IC

Research Report

No. 80-1

BAND EFFICIENT DIGITAL ANGLE MODULATION SIGNALS

PREPARED FOR

THE DEPARTMENT OF COMMUNICATIONS
UNDER DSS CONTRACT No. OSU79-00072

BY

P.H. WITTKE

G.S. DESHPANDE

Queen's University at Kingston
Department of Electrical Engineering

BAND EFFICIENT DIGITAL ANGLE MODULATION SIGNALS

Final Report

Prepared for

The Department of Communications

Under DSS Contract No. OSU79-00072

by

P.H. Wittke G.S. Deshpande

Department of Electrical Engineering

Queen's University

Kingston, Ontario, Canada

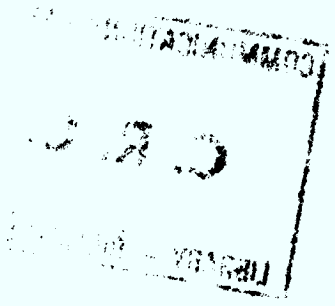
Scientific Authority

W.L. Nowland J.L. Pearce

Communications Research Centre

Ottawa

February 1980



ABSTRACT

Constant envelope modulation signals that are spectrally efficient and possess good error performance capability are desirable in communication systems. Previous experimental and simulation results have shown correlative encoded frequency shift keying to be attractive from a spectral occupancy point of view. In this report, a straightforward method of obtaining an explicit expression for the spectrum of correlative encoded FM signal is given, using Rowe and Prabhu's method. An exhaustive investigation of the spectrum for modulation by a second order polynomial and rectangular pulse shaping has been carried out and the values of the encoding coefficients, for which the spectrum is compact, are shown. Results are given also for raised-cosine pulse shaping. An investigation of the error performance for these modulations is carried out. The results indicate that the encoding polynomial $(1+2D+D^2)/4$ with rectangular pulse shaping and modulation index, η, equal to 0.5π is spectrally 47% more efficient than MSK (Minimum Shift Keying) at an expense of 1.1 dB in SNR. Other correlative encodings are shown that are spectrally somewhat less efficient than the above polynomial but promise better error performance than MSK.

The problem of designing the baseband pulse shapes in MSK-type signals to minimize the fraction of out-of-band power is considered. The optimum pulse shapes and their spectral properties are presented for a range of channel bandwidths up to three times the bit rate. The optimization yields pulse shapes which attain a lower out-of-band power than the well-known MSK-type modulations, such as, MSK, OKQPSK (Offset Keyed Quadrature Phase Shift Keying) and SFSK (Sinusoidal Frequency Shift Keying). The improvement is not dramatic. However, the optima obtained in this report provide a bound on what is achievable and a useful basis for comparison.

MSK-type signals are a subclass of angle modulated signals, for which the modulation index, η, is 0.5π. The more general problem of shaping the baseband pulse with an arbitrary modulation index is treated also. The spectrum of the angle modulated signal is approximated by the power spectrum of a time-limited waveform. The time-limited waveform is then expressed in terms of an infinite series of spheroidal wave functions and is shaped to minimize the fraction of out-of-band power. Finally, the baseband pulse corresponding to the optimized time-limited waveform is obtained. To demonstrate the technique, results are given for modulation indices $\eta = 0.5\pi$, 0.6π, 0.7π and 0.8π.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

The increasing demand for communications has put the allocation of bandwidth in radio systems at a premium. The problem can be alleviated by the following techniques [1,2]:

i) New allocations at higher frequencies

ii) Frequency reuse techniques

iii) Data compression by source encoding

iv) Spectrally efficient modulation.

The last technique as applied to data transmission is studied in this report. The primary objective of spectrally efficient data transmission is to transmit data at a specified rate and error performance in as small a channel bandwidth as possible. The investigation is further restricted to constant envelope signals. Constant envelope signals are desirable in communication systems due to their immunity to fading and non-linear distortion.

Constant envelope signals can be generated by modulating the angle of a carrier. In particular, a digital frequency modulated signal, $x(t)$, can be written as

$$x(t) = \cos\left(2\pi f_c t + \int_{-\infty}^{t} \sum_{k=-\infty}^{\infty} a_k h(u-kT)\,du\right) \qquad (1.1)$$

In the above equation, f_c is the carrier frequency and h(t) is the baseband pulse shape. The sequence of random variables a_k are the symbols to be transmitted. They are drawn from a finite alphabet $\{a_i\}$ and one is transmitted every T seconds. The modulation index in digital FM is linearly proportional to the magnitude of the baseband pulse h(t). In conventional FSK (Frequency Shift Keying), the modulation index h is defined as $h = 2f_dT$, where f_d is the peak instantaneous frequency deviation. The baseband pulse in FSK is rectangular and therefore the instantaneous frequency is constant over the bit interval. With baseband pulse shaping the instantaneous frequency varies during the bit interval. In this investigation, the modulation index, η, will be defined as the maximum phase accumulation during a bit interval and therefore is given by

$$\eta = \max_i |a_i| \sum_{k=-\infty}^{\infty} \left| \int_{kT}^{(k+1)T} h(t)dt \right| \qquad (1.2)$$

The maximum in the above equation is over the symbol alphabet $\{a_i\}$. In conventional binary FSK, h(t) is a rectangular pulse of height $2\pi f_d$ extending over one bit interval and the random variable a_k can assume a value of ± 1. For this special case, the modulation index η is the same as $h\pi$.

The signal x(t) can be written also as

$$x(t) = \cos(2\pi f_c t + \sum_{k=-\infty}^{\infty} a_k g(t-kT)) \qquad (1.3)$$

where

$$g(t) = \int_{-\infty}^{t} h(u)du \tag{1.4}$$

The class of baseband pulses considered in this report include functions g(t) with step changes. A step change in g(t) corresponds to an unbounded instantaneous frequency which is not considered in conventional digital FM. Therefore, when g(t) is allowed step changes, the modulated signal, x(t), is referred to as angle modulated rather than frequency modulated.

Channel bandwidth can be defined as the span of frequencies which contains a specified fraction of the total modulated signal power. The fraction of power in the band $[f_c-W, f_c+W]$ is given by

$$P_{in} = \frac{\int_{f_c-W}^{f_c+W} P_x(f)df}{\int_{0}^{\infty} P_x(f)df} \tag{1.5}$$

and the fraction of out-of-band power is given by

$$P_{out} = 1 - P_{in} \tag{1.6}$$

where, $P_x(f)$ is the power spectral density of x(t). A task in the design of a transmission system would be to choose a modulation index and shape the baseband pulses such that for a specified P_{in}, say $P_{in} = 0.99$, the channel bandwidth is minimized and a specified error performance, P_e, is maintained.

Previous experimental and simulation results [6] have shown that the channel bandwidth, W, can be reduced by correlatively encoding the input data symbols a_k. It will be seen that a correlatively encoded digital FM signal can always be represented by an FM signal in which the input data symbols are independent and the baseband pulse, h(t), extends over more than one bit interval. In this report, an analytical method is developed to obtain the spectrum of correlative encoded FM signals. The spectra of correlative encoded FM are investigated for various baseband encoding polynomials and pulse shapings. The encoding polynomials that yield small channel bandwidths are indicated. Also, an investigation of the error performance of these signals is carried out. Signals that yield a compact spectrum at a small penalty in SNR (Signal to Noise Ratio) are pointed out. When the baseband pulse, h(t), extends over only one bit interval and the input data symbols are independent, the optimum baseband pulse shaping to minimize the channel bandwidth is derived.

1.1 Correlative Encoded FM

Correlative encoding or partial response techniques are used in data transmission systems employing linear modulation to obtain a desirable signal spectrum [3-5]. As well, the spectrum of the modulated signal in digital FM can be manipulated by correlatively encoding the modulating signal. However, since the modulation is nonlinear, a modulating signal with a compact spectrum does not necessarily lead to a modulated signal with desirable spectral properties. The

application of duobinary encoding to data transmission by FM was first introduced by Lender [3] in 1963. Recently Melvin and Middlestead [6] have investigated the power density spectrum of correlative encoded FSK, experimentally and by computer simulation. Their results show correlative encoded FSK modulation to be superior to conventional FSK in terms of spectral conservation. Duobinary is the only correlative encoding for which the FSK spectrum has been obtained theoretically [7,8]. Garrison [9] has developed a computational technique to obtain the power spectral density for correlative encoded FM. The modulating pulses are approximated by a staircase function and the resulting FM spectrum is computed. Experimental and computer simulation studies can be time consuming, and in an investigation of partial response encoded FM, a much larger class of encoding polynomials than in the case of a linear modulation must be considered, for there are no simple properties that eliminate all but a few polynomials from consideration.

A method of obtaining a closed form expression for the spectrum of digital FM with partial response encoding is given in Chapter 2. The spectra of the modulated signals, for several encoding polynomials and pulse shapings, are examined in Chapter 3. An exhaustive investigation of the spectrum for modulation by a second order polynomial has been carried out and the values of the encoding coefficients for which the spectrum is compact are shown. The channel bandwidths required to transmit 99% and 99.9% of the modulated signal power are tabulated.

An upper bound on the performance of correlative encoded digital FM can be given in terms of the minimum Euclidean distance. An exhaustive investigation of the minimum Euclidean distance is carried out, in Chapter 4, for modulation by a second order encoding polynomial. Correlative encoded FM signals that are spectrally efficient at a small penalty in SNR are pointed out in Chapter 7.

1.2 Optimum Pulse Shaping in MSK-Type Signals

MSK-type signals are a subclass of angle modulated signals that have received considerable attention [10-26]. When the baseband pulse, $h(t)$, in an angle modulated signal extends over only one bit interval and the modulation index, η, is equal to 0.5π, the modulated signal is a MSK-type signal. Angle modulation involves a nonlinear transformation of the baseband signal. However, MSK-type signals can be represented by a linear quadrature carrier modulation [11,12] and can be viewed as a degenerate case.

MSK (Minimum Shift Keying)was patented by Doelz and Heald [10] in 1961. MSK is known also as FFSK (Fast Frequency Shift Keying) [12] and FM-PSK (Frequency Modulation Phase Shift Keying) [18]. These modulations have been commended for their desirable spectral properties. Amoroso [17] pointed out that the asymptotic behaviour of MSK can be improved even more by having continuous rate of change of the in-phase and the quadrature envelopes. He suggested a modulation known as SFSK (Sinusoidal Frequency Shift Keying). Gronmeyer and McBride [19] studied MSK and OKQPSK (Offset Keyed Quadrature

Phase Shift Keying) in detail and compared their performances over bandlimited channels. They concluded that OKQPSK outperforms MSK only when the channel bandwidth is highly restricted.

A number of authors have considered the effect of pulse shaping on the spectra of MSK-type signals. Rabzel and Pasupathy [23] proposed a family of M^{th} order shaping functions which lead to a spectral rolloff as $|f|^{-(4M+4)}$. For large values of M such signals have excellent rolloff characteristics, but they cause severe adjacent channel interference when a small channel bandwidth is specified. Kalet and White [21,22] examined the interchannel interference for MSK-type modulation. Eaves and Wheatley [25] have proposed a family of wave shapes and optimized the parameters by a computer search procedure to minimize the interchannel interference.

The problem of baseband pulse shaping to minimize the fraction of out-of-band power, P_{out}, for a given channel bandwidth is treated in detail in Chapter 5. Rabow [13] has considered this problem with an assumption of symmetry in the baseband pulses and has given the optimum parameters for a particular value of the channel bandwidth. Boutin et al [24] have recently reconsidered this problem pointing out Rabow's work. Prabhu [18] has obtained a lower bound, that is not necessarily achievable, on P_{out} for a given channel bandwidth.

In Chapter 5, baseband pulse shapes are derived that minimize the fraction of out-of-band power, P_{out}. Results are presented for a range of channel bandwidths of practical

interest. The spectra of the MSK-type signals obtained through optimization are compared with the spectra of the well known schemes, such as MSK, OKQPSK and SFSK.

1.3 Optimum Pulse Shaping in Angle Modulated Signals

As mentioned previously, MSK-type signals may be considered as a special linear case of digital FM. In Chapter 6, the more general problem of shaping the baseband pulse to minimize the fraction of out-of-band power for digital FM with an arbitrary modulation index is considered. The modulated signal is referred to as angle modulated because baseband pulses which cause unbounded instantaneous frequency are permitted. The baseband pulse, $h(t)$, is allowed to extend over only one bit period and for simplicity only baseband pulses which produce symmetrical phase variation are considered.

Methods of obtaining the spectra of digital FM signals and band occupancy of FSK signals were investigated by several authors [27-32] in the 1960's. More recently, Rowe and Prabhu [33] have presented a method for determining the power spectrum of digital FM signals. The result is in terms of products of matrices, which will prove most convenient for the purpose of this thesis. Using Rowe and Prabhu's method it is shown in Chapter 6, that the spectrum of an angle modulated signal can be approximated by the power spectrum of a time-limited waveform. The time-limited waveform is then expressed in terms of an infinite series of spheroidal wave functions. The coefficients of the spheroidal wave functions

are optimized to minimize the out-of-band power, P_{out}. Finally the baseband pulse corresponding to the optimum time-limited waveform is obtained. Optimum baseband pulse shapes are computed for several modulation indices and a range of channel bandwidths. The results are presented in Chapter 6.

1.4 Outline of the Report

In conclusion, an outline of this study is as follows. A method to obtain the spectrum of correlative encoded digital FM signals is presented in Chapter 2. In Chapter 3, the spectra of correlative encoded digital FM signals are examined for a range of encoding polynomials and pulse shapes. An exhaustive investigation of the spectrum for modulation by a second order encoding polynomial is carried out. The error performance of correlative encoded FM is studied in Chapter 4.

Baseband pulse shapes in MSK-type signals are derived, in Chapter 5, to minimize the fraction of out-of-band power. An achievable lower bound on the fraction of out-of-band power is obtained for a range of channel bandwidths of practical interest. The more general problem of shaping the baseband pulses in angle modulated signals with an arbitrary modulation index is considered in Chapter 6.

The bandwidth efficiencies of various modulations, investigated in Chapters 2, 3, 5 and 6, are summarized in Chapter 7. Correlative encoded FM signals that are spectrally efficient, and as well possess good error performance capability are pointed out. Finally, the contributions made in this study are summarized and the scope for further work is suggested.

Chapter 2

THE SPECTRUM OF CORRELATIVE ENCODED DIGITAL FM

Previous experimental and simulation results have shown correlative encoded frequency shift keying (FSK) to be attractive from a spectral occupancy point of view. With the exception of duobinary encoding, the spectrum of correlative encoded FM has not been given theoretically. A method of obtaining an explicit expression for the spectrum of a correlative encoded digital FM signal is presented in this chapter using the results of Rowe and Prabhu [33] on FM spectra.

2.1 A Correlative Encoded Digital FM Modulator

A correlative encoded FM modulation can be represented by the model shown in Figure 2.1. The block D corresponds to a delay of one bit period. Such a correlative scheme has a partial response system (PRS) polynomial [5]

$$(k_0 + k_1 D + k_2 D^2 + \ldots k_m D^m)/C \qquad (2.1)$$

where $C = \sum_{\ell=0}^{m} |k_\ell|$

For convenience a normalization of the polynomial has been

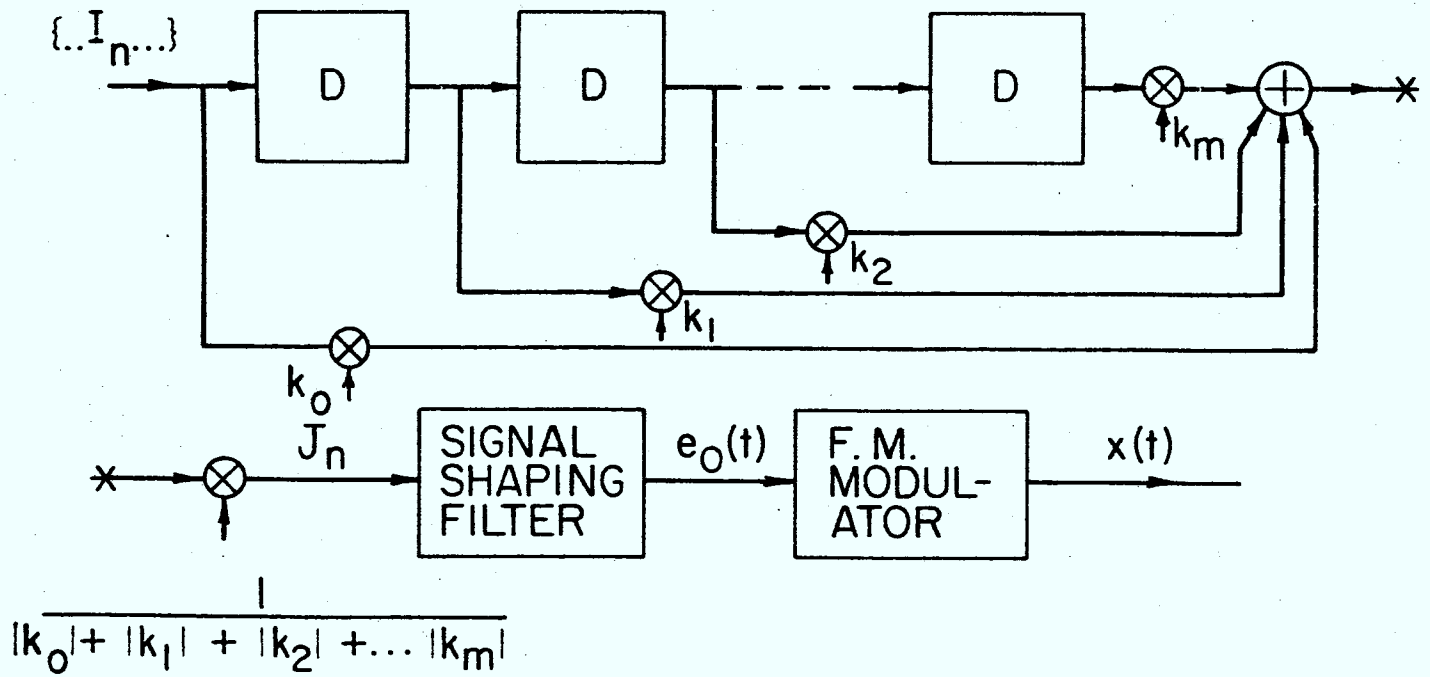


Figure 2.1 Correlative Encoded Digital FM Modulator

carried out. In Figure 2.1, the I_n are independent identically distributed random variables which can assume a finite number L, of discrete values. The correlated variables J_n are given by

$$J_n = \sum_{\ell=0}^{m} k_\ell I_{n-\ell}/C \tag{2.2}$$

Each J_n can assume at most $L^{(m+1)}$ distinct values. It is assumed that the signal shaping filter in the modulator produces an arbitrary waveform s(t) of finite duration, such that

$$s(t) = 0;\ t < 0 \text{ and } t > qT \tag{2.3}$$

where q is some integer. The modulating signal $e_o(t)$ can be written as

$$e_o(t) = \sum_{n=-\infty}^{\infty} J_n s(t - nT) \tag{2.4}$$

and the frequency modulated signal x(t) is given by

$$x(t) = \cos[2\pi f_c t + \int_{-\infty}^{t} e_o(\tau)d\tau] \tag{2.5}$$

The depth of modulation is taken into account by a scaling of s(t). The problem of finding the spectral properties of x(t) is approached using Rowe and Prabhu's method [33].

2.2 Rowe and Prabhu's Method

Rowe and Prabhu [33] have proposed a matrix method

to find the spectrum of a carrier, frequency modulated by a random baseband pulse train in which the signalling pulse duration is finite and the signal pulses may overlap. The symbols transmitted must be statistically independent. A summary of the results required here, will be given using their notation.

$$\text{Let} \qquad x(t) = \cos\left(2\pi f_c t + \int_{-\infty}^{t} e_o(u)du\right) \tag{2.6}$$

$$= \mathrm{Re}[v(t)\exp(j2\pi f_c t)]$$

where $$v(t) = \exp\left(j \int_{-\infty}^{t} e_o(u)du\right)$$

and $$e_o(t) = \sum_{k=-\infty}^{\infty} h_{s_k}(t - kT)$$

In the above expressions v(t) is the phasor associated with x(t). The notational conventions used are explained in the footnote. In (2.6) s_k is a random variable which can take on M values, and $h_{s_k}(t)$ is one of M possible modulating waveforms. The expression for $e_o(t)$ can be written as the inner product of a random vector $\underline{a_\nu}$ times a vector of all the possible modulating waveforms $\underline{h(t)}$ Thus

Footnote: The following notational conventions of Rowe and Prabhu [33] are adopted:

(i) Row and column vectors are distinguished by $\llcorner$ and], respectively.
(ii) Matrix multiplication is indicated by ·, Kronecker matrix products by X.
(iii) The transpose of a matrix is indicated by ' .
(iv) The Hermitian transpose of a matrix is indicated by †.
(v) Multiple Kronecker products are indicated by ΠX and the Kronecker power is indicated by an integer exponent enclosed in square brackets.

$$e_o(t) = \sum_{k=-\infty}^{\infty} \underline{a_k} \cdot h(t - kT)] \tag{2.7}$$

where

$$\underline{a_k} = [a_k^{(1)} \; a_k^{(2)} \; \ldots \; a_k^{(M)}]$$

$$a_k^{(j)} = 1 \quad \text{if } s_k = j$$

$$= 0 \quad \text{otherwise}$$

and

$$\underline{h(t)} = h(t)]' \equiv [h_1(t) \; h_2(t) \; \ldots \; h_M(t)]$$

Then the spectral density of the phasor v(t) associated with x(t) is given by

$$P_v(f) = \frac{1}{T} \underline{R(f)} \cdot (A + A^{\dagger}) \cdot R^{*}(f)] \tag{2.8}$$

where

$$A = \frac{1}{2} w_d^{[K]}$$

$$+ \sum_{n=1}^{K-1} e^{-j2\pi nfT} \{\underline{q(U_K)} \cdot w_d\}^{[n]} X \; w_d^{[K-n]} \; X \; w]^{[n]}$$

$$+ \frac{\{e^{-j2\pi fT} \; w] \cdot \underline{q(U_K)} \cdot w_d\}^{[K]}}{1 - e^{j2\pi fT} \underline{w} \cdot q(U_K)]} \tag{2.9}$$

and

$$R(f)] = \int_{-\infty}^{\infty} e^{-j2\pi ft}\, r(t)]dt \qquad (2.10)$$

Here K denotes the number of bit periods over which the modulating pulse extends. A brief explanation of the terms appearing in the above expressions follows

$$\underline{w} = [w_1\ w_2\ \cdots\ w_M] \qquad (2.11)$$

where

$$w_i = \text{Prob}\ \{s_k = i\}$$

and, w_d is an M x M diagonal matrix with diagonal terms $w_1, w_2, w_3 \cdots w_M$.

The signalling baseband pulses $h_i(t)$ extend over K bit periods, and the method requires a centering as follows

$$h(t)] = 0] \quad t < L_K, \quad t > U_K \qquad (2.12)$$

$$L_K = \begin{cases} -(K-1)T/2, & K \text{ odd} \\ -KT/2\ , & K \text{ even} \end{cases}$$

$$U_K = \begin{cases} (K+1)T/2, & K \text{ odd} \\ KT/2\ , & K \text{ even} \end{cases}$$

The vector r(t)] in the expression (2.10) is given by

$$r(t)] = \begin{cases} \prod_{i=-(K-1)/2}^{(k-1)/2} X\ q(t - iT)] , & K \text{ odd} \\ \prod_{i=-(k-2)/2}^{K/2} X\ q(t - iT)], & K \text{ even} \end{cases} \tag{2.13}$$

and

$$\underline{q(t)} = \left[\exp(j \int_{L_K}^{t} h_1(u)du) \quad \exp(j \int_{L_K}^{t} h_2(u)du) \right.$$

$$\left. \ldots \exp(j \int_{L_K}^{t} h_M(u)du) \right] \tag{2.14}$$

$$L_K < t < U_K$$

$$= 0 \qquad \text{elsewhere}$$

Equation (2.8) gives the spectral density when the following condition, a necessary condition for spectral lines to exist, is not satisfied.

$$| \underline{w} \cdot q(U_K)] | = 1 \tag{2.15}$$

When the above condition is satisfied the expressions for the line component and the continuous component of the spectrum are given by [eq. 72, 33] and [eq. 73, 33] respectively

2.3 Correlative Encoded FM Spectrum

The modulating signal for correlative encoding is given

by

$$e_o(t) = \sum_{n=-\infty}^{\infty} J_n \, s(t - nT) \tag{2.16}$$

In this expression the J_n are not statistically independent, but we can write $e_o(t)$ as

$$e_o(t) = \frac{1}{C} \sum_{n=-\infty}^{\infty} \sum_{j=0}^{m} k_j \, I_{n-j} \, s(t - nT)$$

$$= \sum_{n=-\infty}^{\infty} I_n \, b(t - nT) \tag{2.17}$$

where

$$b(t) = \frac{1}{C} \sum_{i=0}^{m} k_i \, s(t - iT) \tag{2.18}$$

The function b(t) extends over (m + q) bit intervals. Essentially the memory in the correlative encoder has been taken into account by redefining a new equivalent baseband pulse. Thus it can be seen from (2.17) that a correlative encoded FM signal can be viewed as a digital FM modulation in which the baseband pulses extend over more than one bit interval and the input data symbols are independent. Rowe and Prabhu's method can be applied directly to find the spectrum of the modulated signal because the I_n in (2.17) are statistically independent. If the L discrete levels that I_n can assume are denoted by $(\ell_1, \ell_2, \ell_3 \ldots \ell_L)$, $e_o(t)$ can be written in the matrix notation as

$$e_o(t) = \sum_{k=-\infty}^{\infty} \underline{a_k} \cdot h(t - kT)] \qquad (2.19)$$

where

$$\underline{a_k} = [a_k^{(1)} \; a_k^{(2)} \; . \; . \; . \; a_k^{(L)}]$$

is the vector with one element $a_k^{(j)} = 1$ and the remaining elements zero, as described above. The particular j corresponds to the waveform $h_j(t)$ transmitted at time kT. The components of the vector

$$\underline{h(t)} = [h_1(t) \; h_2(t) \; . \; . \; . \; h_L(t)]$$

are

$$h_i(t) = \ell_i \, b(t) \quad \text{for } i = 1, 2, \ldots L \qquad (2.20)$$

The results of Section 2.2 can be applied directly to compute the spectrum of correlative encoded digital FM signal modulated by the baseband signal $e_o(t)$ given by (2.19).

2.4 Definition of Modulation Index

As mentioned in Chapter 1, modulation index is linearly proportional to the magnitude of the baseband modulating signal. In conventional FSK, the modulation index h is defined as $h = 2f_dT$, where f_d is the peak instantaneous frequency deviation of the carrier. The baseband pulse shape is rectangular in conventional FSK and therefore the instant-

aneous frequency is constant over the bit interval. With baseband pulse shaping the instantaneous frequency varies during the bit interval. In the current study the modulation index, η, is defined as the maximum phase accumulation during a bit interval. Thus for conventional FSK, the modulation index, η, is given by $\eta = h\pi = 2\pi f_d T$. In this report the baseband pulses are scaled so that the modulation index is η. In terms of the equivalent baseband pulse, $b(t)$, η is given by

$$\eta = \max_{i=1,2\ldots L} |\ell_i| \sum_{k=-\infty}^{\infty} \left| \int_{kT}^{(k+1)T} b(t)dt \right| \qquad (2.21)$$

2.5 Binary Correlative Encoded FM Spectrum for a Second Degree Polynomial

The spectrum of a binary correlative encoded FM signal with rectangular pulse shaping and a second degree encoding polynomial is given in this section. In binary correlative encoding the input random variable I_n assumes only two discrete values, and let these values be $\ell_1 = 1$ and $\ell_2 = -1$. The shaping pulse $s(t)$ is a rectangular pulse over one bit period, that is, $q = 1$, and the encoding polynomial is of degree two.

$$(k_o + k_1 D + k_2 D^2)/C \qquad (2.22)$$

From the development of equation (2.19), the modulating signal $e_o(t)$ can then be written as

$$e_o(t) = \sum_{k=-\infty}^{\infty} \underline{a_k} \cdot h(t - kT)] \qquad (2.23)$$

where, $\underline{a_k}$ and $\underline{h(t)}$ are each of dimension 2. If the assumed rectangular s(t) of height $2\pi f_d$ is substituted in (2.18) and the resulting expression for b(t) is in turn substituted in (2.20) then, after the centering given by (2.12),

$$h_1(t) = b(t) = \frac{2\pi f_d}{C} \sum_{i=-1}^{+1} k_{i+1}\, p(t - iT) \qquad (2.24)$$

where

$$p(t) = 1 \quad 0 \leq t < T$$

$$= 0 \quad \text{elsewhere}$$

Also $h_2(t) = -h_1(t)$

From the expressions for $h_1(t)$ and $h_2(t)$ it is seen that the maximum deviation in the instantaneous frequency from the carrier frequency is f_d. The modulation index η defined in the last section is then given by $\eta = 2\pi f_d T$. Traditionally for conventional FSK, the modulation index, h, is defined as $h = 2f_d T$. When k_1 and k_2 in the PRS polynomial are zero, the modulated signal reduces to conventional FSK and η/π and h are identical.

If the binary input symbols are assumed to be equiprobable,

$$\underline{w} = [0.5 \quad 0.5] \tag{2.25}$$

and

$$w_d = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$$

The necessary condition for the spectral lines to exist, given by equation (2.15), reduces to

$$|\cos\{(k_o + k_1 + k_2)\eta/C\}| = 1 \tag{2.26}$$

When the spectrum does not contain discrete spectral lines, the power spectrum is given by equation (2.8). Since, $h_1(t)$ and $h_2(t)$ are known, $\underline{R(f)}$ and A can be obtained from (2.10) and (2.9) respectively. The imaginary part of equation (2.8) vanishes and the real part gives the required mean power spectral density. The expression for the spectral density in terms of matrix products in the form of equation (2.8) is convenient to program on the computer to obtain numerical results. However, an explicit algebraic expression can be obtained by expanding the matrix products. The vector $\underline{R(f)}$ for a second degree polynomial is given in Appendix A. The fact that the final expression is even with respect to both f and η is useful in partly reducing the tediousness of the algebraic expansion of the matrix products. When $k_2 = 0$, the expression reduces to

$$P_v(f)/T = \sum_{i=1}^{6} F_i(f,\eta) + \sum_{i=1}^{4} F_i(f,-\eta) \qquad (2.27)$$

where

$$F_1(f,\eta) = \nu^2(x_1)[1 + \cos(f_r - \eta) + z_1/4z_4]/4$$

$$F_2(f,\eta) = \nu^2(x_2)[1 + z_3/4z_4]/4$$

$$F_3(f,\eta) = \nu(x_1)\nu(x_2)[\cos\{f_r - k_1\eta/(k_1+1)\} + z_3\cos\{k_1\eta/(1+k_1)\}/2z_4]/4$$

$$F_4(f,\eta) = \nu(x_1)\nu(x_3)[\cos\{f_r - \eta/(k_1+1)\} + z_2\cos\{\eta/(1+k_1)\}/2z_4]/4$$

$$F_5(f,\eta) = \nu(x_2)\nu(x_3)[\cos f_r + z_3\cos\eta/4z_4]/2$$

$$F_6(f,\eta) = \nu(x_1)\nu(x_4)z_3\cos\{(1-k_1)\eta/(1+k_1)\}/8z_4$$

and

$$x_1 = (f_r - \eta)/2 \quad ; \quad x_2 = \{f_r - (k_1 - 1)\eta/(k_1 + 1)\}/2$$

$$x_3 = \{f_r + (k_1-1)\eta/(k_1+1)\}/2 \quad ; \quad x_4 = (f_r + \eta)/2$$

f = (frequency - f_c) ; f_c = carrier frequency

$f_r = 2\pi f T$; $\eta = 2\pi f_d T$

$$z_1 = 2\cos(2f_r - 2\eta) - 2\cos\eta\cos(f_r - 2\eta)$$

$$z_2 = 2\cos(2f_r - \eta) - 2\cos\eta\cos(f_r - \eta)$$

$$z_3 = 2\cos 2f_r - 2\cos\eta\cos f_r$$

$$z_4 = 1 - 2\cos f_r\cos\eta + \cos^2\eta$$

and

$$\nu(x) = \sin(x)/x$$

$k_o = 1$ is assumed in the above expression without any loss of generality

If $k_1 = 0$ is substituted in the above expression, it reduces to equation (48) of Bennett and Rice [27]. When $k_1 = 1$, we have duobinary FSK and the expression agrees with equation (7) of Von Baeyer and Tjhung [7]. It agrees also with equation (5.10) of Roth [8] after correcting the typographical errors in his expression.

2.6 Summary

A method for obtaining the spectrum of a correlative encoded digital FM signal was presented in this chapter. This technique was applied to obtain the spectrum of a signal modulated by a second order PRS encoding polynomial with rectangular pulse shaping. The result is a generalization of

previous results for FSK spectra and agrees with the previously known results for the special cases of duobinary and conventional FSK modulation.

Chapter 3

INVESTIGATION OF CORRELATIVE ENCODED FM SPECTRA

A method of obtaining the spectrum of a correlative encoded digital FM signal was presented in Chapter 2. Spectra are computed for various correlative encoding polynomials and modulation indices, and the results are presented in this chapter. An exhaustive investigation of the spectrum for modulation by a second order polynomial has been carried out and the values of the encoding coefficients, for which the spectrum is compact, are shown. The spectrum of the modulated signal is determined by the encoding polynomial as well as the pulse shape. In this chapter, results are given for rectangular and raised-cosine pulse shapings.

3.1 Equivalent PRS Polynomials

In an investigation of partial response encoded FM, a much larger class of encoding polynomials must be considered than for a linear modulation. There are no simple properties of the polynomial that determine the predominant spectral characteristics and eliminate all but a few polynomials from consideration. In an investigation of the spectra of partial response FM, the effort involved can be reduced by noting that a partial response encoding polynomial $\frac{1}{C}\sum_{i=0}^{m} k_i D^i$ and its

reciprocal $\frac{1}{C}\sum_{i=0}^{m} k_{m-i}D^i$ yield identical spectra. The argument follows from the fact that a modulating waveform $e_o(t)$ and its time reversal $e_o(-t)$ lead to identical FM spectra. Thus if the shaping pulse s(t) is symmetrical, an encoding polynomial and its reciprocal yield identical spectra. The proof is given in Appendix B.

3.2 Binary Correlative Encoding with Rectangular Pulse Shaping

The spectrum of a binary correlative encoded FM, with rectangular pulse shaping and a second order PRS polynomial was given in Section 2.5. The spectra for different values of the encoding coefficients k_1 and k_2 are examined exhaustively here, for a modulation index of 0.5π. Throughout the following investigation it is assumed that k_0 is unity without any loss of generality.

Let us consider first the spectrum of a signal modulated by a first order polynomial, that is $k_2 = 0$. Melvin and Middlestead [6] have obtained results experimentally and by computer simulation for the case $k_1 = 1$. Even though their simulation and experimental results are in good agreement, they note that their results do not agree with the analytical results obtained by Roth [8]. This is due to the omission of two terms in the printing of Roth's result. The results obtained here agree with the observations of Melvin and Middlestead.

It was pointed out in Section 3.1 that the reciprocal encoding polynomials $\frac{1}{C}\sum_{i=0}^{m} k_i D^i$ and $\frac{1}{C}\sum_{i=0}^{m} k_{m-i}D^i$ yield

identical spectra. Thus the spectra for $k_1 = k$ and $k_1 = 1/k$ are identical and hence, it is sufficient to examine the spectrum for k_1 ranging from -1 to +1. When $k_1 = -1$, the spectrum contains discrete spectral lines, and the spectrum contains sharp spikes for values of k_1 close to -1. For positive values of k_1, the spectrum is most compact for $k_1 = 1$.

Let us now consider the signal modulated by a second order polynomial. From the result on equivalent encoding polynomials, it follows that for a second order polynomial the encoding coefficients (k_1,k_2) and $(k_1/k_2, 1/k_2)$ yield identical spectra. Therefore, it is sufficient to examine the region $-\infty < k_1 < \infty$ and $-1 \leq k_2 \leq 1$. For large values of k_1, the polynomial can be approximated by $k_1 D/|k_1|$, which asymptotically approaches the case of digital FM without correlative encoding.

The spectrum was computed for various values of k_1, k_2 in the region $-5 \leq k_1 \leq 5$ and $-1 \leq k_2 \leq 1$, and the results were applied to the other values of k_1, k_2 using the reciprocal polynomial result. The results are shown in Figure 3.1. The points marked E are the values of k_1,k_2 for which discrete spectral lines exist for all values of η. The spectrum contains sharp spikes for the values of k_1, k_2 marked D. The spectrum is most compact for the values of k_1, k_2 in the area marked A. The points marked B have less compact spectra but comparable to the spectrum of the polynomial $(1 + D)/2$. The spectra are least compact at the points marked C, but are comparable to MSK spectra.

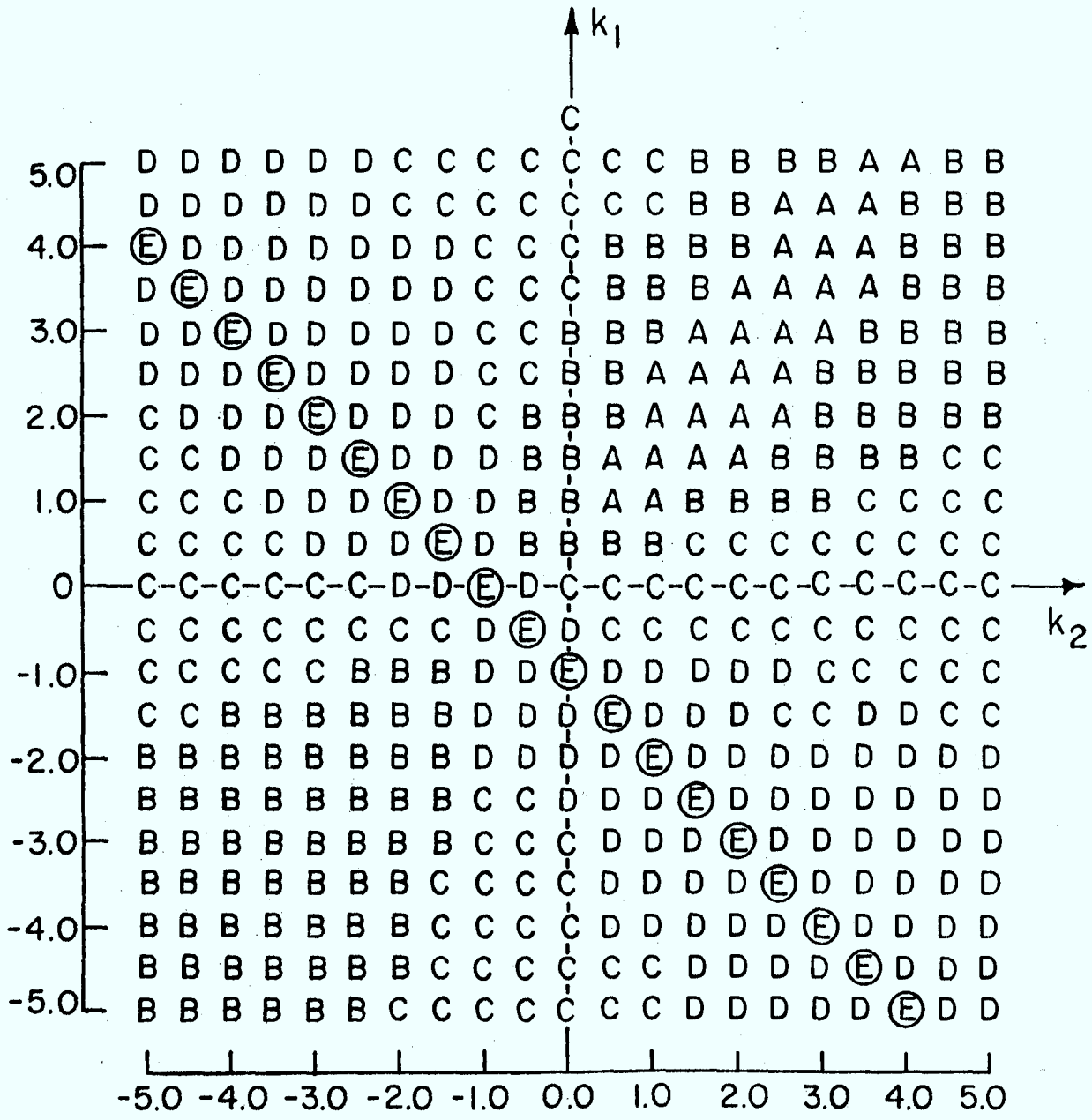


Figure 3.1 Behaviour of Spectrum with Rectangular s(t), q=1 and PRS Polynomial $(1+k_1D+k_2D^2)/C$

The PRS polynomials 1, $(1 + D)/2$ and $(1 + D + D^2)/3$ yield the most compact spectrum for the zeroth, first and second order polynomials respectively. As well they give the least number of output levels. The spectra for these polynomials for various modulation indices are given in Figures 3.3, 3.4 and 3.5. The equivalent baseband pulse shapes, b(t), for these signals, given by (2.18), are shown in Figures 3.2a, b, and c respectively. The spectra were computed by evaluating the expression (2.8). The vectors $\underline{R(f)}$, appearing in (2.8), associated with the equivalent baseband waveforms in Figures 3.2a, b and c are given in Appendix A.

