UNCLASSIFIED

AD NUMBER

AD883415

LIMITATION CHANGES

TO:

Approved for public release; distribution is unlimited.

FROM:

Distribution authorized to U.S. Gov't. agencies only; Test and Evaluation; 11 MAY 1971. Other requests shall be referred to Office Naval Research, 875 North Randolph Street, John J. Kingman Road, Arlington, VA 22203-1995.

AUTHORITY

ONR ltr, 29 Aug 1973

THIS PAGE IS UNCLASSIFIED

AD883415

(12)

PROCESSING OF DATA FROM SONAR SYSTEMS

Volume VIII

Office of Naval Research (Code 466)
ARLINGTON VA 22217

GENERAL DYNAMICS
Electric Boat Division

Best
Available
Copy

V

GENERAL DYNAMICS
Electric Boat Division

12

Reproduction in whole
or in part is permitted for
any purpose of the United
States Government.

Distribution limited to U.S. Gov't. agencies only;
Test and Evaluation; 11 MAY 1971. Other requests
for this document must be referred to
Office of Naval Research (Code 466).
ARLINGTON, VA. 22217

PROCESSING OF DATA FROM SONAR SYSTEMS

Volume VIII

This research was sponsored by the Office of Naval Research under Contract N00014-68-C-0392 (ONR Contract Authority Identification Number NR-286-001-1) with General Dynamics Corporation, Electric Boat division. The work was performed under subcontract by Yale University.

Examined: _____
J. W. Herring
Project Engineer

Approved: _____
Dr. A. J. van Woerkom
Manager of Scientific Research

DDC
RECEIVED
MAY 11 1971
D

U417-70-065
October 15, 1970

496

ABSTRACT

Volume VIII deals with the following topics:

1. Application of Distribution - Free Tolerance Regions to Pattern Recognition

Pattern recognition is needed to identify sonar signatures as to the type of target by which they are generated. Distribution - free methods are desirable in this context since the probability distributions underlying the pattern classes are frequently unknown, and it is desirable to establish some upper bound on at least the expected false-alarm probability. The recognition method developed has actually been applied to the recognition of speech wave forms since these are more easily obtainable than sonar signatures yet possess some of the same characteristics.

2. Passive Detection and Tracking using Surface Scattered Signals

Signals reflected from irregular time varying boundaries such as the sea surface undergo distortion which limits their detectability and useability for tracking. The properties of this distortion for correlator processing are related to the statistical constraints placed upon the time variation and irregularity of the boundary. Two propagation geometries are analysed. The first deals with the crosscorrelation of surface reflected and direct transmission paths, and the second with the crosscorrelation of surface scattered signals received at two different locations. This analysis assumes that the signal generated at the target and the background noise are both gaussian random variables. Three models of the scattering mechanism are proposed and two are analysed in detail. In all cases the correlator output is shown to exhibit very persistent fluctuations

due to the scattering. The existence of these fluctuations is related to the non-gaussian nature of the scattered signals. The fourth order cumulant is computed to show that well spaced scattered signal samples may be dependent even when they are uncorrelated. Results are presented for low pass signal spectra and are investigated as a function of bandwidth.

CONTENTS

Report	Title	Page
	Abstract	iii
	Foreword	
I	Introduction	1
II	Application of Distribution Free Tolerance Regions to Pattern Recognition	2
III	Passive Detection and Tracking Using Surface Scattered Signals	6
42	Application of Distribution-Free Tolerance Regions to Pattern Recognition	Appendix A
43	Passive Detection and Tracking Using Surface Scattered Signals	Appendix B

FOREWORD

This is the eighth in a series of reports describing work performed by Yale University under a subcontract with Electric Boat division of General Dynamics, prime contractor. The Office of Naval Research is sponsor of this research under Contract N00014-68-C-0392 (Contract Authority Identification Number NR 286-001-1). LCDR J. F. Lyding is Project Officer for ONR. Mr. J. W. Herring is Project Engineer for Electric Boat division under the direction of Dr. A. J. van Woerkom, Manager of Scientific Research.

1. Introduction

This report is the second of two volumes dealing with work completed under contract 8050-31-55001 between Yale University and the Electric Boat Company during the period from July 1, 1968 to April 30, 1970. More detailed discussions of the results are contained in the two progress reports Nos. 42 and 43 which are appended. The companion volume (Vol. VII of this series) covers work done during the same time period and contains results submitted originally in progress reports No. 38 through 41. The present volume is concerned with pattern recognition and detection of surface scattered signals, and it therefore represents something of a departure from previous work, where the emphasis was mainly a signal processing.

The interest in pattern recognition arose initially from the desire of identifying target types from their sonar signatures, i.e. to determine whether a received signal was generated by a ship, or a submarine, or possibly a school of fish. Pattern recognition is however still a rather inexact discipline relying rather heavily on ad hoc procedures. Hence the approach taken in a given case depends very strongly on the nature of the application, and in order to apply pattern recognition techniques to sonar signature discrimination it would have been necessary to have had on hand representative samples of sonar signals. These proved to be not easily available, and it was decided therefore to transform the recognition problem into a speech recognition problem, on the supposition that speech waveforms would be roughly equivalent to sonar waveforms. This kind of equivalence would of course exist only for signals from single hydrophones and information contained in the spatial distribution of sonar signals from different types of targets is therefore discarded. An initial attempt at signature discrimination would however probably not have included

spatial properties in any case, since these call for entirely different approaches some of which are being considered in current research.

Surface-scattered signals must be considered in sonar detection and communication systems because in many cases signals will be transmitted to the receiving array not only by the direct path but also by reflection from the surface (and the bottom). In fact under certain shadowing conditions the surface-reflected path may be the only one transmitting significant energy. In past work the characteristics of the propagation path have been largely ignored; i.e. only the most elementary propagation models were used. While many important and valid results were obtained this way it has always been clear that many other effects observed in sonar systems could only be analyzed by considering more sophisticated models. The only previous effort in this direction is contained in Progress report No. 13 (Appendix F of Volume II) where the effect of volume inhomogeneities in producing errors in the bearing estimate were considered. The surface scattering path studied in the present volume represents another effort at a more realistic characterization of the propagation path.

II. Application of Distribution-Free Tolerance Regions to Pattern Recognition

As noted in the introduction, the sonar signature classification problem has been converted into a speaker recognition problem. The formulation of the problem dealt with in progress report No. 42 is that the system is to recognize a main speaker with a fixed expected false alarm probability. Any test speaker who is not the main speaker is considered to be an impostor and a false alarm is defined as the error committed when an impostor is classified as the main speaker. In addition

to fixing the probability of false alarm the system should correctly recognize the main speaker as often as possible.

It is assumed that the probability distributions governing the class distributions are unknown. Hence with a finite sample size it is impossible to make precise statements about any of the error probabilities. The method of Distribution-Free Tolerance Regions makes it possible to fix the expected probability of one kind of error, here chosen to be the false alarm. It does not guarantee anything about the errors of the other kind - different choices of tolerance regions resulting in different error rates in any given sample. Thus the best that can be done is to select a method that appears to have desirable properties and can therefore be expected to do a good job of maximizing the probability of correct classification.

It is assumed that much of the information for the recognition of speakers is contained in the transition between phonemes as well as in the phonemes themselves. For this reason a simple word which contained a dipthong was analyzed by calculating many short-term spectra over the length of the word. These spectra were used to form the measurement space in which the decision regions were constructed.

Recordings of 225 utterances by each of three speakers, 25 utterances by each of 26 impostors, and 10 utterances by each of 30 impostors were used. These utterances consisted of the sentence "My code is - ". digitized into 10 bit accuracy at 8000 samples per second. Only the word "my" was actually analyzed.

A short-term spectrum was calculated from 256 samples of the waveform by a fast Fourier transform. Hence, each spectrum consisted of 128 unique frequency components. The questions of 1) how many spectra to use and 2) how coarse each spectrum should be were investigated by forming

3-different 256-dimensional measurement spaces. The first measurement space was made up of 4 spectra with each spectrum having 64 frequency components. The second space consisted of 8 spectra with each spectrum having 32 components. The third space consisted of 16 spectra with each spectrum having 16 components. The number of frequency components per spectrum was reduced from 128 to 64, 32, and 16 by simple averaging. The amplitude of each spectrum was normalized to make the energy content of the word constant. The length of the word "my" varied from 900 time samples (approximately 110 mscc) to 3350 time samples (approximately 420 ms) according to the particular speaker and the particular utterance involved. Typical variation of the length of "my" by the same speaker was from 1500 to 2200 time samples. This variation was normalized by placing the spectra uniformly across the word "my". Therefore, in the case of the typical speaker eight spectra would approximately cover the word with little if any overlap.

In the method of Distribution Free Tolerance Regions (DFTR) the sample space is separated into statistically equivalent blocks by means of a set of ordering functions. The general procedure is described in chapter 2 of progress report No. 42, and a more detailed description of the ordering functions used in the speaker-verification experiment is given in chapter 3. The union of a certain number of these blocks forms the acceptance region R_A; if a new sample falls into this region it is classified as being a number of class A (here taken to be the class of main-speaker samples). If the number of main speaker training samples is n_a and if m is the number of statistically equivalent blocks combined to form the acceptance region R_A then

$$m = \alpha(n_a + 1)$$

where α is the expected value of false alarm probability that is to be achieved. By chosing the ordering functions to be hyperspheres expanding from each one of the n_a main-speaker training samples one is assured that at least all of the training samples lie in the region R_A. Hopefully this procedure will therefore do well on main-speaker test samples as well.

The ordering functions that are combined to form the acceptance region R_A are formed by ordering the impostor training samples. By ordering the main-speaker samples as well an estimate of the correct classification can be obtained. The expected value of the probability of correct classification obeys the inequality

$$E[P_{cc}] > \frac{b}{b+1}$$

where b is the number of complete blocks in region R_A that can be formed by ordering the main-speaker samples.

To test the system 40 of the 225 available main-speaker samples and 208 of the 6500 impostor samples were used as training samples. The acceptance region was composed of seven blocks giving an expected false alarm probability of 7/209 or .0335. A summary of the major results is given in table 4.4 of progress report 42.

The measured false alarm rate for several different ways of forming the acceptance regions is generally within one standard deviation of the expected value of .0335. Also the probability of miss has roughly the same order of magnitude and turned out to be slightly better for the sample space made up of eight short-time spectra, 32 frequency components per spectrum, than for the other sample spaces.

The procedure was compared with the conceptually much simpler nearest-neighbor method, where a test point is classified according to the class of the nearest training point. The nearest-neighbor method involves

much more computation time than the DFTR method and the sample space was therefore arbitrarily reduced to 48 dimensions. It was found that the reduction in dimensionality increased the miss rate by almost a factor of 10 for the same expected false alarm rate in the DFTR method. The nearest-neighbor method shows a much smaller error rate than the DFTR method; this is to be expected since it utilizes information about all the samples while the DFTR method only uses information about samples that have been ordered. On the other hand the nearest neighbor method cannot be set up to guarantee a specified expected false alarm rate, and it also takes much more computation time. Thus if computation time is a factor the DFTR method is definitely superior. The DFTR method is easy to program, and once the system is trained checking out a new test sample only takes a few seconds of IBM7094 time. The system therefore appears to have practical usefulness.

III. Passive Detection and Tracking Using Surface Scattered Signals

The major effect of surface scattering considered in Progress report No. 43 is the decorrelation produced in the signals received by pairs of hydrophones. A system consisting of two hydrophones is therefore postulated, and it is assumed that the hydrophone signals are processed by a simple cross-correlator as shown in either Fig. 3.1-1 or 3.4-1 of Progress report No. 43.

Three different propagation models are considered. The first considers a direct channel to one of the hydrophones and a surface scattering channel modelled as a random amplitude and delay model to the second hydrophone. The transfer function for the direct channel is

$$H_d(\omega) = \frac{e^{-j\omega R/c}}{R}$$

where R is the line-of-sight distance from target to receiver and c is the

speed of sound. The transfer function for the scattering channel is time varying and has the form

$$H_s(\omega,t) = A(t)e^{-j\omega\tau(t)}$$

where the amplitude function A(t) and the delay function $\tau(t)$ are considered to be independent stationary Gaussian random processes whose variation is slow relative to the signal bandwidth.

In the second model both channels between the transmitter and the two hydrophones are assumed to be random amplitude and delay models with

$$H_1(\omega,t) = A_1(t)e^{-j\omega\tau_1(t)};\ H_2(\omega,t) = A_2(t)e^{-j\omega\tau_2(t)}$$

The amplitude function $A_1(t)$ and the delay function $\tau_1(t)$ are assumed to be jointly independent stationary Gaussian random processes, as are $A_2(t)$ and $\tau_2(t)$; however $A_1(t)$ and $A_2(t)$ are jointly dependent, as are $\tau_1(t)$ and $\tau_2(t)$.

In the third model one direct channel and one surface scattering channel are again assumed, but now the surface is itself modelled as a random sine wave of the form

$$(x,y,t) = h(t)\sin[q_s x\cos\alpha_s + q_s y\sin\alpha_s - \Omega_s t - \chi(t)]$$

where h(t) and $\chi(t)$ are random waveheight and positional phase parameters that are supposed to be very slowly varying. The parameters q_s, α_s, and Ω_s, are the magnitude and orientation of the propagation vector and temporal frequency of the surface respectively.

In all cases the signal x(t) at the transmitter is assumed to be a stationary zero - mean Gaussian random process having a power spectral density $S_{xx}(\omega)$. Noise signals $n_1(t)$ and $n_2(t)$ are assumed to add to the signals entering the hydrophones; these are wide-range stationary Gaussian random processes that are independent from the signal x(t), but not

necessarily jointly independent.

It is to be noted that in all cases a Gaussian signal is operated upon by a random channel function; it is therefore not suprising that the signal received by the hydrophones is no longer gaussian. As a formal demonstration of this fact the fourth-order cumulant of the received signal has been computed and turns out to be non zero under several different input conditions, and for all three of the assumed scattering models. (See fig. 4.4-1).

The correlator output signal is denoted by $\Xi(\tau,T,\rho)$ where τ is the delay introduced in one of the channels to "steer" the two-hydrophones in the direction of the target, T is the averaging time, and ρ is the time at which averaging starts. In the absence of scattering and of noise Ξ would show a sharp peak at $\tau = \tau_o$, where τ_o is the "correct" steering delay. As a result of scattering and/or noise the location of the peak becomes a random variable depending on the instantaneous scattering conditions during the averaging period T, and the height and sharpness of the peak are also reduced so that under severe scattering no clear peak is discernible. (See Fig. 4.2-4).

Several criteria may be employed to evaluate the performance of the correlator. One of these is to form the likelihood ratio of the correlator output for $\tau = \tau_o$ and to assume that a signal is present (hypothesis H_1 is ture) if the likelihood ratio exceeds some threshold; otherwise hypothesis H_o - noise only - is assumed to be correct. For reasonable integration times Ξ is approximately Gaussian with zero mean and variance σ_0^2 if H_o is true, and with a mean μ_1 and variance σ_1^2 if H_1 is true. It is then a straight forward matter to compute the false alarm and miss probabilities, and this is done ir Eqs. (3.8-10) and (3.8-11). The two error probabilities

are seen to depend only on the two normalized standard deviations of Ξ defined by $d_o = \sigma_o/\mu_1$ and $d_1 = \sigma_1/\mu_1$. Unfortunately the dependence is rather complex and not easily visualized. Furthermore this definition of error probabilities is somewhat misleading since a small shift of the peak could result in a very marked reduction of the magnitude of Ξ at $\tau = \tau_o$. This would result in rejection of the hypothesis H_1 even though the peak might still be quite clearly discernible.

Another criterion for evaluating the system may be obtained by considering the location of the peak of Ξ to be an estimate of the true value of τ. Then the variance of this location would be an indication of the accuracy of the estimate. If it is assumed that there is only a single peak, then this criterion is equivalent to computing the variance of the zero crossing of the derivative $\frac{\partial}{\partial\tau}$; specifically, the quantity of interest is this variance normalized with respect to the mean slope of $\frac{\partial}{\partial\tau}$ at $\tau=\tau_o$:

$$\sigma^2_{\tau_o} = \left.\frac{\mathrm{Var}\,[\Xi']}{\left[\frac{\partial\Xi'}{\partial\tau}\right]^2}\right|_{\tau = \tau_o}$$

Actually, it is just as easy to compute the normalized autocorrelation function of Ξ, whose value for zero argument is then $\sigma^2_{\tau_o}$. A general expression for this function, denoted by $R_{\tau_o}(\mu)$, is given in Eq. (4.2-13); unfortunately it is rather complex.

As a third criterion the normalized variance d_1^2 may itself be used since it contains essentially the same information as $\sigma^2_{\tau o}$. A general expression for d_1^2 is given in Eq. (4.2-8), and is seen to be quite similar to the expression for $R_{\tau_o}(\mu)$.

In performing the computation for these expressions is assumed that signal and noise spectra, filter transfer functions, and the spectra of amplitude and delay all have a Gaussian shape; thus the signal spectrum is

given by

$$S_{xx}(\omega) = \frac{\sqrt{2\pi}\, P_x}{\Omega_x} \exp\left[- \tfrac{1}{2} \frac{\omega^2}{\Omega_x^2} \right]$$

The autocorrelation function of the random delay is

$$R_{\tau\tau}(\mu) = \sigma_\tau^2 \exp\left[- \tfrac{1}{2} \Omega_\tau^2 \mu^2 \right]$$

etc. (See Eqs. 4.1-17, 4.1-18, 4.2-2, and 4.2-3).

Curves for d_1 as a function of T are shown in Figs. 4.2-1, 4.2-2, 4.2-3, and 4.2-5. These curves are for the first of three propagation models, but qualitatively the curves for the other two models are similar. The most striking feature of all of these curves is a well-defined plateau over which d_1 decreases only very little as T increases. For values of T smaller or larger than the plateau values the dependence of d_1 is proportional to $1/\sqrt{T}$ as might be expected from rather general statistical considerations.

The existence of the plateau is another indication of the fact that the signals received by the hydrophones is non Gaussian. Qualitatively the plateau is a result of the fluctuations in the instantaneous estimate of τ that results from the random delays produced by the surface scattering. For small integration times an increase in T tends to eliminate fluctuations due to noise and the randomness of the target signal and to produce a better definition of the peak in the output. Hence d_1 decreases. However d_1 also measures the fluctuations of the peak that result from scattering, and when T has become large enough to eliminate noise and signal effects from the peak, the fluctuation due to scattering still persist. Hence d_1 remains essentially constant until the integration time has become so large that the scattering fluctuation are also being "washed out".

For the first propagation model (single random amplitude and delay channel) the level of the plateau in d_1^2 is

$$\left[1 + \left(\frac{\sigma_A}{A_c}\right)^2\right]\left[\frac{\alpha_m^2}{\sqrt{\alpha_m^4 - \sigma_\tau^4}}\right] - 1$$

where σ_A^2 is the variance of the amplitude fluctuation A_c is the mean amplitude of the channel, σ_τ^2 is the variance of the delay fluctuation, and

$$\alpha_m^2 = \frac{1}{\Omega_m^2} = \sigma_\tau^2 + \frac{1}{\Omega_x^2} + \frac{1}{2\Omega_1^2} + \frac{1}{2\Omega_2^2}$$

Here Ω_x is the bandwidth of the signal and Ω_1 and Ω_2 are the bandwidths of the two filters $H_1(\omega)$ and $H_2(\omega)$ used in the correlator. Note that the plateau will exist even if there is no random delay: this is because d_1 measures the total variation of the output peak, not only its motion along the τ axis. The plateau does disappear if both σ_A and σ_τ are zero, as would be expected. It also becomes less pronounced as the signal-to-noise ratio decreases, since the received signal consists then of mostly noise and tends to be Gaussian.

A similar plateau is found in the expression of R_τ (0) which is a better indicator of the tracking error than d_1. The plateau here is

$$\left(1 + \frac{\sigma_A^2}{A_c^2}\right)\left[\frac{\sigma_\tau \alpha_m^6}{\sqrt{\alpha_m^4 - \sigma_\tau^4}}\right]$$

Note that this plateau disappears when $\sigma_\tau = 0$ since amplitude fluctuations do not affect the tracking error in that case.

Similar plateaus are found in the other propagation models. Expressions corresponding to the two given here are Eqs. 4.4-5 and 4.4-8 respectively for the two-scattering channel model and Eqs. 5.3-12 and 5.3-15 for the random sine-wave model.

In the random sine-wave model the height of the plateau can be related to an effective Rayleigh parameter.

$$\Omega_{fx}/c = \frac{\sqrt{2}\ \sigma_h\ \mathrm{Sin}\ \psi\ \Omega_{fx}}{c}$$

where σ_h is the variance of the amplitude, ψ is the grazing angle, c in the sound velocity in water and $\Omega_{fx} = [1/\Omega_f^2 + 1/\Omega_x^2]^{-\frac{1}{2}}$ where Ω_f is the filter bandwidth (assumed to be indentical in the two correlator channels) and Ω_x is the signal bandwidth, suitably defined (See Eq. 5.3-1). The critical value of this parameter for both d_1^2 and $R_{\tau_o}(0)$ is roughly unity: the plateau is small for lesser values, and rises steeply for larger values. It is interesting to note the appearance of the Rayleigh parameter in this context since it is generally a good measure of relative surface roughness.

The height of the plateau can be reduced and the performance of the system improved by reducing the filter bandwidth Ω_f. It is clear from the expression for α_m (where $\Omega_1 = \Omega_2 = \Omega_f$) that this makes $\alpha_m >> \sigma_\tau$ and therefore reduces the second term in the expression for the plateau level. Physically the effect of reducing the bandwidth of the filter is to screen out some of the fluctuation, and it seems reasonable that this should improve the performance if not carried too far. It could also be expected that extreme reduction of filter bandwidth would result again in a worsening of performance, and this is clearly shown in Figs. 4.2-8 and 4.2-9.

Additional results contained in progress report No. 43 deal with other aspects of the scattering transfer functions for the three models. Expressions have been obtained for the interfrequency correlation function, frequency spreading function, and other moments that will be useful in signal design and receiver optimization. These expressions are all rather complex, and details must be obtained from computer calculations. In general, however, the results for all three models in regard to these functions are qualitatively similar.



APPLICATION OF DISTRIBUTION-FREE TOLERANCE REGIONS TO PATTERN RECOGNITION

by

Guy W. Beakley

Progress Report No. 42

General Dynamics/Electric Boat Research

April 1970

DEPARTMENT OF ENGINEERING
AND APPLIED SCIENCE

YALE UNIVERSITY

ABSTRACT

The purpose of this work is to investigate a nonparametric classification procedure based on distribution-free tolerance regions. The procedure is one which gives some knowledge about how well the classifier is expected to perform. This is achieved by using only one sample of statistically independent observations from each class.

The approach, which is called the hypersphere DFTR approach, is formulated in a two class problem. The proposed recognition system is one which can be designed for a given expected false alarm probability or for a given confidence that the false alarm probability is less than a given amount. A few procedures are presented which have certain desirable properties and which appear to do a good job of minimizing the miss probability.

Three principal DFTR procedures are presented. The small and large sample properties of these procedures are investigated and the procedures are compared. A procedure for obtaining a measure of the miss probability is also presented.

The procedures are tested in an automatic speaker verification experiment. A comparison is made of the test false alarm rate with the 95% upper tolerance limit on the false alarm probability and also with the expected false alarm probability. In the experiment all test false alarm rates fell below the 95% upper tolerance limit. The average test false alarm for the 21 different cases studied here was approximately equal to 0.8 of the average expected false alarm probability.

A comparison is made of the test miss rate with a measure of the miss probability that was obtained by using a tolerance region approach. In the speaker verification experiment all test miss rates fell below the 95% upper tolerance limit. For the 21 different cases studied, the average test miss rate was equal to 0.82 of the average expected miss rate.

Finally, the probability of error for the hypersphere DFTR procedure is theoretically compared with the probability of error for the nearest-neighbor rule without assuming the form of the class probability distributions.

TABLE OF CONTENTS

		Page
Chapter 1.	INTRODUCTION	A-1
	Review of Contents	A-9
Chapter 2.	DISTRIBUTION-FREE TOLERANCE REGIONS AND CLASSIFICATION	A-11
	A Brief Review of Distribution-Free Tolerance Regions	A-11
	Classification Using Distribution-Free Tolerance Regions	A-16
Chapter 3.	THE HYPERSPHERE DFTR APPROACH	A-23
3.1.	Summary	A-23
3.2.	Introduction	A-23
3.3.	Application of Distribution-Free Tolerance Regions to Classification	A-25
3.4.	Discussion of Practical Ordering Functions	A-32
3.5.	Programming on a Digital Computer	A-48
3.6.	Large Sample Properties	A-50
3.7.	An Optimum Ordering Procedure	A-59
3.8.	A Further Comparison of the AHE and CHS Procedures	A-62
3.9.	Measure of the Miss Probability	A-67
3.10.	An Ordering Procedure which Gives Distribution-Free Measures of Both the False Alarm and Miss Errors	A-78
3.11.	Multi-class Problem	A-82
Chapter 4.	AUTOMATIC SPEAKER VERIFICATION	A-97
4.1.	Introduction	A-97
4.2.	Speaker Recognition Review	A-98
4.3.	Experimental Setup	A-103
4.4.	Preliminary Study: Speaker Recognition by Analysis of Phonemes	A-105
4.5.	Measurements for the Speaker Verification Experiment	A-110
4.6.	Data	A-117
4.7.	Decision Regions	A-118
4.8.	Reduced Measurement Space	A-126
4.9.	Comparison of Expected False Alarm Probability with Test False Alarm Rate	A-131

4.10.	Comparison of a Measure of the Expected Miss Rate with the Test Miss Rate	A-136
4.11.	7040-7094 Computer Time	A-141
Chapter 5.	THEORETICAL COMPARISON OF THE PROBABILITY OF ERROR FOR THE AHE-DFTR PROCEDURE WITH THE PROBABILITY OF ERROR FOR THE NEAREST-NEIGHBOR RULE	A-144
5.1.	Summary	A-144
5.2.	The Nearest-Neighbor Rule	A-146
5.3.	The AHE-DFTR Procedure	A-151
5.4.	Intuitive Investigation	A-164
Chapter 6.	CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK	A-167
Appendix A.	Theory of Distribution-Free Tolerance Regions	A-173
A.1.	One Dimensional Theory	A-173
A.2.	Generalizations	A-178
A.3.	D-Dimensional Theory	A-183
A.4.	Wald's Ordering	A-184
A.5.	A General Ordering Procedure	A-189
A.6.	Discontinuities	A-196
A.7.	Other Extensions	A-198
Appendix B.	Properties of the Beta Distribution	A-199
Appendix C.	Discussion of $E\left\{\int_{R_2} dF_1(x)\,dx\right\} = \alpha$	A-201
Appendix D.	Review of Classification Methods in Pattern Recognition Given Training Samples of Known Classification	A-204
I.	Optimum Solution with Assumed Probability Densities	A-204
II.	Estimation or Approximation of the Probability Densities	A-208
III.	Estimation or Approximation of the Class Discriminating Boundaries	A-213
IV.	Other Intuitive Criteria	A-220
Bibliography		A-228

LIST OF ILLUSTRATIONS

Figure		Page
1.	Model of a pattern recognizer	A-4
2.1.	Wald's method of successive elimination for forming distribution-free tolerance regions	A-14
2.2.	A general ordering technique for forming distribution-free tolerance regions	A-14
2.3.	An example by Quesenberry and Gessaman (1968)	A-21
3.1.	A hypersphere ordering	A-29
3.2.	All hyperspheres expand (AHE)	A-34
3.3.	Ordered hyperspheres constant (OHC)	A-38
3.4.	All hyperspheres are stopped by one P_1 observation	A-40
3.5.	Conditioned hyperspheres stop (CHS)	A-43
3.6.	Composite of the AHE, OHC, and CHS ordering procedures	A-45
3.7.	A comparison in favor of the AHE procedure	A-47
3.8.	A comparison in favor of the CHS procedure	A-47
3.9.	A one-dimensional example	A-52
3.10.	Length of the accept region on the positive real line	A-52
3.11.	Length of the accept region on the negative real line	A-57
3.12.	Optimum ordering for disjoint probability density functions	A-57
3.13.	Average length of an interval surrounding a P_2 observation (250 observations from each class)	A-63
3.14.	Average length of an interval surrounding a P_2 observation (500 observations from each class)	A-65

		Page
3.15.	Average length of an interval surrounding a P_2 observation (1000 observations from each class)	A-66
3.16.	A situation with a possibly large miss rate	A-68
3.17.	An unsuitable ordering for measuring the miss rate	A-68
3.18.	The OHC-R and CHS-R procedures	A-77
3.19.	Example of the ordering of Section 3.10	A-80
3.20.	A two-class hyperplane approach	A-83
3.21.	A two-class hypersphere approach	A-83
3.22.	A three-class hypersphere approach	A-89
3.23.	An alternate three-class hypersphere approach	A-89
3.24.	A hypersphere approach for criterion A	A-93
3.25.	A hypersphere approach for criterion B	A-93
4.1.	Chart for obtaining measurements	A-112
4.2.	Four utterances of "My Code..."	A-115
5.1.	The transformation $Z = \min(Z_1, Z_2)$	A-150
5.2.	The transformation $Z = \min(\lvert X_1 - Y_1 \rvert, \lvert X_2 - Y_1 \rvert)$	A-150
5.3.	The transformation for equation 5.64	A-160
A-1.	Ordering from a fixed point	A-181
A-2a.	Wald's ordering procedure	A-181
A-2b.	Not distribution-free (assuming dependent variables)	A-181
A-3.	Illustration of a general ordering procedure	A-197
D-1.	Optimum recognition system	A-205
D-2.	A partitioning with linear decision functions	A-215
D-3.	Discriminant functions	A-223
D-4.	Elementary decision rules	A-223

LIST OF TABLES

Table		Page
3.1.	Expected Value and Standard Deviation for $\beta = 0.05$ and $\gamma = 0.95$ and for Different m and n	A-32
3.2.	A Comparison of the AHE and CHS Procedures	A-58
4.1.	Text for Phoneme Analysis	A-106
4.2.	Impostor Data	A-114
4.3.	Main Speaker — RF Data	A-114
4.4.	Three Different 256-Dimensional Measurement Spaces	A-121
4.5.	Test Errors in 3 Different 256-Dimensional Measurement Spaces	A-124
4.6.	48 Dimensional Space, Two Main Speakers	A-127
4.7.	Comparison of Expected False Alarm Probability with Test False Alarm Rate	A-133
4.8a.	Comparison of a Measure of the Expected Miss Probability with the Test Miss Rate in the 256-Dimensional Spaces	A-138
4.8b.	Comparison of a Measure of the Expected Miss Probability with the Test Miss Rate for the Space of 6 Spectra with 8 Components/Spectrum	A-139
B-1.	Comparison of the Mean, Median, and Mode for Two Different Values of m and n	A-200

Chapter 1

INTRODUCTION

The ability to recognize and to respond to visual, auditory, or other patterns can be regarded as a prerequisite for any intelligent behavior, and it is, in fact, possessed by all living things to some degree. In general it can probably be said that the more intelligent an animal is the greater is the repertoire of patterns that it can recognize. Certainly the number of patterns that can be recognized by human beings, ranging over auditory patterns such as speech sound, music, sounds of nature, etc. to visual patterns such as those made by objects, faces, letters, etc. is so vast as to defy enumeration.

Since the early days of computers, attempts have been made to enable these supposedly intelligent machines to recognize patterns also. To some extent these attempts have been quite successful. Every computer possesses the ability to recognize the symbols of its machine-language alphabet, and the developments in computer software over the last dozen years have shown that computers can be made to recognize rather intricate input patterns that seem quite far removed from the basic machine language.

A major difference between the ability of computers and of living beings to recognize patterns appears, however, to be that the latter can recognize patterns that they have never observed before while the former

can generally not do this. Thus a person has no difficulty in recognizing an object to be, say, a glass, even though the precise shape or color may be quite different from any that he has seen before. Apparently the human pattern recognizer is able to react to general features that categorize the pattern without being put off by details that are somehow understood to be irrelevant. This ability is one that, so far, machines possess only very imperfectly.

There are many machine pattern recognition tasks which up to now have not been solved satisfactorily. Some of these are recognition of a person from his handwriting, or from his voice, or from his picture; recognition of spoken messages regardless of the speaker; and recognition of complex structural images from pictures.

Automatic pattern recognition has been attempted in many fields. For instance, in medicine pattern recognition is generally used by the doctor for diagnosis. Machine recognition is being investigated for such seemingly applicable tasks as the analysis of electrocardiograms, electroencephalograms and blood cell photos. Other examples of areas in which machine pattern recognition is being applied are physics, geology, and meteorology. In physics, automatic pattern recognition is being used for particle tracking in bubble chambers. Recognition of the location of oil deposits through seismic and magnetic signal analysis is being attempted by geologists. Meteorologists are investigating weather prediction through the analysis of cloud photographs.

Model

An often-used model for a pattern recognizer was proposed by Marill and Green (1960). This model is shown in Figure 1. It consists of two important parts, the receptor and the categorizer (or classifier).

The receptor transforms the input data, which might be the motion of a transducer or the output of an optical scanner, into a measurement space of high dimensionality in which observations from the same class cluster and observations from unlike classes separate. The transformation may be linear or nonlinear, information preserving or destroying. The categorizer determines the decision regions in the measurement space and tests the proximity of an unclassified observation to these regions.

Decision Theory

Fundamental to the design of the categorizer is statistical decision theory, c.f. Wald (1950), Blackwell and Girshick (1954), Anderson (1958). We briefly review some of this theory that is applicable to our problem.

Suppose an observation is to be classified into one of several classes. The observation is represented by a measurement vector $\underline{v}$ in measurement space ν. The classification procedure can be described as a mapping of measurement space ν into the $i = 1, \ldots, K$ classes. Let R_i be the region of the measurement space which is mapped into class i, $i = 1, \ldots, K$. If a new observation falls in R_i, it is classified into class i. Let ξ_i be the a priori probability that the observation

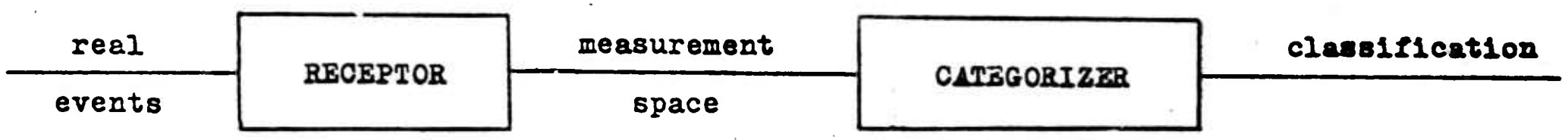


Figure 1. Model of a Pattern Recognizer.

belongs to class i, let $f_i(\underline{v})$ be the probability density function of the observation $\underline{v}$, assuming that it is a member of class i, and let $C_i(j)$ be the cost of deciding that $\underline{v}$ is a member of class j when it is a member of class i.

The expected risk or loss in making decisions is

$$\sum_{i=1}^{K} \xi_i \sum_{j=1}^{K} C_i(j) \int_{R_j} f_i(\underline{v})\, d\underline{v} \,. \tag{1.1}$$

According to the Bayes criterion the expected loss is minimized by deciding that $\underline{v}$ belongs to class k when

$$\sum_{i=1}^{K} \xi_i f_i(\underline{v})\, C_i(k) \leq \sum_{i=1}^{K} \xi_i f_i(\underline{v})\, C_i(j) \tag{1.2}$$

for all j.

Suppose the cost of making a correct decision is zero and the cost of making an incorrect decision is equal to C. Then $C_i(j) = \begin{cases} C & j \neq 1 \\ 0 & j = i \end{cases}$.

By subtracting $C \sum_{\substack{i=1 \\ i \neq j,k}}^{K} \xi_i f_i(\underline{v})$ from both sides of equation 1.2 and by dividing through by C we obtain the decision rule for deciding in favor of class k. The decision is made that $\underline{v}$ belongs to class k when

$$\xi_j f_j(\underline{v}) \leq \xi_k f_k(\underline{v}) \tag{1.3}$$

for all j.

It is sometimes convenient to formulate the decision rule in terms of the likelihood ratio,

$$L_{kj} = \frac{f_k(\underline{v})}{f_j(\underline{v})} \,.$$

Rewriting equation 1.3 in terms of the likelihood ratio, one decides in favor of class k if

$$L_{kj} \geq \frac{\xi_j}{\xi_k} \tag{1.4}$$

for all j .

For a two-class problem, one simply compares the likelihood ratio with a constant. Many criteria yield decision rules involving likelihood ratio comparisons. Some of these are the Neyman-Pearson criterion, the Ideal Observer criterion, and the Minimax criterion, c.f. Van Trees (1968).

In the following chapters the two-class problem is discussed at length. For convenience, the two conditional probabilities of error will be defined as follows. The conditional probability of deciding in favor of class 2 when the observation belongs to class 1 (false acceptance of class 2) is called the false alarm probability. The mathematical notation for this conditional probability is

$$P_{FA} = \int_{R_2} f_1(\underline{v})\, d\underline{v} . \tag{1.5}$$

The conditional probability of deciding in favor of class 1 when the observation belongs to class 2 is called the miss probability. It is denoted by

$$P_M = \int_{R_1} f_2(\underline{v})\, d\underline{v} . \tag{1.6}$$

The use of the terms "false alarm" and "miss" implies that classes 2

and 1 are associated respectively with the occurrence or nonoccurrence of some event (such as the presence of a target on a radar screen). These terms are more appropriate for this thesis than the less descriptive "error of the first kind" and "error of the second kind" because the major problem dealt with here involves a main class and an impostor class, c.f. Section 4.1.

Classification Methods

In general, the conditional joint probability densities, $f_i(\underline{v})$, are not known in pattern recognition problems. Usually the only information available for designing the pattern recognizer is a limited set of properly classified data. In this case an optimum solution in the Bayes sense is not applicable.

Let us briefly define some of the terminology which distinguishes the classification procedures. A machine is said to "learn" when it is able to improve its performance by benefiting from its past experience. The period of constructing decision regions is called the training period. This is distinguished from the recognition or test period in which observations of unknown classification are classified. Pattern recognition may be accomplished by supervised or unsupervised learning. In the former, the training observations are of known classification and in the latter, which is sometimes called learning without a teacher, the training observations are of unknown classification.

Some of the methods which have been proposed for solving the

classification problem in pattern recognition are:

(1) Optimum Solution with Assumed Probability Densities: Functional forms for the conditional densities $f_i(\underline{v})$, $i = 1, \ldots, K$ are assumed to be known. Some of the parameters of these densities are often estimated from training observations. The optimum decision regions are found for these assumptions. Then new observations are classified. If the results are not satisfactory, the assumptions are revised and new decision regions are formed.

(2) Estimation or Approximation of the Probability Densities: The class probability densities are estimated or approximated using the training observations. A new observation is then classified according to Bayes rule, where the estimated probability densities are substituted for the true probability densities.

(3) Estimation or Approximation of the Class Discriminating Boundaries: A structure for the boundary which partitions the measurement space into decision regions is assumed. The structures for the discriminating boundary range from the simple hyperplane to complex surfaces of the form $\sum_{i=1}^{m} w_i \phi_i$ where the ϕ_i are functions of the measurement space and m is finite. The discriminating boundary is then trained for the "best" results. "Best" results usually means minimum probability of error when the class probability densities are assumed or minimum misclassification of the training observations when the class probability densities are not assumed.

(4) Other Intuitive Criteria: These include such approaches as maximization of entropy, minimization of intraclass distance around characteristic points of the classes, the Fisher Criterion, and the nearest-neighbor rule.

Appendix D gives a review of these classification methods along with the associated references. It should be noted that a comparison of the above approaches is very difficult since the criterion for a good pattern recognizer varies from author to author and since the data sets on which the pattern recognizers are tested are usually different.

Review of the Contents

It is the intention of the present work to investigate a nonparametric classification procedure based on distribution-free tolerance regions. This procedure is one which gives some knowledge about how well the classifier is expected to perform. This is achieved by using only one sample of statistically independent observations from each class. The classification procedure is then applied to a practical pattern recognition problem.

In Chapter 2 a brief review of the theory of distribution-free tolerance regions is presented. A detailed study of this theory is made in Appendix A. A review of existing methods for applying the theory of distribution-free tolerance regions to classification problems also appears in Chapter 2.

The effectiveness of certain methods for constructing distribution-

free tolerance regions for classification purposes is investigated in Chapter 3. The approach, which is called the hypersphere DFTR approach, is formulated in a two-class problem. The proposed recognition system is one which can be designed for a given expected false alarm probability or for a given confidence that the false alarm probability is less than a given amount. It is assumed that the only information available for designing the recognizer is a properly labeled sample of statistically independent observations from each class. A few procedures are presented which have certain desirable properties and which appear to do a good job of minimizing the miss probability. A procedure for obtaining a measure of the miss probability is also presented. The extension of the hypersphere DFTR procedure to the multiclass problem is also discussed.

Chapter 4 reports on an automatic speaker verification system and its use in testing the hypersphere DFTR classification schemes.

A theoretical comparison of the probability of error for a hypersphere DFTR procedure with the probability of error for the nearest-neighbor rule is presented in Chapter 5.

Chapter 6 presents conclusions and lists suggestions for further study.

Chapter 2

DISTRIBUTION-FREE TOLERANCE REGIONS AND CLASSIFICATION

Introduction

Existing classification methods which involve the theory of distribution-free tolerance regions are discussed in this chapter. The chapter begins with a brief review of the theory of distribution-free tolerance regions. For further details on this subject, see Appendix A. Later in Chapter 2 classification procedures by Anderson (1966) are presented. Next, the use of statistically equivalent blocks and the empirical Bayes approach by Patrick (1966) is discussed. Later, a method by Quesenberry and Gessaman (1968), which involves regions of indecision, is presented.

A Brief Review of Distribution-Free Tolerance Regions

Suppose n_1 independent observations, $X_1, X_2, \ldots, X_{n_1}$, are available from a population with continuous univariate probability density $f_1(x)$. Let $X_{(1)} < X_{(2)} < \ldots < X_{(n_1)}$ denote the observations arranged in ascending order of magnitude. It was first shown by Wald (1941) that the amount of probability in $(X_{(r)}, X_{(n_1-r+1)})$ is distribution-free. Hence a statement such as

$$E\left\{\int_{X_{(r)}}^{X_{(n_1-r+1)}} f_1(x)\,dx\right\} = \alpha \qquad 0 \le \alpha \le 1 \qquad (2.1a)$$

or

$$P_r\left\{\int_{X_{(r)}}^{X_{(n_1-r+1)}} f_1(x)\,dx < \beta\right\} = \gamma \qquad 0 \le \beta \le 1,\ 0 \le \gamma \le 1 \qquad (2.1b)$$

can be made even though the density $f_1(x)$ is unknown.

The theory of distribution-free tolerance regions was extended for multivariate distributions by Wald (1943). He formed distribution-free tolerance regions by successive elimination of sample regions of the multidimensional space. For example, suppose a statement is to be made about the amount of probability in the "center" of the two-dimensional distribution $F_1(x_1, x_2)$. Let the independent observations $X_i = \begin{pmatrix} X_{i1} \\ X_{i2} \end{pmatrix}$, $i = 1, \ldots, n_1$ be arranged in ascending order of magnitude of the 1st variate, x_1. Denote the ordered variate values by $X_{(1)1} < X_{(2)1} < \ldots < X_{(n_1)1}$. Let r be an integer which is less than $n_1/2$. Let the region for which the first variate x_1 is less than or equal to the r^{th} smallest first variate of the n_1 observations be "eliminated" from the space. That is, the space is partitioned into two regions. One region, the region $\{\underline{x}: x_1 \le X_{(r)1}\}$ will not be considered in further ordering of the observations, hence it is "eliminated". Since the stated interest is in the center portion of $F_1(x_1, x_2)$, eliminate

the region for which $x_1 \geq X_{(n_1-r)1}$. Further eliminate the region for which the second variate is less than or equal to the s^{th} smallest second variate of the remaining observations. Here, s is an integer which is less than $(n_1-2r)/2$. Also eliminate the region $\{\underline{x}:x_2 \geq X_{(n_1-2r-2s)2}\}$. The remaining region $\{\underline{x}:X_{(r)1} < x_1 < X_{(n_1-r)1},\ X_{(s)2} < x_2 < X_{(n_1-2r-2s)2}\}$ is distribution-free. Figure 2.1 shows such a region R_2 for $s = r = 2$.

The procedures for forming distribution-free tolerance regions have been generalized in papers by Scheffé and Tukey (1945), Tukey (1947), Tukey (1948), Fraser and Wormleighton (1951), Fraser (1951), Fraser (1953), and Kemperman (1956). A particularly useful generalization is the following: Suppose that n_1 independent observations are available from a continuous D-dimensional cumulative distribution function $F_1(x_1, x_2, \ldots, x_D) = F_1(\underline{x})$. Let $h_i(\cdot)$, $i = 1, \ldots, n_1$ be n_1 functions such that $h_1(\underline{X}), \ldots, h_{n_1}(\underline{X})$ are random variables with a continuous joint distribution function. The functions $h_i(\underline{x})$ are called ordering functions. They are used to partition the sample space into $n_1 + 1$ mutually exclusive and exhaustive sample regions called "statistically equivalent blocks." The regions are hereafter called simply "blocks." Let $\underline{X}_{(1)}$ be the observation which yields the smallest value for the first ordering function, $h_1(\underline{x})$. Then the $n_1 + 1$ blocks $B_1, \ldots, B_{n+1}$ can be defined as follows:

$$B_1 = \{\underline{x}:h_1(\underline{x}) \leq h_1(\underline{X}_{(1)})\}$$

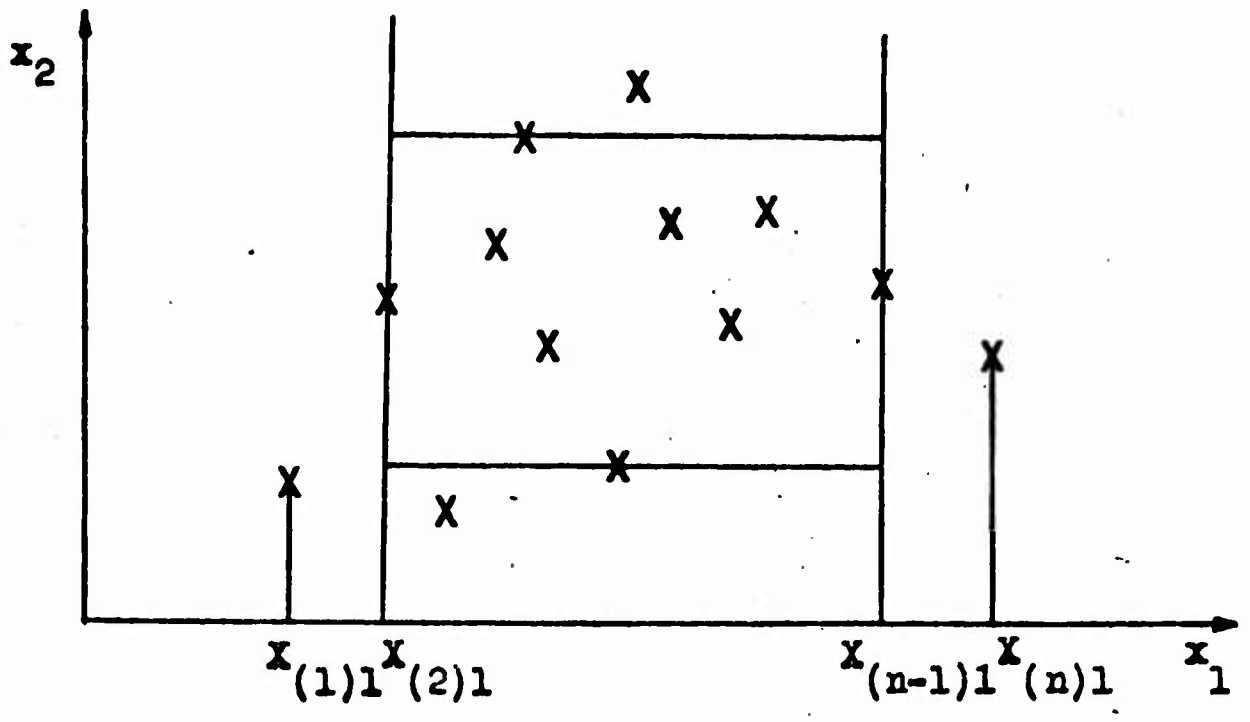


Figure 2.1. Wald's Method of Sucessive Elimination for Forming Distribution-Free Tolerance Regions.

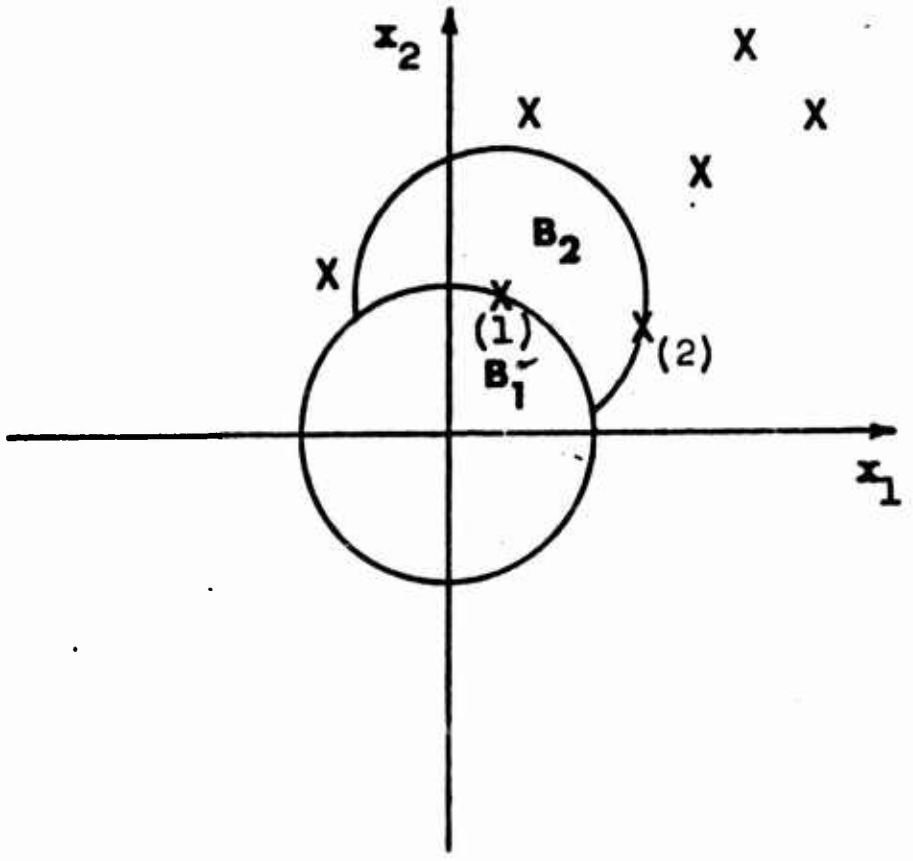


Figure 2.2. A General Ordering Technique for Forming Distribution-Free Tolerance Regions.

where $\underline{X}_{(1)}$ is the observation for which

$$h_1(\underline{X}_{(1)}) = \min_{1 \le i \le n_1} h_1(\underline{X}_i) ,$$

$$B_2 = \{\underline{x}: h_2(\underline{x}) \le h_2(\underline{X}_{(2)}),\ h_1(\underline{x}) > h_1(\underline{X}_{(1)})\} \tag{2.2}$$

where $\underline{X}_{(2)}$ is the observation $\underline{X}$, excluding $\underline{X}_{(1)}$, for which $h_2(\underline{x})$ is minimum,

$$h_2(\underline{X}_{(2)}) = \min_{\substack{1 \le i \le n_1 \\ i \ne (1)}} h_2(\underline{X}_i) .$$

$$\vdots$$

$$B_k = \{\underline{x}: h_k(\underline{x}) \le h_k(\underline{X}_{(k)}),\ h_{k-1}(\underline{x}) > h_{k-1}(\underline{X}_{(k-1)}), \ldots,$$
$$h_1(\underline{x}) > h_1(\underline{X}_{(1)})\} \tag{2.3}$$

where $\underline{X}_{(k)}$ is the observation, excluding $\underline{X}_{(1)} \ldots \underline{X}_{(k-1)}$, for which $h_k(\underline{x})$ is minimum,

$$h_k(\underline{X}_{(k)}) = \min_{\substack{1 \le i \le n_i \\ i \ne (1), \ldots, (k-1)}} h_k(\underline{X}_i) , \tag{2.4}$$

and B_{n_1+1} is the space which remains after the n_1 blocks have been formed,

$$B_{n_1+1} = X - \bigcup_{i=1}^{n_1} B_i . \tag{2.5}$$

All ordering functions subsequent to h_k may depend on the blocks previously formed, all known boundary observations, and on certain sets of indices. For example, let $h_1(\underline{x}) = |\underline{x}|$. Let $\underline{X}_{(1)}$ be

the observation that minimizes h_1, $h_1(\underline{X}_{(1)}) = \min_{1 \le i \le n_1} h_1(\underline{X}_i)$. Then the first block B_1 consists of the region inside the hypersphere $|\underline{x}| = |\underline{X}_{(1)}|$. The block B_1 along with the observation $\underline{X}_{(1)}$ is now eliminated from the sample space. The information found in the 1st ordering (e.g. the location of $\underline{X}_{(1)}$, the size of B_1, etc.) can be used to form subsequent blocks. For example, the ordering function $h_2(\underline{x}) = |\underline{x} - \underline{X}_{(1)}|$ can be used to order the second observation and form the second block. Let $\underline{X}_{(2)}$ be the observation among the remaining $n_1 - 1$ observations which yields the smallest value for $h_2(\underline{x})$. Then

$$B_2 = \{\underline{x}: h_2(\underline{x}) \le h_2(\underline{X}_{(2)}),\ h_1(\underline{x}) > h_1(\underline{X}_{(1)})\}.$$

The formation of the first two blocks in this example is illustrated in Figure 2.2 for a two-dimensional vector $\underline{x}$.

Note that the first block could have been formed by choosing the observation which gives the largest value for $h_1(\underline{x})$. Of course, this block would not be the same as B_1 of Figure 2.2. Distribution-free tolerance regions can even be formed by choosing the r^{th} smallest or r^{th} largest value for h_1, c.f. Fraser (1957).

Classification Using Distribution-Free Tolerance Regions

Anderson (1966) proposes various multivariate statistical techniques based on the properties of statistically equivalent blocks. He presents procedures for (1) testing the hypothesis that an unknown cumulative distribution is a specified one, (2) testing the hypothesis

that two unknown distributions are identical, and (3) classifying an observation into one of two populations. We are interested in the classification techniques. Let $\underline{X}_1,\ldots,\underline{X}_n$ be n independent vector observations from a population with distribution $F(\underline{x})$ and $\underline{Y}_1,\ldots,\underline{Y}_m$ be m independent vector observations from a population $G(y)$, where $F(\underline{x})$ is assumed different from $G(\underline{y})$. Let $\underline{V}$ be a new observation which is drawn from one of the two populations. The observation $\underline{V}$ is to be classified into one of the populations. Anderson mentions several nonparametric classification procedures based on ordering the observations.

In one procedure, the blocks are formed by ranking the pooled $\underline{X}$ and $\underline{Y}$ observations. Consider the block in which $\underline{V}$ falls. The observation $\underline{V}$ is classified according to the majority of observations defining the block. For example, suppose $\underline{V}$ falls in a block which has four sides. If three of these four sides are drawn through $\underline{X}$ observations, then $\underline{V}$ is classified as an $\underline{X}$ observation. In another procedure the $\underline{X}$ and $\underline{Y}$ observations are ordered separately. Consider the $\underline{X}$-block and the $\underline{Y}$-block that $\underline{V}$ falls in. If there are fewer $\underline{Y}$ observations in the $\underline{X}$-block than $\underline{X}$ observations in the $\underline{Y}$-block, $\underline{V}$ is classified as an $\underline{X}$ observation. This procedure is similar to the k_n-nearest neighbor rule which is discussed in Appendix D. Other similar classification procedures are mentioned. Anderson points out that some of these classification procedures can be made to depend on n and m so as n and m increase the probabilities of misclassification

will converge to the probabilities of a procedure based on the likelihood ratio.

Patrick (1966) and Patrick and Fisher (1967) present a general classification approach which they refer to as an empirical Bayes approach for distribution-free minimum conditional risk learning systems. This approach involves the construction of distribution-free tolerance regions for each class. Classification is obtained by comparing the volumes of the tolerance regions for the different classes. For example, consider the tolerance regions in which a new observation $\underline{V}$ falls. Note that each class has been ordered separately and for every $\underline{V}$ there is one tolerance region to be considered for each class. The observation $\underline{V}$ is classified into the class whose tolerance region has the smallest volume, with appropriate compensation being made for the loss functions and the a priori class probabilities.

The approaches of Anderson and of Patrick do not use the blocks to obtain an estimate of how well the classifier will perform. Since the decision regions contain partial blocks, this information cannot be obtained accurately from classifiers of their design. This fact becomes clearer as we study Chapter 3.

A different use for distribution-free tolerance regions in classification is made by Quesenberry and Gessaman (1968). Emphasis is placed upon the control of the distribution of the conditional probabilities of error, i.e. the false alarm probability and the miss probability in the two-class problem. This approach requires a region in the

measurement space which is commonly called a reject region or a deferred decision region. If a new observation falls in this region, no decision is made. The problem with their approach is that the size of this region depends on the location of the observations from the various classes and on the ordering functions chosen. No control is exercised over the size of the reject region. Hence if the distributions are "close" together or if the ordering functions are unhappily chosen, the probability of not making a decision can be large.

Quesenberry and Gessaman's procedure involves the construction of a distribution-free tolerance region A_j containing a_j blocks for each distribution F_j, $j = 1, \ldots, K$. For each set A_j there is a complement set $\bar{A}_j = \chi - A_j$. Let R_j be the region in which the decision is made that the new observation comes from distribution F_j. Let R_j be given by

$$R_j = \bar{A}_j \bigcap_{\substack{i=1 \\ i \neq j}}^{K} A_i$$

Let R_o be the region in which no decision is made. Let R_o be given by

$$R_o = \left(\bigcap_{i=1}^{K} A_i \right) \cup \left(\bigcap_{i=1}^{K} \bar{A}_i \right)$$

The probability of deciding any class other than class j when the new observation is from class j is controlled since there are no more than a_j blocks in the regions for deciding any other class. These ideas become more transparent as this approach and the hypersphere DFTR

approach are investigated in Section 3.11.

The choice of ordering functions is left to the person who implements the classifier. Quesenberry and Gessaman give examples of appropriate ordering functions for (1) two distributions with a monotone likelihood ratio and (2) two univariate normal distributions. The third example which was given is repeated below. Suppose two classes are represented in a two-dimensional space by two distributions, both of which are thought to be unimodal. A reasonable choice for $\bar{A}_j$ is a bounded convex region containing $(n_j - a_j + 1)$ blocks. This can be accomplished in many ways. Figure 2.3 shows an artificial example which was given to illustrate the tolerance region approach. The data was generated by drawing samples of size $n_1 = n_2 = 81$ from bivariate normal distributions P_1 and P_2 with mean vectors

$$(\mu_{11}, \mu_{12}) = (0, 0), \quad (\mu_{21}, \mu_{22}) = (3, 0)$$

and covariance matrices

$$\sum_1 = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \quad \sum_2 = \begin{bmatrix} 1 & 2 \\ 2 & 9 \end{bmatrix}.$$

An ordering which was suggested by Tukey (1947) was then used to construct the tolerance regions. Figure 2.3 is the resulting figure for a probability of .90 that the conditional probability of either error is less than 0.14. The region for deciding class 1 is R_1; the region for deciding class 2 is R_2, and the region for making no decision is R_0.

A problem with the use of distribution-free tolerance regions for estimation of how well the classifier will perform is that the ordering

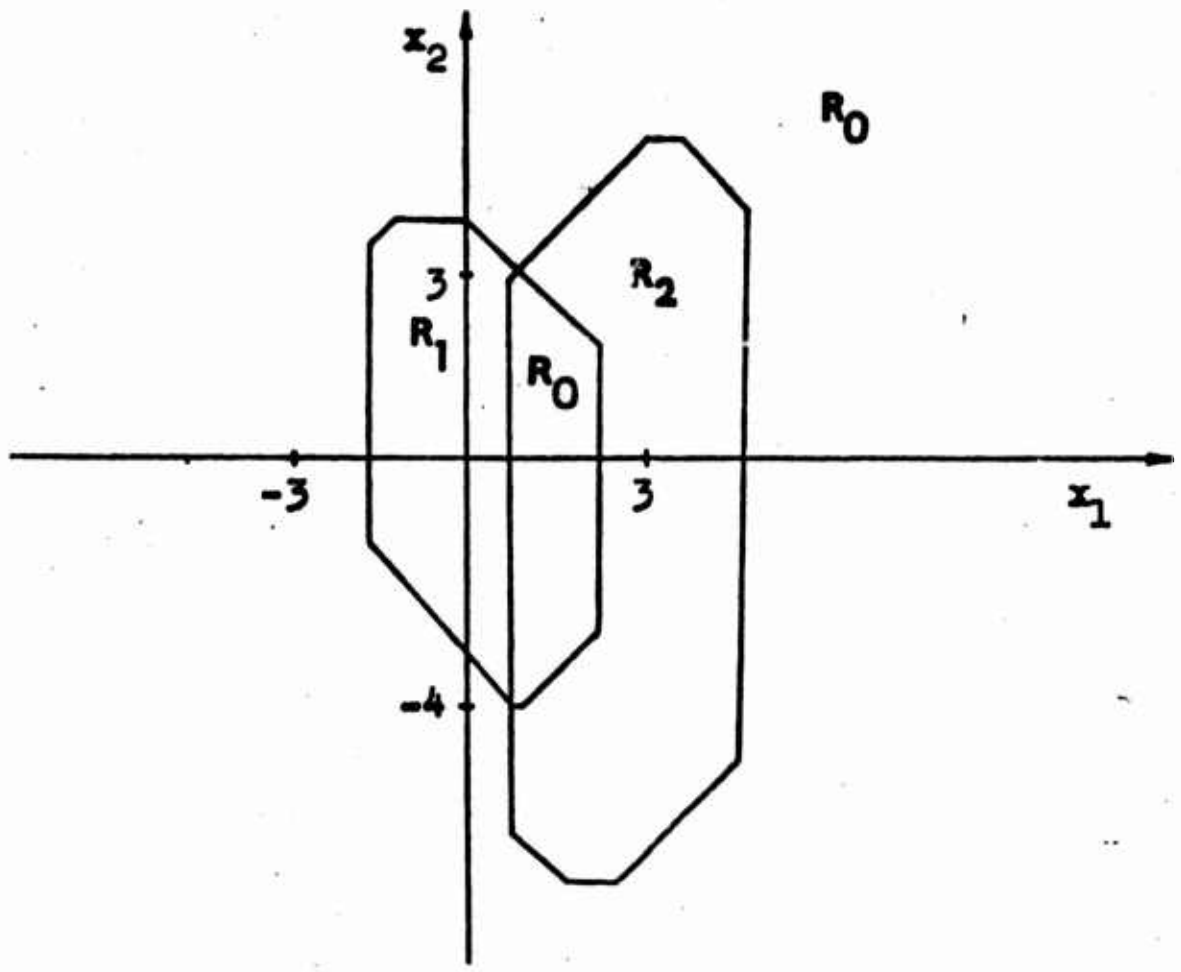


Figure 2.3. An Example by Quesenberry and Gessaman (1968).

functions and the blocks to be used in the decision regions should be chosen without any knowledge of the outcome of the observations. (They can, of course, be based on any a priori knowledge.) Hence if nothing is known about the distributions, a classifier which yields very poor results can be obtained.

In the following chapter, ordering procedures are presented for the case where nothing is known about the class probability distributions. Use is made of the fact that the location of the observations of one class can be used to order the observations of the other classes. Hence the decision regions can conform to the "shape" of the classes. These ordering procedures create decision regions suitable for multimodal class distributions. This is, of course, not the case with the ordering of Figure 2.8. Furthermore, the procedures of the next chapter do not yield a reject region, R_0. Further comparison of this approach with the one of Quesenberry and Gessaman is made in Section 3.11.

Chapter 3

THE HYPERSPHERE DFTR APPROACH

3.1. Summary

The effectiveness of certain methods for constructing distribution-free tolerance regions for classification purposes is investigated in this chapter. The approach is first formulated in a two class problem. The proposed recognition system is one which can be designed for a given expected false alarm probability (probability of misrecognizing a class 1 event as a class 2 event) or for a given confidence that the probability of false alarm is less than a given amount. It is assumed that the only information available for designing the recognizer is a properly labeled sample of statistically independent observations from each class. A few procedures are presented which have certain desirable properties and which appear to do a good job of minimizing the miss probability (probability of misrecognizing a class 2 event as a class 1 event). A procedure for obtaining a measure of the miss probability is also discussed.

3.2. Introduction

Let $\{P_i / i \in \Omega\}$, where $\Omega = \{1, \ldots, K\}$ is a finite parameter space, be a class of probability measures defined over measure space (X, A, μ). Based on n_i statistically independent observations from

P_i, $i = 1, \ldots, K$ a method is sought for classifying an unknown observation x into one of the K classes described by P_i.

Let us consider the case when $K = 2$. Suppose that the probability density functions exist and are defined by

$$P_i(X \leq x) = F_i(x) = \int_{-\infty}^{x} f_i(z)\, d\mu(z) \qquad i = 1, 2 \tag{3.1}$$

Suppose further that the a priori probability that the observation x belongs to class i is ξ_i; clearly $\sum_{i=1}^{2} \xi_i = 1$. Using the Bayes Criterion, one decides that x belongs to class 1 if

$$\frac{f_1(x)}{f_2(x)} > \frac{\xi_2}{\xi_1} \frac{[C_2(1) - C_2(2)]}{[C_1(2) - C_1(1)]} . \tag{3.2}$$

$C_i(j)$ is the cost of classifying an observation from class i into class j.

If the a priori probabilities are unknown, one can use the Neyman-Pearson criterion and maximize

$$\int_{R_2} dF_2(x) \tag{3.3}$$

subject to the condition that

$$\int_{R_2} dF_1(x) \leq \alpha \qquad 0 \leq \alpha \leq 1 . \tag{3.4}$$

It is well known that this criterion also yields a likelihood-ratio test. That is, one decides that x belongs to class 1 if

$$\frac{f_1(x)}{f_2(x)} > L \tag{3.5}$$

where L is such that equation 3.4 is satisfied.

In the pattern recognition problem considered here it is assumed that the probability densities $f_i(x)$, $i = 1, 2$ are unknown. The following analogue of the Neyman-Pearson criterion evolves naturally. Given n_i statistically independent observations from P_i, $i = 1, 2$ it is desirable to maximize

$$\int_{R_2} dF_2(x) \tag{3.6}$$

subject to the condition

$$E\left\{\int_{R_2} dF_1(x)\right\} \leq \alpha \qquad 0 \leq \alpha \leq 1 \tag{3.7a}$$

or subject to the condition

$$Pr\left\{\int_{R_2} dF_1(x) \leq \beta\right\} \geq \gamma \qquad \begin{array}{l} 0 \leq \beta \leq 1 \\ 0 \leq \gamma \leq 1 \end{array} \tag{3.7b}$$

Conditions (3.7a) or (3.7b) can be established even though $F_1(x)$ is unknown. This is done through the theory of distribution-free tolerance regions. A tolerance region with the property, $E\{\int_{R_2} dF_1(x)\} = \alpha$, is known as an α-expected tolerance region. A tolerance region with the property, $Pr\{\int_{R_2} dF_1(x) \geq \beta\} = \gamma$, is known as a β content tolerance region at level γ. It should be noted that $E\{\int_{R_2} dF_1(x)\} = \alpha$ can be considered an α-confidence statement that a new observation from $F_1(x)$ will fall in R_2. This fact is demonstrated in Appendix C.

3.3. Application of Distribution-Free Tolerance Regions to Classification

As previously stated we would like to maximize

$$\int_{R_2} dF_2(\underline{x}) \tag{3.6}$$

under one of the following constraints:

$$E\left\{\int_{R_2} dF_1(\underline{x})\right\} \leq \alpha \qquad 0 \leq \alpha \leq 1 \tag{3.7a}$$

or

$$Pr\left\{\int_{R_2} dF_1(\underline{x}) \leq \beta\right\} \geq \gamma \qquad 0 \leq \beta \leq 1 \qquad 0 \leq \gamma \leq 1 \tag{3.7b}$$

The problem is to order the observations from $F_1(x)$ so that R_2 consists of the number of blocks "m" required to satisfy (3.7a) or (3.7b) and so that $\int_{R_2} dF_2(\underline{x})$ is maximized. The number of blocks needed to satisfy equation 3.7b can be found by consulting tables of the Beta distribution, tables of the cumulative binomial distribution, tables by Somerville (1958), or graphs by Murphy (1948). The number of blocks needed to satisfy equation 3.7a can be obtained from the equation $E\{\int_{R_2} dF_1(\underline{x})\} = m/(n_1+1)$. Therefore, if equation 3.7a is to be satisfied, m is the largest integer less than or equal to $(n_1+1)\,\alpha$.

The blocks should be constructed so that $\int_{R_2} dF_2(\underline{x})$ is maximized. It is assumed that the only information given about $F_1(\underline{x})$ or $F_2(\underline{x})$ is that they are continuous cumulative distribution functions. Therefore, given only a finite number of observations from each class, one can never be certain that $\int_{R_2} dF_2(\underline{x})$ is maximized. A likely approach is to construct R_2 so that it contains as many P_2 observations as possible. All of the P_2 observations can be contained in R_2 if the P_1 observa-

tions, $\underline{X}_1^{(1)}, \ldots, \underline{X}_{n_1}^{(1)}$ are ordered by functions which are in some sense centered about all n_2 of the P_2 observations, $\underline{X}_1^{(2)}, \ldots, \underline{X}_{n_2}^{(2)}$. Then R_2 is made up of the first m blocks established by the ordering.[1]

To accomplish this ordering consider the continuous functions $d_{kj}(\underline{x}, \underline{X}_k^{(2)})$ of the arbitrary observation vector $\underline{x}$, where $k=1, \ldots, n_2$, $j=1, \ldots, n_1$. These functions are basically "distance" functions that satisfy the following conditions:

$$\begin{aligned} &1. \quad d_{kj}(\underline{x}, \underline{X}_k^{(2)}) \geq 0 \\ &2. \quad d_{kj}(\underline{X}_k^{(2)}, \underline{X}_k^{(2)}) = 0 \end{aligned} \tag{3.8}$$

A simple example is

$$d_{kj}(\underline{x}, \underline{X}_k^{(2)}) = |\underline{x} - \underline{X}_k^{(2)}| \tag{3.9}$$

(Note that the subscript k is used to label possibly differing functions which can be associated with each of the observations from the class P_2. The subscript j is used to label possibly differing functions used to form successive blocks. The need for such functions is illustrated in the next section.)

The $n+1$ blocks $B_1, B_2, \ldots, B_{n+1}$ can now be defined as follows: First, let $d_1(\underline{X}_{(1)}^{(1)}$ be the smallest distance between points from the two different classes; i.e.

---

1. The idea of using ordering functions which are centered by the P_2 observations was suggested by Professor I. R. Savage.

$$d_1(\underline{X}_{(1)}^{(1)}) = \min_{1 \le k \le n_2} \min_{1 \le i \le n_1} d_{k1}(\underline{X}_i^{(1)}, \underline{X}_k^{(2)}) \qquad (3.10)$$

Define the regions L_{k1}, $k = 1, \ldots, n_2$ by

$$L_{k1} = \left\{ \underline{x} : d_{k1}(\underline{x}, \underline{X}_k^{(2)}) \le d_1(\underline{X}_{(1)}^{(1)}) \right\} \qquad (3.11)$$

For two-dimensional vectors $\underline{x}$ and a metric as given by equation 3.9, the regions L_{k1} are seen to be circles centered at the points $\underline{X}_i^{(2)}$ with radius $d_1(\underline{X}_{(1)}^{(1)})$. (see Fig. 3.1) The definition of $d_1(\underline{X}_{(1)}^{(1)})$ is such that at least one of these circles contains a point from class 1 on its circumference (the point $\underline{X}_2^{(1)}$ in the figure). The probability that there are more than one such points is assumed negligible. This point is labeled $\underline{X}_{(1)}^{(1)}$ and is said to be ordered. It is clear that for n-dimensional vectors and for the metric of equation 3.9, the regions L_{k1} are hyperspheres.

The first block B_1 is now defined as the union of all the regions L_{k1}:

$$B_1 = \bigcup_{k=1}^{n_2} L_{k1} \qquad (3.12)$$

To obtain the second block the distance $d_2(\underline{X}_{(2)}^{(1)})$ is defined by

$$d_2(\underline{X}_{(2)}^{(1)}) = \min_{1 \le k \le n_2} \min_{\substack{1 \le i \le n_1 \\ i \ne (1)}} d_{k2}(\underline{X}_i^{(1)}, \underline{X}_k^{(2)}) \qquad (3.13)$$

The implication of the subscript "2" of $d_{k2}(\cdot)$ is that $d_{k2}(\cdot)$ and $d_{k1}(\cdot)$ can be completely different functions. The regions L_{k2} are then defined as before by

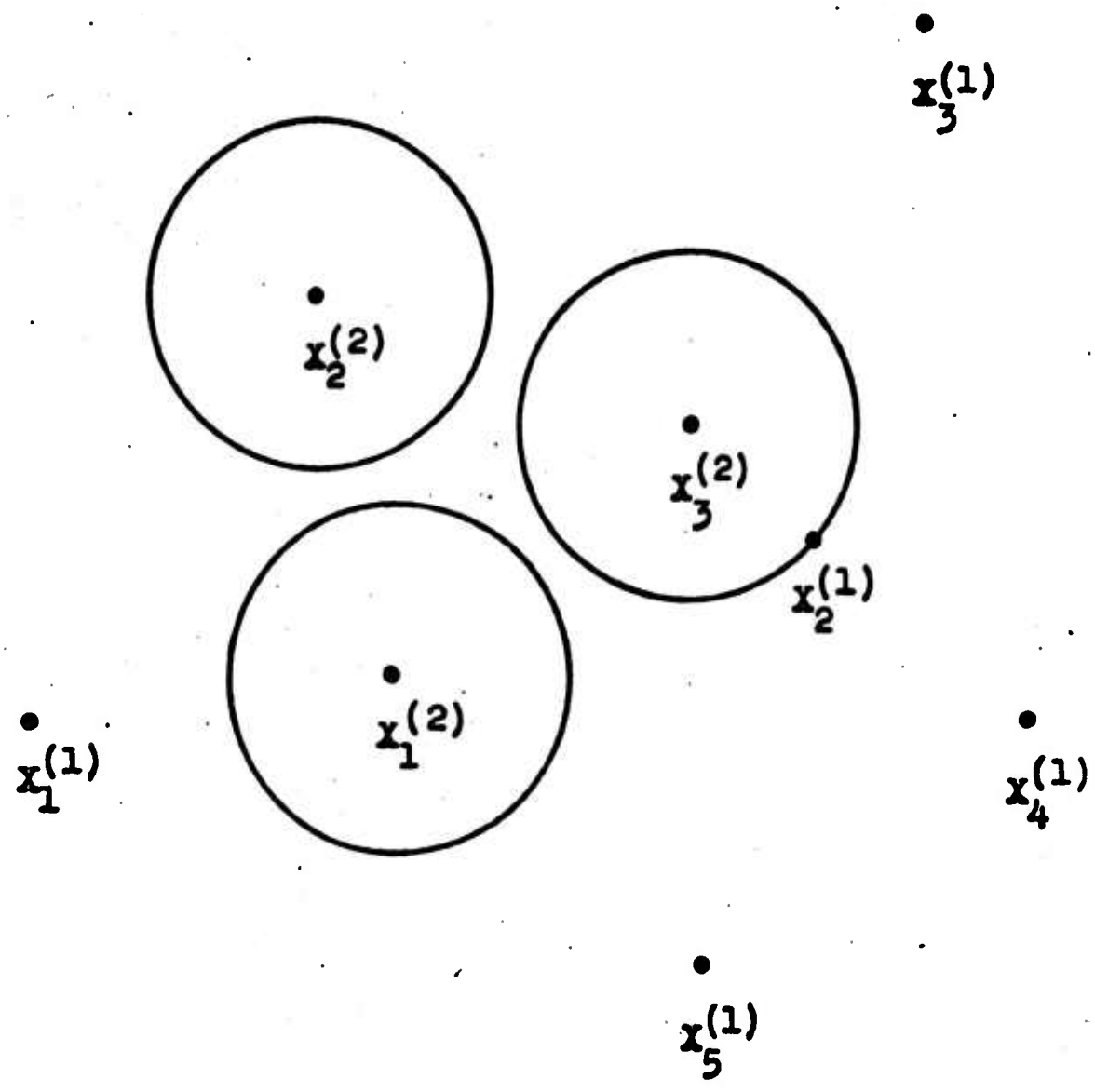


Figure 3.1. A Hypersphere Ordering.

$$L_{k2} = \left\{ \underline{x} : d_{k2}(\underline{x}, \underline{X}_k^{(2)}) \leq d_2(\underline{X}_{(2)}^{(1)}) \right\} \quad k = 1, \ldots, n_2 \tag{3.14}$$

If $d_{k2}(\cdot)$ and $d_{k1}(\cdot)$ are both of the form of equation 3.9 then, for two-dimensional vectors the regions L_{k2} are circles extending to the next closest point of class 1, which is labeled $\underline{X}_{(2)}^{(1)}$.

The second block B_2 is now given by

$$B_2 = \left(\bigcup_{k=1}^{n_2} L_{k2} \right) \cap \overline{B}_1 \tag{3.15}$$

where $\overline{B}_1$ is the set of points not contained in B_1. In our example B_2 would consist of the union of all the annular areas between the circles of radius $d_1(\underline{X}_{(1)}^{(1)})$ and $d_2(\underline{X}_{(2)}^{(1)})$.

This procedure is continued, and therefore the r^{th} block is given by

$$B_r = \left(\bigcup_{k=1}^{n_2} L_{kr} \right) \bigcap_{i=1}^{r-1} \overline{B}_i \tag{3.16}$$

where

$$L_{kr} = \left\{ \underline{x} : d_{kr}(\underline{x}, \underline{X}_k^{(2)}) \leq d_r(\underline{X}_r^{(1)}) \right\} \qquad k = 1, \ldots, n_2$$

and where

$$d_r(\underline{X}_{(r)}^{(1)}) = \min_{1 \leq k \leq n_2} \min_{\substack{1 \leq i \leq n_1 \\ i \neq (1) \ldots (r-1)}} d_{kr}(\underline{X}_i^{(1)}, \underline{X}_k^{(2)}) .$$

The $(n_1 + 1)^{th}$ block is

$$B_{n_1+1} = \chi - \bigcup_{i=1}^{n_1} B_i = \chi \cap \left(\bigcap_{i=1}^{n_1} \overline{B}_i \right) . \tag{3.17}$$

It is convenient to think of the blocks as being generated by hyperspheres (or other hypervolumes, depending on the form of $d_{kr}(\underline{x}, \underline{X}_k^{(2)})$ <u>expanding</u>

from the centers $\underline{X}_k^{(2)}$, $k = 1, \ldots, n_2$. The expansion continues until the first observation of class 1 is reached; this observation is thereby ordered (i.e. given rank order (1)) and the resulting volume is the first block. Further blocks are generated by further expansions to the remaining $(n_1 - 1)$ P_1 observations.

The region R_2 is the union of the first m blocks formed by this ordering,

$$R_2 = \bigcup_{i=1}^{m} B_i \ .$$

The value for m is obtained from the constraints on the design of the classifier. For example, suppose a classifier is to be designed in which one has 95% confidence that the false alarm probability will be less than 0.05. Then β and γ in equation 3.7b are equal to 0.05 and 0.95, respectively. One of the variables n_1, the number of P_1 training observations, or m, the number of blocks used to construct R_2, is now fixed. The value of the other variable can be found from graphs by Murphy (1948) or from tables by Somerville (1958). For example, we find from Murphy for $\beta = 0.95$, and $\gamma = .95$, and $n_1 = 210$ observation from P_1, that 6 blocks may be used to construct R_2. These numbers give an expected value for the false alarm probability of 0.0284 with a standard deviation of 0.0114. Table 3.1 shows the mean and standard deviation for 3 values of m and n_1 which satisfy the condition $\beta = 0.05$ and $\gamma = 0.95$. As seen from the table, when many observations are available from P_1, the expected false alarm can be higher for the same β and γ than when

few observations are available from P_1.

When a new observation $\underline{V}$ is to be classified, the following rule is used. If $\underline{V}$ falls in R_2, $\underline{V}$ is classified as a member of class 2. Otherwise $\underline{V}$ is classified as a member of class 1.

Table 3.1. Expected Value and Standard Deviation for $\beta = 0.05$ and $\gamma = 0.95$ and for Different m and n.

$Pr\{\int_{R_2} dF_1(x) \leq .05\} = .95$		
	$E\{\int_{R_2} dF_1(x)\}$	$\sigma\{\int_{R_2} dF_1(x)\}$
$n_1 = 430,\ m = 15$	0.0348	0.0088
$n_1 = 210,\ m = 6$	0.0284	0.0114
$n_1 = 58,\ m = 1$	0.0170	0.0167

3.4. Discussion of Practical Ordering Functions

The purpose of this section is to discuss some simple ordering procedures based on the idea of expanding functions from the P_2 observations. The relative merits of these procedures when applied to a problem with a limited sample size are investigated. For simplicity, let the ordering functions be defined as follows:

$$d_{kj}(\underline{x}, \underline{X}_k^{(2)}) = a_{kj} |\underline{x} - \underline{X}_k^{(2)}| \qquad a_{kj} \geq 1 \qquad k = 1, \ldots, n_2 \quad j = 1, \ldots, n_1 \tag{3.18}$$

Then the P_1 observations are ordered by hyperspheres which expand from the P_2 observations.

Distribution-free tolerance regions can be formed by any of the following three ordering procedures:

(1) All Hyperspheres Expand (AHE)

Hyperspheres expand at the same rate from all P_1 observations until n_1+1 blocks have been formed. The first m of these blocks make up region R_2. Since the hyperspheres expand at the same rate, let $a_{kj}= 1$, $k = 1,\ldots,n_2$, $j = 1,\ldots,n_1$. Then the ordering functions are given by

$$d_{kj}(\underline{x}, \underline{X}_k^{(2)}) = |\underline{x} - \underline{X}_k^{(2)}| \qquad \begin{matrix} j = 1,\ldots,n_1 \\ k = 1,\ldots,n_2 \end{matrix} \tag{3.19}$$

The statistically equivalent blocks are described by equations 3.10 through 3.17. This procedure is illustrated in Figure 3.2 for a two-dimensional vector $\underline{x}$, $m = 3$ blocks, $n_2= 6$, $n_1= 33$ with the observations $\underline{X}_1^{(1)},\ldots,\underline{X}_{n_1}^{(1)}$ represented by O's and the observations $\underline{X}_1^{(2)},\ldots,\underline{X}_{n_2}^{(2)}$ represented by X's.

At times the number of blocks with which region R_2 is formed may be small with respect to the number of blocks that are needed for a reasonably low miss rate when using the above ordering procedure. This situation can be a direct result of having few P_2 training observations with at least m P_1 training observations being relatively close to the P_2 observations. For example, notice the two-dimensional example of Figure 3.16. The circular regions surrounding the P_2 observations are

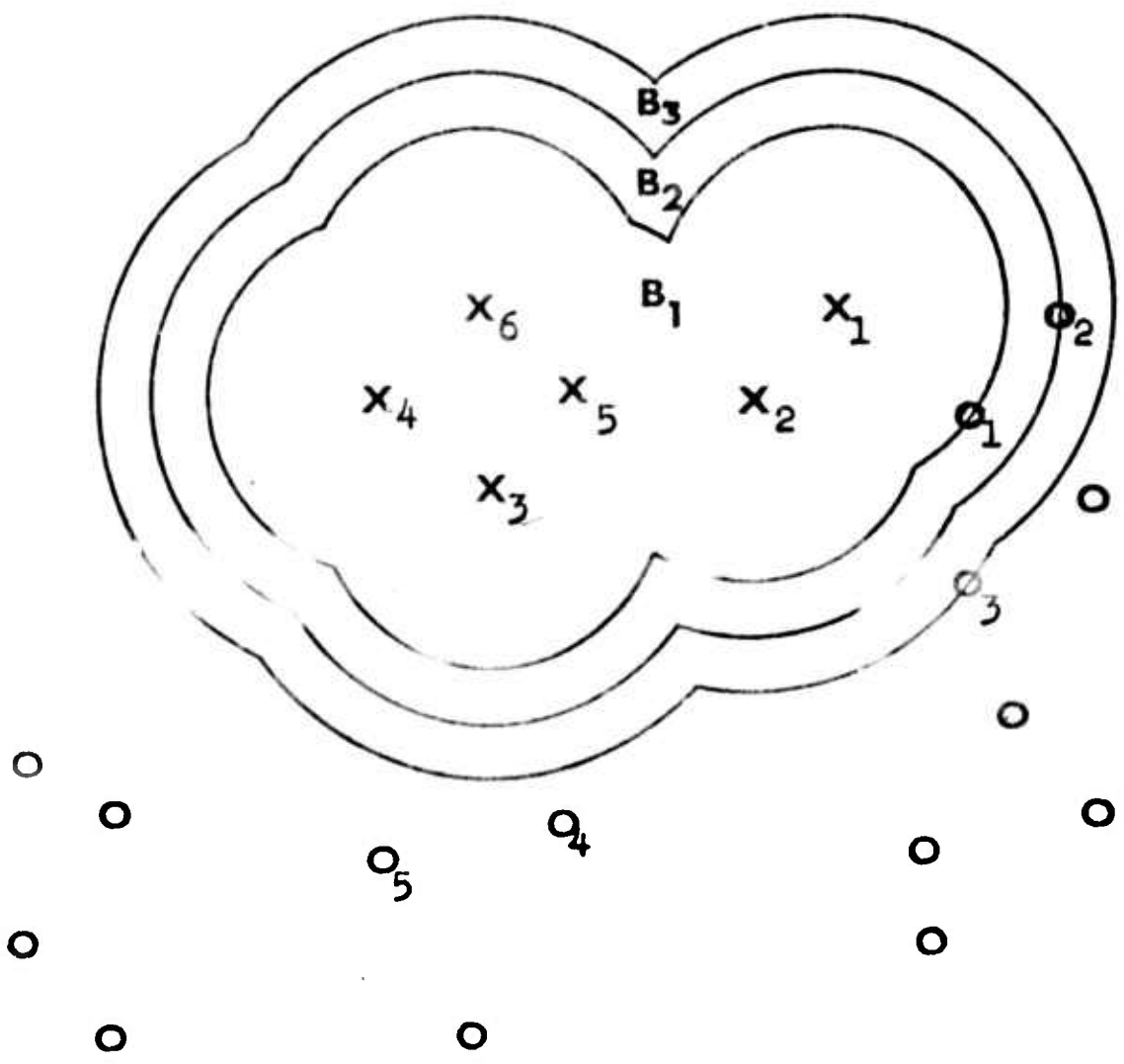


Figure 3.2. All Hyperspheres Expand (AHE).

unconnected. In this case a procedure, which assumes that more P_1 observations may be found in the vicinity of the P_1 observations which have previously been ordered, allows the regions centered by the P_2 observations to expand into regions which are connected and which have a larger volume than the regions produced by the AHE ordering procedure. The following two procedures use the information from previously ordered observations to allow region R_2 to expand faster in directions away from the clustered P_1 observations than toward them.

(2) Ordered Hyperspheres Slowed

With this procedure the hyperspheres which order the P_1 observations are not allowed to expand as rapidly in subsequent orderings as the other hyperspheres. The functions for ordering the first P_1 observation are the same as those for procedure (1). That is,

$$d_{k1}(\underline{x}, \underline{X}_k^{(2)}) = |\underline{x} - \underline{X}_k^{(2)}| \qquad k = 1, \ldots, n_2 \tag{3.20}$$

The functions for ordering the second P_1 observation are given by

$$d_{k2}(\underline{x}, \underline{X}_k^{(2)}) = a_{k2} |\underline{x} - \underline{X}_k^{(2)}| \qquad k = 1, \ldots, n_2 \tag{3.21}$$

where

$$a_{k2} > a_{k1} = 1 \qquad k = (1)$$

$$a_{k2} = a_{k1} = 1 \qquad \text{otherwise .}$$

Note that $k = (1)$ is any k which satisfies

$$d_1(\underline{X}_{(1)}^{(1)}) = a_{k1} |\underline{X}_{(1)}^{(1)} - \underline{X}_k^{(2)}| \ . \tag{3.22}$$

The increase of $a_{(1)1}$ to $a_{(1)2}$ may be viewed as a decrease in the rate at which the ordering hypersphere expands from the P_2 observation, $\underline{X}^{(2)}_{(1)}$, in search of the next P_1 observation.

The functions for ordering the r^{th} P_1 observation are given by

$$d_{kr}(\underline{x}, \underline{X}^{(2)}_k) = a_{kr} |\underline{x} - \underline{X}^{(2)}_k| \tag{3.23}$$

where

$$a_{kr} > a_{k, r-1} \qquad k = (1), \ldots, (r-1)$$

$$a_{kr} = a_{k, r-1} \qquad \text{otherwise .}$$

$k = (j)$ is any k which satisfies

$$d_j(\underline{X}^{(1)}_{(j)}) = a_{kj} |\underline{X}^{(1)}_{(j)} - \underline{X}^{(2)}_k| . \tag{3.24}$$

The adjustment of the multiplicative constant a_{kr}, $k = (1), \ldots, (r-1)$ is quite arbitrary. In the speaker recognition experiment to be discussed in Chapter 4 the increase of $a_{k, r-1}$ to a_{kr} was made very large so that the differences in the three ordering procedures would become evident. Suppose a_{kr} is determined by

$$a_{kr} = (N a_{k, r-1})^N \qquad k = (1), \ldots, (r-1)$$

$$a_{kr} = a_{k, r-1} \qquad \text{otherwise .} \tag{3.25}$$

Suppose N is a large number. This causes the hyperspheres which order the P_1 observations essentially to stop expanding in relation to the hyperspheres which have not ordered a P_1 observation. The ordering procedure for this case will be called Ordered Hyperspheres Constant, OHC.

Suppose all the hyperspheres have ordered a P_1 observation. Then $m > n_2$. In ordering the $(n_2+1)^{th}$ P_1 observation the above procedure causes all hyperspheres to expand at the same rate. Whenever the $(n_2+1)^{th}$ P_1 observation is located, the hypersphere which orders this observation stops expanding in relation to the other hyperspheres. This is because for $m = n_2$ all a_{km} are now large, and therefore the $(n_2+1)^{th}$ a_k is again $(Na_{k,r-1})^N$ larger than the others. Figure 3.3 illustrates the OHC procedure for the same sample set as used in Figure 3.2, where the X's and O's again refer to the P_2 and P_1 observations, respectively.

The first block for the OHC procedure is the same as the first block for the AHE procedure. However, in this example the second block for the OHC procedure differs from the second block for the AHE procedure. This is because the hyperspheres (circles in the figure) expand from all X's except X_1 in search for a new P_1 observation. The observation, which is found is O_3 and it is intersected by the circle expanding from X_2. Then, in forming the third block, circles expand from all X's except X_1 and X_2. Observation O_4 is found and block B_3 is formed.

(3) Conditioned Hyperspheres Slowed

With this procedure the growth of hyperspheres which intersect the P_1 observations is slowed even if these observations have been previously ordered by other hyperspheres.

The ordering functions for $j = 1$ are equivalent to the ordering

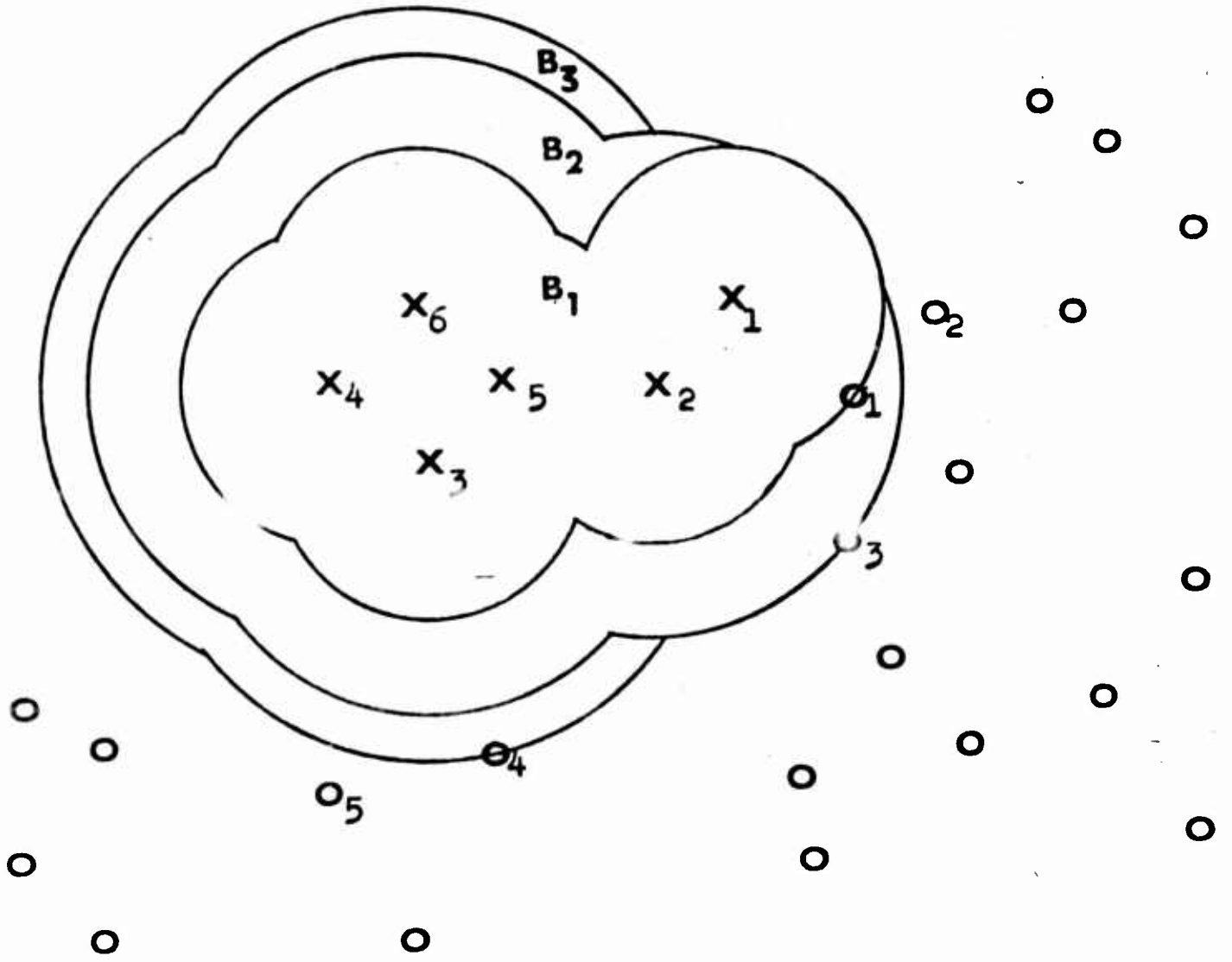


Figure 3.3. Ordered Hyperspheres Constant (OHC).

functions for $j = 1$ for the previous two procedures. The remaining ordering functions can be different from those of the previous two procedures because here an ordering hypersphere is slowed if it comes into contact with $\underline{X}_{(1)}^{(1)}$. Let $\underline{X}_{[1]}^{(1)}$ be the observation for which

$$d_2(\underline{X}_{[1]}^{(1)}) = \min_{1 \le k \le n_2} \min_{1 \le i \le n_1} a_{k2} |\underline{X}_i^{(1)} - \underline{X}_k^{(2)}| \qquad (3.26)$$

where

$$a_{k2} > a_{k1} = 1 \qquad k = (1)$$

$$a_{k2} = a_{k1} = 1 \qquad \text{otherwise}$$

and $k = (1)$ is any k which satisfies equation 3.22. Equation 3.26 can be viewed as telling us that during the second ordering the expanding hyperspheres have intersected a P_1 observation, $\underline{X}_{[1]}^{(1)}$. But this observation might well be the P_1 observation $\underline{X}_{(1)}^{(1)}$ which has already been ordered. For example, in Figure 3.4, the observation which satisfies equation 3.26 is $O_{(1)}$. But this observation has already been ordered. Hence another P_1 observation has to be found before the second block is completed.

Therefore, if $\underline{X}_{[1]}^{(1)} \neq \underline{X}_{(1)}^{(1)}$, denote $\underline{X}_{[1]}^{(1)}$ by $\underline{X}_{(2)}^{(1)}$ and B_2 is given by equations 3.14 and 3.15. If $\underline{X}_{[1]}^{(1)} = \underline{X}_{(1)}^{(1)}$, a block has not been completed. Let

$$G_{k2} = \left\{ \underline{x} : d_{k2}(\underline{x}, \underline{X}_k^{(2)}) \le d_k(X_{[1]}^{(1)}) \right\} . \qquad (3.27)$$

Now let

$$d_{k3}(\underline{x}, \underline{X}_k^{(2)}) = a_{k3} |x - \underline{X}_k^{(2)}| \qquad (3.28)$$

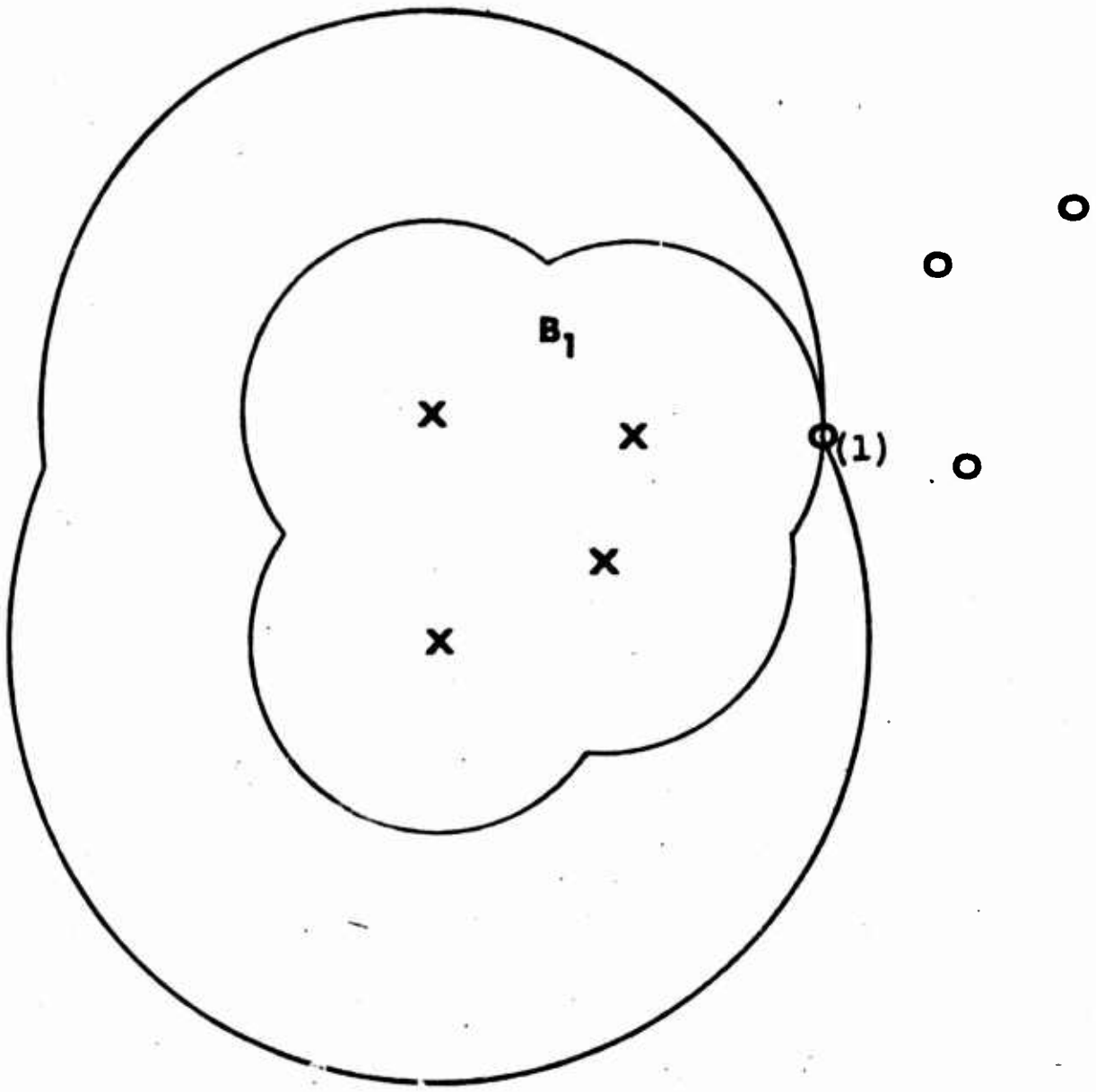


Figure 3.4. All Hyperspheres Stopped by One P_1 Observation.

where

$$a_{k3} > a_{k2} \qquad k = (1), [1]$$

$$a_{k3} = a_{k2} \qquad \text{otherwise} . \tag{3.29}$$

Note $k = [1]$ is any k for which

$$d_2(\underline{X}^{(1)}_{[1]}) = a_{k2} |\underline{X}^{(1)}_{[1]} - \underline{X}^{(2)}_{k}| . \tag{3.30}$$

Now let

$$d_3(\underline{X}^{(1)}_{[1]}) = \min_{1 \le k \le n_2} \min_{1 \le i \le n_1} d_{k3}(\underline{X}^{(1)}_{i}, \underline{X}^{(2)}_{k}) \tag{3.31}$$

and

$$G_{k3} = \left\{ \underline{x} : d_{k3}(\underline{x}, \underline{X}^{(2)}_{k}) \le d_3(\underline{X}^{(1)}_{[2]}) \right\} . \tag{3.32}$$

This procedure is continued until $\underline{X}^{(1)}_{[j]} \neq \underline{X}^{(1)}_{(1)}$ or $(j+1) \ge n_2$. If $\underline{X}^{(1)}_{[1]} \neq \underline{X}^{(1)}_{(1)}$ and $(j+1) < n_2$, let $\underline{X}^{(1)}_{[j]} = \underline{X}^{(1)}_{(2)}$. The second block is

$$B_2 = \left(\bigcup_{r=1}^{j} \bigcup_{k=1}^{n_2} G_{k, r+1} \right) \cap \bar{B}_1 . \tag{3.33}$$

If $j+1 \ge n_2$ and $\underline{X}^{(1)}_{[i]} = \underline{X}^{(1)}_{(1)}$, $i = 1, \ldots, n_2$, let

$$d_{n_2}(\underline{X}^{(1)}_{(2)}) = \min_{1 \le k \le n_2} \min_{\substack{1 \le i \le n_1 \\ i \neq (1)}} d_{kn_2}(\underline{X}^{(1)}_{i}, \underline{X}^{(2)}_{k}) \tag{3.34}$$

and

$$G_{kn_2} = \left\{ \underline{x} : d_{kn_2}(\underline{x}, \underline{X}^{(2)}_{k}) \le d_{n_2}(\underline{X}^{(2)}_{(2)}) \right\}$$

Then

$$B_2 = \left(\bigcup_{r=1}^{n_2} \bigcup_{k=1}^{n_2} G_{k, r+1} \right) \cap \bar{B}_1 . \tag{3.35}$$

Note that the restriction, $(j+1) \geq n_2$, is necessary to eliminate the possibility that the procedure enters an infinite loop. The condition $(j+1) \geq n_2$ was chosen especially for the case where the hyperspheres stop when they intersect a P_1 observation (equation 3.25, where N is a very large number). In this case all hyperspheres are allowed to expand until they intersect the P_1 observation which normally would cause the procedure to enter an infinite loop. Then they are allowed to expand past this observation. For example, consider Figure 3.4. The observation $O_{(1)}$ will stop every expanding hypersphere. The condition $(j+1) \geq n_2$ allows all four hyperspheres to expand to $O_{(1)}$ and then to expand past $O_{(1)}$ to form the second block.

The extension of this procedure for the formation of m blocks is straightforward. Figure 3.5 illustrates the procedure for the same sample set as used in Figures 3.2 and 3.3 and for N equal to a very large number and

$$a_{kr} = (N a_{k,\,r-1})^N \qquad k = (1),\ [1],\ [2], \ldots, [j]$$

$$a_{kr} = a_{k,\,r-1} \qquad \text{otherwise.} \qquad (3.36)$$

This ordering is called CHS, Conditioned Hyperspheres Stop. In this example the second block for the CHS procedure differs from the second block for the OHC procedure. In forming the second block the hyperspheres are expanding from all X's except X_1. When the hypersphere expanding from X_2 intersects O_1, it stops in the CHS procedure, even though a block has not been completed. Hyperspheres continue to expand from

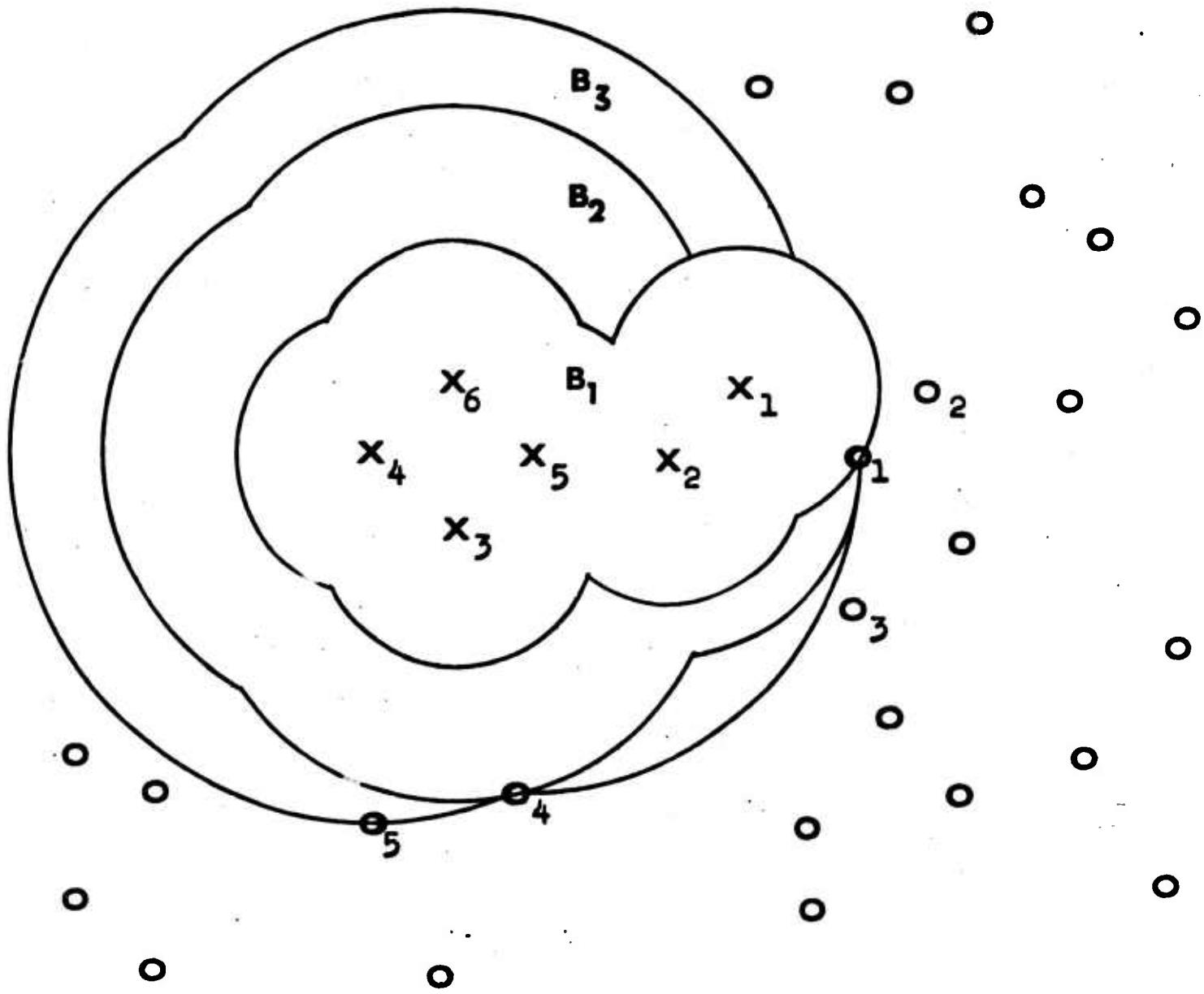


Figure 3.5. Conditioned Hyperspheres Stop (CHS).

observations X_3, X_4, X_5, and X_6 in search of a new O observation. The hypersphere which expands from X_3 intersects O_4 and the second block is complete. Then hyperspheres expand from X_4, X_5, and X_6 in search of a new observation. The procedure is continued in this manner.

The resulting region R_2 (for m = 3 blocks) is shown in Figure 3.6 for each of the three procedures. In this particular example it is seen that the AHE procedure produces a region R_2 which has expanded into the O's whereas the CHS procedure produces a region R_2 which has been stopped by the O's and has expanded in a direction away from the ordered O's.

It is evident that these three procedures are not the only procedures that can be formulated when hyperspheres expand from the P_2 observations. For example, one might decide to slow the expansion of any ordering hypersphere which is in the vicinity of an ordered P_1 observation. However, this procedure would bias the estimate of the miss probability, which is discussed in section 3.9.

A comparison of the three ordering procedures requires iteration of all possible sample sets. Nevertheless, some general observations can be made.

1) The AHE procedure is probably preferable to the OHC procedure, which is probably preferable to the CHS procedure, if spurious P_1 observations are involved. This is because hyperspheres in the AHE procedure, and in the OHC procedure to a lesser extent, continue to expand

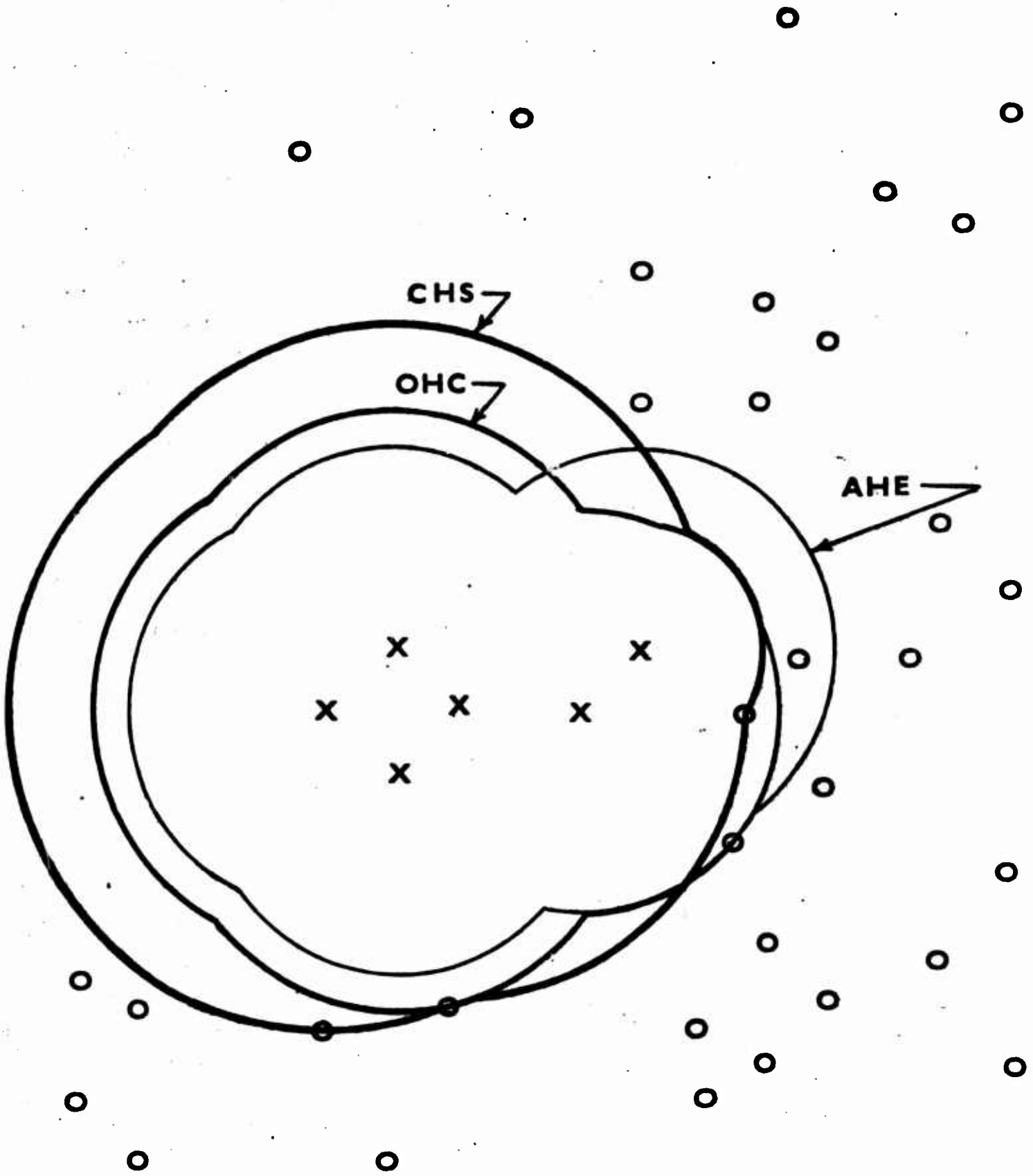


Figure 3.6. Composite of the AHE, OHC, and CHS Ordering Procedures.

past the peripheral P_1 observations. For example, consider Figure 3.7 where two blocks make up R_2. The resulting R_2 for the AHE and the CHS procedures is shown in the figure. Notice that the hyperspheres which expand from X_1 and X_2 stop when they intersect $O_{(1)}$ in the CHS procedure. They, of course, do not stop in the AHE procedure.

The area of R_2 for the AHE procedure is equal to the area of R_2 for the CHS procedure plus the crosshatched area. Hence the miss probability in this case is less for the AHE procedure than for the CHS procedure.

2) The CHS procedure is probably preferable to the OHC procedure which is probably preferable to the AHE procedure if the P_1 observations are tightly clustered in two or more clusters and the clusters are different distances from the P_2 observations. This is because the hyperspheres in the CHS procedure, and in the OHC procedure to a lesser extent, expand more in directions away from the ordered P_1 observations than do the hyperspheres of the AHE procedure. For example, consider Figure 3.8 where two blocks are formed with the AHE and the CHS ordering procedures. The area of the CHS procedure is equal to the area of the AHE procedure plus the crosshatched region minus the shaded area. Since the crosshatched area is much larger than the shaded area, one may feel that the miss probability in this case is less for the CHS procedure than for the AHE procedure.

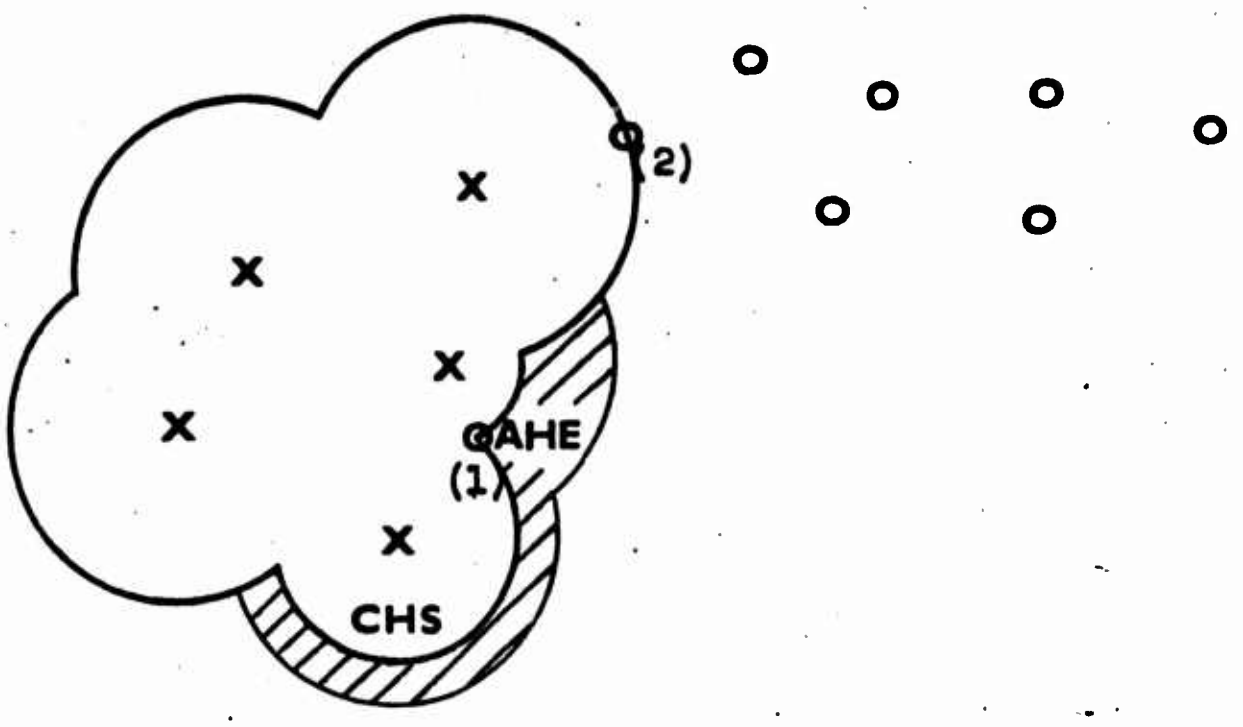


Figure 3.7. A Comparison in Favor of the AHE Procedure.

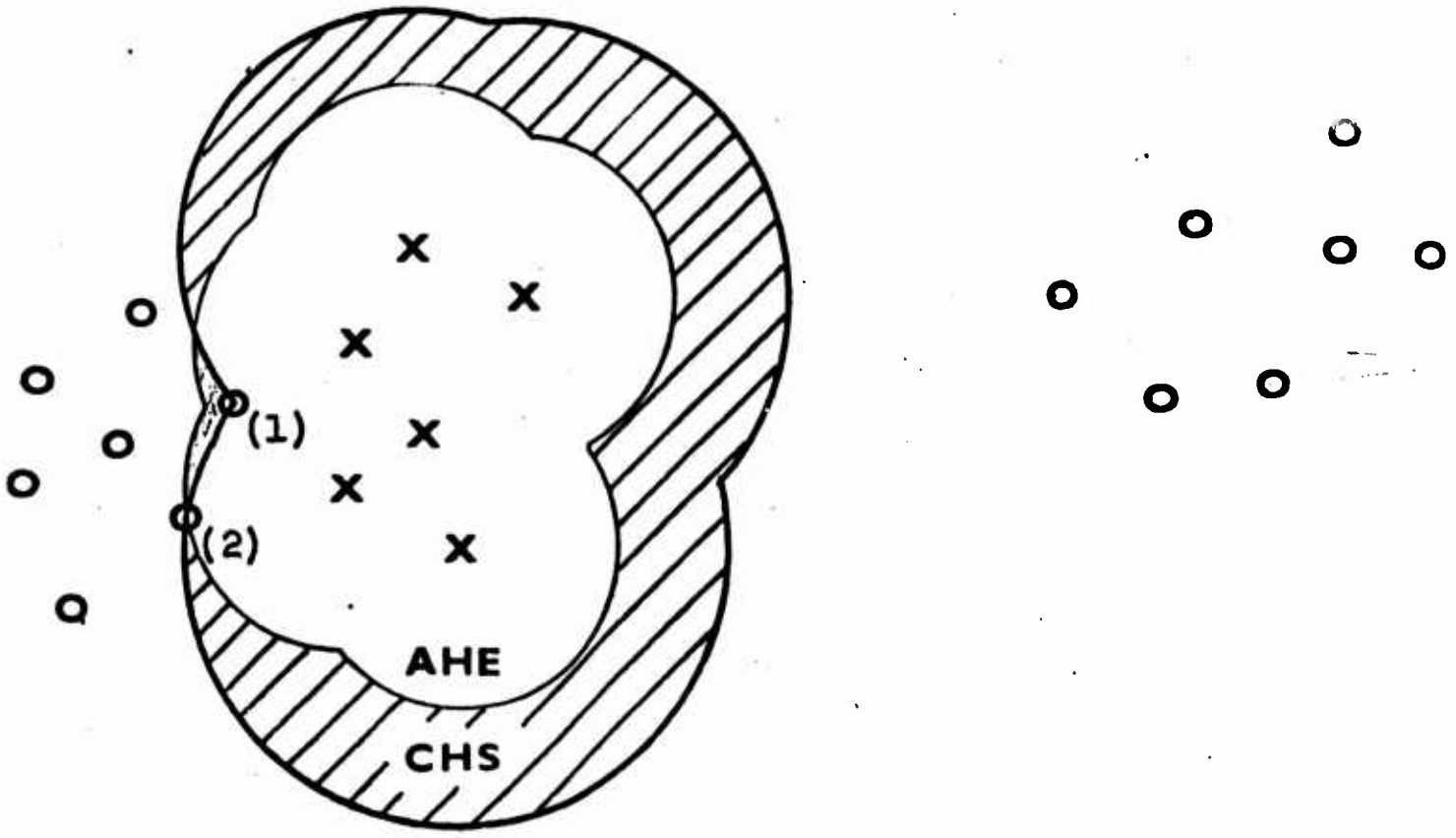


Figure 3.8. A Comparison in Favor of the CHS Procedure.

