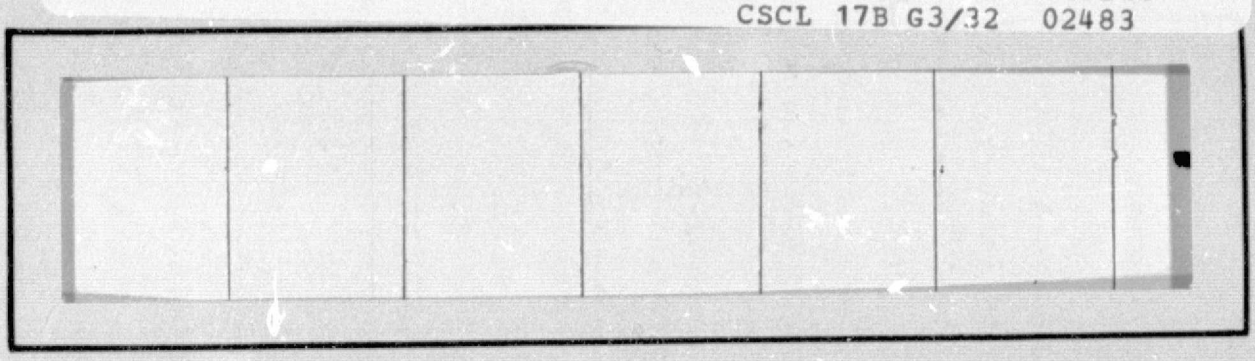


Axiomatix

Marina del Rey • California

INTEGRATED SOURCE AND CHANNEL ENCODED DIGITAL COMMUNICATION SYSTEM DESIGN STUDY

FINAL REPORT

Contract No. NAS 9-13467
Exhibit D

Prepared for

NASA Lyndon B. Johnson Space Center
Houston, Texas 77058

Prepared by

Gaylord K. Huth

Axiomatix
13900 Panay Way, Suite 110M
Marina del Rey, California 90291

Axiomatix Report No. R7607-3
July 31, 1976

TABLE OF CONTENTS

Appendix

1.0 INTRODUCTION

This report summarizes the results of several study tasks pertaining to various aspects of the Space Shuttle communication system. These tasks performed can be logically divided into the following six categories:

1. Phase multiplexing for two- and three-channel data transmission.
2. Effects of phase noise on the performance of coherent communication links.
3. Analysis of command system performance.
4. Error correcting code tradeoffs.
5. Signal detection and angular search procedure for the Shuttle Ku-band communication system.
6. False lock performance of Costas loop receivers.

A summary of the phase multiplexing techniques for the Ku-band return link is presented in Section 2.0 with detailed analyses in Appendices A, B, and C. The effects of phase noise on communication system performance is summarized in Section 3.0 with details in Appendix D. Analysis of the Range Safety command system is analyzed in Section 4.0. Techniques for providing command authentication for Air Force missions are summarized in Section 4.0 with details in Appendix E. Error correcting coding tradeoffs for the 50 Mbps Ku-band return link are discussed in Section 5.0. Detailed analysis for predicting performance of the Shuttle convolutional coded communication links is presented in Appendix F, with the tradeoffs that were considered in the final recommendation of the 50 Mbps coded system identified in Appendix G. The signal detection analysis and angular search procedure required to point the Orbiter Ku-band antenna at the TDRS are discussed in Section 6.0 with details in Appendices H and I. Finally, the potential false carrier lock conditions for the S-band and Ku-band link are summarized in Section 7.0 with the detailed analysis in Appendix J.

2.0 PHASE MULTIPLEXING FOR TWO- AND THREE-CHANNEL DATA TRANSMISSION

Phase multiplexing offers a technique of accommodating multiple data channels with independent data rates and clocks as required on the Ku-band return link. Alternatives to phase multiplexing are time-division multiplexing and frequency-division multiplexing. Time division multiplexing at the 100 Mbps data rate encountered on the Ku-band return link is extremely complex because of the requirement of reclocking the data. Frequency-division multiplexing requires large bandwidths to avoid intermodulation products. The available bandwidth through the TDRS for the Ku-band return link is 225 MHz, which is just large enough to pass the 100 Mbps data rate, let alone frequency multiplexing additional data rates into this bandwidth.

Appendices A, B, and C investigate the phase multiplexing modulation and demodulation. Appendix A investigates the general phase multiplexing techniques including conventional phase modulation with subcarriers, interplex, and hybrid interplex-quadrature modulation. Phase multiplexing using quadrature carriers is shown to have potential implementation advantages over both a conventional PM subcarrier approach and the interplex approach which has been utilized for deep space communications. To illustrate this potential implementation advantage, the particular cases of two- and three-channel phase multiplexing using various techniques are compared in terms of implementation and performance. It is demonstrated in Appendix A that either the interplex or the hybrid phase multiplexing technique for transmission of three independent and asynchronous digital data channels offers a significant performance advantage over a conventional PM subcarrier design. The hybrid phase multiplexing technique differs from interplex primarily in terms of implementation. Whereas interplex involves data channel combining at baseband and utilizes a linear phase modulator at IF, the hybrid technique involves data channel combining at IF and utilizes quadrature carrier modulation. Since the performance characteristics of these two

approaches are essentially identical, it appears that the selection of a technique for any particular application should be dictated only by equipment design considerations.

Appendix B investigates candidate receivers for a two-channel phase multiplexed Ku-band communication signal. It is shown that the fourth-power loop makes very ineffective use of the total available tracking power when the power division is near 4:1. An additional disadvantage is that the stable lock points occur at multiples of $\pi/4$, thereby allowing the possibility that the two data channels become interchanged. This could be detected and corrected in the system design, but only at the expense of additional complexity.

The Costas loop which tracks on the high data rate signal of the unbalanced QPSK waveform is shown to perform satisfactorily. Appropriate error rate computations show that the Costas loop considered performs within a few tenths of a dB of the ideal receiver. Also, it is shown that the Costas loop has stable lock points at multiples of π which guarantee that the two data channels cannot be sent to the incorrect data demodulator. Although it is usually expected that the power in both channels will always be present even if the data rates are reduced to zero, should the power in the high data rate channel fail, then the sign of the tracking signal is reversed and the stable lock angles are at $\pi/2 \pm k\pi$, where k is $1, 2, 3, \ldots$. If this occurs, then the low data rate link will be processed into the high data rate demodulator. If the power in the low data rate link goes to zero, then the stable lock angles remain unchanged, so that the high data rate link continues to be processed into the high data rate demodulator.

Appendix C presents the analysis of the recommended modulator and demodulator for the Ku-band return link. The exact analytical expression for the hard-limited three-channel hybrid phase multiplexed signal is derived From this analytical representation of the modulation, the performance of the Costas loop demodulator is analyzed. Analysis is presented that allows the optimization of arm filters of the Costas loop in terms of the high data rate and filter type. This optimization approach is new, since the usual analysis of the Costas loop considers that the arm filters are sufficiently

wide that the data modulation is undistorted. It is shown that the optimum choice of the arm filters is narrow enough to actually distort the signal but the loop signal-to-noise ratio is maximized by limiting the noise that is passed through the filters.

The result of carrier tracking phase error in the receiver introduces crosstalk from one data channel into another. In particular, the low data rate channels introduce crosstalk into the high rate channel in proportion to their relative received powers and the carrier tracking loop phase error. While Appendix A presents the crosstalk power as a function of fixed carrier tracking loop phase error, Appendix C presents the average crosstalk as evaluated by considering the loop phase error probability density function. The average crosstalk is important because it indicates the actual system performance degradation that can be expected.

An area for further study of the phase multiplexing technique is numerical integration of the integrals in Appendix C that allow optimization of Costas loop tracking and performance degradation due to channel crosstalk. In addition, a similar analysis to Appendix C needs to be performed on the subcarrier tracking loop. Finally, an analysis of the false lock characteristics of the Costas loop with three-channel modulation needs to be performed similar to the analysis for bi-phase data in Appendix J.

3.0 EFFECTS OF PHASE NOISE ON PERFORMANCE OF COHERENT COMMUNICATION LINKS

Transmitter phase noise limits the minimum probability of error achieved by a coherent phase-shift-keyed (PSK) communication system. Appendix D systematically considers the effects of transmitter phase noise on the performance of two-way coherent communication links and indicates design approaches that may be used to reduce or minimize such effects. As a specific example of phase noise, the frequency synthesizer necessary for the S-band payload interrogator in the DSN mode is considered. For this transmitter phase noise spectrum, the resulting degradation in bit error rate performance for baseband and PSK subcarrier channels is determined. For the forward portion of the two-way link, the transmitted phase noise is due primarily to jitter introduced by the oscillators and phase detectors used in the frequency synthesizer to generate the carrier frequency. For the return link, the transmitted phase noise is due primarily to the turn-around of the phase noise which is not tracked out by the forward link receiver carrier tracking loop.

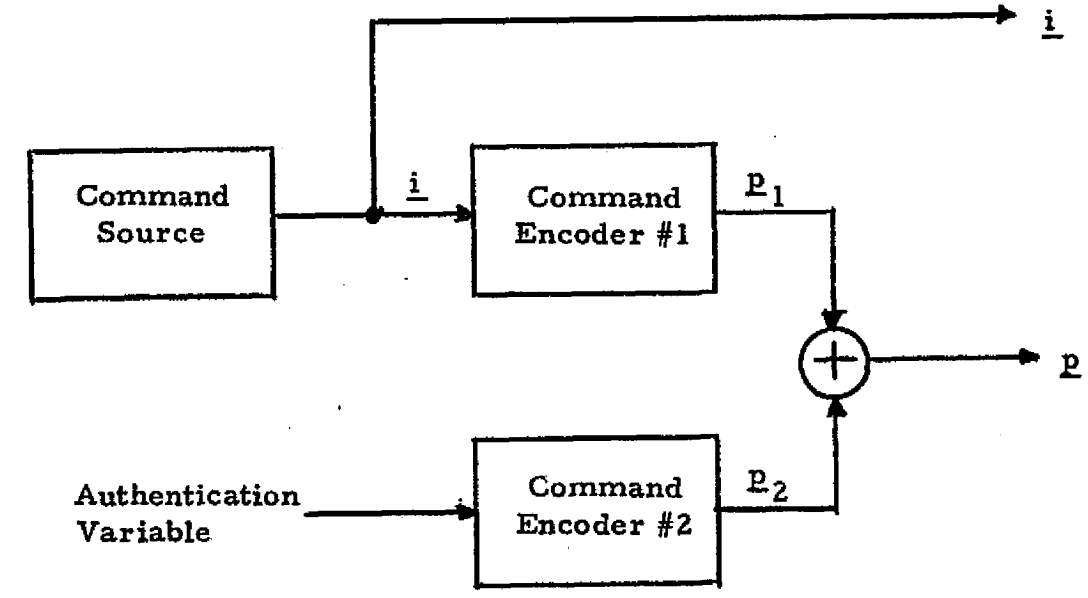
The analysis of the two-link with transmitter phase noise presented in Appendix D shows that there is a trade-off between the effect of phase noise on the forward link and the phase noise turned around for the return link. As a final result, it is found that at bit error probability of 10^{-4}, the performance degradation of the S-band payload interrogator in the DSN mode communicating with the DSN standard transponder is less than 0.1 dB and 0.2 dB for the forward link and the total two-way link, respectively.

4.0 COMMAND SYSTEM CONSIDERATIONS

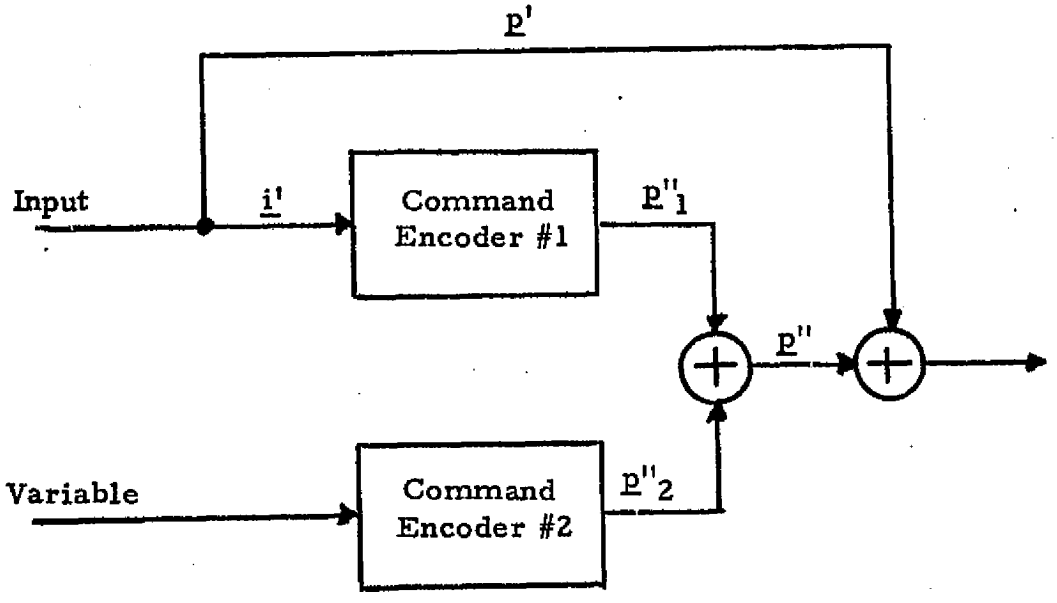
For Air Force missions, modifications are proposed in Appendix E to the Shuttle command system to provide command authentication. Three possible modifications are proposed. In each case, an authenticator variable up to 50 bits long is used to modify the parity portion of the command codeword. Of the three modifications proposed, the simplest of modulo-2 adding the authentication variable to the parity is considered too easy to jam to be of much utility. The other two modifications were considered to be hard to jam. The simplest of these two techniques is presented in Figure 1, where the information word and the variable are encoded by completely different (127, 50) BCH encoders. The two parities produced are modulo-2 added to form $\underline{p}$. The codeword $\underline{x} = (\underline{i}, \underline{p})$ is transmitted. The same process is performed onthe received word in the error detector; the resultant parity $\underline{p}''$ is compared with the received parity $\underline{p}'$. The remainder of Appendix E analyzes the susceptibility of jamming of this recommended modification and is found relatively impossible to jam.

An additional analysis that was performed on command systems was the performance analysis of the Range Safety Command System. A typical message format for the command system is shown in Figure 2. The subcarrier frequencies (tones) are from 7.35 kHz to 13.65 kHz. Each character is made up of two tones transmitted simultaneously. The secure address is composed of 9 characters and 2 characters comprising the functional selection.

To analyze the possibility of jamming the command system, an assumption must be made about the demodulator. It is assumed that, when the secure address is locked into the command system, thresholds are set such that, at any given time if any tones are received that are greater than the thresholds besides the expected two tones making up the proper character, then the command will be rejected. If these thresholds are not set, then a jammer could cause a command to be incorrectly accepted by placing jamming tones at each of the seven possible subcarriers. Using these



(a) Encoder



(b) Detector

Figure 1. Command Authentication Technique

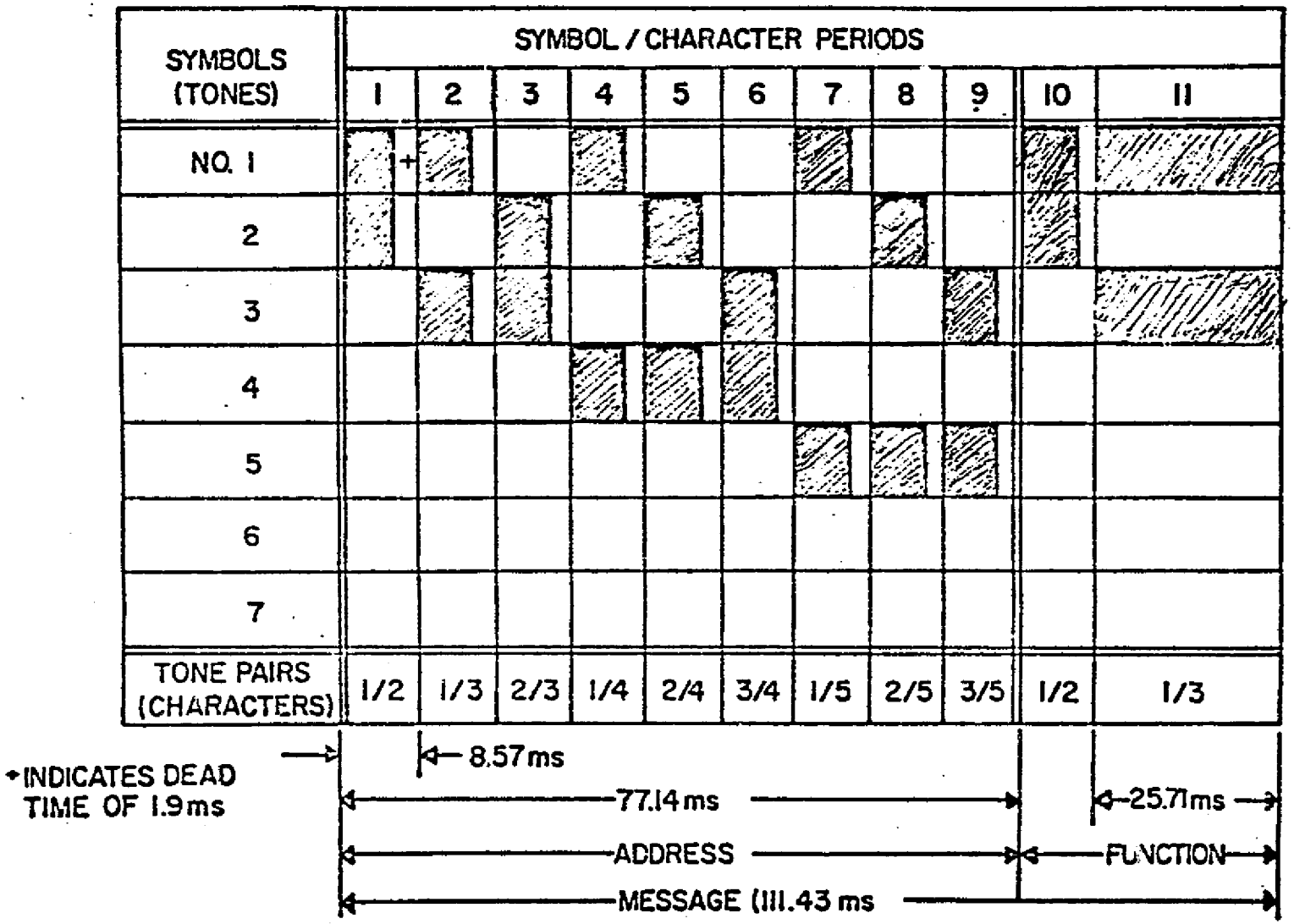


Figure 2. Typical Message Format

thresholds, the command system will accept a command if it received only two tones per character and only the correct characters corresponding to the secure address. Similarly, the command is accepted only if there are only two tones per character for the function selection. It should be noted that, with the use of these thresholds, it is easy to deny access to the command system by the jammer simply placing jamming tones at each of the seven possible subcarriers causing all commands to be rejected. This type of jamming is not considered a threat, however, because of the small utility to the jammer.

With the demodulator assumption of the use of thresholds, it is straightforward to calculate the possibility of a jammer causing an incorrect command to be accepted. Assume a technically competent jammer (i.e., possession of the encoder, transmitter, receiver, and decoder). There are seven possible tones taken two at a time (only two) corresponding to 21 possible characters. For a nine-character secure address, there are 21^9 or 7.9×10^{11} possibilities. To try all possible secure addresses, it would require 2.4×10^7 hours of continuous transmission. In 30 minutes, 1.6×10^4 possibilities could be tried or the probability of trying the correct secure address is 2×10^{-8}.

5.0 ERROR CORRECTING CODE TRADEOFFS

To design a convolutional coded communication system, techniques are required to predict probability of bit error (P_b) versus energy per bit to single-sided noise spectral density (E_b/N_0) for the modulation/demodulation scheme employed. In addition, it would be desirable to predict the nature of the errors out of the decoder. That is, do the errors occur in bursts and, if so, what is the probability of an error burst of a given length.

There are three basic analytical techniques available for predicting P_b versus E_b/N_0. The first technique uses Gallager's exponential bound for random codes. This bound allows quick comparisons of various modulation/demodulation and encoding/decoding schemes in terms of the basic coding parameters of code rate (R) and encoding constraint length (K). Therefore, it can be used as an initial optimization technique.

The second technique is to analyze the distance between codewords using a generating function (transfer function) developed by Viterbi. This analysis technique has been used to obtain tight upper bounds on the error probability for the binary symmetric channel (BSC), the additive white Gaussian noise (AWGN) channel, and general memoryless channels. These bounds can be used to optimize the communication system parameters for a particular convolutional code and Viterbi decoding. While the bounds on error probability using the transfer function approach are very precise, the bounds are difficult to calculate for long constraint codes. Also, for even moderate constraint lengths, the transfer function bounds are evaluated in matrix form and the individual codeword distances are not available to determine the effect of decoder memory truncation or to determine the statistics of the errors at the output of the decoder.

The third technique uses an approximation of the transfer function developed by Huth and Weber, which is easy to calculate even for long constraint length codes. The resulting bounds are as tight as Viterbi transfer function bounds and the individual codeword distances appear in the bound

so that it is possible to determine decoder memory truncation effects and output error statistics.

Appendix F compares the three approaches for predicting performance and presents results for Viterbi decoding bit error probability, burst error probability and decoder memory length. The predicted performance with the Huth-Weber approximations is compared with actual system tests and found to be very accurate.

Error correcting coding is required for 50 Mbps data link from the Shuttle Orbiter through the Tracking and Data Relay Satellite System (TDRSS) to the ground because of severe power limitations. Convolutional coding has been chosen because the decoding algorithms (sequential and Viterbi) provide significant coding gains at the required bit error probability of 10^{-6} and can be implemented at 50 Mbps with moderate hardware. While a 50 Mbps sequential decoder has been built, the highest data rate achieved for a Viterbi decoder is 10 Mbps. Thus, five multiplexed 10 Mbps Viterbi decoders must be used to provide a 50 Mbps data rate. Appendix G discusses the tradeoffs which were considered in selecting the multiplexed Viterbi decoder approach for this application.

It is felt that the selected 50 Mbps encoding/decoding approach offers a reasonable compromise between performance and system complexity. The coding gain provided by this approach is relatively high (5.3 dB at a probability of error of 10^{-6}), the system degrades gracefully (a coding gain of 2.7 dB is available at a probability of error of 10^{-2}), the spaceborne encoder is extremely simple to implement, and the ground-based decoders are available commercially and can be combined in a straightforward manner. Such an approach is not limited to 50 Mbps data rates. It appears reasonable to extend the parallel 10 Mbps Viterbi decoder concept to systems requiring operation at data rates well in excess of 100 Mbps, although reliability becomes of increasing concern as the number of parallel decoders becomes large.

6.0 SIGNAL DETECTION AND ANGULAR SEARCH PROCEDURE FOR SHUTTLE KU-BAND COMMUNICATION SYSTEM

To establish the Ku-band communication link between the Shuttle Orbiter and the TDRS, the antennas of these two satellites must be pointed at each other. Various techniques for pointing the TDRS antenna at the Orbiter have been considered, and the latest information available to Axiomatix indicates that initial pointing can be accomplished without requiring the Shuttle antenna axis to be in perfect alignment with the TDRS antenna axis. During this initial pointing, however, the effective EIRP available from the TDRS is only 40 dBw, instead of the nominal 50 dBw. But, before the TDRS makes its final pointing correction, the Shuttle antenna must search out the random designation errors which exist between its computer estimated line-of-sight (LOS) to the TDRS and its actual pointing direction. Appendix H analyzes several procedures for performing this search and estimates the time requirements for each.

Since the object of any search is to increase the energy of the received signal, there must be a signal detector to stop the search when the signal energy has exceeded a preset value or threshold. A simple and fast circuit for determining this is a square-law device followed by an integrating (LPF) filter and a threshold crossing indicator. Ideally, the predetection IF filter is matched to the RF spread spectrum; however, a practical Butterworth or Chebishev filter will degrade the predetection SNR only slightly. For B_{IF} = 20 MHz, this loss is 0.9 dB, so for P_{rec}/N_0 = 64.8 dB-Hz (assumes EIRP = 40 dBw), the predetection SNR is about -7.6 dB. Appendix H computes the required integration times for various values of probability of detection, P_D, and for a false alarm probability, $P_{fa} = 10^{-9}$. Assuming that the design is to detect the signal within a 1 dB beamwidth with P_D = 0.999, then the corresponding integration time is 220 microseconds, which is short compared to the dwell times of mechanically scanned antennas.

Having established the criteria for detecting the presence of the signal within the 1 dB antenna beamwidth, Appendix H considers the angular search procedure which insures that the designation uncertainty volume is scanned to provide a predetermined probability of true LOS intercept. The Shuttle antenna beam is spiraled out from its designated LOS to search for the true LOS. The spiral scan procedure is analyzed in Appendix I in terms of dwell time, total scan time, and scan rate. Several numerical examples pertinent to the Shuttle Ku-band radar/communication system are presented.

When the energy detector commands the Shuttle antenna to stop scanning and to start tracking, there may be a pull-in transient. A conservative estimate is that the pull-in time should be approximately one third of the inverse of the servo loop bandwidth. For a 1 Hz loop, this implies that about 0.3 second will be required for the servo loop to go from scan to track. Consequently, the time that the true LOS stays within the moving antenna beam, called the dwell time, must be at least 0.3 second.

The analysis presented in Appendixes H and I indicate that the times required to search the uncertainty angles for TDRS antenna acquisition are short compared to radar search times. This is due to the fact that the SNR available from TDRS for signal detection during acquisition is much higher than the SNR of radar targets. Consequently, the scan time limitation for TDRS acquisition is determined by servo considerations rather than signal detection. The approach for radar search is the same as that for TDRS acquisition, and only the dwell time requirements are different. We therefore recommend that the radar search pattern be adopted for TDRS acquisition.

7.0 FALSE LOCK PERFORMANCE OF COSTAS LOOP RECEIVERS

False lock can occur in Costas loop receivers when the frequency search requirement due to Doppler and oscillator instabilities is large enough to include false lock points. Appendix J demonstrates under what conditions the Costas loop can exhibit a false lock condition in the sense that the loop can lock up on a data sideband. In particular, for the case of random data, the loop can theoretically lock at a frequency which is any integer multiple of half the data rate. The required frequency search for both the S-band and Ku-band is larger than half of the data rate. Therefore, false lock conditions are expected on both sides.

Appendix J derives an expression for the DC output of the lock detector in terms of the Fourier transform of the symbol pulse shape, the squared magnitude of the lowpass arm filter transfer function, and the loop phase error. This expression is then evaluated at the true lock and false lock conditions and numerical results are given for the case of Manchester coded data and a single pole (RC) arm filter. It is found that, by widening the arm filter bandwidths beyond the optimum value which yields minimum squaring loss for the Costas loop, the separation between the true lock and false lock voltages at the lock detector increases. Therefore, the true lock case can be more readily distinguished from the false lock condition. However, widening the arm filter bandwidth also allows more noise to pass through to the lock detector, thus degrading the signal-to-noise ratio at the input to the threshold detector. Therefore, widening the arm filter bandwidth is potentially a solution to the false lock detection problem up to the point where the signal-to-noise ratio which distinguishes between true lock and out-of-lock (as opposed to false lock) becomes too low for a specified level of lock detection performance.

APPENDIX A

PHASE MULTIPLEXING FOR TWO- AND THREE-CHANNEL DATA TRANSMISSION

Gaylord K. Huth

APPENDIX A

PHASE MULTIPLEXING FOR TWO- AND THREE-CHANNEL DATA TRANSMISSION

Sergei Udalov

This appendix considers various techniques for simultaneous transmission of two or three independent and asynchronous digital data channels. A phase-multiplexing technique, using quadrature carriers, is shown to have potential implementation advantages over both a conventional PM subcarrier approach and the interplex approach which has been utilized for deep space communications. The phase-multiplexing technique is a hybrid approach which has some of the features of both quadrature modulation and interplex methods. Performance and implementation considerations are discussed for each of the various approaches.

1.0 INTRODUCTION

Phase-multiplexing offers a technique for accommodating multiple channels with independent data rates and clocks. This technique is especially important when high data rates are involved. Time-division multiplexing at high data rates is complex because of the requirement for reclocking the data. Frequency-division multiplexing requires large bandwidths to avoid intermodulation products. This appendix investigates the general phase multiplexing techniques, including conventional phase modulation with subcarriers, interplex [1-4], and hybrid interplex-quadrature modulation. The particular cases of two- and three-channel phase multiplexing using the various techniques are compared in terms of implementation and performance.

2.0 PHASE MULTIPLEXING PRINCIPLES

2.1 Conventional Phase Multiplexing

The general form of a phase-modulated RF signal of carrier frequency ω_c is

$$z(t) = \sqrt{2P}\ \sin\left[\omega_c t + \theta(t)\right] . \tag{1}$$

In the conventional modulation system, adapting the nomenclature used in [1], the combined phase modulation signal resulting from N channels is defined by the term:

$$\theta(t) = \sum_{n=1}^{N} \theta_n S_n(t) , \tag{2}$$

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

$$S_n(t) = d_n(t)\ sq\ (\omega_n t) = \pm 1 , \tag{3}$$

which defines the nth data stream biphase modulated on a squarewave subcarrier of frequency ω_n.

To show how the total RF power of the conventional PSK/PM system gets distributed among the various components, we shall consider the relatively simple cases of a three-channel and two-channel system.

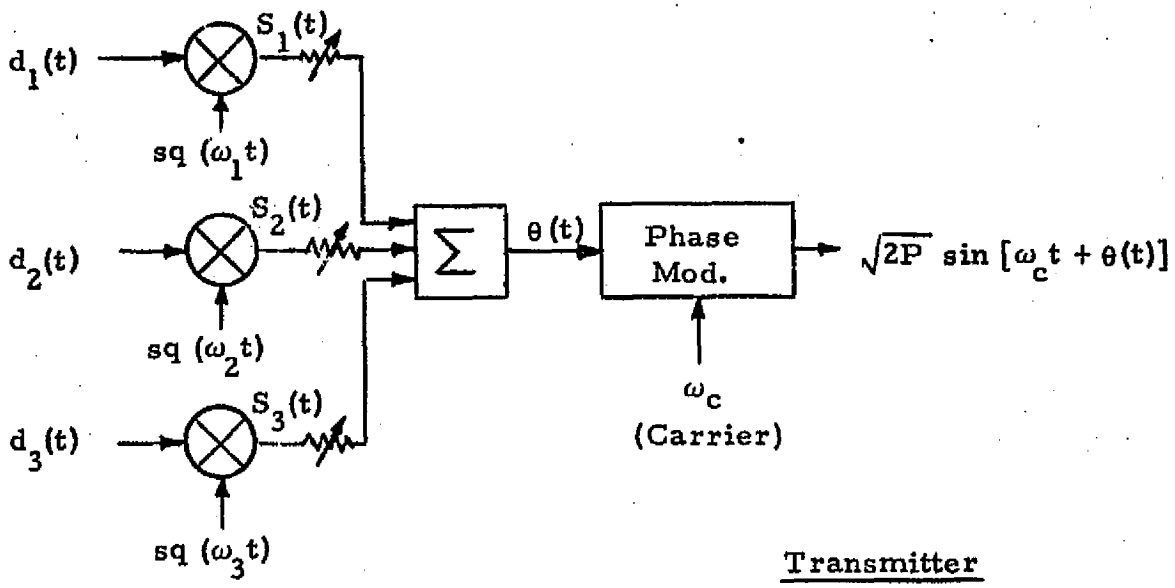
The modulator for the three-channel conventional PSK/PM system is shown in Figure 1. The simplified block diagram for the corresponding receiver is also shown here. For the case of three channels, the received phase modulated RF signal is

$$z(t) = \sqrt{2P}\ \sin\left[\omega_c t + \theta_1 S_1(t) + \theta_2 S_2(t) + \theta_3 S_3(t)\right] + n(t) , \tag{4}$$

where n(t) is assumed to be white Gaussian process of spectral density N_0 and zero mean.

Equation (4) can be rewritten in the form which shows the two orthogonal components of the RF signal:

$$\begin{aligned} z(t) &= \sqrt{2P}\ \sin\omega_c t\ \cos\left[\theta_1 S_1(t) + \theta_2 S_2(t) + \theta_3 S_3(t)\right] \\ &\quad + \sqrt{2P}\ \cos\omega_c t\ \sin\left[\theta_1 S_1(t) + \theta_2 S_2(t) + \theta_3 S_3(t)\right] + n(t) \\ &= \sqrt{2P}\ \sin\omega_c t\ \cos\theta(t) + \sqrt{2P}\ \cos\omega_c t\ \sin\theta(t) + n(t) . \end{aligned} \tag{5}$$



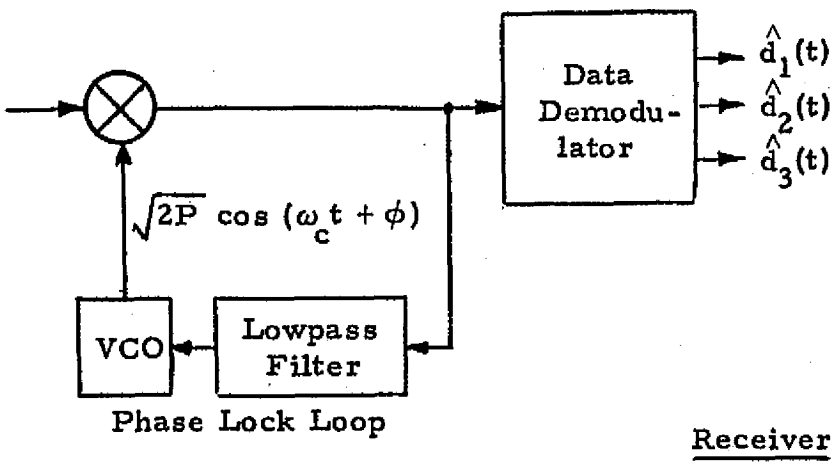


Figure 1. Conventional Three-Channel PSK/PM System

The trigonometric manipulations of the cos $\theta(t)$ and sin $\theta(t)$ terms results in detailed expansion of (5) into its components. Thus,

$$z(t) = \sqrt{2}\cos\omega_c t[\sqrt{P_1}\,S_1(t) + \sqrt{P_2}\,S_2(t) + \sqrt{P_3}\,S_3(t) - \sqrt{P_{cm}(1,2,3)}\,S_1(t)S_2(t)S_3(t)]$$

$$+ \sqrt{2}\sin\omega_c t\,[\sqrt{P_c} - \sqrt{P_{cm}(2,3)}\,S_2(t)S_3(t) - \sqrt{P_{cm}(1,2)}\,S_1(t)S_2(t)$$

$$- \sqrt{P_{cm}(1,3)}\,S_1(t)S_3(t)] + n(t)\,, \qquad (6)$$

where

$$P_1 = P\sin^2\theta_1\cos^2\theta_2\cos^2\theta_3 = \text{Power in channel 1}$$

$$P_2 = P\cos^2\theta_1\sin^2\theta_2\cos^2\theta_3 = \text{Power in channel 2}$$

$$P_3 = P\cos^2\theta_1\sin^2\theta_2\sin^2\theta_3 = \text{Power in channel 3}$$

$$P_c = P\cos^2\theta_1\cos^2\theta_2\cos^2\theta_3 = \text{Unmodulated carrier power}$$

$$P_{cm}(1,2,3) = P\sin^2\theta_1\sin^2\theta_2\sin^2\theta_3 = \text{Third-order cross-modulation power}$$

$$P_{cm}(1,2) = P\sin^2\theta_1\sin^2\theta_2\cos^2\theta_3$$

$$P_{cm}(2,3) = P\cos^2\theta_1\sin^2\theta_2\sin^2\theta_2 = \text{Second-order cross-modulation powers}\,.$$

$$P_{cm}(1,3) = P\sin^2\theta_1\cos^2\theta_2\sin^2\theta_3 \qquad (7)$$

Besides the obvious significance of wasted RF power, the role of the cross-modulation terms can be examined further by considering the signal recovery process in the receiver. As shown in Figure 1, the receiver consists of a phase-locked loop which tracks the phase and frequency of the unmodulated carrier component and, at the same time, coherently recovers the multiplexed subcarriers. The subcarriers are then applied to data demodulators for baseband data recovery. Thus, as shown, the incoming signal z(t) of (6) is multiplied by the local reference $\sqrt{2}\cos(\omega_c t + \phi)$, where ϕ is the residual RF carrier phase error. The result is:

$$y(t) = \cos\phi[\sqrt{P_1}\,S_1(t) + \sqrt{P_2}\,S_2(t) + \sqrt{P_3}\,S_3(t) - \sqrt{P_{cm}(1,2,3)}\,S_1(t)S_2(t)S_3(t)]$$

$$+ \sin\phi[\sqrt{P_c} - \sqrt{P_{cm}(2,3)}\,S_2(t)S_3(t) - \sqrt{P_{cm}(1,2)}\,S_1(t)S_2(t)$$

$$- \sqrt{P_{cm}(1,3)}\,S_1(t)S_3(t)] + n'(t)\,. \qquad (8)$$

From (8), it becomes evident that, if the frequency components of the second-order terms fall within the bandwidth of the lowpass filter, carrier tracking capability of the receiver will be impaired. Furthermore, the third-order term's spectrum, depending on its nature, can degrade either the carrier tracking or data recovery, or both. The selection of the subcarrier frequencies and the corresponding data rates thus becomes an important design consideration.

Equations (7) indicate that, for the case of the conventional PSK/PM, the unmodulated carrier power cannot be fully suppressed, such as for Costas loop reception, without suppressing all channels except channel 1. Furthermore, as pointed out in [1], the conventional PSK/FM system does not make use of the cross-modulation terms which do contain information which could be utilized if properly recovered. In [1], it is also shown that for less than five channels, a system called Interplex (see also [2-5]) is more efficient. Since, in this appendix, we are concerned primarily with two- and three-channel multiplexing, we will therefore concentrate on the interplex and related systems.

2.2 Interplex Modulation

With an interplex implementation of a PSK/PM system, 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

$$\theta(t) = \theta_1 S_1(t) + \sum_{n=2}^{N} \theta_n S_1(t) S_n(t) \ . \qquad (9)$$

Figure 2 shows the implementation of a three-channel interplex modulator. The corresponding equation for the received RF signal is

$$z(t) = \sqrt{2P} \sin [\omega_c t + \theta_1 S_1(t) + \theta_2 S_1(t)S_2(t) + \theta_3 S_1(t)S_3(t)] + n(t) \qquad (10)$$

where all of the terms are as defined earlier.

Rewriting (10) in terms of its orthogonal RF components, one obtains

$$\begin{aligned} z(t) &= \sqrt{2P} \sin \omega_c t \cos [\theta_1 S_1(t) + \theta_2 S_1(t)S_2(t) + \theta_3 S_1(t)S_3(t)] \\ &\quad + \sqrt{2P} \cos \omega_c t \sin [\theta_1 S_1(t) + \theta_2 S_1(t)S_2(t) + \theta_3 S_1(t)S_3(t)] + n(t) \\ &= \sqrt{2P} \sin \omega_c t \cos \theta^*(t) + \sqrt{2P} \cos \omega_c t \sin \theta^*(t) + n(t) \ , \qquad (11) \end{aligned}$$

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

Following the same procedure as for conventional PSK/PM, we expand (11) into its power term components. This results in

$$\begin{aligned} z(t) &= \sqrt{2} \cos \omega_c t [\sqrt{P_1} S_1(t) - \sqrt{P_{cm}(1,2,3)} S_1(t)S_2(t)S_3(t) + \sqrt{P_{cm}(1,2)} S_1(t)S_2(t) \\ &\quad + \sqrt{P_{cm}(1,3)} S_1(t)S_3(t)] + \sqrt{2} \sin \omega_c t [\sqrt{P_c} - \sqrt{P_{cm}(2,3)} S_2(t)S_3(t) \\ &\quad - \sqrt{P_2} S_2(t) - \sqrt{P_3} S_3(t)] + n(t) \qquad (12) \end{aligned}$$

where

$P_1 = P \sin^2\theta_1 \cos^2\theta_2 \cos^2\theta_3$ = Power in channel 1

$P_2 = P \sin^2\theta_1 \sin^2\theta_2 \cos^2\theta_3$ = Power in channel 2

$P_3 = P \sin^2\theta_1 \cos^2\theta_2 \sin^2\theta_3$ = Power in channel 3

$P_c = P \cos^2\theta_1 \cos^2\theta_2 \cos^2\theta_3$ = Unmodulated carrier power

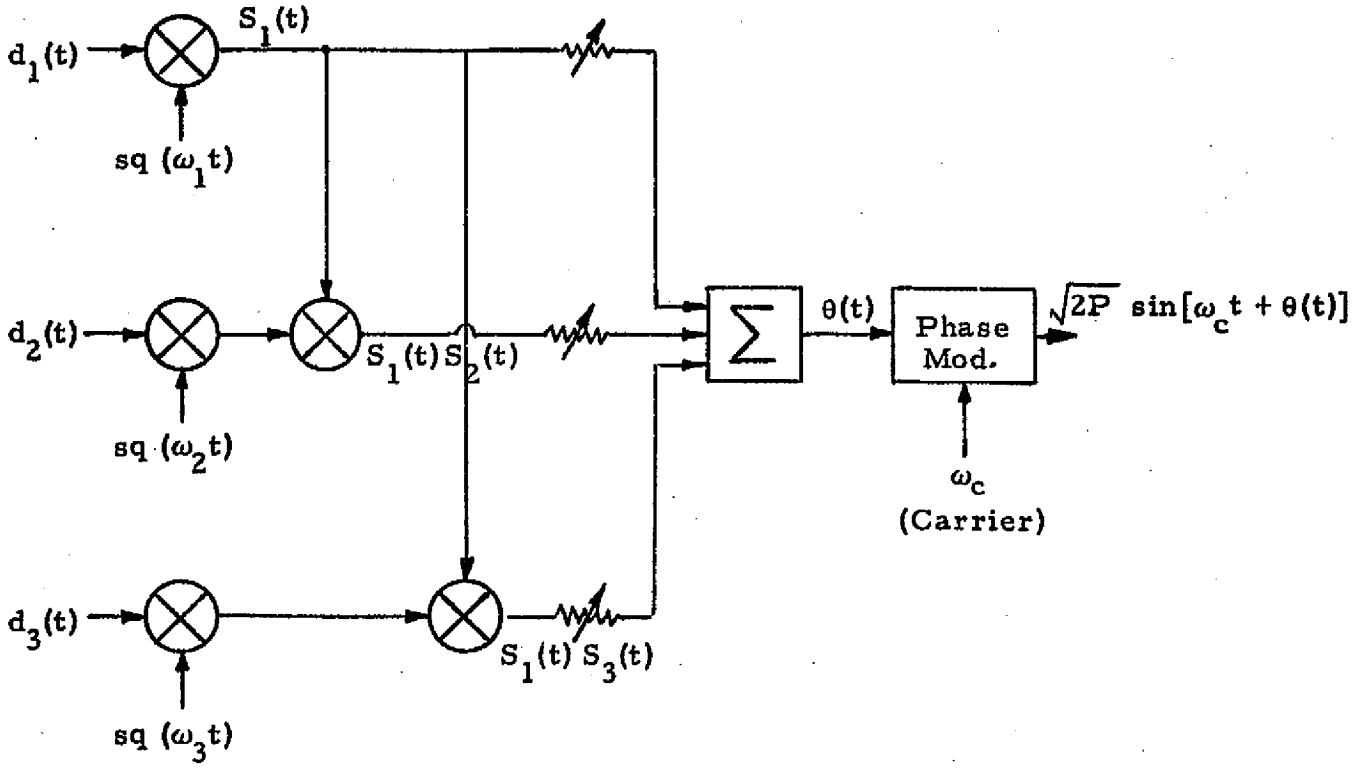


Figure 2. Three-Channel Interplex Modulator

$$P_{cm}(1,2,3) = P \sin^2\theta_1 \sin^2\theta_2 \sin^2\theta_3 \quad = \text{Third-order cross-modulation power}$$

$$\left.\begin{aligned} P_{cm}(1,2) &= P \cos^2\theta_1 \sin^2\theta_2 \cos^2\theta_3 \\ P_{cm}(2,3) &= P \cos^2\theta_1 \sin^2\theta_2 \sin^2\theta_3 \\ P_{cm}(1,3) &= P \cos^2\theta_1 \cos^2\theta_2 \sin^2\theta_3 \end{aligned}\right\} = \text{Second-order cross-modulation powers} \qquad (13)$$

The obvious significance of (13) is that letting $\theta_1 = \pm 90^\circ$ provides the following advantages:

(a) power in all three channels is maximized,

(b) unmodulated carrier is completely suppressed (Costas loop tracking can be utilized), and

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

For the case of $\theta_1 = \pm 90^\circ$, the signal portion of (12) can be written as

$$z(t) = \sqrt{2P} \cos \omega_c t \left[S_1(t) \cos\theta_2 \cos\theta_3 - S_1(t)S_2(t)S_3(t) \sin\theta_2 \sin\theta_3\right]$$

$$- \sqrt{2P} \sin \omega_c t \left[S_2(t) \sin\theta_2 \cos\theta_3 + S_3(t) \cos\theta_2 \sin\theta_3\right]. \qquad (14)$$

Equation (14) is useful for providing a geometrical representation of the three-channel interplex PSK/PM signal.

Figure 3 shows such representatior for two of the eight possible vector states of the composite signal. The salient feature of the vector diagram of Figure 3 is that the cross-modulation term, although interfering with the amplitude of the main channel signal, keeps the envelope of the composite signal constant. This can also be proved by showing that the sum of the squared bracketed terms of (14) is constant, regardless of the polarities of $S_n(t)$.

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 there is no spectral overlap between the three

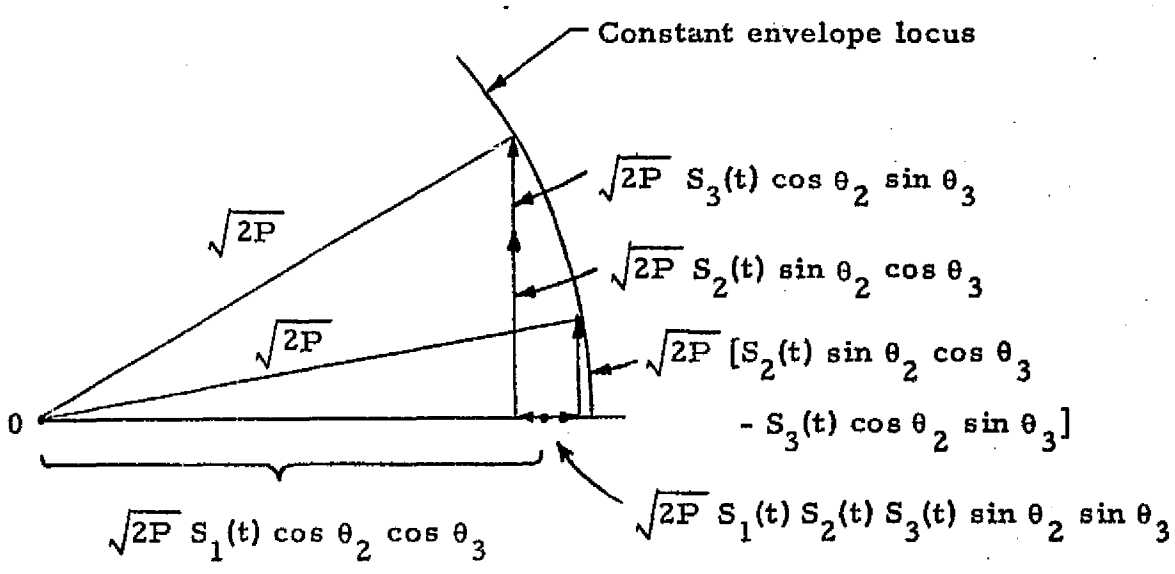


Figure 3. Geometrical Representation of a Three-Channel Interplex PSK/PM Signal

channels, there will be no interference regardless of the relationship shown in Figure 3.

On the other hand, 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 (13). Also, depending on the demodulation implementation, channel 1 can interfere with channels 2 and 3 if there is a spectrum overlap and carrier tracking error is not zero. This type of interference is discussed in other sections of this appendix.

The two-channel interplex system power allocations can be derived from (13) by setting $\theta_3 = 0$. This results in

$$\begin{aligned} P_1 &= P \sin^2\theta_1 \cos^2\theta_2 && = \text{Power in channel 1} \\ P_2 &= P \sin^2\theta_1 \sin^2\theta_2 && = \text{Power in channel 2} \\ P_c &= P \cos^2\theta_1 \cos^2\theta_2 && = \text{Unmodulated carrier power} \\ P_{cm} &= P \cos^2\theta_1 \sin^2\theta_2 && = \text{Cross-modulation power} \end{aligned} \tag{15}$$

which, of course, is the same as the equation set (15) in [1]. Furthermore, setting $\theta_1 = \pm 90^\circ$ for maximum efficiency reduces the two-channel interplex system to a quadriphase system with unequal powers in the two orthogonal components. Thus, (14) becomes simply

$$z_u(t) = \sqrt{2P}\cos\omega_c t\,[S_1(t)\cos\theta_2] - \sqrt{2P}\sin\omega_c t\,[S_2(t)\sin\theta_2] \tag{16}$$

where the subscript u indicates that we are now dealing with an unequal power division quadriphase modulation. The corresponding polar form of this equation is

$$z_u(t) = \sqrt{2P}\cos[\omega_c t + S_1(t)\,S_2(t)\,\theta_2]\,. \tag{17}$$

2.3 Equivalence of Two-Channel Interplex ($\theta_1 = \pm 90^\circ$) and Quadriphase Modulation (QPSK)

Consider now a signal which results from addition of two quadrature carriers, each of which is amplitude-modulated (DSB-SC) by a separate bipolar data sequence. The block diagram for generation of such a composite signal is shown in Figure 4. The resultant QPSK signal is given by

$$f(t) = A\, a(t) \cos \omega_c t - B\, b(t) \sin \omega_c t \tag{18}$$

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 comparison consistency with interplex.) It is customary to reduce (18) to an equivalent polar form,

$$f(t) = R \cos (\omega_c t + \theta) . \tag{19}$$

For this, we use trigonometric identity:

$$\cos (\omega_c t + \theta) = \cos \omega_c t \cos \theta - \sin \omega_c t \sin \theta . \tag{20}$$

Thus,

$$\begin{aligned} f(t) &= A\, a(t) \cos \omega_c t - B\, b(t) \sin \omega_c t \\ &= R \cos \omega_c t\, [a(t) \cos \theta] - R \sin \omega_c t\, [b(t) \sin \theta] . \end{aligned} \tag{21}$$

Note that (20) is of the same form as (16). Furthermore, for (20) to be true for all values of t, the coefficients of $\cos \omega_c t$ and $\sin \omega_c t$ must be such that

$$\begin{aligned} A\, a(t) &= R\, a(t) \cos \theta \\ B\, b(t) &= R\, b(t) \sin \theta \end{aligned}$$

$$\begin{aligned} R &= \sqrt{A^2 + B^2} \\ \theta &= \tan^{-1} \left[\frac{B\, b(t)}{A\, a(t)} \right] . \end{aligned} \tag{22}$$

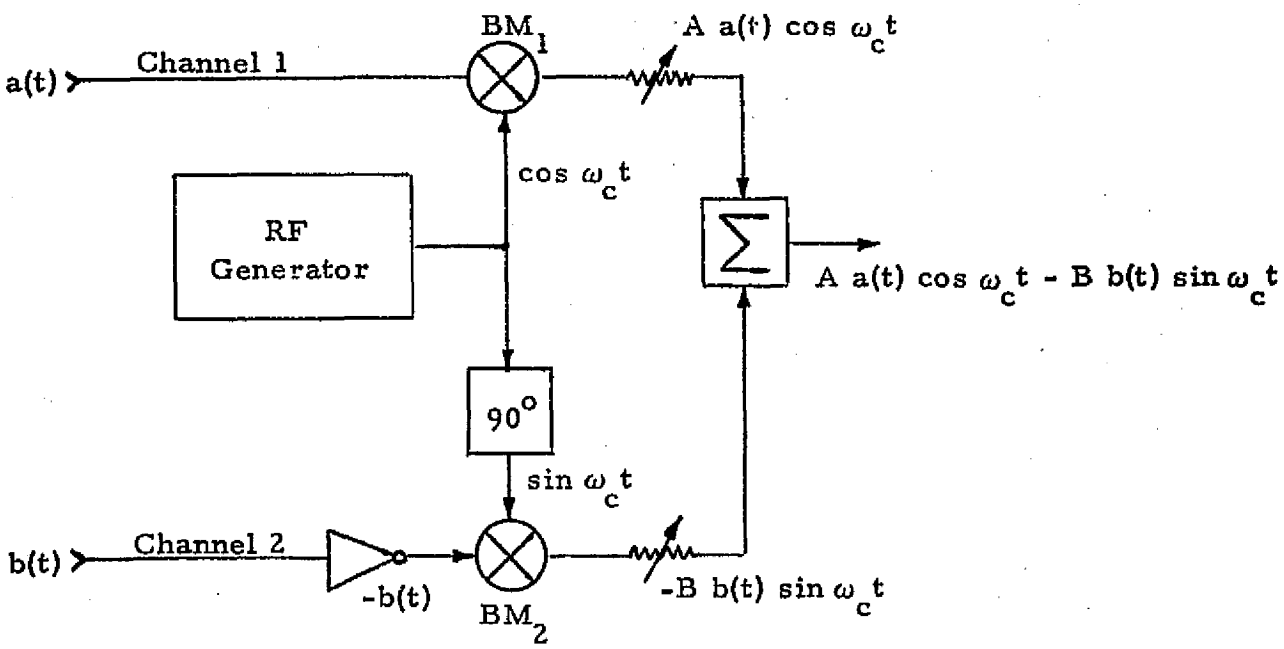


Figure 4. Two-Channel Quadriphase Modulator

Hence, (18) can be written as

$$f(t) = \sqrt{A^2 + B^2}\ \cos\left\{\omega_c t + \tan^{-1}\left[\frac{B\ b(t)}{A\ a(t)}\right]\right\} . \qquad (23)$$

The equivalence with two-channel ($\theta_1 = \pm 90^\circ$) interplex is estallished by conversion coefficients

$$A = R \cos\theta_2 = \sqrt{2P}\cos\theta_2$$

$$B = R \sin\theta_2 = \sqrt{2P}\sin\theta_2$$

$$A/B = \tan\theta_2$$

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

$$b(t) = S_2(t) . \qquad (24)$$

Thus, we have shown that, 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 with off-the-shelf components. Additional similarities and differences are discussed in other sections of this appendix.

2.4 Three-Channel PM Multiplexer/Modulator Implementation Based on Quadrature Carrier Summation

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 [6-8]. Such a three-channel multiplexer/modulator is shown in Figure 5. 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 other hand, the similarity with quadriphase modulation is due to the fact that two quadrature components of the same RF carrier are used to synthesize the composite signal.

As shown in Figure 5, the modulator consists of two balanced mixers, BM_1 and BM_2, which are supplied, respectively, by the two orthogonal components of an RF carrier. BM_1 accepts the primary (wideband) data stream a(t) and thus provides a constant envelope biphase modulated carrier to the summing network. BM_2 accepts a linear sum of baseband secondary data stream b(t) and a square wave subcarrier biphase modulated by the third data stream c(t). The output of BM_2 is then a double sideband, suppressed carrier whose amplitude and phase vary according to the instantaneous sum of the two signals at its modulation input terminal.

The outputs of BM_1 and BM_2 are summed and applied to an amplitude limiter. In this limiter, the amplitude modulation of BM_2 output is converted into a phase modulation which adds to the antipodal modulation of the main carrier vector. The addition is such that it provides an orthogonal carrier whose amplitude and phase vary in accordance with the output of BM_2. The amplitude limiter, however, keeps the amplitude of the composite signal constant. This process can be understood by referring to the vector diagram shown in Figure 6. Before limiting, the composite signal consists of an orthogonal addition of $|A'|$ and $|B' \pm C'|$, where A', B' and C' are the amplitudes of the three data streams prior to limiting. After limiting, the amplitude of the A vector gets modulated, the nature and the degree of modulation being determined by the instantaneous polarities of the B and C vectors. This modulation is similar to the third-order cross-modulation

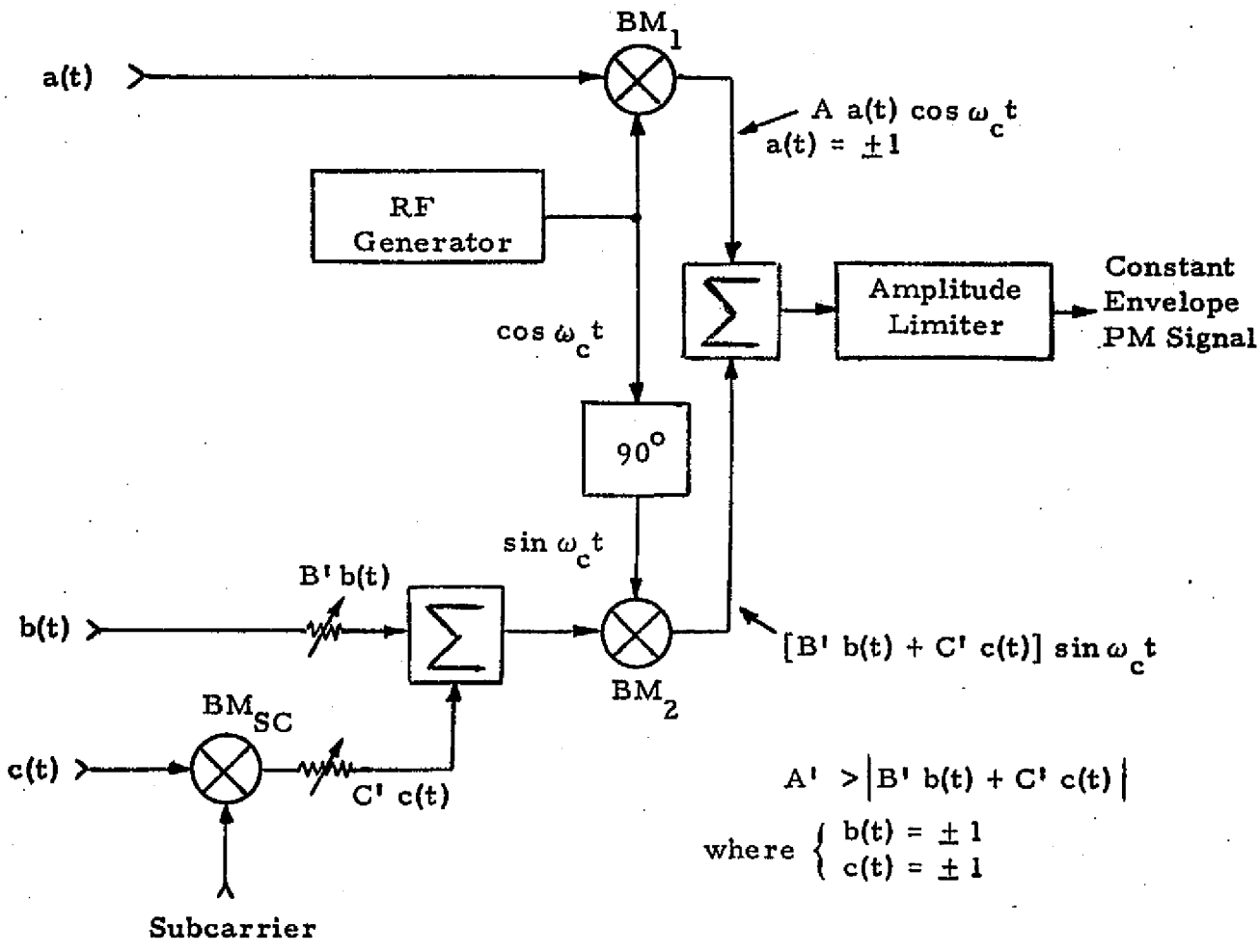


Figure 5. Three-Channel PM Multiplexer/Modulator Implementation Based on Quadrature Carrier Summation

generated by a three-channel interplex signal. The magnitude of this modulation can be computed by considering the terms:

$$A_{max} = R \cos\left[\tan^{-1}\left(\frac{B' - C'}{A'}\right)\right] \tag{25}$$

$$A_{min} = R \cos\left[\tan^{-1}\left(\frac{B' + C'}{A'}\right)\right] \tag{26}$$

$$A_{ave} = R \cos\left[\tan^{-1}\left(\frac{B'}{A'}\right)\right] \tag{27}$$

where R is the magnitude of the constant envelope of the RF signal out of the amplitude limiter and A_{max}, A_{min}, and A_{ave} are as indicated in Figure 6.

The cross-modulation term's amplitude, normalized to the average amplitude of the A vector, is approximated by

$$A_{ave} = \frac{A_{max} - A_{min}}{2 A_{ave}}$$

$$= \frac{\cos\left[\tan^{-1}\left(\frac{B' - C'}{A'}\right)\right] - \cos\left[\tan^{-1}\left(\frac{B' + C'}{A'}\right)\right]}{2 \cos\left[\tan^{-1}\left(\frac{B'}{A'}\right)\right]} . \tag{28}$$

This term is essentially equivalent to the corresponding term obtained from (13) for the interplex system. The latter term is

$$\Delta(S_1) = \sqrt{\frac{P_{cm}(1,2,3)}{P_1}} = \tan\theta_2 \tan\theta_3 , \tag{29}$$

where S_1 stands for the amplitude of the cross-modulation term superimposed on the primary data stream $S_1(t)$ of the three-channel interplex system with $\theta_1 = \pm 90^\circ$. Note that both (28) and (29) are third-order cross-modulation terms and, as shown below, these terms are approximately equal provided the primary/secondary channel power division ratios are 6 dB (4 to 1) and up.