Spectra for binary correlative encoded FM signals were also obtained experimentally. The experimental results were in good agreement with the theoretical results. Some illustrative spectra are shown in Figure 3.6.

3.3 Binary Correlative Encoding with Raised-Cosine Pulse Shaping

The spectrum of a correlative encoded FM signal depends on the encoding polynomial and the shaping pulse s(t). In this section, the spectra of the modulated signal is investigated when s(t) is a raised-cosine pulse extending over one, two and three bit intervals. Raised-cosine pulse shaping can be used to model baseband pulses that have a smooth transition from on to off. Since the polynomials 1, $(1+D)/2$ and $(1+D+D^2)/3$ provided the most compact spectra for rectangular s(t), we have studied the same polynomials for the raised-cosine pulse shaping.

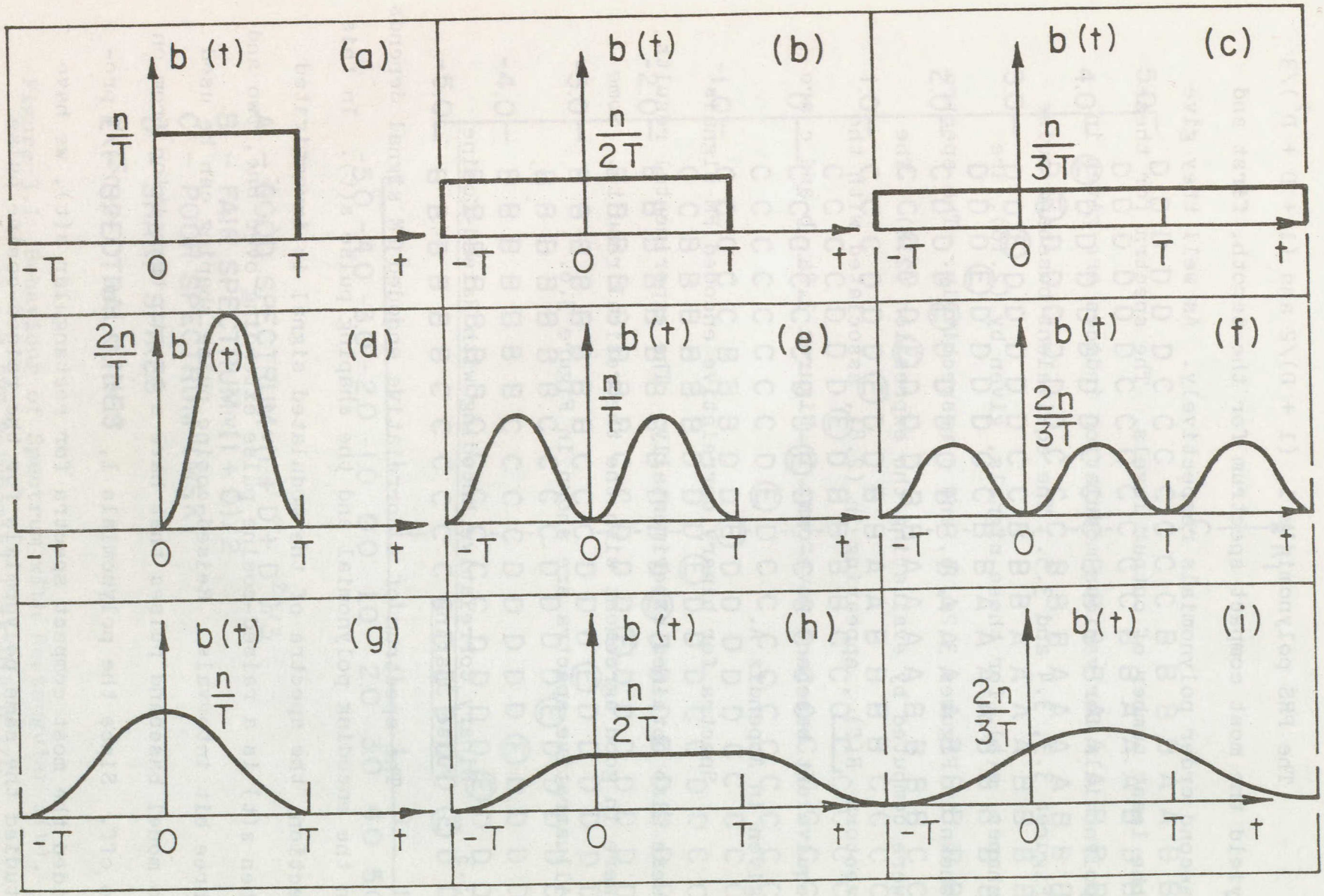


Figure 3.2 Equivalent Baseband Pulse Shapes, b(t), for the Modulations Investigated

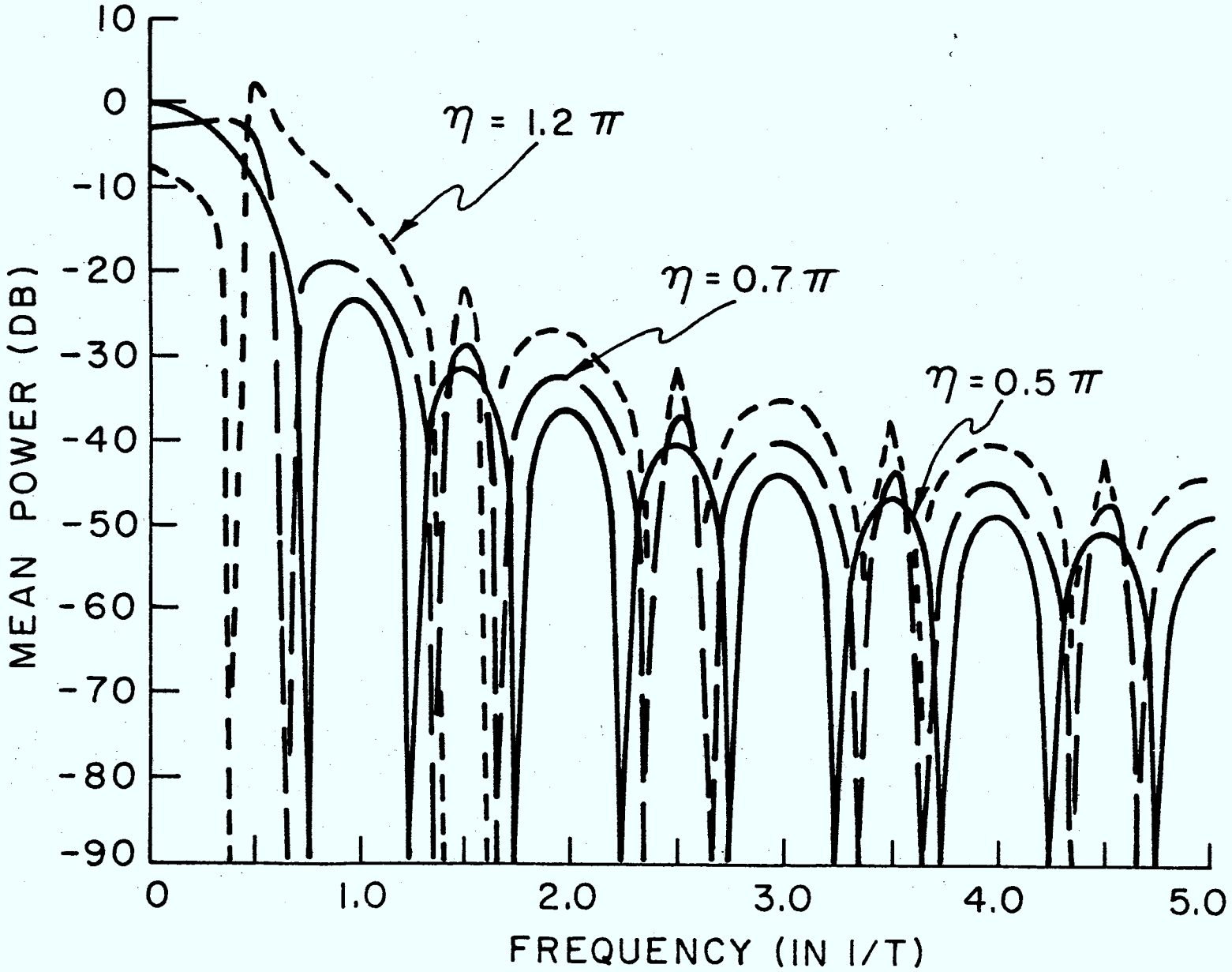


Figure 3.3 Spectrum with Rectangular s(t), q=1 and PRS Polynomial 1

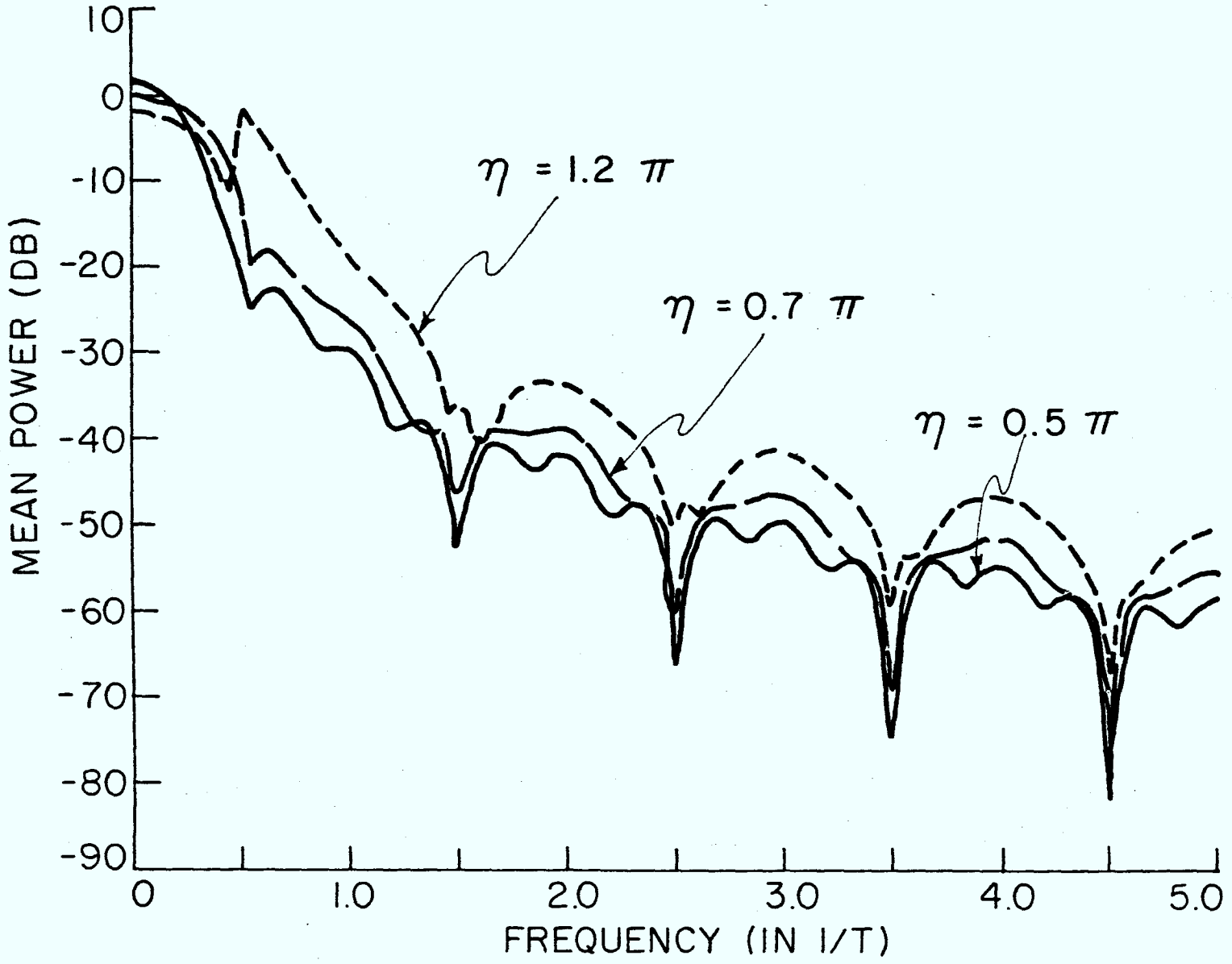


Figure 3.4 Spectrum with Rectangular s(t), q=1 and PRS polynomial (1+D)/2

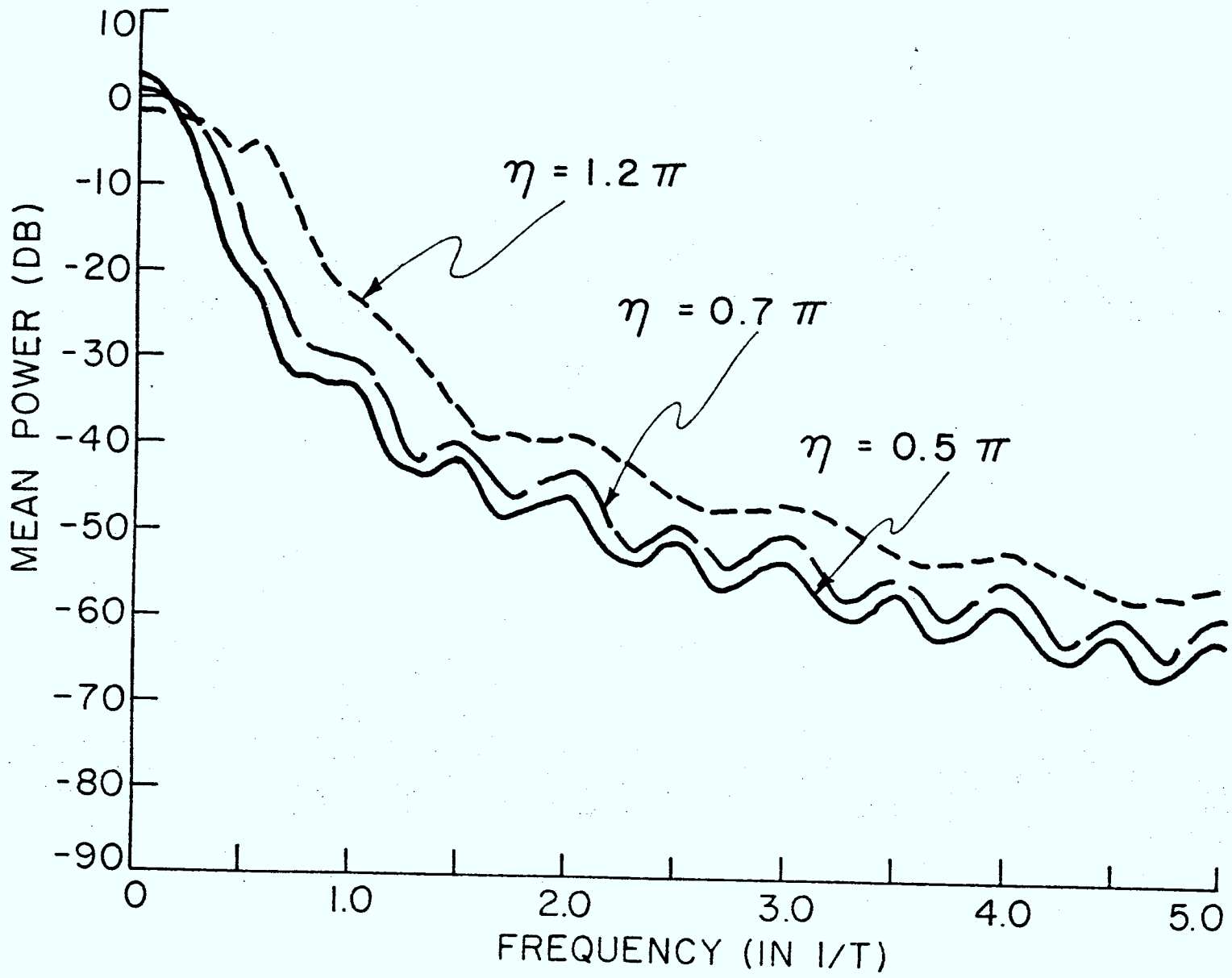
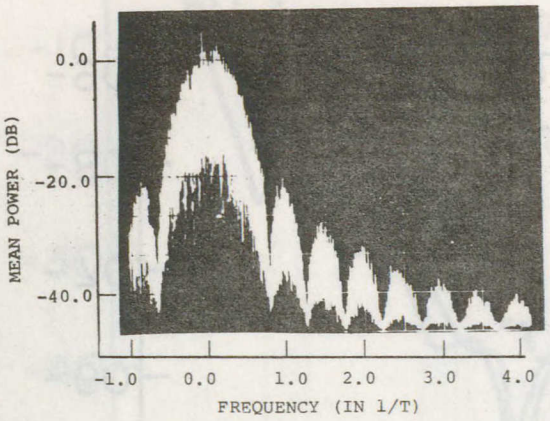
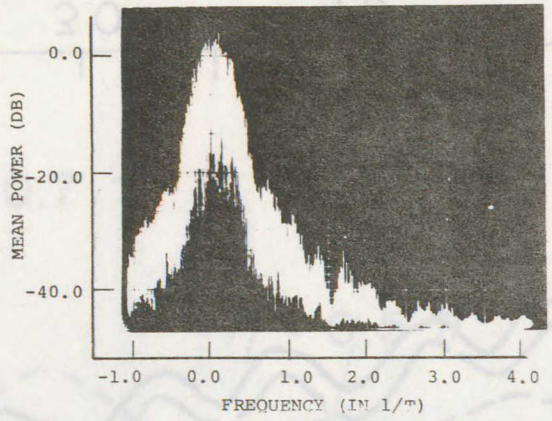


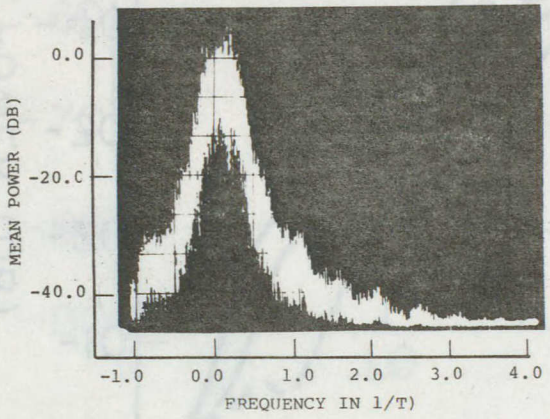
Figure 3.5 Spectrum with Rectangular s(t), q=1 and PRS Polynomial $(1+D+D^2)/3$



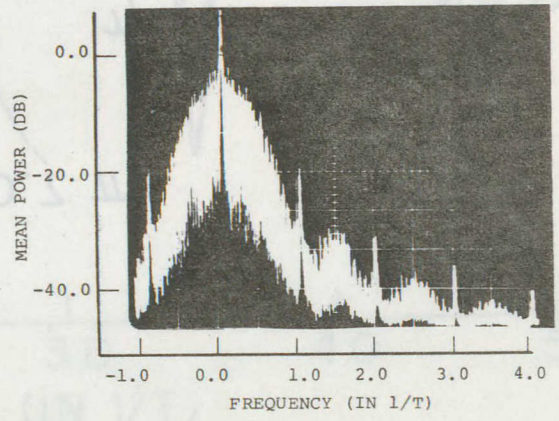
a) SPECTRUM WITH RECTANGULAR S(t) AND PRS=1, $\eta=0.5\pi$



b) SPECTRUM WITH RECTANGULAR S(t) AND PRS=$(1+D)/2$, $\eta=0.5\pi$



c) SPECTRUM WITH RECTANGULAR S(t) AND PRS=$(1+D+D^2)/3$, $\eta=0.5\pi$



d) SPECTRUM WITH RECTANGULAR S(t) AND PRS=$(1-D^2)/2$, $\eta=0.5\pi$

Figure 3.6 Spectrum Obtained Experimentally for Various PRS Polynomials with Rectangular s(t) and q=1

Let us first consider the signal when s(t) is a raised-cosine pulse extending over one bit interval.

$$\left.\begin{aligned} s(t) &= \frac{\eta}{T}\,(1 - \cos 2\pi t/T) && 0 \le t < T \\ &= 0 && \text{elsewhere} \end{aligned}\right\} \qquad (3.1)$$

The spectra for the polynomials 1, (1+D)/2 and $(1+D+D^2)/3$ with the above s(t) are given in Figures 3.7, 3.8 and 3.9. The corresponding equivalent baseband pulses b(t), given by (2.18), are shown in Figures 3.2d, e and f. The vectors $\underline{R(f)}$ associated with these baseband pulses are given in Appendix A.

We next consider shaping pulse, s(t), extending over two bit intervals, that is q = 2. When s(t) is a rectangular pulse with q = 2 and polynomial 1, it is identical to rectangular s(t) with q = 1 and the polynomial (1+D)/2. In addition for rectangular s(t) with q = 2 and polynomial (1+D)/2, and q = 1 and polynomial $(1+2D+D^2)/4$ the spectra are identical. When s(t) is a raised-cosine pulse with q = 2, it is given by

$$\left.\begin{aligned} s(t) &= \frac{\eta}{2T}\,(1 + \cos \pi t/T) && -T \le t < T \\ &= 0 && \text{elsewhere} \end{aligned}\right\} \qquad (3.2)$$

The spectra with the above s(t) and the polynomials 1 and (1+D)/2 are shown in Figures 3.10 and 3.11. The corresponding equivalent baseband pulse shapes, b(t), are shown in Figures

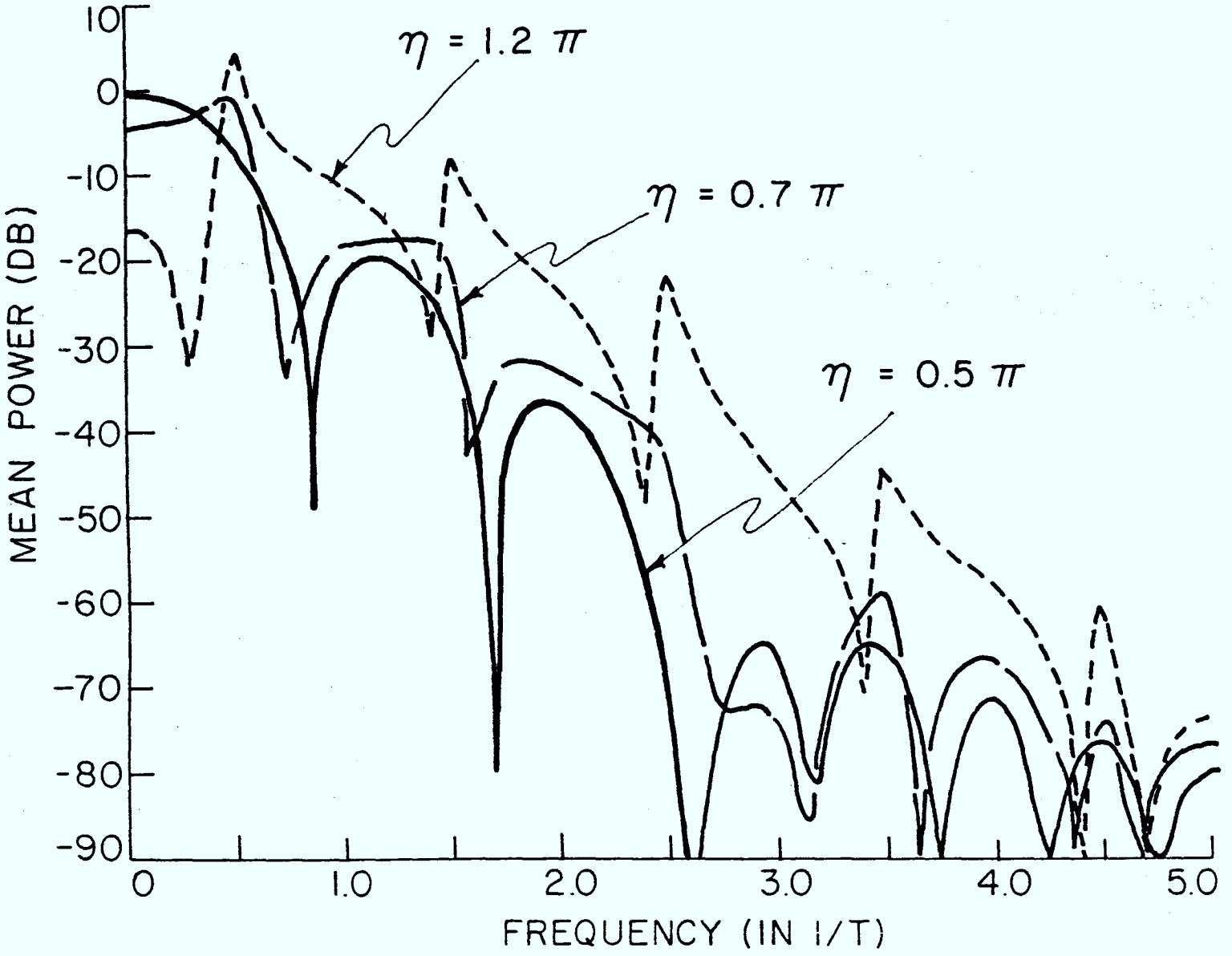


Figure 3.7 Spectrum with Raised-Cosine s(t), q=1 and PRS Polynomial 1

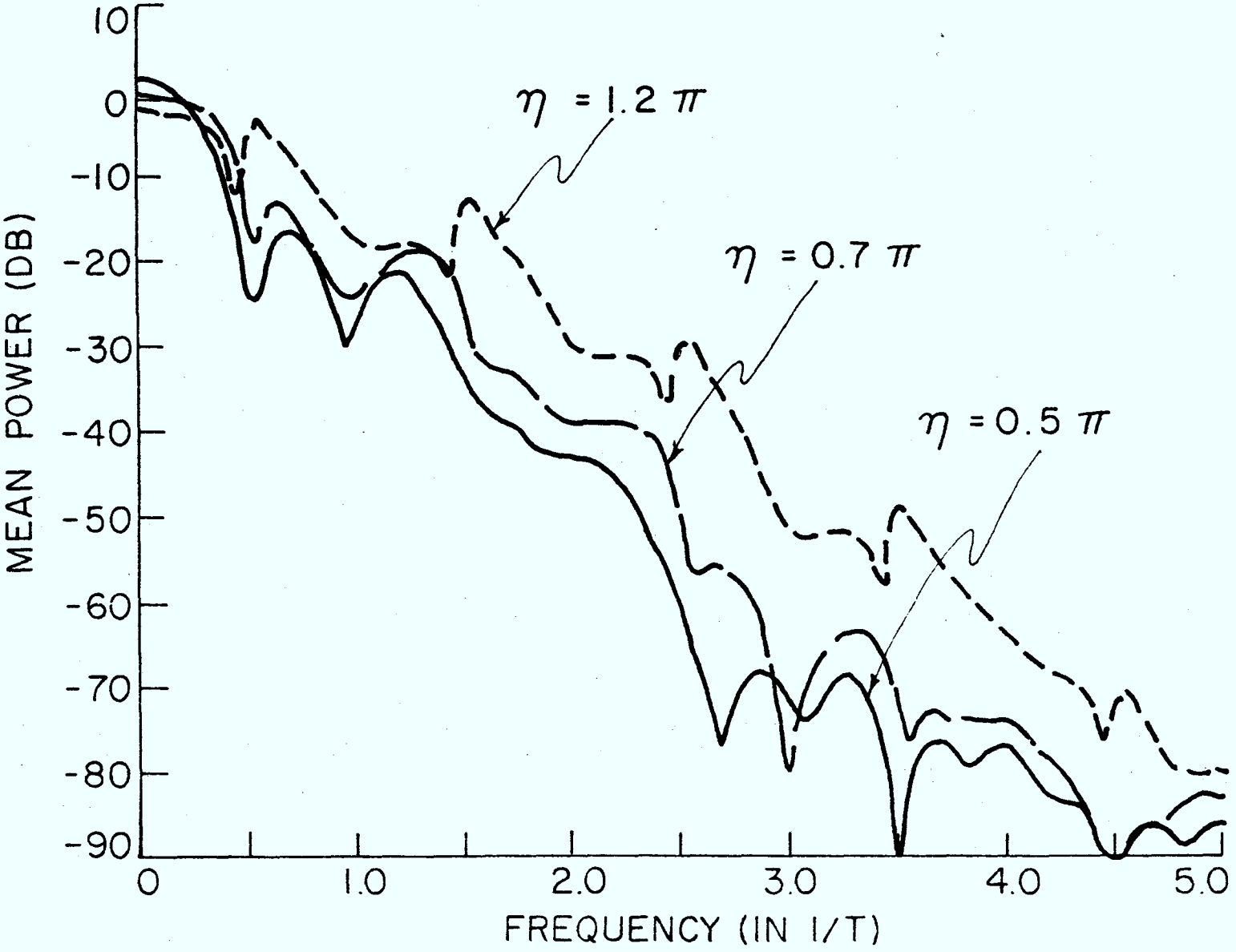


Figure 3.8 Spectrum with Raised-Cosine s(t), q=1 and PRS Polynomial (1+D)/2

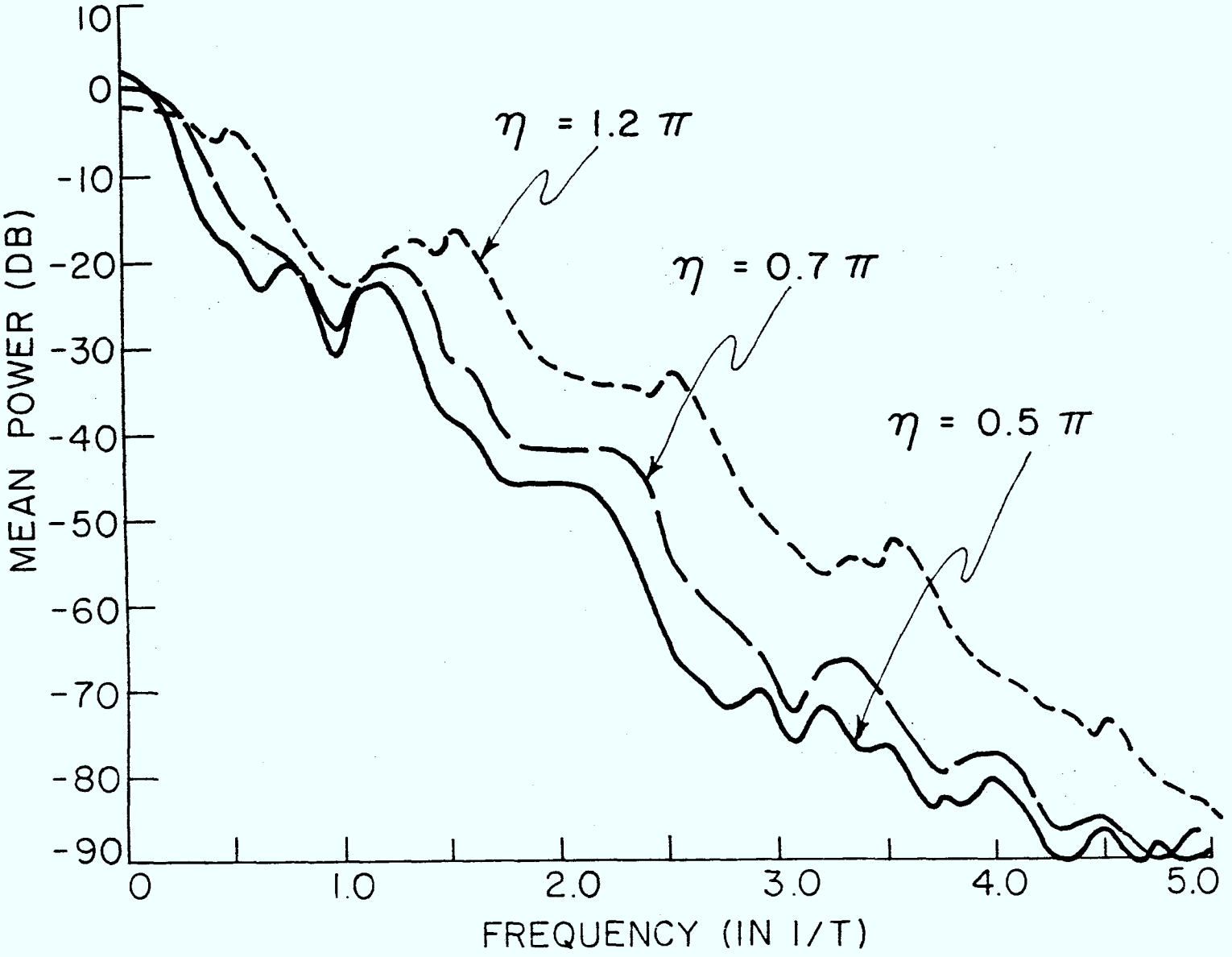


Figure 3.9 Spectrum with Raised-Cosine s(t), q=1 and PRS Polynomial $(1+D+D^2)/3$

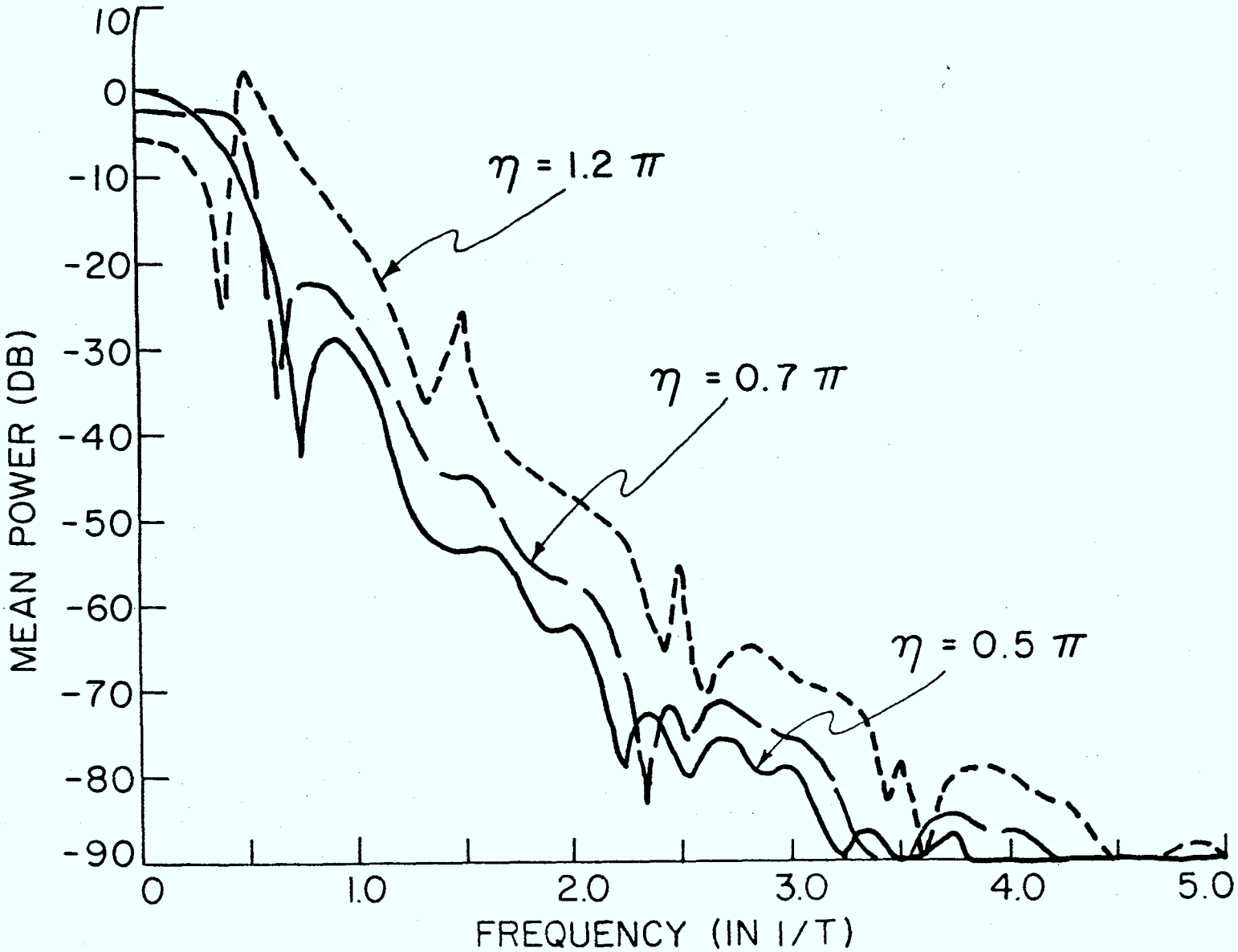


Figure 3.10 Spectrum with Raised-Cosine s(t), q=2 and PRS Polynomial 1

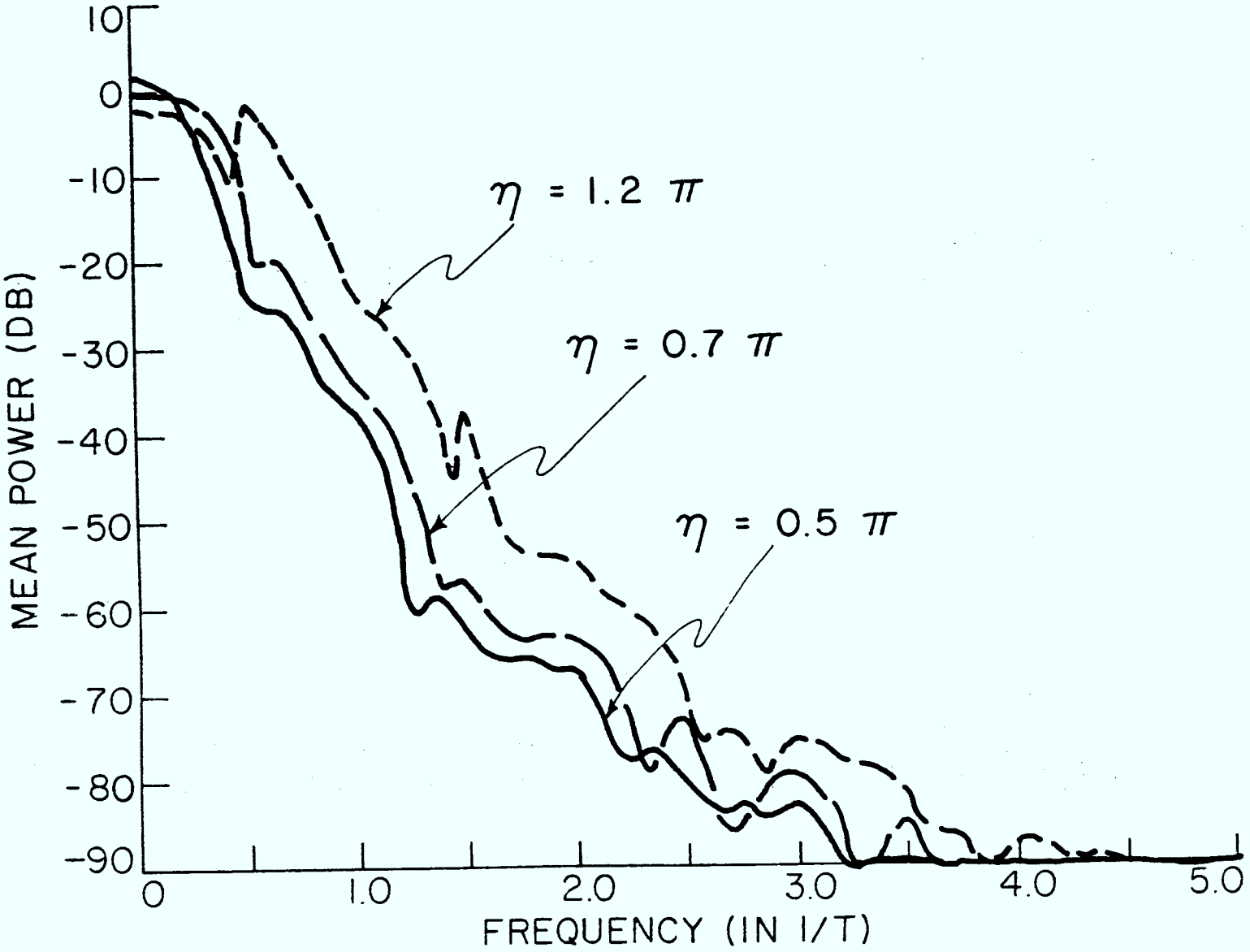


Figure 3.11 Spectrum with Raised-Cosine s(t), q=2 and PRS Polynomial (1+D)/2

3.2g and h. The vectors $\underline{R(f)}$ associated with these baseband pulses are given in Appendix A.