3.5. Programming on a Digital Computer

Note that these three procedures are very easily programmed on a digital computer. One simply calculates the distance between every P_1 and P_2 training observation. Let the distance between $\underline{X}_i^{(1)}$ and $\underline{X}_j^{(2)}$ be denoted by D_{ij}. These distances are arranged in a matrix D, where the i^{th} row consists of the distances between $\underline{X}_i^{(1)}$ and $\underline{X}_j^{(2)}$, $j = 1, \ldots, n_2$.

Consider first the AHE procedure. A search is made through the elements of the matrix for the smallest distance. Let this distance be D_{kr}. Then the k^{th} P_1 observation is ordered and a block is formed. The k^{th} row is multiplied by the largest number available on the machine. This removes the k^{th} P_1 observation from further ordering. A search is now made through the elements of the new matrix for the smallest distance. This procedure is continued until m blocks are formed.

In the ordered hyperspheres slowed procedure a search is made through the elements of the matrix D for the smallest distance as before. Let this distance be D_{kr}. The k^{th} row is then multiplied by the largest number available on the machine. Thus far, the two procedures are the same. Now the r^{th} column is multiplied by a number which controls the rate at which the hypersphere expands from the r^{th} P_2 observation. A search is made through the elements of the new matrix for the smallest distance. This procedure is continued with the columns and the rows corresponding to the smallest distance being multiplied by the appropriate numbers after each block is formed.

Now consider the conditioned hypersphere slowed procedure. As in the previous two procedures a search is made for the smallest distance. Let this distance be D_{kr}. The r^{th} column is now multiplied by a number which controls the rate at which the hypersphere expands from the r^{th} P_2 observation. Unlike the previous two procedures, the k^{th} column remains untouched. A search is made through the elements of the new matrix for the smallest distance. Let this distance be D_{st}. If $s = k$, a block has not been formed. In any case the t^{th} column is multiplied by a number which controls the rate at which the hypersphere expands from the t^{th} observation. This procedure is continued until a smallest distance D_{gh} is found so that $g \neq k$ or until a restriction on the number of iterations is reached ($(j+1) \geq n_2$ in the above discussion). If a smallest distance D_{gh} is found so that $g \neq k$, the second block is formed. The h^{th} column is multiplied by a number which controls the rate at which the hypersphere expands from the h^{th} P_2 observation and the procedure is continued. If the restriction on the number of iterations is reached, the r^{th} row is multiplied by the largest number available on the machine. A search is then made through the elements of the new matrix for the smallest distance. When this distance is found, the second block is formed. The procedure is continued in this fashion until m blocks are formed.

Note that when a new observation is to be classified, the following information must be stored in the computer for the various procedures.

1) AHE – all P_2 observations and the value of the m^{th} order statistic.

2) OHC – all P_2 observations and the value of the m order statistics along with the indices of the P_2 observations corresponding to the m order statistics.

3) CHS – all P_2 observations and the values of the distances along with the corresponding indices of the P_2 observations found in the ordering process. This procedure requires at least as much storage as the OHC procedure.

It is obvious that the information to be stored can be reduced further by clustering the P_2 observations and using representative points for the clusters (for example the means of the clusters) to order the P_1 observations. However, this approach biases the estimate of the miss probability as seen in section 3.9.

On the other hand, in a situation where too few P_2 observations are available, (see Fig. 3.16) one can sometimes cause the regions of R_2 to be connected by adding "fictitious" P_2 observations between the P_2 observations whose nearest P_2 observation is furthest away. Hyperspheres expand from these "fictitious" P_2 observations in the same manner as they did from the "real" P_2 observations.

3.6. Large Sample Properties

Three methods were proposed in section 3.4 for the classification of observations from two different classes. The large sample properties

of these methods are investigated in this section. The goal for the nonparametric method is the emulation of the Neyman-Pearson rule, which was stated in equations 3.3 and 3.4.

Let n_1, the number of observations from P_1, and n_2, the number of observations from P_2, approach infinity such that n_1/n_2 is bounded away from zero and infinity. Let the classifier be designed for

$$E\left\{\int_{R_2} dF_1(\underline{x})\right\} = \frac{m}{n_1+1} = \alpha$$

where $m \leq n_1+1$. Hence m approaches infinity while $m/(n_1+1) = \alpha$.

The false alarm probability converges in probability to the desired value α as n_1 approaches infinity, i.e.

$$\int_{R_2} dF_1(\underline{x}) \longrightarrow \alpha .$$

This follows directly from the Tchebycheff inequality since the variance of $\int_{R_2} dF_1(\underline{x})$, considered as a random variable, approaches zero as n_1 approaches infinity.

We now wish to determine the outcome of $\int_{R_2} dF_2(\underline{x})$ as n_1 and n_2 approach infinity. For simplicity, suppose that the probability densities $f_1(\underline{x})$ and $f_2(\underline{x})$ are continuous, univariate, unimodal, and nonzero everywhere. Furthermore, let $n_1 = n_2 = n$. Consider the DFTR method called All Hyperspheres Expand (AHE).

Figure 3.9 gives a general picture of the situation to be discussed. As in the previous examples, the P_2 observations are represented as

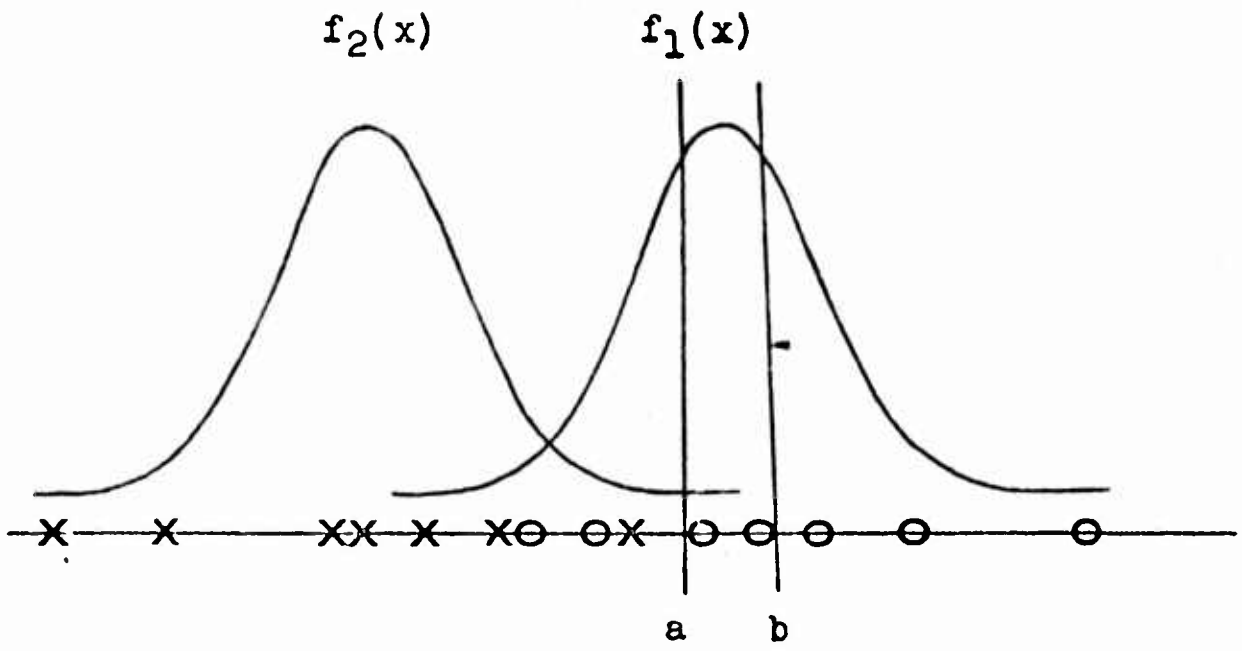


Figure 3.9. A One-Dimensional Example.

Length

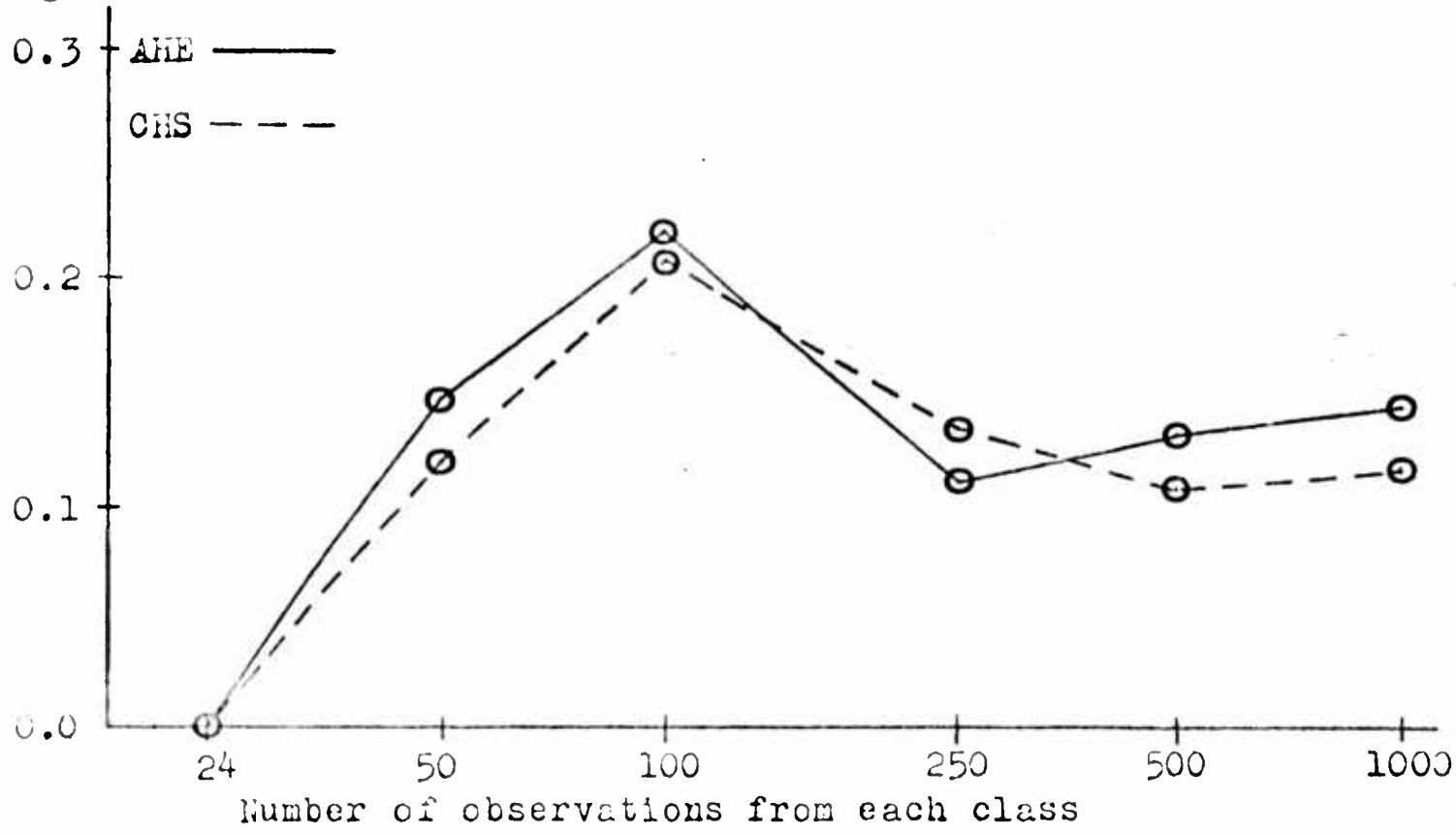


Figure 3.10. Length of the Accept Region on the Positive Line.

X's and the P_1 observations as O's .

As n becomes large, the ratio of P_1 observations to P_2 observations in any small region R approaches the average likelihood ratio existing in this region, $\langle f_1(x)/f_2(x) \rangle_R$. Since $f_1(x)$ and $f_2(x)$ were assumed to be continuous, it can be assumed that as R becomes very small $\langle f_1(x)/f_2(x) \rangle$ is approximately constant in R .

Consider a small interval [a,b] under the peak of $f_1(x)$. Suppose that α and $f_1(x)$ are such that in this region

$$\alpha << \int_a^b f_1(x)\,dx . \tag{3.37}$$

The number of P_1 observations (O's in Fig. 3.9) in [a,b] approaches $n\int_a^b f_1(x)\,dx$ as $n \to \infty$, by Tchebycheff's theorem. Since the number of blocks "m" is approximately equal to αn , equation 3.37 implies that

$$m << n\int_a^b f_1(x)\,dx = \text{number of O's in } [a,b] .$$

By the assumption that $f_2(x)$ is nonzero everywhere, there are X's in the interval [a,b] with probability 1 , (as $n \to \infty$) . Suppose that in the interval [a,b]

$$f_2(x)/f_1(x) << \alpha .$$

This means that $mf_1(x) >> nf_2(x)$. Therefore on the average each X is surrounded by many more than m O's. Then the region $R_2 = \bigcup_{i=1}^{m} B_i$ consists largely of short, unconnected intervals centered on the X's. We now wish to determine the average length of these intervals. The average number of O's in [a,b] is $n\int_a^b f_1(x)\,dx$. Therefore

$(b-a)/n\int_a^b f_1(x)\,dx$ is the average distance between the O's. Since each X is surrounded by many more than m O's, the length of each interval can be no more than $\frac{m(b-a)}{n\int_a^b f_1(x)\,dx}$. The length of each interval is in fact equal to $\frac{m(b-a)}{n\int_a^b f_1(x)\,dx}$ if the m closest O's to an X are as likely to be in [a,b] as in all other regions. In this case as $n \to \infty$ the length of each interval in [a,b] is $\alpha(b-a)/\int_a^b f_1(x)\,dx$. Since this is a finite number, there would be nonzero sections in areas where $f_1(x)/f_2(x)$ is very small. This is not the case when R_2 is determined by $f_1(x)/f_2(x) > C$ where C is a threshold.

Note that all assumptions are such as to minimize the extent of region R_2 in places where $f_1(x)$ is large. Thus if these assumptions are removed, the result holds a fortiori. For example, consider the assumption

$$f_2(x)/f_1(x) << \alpha$$

in the interval [a,b]. If this assumption is not made, we cannot say that each X in [a,b] is a nucleus for a small section of region R_2. In fact, several blocks in [a,b] may coalesce into connected intervals. However, this only increases the extent of region R_2 in areas that would be excluded by a likelihood ratio test.

The fact that the AHE-DFTR test does not, in general, approach the likelihood ratio test can also be demonstrated as follows. When the AHE procedure is used, the lengths of the intervals surrounding all

P_2 observations are equal. As $n \to \infty$ the number of P_2 observations in a small interval of length e around the maximum of $f_2(x)$ is greater than the number of P_2 observations in a small interval of length e around any other point of $f_2(x)$. Hence if the intervals surrounding the P_2 observations coalesce, they would most likely coalesce in the region where $f_2(x)$ is a maximum. Using the likelihood ratio procedure, the smallest accept region is in the vicinity of the maximum of $f_2(x)/f_1(x)$. The maximum of $f_2(x)$ and the maximum of $f_2(x)/f_1(x)$ do not necessarily occur at the same point. Hence, in general, the AHE procedure does not approach the likelihood ratio procedure as n approaches infinity.

It is not known at this time how to determine the large sample properties of the OHC or the CHS procedure. Therefore, an example was simulated on the computer. The probability densities were arbitrarily chosen to be

$$f_1(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x - 1.75)^2\right\}$$

and

$$f_2(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x + 1.75)^2\right\}$$

Equal a priori probabilities and equal costs of misrecognition were assumed. The optimum decision in a Bayes sense is a decision in favor of P_1 if a new observation has a value which is less than zero. This yields a false alarm rate, $P_{FA} = 0.0401$. The length of the accept region (R_2) to the positive side of zero and the length of the accept region to the negative side of zero were then found for samples of 24, 49, 99,

250, 500, and 1000 observations from each class for both the AHE and the CHS procedures. The value of m was chosen so that

$$E \{ \int_{R_2} dF_1(x) \} \sim 0.0401 .$$

Therefore values of m of 1, 2, 4, 10, 20, and 40 were used for the samples of 24, 49, 99, 250, 500, and 1000, respectively.

The accept region R_2 for the likelihood ratio criterion is the negative real line. If the DFTR procedures are to approach the likelihood ratio procedure as $n \to \infty$, the length of the accept region on the positive real line should approach zero. The length of the accept region on the positive real line for both the AHE and CHS DFTR procedures is shown in Figure 3.10. The results are not definitive since they are based on one trial. However, they indicate that the length of the accept region on the positive real line approaches some value other than zero for both DFTR procedures.

The length of the accept region on the negative real line is shown in Figure 3.11 for both procedures. If the DFTR tests approach the likelihood ratio test, these curves should continually increase as $n \to \infty$. Again the results are not definitive. However, it appears that the curves approach some finite value rather than infinity.

Both Figures 3.10 and 3.11 indicate that the CHS procedure performs better than the AHE procedure for this particular situation. The length of the accept region on the positive real line for the CHS procedure is less than or equal to the corresponding length for the AHE

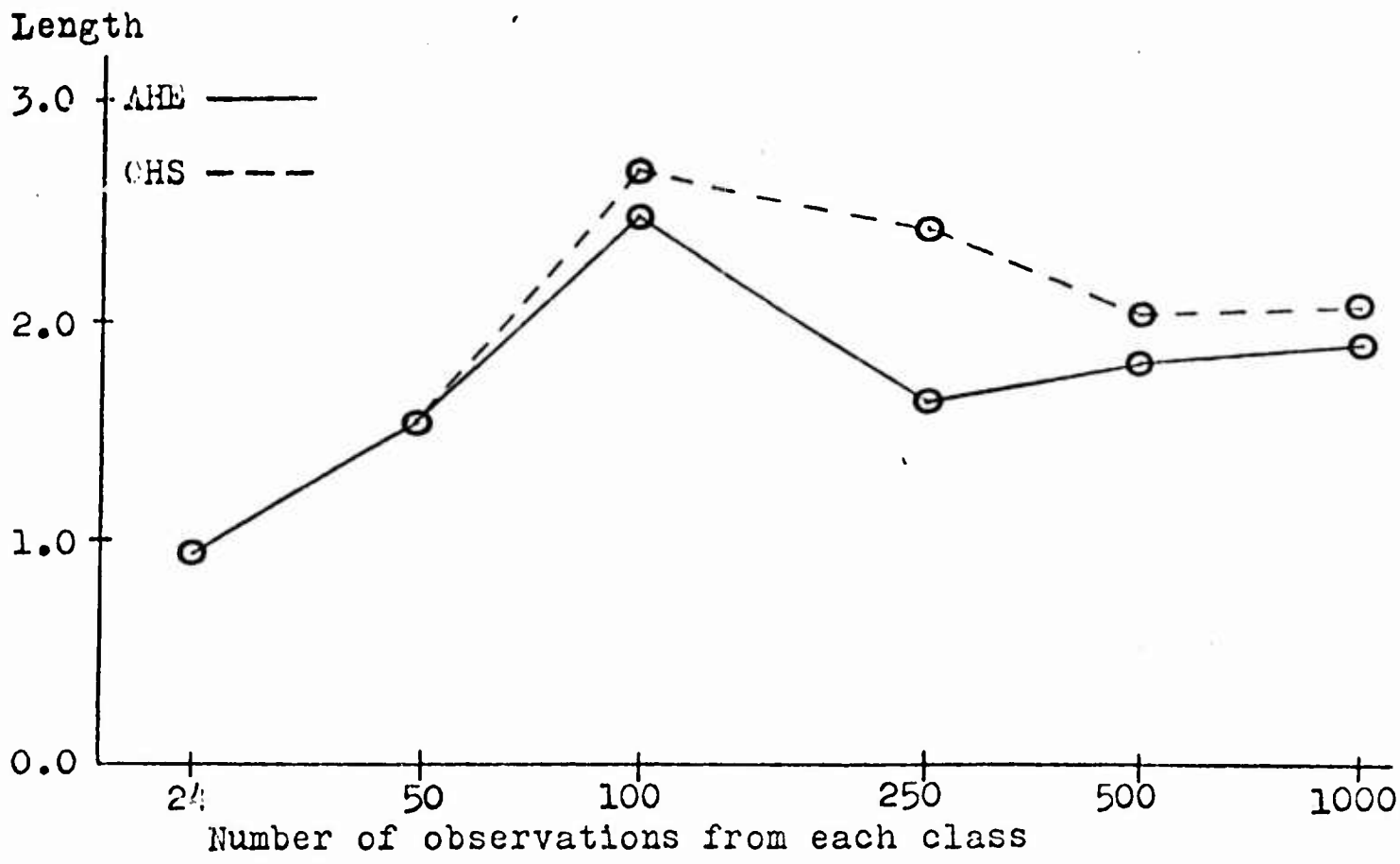


Figure 3.11. Length of the Accept Region on the Negative Line.

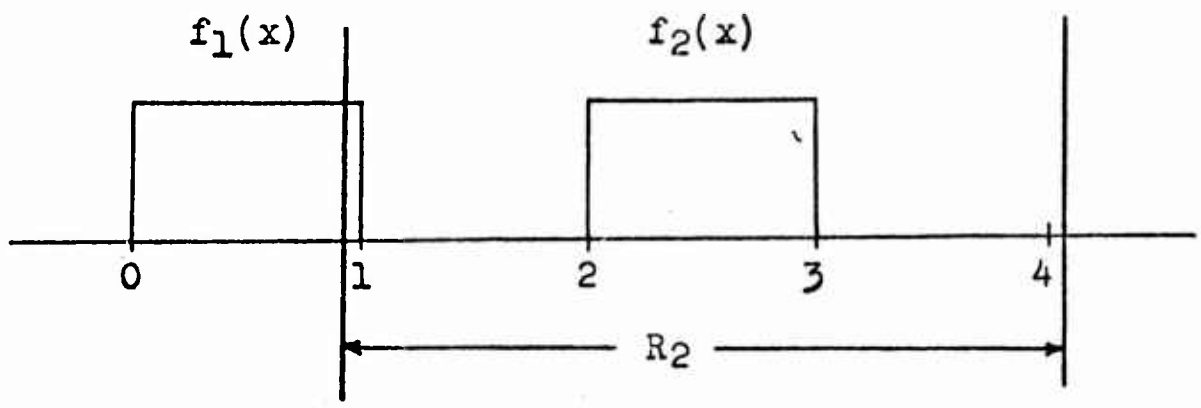


Figure 3.12. Disjoint Probability Density Functions.

procedure for all samples except n = 250. The length of the accept region on the negative real line for the CHS procedure is greater than or equal to the corresponding length for the AHE procedure for all sample sizes.

The ratio of the length of the accept region on the negative real line to the length of the accept region on the positive real line gives some indication of the relative performance of these procedures. These ratios labeled R_{AHE} for the AHE procedure and R_{CHS} for the CHS procedure are shown in Table 3.2 for the various sample sizes.

Table 3.2. A Comparison of the AHE and CHS Procedures

Number of Training Observations	R_{AHE}	R_{CHS}
49	11.0	13.8
99	11.4	12.9
250	15.4	19.2
500	14.2	19.1
1000	14.0	19.0

One sees that these ratios are not steadily increasing with sample size as they should if the DFTR procedures are to approach a likelihood ratio procedure. Note that R_{CHS} is greater than R_{AHE} for all sample sizes.

3.7. An Optimum Ordering Procedure

An ordering which does approach the Neyman-Pearson procedure is easily obtained if the class probability distributions are known. This is achieved by simply using ordering functions equal to $f_1(x)/f_2(x)$. That is, the following ordering functions are used:

$$h_1(x) = h_2(x) = \ldots = h_{n_1}(x) = f_1(x)/f_2(x) \; . \tag{3.38}$$

The fact that a DFTR procedure which uses the above ordering functions approaches a Neyman-Pearson test is easily demonstrated. When the Neyman-Pearson criterion is used, the accept region is

$$R_2 = \{ x : f_1(x)/f_2(x) < C \}$$

where C is determined so that

$$\int_{R_2} dF_1(x) = \alpha \; .$$

Let R_2' be the region obtained when statistically equivalent blocks are formed by the likelihood ratio ordering functions in equation 3.38 . By the Tchebycheff inequality

$$\int_{R_2'} dF_1(x) \xrightarrow[p]{} \alpha$$

where

$$R_2' = \bigcup_{i=1}^{m} B_i \; ,$$

m is the largest integer satisfying

$$m \leq (n_1 + 1) \; ,$$

and the blocks are given by

$$B_1 = \left\{x : h_1(x) \leq \min_{1 \leq i \leq n_1} h_1(x_i)\right\}$$

$$\vdots$$

$$B_m = \left\{x : h_m(x) \leq \min_{\substack{1 \leq i \leq n \\ i \neq (1) \ldots (m-1)}} h_m(x_i)\right\} \bigcap_{i=1}^{m-1} B_i \ .$$

Let $\min_{\substack{1 \leq i \leq n_1 \\ i \neq (1) \ldots (m-1)}} h_m(x_i) = C_m$.

Since $h_1(x) = h_2(x) = \ldots = h_m(x)$,

$$R_2 = \{ x : h_m(x) \leq C_m \} = \{ x : f_1(x)/f_2(x) \leq C_m \} \ .$$

It now remains for us to show that $C_m \xrightarrow{p} C$. But this has to be true since

$$\int_{\{ x : f_1(x)/f_2(x) \leq C_m \}} dF_1(x) \xrightarrow{p} \int_{\{ x : f_1(x)/f_2(x) \leq C \}} dF_1(x)$$

and $f_1(x) > 0$ and $f_2(x) > 0$ for all x . Therefore the DFTR procedure with the likelihood ratio ordering functions approaches a Neyman-Pearson test. For example, let

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2} \quad \text{and} \quad f_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x+\mu)^2}$$

where $\mu > 0$.

Then

$$f_1(x)/f_2(x) = e^{2\mu x} \ .$$

Using $e^{2\mu x}$ as an ordering function, the blocks are

$$B_1 = (-\infty, X_{(1)}]$$

$$B_2 = (X_{(1)}, X_{(2)}]$$

$$\vdots$$

$$B_m = (X_{(m-1)}, X_{(m)}]$$

Therefore $R_2' = (-\infty, X_{(m)}]$.

The accept region to satisfy the Neyman-Pearson criterion is $(-\infty, z)$ where

$$\int_{-\infty}^{z} f_1(x)\,dx = F_1(z) = \alpha .$$

By Tchebycheff's inequality

$$\int_{-\infty}^{X_{(m)}} f_1(x)\,dx = F_1(X_{(m)}) \xrightarrow[p]{} \alpha .$$

Since $f_1(x) > 0$ for all x, ($F_1(x)$ is a monotone increasing function) $X_{(m)} \xrightarrow[p]{} z$. Therefore the accept region for the DFTR procedure using the likelihood ratio ordering functions converges in probability to the accept region for the likelihood ratio procedure.

It should be noted that the AHE, OHC, and CHS tests approach Neyman-Pearson tests in the limit if the class probability densities are disjoint. That is,

$$\int_{R_2} dF_1(x) \xrightarrow[p]{} \alpha \tag{3.39}$$

and

$$\int_{R_2} dF_2(x) = 1 \tag{3.40}$$

when $$\int_{\text{all } x} f_1(x)\, f_2(x)\, dx = 0 .$$

For example, consider the configuration of Figure 3.12. $f_1(x)$ is uniform over $[0,1]$ and $f_2(x)$ is uniform over $[2,3]$. The region R_2 produced by the AHE ordering procedure as n_1 and n_2 approach infinity is as shown in the figure. It is easily seen that conditions (3.39) and (3.40) are satisfied here.

3.8. A Further Comparison of the AHE and CHS Procedures

A further comparison of the AHE and CHS procedures can be obtained by observing the average length of an interval which surrounds a P_2 observation as the number of blocks used to form the tolerance region varies. This is an appropriate comparison because we believe that the additional volume (length) of the accept region that the CHS procedure produces over the AHE procedure, if any, is located so that the probability of correct detection increases and so that the probability of false alarm stays approximately constant (for large n). In fact, the comparison of the length of the accept region on the positive and negative real line in Figures 3.10 and 3.11 seems to demonstrate this.

The average length of an interval surrounding a P_2 observation versus the number of blocks used to form R_2 is shown in Figures 3.13, 3.14, and 3.15 for samples of 250, 500, and 1000 observations. For example, let us consider Figure 3.13. For 10 blocks used to form R_2 the length of an interval surrounding a P_2 observation for the AHE

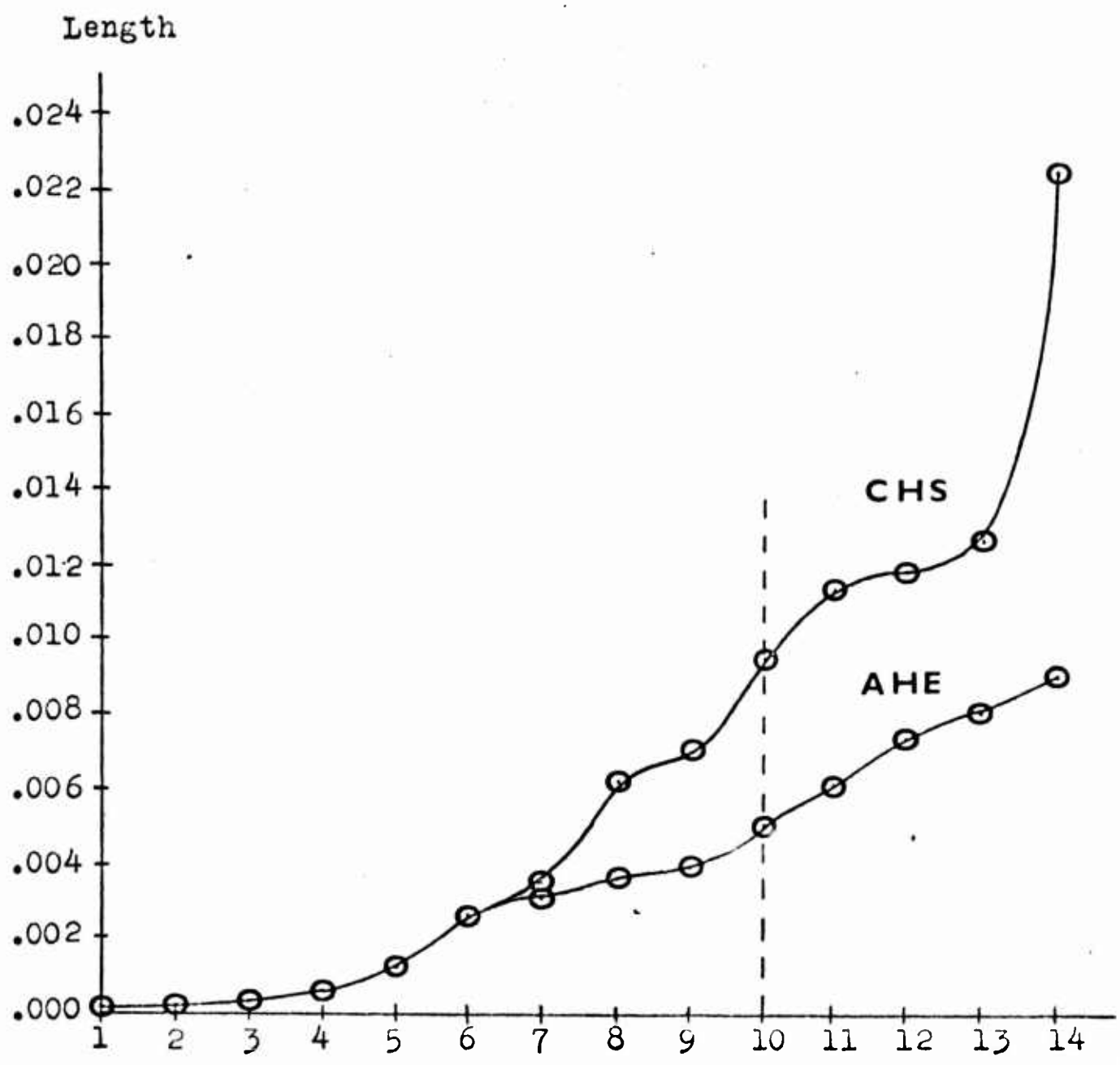


Figure 3.13. Average Length of an Interval Surrounding a P_2 Observation. (250 Observations from Each Class)

procedure is .00495. For the CHS procedure, 241 of the intervals have a length of .0095. The other 9 intervals vary in length from .0002 to .0069. This gives an average interval length of .00925.

As seen in Figure 3.13 no benefit is obtained in using the CHS procedure over the AHE procedure for $m \leq 6$ blocks in R_2. As more blocks are added, the average length of an interval surrounding a P_2 observation for the CHS procedure becomes larger than the length of an interval for the AHE procedure. As still more blocks are added, a point is reached where the average length of an interval for the CHS procedure becomes much larger than the length of an interval for the AHE procedure. This is the point at which most of the P_1 observations in the areas where the P_1 and P_2 observations are highly confused have been ordered.

Similar curves for samples of 500 and 1000 observations from each class are shown in Figures 3.14 and 3.15, respectively. These curves seem to indicate that a good deal of the benefit of the CHS procedure over the AHE procedure had not been revealed for $m/(n+1) = 0.0401$ ($m = 20$ in the 500 sample experiment and $m = 40$ in the 1000 sample experiment). This value of $m/(n+1)$ was, of course, used to obtain the results of Figures 3.10 and 3.11. If more blocks had been used to obtain these figures, it is likely that a larger improvement in the CHS performance over the AHE performance would be noted for $n = 500$ and $n = 1000$.

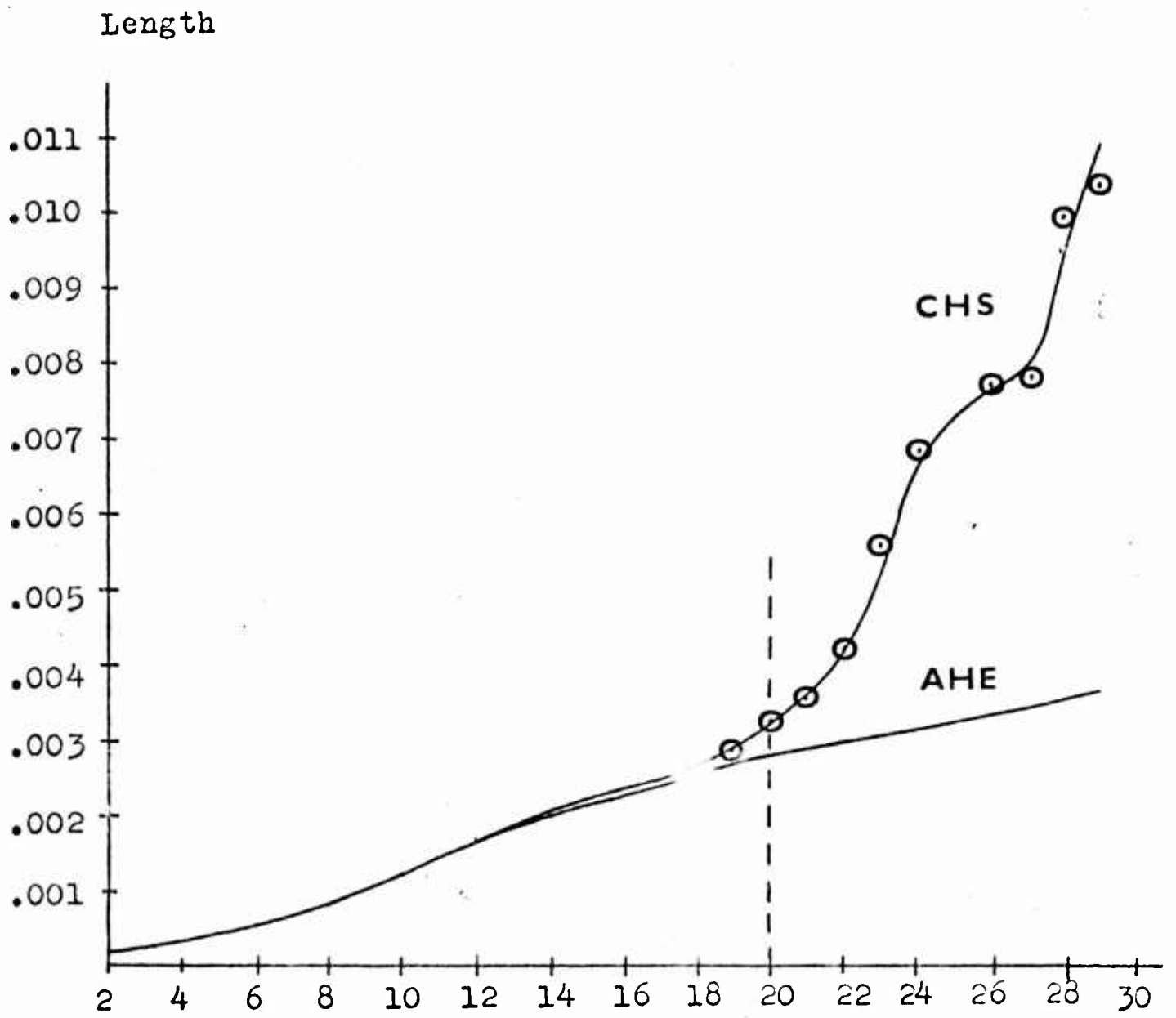


Figure 3.14. Average Length of an Interval Surrounding a P_2 Observation. (500 Observations from Each Class)

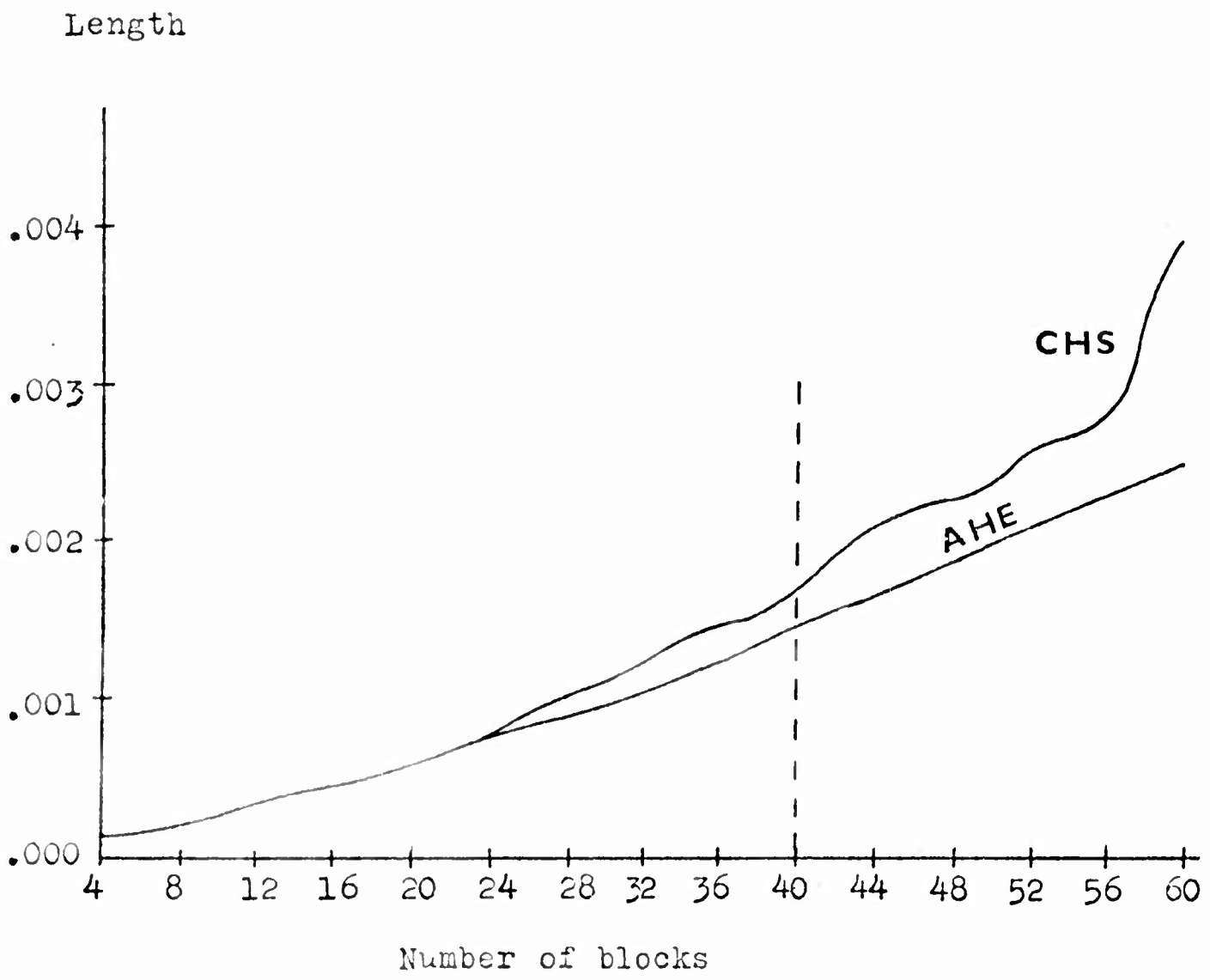


Figure 3.15. Average Length of an Interval Surrounding a P_2 Observation. (1000 Observations from Each Class)

3.9. Measure of the Miss Probability

A recognition system has been proposed which classifies with a given expected false alarm probability (or with a given confidence that the false alarm probability is less than a given quantity). It also correctly classifies all P_2 training observations. Nevertheless, one can find situations in which the classifier may perform poorly. For example, consider the two-dimensional measurement space of Figure 3.16. A classifier is designed using the AHE procedure for an expected false alarm probability of 0.20. Region R_2 consists of two blocks and is the region inside the three circles which are centered by the X's. Since none of the regions encircling the X's are connected, one feels that the miss probability could be quite large.

If this classifier is to be used in a practical problem, a measure of the expected miss error is needed. Then if the expected miss error is much larger than desired, one can redesign the classifier by using more P_1 observations, by using more P_2 observations, or by using a different measurement space.

Suppose a classifier has been designed by one of the methods previously discussed. The P_1 observations have been ordered with hyperspheres which expand from each of the P_2 observations. Now suppose the P_2 observations are ordered, thus forming blocks with respect to $F_2(\underline{x})$. The number of $F_2(\underline{x})$ blocks which are contained in region R_2 can be counted and statements such as

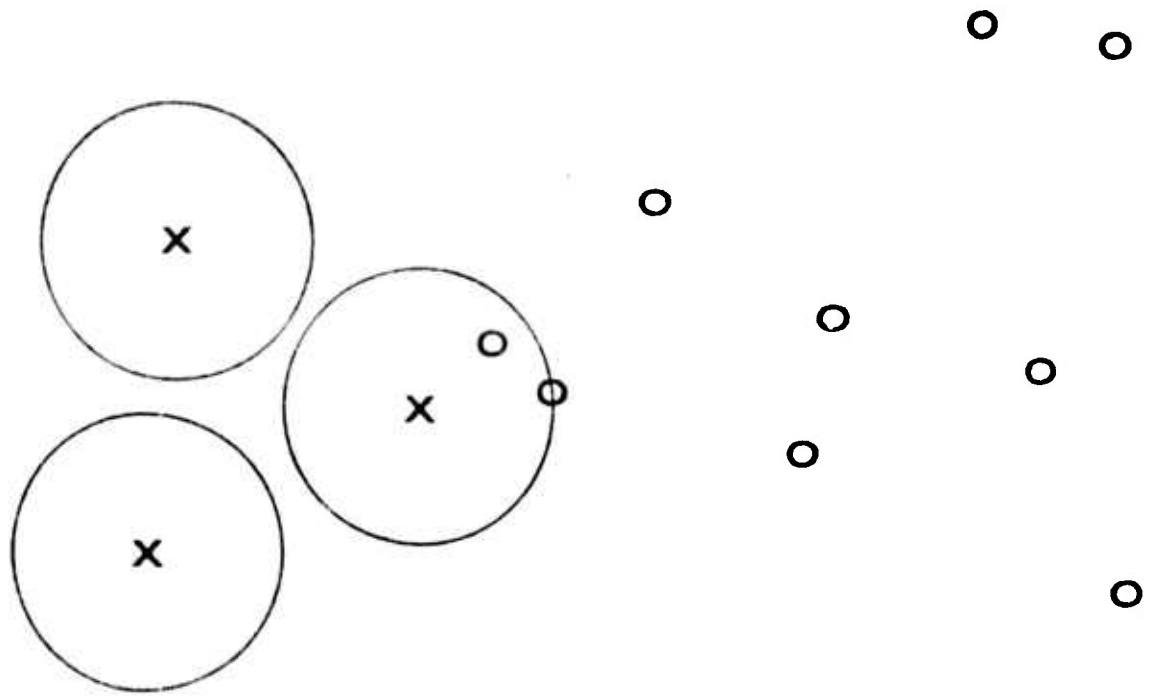


Figure 3.16. A Situation with a Possibly Large Miss Rate.

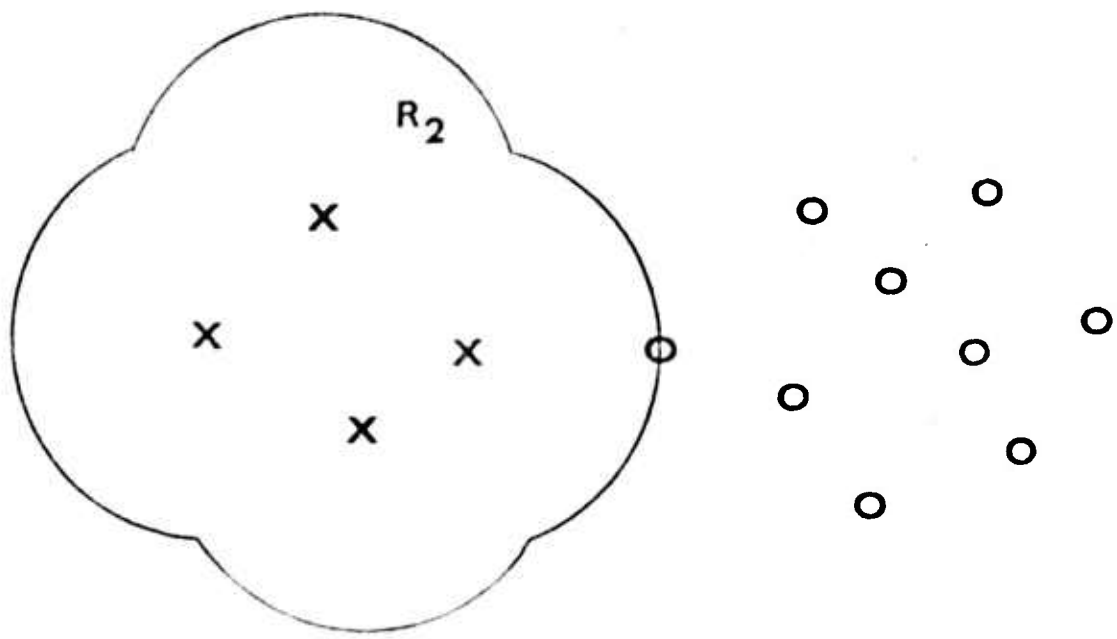


Figure 3.17. An Unsuitable Ordering for Measuring the Miss Rate.

$$E\left\{\int_{\bar{R}_2} dF_2(\underline{x})\right\} \leq \omega \tag{3.41}$$

or

$$Pr\left\{\int_{\bar{R}_2} dF_2(\underline{x}) \leq \emptyset\right\} \geq \nu \tag{3.42}$$

can be made. The quantities ω, $\emptyset$, and ν are determined from n_2, the number of P_2 observations used to design the classifier, and from b, the number of $F_2(\underline{x})$ blocks inside R_2. This procedure may not be distribution-free, as will be discussed later in this section.

One can immediately see that if the functions for ordering the P_2 observations are not judiciously chosen, a very poor estimate of the expected miss probability may be obtained. For example, consider Figure 3.17.

As in the previous figures, the P_1 observations are represented by O's and the P_2 observations by X's. The region R_2, as shown, was constructed for an expected false alarm probability of 0.1. Suppose a measure of the expected miss probability is desired for this classifier. The X's can now be ordered so that this measure can be obtained. Suppose the X's are ordered by hyperspheres which expand from all of the O's. However, none of the blocks which are formed by this ordering will lie entirely in region R_2. Since the theory of distribution-free tolerance regions gives no information about the cumulative distribution contained in a partial block, this ordering procedure is useless for

making a statement about the cumulative distribution $F_2(x)$ in R_2.

Because of the procedure used to order the P_1 observations, all P_2 observations must lie in R_2. Hence, all blocks formed by ordering the P_2 observations must have subsets which are contained in R_2. Since the theory of distribution-free tolerance regions gives no information about the amount of probability in a partial block, good estimates of the cumulative distribution $F_2(\underline{x})$ in R_2 are made only if the ordering procedure is such to allow the blocks which are formed to be contained in R_2.

Consider the following procedure for ordering the P_2 observations. The procedure consists essentially of first locating a P_2 observation. A search is then made for another P_2 observation with a hypersphere which expands from the first P_2 observation. When the second P_2 observation is found, hyperspheres expand from both P_2 observations in search of a third P_2 observation. This process is continued until n_2+1 blocks are formed. The number of blocks contained in R_2 is counted, thereby giving numbers for ω, $\emptyset$, and ν in equations 3.41 and 3.42.

The form of the first ordering function $h_1(\underline{x})$ is arbitrary. For simplicity, let $h_1(\underline{x})$ be linear;

$$h_1(\underline{x}) = \underline{A}^T \underline{x} \tag{3.43}$$

where $\underline{A}$ is a vector constant. Let $\underline{Y}_{(1)}^{(2)}$ be the P_2 observation which satisfies

$$h_1(\underline{Y}_{(1)}^{(2)}) = \min_{1 \le i \le n_2} \underline{A}^T \underline{X}_i^{(2)}$$

Here $\underline{Y}_{(1)}^{(2)}$ is used rather than $\underline{X}_{(1)}^{(2)}$ to avoid the confusion with $\underline{X}_{(1)}^{(2)}$ which was used when the P_1 observations were being ordered. Then a block C_1 is formed, where

$$C_1 = \left\{ \underline{x} : h_1(\underline{x}) \le h_1(\underline{Y}_{(1)}^{(2)}) \right\} \tag{3.43a}$$

But C_1 is not contained in R_2.

Note that a function such as

$$h_1(\underline{x}) = |\underline{x} - \underline{A}|$$

can be used as the first ordering function. Then C_1 may or may not be contained in R_2 depending on the choice of the vector $\underline{A}$. Since $\underline{A}$ cannot be chosen to guarantee that $C_1 \subset R_2$, the estimate of the miss probability will vary for a given R_2 with the choice of $\underline{A}$. Moreover, it is unlikely in an unbounded space and without any knowledge of the location of the P_2 observations that the vector $\underline{A}$ will be chosen so that the first block is contained in the bounded region R_2.

To form the second block a search is made for a P_2 observation with a hypersphere which expands from $\underline{Y}_{(1)}^{(2)}$. Let the second ordering function be given by

$$h_{21}(\underline{x}, \underline{Y}_{(1)}^{(2)}) = |\underline{x} - \underline{Y}_{(1)}^{(2)}| \quad . \tag{3.44}$$

Let $\underline{Y}_{(2)}^{(2)}$ be such that

$$h_2(\underline{Y}_{(2)}^{(2)}) = \min_{\substack{1 \le i \le n_2 \\ \underline{X}_i^{(2)} \neq \underline{Y}_{(1)}^{(2)}}} h_{21}(\underline{X}_i^{(2)}, \underline{Y}_{(1)}^{(2)}) . \tag{3.45}$$

Then the second block is given by

$$C_2 = \left\{ \underline{x} : h_{21}(\underline{x}, \underline{Y}_{(1)}^{(2)}) \le h_2(\underline{Y}_{(2)}^{(2)}) , \; h_1(\underline{x}) > h_1(\underline{Y}_{(1)}^{(2)}) \right\} . \tag{3.46}$$

Let the functions for ordering the third P_2 observation be given by

$$h_{3j}(\underline{x}, \underline{Y}_{(j)}^{(2)}) = |\underline{x} - \underline{Y}_{(j)}^{(2)}| \qquad j = 1, 2 . \tag{3.47}$$

Let

$$h_3(\underline{Y}_{(3)}^{(2)}) = \min_{\substack{1 \le i \le n_2 \\ \underline{X}_i^{(2)} \neq \underline{Y}_{(1)}^{(2)}, \; \underline{Y}_{(2)}^{(2)}}} \; \min_{j=1,2} h_{3j}(\underline{X}_i^{(2)}, \underline{Y}_{(j)}^{(2)}) . \tag{3.48}$$

and

$$K_{3j} = \left\{ \underline{x} : h_{3j}(\underline{x}, \underline{Y}_{(j)}^{(2)}) \le h_3(\underline{Y}_{(3)}^{(2)}) \right\}, \qquad j=1,2 . \tag{3.49}$$

The third block is given by

$$C_3 = \left(\bigcup_{j=1}^{2} K_{3j} \right) \cap \left(\bigcap_{i=1}^{2} \overline{C}_i \right) . \tag{3.50}$$

The functions for ordering the r^{th} P_2 observation are given by

$$h_{rj}(\underline{x}, \underline{Y}_{(j)}^{(2)}) = |\underline{x} - \underline{Y}_{(j)}^{(2)}| \qquad j=1, \ldots, r-1 . \tag{3.51}$$

Let

$$h_r(\underline{Y}_{(r)}^{(2)}) = \min_{\substack{1 \le i \le n_2 \\ \underline{X}_i^{(2)} \neq \underline{Y}_{(1)}^{(2)}, \ldots, \underline{Y}_{(r-1)}^{(2)}}} \; \min_{1 \le j \le r-1} h_{rj}(\underline{X}_i^{(2)}, \underline{Y}_{(j)}^{(2)}) \tag{3.52}$$

and

$$K_{rj} = \left\{ \underline{x} : h_{rj}(\underline{x}, \underline{Y}_{(j)}^{(2)}) \leq h_r(\underline{Y}_{(r)}^{(2)}) \right\}, \qquad j=1,\ldots,r-1 . \tag{3.53}$$

Then the r^{th} block is given by

$$C_r = \left(\bigcup_{j=1}^{r-1} K_{rj} \right) \cap \left(\bigcap_{i=1}^{r-1} \overline{C}_i \right) . \tag{3.54}$$

If the AHE procedure has been used to design the classifier, the regions L_{kr} (equation 3.16) have the same radius for a particular k and for all $r=1,\ldots,n_2$. Suppose m blocks were used to form R_2. If

$$h_2(\underline{Y}_{(2)}^{(2)}) \leq d_m(\underline{X}_{(m)}^{(1)}) ,$$

the entire block

$$C_2 = \{ \underline{x} : h_{21}(\underline{x}, \underline{Y}_{(1)}^{(2)}) \leq h_{21}(\underline{Y}_{(2)}^{(2)}) , h_1(\underline{x}) > h_1(\underline{Y}_{(1)}^{(2)}) \}$$

is contained in R_2. For the general case it is easily seen that if

$$h_i(\underline{Y}_{(i)}^{(2)}) \leq d_m(\underline{X}_{(m)}^{(1)}) , \tag{3.55}$$

the entire block C_i, $i=1,\ldots,n_2$, is contained in R_2. It is also seen that if

$$h_i(\underline{Y}_{(i)}^{(2)}) > d_m(\underline{X}_{(m)}^{(1)}) , \tag{3.56}$$

the block C_i may or may not be contained in R_2. If equation 3.56 is satisfied, one must find a point in C_i which does not lie in R_2 to show that $C_1 \not\subset R_2$. However, to show that C_i is contained in R_2 one must show that all points which lie in C_i also lie in R_2. This is, of

course, no easy task. A conservative statement about the miss probability can be made by counting the number of times equation 3.55 is satisfied for $i=1,\ldots,n_2$.

Let the total number of blocks contained in R_2 (from equation 3.55) be b. Then

$$E\left\{\int_{\overline{R}_2} dF_2(\underline{x})\right\} \leq \frac{n_2 - b + 1}{n_2 + 1} \tag{3.57}$$

and the variance σ^2 is given by

$$\sigma^2\left\{\int_{\overline{R}_2} dF_2(\underline{x})\right\} \leq \begin{cases} \dfrac{b(n_2+1-b)}{(n_2+1)^2(n_2+2)} & b \geq \dfrac{n_2+1}{2} \\ \dfrac{(b+1)(n_2-b)}{(n_2+1)^2(n_2+2)} & b < \dfrac{n_2+1}{2} \end{cases}.$$

The values of ϕ and ν in equation 3.42 can be found by consulting Murphy's graphs (1948).

Suppose equation 3.43 is used to form the first block C_1. This block is not contained in R_2. Suppose also that R_2 is bounded. Note that R_2 is always bounded except in the trivial case where R_2 is the entire measurement space. If R_2 is bounded, the last block C_{n_2+1} is not contained in R_2. Therefore, under normal conditions,

$$b \leq n_2 - 1 .$$

This leads to a minimum expected miss probability

$$E\left\{\int_{\overline{R}_2} dF_2(\underline{x})\right\} \geq \frac{2}{n_2+1} \tag{3.58}$$

To be certain that the P_2 blocks are distribution-free, the ordering functions and region R_2 should be chosen before the P_2 observations are taken. This, of course, can not be done here because R_2 is determined from the P_2 observations. Hence this ordering procedure may produce an estimate of the probability of a miss which is not distribution-free. This procedure is used here only for obtaining a rough estimate of the probability of a miss, so that the classifier can be redesigned if the estimate is much poorer than the desired miss probability.

We expect the estimate to perform this function adequately because it measures how well the hyperspheres making up R_2 are connected. For example, any time a P_2 block is counted as contained in R_2 by equation 3.55 we know that one of the hyperspheres making up R_2 contains the center of another of the hyperspheres making up R_2. These two hyperspheres are certainly connected.

The OHC-R and CHS-R Procedures

If the OHC or the CHS procedure has been used to design the classifier, all of the hyperspheres which make up R_2 do not have the same radius. Therefore, the number of P_2 blocks contained in R_2 varies with the P_2 observation with which the ordering starts. Furthermore, with classifiers of the OHC and CHS designs, the i^{th} block is not necessarily contained in R_2 when

$$h_i\left(\underline{Y}^{(2)}_{(i)}\right) \leq d_m\left(\underline{X}^{(1)}_{(m)}\right) .$$

Figure 3.18 illustrates this fact with the X's representing the P_2 observations. Suppose R_2 is the union of the area inside the four hyperspheres which are shown with thick lines. The boundaries for the C_i blocks are shown with narrow lines. Suppose the ordering function for ordering the 1st observation is given by $h_1(\underline{x}) = x_1$. Then the observation X_1 in the figure is the first ordered P_2 observation, $\underline{Y}_{(1)}^{(2)}$. A hypersphere is then allowed to expand from X_1 in search for a new P_2 observation. The observation $\underline{Y}_{(2)}^{(2)} = X_2$ is located and block C_2 is defined. Block C_2 is contained in R_2. Now hyperspheres expand from X_1 and X_2 in search of a new P_2 observation. The observation $\underline{Y}_{(3)}^{(2)} = X_3$ is located. Note that

$$h_3(\underline{Y}_{(3)}^{(2)}) \leq d_m(\underline{X}_{(m)}^{(1)}) .$$

However, this block is not contained in R_2 because an area outside of the hypersphere which makes up R_2 and surrounds X_2 (the shaded area) is included in block C_3.

We reason here that our foremost concern is that we measure whether the hyperspheres are connected. Using this reasoning, we should use the radii of the hyperspheres which make up R_2 in ordering the P_2 observations. These ordering procedures are called the OHC-R and the CHS-R procedures. The ordering functions are given by

$$h_{rj}(\underline{x}, \underline{Y}_{(j)}^{(2)}) = \frac{d_m(\underline{X}_{(m)}^{(1)})}{d_j(\underline{Y}_{(j)}^{(2)})} \left| \underline{x} - \underline{Y}_{(j)}^{(2)} \right| \qquad (3.59)$$

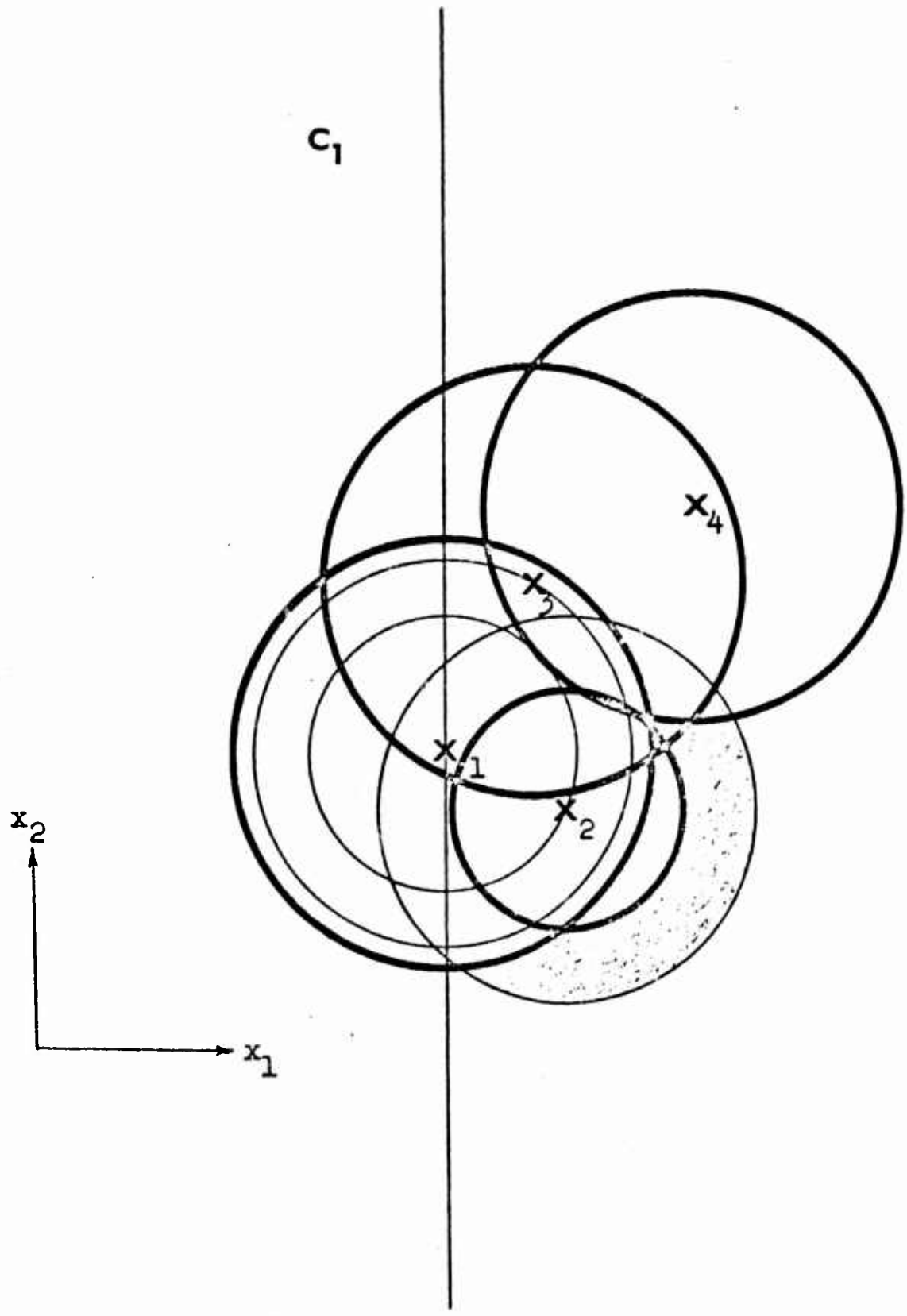


Figure 3.18. The OHC-R and CHS-R Procedures.

where $d_m(\underline{X}^{(1)}_{(m)})$ is the radius of the largest hypersphere making up R_2 and $d_j(\underline{Y}^{(2)}_{(j)})$ is the radius of the hypersphere making up R_2 and centered by $\underline{Y}^{(2)}_{(j)}$. This equation simply alters the rate of expansion of the hypersphere from the observation $\underline{Y}^{(2)}_{(j)}$ so that a hypersphere expanding from any other P_2 observation reaches the hypersphere of R_2 which was centered at that observation at the same time as the hypersphere which expands from $\underline{Y}^{(2)}_{(j)}$ reaches the hypersphere of R_2 which was centered at $\underline{Y}^{(2)}_{(j)}$.

3.10. An Ordering Procedure which Gives Distribution-Free Measures of Both the False Alarm and Miss Errors

It was concluded in the last section that the procedure given there for obtaining a measure of the miss probability may not be independent of the distribution. An ordering procedure is presented in this section which does give distribution-free measures of both the miss probability and the false alarm probability. However, the ordering procedure can at times yield very poor classification regions as will be seen later.

As before, we wish to fix the confidence that the false alarm rate is less than a given quantity. To obtain a distribution-free estimate of both the miss probability and the false alarm probability, we must specify the ordering functions before the outcome of the observations is known, c.f. Murphy (1948). This can be done by using the ordering functions of equations 3.43 through 3.54. However, now these functions (with one

exception) are used to simultaneously order both the P_1 and the P_2 observations. The exception is the first ordering function which is used to order a P_1 observation. The first block is assigned to $R_1 = X - R_2$. The remaining blocks are assigned to region R_2 until the number of P_1 blocks assigned to R_2 is m. The number m is determined from the desired false alarm probability. We will illustrate the ordering procedure by the two-dimensional example of Figure 3.19. Here the X's represent the P_2 observations and the O's the P_1 observations. The number of P_2 observations is 6; the number of P_1 observations is 9. Suppose that $m = 1$ gives the expected false alarm probability that is desired. The first ordering function is used to order a P_2 obser-tation. Suppose $\underline{A}$ in equation 3.43 is chosen so that

$$h_1(\underline{x}) = (1 \quad 0)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

The first block $C_1 = \{\underline{x} : h_1(\underline{x}) \leq h_1(X_{(1)})\}$ (see Fig. 3.19) is assigned to $R_1 = X - R_2$. The ordering is continued, with equations 3.44 through 3.54 being used to order both the P_1 and P_2 observations. This ordering proceeds as follows. A hypersphere expands from $X_{(1)}$ in search of an observation from either class. The block is to be added to R_2 regardless of whether the observation is a P_1 observation or a P_2 observation. Luckily, in our example a P_2 observation $X_{(2)}$ is found. Next, hyperspheres expand from both $X_{(1)}$ and $X_{(2)}$ in search of an observation from either class. A P_2 observation $X_{(3)}$ is found. One

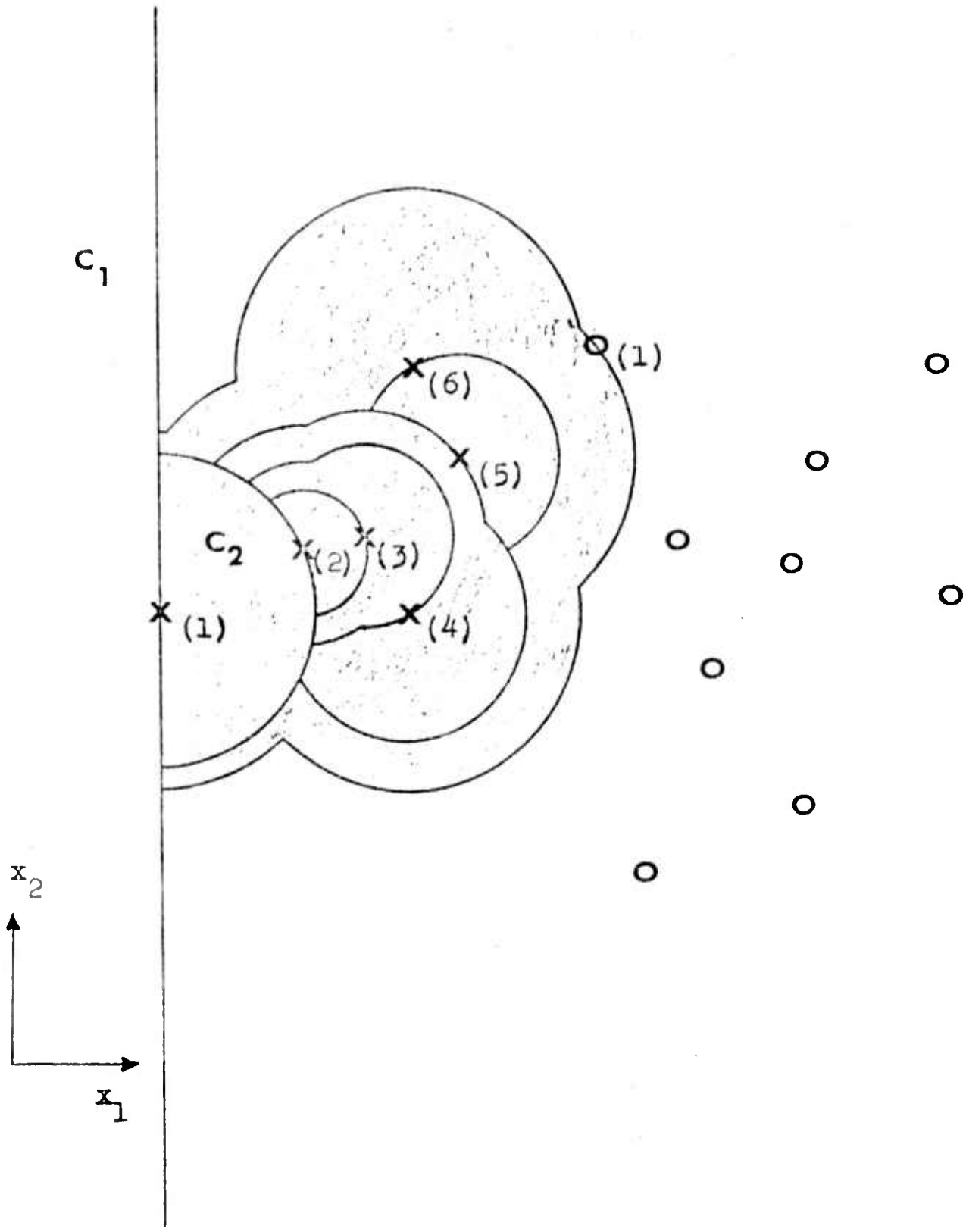


Figure 3.19. Example of the Ordering of Section 3.10.

may trace through the remaining ordering and find that $X_{(4)}$, $X_{(5)}$, $X_{(6)}$ and $O_{(1)}$ are located in that order. The region R_2 is then the shaded region of Figure 3.19. It contains one complete P_1 block, five complete P_2 blocks, and part of another P_2 block. Hence the expected false alarm probability is 1/10 and the expected miss probability is less than 2/7.

Note that the ordering procedure illustrated above will generally yield poor results if $F_1(\underline{x})$ is a multimodal distribution. This is easily seen by noting that a cluster of X's in the lower right corner of Figure 3.19 would not be included in R_2 before $O_{(1)}$ is found and the ordering is curtailed. However, one does know from the number of P_2 blocks in R_2 that the classification regions will yield a poor miss probability. Then some other classification rule can be used.

The ordering procedure can be changed so that it is more acceptable for multimodal situations. For example, if at any time in the process (before m P_1 observations are ordered) more P_1 blocks are being formed than P_2 blocks, a hyperplane can be used to search for a P_2 observation. The P_2 block which is formed is assigned to region R_1. A hypersphere is then allowed to expand from the new P_2 observation and blocks are again added to region R_2. We have essentially searched for another mode of the $F_2(\underline{x})$ distribution.

The procedure presented in this section is somewhat different from the AHE, OHC, and CHS procedures presented in section 3.4. The procedure of this section gives distribution-free measures of both the

miss probability and the false alarm probability. However, it does not maximize the number of P_2 training observations which are recognized as members of class 2. Furthermore, there may be some difficulty in making this procedure work well for multimodal distributions. For this reason, the procedure was not used in the automatic speaker verification experiment of the next chapter.

3.11. Multi-class Problem

The extension of the ordering procedures of this chapter so that they are applicable for more than two classes is investigated in this section. Suppose there exist K classes where $K > 2$. Using distribution-free tolerance regions, Quesenberry and Gessaman (1968) propose a classifier in which all probabilities of misclassification are specified. This approach is permitted through the use of a region in which no decision is made. This region will be called the rejection region. To see how this approach differs from the hypersphere DFTR approach, consider a 2 class example. Suppose the distributions are bivariate and unimodal. Quesenberry and Gessaman suggest that a reasonable ordering might be one which yields a bounded convex region for each class. Suppose the ordering of Figure 2.1 is used to order both the P_1 and the P_2 observations. The resulting decision regions might appear as shown in Figure 3.20. Region R_1 is the region obtained by ordering the P_1 observations and R_2 is the region obtained by ordering the P_2 observations. The decision rule is such that a new observation $\underline{V}$ is classified

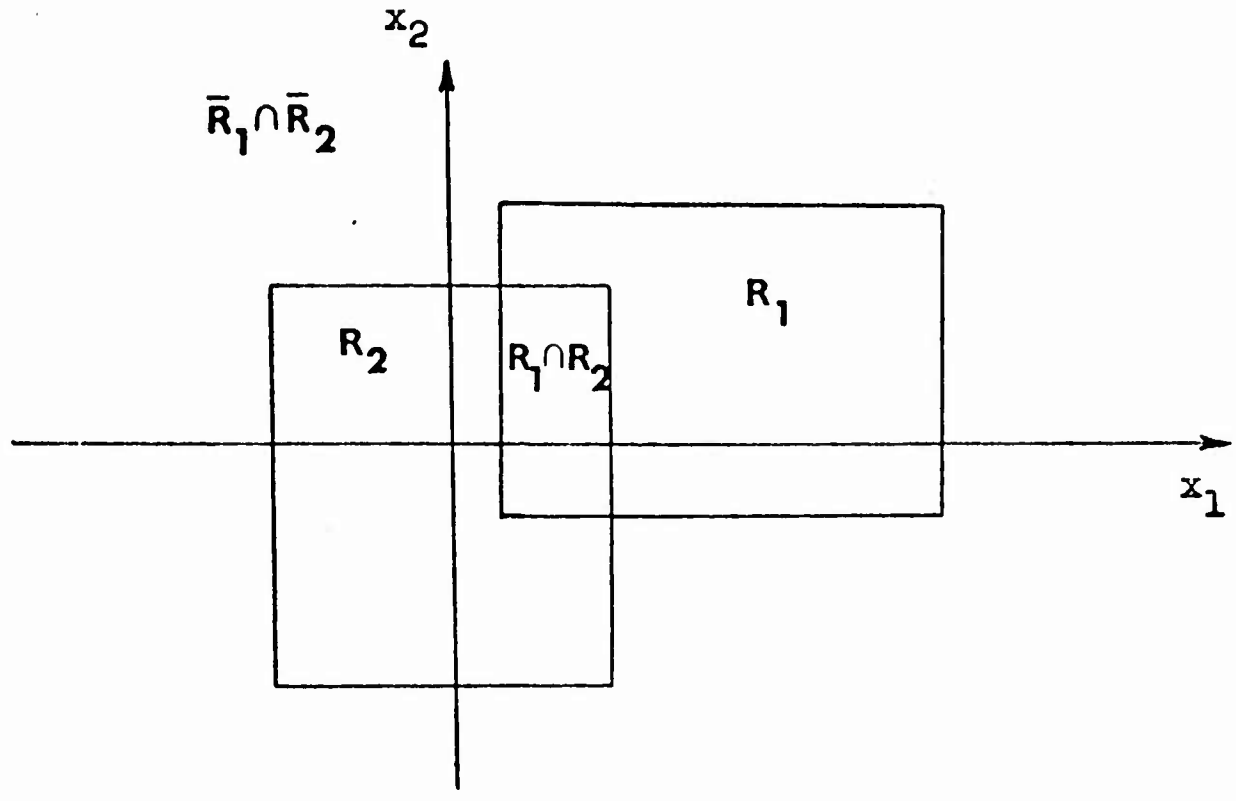


Figure 3.20. A Two-Class Hyperplane Approach.

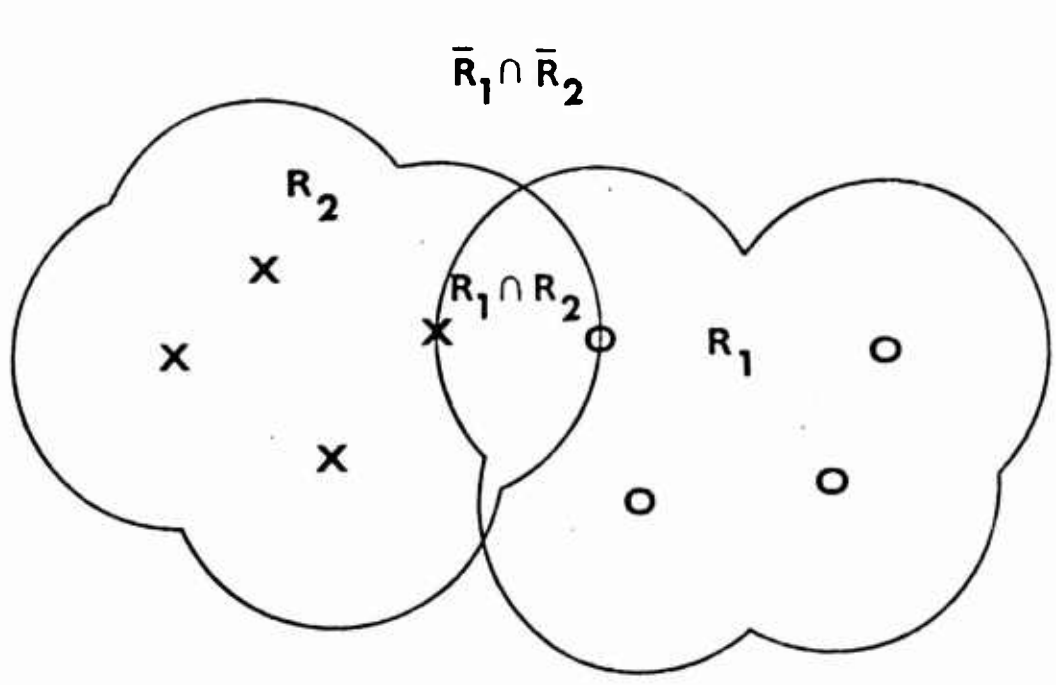


Figure 3.21. A Two-Class Hypersphere Approach.

as a P_1 observation if $\underline{V} \in \{R_1 - (R_1 \cap R_2)\}$. It is classified as a P_2 observation if $\underline{V} \in \{R_2 - (R_1 \cap R_2)\}$. No decision is made if $\underline{V} \in \{(R_1 \cap R_2) \cup (\bar{R}_1 \cap \bar{R}_2)\}$. The difficulty with this procedure is that a large rejection region, $(R_1 \cap R_2) \cup (\bar{R}_1 \cap \bar{R}_2)$, may result according to the ordering functions chosen and the proximity of the classes. Note also for a many-variate problem, an ordering like the one of Figure 3.20 requires many blocks to be removed from both R_1 and R_2 if R_1 and R_2 are to be closed regions.

The conditional probabilities of error are specified as follows. Let the probability that $\underline{V}$ is classified as a P_j observation when it is a P_i observation be denoted by $p(j/i)$ where $i=1,2$, and $j=1,2$. Let R_1 and R_2 be such that

$$\Pr\left\{\int_{\bar{R}_1} d F_1(\underline{x}) \leq \beta_1\right\} = \gamma_1 \tag{3.60}$$

and

$$\Pr\left\{\int_{\bar{R}_2} d F_2(\underline{x}) \leq \beta_2\right\} = \gamma_2.$$

Note that

$$p(2/1) = \int_{R_2 - R_1 \cap R_2} d F_1(\underline{x}) \tag{3.61}$$

and

$$p(1/2) = \int_{R_2 - R_1 \cap R_2} d F_2(\underline{x}) .$$

Since $(R_2 - R_1 \cap R_2) \subset \bar{R}_1$ and $(R_1 - R_1 \cap R_2) \subset \bar{R}_2$,

$$p(2/1) \leq \int_{\bar{R}_1} d F_1(\underline{x}) \tag{3.62}$$

and

$$p(1/2) \leq \int_{R_2} d F_2(\underline{x}) .$$

Then (3.63)

$$Pr\left\{p(2/1) \leq \beta_1 \right\} \geq \gamma_1$$

and

$$Pr\left\{p(1/2) \leq \beta_2 \right\} \geq \gamma_2 .$$

Of course, expanding hyperspheres can be used for ordering both the P_1 and P_2 observations when one wishes to specify both probabilities of error and when one is willing to accept a region in which no decision is made. For example, see Figure 3.21. Here the hyperspheres (circles in the figure) expand from the P_2 observations and order the P_1 observations, thus forming R_2. Also, hyperspheres expand from the P_1 observations and order the P_2 observations, thus forming R_1. Suppose R_1 and R_2 are such that

$$Pr\left\{\int_{R_2} d F_1(\underline{x}) \leq \beta_1 \right\} \geq \gamma_1$$

and (3.64)

$$Pr\left\{\int_{R_1} d F_2(\underline{x}) \leq \beta_2 \right\} = \gamma_2 .$$

Since $(R_1 - R_1 \cap R_2) \subset R_1$ and $(R_2 - R_1 \cap R_2) \subset R_2$, equation 3.63 is also satisfied by the hypersphere DFTR approach. Note that equation 3.64 can be rewritten as

$$\Pr\left\{\left[\int_{R_2 - R_1 \cap R_2} dF_1(\underline{x}) + \int_{R_1 \cap R_2} dF_1(\underline{x})\right] \leq \beta_1\right\} = \gamma_1$$

and

$$\Pr\left\{\left[\int_{R_1 - R_1 \cap R_2} dF_2(\underline{x}) + \int_{R_1 \cap R_2} dF_2(\underline{x})\right] \leq \beta_2\right\} = \gamma_2 \,. \tag{3.65}$$

Hence equation 3.63 is satisfied if all or part of the region $R_1 \cap R_2$ is adjoined to either R_1 or R_2. This means that there is no need to designate $R_1 \cap R_2$ as a rejection region when using the expanding hypersphere approach. Another decision rule, for example the nearest-neighbor rule, Fix and Hodges (1951), can be used to classify an event which occurs in region $R_1 \cap R_2$. In this case the decision rule is

$$d_2 \quad \begin{cases} \text{if } \underline{V} \in R_2 - R_1 \cap R_2 \\ \text{or if } \underline{V} \in R_1 \cap R_2 \text{ and } \min_i |\underline{V} - \underline{X}_i^{(1)}| \geq \min_i |\underline{V} - \underline{X}_i^{(2)}| \end{cases}$$

$$d_1 \quad \begin{cases} \text{if } \underline{V} \in R_1 - R_1 \cap R_2 \\ \text{or if } \underline{V} \in R_1 \cap R_2 \text{ and } \min_i |\underline{V} - \underline{X}_i^{(1)}| < \min_i |\underline{V} - \underline{X}_i^{(2)}| \end{cases} \tag{3.66}$$

$$d_0 \qquad \text{if } \underline{V} \in \bar{R}_1 \cap \bar{R}_2$$

where d_1 means the decision is made that $\underline{V}$ is a P_1 observation, d_2 means the decision is made that $\underline{V}$ is a P_2 observation, and d_0 means that no decision is made.

On the other hand, for the ordering procedure of Figure 3.20, the adjoining of the region $\bar{R}_1 \cap \bar{R}_2$ to R_1 or R_2 also satisfies equation 3.63. This follows because $\bar{R}_1$ and $\bar{R}_2$ were constructed to satisfy

equation 3.60. Then $\bar{R}_1 \cap \bar{R}_2$ need not necessarily be designated a rejection region for the procedure in Figure 3.20. Note also that this procedure need not necessarily yield a rejection region at all if $R_1 \cap R_2 = \emptyset$, the null set. This is not the case for the hypersphere procedure of Figure 3.21. This follows because the hyperspheres are bounded by the observations of the opposite class. Hence $\bar{R}_1 \cap \bar{R}_2 = \emptyset$ only if $\bar{R}_1 = \emptyset$ or $\bar{R}_2 = \emptyset$, a trivial case.

Now consider the hypersphere-DFTR approach for $K > 2$ classes. A multiclass decision rule of the Neyman-Pearson type is used. The particular ordering procedure to follow, and ultimately, whether a rejection region is needed or not, depend on the desired outcome of the classifier. For example, consider the three class problem. Let the probability that $\underline{V}$ is classified as a P_j observation when $\underline{V}$ is a P_i observation be denoted by $p(j/i)$, $i = 1,2,3$, $j = 1,2,3$. There are nine conditional probabilities of classification which obey the following three equations.

$$\begin{aligned} p(1/1) + p(2/1) + p(3/1) &= 1 \\ p(1/2) + p(2/2) + p(3/2) &= 1 \\ p(1/3) + p(2/3) + p(3/3) &= 1 \end{aligned} \tag{3.67}$$

For some problems the following criterion may be desirable.

$$\begin{aligned} &\text{(a)} \quad \Pr\{[p(2/1) + p(3/1)] \leq \beta_1\} \geq \gamma_1 \\ &\text{(b)} \quad \Pr\{[p(1/2) + p(3/2)] \leq \beta_2\} \geq \gamma_2 \\ &\text{(c)} \quad \text{maximize } p(3/3); \text{ (which also minimizes } p(1/3) + p(2/3)\text{)} \end{aligned} \tag{3.68}$$

By analogy to the 2 class hypersphere DFTR approach, hyperspheres expand simultaneously from the P_2 and the P_3 observations to order the P_1 observations. Also, hyperspheres expand simultaneously from the P_1 and P_3 observations to order the P_2 observations. Let the P_1, P_2, and P_3 observations be represented in Figure 3.22 by O's, X's, and •'s, respectively. Regions R_3^1 and R_2^1 are formed about the P_3 and P_2 observations, respectively, when a P_1 observation is ordered. Regions R_3^2 and R_1^2 are formed about the P_3 and P_1 observations, respectively, when a P_2 observation is ordered. Since one would like to maximize $p(3/3)$ and satisfy (a) and (b) in equation 3.68, the following decision rule is a logical choice.

$$
\begin{aligned}
d_3 &: \text{if } \underline{V} \in R_3^1 \cap R_3^2 \\
d_2 &: \text{if } \underline{V} \in R_2^1 \cap \bar{R}_1^2 \cap \overline{R_3^1 \cap R_3^2} \\
d_1 &: \text{if } \underline{V} \in R_1^2 \cap \bar{R}_2^1 \cap \overline{R_3^1 \cap R_3^2} \qquad (3.69) \\
d_0 &: \begin{cases} \text{if } \underline{V} \in \bar{R}_1^2 \cap \bar{R}_2^1 \cap \overline{R_3^1 \cap R_3^2} \\ \text{or if } \underline{V} \in R_1^2 \cap R_2^1 - R_3^1 \cap R_3^2 \cap R_1^2 \cap R_2^1 \qquad \text{(e)} \end{cases}
\end{aligned}
$$

Note that a rejection region is necessary in this case. The crosshatching in Figure 3.21 illustrates the classification regions for this decision rule with the nearest-neighbor rule being used in place of condition (e) above.

In the region with vertical crosshatching, the decision is made in favor of class 3. In the region with northeast (NE) crosshatching, the decision is made in favor of class 1. In the region with northwest (NW)

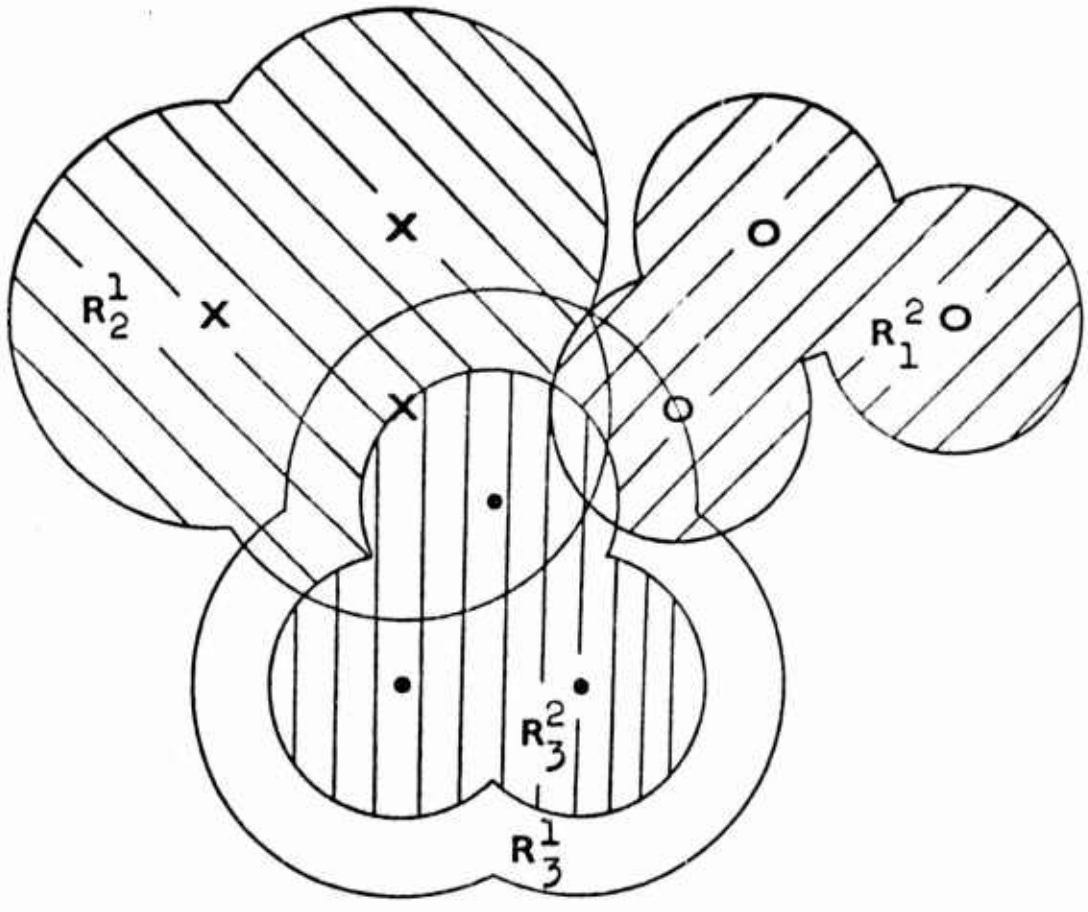


Figure 3.22. A Three-Class Hypersphere Approach.

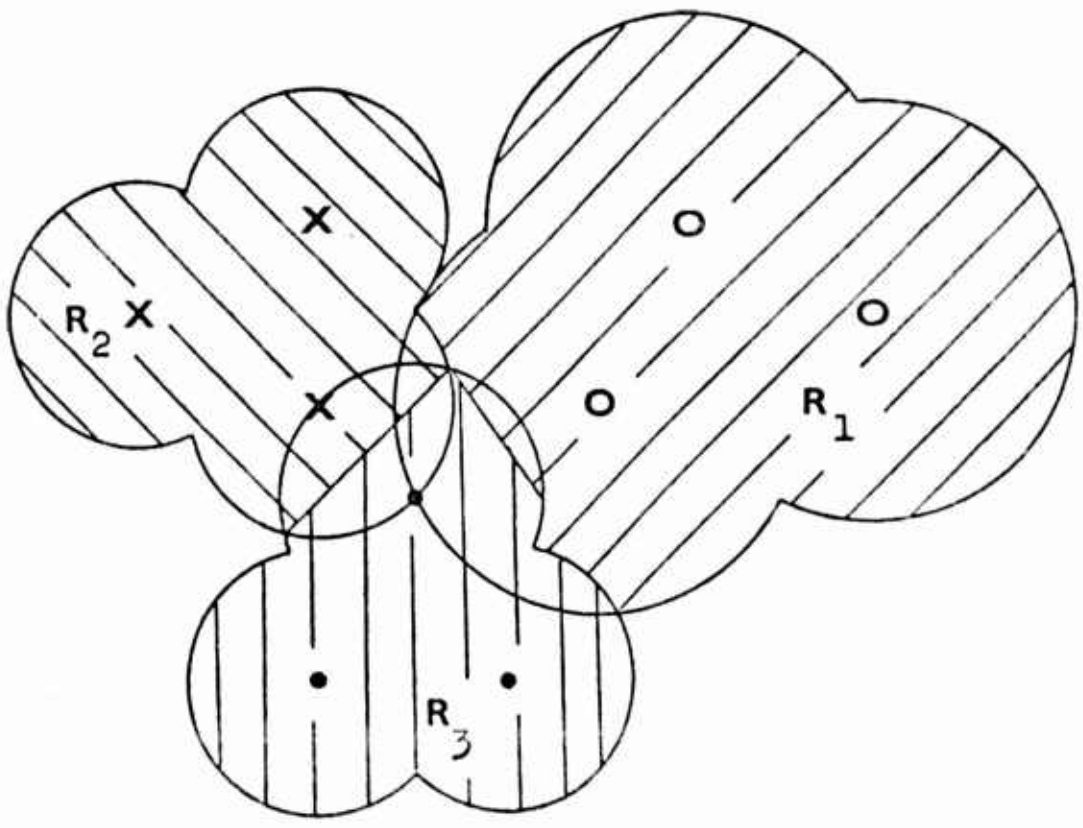


Figure 3.23. An Alternate Three-Class Hypersphere Approach.

crosshatching, the decision is made in favor of class 2. The region with no crosshatching is the rejection region. Therefore, the decision rule is

$$\begin{aligned} &d_3 \quad \text{if } \underline{V} \in \text{(the region with vertical crosshatching)} \\ &d_2 \quad \text{if } \underline{V} \in \text{(the region with NW crosshatching)} \\ &d_1 \quad \text{if } \underline{V} \in \text{(the region with NE crosshatching)} \\ &d_0 \quad \text{if } \underline{V} \in \text{(the region with no crosshatching)} \end{aligned} \tag{3.70}$$

Generalizing this approach to a problem with K classes, the criterion is

$$\begin{aligned} &\Pr\{\sum_{\substack{i \\ i\neq 1}} p(i/1) \leq \beta_1\} \geq \gamma_1 \\ &\Pr\{\sum_{\substack{i \\ i\neq 2}} p(i/2) \leq \beta_2\} \geq \gamma_2 \\ &\qquad\vdots \\ &p\ \{\sum_{\substack{i \\ i\neq K-1}} p(i/K-1) \leq \beta_{K-1}\} \geq \gamma_{K-1} \\ &\text{maximize } p(K/K) \ . \end{aligned} \tag{3.71}$$

It is seen that the P_K observations are encircled by K-1 regions R_K^1, R_K^2, ... R_K^{K-1} and the other observations are encircled by K-2 regions. The resulting decision regions are complex and probably very conservative. Hence this does not seem to be a desirable ordering procedure for K classes.

We now consider an ordering procedure which is better suited to the hypersphere DFTR approach. In this procedure, a hypersphere expands from each observation of a particular class and orders the observations of all other classes. This can be done for all K classes. Hence following probabilities of error are specified.

$$\begin{aligned}
&\Pr\{ \sum_{\substack{i \\ i\neq 1}} p(1/i) \leq \beta_1 \} \geq \gamma_1 \\
&\Pr\{ \sum_{\substack{i \\ i\neq 2}} p(2/i) \leq \beta_2 \} \geq \gamma_2 \qquad (3.72) \\
&\qquad\qquad \vdots \\
&\Pr\{ \sum_{\substack{i \\ i\neq K}} p(K/i) \leq \beta_K \} \geq \gamma_K
\end{aligned}$$

Figure 3.23 illustrates this procedure for the two-dimensional case, for $K = 3$ classes, and for the same samples as shown in Figure 3.22. Region R_1 is obtained by hyperspheres which expand from the P_1 observations to order the P_2 and P_3 observations. The region R_1 in the figure is completed when either a P_2 or a P_3 observation is found. Likewise, region R_2 is completed when the hyperspheres expanding from the P_2 observations intersect either a P_1 or a P_3 observation. Region R_3 is completed when the hyperspheres expanding from the P_3 observations intersect either a P_1 or a P_2 observation. The decision rule is given by equation 3.70.

A criterion which is well suited for the hypersphere DFTR approach is

$$\Pr\{\sum_{\substack{i \\ i \neq K}} p(i/K) \leq \beta_K\} \geq \gamma_K$$

$$\begin{aligned} &\text{maximize} \quad p(1/1) \\ &\text{maximize} \quad p(2/2) \\ &\qquad \vdots \\ &\text{maximize} \quad p(K-1/K-1) \end{aligned} \tag{3.73}$$

Here hyperspheres expand simultaneously from the $P_1, P_2, \ldots, P_{K-1}$ observations to order the P_K observations. An example for the 3 class problem is shown in Figure 3.24 for the same samples as Figures 3.22 and 3.23. Note that a rejection region is unnecessary for this case.

If one wishes to specify the individual errors, $p(i/j)$, the following criterion works well with the hypersphere DFTR procedure and is easily extendable to K classes.

$$\begin{aligned} &\Pr\{p(3/2) \leq \beta_1\} \geq \gamma_1 \ , \quad \text{maximize} \quad p(2/2) \\ &\Pr\{p(2/1) \leq \beta_1\} \geq \gamma_2 \ , \quad \text{maximize} \quad p(1/1) \\ &\Pr\{p(1/3) \leq \beta_3\} \geq \gamma_3 \ , \quad \text{maximize} \quad p(3/3) \end{aligned} \tag{3.74}$$

This case is represented in Figure 3.25. The decision rule is

$$\begin{aligned} &d_3 \qquad \text{if } \underline{V} \in R_3 \cap \bar{R}_1 \cap \bar{R}_2 \\ &d_2 \qquad \text{if } \underline{V} \in R_2 \cap \bar{R}_1 \cap \bar{R}_3 \\ &d_1 \qquad \text{if } \underline{V} \in R_1 \cap \bar{R}_3 \cap \bar{R}_2 \\ &d_0 \qquad \text{otherwise .} \end{aligned} \tag{3.75}$$

One may, of course, use the nearest-neighbor rule in $R_1 \cap R_2$, $R_1 \cap R_3$,

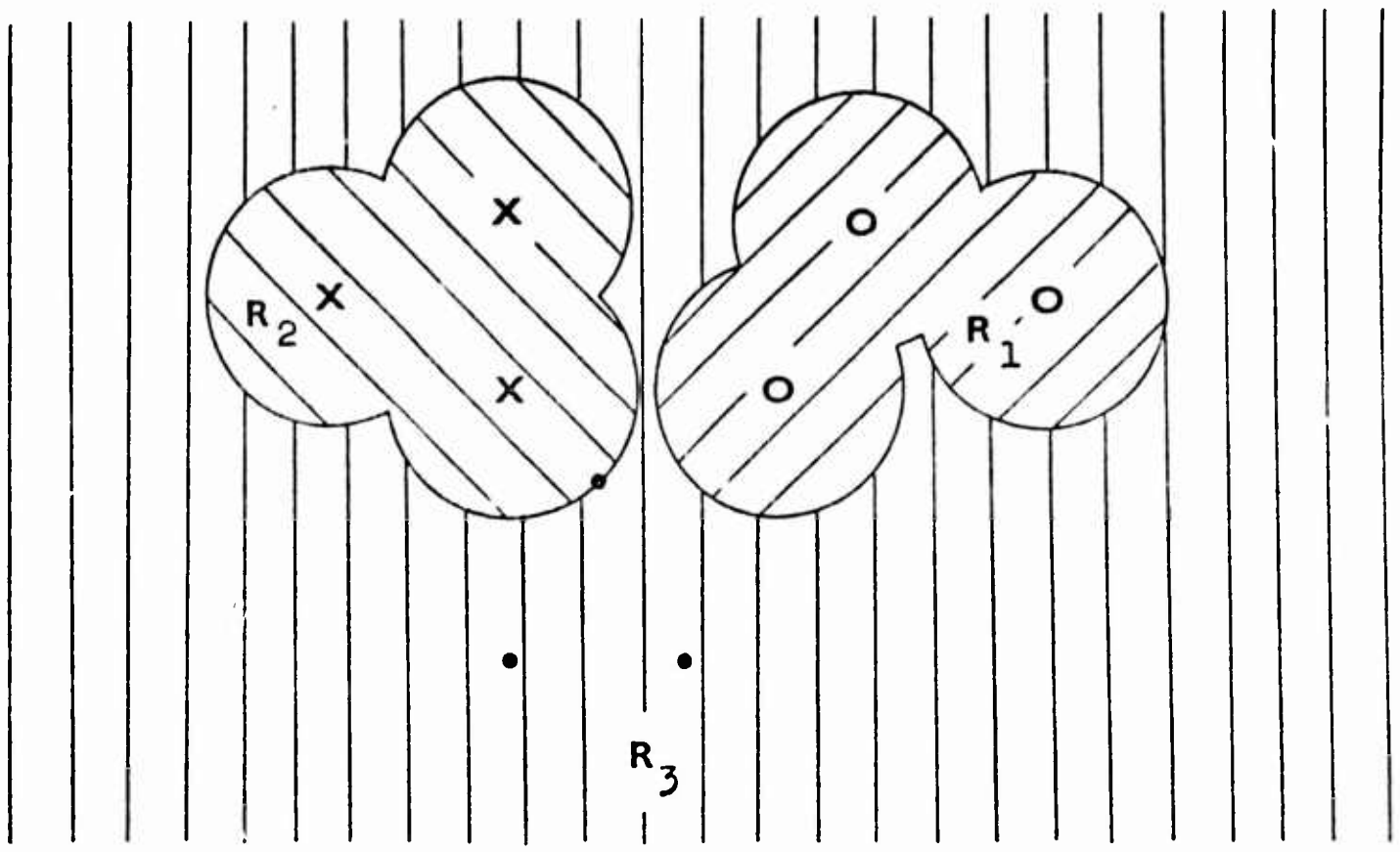


Figure 3.24. A Hypersphere Approach for Criterion A.

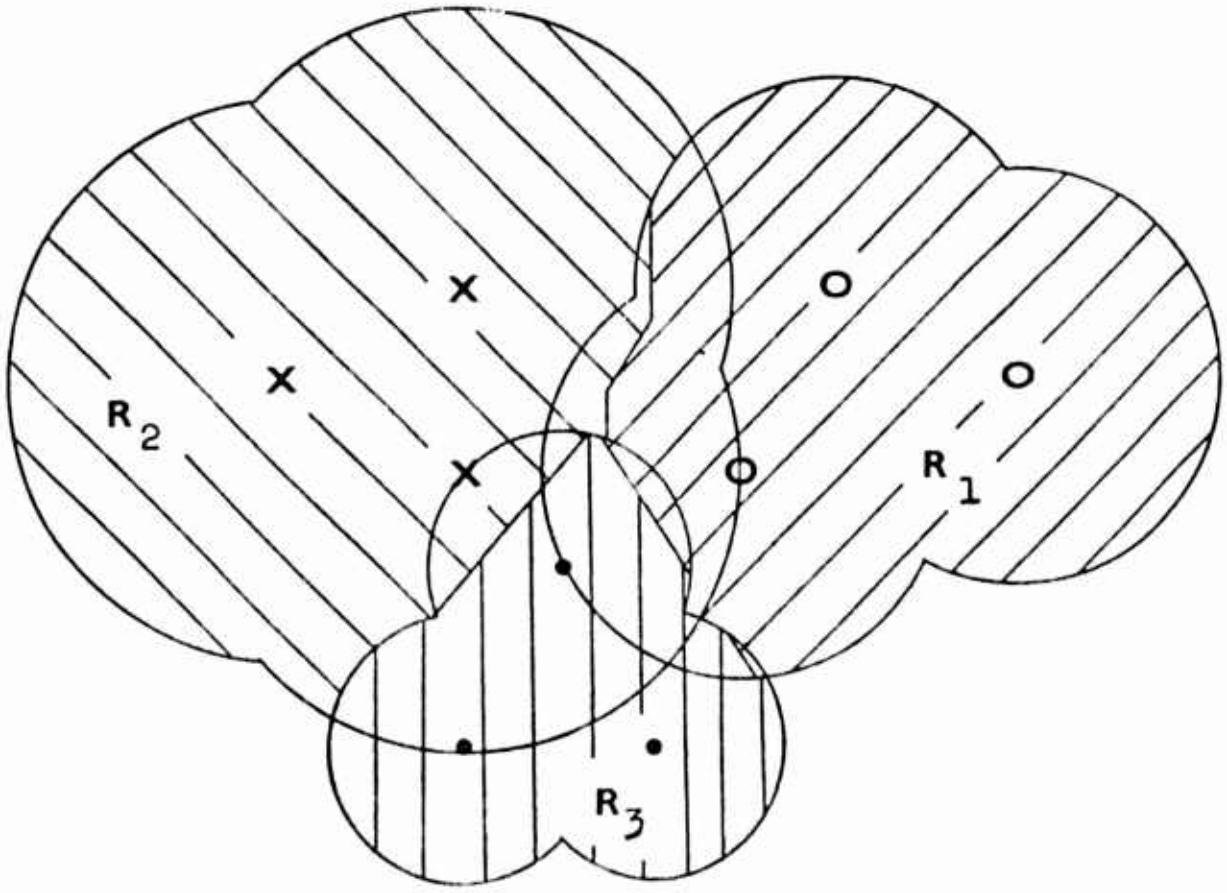


Figure 3.25. A Hypersphere Approach for Criterion B.

and $R_2 \cap R_3$ as shown in the figure.

In summarizing the utility of the hypersphere DFTR approach in the K class problem, it may be stated that

(1) If one wishes to specify the conditional probabilities of error, $\sum_{i \neq 1} p(i/1)$, $\sum_{i \neq 2} p(i/2), \ldots, \sum_{i \neq K} p(i/K)$ where K is large and if one is willing to tolerate a rejection region, another ordering procedure (for example that of Figure 3.20) might be more appropriate than the hypersphere DFTR procedure. This, of course, depends on the proximity of the classes, the modality of the class distributions, and the dimensionality of the space.

(2) For other situations the hypersphere-DFTR approach might be more appropriate because

(a) The rejection region is unnecessary for certain criteria.

(b) The classification regions assume the shape of the observations; hence the approach works well for multimodal class probability distributions.

(c) The minimum number of blocks and hence, the specification of error rates is not dependent on the dimensionality of the space, if a bounded region is desired.

A comment should be made about item (c) above. One may use a hypersphere ordering where the hypersphere contracts around each class. In this case the approach is independent of the dimensionality of the space and is eminantly suited for the criterion given in equation 3.71.

The problem is about what point or points should the hypersphere(s) be centered. Of course, one could use the median of a certain variate to center a contracting hypersphere. But many times this is not suitable for that median does not "center" the class. One may consider using the sample mean of the observations. However, distribution-free tolerance regions are not formed for all distributions when the sample mean is used. Nevertheless, if the class of probability distributions governing the populations is restricted, the sample mean can be used. For example McKay (1935) showed that in a normal population, the order statistics measured from the sample mean are distributed independently of the sample mean.

Summary

A classification procedure has been presented which seems reasonable when nothing is known about the class probability distributions and when it is desirable to specify some of the conditional probabilities of error. It is assumed that a properly classified sample of independent observations is available from each class. This approach is advocated because:

(1) Appropriate decision regions are formed for multimodal probability distributions.

(2) The approach is independent of the dimensionality of the sample space.

(3) The approach is very simple to program on a digital computer.

(4) Automatic data reduction results when this approach is used.

(5) No rejection region is required for certain error criteria.

(6) The approach indicates when the classification system should be redesigned because of expected probabilities of error which are too large.

Chapter 4

AUTOMATIC SPEAKER VERIFICATION

4.1. Introduction

The purpose of this section is to report on an automatic speaker verification system and its use in testing the hypersphere DFTR classification schemes. A speaker verification system is one which tests the purported identity of a speaker from a sample of the speaker's voice. An automatic speaker verification system accomplishes this without any human intervention in the decision process. For example, suppose an automatic speaker verification system is to be used for allowing a person entrance through a company gate. The test subject might be required to push a button beside the name of the person he purports to be. The test subject then says a required phase, for example, "My name is speaker X." A computer then identifies him as the main speaker (the speaker whose identity the test subject has assumed) or an impostor.

This paper uses the terms speaker verification and speaker recognition in accordance with their use in the literature. In a speaker verification system the decision is made whether the test speaker is Speaker 1 or not Speaker 1. In a speaker recognition system the decision is made whether the test speaker is Speaker 1, or Speaker 2, . . . , or Speaker K.

It is assumed in this experiment that the test speaker desires recognition as the main speaker. However, no minicry was involved in these tests.

4.2. Speaker Recognition Review

The relationship between speaker recognition — the recognition of the identity of a speaker from his speech — and speech recognition — the recognition of the content of the speech no matter who is talking — is quite interesting. In the first case it is the similarity in the voices that makes the recognition process difficult while in the latter case it is the difference in the voices that makes the recognition process difficult. Hopefully, what is learned about speaker recognition can be employed to improve speech recognition. The ideal speech processor is one which can extract the differences in the voices for speaker identification and use the similarities in the voices for speech recognition.

The literature contains many papers on speaker recognition. In some of the papers the recognition is done by humans, c.f. Pollack et al. (1954); in some of the papers by a combination of humans and machine, c.f. Kersta (1962a) (1962b); and in some of the papers by machine alone, c.f. Luck (1969).

Pollack, Pickett, and Sumby (1954) tested the ability of humans to identify speakers. Emphasis was placed on recognition accuracy versus duration of speech. For 8 different male speakers, the listeners were able to correctly identify the speakers 70% of the time when a

monosyllabic word was spoken. Recognition rates increased to 80% for a speech sample of .65 seconds duration and to 90% for a speech sample of 1 second duration.

Compton (1963) studied the ability of 15 listeners to recognize 9 speakers from recordings of vowels as the vowel duration was varied. The recognition rates for the vowel with IPA (International Phonetic Association) symbol /i/ ranged from 36% for a speech duration of 25 milliseconds to 57% for a speech duration of 1500 milliseconds. He found that a shorter bandwidth required a greater duration of speech for the same recognition rate. He also found that attenuation of the frequencies below 1020 Hz did not affect the ability of the listeners to recognize the speakers.

Bricker and Pruzansky (1966) conducted a speaker recognition experiment with 10 speakers and 16 listeners, all of whom had worked together for at least two years. They used excerpted vowels, consonant-vowel sequences, monosyllabic words, disyllabic nonsense words and sentences. One of their results was that the recognition rate improved directly with the number of phonemes in a speech sample, even when the duration of the speech sample was controlled. The recognition accuracy ranged from 56% for vowels of 117 milliseconds duration to 98% for sentences of 2.4 seconds duration. They also reported on a computer recognition system which used 60 measurements of each vowel. Here they obtained a recognition rate of 79%. This is a 23% improvement over the result obtained by the listeners for this short duration speech

segment.

L. G. Kersta (1926a), (1962b), (1965), (1966) has employed many techniques for extracting speaker identity from spectrograms or "voice prints" as he calls them. Spectrograms are two-dimensional pictures of the speech showing the speech magnitude versus frequency and time. This is done in two dimensions by allowing the blackness of the spectrogram to be proportional to the magnitude of the speech. Kersta achieves remarkable recognition rates. Some examples are: 97% or better in (1962a), 99% in (1962b), 96% in (1965), better than 90% for 120 speakers in (1966). The references (1962a) and (1962b) report on the training of people to identify the speakers from "voice prints." The references (1965) and (1966) deal with computer recognition of the speakers from "voice prints."

Kersta's work has been criticized by various authors. Ladefoged and Vanderslice (1968) in a 17-page paper criticize Kersta's technique as being more of an art than a science. They list evidence that at times the spectrograms that Kersta uses for identification are readily confused with spectrograms of different people that Ladefoged and Vanderslice have obtained. Young and Cambell (1967) trained observers to identify speakers from monosyllables by Kersta's method of visually comparing spectrograms. The training and test words were spoken in different contexts. Rather poor recognition rates (78% for words in the same context, 37.3% for words in different context) were reported.

They listed the techniques which were employed and possible reasons why their recognition rates were much poorer than Kersta's.

Pruzansky (1963) reported on an automatic speaker recognition system which used energy-time-frequency patterns. Seven band-pass filters were used to obtain the frequency components of the measurement space. The time components were obtained by sampling the output of each filter at 10 msec intervals. Seven male and three female speakers repeated the required speech four times. Ten words were excerpted for analysis. Three utterances of each word by each speaker were used to form a reference vector for each word and each speaker. These three utterances plus the 4^{th} utterance were used to test the system. A test observation was classified into the class whose reference vector gave maximum correlation with the test vector. A recognition rate of 89% was obtained. Pruzansky and Matthews (1964) investigated a method for reducing the number of features used to recognize the speakers. Energy-time-frequency patterns were again used. This time 7 utterances were taken. The first 3 utterances were used to form the reference patterns and the last 4 utterances were used as test observations. A test observation was classified into the class whose reference vector was closest to the test vector. A recognition rate of approximately 90% was obtained. Recognition rates of 90% were also achieved by Hargreaves and Starkweather (1963).

Li, Dammann, and Chapman (1966) reported on an automatic

speaker verification system. Using one main speaker and 10 impostors they received recognition rates from 80% to over 90% depending on the phrase which was used. For the phrase "My name is _____" they received a recognition rate of approximately 85%.

Glenn and Kleiner (1968) used power spectra which were produced during nasal phonations to recognize speakers. A total of 20 male and 10 female speakers were used. Ten occurrences were used to form a reference vector. Ten occurrences of /n/ from a different list were used to test the system. A test observation was classified into the class whose reference vector gave maximum correlation with the test vector. The recognition rate increased proportionally to the number of observations which were used to find the reference vector. For one observation used as the reference vector the recognition rate was 43%; for two observations used to calculate the reference vector, the recognition rate was 68%; for 5 observations used to calculate the reference vector, the recognition rate was 82%; and for 10 observations used to calculate the reference vector, the recognition rate was 93%.

Das (1969) reported on a speaker verification system which employed 6 main speakers and 13 impostors. Fifty training observations were taken from each main speaker. Ninety observations were used to test each main speaker and 30 observations were used to test each impostor. Approximately 1600 measurements were obtained from each speech segment. The number of measurements (dimension of the

measurement space) was reduced to 200 through the use of analysis of variance. The recognition rates ranged from 91.4% to 98.6% with the average recognition rate being 95.4%.

Speaker recognition studies have not been limited to the speakers of English. Solzhenitsyn in The First Circle mentions work in this area during the Stalin Era. In more recent times Ramishvilli (1966) reports on an automatic speaker recognition system which achieves a recognition rate greater than 90%.

It should be noted that in most of these experiments that some phase of the process was done by the human. For example in many cases the words or phonemes were excerpted manually. The speaker verification experiment which follows has the advantage that it can be completely automated.

4.3. Experimental Setup

The experimental work for this thesis was done on equipment at the Applied Research Laboratory (ARL), Sylvania Electronic Systems, Waltham, Massachusetts. The equipment for the preliminary study on phonemes for speaker recognition and the data for the speaker verification project were graciously furnished by Dr. James E. Luck of that laboratory. The equipment consists of a computer program (see Luck (1968a), (1968b)), a Control Data CDC 3200 computer, a Texas Instruments 846 11 bit A/D converter, a remote control unit, and other associated equipment.

Two speaker recognition experiments, a phoneme extraction experiment and a speaker verification experiment, were done. The speech data for the phoneme experiment was taken in a soundproof room at the Audio-Visual Center at Yale University on Ampex professional equipment. It was then played back at ARL on an Ampex PR-10 tape recorder and an EDIT computer program, Luck (1968a), was used to record the data on digital tape.

The speaker verification data was taken at ARL from a microphone in a soundproof room. The speech was band limited, A/D converted and stored on digital tape. The band-pass filter was flat to 3k Hz and down 25 db at 4k Hz. The A/D conversion rate was 8000 samples per second at 10 bit accuracy. The data was recorded on digital tape in 2000-24 bit word records. Each computer word contained two 12 bit samples. The digital tapes were converted at the Yale University Computer Center for use on the IBM 7040-7094 DCS system.

To be able to use the DFTR classification procedures, the utterances must be independent. This was insured by Dr. Luck's data gathering procedure. A command from the teletype instructs the computer to accept data. At the same time an indicator light tells the speaker to repeat the test sentence. Then the speaker says the sentence and the computer processes the data. There is a one-minute delay before the indicator light requests another utterance.

Dr. Luck's system for recording the data is such that the samples

from the A/D converter are temporarily stored from the moment that the indicator light is turned on. When the amplitude of a sample exceeds a certain threshold, the data which follows that sample along with the data just prior to that sample are stored. A total of 8000 samples are stored per utterance.

4.4. Preliminary Study; Speaker Recognition by Analysis of Phonemes

As a first step into speaker recognition we decided to study the use of phonemes for the recognition of speakers. The phonemes to be employed were those that required different lip, jaw, and tongue positions and various sections of the nasal and oral cavities. It was hoped that the different physical constraints placed upon the speech by the different speakers would yield the speaker's identity.

Fifteen adult males were recorded while repeating the words of Table 4.1. They were requested to speak normally and to pause between words. It should be noted that the words were not excerpted from connected speech.

The words which contained the vowels were chosen because they exhibit the various tongue positions. Consider the words "beat, bit, bait, bet, bat." The tongue is positioned toward the front of the mouth for the vowels in these words. For the "ea" of "beat" the vertical position of the tongue is high in the mouth. The vertical position of the tongue is lower for the vowel in "bit," lower still for the vowel in "bait" and lowest for the vowel in "bat." The horizontal position of the tongue

Table 4.1. Text for Phoneme Analysis

Vowels	Stops
beat	pop
bit	bob
bait	tat
bet	dad
bat	kick
but	gig
do	
foot	Sonorants
dough	mum
bought	none
dot	lung
	lull
Fricatives	Dipthongs
fluff	my
verve	how
sauce	toy
zoos	amuse
mesh	
measure	
thin	
then	

during the vowels in "do, foot, dough, bought, dot" is in the back of the mouth. The vertical position of the tongue ranges from high for the vowel in "do" to low for the vowel in "dot."

Words using the consonants were included to see if the information they exhibit for speaker recog ition is sufficient to justify the higher bandwidth necessary to accommodate them. Diphthongs were included so that the information about the identity of the speakers in the transition between phonemes could be investigaged.

