ADAPTIVE MULTICHANNEL RECEPTION OF BINARY SIGNALS

by

B. Shawaf

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of Califirnia Department of Engineering Los Angeles, California 90024	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

ADAPTIVE MULTICHANNEL RECEPTION OF BINARY SIGNALS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)

B. Shawaf

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
July 1969	123	

8a. CONTRACT OR GRANT NO. AF-AFOSR 68-1408 b. PROJECT NO. 9749-01 c. 6144501F d. 681304	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) **AFOSR 69-1821TR**

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES **THESIS** **University Of California, Los Angeles**	12. SPONSORING MILITARY ACTIVITY Air Force Office of Scientific Research (SRMA) 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

The dissertation considers the optimization problem of adaptive multichannel reception of binary signals which are corrupted multiplicatively by random amplitude parameters with no known a priori distributions, and additively by white Gaussian noise processes of known spectral densities which we allow to be different among the channels.
A distinction is made between two modes of operation: "multiple processing," in which the unknown amplitude parameters are estimated and used "adaptively" to adjust weighting factors in a composite decision rule, and based on this, a decision is made within each observation interval; the other, in which the processing is carried out independently at each receiver, arriving at an ultimate decision using "majority logic" combining. This latter mode is found to yield consistently higher error probability but may be simpler in implementation than the former in that the requirement for estimation is obviated.
In the proofs we appeal to geomatrical ideas, and evaluate the error probabilities by interpreting the aspect and variations of the decision hypersurface with respect to the coordinate axes of N-dimensional Euclidean space.
The results show that under certain schemes of multiple processing operation, the error probability is a function of the total received energy over all channels normalized by their respective noise spectral densities, and of a centrality parameter which is a measure of the angle by which the amplitude parameters deviate from the case of being all equal. Under some other (continued)

DD FORM 1 NOV 65 1473

Security Classification

schemes of multiple processing, the error probability is found to be sensitive to the distribution of energy among the various channels. It is shown that multiple processing can be designed to provide better performance than majority combining even in the absence of complete knowledge of the amplitude parameters, and indeed, this improvement may be appreciable under certain circumstances.

AFOSR 69-1821TR

UNIVERSITY OF CALIFORNIA

Los Angeles

Adaptive Multichannel Reception of Binary Signals

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Engineering

by

Bassim Shawaf

Committee in charge:

Professor A.V. Balakrishnan, Chairman
Professor Jack W. Carlyle
Professor Andrew J. Viterbi
Professor John W. Green
Professor Charles J. Stone

1968

This re... supported by the Applied M... Division, AFOSR, ... under Contract/Grant AF-AFOSR 68-1408

The dissertation of Bassim Shawaf is approved,
and it is acceptable in quality and form
for publication on microfilm:

Charles J. Stone
____

Jack W. Carlyle
____

John W. Green
____

Andrew J. Viterbi
____

A. V. Balakrishnan
____
Committee Chairman

University of California, Los Angeles

1968

TO CAROLINE

Leith, Raad and Aiman

in lieu of time missed

TABLE OF CONTENTS

LIST OF SYMBOLS

$s_j(t)$	jth. signaling waveform
$n_i(t)$	noise process on ith. path
E	energy parameter
ρ_{jk}	correlation coefficient between jth. and kth. signals
D_i	power spectral density of white noise process n_i
$R_N(\sigma,t)$	autocorrelation function of noise process
a_i	amplitude parameter on ith. path - (random)
$\hat{a}_i$	estimate of a_i
$\phi_k(t)$	orthonormal functions - solutions to integral equation
$\delta(t)$	delta function
$[x,y]$	inner product
γ_j, ϵ_j	weight factors
$Q, Q_{\underline{\epsilon}}$	matrices
T	matrix - rotation of axes transformation
V_i, W_i, x_i	random variables
m_i, α_i	expected values of random variables
M, A	vectors of expected values - POLAR AXES
Λ	covariance matrix of random variables
B	vector [1, 1, . . . , 1] - MAJOR AXIS
$p(.,.)$	probability density function
$C_V(.)$	characteristic function of random variable V

$\mathrm{erfc}(.)$	complementary error function
$R^{\pm}$	a certain ratio parameter
P_e	average probability of error
η	real parameter
Z_η	composite decision rule
$\theta, \Phi, \mu, \nu, \psi$	angles
$D_n(z)$	Parabolic Cylinder function
$\Gamma(n)$	Gamma function
$h_\eta(\theta, \Phi_k, ..)$	equation of hypersurface in $E^{(N)}$
$I_{N,i}(a,b)$	a certain integral with limits a to b, defined in equation (4.60)
$\omega^{(i)}(z)$	ith. derivative of a function of z
λ	Lagrange multiplier
$w_i^{e_i}$	a certain probability
$\delta_{k,i}$	modified Kronecker delta, allowing for non-integer subscripts
$d(.)$	weight of a vector
$C(A \sim \alpha_k)$	function of elements of vector A except the kth. element
$\underline{\sigma}, \underline{\tau}$	basis vectors
$P_i(z), f_i(z), g_i(z)$	polynomial functions of z
$*$	transpose
$\lVert(.)\rVert^2$	Euclidean norm squared
$\sum$	summation sign
Π	product sign

LIST OF FIGURES

ACKNOWLEDGEMENTS

I wish to express my indebtedness to a distinguished teacher, Professor A.V. Balakrishnan for his advice and guidance during this investigation and throughout the period of my studies. The kind help and encouragement of Professors Jack Carlyle, Andrew Viterbi, John Green, and Charles Stone, of my committee, are highly appreciated.

I also wish to thank Dr. Robert Price for a brief, though most illuminating discussion.

This investigation was supported by various agencies while the author was a Post Graduate Research Engineer at the University of California at Los Angeles. Notable among them are NASA, through JPL contract: Space Communications, No. 4-482542-26279, and AFOSR contract: Information Systems, No. 4-442542-22545.

And last but not least, I wish to thank my brothers to whom I owe the elegance in the computation and graphing of the results in a short time.

VITA

December 24, 1935	Born, Baghdad, Iraq
1960 ------ B.Sc.	University of London, London, England
1960 - 1962	Communication Engineer, Iraq Petroleum Co., Iraq
1964 ------ M.S.	University of California, Los Angeles
1964 - 1968	Research Assistant and Post Graduate Research Engineer, University of California, Los Angeles
1965 - 1968	Assistant Professor and Associate Professor, California State College at Long Beach

PUBLICATIONS

"Adaptive Multichannel Reception with M'ary Signals," Jet Propulsion Laboratory, Space Program Summary, 37-43, Vol.IV, February 1967

"Adaptive Multichannel Reception of Binary Signals," Jet Propulsion Laboratory, Space Program Summary, 37-50, Vol.III, June 1968

"Adaptive Multichannel Reception of Binary Antipodal Signals," To be published, JPL - SPS

FIELDS OF STUDY

Major Field: Engineering

Studies in Communication Systems

Professor A.V. Balakrishnan

Studies in Control Systems

Professor C.T. Leondes

Studies in Applied Mathematics

Professor A.J.A. Morgan

ABSTRACT OF THE DISSERTATION

Adaptive Multichannel Reception
of Binary Signals

by

Bassim Shawaf

Doctor of Philosophy in Engineering
University of California, Los Angeles, 1968
Professor A.V. Balakrishnan, Chairman

The dissertation considers the optimization problem of adaptive multichannel reception of binary signals which are corrupted multiplicatively by random amplitude parameters with no known a priori distributions, and additively by white Gaussian noise processes of known spectral densities which we allow to be different among the channels.

A distinction is made between two modes of operation: "multiple processing," in which the unknown amplitude parameters are estimated and used "adaptively" to adjust weighting factors in a composite decision rule, and based on this, a decision is made within each observation interval; the other, in which the processing is carried out independently at each receiver, arriving at an

ultimate decision using "majority logic" combining. This latter mode is found to yield consistently higher error probability but may be simpler in implementation than the former in that the requirement for estimation is obviated.

In the proofs we appeal to geometrical ideas, and evaluate the error probabilities by interpreting the aspect and variations of the decision hypersurface with respect to the coordinate axes of N-dimensional Euclidean space.

The results show that under certain schemes of multiple processing operation, the error probability is a function of the total received energy over all channels normalized by their respective noise spectral densities, and of a centrality parameter which is a measure of the angle by which the amplitude parameters deviate from the case of being all equal. Under some other schemes of multiple processing, the error probability is found to be sensitive to the distribution of energy among the various channels. It is shown that multiple processing can be designed to provide better performance than majority combining even in the absence of complete knowledge of the amplitude parameters, and indeed, this improvement may be appreciable under certain circumstances.

CHAPTER I

INTRODUCTION

Whereas the subject of self-adjusting or "adaptive" systems has been widely discussed in the literature insofar as it pertains to Control systems, the subject of adaptive Communication has not found as wide an exposure. At least, adaptivity,where it relates to Communication systems has been a fairly recent phenomenon. The most noteworthy of these exposures are Price's [16] and Balakrishnan's [2] not to exclude the many other workers in the field.

1.1 THE ADAPTIVE COMMUNICATION PROBLEM

The problem of optimum reception of data over transmission media which are subject to random disturbances has been largely discussed. The solution to the problem of combating fading in a corrupted signal is well known and consists of using several replicas of the transmitted signal which have been subject to independent disturbances, in order to achieve, by some processing scheme, a consistently better performance rate. These processing schemes may be classified under the general title of "diversity reception" [19].

The concept of an adaptive communication system may be described as a system which monitors its own performance and adjusts its structure in some fashion to optimize its performance according to some criterion of performance chosen a priori. The criterion we consider is the minimization of the error probability. Central to the idea of adaptivity is, therefore, the simultaneous detection of the signal, and the estimation of the unknown parameter(s) in the received waveforms, which reflect the unknown channel parameter(s). These estimates are then used in the signals processing as if they had perfect accuracy.

Middleton[17] and Middleton and Esposito [18] consider the problem of simultaneous detection and estimation of signal parameters in noise from the point of view of radar detection and with the criterion of minimizing the average Bayes risk.

Balakrishnan's thorough tutorial paper [2] discusses the theory of adaptive methods in communication systems and defines the characteristics and features a system must possess in order to be termed adaptive. He considers three categories of adaptive communication systems. One, is where the adaptive feature is at the transmitter, an example of this is a system which recognizes loss of performance (by feedback possibly) and adjusts its mode of operation, by repeating the data. Another is where the adaptive feature is at the receiver, an example of which is the present investigation. A third, is a combination of both previous categories. This system recognizes for example, redundant data by a self-monitoring function and adjusts its transmission by suppressing the redundant data. The theoretical evaluation of such a system has been tackled by Davisson [22], and in subsequent publications [23] and [24].

Price [16], described the operation of an adaptive receiver for binary signals over a multichannel link. He considers, as we do, a pair of equal energy waveforms that are either orthogonal or antipodal. He determines the probability of decision error as a function of the total energy received over all channels, which he finds independent of the distribution of signal strength among the channels. His estimates of the channel parameters go over several past observation intervals. Subsequently, Hingorani [26], [27], Bello [28] and Proakis [29], [30] and [32] found similar expressions for the error probability for multichannel binary signaling over channels characterized by specular and Rayleigh-fading components.

1.2 THE PROBLEM AND OBJECTIVES

This study is devoted to a theoretical investigation of the adaptive communication problem discussed by Balakrishnan

[2]. We consider this as a major reference in that it is the point of departure for this study, whose main objective is the adaptive optimization of the reception of binary information via a multiple channel link.

We treat the problem with a particular view to multiple receiver operation (space diversity), although the analysis may easily be generalized to other forms of diversity.

The specific problem we consider is a P.C.M. system which uses either of two modes of transmission PCM-FM or PCM-PM. Within each interval [0,T] it transmits a signal of the form,

$$s(t) = S \cos [\omega_c t + M(t)]$$

where M(t) is a narrow-band signal proportional to the modulating waveform. This is received at the i^{th} receiver coherently, in the presence of white Gaussian noise $n_i(t)$ of known power spectral density D_i; in addition, it is subject to a random multiplicative change in amplitude due to channel fading.

In view of the advanced stage of development of phase-locked loops [25], the assumption that the reception is coherent, is made justifiable. Thus the received waveforms may be expressed as,

$$y_i(t) = a_i \; s_j(t) + n_i(t) \qquad 0 \leq t \leq T$$
$$j = 1,2 \; ; \; i = 1,\ldots,N$$

One of the objectives of the study is the derivation of the optimum decision and estimation rules, which are then used to evaluate the performance of the system under two basic modes of operation: "multiple processing" in which the set of received waveforms over all channels is considered as a multiple stochastic process; and "majority logic" combining, in which a decision is made at each receiver independently of the others, and using a majority combining scheme to arrive at a decision within each observation interval.

We utilize largely geometric ideas in the evaluation of system performance. We find the approach both appealing and instructive to explain the major divergence in conclusion with the work of Price [16]. One major conclusion of this study is that the probability of error is not only a function of the total normalized energy received over all N-channels as Price concludes, but it, generally, is a function of how this energy is distributed among the various channels. Our conclusions coincide with those of Price only in two specific cases

(i) that the signal strngths (amplitude parameters) are known exactly and without any uncertainty for each channel.

(ii)that they are unknown in magnitude, but they happen to be equal to each other.

Considering, though, that the actual magnitudes are immaterial, as we shall presently see, but all that is required is their ratios, case (iii) above reduces to case (ii). With any uncertainty as to the ratios of the amplitude parameters, our conclusion is that the probability of error is a function of how the total received energy is distributed among the channels.

The other major conclusion is that although multiple processing is superior to majority combining, the degree of weighting used in the adaptive decision rule is of little significance, and it turns out that, at least where antipodal signals are concerned, adaption does not consistently yield better performance.

1.3 ORGANIZATION OF THE DISSERTATION

This investigation is organized into six chapters and three appendices.

In chapter II, we derive the optimum decision rule and the optimal estimates of the unknown amplitude parameters. We derive the estimates within each observation interval, and use them in the decision rule to form what shall be referred to in the sequel as the composite decision rule.

In chapter III, we utilize the forms of the composite decision rule to undertake an error analysis of the problem and to obtain a simple bound on the average probability of error.

Chapter IV, is concerned with the multiple processing mode for antipodal signals. We present a detailed discussion of the "estimation" phase of the problem and the resulting composite decision rule. The error probability is obtained by geometrically interpreting the aspect and variations of the decision hypersurface in Euclidean space.

The chapters, II through IV pertain to the "multiple processing" mode of operation. The "majority logic" combining mode is the subject of chapter V, in which we present the conditions under which majority logic yields the maximum and minimum probability of error.

We present a few miscellaneous results and the summary of conclusions in chapter VI.

In chapter IV, we utilize a generalized spherical coordinate transformation; a summary of which appears in Appendix "A".

In the same chapter, we discuss an example in which a difference in probability content arises. We discuss this difference and its asymptotic behavior in Appendix "B".

In Appendix "C", we list some recursion formulae which are used in the evaluation of a certain integral.

CHAPTER II

MATHEMATICAL STATEMENT OF THE PROBLEM

2.1 INTRODUCTION

In this chapter we shall describe, mathematically, the problem to be considered, and set forth the assumptions which will enter in the derivation of the decision processing.

Based on the discussion of the introductory chapter, in which the idea of diversity was described; we shall, in this chapter, examine the problem with a particular view to receiver diversity. Nothing in the derivations, however, will prohibit its extension to other forms of diversity.

Much of the initially derived results will be based on considering general M'ary signaling. The optimum decision and estimation rules arrived at in this fashion will then be specified to the binary signaling case under investigation.

The initial formulation, due to Balakrishnan [2] and [3] has already appeared in the literature. We present it here for the sake of establishing uniform notation and consistent terminology.

2.2 DESCRIPTION OF THE PROBLEM

We consider, for example, a PCM system employing "space" or "receiver" diversity. That is to say, we consider the problem of communication by one of a set of M-possible equal energy signaling waveforms $\{s(t)\}$, via a multiple link (or channel or receiver) system, and received coherently at the N receiving terminals (or stations or antennas).

Specifically, we consider that we have available a set $\{Y(t)\}$ of N corrupted versions of a transmitted signal $s_j(t)$, where each component of the set is of the form,

$$y_i(t) = a_i\, s_j(t) + n_i(t) \quad ; \; 0 \leq t \leq T \qquad (2.1)$$
$$i = 1, \ldots, N$$
$$j = 1, \ldots, M$$

in which $y_i(t)$ is the received waveform at the i^{th} receiver and $s_j(t)$ is the transmitted information-bearing signal, whose waveform is known at the receiver. The transmitted signal is subjected to a multiplicative disturbance a_i, and to an additive noise process $n_i(t)$ on the i^{th} channel.

We take the set $\{N(t)\}$ of the noise processes to be composed of components of white Gaussian noise processes, independent of each other, and with known power spectral densities D_i, and we allow the components $\{D_i\}$ to be unequal.

The transmitted signals $s_j(t)$, $j = 1,\ldots,M$ will be taken to be equal energy signals, where the energy, E, is defined by,

$$E = [s_j, s_j] \tag{2.2}$$

and we take their correlation coefficients ρ_{ij} to be defined by,

$$\rho_{ij} = [s_i, s_j] / E \tag{2.3}$$

wherein, the inner product notation,

$$[f,g] = \int_0^T f(t)\, g(t)\, dt$$

has been used.

We take the viewpoint, initially that the amplitude factors $\{a_i\}$ are known, and proceed to derive the optimal receiver structure.

2.3 OPTIMAL RECEIVER STRUCTURE

The optimal receiver which minimizes the probability of error using all the available versions of $y_i(t)$ can be derived easily by considering the waveform set $\{Y(t)\}$ as a "multiple stochastic process" [1]. Thus setting,*

$$\begin{aligned} Y(t) &= [y_1(t),\ y_2(t),\ \ldots\ ,\ y_N(t)] \\ S_j(t) &= [a_1 s_j(t),\ a_2 s_j(t),\ \ldots\ ,\ a_N s_j(t)] \\ N(t) &= [n_1(t),\ n_2(t),\ \ldots\ ,\ n_N(t)] \end{aligned} \tag{2.4}$$

*Here, and in the sequel, all vectors shall be understood to be column vectors.

Let,

$$R_N(\sigma,t) = E\,[\,N(\sigma)\,N^*(t)\,] \tag{2.5}$$

the * denoting conjugate transpose.

Then, with $[\,\phi_k(t)\,]$, the set of (vector) orthonormal functions corresponding to the covariance kernel $R_N(\sigma,t)$ of the integral equation,

$$\int_0^T R_N(\sigma,t)\,\phi_k(t)\,dt = \lambda_k\,\phi_k(\sigma) \tag{2.6}$$

it is well known that each member of the signal set may be expressed as,

$$S_j(t) = \sum_{k=1}^{\infty} b_{jk}\,\phi_k(t) \tag{2.7}$$

Ordinarily, $b_{jk} = 0$ for $k > M$, so that the signals $S_j(t)$ have a finite number of degrees of freedom, M.

In terms of the same set of orthonormal functions $\{\,\phi_k(t)\}$, one may write,

$$Y(t) = \sum_{k=1}^{\infty} y_k\,\phi_k(t) \tag{2.8}$$

$$N(t) = \sum_{k=1}^{\infty} n_k\,\phi_k(t) \tag{2.9}$$

where,

$$y_k = \int_0^T Y^*(t)\,\phi_k(t)\,dt \tag{2.10}$$

$$n_k = \int_0^T N^*(t)\,\phi_k(t)\,dt \tag{2.11}$$

thus,

$$y_k = b_{jk} + n_k \quad ;\ k = 1, \ldots, M \tag{2.12}$$

so that we have a reduction to a countable number of coordinates.

Assuming now, the signals $s_j(t)$ are equally likely, consistent with their being equal energy; it follows that we should decide on the m^{th} signal if,

$$\sum_k (y_k - b_{mk})^2/\lambda_k = \min_j \sum_k (y_k - b_{jk})^2/\lambda_k \qquad (2.13)$$

or, equivalently,

$$\sum_k b_{mk}^2/\lambda_k - 2\sum_k y_k b_{mk}/\lambda_k = \min_j \sum_k b_{jk}^2/\lambda_k - 2\sum_k y_k b_{jk}/\lambda_k \qquad (2.14)$$

or, in terms of the time functions,

$$\int_0^T h_m^*(t) S_m(t)\, dt - 2\int_0^T h_m^*(t) Y(t)\, dt \qquad (2.15)$$

$$= \min_j \int_0^T h_j^*(t) S_j(t)\, dt - 2\int_0^T h_j^*(t) Y(t)\, dt$$

where $h_j(t)$ is the solution to,

$$\int_0^T R_N(\sigma, t)\, h_j(t)\, dt = S_j(\sigma) \qquad (2.16)$$

In the case we consider, where the noise processes are white Gaussian, the covariance matrix is,

$$R_N(\sigma, t) = D \;\; \delta(\sigma - t) \qquad (2.17)$$

where D is a diagonal matrix whose entries are the positive power spectral densities D_i, $i = 1, \ldots, N$; corresponding to the possibly different noise temperatures at the receivers.

Hence, the solution to (2.16) is,

$$h_j(t) = \left| \frac{a_1}{D_1} s_j(t), \frac{a_2}{D_2} s_j(t), \ldots, \frac{a_N}{D_N} s_j(t) \right| \qquad (2.18)$$

and we have,

$$\int_0^T h_j^*(t)\, S_j(t)\, dt = E \sum_{i=1}^{N} \frac{a_i^2}{D_i} \quad ; \; j=1,\ldots,M \tag{2.19}$$

since we assume equal energy signals.

Substituting into our decision rule (2.15), they simplify to: decide on the m^{th} signal if,

$$\int_0^T Y^*(t)\, D^{-1} S_m(t)\, dt = \max_j \int_0^T Y^*(t)\, D^{-1} S_j(t)\, dt \tag{2.20}$$

or equivalently,

$$\int_0^T \left[\sum_{i=1}^{N} \frac{a_i}{D_i}\, y_i(t)\right] s_m(t)\, dt = \max_j \int_0^T \left[\sum_{i=1}^{N} \frac{a_i}{D_i}\, y_i(t)\right] s_j(t)\, dt$$

or in terms of the inner product notation,

$$\left[\sum_{i=1}^{N} \frac{a_i}{D_i}\, y_i \, , \, s_m\right] = \max_j \left[\sum_{i=1}^{N} \frac{a_i}{D_i}\, y_i \, , \, s_j\right] \tag{2.20a}$$

The interpretation of the above decision rule, is that each received waveform $y_i(t)$ has to be weighted by the corresponding signal strength-to-noise ratio as reflected in a_i/D_i; a result which is intuitively satisfying.

Actually, the decision rule does not require the exact knowledge of all the parameters $[a_i]$; only their relative ratios enter in the weighting. Even these, however, may not be available in a practical situation. We, therefore, proceed to remove this restriction in the next section.

2.4 ADAPTIVE FEATURE

We observe that the set of amplitude factors $[a_i]$, or their relative ratios are, in a diversity system, usually unknown or are subject to variations which may be regarded as slow within the observation interval [0,T], but may vary considerably over periods longer than this.

Thus, in the absence of knowledge of the actual values of the a_i's we are led to including in the system a "learning" phase wherein the a_i's are estimated according to some criterion, and then an "adaptive" phase wherein the weighting factors a_i/D_i are adjusted according to the estimates $[\ \hat{a}_i\ ]$ of $[\ a_i\ ]$, which we use in the decision rule as if they were perfectly accurate, yielding the decision rule: decide on the m^{th} signal if,

$$[\sum_i^N \frac{\hat{a}_i}{D_i} y_i\ ,\ s_m] = \max_j\ [\sum_i^N \frac{\hat{a}_i}{D_i} y_i\ ,\ s_j] \qquad (2.21)$$

We observe briefly, in passing, that in the above decision rule (2.21), when N = 1, the rule reduces to,

$$[y_i\ ,\ s_m] = \max_j\ [y_i\ ,\ s_j] \qquad (2.22)$$

that is, for a single channel, the amplitude parameter a_i is not required for optimal processing. This suggests that if each receiver were to independently make a decision, then use some form of majority logic combining for the decisions of all the independent receivers; one may arrive at an ultimate decision as to which signal was present, without the necessity to incorporating an estimation phase. We leave the discussion of this subject, however, to chapter V.

Henceforth, we shall consider the set $[a_i]$ to be unknown, and therefore, we seek an estimate for each a_i, the form of which we discuss in the next section. For this purpose, we assume each a_i is constant (or slowly varying) within the signaling interval $[0,T]$, and we consider it to be a non-negative real number, i.e. $0 < a_i < \infty$. Moreover, no a priori distribution shall be assigned to the elements of the set of numbers $[a_i]$.

2.5 LEARNING PHASE - ESTIMATION OF $[a_i]$

Since we take the point of view that no a priori information is available (or assumed) about the set $[a_i]$, the problem of estimation of the parameters a_i may conveniently be handled by using the Maximum Likelihood criterion. This requires, setting for each i, $i=1,\ldots,N$

$$\frac{\partial}{\partial a_i} \log p(y_i \mid a_i) = 0 \qquad (2.23)$$

Based on our assumption that the signals are equally likely, and if we further assume that the a_i's are mutually independent, we have,

$$p(y_i|a_i) = \sum_{j=1}^{M} \frac{1}{M} p(y_i| a_i,s_j) \qquad (2.24)$$

but

$$p(y_i|a_i,s_j) = p(n_i = y_i - a_i s_j) \qquad (2.25)$$

and since the noise processes are white Gaussian, all this amounts to taking for each i,

$$\frac{\partial}{\partial a_i} \log\left\{ \sum_{j=1}^{M} \frac{1}{M} \exp\left[- \frac{1}{2D_i} \int_0^T (y_i(t) - a_i s_j(t))^2\, dt\right]\right\} = 0 \qquad (2.26)$$

or, since the log is monotonic,

$$\frac{\partial}{\partial a_i}\left\{ \sum_{j=1}^{M} \frac{1}{M} \exp\left[- \frac{1}{2D_i} \int_0^T (y_i(t) - a_i s_j(t))^2\, dt\right]\right\} = 0 \qquad (2.26a)$$

which yields upon simplification,

$$\sum_{j=1}^{M}\left[\int_0^T (y_i(t)-a_i s_j(t)) s_j(t) dt\right] \exp\left[- \frac{1}{2D_i}\int_0^T (y_i(t)-a_i s_j(t))^2 dt\right] = 0 \qquad (2.27)$$

Using inner product notation, and separating terms, we find that the expression which each $\hat{a}_i$ must satisfy reduces to,

$$\sum_j [y_i,s_j] \exp\left\{- \frac{1}{2D_i}([y_i,y_i]-2\hat{a}_i[y_i,s_j]+\hat{a}_i^2[s_j,s_j])\right\} =$$

$$\sum_j \hat{a}_i [s_j,s_j] \exp\left\{- \frac{1}{2D_i}([y_i,y_i]-2\hat{a}_i[y_i,s_j]+\hat{a}_i^2[s_j,s_j])\right\} \qquad (2.28)$$

and on multiplying both sides by

$$\exp\left\{ \frac{1}{2D_i}([y_i,y_i] + \hat{a}_i^2 [s_j,s_j])\right\}$$

since it is independent of j, we have,

$$\sum_{j=1}^{M}[y_i,s_j]\exp\left\{\frac{\hat{a}_i}{D_i}[y_i,s_j]\right\} = \hat{a}_i E \sum_{j=1}^{M}\exp\left\{\frac{\hat{a}_i}{D_i}[y_i,s_j]\right\} \tag{2.29}$$

or,

$$\hat{a}_i = \frac{\sum_{j=1}^{M}\exp\left\{\frac{\hat{a}_i}{D_i}[y_i,s_j]\right\}\cdot[y_i,s_j]}{E\sum_{j=1}^{M}\exp\left\{\frac{\hat{a}_i}{D_i}[y_i,s_j]\right\}} \tag{2.30}$$

i.e. the maximum likelihood estimate $\hat{a}_i$ of a_i is given by the non-trivial root of (2.30).