Finally we consider s(t) extending over three bit periods. When s(t) is a rectangular pulse with q = 3 and polynomial 1, the modulated signal is identical to s(t) being rectangular with q = 1 and the polynomial $(1+D+D^2)/3$. Spectra are given in Figure 3,12 for s(t) a raised-cosine pulse with q = 3 and polynomial 1. The corresponding baseband pulse b(t) is shown in Figure 3.2i. The vector $\underline{R(f)}$ associated with this baseband pulse is given in Appendix A also.

3.4 Discussion

The behaviour of the spectra observed for correlative encoded FM with rectangular and raised-cosine pulse shapings and encoding polynomials 1, $(1+D)/2$ and $(1+D+D^2)/3$ are summarized and discussed here. To compare the band occupancy of the signals modulated by the equivalent baseband pulse shapes shown in Figure 3.2, the bandwidths required to transmit 99% and 99.9% of the modulated signal power are listed in Table 3.1. The waveforms in Figures 3.2a, b and c correspond to PRS polynomials 1, $(1+D)/2$ and $(1+D+D^2)/3$ with rectangular pulse shaping. The spectra for these waveforms become progressively more compact for higher order polynomials. The waveforms in Figures 3.2d, e and f correspond to PRS polynomials 1, $(1+D)/2$ and $(1+D+D^2)/3$ with raised-cosine pulse shaping. Compared to rectangular shaping (Figures 3.2a, b and c), raised-cosine shaped waveforms (Figures 3.2d, e and f) yield spectra

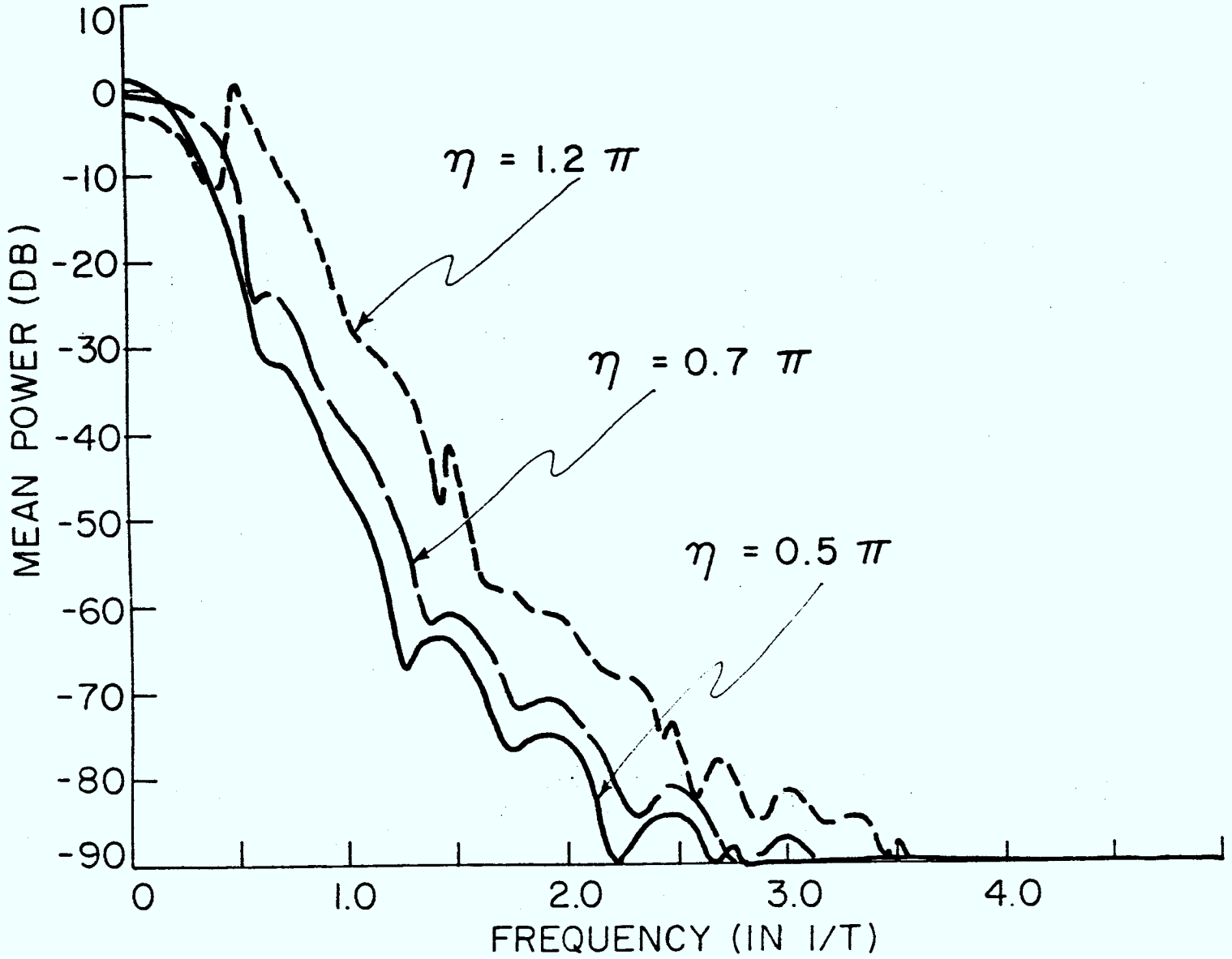


Figure 3.12 Spectrum with Raised-Cosine s(t), q=3 and PRS Polynomial 1

Baseband pulse shape, b(t), as shown in Figure	Bandwidth in units of 1/Bit Period							
	$\eta = 0.5\pi$		$\eta = 0.7\pi$		$\eta = 0.8\pi$		$\eta = 1.2\pi$	
	99%	99.9%	99%	99.9%	99%	99.9%	99%	99.9%
3.2a	1.18	2.80	1.80	3.14	2.00	3.80	2.30	4.40
3.2d	2.20	2.92	2.74	3.40	2.90	3.80	3.80	5.10
3.2b	0.92	1.80	1.30	2.20	1.44	2.30	2.00	3.00
3.2e	1.60	2.74	2.50	3.05	2.70	3.25	3.30	4.60
3.2g	1.10	1.65	1.20	2.00	1.50	2.08	1.96	2.60
3.2c	0.80	1.35	1.00	1.80	1.16	2.00	1.60	2.54
3.2f	1.48	2.60	2.00	2.90	2.40	3.00	3.12	3.80
3.2h	0.86	1.40	1.10	1.60	1.24	1.74	1.70	2.24
3.2i	1.00	1.20	1.06	1.50	1.20	1.60	1.70	2.04

Table 3.1 Bandwidth Utilization for Various Encoding Schemes

with better asymptotic behaviour, but require a wider bandwidth to transmit 99% of the modulated signal power. The equivalent baseband pulses shown in Figures 3.2 c and i appear to have better spectral behaviour than the other pulse shapes in Figure 3.2.

3.5 Quaternary Correlative Encoding with Rectangular Pulse Shaping

The method described in Chapter 2 can be applied directly to find the FM spectra for multilevel modulation signals and correlative encoding. As an example, we will consider quaternary modulation. Let the random variables I_n assume four discrete levels, that is L = 4 and let the four levels be $\ell_i = 5-2i$ for i=1,2,3,4. Suppose the encoding polynomial is (1+D)/2. This polynomial yielded good results for binary modulation. The modulating signal can be written as

$$e_o(t) = \sum_{k=-\infty}^{\infty} \underline{a_k} \cdot h(t-kT)] \qquad (3.3)$$

where $\underline{a_k}$ and $\underline{h(t)}$ each now have four components, and

$$h_i(t) = \ell_i b(t) = \ell_i (2\pi f_d/6)\{P(t+T)+P(t)\}$$

$$\text{for } i=1,2,3,4 \qquad (3.4)$$

where

$$P(t) = 1 \qquad \text{for } 0 \leq t < T$$

$$= 0 \qquad \text{elsewhere}$$

From the definitions of $h_i(t)$, it can be seen that the maximum deviation of the instantaneous frequency has been set to f_d. Then as before $\eta = 2\pi f_d T$. The encoded baseband will have 7 levels, and the corresponding frequency deviations are $\{f_d(1 - i/3), i=0,1,. . .6\}$. If all the four input levels are equiprobable

$$\underline{w} = [1/4 \quad 1/4 \quad 1/4 \quad 1/4] \tag{3.5}$$

and

$$w_d = 1/4\ [ID] \tag{3.6}$$

where [ID] is a (4x4) identity matrix.

The necessary condition for the spectral lines to exist, given by equation (2.15), reduces to

$$|\cos\eta/2\ \cos\eta| = 1 \tag{3.7}$$

When there are no discrete spectral lines, the mean power spectral density is given by (2.8). A computer program was written to evaluate this expression and the results are given in Figure 3.13. The spectrum has favourable asymptotic behaviour.

3.6 Summary

Spectra of correlative encoded digital FM signals for a range of encoding polynomials have been presented in this

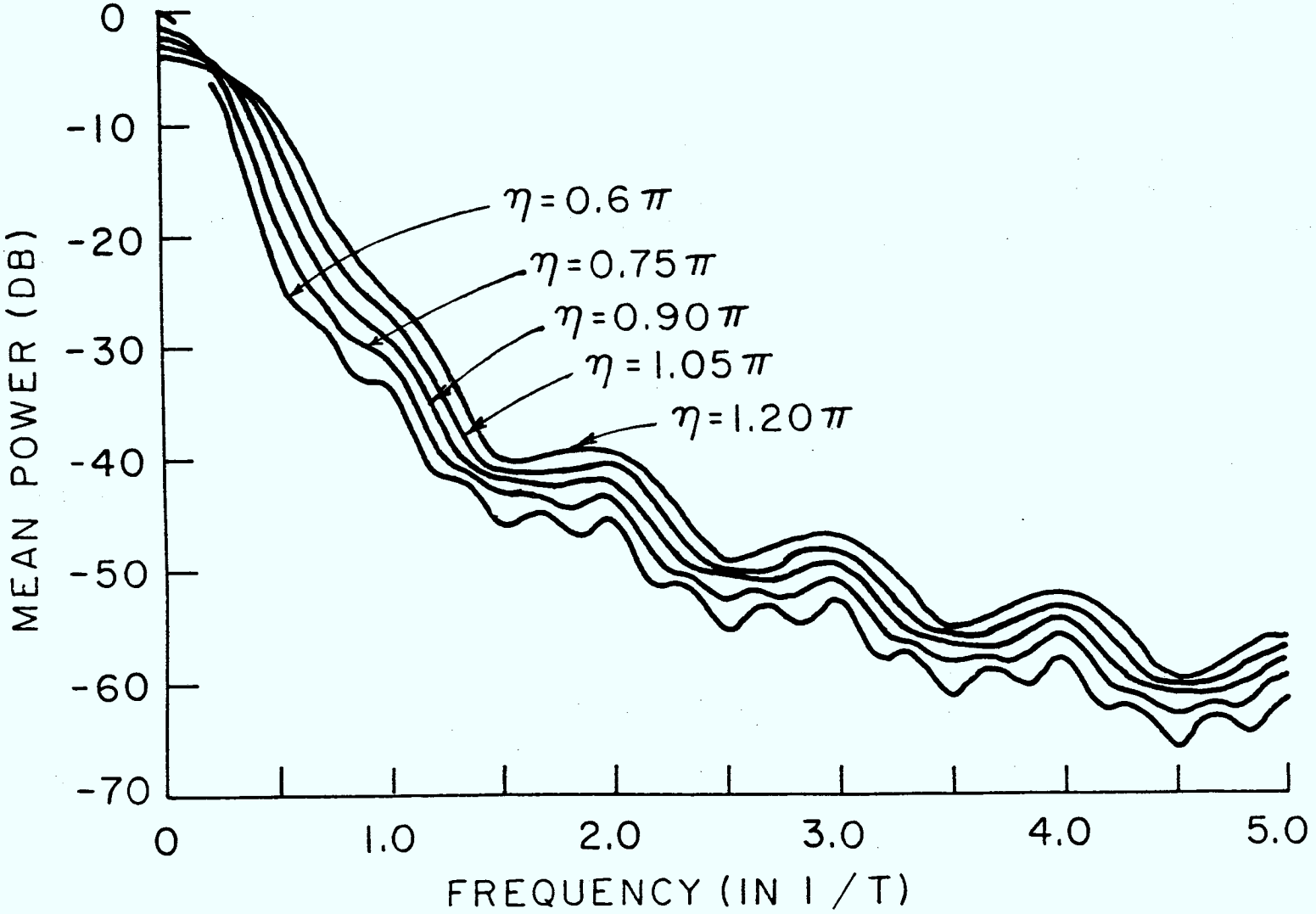


Figure 3.13 Spectrum for Quaternary Correlative Encoding with Rectangular s(t), q=1 and PRS Polynomial (1+D)/2

chapter. It is evident from an examination of the spectra that correlative encoded FM is superior to conventional FSK in terms of spectral conservation.

A thorough investigation of the spectrum of a second order polynomial has been presented, and the values of the encoding coefficients k_1, k_2 for which the spectrum is compact are shown. Spectra are given for various PRS encoding polynomials and modulation indices, for raised-cosine as well as rectangular pulse shaping. The bandwidths required to transmit 99% and 99.9% of the modulated signal power are tabulated for all the schemes considered.

The method is applicable to digital FM by multilevel modulating signals which have been correlatively encoded. As an example, results are given for a quaternary modulating signal with duobinary encoding.

Chapter 4

ERROR PERFORMANCE OF CORRELATIVE ENCODED DIGITAL FM

The spectral occupancy of correlative encoded FM was investigated in Chapters 2 and 3. The other major consideration in the choice of modulation is the SNR required to attain the desired level of error performance. The error performance of correlative encoded FM signals is investigated in this chapter. In partial response signalling a degradation is observed when conventional detectors are used. As an example, the error performance degradation for partial response signalling and discriminator detection is given in Section 4.1. Then we will turn to the performance with newer detectors such as the Viterbi decoder. In the analysis of communication systems, an upper bound on the performance of the optimum detector can be obtained in terms of the minimum Euclidean distance, using the union bound. The upper bound is evaluated for several correlative encoded FM signals in Section 4.2.

4.1 Discriminator Detection

A correlative encoded FM system with discriminator detection can be modelled by the block diagram shown in Figure 4.1. The performance of a discriminator detector has been

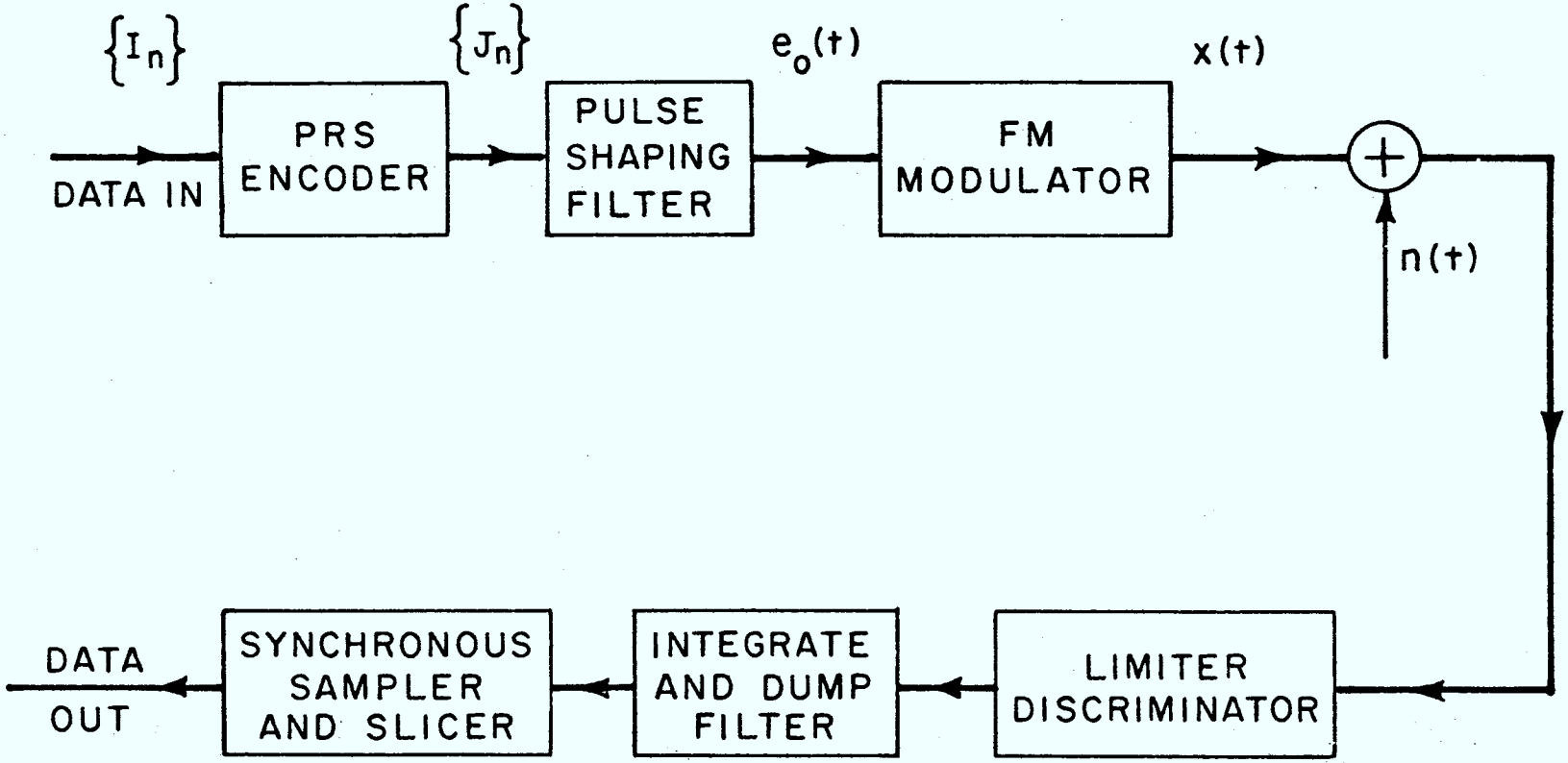


Figure 4.1 Block Diagram for Correlative Encoded FM with Discriminator Detection

investigated by several authors [35-39] for conventional FSK. Even though the noise in the channel is additive gaussian, the noise at the output of the limiter discriminator has a complicated distribution, because of nonlinearity of the limiter discriminator. The distribution of the noise at the output of the discriminator was given by Rice [34] using the theory of "clicks". This theory was adopted by Mazo and Salz [35] to analyze the performance of a discriminator detector for FSK modulation. Tjhung and Wittke [38] included the effects of filtering on the signal and obtained the parameters of the pre-detection filter to optimize the performance for various modulation indices. Papantoni-Kazakos and Kaz [39] gave a theoretical analysis for the performance when the channel introduces an arbitrary distortion causing intersymbol interference.

We next consider the degradation introduced in discriminator detection by correlatively encoding the baseband signal in FSK. Swartz [40] has derived an upper limit on the error rate for modified duobinary encoded FSK with binary data input. Lender [41] has given experimental results on the performance of modified duobinary encoded FSK with quaternary data input. We consider the performance of precoded duobinary FSK with discriminator detection. In the analysis, the noise, n(t), is assumed bandlimited but the signal, x(t), is assumed to be undistorted at the input to the detector. The modulated signal, x(t), is of amplitude A. The noise, n(t), has a flat power spectrum with two-sided power density of N_o watts/Hz,

and is bandlimited to B Hz. Thus, the SNR at the input to the discriminator is $\rho = A^2/(4BN_o)$. Even though the error rates for FSK are well known, they are reproduced in Figure 4.2 so that they can be used for comparison with duobinary FSK.

The performance of duobinary FSK (PRS encoding polynomial (1+D)/2 with rectangular baseband pulse shaping) with precoding will now be considered. During any bit interval, the modulated signal, x(t), has an instantaneous frequency of either f_c, (f_c+f_d) or (f_c-f_d). Here, f_c is the carrier frequency. The frequency deviation f_d, for a modulation index η, is given by

$$f_d = \frac{\eta}{2\pi T} \tag{4.1}$$

Let the output of the synchronous sampler be r. In duobinary encoding 3 signal levels are possible. For details on precoded duobinary transmission see Lucky, Salz and Weldon [42, Chapter 4]. The detection can be treated as a ternary hypothesis problem.

$$\left.\begin{array}{lll} H_+ \; ; & r = \eta + n_+ & \\ H_0 \; ; & r = n_o & \\ H_- \; ; & r = \eta + n_- & \end{array}\right\} \tag{4.2}$$

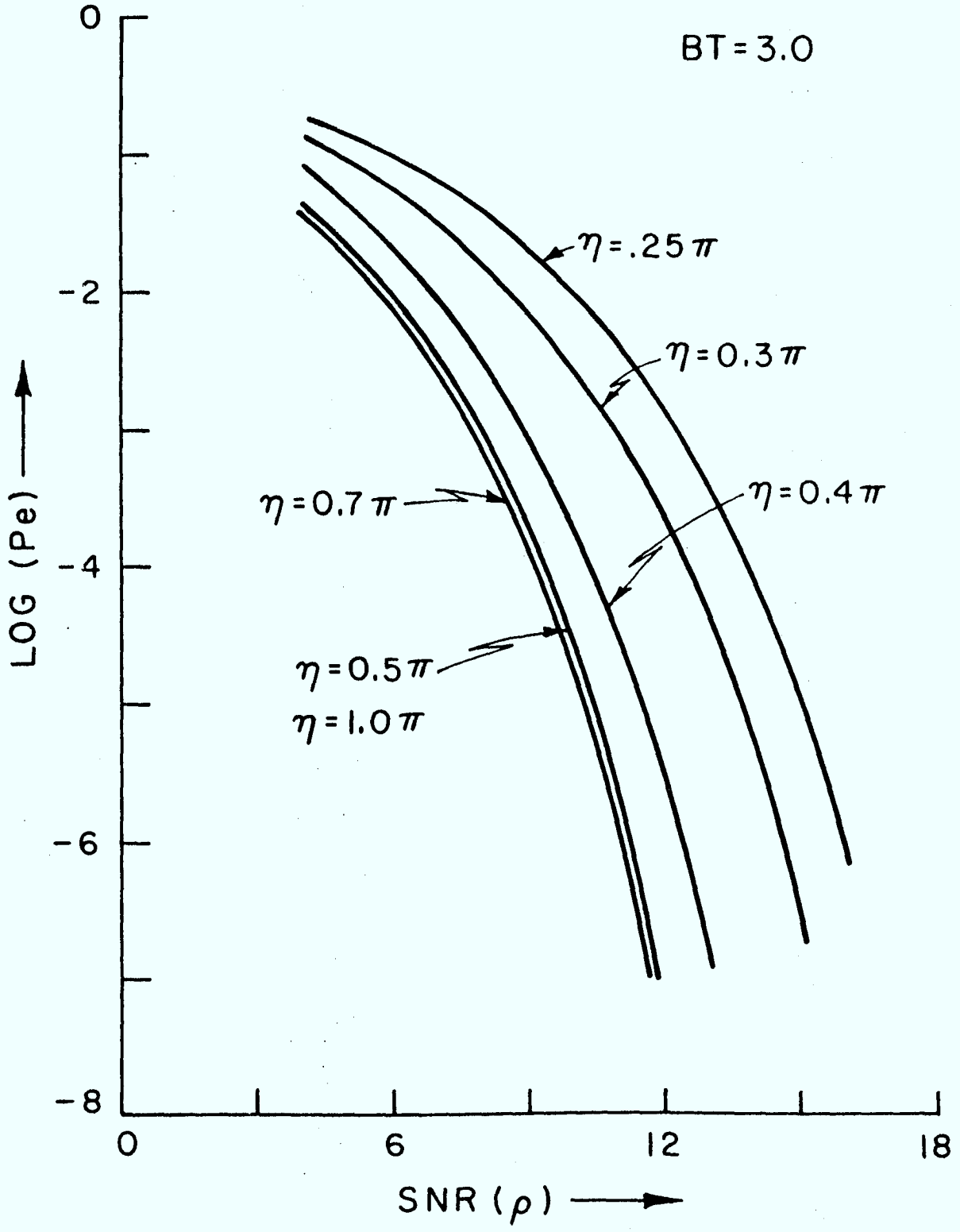


Figure 4.2 Probability of Error for FSK with Discriminator Detection

Different symbols are used for the noise in the three cases due to the dependency of the noise on signal frequency. Given the observation r, a decision has to be made between the three hypotheses. Due to precoding, H_+ and H_- correspond to a transmitted data symbol, 1, and H_o corresponds to a data symbol, -1. The noises n_+, n_o and n_- consist of two components. The first component is a continuous random variable which can take any value in the interval $[-\pi,\pi]$. The second component is a discrete random variable which assumes only discrete values of $\pm 2\pi m$, where m is the number of clicks during the bit interval. The probability density for these components can be calculated directly with the method used by Mazo and Salz [35]. The decision rule for the receiver is given by;

$$\left.\begin{array}{ll} \text{i)} & \text{Output is bit 1: if } |r| \geq \eta/2 \\ \text{ii)} & \text{Output is bit -1: if } |r| < \eta/2 \end{array}\right\} \qquad (4.3)$$

The probability of error with the above decision rule was calculated and the results are given in Figure 4.3. For $\eta = 0.5\pi$, the curves show duobinary FSK to be about 4.5 dB inferior to FSK. However, in the calculation of the curves, the noise bandwidth, B, was assumed to be the same for both the schemes. Duobinary FSK has a 99% channel bandwidth of 0.92/T, compared to 1.18/T for FSK. Thus the pre-detection filter for duobinary FSK would be narrower than for conventional FSK. This gives duobinary FSK an advantage of 1.1 dB in pre-detection SNR. Thus duobinary FSK is approximately 3.4 dB inferior to FSK.

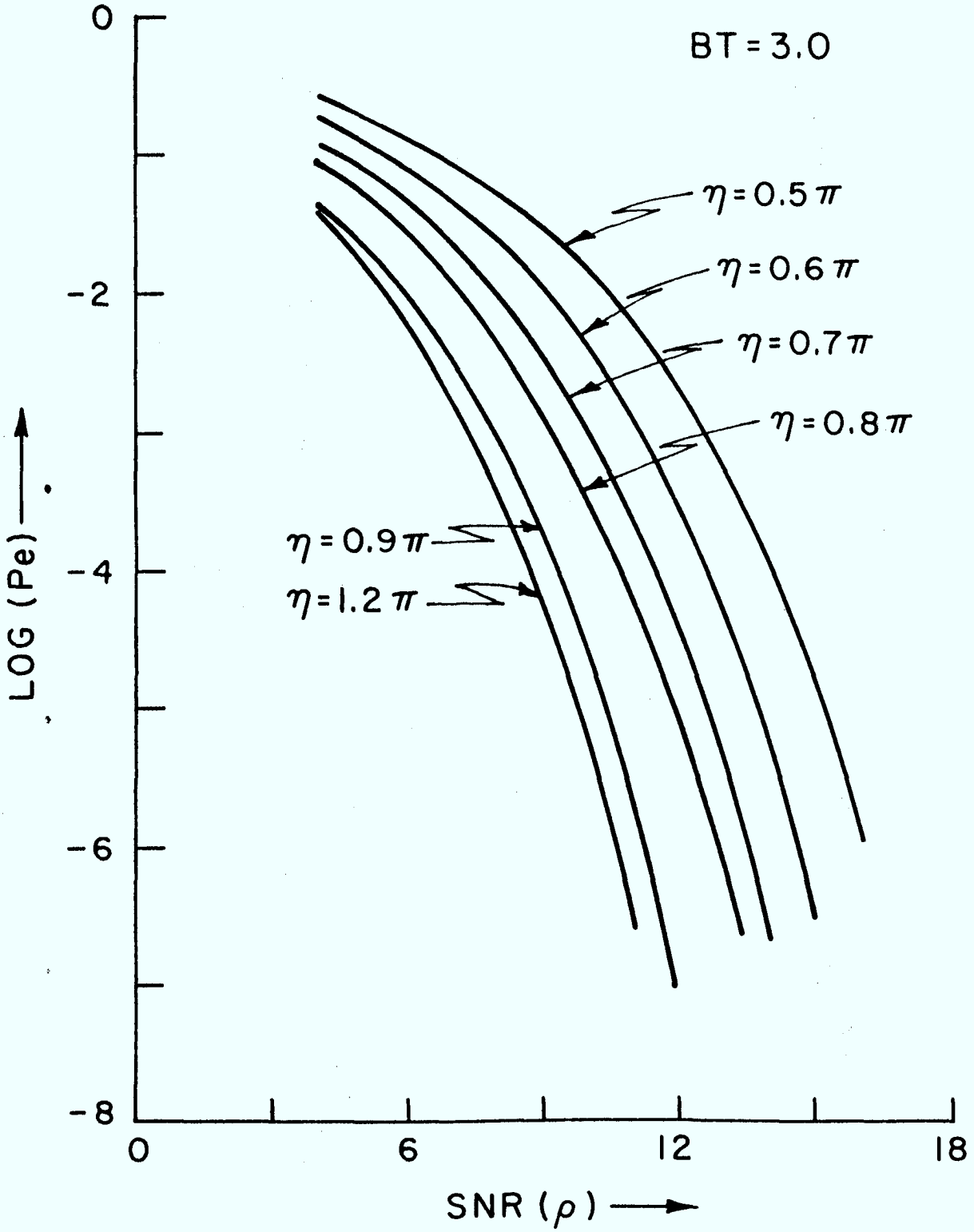


Figure 4.3 Probability of Error for Duobinary FSK with Discriminator Detection

It is well known that the continuous phase in digital FM imposes intersymbol interference. Correlative encoding of the baseband signal introduces an additional intersymbol influence. The discriminator detector makes decisions based on the observation during a single bit interval. Better performance can be obtained by observing over a longer period of time. Such detectors are considered in the next section.

4.2 Optimum Detection

The phase continuity in digital FM introduces intersymbol influence. Therefore, performance gain can be obtained by observing the signal over several bit intervals and making a decision after a delay. This was observed by several authors in the early 70's [12, 43-46]. It is known that at high SNR, a tight union bound can be obtained on the performance of a digital FM receiver [48, 49]. The behaviour of this bound is dominated by the minimum Euclidean distance [57]. Aulin, Rydbeck and Sundberg have computed the minimum Euclidean distance for various correlative encoded FM schemes. The results are given in a number of technical reports [52-56]. They have also carried out an exhaustive investigation of the minimum Euclidean distance for a first order encoding polynomial.

In this section, the derivation of the upper bound on the error performance of a receiver, which makes decisions after a delay, is summarized [48]. The validity of the assumptions made in the derivation are discussed. It is

pointed out that reciprocal encoding polynomials have identical minimum Euclidean distance and so the work in exhaustive studies can be reduced. In Chapter 3, an exhaustive investigation of the band occupancy of correlative encoded digital FM encoded by a second order polynomial was carried out. An exhaustive investigation of the minimum Euclidean distance for the same signals is carried out here. The results on the band occupancy and the minimum Euclidean distance are combined in Chapter 7 to seek encoding polynomials which are spectrally efficient with a small penalty in SNR.

A correlative encoded digital FM signal can be written as

$$x(t) = \cos\left(2\pi f_c t + \int_{LT}^{t} \sum_{k=-\infty}^{\infty} a_k h(u-kT)du + \phi_L\right) \quad \text{for } t > LT \qquad (4.4)$$

In the above equation, ϕ_L is the phase of the modulated signal at $t = LT$. It is assumed that additive white gaussian noise is introduced in the transmission channel. Therefore, the signal at the input of the receiver is given by

$$r(t) = x(t) + n(t) \qquad (4.5)$$

where $n(t)$ is white gaussian noise with two-sided spectral density of N_o watts/Hz. We will consider the detector that observes $r(t)$ over the interval $LT < t \leq (L+N)T$ and makes

the decision on a_L to minimize the probability of error. The investigation is restricted to coherent detection, which assumes perfect knowledge of ϕ_L. In non-coherent detectors, ϕ_L is treated as a random variable, usually with uniform distribution over the interval $[-\pi,\pi]$. Non-coherent detectors have been analyzed for conventional FSK schemes [44, 48, 49].

The signal x(t), over the interval $LT < t \leq (L+N)T$, depends on the data sequence a_L, $a_{L+1} \cdots a_{L+(N-1)}$ and the phase ϕ_L. If the baseband pulse, h(t), in equation (4.4) extends over m bit intervals, x(t) depends also on the (m-1) data symbols prior to t = LT, that is a_{L-1}, $a_{L-2} \cdots$, $a_{L-(m-1)}$. In the analysis perfect knowledge of ϕ_L and the data symbols a_{L-1}, $a_{L-2} \cdots a_{L-(m-1)}$ is assumed. The validity of this assumption is discussed later. The signal, x(t), can assume any one of 2^N possible waveforms depending on the data sequence in the interval $LT < t \leq (L+N)T$. Let each of these waveforms be denoted by $s(t, a_L, A_k)$, where A_k represents a particular data sequence a_{L+1}, $a_{L+2} \cdots$, $a_{L+(N-1)}$. The receiver observes $s(t, a_L, A_k)$ in the presence of noise and has to make a decision on the data symbol a_L. The problem posed in this format is treated in [47] as a composite hypothesis problem. The solution to the composite hypothesis problem is known to be a maximum likelihood receiver. Using the arguments of Osborne and Luntz [48], the likelihood ratio, ℓ, can be reduced to

$$\ell = \frac{\sum_{k=1}^{n} \exp\left(\frac{2}{N_o} \int_{LT}^{(L+N)T} r(t)s(t,1,A_k)dt\right)}{\sum_{k=1}^{n} \exp\left(\frac{2}{N_o} \int_{LT}^{(L+N)T} r(t)s(t,-1,A_k)dt\right)} \tag{4.6}$$

The decision rule for the detector, that yields a minimum probability of error, is given by

$$\ell \underset{-1}{\overset{1}{\gtrless}} 1 \tag{4.7}$$

4.2.1 Bound on the Performance with Optimum Detection

The exact performance of the detector given by the decision rule (4.7) is difficult to compute. However, tight performance bounds can be obtained. Osborne and Luntz [48] have obtained two bounds for conventional FSK modulation. One bound is tight for low SNR, whereas the other is tight for high SNR. Only the bound for high SNR is considered here. Let

$$\xi_{a_L,k} = \frac{2}{N_o} \int_{LT}^{(L+N)T} r(t)s(t,a_L,A_k)dt \tag{4.8}$$

Then, the likelihood ratio is given by

$$\ell = \frac{\sum_{k=1}^{n} \exp(\xi_{1,k})}{\sum_{k=1}^{n} \exp(\xi_{-1,k})} \tag{4.9}$$

For large SNR

$$\sum_{k=1}^{n} \exp(\xi_{1,k}) \simeq \exp(\xi_{1,p}) \tag{4.10}$$

and

$$\sum_{k=1}^{n} \exp(\xi_{-1,k}) \simeq \exp(\xi_{-1,q}) \tag{4.11}$$

where $\xi_{1,p}$ and $\xi_{-1,q}$ are the largest of $\xi_{1,k}$ and $\xi_{-1,k}$, $k=1,2,\ldots n$ respectively. With the above approximation, the decision rule (4.7) reduces to

$$\max_k [\xi_{1,k}] \underset{-1}{\overset{1}{\gtrless}} \max_k [\xi_{-1,k}] \tag{4.12}$$

For high SNR, a tight upper bound can be obtained on the probability of error of the above receiver, using the union bound [48, 49].

$$P_e < \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{n} Pr(\xi_{1,j} < \xi_{-1,k} \mid a_L=1, A_j) \tag{4.13}$$

The variables $\xi_{a_L,k}$ are gaussian random variables. Therefore,

the $Pr(\xi_{1,j} < \xi_{-1,k} \mid a_L=1, A_j)$ can be expressed in terms of the Euclidean distance, $d_{j,k}$, between the signals $s(t,1,A_j)$ and $s(t,-1,A_k)$.

$$d_{j,k}^2 = \int_{LT}^{(L+N)T} \{s(t,1,A_j) - s(t,-1,A_k)\}^2 \, dt \qquad (4.14)$$

and

$$Pr[\xi_{1,j} < \xi_{-1,k} \mid a_L=1, A_j) = Q\left[\frac{d_{j,k}}{\sqrt{N_o}}\right] \qquad (4.15)$$

where, Q is the error function given by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{u^2}{2}} \, du \qquad (4.16)$$

Therefore, P_e in equation (4.13) can be written as

$$P_e < \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{n} Q\left[\frac{d_{j,k}}{\sqrt{N_o}}\right] \qquad (4.17)$$

Let d_{min} be the smallest Euclidean distance of all $d_{j,k}$, that is

$$d_{min} = \min_{j,k} [d_{j,k}] \qquad (4.18)$$

For large SNR, the behaviour of the upper bound in equation (4.17) is dominated by the minimum Euclidean distance, d_{min} [57].