The first experiment was a spectral comparison of the 11 vowels listed in Table 4.1 for the different speakers. Each of the 15 speakers said each of the 11 words containing the vowels. Each word was said once and contributes an utterance. Thus there were 165 utterances for analysis.

The speech signals were filtered to pass frequencies up to 4kHz and digitized at 10,000 samples per second. The vowels were isolated by means of the digital computer program EDIT, Luck (1968a). This program enables one to examine digital signals in great detail. Segments of the speech can be heard by the researcher and simultaneously observed on an oscilloscope. The speech window can be lengthened or shortened as desired and the speech which is observed in the window can then be transferred to digital tape.

The Fast Fourier Transform, Cooley and Tukey (1965), Cochran et. al. (1967), was used to obtain the spectral components of the speech.

Four short-term spectra of each vowel were calculated and plotted by the computer. The first spectrum was calculated from the first 256 samples of the isolated vowel. The second spectrum was calculated from the next 256 samples, the third spectrum from the next 256 samples and the fourth spectrum from the next 256 samples. The total energy in each spectrum was made the same. From inputs x_t and y_t, $1 \leq t \leq 256$, the FFT program computes a_f and b_f where

$$a_f + jb_f = \sum_{t=1}^{256} (x_t + jy_t)\, e^{-j\frac{2\pi tf}{256}} \qquad f=1,\ldots,256 \qquad (4.1)$$

Note that x_t is the amplitude of the speech signal at time t and y_t, the imaginary part, is zero for all t. Since the imaginary part is zero, there is symmetry in a_f and b_f about the midfrequency. Hence, only 128 unique a_f and 128 unique b_f result. Let c_f be defined by

$$c_f = \sqrt{a_f^2 + b_f^2} \qquad f=1,\ldots,128\,. \qquad (4.2)$$

The qualitative result from the experiment was that the difference in spectra for two different speakers saying the same vowel was not much greater than the difference in spectra for the same speaker taken at different times during the utterance. This was concluded from visually observing the spectra and by deterministically comparing the spectra. The spectra were deterministically compared as follows. Let the c_f, $f=1,\ldots,128$, be the coordinates of a measurement space.

Then each spectrum can be represented as a vector in this 128-dimensional measurement space. The distance was calculated between the various spectra. In all cases the closest spectrum to any given spectrum was one from the same utterance. That is, the closest spectrum to the third spectrum of one speaker was either the first, second, or fourth spectrum of the speaker. However, in many cases the distance between the first and the fourth spectrum was greater than the distance between spectra of different persons.

This experiment was conducted because in most automatic speaker recognition systems the same point in each phoneme is not located each time that the phoneme is uttered.* The purpose of this test was to determine how this variation in locating the phoneme affected speaker recognition. Since the experiment was not very encouraging and since the EDIT program required considerable computer time for the isolation of the phonemes, we decided to use the information in the transitions between phonemes in addition to the short-term spectra of the phonemes for the speaker verification experiment.

---

* Note that the phonemes here were not extracted automatically but rather they were extracted manually through the EDIT computer program. In the usual automatic speaker recognition system the phoneme is isolated by an approximate method. For example, many times a vowel is located by finding the point of largest amplitude in speech which has been low-pass filtered. Hence in many trials one is likely to isolate many different segments of the phoneme.

4.5. Measurements for the Speaker Verification Experiment

It is hypothesized that different speakers differ in some of the following aspects: the size and shape of the nasal and oral cavities, the placement of the teeth, the tongue mass, and the manner in which the tongue, lips, and jaws are used in speaking. It is assumed that much of the information for the recognition of speakers is contained in the transition between phonemes as well as in the phonemes themselves. For this reason, a simple word which contained a diphthong was analyzed by calculating many short-term spectra over the length of the word. These spectra were used to form the measurement space in which the decision regions were constructed.

Recordings of 225 utterances by each of three speakers and 25 utterances by each of 26 impostors were furnished by Dr. James Luck. These utterances consisted of the sentence "My code is (and then the speaker's initials)" digitized into 10 bit accuracy at 8000 samples per second. Only the word "my" was used in this analysis.

The word "my" was considered a good word for the speaker recognition experiment for the following reasons. First, the nasal /m/ is thought to give good measurements for speaker identification because of the relatively fixed influence of the nasal cavity. See Glenn and Kleiner (1968) or Wolf (1969). Also, the word "my" contains a diphthong (/aI/) and is therefore good for testing the usefulness for speaker identification purposes of the information in the transitions between phonemes.

Furthermore, the word "my" is a good word for our purposes because very little information is lost from the word when it is filtered at 3.4 kHz. Notice that the stop constant "c" in "code" allows the word "my" to be isolated by simple amplitude detection. Therefore, the recognition system is easily automated.

To extract the speaker identity information in the transition between phonemes, many short-term spectra were calculated over the duration of "my." The questions to be answered were: (1) How many spectra should be used? (2) What should be the frequency resolution of each spectrum?

These questions were answered by utilizing three different 256-dimensional measurement spaces. The first measurement space was made up of four spectra with each spectrum having 64 frequency components. The second space consisted of eight spectra with each spectrum having 32 frequency components. The third space consisted of 16 spectra with each spectrum having 16 frequency components. Later a reduced space consisting of six spectra with each spectrum having eight frequency components was used.

Now consider in detail the procedure for obtaining the measurements which were used in the experiments. A chart outlining the procedure is shown in Figure 4.1. The data from Dr. Luck was stored on digital tape in 2000-24 bit words per record. Each word contained two 12 bit samples of the speech. The data was blocked at Yale University

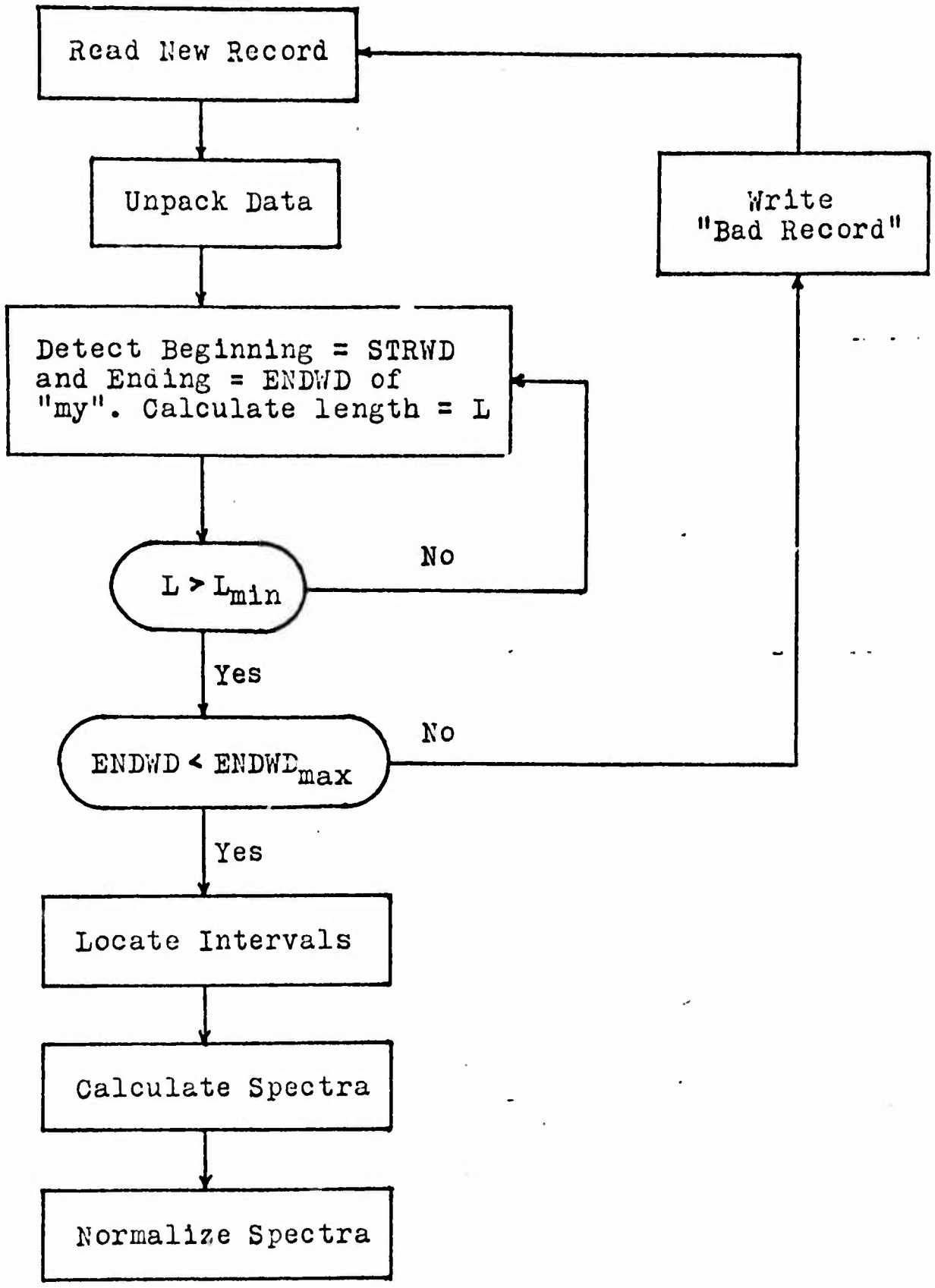


Figure 4.1. Chart for Obtaining Measurements.

into 460 words per record so that it could be read on the 7040-7094 DCS system. The samples were then unpacked and placed in a 36 bit word.

The beginning and ending of the word "my" were located by amplitude detection. The thresholds for the detection were obtained from a preliminary study on the first five utterances by the first three speakers of the impostor data. See Table 4.2. It should be noted that some difficulty was incurred in finding the beginning of "my" for the second speaker. This was due to a low amplitude guttural sound produced by this speaker. Even by inspection it was rather difficult to decide where the /m/ begins.

To ascertain that the word "my" was detected, a minimum acceptable word length and an acceptable word position in the record was established. If the length of the word was too short, a search was made for a longer word. The occurrence of a short word was sometimes due to an extraneous noise made during the recording. The computer was programmed so that if a word of acceptable length was found but the word was cut off by the end of the record, the computer rejected this word. It then proceeded to read another utterance from the input type.

Location of the Intervals

Figure 4.2 shows four typical amplitude versus time waveforms of the filtered speech "My code." The top waveform includes the beginning of the /d/ in "code." The bottom wave form contains none of

Table 4.2. Impostor Data

Speaker Number	Speaker Name	Date Uttered	Speaker Number	Speaker Name	Date Uttered
1	R. Freudberg	7-30-68	15	G. Briskman	7-3-68
2	J. Luck	8-6-68	16	G. Bethoney	7-3-68
3	J. DeLellis	7-3-68	17	C. Mariano	7-3-68
4	H. Shaffer	7-3-68	18	R. Pike	7-3-68
5	C. Howard	7-3-68	19	J. Stoddard	7-3-68
6	H. Manley	7-10-68	20	D. Kinsley	7-3-68
7	K. Lang	7-3-68	21	R. Hasselboum	7-3-68
8	L. Abraham	7-3-68	22	T. MacDonald	7-3-68
9	G. Cummings	7-3-68	23	J. Waggett	7-3-68
10	F. Cassidy	7-3-68	24	S. Free	7-3-68
11	R. Lucy	7-3-68	25	G. Beakley	10-16-68
12	A. Levesque	7-3-68	26	W. Wright	10-16-68
13	J. Boucher	7-3-68	27	B. Fitzgerald	10-16-68
14	H. Halewisn	7-3-68			

Table 4.3. Main Speaker – RF-Data

Sitting Number	Date	Time
1	7-3-68	11:30 a.m.
2	7-3-68	12:15 a.m.
3	7-3-68	1:35 p.m.
4	7-10-68	3:15 p.m.
5	7-11-68	9:15 a.m.
6	7-16-68	10:00 a.m.
7	7-18-68	?
8	7-25-68	11:00 a.m.

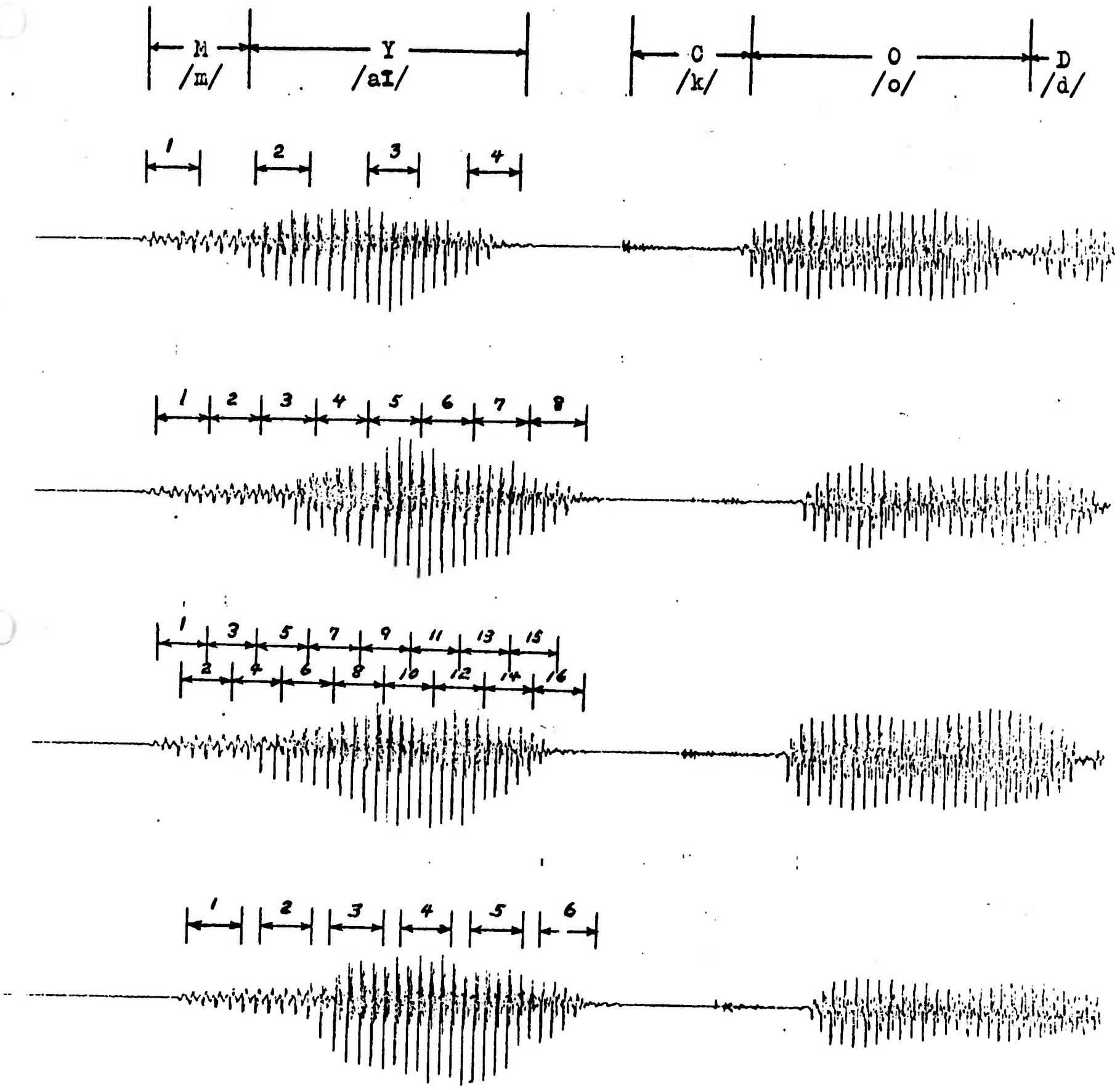


Figure 4.2. Four Utterances of "My code...".

the /d/. Notice that the /c/ has been filtered and is barely visible in the figure.

Above each waveform is a set of numbered intervals. These intervals show where the spectra are calculated for the four different measurement spaces which were used in the experiments. The intervals above the top three waveforms show the regions analyzed for the three different 256 dimensional measurement spaces. The information in the regions outside these intervals is not used. Notice that there is considerable overlap of the intervals when 16 spectra are calculated.

Each interval is of a fixed duration, 256 time samples or approximately 32 milliseconds. The intervals are placed uniformly across "my." The first interval is placed at the "beginning" of the word and the last interval is placed at the "ending" of the word.

A spectrum is calculated from the 256 time samples in each interval by means of the Fast Fourier Transform (FFT). This means that equations 4.1 and 4.2 are used to obtain $c_1, \ldots, c_{128}$ for each interval. Now let

$$d_{f'} = \frac{1}{M} \sum_{f' = M*(f'-1)+1}^{M*f'} c_{f'} \qquad f' = 1, \ldots, 128/M \qquad (4.3)$$

where M is a positive integer. The $d_{f'}$ are used to form the coordinates of the measurement space. For example, let us consider the first measurement space where four spectra are used, with each spectrum having 64 frequency components. Here the integer M of equation 4.3 is equal

to 2 . This gives 64 frequency components, $d_1, \ldots, d_{64}$, for each spectrum. Let us denote the measurements from all 4 spectra by $d_1, \ldots, d_{256}$. Each one of these measurements is used as a coordinate of the first measurement space. Then each utterance of the word "my" is represented by a vector in this space. In these experiments all vectors are normalized to have the same length. This is done to eliminate the variation in amplitude of the speech.

For the second and third measurement spaces, the integer M in equation 4.3 was set equal to 4 and 8 , respectively. Since 8 spectra are calculated for the second measurement space and 16 for the third space, all three spaces have 256 dimensions. Note that the first space considers spectral detail more important than spectral variation in time. The third space does the reverse. That is, it considers spectral variation in time more important than spectral detail. One object of the experiment is to determine the relative importance of these factors for the recognition of speakers.

4.6. Data

Table 4.2 shows the 27 speakers who were recorded on the impostor tape and the date of the recordings. Each speaker recorded 25 utterances of "My code is ____" at one sitting. The first 8 utterances were used as training observations. The last 17 utterances were used as test observations. The number 8 was chosen primarily with the results of Glenn and Kleiner (1968) in mind. It was desirable to use a sufficient

number of training observations to get a good representation of the speaker but nevertheless to have enough unused observations to extensively test the verification system.

The computer program was unable to locate the word "my" in 3 out of the 675 utterances on the impostor tape. In the training phase this happened for the 2^{nd} utterance of the 24^{th} speaker. Hence, the 9^{th} utterance of the 24^{th} speaker was used in its place as a training observation. A total of 8*26 = 208 impostor observations were used in the training phase. The computer program was unable to locate "my" for the 21^{st} utterance by the 4^{th} speaker and the 21^{st} utterance by the 9^{th} speaker. Hence a total of 17*26 - 3 = 439 impostor observations were used in the test phase.

Table 4.3 shows the main speaker data for R. Freudberg (RF). Twenty-five utterances by RF were recorded at each sitting. The first 8 utterances at the first 5 sittings were used as main speaker training observations. The computer program was able to locate the word "my" for all RF utterances. Therefore 5*8 = 40 main speaker training observations and 185 main speaker test observations were used. Note there are 25 main speaker test observations listed with the impostor data.

4.7. Decision Regions

The decision procedures which are used in the speaker verification experiment are summarized in this section. They were completely described in Chapter 3. Let class 1 be the class of the impostors and

class 2 be the class of the main speaker. If a test observation falls in region R_2 it is classified as the main speaker. Otherwise it is classified as an impostor.

(1) AHE (All Hyperspheres Expand) – All hyperspheres expand until n_1+1 blocks have been found. The first m of these blocks are used to form region R_2.

(2) OHC (Ordered Hyperspheres Constant) – All hyperspheres expand until the first block has been formed (a class 1 observation has been ordered). All hyperspheres except the one which ordered the class 1 observation expand until another block is formed. This procedure is continued with each hypersphere stopping after it orders a class 1 observation.

(3) CHS (Conditioned Hyperspheres Stop) – The hyperspheres which order the observations stop as they did in the OHC procedure. However, the procedures differ when an expanding hypersphere intersects a class 1 observation which has already been ordered by another hypersphere. In the CHS procedure the hypersphere stops even though the block has not been completed.

The number of blocks, m , which are used to form the region R_2 should be chosen so that the classifier will produce the desired false alarm rate. In this experiment there was also another objective. There should be a sufficient number of blocks in R_2 so that the differences in

the above ordering procedures become evident. The number m = 7 was thought to be a good compromise between a number small enough for a respectable false alarm rate and a number large enough to exhibit differences in the ordering procedures. For $m = 7$ and $n_1 = 208$ the following parameters are obtained for the classifier. P_{FA} is the false alarm probability, $Pr(\cdot)$ is the probability of $(\cdot)$, $E(\cdot)$ is the expected value of $(\cdot)$, $\sigma(\cdot)$ is the standard deviation of $(\cdot)$.

$$Pr(P_{FA} \leq .055) = .95$$

$$E(P_{FA}) = .0335$$

$$\sigma(P_{FA}) = .0124$$

Table 4.4 shows the results which were obtained when the three DFTR procedures were trained and tested in the three different 256 dimensional spaces. The test false alarm rate, the test miss rate, and the total error rate are listed for each procedure. For example, for the AHE procedure in measurement space 1, 15 out of 439 impostor test observations were classified as the main speaker and 7 out of 185 main speaker test observations were classified as an impostor. This gives a total error rate of 22/624 = .0352. The average error rate obtained for all three ordering procedures in measurement space 1 was .0347. This compared with an average error rate of .0251 for measurement space 2 and an average error rate of .0336 for measurement space 3. Along with having the lowest average error rate, measurement space 2 also had the lowest error rate for each ordering procedure. It was

Table 4.4. Three Different 256-dimensional Measurement Spaces

Training Data

40 Main Speaker Training Samples	208 Impostor Training Samples
$Pr(P_{FA} \leq .07) = .99$	$Pr(P_{FA} \leq .055) = .95$
$E(P_{FA}) = .0355$	
$\sigma(P_{FA}) = .0124$	

Test Results

185 Main Speaker Test Samples — 439 Impostor Test Samples

Measurement Space 1 - 4 spectra, 64 components/spectrum

Procedure	False Alarm	Miss	Total Error
(1) AHE	15/439 = .0342	7/185 = .0379	.0352
(2) OHC	15/439 = .0342	7/185 = .0379	.0352
(3) CHS	14/439 = .0319	7/185 = .0379	.0337

Measurement Space 2 - 8 spectra, 32 components/spectrum

Procedure	False Alarm	Miss	Total Error
(1) AHE	12/439 = .0274	7/185 = .0379	.0305
(2) OHC	10/439 = .0228	6/185 = .0324	.0256
(3) CHS	9/439 = .0205	3/185 = .0162	.0192

Measurement Space 3 - 16 spectra, 16 components/spectrum

Procedure	False Alarm	Miss	Total Error
(1) AHE	7/439 = .0159	26/185 = .1405	.0529
(2) OHC	4/439 = .0091	14/185 = .0756	.0288
(3) CHS	8/439 = .0181	4/185 = .0216	.0192

therefore judged to be the best space. One may speculate why this is true. First, the intervals in which the 8 spectra are calculated usually cover the entire utterance of "my" for the main speaker and for most of the impostors. See Figure 4.2. Therefore much of the information for the recognition of the main speaker and most of the impostors should be contained in this space. One can argue that the space of 16 spectra with 16 components per spectrum emphasized the time element too strongly. Hence an impostor who changes phonemes at the same rate as the main speaker may be easily classified as the main speaker. (This seems to be the case with Impostor 6 as will be discussed later.) Also one may expect the "time features" to vary more from sitting to sitting than the "frequency features." Hence one might expect a high miss rate in measurement space 3 since the main speaker was recorded at 9 different sittings over a period of a month. This high miss rate is seen in Table 4.4 for the AHE and OHC procedures. Measurement space 1 does not perform as well as the other two spaces because (1) less information is extracted from the speech waveform, (2) a large variation in the placement of the second and third spectra can occur over different utterances and (3) only two adjacent frequency components were averaged, thus possibly resulting in excessive spectral fluctuations.

Now compare the three ordering procedures. Very little difference is noted in the total error rates for the three procedures in measurement space 1. In the other two spaces the CHS procedure and the OHC procedure

easily outperform the AHE procedure. This might have been expected since the CHS and the OHC procedures use the information that becomes available during the ordering to select subsequent ordering functions. Hence the region R_2 for these procedures is somewhat shaped by the observations.

The average error rate for the AHE ordering procedure in the three spaces is .0395. This compares with an average error rate of .0299 for the OHC procedure and an average error rate of .0240 for the CHS procedure.

A detailed study was made of the training and test data in the three spaces. Table 4.5 shows the errors made by the particular speakers in these spaces. The errors made by the main speaker are listed by the number of the sitting at which the speech was recorded. MS9 is the sitting of the main speaker which was recorded on the impostor tape. The main results of this study are:

(1) There is a definite correlation between the impostors which are closest in Euclidean distance to the main speaker in the training phase (this was determined from the ordering) and the impostors which made the errors in the test phase. This indicates that 8 utterances are probably sufficient to represent a speaker at one sitting. For example, the four closest impostors in the training phase in measurement space 1 were impostors 16, 13, 12, and 15, in that order. The impostors who made the most errors during the test phase were impostors 12, 16, 10,

Table 4.5. Test Errors in 3 Different 256-dimensional Measurement Spaces

	Space 1			Space 2			Space 3		
	4 spectra 64 components			8 spectra 32 components			16 spectra 16 components		
IMPOSTORS	AHE	OHC	CHS	AHE	OHC	CHS	AHE	OHC	CHS
IM 2	-	-	-	1	-	-	-	-	-
IM 6	1	-	-	1	1	-	5	4	5
IM 9	-	-	-	1	-	-	2	-	-
10	2	2	2	-	-	2	-	-	1
11	-	1	1	-	-	1	-	-	-
12	4	4	4	1	1	1	-	-	-
13	2	1	1	4	2	2	-	-	1
15	1	-	-	-	-	1	-	-	-
16	2	2	2	4	6	2	-	-	1
17	1	1	-	-	-	-	-	-	-
18	1	1	1	-	-	-	-	-	-
21	-	1	1	-	-	-	-	-	-
22	1	2	2	-	-	-	-	-	-
Main Speakers									
MS 1	-	-	-	-	-	-	1	-	-
2	-	-	-	-	-	1	2	2	1
3	1	1	1	2	2	-	2	-	-
5	-	-	-	-	-	-	1	1	1
6	3	3	3	1	1	-	3	2	-
7	-	-	-	1	1	-	1	-	-
8	-	-	-	2	1	1	4	2	-
9	3	3	3	1	1	1	12	7	2

22, and 13. The four closest impostors in the second space were impostors 13, 16, 18, and 12. The most errors were made by impostors 16, 13, and 12. The four closest impostors in the 3rd space were impostors 6, 16, 13, and 9. Impostor 6 made the most errors in that space.

(2) The impostors which were relatively good substitutes for the main speaker in one space were not necessarily good substitutes for the main speaker in one of the other spaces. For example, the observations of impostor 12 were near the observations of the main speaker in measurement space 1 but were not in measurement space 3.

(3) There was a tendency for main speaker observations of the same sitting to cluster more than main speaker observations of different sittings. This was manifest in smaller error rates for MS1, MS2, MS3, MS4, and MS5 than for MS6, MS7, MS8, and MS9, a situation which occurred for all cases except the CHS procedure in measurement space 3. Furthermore, the closest main speaker training observation to a main speaker test observation from MS1,...,MS5 was usually a training observation which was recorded at the same sitting. This occurred for over 70% of the test observations from MS1,...,MS5.

(4) The AHE procedure in measurement space 3 yielded a relatively poor miss rate. This was because the training observations from Impostor 6 were near those of the main speaker. In the training phase the ordering hyperspheres for the AHE procedure continued to expand into the training observations of Impostor 6. Hence the hyperspheres were stopped before

they became large enough to yield a good miss rate. For the CHC and CHS procedures, the hyperspheres which were expanding into the training observations of Impostor 6 were stopped, allowing the remaining hyperspheres to become much larger than the hyperspheres for the AHE procedure. Here is a case where the OHC and the CHS procedures clearly perform better than the AHE procedure.

4.8. Reduced Measurement Space

In this section the hypersphere DFTR approaches are experimentally compared with other classification methods using two different main speakers. Because of the excessive computing time for these experiments in a 256-dimensional measurement space, it was decided to reduce the size of the space. For example, the use of the nearest neighbor rule in a 256-dimensional space would have taken approximately 10 hours of computing time. This is partially due to the fact that the pertinent data for the nearest-neighbor rule would not fit into the core of the IBM 7094.

The space which gave the best results in the 3 space experiment was that space consisting of 8 spectra with 32 components per spectrum. It was decided that approximate coverage of the word "my" with spectra was important. A compromise of 6 spectra placed uniformly across "my" was reached for the reduced dimensional space. The number of components per spectrum was reduced to 8, making the reduced space a 48-dimensional space.

Table 4.6 shows the results for this space when two different main

Table 4.6. 48-dimensional Space, Two Main Speakers

6 spectra, 8 components/spectrum

Decision Procedure	RF = Main Speaker			JD = Main Speaker		
	False Alarm	Miss	Total Error	False Alarm	Miss	Total Error
$Pr(P_{FA} \leq .055) = .95$						
$E(P_{FA}) = .0335$						
(1) AHE	.0319	.2972	.1105	.0297	.3162	.0900
(2) OHC	.0297	.1837	.0754	.0297	.2991	.0863
(3) CHS	.0342	.1243	.0609	.0297	.2991	.0863
$Pr(P_{FA} \leq .085) = .95$						
$E(P_{FA}) = .0574$						
(1) AHE	.0550	.2000	.0977	.0388	.1710	.0665
(2) OHC	.0455	.0650	.0512	.0410	.1625	.0665
(3) CHS	.0387	.0595	.0448	.0434	.1368	.0629
NN Rule	.0205	.0216	.0209	.0369	.1111	.0522
5 MS, 26 IM	.0228	.0216	.0224			
1 MS, 26 IM	.0137	.0973	.0385			
1 MS, 1 IM	.1985	.0000	.1395			

speakers were used. The first row shows the DFTR results when 7 blocks were used to form R_2. The next row shows the results when 12 blocks were used to form R_2. The results for the different DFTR procedures varied less for speaker JD than for speaker RF. The average error rate over all three DFTR procedures was approximately equal for each speaker. The average error for the 7 block DFTR problem was .0823 for RF and .0875 for JD. The average error for the 12 block DFTR problem was .0646 for RF and .0653 for JD.

The nearest-neighbor (NN) rule was then used to classify the observations.* The nearest-neighbor rule performed better than the best DFTR procedure for both speakers. This might have been expected from the nature of the decision rules. Both the NN and the DFTR procedures use the information about all main speaker training observations. However, the DFTR procedures use less information about the impostor training observations than the NN rule. The AHE-DFTR procedure uses only the information about the m^{th} impostor training observation to obtain the decision regions. The OHC and the CHS procedures use the information about the first m impostor training observations to obtain the decision regions. The NN procedure, however, uses information about all the impostor training observations to obtain the decision regions. This allows the main speaker observations to fluctuate more in certain

* See Chapter 5 for details of the NN rule.

directions before errors are made than the DFTR procedures allow. Other differences in the decision rules may tend to cancel. This, of course, depends on the situation. For example, the NN rule allows the impostor test observations which are near the main speaker observations to flucuate more before errors result than the DFTR rules allow. However, the DFTR rule allows the main speaker observations to fluctuate more toward the nearest impostor observations before errors result. This is because the hyperspheres forming R_2 in the DFTR procedures actually intersect the closest impostor training observations. But the hyperplanes forming R_2 in the NN procedure are constructed midway between each main speaker and each impostor training observation.

Considerations that tend to favor the DFTR rule over the NN rule are:

(1) The NN rule takes 5 times longer to test an observation than the DFTR rule. (This was for 40 main speaker training observations and 208 impostor training observations in a 48-dimensional space.) Thus if the rules are compared on an equal computing time basis, the DFTR rules are, in fact, superior in performance. Note further that the CHS-DFTR procedure in measurement space 2 is superior in both error rate and computing time to the NN rule in the reduced space. (The DFTR procedures in measurement space 2 took 2 times longer to test than the DFTR procedures in the reduced space. See section 4.11.)

(2) The DFTR rule allows the machine designer to know how well the machine is expected to perform; i.e. it gives information about

the expected false alarm and the expected miss rates.

(3) The DFTR rule requires storage of the main speaker training observations and approximately m impostor training observations (more for the CHS procedure; less for the AHE procedure) whereas the NN rule requires storage of the main speaker training observations and all of the impostor training observations .

Other decision rules were applied to the data with RF as the main speaker. The rule labeled 5 MS, 26 IM in Table 4.6 was obtained as follows. A sample mean was calculated from the 8 main speaker training observations at each sitting. These 5 sample means were used as main speaker reference observations. The sample mean was then calculated from the training observations of each impostor. This gave 26 impostor reference observations. The nearest-neighbor rule was then employed using the 5 main speaker reference observations and the 26 impostor reference observations. Very good results were obtained from this procedure as is seen in Table 4.6. This indicates that the observations from each sitting were well clustered.

For the rule labeled 1 MS, 26 IM, one reference vector was calculated from the main speaker training observations. The 26 impostor reference observations were again used. The nearest-neighbor rule was employed and a rather good error rate resulted. Note that the miss rate jumped from .0216 when 5 main speaker reference observations were used to .0973 when 1 main speaker reference observation was used. This

further indicates that the main speaker observations from all sittings are not clustered as well as the main speaker observations from one sitting.

For the rule labeled 1 MS, 1 IM the impostor training observations were averaged to obtain one impostor reference observation. A hypersphere was then constructed midway between the main speaker reference observation and the impostor reference observation. Rather poor error rates resulted as seen from the table.

4.9. Comparison of Expected False Alarm Probability with Test False Alarm Rate

One of the most important purposes of the speaker recognition experiment was the comparison of the probability of false alarm for which the pattern recognizer was designed and the false alarm rate which was obtained from the test observations. In one experiment a pattern recognizer was designed using 208 impostor training observations, and seven blocks to form region R_2. Hence this pattern recognizer had the following characteristics; where P_{FA} denotes the false alarm probability, $Pr(\cdot)$ denotes probability, $E(\cdot)$ denotes expected value, and $\sigma(\cdot)$ denotes standard deviation.

$$Pr(P_{FA} \leq .055) = .95$$

$$E(P_{FA}) = .0335$$

$$\sigma(P_{FA}) = .0124$$

The test results for the pattern recognizers designed by the three hypersphere DFTR procedures are shown in Table 4.7. The last two rows in this table show the results for pattern recognizers which were designed using 12 blocks to form R_2. For this case the characteristics were

$$\Pr(P_{FA} \leq .85) = .95$$

$$E(P_{FA}) = .0574$$

$$\sigma(P_{FA}) = .0160 .$$

Consider the results in the top row of the table. Here a 256-dimensional space consisting of 4 spectra with each spectrum having 64 frequency components was used. The main speaker was RF and $m = 7$ blocks were used to form region R_2. Using the training sample, a pattern recognizer was designed in which we are 95% confident that the false alarm rate is less than .055. The expected false alarm rate is .0335. Using 439 impostor test observations, a false alarm rate of .0342 was obtained. A 95% upper confidence limit can be obtained from these test results. For this particular example we are 95% confident that the false alarm rate is less than .055.

The confidence limit on the test results can be obtained as follows. The test observations are assumed to be independent. Let n be the total number of impostor test observations. Let n_{FA} be the number of impostor test observations which are erroneously classified as the main speaker. The distribution for n_{FA} is given by

$$\binom{n}{n_{FA}} P_{FA}^{n_{FA}} (1 - P_{FA})^{n - n_{FA}} .$$

Table 4.7. Comparison of Expected False Alarm Probability with Test False Alarm Rate

Situation	TRAINING SAMPLE		TEST SAMPLE	
	95% Upper Tolerance Limit on P_{FA}	$E(P_{FA})$	False Alarm Rate	95% Upper Confidence Limit on Test False Alarm Rate
4 spectra RF=MS, m=7				
(1) AHE	.055	.0335	.0342	.055
(2) OHC	.055	.0335	.0342	.055
(3) CHS	.055	.0335	.0319	.05
8 spectra RD=MS, m=7				
(1) AHE	.055	.0335	.0274	.045
(2) OHC	.055	.0335	.0228	.04
(3) CHS	.055	.0335	.0205	.035
16 spectra RF=MS, m=7				
(1) AHE	.055	.0335	.0159	.03
(2) OHC	.055	.0335	.0091	.02
(3) CHS	.055	.0335	.0181	.03
6 spectra RF=MS, m=7				
(1) AHE	.055	.0335	.0319	.05
(2) OHC	.055	.0335	.0297	.05
(3) CHS	.055	.0335	.0342	.055
6 spectra JD=MS, m=7				
(1) AHE	.055	.0335	.0297	.05
(2) OHC	.055	.0335	.0297	.05
(3) CHS	.055	.0335	.0297	.05
6 spectra RF=MS, m=12				
(1) AHE	.085	.0574	.0550	.08
(2) OHC	.085	.0574	.0455	.07
(3) CHS	.085	.0574	.0387	.06
6 spectra JD=MS, m=12				
(1) AHE	.085	.0574	.0388	.06
(2) OHC	.085	.0574	.0410	.065
(3) CHS	.085	.0574	.0434	.065

See Cramer (1946). The maximum likelihood estimate of P_{FA} is obtained by equating the partial derivative of the above equation with respect to P_{FA} to zero. The maximum likelihood estimate of P_{FA} is given by

$$\hat{P}_{FA} = n_{FA}/n_2 .$$

This is precisely the test false alarm rate. A $100\tau\%$ upper confidence limit can be defined, with $p(\cdot)$ denoting the probability density of $(\cdot)$.

$$\Pr(P_{FA} \leq \theta) = \int_0^{\theta} p(\hat{P}_{FA}/P_{FA})\, dP_{FA} = \tau$$

From this equation it is possible to find θ for a given τ. The values of θ for $\tau = .95$, which are listed in Table 4.7, were obtained from graphs found in Crow, Davis, and Maxfield (1960).

The most important results from this table are:

(1) All test false alarm rates lie in the 95% upper tolerance limit.

(2) In all cases we are 95% confident that the true false alarm rate lies within the limit obtained from the 95% upper tolerance regions. A note of clarification needs to be made about this. Assuming that the test false alarm rate stays the same, the upper confidence limit on the false alarm rate (at the same confidence level) will decrease as more observations are tested. Since more observations were used in the test phase than in the training phase, we want the upper confidence limit on the false alarm rate at confidence level 95% to be less than the upper tolerance level on the false alarm probability at tolerance level 95%. This is the result

which was obtained.

(3) (a) All test false alarm rates are less than the expected false alarm rate plus the standard deviation.

(b) In 18 out of 21 cases the test false alarm rate is less than the expected false alarm rate.

(c) In 6 out of 21 cases the test false alarm rate is as close as one observation out of 439 to the expected false alarm rate. That is, if the test false alarm rate is less than $E\{P_{FA}\}$, one more impostor which is classified as the main speaker will cause the test false alarm rate to be greater than $E\{P_{FA}\}$. Likewise, if the test false alarm rate is greater than $E\{P_{FA}\}$, one less impostor which is erroneously classified will cause the test false alarm rate to be less than $E\{P_{FA}\}$.

(d) In 14 out of 21 cases the test false alarm rate falls in the interval $[E\{P_{FA}\} - \sigma\{P_{FA}\},\ E\{P_{FA}\} + \sigma\{P_{FA}\}]$.

(4) (a) In the 15 cases where $E\{P_{FA}\} = .0335$, the average test false alarm rate, FA_{AV}, obtained by summing the 15 test false alarm rates and dividing by 15 was

$$FA_{AV} = .0267 .$$

This gives the following relation

$$FA_{AV} \sim .8\, E\{P_{FA}\} .$$

(b) In the 6 cases where $E\{P_{FA}\} = .0574$, FA_{AV} was found to be .0438. This gives

$$FA_{AV} \sim .76\, E\{P_{FA}\}.$$

(5) There are no significant differences in the relationship of the average test false alarm rate to the expected false alarm for the various DFTR procedures. The results are:

For the AHE procedure, $FA_{AV} \sim .82\, E\{P_{FA}\}$

For the OHC procedure, $FA_{AV} \sim .75\, E\{P_{FA}\}$

For the CHS procedure, $FA_{AV} \sim .78\, E\{P_{FA}\}$.

4.10. Comparison of a Measure of the Expected Miss Rate with the Test Miss Rate

A method was presented in Chapter 3 for obtaining a measure of the expected miss probability. The method had straightforward application to a classifier constructed by the AHE procedure since all hyperspheres making up region R_2 have the same radius. However, for the OHC and CHS procedures this is not the case. For these procedures the number of P_2 blocks contained in R_2 varies with the P_2 observation with which the ordering starts. An obvious way to overcome this dilemma is to average over the number of blocks which are obtained when each P_2 observation is used to start the ordering. This, however, is usually a prohibitively time consuming process. The measures of the expected miss probability for the classifiers designed by the OHC and CHS procedures were obtained principally by two different methods. One method, labeled OHC-R or CHS-R in Tables 4.8a and 4.8b, uses the information about the radii of the hyperspheres making up R_2 to order the P_2

observations. The method labeled OHC or CHS (no R) in Tables 4.8a and 4.8b does not use information about these radii in ordering the P_2 observations. The ordering functions for the CHC and the CHS procedure are given by equation 3.51. The ordering functions for the OHC-R and the CHS-R procedures are given by equation 3.59.

Table 4.8a compares the measure of the miss probability obtained in the training phase with the miss rate obtained in the test phase for the 3 different 256-dimensional spaces. Table 4.8b compares these quantities for the experiments in the 48-dimensional space. Consider the first row of Table 4.8a. For the AHE procedure with main speaker RF and seven P_1 blocks used to form R_2, 37 P_2 blocks were found in region R_2. Note that there is also at least part of an additional P_2 block in R_2. The expected miss probability is .0976 for 37 blocks and .0732 for 38 blocks. The miss rate for the 185 main speaker test observations was .0379. The 95% upper tolerance limit on the miss probability using 37 blocks was .18. The 95% upper confidence limit on the test miss rate was .07. In the following, the larger number in the column of expected miss probabilities will be denoted by E^+ and the smaller number by E^-. Let the standard deviation corresponding to the number of complete blocks in R_2 be σ^+. The important results are:

(1) The expected miss probability for the OHC and the CHS procedures is much too conservative if the information about the radii of the hyperspheres making up R_2 is not used in ordering the P_2 observations.

Table 4.8a. Comparison of a Measure of the Expected Miss Probability with the Test Miss Rate in the 256-Dimensional Spaces

Situation	TRAINING SAMPLE			TEST SAMPLE	
	Number of Blocks	95% Upper Tolerance Limit on P_m	$E\{P_m\}$	Miss Rate	95% Upper Confidence Limit on the Miss Rate
4 spectra RF=MS, M=7					
(1) AHE	37	.18	.0732-.0976	.0379	.07
(2) OHC	35	.24	.1220-.1463	.0379	.07
OHC-R	37	.18	.0732-.0976		
(3) CHS	35	.24	.1220-.1463	.0379	.07
CHS-R	37	.18	.0732-.0976		
8 spectra RF=MS, M=7					
(1) AHE	38	.15	.0488-.0732	.0379	.07
(2) OHC	36	.22	.0976-.1220	.0324	.065
OHC-R	39	.115	.0244-.0488		
(3) CHS	36	.22	.0976-.1220	.0162	.04
CHS-R	39	.115	.0244-.0488		
16 spectra RF=MS, M=7					
(1) AHE	37	.18	.0732-.0976	.1405	.20
(2) OHC	36	.22	.0976-.1220	.0756	.12
OHC-R	38	.15	.0488-.0732		
(3) CHS	36	.22	.0976-.1220	.0216	.05
CHS-R	39	.115	.0224-.0488		

Table 4.8b. Comparison of a Measure of the Expected Miss Probability with the Test Miss Rate for the Space of 6 Spectra with 8 Components/Spectrum

Situation	TRAINING SAMPLE			TEST SAMPLE	
	Number of Blocks	95% Upper Tolerance Limit on P_m	$E\{P_m\}$	Miss Rate	95% Upper Confidence Limit on the Miss Rate
RF=MS, M=7					
(1) AHE	32	.32	.1951-.2195	.2972	.37
(2) OHC					
OHC-R	32	.32	.1951-.2195	.1837	.25
(3) CHS	32	.32	.1951-.2195		
CHS-R	35	.24	.1220-.1463	.1243	.18
JD=MS, M=7					
(1) AHE	18	.68	.5366-.5610	.3162	.38
(2) OHC	20	.64	.4878-.5122		
OHC	20	.64	.4878-.5122	.2991	.37
(3) CHS	20	.64	.4878-.5122		
CHS-R	20	.64	.4878-.5122	.2991	.37
RF=MS, M=12					
(1) AHE	36	.22	.0976-.1220	.2000	.26
(2) OHC	32	.32	.1951-.2195		
OHC-R	38	.15	.0488-.0732	.0650	.11
(3) CHS	32	.32	.1951-.2195		
CHS-R	37	.18	.0732-.0976	.0595	.10
JD=MS, M=12					
(1) AHE	33	.30	.1707-.1951	.1710	.225
(2) OHC	20	.64	.4878-.5122		
OHC-R	34	.27	.1463-.1701	.1625	.22
(3) CHS	20	.64	.4878-.5122		
CHS-R	35	.24	.1220-.1463	.1368	.19

Let M_{AV} denote the average test miss rate for the 7 cases in Tables 8a and 8b. Let $E\{P_m\}_{AV}$ be equal to $(E^+ + E^-)/2$ for the same 7 cases. For the OHC and the CHS procedures, the following results are obtained.

$$\text{OHC} \qquad M_{AV} = .48 \quad E\{P_m\}_{AV}$$

$$\text{CHS} \qquad M_{AV} = .40 \quad E\{P_m\}_{AV}$$

These are compared with the following results for OHC-R and CHS-R procedures.

$$\text{OHC-R} \qquad M_{AV} = .84 \quad E\{P_m\}_{AV}$$

$$\text{CHS-R} \qquad M_{AV} = .68 \quad E\{P_m\}_{AV} \ .$$

The result for the AHE procedure is

$$M_{AV} = .94 \quad E\{P_m\}_{AV} \ .$$

It is therefore concluded that for a realistic estimate of the expected probability of a miss that the radii of the hyperspheres making up R_2 should be used in ordering the P_2 observations. Hence from this point on we will only consider the results for the AHE, OHC-R, and CHS-R procedures.

It should be noted that the lowest expected miss rate that can be achieved with the 40 main speaker training sample occurs for $E^- = .022$ and $E^+ = .0488$. This occurs when 39 P_2 blocks are contained in R_2. Note that 39 P_2 blocks are counted in R_2 for the OHC-R and CHS-R procedures in the space consisting of 8 spectra with 32 components per

spectrum and for the CHS-R procedure in the space consisting of 16 spectra with 16 components per spectrum. In both spaces the test miss rate is lower than expected. This is part of the reason why the P_2 ordering procedure gives pessimistic results for a classifier of the CHS design. Other important results are:

(2) All test miss rates lie in the 95% upper tolerance limit.

(3) In 17 out of 21 cases we are 95% confident that the true miss rate lies within the limit obtained from the 95% upper tolerance regions.

(4) (a) In 19 out of 21 cases the test miss rate is less than $E^+ + \sigma^-$.

(b) In 16 out of 21 cases the test miss rate is less than E^+.

(c) In 6 out of 21 cases the test miss rate, M, satisfies $E^- < M < E^+$.

(d) In 16 out of 21 cases the test miss rate satisfies $E^- - \sigma^+ < M < E^+ + \sigma^+$.

4.11. 7040-7094 Computer Execution Time

The approximate 7040-7094 DCS execution times for the speaker verification experiments are discussed in this section. The execution time for the training phase is approximately equal to the time for calculation of the spectra plus the time for calculation of the distances between the P_1 and P_2 observations and for ordering them. Let " represent seconds in the following equations. Let MS be the number of main speaker training observations and IM be the number of impostor training observations. First, consider the space consisting of 6 spectra with 8

components per spectrum. The training time is given by

$$\text{Training Time} \cong 4.4''*(MS+IM) + .035''*(MS*IM)$$

For 40 main speaker training observations and 208 importor training observations, a training time of $23\frac{1}{2}$ minutes was required. During the test phase, it is only necessary to calculate the spectra for the test utterance and to compare these spectra with those of the main speaker training utterances. Let UT be the number of speakers which are tested. Then the testing time is given by

$$\text{Testing Time} \cong (4.4'' + .035''*MS)*UT .$$

For one test speaker and 40 main speaker training observations, the testing time is 5.8 seconds. For the 624 test speakers, a total test time of approximately 60 minutes was required.

For the space consisting of 4 spectra with 64 components per spectrum,the following equations hold.

$$\text{Training Time} \cong 2.9''*(MS+IM) + .175''*(MS*IM)$$

$$\text{Testing Time} \cong (2.9'' + .175''*MS)*UT$$

For 40 main speaker training observations and 208 impostor training observations, a training time of $36\frac{1}{2}$ minutes was required. For 624 test speakers, a total test time of approximately 72 minutes was required. The time required to test one speaker was approximately 9.9 seconds.

For the space consisting of 8 spectra with 32 components per spectrum, the following equations hold.

$$\text{Training Time} \cong 4.9'' * (MS + IM) + .175'' * (MS * IM)$$

$$\text{Testing Time} \cong (4.9'' + .175'' * MS) * UT$$

For 40 main speaker training observations and 208 impostor training observations, a training time of approximately 45 minutes was required. For 624 test speakers, the total test time was approximately 120 minutes. The time required to test one speaker was approximately 11.9 seconds.

For a space consisting of 16 spectra with each spectrum having 16 components, the following equations hold.

$$\text{Training Time} \cong 7.4'' * (MS + IM) + .175'' * (MS * IM)$$

$$\text{Testing Time} \cong (7.4'' + .175'' * MS) * UT$$

For 40 main speaker training observations and 208 impostor training observations, a training time of 55 minutes was required. For 624 test speakers, the total test time was approximately 150 minutes. The time required to test one speaker was approximately 14.4 seconds.

On the whole, these execution times seem satisfactory for a practical automatic speaker verification system. However, the IBM 7094 computer is probably larger and faster than a computer which is likely to be used in an automatic speaker verification system.

Chapter 5

THEORETICAL COMPARISON OF THE PROBABILITY OF ERROR FOR THE AHE-DFTR PROCEDURE WITH THE PROBABILITY OF ERROR FOR THE NEAREST NEIGHBOR RULE

5.1. Summary

In this chapter the small sample performance of the AHE-DFTR procedure is investigated and compared to the performance of the nearest-neighbor rule for the two-class situation.

Suppose n_1 independent observations are available from class 1 and n_2 independent observations are available from class 2, both on the real line. Let P_{FA}, the false alarm probability, be the conditional probability that a new observation V is classified into class 2 given that V belongs to class 1. Let P_M, the miss probability, be the conditional probability that V is classified into class 1 given that V belongs to class 2. See equations 1.5 and 1.6. The following results are obtained without specifying the class probability distributions.

(1) The false alarm probability and the miss probability are derived for the m-block AHE-DFTR procedure when n_1 independent observations are available from class 1 and one observation is available from class 2.

(a) For the above conditions the nearest-neighbor (NN) rule

and the one-block DFTR procedures yield identical miss probabilities.

(b) This leads to the result that under these conditions the miss probability for the m-block AHE-DFTR procedure, $m > 1$, is less than or equal to the miss probability for the NN procedure.

(2) For $n_2 > 1$ independent observations from class 2 and n_1 independent observations from class 1 an intuitive comparison is made of the miss probabilities for the NN and AHE-DFTR procedures. The result is that the miss probability for the NN rule is less than or equal to the miss probability for the one-block AHE-DFTR procedure.

(3) For one observation from class 2 and one observation from class 1 it is shown that th DFTR false alarm probability P_{FA} is equal to one-half. Furthermore, for two independent observations from class 2 and one observation from class 1 it is shown that $P_{FA} = \frac{1}{2}$. This indicates that the number of class 2 observations does not affect P_{FA}, a result expected from a consideration of distribution-free tolerance region theory.

(4) It is further shown that the false alarm probability for the DFTR rule can be less than the false alarm probability for the nearest-neighbor rule.

The results for P_{FA} for the DFTR procedures hold regardless of the dimensionality of the space. The comparison of P_M for the NN rule with P_M for the DFTR procedure holds regardless of the dimensionality of the space as long as the same metric is used for both procedures.

5.2. The Nearest-Neighbor Rule

Under the nearest-neighbor rule a new observation is classified into the class of its nearest neighbor. The nearest neighbor is that observation which is closest to the new observation in some metric distance. The metric distance used in this chapter is Euclidean distance.

Suppose there exist two classes, class 1 described by the probability distribution $F_1(z)$, and class 2, described by $F_2(z)$. Let z be a variable on the real line. Suppose one observation is available from each class. Let the observations be denoted by X_1 and Y_1, where X_1 is from the population with distribution F_2 and Y_1 is from the population with distribution F_1. Let V be a new observation. Now let P_{FA} be the conditional probability that V is classified into class 2 given that V belongs to class 1 and let P_M be the conditional probability that V is classified into class 1 given that V belongs to class 2. Also let

$$Z_1 = |X_1 - V| \tag{5.1}$$

and

$$W_1 = |Y_1 - V| . \tag{5.2}$$

Then P_M is the probability that $W_1 < Z_1$ given $V \sim F_2$, where $V \sim F_2$ denotes that V is from the probability distribution F_2. Likewise, P_{FA} is the probability that $W_1 > Z_1$ given $V \sim F_1$. The probability distribution of Z_1 given that $V = v$ is

$$F_{Z_1}(z_1/v) = \begin{cases} F_2(v+z_1) - F_2(v-z_1) & z_1 > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{5.3}$$

The probability density of Z_1 given that $V = v$ is

$$f_{Z_1}(z_1/v) = \begin{cases} f_2(v+z_1) + f_2(v-z_1) & z_1 \geq 0 \\ 0 & \text{otherwise} . \end{cases} \tag{5.4}$$

A similar distribution and density are obtained for W_1, where W_1 is substituted for Z_1, w_1 is substituted for z_1, and 1 is substituted for 2. Now let

$$U = Z_1 - W_1 . \tag{5.5}$$

The miss probability is the probability that $U > 0$ given $V \sim F_2$.

The probability distribution for U given $V = v$ is

$$F_U(u/v) = \int_0^\infty dw_1 \int_0^{u+w_1} dz_1 f_{Z_1 W_1}(z_1, w_1/v) . \tag{5.6}$$

Since Z_1 and W_1 are independent given $V = v$,

$$F_U(u/v) = \int_0^\infty dw_1 F_{Z_1}(u+w_1/v) f_{W_1}(w_1/v) . \tag{5.7}$$

The miss probability is given by

$$P_M = \int_{-\infty}^\infty f_2(v)dv \, [1 - F_U(0/v)] \tag{5.8}$$

where

$$F_U(0/v) = \int_0^\infty dw_1 [F_2(v+w_1) - F_2(v-w_1)][f_1(v+w_1) + f_1(v-w_1)] .$$

Since

$$\int_0^\infty dw_1 [f_1(v+w_1) + f_1(v-w_1)] = 1 \tag{5.9}$$

the miss probability is given by

$$P_M = \int_{-\infty}^\infty f_2(v)dv \int_0^\infty dw_1 [1 - F_2(v+w_1) + F_2(v-w_1)][f_1(v+w_1) + f_1(v-w_1)] \tag{5.10}$$

A similar analysis can be made for the false alarm probability with the result

$$P_{FA} = \int_{-\infty}^{\infty} f_2(v)dv \int_0^{\infty} dw_1 [F_2(v+w_1) - F_2(v-w_1)][f_1(v+w_1) + f_1(v-w_1)] \qquad (5.11)$$

Equivalent results for P_M and P_{FA} can be obtained by letting $U = W_1 - Z_1$. These results are listed below because they are used in a later analysis.

$$P_M = \int_{-\infty}^{\infty} f_2(v)dv \int_0^{\infty} dz_1 [F_1(v+z_1) - F_1(v-z_1)][f_2(v+z_1) + f_2(v-z_1)] \qquad (5.12)$$

$$F_A = \int_{-\infty}^{\infty} f_1(v)dv \int_0^{\infty} dz_1 [1-F_1(v+z_1) + F_1(v-z_1)][f_2(v+z_1) + f_2(v-z_1)] \qquad (5.13)$$

The fact that equations 5.12 and 5.10 and equations 5.13 and 5.11 are respectively identical can be seen by integrating equations 5.10 and 5.11 by parts and using equation 5.9.

Suppose n_2 statistically independent observations $X_1, \ldots, X_{n_2}$ are available from class 2 and n_1 statistically independent observations $Y_1, \ldots, Y_{n_1}$ are available from class 1. Let V be a new observation. Now let

$$Z_i = |X_i - V| \qquad i = 1, \ldots, n_2 \qquad (5.14)$$

and

$$W_i = |Y_i - V| \qquad i = 1, \ldots, n_1 \qquad (5.15)$$

Also let

$$Z = \min_{1 \le i \le n_2} Z_i \qquad (5.16)$$

and

$$W = \min_{1 \le i \le n_1} W_i \qquad (5.17)$$

P_M is the probability that $W < Z$ given $V \sim F_2$. P_{FA} is the probability that $W > Z$ given $V \sim F_1$.

Consider the problem of finding the probability distribution for Z when $Z = \min(Z_1, Z_2)$. Figure 5.1 shows the region for $\min(z_1, z_2) \leq z$. Then

$$F_Z(z) = F_{Z_1}(z) + F_{Z_2}(z) - F_{Z_1 Z_2}(z, z). \tag{5.18}$$

If the variables Z_1 and Z_2 are independent and identically distributed with distribution F_{Z_1}, the probability distribution for Z is given by

$$F_Z(z) = 1 - [1 - F_{Z_1}(z)]^2. \tag{5.19}$$

Suppose $Z = \min(Z_1, Z_2, Z_3)$. Then

$$F_Z(z) = F_{Z_1}(z) + F_{Z_2}(z) + F_{Z_3}(z) - F_{Z_1 Z_2}(z, z) - F_{Z_1 Z_3}(z, z) - F_{Z_2 Z_3}(z, z) + F_{Z_1 Z_2 Z_3}(z, z, z). \tag{5.20}$$

If the variables Z_1, Z_2, and Z_3 are independent and identically distributed with distribution F_{Z_1},

$$F_Z(z) = 1 - [1 - F_{Z_1}(z)]^3 \tag{5.21}$$

These equations are easily extended for $Z = \min_{1 \leq i \leq n_2} Z_i$.

Again consider finding the distributions of Z and W given in equations 5.14 through 5.17. The probability distribution for Z_1, given $V = v$, is given by equation 5.3. Since $Z_1, Z_2, \ldots, Z_{n_2}$ are independent

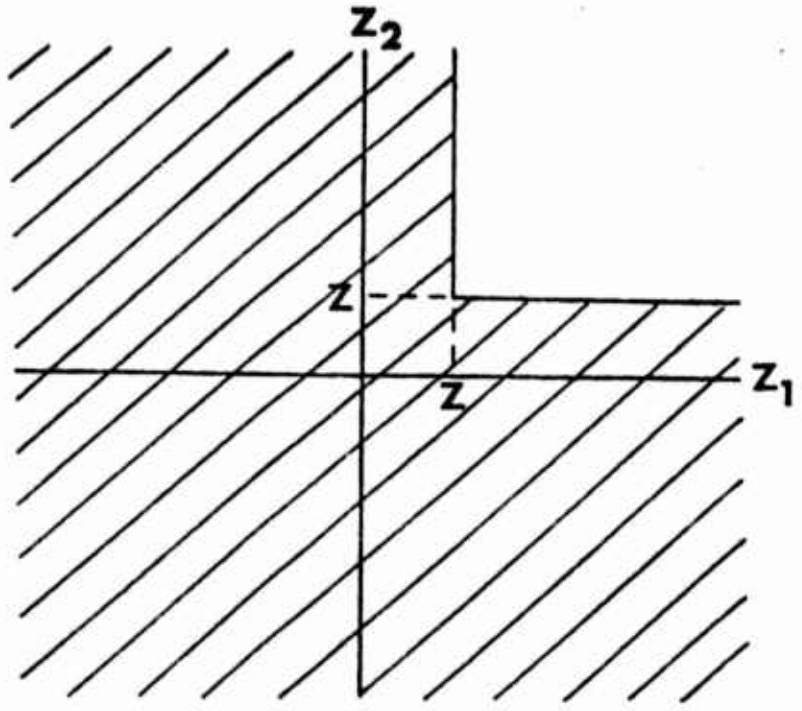


Figure 5.1. The Transformation $Z = \min(Z_1, Z_2)$.

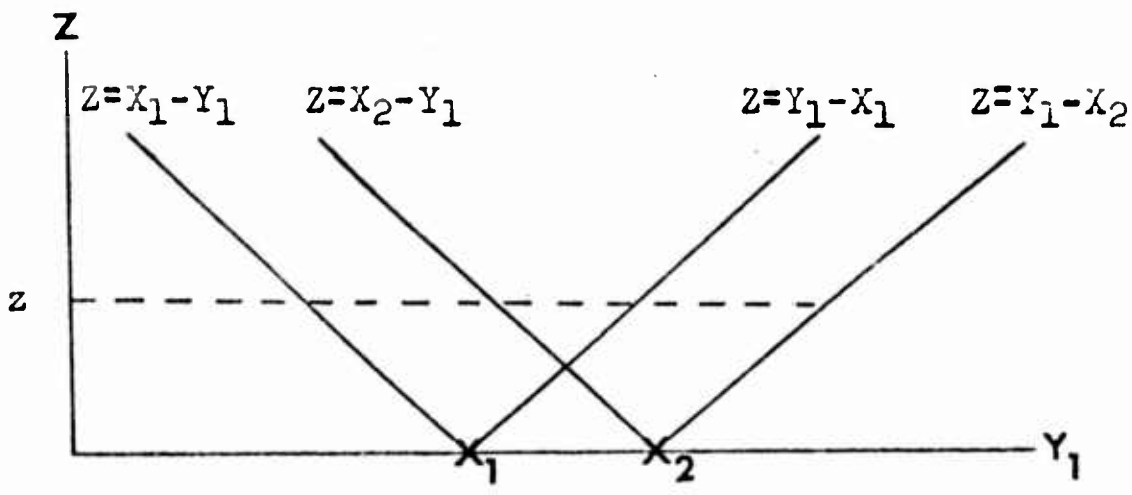


Figure 5.2. The Transformation $Z=\min(|Y_1-X_1|,|Y_1-X_2|)$.

and identically distributed given $V = v$, the probability distribution of Z given $V = v$ is

$$F_Z(z/v) = 1 - [1 - F_{Z_1}(z/v)]^{n_2} . \tag{5.22}$$

The probability density for W in equation 5.17 is easily seen to be

$$f_W(w/v) = n_1[1 - F_{W_1}(w/v)]^{n_1-1} f_{W_1}(w/v) \tag{5.23}$$

Let $U = Z - W$. Then

$$F_U(u/v) = \int_0^\infty dw \int_0^{u+w} dz\, f_{W,Z}(w, z/v) \tag{5.24}$$

and

$$P_M = \int_{-\infty}^{\infty} f_2(v)\, dv\, [1 - F_U(0/v)] \tag{5.25}$$

When the proper substitutions are made, one finds

$$P_M = n_1 \int_{-\infty}^{\infty} f_2(v)\, dv \int_0^\infty dw[1 - F_2(v+w) + F_2(v-w)]^{n_2} *$$

$$[1 - F_1(v+w) + F_1(v-w)]^{n_1-1} [f_1(v+w) + f_1(v-w)] . \tag{5.26}$$

By a similar procedure one obtains

$$P_{FA} = n_2 \int_{-\infty}^{\infty} f_1(v)\, dv \int_0^\infty dz[1 - F_1(v+z) + F_1(v-z)]^{n_1} *$$

$$[1 - F_2(v+z) + F_2(v-z)]^{n_2-1} [f_2(v+z) + f_2(v-z)] . \tag{5.27}$$

5.3. The AHE-DFTR Procedure

Now consider the AHE-DFTR procedure. Suppose one observation is available from each class. Let $V = v$ be a new observation. By the

AHE-DFTR procedure the observation V is assigned to class 1 if

$$|V - X_1| > |Y_1 - X_1| \quad . \tag{5.28}$$

Let

$$Z = |Y_1 - X_1| \tag{5.29}$$

and

$$W = |V - X_1| \, . \tag{5.30}$$

Then a miss error is made when $W > Z$ given $V \sim F_2$. Let $U = W - Z$, then a miss error is made when $U > 0$ given $V \sim F_2$. Proceeding as before with $V \sim F_2$,

$$F_U(u/x_1) = \int_0^\infty F_W(u+z/x_1)\, f_Z(z/x_1)dz \, , \tag{5.31}$$

where

$$F_W(u+z/x_1) = \begin{cases} F_2(x_1+u+z) - F_2(x_1-u-z) & w \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$f_Z(z/x_1) = \begin{cases} f_1(x_1+z) + f_1(x-z) & z \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$P_M = \int_{-\infty}^{\infty} f_2(x_1)\, dx_1 \int_0^\infty [1 - F_2(x_1+z) + F_2(x_1-z)][f_1(x_1+z) + f_1(x_1-z)]\, dz \, . \tag{5.32}$$

This equation is identical to equation 5.10. Therefore, given one sample from each class, the AHE-DFTR procedure has the same miss probability as the NN rule.

Consider the error when V is classified into class 2 and $V \sim F_1$.

Now

$$F_W(w/x_1) = \begin{cases} F_1(x_1+w) - F_1(x_1-w) & w \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad (5.33)$$

and

$$F_U(u/x_1) = \int_0^\infty [F_1(x_1+u+z) - F_1(x_1-u-z)][f_1(x_1+z) + f_1(x_1-z)]\,dz, \qquad (5.34)$$

and

$$P_{FA} = \int dx_1\, f_2(x_1) F_U(0/x_1)\,. \qquad (5.35)$$

Then

$$P_{FA} = \int_{-\infty}^{\infty} dx_1 f_2(x_1) \int_0^\infty [F_1(x+z) - F_1(x_1-z)][f_1(x_1+z) + f_1(x_1-z)]\,dz\,. \qquad (5.36)$$

Noting that the last integral is of the form $\int u\,du$, the following result is obtained.

$$P_{FA} = \frac{1}{2} \qquad (5.37)$$

This is as expected from the DFTR theory since one block is used to form region R_2.

n_1 – Class 1 Observations, 1 – Class 2 Observation

Now suppose n_1 independent identically distributed observations are available from class 1 and one observation is available from class 2. Let

$$Z_i = |Y_i - X_1| \qquad i = 1, \ldots, n_1 \qquad (5.38)$$

and

$$W_1 = |V - X_1| \qquad (5.39)$$

Let

$$Z = \min_{1 \leq i \leq n_1} Z_i \, . \tag{5.40}$$

Suppose $V \sim F_2$. Then a miss error is made when $W > Z$. Let $U = W - Z$. Then

$$F_{Z_i}(z_i/x_1) = \begin{cases} F_1(x_1 + z_i) - F_1(x_1 - z_i) & z_i \geq 0 \quad i = 1, \ldots, n_1 \\ 0 & \text{otherwise} \end{cases} \tag{5.41}$$

and

$$F_W(w/x_1) = \begin{cases} F_2(x_1 + w) - F_2(x_1 - w) & w \geq 0 \\ 0 & \text{otherwise} \, . \end{cases} \tag{5.42}$$

Since the Z_i conditioned on X_1 are independent, it is seen from equation 5.21 that

$$F_Z(z/x_1) = 1 - [1 - F_{Z_1}(z/x_1)]^{n_1} \, . \tag{5.43}$$

$F_U(u/x_1)$ is equal to

$$F_U(u/x_1) = \int_0^\infty dz \int_0^{u+z} dw \, f_W(w/x_1) \, f_Z(z/x_1) \, . \tag{5.44}$$

Substituting the density for $f_Z(z/x_1)$ into equation 5.44 and integrating over w , it is found that

$$F_U(u/x_1) = \int_0^\infty dz \, F_W(u + z/x_1) \, n_1 [1 - F_{Z_1}(z/x_1)]^{n_1 - 1} f_{Z_1}(z/x_1) \, . \tag{5.45}$$

Substituting $u = 0$ in equation 5.45 and noting that

$\int_0^\infty n_1 [1 - F_{Z_1}(z/x_1)]^{n_1 - 1} f_{Z_1}(z/x_1) \, dz = 1$ one obtains

$$P_M = \int_{-\infty}^{\infty} dx_1 f_2(x_1) \int_0^{\infty} dz[1 - F_2(x_1+z) + F_2(x_1-z)] *$$

$$n_1[1 - F_1(x_1+z) + F_1(x_1-z)]^{n_1-1} [f_1(x_1+z) + f_1(x_1-z)]. \qquad (5.46)$$

Note that equation 5.26 with $n_2 = 1$ is identical to equation 5.46. Therefore, for n_1 independent observations from class 1 and one observation from class 2, the NN rule and the one block DFTR procedure give identical miss probabilities.

Let region R_2 be the union of more than one block $(m > 1)$. Region R_2 becomes larger as each block is added. Therefore, for $m > 1$, n_1 class 1 observations, and one class 2 observation, the AHE miss probability is less than or equal to the NN miss probability,

$$P_M^{DFTR} \leq P_M^{NN} . \qquad (5.47)$$

We next consider the false alarm probability, P_{FA}. Now

$$F_W(w/x_1) = \begin{cases} F_1(x_1+w) - F_1(x_1-w) & w \geq 0 \\ 0 & \text{otherwise} . \end{cases} \qquad (5.48)$$

Then

$$F_U(0/x_1) = \int_0^{\infty} dz\,[F_2(x_1+u+z) - F_2(x_1-u-z)]\, n_1$$

$$[1 - F_1(x_1+z) + F_1(x_1-z)]^{n_1-1} [f_1(x_1+z) + f_1(x_1-z)]. \qquad (5.49)$$

Integrating by parts one obtains

$$P_{FA} = \frac{1}{n_1+1} . \qquad (5.51)$$

This is the result predicted by the DFTR theory since one block out of

the possible n_1+1 blocks is used to form region R_2.

Let region R_2 be the union of m blocks where the AHE ordering procedure is used. Let Z be the m^{th} smallest of the Z_i, $i = 1, \ldots, n_1$. Denote this by

$$Z = \min_{1 \le i \le n_1} (m)\, Z_i . \tag{5.52}$$

From the theory of order statistics, (c.f. Kendall and Stuart (1958), p. 252) the probability density for Z is given by

$$f_Z(z) = \frac{n_1!}{(m-1)!\,(n_1-m)!}\, F_{Z_1}(z)^{m-1}[1 - F_{Z_1}(z)]^{n_1-m} f_{Z_1}(z) . \tag{5.53}$$

Substituting equation 5.53 into equation 5.35 one obtains P_M when m blocks are used to form R_2 by the AHE procedure.

$$P_M(m) = \int_{-\infty}^{\infty} dx_1 f_2(x_1) \int_0^{\infty} dz[1 - F_2(x_1+z) + F_2(x_1-z)]\, \frac{n_1!}{(m-1)!\,(n_1-m)!}$$
$$[F_1(x_1+z) - F_1(x_1-z)]^{m-1}[1 - F_1(x_1+z) + F_1(x_1-z)]^{n_1-m}[f_1(x_1+z) + f_1(x_1-z)] \tag{5.54}$$

Similarly, P_{FA} for m blocks is

$$P_{FA}(m) = \int_{-\infty}^{\infty} dx_1 f_2(x_1) \int_0^{\infty} dz[F_1(x_1+z) - F_1(x_1-z)]\, \frac{n_1!}{(m-1)!\,(n_1-m)!}$$
$$[F_1(x_1+z) - F_1(x_1-z)]^{m-1}[1 - F_1(x_1+z) + F_1(x_1-z)]^{n_1-m}[f_1(x_1+z) + f_1(x_1-z)]. \tag{5.55}$$

The integral over z is simply the expected value of a random variable from a Beta distribution with parameters m and $n_1 - m + 1$. Then m-block false alarm probability is equal to

$$P_{FA}(m) = \frac{m}{n_1+1} \quad . \tag{5.56}$$

DFTR Procedure for $n_2 > 1$

We now use the same techniques to investigate the DFTR probability of error when more than one class 2 observation is used. The equations become very complex and can not be carried through in general (i.e. without assuming some probability density function). This section is included only to show some of the difficulty which is involved. An intuitive investigation into the problem is made in the next section.

Suppose two independent observations X_1 and X_2 are available from class 2 and one observation Y_1 is available from class 1. Let X_1 and X_2 be given, with $X_1 < X_2$. Let

$$\begin{aligned} Z_i &= |Y_1 - X_i| \qquad i = 1,2 \\ Z &= \min(Z_1, Z_2) \\ W_1 &= |V - X_i| \qquad i = 1,2 \\ W &= \min(W_1, W_2) \end{aligned} \tag{5.57}$$

The distribution of Z given $X_1 < X_2$ is, from equation 5.18,

$$F_Z(z) = F_{Z_1}(z) + F_{Z_2}(z) - F_{Z_1Z_2}(z,z). \tag{5.58}$$

$F_{Z_1Z_2}(z,z)$ has a nonzero value only when $z > \frac{x_2 - x_1}{2}$. $F_{Z_1Z_2}(z,z)$ is equal to the F_1 distribution in the region of overlap of the intervals expanding from each x.

Since $X_1 < X_2$

$$F_{Z_1 Z_2}(z, z/x_1 < x_2) = \begin{cases} F_1(x_1+z) - F_1(x_2-z) & z > \dfrac{x_2 - x_1}{2} \\ 0 & \text{otherwise.} \end{cases} \quad (5.59)$$

Then

$$F_Z(z/x_1 < x_2) = \begin{cases} F_1(x_2+z) - F_1(x_1-z) & z \geq \dfrac{x_2 - z_1}{2} \\ F_1(x_2+z) - F_1(x_2-z) + F_1(x_1+z) - F_1(x_1-z) & 0 \leq z \leq \dfrac{x_2 - x_1}{2} \\ 0 & \text{otherwise.} \end{cases} \quad (5.60)$$

The above equation could have been obtained directly from Figure 5.2 .

Error E_1 is obtained when $W < Z$ and $V \sim F_1$. In this case

$$F_W(w/x_1 < x_2) = \begin{cases} F_1(x_2+w) - F_1(x_1-w) & w \geq \dfrac{x_2 - x_1}{2} \\ F_1(x_2+w) - F_1(x_2-w) + F_1(x_1+w) - F_1(x_1-w) & 0 \leq w \leq \dfrac{x_2 - x_1}{2} \\ 0 & \text{otherwise.} \end{cases} \quad (5.61)$$

Let

$$U = W - Z \quad (5.62)$$

and

$$\frac{dF_Z(z/x_1 < x_2)}{dz} = \begin{cases} f_Z^+ & z \geq m \\ f_Z^- & 0 \leq z \leq m \end{cases}$$

$$\frac{dF_W(w/x_1 < x_2)}{dw} = \begin{cases} f_W^+ & w \geq m \\ f_W^- & 0 \leq w \leq m \end{cases} \quad (5.63)$$

where

$$m = \frac{x_2 - x_1}{2}$$

From Figure 5.3 the distribution for $U \leq 0$ given $x_1 < x_2$ is

$$F_U(0/x_1 < x_2) = \int_0^m dz \int_0^z dw\, f_W^- f_Z^- + \int_m^\infty dz \int_0^m dw\, f_Z^+ f_W^- + \int_m^\infty dz \int_m^\infty dw\, f_Z^+ f_W^+ \tag{5.64}$$

These integrals are easily solved. Consider the first double integral.

$$\int_0^m dz \int_0^z dw\, f_W^- f_Z^- = \int_0^m dz [F_1(x_2+z) - F_1(x_2-z) + F_1(x_1+z) - F_1(x_1-z)] * [f_1(x_2+z) + f_1(x_2-z) + f_1(x_1+z) + f_1(x_1-z)] \tag{5.65}$$

Recognizing that this integral is of the form $\int u\,du$ one obtains

$$\int_0^m dz \int_0^z dw\, f_W^- f_Z^- = \frac{1}{2}[F_1(x_2+m) - F_1(x_2-m) + F_1(x_1+m) - F_1(x_1-m)]^2 \tag{5.66}$$

Consider the second double integral of equation 5.64. By straightforward integration one obtains

$$\int_m^\infty dz \int_0^m dw\, f_Z^+ f_W^- = [F_1(x_2+m) - F_1(x_2-m) + F_1(x_1+m) - F_1(x_1-m)] * [1 - F_1(x_2+m) + F_1(x_1-m)] . \tag{5.67}$$

Now consider the final double integral of equation 5.64. Integrating with respect to w one obtains

$$\int_m^\infty dz \int_m^z dw\, f_Z^+ f_W^+ = \int_m^\infty dz\, [F_1(x_2+z) - F_1(x_1-z) \cdot F_1(x_2+m) + F_1(x_1-m)] * [f_1(x_2+z) + f_1(x_1-z)] . \tag{5.68}$$

Let

$$u = F_1(x_2+z) - F_1(x_1-z)$$

$$du = f_1(x_2+z) + f_1(x_1-z) .$$

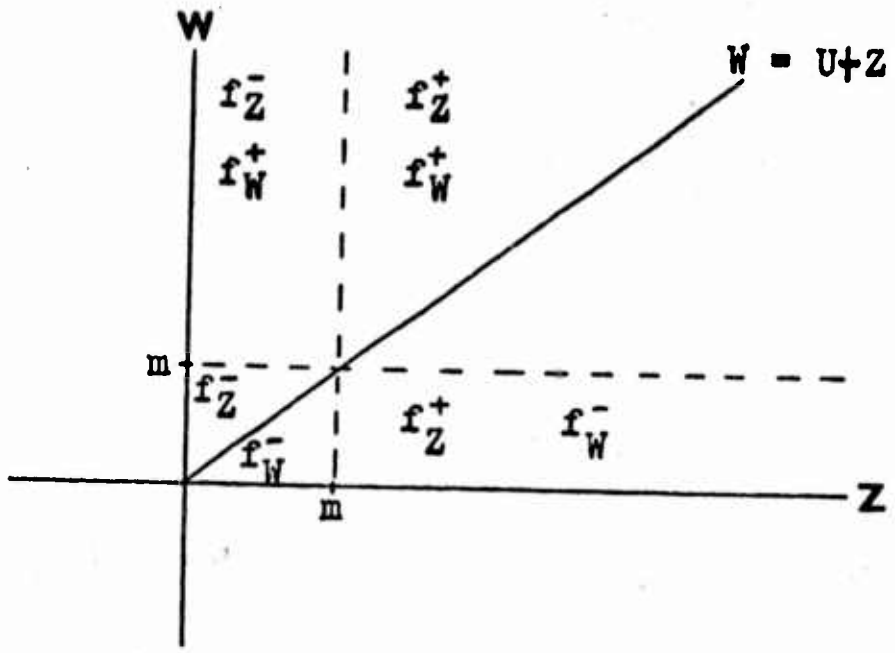


Figure 5.3. The Transformation for Equation 5.64.

Then the integral becomes

$$\int_m^\infty dz \int_m^z dw f_Z^+ f_W^+ = \int_{F_1(x_2+m) - F_1(x_1-m)}^1 u\,du - [F_1(x_2+m) - F_1(x_1-m)][1 - F_1(x_2+m) + F_1(x_1-m)] \tag{5.69}$$

After integrating and collecting terms, one obtains

$$\int_m^\infty dz \int_m^z dw f_Z^+ f_W^+ = \frac{1}{2}[1 - F_1(x_2+m) + F_1(x_1-m)]^2 . \tag{5.70}$$

Let

$$F_Y = F_1(x_2+m) - F_1(x_1-m)$$

$$F_X = F_1(x_1+m) - F_1(x_2-m) . \tag{5.71}$$

Then

$$F_U(0/x_1 < x_2) = \frac{1}{2}\{(F_X+F_Y) + (1-F_X)^2 + 2(F_X+F_Y)(1-F_X)\}$$

$$= \frac{1}{2}(F_Y+1)^2 .$$

Note that

$$x_1+m = x_1 + \frac{x_2-x_1}{2} = \frac{x_1+x_2}{2}$$

and

$$x_2 - m = x_2 - \frac{x_2-x_1}{2} = \frac{x_1+x_2}{2} .$$

Therefore

$$F_Y = F_1\left(\frac{x_1+x_2}{2}\right) - F_1\left(\frac{x_1+x_2}{2}\right) = 0 . \tag{5.72}$$

Then

$$F_U(0/x_1 < x_2) = \frac{1}{2} \tag{5.73}$$

and

$$P_{FA} = 2\int_{-\infty}^{\infty} dx_2\, f_2(x_2) \int_{-\infty}^{x_2} dx_1\, f_2(x_1)\,\frac{1}{2}$$

$$= \frac{1}{2} . \tag{5.74}$$

This is, of course, the answer which was expected from a consideration of DFTR theory. This answer would be expected regardless of the number of class X observations used to design the classifier.

We now wish to obtain an expression for the miss probability for one class 1 observation and two class 2 observations. $F_W(w)$ is now given by

$$F_W(w) = \begin{cases} F_2(x_2+w) - F_2(x_1-w) & w \geq \frac{x_2-x_1}{2} \\ F_2(x_2+w) - F_2(x_2-w) + F_2(x_1+w) - F_2(x_1-w) & 0 \leq w \leq \frac{x_2-x_1}{2} \\ 0 & \text{otherwise.} \end{cases} \quad (5.75)$$

$F_U(0/x_1 < x_2)$ is given by equation 5.64. The first double integral in that equation is equal to

$$\int_0^m dz \int_0^z dw\, f_W^- f_Z^- = \int_0^m dz[F_2(x_2+z) - F_2(x_2-z) + F_2(x_1+z) - F_2(x_1-z)] * [f_1(x_2+z) + f_1(x_2-z) + f_1(x_1+z) + f_1(x_1-z)] . \quad (5.76)$$

The second double integral is equal to

$$\int_m^\infty dz \int_0^m dw\, f_W^- f_Z^+ = [F_2(x_2+m) - F_2(x_2-m) + F_2(x_1+m) - F_2(x_1-m)] * [1 - F_1(x_2+m) + F_1(x_1-m)] . \quad (5.77)$$

The last double integral is equal to

$$\int_m^\infty dz \int_m^z dw\, f_W^+ f_Z^+ = \int_m^\infty dz[f_1(x_2+z) + f_1(x_1-z)][F_2(x_2+z) - F_2(x_1-z) - F_2(x_2+m) + F_2(x_1-m)]$$
$$= -[F_2(x_2+m) - F_2(x_1-m)][1 - F_1(x_2+m) + F_1(x_1-m)]$$
$$+ \int_m^\infty dz[f_1(x_2+z) + f_1(x_1-z)][F_2(x_2+z) - F_2(x_1-z)] . \quad (5.78)$$

Combining terms and substituting $m = \frac{x_2 - x_1}{2}$, one obtains

$$F_U(0/x_1 < x_2) = \int_0^{\frac{x_2 - x_1}{2}} dz[F_2(x_2+z) - F_2(x_2 - z) + F_2(x_1+z) - F_2(x_1 - z)][f_1(x_2+z)$$

$$+ f_1(x_2 - z) + f_1(x_1+z) + f_1(x_1 - z) + \int_{\frac{x_2 - x_1}{2}}^{\infty} dz[F_2(x_2+z) - F_2(x_1 - z)] *$$

$$[f_1(x_2+z) + f_1(x_1 - z)]$$

$$F_U(0/x_1 < x_2) = \int_0^{\infty} dz[F_2(x_2+z) - F_2(x_1 - z)][f_1(x_2+z) + f_1(x_1 - z)]$$

$$+ \int_0^{\frac{x_2 - x_1}{2}} dz[F_2(x_2+z) - F_2(x_1 - z)][f_1(x_1+z) + f_1(x_2 - z)]$$

$$+ \int_0^{\frac{x_2 - x_1}{2}} dz[F_2(x_1+z) - F_2(x_2 - z)][f_1(x_2+z) + f_1(x_1 - z)]$$

$$+ \int_0^{\frac{x_2 - x_1}{2}} dz[F_2(x_1+z) - F_2(x_2 - z)][f_1(x_1+z) + f_1(x_2 - z)]. \quad (5.79)$$

P_M is given by

$$P_M = 2 \int_{-\infty}^{\infty} dx_2 F_2(x_2) \int_{-\infty}^{x_2} dx_1 f_2(x_1)[1 - F_U(0/x_1 < x_2)] . \quad (5.80)$$

This equation has not been reduced so that a meaningful comparison can be made between P_M^{DFTR} and P_M^{NN} without any assumption on the probability distributions. Hence an intuitive comparison is made below of the miss probabilities for the DFTR and NN procedure when n_2 class 2 observations and n_1 class 1 observations are available.

5.4. Intuitive Investigation

Intuitive justification of these results can be easily produced. For example, consider again the case when one observation is available from each class. Let a new observation be denoted by V. The NN rule compares $|V - X_1|$ with $|V - Y_1|$. The DFTR procedure compares $|X_1 - Y_1|$ with $|X_1 - V|$. Let $V \sim F_2$ and replace V by X. Then

$$P_M^{DFTR} = Pr[(|X_1 - Y_1| - |X - X_1|) < 0]$$

and

$$P_M^{NN} = Pr[(|X - Y_1| - |X - X_1|) < 0] \quad . \tag{5.81}$$

Note that the random variable $|X_1 - X| - |X_1 - Y_1|$ conditioned on X_1 has the same statistics as the random variable $|X - X_1| - |X - Y_1|$ conditioned on X. Since X_1 and X are independent identically distributed random variables, P_M for the DFTR rule is equal to P_M for the nearest-neighbor rule, a result which was obtained formally in the first part of this chapter. Now let $V \sim F_1$ and replace V by Y. Then the false alarm probability for the DFTR procedure is

$$P_{FA}^{DFTR} = Pr[(|X_1 - Y_1| - |X_1 - Y|) < 0] = \frac{1}{2} , \tag{5.82}$$

another result previously obtained.

Now consider a problem for which no results were obtained. The problem is to compare P_M for the DFTR rule and P_M for the nearest-neighbor rule when more than one observation is available from class 2. Specifically, suppose two observations are available from

class 2 and one observation is available from class 1. Let $V \sim F_2$ and replace V by X. Then

$$P_M^{DFTR} = Pr\,[\min(|X_1 - Y_1|, |X_2 - Y_1|) - \min(|X - X_1| X - Y_2|) < 0] \quad (5.83)$$

and

$$P_M^{NN} = Pr[\,|X - Y_1| - \min(|X - X_1|, |X - X_2| < 0\,] \quad (5.84)$$

The term immediately preceding the inequality sign is the same in both equations. Note that

$$Pr[\min(|X_1 - Y_1|\,, |X_2 - Y_1|) < 0] \geq Pr[|X - Y_1| < 0\,].$$

The dependency in the terms, $\min(|X_1 - Y_1|\,, |X_2 - Y_1|)$ and $\min(|X - X_1|\,, |X - X_2|)$ in the first equation is through X_1 and X_2. The dependency in the terms $|X - Y_1|$ and $\min(|X - X_1|\,, |X - X_2|)$ of the second equation is through X. Since X, X_1, and X_2 are independent identically distributed random variables, the conclusion is made that

$$P_M^{NN} \leq P_M^{DFTR}.$$

For the general case of n_1 observations from class 1 and n_2 observations from class 2, P_M for the DFTR procedure is

$$P_M^{DFTR} = Pr\,[\min_{\substack{1 \leq i \leq n_2 \\ 1 \leq j \leq n_1}} |X_i - Y_j| - \min_{1 \leq i \leq n_2} |X - X_i| < 0\,] \quad (5.85)$$

and P_M for the nearest-neighbor rule is

$$P_M^{NN} = Pr\left[\min_{1 \le j \le n_1} |X - X_j| - \min_{1 \le i \le n_2} |X - X_i| < 0 \right] \tag{5.86}$$

Therefore, the general conclusion is that for $n_2 > 1$

$$P_M^{NN} \le P_M^{DFTR} . \tag{5.87}$$

This is because there are $(n_2 - 1)n_1$ more random variables of the same distribution from which to find a minimum which is less than zero in the DFTR case than in the NN case. This further leads one to expect a larger difference in P_M^{NN} and P_M^{DFTR} as the number of class 2 observations is increased.

Note that the false alarm probability for the DFTR rule can be smaller than the false alarm probability for the NN rule. For example, suppose two observations are available from each class. Then

$$P_{FA}^{DFTR} = 0.3333 . \tag{6.88}$$

If the class probability densities are univariate normals with equal variances and with the distance between means equal to the variances, Fix and Hodges (1952) obtained

$$P_{FA}^{NN} = 0.4086 . \tag{5.89}$$

Chapter 6

CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK

A study has been made of the application of distribution-free tolerance regions to pattern recognition. Some procedures have been presented here for designing a pattern verification system with a given confidence that the false alarm probability will be less than a desired quantity. These procedures maximize the number of main class training observations which are correctly classified. In addition, a method has been given for obtaining a measure of the miss probability. The procedures have been successfully applied to a speaker verification problem.

The advantages of the hypersphere DFTR classification procedure are the following: (1) The hypersphere DFTR classification procedure gives information about how well the classifier is expected to perform. This is done without any knowledge of the class probability distributions and with only one sample of independent observations from each class. (2) The procedure is able to form very complicated, unconnected decision regions. Hence it is useful when multimodal class probability distributions are involved. (3) The hypersphere DFTR procedure is very easily programmed on a digital computer. (4) The procedure is independent of the dimensionality of the measurement space. (5) It offers automatic reduction of the data which must be stored in the computer.

It should, however, be noted that this is a distribution-free procedure and hence can be expected to be quite inefficient when compared to

a procedure based upon a priori knowledge of the class probability distributions. The hypersphere DFTR procedure is most applicable to the situation in which (1) nothing is known about the probability distributions and (2) information about the expected probability of error is desired without using a test sample. The only requirements for using the hypersphere DFTR approach are (1) a properly labeled sample of independent observations must be available from each class and (2) the class probability distributions must be stationary.

The hypersphere DFTR procedures were applied in an automatic speaker verification experiment. Good results were obtained by the use of many short-term spectra of the word "my". Error rates as low as 1.92% were obtained when a 256-dimensional measurement space was used. In a 48-dimensional measurement space, error rates as low as 4.48% were obtained.

Three different ordering procedures were developed and are described in detail in Chapter 3. They were tested on the speaker verification data. Of the three methods, the CHS procedure gave lower error rates on the average than the OHC procedure and the OHC procedure gave lower error rates on the average than the AHE procedure.

A comparison was made of the false alarm rate which was obtained in the speaker verification tests with the 95% upper tolerance level on the false alarm probability which was predicted with the DFTR approach. All test false alarm rates fell below the 95% upper tolerance limit. The average test false alarm for the 21 different cases studied here was

approximately equal to 0.8 of the average expected false alarm probability predicted from the DFTR approach.

A comparison was made of the test miss rate with a measure of the miss probability that was obtained by using a tolerance regions approach. All test miss rates fell below the 95% upper tolerance limit. For the 21 different cases studied, the average miss rate was equal to 0.94 of the average expected miss rate for the AHE procedure. The average miss rate was equal to 0.84 of the average expected miss rate for the OHC-R procedure. The average expected miss rate was equal to 0.68 of the average expected miss rate for the CHS-R procedure.

Suggestions for Further Work

There are many mathematical questions which were not resolved in Chapter 5. For example no general expression was obtained for the m block miss probability $P_M(m)$ for the AHE procedure for n_2 class 2 observations and n_1 class 1 observations. If this expression, and similar ones for the OHC and CHS procedures could be obtained, $P_M(m)$ for the AHE, OHC, and CHS procedures could be compared. This would probably require assumption of some underlying probability density functions. One would also like to compare the miss probability for the m block AHE procedure with the miss probability for the nearest-neighbor rule. Specific questions for which answers are needed are

(1) Is the total probability of error, $\xi_1 P_{FA} + \xi_2 P_M$, for the DFTR rule ever less than the total probability of error for the nearest-neighbor rule

when the a priori probabilities are $\xi_1 = \frac{n_1}{n_1+n_2}$ and $\xi_2 = \frac{n_2}{n_1+n_2}$?

It appears that the answer to this question may be no. This follows from the fact that the NN procedure uses information about all observations from both classes whereas the DFTR procedure uses information about all observations from class 2 but only information about those observations from class 1 which are used in forming the m blocks of region R_2.

The second question for which one would like an answer is: What is the value of m to guarantee that the miss probability for the DFTR procedure is less than the miss probability for the nearest-neighbor rule? The answer for n_1 class 1 observations and 1 class 2 observations has been obtained. For this case $m = 2$. However, no result has been obtained for $n_2 > 1$. Consider an example where two observations are available from each class. Then the nearest-neighbor miss probability is

$$P_M^{NN} = Pr[\min(|X-X_1|, |X-X_2|) > \min(|X-Y_1|, |X-Y_2|)] \tag{6.1}$$

and the one-block DFTR miss probability is

$$P_M^{DFTR}(1) = Pr[\min(|X-X_1|, |X-X_2|) > \min(|X_1-Y_1|, |X_1-Y_2|, |X_2-Y_1|, |X_2-Y_2|)] \tag{6.2}$$

and the two-block DFTR miss probability is

$$P_M^{DFTR}(2) = Pr[\min(|X-X_1|, |X-X_2|) > \max\{\min(|X_1-Y_1|, |X_2-Y_1|), \min(|X_1-Y_2|, |X_2-Y_2|)\}] . \tag{6.3}$$

It was previously concluded in Chapter 5 that $P_M^{NN} \leq P_M^{DFTR}(1)$. Now compare the NN probability with two-block AHE-DFTR miss probability. Suppose the terms on the right side of the inequalities of equations 6.1 and 6.3 are compared. It is easily seen that

$$\Pr[\min(|X - Y_1|, |X - Y_2|) \leq w] = \int_{-\infty}^{\infty} f_1(y)\,dy \{1 - [1 - F_2(y + w) + F_2(y - w)]^2\} \tag{6.4}$$

and

$$\Pr[\min(|X_1 - Y_1|, |X_2 - Y_1|) \leq w] = \int_{-\infty}^{\infty} f_2(x)\,dx \{1 - [1 - F_1(x + w) + F_1(x - w)]^2\} \tag{6.5}$$

These two terms cannot be compared without some knowledge of the probability distributions. Suppose both distributions are normal with equal variances. Then one would expect little difference in the (6.4) and (6.5). Likewise one would expect little difference in

$$\Pr[\min(|X - Y_1|, |X - Y_2|) \leq w] \quad \text{and} \quad \Pr[\min(|X_1 - Y_2|, |X_2 - Y_2|) \leq w].$$

Since equations 6.3 involved $\max\{\min(|X_1 - Y_1|, |X_2 - Y_1|), \min(|X_1 - Y_2|, |X_2 - Y_2|)\}$, one expects that $P_M^{DFTR}(2) \leq P_M^{NN}$ for two normal distributions with equal variances.

Many of the ordering procedures presented in Chapter 3 were not tested experimentally. For example, the ordering procedure which gives both the expected false alarm probability and the expected miss probability (section 3.10) was not tested. The performance of this procedure would be of considerable interest. In addition, an experimental comparison of

the hypersphere DFTR procedures with classification procedures which were not tested here would be desirable.

Appendix A

THEORY OF DISTRIBUTION-FREE TOLERANCE REGIONS

Summary

This section contains a discussion of the theory of distribution-free tolerance regions (DFTR), especially that theory which can be applied to pattern recognition. At first, the discussion is limited to one dimensional, continuous probability distribution functions. The object is to make the statement that with probability γ $(0<\gamma<1)$ at least $100\beta\%$ $(0<\beta<1)$ of an unknown probability distribution is contained in the interval between certain order statistics.

Next, techniques are considered for the construction of regions in D dimensional space in which at least $100\beta\%$ of an unknown D dimensional distribution is contained with probability γ. Results for discontinuous distributions are then discussed.

A.1. One Dimensional Theory

Suppose X is a random variable with a continuous distribution function $F(x)$. The probability that X is less than or equal to x is denoted by

$$\Pr(X \leq x) = \int_{-\infty}^{x} f(y)\,dy = F(x). \qquad \text{(A-1)}$$

Let the differential form of the above equation,

$$\Pr(x - \delta x \leq X < x) = dF(x)\,, \qquad \text{(A-2)}$$

be called the probability element of X.

Suppose $(X_1, X_2, \ldots, X_n)$ is a sample of n statistically independent observations from a population with continuous distribution $F(x)$. Population is used here in the usual statistical sense to mean the totality of possible outcomes of an experiment. The probability element of the sample is

$$\Pr(x_1 - \delta x_1 \leq X_1 < x_1, \ldots, x_n - \delta x_n \leq X_n < x_n) = \prod_{i=1}^{n} dF(x_i) . \qquad (A-3)$$

Let $Y_i = F(X_i)$, $i = 1, \ldots, n$. The probability element of the Y_i is

$$P(y_1 - \delta y_1 \leq Y_1 < y_1, \ldots, y_n - \delta y_n \leq Y_n < y_n) = \begin{cases} \prod_{i=1}^{n} dy_i & 0 < y_i < 1 \\ 0 & \text{otherwise} . \end{cases}$$

Suppose the observations $X_1, X_2, \ldots, X_n$ are arranged in order of increasing magnitude. The ordered observations, $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ where $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$, are called order statistics of the sample. The intervals $(-\infty, X_{(1)}]$, $(X_{(1)}, X_{(2)}], \ldots, (X_{(n)}, \infty)$ are called sample blocks. The random variables $F(X_{(1)}), F(X_{(2)}) - F(X_{(1)}), \ldots, 1 - F(X_{(n)})$ are called coverages of these blocks. Notice that the coverage of a given block is the amount of probability from the distribution function $F(x)$ in that block. Since $F(x)$ is assumed continuous, $\Pr\{X_{(i-1)} = X_{(i)}\} = 0$, $i = 2, \ldots, n$. Therefore the $\leq$ sign between the order statistics can be replaced by the $<$ sign.

The probability element of the ordered random variables $Y_{(i)} = F(X_{(i)})$, $i = 1, 2, \ldots, n$ is

$$\begin{cases} K\,dy_{(1)}\,dy_{(2)}\cdots dy_{(n)} & 0 < y_{(1)} < y_{(2)} < \cdots < y_{(n)} < 1 \\ 0 & \text{otherwise .} \end{cases} \tag{A-4}$$

To find the value of the constant K, equation (A-4) is integrated over its region of definition. Since $F(x)$ is a continuous nondecreasing function of x, the $y_{(i)}$'s and the $x_{(i)}$'s have the same order. Therefore

$$\int_0^1 \int_0^{y_{(n)}} \cdots \int_0^{y_{(2)}} K\,dy_{(1)}\,dy_{(2)} \cdots dy_{(n)} = \frac{K}{n!} = 1 \tag{A-5}$$

and $K = n!$

Eventually the statement,

$$\Pr\{[F(X_{(s)}) - F(X_{(r)})] \geq \beta\} = \gamma\,, \tag{A-6}$$

is made where r and s are positive integers with $r < s \leq n$. Therefore the marginal, joint distribution of $F(X_{(r)})$ and $F(X_{(s)})$ must be found. The probability element of $Y_{(r)} = F(X_{(r)})$ and $Y_{(s)} = F(X_{(s)})$ is obtained by integrating equation (A-5) with respect to $y_{(1)},\ldots,y_{(r-1)}$ over the region $0 < y_{(1)} < \ldots < y_{(r)}$, with respect to $y_{(r+1)},\ldots,y_{(s-1)}$ over the region $y_{(r)} < \ldots < y_{(s)}$, and with respect to $y_{(s+1)},\ldots,y_{(n)}$ over the region $y_{(s)} < \ldots < y_n$.

$$\Pr\{y_{(r)} - \delta y_{(r)} \leq Y_{(r)} < y_{(r)}\,,\ y_{(s)} - \delta y_{(s)} \leq Y_{(s)} < y_{(s)}\} =$$

$$n!\,dy_{(r)}dy_{(s)} \int_{y_{(s)}}^{1} \cdots \int_{y_{(s)}}^{y_{(s+2)}} dy_{(s+1)} \cdots dy_{(n)} \int_{y_{(r)}}^{y_{(s)}} \cdots \int_{y_{(r)}}^{y_{(r+2)}} dy_{(r+1)}$$

$$\cdots dy_{(s-1)} \int_0^{y_{(r)}} \cdots \int_0^{y_{(2)}} dy_{(1)} \cdots dy_{(r-1)} =$$

$$\frac{n!\,dy_{(r)}dy_{(s)}}{(r-1)!\,(s-r-1)!\,(n-s)!}\, y_{(r)}^{r-1} [y_{(s)} - y_{(r)}]^{s-r-1} [1 - y_{(s)}]^{n-s}$$

$$0 < y_{(r)} < y_{(s)} < 1 \tag{A-7}$$

The probability element of the random variable , $Y_{(s)} - Y_{(r)}$, is desired. Therefore let

$$W = Y_{(s)} - Y_{(r)}$$

and

$$Z = Y_{(r)} .$$

The Jacobian of the transformation is one. Hence, probability element of W and Z is

$$\frac{n!\, dz\, dw\, z^{r-1}}{(r-1)!\,(s-r-1)!\,(n-s)!} w^{s-r-1} [1-w-z]^{n-s} \qquad 0 < z < w+z < 1 .$$

Integrating z over the range $[0,1-w]$, the marginal distribution of $W = F(X_{(s)}) - F(X_{(r)})$ is obtained.

$$\Pr\{w-\delta w \leq W < w\} = \frac{n!\, w^{s-r-1} dw}{(r-1)!\,(s-r-1)!\,(n-s)!} \int_0^{1-w} z^{r-1} (1-w-z)^{n-s}\, dz \qquad (A-8)$$

Letting $z = (1-w)t$, this equation reduces to

$$\Pr\{w-\delta w \leq W < w\} = \frac{n!\, w^{s-r-1} dw(1-w)}{(r-1)!\,(s-r-1)!\,(n-s)!} \int_0^1 (1-w)^{r-1} (1-w)^{n-s} *$$

$$(1-t)^{n-s}\, t^{r-1}\, dt .$$

But

$$\int_0^1 (1-t)^{n-s}\, t^{r-1}\, dt = \frac{(n-s)!\,(r-1)!}{(n-s+r)!} .$$

Therefore

$$\Pr\{w-\delta w \leq W < w\} = \frac{n!}{(s-r-1)!\,(n-s+r)!} w^{s-r-1} (1-w)^{n-s+r}\, dw$$

$$0 \leq w \leq 1 . \qquad (A-9)$$

Equation (A-9) is recognized as the Beta distribution.

The probability that $F(X_{(s)}) - F(X_{(r)})$ is greater than or equal to β is given by

$$\Pr\{[F(X_{(s)}) - F(X)_{(r)})] \geq \beta\} = \int_{\beta}^{1} \frac{n!\ w^{s-r-1}(1-w)^{n-s+r}}{(s-r-1)!\,(n-s+r)!}\, dw$$

$$= 1 - I_{\beta}(s-r,\ n-s+r+1) \qquad \text{(A-10)}$$

The function $I_{\beta}(p, q)$ is called the Incomplete Beta Function and its values are tabulated in the literature [Pearson (1934)]. Notice that $s-r$ appears symmetrically in the above equation. Let $m = n + 1 - (s-r)$ be the number of intervals which are excluded from the region in which we want to contain at least β of the population. Then $\Pr\{[F(X_{(s)}) - F(X_{(r)})] \geq \beta\} = \gamma$ is given by

$$\gamma = \int_{\beta}^{1} \frac{n!}{(n-m)!\,(m-1)!}\, w^{n-m}(1-w)^{m-1}\, dw = I_{\beta}(n+1-m, m)\ . \qquad \text{(A-11)}$$

Given three of the four variables n, m, β, γ one can solve for the fourth. Murphy (1948) has constructed graphs of β versus $n (n \leq 500)$ for various m and for confidences $\gamma = 0.90,\ 0.95,\ 0.99$. Sommerville (1958) has tabulated m for values of $n = 50$ to $n = 1000$; $\beta = 0.50,\ 0.75,\ 0.90,\ 0.95,\ 0.99$; and $\gamma = 0.50,\ 0.75,\ 0.90,\ 0.95,\ 0.99$. If these graphs and tables do not contain the desired values for n, m, β, or γ, one can calculate n by the following approximation due to Scheffé and Tukey (1945),

$$n \sim \left[\frac{1}{4}\, \chi^2_{2m;1-\gamma}\, (1+\beta)/(1-\beta) + \frac{1}{2}\,(m-1)\right] \qquad \text{(A-12)}$$

where $\chi^2_{2m;1-\gamma}$ is the point exceeded with probability $1-\gamma$ for the Chi-squared distribution with $2m$ degrees of freedom. For γ and β in the range (0.9, 1.0) the approximation error is generally less than one tenth of one percent. If one wishes to calculate β for a desired γ, m, and large n the following approximation may be used.

$$\beta \sim \left[\frac{\sqrt{(\chi^2_{2m;1-\gamma} - 2m^2 + 16n(n-m))} - (\chi^2_{2m;1-\gamma} - 2m)}{4n}\right]^2 \qquad \text{(A-13)}$$

A.2. Generalizations

As one might expect, the sum of any m of the n+1 blocks determined by the ordered observations gives the result (A-11). This is easily seen by examining the distribution of the coverages. From equation (A-4), the probability element of $Y_{(i)} = F(X_{(i)})$, $i = 1, \ldots, n$ is

$$\begin{cases} n!\, dy_{(1)}\, dy_{(2)} \cdots dy_{(n)} & 0 < y_{(1)} < \ldots < y_{(n)} < 1 \\ 0 & \text{otherwise .} \end{cases}$$

Let the coverages be denoted by U_i, $i = 1, \ldots, n+1$. Let u_i, $i = 1, \ldots, n+1$ be the variates corresponding to the random variables U_i, $i = 1, \ldots, n+1$. Let $u_1 = y_{(1)}$, $u_2 = y_{(2)} - y_{(1)}, \ldots, u_i = y_{(i)} - y_{(i-1)}, \ldots, u_n = y_{(n)} - y_{(n-1)}$ and $u_{n+1} = 1 - u_1 - \ldots - u_n$. Since the magnitude of the Jacobian of the transformation from $y_{(1)}, \ldots, y_{(n)}$ to $u_1, \ldots, u_n$ is one, the probability element of the coverages is

$$\begin{cases} n!\, du_1 du_2 \cdots du_n & 0 < u_i, i = 1, \ldots, n; \ \sum_{i=1}^{n+1} u_i = 1 \\ 0 & \text{otherwise.} \end{cases} \qquad \text{(A-14)}$$

Equation (A-14) is completely symmetrical with respect to the u_i. This means that coverage of one block has the same properties as the coverage of any other block. Because of the symmetry, the distribution of the sum of any t coverages is the same as the distribution of the sum of the first t coverages. The sum of the first t coverages is $\sum_{i=1}^{t} U_i = F(X_{(t)}) = Y_{(t)}$. Therefore, the probability element of the sum of any t coverages is given by

$$n!\, dy_{(t)} \int_0^{y_{(t)}} \cdots \int_0^{y_{(2)}} dy_{(1)} \cdots dy_{(t-1)} \int_{y_{(t)}}^{1} \cdots \int_{y_{(n-1)}}^{1}$$

$$dy_{(n)} \cdots dy_{(t+1)} = \frac{n!}{(t-1)!\,(n-r)!}\, y_{(t)}^{t-1} (1-y_{(t)})^{n-t}\, dy_{(t)} \qquad \text{(A-15)}$$

Let $m = n+1-t$, be the number of blocks which are eliminated from the interval of interest, and denote $y_{(n+1-t)}$ by w. The probability element of the coverage of the intervals between any $n+1-m$ order statistics is given by

$$\frac{n!}{(n-m)!\,(m-1)!}\, w^{n-m} (1-w)^{m-1}\, dw. \qquad 0 \le w \le 1 \qquad \text{(A-16)}$$

Integrating equation (A-16) over $[\beta, 1]$ one obtains equation (A-11).

The idea of coverage is more general than it first appears. Let the observations again be labeled $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ where

$X_{(1)} < X_{(2)} < \ldots < X_{(n)}$. The blocks need not necessarily be defined as $(-\infty, X_{(1)}]$, $(X_{(1)}, X_{(2)}], \ldots, (X_{(n)}, \infty)$. For example consider the blocks $[\zeta, X_{(j)}]$ $(X_{(j)}, X_{(j+1)}], \ldots, (X_{(n-1)}, X_{(n)}]$, $\{(X_{(n)}, \infty) \cup (-\infty, X_{(1)}]\}$ $(X_{(1)}, X_{(2)}], \ldots, (X_{(j-2)}, X_{(j-1)}]$, $(X_{(j-1)}, \zeta)$ where ζ is some number on the real line and j is a positive integer with $j \leq n$. These blocks are formed by ordering the observations in the following manner. See Figure A-1. The blocks are numbered (1) , (2) , ... in the order that they are formed. The first block is formed by searching for the observation whose value is closest to ζ while being greater than ζ . The second block is formed by searching for the second closest observation to ζ which is also greater than ζ . The procedure is continued until the largest observation $X_{(n)}$ is found. A search is made for an observation greater than $X_{(n)}$ and none is found. To complete the block, a search is made from minus infinity for the smallest observation. When this observation is found, the block $(X_{(n)}, \infty) \cup (-\infty, X_{(1)})$ is completed. The last $j-2$ blocks are formed by searching for subsequently larger observations. The $(n+1)^{th}$ block is $(X_{(j-1)}, \zeta)$. Let the coverages be defined as follows

$$U_1 = F(X_{(j)}) - F(\zeta), \; U_2 = F(X_{(j+1)}) - F(X_{(j)}), \ldots, U_{n-j+1}$$
$$= 1 - F(X_{(n)}) + F(X_{(1)}), \ldots, U_n = F(X_{(j-1)}) - F(X_{(j-2)}) .$$

The Jacobian of the transformation from the random variables $F(X_{(1)}), \ldots, F(X_{(n)})$ to the coverages $U_1, \ldots, U_n$ is one. Hence, the probability element of the coverages is

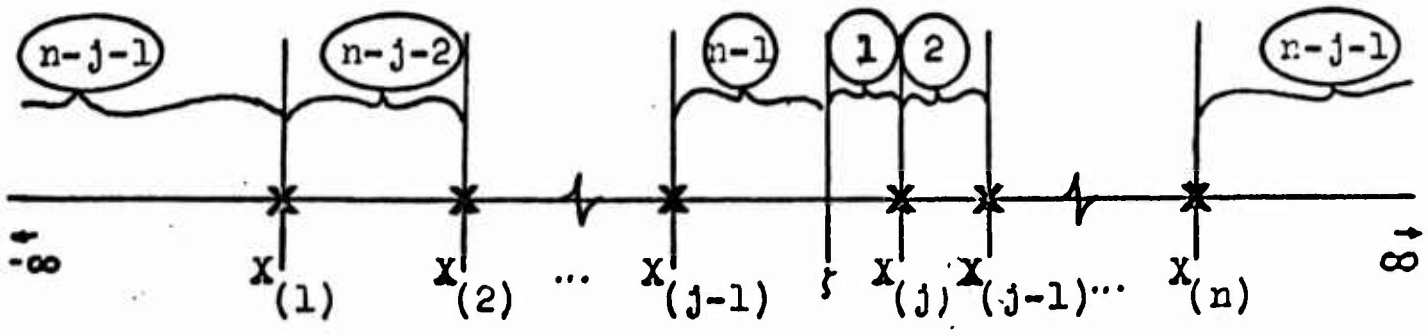


Figure A-1. Ordering From a Fixed Point.

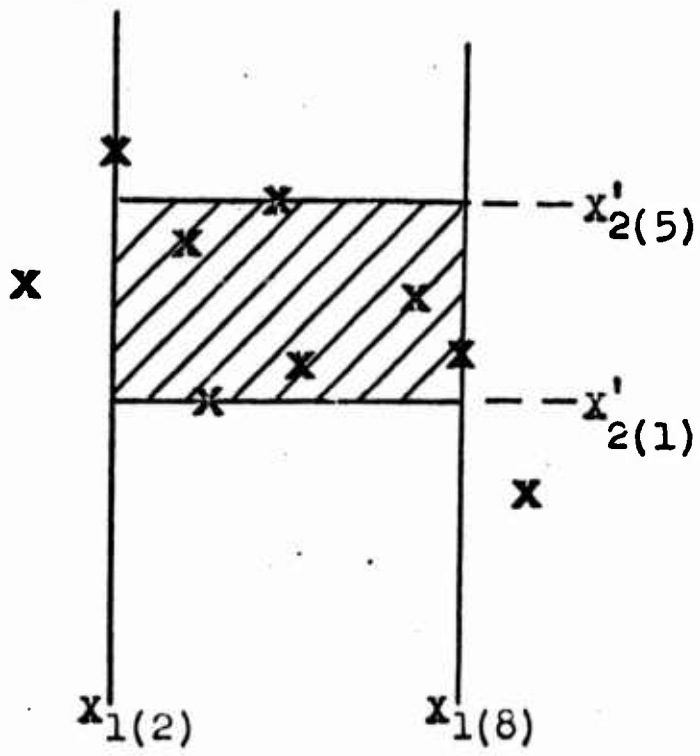


Figure A-2a. Wald's Ordering Procedure.

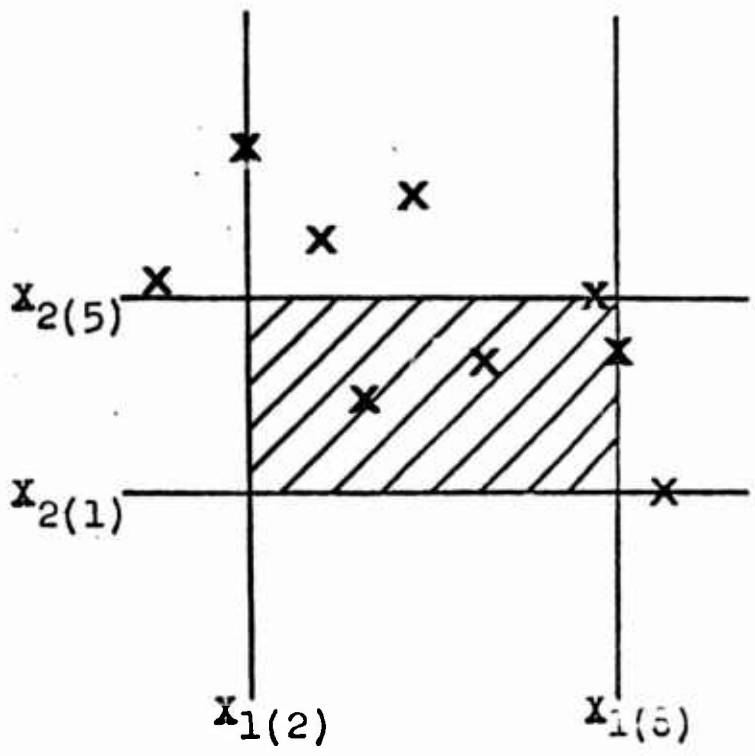


Figure A-2b. Not Distribution-Free. (Assuming Dependent Variables)

$$\begin{cases} n!\, du_1 du_2 \cdots du_n & 0 \le u_i,\ i = 1, \ldots, n; \ \sum_{i=1}^{n+1} u_i = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{A-17}$$

Therefore, this ordering procedure produces distribution-free tolerance regions.

It should be noted that only order statistics yield distribution-free tolerance regions. See Robbin (1944).

Some warning seems appropriate on the selection of intervals in which β of the distribution is to be contained with probability γ. The ordering functions and the blocks should, in general, be prescribed before the sample is taken. For example, suppose that one observation, X_1, is taken from the population described by $F(x)$ on the real line. If the location of X_1 is not known, one can search for X_1 from any point ζ as before and have no reason to believe that X_1 will be found before one-half the probability measure has been covered. However, if the location of X_1 is known and if an ordering procedure is devised so a "search" is made toward X_1 from a point ζ which is a very small distance to the left of X_1, the amount of probability in (ζ, X_1) can be made much less than the amount of probability in the other "block" $(-\infty, \zeta) \cup (X_1, \infty)$. Hence statistically equivalent blocks are not formed in this example. The blocks can be chosen, however, on the basis of a previous sample when the results are to be applied to a future sample. It is seen later that subsequent ordering functions may depend on the location of the observations that have previously been ordered. But the

ordering functions should not, in general, depend on the location of any unordered observations.

A. 3. D-Dimensional Theory

Thus far, only one-dimensional, continuous distribution functions have been discussed. Multi-dimensional distribution-free tolerance regions are defined as follows:

Definition: Suppose a sample of size n is drawn from a continuous, D-dimensional distribution function $F(x_1, x_2, \ldots, x_D)$. Region R is a D-dimensional distribution-free tolerance region if the amount of probability from the distribution $F(x_1, x_2, \ldots, x_D)$ in region R does not depend on $F(x_1, \ldots, x_D)$.

Suppose X_1 and X_2 are two random variables which are described by the bivariate distribution function $F(x_1, x_2)$. Our task is to construct distribution-free tolerance regions in space $\{x_1, x_2\}$. Suppose a sample of n observations is available from $F(x_1, x_2)$. It is evident from the one-dimensional study that n lines, each of which intersects a different observation and lies parallel to the x_1 axis, divide the space into statistically equivalent blocks. The same is true, of course, for n lines passing through each observation and being parallel to the x_2 axis.

A more general ordering for producing distribution-free tolerance regions also becomes evident. Suppose the observations are ordered with the function $h(x_1, x_2)$. Let $V = h(X_1, X_2)$. As long as the distribution

function of V, $F_V(v)$, is continuous, it is seen from the one-dimensional derivation that the amount of probability between the order statistics $V_{(1)}, V_{(2)}, \ldots, V_{(n)}$ is independent of $F_V(v)$. Hence, the coverages defined by the n identical curves, each one passing through one of the n observations, are distribution-free. These methods for ordering the observations are easily extended for distribution functions of more than two variates.

A.4. Wald's Ordering

A more general ordering was proposed by Wald (1943). Let $(X_1, X_2, \ldots, X_D)$ be a set of D random variables with continuous probability density $F(x_1, x_2, \ldots, x_D)$. Take a sample of n independent observations and denote the α^{th} observation of X_i by $X_{i\alpha}(i = 1, \ldots, D; \alpha = 1, \ldots, n)$. The problem is to construct D pairs of functions, $L_i(X_{11}, \ldots, X_{Dn})$ and $M_i(X_{11}, \ldots, X_{Dn})$, $i = 1, \ldots, D$, so that the distribution of the statistic

$$Q = \int_{L_D}^{M_D} \cdots \int_{L_1}^{M_1} f(x_1, \ldots, x_D)\, dx_1 \cdots dx_D \qquad (A\text{-}18)$$

is independent of $f(x_1, \ldots, x_D)$. The following construction procedure satisfies this requirement. Let the $X_{1\alpha}$ be arranged in order of increasing magnitude, $X_{1(1)} < X_{1(2)} < \ldots < X_{1(n)}$. Choose $L_1 = X_{1(r_1)}$ and $M_1 = X_{1(s_1)}$ where r_1 and s_1 are positive integers with $r_1 < s_1 \leq n$. Next consider only the observations for which $X_{1(r_1)} < X_{1j} < X_{1(s_1)}$. Arrange these observations in order of increasing magnitude of the second coordinate, $X'_{2(1)} < X'_{2(2)} < \ldots < X'_{2(s_1 - r_1 - 1)}$. The prime distinguishes

the observations for which $X_{1(r_1)} < X_{1j} < X_{1(s_1)}$ from the original n observations. Let $L_2 = X'_{2(r_2)}$ and $M_2 = X'_{2(s_2)}$ where r_2 and s_2 are positive integers with $r_2 < s_2 \leq s_1 - r_1 - 1$. The construction procedure is continued in this manner until all D variates are exhausted. A two-variate example is shown in Figure A-2a with $r_1 = 2$, $s_1 = 8$, $r_2 = 1$, $s_2 = 5$.