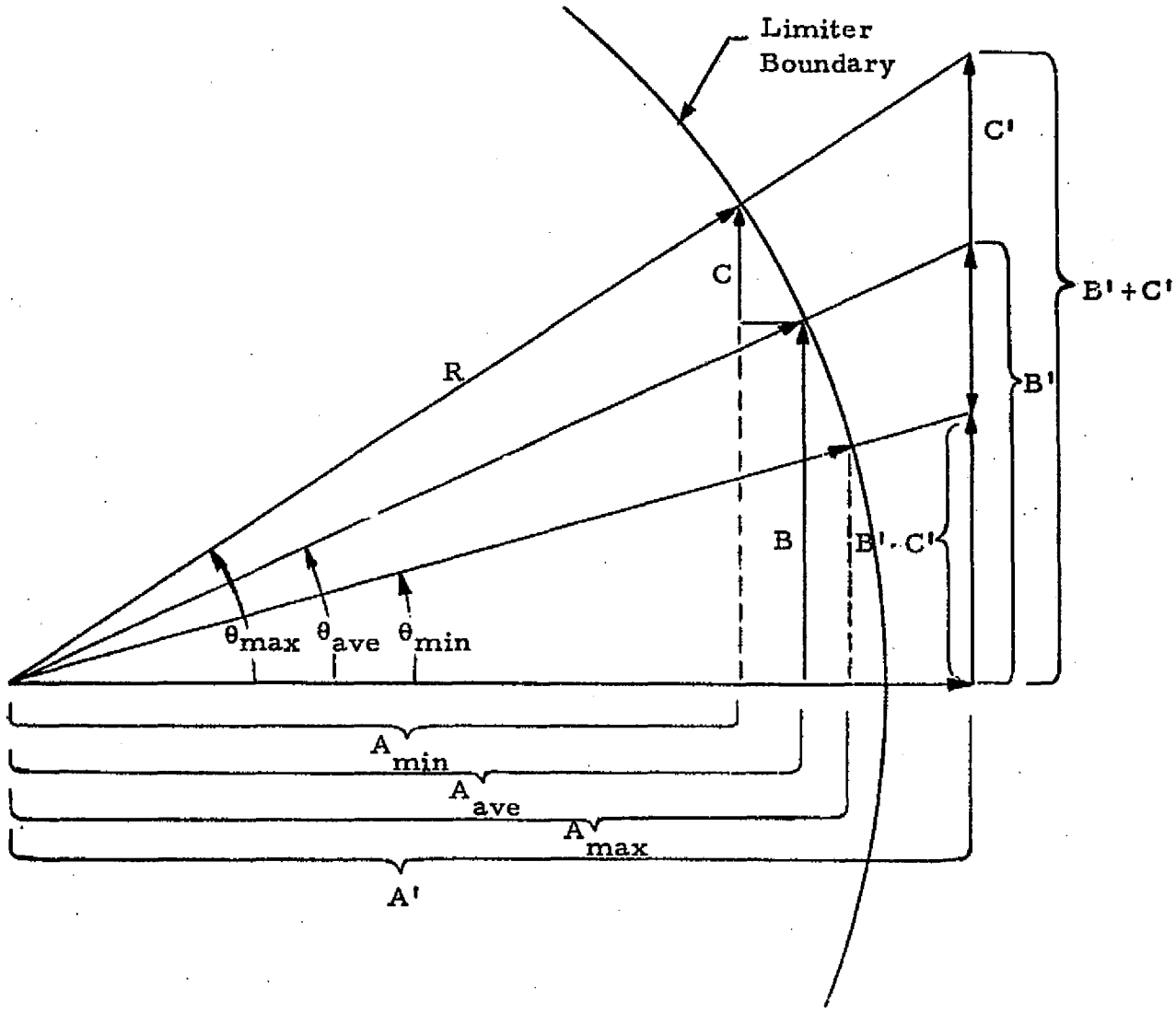


Figure 6. Vectorial Addition of Two Quadriphase Carriers Modulated by Three Data Streams

2.5 Demodulation

For both the interplex and the hybrid phase multiplexing, it has been shown that the antipodal modulation should be used for the primary channel, i.e., the channel which carries the highest data rate and thus contains the largest fraction of the total RF power. For such power allocation, combined with antipodal modulation, a Costas loop demodulation can be used for carrier tracking and data demodulation. As described in [7] and in Appendix B, such a demodulator, although not optimal, provides good performance and yet is relatively simple to implement. Such a demodulator is shown in Figure 7. With the carrier frequency and phase lock established within residual phase tracking error ϕ, the demodulated primary data stream appears at the output of the balanced mixer BM_1. The secondary data streams which are modulated on the orthogonal component of the carrier appear at the output of BM_2. These signals are then processed in their respective data recovery circuits and the data streams recovered.

2.5.1 Two-Channel Quadriphase Input Demodulation

Assuming that the frequency lock is perfect, i.e., $\omega_c' = \omega_c$, and that only the residual carrier phase error ϕ is present ($\phi \ll 45^\circ$), the expressions for the outputs of BM_1 and BM_2 in Figure 7 can be obtained for the baseband terms by neglecting the $2\omega_c$ terms of the appropriate trigonometric expansions. Thus, the baseband output of BM_1 is

$$y_1(t) = [A\, a(t) \cos \omega_c t + B\, b(t) \sin \omega_c t + n(t)] \cdot \sqrt{2} \cos(\omega_c t + \phi)$$

$$= \underbrace{\frac{1}{\sqrt{2}} [A\, a(t) + n_c(t)] \cos\phi}_{\text{Channel 1 data and noise}} + \underbrace{\frac{1}{\sqrt{2}} [B\, b(t) + n_s(t)] \sin\phi}_{\text{Channel 2 crosstalk and noise}} \qquad (30)$$

Similarly, the baseband output of BM_2 is

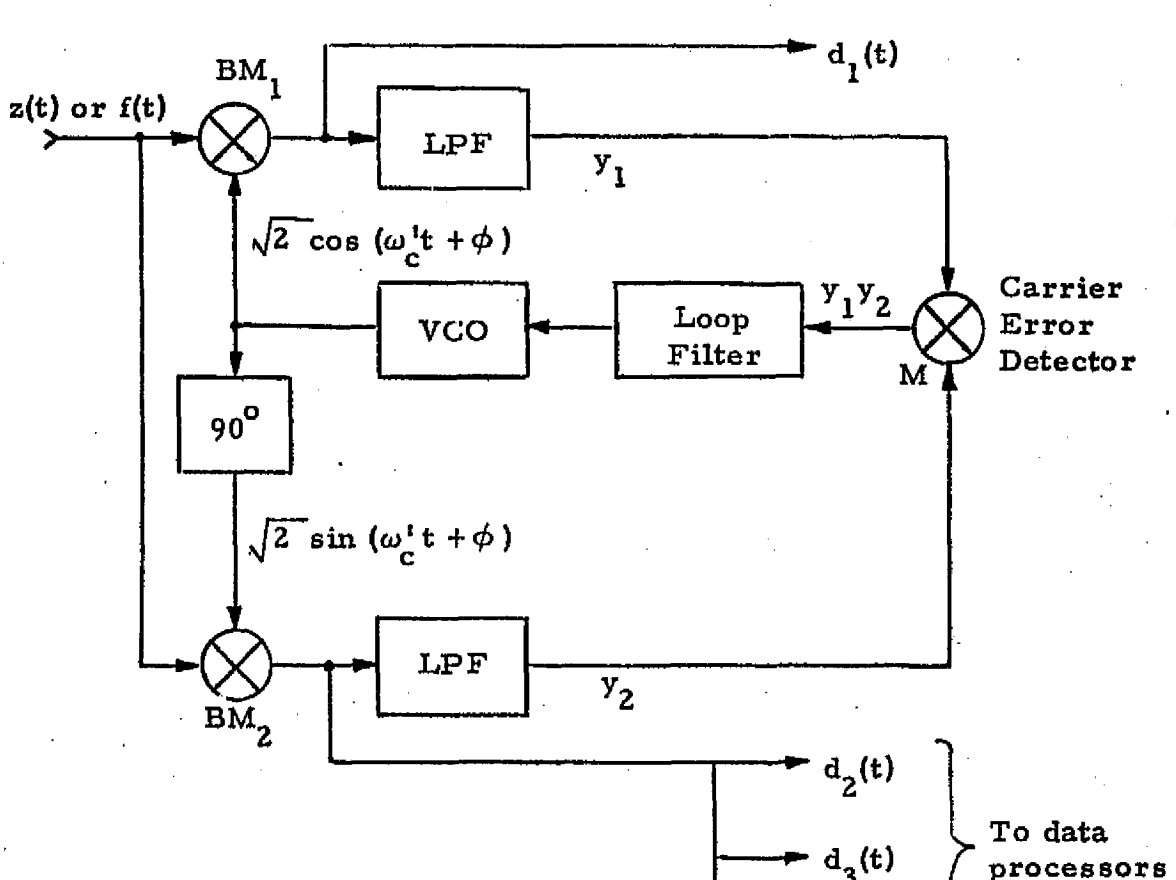


Figure 7. Costas Loop Demodulator for Interplex or Quadriphase Modulation

$$y_2(t) = \underbrace{\frac{1}{\sqrt{2}}\,[A\,a(t) + n_c(t)]\,\sin\phi}_{\text{Channel 1 crosstalk and noise}} - \underbrace{\frac{1}{\sqrt{2}}\,[B\,b(t) + n_s(t)]\,\cos\phi}_{\text{Channel 2 data and noise}} . \qquad (31)$$

The terms $n_c(t)$ and $n_s(t)$ in both (30) and (31) are the projections of the bandpass noise in the cosine and sine directions of the composite modulation process.

From (30) and (31), it is evident that the penalty paid for using only the major component of the quadriphase signal is the existence of the crosstalk between the channels when the tracking error ϕ is not zero. Also, as is shown below, the power available for carrier tracking is also reduced, but this reduction is not significant (less than 1 dB) if power division ratio between the two channels is 6 dB or more.

Let $D_k(\ell)$ be the ratio of the power in the kth channel to the crosstalk caused by the ℓth channel data. Also let B_1 and B_2 be the bandwidths (to the first null) of the A a(t) and B b(t) signals, respectively. Now if we assume that $B_1 > B_2$, which is generally the case, we obtain, neglecting the noise terms in (30) and (31),

$$D_1(2) = \frac{A^2[a(t)]^2}{B^2[b(t)]^2}\left(\frac{\cos\phi}{\sin\phi}\right)^2 = \frac{P_1}{P_2}\cot^2\phi \qquad (32)$$

$$D_2(1) = \frac{B^2[b(t)]^2}{A^2[a(t)]^2}\left(\frac{B_1}{B_2}\right)\left(\frac{-\cos\phi}{\sin\phi}\right)^2 = \left(\frac{P_2}{P_1}\right)\left(\frac{B_1}{B_2}\right)\cot^2\phi \qquad (33)$$

where P_1 and P_2 are the powers of the respective carrier components and ϕ is the residual carrier tracking error.

Note that in (32), the bandwidth ratio does not appear because as long as B_2 is located within the spread of B_1, all of the power P_2 contributes to the crosstalk. However, because the rate of b(t) is less than that of a(t), the crosstalk appears as a bias which shifts the mean of the A a(t) signal.

On the other hand, the $D_2(1)$ crosstalk has a general nature of filtered pseudorandom signal and thus is a noise-like process with statistics being determined by the properties of the a(t) and b(t) data streams.

The carrier tracking error is provided by multiplying y_1 and y_2 in the carrier error detector M and lowpass filtering. The output of the carrier error detector is thus:

$$e(t) = y_1 y_2 = \frac{1}{4}[A^2 - B^2 + 2A\,a(t)\,n_c(t) - 2B\,b(t)\,n_s(t) + n_c^2(t) - n_s^2(t)] \sin 2\phi - \frac{1}{2}[AB\,a(t)b(t) + B\,b(t)n_c(t) + A\,a(t)n_s(t) + n_c(t)n_s(t)] \cos 2\phi \,. \tag{34}$$

Note that, in the absence of noise, (34) reduces to

$$e(t) = \frac{A^2 - B^2}{2} \sin 2\phi - \frac{AB\,a(t)\,b(t)}{2} \cos 2\phi \,. \tag{35}$$

The first term in (35) is the carrier tracking DC term, which is reduced from its full possible value by not utilizing the power in the second channel. This reduction is negligible, however, for channel power division ratios of 6 dB or more. The second term represents wideband crosstalk which enters the carrier tracking loop and is a function of spectral properties of a(t) and b(t). The effect of this term on the carrier tracking performance is, furthermore, the function of the carrier tracking loop bandwidth. In general, the faster the data streams a(t) and b(t) and the narrower the loop bandwidth, the smaller is the interference.

A typical situation may involve Manchester type data for both data streams, resulting in spectrums shown qualitatively in Figure 8. The convolution of the two spectra results in the cross-product spectrum shown in Figure 9. From this figure, it is evident that, in the region of the low frequencies, where the carrier tracking bandwidth lies, the spectral density of the cross-product signal is relatively flat.

The actual magnitude of the spectral density in the low frequency region is given in Appendix B by

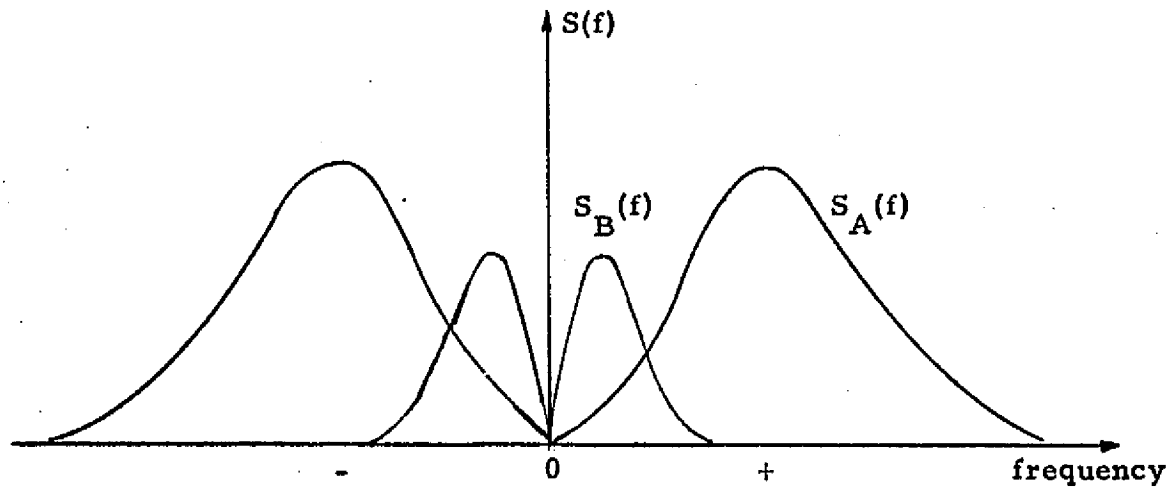


Figure 8. Manchester Spectrums for Channel 1 and Channel 2

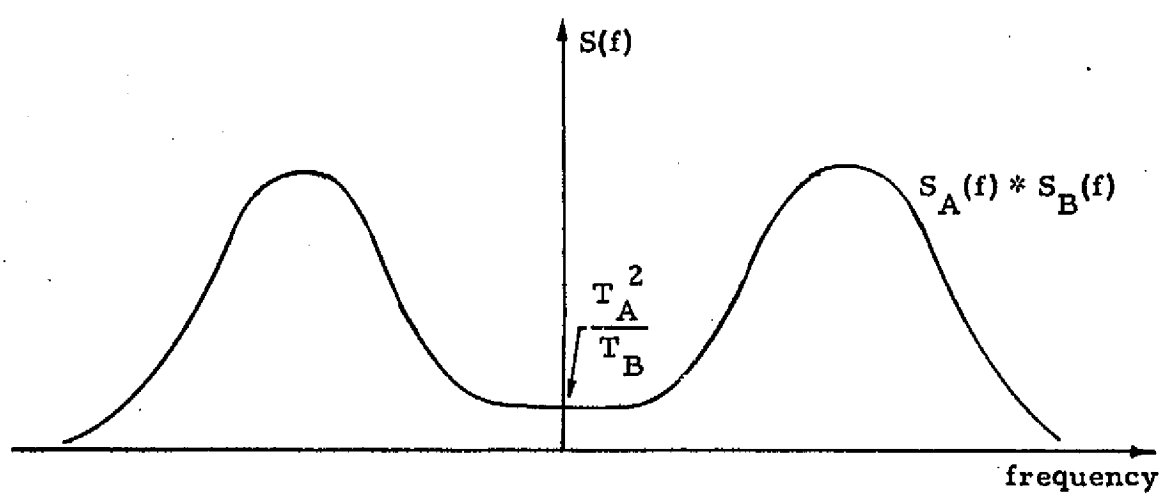


Figure 9. Carrier Loop Cross-Product Spectrum for Case of Two Manchester Data Streams

$$S(0) = \frac{T_A^2}{2\,T_B} \tag{36}$$

where $S(0)$ = spectral density of $a(t)\,b(t)$ at $f = 0$

T_A = symbol time of $a(t)$

T_B = symbol time of $b(t)$.

The expression for the signal-to-crosstalk power in the carrier tracking loop, for the noiseless case and both data streams Manchestered, is thus

$$\frac{P}{S(0)\,B_L} = \frac{(A^2 - B^2)^2}{2\,A^2B^2} \cdot \frac{R_A^2}{R_B\,B_L} \quad \text{for } R_A \geq 2R_B \tag{37}$$

where R_A = Channel 1 data rate

R_B = Channel 2 data rate

A = Channel 1 signal amplitude

B = Channel 2 signal amplitude

B_L = Carrier tracking loop noise bandwidth .

For the case where $a(t)$ is biphase modulated and only $b(t)$ is Manchestered, the spectral density in the low frequency region is expressed in Appendix B by

$$S(0) = T_A\left(1 - \frac{T_A}{T_B}\right) \tag{38}$$

and, correspondingly, the signal-to-crosstalk power in the carrier tracking loop is

$$\frac{P}{S(0)\,B_L} = \frac{(A^2 - B^2)^2}{4\,A^2B^2} \cdot \frac{R_A^2}{(R_A - R_B)\,B_L} \quad \text{for } R_A \geq 2R_B \tag{39}$$

where all the symbols in (38) and (39) are the same as previously defined.

2.5.2 Two-Channel Interplex Input Demodulation

A two-channel, quadrature interplex signal can also be recovered by the Costas demodulator of Figure 7. The corresponding expressions for the y_1 and y_2 signals are obtained by multiplying $z_u(t)$ of (16) with the cosine and sine components of the locally generated replica of the RF carrier. Thus, one obtains

$$y_1(t) = \underbrace{\sqrt{P}\,[S_1(t)\cos\theta_2 + n_c(t)]\cos\phi}_{\text{Channel 1 data and noise}} + \underbrace{\sqrt{P}\,[S_2(t)\sin\theta_2 + n_s(t)]\sin\phi}_{\text{Channel 2 crosstalk and noise}} \quad (40)$$

and

$$y_2(t) = \underbrace{\sqrt{P}\,[S_1(t)\cos\theta_2 + n_c(t)]\sin\phi}_{\text{Channel 1 crosstalk and noise}} - \underbrace{\sqrt{P}\,[S_2(t)\sin\theta_2 + n_s(t)]\cos\phi}_{\text{Channel 2 data and crosstalk}} \quad (41)$$

For a noiseless case, it is a simple matter to derive from (40) and (41) the data crosstalk equations similar to (32) and (33). These are

$$D_1(2) = \left(\frac{\cos^2\theta_2}{\sin^2\theta_2}\right)\left(\frac{\cos\phi}{\sin\phi}\right)^2 = \cot^2\theta_2\cot^2\phi \quad (42)$$

$$D_2(1) = \left(\frac{\sin^2\theta_2}{\cos^2\theta_2}\right)\left(\frac{-\cos\phi}{\sin\phi}\right)^2\left(\frac{B_1}{B_2}\right) = -\left(\frac{B_1}{B_2}\right)\tan^2\theta_2\cot^2\phi\,. \quad (43)$$

We can also derive the equivalent expression for the carrier tracking error, which for the two-channel interplex is

$$e(t) = y_1y_2 = \frac{P}{2}\,[\cos 2\theta_2\sin 2\phi - S_1(t)\,S_2(t)\sin 2\theta_2\cos 2\phi]\,. \quad (44)$$

Note that (42), (43), and (44) can be derived directly from (32), (33) and (35) by use of the conversion coefficients of (24).

2.5.3 Three-Channel Interplex Demodulation

When the input to the Costas demodulator in Figure 7 is a three-channel interplex signal defined by (14), the corresponding $y_1(t)$ and $y_2(t)$ signals are as shown below. The $y_1(t)$ signal is

$$y_1(t) = \sqrt{P}\,[\underbrace{S_1(t)\cos\theta_2\cos\theta_3}_{\text{Channel 1 data}} - \underbrace{S_1(t)\,S_2(t)\,S_3(t)\sin\theta_2\sin\theta_3}_{\text{Cross-modulation}} + \underbrace{n_c(t)}_{\text{Noise}}]\cos\phi$$

$$+\sqrt{P}\,[\underbrace{S_2(t)\sin\theta_2\cos\theta_3}_{\substack{\text{Channel 2 data}\\ \text{crosstalk}}} + \underbrace{S_3(t)\cos\theta_2\sin\theta_3}_{\substack{\text{Channel 3 data}\\ \text{crosstalk}}} + \underbrace{n_s(t)}_{\text{Noise}}]\sin\phi \qquad (45)$$

Equation (45) shows that, even if the carrier phase tracking is perfect (i.e., $\phi = 0$), a small third-order cross-modulation term contaminates S_1 data output. This is not the case for the y_2 output where data crosstalk occurs only for $\phi \neq 0$:

$$y_2(t) = \sqrt{P}\,[\underbrace{-S_2(t)\sin\theta_2\cos\theta_3}_{\text{Channel 2 data}} - \underbrace{S_3(t)\cos\theta_2\sin\theta_3}_{\text{Channel 3 data}} - \underbrace{n_s(t)}_{\text{Noise}}]\cos\phi$$

$$+\sqrt{P}\,[\underbrace{S_1(t)\cos\theta_2\cos\theta_3}_{\substack{\text{Channel 1 data}\\ \text{crosstalk}}} - \underbrace{S_1(t)S_2(t)S_3(t)\sin\theta_2\sin\theta_3}_{\text{Cross-modulation}} + \underbrace{n_c(t)}_{\text{Noise}}]\sin\phi\,. \qquad (46)$$

Let $D_k(\ell, m)$ be the ratio of the power in the kth data channel to the crosstalk caused by either the ℓ th or mth channel, or both. Also let B_n be the bandwidth (to the first null) of the nth channel. Equations (45) and (46) can now be used to define the $D_k(\ell, m)$ ratios. Thus, we have for the higher rate data channel:

$$D_1(2) = \cot^2\theta_2 \cot^2\phi \tag{47}$$

$$D_1(3) = \cot^2\theta_3 \cot^2\phi \tag{48}$$

$$D_1(2,3) = \cot^2\theta_2 \cot^2\theta_3 \tag{49}$$

and, for the secondary, lower rate data channels:

$$D_2(1) = \frac{B_1}{B_2} \tan^2\theta_2 \cot^2\phi \tag{50}$$

$$D_3(1) = \frac{B_1}{B_3} \tan^2\theta_3 \cot^2\phi \tag{51}$$

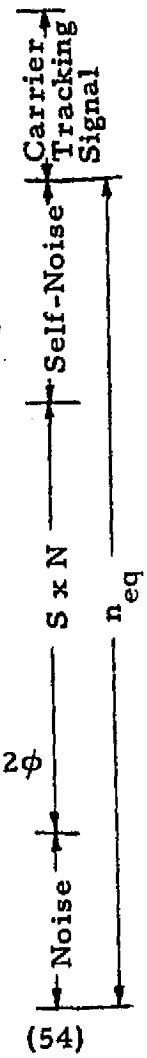
$$D_2(1,3) = \frac{B_{13}}{B_2} \cot^2\theta_3 \cot^2\phi \tag{52}$$

$$D_3(1,2) = \frac{B_{12}}{B_3} \cot^2\theta_2 \cot^2\phi\,. \tag{53}$$

Note that in (50) through (53), the data bandwidth ratios are included. These ratios, however, provide only the first-order approximations, particularly in (52) and (53) where B_{12} and B_{13} are bandwidths determined by convolving B_1 with B_2 and B_1 with B_3, respectively. Therefore, to obtain an accurate bandwidth ratio improvement factor, each specific combination of data rates and formats must be examined in detail. This is also an area where a great deal of optimization can be obtained for any particular link by selecting the appropriate modulation formats for the component data streams of the composite multi-channel signal.

The carrier tracking error signal which is applied to the loop filter is, prior to filtering, the product of (45) and (46). Thus, for a three-channel interplex signal, we have

$$
\begin{aligned}
e_c(t) &= y_1(t)\, y_2(t) \\
&= \frac{P}{2} \cos 2\theta_2 \cos 2\theta_3 \sin 2\phi \\
&\quad - \frac{P}{2} [S_2(t) S_3(t) \sin 2\theta_2 \sin 2\theta_3] \sin 2\phi \\
&\quad - \frac{P}{2} [S_1(t) S_2(t) \sin 2\theta_2 \cos 2\theta_3 + S_1(t) S_3(t) \cos 2\theta_2 \sin 2\theta_3] \cos 2\phi \\
&\quad - P[S_1(t) \cos\theta_2 \cos\theta_3 n_s - S_1(t) S_2(t) S_3(t) \sin\theta_2 \sin\theta_3 n_s \\
&\qquad + S_2(t) \sin\theta_2 \cos\theta_3 n_c + S_3(t) \cos\theta_2 \sin\theta_3 n_c] \cos 2\phi \\
&\quad + P[S_1(t) \cos\theta_2 \cos\theta_3 n_c - S_2(t) \sin\theta_2 \cos\theta_3 n_s \\
&\qquad - S_3(t) \cos\theta_2 \sin\theta_3 n_s - S_1(t) S_2(t) S_3(t) \sin\theta_2 \sin\theta_3 n_c] \sin 2\phi \\
&\quad - P(n_c n_s) \cos 2\phi \\
&\quad + \frac{P}{2} (n_c^2 - n_s^2) \sin 2\phi \, .
\end{aligned}
\tag{54}
$$



A similar expression can be obtained for the hybrid multiplexer signal by substituting the appropriate magnitudes of the A, B and C components. Equation (54) and its equivalent for the hybrid multiplexer can then be used to obtain carrier tracking performance as a function of power division ratios, data formats, data rates, and the P/N_0 ratio.

2.6 Numerical Example for a Three-Channel System

Let us consider a numerical example for a three-channel system which may be typical of a space communication link. The nominal data rate bandwidths, formats and power divisions may be as shown in Table 1.

Table 1. Bandwidth, Data Formats, and Power Assignments for the Numerical Sample

Channel	Bandwidth (MHz)	Data Format	Power Assignment (%)
1	100	Biphase, baseband	80
2	4	Bi-ϕ-L, baseband	18
3	0.4	Bi-ϕ-L, subcarrier	2

Note that, in this example, we arbitrarily have assigned more power to the secondary channels (2 and 3) than would be required if power division had been carried out based on equal error rates (same E_b/N_0) in all three channels. Also, to minimize the cross-talk between channels 2 and 3, the baseband data of the latter channel is assumed to be riding on a squarewave subcarrier whose frequency is well above the high end of channel 2 data baseband spectrum. On the other hand, the frequency of the channel 3 subcarrier should be low enough to allow at least the 3rd and 5th harmonics to pass through the RF channel without significant attenuation.

2.6.1 Interplex Implementation

Assuming that a three-channel interplex system is used, the modulation angle assignments are as follows:

$$\theta_1 = 90^\circ\ (\pm)$$
$$\theta_2 = \tan^{-1}\sqrt{P_2/P_1} = \tan^{-1}\sqrt{18/80} \cong \tan^{-1}(0.47) \cong 25.4^\circ\ (\pm)$$
$$\theta_3 = \tan^{-1}\sqrt{P_3/P_1} = \tan^{-1}\sqrt{2/80} \cong \tan^{-1}(0.16) \cong 9^\circ\ (\pm) \qquad (55)$$

From these angle assignments, the data-to-crosstalk ratios $D_k(\ell, m)$ as a function of carrier tracking error ϕ can be computed using (47) through (53). These are plotted in Figure 10. From the plots, it is evident that $D_1(2)$, $D_2(1)$, and $D_3(1)$ are affected most by high values of ϕ. But, because the

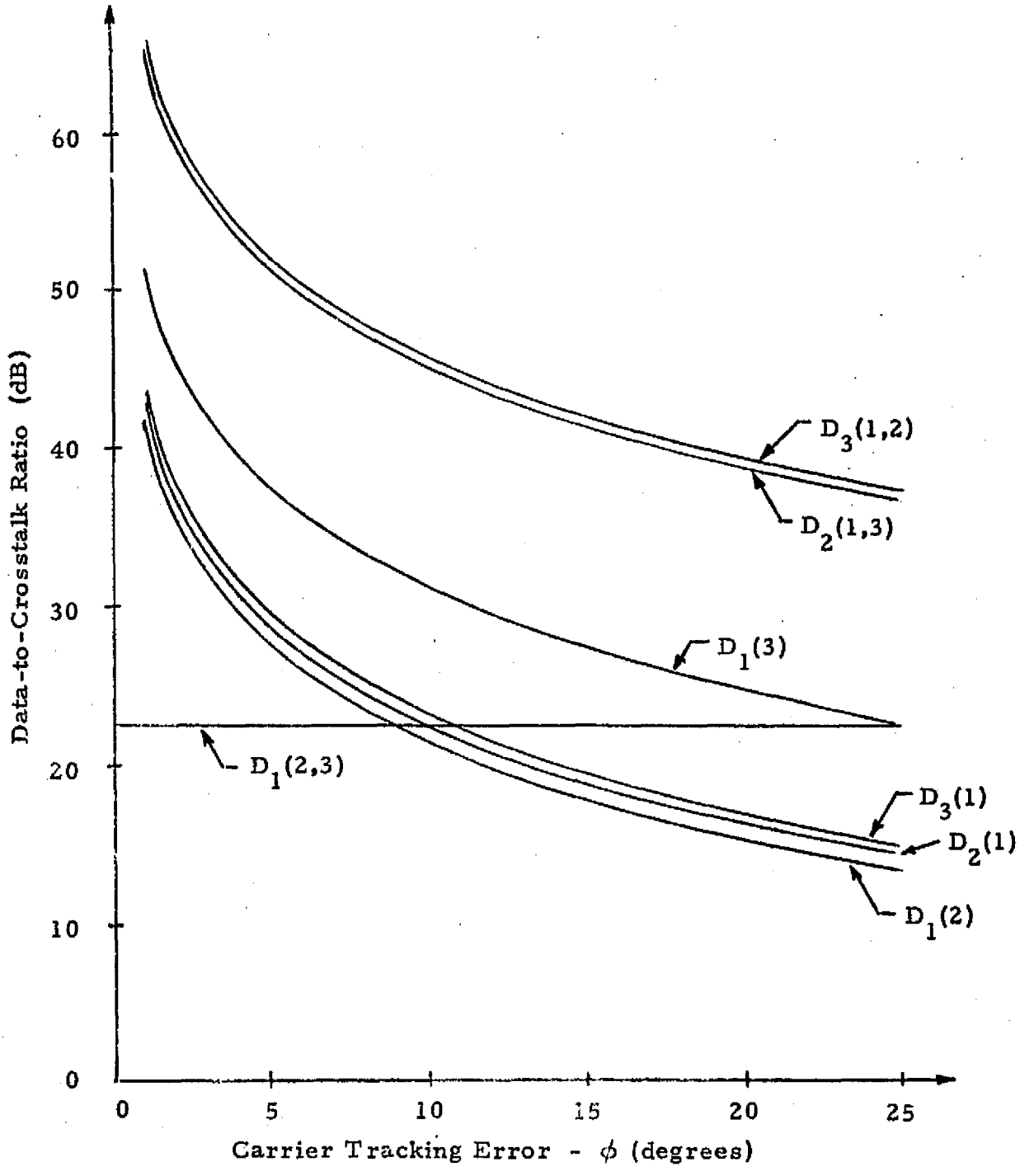


Figure 10. Data-to-Crosstalk Ratios $D_k(\ell, m)$ vs. Carrier Tracking Error ϕ

channel 1 data rate is much higher than the channel 2 data rate, $D_1(2)$ in this case is the ratio of the $S_1(t)$ signal power to the time-varying bias due to $S_2(t)$ data. The same is true of $D_1(3)$. Consequently, at least in principle, both of these ratios can be improved by the use of coarse estimates of $S_2(t)$ and $S_3(t)$. $D_2(1)$ and $D_3(1)$, on the other hand, are determined by filtered $S_1(t)$ data stream and thus their crosstalk noise floor is a fluctuating signal with statistics approaching that of a Gaussian process. Thus, these ratios can be improved by optimal filtering of $S_2(t)$ and $S_3(t)$ data streams. Of course, once bit synchronization is established, such optimum filtering will be the integrate-and-dump detectors for the $S_2(t)$ and $S_3(t)$ data streams.

2.6.2 Hybrid Modulator Implementation

To achieve the same power division between the three channels as was assumed for the preceding example for interplex, the relative amplitudes of the respective data streams will be as follows:

$$\left.\begin{aligned} A' &= 1.0 \\ B' &= 0.474 \\ C' &= 0.158 \end{aligned}\right\} \text{ normalized with respect to } A' \qquad (56)$$

Since we are primarily interested in comparing the third-order terms (i.e., the terms independent of carrier tracking error ϕ, we compute the ΔA_{ave} term by substituting values of (56) into (28). This results in

$$\Delta A_{ave} = 0.081 . \qquad (57)$$

The equivalent $D_A(B, C)$ ratio for the hybrid implementation is therefore

$$\begin{aligned} D_A(B, C) &= 20 \log_{10} (1/\Delta A_{ave}) \\ &= 20 \log_{10} (1/0.081) = 21.8 \text{ dB} . \end{aligned} \qquad (58)$$

The corresponding term, i.e., $D_1(2, 3)$ for the three-channel interplex system is 22.5 dB, which is only 0.7 dB above the value shown in (58). Considering the absolute values of the $D_A(B, C)$ and $D_1(2, 3)$ ratios (≈ 20 dB),

this is a small difference in view of potential implementation simplicity of the hybrid modulator/multiplexer.

3.0 CONCLUSIONS

It has been demonstrated that either the interplex or the hybrid phase-multiplexing technique for transmission of three independent and asynchronous digital data channels offers a significant performance advantage over a conventional PM subcarrier design. The hybrid phase-multiplexing technique differs from interplex primarily in terms of implementation. Whereas interplex involves data channel combining at baseband and utilizes a linear phase modulator at IF, the hybrid technique involves data channel combining at IF and utilizes quadrature carrier modulation. Since the performance characteristics of these two approaches are essentially identical, it appears that the selection of a technique for any particular application should be dictated only by equipment design considerations.

REFERENCES

1. S. Butman and U. Timor. "Interplex—An Efficient PSK/PM Telemetry System," IEEE Transactions on Communications, Vol. COM-20, No. 3, June 1972, pp. 415-419.

2. J. LaFrieda. "Optimum Modulation Indexes and Maximum Data Rates for the Interplex Modem," JPL Space Programs Summary 37-64, Vol. III, June 1970, pp. 23-27.

3. U. Timor and S. Butman. "Suppressed-Carrier Two-Channel Interplex Modulation System," JPL Space Programs Summary 37-64, Vol. III, June 1970, pp. 27-31.

4. S. Butman and U. Timor. "Interplex—An Efficient Two-Channel Telemetry System for Space Exploration," JPL Space Programs Summary 37-62, Vol. III, February 1970.

5. S. Udalov. "Feasibility Study of an Interplex Modulation System for the Orbiter's Ku-Band Link" (under NASA Contract No. NAS 9-13467), Axiomatix Report No. R7410-5, October 7, 1974.

6. R. W. Allen and B. H. Batson. "A Variable-Data-Rate, Multi-Mode Quadriphase Modem," Proc. Int. Comm. Conf., 1973.

7. S. Udalov. "Multiplexing Modulation Formats for the Orbiter's Ku-Band Return Link" (under NASA Contract No. NAS 9-13467), Axiomatix Report No. R7502-1, February 13, 1975.

8. C. L. Weber. "Candidate Receivers for Unbalanced QPSK," ITC, September 1976, Los Angeles, California.

APPENDIX B

FEASIBILITY OF CANDIDATE RECEIVERS FOR KU-BAND COMMUNICATION SIGNALS TO AND FROM SHUTTLE

Charles L. Weber

APPENDIX B

FEASIBILITY OF CANDIDATE RECEIVERS FOR KU-BAND COMMUNICATION SIGNALS TO AND FROM SHUTTLE

Charles L. Weber

1.0 INTRODUCTION

Signal waveforms for the Orbiter Ku-band relay communication links have been described in [3], along with the link power budgets. Data multiplexing, encoding, and modulation schemes for both the forward and return links have also been described in general terms.

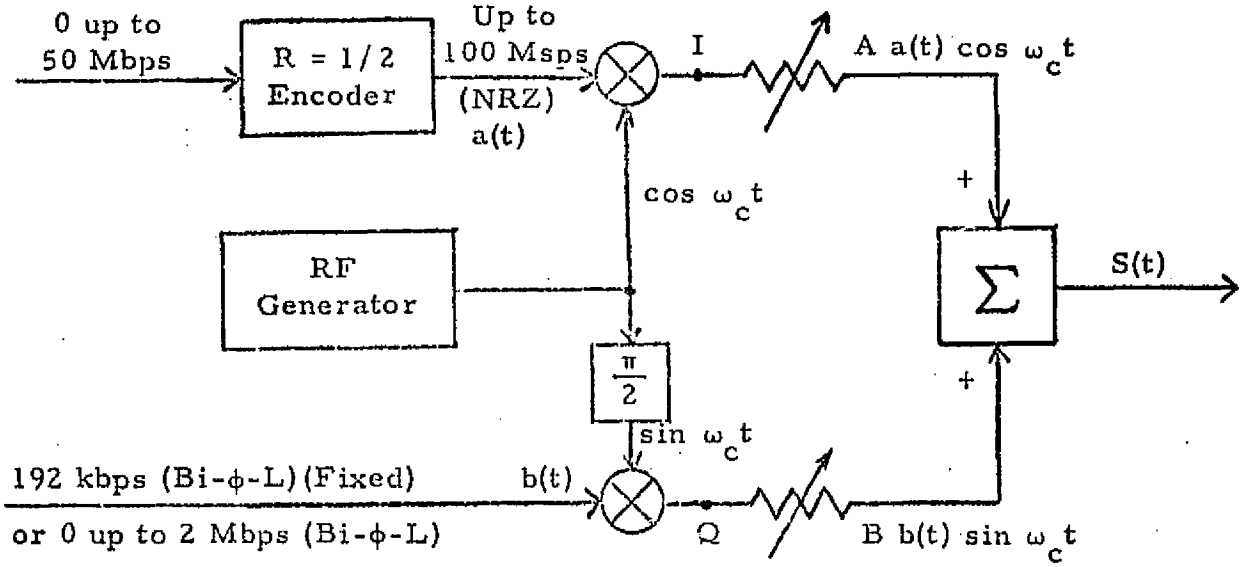
We consider various candidate receivers for the unbalanced QPSK signal format. The basic modulator configuration is shown in Figure 1 for both the return and forward links. In all cases the transmitted signal is

$$S(t) = A\, a(t) \cos \omega_c t + B\, b(t) \sin \omega_c t \qquad (1)$$

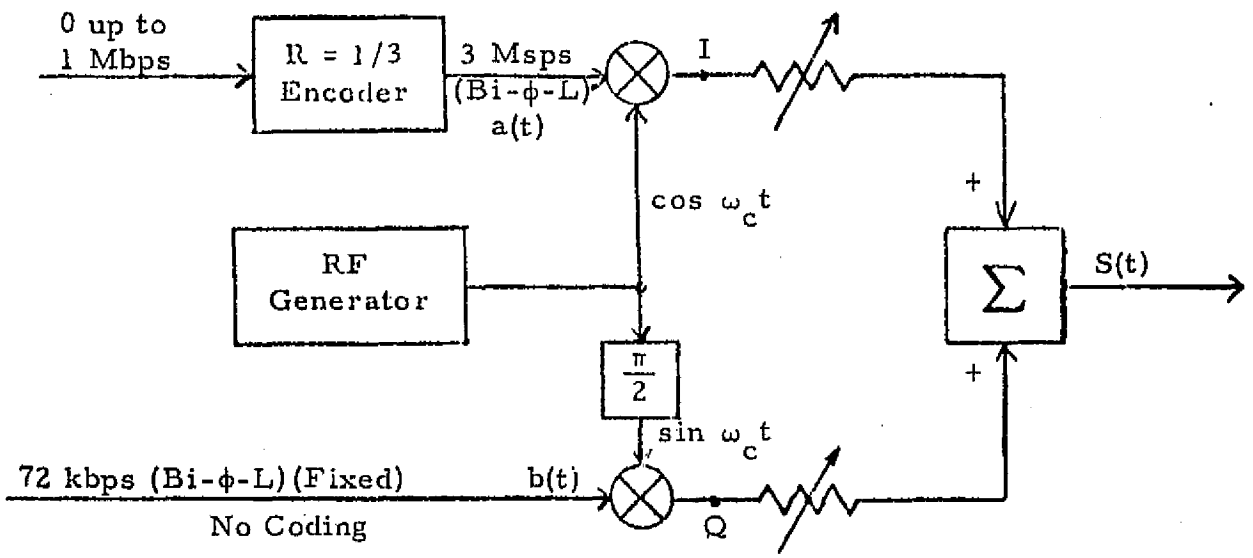
where a(t) represents the encoded data of the high rate, or I, channel and b(t) represents the encoded data in the low rate, or Q, channel.

Referring to Figure 1a, in the I channel, a(t) is a sequence of ±1's of NRZ data which has symbol rate up to 100 Msps. It represents rate R = 1/2 convolutionally encoded data which has variable rate from zero up to 50 Mbps. In the Q channel, b(t) is either at the fixed rate of 192 kbps of Bi-ϕ-L signals, or it is variable rate from zero to 2 Mbps of Bi-ϕ-L signals. The data rates for the two links are completely independent and may be either synchronous or asynchronous.

The proposed power division for the return link is 80 percent and 20 percent for the high data rate and low data rate links, respectively. With this power division ratio of 4:1, the power available to the high data rate link will be reduced by about 1 dB with respect to the total power. The purpose of unequal power ratio is to cause the symbol energy, and therefore the error rates, in the two channels to be closer to the same value.



(a) Return Link



(b) Forward Link

Figure 1. Unbalanced Quadriphase Modulators

The present design calls for a fixed 4:1 power ratio in both the forward and return links. From an error rate point of view, this overpowers the low data rate channels. However, the maximum to be gained by placing all the power in the high data rate channel is only 0.97 dB. This power ratio of 4:1 is considered satisfactory, and we shall emphasize this power ratio throughout the present study.

In Section 2.0, the fourth-power receiver tracking loop is considered. When the power division in the unbalanced QPSK signal is 4:1, it is shown that this receiver tracking loop is an unattractive choice.

In Section 3.0, the Costas loop is evaluated as a candidate receiver for both the forward and return links. In this appendix, we restrict attention to design of the Costas loop receiver which is tracking primarily on the high data rate signal. In sections 3.1 and 3.2, it is shown that the Costas loop quadriphase demodulator in Figure 6 performs with an RMS phase error of less than 10 degrees over a variety of tracking loop bandwidths and over a variety of rates in both the high and low data rate channels. The performance curves are shown in Figure 7, where it is seen that this performance is attained for a power division of 70/30 up to 100/0.

In sections 3.3 and 3.4, similar results are obtained for the forward link Costas loop shown in Figure 5.

In section 3.5, the return link Costas loop is examined when the rate in the high data rate link goes to zero, but the power remains on. The tracking is then carried out primarily on a monochromatic tone and, as a result, performance improves with respect to that reported in the earlier sections of 3.0.

It should also be noted that, if the power in the high data rate link is turned off, tracking is then carried out only on the low data rate link. Without any alterations, the Costas loops in Figures 5 and 6 will shift their reference by $\pi/4$ radians, thereby sending the low data rate signal into the high data rate demodulator. This problem can be easily overcome, but a better solution may be to not remove the power in either of the data links without removing both.

An additional candidate receiver is the four-phase Costas loop receiver [4]. This receiver requires significantly more equipment for implementation with little anticipated improvement in performance. This receiver would also exhibit phase ambiguities of multiples of $\pi/4$ radians, so that an algorithm would have to be developed to process the proper data links into the appropriate demodulator. For these reasons, this receiver was not considered in this feasibility study.

In the last section, some approximate error rate computations are presented that show that performance is within a few tenths of a dB of ideal performance.

2.0 FOURTH-POWER RECEIVER TRACKING LOOP

A block diagram of the fourth-power tracking loop is shown in Figure 2. The signal at the input to the fourth-power device is

$$y(t) = s(t) + n(t) , \tag{2}$$

when s(t) is given by (1) and the narrowband noise n(t) is given by

$$n(t) = \sqrt{2}\,[n_c(t) \cos \omega_c t + n_s(t) \sin \omega_c t] . \tag{3}$$

The independent baseband noises $n_c(t)$ and $n_s(t)$ each have one-sided power spectral density equal to that of n(t), namely, N_0 watts/Hz. The bandwidth of $n_c(t)$ and $n_s(t)$ is $B_{IF}/2$, where B_{IF} is the bandwidth of the IF filter H(p), as shown in Figure 3.

The difficulties associated with the fourth-power tracking loop can be demonstrated without considering the additive noise when the power division is 4:1.

With signal only, therefore, the output of the fourth-power device in Figure 2 is

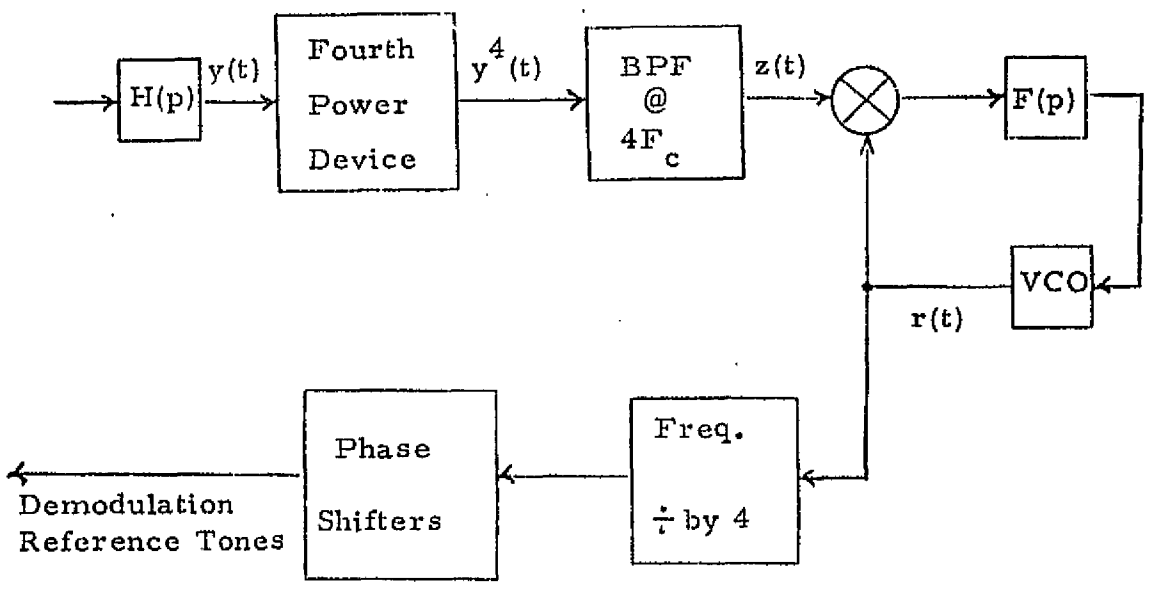


Figure 2. Fourth-Power Receiver Tracking Loop

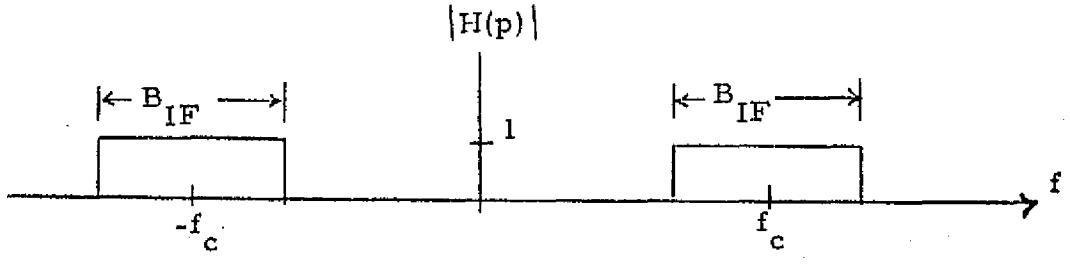


Figure 3. IF Filter Transfer Function

$$y^4(t) = s^4(t)$$

$$= A^4 \cos^4 \omega_c t + 4A^3B\, a(t)b(t) \cos^3 \omega_c t \sin \omega_c t$$

$$+ 6A^2B^2 \cos^2 \omega_c t \sin^2 \omega_c t + 4AB^3\, a(t)b(t) \cos \omega_c t \sin^3 \omega_c t$$

$$+ B^4 \sin^4 \omega_c t$$

$$= \left(\frac{A^4+B^4}{8}\right) \cos (4\omega_c t) + \left(\frac{A^4-B^4}{8}\right) \cos (2\omega_c t) + 3\left(\frac{A^4+B^4}{8}\right)$$

$$+ \frac{1}{2} a(t)b(t)\{A^3B[\sin (4\omega_c t) - \sin (2\omega_c t)] - AB^3[\sin (4\omega_c t) + \sin (2\omega_c t)]\}$$

$$+ \frac{3}{2}(A^3B + AB^3) \sin (2\omega_c t) + \frac{3}{2} A^2B^2\left[1 - \frac{1}{2}\left(1 + \cos (4\omega_c t)\right)\right] . \qquad (4)$$

After passing the BPF centered at $4f_c$ in Figure 2, the result is

$$z(t) = \left[\frac{A^4 + B^4 - 6A^2B^2}{8}\right] \cos (4\omega_c t) + \frac{a(t)b(t)}{2}(A^3B - AB^3) \sin (4\omega_c t) . \qquad (5)$$

The first term in (5) is the tracking term, and the second term is a self-noise term. In terms of the average power in each link, namely,

$$P_A = A^2/2 , \qquad P_B = B^2/2 , \qquad (6)$$

z(t) can be expressed as

$$z(t) = \frac{1}{2}[P_A^2 + P_B^2 - 6P_AP_B] \cos (4\omega_c t) + 2\, a(t)b(t) \sqrt{P_A P_B}\, (P_A - P_B) \sin (4\omega_c t) . \qquad (7)$$

The tracking term contributes a spectral component at $4\omega_c t$. The self-noise term contributes a spectral density at $4\omega_c t$. The level of this spectral density within the receiver tracking loop is dependent on whether the forward or return link is considered, since the waveforms are different.

The amplitude of the tracking term is the coefficient of cos $(4\omega_c t)$ in (7). The power in that term is

$$P_{track} = \frac{1}{8}[P_A^2 + P_B^2 - 6P_A P_B]^2 . \tag{8}$$

The self-noise power in the tracking loop is the equal to the spectral density of the second term in (7) at $4f_c$ times the tracking loop bandwidth. This self-noise power can be expressed as

$$P_{SN} = 4P_A P_B (P_A - P_B)^2 S(0) B_L , \tag{9}$$

where B_L is the tracking loop noise bandwidth and S(0) is the two-sided spectral density of the signal a(t) b(t) at f = 0. In Addendum A, we show the spectral density of a(t) b(t) at f = 0 is

$$S(0) = T_A \left(1 - \frac{T_A}{T_B}\right) \tag{10}$$

when a(t) is NRZ and b(t) is Bi-ϕ-L, and

$$S(0) = \frac{T_A^2}{2T_B} \tag{11}$$

when a(t) and b(t) are both Bi-ϕ-L. In (10) and (11), T_A and T_B are the symbol times of a(t) and b(t), respectively.

In order to obtain a preliminary evaluation of the performance of the fourth-power tracking loop, let the total power be designated as P_T, so that

$$\begin{aligned} P_A &= \alpha P_T \\ P_B &= (1 - \alpha) P_T . \end{aligned} \tag{12}$$

Upon substitution of (12) into (8) and (9), the ratio of tracking power to self-noise power within the tracking loop can be written as

$$\frac{P_{track}}{P_{SN}} = \frac{1}{32} \frac{[P_A^2 + P_B^2 - 6P_A P_B]^2}{P_A P_B [P_A - P_B]^2 B_L 2S(0)}$$

$$= \frac{1}{32} \frac{N}{D[2S(0)]} \tag{13}$$

where

$$N = [\alpha^2 + (1-\alpha)^2 - 6\alpha(1-\alpha)]^2 \tag{14}$$

and

$$D = \alpha(1-\alpha)[\alpha - (1-\alpha)]^2 . \tag{15}$$

The parameter α represents the fraction of total power in the high data rate link, which assumes values over $0.5 \leq \alpha \leq 1$. The numerator N in (13) and (14) represents the variation in tracking power as a function of the power division and is shown in Figure 4 as the dashed line. The denominator D represents the variation in self-noise spectral density as a function of the power division. The denominator is also shown in Figure 4, where it can be seen that D reaches a maximum and N has a minimum of 0 near $\alpha = 0.85$.

The signal power to self-noise spectral density,

$$\frac{N}{32D} = B_L 2S(0) \frac{P_{track}}{P_{SN}}, \tag{16}$$

is also shown in Figure 4, where it is seen that the worst power division values are within [0.75, 0.95]. We see that the fourth power tracking loop makes very ineffective use of the total available tracking power when the power division is in the vicinity of 0.8. An additional disadvantage of the fourth-power tracking loop is that the stable lock points occur at multiples of $\pi/4$, thereby allowing the possibility that the two data channels become interchanged. This could be detected and corrected in the system design, but only at the expense of additional complexity.

With these results, we conclude that the fourth-power tracking loop should not be recommended for the demodulation of an unbalanced 4:1 QPSK signal.

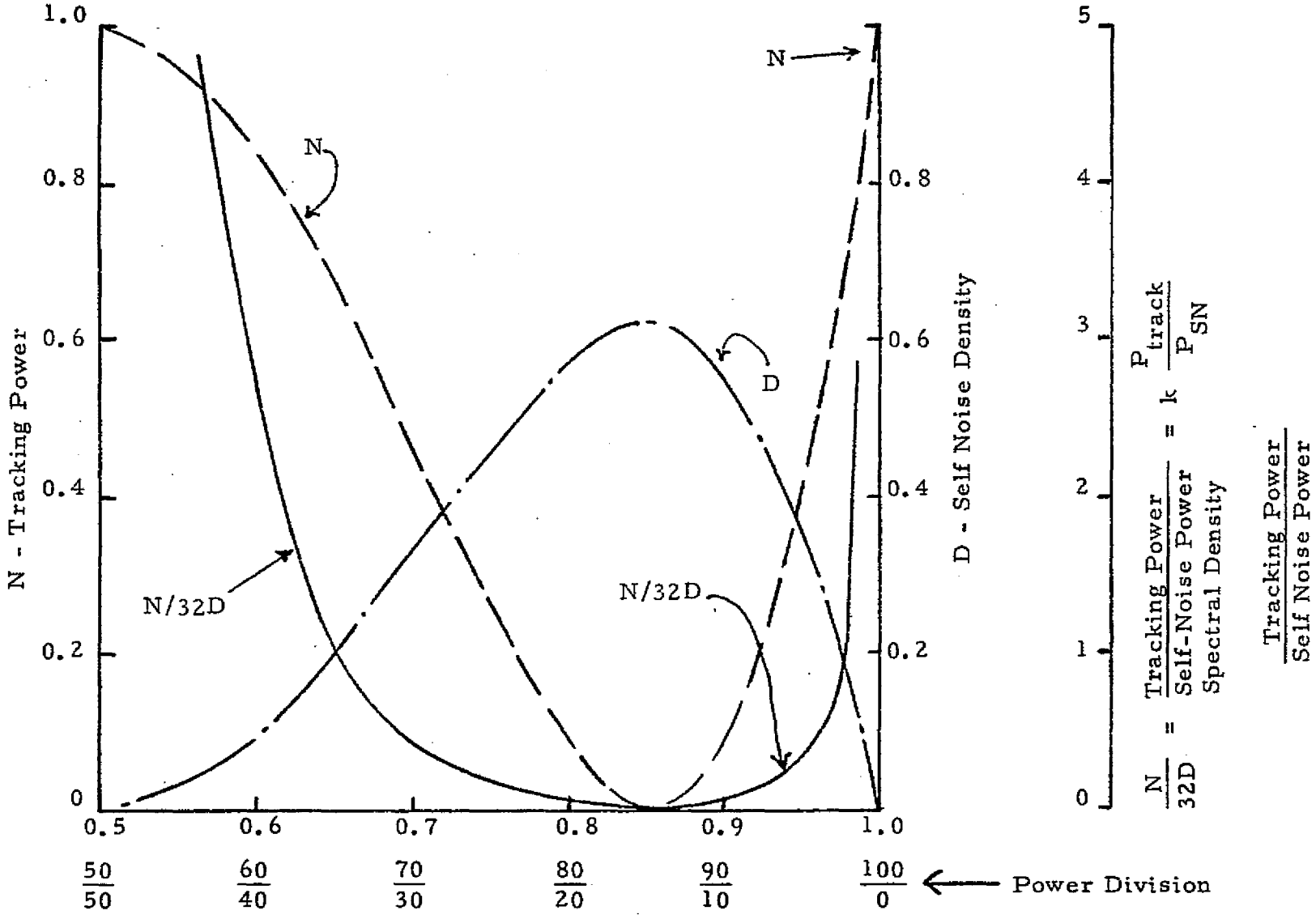


Figure 4. Tracking Power, Self Noise Density and Ratio of Tracking Power to Self Noise Density for the Fourth Power Tracking Loop.

3.0 COSTAS RECEIVER

In this section, we consider the Costas receiver for demodulating an unbalanced QPSK signal. Block diagrams of the receiver as shown in Figures 5 and 6 for the forward and return links, respectively.

The received signal is as described in (1) and (2), where we elect to model the noise [6, 7] as

$$n(t) = \sqrt{2}\,[n_c(t)\cos(\omega_c t + \phi) + n_s(t)\sin(\omega_c t + \phi)] \tag{17}$$

where $n_c(t)$ and $n_s(t)$ have the same statistics as described in (3).

We initially develop a generalized description of the receiver operation, then consider several specific cases.

