INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeib Road
Ann Arbor, Michigan 48106
NEFF, John Alexander, 1942-
DETECTION OF MAXIMAL LENGTH BINARY TIME
SEQUENCES IN ADDITIVE GAUSSIAN NOISE:
APPLICATION OF SEQUENTIAL LIKELIHOOD RATIO
DETECTION.

The Ohio State University, Ph.D., 1973
Engineering, electrical

University Microfilms, A XEROX Company, Ann Arbor, Michigan
DETECTION OF MAXIMAL LENGTH BINARY TIME SEQUENCES
IN ADDITIVE GAUSSIAN NOISE: APPLICATION OF
SEQUENTIAL LIKELIHOOD RATIO DETECTION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

by


* * * * *

The Ohio State University
1973

Approved by

Robert B. Lackey
Adviser
Department of Electrical Engineering
ACKNOWLEDGMENTS

My foremost appreciation is extended to Dr. Charles P. Hatsell, a former colleague of mine at the Air Force Institute of Technology (AFIT), for his initial encouragement of my research of maximal length sequences and for his many helpful suggestions. I consider it an honor to have been associated with such an individual as Dr. Hatsell with his extensive knowledge of so many fields of science and his outstanding capacity for research.

My appreciation is also given to Captain John R. Green, a student of mine at AFIT, for his assistance in the study of the detection of the length of the maximal length sequences. It was through his interest in radar signal problems that the detection scheme considered in this thesis was formulated. The research problem of how to detect maximal length sequences used in the coding of biphase modulated radar waveforms was presented to Captain Green and myself by personnel of the Electronic Warfare Branch, Air Force Avionics Laboratory, Wright-Patterson AFB, Ohio.

I would like to thank both my adviser, Dr. Robert B. Lackey, for his continued confidence in my ability to
complete my thesis research, and Professor C. E. Warren for his many fine suggestions for revision of the rough draft. I would also like to thank Dr. Robert Fontana and Professor John J. D'Azzo of AFIT and Dr. Carroll N. Day of the Aerospace Medical Research Laboratory for their assistance in adjusting my workload so that I could complete my research work. Thanks also goes to Mrs. Jane Manemann of AFIT for her excellent and diligent work in typing the final manuscript.

Last, but not least, I owe a great deal of gratitude to my wife and son for tolerating the many evenings and weekends that we could not spend together so that I could complete my degree requirements.
VITA

September 5, 1942 . . . Born - Bellaire, Ohio
1964 . . . . . . . B.A., Ohio Wesleyan University, Delaware, Ohio
1964 . . . . . . . Assistant Physicist, Riegal Textile Corporation, Ware Shoals, South Carolina
1965-1968 . . . . Teaching Associate, Department of Mathematics, The Ohio State University, Columbus, Ohio
1967 . . . . . . . M.Sc., The Ohio State University, Columbus, Ohio
1969-1973 . . . . Assistant Professor, Department of Electrical Engineering, Air Force Institute of Technology, Wright-Patterson Air Force Base, Dayton, Ohio
1973 . . . . . . . Bio-Control Engineer, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio

PUBLICATIONS


FIELDS OF STUDY

Major Field: Electrical Engineering

Studies in Circuits and Communication. Professors C. Earl Warren and William C. David
Studies in Control Systems. Professor John Bacon

Studies in Electromagnetism. Professor Albert F. Prebus

# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .......................................................... ii
VITA .................................................................................. iv
LIST OF TABLES ................................................................... vii
LIST OF ILLUSTRATIONS ..................................................... viii
LIST OF SYMBOLS ................................................................ xi

Chapter

I. INTRODUCTION .............................................................. 1
II. MAXIMAL LENGTH BINARY SEQUENCES ......................... 9
III. DETERMINATION OF THE SAMPLING RATE .................. 17
IV. STATISTICAL DETECTION THEORY .............................. 40
V. DEVELOPMENT OF THE OPTIMAL SEQUENCE DETECTOR ........ 49
VI. CLIPPER CROSS CORRELATOR DETECTOR - SUBOPTIMAL SEQUENCE DETECTOR .... 68
VII. PERFORMANCE OF THE OPTIMAL AND SUBOPTIMAL DETECTORS ...................... 79
VIII. SUMMARY AND DISCUSSION OF FUTURE RESEARCH ............. 109