The above formulation and the result (2.30) are due to Balakrishnan[2]. Based on this we shall, in the sequel, modify the form of the estimate to make it more manageable for insertion in the decision rule and ultimately, to effect the error analysis.

2.6 VARIANTS ON THE FORM OF THE ESTIMATE

Denoting the quantity $\exp\left\{\frac{\hat{a}_i}{D_i}[y_i,s_j]\right\}$ by γ_{ij}, the estimate a_i of (2.30) may be written,

$$\hat{a}_i = \frac{\sum_j \gamma_{ij}[y_i,s_j]}{\sum_j \gamma_{ij}[s_j,s_j]} \tag{2.31}$$

Clearly, the role of the γ_{ij} is to weight the output $[y_i,s_j]$ of each filter matched to the different possible signals at each receiver according to the relative importance of that output, i.e. it attaches more weight to those outputs for which the ratio a_i/D_i is higher, and which respond strongly to the input.

The form (2.31) of the estimate clearly gives a biased estimate, and motivated by the fact that we wish to obtain an unbiased estimate, we consider instead,

$$\hat{a}_i = \frac{\sum_j \gamma_{ij} [v_i, s_j]}{\frac{1}{M} \sum_k \sum_j \gamma_{ij} [s_k, s_j]} \qquad (2.31a)$$

where, here we take the γ_{ij} to be arbitrary weight functios whose sum for each i is unity. The choice of the γ_{ij} will usually have to be made a priori as part of the signal design problem associated with the kind of diversity problem which we consider. Furthermore, since, by our choice, there is no reason why they should depend on the path or channel considered, we shall suppress their dependence on i.

It is quite difficult to obtain an explicit solution to (2.30), and so Balakrishnan [3] proposed a variant on the estimate, motivated essentially by the form of (2.31). He instead considers,

$$\hat{a}_i = \frac{\sum_j [v_i, s_j]}{\frac{1}{M} \sum_k \sum_j [s_k, s_j]} \qquad (2.32)$$

where the denominator was chosen in order to make the estimate unbiased. This estimate is Gaussian with variance given by,

$$\mathrm{Var}(\hat{a}_i) = \frac{M^2 D_i}{\sum_k \sum_j [s_k, s_j]} \tag{2.33}$$

The form of the estimate given by (2.32) may, however, be justified by the following arguments

If it was known at the receiver which signal was sent, or if there was only one signal as in on-off signaling, then it is well known that the maximum likelihood estimate for the amplitude parameter a_i is given by,

$$\hat{a}_i = \frac{[y_i, s]}{[s\ ,s]} \tag{2.34}$$

and since, $y_i = a_i s + n_i$, we have,

$$\hat{a}_i = a_i + \frac{[n_i, s]}{[s\ ,s]} \tag{2.35}$$

so the estimate $\hat{a}_i$ of a_i is the true value a_i plus a term due to the noise, and since the noise is a zero-mean process, it is seen that the estimate (2.35) is unbiased.

The above was predicated upon the receiver knowing exactly which signal was sent. Since we assume the signals equally likely, it is natural to remove the conditioning by weighting the conditional estimates by the relative frequency of occurence of the signals, and thus arrive at,

$$\hat{a}_i = \frac{1}{M} \sum_{j=1}^{M} \frac{[y_i, s_j]}{[s_j, s_j]} \tag{2.36}$$

which is clearly biased, with

$$E(\hat{a}_i) = a_i \cdot \frac{1}{M^2} \sum_k \sum_j [s_k, s_j] \tag{2.37}$$

and to preserve unbiasedness, we take,

$$\hat{a}_i = \frac{\sum_j [y_i, s_j]}{\frac{1}{M} \sum_k \sum_j [s_k, s_j]} \tag{2.38}$$

exactly the form proposed by Balakrishnan [3].

2.7 SPECIFICATION TO THE BINARY SIGNALING CASE

As a prelude to using some of the results derived in this chapter, we shall specify their forms when the signaling is by two basic waveforms $s_1(t)$ or $s_2(t)$.

The decision rule then reduces to: decide on s_1 if,

$$\sum_{i=1}^{N} \frac{\hat{a}_i}{D_i} \left\{ [y_i, s_1] - [y_i, s_2] \right\} > 0 \tag{2.39}$$

and on s_2 otherwise.

For $\hat{a}_i$ in (2.39), we are to use any estimate given by the maximum likelihood criterion or any of its variants. In chapter III, we shall use either one of the forms,

$$\hat{a}_i = \frac{1}{E} \left\{ \gamma_1 [y_i, s_1] + \gamma_2 [y_i, s_2] \right\} \tag{2.40}$$

corresponding to (2.31), and in which the γ_j's are taken to be arbitrary weight factors chosen to minimize the probability of decision error. Clearly, there is no loss of generality by taking,

$$\gamma_1 + \gamma_2 = 1$$

or we consider an estimate of the form,

$$\hat{a}_1 = \frac{\gamma_1[v_1, s_1] + \gamma_2[v_1, s_2]}{\frac{1}{2} E (1 + \rho)} \qquad (2.41)$$

corresponding to the estimate of (2.31a).

In chapter IV we shall use another estimate for a_1 in the case of antipodal signals, corresponding to $\rho = -1$ in (2.41), which estimate is obviously indeterminate for $\rho = -1$. We shall defer discussion of this estimate until chapter IV.

CHAPTER III

MULTIPLE PROCESSING OPERATION

In this chapter, we shall use the results of the previous chapter in order to build the composite decision rule.

We shall derive the statistics of the decision variables and formulate expressions for the probability of decision error.

We shall also describe the two basic modes of operation, between which we subsequently make a comparison to determine the better mode.

3.1 COMPOSITE DECISION RULE

We determined in the previous chapter that the optimum decision rule for binary signaling is: Decide on s_1 if

$$\sum_{i=1}^{N} \frac{a_i}{D_i} \left\{ [y_i, s_1] - [y_i, s_2] \right\} > 0 \qquad (3.1)$$

and on s_2 otherwise.

We also noted that in the absence of complete knowledge of the $[a_i]$ or their relative ratios, we should estimate a_i and use in the decision rule those estimates as if they had perfect accuracy, in which case we have,

$$\sum_{i=1}^{N} \frac{\hat{a}_i}{D_i} \left\{ [y_i, s_1] - [y_i, s_2] \right\} > 0 \qquad (3.1a)$$

where for $\hat{a}_i$ we use any of those estimates discussed in chapter II.

Upon insertion of the estimate $\hat{a}_i$ in the decision rule, we obtain what shall be referred to in the sequel as the composite decision rule. For example, upon inserting (2.40) into (3.1a) and simplifying, we get: decide on s_1 if,

$$Z=\sum_{i=1}^{N} \frac{1}{ED_i} \left\{ \gamma_1 [y_i, s_1]^2 - \gamma_2 [y_i, s_2]^2 + (\gamma_2 - \gamma_1)[y_i, s_1][y_i, s_2] \right\} > 0 \qquad (3.2)$$

and on s_2 otherwise.

We refer to this mode of operation, whereby a single decision is made at the end of each observation interval based on the single decision variable Z in (3.2), as MULTIPLE PROCESSING, abbreviated (M.P.)

On the otherhand, there exists always the simpler mode of operation, viz. the single channel operation, whereby each receiver makes a decision independent of the others, and using majority logic to arrive at the ultimate decision at the end of each observation interval. We refer to this second mode as MAJORITY LOGIC operation, abbreviated (M.L.) the advantage, if it can be called as such, with this mode of operation is that the need for the estimate of the amplitude parameter is obviated.

3.2 PERFORMANCE ANALYSIS

Considering equally likely signals, it is easy to see, by symmetry, that the average probability of error is the probability of error conditioned on s_1 being sent. Thus we may write,

$$P_e = P\left\{ Z < 0 \mid s_1 \right\} \qquad (3.3)$$

where Z is the composite decision variable.

Prior to the determination of the average probability of error, we seek to evaluate the statistics of the decision variables, thus let,

$$x_{ij} = \frac{[y_i, s_j]}{(E\, D_i)^{1/2}} \qquad (3.4)$$

then the set $[x_{ij}]$, $j=1,2$; $i=1,\ldots,N$ are Gaussian random variables, being linear transformations of the set $[y_i]$ which are Gaussian. The conditional expectations and covariances of the $[x_{ij}]$ - conditioned on s_1 sent - are,

$$\bar{x}_{ij} = \begin{cases} \left(\frac{E}{D_i}\right)^{\frac{1}{2}} a_i & , \ j = 1 \\ \left(\frac{E}{D_i}\right)^{\frac{1}{2}} \rho\, a_i & , \ j = 2 \end{cases} \qquad (3.5)$$

$$\text{Cov}\,[x_{ij}\; x_{mk}] = \begin{cases} 0 & i \neq m \\ \left.\begin{array}{ll} 1 & , j = k \\ \rho & , j \neq k \end{array}\right\} & i = m \end{cases} \qquad (3.6)$$

Thus the Gaussian random variables x_{ij} are statistically independent for different channels; and for the same channel, they have unity variances and covariance equal to the correlation coefficient of the signals.

We now focus our attention on the determination of the average probability of error. We note that the composite decision variable Z may be written as a quadratic form by defining the vector X as,

$$X = \left\{ x_{11}, x_{21}, \ldots\ldots, x_{N1}; x_{12}, x_{22}, \ldots\ldots, x_{N2} \right\} \qquad (3.7)$$

thus,

$$Z = [X, QX]$$

where,

$$Q = \begin{bmatrix} \gamma_1 \, I & \dfrac{\gamma_2 - \gamma_1}{2} \, I \\ \dfrac{\gamma_2 - \gamma_1}{2} \, I & -\gamma_2 \, I \end{bmatrix} \qquad (3.8)$$

with I, the N x N identity matrix.

Hence,

$$P_e = P\left\{ [X, QX] < 0 \mid s_1 \right\} \qquad (3.9)$$

where X is a Gaussian random 2N-vector whose probability density function is,

$$p(X) = (2\pi)^{-N}(\det \Lambda)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(X-M)^{*}\Lambda^{-1}(X-M)\right] \quad (3.10)$$

with mean vector,

$$M = \left\{A_d \,,\, \rho A_d\right\} \quad (3.11)$$

in which A_d is the column N-vector

$$A_d = \left\{\left(\frac{E}{D_1}\right)^{\frac{1}{2}} a_1 \,, \left(\frac{E}{D_2}\right)^{\frac{1}{2}} a_2 \,, \,\ldots\ldots \,, \left(\frac{E}{D_N}\right)^{\frac{1}{2}} a_N\right\} \quad (3.12)$$

and with covariance matrix,

$$\Lambda = \begin{bmatrix} I & \rho I \\ \rho I & I \end{bmatrix} \quad (3.13)$$

One may express (3.9) by invoking the famous inversion formula [11] for the characteristic function $C_V(.)$ of a random variable $V(.)$

$$F_V(v) = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \frac{\operatorname{Im}[e^{-iuv} C_V(u)]}{u}\, du \quad (3.14)$$

where the integral is understood as a Cauchy principal value in the sense,

$$\lim_{\mu \to \infty} \int_{1/\mu}^{\mu}$$

For this purpose, we need the characteristic function of the quadratic form $[X, QX]$ with X distributed as in (3.10) above. This is a well known result and can easily be shown to be,

$$C_{[X,QX]}(u) = \frac{1}{[D(u)]^{N/2}} \exp\left\{ \frac{\|M\|^2}{2} \left[\frac{1+2iu\,\gamma_2(1-\rho^2)}{D(u)} - 1 \right] \right\} \qquad (3.15)$$

where,

$$D(u) = 1 + 2iu(\gamma_2 - \gamma_1)(1-\rho) + u^2(1-\rho^2) \qquad (3.16)$$

Using (3.14), we find that,

$$P_e = F_{[X,QX]}(0) = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \frac{\mathrm{Im}[C_V(u)]}{u}\,du \qquad (3.17)$$

which, upon inserting $\mathrm{Im}[C_V(u)]$ reduces eventually to,

$$P_e = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \exp\left\{ -\frac{\|M\|^2}{2} \left[1 - \frac{1+u^2(1+\mu\nu)}{(1+u^2)^2+u^2\,\nu^2} \right] \right\}$$

$$\cdot \sin\left\{ \frac{\|M\|^2}{2} \frac{u[\mu(1+u^2)-\nu]}{(1+u^2)^2+u^2\,\nu^2} - \frac{N}{2}\,\varnothing(u) \right\}$$

$$\cdot \frac{1}{u[(1+u^2)^2+u^2\,\nu^2]^{N/4}}\,du \qquad (3.18)$$

where,

$$\varnothing(u) = \tan^{-1} \frac{u\nu}{1+u^2}$$

$$\mu = 2\,\gamma_2\sqrt{1-\rho^2}$$

$$\nu = 2(\gamma_2 - \gamma_1)\sqrt{\frac{1-\rho}{1+\rho}}$$

We do not propose to carry this line of approach further, except to note the complexity of the expression, a fact which makes further analysis very difficult. We only refer the interested reader to the work by Balakrishnan and Abrams [7] where they used a similar type of integral for the special case N=2. Stein [8] confines discussion of a similar integral to only special cases of interest.

3.3 DECISION RULE RE-CONSIDERED

In section (3.1) we discussed the multiple processing operation mode with a particular estimate (2.40) used in the decision rule (3.1a) which yielded a composite decision rule (3.2).

In this section we shall make use of a slightly different approach in the decision rule, to obtain a simple bound on the average probability of error.

Noting that the terms in the decision rule (3.1a) and the estimate (2.41) involve sums and differences of random variables that can be made independent by choosing $\gamma_1 = \gamma_2$ we consider here an estimate of the form,

$$\hat{a}_1 = \frac{1}{E(1+\rho)} \left\{ [v_1, s_1] + [v_1, s_2] \right\} \tag{3.19}$$

At this point, it is not clear what to do in the case of antipodal signals. We thus exclude them for the time being and defer discussion of this case until chapter IV.

Inserting the estimate (3.19) into the decision rule (3.1a), we have the new composite decision rule: decide on s_1 if,

$$Z = \sum g\left(\frac{\hat{a}_i}{\sqrt{D_i}}\right) \left\{ \frac{[y_i, s_1] + [y_i, s_2]}{\sqrt{D_i}} \right\} > 0 \qquad (3.20)$$

and on s_2 otherwise.

Here, instead of $\hat{a}_i$ we have put a function $g(\hat{a}_i)$ of it. This function $g(.)$ is only a weight to be attached to each term in the summation, therefore we are not restricted to using the $\hat{a}_i$'s themselves, but may use any function of them provided it has the following basic properties:

(i) Since the decision is based on the sign of a weighted sum of terms, it is imperative that the weighting does not introduce any sign destruction. Hence we require $g(.)$ to be non-negative.

(ii) Since larger values of $\frac{a_i}{\sqrt{D_i}}$ reflect that a channel with that index i has a higher SNR, hence more reliable, we should require $g(.)$ to be monotonic non-decreasing.

We observe that since the estimate considered in (3.19) is unbiased with mean at the true value a_i and variance equal to $2D_i/E(1+\rho)$, we should expect that as D_i increases the probability density function of $\hat{a}_i$ flattens. Hence $\hat{a}_i$ will be negative with non-zero probability which increases with D_i. We may thus consider one possible choice for $g(.)$

$$g(\hat{a}_i) = \beta_i \; \hat{a}_i \; ; \quad \beta_i = \begin{cases} 1 & \text{if } a_i > 0 \\ 0 & \text{if } a_i < 0 \end{cases} \qquad (3.21)$$

which says that we use in the decision processing only that data arriving via the channels [i] for which $[v_i,s_i]+[v_i,s_2]$ is greater than zero, and to ignore the others as being too noisy to be relied upon.

The above does not preclude using the data from all channels, whether a_i be positive or negative. We may do this by considering,

$$g(\hat{a}_i) = \exp\left\{[v_i,s_i] + [v_i,s_2]\right\} \qquad (3.22)$$

which gives a weight larger than one for the data from the channels [i] for which $[v_i,s_i]+[v_i,s_2]$ is greater than zero, and a weight less than one otherwise.

3.4 A BOUND ON P_e

For normalization purposes, we shall slightly modify the decision quantity Z of (3.20) to read,

$$Z=\sum_{i=1}^{N} g\left[\frac{[v_i,s_i]+[v_i,s_2]}{(2ED_i(1+\rho))^{1/2}}\right]\left\{\frac{[v_i,s_i]-[v_i,s_2]}{(2ED_i(1-\rho))^{1/2}}\right\} \qquad (3.23)$$

Considering we have equally likely signals, we again have,

$$P_e = P\left\{Z < 0\right\} \qquad (3.24)$$

Now, defining,

$$V_i = \frac{[v_i,s_i]+[v_i,s_2]}{(2ED_i(1+\rho))^{1/2}} \qquad (3.25)$$

and,

$$W_1 = \frac{[y_1, s_1]-[y_1, s_2]}{(2ED_1(1-\rho))^{1/2}} \tag{3.26}$$

It can be seen that V_1 and W_1 are Gaussian random variables with conditional means and variances given by,

$$\bar{V}_1 = \sqrt{\frac{1+\rho}{2}}\; m_1 \tag{3.27}$$

$$\bar{W}_1 = \sqrt{\frac{1-\rho}{2}}\; m_1 \tag{3.28}$$

$$Var(V_1) = Var(W_1) = 1 \tag{3.29}$$

where we define,

$$m_1 = \sqrt{\frac{E}{D_1}}\; a_1 \tag{3.30}$$

To show that V_1 and W_1 are independent for all 1, consider,

$$V_1+W_1 = \frac{1}{\sqrt{2}}\left\{ \frac{x_{11} + x_{12}}{\sqrt{1+\rho}} + \frac{x_{11} - x_{12}}{\sqrt{1-\rho}} \right\}$$

$$= \frac{1}{\sqrt{2(1-\rho^2)}}\left\{x_{11}(\sqrt{1-\rho} + \sqrt{1+\rho}) + x_{12}(\sqrt{1-\rho} + \sqrt{1+\rho})\right\} \tag{3.31}$$

where x_{1j} is defined in (3.4), and therefore,

$$Var(V_1+W_1) = \frac{1}{2(1-\rho^2)}\Big\{(\sqrt{1-\rho} + \sqrt{1+\rho})^2 Var(x_{11})$$

$$+ (\sqrt{1-\rho} - \sqrt{1+\rho})^2 Var(x_{12})$$

$$+ 2(\sqrt{1-\rho} + \sqrt{1+\rho})(\sqrt{1-\rho} - \sqrt{1+\rho}) Cov(x_{11}, x_{12})\Big\}$$

$$= 2 \tag{3.32}$$

where we have used (3.6). On the other hand, it is known,

$$\begin{aligned} Var(V_i+W_i) &= Var(V_i) + Var(W_i) + 2Cov(V_i,W_i) \\ &= 2 + 2\ Cov(V_i,W_i) \end{aligned} \tag{3.33}$$

From (3.32) and (3.33) we conclude the random variables V_i and W_i are uncorrelated, and since they are Gaussian, they are statistically independent for all i. Moreover, there is complete statistical independence for all these variables over the different indices [i] since the noise processes are independent, as can be seen from (3.6).

Thus, in terms of V_i and W_i, we may write,

$$P_e = P\left\{ \sum g(V_i)\ W_i < 0 \mid s_1 \right\} \tag{3.34}$$

Now, given that each $V_i = \beta_i$, the decision quantity Z is a weighted sum of independent Gaussians, hence Gaussian itself, and the conditional probability that this sum be less than zero given the set $[\beta_i]$ will be,

$$P\left\{ Z < 0 \mid s_1, \underline{V} \right\} = erfc\left\{ \sqrt{\frac{1-\rho}{2}}\ \frac{\sum m_i g(\beta_i)}{\left[\sum g^2(\beta_i)\right]^{1/2}} \right\} \tag{3.35}$$

where we define,

$$erf(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-y^2/2)\ dy \tag{3.36}$$

$$erfc(x) = 1 - erf(x) \tag{3.37}$$

Hence,

$$P_e = \int_{-\infty}^{\infty} \cdots\cdots \int_{-\infty}^{\infty} \operatorname{erfc}\left\{ \sqrt{\frac{1-\rho}{2}} \frac{\sum m_i g(\beta_i)}{[\sum g^2(\beta_i)]^{1/2}} \right\} p_{\underline{V}}(\underline{\beta})\, d\underline{\beta} \tag{3.38}$$

where,

$$p_{\underline{V}}(\underline{\beta}) = \prod_{i \in I}^{\sim} p_{V_i}(\beta_i) \tag{3.39}$$

by independence of the V_i's over the index set [I], and where,

$$p_{V_i}(\beta_i) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} (\beta_i - \bar{\beta}_i)^2 \right\} \tag{3.40}$$

with $\bar{\beta}_i$, the expected value of V_i given by (3.27).

It is possible now to bound P_e from above and from below; for since,

$$\sum_i m_i\, g(\beta_i) < [\sum_i m_i^2]\, [\sum_i g^2(\beta_i)] \tag{3.41}$$

by the Schwarz inequality, we have,

$$P_e \geq \operatorname{erfc}\left\{ \sqrt{\frac{1-\rho}{2}} [\sum_i m_i^2]^{\frac{1}{2}} \right\} \tag{3.42}$$

On the otherhand, for any set of non-negative real numbers x_i, it is always true that,

$$\sum_i x_i^2 \leq [\sum_i x_i]^2 \tag{3.43}$$

hence,

$$\frac{\sum_i m_i\, g(\beta_i)}{[\sum_j g^2(\beta_j)]^{1/2}} \geq \frac{\sum_i m_i\, g(\beta_i)}{\sum_j g(\beta_j)} \tag{3.44}$$

and since erfc(x) is monotonically decreasing in x, then,

$$\operatorname{erfc}\left\{\sqrt{\frac{1-\rho}{2}}\ \frac{\sum_i m_i g(\beta_i)}{[\sum_j g^2(\beta_j)]^{1/2}}\right\} < \operatorname{erfc}\left\{\sqrt{\frac{1-\rho}{2}}\ \frac{\sum_i m_i g(\beta_i)}{\sum_j g(\beta_j)}\right\} \tag{3.45}$$

Moreover, since erfc(x) is convex for non-negative arguments, therefore the right-hand side of (3.45) may further be bounded by,

$$\sum_i \frac{g(\beta_i)}{\sum_j g(\beta_j)} \operatorname{erfc}\left\{\sqrt{\frac{1-\rho}{2}}\ m_i\right\}$$

Therefore,

$$P_e \leq \sum_i \operatorname{erfc}(R^- m_i)\ E\left\{\frac{g(\beta_i)}{\sum_j g(\beta_j)}\right\} \tag{3.46}$$

in which we use,

$$R^{\pm} = \sqrt{\frac{1 \pm \rho}{2}} \tag{3.47}$$

and where the expectation in (3.46) is with respect to the independent Gaussian random variables β_i, each with mean given by (3.27) and unity variance. Now letting,

$$\epsilon_i = E\left\{\frac{g(\beta_i)}{\sum_j g(\beta_j)}\right\} \tag{3.48}$$

then clearly each ϵ_i is bounded by $0 < \epsilon_i < 1$.

It is quite difficult to obtain a general expression for the expectation in (3.48). However, for the class of candidate functions we use for g(.), for example $g(\beta_i) = \exp(\beta_i)$, it may be shown that as a first approximation, the expectation may be taken as,

$$\epsilon_i \approx \frac{g(\bar{\beta}_i)}{\sum g(\bar{\beta}_i)} \tag{3.49}$$

with the second term of the order of,

$$o\left(\frac{1}{g(\max_i \bar{\beta}_i)}\right)$$

Thus from (3.46) we may write,

$$P_e \leq \sum_i \epsilon_i \;\mathrm{erfc}(R^- m_i) \tag{3.50}$$

Combining (3.42) and (3.50) we have,

$$\mathrm{erfc}\left\{R^- \left[\sum_i m_i^2\right]^{1/2}\right\} \leq P_e \leq \sum_i \epsilon_i \;\mathrm{erfc}\left\{(R^- m_i)\right\}$$

$$< \max_i \;\mathrm{erfc}\left\{(R^- m_i)\right\}$$

$$= \mathrm{erfc}\left\{(R^- \min_i m_i)\right\} \tag{3.51}$$

where,

$$\epsilon_i \approx \frac{g(R^+ m_i)}{\sum_j g(R^+ m_j)} \tag{3.52}$$

The bound in expression (3.51) is the bound we seek. It applies to the "multiple processing" mode of operation. To effect a comparison with the "majority logic" mode would be involved at this point. Rather than obscure the result with too many terms, we refer only to the case of two-channel operation. In chapter VI we show that for

N = 2, the probability of error under M.L. may be written,

$$P_e = \frac{1}{2}\,\mathrm{erfc}(R^- m_1) + \frac{1}{2}\,\mathrm{erfc}(R^- m_2) \qquad (3.53)$$

For the same situation, the bound of (3.51) reduces to,

$$P_e \leq \epsilon_1\,\mathrm{erfc}(R^- m_1) + \epsilon_2\,\mathrm{erfc}(R^- m_2) \qquad (3.54)$$

It is evident that the bound of (3.54) is smaller than the right side of (3.53) by virtue of the fact that the larger weight is attached to the smaller of the two quantities $\mathrm{erfc}(R^- m_1)$ and $\mathrm{erfc}(R^- m_2)$. The bound of (3.54) and the error probability of (3.53) coincide if $m_1 = m_2$.

