1

49471

0-315-0164-0

National Library
of CanadaBibliothèque nationale
du CanadaCANADIAN THESES
ON MICROFICHETHÈSES CANADIENNES
SUR MICROFICHE

NAME OF AUTHOR/NOM DE L'AUTEUR John Sterling Bird

TITLE OF THESIS/TITRE DE LA THÈSE The Application of Circular Symmetric Signal Theory to the
Detection of Communications and Radar Signals in Noise and
Clutter.

UNIVERSITY/UNIVERSITÉ Carleton University

DEGREE FOR WHICH THESIS WAS PRESENTED/ Doctor of Philosophy (Electrical Engineering)
GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE

YEAR THIS DEGREE CONFERRED/ANNÉE D'OBTENTION DE CE DEGRÉ 1980

NAME OF SUPERVISOR/NOM DU DIRECTEUR DE THÈSE Professor D.A. George

Permission is hereby granted to the NATIONAL LIBRARY OF
CANADA to microfilm this thesis and to lend or sell copies
of the film.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈ-
QUE NATIONALE DU CANADA de microfilmer cette thèse et
de prêter ou de vendre des exemplaires du film.

The author reserves other publication rights, and neither the
thesis nor extensive extracts from it may be printed or other-
wise reproduced without the author's written permission.

L'auteur se réserve les autres droits de publication; ni la
thèse ni de longs extraits de celle-ci ne doivent être imprimés
ou autrement reproduits sans l'autorisation écrite de l'auteur.

DATED/DATE

May 16, 1980

SIGNED/SIGNE

John S Bird

PERMANENT ADDRESS/RÉSIDENTIE FIXE

655 Walkley Road, Apt. 44Ottawa, OntarioK1V 9P1

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

**THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED**

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

**LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE**

THE APPLICATION OF CIRCULAR SYMMETRIC SIGNAL THEORY
TO
THE DETECTION OF COMMUNICATIONS AND RADAR SIGNALS
IN
NOISE AND CLUTTER

by

JOHN STERLING BIRD

A thesis submitted to the Faculty of Graduate Studies
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy.

Carleton University

Ottawa, Ontario

March 27, 1980

The undersigned recommend to the Faculty of Graduate
Studies acceptance of the thesis

"The Application of Circular Symmetric Signal Theory to
the Detection of Communications and Radar Signals in
Noise and Clutter,"

submitted by John Sterling Bird in partial fulfilment of the
requirements for the degree of Doctor of Philosophy.

Thesis Supervisor

Chairman, Department of Systems Engineering

Carleton University

Ottawa, Ontario

May 15, 1980

ABSTRACT

This thesis has two parts. Part 1 develops a general approach to detection problems that are encountered in communications, radar and sonar. The concepts of complex signal notation, phase plane diagrams and circular symmetric random vectors are merged to provide a conceptually simple and mathematically concise approach. The mathematical expressions for various system performance probabilities are in the form of Bessel integrals which in the past have been difficult to solve. An efficient and easily implemented technique for solving these integrals using the Fourier-Bessel series is presented. Several examples of system performance calculations are given to illustrate the utility of the general approach. These include calculating the probability of error for CPSK and multi-level ASK communications systems operating in cochannel interference and Gaussian noise. It is shown how the Rosenbaum and Glave upper bounds for the probability of error of M-ary CPSK operating in peak limited interference and Gaussian noise can be easily calculated. In addition, a classical radar/sonar detection problem requiring the calculation of the notorious radar Q-function is solved. In summary, this approach provides a general framework by which many detection problems may be solved.

Part 2 deals with the important problem of ground clutter in airspace surveillance radars. It is shown how new advances in

technology, specifically the phased array antenna and digital signal processing techniques, can be applied to deal effectively with ground clutter. Signal models for targets, clutter and noise are derived which are consistent with the step-scan of a phased array antenna rather than the continuous scan of a conventional "dish" antenna. A receiver structure is developed that is simple and essentially optimum for the airspace surveillance application. The conventional techniques for suppressing ground clutter, variable thresholding and pulse cancellation, are discussed in the light of the signal models and receiver structure. A new technique that takes advantage of the changed nature of the signals due to the step-scan is then discussed. This technique consists of building a coherent clutter map of the ground clutter environment using a recursive estimator and then using this map to remove the clutter from future signals. This new technique is compared to the conventional ones by means of a performance measure called the signal-to-noise ratio penalty. This performance measure reflects the scheme's target detection ability with reference to the optimum detector. It is found that the coherent clutter map provides essentially optimum performance for all target Doppler frequencies, outperforming the conventional clutter suppression techniques by a significant margin. It is recommended that experimental work be done to verify the conclusions of this feasibility study.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. D.A. George for suggesting radar detection theory as an area of study and for his guidance throughout the preparation of this thesis. In addition to this academic support Dr. George provided much-appreciated financial assistance.

I thank my wife, Lois, and children, Kevin and Karen, for their constant encouragement throughout the period of study. In addition, I thank my wife for her skillful typing of the final version of this thesis.

Finally, I thank the Public Service of Canada and specifically the Department of National Defense for granting the educational leave and financial support that made this endeavor possible.

TABLE OF CONTENTS

	Page
CHAPTER 1 INTRODUCTION	1.1
1.1 Overview	1.1
1.1.1 Radar, Why More Research	1.2
1.1.2 A Unified Approach to Detection Problems	1.3
1.1.3 Organization of the Thesis	1.5
1.2 Radar and The Ground Clutter Problem	1.6
1.2.1 Radar Principles	1.7
1.2.2 Clutter, A Basic Problem	1.14
1.2.3 Ground Clutter Suppression	1.23
1.2.3.1 Doppler Separation	1.26
1.2.3.2 Variable Thresholding	1.31
1.2.4 A Coherent Clutter Map - A New Concept	1.35
 PART 1 UNIFIED APPROACH TO DETECTION PROBLEMS	
CHAPTER 2 SYSTEM, SIGNALS AND PROBABILITY DISTRIBUTIONS	2.1
2.1 System and Signals	2.1
2.2 Spherical Symmetric Random Vectors	2.9
2.2.1 Probability Distributions and Functions	2.9
2.2.2 The Characteristic Function and Related Transform Pairs	2.10
2.2.3 The Significance of a Dimension Change	2.14
2.2.4 Properties of the Second Moment	2.16
2.2.5 Displaced Spherical Symmetric Distributions	2.17
2.3 Chapter Summary	2.21
2.4 Mathematical Details	2.22
2.4.1 Displaced Spherical Symmetric Distributions	2.22

	Page
CHAPTER 3 CALCULATION TECHNIQUE	3.1
3.1 The Problem and an Introduction to a New Solution	3.1
3.1.1 The Problem	3.1
3.1.2 Traditional Solutions	3.2
3.1.3 Proposed Solution	3.3
3.2 The New Technique	3.5
3.2.1 Description	3.5
3.2.2 Error Analysis	3.8
3.3 Mathematical Details	3.14
3.3.1 The Fourier-Bessel Series	3.14
3.3.2 Bound on $Q(R)$	3.16
CHAPTER 4 EXAMPLES	4.1
4.1 Calculation of Distribution and Probability Functions	4.1
4.1.1 Gaussian Signals	4.2
4.1.2 Constant Amplitude Random Phase Vector	4.11
4.1.3 Gaussian Plus N Constant Amplitude Random Phase Vectors	4.19
4.2 Performance Calculations	
4.2.1 Probability of Error for Phase Shift Keyed Communications Systems	4.30
4.2.2 Probability of Error for Amplitude Shift Keyed Communications Systems	4.42
4.2.3 A Classical Radar/Sonar Detection Problem	4.51
4.3 Mathematical Details	4.56
4.3.1 Bounding of e_{t2}	4.56
4.3.2 Probability of Error for Multi-level ASK	4.58
4.3.3 The Radar Q-function	4.58
CHAPTER 5 CONCLUSIONS TO PART 1	5.1

	Page
PART 2	
GROUND CLUTTER IN AIRSPACE SURVEILLANCE RADARS AND THE COHERENT CLUTTER MAP	
CHAPTER 6	
SIGNAL MODELS, RECEIVER STRUCTURES AND DECISION PROCESSES	6.1
6.1 Signal Models	6.1
6.1.1 Target and Clutter Signal Model	6.3
6.1.2 Noise Signal Model	6.8
6.2 Receiver Structure	6.9
6.2.1 Detection of a Single Point Target in Gaussian Noise	6.10
6.2.2 Simplified Receiver Structure	6.14
6.2.3 Response to Complicated Targets and Clutter	6.20
6.3 The Decision Process	6.22
6.3.1 Doppler Processing	6.23
6.3.2 Decision Making	6.27
6.3.3 False Alarms	6.30
6.3.4 A Gaussian Target in Gaussian Noise	6.37
6.4 Chapter Highlights	6.43
CHAPTER 7	
THE GROUND CLUTTER PROBLEM: MODELS AND THE OPTIMUM SOLUTION	7.1
7.1 Ground Clutter, Target and Noise Models	7.1
7.1.1 Ground Clutter	7.2
7.1.2 Target Model	7.3
7.1.3 Noise	7.5
7.2 The Detection Problem and the Optimum Solution	7.5
7.2.1 Single Return	7.6
7.2.2 Multiple Returns	7.14
7.3 Highlights of Chapter 7	7.20

	Page
7.4 Mathematical Details	7.21
7.4.1 Optimum Detector: General Target and Noise Correlation	7.22
7.4.2 Optimum Detector for a Perfectly Correlated Target in Uncorrelated Gaussian Noise	7.25
CHAPTER 8 CONVENTIONAL GROUND CLUTTER SUPPRESSION SCHEMES	8.1
8.1 Variable Thresholding	8.2
8.1.1 System Model	8.7
8.1.2 Performance	8.11
8.2 Pulse Cancellers	8.25
8.2.1 Mathematical Form of Pulse Cancellers	8.28
8.2.2 Performance of Canceller Scheme	8.33
8.3 Chapter Highlights	8.39
8.4 Mathematical Details	8.39
8.4.1 Limiting Value of the SNR Penalty	8.40
8.4.2 Upper and Lower Bounds for T_A	8.43
CHAPTER 9 THE COHERENT CLUTTER MAP	9.1
9.1 The Concept and Principle of Operation	9.1
9.2 Performance	9.9
9.3 Chapter Highlights	9.16
9.4 Mathematical Details	9.17
9.4.1 Determining σ_E^2	9.18
CHAPTER 10 CONCLUSIONS TO PART 2	10.1
10.1 Need and Objective	10.1
10.2 Summary and Comparisons	10.2
10.2.1 Summary of Models	10.2
10.2.2 Comparison	10.4
10.3 Practical Considerations	10.7
10.3.1 Radar System Considerations	10.8
10.3.2 The Study and Recommendations	10.10

	Page
CHAPTER 11 THESIS CONCLUSIONS	11.1
REFERENCES	R.1
APPENDIX A NONCOHERENT INTEGRATION	A.1
A.1 The Detectors	A.2
A.2 Fluctuation Statistics	A.3
A.3 A Special Case	A.5

LIST OF TABLES

Two tables appear in the text and are referred to often.

For ease of future reference the page numbers of these tables are given here.

Table 2.1

Page 2.12

Table 3.1

Page 3.7

NOTATION

complex quantity: denoted by a bar over the symbol e.g., $\bar{y}$

vector: denoted by a bar under the symbol e.g., $\underline{y}$

matrix: denoted by a bold face symbol e.g., $\mathbf{A}$ instead of A

Commonly used mathematical functions

$\cos(\cdot)$ trigonometric cosine

$\sin(\cdot)$ trigonometric sine

$\ln(\cdot)$ natural logarithm (base e)

$\exp(\cdot)$ exponential (base e)

$J_n(\cdot)$ ordinary Bessel function of order n

$I_n(\cdot)$ modified Bessel function of order n

$\Gamma(\cdot)$ Gamma function

CHAPTER 1

INTRODUCTION

This thesis initially was intended to deal with a specific problem in radar target detection. During the course of the research, however, I realized that the approach being taken was applicable to many detection problems encountered in communications, radar and sonar. In particular, systems performance evaluations can be handled efficiently with a common method. Therefore the thesis is arranged in an unusual manner. It consists of two parts: Part 1 deals with the general detection problem and the unifying approach, and Part 2 deals with the initial radar problem.

This first chapter provides an overview of the thesis contents and organization, as well as the background required for an appreciation of the radar problem.

1.1 Overview

This section summarizes the contents and organization of the thesis. It is divided into three subsections. The first gives a brief description of the radar problem and the reason for its importance. The second subsection introduces the general detection problem and comments on the significance of the solution. The final subsection addresses the organization of the thesis by outlining the major results and guiding the reader to the appropriate locations.

1.1.1 Radar, Why More Research

It would seem that a subject with as long a history as radar would be exhausted. This is not the case however. New technological developments continue to remove former constraints on radar designs. In this thesis the specific application of radar to airspace surveillance is considered. This is perhaps the single most important use of radar in North America (Muehe et al. 1974). In discussing airspace surveillance radars Muehe comments on the problems involved in automating such systems: "The principle difficulties are either an excessive number of false alarms which overload the automatic computers or else aircraft detection so spotty that the computer cannot make long continuous tracks." Continuing he states the reason for these difficulties: "The poor performance previously mentioned is due to the competition of aircraft radar returns with so-called 'clutter' returns and the response of circuits used in the radar to overcome these clutter returns." Having identified the problem as clutter he goes on to say that, "By far the largest undesired radar reflections come from fixed objects on the ground."

In this thesis I address the problem of these clutter returns from the ground and how their effect may be minimized by the use of advanced radar and digital signal processing technology. The specific advances to be taken advantage of are the development of electronically steered phased array antennas

and the large memory capabilities of present digital systems. It will be shown that, in the light of these developments, a processing concept called a coherent clutter map is essentially optimum for the removal of ground clutter.

1.1.2 A Unified Approach to General Detection Problems

As a result of researching methods for minimizing ground clutter a method of approaching a large class of detection problems was developed. It finds application in diverse detection problems, from calculating the probability of error in communications systems to calculating the detection probability of classical and modern radar and sonar systems. The approach consists of bringing together concepts that have been used successfully in isolated areas and coupling the result with an efficient calculation technique. The phase plane diagram has long been a tool of communications researchers but has not been popular in the fields of radar and sonar. The reason for this is the assumption that radar and sonar signals are intrinsically noncoherent and therefore a phase representation is not helpful. Although the assumption is correct to a point, the conclusion is not. When coupled with the concept of two-dimensional circular symmetric probability distributions the phase plane diagram becomes a valuable tool in formulating the solutions to detection problems involving both coherent and noncoherent signals.

Circular symmetric probability distribution theory provides a conceptually simple and mathematically concise way of handling

many detection problems, but there is a difficulty in that the mathematical expressions for performance probabilities are not easy to evaluate. Typically these expressions contain a Bessel integral that is difficult to solve. The method traditionally employed uses an expansion involving the moments of the circular symmetric probability distribution. These expansions are almost invariably complicated and require the use of special functions, such as the generalized Laguerre polynomial and the Hermite polynomial (Goldman 1974, Rosenbaum 1970). As well, there are problems in numerically handling the expansions for some parameter values as terms can become very large and alternate in sign (Esposito and Wilson 1973, Goldman 1974). The methods used for terminating the series expansions are usually ad hoc with very little information given about the accuracy of the result. Finally, since the expansions require knowledge of the moments of the resulting circular symmetric probability distribution these moments must be determined, and that is not always an easy task (Rosenbaum 1970).

In this thesis I present a method by which most of the Bessel integrals derived from detection problems can be solved easily. It involves the use of the Fourier-Bessel series and the characteristic function of the distribution. The terms of the Fourier-Bessel series are well-behaved and in most cases the characteristic function is easily determined. Although this method is approximate the error can be made arbitrarily small. This calculation technique is conceptually simple, easy to

program, computationally efficient, and need not be tailored to the specific problem being solved.

1.1.3 Organization of the Thesis

As was noted, the main body of the thesis is divided into two parts. Part 1 deals with the general detection problem and the probability calculation technique. Several examples of probability calculations and communications systems performance calculations are given. Specifically, probabilities of error are found for PSK (phase shift keyed) and ASK (amplitude shift keyed) communications systems operating in the presence of cochannel interference. It is shown how the Rosenbaum and Glave (1974) upper bounds for the probability of error for M-ary PSK systems operating in peak limited interference can be found easily using the general approach. The difficult and apparently unsolved problem of calculating the probability of error for multi-level ASK communication systems operating in the presence of cochannel interference and Gaussian noise is also solved. Further, a fundamental radar/sonar detection problem is described and a simple example given.

Part 2 begins by providing the mathematical framework for the radar detection problem, and building the specific model of interest. The conventional schemes for clutter suppression are then compared with each other and finally with the proposed scheme, the coherent clutter map. These comparisons are facilitated by a novel technique which illustrates the strengths

and weaknesses of each scheme virtually at a glance. During the course of this development it is necessary to compute the radar Q-function to a high degree of accuracy, for which the calculation technique of Part 1 is utilized.

To further aid the reader in gaining an appreciation of the topics discussed, a section entitled "Mathematical Details" has been added to some chapters. This permits more detail in the mathematical developments than would be normal if they appeared in the discussion. It also permits the mathematical details to be set aside for later study.

Those readers who are only interested in the general approach to detection problems should now proceed to Part 1 of the thesis. Those interested in the radar problem should continue to section 1.2 of this chapter and then proceed to Part 2 of the thesis. If interested in both, the reader should follow the natural order of the thesis.

1.2 Radar and the Ground Clutter Problem

This section provides both an introduction to radar and a qualitative description of the specific radar problem to be addressed in detail in Part 2. The radar introduction assumes that the reader is familiar with the basic signal processing concepts that are presented in first level graduate courses. The discussion proceeds from the fundamentals to a description of the problem, ground clutter in airspace surveillance radars. Various conventional schemes for suppressing ground clutter are

then discussed. Finally the new suppression scheme, the coherent clutter map, is introduced.

1.2.1 Radar Principles

Radar is an electronic system that detects the presence and determines the location of distant objects. It does this by transmitting a radio frequency signal modulated in some manner and detecting the returned (or echoed) signal. Radar extends the capability of man's senses to observe his environment, particularly the sense of vision. Although it does not have the resolution or color perception of the eye it does have the capability of "seeing" in conditions that the eye cannot, for example darkness, fog, snow, and rain. Radar also has the advantage of being able to measure the distance, or range, of the reflecting object. This is probably its most important attribute.

Figure 1.1 illustrates a simple form of radar. It consists of a transmitter, a transmit antenna that emits electromagnetic energy, a receive antenna, and a receiver. A portion of the signal that illuminates the target is reradiated in all directions and part of this is collected by the receive antenna. It is this very weak received signal that is then processed to extract target information. Both the range and bearing of the target can be measured by most radars.

figure 1.1 a simple form of radar.

The range of the target is determined by measuring the time delay between transmitting the energy and receiving the echo. Figure 1.2 illustrates the timing of a pulse radar. At time $t=0$ a signal of duration T is transmitted by the radar. Another signal is transmitted Δ seconds later where Δ is the interpulse period (IPP) and $1/\Delta$ is the pulse repetition frequency (PRF). The receiver is gated to receive energy during a portion of the interpulse period. This is called the receive window. Since some form of pulse compression (matched filtering) is usually used, the range window is shorter than the receive window by T , the length of the signal pulse. If our target of interest lies between ranges $R_{\min}$ and $R_{\max}$ the range window is defined by

$$t_{\min} = \frac{2 R_{\min}}{c} \quad (1.1)$$

and

$$t_{\max} = \frac{2 R_{\max}}{c} \quad (1.2)$$

where c is the speed of propagation of electromagnetic waves (3×10^8 m/s).

The other parameter measured is the direction of the target as determined from the direction of the wavefront of the received signal. This is usually measured by determining the pointing angle of a highly directional receive antenna when the target echo was actually received. Beamwidths of modern reflector type antennas used for radar vary from a fraction of a degree to about 10 degrees depending on the size of the antenna and the frequency

figure 1.2 timing for a pulse radar.

of operation.

In some radars, the relative velocity between the radar and the target may be determined from the shift in carrier frequency that takes place because of the velocity difference. This Doppler effect causes the signal reflected from the moving target to be shifted in frequency by an amount f_D .

$$f_D = \frac{2 v_r}{\lambda} \quad (1.3)$$

where v_r is the relative (or radial) velocity between the target and the radar (+ve towards the radar and -ve away from the radar) and λ is the wavelength at the carrier radio frequency.

Some radars do not use pulsed signals but transmit continuously and use separate transmit and receive antennas. These radars cannot be used to obtain range information but have been used very successfully for determining target speed.

The operating frequencies of radars cover a band from about 25 MHz to 70 GHz, a spread of more than 11 octaves. Generally, high resolution radars operate at the higher frequencies and low resolution radars at the lower frequencies. Because high resolution requires narrow beamwidths and since beamwidth is inversely proportional to antenna width, measured in wavelengths, high frequencies are required so that the antennas are of manageable size. For long range search radars power is of prime importance and, since it is easier to obtain high power at lower frequencies, these types of radars generally operate in the lower frequency bands.

This thesis is concerned mainly with radars used for airspace surveillance. Surveillance radars are usually large stationary installations which detect and track aircraft. Since range, direction, and target velocity are the primary target features to be determined, pulsed radars with a facility for determining Doppler information are used for this application.

Figure 1.3 illustrates the basic configuration of such a system. The transmitter produces the radio frequency (RF) signal that is to be broadcast as well as timing and phase information necessary for processing the received echo. The duplexer gates the antenna so that it is either connected to the receiver or the transmitter. The antenna is usually a mechanically rotated antenna with a beamwidth in the order of 1 to 7 degrees in azimuth. The RF and IF (intermediate frequency) amplifiers increase the power level of the returned signal and the synchronous demodulator decomposes it into its in-phase (I) and quadrature (Q) components. A filter matched to the incoming signal with the desired Doppler shift then increases the signal-to-noise ratio (SNR). Next, the signal is passed to the detector where the vector magnitude of the I and Q channels is found, and then displayed. The display receives a signal that is proportional to the target and noise energy in the return and changes with time. The noise is due to the standard thermal, cosmic and other sources that affect all radio receivers.

figure 1.3 a block diagram of a simple pulsed radar

Two types of displays have been frequently used, the A-scope and the plan position indicator (PPI). The A-scope plots the received energy as a function of time as shown in figure 1.4a. The PPI introduces the angle variable and plots the energy as brightness on the screen as shown in figure 1.4b.

The decision on the presence or absence of a target depends on whether or not the energy level crosses a predetermined threshold as shown in figure 1.5. If the energy level crosses the threshold, a target is declared present at that range and azimuth. This threshold is determined from the probability distribution of the noise and the desired false alarm rate. A false alarm occurs whenever the energy level crosses the threshold due to noise alone. If the threshold is increased there will be fewer false alarms but weak targets will not be detected. If the threshold is decreased there will be more false alarms but weak targets will be more readily detected. This creates a trade-off situation where we exchange a low number of false alarms for target detectability.

Unfortunately, noise is not the only unwanted signal in a radar system. Echoes from objects not considered to be targets such as buildings, hills, rain drops and birds also interfere with target detection.

1.2.2 Clutter, A Basic Problem

Clutter is defined as a conglomeration of unwanted radar echoes (Definition of Terms, 1954; Proc. IRE). The term clutter

figure 1.4 radar displays: (a) A-scope, (b) plan position indicator (PPI).

figure 1.5 A-scope display of a detected target.

is very appropriate as these unwanted echoes do "clutter" a radar display and make target detection difficult. Unlike thermal noise, clutter energy increases as the output power of the radar increases; therefore increasing the radar output will not solve the clutter problem.

What is to be classified as clutter depends upon the application of the radar. For example, with a radar used for terrain mapping, the ground is the target and an aircraft is clutter. In the case of airspace surveillance radars, aircraft, missiles and the like are targets whereas buildings, hills, rain, birds etc. are clutter.

Clutter is usually divided into two classes: isolated and composite. Isolated clutter is a return from things like a building or chimney that stand alone and provide an identifiable echo. Isolated clutter is not correlated in range and azimuth. In other words, the amount of energy received from ranges and azimuths close to that of the clutter object is not related to the amount of energy received from the clutter object. Isolated clutter may be the result of a stationary object as in the case of buildings or a moving object such as an automobile. Composite clutter is due to a collection of reflecting objects. Examples of composite clutter are groups of buildings, a rainstorm (many rain droplets) and trees in a forest. Composite clutter tends to be correlated over range and azimuth and may be the result of both stationary and moving objects.

In addition to attributing clutter to land, sea or weather, special names have been given to certain other types of clutter. Chaff consists of many small pieces of reflecting material, usually aluminum, deliberately released by an aircraft to simulate real targets and confuse military radar. Angel is the name given to almost any clutter source that is not readily explained, and investigations show that this term has been used to describe many types of reflectors including birds and atmospheric disturbances.

The main difficulty with clutter is that it competes with the target for detection. Figure 1.6 illustrates an aircraft flying towards a radar. In addition to the aircraft's echo an echo from the building is received; this is ground clutter. The fixed threshold, shown in figure 1.7a, must be large enough to prevent the building from being shown as a target, but this prevents the detection of the actual target. An obvious solution is to make the threshold a variable depending on range as shown in figure 1.7b. This method for handling clutter is known as variable thresholding.

In this example the aircraft is flying directly toward the radar so the Doppler shift of the aircraft return could be used to separate the target return from the clutter as shown in figure 1.8. Here the output of a detector preceded by a filter matched to the Doppler shift of the target is displayed. The return from the building is still visible as it is not possible in practice to design a perfect band reject filter. However, the smaller response to the building permits the threshold to

figure 1.6 illustration of ground clutter (building) and target.

figure 1.7 detection of target in ground clutter: (a) fixed threshold, (b) variable threshold.

figure 1.8 radar response to target for a filter matched to the target's Doppler frequency.

be reduced and the target detected. On the other hand, if the aircraft was flying at right angles to the direction of propagation of the radar pulse, there would be no relative velocity between the radar and the aircraft and hence no Doppler shift.

This leads to an important and difficult detection problem:

important because most aircraft at some point are tangential to the radar, and difficult because Doppler separation of target and clutter is not possible.

Perhaps the oldest method for handling ground clutter is to ignore detections that occur in areas of known clutter objects, but this has the drawback that targets in the same radar space location are also ignored.

In spite of its impact on target detection the intrinsic nature of ground clutter returns has not been studied in any detail. Irabu et al. (1977) comments on this lack of study noting: "The characteristics of backscattered signals from ground clutter are not yet well known." The reason for this appears to be that clutter measurements are very much a function of the radar used to make the measurements. Factors such as scan rate, pulse length, power and phase stability confuse the results. Of particular significance is whether or not the antenna is being continuously rotated. Virtually all literature on ground clutter is based on rotating antenna radars rather than the newer and less common stepped-scan radars. It is possible however, to glean some facts about ground clutter from studies conducted with other objectives in mind.

Muehe et al. (1974) make some interesting and relevant comments concerning ground clutter. Addressing the problem of ground clutter in radars using a phased array antenna (electronically stepped instead of scanned) they state:

"Its (ground clutter) natural or intrinsic spectrum is very narrow compared to the spectral spread caused by antenna scanning motion."

"Because the antenna is not scanned and ground clutter can be considered to be time-stationary (constant voltage) the clutter spectrum is just an impulse at zero frequency."

"It (ground clutter spectrum) is so narrow in fact that it is possible to completely separate the aircraft with slow velocities from the ground clutter".

From these comments we deduce that returns from stationary ground clutter are constant in both phase and amplitude. Further evidence for this conclusion is given by Muehe et al. (1974) when discussing the results of an experiment with a phased array radar. A simple d.c. filter was used to suppress ground clutter with the result that: "No ground clutter was seen." This confirms the expected nature of ground clutter.

Less direct evidence of the stability of ground clutter is available from other studies. Von Schlachta (1975) recognized the need for making more complete measurements of radar target returns, including phase stability, in order to properly predict the performance of phased array radars. His measurements show a remarkable phase stability in aircraft target returns that

stretches over tens of seconds. Extrapolating these results to stationary targets or clutter one would expect to observe similar long term stability in the returns.

Measurements of the amplitude of ground clutter returns made using noncoherent radar techniques also indicate the possibility of long term ground clutter stability. For example, Long (1975) in describing some of D.E. Kerr's work, comments on the use of the Rician distribution to describe certain types of ground clutter. This distribution is that of the envelope of a sine wave plus Gaussian noise with one of the parameters being the ratio of the sine wave power to noise power. From Kerr's measurements this power ratio was approximately 30 for rocky terrain. It is of interest that this same Rician distribution would have been measured whether the sine wave was phase stable or not (as is shown in section 2.2.5 of this thesis).

There is a definite need for more complete ground clutter measurements, including its phase stability, in order to determine the performance of the radars of the future. From the measurements that have been made for different purposes and from comments of researchers familiar with the problem we may, however, infer that ground clutter possesses long term amplitude and phase stability.

1.2.3 Ground Clutter Suppression

Numerous systems for suppressing ground clutter have evolved through the years. These systems were designed to

handle ground clutter returns received from a continuous scan antenna. However, the constraints imposed by the continuous scan severely limit the effectiveness of the clutter suppression systems. The two major causes for the limitation are limited dwell time and beampattern modulation.

Dwell time is that time that a target (or clutter element) is within the beamwidth of the antenna. If the antenna is continuously scanned a limited number of returns are received. The number of returns depends upon the beamwidth and the rate of rotation. Faster rotation or smaller beamwidths result in fewer returns. For target processing this limited dwell time affects the number of returns that can be processed coherently; target movements from scan to scan are significant thus destroying the possibility of coherently processing returns from different scans. For perfectly stationary ground clutter, on the other hand, continuous scan causes the return to be different for each pulse transmitted. To keep track of the ground clutter returns coherently over times beyond one scan would require that the mechanical scan mechanism be perfectly synchronized with the transmitter and that the returns for each pulse transmitted be stored. Both these requirements necessitate systems of excessive complexity and cost. Therefore practical radars usually confine coherent processing to the returns received during a single scan.

Beampattern modulation limits the performance of clutter suppression schemes by destroying the constant nature of the

clutter return. The maximum response for the antenna is usually on the beam axis. This response falls off to one-half its on-axis response at an angle (from the axis) of one-half the beamwidth. Therefore as the beam sweeps past a target or clutter element the response at first builds up and then tapers off, effectively amplitude modulating the returns. This amplitude modulation limits the coherent processing capabilities of the radar.

Since continuous scan antennas have been used for nearly all surveillance radars, except for the recent appearance of phased array radars, clutter suppression systems have been designed under the constraints imposed by such antennas. These systems can be grouped into two major classes: those that use Doppler separation of target and clutter to enhance target detection, and those that use a variable threshold. Doppler separation becomes possible when there is a radial velocity difference between the clutter source and the target. This translates into a difference in frequency between the two returns by an amount equal to the difference of their respective Doppler shifts. Doppler signal processing systems take advantage of this difference to enhance target detection. If there is no radial velocity difference or if the clutter is so strong as to be insufficiently suppressed by the Doppler processing, then a variable threshold is used.

These two techniques will now be discussed in more detail leaving the mathematical complexities until later.

1.2.3.1 Doppler Separation

As noted, a difference in radial velocities between clutter source and target may be used to enhance target detection. If we are interested in all returns that indicate a target with a radial velocity outside that band occupied by the clutter, the velocity filter shown in figure 1.9a may be devised. This is a clutter-filter system. Alternatively, if it is known that the target of interest has a radial velocity that is between certain limits, we may design a velocity filter that accepts only targets with these velocities as shown in figure 1.9b. This is a target-filter system.

Clutter-filter systems usually employ a transversal filter (they are transversal because a finite number of target returns are available) as shown in figure 1.10. There is one such filter for each of the in-phase and quadrature channels, and they are usually placed just after the synchronous detector. These filters are generally designed to reject stationary clutter sources, i.e., the tap gains satisfy

$$\sum_{i=0}^{N-1} a_i = 0 \quad (1.4)$$

This constraint provides a zero response to constant clutter returns. The effect of beampattern modulation is usually ignored.

The use of the transversal filter to reject clutter is very widespread. The filter is usually in the form of a two pulse canceller ($N=2$) or a three pulse canceller ($N=3$). Longer

figure 1.9 filter responses for separating clutter from targets: (a) clutter filter,
(b) target filter;

figure 1.10 a transversal filter used for clutter removal. Δ is the interpulse period and $a_0, a_1, \dots, a_{N-1}$ are the filter coefficients.

filters have not been widely used in operational radars as yet, and still remain very much an object of research. Part of the reason for not using them is the need for longer term clutter correlation and longer dwell times (more pulses) than are available with a continuously rotating antenna.

One of the main disadvantages of clutter reject filters, or pulse cancellers, is that their velocity response is periodic with a period equal to $2\lambda/\Delta$ as shown in figure 1.11. This gives rise to blind speeds. For example, a typical L-band radar operating at 1250 MHz and a PRF of 360 pulses per second would not be able to detect targets travelling at radial velocities close to $175n$ km/h, where $n = 0, \pm 1, \pm 2, \dots$. Also, targets which do not exhibit any Doppler shift (i.e., those flying at right angles to the radar beam) would be removed by the filter. Clutter filters however, do perform well when the target and clutter situation is suitable.

The target-filter system, by comparison, is one in which only a specific band of radial velocities is selected. The filter is usually matched to the expected radial velocity (Doppler frequency) of the target. In practice it is often implemented as a transversal filter similar to the clutter reject filter but with different filter coefficients.

Two problems with the target filter are "leakage" and zero radial velocity targets. Since in practice the filters are not ideal some clutter energy is picked up from strong clutter sources with radial velocities outside the band of

figure 1.11 an illustration of the periodic nature of the velocity response of a clutter filter. λ is the wavelength and Δ is the interpulse period.

interest. In particular, stationary clutter sources often have large returns which contaminate the target filter. Therefore this filter is often preceded by a pulse canceller to reduce the size of the stationary component of the return (Muehe et al. 1974). Given the possibility of large clutter sources at zero radial velocity, detection of targets at this velocity is a problem which the target filter cannot solve. To handle this, variable threshold techniques must be considered."

1.2.3.2 Variable Thresholding

When Doppler separation of the clutter and targets is not possible, variable threshold techniques prove useful in improving target detection. As was mentioned in section 1.2.1, the threshold is that predetermined energy level above which a target is declared present. Figure 1.12 illustrates the advantages to be gained with a variable threshold. In region "A" a target is not as detectable using a fixed threshold as it would be using a variable threshold because the fixed threshold is set to account for the large isolated clutter. In region "B", dominated by a rainstorm, the detectability is also enhanced by the use of a variable threshold.

Determining the variable threshold is not easy and much research continues on this subject. The techniques developed usually fall into one of two classes: those that use information from neighboring ranges and azimuths to set the threshold, and those that use information from the same range and azimuth at

figure 1.12 an illustration of how variable thresholding enhances target detection:
region A (low clutter) variable threshold low, region B (larger clutter)
variable threshold still lower than fixed threshold.

earlier times to set the threshold. The former can be referred to as area thresholding and the latter temporal thresholding.

Area thresholding as practiced suffers from several problems. It is detrimental with isolated clutter as neighboring areas do not contain relevant information. Also, a target in a neighboring area greatly reduces its effectiveness. Rickard and Dillard (1977) address this problem and make refinements in the concept to adjust for this difficulty.

Temporal thresholding has the advantage that, by using data from the range and azimuth of interest, it can handle isolated clutter and nearby targets. However, some means of determining when to stop using the past data must be developed in order that the threshold is not raised as the target appears. Blyth and Treciokas (1977) give some results using temporal thresholding to suppress clutter.

If the thresholds are being continually updated to adjust to variable clutter situations such as rainstorms, the term adaptive threshold may be used to describe both the area and temporal thresholding techniques.

The concept of a clutter map is closely related to variable thresholding. If the variable threshold was plotted as shown in figure 1.13a, it is evident that high thresholds mean large clutter returns and low thresholds mean small clutter. This data plotted as shown in figure 1.13b yields the clutter map concept. In practice the actual clutter levels are usually stored in map form and the thresholds derived from these.

figure 1.13 variable threshold: (a) A-scope display, (b) PPI display.

A crude form of clutter map simply blanks the spot on the PPI wherever the clutter energy exceeds a certain threshold. In this case clutter is considered to be anything that does not move from scan to scan. This is sometimes called area moving target indicator or area MTI. The obvious disadvantage here is that targets in the same area as the clutter are also blanked.

In all, the clutter map is a fairly effective method of combatting that clutter which cannot be separated from the target by Doppler processing. However, information that might be used to further improve detection is thrown away. In particular, clutter sources that are very stable such as buildings and rocky hills with little vegetation have very stable returns. Consequently, information about the clutter phase could be used to improve target detection. The next section introduces this prospect.

1.2.4 A Coherent Clutter Map - A New Concept

Clutter maps have been used successfully to enhance target detection in the presence of clutter when the clutter and target cannot be separated by Doppler processing. Potential phase information is, however, not utilized by such clutter maps. If a radar environment is absolutely stationary and the antenna is not rotating we would expect a certain return. A week later we would expect the same return. The point is that the stationary environment results in the same return time after time. Now if

superimposed upon this stationary environment is a radar environment in motion such as trees blown in the wind, rain, airplanes, cars, etc. we have a total radar environment made up of two components: a component that never changes and a variable component. If we could construct a coherent map of the stationary component it could be removed from the radar return and target detection would be enhanced.

The residue left after the stationary component of the radar environment had been removed could then be processed with the conventional systems presented earlier. These are designed to produce good results in nonstationary radar environments.

This coherent removal of clutter is not feasible with systems using continuous scan antennas as it requires that the clutter returns be coherent over many scans. However, in recent years phased array antennas have been introduced. These antennas do not rotate but steer their beams electronically, stepping from one direction to another in a matter of microseconds (Burgess 1977). The versatility of the phased array radar enables it to operate as a multifunction radar incorporating such tasks as target searching and active tracking in a single radar. The main benefit of the phased array, as far as clutter suppression is concerned, is that the clutter appears in its true nature. Therefore, stationary clutter possesses long term coherency.

Figure 1.14 illustrates the general system concept of coherent clutter removal as first proposed by George (1978).

C_I and C_Q are the in-phase and quadrature components of the

figure 1.14 coherent clutter removal. C_I and C_Q are the in-phase and quadrature components of the stationary clutter.

stationary part of the clutter for a given range and azimuth. They are stored in the form of a coherent clutter map and coherently subtracted from the signal. The residual signal is then processed to further enhance target detection and the final result is then appropriately displayed.

The advantages of such a system over a system that assumes and indeed creates a nonstationary environment are twofold. First, there are no blind speeds as only the stationary components of the clutter are removed. Second, detection of slow and stationary targets is enhanced due to the removal of the stationary clutter that prevents optimal detection. These advantages would greatly increase the ability of a radar to detect targets in stationary ground clutter, especially when targets are travelling at right angles to the antenna beam.

In practice the stationary component of the radar environment would not be known and would have to be estimated. Also, the stationary component might actually be pseudo-stationary in the sense that it could change with season, time of day, sun spot activity, etc. The important system factors are: how well the stationary component can be estimated, and whether the estimation can keep pace with a slowly varying environment. In essence such a system requires long term coherency in the clutter, and evidence of this was given in section 1.2.2.

The coherent clutter map requires a large memory to store clutter values for each range and azimuth of concern, but with

the recent and expected advances in solid state technology this memory requirement should not be a problem (Giaccari and Nucci 1979).

This concludes the qualitative discussion of the radar problem which is addressed in detail in Part 2 of this thesis. Readers interested in only the radar problem should proceed directly to Part 2. Those interested in the general approach to the detection problems as well as the radar problem should proceed to Part 1.