Referring to Figures 5 and 6, the reference signals are given by

$$r(t) = \sqrt{2}\cos(\omega_c t + \phi) \tag{18}$$

and

$$r'(t) = \sqrt{2}\sin(\omega_c t + \phi)\,. \tag{19}$$

Under the assumption that the IF filter preceding the Costas loop receiver is sufficiently wide to pass both channels of the received signal without distortion, then the baseband outputs, $e_c(t)$ and $e_s(t)$, of the mixers, neglecting the $2f_c$ terms, are given by

$$e_c(t) \triangleq y(t)\,r(t)\Big|_{\text{video}}$$

$$= (\sqrt{2}/2)[A\,a(t)\cos\phi - B\,b(t)\sin\phi] + n_c(t) \tag{20}$$

and

$$e_s(t) \triangleq y(t)\,r'(t)\Big|_{\text{video}}$$

$$= (\sqrt{2}/2)[A\,a(t)\sin\phi + B\,b(t)\cos\phi] + n_s(t)\,. \tag{21}$$

3.1 The Return Link Costas Loop Receiver

In considering the return link, we first consider the bandwidth of the LPF to be

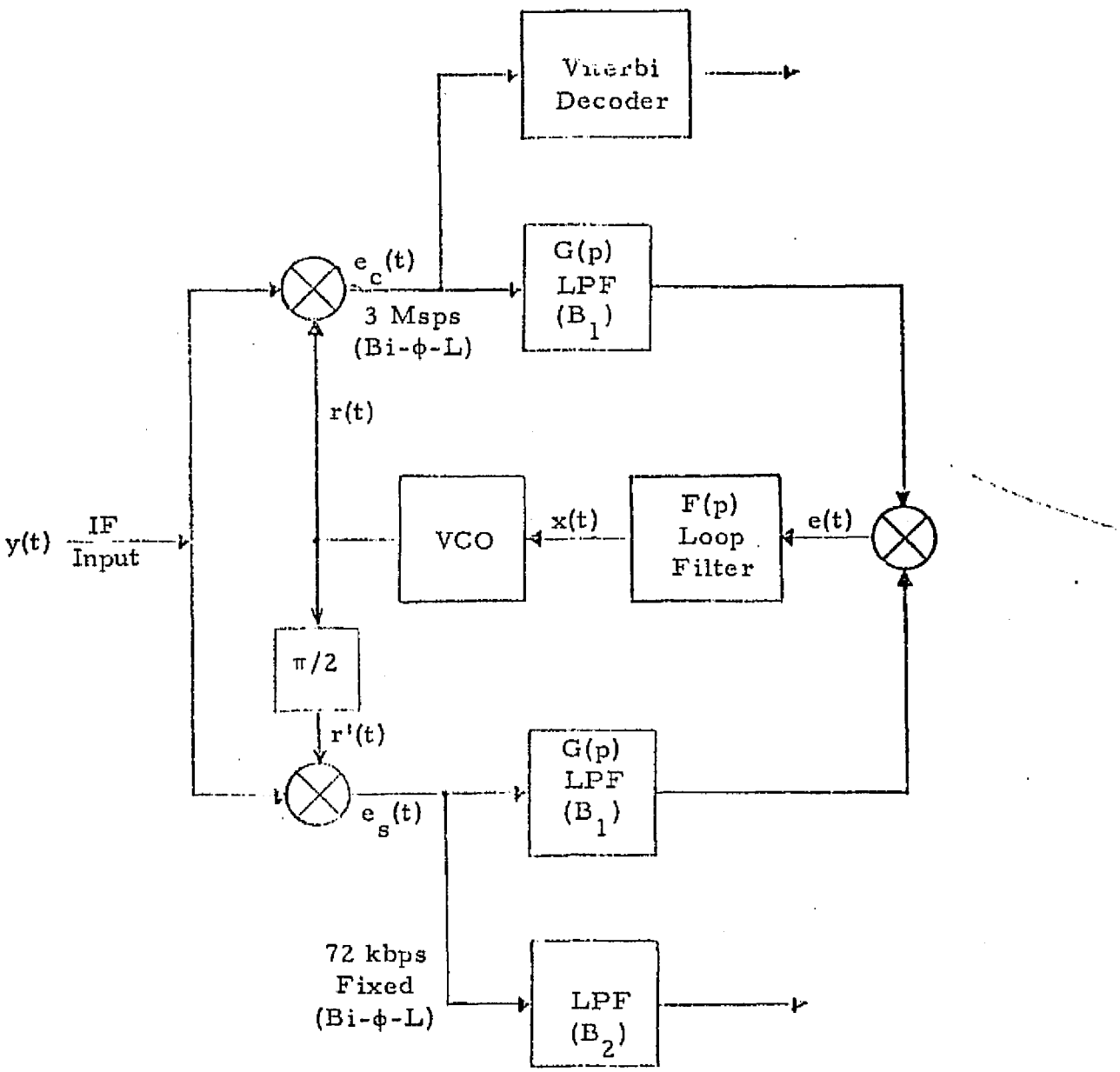


Figure 5. Costas Loop Quadriphase Demodulator for the Forward Link

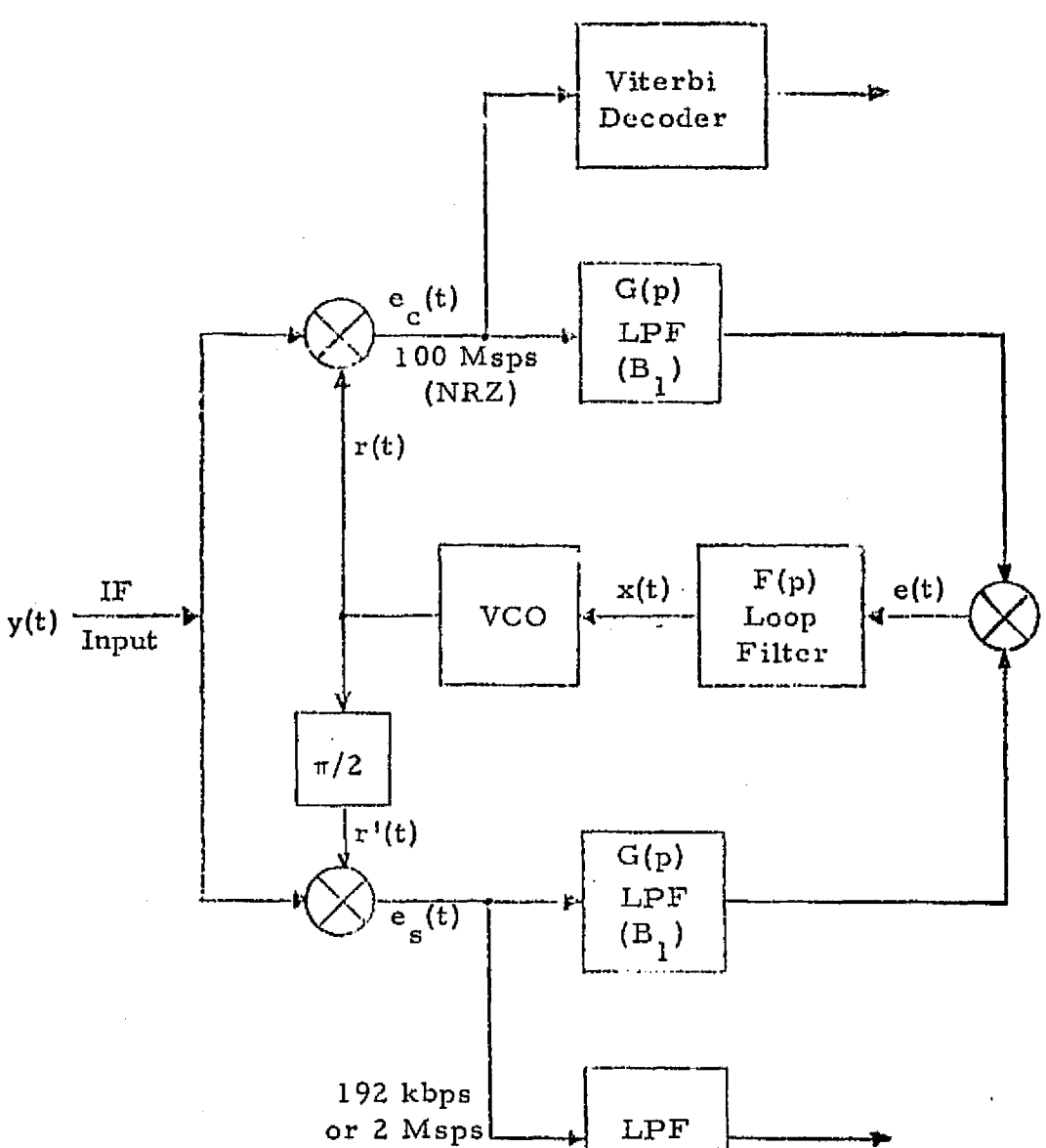


Figure 6. Costas Loop Quadriphase Demodulator for the Return Link

$$B_1 = 100 \text{ MHz}, \tag{22}$$

thereby allowing the passage of the signal portions of $e_c(t)$ and $e_s(t)$ without significant distortion. This is equivalent to saying the tracking loop is going to function primarily on the high-rate, high-powered channel. The dynamic error signal e(t) is then given by

$$\begin{aligned} e(t) &\triangleq e_c(t)\, e_s(t) \\ &= \left(\frac{P_A - P_B}{2}\right) \sin 2\phi \quad \left.\begin{matrix}\text{Tracking} \\ \text{Signal}\end{matrix}\right\} \\ &\quad + \sqrt{P_A P_B}\, a(t) b(t) \cos 2\phi \quad \left.\begin{matrix}\text{Self-Noise}\end{matrix}\right\} \\ &\quad + n_c(t)\, n_s(t) \quad \left.\begin{matrix}\text{Noise}\end{matrix}\right\} \\ &\quad + n_c(t)\left[\sqrt{P_A}\, a(t) \sin\phi + \sqrt{P_B}\, b(t) \cos\phi\right] \\ &\quad + n_s(t)\left[\sqrt{P_A}\, a(t) \cos\phi - \sqrt{P_B}\, b(t) \sin\phi\right] \quad \left.\begin{matrix}\text{S x N}\end{matrix}\right\} \end{aligned} \Bigg\} \triangleq n_{eq}(t) \tag{23}$$

where, in addition to the desired tracking signal, self-noise, noise, and (signal x noise) terms are generated, which we designate as the equivalent loop noise, $n_{eq}(t)$.

Observation of the tracking signal $[(P_A - P_B)/2] \sin 2\phi$ in (23) provides the following:

(a) The stable lock points are at $\phi = k\pi$, with the result that the two data links cannot be sent to the incorrect data demodulator. The high data rate link will always go to the high data rate demodulator, and the same for the low data rate link. This assumes that the power in both channels is always present, which is the most likely case.

(b) Although it is expected that the power in both channels will always be "on" even if the data rates are reduced to zero, should the power in the high data rate channel fail, i.e., $P_A \to 0$, then the sign of the tracking signal is reversed and the stable lock angles are at $\pi/2 \pm k\pi$. If this

occurs, then the low data rate link will be processed into the high data rate demodulator.

(c) If the power in the low data rate link goes to zero, i.e., $P_B \to 0$, the stable lock angles remain unchanged, so that the high data rate link continues to be processed into the high data rate demodulator.

The statistics of the equivalent noise, $n_{eq}(t)$, are obtained by assuming that the tracking loop bandwidth is much smaller than B_1. The phase error $\psi(t)$ will then be slowly varying with respect to the other random processes present in e(t). The first and second conditional moments are given by

$$E[n_{eq}(t)|\phi] = 0 \tag{24}$$

and

$$\begin{aligned} R_{n_{eq}}(\tau) &= E[n_{eq}(t)\, n_{eq}(t+\tau)|\phi] \\ &= P_A P_B R_a(\tau) R_b(\tau) \cos^2(2\phi) \\ &\quad + R_N^2(\tau) + R_N(\tau)[P_A R_a(\tau) + P_B R_b(\tau)], \end{aligned} \tag{25}$$

where $R_a(\tau)$ and $R_b(\tau)$ are the autocorrelation functions of a(t) and b(t), respectively, as discussed in Addendum A.

The spectral density of $n_{eq}(t)$ is essentially flat over the bandwidth of the tracking loop, so that the significant statistic when considering tracking loop performance is the spectral density at f = 0. The spectral density of the terms in (25) can be shown [5] to be

$$R_N^2(\tau) \to S(0) = \int_{-B_1}^{B_1} (N_0/2)^2 \, df = \frac{N_0^2 B_1}{2} \tag{26}$$

$$R_N(\tau)[P_A R_a(\tau) + P_B R_b(\tau)] \to S(0) \approx \frac{N_0}{2}(P_A + P_B) \tag{27}$$

$$P_A P_B R_a(\tau) R_b(\tau) \cos^2(2\phi) \to S(0) \approx P_A P_B T_A \left(1 - \frac{T_A}{T_B}\right) \tag{28}$$

where, in (28), we have imposed the upper bound, $\cos^2(2\phi) \leq 1$. In addition, we have used the fact that the spectral density of $R_a(\tau)\, R_b(\tau)$ at $f = 0$ is given by $N_0' = 2\,T_A[1 - (T_A/T_B)]$ for the return link which consists of NRZ waveforms in the high data rate channel and Bi-ϕ-L waveforms in the low data rate link. This result is derived in Part I of Addendum A.

Combining these results, the two-sided spectral density of $n_{eq}(t)$ is

$$S_{n_{eq}}(0) \triangleq \frac{N_{0_{eq}}}{2} = \frac{N_0^2\, B_1}{2} + \frac{N_0}{2}(P_A + P_B) + P_A\, P_B\, T_A \left(1 - \frac{T_A}{T_B}\right) \tag{29}$$

We shall base our performance analyses on a linearized assumption. We shall see that, at the signal-to-noise ratios which are anticipated, the Costas loop is well within its linearized region. When such is the case, the variance of the phase error is given by [7]:

$$\sigma_\phi^2 = \frac{1}{4}\,\sigma_{2\phi}^2 = \frac{1}{4\rho}\,(\mathrm{rad}_1)^2 \tag{30}$$

where ρ is the signal-to-noise ratio in the Costas loop, which is given by

$$\rho = \frac{P_c}{N_{0_{eq}}\, B_L} \tag{31}$$

where P_c is the available power in watts for tracking, B_L is the one-sided equivalent noise bandwidth of the tracking loop in Hz, and $N_{0_{eq}}$ is the one-sided spectral density of the equivalent additive noise in the loop in watts/Hz.

In our receiver, $N_{0_{eq}}$ is given by (29), and from (23):

$$P_c = (P_A - P_B)^2/4 \ . \tag{32}$$

Upon substitution of (29) and (32) into (31), the Costas loop signal-to noise ratio is given by

$$\rho = \frac{(P_A - P_B)^2/4}{N_0 B_L \left[N_0 B_1 + P_A + P_B + \dfrac{2\,P_A\, P_B}{N_0}\, T_A \left(1 - \dfrac{T_A}{T_B}\right)\right]} \ . \tag{33}$$

Using the power division parameter α, as given by (12),

$$\rho = \frac{P_T}{4N_0B_L}\left[\frac{(2\alpha - 1)^2}{1 + \frac{N_0B_1}{P_T} + \frac{2P_T\,\alpha(1-\alpha)}{N_0}T_A\left(1 - \frac{T_A}{T_B}\right)}\right]. \qquad (34)$$

The performance of the return link receiver using a Costas tracking loop and tracking on the high-powered channel is given by (30) and (34).

3.2 Performance Computation - Return Link

The performance of the return link Costas loop using the high data rate channel is shown in Figure 7. The standard deviation of the phase error is shown versus the power division parameter α. Two sets of curves are shown. Set (1) shows performance for the symbol rates R_A = 384 ksps and R_B = 192 ksps. This is a worst set of rates from the point of view of maximizing the standard deviation of the phase error. As shown, the two curves are for loop bandwidths of 5 kHz and 10 kHz. As $\alpha \rightarrow 0.5$, the Costas loop is unable to track and $\sigma \rightarrow \infty$. As $\alpha \rightarrow 1$, the signal becomes single channel, the high data rate channel, and the self-noise goes to zero. The effect of the self-noise can be determined by comparing the standard deviation at some α to that at $\alpha = 1$. The increase is due to the effects of the self-noise, and the fact that the available tracking power is also decreasing.

The second set of curves in Figure 7 is for the best set of rates in the return link. The improvement due to the resulting reduction in self-noise at the higher symbol rates is seen to be approximately an order of magnitude in σ. The points on the curves corresponding to $\alpha = 0.8$, a 4-to-1 power division, have been emphasized by dots, since this is the most likely choice of power division.

A summary of the various parameters used in the forward link and return link computations is shown in Table I.

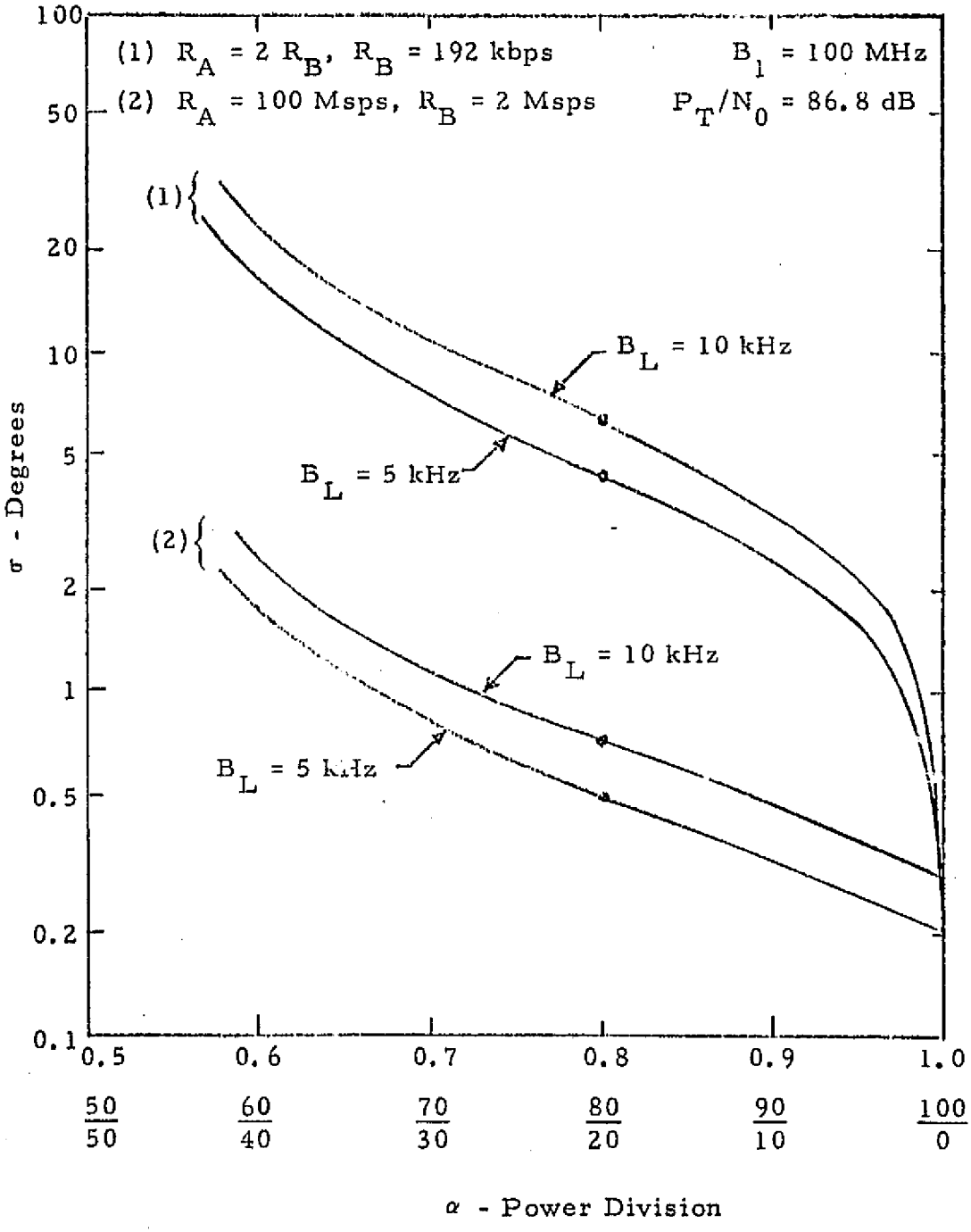


Figure 7. RMS Phase Error Versus Power Division for Unbalanced QPSK for the Return Link

TABLE I

FORWARD AND RETURN LINK PARAMETERS

	Forward Link	Return Link
Channel A	0 to 1 Mbps R = 1/3 0 to 3 Msps Bi-ϕ-L	0 to 50 Mbps R = 1/2 0 to 100 Msps NRZ
Channel B	72 kbps Fixed Bi-ϕ-L No coding	192 kbps Fixed or 2 Mbps Bi-ϕ-L
One-Sided Self-Noise Spectral Density at f = 0	$N_0' = T_A^2/T_B$ Bi-ϕ-L and Bi-ϕ-L	$N_0' = 2T_A\left(1 - \frac{T_A}{T_B}\right)$ Bi-ϕ-L and NRZ
P_T/N_0	73.8 dB-Hz [+]	86.8 dB-Hz [+*]
When Tracking on Channel A	B_l = 6 MHz (Curves 3 and 4)	B_l = 100 MHz (Curves 1 and 2)
When Tracking on Channel B	B_l = 72 kHz	B_l = 4 MHz

*Includes all cases except symbol sync.

+Reference B. H. Batson power budgets.

3.3 Forward Link Costas Loop Receiver

We now consider the forward link Costas loop receiver. The development is similar, with the following exceptions:

(a) The two data links are both Bi-ϕ-L waveforms, with the result that the self-noise spectral density at f = 0 is given by

$$N_0' = T_A^2/T_B , \tag{35}$$

as given in Table I and derived in Addendum A, Part II.

(b) The IF bandwidths are set at B_1 = 6 MHz, which will pass the signal essentially undistorted at all anticipated rates.

(c) The carrier power-to-noise spectral density is 73.8 dB-Hz, as shown in Table I.

The performance is still given by (30) and (31), but the equivalent spectral density is now given by

$$S_{n_{eq}}(0) = \frac{N_{0_{eq}}}{2} = \frac{N_0^2 B_1}{2} + \frac{N_0}{2}(P_A + P_B) + \frac{P_A P_B T_A^2}{2 T_B} . \tag{36}$$

The resulting signal-to-noise ratio in the tracking loop bandwidth is then given by

$$\rho = \frac{P_T}{4 N_0 B_L}\left[\frac{(2\alpha - 1)^2}{1 + \frac{N_0 B_1}{P_T} + \frac{P_T T_A^2}{N_0 T_B}\alpha(1-\alpha)}\right] . \tag{37}$$

The performance of the Costas loop for the forward link employing the high data rate link is then given by (37) and (30).

3.4 Performance Computation - Forward Link

The performance of the forward link Costas loop using the high data rate link for tracking is shown in Figure 8. The RMS phase error in degrees is again shown versus the power division for the unbalanced QPSK signal.

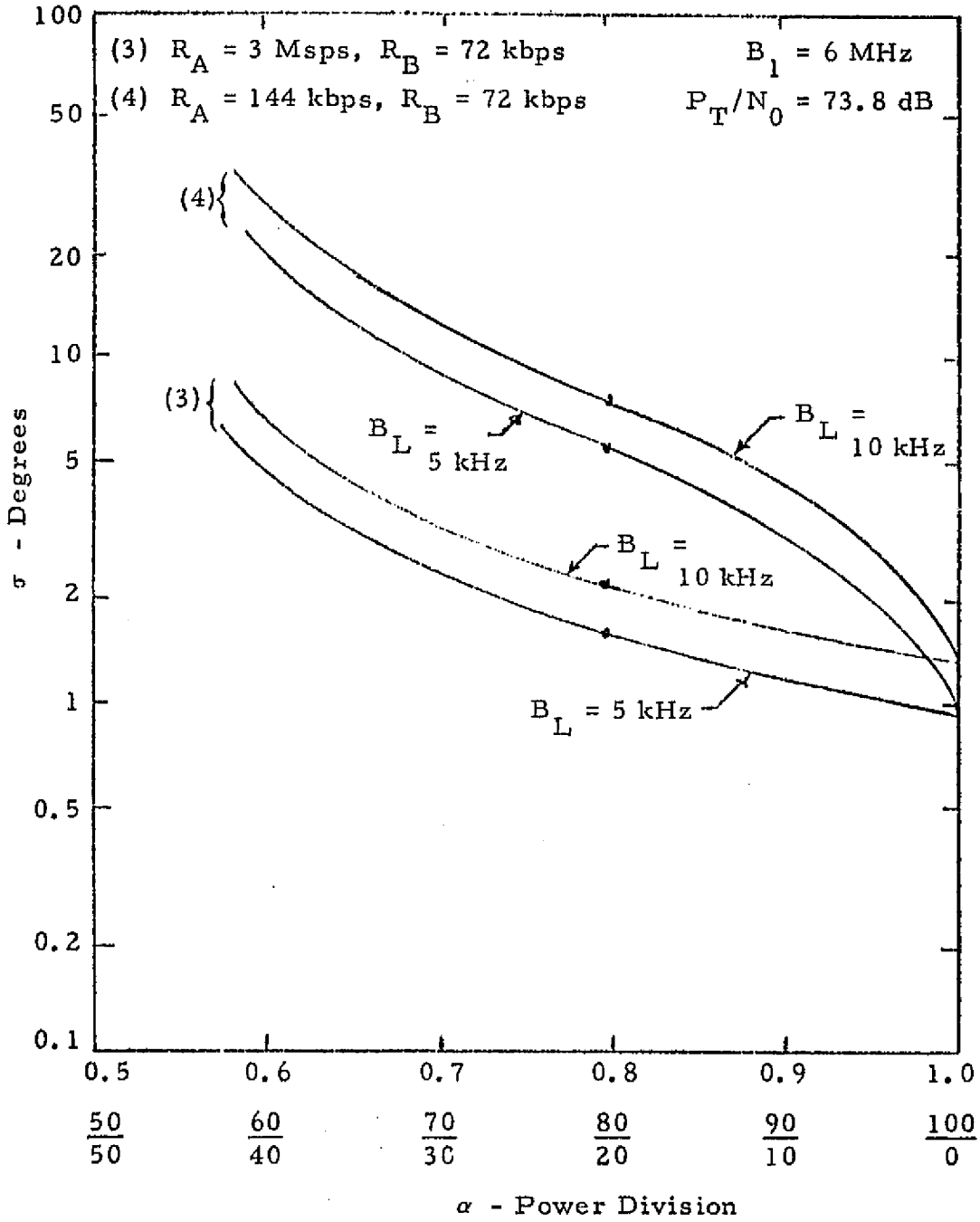


Figure 8. RMS Phase Error Versus Power Division for Unbalanced QPSK for the Forward Link

The set of curves designated (3) gives performance for the best set of rates from the RMS phase error point of view. In both the forward and return links, this corresponds to the highest rate possible, since this minimizes the spectral density of the self-noise. The set of curves (4) gives tracking performance for the worst set of rates.

The variation in performance over the possible sets of rates is not as large for the forward link as with the return link. Throughout, we have imposed the restriction that $R_A \geq 2R_B$, since this restriction was used in the development of the self-noise spectral densities in Addendum A.

For both the forward and return links at $\alpha = 0.8$, it is seen that the Costas loop receiver performs satisfactorily over the entire range of anticipated rates. We shall consider the case when the rate in the high data rate link goes to zero separately.

3.5 Return Link Costas Loop Performance When $R_A = 0$

The performance of the Costas loop receiver is now considered for the return link when the rate in the high data rate link is set equal to zero. That is, we set

$$R_A = 0 . \tag{38}$$

The powers P_A and P_B remain present, so that the predominant signal for tracking purposes is a monochromatic tone in the high data rate link.

The dynamic error signal when $R_A = 0$ and $B_1 = 100$ MHz can be determined directly from (23) by setting $a(t) = 1$, so that

$$\left.\begin{aligned} e(t) = {} & \frac{P_A - P_B}{2} \sin 2\phi \\ & + \sqrt{P_A P_B}\; b(t) \cos 2\phi \\ & + n_c(t)\, n_s(t) \\ & + n_c(t)\left[\sqrt{P_A} \sin\phi + \sqrt{P_B}\; b(t) \cos\phi\right] \\ & + n_s(t)\left[\sqrt{P_A} \cos\phi - \sqrt{P_B}\; b(t) \sin\phi\right] . \end{aligned}\right\} = n_{eq}(t) \tag{39}$$

By using the same approach as that used in Section 3.1, the auto-correlation function of $n_{eq}(t)$ is now given by

$$R_{eq}(\tau) = P_A P_B R_b(\tau) + R_N^2(\tau) + R_N(\tau)[P_A + P_B R_b(\tau)] . \quad (40)$$

The spectral density of $R_N^2(\tau)$ is given by (26). From (27), the spectral density of $P_B R_N(\tau) R_b(\tau)$ is $(N_0/2)P_B$. The spectral density of $P_A R_N(\tau)$ is $(N_0/2)P_A$. For b(t) a Bi-ϕ-L waveform, the spectral density of $P_A P_B R_b(\tau)$ is (see e.g. [4]):

$$P_A P_B T_s \frac{\sin^4(\pi f T_B/2)}{(\pi f T_B/2)^2} . \quad (41)$$

In the vicinity of f = 0, this spectral density is negligibly small, and can therefore be neglected. The spectral density of $n_{eq}(t)$ at f = 0 is then given by

$$S_{n_{eq}}(0) = \frac{N_{0_{eq}}}{2} = \frac{N_0^2 B_1}{2} + (P_A + P_B) N_0/2 \quad (42)$$

and the corresponding Costas loop signal-to-noise ratio is

$$\rho = \frac{P_T}{4 N_0 B_L} \left[\frac{(2\alpha - 1)^2}{1 + \frac{N_0 B_1}{P_T}} \right] . \quad (43)$$

The phase error variance of the Costas loop under the assumption of linearity is then related to ρ by (30). The result is plotted in Figure 9 for the return link for loop bandwidths of 5 kHz and 10 kHz. It can be seen that performance is uniformly better than when

$$384 \text{ ksps} < R_A < 100 \text{ Msps} , \quad (44)$$

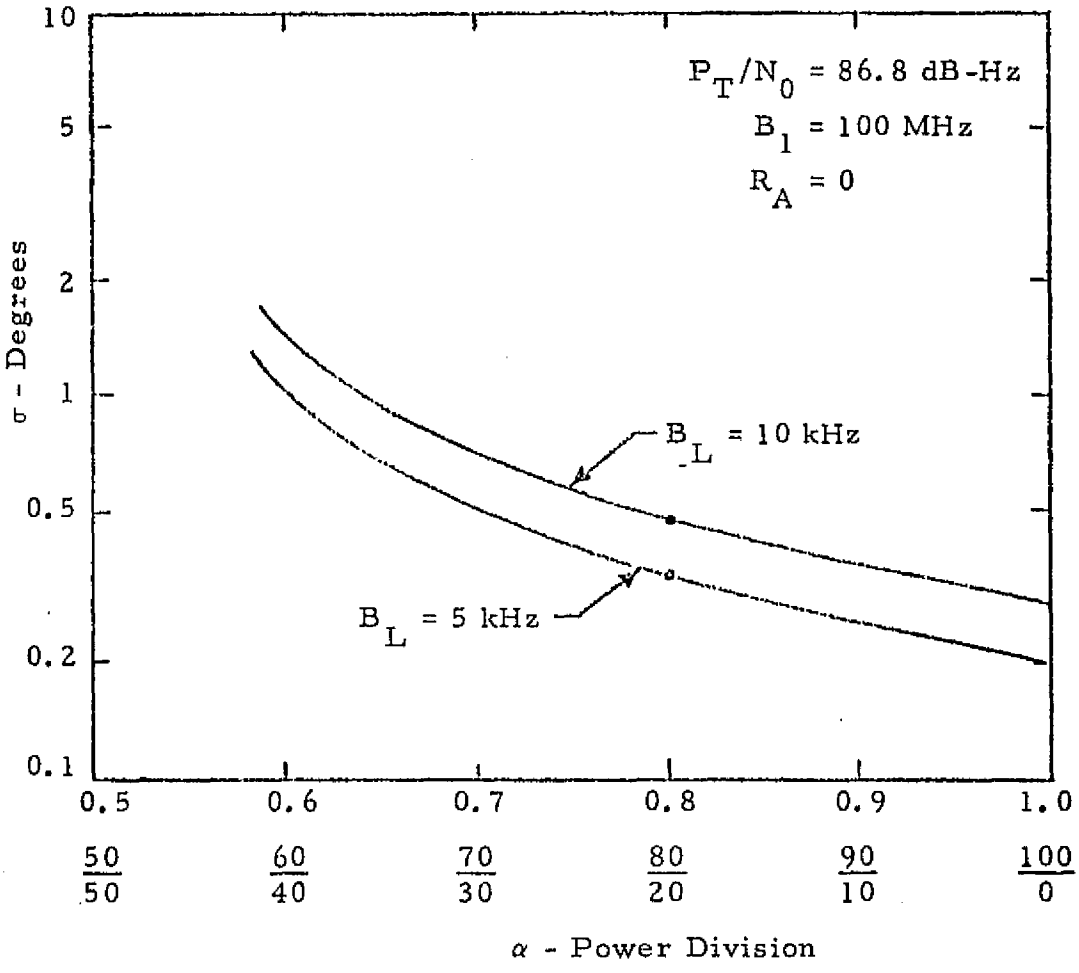


Figure 9. RMS Phase Error Versus Power Division for Unbalanced QPSK for the Return Link When R_A is Zero

which is shown in Figure 7. Since this is the case, we henceforth consider only the case where the range of rates is as shown in Figure 7.

4.0 ERROR RATE PERFORMANCE

The resulting error rate for the forward link has been approximated. The tracking loop is assumed to be operating at sufficiently high signal-to-noise ratio so that the linearized assumption that we have made throughout can be maintained. The phase error is then approximately constant throughout the symbol time of T_A seconds and is approximately a zero mean Gaussian random variable with variance

$$\sigma_\phi^2 = \frac{1}{4\rho} \tag{45}$$

where ρ is given by (37). The symbol energy-to-noise spectral density is

$$\frac{E_b}{N_0} = \frac{P_T T_A}{N_0} \ . \tag{46}$$

In the computation presented in Figure 10, we have set

$$\begin{aligned} R_A &= \frac{1}{T_A} = 3 \text{ Mbps} \\ B_L &= 0.01\ R_A = 30 \text{ kHz} \\ B_1 &= 6 \text{ MHz} \\ \alpha &= 0.8 \ . \end{aligned} \tag{47}$$

The RMS phase error is then

$$\sigma_\phi = 3.8^\circ \text{ RMS, when } \frac{P_T}{N_0} = 73.8 \text{ dB-Hz} \ . \tag{48}$$

The expression for the error rate used in Figure 10 is

$$P_e = \int_{-\pi/2}^{\pi/2} \text{erfc}\left[\sqrt{\frac{2E_b}{N_0}}\cos\phi\right] G(\phi;\, a;\, \sigma_\phi^2)\, d\phi \tag{49}$$

where erfc( ) is the complementary error function

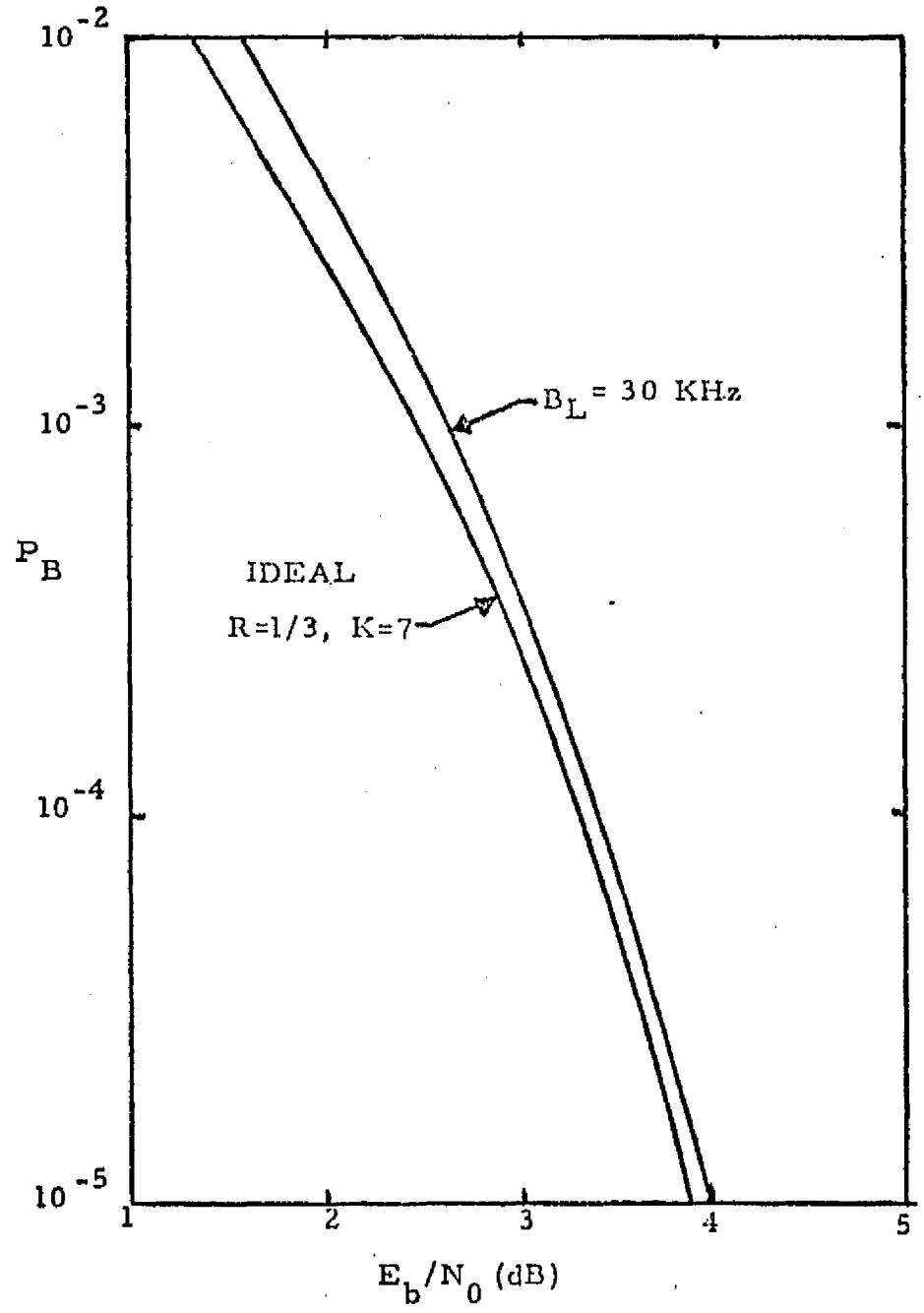


Figure 10. Effects of Costas Loop Tracking of Unbalanced QPSK

$$\text{erfc}(x) = \int_x^\infty \frac{e^{-1/2\, u^2}}{\sqrt{2\pi}}\, du \tag{50}$$

and $G(\phi;\, a\,;\, \sigma_\phi^2)$ is the Gaussian pdf of ϕ with mean a and variance σ_ϕ^2. The result was shifted by 5.3 dB to the left to represent the ideal gain obtained from Viterbi decoding of an R = 1/3, K = 7 convolutional code. The result is compared to

$$P_{e_{ideal}} = \text{erfc}\sqrt{\frac{2\, E_b}{N_0}} \tag{51}$$

which was also shifted to the left by the same amount. These approximate results show that only a few tenths of a dB are lost from ideal when employing the Costas loop which is tracking the high data rate channel. This rough analysis does not account for (1) differential encoding, (2) interchannel interference, and (3) symbol synchronization.

It is felt, however, that all of these effects would not alter the fact that the Costas loop considered in this study should remain as a viable candidate for both the forward and return link receivers.

REFERENCES

1. R. W. Allen and B. H. Batson. "Variable Data Rate Multimode Quadriphase Modem," 1973 Proceedings of the International Communications Conference.

2. S. Udalov. "Multiplexing Modulation Formats for the Orbiter's Ku-Band Return Link," Axiomatix Report No. R7502-1, February 13, 1975.

3. B. H. Batson. "Signal Designs for Orbiter Wideband Relay Communications Links," NASA Lyndon B. Johnson Space Center, Report No. EJ5-74-480, October 1, 1974.

4. W. C. Lindsey and M. K. Simon. Telecommunication Systems Engineering, Prentice Hall, 1973.

5. C. L. Weber. "Shuttle Monopulse System for Ku-Band Communication Signal from TDRS," Axiomatix Report No. R7508-3, August 22, 1975.

6. W. C. Lindsey. Synchronization Systems in Communication and Control, Prentice Hall, 1972.

7. A. J. Viterbi. Principles of Coherent Communication, McGraw-Hill, 1966.

8. I. S. Gradshteyn and I. M. Ryzhik. Table of Integrals, Series, and P Products, Academic Press, 1965, p. 446.

ADDENDUM A

Part I. Spectrum of Product of an NRZ Data Stream and a Bi-φ-L Data Stream.

In this first part of Addendum A, we determine the spectrum of the product of an NRZ waveform and a Bi-φ-L (Manchester) waveform. In the computation we impose the following restrictions:

i) The high rate signal has rate $R_A = 1/T_A$ and is an NRZ signal.

ii) The low rate signal has rate $R_B = 1/T_B$, is a Bi-φ-L signal, and

$$R_A \geq 2R_B \tag{A-1}$$

or equivalently

$$T_A < \frac{1}{2T_B} \tag{A-2}$$

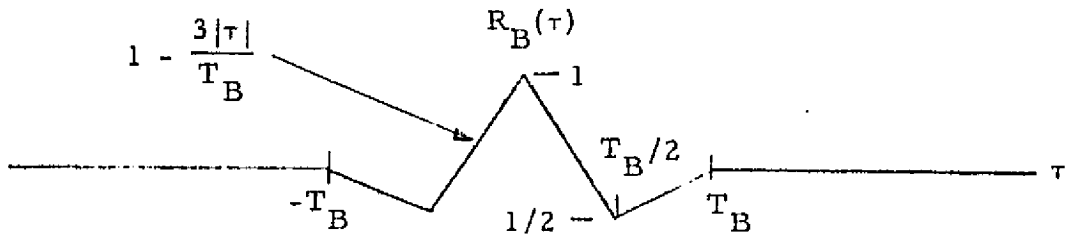
iii) The two sequences are independent and are operating asynchronously.

In Figure A-1, the autocorrelation functions $R_A(\tau)$ and $R_B(\tau)$ are shown for the case $T_A = \frac{1}{2}T_B$.

The autocorrelation of the product of NRZ and Bi-φ-L under the above assumptions is given by

$$R(\tau) = R_A(\tau)R_B(\tau) = \left(1 - \frac{3|\tau|}{T_B}\right)\left(1 - \frac{|\tau|}{T_A}\right), \quad |\tau| \leq T_A$$

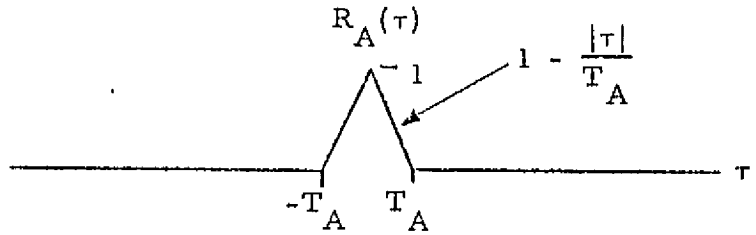
$$= \begin{cases} 1 - |\tau|\left(\frac{3}{T_B} + \frac{1}{T_A}\right) + \frac{3\tau^2}{T_A T_B}, & |\tau| \leq T_A \\ 0, & \text{otherwise} \end{cases} \tag{A-3}$$



Bi-ϕ-L

$R_A \geq 2 R_B$

$T_A \leq T_B/2$



NRZ

Figure A-1. The Autocorrelation Functions $R_A(\tau)$ and $R_B(\tau)$.

The power spectral density is given by

$$S(\omega) = \mathcal{F}\{R(\tau)\}$$

$$= 2\int_0^{\infty} \cos(\omega\tau)R(\tau)d\tau$$

$$= 2\int_0^{T_A} \cos(\omega\tau)\left[1 - \tau\left(\frac{3}{T_B} + \frac{1}{T_A}\right) + \frac{3\tau^2}{T_A T_B}\right] d\tau$$

$$= \frac{2}{\omega^2}\left\{\frac{3}{T_B} + \frac{1}{T_A} + \left(\frac{3}{T_B} - \frac{1}{T_A}\right)\cos \omega T_A - \frac{6 \sin \omega T_A}{T_A T_B \omega}\right\}$$

$$= \frac{12}{T_B \omega^2}\left[\cos^2\left(\frac{\omega T_A}{2}\right) - \frac{\sin \omega T_A}{\omega T_A}\right] + T_A \frac{\sin^2(\omega T_A/2)}{(\omega T_A/2)^2}$$

(A-4)

This spectral density is essentially flat over the bandwidth of the tracking loop. The two-sided power spectral density at $\omega = 0$, is

$$S(0) = \lim_{\omega \to 0} S(\omega) = T_A \left(1 - \frac{T_A}{T_B}\right) \quad \text{watts/Hz} \qquad \text{(A-5)}$$

Part II. Spectrum of the Product of Two Bi-φ-L Data Streams.

In Figure A-2, the autocorrelation functions of two Bi-φ-L waveforms are shown where the assumption is made that $T_A \leq \frac{1}{2}T_B$, and the data streams are statistically independent and asynchronous. The autocorrelation of the product of the two Bi-φ-L signals is then equal to the product of the

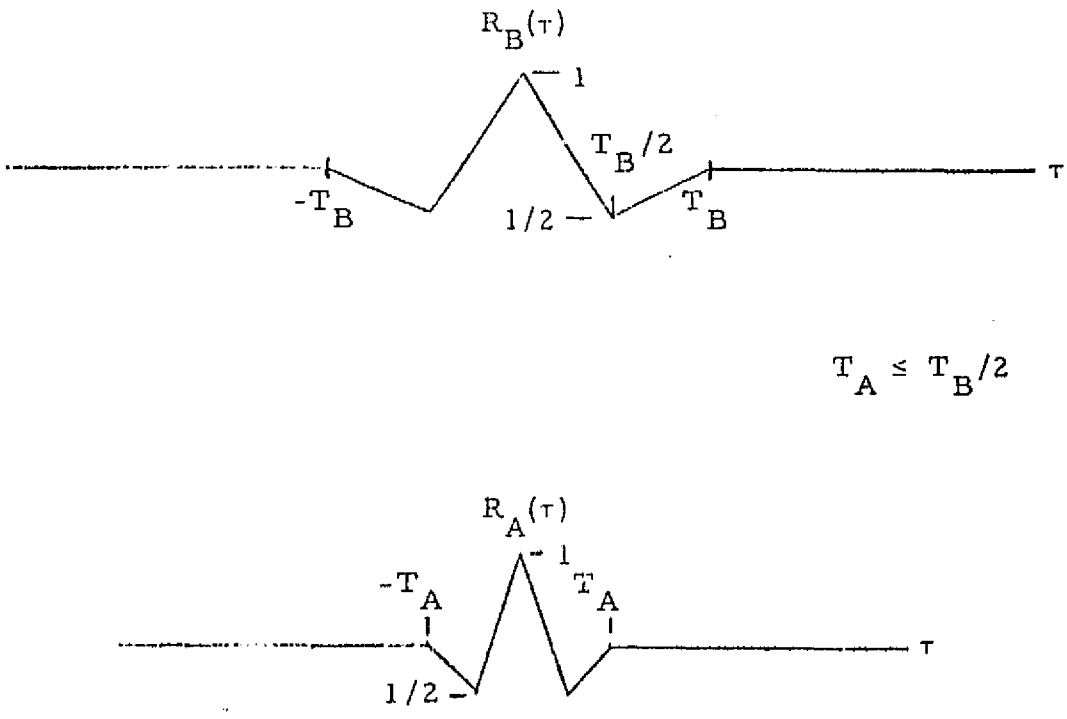


Figure A-2. The Autocorrelation Function of Two Bi-φ-L Signals.

individual autocorrelation functions. Hence

$$R(\tau) = R_A(\tau)R_B(\tau)$$

$$= \left(1 - \frac{3|\tau|}{T_A}\right)\left(1 - \frac{3|\tau|}{T_B}\right)$$

$$= 1 - 3|\tau|\left(\frac{1}{T_A} + \frac{1}{T_B}\right) + \frac{9\tau^2}{T_A T_B} \quad \text{when } |\tau| \le \frac{T_A}{2} \qquad \text{(A-6)}$$

and

$$R(\tau) = \left(\frac{|\tau|}{T_A} - 1\right)\left(1 - \frac{3|\tau|}{T_B}\right)$$

$$= |\tau|\left(\frac{1}{T_A} + \frac{3}{T_B}\right) - \left(1 + \frac{3\tau^2}{T_A T_B}\right) \quad \text{when } \frac{T_A}{2} \le |\tau| \le T_A \qquad \text{(A-7)}$$

where we still assume $R_A \ge 2R_B$. The spectral density of the product waveform is obtained from

$$S(\omega) = \mathcal{F}\{R(\tau)\}$$

$$= 2\int_0^{T_A/2} \left\{1 - 3\tau\left(\frac{1}{T_A} + \frac{1}{T_B}\right) + \frac{9\tau^2}{T_A T_B}\right\} \cos \omega\tau \, d\tau$$

$$+ 2\int_{T_A/2}^{T_A} \left\{\tau\left(\frac{1}{T_A} + \frac{3}{T_B}\right) - \left(1 + \frac{3\tau^3}{T_A T_B}\right)\right\} \cos \omega\tau \, d\tau \qquad \text{(A-8)}$$

These integrals can be evaluated by parts, which after some algebraic manipulation

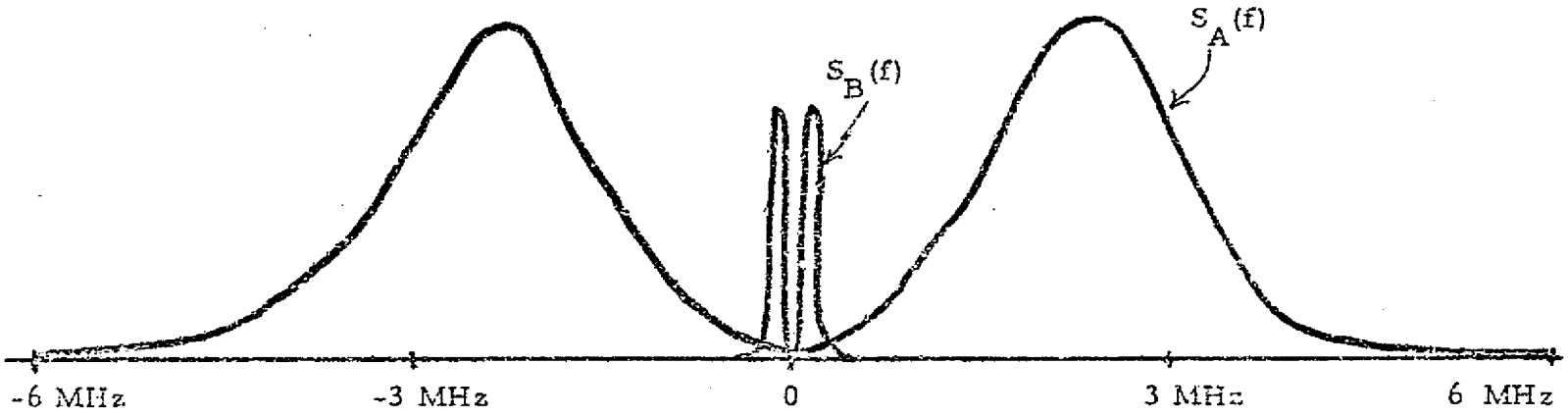
$$S(\omega) = T_A\left(\frac{\sin^4(\omega T_A/4)}{(\omega T_A/4)^2}\right) + \frac{3T_A^2}{2T_B}\cos(\omega T_A/2)\left(\frac{\sin^2(\omega T_A/4)}{(\omega T_A/4)^2}\right)$$

$$+ \frac{12}{\omega^2 T_B}\left[1 - 2\frac{\sin(\omega T_A/2)}{(\omega T_A/2)} + \frac{\sin(\omega T_A)}{(\omega T_A)}\right] \qquad \text{(A-9)}$$

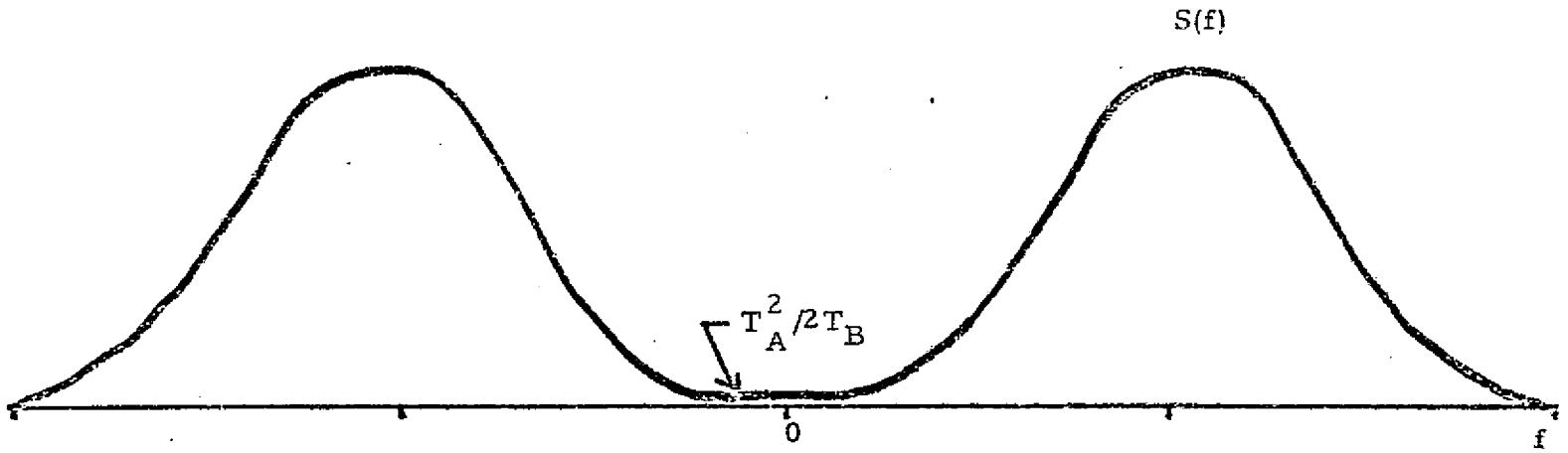
This spectral density is shown in Figure A-3 along with the original Bi-φ-L (Manchester) spectrums for the forward link. In this case also, it can be seen that the spectral density is very flat near the origin. The spectral density at f = 0 is given by

$$\lim_{\omega \to 0} S(\omega) = \frac{1}{2}\frac{T_A^2}{T_B} \qquad \text{(A-10)}$$

These spectral densities are used to determine the effect of self noise on the performance of various candidate QPSK tracking receivers.



MANCHESTER SPECTRUMS FOR THE FORWARD LINK



PRODUCT SPECTRUM OF TWO BI-φ-L WAVEFORMS

Figure A-3.

APPENDIX C

MODULATOR POWER ALLOCATION AND CARRIER TRACKING ANALYSES FOR THREE-CHANNEL ORBITER

Marvin K. Simon

APPENDIX C

MODULATOR POWER ALLOCATION AND CARRIER TRACKING ANALYSES FOR THREE-CHANNEL ORBITER Ku-BAND RETURN LINK

by Marvin K. Simon

1.0 THREE-CHANNEL PHASE MULTIPLEXING MODULATION

Consider the modulator block diagram illustrated in Figure C-1, whose purpose is 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 structure of the modulator is such as to form an unbalanced quadriphase signal wherein the high rate data signal, $\sqrt{P_1}\, m_1(t)$, is bi-phase modulated on the in-phase 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 squarewave subcarriers $Sq_2(t)$ and $Sq_3(t)$), are added and bi-phase modulated onto the quadrature carrier, $\sqrt{2}\, \cos \omega_0 t$. The sum of the in-phase and quadrature modulated carriers, i.e., the unbalanced quadriphase signal s(t) is then bandpass, hard limited and power amplified.

In view of the above, the signal generated by the modulator of Figure C-1 is

$$s(t) = \sqrt{2}\, [C(t) \cos \omega_0 t + S(t) \sin \omega_0 t] \tag{C-1}$$

where

$$\begin{aligned} C(t) &= \sqrt{P_2}\, Sq_2(t)\, m_2(t) + \sqrt{P_3}\, Sq_3(t)\, m_3(t) \\ &\triangleq \sqrt{P_2}\, s_2(t) + \sqrt{P_3}\, s_3(t) \\ S(t) &= \sqrt{P_1}\, m_1(t) \triangleq \sqrt{P_1}\, s_1(t) \end{aligned} \tag{C-2}$$

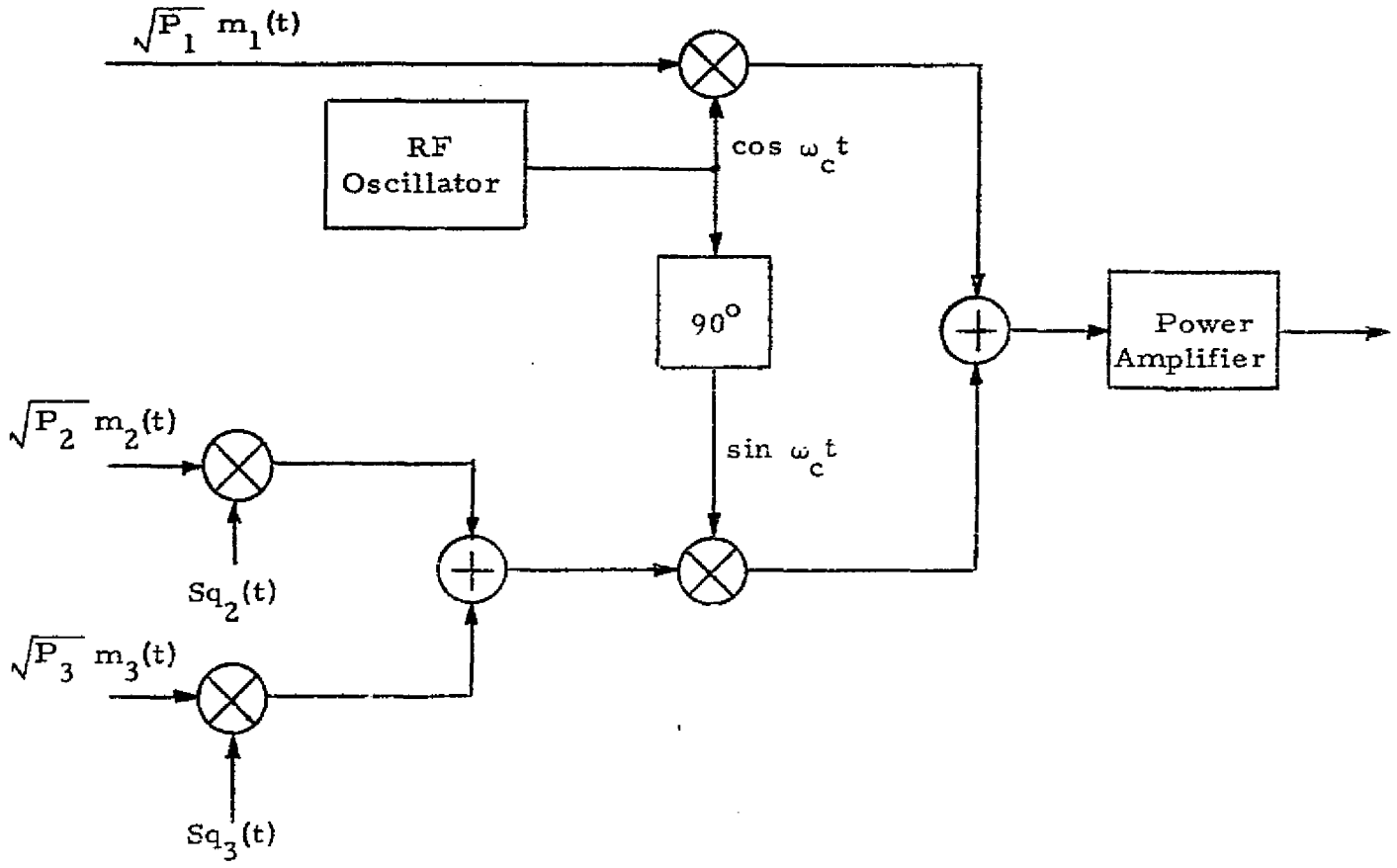


Figure C-1. Three-Channel Quadrature Multiplex Modulator

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

$$s(t) = \sqrt{2}\ V(t) \sin\left(\omega_0 + \phi(t)\right) \tag{C-3}$$

where

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

$$\phi(t) = \tan^{-1} \frac{C(t)}{S(t)} \ . \tag{C-4}$$

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

$$z_1(t) = \sqrt{2P}\ \sin\left(\omega_0 t + \phi(t)\right) \tag{C-5}$$

where P is the total power in the first zone after amplification. In terms of in-phase and quadrature components, equation (C-5) can be rewritten as

$$z_1(t) = \sqrt{2P}\left[\frac{S(t)}{V(t)} \sin \omega_0 t + \frac{C(t)}{S(t)} \cos \omega_0 t\right] . \tag{C-6}$$

From (C-2) and (C-4), we have that

$$V(t) = \sqrt{P_1 s_1^2(t) + P_2 s_2^2(t) + P_3 s_3^3(t) + 2\sqrt{P_2 P_3}\ s_2(t)\, s_3(t)}$$

$$= \sqrt{P_1 + P_2 + P_3 + 2\sqrt{P_2 P_3}\ s_2(t)\, s_3(t)} \ . \tag{C-7}$$

Let $P_T \triangleq P_1 + P_2 + P_3$ denote the total input power. Then,

$$V(t) = \sqrt{P_T\left(1 + 2\sqrt{P_2/P_T}\sqrt{P_3/P_T}\ s_2(t)\, s_3(t)\right)} \tag{C-8}$$

or, equivalently,

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

$$\left. - s_2(t)\, s_3(t)\, \underbrace{\frac{1}{2}\left[\frac{1}{\sqrt{1-2\sqrt{P_2/P_T}\sqrt{P_3/P_T}}} - \frac{1}{\sqrt{1+2\sqrt{P_2/P_T}\sqrt{P_3/P_T}}}\right]}_{C_2} \right\} \tag{C-9}$$

Substituting (C-2) and (C-9) into (C-6) gives

$$z_1(t) = \sqrt{2P}\left\{\left[C_1\sqrt{P_2/P_T} - C_2\sqrt{P_3/P_T}\right] s_2(t) + \left[C_1\sqrt{P_3/P_T} - C_2\sqrt{P_2/P_T}\right] s_3(t)\right\} \cos \omega_0 t$$

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

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

$$+ \sqrt{2}\left[\sqrt{\tilde{P}_2}\, s_2(t) + \sqrt{\tilde{P}_3}\, s_3(t)\right] \cos \omega_0 t \tag{C-10}$$

where

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

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

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

$$\tilde{P}_d = (P/P_T)\, P_1\, C_2^2 . \tag{C-11}$$

If we arbitrarily set the output level of the power amplifier equal to the total input power level, i.e., $P = P_T$, then for fixed values of P_1/P_T, P_2/P_T, and P_3/P_T, one can compute from (C-11) 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 (C-10). For example, suppose 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, we have that $P_1/P_T = 0.8$, $P_2/P_T = 0.16$, and $P_3/P_T = 0.04$. Then,

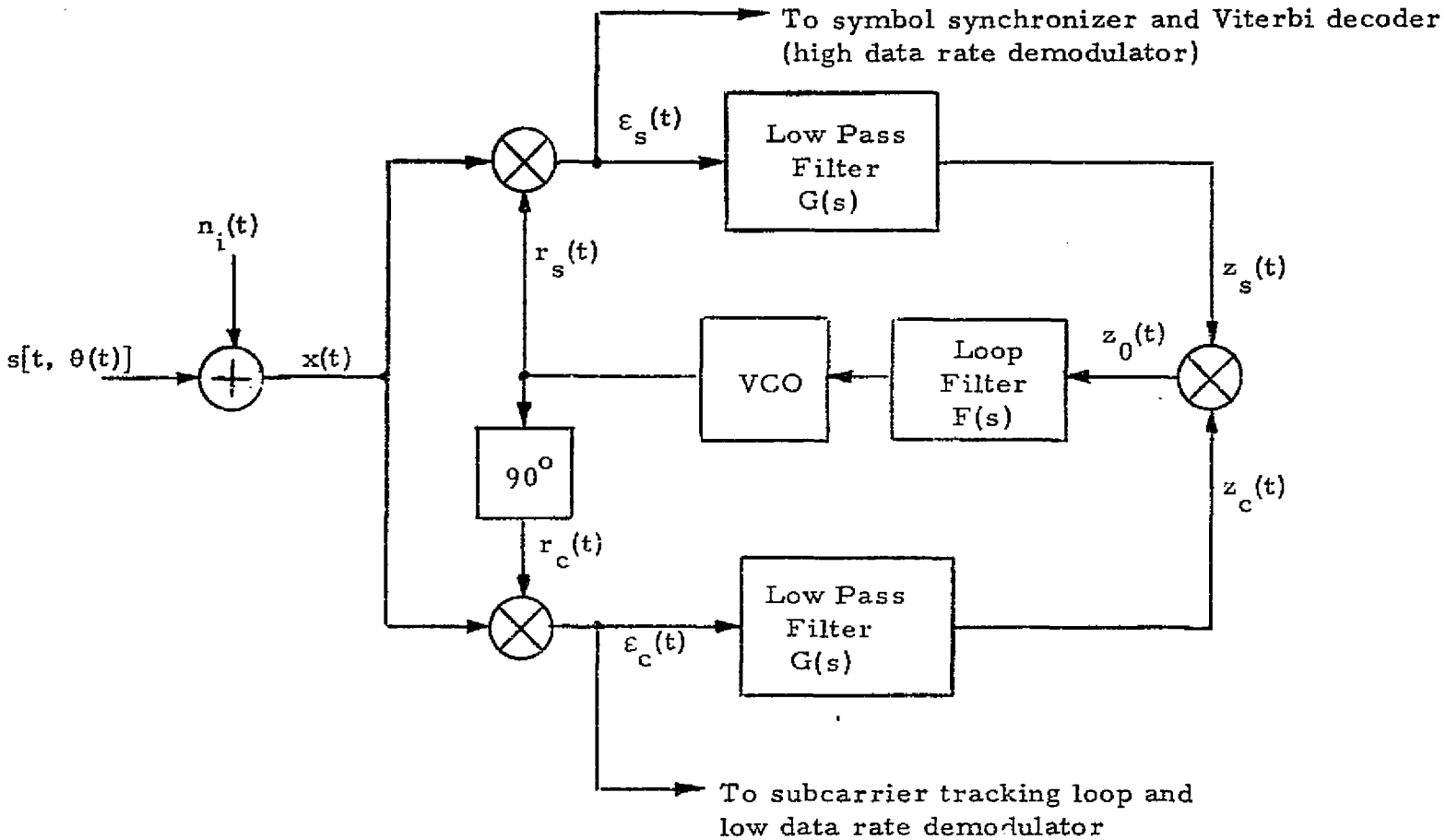


Figure C-2. Costas Loop for Carrier Tracking

from (C-11), 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 $\tilde{P}_d/P_T = 0.00529$.