The weakness in the application of the above upper bound is the assumption that ϕ_I and the data symbols a_{L-1}, $a_{L-2}, \cdots, a_{L-(m-1)}$ are known. This knowledge supposes that the data symbols prior to t = LT are detected correctly. Thus, P_e is the probability of an initial error. An initial error could well result in a run of errors. In that case, P_e would be the probability that a run of errors will occur. This phenomenon of error propagation is difficult to analyze analytically. Anderson and Taylor [51] have discussed this problem for a somewhat similar type of modulation scheme known as "Multi-h Codes". They have used a bound on P_e based on the minimum Euclidean distance, and have simulated the detector using the Viterbi algorithm. For error rates below 10^{-5}, the performance of the implemented detector was within 0.2 dB of the error bound. A similar behaviour could be expected for a correlative encoded digital FM detector at high SNR.

To summarize the above results, an estimate on the error performance of a correlative encoded digital FM scheme can be obtained in terms of the minimum Euclidean distance.

4.2.2 Computation of the Minimum Euclidean Distance

The Euclidean distance $d_{j,k}$ is given by equation (4.14). If the carrier frequency, f_c, is large compared to the bit rate, it can be shown that

$$d_{j,k}^2 = NT - \int_{LT}^{(L+N)T} \cos\left(\int_{LT}^{t} \sum_{i=-\infty}^{\infty} (b_i - c_i) h(u-iT)\,du \right) dt \qquad (4.19)$$

The data symbols b_i and c_i correspond to data sequences of $s(t,1,A_j)$ and $s(t,-1,A_k)$ respectively. From the above equation it is clear that the minimum Euclidean distance depends neither on the initial phase, ϕ_L, nor the data symbols $a_{L-1}, a_{L-2}, \cdots, a_{L-(m-1)}$. Let a sequence of symbols e_i be defined as

$$e_i = b_i - c_i \tag{4.20}$$

Since the data symbols b_i and c_i can take values of ± 1, e_i can be either ± 2 or 0. The Euclidean distance can be written as

$$d^2_{j,k} = NT - \int_{LT}^{(L+N)T} \cos\left(\int_{LT}^{t} \sum_{i=-\infty}^{\infty} e_i h(u-iT)du\right)dt \tag{4.21}$$

To compute the minimum Euclidean distance, d_{min}, set $e_L = 2$ and find $d_{j,k}$ for all possible sequences $e_{L+1}, e_{L+2}, \cdots, e_{L+(N-1)}$ and choose the smallest Euclidean distance.

The Euclidean distance can be visualized with the aid of a phase tree. A phase tree consists of phase paths for all the possible data sequences. The phase tree for the polynomial $(1+2D+D^2)/4$ with rectangular pulse shaping is shown in Figure 4.4. Phase path merges in the phase tree provide an upper bound on the minimum Euclidean distance. The phase paths corresponding to the data sequences 1-1 1-1 and -111-1, shown in solid line in Figure 4.4, merge after 4 bit intervals.

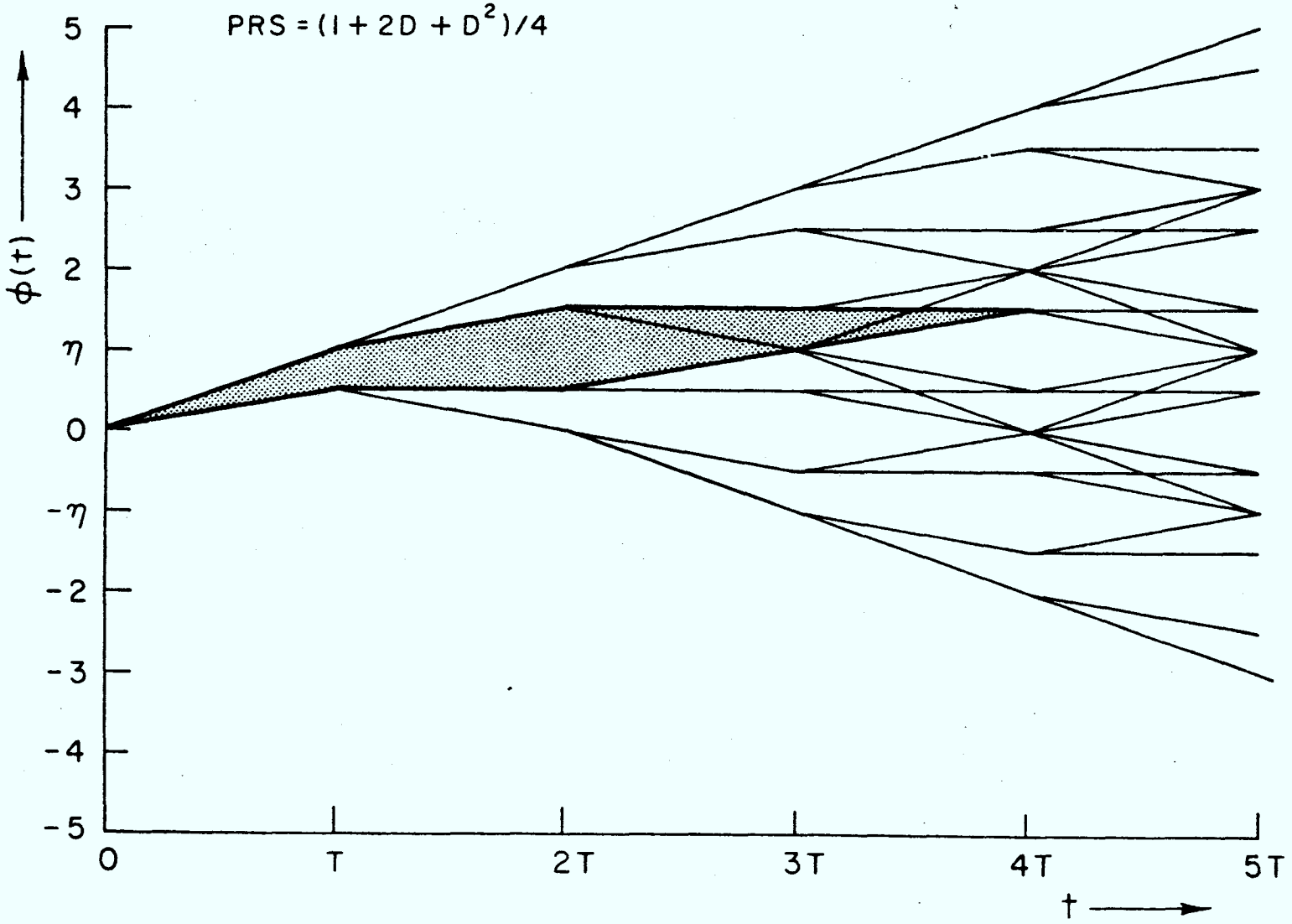


Figure 4.4 Phase Tree for PRS Polynomial $(1+2D+D^2)/4$ with Rectangular s(t) and q=1

Therefore, the Euclidean distance between the signals corresponding to the data sequences -111-1xxx. . . and 1-11-1xxx. . . cannot be increased by observing the signal any longer than 4 bit intervals. Thus the Euclidean distance between signals with merging phase paths provides an upper bound on the minimum Euclidean distance.

The square of the minimum Euclidean distance, d^2_{min}, is shown in Figure 4.5 for various values of N along with the upper bound based on the merge argument for encoding polynomial $(1+2D+D^2)/4$ with rectangular pulse shaping. We next consider an exhaustive investigation of d_{min} for a second order polynomial with rectangular pulse shaping and $\eta = 0.5\pi$. The phase tree for the polynomial $\frac{1}{C}\sum_{i=0}^{m} k_i D^i$ can be obtained by traversing the phase tree for $\frac{1}{C}\sum_{i=0}^{m} k_{m-i} D^i$ in the reverse direction. Therefore, the d_{min} for the polynomial $\frac{1}{C}\sum_{i=0}^{m} k_i D^i$ and its reciprocal $\frac{1}{C}\sum_{i=0}^{m} k_{m-i} D^i$ are identical. The proof for this equivalence is given in Appendix B. It follows from this result that for a second degree polynomial, the encoding coefficients (k_1,k_2) and $(k_1/k_2,1/k_2)$ yield identical d_{min}. Therefore, to exhaustively investigate a second order polynomial it is sufficient to investigate the region $-\infty < k_1 < \infty$ and $-1 \leq k_2 \leq 1$. The minimum Euclidean distance, for N=7, was computed for various values of k_1,k_2 in the region $-5 \leq k_1 \leq 5$ and $-1 \leq k_2 \leq 1$ and the results were applied to other values of k_1,k_2 using the reciprocal polynomial result. The results are presented in Figure 4.6.

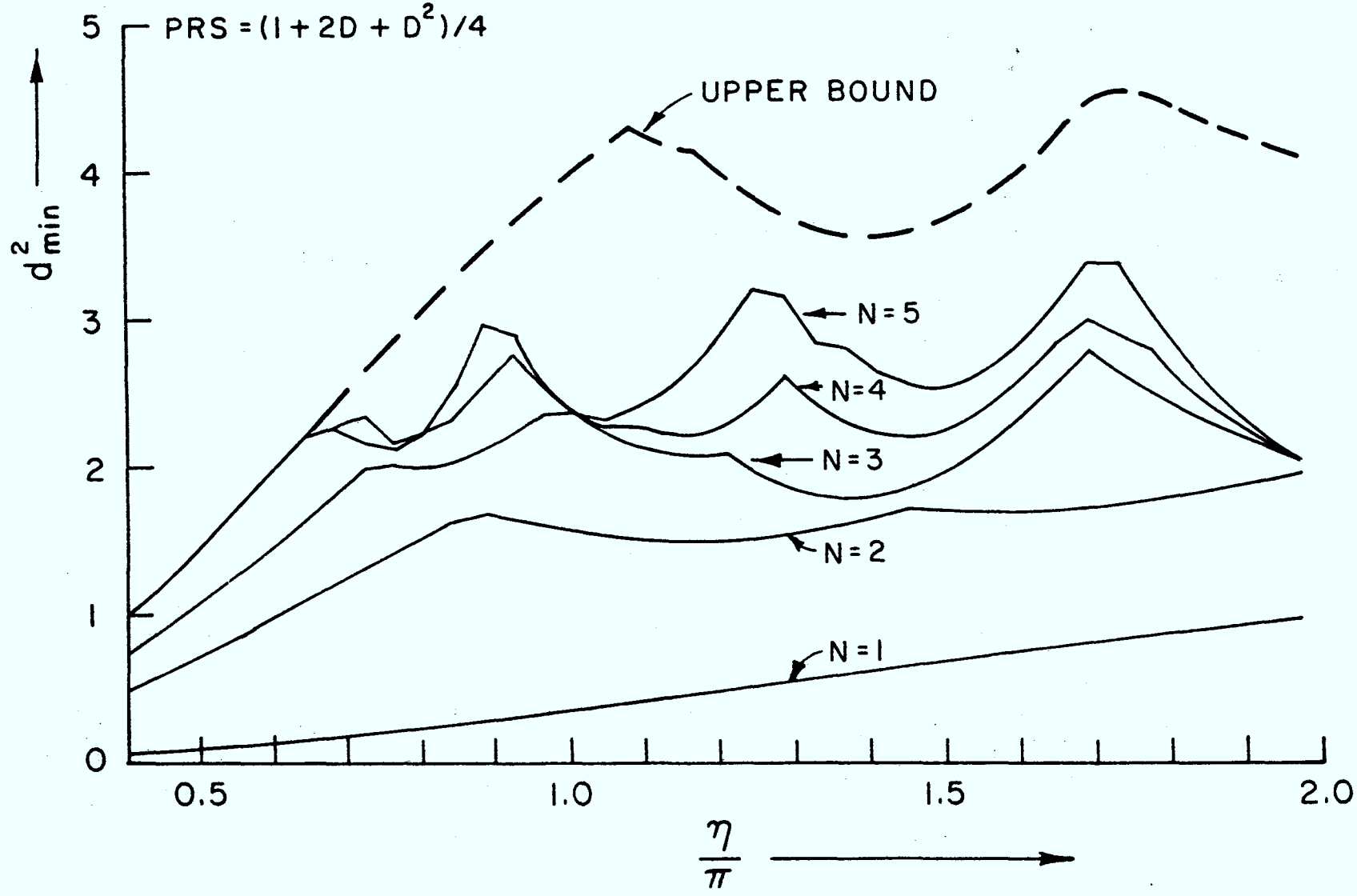


Figure 4.5 d^2_{min} for PRS Polynomial $(1+2D+D^2)/4$ with Rectangular s(t) and q=1

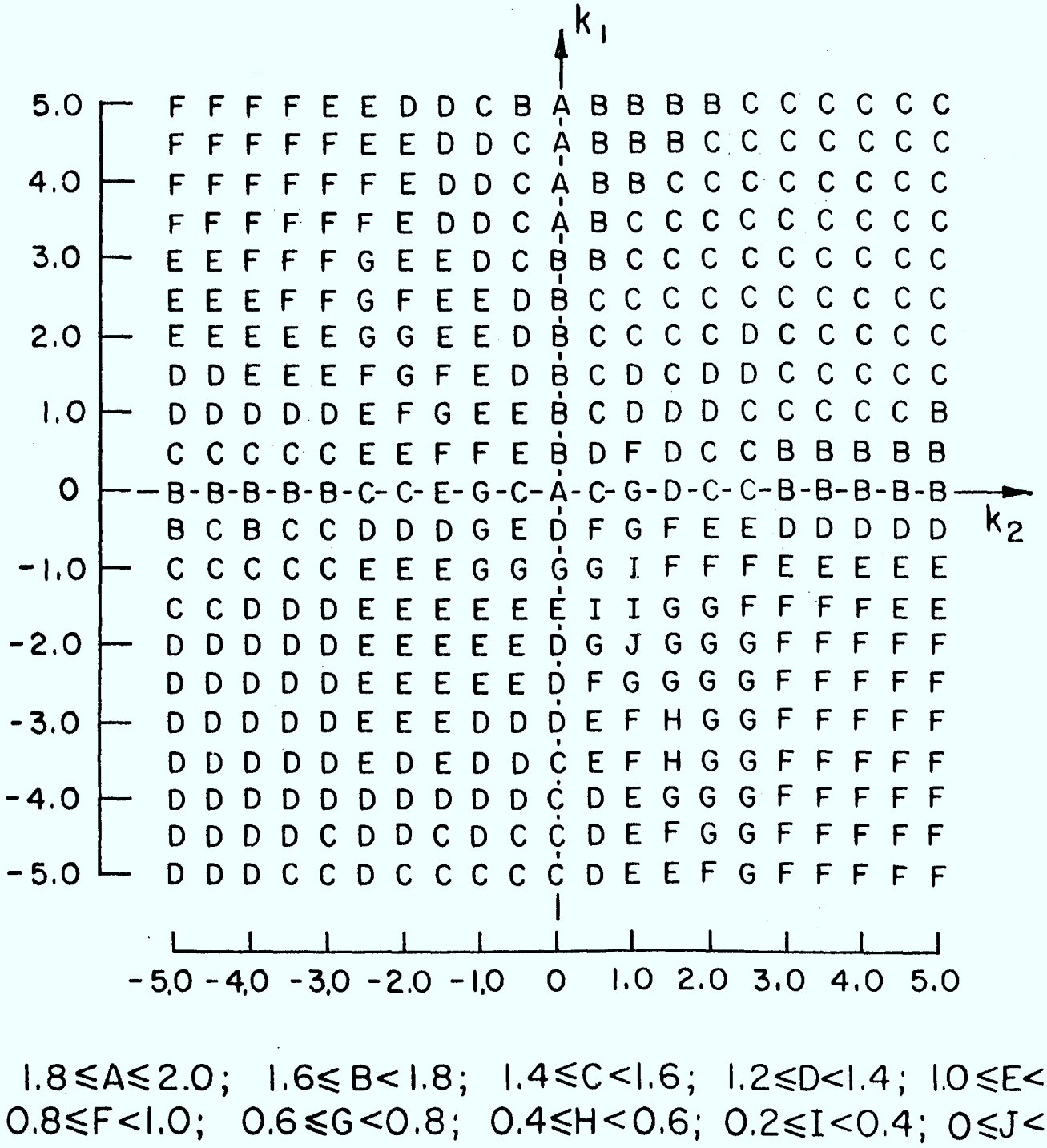


Figure 4.6 d^2_{min} for PRS Polynomial $(1+k_1D+k_2D^2)/C$ with Rectangular s(t), q=1 and $\eta = 0.5\pi$

4.2.3 Complexity of the Receiver

If the detector is implemented according to the decision rule in equation (4.12), the receiver structure will be complicated. The use of Viterbi algorithm to simplify the receiver structure has been suggested by several authors [45,46,48-50]. To demonstrate the complexity of the receiver, duobinary FSK with rectangular pulse shaping and $\eta = 0.5\pi$ will be considered. When the modulation index, η, is such that, η/π is a rational number, the phase tree collapses into a phase trellis. The object of the receiver is to find the path through the trellis which most closely represents the received signal. The Viterbi algorithm is ideally suited for this task. Details can be obtained from [46].

The phase trellis for the duobinary FSK, with $\eta = 0.5\pi$, is shown in Figure 4.7. It has 4 nodes and 4x3 = 12 branches. The detector maintains 4 path histories and updates them every bit interval after finding the 12 branch metrics. Let the branch metrics be denoted by b_{mn}, where $m(m=0,\pm1,2)$ refers to a node and $n(n=0,\pm1)$ refers to a branch emanating from the node. If $r(t)$ is the received signal, the branch metric is given by

$$b_{mn} = \int_0^T r(t)\cos\left(2\pi f_c t + \frac{n\pi t}{2T} + \frac{\pi m}{2}\right) dt \qquad (4.22)$$

It would appear that 12 correlators are required to implement the detector. However, using trigonometric identities b_{mn} can be written as

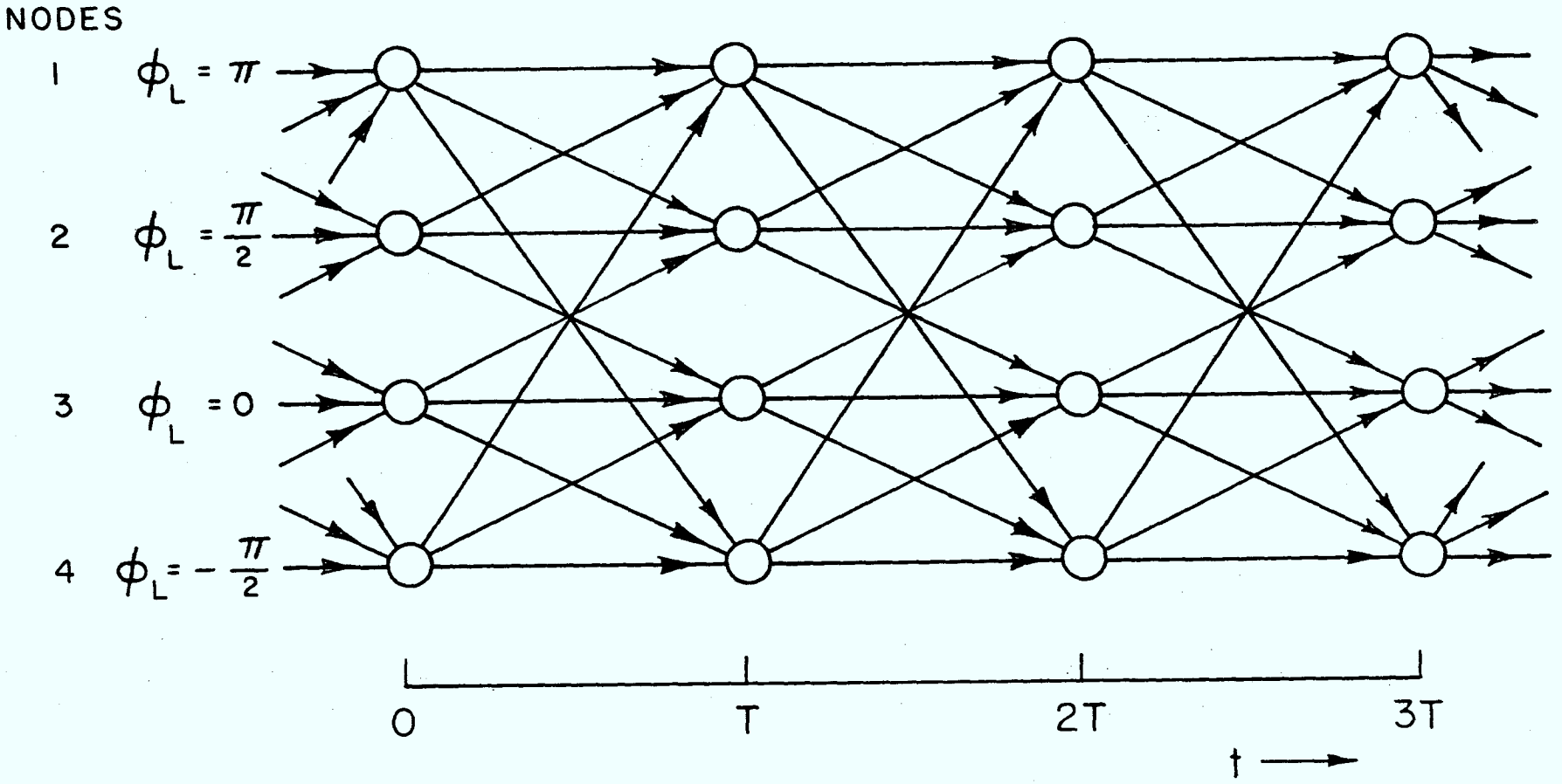


Figure 4.7 Phase Trellis for PRS Polynomial (1+D)/2 with Rectangular s(t), q=1 and $\eta = 0.5\pi$

$$b_{mn} = X_n \cos\frac{\pi m}{2} + Y_n \sin\frac{\pi m}{2} \qquad (4.23)$$

where

$$x_n = \int_0^T r(t)\cos(2\pi f_c t + \frac{n\pi t}{2T})dt \qquad (4.24)$$

and

$$Y_n = -\int_0^T r(t)\sin(2\pi f_c t + \frac{n\pi t}{2T})dt \qquad (4.25)$$

Equation (4.23) shows that all the branch metrics can be obtained with 6 correlators. Schonoff et. al.[50] show that if f_c is an integral multiple of the bit rate, X_n and Y_n can be obtained by hetrodyning the output of a single correlator by $\cos 2\pi f_c t$ and $\sin 2\pi f_c t$. Thus, the detector for duobinary FSK requires 3 correlators and subsequent software. Based on similar arguments it can be shown that the polynomials $(1+D+D^2)/3$ and $(1+2D+D^2)/4$ would require 4 and 5 correlators respectively.

4.3 Summary

A preliminary investigation of error performance of correlative encoded digital FM schemes was carried out. Duobinary FSK appeared approximately 3.4 dB inferior to conventional FSK with discriminator detection.

The optimum detector for correlative encoded digital

FM observes the signal over several bit intervals before making the decision. It is known that a bound on the error performance of such a detector can be obtained from the minimum Euclidean distance. An exhaustive investigation of the minimum Euclidean distance for a second order encoding polynomial was carried out for $\eta = 0.5\pi$. It was pointed out that any encoding polynomial and its recriprocal possess identical d_{min}.

A method of implementing the optimum detector using Viterbi algorithm was discussed briefly. The detectors for encoding polynomials 1, $(1+D)/2$, $(1+D+D^2)/3$ and $(1+2D+D^2)/4$ require 2,3,4 and 5 correlators respectively. The complexity of the subsequent software depends on the total number of nodes and branches in the trellis. At low bit rates this will pose no problem. However, at high rates the implementation of these detectors may be unrealistic at this time.

Chapter 5

OPTIMUM PULSE SHAPING IN MSK-TYPE SIGNALS

MSK-type signals can be defined as angle modulated signals, in which the baseband pulse h(t) extends over only one bit interval, and the modulation index, η, is equal to 0.5π. Note that angle modulated signals are the same as frequency modulated signals which are allowed unbounded instantaneous frequencies. In general, angle modulated signals are difficult to analyze because they involve a nonlinear transformation of the baseband signal. However, MSK-type signals - a special subclass of angle modulated signals - can be represented by a linear quadrature carrier modulation as shown in Figure 5.1. The linearity makes it possible to analyze and implement them with greater ease.

In this chapter, the baseband pulses in MSK-type signals are optimized to minimize the fraction of out-of-band power. Results are presented for a range of channel bandwidths of practical interest. The spectra of the MSK-type signals obtained through optimization are compared with the spectra of well-known modulations, such as MSK, SFSK and OKQPSK.

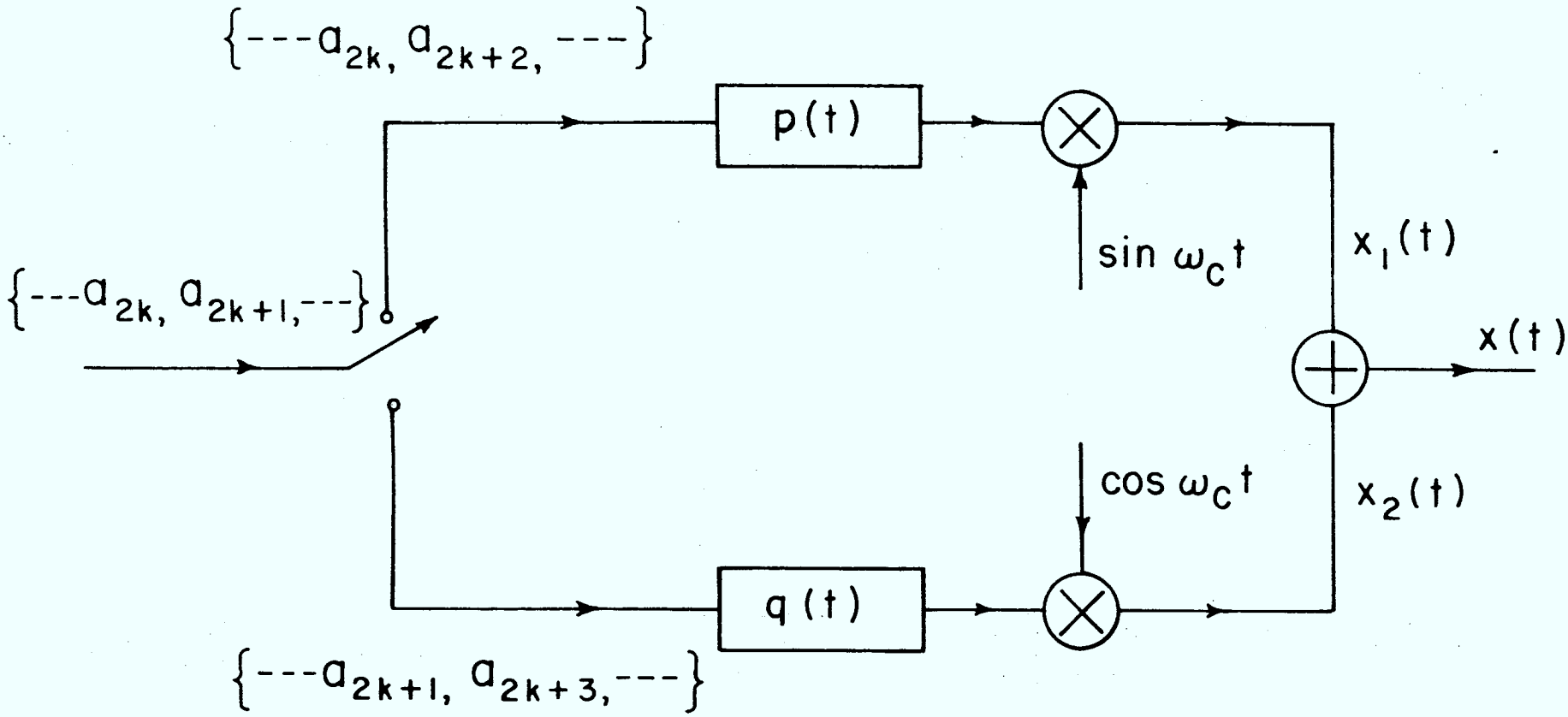


Figure 5.1 Quadrature Generation of MSK-type Signals

5.1 Equivalent Quadrature Carrier Modulation

In general, an angle modulated signal can be written in terms of in-phase and quadrature components. For $\eta = 0.5\pi$, the structure reduces to the linear form as shown in Figure 5.1 [20]. These signals are termed MSK-type signals. The input binary data bits a_k are independent and can assume a value of ± 1. Alternate bits are modulated on I and Q channels as shown in Figure 5.1. The impulse response of the equivalent baseband shaping filters p(t) and q(t) are non-zero over two bit periods, 2T, and the baseband pulses in the I and Q channels are offset or staggered by one bit interval T. The modulated signal x(t) is given by

$$x(t) = x_1(t) + x_2(t) \tag{5.1}$$

where

$$x_1(t) = \sum_{k=-\infty}^{\infty} a_{2k}\, p(t - 2kT) \sin \omega_c t$$

$$x_2(t) = \sum_{k=-\infty}^{\infty} a_{2k+1}\, q(t - (2k+1)T) \cos \omega_c t$$

and

$$p(t) = q(t) = 0 \quad \text{for } t > 2T \quad \text{and} \quad t < 0$$

The envelope e(t) of the transmitted signal x(t) can be written as

$$e(t) = \begin{cases} \sqrt{p^2(t - 2kT) + q^2(t - (2k - 1)T)} & 2kT \leq t < (2k + 1)T \\ \sqrt{p^2(t - 2kT) + q^2(t - (2k + 1)T)} & (2k + 1)T \leq t < (2k + 2)T \end{cases} \quad (5.2)$$

And the phase $\phi(t)$ of $x(t)$ is

$$\phi(t) = \begin{cases} \tan^{-1}[a_{2k}p(t - 2kT) / a_{2k-1}\, q(t-(2k - 1)T)] & 2kT \leq t < (2k + 1)T \\ \tan^{-1}[a_{2k}\, p(t - 2kT) / a_{2k+1}\, q(t-(2k + 1)T)] & (2k + 1)T \leq t < (2k + 2)T \end{cases} \quad (5.3)$$

For the modulated signal $x(t)$ to be a constant-envelope signal of amplitude A, $p(t)$ and $q(t)$ have to satisfy the following constraints

$$\begin{cases} p^2(t) + q^2(t + T) = A^2 & 0 \leq t < T \\ p^2(t) + q^2(t - T) = A^2 & T \leq t < 2T \end{cases} \quad (5.4)$$

Under these constraints the modulated signal $x(t)$ can be written in the form of an angle-modulated signal as

$$x(t) = A \cos(\omega_c t - \phi(t)) \tag{5.5}$$

If x(t) is transmitted over an additive white gaussian channel, the performance does not depend on the pulse shapes p(t) and q(t) but only on their total energy. In the model, the odd bits are transmitted through the I channel and even bits are transmitted through the Q channel. Since it is desirable to have the same error performance for all bits, the power would be divided equally between the I and Q signals in the model. Thus if

$$\int_0^{2T} p^2(t)dt = A^2T \tag{5.6}$$

then

$$\int_0^{2T} q^2(t)dt = A^2T \tag{5.7}$$

The symbols in the input data stream a_k are assumed to be statistically independent. Therefore the power spectral density of x(t) can be expressed as

$$P_x(f) = P_{x_1}(f) + P_{x_2}(f) \tag{5.8}$$

In the above notation $P_x(f)$ denotes the power spectral density of x(t). $P_x(f)$ can be expressed also in terms of the power

spectra of the baseband signals

$$P_x(f) = \frac{1}{2}\{P_p(f-f_c) + P_p(f+f_c) + P_q(f-f_c) + P_q(f+f_c)\} \quad (5.9)$$

If f_c is sufficiently large compared to the bit rate, $P_x(f)$ will be concentrated around $\pm f_c$, and hence it is sufficient to investigate the baseband equivalent spectral density

$$P_{x_b}(f) = P_p(f) + P_q(f) \quad (5.10)$$

The baseband pulses p(t) and q(t) are time-limited and hence their spectra cannot be strictly bandlimited. Therefore, in the suggested modulation adjacent channel interference is inevitable. Let the channel bandwidth be 2W Hz. Then the fraction of in-band power of x(t) in the band (f_c-W, f_c+W) is given by

$$P_{in} = \int_{-W}^{W} P_{x_b}(f)df \; / \; (A^2/2) \quad (5.11)$$

And the fraction of out-of-band power is

$$P_{out} = 1 - P_{in} \quad (5.12)$$

The fraction of out-of-band power, P_{out}, is a good measure of

the adjacent channel interference. Therefore, designing an MSK-type signal involves shaping p(t) and q(t), which satisfy constraints (5.4) and (5.6), such that P_{out} is minimized or P_{in} is maximized for a given channel bandwidth. Rabow [13] has treated this optimization problem with the assumptions that p(t) = q(t) and p(T - t) = p(T + t), and has given the optimum parameters and wave shapes for one particular value of the time-bandwidth product. Boutin et al. [24] have recently reconsidered this problem pointing out Rabow's work. Prabhu [18] has obtained a lower bound on P_{out}, that is not necessarily achievable. The optimization without Rabow's assumptions is considered next.

Maximization of the fractional energy in a given bandwidth for time-limited signals has been studied by Slepian, Pollak and Landau [59,60] with the aid of prolate spheroidal wave functions. The relevant properties of the spheroidal wave functions are presented in Appendix C.

5.2 Optimum Baseband Pulse Shaping

The problem posed in the last section can be summarized as follows:

find p(t) and q(t) to maximize P_{in} for a given W, with the following constraints

$$p(t) = q(t) = 0 \qquad |t| > T \tag{5.13}$$

$$p^2(t) + q^2(t + T) = A^2 \qquad -T \leq t < 0 \tag{5.14}$$

$$p^2(t) + q^2(t - T) = A^2 \qquad 0 \leq t < T \tag{5.15}$$

and

$$\int_{-T}^{T} p^2(t)dt = \int_{-T}^{T} q^2(t)dt = A^2T \tag{5.16}$$

The pulses p(t) and q(t) have been shifted to extend over the interval (-T,T) instead of (0, 2T) as in section 5.1 for the sake of convenience. Since the spheroidal wave functions, $\psi_i(t)$, are a complete set over the subspace of functions strictly time-limited to interval [-T,T], p(t) and q(t) can be expressed in terms of $\psi_i(t)$. Let

$$p(t) = \sum_{n=0}^{\infty} a_n\psi_n(t) \qquad |t| \leq T \tag{5.17}$$

and

$$q(t) = \sum_{n=0}^{\infty} b_n\psi_n(t) \qquad |t| \leq T \tag{5.18}$$

The parameter c for the spheroidal wave functions is given by $c = 2\pi WT$. From the orthogonal properties of $\psi_i(t)$, it follows that

$$\left.\begin{aligned} &\int_{-T}^{T} p^2(t)dt = A^2T = \sum_{n=0}^{\infty} a_n^2\,\lambda_n \\ &\text{and} \\ &\int_{-T}^{T} q^2(t)dt = A^2T = \sum_{n=0}^{\infty} b_n^2\lambda_n \end{aligned}\right\} \tag{5.19}$$

The fraction of in-band power P_{in} can then be written as

$$P_{in} = \{ \sum_{n=0}^{\infty} (a_n^2 \lambda_n^2 + b_n^2 \lambda_n^2) \} / (2A^2T) \qquad (5.20)$$

The coefficients a_n can be expressed in terms of p(t) as

$$a_n \lambda_n = \int_{-T}^{T} p(t) \psi_n(t) dt \qquad (5.21)$$

The coefficients b_n can also be expressed in terms of p(t) as

$$b_n \lambda_n = \int_{-T}^{T} q(t) \psi_n(t) dt$$

$$= \int_{-T}^{0} q(t) \psi_n(t) dt + \int_{0}^{T} q(t) \psi_n(t) dt$$

$$= \int_{-T}^{0} \sqrt{A^2 - p^2(x)} \; \psi_n(x + T) dx$$

$$+ \int_{0}^{T} \sqrt{A^2 - p^2(x)} \; \psi_n(x - T) dx \qquad (5.22)$$

The fraction of in-band power P_{in} can now be expressed as a functional of p(t)

$$P_{in} = \sum_{n=0}^{\infty} \left[\left\{ \int_{-T}^{T} p(t)\psi_n(t)dt \right\}^2 + \left\{ \int_{-T}^{0} \sqrt{A^2 - p^2(t)}\ \psi_n(t+T)dt + \int_{0}^{T} \sqrt{A^2 - p^2(t)}\ \psi_n(t-T)dt \right\}^2 \right] \Big/ (2A^2T) \tag{5.23}$$

Let $f(t)$ be a $p(t)$ which maximizes P_{in}. Then if

$p(t) = f(t) + \varepsilon\, \Delta(t)$

$$\lim_{\varepsilon \to 0} \frac{\partial P_{in}}{\partial \varepsilon} = 0 \qquad \forall\ \Delta(t) \tag{5.24}$$

$$\lim_{\varepsilon \to 0} \frac{\partial P_{in}}{\partial \varepsilon} = \left\{ \sum_{n=0}^{\infty} 2a_n\lambda_n \int_{-T}^{T} \Delta(t)\psi_n(t)dt - \sum_{n=0}^{\infty} 2b_n\lambda_n \left[\int_{-T}^{0} \frac{f(t)\Delta(t)\psi_n(t+T)dt}{\sqrt{A^2 - f^2(t)}} + \int_{0}^{T} \frac{f(t)\Delta(t)\psi_n(t-T)}{\sqrt{A^2 - f^2(t)}}\, dt \right] \right\} \Big/ (2A^2T) = 0 \tag{5.25}$$

Therefore

$$\left.\begin{aligned} \sum_{n=0}^{\infty} a_n\lambda_n\psi_n(t) &= \frac{f(t)}{\sqrt{A^2 - f^2(t)}} \sum_{n=0}^{\infty} b_n\lambda_n\psi_n(t + T) \\ & \qquad -T \leq t < 0 \\ \sum_{n=0}^{\infty} a_n\lambda_n\psi_n(t) &= \frac{f(t)}{\sqrt{A^2 - f^2(t)}} \sum_{n=0}^{\infty} b_n\lambda_n\psi_n(t - T) \\ & \qquad 0 \leq t < T \end{aligned}\right\} \qquad (5.26)$$

In the above equations the coefficients a_n and b_n are functions of f(t) and are given by (5.21, 5.22) with p(t) = f(t). Even though it is not possible to find a closed form expression for f(t) which satisfies equations (5.26), the solution can be obtained iteratively as follows.

i) Assume f(t) to be an arbitrary p(t) which satisfies the constraint (5.16).

ii) Compute the coefficients a_n and b_n using (5.21) and (5.22).

iii) Using the values of a_n and b_n computed in step ii), obtain the f(t) which satisfies the equations (5.26).

iv) Return to step ii). The final solution is obtained when a set of coefficients a_n, b_n and a function f(t) are found that satisfy equations (5.26).

The optimum f(t) was computed iteratively for a range of c of practical interest from c = 1.0 to 10.0, which corresponds to a range of time-bandwidth products of approximately 0.3 to 3.2. In all the cases considered it was found that p(t) = q(t) and also p(t) = p(-t), which Rabow [13] assumed. This implies

$$\left.\begin{aligned} & a_{2n} = b_{2n} \\ \text{and} \\ & a_{2n+1} = b_{2n+1} = 0 \end{aligned}\right\} \quad n = 0, 1, 2 \ldots \qquad (5.27)$$

Under these conditions the equations (5.26) reduce to

$$\sum_{n=0}^{\infty} a_{2n}\lambda_{2n}\psi_{2n}(t) = \frac{f(t)}{\sqrt{A^2 - f^2(t)}} \sum_{n=0}^{\infty} a_{2n}\lambda_{2n}\psi_{2n}(T - t)$$

$$0 \leq t < T/2 \qquad (5.28)$$

The above equation is identical to Rabow's [13] equation (8).