PART 1

UNIFIED APPROACH TO DETECTION PROBLEMS

PART 1

This part of the thesis, containing chapters 2 through 5, deals with the general detection problem. Chapter 2 describes the general form of this problem and then merges the concepts of complex signal notation, phase plane representations and circular symmetric probability theory to provide a framework for a solution. This method of approach has received limited attention in the past due to the difficulty of evaluating the resulting probability integrals, and chapter 3 presents a calculation technique that effectively removes this bottleneck. In chapter 4 several examples of the use of the approach and the calculation technique are given. These include the calculation of fundamental probability distributions and functions as well as the calculation of performance for several communications and radar/sonar systems. Chapter 5 concludes Part 1 by reviewing the concepts and major results.

Extensive use is made of complex signal notation throughout Part 1. The reader who is not familiar with this notation should read Van Trees (1971, appendix) or a like reference before continuing.

CHAPTER 2

SYSTEM, SIGNALS AND PROBABILITY DISTRIBUTIONS

This chapter presents the basic form of the detection problem along with a framework for solution. The building blocks for this framework are complex signal notation, the phase plane diagram and circular symmetric probability theory. I begin with a bandpass model which describes a wide class of signal detection systems. Next, it is shown that the output of such a model can be described conveniently by a phase plane diagram and, then, that many of the signals used to represent messages, targets, interference and noise have circular symmetric probability distributions in the phase plane. This naturally leads to the discussion of circular symmetric probability theory and the expression of probabilities in terms of this theory.

2.1 System and Signals

Whenever a signal to be detected is narrowband¹ and of finite duration, the usual processing model is a filter followed by a synchronous demodulator. This system is illustrated in figure 2.1. The bandpass filter when possible is a matched filter.

¹ The bandwidth of the signal is much less than its center frequency.

figure 2.1 block diagram of system used to model detection problems.

If the detection process is coherent both branches of the synchronous demodulator are not required, but I shall include both in the analysis because this is of conceptual assistance.

The outputs of the synchronous demodulator are sampled at some time t_0 (usually the maximum response time for the signal). These samples are then examined and a decision is made as to the presence or absence of the signal. We shall be concerned with the nature of these samples and the decision making process.

In complex envelope notation the output of the synchronous demodulator is $\bar{y}(t_0)$ where

$$\bar{y}(t_0) = y_R(t_0) - j y_I(t_0) \quad (2.1)$$

and the bar on y indicates that y is complex. Later I shall find it convenient to refer to the same quantity as a vector. In that case the bar will be moved to the bottom of the symbol. The vector components will be $y_R(t_0)$ and $y_I(t_0)$.

From the theory of narrowband signals and bandpass systems (Van Trees 1971, appendix) the output $\bar{y}(t_0)$ of the synchronous demodulator is

$$\bar{y}(t_0) = \int_{-\infty}^{\infty} \bar{H}(\omega) \bar{X}(\omega) \exp(j\omega t_0) \frac{d\omega}{2\pi} \quad (2.2)$$

where $\bar{H}(\omega)$ and $\bar{X}(\omega)$ are the Fourier transforms of $\bar{h}(t)$ and $\bar{x}(t)$ respectively. $\bar{h}(t)$ and $\bar{x}(t)$ in turn are the complex representations of the filter response $h(t)$ and the input signals $x(t)$ such that

$$h(t) = \operatorname{Re} (2 \bar{h}(t) \exp(j\omega_c t)) \quad (2.3)$$

and

$$x(t) = \operatorname{Re}(\sqrt{2} \bar{x}(t) \exp(j\omega_c t)) \quad (2.4)$$

where ω_c is the carrier frequency. Without loss of generality I shall assume that $t_0 = 0$ and, further, for convenience of notation the dependence on t_0 will be suppressed.

We may assume that the input signal $x(t)$ is in the following general form.

$$x(t) = \operatorname{Re}(\sqrt{2} b \exp(j\phi) \bar{f}(t) \exp(j\omega_c t)) \quad (2.5)$$

where b is the signal amplitude, ϕ the phase and $\bar{f}(t)$ the normalized complex shape.¹ Therefore

$$\bar{x}(t) = b \exp(j\phi) \bar{f}(t) \quad (2.6)$$

and

$$\bar{y} = b \exp(j\phi) \int_{-\infty}^{\infty} \bar{H}(\omega) \bar{F}(\omega) \frac{d\omega}{2\pi} \quad (2.7)$$

where $\bar{F}(\omega)$ is the Fourier transform of $\bar{f}(t)$.

The integral in equation 2.7 is simply a complex constant which may be written in the form $d \exp(j\beta)$. Therefore the output $\bar{y}$ becomes

$$\bar{y} = r \exp(j\gamma) \quad (2.8)$$

$$\text{where } r = bd \quad (2.9)$$

$$\text{and } \gamma = \phi + \beta. \quad (2.10)$$

¹ The complex signal shape is normalized in the sense that

$$\int_{-\infty}^{\infty} |\bar{f}(t)|^2 dt = \int_{-\infty}^{\infty} |\bar{F}(\omega)|^2 \frac{d\omega}{2\pi} = 1$$

where $\bar{F}(\omega)$ is the Fourier transform of $\bar{f}(t)$.

If the filter is matched¹ to $\bar{f}(t)$ then $d = 1$ and $\beta = 0$ and we have

$$\bar{y} = b \exp(j\phi) \quad (2.11)$$

The output $\bar{y}$ may be diagrammatically represented as shown in figure 2.2, and may also be represented in vector notation as

$$\underline{y} = \begin{bmatrix} y_R \\ y_I \end{bmatrix} = \begin{bmatrix} r \cos(\gamma) \\ r \sin(\gamma) \end{bmatrix} \quad (2.12)$$

If the random variables r and γ are independent and γ is uniformly distributed from $-\pi$ to π , then $\underline{y}$ is by definition a circular symmetric random vector (Goldman 1974). In other words, the two dimensional joint probability distribution of y_R and y_I is circular symmetric about the origin.

The two conditions for $\underline{y}$ to be a circular symmetric random vector are met by a large class of signals encountered in detection problems. For most interference, noise, and message signals the amplitude and phase are not related. Also the phase of interference and noise signals is seldom known and is usually modelled as being uniformly distributed from $-\pi$ to π . This is also true for non-coherent message signals.

As a simple example of this form of signal representation consider a signal $x(t)$ whose amplitude is constant and whose phase is uniformly random. For this signal r is constant and γ is a random variable distributed from $-\pi$ to π . Figure 2.3 illustrates how this signal may be represented diagrammatically. The circle with radius a represents the probability distribution

¹ matched meaning $\bar{H}(\omega) = \bar{F}^*(\omega)$ where $*$ indicates the complex conjugate.

figure 2.2 polar coordinate representation of signal.

figure 2.3 diagrammatic representations of a constant amplitude uniformly random phase signal: (a) phase plane representation, (b) three-dimensional representation.

shown in figure 2.3b. The distribution in functional form is written

$$f(y_R, y_I) = \frac{1}{2\pi a} \delta(\sqrt{y_R^2 + y_I^2} - a) \quad (2.13)$$

where $\delta(\cdot)$ is the delta function. In polar coordinates

$$f(r, \theta) = \frac{1}{2\pi a} \delta(r - a) \quad (2.14)$$

$$\text{where } r = \sqrt{y_R^2 + y_I^2} \quad (2.15)$$

Many detection problems can be visualized with the use of this method of representing random signals and their distributions. The phase plane approach permits one to qualitatively investigate a problem and easily make general statements about the solution. This process is demonstrated in chapter 4 when I consider examples of system performance evaluations.

From this discussion it seems that probability theory dealing with circular symmetric distributions would be helpful in solving detection problems. This is indeed the case. Specifically, distribution theory applied to both one and two-dimensional symmetrically distributed vectors can be utilized effectively. The one-dimensional distributions apply to coherent detection problems while the two-dimensional distributions apply to non-coherent detection problems. In the next section of this chapter the important concepts of spherical symmetric¹ distribution theory as related to signal detection problems are presented.

¹ Spherical symmetric refers to the general dimension case while circular symmetric is reserved for two-dimensions.

2.2 Spherical Symmetric Random Vectors

I demonstrated in the previous section how circular symmetric probability distributions arise naturally in signal detection problems. In this section I will present those aspects of the probability theory of the general spherical symmetric case which are important for solving detection problems. Many of the results to be presented come from Lord (1954). The mathematical details leading to these results will not be repeated here. However, the results will be given a heuristic justification and a signal vector interpretation. A new result concerning displaced spherical symmetric vectors will be treated in detail.

2.2.1 Probability Distributions and Functions

Spherical symmetric random vectors have probability distributions that are functions of the vector length only. This means that, in polar coordinates, the probability density is a function of r , the vector length. This function will be denoted as $f_s(r)$. For an s -dimensional vector we have

$$p(\underline{y}) = p(r, \theta_1, \theta_2, \dots, \theta_{s-1}) = f_s(r) \quad (2.16)$$

where $\theta_1, \theta_2, \dots, \theta_{s-2}$ run from 0 to π and θ_{s-1} from 0 to 2π .

The probability $p_s(r)$ that the vector length will lie between r and $r + dr$ is related to $f_s(r)$ by¹

¹ Note that the surface of a unit sphere in s -dimensions is

$$\Sigma_s = \frac{2 \pi^{\frac{s}{2}}}{\Gamma(\frac{s}{2})}$$

$$p_s(r) = \frac{2 \pi^{\frac{s}{2}} r^{s-1}}{\Gamma(\frac{s}{2})} f_s(r), \quad r \geq 0 \quad (2.17)$$

Where s is the dimension of the vector and $\Gamma(\cdot)$ is the gamma function. The cumulative probability function of the vector length is defined as

$$P_s(r) = \int_0^r p_s(u) du \quad (2.18)$$

and it is this function that yields performance measures for various systems. Unfortunately, it is usually not easy to determine $p_s(r)$ and in turn integrate it to find $P_s(r)$ because the distributions have awkward functional forms. It is possible, however, to approach the problem via the characteristic function.

2.2.2 The Characteristic Function and Related Transform Pairs

The characteristic function for an s -dimensional random vector is defined as the s -dimensional Fourier transform of the probability distribution. Since the vectors I am considering are spherical symmetric, the Fourier transform reduces to a Bessel integral which may be expressed in terms of any of the three functions $f_s(r)$, $p_s(r)$ or $P_s(r)$ (Lord 1954). Denoting the characteristic function as $\Phi(\rho)$ where ρ is the radial frequency we have

$$\Phi(\rho) = \frac{(2\pi)^{\frac{s}{2}}}{\rho^{\frac{s}{2}-1}} \int_0^\infty f_s(r) r^{\frac{s}{2}} J_{\frac{s}{2}-1}(\rho r) dr \quad (2.19)$$

$$\Phi(\rho) = \frac{\Gamma(\frac{s}{2}) 2^{\frac{s}{2}-1}}{\rho^{\frac{s}{2}-1}} \int_0^\infty \frac{p_s(r)}{r^{\frac{s}{2}-1}} J_{\frac{s}{2}-1}(rp) dr \quad (2.20)$$

and

$$\Phi(\rho) = \frac{\Gamma(\frac{s}{2}) 2^{\frac{s}{2}-1}}{\rho^{\frac{s}{2}-2}} \int_0^\infty \frac{p_s(r)}{r^{\frac{s}{2}-1}} J_{\frac{s}{2}}(rp) dr \quad (2.21)$$

where $J_n(\cdot)$ is the n^{th} order Bessel function, $\Gamma(\cdot)$ is the gamma function and s is the dimension of the vector.

The following transforms can be used to obtain the probability expressions from the characteristic function:

$$f_s(r) = \frac{1}{(2\pi)^{\frac{s}{2}} r^{\frac{s}{2}-1}} \int_0^\infty \Phi(\rho) \rho^{\frac{s}{2}} J_{\frac{s}{2}-1}(rp) d\rho \quad (2.22)$$

$$p_s(r) = \frac{r^{\frac{s}{2}}}{\Gamma(\frac{s}{2}) 2^{\frac{s}{2}-1}} \int_0^\infty \Phi(\rho) \rho^{\frac{s}{2}} J_{\frac{s}{2}-1}(rp) d\rho \quad (2.23)$$

$$p_s(r) = \frac{r^{\frac{s}{2}}}{\Gamma(\frac{s}{2}) 2^{\frac{s}{2}-1}} \int_0^\infty \Phi(\rho) \rho^{\frac{s}{2}-1} J_{\frac{s}{2}}(rp) d\rho \quad (2.24)$$

Now for most detection problems the dimension of the signal vectors is either one or two, but for the special case of noncoherent integration higher dimensions are required as is discussed in appendix A. For convenient future reference, table 2.1 gives the above transforms in one and two dimensions.

Table 2.1

Bessel transforms for one and two dimensions

$f_s(r)$: probability function for the vector written as a function of r

$p_s(r)$: probability function for the vector length r

$P_s(r)$: cumulative probability function for the vector length r

$\Phi(\rho)$: the characteristic function

One dimension

$$\Phi(\rho) = 2 \int_0^\infty f_1(r) \cos(\rho r) dr \qquad f_1(r) = \frac{1}{\pi} \int_0^\infty \Phi(\rho) \cos(\rho r) d\rho$$

$$\Phi(\rho) = \int_0^\infty p_1(r) \cos(\rho r) dr \qquad p_1(r) = \frac{2}{\pi} \int_0^\infty \Phi(\rho) \cos(\rho r) d\rho$$

$$\Phi(\rho) = \rho \int_0^\infty P_1(r) \sin(\rho r) dr \qquad P_1(r) = \frac{2}{\pi} \int_0^\infty \frac{\Phi(\rho)}{\rho} \sin(\rho r) d\rho$$

Two dimensions

$$\Phi(\rho) = 2\pi \int_0^\infty f_2(r) r J_0(\rho r) dr \qquad f_2(r) = \frac{1}{2\pi} \int_0^\infty \Phi(\rho) \rho J_0(\rho r) d\rho$$

$$\Phi(\rho) = \int_0^\infty p_2(r) J_0(\rho r) dr \qquad p_2(r) = r \int_0^\infty \Phi(\rho) \rho J_0(\rho r) d\rho$$

$$\Phi(\rho) = \rho \int_0^\infty P_2(r) J_1(\rho r) dr \qquad P_2(r) = r \int_0^\infty \Phi(\rho) J_1(\rho r) d\rho$$

The Bessel functions of order $-1/2$ and $1/2$ reduce to trigonometric functions so for $s = 1$ the transforms are trigonometric integrals.

The characteristic function is important because the characteristic function of a random vector made up of a sum of independent random vectors is the product of the individual characteristic functions. More specifically, if

$$\underline{y} = \sum_{n=1}^N \underline{y}_n \quad (2.25)$$

where $\underline{y}_n$ is a spherical symmetric vector, then

$$\Phi_{\underline{y}}(\rho) = \prod_{n=1}^N \Phi_{\underline{y}_n}(\rho) \quad (2.26)$$

From equation 2.26 it is seen that the sum of spherical symmetric random vectors is also spherical symmetric (i.e., $\Phi_{\underline{y}}$ depends only on ρ).

From the point of view of signal vectors the property described by equations 2.25 and 2.26 is important. It permits us to easily determine the characteristic function of a number of signal vectors combined in a linear fashion provided the individual characteristic functions are known. For example, if one vector represented interference and another noise, the characteristic function for the combined interference and noise vector is found by multiplying the characteristic functions of the two vectors.

Usually the characteristic functions for individual signal vectors are found readily and, therefore, so is the characteristic function of the composite. Problems do arise, though, when

transforming back to a probability expression. Inverse transforms for simple cases have been the object of much research but with no real breakthroughs (Esposito and Wilson 1973, Goldman 1974). In chapter 3 I present a technique that is both simple and efficient for solving these transforms. Specifically, the calculation of $P(r)$ is treated in great detail as system performance measures depend on this function.

Before discussing the calculation technique however, three important concepts related to spherically symmetric distributions will be discussed: the significance of the dimension change, the properties of the second moments, and the effect of displacement.

2.2.3 The Significance of a Dimension Change

As was pointed out by Lord (1954), decreasing the dimension of the probability space is equivalent to projecting the spherical symmetric distribution onto a subspace. The distribution in the subspace is again spherical symmetric and has the same characteristic function as the previous distribution.

This projection process is illustrated in figure 2.4. A circular symmetric distribution in two dimensions is projected onto one dimension where it becomes an even distribution with the same characteristic function. This process is important when considering error probabilities in CPSK systems as the noise and interferences processes are usually two-dimensional while the decision process is one-dimensional; as will be demonstrated in an example in chapter 4.

figure 2.4 a projection of a two-dimensional distribution onto one dimension

2.2.4 Properties of the Second Moment

The second moment of r is defined as

$$m \triangleq \int_0^{\infty} p_s(r) r^2 dr \quad (2.27)$$

and Lord (1954) showed that the second moments are additive.

That is, the second moment of a sum of independent spherically symmetric vectors is the sum of the individual second moments.

This is, of course, a general property of signal vectors where the second moment is the average signal energy, and the total average energy of a sum of independent signals is simply the sum of the individual average signal energies. For example, if the signal has constant amplitude and uniformly random phase then from equations 2.14 and 2.17 with $s = 2$ we have

$$p_2(r) = \delta(r-a) \quad (2.28)$$

where a is the signal amplitude. Then

$$m_c = \int_0^{\infty} \delta(r-a) r^2 dr = a^2 \quad (2.29)$$

where the subscript c denotes the constant amplitude.

For a slightly different approach consider a Gaussian signal vector. In the notation defined early in section 2.1, this vector is defined as one for which y_R and y_I are independent Gaussian random variables with equal variances and zero means.

For such a vector we have

$$E\{r^2\} = m_G = E\{y^T y\} = E\{y_R^2\} + E\{y_I^2\} \quad (2.30)$$

where subscript G denotes the Gaussian case and superscript T denotes the vector transpose. Since both y_R and y_I are Gaussian

with equal variances σ^2 and zero means we have

$$m_G = 2 \sigma^2 \quad (2.31)$$

where m_G is the average energy in the Gaussian signal.

To continue this very simple example, suppose we want the ratio of energy due to the constant amplitude signal to that due to the Gaussian signal, the signal-to-noise ratio perhaps. This is defined as

$$\frac{m_c}{m_G} = \frac{a^2}{2 \sigma^2} \quad (2.32)$$

Then if we wanted the average energy for the two signals added together we would just sum m_c and m_G to give

$$m = a^2 + 2 \sigma^2 \quad (2.33)$$

Later we shall see that the additive property of the second moment and its relationship to the average energy makes it possible to determine energy ratios for relatively complex signal combinations by straightforward inspection.

2.2.5 Displaced Spherical Symmetric Distributions

Signals that are perfectly coherent do not have a circular symmetric distribution in the phase plane. Their distributions are represented by an impulse at the appropriate amplitude and phase. When this signal is added to one represented by a circular symmetric distribution the effect is to displace the circular symmetric distribution, as illustrated in figure 2.5. The resulting distribution is not circular symmetric.

figure 2.5 convolution of a displaced impulse distribution with a general circular symmetric distribution.

Since coherent signals are encountered in many detection problems a method for handling them is required. As mentioned in section 2.2.3, a change to one dimension can sometimes resolve the problem, but when this approach cannot, it becomes necessary to handle the displaced symmetric distribution in two dimensions. More specifically, some performance calculations require the determination of the probability that the signal vector length will be smaller (or larger) than some constant, and circular symmetric probability theory cannot in general handle this for the non-symmetric case. However, if the distribution which was displaced was originally circular symmetric it can.

Consider the two sets of displaced distributions shown in figures 2.6 and 2.7. In figure 2.6 the distribution which was displaced is circular symmetric while in figure 2.7 it is not. Suppose we are interested in calculating the probability that the vector length is less than a constant T_h . In other words, we require the probability that the vector lies inside a circle of radius T_h . It can be seen from figure 2.6 that if the displaced distribution is circular symmetric, the desired probability does not change with the orientation of the displacement. On the other hand if the displaced distribution is not circular symmetric the orientation is important (figure 2.7). Since the orientation is irrelevant for the circular symmetric case the probability of being inside the circle is the same if

figure 2.6 an illustration showing that the orientation of the displacement does not affect the probability of being inside circle T_h when the distribution being displaced is circular symmetric.

figure 2.7 an illustration showing that the orientation of the displacement does affect the probability of being inside circle T_h when the distribution being displaced is not circular symmetric.

a constant amplitude (equal to the displacement amplitude) uniformly random phase vector is added to the circular symmetric vector instead of the fixed displacement. This vector is circular symmetric and hence the result is circular symmetric and the theory for circular symmetric random vectors can be used to calculate the desired probability.

In summary, if we require the probability of being inside (or outside) a certain radius T_h for a displaced circular symmetric distribution we do the following. Add the circular symmetric vector to be displaced to a constant amplitude (amplitude equal to the displacement) uniformly random phase vector and use the theory developed for circular symmetric random vectors. This gives rise to an equivalent circular symmetric distribution (equivalent in the sense that it gives the same probability result as the displaced distribution) which will be denoted by a subscript e.

Although this discussion was confined to the two-dimensional case the result is true in general. A formal proof of the general result is given in section 2.4.

2.3 Chapter Summary

This chapter has presented the basic form of the detection problem along with a framework for solution. It was shown how the output of a simple bandpass model describing many signal detection systems is conceptually represented by a phase plane

diagram which leads directly to the concept of circular symmetric random vectors. The relevant probability theory associated with these vectors was then presented.

In the discussion the importance and difficulty of calculating $P(r)$ was explained. The calculation technique to be presented in the next chapter eliminates this difficulty and makes the general approach viable.

2.4 Mathematical Details

This section presents the mathematical detail leading to the proof of the statements made in subsection 2.2.5 concerning displaced spherical symmetric random vectors. It is recommended that readers not interested in this detail disregard this section to preserve the continuity of the thesis discussion.

2.4.1 Displaced Spherical Symmetric Distributions

If a spherical symmetric probability distribution is displaced from the origin, the equivalent distribution for r (the distance from the origin) may be found by adding a constant amplitude (equal to the displacement) uniformly random oriented vector to the spherical symmetric vector. To show that this is true I will first integrate the displaced distribution over the inside of a sphere of radius T_h and from the result identify $p_{es}(r)$, the equivalent spherical symmetric distribution. Next, I will add a constant amplitude uniformly

random oriented vector to the spherically symmetric vector (i.e., convolve their distributions) and show that this approach likewise yields $p_{es}(r)$.

Integrating Over a Sphere

To find the probability of being inside a sphere of radius T_h we must solve the following integral

$$P(r < T_h) = \int_0^{T_h} \int_{\Sigma_s} p(r, \theta) r^{s-1} d\Sigma_s dr \quad (2.34)$$

where Σ_s is the surface of the unit sphere defined as

$$\Sigma_s \triangleq \frac{2\pi^{\frac{s}{2}}}{\Gamma(\frac{s}{2})} \quad (2.35)$$

and θ is the orientation vector with components $\theta_1, \theta_2, \dots, \theta_{s-1}$

where $\theta_1, \theta_2, \dots, \theta_{s-2}$ run from 0 to π and θ_{s-1} runs from 0 to 2π .

Now, $p(r, \theta)$ may be expressed in linear coordinates

$$p(r, \theta) = p(x_1, x_2, x_3, \dots, x_s) \quad (2.36)$$

where x_n is the n^{th} coordinate, and since the distribution is spherical symmetric except for a displacement $\underline{a}$ we may write

$$p(x_1, x_2, \dots, x_s) = f_s(\sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2 + \dots + (x_s - a_s)^2}) \quad (2.37)$$

where $a_1, a_2, \dots, a_s$ are the components of the vector $\underline{a}$.

But the argument for $f_s(\cdot)$ in polar coordinates is

$$r^2 + a^2 - 2 r a \cos(\theta) \quad (2.38)$$

where $r^2 = x_1^2 + x_2^2 + \dots + x_s^2$, $a^2 = a_1^2 + a_2^2 + \dots + a_s^2$ and

$r a \cos(\theta) = x_1 a_1 + x_2 a_2 + \dots + x_s a_s$. Since the orientation

of $\underline{a}$ is irrelevant we may choose $\theta = \theta_1$, and from equation 2.34

$$P(r < T_h) = \int_0^{T_h} \left(\int_{\Sigma_s} f_s(\sqrt{r^2 + a^2 - 2ra \cos(\theta_1)}) r^{s-1} d\Sigma_s \right) dr \quad (2.39)$$

By inspection the expression inside the brackets is $p_{es}(r)$. Hence

$$p_{es}(r) = r^{s-1} \int_{\Sigma_s} f_s(\sqrt{r^2 + a^2 - 2ra \cos(\theta_1)}) d\Sigma_s \quad (2.40)$$

Now I shall obtain the same expression for $p_{es}(r)$ by the process of adding a uniformly random oriented constant amplitude vector to the spherical symmetric vector.

Adding a Constant Length Uniformly Random Oriented Vector

Adding two spherical symmetric distributions results in the convolution of their distributions. Denoting the circular symmetric distribution to be displaced as $p_f(\underline{x})$ and the distribution for the constant length uniformly random oriented vector as $p_d(\underline{x})$ we have

$$p(r, \theta) = p_d(\underline{x}) * p_f(\underline{x}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_d(\underline{u}) p_f(\underline{x} - \underline{u}) d\underline{u} \quad (2.41)$$

Since $p_d(\underline{x})$ and $p_f(\underline{x})$ are both spherical symmetric this integral can be written in general spherical coordinates as

$$p(r, \theta) = \int_0^{\infty} \int_{\Sigma_s} d_s(r) f_s(\sqrt{r^2 + x^2 - 2xr \cos(\theta)}) r^{s-1} d\Sigma_s dr \quad (2.42)$$

where $r^2 = u_1^2 + u_2^2 + \dots + u_s^2$, $x^2 = x_1^2 + x_2^2 + \dots + x_s^2$, θ is the angle between $\underline{x}$ and $\underline{u}$, and $d_s(r)$ and $f_s(r)$ are defined by equation 2.16. The integral does not depend on the direction of $\underline{x}$ so we can

take it to be along θ_1 and obtain

$$p(r, \theta) = \int_0^\infty r^{s-1} d_s(r) \int_{\Sigma_s} f_s(\sqrt{r^2 + x^2 - 2xr \cos(\theta_1)}) d\Sigma_s dr \quad (2.43)$$

Now we know that

$$\int_0^\infty \int_{\Sigma_s} r^{s-1} d_s(r) d\Sigma_s dr = 1 \quad (2.44)$$

because the total probability in the s-dimension sphere must be 1, and since

$$d_s(r) = C \delta(r-a) \quad (2.45)$$

where C is some constant and a is the length of the vector, we have by substituting equation 2.45 into 2.44

$$d_s(r) = \frac{\Gamma(\frac{s}{2})}{2 a^{s-1} \pi^{\frac{s}{2}}} \delta(r-a) \quad (2.46)$$

Therefore equation 2.43 becomes

$$p(r, \theta) = \frac{\Gamma(\frac{s}{2})}{2 \pi^{\frac{s}{2}}} \int_{\Sigma_s} f_s(\sqrt{r^2 + x^2 - 2xr \cos(\theta_1)}) d\Sigma_s \quad (2.47)$$

Using equation 2.17 to find $p_s(r)$ we have

$$p_s(r) = r^{s-1} \int_{\Sigma_s} f_s(\sqrt{r^2 + x^2 - 2xr \cos(\theta_1)}) d\Sigma_s \quad (2.48)$$

which is the same as the expression for $p_{es}(r)$ (equation 2.40)

Therefore

$$p_s(r) = p_{es}(r) \quad (2.49)$$

Q.E.D.

CHAPTER 3

CALCULATION TECHNIQUE

In the last chapter it was shown how circular symmetric probability distributions arise naturally from signal detection problems. The probability expressions were in the form of Bessel integrals. These integrals are difficult to solve by conventional techniques, thus limiting the usefulness of the general approach. In this chapter I present a simple and efficient technique for solving these Bessel integrals. This technique reduces the solution of many formerly difficult problems to the point of triviality.

3.1 The Problem and an Introduction to a New Solution

This section is divided into three subsections. The first discusses the problem that must be solved to make the general approach to detection problems viable. The next subsection briefly outlines previous solutions and their difficulties while the final subsection introduces the new solution along with its history.

3.1.1 The Problem

The probability distribution $p_s(r)$ and the cumulative probability function $P_s(r)$ were defined for spherical symmetric random vectors in chapter 2. In that chapter the following Bessel

integrals were shown to represent these functions.

$$p_s(r) = \frac{r^{\frac{s}{2}}}{\Gamma(\frac{s}{2}) 2^{\frac{s}{2}-1}} \int_0^\infty \Phi(\rho) \rho^{\frac{s}{2}} J_{\frac{s}{2}-1}(r\rho) d\rho \quad (3.1)$$

$$P_s(r) = \frac{r^{\frac{s}{2}}}{\Gamma(\frac{s}{2}) 2^{\frac{s}{2}-1}} \int_0^\infty \Phi(\rho) \rho^{\frac{s}{2}-1} J_{\frac{s}{2}}(r\rho) d\rho \quad (3.2)$$

where $\Phi(\rho)$ is the characteristic function for the s -dimension probability distribution.

Of the two expressions, $P_s(r)$ is the most important since it is from $P_s(r)$ that performance measures such as probabilities of detection and error are determined. Therefore it is necessary to be able to efficiently compute $P_s(r)$ with sufficient accuracy. It is also desirable to be able to calculate $p_s(r)$ but in a qualitative sense since no system performance measures are directly derived from $p_s(r)$. The calculation technique presented in this chapter is applicable to calculating both $P_s(r)$ and $p_s(r)$, however in the error analysis I shall concentrate on the expression for $P_s(r)$.

For most problems the Bessel integrals of equation 3.1 and 3.2 are not easy to solve. Previous solutions have been numerical in nature and good only for the problem at hand, as discussed in section 1.1.2.

3.1.2 Traditional Solutions

The traditional method employed in solving the Bessel integrals is to use an expansion involving the moments of the

spherical symmetric vectors. These expansions are complicated and require the use of special functions such as the generalized Laguerre polynomial and the Hermite polynomial (Goldman 1974, Rosenbaum 1970). Also, there are problems in numerically handling these expansions for some parameter values as the terms can become very large and alternate in sign (Esposito and Wilson 1973, Goldman 1974). The methods used for terminating the series expansions are usually ad hoc and very little information about the accuracy of the results is given. Also, since the expansions require knowledge of the moments of the resulting spherical symmetric vector, these moments must be determined. This is not always an easy task. In fact, this may become a major problem in itself (Rosenbaum 1970). The difficulty in determining moments and the numerical problems with the series makes this solution difficult. Often new tricks must be employed for each problem (Esposito and Wilson 1973).

3.1.3 Proposed Solution

In this chapter I present a technique that is both simple and efficient and may be applied directly to a wide variety of problems. It involves an assumption that the resulting spherical symmetric vector is amplitude limited. The Fourier-Bessel series is then used to calculate the desired probability. Although the method is not exact as the assumption is often violated (for example, by Gaussian noise) the error can be bounded, and two parameters can be adjusted to obtain a result that is arbitrarily

close to the desired answer. This method uses the characteristic function for the spherical symmetric random vector which can usually be easily found. Also, the terms of the Fourier-Bessel series are well behaved and therefore do not present the numerical difficulties encountered in the moment expansion approach. The computer implementation of the calculation technique is straightforward and requires little programming expertise. A typical subprogram for calculating probabilities can be written in about 50 lines of Fortran code with an execution time in the order of a few milliseconds.

The Fourier-Bessel series was first suggested by Barakat (1974), for the amplitude limited distributions found in applications involving laser speckle patterns. Jao and Elbaum (1978) then used the series to determine the first-order statistics of a non-Rayleigh fading signal. In their application the assumption of amplitude limited vectors was violated, but through careful choice of series parameters they were able to obtain results for smooth distributions that were accurate to four decimal places when using single precision floating point arithmetic on a 32-bit-word computer. This was adequate for their purposes but for applications where one needs precision to 10 or 12 decimal places, such as in the calculation of the probability of error for communication systems, a systematic error bounding technique is required. This chapter presents such a technique.

3.2. The New Technique

In this section the new technique for solving the Bessel integrals described in the last section is presented. First a detailed description of the technique is given. This is followed by an error analysis which permits one to bound the accuracy of the result.

3.2.1 Description

Since the probability distribution is spherically symmetric it may be expressed exactly, out to some radius R , by a Fourier-Bessel series. The resulting expressions for $p_s(r)$ and $P_s(r)$ are as follows (with mathematical details given in section 3.3.1).

$$p_s(r) = \frac{r^{\frac{s}{2}}}{2^{\frac{s}{2}-2} R^{\frac{s}{2}+1} \Gamma(\frac{s}{2})} \sum_{k=1}^{\infty} \frac{\lambda_k^{\frac{s}{2}-1} F_k}{J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}-1}(\lambda_k \frac{r}{R}) \quad (3.3)$$

$P_s(r)$ is found by integrating this series term by term.

$$P_s(r) = \frac{r^{\frac{s}{2}}}{2^{\frac{s}{2}-2} R^{\frac{s}{2}} \Gamma(\frac{s}{2})} \sum_{k=1}^{\infty} \frac{\lambda_k^{\frac{s}{2}-2} F_k}{J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}}(\lambda_k \frac{r}{R}) \quad (3.4)$$

λ_k is the k^{th} zero of $J_{\frac{s}{2}-1}(\cdot)$ and F_k is the Fourier-Bessel coefficient defined as

$$F_k \triangleq \frac{\Gamma(\frac{s}{2})}{\lambda_k^{\frac{s}{2}-1}} \int_0^R \frac{P_s(r)}{r^{\frac{s}{2}-1}} J_{\frac{s}{2}-1}(\lambda_k \frac{r}{R}) dr \quad (3.5)$$

The Fourier-Bessel coefficient is identical to the characteristic function (ρ replaced by $\frac{\lambda_k}{R}$) if $p_s(r)$ is zero for $r > R$. If $p_s(r)$ is not zero past R , the two are not equal, however if $p_s(r)$, $r > R$, is very small F_k and $\phi(\frac{\lambda_k}{R})$ are close. This results in a small error in $P_s(r)$ when $\phi(\frac{\lambda_k}{R})$ is used instead of F_k for the calculation. The error can be made arbitrarily small by increasing R .

For practical calculations we also have another source of error in that the series must be truncated at some number of terms, say M . This causes a truncation error that must be considered along with the error caused by using a finite R .

Given these considerations, our estimates of $p_s(r)$ and $P_s(r)$, using a truncated series and the characteristic function instead of F_k , are:

$$\hat{p}_s(r) = \frac{r^{\frac{s}{2}}}{2^{\frac{s}{2}-2} R^{\frac{s}{2}+1} \Gamma(\frac{s}{2})} \sum_{k=1}^M \frac{\lambda_k^{\frac{s}{2}-1} \phi(\frac{\lambda_k}{R})}{J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}-1}(\lambda_k \frac{r}{R}) \quad (3.6)$$

and

$$\hat{P}_s(r) = \frac{r^{\frac{s}{2}}}{2^{\frac{s}{2}-2} R^{\frac{s}{2}} \Gamma(\frac{s}{2})} \sum_{k=1}^M \frac{\lambda_k^{\frac{s}{2}-2} \phi(\frac{\lambda_k}{R})}{J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}}(\lambda_k \frac{r}{R}) \quad (3.7)$$

and Table 3.1 shows the form of these series for $s = 1$ and 2 .

What remains is to bound the error for $P_s(r)$ in such a manner that the bound may be used to determine appropriate values for R and M .

Table 3.1

Estimates for $p_s(r)$ and $\hat{p}_s(r)$ for $s = 1$ and 2.

One dimension ($s = 1$)

$$\hat{p}_1(r) = \frac{2}{R} \sum_{k=1}^M \phi\left(\frac{(2k-1)\pi}{2R}\right) \cos\left(\frac{(2k-1)\pi r}{2R}\right)$$

$$\hat{p}_1(r) = \frac{4}{\pi} \sum_{k=1}^M \phi\left(\frac{(2k-1)\pi}{2R}\right) \frac{1}{(2k-1)} \sin\left(\frac{(2k-1)\pi r}{2R}\right)$$

Two dimensions ($s = 2$)

$$\hat{p}_2(r) = \frac{2r}{R^2} \sum_{k=1}^M \frac{\phi\left(\frac{\lambda_k}{R}\right)}{J_1^2(\lambda_k)} J_0\left(\lambda_k \frac{r}{R}\right)$$

$$\hat{p}_2(r) = \frac{2r}{R} \sum_{k=1}^M \frac{\phi\left(\frac{\lambda_k}{R}\right)}{\lambda_k J_1^2(\lambda_k)} J_1\left(\lambda_k \frac{r}{R}\right)$$

3.2.2 Error Analysis

The error in the estimate $\hat{P}_s(r)$ may be determined as follows.

$$e = P_s(r) - \hat{P}_s(r) = \frac{r^{\frac{s}{2}}}{2^{\frac{s}{2}-2} R^{\frac{s}{2}} \Gamma(\frac{s}{2})} \left[\sum_{k=1}^{\infty} \frac{\lambda_k^{\frac{s}{2}-2} F_k}{J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}}(\lambda_k \frac{r}{R}) - \sum_{k=1}^M \frac{\lambda_k^{\frac{s}{2}-2} \phi(\frac{\lambda_k}{R})}{J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}}(\lambda_k \frac{r}{R}) \right] \quad (3.8)$$

By adding and subtracting the expression

$$\frac{r^{\frac{s}{2}}}{2^{\frac{s}{2}-2} R^{\frac{s}{2}} \Gamma(\frac{s}{2})} \sum_{k=M+1}^{\infty} \frac{\lambda_k^{\frac{s}{2}-2} \phi(\frac{\lambda_k}{R})}{J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}}(\lambda_k \frac{r}{R})$$

the error may be written as the sum of two error terms as follows

$$e = \frac{r^{\frac{s}{2}}}{2^{\frac{s}{2}-2} R^{\frac{s}{2}} \Gamma(\frac{s}{2})} \sum_{k=1}^{\infty} \frac{\lambda_k^{\frac{s}{2}-2} (F_k - \phi(\frac{\lambda_k}{R}))}{J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}}(\lambda_k \frac{r}{R}) + \frac{r^{\frac{s}{2}}}{2^{\frac{s}{2}-2} R^{\frac{s}{2}} \Gamma(\frac{s}{2})} \sum_{k=M+1}^{\infty} \frac{\lambda_k^{\frac{s}{2}-2} \phi(\frac{\lambda_k}{R})}{J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}}(\lambda_k \frac{r}{R}) \quad (3.9)$$

The first sum is the error caused by using a finite R ; it will be denoted as e_{Rs} . The second sum is the truncation error and will be denoted as e_{ts} . e_{Rs} tells us how close the estimate would be

if no truncation was involved and e_{ts} gives the error in the estimate caused by truncation. The net error is the sum of the two. We will now consider the two errors separately.

Truncation error e_{ts} for two dimensions becomes

$$e_{t2} = \frac{2r}{R} \sum_{k=M+1}^{\infty} \frac{\phi\left(\frac{\lambda_k}{R}\right) J_1\left(\lambda_k \frac{r}{R}\right)}{\lambda_k J_1^2(\lambda_k)} \quad (3.10)$$

and this may be bounded by

$$|e_{t2}| \leq \frac{2r}{R} \sum_{k=M+1}^{\infty} \frac{\left|\phi\left(\frac{\lambda_k}{R}\right)\right| \left|J_1\left(\lambda_k \frac{r}{R}\right)\right|}{\lambda_k J_1^2(\lambda_k)} \quad (3.11)$$

Since the term $\lambda_k J_1^2(\lambda_k)$ is not smaller than 0.6, which may be verified by tables for small k and by the asymptotic form for large k , and since $|J_1(\cdot)| \leq 1/\sqrt{2}$ we have

$$\frac{2 \left|J_1\left(\lambda_k \frac{r}{R}\right)\right|}{\lambda_k J_1^2(\lambda_k)} \leq 1.25 \quad (3.12)$$

Also, the largest value of r is R , so for two dimensions

$$|e_{t2}| < 1.25 \sum_{k=M+1}^{\infty} \left|\phi\left(\frac{\lambda_k}{R}\right)\right| \quad (3.13)$$

A similar argument for one dimension yields

$$|e_{t1}| < \frac{4}{\pi} \sum_{k=M+1}^{\infty} \frac{\left|\phi\left(\frac{(2k-1)\pi}{2R}\right)\right|}{2k-1} \quad (3.14)$$

The trigonometric forms for $J_{\frac{1}{2}}(\cdot)$ and $J_{-\frac{1}{2}}(\cdot)$ were used to obtain this expression.