The key to Wald's successful ordering, as will be seen later, is the successive elimination of blocks which have been formed. Suppose rather than use Wald's ordering, we let $L_i = X_{i(r_i)}$ and $M_i = X_{i(s_i)}$, $i = 1, \ldots, D$ where r_i and s_i denote positive integers with $r_i < s_i \leq n$. Therefore for $i = 2, \ldots, D$ we do not eliminate from subsequent ordering those observations which have previously been ordered. This ordering does not yield a region whose coverage is independent of the distribution $F(x_1, \ldots, x_D)$ for all $r_i < s_i \leq n$ when $F(x_1, x_2, \ldots, x_D) \neq F_1(x_1)F_2(x_2) \ldots F_D(x_D)$. For example, the crosshatched region of Figure A-2b is not a distribution-free tolerance region when the random variables X_1 and X_2 are statistically dependent.

However, if the random variables $X_1, X_2, \ldots, X_D$ are statistically independent, each variable can be ordered separately. Then the coverage of the region defined by $X_{i(r_i)}$ and $X_{i(s_i)}$, $i = 1, \ldots, D$ is given by the product of the coverages of each variable. For example, consider a bivariate distribution $F(x_1, x_2)$ which satisfies the relationship $F(x_1, x_2) = F_1(x_1)F_2(x_2)$. If $U = F_1(X_{1(s_1)} - F_1(X_{1(r_1)})$ and $V = F_2(X_{2(s_2)}) - F_2(X_{2(r_2)})$ the probability distribution of $W = U \cdot V$ can be calculated by integrating

$$\int_0^1 \int_0^{w/v} \frac{n!}{(s_1-r_1-1)!\,(n-s_1+r_1)!\,(s_2-r_2-1)!\,(n-s_2+r_2)!}\, u^{s_1-r_1-1}$$

$$v^{s_2-r_2-1}\,(1-u)^{n-s_1+r_1}\,(1-v)^{n-s_2+r_2}\,du\,dv\,. \qquad \text{(A-19)}$$

Now consider Wald's proof that his ordering scheme produces distribution-free tolerance regions. Consider the bivariate case where the random variables are statistically dependent and the joint probability density function is continuous. The object is to show that the distribution of the statistic

$$Q = \int_{L_2}^{M_2} \int_{L_1}^{M_1} f(x_1, x_2)\,dx_1\,dx_2\,, \qquad \text{(A-20)}$$

where L_1, M_1, L_2, and M_2 are given by Wald's ordering, is independent of the probability density function, $f(x_1, x_2)$.

Make the following definitions:

$$P \equiv \int_{-\infty}^{\infty} \int_{L_1}^{M_1} f(x_1, x_2)dx_1\,dx_2 \equiv F_1(M_1) - F_1(L_1) \qquad \text{(A-21)}$$

$$\overline{P} \equiv \frac{\int_{L_2}^{M_2} \int_{L_1}^{M_1} f(x_1, x_2)\,dx_1\,dx_2}{\int_{-\infty}^{\infty} \int_{L_1}^{M_1} f(x_1, x_2)\,dx_1\,dx_2} \qquad \text{(A-22)}$$

P is the probability that X_1 lies between L_1 and M_1. $\overline{P}$ is the probability that X_2 lies between L_2 and M_2 given that X_1 lies between L_1 and M_1. It is evident that

$$Q = P\overline{P} \tag{A-23}$$

If $L_1 = X_{1(r_1)}$ and $M_1 = X_{1(s_1)}$ where r_1 and s_1 are positive integers with $r_1 < s_1 \leq n$, it is clearly seen from the one-dimensional discussion that the probability element of P is given by

$$\frac{n!}{(s_1-r_1-1)!\,(n-s_1+r_1)!}\, P^{s_1-r_1-1}\,(1-P)^{n-s_1+r_1}\,dP \tag{A-24}$$

Let $L_2 = X'_{2(r_2)}$ and $M_2 = X'_{2(s_2)}$ as defined in Wald's construction procedure. Since $X'_{2(1)}, \ldots, X'_{2(s_1-r_1-1)}$ can be considered as s_1-r_1-1 independent observations on random variable X_2 under the condition that $L_1 < X_1 < M_1$, the probability element of $\overline{P}$ is given by

$$\frac{(s_1-r_1-1)!}{(s_2-r_2-1)!\,(s_1-r_1-s_2+r_2-1)!}\,(\overline{P})^{s_2-r_2-1}\,(1-\overline{P})^{s_1-r_1-1-s_2+r_2}\,d\overline{P} \tag{A-25}$$

Note that equation (A-25) does not involve L_1 and M_1. Hence, the joint probability element of P and $\overline{P}$ is given by the product of equations (A-24) and (A-25).

$$\frac{n!}{(n-s_1+r_1)!\,(s_2-r_2-1)!\,(s_1-r_1-s_2+r_2-1)!}\, P^{s_1-r_1-1}\,(1-P)^{n-s_1+r_1}$$

$$(\overline{P})^{s_2-r_2-1}\,(1-\overline{P})^{s_1-r_1-1-s_2+r_2}\,dP\,d\overline{P} \tag{A-26}$$

Then the joint probability of P and $Q = P\overline{P}$ is given by

$$K(1-P)^{n-s_1+r_1} Q^{s_2-r_2-1} (P-Q)^{s_1-r_1-1-s_2+r_2} dP\,dQ$$

where K is the multiplicative constant of equation A-26. By integrating P over the interval $[Q,1]$ we obtain

$$KdQ\ Q^{s_2-r_2-1} \int_Q^1 (1-P)^{n-s_1+r_1} (P-Q)^{s_1-r_1-1-s_2+r_2} dP\ .$$

Let $R = P-Q$. The value of the integral is then

$$\int_Q^1 (1-P)^{n-s_1+r_1} (P-Q)^{s_1-r_1-1-s_2+r_2} dP = \int_Q^{1-Q} (1-Q-R)^{n-s_1+r_1}$$

$$R^{s_1-r_1-1-s_2+r_2} dR\ .$$

Let $R = (1-Q)\,T$. Then the above equation reduces to

$$(1-Q)^{n-1-s_2+r_2} (1-Q) \int_0^1 (1-T)^{n-s_1+r_1} T^{s_1-r_1-1-s_2+r_2} dT\ .$$

Integrating with respect to T one finds that the probability element of Q is

$$\frac{n!}{(s_2-r_2-1)!\,(n-s_2+r_2)!}\ Q^{s_2-r_2-1} (1-Q)^{n-s_2+r_2} dQ\ . \qquad \text{(A-27)}$$

Since equation A-27 is independent of $f(x_1, x_2)$, Wald's method of ordering produces distribution-free tolerance regions.

It is convenient to think of Wald's ordering as eliminating blocks from the region of interest. The 1st ordering ($X_{1(r_1)}$ and $X_{1(s_1)}$)

eliminates the r_1 blocks whose X_1 coordinate is less than $X_{1(r_1)}$ and the $n+1-s_1$ blocks whose X_1 coordinate is greater than $X_{1(s_1)}$. The 2^{nd} ordering eliminates r_2 blocks and $s_1-r_1-s_2$ blocks for a total elimination of $n+1-s_2+r_2$ blocks. By substituting $m = n+1 - s_2+r_2$ into equation A-27 one obtains equation A-16, the equation for the probability element for the one-dimensional case.

A.5. A General Ordering Procedure

Consider a more general method for ordering the observations. Assume that a sample $(X_{1\alpha}, X_{2\alpha}; \alpha = 1, \ldots, n)$ is available from a continuous two-dimensional distribution $F(x_1, x_2)$. Let $w_k = h_k(x_1, x_2)$; $k = 1, \ldots, n$ be n functions, possibly alike, possibly distinct, such that $h_1(X_1, X_2), \ldots, h_n(X_1, X_2)$ are random variables with a continuous joint distribution function. Since $F(x_1, x_2)$ is assumed to be continuous, the functions $w_k, k = 1, \ldots, n$ are continuous almost everywhere. These functions are called ordering functions. They are used to form distribution-free tolerance regions in the following manner. The first ordering function $h_1(x_1, x_2)$ is used to select the observation $(X_{1(1)}, X_{2(1)})$ which satisfies

$$\max_{\alpha=1, \ldots, n} h_1(X_{1\alpha}, X_{2\alpha}) = h_1(X_{1(1)}, X_{2(1)}) . \qquad \text{(A-28)}$$

Let the region of the sample space for which $h_1(x_1, x_2) > h_1(X_{1(1)}, X_{2(1)}$ be called block B_1, and the region of the sample space for which $h_1(x_1, x_2) = h_1(X_{1(1)}, X_{2(1)})$ be called T_1. Then

$$B_1 = \left\{(x_1, x_2) : h_1(x_1, x_2) > h_1(X_{1(1)}, X_{2(1)}) \right\}$$

and

$$T_1 = \left\{(x_1, x_2) : h_1(x_1, x_2) = h_1(X_{1(1)}, X_{2(1)}) \right\} . \qquad \text{(A-29)}$$

Eliminate the block B_1 and the cut T_1 from the sample space. Note that the observation $(X_{1(1)}, X_{2(1)})$ is also eliminated. Select from the remaining n-1 observations the one which satisfies

$$\max_{\substack{\alpha=1,\ldots,n \\ \alpha \neq (1)}} h_2(X_{1\alpha}, X_{2\alpha}) = h_2(X_{1(2)}, X_{2(2)}) . \qquad \text{(A-30)}$$

Let the region of the sample space for which $h_2(x_1, x_2) > h_2(X_{1(2)}, X_{2(2)})$ and $h_1(x_1, x_2) < h_1(X_{1(1)}, X_{2(1)})$ be called block B_2 and the region for which $h_2(x_1, x_2) = h_2(X_{1(2)}, X_{2(2)})$ and $h_1(x_1, x_2) < h_1(X_{1(1)}, X_{2(1)}$ be called cut T_2. Then

$$B_1 = \left\{(x_1, x_2) : h_2(x_1, x_2) > h_2(X_{1(2)}, X_{2(2)}),\ h_1(x_1, x_2) < h_1(X_{1(1)}, X_{2(1)}) \right\}$$

and

$$T_1 = \left\{(x_1, x_2) : h_2(x_1, x_2) = h_2(X_{1(2)}, X_{2(2)}),\ h_1(x_1, x_2) < h_1(X_{1(1)}, X_{2(1)}) \right\} \qquad \text{(A-31)}$$

This procedure is continued until n blocks are eliminated from the sample space. The sample space is thereby partitioned into n+1 mutually exclusive and exhaustive blocks by the n cuts.

It is now shown that this procedure produces distribution-free

tolerance regions. Denote the portion of the distribution contained in block B_i (i.e. the coverage of B_i) by U_i^*, $i = 1, \ldots, n$. To prove that the coverages U_i^*, $i = 1, \ldots, n$ are distribution-free it is sufficient to prove that the probability density of the coverages is uniform. The proof given here follows the one by Wilks (1962). Consider two random variables X_1 and X_2 with distribution $F(x_1, x_2)$. The probability element for sample $(X_{1\alpha}, X_{2\alpha}\ ; \alpha = 1, \ldots, n)$ is

$$\prod_{\alpha=1}^{n} dF(x_{1\alpha}, x_{2\alpha}) \qquad \text{A-32)}$$

Let $(X_{1(1)}, X_{2(1)})$ be the observation which yields the largest value for $h_1(x_1, x_2)$. Let $(X_{1\alpha}^{(1)}, X_{2\alpha}^{(1)}; \alpha = 1, \ldots, n)$ be the set of observations obtained by deleting $(X_{1(1)}, X_{2(1)})$ from the sample space. The probability element of $(X_{1(1)}, X_{2(1)})$ and $(X_{1\alpha}^{(1)}, X_{2\alpha}^{(1)}; \alpha = 1, \ldots, n\text{-}1)$ is

$$n dF(x_{1(1)}, x_{2(1)}) \prod_{\alpha=1}^{n-1} dF(x_{1\alpha}^{(1)}, x_{2\alpha}^{(1)}) \qquad \text{(A-33)}$$

Let U_1' be the coverage associated with the set of points (x_1, x_2) for which $h_1(x_1, x_2) > h_1(x_{1(1)}, x_{2(1)})$. The probability element of U_1' is

$$n(1-u_1')^{n-1}\, du_1' \qquad \text{(A-34)}$$

where

$$u' = \int_{B_1} dF(x_1, x_2) \quad \text{and} \quad du_1' = dF(x_{1(1)}, x_{2(1)}) .$$

The probability element of the conditional random variable

$(X_{1\alpha}^{(1)}, X_{2\alpha}^{(1)} / X_{1(1)}, X_{2(1)};\ \alpha = 1, \ldots, n)$ is given by the ratio of equation A-33 to equation A-34. This ratio reduces to

$$\prod_{\alpha=1}^{n-1} dF^{(1)}(x_{1\alpha}^{(1)}, x_{2\alpha}^{(1)})$$

where

$$F^{(1)}(x_{1\alpha}^{(1)}, x_{2\alpha}^{(1)}) = \frac{F(x_{1\alpha}^{(1)}, x_{2\alpha}^{(1)})}{1-u_1'} .$$

Therefore given $(X_{1(1)}, X_{2(1)}) = (x_{1(1)}, x_{2(1)})$ the remaining observations of the original sample behave like a sample of n-1 observations from the distribution

$$F^{(1)}(x_1, x_2) = \frac{F(x_1, x_2)}{1-u_1'} . \qquad \text{(A-36)}$$

Now let $(X_{1(2)}, X_{2(2)})$ be the observation among the remaining n-1 observations which yields the largest value for $h_2(x_1, x_2)$. Let $(X_{1\alpha}^{(2)}, X_{2\alpha}^{(2)};\ \alpha = 1, \ldots, n-2)$ be the set of n-2 observations obtained by deleting $(X_{1(2)}, X_{2(2)})$ from the n-1 observations. The probability element of $(X_{1(2)}, X_{2(2)})$ and $(X_{1\alpha}^{(2)} X_{2\alpha}^{(2)};\ \alpha = 1, \ldots, n-2)$ is

$$(n-1)\, dF^{(1)}(x_{1(2)}, x_{2(2)}) \prod_{\alpha=1}^{n-2} dF^{(1)}(x_{1\alpha}^{(2)}, x_{2\alpha}^{(2)}) . \qquad \text{(A-37)}$$

Let B_2 be the region for which $h_2(x_1, x_2) > h_2(X_{1(2)}, X_{2(2)})$ and $h_1(x_1, x_2) < h_1(X_{1(1)} X_{2(1)})$. Let

$$u_2' = \int_{B_2} dF^{(1)}(x_1, x_2) \qquad \text{(A-38)}$$

be the conditional coverage associated with B_2. Then the probability element of U_2' is given by

$$(n-1)\,(1-u_2')^{n-2}\,du_2' \,. \tag{A-39}$$

The probability element of $(X_{1\alpha}^{(2)}, X_{2\alpha}^{(2)};\ \alpha = 1, \ldots, n-2)$ is

$$\prod_{\alpha=1}^{n-2} dF^{(2)}\,(x_{1\alpha}^{(2)}, x_{2\alpha}^{(2)}) \tag{A-40}$$

where

$$F^{(2)}\,(x_{1\alpha}^{(2)}, x_{2\alpha}^{(2)}) = \frac{F^{(1)}\,(x_{1\alpha}^{(2)}, x_{2\alpha}^{(2)})}{1-u_2'}$$

Therefore given $(x_{1(2)}, x_{2(2)})$, the remaining n-2 observations behave like a sample from

$$F^{(2)}\,(x_1, x_2) = \frac{F^{(1)}\,(x_1, x_2)}{1-u_2'} \,. \tag{A-41}$$

Continuing in the above manner, one concludes that the conditional coverages $U_1', U_2', \ldots, U_n'$ have the probability element

$$n!\,(1-u_1')^{n-1}\,(1-u_2')^{n-2} \ldots (1-u_n')^{1}\ du_1' \ldots du_n' \,. \tag{A-42}$$

The conditional coverages $(U_i';\ i = 1, \ldots, n)$ are related to the coverages $(U_i^*;\ i = 1, \ldots, n)$ by the following expressions:

$$\begin{aligned} U_1' &= U_1^* \\ U_2' &= \frac{U_2^*}{1-U_1^*} \\ &\vdots \\ U_n' &= \frac{U_n^*}{1-U_1^*-\ldots-U_{n-1}^*} \end{aligned} \tag{A-43}$$

Rewriting equation A-42 in terms of the U_i^* one finds that the probability element of the coverages is

$$n!\, du_1^* \ldots du_n^* . \qquad \text{(A-44)}$$

Since this equation is the same as equation A-11, the proof is completed.

The procedure for forming distribution-free tolerance regions has been further generalized to permit the subsequent ordering functions to depend in any way on the information gained in the application of previous ordering functions. See Fraser (1953), Kemperman (1956).

The following procedure is one example of a particularly general ordering. See Fraser (1957). Let $h_1(\underline{x}), \ldots, h_n(\underline{x})$ be n real-valued measurable functions of $\underline{x} = (x_1, \ldots, x_D)$. Let $\max_{i=1,\ldots,n}{}^{(r_j)} h_j(\underline{X}_i)$ denote the r_j^{th} largest value of h_j. Note that r_j must be an integer which is less than n. Let the observations be ordered as follows. First, locate the observation which gives the r_1^{th} largest value for h_1. Denote this observation by $\underline{X}_{(1)}$. Then

$$h_1(\underline{X}_{(1)}) = \max_{i=1,\ldots,n}{}^{(r_1)} h_1(\underline{X}_i) \qquad 1 \leq r_1 \leq n_1 \qquad \text{(A-45)}$$

The sample space is partitioned into two subspaces,

$$S_{1 \ldots r_1} = \left\{ \underline{x} : h_1(\underline{x}) > h_1(\underline{X}_{(1)}) \right\} \qquad \text{(A-46a)}$$

and

$$S_{r_1+1 \ldots n+1} = \left\{ \underline{x} : h_1(\underline{x}) < h_1(\underline{X}_{(1)}) \right\} , \qquad \text{(A-46b)}$$

by means of the cut

$$T_{r_1} = \left\{ \underline{x} : h_1(\underline{x}) = h_1(\underline{X}_{(1)}) \right\} . \qquad \text{(A-46c)}$$

The observation $\underline{X}_{(1)}$ is eliminated from any further ordering. The two subspaces $S_{1\ldots r_1}$ and $S_{r_1+1\ldots n+1}$ are treated separately in further ordering. There are r_1-1 observations in subspace $S_{1\ldots r_1}$ and $n-r_1$ observations in subspace $S_{r_1+1\ldots n+1}$. Suppose next we locate the observation which gives the r_2^{th} largest value of $h_2(\underline{x})$ in $S_{1\ldots r_1}$ where $1 \leq r_2 \leq r_1-1$. Denote this observation by $\underline{X}_{(2)}$. Then

$$h_2(\underline{X}_{(2)}) = \max_{\underline{X}_i \in S_{1\ldots r_1}} (r_2)\, h_2(\underline{X}_i) \qquad 1 \leq r_2 \leq r_1 - 1 \ . \qquad \text{(A-47)}$$

(Note that $h_2(\underline{x})$ can depend on $\underline{X}_{(1)}$.) The subspace $S_{1\ldots r_1}$ is partitioned into subspaces

$$S_{1\ldots r_2} = \left\{\underline{x} : h_2(\underline{x}) < h_2(\underline{X}_{(2)}),\ h_1(\underline{x}) > h_1(\underline{X}_{(1)}) \right\} \qquad \text{(A-48a)}$$

and

$$S_{r_2+1\ldots r_1} = \left\{\underline{x} : h_2(\underline{x}) > h_2(\underline{X}_{(2)}), h_1(\underline{x}) > h_1(\underline{X}_{(1)}) \right\} \qquad \text{(A-48b)}$$

by means of cut

$$T_{r_2} = \left\{\underline{x} : h_2(\underline{x}) = h_2(\underline{X}_{(2)})\ ,\ h_1(\underline{x}) > h_1(\underline{X}_{(1)}) \right\} . \qquad \text{(A-48c)}$$

The observation $\underline{X}_{(2)}$ is eliminated from further ordering. Next we order in one of the three subspaces $S_{1\ldots r_2}$, $S_{r_2+1\ldots r_1}$, or $S_{r_1+1\ldots n+1}$. This procedure is continued until $n+1$ blocks are formed.

The two-dimensional example shown in Figure A-3 illustrates this ordering procedure. The coordinates of the space are labeled x_1 and x_2 and arrows show the direction of increasing magnitude. The three observations are labeled with $\underline{X}$'s. The parameters for the first

ordering are $h_1(\underline{x}) = x_1$ and $p_1 = 2$. This divides the space into S_{12} and S_{34}. The parameters $h_2(\underline{x}) = |\underline{x} - \underline{X}_{(1)}|$ and $p_1 = 1$ are used to divide S_{12} into S_1 and S_2. The parameters $h_3(\underline{x}) = x_2$ and $p_3 = 1$ are used to divide S_{34} into S_3 and S_4. A total of four blocks S_1, S_2, S_3, and S_4 result.

A.6. Discontinuities

Distribution-free tolerance regions can be constructed when the original distribution has a countable number of discontinuities. In dealing with discontinuous distributions, problems arise from the finite probabilities associated with the cuts. In addition, the construction procedure must incorporate a method for handling ties.

It has been shown in Scheffé and Tukey (1945), Tukey (1948), and Fraser and Wormleighton (1951) that if all cuts adjacent to the blocks of interest are included in R, the region of interest, the following statement can be made

$$\Pr\left\{ \int_R dF(\underline{x}) \geq \beta \right\} \geq \gamma \qquad \text{(A-49)}$$

where β and γ are determined from the number of blocks contained in R. For example if S_4 in Figure A-3 is the region of interest, the cut $T_1 = \left\{\underline{x} : x_1 = \underline{X}_{1(1)}\right\}$ and the cut $T_3 = \left\{\underline{x} : x_2 = \underline{X}_{2(3)},\ x_1 \leq \underline{X}_{1(1)}\right\}$ must be included in R so that statement (A-49) can be made.

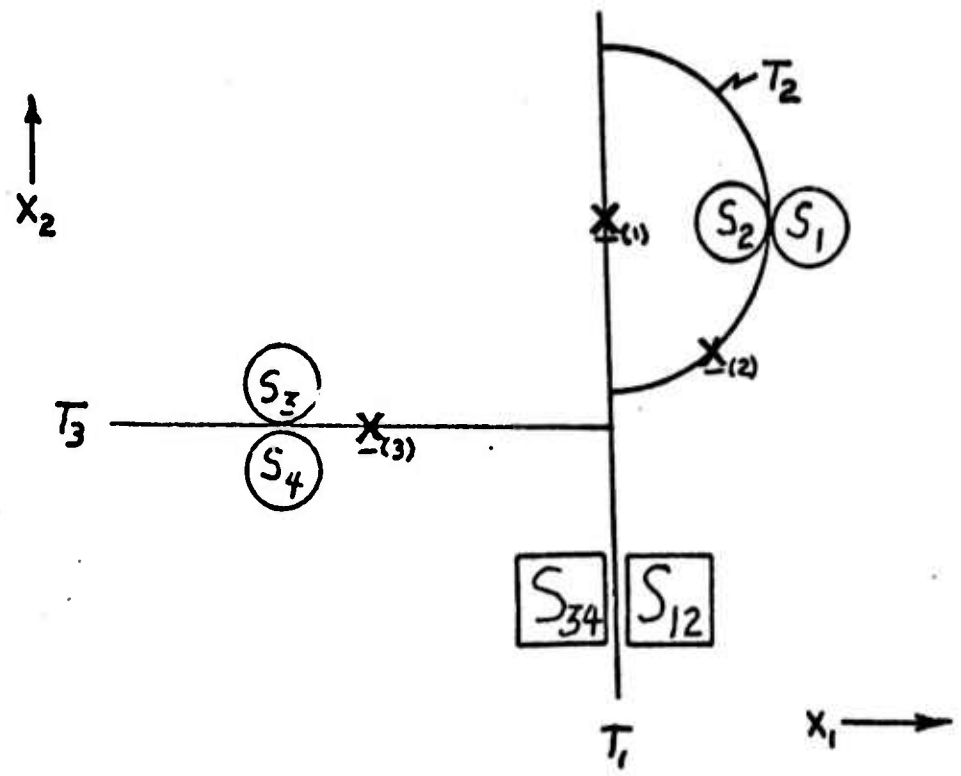


Figure A-3. Illustration of a General Ordering Procedure

A.7. Other Extensions

A tolerance region theory which uses statistically equivalent blocks has been developed for cases in which the class of probability distribution functions is limited. As might be expected, when the kind of distribution function under consideration is restricted, more blocks can be eliminated from the region of interest for the same γ and β than can be eliminated when nothing is known about the distribution. A frequently considered class of distributions is the class for which $\frac{f(x)}{1-F(x)}$, the hazard rate, is a monotone function. See Hanson and Koopmans (1964) and Barlow and Proschan (1966) for further information.

Appendix B

PROPERTIES OF THE BETA DISTRIBUTION

The Beta distribution is given by the expression

$$f(x) = \frac{\Gamma(n+1)}{\Gamma(n-m+1)\,\Gamma(m)}\, x^{n-m}(1-x)^{m-1} \qquad 0 \le x \le 1 . \qquad \text{(B-1)}$$

The distribution is symmetric about $\frac{1}{2}$ when $m = \frac{n+1}{2}$. It is skewed toward the value 1 when n-m is greater than m-1 and toward the value 0 when n-m is less than m-1. The mean of the Beta distribution is

$$\mu = \frac{n-m+1}{n+1} \qquad \text{(B-2)}$$

and the variance is

$$\sigma^2 = \frac{m(n-m+1)}{(n+1)^2\,(n+2)} = \frac{\mu(1-\mu)}{(n+2)} . \qquad \text{(B-3)}$$

At times one may use the median or the mode rather than the mean to predict the coverage. The mode is easily found by differentiating equation B-1 and setting the differential equal to zero. The mode is found to be

$$\text{MODE} = \frac{n-m}{n-1} . \qquad \text{(B-4)}$$

The median is given by

$$\int_0^{\beta} \frac{\Gamma(n+1)}{\Gamma(n-m+1)\,\Gamma(m)}\, \lambda^{n-m}(1-\lambda)^{m-1}\, dx =$$

$$I_{\beta}(n-m+1, m) = \frac{1}{2} . \qquad \text{(B-5)}$$

One can find values of β in tables of 50 percentage points of the Beta function or from tables of the cummulative Binomial distribution by using the relation

$$\frac{1}{2} = \sum_{s=m}^{n} \binom{n}{s} (1-\beta)^s \beta^{n-s} .$$

For values of n-m greater than m-1 , the mean is less than the median which is less than the mode. For example, consider the following table.

Table B-1. Comparison of the Mean, Median and Mode for Two Different Values of m and n.

Conditions	Mean	Median	Mode
n=65, m=6	0.90909	0.91321	0.92188
n=60, m=1	0.98361	0.98851	1.00000

Appendix C

Let R_2 be a distribution-free tolerance region on the continuous cumulative probability distribution function $F_1(x)$. The purpose of this appendix is to demonstrate that the expected value of $\int_{R_2} dF_1(x)$,

$$E\left\{\int_{R_2} dF_1(x)\right\} = \alpha \qquad \text{(C-1)}$$

can be considered an α-confidence statement that a new observation from $F_1(x)$ will fall in R_2. In other words, equation C-1 is equivalent to saying that the probability is α that a new observation from $F_1(x)$ will fall in R_2.

Let $X_{(r)}$ and $X_{(s)}$ be the r^{th} and s^{th} order statistics, respectively, from the distribution $F_1(x)$, $r < s \leq n_1$. Let Y be a new observation from this distribution. The probability that Y falls in $(X_{(r)}, X_{(s)}]$ is

$$\Pr(X_{(r)} < Y < X_{(s)}) = \int_{-\infty}^{\infty} dx_{(s)} \int_{-\infty}^{x_{(s)}} dx_{(r)} \int_{x_{(r)}}^{x_{(s)}} dy\, f_1(y)\, f(x_{(r)}, x_{(s)}) \qquad \text{(C-2)}$$

The integral with respect to y is

$$\int_{x_{(r)}}^{x_{(s)}} f_1(y)\, dy = F_1(x_{(s)}) - F_1(x_{(r)}) . \qquad \text{(C-3)}$$

The probability element, $f(x_{(r)}, x_{(s)})\, dx_{(r)}\, dx_{(s)}$, from equation A-7, Appendix A is

$$\frac{\Gamma(r+1)\, dF_1(x_{(r)})\, dF_1(x_{(s)})}{\Gamma(r)\,\Gamma(n+1-s)\,\Gamma(s-r)} \left[F_1(x_{(r)})\right]^{r-1} *$$

$$\left[F_1(x_{(s)}) - F_1(x_{(r)})\right]^{s-r-1} \left[1 - F_1(x_{(s)})\right]^{n_1-s} . \qquad \text{(C-4)}$$

Making the transformation $W = F_1(X_{(s)}) - F_1(X_{(r)})$ and $Z = F_1(X_{(r)})$ and proceeding as in equations A-8 and A-9 of Appendix A, one obtains

$$Pr(X_{(r)} < Y \leq X_{(s)}) = \int_0^1 w * \frac{\Gamma(n_1+1)}{\Gamma(s-r)\,\Gamma(n_1-s+r+1)} w^{s-r-1} (1-w)^{n_1-s+r} \, dw. \qquad \text{(C-5)}$$

Note that

$$\frac{\Gamma(n_1+1)}{\Gamma(s-r)\,\Gamma(n_1-s+r+1)} w^{s-r-1} (1-w)^{n_1-s+r} \, dw$$

is the probability density function for the coverage of $(X_{(r)}, X_{(s)}]$ (equation A-9). Then the right side of equation C-5 is the expected coverage of $(X_{(r)}, X_{(s)}]$. And, in fact, by recognizing that

$$\int_0^1 w^{s-r} (1-w)^{n_1-s+r} \, dw = \frac{\Gamma(s-r+1)\,\Gamma(n_1-s+r+1)}{\Gamma(n_1+2)}$$

we obtain

$$Pr(X_{(r)} < Y \leq X_{(s)}) = \frac{s-r}{n_1+1} ,$$

the expected coverage of $(X_{(r)}, X_{(s)}]$. Since the blocks are statistically equivalent, the result holds for any $s-r$ blocks. Hence, it can be generalized that

$$E\left\{\int_{R_2} dF_1(x)\right\} = \alpha$$

where R_2 is a distribution-free tolerance region, can be considered an α-confidence statement that a new observation from $F_1(x)$ will fall in R_2.

Appendix D

REVIEW OF CLASSIFICATION METHODS IN PATTERN RECOGNITION GIVEN TRAINING SAMPLES OF KNOWN CLASSIFICATION

The following is a review of classification methods along the guidelines discussed in Chapter 1.

I. Optimum Solution with Assumed Probability Densities

Using this approach the machine designer assumes the form of the a priori and conditional probabilities. Hence "optimum" recognition is achieved if the assumptions are correct. A model for the optimum recognition system is shown in Figure D-1. In this figure $\underline{v}$ is a vector in measurement space, ξ_i is the a priori probability of a class i event, $f_i(\underline{v})$ is the probability density function of $\underline{v}$ given that $\underline{v}$ is a member of class i and $C_i(j)$ is the cost of deciding that $\underline{v}$ is a member of class j when it is a member of class i.

The optimum recognition system is considerably simplified if certain assumptions are made about the $C_i(j)$. Suppose the cost of making an incorrect decision is equal to one and the cost of making a correct decision is equal to zero. Then the optimum recognition system consists of the probability density computer $\xi_i f_i(\underline{v})$, $i = 1, \ldots, K$ and a maximum selector. The inputs to the extremum selector (whether maximum or minimum) are called discriminant functions.

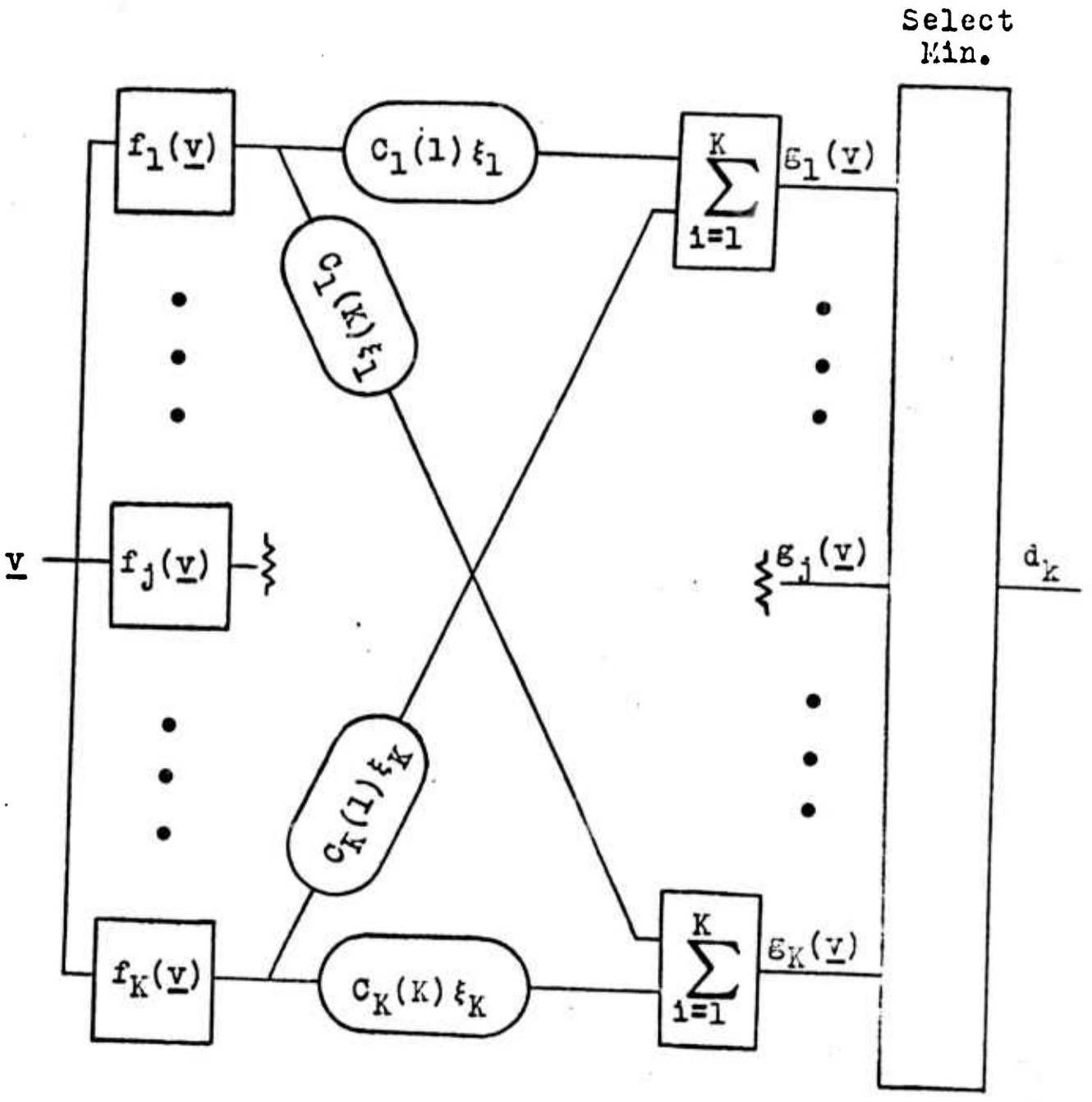


Figure D-1. Optimum Recognition System.

Let the discriminant functions be denoted by $g_i(\underline{v})$, $i = 1, \ldots, K$. Then for the above cost assumptions an optimum recognition system classifies $\underline{v}$ into class k if $g_k(\underline{v}) \geq g_i(\underline{v})$ for all i.

Consider the problem of finding the discriminant functions, which are optimum in the Bayes sense, when there are K Gaussian classes with covariance matrices $\underline{\Sigma}_i$ and mean vectors $\underline{\mu}_i$, $i=1, \ldots, K$. Assume also that the cost of an incorrect decision is one and that the cost of a correct decision is zero. The discriminant functions are

$$g_i(\underline{v}) = \frac{\xi_i}{(2\pi)^{K/2} |\underline{\Sigma}_i|^{1/2}} \exp \left[-\frac{1}{2} (\underline{v} - \underline{\mu}_i)^T \underline{\Sigma}_i^{-1} (\underline{v} - \underline{\mu}_i) \right] \qquad i = 1, \ldots, K \qquad \text{(D-1)}$$

where $(\cdot)^T$ denotes the transpose of $(\cdot)$. Let $h_i(\underline{v}) = \log g_i(\underline{v})$. Since the log function is a monotonically increasing function of its argument, the log of $g_i(\underline{v})$ can be used instead of $g_i(\underline{v})$ without any change in the decision. In the following log $g_i(\underline{v})$ and $g_i(\underline{v})$ will both be referred to as discriminant functions. The log of $g_i(\underline{v})$ for the Gaussian case is

$$h_i(\underline{v}) = \log \xi_i - \frac{K}{2} \log (2\pi) - \frac{1}{2} \log |\underline{\Sigma}_i| - \frac{1}{2} \left[(\underline{v} - \underline{\mu}_i)^T \underline{\Sigma}_i^{-1} (\underline{v} - \underline{\mu}_i) \right] \qquad i=1, \ldots, K . \qquad \text{(D-2)}$$

Therefore, the optimum discriminant functions for Gaussian patterns are quadratic functions.

Consider the case where the covariance matrices are equal. Keeping only the terms in equation D-2 which depend on i, the discrim-

inant functions become

$$h_i(\underline{v}) = \log \xi_i + \underline{v}^T \underline{\Sigma}^{-1} \underline{\mu}_i - \frac{1}{2} \underline{\mu}_i^T \underline{\Sigma}^{-1} \underline{\mu}_i$$

$$i=1,\ldots,K . \qquad \text{(D-3)}$$

These functions are hyperplanes in $\underline{v}$ space. Note that if $K=2$, only one discriminant function, $h(\underline{v}) = h_1(\underline{v}) - h_2(\underline{v})$, is needed. If $h(\underline{v}) > 0$, $\underline{v}$ is classified into class 1. If $h(\underline{v}) < 0$, $\underline{v}$ is classified into class 2. For $K = 2$, $h(\underline{v})$ is given by

$$h(\underline{v}) = h_1(\underline{v}) - h_2(\underline{v}) = \log \frac{\xi_1}{\xi_2} \underline{v}^T \underline{\Sigma}^{-1} (\underline{\mu}_1 - \underline{\mu}_2) -$$

$$\frac{1}{2} \underline{\mu}_1 \underline{\Sigma}^{-1} \underline{\mu}_1 + \frac{1}{2} \underline{\mu}_2^T \underline{\Sigma}^{-1} \underline{\mu}_2 . \qquad \text{(D-4)}$$

The decision boundary, which is obtained by setting $h(\underline{v})$ equal to zero, is a hyperplane which bisects the line connecting the means. The hyperplane is inclined to this line at an angle determined by the covariance matrix and the relative positions of the means. If the a priori probabilities of each class are equal and the covariance matrices are proportional to the identity matrix (i.e. $\underline{\Sigma}_i = \gamma \underline{I}$, where $i = 1,2$ and γ is a constant), equation D-4 becomes

$$\frac{h(v)}{\alpha} = \underline{v}^T (\underline{\mu}_1 - \underline{\mu}_2) - \frac{1}{2} |\underline{\mu}_1|^2 + \frac{1}{2} |\underline{\mu}_2|^2 . \qquad \text{(D-5)}$$

This is the equation for a hyperplane which is the perpendicular bisector of the line connecting the means of the two distributions.

If slight changes are made in the assumptions, different decision

boundaries result. For example, if the class densities are spherically symmetric Gaussian densities which differ only in location and scale, the hypersphere is the optimum boundary. Cooper (1962), (1963) further investigates the hyperplane and hypersphere as decision boundaries.

If the mean and/or the variance is unknown, a "good" estimate for these statistics can be used in place of the actual mean and variance. Such an estimate for the mean is

$$\hat{\underline{\mu}}(j) = \frac{1}{N_j} \sum_{i=1}^{N_j} \underline{v}_i(j) \qquad \text{(D-6)}$$

where $\hat{\underline{\mu}}(j)$ is the estimate of the mean vector of the j^{th} class, N_j is the number of observations of the j^{th} class, and $\underline{v}_i(j)$ is the i^{th} observation vector of the j^{th} class. A "good" estimate for the covariance matrix is

$$\hat{\underline{\Sigma}}(j) = \frac{1}{N_j - 1} \sum_{i=1}^{N_j} [\underline{v}_i(j) - \hat{\underline{\mu}}(j)][\underline{v}_i(j) - \hat{\underline{\mu}}(j)]^T . \qquad \text{(D-7)}$$

II. Estimation or Approximation of the Probability Densities

Many of the authors in the statistical literature consider the class of probability density estimators,

$$\hat{f}_n(v) = \frac{1}{n} \sum_{j=1}^{n} w(v, V_j) . \qquad \text{(D-8)}$$

$\hat{f}_n(v)$ is the estimate of the probability density $f(v)$ when n observations $V_1, \ldots, V_n$ are available from $f(v)$. The function $w(v, V_j)$ might take a form such as

$$w\left(\frac{v-V_j}{h}\right) = \begin{cases} \frac{1}{2} & \text{if } \left|\frac{v-V_j}{h}\right| \leq 1 \\ 0 & \text{if } \left|\frac{v-V_j}{h}\right| > 1 \end{cases} \qquad \text{(D-9)}$$

where h is some function of n. Other weighting functions are given in Table 1, p. 1068 of Parzen (1962).

Fix and Hodges (1951) use the following equation to estimate the probability density,

$$\hat{f}_n(v) = \frac{\{F_n(v+h) - F_n(v-h)\}}{2h} \qquad \text{(D-10)}$$

where

$$F_n(v) = \frac{\text{(the number of observations} < v)}{n} .$$

The parameter h is a function of n which approaches zero as n approaches infinity. The form of h which gives the "best" results has not been determined. Rosenblatt (1956) lets $h = \beta n^{-\alpha}$ and obtains the β and α which minimize the expected mean square error.

Loftsgaarden and Quesenberry (1965) propose an estimator similar to (D-10) for estimating the density function. Their estimator is

$$\hat{f}_n(v) = \frac{k(n) - 1}{2h} . \qquad \text{(D-11)}$$

where $k(n)$ is some integer which is less than n. That is, they specify some number $k(n)$. They then calculate the distance between v and the $k(n)^{th}$ closest observation to v. This distance is substituted for h.

Then rather than specify some distance h and find the number of observations k, as Fix and Hodges do, they specify some number of observations k and find the distance h. Consistency is shown for this estimator when $k(n) \longrightarrow \infty$ and $k(n)/n \longrightarrow 0$ as $n \longrightarrow \infty$. They indicate that $k(n) = n^{1/2}$ gave good results on some empirical work.

These concepts are readily applicable for estimating the density function for a D-dimensional vector $\underline{v}$. For example, if Euclidean distance is used, equation D-10 becomes

$$\hat{f}_n(\underline{v}) = \frac{\left[\begin{array}{l}\text{the number of observations contained} \\ \text{in a hypersphere of radius } r \text{ from } \underline{v}\end{array}\right]}{n\left[2r^D \pi^{D/2} \Big/ D\,\Gamma(D/2)\right]} . \qquad \text{(D-12)}$$

Loftsgaarden and Quesenberry's estimator becomes

$$\hat{f}_n(\underline{v}) = \frac{[k(n) - 1]}{n\left[2r^D \pi^{D/2} \Big/ D\,\Gamma(D/2)\right]} . \qquad \text{(D-13)}$$

One sees that these procedures are very easily used in classification. Rather than compute the class probability densities at a point, one need only count the number of observations of each class within an appropriate radius from the new observation $\underline{V}$. Under certain costs, a priori probabilities, and number of observations from each class, the new observation is assigned to the class which is most heavily represented by these observations. This procedure, called the k_n nearest-neighbor rule, is further discussed in the last section of this appendix.

Sebestyen (1962a), (1962b) proposes a histogram approach which

involves the approximation of the probability density with many Gaussian subdensities. Consider this method which is called "adaptive sample set construction" by Sebestyen. Suppose we wish to distinguish events of class F from those of class G. Let us approximate each class probability density with many spherically-symmetric multivariate Gaussian densities. Suppose we use K_F subclasses to represent class F and K_G subclasses to represent class G. The decision is made in favor of class F if

$$\sum_{i=1}^{K_F} p(F_i)\, p_{F_i}(\underline{v}) > C \sum_{j=1}^{K_G} p(G_j)\, p_{G_j}(\underline{v}) \qquad \text{D-14)}$$

where C is a constant, $p(F_i)$ is the apriori probability of subclass F_i, and $p_{F_i}(\underline{v})$ is the conditional probability of subclass F_i. Since the apriori probability of a certain subclass is not known, it is estimated. Letting M_{F_i} be the number of observations in F_i and M_{G_j} be the number of observations in subclass G_j, the decision rule for deciding in favor of class F is

$$\sum_{i=1}^{K_F} M_{F_i} \exp\left\{\frac{-\sum_{n=1}^{D}[v_n - m_n(F_i)]^2}{2\sigma^2}\right\} > C' \sum_{j=1}^{K_G} M_{G_j}$$

$$\exp\left\{\frac{-\sum_{n=1}^{D}[v_n - m_n(G_j)]^2}{2\sigma^2}\right\} . \qquad \text{(D-15)}$$

Here, $m_n(F_i)$ is the n^{th} coordinate of the mean of subclass F_i, D is the number of coordinates, and σ^2 is the variance. The selection of the

means of the subclasses can be made as follows. Let us introduce the first training observation $\underline{V}^1$. Suppose that $\underline{V}^1$ is a member of class F. We assign $\underline{V}^1$ to subclass F_1 with mean $\underline{m}_{F_1} = \underline{V}^1$ and variance σ^2. The value of the variance σ^2 is arbitrarily chosen. At this point in the procedure, M_{F_1} in the above equation is equal to one. Now we introduce the second observation $\underline{V}^2$. If $\underline{V}^2$ is a member of class F and within a radius T (its value is arbitrary) of $\underline{m}_{F_1}$, we set $M_{F_1} = 2$ and let $\underline{m}_{F_1}$ equal the mean of the 1st two observations. If $\underline{V}^2$ is a member of class F but lies outside the sphere of radius T with center at $\underline{V}^1$, we assign $\underline{V}^2$ to subclass F_2 with mean $\underline{m}_{F_2} = \underline{V}^2$ and variance σ^2. If $\underline{V}^2$ is a member of class G, we assign the sample to subclass G_1 with mean $\underline{m}_{G_1} = \underline{V}^2$ and variance σ^2. This process is continued until all training observations are exhausted. It is seen that this procedure approximates the class probability densities by many Gaussian subdensities. The degree of approximation depends on the original class probability densities, the variance σ^2, the radius T, and the order in which the samples are introduced. Waltz and Fu (1965) have used this general idea along with the gradual reduction of each subset radius to facilitate a more precise boundary.

Sebestyen and Edie (1966) have devised a scheme for estimating a multi-dimensional density using hyperellipsoids as estimation cells. Their method allows the size and the shape of the histogram cells to be influenced by the local distribution of the data. The initial size and shape of the first histogram cell is chosen arbitrarily. The updating and

generating of new cells is similar to Sebestyen's sample set construction. A difference arises in the handling of observations which fall outside existing cells but nevertheless "close" to the boundary. These events are stored for processing at a later time when the average number of elements per cell reaches a certain threshold. The cell size may increase or decrease as more training observations are received for classification. One problem with this procedure is that there are many variables which must be determined by trial and error. When the initial training is finished, one may decide that too few or too many cells have been generated. Then certain parameters must be changed and the training procedure redone.

Aizerman et al. (1964b) use the method of potential functions (orthogonal functions) to approximate an unknown probability density. They assume that the probability $f(v)$ exists and that a finite number N of orthonormal functions $\phi_i(v)$ can be selected so that $f(v) = \sum_{i=1}^{N} C_i \phi_i(v)$. A training algorithm is proposed so that, as the number of independent, identically distributed training observations approaches infinity, the resulting function will converge in probability to $f(v)$.

III. Estimation or Approximation of the Class Discriminating Boundaries

The object here is to assume a form for the decision boundary and to locate the boundary so that the best possible recognition is obtained. A simple boundary like a hyperplane or hypersphere is usually employed.

Let us first consider the use of the hyperplane. This has been

thoroughly treated in the literature, Highleyman (1962), Cooper (1962), (1963), Albert (1963), Peterson and Mattson (1966), Wolff (1966). We have seen that the hyperplane is the optimum boundary for discriminating between 2 classes which are described by normal distributions with equal covariance matrices and different means. It can also be seen that the hyperplane is the optimum boundary for two classes which are equally probably a priori, have equal costs of misrecognition, and have probability densities which are ellipsoidally symmetric with equal eccentricities and monotonically decreasing from the mean, c.f. Cooper (1962).

It turns out that the hyperplane is the optimum discriminant for other cases. For example, suppose we wish to discriminate between two classes where the sample vectors $\underline{v}$ consist of D binary components, either zero or one. If the components of $\underline{v}$ are statistically independent, a linear discriminant function is optimum, c.f. Minsky (1961), Nilson (1965).

One needs, in general, $\frac{K(K-1)}{2}$ hyperplanes to separate K classes. Some of these hyperplanes may not be needed depending on the location and shape of the pattern classes. For example, it may be possible to separate each class from all the remaining classes. In this case only K-1 hyperplanes are required.

Suppose that we wish to distinguish between 3 classes by using hyperplanes. A general two-dimensional situation is shown in Figure D-2. The classes are represented by the circles labeled (1), (2), and (3). Boundary B_{ij} separates class i from class j. Two problems become

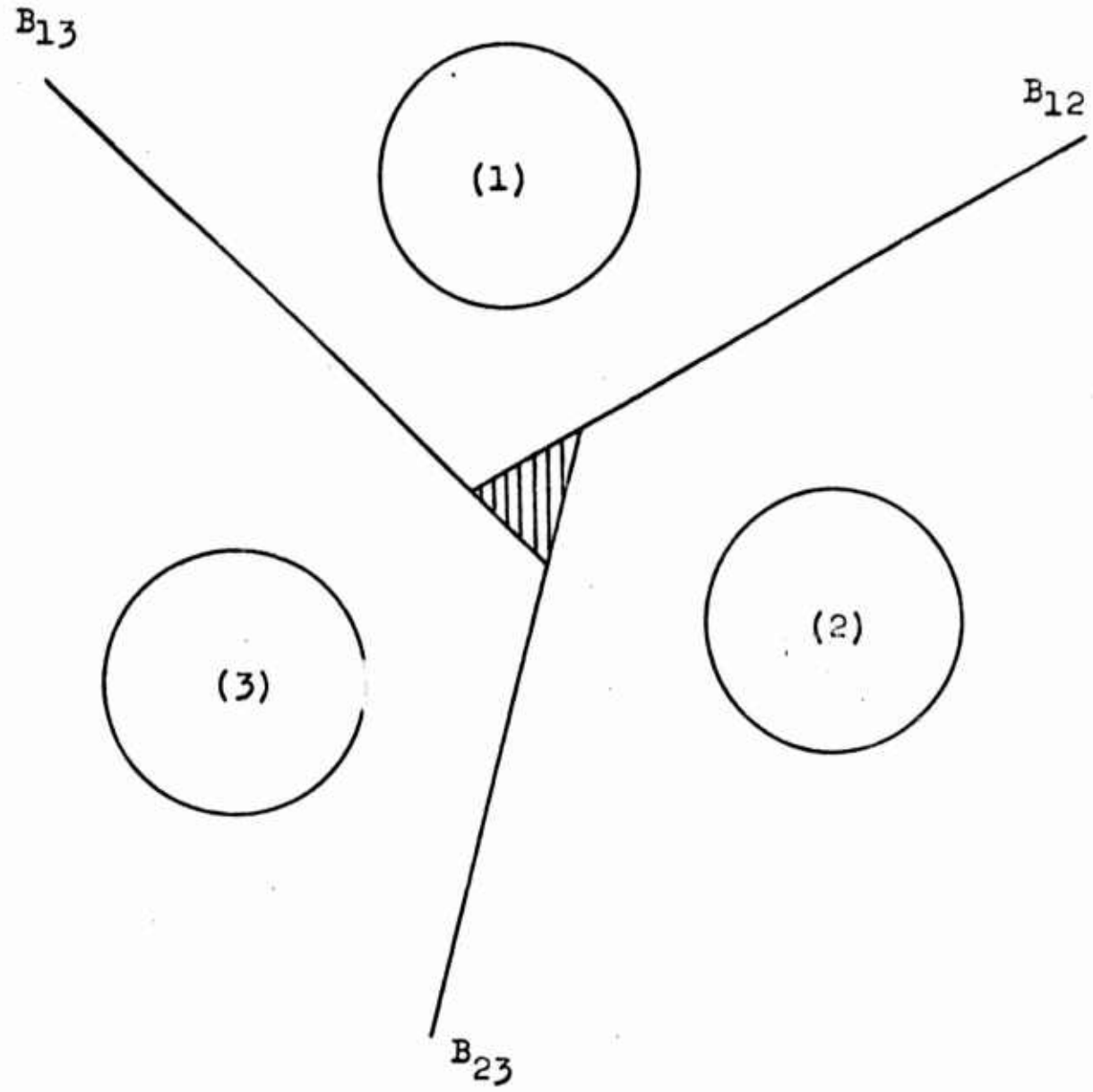


Figure D-2. A Partitioning with Linear Decision Functions.

evident. (1) For best results, the 3 hyperplanes (lines in the figure) cannot be positioned independently of one another. (2) A region may appear in which an observation may be classified as belonging to any of the three classes. This region is known by many names such as void region, region of indecision, deferred decision region, or reject region. This is the crosshatched region of Figure D-2.

Consider problem (1). Suppose the objective is to minimize the number of training observations which are misclassified. For best performance, the boundaries should be determined simultaneously. However, the simultaneous location of $\frac{K(K-1)}{2}$ hyperplanes for minimum misclassification is often very difficult. Hence each hyperplane is usually positioned sequentially. After the hyperplanes are located, the results may not be as good as expected. In this case Highleyman (1962) suggests an iterative procedure in which all subsequent hyperplanes are located by using only the observations which are correctly classified by the previously located hyperplanes. That is, if B_{12} in Figure D-2 is located first, then B_{23} is located using only the observations of class 2 which are correctly classified by B_{12}.

Now consider the void region in which a new observation may be classified into class 1, 2, or 3. For example, an observation in the crosshatched region of Figure D-2 lies to the class 1 side of B_{12} and to the class 3 side of B_{13}. Therefore if this observation is compared to B_{12} and then to B_{13}, it is classified into class 3. However, if the observation is compared to B_{23} and then B_{12}, it is classified into

class 1. Note that a void region will not occur if the class probability densities are Gaussian with equal variances and different means. This is because the optimum decision hyperplanes in this case are the perpendicular bisectors of the lines joining the means of the classes. Note also that if the observations are transformed into a space where the likelihood ratios act as coordinates, no void region results when Bayes criterion is used, c.f. Van Trees (1968).

It was shown in equations D-1, D-2, and D-3 that the hyperplane is the optimum decision surface for separating two classes which have Gaussian distributions with different means and equal covariance matrices. Anderson and Bahadur (1962) have investigated linear procedures for classifying observations from Gaussian distributions with unequal covariance matrices. They give methods for constructing a hyperplane which minimizes one probability of misclassification, given the other, and for constructing the optimum hyperplane when a minimax criterion is used.

Now assume that $f_i(\underline{v})$ is unknown as it is in most pattern recognition problems. Highleyman (1962) proposes that the optimum hyperplane be determined by a search through a set of hyperplanes for one which minimizes the maximum likelihood estimate of the expected risk. He suggests that the expected risk be estimated by

$$\frac{C_i(j)\, e_i(j) + C_j(i)\, e_j(i)}{N},$$

where $e_i(j)$ is the number of observations from class i which are classified into class j and N is the total number of observations which are

used in the estimate. If the apriori probabilities ξ_i are known beforehand, the number of class i observations, n_i, which are used in making the estimate should be determined by $n_i = \xi_i N$.

It is seen that the estimate of the expected risk is a discontinuous function because it is determined by a finite number of discrete observations. This prohibits the use of a gradient method in a search for the minimum risk. Highleyman chooses to approximate the abrupt change in risk, when the hyperplane is moved from one side of an observation to the other, by a continuous function of distance from the observation to the hyperplane. A convenient approximation for this step function is the Gaussian cumulative distribution with mean at the hyperplane location and variance σ^2. The risk function, which is the sum of these functions over all observations, is then minimized with respect to the hyperplane coordinates by the method of steepest descent. The variance σ^2 is reduced and the process repeated until the desired recognition accuracy is achieved.

Wolff (1966) states that the method of steepest descent contains the inherent disadvantage that a relative instead of an absolute minimum may be obtained. He employs a variant of the "creeping random method" by Brooks (1958) which Wolff states is more appropriate for discontinuous functions and has less chance of yielding a relative minimum. The equation of a hyperplane in D dimensions is $\underline{w}^T \underline{v} = 0$, where $\underline{w}^T = [w_0\ w_1 \ldots w_D]$ and $\underline{v} = \begin{bmatrix} 1 \\ v_1 \\ \vdots \\ v_D \end{bmatrix}$. Wolff's method consists of first

selecting a starting position for the hyperplane. Let the initial weight vector $\underline{w}$ be $\underline{w}^0$. The error rate is then determined for this hyperplane. The hyperplane is given a displacement (rotation and/or translation) and the error rate is calculated for the hyperplane in its new position. Let the new position be described by $\underline{w}^1 = \underline{w}^0 + \Delta\underline{w}^1$. If the error rate in the new position is less than the error rate in the former position, the hyperplane is given a displacement from the new position. In this case $\underline{w}^2 = \underline{w}^1 + \Delta\underline{w}^2$. Otherwise the hyperplane is given a displacement from the former position ($\underline{w}^1 = \underline{w}^0 + \Delta\underline{w}^2$). The increments of displacement $\Delta\underline{w}^1$, $\Delta\underline{w}^2$,... are chosen from a random number generator. Wolff uses a Gaussian distribution with zero mean and variance σ^2 to generate the random numbers. The variance is held constant in the early stage of the process and then reduced during the final stages.

The displacement of the hyperplane depends upon its starting position $\underline{w}^i$ as well as the increment $\Delta\underline{w}^{i+1}$. For certain $\underline{w}^i$, the increment $\Delta\underline{w}^{i+1}$ needs to be larger than it would for other $\underline{w}^i$ for the same relative displacement. Wolff eliminates this problem by describing the hyperplane by a point on a unit sphere in $\underline{w}$ space. Such a point is given by

$$w_i = \begin{cases} \cos(\phi_1) & i = 0 \\ \cos(\phi_{i+1}) \prod_{r=1}^{i} \sin(\phi_r) & i \neq 0, D \\ \prod_{r=1}^{D} \sin(\phi r) & i = D \end{cases} .$$

He then uses the creeping random method to increment the ϕ. Questions remain about (1) the rate of convergence, (2) termination of the search, and (3) the best statistics for the random number generator when the creeping random method is used.

IV. Other Intuitive Criteria

Many ad hoc optimization criteria have been proposed for the solution of pattern recognition problems. Most of these criteria are intuitively appealing and offer a so-called "optimum" solution without the use of the class probability densities.

Consider a criterion which was proposed by R. A. Fisher (1925). Suppose there are two classes in D-dimensional space. Suppose it is desirable to project these classes onto a line so that the "distance" between the classes is as large as possible. A threshold can be set along this line and classification of new observations begun.

To make the "distance" between the classes large, Fisher maximizes the scatter between the classes while keeping the scatter among observations of the same class constant. Let $\underline{W}$ be the linear transformation to do this. The problem then is to maximize

$$\underline{V}_i \sum_{\epsilon \text{ class } 1} (\underline{W}^T \underline{V}_i - \underline{W}^T \underline{V}_j)^2$$
$$\underline{V}_j \epsilon \text{ class } 2$$

while constraining the following equation to be equal to a constant.

$$\sum_{\substack{\underline{V}_i \in \text{class 1} \\ \underline{V}_j \in \text{class 1}}} (\underline{W}^T \underline{V}_i - \underline{W}^T \underline{V}_j)^2 + \sum_{\substack{\underline{V}_i \in \text{class 2} \\ \underline{V}_j \in \text{class 2}}} (\underline{W}^T \underline{V}_i - \underline{W}^T \underline{V}_j)^2 = C \ .$$

The solution $\underline{W}$ is the eigenvector of the largest eigenvalue of

$$(BA^{-1} - \lambda I) = 0 \ ,$$

where

$$A = \sum_{\underline{V}_i \in \text{class 1}} (\underline{V}_i - \underline{m}_1)(\underline{V}_i - \underline{m}_1)^T + \sum_{\underline{V}_1 \in \text{class 2}} (\underline{V}_i - \underline{m}_2)(\underline{V}_i - \underline{m}_2)^T$$

is the intraset scatter matrix, and B is the interset scatter matrix,

$$B = \sum_{\text{all classes}} \left(\underline{V}_i - \frac{\underline{m}_1 + \underline{m}_2}{2}\right)\left(\underline{V}_i - \frac{\underline{m}_1 + \underline{m}_2}{2}\right)^T - A \ .$$

$\underline{m}_1$ and $\underline{m}_2$ are the sample means of class 1 and class 2, respectively.

Sebestyen (1961), (1962a) advocates maximizing the scatter between classes (interset distance) while keeping the total scatter constant (sum of interset and intraset distances). As expected, this method also yields an eigenvalue problem $(BA^{-1} - \lambda I) = 0$ where A and B are now given by

$$A = \sum_{\text{all classes}} (\underline{V}_i - \underline{V}_j)(\underline{V}_i - \underline{V}_j)^T$$

and

$$B = \sum_{\substack{\underline{V}_i \in \text{class 1} \\ \underline{V}_j \in \text{class 2}}} (\underline{V}_i - \underline{V}_j)(\underline{V}_i - \underline{V}_j)^T \ .$$

One problem with these methods is that the inverse of a large matrix has to be computed. If one decides to minimize the total scatter while con-

straining $\underline{W}$ (e.g. $\sum_i W_i^2 = 1$), A^{-1} disappears from the equation. However, this approach does not yield a suitable solution to such a simple problem as shown in Figure D-3. The ellipses labeled (1) and (2) represent the contours of constant probability for the classes. A line which suitably discriminates between the two classes is labeled as the "ideal discriminant". This line is contrasted with the line which minimizes the total scatter.

Sebestyen (1962a) offers a nonlinear approach to discriminating between patterns. He approximates a generalized discriminant with a polynomial function of the coordinates of the measurement space. A search is performed over polynomials of various order starting with the first order polynomial and continuing with higher order polynomials until a suitable categorizer is found. However, for a high dimensional space and a high order polynomial, this approach can be very time consuming.

Widrow and Hoff (1960) introduce a performance criterion which states that in a two class problem the distance within the classes should be minimized about two fixed points. They devise an iterative procedure for the linear separation of binary patterns. Patterson and Womack (1966) use this criterion and a nonlinear discriminant function. They assume a discriminant function of the form $u(\underline{W}, \underline{V}) = \sum_{i=1}^{n} W_i \phi_i(\underline{V})$, where the $\phi_i(\underline{v})$ are given. They train the machine to approximate a discriminant function which maps class 1 observations to point K_1 and class 2 observations to point $-K_2$. A search technique is used to minimize the mean square deviation of $u(\underline{W}, \underline{V})$ from these points. That is, they minimize

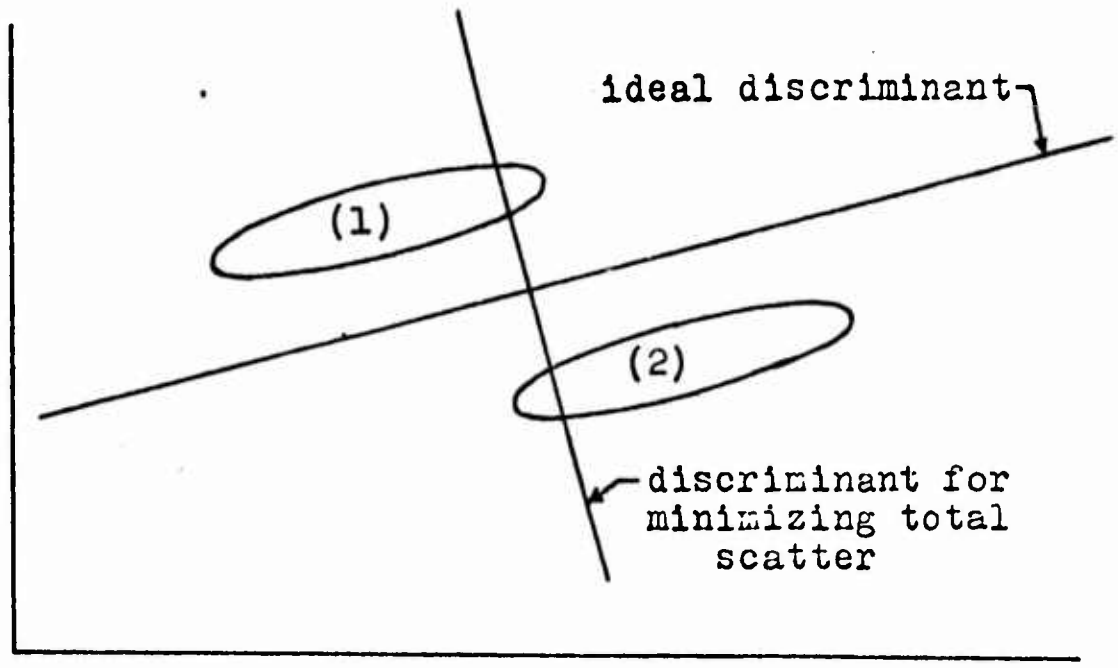


Figure D-3. Discriminant Functions.

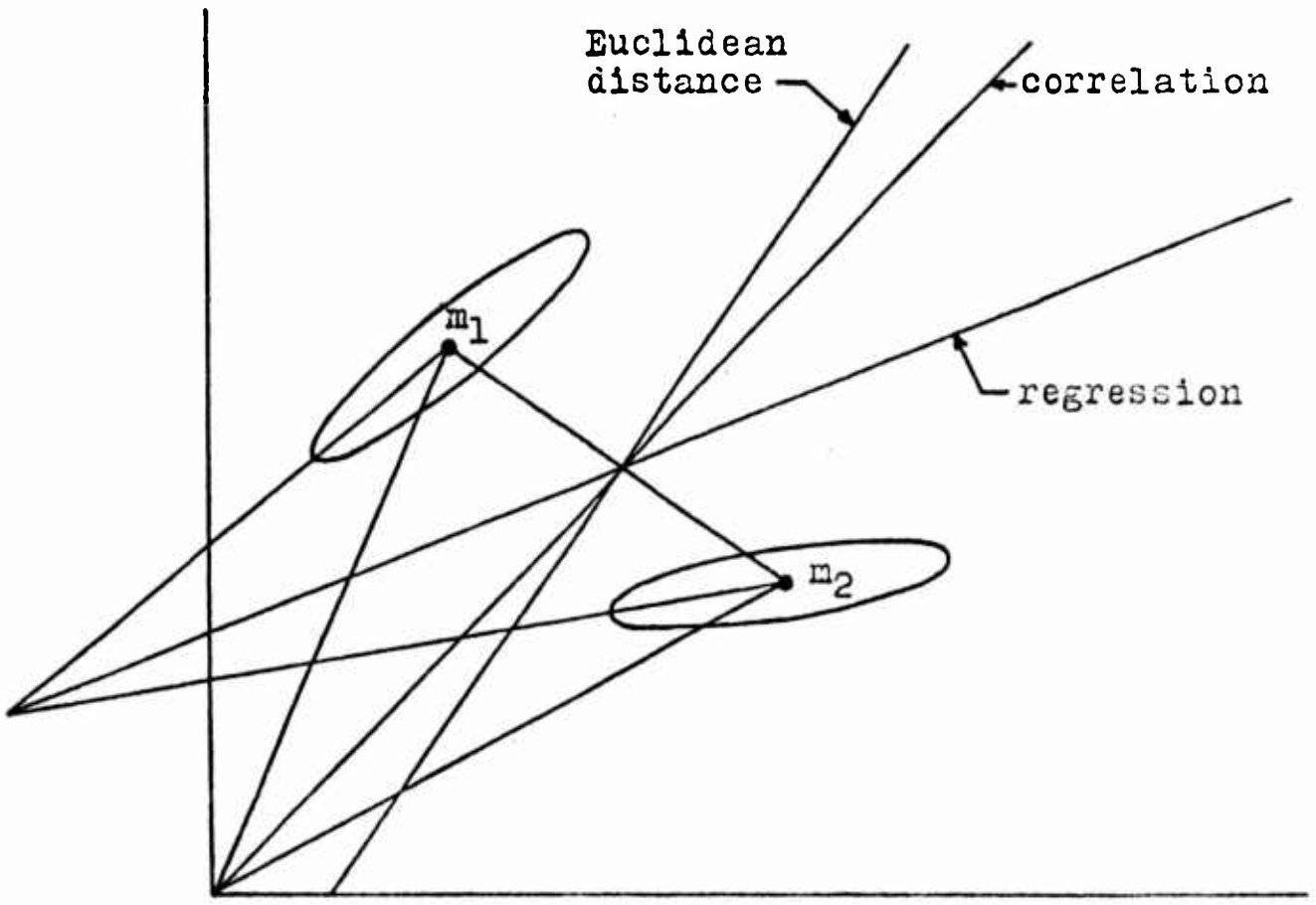


Figure D-4. Elementary Decision Rules.

$$\langle [u(\underline{W}, \underline{V}) - K_1]^2 \rangle_{\text{class 1}} + \langle [u(\underline{W}, \underline{V}) + K_2]^2 \rangle_{\text{class 2}}$$

where $\langle \cdot \rangle_{\text{class i}}$ denotes the average over the class i observations.

There are many simple decision rules which are based upon regression, Euclidean distance, or correlation. For example, a mean square regression line (curve, surface, etc.) can be determined for each class. One then decides that a new observation $\underline{V}$ is a member of a certain class if it is closer to the regression line of that class than to the regression line of any other class.

Another simple decision rule is based upon the Euclidean distance from $\underline{V}$ to a characteristic point of the classes (e.g. the class sample means). The rule decides that $\underline{V}$ belongs to a certain class if $\underline{V}$ is closer to the characteristic point of that class than to the corresponding characteristic point of any other class.

In still another rule it is decided that $\underline{V}$ belongs to class 1 if its dot product (i.e. correlation) with the sample mean of that class is larger than its dot product with the sample mean of any other class. Figure D-4 illustrates these procedures based on the sample means $\underline{m}_1$ and $\underline{m}_2$ of class 1 and class 2, respectively. One can see that these simple schemes only work well for certain class configurations.

Fix and Hodges (1951) propose that an unknown observation $\underline{V}$ be classified as a member of the class of its nearest neighbor (as described by an arbitrary metric). This is the famous nearest-neighbor (NN) rule. A more complicated rule, the k_n nearest-neighbor (k_n- NN) rule,

assigns an unclassified point to the class most heavily represented among its k_n-nearest neighbors. Fix and Hodges establish the consistency of the k_n-NN rule for sequences $k_n \longrightarrow \infty$ such that $\frac{k_n}{n} \longrightarrow 0$. Also, Fix and Hodges (1952) present a numerical investigation of the small sample performance of the NN rule and the 3-NN ($k_n=3$) rule under the assumption of normal statistics.

Cover and Hart (1967) show that the probability of error for the nearest-neighbor rule is less than twice the Bayes probability of error, based on an infinite sample. They further demonstrate that the NN rule is admissible among the k_n-NN rules for certain classes of distributions. These are the classes of distributions for which the distance between any two elements of the same class is less than the distance between any two elements of different classes. This, of course, rules out any possibility of overlap between the classes. Consider a demonstration of the admissibility of the NN rule among the k_n-NN rules. Suppose one class is uniformly distributed over the interval $[-2, -1]$ and the other class is uniformly distributed over the interval $[1, 2]$, both on the real line. Suppose that the apriori probability of occurrence of each class is equal to $\frac{1}{2}$. Let n training observations be taken. Suppose that a new observation, which is to be classified, falls in the interval $[1, 2]$. To make an error by the NN rule all of the n training observations must fall in $[-2, -1]$. The probability of this occurring is $(\frac{1}{2})^n$. To make an error by the k_n-NN rule, where k_n is odd, $(k_n-1)/2$ or more of the n

observations must fall in $[-2, -1]$. The probability that $(k_n - 1)/2$ of the n observations fall in $[-2, -1]$ is $(\frac{1}{2})^n \sum_{j=1}^{(k_n-1)/2} \binom{n}{j}$. This is, of course, greater than $(\frac{1}{2})^n$. Thus admissiblity is proved for this set of distributions.

The nearest-neighbor rule is sensitive to spurious information since an annulus is formed about an observation from one class which falls in a region surrounded by observations from different classes. Sebestyen (1962c) mentions two rules which eliminate or greatly reduce the possibility that an annulus forms around spurious observations. These rules have the same effect as the k_n-NN rule in that more than one nearest neighbor is considered. Let $d(\underline{V}, \underline{f}_m)$ be some distance measure from point $\underline{V}$ to the m^{th} element of class F . The 1^{st} rule decides that $\underline{V}$ is a member of class F if

$$\sum_{m=1}^{M_F} \frac{1}{d^k(\underline{V}, \underline{f}_m)} \qquad \sum_{s=1}^{M_G} \frac{1}{d^k(\underline{V}, \underline{g}_s)}$$

where M_F is the number of elements in class F and k is an arbitrary number chosen to determine the neighborhood of $\underline{V}$ which is to influence the decision. The second rule decides in favor of class F if

$$\sum_{m=1}^{M_F} \frac{1}{1 + \left(\frac{d(\underline{V}, \underline{f}_m)}{r} \right)^k} > \sum_{s=1}^{M_G} \frac{1}{1 + \left(\frac{d(\underline{V}, \underline{g}_s)}{r} \right)^k} .$$

Note that for Euclidean distance and for a large k, this rule essentially counts the number of elements from class F contained in a radius r from $\underline{v}$ and compares this with the number of elements from class G in that same radius.

As has been seen, many classification procedures are available for use in pattern recognition. Since these decision procedures are difficult to compare, some ambiguity is involved in the choice of a decision procedure. The procedure to use depends on the amount of information available, the desired complexity of the decision procedure, and personal preference.

BIBLIOGRAPHY

A. Albert, "A Mathematical Theory of Pattern Recognition," Ann. Math. Statist., Vol. 34, No. 1, pp. 284-299 (1963).

T. W. Anderson, An Introduction to Multivariate Statistical Analysis, John Wiley, N. Y., (1958).

T. W. Anderson and R. Bahadur, "Classification into Two Multivariate Normal Distributions with Different Covariance Matrices," Ann. Math. Statist., Vol. 33, pp. 420-431 (1962).

T. W. Anderson, "Some Nonparametric Multivariate Procedures Based on Statistically Equivalent Blocks," Proc. of an International Symposium on Multivariate Analysis, P. R. Krishnaiah, Ed., Academic Press, N. Y., (1966).

R. E. Barlow and F. Proschan, "Tolerance and Confidence Limits Based on Failure Rate," Ann. Math. Statist., Vol. 37, pp. 1593-1601 (1966).

D. Blackwell and M. Girshick, Theory of Games and Statistical Decisions, John Wiley, N. Y., (1954).

R. H. Bolt et. al., "Identification of a Speaker by Speech Spectrograms," Science, Vol. 166, No. 3903, pp. 338-343 (Oct. 1969).

P. D. Bricker and S. Pruzansky, "Effects of Stimulus Content and Duration on Talker Identification," J. Acoust. Soc. Am., Vol. 40, No. 6, pp. 1441-1449 (1966).

S. H. Brooks, "A Discussion of Random Methods for Seeking Maxima," Operations Research, Vol. 6, pp. 244 (1958).

Y. T. Chien and K. S. Fu, "A Modified Sequential Recognition Machine Using Time-varying Stopping Boundaries," IEEE Trans. Information Theory, IT-12, pp. 206-214 (April 1966).

C. K. Chow, "An Optimum Character Recognition System Using Decision Functions," IRE Trans. on Electronic Computers, EC-6, pp. 247-254 Dec. (1957).

William T. Cochran et. al., "What is the Fast Fourier Transform," IEEE Trans. on Audio and Electroacoustics, Vol. AU-15, No. 2, pp. 45-55 (1967).

A. J. Compton, "Effects of Filtering and Vocal Duration upon the Identification of Speakers, Aurally," J. Acoust. Soc. Am., Vol. 35, No. 11, pp. 1748-1752 (1963).

J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," Math. of Comput., Vol. 19, pp. 297-301 (1965).

P. W. Cooper, "The Hyperplane in Pattern Recognition," Cybernetics, Vol. 5, No. 4 (1968a).

_____, "The Hypersphere in Pattern Recognition," Info. and Control, Vol. 5, No. 4, pp. 324-346 (1962b).

_____, "Hyperplanes, Hyperspheres, and Hyperquadrics as Decision Boundaries," Computers and Information Sciences, Spartan Books, Washington, D.C., (1963).

T. M. Cover and P. E. Hart, "Nearest-Neighbor Pattern Classification," IEEE Trans. on Info. Theory, Vol. IT-13, No. 1, pp. 21-27 (1967).

H. Cramér, Mathematical Methods of Statistics, Princeton University Press, Princeton, N. J., (1946).

E. Crow, F. Davis, and M. Maxfield, Statistics Manual, Dover, N. Y. (1960).

Subtra K. Das, "A Method of Decision Making in Pattern Recognition," IEEE Trans. on Computers, C-18, No. 4, pp. 329-333 (April 1969).

H. B. Dwight, Tables of Integrals and Other Mathematical Data, 4th ed., The Macmillan Co., N. Y. (1961).

M. A. Eiserman, E. M. Braverman, and L. I. Boxonoer, "Theoretical Principles of Potential Functions in the Problem of Teaching Automata to Recognize Classes," Automation and Remote Control, Vol. 25, No. 6, pp. 821-837 (1964a).

_____, "The Probability Problem in Teaching Automata to Recognize Classes and the Method of Potential Functions," Automation and Remote Control, Vol. 25, No. 9, pp. 1175-1190 (1964b).

R. A. Fisher, Statistical Methods for Research Workers, Hafner, New York, N.Y., 13th ed., (1963); 1st ed., (1925).

E. Fix and J. L. Hodges, Jr., "Discriminatory Analysis: Nonparametric Discrimination; Consistency Properties," USAF School of Aviation Medicine, Randolph Field, Texas; Project 21-49-004, Rept. 4 Contract AS41 (128)-31, (1951).

_____, "Discriminatory Analysis: Nonparametric Discrimination; Small Sample Performance," USAF School of Aviation Medicine, Project No. 21-49-004, Rept. No. 11, (1952).

D.A.S. Fraser and R. Wormleighton, "Nonparametric Estimation: IV," Ann. Math. Statist., Vol. 22, pp. 294-298 (1951).

D.A.S. Fraser, "Sequentially Determined Statistically Equivalent Blocks," Ann. Math. Statist., Vol. 22, pp. 372-381 (1951).

_____, "Nonparametric Tolerance Regions," Ann. Math. Statist., Vol. 24, pp. 44-55 (1953).

_____, Nonparametric Methods in Statistics, Wiley, N.Y., (1957).

K. S. Fu and C. H. Chen, "A Sequential Decision Approach to Problems in Pattern Recognition and Machine Learning," presented at the 3rd Symposium on Discrete Adaptive Processes, Chicago, Ill., (1964).

K. S. Fu and Y. T. Chien, "Sequential Recognition Using a Nonparametric Ranking Procedure," IEEE Trans. Information Theory, Vol. IT-18, pp. 484-492 (1967).

K. S. Fu, Sequential Methods in Pattern Recognition and Machine Learning, Academic Press, N. Y., (1968).

J. W. Glenn and N. Kleiner, "Speaker Identification Based on Nasal Phonation," J. Acoust. Soc. Am., Vol. 43, pp. 368-372 (1968).

Irwin Guttman, "Tolerance Regions, a Survey of Its Literature. Introduction and Distribution-free Tolerance Regions," Report No. 123, Dept. of Statistics, University of Wisconsin, Madison, Wis., (1967a).

_____, "Tolerance Regions, a Survey of Its Literature, II. β-expectation Tolerance Regions," Report No. 124, Dept. of Statistics, University of Wisconsin, Madison, Wis., (1967b).

_____, "Tolerance Regions, a Survey of Its Literature III. β-content Tolerance Regions," Report No. 125, Dept. of Statistics, University of Wisconsin, Madison, Wis., (1967c).

Irwin Guttman, "Construction of β-content Tolerance Regions at Confidence Level γ for Large Samples from the k-variate Normal Distribution," Technical Report No. 128, Dept. of Statistics, University of Wisconsin, Madison, Wis., (1968).

D. L. Hanson and L. H. Koopmans, "Tolerance Limits for the Class of Distributions with Increasing Hazard Rates," Ann. Math. Statist., Vol. 35, pp. 1561-1570 (1964).

W. A. Hargreaves and J. A. Starkweather, "Recognition of Speaker Identity," Lang. and Speech, Vol. 6, pp. 63-67 (1963).

J. K. Hawkins, "Self-organizing Systems – A Review and Commentary," Proc. IRE, Vol. 49, No. 1, pp. 31-48 (1961).

W. H. Highleyman, "Linear Decision Functions with Applications to Pattern Recognition," Proc. IRE, Vol. 50, No. 6, pp. 1501-1514 (1962).

Gary L. Holmgren, "Speaker Recognition, Speech Characteristics, Speech Evaluation, and Modification of Speech Signal – A Selected Bibliography." IEEE Trans. on Audio and Electroacoustics, Vol. AU-14, No. 1, pp. 32-39 (1966).

R. L. Kashyap and C. C. Blaydon, "Estimation of Probability Density and Distribution Functions," IEEE Trans. on Info. Theory, IT-14 (July 1968); pp. 549-556 (1968).

J. H. B. Kemperman, "Generalized Tolerance Limits," Ann. Math. Statist., Vol. 27, pp. 180-186 (1956).

M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, Vol. I, Hafner, N. Y., (1958).

_____, The Advanced Theory of Statistics, Vol. II, Hafner, N. Y. (1961).

L. G. Kersta, "Voiceprint Identification," paper presented at the 64th meeting of the Acoustical Soc. of Am., Vol. 34, p. 725(A), (1962a). See also Nature, Vol. 196, pp. 1253-1257 (1962).

_____, "Voiceprint Identification Infallibility," paper presented at the 64th meeting of the Acoustical Soc. of Am., Vol. 34, p. 1978 (1962b).

_____, "Voiceprint Classification," J. Acoust. Soc. Am., Vol. 37, p. 1217(A), (1965).

L. G. Kersta, "Voiceprint Classification for an Extended Population," J. Acoust. So Am., Vol. 39, p. 1239(A), (1966).

Peter Ladefoged and Ralph Vanderslice, "The Voiceprint Mystique," Working Papers in Phonetics (Dept. of Linguistics, Univ. of California, Los Angeles), pp. 126-142 (1967).

K. P. Li, J. E. Damman, and W. D. Chapman, "Experimental Studies in Speaker Verification, Using an Adaptive System," J. Acoust. Soc. Am., Vol. 40, pp. 966-978 (1966).

D. O. Loftsgaarden and C. P. Quesenberry, "A Nonparametric Estimate of a Multivariate Density Function," Ann. Math. Statist., Vol. 36, pp. 1049-1051 (1965).

James E. Luck, "A Study of Spectral Speech Data – An Examination of the Segmentation and Recognition Problem," doctoral dissertation, Yale University, New Haven, Conn., (1967).

_____, "Computer Program EDIT," research note 740, Project 507-01, Applied Research Laboratory, Sylvania Electronic Systems, Waltham, Mass., (1968a).

_____, "Progress Report on an Automatic Speaker Verification System," research note 749, Projects 711, 507-01, Applied Research Laboratory, Sylvania Electronic Systems, Waltham, Mass., (1968b).

_____, "Description of a Completely Automatic Speaker Verification System," paper presented at the 77th Meeting of the Acoustical Soc. of Am., Vol. 46, p. 90 (1969).

Bertil Malmberg, Phonetics, Dover, N. Y., (1963).

T. Marill and D. M. Green, "Statistical Recognition Functions and the Design of Pattern Recognizers," IRE Trans. on Electronic Computers, EC-9, 472-477 (1960).

M. Minsky, "Steps Toward Artificial Intelligence," Proc. IRE, Vol. 49, No. 1, pp. 8-30 (1961).

R. B. Murphy, "Nonparametric Tolerance Limits," Ann. Math. Statist., Vol. 19, pp. 581-588 (1948).

V. K. Murthy, "Estimation of Probability Density," Ann. Math. Statist., Vol. 36, pp. 1027-1031 (1965).

N. J. Nilsson, Learning Machines, McGraw-Hill, New York (1965).

E. Parzen, "On Estimation of a Probability Density and Modes," Ann. Math. Statist., Vol. 33, pp. 1065-1076 (1962).

E. A. Patrick, "Distribution-free Minimum Conditional Risk Learning Systems," 1967 Princeton Conference on Info. Theory, Purdue University, Report No. TR-EE-66-18, (1966).

E. A. Patrick and F. P. Fischer II, "Introduction to the Performance of Distribution-free, Minimum Conditional Risk Learning Systems," Purdue University, Report No. TR-EE-67-12, (1967).

J. D. Patterson and B. F. Womack, "An Adaptive Pattern Classification System," IEEE Trans. on Systems Science and Cybernetics, Vol. SSC-2, No. 1, pp. 62-67 (1966).

E. Paulson, "A Note on Tolerance Limits," Ann. Math. Statist., Vol. 14, pp. 90-93 (1943).

Karl Pearson, editor, Tables of the Incomplete Beta Function, Cambridge University Press (1934).

D. W. Peterson and R. L. Mattson, "A Method of Finding Discriminant Functions for a Class of Performance Criteria," IEEE Trans. on Info. Theory, Vol. IT-12, No. 3, pp. 380-387 (1966).

I. Pollack, J. M. Pickett and W. H. Sumby, "On the Identification of Speakers by Voice," J. Acoust. Soc. Am., Vol. 26, No. 3, pp. 403-406 (1954).

S. Pruzansky, "Pattern-matching Procedure for Automatic Talker Recognition, J. Acoust. Soc. Am., Vol. 35, No. 3, pp. 354-358 (1963).

J. Pruzansky and M. V. Mathews, "Talker Recognition Procedure Based on Analysis of Variance," J. Acoust. Soc. Am., Vol. 36, pp. 2041-2047 (1964).

C. P. Quesenberry and M. P. Gessaman, "Nonparametric Discrimination Using Tolerance Regions," Ann. Math. Statist., Vol. 39, No. 2, pp. 664-673 (1968).

G. S. Ramishvilli, "Automatic Voice Recognition," Eng. Cybern., No. 5 (1966).

H. Robbins, "On Distribution-free Tolerance Limits in Random Sampling," Ann. Math. Statist., Vol. 15, p. 214 (1944).

F. Rosenblatt, "On the Convergence of Reinforcement Procedures in Simple Perceptrons," Cornell Aeronautical Lab Report VG-1196-G-4, Buffalo, N. Y., (Feb. 1960). See also VG-1196-G-1,2,3.

_____ ___, Principles of Neutodynamics: Perceptrons and the Theory of Brain Mechanisms, Sparton, Washington, D.C., (1961).

M. Rosenblatt, "Remarks on Some Nonparametric Estimates of a Density Function," Ann. Math. Statist., Vol. 27, pp. 832-837 (1956).

H. Scheffé and J. W. Tukey, "Nonparametric Estimation: I. Validation of Order Statistics," Ann. Math. Statist., Vol. 16, pp. 187-192 (1945).

G. S. Sebestyen, "Recognition of Membership in Classes," IRE Trans. on Info. Theory, Vol. IT-7, pp. 44-50 (Jan. 1961).

_____, Decision Making Processes in Pattern Recognition, The Macmillan Co., New York (1962a).

_____, "Pattern Recognition by Adaptive Process of Sample Set Construction," IRE Trans. on Info. Theory, Vol. IT-8, pp. S82-S91 (Sept. 1962b).

G. Sebestyen and J. Edie, "An Algorithm for Nonparametric Pattern Recognition," IEEE Trans. on Electronic Computers, Vol. EC-15, No. 6 (Dec. 1966).

A. I. Solzhenitsyn, The First Circle, Harper & Row, New York (1968).

P. N. Somerville, "Tables for Obtaining Nonpara netric Tolerance Limits," Ann. Math. Statist., Vol. 29, pp. 599-601 (1958).

L. V. Surgent and L. H. Yost, "Measurement of Speaker Recognition," paper presented at 61st Meeting of Acoustical Soc. Am., New York (1961).

J. W. Tukey, "Nonparametric Estimation: II. Statistically Equivalent Blocks and Tolerance Regions – The Continuous Case," Ann. Math. Statist., Vol. 18, pp. 529-539 (1947).

_____, "Nonparametric Estimation: III. Statistically Equivalent Blocks and Multivariate Tolerance Regions – The Discontinuous Case," Ann. Math. Statist., Vol. 19, pp. 30-39 (1948).

H. L. Van Trees, Detection, Estimation, and Modulation Theory, Part 1, John Wiley, New York (1968).

A. Wald, Sequential Analysis, John Wiley, New York (1947).

_____, Statistical Decision Functions, John Wiley, New York (1950).

_____, "An Extension of Wilks' Method for Setting Tolerance Limits," Ann. Math. Statist., Vol. 14, pp. 45-55 (1943).

M. D. Waltz and K. S. Fu., "A Heuristic Approach to Reinforcement Learning Control Systems," IEEE Trans. on Automatic Control, Vol. AC-10 (1965).

G. S. Watson and M. R. Leadbetter, "On the Estimation of the Probability Density," Ann. Math. Statist., Vol. 34, pp. 480-491 (1963).

P. Whittle, "On the Smoothing of Probability Density Functions," J. Royal Statist. Soc., B, Vol. 20, pp. 334-343 (1958).

B. Widrow and F. M. Hoff, "Adaptive Switching Circuits," 1960 Wescon Conv. Record, Part 4, pp. 94-140 (1960).

B. Widrow, "Generalization and Information Storage in Networks of Adaline 'Neurons'," in Yovits, Jacobi and Goldstein Self-organizing Systems 1962 , pp. 435-461, Spartan, Washington, D.C., (1962).

S. S. Wilks, "Determination of Sample Sizes for Setting Tolerance Limits," Ann. Math. Statist., Vol. 12, pp. 91-96 (1941).

_____, "Statistical Prediction with Special Reference to the Problem of Tolerance Limits," Ann. Math. Statist., Vol. 13, p. 400 (1942).

_____, "Order Statistics," Bull. Amer. Math. Soc., Vol. 54, p. 6 (1948).

_____, Mathematical Statistics, John Wiley, New York (1962).

J. J. Wolf, "Acoustic Measurements for Speaker Recognition," paper presented at the 77th meeting of the Acoustical Soc. of Am., Vol. 46, p.90 (1969).

A. C. Wolff, "The Estimation of the Optimum Linear Decision Function with a Sequential Random Method," IEEE Trans. on Info. Theory, Vol. IT-12, No. 3, pp. 312-315 (1966).

M. A. Young and R. A. Cambell, "Effects of Content on Talker Identification," J. Acoust. Soc. Am., Vol. 42, No. 6, pp. 1250-1254 (1967).



PASSIVE DETECTION AND TRACKING USING SURFACE SCATTERED SIGNALS

by

John F. McDonald

Progress Report No. 43

General Dynamic/Electric Boat Research

April 1970

DEPARTMENT OF ENGINEERING
AND APPLIED SCIENCE

YALE UNIVERSITY

ABSTRACT

Signals reflected from irregular time varying boundaries such as the sea surface undergo distortion which limits their detectability and useability for tracking. The properties of this distortion for correlator processing are herein related to the statistical constraints placed upon the time variation and irregularity of the boundary. Two propagation geometries are analysed. The first deals with the cross-correlation of surface reflected and direct transmission paths, and the second with the cross-correlation of surface scattered signals received at two different locations. This analysis assumes that the signal generated at the target and the background noise are both gaussian random variables. Three models of the scattering mechanism are proposed and two are analysed in detail. In all cases the correlator output is shown to exhibit very persistent fluctuations due to the scattering. The existence of these fluctuations is related to the non-gaussian nature of the scattered signals. The fourth order cumulant is computed to show that well spaced scattered signal samples may be dependent even when they are uncorrelated. Results are presented for Low pass signal spectra and are investigated as a function of bandwidth. When the receiver is constrained to be steered "on target" only the signal energy that is coherent between various paths contains information useful for detection or tracking. However, when the receiver is not so constrained, signal scattering of a delay modulated nature is shown to be useful for detection.

Table of Contents

ABSTRACT iii

CHAPTER I, INTRODUCTION

1.0 Preliminary Remarks B-1

1.1 Description of the General Problem and Preview of the Results B-7

1.2 A Brief Historical Summary of Surface Scattering and the Motivation for the Current Research B-12

1.3 Scattering from Time Varying Surfaces, The Extended Kirchhoff Integral Equation B-20

1.4 The Approximate First Order Solution of the Kirchhoff Integral Equation B-27

CHAPTER II, PRELIMINARY DISCUSSION OF SIGNAL AND SYSTEM PROPERTIES

2.0 Introduction B-35

2.1 Concerning the Nature of the Target Signal B-35

2.2 System Function Description for Linear Time Varying Filters B-42

2.3 Cascaded Linear Time Varying Systems B-51

2.4 Stationary Stochastic System Correlations B-54

2.5 Time Invariant and Slowly Varying Systems B-66

CHAPTER III, STATISTICS OF THE FINITE-TIME CORRELATOR

3.0 Introduction B-68

3.1 Cross-Correlation of Direct and Surface Reflected Paths B-69

3.2 Second Order Statistics for the Output of the Correlator B-72

3.3 Correlator Fluctuation for the Direct and Surface Reflected Multipath Processor B-76

3.4 Two Receiver Array Cross-Correlator Processing B-81

3.5 Cross-Correlator Fluctuation for the Two Receiver Array B-84

3.6 Correlator Tracking Error B-92

3.7 Concerning the Departure of the Statistics of the Scattered Signals from Gaussian. B-95

3.8 The Two Sided Likelihood Decision Scheme for the Correlator Detector B-99

CHAPTER IV

4.0 Introduction B-104

4.1 First and Second Order Statistics for the Random Amplitude and Delay Model B-105

4.2 Multipath Correlator Fluctuations for the Random Amplitude and Delay Model B-109

4.3 Second and Fourth Order Cross System Statistics for the Random Amplitude and Delay Model B-125

4.4 Array Correlator Fluctuations for the Two Channel Random Amplitude and Delay Model B-129

CHAPTER V, THE RANDOMIZED SINUSOIDAL SURFACE MODEL

5.0 Introduction B-137

5.1 Gulin's Solution for $H(\omega,t)$ and the Associated Impulse Response B-138

5.2 First and Second Order Statistics of the Sinusoidal Boundary Model B-143

5.3 Multipath Correlator Fluctuations for the Random Sinusoidal Boundary B-154

5.4 Second and Fourth Order Cross System Statistics for the Two Channel Random Sinusoidal Boundary B-162

CHAPTER VI, POSSIBLE EXTENSIONS OF THE PRESENT WORK AND SUGGESTIONS FOR FUTURE RESEARCH

6.0 Extensions to the Fully Random Boundary B-173

6.1 Other suggested Topics for Future Research B-183

APPENDIX A, MORGAN'S DERIVATION OF EQUATION (1.3-14). B-184

APPENDIX B, PLANE WAVE EXPANSION FOR SURFACE SCATTERING B-186

APPENDIX C, SPECULAR POINT EXPANSIONS FOR r and r' B-189

APPENDIX D, REFINEMENTS IN THE CRITERIA FRO DETECTABILITY B-196

APPENDIX E, VARIANCE INTEGRALS FOR THE RANDOM AMPLITUDE AND DELAY MODEL B-201

APPENDIX G, TABULATION OF THE FIRST FEW $G_{n,\ell,m}(p,q,\tau)$ B-209

APPENDIX H, VARIANCE INTEGRALS FOR THE RANDOM SINUSOIDAL BOUNDARY MODEL B-212

APPENDIX I, NUMERICAL EVALUATION OF THE $G_{13}^{31}(z)$ B-215

APPENDIX J, EXTENSIONS TO LARGE ARRAYS B-222

CHAPTER I

INTRODUCTION

1.0 Preliminary Remarks

Two important techniques that have been used to improve the performance of Sonar detection and communication systems are multipath and space diversity signal processing. Multipath signal processing capitalizes on the signal replication or echoing that characterizes propagation from source to receiver along many paths. Similarly, space diversity processing takes advantage of the signal replication which occurs at an array of spatially separated receivers when transmission is from a common source.

Multipath Sonar processing is generally difficult to implement in practice since it usually requires a detailed knowledge of the propagation geometry. When this information is not available or when it is difficult to estimate, multipath effects are more often regarded as a hindrance than as an aid. However, multipath propagation has been studied with great interest for range and depth estimation in tracking. Furthermore, in certain receivers concerned with the detection of signals of unknown spectrum, multipath replication is the sole distinguishing feature which can be used to discriminate between targets and noise.[1]

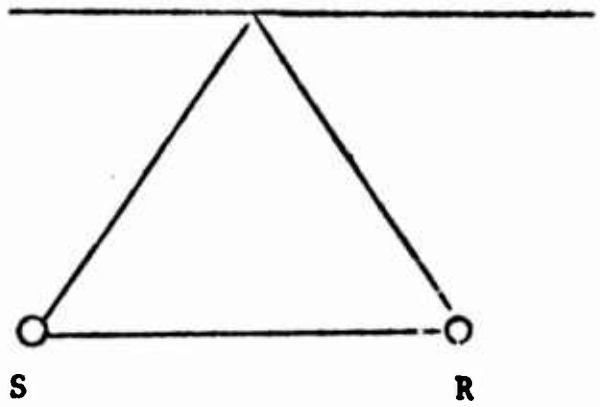
Space diversity processing on the other hand is more commonly exploited and more thoroughly understood. Here the design of signal processors is not as dependent on knowledge of specific propagation geometry and it is therefore less sensitive, more flexible, and easier to implement. Moreover, such signal processing permits discrimination against non-directional background noise. This noise rejection can frequently be improved by simply increasing the number of receiving

elements.[2]

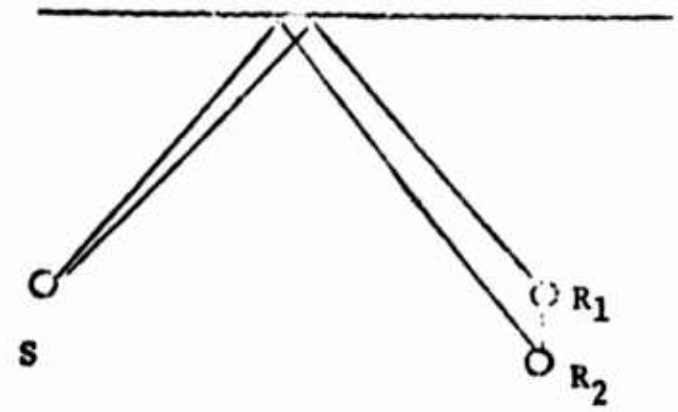
Clearly, propagation geometries exist which include both multipath and space diversity replication. Figure 1.0-1 illustrates some typical propagation geometries. Case (c) falls into this mixed space and multipath category. The receivers in these examples are assumed to be single-site sensors located at various points in space. Each sensor is assumed to be directional enough to select either by design or by accident certain ray paths for reception. Case (b) is an example which illustrates two single-site receivers which suppress the direct paths of transmission.

When the reflections from the boundary do not change the incident signal (except possibly for sign) then the replication is termed perfect. This occurs when sound reflects from a completely smooth air-water interface when the sound is incident from within the water medium. In this case the reflection is locally characterized by a pressure-release boundary condition.[3]

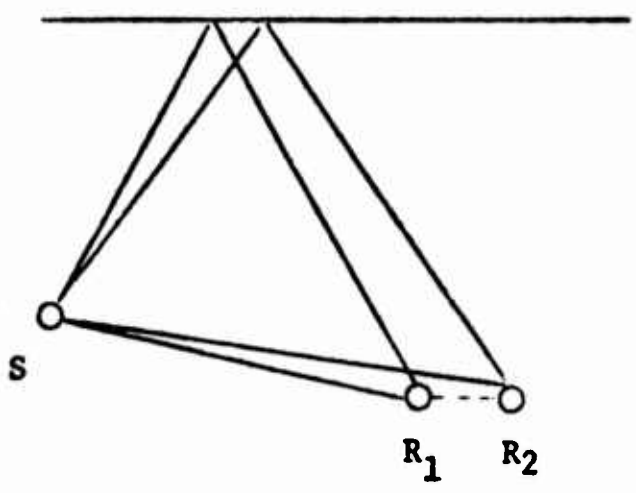
However, when the boundary is deformed spatially in some random fashion then the signal replication becomes distorted. The spatial deformations produce frequency dependent interference effects. At certain frequencies and locations these interference effects superpose constructively to enhance the strength of transmission. At other frequencies or locations, however, the interference is found to be destructive producing poor transmission.



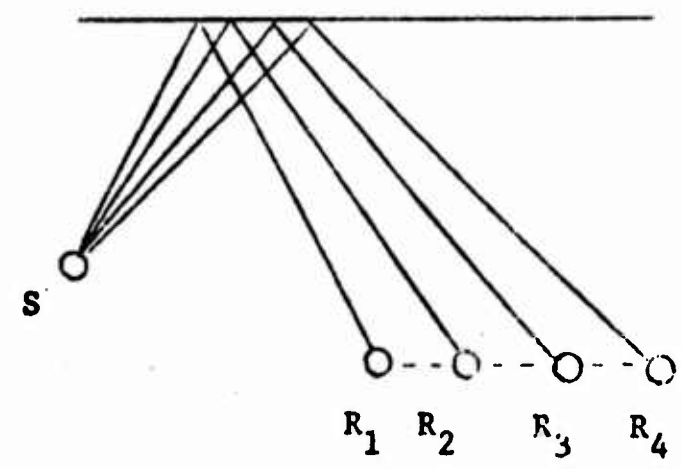
(a) Single Multipath Receiver Using Direct and Surface Reflected Paths



(b) Two Element Vertical Array Using Only Surface Reflected Paths



(c) Two Element Horizontal Array Using Direct and Surface Reflected Paths



(d) Multi-receiver Array Using Only Surface Paths

Figure 1.0-1 Typical Propagation Geometries Involving Surface Scattering

Constructive interference occurs at any given frequency when transmissions over various paths reflecting from randomly oriented facets of the irregular boundary arrive at the receiver in phase with each other. For far-field reception this results from spatial periodicities in the boundary deformations which give rise to effective path length differences which are multiples of the wavelength λ of the radiation.[4] Similarly, the degree of power loss during destructive interference is determined by the likelihood that the boundary facets position and align themselves in such a way as to consistently divert energy away from the receiver. From the point of view of multiple transmission paths, effective path length differences produce signal cancellation. The properties of the scattered radiation are thus seen to be linked to the statistical properties of the boundary deformations. In particular, the two dimensional space spectrum describing the harmonic content of the surface irregularities plays an important role.[5]

In connection with this purely spatial redistribution of reflected energy, the interference due to surface deformations introduces frequency selective transmission properties. These produce spectral alterations in the scattered signal which invariably cause correlation degradation and consequently poorer signal detectability. In addition, the spatial extent of the active scattering area produces a general spreading or smearing of signal correlation over arrival time.[6]

The irregular surfaces considered here are assumed to be instantaneous realizations of a stochastic ensemble of such surfaces. Each surface ensemble member generally produces a different scattered signal. If the boundary deformations change in shape as a function of time then the scattered signal exhibits fluctuations due to this

motion. For severe scattering this fluctuation can turn into deep fading. When the time variations are very slow, signal processors operating on the scattered signals are occasionally confronted with low signal to noise ratios for long periods of time.

If the time variations of the surface is assumed to satisfy the two dimensional wave equation then the periodic or harmonic components of the surface propagate at constant velocities in various directions. If the incident radiation is monochromatic these surface motions generate side bands near the frequency of the incident radiation which are easily observed.[7] In this case a discrete harmonic space component of the surface with temporal period of $2\pi/\Omega$ which produces constructive interference due to transmission path length differences of $n\lambda$ generates signal sidebands at

$$\omega = \left[\frac{2\pi c}{\lambda}\right] \pm n\Omega \qquad (1.0\text{-}1)$$

where c is the signal propagation velocity. Similarly, surfaces with diffuse space components produce a general smearing of the signal spectrum.

In agreement with the generally accepted definition,[8] the constructive superposition at a certain location and frequency due to transmission with path length differences at $n\lambda$ is termed a spectral order. We call n the order of interference. It is clear that the number and strength of orders that are "seen" by the receiver depend on the directivity and orientation of the sensors. With highly directive sensors it is possible to observe single orders.[9]

At higher acoustic frequencies the number and spatial density of these orders increases. Consequently, the degree of frequency smear or shifting which can occur in accordance with Equation (1.0-1) becomes larger. Furthermore, the high frequency components of the incident signal which constitute the fine structure of the signal correlation are especially susceptible to temporal smearing. Hence, it is desirable to have a simple criterion by which one can judge whether significant time or frequency smear occurs during reflection. The most universally accepted criterion is that of Rayleigh.[10] It states that if the angle of grazing for incident radiation is ψ_i and if the maximum height of the surface irregularities is h then the surface appears smooth if

$$kh \, Sin(\psi_i) << 1 \qquad (1.0\text{-}2)$$

where $k = 2\pi/\lambda$. Although many theoretical efforts have been made to refine this crude rule-of-thumb it remains one of the most satisfactory criteria available as will be shown in chapter 5.

1.1 Description of the General Problem and Preview of the Results

This report deals with the passive sonar detection of targets generating zero mean Gaussian noise-like signals. The detector is assumed to use echos or replicas of the signals which are distorted by scattering from an irregular time-varying pressure release surface. The target is assumed to radiate in an omni-directional manner into a uniform, isovelocity medium which is characterized by rectilinear or straight line ray propagation.[11] The radiated signals are presumed to travel along a small number of paths to one or possibly two receivers with some of these paths reflecting from the surface. The received signals are also assumed to be corrupted by broad-band, additive Gaussian background noise.

Attention is primarily focused on evaluating the effect of the slowly varying irregular surface on the performance of correlator detectors and trackers. Scatter degradation is computed as a function of two principle design parameters:

1. Correlator integration time.
2. Signal processing bandwidth.

The technique used to evaluate performance is to compute the mean and variance of the correlator output under the two hypotheses of target present and target absent. The general nature of the propagation geometry is considered known.

It is shown that slow time variations in the irregularities of the surface manifest themselves as persistent fluctuations in the output of correlators operating on the scattered signals. These fluctuations are

noticeable until the correlator integration interval becomes much larger than the relaxation times that describe the temporal behavior of the surface. Only at this point does the correlator reliably "average out" these variations over the ensemble of possible surface deformations. Similar fluctuations are also exhibited in estimates of such parameters as target bearing, range or depth which are derived from the correlator output.

The most distinguishing feature of the treatment presented in this report is the explicit inclusion of the time-variation of scattering. A comprehensive description of the scatter dynamics or relaxation mechanisms becomes necessary when the scattering is strong enough to force long processing intervals in persuit of reliability. A distinction is drawn between long and short term correlator fluctuations.

Of course, when the input signal to noise ratio is high and the scattering weak, short integration times may be used. The correlator fluctuations are still apparent, but the physics of the scattering may now be considered "frozen" for the duration of the processing interval. The time variation of the scattering then becomes important only onsofar as it describes the average lengths of time during which the signal correlation "fades" and remains uniformly weak, the smear due to drift being negligible. The "frozen" model is, however, just a special case of the dynamic or time-varying model.

With regard to the signal processing bandwidth it has been noted that higher frequency signals tend to be more severely decorrelated by scattering. Furthermore, signal to noise ratio tends to decrease

with increasing frequency. Therefore, it is desirable to be able to set the working bandwidth at a value which allows enough usable energy to be processed while rejecting excess background noise and unusable, badly decorrelated signal. This value generally depends on the statistical roughness of the surface irregularities.

The exact definition of the optimal processing bandwidth is more complex for array processing. It is interrelated with the question of the placement of sensors and interpath coherence. For if two sensors are placed in close proximity, the scattered energy received by them tends to come from nearly identical or overlapping scatter facets. Even heavily scattered signals can contribute to sensor cross-correlation under such circumstances.

On the other hand, it is a well known fact that background noise cross-correlation at widely separated sensors tends to be small. This effect is often used to obtain increased array gain while maintaining simplicity of processor design. It is important to remember, however, that scattered signal correlation also drops off with increasing separation. For very large separations the sensors receive signals scattered by completely independently positioned and oriented facets.

Three basic propagation geometries are investigated:

1. Single-site reception of direct path and surface reflected path (Figure 1.0-1a). In this case the signals transmitted over the two paths are assumed to be separable through use of directional sensors.

2. Two site reception of surface reflected paths (Figure 1.0-1b). The direct paths are assumed to be suppressed in order to prevent the analysis From becoming unnecessarily complicated.

3. N-site reception of surface-reflected path (Figure 1.0-1d). Again the direct paths are suppressed.

The first case is of interest in range and depth estimation while case (2) & (3) is important in bearing estimation. The third case is examined only in a very brief manner in Appendix J.

Three different models are examined for surface scattering:

1. Random time-variable delay and amplitude modulation. This model is not particularly realistic, but it is suitable for determining the effect of the time-variation of the scattering. (Chapter 4)

2. Randomized one-dimensional surface corrugation scattering. This model is slightly more realistic than model (1) while remaining reasonably tractable. This surface is perfectly correlated along the direction parallel to the corrugation. Randomization of the model is handled by stochastic parameters. (Chapter 5)

3. Two dimensionally irregular surface scattering. In this case the boundary deformations are described statistically by a two dimensionally stationary spatial correlation function. (Chapter 6)

Although the simpler models (1) and (2) are not entirely satisfactory representatives of real physical scattering, the results derived for them do share certain overall similarities the fully stochastic model of (3). Moreover, it is possible to obtain some results in closed form for the simpler models which are at the time of this writing unattainable for the more realistic case. Model (3) is discussed only briefly .

The analyses of models (2) and (3) are approximate. The approximations used, however, enjoy a fair degree of popular acceptance in the literature. To some extent there is experimental justification for this optimism,[12] but the agreement rapidly becomes only qualitative when the restrictions on the approximations are

violated. Formally, the results of this report apply only for weak to moderate scattering observed in the far-field and at angles of grazing large enough so that over-shadowing and multiple scattering do not occur. This also implies that most of the scattering comes from the vicinity of the surface near the specular point. It is also assumed that the surface displacements and slopes are small and the radii of curvature are large compared to λ. This in turn implies that the surface roughness is very slight and that the incident radiation is of low to moderate frequency.

Although all of the applications actually analysed in chapters 4 and 5 only low-pass filtering is examined, chapter 3 is sufficiently general to handle narrowband problems. Applications of the formulas in chapter 3 to the very narrowbandwidth case must, however, be prefaced by the comments in chapter 2. In this limit frequency smearing must be properly taken into account.

Additional material is presented in chapter 2 which is important in applying the general results of chapter 3 to multiple bounce scatter propagation. In this case the frequency spreading due to a given reflection interacts with that of successive reflections.

1.2 A Brief Historical Summary of Surface Scattering and the Motivation for the Current Research

It is not the purpose of this report to present a comprehensive summary of the development of the theory of scattering from irregular surfaces. Beckmann and Spizzichino[13] have provided a very complete bibliography covering publications up to 1963. The article by Lysanov[14] compiles a listing of the Russian effort up to 1958. Many of the more recent articles on scattering are refinements of these earlier treatments. We shall confine our attention here to a discussion of some of the simpler concepts which are building blocks leading to the point of view developed in this report.

The first attempts to analyze rough surface scattering were preceded by a long period of study of diffraction and interference effects by Fresnel[15] and Fraunhofer.[16] The analysis of the Fraunhofer parallel slit problem demonstrated the relation between spatial periodic components describing apertures in a screen and the position of interference maxima for waves passing through the apertures.

The analysis by Payleigh[17] for the reflection of sound from a diffraction grating appears to be the earliest attack on the problem of scattering from surface deformations. This analysis assumes an incident radiation which is in the form of a monochromatic plane wave. The grating is assumed to be in the form of a periodic corrugation ζ with, for example, the direction of the corrugation along the x-axis:

$$\zeta(x) = \zeta(x + \Lambda) \qquad (1.2\text{-}1)$$

where Λ is the spatial period of the corrugation. The incoming

radiation is assumed incident along the x-axis at an angle ψ_i . The receiver is placed in the infinitely removed far-field observing the reradiation along the x-axis at angle ψ_r . The path length difference for two rays connecting the source and receiver which impinge on the corrugation at similar points separated by Λ (see Figure 1.2-1) is given by

$$BC - AD = \Lambda (\text{Cos } \psi_r - \text{Cos } \psi_i) \qquad (1.2\text{-}2)$$

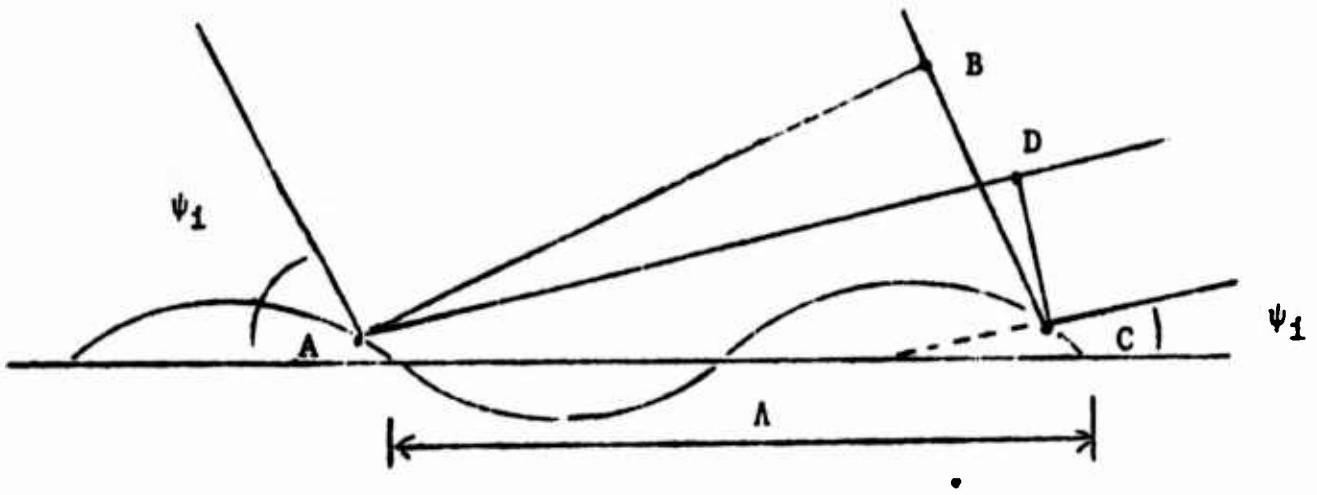


Figure 1.2-1 The Diffraction Grating Problem

For an interference maximum of reradiation to occur at an angle of ψ_{rm} this path length difference should be a multiple of the wavelength λ so that

$$\text{Cos } \psi_{rm} = \text{Cos } \psi_i + m\lambda/\Lambda \qquad (1.2\text{-}3)$$

This relation is called the grating equation and the angles ψ_{rm} (for $m = 0, \pm 1, \pm 2, \ldots$) are termed Bragg angles.[18] From the requirement that $|\text{Cos}(\psi_{rm})| \leq 1$ it is seen that propagating orders exist in the reradiating field for values of m satisfying the inequalities

$$-\frac{\Lambda}{\lambda} (1 + \text{Cos } \psi_i) \leq m \leq \frac{\Lambda}{\lambda} (1 - \text{Cos } \psi_i) \qquad (1.2\text{-}4)$$

For $\Lambda/\lambda << 1$ only the 0-th order propagates. This order is termed the specular order or mode. For high frequencies (small values of λ) many propagating orders are possible. Orders for values of m not satisfying Equation (1.2-4) are said to be in cutoff and they correspond to surface waves that do not carry energy away from the boundary.

Formal extension of path length difference analysis for interference to two dimensionally periodic surfaces presents no difficulty.[20]

An analysis of the scattering of sound from two dimensionally random surfaces is presented by Eckart.[21] This treatment clearly demonstrates the connection between the interference maxima and the surface spectrum. His analysis begins with the well known Kirchhoff integral theorem for monochromatic reradiation:

$$\hat{p}_s(P) = \frac{1}{4\pi}\iint_S \{\frac{\partial \hat{p}_s}{\partial n}\frac{e^{jkr}}{r} - \hat{p}_s \frac{\partial}{\partial n}(\frac{e^{jkr}}{r})\}\, dS \qquad (a) \qquad (1.2\text{-}5)$$

where

$$p_s(x,y,z,t) = \hat{p}_s(x,y,z)\, e^{-j\omega t} \qquad (b)$$

Here $\hat{p}_s$ is the reradiation amplitude, k is the wave number $2\pi/\lambda$, S is the scattering surface, r is the distance from the receiver point P to a point on the surface, and $\partial/\partial n$ is the derivative along the outward normal of the surface S.

Eckart assumes that the surface S is illuminated by an incident wave generated by a directional monochromatic point source with a finite beam-width:

$$\hat{p}_i(x,y,z) = B_e(\theta,\phi)\frac{e^{jkr'}}{r'} \qquad (1.2\text{-}6)$$

Here $B_e(\theta,\phi)$ is the beam function, r' is the distance from the source point Q, and θ and ϕ are the polar and azimuthal angles for a spherical coordinate system centered at Q.

The boundary condition on the surface S is assumed to be a pressure release constraint:

$$\hat{p}_i + \hat{p}_s = 0 \quad [\text{on } S] \tag{1.2-7}$$

After making a series of approximations which are explored in greater depth in Section 1.4, Eckart arrives at an approximate solution to Equation (1.2-5) subject to Equations (1.2-6) and (1.2-7) given by

$$\hat{p}_s(P) \approx \frac{1}{4\pi} \iint_S \frac{B_e(\theta,\phi)}{rr''} \frac{\partial}{\partial z} \left[e^{jk(r+r')}\right] dx\, dy \tag{1.2-8}$$

Eckart assumed the beam pattern to be narrow enough to permit linearization of the path length $r+r'$ as follows:

$$r + r' \approx r_o + r_o' + (a_s^o x + b_s^o y + c_s^o \zeta) \tag{1.2-9}$$

where r_o and r_o' are the values of r and r' near the center of the illuminated area which is considered the origin of the x,y,z coordinate system and a_s^o, b_s^o, c_s^o are respectively the sums of the x,y,z direction cosines of r_o and r_o'. On executing the partial derivative with respect to z in Equation (1.2-8) and replacing the reciprocal of rr' by its value near the center of the illuminated area we have

$$\hat{p}_s(P) \approx \frac{(jkc_s^o)}{4\pi} \frac{e^{jk(r_o+r_o')}}{r_o r_o'} \iint B_e(\theta,\phi)\, e^{jk(a_s^o x + b_s^o y + c_s^o \zeta)} \, dx\, dy \tag{1.2-10}$$

where ζ, θ, and ϕ are all functions of x and y. The result

shows that the scattered signal is approximately composed of a superposition of replicas of the incident wave which are randomly delayed, attenuated and time differentiated (as may be seen from the factor jk in Equation 1.2-10). The time derivative should not be surprising since the result is in the form of a density over arrival times.[22]

Finally, by forming the mean square magnitude for $\hat{p}_s(P)$ Eckart obtained the desired relationship between the reradiation and the statistical properties of the surface. Assuming the surface perturbation $\zeta(x,y)$ to be a stationary random variable[23] in x and y we have from Equation (1.2-10)

$$\overline{p_s(P)\ p_s(P)^*} = \left[\frac{kc_s^o}{4\pi r_o'}\right]^2 \iint J(\xi,\eta)\ Q_2(-kc_s^o, kc_s^o, \xi, \eta) \tag{1.2-11}$$
$$e^{-jk(a_s^o\xi + b_s^o\eta)}\ d\xi\ d\eta$$

Here the asterisk denotes the complex conjugate and the over-bar indicates ensemble averaging. The function $Q_2(u,v,\xi,\eta)$ is the second order characteristic function corresponding to the two sample probability distribution $F[\zeta(x,y), \zeta(x+\xi, y+\eta)]$:

$$Q_2(u,v,\xi,\eta) = \overline{\exp(-j[u\ \zeta(x,y) + v\ \zeta(x+\xi, y+\eta)])} \tag{1.2-12}$$
$$= \iint_{R_{xy}} e^{-j[u\zeta(x,y) + v\zeta(x+\xi, y+\eta)]}\ dF[\zeta(x,y), \zeta(x+\xi, y+\eta)]$$

The function $J(\xi,\eta)$ is the convolution of the beam patterns projected onto the surface:

$$J(\xi,\eta) = \frac{1}{(r_o)}^2 \iint B_e[\theta(x,y),\phi(x,y)]$$

$$B_e[\theta(x+\xi,y+\eta),\phi(x+\xi,y+\eta)]\ dx\ dy \qquad (1.2\text{-}13)$$

The region R_{xy} in Equations (1.2-12) and (1.2-13) is the infinite x-y plane.