5.3 Numerical Results

The parameters of the optimum pulse shape for a range of time-bandwidth products are presented in this section. For a given $c = 2\pi WT$, the spheroidal wave functions were generated by solving the differential equation (C.9) of Appendix C. The values of χ_i in equation (C.9) were obtained from [61]. This reference also contains the values of λ_i. The differential equation was solved numerically and the solutions normalized such that

$$\int_{-T}^{T} \psi_i^2(t)dt = \lambda_i \tag{5.29}$$

For illustration, $\psi_i(t)$ are shown in Figure 5.2, for i=0,2,4 and c=8.0.

As mentioned in the last section, for all the time-bandwidth products considered, the optimum baseband pulse shapes satisfied the symmetry conditions (5.27). Therefore, we present the results with the assumptions (5.27) without loss of generality. To solve equation (5.28), f(t) was approximated by a truncated series

$$f(t) \simeq \sum_{n=0}^{5} a_{2n}\psi_{2n}(t) \tag{5.30}$$

Equation (5.28) was solved iteratively and the optimum coefficients a_{2n} and f(t) were obtained. The mean square error due to truncation can be expressed as

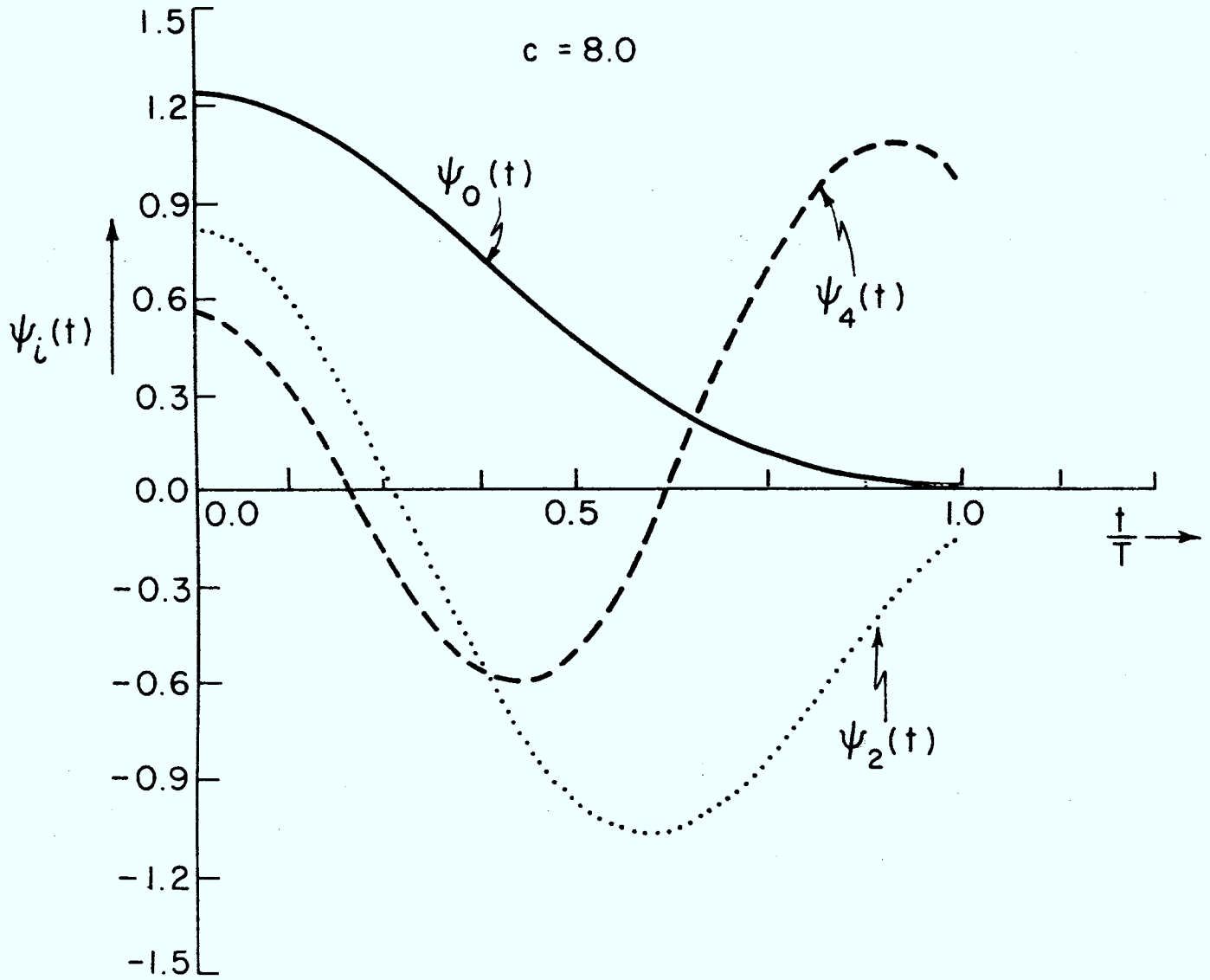


Figure 5.2 Spheroidal Wave Functions, $\psi_i(t)$, for c = 8.0

$$E = \int_{-T}^{T} (f(t) - \sum_{n=0}^{5} a_{2n}\psi_{2n}(t))^2 dt$$

$$= A^2T(1 - \sum_{n=0}^{5} a_{2n}^2\lambda_{2n}) \qquad (5.31)$$

Six terms were found sufficient, and the normalized mean square error (E/A^2T) in all the cases was less than 2.5×10^{-5}.

All the computations were carried out with the assumption that $A^2T = 1$. The optimum parameters for various values of c are listed in Table 5.1. Some of the coefficients a_{2n} are very large in the table. This is due to the normalization (5.29) of $\psi_{2n}(t)$. For example, for c = 1.0 $a_{10} = 4.8 \times 10^9$, but then $\lambda_{10} = 1.03 \times 10^{-25}$ and therefore the contribution of $a_{10}\psi_{10}(t)$ is still very small. In fact, $a_{10}\psi_{10}(t)$ has a total energy of only $a_{10}^2\lambda_{10} = 2.4 \times 10^{-6}$.

The minimum P_{out} computed for various time-bandwidth products is shown as the lower bound in Figure 5.3. This is an achievable lower bound. The lower bound obtained by Prabhu [18] is also shown in the same figure. Prabhu's lower bound is based on the following argument. Since p(t) is a time-limited waveform, the fraction of in-band power of p(t) is always less than or equal to the fraction of in-band power of the zeroth order spheroidal wave function $\psi_0(t)$. Therefore, P_{in} as a function of bandwidth is bounded above by the eigenvalue λ_0 of $\psi_0(t)$. For time-bandwidth products less than 1.0, $\psi_0(t)$ almost satisfies the constraints that p(t) has

$c = 2\pi WT$	Time-band Width Product $(2WT) = \frac{c}{\pi}$	Coefficients						Mean-Square Error E $\times 10^{-5}$	P_{in}	P_{out}
		a_o	a_2	a_4	a_6	a_8	a_{10}			
1.0	0.318	1.32147	-.05820	45.555	10300.0	5.34×10^6	4.82×10^9	0.328	.57250	.42750
2.0	0.637	1.06539	-.00457	4.2421	282.0	36900.0	8.59×10^6	2.14	.88010	.11990
3.0	0.955	1.01227	.00378	.23983	7.2073	592.0	79000.0	1.17	.97570	.02430
4.0	1.273	1.00161	-.02388	-.1972	-4.570	-91.170	-467.70	0.217	.99510	.00490
5.0	1.592	0.99565	-.10015	-.1208	-1.4778	-18.605	-69.309	0.212	.99650	.00350
6.0	1.910	0.98442	-.18077	-.0322	-.17497	-1.2435	.5911	0.076	.99783	.00217
7.0	2.228	0.97541	0.22192	-.0073	-.01279	-.12159	-2.3097	0.025	.99927	.00073
8.0	2.547	0.96842	-.24966	-.0026	-.00093	-.04108	-.8285	0.052	.99981	.00019
9.0	2.865	0.96162	-.27443	-.0028	.000825	-.00433	.09705	0.025	.99995	.00005
10.0	3.183	0.95324	-.30214	.00590	.003421	.00844	.2226	0.188	.99998	.00002

Table 5.1 Parameters of the Optimum Pulse Shape for Various Time-Bandwidth Products

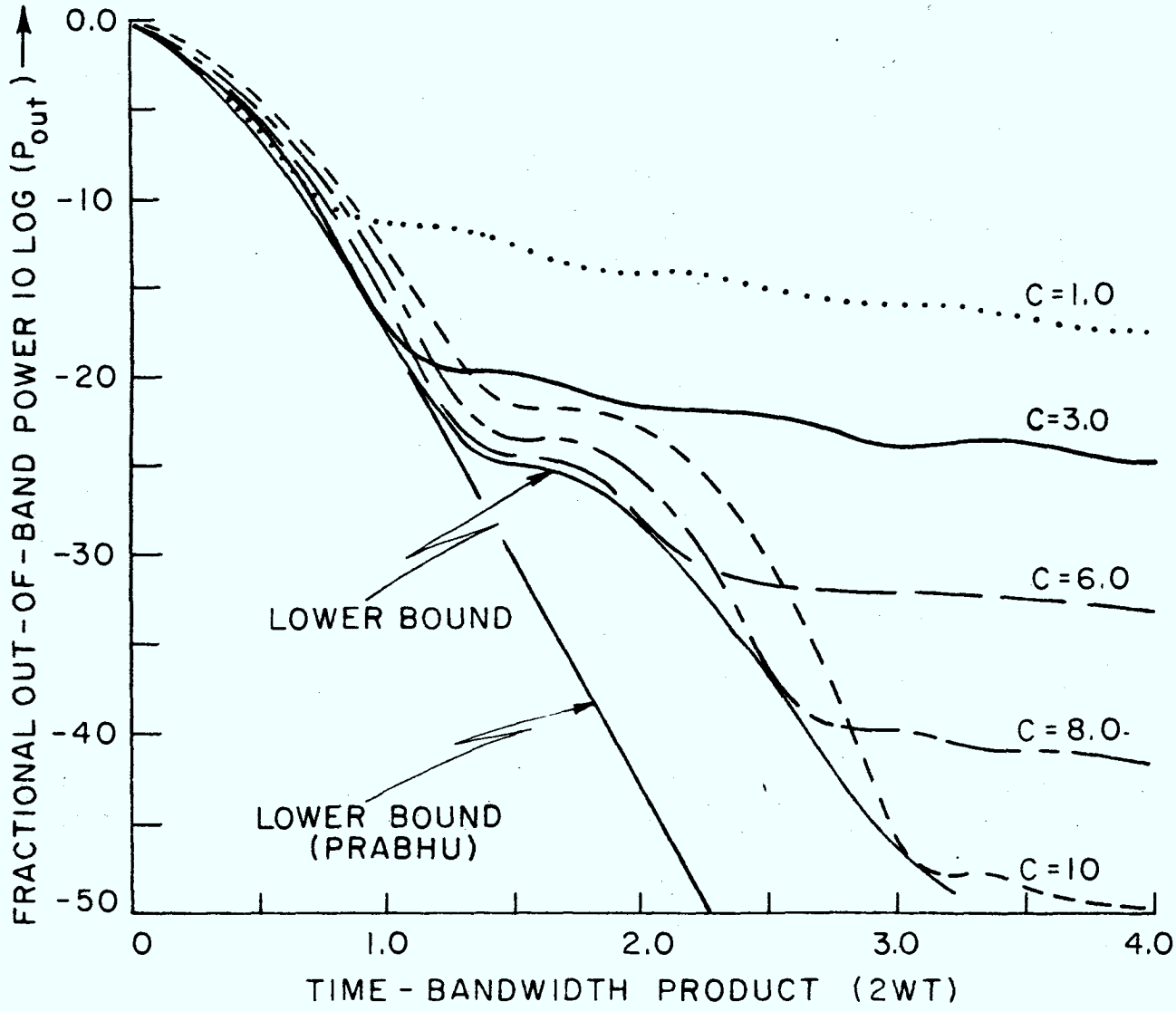


Figure 5.3 Variation of Fraction of Out-of-Band Power for the Optimum Pulse Shapes at c = 1.0, 3.0, 6.0, 8.0 and 10.0

to satisfy to be a baseband pulse of an MSK-type signal. Therefore, Prabhu's bound is the same as the achievable lower bound, for time-bandwidth products less than 1.0.

There is no single pulse shape p(t) which can attain the lower bound at all time-bandwidth products. The variation of the fraction of out-of-band power is plotted in Figure 5.3 for the pulse shapes p(t) optimized for c = 1.0, 6.0, 8.0 and 10.0. As expected they achieve the lower bound at the values of the time-bandwidth product for which they are designed. The spectra of the modulated signal with the optimized p(t) are shown in Figure 5.4. As the pulse is designed for larger values of c, the main lobe in the spectrum becomes wider and the asymptotic behaviour improves.

The optimum pulse shapes for various values of c are shown in Figure 5.5. As $c \to 0$, p(t) approaches a rectangular pulse whereas for large values of c the transition is smooth. The phase trajectories $\phi(t)$ that occur with the optimum pulses are shown in Figure 5.6. As $c \to 0$, the phase makes an abrupt jump of $\pi/4$ radians. As c increases, the magnitude of the jump decreases and for large values of c, $\phi(t)$ has a smooth variation in time.

5.4 Discussion

A number of modulations have been proposed in the past based on the pulse shape p(t). Well-known examples are Minimum Shift Keying (MSK), Offset-Keyed Quadrature Phase Shift Keying (OKQPSK) and Sinusoidal Frequency Shift Keying (SFSK). The variation of P_{out} as a function of bandwidth for

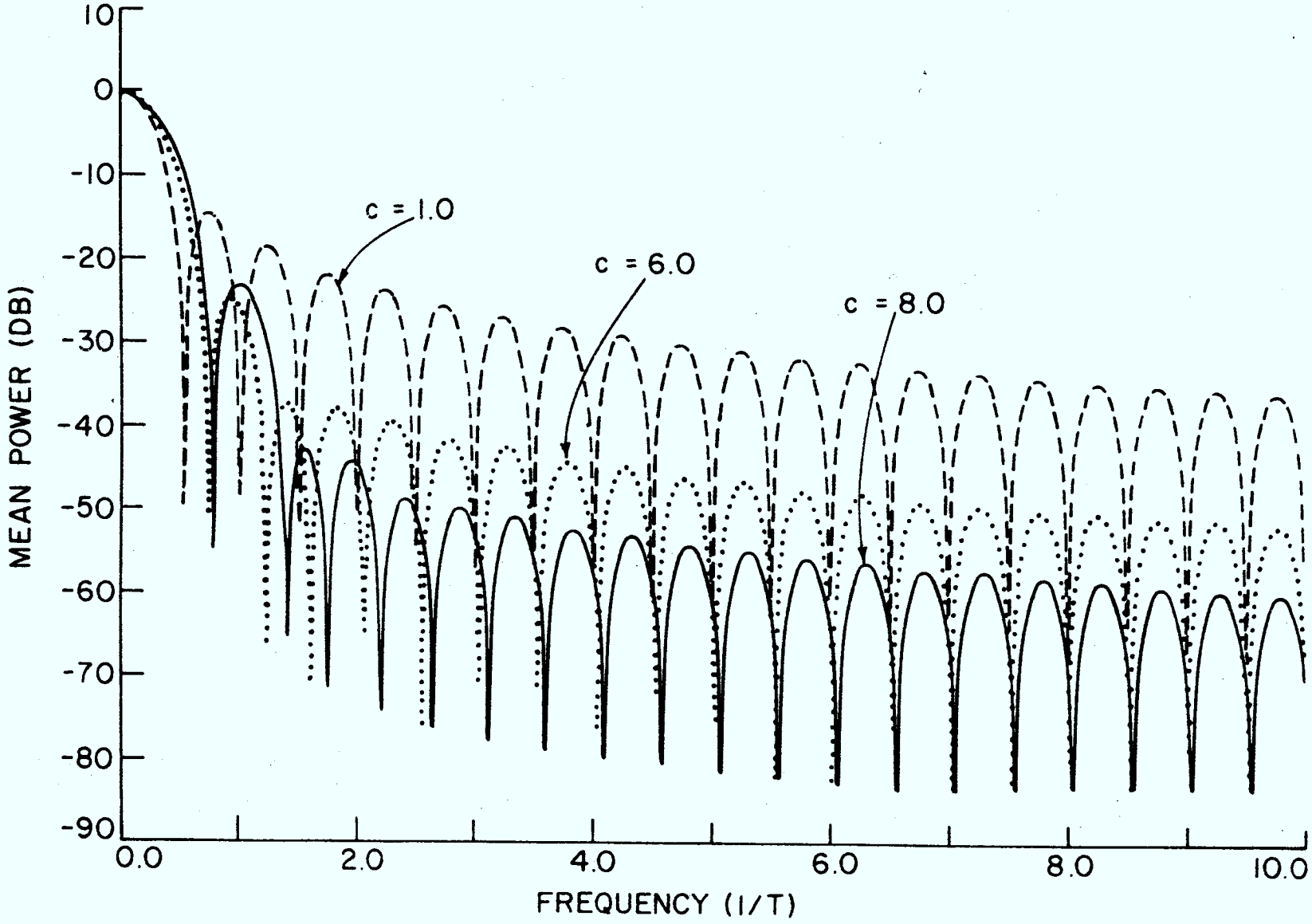


Figure 5.4 Power Spectra for the Optimum Pulse Shapes at c = 1.0, 6.0 and 8.0

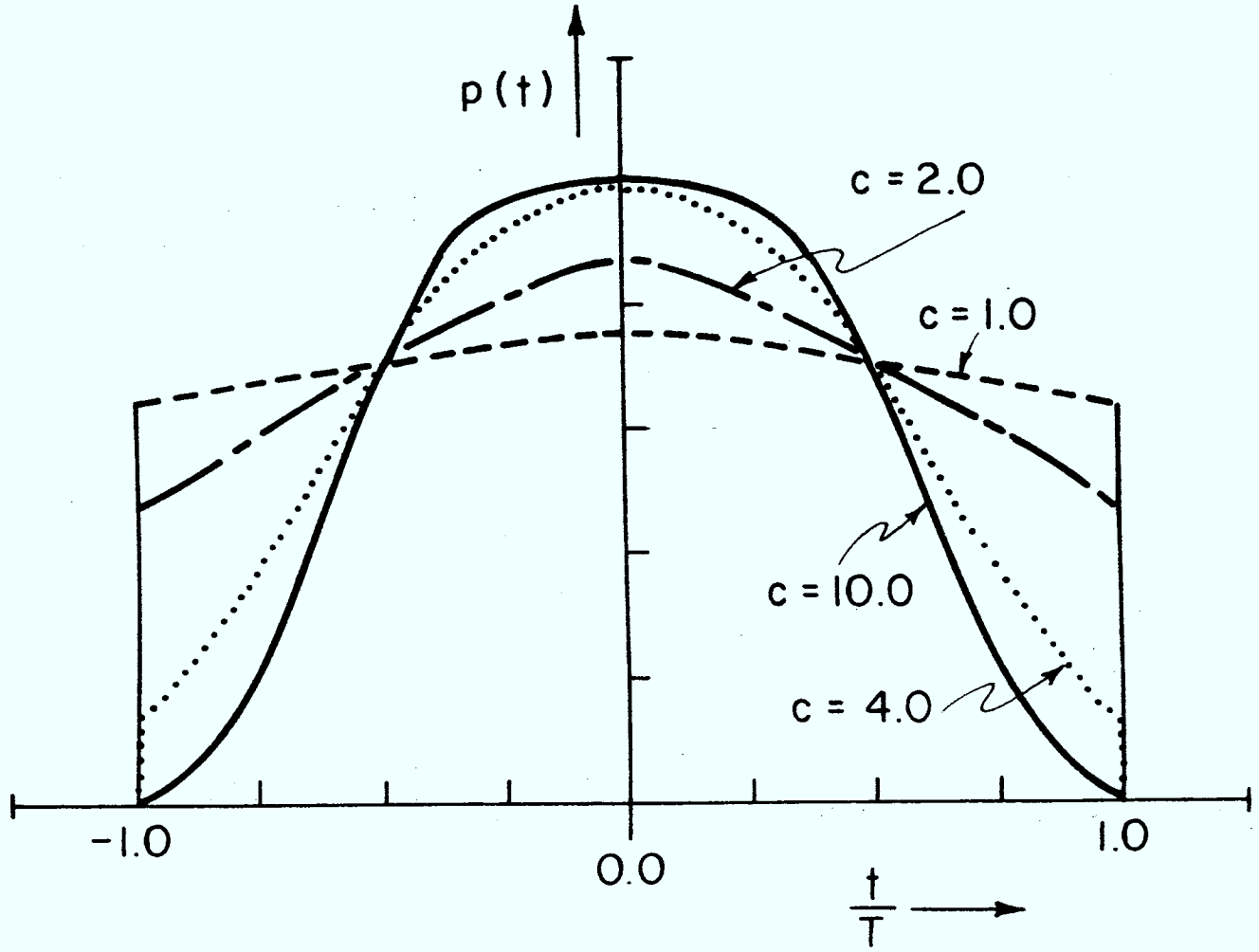


Figure 5.5 Optimum Pulse Shapes at c = 1.0, 2.0, 4.0 and 10.0

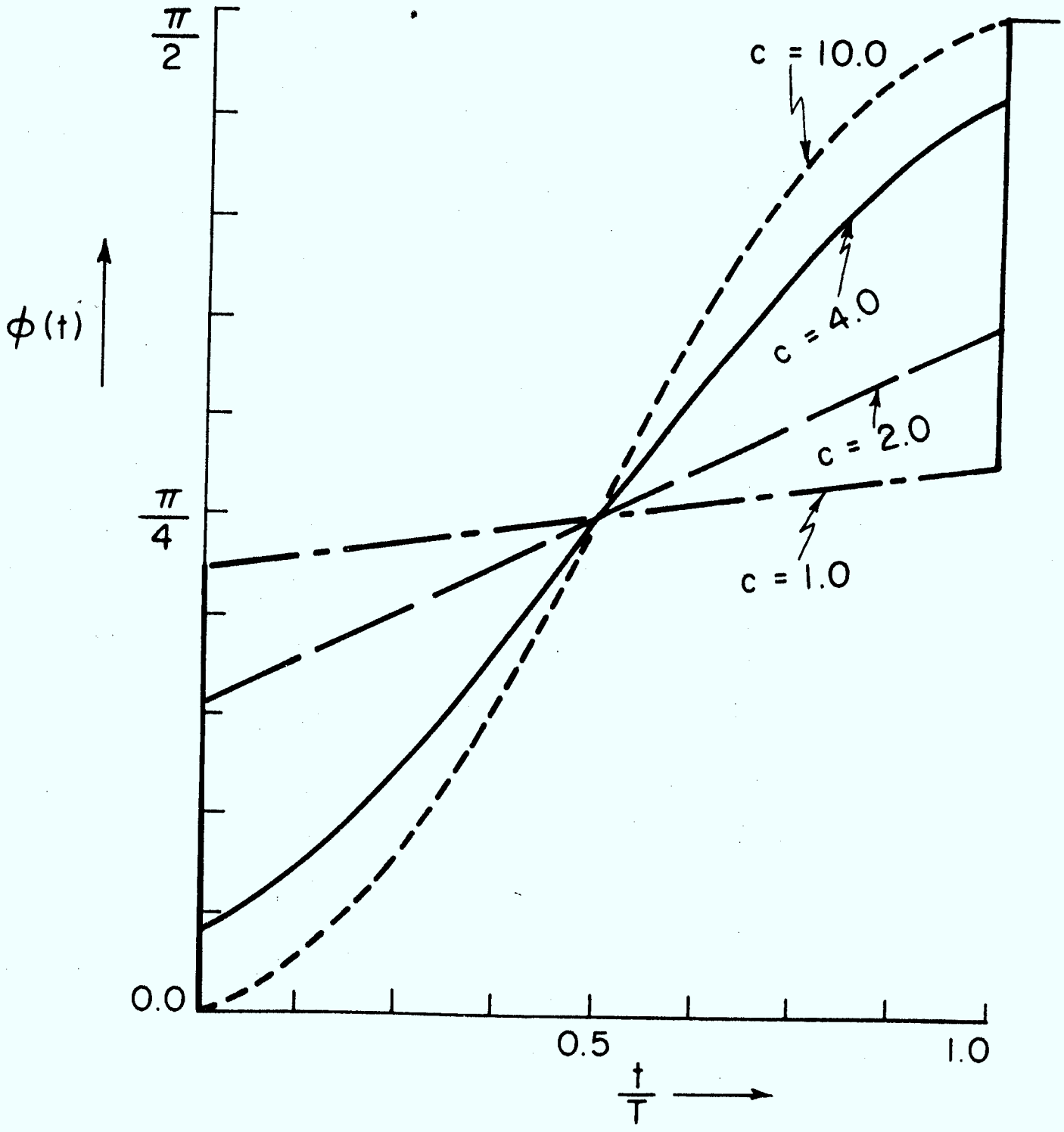


Figure 5.6 Optimum Phase Trajectories at
c = 1.0, 2.0, 4.0 and 10.0

these modulations is shown in Figure 5.7 along with the lower bound. The pulse shape in OKQPSK is given by

$$\begin{aligned} p(t) &= 1 \qquad |t| \leq T \\ &= 0 \qquad \text{elsewhere} \end{aligned} \tag{5.32}$$

This is the optimum pulse shape at c = 0. Therefore, OKQPSK is close to the bound at low time-bandwidth products, and hence can be expected to perform better than other pulse shapes for highly restricted channel-bandwidths. This is consistent with the observation of Gronmeyer and McBride [19] that OKQPSK outperforms MSK for very narrow channel bandwidths.

The fraction of out-of-band power P_{out} for MSK follows the lower bound closely for bandwidths up to 1.2 times the bit rate. The 99% bandwidth for MSK is 1.18/T, whereas the minimum attainable is 1.12/T. The closeness of MSK to optimality was pointed out by Prabhu [18]. SFSK spectrum has good asymptotic properties but its fraction of out-of-band power is not close to the minimum at any channel bandwidth.

The variation of out-of-band power of the signal, modulated by the optimum pulse at c = 8.0, is also shown in Figure 5.7. It closely follows the lower bound up to the time-bandwidth product of 2.5 and is spectrally superior to both MSK and SFSK.

5.5 Summary

A method of shaping the baseband pulses in constant-

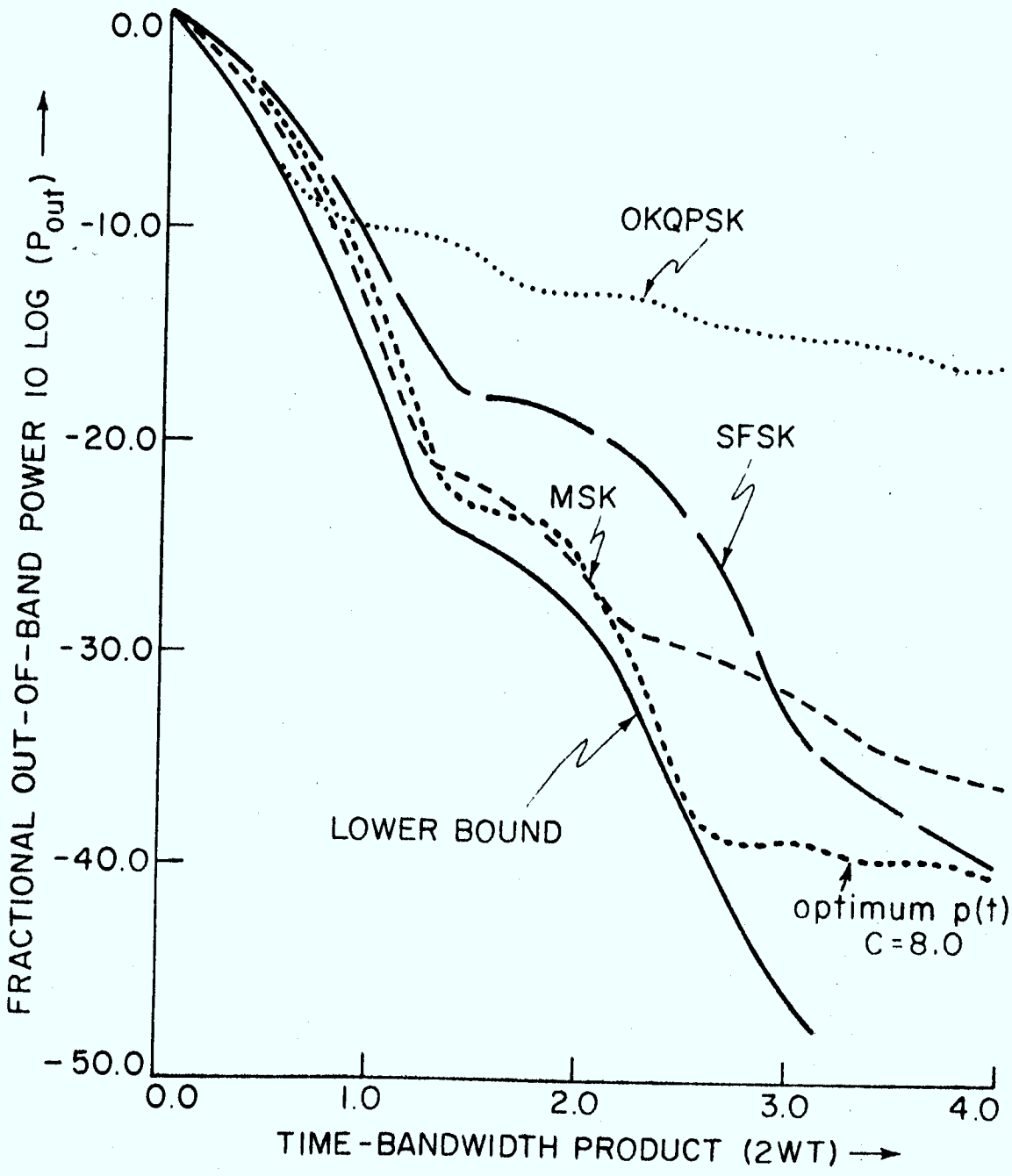


Figure 5.7 Variation of Fraction of Out-of-Band Power for Various Modulation Schemes

envelope, quadrature carrier modulation to maximize the fractional power in a given bandwidth has been presented. It was found that the optimum pulse shapes were identical in both the in-phase and quadrature channels, and also they were symmetric and even. The least possible fraction of out-of-band power, P_{out}, was computed for a range of time-bandwidth products of practical interest. The optimum pulse shapes and their associated properties were examined for a range of values of the parameter c (a quantity proportional to the time-bandwidth product). The variation of the out-of-band fraction of power for the optimum pulse shapes designed for various values of c was plotted. The power spectra of the signal modulated by these pulse shapes were also given. As c increases the main lobe of the spectra becomes wider and the asymptotic behaviour improves. The optimum pulse shape tends to a rectangular waveform as c tends to zero. For large values of c the pulse has a smooth transition. The phase trajectory for the optimum shaping has an abrupt phase jump of $\pi/4$ radians for very small values of c. As c increases the magnitude of the jump decreases and for large values of c, the phase changes gradually with time.

A number of modulations have been proposed in the past based on the baseband pulse shapes in a quadrature carrier modulation. The out-of-band fraction of power for the well known modulations, such as MSK, SFSK and OKQPSK were compared with results from the optimization. The optimization yields pulse shapes which attain a lower out-of-band power. The

improvement is not dramatic. However, the lower bound obtained in this chapter provides a useful limit on what is achievable and a basis for comparison.

Chapter 6

OPTIMUM PULSE SHAPING IN DIGITAL ANGLE MODULATED SIGNALS

Angle modulated signals were defined in Chapter 1, and are given by equation (1.1) or equivalently (1.3). As mentioned previously, the difference between phase continuous conventional digital FM signals and what are called "digital angle modulated signals" here, is that, in angle modulation baseband pulses which cause phase jumps are permitted. In terms of equivalent FM signals, a phase jump would result in an unbounded instantaneous frequency. Phase jumps occur also in digital phase modulated signals, such as PSK, QPSK, etc. However, in angle modulated signals the phase of the carrier can accumulate after every symbol transmitted and cause intersymbol influence.

A technique for obtaining the baseband pulse shapes in angle modulated signals that minimize the fraction of out-of-band power is presented in this chapter. The consideration is restricted to angle modulation, when the baseband pulse, $h(t)$, is symmetric and extends over only one bit interval. When the modulation index, η, is equal to 0.5π, a linear rep-

resentation for the modulation is possible. This special case was treated in detail in Chapter 5. Here, the more general problem of shaping the baseband pulse with an arbitrary modulation index, η, is treated.

It is shown that the spectrum of an angle modulated signal can be approximated by the power spectrum of a time-limited waveform. The time-limited waveform is then expressed as a series of spheroidal wave functions and the waveform is shaped to minimize the out-of-band power. The corresponding baseband pulse is obtained from the time-limited waveform. Examples are calculated for a range of time-bandwidth products and several modulation indices.

6.1 The Spectrum of Digital Angle Modulated Signal

The signal under consideration can be written as

$$x(t) = \cos(2\pi f_c t + \int_{-\infty}^{t} e_o(u)du)$$

$$= \mathrm{Re}[v(t)\ \exp(j2\pi f_c t)] \qquad (6.1)$$

where the integral of $e_o(t)$ may contain step changes, and the phasor $v(t)$ is given by

$$v(t) = \exp(j \int_{-\infty}^{t} e_o(u)du) \qquad (6.2)$$

The spectrum of the modulated signal, $x(t)$, can be obtained using Rowe and Prabhu's method [33], details of which were

discussed in Section 2.2. In their notation, the baseband signal, $e_o(t)$, is expressed as the inner product of a random vector $\underline{a_k}$ and a vector of modulating waveforms $\underline{h(t)}$. Thus

$$e_o(t) = \sum_{k=-\infty}^{\infty} \underline{a_k} \cdot h(t-kT)] \tag{6.3}$$

The vectors $\underline{a_k}$ and $\underline{h(t)}$ are each of dimension two for binary modulation. The analysis will be restricted to modulations with antipodal baseband pulses. Therefore,

$$\underline{a_k} = [a_k^{(1)} \quad a_k^{(2)}] \tag{6.4}$$

and

$$\underline{h(t)} = [h(t) \quad -h(t)] \tag{6.5}$$

The baseband pulse, h(t), is allowed to extend over only one bit interval. Hence

$$h(t) = 0 \text{ for } t < 0 \text{ and } t > T \tag{6.6}$$

The modulated signal, x(t), can be expressed also as

$$x(t) = \cos(2\pi f_c t + \sum_{k=-\infty}^{\infty} \underline{a_k} \cdot g(t-kT)]) \tag{6.7}$$

where

$$\underline{g(t)} = [g(t) \quad -g(t)]$$

and

$$g(t) = \int_{-\infty}^{t} h(u)du$$

For a modulation index, η, the phasor $v(t)$ goes through a total phase shift of $\pm \eta$ radians over each bit interval. Thus

$$\left.\begin{aligned} g(t) &= 0 \qquad \text{for } t \leq 0 \\ &= \eta \qquad \text{for } t \geq T \end{aligned}\right\} \tag{6.8}$$

The function $g(t)$ will be allowed step changes. Using Rowe and Prabhu's method, the spectral density of the phasor, $v(t)$, can be written as

$$P_v(f) = \frac{1}{T} \underline{R(f)} \cdot (A+A^{\dagger}) \cdot R^{*}(f)] \tag{6.9}$$

where

$$A = \frac{1}{2} w_d + \frac{\exp(-j2\pi fT)\ w] \cdot \underline{q(U_K)} \cdot w_d}{1 - \exp(j2\pi fT)\ \underline{w} \cdot q(U_K)]}$$

and

$$R(f)] = \int_{-\infty}^{\infty} \exp(-j2\pi ft)\ r(t)]dt$$

The functions appearing in (6.9) were defined in Section 2.2.