Equations 3.13 and 3.14 provide us with an upper bound on the truncation error. This bound is easily calculated since we know the characteristic function $\Phi(\cdot)$.

We shall now consider the error caused by using a finite R . Replacing F_k and $\Phi\left(\frac{\lambda_k}{R}\right)$ by their integral forms (equation 3.5 and 2.20) e_{Rs} becomes

$$e_{Rs} = -r^{\frac{s}{2}} 2R \sum_{k=1}^{\infty} \frac{\int_R^{\infty} \frac{J_{\frac{s}{2}-1}\left(\lambda_k \frac{x}{R}\right)}{x^{\frac{s}{2}-1}} p_s(x) dx}{\lambda_k J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}}\left(\lambda_k \frac{r}{R}\right) \quad (3.15)$$

The summation and the integration order will now be interchanged, although the validity of this interchange cannot be strictly established except for the case $s = 1$. For larger s there is some question as to the convergence of the resulting sum. For the purpose of illustration, however, I shall continue as if the interchange is valid as the final result is not affected. This interchange gives

$$e_{Rs} = -r^{\frac{s}{2}} 2R \int_R^{\infty} \frac{p_s(x)}{x^{\frac{s}{2}-1}} \left[\sum_{k=1}^{\infty} \frac{J_{\frac{s}{2}-1}\left(\lambda_k \frac{x}{R}\right) J_{\frac{s}{2}}\left(\lambda_k \frac{r}{R}\right)}{\lambda_k J_{\frac{s}{2}}^2(\lambda_k)} \right] dx \quad (3.16)$$

and we see that e_{Rs} depends upon the function inside the square brackets (unless $p_s(x) = 0$ for $x > R$, and then $e_{Rs} = 0$). In one

dimension this function is well defined and we have

$$e_{R1} = - \int_R^\infty p_1(x) \left[2 \sum_{k=1}^{\infty} \frac{\cos(\frac{(2k-1)\pi x}{2R}) \sin(\frac{(2k-1)\pi r}{2R})}{\frac{\pi}{2}(2k-1)} \right] dx \quad (3.17)$$

The function in the square brackets is the Fourier series for the function shown in figure 3.1. Since this function is bounded by the value 1, we may write

$$|e_{R1}| \leq \int_R^\infty p_1(x) dx \quad (3.18)$$

However, it is often easier to define the distributions in two dimensions as that is the natural dimension of the signal space.

If we define $Q(R)$ as the following integral

$$Q(R) \triangleq \int_R^\infty p_2(x) dx \quad (3.19)$$

then by a simple consideration of the corresponding areas of integration we obtain the inequality

$$Q(R) \geq \int_R^\infty p_1(x) dx \quad (3.20)$$

and so a bound for e_R in one dimension is

$$|e_{R1}| \leq Q(R) \quad (3.21)$$

Now this is not entirely satisfactory as we have a bound on e_{R1} that depends on the very probability function we are calculating, but it may be shown (section 3.3.2) that if the desired distribution is made up of a convolution of k distributions (the resulting vector is a sum of k independent random vectors)

figure 3.1 illustration of the periodic function described by the Fourier series in equation 3.17.

then

$$Q(x_0) \leq \sum_{i=1}^k Q_i(x_{i-1} - x_i) \quad x_0 > x_1 > x_2 > \dots > x_k = 0 \quad (3.22)$$

where $Q_i(\cdot)$ is the function defined by equation 3.19 for the i^{th} distribution. Therefore, the bound on e_R for one dimension becomes

$$|e_{R1}| \leq \sum_{i=1}^k Q_i(x_{i-1} - x_i) \quad x_0 = R > x_1 > x_2 > \dots > x_k = 0 \quad (3.23)$$

which yields a bound in terms of the component distributions which are usually easy to handle.

The remaining task is to choose the $x_{k-1}, \dots, x_1$ such that the error is within a desired bound. This process is quite simple as is demonstrated in chapter 4.

For two dimensions the error is

$$e_{R2} = -2 \pi R \int_R^\infty p_2(x) \left[\sum_{k=1}^\infty \frac{J_0(\lambda_k \frac{x}{R}) J_1(\lambda_k \frac{x}{R})}{\lambda_k J_1^2(\lambda_k)} \right] dx \quad (3.24)$$

but the convergence of the series in brackets for $x > R$ is in question. This is a difficult mathematical problem which has not been developed in the mathematical literature¹. To circumvent this difficulty an effective numerical method for bounding e_{Rs} was developed. This will now be described.

¹ I would like to acknowledge the help I received from Professor Mizanur Rahman of the Carleton University math department in establishing this fact and thank him for his encouragement concerning the approach to follow.

Since the two sources of error, finite R and truncation, are separable we may make e_{t2} so small that it is negligible compared to the accuracy of the calculations being made. Then we may choose an R that appears reasonable and calculate the estimate $\hat{P}_2(r)$. R is then increased and $\hat{P}_2(r)$ calculated again. By comparing this result with the previous estimate, one may test for convergence by looking for changes in the significant digit of interest. Results from calculating known probabilities for distributions that decay in an exponential manner past radius R indicate that a rule of thumb is that e_{R2} is bounded by $Q(R)$. Therefore, as a good initial choice one would choose an R such that $Q(R)$ is less than the acceptable error, where $Q(R)$ may be bounded using equation 3.22. The use of this technique will be demonstrated in chapter 4.

3.3 Mathematical Details

This section presents the mathematical detail leading to some of the equations used in chapter 3. It is recommended that readers not interested in this detail disregard this section to preserve the continuity of the thesis discussion.

3.3.1 The Fourier-Bessel Series

Consider an s -dimensional spherical symmetric probability distribution. The characteristic function is expressed as

$$\phi(\rho) = \frac{\Gamma(\frac{s}{2})}{\frac{s}{2}-1} \int_0^{\infty} \frac{p_s(r)}{r^{\frac{s}{2}-1}} J_{\frac{s}{2}-1}(\rho r) dr \quad (3.25)$$

and the Fourier-Bessel coefficient F_k is

$$F_k = \frac{\Gamma(\frac{s}{2})}{\frac{s}{2}-1} \int_0^R \frac{p_s(r)}{r^{\frac{s}{2}-1}} J_{\frac{s}{2}-1}(\lambda_k \frac{r}{R}) dr \quad (3.26)$$

We must now find an expression for $p_s(r)$ in terms of the Fourier-Bessel coefficients F_k . Changing the variable of integration in equation 3.26 to $\frac{r}{R}$ gives

$$\left(\frac{\lambda_k}{2R}\right)^{\frac{s}{2}-1} \frac{F_k}{\Gamma(\frac{s}{2})} = \frac{1}{\int_0^1} \frac{u p_s(uR) R^2}{(uR)^{\frac{s}{2}}} J_{\frac{s}{2}-1}(u \lambda_k) du \quad (3.27)$$

By definition of the Fourier-Bessel series (Spiegel 1968, equation 24.109)

$$\frac{p_s(uR) R^2}{(uR)^{\frac{s}{2}}} = \sum_{k=1}^{\infty} \frac{2}{J_{\frac{s}{2}}^2(\lambda_k)} \left(\frac{\lambda_k}{2R}\right)^{\frac{s}{2}-1} \frac{F_k}{\Gamma(\frac{s}{2})} J_{\frac{s}{2}-1}(u \lambda_k) \quad (3.28)$$

and letting $r = uR$, we have

$$p_s(r) = \frac{r^{\frac{s}{2}}}{2^{\frac{s}{2}-2} R^{\frac{s}{2}+1} \Gamma(\frac{s}{2})} \sum_{k=1}^{\infty} \frac{\lambda_k^{\frac{s}{2}-1} F_k}{J_{\frac{s}{2}}^2(\lambda_k)} J_{\frac{s}{2}-1}(\lambda_k \frac{r}{R}) \quad (3.29)$$

which is the desired result.

To obtain $P_s(r)$ we integrate $p_s(r)$ term by term.

$$P_s(r) = \int_0^r p(u) du = \frac{1}{2^{\frac{s}{2}-2} R^{\frac{s}{2}+1} \Gamma(\frac{s}{2})} \sum_{k=1}^{\infty} \frac{\lambda_k^{\frac{s}{2}-1} F_k}{J_{\frac{s}{2}}^2(\lambda_k)} \int_0^r u^{\frac{s}{2}-1} J_{\frac{s}{2}-1}(\lambda_k \frac{u}{R}) du \quad (3.30)$$

The integral, after a change in variables, is tabulated in Spiegel (1968, equation 24.84) as

$$\int_0^r u^{\frac{s}{2}-1} J_{\frac{s}{2}-1}(\lambda_k \frac{u}{R}) du = \frac{R}{\lambda_k} r^{\frac{s}{2}} J_{\frac{s}{2}}(\lambda_k \frac{r}{R}) \quad (3.31)$$

so

$$P_s(r) = \frac{r^{\frac{s}{2}}}{R^{\frac{s}{2}} 2^{\frac{s}{2}-2} \Gamma(\frac{s}{2})} \sum_{k=1}^{\infty} \frac{\lambda_k^{\frac{s}{2}-2} F_k}{J_{\frac{s}{2}}(\lambda_k)} J_{\frac{s}{2}}(\lambda_k \frac{r}{R}) \quad (3.32)$$

Q.E.D.

3.3.2 Bound on $Q(R)$

Let us assume that $f(r)$ is the result of convolving two circular symmetric distributions, $f_1(r)$ and $f_2(r)$. Then by using the definition of convolution in two dimensions we have

$$f(r) = f_1(r) * f_2(r) = \int_0^{\infty} \int_0^{2\pi} f_1(x) f_2(\sqrt{r^2 + x^2 - 2rx \cos(\theta)}) x d\theta dx \quad (3.33)$$

Defining $p_{2e}(r|x)$ according to equation 2.40 we have

$$p_{2e}(r|x) \triangleq r \int_0^{2\pi} f_2(\sqrt{r^2 + x^2 - 2rx \cos(\theta)}) d\theta \quad (3.34)$$

which is a circular symmetric distribution obtained by displacing $f_2(r)$ by x and then integrating on a circle of radius r .

Therefore

$$f(r) = \int_0^{\infty} f_1(x) x \frac{p_{2e}(r|x)}{r} dx \quad (3.35)$$

Now

$$\begin{aligned} Q(z) &= \int_z^{\infty} \int_0^{2\pi} f(r) r d\theta dr \\ &= \int_z^{\infty} \int_0^{2\pi} \int_0^{\infty} f_1(x) x \frac{p_{2e}(r|x)}{r} dx r d\theta dr \end{aligned} \quad (3.36)$$

resulting in

$$Q(z) = \int_0^{\infty} 2\pi x f_1(x) \int_z^{\infty} p_{2e}(r|x) dr dx \quad (3.37)$$

Defining

$$Q_{2e}(z|x) \triangleq \int_z^{\infty} p_{2e}(r|x) dr \quad (3.38)$$

we have

$$Q(z) = \int_0^{\infty} p_1(x) Q_{2e}(z|x) dx \quad (3.39)$$

and by defining x_1 as some x less than z we may write

$$Q(z) = \int_0^{x_1} p_1(x) Q_{2e}(z|x) dx + \int_{x_1}^{\infty} p_1(x) Q_{2e}(z|x) dx \quad (3.40)$$

2

Now for the condition $x_1 < z$ we have the inequality

$$Q_{2e}(z|x) \leq Q_2(z - x_1) \quad \text{for } x \leq x_1 \quad (3.41)$$

To verify this inequality consider figure 3.2. In this figure the distribution is centered around the point x_1 , therefore, the probability of being outside contour C_1 is $Q(z|x_1)$. This is obviously smaller than the probability of being outside contour C_2 which is $Q_2(z - x_1)$. If the distribution is centered at any $x < x_1$ then we may still conclude that the probability outside C_1 is smaller than outside C_2 (C_2 being centered at x with radius $z - x_1$). Noting also that

$$Q_{2e}(z|x) \leq 1 \quad (3.42)$$

we may rewrite equation 3.40 as the following inequality:

$$Q(z) \leq \int_0^{x_1} p_1(x) dx Q_2(z - x_1) + \int_{x_1}^{\infty} p_1(x) dx \quad (3.43)$$

The first integral is less than or equal to 1 and the second is just $Q_1(x_1)$ so we may write

$$Q(z) \leq Q_2(z - x_1) + Q_1(x_1) \quad (3.44)$$

If more than two distributions are convolved to obtain $f(r)$ we may easily extend equation 3.44 to include the others. Suppose $f_1(r)$ is actually the convolution of two distributions $f_3(r)$ and $f_4(r)$ then from the same argument as above

$$Q_1(x_1) \leq Q_3(x_1 - x_2) + Q_4(x_2) \quad x_2 < x_1 \quad (3.45)$$

figure 3.2 contours to demonstrate inequality 3.41.

Substituting this bound for $Q_1(x_1)$ into equation 3.44 we have

$$Q(z) \leq Q_2(z - x_1) + Q_3(x_1 - x_2) + Q_4(x_2) \quad z > x_1 > x_2 \geq 0 \quad (3.46)$$

Again we could break up $Q_4(x_2)$ in a similar way to account for the convolution of four distributions. By induction (after rearranging the subscripts) then

$$Q(x_0) \leq \sum_{i=1}^N Q_i(x_{i-1} - x_i) \quad x_0 > x_1 > x_2 > \dots > x_N = 0 \quad (3.47)$$

where N is the number of distributions convolved, or equivalently is the number of circular symmetric random vectors added together.

CHAPTER 4

EXAMPLES

In chapter 2 I showed how a large class of detection problems could be described with the phase plane and circular symmetric probability distribution theory. The resulting probability expressions were in the form of Bessel integrals. A technique for efficiently solving these integrals was presented in chapter 3. In chapter 4 I will present several examples demonstrating the utility of the general approach to detection problems.

Chapter 4 is divided into two sections. The first section demonstrates the use of the calculation technique in determining the probability distribution and cumulative probability function of several significant signals. The second section shows how the results obtained from these examples can be related to systems performance calculations.

4.1 Calculation of Distribution and Probability Functions

In this section the probability density functions and cumulative probability functions are calculated for circular symmetric random vectors that commonly appear in signal detection problems. First I treat the case of Gaussian noise, demonstrating both the two-dimensional and one-dimensional distributions. Next the constant amplitude random phase vector is considered. Finally,

the distributions for a linear combination of a Gaussian signal and an arbitrary number of constant amplitude random phase signals are found.

4.1.1 Gaussian Signals

As was mentioned in chapter 2, Gaussian signals can be described as a zero mean complex Gaussian random variable. In vector notation, this is a Gaussian random vector with two independent components of equal variance and zero mean. The two-dimensional probability distribution for this vector is shown in figure 4.1. The axes, x_1 and x_2 , represent the quadrature components of the bandpass signal, and the probability distribution is

$$f(x_1, x_2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x_1^2 + x_2^2}{2\sigma^2}\right) \quad (4.1)$$

where σ^2 is the single component variance.

Changing to polar coordinates and using equation 2.17 the expression for $p_2(r)$ (the probability distribution of the length of the vector) becomes

$$p_2(r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) \quad (4.2)$$

This is the well known Rayleigh distribution. From table 2.1, the characteristic function is

$$\Phi(\rho) = \int_0^\infty \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) J_0(\rho r) dr \quad (4.3)$$

figure 4.1 - two-dimensional probability distribution for a Gaussian signal vector:
(a) phase plane representation, (b) conceptual representation.

and equation 11.4.29 of Abramowitz and Stegun (1964) yields

$$\Phi(\rho) = \exp\left(\frac{-\sigma^2 \rho^2}{2}\right) \quad (4.4)$$

Equation 4.2. may be integrated directly to find the cumulative probability function, but to illustrate the use of the characteristic function the appropriate transform from table 2.1 will be utilized. This yields

$$P_2(r) = r \int_0^{\infty} \exp\left(\frac{-\sigma^2 \rho^2}{2}\right) J_1(r\rho) d\rho \quad (4.5)$$

Equation 1 of section 6.618 of Gradshteyn and Ryzhik (1965) then gives

$$P_2(r) = 1 - \exp\left(\frac{-r^2}{2\sigma^2}\right) \quad (4.6)$$

and so for two dimensions the probability density function and the cumulative probability function can be found in closed form.

It is illustrative however, to calculate these functions using the technique developed in chapter 2, and later in this subsection I shall discuss the one-dimension case where the cumulative probability function cannot be found as easily as equation 4.6.

For calculating $p_2(r)$ and $P_2(r)$ the expressions from table 3.1 may be used. Substituting for $\Phi(\rho)$ as defined in equation 4.4 yields

$$p_2(r) = \frac{2r}{R^2} \sum_{k=1}^M \frac{\exp\left(\frac{-\sigma^2 \lambda_k^2}{2R^2}\right)}{J_1^2(\lambda_k)} J_0\left(\lambda_k \frac{r}{R}\right) \quad (4.7)$$

and

$$P_2(r) = \frac{2r}{R} \sum_{k=1}^M \frac{\exp\left(\frac{-\sigma^2 \lambda_k^2}{2R^2}\right)}{\lambda_k J_1^2(\lambda_k)} J_1\left(\lambda_k \frac{r}{R}\right) \quad (4.8)$$

Now it is important that the error in $P_2(r)$ be bounded below the desired level of significance because it is from $P_2(r)$ that performance calculations are made. To do this the technique outlined in section 3.2.2 for the two-dimensional case is used. First the truncation error e_{t2} is determined and made negligible. Then R is adjusted using a numerical convergence check to ensure that e_{R2} is also negligible.

Consider the truncation error. From equation 3.13 we have

$$|e_{t2}| < 1.25 \sum_{k=M+1}^{\infty} \left| \Phi\left(\frac{\lambda_k}{R}\right) \right| \quad (4.9)$$

After substituting the expression for the characteristic function from equation 4.4 we have

$$|e_{t2}| < 1.25 \sum_{k=M+1}^{\infty} \exp\left(\frac{-\sigma^2 \lambda_k^2}{2R^2}\right) \quad (4.10)$$

For the purpose of bounding an approximation for λ_k of $(k - 0.25)\pi$ is appropriate since it is almost equal to λ_k although slightly smaller. Therefore

$$|e_{t2}| < 1.25 \sum_{k=M+1}^{\infty} \exp\left(\frac{-\sigma^2 (k-0.25)^2 \pi^2}{2R^2}\right) \quad (4.11)$$

and it may be shown (Mathematical Details 4.3.1) from equation 4.11 that the following inequality is also satisfied:

$$|e_{t2}| < 1.25 \exp \left(\frac{-\sigma^2 \pi^2}{2R^2} (M + 0.75)^2 \right) \left(\frac{1}{1 - \exp \left(\frac{-\sigma^2 \pi^2}{R^2} (M + 3) \right)} \right) \quad (4.12)$$

Figure 4.2 gives this bound as a function of M for various values of the parameter R/σ , and it can be seen that the series converges very rapidly. The significance of the parameter R/σ will become clear in subsection 4.1.3.

Now we must bound e_{R2} . The procedure for this is to choose a reasonable value for R and then a value for M that guarantees e_{t2} will be negligible. Next, $P_2(r)$ is calculated and R increased and calculated again. The two estimates are then tested for convergence to the level of significance desired.

This procedure need only be done a few times when dealing with the same type of distribution as the value of R that guarantees negligible e_{R2} soon becomes obvious. For all cases for which I have made calculations, choosing an R such that $Q(R)$ (defined by equation 3.19) is much less than the accuracy of the double precision arithmetic has been sufficient to make e_{R2} negligible. For this particular case if we choose

$$R = 10 \sigma \quad (4.13)$$

then

$$|e_{R2}| < Q(R) \approx 2(10)^{-22} \quad (4.14)$$

Since in this example the value for $P_2(r)$ is available in closed form, we may compare the results obtained using the Fourier-

figure 4.2 plot of the upper bound on $|e_{t2}|$ for various values of the parameter R/σ . The points marked $\times$ are the measured errors.

Bessel series with the actual answer. The points "x" in figure 4.2 show the largest error obtained with $R = 10 \sigma$. For this value of R , e_{R2} is negligible so the error is attributed entirely to truncation. It is also seen that, for this case, the bound on e_{t2} (equation 4.12) is relatively tight.

Figure 4.3 shows both $p_2(r)$ and $P_2(r)$ for $\sigma^2 = 1$. Both are indistinguishable from the closed form solutions.

I shall now consider the one-dimensional case. As was described in section 2.2.3, the one-dimensional distribution is a projection of the two-dimensional distribution onto one axis. The negative axis portion of the resulting even distribution is then folded over onto the positive axis. The distribution and probability function for the length of the one-dimensional vector are found using the expressions of table 2.1:

$$P_1(r) = \frac{2}{\pi} \int_0^{\infty} \exp\left(\frac{-\sigma^2 \rho^2}{2}\right) \cos(r\rho) d\rho \quad (4.15)$$

$$P_1(r) = \frac{2}{\pi} \int_0^{\infty} \exp\left(\frac{-\sigma^2 \rho^2}{2}\right) \frac{\sin(r\rho)}{\rho} d\rho \quad (4.16)$$

Equation 4 from section 3.896 and equation 6 from section 3.952 of Gradshteyn and Ryzhik (1965) yields

$$p_1(r) = \frac{1}{\sigma} \frac{2}{\pi} \exp\left(\frac{-r^2}{2\sigma^2}\right) \quad r \geq 0 \quad (4.17)$$

and

$$P_1(r) = 2 \operatorname{erf}(r) - 1 \quad r \geq 0 \quad (4.18)$$

figure 4.3 probability distribution and cumulative probability function for the length of a two-dimensional Gaussian random vector with independent components that have equal variances ($\sigma^2 = 1$) and zero means.

where

$$\text{erf}(r) \triangleq \frac{1}{2\pi\sigma^2} \int_{-\infty}^r \exp\left(\frac{-r^2}{2\sigma^2}\right) dr \quad (4.19)$$

Again we find that $p_1(r)$ is in an elementary mathematical form, however, $P_1(r)$ is not. Calculating $\text{erf}(\cdot)$ is not easy in that one must resort to the use of polynomial approximations, power series or limiting values. With the use of the Fourier-Bessel series approach however, the calculation is easy. From table 3.1 the Fourier series expressions for $p_1(r)$ and $P_1(r)$ are

$$p_1(r) = \frac{2}{R} \sum_{k=1}^M \exp\left(\frac{-\sigma^2 (2k-1)^2 \pi^2}{2R^2}\right) \cos\left(\frac{(2k-1)\pi r}{2R}\right) \quad (4.20)$$

and

$$P_1(r) = \frac{4}{\pi} \sum_{k=1}^M \frac{1}{(2k-1)} \exp\left(\frac{-\sigma^2 (2k-1)^2 \pi^2}{2R^2}\right) \sin\left(\frac{(2k-1)\pi r}{2R}\right) \quad (4.21)$$

The error involved in using equation 4.21 is easily bounded.

The truncation error e_{t1} is obtained from equation 3.14

$$|e_{t1}| \leq \frac{4}{\pi} \sum_{k=M+1}^{\infty} \frac{1}{(2k-1)} \exp\left(\frac{-\sigma^2 (2k-1)^2 \pi^2}{8R^2}\right) \quad (4.22)$$

By the same method used to reduce equation 4.11 it can be shown that

$$|e_{t1}| \leq \frac{4}{\pi(2M+1)} \exp\left(\frac{-\sigma^2 \pi^2 (M+0.5)^2}{2R^2}\right) \left(\frac{1}{1 - \exp\left(\frac{-\sigma^2 \pi^2 (M+0.5)^2}{R^2}\right)} \right) \quad (4.23)$$

Figure 4.4 shows this bound as a function of M for various values of the parameter R/σ . Again it can be seen that the series converges quickly.

e_{R1} may be bounded by appealing to equation 3.23

$$|e_{R1}| \leq Q(R) \quad (4.24)$$

Letting

$$R = 10 \sigma \quad (4.25)$$

we have, as before,

$$|e_{R1}| < Q(R) \approx 2(10)^{-22} \quad (4.26)$$

and this error is negligible with double precision arithmetic.

With R set by equation 4.25 it can be seen from figure 4.4 that as few as 30 terms in the sum will ensure double precision accuracy.

Figure 4.5 shows both $p_1(r)$ and $P_1(r)$ for this example. To obtain $\text{erf}(r)$ from $P_1(r)$ one simply solves equation 4.18 for $\text{erf}(r)$ using the numerical result for $P_1(r)$.

4.1.2 Constant Amplitude Random Phase Vector

In this subsection I deal with a random vector whose length is constant but whose phase is uniformly distributed from $-\pi$ to π . The two-dimensional distribution for this vector is shown in figure 4.6. The distribution is an infinitely thin fence (a circular delta function) about the origin. The radius

figure 4.4. plot of upper bound on $|e_{t1}|$ for various values of the parameter R/σ .

figure 4.5 probability distribution and cumulative probability function for the length of a one-dimensional Gaussian random vector with variance equal to one and a zero mean.

figure 4.6 diagrammatic representations of a constant amplitude uniformly random phase signal: (a) phase plane representation, (b) three-dimensional representation.

is equal to the amplitude of the vector. In polar coordinates the distribution may be written as

$$f(r) = \frac{1}{2\pi a} \delta(r-a) \quad (4.27)$$

where $\delta(\cdot)$ is the delta function and a is the vector amplitude.

From equation 2.17 $p_2(r)$ is

$$p_2(r) = \delta(r-a) \quad (4.28)$$

The characteristic function found with the appropriate transform from table 2.1 is

$$\Phi(\rho) = J_0(a\rho) \quad (4.29)$$

Therefore from the same table

$$P_2(r) = r \int_0^{\infty} J_0(a\rho) J_1(r\rho) d\rho \quad (4.30)$$

Equation 3 of section 6.512 of Gradshteyn and Ryzhik (1965) gives the expected result

$$P_2(r) = u(r-a) \quad (4.31)$$

where $u(\cdot)$ is the unit step function.

Now $p_2(r)$ and $P_2(r)$ are not continuous functions, so one would expect the Fourier-Bessel series to converge rather slowly. Indeed it does, and equation 3.13 cannot be used to bound e_{t2} because the sum diverges. On the positive side, e_{R2} is zero because the distribution does not extend past radius a .

The difficulty in bounding e_{t2} is not a problem in practical

calculations because the constant amplitude random phase vector never appears by itself. It is invariably added to a vector with a continuous distribution, such as Gaussian, that smooths the discontinuities. It is instructive however, to use the calculation technique to calculate $p_2(r)$ and $P_2(r)$ as this provides insight into the effect of the number of terms.

Figures 4.7 and 4.8 show $p_2(r)$ and $P_2(r)$ respectively as calculated using the Fourier-Bessel series and the characteristic function of equation 4.29. The amplitude a was $\sqrt{2}$ and R was set at 10. As M (the number of terms in the series) increases the series approaches the exact function and the oscillations increase in frequency. The approximation for $P_2(r)$ is better than that for $p_2(r)$ because $P_2(r)$ has only a finite discontinuity.

In one-dimension we can again obtain expressions for the distribution and probability function. From table 2.1 and the characteristic function we have

$$p_1(r) = \frac{2}{\pi} \int_0^{\infty} J_0(a\rho) \cos(r\rho) \cdot d\rho \quad (4.32)$$

and

$$P_1(r) = \frac{2}{\pi} \int_0^{\infty} J_0(a\rho) \frac{\sin(r\rho)}{\rho} d\rho \quad (4.33)$$

Equation 2 of section 6.671 and 1 of section 6.693 of Gradshteyn and Ryzhik (1965) result in

$$p_1(r) = \begin{cases} \frac{2}{\pi(a^2 - r^2)^{3/2}} & 0 \leq r < a \\ 0 & r > a \end{cases} \quad (4.34)$$

figure 4.7 probability density function for the length of a constant amplitude random phase vector as calculated by a truncated Fourier-Bessel series. The amplitude is $\sqrt{2}$ and R is 10.

figure 4.8: cumulative probability function for the length of a constant amplitude random phase vector as calculated by a truncated Fourier-Bessel series. The amplitude is $\sqrt{2}$ and R is 10.

$$P_1(r) = \begin{cases} \frac{2}{\pi} \arcsin\left(\frac{r}{a}\right) & 0 \leq r \leq a \\ 1 & r > a \end{cases} \quad (4.35)$$

The Fourier-Bessel series approximations to these functions for various values of M are shown in figures 4.9 and 4.10. Again we see some oscillation about the actual function. The approximation for $P_1(r)$ seems better than that for $P_2(r)$ for the same number of terms, because the one-dimensional probability function does not have a jump discontinuity.

In the next subsection, I shall consider a mixture of constant amplitude random phase vectors with a Gaussian vector.

4.1.3 Gaussian Plus N Constant Amplitude Random Phase Vectors

In this subsection I determine the distribution and probability function of a vector made up of a linear combination of N constant amplitude random phase vectors and a Gaussian vector. This linear combination may be written as

$$\underline{x} = \underline{x}_G + \sum_{n=1}^N \underline{x}_{cn} \quad (4.36)$$

where $\underline{x}_G$ is a two-dimensional circular symmetric Gaussian random vector and $\underline{x}_{cn}$ is the n^{th} constant amplitude random phase vector.

Using the characteristic functions for $\underline{x}_G$ and $\underline{x}_{cn}$ developed in the previous two subsections we may easily write the characteristic function for $\underline{x}$ as

$$\Phi(\rho) = \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) \prod_{n=1}^N J_0(a_n \rho) \quad (4.37)$$

figure 4.9 probability density function for the length of a constant amplitude random phase vector projected onto one dimension as calculated by a truncated Fourier-Bessel series. The amplitude is $\sqrt{2}$ and R is 10.

figure 4.10 cumulative probability function of a constant amplitude random phase vector projected onto one dimension as calculated by a truncated Fourier-Bessel series. The amplitude is $\sqrt{2}$ and R is 10.

where a_n is the amplitude of the n^{th} constant amplitude random phase vector.

With this characteristic function and the expressions for $p_2(r)$ and $P_2(r)$ in table 2.1, we can write

$$p_2(r) = r \int_0^\infty \exp\left(\frac{-\sigma^2 \rho^2}{2}\right) \prod_{n=1}^N J_0(a_n \rho) \rho J_0(r\rho) d\rho \quad (4.38)$$

and

$$P_2(r) = r \int_0^\infty \exp\left(\frac{-\sigma^2 \rho^2}{2}\right) \prod_{n=1}^N J_0(a_n \rho) J_1(r\rho) d\rho \quad (4.39)$$

These integrals are not easily solved even for small N . One approach is to expand the integrals in terms of the moments of the distribution. This leads to series involving generalized Laguerre polynomials (Goldman 1974, Esposito and Wilson 1973). These polynomials must be evaluated recursively making the numerical computations relatively complicated. In addition, the resulting series are difficult to evaluate for some parameter values because the terms become very large and alternate in sign (Esposito and Wilson 1973). Despite the interest and research there is still a need for a simple, yet efficient technique to solve this problem.

I shall now show that the technique developed in chapter 3 meets this need.

Using the Fourier-Bessel series for $p_2(r)$ and $P_2(r)$ from table 3.1

$$p_2(r) = \frac{2r}{R^2} \sum_{k=1}^M \frac{\exp\left(\frac{-\sigma^2 \lambda_k^2}{2R^2}\right)}{J_1^2(\lambda_k)} \prod_{n=1}^N J_0\left(a_n \frac{\lambda_k}{R}\right) J_0\left(\lambda_k \frac{r}{R}\right) \quad (4.40)$$

$$P_2(r) = \frac{2r}{R} \sum_{k=1}^M \frac{\exp\left(\frac{-\sigma^2 \lambda_k^2}{2R^2}\right)}{\lambda_k J_1^2(\lambda_k)} \prod_{n=1}^N J_0\left(a_n \frac{\lambda_k}{R}\right) J_1\left(\lambda_k \frac{r}{R}\right) \quad (4.41)$$

and what remains to be done is to determine R and M so that the error in $P_2(r)$ is negligible.

First I shall consider the truncation error e_{t2} . From equation 3.13 we have

$$|e_{t2}| < 1.25 \sum_{k=M+1}^{\infty} \exp\left(\frac{-\sigma^2 \lambda_k^2}{2R^2}\right) \left| \prod_{n=1}^N J_0\left(a_n \frac{\lambda_k}{R}\right) \right| \quad (4.42)$$

But

$$\left| J_0\left(a_n \frac{\lambda_k}{R}\right) \right| \leq 1 \quad (4.43)$$

so

$$|e_{t2}| < 1.25 \sum_{k=M+1}^{\infty} \exp\left(\frac{-\sigma^2 \lambda_k^2}{2R^2}\right) \quad (4.44)$$

This is the same as the bound obtained for the Gaussian vector (equation 4.10), and so the curves of figure 4.2 may be used to bound the truncation error.

Now I shall consider the error e_{R2} . As described for the Gaussian case, this error is bounded by a numerical technique. The initial choice for R was recommended to be such that $Q(R)$ is smaller than the acceptable error e_{R2} . To guarantee this for the multiple vector case we must appeal to equation 3.22. Taking the Gaussian Q -function out of the sum, we have

$$Q(R) \leq Q_G(R - x_0) + \sum_{n=1}^N Q_n(x_{n-1} - x_n) \quad (4.45)$$

where $x_0 > x_1 > x_2 > \dots > x_k = 0$. If the difference $x_{n-1} - x_n$ is made equal to a_n (the amplitude of the n^{th} vector) then the right hand sum is zero and we have

$$Q(R) \leq Q_G\left(R - \sum_{n=1}^N a_n\right) \quad (4.46)$$

If we set

$$R = \sum_{n=1}^N a_n + 10\sigma \quad (4.47)$$

then we are assured that

$$Q(R) \leq Q_G(10\sigma) \approx 2(10)^{-22} \quad (4.48)$$

Choosing R according to equation 4.47 has proven sufficient to make the error e_{R2} negligible for double precision calculations.

It should be noted that R is increased over that used for the Gaussian case by the sum of the amplitudes a_n . This must be reflected in the choice of M . Referring back to figure 4.2, we see that as the ratio R/σ increases more terms are required to maintain accuracy. Convergence is still fast, however, even for relatively large values of this ratio. In terms of signals a ratio R/σ of 50 corresponds to a constant amplitude random phase signal to Gaussian signal power ratio of 29dB. Only 140 terms are required at this power ratio to ensure that the error is smaller than 10^{-15} .

The effect that the number of constant amplitude random phase vectors has on the shape of the probability distribution is of interest. Figure 4.11 shows the distribution for N equal

figure 4.11 the probability density function of the length of a signal vector made up of the sum of N random phase constant amplitude signal vectors and a Gaussian signal vector. The power ratio between the constant amplitude vectors and the Gaussian vector is 20 dB and the total average power is normalized to one.

amplitude random phase vectors added to a Gaussian vector. The means of the distributions were normalized to one to facilitate the comparison, and the ratio of the power of the constant amplitude random phase signals to the Gaussian signal power was maintained at 100 (20 dB). At this power ratio a single constant amplitude signal causes the distribution to be highly concentrated about the amplitude of that signal. As a result of the two-dimensional central limit theorem (Lord 1954) the distribution approaches a Rayleigh distribution as more constant amplitude signals are added. It only takes five constant amplitude random phase vectors (signals) to produce a distribution that appears to be Rayleigh (except, of course, on the tail).

The significance of the different distribution shapes may be illustrated by assuming the constant amplitude random phase vectors are the result of some kind of interference (later we shall see they are used to model cochannel interference in digital communication systems). Often performance calculations are made using the assumption that the corrupting signals are Gaussian. If, however, the corrupting signals are a mixture of Gaussian noise and interference the distributions can deviate significantly from the Gaussian-Rayleigh assumption.

Returning now to the discussion of the mathematical form of the distributions I shall discuss the one-dimensional distribution $p_1(r)$ and probability function $P_1(r)$. The integrals for these functions, like those for the two-dimensional case, are difficult to solve. From table 2.1

$$p_1(r) = \frac{2}{\pi} \int_0^{\infty} \exp\left(\frac{-\sigma^2 \rho^2}{2}\right) \prod_{n=1}^N J_0(a_n \rho) \cos(r\rho) d\rho \quad (4.49)$$

and

$$P_1(r) = \frac{2}{\pi} \int_0^{\infty} \exp\left(\frac{-\sigma^2 \rho^2}{2}\right) \prod_{n=1}^N J_0(a_n \rho) \frac{\sin(r\rho)}{\rho} d\rho \quad (4.50)$$

The Fourier-Bessel series for $p_1(r)$ and $P_1(r)$ from table

3.1 are

$$p_1(r) = \frac{2}{R} \sum_{k=1}^M \exp\left(\frac{-\sigma^2 (2k-1)^2 \pi^2}{2R^2}\right) \prod_{n=1}^N J_0\left(\frac{a_n (2k-1)\pi}{R}\right) \cos\left(\frac{(2k-1)\pi r}{2R}\right) \quad (4.51)$$

and

$$P_1(r) = \frac{4}{\pi} \sum_{k=1}^M \frac{1}{(2k-1)} \exp\left(\frac{-\sigma^2 (2k-1)^2 \pi^2}{2R^2}\right) \prod_{n=1}^N J_0\left(\frac{a_n (2k-1)\pi}{R}\right) \sin\left(\frac{(2k-1)\pi r}{2R}\right) \quad (4.52)$$

R and M must be determined to ensure that the error in $P_1(r)$ is negligible. Again by using the fact that

$$|J_0(x)| < 1 \quad (4.53)$$

the truncation error bound can be shown to satisfy equation 4.23 and therefore figure 4.4 can be used to bound this error. Also, for the one-dimensional case we have (equation 3.21)

$$|e_{R1}| \leq Q(R) \quad (4.54)$$

and from equation 4.48 we see that

$$Q(R) < Q_G(10\sigma) \approx 2(10)^{-22} \quad (4.55)$$

if

$$R = \sum_{n=1}^N a_n + 10 \sigma \quad (4.56)$$

Therefore, for the one-dimensional case, R may be chosen in the same way as it was for the two-dimensional case.

Figure 4.12 shows $p_1(r)$ for different numbers of constant amplitude vectors (signals). Again the total signal power is normalized to one and the constant amplitude to Gaussian signal power ratio is 20 dB. For a single constant amplitude vector the distribution is concentrated about its amplitude. As the number increases however, the distribution assumes the Gaussian shape. This is a result of the two-dimensional distribution becoming Gaussian which in turn projects to a Gaussian distribution in one-dimension. The importance of these distributions will become clear in the next section.

4.2 Performance Calculations -

In this section the results of chapters 2 and 3 and section 4.1 are combined for the purpose of making performance calculations for various systems.

First the problem of calculating the probability of error for a CPSK communication system, operating in Gaussian noise and cochannel interference, is considered. Probability of error curves are found for various levels of interference and numbers of interferers. It is shown how the Rosenbaum and Glave bounds

figure 4.12 the probability density function of the length of a signal vector projected onto one dimension. The signal vector is made up of the sum of N random phase constant amplitude signal vectors and a Gaussian signal vector. The power ratio between the constant amplitude vectors and the Gaussian vector is 20 dB and the total average power is normalized to one.

(1974 and 1975) for M-ary CPSK operating in peak limited interference can be easily calculated using the technique of chapter 3.

The second problem to be considered is the calculation of probability of error for multi-level ASK communication systems. This is not an easy problem even when the corrupting signal is Gaussian noise alone. With the technique of chapter 3 however, the probability of error is easily determined for both this and the more general case of Gaussian noise plus cochannel interference.