APPENDIX

A. MONTE CARLO METHOD ................................................. 114
B. CORRELATION OF NOISE SAMPLES ............................... 124

BIBLIOGRAPHY ................................................................... 129
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Number of Binary Waveforms for the Various Sampling Rates</td>
<td>18</td>
</tr>
<tr>
<td>2.</td>
<td>Number of Autocorrelation Ratios for the Various Sampling Rates</td>
<td>22</td>
</tr>
<tr>
<td>3.</td>
<td>The Five Likelihood Ratios for the 6-hypotheses 31 Bit SLR Detector Following Each of the First Four Sequence Repetitions in an Average Trial Run for S/N = 1</td>
<td>85</td>
</tr>
<tr>
<td>4.</td>
<td>The Five Likelihood Ratios for the 6-hypotheses 31 Bit SLR Detector Following Each of the First Four Sequence Repetitions in the &quot;Worst&quot; Trial Run Out of 1116 Trials for S/N = 1</td>
<td>86</td>
</tr>
<tr>
<td>5.</td>
<td>Comparison of the Probabilities of Error for the Optimal Detection of a Hard-clipped Signal to That for the Optimal Detection of a Non-clipped Signal for the 15 Bit and 31 Bit Optimal &quot;M&quot; Sequence Detectors</td>
<td>107</td>
</tr>
</tbody>
</table>
# LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Transmitted Signal of a Pulsed Radar</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>Modulating Envelope for Pulsed Radar</td>
<td>1</td>
</tr>
<tr>
<td>3.</td>
<td>Binary Waveform Used to Modulate Envelope of Pulsed Radar</td>
<td>2</td>
</tr>
<tr>
<td>4.</td>
<td>Biphase Modulated Radar Waveform</td>
<td>3</td>
</tr>
<tr>
<td>5.</td>
<td>Block Diagram of Proposed Biphase Modulated Radar System</td>
<td>5</td>
</tr>
<tr>
<td>6.</td>
<td>Typical Output Waveform from Demodulator of Biphase Modulated Radar Receiver</td>
<td>6</td>
</tr>
<tr>
<td>7.</td>
<td>Binary Sequence Generator</td>
<td>11</td>
</tr>
<tr>
<td>8.</td>
<td>The Two 15 Bit &quot;M&quot; Sequences and Their Respective Shift Register Generators</td>
<td>16</td>
</tr>
<tr>
<td>9.</td>
<td>The Noise-free Autocorrelation of the 63 Bit Maximal Length Sequence Number M-147</td>
<td>21</td>
</tr>
<tr>
<td>10.</td>
<td>Ranges of the Average Autocorrelation Ratios Versus the S/N Ratio for the Sequences of Various Lengths Sampled 31 Samples Per Sequence</td>
<td>29</td>
</tr>
<tr>
<td>11.</td>
<td>Same as figure 10 except &quot;... 63 Samples Per Sequence&quot;</td>
<td>30</td>
</tr>
<tr>
<td>12.</td>
<td>Same as figure 10 except &quot;... 127 Samples Per Sequence&quot;</td>
<td>31</td>
</tr>
<tr>
<td>13.</td>
<td>Same as figure 10 except &quot;... 255 Samples Per Sequence&quot;</td>
<td>32</td>
</tr>
<tr>
<td>14.</td>
<td>Same as figure 10 except &quot;... 511 Samples Per Sequence&quot;</td>
<td>33</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>15. Comparison Between R(0) and R(Δ), Where Δ Represents the Time Interval Between Samples, for a Correctly Sampled Sequence and for an Oversampled Sequence</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>16. Basic Signal Detection Problem</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>17. Model of Binary Hypotheses Decision Process</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>18. The Binary Time Waveforms of the 15 Shifted Versions of Sequence M-23</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>19. The Binary Time Waveforms of the 15 Shifted Versions of Sequence M-31</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>20. Optimal Detector</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>21. Usual CCC Configuration</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>22. Modified CCC Configuration</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>23. Optimal Detector for Hard-Clipped Signal</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>24. The 5 Likelihood Ratios of the 31 Bit Detector Versus the Number of Samples Processed by the Receiver for an Average Trial Run Under S/N = 1</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>25. The 5 Likelihood Ratios of the 31 Bit Detector Versus the Number of Samples Processed by the Receiver for the &quot;Worst&quot; Trial Run Out of 1116 Trials Under S/N = 1</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>26. The Classification Probability of the Shifted Version Actually Received and the 2 Out of the Other 30 Classification Probabilities that Converged to Zero the Least Rapidly for an Average Trial Run Under S/N = 1</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>27. The Classification Probability of the Shifted Version Actually Received and the 2 Out of the Other 30 Classification Probabilities that Converged to Zero the Least Rapidly for the &quot;Worst&quot; Trial Out of 1116 Trials Using S/N = 1</td>
<td>94</td>
<td></td>
</tr>
</tbody>
</table>
28. Probability of Error Versus the Number of Samples Processed by the Receiver for the 15 Bit SLR Receiver Detecting an "M" Sequence and Its Shifted Version in Additive Gaussian Noise ............... 97

