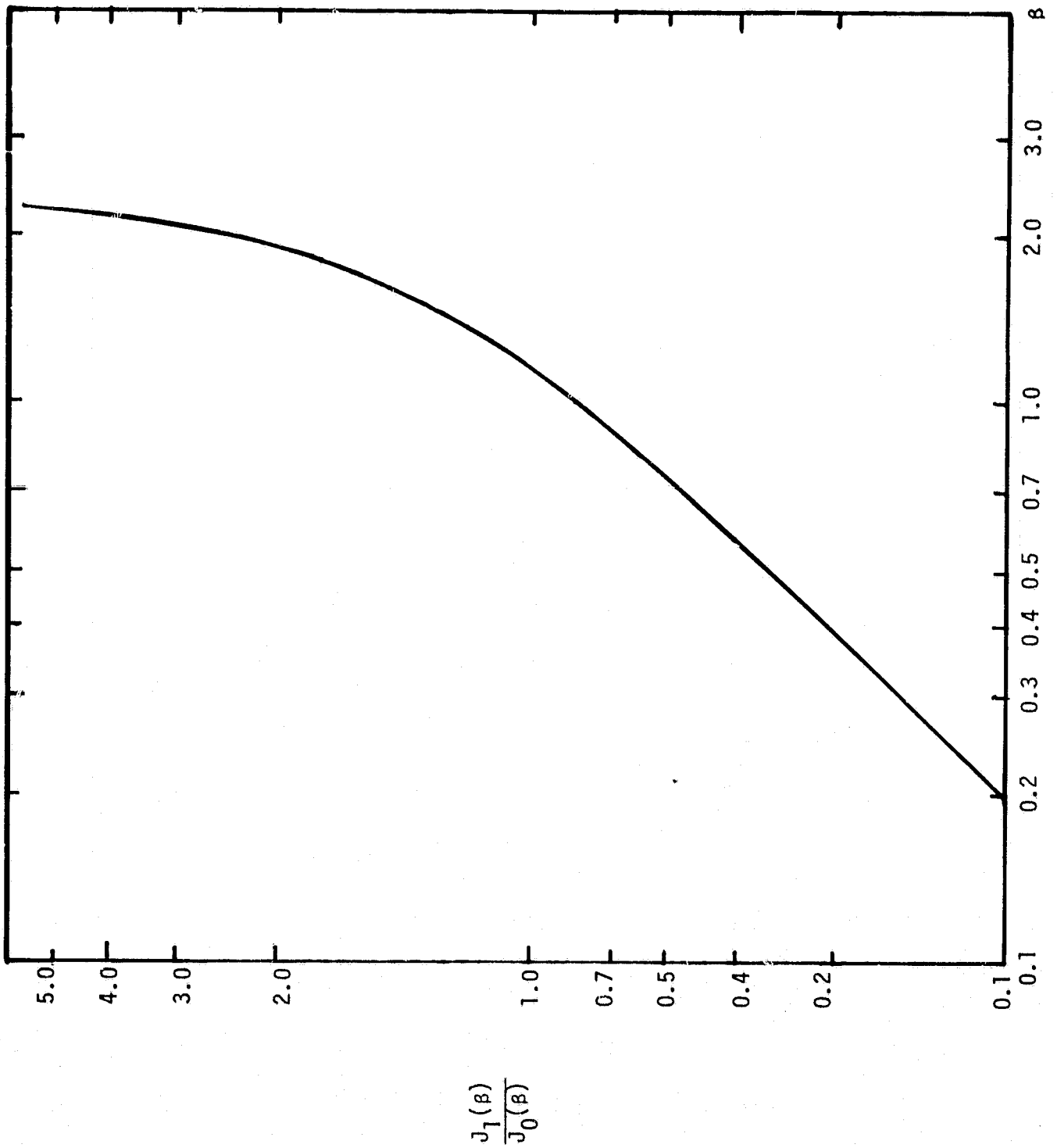


Figure D.1

where the numerical coefficient is based upon the carrier-tracking loop maximum in-lock bandwidth* and β is the carrier phase deviation ($\beta \leq 71.5°$).

In order to keep carrier loop phase slewing to less than 15° during a string of transitionless data bits, the maximum number of such bits shall be

$$\text{Maximum number of bits without transition} = 9.3 \times 10^{-5} R_b . \qquad (8)$$

This transitionless period must be followed by a reasonable number of transitions in such a pattern that the slewing error is negated within a period of bits equal to five times the transitionless period.

To avoid the problem of bit slewing, Manchestering of the bits is recommended. Given Manchestered bits, condition (8) is no longer applicable, and the minimum bit rate allowed is the maximum of that calculated from (9) and (10):

$$R_b \geq 300 \tan^2 (\beta) \qquad (9)$$

$$R_b \geq 7.1 \quad 10^3 \sqrt{\tan (\beta)} \qquad (10)$$

*See section 1.2.