Finally, we consider the classical radar/sonar problem of detecting a constant amplitude random phase signal in the presence of Gaussian noise. To do this the so-called radar Q-function must be calculated. This is difficult and for the most part approximations are used. If a high degree of accuracy is required one must resort to mathematically complicated series such as those used by Goldman (1974) and Mitchell and Walker (1971). These series however suffer convergence problems for large signal values. Fortunately, the technique developed in chapter 3 can be used to evaluate the Q-function to a high degree of accuracy simply and efficiently.

4.2.1 Probability of Error for Phase Shift Keyed Communications

Systems

As a first example I shall consider the calculation of the probability of error and probability of error bounds for coherent

phase shift keyed (CPSK) communication systems. The corrupting signal will be made up of N cochannel interferers and Gaussian noise. The cochannel interferers will be modelled as constant amplitude uniformly random phase circular symmetric vectors (Rosembaum 1970, Goldman 1974). The received signal will be represented in complex envelope notation as follows

$$\bar{s}_r = b \exp(j \phi_\ell) + \sum_{i=1}^k a_i \exp(j \gamma_i) + \bar{n} \quad (4.57)$$

where a_i is the amplitude of the i^{th} interferer and γ_i the phase. The phases are uniformly distributed from $-\pi$ to π and independent of each other. $\bar{n}$ is the Gaussian noise term and b the amplitude of the message signal. The message information is carried in the phase ϕ_ℓ defining the signal point in the phase plane. For a given message the corrupting signal results in a circular symmetric vector radiating from the signal point. An error occurs if this interfering vector crosses a decision boundary.

For binary CPSK the message phases are $-\pi$ to π . The decision boundary and probability distributions for this case are illustrated in figure 4.13. Also shown in the figure is the process of projecting the distribution onto the one-dimensional signal vector. The probability of error is the probability that the resulting signal vector crosses the decision boundary which may be expressed as

$$P_e = \frac{1}{2} (1 - P_1(b)) \quad (4.58)$$

where $P_1(b)$ is the probability that the length of the interfering

figure 4.13 signal points and probability distributions for binary CPSK.. In this figure it is assumed that phase $\phi_c = \pi$ was transmitted. The two-dimensional circular symmetric probability distribution centered about message point b is projected onto one dimension

vector projected onto the signal vector will be less than or equal to b . The problem is to determine $P_1(b)$ given the nature of the interference. This is done by finding the characteristic function of the interfering distribution and then using the Fourier series. From equation 4.50

$$P_1(b) = \frac{2}{\pi} \int_0^{\infty} \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) \prod_{n=1}^N J_0(a_n \rho) \frac{\sin(b\rho)}{\rho} d\rho \quad (4.59)$$

and the Fourier series approximation for $P_1(b)$ (see equation 4.52) is

$$P_1(b) = \frac{4}{\pi} \sum_{k=1}^M \frac{1}{(2k-1)} \exp\left(-\frac{\sigma^2 (2k-1)^2 \pi^2}{2R^2}\right) \prod_{n=1}^N J_0\left(\frac{a_n (2k-1)\pi}{R}\right) \sin\left(\frac{(2k-1)\pi b}{2R}\right) \quad (4.60)$$

where R and M may be determined as was shown in section 4.1.3, namely

$$R = \sum_{n=1}^N a_n + 10 \sigma \quad (4.61)$$

and M chosen from figure 4.4.

Figure 4.14 shows the probability of error for a binary CPSK system corrupted by N equal amplitude cochannel interferers and Gaussian noise. The interference to noise ratio defined as

$$\text{INR}_{\text{dB}} = 10 \log\left(\frac{Na^2}{2\sigma^2}\right) \quad (4.62)$$

is 0 dB.

The largest system degradation occurs in the transition from zero to one cochannel interferer. As N increases the

figure 4.14 probability of error for a binary CPSK system corrupted by Gaussian noise and N cochannel interferers. $\text{INR}_i = 0$ dB.

probability distribution of the cochannel interferers approaches a two-dimensional Gaussian distribution as a result of the two-dimensional central limit theorem (Lord 1954). Thus for large N and an INR equal to 0 dB, the probability of error approaches that of a system corrupted by Gaussian noise alone but with the signal-to-noise ratio reduced by 3 dB. Similar results are obtained for other values of INR with the limiting case defined by the corresponding equivalent Gaussian distribution.

Figure 4.15 shows the probability of error for a binary CPSK system for a single cochannel interferer and various values of the interference-to-noise ratio. As the interference-to-noise ratio increases the probability of error increases. The increase in error probability is most rapid when the interference-to-noise ratio is of the same order as the signal-to-noise ratio.

Exact calculations of the probability of error are useful if the corrupting distributions are known. Often we do not have the necessary information however, and we must resort to establishing reasonable bounds based on partial knowledge. Rosebaum and Glave (1974) have developed an upper bound on the probability of error for CPSK systems corrupted by amplitude limited circular symmetric interference and Gaussian noise. A parameter used in their bound is the peak to r.m.s. value of the interference. This ratio is called the peak factor and is defined as

$$PF \triangleq 20 \log \left(\frac{R_I}{\sigma_I} \right) \quad (4.63)$$

figure 4.15 probability of error for a binary CPSK system corrupted by Gaussian noise and one cochannel interferer. INR is the interference-to-noise ratio.

where R_I is the peak value and σ_I the r.m.s. value of the interference.

To develop the bound the probability of error for a binary CPSK system operating in Gaussian noise was determined, conditioned on the magnitude of the interfering vector. Then the distribution of the magnitude that maximized the probability of error, subject to the interference having a specific peak factor, was found. This distribution was¹

$$p(r) = p \delta(r - R_I) + (1 - p) \delta(r) \quad (4.64)$$

where

$$p = \frac{\sigma_I}{R_I} = \frac{1}{\frac{PF}{10^{20}}} \quad (4.65)$$

The characteristic function for this distribution is

$$\Phi(\rho) = p J_0(R_I \rho) + (1 - p) \quad (4.66)$$

Therefore, when added to Gaussian noise, the characteristic function for the distribution that maximizes the probability of error for a given peak factor is

$$\Phi(\rho) = \exp\left(\frac{-\sigma^2 \rho^2}{2}\right) (p J_0(R_I \rho) + (1 - p)) \quad (4.67)$$

Using equation 4.58 and the expression for $P_1(b)$ from table 2.1 we may express the maximum probability of error as

¹ There are conditions to be met for this to be the maximizing distribution but they do not affect this discussion. The interested reader is referred to the Rosenbaum and Glave (1974) paper.

$$P_{\text{emax}} = \frac{p}{2} \left(1 - \frac{2}{\pi} \int_0^{\infty} \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) J_0(R_I \rho) \frac{\sin(r\rho)}{\rho} d\rho \right) + \frac{(1-p)}{2} \left(1 - \frac{2}{\pi} \int_0^{\infty} \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) \frac{\sin(r\rho)}{\rho} d\rho \right) \quad (4.68)$$

The first term here, except for the p , is the probability of error for a single interferer with a peak factor of 0 dB, and the second term, except for $(1-p)$, is the probability of error for Gaussian noise alone. Therefore, P_{emax} is a linear combination of the probabilities of error for Gaussian noise plus a single interferer and Gaussian noise alone. The coefficients for the linear combination depend upon the peak factor of the amplitude limited interference. This separation of probabilities permits one to make rapid calculations of the Rosenbaum and Glave upper bound for different peak factors knowing only the probability of error for a peak factor of 0 db and a peak factor of ∞ (Gaussian noise alone). This procedure will be demonstrated presently, but first I will outline the application of the above approach to an extended form of the bound.

Glave and Rosenbaum (1975) extended their bound to include two identifiable amplitude limited circular symmetric interferers. Here a peak factor was defined for each interference.

$$PF_1 = 20 \log\left(\frac{R_1}{\sigma_1}\right) \quad (4.69)$$

$$PF_2 = 20 \log\left(\frac{R_2}{\sigma_2}\right) \quad (4.70)$$

The probability of error was found conditioned on the amplitude distribution of both interferers. The joint distribution of the amplitudes was then found that maximized the probability of error under the constraint of PF_1 and PF_2 being satisfied. Since the interferers were assumed to be independent the joint distribution of the amplitude was the product of the two distributions¹

$$p_{r1}(r) = p_1 \delta(r - R_1) + (1 - p_1) \delta(r) \quad (4.71)$$

$$p_{r2}(r) = p_2 \delta(r - R_2) + (1 - p_2) \delta(r) \quad (4.72)$$

where

$$p_1 = \frac{\sigma_1}{R_1} = \frac{1}{\frac{PF_1}{10^{20}}} \quad (4.73)$$

and

$$p_2 = \frac{\sigma_2}{R_2} = \frac{1}{\frac{PF_2}{10^{20}}} \quad (4.74)$$

The bound derived in this manner takes advantage of knowing the interference consists of two or more interferers, and so the bound is tighter than the single-interferer bound. This bound may be decomposed in a manner similar to that of equation 4.68. The product of the characteristic functions for $p_{r1}(r)$ and $p_{r2}(r)$ is

¹ Again there are conditions to be met for this distribution to be the maximizing distribution and the interested reader is referred to the Glave and Rosenbaum (1975) paper.

$$\begin{aligned} \Phi(\rho) = & p_1 p_2 J_0(R_1 \rho) J_0(R_2 \rho) + p_1 (1 - p_2) J_0(R_1 \rho) \\ & + (1 - p_1) p_2 J_0(R_2 \rho) + (1 - p_1) (1 - p_2) \end{aligned} \quad (4.75)$$

By inspection, the probability of error due to this interference and Gaussian noise is a linear combination of four probabilities of error: P_e for two interferers with amplitudes R_1 and R_2 , P_e for a single interferer of amplitude R_1 , P_e for a single interferer of amplitude R_2 , and P_e for Gaussian noise alone. The multiplicative coefficients are $p_1 p_2$, $p_1 (1 - p_2)$, $(1 - p_1) p_2$, and $(1 - p_1) (1 - p_2)$ respectively. Therefore, the double interferer bound for different values of PF_1 and PF_2 can be found from a linear combination of four probability of error curves. Each of these curves is easily determined using the calculation technique described in chapter 3.

Rosenbaum and Glave's single and double interference bounds may be used to obtain upper bounds for M-ary CPSK systems as the probability of error for these schemes must be less than twice that found for binary CPSK (Rosenbaum and Glave 1974). The signal-to-noise ratio however, must be scaled by an amount

$$CF \triangleq -20 \log \left(\sin \left(\frac{\pi}{M} \right) \right) \quad (4.76)$$

Figure 4.16 shows the process of obtaining the single interference bound for a M-ary CPSK scheme. In this example $INR = 0$ dB and $PF = 20$ dB. First a linear combination of the two P_e curves for Gaussian noise and Gaussian noise plus a single interferer is found. The coefficients for this linear combination

figure 4.16 illustration showing how the single interferer probability of error bound is obtained for M-ary CPSK when INR = 0 dB and PF = 20 dB.

are determined from the peak factor. The amplitude limit of the interference (i.e., the algebraic sum of the components) determines the INR. The result is the dashed curve and this is the bound for the binary CPSK. Doubling this bound and scaling the signal-to-noise ratio axis results in the M-ary CPSK bound.

This example illustrated the usefulness of circular symmetric probability distribution theory in solving a particular detection problem in coherent phase shift keyed communication systems. The practical necessity of calculating probabilities has limited this approach in the past. The calculation technique presented in chapter 3 permits this problem to be solved easily and, what is perhaps more important, within a general framework that may be used for many detection problems.

4.2.2 Probability of Error for Amplitude Shift Keyed Communication Systems

The second example is that of calculating the probability of error for multi-level amplitude shift keyed (ASK) schemes operating in the presence of Gaussian noise and cochannel or adjacent channel interference. The interference is again modelled by I constant amplitude uniformly random phase vectors. In complex envelope notation the received signal $\bar{s}_r$ is

$$\bar{s}_r = b_\ell \exp(j\phi) + \sum_{i=1}^I a_i \exp(j\gamma_i) + \bar{n} \quad (4.77)$$

where a_i is the amplitude and γ_i the phase of the i^{th} cochannel interferer. The γ_i 's are uniformly distributed from $-\pi$ to π and

independent of each other, $\bar{n}$ is the Gaussian noise, b_ℓ and ϕ are the amplitude and phase of the message signal, and ϕ is a random variable with a uniform distribution from $-\pi$ to π .

The message information is carried in the amplitude b_ℓ which defines a message radius in the phase plane. Since all three components of $\bar{s}_r$ (the signal, the interference, and the noise) have circular symmetric distributions so does $\bar{s}_r$. In the absence of interference and noise the signals are rings in the phase plane as shown in figure 4.17. To be specific the separation between signal amplitudes is chosen to be a constant D . The decision boundaries for the minimum distance decision rule are shown. They lie one-half way between the signal radii. If $P_2^\ell(r)$ is defined as the probability that $|\bar{s}_r|$ is less than or equal to r given the signal ℓ is transmitted then, as shown in subsection 4.3.2, the probability of error is

$$P_e = \frac{N-1}{N} - \frac{1}{N} \sum_{\ell=1}^{N-1} P_2^\ell \left(\frac{D}{2} (2\ell - 1) \right) + \frac{1}{N} \sum_{\ell=2}^N P_2^\ell \left(\frac{D}{2} (2\ell - 3) \right) \quad (4.78)$$

where N is the number of signal levels. It is assumed that each signal amplitude is equally likely.

It now remains to determine the probability $P_2^\ell(r)$ for arbitrary ℓ and r . Using equation 4.37, the characteristic function for the distribution of $|\bar{s}_r|$ given ℓ is

$$\Phi(\rho) = J_0((\ell - 1)D\rho) \prod_{i=1}^I J_0(a_i \rho) \exp\left(\frac{-\sigma^2 \rho^2}{2}\right) \quad (4.79)$$

figure 4.17 signal radii and decision radii for 4-level ASK

From table 2.1

$$P_2^{\ell}(r) = r \int_0^{\infty} \phi(\rho) J_1(r\rho) d\rho \quad (4.80)$$

where $\phi(\cdot)$ is defined by equation 4.79. For arbitrary ℓ this integral cannot be evaluated, but with the use of the Fourier-Bessel series it can be determined numerically. From table 3.1 we have

$$P_2^{\ell}(r) = \frac{2r}{R} \sum_{k=1}^M \frac{\phi\left(\frac{\lambda_k}{R}\right)}{\lambda_k J_1^2(\lambda_k)} J_1\left(\lambda_k \frac{r}{R}\right) \quad (4.81)$$

where λ_k is the k^{th} zero of $J_0(\cdot)$ and $\phi(\cdot)$ is defined by equation 4.79.

R and M must be determined so that the error in $P_2^{\ell}(r)$ is within the desired bound. This problem is identical to the two-dimensional problem addressed in subsection 4.1.3. It was found there that the error was negligible if R was chosen such that

$$R_{\ell} = (\ell - 1) D + \sum_{i=1}^I a_i + 10 \sigma \quad (4.82)$$

and M was chosen from figure 4.2 according to the accuracy desired.

Figures 4.18, 4.19 and 4.20 show the probability of error which results for a multi-level ASK system operating in Gaussian noise and cochannel interference. The following definitions for the signal-to-noise ratio (SNR) and the interference-to-noise ratio (INR) are used:

$$\text{SNR} = \frac{1}{2N\sigma^2} \sum_{\ell=1}^N b_{\ell}^2 \quad (4.83)$$

figure 4.18 probability of error for N-level ASK corrupted by Gaussian noise.

figure 4.19 probability of error for 2-level ASK corrupted by a single cochannel interferer and Gaussian noise.

figure 4.20 probability of error for 2-level ASK corrupted by
 I equal amplitude cochannel interferers. INR = 0 dB.

$$\text{INR} = \frac{1}{2\sigma^2} \sum_{i=1}^L a_i^2 \quad (4.84)$$

Figure 4.18 shows the probability of error for multi-level ASK when corrupted by Gaussian noise only. Changing from 2 levels to 3 levels causes a SNR penalty of about 5 dB but each message contains more information. Changing to 4 levels causes an additional 3 dB loss in SNR. Comparing the 4-level scheme with the 2-level scheme for equal information transmission, it is evident that it is better to use 2 levels and transmit two bits than to use 4 levels. The dashed curve shows the probability of having an error in a sequence of two messages for a 2-level scheme.

Figure 4.19 illustrates the loss in performance for a 2-level ASK scheme operating in the presence of Gaussian noise and a single cochannel interferer. As the interference increases the probability of error curve moves to the right and becomes steeper. This reflects the fact that the interferer is spreading the distribution about the signal point in a ring and that the tail outside this ring is steeper than the tail of the Gaussian distribution at these signal-to-noise ratios.

In figure 4.20 the effect on a 2-level ASK scheme of increasing the number of equal amplitude interferers is shown. For the same interference-to-noise ratio the probability of error curve moves to the right as the number of interferers increases. This degradation is limited however, because the distribution of the interferers approaches a Gaussian distribution due to the two-dimensional central limit theorem (Lord 1954). Therefore, as I

gets large the probability of error curve is the same as that for no interference but shifted to the right by the number of dB indicated by the interference-to-noise ratio. For example, in figure 4.20 the interference-to-noise ratio is 0 dB. This indicates that the noise and interference have equal average power. Therefore in the limiting case (large I) the no interference curve should be shifted to the right by 3 dB. This curve is shown by a dashed line in the figure.

This concludes the discussion of ASK communication systems. In summary, I have shown how the theory of circular symmetric probability distributions and the calculation technique developed in chapter 3 are used to calculate the probability of error for multi-level ASK operating in the presence of Gaussian noise and cochannel interference. This discussion was not intended to provide a detailed analysis of the problem but to illustrate the utility of the new approach in attacking it. It is interesting to note that a literature search has not revealed any published papers that treat this specific problem. This is probably due to the lack of ability to calculate the necessary probabilities since the problem is of practical significance.

I shall now proceed to the final example, a classical radar/sonar detection problem.

4.2.3 A Classical Radar/Sonar Detection Problem

In this subsection I consider the problem of detecting a constant amplitude random phase target signal in the presence of Gaussian noise. In the phase plane the situation may be represented as shown in figure 4.21. If there is no target the distribution will be circular symmetric Gaussian. If a target is present the addition of the constant amplitude random phase signal will cause the distribution to move out radially from the origin. This distribution is known as a Rician distribution.

For the purpose of deciding the presence or absence of a target a circular threshold is established with radius T_h . If the received signal vector has a length greater than T_h a target is declared present and if the length is less than T_h no target is declared. T_h is set to provide a given probability of false alarm P_F (the probability that the signal vector length exceeds T_h when no target is present). The reason for this decision rule is discussed in chapter 6 of part 2.

The performance measure is the probability of detection P_D which is the probability that the signal vector length exceeds T_h given the target is present. In order to determine P_D we must calculate the probability of being outside the circular contour for the target present condition.

First however, I shall consider P_F and the setting of the threshold. Since the distribution is Gaussian in the absence of a target, P_F may be found directly from equation 4.6

figure 4.21 phase plane representation of the probability distributions for the target absent case and the target present case.

$$P_F = \exp \frac{-T_h^2}{2\sigma^2} \quad (4.85)$$

Solving for T_h we have

$$T_h = \left(2\sigma^2 \ln \left(\frac{1}{P_F} \right) \right)^{1/2} \quad (4.86)$$

Therefore to obtain a given probability of false alarm, T_h should be set according to equation 4.86, and P_D is determined for that threshold setting.

Since the target signal has constant amplitude and uniform random phase, equation 4.39 with $N=1$ may be used to find P_D , yielding

$$P_D = 1 - T_h \int_0^\infty \exp \left(-\frac{\sigma^2 \rho^2}{2} \right) J_0(a\rho) J_1(T_h \rho) \rho \, d\rho \quad (4.87)$$

where a is the amplitude of the target signal. In subsection 4.3.3 it is shown that the right side of equation 4.87 is the radar Q-function although this is not obvious by inspection.

All that remains now is to use the Fourier-Bessel series with appropriate values of R and M to calculate P_D . For this case the Fourier-Bessel series is

$$P_D = 1 - \frac{2 T_h}{R} \sum_{k=1}^M \frac{\exp \left(-\frac{\sigma^2 \lambda_k^2}{2R^2} \right)}{\lambda_k J_1^2(\lambda_k)} J_0 \left(\frac{a \lambda_k}{R} \right) J_1 \left(\lambda_k \frac{T_h}{R} \right) \quad (4.88)$$

To make the error in P_D negligible it is sufficient to choose R such that

$$R = a + 10 \sigma \quad (4.89)$$

and to choose M from figure 4.2 to obtain the desired accuracy.

Figure 4.22 shows the resulting probability of detection for the constant amplitude random phase target as a function of signal-to-noise ratio (SNR) with P_F as a parameter. SNR is defined as

$$\text{SNR}_{\text{dB}} = \triangle 10 \log \left(\frac{a^2}{2\sigma^2} \right) \quad (4.90)$$

These results are identical to those published by Meyer and Mayer (1973).

As SNR increases the probability of detection increases. There is a trade-off between false alarm probability and detection probability. If we are willing to accept a P_F of 10^{-2} P_D is higher for a given SNR than if we are only willing to accept a P_F of 10^{-10} .

The Fourier-Bessel series provides a simple, well-behaved method for calculating the radar Q-function. The accuracy of the result can be adjusted to suit the application. As few as 140 terms provide double precision accuracy for signal-to-noise ratios as high as 30 dB.

On a general note, it is evident that the addition of additional constant amplitude random phase signals does not make the detection problem any more difficult to solve if the techniques outlined in this example are used. The only complication that arises is in the determination of the threshold T_h . If the corrupting signal is a combination

figure 4.22 probability of detecting a constant amplitude random phase target in Gaussian noise.

of Gaussian noise and constant amplitude random phase vectors then T_h cannot be expressed as a function of P_F but must be determined iteratively. This procedure is not difficult and was used successfully to solve the radar problem discussed in part 2 of this thesis.

Those readers who are familiar with the classical radar detection theory as developed by Marcum (1960) and Swerling (1960) and who are interested in its connection with the theory presented here should refer to thesis Appendix A. In this appendix the topic of noncoherent integration is discussed using the radar oriented notation of part 2. Those readers not familiar with the theory but interested in the connection are encouraged to read chapters 6 and 7 of part 2 before proceeding to the appendix.

4.3 Mathematical Details

This section presents the mathematical detail leading to some of the equations used in chapter 4. Readers not interested in this detail may disregard the section to preserve the continuity of the thesis discussion.

4.3.1 Bounding of e_{t2}

From equation 4.11 we have

$$|e_{t2}| < 1.25 \sum_{k=M+1}^{\infty} \exp(-w(k - \frac{1}{4})^2) \quad (4.91)$$

where

$$w = \frac{\sigma^2 \pi^2}{2R} \quad (4.92)$$

Letting

$$i = k - (M + 1) \quad (4.93)$$

the sum becomes

$$|e_{t2}| < 1.25 \sum_{i=0}^{\infty} \exp(-w (M + \frac{3}{4} + i)^2) \quad (4.94)$$

By multiplying out the square, and taking terms independent of i outside the sum, we obtain

$$|e_{t2}| < 1.25 \exp(-w (M + \frac{3}{4})^2) \sum_{i=0}^{\infty} \exp(-w (i^2 + 2iM + \frac{3}{2}i)) \quad (4.95)$$

But

$$\exp(-w (2M + \frac{3}{2}) i) > \exp(-w (i^2 + 2iM + \frac{3}{2}i)) \quad (4.96)$$

so

$$|e_{t2}| < 1.25 \exp(-w (M + \frac{3}{4})^2) \sum_{i=0}^{\infty} \exp(-w 2 (M + 3) i) \quad (4.97)$$

The sum is now in the form of a geometric series, therefore

$$|e_{t2}| < 1.25 \exp(-w (M + \frac{3}{4})^2) \left(\frac{1}{1 - \exp(-w 2 (M + 3))} \right) \quad (4.98)$$

which is the desired bound.

4.3.2 Probability of Error For Multi-level ASK

Let P_{el} be the probability of error given that the ℓ^{th} level was transmitted. Then by inspection of figure 4.17 we have

$$P_{e1} = 1 - P_2^1 \left(\frac{D}{2} \right) \quad (4.99)$$

$$P_{e2} = P_2^2 \left(\frac{D}{2} \right) + 1 - P_2^2 \left(\frac{3D}{2} \right) \quad (4.100)$$

$$P_{ek} = P_2^k \frac{(2k-3)D}{2} + 1 - P_2^k \frac{(2k-1)D}{2} \quad (4.101)$$

$$P_{eN} = P_2^N \frac{(2N-3)D}{2} \quad (4.102)$$

Assuming that each level occurs with a probability of $\frac{1}{N}$, the total probability of error may be obtained by the sum of the P_{ek} 's divided by N .

$$P_e = \frac{1}{N} \sum_{\ell=1}^N P_{e\ell} = \frac{1}{N} \sum_{\ell=2}^N P_2^{\ell} \left(\frac{(2\ell-3)D}{2} \right) + \frac{N-1}{N} - \frac{1}{N} \sum_{\ell=1}^{N-1} P_2^{\ell} \left(\frac{(2\ell-1)D}{2} \right) \quad (4.103)$$

which is the desired result.

4.3.3 The Radar Q-function

Since

$$P_D = 1 - P_2(r) \quad (4.104)$$

where

$$r = T_h \quad (4.105)$$

we have from equation 4.87

$$P_2(r) = r \int_0^{\infty} \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) J_0(a\rho) J_1(r\rho) d\rho \quad (4.106)$$

Therefore

$$p_2(r) = \frac{dP_2(r)}{dr} = r \int_0^{\infty} \rho \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) J_0(a\rho) J_0(r\rho) d\rho \quad (4.107)$$

This is Weber's second exponential integral, the solution to which is in Watson (1962, page 395).

$$p_2(r) = \frac{r}{\sigma^2} \exp\left(-\frac{(a^2 + r^2)}{2\sigma^2}\right) I_0\left(\frac{ar}{\sigma^2}\right) \quad (4.108)$$

where $I_0(\cdot)$ is the modified Bessel function of order zero.

Therefore we have

$$P_2(r) = \int_0^r \frac{x}{\sigma^2} \exp\left(-\frac{(a^2 + x^2)}{2\sigma^2}\right) I_0\left(\frac{ax}{\sigma^2}\right) dx \quad (4.109)$$

or

$$P_2(r) = 1 - \int_r^{\infty} \frac{x}{\sigma^2} \exp\left(-\frac{(a^2 + x^2)}{2\sigma^2}\right) I_0\left(\frac{ax}{\sigma^2}\right) dx \quad (4.110)$$

The integral in equation 4.110, which is equal to P_D , is the radar Q-function (Meyer and Mayer 1973).

CHAPTER 5

CONCLUSIONS TO PART 1

Part 1, containing chapters 2 through 5, dealt with the general detection problem. In chapter 2 the problem was developed and it was found that merging the concepts of complex signal notation, phase plane representations and circular symmetric probability theory yielded a general framework for solution. This framework provides a conceptually simple method for investigating detection problems and obtaining mathematical expressions for performance measures.

Despite the ease in obtaining these mathematical expressions they are invariably hard to evaluate by conventional techniques. In chapter 3 I presented a new technique that was not only conceptually simple but very efficient. It employed the characteristic function and a Fourier-Bessel series to estimate the desired probabilities to an arbitrary degree of accuracy. Unlike other techniques that have been used in the past this technique needs no special considerations when implemented. The terms in the series are well behaved and the number of terms needed is well defined.

As a result of combining the general approach with an efficient calculation technique it is possible to solve many classical detection problems as well as those of current interest. This was demonstrated in chapter 4 with several examples. The

first example was the calculation of the probability of error of a PSK communication system operating in cochannel interference and Gaussian noise. It was found that the presence of the cochannel interferers caused the error probability to increase but not so much as it would have if the interference had been Gaussian. In addition it was shown how the Rosenbaum and Glave upper bounds for probability of error could be simply determined using the new calculation technique. It was found that the bounds for arbitrary peak to r.m.s. interference ratios could be determined by linearly combining two error curves: one for no interference and one for a peak to r.m.s. ratio of 0 dB. Upper bounds on the probability of error for M-ary CPSK were found by simply doubling the upper bounds for binary CPSK.

This example illustrated the reason for considering the single dimension symmetric distribution.

The second example examined a two-dimension detection problem: that of calculating the probability of error for a multi-level ASK communication system operating in cochannel interference and Gaussian noise. This is a difficult problem when approached in a conventional manner, however, with the use of the new calculation technique it is easily solved. Again it was determined that the interference increased the probability of error, but not as much as it would have if the interference had been Gaussian.

In the third example I moved away from communications applications to the classical radar/sonar problem of detecting

a constant signal in Gaussian noise. This required the calculation of the notorious radar Q-function. It was shown how this function can be calculated easily and efficiently using the new calculation technique. Through demonstrating the usefulness of the approach to this classical problem it was determined that detection problems involving any number of Gaussian and constant amplitude uniformly random phase signals could be solved with similar ease.

In summary, the unified approach to detection problems presented in Part 1 provides a conceptually simple, mathematically concise and numerically efficient method of solving a wide variety of detection problems.

This concludes the discussion of the general approach to detection problems. The reader who is interested in the ground-clutter-in-radar problem described in chapter 1 should now proceed to Part 2.

PART 2

GROUND CLUTTER IN AIRSPACE SURVEILLANCE RADARS
AND THE COHERENT CLUTTER MAP

PART 2

This second part of the thesis, containing chapters 6 through 10, deals with the problem of ground clutter in step-scan airspace surveillance radars. Chapter 6 describes the signal models, receiver structures, and decision processes that are consistent with the radar application. In chapter 7 the signal models are further tailored to the problem of aircraft detection in the presence of ground clutter and an optimum decision process is derived. With the optimum processor as a basis, chapter 8 compares the performance of the two general classes of conventional clutter suppression schemes: variable thresholding and pulse cancellation. Following this, chapter 9 describes in detail a practical approach to the optimum, the coherent clutter map, and compares its performance with that of the optimum processor. Finally chapter 10 compares the conventional clutter suppression schemes with the coherent clutter map and makes conclusions and recommendations based on this comparison.

CHAPTER 6

SIGNAL MODELS, RECEIVER STRUCTURES AND DECISION PROCESSES

This chapter presents three fundamental aspects of radar detection (returned signals, receiver structures and decision processes) in relation to the application to airspace surveillance. This material is not new but is often over-looked, sometimes leading to unfounded conclusions. Therefore, readers not familiar with the concepts and constraints of airspace surveillance radars should read the chapter in detail. Others should skim the chapter quickly so that later chapters will appear in proper context.

Figure 6.1 illustrates the fundamental aspects in the order they are discussed. First a realistic mathematical model for the returned signals is developed. This model leads naturally to a receiver structure that is both practical and essentially optimum. Finally, the decision process is described.

Throughout this chapter, and the remainder of the thesis, complex notation is used to represent bandpass signals and systems. It is suggested that those readers not familiar with this notation read Van Trees (1971, appendix) or a like reference before continuing.

6.1 Signal Models

Three types of signals are assumed to be collected by the receiving antenna: those due to targets, those due to clutter, and those due to random noise. The signals due to targets and clutter

figure 6.1 fundamental aspects of radar detection and the order of discussion.

are a result of essentially the same reflection process, but from different objects, so they will be modelled in the same way but with different parameters. This model will now be discussed.

6.1.1 Target and Clutter Signal Model

In the case of surveillance radars, where it may be assumed that the reflecting surfaces of targets and clutter do not change significantly during the time duration of the radar pulse, targets and clutter can be represented by a collection of point targets or clutter distributed in space (Van Trees 1971, chapters 9 and 12, and Nitzberg 1978). The meaning of the term point target or more generally point reflector will become more concrete as the discussion continues. Essentially, it refers to an individual reflector with no depth or spatial character. The spatial character is introduced by considering a number of point reflectors.

The nature of signals returned from targets and clutter will now be discussed in detail. This development closely parallels that found in Van Trees (1971, chapter 9).

Consider for the present that the transmitter emits a single continuous sinusoid

$$s_t(t) = \sqrt{2P_t} \cos(\omega_c t) = \sqrt{2} \operatorname{Re}(\sqrt{P_t} \exp(j\omega_c t)) \quad (6.1)$$

where $s_t(t)$ is the transmitted signal, P_t the transmitted power,

ω_c the carrier frequency, and $\text{Re}(\cdot)$ designates "the real part of". Then, if there is a stationary point target (or clutter object) at some range R_a , the received signal will have the following structure

$$s_r(t) = \sqrt{2} \text{Re}(\sqrt{P_t} g \exp(j\omega_c(t - \tau) + j\theta)) \quad (6.2)$$

This is a delayed, scaled, and phase shifted replica of the transmitted signal. The attenuation constant g includes the effects of the transmitting antenna gain, the two-way propagation loss, the radar cross-section (or effective size) of the target, and the receiving antenna gain. The phase angle θ is a random phase incurred in the reflection process. The phase shift may sometimes, for convenience of notation, be absorbed into the attenuation constant. When this is done the modified attenuation constant will be denoted as $\bar{g}$. The delay τ is the round trip delay which is related to range as follows

$$\tau = \frac{2 R_a}{c} \quad (6.3)$$

where R_a is the range, and c the speed of propagation of electromagnetic waves.

Two basic assumptions are now made concerning the reflection process. First, it is assumed to be frequency independent. That is if we transmit

$$s_t(t) = \sqrt{2} \text{Re}(\sqrt{P_t} \exp(j\omega_c t + j\omega t)) \quad (6.4)$$

we receive

$$s_r(t) = \sqrt{2} \text{Re}(\sqrt{P_t} \bar{g} \exp(j(\omega_c + \omega)(t - \tau))) \quad (6.5)$$

Second, the reflection process is assumed to be linear so that if we transmit

$$s_t(t) = \sqrt{2} \operatorname{Re}(\sqrt{E_t} \bar{f}(t) \exp(j\omega_c t)) \quad (6.6)$$

where $\bar{f}(t)$ is a normalized¹ finite energy complex envelope pulse, and E_t the energy in the pulse (or equivalently

$$s_t(t) = \sqrt{2} \operatorname{Re}(\sqrt{E_t} \exp(j\omega_c t) \int_{-\infty}^{\infty} \bar{F}(\omega) \exp(j\omega t) \frac{d\omega}{2\pi}) \quad (6.7)$$

where $\bar{F}(\omega)$ is the Fourier transform of $\bar{f}(t)$) we receive

$$s_r(t) = \sqrt{2} \operatorname{Re}(\sqrt{E_t} \bar{g} \exp(j\omega_c(t - \tau)) \int_{-\infty}^{\infty} \bar{F}(\omega) \exp(j\omega(t - \tau)) \frac{d\omega}{2\pi}) \quad (6.8)$$

or equivalently :

$$s_r(t) = \sqrt{2} \operatorname{Re}(\sqrt{E_t} \bar{g} \exp(j\omega_c(t - \tau)) \bar{f}(t - \tau)) \quad (6.9)$$

With these two assumptions in hand the problems of a pulsed signal and a moving target may be considered.

Let a point target have some constant radial velocity v_r so that the range is now a function of the time t .

$$R(t) = R_a - v_r t \quad (6.10)$$

The signal returned from this target is

$$s_r(t) = \sqrt{2} \operatorname{Re}(\sqrt{E_t} \bar{g} \bar{f}(t - \tau(t)) \exp(j\omega_c(t - \tau(t)))) \quad (6.11)$$

¹ normalized in the sense that $\int_{-\infty}^{\infty} |\bar{f}(t)|^2 dt = 1$

where $\tau(t)$ is the round trip delay time. A signal received at time t was reflected from the target at time $t - \frac{\tau(t)}{2}$ and at that time the range was

$$R(t - \frac{\tau(t)}{2}) = R_a - v_r (t - \frac{\tau(t)}{2}) \quad (6.12)$$

However, by definition we have

$$\tau(t) = \frac{2}{c} R(t - \frac{\tau(t)}{2}) \quad (6.13)$$

so substituting equation 6.12 into equation 6.13 gives

$$\tau(t) = \frac{\frac{2}{c} R_a}{1 + \frac{v_r}{c}} - \frac{\frac{2}{c} v_r t}{1 + \frac{v_r}{c}} \quad (6.14)$$

For target velocities of interest $v_r/c \ll 1$ and

$$\tau(t) \approx \frac{2R_a}{c} - \frac{2}{c} v_r t = \tau - \frac{2}{c} v_r t \quad (6.15)$$

Finally, by substituting equation 6.15 into equation 6.11 we obtain the received signal

$$s_r(t) = \sqrt{2} \operatorname{Re}(\sqrt{E_t} \bar{g} \bar{f}(t - \tau + \frac{2}{c} v_r t) \exp(j\omega_c(t - \tau + \frac{2}{c} v_r t))) \quad (6.16)$$

The phase shift from the term $\exp(-j\omega_c \tau)$, which depends on target range, may be absorbed into the attenuation constant.

It is seen that the target velocity has two effects. The first is a compression or stretching, depending on the sign of v_r , of the time scale of the complex envelope. In most cases we can ignore this. To demonstrate, consider the error in plotting $f(t)$ instead of $f(t - \frac{2}{c} v_r t)$. The maximum difference in the arguments occurs at the end of the pulse (say at time T) and

equals $\frac{2}{c} v_r T$. The resulting error in amplitude is a function of the signal bandwidth. If the bandwidth is BW the signal does not change appreciably in a time equal to BW^{-1} . Therefore if

$$\frac{2}{c} v_r T \ll \frac{1}{BW}$$

or

$$BW T \ll \frac{c}{2v_r} \quad (6.17)$$

we may ignore the time-scale change. This inequality is strictly satisfied in surveillance radar situations.

The second effect is a shift of the carrier frequency. This is called the Doppler shift and is denoted as ω_D , where

$$\omega_D = \omega_c \frac{2}{c} v_r$$

or

$$f_D = f_c \frac{2}{c} v_r = \frac{2}{\lambda} v_r \quad (6.18)$$

where λ is the wavelength. The received waveform neglecting compression and absorbing the phase shift into $\bar{g}$ may now be written as

$$s_r(t) = \sqrt{2} \operatorname{Re}(\sqrt{E_t} \bar{g} \bar{f}(t - \tau) \exp(j\omega_c t + j\omega_D t)) \quad (6.19)$$

For the more general case of a conglomeration of point targets we have

$$s_r(t) = \sqrt{2} \operatorname{Re}(\sqrt{E_t} \sum_{i=1}^k \bar{g}_i \bar{f}(t - \tau_i) \exp(j\omega_c t + j\omega_{Di} t)) \quad (6.20)$$

where k is the number of point targets, $\bar{g}_i$ the attenuation constant for the i^{th} reflector, τ_i the delay for the i^{th} reflector and ω_{Di} the Doppler shift for the i^{th} reflector.

This mathematical representation of the signal received from either a target or clutter will be used in section 6.2 when discussing the form of the receiver to be used in the evaluation of the various clutter suppression schemes. Attention is now turned to noise, another factor besides clutter, that interferes with target detection.

6.1.2 Noise Signal Model

The noise background in which a radar operates is modelled quite well as being white, Gaussian and additive. It may be expressed in complex bandpass notation as follows

$$n(t) = \sqrt{2} \operatorname{Re}(\bar{n}(t) \exp(j\omega_c t)) \quad (6.21)$$

where $n(t)$ is the bandpass noise and $\bar{n}(t)$ the complex envelope of the bandpass noise.

$$\bar{n}(t) = n_R(t) - jn_I(t) \quad (6.22)$$

where $n_R(t)$ is the real part of the complex envelope and $n_I(t)$ is the negative of the imaginary part. These components are white, Gaussian and independent (Van Trees 1971, appendix).