Eckart assumes ζ to be a Gaussian random process yielding

$$Q_2(u,v,\xi,\eta) = e^{-\frac{1}{2}[u^2h^2+2uv\Psi_\zeta(\xi,\eta)+v^2h^2]} \qquad (1.2\text{-}14)$$

where h is the mean square value of the (zero mean) deformation ζ and

$$\Psi_\zeta(\xi,\eta) = \overline{\zeta(x,y)\ \zeta(x+\xi,y+\eta)} \qquad (1.2\text{-}15)$$

is the correlation function for the deformation at two points on the surface separated by displacements ξ and η. The dependence of Ψ only on these displacements is a result of the assumption of spatial stationarity of ζ.

By substituting Equation (1.2-14) into (1.2-11) Eckart obtains

$$\overline{\hat{p}_s(P)\ \hat{p}_s(P)^*} = \left[\frac{kc_s^o}{4\pi r_o'}\right]^2 \exp[-k^2h^2c_s^{o2}] \iint_{R_{xy}} J(\xi,\eta)$$

$$\exp[k^2c_s^{o2}\ \psi_\zeta(\xi,\eta)]\ e^{-jk(a_s^o\xi+b_s^o\eta)}\ d\xi\ d\eta \qquad (1.2\text{-}16)$$

Ignoring the beam pattern convolution function $J(\xi,\eta)$ for the moment, it is seen that the directional and frequency selective properties are determined on the average by the space transform of $\exp[k^2c_s^{o2}\ \Psi_\zeta(\xi,\eta)]$.

In a similar manner Eckart obtains the mean of $p_s(P)$

$$\overline{\hat{p}_s(P)} = Q_1(kc_s^o)\left\{\frac{(jkc_s^o)}{4\pi}\frac{e^{jk(r_o+r_o')}}{r_o\, r_o'}\iint_{R_{xy}} B_e(\theta,\phi)\right. \qquad (1.2\text{-}17)$$

$$\left. e^{jk(a_s^o x+b_s^o y)}\, dx\, dy\right\}$$

where the quantity in braces is identical to the mirror reflection term $\hat{p}_s(P)$ for $\zeta = 0$ (as may be seen from Equation (1.2-10) and $Q_1(u)$ is the one dimensional characteristic function for ζ :

$$Q_1(u) = \overline{\exp(ju\zeta)} \qquad (1.2\text{-}18)$$

Again, by assuming Gaussian statistics for the variable ζ Eckart finds

$$\overline{\hat{p}_s(P)} = [\hat{p}_s(P)]_{\zeta=0}\; e^{-k^2h^2c_s^{o2}/2} \qquad (1.2\text{-}19)$$

The mean value $\overline{\hat{p}_s(P)}$ is termed the coherent reradiation by Eckart. While Equation (1.2-10) relies heavily on certain approximations, there is good experimental evidence that Equation (1.2-19) is very accurate when the restrictions on the approximations are not violated.[24]

In particular, assuming the incident beam to have an angle of grazing ψ_i and examining the forward coherent reradiation into the specular direction defined by

$$a_s^o = 0 \qquad \text{(a)}$$

$$b_s^o = 0 \qquad \text{(b)} \qquad (1.2\text{-}20)$$

$$c_s^o = -2 \sin \psi_i \qquad \text{(c)}$$

we have from Equation (1.2-19) the coherent reflection coefficient R :

$$R = \frac{\overline{[\hat{p}_s(P)]}}{[\hat{p}_s(P)]_{\zeta=0}} = \exp(-2\, k^2 h^2 \mathrm{Sin}^2 \psi_1) \qquad (1.2\text{-}21)$$

It can be seen that this is a function only of the Rayleigh parameter given in Equation (1.0-2).

Unfortunately, if we choose to violate restrictions used in obtaining these results this agreement breaks down. For example, if we wish to examine a passive detection problem, a narrow beam pattern for the source at the target is unlikely. If we set $B(\theta,\phi) = 1$ and consider reradiation from a flat surface ($\zeta = 0$) then in the specular direction defined by Equations (1.2-20), (1,2-10) yields

$$\hat{p}_s(P) \to \infty \; ! \qquad (1.2\text{-}22)$$

The difficulty can be traced to the linearization in Equation (1.2-9). This problem is examined further in Section 1.4.

1.3 Scattering from Time Varying Surfaces

The Extended Kirchhoff Integral Equation

The Kirchhoff formula, Equation (1.2-5), which forms the basis of the Eckart analysis applies strictly only to the case of scattering from time invariant surfaces. Eckart tacitly assumes that if the surface motion is slow compared to the period λ/c of the radiation then Equation (1.2-5) still applies but with the boundary

$$[S(t):\ z = \zeta(x,y,t)] \tag{1.3-1}$$

which is explicitly time varying. Eckart's analysis of the mean and mean square of $p_s(P)$ did not require an explicit description of the time varying behavior. However, we are primarily interested in examining space-time correlation functions and the temporal fluctuations in their estimates. Therefore, a few qualifying remarks are warranted with regard to this approximation.

The Kirchhoff formula is essentially a restatement of the well known wave equation for small pressure disturbance in isotropic elastic media

$$\nabla^2 p = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \tag{1.3-2}$$

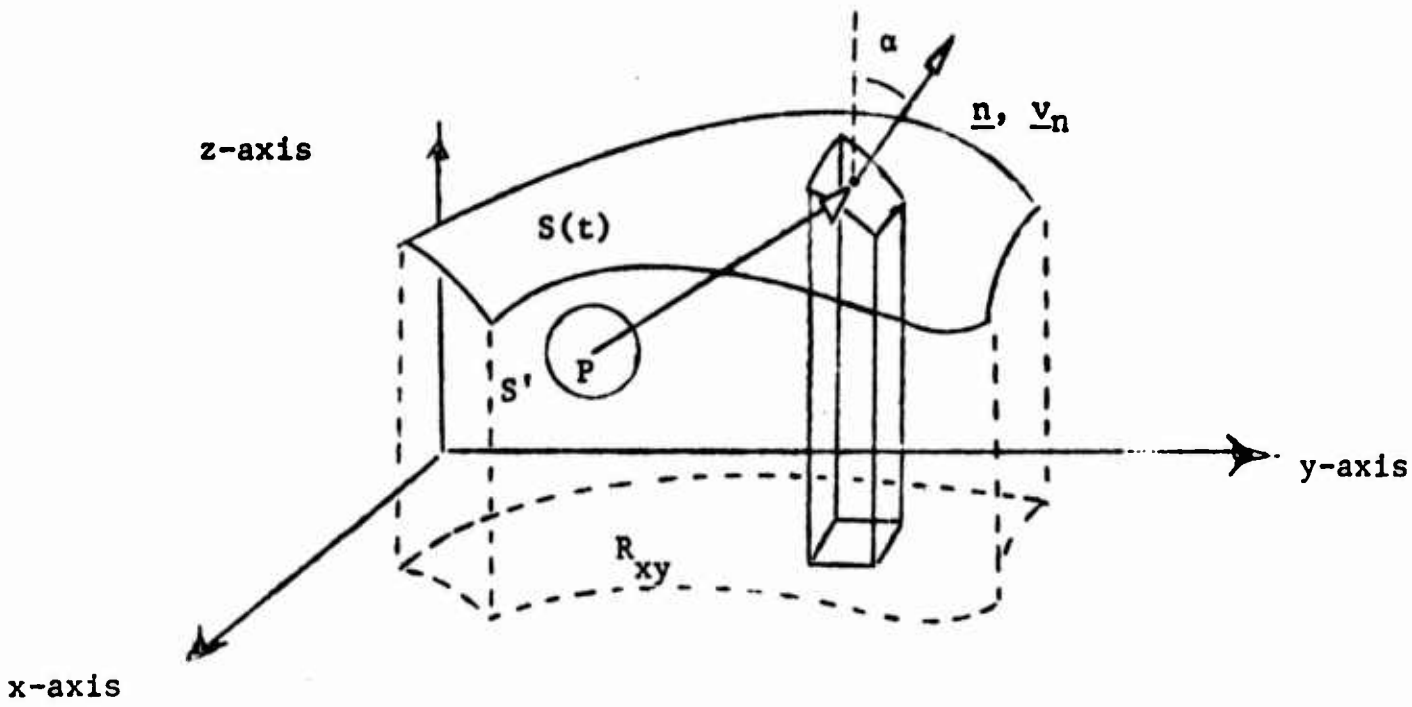
in terms of an integral equation in which the integral is over the surface $S(t)$. The extension of the basic Kirchhoff formula, Equation (1.2-5), to the case of time varying boundaries was carried out by W. R. Morgans in 1929.[25] We now briefly review his solution.

Mathematical restrictions placed on the pressure p inside the volume V enclosed by the surface $S(t)$ (see Figure 1.3-1) which are necessary for the extended Kirchhoff formula to hold are equivalent to

the requirement that V be source-free. Consequently, the solution to Equation (1.3-2) is broken into two parts:

$$\boxed{p(t) = p_i(t) + p_s(t)} \qquad (1.3\text{-}3)$$

where both $p_i(t)$ and $p_s(t)$ are solutions to Equation (1.3-2). The solution $p_i(t)$ is that due to the source of the radiation but in an unbounded medium, while $p_s(t)$ is the solution which must be added to $p_i(t)$ in order to satisfy the boundary condition on $S(t)$. It is assumed that the incident radiating field $p_i(t)$ is known so that there remains only the problem of obtaining the integral equation for $p_s(t)$.



Geometry for Time Variable Surface

in the Kirchhoff Formula

Figure 1.3-1

The Kirchhoff integral equation for time varying surfaces is obtained in a manner similar to that used in the time invariant case.[26,27,28] One begins by considering the wave equations for the scattered wave p_s and an auxiliary solution q which is used in the derivation:

$$\text{(a)} \quad \nabla^2 p_s = \frac{1}{c^2} \frac{\partial^2 p_s}{\partial t^2} \qquad \text{(b)} \quad \nabla^2 q = \frac{1}{c^2} \frac{\partial^2 q}{\partial t^2} \tag{1.3-4}$$

Multiplying Equation (1.3-4b) by p_s, Equation (1.3-4a) by q, and then taking the difference we have

$$q\,\nabla^2 p_s - p_s\,\nabla^2 q = [\nabla\cdot(q\,\nabla p_s) - \nabla p_s\cdot\nabla q] - [\nabla\cdot(p_s\,\nabla q) - \nabla q\cdot\nabla p_s]$$

$$= \nabla\cdot[q\,\nabla p_s - p_s \nabla q] \tag{1.3-5}$$

$$= q\,\frac{\partial^2 p_s}{\partial t^2} - p_s\,\frac{\partial^2 q}{\partial t^2} = \frac{\partial}{\partial t}\left\{q\,\frac{\partial p_s}{\partial t} - p_s\,\frac{\partial q}{\partial t}\right\}$$

We now consider the point P and the time t_o at which we ultimately desire the value of p_s. Following Eckart's notation we let r denote the distance from P. We surround P with a small sphere S' (see Figure 1.3-1) and designate the volume below $S(t)$ and outside S' as V'. Integrating Equation (1.3-5) over V' and using Green's Theorem

$$\frac{1}{c^2} \iiint_{V'} \frac{\partial}{\partial t}\left(q\,\frac{\partial p_s}{\partial t} - p_s\,\frac{\partial q}{\partial t}\right) dV = \tag{1.3-6}$$

$$\iint_{S'} \left(q\,\frac{\partial p_s}{\partial n} - p_s\,\frac{\partial q}{\partial n}\right) dS + \iint_{S(t)} \left(q\,\frac{\partial p_s}{\partial n} - p_s\,\frac{\partial q}{\partial n}\right) dS$$

Next, the derivation of the Kirchhoff formula requires the integration of Equation (1.3-6) with respect to t and it is at this point that the time variation of $S(t)$ must be taken into account. This may be done by using the integrated form of the Stoke's total derivative. For any function $f(x,y,z,t)$

$$\frac{\partial}{\partial t}\iiint_{V'} f\, dV = \iiint_{V'} \frac{\partial f}{\partial t}\, dV + \iint_{S(t)} f\, v_n\, dS \qquad (1.3\text{-}7)$$

where v_n is the velocity of an element of the surface, dS, taken along the outward normal, $\underline{n}$.

Applying Equation (1.3-7 to (1.3-6) and integrating t over $-\infty$ to $+\infty$ we have

$$\frac{1}{c^2}\iiint_{V'}\left(q\,\frac{\partial p_s}{\partial t} - p_s\,\frac{\partial q}{\partial t}\right) dV \Bigg|_{t=-\infty}^{t=+\infty} =$$

$$\int_{-\infty}^{+\infty} dt\,\left\{\frac{1}{c^2}\iint_{S(t)} v_n\left(q\,\frac{\partial p_s}{\partial t} - p_s\,\frac{\partial q}{\partial t}\right) dS + \iint_{S(t)}\left(q\,\frac{\partial p_s}{\partial n} - p_s\,\frac{\partial q}{\partial n}\right) dS \right. \qquad (1.3\text{-}8)$$

$$\left. + \iint_{S'}\left(q\,\frac{\partial p_s}{\partial n} - p_s\,\frac{\partial q}{\partial n}\, dS\right)\right\}$$

The volume integral in Equation (1.3-8) may be made zero by forcing q to be zero as $t \to \pm\infty$. On specializing q to be the function

$$q(r,t) = W(ct - ct_o + r)/r \qquad (1.3\text{-}9)$$

where $W(\xi)$ is a very narrow pulse centered around $\xi = 0$, and which is normalized

$$\int_{-\infty}^{+\infty} W(\xi)\, d\xi = 1 \qquad (1.3\text{-}10)$$

it becomes possible to evaluate the integral over S'. If this sphere is centered on P and shrunk in size then the integral over S' becomes

$$\frac{4\pi r^2}{c}(-1/r^2)\int_{-\infty}^{+\infty} W(ct-ct_o+r)\, p_s(r=0,t)\, c\, dt + o(r) \qquad (1.3\text{-}11)$$

$$+\frac{4\pi}{c} p_s(P,t_o)$$

Equation (1.3-8) then reduces to

$$p_s(P,t_o) = \frac{c}{4\pi}\iint_{R_{xy}} dx\, dy \qquad (1.3\text{-}12)$$

$$\{\int_{-\infty}^{+\infty} [q\,(\frac{\partial p_s}{\partial n} + \frac{v_n}{c^2}\frac{\partial p_s}{\partial t}) + p_s\,(\frac{\partial q}{\partial n} + \frac{v_n}{c^2}\frac{\partial q}{\partial t})]\, H(x,y,t)\, dt\}$$

where $H(x,y,t)$ is the secant of the angle to the z-axis of $\underline{n}(x,y,t)$:

$$\boxed{H(x,y,t) = \text{Sec}\,\alpha = \sqrt{1 + (\frac{\partial \zeta}{\partial x})^2 + (\frac{\partial \zeta}{\partial y})^2}} \qquad (1.3\text{-}13)$$

There remains now only the problem of performing the integral over t in Equation (1.3-12). The details of this manipulation are presented in Appendix A. The result is

$$p_s(P,t) = \iint_{R_{xy}} \frac{dx\, dy}{4\pi}\,[\frac{H}{r(1+\dot{r}/c)}\,\{\frac{\partial p_s}{\partial n} + \frac{v_n}{c^2}\frac{\partial p_s}{\partial t} + \frac{\left(\frac{\partial r}{\partial n} - \frac{\dot{r}v_n}{c^2}\right)}{(1+\dot{r}/c)}\frac{p_s}{r} \qquad (1.3\text{-}14)$$

$$+\frac{1}{Hc}\frac{d}{dt}[(\frac{\partial r}{\partial n} + \frac{v_n}{c})\,\frac{H\,p_s}{(1+\dot{r}/c)}]\}]_{t=t_o-\frac{r}{c},\ z=\zeta(x,y,t)}$$

The integrand is evaluated at the retardation time $t_o - r/c$ where r is the distance to the portion dS of the surface lying above $dx\,dy$ when it is positioned to give a disturbance arriving at P at t_o. The determination of r for a given x,y,t_o is itself a formidable problem.

Ignoring terms of the order v_n/c, $\dot{r}/c$, $\dot{v}_n/c$, $\dot{H}/c$ and $(\partial \dot{r}/\partial n)/c$ which are all quite small for the problems of interest in this report we see that Equation (1.3-14) becomes

$$p_s(P,t_o) = \iint_{R_{xy}} \frac{dz\,dy}{4\pi} \left[\frac{H}{r}\left\{\frac{\partial p_s}{\partial n} + \frac{\partial r}{\partial n}\frac{p_s}{r} + \frac{1}{c}\frac{\partial r}{\partial n}\frac{\partial p_s}{\partial t}\right\}\right]_{t=t_o-r/c,\, z=\zeta(x,y,t)} \tag{1.3-15}$$

We desire a comparison of this result with the Kirchhoff formula for monochromatic radiation Equation (1.2-5). Therefore, we assume that the scattered field which varies due to the surface motion is quasi-monochromatic or narrow band:

$$p_s(x,y,z,t) = \hat{p}_s(x,y,z,t)\, e^{-j\omega t} \tag{1.3-16}$$

where $\hat{p}_s(x,y,z,t)$ is the slowly varying (complex) amplitude. Inserting this into Equation (1.3-15) we have

$$\hat{p}_s(P,t_o) = \iint_{R_{xy}} \frac{dx\,dy}{4\pi}\left[\frac{H}{r}\left\{\frac{\partial \hat{p}_s}{\partial n} + \frac{\partial r}{\partial n}\frac{\hat{p}s}{r} + \frac{1}{c}\frac{\partial r}{\partial n}\left[\frac{\partial \hat{p}_s}{\partial t} - j\omega\,\hat{p}_s\right]\right\}\right] e^{jkr} \tag{1.3-17}$$
$$t=t_o-r/c,\quad z=\zeta(x,y,t)$$

We assume that the slow variations of the amplitude satisfy the inequality suggested by Eckart:

$$\boxed{\frac{1}{\hat{p}_s}\frac{\partial \hat{p}_s}{\partial t} << \omega = \frac{2\pi c}{\lambda}} \tag{1.3-18}$$

Taking advantage of this we find

$$\hat{p}_s(P,t_o) = \iint_{R_{xy}} \frac{dx\ dy}{4\pi}\left[H\ \left\{\frac{e^{jkr}}{r}\frac{\partial \hat{p}_s}{\partial n} + \hat{p}_s\ \left[\frac{\partial}{\partial n}\left(\frac{-1}{r}\right) - \frac{j\omega}{rc}\right] e^{jkr}\right\}\right]_{\substack{t=t_o-r/c \\ z=\zeta(x,y,t)}} \tag{1.3-19}$$

$$= \iint_{R_{xy}} \frac{dx\ dy}{4\pi}\left[H\ \left\{\frac{e^{jkr}}{r}\frac{\partial \hat{p}_s}{\partial n} - \hat{p}_s\ \frac{\partial}{\partial n}\left(\frac{e^{jkr}}{r}\right)\right\}\right]_{\substack{t=t_o-r/c \\ z=\zeta(x,y,t)}}$$

Recalling that $ds = H\ dx\ dy$ we see that the result is indeed similar to the time invariant Kirchhoi. formula as Eckart suggests, but in some cases it is important to remember the need for the retardation time in the integrand. This occurs when differences in retardation times within the active scattering area become comparable to the characteristic period of the surface time variations.

1.4 The Approximate First Order Solution of the Kirchhoff Integral Equation

In this section we examine more closely the transition from the Kirchhoff formula, Equation (1.2-5), to Eckart's approximate solution, Equation (1.2-6). We also examine the works of a few other authors which either contrast with the Eckart approach or improve it.

From the Dirichlet or pressure release condition assumed by Eckart we have

$$\hat{p}_s(x,y,z,t) = -\hat{p}_i(x,y,z) \quad [z=\zeta(x,y,t)] \tag{1.4-1}$$

Inserting this directly into Equation (1.3-19)

$$\hat{p}_s(P,t_o) = \frac{1}{4\pi} \iint_{S_r(t_o)} \hat{p}_i(x,y,z)\, \frac{\partial}{\partial n}\left(\frac{e^{jkr}}{r}\right) + \frac{e^{jkr}}{r}\, \frac{\partial}{\partial n}\, \hat{p}_s(x,y,z,t_o-r/c)\, dS \tag{1.4-2}$$

where $S_r(t_o)$ is the appropriately retarded surface. Unfortunately, the partial derivative of $\hat{p}_s$ with respect to n in the integrand of Equation (1.4-2) is unknown a-priori, and it cannot be assigned in an independent manner from the constraint Equation (1.4-1). Therefore, although $\hat{p}_i$ can usually be found quite easily, the problem of solving the integral equation (1.4-2) for $\hat{p}_s$ remains a difficult problem.

Currently, the only serious attempt to solve integral equations of the form (1.4-2) in a straightforward manner is given by Uretsky.[29] His analysis considers only the case of a plane wave reflected from a time invariant sinusoidal corrugation. The technique assumes that the normal derivative of $\hat{p}_s$ is periodic on the surface along the direction of the corrugation and a Fourier expansion is applied. This results in an infinite set of equations in an infinite number of coefficients. Uretsky

obtains the first few of these coefficients using approximate numerical methods. This technique, however, does not lend itself to the problem of scattering from arbitrary or random boundaries.

Other authors such as Rice,[30] or Marsh, Schulkin and Kneale[31] begin by extending the Rayleigh approach to the case of random surfaces. They assume a plane wave expansion for $\hat{p}_s$ with an unknown but position invariant space spectral density. These authors attempt to match the boundary condition (1.4-1) without making explicit reference to the Kirchhoff formula. This technique, however, also leads to an infinite set of equations for various components of the unknown spectral density and the solution requires conditions that imply weak scattering.

Although the plane wave expansion approach appears to be more attractive than the Uretsky technique from the point of view of flexibility it suffers from a serious defect. Lord and Murphy[32] show that the assumption of a constant spatial density of plane waves is in contradiction with the Kirchhoff formula. This difficulty was first noted by Lippmann[33] and later explained by Meecham[34] in an article predating the efforts of Marsh, Schulkin and Kneale. Meecham examines the problem of scattering from corrugations. Using an argument which is easily generalized to two dimensionally random surfaces (see Appendix A) Meecham shows that although a plane wave expansion exists, the density of the expansion becomes position dependent in the neighborhood of the surface deformations. This fact prevents the determination of the far-field wave density from the boundary condition Equation (1.4-1).

Alternatively, Mintzer[35] proposes a modification of Eckart's analysis which is essentially a perturbation-iteration scheme. This results in a

very flexible solution which is applicable for time varying random surfaces and which is theoretically more accurate than that obtained from plane wave expansions. In this treatment Equation (1.4-2) is rewritten as follows:

$$\hat{p}_s(P,t_o) = \frac{1}{4\pi} \iint_{S_r(t_o)} \frac{\partial}{\partial n}\left(\frac{e^{jkr}}{r}\, \hat{p}_i(x,y,z)\right) dS$$

$$+ \frac{\delta}{4\pi} \iint_{S_r(t_o)} \frac{e^{jkr}}{r}\, \frac{\partial}{\partial n}\left(\hat{p}_s(x,y,z,t_o-r/c) - \hat{p}_i(x,y,z)\right) dS \qquad (1.4\text{-}3)$$

where δ is a parameter of "smallness" which is 1 for the physical problem of interest. However, we consider the general problem of solving Equation (1.4-3) with arbitrary $\delta \leq 1$ and use δ to drive the iterative solution process.

The choice of the second term as the additive perturbation is motivated by the fact that it is identically zero for a flat, perfectly reflecting surface. In this case the boundary acts as a plane of symmetry or "mirror" surface which reflects "images" of the sources generating $\hat{p}_i$. This term is also known to be small when the amount of overshadowing of the surface S on itself is slight.[35] It should be noted that multiple scattering and other effects (such as a certain amount of diffraction) are also included in this term. Nevertheless, we broadly refer to this term as the "shadowing" term even when the frequency of radiation is too low to cast distinct shadows. The condition of slight overshadowing requires generally large angles of grazing and small surface slopes.

The standard perturbation-iteration approach to the solution of (1.4-3) is to assume that the solution can be grouped into terms of decreasing order of magnitude in δ:[36]

$$\boxed{\hat{p}_s = \hat{p}_s^{(0)} + \delta\, \hat{p}_s^{(1)} + \delta^2\, \hat{p}_s^{(2)} + \ldots} \qquad (1.4\text{-}4)$$

Substituting (1.4-4) into (1.4-3) and collecting terms of equal order in δ we have

$$\hat{p}_s^{(0)}(P,t_o) = \frac{1}{4\pi} \int_{S_r} \int_{(t_o)} \frac{\partial}{\partial n}\left(\frac{e^{jkr}}{r}\, \hat{p}_i(x,y,z)\right) dS \qquad (1.4\text{-}5)\ (a)$$

$$\hat{p}_s^{(1)}(P,t_o) = \frac{1}{4\pi} \int_{S_r(t_o)} \int \frac{e^{jkr}}{r} \frac{\partial}{\partial n}\left(\hat{p}_s^{(0)}(x,y,z,t_o - r/c) - \hat{p}_i(x,y,z)\right) dS \qquad (b)$$

$$\hat{p}_s^{(m)}(P,t_o) = \frac{1}{4\pi} \int_{S_r(t_o)} \int \frac{e^{jkr}}{r} \frac{\partial}{\partial n}\left(\hat{p}_s^{(m-1)}(x,y,z,t_o - r/c)\right) dS \qquad (c)$$

[for $m \geq 2$]

Thus, in principle the first iterative solution is determined by the incoming or free-space solution $\hat{p}_i$, and each successive iteration is obtained recursively from the next lower solution. The technique is, however, limited by two drawbacks. First, there is no assurance that the expansion in (1.4-4) converges with $\delta = 1$ for a given boundary or frequency. Second, even the integral for $\hat{p}_s^{(0)}$ cannot be performed exactly owing to the complexity of its integrand, especially for P in the near field.

Despite these limitations the perturbation-iteration method enjoys widespread acceptance in the literature. Frequently the iteration is

stopped after only the first term is obtained:

$$\hat{p}_s \simeq \hat{p}_s^{(0)} \tag{1.4-6}$$

This approximation which is exact for the flat surface tends to yield the correct frequency behavior even for mild scattering.[37] It is in this approximation that Eckart obtained his results. We shall limit ourselves in this report to the situations for which (1.4-6) is sufficiently accurate.

We now proceed to evaluate (1.4-5a). The normal derivative is conveniently computed in operator form as follows:

$$\frac{\partial}{\partial n} = \underline{n} \cdot \nabla = \frac{-\frac{\partial \zeta}{\partial x}\frac{\partial}{\partial x} - \frac{\partial \zeta}{\partial y}\frac{\partial}{\partial y} + \frac{\partial}{\partial z}}{H(x,y,t)} \tag{1.4-7}$$

where $H(x,y,t)$ is the secant of the angle between $\underline{n}$ and the vertical and $\partial\zeta/\partial x$ and $\partial\zeta/\partial y$ are the x and y slopes of the surface

Next we assume a directional point source located at the point Q (see Equation (1.2-5)). Examining the normal derivative in (1.4-5a) we have

$$\frac{\partial}{\partial n}\left[B_e(\theta,\phi)\,\frac{e^{jk(r+r')}}{rr'}\right] =$$

$$\frac{e^{jk(r+r')}}{rr'}\left[\frac{\partial B_e(\theta,\phi)}{\partial n} + B_e(\theta,\phi)\left\{\frac{\partial}{\partial n}[jk(r+r')] - \frac{1}{r}\frac{\partial r}{\partial n} - \frac{1}{r'}\frac{\partial r'}{\partial n}\right\}\right] \tag{1.4-8}$$

We consider only the case of wide beam patterns so that $\partial B_e(\theta,\phi)/\partial n$ can be ignored, and we limit ourselves to the far field where

$$k >> \frac{1}{r}, \frac{1}{r'}, \cdot \tag{1.4-9}$$

Hence, (1.4-5a) becomes with the use of (1.4-7) and (1.4-8) together with the relation $dS = dx\,dy\,H$

$$\hat{p}_s(P,t_o) = \iint_{R_{xy}} \frac{dx\,dy}{4\pi}\left[\frac{e^{jk(r+r')}}{rr'}\left\{\frac{-\partial\zeta}{\partial x}\frac{\partial}{\partial x}\frac{-\partial\zeta}{\partial y}\frac{\partial}{\partial y}\frac{+\partial}{\partial z}\right\}[jk(r+r')]\right]_{\substack{t=t_o-r/c \\ z=\zeta(x,y,t)}} \quad (1.4\text{-}10)$$

Following Gulin[38] and x,y,z coordinate system is selected in such a way that both the source point Q and the point of observation P lie in the plane defined by y=0, as shown in Figure 1.4-1. The origin of the coordinate system is chosen to lie in the plane $\zeta = 0$ (the plane defined by mean of the random deformations) and to be located at the specular point for wide beamwidth incident radiation. The beam pattern is here assumed to be aimed directly at the specular point. This point can be determined by finding the intersection of the plane $\zeta = 0$ and the line PQ' where Q' is the flat surface image point for the source.

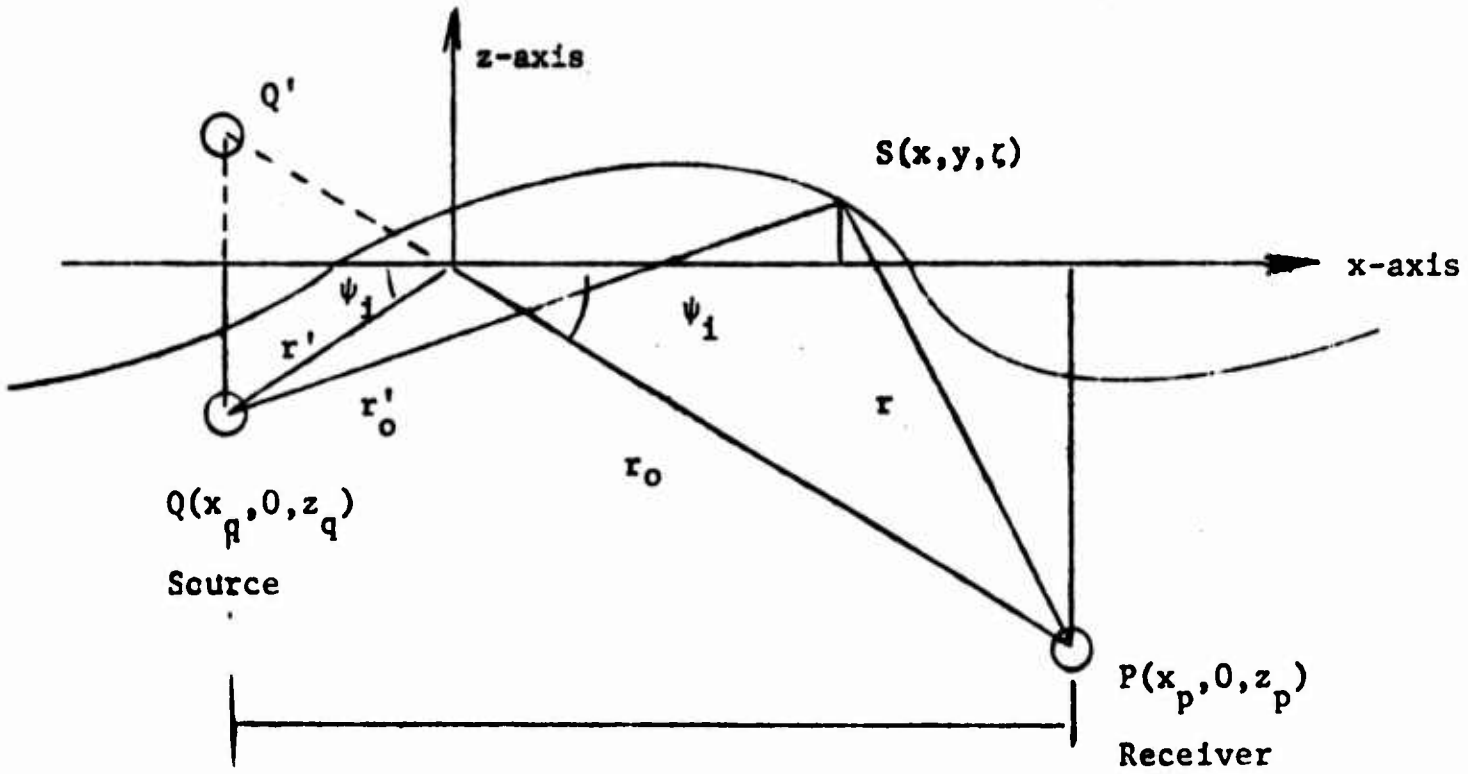


Figure 1.4-1

A key assumption which must be made in order to evaluate (1.4-10) is that the most important region of the integration is in the vicinity of the specular point even when the beam pattern is not very narrow. Beckmann[39] observes that while most of the integrand varies slowly over the surface, the exponent is very oscillatory. The most important contribution to the integral comes from within the first Fresnel zone of this spatial oscillation. The larger higher order Fresnel zones become more closely spaced and produce successively smaller contributions of alternating sign which tend to cancel. Provided the angle of grazing Ψ_i for the incident beam is large, the first zone is rather circular and centered at the specular point. At lower grazing angles the zones become ellipses elongated along the x-axis so that for wide beam patterns there are significant contributions made far from the specular point.

For large grazing angles, overshadowing is not important and the localized nature of the active scattering region eliminates the need for using different retardation times in the integrand of (1.4-10). It also becomes permissible to expand r and r' about the specular point. Retaining terms to quadratic order in x/r, x/r', y/r, and y/r' in the exponent of (1.4-10) we obtain (see Appendix C) a result which is slightly more accurate than that given by Gulin[38] for scattering into the specular direction:

$$\hat{p}_s(P,t_o) = \frac{jk\, e^{jk(r_o+r'_o)}}{4\pi\, r_o r'_o} \iint_{R_{xy}} B_e(\theta,\phi)\left[\frac{1}{4}\, a_s(x,y)\, (c_s^o)^2 \left(\frac{\partial \zeta}{\partial x}\right)\right. \qquad (1.4\text{-}11)$$

$$\left. + b_s(x,y) \left(\frac{\partial \zeta}{\partial y}\right) + c_s^o + [a^o - a'^o - a(x,y) + a'(x,y)]a^o\, c_s^o\right]$$

$$e^{jk[y^2/Re + \frac{1}{4}(x\, c_s^o)^2/Re - c_s^o\, \zeta(x,y,t_o - r_o/c)]}\, dx\, dy$$

Here a(x,y), b(x,y), c(x,y) are the x,y,z direction cosines for r and a'(x,y), b'(x,y), c'(x,y) are the x,y,z direction cosines for r'. The subscript "s" on a, b, or c denotes the sum of the primed and unprimed quantities and the superscript "o" denotes evaluation of these quantities at the specular point. The distances r_o and r'_o are the distances from the specular point to P and Q and

$$\boxed{R_e = \frac{2r_o r'_o}{r_o + r'_o}} \qquad (1.4\text{-}12)$$

Now <u>at</u> the specular point a_s, b_s, and c_s satisfy Equation (1.2-20). Assuming that these quantities do not vary appreciably in the active region of scattering (see Appendix C) we have Gulin's result

$$\boxed{\hat{p}_s(P,t_o) = \frac{jk\, c_s^o\, e^{jk(r_o+r'_o)}}{4\pi\, r_o r'_o} \iint dx\, dy\, B_e(\theta,\phi)\, e^{jk\left[y^2/R_e + \frac{1}{4}(x\, c_s^o)^2/R_e - c_s^o \zeta(x,y,t_o - r_o/c)\right]}} \qquad (1.4\text{-}13)$$

Finally, we examine the special case of perfect reflection and wide beamwidth. Setting $\zeta \to 0$, $B_e(\theta,\phi) \to 1$

$$\hat{p}_s(P,t_o) \to \frac{jk\, c_s^o\, e^{jk(r_o+r'_o)}}{4\pi\, r_o r'_o} \int_{-\infty}^{\infty} e^{jky^2/R_e}\, dy \int_{-\infty}^{\infty} e^{jk(x\, c_s^o)^2/4R_e}\, dx$$

$$= \frac{jk\, c_s^o\, e^{jk(r_o r'_o)}}{4\pi\, r_o r'_o} \left(\frac{2jR_e\pi}{k\, c_s^o}\right) = \frac{-\, e^{jk(r_o r'_o)}}{(r_o + r'_o)} \qquad (1.4\text{-}14)$$

The result is, of course, exact for all frequencies in spite of the fact that at low frequencies (large λ) the Fresnel zones are large. It is seen that Gulin's solution is superior to Eckart's for passive detection problems because Eckart's solution diverges in this limit.

CHAPTER 2

PRELIMINARY DISCUSSION OF SIGNAL AND SYSTEM PROPERTIES

2.0 Introduction

In this chapter the properties assumed for the signals and systems used in this report are summarized. Both signals and corrupting background noises are taken to be stationary random variables although for the most part only wide-sense stationarity need be assumed. Since the fluctuation of the output of finite-time correlators is examined in Chapter 3 it is desirable to assume that the background noise is Gaussian. This simplifies certain fourth order moments that arise in the computation of the variance of the fluctuation. For a similar reason it is desirable to assume that the target signals are Gaussian before they are scattered.

The scattering mechanisms are assumed to be in the form of random, linear, time-varying systems. The question of the various degrees of stationarity of these systems required for different problems is examined. Several equivalent representations for the effect of scattering are discussed. Each representation is useful in various applications. Some attention is focused on the problem of cascading two or more random, linear, time-varying systems. This problem becomes important in the processing of signals and in the representation of multiple scattering when the frequency smear of a particular scattering system is comparable to the bandpass properties of systems following it in the cascade arrangement. While the emphasis of the remaining chapters is placed on wide band detection, these latter results are needed for analysis of narrowband detectors.

2.1 Concerning the Nature of the Target Signal

We are restricting out attention to the passive detection of targets which generate real, stationary Gaussian signals of zero mean. Consider such a signal, $x(t)$ for $-\infty < t < \infty$. It is well-known that $x(t)$ can be expressed in terms of a Fourier Transform written in Stieltjes form[42] as follows*

$$x(t) = \frac{1}{2\pi} \int_{\omega} \exp(j\omega t)\, dz(\omega) \qquad (2.1\text{-}1)$$

where $z(\omega)$ is termed the spectral distribution corresponding to $x(t)$. From the real nature of the signal $x(t)$ it is seen that the increments of the distribution possess conjugate odd symmetry:

$$-dz(-\omega) = dz(\omega)^* \qquad (2.1\text{-}2)$$

Using (2.1-1) one may write the correlation between the value of x at two times t and t'

$$R(t,t') = \overline{x(t)\, x(t')^*} = \int_{\omega}\int_{\omega'} \exp(j\omega t - j\omega' t')\, \frac{\overline{dz(\omega)\, dz(\omega')^*}}{(2\pi)^2} \qquad (2.1\text{-}3)$$

where the overbar denotes ensemble averaging. By the hypothesis of statistical stationarity

$$R(t,t') = R(t-t') \qquad (2.1\text{-}4)$$

This functional dependence of the signal correlation on only the time

---

*Following a convention used in the literature, the absence of limits on an integral denotes integration throughout the range of the non-zero integrand while the symbolic label beneath the integral sign scans the domain $[-\infty,\infty]$ in a positive sense.

displacement $t-t'$ suggests a relation which can be written symbolically as follows:

$$\overline{dz(\omega)\ dz(\omega')^*} = 2\pi\ dZ(\omega)\ \delta(\omega-\omega')\ d\omega' \qquad (2.1\text{-}5)$$

where $\delta(\omega)$ is the Dirac delta function and $Z(\omega)$ is the power spectrum distribution function corresponding to $R(\tau)$:

$$R(\tau) = \frac{1}{2\pi} \int_{\omega} \exp(j\omega\tau)\ dZ(\omega) \quad . \qquad (2.1\text{-}6)$$

The random process $z(\omega)$ is said to have orthogonal increments when Equation (2.1-5) is satisfied.

Although this argument is rather loosely presented and its conclusions somewhat tersely stated, these results are derived in a more rigorous manner by Kolmogorov,[43] Doob,[44] and Yaglom.[45] The distribution function $Z(\omega)$ can be shown to consist of three components; a continuous component, a component with discrete jump discontinuities, and a continuous component with a derivative which vanishes almost everywhere.[46] Since the distribution $z(\omega)$ is a random variable it can be very discontinuous for certain ensemble realizations.[45] These discontinuities in $z(\omega)$ may exist in individual realizations even when $Z(\omega)$ is completely continuous or when, in fact, $Z(\omega)$ is differentiable and possesses a spectral density:

$$S_{xx}(\omega) = \frac{dZ(\omega)}{d\omega} \quad . \qquad (2.1\text{-}7)$$

When this spectral density exists Equation (2.1-5) may be rewritten as follows:

$$\overline{dz(\omega)\ dz(\omega')^*} = 2\pi\ S_{xx}(\omega)\ \delta(\omega-\omega')\ d\omega\ d\omega' \quad . \qquad (2.1\text{-}8)$$

The distribution $z(\omega)$ is determined to within a constant by the inverse of Equation (2.1-1).[45]

$$z(\omega) = \lim_{T\to\infty} \int_{-T}^{T} \frac{\exp(-j\omega t) - 1}{-jt}\ x(t)\ dt \quad . \qquad (2.1\text{-}9)$$

Since the mean of $x(t)$ is zero, using this definition we have

$$\overline{z(\omega)} = 0 \quad . \qquad (2.1\text{-}10)$$

Given any finite collection of x-samples, x_i for $i=1,2,\ldots,N$ we have the following probability density as a consequence of the Gaussian hypothesis

$$f(\underline{x}) = \frac{1}{\sqrt{(2\pi)^N\ |R|}} \exp\left(-\tfrac{1}{2}\ \underline{x}^T\ \underline{R}^{-1}\ \underline{x} \right) \qquad (2.1\text{-}11)$$

where $\underline{x}^T$ is the row vector $(x_1,x_2,\ldots,x_N)$ and $\underline{R}$ is the matrix of elements $\overline{\underline{x}\ \underline{x}^T}$. The characteristic function or moment generating function corresponding to this density is then

$$Q(\underline{u}) = \overline{\exp(-j\ \underline{x}^T\ \underline{u})} = \exp\left(-\tfrac{1}{2}\ \underline{u}^T\ \underline{R}\ \underline{u} \right) \qquad (2.1\text{-}12)$$

It is easily seen that any linear combination of x samples is also a Gaussian random variable.[48] In particular, the spectral distribution is a continuous linear combination of x values as shown by Equation (2.1-10) and it is therefore Gaussian.

The Gaussian nature of $z(\omega)$ can be used to simplify various higher order moments of the increments $dz(\omega)$. Since

$$\overline{dz(\omega)} = 0 \tag{2.1-13}$$

odd order moments of the form

$$\overline{dz(\omega)\ dz(\omega')^*\ dz(\omega'')} \tag{2.1-14}$$

vanish identically.[49] Furthermore, higher order even moments can be expressed in terms of linear combinations of products of second order moments. For the fourth order moment we have[49]

$$\begin{aligned}\overline{dz(\omega)\ dz(\omega')^*\ dz(\omega'')\ dz(\omega''')^*} = \\ \overline{dz(\omega)\ dz(\omega')^*} \times \overline{dz(\omega'')\ dz(\omega''')^*} + \\ \overline{dz(\omega)\ dz(\omega'')} \times \overline{dz(\omega''')^*\ dz(\omega')^*} + \\ \overline{dz(\omega)\ dz(\omega''')^*} \times \overline{dz(\omega')^*\ dz(\omega'')}\end{aligned} \tag{2.1-15}$$

The second order moments of the first and third term in Equation (2.1-15) may be simplified by direct application of Equation (2.1-8). To reduce the second term we invoke Equation (2.1-2):

$$\begin{aligned}\overline{dz(\omega)\ dz(\omega'')} &= -\overline{dz(\omega)\ dz(-\omega'')^*} \\ &= 2\pi\ S_{xx}(\omega)\ \delta(\omega+\omega'')\ d\omega\ d\omega''\end{aligned} \tag{2.1-16}$$

and

$$\begin{aligned}\overline{dz(\omega''')^*\ dz(\omega')^*} &= -\overline{dz(-\omega''')\ dz(\omega')^*} \\ &= 2\pi\ S_{xx}(\omega''')\ \delta(\omega'''+\omega')\ d\omega'''\ d\omega'\end{aligned} \tag{2.1-17}$$

Therefore, (2.1-15) becomes

$$\begin{aligned}\overline{dz(\omega)\ dz(\omega')^*\ dz(\omega'')\ dz(\omega''')^*} &= \\ = (2\pi)^2\ S_{xx}(\omega)\ [&\delta(\omega-\omega')\ S_{xx}(\omega'')\ \delta(\omega''-\omega''') + \\ &\delta(\omega+\omega'')\ S_{xx}(\omega''')\ \delta(\omega'''+\omega') + \\ &\delta(\omega-\omega''')\ S_{xx}(\omega')\ \delta(\omega'-\omega'')]\ d\omega\ d\omega'\ d\omega''\ d\omega'''\end{aligned} \tag{2.1-18}$$

Finally, by entirely analogous reasoning we may represent any pair of jointly stationary random variables by using orthogonal increments:

$$y(t) = \frac{1}{2\pi}\int_{\omega} e^{j\omega t}\ dz_y(\omega) \quad ; \quad x(t') = \frac{1}{2\pi}\int_{\omega'} e^{j\omega' t'}\ dz_x(\omega') \tag{2.1-19}$$

where

$$\overline{dz_x(\omega)\ dz_y(\omega')^*} = 2\pi\ dZ_{xy}(\omega)\ \delta(\omega-\omega')\ d\omega' \tag{2.1-20}$$

and where $Z_{xy}(\omega)$ is the cross-spectral distribution corresponding to the cross-correlation function for x and y :

$$R_{xy}(\tau) = \overline{x(t)\ y(t-\tau)} = \frac{1}{2\pi}\int_{\omega} \exp(j\omega\tau)\ dZ_{xy}(\omega) \tag{2.1-21}$$

Assuming $Z_{xy}(\omega)$ to possess a spectral density $S_{xy}(\omega)$:

$$S_{xy}(\omega) = \frac{dZ_{xy}(\omega)}{d\omega} \tag{2.1-22}$$

we have, following Equation (2.1-18)

$$\overline{dz_x(\omega)\ dz_x(\omega')^*\ dz_y(\omega'')\ dz_y(\omega''')^*} =$$

$$= (2\pi)^2 \left[S_{xx}(\omega)\ \delta(\omega-\omega')\ S_{yy}(\omega'')\ \delta(\omega''-\omega''') + \right.$$

$$S_{xy}(\omega)\ \delta(\omega+\omega'')\ S_{xy}(\omega''')\ \delta(\omega'''+\omega') +$$

$$\left. S_{xy}(\omega)\ \delta(\omega-\omega''')\ S_{yx}(\omega')\ \delta(\omega'-\omega'')\right] d\omega\ d\omega'\ d\omega''\ d\omega''' \qquad (2.1\text{-}23)$$

2.2 System Function Description for Linear Time Varying Filters

As indicated in Chapter 1 a truly accurate and complete description of the effect of random surface scattering on reflected signals is very difficult to obtain. Perhaps the only obvious remark which one can initially make with some assurance is that at the power levels commonly encountered in passive detection the phenomenon of surface scattering is linear. This is simply due to the linearity of the basic wave equation (1.3-2). Armed with this knowledge alone we may draw several very interesting conclusions with regard to optimal and sub-optimal receiver design. However, any effort to compute parameters in such designs or to arrive at some sort of understanding of the degree of the optimality attained ultimately requires the construction of a realistic scattering model from the underlying physics.

In this section we examine some of the well-known properties of surface scattering that are due to linearity. In particular, we study extensions of the techniques used by Ellinthorpe and Nuttal,[50] Lindenlaub,[51] and Kailath[17] to describe the dynamic input-output relationships for Linear Time-varying Filters (hereinafter referred to as LTVF's). This notation serves as a vehicle for stating those properties of scattering which one can advance without making reference to a specific physical model. The eventual goal is, however, to find approximate physical counterparts for the notational conveniences.

We recall that the outputs $y(t)$ of a LTVF can be related to their corresponding inputs by a linear functional relation of the form

$$y(t) = \int_{\sigma} h(\sigma,t)\, x(t-\sigma)\, d\sigma \qquad (2.2\text{-}1)$$

In particular, we have the following function pair which satisfies this mapping

$$x(t) = \delta(t-t_o)$$
$$y(t) = h(t-t_o,t) \qquad (2.2\text{-}2)$$

which identifies $h(\sigma,t)$ as the response of the LTVF at time t due to a unit strength impulse applied σ units of time previously. The parameter σ is therefore an "elapsed-time" or memory variable. For a given value of σ, $h(\sigma,t)$ fluctuates slowly as a function of t for the class of scattering systems of interest in this report. Formally we identify these slow fluctuations with the surface motion. On the other hand, for fixed t the variation of $h(\sigma,t)$ with σ is much more rapid and generally of short duration.

By substituting $\nu = t-\sigma$ into Equation (2.2-1) we obtain a slightly different but equivalent representation of the input-output relation:

$$y(t) = \int_{\nu} h'(\nu,t)\, x(\nu)\, d\nu \qquad (2.2\text{-}3)$$

In this representation the new weighting function $h'(\nu,t)$ is the reflected and translated impulse response

$$\boxed{h'(\nu,t) = h(t-\nu,t)} \qquad (2.2\text{-}4)$$

For this representation we have the function pair

$$x(t) = \delta(t-t_o)$$
$$y(t) = h'(t_o,t) \qquad (2.2\text{-}5)$$

Therefore, $h'(\nu,t)$ is the impulse response at time t due to excitation at time ν. Whereas σ is a relative time variable, the parameter

ν is an absolute time. Although the easy identification of the "fast" and "slow" axes is an attractive feature in the unprimed impulse response, Lindenlaub indicates that for certain applications the primed notation is more advantageous.

We now consider three classes of Fourier transforms on the primed and unprimed weighting functions; primary axis transforms, secondary axis transforms, and dual or combined transforms:

A) Primary or First Axis Transforms

In both the primed and unprimed systems of notation the response of the LTVF to sinusoidal excitation assumes a role of importance. This is primarily due to the fact that the sinusoidal response is frequently more readily accessible from physical considerations than the impulse response. Moreover, the approximations used in the analysis of scattering tend to be better at some frequencies than at others. Thus, we define primary axis transforms $H(\omega,t)$ and $H'(\omega,t)$ which are generated by the substitution of

$$x(t) = \exp(+j\omega t) \tag{2.2-6}$$

into Equations (2.2-1) and (2.2-3)

$$H(\omega,t) = y(t)/\exp(+j\omega t) = \int_{\sigma} h(\sigma,t)\, \exp(-j\omega\sigma)\, d\sigma \tag{2.2-7}$$

$$H'(\omega,t) = y(t) = \int_{\nu} h'(\nu,t)\, \exp(+j\omega\nu)\, d\nu \tag{2.2-8}$$

The quantity $H(\omega,t)$ is termed the "instantaneous transfer function" by Zadeh[64] while $H'(\omega,t)$ is termed the sinewave response by Lindenlaub. The positive sign convention used in the exponent in Equation (2.2-6) is opposite to that used in Equation (1.2-5), Equation (1.2-16) and the

general literature on surface scattering, but it leads to slightly more symmetric results in this chapter. The difference in conventions is reconciled at a later point in this report.

The primed and unprimed transforms are related by

$$\boxed{H'(\omega,t) = H(\omega,t)\ \exp(+j\omega t)} \qquad (2.2\text{-}9)$$

For fixed ω both $H'(\omega,t)$ and $H(\omega,t)$ are slowly varying with t. These primary axis transforms can be used to restate Equation (2.2-1), the basic linear functional relation for general excitations:

$$y(t) = \frac{1}{2\pi}\int H'(\omega,t)\ dz_x(\omega) = \frac{1}{2\pi}\int H(\omega,t)\ \exp(+j\omega t)\ dz_x(\omega) \qquad (2.2\text{-}10)$$

where $z_x(\omega)$ is the distribution corresponding to $x(t)$ via Equation (2.1-1).

B) Secondary Axis Transforms

By transforming the second axis of the primed and unprimed impulse response we obtain another useful pair of system functions. These are written in terms of spectral distributions*

$$h(\sigma,t) = \frac{1}{2\pi}\int_{\gamma} \exp(j\gamma t)\ dG(\sigma,\gamma) \qquad (2.2\text{-}11)$$

$$h'(\nu,t) = \frac{1}{2\pi}\int_{\gamma} \exp(j\gamma t)\ dG'(\nu,\gamma) \qquad (2.2\text{-}12)$$

---

*We here adopt the convention that $dG(\sigma,\gamma)$ represents an increment in the distribution G over γ with σ normally an independent parameter. When ambiguity might arise, the notation $d_{\gamma}G(\sigma,\beta)$ denotes $dG(\sigma,\gamma)\cdot(\partial\beta/\partial\gamma)$ assuming the derivative exists.

where if h, h', G, and G' are all well-behaved functions on their primary axes we have

$$G(\sigma,\gamma) = \lim_{T\to\infty} \int_{-T}^{+T} \frac{\exp(-j\gamma t) - 1}{-jt} h(\sigma,t)\, dt \qquad (2.2\text{-}13)$$

$$G'(\nu,\gamma) = \lim_{T\to\infty} \int_{-T}^{+T} \frac{\exp(-j\gamma t) - 1}{-jt} h'(\nu,t)\, dt \qquad (2.2\text{-}14)$$

Substituting Equation (2.2-11) into Equation (2.2-1) we obtain

$$y(t) = \int_{\sigma}\int_{\gamma} x(t-\sigma) \exp(j\gamma t) \frac{dG(\sigma,\gamma)}{2\pi} d\sigma \qquad (2.2\text{-}15)$$

or, applying Equations (2.1-19a,b) and using uniqueness for the increment representation

$$dz_y(\omega) = \int\int_{\sigma\ \gamma} e^{-j(\omega-\gamma)\sigma}\, d_\omega z_x(\omega-\gamma) \frac{dG(\sigma,\gamma)}{2\pi} d\sigma \qquad (2.2\text{-}16)$$

Following Ellinthorpe and Nuttall we observe that the quantity

$$d_\omega z_x(\omega-\gamma) \exp(-j(\omega-\gamma)\sigma) \qquad (2.2\text{-}17)$$

is the spectral increment for a σ time delayed and a γ frequency shifted replica of the input signal $x(t)$. Thus, since Equation (2.2-16) represents the spectral increments of $y(t)$ as a superposition of these terms, $G(\sigma,\gamma)$ measures the extent to which the input is spread along the delay and frequency axes. For this reason $G(\sigma,\gamma)$ is termed the system distributional "spreading" function.

By comparison, on applying similar methods to Equations (2.2-3) and (2.2-14) one obtains

$$dz_y(\gamma) = \int_{\nu} x(\nu) \frac{dG'(\nu,\gamma)}{2\pi} d\nu \qquad (2.2\text{-}18)$$

This relation is considerably simpler than Equation (2.2-16), but the interpretation of $G'(\nu,\gamma)$ is not as intuitively appealing as it is for $G(\sigma,\gamma)$. Following Lindenlaub we refer to $G'(\nu,\gamma)$ simply as the impulse spectral distribution. The relation between the impulse spectrum and the system spreading function can be found by substituting Equation (2.2-4) into Equation (2.2-14) and applying Equation (2.2-11):

$$G'(\nu,\gamma) = \lim_{T\to\infty} \int_{-T}^{+T} \frac{\exp(-j\gamma t)-1}{-jt} \left\{ \frac{1}{2\pi} \int_{\gamma'} e^{j\gamma' t} \, dG\,(t-\nu,\gamma') \right\} dt \qquad (2.2\text{-}19)$$

The complexity of this relation accounts for the simplicity of Equation (2.2-18) and the usefulness of the primed kernel depends on its availability from other considerations.

C) Dual Axis Transforms

Finally, we explore the fourth transform pair in this symmetric class of system functions. These are the primed and unprimed bi-frequency relations:

$$dg(\omega,\gamma) = \int_{\sigma} \exp(-j\omega\sigma) \, dG(\sigma,\gamma) \, d\sigma \qquad (2.2\text{-}20)$$

$$dg'(\omega,\gamma) = \int_{\nu} \exp(+j\omega\nu) \, dG'(\nu,\gamma) \, d\nu \qquad (2.2\text{-}21)$$

The relation of these bi-frequency spectral distributions to the fast axis system functions may be found by transforming Equation (2.2-11) and Equation (2.2-12) on σ and ν respectively:

$$H(\omega,t) = \frac{1}{2\pi} \int_{\gamma} e^{j\gamma t} \, dg(\omega,\gamma) \qquad (2.2\text{-}22)$$

$$h'(\omega,t) = \frac{1}{2\pi}\int_{\gamma} e^{j\gamma t}\, dg'(\omega,\gamma) \qquad (2.2\text{-}23)$$

Furthermore, by applying the inverses of Equation (2.2-7) and Equation (2.2-8) these become

$$h(\sigma,t) = \int_{\omega}\int_{\gamma} e^{j(\gamma t+\omega\sigma)}\, \frac{dg(\omega,\gamma)}{2\pi}\, \frac{d\omega}{2\pi} \qquad (2.2\text{-}24)$$

$$h'(\nu,t) = \int_{\omega}\int_{\gamma} e^{j(\gamma t-\omega\nu)}\, \frac{dg'(\omega,\gamma)}{2\pi}\, \frac{d\omega}{2\pi} \qquad (2.2\text{-}25)$$

On substitution of Equation (2.2-9) into Equation (2.2-23) and comparing with Equation (2.2-22) we find by uniqueness that

$$\boxed{dg'(\omega,\gamma) = d_{\gamma}g(\omega,\gamma-\omega)} \qquad (2.2\text{-}26)$$

The quantity $g(\omega,\gamma)$ is termed the bi-frequency spectral distribution function (following the notation of Ellinthorpe) while $g'(\omega,\gamma)$ is termed the sinewave spectral distribution function (following Lindenlaub). By executing the integration over σ in Equation (2.2-16) and invoking Equation (2.2-20) we obtain the last set of input-output relations for the fundamental system functions:

$$dz_y(\omega) = \frac{1}{2\pi}\int_{\gamma} d_{\gamma}g(\omega-\gamma,\gamma)\; d_{\omega}z_x(\omega-\gamma) \qquad (a)$$

$$= \frac{1}{2\pi}\int_{\omega'} d_{\omega}g(\omega',\omega-\omega')\; d_{\omega'}z_x(\omega') \qquad (b) \qquad (2.2\text{-}27$$

$$= \frac{1}{2\pi}\int_{\omega'} dg'(\omega',\omega)\; dz_x(\omega') \qquad (c)$$

By comparison with the distribution $G(\sigma,\gamma)$ which measures the spread of frequency shifts for various arrival delays, $g(\omega,\gamma)$ measures the total amount of "leakage" of signal at input frequency ω into output frequency γ cycles displaced from ω. It therefore offers a natural scheme for analysis or decomposition of system response into side bands.

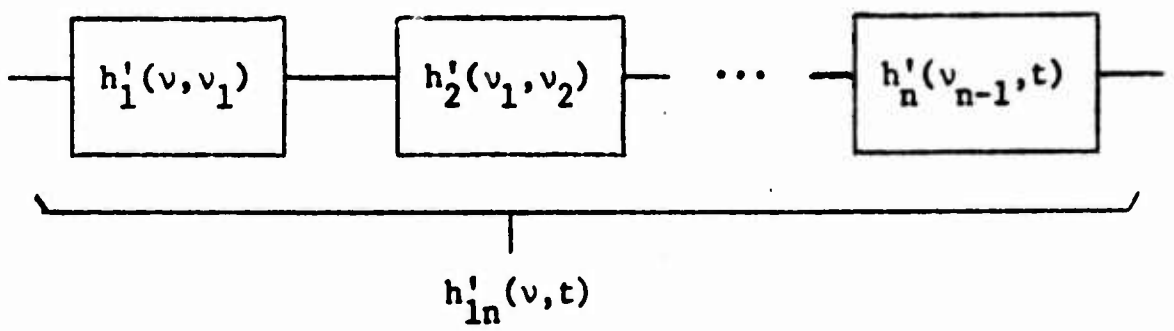
By comparison with the distribution $G(\sigma,\dot{\gamma})$ which measures the spread of frequency shifts for various arrival delays, $g(\omega,\gamma)$ measures the total amount of "leakage" of signal at input frequency ω into output frequency γ cycles displaced from ω. It therefore offers a natural scheme for analysis or decomposition of system response into side bands.

2.3 Cascaded Linear Time Varying Systems

Lindenlaub observes that the input-output working relationships for LTVF's are generally somewhat simpler when expressed in terms of the primed system functions. This feature makes the primed notation especially suitable for analysis of cascaded chains of LTVF's (see Figure 2.3-1). To illustrate this let us consider n LTVF's cascaded in this manner. By repeated application of Equation (2.2-3) one can immediately show that the overall impulse response $h'_{1n}(\nu,t)$ for the cascade is given by

$$h'_{1n}(\nu,t) = \int_{\nu_1}\int_{\nu_2}\cdots\int_{\nu_{n-1}} h'_1(\nu,\nu_1)\,d\nu_1\; h'_2(\nu_1,\nu_2)\,d\nu_2 \cdots d\nu_{n-1}\; h'_n(\nu_{n-1},t) \tag{2.3-1}$$

Although this expression is very symmetric it is clear that generally the n systems in the chain cannot be permuted or interchanged without changing the overall behavior of the chain.



Cascade of n Linear Time Varying Filters

Figure 2.3-1

Since $h'_i(\nu_{i-1},\nu_i)$ is usually not known a-priori in many practical situations a frequency domain counterpart for Equation (2.3-1) is

desirable. By utilizing Equations (2.2-12) and (2.2-25) one obtains the following intermediate result

$$\int_{\nu_1} h_1'(\nu,\nu_1)\, d\nu_1\, h_2'(\nu_1,\nu_2) = \int_{\gamma_1}\int_{\gamma_2} \frac{1}{2\pi}\, dG_1'(\nu,\gamma_1)\, \frac{1}{2\pi}\, dg_2'(\gamma_1,\gamma_2)\, e^{j\gamma_2\nu_2} \tag{2.3-2}$$

Through continued application of Equation (2.2-25) and a final use of Equation (2.3-23) we obtain

$$H_{1n}'(\omega,t) = \int_{\gamma_1}\int_{\gamma_2}\cdots\int_{\gamma_n} dg_1'(\omega,\gamma_1)\, \frac{1}{2\pi}\, dg_2'(\gamma_1,\gamma_2)\, \frac{1}{2\pi} \cdots \frac{1}{2\pi}\, dg_{n-2}'(\gamma_{n-2},\gamma_{n-1})\, \frac{1}{2\pi}\, H_n'(\gamma_{n-1},t) \tag{2.3-3}$$

Clearly, many alternative representations for cascaded systems are possible. In particular, it is occasionally useful to force the first system in the chain to be represented by its sinewave response. For a two element chain we have the especially simple result due to Lindenlaub:

$$H_{12}'(\omega,t) = \int_{\nu_1} H_1'(\omega,\nu_1)\, h_2'(\nu_1,t)\, d\nu_1 \tag{2.3-4}$$

Since the scattering system is usually the first element in a typical chain, it is very advantageous to have the leading element a primary axis transform.

As in Equation (2.3-3) we may extend Equation (2.3-4) by sequential

application of Equation (2.2-25) followed by a final use of Equation (2.2-8) with the following result for long chains:

$$H'_{1n}(\omega,t) = \int_{\nu_1}\int_{\gamma_2}\cdots\int_{\gamma_{n-1}} H'_1(\omega,\nu_1)\,\frac{d\nu_1}{2\pi}\,dG'_2(\nu_1,\gamma_2)\,\frac{1}{2\pi}\,dg'_3(\gamma_2,\gamma_3) \cdots \frac{1}{2\pi}\,dg'_{n-1}(\gamma_{n-2},\gamma_{n-1})\,\frac{1}{2\pi}\,H'_n(\gamma_{n-1},t) \tag{2.3-5}$$

Finally, by consecutive application of Equation (2.2-27) one can easily generate still another relation:

$$dg'_{1n}(\omega,\gamma) = \int_{\gamma_1}\int_{\gamma_2}\cdots\int_{\gamma_{n-1}} dg'_1(\omega,\gamma_1)\,\frac{1}{2\pi}\,dg'_2(\gamma_1,\gamma_2)\,\frac{1}{2\pi} \cdots \frac{1}{2\pi}\,dg'_n(\gamma_{n-1},\gamma) \tag{2.3-6}$$

2.4 Stationary Stochastic System Correlations

In section 2.2 the properties of general LTVF's are examined subject to general excitations. No explicit use is made there, however, of the possible random nature of either the signals or the systems. In this section we investigate various properties of Random Linear Time-Varying Filters (RLTVF's) subject to excitation by stationary random variables. The loss of signal correlation on passage through RLTVF's is measured from several points of view. The degradation of the cross-correlation between the input and output is determined. The deterioration of the auto-correlation of the output with respect to the input is obtained. In addition, we examine the cross-correlation between outputs of RLTVF's taken in parallel and in series cascade.

We distinguish in the following between averages carried out over the ensemble of excitations, the ensemble of random or scattering filters, and the union of these two sets. Ensemble averaging over excitations is denoted by an overbar labeled "x" while RLTVF ensemble averaging is denoted by "s" .

Also discussed is the definition used for system stochastic stationarity and a variety of useful system correlation functions.

A) The Input-to-Output Cross Decorrelation Function

The simplest stochastic input-output relation for a single RLTVF concerns the input-output cross-correlation:

$$R_{xy}(\tau) = \overline{x(t+\tau)\, y(t)} = \int \overline{\overline{h(\sigma,t)}^{s}\, \overline{x(t+\tau)\, x(t-\sigma)}^{x}}\, d\sigma \qquad (2.4\text{-}1)$$

$$= \int_{\sigma} h_c(\sigma)\, R_{xx}(\tau-\sigma)\, d\sigma \qquad (a)$$

$$= \int_{\omega} H_c(\omega)\, S_{xx}(\omega)\, \exp(+j\omega\tau)\, \frac{d\omega}{2\pi} \qquad (b)$$

$$= \int_{\sigma'} D_c(\tau,\sigma')\, R_{xx}(\sigma')\, d\sigma' \qquad (c)$$

where

$$\text{(a)} \quad h_c(\sigma) = \overline{h(\sigma,t)}^{s} \quad ; \quad \text{(b)} \quad H_c(\omega) = \overline{H(\omega,t)}^{s} \tag{2.4-2}$$

are respectively the so-called coherent impulse response and coherent transfer functions for the RLTVF . These are related by

$$h_c(\sigma) = \int_{\omega} H_c(\omega)\ e^{j\omega\sigma}\ \frac{d\omega}{2\pi} \tag{2.4-3}$$

Equations (2.4-2), of course, assume system stationarity at least in the mean.

Because scattering invariably causes loss of correlation we refer to

$$\boxed{D_c(\tau,\sigma') = h_c(\tau-\sigma')} \tag{2.4-4}$$

as the input-to-output cross decorrelation function or the "straight-through" coherency degradation.

B) The Output Auto Decorrelation Function

The second stochastic input-output relation for a single RLTVF concerns the output auto-correlation:

$$R_{yy}(t,t') = \overline{y(t)\ y(t')^*} = \tag{2.4-5}$$

$$\left(\frac{1}{2\pi}\right)^2 \int_{\omega}\int_{\omega'} \overline{H(\omega,t)\ H^*(\omega',t')}^{s}\ \exp(j\omega t - j\omega' t')\ \overline{dz_x(\omega)\ dz_x(\omega')^*}^{x}$$

Assuming that $x(t)$ is a stationary random variable in the wide-sense then from Equations (2.1-5) and (2.1-8)

$$R_{yy}(t,t') = \frac{1}{2\pi}\int_{\omega} \Phi(\omega,\omega,t,t')\ \exp(j(t-t')\omega)\ dZ_{xx}(\omega)$$
$$= \int_{\omega} \Phi(\omega,\omega,t,t')\ \exp(j(t-t')\omega)\ S_{xx}(\omega)\ \frac{d\omega}{2\pi} \qquad (2.4\text{-}6)$$

where the latter equality holds in a spectral density $S_{xx}(\omega)$ exists. Here we define

$$\boxed{\Phi(\omega,\omega',t,t') = \overline{H(\omega,t)\ H^{*}(\omega',t')}^{s}} \qquad (2.4\text{-}7)$$

as the interfrequency system correlation function. It can be seen that a sufficient condition for the output of the RLTVF, y(t), to be stationary in the wide sense is

$$\Phi(\omega,\omega,t,t') = \Phi(\omega,\omega,t-t') \qquad (2.4\text{-}8)$$

We call such a system a wide sense stationary system (WSS). For this class of systems we have

$$\boxed{R_{yy}(\tau) = \frac{1}{2\pi}\int_{\omega} \Phi(\omega,\omega,\tau)\ e^{j\omega\tau}\ dZ_{xx}(\omega)} \qquad (2.4\text{-}9)$$

Note that since the mean of x(t) is assumed to be zero, y(t) is also zero mean.

Now we define a more restrictive class of RLTVF's which satisfy a broader type of stationarity:

$$\boxed{\Phi(\omega,\omega',t,t') = \Phi(\omega,\omega',t-t')} \qquad (2.4\text{-}10)$$

Such systems are here termed interfrequency-wide-sense stationary (IWSS). All of the random scattering systems to be studied satisfy this condition.

An important property of this class of systems is that one may define for it a new spectral distribution:

$$\Phi(\omega,\omega',\tau) = \frac{1}{2\pi}\int_{\omega''} e^{j\omega''\tau}\, d_{\omega''}A(\omega,\omega',\omega'') \qquad (2.4\text{-}11)$$

We refer to the function A as the "tri-frequency" spectral distribution. By substituting Equation (2.4-11) into Equation (2.4-9) we obtain a stochastic input-output relation

$$R_{yy}(\tau) = \left(\frac{1}{2\pi}\right)^2 \int_{\omega}\int_{\omega''} e^{j(\omega+\omega'')\tau}\, d_{\omega''}A(\omega,\omega,\omega'')\, dZ_{xx}(\omega) \qquad (2.4\text{-}12)$$

$$= \frac{1}{2\pi}\int_{\omega'} e^{j\omega'\tau}\left\{\frac{1}{2\pi}\int_{\omega} d_{\omega'}A(\omega,\omega,\omega'-\omega)\, dZ_{xx}(\omega)\right\}$$

By the uniqueness of the spectral increment representation this becomes*

$$dZ_{yy}(\omega') = \frac{1}{2\pi}\int_{\omega} d_{\omega'}A(\omega,\omega,\omega'-\omega)\, dZ_{xx}(\omega) \qquad (2.4\text{-}13)$$

From the similarity between Equations (2.4-13) and (2.2-27) is seen that the function $A(\omega,\omega,\omega'-\omega)$ operates on the input power spectral increments $dZ_{xx}(\omega)$ in the same way that $g(\omega,\omega'-\omega)$ operates on $dz_x(\omega)$. That is,

---

*$d_{\omega'}A(\omega,\omega,\omega'-\omega)/d\omega' = \Gamma(\omega,\omega')$ is called a bi-frequency function by Zadeh and others in the literature.

A measures the average amount of input power at frequency ω which is parametrically "pumped" into output frequency ω .

This relationship between the tri-frequency and bi-frequency functions can be obtained by a more direct method. Starting with Equation (2.2-22) we have

$$\phi(\omega,\omega',t-t') = \left(\frac{1}{2\pi}\right)^2 \int_{\gamma}\int_{\gamma'} e^{j(\gamma t-\gamma' t')} \overline{dg(\omega,\gamma)\ dg^*(\omega',\gamma')} \qquad (2.4\text{-}14)$$

This in turn suggests the relation

$$\boxed{\overline{dg(\omega,\gamma)\ dg^*(\omega',\gamma')} = 2\pi\ \delta(\gamma-\gamma')\ dA(\omega,\omega',\gamma)} \qquad (2.4\text{-}15)$$

That is, the increments in the bi-frequency distribution function are orthogonal even at different values of the fast axis transform frequencies.

The orthogonality of the bi-frequency function increments obviously carries over to the increments of the system spreading function distribution. Inverting Equation (2.2-20) we find

$$\overline{dG(\sigma,\gamma)\ dG^*(\sigma',\gamma')} =$$

$$\left(\frac{1}{2\pi}\right)^2 \int_{\omega}\int_{\omega'} \exp(-j\omega\sigma+j\omega'\sigma')\ \overline{dg(\omega,\gamma)\ dg^*(\omega',\gamma')}\ d\omega\ d\omega' \qquad (2.4\text{-}16)$$

$$= 2\pi\ \delta(\gamma-\gamma')\ dB(\sigma,\sigma',\gamma)$$

where

$$dB(\sigma,\sigma',\gamma) = \int_{\omega}\int_{\omega'} dA(\omega,\omega',\gamma)\ \exp\{-j(\omega\sigma-\omega'\sigma')\}\ \frac{d\omega}{2\pi}\ \frac{d\omega'}{2\pi} \qquad (2.4\text{-}17)$$

This result may be combined with Equation (2.2-16) to yield a power spectrum

spreading function relationship

$$\overline{dz_y(\omega)\ dz_y(\omega')^*} = 2\pi\ \delta(\omega-\omega')\ dZ_{yy}(\omega) =$$

$$\int_{\sigma}\int_{\sigma'}\int_{\gamma}\int_{\gamma'} 2\pi\ dZ_{xx}(\omega-\gamma)\ \delta(\omega-\omega'-\gamma+\gamma')\ \left(\frac{1}{2\pi}\right)^2\ \overline{dG(\sigma,\gamma)\ dG^*(\sigma',\gamma')}$$

$$\exp\{-j(\omega-\gamma)\sigma+j(\omega'-\gamma')\sigma'\}\ d\sigma\ d\sigma' \qquad (2.4\text{-}18)$$

Using Equation (2.4-16) and integrating Equation (2.4-18) with respect to ω' and γ'

$$dZ_{yy}(\omega) = \int_{\gamma}\int_{\sigma}\int_{\sigma'} dZ_{xx}(\omega-\gamma)\ \exp\{-j(\omega-\gamma)(\sigma-\sigma')\}\ dB(\sigma,\sigma',\gamma)\ d\sigma\ d\sigma' \qquad (2.4\text{-}19)$$

$$= \int_{\gamma}\int_{\sigma''} dZ_{xx}(\omega-\gamma)\ \exp\{-j(\omega-\gamma)\sigma''\}\ \frac{dC(\sigma'',\gamma)}{2\pi}\ d\sigma''$$

where $\sigma'' = \sigma'-\sigma$ and

$$dC(\sigma'',\gamma) = \int_{\sigma} dB(\sigma,\sigma+\sigma'',\gamma)\ d\sigma \qquad (2.4\text{-}20)$$

Here C might be termed the power spectrum "smearing" function. Equation (2.4-20) shows that the output signal correlation can be represented as a distribution replicas of the input correlation which are shifted in frequency and over arrival times.

This leads us immediately to the most enlightening stochastic input-output relation. From Equation (2.4-19) we have

$$R_{yy}(\tau) = \int_{\sigma}\int_{\omega}\int_{\gamma} dZ_{xx}(\omega-\gamma)\ \exp\{j(\omega-\gamma)(\tau-\sigma)\}\ \left(\frac{1}{2\pi}\right) e^{j\gamma\tau}\ dC(\sigma,\gamma)\ d\omega\ d\sigma \qquad (2.4\text{-}21)$$

$$= \int_{\sigma'} D_a(\tau,\sigma')\ R_{xx}(\sigma')\ d\sigma'$$

where D_a is the auto-decorrelation function:

$$D_a(\tau,\sigma') = \int_\gamma \exp(j\gamma\tau)\, \frac{dC(\tau-\sigma',\gamma)}{2\pi}$$

$$= \frac{1}{2\pi} \int_\gamma \int_{\sigma''} \exp(j\gamma\tau)\, dB(\sigma'',\sigma''+\tau-\sigma',\gamma)\, d\sigma'' \qquad (2.4\text{-}22)$$

$$= \frac{1}{2\pi} \int_{\sigma''} \int_\gamma \int_\omega \int_{\omega'} \exp\{j[\gamma\tau+\omega\sigma''-\omega'(\sigma''-\tau-\sigma')]\}\, dA(\omega,\omega',\gamma)\, \frac{d\omega}{2\pi}\, \frac{d\omega'}{2\pi}\, d\sigma''$$

Performing first the integral over σ'', then the integral over ω' and finally applying the definition of the tri-frequency function given in Equation (2.4-11) we find

$$\boxed{D_a(\tau,\sigma') = \int_\omega \Phi(\omega,\omega,\tau)\, e^{+j\omega(\tau+\sigma')}\, \frac{d\omega}{2\pi}} \qquad (2.4\text{-}23)$$

Alternatively, the auto-decorrelation function can be written in terms of the second moment of the impulse response:

$$D_a(\tau,\sigma') = \int_\omega \overline{H(\omega,t)\, H^*(\omega,t-\tau)}\, \exp\{+j\omega(\tau+\sigma')\}\, \frac{d\omega}{2\pi}$$

$$= \int_\omega \int_\sigma \int_{\sigma''} \overline{h(\sigma,t)\, h(\sigma'',t-\tau)}\, \exp\{j\omega(\sigma-\sigma''+\tau+\sigma')\}\, \frac{d\omega}{2\pi}\, d\sigma\, d\sigma'' \qquad (2.4\text{-}24)$$

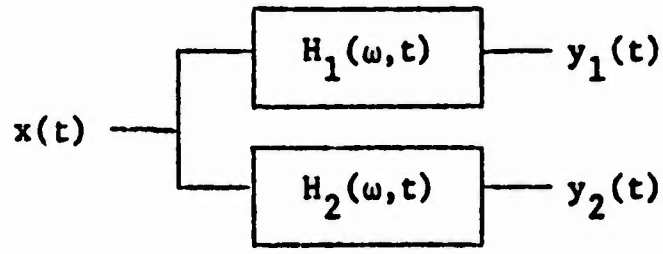
Again, by performing first the integral over ω and then over σ'' this reduces to

$$\boxed{D_a(\tau,\sigma') = \int_\sigma \overline{h(\sigma,t)\, h(\sigma+\tau+\sigma',t-\tau)}\, d\sigma} \qquad (2.4\text{-}25)$$

Equation (2.4-23) is for our purposes more useful than Equation (2.4-25). However, the latter relation is important because it demonstrates that the smearing of the signal correlation function can be noticeable even when the instantaneous impulse response tends to be narrower than the input correlation width. This occurs when the slow system variations tend to translate the instantaneous centroid of the impulse response along the delay axis.

C) Parallel RLTVF Cross Decorrelation

By entirely analogous reasoning we may extend all of the single RLTVF decorrelation equations to the problem of describing the decorrelation of the outputs of two RLTVF's in parallel and excited from a common source (see Figure 2.4-1).



Parallel RLTVF Decorrelation

Figure 2.4-1

Here one distinguishes between ensemble averages carried out over the random parameters of $H_1(\omega,t)$ and $H_2(\omega,t)$ by denoting the former by "s_1" and the latter with "s_2". Again one defines a system correlation function:

$$\Phi_{12}(\omega,\omega',t,t') = \overline{H_1(\omega,t)\ H_2(\omega',t')^*}^{\,s_1 s_2} \qquad (2.4\text{-}26)$$

$$= \Phi_{21}(\omega',\omega,t',t)^*$$

One defines interfrequency wide-sense cross stationarity (IWSCS) by

$$\Phi_{12}(\omega,\omega',t,t') = \Phi_{12}(\omega,\omega',t-t') \qquad (2.4\text{-}27)$$

It is not difficult to see that all of Equations (2.4-5) through (2.4-25) may be rewritten with the appropriate '12' subscripts. The relations of primary interest are summarized here:

$$dZ_{y_1y_2}(\omega) = \frac{1}{2\pi}\int_{\omega'} d_\omega A_{12}(\omega',\omega',\omega-\omega')\, dZ_{xx}(\omega') \qquad (2.4\text{-}28)$$

$$\Phi_{12}(\omega,\omega',\sigma) = \frac{1}{2\pi}\int_{\omega''} \exp(+j\omega''\sigma)\, dA_{12}(\omega,\omega',\omega'') \qquad (2.4\text{-}29)$$

$$R_{y_1y_2}(\tau) = \frac{1}{2\pi}\int_{\omega} \Phi_{12}(\omega,\omega,\tau)\exp(j\omega\tau)\, dZ_{xx}(\omega) \qquad (2.4\text{-}30)$$

$$= \int_{\sigma'} D_{12}(\tau,\sigma')\, R_{xx}(\sigma')\, d\sigma' \qquad (2.4\text{-}31)$$

$$D_{12}(\tau,\sigma) = \int_{\omega} \Phi_{12}(\omega,\omega,\sigma)\exp\{+j\omega(\tau+\sigma)\}\frac{d\omega}{2\pi} \qquad (2.4\text{-}32)$$

$$= \int_{\sigma'} \overline{h_1(\sigma',t)\, h_2(\sigma'+\tau+\sigma,t-\tau)}\, d\sigma' \qquad (2.4\text{-}33)$$

Here $A_{12}(\omega,\omega',\omega'')$ and $D_{12}(\tau,\sigma)$ are termed the cross tri-frequency function and the output cross decorrelation function respectively.

D) Series Cascade Output Auto-Decorrelation

Finally, we examine the problem of cascading RLTVF's. For the two filter cascade shown in Figure 2.4-2 we have from Equation (2.3-3)

$$H_{ab}(\omega,t) = H'_{ab}(\omega,t)\ e^{-j\omega t} = \frac{e^{-j\omega t}}{2\pi}\int_{\gamma} H'_b(\gamma,t)\ dg'_a(\omega,\gamma)$$

$$= \frac{e^{-j\omega t}}{2\pi}\int_{\gamma} H_b(\gamma,t)\ e^{+j\gamma t}\ dg_a(\omega,\gamma-\omega) \qquad (2.4\text{-}34)$$

where Equations (2.2-9) and (2.2-26) are used. From this the system inter-frequency correlation function for the cascade may be computed. Assuming independence between the random parameters of systems a and b we have

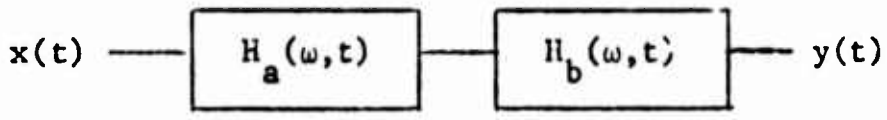
$$\Phi_{ab}(\omega,\omega',\sigma) = \overline{H_{ab}(\omega,t)\ H_{ab}(\omega',t-\sigma)^*} =$$

$$\{(\frac{1}{2\pi})^2 \int_{\gamma}\int_{\gamma'} H_b(\omega,t)\ H_b(\omega',t-\sigma)^*\ \exp\{j(\gamma-\gamma')t+j\gamma'\sigma\} \qquad (2.4\text{-}35)$$

$$\overline{dg_a(\omega,\gamma-\omega)\ dg_a(\omega',\gamma'-\omega')^*}\}\ \exp\{-j(\omega-\omega')t-j\omega'\sigma\}$$

Invoking Equations (2.4-7), (2.4-15) and performing the integration over γ' this becomes

$$\Phi_{ab}(\omega,\omega',\sigma) = \frac{\exp\{-j\omega\sigma\}}{2\pi}\int_{\gamma} \Phi_b(\gamma,\gamma-\omega+\omega',\sigma)\ \exp(+j\gamma\sigma)\ d_\gamma A_a(\omega,\omega',\gamma-\omega) \qquad (2.4\text{-}36)$$



Cascade RLTVF Decorrelation

Figure 2.4-2

The cascade tri-frequency function and auto-decorrelation function are then readily generated. For example, by rewriting Equation (2.4-11) as follows:

$$\phi_b(\gamma, \gamma-\omega+\omega', \sigma) = \frac{1}{2\pi} \int_{\omega''} e^{j(\omega''+\omega-\gamma)\sigma} \, d_{\omega'',A_b}(\gamma, \gamma-\omega+\omega', \omega''+\omega-\gamma) \qquad (2.4\text{-}37)$$

and substituting the result into Equation (2.4-36) we have

$$d_{\omega'',A_{ab}}(\omega,\omega',\omega'') = \frac{1}{2\pi} \int_{\gamma} d_{\gamma A_a}(\omega,\omega',\gamma-\omega) \, d_{\omega'',A_b}(\gamma,\gamma-\omega+\omega',\omega''+\omega-\gamma) \qquad (2.4\text{-}38)$$

Finally, the auto-decorrelation function for the cascade is particularly simple to obtain. Using Equation (2.4-21) we have

$$D_{ab}(\tau,\sigma') = \int_{\sigma''} D_b(\tau,\sigma'') \, D_a(\sigma'',\sigma') \, d\sigma'' \qquad (2.4\text{-}39)$$

where $D_a(\sigma'',\sigma)$ and $D_b(\tau,\sigma'')$ are the auto-decorrelation functions for systems a and b respectively.

Another useful relation which is effectively the dual of Equation (2.4-36) may be obtained by starting with Equation (2.3-4):

$$\begin{aligned} \phi_{ab}(\omega,\omega',\sigma) &= e^{-j(\omega-\omega')t+j\omega'\sigma} \, \overline{H'_{ab}(\omega,t) \, H'_{ab}(\omega',t-\sigma)^*} \\ &= e^{-j(\omega-\omega')t-j\omega'\sigma} \int_{\nu} \int_{\nu'} \overline{H_a(\omega,\nu) \, H_a(\omega',\nu')^*} \, e^{+j(\omega\nu-\omega'\nu')} \\ &\qquad \overline{h_b(t-\nu,t) \, h_b(t-\sigma-\nu',t-\sigma)} \, d\nu \, d\nu' \end{aligned} \qquad (2.4\text{-}40)$$

or

$$\Phi_{ab}(\omega,\omega',\sigma) = e^{-j\omega'\sigma} \int_{\nu''} \Phi_a(\omega,\omega',\nu'')\, e^{+j\omega'\nu''}\, T_b(\omega-\omega',\sigma,\nu'')\, d\nu'' \qquad (2.4\text{-}41)$$

where

$$T_b(\gamma,\sigma,\nu) = \int_{\sigma'} e^{j\gamma\sigma'}\, \overline{h_b(\sigma',t)\, h_b(\sigma'-\sigma+\nu,t-\sigma)}\, d\sigma' \qquad (2.4\text{-}42)$$

might be termed the extended auto-decorrelation function for system b because

$$T_b(0,\tau,\sigma) = D_b(\tau,\sigma) \qquad (2.4\text{-}43)$$

Rewritten in the frequency domain Equation (2.4-42) becomes

$$T_b(\gamma,\sigma,\nu) = \left(\frac{1}{2\pi}\right)^2 \int_{\sigma'}\int_{\gamma'}\int_{\gamma''} e^{j\gamma\sigma'}\, \Phi_b(\gamma',\gamma'',\nu)$$

$$\exp\{j\gamma'\sigma' - j\gamma''(\sigma'+\sigma+\nu)\}\, d\sigma'\, d\gamma'\, d\gamma'' \qquad (2.4\text{-}44)$$

$$= \int_{\gamma''} \Phi_b(\gamma''-\gamma,\gamma'',\nu)\, e^{j\gamma''(\nu+\sigma)}\, \frac{d\gamma''}{2\pi}$$

where the last equality is obtained by first integrating over σ' and then over γ'. The relative usefulness of the dual Equations (2.4-41) and (2.4-36) depends on the ease of integration.

2.5 Time Invariant and Slowly Varying Systems

Before closing this chapter we simplify various quantities derived in sections 2.2 and 2.4 for the case of time-invariance. First we have the specialized versions of the eight fundamental system functions:

$$\text{a)}\quad h(\sigma,t) = h_f(\sigma) \qquad \text{b)}\quad h'(\nu,t) = h_f(\nu-t) \tag{2.5-1}$$

$$\text{c)}\quad H(\omega,t) = H_f(\omega) \qquad \text{d)}\quad H'(\omega,t) = H_f(\omega)\, e^{+j\omega t}$$

$$\text{e)}\quad dG(\sigma,\gamma) = h_f(\sigma)\, \delta(\gamma)\, d\gamma \qquad \text{f)}\quad dG'(\nu,\gamma) = H_f(\gamma)\, e^{-j\gamma\nu}\, d\gamma$$

$$\text{g)}\quad dg(\omega,\gamma) = 2\pi\, \delta(\gamma)\, H_f(\omega)\, d\gamma \qquad \text{h)}\quad dg'(\omega,\gamma) = 2\pi\, \delta(\gamma-\omega)\, H_f(\omega)\, d\gamma$$

where the "fixed" or static impulse response is related to its transfer function by the usual formula

$$h_f(\sigma) = \int_{\omega} H_f(\omega)\, \exp(+j\omega\sigma)\, \frac{d\omega}{2\pi} \tag{2.5-2}$$

Thus, Equations (2.5-1) (e) through (h) show that there exists spreading only along the "delay" axis and no frequency "leakage" when the system is time invariant. For a very slowly time-varying system the frequency smear therefore tends to be narrow.

For the completely time invariant system the input-output relations reduce to the familiar forms:

$$\text{a)}\quad y(t) = \int_{\sigma} h_f(\sigma)\, x(t-\sigma)\, d\sigma \qquad \text{b)}\quad y(t) = \int_{\omega} H_f(\omega)\, e^{j\omega t}\, dz_x(\omega) \tag{2.5-3}$$

The stochastic system correlation function and distributions are reduced for time invariance as follows:

a) $$\Phi(\omega,\omega',\sigma) = \Phi_f(\omega,\omega') = \overline{H_f(\omega)\ H_f(\omega')^*}$$

b) $$dA(\omega,\omega',\omega'') = \Phi_f(\omega,\omega')\ \delta(\omega'')\ d\omega''$$

c) $$dB(\sigma,\sigma',\gamma) = \overline{h_f(\sigma)\ h_f(\sigma')}\ \delta(\gamma)\ d\gamma$$

d) $$T(\gamma,\tau,\sigma) = \int_{\sigma'} e^{j\gamma\sigma'}\ \overline{h_f(\sigma')\ h_f(\sigma'+\tau+\sigma)}\ d\sigma' = T_f(\gamma,\sigma+\tau) \qquad (2.5\text{-}4)$$

e) $$D_a(\tau,\sigma) = T_f(0,\sigma+\tau)$$

f) $$dC(\sigma'',\gamma) = T_f(0,\sigma'')\ \delta(\gamma)\ d\gamma$$

For purely non-random filters the overbar is deleted.

From Equation (2.5-4b) it is seen that in the limit of very slowly time-varying systems the tri-frequency distribution weights the point $\omega''=0$ very heavily. Thus, in Equation (2.4-36) the distribution A_a for the first system of a two system cascade weights the point $\gamma=\omega$ heavily. Therefore, provided the function $\Phi_b(\gamma,\gamma-\omega+\omega',\sigma)$ is suitably smooth in the vicinity of this point we have the approximate result that

$$\boxed{\Phi_{ab}(\omega,\omega',\sigma) \rightarrow \Phi_a(\omega,\omega',\sigma)\ \Phi_b(\omega,\omega',\sigma)} \qquad (2.5\text{-}5)$$

In this limit, of course, the order of the systems in the chain may be reversed. In fact, to the same limit we have

$$H_{ab}(\omega,t) \rightarrow H_a(\omega,t)\ H_b(\omega,t) \qquad (2.5\text{-}6)$$

CHAPTER 3

STATISTICS OF THE

FINITE TIME CORRELATOR OUTPUT

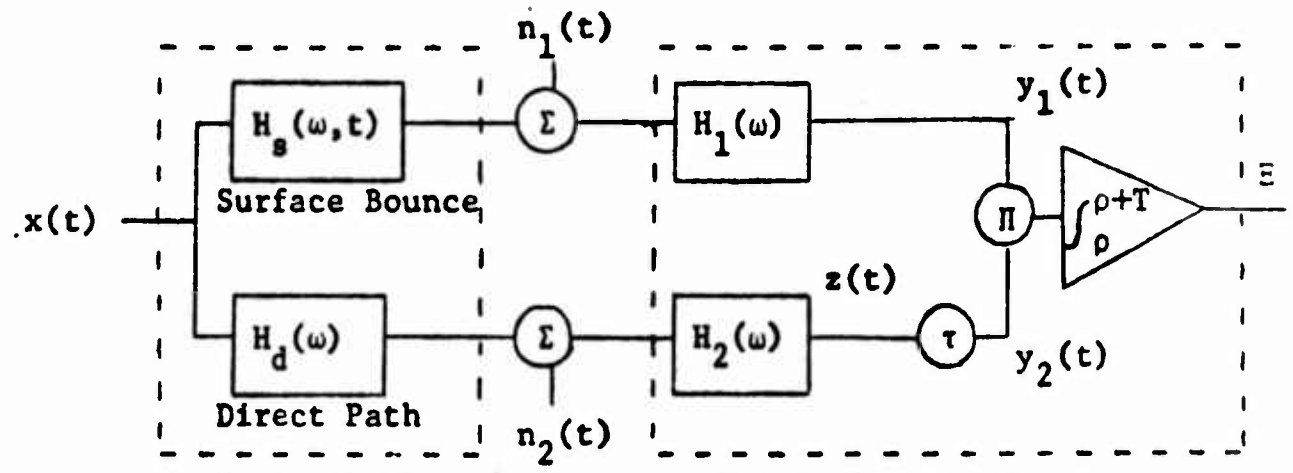
3.0 Introduction

In this chapter practical multipath and array correlator detector-trackers are considered. The mean and variance of the correlator output are determined. The correlator input consists of signals which may be degraded by surface scattering and additive, Gaussian, stationary background noise. The signals are assumed to be Gaussian and stationary when they are emitted at the target. The scattering is represented by passage of the signal through an interfrequency wide-sense stationary RLTVF. Although the IWSS assumption guarantees that the scattered signals retain their stationarity, it is not safe to assume that they remain Gaussian processes. The departure from Gaussian statistics is examined.

It is shown that correlator output exhibits fluctuations that include contributions due to the noise, to the randomness of the target signal, and to the time-variation and randomness of the scattering. These fluctuations are termed estimation noise since the output of the correlator operating during the finite interval $[0,T]$ is only an estimate of the ergodic mean value obtained as $T \to \infty$. Of the three components of the output variance, the fluctuation due to the slow time-variation of the scattering is by far the most persistent. This component occasionally forces the use of extremely long processing intervals. Expressions for the error probabilities of a two sided test on correlator output are derived.

3.1 Cross-correlation of Direct and Surface Reflected Paths

In this section we consider a propagation geometry in which both direct and surface reflected paths are received by a single receiver (see Figure 1.0-1a). It is assumed that each path can be processed separately by the receiver. This might be accomplished by using multi-directional sensors, but the exact nature of the separation will not concern us at this point. A block diagram for the situation is shown in Figure 3.1-1.



Direct vs. Surface Reflected

Cross-correlator Processing

Figure 3.1-1

The two channels are assumed to be corrupted by independent noises $n_1(t)$ and $n_2(t)$. The filters $H_1(\omega)$ and $H_2(\omega)$ have been included in order to provide pre-processing before correlation. The variable delay parameter τ has been included as a "search" parameter for estimating the multipath replication delay. The direct path is assumed to transmit a delayed and attenuated version of the emitted signal

$$H_d(\omega) = \frac{e^{-j\omega R_d/c}}{R_d} \qquad (3.1\text{-}1)$$

where R_d is the line-of-sight distance from target to receiver. Following Equation (2.4-34) we define the cascade responses

$$H_{s1}(\omega,t) = \frac{e^{-j\omega t}}{2\pi} \int_{\gamma} H_1(\gamma)\ e^{j\gamma t}\ dg_s(\omega,\gamma-\omega) \tag{3.1-2}$$

$$H_{d2}(\omega) = H_d(\omega)\ H_2(\omega) \tag{3.1-3}$$

The correlator operates during the interval $[\rho,\rho+T]$ to yield the output $\Xi(\tau,T,\rho)$ at time $\rho+T$. Since the correlator acquires random data for only a finite time T, Ξ is a random variable. Its mean is easily computed (using Equation (2.2-10)):

$$\overline{\Xi(\tau,T,\rho)} = \frac{1}{T}\int_{\rho}^{\rho+T} \overline{y_1(t)\ y_2(t)^*}\ dt =$$

$$\frac{1}{T}\int_{\rho}^{\rho+T} \{\frac{1}{2\pi}\int_{\omega} \overline{[H_1(\omega)\ dz_{n1}(\omega) + H_{s1}(\omega,t)\ dz_x(\omega)]\ \exp(j\omega t)}$$

$$\overline{\{\frac{1}{2\pi}\int_{\omega'} [H_2(\omega')\ dz_{n2}(\omega') + H_{d2}(\omega')\ dz_x(\omega')]\ \exp\{j\omega'(t-\tau)\}\}^*}\ dt \tag{3.1-4}$$

On expanding the product:

$$\overline{\Xi(\tau,T,\rho)} = \frac{1}{T}\int_{\rho}^{\rho+T} \{(\frac{1}{2\pi})^2 \int_{\omega}\int_{\omega'} [H_1(\omega)\ [H_2(\omega')^*\ \overline{dz_{n1}(\omega)\ dz_{n2}(\omega')^*}$$

$$+ H_{d2}(\omega')^*\ \overline{dz_{n1}(\omega)\ dz_x(\omega')^*}] + \overline{H_{s1}(\omega,t)}\ [H_2(\omega')^*\ \overline{dz_x(\omega)\ dz_{n2}(\omega')^*}$$

$$+ H_{d2}(\omega')^*\ \overline{dz_x(\omega)\ dz_x(\omega')^*}]]\ \exp\{j(\omega-\omega')t+j\omega'\tau\}\}\ dt \tag{3.1-5}$$

Invoking the independence of $n_1(t)$, $n_2(t)$ and $x(t)$ and applying Equation (2.1-5) this becomes

$$\overline{\Xi(\tau,T,\rho)} = \left[\frac{1}{T}\int_{\rho}^{\rho+T} dt\right]\left[\int_{\omega} H_d(\omega)^* H_2(\omega)^* \overline{H_{s1}(\omega,t)}\,\frac{e^{j\omega\tau}}{2\pi}\,dZ_{xx}(\omega)\right] \tag{3.1-6}$$

The mean of $H_{s1}(\omega,t)$ can be obtained by averaging Equation (3.1-2) and applying the following relation (which may be verified by insertion in Equation (2.2-22))

$$\overline{dg_s(\omega,\gamma-\omega)} = 2\pi\, H_c(\omega)\,\delta(\gamma-\omega)\,d\gamma \tag{3.1-7}$$

with the result that

$$\overline{H_{s1}(\omega,t)} = H_c(\omega)\,H_1(\omega) \tag{3.1-8}$$

Finally, substituting this into Equation (3.1-6) we have

$$\boxed{\overline{\Xi(\tau,T,\rho)} = \int_{\omega} H_c(\omega)\,H_d(\omega)^*\,H_1(\omega)\,H_2(\omega)^*\,\frac{e^{+j\omega\tau}}{2\pi}\,dZ_{xx}(\omega)} \tag{3.1-9}$$

3.2 Second Order Statistics for the Output of the Cross-correlator

In this section we regard the output $\Xi(\tau,T,\rho)$ of the cross-correlator shown in Figure (3.1-1) as a time series which is a random function of the initial starting time ρ for the integrator. This output fluctuates about its mean, causing uncertainty about the possibility of the existence of target signal correlation. The most general second order statistic that is of interest in analysis of this fluctuation is the cross-covariance of Ξ itself at two values τ and τ' for the replication delay parameter and for two values ρ and ρ' for the starting time:

$$\Delta(\tau,\tau',T,\rho,\rho') = \overline{\Xi(\tau,T,\rho)\ \Xi(\tau',T,\rho')} - \overline{\Xi(\tau,T,\rho)}\ \overline{\Xi(\tau',T,\rho')} \qquad (3.2\text{-}1)$$

This quantity determines the spectrum of the output fluctuation, the persistence and magnitude of these fluctuations, and the degree of dependence of fluctutations at different values of the replication delay parameter.

In this section we focus attention on the computation of the first term in Equation (3.2-1). The method used is essentially that given by Davenport and Root,[54] Laning and Battin,[55] and Bendat[56] (see also Usher[57]). From the definition of Ξ we have

$$\overline{\Xi(\tau,T,\rho)\ \Xi(\tau',T,\rho')} = \left(\frac{1}{T}\right)^2 \int_{\rho}^{\rho+T} \int_{\rho'}^{\rho'+T} \Pi(\tau,\tau',t,t')\ dt\ dt' \qquad (3.2\text{-}2)$$

where we define

$$\Pi(\tau,\tau',t,t') = \overline{y_1(t)\ y_1(t')\ z(t-\tau)\ z(t'-\tau')} \qquad (3.2\text{-}3)$$

At this point it is assumed that this fourth order moment is stationary, that is, we assume that the relation

$$\Pi(\tau,\tau',t,t') = \Pi(\tau,\tau',t-t') \tag{3.2-4}$$

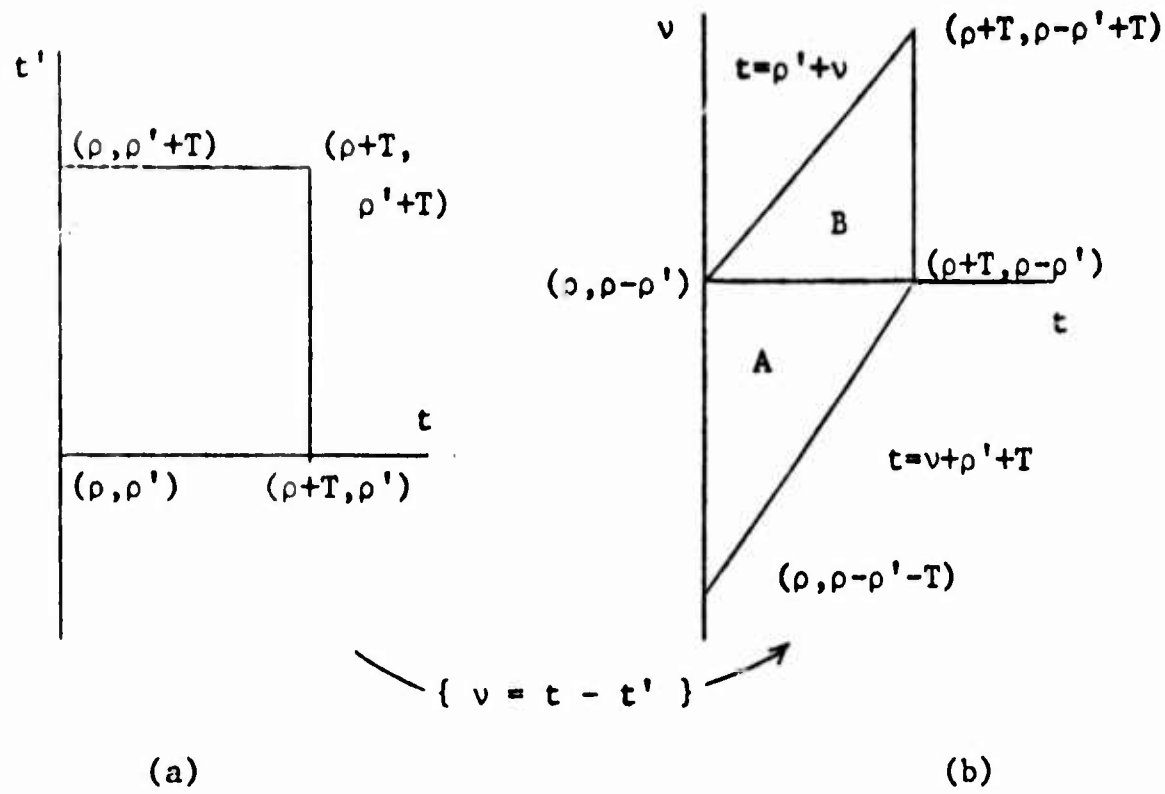
holds. We shall return to this assumption in sections 3.3 and 3.5 to determine the conditions under which Equation (3.2-4) may be applied. When Equation (3.2-4) is valid it suggests the change of variable

$$\begin{aligned} \nu &= t-t' \\ d\nu &= -dt' \end{aligned} \tag{3.2-5}$$

Substituting this into Equation (3.2-2) we have

$$\overline{\Xi(\tau,T,\rho)\ \Xi(\tau',T,\rho')} = \left(\frac{1}{T}\right)^2 \int_{\rho}^{\rho+T} \left\{ \int_{t-\rho'-T}^{t-\rho'} \Pi(\tau,\tau',\nu)\ d\nu \right\} dt \tag{3.2-6}$$

Next, noting that the integration over t is between fixed limits while the integrand is a function of only ν we seek to reverse this situation by interchanging the order of integration. Unfortunately, this forces us to break the integral up into two terms, one for the range $\nu < \rho - \rho'$ (area A in Figure 3.2-1b) and the other for the range $\nu > \rho - \rho'$ (area B)



Regions of Integration

Before (a) and After (b) Transformation

Figure 3.2-1

Executing the interchange of order we have:

$$\overline{\Xi(\tau,T,\rho)\ \Xi(\tau',T,\rho')} =$$

$$\left(\frac{1}{T}\right)^2 \left\{ \int_{\rho-\rho'-T}^{\rho-\rho'} d\nu \int_{\rho}^{\nu+\rho'+T} dt\ \Pi(\tau,\tau',\nu) + \right. \qquad (3.2\text{-}7)$$

$$\left. \int_{\rho-\rho'}^{\rho-\rho'+T} d\nu \int_{\nu+\rho'}^{\rho+T} dt\ \Pi(\tau,\tau',\nu) \right\}$$

The integration over t is now easily performed giving the result

$$\overline{\Xi(\tau,T,\rho)\ \Xi(\tau',T,\rho')} =$$

$$\left(\frac{1}{T}\right)^2 \left\{ \int_{\rho-\rho'-T}^{\rho-\rho'} [T+\nu-(\rho-\rho')]\ \Pi(\tau,\tau',\nu)\ d\nu + \right. \tag{3.2-8}$$

$$\left. \int_{\rho-\rho'}^{\rho-\rho'+T} [T-\nu+(\rho-\rho')]\ \Pi(\tau,\tau',\nu)\ d\nu \right\}$$

$$= \frac{1}{T}\left[V_1(\rho-\rho'+T) - V_1(\rho-\rho'-T)\right]$$

$$- \left(\frac{1}{T}\right)^2 \left[V_2(\rho-\rho'-T) - 2\,V_2(\rho-\rho') + V_2(\rho-\rho'+T)\right]$$

$$+ \frac{(\rho-\rho')}{T^2}\left[V_1(\rho-\rho'-T) - 2\,V_1(\rho-\rho') + V_1(\rho-\rho'+T)\right]$$

which is a function of $(\rho-\rho')$. Here we define

$$\text{a)}\quad V_1(\xi) = \int_0^{\xi} \Pi(\tau,\tau',\nu)\ d\nu \qquad \text{b)}\quad V_2(\xi) = \int_0^{\xi} \nu\ \Pi(\tau,\tau',\nu)\ d\nu \tag{3.2-9}$$

Equation (3.2-8) simplifies greatly for the case $\rho = \rho'$ and $\tau = \tau'$. In this case we may interchange t and t' in Equation (3.2-3) yielding

$$\Pi(\tau,\ \tau,\nu) = \Pi(\tau,\ \tau,-\nu) \tag{3.2-10}$$

Applying this to the first integral of Equation (3.2-8) we obtain the well-known result

$$\overline{\Xi(\tau,T,\rho)\ \Xi(\tau,T,\rho)} = 2\left(\frac{1}{T}\right)^2 \int_0^T [T-\nu]\ \Pi(\tau,\tau,\nu)\ d\nu \tag{3.2-11}$$

$$= 2\left[\frac{1}{T}\,V_1(T) - \left(\frac{1}{T}\right)^2 V_2(T)\right]$$

3.3 Correlator Fluctuation for Direct and Surface Reflected Multipath Processor

In this section we compute the fourth order product function $\Pi(\tau,\tau',t,t')$ defined in Equation (3.2-3) for the direct vs. surface reflected path processor shown in Figure 3.1-1. Rewriting Equation (3.2-3) in the frequency domain proves to be useful in later computations. Thus, we have from Equation (3.2-3)

$$\Pi(\tau,\tau',t,t') =$$

$$\left(\frac{1}{2\pi}\right)^4 \left\{ \int_{\omega} \int_{\omega'} \int_{\omega''} \int_{\omega'''} e^{j(\omega+\omega'')t-j(\omega'+\omega''')t'-j(\omega''\tau-\omega'''\tau')} \left[\right.\right.$$

$$\left[H_1(\omega)\, dz_{n1}(\omega) + H_{s1}(\omega,t)\, dz_x(\omega) \right]$$

$$\left[H_1(\omega')\, dz_{n1}(\omega') + H_{s1}(\omega',t)\, dz_x(\omega') \right]^*$$

$$\left[H_2(\omega'')\, dz_{n2}(\omega'') + H_{d2}(\omega'')\, dz_x(\omega'') \right] \qquad (3.3\text{-}1)$$

$$\left.\left. \left[H_2(\omega''')\, dz_{n2}(\omega''') + H_{d2}(\omega''')\, dz_x(\omega''') \right]^* \right] \right\}$$

On expanding the product of the four factors in Equation (3.3-1) we obtain 16 terms. By the mutual independence and zero mean of $n_1(t)$, $n_2(t)$ and $x(t)$ all but 4 of these terms are found to be zero. The remaining terms are as follows:

$$\Pi(\tau,\tau',t,t') =$$

$$\left(\frac{1}{2\pi}\right)^4 \{ \int_{\omega} \int_{\omega'} \int_{\omega''} \int_{\omega'''} e^{j(\omega+\omega'')t-j(\omega'+\omega''')t'-j(\omega''\tau-\omega'''\tau')}$$

$$H_1(\omega)\, H_1(\omega')^*\, \overline{dz_{n1}(\omega)\, dz_{n1}(\omega')^*}\, [$$

$$H_2(\omega'')\, H_2(\omega''')^*\, \overline{dz_{n2}(\omega'')\, dz_{n2}(\omega''')^*} +$$

$$H_{d2}(\omega'')\, H_{d2}(\omega''')^*\, \overline{dz_x(\omega'')\, dz_x(\omega''')^*}\,] \qquad (3.3\text{-}2)$$

$$+ \overline{H_{s1}(\omega,t)\, H_{s1}(\omega',t')^*}\, \Big[$$

$$H_2(\omega'')\, H_2(\omega''')^*\, \overline{dz_x(\omega)\, dz_x(\omega')^*}\, \overline{dz_{n2}(\omega'')\, dz_{n2}(\omega''')^*}$$

$$+ H_{d2}(\omega'')\, H_{d2}(\omega''')^*\, \overline{dz_x(\omega)\, dz_x(\omega')^*\, dz_x(\omega'')\, dz_x(\omega''')^*}\, \Big] \}$$

We now make the following assumptions:

$$\left.\begin{array}{l} \text{(1) The noises } n_1(t) \text{ and } n_2(t) \text{ are wide sense stationary.} \\ \\ \text{(2) The signal } x(t) \text{ emitted at the target is a Gaussian process and stationary.} \\ \\ \text{(3) The scattering RLTVF is interfrequency-wide sense stationary (IWSS).} \end{array}\right\} \qquad (3.3\text{-}3)$$

Therefore, using Equation (2.1-8) and (2.1-18)

$$\Pi(\tau,\tau',t,t') =$$

$$\left(\frac{1}{2\pi}\right)^2 \int_{\omega} \int_{\omega'} \int_{\omega''} \int_{\omega'''} e^{j(\omega+\omega'')t - j(\omega'+\omega''')t' - j(\omega''\tau - \omega'''\tau')} \Big($$

$$|H_1(\omega)|^2 S_{n_1 n_1}(\omega)\ \delta(\omega-\omega')\ \delta(\omega''-\omega''')\ [$$

$$|H_2(\omega'')|^2 S_{n_2 n_2}(\omega'') + |H_{d2}(\omega'')|^2 S_{xx}(\omega'')\ ]$$

$$+ \Phi_{s1}(\omega,\omega',t-t') \Big[\qquad (3.3\text{-}4)$$

$$|H_2(\omega'')|^2 S_{xx}(\omega)\ S_{n_2 n_2}(\omega'')\ \delta(\omega-\omega')\ \delta(\omega''-\omega''')$$

$$+ H_{d2}(\omega'')\ H_{d2}(\omega''')^*\ [$$

$$S_{xx}(\omega)\ \delta(\omega-\omega')\ S_{xx}(\omega'')\ \delta(\omega''-\omega''') +$$

$$S_{xx}(\omega)\ \delta(\omega+\omega'')\ S_{xx}(\omega''')\ \delta(\omega'''+\omega') +$$

$$S_{xx}(\omega)\ \delta(\omega-\omega''')\ S_{xx}(\omega')\ \delta(\omega'-\omega'')\ ] \Big] \Big)\ d\omega\ d\omega'\ d\omega''\ d\omega'''$$

where

$$\Phi_{s1}(\omega,\omega',\mu) = \overline{H_{s1}(\omega,t)\ H_{s1}(\omega',t-\mu)} =$$

$$\frac{e^{-j\omega\mu}}{2\pi} \int_{\gamma} e^{j\gamma\mu}\ H_1(\gamma)\ H_1^*(\gamma-\omega+\omega')\ d_\gamma A_s(\omega,\omega',\gamma-\omega) \qquad (3.3\text{-}5)$$

is the interfrequency system correlation function for the cascade of the scattering and filter #1. Separating the last two terms of Equation (3.3-4) and performing two integrals over frequency for all terms we have:

$$\Pi(\tau,\tau',t,t') =$$

$$\left(\frac{1}{2\pi}\right)^2 \{ \int_{\omega} \int_{\omega''} e^{j(\omega+\omega'')\ (t-t')-j(\tau-\tau')\omega''}\ |H_2(\omega'')|^2 \Big($$

$$[|H_1(\omega)|^2\ S_{n_1n_1}(\omega) + \Phi_{s1}(\omega,\omega,t-t')\ S_{xx}(\omega)]$$

$$[S_{n_2n_2}(\omega'') + |H_d(\omega'')|^2\ S_{xx}(\omega'')] \Big)\ d\omega\ d\omega'' \} \qquad (3.3\text{-}6)$$

$$+ \left(\frac{1}{2\pi}\right)^2 \{ \int_{\omega} \int_{\omega'} \Phi_{s1}(\omega,\omega',t-t')\ H_{d2}(\omega)^*\ H_{d2}(\omega')\ S_{xx}(\omega)\ S_{xx}(\omega')$$

$$[\ e^{+j(\omega\tau-\omega'\tau')} + e^{+j(\omega+\omega')\ (t-t')-j(\omega'\tau-\omega\tau')}\ ]\ d\omega\ d\omega'\}$$

Thus, it can be seen that when the assumptions (3.3-3) are satisfied Equation (3.2-4) does indeed hold. Finally, from Equation (3.2-1) and taking $\mu = \rho-\rho')$ we have

$$\begin{aligned}\Delta(\tau,\tau',T,\rho,\rho') &= \Delta(\tau,\tau',T,\rho-\rho') \\ &= \Delta(\tau,\tau',T,\mu)\end{aligned} \qquad (3.3\text{-}7)$$

Substituting Equation (3.3-6) into (3.2-8) we may rewrite Equation (3.2-1) in summary form:

$$\Delta(\tau,\tau',T,\mu) = \left(\frac{1}{T}\right)^2 \{ \Big[\int_{\mu-T}^{\mu} [T+\nu-\mu] + \int_{\mu}^{\mu+T} [T-\nu+\mu] \Big]$$

$$\{[I_1(\nu) + I_2(\nu)]\ I_3(\nu-\tau+\tau') + I_4(\tau,\tau',\nu) + \qquad (3.3\text{-}8)$$

$$I_4(\tau'+\nu,\tau-\nu,\nu)\}\ d\nu\} - I_5(\tau,\tau')$$

where:

$$I_1(\nu) = \int_\omega e^{j\omega\nu} \, |H_1(\omega)|^2 \, S_{n_1 n_1}(\omega) \, \frac{d\omega}{2\pi} \tag{3.3-9}$$

$$I_2(\nu) = \int_\omega e^{j\omega\nu} \, \Phi_{s1}(\omega,\omega,\nu) \, S_{xx}(\omega) \, \frac{d\omega}{2\pi} \tag{3.3-10}$$

$$I_3(\nu) = \int_\omega e^{j\omega\nu} \, |H_2(\omega)|^2 \, [S_{n_2 n_2}(\omega) + |H_d(\omega)|^2 \, S_{xx}(\omega)] \, \frac{d\omega}{2\pi} \tag{3.3-11}$$

$$I_4(\tau,\tau',\nu) = \int_\omega \int_{\omega'} e^{j(\omega\tau - \omega'\tau')} \, \Phi_{s1}(\omega,\omega',\nu) \, H_{d2}(\omega)^* \, H_{d2}(\omega') \, S_{xx}(\omega) \, S_{xx}(\omega') \, \frac{d\omega}{2\pi} \, \frac{d\omega'}{2\pi} \tag{3.3-12}$$

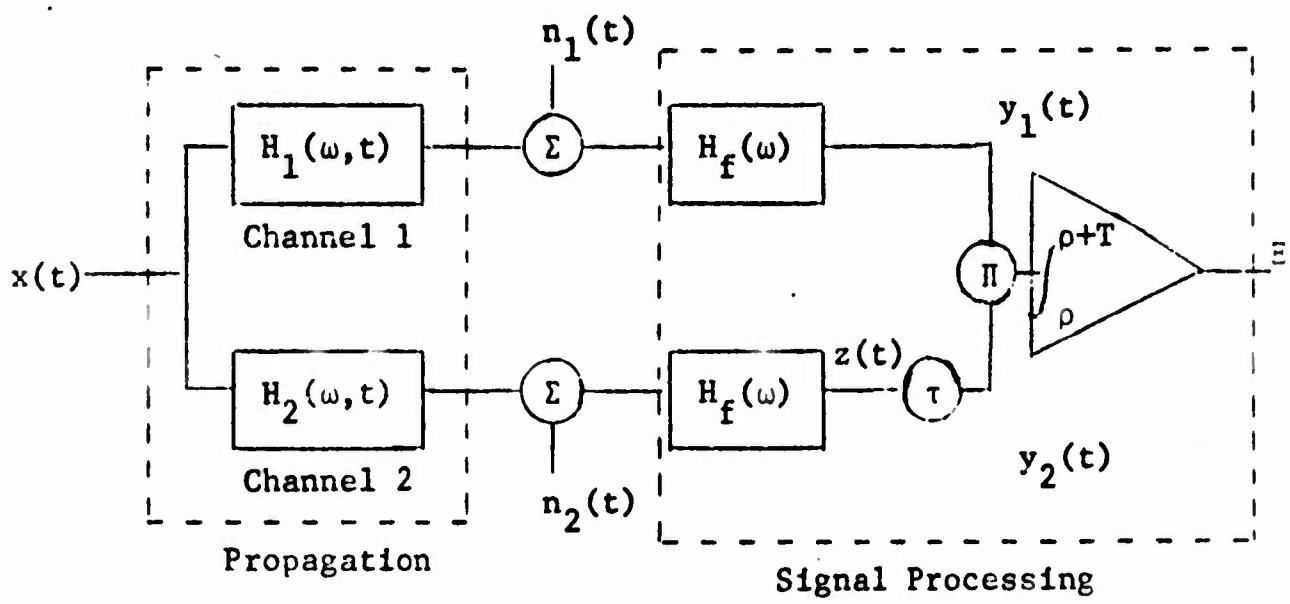
$$I_5(\tau,\tau') = \overline{\Xi(\tau,T,\rho)} \; \overline{\Xi(\tau',T,\rho')} =$$

$$\left[\int_\omega H_c(\omega) \, H_d(\omega)^* \, H_1(\omega) \, H_2(\omega)^* \, e^{+j\omega\tau} \, S_{xx}(\omega) \, \frac{d\omega}{2\pi} \right] \tag{3.3-13}$$

$$\left[\int_{\omega'} H_c(\omega')^* \, H_d(\omega') \, H_1(\omega')^* \, H_2(\omega') \, e^{-j\omega'\tau'} \, S_{xx}(\omega') \, \frac{d\omega'}{2\pi} \right]$$

3.4 Two Receiver Array Cross-correlator Processing

In this section we apply the same technique used in section 3.1 to the problem of two channel array cross-correlation. In this case the propagation geometry for the two channels is assumed similar in nature and statistically correlated. This geometry might consist only of a pair of surface reflected paths (see Figure 1.0-1b), or for a slight increase in complexity, one might include direct transmission (Figure 1.0-1c). The geometry is, in fact, arbitrary for the purposes of this section with the exception of the assumption of statistical symmetry between the two channels. A block diagram for this processor is shown in Figure 3.4-1.