If we assume the binary input symbols to be independent and equiprobable

$$\underline{w} = [1/2 \qquad 1/2] \tag{6.10}$$

and

$$w_d = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \tag{6.11}$$

Since the baseband pulse extends over only one bit period,

$$\left.\begin{aligned} K &= 1 \\ L_K &= 0 \end{aligned}\right\} \tag{6.12}$$

and

$$\left.\begin{aligned} U_K &= T \end{aligned}\right\}$$

The vector $\underline{q(t)}$ is given by

$$\left.\begin{aligned} \underline{q(t)} &= [\exp(jg(t)) \quad \exp(-jg(t))] \qquad 0 \leq t < T \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{elsewhere} \end{aligned}\right\} \tag{6.13}$$

Thus

$$\underline{q(U_K)} = \underline{q(T)} = [e^{j\eta} \qquad e^{-j\eta}] \tag{6.14}$$

The matrix A in equation (6.9) can now be written as

$$A = \frac{1}{4}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{e^{-j2\pi fT}}{4(1-e^{-j2\pi fT}\cos\eta)}\begin{bmatrix} e^{j\eta} & e^{-j\eta} \\ e^{j\eta} & e^{-j\eta} \end{bmatrix} \quad (6.15)$$

and hence

$$A+A^{\dagger} = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2z_4}\begin{bmatrix} z_1 & e^{-j\eta}z_2 \\ e^{j\eta}z_2 & z_3 \end{bmatrix} \quad (6.16)$$

where

$$z_1 = \cos(2\pi fT-\eta)-\cos^2\eta$$

$$z_2 = \cos(2\pi fT)-\cos\eta$$

$$z_3 = \cos(2\pi fT+\eta)-\cos^2\eta$$

$$z_4 = 1 - 2\cos(2\pi fT)\cdot\cos\eta + \cos^2\eta$$

The vector $\underline{R(f)}$ in equation (6.9) is given by

$$\underline{R(f)} = [R(f) \quad R^*(-f)] \quad (6.17)$$

where

$$R(f) = \int_0^T e^{jg(t)}\, e^{-j2\pi ft}\, dt \quad (6.18)$$

because

$$\underline{r(t)} = \underline{q(t)} \quad (6.19)$$

Substituting the above expressions for $\underline{R(f)}$ and $(A+A^{\dagger})$ in equation (6.9), $P_v(f)$ reduces to

$$P_v(f) = \frac{1}{2T} \left\{ R(f)R^*(f)(1+z_1/z_4) \right. \tag{6.20}$$

$$+ R(-f)R(-f)^*(1+z_3/z_4)$$

$$\left. + (z_2/z_4)(e^{-j\eta}R(f)R(-f)+e^{j\eta}R^*(f)R^*(-f)) \right\}$$

To simplify the analysis, only the angle modulated signals with symmetric phase variation are considered. That is

$$h(\tfrac{T}{2} + t) = h(\tfrac{T}{2} - t) \tag{6.21}$$

or equivalently

$$g(\tfrac{T}{2} + t) - \tfrac{\eta}{2} = \tfrac{\eta}{2} - g(\tfrac{T}{2} - t) \tag{6.22}$$

With the above assumption, R(f) in equation (6.18) can be written as

$$R(f) = e^{-j(\pi fT - \frac{\eta}{2})} [R_e(f) + R_o(f)] \tag{6.23}$$

where

$$R_e(f) = \int_{-T/2}^{T/2} \cos z(t)\, e^{-j2\pi ft}\, dt \tag{6.24}$$

$$R_o(f) = \int_{-T/2}^{T/2} j \sin z(t)\, e^{-j2\pi ft}\, dt \qquad (6.25)$$

and

$$z(t) = g(t + \frac{T}{2}) - \frac{\eta}{2} \qquad (6.26)$$

The functions $R_e(f)$ and $R_o(f)$ are even and odd in f respectively. If the expression (6.23) for R(f) is substituted in expression (6.20) for $P_v(f)$, $P_v(f)$ reduces to

$$P_v(f) = \frac{1}{T} \left(R_e(f)X_e(f) + R_o(f)X_o(f)\right)^2 \qquad (6.27)$$

where

$$X_e(f) = 2 \sin \frac{\eta}{2} \cos(\pi fT)/\sqrt{z_4} \qquad (6.28)$$

and

$$X_o(f) = 2 \cos \frac{\eta}{2} \sin(\pi fT)/\sqrt{z_4} \qquad (6.29)$$

6.2 The Spectrum of Equivalent Time-Limited Waveform

The power spectral density of an angle modulated signal, given by equation (6.27), can be approximated by the power spectrum of a time-limited waveform. The functions $X_e(f)$ and $X_o(f)$ appearing in equation (6.27) are shown in Figures 6.1 and 6.2 for various modulation indices. The functions $X_e(f)$ and $X_o(f)$ are even and odd functions in f respectively. They are

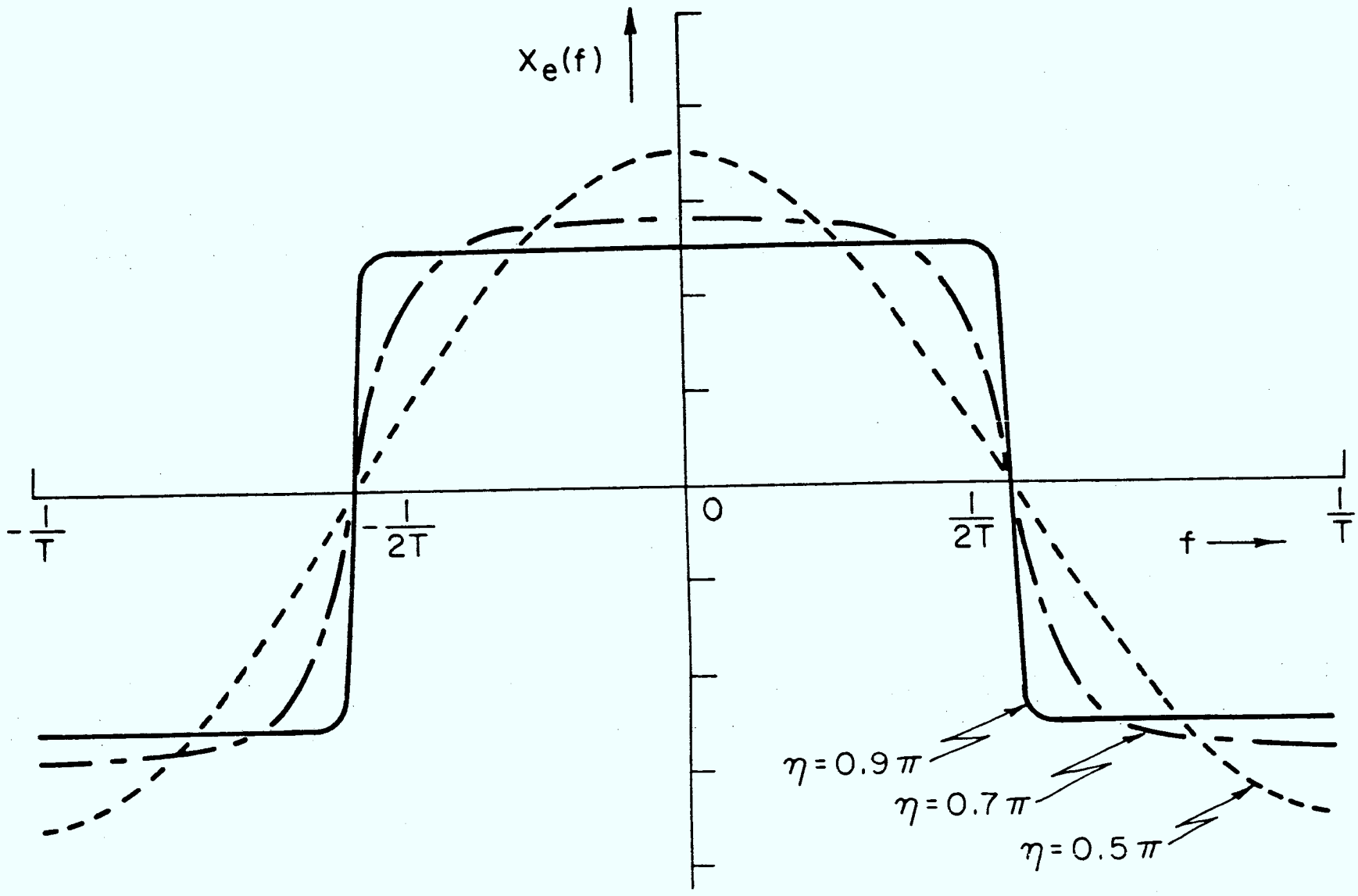


Figure 6.1 The Function $X_e(f)$ for $\eta=0.5\pi, 0.7\pi$ and 0.9π

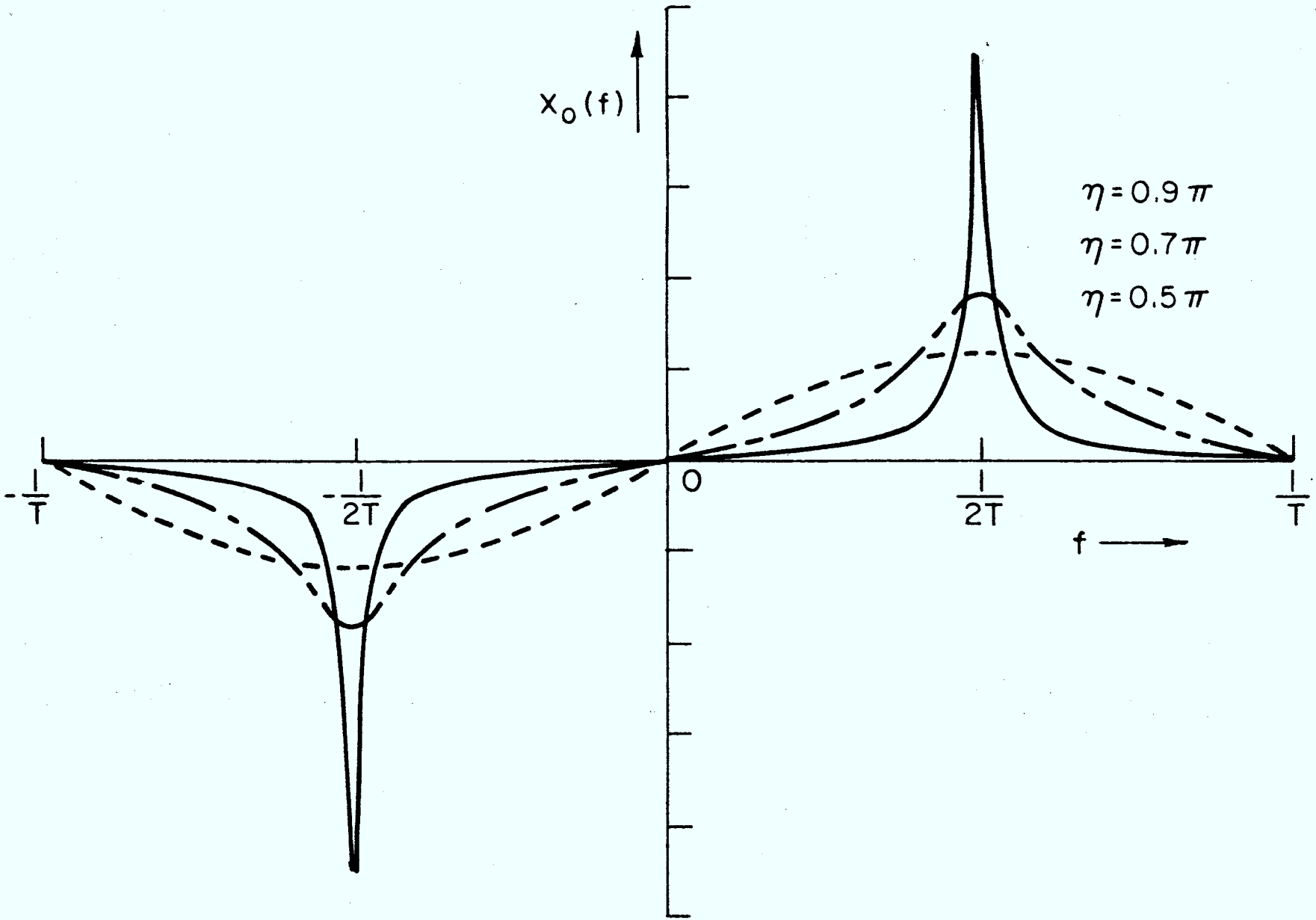


Figure 6.2 The Function $X_c(f)$ for $\eta=0.5\pi, 0.7\pi$ and 0.9π

also periodic with a period 2/T. Therefore, they can be expressed in terms of cosine and sine series as

$$X_e(f) = \sum_{n=1}^{\infty} A_{2n-1} \cos\{\pi fT(2n-1)\} \tag{6.30}$$

and

$$X_o(f) = \sum_{n=1}^{\infty} B_{2n-1} \sin\{\pi fT(2n-1)\} \tag{6.31}$$

Thus, the power spectral density, $P_v(f)$, can be written as

$$P_v(f) = \frac{1}{T}\left\{ \sum_{n=1}^{\infty} A_{2n-1}\cos(\pi fT(2n-1))R_e(f) + \sum_{n=1}^{\infty} B_{2n-1}\sin(\pi fT(2n-1))R_o(f)\right\}^2 \tag{6.32}$$

It can be seen from (6.24) that $R_e(f)$ in the above expression is the Fourier transform of cos z(t), strictly time-limited to the interval $[-\frac{T}{2}, \frac{T}{2}]$.

$$\left.\begin{aligned} R_e(f) &= F\{\cos z(t) \cdot p(t)\} \\ &\text{where} \\ p(t) &= 1 \qquad \text{if } |t| \leq \tfrac{T}{2} \\ &= 0 \qquad \text{elsewhere} \end{aligned}\right\} \tag{6.33}$$

and

$$\cos\{\pi fT(2n-1)\} = F\left\{\frac{1}{2}\left[\delta\left(t+\frac{T(2n-1)}{2}\right)\right.\right.$$

$$\left.\left. + \delta\left(t-\frac{T(2n-1)}{2}\right)\right]\right\} \qquad (6.34)$$

In the above equations, $F\{x(t)\}$ denotes the Fourier transform of $x(t)$ and $\delta(t)$ is the Dirac delta function. Convolving the time-domain waveforms in (6.33) and (6.34), we get

$$R_e(f) \cdot \cos\{\pi fT(2n-1)\}$$

$$= F\left\{\frac{1}{2}\left[\cos z\left(t + \frac{T(2n-1)}{2}\right) \cdot p\left(t + \frac{T(2n-1)}{2}\right)\right]\right.$$

$$\left. + \cos z\left(t-\frac{T(2n-1)}{2}\right) \cdot p\left(t - \frac{T(2n-1)}{2}\right)\right]\right\} \qquad (6.35)$$

Similarly

$$R_o(f) \cdot \sin\{\pi fT(2n-1)\}$$

$$= F\left\{\frac{1}{2}\left[\sin z\left(t + \frac{T(2n-1)}{2}\right) \cdot p\left(t + \frac{T(2n-1)}{2}\right)\right.\right.$$

$$\left.\left. - \sin z\left(t - \frac{T(2n-1)}{2}\right) \cdot p\left(t - \frac{T(2n-1)}{2}\right)\right]\right\} \qquad (6.36)$$

Substituting the expressions (6.35) and (6.36) in (6.32), $P_v(f)$ can be written in terms of the Fourier transform of a waveform $c(t)$. Thus

$$P_v(f) = \frac{1}{T}\,[F\{c(t)\}]^2 \tag{6.37}$$

where

$$\begin{aligned} c(t) = \frac{1}{2}\Bigg[& \sum_{n=1}^{\infty} \{A_{2n-1}\cos z(t + \tfrac{T(2n-1)}{2}) \\ & + B_{2n-1}\sin z(t + \tfrac{T(2n-1)}{2})\}\, p(t + \tfrac{T(2n-1)}{2}) \\ & + \sum_{n=1}^{\infty} \{A_{2n-1}\cos z(t - \tfrac{T(2n-1)}{2}) \\ & - B_{2n-1}\sin z(t - \tfrac{T(2n-1)}{2})\}\, p(t - \tfrac{T(2n-1)}{2})\Bigg] \end{aligned} \tag{6.38}$$

Note that c(t) convolved with itself gives the envelope of the autocorrelation function of the modulated signal x(t). We now introduce a function $d_N(t)$, strictly time-limited to the interval [-NT,NT], given by

$$\left.\begin{aligned} d_N(t) &= c(t) && \text{if } |t| \le NT \\ &= 0 && \text{elsewhere} \end{aligned}\right\} \tag{6.39}$$

The coefficients A_{2n-1} and B_{2n-1} in (6.38) decrease with n as shown in Table 6.1. For $0.5\pi \le \eta \le 0.8\pi$, A_{2n-1} and B_{2n-1} are essentially zero for $n > N$, where N can be seen from Table 6.1. Therefore, c(t) can be approximated by $d_N(t)$.

η	A_{2n-1} coefficients					B_{2n-1} coefficients				
	A_1	A_3	A_5	A_7	A_9	B_1	B_3	B_5	B_7	B_9
0.5π	2.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
0.6π	1.979	-0.282	0.0634	-0.0161	0.00431	1.971	-0.329	0.0781	-0.0204	0.00554
0.7π	1.929	-0.477	0.196	-0.0926	0.0465	1.864	-0.638	0.295	-0.148	0.0774
0.8π	1.870	-0.576	0.310	-0.194	0.131	1.633	-0.838	0.550	-0.386	0.280

Table 6.1 Coefficients A_{2n-1} and B_{2n-1} for $\eta = 0.5\pi, 0.6\pi, 0.7\pi$ and 0.8π

In other words, the spectrum of an angle modulated signal can be approximated by the spectrum of a strictly time-limited waveform $d_N(t)$.

6.3 Optimum Baseband Pulse Shaping

The properties of spheroidal wave functions used in the following optimization are summarized in Appendix C. Let the bandwidth, within which the fraction of power of the modulated signal is to be maximized, be W Hz. Since the spheroidal wave functions, with the parameter $c = 2\pi WNT$, form a complete set in the class of square integrable functions strictly time-limited to the interval $[-NT,NT]$, $d_N(t)$ can be written as

$$d_N(t) = \sum_{m=0}^{\infty} a_{2m}\, \psi_{2m}(t) \qquad \text{for} \quad |t| \leq NT \tag{6.40}$$

Spheroidal wave functions of only even order are included in the above summation because $d_N(t)$ is an even function. From the orthogonal properties of the spheroidal wave functions, it follows that

$$a_{2m}\lambda_{2m} = 2\int_0^{NT} d_N(t)\psi_{2m}(t)dt$$

$$= \sum_{n=1}^{N} \int_{(n-1)T}^{nT} \{A_{2n-1}\cos z(y - \tfrac{T(2n-1)}{2})$$

$$- B_{2n-1} \sin z(y - \tfrac{T(2n-1)}{2})\}\psi_{2m}(y)dy$$

$$= \sum_{n=1}^{N} \int_{-T/2}^{T/2} \{A_{2n-1}\cos\ z(x)$$

$$- B_{2n-1}\sin\ z(x)\}\psi_{2m}(x + \frac{T(2n-1)}{2})dx$$

$$= \int_{-T/2}^{T/2} \{\cos\ z(x) \sum_{n=1}^{N} A_{2n-1}\psi_{2m}\ (x + \frac{T(2n-1)}{2})$$

$$- \sin\ z(x) \sum_{n=1}^{N} B_{2n-1}\psi_{2m}(x + \frac{T(2n-1)}{2})\}dx \qquad (6.41)$$

The function, $z(x)$, is an odd function of x. Therefore

$$a_{2m}\lambda_{2m} = \int_{0}^{T/2} \{\cos\ z(x)\tilde{\psi}_{2m}(x) + \sin z(x)\tilde{\tilde{\psi}}_{2m}(x)\}\ dx \qquad (6.42)$$

where

$$\tilde{\psi}_{2m}(x) = \sum_{n=1}^{N} A_{2n-1}[\psi_{2m}(-x + \frac{T(2n-1)}{2})$$

$$+ \psi_{2m}(x + \frac{T(2n-1)}{2})] \qquad (6.43)$$

and

$$\tilde{\psi}_{2m}(x) = \sum_{n=1}^{N} B_{2n-1}[\psi_{2m}(-x + \frac{T(2n-1)}{2})$$

$$- \psi_{2m}(x + \frac{T(2n-1)}{2})] \qquad (6.44)$$

The total energy of $d_N(t)$ is normalized such that

$$\int_{-\infty}^{\infty} d_N^2(t)dt = \sum_{m=0}^{\infty} a_{2m}^2 \lambda_{2m} = 1 \qquad (6.45)$$

Let P'_{in} be the fraction of power in a bandwidth of W Hz of the waveform $d_N(t)$, and let P'_{out} be the fraction of out-of-band power.

Then

$$P'_{in} = \sum_{m=0}^{\infty} {}^{2}{}_{2m}{}^{2} \lambda_{2m}^{2} \qquad (6.46)$$

and the fraction of out-of-band power is

$$P'_{out} = 1 - P'_{in} \qquad (6.47)$$

The fraction of power within the channel bandwidth, P'_{in}, can now be written as a functional of z(t)

$$P'_{in} = \sum_{m=0}^{\infty} \left\{ \int_{0}^{T/2} (\cos z(x)\tilde{\psi}_{2m}(x) + \sin z(x)\tilde{\tilde{\psi}}_{2m}(x))dx \right\}^2 \qquad (6.48)$$

Let $\bar{z}(x)$ be the function that maximizes P'_{in}. Then, if

$$z(x) = \bar{z}(x) + \varepsilon\Lambda(x)$$

$$\lim_{\varepsilon \to 0} \frac{\partial P'_{in}}{\partial \varepsilon} = 0 \qquad \forall\ \Delta(x) \tag{6.49}$$

$$\lim_{\varepsilon \to 0} \frac{\partial P'_{in}}{\partial \varepsilon} = \sum_{m=0}^{\infty} 2a_{2m}\lambda_{2m} \left\{ \int_0^{T/2} [\sin \bar{z}(x) \cdot \tilde{\psi}_{2m}(x) \right.$$

$$\left. - \cos \bar{z}(x) \cdot \tilde{\tilde{\psi}}_{2m}(x)]\Delta(x)dx \right\} \tag{6.50}$$

Therefore, the optimum $\bar{z}(t)$ is given by

$$\tan \bar{z}(t) \cdot \sum_{m=0}^{\infty} a_{2m}\lambda_{2m}\tilde{\psi}_{2m}(t)$$

$$= \sum_{m=0}^{\infty} a_{2m}\lambda_{2m}\tilde{\tilde{\psi}}_{2m}(t) \qquad 0 \leq t \leq T/2 \tag{6.51}$$

The coefficients a_{2m} in the above equation are functions of $\bar{z}(t)$, given by equation (6.42) with z(t) replaced by $\bar{z}(t)$. As before it is not possible to find a closed form expression for $\bar{z}(t)$, but equation (6.51) can be solved iteratively.

When the modulation index, η, is equal to 0.5π, the modulated signal x(t) is an MSK-type signal and equation (6.51) reduces to equation (5.28) of Chapter 5.

6.4 Numerical Results

The optimum pulse shapes are obtained by iteratively solving (6.51). The spheroidal wave functions required for computation were generated and normalized as explained in Section 5.3 of Chapter 5. To obtain numerical solutions, the infinite series in (6.51) has to be truncated. An estimate on the truncation error can be obtained as follows.

The waveform $d_N(t)$ is normalized in equation (6.45) to have unit energy. Therefore

$$\sum_{m=(M+1)}^{\infty} a_{2m}^2 \lambda_{2m} = (1 - \sum_{m=0}^{M} a_{2m}^2 \lambda_{2m}) \tag{6.52}$$

The eigenvalues, λ_{2m}, are real positive numbers and are ordered in magnitude, that is, $\lambda_0 > \lambda_2 > \lambda_4 > \ldots$ Therefore

$$\sum_{m=(M+1)}^{\infty} a_{2m}^2 \lambda_{2m}^2 \leq \lambda_{2(M+1)}(1 - \sum_{m=0}^{M} a_{2m}^2 \lambda_{2m}) \tag{6.53}$$

Let the quantity to be maximized, P'_{in}, given by equation (6.46), be truncated by

$$P'_{in} \simeq \sum_{m=0}^{M} a_{2m}^2 \lambda_{2m}^2 \tag{6.54}$$

Then, the error due to truncation can be written as

$$\text{Error} = \sum_{m=(M+1)}^{\infty} a_{2m}^2 \lambda_{2m}^2 \leq \lambda_{2(M+1)}(1 - \sum_{m=0}^{M} a_{2m}^2 \lambda_{2m}) \tag{6.55}$$

or more simply

$$\text{Error} \leq \lambda_{2(M+1)} \tag{6.56}$$

The eigenvalue, λ_{2m}, is the fraction of energy of $\psi_{2m}(t)$ within a bandwidth of W Hz. It is pointed out by Slepian and Pollak [59, p.45] that once 2m exceeds $(2c/\pi)$, λ_{2m} fall cff to zero rapidly. In the computation 6 terms were included. The value of λ_{12} is listed for a range of values of c in Table 6.2. The eigenvalue, λ_{12}, is a pessimistic bound on the truncation error, whereas a more accurate bound is given by (6.55). Table 6.2 shows that as the value of the parameter, c, increases more terms have to be included in the truncation (6.54) to maintain a small error.

Another approximation involved in the analysis is approximating c(t) by $d_N(t)$. A measure of goodness of this approximation can be obtained from the fractional energy of c(t) in the interval [-NT,NT].

$$P_{NT} = \frac{\int_{-\infty}^{\infty} d_N^2(t)dt}{\int_{-\infty}^{\infty} c^2(t)dt} \tag{6.57}$$

For $\eta = 0.5\pi$, 0.6π, 0.7π and 0.8π, P_{NT} was greater than the value shown in Table 6.3 for all the values of c from 1.0 to 20.0.

When $\eta = 0.5\pi$, c(t) is strictly time-limited to the interval [-T,T] and therefore identical results were obtained

c	1	5	10	12	14	16	18	20
λ_{12}	2.3×10^{-32}	6.9×10^{-15}	2.1×10^{-7}	1.8×10^{-5}	7.3×10^{-4}	1.5×10^{-2}	1.5×10^{-1}	5.8×10^{-1}

Table 6.2 λ_{12} for a Range of Values of the Parameter c

η	$P_{NT} \geq$			
	N = 1	N = 2	N = 3	N = 4
0.5π	1.000	1.000	1.000	1.000
0.6π	0.977	0.997	0.998	0.998
0.7π	0.920	0.982	0.995	0.997
0.8π	0.850	0.930	0.963	0.979

Table 6.3 P_{NT} for $\eta = 0.5\pi, 0.6\pi, 0.7\pi$ and 0.8π, and N = 1,2,3 and 4

for N = 1, 2, 3 and 4. The final results obtained were identical with the results for MSK-type signals in Chapter 5, Figure 5.3.

The optimum P'_{out} was computed for $\eta = 0.6\pi$, 0.7π and 0.8π, and the results are shown in Figures 6.3, 6.4 and 6.5 respectively. Computations were carried out for N = 1, 2, 3 and 4. As η increases, the waveform c(t) spreads out over a larger interval and it becomes necessary to consider $d_N(t)$ for larger values of N.

The optimum pulse shape $\bar{z}(t)$ can be computed for a given value of c, N and η. Let us define a parameter c_{eq}

$$c_{eq} = \frac{c}{N} \tag{6.58}$$

A baseband pulse optimized for a given value of c and N, maximizes the fractional energy of the modulated signal in a bandwidth of $2W = c_{eq}/(\pi T)$. The out-of-band power variation for the signals modulated by the baseband pulse optimized for N=2 and $\eta=0.7\pi$ are shown in Figure 6.5a for various values of c. The modulated signal optimized for a particular value of c, attains the least possible P_{out} for a channel bandwidth of $2W = c/(\pi NT)$. Therefore, a lower bound on P_{out} can be obtained by joining the out-of-band powers of the signals at the time-bandwidth products for which they were optimized.

One of the approximations made in the optimization is approximating c(t) by $d_N(t)$. Therefore, an improved

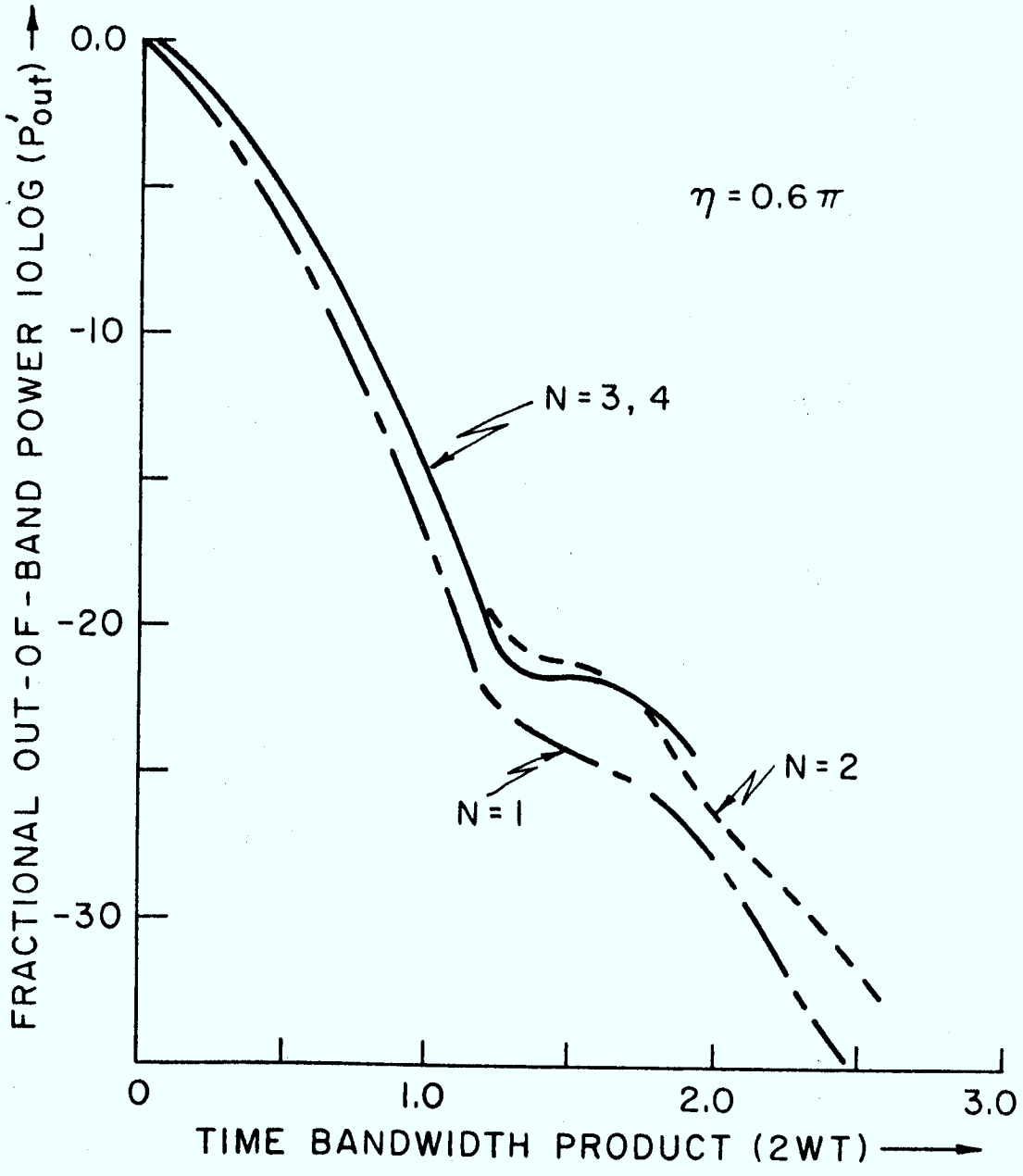


Figure 6.3 Variation of Optimum P'_{out} for $\eta = 0.6\pi$

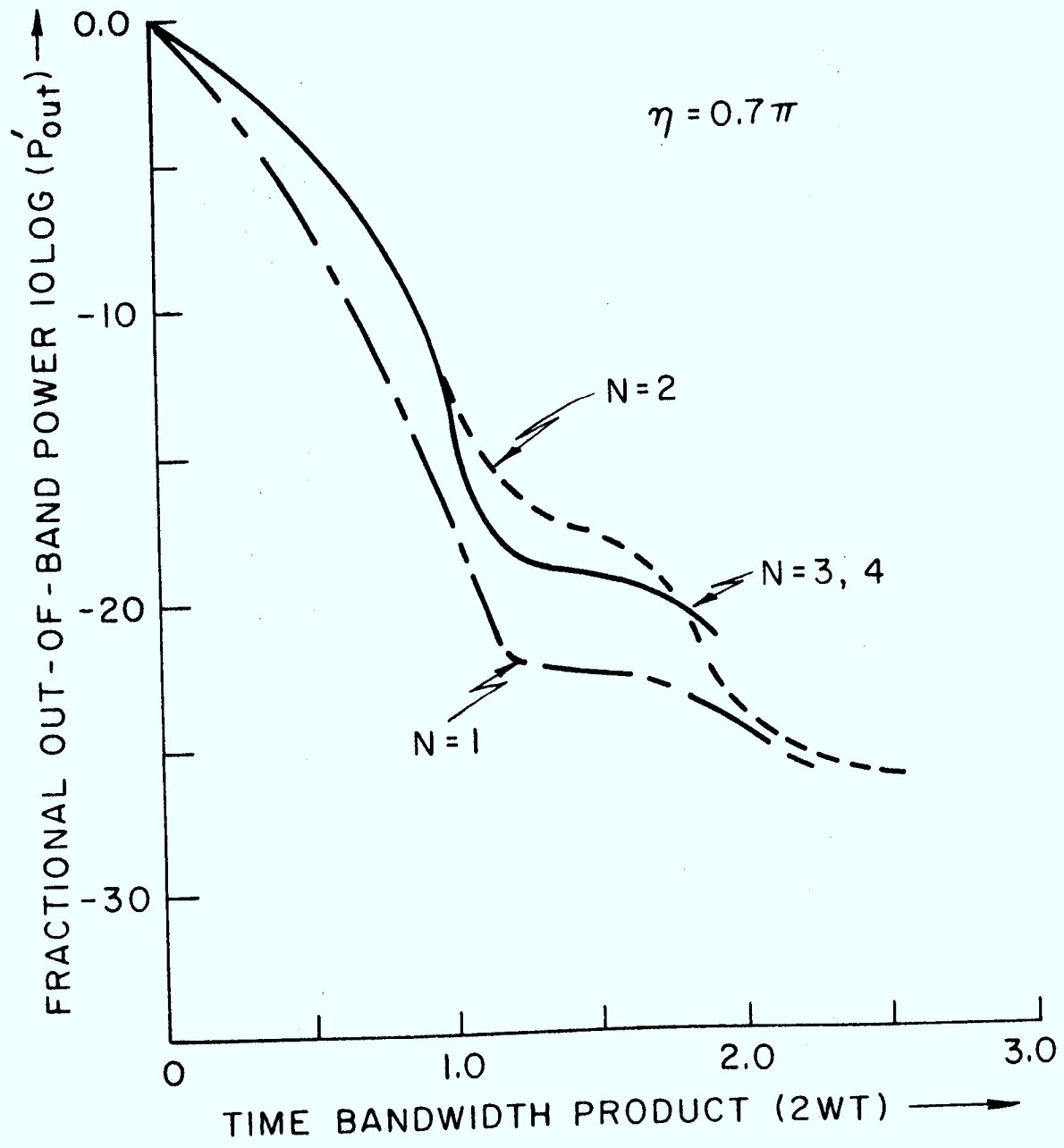


Figure 6.4 Variation of Optimum P'_{out} for $\eta = 0.7\pi$

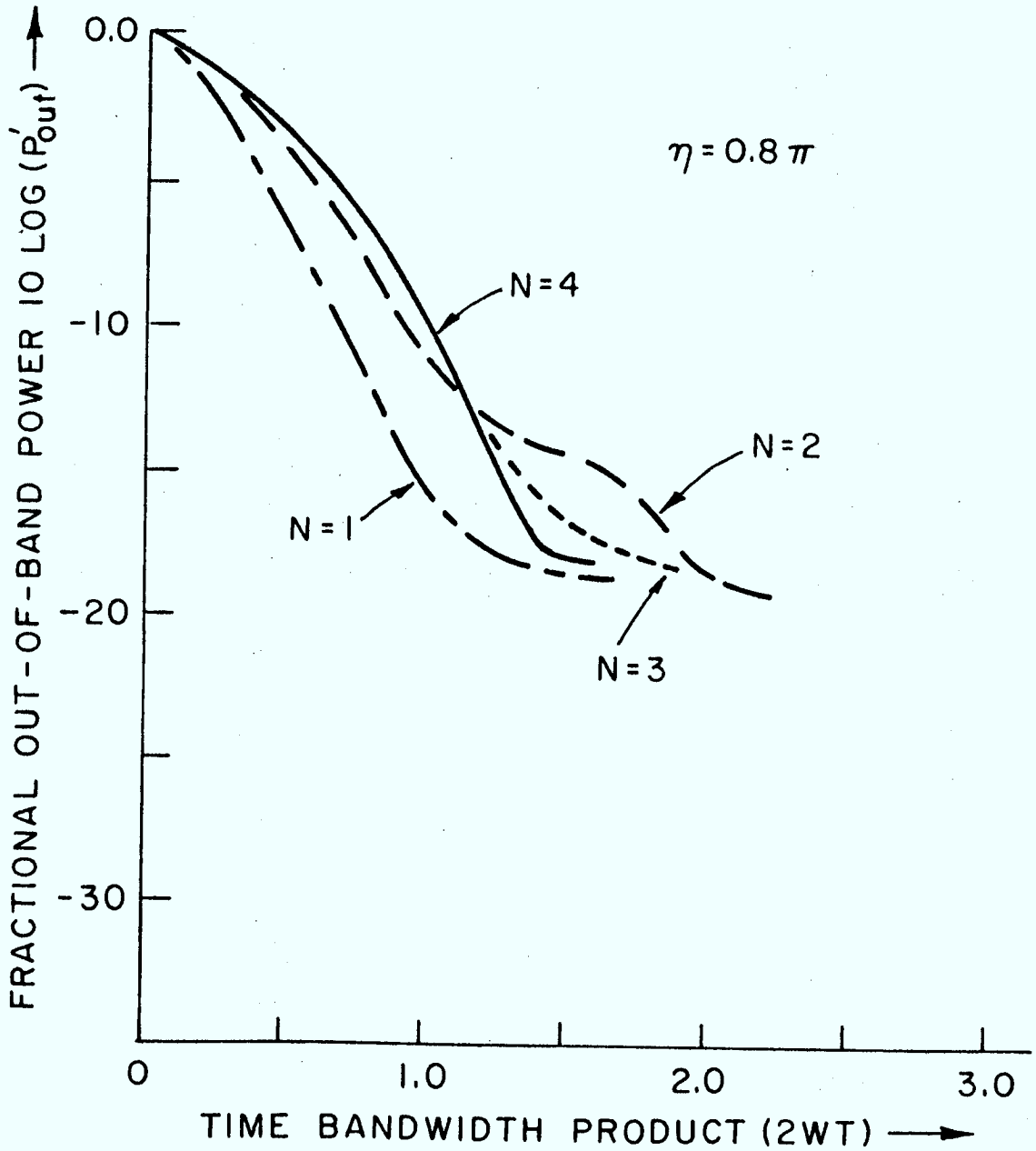


Figure 6.5 Variation of Optimum P'_{out} for $\eta = 0.8\pi$

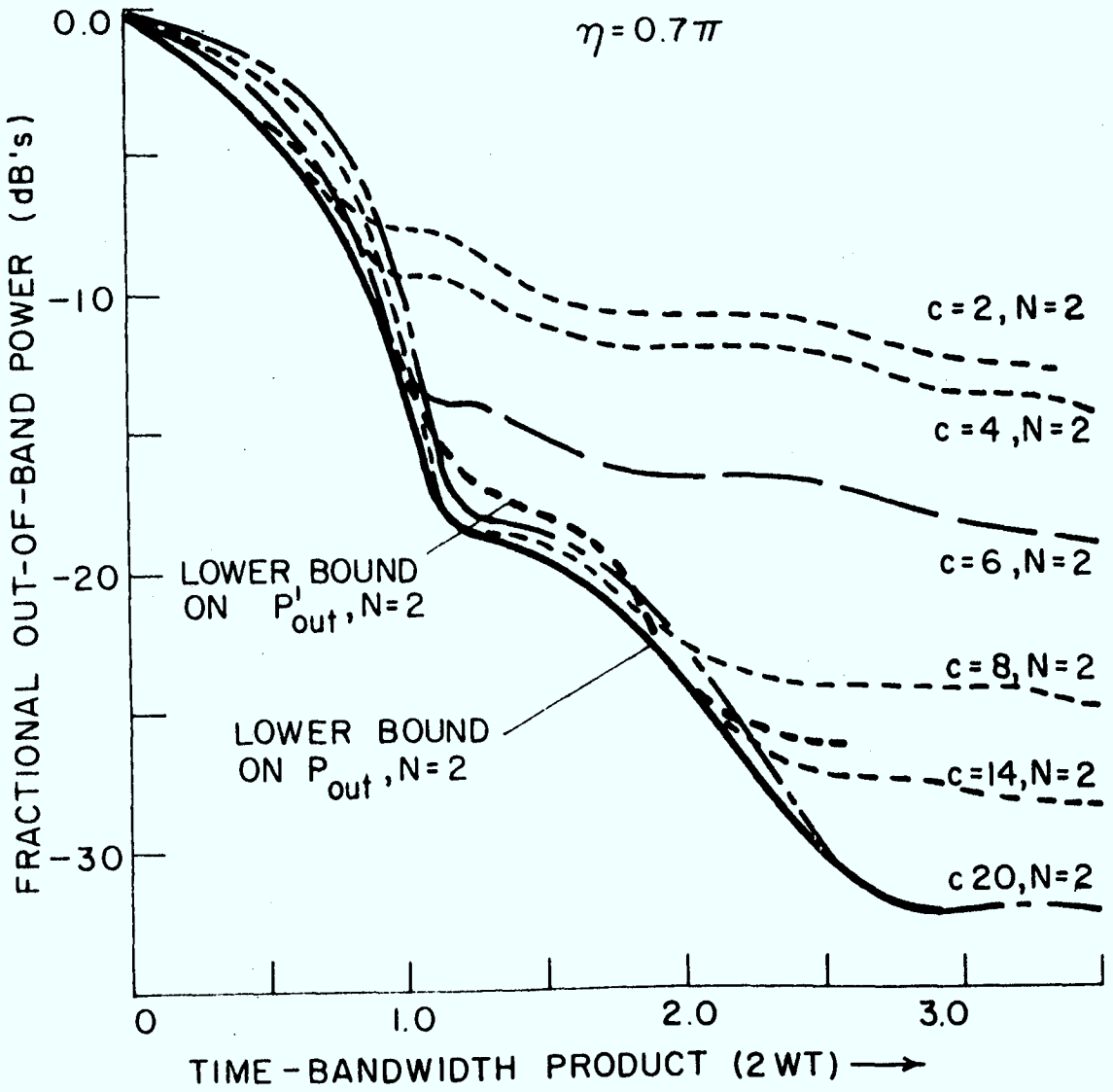


Figure 6.5a Variation of Out-of-Band Power for η = 0.7π

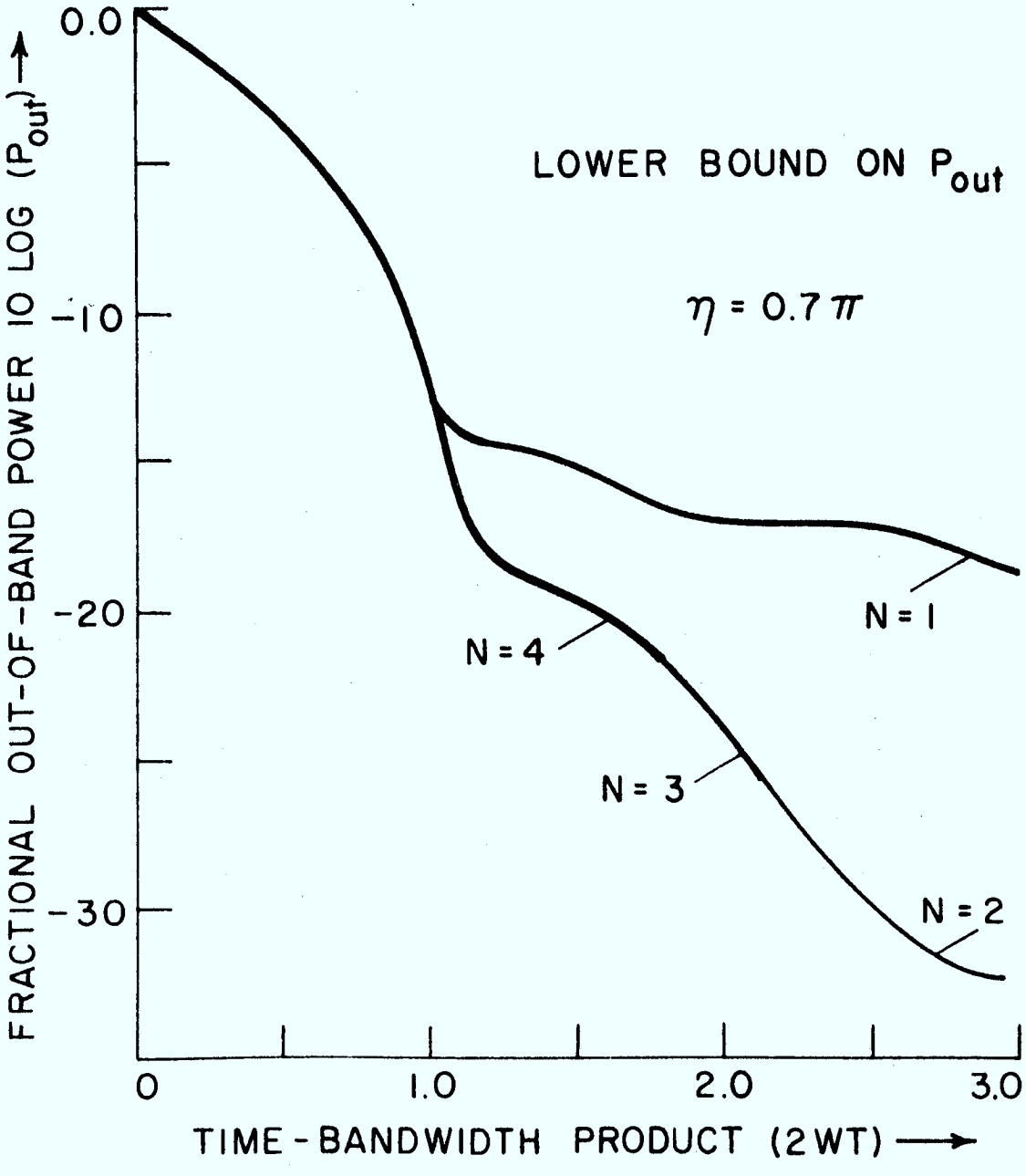


Figure 6.5b Lower Bound on P_{out} for $\eta = 0.7\pi$ and N = 1, 2, 3 and 4

achievable lower bound on P_{out} can be expected as the baseband pulse is optimized for higher values of N. The lower bounds on P_{out} for N = 1,2,3, and 4 are shown in Figure 6.5b for $\eta = 0.7\pi$. It can be seen from this figure that N = 2 provides a sufficiently accurate bound for time-bandwidth products up to 2.5.