Noise, unlike clutter, does not depend on the transmitted energy of the radar but on such things as system noise figures and the level of cosmic RF interference. Early radars were

self-noise limited in their detection performance in that system components were the main source of noise. This is not the case for modern radars due to advances that have taken place in microwave electronics. Today's radars are either clutter-limited or interference-limited over much of their operational range with self-noise being of concern only at extreme distances.

The initial signal processing occurs in the radar receiver. The structure of this receiver is the topic of the next section.

6.2 Receiver Structure

In this section it is shown how the form of the received signals results in a practical structure for a receiver that is essentially optimum for airspace surveillance radar applications. First I derive the optimum receiver for a single Doppler shifted return in Gaussian noise. Then the constraint imposed by the requirements that the radar have both good range and velocity resolution is applied to simplify this receiver. The result is a practical receiver with nearly optimum performance. Finally, the mathematical forms of the signals at the output of the receiver due to complicated targets, clutter and noise are derived. It is these signals that are passed to the decision process which attempts to suppress the interferers and make an optimal decision regarding the presence or absence of a target.

6.2.1 Detection of a Single Point Target in Gaussian Noise

To derive the structure of the optimum receiver it will be assumed that a specific delay τ (range) and Doppler frequency ω_D (radial velocity) have been chosen for investigation and it is desired to know if a target is present at these coordinates in the delay-Doppler plane. A number of terms from the theory of statistical hypothesis testing will be introduced in this section to aid in the development. These concepts will be discussed in greater detail in section 6.3.

The problem is to decide, given a τ and ω_D , whether a target is present or not. This is called a binary hypothesis testing problem. There are two hypotheses: the target absent hypothesis or null hypothesis, and the target present hypothesis or alternative hypothesis. The null hypothesis will be denoted by the letter H and the alternative by the letter A.

In terms of the signal models developed in section 6.1 we have the following problem.

$$\text{under H: } r(t) = n(t)$$

$$\text{under A: } r(t) = s_r(t) + n(t)$$

(6.23)

where $r(t)$ is the received signal, or more fully

$$\text{under H: } r(t) = \sqrt{2} \operatorname{Re}(\bar{n}(t) \exp(j\omega_c t))$$

$$\begin{aligned} \text{under A: } r(t) = & \sqrt{2} \operatorname{Re}(\sqrt{E_t} \bar{g} \bar{f}(t - \tau) \exp(j\omega_c t + j\omega_D t)) \\ & + \sqrt{2} \operatorname{Re}(\bar{n}(t) \exp(j\omega_c t)) \end{aligned}$$

(6.24)

Appealing to the theory of bandpass signal representation

(Van Trees 1971, appendix) this binary hypothesis testing problem may be written in the equivalent complex envelope form as follows

$$\text{under H: } \bar{r}(t) = \bar{n}(t)$$

$$\text{under A: } \bar{r}(t) = \sqrt{E_t} \bar{g} \bar{f}(t - \tau) \exp(j\omega_D t) + \bar{n}(t) \quad (6.25)$$

where $\bar{r}(t)$ is the complex envelope of $r(t)$. In receiver system terms this is equivalent to considering the real and imaginary parts of $\bar{r}(t)$ to be the in-phase and quadrature channel outputs respectively of a synchronous demodulator.

A method must now be found by which a choice may be made between H and A. I shall follow the approach of Van Trees (1971, page 244). Since the noise is white and Gaussian we can expand the received signal using any complete orthonormal set of basis functions and obtain statistically independent coefficients. The signal can be chosen as the first orthonormal function and the resulting coefficient

$$\bar{r} \triangleq \int_{\tau - \frac{T}{2}}^{\tau + \frac{T}{2}} \bar{r}(t) \bar{f}^*(t - \tau) \exp(-j\omega_D t) dt \quad (6.26)$$

will be the minimum sufficient statistic for the hypothesis test.

In equation 6.26 * denotes complex conjugation.

Calling $\bar{r}$ the minimum sufficient statistic simply means that none of the other coefficients resulting from other orthonormal basis functions contain any information about the target return that would aid in the decision process as the functions

are all orthogonal to the returning target signal. Also, the noise terms resulting from the other basis functions bear no information about the noise term in $\bar{r}$ because they are all independent. Therefore, the binary hypothesis testing problem reduces to the following

$$\begin{aligned} \text{under H: } \bar{r} &= \bar{n} \\ \text{under A: } \bar{r} &= \sqrt{E_t} \bar{g} + \bar{n} \end{aligned} \quad (6.27)$$

where

$$\bar{n} = \int_{\tau - \frac{T}{2}}^{\tau + \frac{T}{2}} \bar{n}(t) \bar{f}^*(t - \tau) \exp(-j\omega_D t) dt \quad (6.28)$$

and is a zero-mean complex Gaussian random variable.

The complex quantity $\bar{r}$ was obtained by assuming a specific τ and ω_D . Equation 6.27 will now be rewritten with this functional dependence shown explicitly, giving

$$\begin{aligned} \text{under H: } \bar{r}(\tau, \omega_D) &= \bar{n}(\tau, \omega_D) \\ \text{under A: } \bar{r}(\tau, \omega_D) &= \sqrt{E_t} \bar{g}(\tau, \omega_D) + \bar{n}(\tau, \omega_D) \end{aligned} \quad (6.29)$$

The statistic $\bar{r}(\tau, \omega_D)$ may be generated by either a correlation receiver as shown in figure 6.2a or a matched filter receiver as shown in figure 6.2b. Both receivers are entirely equivalent in terms of performance, however the philosophy of implementation is slightly different. In the correlation receiver, τ is fixed and the Doppler frequency coordinate is searched for a target. In the matched filter receiver ω_D is

(a)

(b)

figure 6.2 receiver structures: (a) correlation receiver, (b) matched filter receiver. $h(u)$ is the impulse response of the matched filter.

fixed and the τ coordinate is searched for a target. Whether one receiver is used over another is simply a matter of practicality for the specific radar application.

6.2.2 Simplified Receiver Structure

In the previous section the optimum receiver for detecting a target in white Gaussian noise was derived. In this section this receiver is considered again but in the context of the radar application of interest, namely surveillance radar. It will be seen that the constraints imposed by the application permit some simplifications of the receiver structure.

When considering a single pulse from a radar there is a trade-off between range resolution and velocity resolution. The longer the pulse the better the velocity resolution but the range resolution suffers and vice versa. This is a well known effect and is best described with the use of the radar ambiguity function, however a detailed study of this is not of prime importance to our discussion. The interested reader is referred to Skolnick (1962, chapter 10) or Van Trees (1971, chapter 10) for more details. In surveillance radar applications range resolution is important so the pulses must be kept short and Doppler resolution is obtained by processing a number of pulse returns. The constraint of a short pulse permits a simplification of the receiver structure developed in section 6.2.1 and given by equation 6.26 as

$$\bar{r} = \int_{\tau - \frac{T}{2}}^{\tau + \frac{T}{2}} \bar{r}(t) \bar{f}^*(t - \tau) \exp(-j\omega_D t) dt \quad (6.30)$$

Since it is assumed that the pulse is so short that there is very little Doppler information in a single return, the expression $\exp(-j\omega_D t)$ does not change much during the pulse duration T , and hence the statistic $\bar{r}$ may be approximated as

$$\hat{\bar{r}} = \exp(-j\omega_D \tau) \int_{\tau - \frac{T}{2}}^{\tau + \frac{T}{2}} \bar{r}(t) \bar{f}^*(t - \tau) dt \quad (6.31)$$

The integral in equation 6.31 is the optimum receiver for a signal with zero Doppler shift, so several returns must be processed together to take advantage of the phase factor preceding the integral so as to obtain a Doppler measurement. The discussion of this will be left to the section on the decision process. Of current interest is the degree to which a filter matched to zero Doppler shift may be considered "almost" matched for other Doppler frequencies in terms of signal energy.

Consider a single return in the absence of noise

$$\bar{r}(t) = \sqrt{E_t} \bar{g} \bar{f}(t - \tau) \exp(j\omega_D t) \quad (6.32)$$

Then the output of the approximately matched filter is

$$\hat{\bar{r}} = \exp(-j\omega_D \tau) \int_{\tau - \frac{T}{2}}^{\tau + \frac{T}{2}} \sqrt{E_t} \bar{g} \bar{f}(t - \tau) \exp(j\omega_D t) \bar{f}^*(t - \tau) dt \quad (6.33)$$

Since τ is arbitrary we may set it to zero without loss of generality, giving

$$\hat{r} = \sqrt{E_t} \bar{g} \int_{-\frac{T}{2}}^{\frac{T}{2}} |\bar{f}(t)|^2 \exp(j\omega_D t) dt \quad (6.34)$$

However, if the filter had been perfectly matched to the Doppler frequency the output would have been

$$\bar{r} = \sqrt{E_t} \bar{g} \int_{-\frac{T}{2}}^{\frac{T}{2}} |\bar{f}(t)|^2 dt = \sqrt{E_t} \bar{g}, \quad (6.35)$$

remembering that

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} |\bar{f}(t)|^2 dt = 1$$

Consequently, the error between the approximate statistic

$\hat{r}$ and $\bar{r}$ is

$$e = \bar{r} - \hat{r} = \sqrt{E_t} \bar{g} \left(1 - \int_{-\frac{T}{2}}^{\frac{T}{2}} |\bar{f}(t)|^2 \exp(j\omega_D t) dt \right) \quad (6.36)$$

For definiteness it will be assumed that the pulse is a gate function between $-T/2$ and $T/2$, i.e., that

$$\bar{f}(t) = \frac{1}{\sqrt{T}} G_T(t) = \begin{cases} \frac{1}{\sqrt{T}} & -\frac{T}{2} < t < \frac{T}{2} \\ 0 & \text{otherwise} \end{cases} \quad (6.37)$$

and so

$$e = \sqrt{E_t} \bar{g} \left(1 - \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \exp(j\omega_D t) dt \right) \quad (6.38)$$

or

$$e = \sqrt{E_t} \cdot g \left(1 - \text{Sa} \left(\frac{T}{2} \omega_D \right) \right) \quad (6.39)$$

where $\text{Sa}(x)$ is the sampling function defined as

$$\text{Sa}(x) = \frac{\sin(x)}{x} \quad (6.40)$$

From this, it is clear that if the product $\omega_D T$ is small so that $\text{Sa} \left(\frac{T}{2} \omega_D \right) \approx 1$ then the error is negligible. An example using typical surveillance radar parameters illustrates that $\omega_D T$ is indeed small.

For a L band radar operating around 1200 MHz the pulse duration is typically 1 μ s. Rearranging the product so that one of the parameters is radial velocity we have

$$\frac{T}{2} \omega_D = \frac{2\pi T}{c} f_c v_r \quad (6.41)$$

where f_c is the carrier frequency, c the propagation speed of electromagnetic waves, T the pulse length, and v_r the radial velocity. Equation 6.41 reduces to

$$\frac{T}{2} \omega_D \approx 0.7(10)^{-5} v_r \quad (6.42)$$

where v_r is expressed in km/h. Then if an upper limit of velocity for any target of interest is set at 2000 km/h,

$$\frac{T}{2} \omega_D \approx 0.014 \quad (6.43)$$

and the error is approximately

$$e \approx \sqrt{E_t} \cdot g (0.0003) \quad (6.44)$$

or about 0.03 %. So for surveillance radars that maintain good

range resolution the loss in performance for using the "almost" matched filter is indeed small.

In radar practice the "almost" matched filter is usually implemented in the IF (intermediate frequency) stage as a bandpass filter matched to the length of the radar pulse (i.e., the bandwidth of the filter is usually $1/T$). If the bandpass filter is not matched to the pulse shape but only to its length an additional loss is incurred. This loss is well documented in tables (Meyer and Mayer 1973, page 10) and must be included in performance calculations as a loss of signal-to-noise ratio if such a filter is used.

For the analysis presented in this thesis it will be assumed that the radar pulses are short compared to the period of the highest Doppler frequency of interest (i.e., the "almost" matched filter is virtually ideal) and that the filter is matched to the pulse shape. The block diagram of such a receiver is shown in figure 6.3. All further signal processing for the purpose of suppressing clutter to improve detection will be done using the in-phase and quadrature outputs of the synchronous demodulator. The resulting mathematical representation of the target signal including the dependence on τ and ω_D is

$$\bar{s}(\tau, \omega_D) = \sqrt{E_t} \bar{g}(\tau, \omega_D) \exp(j\omega_D \tau), \quad (6.45)$$

and for noise is

$$\bar{n}(\tau, \omega_D) = n_R(\tau, \omega_D) - j n_I(\tau, \omega_D) \quad (6.46)$$

where $n_R(\tau, \omega_D)$ and $n_I(\tau, \omega_D)$ are Gaussian and independent.

figure 6.3 receiver structure to be used in subsequent analysis. The filter is matched to a single return with zero Doppler shift. Doppler processing follows the synchronous demodulator and requires a number of returns.

The discussion now turns from the case of a single target in noise to the more practical problem of multiple targets in both clutter and noise.

6.2.3 Response to Complicated Targets and Clutter

I shall now consider the response of the "almost" matched filter receiver to a target made up of a number of reflectors. From equation 6.20, representing such a target, the complex envelope is

$$\bar{s}_r(t) = \sqrt{E_t} \sum_{i=1}^k \bar{g}_i \bar{f}(t - \tau_i) \exp(j\omega_{Di} t) \quad (6.47)$$

The filter output is

$$\bar{s}_o(\tau) = \int_{\tau - \frac{T}{2}}^{\tau + \frac{T}{2}} \sqrt{E_t} \sum_{i=1}^k \bar{g}_i \bar{f}(t - \tau_i) \exp(j\omega_{Di} t) \bar{f}^*(t - \tau) dt \quad (6.48)$$

or assuming that the pulse is short compared with $1/f_D$,

$$\bar{s}_o(\tau) \approx \sqrt{E_t} \sum_{i=1}^k \bar{g}_i \exp(j\omega_{Di} \tau) \int_{\tau - \frac{T}{2}}^{\tau + \frac{T}{2}} \bar{f}(t - \tau_i) \bar{f}^*(t - \tau) dt \quad (6.49)$$

The integral expression in equation 6.49 is the complex autocorrelation of the complex envelope $\bar{f}(t)$ and is a function of the difference between τ and τ_i . It will be denoted as $\phi(\tau - \tau_i)$ ¹

¹ It should be noted that $\phi(\tau - \tau_i)$ is zero if $\tau - \tau_i > T$. Therefore signals separated by a delay greater than T , the pulse length, do not overlap and thus are easily separated. This allows us to define a range bin or target separation distance that corresponds to the distance the pulse travels in time T . This is related to the bandwidth and will be discussed in section 6.3.

and our filter output becomes

$$\bar{s}_o(\tau) = \sqrt{E_t} \sum_{i=1}^k \bar{g}_i \exp(j\omega_{Di}\tau) \phi(\tau - \tau_i) \quad (6.50)$$

If a specific τ is chosen, say τ' , for which the decision process is to be applied, $\phi(\tau - \tau_i)$ may be absorbed into $\bar{g}_i$ to yield

$$\bar{s}_o(\tau') = \sqrt{E_t} \sum_{i=1}^k \bar{g}_i \exp(j\omega_{Di}\tau') \quad (6.51)$$

where $\bar{g}_i$ now includes the effect of reflectors at other ranges.

A similar expression can be found for returns due to clutter and will be expressed as shown in equation 6.52 for M clutter reflectors.

$$\bar{c}_o(\tau') = \sqrt{E_t} \sum_{p=1}^M \bar{b}_p \exp(j\Omega_{Dp}\tau') \quad (6.52)$$

where $\bar{b}$ is the clutter equivalent of $\bar{g}$ in equation 6.51 and Ω_D the clutter equivalent of ω_D .

Since a number of pulses are required to obtain any significant Doppler resolution the pulse number l will be introduced into the notation. To simplify further expressions it will be assumed that the time delay (range) of interest is fixed and known so it will be dropped from the notation. Also, the phase of the initial return is random so the term $\exp(j\omega_{Di}\tau)$ may be absorbed by the the attenuation coefficient $\bar{g}_i$. Only the phase difference $\omega_{Di}\Delta$ from return to return need be noted.

The expressions for the target and clutter with this new notation become

$$\bar{s}(\ell) = \sqrt{E_t} \sum_{i=1}^k \bar{g}_i(\ell) \exp(j\omega_{Di} \Delta \ell) \quad (6.53)$$

$$\bar{c}(\ell) = \sqrt{E_t} \sum_{p=1}^M \bar{b}_p(\ell) \exp(j\Omega_{Dp} \Delta \ell) \quad (6.54)$$

where $\ell = 0, 1, 2, \dots$. The noise term is denoted by $\bar{n}(\ell)$.

Various combinations of detection problems may now be posed using equations 6.53 and 6.54 plus a noise term. The nature of the problem (i.e., the real physical situation it describes) will depend upon the statistics assigned to $\bar{g}_i$ and $\bar{b}_p$ as well as the values of ω_{Di} and Ω_{Dp} . In the next section I show how these signals are manipulated so that a decision concerning the presence or absence of a target may be made.

6.3 The Decision Process

In section 6.2 a receiver structure was derived that was nearly optimum for the case of a single target return in Gaussian noise. This receiver consisted of an antenna, RF amplifier, IF amplifier, "almost" matched filter and synchronous demodulator. For airspace surveillance radars it was shown that the "almost" matched filter performs essentially the same as a matched filter so in subsequent discussion it will simply be called a matched filter.

The responses of this receiver to complicated targets and clutter are expressed in equations 6.53 and 6.54. As was noted, a single return does not have any Doppler information. This information must be obtained from the relative phases of consecutive returns. The processing required to extract the

Doppler information is often considered part of the receiver structure, however in our application it is best described as part of the decision process. The reason for this will be made clear in later chapters when different schemes for suppressing clutter are discussed.

This section is divided into four subsections. The first deals with the basic form of the Doppler processing, the second with decision making, the third with false alarms, and the fourth presents an important illustrative example.

6.3.1 Doppler Processing

As was mentioned in the last section the relative phases of consecutive returns must be used to obtain Doppler information. From equation 6.53 the signal from the receiver for a moving point target and the l^{th} return is

$$\bar{s}(l) = \sqrt{E_t} \bar{g} \exp(j\omega_D \Delta l) \quad (6.55)$$

where $\bar{g}$ is assumed to be constant (i.e., independent of l). In order for consecutive returns to add coherently, producing maximum output, each return must be phase shifted (or rotated) to coincide with some reference return. Figure 6.4 illustrates a simple transversal filter that accomplishes this. It can be seen by inspection that the decision sample taken when the delay line is full is

$$\bar{y} = N \sqrt{E_t} \bar{g} \quad (6.56)$$

If the Doppler filter were set to another frequency, say

figure 6.4 transversal filter that rotates consecutive target returns so that they add coherently.

ω , the response would be

$$\bar{y} = \sum_{l=0}^{N-1} \sqrt{E_t} \bar{g} \exp(j(\omega - \omega_D) l \Delta) \quad (6.57)$$

which is a geometric series with the result that

$$\bar{y} = \sqrt{E_t} \bar{g} \exp(j(\omega - \omega_D) \Delta \frac{N-1}{2}) \frac{\sin((\omega - \omega_D) \frac{N\Delta}{2})}{\sin((\omega - \omega_D) \frac{\Delta}{2})} \quad (6.58)$$

This reduces to equation 6.56 when $\omega = \omega_D$.

The filter response for $N = 8$ is shown in figure 6.5. It is periodic with period 2π and obtains its maximum response of 8 when

$$(\omega - \omega_D) \Delta = n 2\pi \quad (6.59)$$

where $n = 0, \pm 1, \pm 2, \dots$. Submaximums or side lobes are evident, producing varied responses for different values of $\omega - \omega_D$. Other values of N yield similar periodic responses except that the number of submaximums and the size of the maximum response vary directly with N .

This type of filter is often used for Doppler processing and is frequently found in the form of a discrete Fourier transform (DFT). This is equivalent to using N separate filters one tuned to each of the nulls of the zero Doppler ($\omega_D = 0$) filter. With such a scheme there are portions of the Doppler band for which the response is not maximum resulting in what is called a cusping loss (Nathanson 1969). Such losses are not usually considered in detection performance calculations until the final system calculations are made where they are included as a loss in

figure 6.5 Doppler filter response. The target's actual Doppler frequency is ω and the filter is set for a maximum response at $\omega = \omega_D$.

signal-to-noise ratio.

As was mentioned earlier I shall assume that the target's range and velocity have been selected a priori and that the system is tuned in accordance with this selection. The problem remaining is to determine if a target with those parameters is present. This philosophy reflects the processing flexibility of modern radars.

The Doppler filter with coefficients as shown in figure 6.4 is the optimum processor for constant targets in Gaussian noise as is shown in chapter 7. Recently much research activity has centered around altering the coefficients so as to make the processor optimum for other types of interference such as general clutter and other radars. I shall not deal with this aspect of the Doppler filter problem but instead concentrate on the modifications made to the basic Doppler filter to combat ground clutter. These modifications, discussed in chapters 8 and 9, involve both additional processing and special threshold techniques.

I shall now deal with the criteria for deciding on the presence or absence of a target.

6.3.2 Decision Making

Generally in radar detection problems, as described earlier, the decision process is a matter of choosing between whether there is a target or not. In the theory of hypothesis testing this is called a binary hypothesis test. Essentially a statistic is

given which has one probability distribution if there is no target and another one if there is. Given the value of the test statistic, one wants to determine which probability distribution is the most likely and hence which decision should be made.

The two decisions that might be made are called hypotheses. The target absent hypothesis is generally called the null hypothesis and the target present hypothesis the alternative. Essentially one is testing to see if there is enough evidence to reject the null hypothesis and hence accept the alternative.

The theory of hypothesis testing is due originally to Neyman and Pearson (1933) who proposed tests that minimized the chances of error. The notions of cost and risk were later introduced by Wald (1939) whose theory has evolved into what is known today as decision theory.

Both Neyman-Pearson theory and decision theory have the same mathematical root, namely the likelihood ratio. The likelihood ratio is simply the ratio of the conditional probability distributions of the test statistic under the two hypotheses. If we let X be a random variable which is the test statistic, $p(X = x | H)$ be the distribution of X when the null hypothesis is true, and $p(X = x | A)$ be the distribution of X when the alternative hypothesis A is true, then the likelihood ratio is

$$\Lambda(x) \triangleq \frac{p(X = x | A)}{p(X = x | H)} \quad (6.60)$$

In general the statistic X may be a vector.

It can be seen by inspection that if for a given x $p(X = x | A)$ is much greater than $p(X = x | H)$ then $\Lambda(x)$ will be large and the alternative hypothesis is accepted. Conversely if $p(X = x | H)$ is much greater than $p(X = x | A)$ then $\Lambda(x)$ is small and the null hypothesis is accepted. Therefore for large $\Lambda(x)$ the alternative is accepted (null rejected) and for small $\Lambda(x)$ the null is accepted (alternative rejected). The problem one has in making the decision is where to set a threshold on $\Lambda(x)$. For very large or very small values of $\Lambda(x)$ there is no difficulty, but for less decisive values some criteria must be established for making the decision. This is where the two philosophies of hypothesis testing, Neyman-Pearson and decision theory, part company.

In Neyman-Pearson theory the threshold (or decision regions in the vector case) are determined in order to minimize the errors in making the decision. The probability of rejecting the null hypothesis when it is true (false alarm probability in radar) is fixed at some acceptable size and then the probability of accepting the null hypothesis when it is not true (false dismissal) is minimized. Since the minimization maximizes the probability of detection, under Neyman-Pearson theory the radar problem becomes "set the probability of false alarm and maximize the probability of detection".

In decision theory one needs to establish the costs involved in making decisions be they right or wrong. This cost function along with any a priori information on the presence of

a target are combined to bias the threshold (or contour in the vector case) to minimize the cost. The problem with the decision theory approach in a radar application is that costs are rarely known and at best are educated guesses. Also the alternative being true (target present) is a very rare event and so it is almost impossible to establish a priori statistics.

The Neyman-Pearson approach, on the other hand, offers a very concrete criteria. A radar processing system has a specific capability for handling false alarms whether its detection scheme is an operator watching a PPI or a complicated automatic process. Using the Neyman-Pearson approach the probability of detecting a target is maximized for the resources available to handle false alarms. For this reason Neyman-Pearson theory will be used in the remainder of this thesis for calculating radar systems performance. This is not a great restriction however, since the likelihood ratio is still the basis of the decision process and so the same system parameters may be used when considering other decision criteria. For further study on the subject the reader is referred to Helstrom (1960, chapter 3) ✓

6.3.3 False Alarms

In the Neyman-Pearson theory of hypothesis testing as applied to radar, the probability of false alarm is an important parameter. To get an idea of the constraints that must be imposed on this parameter a typical case of a L band surveillance radar will be discussed. First however, the number of opportunities

3

for false alarms will be established and then this interpreted in terms of the radar's ability to handle false alarms. During the course of this discussion the term "detection bin" will be defined.

In radar systems in which automatic processing is done with digital hardware, the radar space is usually divided into areas called detection bins. The quadrature channels of the synchronous demodulator are sampled at a rate equal to the bandwidth of the system which is approximately equal to the reciprocal of the pulse length. Each sample is the output (or test statistic) for a specific detection bin. Figure 6.6a illustrates the detection bins in the range dimension. T is the length of the pulse and the sample is taken in the middle of the bin. If the pulse were rectangular and the bandpass filter matched to this shape the resulting sample for the n^{th} detection would be as illustrated in figure 6.6b. For a target at a range corresponding to a delay of exactly $(n + \frac{1}{2})T$ the output for bin n would be maximum and the output for the adjacent bins would be zero. If however the target were a little further away the results would be as shown with the dashed lines. The maximum would be elsewhere in the bin and the $n + 1$ bin would also contain some target information. This results in a loss in target detectability called range-cusping loss (see Nathanson 1969, page 558). In detection calculations, this loss is ignored when determining the probability of detection for a given signal level and in comparing detection schemes based on the same sampling procedure. It is included later when overall system performance is desired as an

(a)

(b)

figure 6.6 detection bins in the range dimension:- (a) the extent of the detection bins and the sample locations, (b) an illustration of range cusping loss.

effective average loss in signal-to-noise ratio of about 2 to 3 dB.

The detection bin has two dimensions, one in range and the other in azimuth. The azimuth width of a detection bin is one beamwidth. Since the beam is considered to be stepped instantaneously from one direction to the next there is also a cusping loss associated with the azimuthal sampling. This loss, like the range-cusping loss, is ignored in detection calculations and included later as an average loss in signal level when considering the performance of the entire system.

Cusping losses can be reduced substantially by increasing the sample rate in both the range and azimuth dimensions. This increases the correlation of the noise and target information from sample to sample causing a target or false alarm to be simultaneously detected in more than one bin.

Figure 6.7 illustrates the area concept of the detection bin. Figure 6.7a illustrates the detection bin for range sampling at a period of one pulse length and an azimuth period of one beamwidth. Figure 6.7b illustrates the same concept but with twice the sampling rate to reduce cusping losses. Notice that the detection bins remain the same size and overlap. Therefore the information in each bin is correlated with adjacent bins. Increasing the number of samples increases the opportunity for multiple false alarms so that for a specific system a trade-off occurs between cusping losses and false alarms.

Several assumptions are now made to arrive at a realistic

figure 6.7 area concept of detection bins: (a) range sampling period is T , azimuth sampling period is one beamwidth, (b) sampling periods one-half those of (a).

figure for the number of false alarms. It is assumed that individual tests for the presence or absence of a target are made at ranges corresponding to multiples of the pulse length T . This provides a realistic number of samples in the sense that the white Gaussian noise is independent from sample to sample because the signal is bandlimited to $1/T$ and the target returns will be of duration approximately equal to T . It will also be assumed that the antenna is stepped in intervals equal to the beamwidth and that the returns from neighboring beams are independent.

In digital processing of radar signals, detection bins are usually fixed in position and each bin is tested for the presence or absence of a target. In analog processing (such as on a PPI) the bins are not defined by hard boundaries but nevertheless the extent of each bin is the same with position being dependent on the specific noise and target signals at the time. If N_f is defined as the number of opportunities for a false alarm per pulse transmitted we have

$$N_f = \frac{t_{\max}}{T} = \frac{2 R_{\max}}{c T} \quad (6.61)$$

where $t_{\max}$ is the delay for maximum range, T the pulse length, $R_{\max}$ the minimum range, and c the speed of electromagnetic wave propagation.

To find the total number of opportunities for false alarm N_{ft} , N_f must be multiplied by the pulse repetition rate PRF

$$N_{ft} = (\text{PRF}) \frac{2 R_{\max}}{c T} \quad (6.62)$$

Now if each pulse is not processed independently (because of integration, pulse cancellation, etc.) the PRF must be modified to an effective PRF, say PRF^1 . Then

$$N_{ft} = (PRF^1) \frac{2 R_{a_{max}}}{c T} \quad (6.63)$$

For a typical airport surveillance L band radar $R_{a_{max}} \approx 140$ km, $T \approx 1$ μ s and the PRF ≈ 800 pulses per second so

$$N_{ft} \approx \frac{800(2) 140(10)^3}{3(10)^8 1(10)^{-6}} = 7.47 (10)^5 \approx 10^6 \quad (6.64)$$

Then, if ten pulses go into every decision, the N_{ft} reduces to 10^5 . Generally, there are approximately 10^5 to 10^6 opportunities for false alarms per second with a typical surveillance radar. Assuming each decision to be independent and that the radar in question is capable of processing one false alarm per second, then the probability of getting a false alarm from any one bin must be in the order of 10^{-6} .

Usually performance calculations include the probability of false alarm as a parameter so that a particular scheme's merits may be related to the overall system capability of handling false alarms rather than to some particular and perhaps arbitrary value.

This completes the overview of basic radar signal detection. A simple example will now be presented in detail for the purpose of relating the concepts of hypothesis testing to the complex receiver statistic developed in section 6.2.

6.3.4 A Gaussian Target in Gaussian Noise

To this point the concepts of Neyman-Pearson theory have been discussed along with the approximate size of the important parameter, the probability of false alarm. It is the object of this section to relate these ideas to the receiver statistic developed in section 6.2 through the simple yet important example of detecting a Gaussian target in Gaussian noise. Ample use will be made of two-dimensional diagrams of the probability distributions involved in order to present an intuitive as well as theoretical approach to the detection process.

The binary hypothesis testing problem consists of deciding between two hypotheses H and A as follows

$$\text{under } H: \bar{r} = \bar{n}$$

$$\text{under } A: \bar{r} = \sqrt{E_t} \bar{g} + \bar{n} \quad (6.65)$$

where $\bar{g}$ and $\bar{n}$ are independent complex Gaussian random variables.

The probability distribution of $\bar{r}$ under H is

$$p(\bar{r} | H) = \frac{1}{2\pi\sigma_n^2} \exp\left(-\frac{|\bar{r}|^2}{2\sigma_n^2}\right) \quad (6.66)$$

and of $\bar{r}$ under A

$$p(\bar{r} | A) = \frac{1}{2\pi(\sigma_n^2 + \sigma_s^2)} \exp\left(-\frac{|\bar{r}|^2}{2(\sigma_n^2 + \sigma_s^2)}\right) \quad (6.67)$$

where σ_n^2 is the variance of the real or imaginary part of $\bar{n}$ and

σ_s^2 the variance of the real or imaginary part of the target

signal $\bar{s}$.

These distributions are illustrated in figure 6.8. In this figure σ^2 refers to the variance of a component (i.e., the variance of $\text{Re}(\bar{r})$ or $\text{Im}(\bar{r})$). The distributions are circular symmetric with the distribution under the A hypothesis having the larger spread.

The likelihood ratio for this problem is

$$\Lambda(\bar{r}) = \frac{p(\bar{r} | A)}{p(\bar{r} | H)} = \frac{\frac{1}{2\pi(\sigma_n^2 + \sigma_s^2)} \exp\left(-\frac{|\bar{r}|^2}{2(\sigma_n^2 + \sigma_s^2)}\right)}{\frac{1}{2\pi\sigma_n^2} \exp\left(-\frac{|\bar{r}|^2}{2\sigma_n^2}\right)} \quad (6.68)$$

Remembering that if $\Lambda(\bar{r})$ is greater than some constant, say d_0 , then the null hypothesis is rejected and the following statements can be made.

$$\text{Reject H if } \Lambda(\bar{r}) > d_0 \quad (6.69)$$

or

$$\frac{\sigma_n^2}{\sigma_n^2 + \sigma_s^2} \exp\left(\frac{|\bar{r}|^2}{2} \left(\frac{\sigma_s^2}{\sigma_n^2(\sigma_s^2 + \sigma_n^2)}\right)\right) > d_0 \quad (6.70)$$

or

$$\exp\left(\frac{|\bar{r}|^2}{2} \left(\frac{\sigma_s^2}{\sigma_n^2(\sigma_s^2 + \sigma_n^2)}\right)\right) > d_1 = \left(\frac{\sigma_n^2 + \sigma_s^2}{\sigma_n^2}\right) d_0 \quad (6.71)$$

Since $\ln(x)$ is a monotonic increasing function of x , an equivalent test is

$$\text{reject H if } |\bar{r}|^2 > d_2 = \frac{2\sigma_n^2(\sigma_n^2 + \sigma_s^2)}{\sigma_s^2} \ln(d_1) \quad (6.72)$$

(a)

(b)

figure 6.8 probability distributions for the Gaussian target in Gaussian noise: (a) distribution under the H (target absent) hypothesis, (b) under the A (target present) hypothesis.

or reducing it even further, as $|\bar{r}|$ is always positive

$$\text{reject } H \text{ if } |\bar{r}| > d_3 = \sqrt{d_2} \quad (6.73)$$

Therefore a test based on the likelihood ratio simply calculates $|\bar{r}|$ and compares that to a threshold d_3 . This threshold d_3 may be interpreted as a circular contour in the complex plane as shown in figure 6.9. To conform with future notation the threshold contour radius d_3 will be denoted as T_h .

Under H the variance is small so most of the time $|\bar{r}|$ is smaller than T_h , but under A the distribution spreads out and $|\bar{r}|$ is more likely to be larger than T_h .

Calculating the threshold T_h using the Neyman-Pearson criteria is quite simple for this case. Note that under H the distribution of $|\bar{r}|$ does not depend on σ_s^2 , the target variance, so T_h may be calculated to yield a preselected probability of false alarm P_F independent of the signal strength. More specifically

$$P_F = P(|\bar{r}| > T_h \mid H) \quad (6.74)$$

so from equation 6.66

$$P_F = \int_0^{2\pi} \int_{T_h}^{\infty} \frac{1}{2\pi\sigma_n^2} \exp\left(-\frac{r^2}{2\sigma_n^2}\right) r \, dr \, d\theta \quad (6.75)$$

where

$$r = |\bar{r}| \quad (6.76)$$

and solving we have

$$P_F = \exp\left(-\frac{T_h^2}{2\sigma_n^2}\right) \quad (6.77)$$

(a)

(b)

figure 6.9 threshold contour superimposed on the distributions of $\bar{r}$ under the two hypotheses: (a) H (target absent) hypothesis, (b) A (target present) hypothesis.

resulting in

$$T_h = \sqrt{2} \sigma_n \sqrt{\ln\left(\frac{1}{P_F}\right)} \quad (6.78)$$

Therefore the threshold is determined by two parameters, the probability of false alarm and σ_n . By the Neyman-Pearson Lemma (Lehman 1959) the test described by equation 6.73 is most powerful (i.e., has the highest probability of detection for the given probability of false alarm). Also since T_h does not depend on σ_s , the signal strength, the test is most powerful for all signal strengths and therefore it may be called uniformly most powerful.

The probability of detection for this case may be related directly to the probability of false alarm. Following a similar argument as with the probability of false alarm we find

$$P_D = P(|\bar{r}| > T_h \mid A) \quad (6.79)$$

or

$$P_D = \int_0^{2\pi} \int_{T_h}^{\infty} \frac{1}{2\pi(\sigma_n^2 + \sigma_s^2)} \exp\left(-\frac{r^2}{2(\sigma_n^2 + \sigma_s^2)}\right) r \, dr \, d\theta \quad (6.80)$$

resulting in

$$P_D = \exp\left(-\frac{T_h^2}{2(\sigma_n^2 + \sigma_s^2)}\right) \quad (6.81)$$

Substituting equation 6.78 for T_h into equation 6.81

$$P_D = P_F^{\frac{\sigma_n^2}{\sigma_n^2 + \sigma_s^2}} = P_F^{\frac{1}{1 + \frac{\sigma_s^2}{\sigma_n^2}}} \quad (6.82)$$

so for a given P_F the probability of detection depends on the signal-to-noise ratio, σ_s^2/σ_n^2 .

The above example of a radar detection problem, although simple, is very important. The Gaussian target described by $\sqrt{E_t} \bar{g}$ where $\bar{g}$ is a complex Gaussian random variable with independent components is often called a Rayleigh target, due to the name of the probability distribution of $|\sqrt{E_t} \bar{g}|$. Since the noise term has the same distribution, but with a different parameter, it is also Rayleigh distributed so this detection problem is often referred to as detecting a Rayleigh target in Rayleigh noise.

In general the solution to a detection problem using the likelihood ratio is not as easy as the one presented above. Several complications can arise such as distributions that do not possess a closed form or distributions where the likelihood ratio is not monotonic. These problems will not be encountered in this thesis however.

6.4 Chapter Highlights

This chapter has discussed three aspects of the airspace surveillance radar problem. The first was the nature of the signals returned to the radar. It was found that these signals could be described in convenient mathematical forms with two of the important parameters being a complex, possibly random, amplitude and a Doppler frequency shift.

These signal representations led to the second aspect: a receiver structure that was particularly suited to the airspace surveillance radar application. It consisted of an "almost" matched filter followed by a synchronous demodulator.

The third aspect discussed was the decision process. A Doppler filter was described that used the relative phase of consecutive returns to separate signals in Doppler frequency. It is this filter that forms the basic building block of the ground clutter suppression schemes to be discussed in later chapters.

The Neyman-Pearson decision criteria was chosen as most suitable for radar applications because it maximized the probability of detection for a given false alarm probability P_F . An acceptable P_F is determined by the resources of a particular radar system for handling false alarms. A reasonable choice for P_F is 10^{-6} .

Finally, the problem of detecting a Gaussian target in Gaussian noise was treated in detail to illustrate the concepts discussed earlier in the chapter and for reference in later chapters.

In the next chapter, I build on the foundation given in this chapter by discussing the ground clutter problem in detail and deriving the optimum detector.

CHAPTER 7

THE GROUND CLUTTER PROBLEM: MODELS AND THE OPTIMUM SOLUTION

In this chapter a model for the problem of ground clutter in step-scan airspace surveillance radars is presented. This model merges the concepts discussed in section 1.2 concerning ground clutter with the mathematical structure of detection presented in chapter 6.

Chapter 7 is divided into three sections. The first section develops a model for ground clutter consistent with the observations made in section 1.2 concerning step-scan radars. Also target and noise models are presented that realistically represent the airspace surveillance application. In the second section the detection problem is summarized and the optimum solution given. Finally the significant results of the chapter are reviewed and the results of the next two chapters foreshadowed.

7.1 Ground Clutter, Target and Noise Models

In this section I present models for ground clutter, targets and noise that are consistent with airspace surveillance radars. It is assumed that a step-scan antenna such as a phased array is used so that the signals are not modified by antenna motion.

7.1.1 Ground Clutter

Equation 6.54 (repeated here as equation 7.1) gives the general form of the output of the receiver due to clutter.

$$\bar{c}(\ell) = \sqrt{E_t} \sum_{p=1}^M \bar{b}_p(\ell) \exp(j\Omega_{Dp} \Delta \ell) \quad (7.1)$$

where E_t is the transmitted energy, $\bar{b}_p(\ell)$ the complex amplitude coefficient (i.e., including phase) of the ℓ^{th} return from the p^{th} reflector, and Ω_{Dp} the Doppler frequency of the p^{th} reflector. The exponential term represents the consecutive phasing of the returns due to the Doppler shift.