Two Receiver Array

Cross-correlator

Figure 3.4-1

As a special case, when the two receivers are drawn together so that they coalesce, Ξ becomes an auto-correlation estimate.

For the analysis of this section $n_1(t)$ and $n_2(t)$ are <u>not</u> assumed to be independent. We define the cascade responses

$$H_{1f}(\omega,t) = \frac{e^{-j\omega t}}{2\pi} \int H_f(\gamma)\ e^{j\gamma t}\ dz_1(\omega,\gamma-\omega) \qquad (3.4\text{-}1)$$

$$H_{2f}(\omega,t) = \frac{e^{-j\omega t}}{2\pi} \int H_f(\gamma)\ e^{j\gamma t}\ dg_2(\omega,\gamma-\omega) \qquad (3.4\text{-}2)$$

Following Equation (3.1-4) we compute the mean of Ξ

$$\overline{\Xi(\tau,T,\rho)} =$$

$$\frac{\frac{1}{T}\int_{\rho}^{\rho+T} \{\frac{1}{2\pi}\int_{\omega} \overline{[H_f(\omega)\ dz_{n1}(\omega) + H_{1f}(\omega,t)\ dz_x(\omega)]\ \exp(j\omega t)\}}}{\{\frac{1}{2\pi}\int_{\omega'} [H_f(\omega')\ dz_{n2}(\omega') + H_{2f}(\omega',t-\tau)\ dz_x(\omega')]\ \exp\{j\omega'(t-\tau)\}\}^*\ dt} \qquad (3.4\text{-}3)$$

$$= \frac{1}{T}\int_{\rho}^{\rho+T} \{(\frac{1}{2\pi})^2 \int_{\omega}\int_{\omega'} [H_f(\omega)\ [H_f(\omega')^*\ \overline{dz_{n1}(\omega)\ dz_{n2}(\omega')^*} +$$

$$\overline{H_{2f}(\omega',t-\tau)^*\ dz_{n1}(\omega)\ dz_x(\omega')^*}] + \overline{H_{1f}(\omega,t)}\ [H_f(\omega')^*\ \overline{dz_x(\omega)\ dz_{n2}(\omega')^*}]$$

$$+ \overline{H_{1f}(\omega,t)\ H_{2f}(\omega',t-\tau)^*\ dz_x(\omega)\ dz_x(\omega')^*}]\ \exp\{j(\omega-\omega')t+j\omega'\tau\}\}^*\ dt$$

Once again, we assume that the signal $x(t)$ is independent from the background noise in each of the two channels. Also, we assume wide-sense stationary signals and noises, and wide-sense stationary (WSS) scattering. Allowing for the dependence of $n_1(t)$ and $n_2(t)$ we have

$$\overline{\Xi(\tau,T,\rho)} = \qquad (3.4\text{-}4)$$

$$\int_{\omega} [\ [|H_f(\omega)|^2\ dZ_{n_1n_2}(\omega) + \Phi_{1_f2_f}(\omega,\omega,\tau)\ dZ_{xx}(\omega)]\ \frac{e^{j\omega\tau}}{2\pi}\ ]$$

where the parallel system cross-correlation function for the series cascade of scattering and filters is obtained in the same manner as Equation (2.4-36):

$$\Phi_{1_f 2_f}(\omega,\omega,\tau) = \overline{H_{1f}(\omega,t)\ H_{2f}(\omega,t-\tau)^*} = \tag{3.4-5}$$

$$\frac{e^{-j\omega\tau}}{2\pi} \int_{\gamma} H_f(\gamma)\ H_f(\gamma)^*\ e^{j\gamma\tau}\ dA_{12}(\omega,\omega,\gamma-\omega)$$

3.5 Cross-correlator Fluctuation for the Two Receiver Array

We now apply the analysis of section 3.3 to the problem of describing the correlator fluctuation for the two receiver array. The equivalent of Equation (3.3-1b) for this case is

$$\Pi(\tau,\tau',t,t') =$$

$$\left(\frac{1}{2\pi}\right)^4 \{ \int_{\omega} \int_{\omega'} \int_{\omega''} \int_{\omega'''} e^{j(\omega+\omega'')t - j(\omega'+\omega''')t' - j(\omega''\tau - \omega'''\tau')}$$

$$\overline{[\, H_f(\omega)\, dz_{n1}(\omega) + H_{1f}(\omega,t)\, dz_x(\omega)\,]}$$

$$\overline{[\, H_f(\omega')\, dz_{n1}(\omega') + H_{1f}(\omega',t')\, dz_x(\omega')\,]^*} \qquad (3.5\text{-}1)$$

$$\overline{[\, H_f(\omega'')\, dz_{n2}(\omega'') + H_{2f}(\omega'',t-\tau)\, dz_x(\omega'')\,]}$$

$$\overline{[\, H_f(\omega''')\, dz_{n2}(\omega''') + H_{2f}(\omega''',t'-\tau')\, dz_x(\omega''')\,]^*}$$

Once again, on expanding this product we obtain sixteen terms. For this case, however, we are assuming that $n_1(t)$ and $n_2(t)$ are partially correlated. Nevertheless, the signal $x(t)$ remains independent of the noises. Therefore, of the sixteen terms in the product all but eight of these are still found to be zero. Equation (3.5-1) becomes

$$\Pi(\tau,\tau',t,t') = \left(\frac{1}{2\pi}\right)^4 \{ \int_{\omega} \int_{\omega'} \int_{\omega''} \int_{\omega'''} e^{j(\omega+\omega'')t - j(\omega'+\omega''')t' - j(\omega''\tau - \omega'''\tau')}$$

$$[H_f(\omega)\ H_f(\omega')^*\ H_f(\omega'')\ H_f(\omega''')^*$$

$$\overline{dz_{n1}(\omega)\ dz_{n1}(\omega')^*\ dz_{n2}(\omega'')\ dz_{n2}(\omega''')^*}\ ] +$$

$$[H_f(\omega)\ H_f(\omega')^*\ \overline{H_{2f}(\omega'',t-\tau)\ H_{2f}(\omega''',t'-\tau')^*}$$

$$\overline{dz_{n1}(\omega)\ dz_{n1}(\omega')^*}\ \overline{dz_x(\omega'')\ dz_x(\omega''')^*}\ ] +$$

$$[H_f(\omega)\ H_f(\omega'')\ \overline{H_{1f}(\omega',t')^*\ H_{2f}(\omega''',t'-\tau')^*}$$

$$\overline{dz_{n1}(\omega)\ dz_{n2}(\omega'')}\ \overline{dz_x(\omega')^*\ dz_x(\omega''')^*}\ ] +$$

$$[H_f(\omega)\ H_f(\omega''')^*\ \overline{H_{1f}(\omega',t')^*\ H_{2f}(\omega'',t-\tau)}$$

$$\overline{dz_{n1}(\omega)\ dz_{n2}(\omega''')^*}\ \overline{dz_x(\omega')^*\ dz_x(\omega'')}\ ] +$$

(3.5-2)

$$[H_f(\omega')^*\ H_f(\omega'')\ \overline{H_{1f}(\omega,t)\ H_{2f}(\omega''',t'-\tau')^*}$$

$$\overline{dz_{n1}(\omega')^*\ dz_{n2}(\omega'')}\ \overline{dz_x(\omega)\ dz_x(\omega''')^*}\ ] +$$

$$[H_f(\omega')^*\ H_f(\omega''')^*\ \overline{H_{1f}(\omega,t)\ H_{2f}(\omega'',t-\tau)}$$

$$\overline{dz_{n1}(\omega')^*\ dz_{n2}(\omega''')^*}\ \overline{dz_x(\omega)\ dz_x(\omega'')}\ ] +$$

$$[H_f(\omega'')\ H_f(\omega''')^*\ \overline{H_{1f}(\omega,t)\ H_{1f}(\omega',t')^*}$$

$$\overline{dz_{n2}(\omega'')\ dz_{n2}(\omega''')^*}\ \overline{dz_x(\omega)\ dz_x(\omega')^*}\ ] +$$

$$\overline{[H_{1f}(\omega,t)\ H_{1f}(\omega',t')^*\ H_{2f}(\omega'',t-\tau)\ H_{2f}(\omega''',t'-\tau')^*}$$

$$\overline{dz_x(\omega)\ dz_x(\omega')^*\ dz_x(\omega'')\ dz_x(\omega''')^*}\ ] \}$$

We now make the following assumptions:

(1) $n_1(t)$ and $n_2(t)$ are jointly stationary and Gaussian.

(2) The signal $x(t)$ emitted at the target is a stationary Gaussian process.

(3) The systems $H_1(\omega,t)$ and $H_2(\omega,t)$ belong to a class of cross-fourth-order-interfrequency stationary systems (CFOIS) satisfying the relation

(3.5-3)

$$\Phi_{1,2}^{[4]}(\omega,\omega',\omega'',\omega''',\ t,t-\mu,t-\mu',t-\mu'') =$$

$$\overline{H_1(\omega,t)\ H_1(\omega',t-\mu)^*\ H_2(\omega'',t-\mu')\ H_2(\omega''',t-\mu'')^*} = \tag{3.5-4}$$

$$\Phi_{1,2}^{[4]}(\omega,\omega',\omega'',\omega''',\mu,\mu',\mu'')$$

In general, the fourth order system correlation function for the cascade of scattering and filters may be obtained as in Equation (2.4-41) from the definitions (3.5-4), (3.4-1) and (3.4-2):

$$\Phi_{1_f 2_f}^{[4]}(\omega,\omega',\omega'',\omega''',\mu,\mu',\mu'') = e^{-j(\omega'\mu-\omega''\mu'+\omega'''\mu'')}$$

$$\int_{\upsilon'}\int_{\upsilon''}\int_{\upsilon'''} \Phi_{1,2}^{[4]}(\omega,\omega',\omega'',\omega''',\upsilon',\upsilon'',\upsilon''')\ e^{+j(\omega'\upsilon'-\omega''\upsilon''+\omega'''\upsilon''')} \tag{3.5-5}$$

$$T_f^{[4]}(\omega-\omega'+\omega''-\omega''',\mu,\mu',\mu'',\upsilon',\upsilon'',\upsilon''')\ d\upsilon'\ d\upsilon''\ d\upsilon'''$$

where the appropriate extension of Equation (2.4-42) is

$$T_f^{[4]}(\gamma,\mu,\mu',\mu'',\upsilon',\upsilon'',\upsilon''') = \tag{3.5-6}$$

$$\int_{\sigma} e^{+j\gamma\sigma}\ h_f(\sigma)\ h_f(\sigma-\mu+\upsilon')\ h_f(\sigma-\mu'+\upsilon'')\ h_f(\sigma-\mu''+\upsilon''')\ d\sigma$$

When assumptions (3.5-3) hold, using (2.1-18), (2.1-23), (2.1-16) and (2.1-17) we have

$$\Pi(\tau,\tau',t,t') =$$

$$\left(\frac{1}{2\pi}\right)^2 \int_{\omega}\int_{\omega'}\int_{\omega''}\int_{\omega'''} e^{j(\omega+\omega'')t-j(\omega'+\omega''')t'-j(\omega''\tau-\omega'''\tau')} \Big($$

$$H_f(\omega)\, H_f(\omega')^*\, H_f(\omega'')\, H_f(\omega''')^* \Big[$$

$$S_{n_1n_1}(\omega)\, \delta(\omega-\omega')\, S_{n_2n_2}(\omega'')\, \delta(\omega''-\omega''') +$$

$$S_{n_1n_2}(\omega)\, \delta(\omega+\omega'')\, S_{n_1n_2}(\omega''')\, \delta(\omega'''+\omega') +$$

$$S_{n_1n_2}(\omega)\, \delta(\omega-\omega''')\, S_{n_2n_1}(\omega')\, \delta(\omega'-\omega'') \Big]$$

$$|H_f(\omega)|^2\, \Phi_{2f}(\omega'',\omega'',t-t'-\tau+\tau') \qquad (3.5\text{-}7)$$

$$S_{n_1n_1}(\omega)\, \delta(\omega-\omega')\, S_{xx}(\omega'')\, \delta(\omega''-\omega''')$$

$$+ |H_f(\omega)|^2\, \Phi_{1_f2_f}(\omega''',\omega''',\tau')$$

$$S_{n_1n_2}(\omega)\, \delta(\omega+\omega'')\, S_{xx}(\omega''')\, \delta(\omega'+\omega''')$$

$$+ |H_f(\omega)|^2\, \Phi_{2_f1_f}(\omega',\omega',t-t'-\tau)$$

$$S_{n_1n_2}(\omega)\, \delta(\omega-\omega''')\, S_{xx}(\omega')\, \delta(\omega''-\omega')$$

$$+ |H_f(\omega')|^2\, \Phi_{1_f2_f}(\omega,\omega,t-t'+\tau')$$

$$S_{n_2n_1}(\omega')\, \delta(\omega'-\omega'')\, S_{xx}(\omega)\, \delta(\omega-\omega''')$$

$$+ |H_f(\omega''')^2\, \Phi_{2_f1_f}(\omega,\omega,-\tau)^*$$

$$S_{n_1n_2}(\omega''')\, \delta(\omega'+\omega''')\, S_{xx}(\omega)\, \delta(\omega+\omega'')$$

$$+ |H_f(\omega'')|^2 \Phi_{1f}(\omega,\omega,t-t')$$

$$S_{n_2n_2}(\omega'')\ \delta(\omega''-\omega''')\ S_{xx}(\omega)\ \delta(\omega-\omega')$$

$$+ \Phi^{[4]}_{1_f2_f}(\omega,\omega',\omega'',\omega''',t-t',\tau,t-t'+\tau')\ [$$

$$S_{xx}(\omega)\ \delta(\omega-\omega')\ S_{xx}(\omega'')\ \delta(\omega''-\omega''') +$$

$$S_{xx}(\omega)\ \delta(\omega+\omega'')\ S_{xx}(\omega''')\ \delta(\omega'''+\omega') +$$

$$S_{xx}(\omega)\ \delta(\omega-\omega''')\ S_{xx}(\omega')\ \delta(\omega'-\omega'')\ ]\ \Big)\ d\omega\ d\omega'\ d\omega''\ d\omega'''$$

Upon rearranging terms and performing two integrals over frequency for all terms we have

$$\Pi(\tau,\tau',t,t') = \left(\frac{1}{2\pi}\right)^2 \{ \int_{\omega}\int_{\omega''} e^{j(\omega+\omega'')\ (t-t')-j(\tau-\tau')\omega''} \Big($$

$$|H_f(\omega)|^2\ [|H_f(\omega'')|^2\ S_{n_1n_1}(\omega)\ S_{n_2n_2}(\omega'')$$

$$+ \Phi_{2f}(\omega'',\omega'',t-t'-\tau+\tau')\ S_{n_1n_1}(\omega)\ S_{xx}(\omega'')\ ]$$

$$+ \Phi_{1f}(\omega,\omega,t-t')\ |H_f(\omega'')|^2\ S_{n_2n_2}(\omega'')\ S_{xx}(\omega)$$

$$+ \Phi^{[4]}_{1_f2_f}(\omega,\omega,\omega'',\omega'',t-t',\tau,t-t'+\tau')\ S_{xx}(\omega)\ S_{xx}(\omega'')\ \Big)\ d\omega\ d\omega''\ \}$$

$$+ \left(\frac{1}{2\pi}\right)^2 \{ \int_{\omega}\int_{\omega'''} e^{-j(-\omega\tau-\omega'''\tau')}\ \Big($$

$$|H_f(\omega)|^2\ [|H_f(\omega''')|^2\ S_{n_1n_2}(\omega)\ S_{n_1n_2}(\omega''')$$

$$+ \Phi_{1_f2_f}(\omega''',\omega''',\tau')\ S_{n_1n_2}(\omega)\ S_{xx}(\omega''')\ ]$$

$$+ \Phi_{1_f 2_f}(\omega,\omega,+\tau)\ |H_f(\omega''')|^2\ S_{n_1 n_2}(\omega''')\ S_{xx}(\omega)$$

$$+ \Phi^{[4]}_{1_f 2_f}(\omega,-\omega''',-\omega,\omega''',t-t',\tau,t-t'+\tau')\ S_{xx}(\omega)\ S_{xx}(\omega''') \Big\}\ d\omega\ d\omega''' \Big\}$$

$$+ \left(\frac{1}{2\pi}\right)^2 \Big\{ \int_{\omega} \int_{\omega'} e^{+j(\omega+\omega')\ (t-t')-j(\omega'\tau-\omega\tau')} \Big($$

$$|H_f(\omega)|^2\ [\,|H_f(\omega')|^2\ S_{n_1 n_2}(\omega)\ S_{n_2 n_1}(\omega') \qquad (3.5\text{-}8)$$

$$+ \Phi_{2_f 1_f}(\omega',\omega',t-t'-\tau)\ S_{n_1 n_2}(\omega)\ S_{xx}(\omega')\ ]$$

$$+ \Phi_{1_f 2_f}(\omega,\omega,t-t'+\tau')\ |H_f(\omega')|^2\ S_{n_2 n_1}(\omega')\ S_{xx}(\omega)$$

$$+ \Phi^{[4]}_{1_f 2_f}(\omega,\omega',\omega',\omega,t-t',\tau,t-t'+\tau')\ S_{xx}(\omega)\ S_{xx}(\omega') \Big\}\ d\omega\ d\omega' \Big\}$$

It follows again that the assumptions stated imply that $\Pi(\tau,\tau',t,t')$ is a function of the displacement $t-t'$. Therefore, the results of section 3.2 apply.

Rewriting Equation (3.5-8) in summary form we have the equivalent of Equation (3.3-8)

$$\Delta(\tau,\tau',T,\mu) = \left(\frac{1}{T}\right)^2 \Big\{ \Big[\int_{\mu-T}^{\mu} [T+\nu-\mu] + \int_{\mu}^{\mu+T} [T-\nu+\mu] \Big]$$

$$\{M_a(\nu)\ M_a(\nu-\tau+\tau') + K_a(\nu-\tau+\tau')\ M_a(\nu) + K_a(\nu)\ M_a(\nu-\tau+\tau') \qquad (3.5\text{-}9)$$

$$+ M_c(\nu+\tau')\ M_c(\tau-\nu) + K_c(\nu+\tau')\ M_c(\nu-\tau) + K_c(\tau-\nu)\ M_c(-\nu-\tau')$$

$$+ J_1(\tau,\tau',\nu) + J_2(\tau,\tau',\nu) + J_3(\tau,\tau',\nu)\}\ d\nu\Big\} - J_4(\tau,\tau')$$

where we have assumed symmetry between the two channels, defining the quantities

$$S_{nn}(\omega) = S_{n_1 n_1}(\omega) = S_{n_2 n_2}(\omega) \tag{3.5-10}$$

$$\Phi_a(\omega,\omega,\upsilon) = \Phi_{1_f 1_f}(\omega,\omega,\upsilon) = \Phi_{2_f 2_f}(\omega,\omega,\upsilon) \tag{3.5-11}$$

and the integrals

$$M_a(\upsilon) = \int e^{j\omega\upsilon} |H_f(\omega)|^2 S_{nn}(\omega) \frac{d\omega}{2\pi} \tag{3.5-12}$$

$$K_a(\upsilon) = \int e^{j\omega\upsilon} \Phi_a(\omega,\omega,\upsilon) S_{xx}(\omega) \frac{d\omega}{2\pi} \tag{3.5-13}$$

$$M_c(\upsilon) = \int e^{j\omega\upsilon} |H_f(\omega)|^2 S_{n_1 n_2}(\omega) \frac{d\omega}{2\pi} \tag{3.5-14}$$

$$K_c(\upsilon) = \int e^{j\omega\upsilon} \Phi_{1_f 2_f}(\omega,\omega,\upsilon) S_{xx}(\omega) \frac{d\omega}{2\pi} \tag{3.5-15}$$

$$J_1(\tau,\tau',\upsilon) = \int_{\omega} \int_{\omega'} e^{j[\omega\upsilon+\omega'(\upsilon-\tau+\tau')]} \Phi^{[4]}_{1_f 2_f}(\omega,\omega,\omega',\omega',\upsilon,\tau,\upsilon+\tau') S_{xx}(\omega) S_{xx}(\omega') \frac{d\omega}{2\pi} \frac{d\omega'}{2\pi} \tag{3.5-16}$$

$$J_2(\tau,\tau',\upsilon) = \int_{\omega} \int_{\omega'} e^{j(\omega\tau+\omega'\tau')} \Phi^{[4]}_{1_f 2_f}(\omega,-\omega',-\omega,\omega',\upsilon,\tau,\upsilon+\tau') S_{xx}(\omega) S_{xx}(\omega') \frac{d\omega}{2\pi} \frac{d\omega'}{2\pi} \tag{3.5-17}$$

$$J_3(\tau,\tau',\upsilon) = \int_{\omega}\int_{\omega'} e^{j[\omega(\upsilon+\tau')+\omega'(\upsilon-\tau)]} \tag{3.5-18}$$

$$\Phi^{[4]}_{1_f 2_f}(\omega,\omega',\omega',\omega,\upsilon,\tau,\upsilon+\tau')\; S_{xx}(\omega)\; S_{xx}(\omega')\; \frac{d\omega}{2\pi}\,\frac{d\omega'}{2\pi}$$

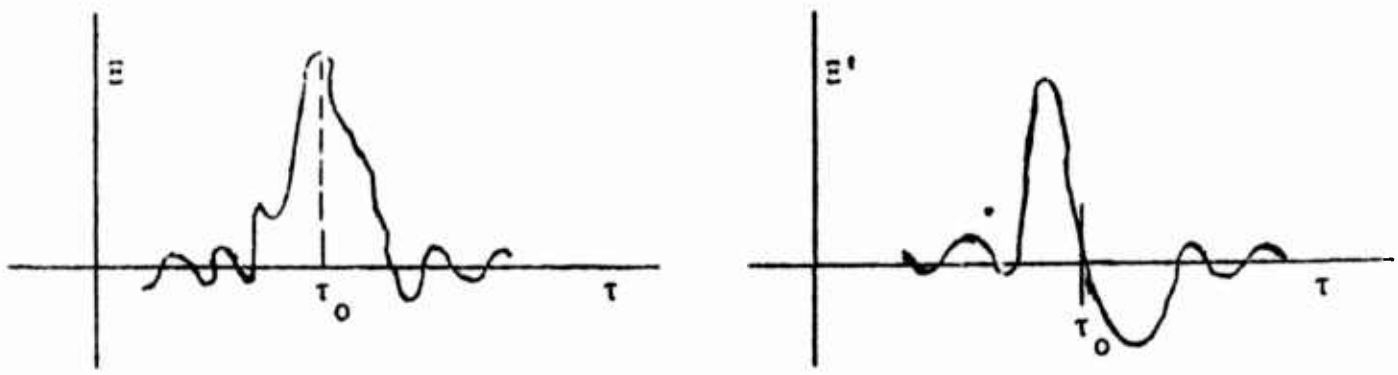
$$J_4(\tau,\tau') = \left(\frac{1}{2\pi}\right)^2 \{ \left[\int \Phi_{1_f 2_f}(\omega,\omega,\tau)\; S_{xx}(\omega)\; e^{j\omega\tau}\, d\omega\right] \tag{3.5-19}$$

$$\left[\int \Phi_{1_f 2_f}(\omega',\omega',\tau')\; S_{xx}(\omega')\; e^{j\omega'\tau'}\, d\omega'\right] \}$$

$$= K_c(\tau)\; K_c(\tau')$$

3.6 Correlator Tracking Error

In both the multipath and array type cross-correlators of Figures 3.1-1 and 3.4-1 a target is detected when the appropriate signal replication produces a peak of the output Ξ which is reliably discernible above the estimation noise (see Figure 3.6-1a). This peak of $\Xi(\tau,T,\rho)$ occurs in the vicinity of the value of the replication delay τ_o for any given value of the "read-out" time $\rho+T$. Since Ξ is a random variable the position of the peak fluctuates around the true value τ_o. Once the target is acquired in the display of Ξ as a function of τ it is usually desirable to obtain and track the peak with increased accuracy.



Ξ and Ξ' as a Function of τ

For a Given ρ

Figure 3.6-1

A suitable method of measuring the accuracy of this scheme of estimating τ_o is to obtain the variance of the location of the main peak in Ξ.[60,61,62] This is equivalent to obtaining the variance of the corresponding zero crossing in the derivative of Ξ

$$\Xi'(\tau,T,\rho) = \frac{\partial}{\partial\tau}\,\Xi(\tau,T,\rho) \qquad (3.6\text{-}1)$$

This assumes that the derivative exists in the vicinity of this null. Furthermore, it is assumed that only one isolated null exists near τ_o. This in turn requires appropriate limitations on the bandwidth of the signals and noises (which are normally satisfied).

Asymptotically with large T we expect the variance of the zero crossing to approach the variance of Ξ' divided by the mean slope of Ξ' at the null ξ squared:

$$\sigma^2_{\tau_o} \to \left.\frac{\mathrm{Var}[\Xi']}{\left[\frac{\partial \overline{\Xi'}}{\partial \tau}\right]^2}\right|_{\tau=\tau_o}$$

$$= \left.\frac{\overline{\frac{\partial \Xi(\tau,T,\rho)}{\partial \tau}\frac{\partial \Xi(\tau',T,\rho)}{\partial \tau'}} - \overline{\frac{\partial \Xi(\tau,T,\rho)}{\partial \tau}}\,\overline{\frac{\partial \Xi(\tau',T,\rho)}{\partial \tau'}}}{\left[\frac{\partial^2 \overline{\Xi(\tau,T,\rho)}}{\partial \tau^2}\right]^2}\right|_{\tau=\tau'=\tau_o}$$

$$= \left.\frac{\frac{\partial^2 \Delta(\tau,\tau',T,0)}{\partial \tau\,\partial \tau'}}{\left[\frac{\partial^2 \overline{\Xi(\tau,T,\rho)}}{\partial \tau^2}\right]^2}\right|_{\tau=\tau'=\tau_o} \qquad (3.6\text{-}2)$$

To be valid this formula requires T to be large enough so that the fluctuation of the zero crossing is smaller than the width of the correlation peak. This permits the linearization implied by Equation (3.6-2) (see Figure 3.6-1).

By entirely similar reasoning we may consider the zero-crossing to be a function of ρ and compute an auto-correlation function for this

random variable:

$$R_{\tau_o}(\mu) + \left. \frac{\dfrac{\partial^2 \Delta(\tau,\tau',T,\mu)}{\partial\tau\ \partial\tau'}}{\left| \dfrac{\partial^2 \Xi(\tau,T,\rho)}{\partial\tau^2} \right|^2} \right|_{\tau=\tau'=\tau_o} \tag{3.6-3}$$

This quantity measures the persistance of the tracking errors in the same way that Δ measures the persistance of fluctuations in Ξ .

3.7 Concerning the Departure of the Statistics of the Scattered Signals from the Gaussian

A key assumption made in sections 3.3 and 3.5 is that the target signal $x(t)$ is Gaussian, i.e. its time samples are statistically described by the probability density Equation (2.1-11). After passage of such a signal through a random scattering system the joint probability density of the scattered signal samples is usually no longer Gaussian. Making certain assumptions about convergence, the joint density for N samples of a non-Gaussian process can be written in terms of an Edgeworth expansion[65] as follows:

$$f(y_1,y_2,\ldots,y_N) =$$

$$\exp\left\{\sum_{\nu\geq 3}^{\infty}(-1)^{\nu}\kappa_{\nu_1\nu_2\ldots\nu_N}\prod_{i=1}^{N}\left[\frac{1}{(\nu_i)!}\left(\frac{\partial}{\partial y_i}\right)^{\nu_i}\right]\right\}\theta(y_1,y_2,\ldots,y_N) \qquad (3.7\text{-}1)$$

where all possible permutations of the ν_i are taken subject to

$$\nu = \sum_{i=1}^{N}\nu_i \qquad (3.7\text{-}2)$$

and where $\theta(y_1,y_2,\ldots,y_N)$ is the Gaussian density given by Equation (2.1-11) using the correlation matrix $\underline{R} = \overline{\underline{y}\,\underline{y}^T}$.

The constants $\kappa_{\nu_1\nu_2\ldots\nu_N}$ are termed the cumulants for the joint density f and the parameter ν denotes the order of the cumulants. Cumulants of order 1 are defined as the means of the y_i (these are zero in the present context). Cumulants of order 2 are correlations between samples and are elements of the matrix $\underline{R}$. Cumulants of order 3 or higher can be obtained as in the case of one-dimensional probability

densities from the various moments:

$$\mu_{\iota_1\iota_2\cdots\iota_N} = \int_{y_1}\int_{y_2}\cdots\int_{y_N}\prod_{k=1}^{N}[y_k^{\iota_k}]\, f(y_1,y_2,\ldots,y_N)\, dy_1\, dy_2\cdots dy_N \qquad (3.7\text{-}3)$$

This is accomplished by forming the characteristic function corresponding to Equation (3.7-1):

$$Q(u_1,u_2,\ldots,u_N) =$$

$$1 + \sum_{\iota\geq 1}^{\infty}\mu_{\iota_1\iota_2\cdots\iota_N}\prod_{k=1}^{N}\left[\frac{1}{(\iota_k)!}(ju_k)^{\iota_k}\right] = \qquad (3.7\text{-}4)$$

$$\exp\left\{\sum_{\nu\geq 1}\kappa_{\nu_1\nu_2\cdots\nu_N}\prod_{i=1}^{N}\left[\frac{1}{(\nu_i)!}(ju_i)^{\nu_i}\right]\right\}$$

where all permutations of the ν_i and ι_k are taken subject to the constraints Equation (3.7-2) and

$$\iota = \sum_{k=1}^{N}\iota_k \qquad (3.7\text{-}5)$$

By taking the log of Equation (3.7-4) and equating coefficients of equal products of the u_i the relations giving the cumulants in terms of the moments are obtained.

When all cumulants of order 3 or greater are zero the density f reduces to the Gaussian function θ. Hence the cumulants of order ≥ 3 measure the departure of f from θ. Because scattering is modeled here as a linear process, the odd order moments and cumulants of f are zero due to Equation (2.1-14) and kindred relations. Therefore, the

fourth order cumulants are the first to produce non-Gaussian behavior. The next group to contribute are cumulants of sixth order.

We consider here only the nature of the fourth order cumulant. It suffices to consider a four sample density (N = 4) in order to obtain the general form of the relation between the cumulants and the moments. We have from Equation (3.7-4)

$$\kappa_{1111} = \mu_{1111} - \mu_{0011}\,\mu_{1100} - \mu_{0101}\,\mu_{1010} - \mu_{1001}\,\mu_{0110} \qquad (3.7\text{-}6)$$

Forcing the two channels in Figure 3.4-1 to be the same, particularizing $H_f(\omega)$ to be 1, and taking the four samples at times t, t', $t-\tau$, $t'-\tau'$ we have

$$\mu_{1111} = \Pi(\tau,\tau',t,t') \qquad \text{(a)} \qquad (3.7\text{-}7)$$

$$\mu_{1100} = \overline{\Xi(t-t',T,\rho)} \qquad \text{(b)}$$

$$\mu_{0011} = \overline{\Xi(t-t'-\tau+\tau',T,\rho)} \qquad \text{(c)}$$

$$\mu_{1010} = \overline{\Xi(\tau,T,\rho)} \qquad \text{(d)}$$

$$\mu_{0101} = \overline{\Xi(\tau',T,\rho)} \qquad \text{(e)}$$

$$\mu_{1001} = \overline{\Xi(t-t'+\tau',T,\rho)} \qquad \text{(f)}$$

$$\mu_{0110} = \overline{\Xi(t'-t+\tau,T,\rho)^*} \qquad \text{(g)}$$

Substituting these into Equation (3.7-6) and using Equation (3.5-8) we have

$$\kappa_{1111}(\nu,\tau,\tau') = \left(\frac{1}{2\pi}\right)^2 \{ \int_{\omega} \int_{\omega'} S_{xx}(\omega)\, S_{xx}(\omega')$$

$$e^{j(\omega+\omega')\nu - j(\tau'-\tau)\omega'} \left[\Phi^{[4]}(\omega,\omega,\omega',\omega',\nu,\tau,\nu+\tau') - \Phi(\omega,\omega,\nu)\, \Phi(\omega',\omega',\nu-\tau+\tau')\right] +$$

$$e^{j(\omega\tau+\omega'\tau')} \left[\Phi^{[4]}(\omega,-\omega',-\omega,\omega',\nu,\tau,\nu+\tau') - \Phi(\omega,\omega,\tau)\, \Phi(\omega',\omega',\tau')\right] + \qquad (3.7\text{-}8)$$

$$e^{j(\omega+\omega')\nu - j(\omega'\tau-\omega\tau')} \left[\Phi^{[4]}(\omega,\omega',\omega',\omega,\nu,\tau,\nu+\tau') - \Phi(\omega',\omega',\nu-\tau)\, \Phi(\omega,\omega,\nu+\tau')\right] d\omega\, d\omega' \}$$

$$= [J_1(\tau,\tau',\nu) - K_a(\nu)\, K_a(\nu-\tau+\tau')]$$

$$+ [J_2(\tau,\tau',\nu) - K_a(\tau)\, K_a(\tau')]$$

$$+ [J_3(\tau,\tau',\nu) - K_a(\nu+\tau')\, K_a(\tau-\nu)]$$

The integrals for J_1, J_2, J_3, and K_a are evaluated for various models for scattering in the remaining chapters of this report. In general, the result for $\kappa_{1111}(\nu,\tau,\tau')$ is found to be non-zero. The scattered signal is therefore a non-Gaussian process.

3.8 The Two-Sided Likelihood Ratio Decision Scheme for the Correlator Detector

It is clear from the discussion of section 3.7 that since the probability density for the scattered signal is generally non-Gaussian, joint statistics for a two channel array or multipath configurations are also non-Gaussian. Evidence presented in the remainder of this report indicates the importance of taking into account this departure from Gaussian statistics in describing the scattered signals. The design of optimal detectors for such signals is not a trivial task, and we shall not undertake this problem.

In lieu of a formal attack on the optimal detector design problem we regard the detector structures of Figures 3.1-1 and 3.4-1 as essentially fixed and inquire how best to decide whether a target is present from the output $\Xi(\tau,T,\rho)$. We begin by examining the case for which the replication delay parameter τ is fixed or "steered" to the correct value τ_o given that a target is present. We call this the "on-target" detection problem since we effectively evaluate the ability of the detector to function properly while it is "looking" directly at the target.

Given the correlator output $\Xi(\tau_o,T,\rho)$ one must decide between two hypotheses:

H_0: The target is not present.
(The correlator output is due to noise only.)

H_1: The target is present.
(The correlator output contains a component due to the presence of signal.)

We assume that the correlator integration interval T is large enough so that $\Xi(\tau_o,T,\rho)$ tends to be a Gaussian random variable[68] under either

hypothesis. We then have the following probability densities for Ξ :

$$H_0: \qquad f_0(\Xi) = \frac{1}{\sqrt{2\pi}\,\sigma_0} \exp\{-\tfrac{1}{2}(\Xi - \bar{\Xi}_0)^2/\sigma_0^2\} \tag{3.8-1}$$

$$H_1: \qquad f_1(\Xi) = \frac{1}{\sqrt{2\pi}\,\sigma_1} \exp\{-\tfrac{1}{2}(\Xi - \bar{\Xi}_1)^2/\sigma_1^2\} \tag{3.8-2}$$

As a practical matter $\bar{\Xi}_0$ is either zero or approximately zero for most problems considered, so we set it to zero here.

Now the likelihood ratio procedure[69] for deciding between the two alternative hypotheses H_0 and H_1 given a measured value for Ξ is to decide in favor of H_0 if

$$\frac{f_0(\Xi)}{f_1(\Xi)} \geq K_{th} \tag{3.8-3}$$

and in favor of H_1 if

$$\frac{f_0(\Xi)}{f_1(\Xi)} < K_{th} \tag{3.8-4}$$

The performance of such a decision scheme depends on the choice of the threshold value K_{th} and on the parameters σ_0, $\bar{\Xi}_1$, and σ_1 which are determined by the statistics of the signals, noises and scattering. These latter parameters may be computed from Equations (3.1-9) and (3.3-8) for the multipath detector or (3.4-4) and (3.5-9) for the two-element array detector.

The critical values of Ξ for which equality holds in Equation (3.8-3) are obtained by substituting Equations (3.8-1) and (3.8-2) and taking the natural logarithm yielding

$$\Xi^2 \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right) + 2\,\Xi\,\frac{\bar{\Xi}_1}{\sigma_1^2} - \frac{\bar{\Xi}_1^2}{\sigma_1^2} = -2\,\mathrm{Ln}\left(\frac{K_{th}\sigma_0}{\sigma_1}\right) \qquad (3.8\text{-}5)$$

The solution for the roots of this quadratic leads to the critical values:

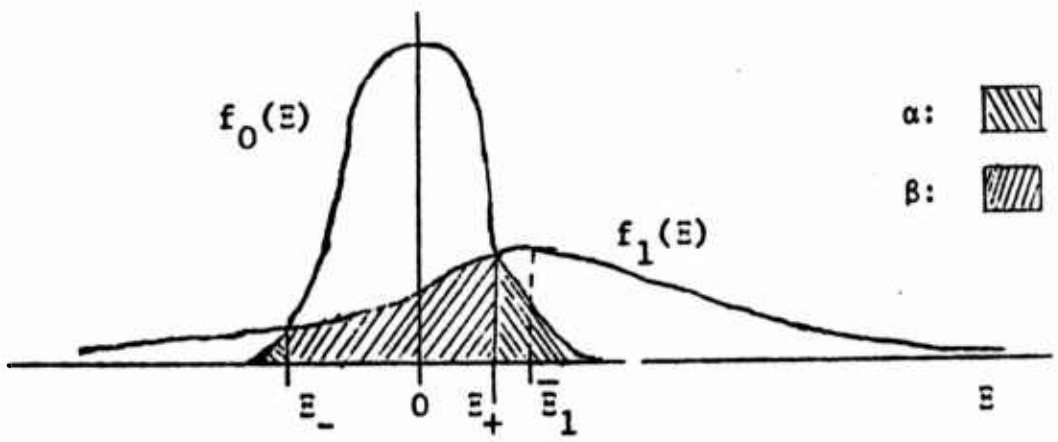
$$\Xi_{\pm} = \frac{\bar{\Xi}_1\, d_0\, d_1}{d_0^2 - d_1^2}\left[\frac{d_0}{d_1} \mp \sqrt{1 + 2\,[d_0^2 - d_1^2]\,\ell n\left(\frac{K_{th}\, d_0}{d_1}\right)}\,\right] \qquad (3.8\text{-}6)$$

where

$$d_0 = \frac{\sigma_0}{\bar{\Xi}_1} \quad ; \quad d_1 = \frac{\sigma_1}{\bar{\Xi}_1} \qquad (3.8\text{-}7)$$

are the standard deviations of the correlator output under the two hypotheses H_0 and H_1 normalized by the mean under H_1.

In general, the variance σ_1 is greater than σ_0 (see Figure (3.8-1)). The special case of $d_1 \to d_0$ in Equation (3.8-6) for which one critical value tends to ∞ arises only in the case of very weak signal-to-noise ratios. Therefore, the two critical values divide



The Decision Regions for the Two-Sided Detector (K_{th}=1)

Figure 3.8-1

the range $-\infty < \Xi < \infty$ into three decision intervals:

$$\begin{aligned} -\infty < \Xi < \Xi_- \quad & \text{Decide in favor of } H_1 . \\ \Xi_- < \Xi < \Xi_+ \quad & \text{Decide in favor of } H_0 . \\ \Xi_+ < \Xi < \infty \quad & \text{Decide in favor of } H_1 . \end{aligned} \tag{3.8-8}$$

Given any threshold value K_{th} and the parameters d_0 and d_1 one may evaluate the decision scheme Equation (3.8-8) by computing the probability of incorrect decisions. The probability α of committing a type I error (false alarm)[70] is equal to the probability of Ξ incorrectly exceeding one of the critical values $\Xi_\pm$:

$$\alpha = \int_{-\infty}^{\Xi_-} f_0(\Xi)\, d\Xi + \int_{\Xi_+}^{+\infty} f_0(\Xi)\, d\Xi$$

$$= 1 - \frac{1}{2}\left[\theta\left\{\frac{\Xi_+}{\sqrt{2}\,\sigma_0}\right\} - \theta\left\{\frac{\Xi_-}{\sqrt{2}\,\sigma_0}\right\} \right] \tag{3.8-9}$$

$$= 1 - \frac{1}{2}\left[\theta\left\{\frac{d_1}{\sqrt{2}\,[d_0^2 - d_1^2]} \left(\frac{d_0}{d_1} - \sqrt{1 + 2\,[d_0^2 - d_1^2]\,\ell n\left(\frac{K_{th}\, d_0}{d_1}\right)}\right)\right\} \right.$$

$$\left. - \theta\left\{\frac{d_1}{\sqrt{2}\,[d_0^2 - d_1^2]} \left(\frac{d_0}{d_1} + \sqrt{1 + 2\,[d_0^2 - d_1^2]\,\ell n\left(\frac{K_{th}\, d_0}{d_1}\right)}\right)\right\} \right]$$

where $\theta(z)$ is the error function[71]:

$$\theta(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-y^2}\, dy \tag{3.8-10}$$

Similarly, the probability β of a type II error or false dismissal is determined by the probability that Ξ incorrectly lies within the critical levels $\Xi_{\pm}$ when in fact a target is present:

$$\beta = \int_{\Xi-}^{\Xi+} f_1(\Xi)\, d\Xi$$

$$= \frac{1}{2}\left[\Theta\left\{ \frac{d_0}{\sqrt{2}(d_0^2 - d_1^2)}\left(\frac{d_1}{d_o} - \sqrt{1 + 2\,[d_0^2 - d_1^2]\,\ell n\left(\frac{K_{th}}{d_1}\frac{d_0}{}\right)}\right)\right\}\right. \tag{3.8-11}$$

$$\left. - \Theta\left\{ \frac{d_0}{\sqrt{2}(d_0^2-d_1^2)}\left(\frac{d_1}{d_o} + \sqrt{1 + 2\,[d_0^2 - d_1^2]\,\ell n\left(\frac{K_{th}\, d_0}{d_1}\right)}\right)\right\}\right]$$

In general, the "on-target" normalized variances d_0^2 and d_1^2 are overly pessimistic indicators of the performance of the correlator detectors. This is true because the peak of measured correlation $\Xi(\tau,T,\rho)$ usually does not occur precisely at the correct location τ_o (see section 3.6). Consequently, a detector which initiates a search for the peak in the vicinity of τ_o performs somewhat better. A procedure for evaluating this "search" detector for certain classes of signals is discussed in Appendix D.

Finally we note the importance of using the two-sided test Equation (3.8-8) in cases of moderate or strong scattering. In such situations Ξ_1 may be quite small while σ_1 may be much greater than σ_0 leading to sign changes in the peak of correlation (see Figure (3.8-1). Failure to use the two-sided test when this occurs can produce sizable losses in detectability.

CHAPTER 4

THE RANDOM AMPLITUDE AND DELAY MODEL FOR SCATTERING

4.0 Introduction

In this chapter surface scattering is modeled as a time-varying system with the transfer function

$$H(\omega,t) = A(t)\, e^{-j\omega\tau(t)} \qquad (4.0\text{-}1)$$

where the amplitude A(t) and delay $\tau(t)$ are considered independent, stationary, Gaussian random variables. This simple model is used by Price,[73] Green,[74] Turin,[75] and others in studies of communication through ionospheric and tropospheric channels. It serves as an example which produces certain effects on correlator outputs which are typical of more complex and realistic models.

The mean and variance of the outputs from multipath and array cross-correlators are computed using the formulas derived in Chapter 3. It is shown that as the correlator integration time is increased from zero the output variance at first decreases rapidly. However, at the point where reliable correlation begins to emerge in the absence of scattering the variance passes through a "settling" period during which averaging over slowly varying scattering fluctuations takes place. The persistence of these fluctuations is computed and related to the 4-the order cumulant of the scattered signal.

4.1 First And Second Order System Statistics For The Random Amplitude And Delay Model

It is assumed that the amplitude and delay in (4.0-1) have the mean values

$$\text{a)}\quad \overline{A(t)} = \frac{\eta_s}{r_o + r_o'} = A_c \qquad \text{b)}\quad \overline{\tau(t)} = \frac{r_o + r_o'}{c} = \tau_s \qquad (4.1\text{-}1)$$

where the distances r_o and r_o' are defined in Figure 1.4-1 and η_s is a constant less than one. We define the correlation functions

$$\text{a)}\quad R_{AA}(\mu) = \overline{A(t)\ A(t-\mu)} \qquad \text{b)}\quad R_{\tau T}(\mu) = \overline{\tau(t)\tau(t-\mu)} - \tau_s^2 \qquad (4.2.1\text{-}2)$$

where for convenience R_{AA} is defined as a non-central moment. Since A(t) and $\tau(t)$ are independent

$$\overline{H(\omega,t)} = H_c(\omega) = \overline{A(t)}\ \overline{e^{-j\omega\tau(t)}} = A_c\ \ Q_\tau^{(1)}(\omega) \qquad (4.1\text{-}3)$$

where $Q_\tau^{(1)}(\omega)$ is the one dimensional characteristic function for $\tau(t)$. From the assumed stationarity of τ, $Q_\tau^{(1)}(\omega)$ is time invariant.

In a similar manner,

$$\phi(\omega,\omega',\mu) = \overline{H(\omega,t)\ H(\omega',t-\mu)^*} =$$

$$\overline{A(t)\ A(t-\mu)}\ \ \overline{\exp\{-j\omega\tau(t)+j\omega'\tau(t-\mu)\}} = R_{AA}(\mu)\ Q_\tau^{(2)}(\omega,-\omega',\mu) \qquad (4.1\text{-}4)$$

where $Q_\tau^{(2)}(\omega\ \omega',\mu)$ is the characteristic function for the two-dimensional probability distribution for τ. Stationarity of τ and A therefore guarantees that the scattering system will be interfrequency wide sense stationary (IWSS). However, in the case of $\tau(t)$ the requirement is

for stationarity of the two sample distribution whereas A(t) need only be wide sense stationary.

Although the instantaneous impulse reponse corresponding to (4.0-1) is of infinitesimal duration,

$$h(\nu,t) = A(t)\ \delta(\nu-\tau(t)) \qquad (4.1\text{-}5)$$

the average or coherent response is of finite width :

$$h_c(\nu) = \int e^{j\omega\nu}\ H_c(\omega)\ \frac{d\omega}{2\pi} = A_c\ f_\tau(\nu) \qquad (4.1\text{-}6)$$

where $f_\tau(\nu)$ is the probability density for τ. Immediate corollaries are obtained from (2.4-1) and (2.4-4)

$$\text{(a)}\quad D_c(\mu,\nu) = A_c\ f_\tau(u-\nu) \qquad \text{b)}\quad R_{xy}(\mu) = A_c \int R_{xx}(\mu-\nu) f_\tau(\nu)\ d\nu \qquad (4.1\text{-}7)$$

Similarly,

$$h(\nu,t)\ h(\nu',t-\mu) = \int_\omega \int_{\omega'} e^{j\omega\nu - j\omega'\nu'}\ H(\omega,t)\ H(\omega',t-\mu)\ \frac{d\omega}{2\pi}\ \frac{d\omega'}{2\pi}$$

$$= R_{AA}(\mu)\ f_\tau\ (\nu,\nu',\mu) \qquad (4.1\text{-}8)$$

where $f_\tau(\nu,\nu',\mu)$ is the two sample probability density corresponding to $Q_\tau^{(2)}(\omega,\omega',\mu)$. The auto decorrelation function is also readily found from (2.4-23)

$$D_a(\mu,\nu) = R_{AA}(\mu) \int_\omega e^{j\omega[\tau(t)-\tau(t-\mu)]}\ e^{j\omega(\mu+\nu)}\ \frac{d\omega}{2\pi}$$

$$= R_{AA}(\mu)\ f_{\delta\tau}\ (\mu+\nu,\mu) \qquad (4.1\text{-}9)$$

where $f_{\delta\tau}(\nu,\mu)$ is the probability density for the difference ν between the values of τ at times t and t-μ. It follows from (2.4-21) that

$$R_{yy}(\mu) = R_{AA}(\mu) \int_{\mu} R_{xx}(\mu-\nu) \; f_{\delta\tau}(\nu,\mu) \; d\nu \qquad (4.1\text{-}10)$$

For the special case of Gaussian τ we have

$$H_c(\omega) = A_c \; \exp\{-1/2(\omega^2\sigma_\tau^2) - j\omega\tau_s\} \qquad (4.1\text{-}11)$$

$$\phi(\omega,\omega',\mu) = R_{AA}(\mu) \; \exp\{-1/2[(\omega^2+\omega'^2)\sigma_\tau^2 - 2\omega\omega' R_{\tau\tau}(\mu)] - j(\omega-\omega')\tau_s\} \qquad (4.1\text{-}12)$$

$$D_c(\mu,\nu) = \frac{\exp\{-1/2(\mu-\nu-\tau_s)^2/\sigma_\tau^2\}}{\sqrt{2\pi}\;\sigma_\tau^2} \qquad (4.1\text{-}13)$$

$$D_a(\mu,\nu) = R_{AA}(\mu) \; \frac{\exp\{-1/2(\mu+\nu)^2/[\sigma_\tau^2 - R_\tau(\mu)]\}}{\sqrt{2\pi\,[\,\sigma_\tau^2 - R_\tau(\mu)]}} \qquad (4.1\text{-}13)$$

For this case the trifrequency spectral density becomes

$$\frac{dA(\omega,\omega',\omega'')}{d\omega''} = \exp\{-\tfrac{1}{2}\,[(\omega^2+\omega'^2)\sigma_\tau^2] - j(\omega-\omega')\tau_s\} \sum_{n=0}^{\infty} \frac{(\omega\omega')^n}{n!} \; s_n(\omega'') \qquad (4.1\text{-}14)$$

where

$$s_n(\omega'') = \int_{\mu} e^{-j\omega''\mu} \; R_{AA}(\mu) \; [R_{\tau\tau}(\mu)]^n \; d\mu \qquad (4.1\text{-}15)$$

Similarly, applying (2.4-20) to (2.4-17) and using (4.1-14) we

obtain the power spectrum smearing density :

$$\frac{dC(\nu,\gamma)}{d\gamma} = \int_{\omega} \exp(j\omega\nu)\ \frac{dA(\omega,\omega,\gamma)}{d\gamma}\ d\omega =$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_\tau} \sum_{n=o}^{\infty} \frac{s_n(\gamma)}{n!} \left(\frac{-2}{\sigma_\tau^2}\right)^n H_e^{(2n)}\left(\frac{\nu}{\sqrt{2}\,\sigma_\tau}\right) \exp\left(\frac{-\nu^2}{2\,\sigma_\tau^2}\right) \qquad (4.1\text{-}16)$$

where $H_e^{(n)}(z)$ is the m-th order Hermite polynomial.[73]

In order to interpret (4.1-14) and (4.1-16) let

$$R_{AA}(\mu) = \sigma_A^2 \exp(-1/2\mu^2\Omega_A^2) + A_c^2 \qquad (4.1\text{-}17)$$

$$R_{\tau\tau}(\mu) = \sigma_\tau^2 \exp(-1/2\mu^2\Omega_\tau^2) \qquad (4.1\text{-}18)$$

From the definition of $s_n(\gamma)$ in (4.1-15) we have

$$s_n(\gamma) = \frac{\sigma_\tau^2}{\sqrt{2\pi}\ n\ \Omega_\tau} \left[\frac{\sigma_A^2\ e^{-\frac{1}{2}\gamma^2/\Omega_A^2}}{\sqrt{\left(\frac{\Omega A}{n\Omega_\tau}\right)^2 + 1}} + A_c^2 \right] \exp\{\tfrac{1}{2}\gamma^2/(n\Omega_\tau)^2\} \qquad (4.1\text{-}18)$$

It is seen that the frequency smear for larger values of n is wider. From this we can conclude that the higher order and more intricate Hermite functions in (4.1-16) represent the components of scattering which exhibit the greatest fading. From equation (4.1-14) it can be seen that these components are excited by the higher frequency components of the output. The wider frequency smear at higher frequencies is characteristic of delay modulation.

4.2 Multipath Correlator Fluctuations for the Random Amplitude and Delay Model

Following section 3.8 we define normalized covariances

$$d_j^2(\tau,T,\mu) = \frac{\Delta(\tau,\tau,T,\mu)\big|_{H_j}}{[\Xi(\gamma\ T\ \rho)\ ]^2} \qquad j=0,1 \tag{4.2-1}$$

where $\Delta(\tau,\tau,T,\mu)$ must be evaluated under hypothesis H_0 (noise only) and H_1 (signal plus noise present) while $\overline{\Xi(\tau,T,\mu)}$ is understood to be computed for H_1. We compute here the normalized covariances for the configuration shown in Figure 3.1-1 assuming for simplicity that the spectra for the signal and background noise are given respectively by

$$\text{a)} \quad S_{xx}(\omega) = \frac{\sqrt{2\pi}\ P_x}{\Omega_x}\ \exp(-\tfrac{1}{2}\omega^2/\Omega_x^2) \tag{4.2-2}$$

and

$$\text{b)} \quad s_{n_1n_1}(\omega) = \frac{\sqrt{2\pi}\ P_{n1}}{\Omega_{n1}}\ e^{-\frac{1}{2}\omega^2/\Omega_{n1}^2} \qquad \text{c)} \quad S_{n_2n_2}(\omega) = \frac{\sqrt{2\pi}\ P_{n2}}{\Omega_{n2}}\ e^{-\frac{1}{2}\omega^2/\Omega_{n2}^2}$$

Also for simplicity we assume that the filters $H_1(\omega)$ and $H_2(\omega)$ are as follows

$$\text{a)} \quad H_1(\omega) = e^{-\frac{1}{2}\omega^2/\Omega_1^2} \qquad \text{b)} \quad H_2(\omega) = e^{-\frac{1}{2}\omega^2/\Omega_2^2} \tag{4.2-3}$$

Here P_x, P_{n1}, P_{n2} are the signal and noise powers respectively. It is assumed that the bandwidths Ω_x, Ω_1, Ω_2, Ω_{n1}, and Ω_{n2} are all very large compared with the fading bandwidths of the functions $s_n(\gamma)$ in (4.1-15) so that equation (2.5-5)may be used.

The direct path transfer function is written as

$$H_d(\omega) = A_d \; e^{-j\omega\tau_d}$$

where $\tau_d = R_d/c$ and $A_d = 1/R_d$. Using these definitions (and assuming the statistics for A(t) & τ(t) discussed in section 4.1) the mean output of the correlator under hypothesis H_1 becomes

$$\overline{\Xi(\tau,T,\rho)} = A_d \, A_c \, P_x \left(\frac{\Omega_m}{\Omega_x}\right) e^{\frac{1}{2}\Omega_m^2(\tau-\tau_o)^2} \qquad (4.2\text{-}5)$$

where $\tau_o = \tau_s - \tau_d$ and

$$\frac{1}{\Omega_m^2} = \sigma_\tau^2 + \frac{1}{\Omega_x^2} + \frac{1}{2\,\Omega_1^2} + \frac{1}{2\,\Omega_2^2} \; . \qquad (4.2\text{-}6)$$

The mean exhibits a peak of magnitude $A_d A_c P_x \Omega_m/\Omega_x$ at $\tau = \tau_o$ corresponding to the true target laocation.

Let us assume that $R_{AA}(\mu)$ and $R_{\tau\tau}(\mu)$ are as given in (4.1-17) and (4.1-18). Then using the results of Appendix E.1 we have the following result for the "on-target" normalized covariance for the null hypothesis H_0 :

$$d_0^2(\tau_o,T,\mu) =$$

$$\frac{\Omega_x^2/\Omega_m^2}{T^2 A_d^2 A_c^2}\left(\frac{P_{n1}}{P_x}\right)\left(\frac{P_{n2}}{P_x}\right)\left(\frac{\Omega_{1n1}}{\Omega_{n1}}\right)\left(\frac{\Omega_{2n2}}{\Omega_{n2}}\right) \int_{\mu-T}^{\mu} [T+\nu-\mu] \qquad (4.2\text{-}7)$$

$$+ \int_{\mu}^{\mu+T} [T-\nu+\mu] \qquad \exp\{-\tfrac{1}{2}\nu^2 [\Omega_{1n1}^2 + \Omega_{2n2}^2]\} \; d\nu$$

Similarly, we have the "on-target" normalized covariance for the hypothesis H_1 :

$$d_1^2(\tau_o, T, \mu) =$$

$$\frac{\Omega_x^2/\Omega_m^2}{T^2 A_d^2 A_c^2} \left[\int_{\mu-T}^{\mu} [T+\nu-\mu] + \int_{\mu}^{\mu+T} [T-\nu+\mu] \right] \Bigg(\left[\left(\frac{P_{n1}}{P_x} \right) \left(\frac{\Omega_{2n2}}{\Omega_{n1}} \right) \exp\{ -\frac{1}{2} \nu^2 \Omega_{1n1}^2 \} \right.$$

$$+ \frac{R_{AA}(\nu)}{\Omega_x \sqrt{\frac{1}{\Omega_{1x}^2} + 2[\sigma_\tau^2 - R_{\tau\tau}(\nu)]}} \exp\left\{ \frac{-\nu^2}{2} \left[\left[\frac{1}{\Omega_{1x}^2} + 2[\sigma_\tau^2 - R_{\tau\tau}(\nu)] \right]^{-1} \right\} \right]$$

$$\left[\left(\frac{P_{n2}}{P_x} \right) \left(\frac{\Omega_{2n2}}{\Omega_{n1}} \right) e^{-1/2 \nu^2 \Omega_{2n2}^2} + A_d^2 \left(\frac{\Omega_{1x}}{\Omega_x} \right) e^{-1/2 \nu^2 \Omega_{1x}^2} \right] \tag{4.2-8}$$

$$+ \frac{R_{AA}(\nu) A_d^2}{\Omega_x^2 \sqrt{\alpha_m^4 - R_{\tau\tau}^2(\nu)}} \exp\left\{ \frac{-\nu^2}{2} \left[\alpha_m^2 - R_{\tau\tau}(\nu) \right]^{-1} \right\}$$

$$+ \left(\frac{A_c^2 A_d^2}{\Omega_x^2/\Omega_m^2} \right) \left[\frac{R_{AA}(\nu)}{A_c^2} \frac{\alpha_m^2}{\sqrt{\alpha_m^4 - R_{\tau\tau}^2(\nu)}} - 1 \right] \Bigg) d\nu$$

The effective bandwidths Ω_{1x}, Ω_{1n1}, and Ω_{2n2} are defined in Appendix E.1 and

$$\alpha_m = 1/\Omega_m \tag{4.2-9}$$

We now restrict our attention to the special case of

$$P_{n1} = P_{n2} = P_n \qquad \text{(a)}$$

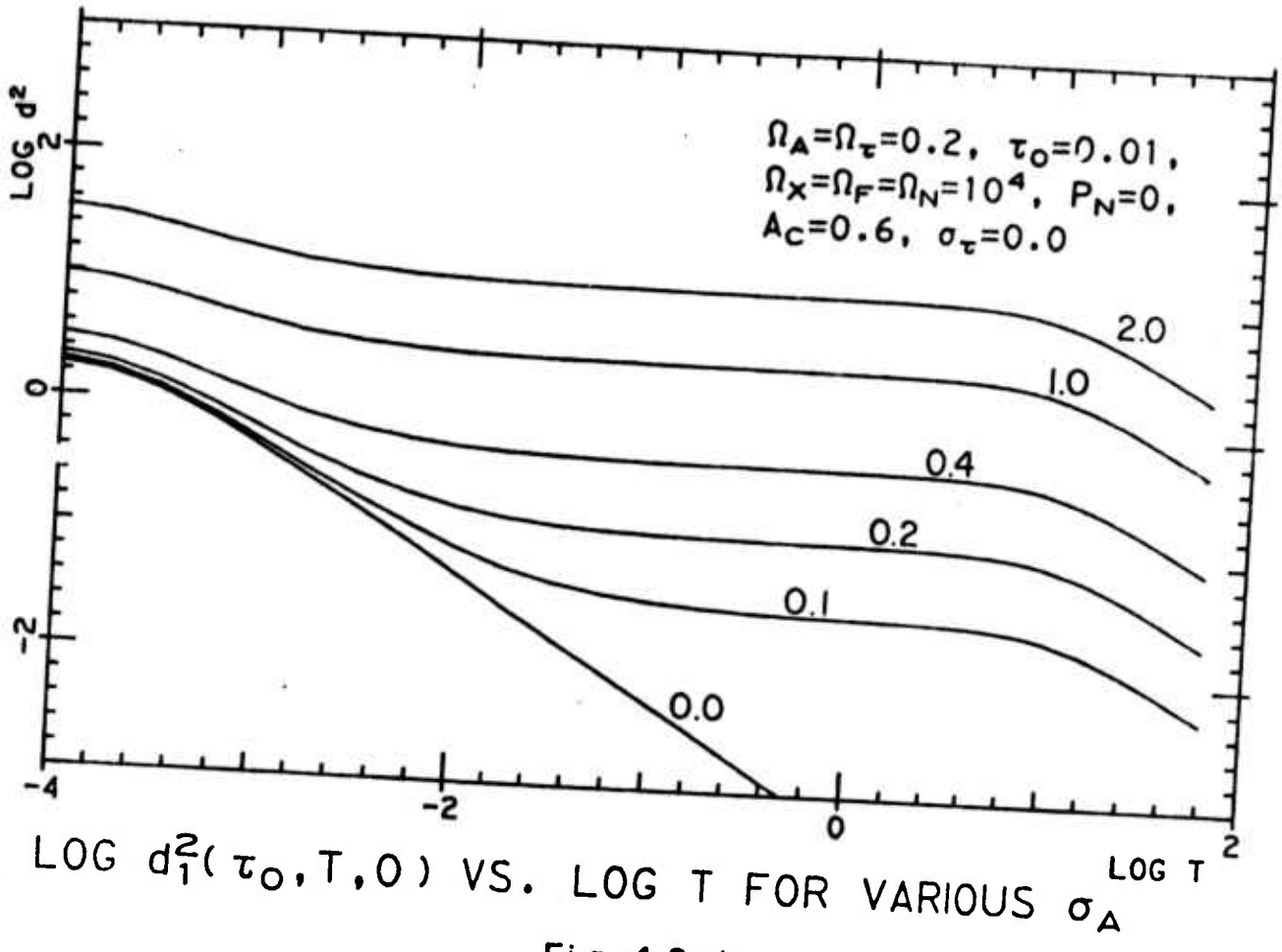
$$\Omega_{n1} = \Omega_{n2} = \Omega_n \qquad \text{(b)}$$

$$\Omega_1 = \Omega_2 = \Omega_f \qquad \text{(c)}$$

We may easily compute the normalized covariances for various numerical examples, but exact closed-form expressions are not readily obtained for all of the integrals involved.

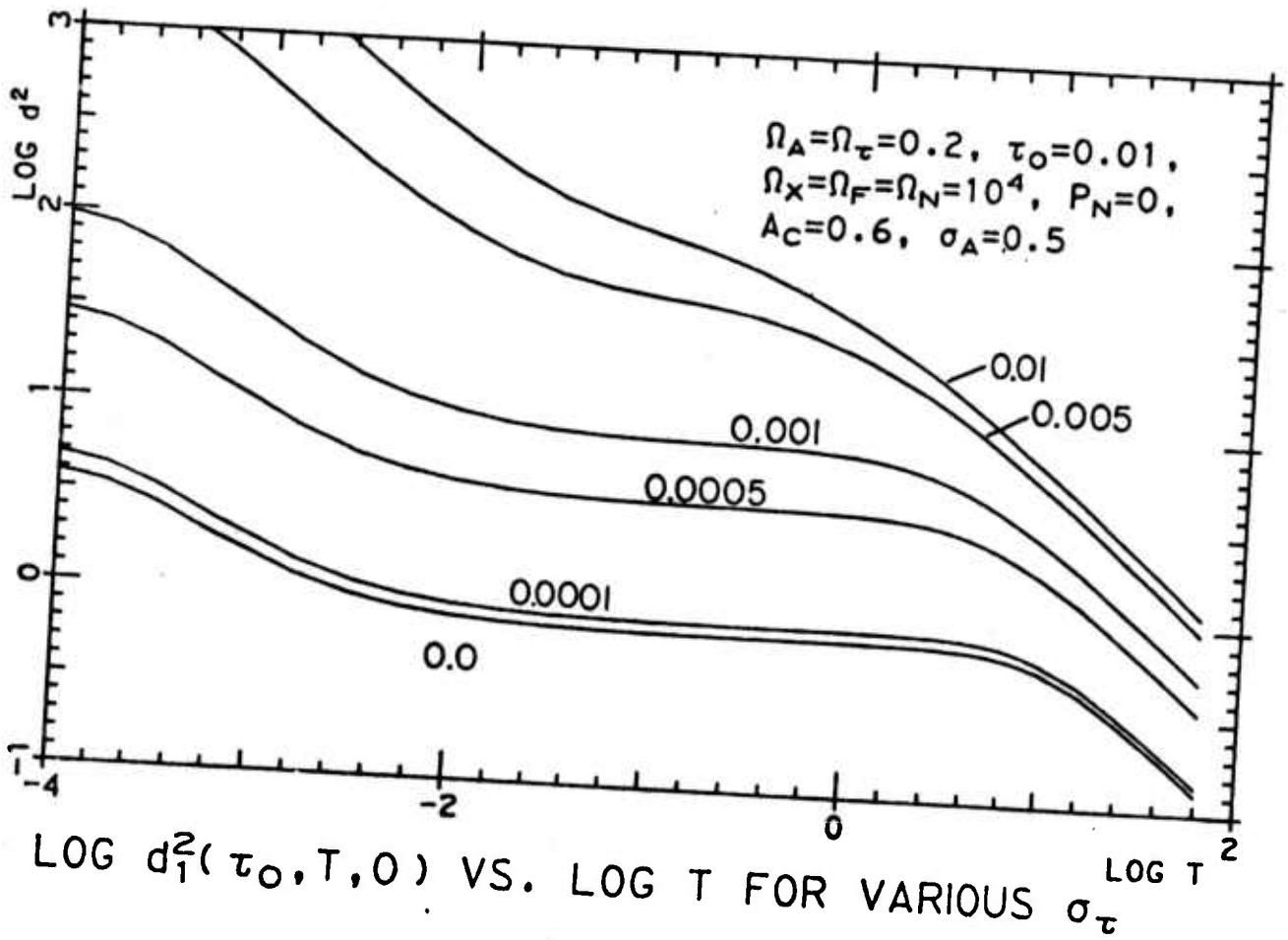
For example, Figure 4.2-1 displays $d_1^2(\tau_o,T,0)$ for $\sigma_\tau = 0$ and for various values of σ_A using typical values for other relevant parameters. Figure 4.2-2 shows $d_1^2(\tau_o,T,0)$ with $\sigma_A = 0.5$ while σ_τ is varied. In either example it is seen that as the integration time T is increased from 0 normalized variance at first decreases somewhat. However, this initial progress is halted just when reliable correlation begins to emerge for certain realizations of the random scattering system. At this point large uncertainties exist in the correlator output which reflect the variability of the scattering. These uncertainties show no reduction with increasing T until T becomes comparable with the time constants of the scattering mechanisms. Once T is increased beyond this point the normalized variance again decreases rapidly. It is during this period that the correlator averages over different scattering ensemble members.

The covariance $d_1(\tau_o,T,\mu)$ of the correlator output $\Xi(t,T,\rho)$ at two different times ρ and $\rho + \mu$ is plotted in Figure 4.2-3.



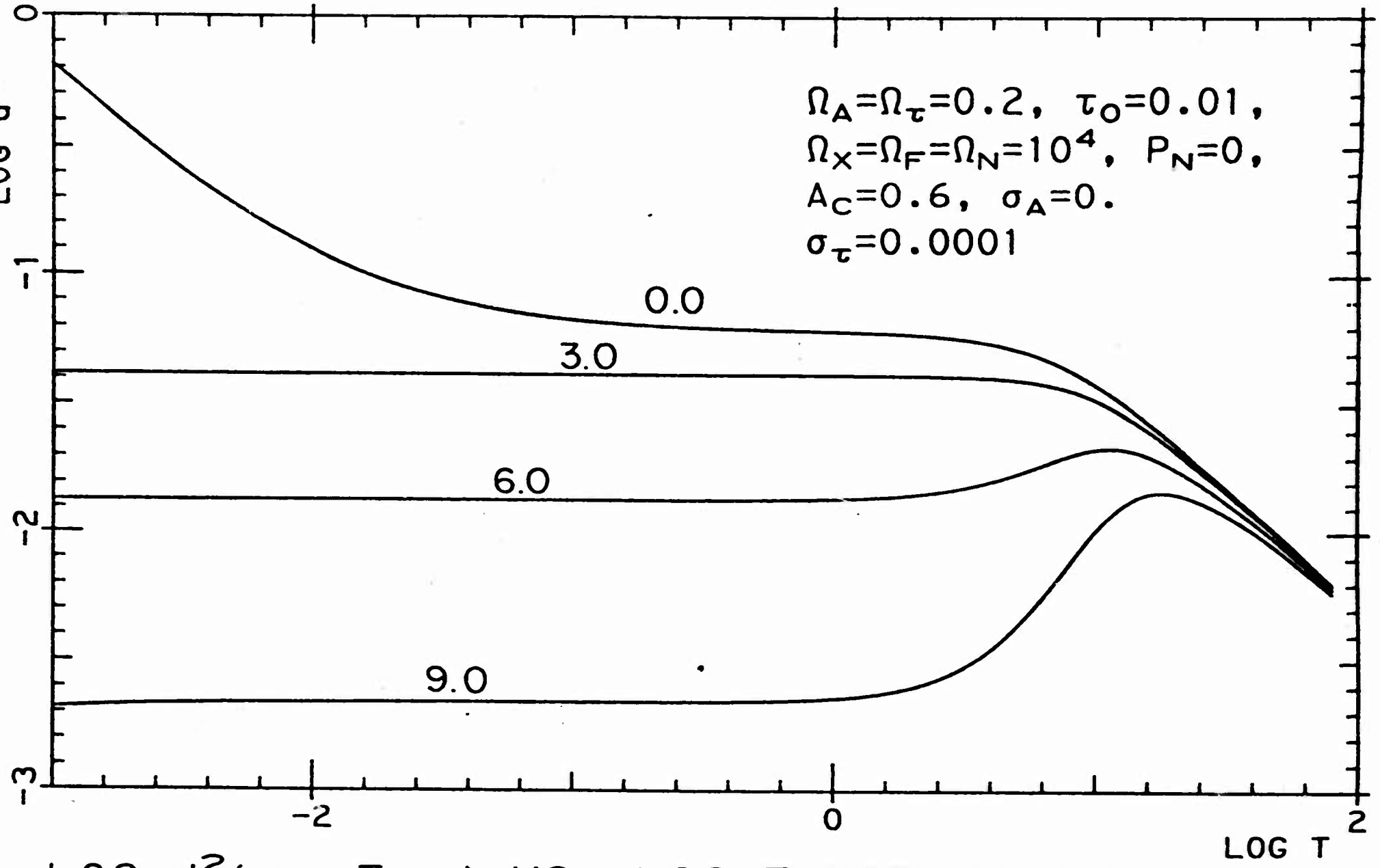
LOG $d_1^2(\tau_o, T, 0)$ VS. LOG T FOR VARIOUS σ_A

Fig. 4.2-1



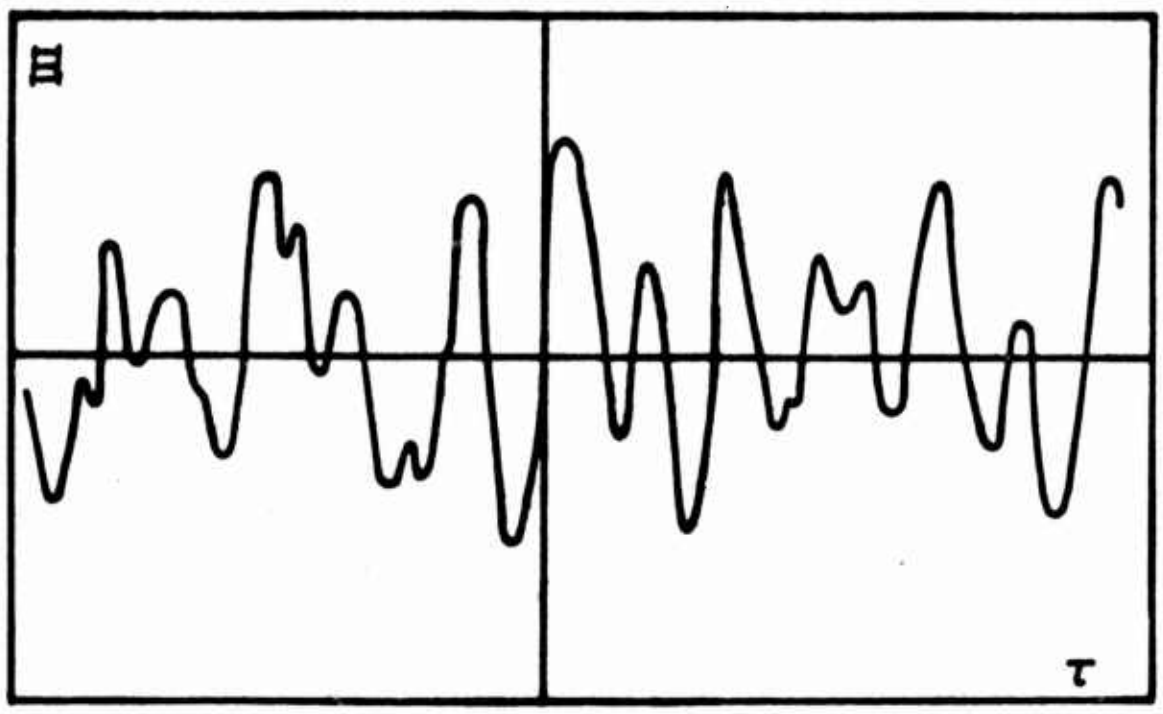
LOG $d_1^2(\tau_o, T, 0)$ VS. LOG T FOR VARIOUS σ_τ

Fig. 4.2-2



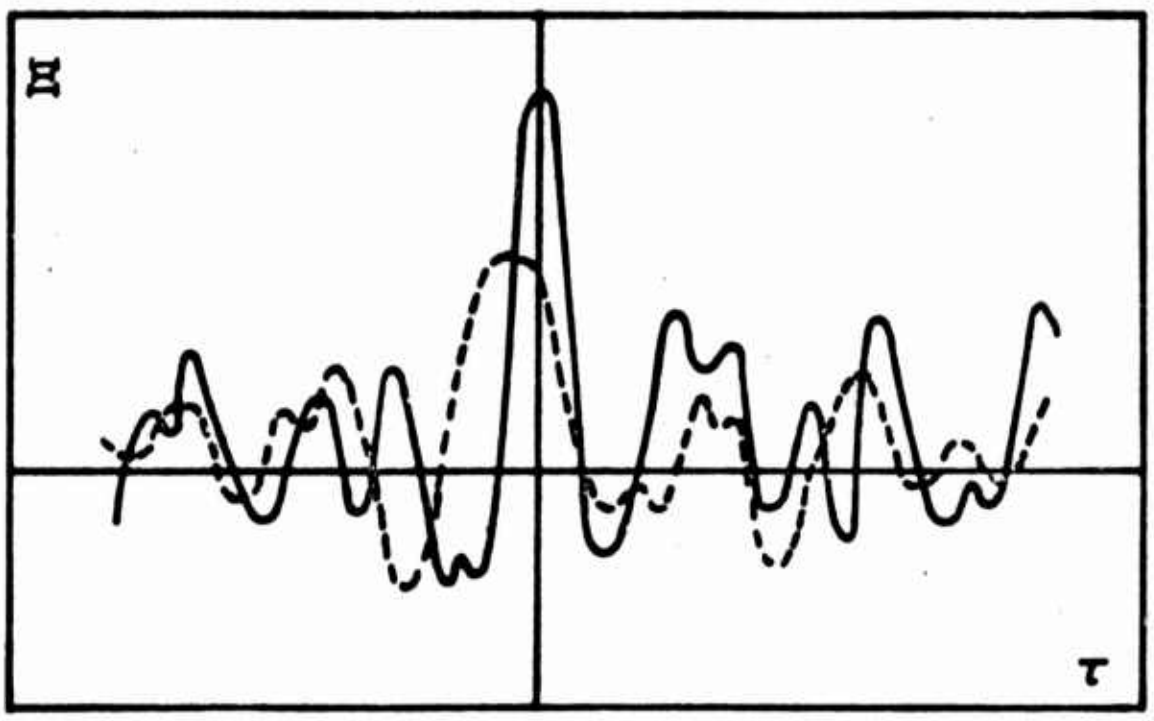
LOG $d_1^2(\tau_o, T, \mu)$ VS. LOG T FOR VARIOUS μ

Fig. 4.2-3



SAMPLE OF Ξ FOR SMALL T

Fig. 4.2-4a



SAMPLES OF Ξ FOR LARGE T

Fig. 4.2-4b

It can be seen that the correlator output for two different starting times ρ and $\rho+\mu$ rapidly becomes decorrelated for small T. As T is increases to the point where discernable correlation begins to emerge, the fluctuations of the output become extremely persistent. A second sample of Ξ initiated at $\rho+\mu$ will not exhibit independent fluctuations since the scattering systemhas not changed substantially. Sample correlator outputs in these two regions of behavior are shown in Figure 4.2-4. For this range of T the fluctuation of the correlator output appears to be localized in the vicinity of τ_o. This may be verified by considering $d_1^2(\tau,T,\mu)$ for $\tau \neq \tau_o$ in Equation (4.2-8). For much larger values of T the fluctuations will be still more persistent but smaller magnitude as more variations of the scattering are included in the processing interval.

A careful examination of the terms in Equation (4.2-8) shows that most important contribution to the output variance in the region of persistent fluctuation arises from the two integrals $I_4(\tau,\tau',\nu)$ and $I_5(\tau,\tau')$ in (3.3-8). The magnitude of this plateau of uncertainty for the model we are considering is approximately

$$\left[1+\left(\frac{\sigma_A^2}{A_c^2}\right)\right] \frac{\alpha_m^2}{\sqrt{\alpha_m^4 - \sigma_\tau^4}} - 1 \qquad (4.2\text{-}11)$$

This is alsc approximately true in cases of high to moderate signal to noise ratios. In Figure 4.2-5 $d_1^2(\tau_o,T,0)$ is plotted for various signal to noise ratios (P_n/P_x). Only in the case of very weak ratios does the plateau become masked by noise. In Figure 4.2-6 $d_o^2(\tau_o,T,0)$ is plotted for the same sequence of signal to noise ratios. In roughly

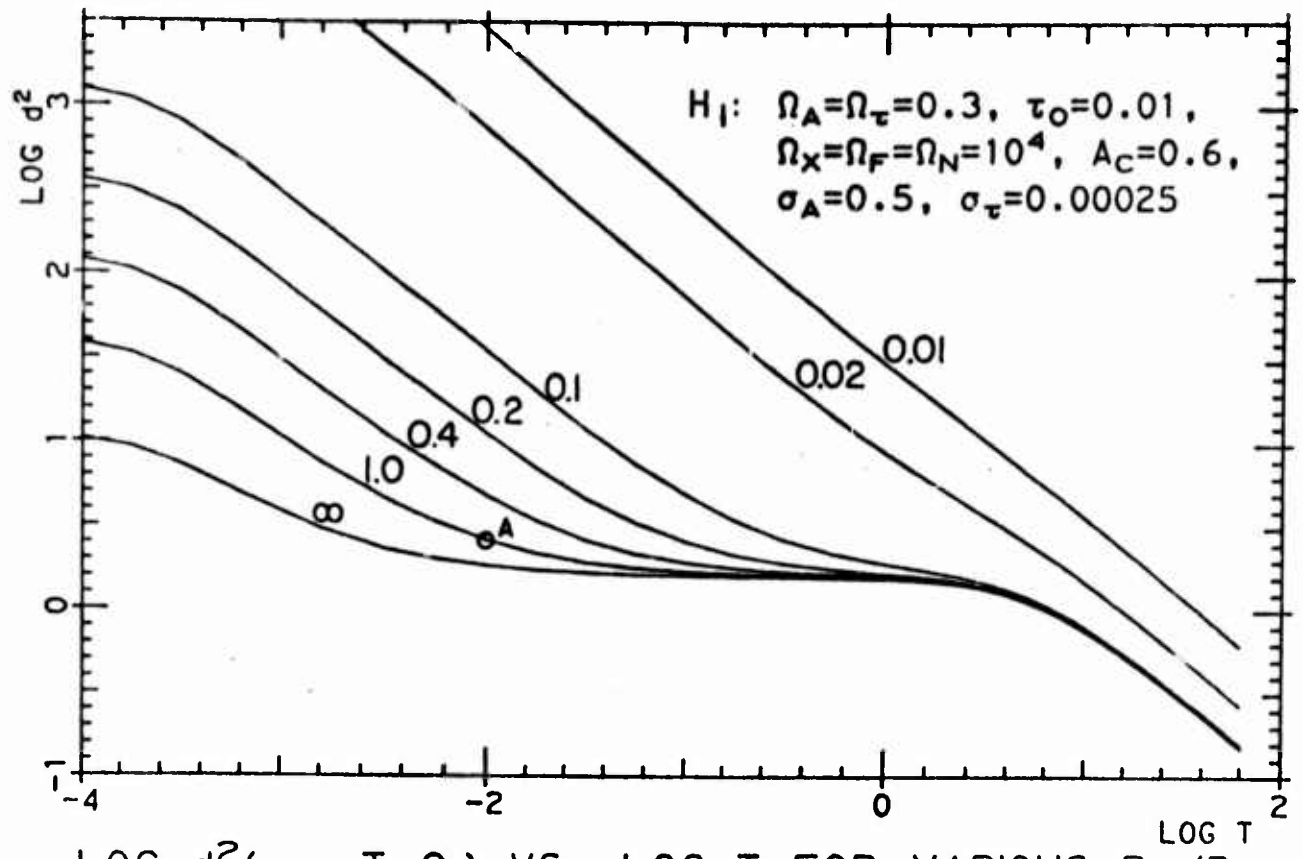
the same range of T that (4.2-11) is valid we have approximately

$$d_o(\tau_o,T,0) \rightarrow \frac{1}{T}\,\frac{1}{\Omega_m^2}\left(\frac{P_n}{\Omega_n}\,\frac{\Omega_x}{P_x}\right)^2 \frac{\sqrt{2}}{A_d^2 A_c^2}\,\frac{\Omega_{1n1}\,\Omega_{2n2}}{\sqrt{\Omega_{1n1}^2+\Omega_{2n2}^2}} \qquad (4.2\text{-}12)$$

Assuming that Ξ tends to be Gaussian we may use these curves for d_o^2 and d_1^2 to compute for a given T the receiver operating characteristic (a plot of α vs. β parameterized by K_{th}). Figure 4.2-7 gives one such curve corresponding to the circled points in Figures 4.2- 5and 4.2-6. Also given for comparison is the related curve for the case of no scattering (obtained by setting $\sigma_\tau = \sigma_A = 0$ and $A_c = 1$). In addition, the corresponding curves for these two cases are plotted for a one sided detector.

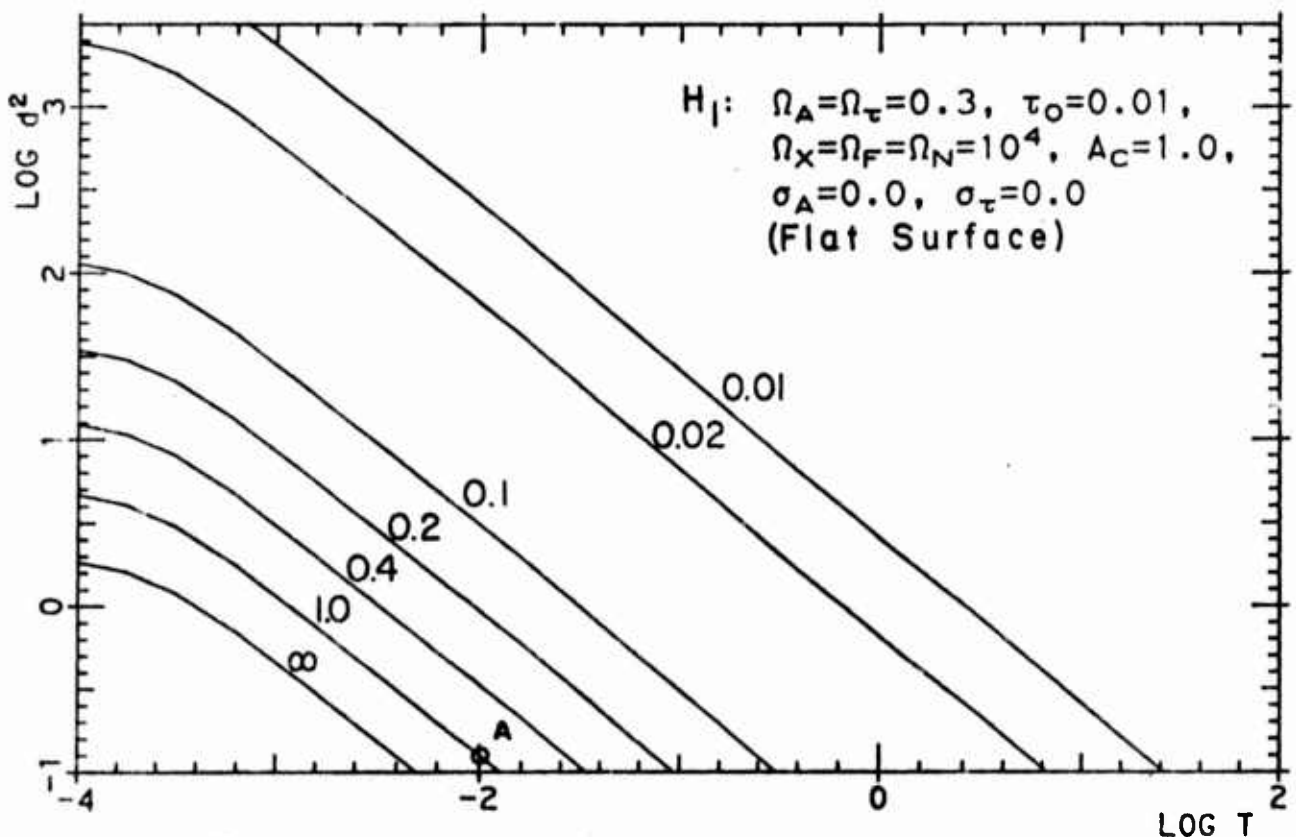
It is clear that with the choice of parameter values selected for these curves the scattering significantly increases the false dismissal probability β for any fixed false alarm probability α. It can also be seen that under the assumption that Ξ is Gaussian the one sided detector produces still larger values of β. Furthermore, these latter increases result in a receiver operating characteristic which is not everywhere convex upward. This is clear manifestation that the one sided test departs markedly from the likelihood strategy in such situations. However, it is important to stress that this result is due to the assumption of symmetry in $f_o(\Xi)$ and $f_1(\Xi)$ under the hypothesis that these functions are Gaussian. Given that A(t)is itself Gaussian this hypothesis is not unreasonable.

We turn briefly to the question of parameter optimization in the adopted receiver design. As might be expected, the rather



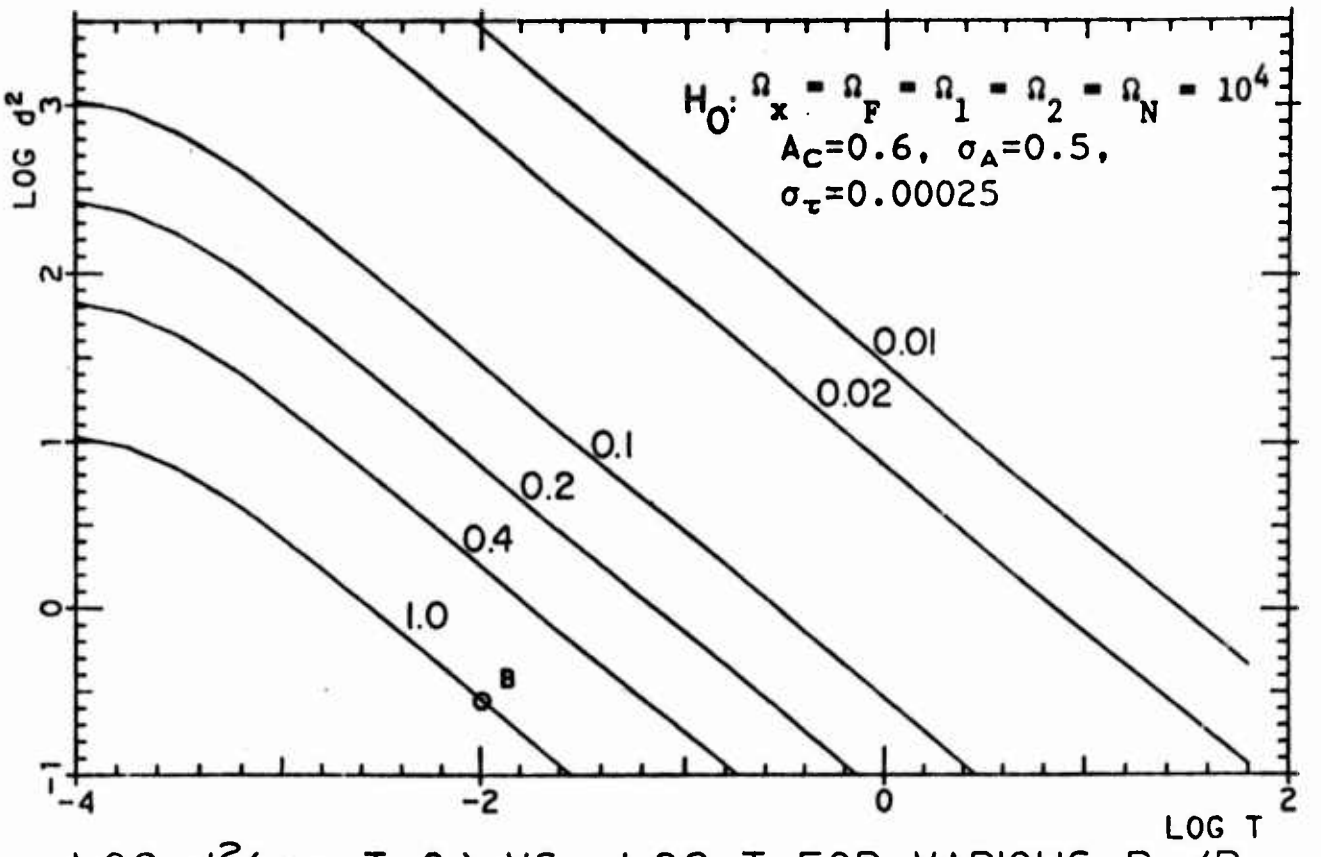
LOG $d_1^2(\tau_o, T, 0)$ VS. LOG T FOR VARIOUS P_X/P_N

Fig. 4.2-5a



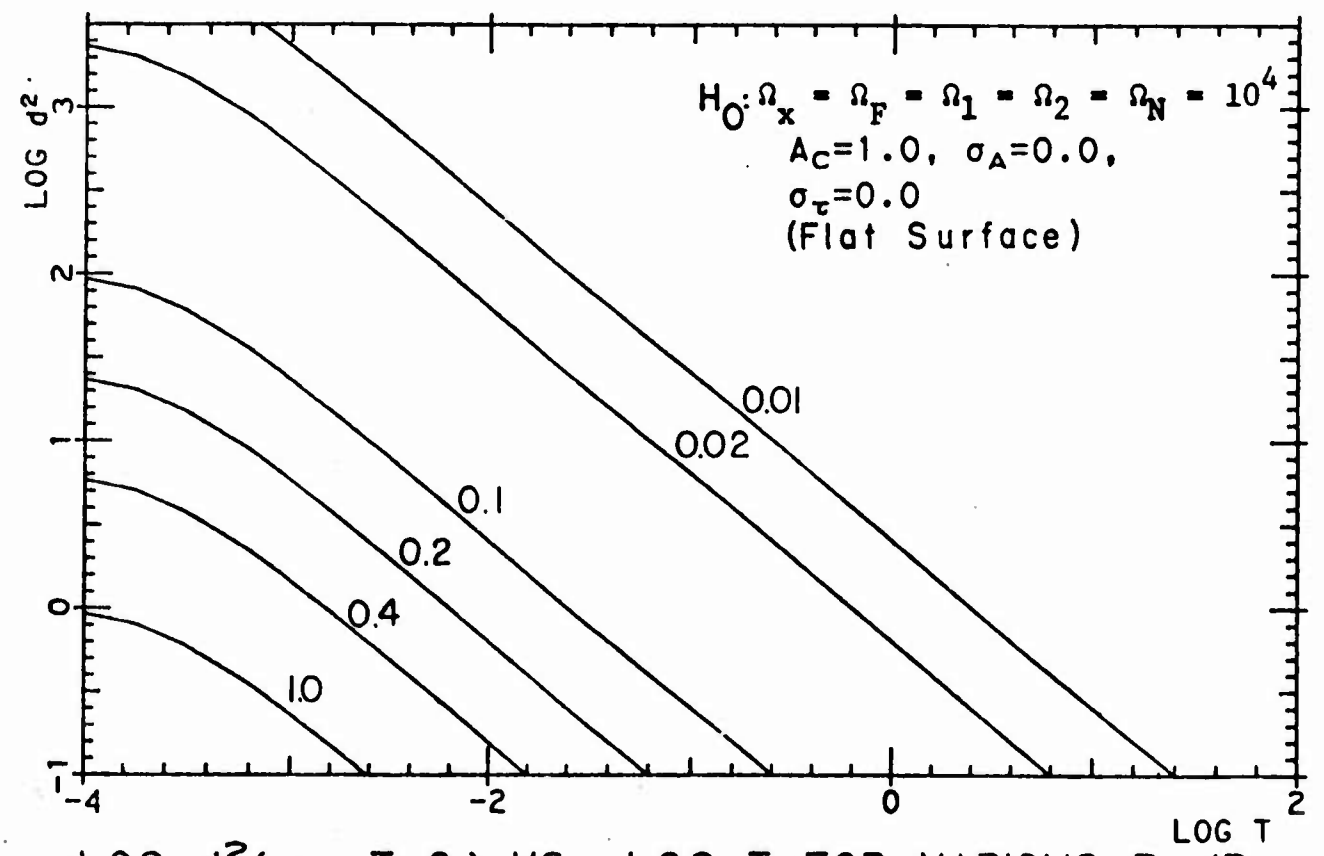
LOG $d_1^2(\tau_o, T, 0)$ VS. LOG T FOR VARIOUS P_X/P_N

Fig. 4.2-5b



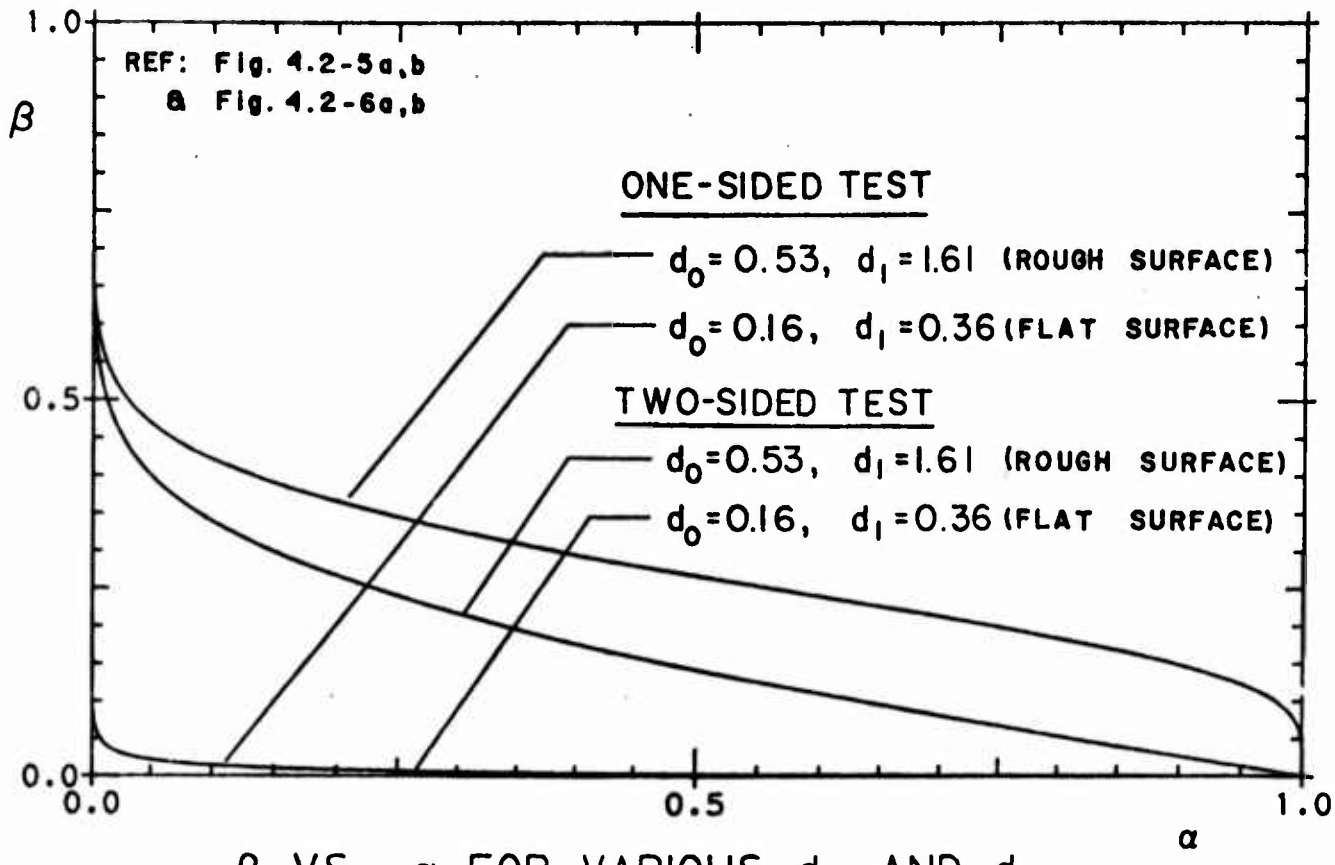
LOG $d_0^2(\tau_0, T, 0)$ VS. LOG T FOR VARIOUS P_X/P_N

Fig. 4.2-6a



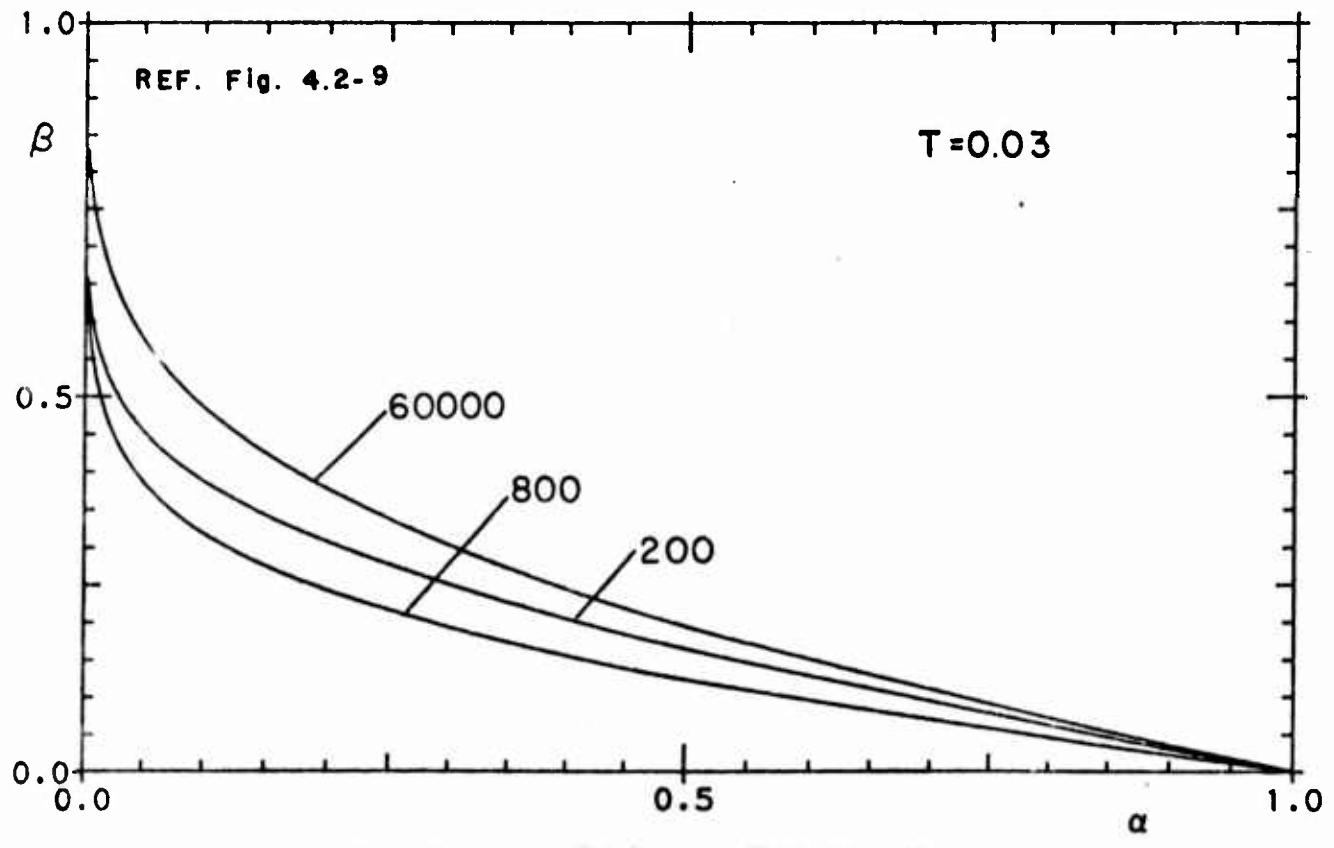
LOG $d_0^2(\tau_0, T, 0)$ VS. LOG T FOR VARIOUS P_X/P_N

Fig. 4.2-6b



β VS. α FOR VARIOUS d_0 AND d_1

Fig. 4.2-7

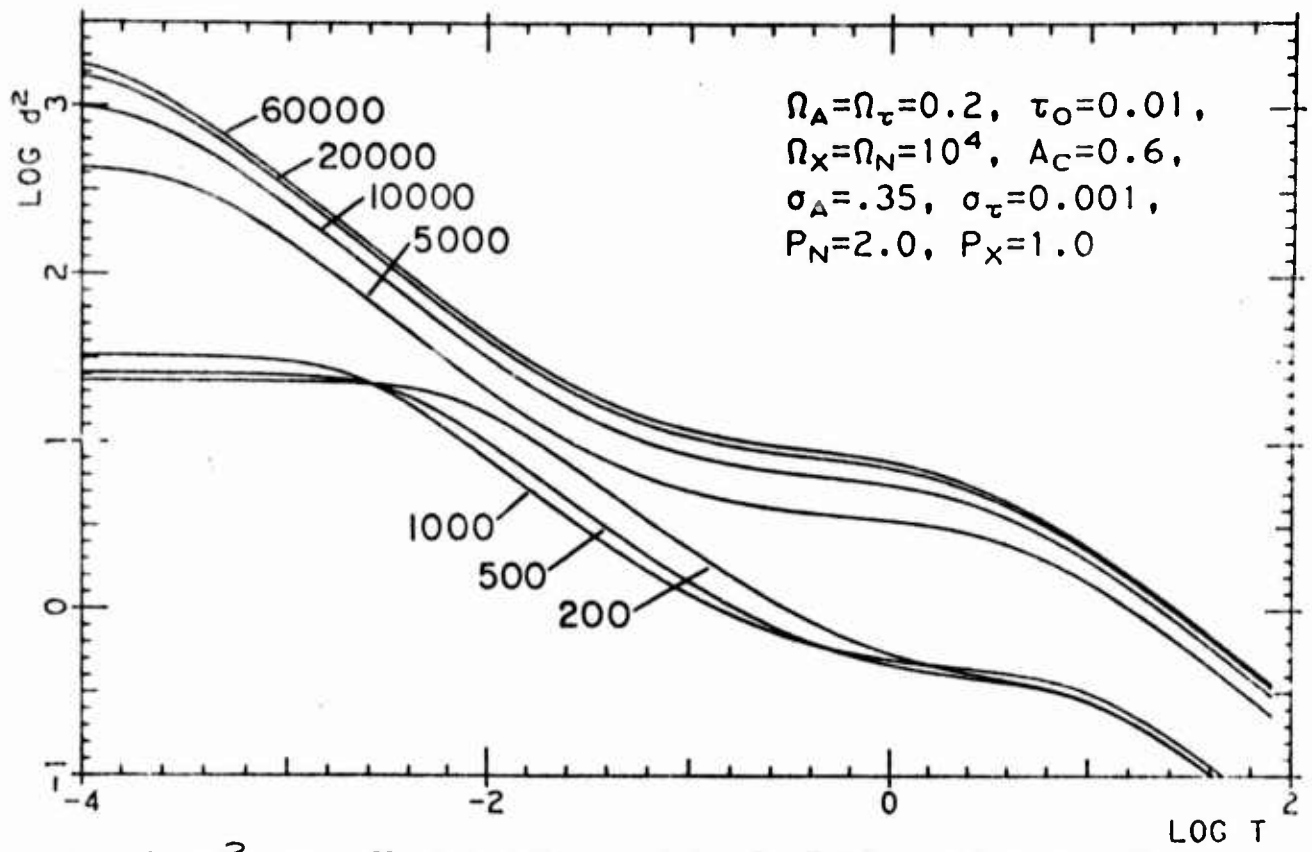


β VS. α FOR VARIOUS Ω_F

Fig. 4.2-8

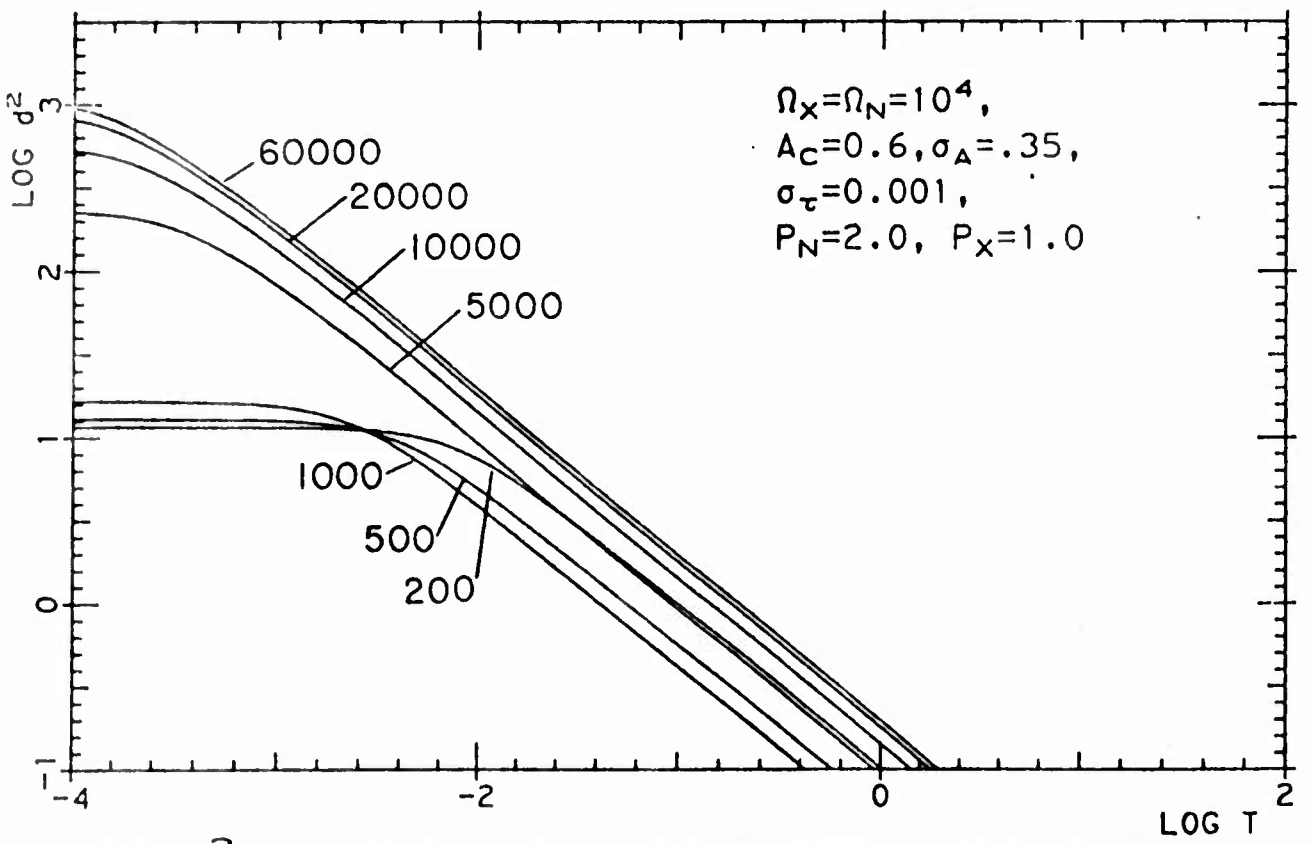
simplistic model of (4.0-1) presents us with few alternatives. For example, it is clear that there is nothing one can do to filter out the effect the slow amplitude variations A(t). This might not be the case were we to assume A(t) to be frequency dependent, but we choose to ignore such effects in this chapter. On the other hand, assuming that our purpose is to improve "on-target" detectability, it is clear that the plateau of normalized variance in (4.2-11) can be reduced by decreasing the working bandwidth Ω_F until $\alpha_m >> \sigma_\tau$. This strategy effectively screens out the incoherent delay modulated signal fluctuations. The receiver operating characteristics of Figure 4.2-8 demonstrate , however, that if Ω_F is made too small, performance can actually worsen. Figure 4.2-9 illustrates how in practice the plateau of Equation (4.2-11) can become masked by the other terms in Equation (4.2-8) before the theoretical minimum of σ_A^2/A_c^2 can be achieved for d_1^2.

Within the limitation just described it is seen that the strategy of screening out incoherent signal definitely leads to an improvement of detectability when the receiver is constrained to be steered on target. When this constraint is removed (see Appendix D) a certain information content is found to exist in the incoherent scattered signal which is due to the non-gaussian nature of this signal. The unconstrained receiver improves performance by attempting to estimate the location of the delay modulated peak of correlation. The magnitude of this peak is always greater than or equal to the magnitude of Ξ at τ_o.



LOG $d_1^2(\tau_o, T, 0)$ VS. LOG T FOR VARIOUS Ω_F

Fig. 4.2-9a



LOG $d_o^2(\tau_o, T, 0)$ VS. LOG T FOR VARIOUS Ω_F

Fig. 4.2-9b

Although some of the incoherent scattered signal contains information useful for detection purposes, clearly none of it is useful for tracking in this case. The covariance $R_{\tau_o}(\mu)$ of the peak location estimate is obviously a suitable, straightforward measure of the magnitude and persistence of traking errors. By applying the results of Appendix E.1 to Equation (3.6-3) we have

$$R_{\tau_o}(\mu) = \frac{(\Omega_x/\Omega_m^3)^2}{T^2 A_d^2 A_c^2} \left[\int_{\mu-T}^{\mu} [T+\nu-\mu] + \int_{\mu}^{\mu+T} [T-\nu+\mu] \right] \Bigg(\left[\left(\frac{P_{n1}}{P_x} \right) \left(\frac{\Omega_{1n1}}{\Omega_{n1}} \right) e^{-1/2\nu^2\Omega_{1n1}^2} \right.$$

$$+ \frac{R_{AA}(\nu)}{\Omega_x \sqrt{\frac{1}{\Omega^2} + 2[\sigma_\tau^2 - R_{\tau\tau}(\nu)]}} \exp\left\{ \frac{-\nu^2}{2} \left[\frac{1}{\Omega_{1x}^2} + 2[\sigma_\tau^2 - R_{\tau\tau}(\nu)]^{-1} \right] \right\} \Bigg] x$$

$$\left[\left(\frac{P_{n2}}{P_x} \right) \left(\frac{\Omega_{2n2}^3}{\Omega_{n2}} \right) [\nu^2\Omega_{2n2}^2 - 1] e^{-1/2\nu^2\Omega_{2n2}^2} \right. \qquad (4.2\text{-}13)$$

$$\left. + A_d^2 \left(\frac{\Omega_{1x}^3}{\Omega_x} \right) [\nu^2\Omega_{1x}^2 - 1] e^{-1/2\nu^2\Omega_{1x}^2} \right] (-1)$$

$$+ \frac{A_d^2 R_{AA}(\nu)}{\Omega_x^2 (\sqrt{\alpha_m^4 - R_{\tau\tau}^2(\nu)})^3} \left[R_{\tau\tau}(\nu) + \right.$$

$$\left. \exp(-[\alpha_m^2 - R_{\tau\tau}(\nu)]^{-1} \nu^2) \{ R_{\tau\tau}(\nu) - \nu^2 [\alpha_m^2 + R_{\tau\tau}(\nu)]^2 \} \right] \Bigg)$$

As can be seen from equation 4.2-13 the covariance of the tracking error exhibits a plateau of uncertainty of approximately

$$\left(1 + \frac{\sigma_A^2}{A_c^2}\right)\sigma_\tau^2 \quad \frac{a_m^6}{\left(\sqrt{a_m^4 - \sigma_\tau^4}\right)^3} \qquad (4.2\text{-}14)$$

which is also reduced by increasing a_m until it is much greater than σ_τ yielding a minimum of $(1 + \sigma_A^2/A_c^2)\ \sigma_\tau^2$.