29. Same as figure 28 except for the 31 Bit SLR Receiver ............... 99

30. Probability of Error Versus the Number of Samples Processed by the 15 Bit SLR Receiver Designed to Detect an "M" Sequence and Its Shifted Version in Additive Gaussian Noise that has been Hard-Clipped to the Binary Levels of the "M" Sequence ............... 103

31. Same as figure 30 except for the 31 Bit SLR Receiver ............... 105

32. The Autocorrelation of 550 of the Random Numbers Used to Simulate Noise Samples in the Simulations of the Sampling Rate Detector and the Sequence Detector ............... 126
LIST OF SYMBOLS

\( \alpha \)  
Decision boundary of hard-clipper

ACR  
Autocorrelation ratio

APP  
A posteriori probability

CCC  
Clipper cross correlator

CDC  
Computer Data Corporation

\( c_m \)  
mth shifted version of maximal length sequence

\( c_{mi}(k) \)  
kth bit of the mth shifted version of the ith hypothesis

\( \delta() \)  
Dirac delta function

\( \Delta \)  
Time interval between time samples

\( E\{ \} \)  
Expected value operator

\( \mathcal{F} \)  
Fourier transform operator

\( f() \)  
Probability density function

\( f(\mid\)  
Conditional probability density function

\( \gamma \)  
Decision threshold

\( H_i \)  
ith statistical hypothesis

\( L \)  
Sequence length (number of bits in binary sequence)

\( L(x) \)  
Likelihood ratio of parameter x

"M"  
Maximal length

\( n(t) \)  
Time waveform of noise signal

\( N \)  
Noise signal vector

\( \xi \)
\begin{align*}
P_i & \quad \text{A priori probability of the } i\text{th hypothesis} \\
P_k(\cdot) & \quad \text{Classification probability following the } k\text{th sample} \\
P_r(\cdot) & \quad \text{Probability operator} \\
r(t) & \quad \text{Time waveform of received signal} \\
R & \quad \text{Received signal vector} \\
R_k & \quad \text{kth sample of received signal} \\
\mathcal{R}_k & \quad \text{The set of samples } \{r_1, r_2, \ldots, r_k\} \text{ of the received signal (a } k\text{ dimensional vector)} \\
R(\cdot) & \quad \text{Autocorrelation function} \\
s(t) & \quad \text{Time waveform of transmitted signal} \\
S & \quad \text{Transmitted signal vector} \\
\mathcal{S}_k & \quad \text{kth sample of transmitted signal} \\
\sigma & \quad \text{Standard deviation of gaussian noise} \\
\text{SLR} & \quad \text{Sequential likelihood ratio} \\
S/N & \quad \text{Signal to noise ratio} \\
\tau_0 & \quad \text{Time width of pulse corresponding to each binary bit} \\
T & \quad \text{Radar pulse width } (T = L\tau_0) \\
\Theta & \quad \text{Multiple hypotheses parameter} \\
w_k & \quad \text{kth sample of hard-clipped received signal} \\
W_k & \quad \text{The set of samples } \{w_1, w_2, \ldots, w_k\} \text{ of the hard-clipped received signal (a } k\text{ dimensional vector)} \\
y_k & \quad \text{kth output from correlator} \\
Y_k & \quad \text{The } k\text{ dimensional vector } (y_1, y_2, \ldots, y_k) \\
Z_i & \quad \text{Decision region for which inclusion of } R \text{ yields the decision that } H_i \text{ is true}
\end{align*}
CHAPTER I