In section 1.2.2 the intrinsic nature of ground clutter was discussed. It was concluded there that, when a step-scan antenna is used, the amplitude and phase of the ground clutter return are constant. In addition, the velocity of this clutter is zero so $\Omega_{Dp} = 0$ for all p . Hence equation 7.1 becomes

$$\bar{c}(\ell) = \sqrt{E_t} \bar{c} \quad (7.2)$$

$$\text{where } \bar{c} = \sum_{p=1}^M \bar{b}_p$$

True ground clutter in most cases has two components: the constant component represented by equation 7.2 and a varying component. This varying component is due to such things as leaves waving in the breeze. It will be assumed that the varying component may be modelled as Gaussian and independent from return to return. Therefore the complete ground clutter model is

$$\bar{c}(\ell) = \sqrt{E_t} \bar{c} + \bar{D}(\ell) \quad (7.3)$$

where $\bar{D}(\ell)$ is a complex Gaussian random variable with zero mean and independent, equal variance components. $\bar{D}(\ell)$ is also independent over the variable ℓ , i.e., independent from return to return.

Having now defined the variable component of ground clutter, for the purpose of system analysis it will be assumed that $\bar{D}(\ell)$ is zero. However, this does not reduce the generality of the result because a noise term with the same characteristics will be included in the analysis. Thus one may simply assume that the varying part of the ground clutter return is all or part of that noise term after making the necessary adjustments in the clutter-to-noise and signal-to-noise ratios.

7.1.2 Target Model

The ℓ^{th} return from a complex target is given by equation 6.53 and is repeated here as equation 7.4:

$$\bar{s}(\ell) = \sqrt{E_t} \sum_{i=1}^k \bar{g}_i(\ell) \exp(j\omega_{Di}\Delta\ell) \quad (7.4)$$

where E_t is the transmitted energy, $\bar{g}_i(\ell)$ the complex amplitude coefficient (i.e., including phase) and ω_{Di} is the Doppler frequency of the i^{th} reflector. It is assumed that all reflectors comprising the target are travelling at the same velocity, so

$\omega_{Di} = \omega_D$ and equation 7.4 becomes

$$\bar{s}(\ell) = \sqrt{E_t} \exp(j\omega_D \Delta \ell) \sum_{i=1}^k \bar{g}_i(\ell) \quad (7.5)$$

Aircraft are considered complex targets made up of many reflectors (i.e., large k). Therefore by virtue of the two-dimensional central limit theorem (Lord 1954) the amplitude coefficient may be modelled as a complex Gaussian random variable with zero mean and independent equal variance components.

Equation 7.5 becomes

$$\bar{s}(\ell) = \sqrt{E_t} \bar{g}(\ell) \exp(j\omega_D \Delta \ell) \quad (7.6)$$

where $\bar{g}(\ell)$ is a complex Gaussian random variable as described above.

As noted before, this target model is often referred to as a Rayleigh model because $|\bar{g}(\ell)|$ is Rayleigh distributed.

Several possibilities are now available for the return-to-return correlation. A model that has been commonly used to describe aircraft returns is Swerling's case 1 (Mayer and Meyer 1973). This model assumes that the amplitudes of the returns are perfectly correlated within a group of say N returns, but uncorrelated between groups. This model best describes returns received by a noncoherent radar as the phase is ignored. Since we are assuming a coherent radar with a facility for Doppler processing, a restricted case 1 Swerling model seems more appropriate. This model assumes that all returns going into a single decision are perfectly correlated, both in amplitude and phase, and that groups of returns are uncorrelated. The lack of

correlation between groups reflects the fluctuating nature of the target while the perfect correlation between returns within a group reflects the short term coherency observed by Von Schlachta (1974).

7.1.3 Noise

The noise is assumed to be a complex Gaussian random variable that is independent from return to return. The real and imaginary components are also independent and have zero mean and equal variances. This represents the thermal noise of the system and will be denoted as $\bar{n}(\ell)$.

As was noted earlier, the noise may actually represent a mixture of thermal noise and the varying component of ground clutter (or, indeed, any other clutter source).

In the next section I show how the models developed in this section can be put together to describe the detection-in-ground-clutter problem.

7.2 The Detection Problem and the Optimum Solution

In this section target detection in ground clutter is discussed in detail. The models developed in section 7.1 are used to describe the situation in terms of the concepts of section 6.3. In this regard it will be shown that a phase plane description offers an intuitive understanding of the problem. An optimum solution is then presented which is to be the basis for comparing future suboptimum solutions.

The problem of detecting a target in ground clutter and noise utilizing a single return will be discussed first, followed by the multiple returns case.

7.2.1 Single Return

For a single return ($\ell=0$), we have from equations 7.2 and 7.6

$$\bar{c}(0) = \bar{c} \quad (7.7)$$

$$\bar{s}(0) = \bar{g} \quad (7.8)$$

where $\sqrt{E_t}$ has been absorbed into the constants.

Figure 7.1 illustrates the distribution of Gaussian noise $\bar{n}(0)$ in the phase plane. The axis of the plane represents the in-phase and quadrature branches of the synchronous demodulator. The probability distribution of the noise is represented by contours in the phase plane.

It is obvious that a vector notation for the signals is also appropriate as was demonstrated in Part 1 of the thesis. The vector has two components and is represented by the same symbol as the complex quantity but with the bar moved to the bottom. Therefore

$$\underline{n}(0) \equiv \bar{n}(0) \quad (7.9)$$

I shall use the vector notation and complex notation interchangeably throughout the remainder of the thesis.

In the absence of a target (the null hypothesis H) the

figure 7.1 phase plane representation of Gaussian noise.

return is made up of Gaussian noise and clutter

$$\text{under H: } \underline{r} = \underline{c} + \underline{n} \quad (7.10)$$

With the target present (alternative hypothesis A) we have

$$\text{under A: } \underline{r} = \underline{s} + \underline{c} + \underline{n} \quad (7.11)$$

The distributions of $\underline{r}$ under the two hypotheses are shown in figure 7.2. Under the null hypothesis the Gaussian distribution of the noise is simply displaced by the constant clutter component $\underline{c}$. With the addition of the target, which also has a Gaussian distribution, the distribution about the constant clutter spreads as the single component variance increases from σ_n^2 (that due to noise alone) to $\sigma_n^2 + \sigma_s^2$ where σ_s^2 is the single component variance due to the target return alone.

To make an optimum decision between H and A we compute the likelihood function described in section 6.3.2, which is

$$\Lambda(\underline{r}) = \frac{\sigma_n^2}{\sigma_n^2 + \sigma_s^2} \exp\left(\frac{\sigma_s^2}{2\sigma_n^2(\sigma_s^2 + \sigma_n^2)} ((x - c_R)^2 + (y - c_I)^2) \right) \quad (7.12)$$

where x and y are components of the vector $\underline{r}$. The alternative A is chosen when the likelihood function is larger than some yet to be determined constant d_0 . This may be expressed as

$$\begin{aligned} \text{If } \Lambda > d_0 & \text{ choose A} \\ \text{If } \Lambda \leq d_0 & \text{ choose H} \end{aligned} \quad (7.13)$$

Assuming for the time being that σ_s^2 is known, equation 7.13 reduces to

(a)

(b)

figure 7.2 probability distributions of $\underline{x}$ under the two hypotheses: (a) the H (target absent) hypothesis, (b) the A (target present) hypothesis.

$$\begin{aligned} \text{If } (x - c_R)^2 + (y - c_I)^2 &> d_1 \text{ choose A} \\ &\leq d_1 \text{ choose H} \end{aligned} \quad (7.14)$$

where d_1 is a constant (to be determined) related to d_0 .

Since the constant d_1 is to be found from the probability that the statistic is larger than d_1 when H is true (the probability of false alarm P_F) and therefore does not depend on σ_s^2 , the signal level does not effect the test. Hence, the test is uniformly most powerful (UMP) for testing H against A.

Figure 7.3 illustrates two equivalent forms of this test.

In figure 7.3a the threshold contour is drawn around the mean point (c_R, c_I) . A more practical scheme as far as radar detectors are concerned is illustrated in figure 7.3b where the offset is subtracted vectorally (coherently) from the statistic before thresholding. The new statistical test is

$$\begin{aligned} \text{If } x'^2 + y'^2 &> d_1 \text{ choose A} \\ &\leq d_1 \text{ choose H} \end{aligned} \quad (7.15)$$

where $x' = x - c_R$ and $y' = y - c_I$. This is the output of a square law detector. An equivalent test is

$$\begin{aligned} \text{If } (x'^2 + y'^2)^{\frac{1}{2}} &> d_1^{\frac{1}{2}} = d_2 \text{ choose A} \\ &\leq d_1^{\frac{1}{2}} = d_2 \text{ choose H} \end{aligned} \quad (7.16)$$

This is the output of a "linear" detector.

Thus, for a single return, the optimum detector for detecting a Gaussian target in Gaussian noise and constant clutter is one that subtracts the constant clutter coherently and then

figure 7.3 an illustration of the two equivalent forms of the hypothesis test: (a) displaced threshold contour, (b) displacement removed from distribution.

applies a noncoherent (independent of vector pointing direction) test to the result.

Calculating the probability of detection for a given false alarm probability is easy for this system. Consider the statistic $\underline{r}'$

$$\underline{r}' = \underline{r} - \underline{c} \quad (7.17)$$

Then the optimum test from equation 7.16 is

$$\begin{aligned} \text{If } |\underline{r}'| &> d_2 \text{ choose A} \\ &\leq d_2 \text{ choose H} \end{aligned} \quad (7.18)$$

The false alarm probability P_F is found by determining the probability that $|\underline{r}'|$ is greater than d_2 given that H is true.

From this the threshold d_2 is determined as a function of P_F and used to calculate the probability of detection P_D . This problem was solved in section 6.3.4 with the result being

$$P_D = P_F \frac{1}{1 + \frac{\sigma_s^2}{\sigma_n^2}} \quad (7.19)$$

Figure 7.4 shows P_D as a function of the signal-to-noise ratio σ_s^2/σ_n^2 for various values of P_F . It may be seen from the figure that as the probability of false alarm is allowed to increase the signal-to-noise ratio required to obtain a specific probability of detection decreases.

It should be noted that this detection scheme is optimum for the statistics given and therefore it is not possible to do any better. In other words, equation 7.19 describes the

figure 7.4 probability of detection for the optimum detector and a single return as a function of signal-to-noise ratio.

$$\Psi_n = \sigma_n^2 I_{2N} \quad (7.35)$$

where I_{2N} is a $2N$ by $2N$ identity matrix.

The covariance matrix for the target vector Ψ_s can be found by considering the definition in equation 7.28.

$$\Psi_s = E \left\{ \begin{bmatrix} g_R(0) \\ g_R(1)\cos\theta - g_I(1)\sin\theta \\ g_R(N-1)\cos((N-1)\theta) - g_I(N-1)\sin((N-1)\theta) \\ g_I(0) \\ g_R(1)\sin\theta - g_I(1)\cos\theta \\ g_R(N-1)\sin((N-1)\theta) + g_I(N-1)\cos((N-1)\theta) \end{bmatrix} \begin{bmatrix} \text{transpose of} \\ \text{same vector} \end{bmatrix} \right\} \quad (7.36)$$

where

$$\theta = \omega_D \Delta \quad (7.37)$$

Since for a perfectly correlated target

$$g_R(0) = g_R(1) = \dots = g_R(N-1) \quad (7.38)$$

$$g_I(0) = g_I(1) = \dots = g_I(N-1) \quad (7.39)$$

and $g_R(l)$ is independent of $g_I(k)$ for all l and k , Ψ_s reduces to

$$\Psi_s = \sigma_s^2 \begin{bmatrix} 1 & \cos\theta & \cos((N-1)\theta) & 0 & \sin\theta & \sin((N-1)\theta) \\ \cos\theta & & & -\sin\theta & & \\ \cos((N-1)\theta) & 1 & & -\sin((N-1)\theta) & 0 & \\ 0 & -\sin\theta & -\sin((N-1)\theta) & 1 & \cos\theta & \cos((N-1)\theta) \\ \sin\theta & & & \cos\theta & & \\ \sin((N-1)\theta) & 0 & \cos((N-1)\theta) & & 1 & \end{bmatrix} \quad (7.40)$$

At this point it is convenient to note that

figure 7.5 optimum system for detecting a perfectly correlated target in Gaussian noise and ground clutter. $\bar{c}$ is the constant clutter return.

I shall now determine the detection performance for the optimum processor shown in figure 7.5 so that it may be used as a basis for comparing various ground clutter suppression schemes. From equations 7.2 and 7.6 the l^{th} return including a noise term when the target is present is

$$\bar{r}(l) = \bar{g} \exp(j\omega_D \Delta l) + \bar{c} + \bar{n}(l) \quad (7.20)$$

Therefore the output $\bar{y}$ is

$$\bar{y} = \sum_{l=0}^{N-1} (\bar{g} \exp(j\omega_D \Delta l) + \bar{c} + \bar{n}(l) - \bar{c}) \exp(-j\omega_D \Delta l) \quad (7.21)$$

or since the rotation of $\bar{n}(l)$ is meaningless (the phases of each $\bar{n}(l)$ are independent) we have

$$\bar{y} = N \bar{g} + \sum_{l=1}^{N-1} \bar{n}(l) \quad (7.22)$$

Equivalently in vector form

$$\underline{y} = N \underline{g} + \sum_{l=1}^{N-1} \underline{n}(l) \quad (7.23)$$

Since $\bar{g}$ is a Gaussian random vector with zero mean and independent equal variance (σ_s^2) components $N\bar{g}$ is also Gaussian, but with component variances $N^2\sigma_s^2$. The noise vectors similarly have zero means, independent equal variance (σ_n^2) components but they are independent of each other so the sum has a component variance of only $N\sigma_n^2$.

The distributions of $\underline{y}$ for the target absent and target present cases are shown in figure 7.6. Under both hypotheses

(a)

(b)

figure 7.6 probability distributions of y (output of the optimum system) for the two hypotheses: (a) the H (target absent) hypothesis, (b) the A (target present) hypothesis.

the distribution is circular symmetric Gaussian, and therefore the performance result of the example in section 6.3.4 is applicable, namely

$$P_D = P_F \frac{1}{1 + \frac{N\sigma_s^2}{\sigma_n^2}} \quad (7.24)$$

where $N^2\sigma_s^2$ has been substituted for σ_s^2 and $N\sigma_n^2$ for σ_n^2 . From equation 7.24 it is obvious that the shape of the detection curves will be the same as those for the single return except shifted to the left on the signal-to-noise ratio axis by an amount $10 \log(N)$ (dB). In other words, the use of multiple returns has increased the effective signal-to-noise ratio. Figure 7.7 shows the detection curves for $N = 1, 2$ and 10 .

Since this performance is for the optimum processor, given the model, no other processing scheme will result in a lower SNR requirement for a given probability of detection. This provides a convenient base from which we can make comparisons. In chapters 8 and 9 a measure of performance will be used called the signal-to-noise ratio penalty. This is simply the increase in SNR required for a particular detection scheme to yield the same probability of detection as the optimum processor.

As a historical note of interest it should be mentioned that a great deal of radar detection theory has been developed around the concept of noncoherent integration of multiple returns because early radars were not capable of coherent processing. This theory is still widely used today and is valuable when

figure 7.7 probability of detection performance of the optimum detector for a perfectly correlated target. The probability of false alarm is 10^{-6} and N is the number of returns that are coherently integrated.

applied properly. Those readers interested in this classical theory and its relationship to the coherent processing considered here are referred to thesis appendix A.

7.3 Highlights of Chapter 7

The purpose of this chapter was to pave the way for the comparisons of the ground clutter suppression schemes to be presented in chapters 8 and 9. To this end, the chapter began by developing an appropriate ground clutter model based on the discussion in chapter 1 and the mathematical structure of chapter 6. The resulting model, ultimately expressed by equation 7.2, was a constant complex quantity (or equivalently a constant two-dimensional vector).

The choice of target model was dictated by the expected form of the target, namely an aircraft. Since aircraft are complicated reflectors they are best modelled as Rayleigh targets. Perfect correlation, both in phase and amplitude, was assumed from return to return within a group of multiple returns to facilitate Doppler processing. The groups of multiple returns were assumed to be uncorrelated to reflect the random nature of the target over long time periods.

The noise was modelled as white and Gaussian representing the ordinary thermal sources. It was remarked however, that the noise might also represent a diffuse or varying component of the clutter return.

Once the models were established the detection problem was described and the optimum detection process derived. This optimum process consisted of coherently subtracting the clutter from the returns and then phasing (or rotating) them so that they added coherently. The resultant was then subjected to a noncoherent threshold test. It was also noted that irrespective of target correlation an optimum processor first removes the clutter coherently. Therefore a clutter suppression scheme such as the coherent clutter map which attempts to do this should perform well when compared with other schemes.

The optimum processor derived in this chapter will serve as a basis for comparing the various conventional clutter suppression schemes to be presented in chapter 8 and the coherent clutter map of chapter 9. The performance measure will be the signal-to-noise ratio penalty defined as the increase in signal-to-noise ratio required for a particular scheme to have the same probability of detection as the optimum processor.

7.4 Mathematical Details

This section presents the mathematical details leading to some results used in chapter 7. It is recommended that the reader not interested in these detailed developments disregard this section to preserve the continuity of the thesis discussion.

Section 7.4 is divided into two subsections. The first deals with the optimum detector structure for general target and noise correlation and the second deals with the structure for a perfectly correlated target and uncorrelated noise.

7.4.1 Optimum Detector: General Target and Noise Correlation

When dealing with a number of returns it is convenient to use a vector notation. Figure 7.8 illustrates the variables involved in such a notation and how they relate to the system being considered. The real and imaginary parts of the returns can be in any order in the multi-return vector. I shall define this vector $\underline{R}$ as follows

$$\underline{R} = \begin{bmatrix} r_R(0) \\ r_R(1) \\ \vdots \\ r_R(N-1) \\ r_I(0) \\ r_I(1) \\ \vdots \\ r_I(N-1) \end{bmatrix} \quad (7.25)$$

where $r_R(\ell)$ is the real part of the ℓ^{th} return and $r_I(\ell)$ the imaginary part.

For our problem

$$\bar{r}(\ell) = \bar{g}(\ell) \exp(j\omega_D \Delta \ell) + \bar{c} + \bar{n}(\ell) \quad (7.26)$$

where $\bar{c}$ is the constant clutter return, $\bar{n}(\ell)$ the noise, and $\bar{g}(\ell)$ the target return which may or may not be present. Since both the target and noise returns are zero mean Gaussian random variables the vector $\underline{R}$ is a Gaussian random vector with the clutter vector as the mean. In mathematical form

figure 7.8 variables used in the vector notation for multiple returns.

$$\underline{\mu} = E\{\underline{R}\} = \begin{bmatrix} c_R \\ c_R \\ c_I \\ c_I \end{bmatrix}$$

(7.27)

where $E\{\cdot\}$ denotes expectation.

The covariance of $\underline{R}$ is

$$\Psi = \text{Cov}(\underline{R}) = E\{(\underline{R} - \underline{\mu})(\underline{R} - \underline{\mu})^T\} \quad (7.28)$$

Now since $\underline{R}$ is a Gaussian random vector we may write the probability distribution as follows

$$p(\underline{R}) = \frac{1}{(2\pi)^N |\Psi|^{1/2}} \exp\left(-\frac{1}{2}(\underline{R}^T - \underline{\mu}^T)\Psi^{-1}(\underline{R} - \underline{\mu})\right) \quad (7.29)$$

For the two hypotheses H and A the mean is the same but the covariance matrix Ψ is different. Let Ψ_H be the covariance matrix when H is true and Ψ_A be the covariance matrix when A is true. Then the likelihood ratio becomes

$$\Lambda(\underline{R}) = \frac{p(\underline{R}|A)}{p(\underline{R}|H)} = \frac{|\Psi_H|^{1/2}}{|\Psi_A|^{1/2}} \exp\left(\frac{1}{2}(\underline{R}^T - \underline{\mu}^T)(\Psi_H^{-1} - \Psi_A^{-1})(\underline{R} - \underline{\mu})\right) \quad (7.30)$$

which results in the following optimum test:

$$\begin{aligned} \text{If } (\underline{R}^T - \underline{\mu}^T)(\Psi_H^{-1} - \Psi_A^{-1})(\underline{R} - \underline{\mu}) &> d \text{ choose } A \\ &\leq d \text{ choose } H \end{aligned} \quad (7.31)$$

Letting

$$\underline{x} = \underline{R} - \underline{\mu} \quad (7.32)$$

and

$$B = \Psi_H^{-1} - \Psi_A^{-1} \quad (7.33)$$

the test may be written more compactly as

$$\begin{aligned} \text{If } \underline{x}^T B \underline{x} &> d \quad \text{choose A} \\ &\leq d \quad \text{choose H} \end{aligned} \quad (7.34)$$

In terms of signals this test indicates that the first step in optimum processing is to subtract the clutter coherently from the returns. Then the quadratic-form of equation 7.34 is calculated and compared to a threshold. The matrix B depends upon the return to return correlation of the target and noise.

Therefore, independent of the target and noise correlation the optimum processor first removes the constant clutter component coherently.

In the next section the specific case of a perfectly correlated target in uncorrelated noise will be considered.

7.4.2 Optimum Detector for a Perfectly Correlated Target in Uncorrelated Gaussian Noise

In order to determine the optimum detector structure for detecting a perfectly correlated target in uncorrelated Gaussian noise the matrix B in equation 7.34 must be determined. I shall first consider the covariance matrix for the noise and signal vectors alone.

By definition the uncorrelated noise has the following covariance matrix

$$\Psi_n = \sigma_n^2 I_{2N} \quad (7.35)$$

where I_{2N} is a $2N$ by $2N$ identity matrix.

The covariance matrix for the target vector Ψ_s can be found by considering the definition in equation 7.28.

$$\Psi_s = E \left\{ \begin{bmatrix} g_R(0) \\ g_R(1)\cos\theta - g_I(1)\sin\theta \\ g_R(N-1)\cos((N-1)\theta) - g_I(N-1)\sin((N-1)\theta) \\ g_I(0) \\ g_R(1)\sin\theta - g_I(1)\cos\theta \\ g_R(N-1)\sin((N-1)\theta) + g_I(N-1)\cos((N-1)\theta) \end{bmatrix} \begin{bmatrix} \text{transpose of} \\ \text{same vector} \end{bmatrix} \right\} \quad (7.36)$$

where

$$\theta = \omega_D \Delta \quad (7.37)$$

Since for a perfectly correlated target

$$g_R(0) = g_R(1) = \dots = g_R(N-1) \quad (7.38)$$

$$g_I(0) = g_I(1) = \dots = g_I(N-1) \quad (7.39)$$

and $g_R(k)$ is independent of $g_I(k)$ for all k and l , Ψ_s reduces to

$$\Psi_s = \sigma_s^2 \begin{bmatrix} 1 & \cos\theta & \cos((N-1)\theta) & 0 & \sin\theta & \sin((N-1)\theta) \\ \cos\theta & & & -\sin\theta & & \\ \cos((N-1)\theta) & 1 & & -\sin((N-1)\theta) & 0 & \\ 0 & -\sin\theta & -\sin((N-1)\theta) & 1 & \cos\theta & \cos((N-1)\theta) \\ \sin\theta & & \cos\theta & & & \\ \sin((N-1)\theta) & 0 & \cos((N-1)\theta) & & 1 & \end{bmatrix} \quad (7.40)$$

At this point it is convenient to note that

$$\Psi_s = \sigma_s^2 A^T C A \quad (7.41)$$

where

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta & 0 & 0 \\ 0 & 0 & \cos((N-1)\theta) & 0 & 0 & \sin((N-1)\theta) \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -\sin\theta & 0 & 0 & \cos\theta & 0 \\ 0 & 0 & -\sin((N-1)\theta) & 0 & 0 & \cos((N-1)\theta) \end{bmatrix} \quad (7.42)$$

and

$$C = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \quad (7.43)$$

Matrices A and C have physical interpretations which will be explained when the optimum detector is found.

Consistent with the above definitions it is easily verified that

$$A^{-1} = A^T \quad (7.44)$$

This property of the matrix A will be useful in the discussion to follow.

Now that we have both Ψ_n and Ψ_s we can write expressions for Ψ_H and Ψ_A .

$$\Psi_H = \Psi_n = \sigma_n^2 I_{2N} \quad (7.45)$$

$$\Psi_A = \Psi_n + \Psi_s = \sigma_n^2 I_{2N} + \sigma_s^2 A^T C A \quad (7.46)$$

Therefore

$$B = \frac{1}{\sigma_n^2} I_{2N} - (\sigma_n^2 I_{2N} + \sigma_s^2 A^T C A)^{-1} \quad (7.47)$$

The problem now is to find the inverse of the last term of equation 7.47. Let

$$L = \sigma_n^2 I_{2N} + \sigma_s^2 A^T C A \quad (7.48)$$

then

$$L = A^T (\sigma_n^2 I_{2N} + \sigma_s^2 C) A \quad (7.49)$$

because of property 7.44 and the distributive principle. The inverse of L , using property 7.44 again, is

$$L^{-1} = A^T (\sigma_n^2 I_{2N} + \sigma_s^2 C)^{-1} A \quad (7.50)$$

I shall now partition the matrix $(\sigma_n^2 I_{2N} + \sigma_s^2 C)$ to provide a convenient form for taking the inverse. This gives

$$\sigma_n^2 I_{2N} + \sigma_s^2 C = \begin{bmatrix} \sigma_n^2 I_N + \sigma_s^2 \underline{e} \underline{e}^T & 0 \\ 0 & \sigma_n^2 I_N + \sigma_s^2 \underline{e} \underline{e}^T \end{bmatrix} \quad (7.51)$$

where I_N is a N by N identity matrix and $\underline{e}$ is a N component vector of all 1's. To find the inverse we need only determine the inverse of $\sigma_n^2 I_N + \sigma_s^2 \underline{e} \underline{e}^T$. This is easily found using theorem 5.22 of Noble (1969) and is

$$(\sigma_n^2 I_N + \sigma_s^2 \underline{e} \underline{e}^T)^{-1} = \frac{1}{\sigma_n^2} I_N - \frac{1}{\sigma_n^2 (\frac{\sigma_n^2}{\sigma_s^2} + N)} \underline{e} \underline{e}^T \quad (7.52)$$

Therefore

$$B = \frac{1}{\sigma_n^2} I_{2N} - A^T \begin{bmatrix} \frac{1}{\sigma_n^2} I_N - \frac{1}{\sigma_n^2 (\frac{n}{2} + N)} \underline{e} \underline{e}^T & 0 \\ 0 & \frac{1}{\sigma_n^2} I_N - \frac{1}{\sigma_n^2 (\frac{n}{2} + N)} \underline{e} \underline{e}^T \end{bmatrix} A \quad (7.53)$$

which reduces to

$$B = A^T \begin{bmatrix} \underline{e} \underline{e}^T & 0 \\ 0 & \underline{e} \underline{e}^T \end{bmatrix} A \left(\frac{1}{\sigma_n^2 (\frac{n}{2} + N)} \right) \quad (7.54)$$

Hence the optimum test is

$$\text{If } \underline{x}^T A^T \begin{bmatrix} \underline{e} \underline{e}^T & 0 \\ 0 & \underline{e} \underline{e}^T \end{bmatrix} A \underline{x} > d_1 \text{ choose A} \quad (7.55)$$

$$\leq d_1 \text{ choose H}$$

where

$$d_1 = \sigma_n^2 \left(\frac{\sigma_n^2}{\sigma_s^2} + N \right) d \quad (7.56)$$

If we let

$$\underline{y} = A \underline{x} = \begin{bmatrix} y_R \\ y_I \end{bmatrix} \quad (7.57)$$

where y_R is the vector of the real part of the returns and y_I

the vector of the imaginary part, then the test becomes

$$\begin{aligned} \text{If } \underline{y}_R^T \underline{e} \underline{e}^T \underline{y}_R + \underline{y}_I^T \underline{e} \underline{e}^T \underline{y}_I &> d_1 \quad \text{choose A} \\ &\leq d_1 \quad \text{choose H} \end{aligned} \quad (7.58)$$

For individual components this is written as

$$\begin{aligned} \text{If } (y_R(0) + y_R(1) + \dots + y_R(N-1))^2 \\ + (y_I(0) + y_I(1) + \dots + y_I(N-1))^2 &> d_1 \quad \text{choose A} \\ &\leq d_1 \quad \text{choose H} \end{aligned} \quad (7.59)$$

We may take the square root of both sides resulting in a test of the vector's length:

$$\begin{aligned} \text{If } ((y_R(0) + y_R(1) + \dots + y_R(N-1))^2 \\ + (y_I(0) + y_I(1) + \dots + y_I(N-1))^2)^{1/2} &> d_2 \quad \text{choose A} \\ &\leq d_2 \quad \text{choose H} \end{aligned} \quad (7.60)$$

where

$$d_2 = (d_1)^{1/2} \quad (7.61)$$

Since this test does not depend on the target strength it is uniformly most powerful (UMP).

Several things should be noted about this optimum detector.

First the vector $\underline{x}$ must be formed by subtracting the clutter coherently from the return.

Next the vector $\underline{y}$ must be formed by pre-multiplying the vector $\underline{x}$ by matrix A . It is easy to verify that this is equivalent to rotating the individual returns so that the relative phase difference of the target returns is removed.

The third step in the processing is a result of the form of the matrix $\underline{e} \underline{e}^T$ which results in the rotated vector components being added coherently.

Finally the length of the resulting vector is calculated and compared with a threshold.

The process described above is exactly that of figure 7.5.

CHAPTER 8

CONVENTIONAL GROUND CLUTTER SUPPRESSION SCHEMES

Conventional ground clutter suppression schemes can be grouped into two major classes: those that use Doppler separation of target and clutter to enhance target detection, and those that use a variable threshold. I shall discuss these two classes in the context of step-scan airspace surveillance radars utilizing the Doppler filter described in subsection 6.3.1. The Doppler filter permits aircraft with a radial velocity component to be separated, to some extent, from the ground clutter. This separation is not complete unless the filter's response to the ground clutter is zero. If this is not the case the target must be detected in the presence of the remaining clutter which requires a variable threshold approach. Also, aircraft flight paths at some point are tangential to the radar resulting in zero relative radial velocity. This makes Doppler processing impossible and variable thresholding a "must". This is the most difficult task to be performed by an airspace surveillance radar: maintaining a continuous track through ground clutter when Doppler separation of clutter and target is not possible.

In this chapter I shall discuss the conventional solutions to the problem of detecting a target when Doppler separation is and is not possible. The first technique to be discussed is variable thresholding which is applicable in the latter situation,

followed by pulse cancellation which is applicable in the former. Comparisons will show the strengths and weaknesses of each technique.

8.1 Variable Thresholding

When Doppler separation of the clutter and target is not possible variable threshold techniques prove useful in improving target detection performance. The threshold is defined as that predetermined energy level above which a target is declared present. As was determined in subsection 6.3.2 the criteria for establishing the threshold value most suitable for radar applications is the Neyman-Pearson criteria. That is, the threshold is set to yield a specified probability of false alarm. In Neyman-Pearson theory however, the likelihood function was used to determine an optimum test statistic. Variable thresholding techniques ignore this step and simply establish a threshold for a specified probability of false alarm for the statistic of interest, usually the energy in the return, and then the detection performance is determined from this.

Figure 8.1 illustrates the merit of a variable threshold over a fixed threshold. The statistic of interest in this case is the energy in the return and it is shown as a function of range for a fixed azimuth. If the threshold is made a function of range it can be lower in areas of little clutter and higher in areas of large clutter, still maintaining an overall preset probability of false alarm. For the fixed threshold most of the

figure 8.1 illustration of variable and fixed thresholding.

false alarms occur around the isolated clutter so the threshold must necessarily be high elsewhere, whereas with variable thresholding the false alarms are spread uniformly over the entire range. Therefore the target in range area "A" is more detectable using a variable threshold than a fixed one. On the other hand, a target at the same range as the building is more likely to be detected using the fixed threshold simply because the fixed threshold is lower here than the variable. It is lower because it is so much higher elsewhere. Since it is usually desirable to have a uniform detection probability over the entire range the variable threshold technique is preferred.

When the statistics of the chosen test statistic match those of the likelihood ratio the variable threshold technique is optimum.

Later, systems are discussed that do some preprocessing of the returns to remove any constant clutter component before a threshold test is performed. This processing is an attempt to get the test-statistic statistics to be as close as possible to the statistics of the likelihood ratio so that the system performs optimally. For variable thresholding this preprocessing is ignored and the detection performance is determined solely on the basis of having established a threshold that yields a given probability of false alarm. The philosophy behind this type of system is that, no matter what the clutter is like, if a threshold can be established that yields the desired P_F the addition of the target energy should lift the test statistic over that threshold with

some increasing probability.

In practice determining a threshold that yields a certain P_F is not easy. This is a subject of much research with the resulting schemes falling into one of two classes: those that determine the threshold from returns of neighboring ranges and azimuths, and those that use previous returns from the same range and azimuth. The former can be referred to as area thresholding, the latter temporal thresholding. Each method has its merits and difficulties as were outlined in subsection 1.2.3.2.

The concept of a clutter map is closely related to variable thresholding. Plotting the variable threshold as shown in figure 8.1 it is evident that high thresholds mean large clutter returns and low thresholds mean very little clutter. Plotted now as shown in figure 8.2 the map concept is introduced. In practice the actual clutter levels are usually stored in map form and the thresholds derived from these levels. The map is usually divided up into detection bins for digital storage.

The noncoherent clutter map approaches the optimum processor when the clutter returns are not coherent in phase. Rainstorms, chaff and angels are examples of this type of clutter. It is also a good solution to the clutter problem in noncoherent radar systems.

Usually the clutter map system is adaptive in the sense that the threshold is updated in time so that moving clutter objects such as storms are tracked over the map. When this is done the process is called adaptive thresholding because the threshold adapts to the changing clutter environment.

figure 8.2 a noncoherent clutter map showing the amount of energy in the returns in contour form.

For targets that are competing with stationary ground clutter the noncoherent clutter map is not optimum. This kind of clutter is phase coherent so phase information may be used to enhance target detection. This is the fundamental idea behind the coherent clutter map discussed in chapter 9. Since clutter generally consists of both types, that which is and that which is not phase stable, a combination of coherent and noncoherent clutter maps may well be the best practical way of handling general clutter. The purpose of this section is to calculate the performance of a noncoherent clutter map system in the presence of phase stable clutter for future comparison with the coherent clutter map. In this way the merits of coherent processing of ground clutter will be determined.

8.1.1 System Model

Noncoherent clutter maps are usually used for detection of targets at zero Doppler frequency. This analysis will not be restricted to this, however, as it is instructive to see the effects of a strong constant clutter component when the system is tuned to other Doppler frequencies.

Figure 8.3 illustrates the system to be analyzed. N returns are received possibly containing target information. Each return is rotated by the amount necessary to compensate for the Doppler shift and then the rotated returns are summed. The magnitude of the resulting complex output $\bar{y}(\ell)$ (the real part being the in-phase component and the imaginary part the quadrature

figure 8.3 Doppler filter followed by the decision process.

component) is then compared with a threshold which is stored in a clutter map and has been predetermined to give a desired probability of false alarm.

The distributions of the output $\bar{y}$ (or equivalently the vector y) for the target present and target absent cases will now be discussed. From the analysis of the optimum processor (equation 7.23) the output y for both target and noise is a Gaussian random vector with zero mean and independent, equal-variance components. For the target the single component variance was $N^2 \sigma_s^2$ and for the noise it was $N \sigma_n^2$.

The response of the filter to constant clutter can be determined from equation 6.58 by setting ω to zero and substituting $\bar{c}$ for $\bar{g}$. After absorbing the terms $\sqrt{E_t}$ and $\exp(-j\omega_D \frac{(N-1)\Delta}{2})$ into the constant $\bar{c}$ we have the clutter response

$$\underline{c'} = \underline{c} \left(\frac{\sin(\omega_D \frac{N\Delta}{2})}{\sin(\omega_D \frac{\Delta}{2})} \right) \quad (8.1)$$

Therefore the output due to clutter is a constant vector scaled by the expression in brackets.

Figure 8.4 shows the phase plane representations of the distributions of y under the two hypotheses. If ω_D is such that $\underline{c'}$ is zero (i.e., $\sin(\omega_D \frac{N\Delta}{2}) / \sin(\omega_D \frac{\Delta}{2}) = 0$) then the distributions are circular symmetric Gaussian under both hypotheses. Otherwise the distributions are displaced by $\underline{c'}$. If the target

(a)

(b)

figure 8.4 phase plane representation of the probability distributions of y under the two hypotheses: (a) the H (target absent) hypothesis, (b) the A (target present) hypothesis.

to be detected has zero Doppler frequency then the filter is tuned by setting ω_D to zero and c' reaches its maximum of N_c . This is the worst case as far as detection using a variable threshold is concerned.

8.1.2 Performance

Two variable threshold techniques will be considered. The first and probably the most widely used assumes the corrupting signal to be Gaussian and sets the threshold accordingly. The second technique involves experimentally adapting the threshold to yield the desired probability of false alarm. In practice this threshold is difficult to determine.

As was demonstrated in section 6.3.4 a threshold is a circular contour about the origin. Figure 8.5 illustrates such a threshold. If the signal vector length exceeds the threshold then a target is declared present, if it does not no target is declared. The probability of false alarm (P_F) is the probability that the signal vector length exceeds the threshold in the absence of a target. The probability of detection (P_D) is the probability that the signal vector length exceeds the threshold when a target is present.

If the clutter and noise signals are assumed to be Gaussian then the threshold is determined by measuring the power in the corrupting return and setting the threshold radius accordingly. For a single return (multiple returns will be considered later) the measured average power would be $c^2 + 2\sigma_n^2$

figure 8.5 threshold contour.

where $c = |c|$. Therefore from equation 6.78 the threshold T_G is

$$T_G = \sqrt{2\left(\frac{c^2}{2} + \sigma_n^2\right) \ln\left(\frac{1}{P_F}\right)} \quad (8.2)$$

where $c^2/2 + \sigma_n^2$ is the equivalent single component variance.

Since the distribution of the corrupting signal is not Gaussian but displaced Gaussian the probability of false alarm will not be P_F .

In order to make probability of detection calculations for this threshold setting we must calculate the probability that the signal vector will be outside the threshold contour when the target is present. This is equivalent to calculating the radar Q-function (for definition see Meyer and Mayer 1973). An accuracy of 4 or 5 places after the decimal is sufficient as we are only interested in detection probabilities greater than about 0.1.

This is not true, however, when considering the second thresholding technique. For that case the probability of the signal vector being outside the contour in the absence of a target must be found in order to establish the threshold for the desired P_F by an iterative process. In order to do this an accuracy of about 10 places after the decimal is required. Using classical techniques this is difficult to achieve especially for large clutter values (large offset) (Mitchell and Walker 1971). It is possible however, by using the technique developed in chapter 3 to make these calculations easily.

This technique is suited to calculating the probability

that a circular symmetric vector lies outside a certain radius. Although for the present problem the signal vector is not circular symmetric about the origin it is circular symmetric about the clutter point, and a simple observation turns the problem into a circular symmetric one.

The probability of being outside a certain circle centered at the origin does not change with the direction of the displacement because the displaced distribution is circular symmetric. It is easily shown (subsections 2.2.5 and 2.4.1) that the same probability is attained whether a simple displacement is made or a constant amplitude uniform random phase vector (amplitude equal to the displacement) is added to the circular symmetric vector to be displaced. Since the constant amplitude uniform random phase vector has a circular symmetric distribution so does the result.