2.0 CARRIER TRACKING LOOP ANALYSIS

The receiver for the three-channel quadrature multiplex signal employs a Costas loop for carrier recovery (Figure C-2). The received signal plus noise, $x(t) = s[t, \theta(t)] + n_i(t)$, is obtained from (C-10) and is of the form

$$x(t) = \sqrt{2}\left[\sqrt{\tilde{P}_1}\, s_1(t) - \sqrt{\tilde{P}_d}\, s_1(t)\, s_2(t)\, s_3(t)\right] \sin \Phi(t) + \sqrt{2}\left[\sqrt{\tilde{P}_2}\, s_2(t) + \sqrt{\tilde{P}_3}\, s_3(t)\right] \cos \Phi(t) + n_i(t) \qquad \text{(C-12)}$$

where $\Phi(t) = \omega_0 t + \theta(t)$, with ω_0 the radian carrier frequency and $\theta(t) = \theta_0 + \Omega_0 t$ the input phase to be estimated, and $n_i(t)$ is the additive channel noise which can be expressed in the form of a narrowband process about the actual frequency of the observed data, i.e.,

$$n_i(t) = \sqrt{2}\left\{N_c(t) \cos \Phi(t) - N_s(t) \sin \Phi(t)\right\} \qquad \text{(C-13)}$$

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

The in-phase and quadrature-phase reference signals (see Figure C-2) are respectively given by

$$r_s(t) = \sqrt{2}\, K_1 \sin \Phi(t) \,, \qquad r_c(t) = \sqrt{2}\, K_1 \cos \Phi(t) \,. \qquad \text{(C-14)}$$

Denoting the in-phase and quadrature-phase detector gains by K_m, then the phase detector outputs are (ignoring second harmonic terms)

$$\varepsilon_s(t) \triangleq K_m\, x(t)\, r_s(t)$$
$$= K_1 K_m \left[\sqrt{\tilde{P}_1}\, s_1(t) - \sqrt{\tilde{P}_d}\, s_1(t)\, s_2(t)\, s_3(t) - N_s(t)\right] \cos\phi(t)$$
$$- K_1 K_m \left[\sqrt{\tilde{P}_2}\, s_2(t) + \sqrt{\tilde{P}_3}\, s_3(t) + N_c(t)\right] \sin\phi(t)$$

$$\varepsilon_c(t) \triangleq K_m\, x(t)\, r_c(t)$$
$$= K_1 K_m \left[\sqrt{\tilde{P}_1}\, s_1(t) - \sqrt{\tilde{P}_d}\, s_1(t)\, s_2(t)\, s_3(t) - N_s(t)\right] \sin\phi(t)$$
$$+ K_1 K_m \left[\sqrt{\tilde{P}_2}\, s_2(t) + \sqrt{\tilde{P}_3}\, s_3(t) + N_c(t)\right] \cos\phi(t) \qquad \text{(C-15)}$$

where $\phi(t) \triangleq \Phi - \hat{\Phi}(t)$ is the loop phase error. After lowpass filtering with in-phase and quadrature-phase arm filters, G(s), the phase detector output signals become, respectively,

$$z_s(t) = K_1 K_m \left[\sqrt{\tilde{P}_1}\, \hat{s}_1(t) - \sqrt{\tilde{P}_d}\, \hat{s}_d(t) - \hat{N}_s(t)\right] \cos\phi(t)$$
$$- K_1 K_m \left[\sqrt{\tilde{P}_2}\, \hat{s}_2(t) + \sqrt{\tilde{P}_3}\, \hat{s}_3(t) + \hat{N}_c(t)\right] \sin\phi(t)$$

$$z_c(t) = K_1 K_m \left[\sqrt{\tilde{P}_1}\, \hat{s}_1(t) - \sqrt{\tilde{P}_d}\, \hat{s}_d(t) - \hat{N}_s(t)\right] \sin\phi(t)$$
$$+ K_1 K_m \left[\sqrt{\tilde{P}_2}\, \hat{s}_2(t) + \sqrt{\tilde{P}_3}\, \hat{s}_3(t) + \hat{N}_c(t)\right] \cos\phi(t) \qquad \text{(C-16)}$$

where $\hat{s}_i(t) \triangleq G(p)\, s_i(t); \quad i = 1, 2, 3$

$\hat{s}_d(t) \triangleq G(p)\, [s_1(t)\, s_2(t)\, s_3(t)]$

$\hat{N}_c(t) \triangleq G(p)\, N_c(t)$

$\hat{N}_s(t) \triangleq G(p)\, N_s(t)$ (C-17)

Multiplying these two lowpass filter outputs (assuming, for simplicity, that this multiplier has unit gain) gives the dynamic error signal

$$
\begin{aligned}
z_0(t) &\triangleq z_c(t)\, z_s(t) \\
&= \frac{K_1^2 K_m^2}{2}\left\{\left[\sqrt{\tilde{P}_1}\,\hat{s}_1(t) - \sqrt{\tilde{P}_d}\,\hat{s}_d(t)\right]^2 - \left[\sqrt{\tilde{P}_2}\,\hat{s}_2(t) + \sqrt{\tilde{P}_3}\,\hat{s}_3(t)\right]^2\right\}\sin 2\phi(t) \\
&+ \frac{K_1^2 K_m^2}{2}\left\{\left[\sqrt{\tilde{P}_1}\,\hat{s}_1(t) - \sqrt{\tilde{P}_d}\,\hat{s}_d(t)\right]\left[\sqrt{\tilde{P}_2}\,\hat{s}_2(t) + \sqrt{\tilde{P}_3}\,\hat{s}_3(t)\right]\right\}\cos 2\phi(t) \\
&+ \frac{K_1^2 K_m^2}{2}\left\{-\hat{N}_c^2(t) + \hat{N}_s^2(t) - 2\left[\sqrt{\tilde{P}_1}\,\hat{s}_1(t) - \sqrt{\tilde{P}_d}\,\hat{s}_d(t)\right]\hat{N}_s(t)\right. \\
&\qquad \left. - 2\left[\sqrt{\tilde{P}_2}\,\hat{s}_2(t) + \sqrt{\tilde{P}_3}\,\hat{s}_3(t)\right]\hat{N}_c(t)\right\}\sin 2\phi(t) \\
&+ \frac{K_1^2 K_m^2}{2}\left\{2\left[\sqrt{\tilde{P}_1}\,\hat{s}_1(t) - \sqrt{\tilde{P}_d}\,\hat{s}_d(t)\right]\hat{N}_c(t)\right. \\
&\qquad \left. - 2\left[\sqrt{\tilde{P}_2}\,\hat{s}_2(t) - \sqrt{\tilde{P}_3}\,\hat{s}_3(t)\right]\hat{N}_s(t) - 2\hat{N}_s(t)\,\hat{N}_c(t)\right\}\cos 2\phi(t)
\end{aligned}
\tag{C-18}
$$

The instantaneous frequency of the VCO output is related to $z_0(t)$ by

$$
\frac{d\hat{\phi}(t)}{dt} = K_V[F(p)\, z_0(t)] + \omega_0 \tag{C-19}
$$

and hence the stochastic integro-differential equation of loop operation becomes

$$
\begin{aligned}
\frac{2d\phi(t)}{dt} &= 2\Omega_0 - KF(p)\left\{\left[\tilde{P}_1\hat{s}_1^2(t) - \tilde{P}_2\hat{s}_2^2(t) - \tilde{P}_3\hat{s}_3^2(t) - 2\sqrt{\tilde{P}_2\tilde{P}_3}\,\hat{s}_2(t)\hat{s}_3(t)\right.\right. \\
&\qquad \left.\left. + \tilde{P}_d\hat{s}_d^2(t) - 2\sqrt{\tilde{P}_1\tilde{P}_d}\,\hat{s}_1(t)\hat{s}_d(t)\right]\sin 2\phi(t)\right. \\
&\quad + \left[\sqrt{\tilde{P}_1\tilde{P}_2}\,\hat{s}_1(t)\hat{s}_2(t) + \sqrt{\tilde{P}_1\tilde{P}_3}\,\hat{s}_1(t)\hat{s}_3(t) - \sqrt{\tilde{P}_d\tilde{P}_2}\,\hat{s}_2(t)\hat{s}_d(t)\right. \\
&\qquad \left.\left. - \sqrt{\tilde{P}_d\tilde{P}_3}\,\hat{s}_3(t)\hat{s}_d(t)\right]\cos 2\phi(t) + v_2[t, 2\phi(t)]\right\}
\end{aligned}
\tag{C-20}
$$

where $K \triangleq K_1^2 K_m^2 K_V$ and

$$v_2[t, 2\phi(t)] = \left\{-\hat{N}_c^2(t) + \hat{N}_s^2(t) - 2\left[\sqrt{\tilde{P}_1}\ \hat{s}_1(t) - \sqrt{\tilde{P}_d}\ \hat{s}_d(t)\right]\hat{N}_s(t)\right.$$
$$\left. - 2\left[\sqrt{\tilde{P}_2}\ \hat{s}_2(t) + \sqrt{\tilde{P}_3}\ \hat{s}_3(t)\right]\hat{N}_c(t)\right\} \sin 2\phi(t)$$
$$+ \left\{2\left[\sqrt{\tilde{P}_1}\ s_1(t) - \sqrt{\tilde{P}_d}\ s_d(t)\right]N_c(t)\right.$$
$$- 2\left[\sqrt{\tilde{P}_2}\ \hat{s}_2(t) + \sqrt{\tilde{P}_3}\ \hat{s}_3(t)\right]\hat{N}_s(t)$$
$$\left. - 2\hat{N}_s(t)\,\hat{N}_c(t)\right\} \cos 2\phi(t) \qquad \text{(C-21)}$$

As is customary in problems of this type, we now decompose each of the signal terms in (C-20) into its mean value plus the variation about this mean, e.g.,

$$\hat{s}_1^2(t) \sin 2\phi(t) = \overline{\hat{s}_1^2(t)} \sin 2\phi(t) + [\hat{s}_1^2(t) - \overline{\hat{s}_1^2(t)}] \sin 2\phi(t) \qquad \text{(C-22)}$$

where the overbar denotes statistical expectation. The bracketed term in (C-22) represents a self-noise due to the modulation, $\hat{s}_1^2(t)$, and such terms would result for each signal component in (C-20) when decomposed as in (C-22). Since all three data rates, $\mathscr{R}_{s_i} \triangleq 1/T_{s_i}$ $(i = 1,2,3)$, are large with respect to the single-sided loop bandwidth B_L, then as has been shown previously for the case of a Costas loop tracking a bi-phase modulated carrier, the self-noise terms can be neglected. Furthermore,

$$\overline{\hat{s}_i^2(t)} \triangleq D_i = \int_{-\infty}^{\infty} S_i(f)\left|G(j2\pi f)\right|^2 df\ ; \quad i = 1, 2, 3$$

$$\overline{\hat{s}_d^2(t)} \triangleq D_d = \int_{-\infty}^{\infty} S_d(f)\left|G(j2\pi f)\right|^2 df \qquad \text{(C-23)}$$

where $S_i(f)$ is the power spectral density of $s_i(t)$; $i = 1, 2, 3$, $\left|G(j2\pi f)\right|^2$ is the square of the magnitude of the arm filter transfer function, and $S_d(f) = S_1(f) * S_2(f) * S_3(f)$, where the asterisk denotes convolution. In

addition, since $m_i(t)$ is independent of $m_j(t)$; $i \neq j = 1, 2, 3$, then

$$\overline{\hat{s}_i(t)\,\hat{s}_j(t)} = 0 \qquad i, j = 1, 2, 3;\ i \neq j$$

$$\overline{\hat{s}_i(t)\,\hat{s}_d(t)} = 0 \qquad i = 1, 2, 3 \tag{C-24}$$

Making use of (C-23) and (C-24) in (C-20), along with the assumption regarding the neglect of the modulation self-noise terms, we get the simplified stochastic equation of loop operation

$$2\frac{d\phi(t)}{dt} = 2\Omega_0 - KF(p)\Big\{[\tilde{P}_1 D_1 - \tilde{P}_2 D_2 - \tilde{P}_3 D_3 + \tilde{P}_d D_d]\sin 2\phi(t) + v_2[t, 2\phi(t)]\Big\}. \tag{C-25}$$

We now proceed to evaluate the noise spectral density of the equivalent delta-correlated noise process, $v_2[t, 2\phi(t)]$, i.e.,

$$N_{sq} = 2\int_{-\infty}^{\infty} R_{v_2}(\tau)\,d\tau \tag{C-26}$$

where

$$R_{v_2}(\tau) \triangleq \overline{v_2[t, 2\phi(t)]\,v_2[t+\tau, 2\phi(t)]}\ . \tag{C-27}$$

Substituting (C-21) into (C-27), we obtain, after considerable algebraic manipulation,

$$R_{v_2}(\tau) = 4\Big\{[\tilde{P}_1 R_{\hat{s}_1}(\tau) + \tilde{P}_2 R_{\hat{s}_2}(\tau) + \tilde{P}_3 R_{\hat{s}_3}(\tau) + \tilde{P}_d R_{\hat{s}_1}(\tau) R_{\hat{s}_2}(\tau) R_{\hat{s}_3}(\tau)] R_{\hat{N}}(\tau) + R_{\hat{N}}^2(\tau)\Big\} \tag{C-28}$$

where

$$R_{\hat{s}_i}(\tau) \triangleq \overline{\hat{s}_i(t)\,\hat{s}_i(t+\tau)}\ ; \quad i = 1, 2, 3$$

$$R_{\hat{N}}(\tau) = \frac{N_0}{2}\int_{\infty}^{\infty} |G(j2\pi f)|^2\, e^{j2\pi f\tau}\,df\ . \tag{C-29}$$

Integrating (C-28) between $-\infty$ and ∞ and noting that

$$\int_{-\infty}^{\infty} R_N^2(\tau)\, d\tau = \int_{-\infty}^{\infty} \left[\frac{N_0}{2}\left|G(j2\pi f)\right|^2\right]^2 df$$

$$\int_{-\infty}^{\infty} R_{\hat{s}_i}(\tau) R_{\hat{N}}(\tau)\, d\tau = \frac{N_0}{2}\int_{-\infty}^{\infty} S_i(f)\left|G(j2\pi f)\right|^4 df \qquad i = 1, 2, 3$$

$$\int_{-\infty}^{\infty} R_{\hat{s}_1}(\tau) R_{\hat{s}_2}(\tau) R_{\hat{s}_3}(\tau) R_{\hat{N}}(\tau)\, d\tau = \frac{N_0}{2}\int_{-\infty}^{\infty} S_{\hat{d}}(f)\left|G(j2\pi f)\right|^2 df \qquad \text{(C-30)}$$

with

$$S_{\hat{d}}(f) \triangleq S_1(f)\left|G(j2\pi f)\right|^2 * S_2(f)\left|G(j2\pi f)\right|^2 * S_3(f)\left|G(j2\pi f)\right|^2 \qquad \text{(C-31)}$$

we get an expression for N_{sq} given by

$$N_{sq} = 4N_0\left\{\sum_{i=1}^{3}\tilde{P}_i\int_{-\infty}^{\infty} S_i(f)\left|G(j2\pi f)\right|^4 df + \tilde{P}_d\int_{-\infty}^{\infty} S_{\hat{d}}(f)\left|G(j2\pi f)\right|^2 df\right\}$$

$$\times\left[1 + \frac{2\int_{-\infty}^{\infty}\left[\frac{N_0}{2}\left|G(j2\pi f)\right|^2\right]^2 df}{N_0\left\{\sum_{i=1}^{3}\tilde{P}_i\int_{-\infty}^{\infty} S_i(f)\left|G(j2\pi f)\right|^4 df + \tilde{P}_d\int_{-\infty}^{\infty} S_{\hat{d}}(f)\left|G(j2\pi f)\right|^2 df\right\}}\right] \qquad \text{(C-32)}$$

Finally, from (C-25), the <u>equivalent loop signal-to-noise ratio</u> ρ' is

$$\rho' = \frac{\left(\tilde{P}_1 D_1 - \tilde{P}_2 D_2 - \tilde{P}_3 D_3 + \tilde{P}_d D_d\right)^2}{N_{sq} B_L} \triangleq \frac{\rho}{4}\mathscr{A}_L \qquad \text{(C-33)}$$

where $\rho = P/N_0 B_L$ is the loop signal-to-noise ratio of a phase-locked (CW) loop operating on the total power P and

$$\mathscr{A}_L = \frac{4N_0}{N_{sq}}\left[\frac{\tilde{P}_1 D_1 - \tilde{P}_2 D_2 - \tilde{P}_3 D_3 + \tilde{P}_d D_d}{P}\right] \qquad \text{(C-34)}$$

is the loop squaring loss. What remains is to express $\mathscr{L}_L$ in terms of the basic system design parameters. Defining the modulation indices η_i; $i = 1, 2, 3$, and η_d by

$$\eta_i \triangleq \frac{\tilde{P}_i}{P}; \qquad i = 1, 2, 3$$

$$\eta_d \triangleq \frac{\tilde{P}_d}{P} \tag{C-35}$$

we obtain, after much simplification,

$$\mathscr{L}_L = \frac{\dfrac{(\eta_1 D_1 - \eta_2 D_2 - \eta_3 D_3 + \eta_d D_d)^2}{\overline{D}}}{\overline{K_D} + \dfrac{K_L}{\rho_i \overline{D}}} \tag{C-36}$$

where

$$\overline{D} \triangleq \eta_1 D_1 + \eta_2 D_2 + \eta_3 D_3 + \eta_d D_d = \text{average filter distortion factor}$$

$$K_L \triangleq \frac{\int_{-\infty}^{\infty} |G(j2\pi f)|^4 \, df}{\int_{-\infty}^{\infty} |G(j2\pi f)|^2 \, df} = \text{filter constant}$$

$$B_i = \int_{-\infty}^{\infty} |G(j2\pi f)|^2 \, df = \text{two-sided arm filter noise bandwidth}$$

$$\overline{K_D} = \frac{\sum_{i=1}^{3} \eta_i \int_{-\infty}^{\infty} S_i(f) |G(j2\pi f)|^4 \, df + \eta_d \int_{-\infty}^{\infty} S_{\hat{d}}(f) |G(j2\pi f)|^2 \, df}{\sum_{i=1}^{3} \eta_i \int_{-\infty}^{\infty} S_i(f) |G(j2\pi f)|^2 \, df + \eta_d \int_{-\infty}^{\infty} S_d(f) |G(j2\pi f)|^2 \, df}$$

$$= \text{combined filter - signal spectrum constant} \tag{C-37}$$

and $\rho_i \triangleq 2P/N_0 B_i$ is the total power-to-noise ratio in the arm filter bandwidth. Alternately, since the bandwidth B_i will be optimally chosen relative

to the high data rate symbol stream at rate $\mathscr{R}_{s_1}$, then

$$\mathscr{L}_L = \frac{\dfrac{(\eta_1 D_1 - \eta_2 D_2 - \eta_3 D_3 + \eta_d D_d)^2}{\overline{D}}}{\overline{K_D} + \left(\dfrac{B_i/\mathscr{R}_{s_1}}{2 R_{d_1}}\right)\dfrac{K_L \eta_1}{\overline{D}}} \qquad \text{(C-38)}$$

where $R_{d_1} \triangleq \tilde{P}_1 T_1/N_0$ is the data signal-to-noise ratio for the high data rate channel.

The simplest case, at least mathematically, to use as an illustration is that of an ideal arm filter for which

$$\left|G(j2\pi f)\right|^2 = \begin{cases} 1\ ; & |f| \le B_i/2 \\ 0 & |f| > B_i/2 \end{cases} \qquad \text{(C-39)}$$

In this situation, we see that $\overline{K_D} = K_L = 1$. Thus, (C-36) simplifies to

$$\mathscr{L}_L = \frac{(\eta_1 D_1 - \eta_2 D_2 - \eta_3 D_3 + \eta_d D_d)^2}{(\eta_1 D_1 + \eta_2 D_2 + \eta_3 D_3 + \eta_d D_d) + \dfrac{1}{\rho_i}} \qquad \text{(C-40)}$$

To simplify (C-40) any further requires specification of the format of the three signal components. As an example, suppose that $s_1(t) = m_1(t)$ is an NRZ modulation of rate $\mathscr{R}_{s_1} = 1/T_{s_1}$. Then,

$$S_1(f) = T_{s_1}\left(\frac{\sin \pi f T_{s_1}}{\pi f T_{s_1}}\right)^2 \qquad \text{(C-41)}$$

For an ideal filter,

$$D_1 = \int_{-B_i/2}^{B_i/2} \left(\frac{\sin \pi f T_{s_1}}{\pi f T_{s_1}}\right)^2 df T_{s_1} = 2 \int_0^{B_i/2 \mathcal{R}_{s_i}} \left(\frac{\sin \pi x}{\pi x}\right)^2 dx \qquad \text{(C-42)}$$

which can be expressed in terms of the tabulated sine integral $S_i(x)$ function. Furthermore, suppose $s_2(t) = m_2(t)Sq_2(t)$ is an NRZ modulation of rate $\mathcal{R}_{s_2} = 1/T_{s_2}$ which is bi-phase modulated on a square wave subcarrier of frequency f_{sc}. Then,

$$S_2(f) = \sum_{n=1,3,5,\ldots}^{\infty} \left(\frac{4}{n\pi}\right)^2 S_{m_2}(f - nf_{sc}) \qquad \text{(C-43)}$$

where

$$S_{m_2}(f) = T_{s_2} \left(\frac{\sin \pi f T_{s_2}}{\pi f T_{s_2}}\right)^2 \qquad \text{(C-44)}$$

Finally, if $s_3(t) = m_3(t)Sq_3(t)$ is a Manchester coded modulation of rate $\mathcal{R}_{s_3} = 1/T_{s_3}$, bi-phase modulated on the quadrature square wave subcarrier $Sq_3(t) = Sq_2(t - 1/4f_{sc})$, then

$$S_3(f) = \sum_{n=1,3,5,\ldots}^{\infty} \left(\frac{4}{n\pi}\right)^2 S_{m_3}(f - nf_{sc}) \qquad \text{(C-45)}$$

where

$$S_{m_3}(f) = T_{s_3} \left[\frac{\sin^4 \frac{\pi f T_{s_3}}{2}}{\left(\frac{\pi f T_{s_3}}{2}\right)^2}\right] \qquad \text{(C-46)}$$

For an ideal filter, D_2 and D_3 can be evaluated by direct numerical integration of (C-43) and (C-46), respectively.

3.0 PROBABILITY DENSITY FUNCTION OF THE LOOP PHASE ERROR

Using Fokker-Planck techniques, the steady-state probability density function (p.d.f.) $p(2\phi)$ of the modulo 2π reduced phase error 2ϕ can be

determined from (C-25). Assuming a loop filter of the form

$$F(s) = \frac{1 + s\tau_2}{1 + s\tau_1} ; \qquad F_1 = \tau_2/\tau_1 \tag{C-47}$$

then (Ref. 1, Ch. 2, and Ref. 2, Chs. 9-10),

$$p(2\phi) = C_0' \exp(\beta 2\phi + a \cos 2\phi) \int_{2\phi}^{2\phi + 2\pi} \exp(-\beta 2x - a \cos 2x)\, dx ;$$

$$|\phi| \le \pi \tag{C-48}$$

where C_0' is a normalization constant

$$\int_{-\pi}^{\pi} p(2\phi)\, d\phi = 2 \int_{-\pi/2}^{\pi/2} p(2\phi)\, d\phi = 1 \tag{C-49}$$

and

$$a = \left(\frac{r+1}{r}\right)\rho' - \frac{1 - F_1}{2 r \sigma_G^2} , \tag{C-50}$$

$$\beta = \left(\frac{r+1}{r}\right)^2 \frac{\rho'}{4 B_L} \left[2\Omega_0 - P[\eta_1 D_1 - \eta_2 D_2 - \eta_3 D_3 + \eta_d D_d] K(1 - F_1) \overline{\sin 2\phi}\right]$$

$$+ a \overline{\sin 2\phi} \tag{C-51}$$

where r = second order loop damping parameter

$= P[\eta_1 D_1 - \eta_2 D_2 - \eta_3 D_3 + \eta_d D_d] K F_1 \tau_2 = 4\zeta^2$

ζ = loop damping

ρ' = effective signal-to-noise ratio in the loop bandwidth

$= \dfrac{\rho \mathcal{A}_L}{4}$

$$\sigma_G^2 = \overline{\sin^2 2\phi} - \overline{(\sin 2\phi)}^2 \tag{C-52}$$

In the linear region of operation, the variance of the loop phase error (tracking jitter) is given by

$$\sigma_{2\phi}^2 = \frac{1}{\rho'} \tag{C-53}$$

where ρ' is found from (C-33) and (C-36); whereas, in the nonlinear region, $\sigma_{2\phi}^2$ must be determined by numerical integration, viz.,

$$\sigma_{2\phi}^2 = 2\int_{-\pi/2}^{\pi/2} (2\phi - 2\bar{\phi})^2 \, p(2\phi)\, d\phi \, . \tag{C-54}$$

4.0 CROSSTALK CONSIDERATIONS

The presence of tracking phase error in the receiver introduces crosstalk from one data channel into another. In particular, we observe from (C-15) that the low data rate channels introduce crosstalk into the high rate channel in proportion to their relative received powers and the carrier tracking loop phase error. Defining $D_k(\ell)\big|_\phi$ equal to the ratio of power in the kth channel to the crosstalk power due to the data in the ℓth channel conditioned on a fixed carrier tracking loop phase error ϕ, we see from (C-15) that

$$D_1(2)\big|_\phi = \frac{\tilde{P}_1 \cos^2\phi}{\tilde{P}_2 \sin^2\phi} = \frac{\tilde{P}_1}{\tilde{P}_2} \cot^2\phi$$

$$D_1(3)\big|_\phi = \frac{\tilde{P}_1 \cos^2\phi}{\tilde{P}_3 \sin^2\phi} = \frac{\tilde{P}_1}{\tilde{P}_3} \cot^2\phi \tag{C-55}$$

where the ratios $\tilde{P}_1/\tilde{P}_2$ and $\tilde{P}_1/\tilde{P}_3$ may be determined from (C-11). The <u>average</u> crosstalks $D_1(2)$ and $D_1(3)$ may then be evaluated by averaging (C-55) over the p.d.f. of (C-48), i.e.,

$$D_1(2) = 2\int_{-\pi/2}^{\pi/2} p(2\phi)\,\frac{\tilde{P}_1}{\tilde{P}_2}\,\cot^2\phi\,d\phi$$

$$D_1(3) = 2\int_{-\pi/2}^{\pi/2} p(2\phi)\,\frac{\tilde{P}_1}{\tilde{P}_3}\,\cot^2\phi\,d\phi\,. \qquad \text{(C-56)}$$

The computations required in (C-56) must be carried out by numerical integration.

REFERENCES:

1. Lindsey, W. C., and Simon, M. K., Telecommunication Systems Engineering (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1973).

2. Lindsey, W. C., Synchronization Systems in Communication and Control (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972).

APPENDIX D

EFFECTS OF PHASE NOISE ON THE PERFORMANCE OF TWO-WAY COHERENT COMMUNICATION LINKS

Gaylord K. Huth

APPENDIX D

EFFECTS OF PHASE NOISE ON THE PERFORMANCE OF TWO-WAY COHERENT COMMUNICATION LINKS

Gaylord K. Huth

This appendix systematically considers the effects of transmitter phase noise on the performance of two-way coherent communications links and indicates design approaches that may be used to reduce or minimize such effects. For typical transmitter phase noise spectra, the resulting degradations in bit error rate performance for baseband and phase-shift-keyed subcarrier channels are determined.

1.0 INTRODUCTION

Figure 1 illustrates functionally the generalized two-way coherent communications link which is considered in this appendix. Each portion (forward and return) of this link is assumed to consist of a baseband digital data channel, summed together with n phase-shift-keyed (PSK) digital data subcarrier channels prior to phase modulation of the transmitted carrier. Of concern here are the effects of transmitted phase noise on the bit error rate performance of these various channels. For the forward portion of the two-way link, the transmitted phase noise is due primarily to jitter introduced by the timing source which generates the carrier frequency. For the return portion of the link, the transmitted phase noise is due primarily to the turnaround of the phase noise which is not tracked out by the forward link receiver carrier reference loop.

2.0 TRANSMITTER PHASE NOISE SPECTRA

As a first approximation, the effects of transmitter phase noise on bit error rate performance of communications links are sometimes determined assuming that the phase noise power spectrum is white, as in Figure 2(a). In practice, the transmitted phase noise spectral characteristics are

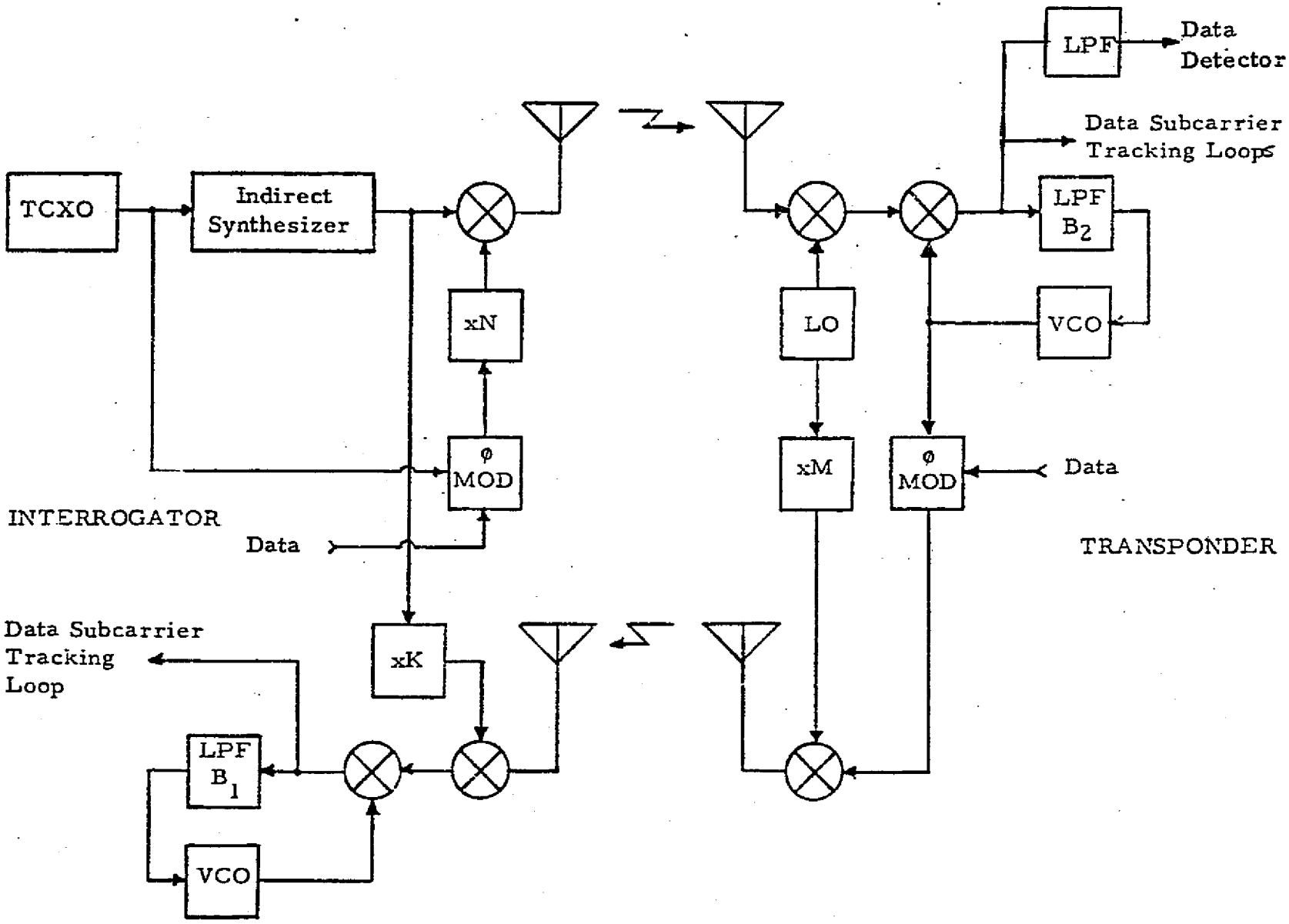
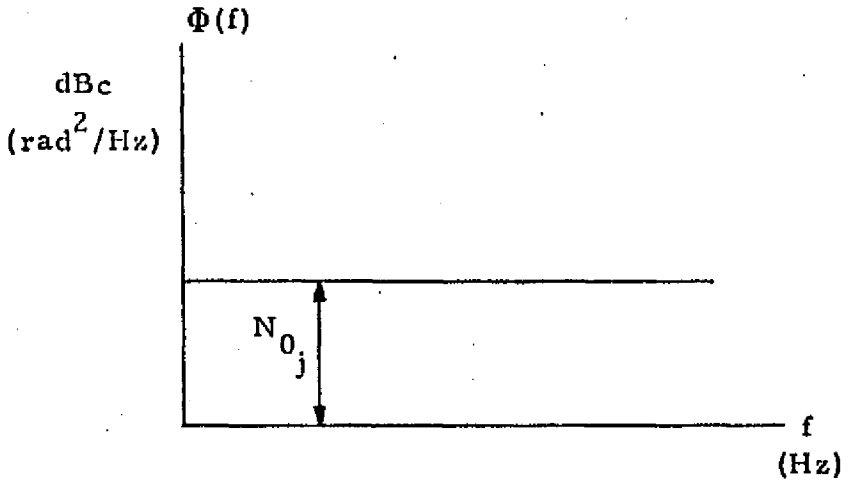
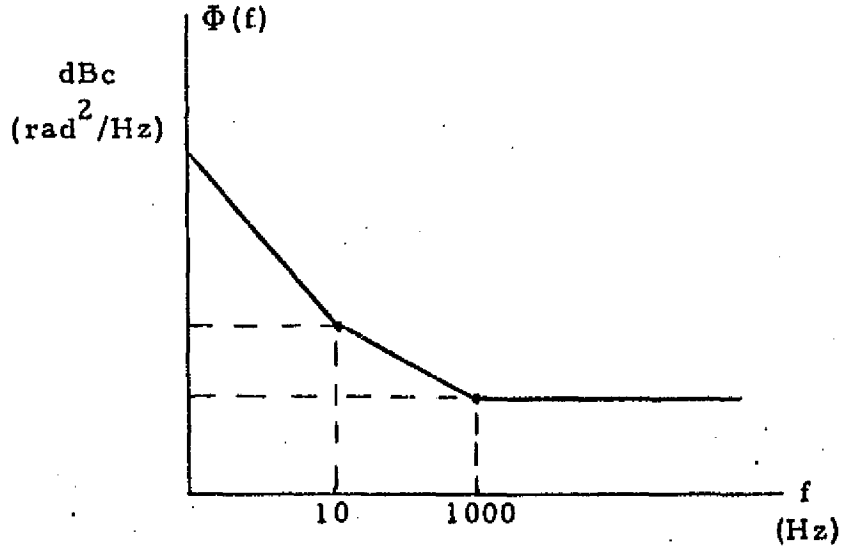


Figure 1. Functional Block Diagram of Interrogator and Transponder



(a) Simplified



(b) Typical

Figure 2. Transmitter Phase Noise Spectra

very dependent on the configuration and characteristics of the circuits used to generate the transmit carrier frequency. Hence, transmitted phase noise spectra may vary considerably from one system to another. Figure 2(b) illustrates the phase noise spectrum typical of what might be encountered over a number of systems. The approach presented subsequently for evaluating the overall effects of transmitted phase noise is general and applies for any spectral characteristics. Only the final numerical results are dependent on the assumed spectral characteristics.

As a specific example of phase noise spectrums, consider the frequency synthesizer necessary for the S-band payload interrogator in the DSN mode presented in Figure 3. Phase noise is generated from each oscillator (i.e., the TCXO, the VCXO, and the VCO) and from each phase detector. These sources of phase noise are accentuated by the multipliers and attenuated by the loop filters. First consider the phase noise generated by the TCXO, the VCXO, and the associated phase detector as modified by the multiplier and loop filter. Figure 4 analyzes each of these sources and combines them to form the total phase noise at the output of the VCXO loop. The asymptotes for the loop filter are also shown. Note that the loop filter attenuates the low frequency response of the VCXO due to the feedback and the high frequency response of the TCXO and phase detector by direct filtering. By adding up the resultant curves of the filtered phase noise source shown in Figure 4, the total output phase noise spectrum is obtained. Note that the loop filter is actually 3 dB down at the 100 Hz cutoff frequency rather than the asymptotes shown and the total phase noise spectrum curve takes this into account.

The next source of phase noise in the system of interest is the output of the VCXO multiplied by 36. If this multiplier is implemented with a phase locked oscillator (PLO), then the effect at the high end of the spectrum can be improved. Figure 5 analyzes the resulting phase noise spectrum using the X36 PLO. First, by multiplying the output VCXO phase noise spectrum obtained in Figure 4 by 36 will give the resulting phase noise

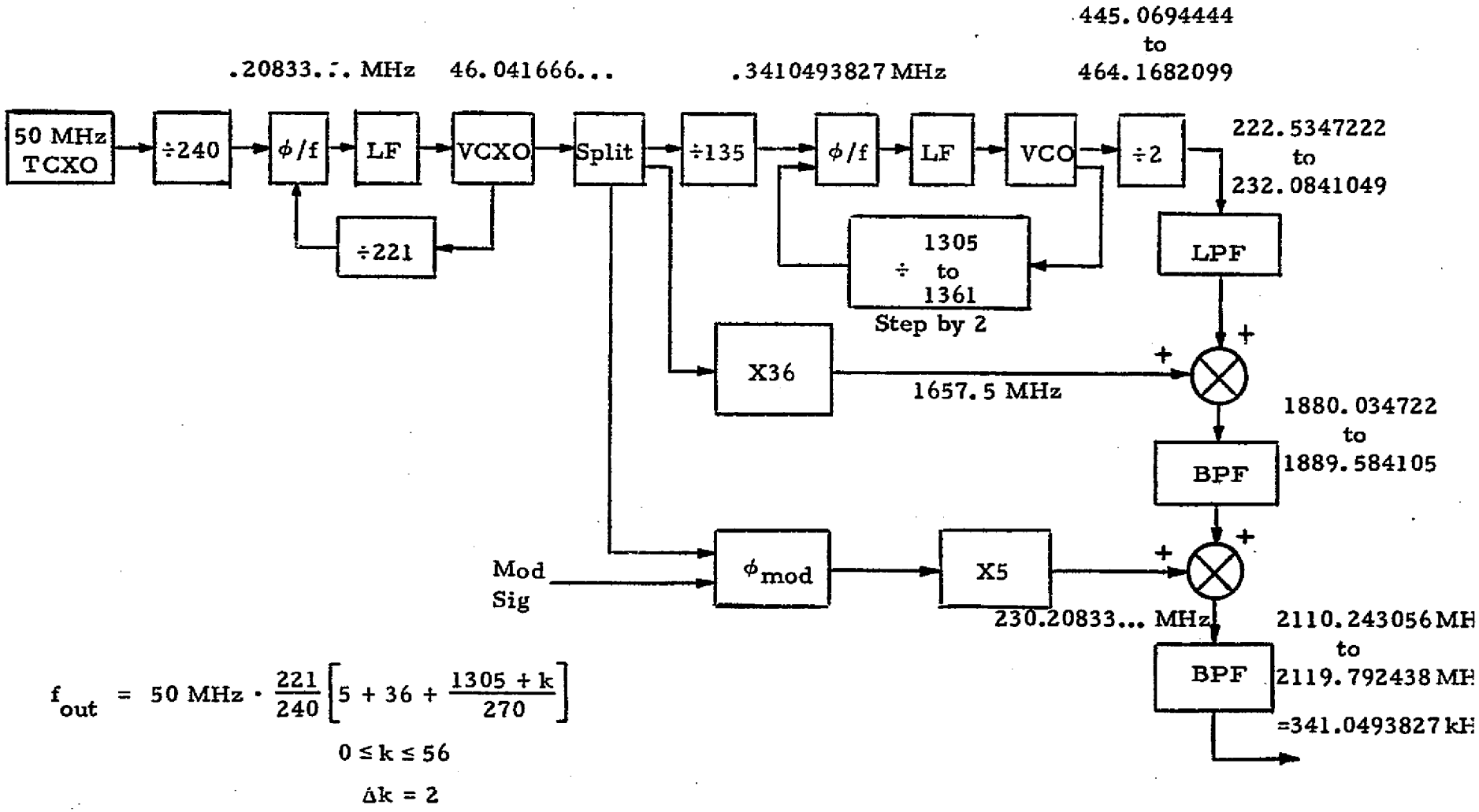


Figure 3. Frequency Synthesizer for S-Band Payload Interrogator in the DSN Mode

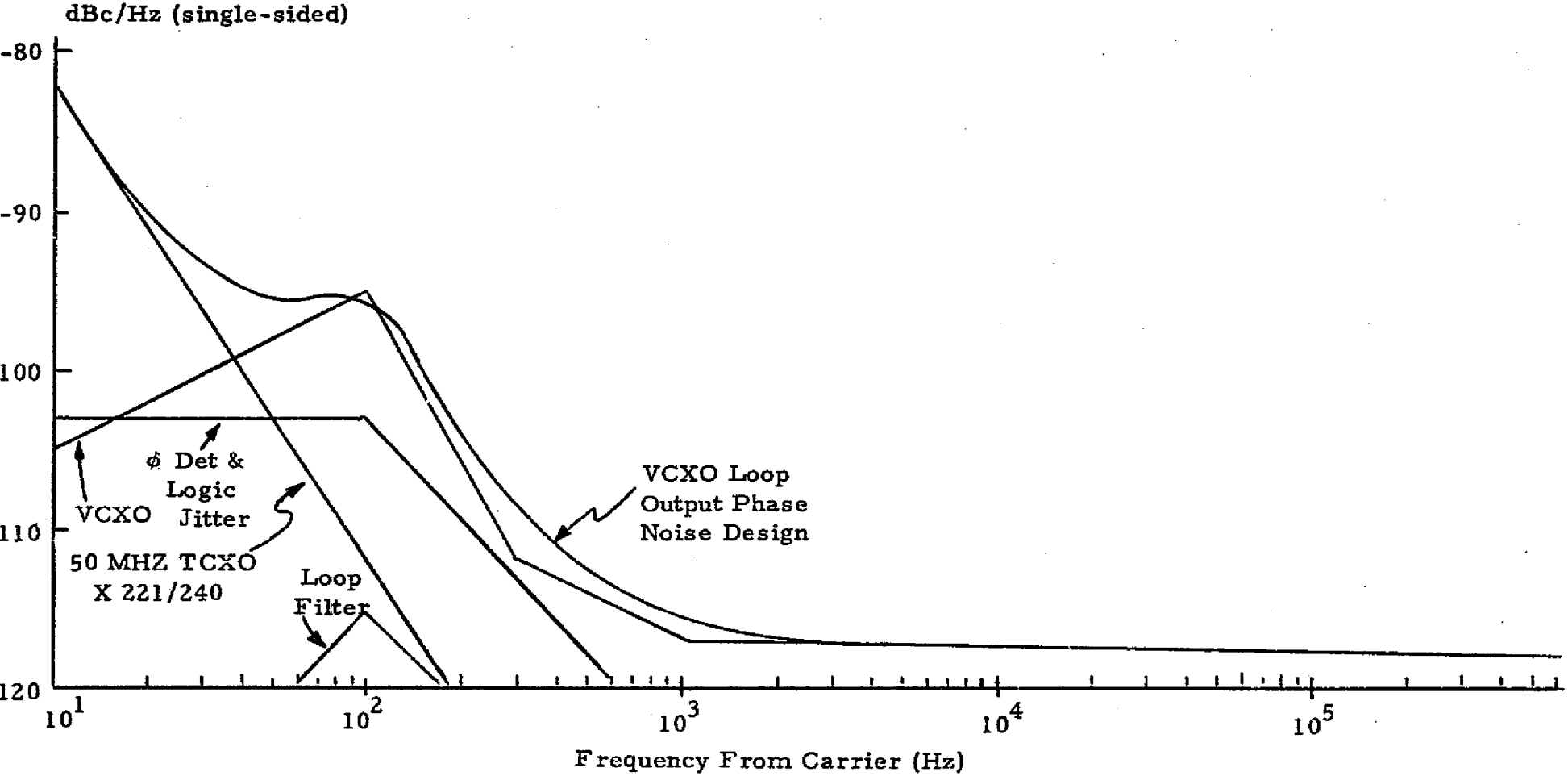
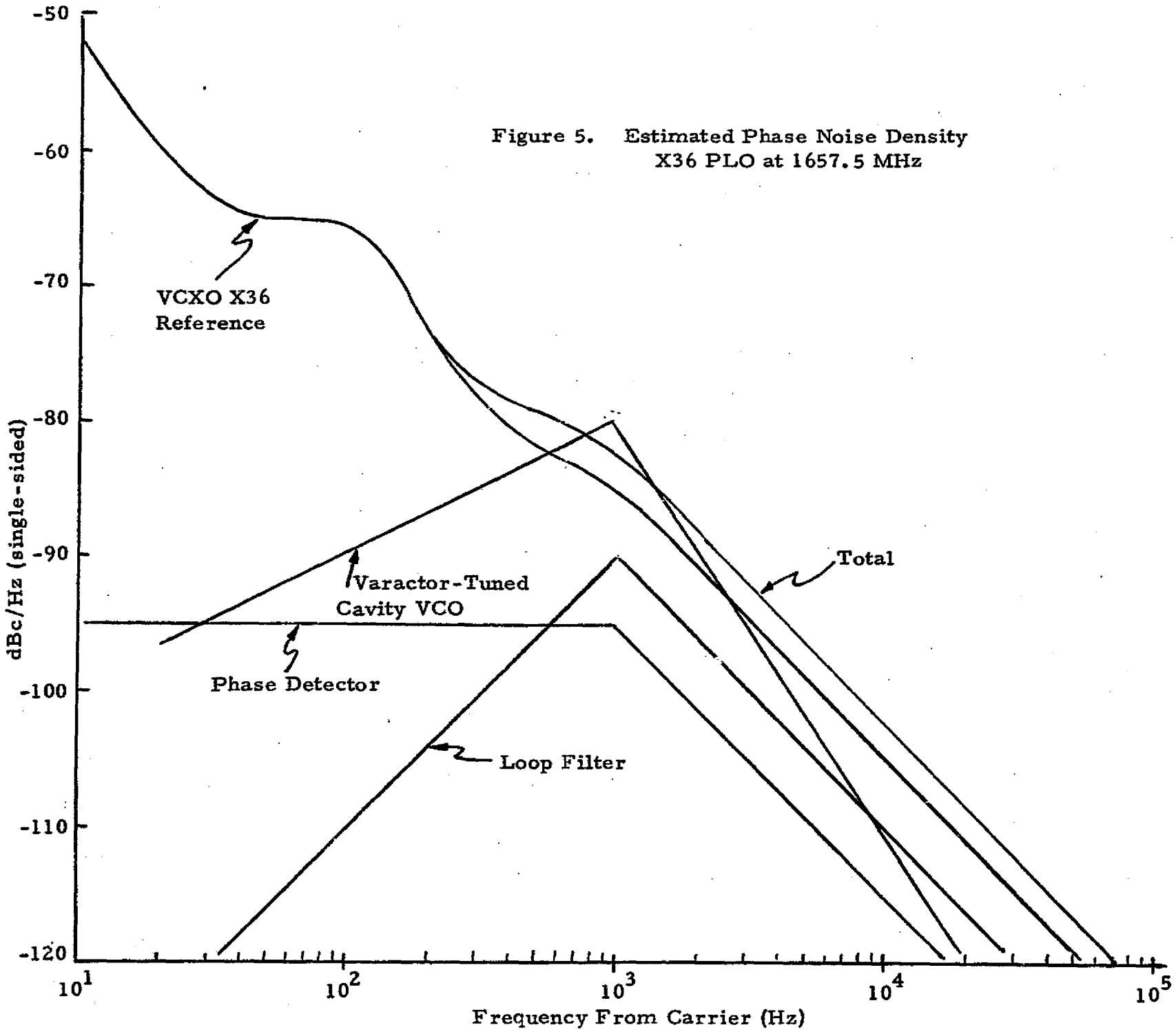


Figure 4. Estimated Phase Noise Density for VCXO Loop

Figure 5. Estimated Phase Noise Density
X36 PLO at 1657.5 MHz



spectrum at the low frequency. The high frequency spectrum will be attenuated by the loop filter. To this filtered VCXO spectrum must be added the effect of the VCO and the phase detected of the PLO as filtered by the loop filter similar to the procedure used in Figure 4.

Using Figures 4 and 5, the total phase noise spectrum for the frequency synthesizer is obtained in Figure 6. First, the VCXO output phase noise from Figure 4 is multiplied by 5 and added in at the data modulation input. Second, the VCXO multiplied by the X36 PLO from Figure 5 is added. Finally, the output of the VCXO is passed through the variable multiplying PLO in the upper channel of Figure 3. The worst case for the low frequency spectrum in this channel is when the multiplication of the VCXO is 1361/135. Again the loop filter attenuates the high frequency of the multiplied VCXO spectrum. The other contributions of the phase noise are the loop filtered phase detector and the VCO, with its low frequency attenuated by the loop filter. By adding all these contributions, the total phase noise spectrum at the output of the frequency synthesizer is shown in Figure 6.

In the remainder of the appendix, the phase noise is referred to as a random process with zero mean and a variance σ_ψ^2. To determine σ_ψ^2 from the total phase noise spectrum in Figure 6, note that, for a zero mean random process,

$$\sigma_\psi^2 = \int_{-\infty}^{\infty} S_\psi(\omega)\, d\omega \tag{1}$$

where $S_\psi(\omega)$ is the power spectral density of the phase noise process. It remains to be shown that the oscillator output power below the carrier (i.e., the center frequency) in Figure 6 is actually the power spectral density of the phase noise process.

A waveform of frequency f_c and constant amplitude A can be expressed as

$$V = A \sin [2\pi f_c t + \theta(t)] \tag{2}$$

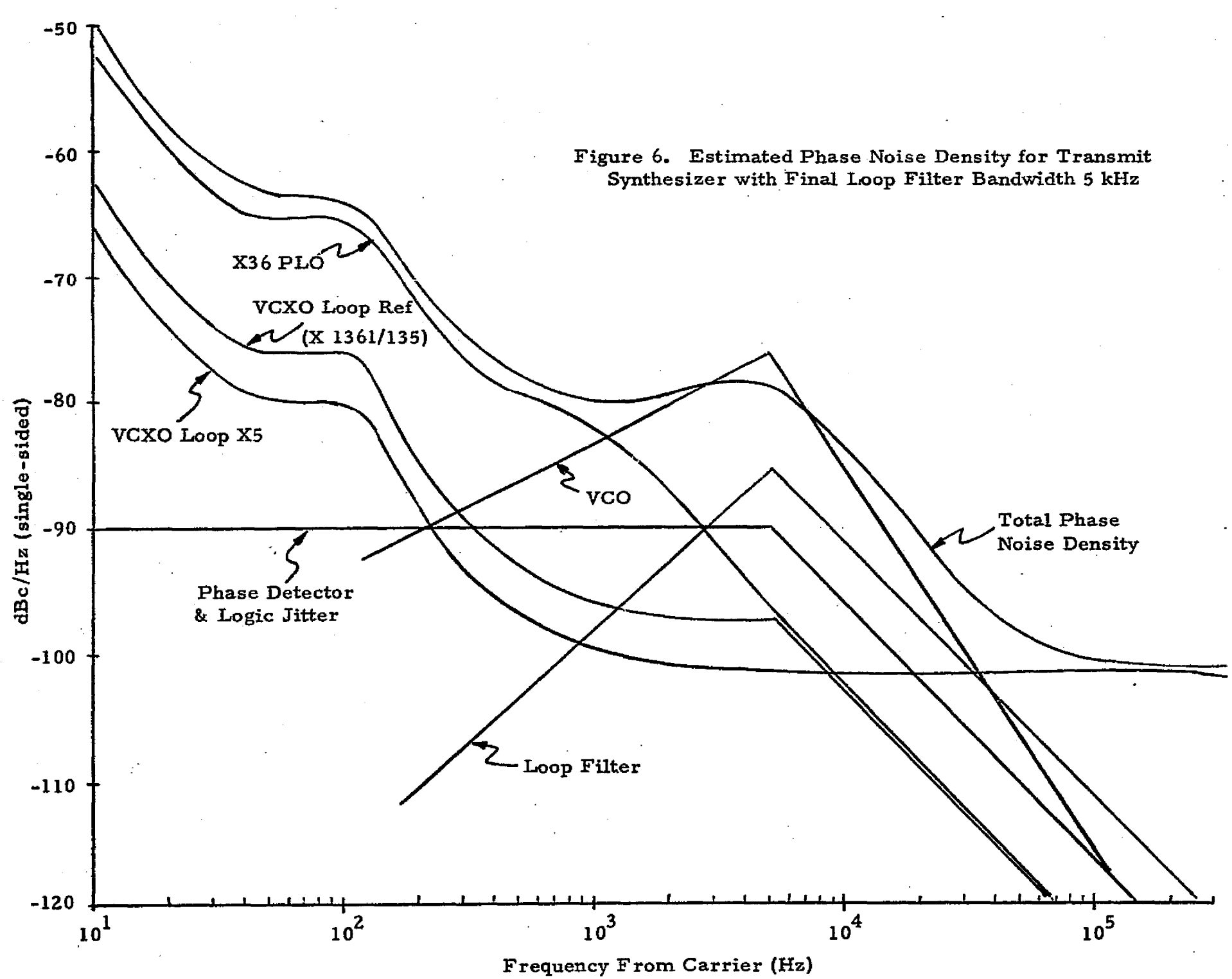


Figure 6. Estimated Phase Noise Density for Transmit Synthesizer with Final Loop Filter Bandwidth 5 kHz

where $\theta(t)$ represents the phase variations about the linearly progressing phase, $2\pi f_c$. The instantaneous frequency is defined by

$$f = f_c + \frac{1}{2\pi}\frac{d\theta(t)}{dt} \quad \text{(Hertz)} . \tag{3}$$

Often the phase or frequency modulation of the carrier can be represented as a sinusoidally varying quantity

$$\theta(t) = \Delta\varphi \sin 2\pi f_m t \tag{4}$$

where $\Delta\varphi$ is the peak phase deviation in radians, and (2) can be written as

$$V = A \sin (2\pi f_c + \Delta\varphi \sin 2\pi f_m t) . \tag{5}$$

Using a Bessel function series expansion, (5) can be written to emphasize the frequency components of a simple phase modulated signal. With $\omega = 2\pi f$,

$$\frac{V}{A} = J_0(\Delta\varphi) \sin \omega_c t + J_1(\Delta\varphi)[\sin (\omega_c + \omega_m)t - \sin (\omega_c - \omega_m)t]$$

$$+ J_2(\Delta\varphi)[\sin (\omega_c + 2\omega_m)t + \sin (\omega_c - 2\omega_m)t] + \ldots . \tag{6}$$

Now assuming that $\Delta\varphi \ll 1$, which is reasonable for a good oscillator, the Bessel function $J_0(\Delta\varphi) \simeq 1$, $J_1 \simeq \frac{\Delta\varphi}{2}$, and J_2, J_3, etc., are negligible. Thus, the waveform appears in the frequency domain as a carrier at f_c with sidebands spaced at $f_c + f_m$ and $f_c - f_m$. The ratio of one sideband amplitude to the carrier amplitude is $\frac{\Delta\varphi}{2}$. Therefore, if the power of spectrum below the carrier is known, then

$$\Delta\varphi = 2\sqrt{\frac{P_{sb}}{P_c}} \tag{7}$$

where P_{sb} is the power in the sideband $f_c + f_m$ and P_c is the power in the carrier. If more than one sideband exists, then the total phase deviation would be the sum of the square root of the ratios of power in each sideband divided by the carrier power. Therefore, the spectrum estimated in

Figure 6 is the power spectral density of the phase noise process and the variance is determined by equation (1).

It will be shown in the following section that the variance of the phase noise process outside the loop bandwidth of transponder affects the performance of the forward link and the phase noise inside the loop bandwidth affects the performance of the return link. The loop bandwidth (double sided) of the DSN transponder is 18 Hz. Therefore, for the forward link, the phase noise spectrum can be integrated from 9 Hz to, say, 1 MHz and then doubled for the two-sided bandwidth to determine the phase variance. For the phase noise power spectral density in Figure 6, the rms phase noise σ_ψ is 2.7° for a double sided bandwidth from ± 9 Hz to ± 1 MHz.

Two additional points must be made about the phase noise spectrum in Figure 6. First, with the final loop bandwidth equal to 5 kHz in the upper channel of Figure 3, spur suppression is difficult with a digital phase detector even though it is fairly easy with a sample-and-hold detector. To alleviate this difficulty, the final loop filter could be decreased in bandwidth to 2 kHz. This would make spur suppression easy with either type of phase detector but would increase the rms phase noise. Figure 7 presents the phase noise power spectral density with the loop bandwidth at 2 kHz. The resulting rms phase noise σ_ψ is 3.4° for a double sided bandwidth from ± 9 Hz to ± 1 MHz. Finally, to simplify the frequency synthesizer, a X36 multiplier could be used instead of the X36 PLO. Figure 8 presents the phase noise spectral density for the X36 multiplier with no PLO. The rms phase noise for this simplified frequency synthesizer is $\sigma_\psi = 5.9^\circ$ for a double sided bandwidth from ± 9 Hz to ± 1 MHz.