CHAPTER IV

MULTIPLE PROCESSING OF ANTIPODAL SIGNALS

In the previous chapter, we excluded consideration of the case when the binary signaling waveforms were antipodal. The reason was the difficulty of reconciling the form of the estimate taken for $\hat{a}_1$ (3.19) with a value of -1 for ρ, the correlation coefficient.

We devote this chapter to the mode of multiple processing of binary antipodal signals, by suggesting a slightly different estimate for $\hat{a}_1$, and using it to form a new composite decision rule.

In the following analysis, we shall appeal to geometric ideas in the determination of the probability of error. The derivation will be based on geometrically interpreting the aspect and variations of the decision hypersurface with respect to the axes of coordinates.

Although the analysis that follows is general, we shall concentrate our attention on odd-dimensional Euclidean space, merely for reasons of convenience and in order to make a meaningful comparison with the other mode of operation of "majority logic", in the next chapter.

4.1 SUGGESTING AN ESTIMATE

In chapter II, we found that the maximum likelihood estimate a_1 for a_1 was given by the root of the equation,

$$\hat{a}_1 = \frac{1}{E} \frac{\gamma_1[v_1,s_1] + \gamma_2[v_1,s_2]}{\gamma_1 + \gamma_2} \qquad (4.1)$$

where,

$$\gamma_j = \exp\left(\frac{\hat{a}_1}{D_1}[v_1,s_j]\right)$$

In the case of antipodal signals we have,

$$[v_1,s_2] = -[v_1,s_1] \qquad (4.2)$$

which reduces (4.1) to,

$$\hat{a}_1 = \frac{1}{E}[v_1,s_1] \tanh\left\{\frac{\hat{a}_1}{D_1}[v_1,s_1]\right\} \qquad (4.3)$$

and the estimate a_1 is contained implicitly in the solution to the transcendental equation (4.3) which is difficult to obtain explicitly.

It is clear, however, that the role of the quantity $\tanh\left\{\frac{\hat{a}_1}{D_1}[v_1,s_1]\right\}$ in (4.3) is to weight the normalized output $[v_1,s_1]/E$, of the single matched filter by a (+) or (-) sign, according as which of the two signals was present, in addition to a magnitude weighting according to the strength of the received signal. Again, motivated by the form of this result (4.3) we consider instead the following estimate,

$$\hat{a}_1 = \frac{1}{E} [y_1, s_1] \, \mathrm{sgn}[y_1, s_1] \tag{4.4}$$

We may explain this approximation by the following arguments: Equation (4.3) may be written as,

$$\frac{1}{C} z = \tanh z \tag{4.5}$$

where we have put,

$$z = \frac{\hat{a}_1}{D_1} [y_1, s_1] \quad \text{and} \quad C = \frac{[y_1, s_1]^2}{ED_1} \tag{4.6}$$

We seek a non-trivial root of (4.5). Hence if we plot the two sides of (4.5) as functions of z, as in fig (4.1), then the root is clearly where the two curves intersect.

From the figure, it is obvious that, if we constrain the root z_o to be positive, then,

$$z_o < C$$

On the otherhand, z_o must satisfy (4.5), and using the simple inequality,

$$\log x \leqslant x - 1$$

it is readily shown that,

$$z_o \geqslant C - 1$$

We see from the above that

$$C - 1 < z_o < C \tag{4.7}$$

that is to say, the root of (4.5) is within one unit interval of C. We choose to take it as being equal to C,

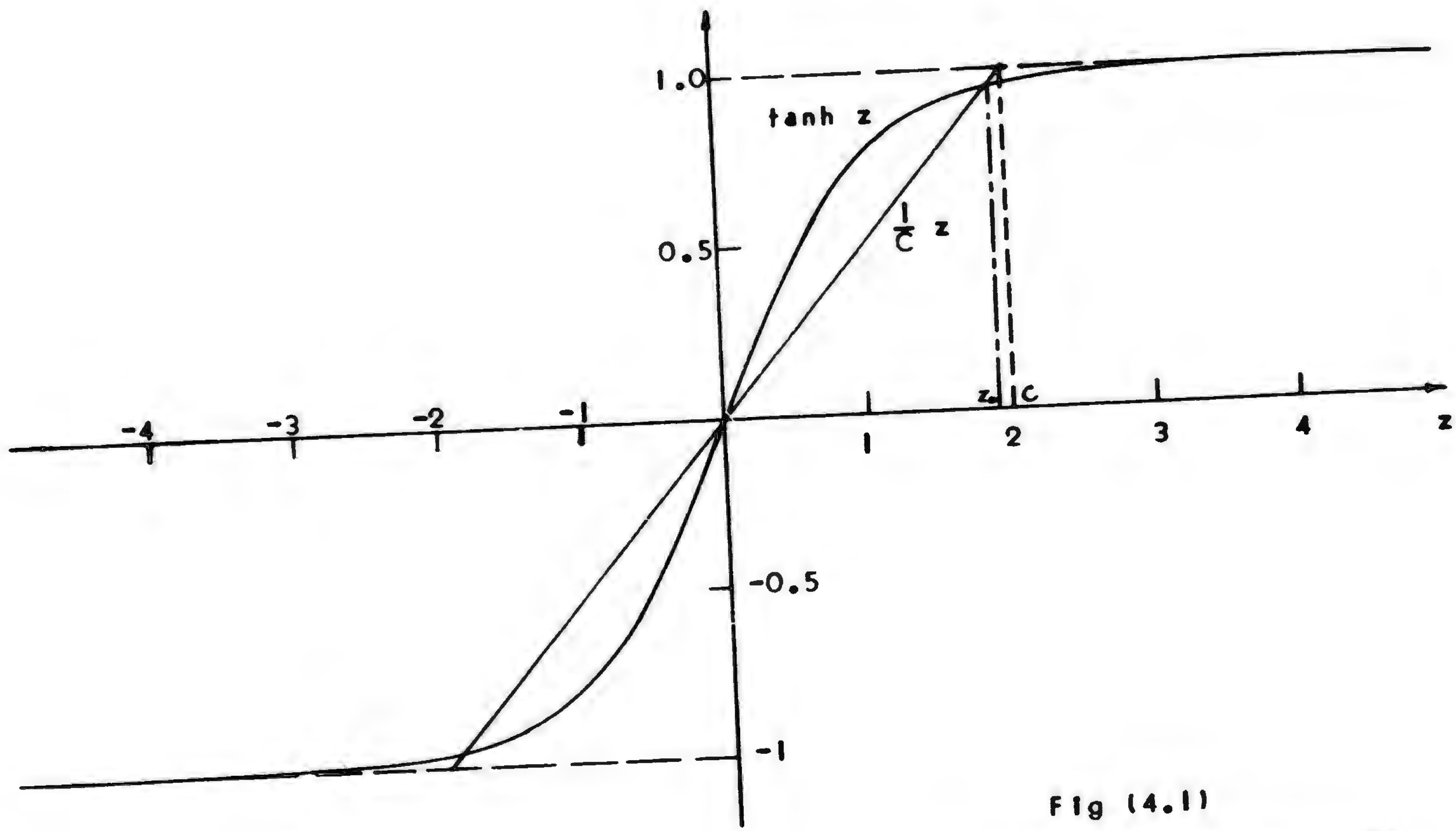


Fig (4.1)
Finding root of eq. (4.5)

or in terms of the filter outputs,

$$\hat{a}_i = \frac{[y_i, s_i]}{E} \tag{4.8}$$

Of course, there will be two roots as can be seen in fig (4.1) according as $[y_i, s_i]$ is positive or negative. To account for both cases, therefore, we introduce $\mathrm{sgn}[y_i, s_i]$ in (4.8) and thus arrive at (4.4).

Furthermore, from the physical point of view, this is a justifiable approximation; for if it was known at the receiver that s_i was sent, then it is well known that an unbiased estimate for the amplitude parameter is the matched filter output normalized by the signal energy as can be seen in (2.34). The filter output, however, will be positive or negative depending on which signal was present. It is reasonable then to take the magnitude as the estimate.

The estimate $\hat{a}_i$ of (4.4) is unbiased with mean at the true value a_i and variance D_i/E, however, it is no longer Gaussian.

4.2 COMPOSITE DECISION RULE

Using (4.2), the composite decision rule of (3.1a) may be written,

$$Z_2 = \sum_{i=1}^{N} \frac{\hat{a}_i}{\sqrt{D_i}} \frac{[y_i, s_i]}{\sqrt{D_i}} \gtrless 0 \qquad (4.9)$$

which upon inserting the estimate a_i of (4.4) reduces to: decide on s_1 if,

$$Z_2 = \sum_{i=1}^{N} \left(\frac{[y_i, s_i]}{\sqrt{ED_i}}\right)^2 \operatorname{sgn}[y_i, s_i] > 0 \qquad (4.10)$$

and on $s_2 = - s_1$ otherwise.

Since $\operatorname{sgn}[y_i, s_i]$ is not affected by changes in E or D_i, (4.10) may simply be written,

$$Z_2 = \sum_{i=1}^{N} x_i^2 \operatorname{sgn} x_i > 0 \qquad (4.10a)$$

where,

$$x_i = \frac{[y_i, s_i]}{\sqrt{ED_i}}$$

In the next chapter, we consider the "majority logic" mode of operation. There, it is shown that the composite decision rule for antipodal signals reduces to,

$$Z_o = \sum_{i=1}^{N} \operatorname{sgn} x_i > 0 \qquad (4.11)$$

These expressions suggest* that one might consider, in general, an alternate composite decision rule of the form,

$$Z_\eta = \sum_{i=1}^{N} |x_i|^\eta \operatorname{sgn} x_i > 0 \qquad (4.12)$$

---

* This was suggested to the author by Dr. Robert Price.

wherein the parameter η may be assigned different values to determine if, possibly, any particular value of leads to a uniformly lower probability of error. We investigate this possibility as we proceed in the determination of the probability of error in the following sections.

Again, considering equally likely signals, the average probability of error is the conditional probability of error given that s_1 was sent. Therefore, conditioned on s_1, we have,

$$P_e = P\left\{ Z_\eta < 0 \;\middle|\; s_1 \right\} \tag{4.13}$$

We recall [c.f. (3.5) and (3.6)] that the random variables $[x_i]$ in (4.12) are independent Gaussian, each with mean,

$$m_i = \sqrt{\frac{E}{D_i}}\; a_i$$

and unity variance.

Thus, the random vector,

$$X = \left\{ x_1 \,,\, x_2 \,,\, \ldots\ldots\ldots \,,\, x_N \right\}$$

is normally distributed in N-dimensional Euclidean space, with mean vector,

$$M = \left\{ \sqrt{\frac{E}{D_1}}\; a_1, \sqrt{\frac{E}{D_2}}\; a_2, \;.\;.\;.\;, \sqrt{\frac{E}{D_N}}\; a_N \right\} \tag{4.14}$$

and covariance matrix I, the N x N identity matrix, and we have,

$$p(X) = (2\pi)^{-N/2} \exp\left\{-\frac{1}{2}(X-M)^{*}(X-M)\right\} \tag{4.15}$$

4.3 PROBABILITY OF ERROR (N = 2)

For the case of two channel operation, the probability of error is simply,

$$P_e = P\left\{|x_1|^{\eta} \operatorname{sgn} x_1 + |x_2|^{\eta} \operatorname{sgn} x_2 < 0 \mid s_1\right\} \tag{4.16}$$

We observe that the above event coincides identically with the event $[\, x_1 + x_2 < 0 \,]$, it being understood that, for the case $\eta = 0$, equality to zero means that either s_1 or s_2 could be present. Thus

$$P_e = P[\, x_1 + x_2 < 0 \mid s_1] \tag{4.17}$$

which is easily shown to be,

$$P_e = \operatorname{erfc}\left[\frac{1}{\sqrt{2}}(m_1 + m_2)\right] \tag{4.18}$$

The significance of this result is discussed in the concluding chapter.

4.4 MULTIPLE CHANNEL OPERATION ($N \geq 3$)

We begin with two basic definitions which will be required in the subsequent discussion.

Definition: The vector M of the means shall be referred to as the POLAR AXIS.

Definition: The vector B in the positive orthant (octant) which is equidistant from the coordinate axes shall be referred to as the MAJOR AXIS.

GEOMETRICAL DESCRIPTION OF THE DECISION HYPERSURFACE

It will be observed that the decision hypersurface which is described by the equation,

$$\sum_{i=1}^{N} |x_i|^{\eta} \operatorname{sgn} x_i = 0 \qquad (4.19)$$

divides the probability volume (or space) into two regions, the boundary between which is a surface passing through the origin of space, and has an undulatory appearance.

In general we have 2^N orthants, one of which is bounded by the surfaces joining the all-positive axes, namely the positive orthant; and one is bounded by the surfaces joining the all-negative axes, namely the negative orthant. The decision hypersurface passes through the remaining $(2^N - 2)$ orthants whose bounding axes are dissimilar in sign.

If N is odd, there will be $\frac{1}{2}(2^N - 2)$ orthants which are bounded by majority positive axes, and an equal number of orthants bounded by majority negative axes.

If N is even; in addition to the positive and negative orthants, there will be $\binom{N}{N/2}$ orthants whose bounding axes are equally positive and negative. This leaves $2^N - 2 - \binom{N}{N/2}$ orthants, half of which are bounded by majority positive axes, and half of which by majority negative axes.

For example, in 3-dimensional Euclidean space, this undulating (or wavy) surface is made of conical segments in each of the six octants whose bounding axes are dissimilar in sign; and with respect to any point in the space, this surface will change its curvature from positive to negative and so on, as it passes from one octant to the adjoining one. Fig (4.2)

Similarly, in 5-dimensional Euclidean space, the decision hypersurface is made of segments of hypercones in each of the thirty orthants whose bounding axes are dissimilar in sign.

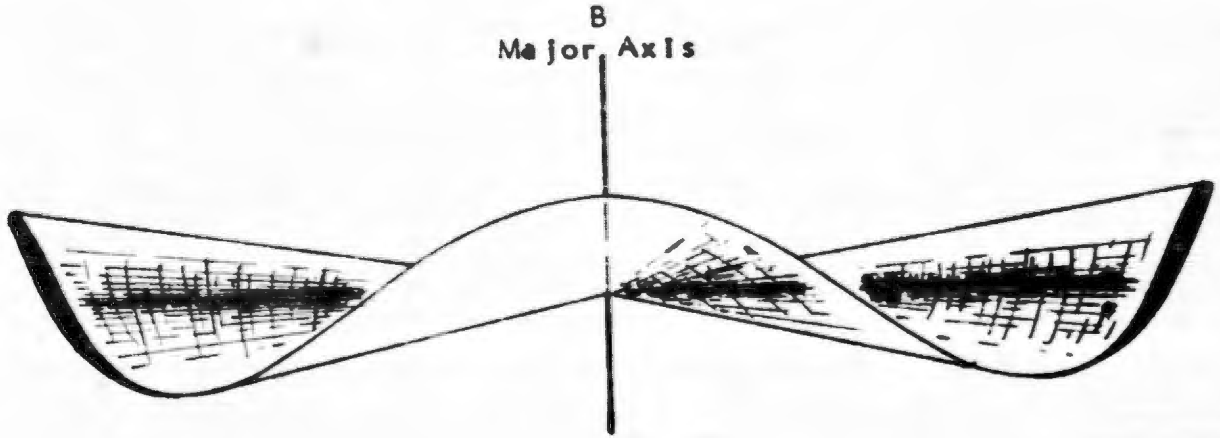


Fig (4.2) Pictorial view of the decision surface in $E^{(3)}$

Furthermore, the decision hypersurface forms an angle, Φ, with the major axis, which oscillates about the value $\frac{\pi}{2}$, and there is complete symmetry with respect to this value, i.e. the decision hypersurface is as much above as it is below the (N - 1)-flat described by $\Phi = \frac{\pi}{2}$.

We shall make use of this feature in evaluating the average probability of error.

DESCRIBING THE ANGLE BETWEEN THE DECISION HYPERSURFACE AND THE MAJOR AXIS

To describe the variations of the decision hypersurface for any η, consider any orthant in $E^{(N)}$ bounded by p positive axes and (N - p) negative axes. In particular, consider the orthant,

$$[\ x_i > 0,\ i=1,2,\ldots\ldots,p\ ;\ x_j < 0,\ j=p+1,\ldots,N\ ]$$

Let this orthant be called "G". The equation of the decision hypersurface in the orthant G, may be written,

$$\sum_{i=1}^{p} |x_i|^{\eta} - \sum_{i=p+1}^{N} |x_i|^{\eta} = 0 \qquad (4.20)$$

or equivalently,

$$\sum_{i=1}^{p} x_i^{\eta} - (-1)^{\eta} \sum_{i=p+1}^{N} x_i^{\eta} = 0 \qquad (4.20a)$$

Consider any point P on the hypersurface (4.20a). Let its coordinates be,

$$
\begin{aligned}
x_1 &= x_1 & \qquad x_{p+1} &= -\delta\, x_1 \\
x_2 &= \beta_1\, x_1 & \qquad x_{p+2} &= -\delta\, \gamma_1\, x_1 \\
x_3 &= \beta_2\, x_1 & \qquad x_{p+3} &= -\delta\, \gamma_2\, x_1 \\
\\
x_p &= \beta_{p-1} x_1 & \qquad x_N &= -\delta\, \gamma_{N-(p+1)} x_1
\end{aligned}
$$

where δ, $[\beta_i]$, $[\gamma_i]$ are real non-negative numbers. Then from the equation of the decision hypersurface (4.20a) it is easily verified that,

$$
\delta = \pm \left[\frac{1 + \sum_{i}^{p-1} \beta_i^{\eta}}{1 + \sum_{i}^{N-p-1} \gamma_i^{\eta}} \right]^{\frac{1}{\eta}} \tag{4.21}
$$

according as η is even or odd integer.

We note that any scalar multiple of the coordinates of the point P results in another point on the hypersurface.

With B, the N-vector whose coordinates are,

$$
B = [\,1,\ 1,\ 1,\ \ldots\ldots\ldots\ldots\ldots,\ 1]
$$

it directly follows that the angle Φ subtended at the origin by the vector P and the major axis B, is given by,

$$
\Phi_G = \cos^{-1}\left[\frac{1}{\sqrt{N}} \frac{\left(1+\sum_{i}^{p-1}\beta_i\right)\left(1+\sum_{i}^{N-p-1}\gamma_i^{\eta}\right)^{\frac{1}{\eta}} - \left(1+\sum_{i}^{p-1}\beta_i^{\eta}\right)^{\frac{1}{\eta}}\left(1+\sum_{i}^{N-p-1}\gamma_i\right)}{\sqrt{\left(1+\sum_{i}^{p-1}\beta_i^2\right)\left(1+\sum_{i}^{N-p-1}\gamma_i^{\eta}\right)^{\frac{2}{\eta}} + \left(1+\sum_{i}^{p-1}\beta_i^{\eta}\right)^{\frac{2}{\eta}}\left(1+\sum_{i}^{N-p-1}\gamma_i^2\right)}} \right]
$$

(4.22)

For any η, it is evident that the angle Φ will be equal to, greater than, or smaller than $\frac{\pi}{2}$ according as $\frac{1+\sum\beta_i}{(1+\sum\beta_i^\eta)^{1/\eta}}$ is respectively equal to, smaller than, or greater than $\frac{1+\sum\gamma_i}{(1+\sum\gamma_i^\eta)^{1/\eta}}$.

For $\eta = 1$, Φ is always equal to $\frac{\pi}{2}$, and the decision hypersurface is simply a hyperplane which passes through the origin orthogonal to the major axis.

For any $\eta \neq 1$, the decision hypersurface is an alpha-ic hypersurface (i.e. of order α), and this surface undulates about the hyperplane corresponding to $\eta = 1$.

On the otherhand, by considering an orthant "H", which is image-through-the-origin to the orthant "G", i.e. the orthant H is that which is bounded by,

$$[\ x_i < 0,\ i=1,2,\ldots,p\ ;\ x_j > 0,\ j=p+1,\ldots,N]$$

It can be shown by similar reasoning that the angle Φ_H subtended at the origin by the major axis B and any vector P on the decision hypersurface, is the complement with respect to π of the angle Φ_G, i.e.

$$\Phi_H = \pi - \Phi_G \tag{4.23}$$

Thus the decision hypersurface for any η in the orthant "H" is a continuation of that in the orthant "G", and where Φ was greater than $\frac{\pi}{2}$ in the orthant "G", it is now smaller than $\frac{\pi}{2}$ in the orthant "H" by the same amount.

In summation, we note that the decision hypersurface for any η, is an envelope of rays (or vectors) that oscillate about $\frac{\pi}{2}$ from the major axis as they describe the range of the other orthogonal coordinates of the N-dimensional generalized spherical coordinates of the space.

We shall determine the average probability of error based on this geometrical interpretation of the decision hypersurface. Hence, if a transformation is made of the probability density function p(X) of (4.15) in terms of generalized spherical coordinates $[r, \theta, \Phi_1, \Phi_2, .., \Phi_{N-2}]$ the error probability will be seen to relate directly to $p(\Phi_1)$, the probability density function of the preferred angle Φ_1 from the polar axis.

4.5 TRANSFORMATION OF p(X) IN TERMS OF GENERALIZED SPHERICAL COORDINATES

In terms of the decision variables $[x_i]$ the probability of error is,

$$P_e = \int_{\mathcal{A}} p(X)\, dX \qquad (4.24)$$

where $\mathcal{A}$ is the region of error,

$$\mathscr{A} = \left\{ X : \sum |x_i|^{\eta - 1} x_i < 0 \right\}$$

Making the orthogonal transformation T (rotation of axes) such that,

$$X = T\,Y \qquad ; \qquad Y = T^* X \tag{4.25}$$

and using the well known relations,

$$T^* T = I \qquad ; \qquad T^* = T^{-1}$$

we have,

$$(X - M)^*(X - M) = Y^* Y + M^* M - 2\, Y^*(T^* M) \tag{4.26}$$

thus the operator T leaves the norm invariant, and we have,

$$p(Y) = (2\pi)^{-N/2} \exp\left\{-\frac{1}{2}(r^2 + \|M\|^2)\right\} \exp\left\{Y^*(T^* M)\right\} \tag{4.27}$$

where,

$$r^2 = X^* X = Y^* Y = \sum y_i^2$$

$$\|M\|^2 = M^* M = \sum m_i^2$$

Introducing a complete orthonormal set of vectors $[\, e_1', e_2', \ldots, e_N' \,]$ with e_1' in the direction of $T^* M$, then $T^* M$ is the polar axis of the distribution (i.e. the axis through the centroid of the distribution.)

Making use of the generalized spherical coordinate change of variables and the Jacobian of transformation associated with it, which are given in appendix 'A', the resulting probability density function is,

$$p(r,\theta,\Phi_1,\ldots,\Phi_{N-2}) = (2\pi)^{-N/2}\ r^{N-1}\prod_{k=1}^{N-2}\sin^{N-1-k}\Phi_k$$

$$\exp\left\{-\tfrac{1}{2}(r^2+\|M\|^2)\right\}\ \exp\left\{r\|M\|\cos\Phi_1\right\}$$

$$r > 0\ ;\ 0 < \theta < 2\pi \qquad (4.28)$$

$$0 < \Phi_k < \pi \quad \text{for } k=1,2,\ldots,N-2$$

since by construction,

$$Y^*(T^*M) = r\|M\|\cos\Phi_1$$

and it follows that,

$$p(r,\theta,\Phi_1,\ldots,\Phi_{N-2}) = p(r,\Phi_1)\ p(\theta)\ \prod p(\Phi_k) \qquad (4.29)$$

where,

$$p(\theta) = \frac{1}{2\pi} \qquad 0 < \theta < 2\pi \qquad (4.30)$$

$$p(\Phi_k) = \sin^{N-1-k}\Phi_k\ ,\ 0 < \Phi_k < \pi\ ,\ k=2,3,\ldots,N-2$$

It is well known [13], however, that,

$$p(r) = \|M\|\left[\frac{r}{\|M\|}\right]^{N/2} \exp\left\{-\tfrac{1}{2}(r^2+\|M\|^2)\right\}\ I_{\frac{N-2}{2}}(r\|M\|) \qquad (4.31)$$

$$r \geq 0$$

On the otherhand, the probability density function along the preferred direction Φ_1 is given by,

$$p(\Phi_1) = K_N'\ \exp\left\{-\tfrac{1}{2}\|M\|^2\right\}\int_0^\infty r^{N-1}\ \exp\left\{-\tfrac{1}{2}r^2+r\|M\|\cos\Phi_1\right\}dr \qquad (4.32)$$

where,

$$K_N' = (2\pi)^{-\frac{N-2}{2}} \prod_{k=2}^{N-2} \int_0^{\pi} \sin^{N-1-k}\Phi_k \, d\Phi_k$$

$$= (2\pi)^{-\frac{N-2}{2}} \frac{1}{\Gamma(\frac{1}{2})\,\Gamma(\frac{N-1}{2})} \tag{4.33}$$

Invoking an identity in [12], we have,

$$\int_0^{\infty} r^{N-1} \exp\left\{-\frac{1}{2} r^2 + r\|M\|\cos\Phi_1\right\} dr$$

$$= \Gamma(N) \exp\left\{\frac{1}{4} \|M\|^2 \cos^2\Phi_1\right\} D_{-N}(-\|M\|\cos\Phi_1) \tag{4.34}$$

Using (4.34) in (4.32) we have,

$$p_N(\Phi_1) = K_N \sin^{N-2}\Phi_1 \exp\left\{-\frac{1}{4}\|M\|^2(1+\sin^2\Phi_1)\right\} D_{-N}(-\|M\|\cos\Phi_1)$$

$$0 \leq \Phi_1 \leq \pi \tag{4.35}$$

with,

$$K_N = \frac{2^{-\frac{N-2}{2}}\,\Gamma(N)}{\Gamma(\frac{1}{2})\,\Gamma(\frac{N-1}{2})} \tag{4.36}$$

where $D_n(z)$ is the Parabolic Cylinder function, and $\Gamma(N)$ is the Gamma function.

The density function $p(\Phi_1)$ of (4.35) may be put in two slightly different forms which we shall have occasion to use subsequently. Thus letting,

$$t = -\cos\Phi_1$$

we have,

$$p_N(t) = K_N\,(1-t^2)^{\frac{N-3}{2}} \exp\left\{-\tfrac{1}{4}\|M\|^2(2-t^2)\right\} D_{-N}(\|M\|\,t)$$

$$|t| < 1 \;;\; N \geq 3 \qquad (4.37)$$

or, provided $\|M\| \neq 0$, letting,

$$z = \|M\|\,t$$

we have,

$$p_N(z) = K_N \frac{\exp(-\|M\|^2/2)}{\|M\|} \left[1 - \left(\frac{z}{\|M\|}\right)^2\right]^{\frac{N-3}{2}} \exp(\tfrac{1}{4} z^2) D_{-N}(z)$$

$$|z| < \|M\| \;;\; N \geq 3 \qquad (4.38)$$

The density function (4.35) or its equivalent (4.38) is plotted for selected values of N in figures (4.3) through (4.6).