Using the notation of Part 1 we may write the probability of being inside a circle of radius r centered at the origin for a circular symmetric distribution as

$$P_2(r) = r \int_0^{\infty} \phi(\rho) J_1(\rho r) d\rho \quad (8.3)$$

where $J_1(\cdot)$ is the first order Bessel function and $\phi(\rho)$ is the characteristic function for the distribution defined as

$$\phi(\rho) = \int_0^{\infty} p_2(r) J_0(\rho r) dr \quad (8.4)$$

where $J_0(\cdot)$ is the zeroth order Bessel function and $p_2(r)$ is the

probability that the vector length lies between r and $r + dr$.

For the Gaussian distribution we have

$$\Phi_G(\rho) = \int_0^{\infty} \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) J_0(r\rho) dr = \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) \quad (8.5)$$

For the constant amplitude uniform random phase vector the distribution $p_2(r)$ is just a circular delta function with radius equal to the amplitude so

$$\Phi_c(\rho) = \int_0^{\infty} \delta(r - c) J_0(r\rho) dr = J_0(c\rho) \quad (8.6)$$

Therefore the characteristic function for the equivalent circular symmetric distribution (equivalent in the sense that it yields the same value for $P_2(r)$ as the displaced distribution) is

$$\Phi_e(\rho) = \Phi_G(\rho) \Phi_c(\rho) = \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) J_0(c\rho) \quad (8.7)$$

and from equation 8.3

$$P_2(r) = r \int_0^{\infty} \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) J_0(c\rho) J_1(r\rho) d\rho \quad (8.8)$$

This integral is a general form of Weber's second exponential integral and has no neat solution as discussed by Watson (1962, page 396).

To calculate $P_2(r)$ the Fourier-Bessel series technique presented in chapter 3 is used. The expression for $P_2(r)$ becomes

$$P_2(r) = \frac{2r}{R} \sum_{k=1}^M \frac{\exp\left(-\frac{\sigma^2 \lambda_k^2}{2R^2}\right) J_0\left(\frac{c \lambda_k}{R}\right)}{\lambda_k J_1^2(\lambda_k)} J_1\left(\lambda_k \frac{r}{R}\right) \quad (8.9)$$

where λ_k is the k^{th} zero of $J_0(\cdot)$. R and M must be determined so that the error in $P_2(r)$ is negligible for the calculations being made. The method for determining R and M is presented in chapter 3 and so will not be repeated here. Generally though, it is sufficient to choose

$$R = c + 10\sigma \quad (8.10)$$

and

$$M = 140. \quad (8.11)$$

to ensure that the error in $P_2(r)$ due to these parameters is less than 10^{-15} for ratios of c/σ as high as 40. The terms in the series are well-behaved and easy to calculate.

The probability of false alarm P_F and the probability of detection P_D are both expressed from equation 8.8 or for calculation purposes 8.9 as $1 - P(T_h)$ where T_h is the threshold. For P_F , σ^2 is replaced by σ_n^2 and for P_D , σ^2 is replaced by $\sigma_n^2 + \sigma_s^2$. Therefore

$$P_F = 1 - T_h \int_0^\infty \exp\left(-\frac{\sigma_n^2 \rho^2}{2}\right) J_0(c\rho) J_1(T_h \rho) d\rho \quad (8.12)$$

and

$$P_D = 1 - T_h \int_0^\infty \exp\left(-\frac{(\sigma_n^2 + \sigma_s^2) \rho^2}{2}\right) J_0(c\rho) J_1(T_h \rho) d\rho \quad (8.13)$$

Figures 8.6 and 8.7 show the probability of detecting the Rayleigh target in the presence of ground clutter for the two threshold techniques and a single return. In figure 8.6 the threshold T_G was set as if the corrupting signal was Gaussian. A P_F of 10^{-6} was used. For figure 8.7 the threshold T_A was set by iteratively determining the threshold, using equation 8.9, that gave a probability of false alarm of 10^{-6} . Since the threshold T_G is larger than T_A the probability of false alarm for T_G is smaller and the resulting probability of detection is also smaller as can be seen by comparing the two figures¹. In the figures, the parameters CNR (clutter-to-noise ratio) and SNR (signal-to-noise ratio) are defined as

$$CNR = \frac{c^2}{2\sigma_n^2} \quad (8.14)$$

and

$$SNR = \frac{\sigma_s^2}{\sigma_n^2} \quad (8.15)$$

For multiple returns (N) the same curves may be used but the parameters CNR and SNR must be scaled

$$CNR' = \frac{CNR}{N} \left(\frac{\sin(\omega_D \frac{N\Delta}{2})}{\sin(\omega_D \frac{\Delta}{2})} \right)^2 \quad (8.16)$$

$$SNR' = N SNR \quad (8.17)$$

¹ Note particularly that the $CNR = -\infty$ dB curve is of course the same in both figures.

figure 8.6 probability of detection using threshold T_G .
Equivalent Gaussian P_F is 10^{-6} .

figure 8.7 probability of detection using threshold T_A .
 P_F is 10^{-6} .

The presence of the ground clutter has a significant effect on the probability of detection. For example, if the clutter-to-noise ratio was 15 dB, using threshold T_G (figure 8.6) results in a 15 dB loss in signal-to-noise ratio compared to the zero clutter ($CNR = -\infty$ dB) case for a probability of detection of 0.9. Using the better threshold T_A (figure 8.7) the signal-to-noise ratio penalty is reduced to 7.6 dB. Thus, if a radar system was capable of handling a P_F of 10^{-6} , using T_G would waste the radar's resources. That is, the actual P_F would be much smaller than the 10^{-6} that the system is capable of handling which in turn would result in an unnecessarily low P_D . On the other hand, T_A which performs better is difficult to determine. Therefore neither solution appears to be particularly attractive.

For comparison purposes it is convenient to express the performance of systems in terms of a signal-to-noise ratio penalty. We can find simple expressions for this penalty for the above two systems by determining the penalty for high signal-to-noise ratios and observing that the expressions are good to within 1 dB for signal-to-noise ratios as low as those required for a P_D of 0.8 in the absence of clutter. In subsection 8.4.2 on mathematical details, it is shown that for large signal-to-noise ratios the signal-to-noise ratio penalty (SNR_p) is, in general,

$$SNR_p = \frac{T_1^2}{T_0^2} \quad (8.18)$$

where T_0 is the threshold radius set for Gaussian noise in the absence of clutter. T_1 is any other threshold radius larger than the length of the clutter vector (this condition holds for both T_G and T_A if $P_F < 0.5$). Therefore the signal-to-noise ratio penalty for large signal-to-noise ratios is the ratio of the areas of the two circular threshold contours. For T_G as determined by equation 8.2, the penalty is

$$SNR_{PG} = \frac{2 \left(\frac{c^2}{2} + \sigma_n^2 \right) \ln \left(\frac{1}{P_F} \right)}{2\sigma_n^2 \ln \left(\frac{1}{P_F} \right)} = CNR + 1 \quad (8.19)$$

The signal-to-noise ratio penalty for T_A is not as easily determined because T_A cannot be found directly. Nevertheless, it is easy to obtain good upper and lower bounds for T_A and therefore upper and lower bounds for SNR_{PA} . In subsection 8.4.3 it is shown that

$$\text{largest} \left[\begin{array}{c} ((CNR)^{\frac{1}{2}} + \frac{1}{\sqrt{2}} \text{erfc}^{-1}(P_F))^2 2\sigma_n^2 \\ 2\sigma_n^2 \ln \left(\frac{1}{P_F} \right) \end{array} \right] < T_A^2 < (CNR^{\frac{1}{2}} + (\ln \left(\frac{1}{P_F} \right))^{\frac{1}{2}})^2 2\sigma_n^2 \quad (8.20)$$

Therefore from equation 8.16

$$\text{largest} \left[\begin{array}{c} \frac{(CNR^{\frac{1}{2}} + \frac{1}{\sqrt{2}} \text{erfc}^{-1}(P_F))^2}{2\sigma_n^2 \ln \left(\frac{1}{P_F} \right)} \\ 1 \end{array} \right] < SNR_{PA} < \frac{(CNR^{\frac{1}{2}} + (\ln \left(\frac{1}{P_F} \right))^{\frac{1}{2}})^2}{\ln \left(\frac{1}{P_F} \right)} \quad (8.21)$$

For P_D 's greater than 0.8 these limiting values for the penalty are within about 1 dB of the actual penalty determined from figures 8.6 and 8.7. This is also true for other probabilities of false alarm. Therefore if the target is assumed to have a probability of detection of 0.8 or greater in the absence of clutter, equations 8.19 and 8.21 may be used to estimate the signal-to-noise ratio penalty incurred by the two thresholding techniques in the presence of clutter.

Again for multiple returns the parameter CNR must be scaled according to equation 8.16.

Illustrating these results graphically, figure 8.8 shows the signal-to-noise ratio penalty as a function of the normalized Doppler frequency $\omega_D \Delta$ for the two thresholds T_G and T_A , $N = 8$, $P_F = 10^{-6}$, and two values of the clutter-to-noise ratio. Three things are to be noted. The first is the periodic structure, as a function of Doppler frequency, of the SNR penalty. The penalty is maximum when the filter clutter residue is maximum and decreases to zero when the filter response to clutter is zero. Submaxima occur between the nulls with the whole pattern repeating with a period of 2π . At Doppler frequencies where the clutter residue is zero the variable threshold techniques reduce to the optimum detector of section 7.2.2 since the Gaussian distributions of the target and noise are no longer displaced in signal space. Increasing N increases the maximum penalty and also the number of nulls.

figure 8.8 signal-to-noise ratio penalty for the two thresholds T_G and T_A . $N = 8$ and $P_F = 10^{-6}$.

The second point to be noted is that the SNR penalty increases as the clutter-to-noise ratio increases. This means that in high clutter regions detection performance is not as good as in low clutter regions possibly resulting in the loss of a target track. The penalty at zero Doppler frequency is finite however, permitting the detection of targets with large signal-to-noise ratios. In addition increasing N increases the clutter-to-noise ratio therefore increasing the penalty.

The final point concerns the relative performance of the thresholds T_G and T_A when considered in the context of practical radars. Since I am assuming that the radar can handle a probability of false alarm P_F of 10^{-6} , there is no advantage to having a smaller P_F . In fact, smaller P_F 's result in high SNR penalties as is evidenced by the penalty for T_G .

In a radar, T_G is relatively easy to establish as one need only measure the average returned energy in the detection bin and apply equation 8.2, however the distribution is not Gaussian so the actual P_F is not the assigned one. For this case it is smaller because the resulting threshold is too large causing the detection probability to fall and an increased SNR penalty.

Threshold T_A provides the correct P_F but must be determined by counting the false alarms which is difficult in most applications. Therefore one has a choice between T_G which is easy to obtain but does not perform well and T_A which is difficult to obtain but performs better, although not well.

The variable threshold approach does not provide optimum performance because the circular symmetric Gaussian distributions are displaced by the ground clutter return. Pulse cancellers attempt to remove this return to enhance target detection and will be our next topic.

8.2 Pulse Cancellers

Pulse cancellers of various forms have been used since World War II for combatting ground clutter in radar systems. The usual name for a system utilizing a pulse canceller is Moving Target Indicator or MTI radar. As the name implies measurement of the actual velocity of the target is not desired, just the separation of moving and stationary objects. The principle of operation is simple. Since stationary objects produce the same return each time a pulse is transmitted and moving objects do not, a number of returns are simply added in such a way that constant returns are eliminated. The simplest pulse canceller subtracts the previous return from the present and then passes the result on for further processing. This is called the two pulse canceller since two returns are involved.

In traditional MTI systems the results of the cancellation process are added noncoherently and a threshold test performed on the net result. Figure 8.9 illustrates this simple form of MTI. The detector takes the magnitude of the complex quantity $\bar{y}(l)$ (the real part is the in-phase component and the imaginary the

figure 8.9 a two pulse canceller MTI system.

quadrature component) and the results are added to form a one-dimensional test statistic $x(\lambda)$. The classical radar detection theory developed by Marcum (1960) and Swerling (1960) is used to analyze this type of system. Appendix A of this thesis summarizes this theory and discusses its connection with the coherent processing we are considering.

For modern radar systems using phased array antennas this type of processing is inferior to what might be accomplished with coherent techniques. For these stationary antennas the correlation time of targets and clutter are extended permitting the use of systems in which the roll of the pulse canceller has changed.

In these new systems, the main concern in detecting moving targets is the leakage of zero Doppler returns due to stationary clutter into higher Doppler frequency filters. The ground clutter returns can be orders of magnitude larger than the target return and so this leakage can degrade performance significantly. To prevent such contamination a pulse canceller is placed in front of the Doppler filter to eliminate the ground clutter before it reaches the filter. This is the philosophy behind the use of a three pulse canceller in Lincoln Laboratory's MTD system (O'Donnell et al. 1974) and a two pulse canceller in a later version of the system (O'Donnell and Muehe, 1979).

This section is divided into two subsections. The first presents the mathematical form of the general pulse canceller and its basic characteristics. The final subsection determines the

performance of a radar system utilizing a two pulse canceller in its new role.

8.2.1 Mathematical Form of Pulse Cancellers

The general form of a pulse canceller of length N is that of a nonrecursive filter shown in figure 8.10. For the most part recursive filters have not been used successfully for clutter suppression because of their long response time and the short correlation time of clutter when observed with a scanning antenna.

This is changing, however, with the advent of phased array antennas.

The impulse response of the filter shown in figure 8.10 is

$$h(\ell) = \sum_{k=0}^{N-1} a_k \delta(\ell - k) \quad (8.22)$$

where $\delta(\ell)$ is the discrete impulse function

$$\delta(\ell) = \begin{cases} 1 & \text{if } \ell = 0 \\ 0 & \text{otherwise} \end{cases} \quad (8.23)$$

Its transfer function is

$$H(z) = \sum_{k=0}^{N-1} a_k z^{-k} \quad (8.24)$$

where $H(z)$ is the z -transform of $h(\ell)$, and the resulting power transfer function is

$$G(\omega) = |H(\omega)|^2 = \left(\sum_{k=0}^{N-1} a_k \cos(k\omega) \right)^2 + \left(\sum_{k=0}^{N-1} a_k \sin(k\omega) \right)^2 \quad (8.25)$$

or

figure 8.10 general form of a pulse canceller.

$$G(\omega) = \left(\sum_{k=0}^{N-1} a_k \right)^2 - 4 \sum_{M=1}^{N-1} \sum_{k=0}^{N-M-1} a_k a_{k+M} \left(\sin\left(\frac{M\omega}{2}\right) \right)^2 \quad (8.26)$$

This is the power transfer function of the general filter.

Pulse cancellers are characterized by having the sum of the coefficients equal to zero. This makes the response to a constant input zero. So we have

$$\sum_{k=0}^{N-1} a_k = 0 \quad (8.27)$$

and therefore the power response becomes

$$G(\omega) = -4 \sum_{M=1}^{N-1} \sum_{k=0}^{N-M-1} a_k a_{k+M} \left(\sin\left(\frac{M\omega}{2}\right) \right)^2 \quad (8.28)$$

The number of returns N used in the canceller is usually small because of the short correlation time of clutter and the fast response time required when using a scanning antenna. The simplest and most widely used canceller is the two pulse canceller shown in figure 8.11. The power response from equation 8.28 with N equal to 2 is

$$G(\omega) = -4 a_0 a_1 \left(\sin\left(\frac{\omega}{2}\right) \right)^2 = 4 \left(\sin\left(\frac{\omega}{2}\right) \right)^2 \quad (8.29)$$

since $a_0 = 1$ and $a_1 = -1$.

Figure 8.12 illustrates this response as a function of normalized Doppler frequency $\omega = \omega_D \Delta / 2\pi$. The response is zero at $\omega_D \Delta = 2\pi n$, $n = 0; \pm 1, \pm 2, \dots$. The resulting blind speeds of pulse cancellers are found from

$$\omega_D = \frac{2\pi n}{\Delta} = \frac{2\pi}{\lambda} v_b \quad (8.30)$$

figure 8.11 two diagrammatic representations of the two pulse canceller

figure 8.12 power response of a two pulse canceller.

resulting in

$$v_b = \frac{n\lambda}{\Delta} \quad n = 0, \pm 1, \pm 2, \dots \quad (8.31)$$

where v_b is the blind speed and λ the carrier wavelength. It can be seen from equation 8.28 that the location of the blind speeds at $\omega_D \Delta = 2\pi n$ is independent of N and a_k . The maximum response occurs at $\omega_D \Delta = \pm(2n-1)\pi$.

8.2.2 Performance of Cancellor Scheme

Consistent with the new role of pulse cancellers discussed earlier I shall now determine the performance of a detection system consisting of a Doppler filter preceded by a two pulse canceller. The two pulse canceller was chosen because it is currently being used in such an application (O'Donnell and Muehe 1979) and also because it exhibits all the characteristics of the pulse canceller family.

The system under consideration is illustrated in figure 8.13. By inspection, the response to constant ground clutter is zero. Additionally since both the target and noise are circular symmetric Gaussian so is the output $\bar{y}$ (or equivalently y). All that remains is to determine the single component variances of the Gaussian distributions.

The variance of a single component of $\bar{y}(\ell)$ for the noise distribution may be determined from Dillard (1975). Denoting this variance as σ_{pn}^2 we have

figure 8.13 two pulse canceller followed by Doppler filter and coherent integration.

$$\sigma_{pn}^2 = 2\sigma_n^2 (N - (N-1)\cos(\omega_D \Delta)) \quad (8.32)$$

where subscript p denotes pulse canceller and n, noise.

The single component variance for the target distribution is most easily determined by moving the pulse canceller to the output of the Doppler filter. This will not affect the results because both systems are linear and time invariant. Assuming the delay line is full and using the first return as a reference, the output after the next return is received is

$$\bar{y} = N \bar{g} \exp(j\omega_D \Delta) - N \bar{g} \quad (8.33)$$

or

$$\bar{y} = 2N \bar{g} \exp(j\omega_D \frac{\Delta}{2}) \sin(\omega_D \frac{\Delta}{2}) \quad (8.34)$$

Since the single component variance of $\bar{g}$ is σ_s^2 and the orientation is irrelevant, the output single component variance is

$$\sigma_{ps}^2 = 4N^2 \sigma_s^2 (\sin(\omega_D \frac{\Delta}{2}))^2 \quad (8.35)$$

where subscript p denotes pulse canceller and s denotes target.

Figure 8.14 illustrates the resulting hypothesis testing problem. Since the distributions are Gaussian under both hypotheses, the signal-to-noise ratio for the optimum processor of chapter 7 can be compared directly to that of the pulse canceller operating in clutter. The signal-to-noise ratio for the pulse canceller scheme is

$$\text{SNR}_{\text{pulse}} = \frac{\sigma_{ps}^2}{\sigma_{pn}^2} = \frac{2N^2 (\sin(\omega_D \frac{\Delta}{2}))^2}{(N - (N-1)\cos(\omega_D \Delta))} \left(\frac{\sigma_s^2}{\sigma_n^2} \right) \quad (8.36)$$

figure 8.14 probability distributions of $\bar{y}$ under the two hypotheses: (a) the H (target absent) hypothesis, (b) the A (target present) hypothesis.

The signal-to-noise ratio for the optimum scheme using an equivalent number of pulses is

$$\text{SNR}_{\text{opt}} = (N + 1) \left(\frac{\sigma_s^2}{\sigma_n^2} \right) \quad (8.37)$$

Therefore the SNR penalty for the pulse canceller scheme is

$$\text{SNR}_{\text{pp}} = \frac{\text{SNR}_{\text{opt}}}{\text{SNR}_{\text{pulse}}} = 1 + \frac{N + \cos(\omega_D \Delta)}{N^2 (1 - \cos(\omega_D \Delta))} \quad (8.38)$$

This penalty is plotted in figure 8.15 for various values of N . Three significant points should be noted.

First, the SNR penalty is independent of the clutter level, hence the pulse canceller scheme performs equally well in strong and weak clutter.

The second point concerns the SNR penalty at zero Doppler frequency and the blind speeds. As can be seen from the figure the penalty goes to infinity for these Doppler frequencies. Therefore aircraft flying tangentially to the radar or radially at one of the blind speeds will not be detected. This is a serious drawback of the pulse canceller schemes.

The final point concerns the number of returns N processed. As N increases the SNR penalty decreases indicating that detection of slow aircraft could be enhanced by increasing the number of pulses processed. The improvement is limited however by the dwell time and the time the moving target is in the detection bin.

figure 8.15 signal-to-noise ratio penalty for the two pulse canceller scheme.

8.3 Chapter Highlights

In this chapter the conventional schemes for suppressing ground clutter were discussed in the context of step-scan airspace surveillance radars. It was found that if the target and ground clutter returns could be separated by Doppler processing then a pulse cancellation scheme worked fairly well. However, when this separation is not possible pulse cancellation is useless and one must resort to variable thresholding techniques. Variable thresholding however, is far from ideal because the SNR penalties are relatively high in moderate clutter and increase as the clutter increases. In addition the most effective threshold T_A is difficult to determine leaving the threshold T_G for which the performance is poor, as a more practicable solution.

In chapter 7 it was shown that the first step in the optimum processor was to coherently remove the ground clutter but, although the pulse canceller is a coherent clutter removal scheme, it is not optimum because it distorts the target returns as was described.

It is possible however, to design a coherent system that takes advantage of long term stability in the clutter to effectively separate targets and clutter. The coherent clutter map, discussed in the next chapter, is such a method.

8.4 Mathematical Details

This section presents the mathematical details leading to some of the results used in chapter 8. It is recommended that

readers not interested in these detailed developments disregard this section to preserve the continuity of the thesis discussion.

Section 8.4 is divided into two subsections. The first deals with the limiting value of the SNR penalty and the second with the upper and lower bounds for T_A .

8.4.1 Limiting Value of the SNR Penalty

The probabilities of detection in the absence of and in the presence of clutter (equations 6.81 and 8.13 combined with result of subsection 4.33) respectively are

$$P_D = \exp\left(-\frac{T_0^2}{2(\sigma_n^2 + \sigma_s^2)}\right) \quad (8.39)$$

and

$$P_D = \int_{T_1}^{\infty} \frac{x}{\sigma_n^2 + \sigma_s^2} \exp\left(-\frac{c^2 + x^2}{2(\sigma_n^2 + \sigma_s^2)}\right) I_0\left(\frac{cx}{\sigma_n^2 + \sigma_s^2}\right) dx \quad (8.40)$$

where $I_0(\cdot)$ is the modified Bessel function of order zero. The two probabilities of detection may be made equal by decreasing σ_s^2 in equation 8.39 by a factor b . Therefore

$$\begin{aligned} & \exp\left(-\frac{T_0^2}{2(\sigma_n^2 + b\sigma_s^2)}\right) \\ &= \int_{T_1}^{\infty} \frac{x}{\sigma_n^2 + \sigma_s^2} \exp\left(-\frac{c^2 + x^2}{2(\sigma_n^2 + \sigma_s^2)}\right) I_0\left(\frac{cx}{\sigma_n^2 + \sigma_s^2}\right) dx \quad (8.41) \end{aligned}$$

Using the series expansion for the right-hand integral that holds for $T_1 > c$ (Helstrom 1960, page 153) we have

$$\begin{aligned} & \exp\left(-\frac{T_0^2}{2(\sigma_n^2 + b\sigma_s^2)}\right) \\ &= \exp\left(-\frac{T_1^2 + c^2}{2(\sigma_n^2 + \sigma_s^2)}\right) \sum_{n=0}^{\infty} \left(\frac{c}{T_1}\right)^n I_n\left(\frac{cT_1}{\sigma_n^2 + \sigma_s^2}\right) \end{aligned} \quad (8.42)$$

solving for b

$$b = \frac{T_0^2}{\sigma_s^2(T_1^2 + c^2)} - \frac{\sigma_n^2}{\sigma_s^2} \quad (8.43)$$

$$\frac{\sigma_s^2(T_1^2 + c^2)}{\sigma_n^2 + \sigma_s^2} - 2\sigma_s^2 \ln\left(\sum_{n=0}^{\infty} \left(\frac{c}{T_1}\right)^n I_n\left(\frac{cT_1}{\sigma_n^2 + \sigma_s^2}\right)\right)$$

Taking the limit as σ_s^2 goes to infinity we have

$$\begin{aligned} b_{\infty} &= \lim_{\sigma_s^2 \rightarrow \infty} (b) \\ &= \frac{T_0^2}{\lim_{\sigma_s^2 \rightarrow \infty} \left(\frac{\sigma_s^2(T_1^2 + c^2)}{\sigma_n^2 + \sigma_s^2} - 2 \lim_{\sigma_s^2 \rightarrow \infty} (\sigma_s^2 \ln\left(\sum_{n=0}^{\infty} \left(\frac{c}{T_1}\right)^n I_n\left(\frac{cT_1}{\sigma_n^2 + \sigma_s^2}\right)\right)) \right)} \\ &= - \lim_{\sigma_s^2 \rightarrow \infty} \left(\frac{\sigma_n^2}{\sigma_s^2} \right) \end{aligned} \quad (8.44)$$

Considering each limit separately

$$\lim_{\sigma_s^2 \rightarrow \infty} \left(\frac{\sigma_n^2}{\sigma_s^2} \right) = 0 \quad (8.45)$$

$$\lim_{\sigma_s^2 \rightarrow \infty} \left(\frac{\sigma_s^2 (T_1^2 + c^2)}{\sigma_n^2 + \sigma_s^2} \right) = T_1^2 + c^2 \quad (8.46)$$

$$\lim_{\sigma_s^2 \rightarrow \infty} \left(\sigma_s^2 \ln \left(\sum_{n=0}^{\infty} \left(\frac{c}{T_1} \right)^n I_n \left(\frac{cT_1}{\sigma_n^2 + \sigma_s^2} \right) \right) \right) = \frac{c^2}{2} \quad (8.47)$$

The first two limits are easily seen. The third one may be determined as follows. Let

$$f(\sigma_s^2) = \sum_{n=0}^{\infty} \left(\frac{c}{T_1} \right)^n I_n \left(\frac{cT_1}{\sigma_n^2 + \sigma_s^2} \right) \quad (8.48)$$

Then the limit becomes

$$\lim_{\sigma_s^2 \rightarrow \infty} \left(\sigma_s^2 \ln (f(\sigma_s^2)) \right) = \lim_{\sigma_s^2 \rightarrow \infty} \left(- \frac{(\sigma_s^2)^2 f'(\sigma_s^2)}{f(\sigma_s^2)} \right) \quad (8.49)$$

where

$$f'(\sigma_s^2) = - \frac{cT_1}{2(\sigma_n^2 + \sigma_s^2)^2} \sum_{n=0}^{\infty} \left(\frac{c}{T_1} \right)^n \left(I_{n-1} \left(\frac{cT_1}{\sigma_n^2 + \sigma_s^2} \right) + I_{n+1} \left(\frac{cT_1}{\sigma_n^2 + \sigma_s^2} \right) \right) \quad (8.50)$$

The limit may be rewritten as

$$\begin{aligned} \lim_{\sigma_s^2 \rightarrow \infty} \left(\frac{(\sigma_s^2)^2 cT_1}{2(\sigma_n^2 + \sigma_s^2)^2} \right) & \lim_{\sigma_s^2 \rightarrow \infty} \left(\left(\sum_{n=0}^{\infty} \left(\frac{c}{T_1} \right)^n I_n \left(\frac{cT_1}{\sigma_n^2 + \sigma_s^2} \right) \right)^{-1} \right. \\ & \times \sum_{n=0}^{\infty} \left(\frac{c}{T_1} \right)^n \left(I_{n-1} \left(\frac{cT_1}{\sigma_n^2 + \sigma_s^2} \right) + I_{n+1} \left(\frac{cT_1}{\sigma_n^2 + \sigma_s^2} \right) \right) \left. \right) \quad (8.51) \end{aligned}$$

The first limit is $cT_1/2$, and the second is c/T_1 as may be found by noting that $I_n(0) = 0$ for $n > 0$, $I_0(0) = 1$ and $I_{-1}(\cdot) = I_1(\cdot)$.

Therefore from equation 8.44

$$b_\infty = \lim_{\sigma_s^2 \rightarrow \infty} (b) = \frac{T_0^2}{T_1^2 + c^2 - 2\left(\frac{cT_1}{2}\left(\frac{c}{T_1}\right)\right)} = \frac{T_0^2}{T_1^2} \quad (8.52)$$

Since the inverse of this limit is the signal-to-noise ratio penalty for high signal-to-noise ratios we have the desired result

$$\text{SNR}_P = \frac{1}{b_\infty} = \frac{T_1^2}{T_0^2} \quad (8.53)$$

8.4.2 Upper and Lower Bounds for T_A

This subsection establishes the upper and lower bounds for T_A shown in equation 8.20.

Upper Bound on T_A

Figure 8.16 shows a displaced Gaussian distribution and two contours C_1 and C_2 . The radius of C_1 is T_0 such that

$$P_F = \exp\left(-\frac{T_0^2}{2\sigma_n^2}\right) \quad (8.54)$$

That is to say, the probability of being outside contour C_1 is P_F . The probability of being outside contour C_2 is necessarily smaller than P_F . So in order that the probability of being outside C_2 be equal to P_F , the radius of C_2 must be made smaller than $c + T_0$ so we have the inequality

figure 8.16 probability distribution and contour structures for determining an upper bound for T_A .

$$T_A < c + T_0 \quad (8.55)$$

But

$$c = \sqrt{2} \sigma_n \text{CNR}^{\frac{1}{2}} \quad (8.56)$$

and

$$T_0 = \sqrt{2} \sigma_n \left(\ln \left(\frac{1}{P_F} \right) \right)^{\frac{1}{2}} \quad (8.57)$$

so

$$T_A^2 < \left(\text{CNR}^{\frac{1}{2}} + \left(\ln \left(\frac{1}{P_F} \right) \right)^{\frac{1}{2}} \right)^2 2 \sigma_n^2 \quad (8.58)$$

Lower Bound on T_A

Figure 8.17 shows a displaced Gaussian distribution and two regions bounded by contour C_2 and line L . The total probability on the shaded side of line L is less than that outside contour C_2 . Denoting these probabilities by P_L and P_{C_2} respectively we have

$$P_L < P_{C_2} \quad (8.59)$$

Now letting P_L be equal to P_F we have

$$P_F < P_{C_2} \quad (8.60)$$

Therefore if P_{C_2} is to be equal to P_F the radius of C_2 must be larger than $c + b$ so

$$c + b < T_A \quad (8.61)$$

figure 8.17 probability distribution and contour structures for determining a lower bound for T_A .

But

$$b = \sigma_n \operatorname{erfc}^{-1}(P_F) \quad (8.62)$$

and

$$c = \sqrt{2} \sigma_n \operatorname{CNR}^{\frac{1}{2}} \quad (8.63)$$

so

$$T_A^2 > \left(\operatorname{CNR}^{\frac{1}{2}} + \frac{1}{\sqrt{2}} \operatorname{erfc}^{-1}(P_F) \right)^2 2\sigma_n^2 \quad (8.64)$$

Also, T_A cannot be smaller than T_0 defined by equation 8.57

so by including this condition we have

$$T_A^2 > \text{largest} \left[\begin{array}{l} \left(\operatorname{CNR}^{\frac{1}{2}} + \frac{1}{\sqrt{2}} \operatorname{erfc}^{-1}(P_F) \right)^2 2\sigma_n^2 \\ 2\sigma_n^2 \ln\left(\frac{1}{P_F}\right) \end{array} \right] \quad (8.65)$$

which completes the development of the upper and lower bounds.

CHAPTER 9

THE COHERENT CLUTTER MAP

As was explained in chapter 1, conventional ground clutter suppression systems were designed under the constraints imposed by a continuous-scan radar. With the advent of step-scan radars however, these constraints have been lifted leaving the door open for improving radar designs. As far as ground clutter is concerned the major difference is that it now appears with its intrinsic nature uncorrupted by the scan. In practical terms this means that the ground clutter returns from truly stationary objects are the same return after return. Therefore they may be stored and subtracted coherently from future returns to remove the ground clutter. In other words, a coherent clutter map may be built up that describes the ground clutter environment, and the information from this map used to enhance target detection.

In this chapter I discuss the concept of a coherent clutter map, its principle of operation and its performance. It is shown that unlike the conventional schemes the coherent clutter map provides virtually optimum performance across the Doppler frequency band.

9.1 The Concept and Principle of Operation

The coherent clutter map differs from the noncoherent clutter map used for variable thresholding. The coherent clutter

map stores both the in-phase and quadrature components of the return rather than just the magnitude thus permitting the clutter to be removed coherently. Such removal is simply not possible with noncoherent data.

A coherent clutter map of the environment is shown in figure 9.1. Here the in-phase and quadrature maps are shown separately in contour form. Suppose that it is desired to determine if a target is present at range R_a and azimuth θ . The radar could send a pulse in direction θ , wait for the return from range R_a and then subtract the stationary components of the I and Q channels from the return. The result would then be processed for target detection.

In practice the storage of a continuous map of the type shown in figure 9.1 is very difficult. As was described in subsection 6.3.3 the radar environment is usually divided up into segments called detection bins with one sample representing all radar information for that bin. Thus only a sampled version of the continuous map need be stored. The sampled version of the map is illustrated in figure 9.2 with only the points where the samples are taken shown. This discrete map will subsequently be referred to as a coherent clutter map.

Subtracting the coherent clutter map values from the radar returns would effectively eliminate the stationary component of the radar environment. Therefore returns from stationary objects such as buildings, chimneys, rocky hills, etc. would be removed and the radar would operate as if they were not there. This gives the radar a greater capability of detecting targets in situations where detection is ground clutter limited.

In-phase component

Quadrature component

figure 9.1 coherent map of stationary radar environment

In-phase component

Quadrature component

figure 9.2 sampled version of the coherent clutter map showing sample point locations.

To build the coherent clutter map the stationary component of the ground clutter must be estimated in the presence of a non-stationary or diffuse component that is due to the motion of some ground clutter objects such as trees when being blown by the wind.

In practice receiver noise and diffuse clutter returns produce an error in the clutter map which in turn degrades system performance. Therefore the estimator must be capable of making a good estimate despite the corrupting returns.

Another requirement which is imposed by technology is that the estimator must be simple in the sense that the hardware for such a system must not become so overwhelming that it is never used.

In Lincoln Laboratory's implementation of the MTD (moving target detector) (O'Donnel et al. 1974) a noncoherent clutter map was used to set detection bin thresholds. This clutter map used a simple first order mean estimator to estimate the clutter power in each detection bin. This estimator is illustrated in figure 9.3. The estimator parameter n was set at 8 to allow the estimator to adjust to varying clutter conditions. The estimate for each bin was stored as a computer word and updated with every return provided no target was present in that bin.

To use this simple yet effective estimator for the coherent clutter map, two storage elements (computer words) are required for each detection bin: one associated with the in-phase component of the return and the other with the quadrature component. An illustrative representation of the estimator and coherent clutter

44

OF / DE

figure 9.3 a first order mean estimator used to estimate clutter power for a noncoherent clutter map.

map is shown in figure 9.4. The return for detection bin $b(R_a, \theta)$, where R_a is the range of the bin and θ is the azimuth of the bin, is scaled by $1/n$ then added to the former estimate of the constant clutter component $c(R_a, \theta)$ scaled by $1 - 1/n$ to form the new estimate which is stored back again in the clutter map. When it is desired to test the bin for the presence of a target the update procedure is stopped and the current estimate of the constant clutter subtracted from the returns processed for the detection test. The noise introduced by using an estimate of the clutter instead of the actual value must be considered in the detection process.

The parameter n for the coherent clutter map could be larger than that of the Lincoln Laboratory's noncoherent clutter map. The noncoherent clutter map was designed to adapt to changes which take place in the order of minutes, such as the passage of a rainstorm. For this reason n had to be small enough so that the time constant of the estimator was such that it could track the changing clutter. For the coherent clutter map an estimate of the constant component of the clutter is desired, a component that if it varies at all changes on a much longer time scale such as hours, days or even months. So the parameter n can be much larger for the coherent clutter map estimator than for the noncoherent.

The next step in the consideration of the coherent clutter map is to calculate the performance and compare it with that of the optimum detector.

figure 9.4 an illustrative representation of the first order estimators and the coherent clutter map.

9.2 Performance

In section 9.1 the concept of the coherent clutter map and the operating procedure were discussed in a qualitative sense. In this section the performance of the coherent clutter map is evaluated in terms of the signal-to-noise ratio penalty. I shall consider the performance for a single detection bin for which the coherent clutter map processor may be modelled as shown in figure 9.5. The principle of operation is as follows.

The radar is directed to the desired location and the processing system is allowed to reach steady state. In this condition the response to the clutter is zero. This is known as the learning phase. During subsequent operation the radar is periodically directed to the location to provide an update for the estimate. When testing for a target the update on the clutter estimate is inhibited. The current estimate is subtracted from the returns and they are processed by the Doppler filter. If no target is detected the returns may be permitted to enter the estimator for updating.

The estimator parameter n may itself be adjusted to reflect the stability of the ground clutter. For example, if the clutter changed significantly over a period of say 10 minutes the parameter must be smaller than if the change took place in a time span of days. The more stable the clutter is the better the cancellation, however, as will be seen from the following analysis, good performance is obtained even with relatively small values of n .

figure 9.5 model for coherent clutter map processor. $\bar{E}$ is the clutter estimate.

First I shall consider the system response to ground clutter and then to noise. The response for target returns is the same as that of the optimum detector because these returns do not enter the estimator. It will be shown that the system response to the clutter is zero so the output is circular symmetric Gaussian under both hypotheses. Therefore the SNR penalty can be determined directly by comparing the SNR with that of the optimum processor.

The estimator's steady-state response to constant clutter, $\bar{E}_c(l)$, is the constant clutter return $\bar{c}$ so when this is subtracted from the returns to be integrated the constant clutter component is removed. To show that the output of the estimator is indeed $\bar{c}$ consider the step response $\bar{E}_s(l)$ of the estimator, where the subscript s denotes step response

$$\bar{E}_s(l) = \bar{c} \left(1 - \left(1 - \frac{1}{n} \right)^{l+1} \right) u(l) \quad (9.1)$$

where $u(\cdot)$ is the discrete unit step function. Since the term $1 - 1/n$ is less than 1, as l grows large this term disappears and the output of the estimator is $\bar{c}$.

The output of the estimator for uncorrelated complex Gaussian noise at the input is complex Gaussian. The variance σ_E^2 of the real and imaginary parts are the same and in section 9.4 on mathematical details it is shown that

$$\sigma_E^2 = \frac{\sigma_n^2}{2n - 1} \quad (9.2)$$

where σ_n^2 is the variance of the real and imaginary parts of the

noise and the subscript E denotes estimate:

The output $\bar{y}_n$ due to the noise input is

$$\bar{y}_n = \sum_{l=0}^{N-1} (\bar{n}(l) - \bar{E}_n) \exp(-j\omega_D \Delta l) \quad (9.3)$$

or

$$\bar{y}_n = \sum_{l=0}^{N-1} \bar{n}(l) \exp(-j\omega_D \Delta l) - \bar{E}_n \sum_{l=0}^{N-1} \exp(-j\omega_D \Delta l) \quad (9.4)$$

The first term of this equation is the sum of N independent Gaussian noise terms progressively rotated. Since the terms $\bar{n}(l)$ are distributed circular symmetrically and are independent, the rotation is irrelevant. The result is a random vector (or a complex random variable) that is circularly symmetric Gaussian with single component variance $N\sigma_n^2$. The second term is independent of the first since it is made up of returns previous to those in the first term and all the returns are independent. This term is also a Gaussian circular symmetric random vector (or complex random variable). Again the net rotation of the sum of exponentials does not matter but this time because only a single value $\bar{E}_n$ is being rotated.