The lower bounds on P'_{out} and P_{out} for $\eta = 0.6\pi$, 0.7π and 0.8π are shown in Figures 6.6, 6.7 and 6.8 for N = 4. For $\eta = 0.6\pi$, P_{4T} (P_{NT} is defined in equation (6.53)) is almost unity and therefore, the signals modulated by optimum pulse shapes attain the optimum P'_{out} at the time-bandwidth products for which they are optimized. For $\eta = 0.7\pi$ and 0.8π, the actual fraction of out-of-band power, P_{out}, of the signals modulated by the optimum pulses is different from the optimized P'_{out}. This discrepancy is due to the approximation made in the optimization. If N is taken large enough such that $d_N(t)$ closely approximates c(t), the discrepancy can be resolved.

The lower bounds obtained on P_{out} for $\eta = 0.5\pi$, 0.6π, 0.7π and 0.8π are compared in Figure 6.8a. As η approaches π, a sharp drop in P_{out} can be expected at a time-bandwidth product of unity due to the discrete line in the spectrum that occurs at 2WT = 1.

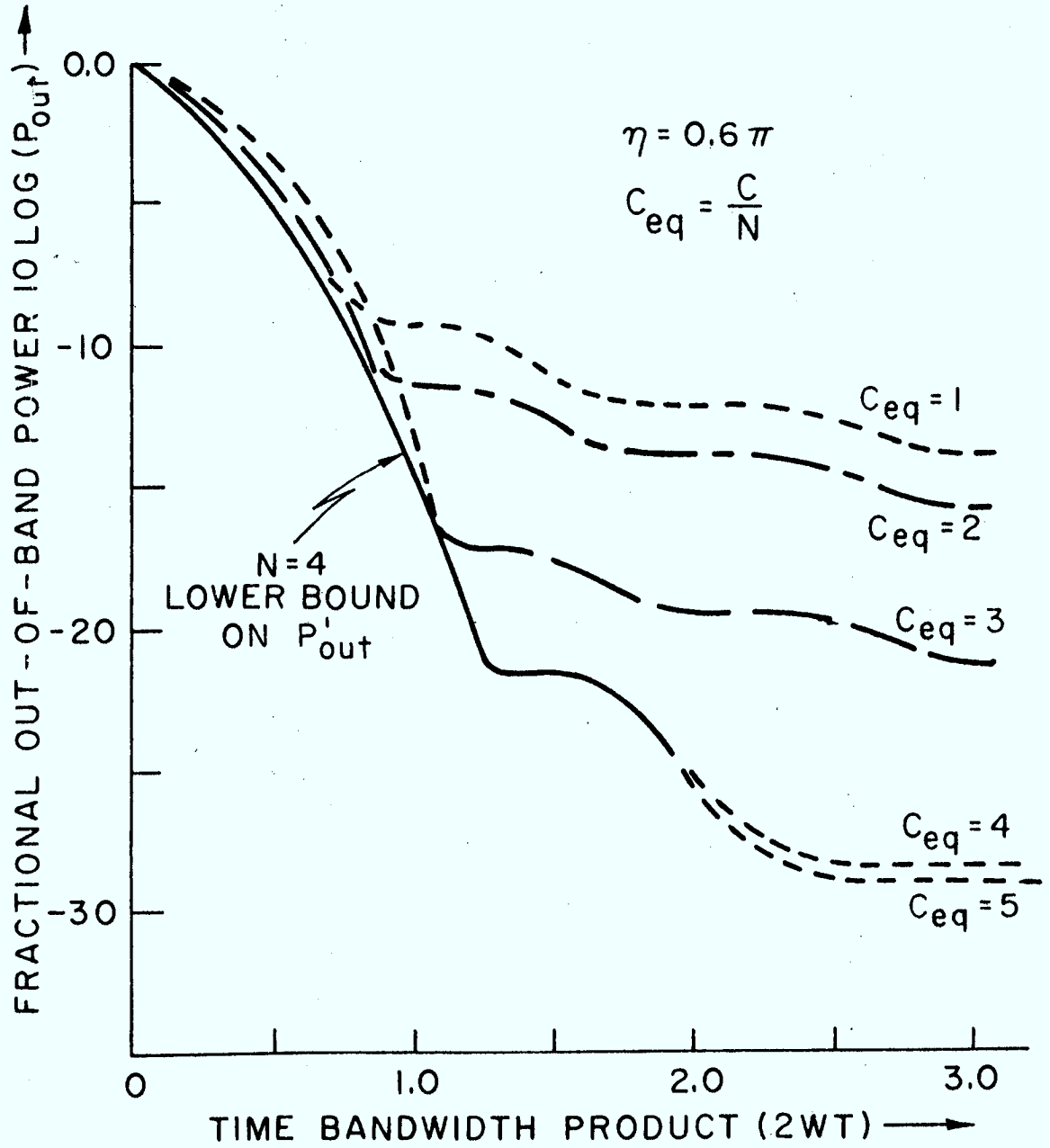


Figure 6.6 Variation of Fraction of Out-of-Band Power for Optimum Pulse Shapes at c_{eq}=1,2,3,4,5 and η=0.6π

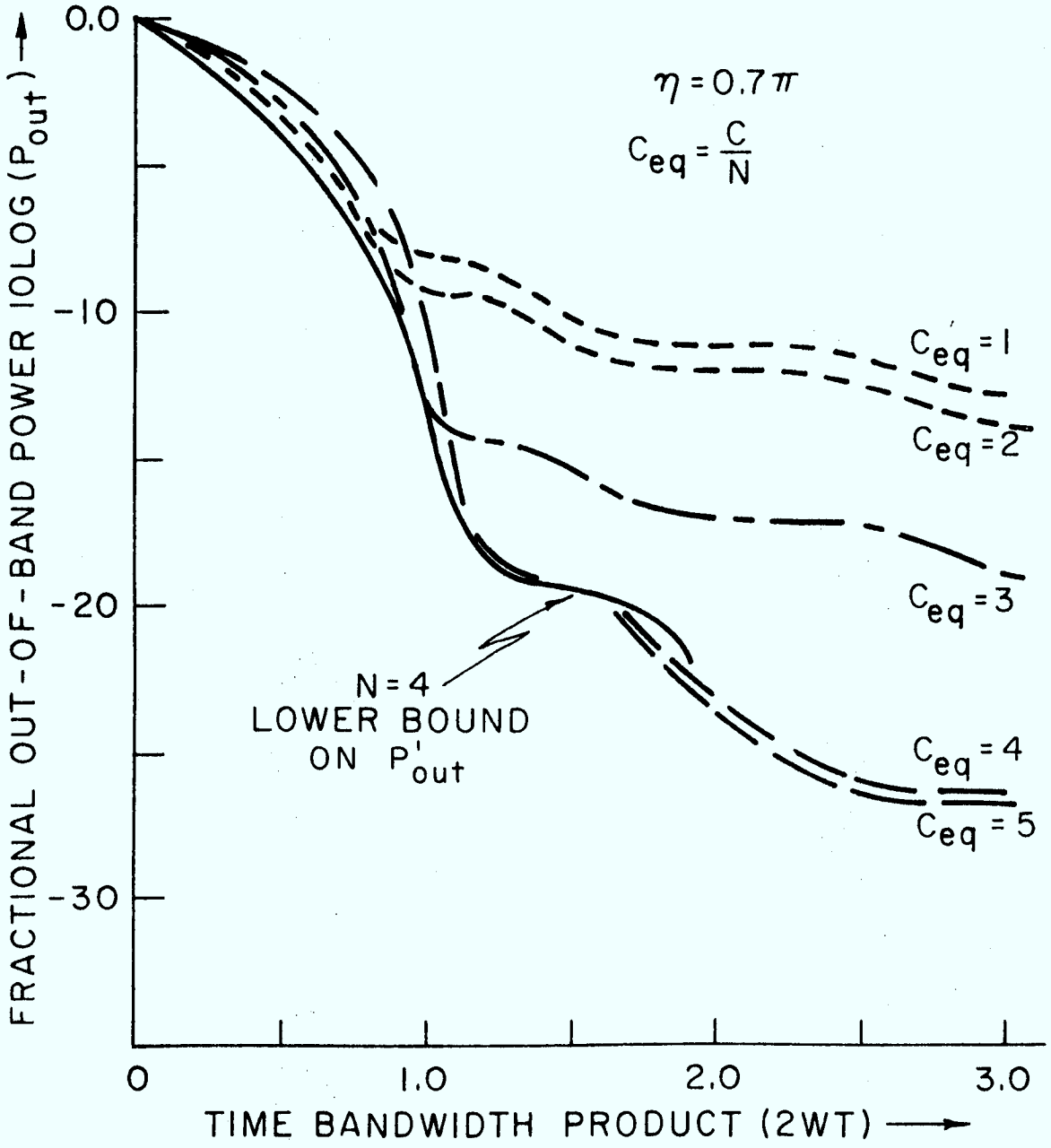


Figure 6.7 Variation of Fraction of Out-of-Band Power for Optimum Pulse Shapes at c_{eq}=1,2,3,4,5 and η=0.7π

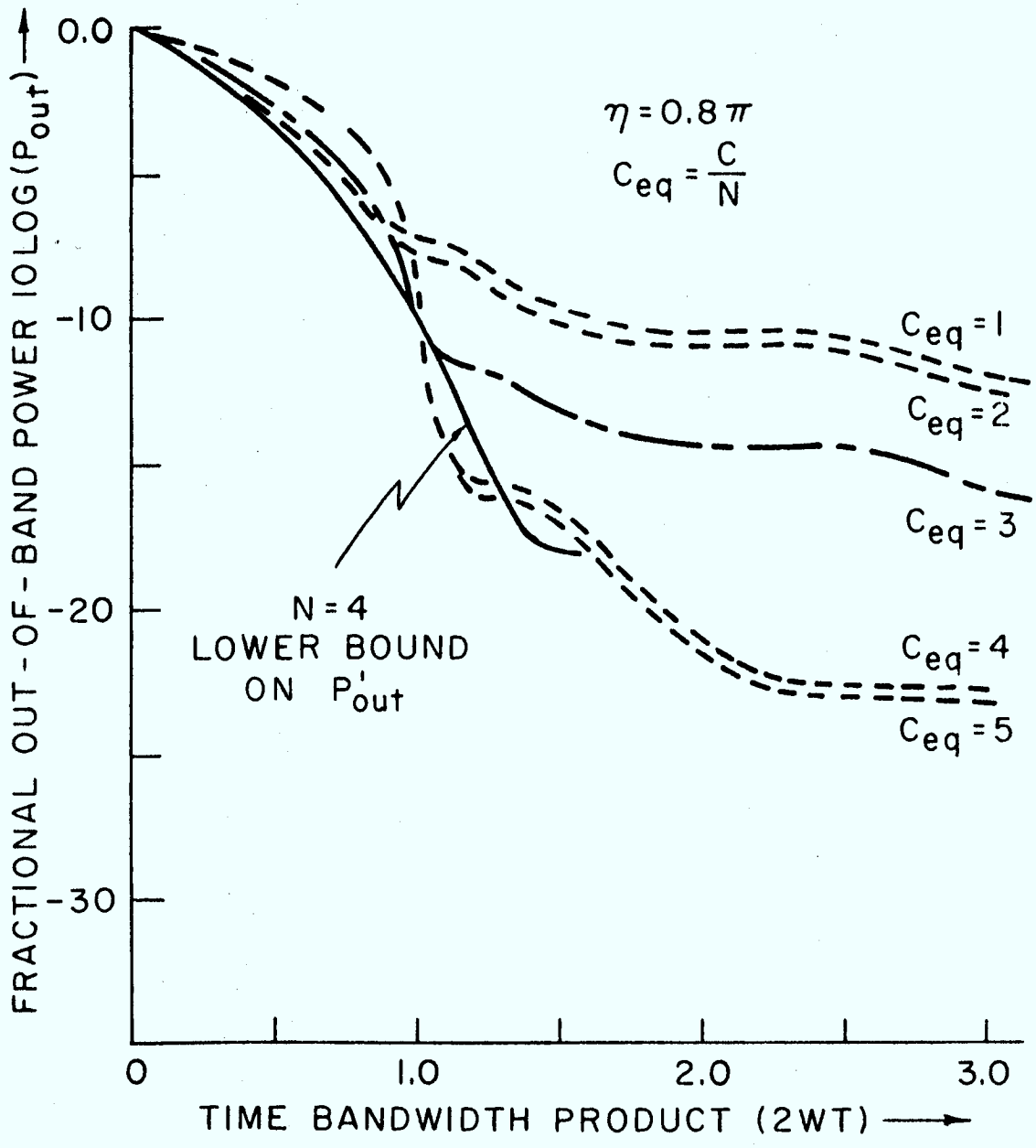


Figure 6.8 Variation of Fraction of Out-of-Band Power for Optimum Pulse Shapes at c_{eq}=1,2,3,4,5 and η=0.8π

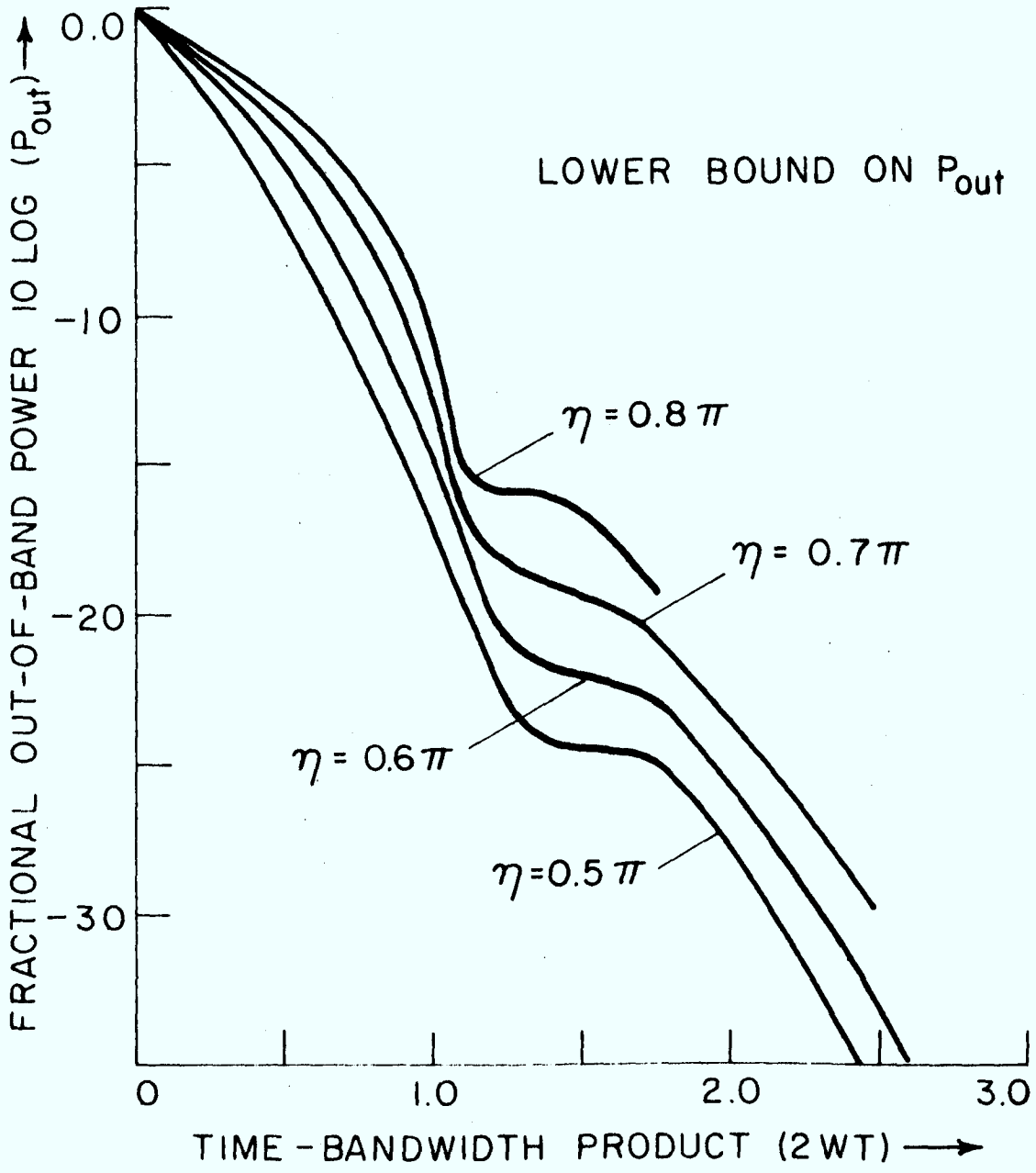


Figure 6.8a Lower Bound on P_{out} for $\eta = 0.5\pi$, 0.6π, 0.7π and 0.8π

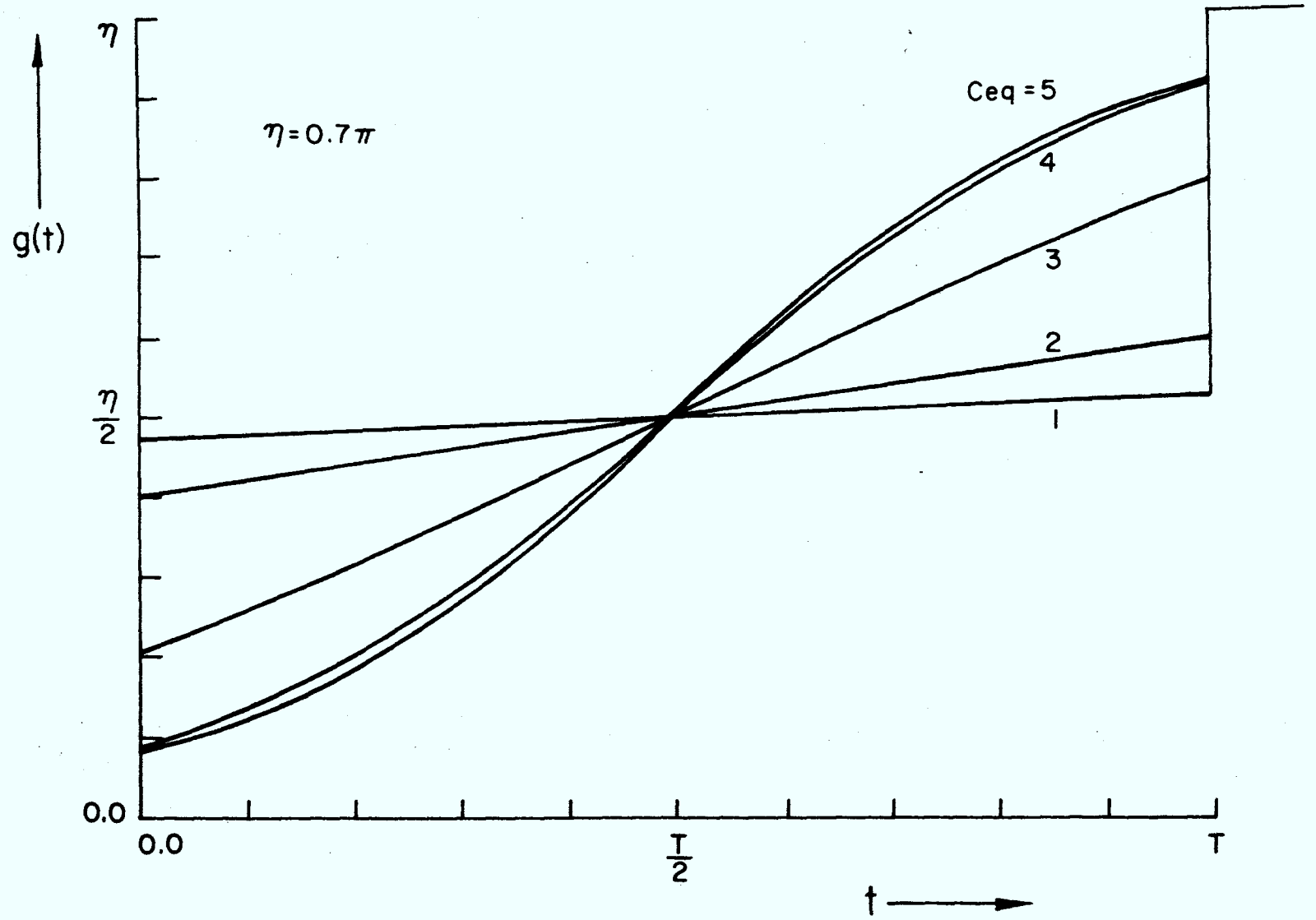


Figure 6.9 Optimum Phase Variation g(t) for $\eta=0.7\pi$, N=4 and c_{eq}=1,2,3,4 and 5

The optimum phase trajectories, $g(t)$, were examined for several modulation indices. The results are shown in Figure 6.9 for $\eta = 0.7\pi$. When the parameter, c, tends to zero, the phase makes an abrupt jump of $\eta/2$ radians. As c increases the magnitude of the jump decreases and the phase changes gradually.

6.5 Summary

A technique for optimizing the baseband pulse shapes in angle modulated signals to minimize the fraction of out-of-band power was presented. The results obtained for angle modulated signals agree with the results for MSK-type signals, when $\eta = 0.5\pi$. The optimum pulse shapes and the fraction of out-of-band power were computed for modulation indices $\eta = 0.6\pi$, 0.7π and 0.8π and the results are given for a range of time-bandwidth products. The phase trajectory for optimum pulse shaping is given for $\eta = 0.7\pi$. When c tends to zero, the phase makes an abrupt jump of $\eta/2$ radians. As c increases the magnitude of the jump decreases and the phase has a gradual variation.

The method of approximating the power spectrum of the modulated signal by the power spectrum of a time-limited waveform, and then optimizing the time-limited waveform to minimize the fraction of out-of-band power appears to be a useful technique. This method may prove useful in the optimization of other modulations.

Chapter 7

CONCLUSIONS

In this report band occupancy of angle modulated signals has been investigated for a range of baseband pulse shapes. The error performance of these modulations has been examined as well. The results on the band occupancy and the error performance are summarized and combined in Section 7.1. Modulation schemes which are spectrally efficient with only a small penalty in SNR are indicated. The new developments in this report are listed in Section 7.2. Finally, the scope for further work is pointed out in Section 7.3.

7.1 Performance of Several Angle Modulated Signals

The bandwidth efficiencies computed for several types of modulations are summarized in Table 7.1. The correlative encoded FM signals are listed according to their equivalent baseband pulse shapes b(t), given by (2.18). The bandwidth efficiency is defined as the ratio of data rate to channel bandwidth in units of bits/sec/Hz. In this definition, the channel bandwidth is the span of frequencies which contains a specified fraction, P_{in}, of the signal power. The bandwidth efficiencies for $P_{in} = 0.99$ and $P_{in} = 0.999$ are given in Table 7.1.

#	Type of Modulation	Bandwidth Efficiency bits/sec/Hz			Error Performance	
		Based on the first null in the Spectrum	Based on 99% of Signal Power	Based on 99.9% of Signal Power	Minimum Euclidean Distance d^2_{min}	$10 \log \frac{d^2_{min}}{d^2_{min}(MSK)}$
1	PSK	0.50	0.05	0.005	2.0	0.0
2	QPSK	1.00	0.10	0.01	2.0	0.0
3	OKQPSK	1.00	0.10	0.01	2.0	0.0
4	MSK (same as Fig.3.2a, $\eta = 0.5\pi$)	0.67	0.85	0.36	2.0	0.0
5	SFSK (same as Fig. 3.2d, $\eta = 0.5\pi$)	0.56	0.45	0.34	2.0	0.0
6	Optimum MSK-type Signal	--	0.89	0.47	2.0	0.0
7	FSK (same as Fig.3.2a) $\eta = 0.7\pi$	--	0.56	0.32	2.4	0.79
8	Optimum possible, $\eta = 0.7\pi$	--	0.60	--	--	--
	Correlative Encoded Digital FM					
9	Figure 3.2b, $\eta = 0.5\pi$	--	1.09	0.56	1.73	-0.63
10	" " $\eta = 0.7\pi$	--	0.77	0.45	2.00	0.0
11	Figure 3.2c, $\eta = 0.5\pi$	--	1.25	0.74	1.19	-2.25
12	" " $\eta = 0.7\pi$	--	1.00	0.56	2.16	0.33
13	Figure 3.2e, $\eta = 0.5\pi$	--	0.63	0.36	--	--
14	" " $\eta = 0.7\pi$	--	0.40	0.33	--	--
15	Figure 3.2f, $\eta = 0.5\pi$	--	0.68	0.38	--	--
16	" " $\eta = 0.7\pi$	--	0.50	0.34	--	--
17	Figure 3.2g, $\eta = 0.5\pi$	--	0.91	0.61	--	--
18	" " $\eta = 0.7\pi$	--	0.83	0.50	--	--
19	Figure 3.2h, $\eta = 0.5\pi$	--	1.16	0.71	--	--
20	" " $\eta = 0.7\pi$	--	0.91	0.63	--	--
21	Figure 3.2i, $\eta = 0.5\pi$	--	1.00	0.83	1.76	-0.56
22	" " $\eta = 0.7\pi$	--	0.94	0.67	2.73	1.35
23	PRS=$(1+2D+D^2)/4$, $\eta = 0.5\pi$ with Rectangular Shaping	--	1.25	0.79	1.55	-1.11

Table 7.1 Performance of Several Digital Angle Modulated Signals

For $P_{in} = 0.99$, the bandwidth efficiency of conventional MSK is only 5% less than the optimum possible with MSK-type signals. This was pointed out by Prabhu [18]. In general, correlative encoded digital FM are spectrally more efficient than MSK. For $\eta = 0.5\pi$ and $P_{in} = 0.99$, the baseband pulse shape in Figure 3.2c, which corresponds to PRS polynomial $(1+D+D^2)/3$ with rectangular pulse shaping, is spectrally 47% more efficient than MSK.

The bandwidth efficiencies of PSK and QPSK are also given in Table 7.1 for the sake of comparison. PSK, QPSK and OKQPSK have power spectra of the form $\{\sin(af)/(af)\}^2$. Since the spectra falls off only as f^{-2}, their bandwidth efficiencies for $P_{in} = 0.99$ and $P_{in} = 0.999$ are very small. Very often in the literature, the channel bandwidth is taken as the bandwidth required to transmit the main lobe of the spectra. The bandwidth efficiencies based on the first null in the spectra are given in Table 7.1 for PSK, QPSK, OKQPSK, MSK and SFSK.

For high SNR a tight upper bound on the error performance of correlative encoded FM can be obtained from the minimum Euclidean distance [57]. The minimum Euclidean distance, d_{min}, is listed in Table 7.1 for the modulations that are spectrally efficient. The penalty in SNR compared to MSK, $10 \log \{d^2_{min} / d^2_{min}\ (MSK)\}$ is also given. The following observations can be made from Table 7.1. In the following comments, the channel bandwidth refers to the bandwidth required to transmit 99% of the signal power.

i) The bandwidth efficiency of MSK is 0.85 bits/sec/Hz and is almost optimum for MSK-type signals.

ii) Correlative encoded FM with the PRS encoding polynomial (1+D)/2 and rectangular pulse shaping (duobinary FSK, Figure 3.2b) with $\eta = 0.5\pi$ is spectrally 28% more efficient than MSK at the expense of 0.6 dB in SNR.

iii) Correlative encoded digital FM with the polynomial $(1+D+D^2)/3$ and rectangular pulse shaping (Figure 3.2c) with $\eta = 0.5\pi$ is spectrally 47% more efficient than MSK with a penalty in SNR of 2.3 dB.

iv) Correlative encoded digital FM with the polynomial $(1+2D+D^2)/4$ and rectangular pulse shaping with $\eta = 0.5\pi$ is spectrally 47% more efficient than MSK at an expense of 1.1 dB in SNR.

v) Digital FM with raised-cosine pulse shaping extending over 3 bit intervals (Figure 3.2i) with $\eta = 0.7\pi$ is spectrally 10% more efficient than MSK with an advantage of 1.4 dB in SNR.

The choice of modulation for a particular application can be a compromise between the spectral efficiency and the SNR required to attain desirable error performance. Also complexity of the receiver may be a consideration. It was pointed out in Chapter 4 that Viterbi algorithm can be used to implement a detector for correlative encoded digital FM. The detectors for the encoding polynomials 1, $(1+D)/2$, $(1+D+D^2)/3$ and $(1+2D+D^2)/4$ require 2, 3, 4 and 5 correlators respectively, and attendant software. At low

bit rates implementation would seem to pose no problem. However, at high bit rates the receiver complexity would be a limiting factor.

7.2 Summary of New Results Presented

This study was directed towards seeking spectrally efficient constant envelope modulation schemes. The contributions made can be summarized as follows:

i) An analytical method of obtaining the spectrum of a correlative encoded digital FM signal was developed. This technique was applied to obtain the spectra for various encoding polynomials with rectangular and raised-cosine pulse shaping. The 99% and 99.9% bandwidths were tabulated. The method is applicable also to digital FM by multilevel modulating signals. As an example, results are given for a quaternary modulating signal with duobinary encoding.

ii) It was shown that reciprocal encoding polynomials yield identical spectra.

iii) An exhaustive investigation of the spectrum for a second order encoding polynomial with rectangular pulse shaping and modulation index $\eta = 0.5\pi$ was carried out. The encoding coefficients for which the spectrum is compact were given.

iv) Optimum baseband pulse shapes were obtained for MSK-type signals to minimize the fraction of out-of-band power for a given bandwidth. Results are presented for a range of channel bandwidths of practical interest. The optimization yields pulse shapes which attain a lower

fraction of out-of-band power than the previously proposed schemes, such as MSK, OKQPSK and SFSK. The improvement is not dramatic. However, the lower bound presented provides a useful indication of what is achievable.

v) A technique for designing the baseband pulse shapes in angle modulated signals that minimize the fraction of out-of-band power was presented. Results are given for modulation indices $\eta = 0.5\pi$, 0.6π, 0.7π and 0.8π.

vi) It is known that an upper bound on the error performance of correlative encoded digital FM can be obtained from the minimum Euclidean distance. It was demonstrated that reciprocal encoding polynomials have identical minimum Euclidean distance.

vii) An exhaustive investigation of the minimum Euclidean distance for a second order encoding polynomial with rectangular pulse shaping and $\eta = 0.5\pi$ was carried out.

viii) The results indicate that second order encoding polynomials can provide up to 47% more spectral efficiency than MSK at an expense of 1.1 dB in SNR. Correlative encoding FM modulations exist which are spectrally more efficient and also possess better error performance capability than MSK.

7.3 Suggestions for Further Investigation

It is evident from this study that correlative encoded digital FM is superior to conventional FSK in terms of spectral conservation. A correlative encoded FM

is completely specified by its equivalent baseband pulse, b(t), given by equation (2.18). When the baseband pulse extends over only one bit interval, pulse shapes were derived that minimize the fraction of out-of-band power. It may be possible to extend this technique to design the baseband pulse shapes that extend over more than one bit period. If this optimization is carried out it will provide an useful indication of the maximum achievable spectral efficiency by correlative encoding.

It is known that an upper bound on the error performance of correlative encoded digital FM can be obtained from the minimum Euclidean distance. It was pointed out in Chapter 4 that an initial error in the detector may well result in a run of errors. In that case, the bound obtained would be a bound on the probability of run of errors. However, the probability of run of errors is a good estimate of the probability of bit error if the average length of run of errors is small. An investigation of the nature of error propagation should be carried out to verify the bounds obtained. This could be done by simulation.

This study shows that spectral efficiency can be increased considerably at a small penalty in SNR, by correlatively encoding the baseband signal in digital FM. Also correlative encoded digital FM modulations exist that are spectrally more efficient than MSK and have a better error performance. However, this error performance can

be realized only if good detectors for these modulations can be implemented economically. It was pointed out that detectors for correlative encoded digital FM can be implemented using Viterbi algorithm. Further investigation is warranted. The possibility of implementing receivers which are economical and which give performance hopefully close to the optimum should be investigated. The problem of carrier synchronization and bit timing recovery should also be considered.

REFERENCES

1. Smith, J.G., "Spectrally Efficient Modulation", International Communications Conference Record, ICC-77, PP.3.1.37 - 3.1.41, 1977.

2. Anderson, C., and Barber, S., "Modulation: The Key to Effective Bandwidth Utilization", Telesis, Vol. 5, No. 6, PP. 172 - 179, December 1977.

3. Lender, A., "The Duobinary Technique for High Speed Data Transmission", IEEE Transactions on Communications and Electronics, PP. 214 - 218, May 1963.

4. Kretzmer, E.R., "Generalization of a Technique for Binary Data Communication", IEEE Transactions on Communications Technology, Vol. COM-14, No. 2, PP. 67 - 68, February 1968.

5. Kabal, P., and Pasupathy, S., "Partial Response Signalling", IEEE Transactions on Communications, Vol. COM-23, No. 9, PP. 921 - 934, September 1975.

6. Melvin, W.J., and Middlestead, R.W., "Power Density Spectrum of M-ary Correlative Encoded MSK", International Communications Conference Record, ICC-77, PP. 3.7.60 - 3.7.63, 1977.

7. Von Baeyer, H.J., and Tjhung, T.T., "Spectra of Biternary Digital FM", Electronics Letters, Vol. 2, No. 2, PP. 54 - 56, February 1966.

8. Roth, D., "Mean Power Density Spectrum of a Continuous FSK Signal Modulated by the Duobinary Sequence", National Telecommunications Conference Record, NTC - 74, PP. 572 - 579, 1974.

9. Garrison, G.J. "A Power Spectral Density Analysis for Digital FM", IEEE Transactions on Communications, Vol. COM-23, No. 11, PP. 1228 - 1243, November 1975.

10. Doelz, M.L., and Heald, E.H., "Minimum Shift Data Communication System", U.S. Patent No. 2 977 417, March 28, 1961 (assigned to Collins Radio Company).

11. van den Elzen, H.C., and van der Wurf, P., "A simple Method of Calculating the Characteristics of FSK Signals with Modulation Index of 0.5", IEEE Transactions on Communications, Vol. COM-20, No. 4, PP. 139 - 147, April 1972.

12. deBuda, R., "Coherent Demodulation of Frequency-Shift Keying with Low Deviation Ratio", IEEE Transactions on Communications, Vol. COM-20, No. 6, PP. 429 - 436, June 1972.

13. Rabow, G., "Amplitude and Time-Limited Functions with Minimum Out-of-Band Energy", IEEE Transactions on Communications, Vol. COm-20, No. 12, PP. 1150 - 1153, December 1972.

14. Tou, F., and Simpson, R. S., "Optimum Deviation Ratio and Observation Interval for Continuous Phase Binary Frequency-Shift Keying", IEEE Transactions on Communications, Vol. COM-21, No. 9, PP. 1067 - 1069, September 1973.

15. Mathwich, H., Balcewicz, J.F., and Hecht, M., "The Effect of Tandem Band and Amplitude Limiting on the E_b/N_o Performance of Minimum (Frequency) Shift Keying (MSK)", IEEE Transactions on Communications, Vol. COM-22, No. 10, PP. 1525 - 1540, October 1974.

16. deBuda, R., "Fast FSK Signals and their Demodulation", Canadian Electrical Engineering Journal, Vol. 1, No. 1, PP. 28 - 34, 1976.