3.0 PHASE ERROR MODEL

As shown in Figure 9, the received signal Z(t) is

$$Z(t) = s(t) + n(t) \tag{8}$$

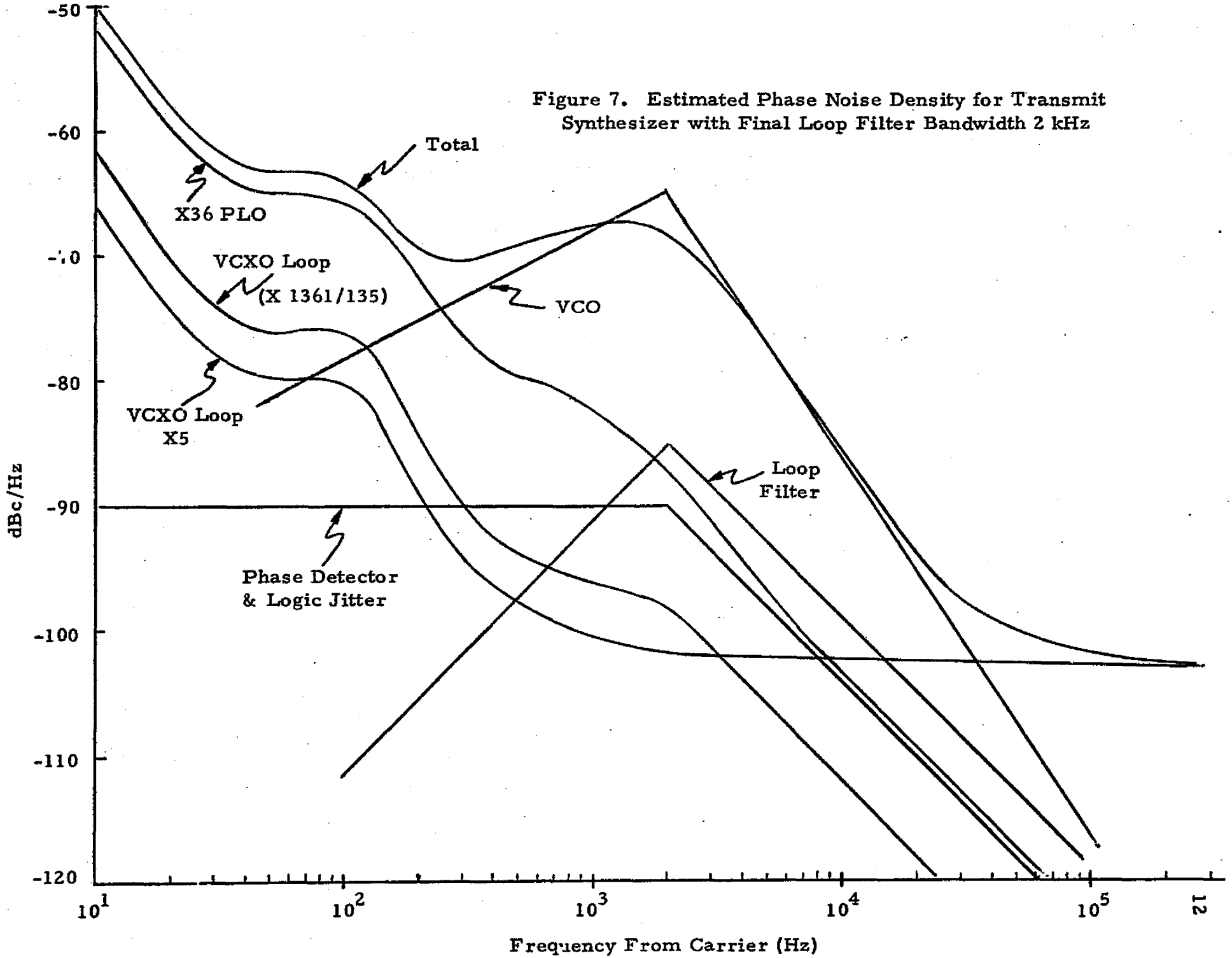
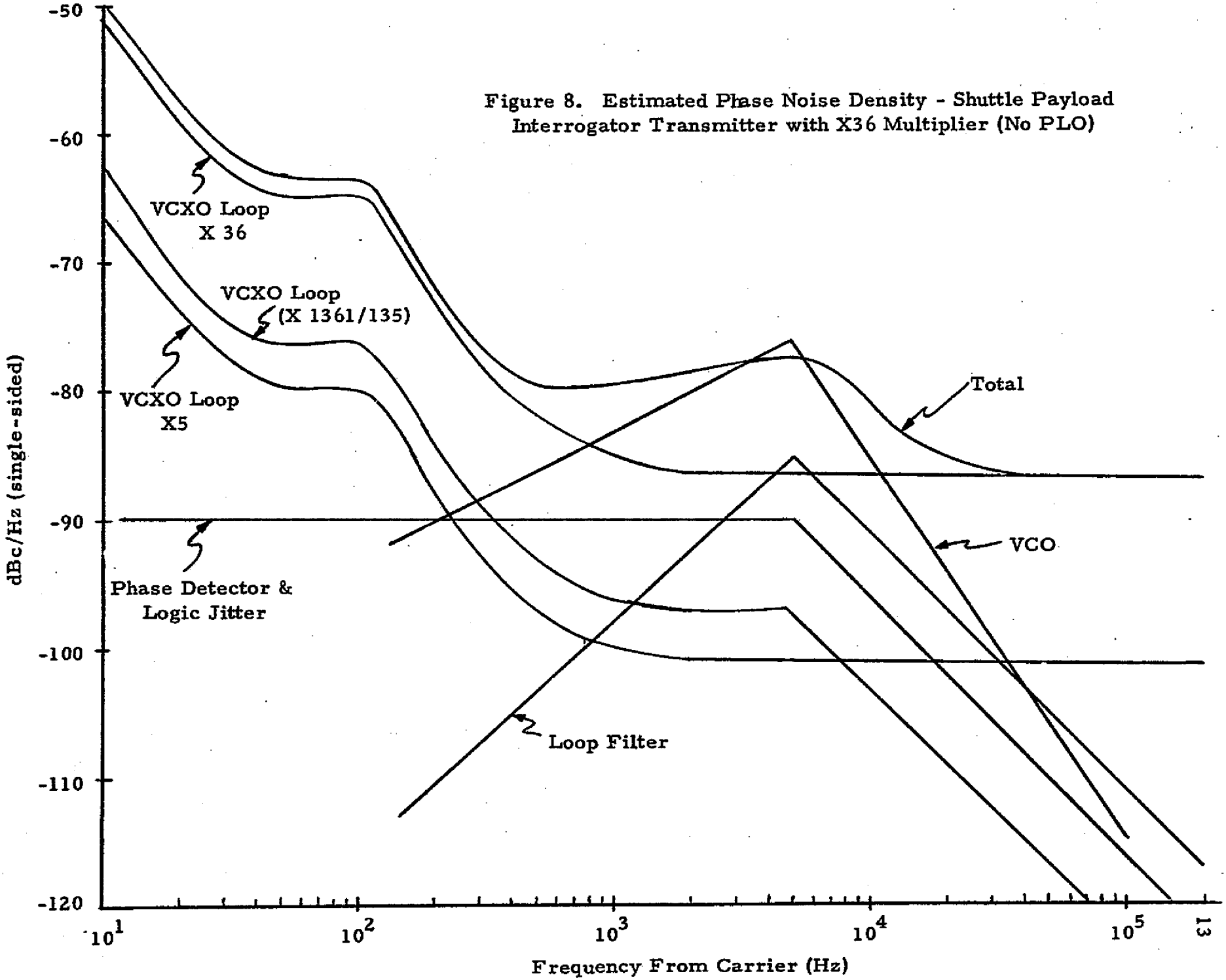


Figure 7. Estimated Phase Noise Density for Transmit Synthesizer with Final Loop Filter Bandwidth 2 kHz

Figure 8. Estimated Phase Noise Density - Shuttle Payload Interrogator Transmitter with X36 Multiplier (No PLO)



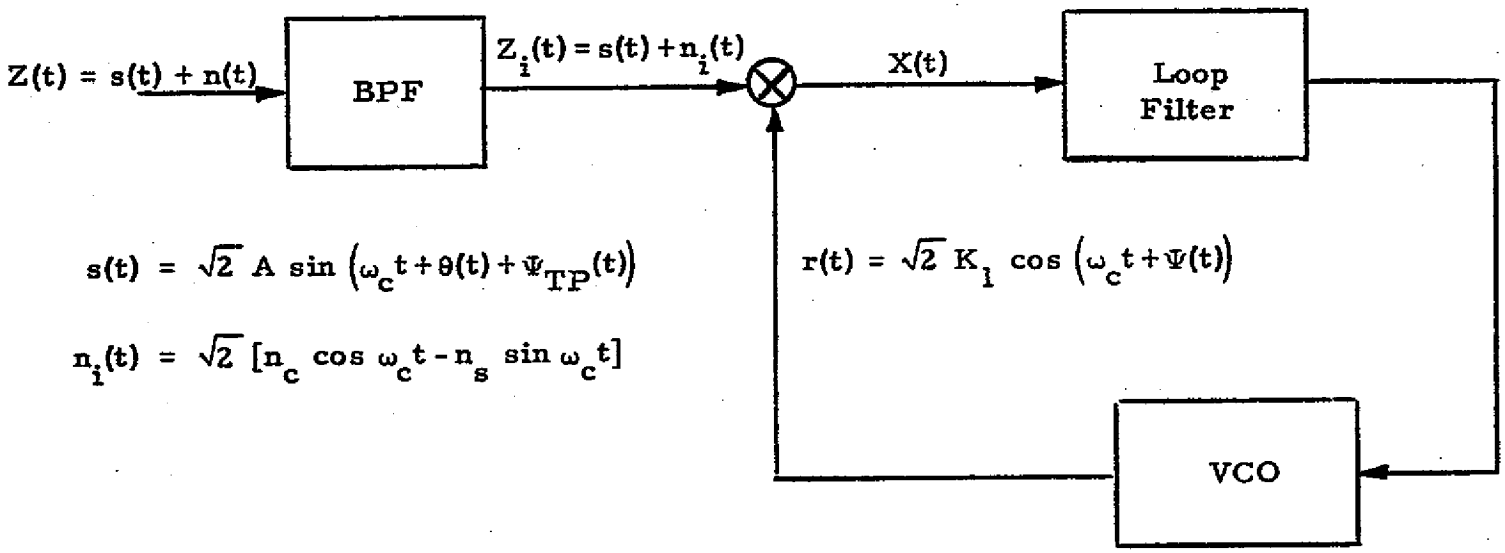


Figure 9. Carrier Tracking Loop

where $s(t) = \sqrt{2}\ A \sin\left(\omega_c t + \theta(t) + \Psi_{TP}(t)\right)$

ω_c = angular carrier frequency

$\theta(t)$ = modulating process

$\Psi_{TP}(t)$ = transmitted phase jitter

A^2 = total received power = P_{rec}

and $n(t)$ = channel white Gaussian noise with flat spectral density N_0 (one-sided) over the frequency range of the receiver.

The bandpass filter with bandwidth B_{IF} centered at carrier frequency is used to reject as much noise as possible to the phase-locked loop (PLL) without distorting the signal. The output of the bandpass filter can then be given by

$$Z_i(t) = s(t) + n_i(t) \tag{9}$$

where the filtered noise $n_i(t)$ can be expressed in the form of a narrowband process [1] about the nominal center frequency of the bandpass filter, i.e.,

$$n_i(t) = \sqrt{2}\ [n_c(t) \cos \omega_c t - n_s(t) \sin \omega_c t]\,. \tag{10}$$

Both $n_c(t)$ and $n_s(t)$ are statistically independent Gaussian noise processes with identical symmetrical spectra corresponding to the lowpass equivalent of $n(t)$, i.e.,

$$\overline{n_c^2(t)} = \overline{n_s^2(t)} = N_0\left(\frac{B_{IF}}{2}\right) = \frac{1}{2}\ \overline{n_i^2(t)}\ . \tag{11}$$

Assuming that frequency lock has been achieved, the expression for the multiplier output can be written as

$$X(t) = K_m Z_i(t)\, r(t) \tag{12}$$

where K_m = multiplier gain

$r(t)$ = output of the voltage controlled oscillator (VCO)

$= \sqrt{2}\ K_1 \cos\left(\omega_c t + \hat{\Psi}(t)\right)$

K_1 = rms voltage of the VCO output

and $\hat{\Psi}(t)$ = PLL estimate of $\Psi_{TP}(t)$.

Since the PLL cannot respond to the double-frequency terms, we have

$$X(t) = AK_1K_m \sin\left(\theta(t) + \Psi_{TP}(t) - \hat{\Psi}(t)\right) + K_1K_m n_c(t) \cos\hat{\Psi}(t) + K_1K_m n_s(t) \sin\hat{\Psi}(t) \ . \tag{14}$$

Let $\Psi(t) = \Psi_{TP}(t) - \hat{\Psi}(t)$ = phase error; we obtain

$$X(t) = K_1K_m \left[\left(A\cos\theta(t) - N_s(t)\right)\sin\Psi(t) + \left(A\sin\theta(t) + N_c(t)\right)\cos\Psi(t)\right] \tag{15}$$

where

$$N_c(t) = n_c(t)\cos\Psi_{TP}(t) + n_s(t)\sin\Psi_{TP}(t) \tag{16}$$

and

$$N_s(t) = n_s(t)\cos\Psi_{TP}(t) - n_c(t)\sin\Psi_{TP}(t) \tag{17}$$

The modulating process can be generalized as

$$\theta(t) = \beta_d d(t) + \sum_{i=1}^{m} \beta_{si} \sin\left(\omega_{si}(t) + \theta_{si}(t)\right) \tag{18}$$

where β_d = baseband data channel modulation index

$d(t) = \pm 1$ = baseband digital data, bi-ϕ-L format

R_d = data rate of $d(t)$

β_{si} = subcarrier channel i modulation index

ω_{si} = channel i angular subcarrier frequency

and $\theta_{si}(t)$ = subcarrier channel i modulating process .

By substituting (18) into (15) and noticing that the loop is narrow-band (as compared to R_d) and cannot respond to high frequency terms, we obtain

$$X(t) = K_1K_m\left[A' \sin\Psi(t) + I\left(t, \Psi(t)\right) + N\left(t, \Psi(t)\right)\right] \ . \tag{19}$$

where $$A' = A \cos \beta_d \prod_{i=1}^{m} J_0(\beta_{si}) \tag{20}$$

$J_0(X)$ = Bessel function of first kind of zero order

$$I\left(t, \Psi(t)\right) = A\, d(t) \sin \beta_d \prod_{i=1}^{m} J_0(\beta_{si}) \cos \Psi(t) \tag{21}$$

and $$N\left(t, \Psi(t)\right) = N_c(t) \cos \Psi(t) - N_s(t) \sin \Psi(t) . \tag{22}$$

From (19), it can be seen that the first term is the loop phase error control signal, the second term is phase distortion due to modulating signal d(t), and the third term is the noise. Under certain physically justifiable assumptions, the noise $N\left(t, \Psi(t)\right)$ can be shown [2] to be approximately a zero mean white Gaussian noise process with the same spectral density as that of the original additive process $n_i(t)$.

The stochastic integro-differential equation that relates the VCO phase output to X(t) is given by

$$\hat{\Psi}(t) = \frac{K_V F(P)}{P} X(t) + \Psi_2(t) \tag{23}$$

where $P = \frac{d}{dt}$ = Heaviside operator

$F(P)$ = loop filter transfer function

K_V = gain of the VCO (radians/volt-sec)

and $\Psi_2(t)$ = VCO self-generated phase jitter.

Since $\Psi(t) = \Psi_{TP}(t) - \hat{\Psi}(t)$, we rewrite equation (23) to

$$\Psi(t) = \Psi_{TP}(t) - \frac{K_V F(P)}{P} X(t) - \Psi_2(t) \tag{24}$$

or

$$\Psi(t) = \Psi_{TP}(t) - \frac{A'K F(P)}{P} \left[\sin \Psi(t) + \frac{I\left(t, \Psi(t)\right)}{A'} + \frac{N\left(t, \Psi(t)\right)}{A'}\right] - \Psi_2(t) \tag{25}$$

where $K = K_1 K_m K_V$ = open loop gain.

A mathematic equivalent phase error model can be conveniently represented by the block diagram of Figure 10.

For practical purposes, the carrier tracking loop is usually designed with sufficient care so that the tracking phase error $\Psi(t)$ will be small ($< \pi/6$). If this is the case, then the linear model can be applied (i.e., $\sin \Psi(t) \cong \Psi(t)$; $\cos \Psi(t) \cong 1$; $N_c(t) \gg N_s(t)\Psi(t)$), and (25) can be reduced to

$$\Psi(t) \cong \Psi_{TP}(t) - \frac{A'K\,F(P)}{P}\left[\Psi(t) + \frac{I'(t)}{A'} + \frac{N_c(t)}{A'}\right] - \Psi_2(t) \tag{26}$$

where

$$I'(t) = A\,d(t) \sin \beta_d \prod_{i=1}^{m} J_0(\beta_{si}) . \tag{27}$$

A slight rearrangement of equation (26) relates the phase error to the input processes:

$$\Psi(t) = [1 - H(P)]\Psi_{TP}(t) - [1 - H(P)]\Psi_2(t) - H(P)\frac{I'(t)}{A'} - H(P)\left(\frac{N_c(t)}{A'}\right) \tag{28}$$

where

$$H(P) = \frac{A'K\,F(P)}{P + A'K\,F(P)} \tag{29}$$

is the closed-loop transfer function and the two-sided noise loop bandwidth is defined by [3],

$$2B_L = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left|H(j\omega)\right|^2 d\omega . \tag{30}$$

From (28), we can see that the total phase error is contributed by two in-band processes, $I'(t)$ and $N_c(t)$, and two out-band processes, $\Psi_{TP}(t)$ and $\Psi_2(t)$.

Generally speaking, the data rate of $d(t)$, R_d, is designed well above the loop bandwidth B_L. Therefore, since the data is bi-ϕ-L with no d.c. component, the contribution to the phase error by $d(t)$ can be neglected. It is also well known that the probability density function of Ψ due to

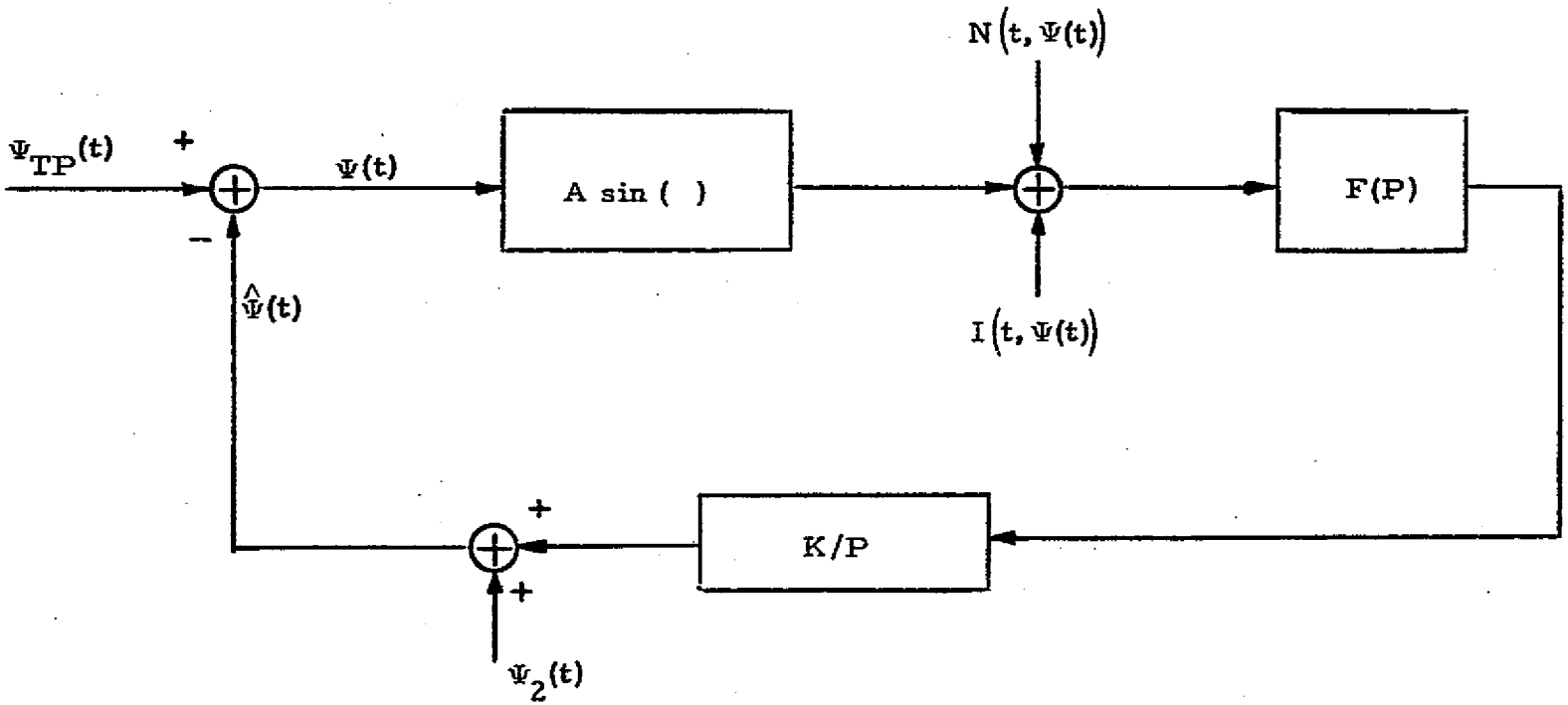


Figure 10. An Equivalent Phase Error Model

noise $N_c(t)$ for a second-order loop can be expressed as [4]:

$$p_{N_c}(\Psi) = \frac{e^{\alpha \cos \Psi}}{2\pi I_0(\alpha)}, \qquad -\pi < \Psi < \pi \tag{31}$$

where

$$\alpha = \frac{1}{\sigma^2_{N_c}} \tag{32}$$

$\sigma^2_{N_c}$ = variance of phase error due to $N_c(t)$

and $I_0(x)$ = modified Bessel function of zero order.

From (28), we obtain

$$\sigma^2_{N_c} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 \frac{S_{N_c}(\omega)}{A'^2} \, d\omega \tag{33}$$

where $S_{N_c}(\omega)$ is the power spectral density of $N_c(t)$ and is $N_0/2$ in the loop bandwidth B_L. Thus,

$$\sigma^2_{N_c} = \frac{N_0}{2A'^2} \cdot 2B_L = \frac{N_0 B_L}{A'^2} \tag{34}$$

Since $A' = A \cos \beta_d \prod_{i=1}^{m} J_0(\beta_{si})$ and $A^2 = P_{rec}$, we have

$$\sigma^2_{N_c} = \frac{N_0 B_L}{P_{rec} \cos^2 \beta_d \prod_{i=1}^{m} J_0^2(\beta_{si})} \tag{35}$$

and

$$\alpha = \frac{1}{\sigma^2_{N_c}} = \frac{P_{rec} \cos^2 \beta_d \prod_{i=1}^{m} J_0^2(\beta_{si})}{N_0 B_L} \tag{36}$$

which is the signal-to-noise ratio (SNR_L) in the loop bandwidth. It should be noted here that the numerator is the total power allocated to the carrier component.

For large SNR_L, say greater than 12 dB, we can simplify (31) to

$$p_{N_c}(\Psi) = \frac{1}{\sqrt{2\pi}\,\sigma_{N_c}} \exp\left[-\frac{\Psi^2}{2\sigma_{N_c}^2}\right], \tag{37}$$

a Gaussian distribution with mean zero and variance $\sigma_{N_c}^2$.

Since $\Psi_{TP}(t)$ and $\Psi_2(t)$ can be assumed to be zero-mean Gaussian processes, we conclude that the phase error $\Psi(t)$ is also a zero-mean Gaussian process if $SNR_L \geq 12$ dB and $R_d \gg B_L$. The variance of (t) under these conditions can be computed as

$$\sigma_\Psi^2 = \sigma_{\Psi_{TP}}^2 + \sigma_{N_c}^2 + \sigma_{\Psi_2}^2 \tag{38}$$

where

$$\sigma_{\Psi_{TP}}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left|1 - H(j\omega)\right|^2 S_{\Psi_{TP}}(\omega)\, d\omega \tag{39}$$

$$\sigma_{\Psi_2}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left|1 - H(j\omega)\right|^2 S_{\Psi_2}(\omega)\, d\omega . \tag{40}$$

$S_{\Psi_{TP}}(\omega)$ and $S_{\Psi_2}(\omega)$ are power spectral densities of $\Psi_{TP}(t)$ and $\Psi_2(t)$, respectively.

Now, if we rewrite equation (28) in terms of $\hat{\Psi}(t)$, we have

$$\begin{aligned}\Psi(t) &= \Psi_{TP}(t) - \hat{\Psi}(t) \\ &= [1 - H(P)]\Psi_{TP}(t) - [1 - H(P)]\Psi_2(t) - H(P)\left(\frac{I'(t)}{A'}\right) - H(P)\left(\frac{N_c(t)}{A'}\right)\end{aligned} \tag{41}$$

or

$$\hat{\Psi}(t) = H(P)\Psi_{TP}(t) + [1 - H(P)]\Psi_2(t) + H(P)\left(\frac{I'(t)}{A'}\right) + H(P)\left(\frac{N_c(t)}{A'}\right) . \tag{42}$$

Notice that $\hat{\Psi}(t)$ is the turnaround phase jitter to be imposed on the return link. From (42), we can see that $\hat{\Psi}(t)$ is due to contributions by the "in-band" transmitted jitter, thermal noise and modulation interference, and out-band contribution by the VCO phase jitter. The narrower the loop bandwidth B_L, the higher the phase error, but the less turnaround phase jitter and phase error contributed by the thermal noise, and vice versa. Therefore, a tradeoff among the loop bandwidth, incoming transmitted phase jitter, and thermal noise can be made to equalize the performance of both links to the limit of the validity of the linear model in equation (26).

4.0 DATA DETECTION AND BIT-ERROR PROBABILITY

If we pass the output of the PLL multiplier through a divider with the gain $1/(K_1 K_m)$ and a lowpass filter with the bandwidth just to pass d(t) as shown in Figure 11, the output of the lowpass filter can then be expressed as

$$d'(t) = A \sin \beta_d \left[\prod_{i=1}^{m} J_0(\beta_{si})\right] \cos \Psi(t)\, d(t) + A \cos \beta_d \left[\prod_{i=1}^{m} J_0(\beta_{si})\right] \sin \Psi(t) + N'\left(t, \Psi(t)\right) \qquad (43)$$

where $N'\left(t,\Psi(t)\right)$ is the filtered version of $N\left(t,\Psi(t)\right)$ as given by (22). The first term is the desired signal, the second term is the interference due to loop error, and the third term is the noise. We can approximate $\sin \Psi(t) \cong \Psi(t)$. Since $\Psi(t)$ is also a lowpass Gaussian process, we can rewrite (43) as

$$d'(t) = A \sin \beta_d \prod_{i=1}^{m} J_0(\beta_{si}) \cos \Psi(t)\, d(t) + n'(t) \qquad (44)$$

where $n'(t)$ is a zero-mean Gaussian process with the power (N_p) in the data rate (R_d) bandwidth to be given by

$$N_p = N_0 R_d + \sigma_\Psi^2 A^2 \cos^2 \beta_d \prod_{i=1}^{m} J_0^2(\beta_{si}) . \qquad (45)$$

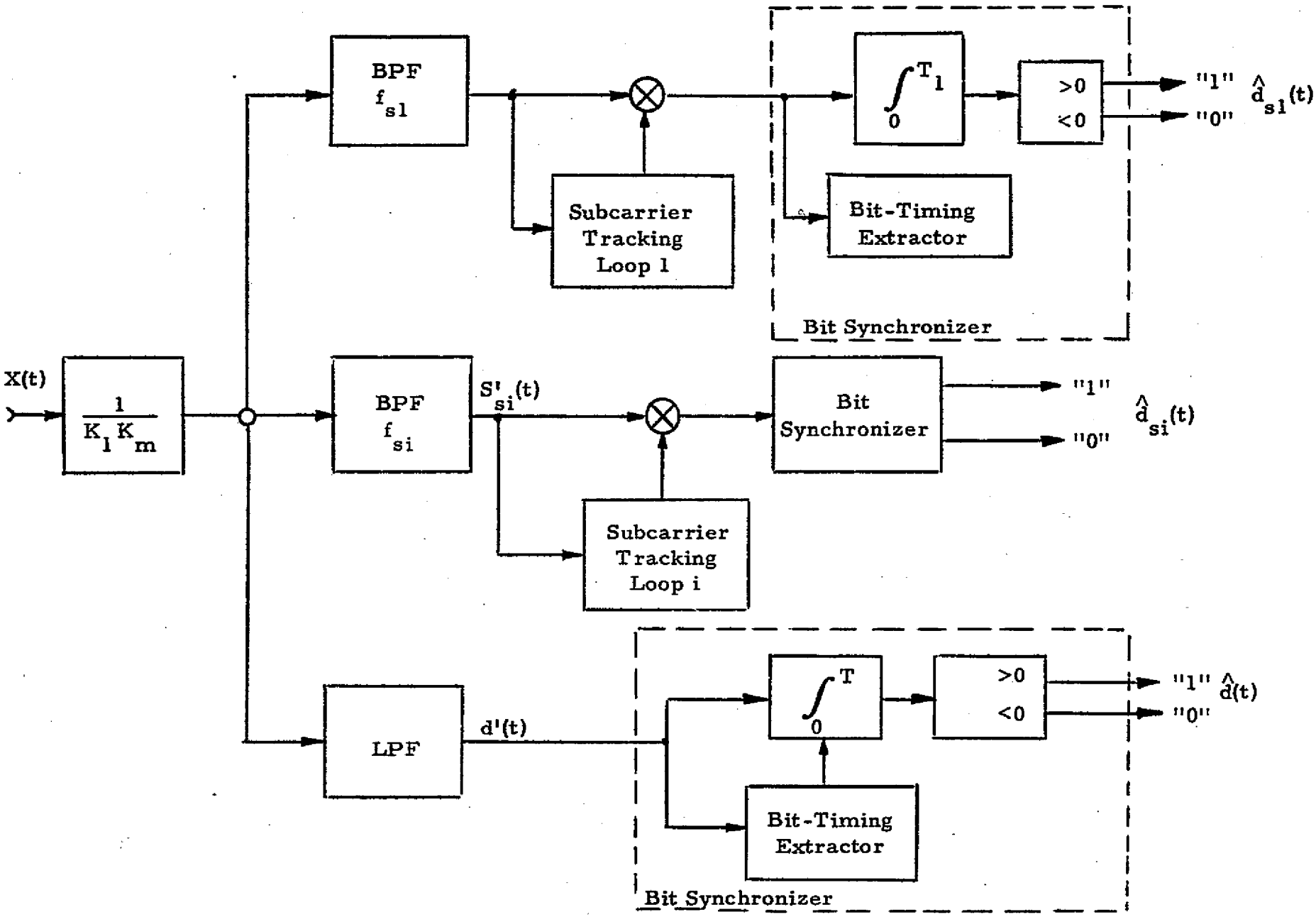


Figure 11. Data Demodulation Configuration

We assume that all the power of $\Psi(t)$ is in the data rate bandwidth.

For $R_d \gg B_L$, the most practical designs, we can also assume $\Psi(t)$ to be constant during the data period $(1/R_d)$. In this case the conditional bit-error probability by detecting d(t) with a matched filter (integrate-and-dump circuit) can be expressed as [5]:

$$P_e(\Psi) = \frac{1}{2} \operatorname{erfc}\left[\sqrt{\frac{P'_{rec} \cos^2\Psi}{N_p}}\right] \tag{46}$$

where erfc (x) = complementary error function = $\frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$ and

$P'_{rec} = A^2 \sin^2 \beta_d \prod_{i=1}^{m} J_0^2(\beta_{si})$.

Substituting (45) into (46), we have

$$P_e(\Psi) = \frac{1}{2} \operatorname{erfc}\left[\left(\frac{A^2 \sin^2 \beta_d \prod_{i=1}^{m} J_0^2(\beta_{si}) \cos^2\Psi}{N_0 R_d + \sigma_\Psi^2 A^2 \cos^2 \beta_d \prod_{i=1}^{m} J_0^2(\beta_{si})}\right)^{1/2}\right] \tag{47}$$

or

$$P_e(\Psi) = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_{b_d} \cos^2\Psi}{N_0 + N'}}\right) \tag{48}$$

where

$$E_{b_d} = \frac{A^2 \sin^2 \beta_d \prod_{i=1}^{m} J_0^2(\beta_{si}) \cos^2\Psi}{R_d} \tag{49}$$

= Energy per bit of baseband channel

and

$$N' = \frac{\sigma_\Psi^2 A^2 \cos^2 \beta_d \prod_{i=1}^{m} J_0^2(\beta_{si})}{R_d} \tag{50}$$

= Equivalent noise spectral density due to phase error.

Since the probability density function of $\Psi(t)$ is Gaussian distributed

$$p(\Psi) = \frac{1}{\sqrt{2\pi}\,\sigma_\Psi} \exp\left[-\frac{\Psi^2}{2\sigma_\Psi^2}\right], \tag{51}$$

the bit-error probability can now be computed by

$$\begin{aligned} P_e &= \int_{-\pi}^{\pi} p(\Psi)\, P_e(\Psi)\, d\Psi \\ &= \int_{-\pi}^{\pi} \left(\frac{1}{\sqrt{2\pi}\,\sigma_\Psi} \exp\left[-\frac{\Psi^2}{2\sigma_\Psi^2}\right]\right) \frac{1}{2}\, \mathrm{erfc}\left[\sqrt{\frac{E_b}{N}}\, d \cos\Psi\right] d\Psi \end{aligned} \tag{52}$$

where $N = N_0 + N'$ = equivalent total noise spectral density.

Figure 12 presents a set of bit-error probability curves as a function of energy per bit/noise spectral density (E_b/N) for various values of σ_Ψ in degrees. If power spectral densities $S_{\Psi_{TP}}(\omega)$ and $S_{\Psi_2}(\omega)$ are known, σ_Ψ^2 can be determined from (37) and (38).

Using the phase noise power spectral density in Figure 6, the rms phase error $\sigma_{\Psi_{TP}} = 2.7^\circ$. The DSN transponder design requirement [6] is for a maximum of 1° static phase error and $\sigma_{\Psi_2} = 3^\circ$. Therefore, the total phase error will be less than 6.7° and the degradation due to phase noise from Figure 12 at $P_e = 10^{-4}$ is less than 0.1 dB. Considering the return link, the turnaround phase noise is described by (42). For a 18 Hz loop, it may be assumed that a major portion of the phase jitter is contributed by $\sigma_{\Psi_2} = 3^\circ$. At the interrogator receiver, a similar frequency synthesizer is used as in the transmitter. Therefore, at the receiver, there is a received phase noise with $\sigma_\Psi = 3^\circ$ plus $\sigma_{\Psi_s} = 2.7^\circ$ from the receiver frequency synthesizer and $\sigma_\Psi = 3^\circ$ from the receiver phase lock loop. With a total phase error $\sigma_\Psi = 8.7^\circ$, Figure 12 shows that the degradation due to phase noise for the two-way link is less than 0.2 dB for $P_e = 10^{-4}$.

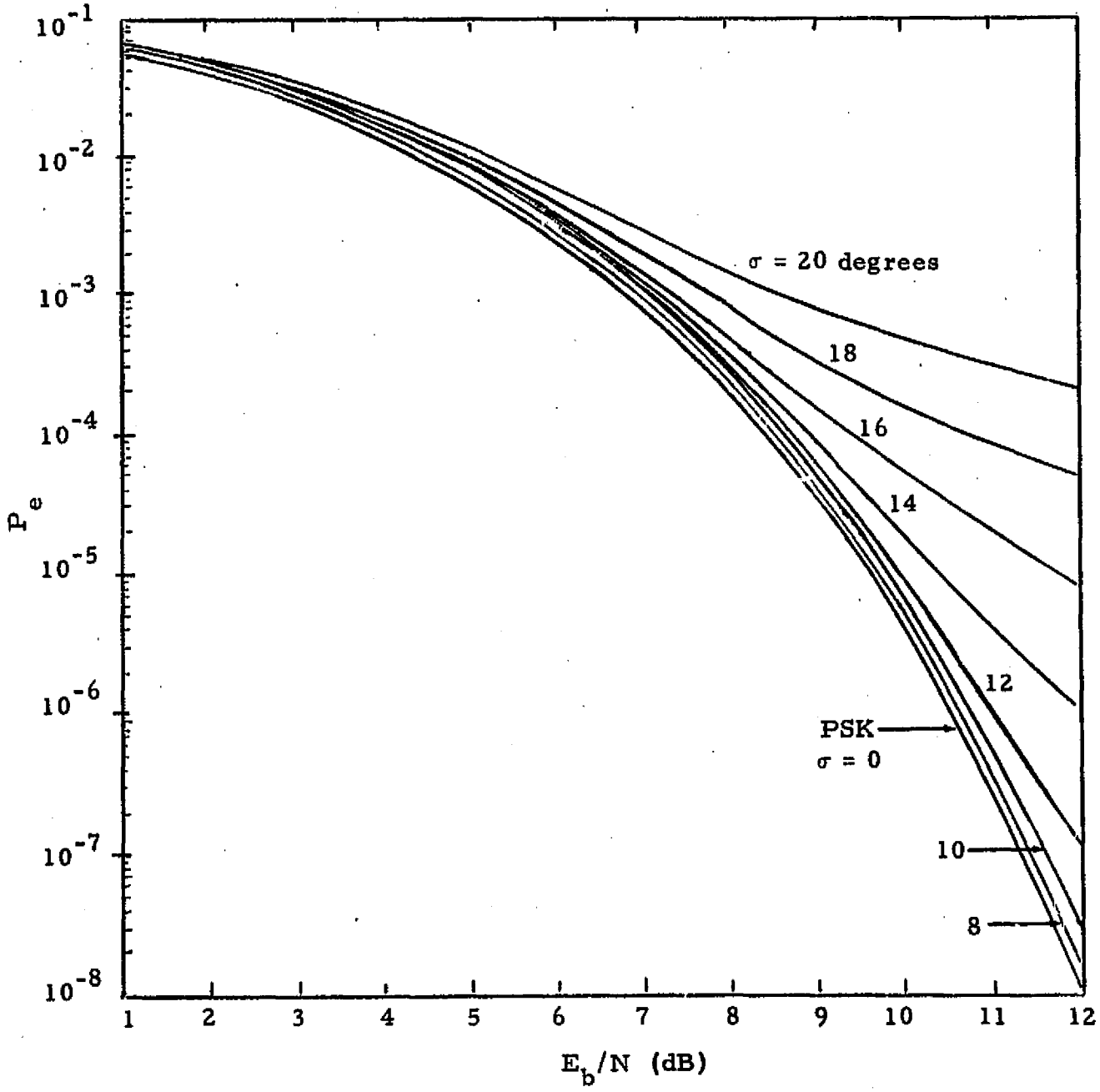


Figure 12. Bit-Error Probability

As shown in Figure 11, the output of the ith subcarrier bandpass filter can be given by

$$S'_{si}(t) = A \cos \beta_d \left(2 J_1(\beta_{si})\right)\left[\prod_{\substack{j=1 \\ j\neq i}}^{m} J_0(\beta_{sj})\right] \sin \left(\omega_{si} t + \theta_{si}(t)\right)$$

$$\cdot \cos \Psi(t) + P(t) + N'\left(t, \Psi(t)\right) \tag{53}$$

where $P(t)$ and $N'\left(t, \Psi(t)\right)$ are the filtered versions of

$$A \sin \beta_d \, d(t) \left(2 J_1(\beta_{si})\right)\left[\prod_{\substack{j=1 \\ j\neq i}}^{m} J_0(\beta_{sj})\right] \sin \left(\omega_{si} t + \theta_{si}(t)\right) \sin \Psi(t)$$

and $N\left(t, \Psi(t)\right)$, respectively.

The first term is the desired signal, the second term is the interference due to $\sin \Psi(t)$, and the third term is the noise. Since $\sin \Psi(t)$ is is a lowpass process, the spectrum of $\sin \Psi(t)$ spilled over the bandpass filter bandwidth is extremely low. Thus, for all practical purposes, the second term in equation (53) can be neglected. In this appendix, we only consider the cases where the subcarriers are phase-shift-keyed by the binary data stream, i.e.,

$$\theta_{si}(t) = d_{si}(t) = \pm 1 .$$

Equation (53) can now be rewritten as

$$S'_{si}(t) = A \cos \beta_d \left(2 J_1(\beta_{si})\right) \left[\prod_{\substack{j=1 \\ j\neq i}}^{m} J_0(\beta_{sj})\right] d_{si}(t) \sin \omega_{si} t \cos \Psi(t) + N'\left(t, \Psi(t)\right) . \tag{54}$$

A subcarrier tracking loop, usually a squaring loop or Costas loop [5], is employed in series with a bit synchronizer (integrate-and-dump circuit) for choherent detection of $d_{si}(t)$. Generally, the subcarrier tracking loop is designed to operate in the high signal-to-noise ratio region and narrow loop filter range (as compared to carrier tracking loop). The phase

error introduced by the subcarrier loop can usually be neglected. Thus, the bit-error probability can be obtained the same way as that of the base-band channel:

$$P_e = \int_{-\pi}^{\pi} \left(\frac{1}{\sqrt{2\pi}\,\sigma_\Psi} \exp\left[- \frac{\Psi^2}{2\sigma_\Psi^2} \right] \right) \frac{1}{2} \operatorname{erfc}\left[\sqrt{\frac{E_{b_{si}}}{N_0}} \cos\Psi \right] d\Psi \qquad (55)$$

where

$$E_{b_{si}} = \frac{A^2 \cos^2 \beta_d \left(2 J_1(\beta_{si})\right)^2 \left[\prod_{\substack{j=1 \\ j \neq i}}^{m} J_0^2(\beta_{sj}) \right]}{R_{si}}$$

$= $ energy per bit of channel i

and $R_{si} =$ data rate of channel i .

Figure 12 can also be utilized to determine the bit-error probability except that N_0 replaces N.

5.0 CONCLUSION

The effect of transmitter phase noise on data transmission for a two-way link has been analyzed. It was shown that there is a tradeoff between the effect of phase noise on the forward link and the phase noise turnaround for the return link. As a final result, the performance degradation to the bit error probability was presented as a function of the total rms phase error in degrees. As a typical example of a two-way link, the S-band payload interrogator in the DSN mode communicating with the DSN standard transponder was analyzed. It was found that at bit error probability of 10^{-4}, the performance degradation was less than 0.1 dB and 0.2 dB for the forward link and the total two-way link, respectively.

REFERENCES

1. Viterbi, A. J. Principles of Coherent Communication. McGraw-Hill, 1966, pp. 7-36.

2. Lindsey, W. C. Synchronization Systems in Communication and Control. Prentice-Hall, Inc., 1972, pp. 81-82.

3. Tausworthe, R. C., "Theory and Practical Design of Phase-Locked Receivers," Jet Propulsion Laboratory Technical Report 32-819, 1966.

4. Viterbi, A. J. Principles of Coherent Communication. McGraw-Hill, 1966, pp. 76-119.

5. Lindsey, W. C., and Simon, M. K. Telecommunication Systems Engineering. Prentice-Hall, Inc., 1973, pp. 302-336.

6. Hershey, L., "Design Requirement - NASA Standard Deep Space Transponder," JPL Design Requirement No. STD-336-M01-DS, June 27, 1975.

APPENDIX E

COMMAND AUTHENTICATION

Gaylord K. Huth

APPENDIX E

COMMAND AUTHENTICATION

Gaylord K. Huth

For Air Force missions, modifications are proposed to the Shuttle command system[1] to provide command authentication. These modifications are shown in Figures 1, 2, and 3. In all figures, there is an authenticator variable that can be up to 50 bits long that is known at both transmitter and receiver. The general approach taken was to modify the parity portion of the command codeword using the authenticator variable.

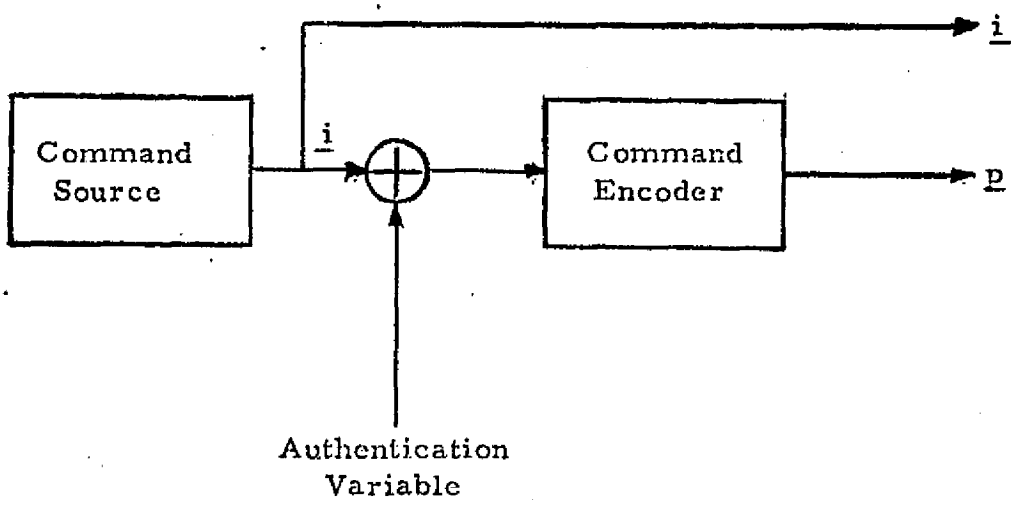
In Figure 1, the variable is modulo-2 added to the command information word $\underline{i}$ generated by the command source. This combination word is then BCH encoded to produce a parity word $\underline{p}$. The codeword $\underline{x}=(\underline{i},\underline{p})$ is then transmitted. Note that the code so produced by a fixed authenticator variable is not cyclic. In the receiver, the error detector encodes the sum of the variable and the received information word $\underline{i}'$ to produce a test parity $\underline{p}''$. This test parity is then compared to the received parity $\underline{p}'$. This is done by simply modulo-2 adding $\underline{p}'$ and $\underline{p}''$. If the sum is zero, the codeword is accepted; otherwise, it is rejected.

In Figure 2, the information word and the variable are encoded by completely different (127, 50) BCH encoders. The two parities produced are modulo-2 added to form $\underline{p}$. The codeword $\underline{x}=(\underline{i},\underline{p})$ is transmitted. The same process is performed on the received word in the error detector; the resultant parity $\underline{p}''$ is compared with the received parity $\underline{p}'$ as in the first method.

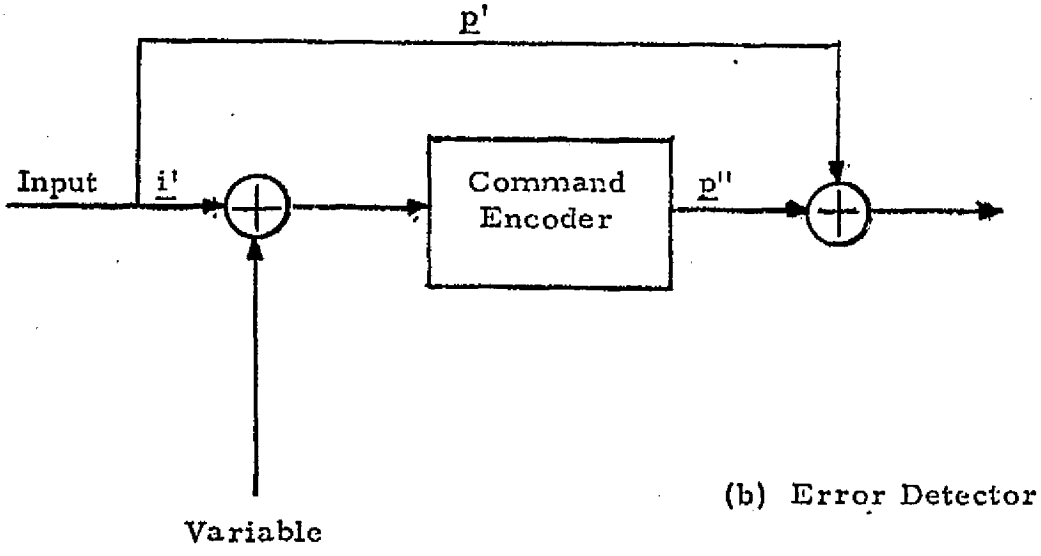
In Figure 3, the information word is encoded and the resulting parity word is permuted in some fashion depending on the variable. The permuted parity word is then transmitted together with the information word. Again,

---

[1] G. K. Huth and B. H. Batson, "A Command Encoding Scheme for A Multiplexed Space Communication Link," NTC'73 Record, Atlanta, Georgia, November 26-28, 1973.

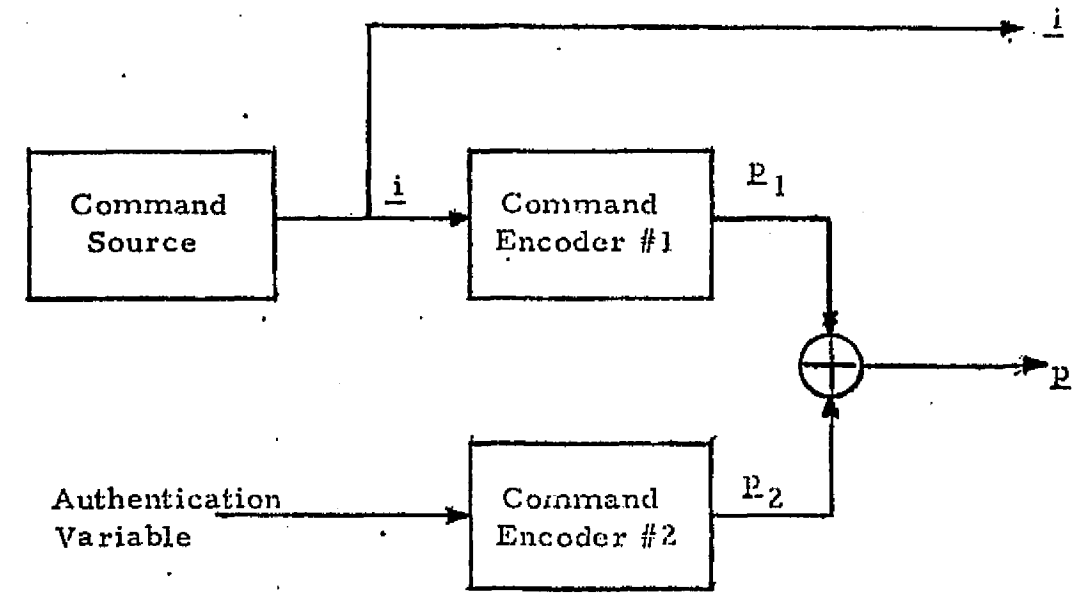


(a) Encoder

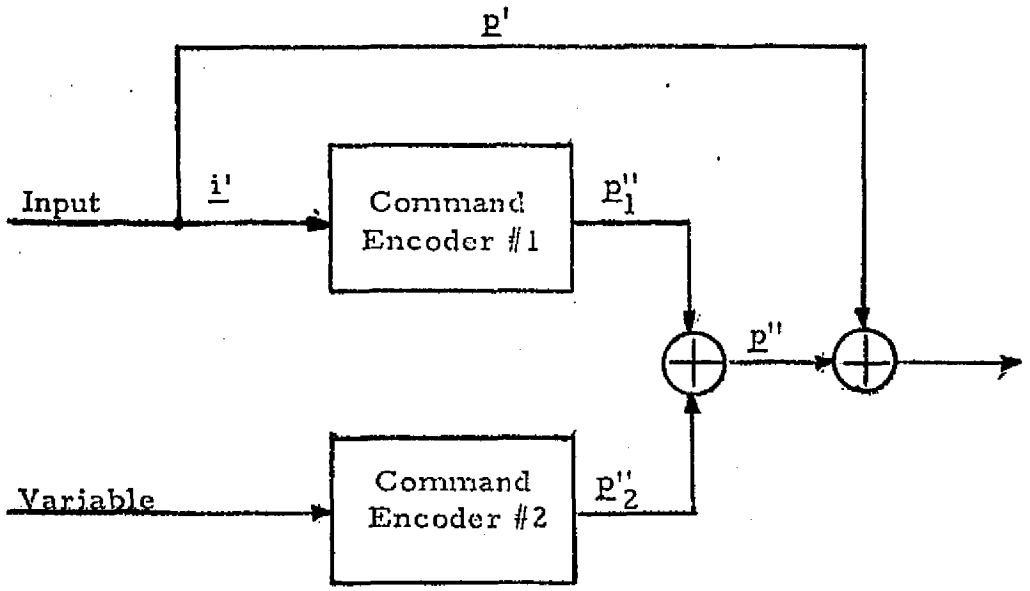


(b) Error Detector

Figure 1. Method No. 1 for Command Authentication

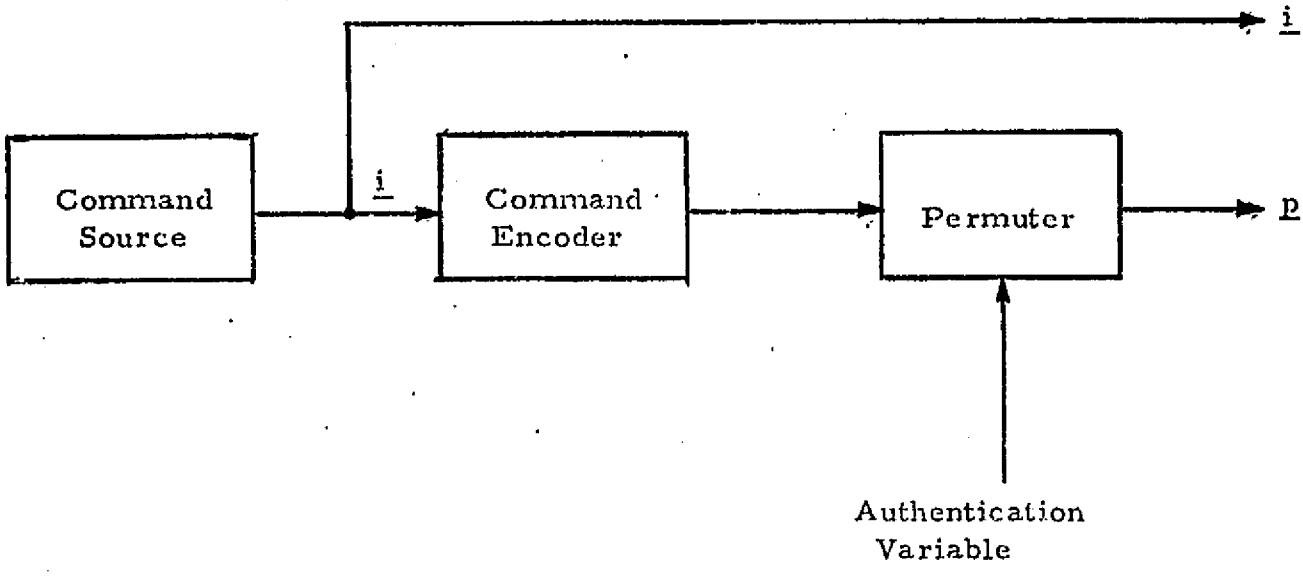


(a) .Encoder

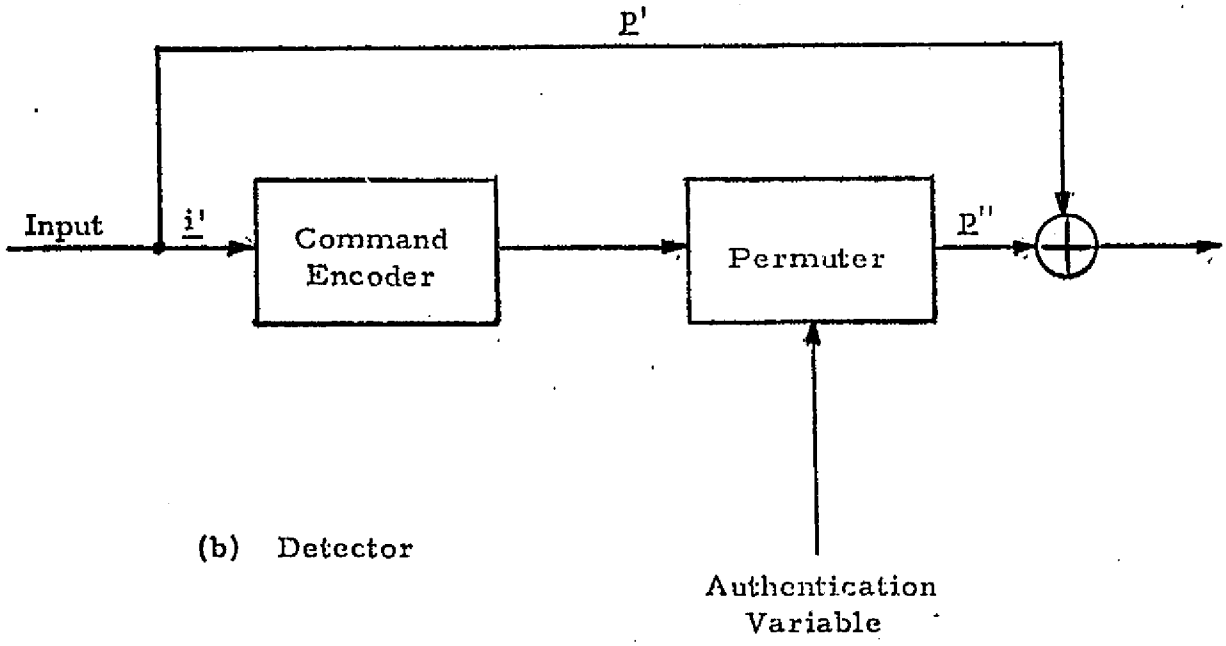


(b) Detector

Figure 2. Method No. 2 for Command Authentication



(a) Encoder



(b) Detector

Figure 3. Method No. 3 for Command Authentication

the detector performs the same encoding process on the received codeword and checks the resulting parity with the received parity.

The first method is easy to jam while the other two methods are harder to jam. The third method is probably the hardest to jam, but is the most complex. The second method is recommended since it is relatively hard to jam and is of considerably less complexity than method 3.

In analyzing the command authentication scheme in Figure 2 against jamming, it is felt that the multiplexing scheme used might provide a susceptibility to jamming. The multiplexing scheme is shown in Figure 4. The synchronization channel is sent as one 32-bit segment to each format frame. The voice channels labeled V1 and V2 in Figure 4 are 128-bit segments sent five times per format frame. For an encoded command word length of 128 bits, Figure 4 indicates that a complete command word is transmitted in four 32-bit segments during a format frame. The basic unencoded command word is 48 bits in length. Hence, with 50 format frames per second, all the data rates are achieved with the command data rate (48 bits/frame) (50 frames/second) equal to 2.4 kbps. It is possible, then, to add 80 bits of redundancy to each 48-bit command word to give the 128 bits allowed for the command word per format frame.

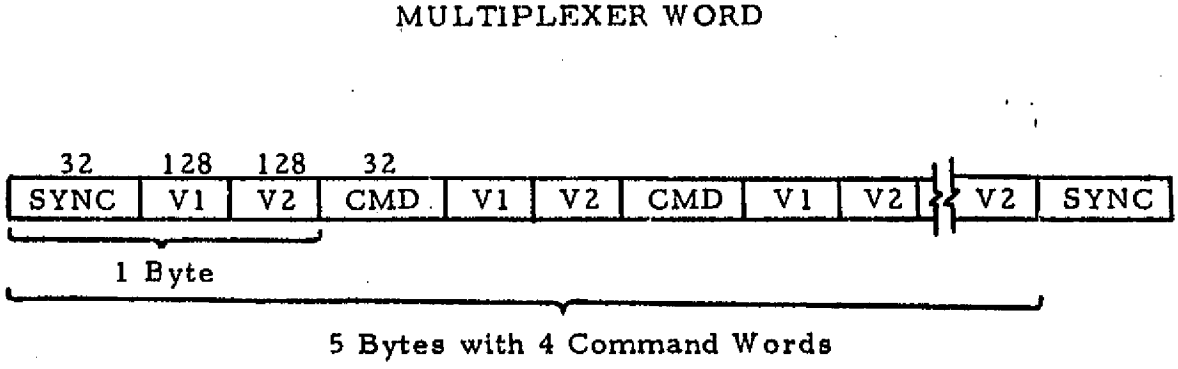


Figure 4. Multiplexer Word Format

A block diagram of the time-division multiplexed system is shown in Figure 5. The Viterbi decoder is switched into the system for the weak link through the TDRS relay satellite, in order to maintain acceptable bit error probabilities on the voice channel. The COMSEC unit may be switched into the system for Air Force missions. The strong direct link will exhibit such a low bit error probability that, even with the error extension of the COMSEC unit, the bit error probability will still be very small. As an example of the error extension phenomenon, consider a hypothetical COMSEC unit which outputs a burst of length 100 containing a 50 percent error rate for each input error. The possibility of jamming occurs by forcing an error before the command information in the multiplexed format so that only the first few information bits are changed due to the error extension. Similarly, the last few information bits in the first 32 bits of the multiplexed format could be changed without the error extension affecting any other bits. The possibility of jamming occurring is whether the parity bits corresponding to the changed information bits can be changed, causing an incorrect command to be accepted. This possibility would occur if the parity bits corresponding to the information bits at the beginning or end of a 32-bit segment are also at the beginning or end of a 32-bit segment so they can be changed in spite of the error extension. To eliminate this jamming possibility, the information bits could be placed at the center of each 32-bit sequence, that is, 12 bits in the center of each 32-bit sequence. In this way, no information could be changed without changing many parity bits due to the error extension and thus rejecting the command. It should be feasible to arrange the information bits and parity bits so that the possibility of jamming is extremely small.