4.3 Second and Fourth Order Cross System Statistics for the Two Channel Random Amplitude and Delay Model

In this section we consider some of the joint statistics for the pair of channels described by the random amplitude and delay system functions :

$$H_1(\omega,t) = A_1(t)\, e^{-j\omega\tau_1(t)} \quad ; \quad H_2(\omega,t) = A_2(t)\, e^{-j\omega\tau_2(t)} \qquad (4.3\text{-}1)$$

where $A_1(t)$ and $\tau_1(t)$ are considered jointly independent, as are $A_2(t)$ and $\tau_2(t)$. However, $A_1(t)$ and $A_2(t)$ are assumed jointly dependent, as are $\tau_1(t)$ and $\tau_2(t)$. All four variables are taken to be jointly Gaussian with the means

$$\overline{A_1(t)} = \overline{A_2(t)} = A_c \quad ; \quad \overline{\tau_1(t)} = \tau_{s1} \quad ; \quad \overline{\tau_2(t)} = \tau_{s2} \qquad (4.3\text{-}2)$$

We again define correlation functions for these parameters :

$$R_{AA}(\mu) = \overline{A_1(t)\, A_1(t-\mu)} = \overline{A_2(t)\, A_2(t-\mu)} \qquad (4.3\text{-}3)$$

$$R_{\tau\tau}(\mu) = \overline{\tau_1(t)\, \tau_1(t-\mu)} - \tau_{s1}^2 = \overline{\tau_2(t)\, \tau_2(t-\mu)} - \tau_{s2}^2 \qquad (4.3\text{-}4)$$

$$R_{Ac}(\mu) = A_1(t)\, A_2(t-\mu) \qquad (4.3\text{-}5)$$

$$R_{\tau c}(\mu) = \overline{\tau_1(t)\, \tau_2(t-\mu)} - \tau_{s1}\tau_{s2} \qquad (4.3\text{-}6)$$

Once again R_{AA} and R_{Ac} are defined as non-central moments for convenience. Assuming stationarity for these parameters we have the following system correlation functions :

$$\Phi_{jj}(\omega,\omega',\mu) = R_{AA}(\mu)\, \exp\{-1/2[(\omega^2+\omega'^2)\sigma_{\tau a}^2 - 2\omega\omega' P_{\tau a}(\mu)] - j(\omega-\omega')\tau_{s_j}\} \qquad (4.3\text{-}7)$$

for j = 1 and 2, and

$$\Phi_{12}(\omega,\omega',\mu) = R_{A_c}(\mu)\ \exp\{-1/2[(\omega^2+\omega'^2)\sigma^2_{\tau_a} - 2\omega\omega' R_{\tau c}(\mu)] - j(\omega\tau_{s1}-\omega'\tau_{s2})\} \tag{4.3-8}$$

where $\sigma^2_{\tau a} = R_{\tau A}(0)$. Obviously, the second order stationarity of relations (4.3-3) through (4.3-6) results in interfrequency wide sense cross stationary channels (IWSCS). This taken together with the assumption that the amplitudes and delays are Gaussian processes implies that the two channels are also jointly cross fourth order interfrequency stationary (CFOIS). Thus,

$$\Phi^{[4]}_{1,2}(\omega,\omega'',\omega''',\mu,\mu,'\mu'') = R^{[4]}_{A_1A_2}(\underline{\mu})\exp\{-1/2\underline{\omega}^T\ \underline{R}_{\tau_1\tau_2}(\underline{\mu})\underline{\omega} - \underline{\omega}^T\ \underline{\tau}s\ \} \tag{4.3-9}$$

where

$$\underline{\mu}^T = (\mu,\mu',\mu'') \tag{4.3-10}$$

$$\underline{\omega}^T = (\omega,-\omega',\omega'',-\omega''') \tag{4.3-11}$$

$$\underline{\tau}_s^T = (\tau_{s1},+\tau_{s1},\tau_{s2},+\tau_{s2}) \tag{4.3-12}$$

$$\underline{R}_{\tau_1\tau_2}(\underline{\mu}) = \begin{bmatrix} R_{\tau a}(0) & R_{\tau a}(\mu) & R_{\tau c}(\mu') & R_{\tau c}(\mu'') \\ R_{\tau a}(\mu) & R_{\tau a}(0) & R_{\tau c}(\mu'-\mu) & R_{\tau c}(\mu''-\mu) \\ R_{\tau c}(\mu') & R_{\tau c}(\mu'-\mu) & R_{\tau a}(0) & R_{\tau a}(\mu''-\mu') \\ R_{\tau c}(\mu'') & R_{\tau c}(\mu''-\mu) & R_{\tau a}(\mu''-\mu') & R_{\tau a}(0) \end{bmatrix} \tag{4.3-13}$$

$$R^{[4]}_{A_1A_2}(\underline{\mu}) = \overline{A_1(t)\ A_1(t-\mu)\ A_2(t-\mu')\ A_2(t-\mu'')} \tag{4.3-14}$$

As special cases we have

$$\Phi^{[4]}_{1,2}(\omega,\omega,\omega',\omega',\nu,\tau,\nu+\tau') = R^{[4]}_{A_1A_2}(\nu,\tau,\nu+\tau') \quad x$$

$$\exp\{-1/2[\alpha_{11}(\nu,\tau,\tau')\omega^2+2\alpha_{12}(\nu,\tau,\tau')\omega\omega'+\alpha_{13}(\nu,\tau,\tau')\omega'^2]\} \tag{4.3-15}$$

$$\Phi^{[4]}_{1,2}(\omega,-\omega',-\omega,\omega',\nu,\tau,\nu+\tau') = R^{[4]}_{A_1A_2}(\nu,\tau,\nu+\tau') \quad x$$

$$\exp\{-1/2[\alpha_{21}(\nu,\tau,\tau')\omega^2+2\alpha_{22}(\nu,\tau,\tau')\omega\omega'+\alpha_{23}(\nu,\tau,\tau')\omega'^2]\} \quad x \tag{4.3-16}$$

$$\exp\{-j(\omega+\omega')(\tau_{s1}-\tau_{s2})\}$$

$$\Phi^{[4]}_{1,2}(\omega,\omega',\omega'\omega,\nu,\tau,\nu+\tau') = R^{[4]}_{A_1A_2}(\nu,\tau,\nu+\tau') \quad x$$

$$\exp\{-1/2[\alpha_{31}(\nu,\tau,\tau')\omega^2+2\alpha_{32}(\nu,\tau,\tau')\omega\omega'+\alpha_{33}(\nu,\tau,\tau')\omega'^2]\} \quad x \tag{4.3-17}$$

$$\exp\{-j(\omega-\omega')(\tau_{s1}-\tau_{s2})\}$$

where

$$\alpha_{11}(\nu,\tau,\tau') = 2[R_{\tau a}(0) - R_{\tau a}(\nu)] \tag{4.3-18}$$

$$2\alpha_{12}(\nu,\tau,\tau') = 2[R_{\tau c}(\tau) + R_{\tau c}(\tau')] \tag{4.3-19}$$

$$- 2[R_{\tau c}(\tau-\nu) + R_{\tau c}(\nu+\tau')]$$

$$\alpha_{13}(\nu,\tau,\tau') = 2[R_{\tau a}(0) - R_{\tau a}(\tau-\nu-\tau')] \tag{4.3-20}$$

$$\alpha_{21}(\nu,\tau,\tau') = 2[R_{\tau a}(0) - R_{\tau c}(\tau)] \tag{4.3-21}$$

$$2\alpha_{22}(\nu,\tau,\tau') = 2[R_{\tau a}(\nu)+R_{\tau a}(\tau-\nu-\tau')]$$

$$- 2[R_{\tau c}(\nu+\tau') + R_{\tau c}(\tau-\nu)] \tag{4.3-22}$$

$$\alpha_{23}(\nu,\tau,\tau') = 2[R_{\tau a}(0) - R_{\tau c}(\tau')] \tag{4.3-23}$$

$$\alpha_{31}(\nu,\tau,\tau') = 2[R_{\tau a}(0) - R_{\tau c}(\nu+\tau')] \tag{4.3-24}$$

$$2\alpha_{32}(\nu,\tau,\tau') = -2[R_{\tau a}(\tau-\nu-\tau')] \tag{4.3-25}$$

$$+ 2[R_{\tau c}(\tau) + R_{\tau c}(\tau')]$$

$$\alpha_{33}(\nu,\tau,\tau') = 2[R_{\tau a}(0) - R_{\tau c}(\tau-\nu)] \tag{4.3-26}$$

These latter parameters satisfy the symmetry relations

$$\alpha_{11}(\nu,\tau_o,\tau_o) = \alpha_{13}(\nu,\tau_o,\tau_o) \tag{4.3-27}$$

$$\alpha_{21}(\nu,\tau_o,\tau_o) = \alpha_{23}(\nu,\tau_o,\tau_o) \tag{4.3-28}$$

Finally, under the Gaussian hypothesis for A_1 and A_2 we have

$$R^{[4]}_{A_1A_2}(\nu,\tau,\nu+\tau') = \tag{4.3-29}$$

$$[R_{A_c}(\tau)\, R_{A_c}(\tau') + R_{A_a}(\tau)\, R_{A_a}(\nu+\tau'-\tau) + R_{Ac}(\nu+\tau')\, R_{Ac}(\nu-\tau)] \cdot$$

$$+ A_c^2 [\, R_{Aa}(\nu+\tau'-\tau) + R_{Ac}(\tau') + R_{Ac}(\nu+\tau')$$

$$+ R_{Aa}(\nu) + R_{Ac}(\tau) + R_{Ac}(\nu-\tau)] + A_c^4$$

4.4 Array Correlator Fluctuations for the Two Channel Random Amplitude and Delay Model

Once again we compute the normalized output variances as in section 4.2. We consider the spectra (4.2-2) and filters (4.2-3) subject to the symmetry assumptions (4.2-10) and introduce the noise cross spectral density

$$S_{n_1 n_2}(\omega) = \frac{\sqrt{2\pi}\ P_{nc}}{\Omega_{nc}}\ e^{-\frac{1}{2}\omega^2/\Omega_{nc}^2} \tag{4.4-1}$$

The mean output of the correlator is then

$$\Xi(\tau,T,\rho) = P_{nc}\ \left(\frac{\Omega_{fnc}}{\Omega_{nc}}\right) e^{-\tau^2\Omega^2_{fnc}} \tag{4.4-2}$$

$$+ \frac{R_{Ac}(\tau)\ P_x}{\Omega_x\sqrt{\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau c}^2 - R_{\tau c}(\tau)\ ]}}\ \exp\{-\frac{(\tau-\tau_o)^2}{2}\ [\ \frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau c}^2 - R_{\tau c}(\tau)]\ ]^{-1}\}$$

Assuming that $\tau_o^2\Omega_{fnc}^2 >> 1$ this mean exhibits a peak at $\tau=\tau_o$ of

$$\frac{R_{Ac}(\tau_o)\ P_x}{\Omega_x\sqrt{\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau c}^2 - R_{\tau_c}(\tau_o)]}}$$

Thus, to this level of approximation using the results of Appendix E.2

$$d_o^2(\tau_o,T,\mu) \approx \frac{\Omega_x^2\left[\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau c}^2 - R_{\tau c}(\tau_o)]\right]}{T^2\ R_{Ac}^2(\tau_o)}\ \left(\frac{P_{na}P_{nc}\Omega_{fna}\Omega_{fnc}}{P_x^2\ \Omega_f^2}\right) \tag{4.4-3}$$

$$\left[\int_{\mu-T}^{\mu}[T+\nu-\mu] + \int_{\mu}^{\mu+T}[T-\nu+\mu]\right][\exp\{-\nu^2\Omega_{fna}^2\} + \exp\{-(\nu^2+\tau_o^2)\Omega_{fnc}^2\}]\ d\nu$$

and

$$d_1^2(\tau,T,\mu) \approx \frac{\Omega_x^2 \left[\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau c}^2 - R_{\tau c}(\tau)] \right]}{T^2 R_{Ac}(\tau_o)}$$

$$\left[\int_{\mu-T}^{\mu} [T+\nu-\mu] + \int_{\mu}^{\mu+T} [T-\nu+\mu] \right] \frac{P_{na}^2}{P_x^2} \frac{\Omega_{fna}^2}{\Omega_x^2} \exp\{-\nu^2\Omega_{fna}^2\}$$

$$+ \frac{2 R_{Aa}(\nu) (P_{na}/P_x) \Omega_{fna}}{\Omega_{na} \Omega_x \sqrt{\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau c}^2 - R_{\tau c}(\nu)]}}$$

$$\exp\left\{ \frac{-\nu^2}{2} \left[\Omega_{fna}^2 + \left[\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau a}^2 - R_{\tau a}(\nu)] \right]^{-1} \right] \right\} \qquad (4.4\text{-}4)$$

$$+ \left(\frac{P_{ca}^2}{P_x^2}\right) \left(\frac{\Omega_{fnc}^2}{\Omega_f^2}\right) \exp\{-(\nu^2+\tau_o^2)\Omega_{fnc}^2\} +$$

$$\frac{\left(\frac{P_{nc}}{P_x}\right) \Omega_{fnc} R_{Ac}(\tau_o-\nu) e^{-\frac{1}{2}(\tau_o+\nu)^2\Omega_{fnc}^2}}{\Omega_{nc}\Omega_x \sqrt{\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau c}^2 - R_{\tau c}(\tau_o-\nu)]}} + \frac{\left(\frac{P_{nc}}{P_x}\right) \Omega_{fnc} R_{Ac}(\tau_o+\nu) e^{-\frac{1}{2}(\tau_o-\nu)^2\Omega_{fnc}^2}}{\Omega_{nc}\Omega_x \sqrt{\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau c}^2 - R_{\tau c}(\tau_c+\nu)]}}$$

$$+ \frac{1}{\Omega_x^2} R^{[4]}(\tau_o,\nu,\nu+\tau_o) \left\{ \frac{\exp\left\{ \frac{-\nu^2}{\alpha_{11}'(\nu,\tau_o,\tau_o) + \alpha_{12}(\nu,\tau_o,\tau_o)} \right\}}{\sqrt{\alpha_{11}'(\nu,\tau_o,\tau_o)^2 - \alpha_{12}(\nu,\tau_o,\tau_o)^2}} \right.$$

$$+ \frac{1}{\sqrt{\alpha'_{21}(\nu,\tau_o,\tau_o)^2 - \alpha_{22}(\nu,\tau_o,\tau_o)^2}}$$

$$\exp\left\{\frac{-\frac{1}{2}\nu^2[\alpha'_{31}(\nu,\tau_o,\tau_o) - 2\alpha_{32}(\nu,\tau_o,\tau_o) + \alpha'_{33}(\nu,\tau_c,\tau_o)]}{\alpha'_{31}(\nu,\tau_o,\tau_o)\,\alpha'_{33}(\nu,\tau_o,\tau_o) - \alpha_{32}(\nu,\tau_o,\tau_o)^2}\right\}$$

$$\Bigg\}\Bigg\}\, d\nu^{-1}$$

$$\sqrt{\alpha'_{31}(\nu,\tau_o,\tau_o)\,\alpha'_{33}(\nu,\tau_o,\tau_o) - \alpha_{32}(\nu,\tau_o,\tau_o)^2}$$

where the α' are defined in Appendix E.2 and use has been made of symmetry relations 4.3-27 and 4.3-28.

The plateau of uncertainty in $d_1^2(\tau_o,T,0)$ for this case is roughly

$$\frac{R^{[4]}_{A_1A_2}(\tau_0,0,\tau_o)}{R^2_{Ac}(\tau_o)} \frac{1 + 2[R_{\tau a}(0) - R_{\tau c}(\tau_o)]\,\Omega^2_{fx}}{\sqrt{1 + 4[R_{\tau a}(0) - R_{\tau c}(\tau_o)]\,\Omega^2_{fx}}} - 1 \quad (4.4\text{-}5)$$

As in the case of the multipath processor the level of uncertainty decreases as the bandwidth of the processed signal decreases. Whereas in the case of the multipath processor this improvement ceased when the working bandwidth decreased below a critical frequency of $1/\sigma_\tau$ the critical frequency for (4.4-5) is roughly $1/[2\{\sigma^2_{\tau a} - R_{\tau c}(\tau_o)\}]^{1/2}$. For incoherent channels ($R_{\tau c}(\tau_o) = 0$) the critical frequency is about the same in both cases. On the other hand, for perfectly correlated channels ($R_{\tau c}(\tau_o) = \sigma^2_{\tau a}$) the critical frequency becomes infinite. For the model considered this behavior might therefore depend on the steering delay, too. Only amplitude fluctuations contribute to the level of the plateau of uncertainty when operating below the critical frequency. However, the other terms of (4.4-4) generally increase with decreasing frequency so that the higher critical frequency available for partially correlated

channels makes array processing attractive.

Similar statements apply to the covariance for the tracking error. In order to shorten the expressions that arise we write

$$\frac{1}{\Omega_a^2(\nu)} = \frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau a}^2 - R_{\tau a}(\nu)]; \quad \frac{1}{\Omega_c^2(\nu)} = \frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau a}^2 - R_{\tau c}(\nu)] \tag{4.4-6}$$

and we contract the notation $\alpha_{jk}(\nu,\tau_o,\tau_o)$ to α_{jk}. Ignoring the derivatives of very slowly varying terms we obtain

$$\begin{aligned} R_{\tau o}(\mu) = {} & \frac{[\Omega x/\Omega_c^3(\tau o)]^2}{T^2 R_{Ac}(\tau o)^2} \left[\int_{\mu+T}^{\mu} [T+\nu-\mu] + \int_{\mu}^{\mu+T} [T-\nu+\mu] \right] \Bigg\{ \\ & \left(\frac{P_{na}^2}{P_x^2}\right)\left(\frac{\Omega^4_{fna}}{\Omega_f^2}\right) [\Omega_{fna}^2 \nu^2 - 1] \exp(-\nu^2\Omega_{fna}^2) \\ & + R_{Aa}(\nu) \left(\frac{P_{na}}{P_x}\right)\left(\frac{\Omega_{fna}}{\Omega_{na}}\right)\left(\frac{\Omega_a(\nu)}{\Omega_x}\right) \exp\{-\tfrac{1}{2}\nu^2 [\Omega_{fna}^2+\Omega_a^2(\nu)]\} \\ \times \quad & \{\Omega_a^2(\nu)[\Omega_a^2(\nu)\nu^2 - 1] + \Omega_{fna}^2[\Omega_{fna}^2\nu^2 - 1]\} \\ & + \left(\frac{P_{ca}^2}{P_x^2}\right)\left(\frac{\Omega_{fnc}^2}{\Omega_f^2}\right) (\tau_o^2 - \nu^2) \exp\{-\tfrac{1}{2} [\nu^2 + \tau_o^2]\} \\ & + \left(\frac{P_{nc}}{P_x}\right)\left(\frac{\Omega_{fnc}}{\Omega_{na}}\right)\left(\frac{\Omega_c(\nu+\tau_o')}{\Omega_x}\right) \times \Omega_{fnc}^2\nu \end{aligned} \tag{4.4-7}$$

$$[R_{Ac}(\nu+\tau_o)\ \Omega_c^2(\nu+\tau_o)\ (\nu+\tau_o) \qquad \exp[-\tfrac{1}{2}\nu^2\Omega_c^2(\nu+\tau_o)\ -\tfrac{1}{2}(\tau_o-\nu)^2\Omega_{fnc}^2\ ]$$

$$+ R_{Ac}(\tau_o-\nu)\,\Omega_c^2(\nu-\tau)(\nu-\tau_o) \quad \exp[-\tfrac{1}{2}\nu^2\Omega_c^2(\tau_o-\nu) - \tfrac{1}{2}(\nu+\tau_o)^2\Omega_{fnc}^2] \quad (-1)$$

$$+ \frac{1}{\Omega_x^2} R^{[4]}_{A_1A_2}(\tau_o,\nu,\nu+\tau_o) \left\{\left[\alpha'_{11} - \frac{\nu^2[\alpha'_{11}+\alpha_{12}]}{[\alpha'_{11}-\alpha_{12}]}\right] \frac{\exp\left\{\frac{-\nu^2}{\alpha'_{11}+\alpha_{12}}\right\}}{[\alpha'^2_{11}-\alpha^2_{12}]}\right. 3/2$$

$$+ \frac{\alpha'_{11}}{[\alpha'^2_{21}-\alpha^2_{22}]} 3/2 +$$

$$\left.\left\{\alpha_{32} + \frac{[\alpha'_{31}-\alpha_{32}][\alpha_{32}-\alpha'_{33}]\nu^2}{[\alpha'_{31}\alpha'_{33}-\alpha^2_{32}]}\right\} \exp\left\{\frac{\frac{1}{2}\nu^2\,[\alpha'_{31}-2\alpha_{32}-\alpha'_{33}]}{[\alpha'_{31}\,\alpha'_{33}-\alpha^2_{32}]}\right\}\right\} d\nu$$

$$[\alpha'_{31}\alpha'_{33}-\alpha^2_{32}]^{3/2}$$

The uncertainty plateau in this case becomes

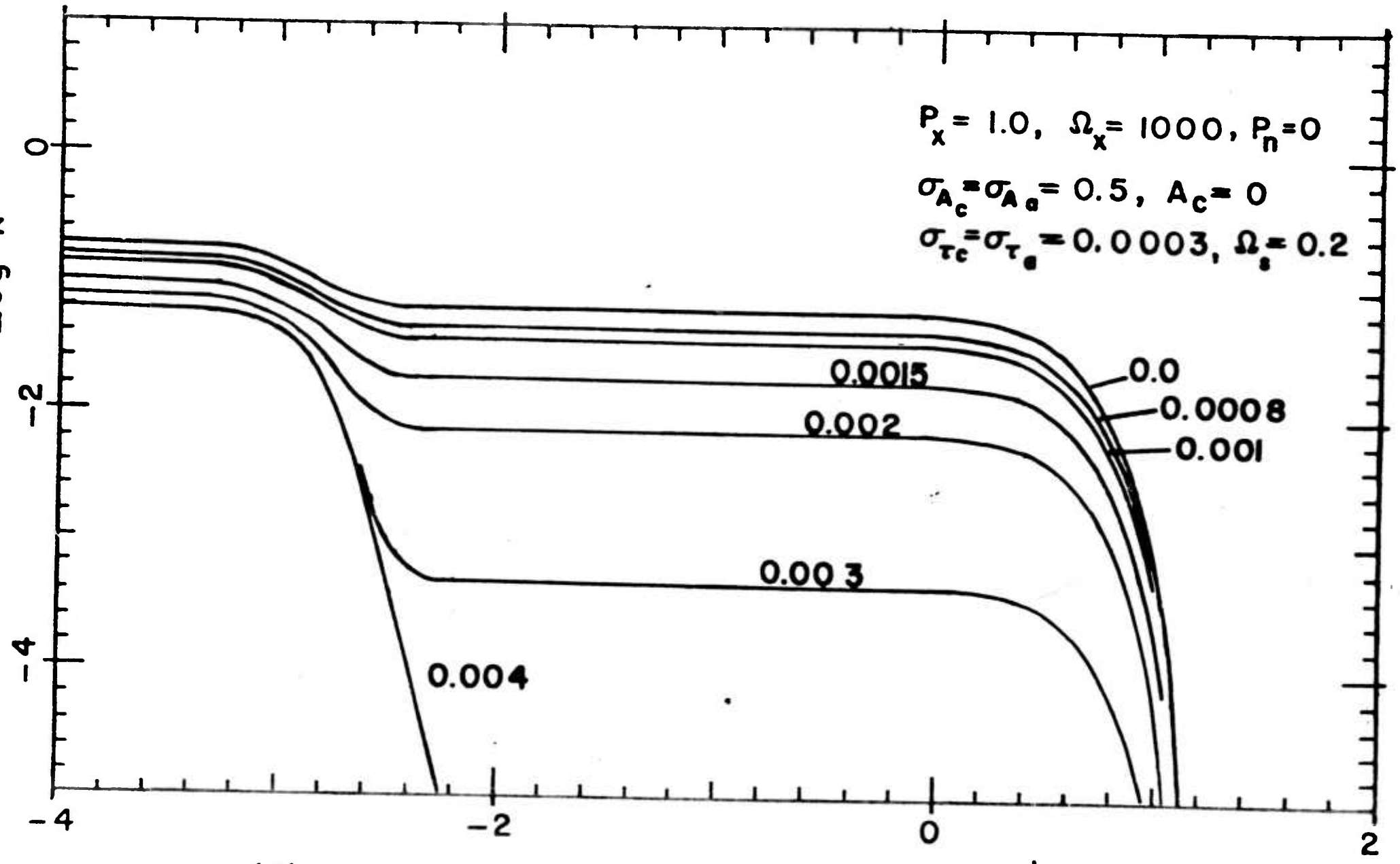
$$\frac{R^{[4]}_{A_1}(\tau_o,0,\tau_o)}{R^2_{Ac}(\tau o)} \quad X \quad \left\{ \frac{[1+2\{\sigma^2_{\tau a}-R_{\tau c}(\tau_0)\}\,\Omega^2_{fx}]^3}{[1+4\{\sigma^2_{\tau a}-R_{\tau c}(\tau_o)\}\,\Omega^2_{fx}]^{3/2}} \right\}$$

$$X \qquad 2[\sigma^2_{\tau a}-R_{\tau c}(\tau_o)] \qquad (4.4\text{-}8)$$

In this case, however, not only does the critical frequency rise with increasing coherence between the two channels, but the magnitude of the plateau below the critical frequency decreases.

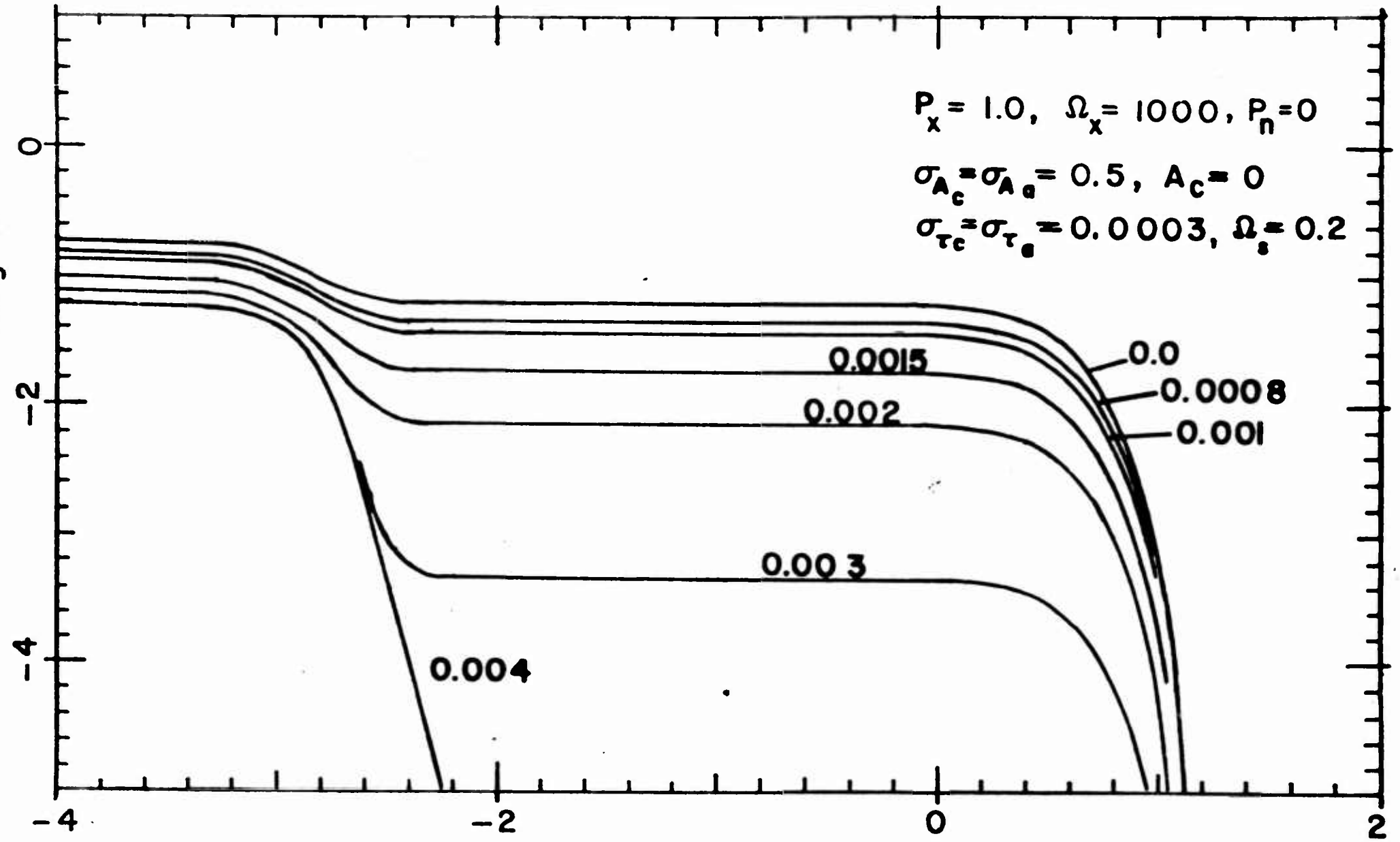
In the limit that $R_{\tau c}(\tau_o) = \sigma^2_{\tau c}$ it is clear that the amplitude fluctuations have no effect on the plateau which, in fact, disappears. The other terms in (4.4-7) do still contribute uncertainty for small T and are affected by amplitude fluctuations unless the background noise terms dominate.

Some of the effects discussed in this chapter are peculiar to the model of frequency insensitive amplitude and delay fluctuations. In particular, the bahavior of the various plateaus of uncertainty as a function of frequency is very simple. However, the root cause for the existence of these persistent levels of uncertainty is the non-gaussian nature of the scattered signals, and this is not a unique property for this simple example. While it is generally necessary to compute all of the cumulants for a non-gaussian density in order to specify it, some insight is gained by examining only the fourth order cumulant $\mu^{[4]}_{1111}(\nu,\tau,\tau')$ which can easily be computed from (3.7-8) and the moments computed in Appendix D. The result for some typical values of the relevant parameters is shown in Figure (4.4-1). The cumulant exhibits a plateau of dependence over a span of time identical to the settling time of the various correlators.



Log $\kappa^{(4)}(\tau, \tau', \nu)$ vs. Log ν FOR VARIOUS $\tau = \tau'$

Fig. 4.4-1



Log $\kappa^{(4)}(\tau, \tau', \nu)$ vs. Log ν FOR VARIOUS $\tau = \tau'$

Fig. 4.4-1

CHAPTER 5

THE RANDOMIZED SINUSOIDAL SURFACE MODEL

5.0 Introduction

Having obtained variance expressions for evaluating the performance of various detectors and trackers using the random amplitude and delay model we now consider the problem of scattering from a sinusoidal boundary. In this case a receiver is excited by reradiation from an illuminated area on a surface of the form

$$\zeta(x,y,t) = h(t)\ \mathrm{Sin}[q_s x \cos\alpha_s + q_s y\ \mathrm{Sin}\ \alpha_s - \Omega_s t - \chi(t)] \quad (5.0.1)$$

where $h(t)$ and $\chi(t)$ are random waveheight and positional phase parameters and are considered to be very slowly varying. The parameters q_s, α_s, and Ω_s are the magnitude and orientation of the propagation vector and temporal frequency for the surface respectively. This model was investigated by Gulin [74] for fixed h and χ.

The primary advantage of this model is that with certain simplifying assumptions the space integrals of equation (1.4-3) can be performed yielding a reliable expression for $H(\omega,t)$ for low to moderate values of the Rayleigh parameter (1.0-2). This means that the results are most relevant for the frequency range of passive detection. The various integrals over frequency required for the computation of various moments of the received signals even with this simplification are still very difficult. Results are presented here in the form of series which converge with reasonable speed.

5.1 Gulin's Solution for H(ω,t) and the Associated Impulse Response

By substituting the equation for the boundary $\zeta(x,y,t)$ into (1.4-13) Gulin obtained the following integral for H(ω,t):

$$H(\omega,t) = \frac{j\omega \sin\psi}{2\,c\,r_o\,r_o'} e^{-j\omega(r_o+r_o')} \sum_{n=-\infty}^{\infty} J_n\left(\frac{2\,h(t-ro/c)\,\mathrm{Sin}\,\psi}{c}\right)$$

$$x\ \exp[\dagger jn\Omega s(t-r_o/c)+jn\chi(t-r_o/c)] \qquad (5.1\text{-}1)$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} B_e(\theta,\phi)\ \exp\{j\frac{\omega}{c}\left(\frac{x^2\sin^2\psi}{R_e}\right)+\frac{y^2}{R_e}\ -jnq(x\cos\alpha + y\sin\alpha)\}\,dx\,dy$$

The space integral is generally not executable unless simplifying assumptions are made about the beam pattern. Consider the case of an omnidirectional receiver for which $B_e(\theta,\phi) \to 1$. The conditions necessary for the expansions of Appendix C in this limit are only marginally satisfied and succeed primarily because of the localized nature of the active scattering region as stated in section 1.4. However, this approximation makes it possible to perform the integral in closed form so that we have

$$H(\omega,t) = \frac{-e^{-j\tau_s\omega}}{r_o + r_o'} \sum_{n=-\infty}^{\infty} J_n[A(t)\omega]\ e^{+j\frac{Bn^2}{\omega}} \qquad (5.1\text{-}2)$$

$$x\ \exp[+jn\Omega_s(t - r_o/c)+jn\chi(t - r_o/c)]$$

where

$$A(t) = \left.\frac{2\,h(t')\,\sin\psi}{c}\right|_{t' = t-r_o/c} \qquad (5.1\text{-}3)$$

$$B = \frac{q_s^2 cR_e}{4} \left[\frac{\cos^2 \alpha}{\sin^2 \psi} + \sin^2 \alpha \right] \qquad (5.1\text{-}4)$$

and

$$\tau_s = \frac{r_o + r'_o}{c} \qquad (5.1\text{-}5)$$

If the waveheight h(t) and phase $\chi(t)$ are slowly varying then it can be seen from (5.1-2) that the time variation of the frequency reponse contains major sidebands at frequencies $n\Omega_s$. Cassedy [59] recognized that each sideband corresponds to a single Bragg order of scattering. In practice, at any given frequency the beamwidth limits the number of excited orders that are received. Furthermore, higher orders are excited at higher frequencies so that for any fixed bandwidth only the first few orders need be retained in (5.1-6). For the applications of this chapter we assume that the model will be bandwidth rather than beamwidth limited.

Additional insight is gained concerning this model if we examine the impulse response corresponding to (5.1-6). Unfortunately, the first power dependence on frequency in the exponent again prevents us from obtaining simple closed form expressions. However, $h(\tau,t)$ can usefully be regarded as a sum of convolutions as follows:

$$h(\tau,t) = \frac{-1}{(r_o + r'_o)} \sum_{n=-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f_1^n(\tau-\rho) f_2^n(\rho) d\rho \right. \qquad (5.1\text{-}6)$$

$$\left. \times \exp\{+jn\Omega_s(t - r_o/c) + jn\chi(t - r_o/c)\} \right] \nu_n$$

where ν_n is 1 for even and j for n odd and where from Watson [78] (p.405,sec.13,42,#4):

$$f_1^n(t) = \int_{-\infty}^{\infty} e^{j\omega(t-\tau_s)} J_n[A(t)\omega] \frac{d\omega}{2\pi}$$

$$= \begin{cases} \dfrac{\mathrm{Cos}\{n \sin^{-1}[(t-\tau_s)/A(t)]\}}{\sqrt{A^2(t) - t^2}} & |t-\tau_s| < A(t) \quad \underline{n \text{ even}} \\ 0 & |t-\tau_s| > A(t) \\ \dfrac{\mathrm{Sin}\{n \sin^{-1}[(t-\tau_s)/A(t)\}}{\sqrt{A^2(t) - t^2}} & |t-\tau_s| < A(t) \quad \underline{n \text{ odd}} \\ 0 & |t-\tau\ | > A(t) \end{cases} \qquad (5.1\text{-}9)$$

and from Erdeli [79] (Vol.1, p.244,#31)

$$f_2^n(t) = \int_{-\infty}^{\infty} e^{j\omega t + j\frac{Bn^2}{\omega}} \frac{d\omega}{2\pi}$$

$$= \begin{cases} \delta(t) + \sqrt{\dfrac{B}{t}}\, J_1(2n\sqrt{Bt}\,) & t = 0 \\ 0 & t < 0 \end{cases} \qquad (5.1\text{-}8)$$

These two functions are sketched in figures (5.1-1a,b). The width of $f_1^n(t)$ is 2A(t) or 4 h(t') Sin ψ /c. This is the difference in travel time between rays reflected from imaginary planes tangent to the upper and lower peaks of the sinusoidal boundary. The singularities at the extremities of this interval suggest the importance of the points of inflection of the surface. Due t he δ - function in (5.1-8) these singularities appear as the result of the convolutions in (5.1-6). The Bessel function in (5.1-8) becomes more oscillatory with increasing n and R_e. In this limit the curvature deviation of the incoming wavefront from planar is negligible over the active scattering region. For large B/A its effect is to introduce very high frequency oscillations which may actually lie outside the processing bandwidth in which case, they may be ignored. Figure (5.1-2) illustrates this effect.

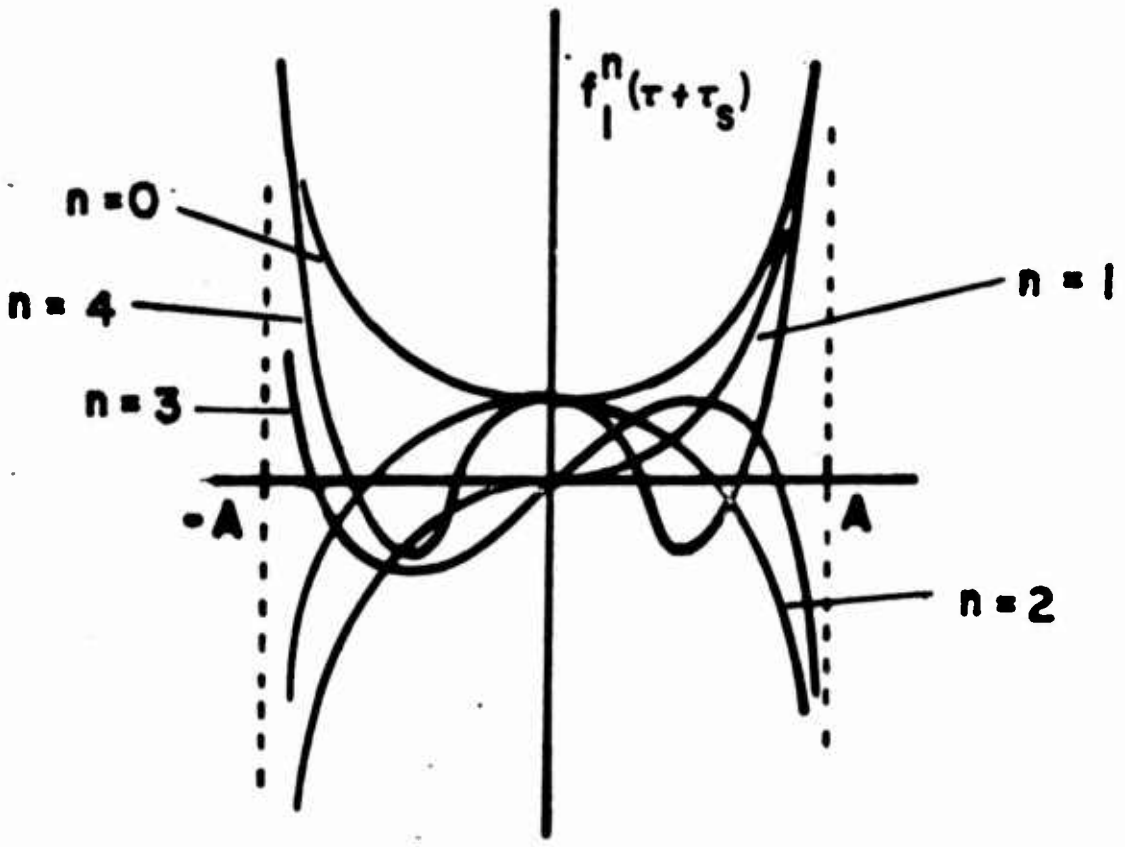


Fig. 5.1-1a

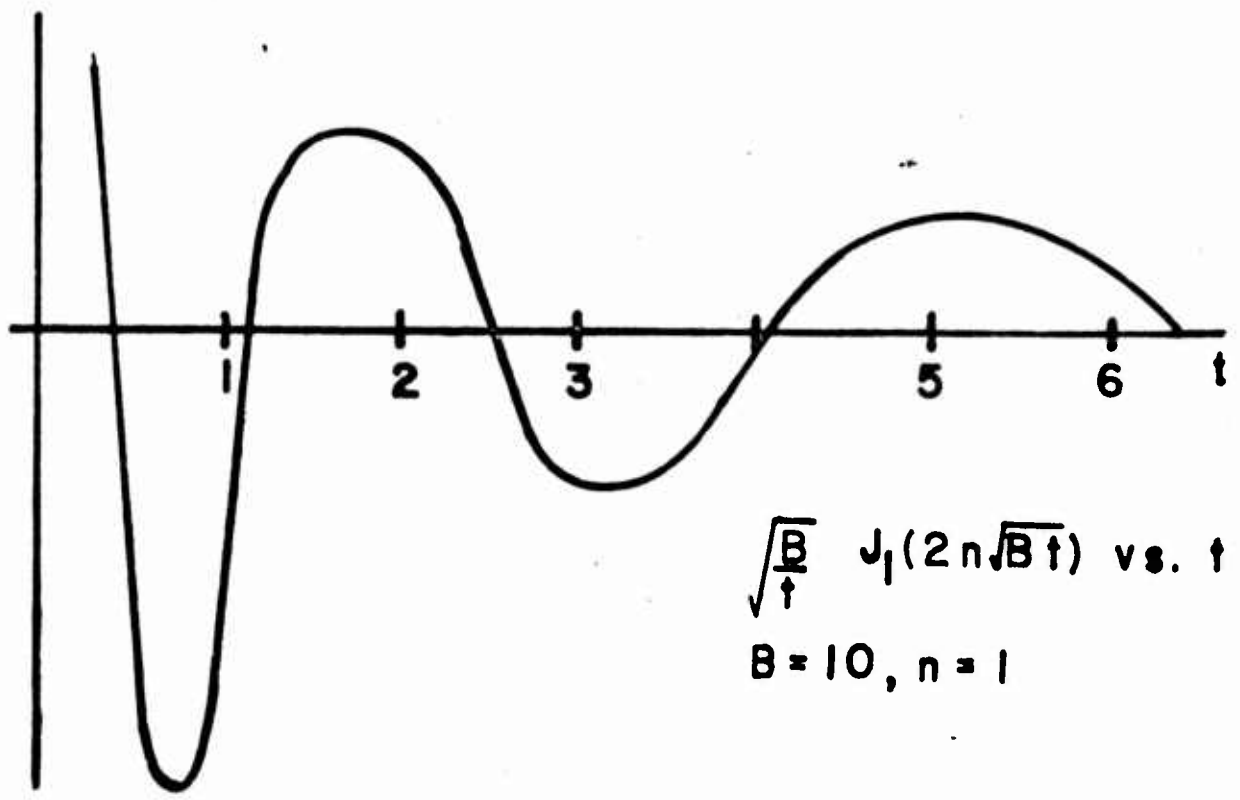


Fig. 5.1-1b

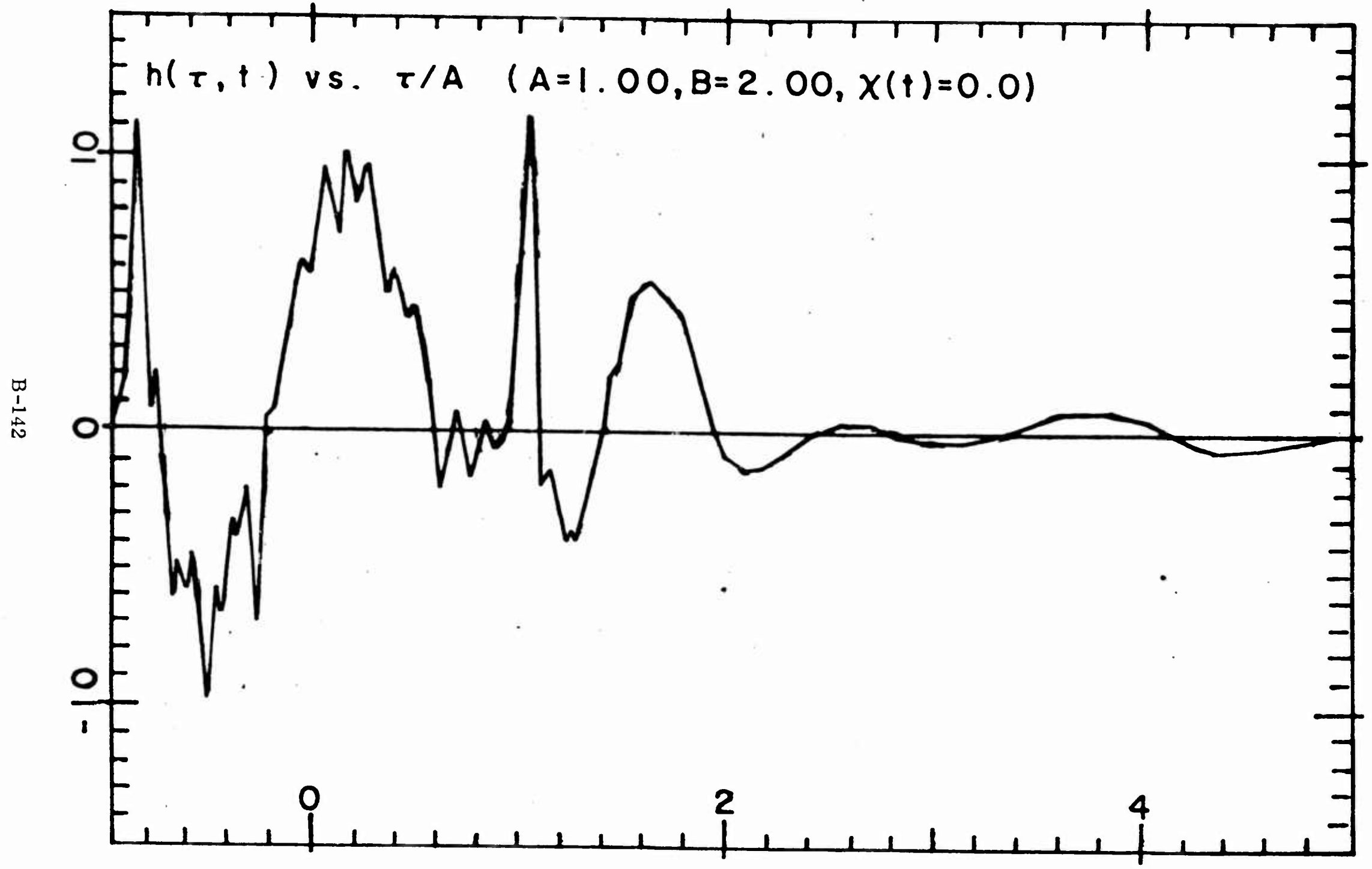


Fig. 5.1-2

5.2 First and Second Order Statistics For the Sinusoidal Boundary Model

We assume that the phase $\chi(t)$ is a sum of two independent random variables:

$$\chi(t) = \chi_o + \chi_1(t) \tag{5.2-1}$$

where χ_o is a random initial phase which uniformly distributed between $[-\pi,\pi]$ and $\chi_1(t)$ is a stationary, zero mean Gaussian process. The waveheight h(t) is also taken to be a stationary, zero mean Gaussian process so that the coherent frequency response becomes

$$H(\omega,t) = H_c(\omega) = \tag{5.2-2}$$

$$\frac{-e^{-j\tau_s\omega}}{(r_o + r'_o)} \int_{-\infty}^{\infty} J_o\left(\frac{2\omega h \sin\psi}{c}\right) \exp\left(\frac{-h^2}{2\sigma_h^2}\right) \frac{dh}{\sqrt{2\pi}\,\sigma_h}$$

$$\frac{-e^{-j\tau_s\omega}}{(r_o + r'_o)} \exp\left(\frac{-\omega^2\sigma_h^2 \sin^2\psi}{c^2}\right) I_o\left(\frac{\omega^2\sigma_h^2 \sin^2\psi}{c^2}\right)$$

The phase $\chi(t)$ and waveheight h(t) are assumed independent. Since computationally this result implicitly contains series representations for the exponential and Bessel functions it is worthwhile to consider an alternative scheme for performing the integral in (5.2-2). By formally writing down the series for an exponential and Bessel function we have for $\frac{1}{2}p - 1 \geq 0$

$$J_{\frac{1}{2}p-1}(z) = (\tfrac{1}{2}z)^{\frac{1}{2}p-1} \exp(-z^2/2p)\, x \tag{5.2-3}$$

$$\left[\sum_{n=o}^{\infty} \sum_{k=o}^{\infty} \frac{1}{n!}\left(\frac{z^2}{2p}\right)^n \frac{1}{k!} \frac{1}{(\frac{1}{2}p)_k}\left(\frac{-1}{4} z^2\right)^k\right] \frac{1}{\Gamma(\frac{1}{2}p)}$$

This series converges very rapidly for values of z up to about 2 or 3 which is sufficient for applications involving weak or moderate scattering. In fact, for some applications one may obtain qualitatively accurate results for small z by stopping with the first term. On rearranging the series we obtain

$$J_{\frac{1}{2}p-1}(z) = \frac{(\frac{1}{2}z)^{\frac{1}{2}p-1}}{\Gamma(\frac{1}{2}p)} \exp\left(\frac{-z^2}{2p}\right) \left[\sum_{m=o}^{\infty} \frac{A(m,p)}{m!} \left(\frac{z}{2p}\right)^m\right] \quad (5.2\text{-}4)$$

where

$$A(m,p) = \sum_{k=o}^{m} \frac{(m)_k}{(\frac{1}{2}p)_k} \frac{(-\frac{1}{2}p)^k}{k!} \quad (5.2\text{-}5)$$

and where

$$(a)_b = \Gamma(a+b)/\Gamma(a) \quad (5.2\text{-}6)$$

In particular, for $J_o(z)$ we have $p = 2$. Hence, (5.2-2) becomes

$$H_c(\omega) = \frac{-e^{-j\omega\tau_s}}{(r_o + r_o')} \frac{1}{\sqrt{1 + \frac{2\omega^2\sigma_h^2 \sin^2\psi}{c^2}}} \quad (5.2\text{-}7)$$

$$\times \sum_{m=o}^{\infty} \frac{1}{2m+1} \frac{(2m)!}{(m!)^2} A(m,2) \frac{\left[\frac{2\omega^2\sigma_h^2 \sin^2\psi}{c^2}\right]^m}{\left[1 + \frac{2\omega^2\sigma_h^2 \sin^2\psi}{c^2}\right]^m}$$

This result converges very rapidly for small values of the Rayleigh parameter $k\sigma_h \sin\psi$, while for large values the frequency behavior is still primarily determined by the first term. Furthermore, when written in this form the result is easier to integrate in various applications. For example, the coherent impulse response corresponding to (5.2-7) is given by

$$h_c(\tau) = \frac{-1}{(r_o+r_o')} \frac{c}{\sqrt{2\pi}\,\sigma_h \operatorname{Sin}\psi} \sum_{m=o}^{\infty} \frac{(-1)^m}{2m+1} \frac{\Gamma(\tfrac{1}{2}-m)(2m)!}{(m!)^2} A(m,2)$$

$$\times \frac{1}{2m} \left[\frac{d^{2m}}{d\xi^{2m}} \xi^m k_m(\xi) \right]_{\xi=(\tau-\tau_s)\frac{c}{\sqrt{2}\,\sigma_h \operatorname{Sin}\psi}} \qquad (\tau \geq \tau_s) \tag{5.2-8}$$

where again the first term is the most important one for low frequency applications.

The same philosophy must be applied in obtaining the second order statistics since without this technique it is not possible to perform the appropriate averages. Hence we have

$$\Phi(\omega,\omega',\tau) = \frac{e^{j(\omega-\omega')\tau_s}}{(r_o+r_o')^2} \sum_{n=-\infty}^{\infty} \frac{}{J_n[A(t)\omega]\, J_n[A(t-\tau)\omega']}$$

$$\times \exp\{jBn^2\left(\frac{1}{\omega} - \frac{1}{\omega'}\right)\}\; Q_\chi(-n,n,\tau)\, e^{jn\Omega_s\tau} \tag{5.1-9}$$

$$= \frac{e^{j(\omega-\omega')\tau_s}}{(r_o+r_o')^2} \sum_{n=o}^{\infty} \frac{1}{(n!)^2} \exp\{jBn^2\left(\frac{1}{\omega} - \frac{1}{\omega'}\right)\}\; Q_{\chi_1}(-n,n,\tau)\, \epsilon_n \operatorname{Cos} n\Omega_s\tau$$

$$\sum_{\iota=o}^{\infty} \sum_{m=o}^{\infty} \frac{1}{[4(n+1)]^{\iota+m}} \frac{A(\iota,2n+1)}{\iota!} \frac{A(m,2n+1)}{m!} \left(\frac{\omega' \operatorname{Sin}\psi\, h(t-\tau)}{c}\right)^{2m+n}$$

$$\left(\frac{\omega \operatorname{Sin}\psi\, h(t)}{c}\right)^{2\iota+n} \exp\{[\omega^2 h^2(t) + \omega'^2 h^2(t-\tau)] \frac{\operatorname{Sin}^2\psi}{(n+1)c2}\}$$

where the stationarity of h(t) and $\chi_1(t)$ has been employed and where $Q_{\chi_1}(n,n',\tau)$ is the characteristic function for the random variables $\chi_1(t)$ and $\chi_1(t-\tau)$. It should be emphasized again that only the terms for low values of ι and m need be retained for low frequency applications. The ϵ_n are 2 for $n \neq 0$ and 1 for n=0.

The remaining average over h(t) and h(t-τ) becomes

$$\left(\frac{\omega \sin\psi}{c}\right)^{2l+n}\left(\frac{\omega' \sin\psi}{c}\right)^{2m+n}\int_{-\infty}^{\infty}\int h_t^{2l+n}\, h_{t-\tau}^{2m+n}\,\frac{dh_t\, dh_{t-\tau}}{2\pi\sqrt{\sigma_h^4-r_h^2(\tau)}}$$

$$e^{-\left\{\left(\frac{\omega^2 \sin^2\psi}{(n+1)c^2}+\frac{\frac{1}{2}\sigma_h^2}{(\sigma_h^4-r_h^2(\tau))}\right)h_t^2-\frac{r_h(\tau)}{(\sigma_h^4-r_h^2(\tau))}\,h_t\,h_{t-\tau}\right.} \qquad (5.2\text{-}10)$$

$$\left.\left(\frac{\omega'^2 \sin^2\psi}{(n+1)\,c^2}+\frac{\frac{1}{2}\sigma_h^2}{(\sigma_h^2-r_h^2(\tau))}\right)h_{t-\tau}^2\right\}$$

where σ_h^2 and $r_h(\tau)$ are the variance and correlation function respectively for h(t). This average can be transformed as follows:

$$D'_{n,l,m}\int_{-\infty}^{\infty}\int h_t^{2l+n}\,h_{t-\tau}^{2m+n}\,\frac{e^{-\frac{1}{2}\left\{\frac{A'_n h_t^2-2B'_n h_t h_{t-\tau}+C'_n h_{t-\tau}^2}{(A'_n C'_n-B'^2_n)}\right\}}}{2\pi\sqrt{A'_n C'_n-B'^2_n}}\,dh_t\,dh_{t-\tau} \qquad (5.2\text{-}11)$$

$$= D'_{n,l,m}\mu'_{2l+n,2m+n}(n)$$

where $\mu'_{i,k}$ is the joint moment of the h(t) and h(t-τ) as computed from the modified Gaussian distribution shown and where

$$D'_{n,l,m} = \left(\frac{\omega \sin\psi}{c}\right)^{2l+n}\left(\frac{\omega' \sin\psi}{c}\right)^{2m+n}\frac{[A'_n C'_n - B'^2_n]^{\frac{1}{2}}}{[\sigma_h^4 - r_h^2(\tau)]^{\frac{1}{2}}} \qquad (5.2\text{-}12)$$

$$A'_n = 2\,\frac{\omega^2 \sin^2\psi}{(n+1)\,c^2} + \frac{\sigma_h^2}{(\sigma_h^4 - r_h^2(\tau))} \qquad (5.2\text{-}13)$$

$$B' = \frac{\frac{r_h(\tau)}{(\sigma_h^4 - r_h^2(\tau))}}{(A'_n C'_n - B'^2_n)^{-1}} \qquad (5.2\text{-}14)$$

$$C'_n = \frac{2\,\frac{\omega'^2 \sin^2\psi}{(n+1)\,c^2} + \frac{\sigma_h^2}{(\sigma_h^4 - r_h^2(\tau))}}{(A'_n C'_n - B'^2_n)^{-1}} \qquad (5.2\text{-}15)$$

and

$$(A'_n\,C'_n - B'^2_n)^{-1} = 4\,\frac{\omega^2\omega'^2 \sin^4\psi}{(n+1)^2\,c4} \qquad (5.2\text{-}16)$$

$$+ \frac{1}{(\sigma_h^4 - r_h^2(\tau))} \left\{ \left\{ [\omega^2+\omega'^2] \left[\frac{2\,\sigma_h^2 \sin^2\psi}{(n+1)\,c^2} \right] + 1 \right\} \right.$$

Thus, (5.2-9) becomes

$$\Phi(\omega,\omega',\tau)$$

$$= \frac{e^{-j(\omega-\omega')\tau_s}}{(r_o+r'_o)^2} \sum_{n=o}^{\infty} \frac{1}{(n!)^2} \exp\left\{ jBn^2 \left(\frac{1}{\omega} - \frac{1}{\omega'} \right) \right\} Q_\chi(-n,n,\tau)\varepsilon_n \cos n\Omega_s\tau \qquad (5.2\text{-}17)$$

$$\sum_{\iota=o}^{\infty} \sum_{m=o}^{\infty} \frac{1}{[4(n+1)]^{\iota+m}} \frac{A(\iota,2n+1)}{\iota!} \frac{A(m,2n+1)}{m!} D'_{n,\iota,m}\mu'_{2\iota+n,2m+n}(n)$$

It should be noted that both $D'_{n,\iota,m}$ and $\mu'_{j,k}(n)$ are functions of τ. The most natural scheme for truncating these series is to retain terms only up to a fixed value for $2(n+\iota+m)$ since this sets the order of the approximation in terms of frequency times waveheight.

Writing down all moments up to order 4 we have

$$
\begin{aligned}
&n=o,\ \iota=o,\ m=o && \mu'_{oo}(o) = 1\\
&n=o,\ \iota=1,\ m=o && \mu'_{2o}(o) = A'_o\\
&n=o,\ \iota=o,\ m=1 && \mu'_{02}(0) = C'_0\\
&n=1,\ 1=o,\ m=o && \mu'_{11}(1) = B'_1\\
&n=o,\ \iota=1,\ m=1 && \mu'_{22}(0) = A'_0C'_0 + 2B'^2_0\\
&n=2,\ \iota=o,\ m=o && \mu'_{22}(22) = A'_2C'_2 + 2B'^2_2\\
&n=1,\ \iota=1,\ m=o && \mu'_{3:}(1) = 3A'_1B'_1\\
&n=1,\ \iota=o,\ m=1 && \mu'_{1:}(1) = 3C'_1B'_1\\
&n=o,\ \iota=2,\ m=o && \mu'_{4C}(0) = A'^2_0\\
&n=o,\ \iota=o,\ m=2 && \mu'_{0:}(0) = C'^2_0
\end{aligned}
\qquad (5.2\text{-}18)
$$

It can be seen that all cf the moments $\mu_{n+2\,\iota,n+2m}(n)$ can be written in the following form:

$$
\mu_{n+2\,\iota,n+2m}(n) = \frac{P_{n,\iota,m}(\omega,\omega',\tau)}{[A'_nC'_n - B'^2_n]^{-(n+\iota+m)}} \qquad (5.2\text{-}19)
$$

where the $P_{n,\iota,m}(\omega,\omega',\tau)$ are polynomials in ω and ω' of order $2(n+\iota+m)$ in frequency and containing only even powers of ω and ω'.

We write them in terms of frequency normalized by the Rayleigh parameter:

$$P_{n,\ell,m}(\omega,\omega',\tau) = \sum_{p,q=0}^{n+\ell+m} G_{n,\ell,m}(p,q,\tau) \left| \frac{2\omega^2\sigma_h^2 \sin^2\psi}{(n+1)c^2} \right|^P \left| \frac{2\omega'^2\sigma_h^2 \sin^2\psi}{(n+1)c^2} \right|^q$$

The first several of the coefficients $G_{n,\ell,m}(p,q,\tau)$ have been tabulated in Appendix G. Using (5.2-20), (5.2-16) and (5.2-12) we have

$$D'_{n,\ell,m}\mu'_{n+2\ell,n+2m}(n) =$$

$$\{(n+1)[1-\rho_h^2(\tau)]\}^{\ell+n+m} \sum_{p,q=0}^{n+\ell+m} G_{n,\ell,m}(p,q,\tau)\xi_n^{2(p+\ell)+n}\xi_n^{2(q+m)+n}$$

$$\tfrac{1}{2}\{\xi_n^2\xi_n'^2[1-\rho_h^2(\tau)] + [\xi_n^2+\xi_n'^2] + 1\}^{\ell+n+m+\frac{1}{2}} \qquad (5.2\text{-}21)$$

where we define the normalized auto-covariance function for h(t):

$$\rho_h(\tau) = \frac{r_h(\tau)}{\sigma^2} \qquad (5.2\text{-}22)$$

and

$$\xi_n = \sqrt{\frac{2}{n+1}}\,\frac{\sigma_h \sin\psi}{c}\,\omega \quad ; \quad \xi'_n = \sqrt{\frac{2}{n+1}}\,\frac{\sigma_h \sin\psi}{c}\,\omega' \qquad (5.2\text{-}23)$$

are the normalized frequencies. As a practical matter, the fourth order dependence on frequency for the term $\xi^2\xi'^2[1-\rho_h^2(\tau)]$ creates certain difficulties when performing various integrals over frequency.

In order to circumvent these difficulties, we rewrite the denominator of (5.2-21) as follows:

$$\frac{1}{\{\xi_n^2\xi_n'^2[1-\rho_h^2(\tau)] + [\xi_n^2+\xi_n'^2] + 1\}^{n+\iota+m+\frac{1}{2}}}$$

$$= \frac{1}{\{ -\xi_n^2\xi_n'^2\rho_h^2(\tau) + (\xi_n^2+1)(\xi_n'^2+1)\}^{n+\iota+m+\frac{1}{2}}} \qquad (5.2\text{-}24)$$

$$= \frac{1}{[(\xi_n^2+1)(\xi_n'^2+1)]^{n+\iota+m+\frac{1}{2}}} \left[1 - \frac{\rho_h^2(\tau)\ \xi_n^2\xi_n'^2}{(\xi_n^2+1)(\xi_n'^2+1)}\right]^{n+\iota+m+\frac{1}{2}}$$

Next we note that

$$0 \leq \frac{\rho_h^2(\tau)\ \xi_n^2\xi_n'^2}{(\xi_n^2+1)(\xi_n'^2+1)} < 1 \qquad (5.2\text{-}25)$$

and for the large τ or for small ξ and ξ' the lower boundary is approached. Using the result that [80]

$$(1-z)^{-a} = {}_1F_0(a;;z) \qquad (5.2\text{-}26)$$

we have that (5.2-25) becomes

$$\frac{1}{\{\xi_n^2\xi_n'^2[1-\rho_h^2(\tau)] + [\xi_n^2+\xi_n'^2] + 1\}^{n+\iota+m+\frac{1}{2}}}$$

$$= \sum_{r=0}^{\infty} \frac{\Gamma(\iota+m+n+r+\frac{1}{2})}{\Gamma(\ +n+m+\frac{1}{2})} \frac{1}{r!} \frac{[\rho_h^2(\tau)\ \xi_n^2\xi_n'^2]^r}{[\xi_n^2+1)(\xi_n'^2+1)]^{n+\iota+m+r+\frac{1}{2}}} \qquad (5.2\text{-}27)$$

Of all the expansions made thus far in the analysis this last one converges least rapidly. However, convergence is worst for the case of $\tau = 0$, and then only for $\xi\xi'$ greater than 1. In this case alternative

procedures can be used to perform the various integrals that arise. For example, by writing

$$\frac{1}{\{\xi_n^2\xi_n'^2[1-\rho_h^2(\tau)] + \xi_n^2 + \xi_n'^2 + 1\}^{n+\iota+m+\frac{1}{2}}}$$

$$= \frac{1}{\{1+\xi_n^2+\xi_n'^2\}^{\iota+m+n+\frac{1}{2}}} \frac{1}{\left\{\frac{1+\{\xi_n^2\xi_n'^2[1-\rho_h^2(\tau)]\}}{(1+\xi_n^2+\xi_n'^2)}\right\}^{\iota+m+n+\frac{1}{2}}} \qquad (5.2\text{-}28)$$

$$= \sum_{r=0}^{\infty} \frac{\Gamma(\iota+m+n+r+\frac{1}{2})}{\Gamma(\iota+m+n+\frac{1}{2})} \frac{1}{r!} \frac{\{-\xi_n^2\xi_n'^2[1-\rho_h^2(\tau)]\}^r}{(1+\xi_n^2+\xi_n'^2)^{\iota+m+n+r+\frac{1}{2}}}$$

The convergence of this series is more rapid for τ near 0 and $\xi\xi'$ greater than unity. Unfortunately, this representation fails to converge at all for frequencies too high for the condition

$$\frac{\xi_n^2\,\xi_n'^2\,[1-\rho_h^2(\tau)]}{(1+\xi_n^2+\xi_n'^2)} < 1 \qquad (5.2\text{-}29)$$

to be satisfied. Therefore, in order to avoid questions of convergence for integrals of these functions over frequency we will use (5.2-27) throughout the rest of the analysis. Computationally (5.2-28) still offers advantages for certain applications. However, an incidental advantage of (5.2-27) is that it is a sum of terms which are factored in such a way that multiple integrals over ξ and ξ' can be performed with greater ease.

It should be noted that the density (5.2-11) is singular for

$\tau = o$ as are the coefficients $G_{n,\iota,m}(p,q,\tau)$, but the product $[1-\rho_h^2(\tau)]^{n+\iota+m} G_{n,\iota,m}(p,q,\tau)$ which arises in (5.2-21) remains finite. Substituting (5.2-27) and (5.2-21) into (5.2-17) we obtain

$$\Phi(\omega,\omega';\tau) \qquad (5.2\text{-}29)$$

$$= \frac{e^{-j(\omega-\omega')\tau_s}}{(r_o+r'_o)^2} \sum_{n=o}^{\infty} \frac{1}{(n!)^2} \exp\left\{jBn^2\left|\frac{1}{\omega} - \frac{1}{\omega'}\right|\right\} Q_{\chi_1}(-n,n,\tau)\varepsilon_n \mathrm{Cos}\, n\Omega_s\tau$$

$$\sum_{\iota=o}^{\infty}\sum_{m=o}^{\infty} \frac{1}{[4(n+1)^{\iota+m}} \frac{A(\iota,2n+1)}{\iota!} \frac{A(m,2n+1)}{m!} \sum_{r=o}^{\infty} \frac{\Gamma(\iota+n+m+r+\frac{1}{2})}{\Gamma(\iota+n+m+\frac{1}{2})} \frac{1}{r!} \sum_{p,q=o}^{n+\iota+m}$$

$$\{(n+1)[1-\rho_h^2(\tau)]\}^{\iota+n+m} G_{n,\iota,m}(p,q,\tau)[\rho_h^2(\tau)]^r \xi_n^{2(r+p+\iota)+n}\xi_n^{2(r+q+m)+n}$$

$$\overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}$$

$$[(\xi_n^2+1)(\xi_n'^2+1)]^{n+\iota+m+r+\frac{1}{2}}$$

for the interfrequency system correlation function, and for comparison with (4.1-14) we have the tri-frequency spectral density:

$$\frac{dA(\omega,\omega',\omega'')}{d\omega''} = \qquad (5.2\text{-}30)$$

$$\frac{e^{-j(\omega-\omega')\tau_s}}{(r_o+r'_o)^2} \sum_{n=o}^{\infty} \frac{1}{(n!)^2} \exp\left\{jBn^2\left|\frac{1}{\omega} - \frac{1}{\omega'}\right|\right\} \varepsilon_n \sum_{\iota=o}^{\infty}\sum_{m=o}^{\infty} \frac{A(\iota,2n+1)A(m,2n+1)}{\iota!\quad m!}$$

$$\sum_{r=o}^{\infty} \frac{\Gamma(\iota+n+m+r+\frac{1}{2})}{\Gamma(\iota+n+m+\frac{1}{2})} \frac{1}{r!} \sum_{p,q=o}^{\infty} S_{n,\iota,m,p,q}(\omega'') \frac{\xi_n^{2(r+p+\iota)+n}\xi_n'^{2(r+q+m)+n}}{[(\xi_n^2+1)(\xi_n'^2+1)]^{n+\iota+m+r+\frac{1}{2}}}$$

where

$$S_{n,\iota,m,p,q,r}(\omega'') = \qquad (5.2\text{-}31)$$

$$\int e^{-j\omega''\tau} Q_{\chi_1}(-n,n,\tau)\{(n+1)[1-\rho_h^2(\tau)]\}^{\iota+n+m}\{\rho_h^2(\tau)\}^r G_{n,\iota,m}(p,q,\tau)\mathrm{Cos}n\Omega_s\tau\, d\tau$$

is the spectrum of the slow variations of the scattering. When the

time variations of the surface positional phase $\chi_1(t)$ and waveheight $h(t)$ are slow compared to the mean translational frequency Ω_s then the presence of the term Cos $n\Omega_s \tau$ in (5.2-32) produces sidebands in $S_{n,\iota,m,p,q,r}(\omega'')$ which is not the case in (4.1-15). Since the ξ_n decrease with increasing n for fixed ω and since $\xi_n \geq 1$, higher order sidebands are excited at higher values of ω.

Similar comments apply to the smearing function

$$\frac{dC(\nu,\gamma)}{d\gamma} = \tag{5.2-32}$$

$$\frac{\sqrt{\pi}}{(r_o+r_o')^2} \sum_{n=o}^{\infty} \frac{\epsilon_n}{(n!)^2} \sum_{\iota=o}^{\infty} \sum_{m=o}^{\infty} \frac{A(\iota,2n+1)A(m,2n+1)}{\iota!\quad m!} \sum_{r=o}^{\infty} \frac{\Gamma(\iota+m+n+r+\frac{1}{2})}{\Gamma(\iota+m+n+r)} \frac{1}{r!} \sum_{p,q=o}^{n+\iota+m}$$

$$S_{n,\iota,m,p,q,r}(\omega'')(\tfrac{1}{2}\nu_n')^{2(n+\iota+m+r)+\frac{1}{2}}$$

$$\times \left(\frac{d}{d\nu'}\right)^{4r+2(n+\iota+m+p+q)} [\nu_n'^{2(n+\iota+m)+\frac{1}{2}} K_{2(n+\iota+m)+\frac{1}{2}}(\nu_n')] \cdot$$

where the scaled delay spread parameter ν_n' is given by

$$\nu_n' = \sqrt{\frac{n+1}{2}}\ \frac{c}{\sigma_h \text{ Sin } \psi}\ \nu$$

That is, the more intricate delay distortions fade more rapidly. Thus, while this model is more complex than the random amplitude and delay model of Chapter 4 because of the detailed nature of $H(\omega,t)$, many of the general properties of the two scattering models are the same.

5.3 Multipath Correlator Fluctuations for the Random Sinusoidal Boundary

The spectra (4.2-2) and filters (4.2-3) were chosen to be Gaussian shaped in order to simplify the various integrals over frequency that arise in the analysis of the random amplitude and delay model. However, the Gaussian shape is not suitable for use with expression (5.2-29) for $\Phi(\omega,\omega',\tau)$ in computing the normalized covariances (4.2-1). This is primarily due to the factor $[(\xi_n^2+1)(\xi_n'^2+1)]^{\frac{1}{2}}$ in the denominator of each of the terms in (5.2-29). The presence of these branch point singularities inevitably leads to Bessel functions or other functions which are poorly tabulated or otherwise unfamilar.

In order to circumvent these difficulties we assume in this case that the spectra are of exponential form:

$$a)\ S_{xx}(\omega) = \frac{\Pi\ P_x}{\Omega_x}\ e^{-|\omega|/\Omega x} \qquad (5.3\text{-}1)$$

$$b)\ S_{n_1n_1}(\omega) = S_{n_2n_2}(\omega) = \frac{\Pi\ P_n}{\Omega_n}\ e^{-|\omega|/\Omega_n}$$

These spectra have roughly the same power in the band from 0 to $\omega = \Omega$ as the spectra in (4.2-2) although they tend to zero less rapidly for large ω. Similarly, we take the filters as

$$H_1(\omega) = H_2(\omega) = e^{-\frac{1}{2}|\omega|/\Omega_f} \qquad (5.3\text{-}2)$$

The direct path transfer function is given by (3.1-1). Using equation (3.4-5) and the expression for the coherent response $H_c(\omega)$.

given in (5.2-2) we have

$$\overline{\Xi(\tau,T,\rho)} = \frac{-1}{2\pi R_d(r_o+r_o')} \int_{-\infty}^{\infty} e^{j\omega(\tau-\tau_o)} \frac{\pi P_x}{\Omega_x} e^{-|\omega|\left(\frac{1}{\Omega_f}+\frac{1}{\Omega_x}\right)}$$

$$\sum_{n=0}^{\infty} \frac{1}{2^{n+1}} \frac{(2n)!}{(n!)^2} A(n,2) \frac{\xi_o^{2n}}{\left[1+\xi_o^2\right]^{n+\frac{1}{2}}} d\omega = \tag{5.3-3}$$

$$\frac{c_o P_x/\Omega_x}{R_d(r_o+r_o')} \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} \frac{(2n)!}{(n!)^2} RE\left\{ \int_0^{\infty} e^{-\xi_o c_o \left[\alpha_{fx}-j(\tau-\tau_o)\right]} \frac{\xi_o^{2n}}{\left[1+\xi_o^2\right]^{n+\frac{1}{2}}} d\xi_o \right\}$$

where the symbol RE stands for the "real part of" the enclosed quantity, and where

$$\alpha_{fx} = \frac{1}{\Omega_{fx}} = \frac{1}{\Omega_f} + \frac{1}{\Omega_x} \tag{5.3-4}$$

$$c_n = \sqrt{\frac{n+1}{2}} \frac{c}{\sigma_h \sin\psi} \tag{5.3-5}$$

$$\tau_o = \tau_s - \tau_d \tag{5.3-6}$$

Now the integral over normalized frequency arises several times in the analysis of this model for the scattering system and so it is worth examining in some detail. Using the contour integral representation given by Slater[82] (p. 25, #1.6.1.6) for the hypergeometric function (5.2-26) we have that for $a > 0$

$$(1+z)^{-a} = \frac{1}{\Gamma(a)} \int_{j\infty}^{j\infty} \Gamma(a+s)\, \Gamma(-s)\, (z)^s \frac{ds}{2\pi j} \tag{5.3-7}$$

where the contour implied must pass from $-j\infty$ to $+j\infty$ between $-a$ and 0. Using this the integral in (5.3-3) becomes

$$\int_0^\infty u^\rho \quad e^{-uz} \quad (1 + u^2)^{-a} \quad du =$$

$$\int_{-j\infty}^{j\infty} \frac{\Gamma(a+s)}{\Gamma(a)} \quad \Gamma(-s) \int_0^\infty e^{-zu} \quad u^{\rho+2s} \quad du \, \frac{ds}{2\pi j} =$$

$$\int_{-j\infty}^{j\infty} \frac{\Gamma(\rho+2s+1)}{z^{\rho+2s+1}} \quad \frac{\Gamma(a+s)}{\Gamma(a)} \quad \Gamma(-s) \quad \frac{ds}{2\pi j} =$$

$$\frac{1}{2\pi\Gamma(a)} \quad \left(\frac{2}{z}\right)^{\rho+1} \int_{-j\infty}^{j\infty} \left(\frac{4}{z^2}\right)^s \quad \Gamma(\tfrac{1}{2}\rho+s+\tfrac{1}{2}) \; \Gamma(\tfrac{1}{2}\rho+s+1) \; \Gamma(a+s) \; \Gamma(-s) \; \frac{ds}{2\pi j}$$

$$= \frac{1}{2\pi \; \Gamma(a)} \quad \left(\frac{2}{z}\right)^{\rho+1} \quad G_{31}^{13}\left(\frac{4}{z^2} \,\middle|\, \begin{matrix} \tfrac{1}{2}\rho+\tfrac{1}{2}, \tfrac{1}{2}\rho+1, a \\ 0 \end{matrix}\right)$$

where the contour is further restricted to pass from $-j\infty$ to $+j\infty$ between 0 and the larger of $-\frac{1}{2}\rho-\frac{1}{2}$ and $-a$. The $G_{pq}^{mn}(x)$ function is called a Meijer function [83] and the integral defining it always converges provided $RE(z)>0$. The contour is normally closed in the left half plane for the purpose of evaluating the integral although an asymptotic representation can be obtained by closing in the right half plane [82]. Since for the applications of interest here the quantity a is either an integer or a half integer, some of the poles of $\Gamma(a+s)$ overlay the poles of $\Gamma(\frac{1}{2}\rho+\frac{1}{2}+s)$ or $\Gamma(\frac{1}{2}\rho+1+s)$ in the integrand of the defining integral thereby making them second order. By a translation and reflection change of variables we can rewrite (5.3-8) in a more compact form:

$$\int_0^\infty u^\rho \; e^{-zu} \; (1+u^2)^{-a} \; du =$$

(5.3-9)

$$\frac{1}{2\pi\,\Gamma(a)} \int_{-j\infty}^{j\infty} \left(\frac{z^2}{4}\right)^s \Gamma(-s)\;\Gamma(\tfrac{1}{2}-s)\;\Gamma(a-\rho-\tfrac{1}{2}-s)\;\Gamma(\rho+\tfrac{1}{2}+s)\;\frac{ds}{2\pi j}$$

$$= \frac{1}{2\pi\,\Gamma(a)}\; G_{13}^{31}\left(\frac{z^2}{4}\middle|_{0,\,\frac{1}{2},\,a-\frac{1}{2}\rho-\frac{1}{2}}^{\frac{1}{2}-\frac{1}{2}\rho}\right)$$

Another obvious property of the Maijer function notation is the simplicity of differentiation:

$$\left(\frac{d}{dz}\right)^m \; G_{13}^{31}\left(\frac{z^2}{4}\middle|_{0,\,\frac{1}{2},\,a-\frac{1}{2}\rho-\frac{1}{2}}^{\frac{1}{2}-\frac{1}{2}\rho}\right) = (-1)^m\, G_{13}^{31}\left(\frac{z^2}{4}\middle|_{0,\,\frac{1}{2},\,a-\frac{1}{2}(\rho+m)-\frac{1}{2}}^{\frac{1}{2}-\frac{1}{2}(\rho+m)}\right) \qquad (5.3\text{-}10)$$

Although the Meijer function serves as a convenient identification of the contour integral, it tells us nothing about how to compute its values for various z. Computation schemes are discussed in Appendix I where series representations are presented.

Nevertheless, using this notation we can express the mean of the output for the multipath correlator using (5.3-3) and (5.3-9):

$$\Xi(\tau,T,\rho) = \frac{-c_o P_x/\Omega_x}{R_d(r_o+r_o')} \sum_{n=0}^{\infty} \frac{1}{2^{n+1}}\,\frac{(2n)!}{(n!)^2}$$

(5.3-11)

$$\mathrm{RE}\left\{\frac{1}{2\pi\,\Gamma(n+\frac{1}{2})}\; G_{13}^{31}\left(\tfrac{1}{4}\left[c_o(\alpha_{fx}-j(\tau-\tau_o)\right]^2\middle|_{0,\;\frac{1}{2},\;0}^{\frac{1}{2}-n}\right)\right\}$$

Roughly speaking, the peak of correlation which occurs at $\tau = \tau_o$ has a width which is on the order of $1/c_o$ for very wide bandwidths

(small α_{fx}). For very small bandwidths the peak width approaches α_{fx} as it should.

The integrals I_k (k=1 to 4) which arise in the expression (3.3-8) for the variance of the output of the correlator have been summarized in Appendix H. Their derivation is straightforward because all multiple integrals that occur can be written as sums of iterated integrals of the form given in (5.3-9). It is unnecessary to write down the full expression for the variance since the presentation of the I_k in Appendix H is already in a format suitable for numerical computation provided a basic program is constructed to implement equation (3.3-8).

Instead, we concentrate on examining the plateau of variance for this model corresponding to (4.2-11) for the random amplitude and delay model. From chapter 4 we know that this plateau is given by

$$\lim_{\upsilon\to 0} \frac{I_4(\tau_o,\tau_o,\upsilon)}{[\Xi(\tau_o,T,\rho)]^2} - 1 = \lim_{\upsilon\to 0}$$

$$\frac{P_x/\Omega_x}{[\Xi(\tau_o,T,\rho)]^2 R_d^2 (r_o+r_o')^2} \sum_{n=0}^{\infty} \int_{-\infty}^{+\infty} \left(\frac{c_n}{n!}\right)^2 Q_{\chi_1}(-n,n,\upsilon)\, \epsilon_n \cos(n\Omega_s\upsilon) \sum_{1=0}^{\infty} \sum_{m=0}^{\infty}$$

$$\frac{1}{[4(n+1)]^{1+m}} \frac{A(1,2n+1)}{1!} \frac{A(m,2n+1)}{m!} \sum_{r=0}^{\infty} \frac{1}{\Gamma(1+m+n+\frac{1}{2})\,\Gamma(1+m+n+r+\frac{1}{2})\,r!}$$

$$\sum_{p,q=0}^{n+1+m} \left\{ \{(n+1)[1-\rho_h^2(\upsilon)]\}^{1+m+n} G_{n,1,m}(p,q,\upsilon) \right\} \qquad (5.3\text{-}12)$$

$$\mathrm{RE}\left\{G_{13}^{31}\left(\tfrac{1}{4}[c_n^2(\alpha_{fx}-j\mu)^2 \,\middle|\, \begin{matrix}\frac{1}{2}-\frac{1}{2}n-r-p-1 \\ 0,\frac{1}{2},n+m+p\end{matrix}\right)\right\} f_2^n(\mu')\, d\mu$$

$$\mathrm{RE}\left\{G_{13}^{31}\left(\tfrac{1}{4}[c_n^2(\alpha_{fx}-j\mu')^2 \,\middle|\, \begin{matrix}\frac{1}{2}+\frac{1}{2}n-r-q-m \\ 0,\frac{1}{2},n+1-q\end{matrix}\right)\right\} f_2^n(\mu')\, d\mu' \; - \; 1$$

where the $f_2^n(\mu)$ are defined by (5.1-8). As mentioned in connection with that definition, the convolutions in (5.3-12) become less important when the receiver and source are at great distance from the active scattering region provided that the grazing angle ψ is not too small.

The plateau (5.3-12) is plotted versus $\sqrt{2}\sigma_h \sin\psi\Omega_{fx}/c$ in figure (5.3-1) for the limit for which the convolutions can be ignored. It can be seen that similar comments can be made about the behavior of the plateau for this model above and below the critical frequency (at roughly $\Omega_{fx} = c_o$) as were made for the simpler model of chapter 4. The plateau is not dependent on the variance $\sigma_{\chi 1}^2$ of the positional phase fluctuations of the surface since the characteristic function $Q_{\chi 1}(-n,n,\upsilon)$ is unity for $\upsilon = 0$. We note, however, that if $\chi_1(t)$ is gaussian we have

$$Q_{\chi 1}(-n,n,\upsilon) = \exp\{-\sigma_{\chi 1}^2 n^2 [1-\rho_{\chi 1}(\upsilon)]\} \qquad (5.3\text{-}13)$$

$$= [Q_{\chi 1}(-1,1,\upsilon)]^{n2}$$

so that the rate of decay of the plateau of uncertainty with increasing υ is enhanced when $\sigma_{\chi 1}^2$ is large. This is due to the improved averaging over the surface fluctuations. The dependence of the critical frequency only on the Rayleigh parameter suggests that the variable h(t) is analogous to a random delay. Again one must conclude that only the coherent signal energy contains useful information if we use the "on-target" normalized correlator output variance as the critereon of performance. This must be qualified as in chapter 4 by the comments made in Appendix D.

Just as in chapter 4 we may make a more confident evaluation of the tracker performance. Using (5.3-11) and (5.3-10) we have

$$\frac{\partial^2}{\partial\tau^2}\left[\Xi(\tau,\,T,\,\rho)\right] = \frac{c_o^3\,P_x/\Omega_x}{R_d(r_o+r_o')}\sum_{n=0}^{\infty}\frac{1}{2^{n+1}}\,\frac{(2n)!}{(n!)^2}$$

$$\mathrm{RE}\left\{\frac{1}{2\pi\,\Gamma(n+\frac{1}{2})}\;G_{13}^{31}\left(\tfrac{1}{4}c_c^2\left[\alpha_{fx}-j(\tau-\tau_o)\right]^2\,\middle|\,\begin{matrix}\frac{1}{2}-n-1\\0,\;\frac{1}{2},\;-1\end{matrix}\right)\right\} \tag{5.3-14}$$

which is the quantity in the denominator of (3.6-3). The appropriate integrals for the evaluation of the covariance of the tracking error, $R_{\tau o}(\mu)$ are also presented in Appendix H. We write here only the plateau of uncertainty:

$$\left[\frac{\partial^2}{\partial\tau^2}\,\Xi(\tau,T,\rho)\right]_{\tau=\tau_o}^{2}\frac{(P_x/\Omega_x)^2}{R_d^2(r_o+r_o')^2}\sum_{n=0}^{\infty}\int_{-\infty}^{+\infty}\left(\frac{C_n}{n!}\right)^2 Q_{\chi 1}(-n,n,\upsilon)\;\varepsilon_n\,\mathrm{Cos}(n\Omega_s\upsilon)$$

$$\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\frac{1}{[4(n+1)]^{1+m}}\,\frac{A(1,2n+1)}{1!}\,\frac{A(m,2n+1}{m!}\sum_{r=0}^{\infty}\frac{1}{\Gamma(1+m+n+\frac{1}{2})\;\Gamma(1+m+n+r+\frac{1}{2})\;r!}$$

$$\sum_{p,q=0}^{n+1+m}\left\{(n+1)\left[1-\rho_h^2(\upsilon)\right]\right\}^{1+m+n}G_{n,1,m}(p,q,\upsilon)\;\left\{\rho_h^2(\upsilon)\right\}^r$$

$$\frac{1}{(2\pi)^2}\;\mathrm{RE}\left\{-jc_n G_{13}^{31}\left(\tfrac{1}{4}\left[c_n(\alpha_{fx}+j\mu)\right]^2\,\middle|\,\begin{matrix}\frac{1}{2}+\frac{1}{2}n-r-p-1\\0,\frac{1}{4},n+m-p\end{matrix}\right)\right\}\,f_2^n(\mu)\;d\mu$$

$$\left\{+jc_n G_{13}^{31}\left(\tfrac{1}{4}\left[c_n(\alpha_{fx}-j\mu')\right]^2\,\middle|\,\begin{matrix}\frac{1}{2}+\frac{1}{2}n-r-q-m\\0,\frac{1}{2},n+1-q\end{matrix}\right)\right\}\,f_2^n(\mu')\;d\mu'$$

The quantity $R_{\tau o}(0)/c_o^2$ is plotted in figure (5.3-2) versus Ω_{fx}/c_o and can be seen to be essentially the same as figure (5.3-1).

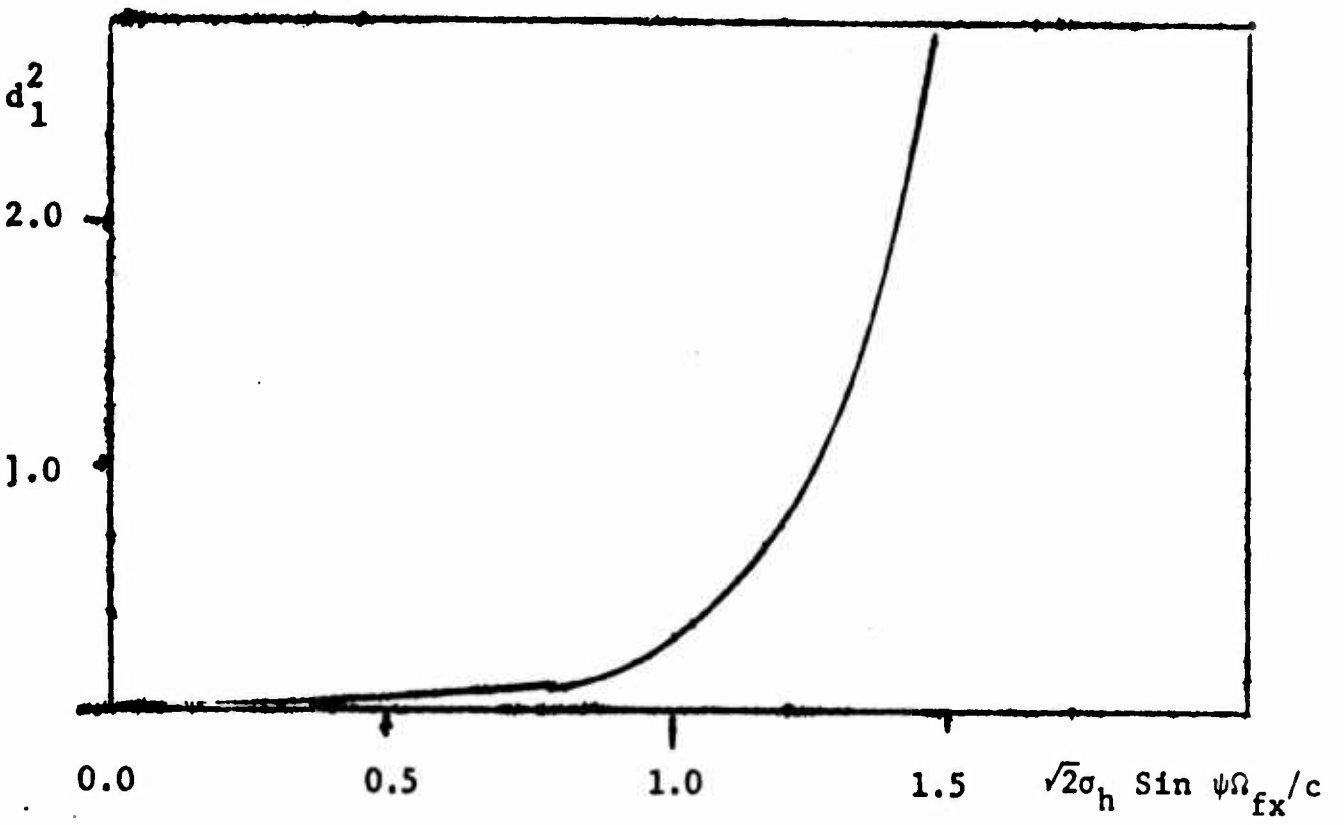


Figure 5.3-1

Normalized Variance Plateau vs. Working Bandwidth

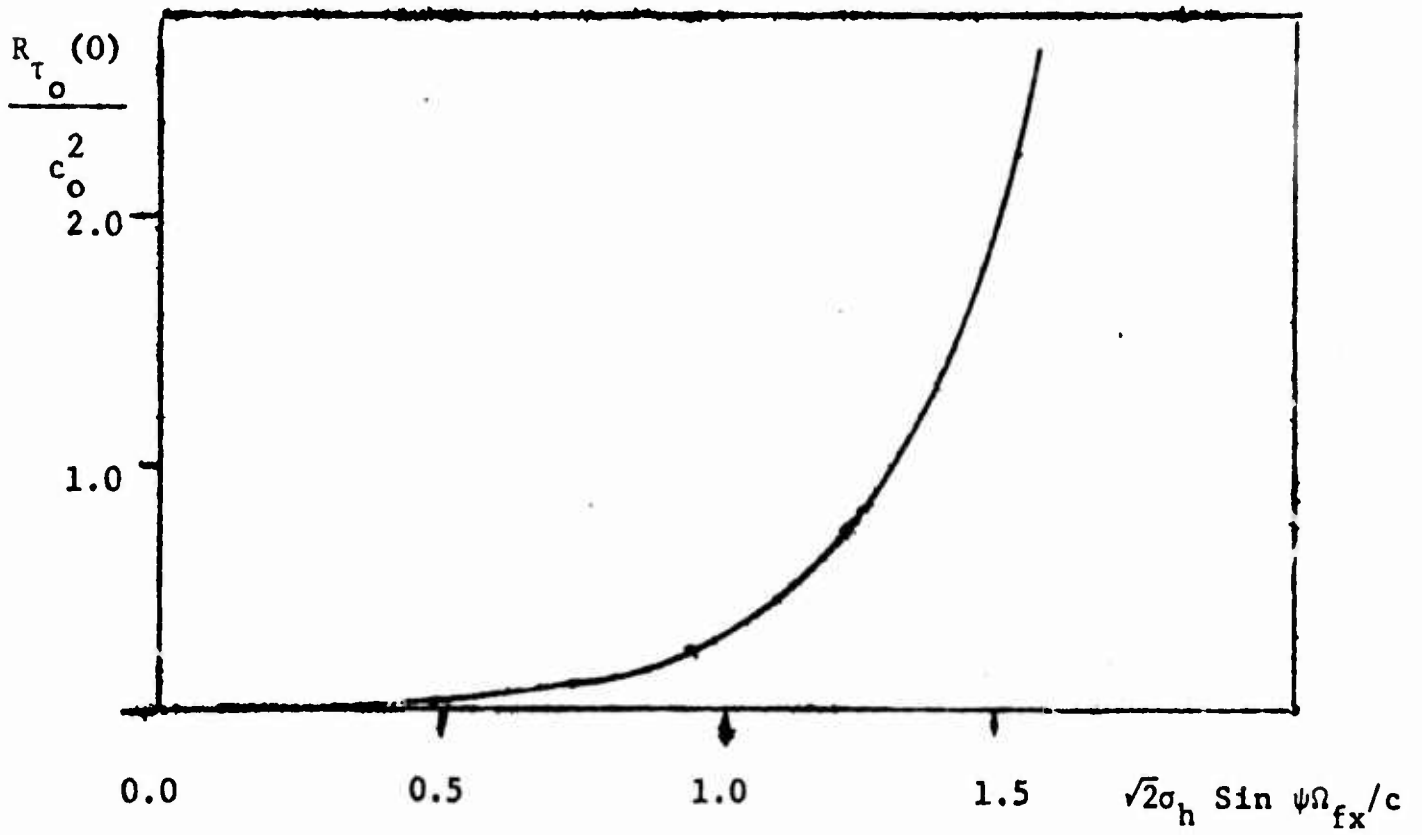


Figure 5.3-2

Normalized Delay Estimate Error vs. Working Bandwidth

5.4 Second and Fourth Order Cross System Statistics for the Two Channel Random Sinusoidal Boundary

In attempting to generalize the single channel transfer function model of (5.1-2) we must take into account the fact that each receiver "sees" different (although possibly overlapping) active regions on the surface. Furthermore, the parameters describing the surface such as the local waveheight, positional phase, or the orientation angle α may be different in these regions. In this section we consider a model in which the local waveheights $h_1(t)$, $h_2(t)$ and positional phases $\chi_1(t)$ and $\chi_2(t)$ are dependent on the relative location of the specular points for the two receivers. The distances from the source to the receivers is considered large compared to the separation between the two receivers.

If the source is located at $Q(x_q,y_q,z_q)$ and if the two receivers are at $P_1(x_{p1},y_{p1},z_{p1})$ and $P_2(x_{p2},z_{p2})$ then the specular points $0_1(x_{o1},y_{o1},0)$ and $0_2(x_{o2},y_{o2},0)$ have coordinates given by

$$x_{oj} = \frac{z_{pj}x_q + x_{pj}z_q}{z_q + z_{pj}} \qquad (5.4\text{-}1a)$$

$$y_{oj} = \frac{z_{pj}y_q + y_{pj}z_q}{z_q + z_{pj}} \qquad (5.4\text{-}1b)$$

The horizontal distance between these two points is thus,

$$\sqrt{(x_{o2}-x_{o1})^2 + (y_{o2}-y_{o1})^2}.$$

This displacement alone would account for a positional phase difference of

$$\chi_s = q_s\{(x_{o2} - x_{o1}) \operatorname{Cos} \alpha_s + (y_{o2} - y_{o1}) \operatorname{Sin} \alpha_s\} \qquad (5.4\text{-}2)$$

However, we consider here a slightly more general model in which the waveheights $h_1(t)$, $h_2(t)$ and positional phases $\chi_1(t)$, $\chi_2(t)$ "seen" by the two receivers are loosely coupled random variables. In order to account for the fixed positional phase displacement (5.4-2) the "initial" phases for the two surface areas are taken as χ_o and $\chi_o + \chi_s$ respectively with χ_o uniformly distributed over the interval $-\pi,\pi$.

The two channels therefore become

$$H_i(\omega,t) = -\frac{e^{-j\tau_{s_i}\omega}}{(r_{oi}+r'_{oi})} \sum_{n=-\infty}^{\infty} J_n\left[A_i(t-r_{oi}/c)\omega\right] e^{-jB_i n^2/\omega}$$
$$\times \exp\left[+jn\Omega_s(t-r_{oi}/c)+jn\chi_i(t-r_{oi}/c)+jn\{\chi_o+(i-1)\chi_s\}\right] \qquad (5.4\text{-}3)$$

for i=1,2

where r_{o1} and r_{o2} are the respective distances of the two receivers from their specular points and where

$$A_i(t) = \frac{2h_i(t)\ \mathrm{Sin}\ \psi_i}{c} \qquad (5.4\text{-}4)$$

$$B_i = \frac{q_s^2 c\ R_{ei}}{4}\left[\frac{\mathrm{Cos}^2\alpha_s}{\mathrm{Sin}^2\psi_i} + \mathrm{Sin}^2\alpha_s\right] \qquad (5.4\text{-}5)$$

$$\tau_{si} = (r_{oi}+r'_{oi})/c \qquad (5.4\text{-}6)$$

$$R_{ei} = \frac{2r_{oi}r'_{oi}}{r_{oi}+r'_{oi}} \qquad (5.4\text{-}7)$$

We make the usual assumptions about the stationarity of the random parameters $h_i(t)$ and $\chi_i(t)$. We further assume that the distances of the source and receivers from the specular points is great enough so

5.4 Second and Fourth Order Cross System Statistics for the Two Channel Random Sinusoidal Boundary

In attempting to generalize the single channel transfer function model of (5.1-2) we must take into account the fact that each receiver "sees" different (although possibly overlapping) active regions on the surface. Furthermore, the parameters describing the surface such as the local waveheight, positional phase, or the orientation angle α may be different in these regions. In this section we consider a model in which the local waveheights $h_1(t)$, $h_2(t)$ and positional phases $\chi_1(t)$ and $\chi_2(t)$ are dependent on the relative location of the specular points for the two receivers. The distances from the source to the receivers is considered large compared to the separation between the two receivers.

If the source is located at $Q(x_q,y_q,z_q)$ and if the two receivers are at $P_1(x_{p1},y_{p1},z_{p1})$ and $P_2(x_{p2},z_{p2})$ then the specular points $0_1(x_{o1},y_{o1},0)$ and $0_2(x_{o2},y_{o2},0)$ have coordinates given by

$$x_{oj} = \frac{z_{pj}x_q + x_{pj}z_q}{z_q + z_{pj}} \qquad (5.4\text{-}1a)$$

$$y_{oj} = \frac{z_{pj}y_q + y_{pj}z_q}{z_q + z_{pj}} \qquad (5.4\text{-}1b)$$

The horizontal distance between these two points is thus,

$$\sqrt{(x_{o2}-x_{o1})^2 + (y_{o2}-y_{o1})^2}.$$

This displacement alone would account for a positional phase difference of

$$\chi_s = q_s\{(x_{o2} - x_{o1}) \operatorname{Cos} \alpha_s + (y_{o2} - y_{o1}) \operatorname{Sin} \alpha_s\} \qquad (5.4\text{-}2)$$

However, we consider here a slightly more general model in which the waveheights $h_1(t)$, $h_2(t)$ and positional phases $\chi_1(t)$, $\chi_2(t)$ "seen" by the two receivers are loosely coupled random variables. In order to account for the fixed positional phase displ :ement (5.4-2) the "initial" phases for the two surface areas are taken as χ_o and $\chi_o + \chi_s$ respectively with χ_o uniformly distributed over the interval $-\pi,\pi$.

The two channels therefore become

$$H_i(\omega,t) = \frac{e^{-j\tau_{s_i}\omega}}{(r_{oi}+r'_{oi})} \sum_{n=-\infty}^{\infty} J_n\left[A_i(t-r_{oi}/c)\omega\right] e^{-jB_i n^2/\omega}$$

$$\times \exp\left[+jn\Omega_s(t-r_{oi}/c)+jn\chi_i(t-r_{oi}/c)+jn\{\chi_o+(i-1)\chi_s\}\right] \qquad (5.4\text{-}3)$$

$$\text{for } i=1,2$$

where r_{o1} and r_{o2} are the respective distances of the two receivers from their specular points and where

$$A_i(t) = \frac{2h_i(t)\ \mathrm{Sin}\ \psi_i}{c} \qquad (5.4\text{-}4)$$

$$B_i = \frac{q_s^2 c\ R_{ei}}{4}\left[\frac{\mathrm{Cos}^2\alpha_s}{\mathrm{Sin}^2\psi_i} + \mathrm{Sin}^2\alpha_s\right] \qquad (5.4\text{-}5)$$

$$\tau_{si} = (r_{oi}+r'_{oi})/c \qquad (5.4\text{-}6)$$

$$R_{ei} = \frac{2r_{oi}r'_{oi}}{r_{oi}+r'_{oi}} \qquad (5.4\text{-}7)$$

We make the usual assumptions about the stationarity of the random parameters $h_i(t)$ and $\chi_i(t)$. We further assume that the distances of the source and receivers from the specular points is great enough so

that $\psi_1 = \psi_2 = \psi$, and that the dimensions of the array are small enough that so that the retardation time difference $(r_{o2} - r_{o1})/c$ is negligible on the time scale of the slow time variations of the waveheight and positional phase variations. The spherical attenuation factor $1/(r_{oi}+r'_{oi})$ is replaced by $1/(r_o+r'_o)$ where r_o and r'_o are nominal values for r_{oi} and r'_{oi}. Thus,

$$\Phi_{ik}(\omega,\omega',\tau) =$$

$$\frac{e^{-j(\tau_{si}\omega-\tau_{sk}\omega')}}{(r_o+r'_o)^2} \sum_{n=0}^{\infty} \frac{1}{(n!)^2} \exp\left\{jn^2\left(\frac{B_i}{\omega} - \frac{B_k}{\omega'}\right)\right\} Q_{\chi i\chi k}(-n,n,\tau)\ \varepsilon_n \tag{5.4-8}$$

$$\mathrm{Cos}\left[n(\Omega_s\tau+\chi_s)\right] \sum_{1=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{[4(n+1)]}\,1+m\ \frac{A(1,2n+1)}{1}\ \frac{A(m,2n+1)}{m!} \sum_{r=0}^{\infty} \frac{\Gamma(1+m+n+r+\frac{1}{2})}{\Gamma(1+m+n+\frac{1}{2})\,r!}$$

$$\sum_{p,q=0}^{n+1+m} \{(n+1)\left[1-\rho^2_{h_ih_k}(\tau)\right]\}^{1+n+m}\ G_{n,1,m}(p,q,\tau)$$

$$\{\rho_{h_ih_k}(\tau)\}^r\ \frac{\xi_n^{2(r+p+1)+n}\ \xi_n'^{2(r+q+m)+n}}{\left[(\xi_n^2+1)\ (\xi_n'^2+1)\right]^{n+1+m+r+\frac{1}{2}}}$$

where we have assumed that the two channels have similar statistical properties (i.e. $\sigma_{h1} = \sigma_{h2} = \sigma_h$). The waveheights $h_i(t)$ are taken to be independent of the positional phase fluctuations $\chi_i(t)$. Again, asserting these parameters to be Gaussian random variables we achieve fourth order cross stationarity. The fourth order system moment is, however, a great deal more complicated owing to a more complex cancellation

of positional phases:

$$\Phi_{12}(\omega,\omega',\omega'',\mu,\mu',\mu'') =$$

$$\overline{H_1(\omega,t)\ H_1(\omega',t-\mu)^*\ H_2(\omega'',t-\mu')\ \ H_2(\omega''',t-\mu'')^*} =$$

$$\frac{e^{-j\tau_{s1}(\omega-\omega')-\ j\tau_{s2}(\omega''-\omega''')}}{(r_o+r_o')^4} \sum_{n=o}^{\infty}\sum_{n=o}^{\infty}\sum_{n=o}^{\infty}\sum_{n=o}^{\infty} \tag{5.4-9}$$

$$\overline{J_{n1}\left(\frac{2\omega\ h_1(t)\ \mathrm{Sin}\ \psi_1}{c}\right)\ J_{n2}\left(\frac{2\omega'\ h_1(t-\mu)\ \mathrm{Sin}\ \psi_1}{c}\right)}$$

$$\cdot\ \overline{J_{n3}\left(\frac{2\omega''h_2(t-\mu'')\ \mathrm{Sin}\ \psi_2}{c}\right)\ J_{n4}\left(\frac{2\omega'''h_2(t-\mu''')\mathrm{Sin}\ \psi_2}{c}\right)}$$

$$\overline{e^{+j(n_1-n_2+n_3-n_4)(\Omega_s t+\chi_o)\ +\ j(n_2\mu-n_3\mu'+n_4\mu'')\ \Omega_s}\ \ e^{+j(n_3-\ n_4)\chi_s}}$$

$$\exp\left\{-jB_1\left(\frac{n_1}{\omega}-\frac{n_2}{\omega'}\right)-jB_2\left(\frac{n_3}{\omega'},\ \ \frac{n_4}{\omega''},\right)\right\} Q^{[4]}_{\chi_1\chi_2}(n_1,-n_2,n_3,-n_4,\mu,\mu',\mu'')$$

where

$$Q^{[4]}_{\chi_1\chi_2}(n_1,n_2,n_3,n_4,\mu,\mu',\mu'') = \tag{5.4-10}$$

$$\overline{\exp\{+j[n_1\chi_1(t)+n_2\chi_1(t-\mu)+n_3\chi_2(t-\mu')+n_4\chi_2(t-\mu'')]\}}$$

is the fourth order characteristic function for the time varying positional phase fluctuations. Carrying out the average over the "initial" phase χ_o we find that only those terms for which

$$n_1 - n_2 + n_3 - n_4 = 0 \tag{5.4-11}$$

are retained, thereby yielding fourth order stationarity. It is also clear that for any four numbers n_1,n_2,n_3,n_4 that satisfy this condition, the numbers $-n_1$, $-n_2,-n_3,-n_4$ satisfy it too. For example, the following are the possible combinations for $|n_1|+|n_2|+|n_3|+|n_4| \leq 2$:

n_1	n_2	n_3	n_4
0	0	0	0
1	1	0	0
0	0	1	1
1	0	0	1
0	1	1	0
1	0	-1	0
0	1	0	-1
-1	-1	0	0
0	0	-1	-1
-1	0	0	-1
0	-1	-1	0
-1	0	1	0
0	-1	0	1

Figure 5.4-1

The proliferation of terms beyond these of second order imposes a practical limit on the usefulness of this model to weak scattering.

In order to apply the expansion (5.2-4) for the Bessel function to (5.4-9) we must first isolate Bessel functions of negative order by utilizing the identity

$$J_{-n}(z) = (-1)^n J_n(z) \tag{5.4-12}$$

This can be rewritten as

$$J_n(z) = j^{(|n|-n)} J_{|n|}(z) \tag{5.4-13}$$

which leaves the Bessel function unaltered if n is positive but

yields the appropriate factor of $(-1)^n$ if n is negative. Using this (5.4-9) becomes

$$\Phi^{[4]}_{1,2}(\omega,\omega',\omega'',\omega''',\nu,\nu'') = \frac{e^{-j\tau_{s1}(\omega-\omega')-j\tau_{s2}(\omega''-\omega''')}}{(r_o + r'_o)^4} \sum_{n_1-n_2+n_3-n_4=0}$$

$$\sum_{m_1=o}^{\infty}\sum_{m_2=o}^{\infty}\sum_{m_3=o}^{\infty}\sum_{m_4=o}^{\infty}\prod_{i=1}^{4}\left[\frac{1^{|n_i|-n_i}}{|n_i|![4(|n_i|+1)]^{m_i}}\ \frac{A(m_i,2|n_i|+1)}{(m_i)!}\right]$$

$$D'_{\underline{n},\underline{m}}\ \mu'_{\underline{n},\underline{m}}(\underline{n})\ \exp\{jB_1|\frac{n_1}{\omega}-\frac{n_2}{\omega'}| + jB_2|\frac{n_3}{\omega''}-\frac{n_4}{\omega'''}|\}\ e^{j(n_3-n_4)\chi_s}$$

$$Q^{[4]}_{\chi 1\chi 2}(n_1,-n_2,n_3-n_4,\nu,\nu',\nu'')\ \exp\{j(n_2\nu-n_3\nu'+n_4\nu'')\} \qquad (5.4\text{-}14)$$

where

$$\underline{\omega} = \{\omega,\omega',\omega'',\omega'''\} \qquad (5.4\text{-}15)$$

and

$$\underline{m} = \{m_1,m_2,m_3,m_4\} \qquad \underline{n} = \{n_1,n_2,n_3,n_4\} \qquad (5.4\text{-}16)$$

The $\mu'_{\underline{n},\underline{m}}(\underline{n})$ are moments computed using the distribution for

$$\underline{h} = \{h_1(t),h_1(t-\nu),h_2(t-\nu'),h_2(t-\nu'')\} \qquad (5.4\text{-}17)$$

modified in a manner anologous to that used in (5.2-11). We assume that $\underline{h}$ is Gaussian with a correlation matrix $\underline{R}_{h1h2}(\underline{\nu})$ which we here abbreviate simply as $\underline{R}$. Again, using transformations similar to those used to obtain equation (5.2-11) we have

$$\mu'_{\underline{n},\underline{m}}(\underline{n}) = \iiiint h_1(t)^{2m_1+|n_1|}\ h_1(t-\nu)^{2m_2+|n_2|}\ h_2(t-\mu')^{2m_3+|n_3|}$$

$$h_2(t-\nu'')^{2m4+|n4|}\ \frac{e^{-\frac{1}{2}\underline{h}^T\underline{R}'^{-1}_{\underline{n}}\underline{h}}}{\sqrt{(2\Pi)^4\ |\underline{R}'_{\underline{n}}|}}\ dh_1(t)\ dh_1(t-\nu)\ dh_2(t-\nu')\ dh_2(t-\nu'') \qquad (5.4\text{-}18)$$

where the correlation matrix $\underline{R'_n}$ (which is a function of both the n_i and of the frequencies $\underline{\omega}$) for the modified density is formed as follows :

$$\underline{R'_n} = (\underline{W_n} + R^{-1})^{-1} = \underline{U_n}^{-1} \tag{5.4-19}$$

Here the matrix $\underline{W_n}$ is a diagonal matrix which contains the dependence on frequency :

$$\begin{vmatrix} \frac{\omega^2}{|n_1|+1} & 0 & 0 & 0 \\ 0 & \frac{\omega'^2}{n_2+1} & 0 & 0 \\ 0 & 0 & \frac{\omega''^2}{|n_3|+1} & 0 \\ 0 & 0 & 0 & \frac{\omega'''^2}{|n_4|+1} \end{vmatrix}$$

The normalizing constant of the modified density $D'_{\underline{n},\underline{m}}(\omega)$ of equation (5.4-14) is given by

$$D'_{\underline{n},\underline{m}} = \prod_{i=1}^{4} \left[\left(\frac{\sin \psi}{c} \right) \omega_i^{\,2m_i+|n_i|} \right] \sqrt{\frac{1}{|\underline{U_n}||R|}} \tag{5.4-21}$$

where $\underline{U_n}$ is the universe of $\underline{R'_n}$. Expressions for the moments generated by equation (5.4-18) may be derived by setting the equivalent order cumulant for $\underline{h}$ equal to zero since the parent distribution is Gaussian. Lanning and Battin[49] present the general moment as a combination of second order moments. From this fact alone we may conclude that similar to equation (5.2-19) one may write these

moments as polynomials divided by the determinant of $\underline{U}_n$ raised to various powers :

$$\mu'_{\underline{n},\underline{m}}(\underline{n}) = \frac{P_{\underline{n},\underline{m}}(\underline{\omega})}{|U_n|_1^{\Sigma \frac{1}{2} |n_i|+m_i}} \qquad (5.4\text{-}22)$$

If we retain only those terms for the values of the n_i tabulated in figure (5.4-1) then only second order moments are needed and these are the elements of the $R'_{\underline{n}}$ matrix itself. Using only to the second order in frequency times waveheight

$$\Phi^{[4]}_{1,2}(\omega,\omega',\omega'',\omega''';\underline{\nu}) = \frac{e^{-j\tau_{s1}(\omega-\omega')+j\tau_{s2}(\omega''-\omega''')}}{(r_o+r'_o)^4} |\underline{R}|^{-\frac{1}{2}} \qquad (5.4\text{-}23)$$

$$\Big[|\underline{U}_{0000}|^{-\frac{1}{2}} \{1+ \frac{\mathrm{Sin}^2\psi}{c^2} [\omega^2 R'_{11}(\underline{0}) + \omega'^2 R'_{22}(\underline{0}) + \omega''^2 R'_{33}(\underline{0}) + \omega'''^2 R'_{44}(\underline{0})]\}$$

$$+ \frac{\mathrm{Sin}^2\psi}{c^2} \{\omega\omega' |\underline{U}_{1100}|^{-\frac{1}{2}} R'_{12}(1100)\, Q^{(4)}_{\chi 1\chi 2}(1,-1,0,0,\underline{\nu}) \mathrm{Cos}(\Omega_s \nu) \exp\{jB_1\left(\frac{1}{\omega}-\frac{1}{\omega'}\right)\}$$

$$+ \omega''\omega''' |\underline{U}_{0011}|^{-\frac{1}{2}} R'_{34}(0011)\, Q^{(4)}_{\chi 1\chi 2}(0,0,1,-1)$$

$$\mathrm{Cos}(\Omega_s(\nu'-\nu'')+2\chi_s) \exp\{jB_2\left(\frac{1}{\omega}-\frac{1}{\omega'''}\right)\}$$

$$+ \omega\omega''' |\underline{U}_{1001}|^{-\frac{1}{2}} R'_{14}(1001)\, Q^{(4)}_{\chi 1\chi 2}(1,0,0,-1,\underline{\nu}) \mathrm{Cos}(\Omega_s \nu''+\chi_s) \exp\{j\left(\frac{B_1}{\omega}-\frac{B_2}{\omega'''}\right)\}$$

$$+ \omega'\omega'' |\underline{U}_{0110}|^{-\frac{1}{2}} R'_{23}(0110)\, Q^{(4)}_{\chi 1\chi 2}(0,-1,1,0,\underline{\nu})$$

$$\mathrm{Cos}(\Omega_s(\nu-\nu')+\chi_s) \exp\{-j\left(\frac{B_1}{\omega'}-\frac{B_2}{\omega''}\right)$$

$$+ \omega\omega'' |\underline{U}_{1010}|^{-\frac{1}{2}} R'_{13}(1010)\, Q^{(4)}_{\chi 1\chi 2}(1,0,-1,0,\underline{\nu}) \mathrm{Cos}(\Omega_s \nu'-\chi_s) \exp\{j\left(\frac{B_1}{\omega}+\frac{B_2}{\omega''}\right)\}$$

$$+ \omega'\omega''' |\underline{U}_{0101}|^{-\frac{1}{2}} R'_{24}(0101)\, Q^{(4)}_{\chi 1\chi 2}(0,-1,0,1,\underline{\nu})$$

$$\mathrm{Cos}(\Omega_s(\nu-\nu'')-\chi_s) \exp\{-j\left(\frac{B_1}{\omega'}+\frac{B_2}{\omega'''}\right)\}\}\Big]$$

where $\underline{\nu} = (\nu,\nu',\nu'')$, and where the $R'_{ij}(\underline{n})$ are elements of the matrix $\underline{R}'$ given by

$$R'_{ij}(\underline{n}) = \frac{|\underline{U}_{\underline{n}}|_{ij}(-1)^{i+j}}{|\underline{U}_{\underline{n}}|} \qquad (5.4\text{-}24)$$

The $|\underline{U}_{\underline{n}}|_{ij}$ are the ij-th minors of the determinant $|\underline{U}_{\underline{n}}|$ and (5.4-24) is a special case of (5.4-22).

Of crucial importance to the performance of various integrals for the variance calculations is the functional dependence of $|\underline{U}|$ on the frequencies $\underline{\omega}$ since this determinant is raised to a fractional power in all of the terms in (5.4-23). We denote the inverse of the correlation matrix for the waveheights $\underline{R}$ by $\underline{M}$ and define normalized frequencies in this case by

$$\xi_{n_1} = \sqrt{\frac{2\beta}{|n_1|+1}}\ \frac{\sigma_h \operatorname{Sin}\psi\ \omega}{c}\ , \qquad (5.4\text{-}25)$$

with similar connections for the pairs (ξ_{n2},ω'), (ξ_{n3},ω'') and (ξ_{n4},ω'''). The constant β is chosen later through convergence considerations.

$$|\underline{U}_{\underline{n}}| = \begin{vmatrix} M_{11}+\dfrac{\xi^2_{n_1}}{\beta\sigma_h^2} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22}+\dfrac{\xi^2_{n_2}}{\beta\sigma_h^2} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33}+\dfrac{\xi^2_{n3}}{\beta\sigma_h^2} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44}+\dfrac{\xi^2_{n4}}{\beta\sigma_h^2} \end{vmatrix}$$

Just as the minor $|\underline{M}|_{ij}$ is formed from the determinant $|\underline{M}|$ by deleting the i-th row and j-th column, we define a second order minor $|\underline{M}|_{ij,kl}$

which is formed by deleting both the i-th and k-th rows and the j-th and ı-th columns from $|\underline{M}|$. For example,

$$|\underline{M}|_{21,43} = \begin{vmatrix} M_{12} & M_{14} \\ M_{32} & M_{34} \end{vmatrix} \tag{5.4-27}$$

Similarly, minors of third order $|\underline{M}|_{ij,kı,mn}$ are the elements of $\underline{M}$ itself. Using this notation,

$$|\underline{U}_{\underline{n}}| = \frac{\xi_{n_1}^2 \xi_{n_2}^2 \xi_{n_3}^2 \xi_{n_4}^2}{\beta^4 \sigma_h^8} + \sum_{i\neq j\neq k} \frac{\xi_{n_i}^2 \xi_{n_j}^2 \xi_{n_k}^2}{\beta^3 \sigma_h^6} |M|_{ii,jj,kk}$$

$$+ \sum_{i\neq j} \frac{\xi_{n_i}^2 \xi_{n_j}^2}{\beta^2 \sigma_h^4} |\underline{M}|_{ii,jj} + \sum_{i} \frac{\xi_{n_j}^2}{\beta \sigma_h^2} |\underline{M}|_{ii} + |M| \tag{5.4-28}$$

where again,

$$\underline{M} = \underline{R}^{-1} \tag{5.4-29}$$

Unfortunately, this polynomial is of eighth order in frequency. Although for the applications of interest in equations (3.5-16) through (3.5-18) the squares of various frequencies are taken equal in pairs, (5.4-28) is still more formidable than the denominator of (5.2-21). However, we may use a series expansion again which is valid for all frequencies and also produces convenient factorization of frequency dependence

$$\frac{1}{(|\underline{U}_{\underline{n}}||\underline{R}|)\nu} = \frac{1}{[(\xi_{n_1}^2+1)(\xi_{n_2}^2+1)(\xi_{n_3}^2+1)(\xi_{n_4}^2+1)]\nu} \times$$

$$\left[1 + \frac{g(\xi_{n1},\xi_{n2},\xi_{n3},\xi_{n4})}{[(\xi_{n_1}^2+1)(\xi_{n_2}^2+1)(\xi_{n_3}^2+1)(\xi_{n_4}^2+1)]} \right]^{\nu} \tag{5.4-30}$$

where $\underline{p} = (p_1,p_2,p_3,p_4)$ and from (5.4-32)

$$[-g(\xi_{n_1},\xi_{n_2},\xi_{n_3},\xi_{n_4})]^r = \sum_{\underline{q}} C_r(\underline{q},\underline{\nu})\, \xi_{n_1}^{2q_1}\, \xi_{n_2}^{2q_2}\, \xi_{n_3}^{2q_3}\, \xi_{n_4}^{2q_4} \tag{5.4-34}$$

We may also write (5.4-21) in terms of the normalized frequencies

$$D_{\underline{n},\underline{m}}(\underline{\omega}) = \sqrt{\frac{1}{|U_n||\underline{R}|}} \prod_{i=1}^{4} \left(\frac{\xi_{h_i}^2}{\beta\sigma_h^2} \right)^{(|n_i|+1)} |n_i| + 2m_i. \tag{5.4-35}$$

Just as in the case of the second order moment, the coefficients in (5.4-32) are singular for $\underline{\nu} = (0,\tau_o,\tau_o)$ where

$$\tau_o = \tau_{s1} - \tau_{s2}. \tag{5.4-36}$$

is the mean path delay difference for the two receivers. However, the product $|\underline{R}|^{\Sigma\frac{1}{2}|n_i|+m_i}\, G_{\underline{n},\underline{m}}(\underline{p},\underline{\nu})$ remains finite.

Finally, we summarize by writing

$$D'_{\underline{n},\underline{m}}\, \mu'_{\underline{n},\underline{m}}(n) = \sum_{r=o}^{\infty} \frac{\Gamma(\Sigma(|n_i|+m_i)+r+\frac{1}{2})}{\Gamma(\Sigma(|n_i|+m_i)+\frac{1}{2})} \frac{1}{r!} |\underline{R}|^{\sum\frac{1}{2}|n_i|+m_i} \tag{5.4-37}$$

$$\sum_{\underline{p}} \sum_{\underline{q}} G_{\underline{n},\underline{m}}(\underline{p},\underline{\nu})\, C_r(\underline{q},\underline{\nu}) \left\{ \prod_{i=1}^{4} \left[\frac{(|n_i|+1)}{\beta\,\sigma_h^2} \right]^{|n_i|+m_i} \xi^{(p_i+m_i+q_i)+|n_i|} \right\}$$

$$\overline{[(\xi_{n_1}^2+1)(\xi_{n_2}^2+1)(\xi_{n_3}^2+1)(\xi_{n_4}^2+1)]_i^{\Sigma(\frac{1}{2}|n_i|+m_i)+r+\frac{1}{2}}}$$

The results may be substituted into (5.4-14) to complete that expression.

CHAPTER 6

POSSIBLE EXTENSIONS OF THE PRESENT WORK AND SUGGESTIONS FOR FUTURE RESEARCH

6.0 Extensions to the Fully Random Boundary

In chapter 5 we have considered the scattering surface to be a sinusoidal boundary with random parameters. More generally we are interested in an arbitrary random boundary $\zeta(x,y,t)$ with known statistical properties. In particular, with the assumption of gaussian statistics, only the spatial and temporal correlation function

$$\Psi_\zeta(\xi,\eta,\tau) = \overline{\zeta(x,y,t)\ \zeta(x-\xi,y-\eta,t-\tau)} \qquad (6.0\text{-}1)$$

need be specified.