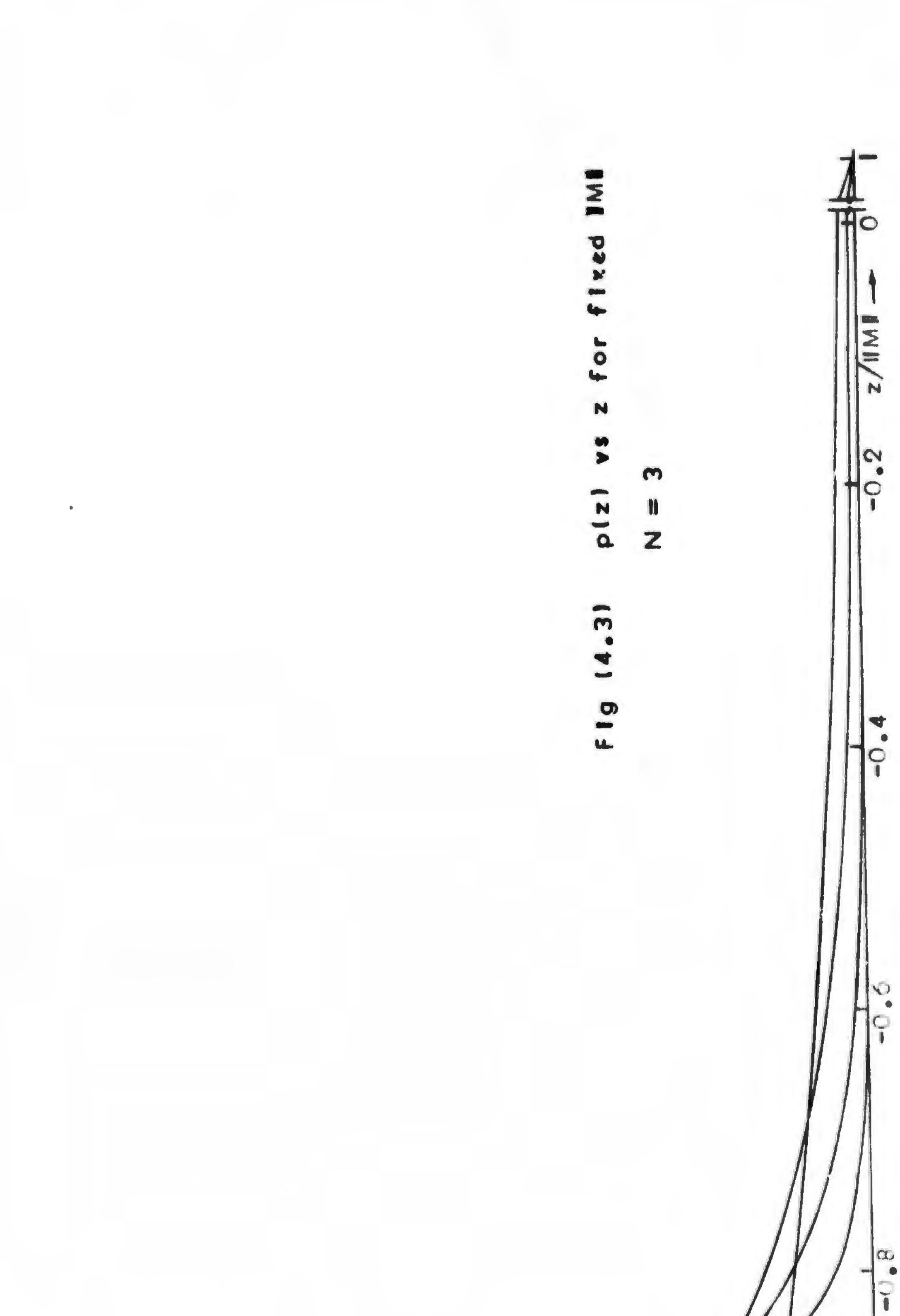


Fig (4.3) p(z) vs z for fixed ‖M‖

N = 3

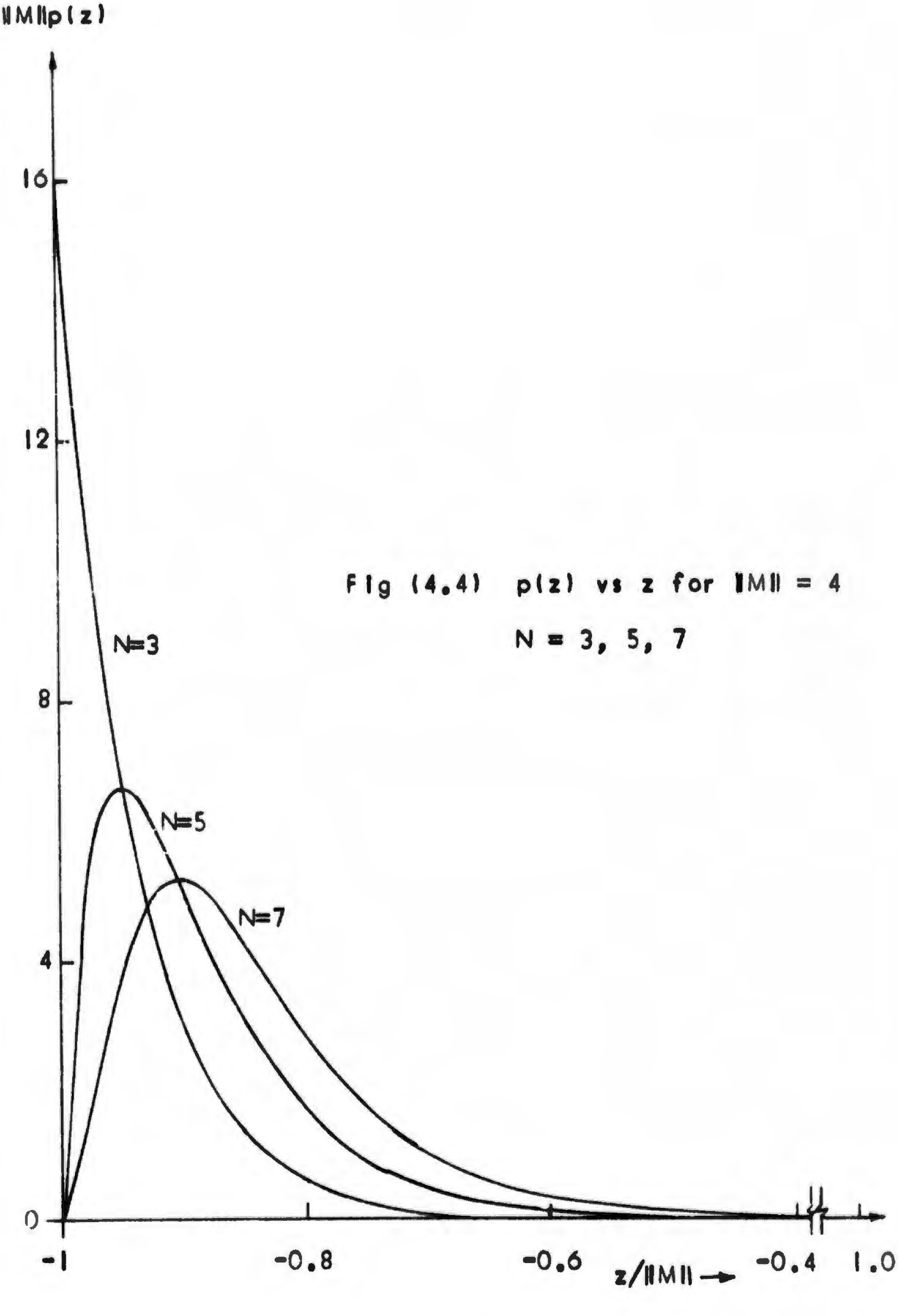


Fig (4.4) p(z) vs z for ‖M‖ = 4

N = 3, 5, 7

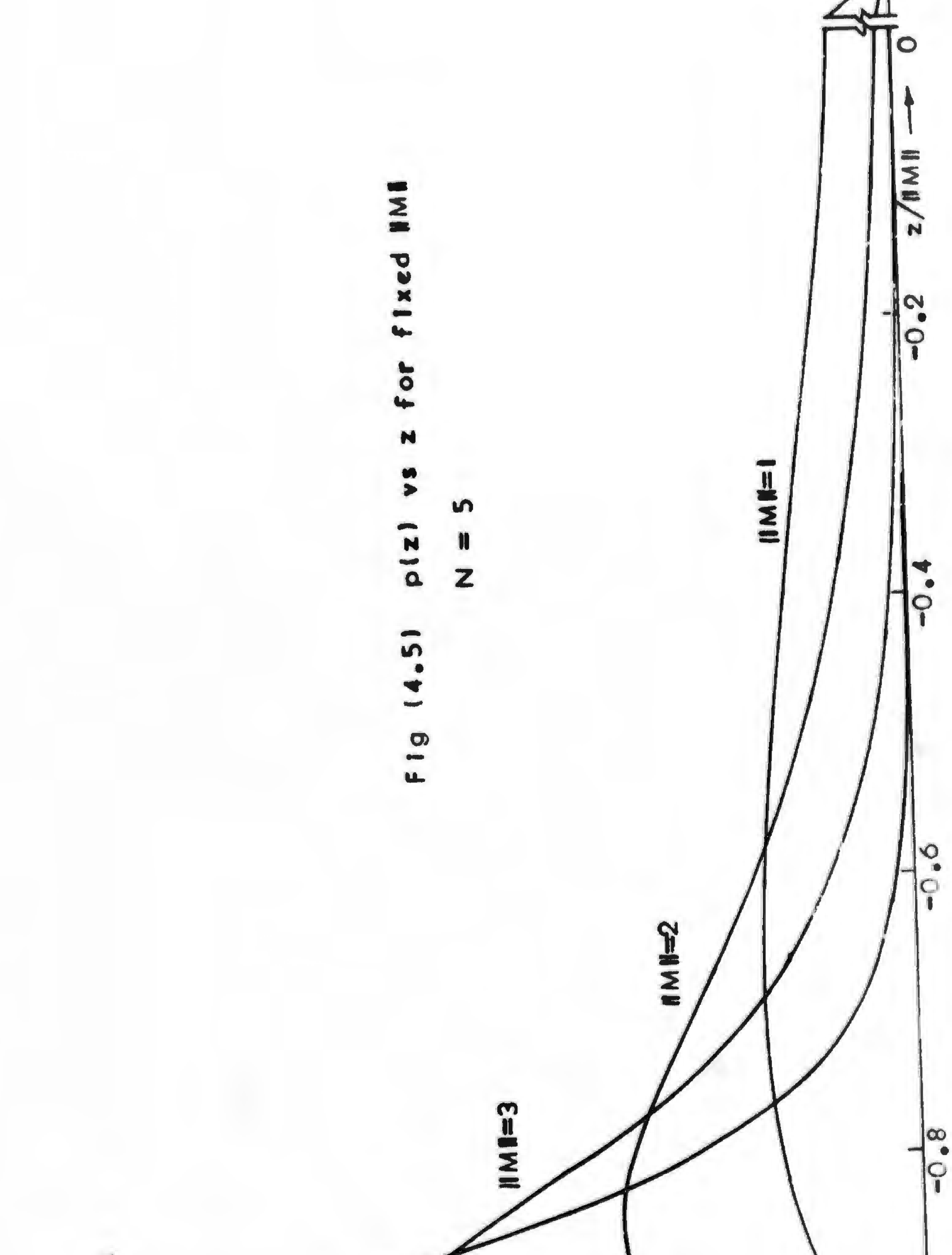


Fig (4.5) p(z) vs z for fixed ‖M‖

N = 5

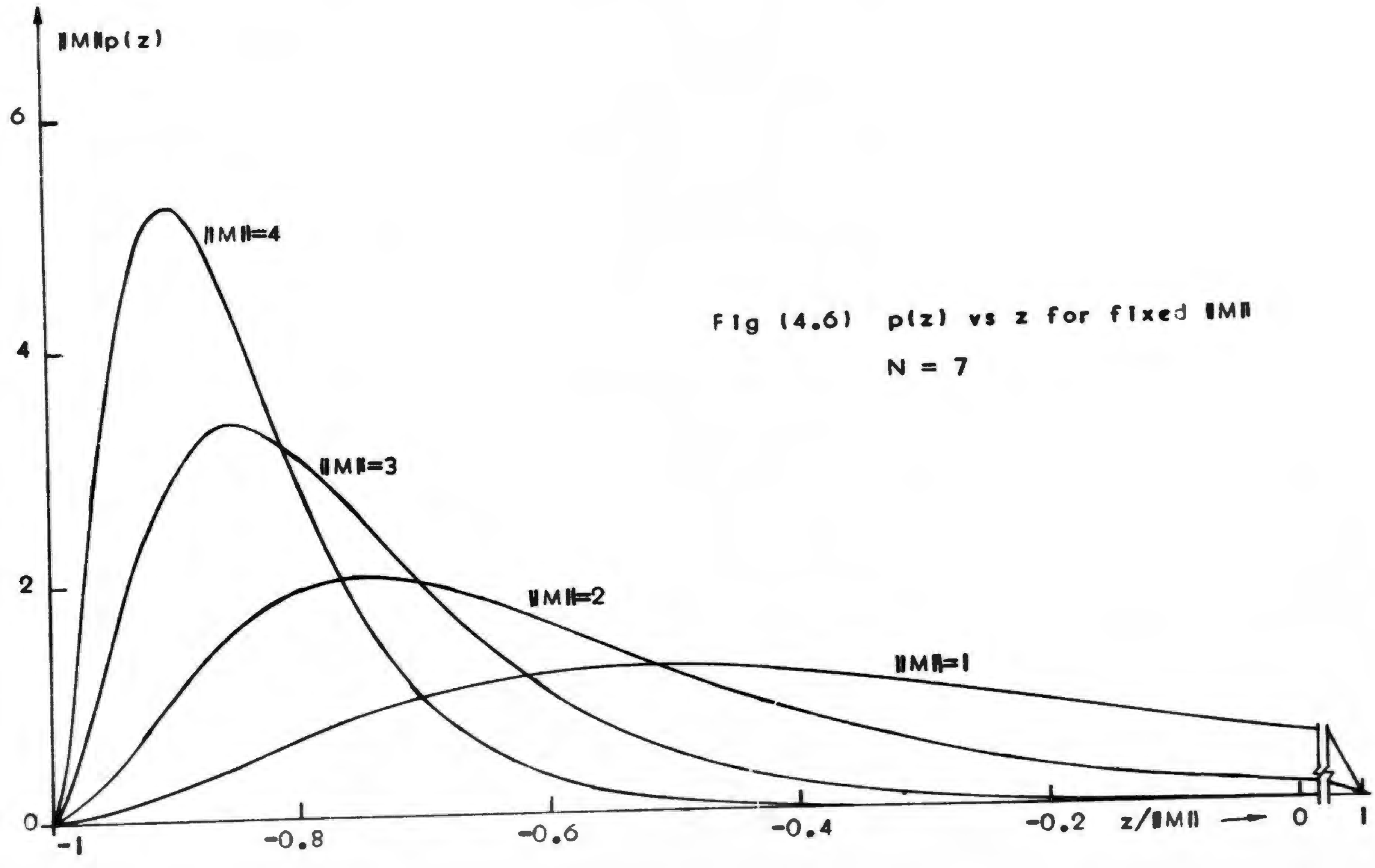


Fig (4.6) p(z) vs z for fixed ‖M‖

N = 7

4.6 REGIONS OF ERROR

We investigate in this section the regions of error corresponding to the general composite decision rule (4.12), repeated below,

$$Z_\eta = \sum |x_i|^\eta \operatorname{sgn} x_i \gtrless 0 \qquad (4.12)$$

4.6.1 REGION OF ERROR ~ $\eta = 1$

Based on the description of the decision hypersurface discussed in section (4.4) and the transformation T introduced in (4.25) we first note that the region of error corresponding to $\eta = 1$, is given by,

$$\mathcal{A} = \left\{ X : \sum x_i < 0 \right\} = \left\{ X : [B,X] < 0 \right\} \qquad (4.39)$$

where B, the major axis, is the vector,

$$B = \left\{ 1, 1, 1, \ldots\ldots, 1 \right\}$$

Thus, for $\eta = 1$, the region of error is,

$$= \left\{ X : [B,TY] < 0 \right\}$$

$$= \left\{ X : [T^{*}B,Y] < 0 \right\}$$

Invoking the generalized spherical coordinate change of variables, let the coordinates of Y be,

$$\begin{aligned}
y_j &= r \left[\prod_{k=1}^{j-1} \sin \Phi_k \right] \cos \Phi_j & j &= 1,\ldots,N-2 \\
& & r &\geq 0 \\
y_{N-1} &= r \left[\prod_{k=1}^{N-2} \sin \Phi_k \right] \cos \theta & 0 &\leq \Phi_k \leq \pi, k=1,..,N-2 \\
y_N &= r \left[\prod_{k=1}^{N-2} \sin \Phi_k \right] \sin \theta & 0 &\leq \theta \leq 2\pi
\end{aligned} \qquad (4.40)$$

similarly, let the coordinates of the vector $F = T^*B$, be given by,

$$f_j = \rho \left[\prod_{k=1}^{j-1} \sin \psi_k \right] \cos \psi_j \qquad j = 1, \ldots, N-2$$

$$f_{N-1} = \rho \left[\prod_{k=1}^{N-2} \sin \psi_k \right] \cos \nu \qquad \rho \geq 0$$

$$\qquad 0 \leq \psi_k \leq \pi, k = 1, .., N-2$$

$$f_N = \rho \left[\prod_{k=1}^{N-2} \sin \psi_k \right] \sin \nu \qquad 0 \leq \nu \leq 2\pi \qquad (4.41)$$

Let the angle subtended at the origin by the major and polar axes be denoted by,

$$\mu = \cos^{-1} \left\{ \frac{[B,M]}{\sqrt{N}\ \|M\|} \right\} \qquad (4.42)$$

Then since e_1' is in the direction of T^*M we have,

$$\psi_1 = \mu \qquad (4.43)$$

and the region of error is that for which,

$$[T^*B,Y] = [F,Y] = \cos\Phi_1 \cos\mu + \sin\Phi_1 \sin\mu \left\{ \cos\Phi_2 \cos\psi_2 + \ldots + \sin\Phi_{N-2} \sin\psi_{N-2} \cos(\Theta - \nu) \right\} \qquad (4.44)$$

is less than zero.

Now, since the only constraint is that the angle between B and M be fixed at μ, it is always possible to suppress the other variables in (4.44) by arbitrarily choosing $\psi_2 = 0$, i.e. choosing it as the reference from which Φ_2 is measured. Thus we may express the region of error as that for which,

$$\cos\Phi_1 \cos\mu + \sin\Phi_1 \sin\mu \cos\Phi_2 < 0 \qquad (4.45)$$

any other choice for ψ_2 only reflects itself in a cyclic shift of reference. The choice we have made corresponds to measuring Φ_2 from the hyperplane formed by B and M, i.e. from the projection of T^*B on the (N-1)-flat described by,

$$\Phi_1 = \frac{\pi}{2}$$

In $E^{(3)}$, corresponding to (4.45) we may write the region of error as that for which,

$$\cos\Phi_1 \cos\mu + \sin\Phi_1 \sin\mu \cos\theta < 0 \qquad (4.45a)$$

which reflects choosing $\nu = 0$ as the reference. These regions may, alternatively, be written as,

$$\left.\begin{array}{ll} \tan\Phi_1 > -\cot\mu\sec\Phi_2 & \begin{cases} \frac{\pi}{2} < \Phi_1 < \pi \\ 0 < \Phi_2 < \frac{\pi}{2} \end{cases} \\ \tan\Phi_1 < -\cot\mu\sec\Phi_2 & \begin{cases} 0 < \Phi_1 < \frac{\pi}{2} \\ \frac{\pi}{2} < \Phi_2 < \pi \end{cases} \end{array}\right. \qquad (4.46)$$

and for $E^{(3)}$,

$$\left.\begin{array}{ll} \tan\Phi_1 > -\cot\mu\sec\theta & \begin{cases} \frac{\pi}{2} < \Phi_1 < \pi \\ \frac{3\pi}{2} < \theta < \frac{\pi}{2} \end{cases} \\ \tan\Phi_1 < -\cot\mu\sec\theta & \begin{cases} 0 < \Phi_1 < \frac{\pi}{2} \\ \frac{\pi}{2} < \theta < \frac{3\pi}{2} \end{cases} \end{array}\right. \qquad (4.46a)$$

4.6.2 REGION OF ERROR~η = 2

For $\eta = 2$, the region of error cannot be expressed in terms of a single expression as was the case with $\eta = 1$.

For this case, each segment of the decision hypersurface must be exemined separately by considering the different orthants through which the hypersurface passes; and for this purpose we introduce the following notation:

Consider a specific orthant, and note the signs of its bounding axes. Let the vector,

$$\underline{e} = \left[\pm 1, \pm 1, \pm 1, \ldots\ldots\ldots, \pm 1\right] \qquad (4.47)$$

be the vector whose elements are either +1 or -1 according as the signs of the bounding axes. Let

$$Q_{\underline{e}} = \begin{bmatrix} \pm 1 & & & & \\ & \pm 1 & & & \\ & & \pm 1 & & \\ & & & \ddots & \\ & & & & \pm 1 \end{bmatrix} \qquad (4.48)$$

be a diagonal matrix, not all of whose diagonal entries are positive or negative, and which corresponds to the vector $\underline{e}$ in the sense that,

$$q_{ii} = \begin{cases} +1 & \text{if } e_i = +1 \\ -1 & \text{if } e_i = -1 \end{cases} \qquad i = 1, \ldots, N \qquad (4.49)$$

Then, in each orthant, the equation of the decision hypersurface may be written,

$$Z_2 = \sum_i x_i^2 \operatorname{sgn} x_i = [X, Q_{\underline{e}} X] = 0 \qquad (4.50)$$

The region of error corresponding to each orthant is then,

$$\begin{aligned} \mathscr{A} &= \left\{ X : [X, Q_{\underline{e}} X] < 0 \right\} \\ &= \left\{ Y : [TY, Q_{\underline{e}} TY] < 0 \right\} \\ &= \left\{ Y : [Y, T^{*} Q_{\underline{e}} TY] < 0 \right\} \end{aligned} \qquad (4.51)$$

Introducing the generalized spherical coordinate change of variables given in (4.40), the quadratic form $[Y, T^{*} Q_{\underline{e}} TY]$ leads then to a quadratic equation in $\tan \Phi_1$, the roots of which delineate the region of error for the orthant.

By cyclically considering all the possible orthants which are bounded by axes not all of whose signs are positive or negative, one arrives at a complete description of the error region.

An example, presented in detail in section (4.8), considers these transformations in $E^{(3)}$.

We remark at this point that since the angle Φ subtended at the origin by the major axis B, and any vector within the decision hypersurface, oscillates about the value $\frac{\pi}{2}$, the decision hypersurface for $\eta = 2$ undulates about and is symmetric with respect to that for $\eta = 1$.

The region of error corresponding to $\eta = 2$, will therefore not uniformly contain or be contained in that corresponding to $\eta = 1$, but will undergo changes from one orthant to another.

For η, other than equal to 1 or 2, there is no mathematically compact way of expressing the region of error. We note, however, that for each orthant, the decision rule is a real-valued continuous function on the connected space which is the orthant. Hence, the decision hypersurface separates the orthant into two subspaces, one on which $Z_\eta(.) > 0$; the other, on which $Z_\eta(.) < 0$.

We do not intend to investigate in detail the merits of the composite decision rule for general η; we simply remark at this juncture that since the decision hypersurface for $\eta \neq 1$ undulates about that for $\eta = 1$, the points of interest can be studied for $\eta = 1$ and $\eta = 2$, the inference for $\eta > 2$ can then be drawn from the quadratic rule.

4.7 PROBABILITY OF ERROR

In terms of the decision variables $[x_i]$, we expressed the average probability of error in (4.24) as,

$$P_e = \int_{\sum |x_i|^{\eta - 1} \; x_i < 0} p(X)\, dX \qquad (4.52)$$

Based, then, on the transformation to generalized spherical coordinates and on the arguments delineating the region of error in $E^{(N)}$, we may express (4.52) in the form,

$$P_e = \int_0^{\infty} dr \int_0^{2\pi} d\theta \int_0^{\pi} d\Phi_2 \cdots \int_0^{\pi} d\Phi_{N-2} \int_{h_\eta(\theta,\Phi_2,\ldots,\Phi_{N-2})}^{\pi} d\Phi_1 \, p(r,\theta,\Phi_1,\Phi_2,\ldots,\Phi_{N-2}) \qquad (4.53)$$

where $h_\eta(\cdot,\cdot,\ldots,\cdot)$ describes the functional relationship governing the variations of the decision hypersurface with the orthogonal coordinates $[\theta,\Phi_1,\Phi_2, \ldots ,\Phi_{N-2}]$.

In $E^{(3)}$, for example, this is simply,

$$P_e = \int_0^{2\pi} d\theta \int_0^{\infty} dr \int_{h_\eta(\theta)}^{\pi} d\Phi_1 \; p(r,\theta,\Phi_1) \qquad (4.54)$$

which upon integrating over r reduces to,

$$P_e = \frac{1}{2\pi} \int_0^{2\pi} d\theta \int_{h_\eta(\theta)}^{\pi} p(\Phi_1) \, d\Phi_1 \qquad (4.54a)$$

For $\eta = 1$, the function $h_1(\theta,\Phi_2, \ldots ,\Phi_{N-2})$, in general, is given by,

$$h_1(\theta,\Phi_2, \ldots ,\Phi_{N-2}) = \Phi_1 = \tan^{-1}(- \cot\mu \, \sec\Phi_2) \qquad (4.55)$$

while in $E^{(3)}$, it is given by,

$$h_1(\theta) = \Phi_1 = \tan^{-1}(- \cot\mu \, \sec\theta) \qquad (4.55a)$$

It is seen, therefore, that the effect of the variation of the decision hypersurface with the other coordinate variables Φ_k, $k \geq 3$, have been suppressed. This is not coincidental, for by construction, we have made all the other variables in the generalized spherical distribution (4.29) insensitive to the energy parameter $\|M\|$. The only variations with $\|M\|$ are with respect to r, which is integrated out, and to Φ_1, which is reflected in the skewness of the density function $p(\Phi_1)$ as a function of the location of the centroid of the distribution.

For $\eta = 2$, the decision hypersurface $h_2(\theta,\Phi_2,\ldots,\Phi_{N-2})$ is a function of all the generalized spherical coordinate variables. The probability of error, strictly speaking, will then depend on the exact variations of the decision hypersurface with these variables. Furthermore, the error probability is not only a function of the angle, μ, by which the polar and major axes diverge, but, to a certain measure, will depend on the relative magnitudes of the elements of the mean vector M. Thus exact evaluation of the error probability may be accomplished computationally in individual cases.