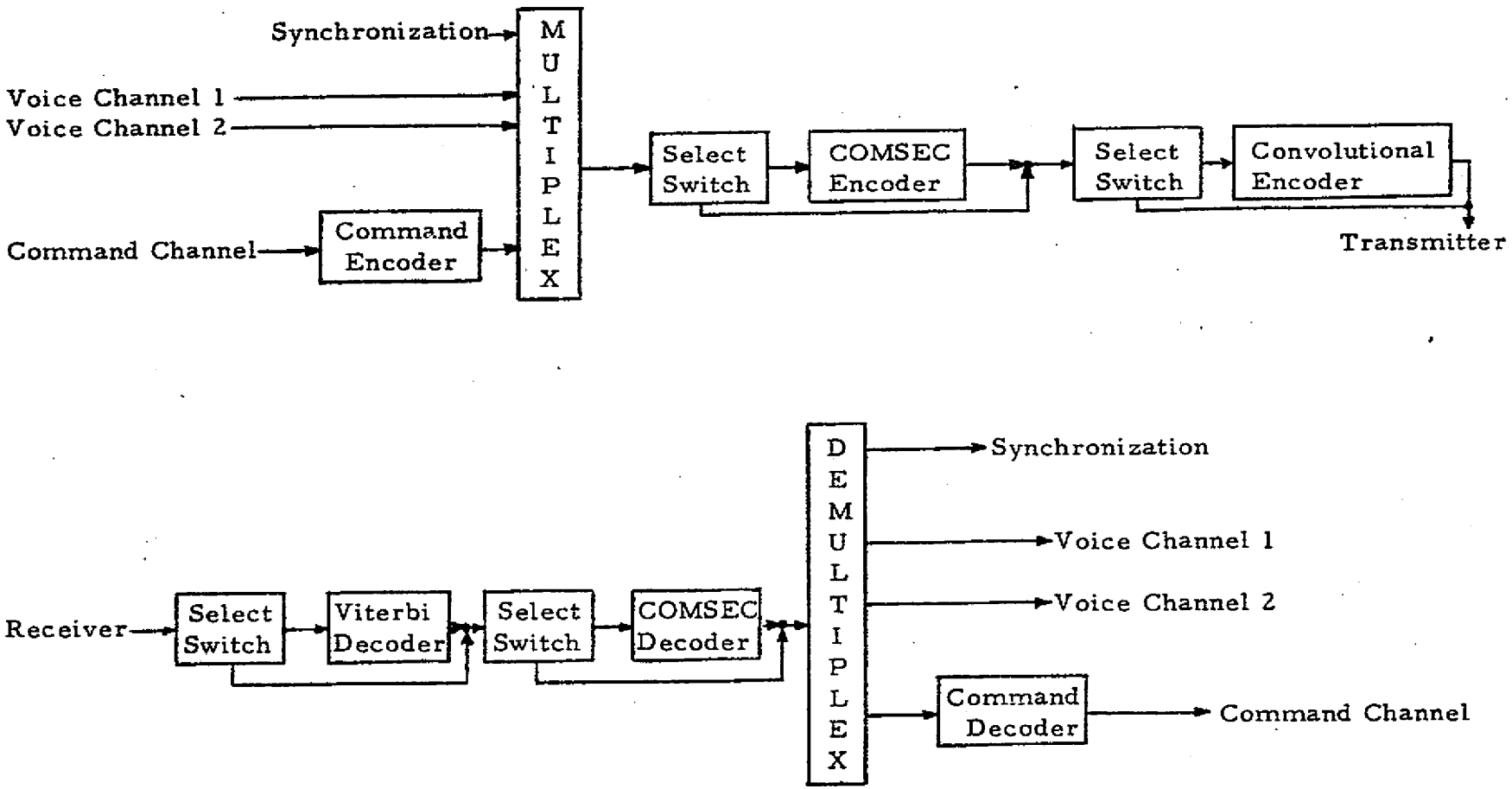


Figure 5. Multiplexed Synchronization, Voice, and Command Channel Block Diagram.

APPENDIX F

PREDICTING PERFORMANCE OF CONVOLUTIONAL CODED COMMUNICATION SYSTEMS

G. K. Huth

APPENDIX F

PREDICTING PERFORMANCE OF CONVOLUTIONAL CODED COMMUNICATION SYSTEMS

G. K. Huth

Summary. -The probability of error for a convolutional coded communication system can be predicted using three approaches. The first approach is using Gallager's [1] exponential bound for random codes. The second technique is the use of transfer function union bounds developed by Viterbi [2]. Finally, an approximation to the transfer function developed by Huth and Weber [3] can be used. This paper compares the three approaches for predicting performance and presents results for Viterbi decoding bit error probability, burst error probability, and decoder memory length.

Introduction. -To design a convolutional coded communication system, techniques are required to predict probability of bit error (P_b) versus energy per bit to single-sided noise spectral density (E_b/N_0) for the modulation/demodulation scheme employed. In addition, it would be desirable to predict the nature of the errors out of the decoder. That is, do the errors occur in bursts and, if so, what is the probability of an error burst of a given length.

There are three basic analytical techniques available for predicting P_b versus E_b/N_0. The first technique is using Gallager's exponential bound for random codes [1]. This bound allows quick comparisons of various modulation/demodulation and encoding/decoding schemes in terms of the basic coding parameters of code rate (R) and encoding constraint length (K). Therefore, it can be used as an initial optimization technique.

The second technique is to analyze the distance between codewords using a generating function (transfer function) developed by Viterbi [2]. This analysis technique has been used to obtain tight upper bounds on the error probability for the binary symmetric channel (BSC), the additive white Gaussian noise (AWGN) channel, and general memoryless channels. These bounds can be used to optimize the communication system parameters for a particular convolutional code and Viterbi decoding. While the bounds on error probability using the transfer function approach are very precise, the bounds are difficult to calculate for long constraint codes. Also, for even moderate constraint lengths, the transfer function bounds are evaluated in matrix form and the individual codeword distances are not available to determine the effect of decoder memory truncation or to determine the statistics of the errors at the output of the decoder.

The third technique uses an approximation of the transfer function developed by Huth and Weber [3], which is easy to calculate even for long constraint length codes. The resulting bounds are as tight as Viterbi transfer function bounds and the individual codeword distances appear in the bound so that it is possible to determine decoder memory truncation effects and output error statistics.

Random Coding Prediction Techniques. -Random coding bounds on error probability were developed by Gallager as easily calculable functions that did not require detailed knowledge of the code distance properties. Gallager's function assumed that the decoder matched its decoding rule to the channel statistics. However, when the actual channel statistics are unknown at a given time, it is desirable: (a) to choose the modulation format and the corresponding demodulation technique; and (b) to optimize the demodulation parameters and the decoder metrics to minimize the probability of error for all possible channel statistics. Considering the decoding metrics, let $d(\underline{y}, \underline{x})$ be a decoding metric function where $d(\underline{y}, \underline{x}) \geq 0$ for all received sequences $\underline{y}$ and all possible transmitted sequences $\underline{x}$. If $\underline{x}_m$ is the transmitted codeword for message m and $\underline{y}$ is the received sequence, then the decoder correctly chooses $\underline{x}_m$ if

$$d(\underline{y}, \underline{x}_m) > d(\underline{y}, \underline{x}_{m'}) \qquad \text{for all } m' \neq m . \tag{1}$$

Let $P_r[\text{error}|m, \underline{x}_m, \underline{y}]$ be the probability of error conditioned on the message m entering the encoder, on the selection of a particular sequence $\underline{x}_m$ as the m^{th} codeword, and on the received sequence $\underline{y}$. This probability can be upper bounded by using a random coding argument similar to that of Gallager by letting the codeword $\underline{x}_m$ be chosen randomly with probability $Q(\underline{x}_m)$. The resultant bound is

$$\bar{P}_{e,m} \leq (M-1) \sum_{\underline{y}} \left[\sum_{\underline{x}_m} Q(\underline{x}_m) \frac{P(\underline{y}|\underline{x}_m)}{\sqrt{d(\underline{y}, \underline{x}_m)}} \right] \left[\sum_{\underline{x}} Q(\underline{x}) \sqrt{d(\underline{y}, \underline{x})} \right] , \tag{2}$$

where M is the number of possible codewords and $P(\underline{y}|\underline{x})$ is the channel transition probability. For a memoryless channel, note that

$$d(\underline{y}, \underline{x}) = \prod_{n=1}^{N} d(y_n, x_n) \qquad \text{and} \qquad P(\underline{y}|\underline{x}) = \prod_{n=1}^{N} P(y_n|x_n) , \tag{3}$$

where y_n takes on the values $0, 1, \ldots, J-1$ and x_n takes on the values $0, 1, \ldots, I-1$. Let $Q(k)$ be an arbitrary probability assignment on the channel input alphabet and let each letter of each codeword of length N be chosen independently so that

$$Q(\underline{x}) = \prod_{n=1}^{N} Q(x_n) . \tag{4}$$

Upon substituting (3) and (4) into (2), we get

$$\bar{P}_{e,m} \leq 2^{-N(R_D - R)} \tag{5}$$

where

$$R_D = -\log_2 \sum_{j=0}^{J-1} \left[\sum_{k=0}^{I-1} Q(k) \frac{P(j|k)}{\sqrt{d(j,k)}} \right] \left[\sum_{i=0}^{I-1} Q(i) \sqrt{d(j,i)} \right] \tag{6}$$

and $2^{NR} = M$. Note that R is the code information rate and R_D is the largest rate such that reliable communication is possible. Furthermore, $R_D = R_{comp} = E_0(1)$ (Gallager's function) when $d(j,k) = P(j|k)$ which

corresponds to a maximum likelihood decoder but requires knowledge of the channel statistics.

Viterbi [2] has derived the bit error probability bound for Viterbi decoding using Gallager's function. For unknown channel statistics, Viterbi's bound can be rewritten using R_D as follows:

$$P_B < \frac{2^{-K R_D/R}}{[1 - 2^{-(R_D/R-1)}]^2}, \tag{7}$$

where K is the convolutional constraint length and R is the code rate.

Using (6) and (7), an initial optimization of the communication system can be performed since the modulation and channel characteristics determine P(j|k) and the decoder metrics such as soft decision, hard decision, or list quantization used in the Viterbi decoder determine d(j, k).

Transfer Function Union Bounds. -The transfer function of a convolutional code, introduced by Viterbi [2] in 1971, is essentially a Hamming weight enumerator of certain types of codewords. Specifically, it enumerates all the closed paths in the code's state diagram [3]. The general form of this function is

$$T(D, N) = \sum_{k=d_f}^{\infty} a_k D^k f_k(N) \tag{8}$$

where $f_k(N)$ is a polynomial of the form

$$f_k(N) = c_{k1} N + c_{k2} N^2 + \ldots c_{kj} N^j + \ldots \tag{9}$$

and d_f is the free distance of the code [2, 3]. The general term in (8) has the following interpretation. For the given code, there are $a_k c_{kj}$ codewords that diverge from the all-zeros path at some point and remerge with it later, and these codewords have Hamming weight k and are caused by input sequences with j "1" bits.

The transfer function has its usefulness in upper bounding the bit error probability of the code. To derive this bound, let P_k be the probability of incorrectly choosing a weight k codeword over the all-zeros codeword at the point of their remergence. If the weight k codeword was caused by j input "1" bits, then the bit error probability for this event is jP_k. If there are $a_k c_{kj}$ paths of this type, the probability of the union of these bit error events is upper bounded by $ja_k c_{kj} P_k$. Finally, the probability of the union of all possible bit error events, P_B, is upper bounded by the sum of the probabilities of these events,

$$P_B < \sum_{k=d_f}^{\infty} a_k P_k \left. \frac{\partial f_k(N)}{\partial N} \right|_{N=1}, \tag{10}$$

From (9), note that the derivative term in (10) merely causes the remaining terms to be multiplied by the exponents of the variable N, and these exponents are just the number of bit errors.

Suppose there existed a bound on each P_k of the form

$$P_k \leq P_{d_f} \exp[-\alpha(k - d_f)x]; \qquad k \geq d_f, \tag{11}$$

where x is the signal-to-noise ratio and α is a function of x and the modulation type. Substituting (11) into (10) and rearranging terms gives

$$P_B < P_{d_f} \exp(\alpha d_f x) \sum_{k=d_f}^{\infty} a_k \exp(-\alpha k x) \left. \frac{\partial f_k(N)}{\partial N} \right|_{N=1} . \qquad (12)$$

Comparing (12) and the derivative of (8) with respect to N yields

$$P_B < P_{d_f} \exp(\alpha d_f x) \left. \frac{\partial T(D,N)}{\partial N} \right|_{N=1,\, D=\exp(-\alpha x)} . \qquad (13)$$

Thus, for the bound in (13), T(D,N) or its derivative need only be evaluated at some fixed values of D and N, a possible task.

As an example of evaluating (8), consider the simple rate R=1/2, constraint length K=3 encoder in Figure 1. The contents of the last two register stages on the right constitute the encoder state, hence there are four states defined by $\underline{a}$=00, $\underline{b}$ = 01, $\underline{c}$ = 10, and $\underline{d}$ = 11. The state transition diagram for this code is shown in Figure 2 where the solid lines mark a transition caused by an input "0", the dotted lines mark a transition caused by an input "1", and the two binary symbols on each line represent the code output for that transition. Looking at this diagram, the first few terms in the transfer function series can be written down. The path $\underline{a}\,\underline{b}\,\underline{c}\,\underline{a}$ has weight 5 and was caused by one input "1". The two paths $\underline{a}\,\underline{b}\,\underline{c}\,\underline{b}\,\underline{c}\,\underline{a}$ and $\underline{a}\,\underline{b}\,\underline{d}\,\underline{c}\,\underline{a}$ have weight 6 and were caused by input sequences containing two input "1" bits, so the next term in the series is $2D^6N^2$.

This counting procedure could be continued, but it is easy to see that it is not too efficient, especially for more complex codes. A better way is to obtain directly a closed form expression for T(D,N). First, redraw Figure 2 so that state $\underline{a}$ is split into both source and sink nodes. Next, label each transition with the dummy variable D raised to the weight of the code output for that transition and N raised to the number of input "1" bits that caused the transition. This construction yields the diagram shown in Figure 3. Now T(D,N) is just the transfer function of this diagram regarded as a signal flow graph. Using standard signal flow graph reduction techniques results in

$$T(D,N) = \frac{D^5 N}{1 - 2DN} = D^5N + 2D^6N + \ldots + 2D^kN^{k-4} + \ldots \qquad (14)$$

which checks with the two terms gotten by counting.

Although the above method is straightforward, it is not suitable for computer computation. For the present example, define the state variables x_a, x_b, x_c, x_d corresponding to the states $\underline{a}$, $\underline{b}$, $\underline{c}$, and $\underline{d}$, respectively. Then, from Figure 3, the following node equations hold:

$$x_a = D^2 x_c; \quad x_b = ND^2 x_a + N x_c; \quad x_c = D x_b + D x_d; \quad x_d = NDx_b + NDx_d . \qquad (15)$$

Next, define the state vector $\underline{x} = (x_b, x_c, x_d)^T$. Then (15) can be rewritten in compact form as

$$\underline{x} = A\underline{x} + B \quad \text{and} \quad x_a = C\underline{x} , \qquad (16)$$

where

$$A = \begin{bmatrix} 0 & N & 0 \\ D & 0 & D \\ ND & 0 & ND \end{bmatrix} \tag{17}$$

is called the state transition matrix, and where

$$B = (ND^2 \ 0 \ 0)^T \quad \text{and} \quad C = (0 \ D^2 \ 0) . \tag{18}$$

Solving for $\underline{x}$ in (16) gives

$$\underline{x} = (I - A)^{-1} B = (I + A + A^2 + \ldots) B \tag{19}$$

where I is the identity matrix. Equation (19) lends itself very easily to computer evaluation. Obviously, the sum in (19) is truncated to a finite number of terms such that the truncation error is negligible compared to the sum. Now, once (19) and (16) have been evaluated, T(D, N) is equal to x_a.

These bounds have been shown to be extremely tight for the AWGN channel. In this case, the bit error probability is bounded by

$$P_B < \Phi\left(\sqrt{\frac{2d_f E_s}{N_0}}\right) \exp\left(\frac{d_f E_s}{N_0}\right) \left.\frac{\partial T(D, N)}{\partial N}\right|_{N=1, D=\exp(-E_s/N_0)} \tag{20}$$

where E_s/N_0 is the binary coded symbol energy per single-sided noise spectral density, and

$$\Phi(x) = \int_x^\infty \frac{\exp(-y^2/2)}{\sqrt{2\pi}} \, dy . \tag{21}$$

For bit error probability less than 10^{-3}, the bound is within 0.1 dB of the actual measured results for an AWGN channel.

Transfer Function Approximations. -The use of the state transition matrix to evaluate the transfer function union bounds loses much of the information contained in the individual terms of the transfer function series notation. An approximation developed by Huth and Weber [3] alleviates the problem and allows additional useful results to be obtained.

It has been shown [4] that the minimum weight D_ℓ of all codewords of length ℓ which have not returned to the all-zeros state (i.e., unmerged) is bounded by

$$\lceil d_0 \ell + b_l \rceil \le D_\ell \le \lfloor d_0 \ell + b_u \rfloor \quad \text{for all } \ell \ge 1 , \tag{22}$$

where d_0 is the minimum average weight per branch, b_l and b_u are biases, $\lceil x \rceil$ is the smallest integer greater than or equal to x, and $\lfloor x \rfloor$ is the largest integer less than or equal to x.

For some length ℓ, the path of smallest weight is lower bounded by $d_0\ell + b_l$. The path with the maximum weight is estimated by $n\ell$, where the code rate is R=1/n. All intermediate weight paths of length ℓ are approximated by appropriate combinations of these two extreme weights. The generating function for the weights of these unmerged codewords is

$$A_\ell(D,N) = ND^{d_0+b_1}(ND^{d_0}+D^n)^a(ND^n+D^{d_0})^b \tag{23}$$

where $a = \left\lceil \frac{\ell-1}{2} \right\rceil$ and $b = \left\lfloor \frac{\ell-1}{2} \right\rfloor$.

The probability of error due to truncation of the Viterbi decoder memory at length ℓ has the effect of causing the Viterbi decoder to choose among unmerged paths to determine the correct path. Thus, the bit error probability of error due to truncation is

$$\left.\frac{\partial A_\ell(D,N)}{\partial N}\right|_{\substack{N=1 \\ D^k=P_k}} = \sum_{i=0}^{a}\binom{a}{i}\sum_{j=0}^{b}\binom{b}{i}(1+i-j+b)P_k \tag{24}$$

where $k = b_1 + \lceil d_0(1+i+j)\rceil + n(\ell - 1 - i - j)$.

Using similar techniques, the probability of error due to the merged paths can be estimated:

$$B(D,N) = \sum_{k=d_f}^{d} a_k D^k f_k(N) + N^{\ell/2} D^{d_0+b_1} \sum_{k=\ell}^{\infty}(D^{d_0}+ND^n)^{k/2} \cdot (ND^{d_0}+D^n)^{k/2}, \tag{25}$$

where the free distance d_f is the minimum distance of the merged paths, and ℓ is defined as $\left\lceil \frac{d-b_1}{d_0} \right\rceil$.

Thus, the approximation to the merged transfer function consists of the first few terms of the transfer function in (8) and an approximation to the higher distance codewords similar to the unmerged codewords. Therefore, the bit error probability due to the merged paths is:

$$\left.\frac{\partial B(D,N)}{\partial N}\right|_{\substack{N=1 \\ D^k=P_k}} = \sum_{k=d_f}^{d} d_k P_k \left.\frac{\partial f_k(N)}{\partial N}\right|_{N=1} + \sum_{k=\ell}^{\infty}\sum_{i=0}^{k/2}\binom{k/2}{i}\sum_{j=0}^{k/2}\binom{k/2}{j}\left(\frac{\ell}{2}+i+\frac{k}{2}-j\right)P_r \tag{26}$$

where $r = b_1 + \lceil d_0(1+i+j)\rceil + n(k-i-j)$.

The total bit error probability is approximated by

$$P_B \approx \left[\frac{\partial A_\ell(D,N)}{\partial N} + \frac{\partial B(D,N)}{\partial N}\right]_{N=1,\, D^k=P_k}. \tag{27}$$

Using (27), the bit error probability can be computed for arbitrary Viterbi decoder path memory as shown in Figure 4. The convolutional code used is the rate 1/3, K=7 code found by Odenwalder [5] to have $d_f = 14$. The measured distance values are $d_0 = 11/17$ and $b_1 = 2$. Figure 4 presents a family of curves for various information bit energy per single-sided noise spectral density (E_b/N_0). Note that the results agree very closely with the results of simulation.

Figure 5 presents the bit error probability versus E_b/N_0 for two decoder memory lengths and compares the predicted results with measured results of a Viterbi decoder operating at the output of a STDN receiver [6]. Also shown are the results for the Viterbi decoder operating on digitally generated Gaussian noise. The Viterbi decoder quantizes the received signal to eight levels while the Huth-Weber approximation is for infinite quantization. Allowing for the 0.2-0.25 dB loss usually associated with 8-level quantization, the Huth-Weber approximation is very close to the measured results for the ideal Gaussian noise at large E_b/N_0. At small E_b/N_0, the union bound technique used to derive the transfer function bound results in a somewhat pessimistic prediction. However, for the region of $P_b < 10^{-3}$, the agreement is very good. Note that measured STDN receiver results are within 0.2 dB of the ideal Gaussian noise when the bit synchronizer degradation is accounted for. Thus, the Huth-Weber approximation is an excellent prediction technique for real channels.

The undetected errors of the Viterbi decoder tend to occur in bursts. This phenomenon has been observed from the operation of Viterbi decoders but also can be predicted from (25). Suppose an incorrect path is chosen by the Viterbi decoder that first differs from the correct path at the i^{th} bit and does not have K-1 identical consecutive bits until the i+L+K-1 bit. Then, the decoder will make burst error of length L. But the sum of undetected errors in (25) is just the sum of the various lengths of paths that remerge with the correct path after L+K-1 bits. Therefore, the probability of a burst error greater than L is the sum of those paths that remerge with the transmitted path in lengths greater than L+K-1. Let P(b≥L) be the probability of a burst error greater than or equal to L. Then, from (25), for $L \le \ell - K+1$,

$$P(b \ge L) < \sum_{i=L+K-1}^{\ell} \sum_{k=d_f}^{ni} g_i(k) P_k + \sum_{i=\ell}^{\infty} \sum_{j=0}^{i} \binom{i}{j} P_s \qquad (28)$$

where $g_i(k)$ is the number of paths of length i and weight k and $s = b_1 + \lceil d_0(1+j) \rceil + n(i-j)$. For $L \ge \ell - K+1$,

$$P(b \ge L) < \sum_{i=L+K-1}^{\infty} \sum_{j=0}^{i} \binom{i}{j} P_s \, . \qquad (29)$$

It is possible to simplify the approximation in (29) for large L by bounding P_s as in (20) for the AWGN channel. Therefore, for large L, P(b≥L) is of the form

$$P(b \ge L) < F \sum_{i=L+K-1}^{\infty} G^i$$

$$< F G^{L+K-1}/1-G \, , \qquad (30)$$

where $F = \Phi\left(\sqrt{2(d_0+b_1) E_s/N_0}\right)$ and $G = \exp(-d_0 E_s/N_0) + \exp(-n E_s/N_0)$.

Conclusions. -Three analytical techniques have been developed to predict the probability of error for a convolutional coded communication

system. Quick comparisons of modulation and coding parameters can be performed with the Gallager random coding exponential bounds. Finer parameter optimization can be obtained by using Viterbi transfer function union bounds and the Huth-Weber approximations to these bounds. The Huth-Weber approximations provide a technique to predict decoder memory truncation degradation and burst error probability, as well as bit error probability. These analytical approximations are important to optimize the modulation/coding parameters at very low probability of error where measurements are impractical. It was noted that the approximations are extremely accurate over the region $10^{-3} \geq P_b \geq 10^{-5}$ where measurements have been made. Since analytical techniques are used to predict performance, communication system optimization can be obtained with little effort.

References.

1. R. G. Gallager, "Information Theory and Reliable Communication," John Wiley and Sons, Inc., New York, N.Y., 1968.

2. A. J. Viterbi, "Convolutional codes and their performance in communication systems," IEEE Trans. on Comm. Tech., Vol. COM-19, No. 5, pp. 751-772; October, 1971.

3. G. K. Huth and C. L. Weber, "Notes on Convolutional Codes" (to be published).

4. G. K. Huth and C. L. Weber, "Minimum weight convolutional codewords of finite length," IEEE Trans. on Info. Theory, Vol. IT-22, No. 2, pp. 243-246; March, 1976.

5. J. P. Odenwalder, "Optimal decoding of convolutional codes," Ph.D. Dissertation, Dept. Eng., University of California, Los Angeles, 1970.

6. Test Report, "System design evaluation test - PSK coded link," No. EE7-75-211, Electronic Systems Test Laboratory, NASA Lyndon B. Johnson Space Center, Houston, Texas.

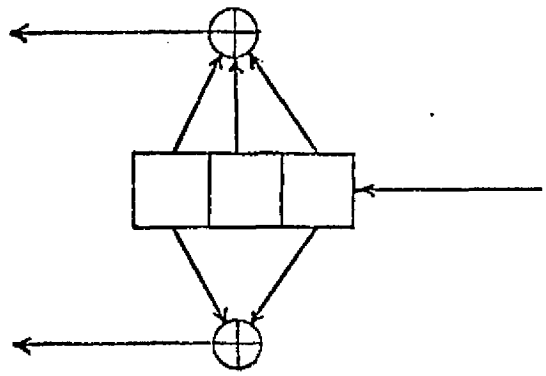


Figure 1. R = 1/2, K = 3 Encoder

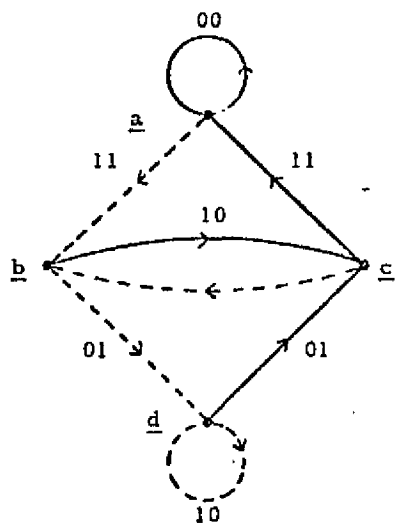


Figure 2. State Transition Diagram for R = 1/2, K = 3 Encoder

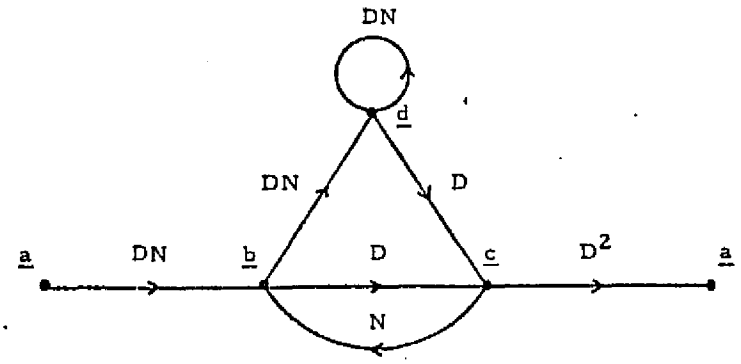


Figure 3. Signal Flow Diagram for R = 1/2, K = 3 Encoder

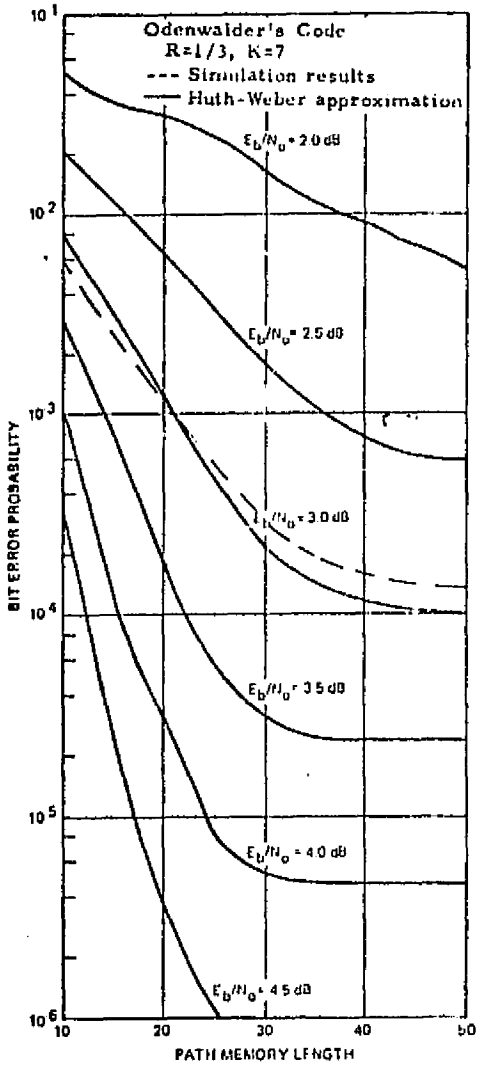


Figure 4. Bit Error Probability versus Decoder Path Memory Length

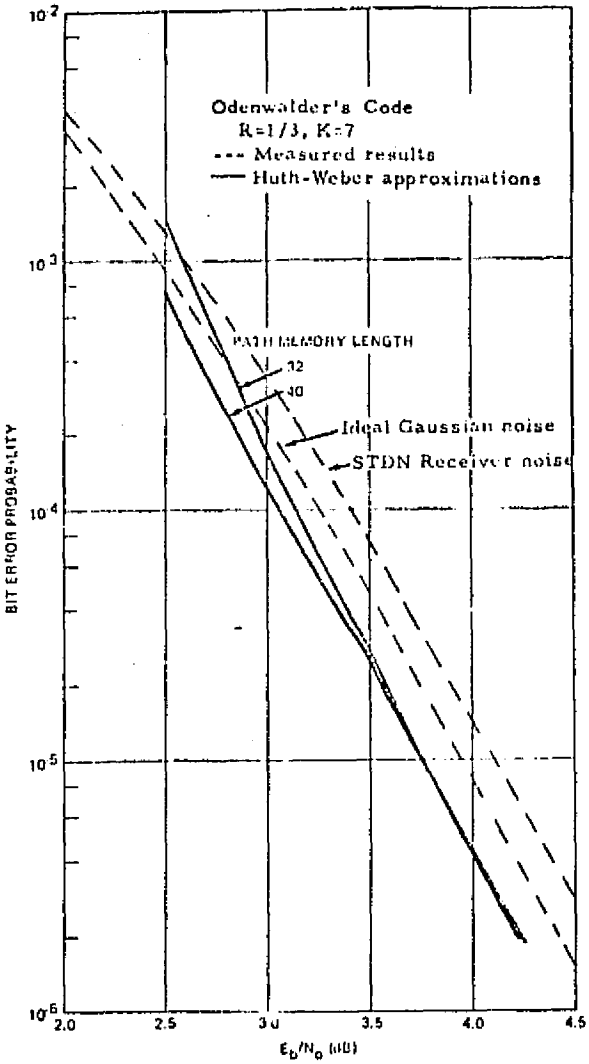


Figure 5. Bit Error Probability versus E_b/N_0

APPENDIX G

CONVOLUTIONAL CODING AT 50 MBPS FOR THE SHUTTLE KU-BAND RETURN LINK

Gaylord K. Huth

APPENDIX G

CONVOLUTIONAL CODING AT 50 MBPS FOR THE SHUTTLE KU-BAND RETURN LINK

Gaylord K. Huth

SUMMARY. - Error correcting coding is required for 50 Mbps data link from the Shuttle Orbiter through the Tracking and Data Relay Satellite System (TDRSS) to the ground because of severe power limitations. Convolutional coding has been chosen because the decoding algorithms (sequential and Viterbi) provide significant coding gains at the required bit error probability of 10^{-6} and can be implemented at 50 Mbps with moderate hardware. While a 50 Mbps sequential decoder has been built, the highest data rate achieved for a Viterbi decoder is 10 Mbps. Thus, five multiplexed 10 Mbps Viterbi decoders must be used to provide a 50 Mbps data rate. This paper discusses the tradeoffs which were considered when selecting the multiplexed Viterbi decoder approach for this application.

INTRODUCTION. - A wideband communications link from the Space Shuttle Orbiter to ground via the NASA TDRSS is currently planned. This link will be used for On-Orbit transmission of up to 50 Mbps of scientific data and, on a time-shared basis, will also accommodate television, analog scientific data, experiment or maintenance tape recorder dumps, etc. The operating frequency will be in the Ku-band (approximately 15 GHz).

In order to accommodate digital data rates on the order of 50 Mbps, under the constraints which exist due to limitations on Orbiter antenna diameter (present planning is for a 20-inch diameter deployable parabolic dish) and on Orbiter power amplifier output (approximately 50 watts), the use of coded, coherent PSK has been selected. The large performance gains associated with convolutional coding, the simple spaceborne encoder design, and the relative simplicity and availability of ground station decoders has led to the choice of convolutional codes over block codes. Due to a restriction in available TDRSS repeater bandwidth to 225 MHz, a code rate of 1/2 has been selected.

This paper addresses the problem of selecting a convolutional encoding and decoding approach that provides the best possible performance without severe penalties in cost, complexity, reliability, etc.

COMPARISON OF SEQUENTIAL AND VITERBI DECODING PERFORMANCE. - Convolutional coding with either sequential or Viterbi decoding has been shown to provide significant performance gains for communications over the memoryless white Gaussian noise channel. Both sequential and Viterbi decoders have been constructed for multi-megabit data rates and are currently available commercially. Soft-decision Viterbi decoders for short constraint length ($K = 7$) convolutional codes are available at information rates up to 10 Mbps, and provide coding gains in the range 4-6 dB. Figure 1 shows the measured performance of the Linkabit LV7015 Viterbi decoder, which operates at a maximum information rate of 100 Kbps. This measured performance coincides with the theoretical performance of the LV7017 Viterbi decoder, which operates at up to 10 Mbps. The measured performance data was obtained in

the Electronic Systems Test Laboratory at the NASA Lyndon B. Johnson Space Center, and reflects the performance achievable for a practical PSK receiver (non-ideal carrier reference) with the degradation due to symbol synchronizer performance corrected out.

Figure 1 also shows the theoretical performance of the Linkabit LS4157 sequential decoder, which is a commercially available hard-decision device originally developed for a 40 Mbps application, but marginally capable of operation at up to 50 Mbps.

As indicated by the performance curve, the coding gain provided by sequential decoding increases with decreasing information rate. This is because the performance of the sequential decoder is primarily determined by the buffer memory size and the speed advantage of the decoding logic (the ratio of computation rate to data rate), and the speed advantage is higher at lower information rates.

Figure 1 indicates that the soft-decision Viterbi decoder significantly outperforms the hard-decision sequential decoder at high error probabilities (greater than about 10^{-5}) and is still somewhat superior at an error probability of 10^{-6}. The almost 2 dB improvement in sequential decoding performance offered by soft-decisions would make sequential decoding a much more favorable candidate for the application being considered, but use of soft-decisions would greatly reduce the maximum information rate at which the decoder could operate. Soft-decision sequential decoders are not commercially available, but the technique is straightforward and has been demonstrated at low information rates.

POSSIBLE 50 MBPS DESIGN APPROACHES. - Based on the performance considerations described above and on the availability of decoder hardware, the following options for the 50 Mbps Shuttle Ku-band link were evaluated:

a. High-speed (> 10 Mbps) Viterbi decoder(s)

b. Multiplexed 10 Mbps Viterbi decoders

c. 50 Mbps hard-decision sequential decoder

d. Multiplexed low-speed (< 50 Mbps) soft-decision sequential decoders

The results of these evaluations are summarized in subsequent sections of this paper.

HIGH-SPEED VITERBI DECODER(S) FOR 50 MBPS OPERATION. - Currently, a Viterbi decoder capable of 50 Mbps operation can not be implemented with off-the-shelf components. The arithmetic unit of the Viterbi decoder is limited to about 35 Mbps using ECL off-the-shelf components and to about 20 Mbps using SOS components. Thus, with off-the-shelf components, at least two Viterbi decoders would have to be operated in parallel on multiplexed data if a 50 Mbps capability is to be provided. If fewer than five parallel Viterbi decoders are used, then a potentially expensive decoder development effort would be required to provide a decoder for operation at information rates greater than 10 Mbps.

The arithmetic unit could achieve 50 Mbps operation if it is implemented with a custom LSI chip using ECL or SOS. For a K = 7 decoder, 64 identical custom arithmetic chips would be required to implement the arithmetic unit. These custom LSI chips would significantly decrease the complexity of the arithmetic unit as well as allowing 50 Mbps operation.

However, this custom LSI chip development would be relatively expensive; also finding a semiconductor manufacturer to develop the custom LSI chip could be extremely difficult without a guaranteed large scale production follow-on.

50 MBPS OPERATION USING MULTIPLEXED 10 MBPS VITERBI DECODERS. - One obvious approach which would provide a 50 Mbps information rate capability without additional decoder development is to utilize five parallel 10 Mbps encoders and five parallel 10 Mbps Viterbi decoders. This implementation is illustrated functionally in Figure 2. The performance provided by a parallel approach such as this would be the same as would be provided by a single 10 Mbps Viterbi decoder.

It should be noted that the encoder portion of Figure 2 can be implemented with a single 31-stage shift register encoder operating at 50 Mbps with no mux/demux units.

HARD-DECISION SEQUENTIAL DECODING AT 50 MBPS. - Sequential decoding of convolutional codes is a powerful decoding technique. This technique has an advantage over Viterbi decoding in that long constraint length convolutional codes can be used since the code constraint length has only an insignificant effect on the decoder hardware complexity. Therefore, undetected errors can be made highly improbable for even small values of E_b/N_0. As noted previously, the performance of sequential decoding is primarily determined by two parameters: the buffer memory size and the speed advantage of the decoding logic (the ratio of computation rate to data rate).

In the Linkabit LS4157 hard-decision 50 Mbps sequential decoder, the received information bits are delayed by a 65,000 bit MOS shift register. The delayed information bits are exclusive-ORed with the output of the sequential decoder, resulting in the corrected information sequence. This large buffer is necessary because the number of computations to decode an information bit is a random variable.

The decoding logic is implemented using the fastest commercially available logic family, MECL III. The operating speed of the decoder logic is 70 megacomputations per second. Nearly all output errors are the result of the sequential decoder not having a correction to the information bit at the end of the buffer soon enough. When this event occurs, the information bit is output with no error correction. For the sequential decoder to "catch up," at least the next 500 bits will be output from the buffer with the channel errors uncorrected. Occasionally, 5000 bits will be output from the buffer without the channel errors corrected. Thus, the errors output from the buffer will tend to occur in long bursts.

50 MBPS OPERATION USING MULTIPLEXED SOFT-DECISION SEQUENTIAL DECODERS.-

From strictly a performance standpoint, this approach is superior to the other three approaches described previously. There are several disadvantages associated with soft-decision sequential decoding, however, and these will be briefly outlined.

The basic arithmetic in a soft-decision sequential decoder is much more complicated than in a hard-decision sequential decoder, because of the increased branch metrics. In addition, a larger buffer memory is required to store the same number of branches as a hard-decision device. The net result is that a 3-bit (8 quantization levels) soft-decision sequential decoder would be three times as complex as a hard-decision sequential

decoder, and would operate at a much lower information rate. A good estimate is that the information rate for a 3-bit soft-decision sequential decoder (using commercially available logic) would be limited to about 8 Mbps, such that six or seven of these decoders could be paralleled to provide a 50 Mbps capability. The performance provided by such an approach would be about 2 dB better than that provided by a single 50 Mbps hard-decision sequential decoder.

A 2-bit (4 quantization levels) soft-decision sequential decoder would only be twice as complex as a hard-decision device and could be made to operate at a somewhat higher information rate (perhaps 12.5 Mbps), such that four decoders could be paralleled to provide a 50 Mbps capability. However, a 2-bit soft-decision sequential decoder provides only a 1.25 dB improvement over hard-decisions, so the performance advantage would not be as significant.

Soft-decision sequential decoding suffers from two additional disadvantages. First, the performance of a soft-decision sequential decoder is highly susceptible to variations in receiver AGC voltage, so somewhat greater care must be exercised in the design of the modem with which the decoder must operate. Secondly, the memory introduced by the costas loop of a PSK receiver results in a pronounced degradation in the performance of a soft-decision sequential decoder; therefore, the use of interleaving (2000 bits or more) may be required in order to realize the full potential performance gain.

DESCRIPTION OF SELECTED APPROACH. - Based on the considerations noted above, the following conclusions were made regarding the four options for 50 Mbps operation which were evaluated:

a. High-Speed (> 10 Mbps) Viterbi Decoder(s) - This approach was considered unacceptable because of (1) the developmental effort required to provide high speed (< 50 Mbps) Viterbi decoders using off-the-shelf components or to provide the custom LSI chip required for a 50 Mbps Viterbi decoder, and because of (2) the fact that the performance would be no better than that achievable using commercially available Viterbi decoders operating at 10 Mbps.

b. Multiplexed 10 Mbps Viterbi Decoders - This approach was considered acceptable and was selected for implementation.

c. 50 Mbps Hard-Decision Sequential Decoder - This approach was considered unacceptable because of the poor performance characteristics of hard-decision sequential decoders.

d. Multiplexed Low-Speed (< 50 Mbps) Soft-Decision Sequential Decoders - This approach was considered unacceptable because of (1) the development costs associated with providing soft-decision sequential decoder hardware, and (2) the relatively high susceptibility of soft-decision sequential decoders to non-ideal conditions such as receiver AGC variations and receiver memory.

The 50 Mbps design approach selected for the Shuttle Ku-band link is summarized in somewhat greater detail in Figures 3 and 4. Figure 3 illustrates the single shift-register implementation for the encoder and also indicates the bit-by-bit interleaving scheme which was selected.

Using the encoder interleaving scheme shown in Figure 3, the parallel decoder configuration is totally self-synchronizing. The demultiplexer

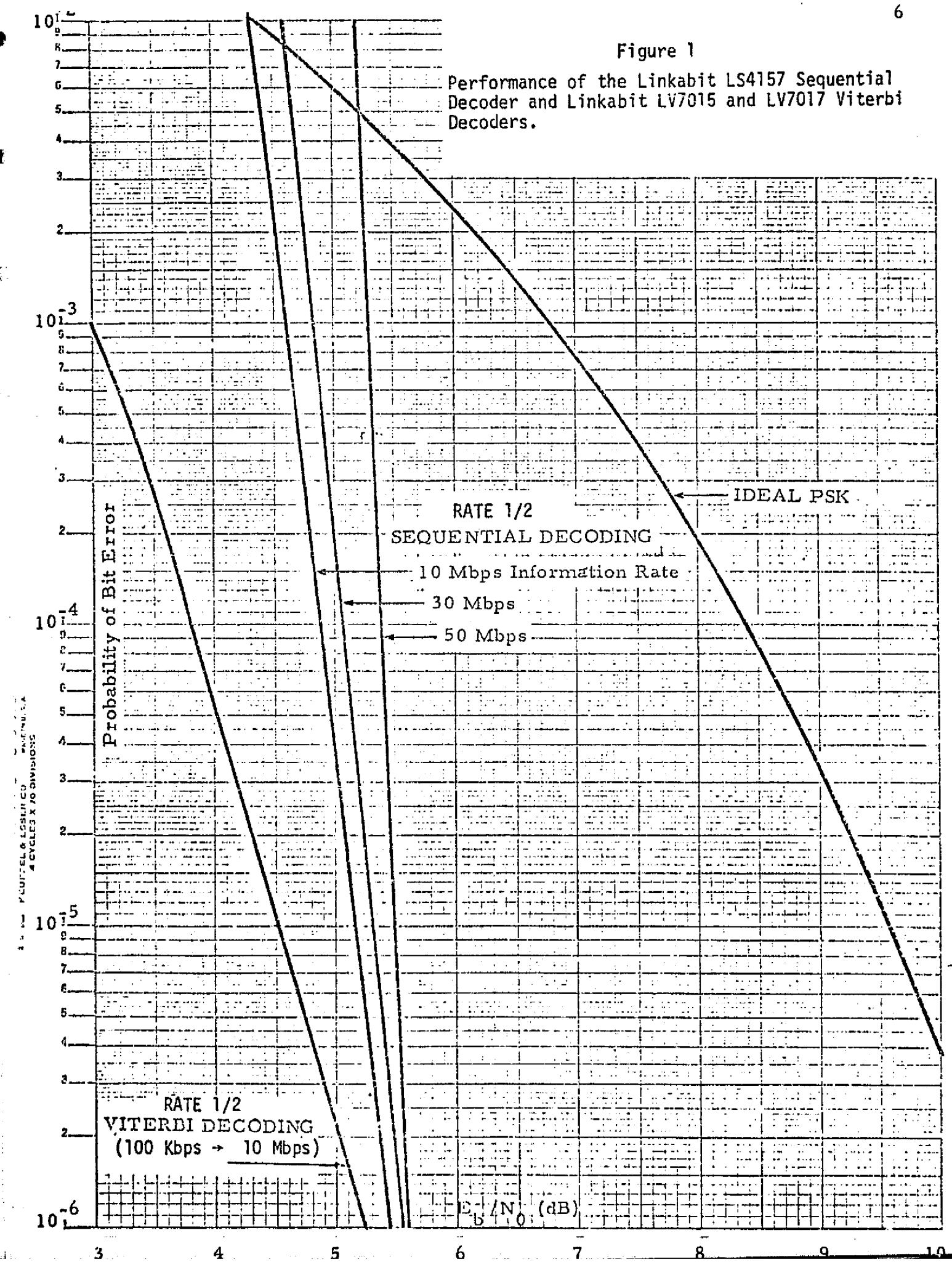
which precedes the parallel decoders is a single decommutator which merely routes every fifth incoming symbol to each decoder. Each decoder establishes its own branch synchronization and the resulting decoded sequences are combined with an output multiplexer. Figure 4 illustrates the details of these operations.

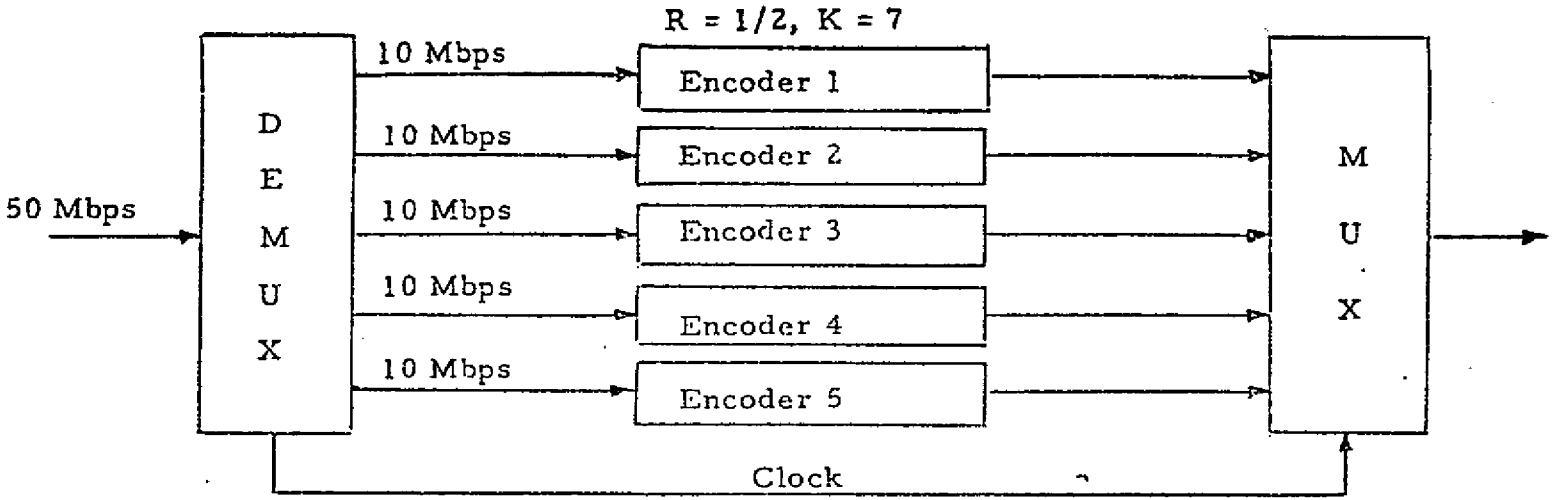
One aspect of the selected parallel Viterbi decoder approach which is of concern is that of reliability. The reliability of a configuration consisting of five decoders operating in parallel is obviously much lower than the reliability of a single component decoder. This decreased reliability may necessitate some sort of fault detection and sparing provisions, possibly even an automatic replacement scheme.

CONCLUSIONS. - It is felt that the selected 50 Mbps encoding/decoding approach offers a reasonable compromise between performance and system complexity. The coding gain provided by this approach is relatively high (5.3 dB at a probability of error of 10^{-6}), the system degrades gracefully (a coding gain of 2.7 dB is available at a probability of error of 10^{-2}), the spaceborne encoder is extremely simple to implement, and the ground-based decoders are available commercially and can be combined in a straightforward manner. Such an approach is not limited to 50 Mbps data rates. It appears reasonable to extend the parallel 10 Mbps Viterbi decoder concept to systems requiring operation at data rates well in excess of 100 Mbps, although reliability becomes of increasing concern as the number of parallel decoders becomes large.

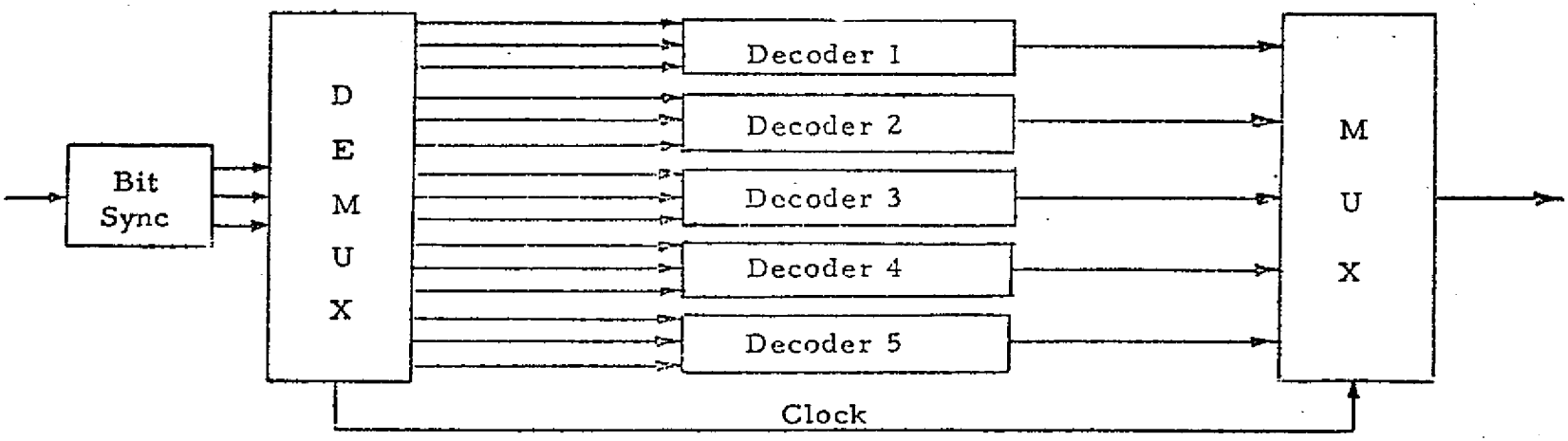
Figure 1

Performance of the Linkabit LS4157 Sequential Decoder and Linkabit LV7015 and LV7017 Viterbi Decoders.





ORBITER 50 MBPS ENCODER



GROUND 50 MBPS DECODER

Figure 2. Functional Configuration for 50 Mbps Operation Using Multiplexed 10 Mbps Viterbi Decoders

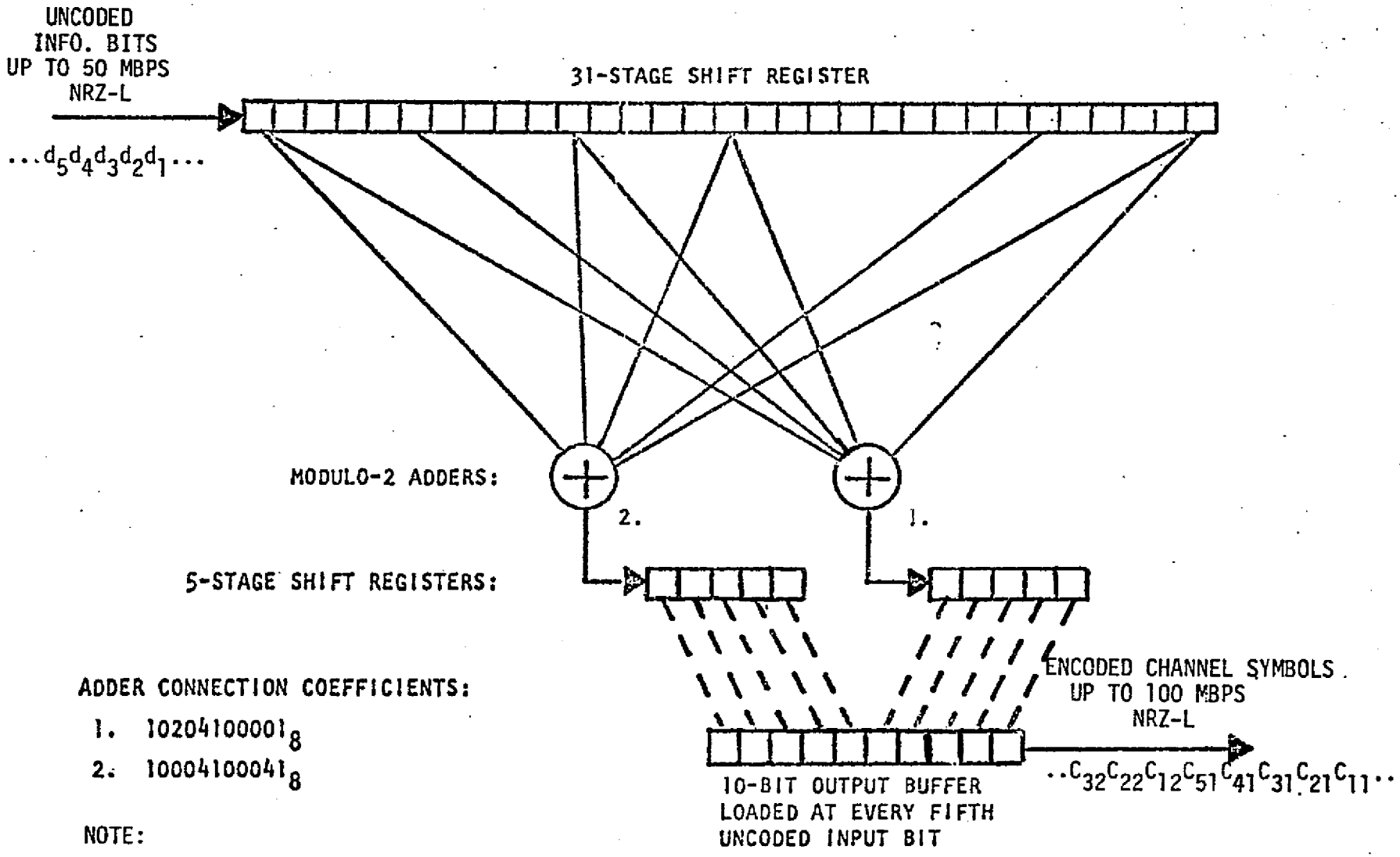
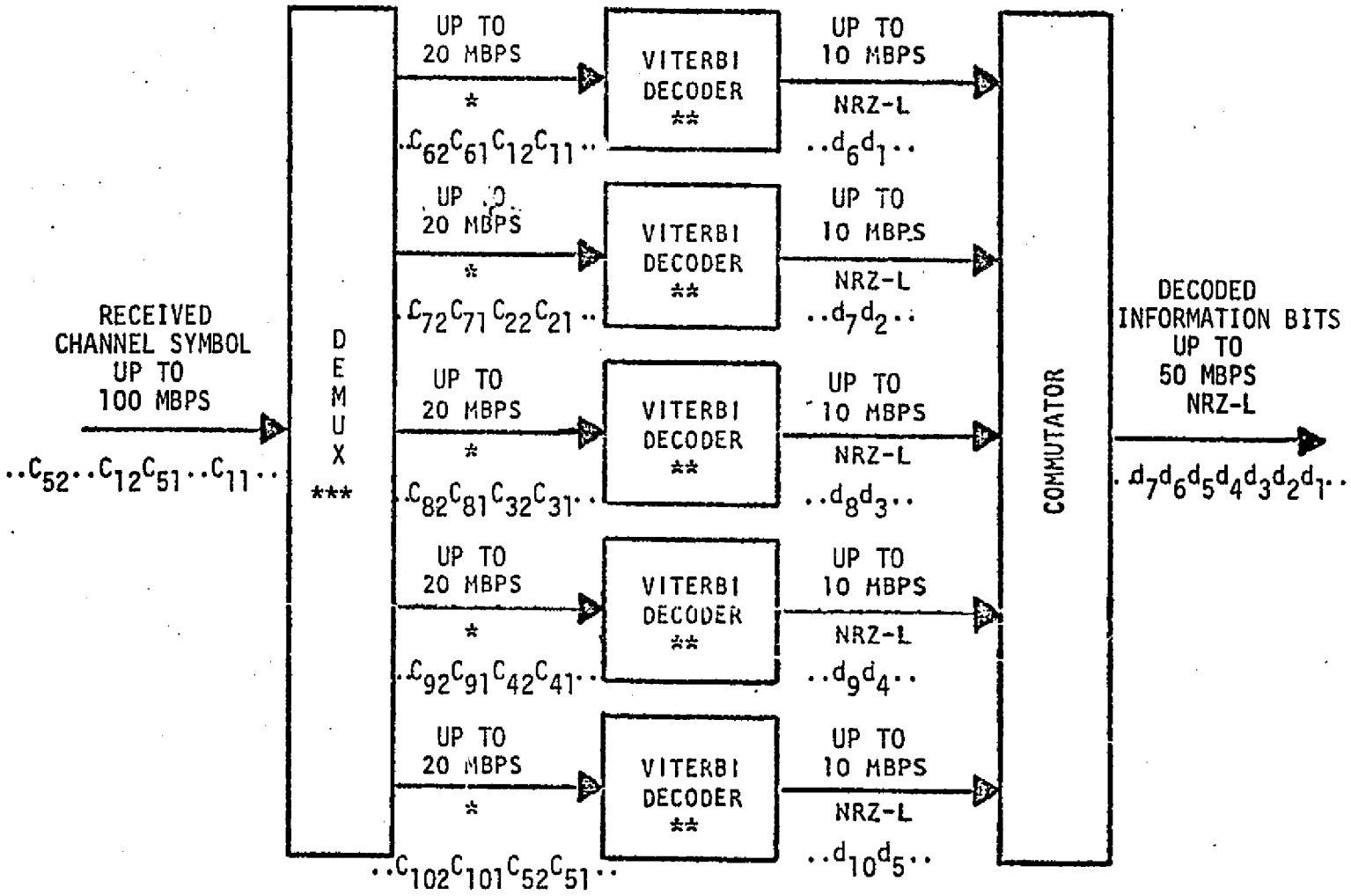


Figure 3. Alternate Encoder Implementation and Interleaving Scheme for 50 Mbps Encoding



* BIT SIGNAL FORMAT: 8-LEVEL SOFT-DECISION

** VITERBI DECODER CONFIGURATION: CODE RATE 1/2
GENERATOR SEQUENCE 35707_8

*** EVERY 5TH CHANNEL SYMBOL ROUTED TO THE SAME DECODER.

Figure 4

Synchronization Strategy for Parallel
Viterbi Decoders

APPENDIX H

SIGNAL DETECTION AND ANGULAR SEARCH PROCEDURE FOR SHUTTLE KU-BAND COMMUNICATION STUDY

Sergei Udalov

APPENDIX H

SIGNAL DETECTION AND ANGULAR SEARCH PROCEDURE FOR SHUTTLE KU-BAND COMMUNICATION STUDY

Sergei Udalov

1.0 INTRODUCTION

To establish the Ku-band communication link between the Shuttle and the TDRS, the antennas of these two satellites must be pointed at each other. Various techniques for pointing the TDRS antenna at the Shuttle have been considered and the latest information available to Axiomatix indicates that initial pointing can be accomplished without requiring the Ku-band Shuttle antenna axis to be in perfect alignment with the axis of the TDRS antenna. During this initial pointing, however, the effective EIRP available from the TDRS to Shuttle is only 40 dBw instead of the nominal 50 dBw. But, before the TDRS makes its final pointing correction, the Shuttle antenna must search out the random designation errors which exist between its computer-estimated line-of-sight (LOS) to TDRS and the actual pointing direction of the Shuttle antenna. In this appendix, we consider the procedures for searching out the residual designation errors of the Shuttle antenna and we estimate the time requirements for these procedures.

The main argument presented here is that the times required to detect the presence of the TDRS signal within an arbitrary beamwidth (1 dB beamwidth is assumed) of the Shuttle antenna are short compared to the typical time requirements for mechanically scanning the LOS uncertainty volume.

Before proceeding with the quantitative analysis, we would like to point out that the overall probability of detecting and acquiring the TDRS signal consists of two terms:

$P_D \equiv$ Probability of detecting the presence of the TDRS signal when the Shuttle antenna is pointed, within a specified offset, at the TDRS, and

$P_V \equiv$ Probability that the true LOS (to TDRS) is contained within the angular volume searched out by the Shuttle antenna.

The P_D is determined by forward link power budget and by the detection procedure used at the Shuttle receiver for determining the presence of the signal when the Shuttle antenna is within a predetermined offset from true

LOS to TDRS. The P_V is determined by such factors as Shuttle antenna beamwidth and beam shape (directivity characteristic), the LOS designation uncertainty, and the angular search pattern used. To a first-order approximation the overall detection probability is

$$P_T \cong P_D \times P_V \tag{1}$$

and with proper selection of detection and search techniques, one can insure that $P_T \geq P_T(\min)$ where $P_T(\min)$ is an arbitrarily selected minimum probability of signal acquisition.

Acquisition in this appendix is defined as termination of the angular search procedure and the initiation of the angular autotrack.

2.0 SIGNAL DETECTION

Before considering the angular scan procedure for the Shuttle antenna, we will analyze the procedure for detecting the presence of the signal within the antenna beamwidth. For this, we will assume that the angular scan over the designation uncertainty volume has been initiated and that the purpose of the signal detection unit is to stop the scan when the received signal exceeds a preset value.

The simplest and the fastest technique for determining the presence of the signal is to use an energy sensor, such as a square law envelope detector followed by an integrating filter and a threshold crossing indicator. Because this technique does not require either the code sync or the carrier lock, it can be applied directly to the incoming spread spectrum signal. The simplified block diagram for implementing such a detection scheme is shown in Figure 1.