The coherent transfer function is easily obtained from (1.4-3) in the limit $B_e(\theta,\phi) \to 1$

$$H_c(\omega) = \frac{-e^{j\omega(r_o+r_o')/c}}{(r_o+r_o')}\ Q_{\zeta_1}(\omega \mathrm{Sin}\psi/c) \qquad (6.0\text{-}2)$$

$$= \frac{-e^{j\tau_s\omega}}{(r_o+r_o')}\ e^{-\frac{\omega^2\sigma_\zeta^2\,\mathrm{Sin}^2\psi}{c^2}}$$

The result is the same as Eckart's reflection coefficient (1.2-21) for directional reception.

Unfortunately, equation (6.0-2) is about the only trivial statement one can make about the fully random surface model. If we advance to second order moments we immediately run into difficulty. From (1.4-13) we have

$$\Phi(\omega,\omega';\tau) = \frac{4\omega\omega'\mathrm{Sin}^2\psi}{c^2} \times e^{-j\tau_s(\omega-\omega')}$$

$$\left(\frac{1}{r_o r_o'}\right)^2 \int_{-\infty}^{\infty}\int\int\int dx_1\, dx_2\, dy_1\, dy_2\, B_e(\theta_1,\phi_1)\, B_e(\theta_2,\phi_2)$$

$$\exp\left\{-j\left[\tfrac{1}{4}(\omega x_1^2 - \omega' x_2^2)\ (c_s^o)^2 + (\omega y_1^2 - \omega' y_2^2)\right]\ \frac{1}{cR_e}\right.$$

$$\frac{}{e^{+j\left[\zeta(x_1,y_1,t)\omega - \zeta(x_2,y_2,t-\tau)\ \omega'\right]\ \mathrm{Sin}\psi/c}}$$

$$dx_1\ dx_2\ dy_1\ dy_2 \qquad (6.0\text{-}3)$$

Again, to simplify matters we take $B_e(\theta,\phi) \to 1$ and assume gaussian statistics for ζ to get

$$\Phi(\omega,\omega';\tau) = \frac{4\omega\omega'\ \mathrm{Sin}^2\psi}{c^2}\ \left(\frac{1}{r_o r'_o}\right)^2$$

$$e^{-\frac{1}{2}\left(\frac{\sigma_\zeta\ \mathrm{Sin}\psi}{c}\right)^2}\ \ (\omega^2+\omega'^2)\ \ e^{-j\tau_s(\omega-\omega')}$$

$$\int\int_{-\infty}^{\infty}\int\int\ \exp\left\{-j\left[(\omega x_1^2 - \omega' x_2^2)\ (c_s^o)^2 + (\omega y_1^2 - \omega' y_2^2)\right]\ \frac{1}{cR_e}\right\}$$

$$\exp\left\{\frac{\omega\omega'}{c^2}\ \Psi_\zeta(x_1-x_2,y_1-y_2,\tau)\ \mathrm{Sin}^2\psi\right\}$$

$$dx_1\ dx_2\ dy_1\ dy_2 \qquad (6.0\text{-}4)$$

The most difficult feature of this integral is the dependence of Ψ_ζ only on the differences

$$x_1 - x_2,\ y_1 - y_2$$

whereas the remainder of the integrand is dependent on the quantities

$$\omega x_1^2 - \omega' x_2^2$$

and

$$\omega y_1^2 - \omega' y_2^2$$

In certain cases transformations or reductions of the integral in (6.0-4) are possible. For example, if only the value of $\Phi(\omega,\omega';\tau)$ for $\omega = \omega'$ is desired, then one may write

$$\omega(x_1^2 - x_2^2) = \omega(x_1 + x_2)(x_1 - x_2) \tag{6.0-5}$$

$$\omega(y_1^2 - y_2^2) = \omega(y_1 + y_2)(y_1 - y_2) \tag{6.0-6}$$

Making the substitutions

$$\begin{aligned} \xi &= x_1 - x_2 & \xi' &= x_1 + x_2 \\ \eta &= y_1 - y_2 & \eta' &= y_1 + y_2 \end{aligned} \tag{6.0-7}$$

we find that

$$\Phi(\omega,\omega,\tau) = \left(\frac{2\omega \sin\psi}{c}\right)^2 \left(\frac{1}{r_o r'_o}\right)^2$$

$$e^{-\left(\frac{\sigma\zeta \sin\psi}{c}\,\omega\right)^2} \int_{-\infty}^{\infty}\int\int\int \tag{6.0-8}$$

$$\exp\left\{-j\left[\xi\xi'(c_s^o)^2 + \eta\eta'\right]\frac{\omega}{cR_e}\right\}$$

$$\times \exp\left\{\left(\frac{\omega \sin\psi}{c}\right)^2 \Psi(\xi,\eta,\tau)\right\}$$

$$d\xi\, d\eta\, d\xi'\, d\eta'$$

One can immediately perform the integral over the primed coordinates:

$$\int_{-\infty}^{\infty}\int e^{-j\left[\xi\xi'(c_s^o)^2 + \eta\eta'\right]\omega/cR_e}\, d\xi'\, d\eta' \tag{6.0-9}$$

$$= \delta\left(\frac{\xi(c_s^o)^2\,\omega}{cR_e}\right)\ \delta\left(\frac{\eta\omega}{cR_e}\right)$$

Thus,

$$\Phi(\omega,\omega,\tau) = \frac{1}{(r_o+r'_o)^2} \; e^{-\left\{\left(\frac{\sigma_\zeta \omega \mathrm{Sin}\psi}{c}\right)^2 \left[1 - p_\zeta(\tau)\right]\right\}} \tag{6.0-10}$$

where

$$p_\zeta(\tau) = \frac{\Psi_{-\zeta}(o,o,\tau)}{\sigma_\zeta^2} \tag{6.0-11}$$

As a corollary one can conclude that

$$\Phi(\omega,\omega,o) = \overline{H(\omega,t)\; H(\omega,t)^*}$$

$$= \frac{1}{(r_o+r'_o)^2} \tag{6.0-12}$$

This last result is true due to the stationarity of ζ and does not depend on the gaussian assumption. It does however, draw very heavily on the truncation of the expansion for the exponent in (1.4-13) at quadratic orders in x and y and at first order in ζ. While (6.0-12) gives the very simplistic average energy transmission which is flat over an infinite bandwidth, it does at least conserve energy. Furthermore, from (2.4-30) it follows that the autocorrelation function for a wideband signal is left largely unaltered by this scattering model.

These statements are quite general and qualitatively correct. However, nothing much can be said about the fluctuation of the correlator output for the multipath detector unless we can obtain $\Phi(\omega,\omega';\tau)$ for $\omega \neq \omega'$ in (3.3-12) for $I_4(\tau,\tau';\nu)$. In this case the transformations (6.0-5) and (6.0-6) are not possible one must be prepared to do each integral over x_1, x_2, y_1 and y_2 separately. This is clear since either ω or ω' may be chosen to be zero independently.

In any case, if it is desired to describe the decrease in the plateau of variance of the correlator output for long integration times one must have an expression for $\Psi(\xi,\eta,\tau)$. This in itself is a non-trivial task.

In general $\Psi_\zeta(\xi,\eta,\tau)$ satisfies the same wave equation as $\zeta(x,y,t)$. If we choose to ignore dispersion in these waves then

$$\left\{ \frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2} - \frac{1}{c^2} \frac{\partial^2}{\partial\tau^2} \right\} \Psi_\zeta(\xi,\eta\,\tau) = 0 \qquad (6.0\text{-}13)$$

Therefore, the behavior of $\Psi(\xi,\eta,\tau)$ as a function of τ for a given $\Psi_\zeta(\xi,\eta,o)$ is that of an initial deformation on an infinite membrane. The progression of $\Psi_\zeta(\xi,\eta,\tau)$ is therefore determined by a superposition of linear surface waves,

$$\Psi_\zeta(\xi,\eta,\tau) = \int_{-\infty}^{\infty}\int E_\zeta(q_x,q_y)\, e^{j\left[q_x x + q_y y - \Omega(q_x,q_y)\,\tau\right]}\, dq_x\, dq_y \qquad (6.0\text{-}14)$$

where

$$\Omega(q_x,q_y) = c_s\sqrt{1-q_x^2-q_y^2} \qquad (6.0\text{-}15)$$

and c_s is the propagation velocity for surface waves. The quantity $E_\zeta(q_x,q_y)$ is the space spectrum [5] for ζ and is determined from $\Psi_\zeta(\xi,\eta,o)$

$$E_\zeta(q_x q_y) = \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty}\int \Psi_\zeta(\xi,\eta,o)\, e^{-j\,(q_x x + q_y y)}\, dq_x\, dq_y \qquad (6.0\text{-}16)$$

The integral (6.0-14) is often difficult to perform.

In order to gain insight into the behavior of Ψ_ζ let us assume that E_ζ is rotationally symmetric in the q_x q_y plane so that

$$E_\zeta(q_x,q_y) = E_\zeta\ (q) \tag{6.0-17}$$

$$q = \sqrt{q_x^2+q_y^2} \tag{6.0-18}$$

We may then write (6.0-14) as

$$\Psi_\zeta(r,\tau) = \int_0^\infty E_\zeta(q_x q_y)\ J_o\left(\frac{rq}{c_s}\right)\cos(q\tau)q\,d\tau \tag{6.0-19}$$

For example, assume that

$$E_\zeta(q) = \frac{e^{-\frac{1}{2}q^2/\Lambda^2}}{q} \tag{6.0-20}$$

Substituting (6.0-20) into (6.0-19) and using the expansion (5.2-3) for the Bessel function

$$\Psi_\zeta(r,\tau) = \sum_{m=0}^{\infty} \frac{A(m,2)}{m!} \left\{ \left(\frac{r}{2c_s}\right)^{2m} \right. \tag{6.0-21}$$
$$\left. \int_0^\infty q^{2m}\ e^{-\frac{q^2}{2}} \left(\frac{1}{\Lambda^2} + \frac{r^2}{2c_s^2}\right)\ \mathrm{Cos}(q\tau)\ dq \right)$$

$$= \sum_{m=0}^{\infty} \frac{A(m,2)}{m!} \left(\frac{r}{2c_s}\right)^{2m} \Gamma(2m+1)$$

$$\tfrac{1}{2}\left(\frac{1}{\Lambda^2} + \frac{r^2}{2c_s^2}\right)^{-m-\frac{1}{2}} \qquad e^{-\frac{\tau^2}{4\left(\frac{1}{\Lambda^2} + \frac{r^2}{2c_s^2}\right)}}$$

$$D_{-2m-1}\left(\frac{i\tau}{\sqrt{\frac{1}{\Lambda^2} + \frac{r^2}{2c_s^2}}}\right) + D_{-2m-1}\left(\frac{-i\tau}{\sqrt{\frac{1}{\Lambda^2} + \frac{r^2}{2c_s^2}}}\right)$$

where the D_{-n} are parabolic cylinder functions. For small r and τ the first term is informative:

$$\Psi_\zeta(r,\tau) \equiv \frac{1}{\sqrt{2\pi}} \qquad \frac{e^{\frac{\tau^2}{2\left(\frac{1}{\Lambda^2} + \frac{r^2}{2c_s^2}\right)}}}{\sqrt{\frac{1}{\Lambda^2} + \frac{r^2}{2c_s^2}}} \qquad (6.0\text{-}22)$$

Hence, a circular ridge of correlation propagates outward with increasing τ.

The combination of this circular symmetry with the cartesean geometry implicit with (6.0-4) presents many difficulties.

A less ambitious approach to (6.0-4) would assume ζ to be a random corrugation rather than a general irregular surface. In that case the correlation function could be written as a sum of a "left" going and a "right" going component just as one treats the propagation of waves on a string. Further simplifications arise if the corrugation is

aligned with either the x or y axis for then one may perform 2 of the space integrals with great ease.

For example, if

$$\Psi_\zeta(\xi,\eta,o) = \Psi_\zeta(\xi) \tag{6.0-23}$$

then (6.0-4) becomes

$$\Phi(\omega,\omega';\tau) = \frac{2\omega \, \mathrm{Sin}^2\psi}{c} \left(\frac{1}{r_o r'_o}\right) \frac{1}{(r_o + r'_o)}$$

$$\exp\left\{ -\tfrac{1}{2}\left(\frac{\sigma_\zeta \, \mathrm{Sin}\psi}{c}\right)^2 \left[\omega^2 + \omega'^2\right] - j\tau_s(\omega-\omega') \right\}$$

$$\int\int_{-\infty}^{\infty} e^{-j\left[x_1^2 - x_2^2\right](c_s^o)^2}$$

(6.0-24)

$$\exp\left\{ \frac{\omega\omega'}{c} \left[\tfrac{1}{2}\Psi_\zeta\left(\left[x_1 - x_2\right] + c\tau\right) \right.\right.$$

$$\left.\left. + \tfrac{1}{2}\Psi_\zeta\left(\left[x_1 - x_2\right] - c\tau\right) \right] \right\} dx_1 \, dx_2$$

We will not pursue this integral further since an adequate treatment would be somewhat lengthy.

Instead, we confine our attention for the remainder of this section to the fourth order moment:

$$\Phi_{1,2}(\omega,\omega',\omega'',\omega''',\mu,\mu',\mu'') = \left(\frac{1}{r_o r'_o}\right)^4$$

$$\frac{16\,\omega\,\omega'\,\omega''\,\omega'''\,\mathrm{Sin}^2\psi_1\,\mathrm{Sin}^2\psi_2}{c^4} \quad e^{-j\tau_{s1}(\omega-\omega') - j\tau_{s2}(\omega''-\omega''')}$$

$$\int_{-\infty}^{\infty}\int\int\int\int\int\int\int\int \; dx_1\, dx_2\, dx_3\, dx_4 \; dy_1\, dy_2\, dy_3\, dy_4$$

$$Be(\theta_1,\phi_1)\; Be^*(\theta_2,\phi_2)\; Be(\theta_3,\phi_3)\; Be^*(\theta_4,\phi_4)$$

$$\exp\{-j\left[\omega x_1^2 - \omega'(x_2+\Delta x)^2\right] \mathrm{Sin}^2\psi_1\}$$

$$\exp\{-j\left[\omega y_1^2 - \omega'(y_2+\Delta y)^2\right]\}$$

$$\exp\{-j\left[\omega'' x_3^2 - \omega''' (x_4+\Delta x)^2\right] \mathrm{Sin}^2\psi_2\}$$

$$\exp\{-j\left[\omega'' y_3^2 - \omega''' (y_4+\Delta x)^2\right]\}$$

$$\exp\{\tfrac{-j}{c}\left[\omega\zeta_1(x_1,y_1,t) - \omega'\zeta_1(x_2+\Delta x, y_2+\Delta y, t-\mu)\right] \mathrm{Sin}\psi_1\}$$

$$\exp\{\tfrac{-j}{c}\left[\omega''\zeta_2(x_3,y_3,t-\mu')\right.$$

$$\left.-\omega'''\zeta_2(x_4+\Delta x, y_4+\Delta y, t-\mu'')\right] \mathrm{Sin}\psi_2 \,.$$

(6.0-25)

where Δx and Δy are the x and y displacements of the specular points for the two receivers. We assume that the grazing angles ψ_1 and ψ_2 and distances Re_1 and Re_2 "seen" by the two receivers may be slightly different.

This 8-fold integral contains the 4-th order characteristic function for ζ_1 and ζ_2. This in turn is a function of the 12 differences in coordinates

$x_1 - x_2$	$y_1 - y_2$
$x_1 - x_3$	$y_1 - y_3$
$x_1 - x_4$	$y_1 - y_4$
$x_2 - x_3$	$y_2 - y_3$
$x_2 - x_4$	$y_2 - y_4$
$x_3 - x_4$	$y_3 - y_4$

Clearly such an integral poses a formidable challenge. Yet in order to be of any real use this integral must be performed for a wide range of frequencies. Furthermore for the applications discussed in chapter 3 this integral must itself be integrable as a function of frequency.

The 8-fold space integral of (6.0-25) may be reduced to a 4-fold integral by restricting attention to corrugations but we will not write this result here.

We have here barely written down the relevant integrals for the fully random model. While at the time of this writing some progress has been made by the author with equation (6.0-4), nothing has yet appeared in the literature approaching a solution of (6.0-25). Both of these integrals are presented only as suggestions for future research or motivations for experimental studies.

6.1 Other Suggestions for Future Research

In closing it would be worthwhile to indicate a few other areas of research which have only been treated briefly in this work. These are listed in ascending order of complexity:

1) Analysis of large arrays. Some relevant material on this problem is presented in Appendix J.

2) Derivation and computation of cumulants of higher order than the fourth. This research would be useful for investigations of optimal receiver design. Since the received signal is non-gaussian, the optimal receiver would presumably be non-quadratic.

3) Narrow band detection. While the applications studied in this work treat low pass types of spectra, the results of chapter 3 are general enough to be used in a study of the effect of the frequency smear due to the time variation of the scattering.

4) Multiple surface reflection. In this case some of the discussions in chapter 2 with respect to cascaded RLTVF's is useful for the narrow band case.

5) Analysis of scattering at low grazing angles. In this case the active scattering region is not localized to the vicinity of the specular point[75,76]. Also, shadowing and multiple scattering become important unless the surface is not very rough.

Of these topics the second one concerning optimal receiver design is probably the most interesting. Presumably such a detector would be self adaptive with respect to the slow axis time variations of the scattering, perhaps identifying some of the coherent properties of the scattering model. It should be possible to show that that the information lost by constraining the correlator detector to be steered "on-target" * would be captured by the optimal detector.

---

*As discussed in Appendix D-

APPENDIX A

MORGAN'S DERIVATION OF EQUATION (1.3-14)

The transition from Equation (1.3-12) to Equation (1.3-14) is facilitated by executing the partial derivative of q with respect to n and rewriting the partial of q with respect to t in terms of the total time derivative. From the definition of the auxiliary wave q in Equation (1.3-9) we have

$$\frac{1}{c}\frac{\partial q}{\partial t} = \frac{W'(\xi)}{r}\Big|_{\xi=ct-ct_o+r} \qquad \frac{\partial q}{\partial r} = -\frac{q}{r} + \frac{1}{c}\frac{\partial q}{\partial t} \tag{A.1-1}$$

where $W'(\xi)$ is the derivative of the narrow pulse function $W(\xi)$ with respect to its argument, and t and r are regarded as independent variables. Next we have

$$\frac{\partial q}{\partial n} = \frac{\partial q}{\partial r}\frac{\partial r}{\partial n} = \left(-\frac{q}{r} + \frac{1}{c}\frac{\partial q}{\partial t}\right)\frac{\partial r}{\partial n} \quad . \tag{A.1-2}$$

and by straightforward application of the chain rule

$$\frac{1}{c}\frac{dq}{dt} = \frac{1}{c}\frac{\partial q}{\partial t} + \frac{\partial q}{\partial r}\frac{1}{c}\frac{dr}{dt} = \frac{1}{c}\frac{\partial q}{\partial t} + \left(-\frac{q}{r} + \frac{1}{c}\frac{\partial q}{\partial t}\right)\frac{1}{c}\dot{r} = \tag{A.1-3}$$

$$(1+\dot{r}/c)\,\frac{1}{c}\frac{\partial q}{\partial t} - \frac{q\dot{r}}{cr}$$

Solving this for $\partial q/\partial t$ and also substituting into Equation (A.1-2)

$$\text{(a)} \quad \frac{1}{c}\frac{\partial q}{\partial t} = \frac{\left(\frac{1}{c}\frac{dq}{dt} + \frac{q\dot{r}}{cr}\right)}{(1+\dot{r}/c)} \qquad \text{(b)} \quad \frac{\partial q}{\partial n} = \frac{\left(\frac{1}{c}\frac{dq}{dt} - \frac{q}{r}\right)\frac{\partial r}{\partial n}}{(1+\dot{r}/c)} \tag{A.1-4}$$

Substituting Equation (A.1-4a,b) into Equation (1.3-12) we have directly

$$p_s(P,t_o) = \iint_{R_{xy}} \frac{dx\,dy}{4\pi} \Bigg\{ \int_{-\infty}^{\infty} \Bigg[q \left(\frac{\partial p_s}{\partial n} + \frac{v_n}{c^2} \frac{\partial p_s}{\partial t} + \frac{\left(\frac{\partial r}{\partial n} - \frac{\dot{r} v_n}{c^2}\right)}{(1+\dot{r}/c)} \frac{p_s}{r} \right) \tag{A.1-5}$$

$$- \frac{1}{c} \frac{dq}{dt} \left\{ \frac{\left(\frac{\partial r}{\partial n} + \frac{v_n}{c}\right)}{(1+\dot{r}/c)} \right\} p_s \Bigg] H\, dt \Bigg\}$$

$$z=\zeta(x,y,y)$$

where H is defined in Equation (1.3-13). Equation (A.1-5) is the sum of two terms one of which is multiplied by q and the other of which is multiplied by dq/dt. Now in the limit of very narrow $W(\xi)$

$$q = \frac{W(\xi)}{r} + \frac{1}{r(1+\dot{r}/c)} \delta\left(t - \left[t_o - \frac{r}{c}\right]\right) \tag{A.1-6}$$

Using this the first term of Equation (A.1-5) is easily reduced to the first three terms of Equation (1.3-14). The second term in Equation (A.1-5) is similarly reducible after integration by parts and utilizing the fact that q is zero for $t = \pm\infty$.

APPENDIX B

THE PLANE WAVE EXPANSION FOR SURFACE SCATTERING

Meecham's derivation of the plane wave expansion for surface scattering is readily extended to surfaces with two-dimensional and even slowly varying irregularities. The analysis proceeds from an expansion given by Levine and Schwinger[40] for the free-space point source Green's function $\exp(jkr)/r$ which occurs in Equation (1.4-3)

$$\frac{e^{jkr}}{r} = \frac{1}{2\pi^2}\iiint \frac{e^{j[k_x(x_s-x_p)+k_y(y_s-y_p)+k_z(z_s-z_p)]}}{(k_x^2+k_y^2+k_z^2-k^2)}\,dk_x\,dk_y\,dk_z$$

$$= \frac{1}{2\pi}\iint \frac{e^{j[k_x(x_s-x_p)+k_y(y_s-y_p)\pm\sqrt{k^2-k_x^2-k_y^2}\,(z_s-z_p)]}}{\sqrt{k^2-k_x^2-k_y^2}}\,dk_x\,dk_y \qquad \text{(B.1-1)}$$

where

$$r = \sqrt{(x_s-x_p)^2+(y_s-y_p)^2+(z_s-z_p)^2} \qquad \text{(B.1-2)}$$

and where in the context of Equation (1.4-3) $P(x_p,y_p,z_p)$ is the observation point and $S(x_s,y_s,z_s)$ is a point on the surface.

The sign in Equation (B.1-1) is assigned as follows. Note that free-space Green's function represents an outgoing wave emanating <u>from</u> P . This wave must be "upgoing" when $z_s > z_p$ and "downgoing" when $z_s < z_p$. The result of the second equality in Equation (B.1-1) is in the form of a plane wave expansion for this outgoing wave provided the "+" sign is used when $z_s > z_p$ and the "-" sign is used when $z_s < z_p$. This position dependence of the space-spectral density of the plane wave

expansion for the point source Green's function leads to a similar dependence in plane wave expansions for surface scattered waves.

Substituting Equation (B.1-1) into (1.4-3) we immediately have the plane wave expansion for the scattered signal received at P

$$\hat{p}_s(P,t_o) = \iint A_+(k_x,k_y,z_p,t_o)\, e^{-j[k_x x_p + k_y y_p + \sqrt{k^2 - k_x^2 - k_y^2}\, z_p]}\, dk_x\, dk_y$$
$$+ \iint A_-(k_x,k_y,z_p,t_o)\, e^{-j[k_x x_p + k_y y_p - \sqrt{k^2 - k_x^2 - k_y^2}\, z_p]}\, dk_x\, dk_y \qquad (B.1\text{-}3)$$

in which the (unknown) wave density is given by

$$A_{\pm}(k_x,k_y,z_p,t_o) = \frac{1}{4\pi} \iint_{S_r^{\pm}(t_o,z_p)} \left(\hat{p}_i \frac{\partial G_{\pm}}{\partial n} + G_{\pm} \frac{\partial}{\partial n} [\hat{p}_s(x,y,z,t_o - r/c)]\right) dS \qquad (B.1\text{-}4)$$

Here $S_r^{\pm}(t_o,z_p)$ is that portion of the appropriately retarded surface $S_r(t_o)$ which lies above z_p (for the "+" sign) or below z_p (for the "-" sign) and

$$G_{\pm}(x_s,y_s,z_s) = \frac{e^{j[k_x x_s + k_y y_s \pm \sqrt{k^2 - k_x^2 - k_y^2}\, z_s]}}{\sqrt{k^2 - k_x^2 - k_y^2}} \left(\frac{i}{2\pi}\right) \qquad (B.1\text{-}5)$$

Equation (B.1-3) shows that when portions of the surface lie above and below the level z_p of the receiver the density of upgoing and downgoing waves becomes position dependent. However, when $z_p < z_s$ for all z_s we have only downgoing waves and the density of these plane

waves becomes position invariant (see Figure B.1-1):

$$A_-(k_x,k_y,z_p,t_o) = 0$$

$$A_+(k_x,k_y,z_p,t_o) = A(k_x,k_y,t_o) \qquad \text{(B.1-6)}$$

This argument which is, of course, only approximate in the case of the time varying surface is seen to be exact for the fixed surface case.



Situations for (a) Position Dependent

and (b) Position Invariant

Plane Wave Expansions

APPENDIX C

SPECULAR POINT EXPANSIONS FOR r AND r'

In order to simplify some of the mathematics necessary for obtaining the expansions for r and r' about the specular point we follow the development used by Gulin.[38] We begin by defining $P(x_p,0,z_p)$ to be the observation point, $Q(x_q,0,z_q)$ to be the source location and $S(x,y,\zeta)$ to be a point on the surface. The origin of coordinates is taken to be at the specular point (see Figure 1.4-1). Let x_1 be the x-axis component of the distance between the source point Q and the surface point S. If L is the x-axis separation between P and Q then

$$x_1 = x - x_q = x + \left(\frac{z_q}{z_q + z_p}\right) L \qquad \text{(C.1-1)}$$

First we expand r and r' to second order in ζ to obtain

$$r = \sqrt{x_1^2 + y^2 + (z_p - \zeta)^2} = \qquad \text{(C.1-2)}$$

$$R - \frac{z_p \zeta}{R} + [1 - (z_p/R)^2] \frac{\zeta^2}{2R} + \ldots$$

$$r' = \sqrt{(L-x_1)^2 + y^2 + (z_q - \zeta)^2} = \qquad \text{(C.1-3)}$$

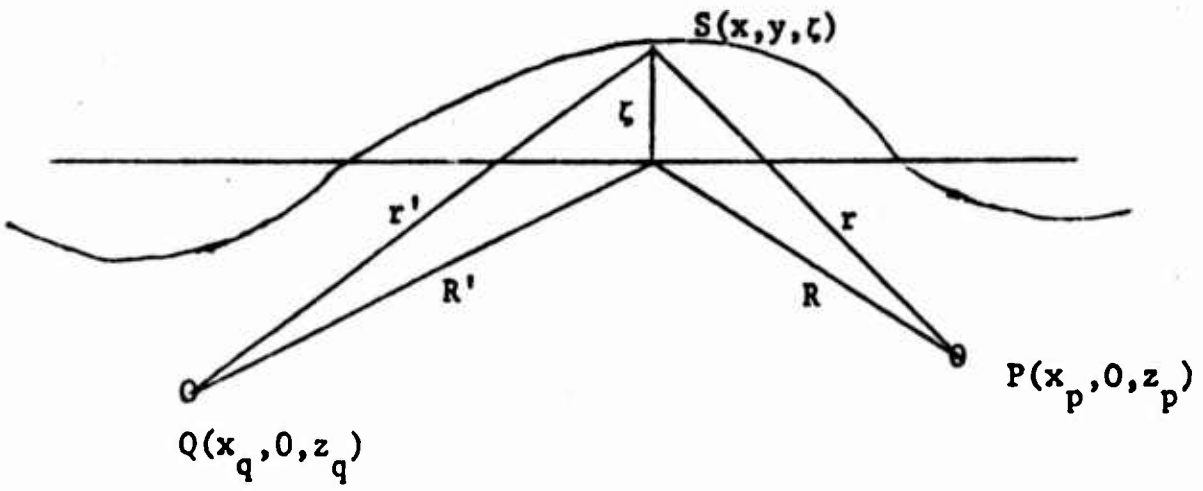
$$R' - \frac{z_q \zeta}{R'} + [1 - (z_q/R')^2] \frac{\zeta^2}{2R'} + \ldots$$

where

$$\text{(a)} \quad R = [x_1^2 + y^2 + z_p^2]^{\frac{1}{2}} \; ; \quad \text{(b)} \quad R' = [(L-x_1)^2 + y^2 + z_q^2]^{\frac{1}{2}} \qquad \text{(C.1-4)}$$

Here R and R' are respectively the distances from the projection of

the surface point S onto the plane $\zeta = 0$ to the points P and Q. (see Figure C.1-1).



The Definitions of R and R'

Figure C.1-1

We next expand R and R' in powers of the displacements x and y from the specular point. First note that if ψ_i is the grazing angle of incident radiation at the specular point then

$$\mathrm{Ctn}(\psi_i) = \frac{L}{z_p + z_q} \tag{C.1-5}$$

Thus, (C.1-4a) becomes

$$R = [(x + z_p\ \mathrm{Ctn}(\psi_i))^2 + y^2 + z_p^2]^{1/2} =$$

$$[z_p^2\ (1 + \mathrm{Ctn}^2(\psi_i))]^{1/2} \cdot [1 + \tfrac{1}{2} \frac{x^2 + 2z_p x\ \mathrm{Ctn}(\psi_i) + y^2}{z_p^2\ [1 + \mathrm{Ctn}^2(\psi_i)]} \tag{C.1-6}$$

$$- \frac{1}{8z_p^4} \frac{(2\ z_p x\ \mathrm{Ctn}(\psi_i))^2}{[1 + \mathrm{Ctn}^2(\psi_i)]^2} + \ldots]$$

Next, defining r_o to be the distance from P to the specular point we have

$$r_o = z_p [1 + Ctn^2(\psi_i)]^{\frac{1}{2}}$$
$$= z_p \, Csc(\psi_i) \qquad \text{(C.1-7)}$$

Substituting this into Equation (C.1-6)

$$R = r_o + \frac{y^2}{r_o} + x \, Cos(\psi_i) + \frac{x^2 \, Sin^2(\psi_i)}{2r_o} + o[\frac{(x^2+y^2)z_{px}}{r_o^4}] \qquad \text{(C.1-8)}$$

In a similar manner from Equation (C.1-4b)

$$R' = [(x - z_q \, Ctn(\psi_i))^2 + y^2 + z_q^2]^{\frac{1}{2}} =$$

$$[z_q^2 \, (1 + Ctn^2(\psi_i))]^{\frac{1}{2}} \cdot [1 + \frac{x^2 - 2z_q x \, Ctn(\psi_i) + y^2}{2z_q^2 \, [1 + Ctn^2(\psi_i)]} \qquad \text{(C.1-9)}$$

$$- \frac{1}{8z_q^4} \frac{(2 \, z_q x \, Ctn(\psi_i))^2}{[1 + Ctn^2(\psi_i)]^2} + \ldots]$$

Finally, defining the distance from Q to the specular point to be r'_o we have

$$r'_o = z_q \, Csc(\psi_i) \qquad \text{(C.1-10)}$$

and

$$R = r'_o + \frac{y^2}{2r'_o} - x \, Cos(\psi_i) + \frac{x^2 \, Sin^2(\psi_i)}{2r'_o} + o[\frac{(x^2+y^2)z_q x}{(r'_o)^4}] \qquad \text{(C.1-11)}$$

Next, examining the reciprocal of r we find

$$\frac{1}{r} = \frac{1}{R}\left[1 + \frac{z_p \zeta}{R^2} + \ldots\right]$$

$$= \frac{1}{r_o}\left[1 - \tfrac{1}{2}\frac{y^2}{r_o^2} + \frac{x\,\mathrm{Cos}(\psi_i)}{r_o} - \frac{x^2\,\mathrm{Sin}^2(\psi_i)}{2\,r_o^2} + \ldots\right] \cdot$$

$$\left[1 + \frac{z_p \zeta}{r_o^2} + \ldots\right] \qquad \text{(C.1-12)}$$

$$= \frac{1}{r_o}\left[1 - \tfrac{1}{2}\frac{y^2}{r_o^2} + \frac{x\,\mathrm{Cos}(\psi_i)}{r_o} - \frac{x^2\,\mathrm{Sin}^2(\psi_i)}{2\,r_o^2} + \frac{z_p \zeta}{r_o^2} + \ldots\right]$$

Similarly,

$$\frac{1}{r'} = \frac{1}{r'_o}\left[1 - \tfrac{1}{2}\frac{y^2}{r'^2_o} - \frac{x\,\mathrm{Cos}(\psi_i)}{r'_o} - \frac{x^2\,\mathrm{Sin}^2(\psi_i)}{2\,r'^2_o} + \frac{z_p \zeta}{r'^2_o} + \ldots\right] \qquad \text{(C.1-13)}$$

Multiplying these two reciprocals together we obtain

$$\frac{1}{r\,r'} = \frac{1}{r_o r'_o}\left[1 + x\,\mathrm{Cos}(\psi_i)\left(\frac{1}{r_o} - \frac{1}{r'_o}\right)\right.$$

$$\left. - \left(\frac{1}{r_o^2} + \frac{1}{r'^2_o}\right)\left[y^2 + x^2(\mathrm{Cos}^2(\psi_i) + \mathrm{Sin}^2(\psi_i)) - z_p\zeta\right] + \ldots\right] \qquad \text{(C.1-14)}$$

where terms up to second order in x and y and up to first order in ζ have been retained.

The sum of r and r' can be found to second order in a similar manner:

$$r + r' = r_o + r'_o + \frac{[y^2 + x^2\,\mathrm{Sin}^2(\psi_i)]}{R_e} - 2\,\zeta\,\mathrm{Sin}(\psi_i)$$

$$+ \zeta^2\left[\left[1 - (z_p/r_o)^2\right]\frac{1}{2r_o} + \left[1 - (z_q/r'_o)^2\right]\frac{1}{2r'_o}\right] + \ldots \qquad \text{(C.1-15)}$$

The second order term in ζ in Equation (C.1-15) will not be of importance in forming $jk(r+r')$ in the exponent of Equation (1.4-10) provided

$$2\pi \underline{\mathrm{Cos}}^2(\psi_i)\,\frac{h^2}{\lambda R_e} \ll 1 \qquad \text{(C.1-16)}$$

where $h = \mathrm{Max}(|\zeta(x,y,t)|)$ and $\lambda = 2\pi/k$.

Next, we evaluate the various partial derivatives of $r+r'$

$$\frac{\partial}{\partial x}[r+r'] = 2\,\mathrm{Sin}^2(\psi_i)\,\left(\frac{x}{R_e}\right) = \left(\frac{x}{r_o} + \frac{x}{r'_o}\right)\mathrm{Sin}^2(\psi_i)$$

$$\simeq -\tfrac{1}{4}\,a_s(x,y)\,(c_s^o)^2 \qquad \text{(C.1-17)}$$

$$\frac{\partial}{\partial y}[r+r'] = 2\,\left(\frac{x}{R_e}\right) = \left(\frac{y}{r_o} + \frac{y}{r'_o}\right) \simeq [b(x,y) + b'(x,y)]\,(-1)$$

$$= -\,b_s(x,y) \qquad \text{(C.1-18)}$$

$$\frac{\partial}{\partial z}[r+r'] = 2\,\mathrm{Sin}(\psi_i) = c_s^o \qquad \text{(C.1-19)}$$

where $a(x,y)$, $b(x,y)$, $c(x,y)$ are the x,y,z direction cosines for r , and $a'(x,y)$, $b'(x,y)$, $c'(x,y)$ are x,y,z direction cosines for r' . The subscript "s" on $a,b,$ or c denotes the sum of primed and unprimed quantities while the superscript "o" denotes the evaluation of these quantities at the specular point. Here we have used the following:

$$\frac{x_p - x}{r_o} \simeq a(x,y),\quad \frac{x_p}{r_o} \simeq a^o\ ;\quad \frac{x_q - x}{r'_o} \simeq a'(x,y),\quad \frac{x_q}{r'_o} \simeq a'^o$$

$$\frac{y}{r_o} \simeq -\,b(x,y)\quad ,\quad \frac{y}{r'_o} \simeq -\,b'(x,y) \qquad \text{(C.1-20)}$$

Using Equations (C.1-14), (1.4-7) and (C.1-17) through (C.1-19) we

obtain the following

$$\frac{H}{r\,r'}\frac{\partial}{\partial n}[r+r'] \approx \frac{1}{r_o\,r'_o}[1 + [a^o - a'^o - a(x,y) + a'(x,y)]\,a^o + \ldots].$$

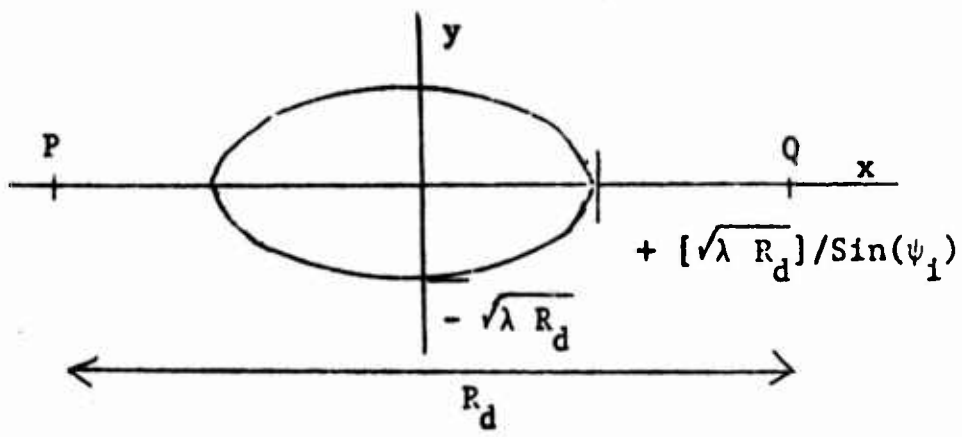
$$[+\tfrac{1}{4}\,a_s(x,y)\,(c_s^o)^2\,(\frac{\partial\zeta}{\partial x}) + b_s(x,y)\,(\frac{\partial\zeta}{\partial y}) + c_s^o] \approx \tag{C.1-21}$$

$$\frac{1}{r_o\,r'_o}[\tfrac{1}{4}\,a_s(x,y)\,(c_s^o)^2\,(\frac{\partial\zeta}{\partial x}) + b_s(x,y)\,(\frac{\partial\zeta}{\partial y}) + c_s^o$$

$$+ [a^o - a'^o - a(x,y) + a'(x,y)]\,a^o\,c_s^o]$$

The result may be used to obtain Equation (1.4-11).

It is of interest here to investigate the magnitude of the terms x/r, x/r', y/r and y/r'. In the vicinity of the edge of the active scattering area which is roughly the first Fresnel zone (see Figure C.1-2) these terms take on their maximum values. If the points P and Q are located at the same depth $r \approx r' \approx R_d/\mathrm{Cos}(\psi_i)$ where R_d is the direct or line-of-sight distance between P and Q. For moderate grazing angles the value of x at the semi-major axis edge of the first Fresnel zone is approximately given by $[\sqrt{\lambda\,R_d}]/\mathrm{Sin}(\psi_i)$.[41] The terms of interest are then of the order of magnitude $[\sqrt{\lambda/R_d}]\,\mathrm{Ctn}(\psi_i)$ and are negligible if Equation (1.4-9) holds. In a similar manner the correction terms in Equations (C.1-8) and (C.1-11) are of order $[\sqrt{\lambda/R_d}]^5\,z_p\,\mathrm{Cos}^4(\psi_i)/[\lambda\,\mathrm{Sin}^3(\psi_i)]$ and can be ignored to the same extent.



The First Fresnel Zone

Figure C.1-2

APPENDIX D

REFINEMENTS IN THE CRITERIA FOR DETECTABILITY

D.1 Peak Displacement Corrections for the Combined Correlator Detector-Tracker

In section 3.8 we examined the problem of detecting a target when the replication delay parameter τ is "steered" to the (known) "on target" condition ($\tau = \tau_o$). In most practical applications, however, the parameter τ_o is not known a-priori. For this reason many standard detectors employ a kind of composite hypothesis test or detector-tracker scheme which emulates the mode of operation of optimum maximum likelihood processors.[69] In this implementation we envisage a large number of parallel correlator detectors each with a different, fixed value of τ covering the range of possible values for τ_o. The processor outputs are scanned simultaneously over τ for a peak of correlation.

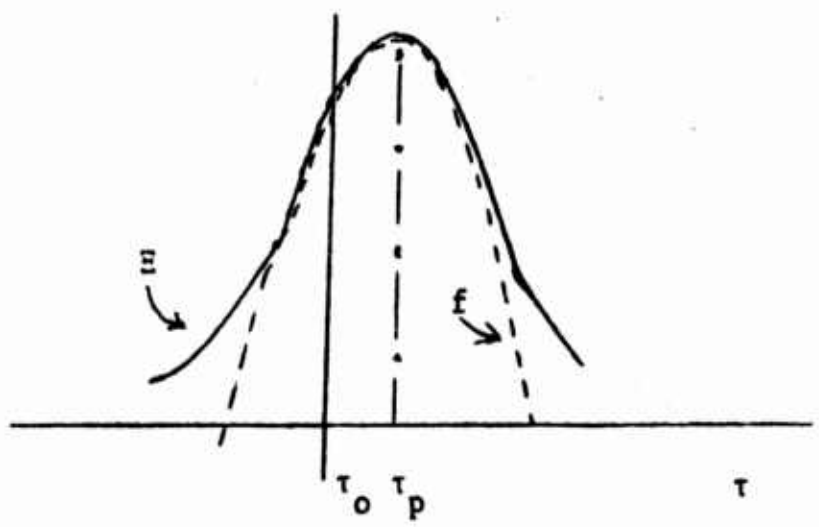
As noted in section 3.6 the peak of correlation is generally not located at the correct value τ_o. This is true even in the absence of scattering, but the presence of a certain amount of delay modulation in most scattering models aggravates this problem. While the effect of scatter delay modulation on tracking errors may be great, the impact on signal detectability need not be as significant. Hence, the "on target" normalized variances

$$d_0^2(\tau_o,T,\mu) = \frac{\Delta(\tau_o,\tau_o,T,\mu)\,|H_o}{[\overline{\Xi(\tau_o,T,\rho)}]^2} \quad ; \quad d_1^2(\tau_o,T,\rho) = \frac{\Delta(\tau_o,\tau_o,T,\mu)\,|H_1}{[\overline{\Xi(\tau_o,T,\rho)}]^2} \qquad (D.1\text{-}1)$$

are usually overly pessimistic indicators of detector performance.

It is desirable to compensate for errors that arise in the evaluation of the correlator detector due to variation in the location of the peak. In this section we examine a set of correction terms that approximately perform this compensation and that are valid in the same limit as the results of section 3.6. That is, we assume that the processing interval T is large enough so that the maximum deviation between the location of the peak, τ_p, in any realization $\Xi(\tau,T,\rho)$ and the mean location τ_o is much less than the signal correlation width. We also assume sufficient signal bandwidth limitation so that the second derivative of Ξ with respect to τ exists. Furthermore, we assume that Ξ'' is essentially fixed in the vicinity of location of the peak at its mean value, i.e. $\Xi''(\tau,T,\rho) \approx \Xi''(\tau_o,T,\rho)$. This in turn implies that the null in $\Xi(\tau,T,\rho)$ at τ_p is well separated from any neighboring zeros.

The correction technique consists essentially in fitting a parabola to the correlation function estimate $\Xi(\tau,T,\rho)$ at $\tau = \tau_o$. The location of the maximum for this parabola is an approximant to the location of the peak τ_p (see Figure D.1-1).



The Quadratic Approximant to Ξ

Figure D.1-1

The equation for the quadratic approximant is

$$f(\tau) = a(\tau-\tau_o)^2 + b(\tau-\tau_o) + c \ . \qquad \text{(D.1-2)}$$

$$\therefore f'(\tau) = 2a(\tau-\tau_o) + b \qquad \text{(D.1-3)}$$

$$f''(\tau) = 2a \ . \qquad \text{(D.1-4)}$$

Fitting this to $\Xi(\tau,T,\rho)$ at $\tau = \tau_o$:

$$a = \tfrac{1}{2}\Xi''(\tau_o,T,\rho) \qquad \text{(D.1-5)}$$

$$b = \Xi'(\tau_o,T,\rho) \qquad \text{(D.1-6)}$$

$$c = \Xi(\tau_o,T,\rho) \qquad \text{(D.1-7)}$$

The approximant has its peak at the location

$$\tau_p = \tau_o - \frac{\Xi'(\tau_o,T,\rho)}{\Xi''(\tau_o,T,\rho)} \qquad \text{(D.1-8)}$$

To the extent that $\Xi''(\tau_o,T,\rho)$ is approximately equal to its mean value this result is equivalent to the discussion of section 3.6.

The value of $f(\tau)$ at the peak is

$$f(\tau_p) = -\frac{[\Xi'(\tau_o,T,\rho)]^2}{2\Xi''(\tau_o,T,\rho)} + \Xi(\tau_o,T,\rho) \qquad \text{(D.1-9)}$$

$$\approx -\frac{[\Xi'(\tau_o,T,\rho)]^2}{2\Xi''(\tau_o,T,\rho)} + \Xi(\tau_o,T,\rho)$$

It should be noted that in order for this analysis to be meaningful the curvature of $\Xi(\tau,T,\rho)$ should be toward the τ-axis at τ_o. Hence, from the following approximate result for the mean height of the peak

$$\overline{f(\tau_p)} \approx \overline{\Xi(\tau_o,T,\rho)} - \frac{\overline{[\Xi'(\tau_o,T,\rho)]^2}}{2\,\overline{\Xi''(\tau_o,T,\rho)}} \qquad \text{(D.1-10)}$$

we have approximately

$$|\overline{f(\tau_p)}| \geq |\overline{\Xi(\tau_o,T,\rho)}| \qquad \text{(D.1-11)}$$

For this analysis to be accurate, however, the correction term should be small compared with $\overline{\Xi(\tau_o,T,\rho)}$.

Similarly, the correction to the variance becomes

$$\mathrm{Var}[f(\tau_p)] =$$

$$\overline{\{[\Xi(\tau_o,T,\rho) - \overline{\Xi(\tau_o,T,\rho)}] - (\frac{[\Xi'(\tau_o,T,\rho)]^2}{2\,\Xi''(\tau_o,T,\rho)} - [\frac{\overline{[\Xi'(\tau_o,T,\rho)]^2}}{2\,\Xi''(\tau_o,T,\rho)}])\}^2}$$

$$\approx \mathrm{Var}[\Xi(\tau_o,T,\rho)] + \frac{\mathrm{Var}\{[\Xi'(\tau_o,T,\rho)]^2\}}{4\,\overline{[\Xi''(\tau_o,T,\rho)]^2}} \qquad \text{(D.1-12)}$$

$$- \overline{[\Xi(\tau_o,T,\rho) - \overline{\Xi(\tau_o,T,\rho)}]\,[\Xi'(\tau_o,T,\rho)]^2}\;\;(\frac{1}{\overline{\Xi''(\tau_o,T,\rho)}})$$

The assumption that the correction to the mean in Equation (D.1-10) is small does not guarantee that the corrections to the variance are small. Indeed, when slowly varying delay modulation dominates any amplitude modulation present in the scattering model, these correction terms become very significant.

A closer examination of Equation (D.1-12) reveals that in order to relate all of the correction terms for the variance of the detector

output to the spectra of the input signals and appropriate system functions one must evaluate many integrals. The primary origin of this complexity is the requirement for the evaluation of the eighth order moment of $\mathrm{Var}\{[\Xi'(\tau_o,T,\rho)]^2\}$. Using the same technique applied in connection with the derivation of Equation (2.1-18) one can show that (2p)-th order moments of the increments $dz_x(\omega)$ give rise under the Gaussian hypothesis to

$$\frac{(2p)!}{2^p \times p!} \qquad \text{(D.1-13)}$$

distinct integrals. The presence of additive noise compounds this difficulty. For the case of $p = 4$ there are a minimum of 105 different integrals!

However, for $p = 3$ there are only 15 integrals for the signal-only terms. Thus, evaluation of the last correction term in Equation (D.1-12) is at least partially tractable. When the scattering fluctuations produce much more amplitude modulation than delay modulation Equation (D.1-12) may be reduced approximately to

$$\mathrm{Var}[f(\tau_p)] \approx \mathrm{Var}[\Xi(\tau_o,T,\rho)] \qquad \text{(D.1-14)}$$

$$- \frac{\overline{\Xi(\tau_o,T,\rho)\ [\Xi'(\tau_o,T,\rho)]^2}}{\overline{\Xi''(\tau_o,T,\rho)}} + \frac{\overline{\Xi(\tau_o,T,\rho)}\ \overline{[\Xi'(\tau_o,T,\rho)]^2}}{\overline{\Xi''(\tau_o,T,\rho)}}$$

The quantities $\overline{\Xi(\tau_o,T,\rho)}$, $\overline{\Xi''(\tau_o,T,\rho)}$ and $\overline{[\Xi'(\tau_o,T,\rho)]^2}$ can be evaluated easily from integrals obtained elsewhere in this report, but the third order moment $\overline{\Xi(\tau_o,T,\rho)\ [\Xi'(\tau_o,T,\rho)]}$ requires a separate development.

APPENDIX E

VARIANCE INTEGRALS FOR RANDOM AMPLITUDE AND DELAY MODEL

E.1 Multipath Processor

The integrals in (3.3-9) through (3.3-13) are evaluated here for the spectra (4.2-2), the filters (4.2-3), and the propagation model described by (4.1-11), (4.1-12), and (4.2-4):

$$I_1(\nu) = \int_\omega e^{j\omega\nu} \frac{\sqrt{2\pi}\, k_1^2\, P_{n1}}{\Omega_{n1}} \exp\{-\frac{\omega^2}{2}\left[\frac{1}{\Omega_1^2} + \frac{1}{\Omega_{n1}^2}\right]\} \frac{d\omega}{2\pi}$$

$$= k_1^2\, P_{n1} \left| \frac{\Omega_{1n1}}{\Omega_{1n}} \right| \exp\{-1/2\nu^2\Omega_{1n1}^2\} \qquad \text{(E.1-1)}$$

where

$$\frac{1}{\Omega_{1n1}} = \sqrt{\frac{1}{\Omega_1^2} + \frac{1}{\Omega_{n1}^2}} \qquad \text{(E.1-2)}$$

$$I_2(\nu) = \int_\omega e^{j\omega\nu} \frac{\sqrt{2\pi}\, k_1^2\, p_x\, R_{AA}(\nu)}{\Omega_x} \exp\left\{\frac{-\omega^2}{2}\left[\frac{1}{\Omega_x^2} + \frac{1}{\Omega_x^2} + 2[\sigma_\tau^2 - R_{\tau\tau}(\nu)]\right]\right\} \frac{d\omega}{2\pi}$$

$$= \frac{R_{AA}(\nu)\, k_1^2\, p_x}{\Omega_x \sqrt{\frac{1}{\Omega_{x1}^2} + 2[\sigma_\tau^2 - R_{\tau\tau}(\nu)]}} \qquad \text{(E.1-3)}$$

$$\exp\left\{\frac{-\nu^2}{2}\left[\frac{1}{\Omega_{1x}^2} + 2[\sigma_\tau^2 - R_{\tau\tau}(\nu)]\right]^{-1}\right\}$$

where

$$\frac{1}{\Omega_{1x}} = \sqrt{\frac{1}{\Omega_1^2} + \frac{1}{\Omega_x^2}} \tag{E.1-4}$$

Similarly,

$$I_3(\nu) = k_2^2\left[P_{n2}\left(\frac{\Omega_{2n2}}{\Omega_{n2}}\right)\exp\{-\tfrac{1}{2}\nu^2\Omega_{2n2}^2\} + A_d^2\, P_x\left(\frac{\Omega_{1x}}{\Omega_x}\right)\exp\{-\tfrac{1}{2}\nu^2\Omega_{1x}^2\}\right] \tag{E.1-5}$$

with

$$\frac{1}{\Omega_{2n2}} = \sqrt{\frac{1}{\Omega_2^2} + \frac{1}{\Omega_{n2}^2}} \tag{E.1-6}$$

Also, applying (4.1-12) and (3.3-12)

$$I_4(\tau,\tau',\nu) = \int_{\omega}\int_{\omega'} e^{j\omega(\tau-\tau_s+\tau d)-j\omega'(\tau'-\tau_s+\tau_d)}\,\frac{2\pi\, k_1^2\, k_2^2\, A_d^2\, P_x^2}{\Omega_x^2}$$

$$R_{AA}(\nu)\,\exp\left\{-\frac{1}{2}\left[(\omega^2+\omega'^2)\alpha_m^2 - 2\omega\omega'\, R_{\tau\tau}(\nu)\right]\right\}\frac{d\omega}{2\pi}\,\frac{d\omega'}{2\pi}$$

$$= \frac{k_1^2\, k_2^2\, A_d^2\, P_x^2}{\Omega_x^2\sqrt{\alpha_m^4 - R_{\tau\tau}^2(\nu)}}\, R_{AA}(\nu)\,\exp\left\{\frac{-1}{2}\left[\alpha_m^4 - R_{\tau\tau}^2(\nu)\right]^{-1}\left[\left[(\tau-\tau_s+\tau_d)^2 + (\tau'-\tau_s+\tau_d)^2\right]\alpha_m^2 - 2(\tau-\tau_s+\tau_d)(\tau'-\tau_s+\tau_d)\, R_{\tau\tau}(\nu)\right]\right\} \tag{E.1-7}$$

where

$$\alpha_m^2 = \sigma_\tau^2 + \frac{1}{\Omega_x^2} + \frac{1}{2\Omega_1^2} + \frac{1}{2\Omega_2^2} \tag{E.1-8}$$

Finally, using (4.2-5) and defining $\Omega_m = 1/\alpha_m$

$$I_5(\tau,\tau') = \frac{k_1^2 k_2^2 A_d^2 A_c^2 p_x^2}{\Omega_x^2} \Omega_m^2 \exp\{-\frac{1}{2}\Omega_m^2 [(\tau-\tau_s+\tau_d)^2 + (\tau'-\tau_s+\tau_d)^2]\} \quad \text{(E.1-9)}$$

Thus, we have

$$\Pi(\tau,\tau',\nu) = k_1^2 k_2^2 p_x^2 \left\{ \left[\left(\frac{P_{n1}}{P_x} \right) \left(\frac{\Omega_n}{\Omega_{n1}} \right) e^{-\frac{1}{2}\nu^2\Omega_{1n1}^2} \right. \right. \quad \text{(E.1-10)}$$

$$+ \frac{R_{AA}(\nu)}{\Omega_x \sqrt{\frac{1}{\Omega_{x1}^2} + 2[\sigma_\tau^2 - R_{\tau\tau}(\nu)]}} \exp\left\{ \frac{-\nu^2}{2} \left[\frac{1}{\Omega_{1x}^2} + 2[\sigma_\tau^2 - R_{\tau\tau}(\nu)] \right]^{-1} \right\} \Bigg] x$$

$$\left[\left(\frac{P_{n2}}{P_x} \right) \left(\frac{\Omega_{2n2}}{\Omega_{n2}} \right) e^{-\frac{1}{2}(\nu-\tau+\tau')^2\Omega_{2n2}^2} + A_d^2 \left(\frac{\Omega_{1x}}{\Omega_x} \right) e^{\frac{1}{2}(\nu-\tau+\tau')^2\Omega_{1x}^2} \right]$$

$$+ \frac{A_d^2 R_{AA}(\nu)}{\Omega_x^2 \sqrt{\alpha_m - R_{\tau\tau}^2(\nu)}} \left[\exp\left\{ -\frac{1}{2} \left[\alpha_m^4 - R_{\tau\tau}^2(\nu) \right]^{-1} \cdot \left[(\tau-\tau_o)^2 \right. \right. \right.$$

$$\left. + (\tau'-\tau_o)^2 \right] \alpha_m^2 - 2(\tau-\tau_o)(\tau'-\tau_o) R_{\tau\tau}(\nu) \Big] \Big\}$$

$$+ \exp\left\{ -\frac{1}{2} \left[\alpha_m^4 - R_{\tau\tau}^2(\nu) \right]^{-1} \left[\left[(\nu+\tau'-\tau_o)^2 \right. \right. \right.$$

$$\left. \left. + (\tau-\nu-\tau_o)^2 \right] \alpha_m^2 - 2(\nu+\tau'-\tau_o)(\tau-\nu-\tau_o) R_{\tau\tau}(\nu) \Big] \right\}$$

where

$$\tau_o = \tau_s - \tau_d \quad \text{(E.1-11)}$$

is the multipath replication delay or steering parameter.

E.2 Array Processing

In this section we evaluate the integrals in (3.5-12) through (3.5-19) for the noise cross-spectrum (4.4-1), the filters and signal spectra of section 4.2 and the fourth order and second order statistics of sections 4.1 and 4.3. Following the derivation of (E.1-1) and (E.1-3),

$$M_a(\nu) = 2\pi\, k_f^2\, P_{na} \left(\frac{\Omega_{fna}}{\Omega_{na}} \right) e^{-\frac{1}{2}\nu^2 \Omega_{fna}^2} \tag{E.2-1}$$

$$K_a(\nu) = \frac{2\pi\, R_{Aa}(\nu)\, k_f^2\, P_x}{\Omega_x \sqrt{\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau a}^2 - R_{\tau a}(\nu)]}} \tag{E.2-2}$$

$$\exp\left\{ \frac{-\nu^2}{2} \left[\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau a}^2 - R_{\tau a}(\nu)] \right]^{-1} \right\}$$

$$M_c(\nu) = 2\pi\, k_f^2\, P_{nc} \left(\frac{\Omega_{fnc}}{\Omega_{nc}} \right) e^{-\frac{1}{2}\nu^2 \Omega_{fnc}^2} \tag{E.2-3}$$

$$K_c(\nu) = \frac{2\pi\, R_{Ac}(\nu)\, k_f^2\, P_x}{\Omega_x \sqrt{\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau c}^2 - R_{\tau c}(\nu)]}} \tag{E.2-4}$$

$$\exp\left\{ \frac{-\nu^2}{2} \left[\frac{1}{\Omega_{fx}^2} + 2[\sigma_{\tau c}^2 - R_{\tau c}(\nu)] \right]^{-1} \right\}$$

where (E.2-5)

a) $\frac{1}{\Omega_{fna}} = \sqrt{\frac{1}{\Omega_f^2} + \frac{1}{\Omega_{na}^2}}$ b) $\frac{1}{\Omega_{fnc}} = \sqrt{\frac{1}{\Omega_f^2} + \frac{1}{\Omega_{nc}^2}}$

and

$$\frac{1}{\Omega_{fx}} = \sqrt{\frac{1}{\Omega_f^2} + \frac{1}{\Omega_x^2}} \tag{E.2-6}$$

Next we define the following modified parameters from (4.3-18) to (4.3-26) :

$$\alpha'_{ij}(\nu,\tau,\tau') = \alpha_{ij}(\nu,\tau,\tau') + \frac{1}{\Omega_{fx}^2} \tag{E.2-7}$$

Using these we obtain

$$J_1(\tau,\tau',\nu) = \int_{\omega}\int_{\omega'} e^{j[\omega\nu+\omega'(\nu-\tau+\tau')]} \frac{P^2}{\Omega_x^2} \quad x \tag{E.2-8}$$

$$\exp\{-\frac{1}{2}[\alpha'_{11}(\nu,\tau,\tau')\omega^2+2\alpha_{12}(\nu,\tau,\tau')\omega\omega'+\alpha'_{13}(\nu,\tau,\tau')\omega'^2]\}\, d\omega\, d\omega' =$$

$$\frac{P_x^2}{\Omega_x^2} \exp\left[\frac{-\frac{1}{2}[\alpha'_{11}(\nu,\tau,\tau')\nu^2-2\alpha_{12}(\nu,\tau,\tau')\nu(\tau-\tau'-\nu)+\alpha'_{13}(\nu,\tau,\tau')(\tau-\tau'-\nu)^2]}{[\alpha'_{11}(\nu,\tau,\tau')\alpha'_{13}(\nu,\tau,\tau') - \alpha_{12}^2(\nu,\tau,\tau')]}\right] - R_{1,2}^{[4]}(\tau,\nu,\nu+\tau')$$

$$\sqrt{\alpha'_{11}(\nu,\tau,\tau')\alpha'_{13}(\nu,\tau,\tau') - \alpha_{12}^2(\nu,\tau,\tau')}$$

Similarly,

$$J_2(\tau,\tau',\nu) = \int_{\omega}\int_{\omega'} e^{j[\omega(\tau-\tau_o)+\omega'(\tau'-\tau_o)]} \frac{P_x^2}{\Omega_x^2} \quad x \tag{E.2-9}$$

$$\exp\{-\frac{1}{2}[\alpha'_{21}(\nu,\tau,\tau')\omega^2+2\alpha_{22}(\nu,\tau,\tau')\omega\omega'+\alpha'_{23}(\nu,\tau,\tau')\omega'^2]\}\, d\omega\, d\omega' =$$

$$\frac{\frac{P_x^2}{\Omega_x^2}\exp\left[\frac{-\frac{1}{2}\left[\alpha'_{21}(\nu,\tau,\tau')(\tau-\tau_o)^2-\alpha_{22}(\nu,\tau,\tau')(\tau-\tau_o)(\tau'-\tau_o)+\alpha'_{23}(\nu,\tau,\tau')(\tau'-\tau_o)^2\right]}{\left[\alpha'_{21}(\nu,\tau,\tau')\alpha'_{23}(\nu,\tau,\tau')-\alpha_{22}^2(\nu,\tau,\tau')\right]}\right]-R_{1,2}^{[4]}(\tau,\nu,\nu+\tau')}{\sqrt{\alpha'_{21}(\nu,\tau,\tau')\alpha'_{23}(\nu,\tau,\tau')-\alpha_{22}^2(\nu,\tau,\tau')}}$$

and finally,

$$J_3(\tau,\tau',\nu)=\int_{\omega}\int_{\omega'} e^{j[\omega(\nu+\tau')+\omega'(\nu-\tau)]}\frac{P_x^2}{\Omega_x^2}\times e^{-j(\omega-\omega)\tau_o} \qquad \text{(E.2-10)}$$

$$\exp\{-\tfrac{1}{2}[\alpha'_{31}(\nu,\tau,\tau')\omega^2+2\alpha_{32}(\nu,\tau,\tau')\omega\omega'+\alpha'_{33}(\nu,\tau,\tau')\omega'^2]\}\,d\omega\,d\omega' =$$

$$\frac{\frac{P_x^2}{\Omega_x^2}\exp\left[\frac{-\frac{1}{2}\left[\alpha'_{31}(\nu,\tau,\tau')(\nu+\tau'-\tau_o)^2-2\alpha_{32}(\nu,\tau,\tau')(\nu+\tau'-\tau_o)(\nu-\tau-\tau_o)+\alpha'_{33}(\nu,\tau,\tau')(\nu-\tau-\tau_o)^2\right]}{\left[\alpha'_{31}(\nu,\tau,\tau')\alpha'_{33}(\nu,\tau,\tau')-\alpha_{32}^2(\nu,\tau,\tau')\right]}\right]R_{1,2}(\tau,\nu,\nu+\tau')}{\sqrt{\alpha'_{31}(\nu,\tau,\tau')\alpha'_{33}(\nu,\tau,\tau')-\alpha_{32}^2(\nu,\tau,\tau')}}$$

the replication time in this case is

$$\tau_o=\tau_{s1}-\tau_{s2} \qquad \text{(E.2-11)}$$

Using these formulas we have

$$\Pi(\tau,\tau',\nu)=P_x^2\left\{\left(\frac{P_{na}^2}{P_x^2}\right)\left(\frac{\Omega_{fna}^2}{\Omega_f^2}\right)\exp\{-\tfrac{1}{2}[\nu^2+(\nu-\tau+\tau')^2]\Omega_{fna}^2\}\right.$$

$$\left.+\frac{R_{Aa}(\nu-\tau+\tau')(Pna/Px)\Omega fna}{\Omega_{na}\,\Omega_x\sqrt{\frac{1}{\Omega_{fx}^2}+\tfrac{1}{2}[\sigma_{\tau a}^2-R_{\tau a}(\nu-\tau+\tau')]}}\right\}\times \qquad \text{(E.2-12)}$$

$$\exp\left\{\frac{-\nu^2\Omega_{fna}^2}{2}\ \frac{-(\nu-(\tau-\tau'))^2}{2}\ \frac{1}{\Omega_{fx}^2}+2[\sigma_{\tau a}^2-R_{\tau a}(\nu-\tau+\tau')]\right\}^{-1}$$

$$+\frac{RAa(\nu)\quad(Pna/Px)\Omega fna}{\Omega_{na}\,\Omega_x\sqrt{\frac{1}{\Omega_{fx}^2}+2[\sigma_{\tau a}^2-R_{\tau a}(\nu)]}}\qquad x \qquad (E.2\text{-}12)$$

$$+\left(\frac{P_{ca}^2}{P_x^2}\right)\left(\frac{\Omega_{fnc}^2}{\Omega_f^2}\right)\ \exp\{-\tfrac{1}{2}[(\nu+\tau')^2+(\nu-\tau)^2]\Omega_{fnc}^2\}$$

$$+\frac{R_{Ac}(\nu+\tau')\ (Pnc/Px)\Omega fnc}{\Omega_{nc}\,\Omega_x\sqrt{\frac{1}{\Omega_{fx}^2}+2[\sigma_{\tau a}^2-R_{\tau c}(\nu+\tau')]}}\qquad x$$

$$\exp\{-\tfrac{1}{2}(\nu-\tau_o+\tau')^2\left[\frac{1}{\Omega_{fx}^2}+2[\sigma_{\tau a}^2-R_{\tau c}(\nu+\tau')]\right]^{-1}-\tfrac{1}{2}(\tau-\nu)^2\Omega_{fnc}^2\}$$

$$+\frac{R_{Ac}(\tau-\nu)\,(Pnc/Px)\Omega fnc}{\Omega_{nc}\,\Omega_x\sqrt{\frac{1}{\Omega_{fx}^2}+2[\sigma_{\tau a}^2-R_{\tau c}(\tau-\nu)]}}\qquad x$$

$$\exp\{-\tfrac{1}{2}(\tau-\tau_o-\nu)^2\left[\frac{1}{\Omega_{fx}^2}+2[\sigma_{\tau a}^2-R_{\tau c}(\tau-\nu)]\right]^{-1}-\tfrac{1}{2}(\nu+\tau')^2\Omega_{fnc}^2\}$$

$$+\frac{1}{\Omega_x^2}\,R_{1,2}^{[4]}(\tau,\nu,\nu+\tau')\ x\quad[$$

$$\exp\left\{\frac{-\frac{1}{2}[\alpha'_{11}(\nu,\tau,\tau')\nu^2-2\alpha_{12}(\nu,\tau,\tau')\nu(\tau-\tau'-\nu)+\alpha'_{13}(\nu,\tau,\tau')(\tau-\tau'-\nu)^2]}{[\alpha'_{11}(\nu,\tau,\tau')\alpha'_{13}(\nu,\tau,\tau')-\alpha^2_{12}(\nu,\tau,\tau')]}\right\}$$

$$\Big/\sqrt{\alpha'_{11}(\nu,\tau,\tau')\alpha'_{13}(\nu,\tau,\tau')-\alpha^2_{12}(\nu,\tau,\tau')}$$

$$\exp\left\{\frac{-\frac{1}{2}[\alpha'_{21}(\nu,\tau,\tau')(\tau-\tau_o)^2-2\alpha_{22}(\nu,\tau,\tau')(\tau-\tau_o)(\tau'-\tau_o)+\alpha'_{23}(\nu,\tau,\tau')(\tau'-\tau_o)^2]}{[\alpha'_{21}(\nu,\tau,\tau')\alpha'_{23}(\nu,\tau,\tau')-\alpha^2_{22}(\nu,\tau,\tau')]}\right\}$$

$$\Big/\sqrt{\alpha'_{21}(\nu,\tau,\tau')\alpha'_{23}(\nu,\tau,\tau')-\alpha^2_{22}(\nu,\tau,\tau')}$$

$$\exp\left\{\frac{-\frac{1}{2}[\alpha'_{31}(\nu,\tau,\tau')(\nu+\tau'-\tau_o)^2-2\alpha_{32}(\nu,\tau,\tau')(\nu+\tau'-\tau_o)(\nu-\tau-\tau_o)+\alpha'_{33}(\nu,\tau,\tau')(\nu-\tau-\tau_o)^2]}{[\alpha'_{31}(\nu,\tau,\tau')\alpha'_{33}(\nu,\tau,\tau')-\alpha^2_{32}(\nu,\tau,\tau')]}\right\}$$

$$\Big/\sqrt{\alpha'_{31}(\nu,\tau,\tau')\alpha'_{33}(\nu,\tau,\tau')-\alpha^2_{32}(\nu,\tau,\tau')}$$

APPENDIX G

TABULATION OF THE FIRST FEW $G_{n,\ell,m}(p,q,\tau)$

The following is a tabulation of the first few coefficients $G_{n,\ell,m}(p,q,\tau)$ of the polynomials $P_{n,\ell,m}(\omega,\omega',\tau)$ defined by equation (5.2-19). All unlisted coefficients for a given polynomial are zero.

$P_{0,0,0}(\omega,\omega',\tau)$

$$G_{0,0,0}(0,0,\tau) = 1 \tag{G.1-1}$$

$P_{0,1,0}(\omega,\omega',\tau)$

$$G_{0,1,0}(0,0,\tau) = \frac{1}{\sigma_h^2\,(1-\rho_h^2(\tau))} \tag{G.1-2}$$

$$G_{0,1,0}(1,0,\tau) = 1 \tag{G.1-3}$$

$P_{0,0,1}(\omega,\omega',\tau)$

$$G_{0,0,1}(0,0,\tau) = G_{0,1,0}(0,0,\tau) \tag{G.1-4}$$

$$G_{0,0,1}(0,1,\tau) = G_{0,1,0}(1,0,\tau) \tag{G.1-5}$$

$P_{1,0,0}(\omega,\omega',\tau)$

$$G_{1,0,0}(0,0,\tau) = \frac{\rho_h(\tau)}{\sigma_h^2\,(1-\rho_h^2(\tau))} \tag{G.1-6}$$

$P_{0,1,1}(\omega,\omega',\tau)$

$$G_{0,1,1}(1,1,\tau) = 1 \qquad \text{(G.1-7)}$$

$$G_{0,1,1}(1,0,\ ) = G_{0,1,1}(0,1,\ ) = \frac{1}{\sigma_h^2\,(1-\rho_h^2(\tau))} \qquad \text{(G.1-8)}$$

$$G_{0,1,1}(0,0,\tau) = \frac{1}{(1-\rho_h^2(\tau))\sigma_h^4} \qquad \text{(G.1-9)}$$

$P_{2,0,0}(\omega,\omega',\tau)$

$$G_{2,0,0}(0,0,\ ) = G_{0,1,0}(0,0,\tau)^2 \qquad \text{(G.1-10)}$$

$$G_{2,0,0}(1,0,\ ) = G_{2,0,0}(0,1,\ ) = G_{0,1,0}(0,0,\tau) \qquad \text{(G.1-11)}$$

$$G_{2,0,0}(1,1,\tau) = 1 \qquad \text{(G.1-12)}$$

$P_{1,1,0}(\omega,\omega',\tau)$

$$G_{1,1,0}(1,0,\tau) = 3\,G_{1,0,0}(0,0,\tau) \qquad \text{(G.1-13)}$$

$$G_{1,1,0}(0,0,\tau) = G_{1,0,0}(0,0,\tau)G_{0,1,0}(0,0,\tau) \qquad \text{(G.1-14)}$$

$P_{1,0,1}(\omega,\omega',\tau)$

$$G_{1,0,1}(0,1,\tau) = 3\,G_{1,0,0}(0,0,\tau) \qquad \text{(G.1-15)}$$

$$G_{1,0,1}(0,0,\tau) = G_{1,1,0}(0,0,\tau) \qquad \text{(G.1-16)}$$

$P_{0,2,0}(\omega,\omega',\tau)$

$$G_{0,2,0}(2,0,\tau) = 1 \qquad \text{(G.1-17)}$$

$$G_{0,2,0}(1,0,\tau) = 2G_{0,1,0}(0,0,\tau) \qquad \text{(G.1-18)}$$

$$G_{0,2,0}(0,0,\tau) = G_{0,1,0}(0,0,\tau)^2 \qquad \text{(G.1-19)}$$

$P_{0,0,2}(\omega,\omega',\tau)$

$$G_{0,0,2}(0,2,\tau) = 1 \qquad \text{(G.1-20)}$$

$$G_{0,0,2}(0,1,\tau) = G_{0,2,0}(1,0,\tau) \qquad \text{(G.1-21)}$$

$$G_{0,0,2}(0,0,\tau) = G_{0,2,0}(0,0,\tau) \qquad \text{(G.1-22)}$$

APPENDIX H

VARIANCE INTEGRALS FOR

THE RANDOM SINUSOIDAL BOUNDARY MODEL

H.1 Multipath Processor

In this section we evaluate (3.3-9) through (3.3-13) for the spectra (5.3-1), the filters (5.3-2) and the random sinusoidal boundary discussed in sections (5.1) through (5.2):

$$I_1(\upsilon) = \frac{P_{n1}\Omega_{fn2}/\Omega_{n1}}{\sqrt{1 + \upsilon^2\Omega_{fn1}^2}} \qquad \text{(H.1-1)}$$

$$I_3(\upsilon) = \frac{P_{n2}\Omega_{fn2}/\Omega_{n2}}{\sqrt{1 + \upsilon^2\Omega_{fn2}^2}} + \frac{1}{R_d}2\,\frac{P_x\,\Omega fx/\Omega x}{\sqrt{1 + \upsilon^2\Omega_x^2}} \qquad \text{(H.1-2)}$$

Using (5.3-8):

$$I_2(\upsilon) = \frac{P_x/\Omega_x}{(r_o+r_o')}2 \sum_{n=0}^{\infty} \frac{c_n}{(n!)}2\, Q_{\chi_i}(n,n,\upsilon)\, \varepsilon_n \cos(n\Omega_s\upsilon) \sum_{1=0}^{\infty} \sum_{m=0}^{\infty}$$

$$\frac{1}{[4(n+1)]}1+m\ \frac{A(1,2n+1)}{1!}\ \frac{A(m,2n+1)}{m!} \sum_{r=0}^{\infty} \frac{\Gamma(1+n+m+r+\frac{1}{2})}{\Gamma(1+m+n+\frac{1}{2})}\ \frac{1}{r!} \sum_{p,q=0}^{n+1+m}$$

$$\{(n+1)\ \ 1-\rho_h^2(\upsilon)\ \}^{1+n+m}\ G_{n,1,m}(p,q,\upsilon)\ \ \rho_h^2(\upsilon)\ ^r \qquad \text{(H.1-3)}$$

$$RE\{\frac{1}{2\pi\ \Gamma\ 2(n+1+m+r)+1}\ G_{31}^{13}\ (\tfrac{1}{4}c_n^2(\alpha_{fx}-j\upsilon)^2\ \Big|_{0,\frac{1}{2},n+1+m-p-q-\frac{1}{2}}^{\frac{1}{2}-(2r+1+m+n+p+q)})\}$$

The corresponding expression for $I_4(\tau,\tau',\upsilon)$ is much more difficult obtain in closed form because of the inverse first power of frequency in the exponent of (5.2-29). As in (5.1-6) this is handled by

convolution with the function $f_2^n(\upsilon)$ defined by (5.1-8). Unfortunately, the convolution must either be executed numerically or taken in the limit of large distances from the boundary. In the latter case the convolution can be ignored as discussed in section (5.1). In the former case we write:

$$I_4(\tau,\tau',\upsilon) = \frac{P_x^2/\Omega_x^2}{R_d^2(r+r')^2} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \int \left(\frac{c_n}{n!}\right)^2 Q_{\chi_1}(n,n,\upsilon)\, \varepsilon_n \cos(n\Omega_s\upsilon)$$

$$\sum_{1,m=0}^{\infty\ \infty} \frac{1}{[4(n+1)]}\, 1+m\, \frac{A(1,2n+1)}{1!}\, \frac{A(m,2n+1)}{m!} \sum_{r=0}^{\infty} \frac{1}{\Gamma(1+n+m+\frac{1}{2})}\, \frac{1}{r!} \sum_{p,q=0}^{n+1+m}$$

(H.1-3)

$$\{(n+1)[1-\rho_h^2(\upsilon)]\}^{1+m+n}\ G_{n,1,m}(p,q,\upsilon)\ [\rho_h^2(\upsilon)]^r\ \frac{1}{(2\pi)^2\Gamma(n+1+m+r+\frac{1}{2})}$$

$$\text{RE}\ \{G_{31}^{13}\ (\tfrac{1}{4}c_n^2\ [\alpha_{fx}-j(\tau'-\mu-\tau_o)]^2\ \Big|_{0,\frac{1}{2},n+m-p}^{\frac{1}{2}+\frac{1}{2}n-r-p-1})\}\ f_2^n(\mu)$$

$$\text{RE}\ \{G_{31}^{13}\ (\tfrac{1}{4}c_n^2\ [\alpha_{fx}+j(\tau-\mu-\tau_o)]^2\ \Big|_{0,\frac{1}{2},n+1-q}^{\frac{1}{2}+\frac{1}{2}n-r-q-m})\}\ f_2^n(\mu')\ d\mu\ d\mu'$$

Expressions (H.1-1) through (H.1-3) can easily be used in conjunction with (3.3-8) to compute the correlator output variance. To compute the tracking error we need the following derivatives:

$$\frac{\partial^2}{\partial\tau\,\partial\tau'}\ I_3(\upsilon-\tau+\tau') =$$

(H.1-4)

$$\left[\frac{P_{n2}}{\Omega_{n2}}\ \Omega_{fn2}^3\ \frac{\left[1 = 2[\Omega_{fn2}(\upsilon-\tau+\tau')]^2\right]}{\left[1 + [\Omega_{fn2}(\upsilon-\tau+\tau')]^2\right]}\right]_{5/2}$$

$$+\ \frac{P_x}{\Omega_x}\Omega_{fx}^3\ \frac{\left[1 - 2[\Omega_{fx}(\upsilon-\tau+\tau')]^2\right]}{\left[1 - [\Omega_{fx}(\upsilon-\tau+\tau')]^2\right]}\ {}_{5/2}$$

and

$$\frac{\partial^2}{\partial\tau\partial\tau'} I_4(\tau,\tau',\upsilon) = \frac{P_x^2/\Omega_x^2}{R_d^2(r_o+r_o')^2} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \frac{c_n^4}{n!^2} Q_{\chi_1}(n,n,\upsilon)\ \epsilon_n \cos(n\Omega_s\upsilon)$$

$$\sum_{1=0}^{\infty} \sum_{m=0}^{\infty} \frac{1}{[4(n+1)]^{1+m}} \frac{A(1,2n+1)}{1!} \frac{A(m,2n+1)}{m!} \sum_{r=0}^{\infty} \frac{1}{\Gamma(1+n+m+\frac{1}{2})} \frac{1}{r!} \sum_{p,q=0}^{n+1+m}$$

(H.1-5)

$$\{(n+1)\left[1-\rho_h^2(\upsilon)\right]\}^{1+m+n} G_{n,1,m}(p,q,\upsilon) \left[\rho_h^2(\upsilon)\right]^r \frac{1}{(2\pi)^2\,\Gamma(n+1+m+r+\frac{1}{2})}$$

$$RE\{-jc_n\, G_{31}^{13}\left(\tfrac{1}{4}c_n^2\left[\alpha_{fx}-j(\tau-\mu-\tau_o)\right]^2 \,\middle|\, \begin{matrix}\frac{1}{2}+\frac{1}{2}n-r-p-1-1\\ 0,\frac{1}{2},n+m-p\end{matrix}\right)\} f_2^n(d\mu)$$

$$RE\{+jc_n\, G_{31}^{13}\left(\tfrac{1}{4}c_n^2\left[\alpha_{fx}+j(\tau'-\mu'-\tau_o)\right]^2 \,\middle|\, \begin{matrix}\frac{1}{2}+\frac{1}{2}n-r-q-m\\ 0,\frac{1}{2},n+1-q\end{matrix}\right)\} f_2^n(d\mu')$$

These can readily be substituted into (5.3-16) to obtain the tracking variance.

Appendix I

Numerical Evaluation of the $G_{13}^{31}(z)$

I.1 Relation to Struve's and Weber's Functions

Following Watson[78] (p. 331, sec. 10.41, #3)

$$\frac{1}{2\pi\ \Gamma(a)}\ G_{13}^{31}\left(\frac{z^2}{4}\ \middle|\ {}^{\frac{1}{2}}_{0,\frac{1}{2},a-\frac{1}{2}}\right) = \int_0^\infty e^{-zu}(1+u^2)^{-a}\,du \qquad (I.1\text{-}1)$$

$$= \tfrac{1}{2}\Gamma(1-a)\ \Gamma(\tfrac{1}{2})\ (\tfrac{1}{2}z)^{a-\frac{1}{2}}\left[S_{-a+\frac{1}{2}}(z)-Y_{-a+\frac{1}{2}}(z)\right]$$

where $S_\nu(z)$ is Struve's function

$$S_\nu(z) = \sum_{m=0}^{\infty} \frac{(-1)^m\ (\frac{1}{2}z)^{\nu+zm+1}}{\Gamma(m+3/2)\Gamma(\nu+m+3/2)} \qquad (I.1\text{-}2)$$

and $Y_\nu(z)$ is Weber's function defined in general by

$$Y_\nu(z) = \frac{J\nu(z)\ \mathrm{Cos}(\nu\pi) - J\text{-}\nu(z)}{\mathrm{Sin}\ (\nu\pi)} \qquad (I.1\text{-}3)$$

For integer values of ν (I.1-2) becomes

$$Y_n(z) = -\frac{(\frac{1}{2}z)^{-n}}{\pi} \sum_{k=0}^{n-1} \frac{(n-k-1)!}{k!}\ (\tfrac{1}{4}z^2)^k + \frac{2}{\pi}\ \ell n(\tfrac{1}{2}z)\ J_n(z) \qquad (I.1\text{-}4)$$

$$\frac{(\frac{1}{2}z)^n}{\pi} \sum_{k=0}^{\infty} \{\psi(k+1)+\psi(n+k+1)\}\ \frac{(-\frac{1}{4}z^2)^k}{k!(n+k)!} = (-1)^n\ Y\text{-}n(z)$$

where

$$J_\nu(z) = (\tfrac{1}{2}z)^\nu \sum_{k=0}^{\infty} \frac{(-\frac{1}{4}z^2)^k}{k!\Gamma(\nu+k+1)} \qquad (I.1\text{-}5)$$

and

$$\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)} \qquad (I.1\text{-}6)$$

is the digamma function. Again for integer values of z we have

$$\psi(n) = -\gamma + \sum_{k=1}^{n-1} k^{-1} \qquad \text{(I.1-7)}$$

where $\gamma = 0.57721566.$

We can extend (I.1-1) by successive differentiation. First we note the following

$$\left(\frac{d}{dz}\right)^p (z)^k = \frac{\Gamma(k+1)}{\Gamma(k-p+1)} z^{k-p} \qquad \text{(I.1-8)}$$

The result is valid for $k>-1$ and is zero for $k-p = -1,-2,-3...$

For example:

$$\left(\frac{d}{dz}\right)^p \left[(\tfrac{1}{2}z)^n S_{-n}(z) \right] = \sum_{k=0}^{\infty} \frac{(-1)^k (\tfrac{1}{2}z)^{2k+1} \Gamma(2k+z)}{\Gamma(k+3/2)\, \Gamma(k-n+3/2)\, \Gamma(2k+2-p)} \frac{1}{z^p} \qquad \text{(I.1-9)}$$

$$\left(\frac{d}{dz}\right)^p \left[(\tfrac{1}{2}z)^n J_n(z) \right] = \sum_{k=0}^{\infty} \frac{(-1)^k (\tfrac{1}{2}z)^{2k+2n} \Gamma(2k+2\nu+1)}{k!\Gamma(n+k+1)\, \Gamma(2k+2\nu-p+1)} \frac{1}{z^p} \qquad \text{(I.1-10)}$$

$$\left(\frac{d}{dz}\right)^p \left[(\tfrac{1}{2}z)^n Y_{-n}(z) \right] = \frac{2(-1)^{n-1}}{\pi} \left[\tfrac{1}{2} \sum_{k=0}^{n-1} \frac{(n-k-1)!\, (\tfrac{1}{2}z)^{2k}\, \Gamma(2k+1)}{k!\quad \Gamma(2k-p+1)} \frac{1}{z^p} \right.$$

$$- \sum_{\nu=0}^{p} \left(\frac{d}{dz}\right)^{p-\nu} \left[\ln(\tfrac{1}{2}z) \right] \left(\frac{d}{dz}\right)^2 \left[(\tfrac{1}{2}z)^n J_n(z) \right] \frac{p!}{\nu!(p-\nu)!}$$

$$\left. + \tfrac{1}{2} \sum_{k=0}^{\infty} \frac{\{\psi(k+1) + \psi(n+k+1)\}\, (-\tfrac{1}{2}z)^{2k+2n}\, \Gamma(2k+2n+1)}{k!\quad (n+k)!\qquad \Gamma(2k+2n-p+1)\qquad z^p} \right] \qquad \text{(I.1-11)}$$

Using (I.1-9) through (I.1-11) and (5.3-9)

$$\frac{1}{2\pi\Gamma(a)} \; G_{13}^{31}\left(\frac{z^2}{4}\,\middle|\,{}^{\frac{1}{2}-\frac{1}{2}p}_{0,\frac{1}{2},a-\frac{1}{2}p-\frac{1}{2}}\right) = \int_0^\infty u^p\, e^{-zu}\, (Hu^2)^{-a}\, du =$$

$$(-1)^p \tfrac{1}{2}\Gamma(-a)\,\Gamma(\tfrac{1}{2})\left(\frac{d}{dz}\right)^p \left\{ (\tfrac{1}{2}z)^{a+\frac{1}{2}} \left[S_{-a-\frac{1}{2}}(z) - Y_{-a-\frac{1}{2}}(z) \right] \right\} \qquad \text{(I.1-12)}$$

Provided $a=n+\frac{1}{2}$ where n is an integer, we may use (I.1-9) through (I.1-11) directly. However, if $a=n$ (I.1-12) is indeterminate because

$$S_{-n-\frac{1}{2}}(z) = Y_{-n-\frac{1}{2}}(z) = \qquad \text{(I.1-13)}$$

$$(-1)^n\, J_{n+\frac{1}{2}}(z)$$

and $\Gamma(-n) = \pm\infty$.

I.2 Evaluation in the Special Case of Integer Values of a.

The limit $a \to n$ in (I.1-12) must be approached with caution. Taking (I.1-13) as a starting point we have from (I.1-2) using $a = n - \varepsilon$

$$S_{-(n+\frac{1}{2})+\varepsilon}(z) = \sum_{k=0}^{n-1} \frac{(-1)^k}{\Gamma(k+3/2)} \frac{(\frac{1}{2}z)^{2k-(n+\frac{1}{2})+1+\varepsilon}}{\Gamma(k-n+1+\varepsilon)}$$

$$+(\tfrac{1}{2}z)^{n+\frac{1}{2}+\varepsilon} (-1)^n \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(k+(n+\frac{1}{2})+1)} \frac{(\frac{1}{2}z)^{2k}}{\Gamma(k+1+\varepsilon)} \qquad \text{(I.2-1)}$$

and from (I.1-3)

$$Y_{-(n+\frac{1}{2})+\varepsilon}(z) =$$

$$\frac{J_{-(n+\frac{1}{2})+\varepsilon}(z)\ (-1)^n \sin(\pi\varepsilon) - J_{n+\frac{1}{2}-\varepsilon}(z)}{(-1)^{n+1} \cos(\pi\varepsilon)}$$

$$= (-1) \frac{\sin(\pi\varepsilon)}{\cos(\pi\varepsilon)} (\tfrac{1}{2}z)^{-(n+\frac{1}{2})+\varepsilon} \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(k+1)} \frac{(\frac{1}{2}z)^{2k}}{\Gamma(-n-\frac{1}{2}+\varepsilon+k+1)}$$

$$+ \frac{(-1)^n}{\cos(\pi\varepsilon)} (\tfrac{1}{2}z)^{(n+\frac{1}{2})-\varepsilon} \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(k+1)} \frac{(\frac{1}{2}z)^{2k}}{\Gamma(n+\frac{1}{2}-\varepsilon+k+1)} \qquad \text{(I.2-2)}$$

Now

$$(\tfrac{1}{2}z)^{\varepsilon} = e^{\varepsilon \ln(\frac{1}{2}z)} \equiv 1 + \varepsilon \ln(\tfrac{1}{2}z) \ldots \qquad \text{(I.2-3)}$$

$$\Gamma(-n+\varepsilon) = \frac{(-1)^n \pi}{\Gamma(1+n-\varepsilon) \sin(\pi\varepsilon)} \qquad \text{(I.2-4)}$$

and

$$\frac{1}{\Gamma(a+\varepsilon)} \equiv \frac{1}{\Gamma(a)} \left[1 - \psi(a)\varepsilon\right] \qquad \text{(I.2-5)}$$

where $a \neq 0,-1,-2,\ldots$ Substituting (I.2-3) through (I.2-5) into (I.2-1) and (I.2-2) we have

$$\lim_{\varepsilon\to 0} \tfrac{1}{2}\Gamma(-n+\varepsilon)\,\Gamma(\tfrac{1}{2})\,(\tfrac{1}{2}z)^{n+\frac{1}{2}-\varepsilon} \left[S_{-(n+\frac{1}{2})+\varepsilon}(z) - Y_{-(n+\frac{1}{2})+\varepsilon}(z)\right]$$

$$= \lim_{\varepsilon\to 0} \frac{\frac{1}{2}\Gamma(\frac{1}{2})\,(\frac{1}{2}z)^{n+\frac{1}{2}}\pi(-1)^n}{\Gamma(1+n)\,\mathrm{Sin}(\pi\varepsilon)} \Bigg[$$

$$\sum_{k=0}^{n-1} \frac{(-1)^k}{\Gamma(k+3/2)} (\tfrac{1}{2}z)^{2k-(n+\frac{1}{2})+1}\,\Gamma(-k+n)\,\mathrm{Sin}(\pi\varepsilon)\,(-1)^{k-n+1}$$

$$+ (\tfrac{1}{2}z)^{n+\frac{1}{2}} \left[1+\varepsilon\ln(\tfrac{1}{2}z)\right] (-1)^n \sum_{k=0}^{\infty} \frac{(-1)^k\,(\frac{1}{2}z)^{2k}}{\Gamma(k+n+\frac{1}{2}+1)\,\Gamma(k+1)} \left[1-\varepsilon\,\psi(k+1)\right]$$

$$+ \frac{\mathrm{Sin}(\pi\varepsilon)}{\mathrm{Cos}(\pi\varepsilon)} (\tfrac{1}{2}z)^{-(n+\frac{1}{2})+\varepsilon} \sum_{k=0}^{\infty} \frac{(-1)^k\,(\frac{1}{2}z)^{2k}}{\Gamma(k+1)\,\Gamma(-n-\frac{1}{2}+k+1)}$$

$$- (\tfrac{1}{2}z)^{n+\frac{1}{2}}\; 1-\varepsilon\,\ln(\tfrac{1}{2}z)\; (-1)^n \sum_{k=0}^{\infty} \frac{(-1)^k\,(\frac{1}{2}z)^{2k}}{\Gamma(k+n+\frac{1}{2}+1)\,\Gamma(k+1)} \left[1+\varepsilon\,\psi(k+n+\tfrac{1}{2}+1)\right] \Bigg]$$

$$= \frac{\frac{1}{2}\pi\Gamma(\frac{1}{2})}{n!} (\tfrac{1}{2}z)^{n+\frac{1}{2}} \Bigg[\frac{-(\frac{1}{2}z)^{-n-\frac{1}{2}}}{\pi} \sum_{k=0}^{n-1} \frac{\Gamma(-k+n)}{\Gamma(k+3/2)} (\tfrac{1}{4}z^2)^{k+\frac{1}{2}}$$

$$+\frac{2}{\pi}\,\ln(\tfrac{1}{2}z)\,J_{n+\frac{1}{2}}(z) + (-1)^n\,J_{-n-\frac{1}{2}}(z)$$

$$- (\tfrac{1}{2}z)^{n+\frac{1}{2}} \sum_{k=0}^{\infty} \frac{(-\frac{1}{4}z^2)^k}{k!\,\Gamma(k+n+\frac{1}{2}+1)} \left[\psi(k+1)+\psi(k+n+\tfrac{1}{2}+1)\right] \Bigg] \qquad \text{(I.2-6)}$$

Finally,

$$\frac{1}{2\pi\,\Gamma(n)}\;G_{13}^{31}\left(\frac{z^2}{4}\,\middle|\,{}^{\frac{1}{2}-\frac{1}{2}p}_{0,\frac{1}{2},n-\frac{1}{2}p-\frac{1}{2}}\right)$$

$$= \int_0^\infty u^p\;\frac{e^{-zu}}{(1+u^2)^n}\;du \qquad \text{(I.2-7)}$$

$$= (-1)^p\,\frac{\pi^{\frac{1}{2}}}{4n!}\left[\sum_{k=0}^{n-1}\frac{\Gamma(-k+n)\;\Gamma(2k+2)\;(\frac{1}{2}z^2)^{k+\frac{1}{2}}}{\Gamma(k+3/2)\;\Gamma(2k+2-p)\;z^p}\right.$$

$$+\;2\sum_{\nu=0}^{p}\left(\frac{d}{dz}\right)^{p-\nu}\left[\ell n(\tfrac{1}{2}z)\right]\left(\frac{d}{dz}\right)^{p}\left[(\tfrac{1}{2}z)^n\,J_n(z)\right]\frac{p!}{\nu!\,(p-\nu)!}$$

$$+\;(-1)^n\pi\sum_{k=0}^{\infty}\frac{(-\frac{1}{4}z^2)^k\;\Gamma(2k+1)}{k!\,\Gamma(-n-\frac{1}{2}+k+1)\;\Gamma(2k-p+1)\;z^p}$$

$$\left.-\sum_{k=0}^{\infty}\frac{(-1)^k\;(\frac{1}{2}z)^{2k+n+\frac{1}{2}}\,\Gamma(2k+n+3/2)\left[\psi(k+1)+\psi(k+n+3/2)\right]}{k!\,\Gamma(n+\frac{1}{2}+k+1)\;\Gamma(2k+n+3/2-p)\;z^p}\right]$$

In computing the digamma function for fractional arguments the following is useful:

$$\psi(n+\tfrac{1}{2}) = -\gamma - 2\ell n 2 + 2\sum_{k=1}^{n}\frac{1}{(2K-1)} \qquad \text{(I.2-8)}$$

I.3 The Asymptotic Expansion

For large z the sories in the two preceeding sections converge too slowly to be useful. Fortunately, a simple asymptotic formula can be used for all values of a. Equation (I.1-8) must be altered, however, to accomodate k<-1:

$$\left(\frac{d}{dz}\right)^p \ (z)^k = \frac{(-1)^p \ \Gamma(p-k)}{\Gamma(-k)} \ z^{k-p} \qquad \text{(T.3-1)}$$

Using this we obtain

$$\int_0^\infty u^p \ e^{-zu} \ (1+u^2)^{-a} \ du = \qquad \text{(I.3-2)}$$

$$\left(\frac{d}{dz}\right)^p \left[\sum_{m=0}^{\infty} \frac{(-1)^m \ (\alpha)m(2m)!}{m! \ z^{2m+1}}\right] =$$

$$\sum_{m=0}^{\infty} \left[\frac{(-1)^{m+p}(\alpha)m \ \Gamma(P+2m+1)}{m! \ z^{2m+1-p}}\right]$$

Appendix J

Extensions to Large Arrays

J.1 The N Element Cross Correlator

The analysis of section 4.3 and 4.4 for a two element cross-correlator is sufficiently general to cover larger arrays. In this section we consider configuration shown in figure (J.1-1)

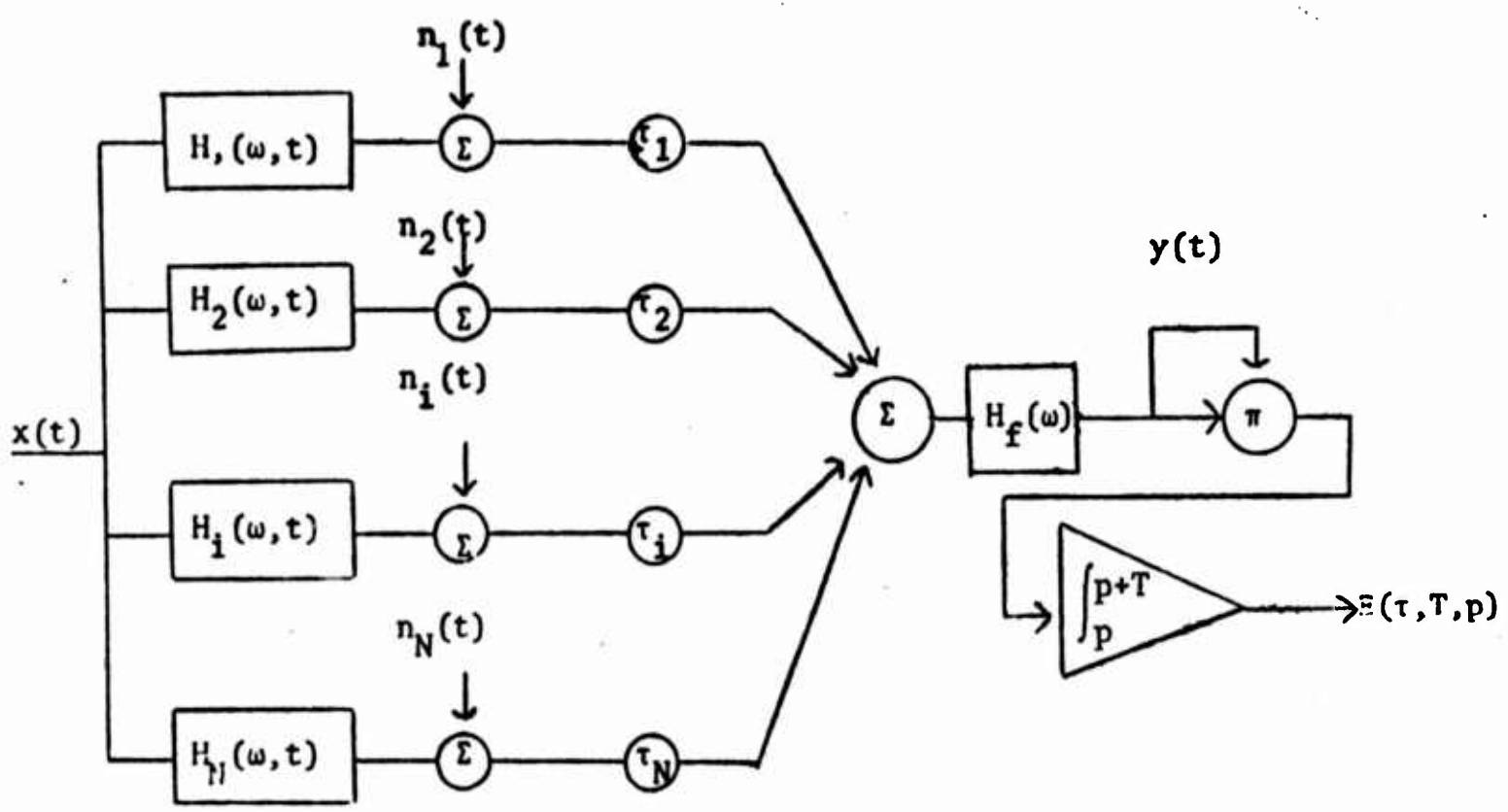


Figure J.1-1

We consider the receivers to be uniformly spaced in a linear array with steering delays τ_i that are multiples of a common steering parameter τ:

$$\tau_i = (i-1)\tau \quad \left[i=1 \ldots,N\right] \qquad \text{(J.1-1)}$$

The filtered output of the summer is thus

$$y(t) = \sum_{i=1}^{N} \left\{ \frac{1}{2\pi} \int_{\omega} \left[H_f(\omega) dz_{n_i}(\omega) + H_{i_f}(\omega, t-\tau_i) dz_x(\omega) \right] e^{-j\omega\tau_i} \right\} \tag{J.1-2}$$

One can readily show that

$$\overline{\Xi(\tau,T,p)} = \sum_{i=i}^{N} \sum_{k=1}^{N} \left\{ \frac{1}{2\pi} \int \left[|H_f(\omega)|^2 \, dz_{n_i n_k}(\omega) \right. \right.$$

$$\left. \left. + \Phi_{i_f k_f}(\omega, \omega, \tau_k - \tau_i) \, dz_{xx}(\omega) \right] e^{j\omega(\tau_k - \tau_i)} \right\} \tag{J.1-3}$$

Furthermore, the general second order moment $\Delta(\tau,\tau',T,\mu)$ is given by

$$\Delta(\tau,\tau',T,\mu) = \sum_{i=1}^{N} \sum_{k=1}^{N} \sum_{\ell=1}^{N} \sum_{m=1}^{N} \left\{ \left(\frac{1}{T}\right)^2 \left[\int_{\mu-T}^{\mu} \left[T+\nu-\mu \right] \right. \right.$$

$$\left. \left. \int_{\mu}^{\mu+T} \left[T-\nu+\mu \right] \right] \right\} \left\{ M_{ik}(\nu-\tau_i+\tau'_k) \, M_{\ell m}(\nu-\tau_\ell+\tau'_m) \right.$$

$$+ M_{ik}(\nu-\tau_i+\tau'_k) \, K_{\ell m}(\nu-\tau_\ell+\tau'_m) + M_{\ell m}(\nu-\tau_\ell+\tau'_m) \, K_{ik}(\nu-\tau_i+\tau'_k)$$

$$+ M_{im}(\nu+\tau'_m-\tau_i) \, M_{\ell k}(\nu+\tau'_k-\tau_\ell) + M_{im}(\nu+\tau'_m-\tau_i) \, K_{\ell k}(\nu+\tau'_k-\tau_\ell)$$

$$+ K_{im}(\nu+\tau'_m-\tau_j) \, M_{\ell k}(\nu+\tau'_k-\tau_\ell) + J^{(1)}_{ik\ell m}(\tau_\ell-\tau_i, \tau_m-\tau'_k, \nu+\tau'_k-\tau_i) \tag{J.1-4}$$

$$+ J^{(2)}_{ik\ell m}(\tau_\ell-\tau_i, \tau'_m-\tau'_k, \nu+\tau'_k-\tau_i) + J^{(3)}_{ik\ell m}(\tau_\ell-\tau_i, \tau_m-\tau'_k, \nu+\tau_k-\tau_i) \Big\} \, d\nu$$

$$\left. - K_{i\ell}(\tau_\ell-\tau_i) \, K_{km}(\tau'_m-\tau'_k) \right\}$$

where τ and τ' are two steering parameters with

$$\tau_i = (i-1)\tau \qquad \tau'_k = (k-1)\tau'$$

and

$$M_{ik}(\nu) = \int_{\omega} e^{j\omega\nu} \; |H_f(\omega)|^2 \; s_{n_i n_k}(\omega) \; \frac{d\omega}{2\pi} \tag{J.1-5}$$

$$K_{ik}(\nu) = \int_{\omega} e^{j\omega\nu} \; \phi_{i_f k_f}(\omega,\omega,\nu) \; S_{xx}(\omega) \; \frac{d\omega}{2\pi} \tag{J.1-6}$$

$$J^{(1)}_{ik\ell m}(\nu,\tau,\tau') = \int_{\omega}\int_{\omega'} e^{j\left[\omega\nu+\omega'(\nu-\tau+\tau')\right]}$$

$$\phi^{(4)}_{i_f k_f \ell_f m_f}(\omega,\omega,\omega',\omega',\nu,\tau,\nu+\tau') \; S_{xx}(\omega)S_{xx}(\omega') \; \frac{d\omega}{2\pi} \; \frac{d\omega'}{2\pi} \tag{J.1-7}$$

$$J^{(2)}_{ik\ell m}(\nu,\tau,\tau') = \int_{\omega}\int_{\omega} e^{j\left[\omega\tau+\omega'\tau'\right]}$$

$$\phi^{(4)}_{i_f k_f \ell_f m_f}(\omega,-\omega',-\omega,\omega',\nu,\tau,\nu+\tau' \mid \; S_{xx}(\omega)S_{xx}(\omega') \; \frac{d\omega}{2\pi} \; \frac{d\omega'}{2\pi} \tag{J.1-8}$$

$$J^{(3)}_{ik\ell m}(\nu,\tau,\tau') = \int_{\omega}\int_{\omega'} e^{j\left[\omega(\nu+\tau')+\omega'(\nu-\tau)\right]}$$

$$\phi^{(4)}_{i_f k_f \ell_f m_f}(\omega,\omega',\omega',\omega,\nu,\tau,\nu+\tau') \; S_{xx}(\omega) \; S_{xx}(\omega') \; \frac{d\omega}{2\pi} \; \frac{d\omega'}{2\pi} \tag{J.1-9}$$

The fourth order moments $\phi^{(4)}$ are formed by anology with (3.5-4) by

$$\phi^{(4)}_{i,k,\ell,m}(\omega,\omega',\omega'',\omega''',\mu,\mu',\mu'') =$$

$$H_i(\omega,t) \; H_k(\omega',t-\mu)^* \; H_\ell(\omega'',t-\mu') \; H_m(\omega''',t-\mu'')^* \tag{J.1-10}$$

REFERENCES

(1) Rosenblatt, M., Time Series Analysis, Ch. 15, John Wiley & Sons, Inc., 1963.

(2) Pierce, J. N. and Stein, S., "Multiple Diversity with Nonindependent Fading", Proc. IRE., Vol. 48, p. 89, Jan. 1960.

(3) Morse, P. M. and Feshbach, H., Methods of Theoretical Physics, McGraw-Hill, 1953.

(4) Fock, V. A., Electromagnetic Diffraction and Propagation Problems, Pergamon Press, 1965.

(5) Lonquet-Higgins, M.S., "Statistical Analysis of a Random, Moving Surface", Phil. Trans. A, 249, pp. 321-387.

(6) Bello, P. A., "Charazterization of Randomly Time-Variant Linear Channels", Vol. IT-11, pp. 360-393, May 1963.

(7) Kingsbury, F. J., Ph.D. Dissertation to be presented to the Faculty of the Engineering and Applied Science Department of Yale University in 1969.

(8) Beckman P., and Spizzichino, The Scattering of Electromagnetic Waves from Rough Surfaces, Pergamon Press, 1963.

(9) Parkins, B. E., "Scattering from the Time-Varying Surface of the Ocean", JASA, Vol. 42, #6, p. 1262.

(10) Rayleigh, Lord, "One the Light Dispersed from Fine Lines Ruled Upon Reflecting Surfaces or Transmitted by Narrow Slits", Phil. Mag., Vol. 14, p. 350-359, 1907.

(11) Officer, C. B., Introduction to the Theory of Sound Transmission, McGraw-Hill, 1958.

(12) LaCase, E. O. and Tamarkin, P., "Underwater Sound Reflection from a Corrugated Surface", JAP, Vol. 27, p. 138, 1956.

(13) Beckman, P. and Spizzichino, A., op. cit., p. 476.

(14) Lysanov, Iu., P., "Bibliography on Surface Scattering", Sov. Phys.-Acoust., Vol. 4, 1, 1958.

(15) Rossi, B. B., Optics, Addison Wesley, 1957.

(16) Born, M. and Wolf, E., Principles of Optics, MacMillan, 1964.

(17) Rayleigh Lord, The Theory of Sound, 3rd ed., MacMillan, London, 1896.

(18) Beckmann, P. and Spizzichino, A., op. cit., p. 34.

(19) Kerker, M. (Ed), Electromagnetic Scattering, MacMillan, 1963.

(20) Rice, S.C., "Reflection of Electromagnetic Waves from Slightly Rough Surfaces", Comm. Pure Appl. Math, Vol. 4, pp.351-378,1951.

(21) Eckart, "The Scattering of Sound from the Sea Surface", JASA, Vol. 25, p. 566, May 1953.

(22) Rosi, B.B., op. cit., Chapter 1.

(23) Cramer, H., Mathematical Methods in Statistics, Princeton University Press, 1946.

(24) Proud,J.M., Beyer, R.T., and Tamarkin, P., "Reflection of Sound from Randomly Rough Surfaces", J. App. Phys., Vol.31, #3,p.543, March 1960.

(25) Morgans, W.R., "The Kirchhoff Formula Extended to Moving Surfaces", Phil. Mag, S7, Vol. 9, #55, p. 141, Jan. 1930.

(26) Baker, B.B. and Copson, E.T., Huygen's Principle, Oxford Univ. Press, 1950.

(27) Stratton, J. A., Electromagnetic Theory, McGraw-Hill, 1941.

(28) Clemmow, P.C., The Plane Wave Representation of Electromagnetic Fields, Pergamon, 1963.

(29) Uretsky, J., "Reflection of a Plane Sound Wave from a Sinusoidal Surface", JASA, Vol. 35, p. 1293, March 1963.

(30) Rice, op. cit.

(31) Marsh, H.W. and Schulkin, M. and Kneale, S.G., "Scattering of Underwater Sound by the Sea Surface", JASA, Vol. 28, #3, May 1956.

(32) Lord, G. and Murphy, S., "Scattering from a Sinusoidal Surface - A Direct Comparison of the Results of Marsh & Uretsky", JASA, Vol. 36, #8, p. 1958, Aug. 1964.

(33) Lippmann, B.A., "Anomolous Reflection from a Ruled Grating", J. Opt. Soc. Am., Vol. 43, p. 408, 1953.

(34) Meecham, W.C., "Fourier Transform Method for the Treatment of the Problem of Reflection of Radiation from Irregular Surfaces", JASA, Vol. 28, #3, May 1961.

(35) Mintzer, D., "Discussion of the Paper by C. Eckart on Sea Surface Scattering", JASA, Vol. 25, p. 1015, July 1953.

(36) Cunningham, W.J., Introduction to Nonlinear Analysis, McGraw-Hill, 1958.

(37) Born, M., and Wolf, E., op. cit.

(38) Gulin, E. P., "Amplitude and Phase Fluctuations of a Sound Wave Reflected from a Sinusoidal Surface", Sov. Phys.-Acoust., Vol. 8, #3, January-March, 1963.

(39) Beckmann, P., and Spizzichino, A., op. cit., p.14.

(40) Levine, H., and Schwinger, J., "On the Theory of Diffraction by an Aperture in an Infinite Plane Screen", Phys. Rev., Vol. 74, p. 958, Oct. 1958.

(41) DuCastel, F., Tropospheric Radiowave Propagation Beyond the Horizon, p.37, Pergamon, 1966.

(42) Shilov, G.E., and Gurevich, B.L., Integral, Measure, and Derivative, a Unified Approach, p. 61, Prentice Hall, 1966.

(43) Kolmogorov, A.,"Interpolation und Extrapolation von stationaren zufalligen Folgen, Bull. Acad. Sci. URSS Ser. Math, 5.3-14, 1941.

(44) Doob, J.L., Stochastic Processes, p. 99, John Wiley, 1962.

(45) Yaglom, A.M., An Introduction to the Theory of Stationary Random Functions, Prentice Hall, 1962.

(46) Cramer, H., "On Harmonic Analysis in Certain Function Spaces", Arkiv. Math, Ast., Fys., Vol. 28, #12.

(47) Whittle, P., Hypothesis Testing in Time Series Analysis, Almquist and Wiksells AB, Upsalla, Sweden, 1951.

(49) Lanning, J., and Battin, R., Random Processes in Automatic Control, p. 82, sec. 2.13, McGraw Hill, 1956.

(50) Ellinthorpe, A.W., and Nuttal, A.H.,"Theoretical and Empirical Results on the Characterization of Undersea Acoustic Channels", Conf. Record, 1st Annual IEEE Communications Convention, 1965.

(51) Lindenlaub, C., "Response Functions of Linear Time-Varying Systems", MIT Lincoln Lab Tech Rept. 47G-2, 9 Sept. 1963.

(52) Batchelor, G.K., The Theory of Homogeneous Turbulence, Camb. Univ. Press, 1956.

(53) Lord, G., "Variance of the Bearing of a Discrete Linear-Difference Array in a Random Transmission Medium", JASA, Vol. 36, #8, Aug. 1964.

(54) Davenport, W.B., and Root, W.L., An Introduction to the Theory of Random Signals and Noise, McGraw Hill, 1958.

(55) Lanning J., and Battin R., op. cit., p. 161, sec.4.3.

(56) Bendat, J.S., Principles and Applications of Random Noise Theory, sec. 7.2, John Wiley, 1958.

(57) Usher, T., "Bearing Uncertainty For Split-Beam SONAR Arrays Due to Scattering in a Random Medium", Prog. Rept. 13, General Dynamics/Electric Boat Research (53 00 10 0231) Yale University, March 1964.

(58) Morse, P.M., and Ingard, K.U., Theoretical Acoustics, McGraw-Hill, 1968.

(59) DeLorenzo, J.D., and Cassedy, E.S., "A Study of the Mechanism of Sea Surface Scattering", IEEE, Vol. AP-14, #5, Sept.1966.

(60) Van Trees, H.L., Detection Estimation and Modulation, Part 1, Chapter 2, John Wiley, 1968.

(61) R.A. McDonald, "Minimum Variance Estimation of Relative Delay Between two Random Signals in Noise", Prog. Rept. #12, General Dynamics/Electric Boat Research (53 00 10 0231), Yale University, March 1964.

(62) Graybill, F.A., An Introduction to Linear Statistical Models, Vol. 1, McGraw-Hill, 1961.

(63) Miller, K.S., Multidimensional Gaussian Distributions, John Wiley, 1964.

(64) Wade, W.D., and Hyde, D.W., "Optimum Diversity Systems for the Detection of Randomly Progating Ocean Acoustic Signals", Final Research Report for Bell Telephone Laboratories, Department of Elctrical Engineering, University of Iowa, May, 1968.

(65) Kendall, M.G., "Rank and Product-Moment Correlation", Biometrica, Vol. 36, p. 177, 1949.

(66) Appel, P., and Kampe de Feriet, Fonctions Hypersheriques et Hyperspheriques Polynomials d' Hermite, Gauthier-Villars, 1926.

(67) Kendall, M.G., and Stuart, A., The Advanced Theory of Statistics, Vol. 1, Hafner, 1963.

(68) Freeman, H., Introduction to Statistical Inference, Addison Wesley, 1963.

(69) Middleton, D., Introduction to Statistical Communication Theory, Chapter 18, McGraw Hill, 1960.

(70) Price, R., "Statistical Theory Applied to Communication Through Multipath Disturbances", Tech. Rept. #34, MIT, Lincoln Lab, Sept. 1953.

(71) Price R., and Green, P.E., "A Communication Technique for Multipath Channels, Proc. IRE, Nov. 27, 1957.

(72) Turin, G., Communication Through Noisy, Random Multipath Channels, Tech. Rept. #116, MIT Lincoln Lab, May 1956.

(73) Erdelyi, A. (Ed.), Higher Trancendental Functions, Vol. 2 p. 192, McGraw Hill, 1953.

(74) Gulin, E.P., "Amplitude and Phase Fluctuations of a Sound Wave Reflected from a Sinusoidal Surface", Sov. Phys.-Acoust., Vol 8, #3, January-March 1963.

(75) Gulin, E.P., "The Correlation of Amplitude and Phase Fluctuations in Sound Waves Reflected from a Statistically Irregular Surface", Sov. Phys.-Acoust., Vol. 8, p. 135, Jan.-March 1963.

(76) Gulin, E.P., "The Correlation of Amplitude and Phase Fluctuations in Sound Waves Reflected from a Statistically Rough Surface", Sov. Phys.-Acoustics, Vol. 8, #4, p.335, April-June 1963.

(77) Whittaker, E.T., and Watson, G.N., A Course of Modern Analysis, (4-th edit.), Cambridge Univ. Press, 1958.

(78) Watson, G.N., Theory of the Bessel Function, Cambridge Univ. Press, (2-nd ed.) 1966-

(79) Erdelyi, A. (Ed.), Tables of Integral Transforms, Vol. 2, McGraw Hill, 1953.

(80) Slater, L., Confluent Hypergeometric Functions, Cambridge Univ. Press, 1960.

(81) Slater, L., Generalized Hypergeometric Functions, p. 17, sec. 1.5, Cambridge Univ. Press, 1966.

(82) Ibid., p. 137, sec. 4.6.2 .

(83) Gradshteyn, I.S., and Ryzhik, I.M., Tables of Integrals, Series, and Products, Academic Press, 1965.

(84) Abramowitz, M., and Stegun, I.A., Handbook of Mathematical Functions, National Bureau of Standards, AMS #55, U.S. Gov. Printing Office, 1964.

(85) Omura, J., and Kailath, T., Some Useful Probability Distributions, Tech. Rept. 7050-6, Stanford Electronics Lab, Sept. 1965.

(86) Hastings, C., Approximations for Digital Computers, Princeton U. Press, 1955.

(87) Kuo, S.S., Numerical Methods and Computers, Addison Wesley, 1965.

(88) Copson, E.T., Asymptotic Expansions, Camb. Math. Tract #55, Cambridge U. Press, 1965.

(89) Jones, D.S., and Klein, M., "Method of Stationary Phase Applied to Double Integrals", J. Math. Phys., 37, 1-28, 1958.

(90) Jolley, L.B.W., Summation of Series, (2-nd ed.), Dover, 1961.

(91) Lancaster, P., Theory of Matrices, Academic Press, 1969.

UNCLASSIFIED
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)	2a. REPORT SECURITY CLASSIFICATION
General Dynamics Corporation Electric Boat division Groton, Connecticut	Unclassified 2b. GROUP

3. REPORT TITLE

PROCESSING OF DATA FROM SONAR SYSTEMS, VOLUME VIII

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

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

Guy W. Beakley and John F. McDonald

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
October 15, 1970	499	-

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
N00014-68-C-0392 b. PROJECT NO.	U417-70-065
c. d.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Each transmittal of this document outside the agencies of the U.S. Government must have prior approval of the Office of Naval Research (Code 466).

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
Reproduction in whole or in part is permitted for any purpose of the United States Government.	Office of Naval Research Washington, D. C.

13. ABSTRACT

Volume VIII deals with the application of distribution-free tolerance regions to pattern recognition and passive detection and tracking using surface-scattered signals.

Pattern recognition is needed to identify sonar signatures as to the type of target by which they are generated. Distribution-free methods are desirable since the applicable probability distributions are frequently unknown, and it is desirable to establish some upper bound on at least the expected false-alarm probability. The recognition method developed has actually been applied to the recognition of speech waveforms since these are more easily obtainable than sonar signatures yet possess some of the same characteristics.

Signals reflected from the sea surface undergo distortion which limits their detectability and usability for tracking. Two propagation geometries are analysed. The first deals with the crosscorrelation of surface-reflected and direct transmission paths, and the second with the crosscorrelation of surface-scattered signals received at two different locations. This analysis assumes that the signal generated at the target and the background noise are both gaussian random variables. Three models of the scattering mechanism are proposed and two are analysed in detail. In all cases the correlator output is shown to exhibit persistent fluctuations due to the scattering. Results are presented for low pass signal spectra and are investigated as a function of bandwidth.