Despite this sensitivity of the error probability to the actual values of the mean parameters, it is desirable to obtain an approximation for the purpose of comparison with the majority logic mode of operation discussed in the next chapter. It is found that because the variations with respect to the energy parameter $\|M\|$ are largely concentrated along Φ_1, that the error probability may be approximated by,

$$P_e \approx P\left[\frac{\pi}{2}+\mu \le \Phi_1 \le \pi\right] + \frac{1}{2} P\left[\frac{\pi}{2}-\mu \le \Phi_1 \le \frac{\pi}{2}+\mu\right] \tag{4.56}$$

In terms of the new variable 'z' of (4.38), equation (4.56) may be written,

$$P_e \approx P[\|M\| \sin\mu < z < \|M\|] + \frac{1}{2} P[-\|M\| \sin\mu < z < \|M\| \sin\mu] \tag{4.57}$$

Now, for 'n' a positive integer, the Parabolic Cylinder function $D_{-n}(z)$ can be expressed [15], in terms of the error function defined in (3.36) and (3.37),

$$D_{-n-1}(z) = \sqrt{2\pi}\,\frac{(-1)^n}{n!} \exp\left(-\frac{1}{4} z^2\right) \frac{d^n}{dz^n}\left\{e^{z^2/2} \operatorname{erfc}(z)\right\} \tag{4.58}$$

Invoking this relationship, p(z) of (4.38) may be expressed as,

$$p(z) = K_N'' \frac{\exp(-\|M\|^2/2)}{\|M\|} \left[1 - \left(\frac{z}{\|M\|}\right)^2\right]^{\frac{N-3}{2}} \frac{d^{N-1}}{dz^{N-1}} \left\{e^{z^2/2} \operatorname{erfc}(z)\right\} \tag{4.59}$$

where,

$$K_N'' = \frac{(-1)^{N-1}}{2^{\frac{N-3}{2}}\,\Gamma\left(\frac{N-1}{2}\right)}$$

Expressions of the probability of error may be put in closed form only for N, odd. We restrict consideration to this case only, and for this purpose we examine the following general integral,

$$I_{N,i}(a,b) = \int_a^b P_i(z) \left[1 - \left(\frac{z}{\|M\|}\right)^2\right]^{\frac{N-3-2i}{2}} \frac{d^{N-1-i}}{dz^{N-1-i}} \left\{ e^{z^2/2} \mathrm{erfc}(z) \right\} dz$$

$$a \leqslant b \leqslant \|M\| \qquad (4.60)$$

where $P_i(z)$ is a polynomial in z of degree i, with $P_o(z)=1$.

On integrating (4.60) once by parts, we let,

$$u_i = P_i(z) \left[1 - \left(\frac{z}{\|M\|}\right)^2\right]^{\frac{N-3-2i}{2}} \qquad (4.61)$$

$$v_i = \frac{d^{N-1-(i+1)}}{dz^{N-1-(i+1)}} \left\{ e^{z^2/2} \mathrm{erfc}(z) \right\} \qquad (4.62)$$

thus,

$$du_i = \left[1 - \left(\frac{z}{\|M\|}\right)^2\right]^{\frac{N-3-2(i+1)}{2}} \left\{ \left[1 - \left(\frac{z}{\|M\|}\right)^2\right] P_i'(z) - (N-3-2i)\frac{z}{\|M\|} P_i(z) \right\} dz \qquad (4.63)$$

Hence,

$$I_{N,i}(a,b) = u_i v_i \Big|_a^b - \int_a^b P_{i+1}(z) \left[1 - \left(\frac{z}{\|M\|}\right)^2\right]^{\frac{N-3-2(i+1)}{2}} \frac{d^{N-1-(i+1)}}{dz^{N-1-(i+1)}} \left\{ e^{z^2/2} \mathrm{erfc}(z) \right\} dz$$

$$= u_i v_i \Big|_a^b - I_{N,i+1}(a,b) \qquad (4.64)$$

where,

$$P_{i+1}(z) = \left[1 - \left(\frac{z}{\|M\|}\right)^2\right] P_i'(z) - (N-3-2i)\frac{z}{\|M\|}P_i(z) \qquad (4.65)$$

The integral we seek for the evaluation of the approximate expression of P_e is $I_{N,0}(a,b)$. Therefore, on performing (N-3) repeated integrations by parts, we finally have,

$$I_{N,0}(a,b) = \sum_{i=0}^{N-4} u_i v_i \Big|_a^b + I_{N,N-3}(a,b) \qquad (4.66)$$

where,

$$I_{N,N-3}(a,b) = \int_a^b P_{N-3}(z) \left[1 - \left(\frac{z}{\|M\|}\right)^2\right]^{-\frac{N-3}{2}} \frac{d^2}{dz^2}\left\{e^{z^2/2}\operatorname{erfc}(z)\right\} dz \qquad (4.67)$$

It can easily be shown, however, that $P_{N-3}(z)$ always reduces to,

$$P_{N-3}(z) = (-1)^{\frac{N-3}{2}} \frac{(N-3)!}{\|M\|^{N-3}} \left[1 - \left(\frac{z}{\|M\|}\right)^2\right]^{\frac{N-3}{2}}$$

with the result that,

$$I_{N,N-3}(a,b) = \frac{(-1)^{\frac{N-3}{2}}(N-3)!}{\|M\|^{N-3}} \int_a^b \frac{d^2}{dz^2}\left\{e^{z^2/2}\operatorname{erfc}(z)\right\} dz$$

$$= \frac{(-1)^{\frac{N-3}{2}}(N-3)!}{\|M\|^{N-3}} I_{3,0}(a,b) \qquad (4.68)$$

But $I_{3,0}(a,b)$ is easily integrated to yield,

$$I_{3,0}(a,b) = \frac{d}{dz}\left\{e^{z^2/2}\operatorname{erfc}(z)\right\}\Big|_a^b \qquad (4.69)$$

We note that as a result of the reduction of $I_{N,0}(a,b)$ we get terms involving up to the (N-2)nd. derivative of $e^{z^2/2}$ erfc(z). Therefore, in order to have a general formulation for these derivatives, consider the function,

$$\omega^{(i)}(z) = f_i(z)\, e^{z^2/2}\, \mathrm{erfc}(z) - \frac{1}{\sqrt{2\pi}}\, g_i(z) \qquad (4.70)$$

where $f_i(z)$ is a polynomial in z of degree i, with $f_0(z) = 1$; and $g_i(z)$ is a polynomial in z of degree (i-1), with $g_0(z) = 0$.

On differentiating $\omega^{(i)}(z)$ once, we obtain,

$$\omega^{(i+1)}(z) = [f_i'(z) + z f_i(z)]\, e^{z^2/2}\mathrm{erfc}(z) - \frac{1}{\sqrt{2\pi}}[f_i(z) + g_i'(z)]$$

$$= f_{i+1}(z)\, e^{z^2/2}\mathrm{erfc}(z) - \frac{1}{\sqrt{2\pi}}\, g_{i+1}(z) \qquad (4.71)$$

Hence, we have the following recursive relations,

$$f_i(z) = z\, f_{i-1}(z) + f_{i-1}'(z) \quad ; \quad f_0(z) = 1 \qquad (4.72)$$

$$g_i(z) = f_{i-1}(z) + g_{i-1}'(z) \quad ; \quad g_0(z) = 0 \qquad (4.73)$$

These polynomials are tabulated in appendix 'C' for selected values of i.

Combining these results, it follows that,

$$P_e \approx K_N'' \frac{e^{-\|M\|^2}}{\|M\|} \left\{ I_{N,0}(\|M\|\sin\mu, \|M\|) + \frac{1}{2}\, I_{N,0}(-\|M\|\sin\mu, \|M\|\sin\mu) \right\} \qquad (4.74)$$

With this formulation we find, for example, that for $N = 3$,

$$P_e \approx \mathrm{erfc}(\|M\|) + \sin\mu \, \exp\left\{-\frac{1}{2}\|M\|^2\cos^2\mu\right\}\left[\frac{1}{2} - \mathrm{erfc}(\|M\|\sin\mu)\right] \tag{4.75}$$

and if $\mu = 0$, the second term vanishes. It is evident that, since the second term is non-negative, the probability of error achieves its minimum when all the channels have received signals of equal normalized energies.

4.8 EXAMPLE

In this section we present a detailed outline of the transformations and the resulting region of error in terms of the spherical coordinates in $E^{(3)}$. We consider three distinct cases for the location of the mean vector.

Case I:

$$M = [m_1, 0, 0], \quad m_1 \neq 0$$
$$B = [1, 1, 1], \quad \mu = \cos^{-1}\sqrt{\tfrac{1}{3}}$$

The region of error corresponding to $\eta = 1$, is that for which $[B,X] < 0$. Clearly, the rotation of axes transformation, T, is not required for this case since M coincides with one of the axes; or we simply let $T = I$ and $Y = X$. Thus, introducing the spherical coordinate change of variables,

$$\left.\begin{aligned} y_1 &= r\cos\Phi_1 \\ y_2 &= r\sin\Phi_1\cos\theta \\ y_3 &= r\sin\Phi_1\sin\theta \end{aligned}\right\} \quad \begin{aligned} &0 \leq r \\ &0 \leq \Phi_1 \leq \pi \\ &0 \leq \theta \leq 2\pi \end{aligned} \tag{4.76}$$

yields,

$$[B,X] = [B,Y] = \cos\Phi_1 \cos\mu + \sin\Phi_1 \sin\mu \cos(\theta - \nu) < 0$$

Choosing $\nu = 0$ as the reference, leads to,

$$\Phi_1 \leq \tan^{-1}\left\{-\cot\mu \sec\theta\right\} \tag{4.77}$$

the equality, giving the equation of the decision surface $h_1(\theta)$.

To determine the variations of the decision surface corresponding to $\eta = 2$, consider the octant bounded by,

$$[\ e_1 > 0,\ e_2 < 0,\ e_3 < 0\ ]$$

The region of error for this octant is that for which $[Y, T^{*}Q_{\underline{e}}TY] < 0$, where,

$$Q_{\underline{e}} = \begin{bmatrix} +1 & & \\ & -1 & \\ & & -1 \end{bmatrix} \qquad ; \qquad T = I$$

Using the change of variables (4.76), this condition corresponds to the region of error being,

$$\mathcal{S} = \left\{ \Phi_1, \theta \ : \ \frac{\pi}{4} < \Phi_1 < \frac{\pi}{2}\ ,\ \frac{3\pi}{4} < \theta < \frac{5\pi}{4} \right\} \tag{4.78}$$

In the image-through-origin octant, it is evident that

$$Q_{\underline{e}} = \begin{bmatrix} -1 & & \\ & +1 & \\ & & +1 \end{bmatrix}$$

and the region of error becomes that for which,

$$\mathscr{A} = \left\{ \Phi_1, \theta \; : \; \frac{3\pi}{4} < \Phi_1 < \pi \; , \; \frac{7\pi}{4} < \theta < \frac{\pi}{4} \right\} \tag{4.79}$$

Consider next the orthant,

$$[\; e_1 > 0, \; e_2 < 0, \; e_3 > 0 \;]$$

Here again, the region of error is that for which,

$[X, Q_{\underline{e}}X] = [Y, Q_{\underline{e}}Y] < 0$, with,

$$Q_{\underline{e}} = \begin{bmatrix} +1 & & \\ & -1 & \\ & & +1 \end{bmatrix}$$

With the change of variables of (4.76), this condition is equivalent to,

$$\mathscr{A} = \left\{ \Phi_1, \theta \; : \; \Phi_1 > \tan^{-1}\sqrt{\frac{-1}{\cos 2\theta}} \; , \; \frac{\pi}{2} < \theta < \frac{3\pi}{4} \right\} \tag{4.80}$$

In the image-through-origin octant, it is evident that,

$$\mathscr{A} = \left\{ \Phi_1, \theta \; : \; \Phi_1 > \tan^{-1}\sqrt{\frac{-1}{\cos 2\theta}} \; , \frac{3\pi}{2} < \theta < \frac{7\pi}{4} \right\} \tag{4.81}$$

Similarly for the octant,

$$[\; e_1 < 0, \; e_2 < 0, \; e_3 > 0 \;]$$

we have,

$$Q_{\underline{e}} = \begin{bmatrix} -1 & & \\ & -1 & \\ & & +1 \end{bmatrix}$$

and the region of error reduces to,

$$\mathcal{S} = \left\{ \Phi_1, \theta \; : \; \Phi_1 > \tan^{-1}\sqrt{\frac{-1}{\cos 2\theta}} \; , \; \frac{\pi}{2} < \theta < \frac{\pi}{4} \right\} \qquad (4.82)$$

and in the image octant,

$$\mathcal{S} = \left\{ \Phi_1, \theta \; : \; \Phi_1 > \tan^{-1}\sqrt{\frac{-1}{\cos 2\theta}} \; , \; \frac{5\pi}{4} < \theta < \frac{3\pi}{2} \right\} \qquad (4.83)$$

These regions are summed up by the figure (4.7) which is a development of the error regions corresponding to $\eta = 1$ and $\eta = 2$, in terms of the integrals over the Φ_1 and θ variables space. The probability of error being the integral of $p(\Phi_1, \theta)$ over the region to the right of the decision surface variations.

The average probability of error may then be written,

$$P_{e_1} = \frac{1}{2\pi} \int_0^{2\pi} d\theta \int_{\tan^{-1}\left\{-\frac{\sec\theta}{\sqrt{2}}\right\}}^{\pi} p(\Phi_1) \, d\Phi_1 \qquad (4.84)$$

$$P_{e_2} = 2 \left\{ \int_0^{\frac{\pi}{4}} p(\theta) d\theta \int_{\frac{3\pi}{4}}^{\pi} p(\Phi_1) d\Phi_1 + \int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} p(\theta) d\theta \int_{\tan^{-1}\sqrt{-\sec 2\theta}}^{\pi} p(\Phi_1) d\Phi_1 + \int_{\frac{3\pi}{4}}^{\pi} p(\theta) d\theta \int_{\frac{\pi}{4}}^{\pi} p(\Phi_1) \, d\Phi_1 \right\} \qquad (4.85)$$

These expressions are graphed in fig (4.8).

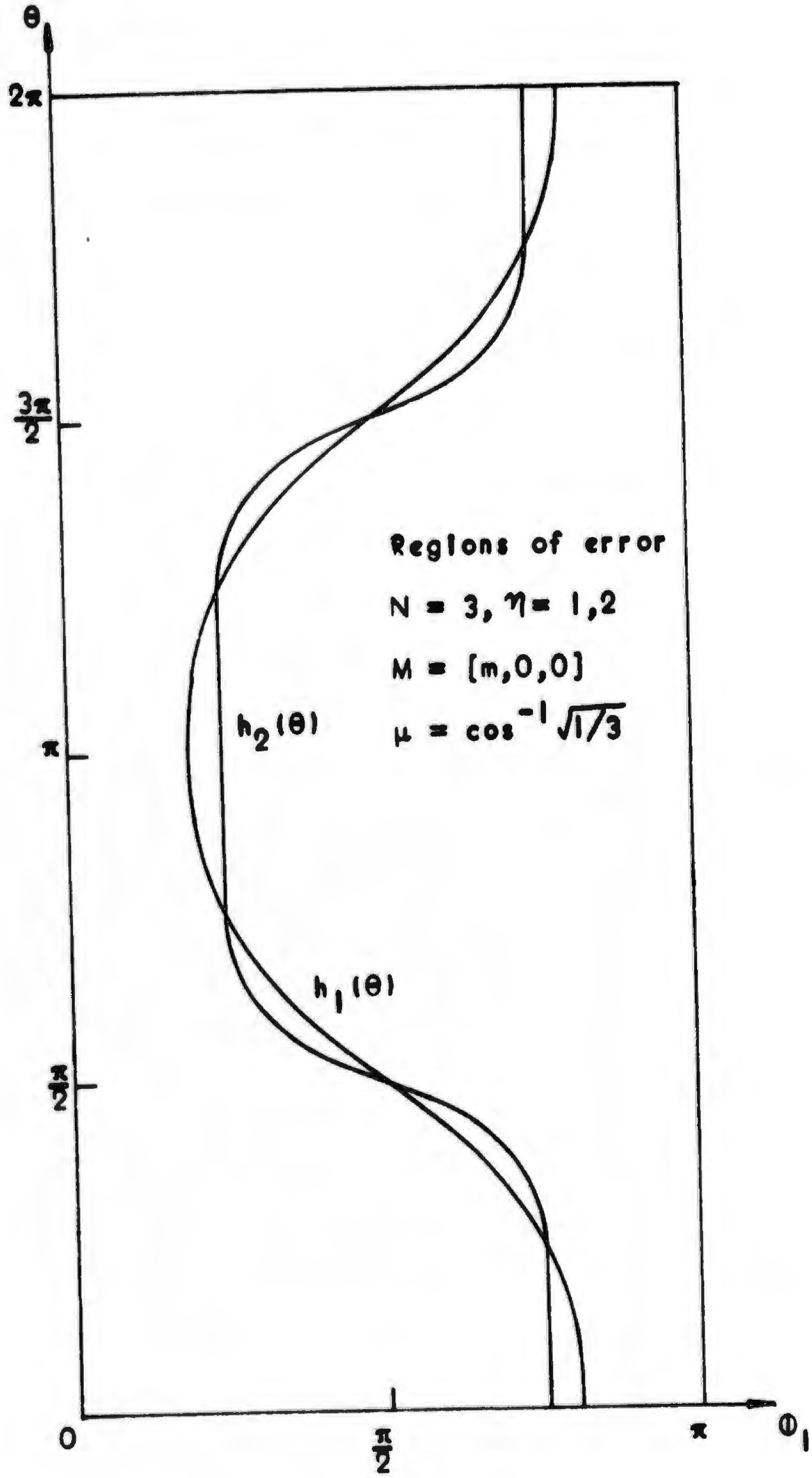

θ
2π
3π/2
π
π/2
0
π/2
π
Φ
Regions of error
N = 3, η = 1,2
M = [m,0,0]
μ = cos⁻¹ √1/3
h2(θ)
h1(θ)

Fig (4.7)

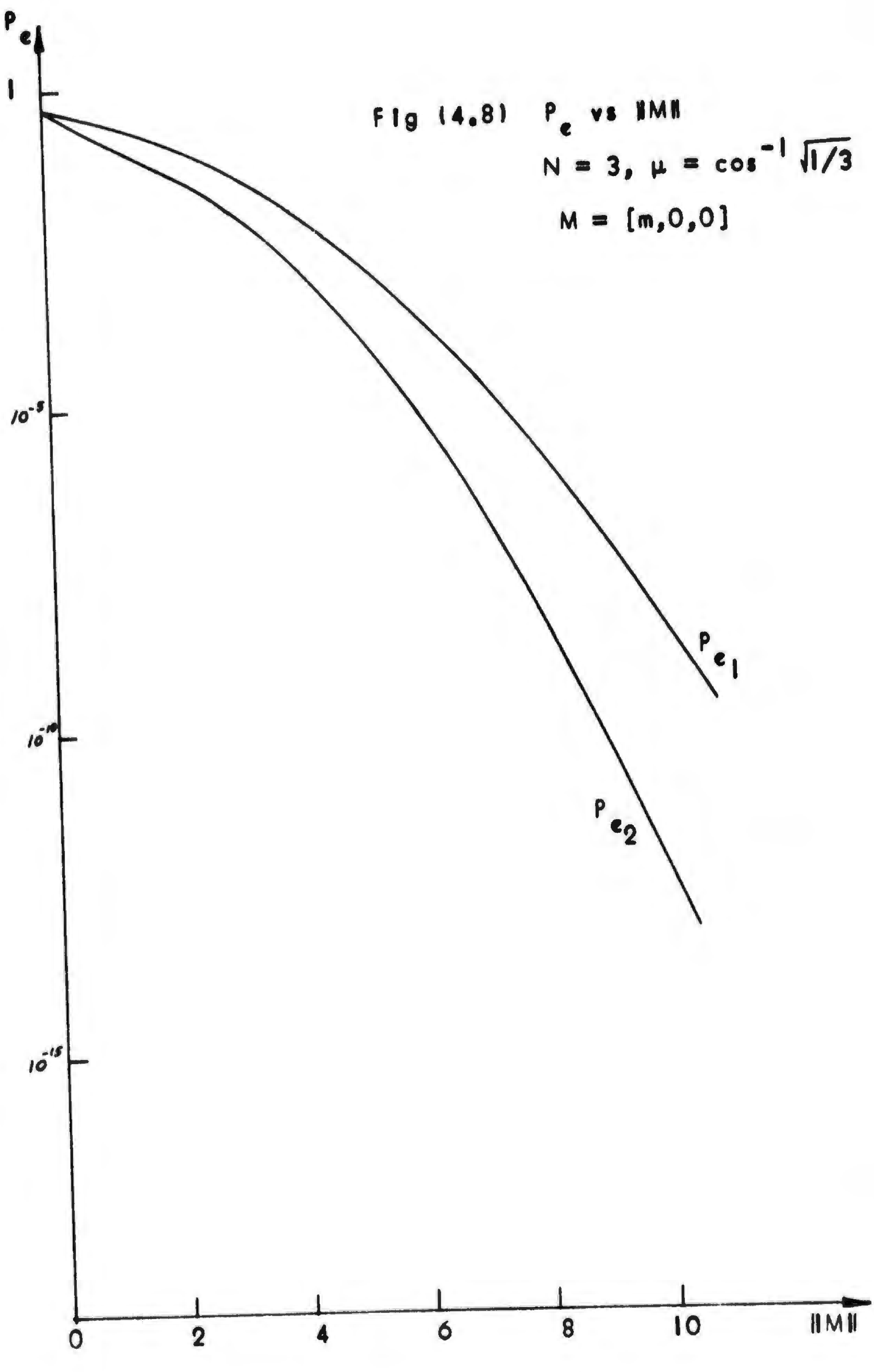


Fig (4.8) P_e vs ‖M‖
N = 3, μ = $\cos^{-1}\sqrt{1/3}$
M = [m,0,0]
P_e
1
10^{-5}
10^{-10}
10^{-15}
P_{e_1}
P_{e_2}
0
2
4
6
8
10
‖M‖

Case 2:

$$M = [\ 0,\ m_2,\ m_3\ ]\quad,\qquad m_2 = m_3 \neq 0$$

$$B = [\ 1,\ 1,\ 1\ ]\qquad\qquad \mu = \cos^{-1}\sqrt{\frac{2}{3}}$$

The transformation T of (4.25) that takes $[e_1, e_2, e_3]$ into $[e_1', e_2', e_3']$ is readily found to be,

$$T = \begin{bmatrix} 0 & 1 & 0 \\ \sqrt{1/2} & 0 & \sqrt{1/2} \\ \sqrt{1/2} & 0 & -\sqrt{1/2} \end{bmatrix} \tag{4.86}$$

It is evident that the region of error corresponding to $\eta = 1$ is given by (4.77) except for the value of μ which is given above. This region is shown in fig (4.9).

The region of error corresponding to $\eta = 2$, is determined in the same manner as with the previous case. Thus consider the octant,

$$[\ e_1 < 0,\ e_2 > 0,\ e_3 > 0\ ]$$

and we have,

$$Q_{\underline{e}} = \begin{bmatrix} -1 & & \\ & +1 & \\ & & +1 \end{bmatrix} \qquad\qquad T^{*}Q_{\underline{e}}T = \begin{bmatrix} +1 & & \\ & -1 & \\ & & +1 \end{bmatrix}$$

and in terms of the spherical coordinate variables, the portion of the decision surface in this octant is described by,

$$h_2(\theta) = \Phi_1 = \tan^{-1}\sqrt{\sec 2\theta}\ ;\ \pi - \tan^{-1}\frac{1}{\sqrt{2}} < \theta < \pi + \tan^{-1}\frac{1}{\sqrt{2}} \tag{4.87}$$

and the same relation applies in the image octant. By a similar procedure we find that for the octant,

$$[\ e_1 > 0,\ e_2 > 0,\ e_3 < 0\ ]$$

$$Q_{\underline{e}} = \begin{bmatrix} +1 & & \\ & +1 & \\ & & -1 \end{bmatrix} \quad ; \quad T^{\#}Q_{\underline{e}}T = \begin{bmatrix} & & +1 \\ & +1 & \\ +1 & & \end{bmatrix}$$

thus,

$$h_2(\theta) = \Phi_1 = \tan^{-1}[-2\tan\theta\ \sec\theta]\ ;\ \frac{\pi}{2} < \theta < \pi - \tan^{-1}\frac{1}{\sqrt{2}} \tag{4.88}$$

and the same relation applies in the image octant.

By symmetry, similar relations apply for the other octants. The region of error corresponding to $\eta = 2$ is summed up in fig (4.9).