The first step in determining the performance of this detector is to determine the power budget for the forward (TDRS to Shuttle) link. Table 1 shows the pertinent parameters of this budget. From this table, we see that the available P_{rec}/N_0 is 64.8 dB-Hz.

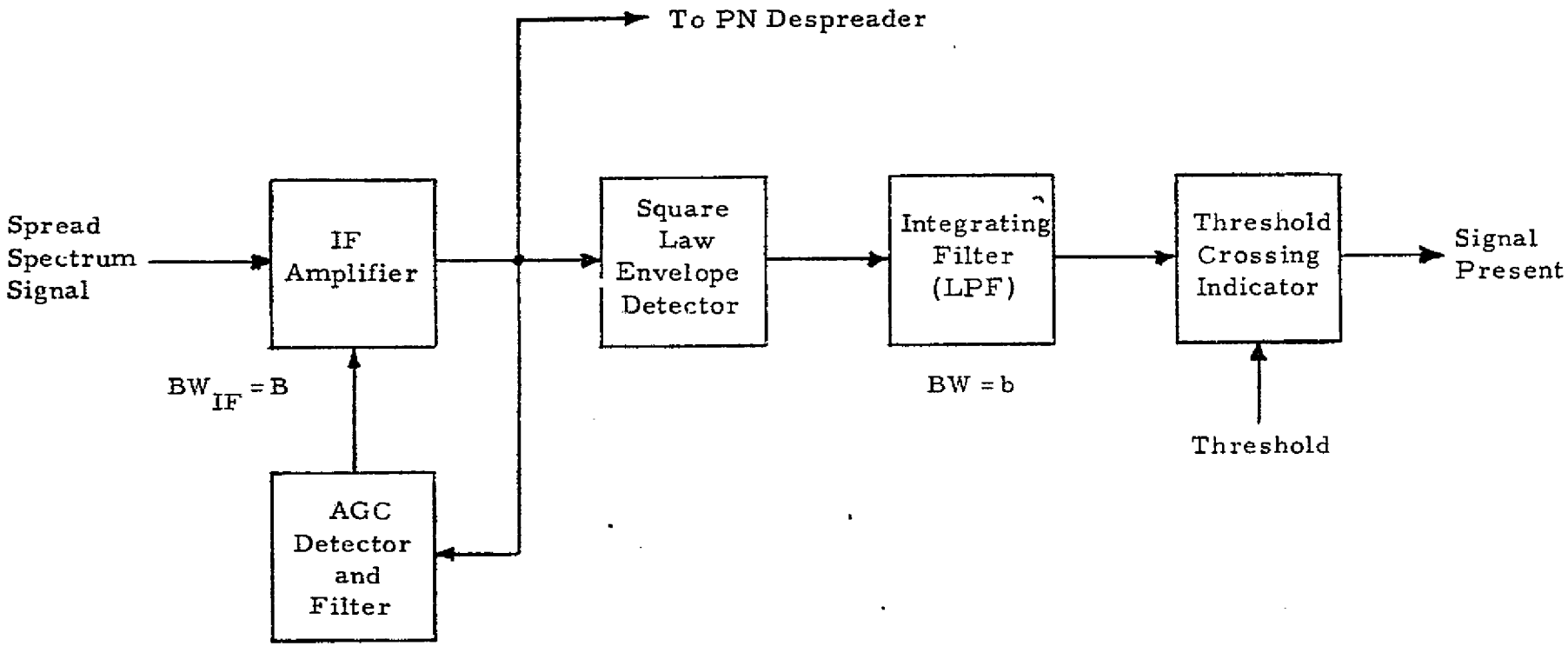


Figure 1. Constant Envelope Signal Detector

Table 1. Forward (TDRS-to-Orbiter) Ku-Band Link Parameters

	Parameter	Units	Value
1.	TDRS EIRP	dBw	40*
2.	Space Loss (13.775 GHz; 22,786 nmi)	dB	-207.7
3.	Orbiter Antenna Gain	dB	34.6**
4.	Total Receive Power, P_{rec} (Sum 1 through 3)	dBw	-133.1
5.	Orbiter System Noise Temp (10 log T_{sys})	dB-K	30.7
6.	Boltzman Constant (10 log 1.38×10^{-23})	dBw/K-Hz	-228.6
7.	Orbiter Noise Spectral Density, N_0 (Sum 5 and 6)	dBw/Hz	-197.9
8.	Total Received Power/Noise Spectral Density, P_{rec}/N_0 (7 minus 8)	dB-Hz	64.8

*Applies to TDRS acquisition phase only.

**Based on a 20-inch diameter parabola.

If we assume a rectangular IF filter* at the receiver, the loss as compared to a matched filter is about 0.9 dB and the $B\tau$ product is 1.37. The IF SNR for the spread spectrum signal of 14.1 Mbps (71.5 dB) is therefore

$$(SNR)_{IF} = 64.8 \text{ dB-Hz} - 0.9 \text{ dB} - 71.5 \text{ dB} = -7.6 \text{ dB} \quad (2)$$

and the corresponding IF bandwidth is

$$B_{IF} = \frac{1.37}{\tau} = 1.37 \times 14.1 \text{ Mbps} = 19.3 \approx 20 \text{ MHz} . \quad (3)$$

The -7.6 dB value for the IF SNR corresponds to the condition when the Shuttle antenna is pointed directly at the TDRS.

To insure that signal detection occurs when the beam of the Shuttle antenna is off the true LOS within a predetermined amount, we have to set the parameters of our detection circuit accordingly. For example, if we want to achieve a specified probability of signal detection within the -1 dB antenna pattern contour, we set our detection circuitry to declare a "signal present" condition when the IF SNR is -8.6 dB. Similarly, if we decided to achieve signal detection within the -3 dB antenna beamwidth contour, we have to adjust the parameters of the signal detection circuit so that the threshold crossing indicator is actuated at -10.6 dB IF SNR.

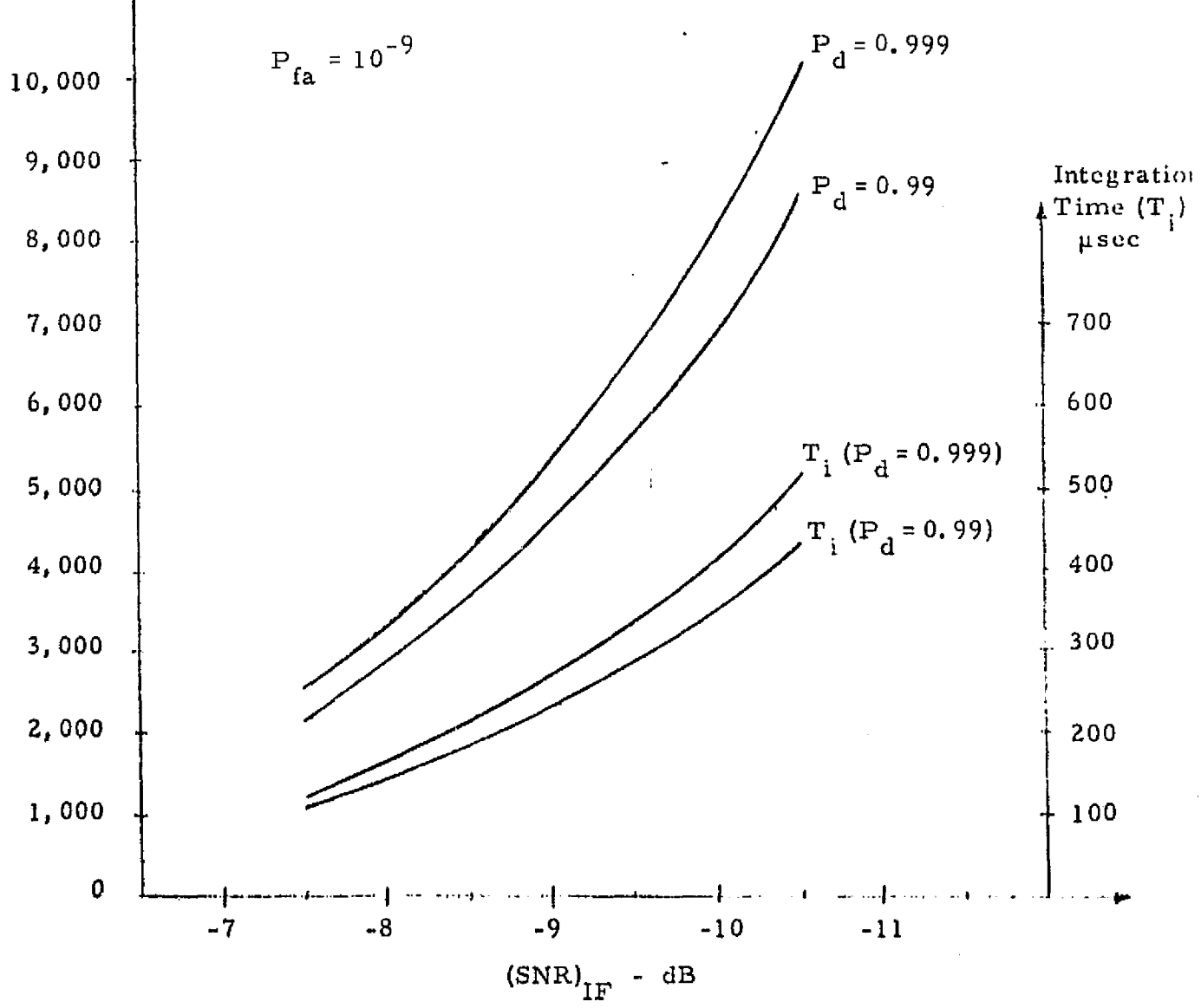
For any given combination of probability of detection, P_D, and the probability of false alarm, P_{fa}, the parameter which we can vary is the number of samples integrated, N. This number is approximately equal to the ratio of IF bandwidth to the bandwidth of the post-detection integrating filter, i.e., $N \cong B/b$. Figure 2 shows N versus IF SNR for two values of P_D and $P_{fa} = 10^{-9}$. The corresponding integration time, T_i, based on the relationship $N \cong B/b$, is also shown in this figure.** From this figure, we

---

*Butterworth or Chebichev design.

**Computational details are given in Addendum A.

Figure 2. Number of Samples Integrated and Approximate Integration Time vs. IF Signal-to-Noise Ratio



see that to achieve specified detection criteria (use $P_D = 0.999$) with the -1 dB antenna pattern ($SNR_{IF} = -8.6$ dB), about 4,100 signal samples must be integrated and the corresponding integration time is about 220 microseconds. As shown in the next section, the integration time is short compared to dwell times achievable with a mechanically scanned antenna. Also, the false alarm time for the parameters used is

$$T_{fa} = \frac{T_i}{P_{fa}} = \frac{220 \times 10^{-6} \text{ sec}}{10^{-9}} = 220 \times 10^3 \text{ seconds}$$

$$\text{or about } \underline{61 \text{ hours}} . \qquad (4)$$

This time is well within any reasonable specification.

We may also consider a special case for which we do not have a paramp in the sum channel of the Shuttle Ku-band receiver. This may be a case where only a single 216 kbps data stream spread by the 14.1 Mbps PN sequence is required to be transmitted to the Shuttle. The noise figure of the sum channel will then be about 8 dB and the corresponding noise temperature is $T_s = (6.3)(290^\circ K) = 1825^\circ K$ or 32.6 dB-K. This will degrade the $(SNR)_{IF}$ by about 1.9 dB, resulting in $(SNR)_{IF}$ of -10.5 dB. From Figure 2 we see that, at this SNR, about 10,300 samples have to be integrated for $P_D = 0.999$ and the integration time is increased to about 515 microseconds. This is still a short time compared to the dwell times encountered with a mechanically scanned antenna.

3.0 ANGULAR SEARCH PROCEDURE

Having established the criteria for detecting the presence of the signal within the 1 dB antenna beamwidth, we now consider the angular search procedure which insures that the designation uncertainty volume is scanned to provide a predetermined probability of true LOS intercept. For this we consider the relationship between the antenna beamwidth and the designation uncertainties. The baseline numbers are

$\theta_B \cong 2.8^\circ \equiv$ Antenna 3 dB beamwidth (20-inch parabol.. dish)

and $\sigma_A = \sigma_E = 0.933^0 \equiv$ Designation uncertainties of Az and El axes (based on $3\sigma_A = 3\sigma_E = 2.8^0$).

It is reasonable to assume that σ_A and σ_E are independent, have zero mean, and are Gaussian-distributed along their respective coordinates. The conversion to polar coordinates then results in a probability distribution of the true LOS around the designated LOS. This distribution has the form:

$$p(\theta,\alpha) = \frac{\theta \exp(-\theta^2/2\sigma_\theta^2)}{2\pi\sigma_\theta^2}, \quad \begin{matrix} 0 \le \theta \le \infty \\ 0 \le \alpha \le 2\pi \end{matrix} \tag{5}$$

where $\sigma = \sigma_A = \sigma_E$, and θ represents the antenna boresight deviation along any redial direction away from the designated LOS. The θ-dependent portion of the two dimensional distribution given by (5), i.e.,

$$p(\theta) = \frac{\theta}{\sigma_\theta^2} \exp(-\theta^2/2\sigma_\theta^2), \quad 0 \le 0 \le \infty \tag{6}$$

is known as the Rayleigh distribution and its general form is shown in the upper part of Figure 3. The angular direction, α, of this distribution, however, can be anywhere in the interval $0 \le \alpha \le 2\pi$. In other words, one can picture the probability distribution of the true LOS as a Rayleigh-shaped "wind vane" which is "pivoted" around the designated LOS and whose direction α is uniformly distributed across the interval from 0 to 2π.

Because we do not know the direction along which the true LOS distribution lies, we have to search out all possible directions and also move the Shuttle antenna beam outwards until the cumulative probability of intercepting the true LOS reaches the predetermined value, such as 0.99 or 0.999, etc. The cumulative probability of intercepting the true LOS as the beam completes successive revolutions and moves "outward" in the θ-direction is

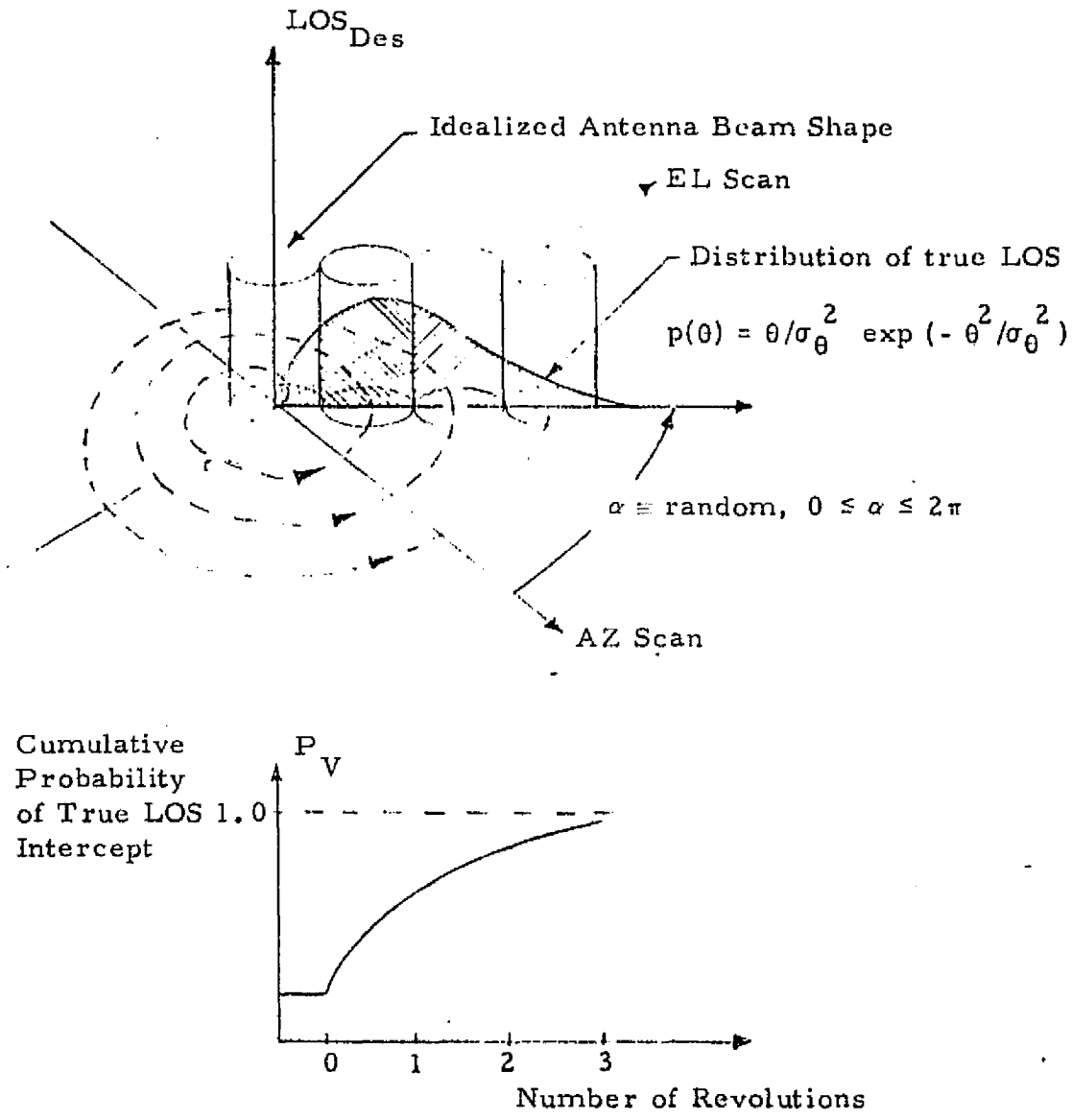


Figure 3. Spiral Scan Provides Means for Interception and Detection of True LOS

$$P_V(\theta_L) = \int_0^{\theta_L} \frac{\theta}{\sigma_\theta^2} \exp\left(-\theta^2/\sigma_\theta^2\right) d\theta = 1 - e^{-\theta_L^2/2\sigma_\theta^2} \qquad (7)$$

where θ_L is the limit of the "outward" beam displacement.

To accomplish the required search, one can use either a stepped circular or a spiral scan. With a stepped circular scan the beam is moved outward along either the Az or the El axis by a preselected angular distance and then a circular scan is initiated. If the true LOS is intercepted during this scan the antenna motion is either terminated immediately after signal detection, so that the autotrack can lock on to the true LOS, or the detection coordinates are stored in a memory and the beam returned to these coordinates upon completion of the scan. If detection is not indicated, the beam is moved outward and a circular scan along a wider radius is performed. This procedure is repeated until the true LOS to TDRS is intercepted and detected.

With a spiral scan, which is more common,* the antenna beam is moved outward from its initial position in a spiral pattern which eventually covers all possible locations of the true LOS up to a predetermined value of the cumulative intercept probability. Figure 3 shows how an idealized antenna beam is moved in a spiral pattern to "cut through" the probability density function of the true LOS. The lower portion of Figure 3 shows how each successive "cut" through the density function contributes to the cumulative probability of true LOS intercept.

Two facts must be noted while we are considering the idealized search depicted in Figure 3. One is that the "height" of the idealized antenna beam shape (shown as a cylinder in the figure) represents the P_D, i.e., the probability of signal detection as determined by the parameters of the detector shown in Figure 1 and the available IF SNR. IF $P_D < 1$, then the contribution of the individual "cuts" through the true LOS distribution must be weighed accordingly. The second fact evident from Figure 3

*Stepped scan requires fast changes in antenna direction and thus increases servo loading.

is that even without scanning there is a finite probability that the true LOS is within the beamwidth regardless of α. The value of this probability depends on the relationship between the antenna beamwidth and the "spread" of the Rayleigh distribution.

First let us consider the Shuttle's antenna pattern. For this we will employ a commonly used assumption that, within a considerable portion of the main lobe, the antenna pattern can be approximated by a Gaussian form:

$$G(\theta) = \exp\left(-2.776\ \theta^2/\theta_B^2\right) \tag{8}$$

where $G(\theta)$ = normalized boresight gain (maximum gain)

θ_B = antenna 3 dB beamwidth

θ = angular deviation from boresight .

The loss, in dB, as a function of θ is therefore .

$$L(\text{dB}) = 10 \log\left[\exp\left(-2.776\ \theta^2/\theta_B^2\right)\right] \quad . \tag{9}$$

The plot of L versus θ for $\theta_B \cong 2.8^\circ$ is shown in the upper left corner of Figure 4. The cumulative probability of true LOS intercept, P_V, versus θ is also shown. From plots of Figure 4 it is evident that using -1 dB antenna pattern contour the initial value of P_V prior to scan is only 0.30. Note that only a half of the antenna pattern is shown prior to scan. As the spiral scan commences the boresight begins to move outward and the -1 dB contour begins to "cut" into the $p(\theta)$, thus contributing to the cumulative probability, $P_V(\theta)$, of true LOS intercept.

Suppose we decide to make $P_V = 0.9999$. From the $P_V(\theta)$ plot of Figure 4 we see that this requires "spiraling out" the -1 dB point from 0.8 to 4.0 degrees. If we spiral out at the rate of 0.8° per 360° turn, the number of turns required, n_t, is

$$n_t = \frac{\text{Distance to maximum } \theta}{\Delta\theta \text{ per turn}} = \frac{4.0^\circ - 0.8^\circ}{0.8^\circ/\text{turn}} = \frac{3.2^\circ}{0.8^\circ/\text{turn}}$$

$$= 4 \text{ turns} . \tag{10}$$

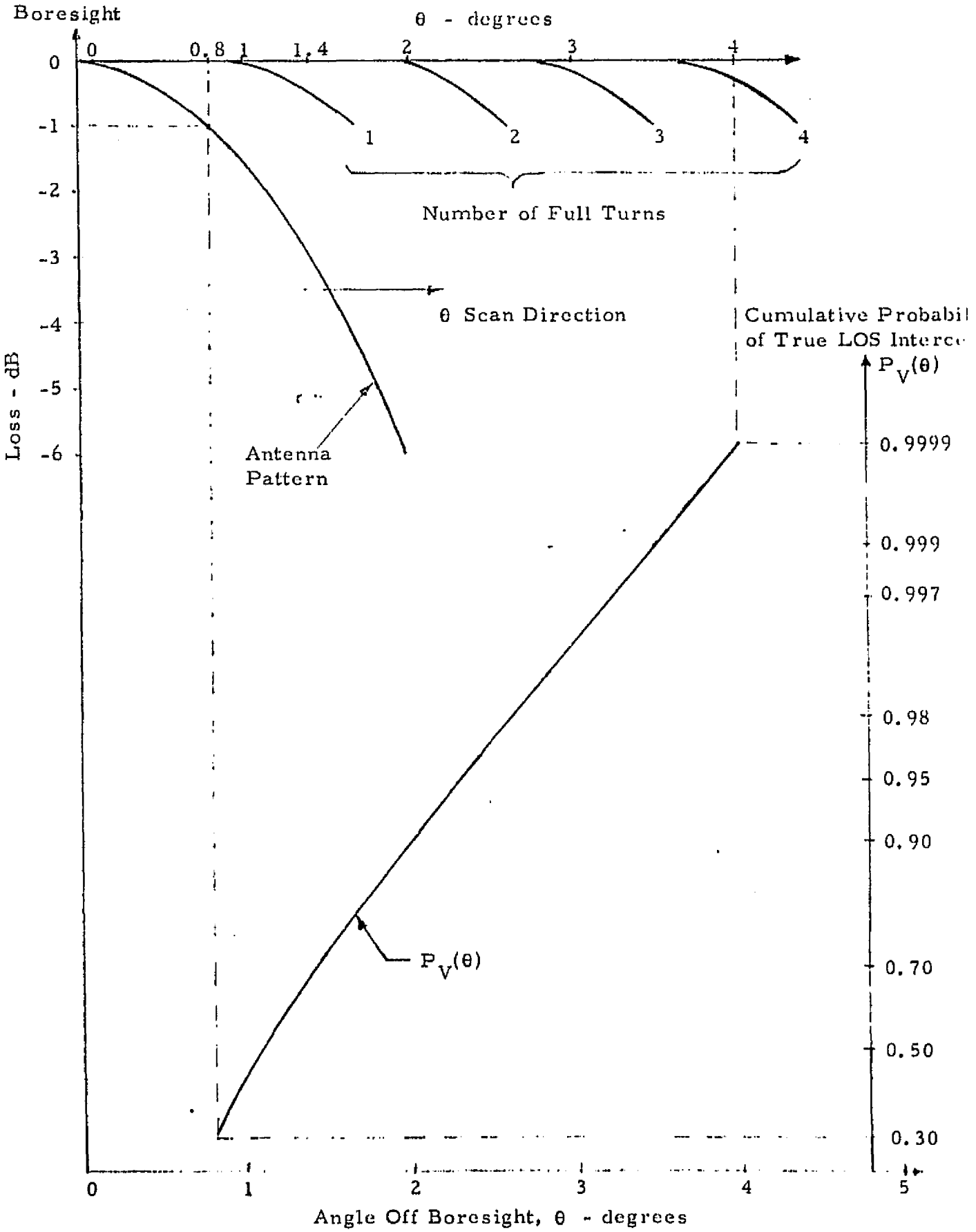


Figure 4. Antenna Pattern and Cumulative Probability of True LOS Intercept as a Function of θ (Angle Off Boresight)

With these parameters, we barely touch the 4^0 contour and thus we are not certain what dwell time we have to stop the scan when we reach our selected limit. Thus, we increase the rate of $\Delta\theta$ advance. If we let $\Delta\theta$ be 0.9^0 per turn, then n_t becomes

$$n_t = \frac{3.2^0}{0.9^0/\text{turn}} = 3.56 \text{ turns} . \tag{11}$$

But we let the scan continue to an even 4 turns. Then the excess is

$$\Delta n_t = 4 - 3.56 = 0.44 \text{ turns} . \tag{12}$$

The time spent beyond the 4^0 limit then becomes

$$\Delta t = \Delta n \times \frac{1}{\text{turn/sec}} = \Delta n \times \frac{1}{f_s} . \tag{13}$$

As explained in the next section, the turn/sec is the scan frequency f_s which determines the dwell time of any antenna beam contour (-1 dB, -3 dB, etc.) and also the time required to spiral out the boresight to any required cone limit. For the example chosen, the -1 dB contour and the boresight are spiraled out to 4.2^0 and 3.6^0, respectively. Also as stated in the next section, if we assume servo loop bandwidth of 1 Hz and we want the servo to pull on the first pass over the true LOS, the scan frequency f_s should be 0.425 Hz. The time of overscan, t_{os}, is then

$$t_{os} = \Delta n_t \times \frac{1}{0.4} = \frac{0.44}{0.425} = 1.04 \text{ seconds} . \tag{14}$$

This time is far in excess of the detection time of about 220 microseconds! The time to complete the four turns at the scan frequency of 0.425 Hz is about 9.4 seconds.

The combined minimum probability of true LOS intercept and detection for the procedure described is

$$P_T(\text{min}) = P_D \times P_V = (0.999)(0.9999) = 0.9989 . \tag{15}$$

The actual probability P_T may be somewhat larger because of the overscan we have allowed for the P_V term. Nevertheless, the procedures outlined above permit one to select various combinations of P_D and P_V to obtain the desired value of P_T(min).

When the energy detector commands the Shuttle antenna scan mechanism to terminate the search, the antenna is stopped and the track mode is initiated. There may be a pull-in transient, but since we have assumed that the antenna is stopped within the -1 dB antenna pattern contour, the servo operation should be in its linear region. A conservative estimate is that the pull-in time should be approximately one third of the inverse of the servo loop bandwidth. For a 1 Hz loop, this implies that about 0.3 second will be required for the servo loop to go from scan to track mode.

At the time when the search is terminated, the value of P_{rec}/N_0 may be 1 dB below the nominal 64.8 dB-Hz value, i.e., it may be about 64 dB-Hz. The recent analysis* of the Shuttle monopulse system indicates that, with the paramp in the sum channel only and no despreading, the maximum RMS error due to thermal noise at about 64 dB-Hz is less than 0.05 degree. Consequently, we can safely assume that the thermal noise is not a limiting factor for establishing a reliable monopulse autotrack at P_{rec}/N_0 values as low as 64 dB-Hz.

When the Shuttle antenna begins to track the TDRS signals, commands to correct the TDRS pointing are sent to the ground via an S-band link. As the TDRS pointing is zeroed in at the Shuttle, the P_{rec}/N_0 available to the Shuttle receiver increases by about 10 dB. The code sync and doppler acquisition then take place and the Ku-band link communication is established.

---

*C. L. Weber, "Shuttle Monopulse System for Ku-Band Communication Signal from TDRS - Part II," Axiomatix Report No. R7509-4 (under Contract NAS 9-14614), September 26, 1975.

4.0 EXTENDED SCAN COVERAGE AND COMPATIBILITY WITH RADAR SEARCH MODE SCAN

In the preceding section, we have considered only the requirements for searching over the residual designation uncertainties of the TDRS LOS acquisition. In practice, however, various factors may increase these uncertainties and a scan over a wider angular area may be required. But the search requirements for some of the radar modes of the integrated radar/communication system do require scan over relatively wide angles and thus the adaptation of the radar scan modes to communication angular search appears feasible. In particular, the $\pm 5^0$ (10^0 cone) radar search mode for the cooperative target should be considered for an adaption to the communication angular search mode. As shown below, scan over wider areas can also be performed for the communication mode at the expense of increased search time. For all of the quantitative discussions which follow, it is assumed that a spiral scan is utilized. The detailed analysis of spiral scan is given in Appendix I.

There are two basic spiral scans to consider. One involves constant angular velocity, i.e., the rate at which the boresight is rotated around its starting position is constant. This results in a progressively increasing tangential velocity as the boresight is spiraled out. As a consequence, the dwell time decreases and eventually it may fall below the value required by the servo to acquire the signal source (TDRS or a radar target) on the first pass. The antenna must then make at least one or more revolutions before acquisition. Figure 5 shows the relationship between the scanned angle (conical coverage) and the dwell time versus scan frequency (rev/sec) and the elapsed scan time. The equations used are

$$\theta = \Delta\theta \, f_s \, t \tag{16}$$

and

$$t_d = \frac{1}{2\pi} \frac{\theta_B}{\Delta\theta} \left(\frac{1}{f_s}\right)^2 \frac{1}{t} \tag{17}$$

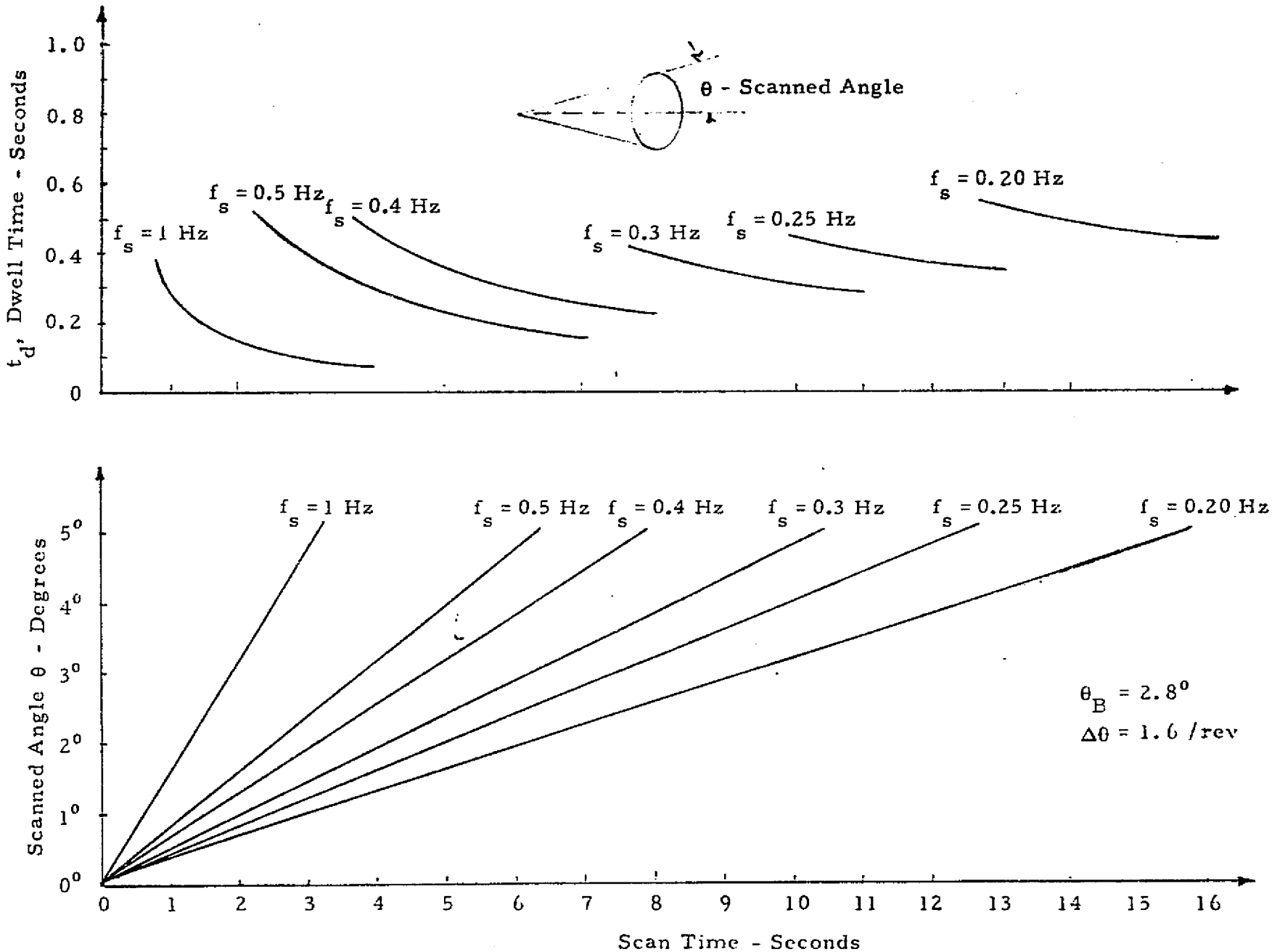


Figure 5. Scanned Angle and Dwell Time Versus Scan Frequency and Scan Time

where θ = scanned angle

f_s = scan frequency

θ_B = 3 dB antenna beamwidth

$\Delta\theta$ = radial advance in deg/rev .

For data in Figure 5, we used $\theta_B = 2.8^0$ and $\Delta\theta = 1.6^0$/sec which corresponds to a -1 dB overlap.

As an example for use of Figure 5, assume that we want to determine the time to scan out a $\pm 5^0$ cone, with a 1 Hz second-order servo loop. A second-order servo requires an acquisition time of about one third of its inverse bandwidth. Therefore, a 1 Hz loop has to be scanned at an 0.3 Hz rate to insure that at the 5^0 limit the dwell time is not less than about 0.3 seconds. The scan time required to reach this limit is then about 10.5 seconds. Slower scan frequencies require longer times to reach any given scan limit but they also provide longer dwell times.

Because the dwell time at the maximum scanned angle is a major factor which affects servo acquisition behavior at the scan limit, the relationship between the parameters which determine this dwell time is important. As derived in Appendix B, the dwell time at maximum scan limit is

$$t_{dm} = \frac{1}{2\pi} \frac{\theta_B}{\theta_m} \frac{1}{f_s} . \tag{18}$$

The time to scan out to the maximum angular limit is

$$t_m = \frac{\theta_m}{\Delta\theta f_s} \tag{19}$$

where $\Delta\theta$ is the radial advance in deg/rev.

Figure 6 shows plots of t_m and t_{dm} versus f_s for various values of θ_m and $\Delta\theta$. The baseline value of $\theta_B = 2.8^0$ is assumed for all plots of t_{dm}. From Figure 6 is is evident that the 0.9^0/rev overlap chosen as an example in Figure 4 is conservative and that a shorter scan time can be obtained, i.e., 5.2 seconds versus 9.4 seconds if the $\Delta\theta$ is increased to 1.6^0/rev.

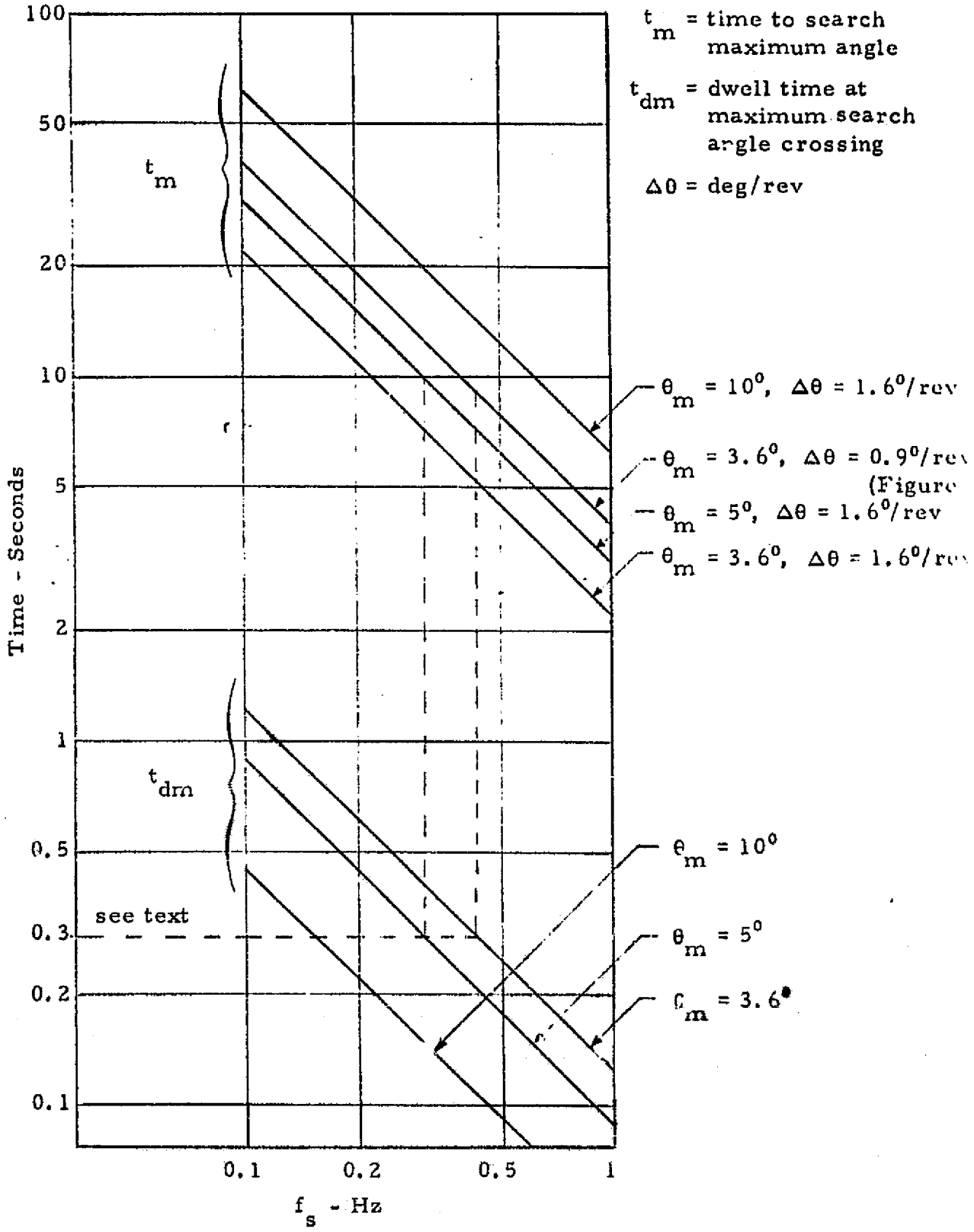


Figure 6. Scan Time to Maximum Search Angle and Dwell Time of Maximum Angle Versus Scan Frequency f_s (Constant Angular Velocity Scan)

The 1.6^0/rev value corresponds to -1 dB worst case beam overlap and thus does not violate the energy detection criterion. This applies to the 1 Hz loop and 0.3 second minimum dwell time assumed as the requirement for servo acquisition.

So far we have discussed the scan which utilizes constant rotational velocity. To achieve constant tangential velocity and constant dwell time over the scan path, the scan frequency has to decrease inversely with time. This, however, implies that at the start of the angular search the scan frequency must be infinite. To avoid this requirement, the actual scan program starts out with a constant angular velocity scan and then switches over to a constant tangential velocity scan. The switchover takes place when the minimum allowable dwell time is reached. For example, from Figure 5 we see that, for f_s = 1 Hz, the 0.3 second dwell time is reached after about 1 second of scan time. If the servo bandwidth remains constant at 1 Hz, compatible with the initial f_s' = 1 Hz, the scan frequency must then decrease to keep the dwell time at a constant value of approximately 0.3 second. The program for the time-variable scan frequency is then

$$f_s(t) = f_s' \qquad 0 \le t \le t_0 \tag{20}$$

$$f_s(t) = f_s' \sqrt{\frac{t_0}{t}} \qquad t_0 \le t \tag{21}$$

where f_s' is the initial constant scan frequency and t_0 is the time after which the scan frequency is slowed down.

Figure 7 shows the application of the time-variable scan frequency. The criterion for the data shown is to achieve $\pm 5^0$ scan limit with dwell time no less than 0.3 seconds and a 1 Hz servo loop. Three cases are shown. Case 1 starts with f_s = 1 Hz constant scan frequency and then switches to constant dwell time scan after t_0 = 1 second. Case 2 starts with f_s = 0.5 Hz and switches over at t_0 = 4 seconds. The Case 3 scan uses a constant f_s value of 0.31 Hz. The most significant result indicated

$\Delta\theta = 1.6^0$/rev (-1 dB overlap)

$\theta_B = 2.8^0$ (3 dB beamwidth)

(1) $f_s = 1$ Hz until $t_0 = 1$ sec

(2) $f_s = 0.5$ Hz until $t_0 = 4$ sec

(3) $f_s = 0.31$ Hz until scan limit

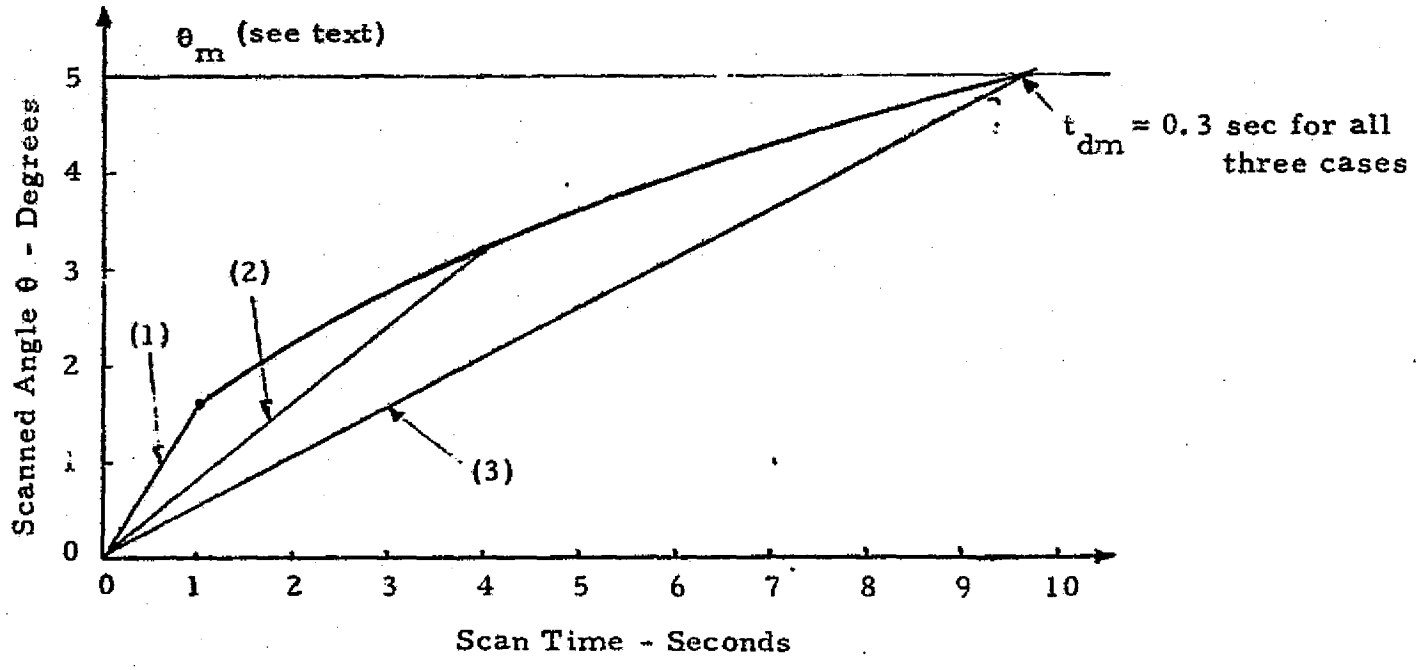


Figure 7. Scan Time to Reach Scan Limit is Independent of Scan Method Used

by this example is that, for a selected dwell time at a given search limit,* the time to reach this limit is independent of the scan method used. But because the constant tangential velocity scan is most efficient from the standpoint of providing uniform dwell time over the major portion of the search area, the switchover to this scan should take place as soon as the minimum allowable dwell time is reached. Thus, for a 1 Hz loop and the $f_s' = 1$ Hz, the switchover should take place at $t_0 = 1$ second.

The data presented so far indicates that times required to search the uncertainty angles for TDRS acquisition are short compared to radar search times. This is due to the fact that SNR available from TDRS for signal detection during acquisition is much higher than the SNR of radar targets. Consequently, the scan time limitation for TDRS acquisition is determined by servo considerations rather than signal detection. The approach for radar search is the same, however, as that for TDRS acquisition and only the dwell time requirements are different. We therefore recommend that the radar search pattern be adapted for TDRS acquisition when the integrated system is in the communication mode.

5.0 CONCLUSIONS

The energy detection method for determining the presence of the spread spectrum TDRS signal within the beamwidth of the scanning Shuttle antenna requires integration times in the order of several orders of magnitude shorter than the dwell times associated with a mechanically scanned antenna. The latter times are typically in the order of seconds or fractions of seconds. Consequently, the time required to search out the initial designation inaccuracies of the Shuttle antenna is determined by the servo drive capabilities and scan pattern rather than the signal detection technique. This applies even to the reduced power mode when the effective EIRP from TDRS is 40 dBw instead of the nominal 50 dBw.

---

*In this case, approximately 0.3 second at a 5^0 limit.

For a Gaussian-distributed, uncorrelated, and zero mean value deviation and azimuth designation errors, the transformation to polar coordinates presented in this appendix provides a convenient analytical technique for estimating the cumulative probability of true LOS intercept as the function of spiral (or stepped circular) scan strategy.

Because the dwell times required for the angular acquisition of the radar targets are typically in excess of the times required to detect TDRS LOS, we recommend that these scan programs be used, with minor modifications, for the establishment of the Ku-band communication link.

ADDENDUM A

SIGNAL INTEGRATION REQUIREMENTS VERSUS SIGNAL-TO-NOISE RATIO

The energy detector used to determine the presence of the spread spectrum TDRS signal within the beamwidth of the scanning Shuttle antenna employs a square law device followed by an integrating low pass filter and a decision threshold device. The square law device is preceded by an IF filter whose bandwidth is approximately matched to the bandwidth of the spread spectrum signal. Because the incoming signal has a constant envelope there is no need to perform its actual sampling at specific intervals to determine the number of pulses integrated required to achieve a given probability of detection versus IF signal-to-noise ratio. In this case, it is the ratio of IF to low pass filter bandwidth which determines the equivalent number of pulses, or more specifically, signal samples integrated. The analysis used for discrete sampling can be applied to this problem, however, to determine the effective number of samples integrated and the corresponding integration time required to achieve a given probability of detection. For this, we use the following equation:*

$$SNR_{IF} = \frac{[\phi^{-1}(P_{fa}) - \phi^{-1}(P_d)]}{\sqrt{N}} \qquad (A-1)$$

where

SNR_{IF} = the rms signal-to-noise ratio in IF

N = number of pulses (or samples) integrated

$\phi^{-1}(P_{fa})$ = function related to false alarm probability P_{fa}

$\phi^{-1}(P_d)$ = function related to probability of detection P_d.

The functions $\phi^{-1}(P_{fa})$ and $\phi^{-1}(P_d)$ are given by Figures 10.7-1 (p. 368) and 10.7-2 (p. 369), respectively, in the reference cited.

---

*DiFranco and Rubin, Radar Detection, Prentice-Hall, Inc., p. 367.

For one of the cases, we select $P_d = 0.999$ and $P_{fa} = 10^{-9}$. For these values, $\phi^{-1}(P_d) = -3.10$ and $\phi^{-1}(P_{fa}) = 5.95$. Substituting in Equation A-1:

$$SNR = \frac{[5.95 - (-3.10)]}{\sqrt{N}} = \frac{9.05}{\sqrt{N}} . \qquad (A-2)$$

Solving for N, we obtain:

$$N = \frac{(9.05)^2}{(SNR)_{IF}^2} = \frac{81.9}{(SNR)_{IF}^2} . \qquad (A-3)$$

For example, for $SNR_{IF} = -10$ dB, the numeric value is 0.1. Substituting this value into A-3 we get:

$$N = \frac{81.9}{(0.1)^2} = \frac{81.9}{0.01} = 8190 \text{ samples integrated} . \qquad (A-4)$$

Since the sample time is approximately equal to the reciprocal of the IF bandwidth, the corresponding integration time will be:

$$T_i \cong \frac{N}{B_{IF}} = \frac{8190}{20 \times 10^6 \text{ MHz}} = 410 \times 10^{-6} \text{ sec}$$

or 410 microseconds . (A-5)

Similar procedures are used for other values of $(SNR)_{IF}$ and also for the case of $P_d = 0.99$. Figure 2 of this report shows the plot of N and corresponding T_i for $(SNR)_{IF}$ ranging from -7.5 dB to -10.5 dB.

APPENDIX I

SPIRAL SCAN ANALYSIS

Gaylord K. Huth
Sergei Udalov

APPENDIX I

SPIRAL SCAN ANALYSIS

by

Gaylord K. Huth
Sergei Udalov

This appendix analyzes the relationship between the dwell time, the total scan time, and the scan rate for the following types of spiral scans:

(1) spiral scan with constant rotational velocity;

(2) spiral scan with constant tangential velocity;

and (3) hybrid scan which starts out with constant rotational velocity and then switches over to constant tangential velocity.

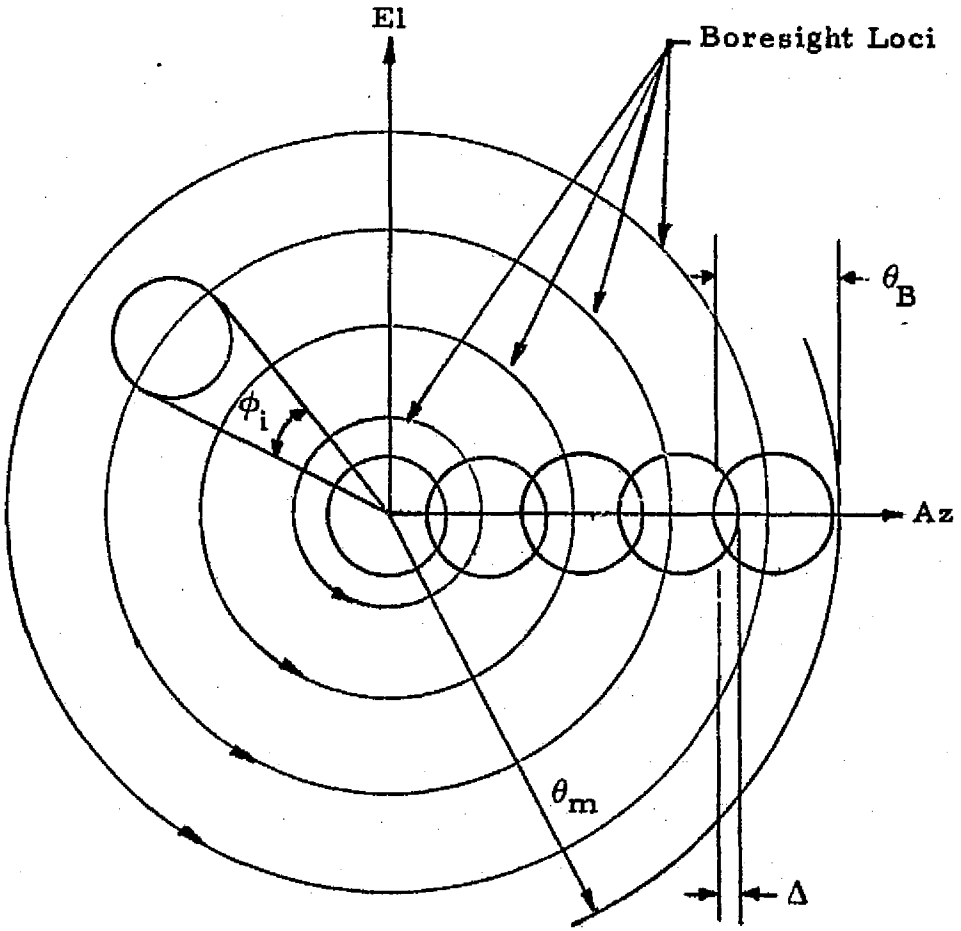
It will be shown that the hybrid spiral scan offers the best compromise between the antenna servo power requirements and the dwell time. Several numerical examples pertinent to the Shuttle Ku-band radar/communication system are provided.

1.0 INTRODUCTORY ANALYSIS - STEPPED CIRCULAR SCAN

Before considering the spiral scan, it is instructive to develop first the relationship between the major parameters for the stepped circular scan, which is a discrete approximation of the continuous spiral scan. Figure 1 defines the basic parameters of the stepped circular scan. After each complete rotation about boresight, the antenna is stepped to the next radius for the next rotation. The number of complete rotations N can be found by noting that

$$\theta_m = \frac{\theta_B}{2} + N\,\theta_B(1 - \Lambda) \qquad (1)$$

where θ_m is maximum scan limit, θ_B is the antenna beamwidth, and Λ is the scan overlap. Equation (1) can be rewritten to determine N:



θ_B = beamwidth (3 dB)

θ_m = maximum scan limit

Δ = beam overlap

ϕ_i = angle of antenna dwell on ith rotation

Figure 1. Stepped Circular Scan Increases the Dimensions of Boresight Loci by Discrete Increments

$$N = \left\lceil \frac{\theta_m - \frac{\theta_B}{2}}{\theta_B(1-\Delta)} \right\rceil \tag{2}$$

where $\lceil x \rceil$ is the largest integer greater than x.

The dwell time on the ith rotation is denoted by t_{d_i} and is given by

$$t_{d_i} = \frac{\phi_i}{f_i\, 2\pi} = \frac{4}{f_i\, 2\pi} \sin^{-1} \frac{\theta_B}{4i\,\theta_B(1-\Delta)} = \frac{2}{f_i\,\pi} \sin^{-1} \frac{1}{4i(1-\Delta)} \tag{3}$$

where f_i is the rotational frequency (i.e., revolutions per second) for the ith revolution and the radius of the ith revolution is $i\,\theta_B(1-\Delta)$. Note that, for reasonable values of Δ (i.e., 30% or less), equation (3) can be rewritten as

$$t_{d_i} \simeq \frac{1}{f_i\, 2\pi\, i(1-\Delta)} \tag{4}$$

with only a 2.2% error or less for i=1 and 0.5% error or less for i=2. The total scan time T_s, i.e., the time required to reach the scan limit boundary θ_m is

$$T_s = t_{d_0} + \sum_{i=1}^{N} \frac{1}{f_i} + N\tau \tag{5}$$

where t_{d_0} is the dwell time on the beamwidth at boresight and τ is the time to step between scan radii. For constant rotational velocity f_c,

$$T_s = t_{d_0} + N\left(\frac{1}{f_c} + \tau\right) = t_{d_0} + \frac{\theta_m - \frac{\theta_B}{2}}{\theta_B(1-\Delta)} \left(\frac{1}{f_c} + \tau\right). \tag{6}$$

For constant tangential velocity,

$$t_{d_0} = t_{d_i} \simeq \frac{1}{f_i\, 2\pi\, i(1-\Delta)} \tag{7}$$

or

$$f_i \simeq \frac{1}{t_{d_0}\, 2\pi\, i(1-\Delta)}\,. \tag{8}$$

Therefore, $$T_s = t_{d_0}\left[1 + 2\pi(1-\Delta)\sum_{i=1}^{N} i\right] + N\tau$$

$$= t_{d_0}\left[1 + \pi(1-\Delta)\,N(N+1)\right] + N\tau \,. \qquad (9)$$

But, using equation (2) for N,

$$N(N+1) = \frac{\theta_m^2 - \theta_m\theta_B\Delta - \frac{\theta_B^2}{4}(1-2\Delta)}{\theta_B^2(1-\Delta)^2} \,. \qquad (10)$$

Hence, for constant tangential velocity,

$$T_s = t_{d_0}\left[1 + \pi\left(\frac{\theta_m^2 - \theta_m\theta_B\Delta - \frac{\theta_B^2}{4}(1-2\Delta)}{\theta_B^2(1-\Delta)}\right)\right] + \frac{\theta_m - \frac{\theta_B}{2}}{\theta_B(1-\Delta)}\tau \,. \qquad (11)$$

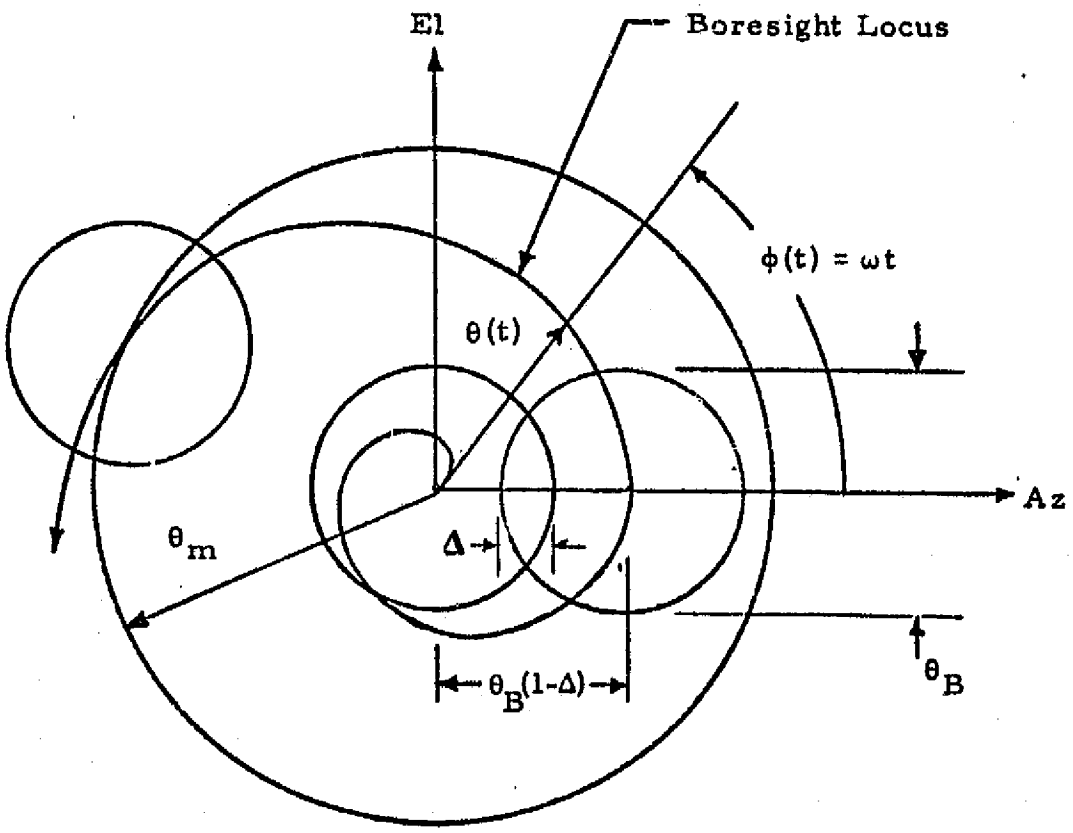
2.0 CONSTANT ROTATIONAL VELOCITY SPIRAL SCAN

With the antenna axis and the monopulse comparator arranged for tracking along two orthogonal axes, such as Az and El, the spiral scan can be implemented by driving each axis with an orthogonal sinusoid of progressively increasing amplitude. The boresight displacement along each of the orthogonal axes is then

$$\Delta Az = \theta(t)\cos\omega t$$
$$\Delta El = \theta(t)\sin\omega t \qquad (12)$$

where $\theta(t)$ is the time-dependent function which determines radial boresight displacement and $\omega = 2\pi f_s$ = angular rate of boresight displacement. The f_s will be referred to henceforth as the "scan frequency." This frequency may be equal to or may be smaller than the servo bandwidth, but it should not exceed the latter.

In polar coordinates, the boresight locus is defined by a spiral $\theta(t) < \phi(t)$ shown in Figure 2. In Figure 2, $\theta_B(1-\Delta)$ is the boresight displacement per one complete turn of $\phi(t)$.



$\theta_B(1 - \Delta)$ = deg/rev

θ_B = beamwidth (3 dB)

θ_m = maximum scan limit

Δ = beam overlap

Figure 2. Spiral Scan Increases the Dimension of the Boresight Locus in a Continuous Manner

Let $\theta(t)$ be proportional to $\phi(t)$. Then, $\theta(t)$ is given by

$$\theta(t) = \frac{\theta_B(1-\Delta)}{2\pi}\phi(t) = \frac{\theta_B(1-\Delta)}{2\pi}\omega t = \theta_B(1-\Delta)f_s t . \quad (13)$$

From (13), we can determine the time T_s required to spiral out the boresight to any given maximum angular (conical) limit, θ_m. Thus,

$$T_s = \frac{\theta_m - \frac{\theta_B}{2}}{f_s\theta_B(1-\Delta)} . \quad (14)$$

Equation (14) can be used directly only if f_s (i.e., scan frequency) is constant. Note that (14) is very close to (6) for the stepped circular scan with τ equal to 0. Thus, (14) applies to a constant angular velocity scan. For a constant tangential velocity, f_s is time-dependent and (14) has to be modified accordingly, as explained in Section 3.0.