INTRODUCTION

This thesis considers the problem of receiver determination of the binary sequence used as the coding scheme in communication systems employing biphase coded waveforms. This detection problem was encountered in the area of radar system design, but applications are envisioned with respect to the design of adaptive receivers.

The concept of biphase modulation and binary coding of radar waveforms may be better understood by referencing figures 1 through 4. Figure 1 illustrates the transmitted signal of a typical pulsed radar, while figure 2 shows the

![Figure 1. Transmitted Signal of a Pulsed Radar.](attachment:figure1.png)

![Figure 2. Modulating Envelope for Pulsed Radar.](attachment:figure2.png)
modulation envelope of the carrier waveform where $T$ is known as the radar pulse width. For radar applications in environments with extreme background noise, signal detection may be aided by modulating the envelope with a binary waveform prior to using this new envelope to modulate the carrier waveform. This binary modulation process may be explained by considering the transformation of a binary sequence of $L$ bits to a binary time waveform. The time waveform will be a sequence of pulses of equal duration $1/L$ and of magnitudes sequenced according to the sequence of binary bits as shown in figure 3 for the sequence $\{1,-1,-1,1,-1,1,1\}$. The resulting transmitted waveform will then be as shown in figure 4 where a change in the binary signal level corresponds to a 180 degree phase shift in the carrier waveform. A class of binary sequences known as maximal length sequences are used as the binary coding scheme because their corresponding binary waveforms have correlation properties which allow better resolution capabilities than those waveforms.

![Figure 3. Binary Waveform Used to Modulate Envelope of Pulsed Radar.](image-url)
The objective of this thesis is to design a receiver that will determine the maximal length sequence used to biphase modulate the carrier waveform prior to transmission given that the transmitted signal is disturbed by additive Gaussian noise and given that the receiver has knowledge of the time duration of the received signal and either the time of the leading edge or the time of the trailing edge of the received signal. Receiver determination of these time parameters is thoroughly treated in literature dealing with radar technology; however, this limited knowledge of the received signal still leaves the necessity of determining the number of bits used in the coding scheme. This knowledge of the number of bits, herein referred to as the sequence length and denoted by $L$, is necessary for the determination of the sampling rate that will be used to sample the received waveform prior to the detection of
the maximal length sequences.

The thesis, then, covers the solution to two different detection problems. The first problem, the determination of sequence length, will be referred to as the sampling rate problem. Knowledge of the number of binary bits, \( L \), in conjunction with knowledge of the signal duration, \( T \), gives knowledge of the maximum number of pulses possible in the binary modulating waveform (resulting from a binary sequence containing no two consecutive bits of equal magnitude). In order that the binary waveform might be able to be reconstructed from its samples, the sampling rate must be at least \( L/T \). It was this minimum sampling rate that was used to design the maximal length sequence detector; thus, the solution to this first problem, the determination of \( L \), must be preliminary to the determination of the coding sequence.

The second problem, the design of the maximal length sequence detector, will be referred to as the sequence detection problem. As stated before, this detector processes samples of the received waveform and determines which maximal length sequence was used by the transmitter modulator. Figure 5 illustrates the location of the sampling rate detector and the sequence detector in the biphase coded radar system. It should be noted that figure 5 is a block diagram of those portions of the system of interest to the two problems to be discussed in this thesis and does
Figure S. Block Diagram of Proposed Biphase Modulated Radar System.

not include other signal processing components such as amplifiers and wide-band filters. A typical waveform, \( x(t) \), that will be processed by both the sampling rate detector and the sequence detector (plus sampler) is shown in figure 6.