The average probability of error for this case may then be written,

$$P_{e_1} = \int_0^{2\pi} p(\theta)d\theta \int_{\tan^{-1}\{-\sqrt{2}\ \sec\theta\}}^{\pi} p(\Phi_1)\ d\Phi_1 \tag{4.89}$$

$$P_{e_2} = 2\left\{ \int_0^{\tan^{-1}\frac{1}{\sqrt{2}}} p(\theta)d\theta \int_{\tan^{-1}\sqrt{\sec 2\theta}}^{\pi} p(\Phi_1)\ d\Phi_1 \right.$$

$$+ \int_{\tan^{-1}\frac{1}{\sqrt{2}}}^{\pi/2} p(\theta)\ d\theta \int_{\tan^{-1}\{-\ 2\tan\theta\ \sec\theta\}}^{\pi} p(\Phi_1)\ d\Phi_1$$

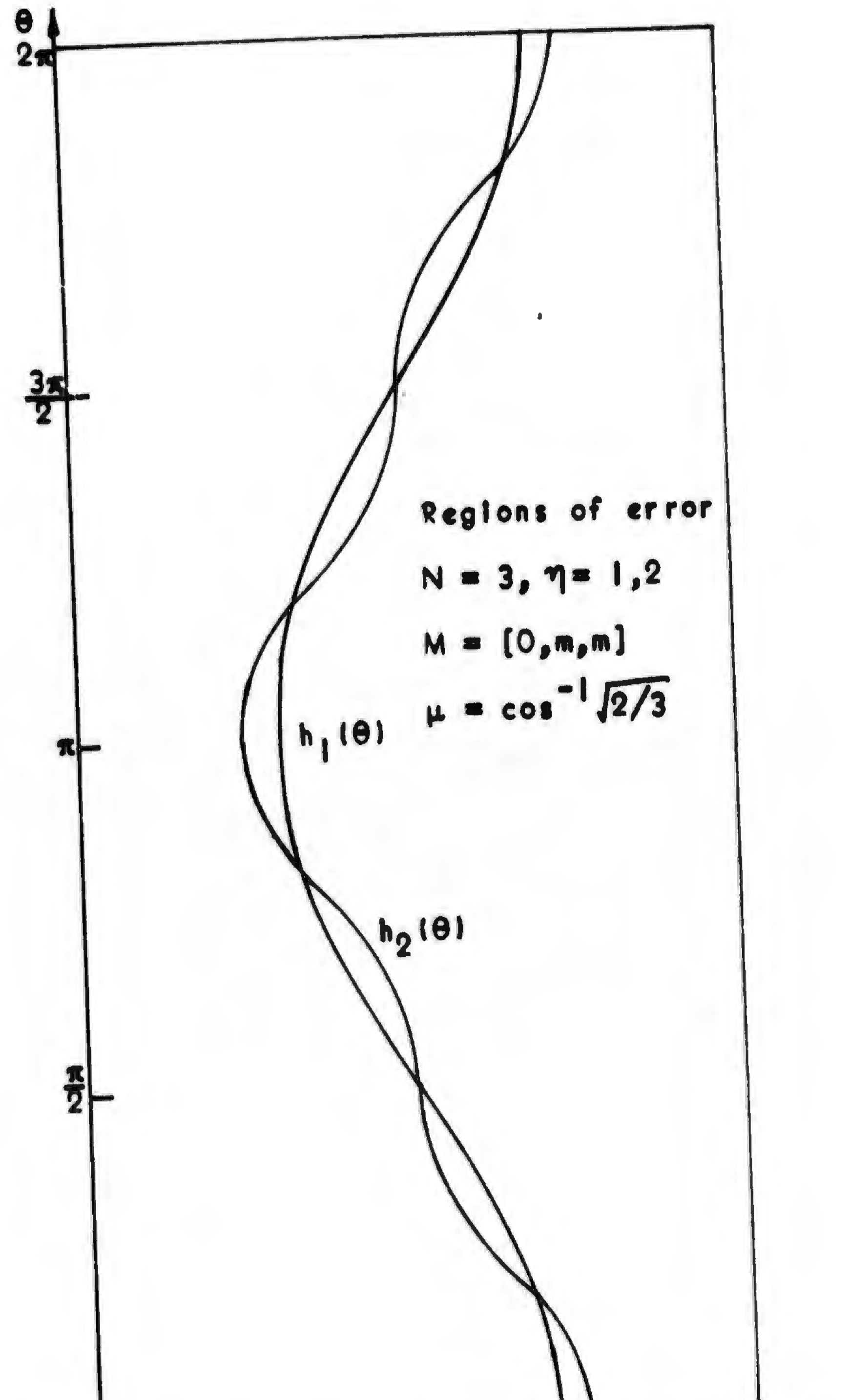

θ
2π
3π/2
π
π/2
0
π/2
π
Φ_1
Regions of error
N = 3, η = 1,2
M = [0,m,m]
μ = cos$^{-1}\sqrt{2/3}$
$h_1(\theta)$
$h_2(\theta)$

Fig (4.9)

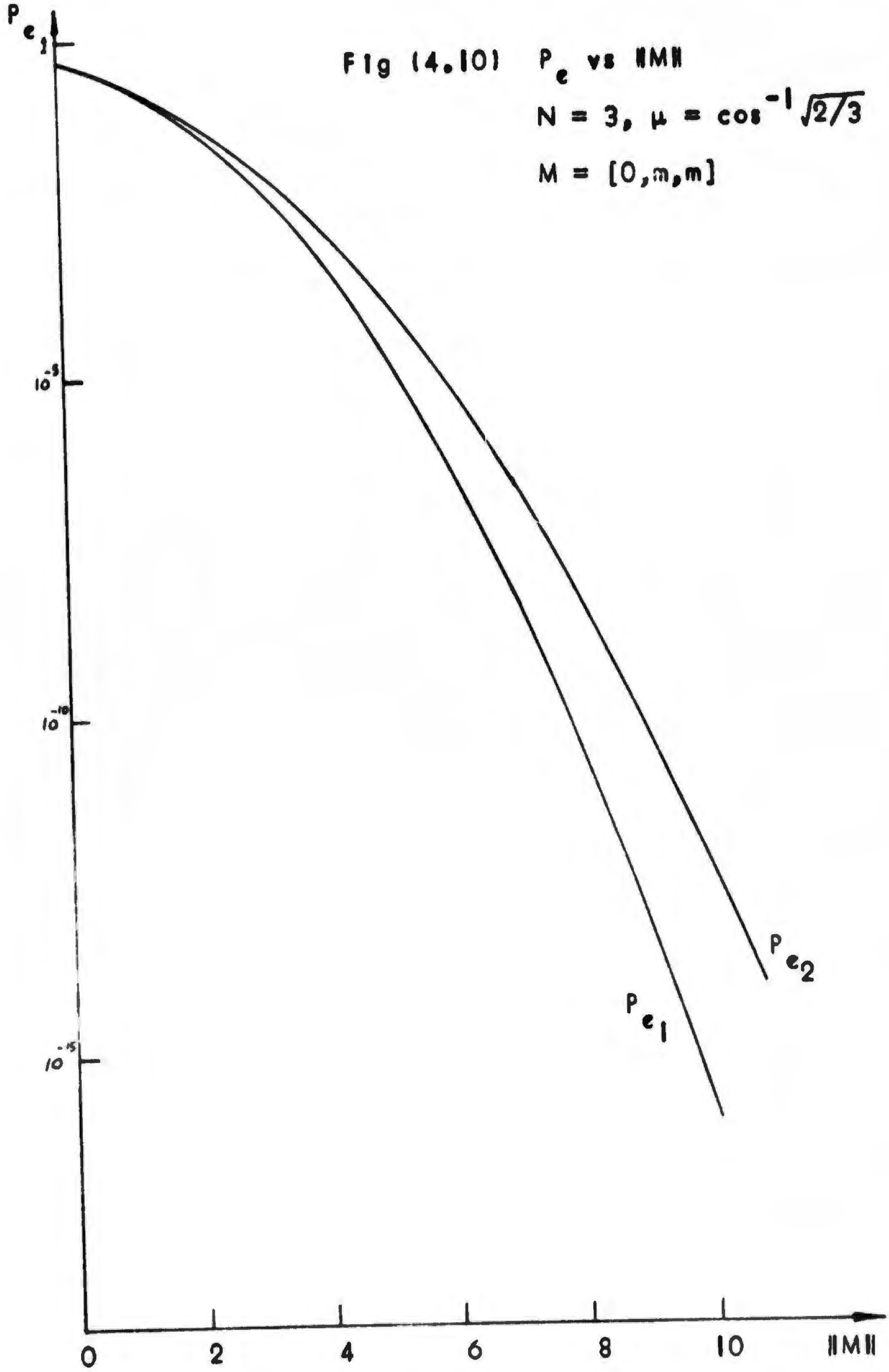


Fig (4.10) P_e vs ||M||
$N = 3,\ \mu = \cos^{-1}\sqrt{2/3}$
$M = [0, m, m]$

$$+\int_{\pi/2}^{\tan^{-1}\frac{-1}{\sqrt{2}}} p(\theta)d\theta \int_{\tan^{-1}\{2\tan\theta\sec\theta\}}^{\pi} p(\Phi_1)\, d\Phi_1$$

$$\left.+\int_{\tan^{-1}\frac{-1}{\sqrt{2}}}^{\pi} p(\theta)d\theta \int_{\tan^{-1}\sqrt{\sec 2\theta}}^{\pi} p(\Phi_1)\, d\Phi_1 \right\} \quad (4.90)$$

These expressions are graphed in fig (4.10).

<u>Case 3:</u>

$$M = [\, m_1, m_2, m_3 \,] \quad , m_1 = m_2 = m_3 = m$$

$$B = [\, 1 \,, 1, 1 \,] \quad , \mu = 0$$

that is the polar and major axes coincide. The rotation of axes transformation for this case may be taken as,

$$T = \begin{bmatrix} \sqrt{1/3} & -\sqrt{1/2} & -\sqrt{1/6} \\ \sqrt{1/3} & \sqrt{1/2} & -\sqrt{1/6} \\ \sqrt{1/3} & 0 & \sqrt{2/3} \end{bmatrix}$$

The decision surface corresponding to $\eta = 1$ is now simply,

$$h_1(\theta) = \Phi_1 = \frac{\pi}{2} \quad (4.91)$$

and that corresponding to $\eta = 2$, is readily found to be,

$$h_2(\theta) = \Phi_1 = \frac{1}{\sqrt{2}} \tan^{-1}\left\{\frac{-(3\cos\theta + \sin\theta) + \sqrt{3\cos\theta(\cos\theta + \sqrt{3}\sin\theta)}}{\sin\theta(\sin\theta - \sqrt{3}\cos\theta)}\right\}$$

$$0 < \theta < \pi/3 \quad (4.92)$$

and for the adjoining octant, it is given by,

$$h_2(\theta)=\Phi_1= \tan^{-1}\left\{\frac{2\sqrt{2}\ \sin\theta + \sqrt{12\ \sin^2\theta - 3}}{3 - 4\ \sin^2\theta}\right\}$$

$$\frac{\pi}{3} < \theta < \frac{2\pi}{3} \qquad (4.93)$$

These two descriptions suffice for this case, for by the symmetry of the variations of the decision surface we have the same functional relationships shifted in θ by $\pm\frac{2\pi}{3}$. Furthermore, we observe that (4.92) is itself (4.93) shifted in θ by $\frac{\pi}{3}$. These relationships are summed up in fig (4.11).

The average probability of error for this case may be expressed as,

$$P_{e_1} = \int_{\frac{\pi}{2}}^{\pi} p(\Phi_1)\ d\Phi_1 \qquad (4.94)$$

$$P_{e_2} = 3\left\{\int_0^{\pi/3} p(\theta)\ d\theta \int_{-\tan^{-1}\left\{\frac{2\sqrt{2}\ \sin\theta + \sqrt{12\sin^2\theta - 3}}{3 - 4\ \sin^2\theta}\right\}}^{\pi} p(\Phi_1)\ d\Phi_1 \right.$$

$$\left. + \int_{\pi/3}^{2\pi/3} p(\theta) d\theta \int_{\tan^{-1}\left\{\frac{2\sqrt{2}\ \sin\theta + \sqrt{12\sin^2\theta - 3}}{3 - 4\ \sin^2\theta}\right\}}^{\pi} p(\Phi_1)\ d\Phi_1\right\} \qquad (4.95)$$

These expressions are graphed in fig (4.12).

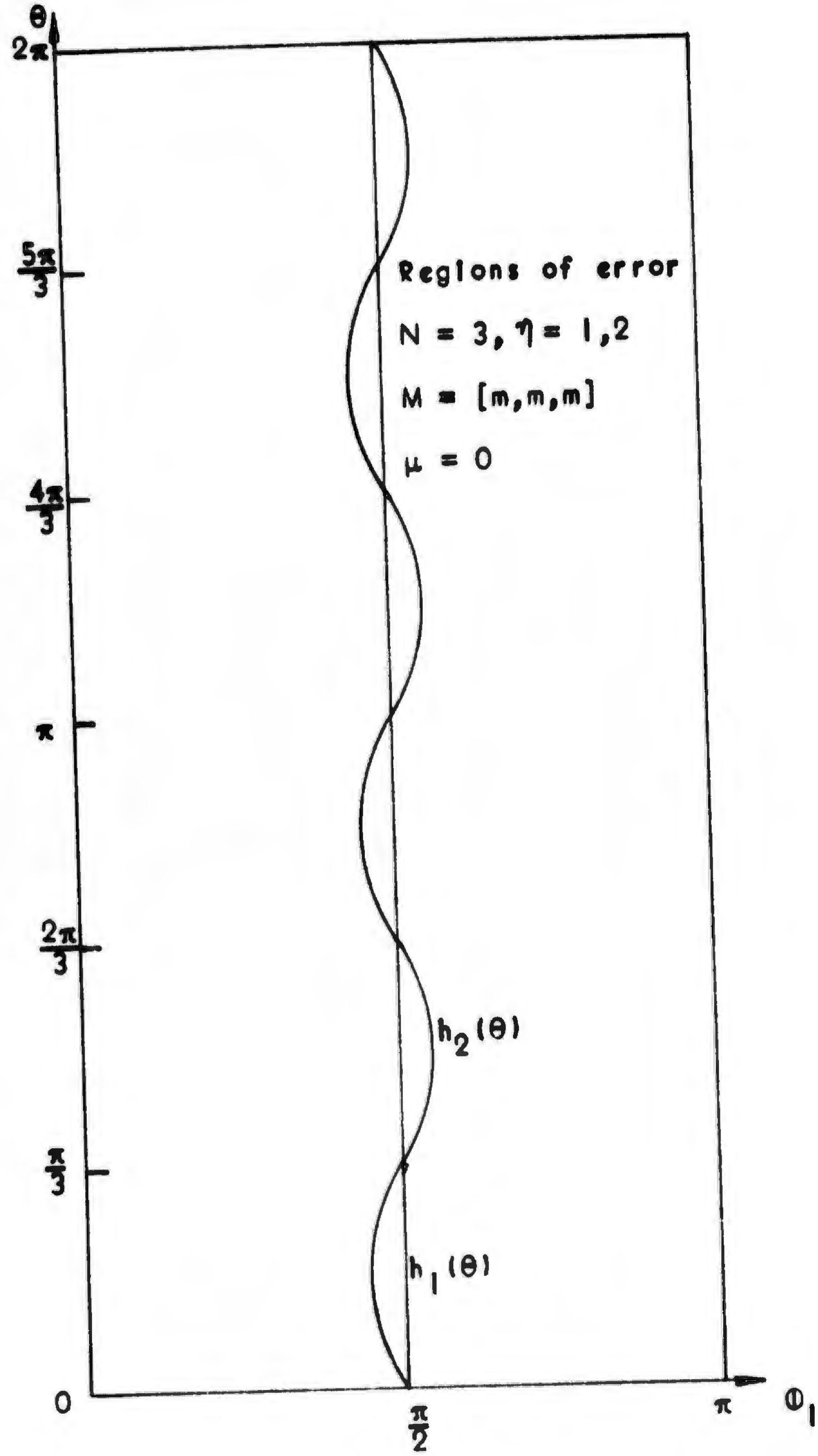

θ
2π
5π/3
4π/3
π
2π/3
π/3
0
π/2
π
Φ1
Regions of error
N = 3, η = 1,2
M = [m,m,m]
μ = 0
h2(θ)
h1(θ)

Fig (4.11)

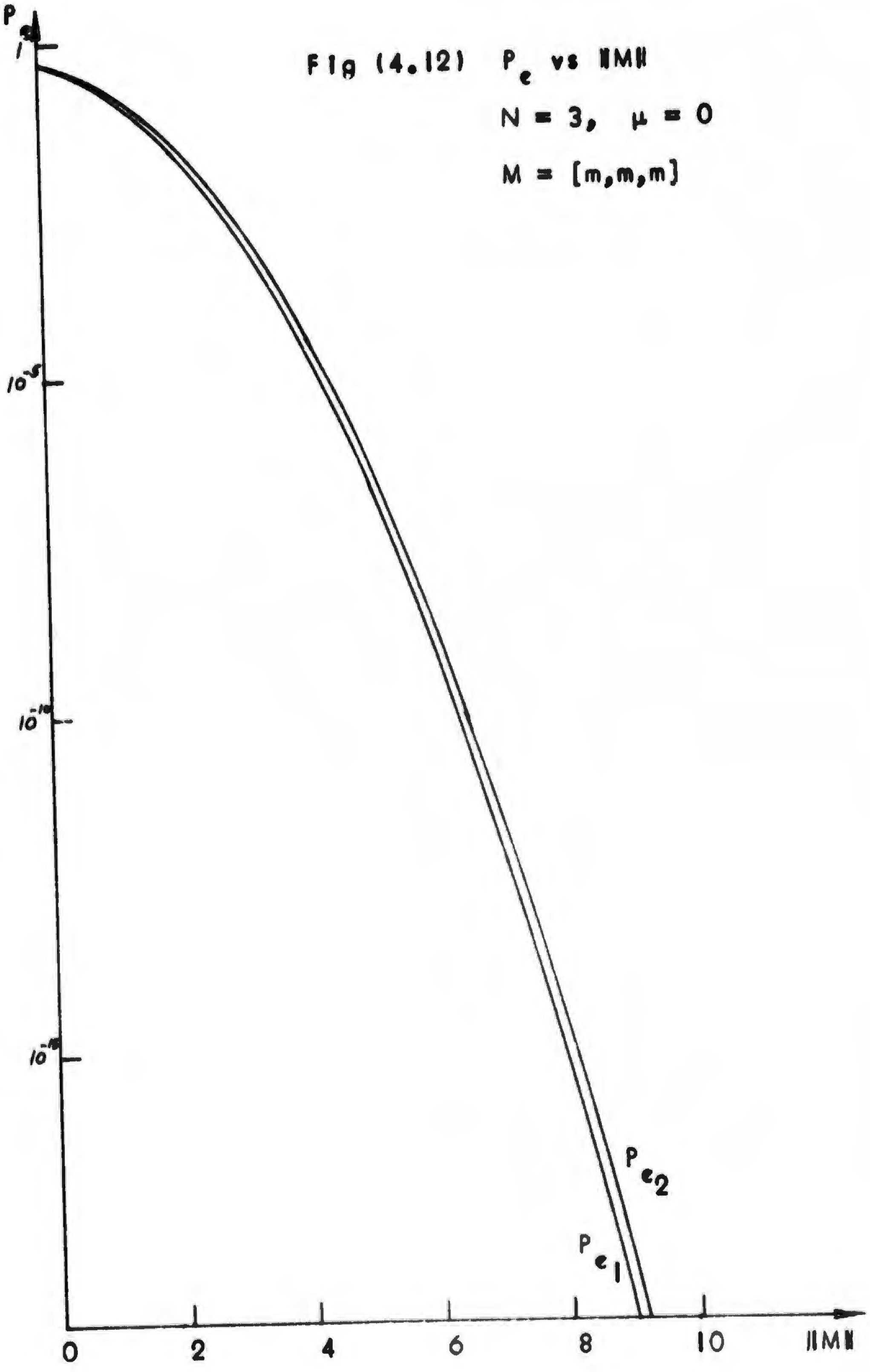


Fig (4.12) P_e vs $\|M\|$

$N = 3, \quad \mu = 0$

$M = [m, m, m]$

4.9 OBSERVATIONS

One of the major steps of this chapter is the solution of the problem of finding the probability that $\sum |x_i|^2 \text{ sgn } x_i$ be less than zero when the x_i's are independent Gaussian variables with unity variances but non-zero means. The geometrical approach we have taken is novel and may be generalized to variables with other distributions or forming other functionals involving the sgn(.) function.

We observe that the error probability under M.P. is a function not only of $\|M\|$, the total normalized energy received over all channels, but in addition is a function of how this total energy is distributed among the channels as reflected in μ. The case, when the signal strengths of all channels are equal, yields the least error probability.

As μ increases from 0, the probability of error also increases. However, there is no uniformly consistent decision rule for all combinations of the signal strengths of the channels. We note that $\eta = 1$ is preferable for certain combinations, while $\eta = 2$ is preferable for others. For the case when $\mu = 0$, we observe that there is little difference between the two decision rules Z_1 and Z_2, although Z_1 yields a slightly lower error probability than does Z_2. A bound on the difference between these probabilities is presented in appendix 'B', for the case $N = 3$, $\mu = 0$.

CHAPTER V

MAJORITY LOGIC OPERATION

Earlier in chapter III, we made note of the simpler, albeit inefficient, mode of operation with multiple channels, namely, "majority logic" operation. It was mentioned that with detecting singly at each receiver and using majority declarations for the ultimate decision, the decision rule reduced to testing whether,

$$[y_i, s_1] - [y_i, s_2] \gtrless 0 \qquad (5.1)$$

resulting in dispensing with the "estimation" phase of the system, a feature which may be desirable in itself, but as we shall presently see, will result in degradation of performance.

5.1 COMPOSITE DECISION RULE AND ERROR PROBABILITY

With a decision made at each receiver independently of the others, according to the rule (5.1), the composite decision rule for the majority logic mode may be expressed as: decide on s_1 if,

$$\sum_{i=1}^{N} \operatorname{sgn} \left\{ [y_i, s_1] - [y_i, s_2] \right\} > 0 \qquad (5.2)$$

and on s_2 otherwise.

Let e_i denote the event,

$$e_i = \frac{1}{2} \left\{ 1 + \text{sgn} \left([y_i, s_i] - [y_i, s_2] \right) \right\} \qquad (5.3)$$

then e_i is either 0 or 1; and conditioned on s_i, we have,

$$p_i = P[\, e_i = 0 \,] = \text{erfc}(\, R^- m_i) = \text{erfc}(\, \alpha_i) \qquad (5.4)$$

where R^- and m_i are given in (3.47) and (3.30) and we have defined,

$$\alpha_i = R^- m_i \qquad (5.5)$$

Define the vector,

$$\underline{e} = [e_1, e_2, \ldots , e_N]$$

where the elements e_i are as in (5.3), then, it is evident that the resulting average probability of error is given by,

$$P_e = \sum_{\{\underline{e}\}} (1 - \frac{1}{2} \delta_{d, \frac{N}{2}}) \prod_{i=1}^{\tilde{N}} w_i^{e_i} \qquad (5.6)$$

where $\delta_{i,j}$ is a modified Kronecker delta allowing for non-integer subscripts; and,

$$w_i^{e_i} = \begin{cases} p_i & \text{if } e_i = 0 \\ q_i = 1 - p_i & \text{if } e_i = 1 \end{cases} \qquad (5.7)$$

The summation is taken over all vectors [$\underline{e}$] with the property that their weights d(.) obey,

$$d(\underline{e}) = \sum_{i i i}^{\sim} e_i < \frac{N}{2} \qquad (5.8)$$

The term in front of the product in (5.6) is to reflect the fact that when N is even, and half the number of channels declare s_1, while the other half declare s_2, either signal can be present.

The number of terms in the summation will be,

$$\begin{cases} 2^{N-1} & \text{if } N \text{ is odd} \\ 2^{N-1} + \binom{N}{N/2} & \text{if } N \text{ is even} \end{cases}$$

Evidently, the average probability of error is minimized by minimizing each term in (5.6) and this is achieved by choosing $\rho = -1$, i.e. antipodal signals.

This is not the only minimizing however, for note that, whereas the average probability of error with multiple processing of antipodal signals according to the composite decision rule Z_1 was found in chapter IV to be a function of the total normalized received energy $\|M\|^2$ and the angle μ by which the individual energies $[m_i^2]$ deviate from the case of being all equal; it is clear that under majority logic operation, it is a function of the individual m_i's.

In order, therefore, to make a meaningful comparison we seek the minimum of (5.6) subject to constraints on $\|M\|$ and μ.

5.2 MAXIMA AND MINIMA OF P_e UNDER CONSTRAINTS

Let A be defined as the vector,

$$A = [\alpha_1, \alpha_2, \alpha_3, \ldots\ldots, \alpha_N]$$

where the α_i's are defined by (5.5).

THEOREM 5.1

For the probability of error given by (5.6), the minimum, under the constraint that $|A|$ be fixed, is attained when all α_i are equal to each other.

PROOF:

Consider,

$$U = P_e + \lambda \|A\| \qquad (5.9)$$

where P_e is that given by (5.6) and λ is the Lagrange multiplier.

Taking the N partial derivatives with respect to α_k, $k = 1, \ldots, N$ and setting them equal to zero, we have,

$$\frac{\partial U}{\partial \alpha_k} = \left\{ \sum_{\substack{\underline{e} \in \{\underline{e}\} \\ \ni e_k = 1}} \prod_{\substack{i=1 \\ i \neq k}}^{N} w_i^{e_i} - \sum_{\substack{\underline{e} \in \{\underline{e}\} \\ \ni e_k = 0}} \prod_{\substack{i=1 \\ i \neq k}}^{N} w_i^{e_i} \right\} \left\{ \frac{-1}{\sqrt{2\pi}} e^{-\alpha_k^2/2} \right\}$$

$$+ 2 \alpha_k \lambda = 0 \quad , k = 1, 2, \ldots, N \qquad (5.10)$$

Solving the set of equations (5.10) for λ, it readily follows that an extremum is attained when all the α_k's are equal to each other.

To show that this is a minimum, let the quantity within the first bracket in (5.10) be denoted by $C(A \sim \alpha_k)$ where $A \sim \alpha_k$ is to indicate that $C(.)$ is a function of all the elements of A except α_k, then,

$$\frac{\partial^2 P_e}{\partial \alpha_k^2} = C(A \sim \alpha_k) \frac{\alpha_k}{\sqrt{2\pi}} e^{-\alpha_k^2/2} \qquad (5.11)$$

which is clearly non-negative, since $C(.)$ is non-negative, and the result follows immediately.

THEOREM 5.2

For N odd, the extrema of the probability of error (5.6) under the two constraints that $\|A\|$ be fixed and μ be fixed, are attained when all α_i except one are equal to each other, the one being determined by the constraints.

Furthermore, a minimum is achieved when the one is smaller than the remainder, and a maximum is achieved when the one is greater than the remainder.

PROOF:

We choose to regard this as an ordinary problem of finding extrema without constraints, by making a simple transformation which incorporates the constraints into the variables.

For this purpose, consider the set of basis vectors $[\underline{\sigma}]$ in $E^{(N)}$, where,

$$\underline{\sigma}_j = [0,0,0,\ldots\ldots,0,\underset{\uparrow \atop j^{th}\ \text{place}}{1},0,\ldots,0] \ , \ j=1,2,\ldots,N$$

is the unit vector in the j^{th} direction. Then, the major axis, B, and the polar axis, A, may be written,

$$B = \sum_{j=1}^{N} \underline{\sigma}_j \tag{5.12}$$

$$A = \sum_{j=1}^{N} \alpha_j \underline{\sigma}_j \tag{5.13}$$

Consider the same space, spanned by the set of vectors $[\underline{\tau}]$ defined as,

$$\underline{\tau}_1 = B \quad ; \quad \underline{\tau}_j = \underline{\sigma}_j \quad , \ j = 2,\ 3,\ \ldots.\ ,\ N \tag{5.14}$$

hence,

$$\underline{\sigma}_1 = \underline{\tau}_1 - \sum_{j=2}^{N} \underline{\tau}_j$$

$$\|\underline{\tau}_1\|^2 = N \quad ; \quad \|\underline{\tau}_j\| = 1 \quad , \quad j = 2,\ 3,\ \ldots,\ N$$

$$[\underline{\tau}_1,\underline{\tau}_j] = 1 \quad , \ j = 2,\ 3,\ \ldots\ ,N$$

$$[\underline{\tau}_j,\underline{\tau}_k] = 0 \quad , \quad j,k \neq 1 \quad , \ j \neq k \tag{5.15}$$

Then, in terms of the new set of vectors $[\ \underline{\tau}\ ]$,

$$A = \alpha_1 \underline{I}_1 + \sum_{j=2}^{N} (\alpha_j - \alpha_1) \underline{I}_j \tag{5.16}$$

and the constraints are,

$$[A, \underline{I}_1] = \|A\| \sqrt{N} \cos \mu = \sum_{i=1}^{N} \alpha_i \tag{5.17}$$

$$\|A\|^2 = \sum_{i=1}^{N} \alpha_i^2 \tag{5.18}$$

be fixed.