As the boresight position spirals out toward a given maximum limit, the dwell time decreases if the ω is held constant. The dwell time can be computed by considering the tangential* velocity along the spiral at the corresponding value of $\theta(t)$. Thus,

$$V_t = \theta(t)\,\omega . \quad (15)$$

The dwell time is then considered as the time required by a predetermined antenna beamwidth contour, θ_B (-1 dB or -3 dB, etc.) to traverse a direction in space as the boresight moves across that direction with velocity V_t,

$$t_d = \frac{\theta_B}{V_t} = \frac{\theta_B}{\theta(t)\,\omega} = \frac{\theta_B}{\theta_B(1-\Delta)f_s\omega t} = \frac{1}{2\pi}\,\frac{\theta_B}{\theta_B(1-\Delta)}\left(\frac{1}{f_s}\right)^2\frac{1}{t} . \quad (16)$$

Of particular importance is the dwell time available at the position where the boresight reaches the maximum scan limit θ_m. The expression

*The radial component due to spiraling out is not included here because it is generally low compared to the tangential component of the total velocity along the spiral.

for the value of this dwell time is obtained by substituting (14) into (16):

$$t_{dm} = \frac{1}{2\pi} \frac{\theta_B}{\theta_B(1-\Delta)} \frac{1}{f_s 2} \frac{\theta_B(1-\Delta)}{\theta_m - \frac{\theta_B}{2}} = \frac{1}{2\pi f_s} \frac{\theta_B}{\theta_m - \frac{\theta_B}{2}} . \qquad (17)$$

Equations (13) and (16) can be used to plot the angle scanned and the corresponding dwell time as a function of time and scan frequency, f_s. Correspondingly, equations (14) and (17) can be used to plot the maximum time to reach the scan limit, θ_m, and the corresponding t_{dm} as the function of f_s.

3.0 CONSTANT TANGENTIAL VELOCITY SPIRAL SCAN

For a spiral scan with constant tangential velocity V_t, from (15) with $\omega = 2\pi f_s$,

$$V_t = 2\pi f_s \theta(t) \qquad (18)$$

where $\theta(t)$ is the instantaneous radius given by (13). Hence,

$$V_t = 2\pi \theta_B(1-\Delta) f_s^2 t . \qquad (19)$$

The total scan time T_s can be found from (5) for the circular scan by letting τ equal zero and the sum be an integral over the continuously changing $\theta(t)$. Therefore, note that maximum value of $\theta(t)$ is

$$\max \theta(t) = \theta_m - \frac{\theta_B}{2} . \qquad (20)$$

The instantaneous number of turns completed, n, is the total radius the spiral has reached divided by the advance per turn. That is,

$$n = \frac{\theta(t)}{\theta_B(1-\Delta)} \qquad (21)$$

and the incremental increase is

C-3

$$dn = \frac{d\theta(t)}{\theta_B(1-\Delta)} \quad . \tag{22}$$

Thus, the sum in (5) becomes the integral

$$T_s = \int_0^{\theta_m - \theta_B/2} \frac{1}{f_s} \frac{d\theta(t)}{\theta_B(1-\Delta)} \quad . \tag{23}$$

Neglecting the radial component due to spiraling out, then from (16),

$$t_d = \frac{\theta_B}{V_t} = \frac{\theta_B}{2\pi f_s \theta(t)} \tag{24}$$

or

$$f_s = \frac{\theta_B}{2\pi t_d \theta(t)} \quad . \tag{25}$$

Substituting (25) into (23) results in

$$T_s = \frac{2\pi t_d}{\theta_B^2(1-\Delta)} \int_0^{\theta_m - \theta_B/2} \theta(t)\, d\theta(t) \quad . \tag{26}$$

Also, note that by substituting (25) into (13), one obtains

$$\theta^2(t) = \frac{\theta_B^2(1-\Delta)t}{2\pi t_d} \quad . \tag{27}$$

Therefore,

$$\theta(t)\, d\theta(t) = \frac{\theta_B^2(1-\Delta)\, dt}{4\pi t_d} \tag{28}$$

and by rearranging (27),

$$t = \frac{2\pi t_d \theta^2(t)}{\theta_B^2(1-\Delta)} \quad . \tag{29}$$

The maximum value of t is

$$\max t = \frac{2\pi t_d(\theta_m - \theta_B/2)^2}{\theta_B^2(1-\Delta)} \quad . \tag{30}$$

Substituting (28) and (30) into (26) gives

$$T_s = \frac{1}{2}\int_0^{\max t} dt$$

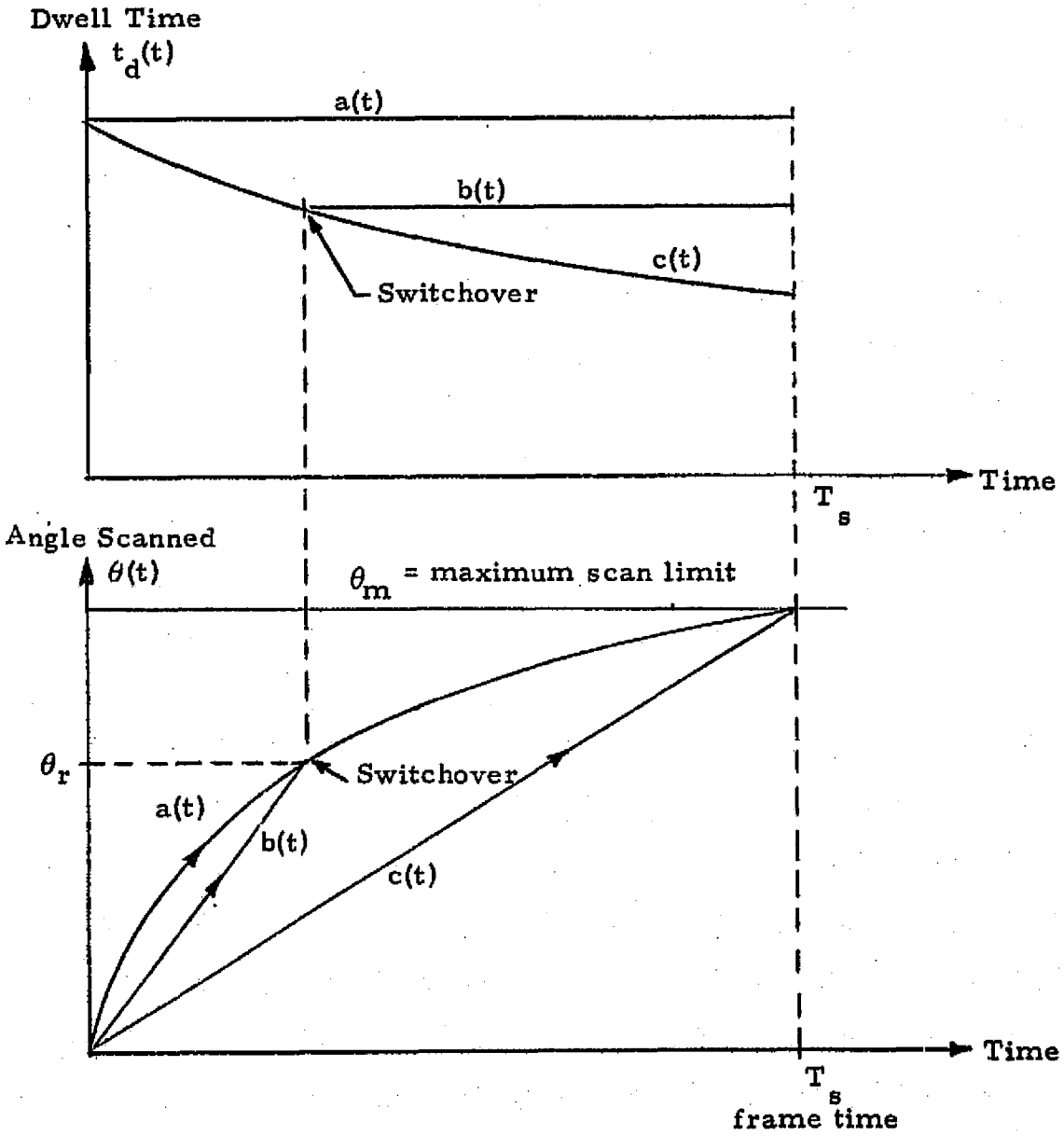
$$= \frac{\pi t_d(\theta_m - \theta_B/2)^2}{\theta_B^2(1-\Delta)} \tag{31}$$

which is very close to the result for stepped circular scan with τ equal to zero.

4.0 HYBRID SPIRAL SCAN

The design of a scan procedure typically requires the maximization of the dwell time for a given total scan time T_s. Figure 3 compares the dwell time for a spiral scan with constant rotational velocity [i.e., scan c(t)] and constant tangential velocity [i.e., scan a(t)]. It can be seen that constant tangential velocity maximizes the dwell time but at $\theta(t) = 0$, it requires infinite rotational velocity as predicted by (25). Therefore, to overcome the infinite rotational velocity at $\theta(t)$ and to increase the dwell time from that obtained by constant rotational velocity, a hybrid scan technique is proposed. The hybrid scan is shown in Figure 3 [i.e., scan b(t)] where initially the scan is performed at a constant rotational velocity and then is switched over to constant tangential velocity. To compute the total scan time of the hybrid scan, let the constant rotational velocity scan switch over to the constant tangential velocity scan at θ_r. Therefore, the radius at switchover is

$$\theta_s(t) = \theta_r - \frac{\theta_B}{2} \quad . \tag{32}$$



Scan a(t): $f_s(0)$ infinite and then progressively decreases to maintain constant t_d.

Scan b(t): f_s constant until switchover, then decreased to maintain constant t_d for the remaining portion of the scan.

Scan c(t): f_s constant to T_s but t_d decreased continuously.

Figure 3. Scanned Angle and Dwell Time for Constant Tangential Velocity a(t), Hybrid b(t), and Constant Rotational Rate c(t) Spiral Scans

The total scan time T_s is composed of the constant rotational velocity scan time T_{sr} and the constant tangential velocity scan time T_{sv}. From (14) and (26),

$$T_s = T_{sr} + T_{sv}$$

$$= \frac{\theta_r - \frac{\theta_B}{2}}{\theta_B(1-\Delta)f_s} + \frac{2\pi t_d}{\theta_B^2(1-\Delta)} \int_{\theta_r - \theta_B/2}^{\theta_m - \theta_B/2} \theta(t)\, d\theta(t) . \qquad (33)$$

Making the substitutions similar to equations (27) to (31),

$$T_s = \frac{\theta_r - \frac{\theta_B}{2}}{\theta_B(1-\Delta)f_s} + \frac{\pi t_d}{\theta_B^2(1-\Delta)} \left[(\theta_m - \theta_B/2)^2 - (\theta_r - \theta_B/2)^2\right] . \qquad (34)$$

For any given initial constant rotational velocity f_s and any final dwell time the following boundary condition from (17) must hold

$$t_d = \frac{1}{2\pi f_s} \frac{\theta_B}{\theta_r - \frac{\theta_B}{2}} \qquad (35)$$

or

$$\theta_r - \frac{\theta_B}{2} = \frac{\theta_B}{2\pi f_s t_d} . \qquad (36)$$

Therefore, for given f_s and t_d, the half-cone angle reached at switchover is

$$\theta_r = \frac{\theta_B}{2}\left[\frac{1}{\pi f_s t_d} + 1\right] \qquad (37)$$

Substituting (36) into (34)

$$T_s = \frac{1}{4\pi(1-\Delta) f_s^2 t_d} + \frac{\pi t_d (\theta_m - \theta_B/2)^2}{\theta_B^2(1-\Delta)} . \qquad (38)$$

Note that the second term in (38) is the same as (31). Thus, for a given t_d, the hybrid scan requires a total scan time greater than the constant tangential velocity scan by the addition of the first term in (38). If $f_s = \infty$, then (38) becomes the constant tangential velocity result in (31). Alternately, t_d can be found for a given T_s by rewriting (38) as

$$\frac{\pi(\theta_m - \theta_B/2)^2}{\theta_B^2(1-\Delta)} t_d^2 - T_s t_d + \frac{1}{4\pi(1-\Delta) f_s^2} = 0 \qquad (39)$$

and solving this quadratic equation for t_d. It is important to point out that solution of (39) for t_d depends not only on T_s but also on such parameters as θ_m, θ_B, and f_s. Because T_s and θ_m are generally fixed for a given mission, maximizing t_d for a particular value of θ_B can be performed by analyzing the relationship between t_d and f_s. Such analysis reveals the minimum value of f_s yielding an acceptable value of t_d. In the next section, several numerical examples of such an analysis are presented, with particular emphasis on Shuttle Ku-band radar/communication system parameters.

5.0 NUMERICAL EXAMPLES

In the preceding sections, the analytical relationships for various spiral scan strategies were developed. It was shown that the hybrid scan provides an optimum compromise between excessive servo requirements and dwell time. We will now use the analytical results to obtain numerical data pertinent to hybrid spiral scan such as may be employed by Shuttle Ku-band radar/communication equipment during acquisition. Specifically, we will use equation (39) to obtain the values of dwell time t_d as a function of f_s for a set of various mission-dependent parameters T_s and θ_m and for antenna sizes ranging from 20 inches to 36 inches in diameter. Specific mission-dependent parameters for Shuttle Ku-band equipment are:

(1) $\theta_m = \pm 30^\circ$

(2) $\theta_m = \pm 20^\circ$

(3) $\theta_m = \pm 10^\circ$

The first two are characteristic of radar acquisition for a passive target. The third ($\theta_m = 10^\circ$) scan limit parameter is typical of cooperative target acquisition in the radar mode and TDRS acquisition in the communication mode.

Figures 4 through 7 show the solution of (39) for t_d as a function of the scan frequency f_s and the mission-dependent parameters T_s and θ_m. The most significant result evident from these figures is that, at the scan frequency of 2 Hz, the dwell times for all cases considered are near their maxima and exhibit only an insignificant increase at higher values of f_s.

As can be expected, the case of $\theta_m = \pm 30^\circ$ and $T_s = 30$ seconds sets the lower limit on f_s for all dish sizes considered. This is because, in this case, the antennas have to scan the maximum value within the minimum allotted time. Hence, the f_s has to be relatively "fast" to spiral the beams across the scan volume determined by the θ_m limits.

Figure 8 shows the minimum dwell time available from the hybrid scan versus antenna dish shize and the sets of mission-dependent parameters used. Essentially, these "minimum" dwell times are taken at $f_s = 2$ Hz, because above this frequency there is no significant gain in the dwell time. The 2 Hz frequency, however, is within the operating range of typical antenna servo drives.

6.0 CONCLUSIONS

Analytical expressions for various types of spiral scans have been derived and the hybrid scan appears to be most efficient from the standpoint of minimizing servo loop loading required to achieve "best" (i.e., longest) dwell times for any given set of parameters. Specific numerical examples

indicate that a 2 Hz scan frequency should meet the dwell time optimization requirements for most missions performed by Shuttle Ku-band radar/communication equipment.

Dish Diameter = 20"

3 dB Beamwidth, θ_B = 2.76° (one way)

Beam Overlap, Δ = 0.3

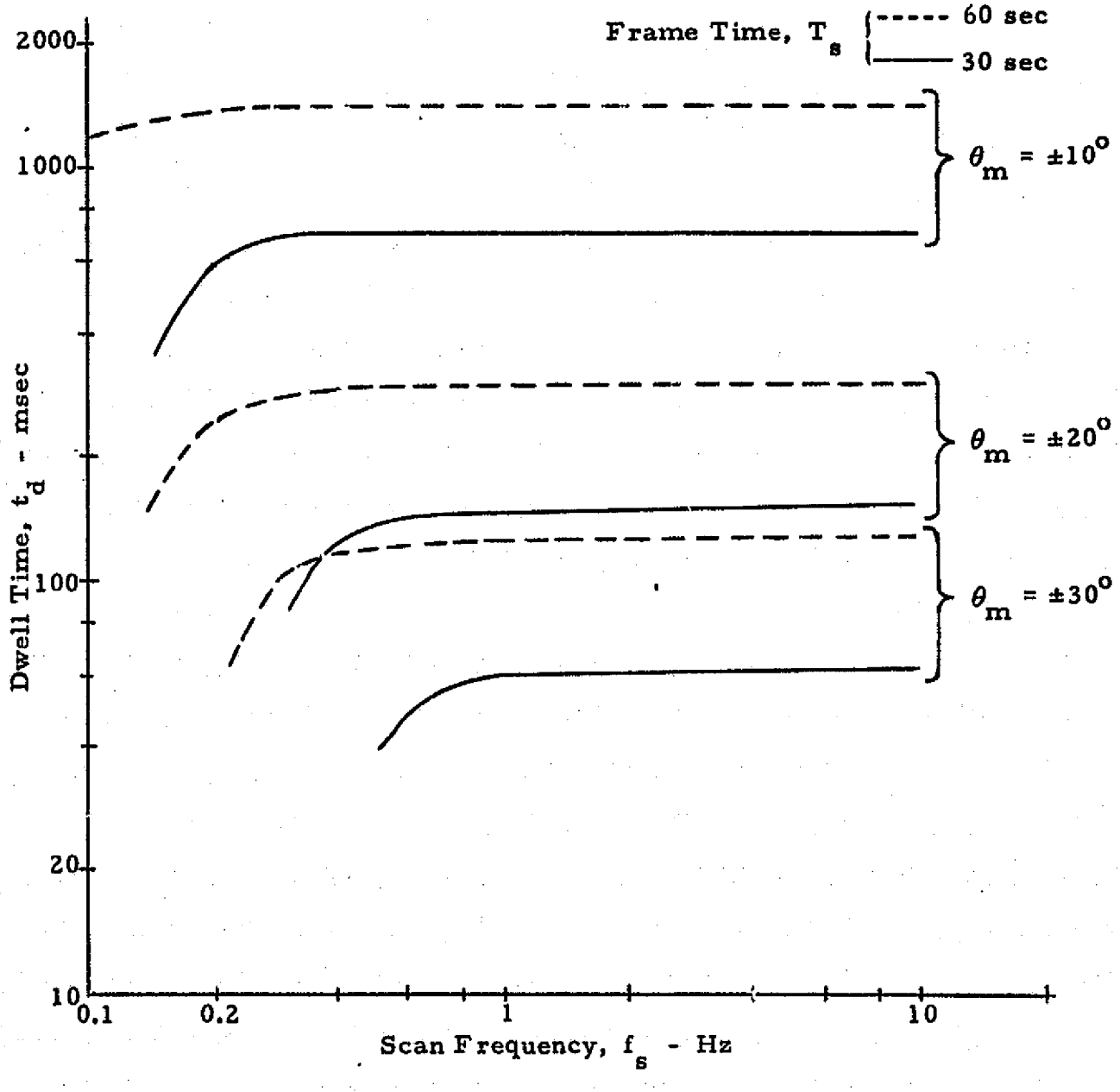


Figure 4. Dwell Time t_d Versus Scan Frequency f_s and Frame Time T_s for 20" Diameter Antenna Dish

Dish Diameter = 26"

3 dB Beamwidth, θ_B = 2.12° (one way)

Beam Overlap, Λ = 0.30

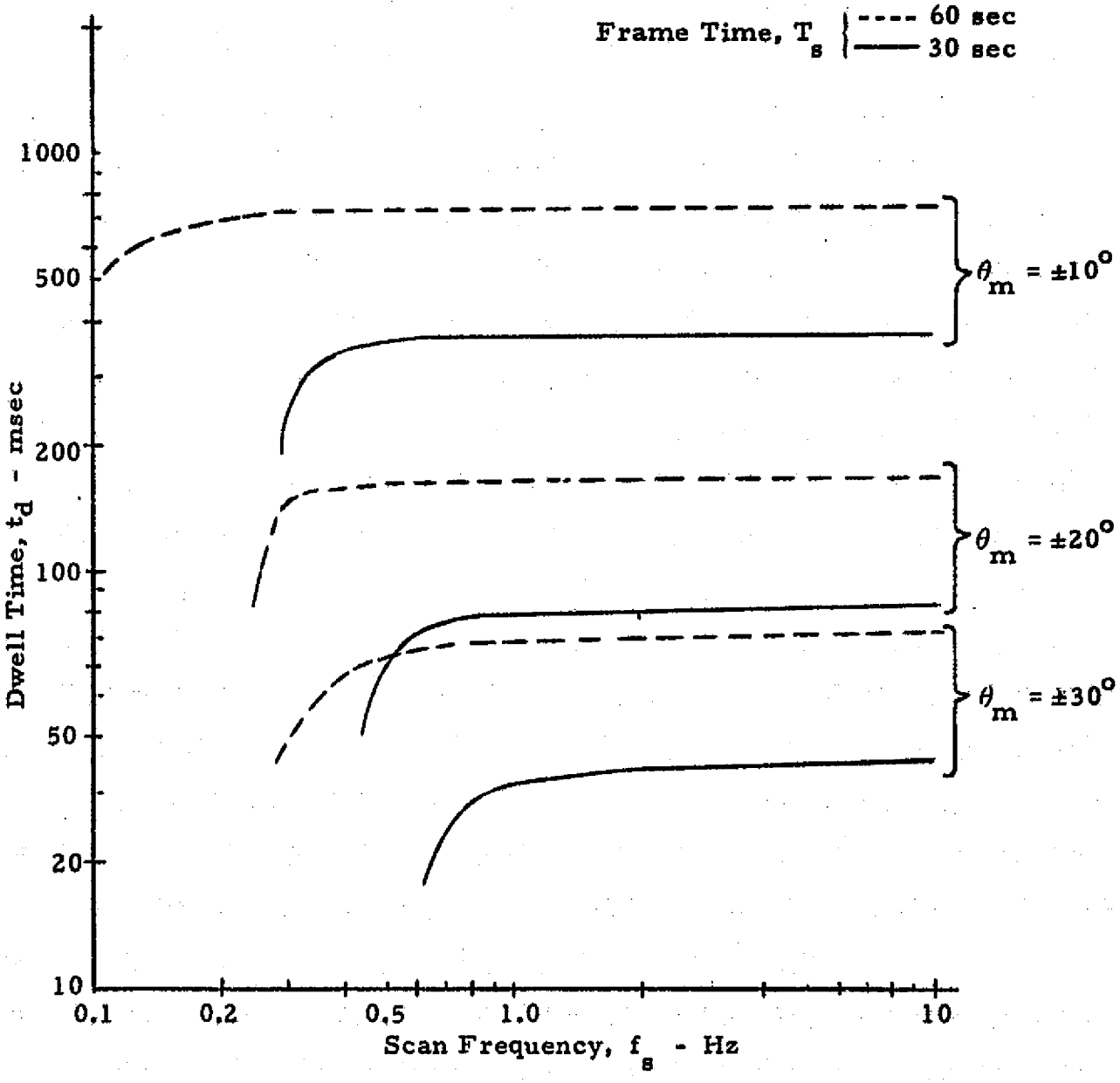


Figure 5. Dwell Time t_d Versus Scan Frequency f_s and Frame Time T_s for 26" Diameter Antenna Dish

Dish Diameter = 32"

3 dB Beamwidth, θ_B = 1.73° (one way)

Beam Overlap, Δ = 0.30

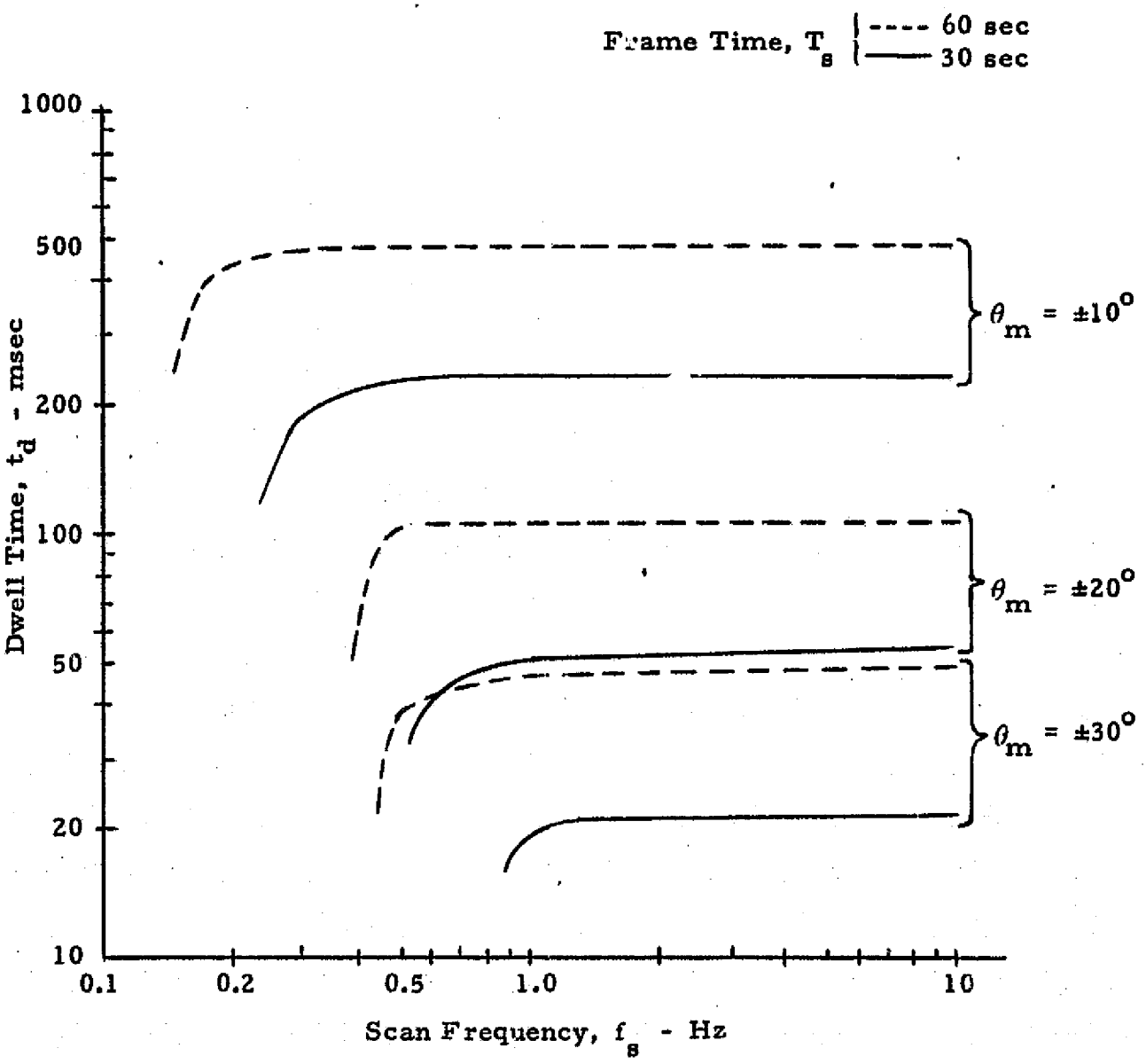


Figure 6. Dwell Time t_d Versus Scan Frequency f_s and Frame Time T_s for 32" Diameter Antenna Dish

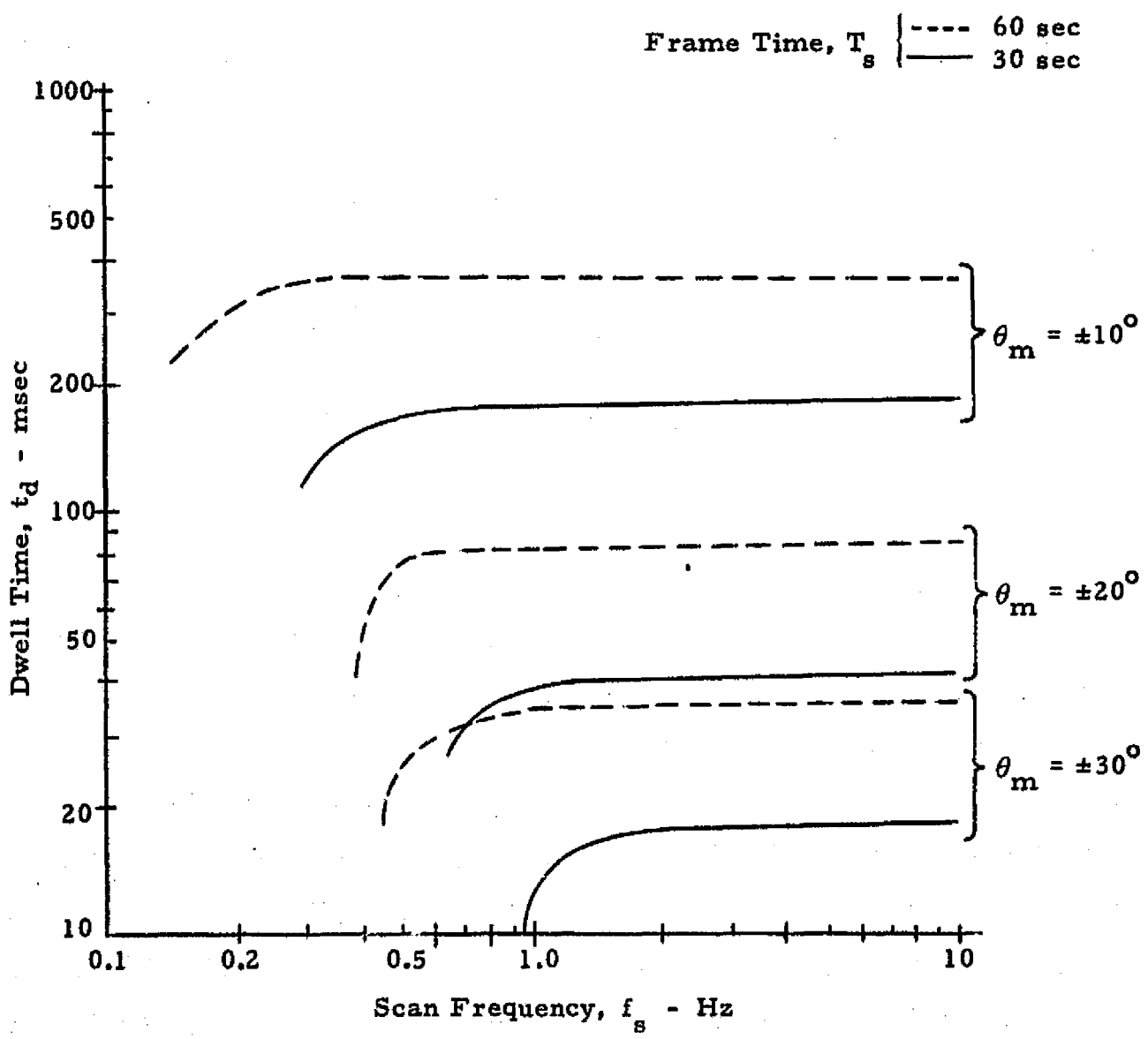


Figure 7. Dwell Time t_d Versus Scan Frequency f_s and Frame Time T_s for 36" Diameter Antenna Dish

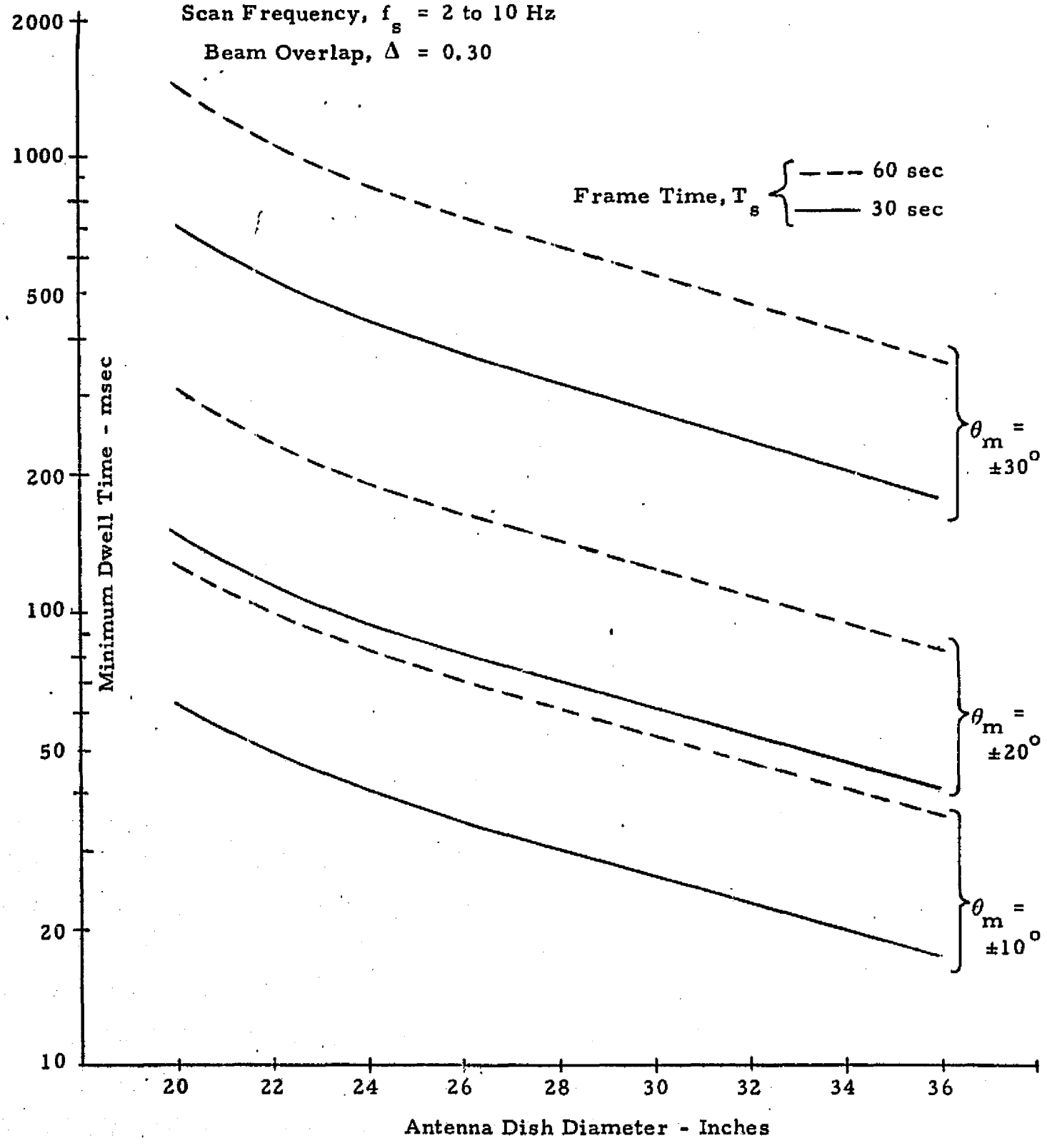


Figure 8. Minimum Dwell Time Versus Dish Diameter, Frame Time, and Maximum Scnn Angle for f_s Range of 2 to 10 Hz

APPENDIX J

FALSE LOCK PERFORMANCE OF COSTAS RECEIVERS

Marvin K. Simon

APPENDIX J

FALSE LOCK PERFORMANCE OF COSTAS RECEIVERS

by Marvin K. Simon

In this appendix, we first demonstrate that a Costas loop can exhibit a false lock condition in the sense that the loop can lock up on a data sideband. In particular, for the case of random data, the loop can theoretically lock at a frequency which is any integer multiple of half the data symbol rate. An expression for the dc output of the lock detector is then derived in terms of the Fourier transform of the symbol pulse shape, the squared magnitude of the lowpass arm filter transfer function and the loop phase error. This expression is then evaluated at the true lock and false lock conditions and numerical results are given for the case of Manchester coded data and a single pole (RC) arm filter.

Consider the Costas loop and associated lock detector illustrated in Figure J-1. The input signal $s[t, \theta(t)]$ is a bi-phase modulated carrier of the form

$$s[t,\theta(t)] = \sqrt{2S}\, m(t) \sin [\omega_0 t + \theta(t)] \tag{1}$$

where S is the average signal power, m(t) is a binary modulation (a ± 1 digital waveform), ω_0 is the radian carrier frequency, and $\theta(t) \triangleq \theta_0 + \Omega_0 t$ is the received carrier phase. The total received signal x(t) is then

$$x(t) = s[t, \theta(t)] + n_i(t) \tag{2}$$

where $n_i(t)$ is the additive channel noise which can be expressed in the form of a narrowband process about the actual frequency of the observed data, viz.,

$$n_i(t) = \sqrt{2} \left\{ N_c(t) \cos [\omega_0 t + \theta(t)] - N_s(t) \sin [\omega_0 t + \theta(t)] \right\} \tag{3}$$

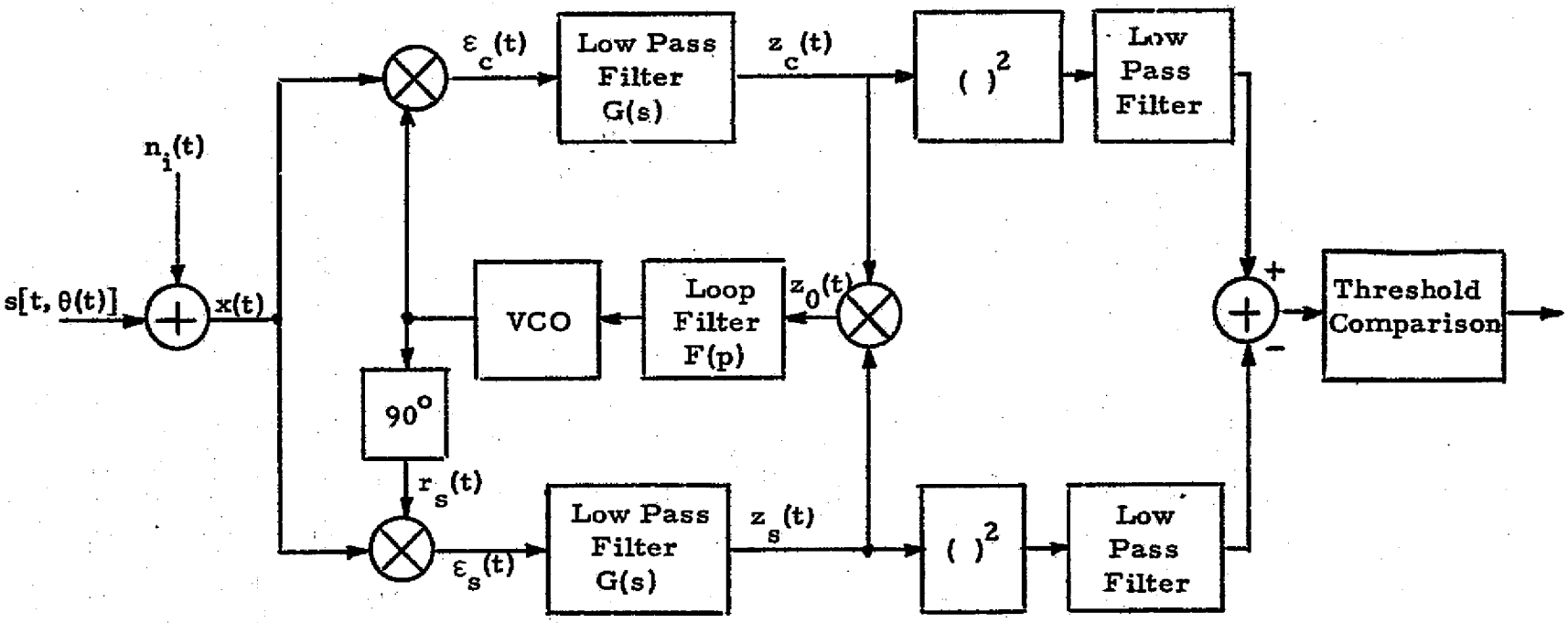


Figure J-1. Costas Loop and Associated Lock Detector

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

For the true lock case, the in-phase and quadrature reference signals $r_s(t)$ and $r_c(t)$, respectively, would be given by

$$r_s(t) = \sqrt{2}\, K_1 \sin[\omega_0 t + \hat{\theta}(t)]$$

$$r_c(t) = \sqrt{2}\, K_1 \cos[\omega_0 t + \hat{\theta}(t)] \tag{4}$$

where $\hat{\theta}(t)$ is the VCO's estimate of $\theta(t)$ and K_1^2 is the reference signal power. In the false lock mode, one would have

$$r_s(t) = \sqrt{2}\, K_1 \sin[(\omega_0 - \omega_f)t + \hat{\theta}(t)]$$

$$r_c(t) = \sqrt{2}\, K_1 \cos[(\omega_0 - \omega_f)t + \hat{\theta}(t)] \tag{5}$$

where ω_f is the false lock radian frequency as yet to be determined.

Denoting the in-phase and quadrature phase detector (multiplier) gains by K_m, then using (3) and (5), the output $\epsilon_c(t)$ of the quadrature phase detector is (ignoring second-harmonic terms)

$$\begin{aligned}\epsilon_c(t) &= K_m\, x(t)\, \sqrt{2}\, K_1 \cos[(\omega_0 - \omega_f)t + \hat{\theta}(t)] \\ &= K_1 K_m [\sqrt{S}\, m(t) - N_s(t)] \sin[\omega_f t + \phi(t)] \\ &\quad + K_1 K_m\, N_c(t) \cos[\omega_f t + \phi(t)]\end{aligned} \tag{6}$$

while the in-phase arm phase detector output is

$$\begin{aligned}\epsilon_s(t) &= K_m\, x(t)\, \sqrt{2}\, K_1 \sin[(\omega_0 - \omega_f)t + \hat{\theta}(t)] \\ &= K_1 K_m [\sqrt{S}\, m(t) - N_s(t)] \cos[\omega_f t + \phi(t)] \\ &\quad - K_1 K_m\, N_c(t) \sin[\omega_f t + \phi(t)]\end{aligned} \tag{7}$$

where $\phi(t) \triangleq \theta(t) - \hat{\theta}(t)$ is the loop phase error. After lowpass filtering with the in-phase and quadrature phase arm filters, these same signals become, respectively,*

$$\overbrace{\hat{m}_s(t;\omega_f,\phi)}^{\longleftarrow\longrightarrow} \qquad \overbrace{\hat{N}_{ss}(t;\omega_f,\phi)}^{\longleftarrow\longrightarrow}$$

$$z_c(t) = K_1K_m\left\{\sqrt{S}\,G(p)\left[m(t)\sin[\omega_f t+\phi(t)]\right] - G(p)\left[N_s(t)\sin[\omega_f t+\phi(t)]\right]\right\}$$

$$\overbrace{\hat{N}_{cc}(t;\omega_f,\phi)}^{\longleftarrow\longrightarrow}$$

$$+ K_1K_m\left\{G(p)\left[N_c(t)\cos[\omega_f t+\phi(t)]\right]\right\}$$

$$\overbrace{\hat{m}_c(t;\omega_f,\phi)}^{\longleftarrow\longrightarrow} \qquad \overbrace{\hat{N}_{sc}(t;\omega_f,\phi)}^{\longleftarrow\longrightarrow}$$

$$z_s(t) = K_1K_m\left\{\sqrt{S}\,G(p)\left[m(t)\cos[\omega_f t+\phi(t)]\right] - G(p)\left[N_s(t)\cos[\omega_f t+\phi(t)]\right]\right\}$$

$$\overbrace{\hat{N}_{cs}(t;\omega_f,\phi)}^{\longleftarrow\longrightarrow}$$

$$- K_1K_m\left\{G(p)\left[N_c(t)\sin[\omega_f t+\phi(t)]\right]\right\} \qquad (8)$$

Here, the signals $\hat{m}_s(t;\omega_f,\phi)$ and $\hat{m}_c(t;\omega_f,\phi)$ are the in-phase and quadrature signals emerging from passage of the data modulated false lock sinusoids, $m(t)\sin[\omega_f+\phi(t)]$ and $m(t)\cos[\omega_f t+\phi(t)]$, respectively, through the lowpass arm filters $G(s)$, and likewise, $\hat{N}_{ss}(t;\omega_f,\phi)$, $\hat{N}_{cc}(t;\omega_f,\phi)$, $\hat{N}_{sc}(t;\omega_f,\phi)$, and $\hat{N}_{cs}(t;\omega_f,\phi)$ are the equivalently filtered versions of the noise components.

Squaring $z_c(t)$ and $z_s(t)$ and differencing these quantities produces a process whose statistical mean (over the data sequence) and time average gives the dc voltage at the lock detector output. From (8), the normalized signal component of this dc voltage is**

---

*In what follows, we shall write differential equations in compact form by introducing the Heaviside operator $p \triangleq d/dt$.

**For convenience, we drop in our notation herein, the dependence of ϕ on time.

$$V_{dc} \triangleq \lim_{T_0 \to \infty} \int_{-T_0/2}^{T_0/2} \left[\overline{\hat{m}_c^2(t; \omega_f, \phi)} - \overline{\hat{m}_s^2(t; \omega_f, \phi)} \right] dt \tag{9}$$

where the overbar denotes statistical mean. It now remains to characterize V_{dc} in terms of the statistical characterization of the modulation m(t) and the filter transfer function characteristic $G(j\omega)$.

The digital modulation m(t) may be represented as

$$m(t) = \sum_{n=-\infty}^{\infty} a_n \, p(t - nT) \tag{10}$$

where $a_n = \pm 1$ is the data symbol sequence and p(t) is the symbol pulse shape, e.g., for Manchester coding,

$$p(t) = \begin{cases} -1 ; & -\frac{T}{2} \le t \le 0 \\ 1 ; & 0 \le t \le \frac{T}{2} \end{cases} \tag{11}$$

Substituting (10) into the definitions of $\hat{m}_c(t; \omega_f, \phi)$ and $\hat{m}_s(t; \omega_f, \phi)$ given in (8) results, after considerable manipulation, in

$$\begin{aligned} \hat{m}_s(t; \omega_f, \phi) &\triangleq G(p)[m(t) \sin(\omega_f t + \phi)] \\ &= \sum_{n=-\infty}^{\infty} a_n \mathcal{F}^{-1} \left\{ G(j\omega) \mathcal{F} [p(t - nT) \sin(\omega_f t + \phi)] \right\} \\ &= \sum_{n=-\infty}^{\infty} a_n \hat{p}_1(t-nT) \frac{e^{j(\omega_f t + \phi)}}{2j} \\ &\quad - \sum_{n=-\infty}^{\infty} a_n \hat{p}_2(t-nT) \frac{e^{-j(\omega_f t + \phi)}}{2j} \end{aligned} \tag{12}$$

where $\mathcal{F}$ denotes Fourier transform with $\mathcal{F}^{-1}$ its inverse and

$$\hat{p}_1(t) = \int_{-\infty}^{\infty} G[j(\omega + \omega_f)]\, P(j\omega)\, e^{j\omega t}\, \frac{d\omega}{2\pi}$$

$$\hat{p}_2(t) = \int_{-\infty}^{\infty} G[j(\omega - \omega_f)]\, P(j\omega)\, e^{j\omega t}\, \frac{d\omega}{2\pi} \qquad (13)$$

Similarly,

$$\hat{m}_c(t;\omega_f,\phi) = G(p)[m(t)\cos(\omega_f t + \phi)]$$

$$= \sum_{n=-\infty}^{\infty} a_n \hat{p}_1(t-nT)\, \frac{e^{j(\omega_f t+\phi)}}{2}$$

$$+ \sum_{n=-\infty}^{\infty} a_n \hat{p}_2(t-nT)\, \frac{e^{-j(\omega_f t+\phi)}}{2} \qquad (14)$$

Squaring (12) and (14) and taking the difference of their statistical means gives

$$\overline{m_c^2(t;\omega_f,\phi)} - \overline{m_s^2(t;\omega_f,\phi)} = \frac{e^{j(2\omega_f t+2\phi)}}{2} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \overline{a_n a_m}\, \hat{p}_1(t-nT)\, \hat{p}_1(t-mT)$$

$$+ \frac{e^{-j(2\omega_f t+2\phi)}}{2} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \overline{a_n a_m}\, \hat{p}_2(t-nT)\, \hat{p}_2(t-mT) \qquad (15)$$

Finally, taking the time average, we get

$$V_{dc} = \frac{1}{2} \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \overline{a_n a_m} \left\{ \lim_{T_0\to\infty} \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} e^{2j(\omega_f t+\phi)}\, \hat{p}_1(t-nT)\, \hat{p}_1(t-mT)\, dt \right.$$

$$\left. + \lim_{T_0\to\infty} \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} e^{-2j(\omega_f t+\phi)}\, \hat{p}_2(t-nT)\, \hat{p}_2(t-mT)\, dt \right\} \qquad (16)$$

where

$$\lim_{T_0\to\infty} \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} e^{2j(\omega_f t+\phi)} \hat{p}_1(t-nT)\,\hat{p}_1(t-mT)\,dt$$

$$= e^{j2\phi} \lim_{T_0\to\infty} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G[j(\omega+\omega_f)]\,P(j\omega)\,G[j(-\omega'+\omega_f)]$$

$$\times\, P(-j\omega')\,\frac{\sin[(\omega-\omega'+2\omega_f)(T_0/2)]}{(\omega-\omega'+2\omega_f)(T_0/2)}\,e^{-j(\omega n-\omega' m)T}\,\frac{d\omega}{2\pi}\,\frac{d\omega'}{2\pi} \qquad (17)$$

and

$$\lim_{T_0\to\infty} \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} e^{-2j(\omega_f t+\phi)} \hat{p}_2(t-nT)\,\hat{p}_2(t-mT)\,dt$$

$$= e^{-j2\phi} \lim_{T_0\to\infty} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G[j(\omega-\omega_f)]\,P(j\omega)\,G[j(-\omega'-\omega_f)]$$

$$\times\, P(-j\omega')\,\frac{\sin[(\omega-\omega'-2\omega_f)(T_0/2)]}{(\omega-\omega'-2\omega_f)(T_0/2)}\,e^{-j(\omega n-\omega' m)T}\,\frac{d\omega}{2\pi}\,\frac{d\omega'}{2\pi}$$

In order to simplify (16) any further, we must consider a specific statistical characterization of the data sequence $\{a_n\}$. For the case of random data, i.e.,

$$\overline{a_n a_m} = \delta_{mn} = \begin{cases} 1\,;\ m=n \\ 0\,;\ m\neq n \end{cases} \qquad (19)$$

The double sum of (16) reduces to a single sum. Recognizing further that

$$\sum_{n=-\infty}^{\infty} e^{-j(\omega-\omega')nT} = \frac{2\pi}{T} \sum_{k=-\infty}^{\infty} \left(\omega-\omega'-\frac{2\pi k}{T}\right) \qquad (20)$$

then,

$$V_{dc} = \lim_{T_0 \to \infty} \frac{1}{2T} \left\{ e^{j2\phi} \sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} G[j(\omega' + \omega_f + \frac{2\pi k}{T})]\, P[j(\omega' + \frac{2\pi k}{T})] \right.$$

$$\times\, G[j(-\omega' + \omega_f)]\, P(-j\omega')\, \frac{\sin[(2\pi\frac{k}{T} + 2\omega_f)(T_0/2)]}{(2\pi\frac{k}{T} + 2\omega_f)(T_0/2)}\, \frac{d\omega'}{2\pi}$$

$$+\, e^{-j2\phi} \sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} G[j(\omega' - \omega_f + \frac{2\pi k}{T})]\, P[j(\omega' + \frac{2\pi k}{T})]$$

$$\left. \times\, G[j(-\omega' - \omega_f)]\, P(-j\omega')\, \frac{\sin[(2\pi\frac{k}{T} - 2\omega_f)(T_0/2)]}{(2\pi\frac{k}{T} - 2\omega_f)(T_0/2)}\, \frac{d\omega'}{2\pi} \right\} \qquad (21)$$

Now suppose that $2\omega_f = \frac{2\pi k_0}{T}$ where $k_0 = 0, 1, 2, \ldots$ is a particular value of k. Then, the value of V_{dc}, namely V_{k_0}, at this frequency is given by

$$V_{k_0} = \frac{1}{2T} \left[e^{j2\phi} \int_{-\infty}^{\infty} \left| G\left[j\left(\omega - \frac{\pi k_0}{T}\right)\right] \right|^2 P\left[j\left(\omega - \frac{2\pi k_0}{T}\right)\right] P(-j\omega)\, \frac{d\omega}{2\pi} \right.$$

$$\left. +\, e^{-j2\phi} \int_{-\infty}^{\infty} \left| G\left[j\left(\omega + \frac{\pi k_0}{T}\right)\right] \right|^2 P\left[j\left(\omega + \frac{2\pi k_0}{T}\right)\right] P(-j\omega)\, \frac{d\omega}{2\pi} \right] \qquad (22)$$

or equivalently

$$\boxed{V_{k_0} = \frac{1}{T}\, \mathrm{Re} \left\{ e^{j2\phi} \int_{-\infty}^{\infty} |G(j\omega)|^2\, P\left[j\left(\omega - \frac{\pi k_0}{T}\right)\right]\, P\left[-j\left(\omega + \frac{\pi k_0}{T}\right)\right] \frac{d\omega}{2\pi} \right\}}$$

$$k_0 = 1, 2, 3, \ldots \qquad (23)$$

where $\mathrm{Re}\{\ \}$ denotes the real part of what is contained within the braces. Note that, for the case of true lock, i.e., $\omega_f = 0$, we get

$$V_0 = \cos 2\phi \left[\frac{1}{T} \int_{-\infty}^{\infty} |G(j\omega)|^2\, |P(j\omega)|^2\, \frac{d\omega}{2\pi} \right] = \cos 2\phi \int_{-\infty}^{\infty} S_m(\omega)\, |G(j\omega)|^2\, \frac{d\omega}{2\pi}$$

(24)

where $S_m(\omega) = \frac{1}{T}|P(j\omega)|^2$ is the power spectral density of the modulation m(t). For any of the false lock frequencies $f_f = \omega_f/2\pi = k_0/2T$, the corresponding lock detector output (signal component) is given by (23).

A particular example will now be given to illustrate the theory just developed. Consider the case where m(t) is a Manchester coded data modulation and G(s) is a single pole RC filter for which

$$|G(j\omega)|^2 = \frac{1}{1 + (\omega/\omega_c)^2} \tag{25}$$

where $\omega_c = 2\pi f_c$ is the radian 3 dB cutoff frequency of the filter and is related to its two-sided noise bandwidth B_i by

$$\omega_c = 2B_i . \tag{26}$$

For the Manchester pulse described by (11), its Fourier transform is given by

$$P(j\omega) = jT\left[\frac{\sin^2 \frac{\omega T}{4}}{\frac{\omega T}{4}}\right] \tag{27}$$

After considerable algebraic manipulation and integral evaluations, we arrive at the following results

$$V_0 = \cos 2\phi\left\{1 - \frac{1}{2\pi f_c T}\left[3 - 4e^{-\pi f_c T} + e^{-2\pi f_c T}\right]\right\} \tag{28}$$

$$V_k = -\cos 2\phi\left\{\left[\frac{1}{1 + \left(\frac{k}{2f_c T}\right)^2}\right]\left[\frac{1 + e^{-2\pi f_c T}}{2\pi f_c T}\right]\right\} ; \quad k = 1, 3, 5, \ldots \tag{29}$$

$$V_k = -\cos 2\phi\left\{\left[\frac{1}{1 + \left(\frac{k}{2f_c T}\right)^2}\right]\left[\frac{3 + 4e^{-\pi f_c T} + e^{-2\pi f_c T}}{2\pi f_c T}\right]\right\} ;$$

$$k = 2, 6, 10, \ldots \tag{30}$$

$$V_k = -\cos 2\phi \left\{ \left[\frac{1}{1 + \left(\frac{k}{2 f_c T} \right)^2} \right] \left[\frac{3 - 4e^{-\pi f_c T} + e^{-2\pi f_c T}}{2\pi f_c T} \right] \right\}$$

$$k = 4, 8, 12, \ldots \qquad (31)$$

At this point, it should be mentioned that, if one was to consider the error signal into the loop filter F(p), i.e.,

$$z_0(t) = z_c(t)\, z_s(t) \qquad (32)$$

where $z_c(t)$ and $z_s(t)$ are given in (8), then the dc value of the normalized signal component in $z_0(t)$, namely,

$$z_{dc} \triangleq \lim_{T_0 \to \infty} \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} \overline{\hat{m}_s(t;\omega_f,\phi)\, \hat{m}_c(t;\omega_f,\phi)}\, dt \qquad (33)$$

would be given by (28) through (31) with $\cos 2\phi$ replaced by $(\sin 2\phi)/2$. In view of this, we may reach the following conclusion. Since, for any $f_c T$, the braced quantity in V_0 of (28) is positive, then as is well known, true lock corresponds to the lock points $\phi = 0, \pm\pi, \pm 2\pi, \pm 3\pi, \ldots$. Furthermore, since the remaining braced quantities in (29) through (31) are also always positive for any $f_c T$, then in the false lock mode, we must have $\cos 2\phi = -1$ or $\phi = \pm\pi/2, \pm 3\pi/2, \pm 5\pi/2, \ldots$. Thus, the lock points for any false lock condition are interleaved midway between those for the true lock state.

Figure J-2 is a plot of V_k (in dB); $k = 0, 1, 2, 3, 4$, versus $f_c T$ for the example given above. We note that, by continuing to widen the arm filter bandwidth beyond the optimum value which yields minimum squaring loss ($f_c T \cong 1.4$), the separation between V_0 and the false lock values continues to increase. This property is desirable from the standpoint of distinguishing false-lock from true-lock. However, widening the arm filter bandwidth also allows more noise to pass through, thus degrading the

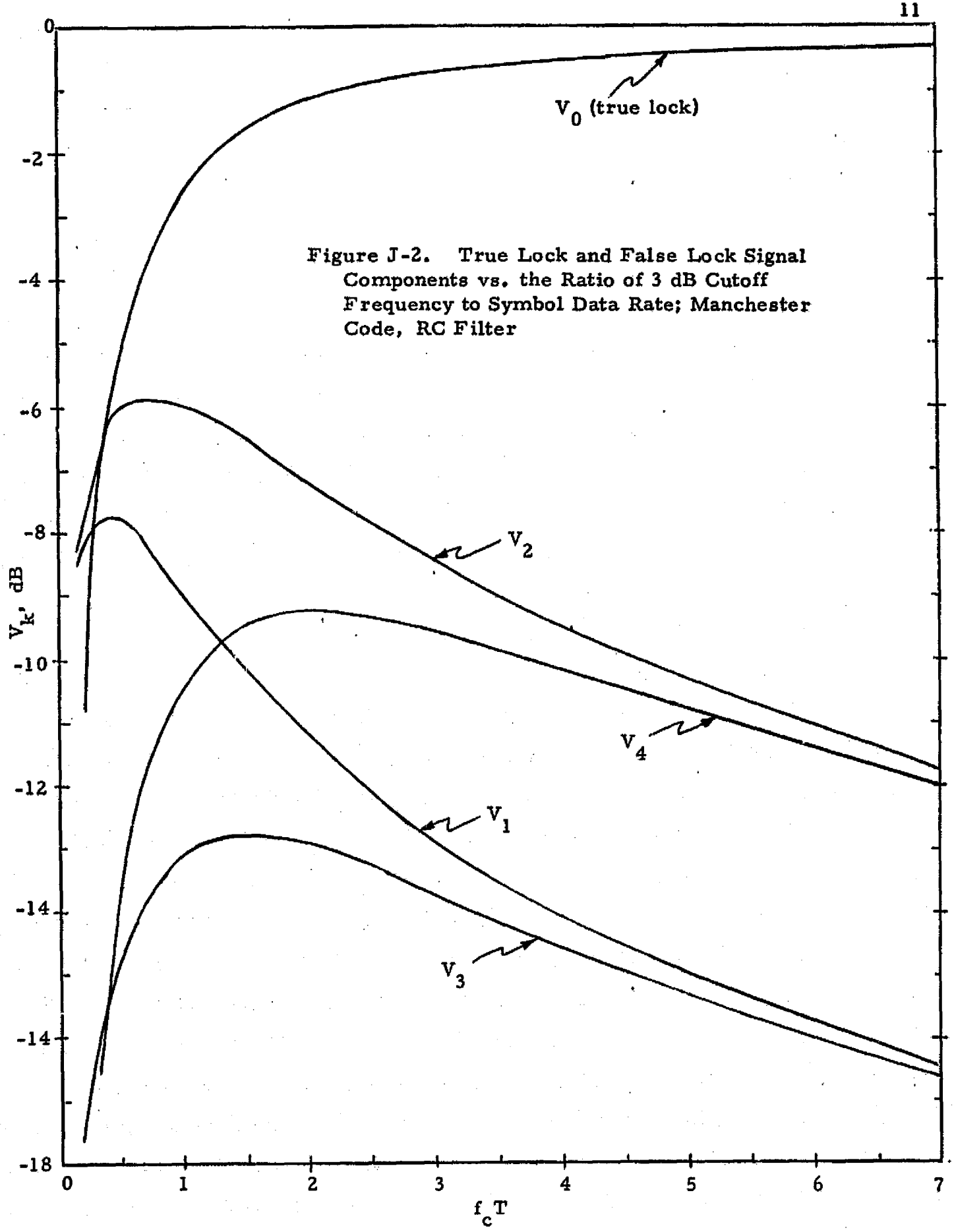


Figure J-2. True Lock and False Lock Signal Components vs. the Ratio of 3 dB Cutoff Frequency to Symbol Data Rate; Manchester Code, RC Filter

signal-to-noise ratio at the input to the threshold detector. Thus, in conclusion, widening the arm filter bandwidth is potentially a solution to the false lock detection problem up to the point where the signal-to-noise ratio which distinguishes between true lock and out-of-lock (as opposed to false lock) becomes too low for a specified level of lock detection performance.