The resulting component variance of the second term is

$$\sigma_e^2 = \frac{\sigma_n^2}{2n-1} \left| \sum_{l=0}^{N-1} \exp(-j\omega_D \Delta l) \right|^2 \quad (9.5)$$

where the subscript e denotes estimate variance for $\bar{y}_n$

or

$$\sigma_e^2 = \frac{\sigma_n^2}{2n-1} \left(\frac{\sin(\omega_D \frac{\Delta N}{2})}{\sin(\omega_D \frac{\Delta}{2})} \right)^2 \quad (9.6)$$

Thus the output $\bar{y}_n$ due to the uncorrelated Gaussian noise is a complex random variable with a Gaussian circular symmetric distribution and single component variance σ^2 equal to the sum of the variance of the two sums.

$$\sigma^2 = N\sigma_n^2 + \sigma_e^2 \quad (9.7)$$

or

$$\sigma^2 = \sigma_n^2 \left(N + \frac{1}{2n-1} \left(\frac{\sin(\omega_D \frac{\Delta N}{2})}{\sin(\omega_D \frac{\Delta}{2})} \right)^2 \right) \quad (9.8)$$

Three things should be noted about the variance σ^2 . First as N , the number of returns integrated, increases the variance increases. Second as n , the estimator parameter, gets larger the variance decreases. This is because the estimator, which is actually a first order recursive filter, has a narrower frequency response for larger n thus decreasing the amount of noise that gets through.

Thirdly as ω_D , the Doppler frequency of interest, changes the variance changes in a periodic manner because of the second term. It is maximum when $\omega_D \Delta$ is a multiple of 2π . In effect the estimator output is a constant term in the Doppler filter input so it receives no attenuation if the filter is set for a maximum response to a constant or what appears to be a constant due to the sampled nature of the system. The maximum variance is

$$\sigma_{\max}^2 = \sigma_n^2 \left(N + \frac{N^2}{2n-1} \right) \quad (9.9)$$

and the minimum is

$$\sigma_{\min}^2 = \sigma_n^2 N \quad (9.10)$$

Now, since the target response for the coherent clutter map is the same as for the optimum processor, the signal-to-noise ratio penalty is found by simply comparing the noise response. Therefore

$$\text{SNR}_{\text{pm}} = \frac{\sigma^2}{\sigma_{\text{opt}}^2} = \frac{\sigma_n^2 \left(N + \frac{1}{2n-1} \left(\frac{\sin(\omega_D \frac{\Delta N}{2})}{\sin(\omega_D \frac{\Delta}{2})} \right)^2 \right)}{N \sigma_n^2} \quad (9.11)$$

or

$$\text{SNR}_{\text{pm}} = 1 + \frac{1}{N(2n-1)} \left(\frac{\sin(\omega_D \frac{\Delta N}{2})}{\sin(\omega_D \frac{\Delta}{2})} \right)^2 \quad (9.12)$$

where the subscript p denotes penalty and m, the coherent clutter map:

Figure 9.6 shows the signal-to-noise ratio penalty as a function of Doppler frequency with the estimator parameter n as a parameter and $N = 8$. The signal-to-noise ratio penalty is independent of the clutter-to-noise ratio so the coherent clutter map performs as well in weak clutter as in strong. Additionally, there are no blind speeds as the signal-to-noise ratio penalty is finite at all Doppler frequencies, but the performance is poorer at low Doppler frequencies because the filter does not reduce the noise from the estimator as it does for the high frequencies.

figure 9.6 the signal-to-noise ratio penalty for the coherent clutter map. n is the estimator (update) parameter.

Increasing n reduces the signal-to-noise ratio penalty. For n as small as 10 the penalty never exceeds ~~1.5~~ dB and for 80% of the Doppler frequency spread, it is below 0.25 dB. If the clutter is fairly stable the signal-to-noise ratio penalty can be made so small as to be negligible.

From equation 9.12 it can be seen that increasing N , the number of returns processed, increases the SNR penalty. For example for $\omega_D = 0$ the penalty reaches a maximum of

$$\text{SNR}_{\text{pm}} = 1 + \frac{N}{2n-1} \quad (9.13)$$

In practical applications one would expect the clutter returns to be coherent over a much longer period of time than the target returns. Since the time for which the target is coherent dictates the practical value of N and the time for the clutter dictates n we have for most applications the possibility of setting n so that

$$n \gg N \quad (9.14)$$

Therefore the SNR penalty can be made negligible.

Increasing N also increases the number of nulls and maximums in the response. This does not affect performance however.

9.3 Chapter Highlights

This chapter has presented the concept of a coherent clutter map. It was observed that step-scan radars preserved the intrinsic nature of ground clutter thus making it possible

to remove the clutter coherently. To do this a coherent clutter map of the radar environment had to be determined by estimating the ground clutter component of radar returns. A simple recursive estimator consisting of two storage elements per detection bin was recommended for this job.

The resulting system built up a coherent clutter map using returns prior to those to be tested for a target. If no target was detected in the returns presently being processed they were allowed to enter the estimator to continue the map update.

The signal-to-noise ratio penalty for such a scheme depended on three variables: the Doppler frequency ω_D , the number of returns processed N and the update parameter n . The penalty was a periodic function of ω_D with one major maximum and a number of submaximums per period. The maximum SNR penalty increased with increasing N and decreased with increasing n . In practical applications it was observed that one could choose $n \gg N$ resulting in a very small SNR penalty.

In conclusion, the coherent clutter map as presented in this chapter deals effectively with ground clutter, providing essentially optimum detection performance across the entire Doppler frequency band.

9.4 Mathematical Details

This section presents the mathematical detail leading to a result used in chapter 9. It is recommended that readers not interested in this detailed development disregard this section

to preserve the continuity of the thesis discussion.

9.4.1 Determining σ_E^2

In this subsection I derive the single component variance for the estimator output when the input is uncorrelated Gaussian noise.

Consider the recursive mean estimator shown in figure 9.7.

The output $y(l)$ of this system is

$$y(l) = \frac{1}{n} x(l) + \left(1 - \frac{1}{n}\right) y(l-1) \quad (9.15)$$

Since the mean of the uncorrelated Gaussian noise input $x(l)$ is zero so is the mean of $y(l)$. Therefore

$$\sigma_E^2 = E\{y(l)^2\} \quad (9.16)$$

where $E\{\cdot\}$ denotes expectation and we have

$$\sigma_E^2 = \frac{1}{n^2} E\{x(l)^2\} + \frac{2}{n}\left(1 - \frac{1}{n}\right) E\{x(l) y(l-1)\} + \left(1 - \frac{1}{n}\right)^2 E\{y(l-1)^2\} \quad (9.17)$$

Since $x(l)$ is uncorrelated with all previous x 's which make up $y(l-1)$ the expectation $E\{x(l) y(l-1)\}$ is zero. In addition, $E\{x(l)^2\}$ is equal to σ_n^2 and in the steady state $E\{y(l-1)^2\}$ is equal to σ_E^2 . So solving equation 9.17 for σ_E^2 we have

$$\sigma_E^2 = \frac{\sigma_n^2}{2n-1} \quad (9.18)$$

which is the desired result.

figure 9.7 recursive mean estimator.

CHAPTER 10

CONCLUSIONS FOR PART 2

This chapter focuses attention on the important results of Part 2. It is divided into three sections: the first summarizes the need for and the objective of the investigation, the second section reviews the results and provides a detailed comparison of the clutter suppression schemes, while the third and final section discusses the practical implications of the results.

10.1 Need and Objective

Perhaps the single most important application for radar is airspace surveillance. With the ever increasing amount of air traffic, successful tracking in difficult clutter conditions is necessary for safe air travel.

For most airspace surveillance radars ground clutter is the major problem. When an aircraft has a relatively large radial velocity, Doppler processing can be successfully used to detect targets in this clutter. However, if the aircraft's track is tangential to the radar there is no Doppler shift so other techniques such as variable thresholding must be used.

These conventional approaches have proved useful in combatting ground clutter but are not entirely satisfactory. Constraints imposed by the continuous scan of the antenna limit

system design possibilities and so curtail performance. With the advent of the new step-scan phased array antennas however, these constraints have been lifted permitting new system concepts to be considered. The most significant change resulting from the step-scan is that the ground clutter returns are no longer corrupted by antenna motion and therefore possess long term stability. It is possible to take advantage of this stability to remove the ground clutter coherently from the return. When this is done the radar operates as if the ground clutter were not present at all.

It was the objective of Part 2 of this thesis to investigate the merits of coherent ground clutter removal for step-scan airspace surveillance radars.

10.2 Summary and Comparisons

This section is divided into two subsections. The first summarizes the main results of the chapters dealing with signal models, receiver structures and decision processors. The second subsection compares the relative merits of the various clutter suppression schemes in terms of the signal-to-noise ratio penalty.

10.2.1 Summary of Models

In order to make meaningful conclusions regarding the performance of various clutter suppression schemes, realistic models of the detection problem had to be formulated. The first

step in this direction was provided by chapter 6 where the signal models, receiver structures and decision processes consistent with the airspace surveillance application were developed.

It was found that a receiver structure consisting of a bandpass filter whose bandwidth is matched to the inverse of the pulse length is almost ideal provided the pulse length is short compared to the period of the highest Doppler frequency. This condition is met for airspace surveillance radars since good range resolution is requisite. Doppler information is usually obtained by comparing the relative phases of consecutive returns.

The decision process consists of two elements: a Doppler filter and a thresholding scheme. Refinements to these basic elements were made to model different clutter suppression schemes. Neyman-Pearson hypothesis testing criterion was used for making decisions because it lends itself readily to the radar application.

In chapter 7 the models developed in chapter 6 were tailored to suit the target in ground clutter detection problem. Ground clutter was modelled as a constant signal vector, the target as a Gaussian signal vector perfectly correlated from return to return, and the noise as a Gaussian signal vector uncorrelated from return to return. In addition the target returns were assumed to be uncorrelated between groups of N returns to reflect the statistical fluctuations due to target motion.

The optimum processor was then derived for these signal

models. It consisted of coherent clutter removal followed by Doppler filtering. The detection performance for this processor, which was to serve as a basis for comparing clutter suppression schemes, was then determined.

10.2.2 Comparisons

In chapter 8 the two major types of conventional ground clutter suppression schemes were discussed: variable thresholding and pulse cancellation. The new scheme (the coherent clutter map) was discussed in chapter 9. Figure 10.1 compares the conventional schemes with the coherent clutter map in terms of the signal-to-noise ratio penalty. Only the maximum penalty for the variable threshold T_A is shown to avoid confusion; the minimum penalty is within 2 dB. For the figure, N was chosen to be 8 and the clutter-to-noise ratio CNR to be 15 dB.

It can be seen that for this moderate value of CNR the SNR penalties for both variable thresholds, T_G and T_A , are large for low Doppler frequencies. However, detection is possible because the penalty is finite unlike that for the pulse canceller. The penalty is higher for T_G than for T_A because T_G was found assuming the clutter and noise distribution to be Gaussian with zero mean which was not actually the case. T_G resulted in a lower probability of false alarm than that assigned and that wastes system resources. On the positive side, T_G is easy to determine making it the most popular choice for practical systems.

T_A provides the assigned probability of false alarm and

figure 10.1 signal-to-noise ratio penalty comparison.
CNR is 15 dB and N is 8.

therefore performs better than T_G but T_A must be determined by counting false alarms which is difficult.

A serious problem with the variable threshold schemes is that the SNR penalty increases with increasing clutter levels (figure 8.8). This means that targets could be lost as they move from areas of weak clutter to areas of strong clutter.

Pulse cancellers on the other hand, have a SNR penalty that is independent of the clutter level making them equally effective in both weak and strong clutter areas. This asset however, is offset by the fact that the SNR penalty goes to infinity at various periodic Doppler frequencies including zero. ~~Therefore targets travelling at the corresponding radial velocities cannot be detected.~~

The radar system design philosophy that has been built around the above mentioned advantages and disadvantages of the conventional clutter suppression schemes is the following (Muehe et al. 1974, O'Donnell et al. 1974, O'Donnell and Muehe 1979). For low speed targets or targets flying tangential to the radar, variable thresholding is used (specifically T_G) to suppress ground clutter. If the target has a significant radial velocity, pulse cancellation is used.

This design philosophy does not take advantage of the long term stability observed in ground clutter when a step-scan antenna is used. Such long term stability permits one to store coherent information about the ground clutter environment and use this information later to enhance target detection. The

coherent clutter map scheme discussed in chapter 9 performs this task.

From figure 10.1 it can be seen that for an estimator parameter as low as 10 the coherent clutter map outperforms any of the conventional schemes across the entire Doppler frequency band. In addition, as was shown in section 9.2, the penalty is independent of clutter level so the coherent clutter map operates equally well in both heavy and light clutter.

In practice the estimation parameter can probably be much larger than 10 since it is expected that the stationary ground clutter will be stable over many returns. Therefore the performance of the coherent clutter map would approach that of the optimum detector for all Doppler frequencies.

These observations imply that the coherent clutter map scheme for suppressing ground clutter does indeed have merit. In the next section some practical aspects of ground clutter suppression are discussed in the light of the performance comparisons just made.

10.3 Practical Considerations

In the last section it was concluded that the coherent clutter map was the best scheme for suppressing ground clutter. In fact, its performance was essentially that of the optimum processor. This section puts this conclusion into context by considering some of the practical aspects of radar systems and of the study itself.

The section is divided into two subsections: the first dealing with radar systems, and the second with the study.

10.3.1 Radar System Considerations

The coherent clutter map is a viable technique for removing ground clutter in radars that have constant ground clutter returns. This implies that the radar must be stationary and employ a step-scan antenna. Fortunately, airspace surveillance radars utilizing phased array antennas meet these requirements thus making it possible, with the use of a coherent clutter map, to improve their tracking performance in a ground clutter environment. In implementing the system however, two important factors must be considered.

The first is the phase stability of the radar. For proper operation the coherent clutter map requires a radar that is coherent and stable. It must be able to recover the phase of the return and relate it faithfully to that of the previous return. This does not pose any great problem for modern radar systems because they are already capable of this type of processing as evidenced by the current use of base band Doppler processing (O'Donnell et al. 1974).

The second factor that must be considered is the need for linear signal processing in the receiver. Nonlinear elements such as logarithmic amplifiers (used to increase dynamic range) and hard limiters must not be used because they distort the

incoming signal and spread its spectrum. Muehe et al. (1974) discuss this aspect of the signal processing and conclude that linear receiver systems are necessary. In addition they built and tested a linear receiver (60 dB dynamic range) which is now being used in their MTD (moving target detector) system.

As far as the logistics of the decision making process is concerned employment of a coherent clutter map has two obvious benefits. First the probability distribution of the test statistic is no longer displaced but is circular symmetric about the origin. This implies, for our model at least, that the easily determined threshold T_G is optimum. For practical situations it probably means that T_G is almost optimum.

The second benefit comes from the ability to choose which detection bins will be processed for ground clutter and which will not. This flexibility is important since only a portion of the radar space is affected by ground clutter and most clutter suppression schemes degrade detection performance even in the absence of clutter. For example, if a pulse canceller were used, tangential targets would be lost in both the clutter and no clutter regions.

In practice, at the first level the clutter map could consist of nothing more than a list of those detection bins requiring correction for ground clutter and the correction. This would keep the memory requirements at a minimum while also providing an opportunity for adding or subtracting bins as the requirements change.

In addition to the above approach which stores a map of the ground clutter environment for the entire radar space, the map could be restricted to specific areas of clutter. For example, a radar tracking an aircraft can predict the next detection bin that the target is most likely to be in. Since the target would not have arrived there yet a small clutter map could be built up around the expected location thus effectively preparing the area for target detection. This kind of clutter mapping would not be much help in the initial detection stages but after a track has been established it could be a great help in following it through a heavy clutter region.

This concludes the subsection on radar systems considerations. I shall now proceed to consider the study itself and the recommendations for future research.

10.3.2. The Study and Recommendations

The objective of this study was to determine if there was any merit in considering a coherent clutter map for removing ground clutter in step-scan airspace surveillance radars. The results conclusively show that the coherent clutter map is worth further consideration. This conclusion, of course, is based on the models used in the study and as with most studies that deal with new concepts assumptions had to be made concerning some aspects of the models.

For example, it was assumed that the target returns were perfectly correlated within a group of N returns. In practice one would not expect to observe perfect correlation as studies such as that done by Von Schlachta (1974) show the correlation to be high but not perfect. With this in mind however, and lacking any definite form for a model, it is expedient to use the ideal. Otherwise, the search for the most realistic model would preclude the study altogether.

The models used in this study are not new but have been the basic building blocks for radar detection studies for years. The reasons why they continue to be useful are: they provide a basis for comparing various systems, they are tractable mathematically, and they are a "reasonable" representation of the real world.

One could propose different statistics for target and noise and invent alternative detection situations to be studied in the context of the coherent clutter map. In fact, the choices are endless. I have simply chosen the most appropriate conventional models to establish the worth of the concept. Further research should build upon this study by experimentally verifying the conclusion.

Experiments for this purpose need not be performed using a step-scan antenna and a coherent clutter map as proposed in the study. Simply holding the antenna stationary mimics the step-scan and reconfiguring a conventional pulse canceller provides a recursive estimator. Therefore a conventional

coherent radar capable of base band coherent processing is adequate to experimentally verify the merit of the coherent clutter map clutter removal scheme.

In addition to the radar systems experiments the coherent nature of ground clutter returns needs to be investigated. The literature is surprisingly sparse on this subject. The radars of the future will need this information to take advantage of advances in signal processing systems.

This concludes Part 2 of the thesis and the discussion of the coherent clutter map. Chapter 11 reviews the objectives of the thesis and lists the accomplishments.

CHAPTER 11

THESIS CONCLUSIONS

This thesis has had two major objectives. The first objective was to present a general approach to detection problems and the second was to investigate the merit of coherent ground clutter removal in airspace surveillance radars. The thesis was correspondingly divided into two parts: the first meeting the first objective and the second the second objective.

Several significant new results were presented in both parts. In Part 1 the major new result was a probability calculation technique that made the general approach to detection problems practical. The general approach merged concepts that have been used successfully in isolated areas to form a systematic method for attacking a large class of detection problems. Specifically, the concepts of the phase plane representation of signals, complex signal notation and circular symmetric probability distribution theory were merged to provide a conceptually simple and mathematically concise approach to detection problems.

I feel, the major reason this approach has been overlooked in the past as a general method for solving detection problems is that the mathematically concise solutions were difficult to evaluate. This difficulty led researchers to consider approaches specifically tailored to their problems thus giving the impression that no general approach existed.

Since the calculation technique presented in chapter 3 alleviates the problem of evaluating the mathematical expressions derived when using the general approach, it now becomes practical and somewhat self-evident.

The calculation technique itself is a nontrivial extension of the Fourier-Bessel series approach first used by Barakat (1974) and then by Jao and Elbaum (1978). Their results were extended in this thesis to arbitrary dimensions and an error bounding technique developed. It is the ability to easily obtain a result to a guaranteed accuracy that makes the technique practical and important.

To demonstrate the utility of the general approach to detection problems several significant examples were considered. In the first example the probability of error for CPSK communications systems operating in cochannel interference and Gaussian noise was found. The results were not new (Rosenbaum 1970 and Goldman 1974) but the ease with which they were obtained was remarkable. In addition it was shown how the Rosenbaum and Glave upper bounds for the probability of error for M-ary CPSK operating in peak limited interference and Gaussian noise could be easily found.

As a second example the difficult and apparently heretofore unsolved problem of calculating the probability of error for multi-level ASK in the presence of cochannel interferers and Gaussian noise was treated, but not in great detail as that was not the intended purpose of the example. Rather the approach to the problem was demonstrated and several interesting sample calculations made.

The ease with which the problem was solved was again remarkable.

As a final example of the general approach to detection problems the probability of detection for a classical radar/sonar target in Gaussian noise was calculated. This target resulted in a constant amplitude random phase signal being added to the noise. Therefore calculating the probability of detection required one to evaluate the well known but notoriously difficult radar Q-function. Again the approach used permitted this problem to be solved easily.

This general approach to detection problems is a worthwhile contribution for several reasons. The first and perhaps the most important is that it permits one to attack these problems in a systematic and general manner alleviating the need for specialized approaches. The general approach is also conceptually simple sometimes making formerly difficult problems seem trivial. In addition to these assets, the chore of actually calculating performance measures is reduced to a single method that requires no particular computer expertise. This puts the method within the reach of most practicing engineers.

The most significant result of Part 2 is the establishment of the feasibility of a new scheme for suppressing ground clutter in airspace surveillance radars. This new scheme, the coherent clutter map, is a result of recognizing that ground clutter returns for step-scan radars are different from those for continuous scan radars. The difference is that for step-scan radars, ground clutter returns possess long term stability thus permitting them to be

stored and removed coherently from future returns.

This observation led to the concept of the coherent clutter map. This map provided coherent ground clutter information about the radar environment. The information is subtracted from future radar returns to enhance target detection. It was found that the coherent clutter map was significantly better than the conventional methods for handling ground clutter and in fact it yielded almost optimum performance.

To make the above mentioned comparisons a performance measure called the signal-to-noise ratio penalty was defined. This penalty was shown to be relevant for each clutter suppression scheme and provided a simple, easily interpreted measure of detection performance. With the use of this measure the relative merits of each scheme were easily determined with the coherent clutter map proving to be the best.

This result is important because of the prominent role radar plays in air traffic control. Ground clutter interferes with target detection causing aircraft tracks to be lost. This is especially true when the aircraft is flying tangential to the radar. Thus there is a great need for a scheme that will reduce the effect of ground clutter.

This study has established the merit of the coherent clutter map scheme for the airspace surveillance application and has shown that further research is warranted.

In conclusion, this thesis provides a practical and powerful general approach for the analysis of detection problems and demonstrates the feasibility of the coherent clutter map for handling ground clutter in step-scan airspace surveillance radars.

REFERENCES

Abramovitz, M. and I. Stegun "Handbook of Mathematical Functions," New York, Dover, 1965.

Barakat, R. "First-order Statistics of Combined Random Sinusoidal Waves With Applications to Laser Speckle Patterns," Opt. Acta., vol. 21, pp. 903-921, 1974.

Blythe, J.H. and R. Treciokas "The Application of Temporal Integration to Plot Extraction," IEE Radar '77, pp. 275-279, October 1977.

Burgess, J.S. "Recent Advances in Radar Technology," IEE Radar '77, pp. 8-11, October 1977.

Dillard, G.M. "Signal-to-Noise Ratio Loss in an MTI Cascaded with Coherent Integration Filters," IEEE International Radar Conference, pp. 117-122, 1975.

Esposito, R. and L.R. Wilson "Statistical Properties of Two Sine Waves in Gaussian Noise," IEEE Trans. Inform. Theory, vol. IT-19, pp. 176-183, 1973.

George, D.A. "Study on the Methods of Radar Plot Extraction, Adaptive Thresholding, and Area Moving Target Indication," A Study Undertaken for the Radar Research Laboratory, Communications Research Centre, Department of Communications, Ottawa, Under Department of Supply and Services Contract #OST77-00122, March 1978.

Giacconi, E. and G. Nucci "A Family of Air Traffic Control Radars," IEEE Trans. AES, vol. AES-15 No. 3, pp. 378-396, May 1979.

Glave, R.E. and A.S. Rosenbaum "An Upper Bound Analysis for Coherent Phase-shift Keying with Cochannel, Adjacent-channel, and Intersymbol Interference," IEE Trans. Commun., vol. COM-23, pp. 586-597, June 1975.

Goldman, J. "Statistical Properties of a Sum of Sinusoids and Gaussian Noise and Its Generalization to Higher Dimensions," Bell Syst. Tech. J., vol. 53, pp. 557-580, April 1974.

Gradshteyn, I.S. and I.M. Ryzhik "Table of Integrals, Series, and Products," Academic Press, 1965, ISBN 0-12-294750-9.

Helstrom, C.W. "Statistical Theory of Signal Detection," Pergamon Press, 1960.

Irabu, T., E. Kiuchi, T. Hagiwara and Y. Tomita "Range-Azimuth Correlation Processor in the Frequency Domain for a Primary Radar," IEE Radar '77, pp. 368-372, October 1977.

Jao, J.K. and M. Elbaum "First-order Statistics of a Non-Rayleigh Fading Signal and Its Detection," Proc. of the IEEE, vol. 66 No. 7, pp. 781-789, July 1978.

Lehman, E.L. "Testing Statistical Hypotheses," New York, Wiley, 1959.

Long, M. "Radar Reflectivity of Land and Sea," Lexington Books, D.C. Heath and Company, 1975.

Lord, R.D. "The Use of Hankel Transforms in Statistics," Biometrika, vol. 41, pp. 44-55, 1954.

Marcum, J.I. "A Statistical Theory of Target Detection by Pulsed Radar, Mathematical Appendix," IRE Trans., vol. IT-6, pp. 145-267, April 1960.

Meyer, D.P. and H.A. Mayer "Radar Target Detection," Academic Press, 1973.

Mitchell, R.L. and J.F. Walker "Recursive Methods for Computing Detection Probabilities," IEEE Trans. Aero. and Elec. Syst., vol. AES-7 No.4, pp. 671-676, July 1971.

Muehe, E., L. Cartledge, W.H. Drury, E.M. Hofstetter, M. Labitt, P.B. McCorison and V.J. Sferrino "New Techniques Applied to Air Traffic Control Radars," Proc. of the IEEE, vol. 62 No.6, pp. 716-723, June 1974.

Nathanson, F.E. "Radar Design Principles," McGraw-Hill Book Company, 1969.

Neyman, J. and E. Pearson "On the Problem of the Most Efficient Tests of Statistical Hypotheses," Phil. Trans. Roy. Soc., vol. 231, pp. 289, 1933.

Nitzberg, R. "Effect of a Few Dominant Specular Reflectors Target Model Upon Target Detection," IEEE Trans. Aero. and Elec. Syst., vol. AES-14 No. 4, pp. 670-673, July 1978.

Noble, B. "Applied Linear Algebra," Prentice-Hall, Inc., 1969, (04122).

O'Donnel, R.M., C.E. Muehe, M. Labitt, W.H. Drury and L. Cartledge "Advanced Signal Processing for Airport Surveillance Radars," EASCON '74, pp. 71-71F, 1974.

O'Donnel, R.M. and C.E. Muehe "Automatic Tracking for Aircraft Surveillance Radar System," IEEE Trans. Aero. and Elec. Syst., vol. AES-15 No. 4, pp. 508-517, July 1979.

Papoulis, A. "Joint Densities with Circular Symmetry," IEEE Trans. Info. Theory, vol. IT-14 No. 1, pp. 164-165, January 1968.

Proc. IRE "IRE Standards on Radio Aids to Navigation: Definition of Terms," proc. IRE, vol. 43, pp. 194, February 1955.

Rickard, J.T. and G.M. Dillard "Adaptive Detection Algorithms for Multiple-target Situations," IEEE Trans. Aero. and Elec. Syst., vol. AES-13 No. 4, pp. 338-342, July 1977.

Rosenbaum, A.S. "Binary PSK Error Probabilities with Multiple Cochannel Interferences," IEEE Trans. Commun. Technol., vol. COM-18, pp. 241-253, June 1970.

Rosenbaum, A.S. and F.E. Glave "An Error Probability Upper Bound for Coherent Phase-shift Keying with Peak-limited Interference," IEEE Trans. Commun., vol. COM-22, pp. 6-16, January 1974.

Skolnich, M.I. "Introduction to Radar Systems," McGraw Hill Book Company, 1962.

Spiegel, M.R. "Mathematical Handbook of Formulas and Tables," Schaum's Outline series, McGraw-Hill Book Company 1968.

Swerling, P. "Probability of Detection for Fluctuating Targets," IRE Trans., vol. IT-6, pp. 269-308, April 1960.

Van Trees, H.L. "Detection, Estimation and Modulation Theory - Part III," John Wiley and Sons, Inc., 1971.

Von Schlachta, K. "Remarks on Target Models for the Design of Radar Systems," IEEE International Radar Conference, pp. 440-445, 1974.

Wald, A. "Contributions to the Theory of Statistical Estimation and Testing of Hypotheses," Ann. Math. Stat., vol. 10, pp. 299, 1939.

Watson, G.N. "A Treatise on the Theory of Bessel Functions," Cambridge at the University Press, 1962.

APPENDIX A

NONCOHERENT INTEGRATION

The subject of noncoherent integration of radar returns has been extensively researched starting with Marcum in the 1940's. The classical papers by Marcum and by Swerling, reprinted in the IRE Transactions on Information Theory in 1960, give most of the results of this work and are widely accepted as the starting point of radar detection theory. This thesis, because it is dealing with coherent radar and the possibility of coherent returns, uses probability distributions that have been developed in the coherent domain of the signal (predetection). Marcum's and Swerling's work dealt mainly with postdetection distributions. The object of this appendix is to show where the distributions derived by Marcum and Swerling fit into the framework of the present discussion. Readers interested in a more complete discussion of this aspect of radar detection theory should refer to Meyer and Mayer (1973).

The first topic to be discussed in this appendix is the relationship between the receiver statistic and noncoherent integration. Then the standard fluctuation statistics due to Swerling are introduced and finally a very interesting special case of noncoherent integration is discussed.

A.1 The Detectors

Noncoherent integration is the summation of a number of returns after detection. Often the returns are squared before they are summed. In terms of our previous discussion, noncoherent integration is the sum of a number of samples of the random variable r where r is the length of the receiver statistic $\bar{r}$ defined by equation 6.26. Let r_i denote the value of r for the i^{th} return. Then the output of the integrator r_t is

$$r_t = r_1 + r_2 + \dots + r_N \quad (\text{A.1})$$

or

$$r_t = |\underline{x}_1| + |\underline{x}_2| + \dots + |\underline{x}_N| \quad (\text{A.2})$$

where $\underline{x}_i$ is the vector receiver statistic or equivalently the complex envelope of the i^{th} return. Therefore

$$r_t = (x_{1R}^2 + x_{1I}^2)^{\frac{1}{2}} + (x_{2R}^2 + x_{2I}^2)^{\frac{1}{2}} + \dots + (x_{NR}^2 + x_{NI}^2)^{\frac{1}{2}} \quad (\text{A.3})$$

where x_{iR} is the real part of the complex envelope and x_{iI} is the imaginary part.

The distribution of r_t cannot be expressed in a useful way in terms of the symmetric or simple nonsymmetric distributions of $\underline{x}$. To obtain the distribution of r_t Marcum (1960) and Swerling (1960) computed the one-dimensional characteristic function of $p(r)$ and then found the product of the characteristic functions and used the inverse transform to obtain $p_{r_t}(r)$. In actual fact it was not the distribution of r_t that was found but of r_d^2 which is defined as

$$r_d^2 = r_1^2 + r_2^2 + \dots + r_N^2 \quad (A.4)$$

or

$$r_d^2 = x_{1R}^2 + x_{1I}^2 + x_{2R}^2 + x_{2I}^2 + \dots + x_{NR}^2 + x_{NI}^2 \quad (A.5)$$

The distributions of r_d were mathematically easier to calculate and the detection performance difference between the two measures was small. r_t is the output of what is called a linear detector and r_d^2 is the output of what is called a square-law detector with integration. Marcum (1960) discussed the relative merits of the two detectors and compares their performance to an optimum detector (found using the likelihood ratio) for specific return statistics. Therefore noncoherent signal processing, and specifically integration, of radar returns involves the combining of the distributions of r or r^2 . Extensive tables of results for various combinations of target statistics are presented by Meyer and Mayer (1973).

A.2 Fluctuation Statistics

Marcum (1960) treated the problem of detecting a constant (nonfluctuating) target in Gaussian noise. Swerling (1960) later introduced target fluctuations into the model so that performance calculations would have better agreement with reality. Using Marcum's results for a constant target (constant signal-to-noise ratio) Swerling proposed various distributions for the signal-to-noise ratio and then integrated the distribution conditioned on this ratio over all possible signal-to-noise ratio values. The family of distributions proposed by Swerling for the signal-to-

noise ratio was the gamma distribution which is more commonly known in radar literature as the chi-square family of distributions. It has the following form.

$$p(y|y_a) = \frac{1}{\Gamma(k)} \left(\frac{k}{y_a} \right)^k y^{k-1} \exp\left(-\frac{ky}{y_a}\right) \quad y > 0 \quad (\text{A.6})$$

where y is the signal-to-noise ratio, y_a the average signal-to-noise ratio, $\Gamma(\cdot)$ the gamma function, and k the target fluctuation parameter. Different values of k correspond to the different Swerling target models.

$0 < k < 1$	Weinstock case (see Meyer and Mayer 1973, page 64)
$k = 1$	Swerling case 1
$k = 2$	Swerling case 3
$k = N$	Swerling case 2
$k = 2N$	Swerling case 4
$k = \infty$	Marcum or nonfluctuating case 0

where N is the number of returns processed.

The Weinstock model was derived from experimental results in order to take care of the special case of a randomly oriented cylinder class of targets where the usual Swerling models did not apply.

Swerling case 1 and 2 apply for targets that can be represented as a number of independent fluctuating reflectors of about equal echoing area. Case 1 is used when echo fluctuations are statistically independent from scan to scan and 2 applies when fluctuations occur from return to return.

Swerling cases 3 and 4 apply for targets that can be represented as one large reflector together with a number of

small reflectors, or as one large reflector subject to small changes in orientation. Case 3 applies when echo fluctuations are statistically independent from scan to scan and case 4 applies when fluctuations occur from return to return.

The Marcum or nonfluctuating case 0 applies to large spherical targets such as balloons.

It should be noted that Marcum's and Swerling's tabulated results can only be applied to the detection of targets in Gaussian noise although the theoretical approach is sound for all distributions.

A.3 A Special Case

An interesting special case where the theory of symmetric distributions can be used in determining the distribution of r_d^2 or equivalently r_d arises when the distributions of the $\underline{x}_i$ are all offset circular symmetric Gaussian random vectors. First consider the expression for r_d^2 :

$$r_d^2 = x_{1R}^2 + x_{1I}^2 + x_{2R}^2 + x_{2I}^2 + \dots + x_{NR}^2 + x_{NI}^2 \quad (A.7)$$

Now suppose that each x_{iR} and x_{iI} may be written as

$$\begin{aligned} x_{iR} &= y_{iR} + \mu_{iR} \\ x_{iI} &= y_{iI} + \mu_{iI} \end{aligned} \quad (A.8)$$

where y_{iR} is a one-dimensional Gaussian random variable with variance σ^2 and zero mean, μ_{iR} is a constant, y_{iI} is a one-dimensional Gaussian random variable with variance σ^2 zero mean and independent of y_{iR} , and μ_{iI} is a constant. If all

the x_i 's are independent of each other then r_d is just the length of a $2N$ -dimensional vector $\underline{x}_t$

$$\underline{x}_t = \underline{y}_t + \underline{\mu}_t = \begin{bmatrix} x_{1R} \\ x_{1I} \\ x_{2R} \\ x_{2I} \\ \vdots \\ x_{NR} \\ x_{NI} \end{bmatrix} = \begin{bmatrix} y_{1R} \\ y_{1I} \\ y_{2R} \\ y_{2I} \\ \vdots \\ y_{NR} \\ y_{NI} \end{bmatrix} + \begin{bmatrix} \mu_{1R} \\ \mu_{1I} \\ \mu_{2R} \\ \mu_{2I} \\ \vdots \\ \mu_{NR} \\ \mu_{NI} \end{bmatrix} \quad (\text{A.9})$$

$\underline{y}_t$ is a $2N$ -dimensional random vector with a spherical symmetric distribution and $\underline{\mu}_t$ is a simple vector shift in location of this distribution. It should be noted that if the underlying distribution is not Gaussian $\underline{y}_t$ will not have a spherical symmetric distribution (Papoulis 1968).

Since $\underline{y}_t$ is a Gaussian spherical symmetric random vector of dimension $2N$ with distribution $p(r, \theta)$ equal to

$$p(r, \theta) = \frac{1}{(2\pi)^N \sigma^N} \exp\left(-\frac{r^2}{2\sigma^2}\right) \quad (\text{A.10})$$

the characteristic function from equation 2.19 is

$$\phi_{\underline{y}_t}(\rho) = \frac{(2\pi)^N}{\rho^{N-1}} \int_0^\infty r^N J_{N-1}(r\rho) \frac{1}{(2\pi)^N \sigma^N} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr \quad (\text{A.11})$$

From Watson (1962, page 394) we obtain

$$\phi_{\underline{y}_t}(\rho) = \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) \quad (\text{A.12})$$

The spherical symmetric distribution of y_t is offset from the origin by the mean vector $\underline{\mu}_t$. This is a simple location shift of the distribution so the characteristic function of the equivalent symmetric distribution (equivalent in the sense that it gives the same distribution for r_d as the offset distribution) can be found by determining the characteristic function of a constant amplitude (equal to $|\underline{\mu}_t|$) random phase 2N-dimensional vector and multiplying it by $\Phi_{y_t}(\rho)$ (subsection 2.2.5). A 2N-dimensional constant amplitude vector has a distribution for r and θ equal to

$$p(r, \theta) = \frac{\Gamma(N)}{2 a^{2N-1} \pi^N} \delta(r-a) \quad (A.13)$$

where a is the offset amplitude equal to $|\underline{\mu}_t|$. So from equation 2.19

$$\Phi(\rho) = \int_0^\infty \int_0^{2\pi} \dots \int_0^{2\pi} J_{N-1}(r\rho) \frac{\Gamma(N)}{2 a^{2N-1} \pi^N} \delta(r-a) dr \quad (A.14)$$

or

$$\Phi(\rho) = \left(\frac{2}{a\rho}\right)^{N-1} \Gamma(N) J_{N-1}(a\rho) \quad (A.15)$$

Therefore the equivalent characteristic function from equations A.12 and A.15 is

$$\Phi_{x_t}(\rho) = \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) \left(\frac{2}{a\rho}\right)^{N-1} \Gamma(N) J_{N-1}(a\rho) \quad (A.16)$$

The probability distribution of r_d may be found from equation 2.23 and is

$$P_N(r_d) = \frac{r_d^N}{a^{N-1}} \int_0^\infty \rho \exp\left(-\frac{\sigma^2 \rho^2}{2}\right) J_{N-1}(r_d \rho) J_{N-1}(a \rho) d\rho \quad (\text{A.17})$$

This is what is called Weber's second exponential integral and has the solution (Watson 1962, page 395)

$$P_N(r_d) = \frac{r_d^2}{a^{N-1} \sigma^2} \exp\left(-\frac{a^2 + r_d^2}{2\sigma^2}\right) I_{N-1}\left(\frac{a r_d}{\sigma^2}\right) \quad (\text{A.18})$$

where $I_n(\cdot)$ is the modified Bessel function of order n . For $N = 1$ this reduces to the Rician distribution.

It is not easy to find a neat expression for the cumulative distribution of r_d as the integral becomes a generalized form of Weber's second exponential integral which possesses no closed form solution. A numerical solution may be found, however, using the Fourier-Bessel series of equation 3.7 and appropriate limits for the parameters R and M .

$$P_N(r_d) \approx \frac{2r_d^2}{a^{N-1} R} \sum_{k=1}^M \frac{\exp\left(-\frac{\lambda_k^2}{2R^2}\right)}{\lambda_k^2 J_N^2(\lambda_k)} J_{N-1}\left(\lambda_k \frac{a}{R}\right) J_N\left(\lambda_k \frac{r_d}{R}\right) \quad (\text{A.19})$$

Therefore for the special case where the components of $\underline{x}_t$ are all Gaussian with variance σ^2 and means μ_{iR} and μ_{iI} where appropriate, the result of noncoherent integration is simply an increase in dimension from the single return case.

It should be noted that the distribution of r_d^2 is related to the distribution of r_d by the function $y = x^2$ so

$$p_N(y) = \left(\frac{y}{x} \right)^{\frac{N-1}{2}} \frac{1}{2\sigma^2} \exp\left(-\frac{x+y}{2\sigma^2} \right) I_{N-1}\left(\frac{(xy)^{\frac{1}{2}}}{\sigma^2} \right) \quad (\text{A.20})$$

where $y = r_d^2$ and $x = a^2$. This is the noncentral chi-square distribution with $2N$ degrees of freedom. When $a = 0$ it becomes the chi-square distribution. Therefore equation A.19 may be used to calculate the cumulative noncentral chi-square distribution.

END

1 5 1 2 8 1

FIN