Furthermore,

$$\begin{aligned} [A, \underline{g}_1] = \alpha_1 &= [A, \underline{I}_1 - \sum_{j=2}^{N} \underline{I}_j] \\ &= \|A\| \sqrt{N} \cos \mu - \sum_{j=2}^{N} \alpha_j \end{aligned} \tag{5.19}$$

and,

$$[A, \underline{g}_j] = [A, \underline{I}_j] = \alpha_j \tag{5.20}$$

In $E^{(N)}$, since there are two constraints, we are left with $(N - 2)$ degrees of freedom; we, therefore, may arbitrarily let α_j, $j=3,\ldots,N$ be the free variables, then α_1 and α_2 are determined by the constraints (5.17) and (5.18)

Setting,

$$\frac{\partial P_e}{\partial \alpha_j} = \sum_{k=1}^{N} \frac{\partial P_e}{\partial \alpha_k} \frac{\partial \alpha_k}{\partial \alpha_j} = 0 \tag{5.21}$$

$$j = 3, \ldots, N$$

and using (5.19) and (5.20), we obtain the set of $(N - 3)$ equations,

$$\left\{ \sum_{\substack{g \in \{g\} \\ \ni e_1 = 1}} \prod_{i=2}^{N} w_i^{e_i} - \sum_{\substack{g \in \{g\} \\ \ni e_1 = 0}} \prod_{i=2}^{N} w_i^{e_i} \right\} \left\{ \frac{1}{\sqrt{2\pi}} e^{-\alpha_1^2/2} \right\}$$

$$+ \left\{ \sum_{\substack{g \in \{g\} \\ \ni e_j = 1}} \prod_{\substack{i=1 \\ i \neq j}}^{N} w_i^{e_i} - \sum_{\substack{g \in \{g\} \\ \ni e_j = 0}} \prod_{\substack{i=1 \\ i \neq j}}^{N} w_i^{e_i} \right\} \left\{ \frac{-1}{\sqrt{2\pi}} e^{-\alpha_j^2/2} \right\} = 0 \quad (5.22)$$

or in short,

$$C(A \sim \alpha_1)\, e^{-\alpha_1^2/2} - C(A \sim \alpha_j)\, e^{-\alpha_j^2/2} = 0$$

$$j = 3, 4, \ldots\ldots, N \quad (5.23)$$

where C(.) is defined in theorem (5.1)

To satisfy the set of equations (5.23) simultaneously we must satisfy the following conditions,

$$C(A \sim \alpha_1) = C(A \sim \alpha_j) \quad , \; j = 3, \ldots, N \quad (5.24)$$

The only way this is possible is when,

$$\alpha_1 = \alpha_j \quad ; \; j = 3, 4, \ldots, N$$

with α_2 determined by (5.18). This proves the first part of the theorem.

Clearly, this choice is not unique, for by a cyclic relabeling of the axes, we can have 2N extrema, N of which will be minima, and N maxima.

Under these conditions, if we take,

$$\alpha_1 = \alpha_1 \quad ; \quad \alpha_j = \alpha_2 \quad , \; j = 2, 3, \ldots\ldots, N$$

and for any given $\|A\|$ and μ, it is easy to determine,

$$\alpha_2 = \frac{\|A\|}{\sqrt{N}} \left\{ \cos \mu - \frac{\sin \mu}{\sqrt{N-1}} \right\}$$

$$\alpha_1 = \|A\| \sqrt{N} \cos \mu - (N-1)\, \alpha_2 \qquad (5.25)$$

Thus, we may write the extrema of the probability of error as,

$$P_e = p_1 \sum_{j=0}^{N_d} \binom{N-1}{j} p_2^{N-1-j} q_2^j + q_1 \sum_{j=0}^{N_d - 1} \binom{N-1}{j} p_2^{N-1-j} q_2^j \qquad (5.26)$$

The second part of the theorem follows immediately by using the condition $\alpha_1 < \alpha_2$ in (5.26) for which we have a minimum, and $\alpha_1 > \alpha_2$ for which we have a maximum.

Theorem (5.2) gives rise to an interesting geometric configuration for the locus of the probability of error. The locus of all vectors A with fixed norm and fixed angle μ about the major axis is an N-dimensional hypercone. The intersection of the hypercone and the 2-dimensional plane orthogonal to the major axis is an (N-1)-dimensional hypersphere. To each point on the hypersphere will correspond a set of the N inner products of A with the axes of coordinates, and to each set will correspond a value for the probability of error which can be visualized as a distance (to some scale) from the center of the hypersphere. The locus of all these points is an (N-1)-dimensional closed surface with N projections (mountains) corresponding to the maxima, and N depressions (valleys) corresponding to the N minima.

On the otherhand, for any value of μ larger than the complement of $\cos^{-1}(\sqrt{1/N})$, we consider only those combinations of the α_i's for which A is entirely within the positive orthant, i.e. we consider only those points belonging to the intersection of the positive orthant and the hypercone of fixed μ; in which case the maxima are still given as before, however, the minima are obtained for the extremal points at the boundary of any of the surfaces bounding the positive orthant.

To illustrate the workings of theorem (5.2) and the associated geometrical configuration of the locus of probability of error, we examine the 3-dimensional case in detail:

Let ψ be the angle subtended at the origin by the major axis and any of the mutually orthogonal coordinate axes, then,

$$\psi = \cos^{-1}(\sqrt{1/3}) \qquad (5.27)$$

Let the positive octant be viewed along the major axis, then the locus of all vectors A with fixed norm and fixed

μ is a cone. The intersection of the cone and the plane surface perpendicular to the major axis is a circle with radius $\|A\| \sin\mu$, as shown in fig (5.1). To every point on the circle there correspond a set of inner products of A and the axes of coordinates. These are given by,

$$[A,x_1] = \alpha_1 = \|A\| \left\{ \cos\psi \cos\mu + \sin\psi \sin\mu \cos\theta \right\}$$
$$[A,x_2] = \alpha_2 = \|A\| \left\{ \cos\psi \cos\mu + \sin\psi \sin\mu \cos(\frac{2\pi}{3} - \theta) \right\}$$
$$[A,x_3] = \alpha_3 = \|A\| \left\{ \cos\psi \cos\mu + \sin\psi \sin\mu \cos(\frac{2\pi}{3} + \theta) \right\} \quad (5.28)$$

where we have introduced θ, the angle between the planes formed by (B and x_1) and (B and A). We have thus incorporated the constraints into the new variable θ, with respect to which we now seek the extrema of,

$$P_e = p_1p_2 + p_2p_3 + p_3p_1 - 2p_1p_2p_3 \quad (5.29)$$

where, $p_i = \mathrm{erfc}(\alpha_i)$

Thus setting,

$$\frac{\partial P_e}{\partial \theta} = \sum_{i=1}^{3} \frac{\partial P_e}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial \theta}$$

to zero, we have,

$$0 = e^{-\alpha_1^2/2} \sin\theta[\mathrm{erfc}(\alpha_2)+\mathrm{erfc}(\alpha_3)-2\mathrm{erfc}(\alpha_2)\mathrm{erfc}(\alpha_3)] \quad (5.30)$$
$$- e^{-\alpha_2^2/2} \sin(\frac{2\pi}{3} - \theta)[\mathrm{erfc}(\alpha_1)+\mathrm{erfc}(\alpha_3)-2\mathrm{erfc}(\alpha_1)\mathrm{erfc}(\alpha_3)]$$
$$+ e^{-\alpha_3^2/2} \sin(\frac{2\pi}{3} - \theta)[\mathrm{erfc}(\alpha_1)+\mathrm{erfc}(\alpha_2)-2\mathrm{erfc}(\alpha_1)\mathrm{erfc}(\alpha_2)]$$

This is satisfied for $\Theta = k\frac{\pi}{3}$, where $k=0,1,2,3,4,5$; which shows that the maxima and minima are sinusoidally periodic with period $\frac{2\pi}{3}$.

The geometrical configuration resulting from this is the 2-dimensional amoeboid pattern, orthogonal to the major axis, with 3 outward projections and 3 inward depressions corresponding to the maxima and minima. Thus if a point on the locus of fixed μ is joined to the major axis (fig (5.1)), then where the line intersects the locus of P_e is a measure (to some scale) of the probability of error obtainable for that combination $[\alpha_1, \alpha_2, \alpha_3]$ corresponding to the point chosen.

For μ larger than the complement of $\cos^{-1}(\sqrt{1/3})$, the maxima of P_e are still found as before, but the minima of P_e are obtained for those combinations of $[\alpha_1, \alpha_2, \alpha_3]$ found by considering the intersection of the cone of fixed μ and the plane surfaces formed by any two of the coordinate axes. It can easily be shown that for given $|A|$ and μ

$$\frac{\pi}{2} - \cos^{-1}(\sqrt{1/3}) < \mu < \cos^{-1}(\sqrt{1/3})$$

the corresponding α_i's are given by,

$$\alpha_i = \alpha_j = \frac{\|A\|}{2}\left[\sqrt{3}\cos\mu \mp \sqrt{2 - 3\cos^2\mu}\right]$$
$$\alpha_k = 0 \qquad i,j,k = 1,2,3 \qquad (5.31)$$

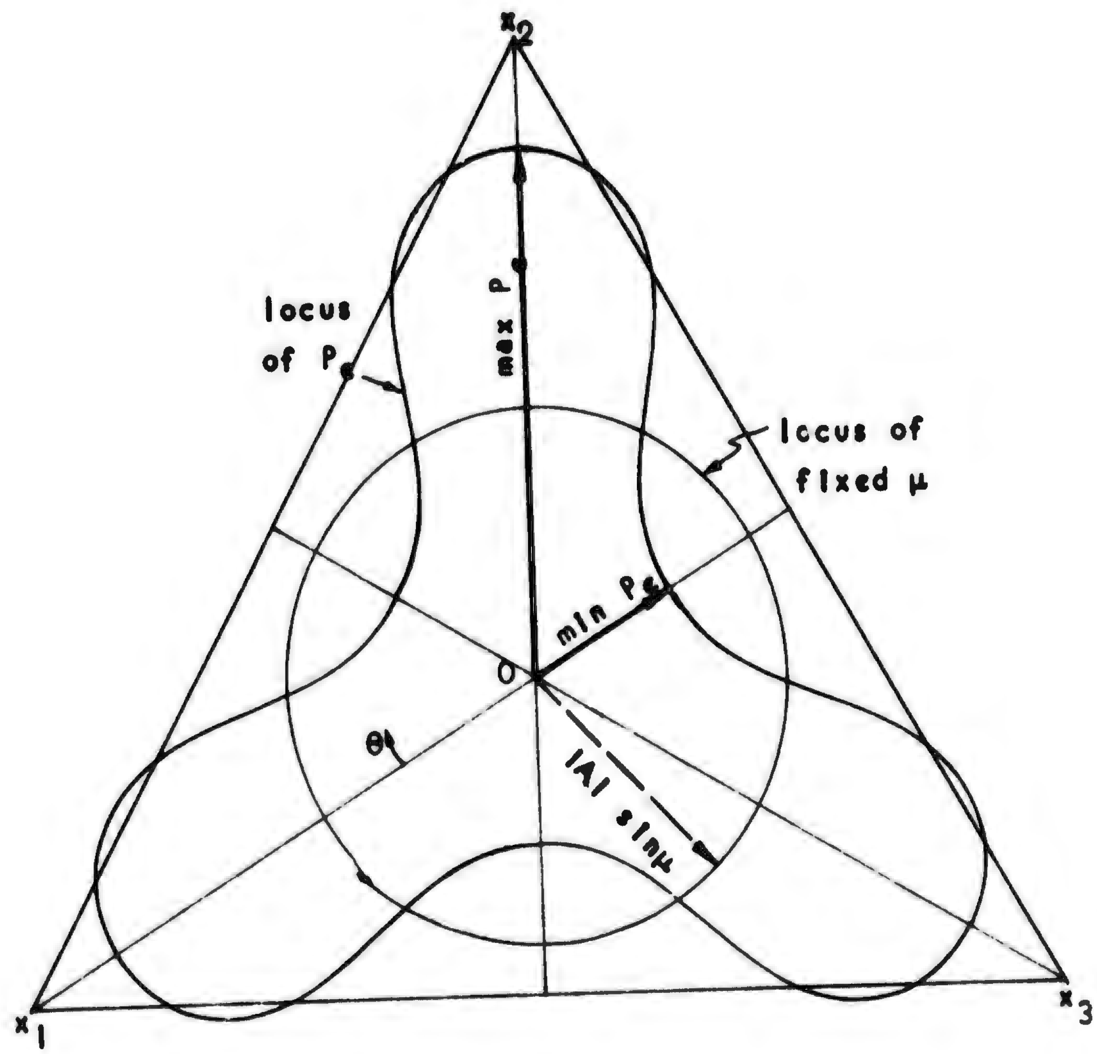


Fig (5.1) Locus of Probability of error
N = 3, Majority logic mode

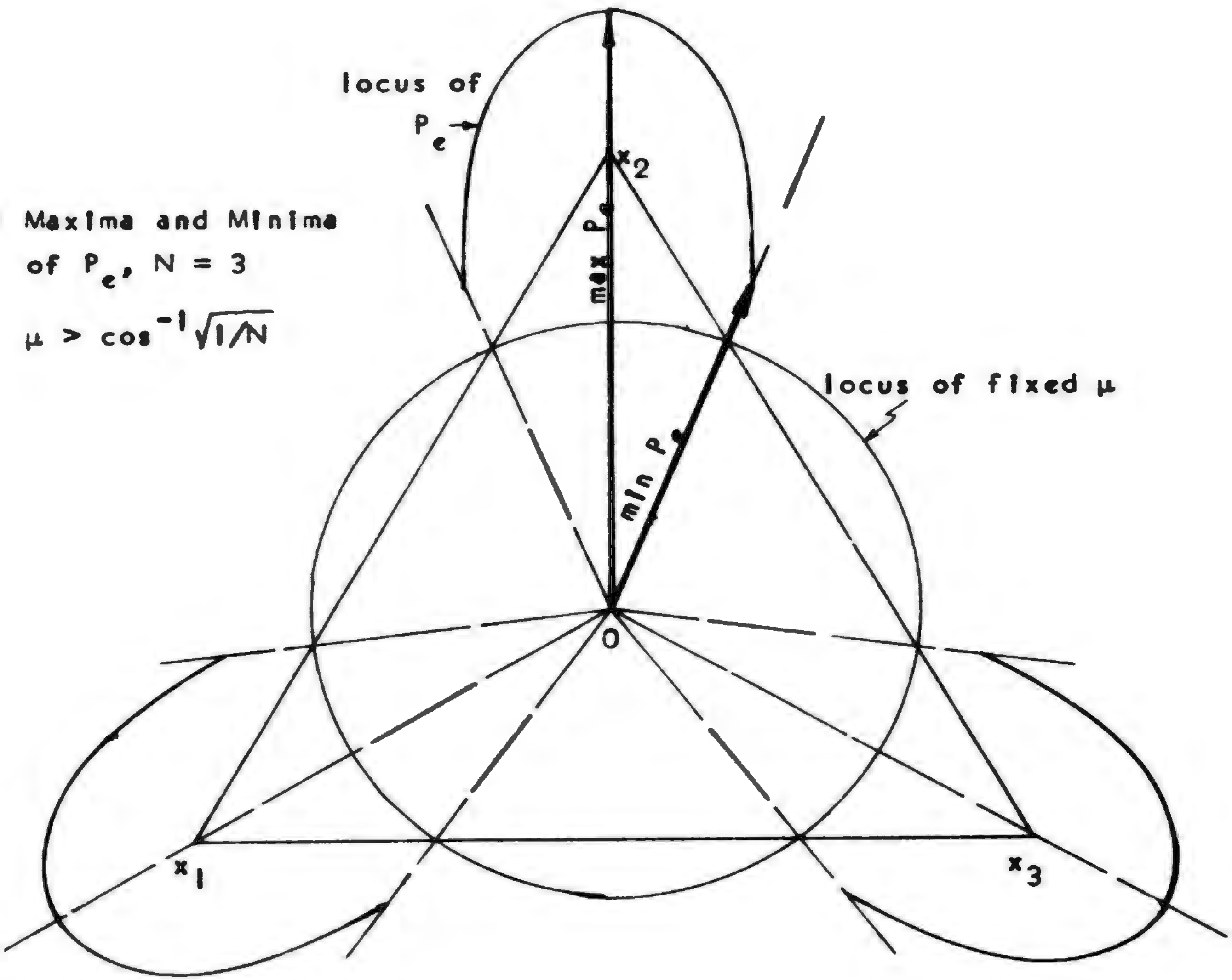


Fig (5.2) Maxima and Minima of P_e, $N = 3$

$\mu > \cos^{-1}\sqrt{1/N}$

which when used in (5.29) yield the minimum probability of error for this case.

5.3 ANTIPODAL SIGNALS

To make the required comparison with the results of the previous chapter for the multiple processing mode of antipodal signals, it is only necessary to use $[y_1, s_2] = -[y_1, s_1]$ in the composite decision rule (5.2) thus reducing it to: decide on s_1 if,

$$\text{sgn}\,[y_1, s_1] > 0 \tag{5.32}$$

and on s_2 otherwise.

The probability of error expressions and their extrema are correspondingly obtained by using $\rho = -1$ in (5.6) in which case the vector A reduces to the vector M of chapter IV.

CHAPTER VI

MISCELLANEA AND CONCLUDING REMARKS

We have, for reasons of keeping a geometrical outlook, omitted to simplify the average error probability expressions corresponding to the case $\eta = 1$. We note that the composite decision rule (4.12) for $\eta = 1$ is simply,

$$Z_1 = \sum_{i=1}^{N} x_i \tag{6.1}$$

with the x_i's as defined in chapter IV, being independent Gaussians each with mean m_i and unity variance. The resulting probability of error is readily found to be,

$$P_{e_1} = \operatorname{erfc}\left(\frac{\sum m_i}{\sqrt{N}}\right) = \operatorname{erfc}(\|M\| \cos \mu) \tag{6.2}$$

Of course, one may compare all these decision rules with the ultimately most optimal composite decision rule, when all the a_i's are known exactly, for which the rule is,

$$Z = \sum \frac{a_i}{\sqrt{ED_i}} \, x_i \gtrless 0 \tag{6.3}$$

resulting in the probability of error,

$$P_e = \operatorname{erfc}(\|M\|) \tag{6.4}$$

6.1 CONCLUDING REMARKS

It is instructive to examine a specific case, namely the two-channel operation, in order to appreciate the meaning of the superiority of multiple processing over majority logic combining. For antipodal binary signals, using majority logic, we have from (5.6)

$$P_{e_{M.L.}} = \frac{1}{2} [\, erfc(m_1) + erfc(m_2)] \qquad (6.5)$$

while for multiple processing we found in (4.18) that,

$$P_{e_{M.P.}} = erfc \left[\frac{1}{\sqrt{2}} (m_1 + m_2) \right] \qquad (6.6)$$

Fig (6.1) shows the erfc(.) sketched for non-negative arguments.

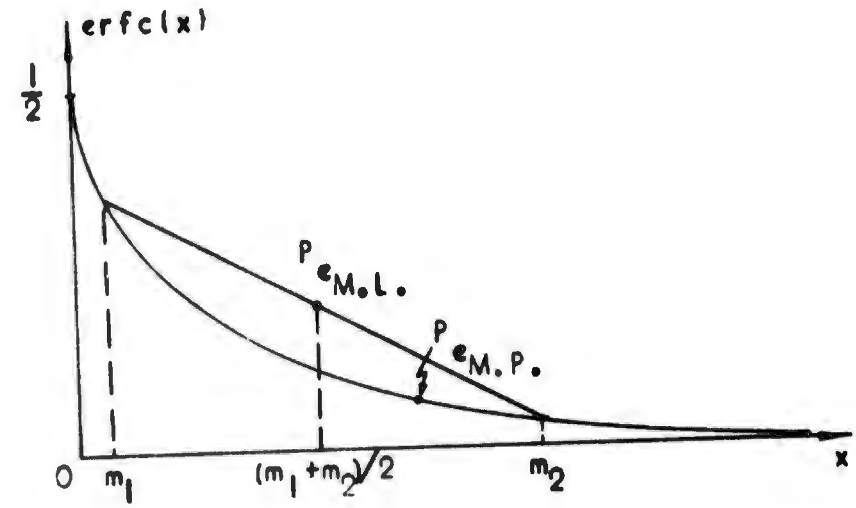


Fig (6.1)

Comparison of the two expressions above (6.5) and (6.6) shows that not only is $P_{e_{M.P.}}$ less than $P_{e_{M.L.}}$ by the convexity of erfc(.), but it is also less than the mid-point of erfc(.) corresponding to (6.5) by virtue of the monotone decreasing character of erfc(.). In short we have,

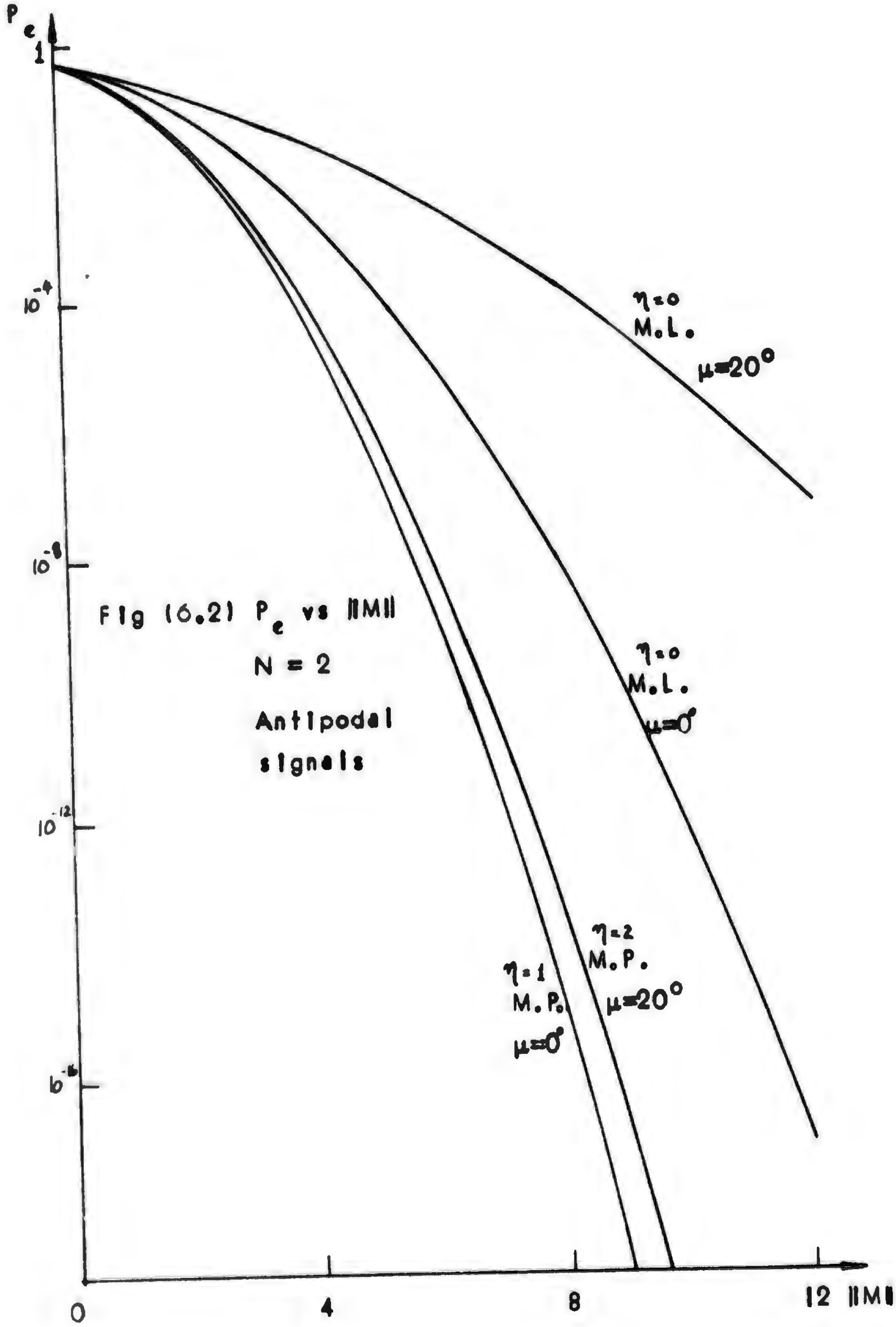
$$\mathrm{erfc}\left[\frac{1}{\sqrt{2}}(m_1 + m_2)\right] \leq \mathrm{erfc}\left[\frac{1}{2}(m_1 + m_2)\right] \leq \frac{1}{2}\left[\mathrm{erfc}(m_1) + \mathrm{erfc}(m_2)\right] \tag{6.7}$$

In figures (6.2) to (6.5) we have summarized the results of the previous two chapters. From the figures we may draw the following conclusions.

First, that under the best of conditions, when all the channel strengths are equal, multiple processing is better than majority logic combining.

Second, that as μ increases, the probability of error increases under both schemes of processing. Majority logic, and multiple processing for other than equal to 1, begin to show the variations between the best and worst cases; yet, multiple processing is better than the best that can be obtained under majority logic.