Both problems were simulated on the computer (CDC 6600) and all error analyses are the results of Monte Carlo techniques applied to the simulated detectors. The solution to the sampling rate problem was formulated through the study and comparison of simulation results
while the solution to the sequence detection problem was analytically determined by the application of optimization techniques and was later simulated for the purpose of performing an error analysis.

Chapter II gives a brief explanation of maximal length sequences while Chapter III discusses the sampling rate problem and its solution. The development of the sampling rate detector closely follows the discussion of Chapter II since it is largely based on special properties of the sequences. The thesis is written in such a manner that one may begin with Chapter IV if the interest is only in the optimal detection of binary sequences given that the receiver has knowledge of the sequence length, L.

Chapter IV gives a brief introduction to statistical detection theory including the development of the three types of statistical hypotheses testing employed in the remainder of the thesis: binary, multiple, and composite.
Chapter V presents the development of the receiver which detects the transmitted binary sequence embedded in a gaussian noise environment with the lowest probability of error (optimal detector). The detector is developed in its sequential form meaning that it optimizes detection following each received sample as opposed to optimizing detection only after reception of all samples of the sequence. This is due to the fact that the probability of error in identifying the transmitted sequence is often an acceptable value following the processing of only a few samples of the received waveform by the sequential detector, thereby eliminating the need to process any remaining samples.

Chapter VI presents the development of a detector which operates with a higher probability of error than the optimum detector but which is a great deal more simplified with respect to implementation. This suboptimum receiver is not sequential in nature, but it is shown to be the optimum detector in the case where the received signal is hard-clipped prior to being detected.

Chapter VII contains graphs depicting the operation of the optimal detector along with a Monte Carlo error analysis for both the optimal and suboptimal detectors. The large amount of computer time required by the Monte Carlo technique limited the error analysis to the shorter maximal length sequences.
Chapter VIII gives a brief summary of conclusions along with recommendations for further study.
CHAPTER II

MAXIMAL LENGTH BINARY SEQUENCES

In the early 1950's, several researchers, such as S. W. Golomb and N. Zierler, became interested in the properties of a special class of binary sequences and placed special emphasis on the autocorrelation property of these sequences. Siebert (Ref. 15), in 1956, combined the knowledge gained by these researchers\(^1\) with the classic radar works of P. M. Woodward and I. L. Davies (Refs. 19, 20, and 21) in his discussion of biphase modulation coding of radar waveforms as a means of coming as close as possible to the ideal radar waveform from the standpoint of accuracy and freedom from ambiguities. That is, the plot of the noise-free cross correlation of the transmitted and received waveforms considered as a function of the time delay \(\tau\) and frequency shift \(\omega\), denoted as \(\psi(\tau,\omega)\), and labeled the ambiguity function, approximates a single central spike of width \(2\pi/T\) in frequency and \(2\pi/W\) in time.

\(^1\)Siebert's paper was based on the unpublished work of Zierler; however, this work was later published in 1959 (Ref. 22). Zierler credits Golomb with simultaneous work in the field but Golomb did not publish his work on binary sequences until 1964 (Ref. 6).
(where T and W are the echo duration and bandwidth in radians/second, respectively) with the remainder of the plot spread out more or less uniformly over a region roughly T seconds wide in time and W radians/second wide in frequency.

Siebert denoted these waveforms as pseudo-random due to the necessity of specifying them by a large set of numbers as opposed to those radar waveforms specified by only a few parameters such as pulse length, repetition rate, etc. Ristenbatt (Ref. 13) describes the "randomness characteristics" that pseudo-random binary sequences possess by comparing them with binary sequences resulting from a coin flipping experiment.

A subset of these pseudo-random sequences, known as maximal length or "M" sequences, produce the best performance when used as the radar coding sequence. That is, for certain given values of the time bandwidth product $\frac{TW}{2\pi}$, namely $\frac{TW}{2\pi} = 2^n - 1$ where n is any positive integer, the ambiguity function $\psi(\tau, \omega)$ is minimized for $\tau \neq 0$ and $\omega \neq 0$. An example of such a coding sequence for $\frac{TW}{2\pi} = 31$ is the following:

1111100110100100001010111011000

Siebert further noted that coding sequences exist which have better performance along the $\tau$ axis than "M" sequences but that such non-maximal length sequences
inevitably seem to have large spikes (i.e., potential ambiguities) for \( \omega \neq 0 \).