17. Amoroso, F., "Pulse and Spectrum Manipulation in the Minimum (Frequency) Shift Keying (MSK) Format", IEEE Transactions on Communications, Vol. COM-24, No. 3, PP. 381 - 384, March 1976.

18. Prabhu, V.K., "Spectral Occupancy of Digital Angle-Modulation Signals", BSTJ, Vol. 55, No. 4, PP. 429 - 453, April 1976.

19. Gronmeyer, S.A., and McBride, A.L., "MSK and Offset QPSK Modulation", IEEE Transactions on Communications, Vol. COM-24, No. 8, PP. 809 - 820, August 1976.

20. Simon, M.K., "A Generalization of Minimum-Shift-Keying Type Signalling Based Upon Input Data Symbol Pulse Shaping", IEEE Transactions on Communications, Vol. COM-24, No. 8, PP. 845 - 856, August 1976.

21. Kalet, I., "A look at Crosstalk in Quadrature-Carrier Modulation Systems", IEEE Transactions on Communications, Vol. COM-25, No. 9, PP. 884 - 892, September 1977.

22. Kalet, I., and White, B.E., "Suboptimal Continuous Shift Keyed (CSK) Demodulation for the Efficient Implementation of Low Crosstalk Data Communications," IEEE Transactions on Communications, Vol. COM-25, No. 9, PP. 1037 - 1041, September 1977.

23. Rabzel, M., and Pasupathy, S., "Spectral Shaping in Minimum Shift Keying Type Signalling", IEEE Transactions on Communications, Vol. COM-26, PP. 189 - 195 January 1978.

24. Boutin, N., Morisette, S., and Dussault, L., "Constant Amplitude PSK Signal with Minimum Out of Band Energy", National Telecommunications Conference Record, NTC-78, PP. 6.3.1 - 6.3.3, December 1978.

25. Eaves, R.E., and Wheatley, S.M., "Optimization of Quadrature - Carrier Modulation for Low Crosstalk and Close Packing of Users", IEEE Transactions on Communications, Vol. COM-27, No. 1, PP. 176 - 186, January 1979.

26. Bazin, B., "A Class of MSK Baseband Pulse Formats with Sharp Spectral Roll-off", IEEE Transactions on Communications, Vol. COM-27, No. 5, PP.826 - 829, May 1979.

27. Bennett, W.R., and Rice, S.O., "Spectral Density and Autocorrelation Functions Associated with Binary FSK", BSTJ, Vol. 42, No. 9, PP. 2355 - 2385, September 1963.

28. Anderson, R.R., and Salz, J., "Spectra of Digital FM", BSTJ, Vol. 44, No. 6, PP. 1165 - 1189, July 1965.

29. Salz, J., "Spectral Density Functions of Multi-level Continuous-Phase FM", IEEE Transactions on Information Theory, Vol. IT-11, No. 7, PP. 429 - 433, July 1965.

30. Shimbo, O., "General Formula for Spectra of Digital F.M. Signals", Proceedings of IEE, Vol. 113, No. 11, PP. 1783 - 1789, November 1966.

31. Wittke, P.H., "On Frequency-Modulation Spectra and the Distribution of a Functional of the Modulating Process", Research Report No. 66-1, Dept. of Electrical Engineering, Queen's University, Kingston, Ontario 1966.

32. Tjhung, T.T., "Band Occupancy of Digital FM Signals", IEEE Transactions on Communications, Vol. COM-12, No. 12, PP. 211 - 216, December 1964.

33. Rowe, H.E., and Prabhu, V.K., "Power Spectrum of a Digital Frequency-Modulation Signal", BSTJ, Vol. 54, PP. 1095 - 1125, August 1975.

34. Rice, S.O., "Noise in FM Receivers", Chapter 25 of "Time Series Analysis", M. Rosenblatt (ed.), John Wiley and Sons, Inc. New York, 1963.

35. Mazo, J.E., and Salz, J., "Theory of Error Rates for Digital FM", BSTJ, Vol. 45, No. 9, PP. 1511 - 1535, November 1966.

36. Klapper, J., "Demodulator Threshold Performance and Error Rates in Angle Modulated Digital Signals", RCA Review, Vol. 27, PP. 226 - 244, June 1966.

37. Schilling, D.L., Hoffman, E., and Nelson, E.A., "Error Rates for Digital Signals Demodulated by an FM Discriminator", IEEE Transactions on Communications, Vol. COM-15, No. 4, PP. 507 - 517, August 1967.

38. Tjhung, T.T., and Wittke, P.H., "Carrier Transmission of Binary Data in a Restricted Band", IEEE Transactions on Communications, Vol. COM-18, No. 8, PP. 295 - 304, August 1970.

39. Papantoni-Kazakos, P., and Kaz, I.M., "The Performance of a Digital FM System with Discriminator; Intersymbol Interference Effects", IEEE Transactions on Communications, Vol. COM-23, No. 9, PP.865 - 878, September 1975.

40. Swartz, T.L., "Performance Analysis of a Three-Level Modified Duobinary Digital FM Microwave Radio System", International Communications Conference Record, ICC-74, PP. 5D1 - 5D4, 1974.

41. Lender, A., "Seven-Level Correlative Digital Transmission Over Radio", International Communications Conference Record, ICC-76, PP. 18.22 - 18.26, 1976.

42. Lucky, R.W., Salz, J., and Weldon, Jr., E.J., "Principles of Data Communications", McGraw-Hill Book Co., Inc., New York, N.Y. 1968.

43. Pearce, J.L., and Wittke, P.H., "Optimum Reception of Digital FM Signals", 1970 IEEE Symposium on Communications, Montreal, Canada, November 1970.

44. Pelchat, M.G., and Adams, S.L., "Noncoherent Detection of Continuous Phase Binary FSK", Proceedings of Int. Conf. Communications, Montreal, P.Q., Canada, 1971.

45. Pelchat, M.G., Davis, R.C., and Luntz, M.B., "Coherent Demodulation of Continuous Phase Binary FSK Signals", Proc. Int. Telemetry Conf., Washington, D.C., 1971.

46. Forney, G.D., "The Viterbi Algorithm", Proc. IEEE, Vol. 61, PP. 268 - 278, March 1973.

47. Van Trees, H.L., "Detection Estimation and Modulation Theory", Part I, Section 4.4, New York Wiley, 1968.

48. Osborne, W.P., and Luntz, M.B., "Coherent and Non-Coherent Detection of CPFSK", IEEE Transactions on Communications, Vol. COM-22, No. 8, PP. 1023 - 1036, August 1974.

49. Schonoff, T.A., "Symbol Error Probabilities for M-ary CPFSK; Coherent and Non-Coherent Detection", IEEE Transactions on Communications, Vol. COM-24, No. 6, PP. 644 - 652, June 1976.

50. Schonoff, T.A., Nicholos, H.E., and Gibbons, H.M., "Use of the MLSE Algorithm to Demodulate CPFSK", International Communications Conference Record, ICC-78, PP. 25.4.1 - 25.4.5, 1978.

51. Anderson, J.B., and Taylor, D.P., "A Bandwidth-Efficient Class of Signal-Space Codes", IEEE Transactions on Information Theory, Vol. IT-24, No. 6, PP. 703 - 712, November 1978.

52. Aulin, T., and Sundberg, C.E., "Binary CPFSK Type of Signalling with Input Data Symbol Pulse Shaping - Error Probability and Spectrum", Technical Report TR-99, Electrical Engineering Department, University of Lund, Sweden, April 1978.

53. Aulin, T., Rydbeck, N., and Sundberg, C.E., "Bandwidth Efficient Digital FM with Coherent Phase Tree Demodulation," Technical Report TR-102, Electrical Engineering Department, University of Lund, Sweden, May 1978.

54. Aulin, T., and Sundberg, C.E., "M-ary CPFSK Type of Signalling with Input Data Symbol Pulse Shaping - Minimum Distance and Spectrum", Technical Report TR-111, Electrical Engineering Department, University of Lund, Sweden, August 1978.

55. Aulin, T., Rydbeck, N., and Sundberg, C.E., "Further Results on Digital FM with Coherent Phase Tree Demodulation - Minimum Distance and Spectrum", Technical Report TR-119, Electrical Engineering Department, University of Lund, Sweden, November 1978.

56. Aulin, T., and Sundberg, C.E., "Minimum Distance Properties of M-ary Correlative Encoded CPFSK", Technical Report TR-120, Electrical Engineering Department, University of Lund, Sweden, November 1978.

57. Aulin, T., Rydbeck, N., and Sundberg, C.E., "Bandwidth Efficient Constant Envelope Digital Signalling with Phase Tree Demodulation", Electronics Letter, Vol. 14, No. 15, PP. 487 - 489, July 1978.

58. Aulin, T., and Sundberg, C.E., "Bounds on the Performance of Binary CPFSK Type of Signalling with Input Data Symbol Pulse Shaping", National Telecommunications Record, NTC-78, PP. 6.5.1 - 6.5.5, December 1978.

59. Slepian, D., and Pollak, H.O., "Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty - I", BSTJ, Vol. 40, No. 1, PP. 43 - 64, January 1961.

60. Landau, H.J., and Pollak, H.O., "Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty -II", BSTJ, Vol. 40, No. 1, PP. 65 - 84, January 1961.

61. Slepian, D., and Sonnenblick, E., "Eigenvalues Associated with Prolate Spheroidal Wave Functions of Zero Order", BSTJ, Vol. 44, No. 10, PP. 1745 - 1759, October 1965.

62. Franks, L.E., "Signal Theory", Chapter 8, Prentice-Hall Inc., Englewood Cliffs, N.J., 1969.

APPENDIX A

EXPRESSIONS FOR THE VECTORS $\underline{R(f)}$

The power spectral density for a correlative encoded digital FM signal, given by equation (2.8) of Chapter 2, is in the form

$$P_v(f) = \frac{1}{T} \left[\underline{R(f)} \cdot (A + A^{\dagger}) \cdot R^{*}(f) \right] \qquad (A.1)$$

The vector $\underline{R(f)}$ in the above equation is determined by $b(t)$, the equivalent baseband pulse shape after correlative encoding. The expressions for $\underline{R(f)}$ for the waveforms of Chapter 3 are given in this appendix. If the waveform extends over K bit intervals, the vector $\underline{R(f)}$ is of dimension $D = 2^K$. The elements of the vector $\underline{R(f)}$ can be written as

$$\underline{R(f)} = [R_1(f) \quad R_2(f) \quad \ldots \quad R_D(f)] \qquad (A.2)$$

The expression for $\underline{R(f)}$ is given by equation (2.10) of Chapter 2. Since $h_2(t) = -h_1(t)$, it follows that

$$R_i(f) = R^{*}_{D+1-i}(-f) \qquad (A.3)$$

Therefore, for each waveform in Figure 3.2 only $R_i(f)$, $i = 1,2,3. \ . \ .(\frac{D}{2} - 1)$ are listed. The following notation is used in the expressions.

$$\nu(y) = \sin(y)/y$$

$$x = 2\pi fT$$

$J_n(y)$ = Bessel function of the first kind of order n

1. Figure 3.2a

$$R_1(f) = \nu(\tfrac{1}{2}(\eta - x))\exp(\tfrac{j}{2}(\eta - x)) \qquad (A.4)$$

2. Figure 3.2b

$$\left.\begin{aligned} R_1(f) &= \nu(\tfrac{1}{2}(\eta - x))\exp(j(\eta - x/2)) \\ R_2(f) &= \nu(x/2)\exp(\tfrac{j}{2}(\eta - x)) \end{aligned}\right\} \qquad (A.5)$$

3. Figure 3.2c

$$\left.\begin{aligned} R_1(f) &= \nu(\tfrac{1}{2}(\eta - x))\exp(\tfrac{j}{2}(3\eta - x)) \\ R_2(f) &= \nu(\tfrac{1}{2}(\tfrac{\eta}{3} - x))\exp(\tfrac{j}{2}(\tfrac{7\eta}{3} - x)) \\ R_3(f) &= \nu(\tfrac{1}{2}(\tfrac{\eta}{3} - x))\exp(\tfrac{j}{2}(\eta - x)) \\ R_4(f) &= \nu(\tfrac{1}{2}(\tfrac{\eta}{3} + x))\exp(\tfrac{j}{2}(\tfrac{\eta}{3} - x)) \end{aligned}\right\} \qquad (A.6)$$

4. Figure 3.2d

$$R_1(f) = \exp(\tfrac{j}{2}(\eta - x))\Big[J_o(\tfrac{\eta}{2\pi})\nu(\tfrac{1}{2}(\eta - x)) + \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{2\pi})\Big\{\nu(\tfrac{\eta}{2} + 2\pi n - \tfrac{x}{2}) + \nu(\tfrac{\eta}{2} - 2\pi n - \tfrac{x}{2})\Big\} + \sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{2\pi})\Big\{\nu(\tfrac{\eta}{2} + (2n-1)\pi - \tfrac{x}{2}) - \nu(\tfrac{\eta}{2} - (2n-1)\pi - \tfrac{x}{2})\Big\}\Big] \qquad (A.7)$$

5. Figure 3.2e

$$\left.\begin{aligned} R_1(f) &= \exp(j(\eta - x/2))\Big[J_o(\tfrac{\eta}{2\pi})\nu(\tfrac{1}{2}(\eta - x)) + \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{2\pi})\Big\{\nu(\tfrac{\eta}{2} + 2\pi n - \tfrac{x}{2}) + \nu(\tfrac{\eta}{2} - 2\pi n - \tfrac{x}{2})\Big\} + \sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{2\pi})\Big\{\nu(\tfrac{\eta}{2} + (2n-1)\pi - \tfrac{x}{2}) - \nu(\tfrac{\eta}{2} - (2n-1)\pi - \tfrac{x}{2})\Big\}\Big] \\ R_2(f) &= \nu(x/2)\exp(\tfrac{j}{2}(\eta - x)) \end{aligned}\right\} \qquad (A.8)$$

6. Figure 3.2f

$$R_1(f) = \exp(\tfrac{j}{2}(3\eta - x))\Big[J_o(\tfrac{\eta}{2\pi})\nu(\tfrac{1}{2}(\eta - x))$$

$$+ \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{2\pi})\Big\{\nu(\tfrac{\eta}{2} + 2\pi n - \tfrac{x}{2})$$

$$+ \nu(\tfrac{\eta}{2} - 2\pi n - \tfrac{x}{2})\Big\}$$

$$+ \sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{2\pi})\Big\{\nu(\tfrac{\eta}{2} + (2n-1)\pi - \tfrac{x}{2})$$

$$- \nu(\tfrac{\eta}{2} - (2n-1)\pi - \tfrac{x}{2})\Big\}\Big]$$

$$R_2(f) = \exp(\tfrac{j}{2}(\tfrac{7\eta}{3} - x))\Big[J_o(\tfrac{\eta}{6\pi})\nu(\tfrac{1}{2}(\tfrac{\eta}{3} - x))$$

$$+ \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{6\pi})\Big\{\nu(\tfrac{\eta}{6} + 2\pi n - \tfrac{x}{2})$$

$$+ \nu(\tfrac{\eta}{6} - 2\pi n - \tfrac{x}{2})\Big\}$$

$$+ \sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{6\pi})\Big\{\nu(\tfrac{\eta}{6} + (2n-1)\pi - \tfrac{x}{2})$$

$$- \nu(\tfrac{\eta}{6} - (2n-1)\pi - \tfrac{x}{2})\Big\}\Big]$$

$$R_3(f) = \exp(\tfrac{j}{2}(\eta - x))\Big[J_o(\tfrac{\eta}{6\pi})\nu(\tfrac{1}{2}(\tfrac{\eta}{3} - x)$$

$$+ \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{6\pi})\Big\{\nu(\tfrac{\eta}{6} + 2\pi n - \tfrac{x}{2})$$

$$+ \nu(\tfrac{\eta}{6} - 2\pi n - \tfrac{x}{2})\Big\}$$

$$+ \sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{6\pi})\Big\{\nu(\tfrac{\eta}{6} + (2n-1)\pi - \tfrac{x}{2})$$

$$- \nu(\tfrac{\eta}{6} - (2n-1)\pi - \tfrac{x}{2})\Big\}\Big]$$

$$R_4(f) = \exp(\tfrac{j}{2}(\tfrac{\eta}{3} - x))\Big[J_o(\tfrac{\eta}{6\pi})\nu(\tfrac{1}{2}(\tfrac{\eta}{3} + x)$$

$$+ \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{6\pi})\Big\{\nu(\tfrac{\eta}{6} + 2\pi n + \tfrac{x}{2})$$

$$+ \nu(\tfrac{\eta}{6} - 2\pi n + \tfrac{x}{2})\Big\}$$

$$+ \sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{6\pi})\Big\{\nu(\tfrac{\eta}{6} + (2n-1)\pi + \tfrac{x}{2})$$

$$- \nu(\tfrac{\eta}{6} - (2n-1)\pi + \tfrac{x}{2})\Big\}\Big] \qquad (A.9)$$

7. Figure 3.2g

$$R_1(f) = \nu(\tfrac{1}{2}(\eta - x))\exp(j(\eta - x/2))$$

$$R_2(f) = \exp(\tfrac{j}{2}(\eta - x))\Big[J_o(\tfrac{\eta}{\pi})\nu(\tfrac{x}{2})$$

$$+ \sum_{n=1}^{\infty}(-1)^n J_{2n}(\tfrac{\eta}{\pi})\Big\{\nu(\tfrac{1}{2}(x + 2\pi n)) + \nu(\tfrac{1}{2}(x - 2\pi n))\Big\}$$

$$- j\sum_{n=1}^{\infty}(-1)^n J_{2n-1}(\tfrac{\eta}{\pi})\Big\{\nu(\tfrac{1}{2}(x + (2n-1)\pi)$$

$$+ \nu(\tfrac{1}{2}(x - (2n-1)\pi))\Big\}\Big] \qquad (A.10)$$

8. Figure 3.2h

$$R_1(f) = \nu(\tfrac{1}{2}(\eta - x))\exp(\tfrac{j}{2}(3\eta - x))$$

$$R_2(f) = \exp(\tfrac{j}{2}(\tfrac{5\eta}{2} - x))\Big[J_o(\tfrac{\eta}{2\pi})\nu(\tfrac{1}{2}(\tfrac{\eta}{2} - x))$$

$$+ \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{2\pi})(-1)^n\Big\{\nu(\tfrac{\eta}{4} + \pi n - \tfrac{x}{2}) + \nu(\tfrac{\eta}{4} - \pi n - \tfrac{x}{2})\Big\}$$

$$- j\sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{2\pi})(-1)^n\Big\{\nu(\tfrac{\eta}{4} + \tfrac{(2n-1)}{2}\pi - \tfrac{x}{2})$$

$$+ \nu(\tfrac{\eta}{4} - \tfrac{(2n-1)}{2}\pi - \tfrac{x}{2})\Big\}\Big]$$

$$R_3(f) = \nu(x/2)\exp(\tfrac{j}{2}(\eta - x))$$

$$R_4(f) = \exp(\tfrac{j}{2}(\tfrac{\eta}{2} - x))\Big[J_o(\tfrac{\eta}{2\pi})\nu(\tfrac{1}{2}(\tfrac{\eta}{2} + x))$$

$$+ \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{2\pi})(-1)^n\Big\{\nu(\tfrac{\eta}{4} - \pi n + \tfrac{x}{2}) + \nu(\tfrac{\eta}{4} + \pi n + \tfrac{x}{2})\Big\}$$

$$-j \sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{2\pi})(-1)^n\Big\{\nu(\tfrac{\eta}{4} + \tfrac{(2n-1)}{2}\pi - \tfrac{x}{2})$$

$$+ \nu(\tfrac{\eta}{4} - \tfrac{(2n-1)}{2}\pi - \tfrac{x}{2})\Big\}\Big] \qquad (A.11)$$

9. Figure 3.2i

$$R_1(f) = \nu(\tfrac{1}{2}(\eta - x))\exp(\tfrac{j}{2}(3\eta - x))$$

$$R_2(f) = \exp(\tfrac{j}{2}(\tfrac{7\eta}{3} - x))\Big[J_o(\tfrac{\eta}{\pi})\nu(\tfrac{1}{2}(\tfrac{\eta}{3} - x))$$

$$+ \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{\pi})\Big\{\exp(\tfrac{j2\pi n}{3})\nu(\tfrac{\eta}{6} + \tfrac{2\pi n}{3} - \tfrac{x}{2})$$

$$+ \exp(\tfrac{-j2\pi n}{3})\nu(\tfrac{\eta}{6} - \tfrac{2\pi n}{3} - \tfrac{x}{2})\Big\}$$

$$+ \sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{\pi})\Big\{\exp(\tfrac{j(2n-1)\pi}{3})\nu(\tfrac{\eta}{6} + \tfrac{(2n-1)\pi}{3} - \tfrac{x}{2})$$

$$- \exp(\tfrac{-j(2n-1)\pi}{3})\nu(\tfrac{\eta}{6} - \tfrac{(2n-1)\pi}{3} - \tfrac{x}{2})\Big\}\Big]$$

$$R_3(f) = \exp(\tfrac{j}{2}(\eta - x))\Big[J_o(\tfrac{\eta}{\pi})\nu(\tfrac{1}{2}(\tfrac{\eta}{3} - x))$$

$$+ \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{\pi})\Big\{\nu(\tfrac{\eta}{6} + \tfrac{2\pi n}{3} - \tfrac{x}{2}) + \nu(\tfrac{\eta}{6} - \tfrac{2\pi n}{3} - \tfrac{x}{2})\Big\}$$

$$- \sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{\pi})\Big\{\nu(\tfrac{\eta}{6} + \tfrac{(2n-1)\pi}{3} - \tfrac{x}{2})$$

$$- \nu(\tfrac{\eta}{6} - \tfrac{(2n-1)\pi}{3} - \tfrac{x}{2})\Big\}\Big]$$

$$R_4(f) = \exp(\tfrac{j}{2}(\tfrac{\eta}{3} - x))\Big[J_o(\tfrac{\eta}{\pi})\nu(\tfrac{1}{2}(\tfrac{\eta}{3} + x))$$

$$+ \sum_{n=1}^{\infty} J_{2n}(\tfrac{\eta}{\pi})\Big\{\exp(\tfrac{j4\pi n}{3})\nu(\tfrac{\eta}{6} - \tfrac{2\pi n}{3} - \tfrac{x}{2})$$

$$+ \exp(\tfrac{-j4\pi n}{3})\nu(\tfrac{\eta}{6} + \tfrac{2\pi n}{3} + \tfrac{x}{2})\Big\}$$

$$+ \sum_{n=1}^{\infty} J_{2n-1}(\tfrac{\eta}{\pi})\Big\{\exp(\tfrac{j(2n-1)\cdot 2\pi}{3})\nu(\tfrac{\eta}{6} - \tfrac{(2n-1)\pi}{3} + \tfrac{x}{2})$$

$$- \exp(\tfrac{-j(2n-1)2\pi}{3})\nu(\tfrac{\eta}{6} + \tfrac{(2n-1)\pi}{3} + \tfrac{x}{2})\Big\}\Big]$$

(A.12)

10. Second Order Encoding Polynomial with Rectangular Pulse Shaping

The expressions for the elements of the vector $\underline{R(f)}$ for an encoding PRS polynomial $(1+k_1D+k_2D^2)/(1+|k_1|+|k_2|)$ with rectangular pulse shaping are given in this section. They are referred to in Section 2.5 of Chapter 2.

$$\left.\begin{aligned}
R_1(f) &= \nu\left(\tfrac{1}{2}(\eta - x)\right)\exp\left(\tfrac{j}{2}\left(\frac{5+3k_1+k_2}{1+k_1+k_2}\eta - x\right)\right) \\
R_2(f) &= \nu\left(\tfrac{1}{2}\left(\frac{-1+k_1+k_2}{1+k_1+k_2}\eta - x\right)\right)\exp\left(\tfrac{j}{2}\left(\frac{3+3k_1+k_2}{1+k_1+k_2}\eta - x\right)\right) \\
R_3(f) &= \nu\left(\tfrac{1}{2}\left(\frac{1-k_1+k_2}{1+k_1+k_2}\eta - x\right)\right)\exp\left(\tfrac{j}{2}(\eta - x)\right) \\
R_4(f) &= \nu\left(\tfrac{1}{2}\left(\frac{-1-k_1+k_2}{1+k_1+k_2}\eta - x\right)\right)\exp\left(\tfrac{j}{2}\left(\frac{-1+k_1+k_2}{1+k_1+k_2}\eta - x\right)\right)
\end{aligned}\right\} \quad (A.13)$$

APPENDIX B

EQUIVALENT ENCODING PRS POLYNOMIALS

It is shown in this appendix that reciprocal encoding polynomials $\frac{1}{C}\sum_{i=0}^{m} k_i D^i$ and $\frac{1}{C}\sum_{i=0}^{m} k_{m-i} D^i$ yield identical spectra and minimum Euclidean distances when the shaping pulse, s(t), in Figure 2.1 is symmetrical. As mentioned in Chapter 2, a correlative encoding is completely specified by its equivalent baseband pulse, b(t), given by (2.18). Let the modulated signal, x(t), and the equivalent baseband pulse, b(t), be denoted by $x_1(t)$ and $b_1(t)$ for the polynomial $\frac{1}{C}\sum_{i=0}^{m} k_i D^i$ and by $x_2(t)$ and $b_2(t)$ for the reciprocal polynomial $\frac{1}{C}\sum_{i=0}^{m} k_{m-i} D^i$. Then

$$x_1(t) = \cos\{2\pi f_c t + \int_{-\infty}^{t} \sum_{k=-\infty}^{\infty} a_k b_1(u-kT)du\} \qquad (B.1)$$

where

$$b_1(t) = \frac{1}{C}\sum_{i=0}^{m} k_i\, s(t-iT)$$

And

$$x_2(t) = \cos\{2\pi f_c t + \int_{-\infty}^{t} \sum_{k=-\infty}^{\infty} a_k b_2(u-kT)du\} \qquad (B.2)$$

where

$$b_2(t) = \frac{1}{C} \sum_{i=0}^{m} k_{m-i}\, s(t-iT)$$

The shaping pulse, s(t), in the above equations extends over the interval [0,qT] and is assumed to be symmetrical.

$$s(t) = s(-t + qT) \tag{B.3}$$

Thus, $b_2(t)$ is a time reversal of $b_1(t)$ but for a time shift

$$b_2(t) = b_1(-t + (m+q)T) \tag{B.4}$$

Therefore, $x_2(t)$ can be written as

$$x_2(t) = \cos\{2\pi f_c t + \int_{-\infty}^{t} \sum_{k=-\infty}^{\infty} a_k b_1(-u + kT + (m+q)T)du\} \tag{B.5}$$

If the carrier frequency, f_c, is large compared to the bit rate, the autocorrelation function for $x_1(t)$ can be written as [31]

$$\begin{aligned} R_{x_1}(\tau) &= \langle x_1(t)\, x_1(t+\tau)\rangle \\ &= \frac{1}{2} \langle \cos\{2\pi f_c \tau + \int_{t}^{t+\tau} \sum_{k=-\infty}^{\infty} a_k b_1(u-kT)du\}\rangle \end{aligned} \tag{B.6}$$

where the angle brackets denote the time average. If we assume the modulating signal to be ergodic we can replace the time

average by an ensemble average. Then

$$R_{x_1}(\tau) = \frac{1}{2} E[\cos\{2\pi f_c \tau + \int_t^{t+\tau} \sum_{k=-\infty}^{\infty} a_k b_1(u-kT)du\}] \quad (B.7)$$

where E[ ] denotes the ensemble average. Similarly

$$R_{x_2}(\tau) = \frac{1}{2} E[\cos\{2\pi f_c \tau + \int_t^{t+\tau} \sum_{k=-\infty}^{\infty} a_k b_1(-u+kT+(m+q)T)du\}] \quad (B.8)$$

After a change of variable in (B.8), $R_{x_2}(\tau)$ can be written as

$$R_{x_2}(\tau) = \frac{1}{2} E[\cos\{2\pi f_c \tau + \int_t^{t+\tau} \sum_{\ell=-\infty}^{\infty} a_{-\ell-(m+q)} b_1(-u-\ell T)du\}]$$

$$= \frac{1}{2} E[\cos\{2\pi f_c \tau + \int_{-t-\tau}^{-t} \sum_{\ell=-\infty}^{\infty} a_{-\ell-(m+q)} b_1(u-\ell T)du\}]$$

$$= \frac{1}{2} E[\cos\{2\pi f_c \tau + \int_{t'}^{t'+\tau} \sum_{k=-\infty}^{\infty} a_{-k-(m+q)} b_1(u-kT)du\}] \quad (B.9)$$

The random variables a_k are independent, identically distributed binary random variables. Therefore, the statistical properties of $R_{x_2}(\tau)$ remain unchanged if the $a_{-k-(m+q)}$ are replaced by a_k. The modulated signals $x_1(t)$ and $x_2(t)$ are

cyclostationary processes. If the usual procedure of time-averaging over one bit period is carried out, $R_{x_2}(\tau)$ is independent of t' in (B.9) and $R_{x_1}(\tau)$ is independent of t in (B.7) [62]. Thus, $R_{x_1}(\tau)$ and $R_{x_2}(\tau)$ are identical. This proves the claimed result on equivalent spectra.

We next show that reciprocal encoding polynomials yield identical minimum Euclidean distances when the signal is observed over a long duration before the decision is made. The Euclidean distance, $d_{j,k}$, between the signals $s(t,1,A_j)$ and $s(t,-1,A_k)$ was given by equation (4.21)

$$d_{j,k}^2 = NT - \int_{LT}^{(L+N)T} \cos \{ \int_{LT}^{t} \sum_{i=-\infty}^{\infty} e_i b(u-iT)du \}dt \qquad (B.10)$$

The relationship between the data sequences of the signals $s(t,1,A_j)$ and $s(t,-1,A_k)$ and the sequence $E_k = \ldots .0,0,e_o, e_1,\ldots$ is given by (4.20). Since, the distance does not depend on LT in (B.10), we will set LT to zero for convenience. Let the sequence $E_k = \ldots .0,0,e_o,e_1,\ldots$ be such that

$$\int_{0}^{t} \sum_{i=-\infty}^{\infty} e_i b(u-iT)du = 2\pi n \quad \text{for } t \geq MT \qquad (B.11)$$

where M and n are integers. If the above condition is satisfied the Euclidean distance, $d_{j,k}$, in (B.10) does not increase for values of N greater than M. When the sequence E_k satisfies (B.11), the phase paths of $s(t,1,A_j)$ and $s(t,-1,A_k)$ merge at t = MT and the distance between them provides an upper bound on the minimum Euclidean distance, d_{min}. From (B.11) it

follows that

$$\sum_{i=-\infty}^{\infty} e_i b(u-iT) \equiv 0 \qquad \text{for } u \geq MT \tag{B.12}$$

The baseband pulse, b(t), extends over (m+q) bit intervals. Therefore, all the elements of the sequence $E_k = \ldots .0,0,e_o,e_1,\ldots e_i,e_{i+1},\ldots$ for $i > J = \{M-(m+q)\}$ should be zero to satisfy (B.12). Therefore

$$2\pi n = \int_0^{MT} \sum_{i=0}^{J} e_i\, b(u-iT)du$$

$$= \sum_{i=0}^{J} e_i \int_0^{MT} b(u-iT)du$$

$$= \sum_{i=0}^{J} e_i B \tag{B.13}$$

where

$$B = \int_0^{(m+q)T} b(u)du$$

$$= \int_{-\infty}^{\infty} b(u)du \tag{B.14}$$

The equation (B.13) gives the condition for a phase merge.

For example, the sequence $E_k = \ldots .0,0,2,-2,0,0,\ldots$ leads to a phase merge at $t = (m+q+1)T$ with a corresponding finite Euclidean distance. Thus, d_{min} is always finite, even if the observation interval is infinite.

The argument for the claimed result is as follows. The phase paths of the two signals which decide d_{min} must merge, otherwise the distance between the signals can be increased by extending the observation interval. Thus, when the observation interval is allowed to be arbitrarily long, the d_{min} is the smallest distance between all signals that have merging paths. It is shown that for every merging phase path for the polynomial $\frac{1}{C}\sum_{i=0}^{m} k_i D^i$, there is a corresponding phase merge for the reciprocal polynomial $\frac{1}{C}\sum_{i=0}^{m} k_{m-i} D^i$ with identical distance. Therefore, reciprocal polynomials yield the same d_{min}.

The equivalent baseband pulses for the reciprocal polynomials are given by $b_1(t)$ and $b_2(t)$. From (B.4) it can be seen that

$$\int_{-\infty}^{\infty} b_1(t)dt = \int_{-\infty}^{\infty} b_2(t)dt = B \qquad (B.15)$$

Let a sequence E_1 be given by

$$E_1 = \ldots .0,0,e_o,e_1,\ldots .e_J,0,0,\ldots \qquad (B.16)$$

where

$$\sum_{i=0}^{J} e_i B = 2\pi n \qquad (B.17)$$

The sequence E_1 corresponds to a phase merge and the distance for the polynomial $\frac{1}{C}\sum_{i=0}^{m} k_i D^i$ is given by

$$d^2_{j,k} = MT - \int_0^{MT} \cos\{\int_0^t \sum_{i=0}^{J} e_i b_1(u-iT)du\}dt \qquad (B.18)$$

where, M and J are related by

$$J = M - (m+q) \qquad (B.19)$$

Now, consider another sequence E_2 given by

$$E_2 = \ldots 0,0,e_J,e_{J-1},\ldots e_1,e_o,0,0, \qquad (B.20)$$

This sequence also satisfies (B.17) and hence corresponds to a phase merge. The Euclidean distance for this phase merge for the reciprocal polynomial $\frac{1}{C}\sum_{i=0}^{m} k_{m-i} D^i$ is given by

$$d^2_{j,k} = MT - \int_0^{MT} \cos\{\int_0^t \sum_{i=0}^{J} e_{J-i} b_2(u-iT)du\}dt \qquad (B.21)$$

After changes in variables i and t in the above expression, it can be written as

$$d^2_{j,k} = MT - \int_0^{MT} \cos\{\int_0^{MT-t} \sum_{i=0}^{J} e_i b_2(u-JT+iT)du\}dt \qquad (B.22)$$

And also

$$\int_{0}^{MT-t} \sum_{i=0}^{J} e_i b_2(u-JT+iT)du$$

$$= 2\pi n - \int_{MT-t}^{MT} \sum_{i=0}^{J} e_i b_2(u-JT+iT)du \qquad (B.23)$$

Therefore

$$d^2_{j,k} = MT - \int_{0}^{MT} \cos\{\int_{MT-t}^{MT} \sum_{i=0}^{J} e_i b_2(u-JT+iT)du\}dt \qquad (B.24)$$

Substituting $b_2(u)$ in terms of $b_1(u)$ as given by (B.4) and from (B.19), we get

$$d^2_{j,k} = MT - \int_{0}^{MT} \cos\{\int_{MT-t}^{MT} \sum_{i=0}^{J} e_i b_1(-u+MT-iT)du\}dt$$

$$= MT - \int_{0}^{MT} \cos\{\int_{0}^{t} \sum_{i=0}^{J} e_i b_1(u-iT)du\}dt \qquad (B.25)$$

The above expression is identical to (B.18). This proves that for every phase merge E_1 for the polynomial $\frac{1}{C} \sum_{i=0}^{m} k_i D^i$, there is a corresponding merge E_2 for the reciprocal polynomial $\frac{1}{C} \sum_{i=0}^{m} k_{m-i} D^i$, with identical Euclidean distance.

APPENDIX C

PROLATE SPHEROIDAL WAVE FUNCTIONS

Slepian, Pollak and Landau [59,60] have made an extensive study of prolate spheroidal wave functions and their application to signal theory. Following is a summary of the properties of these functions relevant to this study, using their notation.

Let L_∞^2 be the class of real functions which are square integrable on the interval $(-\infty,\infty)$

$$\int_{-\infty}^{+\infty} f^2(t)dt < \infty \tag{C.1}$$

Therefore, all the functions f(t) in L_∞^2 possess a Fourier transform whose inverse is the original function.

$$F(t) = \int_{-\infty}^{+\infty} f(t)e^{-j2\pi ft}\, dt \tag{C.2}$$

$$f(t) = \int_{-\infty}^{+\infty} F(f)e^{j2\pi ft}\, df \tag{C.3}$$

Let B be the subclass of L_∞^2 consisting of those f(t), which are strictly bandlimited to W Hz. That is

$$F(f) = 0 \qquad \text{if} \qquad |f| > W \tag{C.4}$$

Let D be the subclass of L_∞^2 consisting of those functions f(t), which are strictly time-limited to the interval $[-\frac{T'}{2}, \frac{T'}{2}]$, that is

$$f(t) = 0 \qquad \text{if} \qquad |t| > \frac{T'}{2} \tag{C.5}$$

For any $T' > 0$ and any $W > 0$, an infinite set of real functions $\psi_0(t), \psi_1(t), \ldots$ and a set of real positive numbers $\lambda_0 > \lambda_1 > \lambda_2 > \ldots$ can be found with the following properties

i) The $\psi_i(t)$ are bandlimited to W Hz, are orthonormal over the interval $(-\infty,\infty)$ and are complete in B.

$$\int_{-\infty}^{\infty} \psi_i(t)\psi_j(t)dt = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{C.6}$$

$$i, j = 0,1,2\ldots$$

ii) In the interval $[-\frac{T'}{2}, \frac{T'}{2}]$, the $\psi_i(t)$ are orthogonal and complete in D.

$$\int_{-T'/2}^{T'/2} \psi_i(t)\psi_j(t)dt = \begin{cases} 0 & i \neq j \\ \lambda_i & i = j \end{cases} \tag{C.7}$$

$$i, j = 0,1,2\ldots$$

iii) The $\psi_i(t)$ are known as spheroidal wave functions, and can be expressed as the solution to the integral equation

$$\lambda_i \psi_i(t) = \int_{-T'/2}^{T'/2} \frac{\sin 2\pi W(t - s)}{\pi(t - s)} \psi_i(s) ds \qquad (C.8)$$

$$i = 0,1,2 \ldots$$

$\psi_i(t)$ can also be expressed as the solution to the differential equation

$$(1 - t^2) \frac{d^2\psi_i(t)}{dt^2} - 2t \frac{d\psi_i(t)}{dt} + (\chi_i - c^2t^2)\psi_i(t) = 0 \qquad (C.9)$$

with the initial conditions

$$\left.\begin{aligned} \psi_i(t)\Big|_{t=0} &= 0 \\ \frac{d\psi_i(t)}{dt}\Big|_{t=0} &= \text{constant} \end{aligned}\right\} \quad \text{if i is odd}$$

$$\left.\begin{aligned} \psi_i(t)\Big|_{t=0} &= \text{constant} \\ \frac{d\psi_i(t)}{dt}\Big|_{t=0} &= 0 \end{aligned}\right\} \quad \text{if i is even}$$

The parameters λ_i and χ_i in the above equations are functions of a parameter c, which is proportional to the time-bandwidth product $c = \pi WT'$. The spheroidal wave function, $\psi_i(t)$, is an odd function of t if i is odd and even for i even.

If we chose z(t), a member of the subspace D, such that

$$\begin{aligned} z(t) &= \psi_i(t) && |t| \leq \frac{T'}{2} \\ &= 0 && \text{elsewhere} \end{aligned} \tag{C.10}$$

The fraction of energy of z(t) in a bandwidth of W Hz is λ_i. Since $\lambda_0 > \lambda_1 > \lambda_2 > \ . \ . \ .$, it follows that within the subspace D, $\psi_0(t)$ has the largest fraction of energy in a bandwidth of W Hz.