As μ increases further, the disparity between the best



P_e
1
10^{-4}
10^{-8}
10^{-12}
10^{-16}
Fig (6.2) P_e vs ||M||
N = 2
Antipodal
signals
$\eta=0$
M.L.
$\mu=20°$
$\eta=0$
M.L.
$\mu=0°$
$\eta=2$
M.P.
$\mu=20°$
$\eta=1$
M.P.
$\mu=0°$
0
4
8
12 ||M||

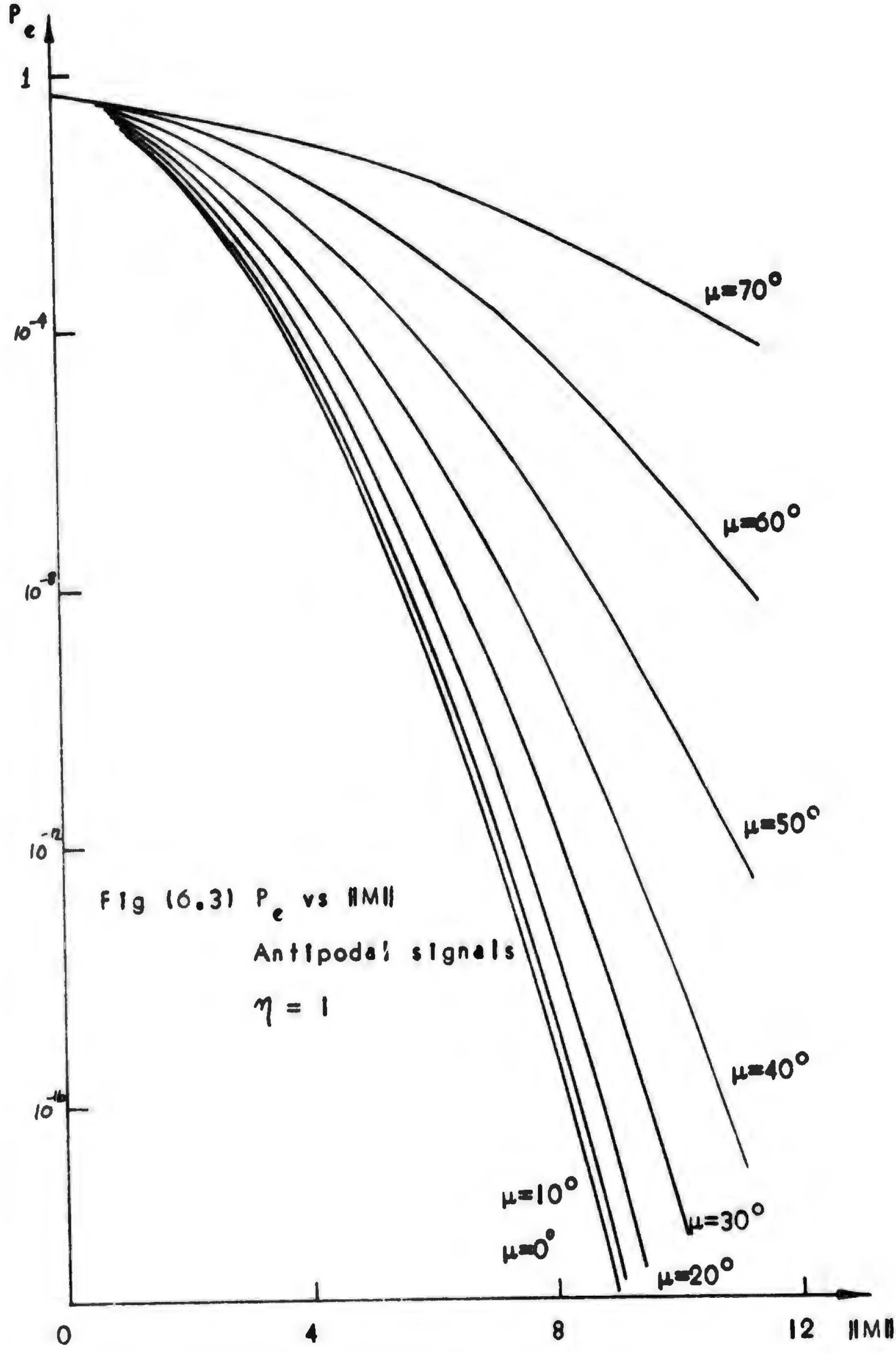


Fig (6.3) P_e vs ‖M‖
Antipodal signals
$\eta = 1$

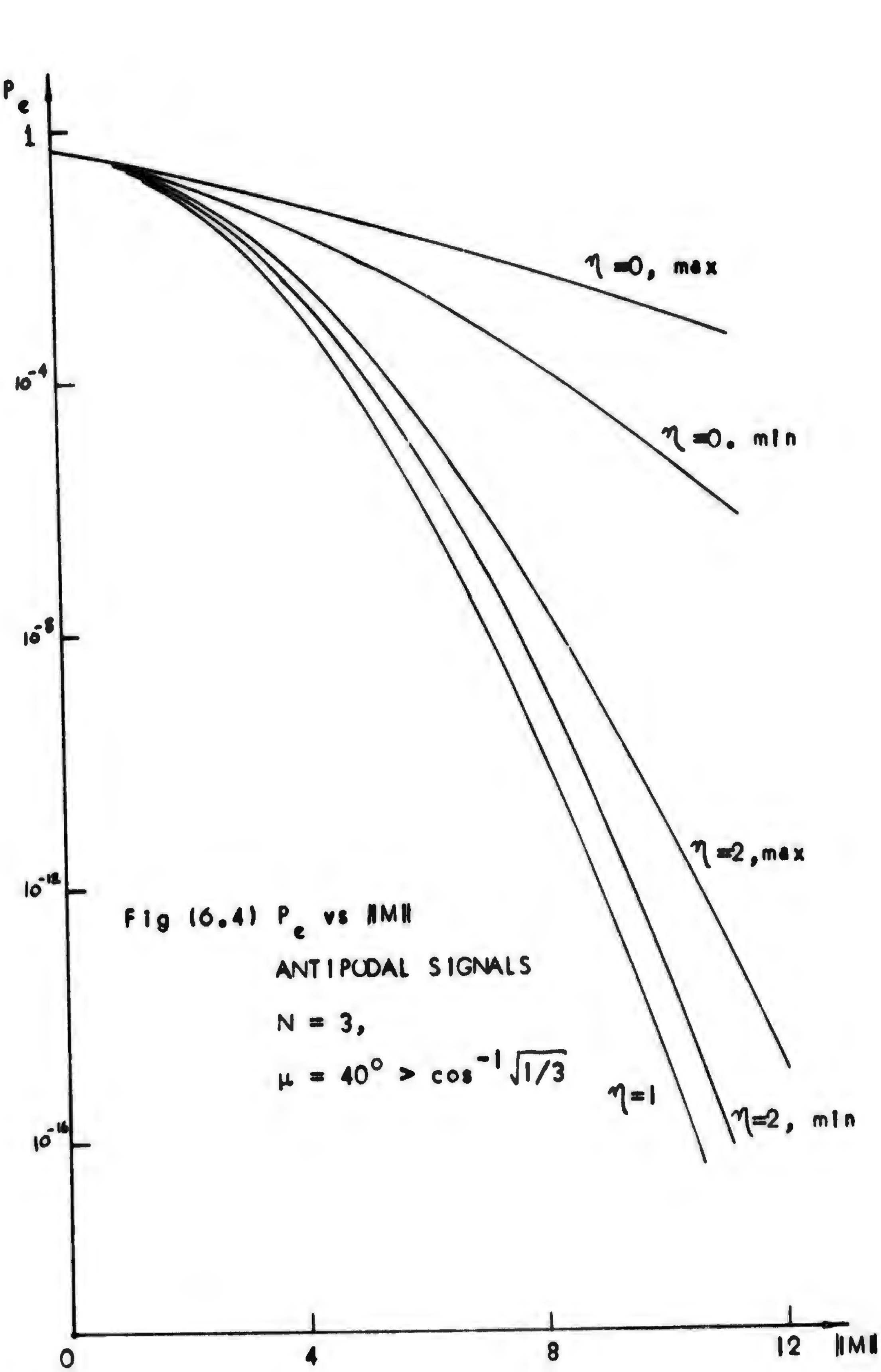


Fig (6.4) P_e vs $\|M\|$

ANTIPODAL SIGNALS

$N = 3$,

$\mu = 40^\circ > \cos^{-1}\sqrt{1/3}$

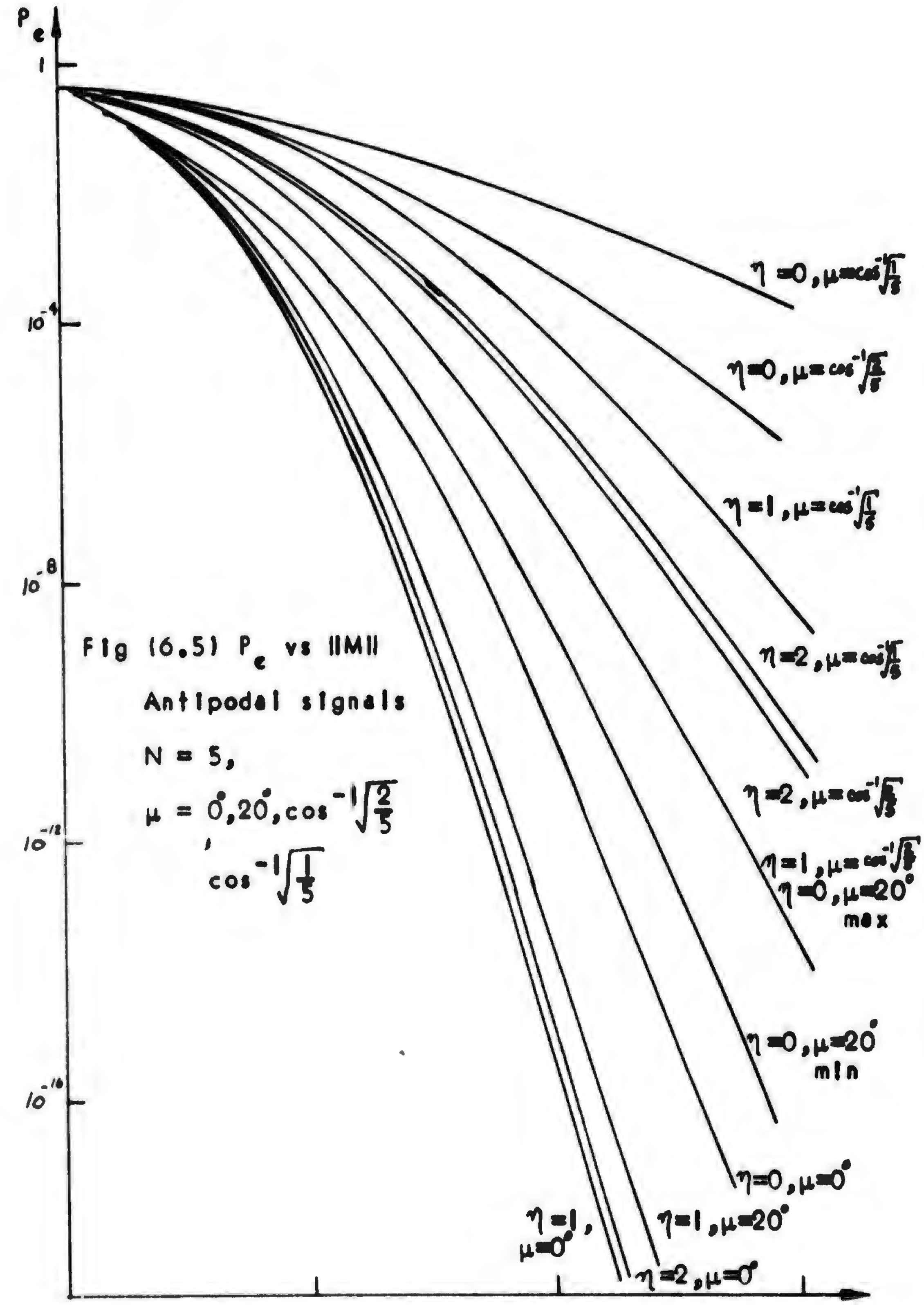


Fig (6.5) P_e vs ||M||
Antipodal signals
N = 5,
$\mu = 0°, 20°, \cos^{-1}\sqrt{\frac{2}{5}}, \cos^{-1}\sqrt{\frac{1}{5}}$

and worst cases of majority logic combining becomes more and more accentuated; and indeed, if the combination of m_i's was such that it yields the maximum (worst) probability of error, then multiple processing offers an appreciable improvement. The degree of this improvement obviously depends on the total normalized received energy.

One conclusion that has been evident throughout the discussion is that the probability of error is,generally, a function of the total normalized received energy and of a centrality factor, which is a measure of the dispersion of the signal strengths of the various channels about the case of being all equal. This is a major divergence from Price's [16] result. We reiterate that the probability of error is a function of just the total normalized energy received over all channels and of no other parameter, only under the following conditions:

(i) that the signal strengths for each channel are known exactly without any equivocation. Or

(ii) that they are unknown but they happen to be equal to each other.

With any uncertainty as to the set of m_i's, the error probability will, by force, reflect in varying degrees, the distribution of the energy among the channels.

Our final observation is that, although the question cannot be conclusively answered, as to whether to introduce adaptivity ($\eta = 2$) into a system or to leave it non-adaptive ($\eta = 1$), we do propose, however, that the outputs of the N channels should be processed multiplely in preference to detecting singly and using majority combining. The multiple processing of the N signals, according to theory is to be carried out at pre-detection level; however, suboptimal post-detection combining is still preferable to majority logic combining.

BIBLIOGRAPHY

[1] Balakrishnan, A.V., "Estimation and Detection Theory for Multiple Stochastic Processes," Journal of Mathematical Analysis and Application, 1, No. 3-4, December, 1960.

[2] Balakrishnan, A.V., "Adaptive Methods in Communication Systems," Transactions of the 3rd. Canadian IEEE Symposium on Communications, Montreal, Canada, September, 1964.

[3] Balakrishnan, A.V., "Coding Modulation and Signal Processing," Presented at the XV General Assembly of the Union Radio Scientifique International, Munich, Germany, September, 1966.

[4] Balakrishnan, A.V., Editor, "Communication Theory," McGraw-Hill, 1968, Chapter 9.

[5] Balakrishnan, A.V., "Effect of Linear and Nonlinear Signal Processing on Signal Statistics," Radio Science, 68D, No. 9, September, 1964.

[6] Fadeev, D.K., and Fadeeva, V.N., "Computational Methods of Linear Algebra," Freeman, 1963.

[7] Balakrishnan, A.V., and Abrams, I.J., "Detection Levels and Error Rates in PCM Telemetry Systems," IRE International Convention Record, Part 5, 1960, pp.37-55.

[8] Schwartz, Mischa, et al, "Communication Systems and Techniques," McGraw-Hill, 1966.

[9] Parzen, Emanuel, "Modern Probability Theory and Its Applications," Wiley, 1960.

[10] Gnedenko, B.V., "The Theory of Probability," 2nd. Edition, Chelsea, 1962.

[11] Kendall, Maurice G., and Stuart, A., "The Advanced Theory of Statistics," Vol. 1 (1958), Vol. 2 (1961), Charles Griffin and Co., London.

[12] Magnus, Wilhelm, and Oberhettinger, F., "Formulas and Theorems for the Functions of Mathematical Physics," Chelsea, 1949.

[13] Miller, Kenneth S., "Multidimensional Gaussian Distributions," Wiley, 1964.

[14] Erdelyi, A., et al, "Higher Transcendental Functions," Vol. 2, McGraw-Hill, 1953.

[15] Kireyeva, I.Ye., and Karpov, K.A., "Tables of Weber Functions," Vol. 1, Pergamon, 1961.

[16] Price, R., "Error Probabilities for Adaptive Multichannel Reception of Binary Signals," M.I.T. Lincoln Lab., Lexington, Mass. Tech. Rept. 258, July, 1962.

[17] Middleton, D., "The Incoherent Estimation of Signal Amplitudes in Normal Noise Background," SIAM Series in Applied Mathematics, M. Rosenblatt, Editor, Wiley, 1963.

[18] Middleton, D., and Esposito, R., "Simultaneous Optimum Detection and Estimation of Signals in Noise," IEEE Trans. on Information Theory,

IT-14, No. 3, May 1968.

[19] Brennan, D.G., "Linear Diversity Combining Techniques," Proc. IRE, June 1959.

[20] Kendall, Maurice G., and Moran, P.A.P., "Geometrical Probability," Hafner Pub. Co., 1963.

[21] Sommerville, D.M.Y., "Introduction to the Geometry of N Dimensions," Dutton and Co., 1929.

[22] Davisson, Lee D., "The Theory of Data Compression," Ph.D. in Engineering, University of California at Los Angeles, June, 1964.

[23] Davisson, L.D., "A Theory of Adaptive Filtering," IEEE Trans. on Information Theory, IT-12, No.2, April, 1968.

[24] Davisson, L.D., "An Approximate Theory of Prediction for Data Compression," IEEE Trans. on Information Theory, IT-13, April, 1967.

[25] Viterbi, Andrew J., "Principles of Coherent Communication," McGraw-Hill, 1966.

[26] Hingorani, G.D., "On Performance of a Class of Binary Receivers," Applied Research Lab., Sylvania Electronic Systems, Waltham, Mass., Research Note 356, August, 1965.

[27] Hingorani, G.D., "Error Rates for a Class of Binary Receivers," IEEE Trans. on Communication Technology, COM-15, April, 1967.

[28] Bello, P.A., "Binary Error Probabilities over Selectively Fading Channels containing Specular Components," IEEE Trans. on Communication Technology, COM-14, August, 1966.

[29] Proakis, J.G., "On the Probability of Error for Multichannel Reception of Binary Signals," IEEE Trans. on Communication Technology, COM-16, February, 1968.

[30] Proakis, J.G., "Probabilities of Error for Adaptive Reception of M-Phase Signals," IEEE Trans. on Communication Technology, COM-16, February, 1968.

[31] Proakis, J.G., Drouilhet, P.R., and Price, R. "Performance of Coherent Detection Systems using Decision-Directed Channel Measurement," IEEE Trans. on Communication Systems, CS-12, March, 1964.

[32] Helstrom, Carl W., "Statistical Theory of Signal Detection," Pergamon, London, 1960.

APPENDIX 'A'

GENERALIZED SPHERICAL TRANSFORMATION

Among the many generalized spherical transformations in N-dimensional Euclidean space $E^{(N)}$, we have occasion to use only one, which we present briefly below, omitting the details which can be found in [13].

Let $[e_1, e_2, \ldots\ldots, e_N]$ be a complete orthonormal set of vectors for $E^{(N)}$, and let X be a vector of norm 'r', with components $x_1, x_2, \ldots\ldots, x_N$ with respect to this basis; then it is possible to choose (N − 1) angles $\Phi_1, \Phi_2, \ldots\ldots, \Phi_{N-2}, \theta$ which are independent of each other, and which, when combined with the norm 'r' completely describe the vector X with respect to the given orthonormal basis. Thus let,

$$
\begin{aligned}
x_j &= r\left[\prod_{k=1}^{j-1}\sin\Phi_k\right]\cos\Phi_j , && j=1,2,\ldots,N-2 \\
x_{N-1} &= r\left[\prod_{k=1}^{N-2}\sin\Phi_k\right]\cos\theta , && k=1,2,\ldots,N-2 \\
x_N &= r\left[\prod_{k=1}^{N-2}\sin\Phi_k\right]\sin\theta , && 0 \leq \Phi_k \leq \pi \\
r &\geq 0 , && 0 \leq \theta \leq 2\pi
\end{aligned}
\qquad (A.1)
$$

in which, the vacuous product is to be interpreted as 1.

The Jacobian J of this transformation is,

$$J_N(r,\theta,\Phi_1,\Phi_2,\ldots,\Phi_{N-2}) = r^{N-1} \prod_{k=1}^{N-2} \sin^{N-1-k}\Phi_k \quad \text{(A.2)}$$

In chapter IV we require the product,

$$\prod_{k=1}^{N-2} \int_0^{\pi} \sin^{N-1-k}\Phi_k \, d\Phi_k$$

which may be shown [13] to be equal to,

$$\frac{\Gamma^{N-3}\left(\frac{1}{2}\right)}{\Gamma\left(\frac{N-1}{2}\right)}$$

APPENDIX 'B'

SIMPLE BOUND ON A PROBABILITY DIFFERENCE

In section (5.8) we examined the case, for $N = 3$, when the polar and major axes were coincident. We found that the probability of error for $\eta = 2$ was higher than that for $\eta = 1$. We present, here, a simple bound for the difference in these probabilities.

With reference to fig (B.1), the difference in probability δ may be written as,

$$\delta = 3 \, (P_I - P_{II}) \qquad (B.1)$$

where P_I and P_{II} are the probability contents with respect to the density $p(\theta, \Phi_1)$ over the regions indicated.

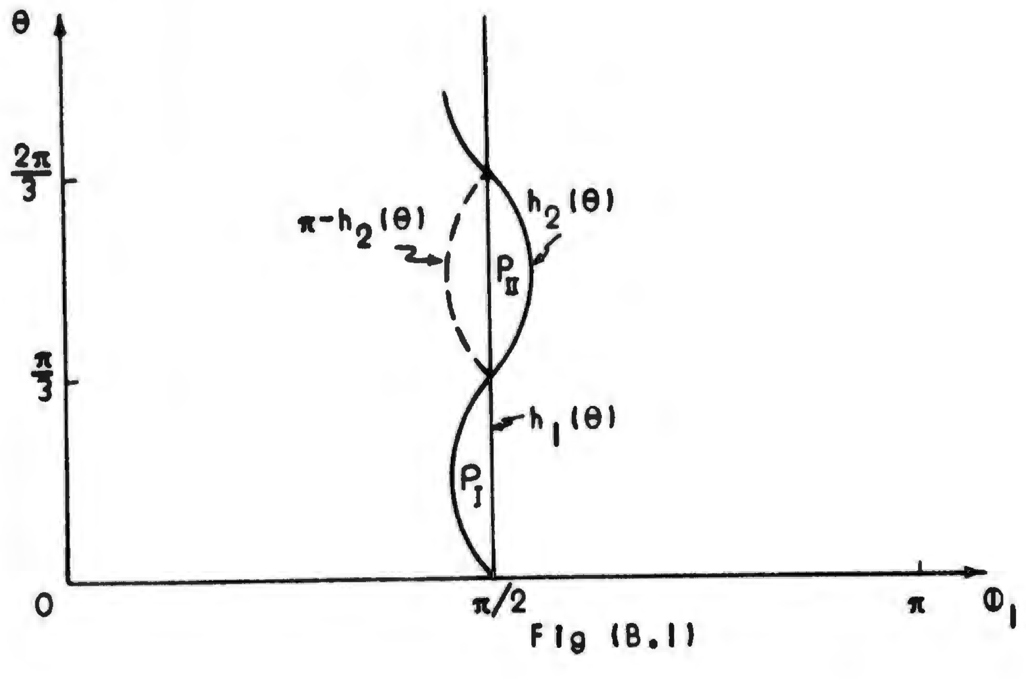


Fig (B.1)

$$P_I = \int_0^{\pi/3} d\theta \int_{h_2(\theta)}^{\pi/2} p(\theta,\Phi_1)\, d\Phi_1$$

$$= \int_{\pi/3}^{2\pi/3} d\theta \int_{\pi-h_2(\theta)}^{\pi/2} p(\theta,\Phi_1)\, d\Phi_1 \qquad (B.2)$$

$$P_{II} = \int_{\pi/3}^{2\pi/3} d\theta \int_{\pi/2}^{h_2(\theta)} p(\theta,\Phi_1)\, d\Phi_1 \qquad (B.3)$$

where $h_2(\theta)$ is the functional relationship of the decision surface relating Φ_1 to θ, for $\eta = 2$.

Using the rotation of axes transformation given in section (4.8) - Case 3; and following it by a spherical coordinate transformation, we arrive at equations (4.92) and (4.93) as descriptions for $h_2(\theta)$

Now it is clear that the difference in probability content $(P_I - P_{II})$ is a function of the uniformity or lack of it (i.e. a function of the slope) of $p(\Phi_1)$ in the region about $\Phi_1 = \pi/2$. We therefore, approximate $p(\Phi_1)$ in this neighborhood by a first order Taylor series expansion,

$$p(\Phi_1) \approx p(\tfrac{\pi}{2}) + (\Phi_1 - \tfrac{\pi}{2})\, p'(\tfrac{\pi}{2})$$

or more simply, using p(t) of (4.37)

$$p(t) \approx p(0) + t\, p'(0) \qquad (B.4)$$

Now, using the identity [12],

$$D_n(0) = \frac{\Gamma(\frac{1}{2})\ 2^{n/2}}{\Gamma(\frac{1-n}{2})}$$

and the identity [14],

$$\frac{d^m}{dz^m}[\exp(z^2/4)\ D_n(z)] = (-1)^m(-n)_m \exp(z^2/4) D_{n-m}(z)$$

where,

$$(a)_m = \frac{\Gamma(a+m)}{\Gamma(a)}$$

It can be shown that,

$$p(t) \approx \exp(-\|M\|^2/2)\ [\ \frac{1}{2} - \sqrt{\frac{2}{\pi}}\ \|M\|\ t] \qquad (B.5)$$

and in terms of t, $h_2(\theta)$ now relates θ to the new variable 't', and is given by,

$$t = \left\{ 1 + \frac{(2\sqrt{2}\sin\theta + \sqrt{12\sin^2\theta - 3})^2}{(3 - 4\sin^2\theta)^2} \right\}^{-\frac{1}{2}} \qquad (B.6)$$

Combining these results, we have,

$$P_I - P_{II} \approx \|M\| \exp(-\frac{\|M\|^2}{2}) \sqrt{\frac{2}{\pi}} \int_{\pi/3}^{2\pi/3} f(\theta)\ d\theta \qquad (B.7)$$

where $f(\theta)$ is the square of the quantity in (B.6)

Hence,

$$\delta = K \|M\| \exp(- \|M\|^2/2) \tag{B.8}$$

where,

$$K = 3\sqrt{\frac{2}{\pi}} \int_{\pi/3}^{2\pi/3} f(\theta)\, d\theta \tag{B.9}$$

It is easily shown that,

$$K < 0.07$$

Thus the behavior of δ as a function of $|M|$ is Rayleigh-like, and as $\|M\|$ increases, the difference vanishes.

APPENDIX 'C'

LISTING OF POLYNOMIALS USED IN THE EVALUATION OF $I_{N,1}(a,b)$

$$f_0(z) = 1$$

$$f_1(z) = z$$

$$f_2(z) = 1 + z^2$$

$$f_3(z) = 3z + z^3$$

$$f_4(z) = 3 + 6z^2 + z^4$$

$$f_5(z) = 15z + 10z^3 + z^5$$

$$f_6(z) = 15 + 45z^2 + 15z^4 + z^6$$

$$f_7(z) = 105z + 105z^3 + 21z^5 + z^7$$

$$f_8(z) = 105 + 420z^2 + 210z^4 + 28z^6 + z^8$$

$$g_0(z) = 0$$

$$g_1(z) = 1$$

$$g_2(z) = z$$

$$g_3(z) = 2 + z^2$$

$$g_4(z) = 5z + z^3$$

$$g_5(z) = 8 + 9z^2 + z^4$$

$$g_6(z) = 33z + 14z^3 + z^5$$

$$g_7(z) = 48 + 87z^2 + 20z^4 + z^6$$

$$g_8(z) = 279z + 185z^3 + 27z^5 + z^7$$