Pseudo-random sequences have also been called shift-register sequences in the literature due to the ability of the binary logic device known as a shift register to generate any one of the sequences provided the register is supplied with a given feedback network. This generation procedure will now be explained since it was used to generate the sequences in the digital simulations and since it yields a logical explanation to the use of the terminology "maximal length".

Consider the digital system shown in figure 7 where the series of binary flip-flops is known as a shift register. The symbol \( \oplus \) denotes an "exclusive-or" network which is nothing more than a modulo 2 adder; i.e., a "carry-free" binary adder. If the initial state of the shift register is taken as 111 (all flip-flops in the same state), then the input to the register will be \( 1 \oplus 1 = 0 \); thus, the next clock pulse will produce 011 as the shift

![Figure 7. Binary Sequence Generator.](image-url)
register state. A continuing analysis will show that, following the sixth clock pulse, the output will be seen to have sequenced in the following manner: 1110010 and the shift register state will once again be 111.

Since the sequence will begin a repeating cycle once the shift register returns to any state previously held, the longest sequence will be one resulting from the register cycling through all possible states. A register composed of n flip flops has $2^n$ possible states; however, the state 000 must be excluded since this state can produce nothing but an all zero sequence. The remaining $2^n - 1$ states may produce a sequence of length $2^n - 1$ bits, and any such sequence will be a maximal length sequence.

The sequence 1110010 resulting from the digital network in figure 7 is maximal length since the shift register sequenced through $2^3 - 1 = 7$ different states (111,011,001, 100,010,101, and 110 representing the binary numbers 7,3,1,4,2,5, and 6, respectively). If the initial state of the register would have been any of the other states besides 111, the resulting sequence would be considered to be the same maximal length sequence--the latter being denoted as a shifted version of the former. That is, a given feedback network will produce only one maximal length sequence but may produce any shifted version of that sequence depending on the initial state of the register.
Not all combinations of flip flop outputs summed modulo 2 and fed back to flip flop #1 will produce maximal length sequences. This may easily be illustrated by noting the sequence of register states resulting from summing modulo 2 the outputs of all three flip flops. If 111 is the initial state, the succeeding states will be 011,101,110,011,101,..., and the resulting non-maximal length sequence will be: 111011011011... containing a three bit repeating cycle. The initial register state used in the generation of maximal length sequences is immaterial with respect to the specific "M" sequence generated, but for the non-maximal length sequences, different initial states may produce different sequences. For example, the initial state of 001 in the above example will produce the sequence 10000... while the initial state of 100 will produce 0011011011... which is not a shifted version of the previous sequence.

Shift register generation of the "M" sequences requires knowledge of which register stages to feed back, and the simplest method of acquiring this knowledge is through the use of primitive polynomials. A polynomial of degree n with the binary coefficients 0 and 1 is primitive if it is not factorable into polynomials of degree less than n (irreducible), is not a factor of the polynomial $x^m - 1$ for any m less than $2^n - 1$, and has all its roots of order $2^n - 1$. Every primitive polynomial yields the
information required to uniquely determine the required feedback to produce one and only one "M" sequence. A very extensive table of both irreducible and primitive polynomials is given by Peterson (Ref. 11), and the connection between these polynomials and the "M" sequences will be explained since the explanation will yield an identification number for each sequence that is convenient in making a distinction between the various sequences.

Consider the set of polynomials having coefficients of only a 0 or 1. Any nth degree polynomial of this set may be uniquely specified by an (n + 1)-tuple of binary numbers consisting of the n + 1 coefficients. The degree of each term with a unity coefficient represents a stage of the shift register that must be fed back to the modulo two adder. For example, the primitive polynomial

\[ x^{10} + x^8 + x^7 + x^6 + x^3 + 1 \]

is represented by (10111001001) and indicates that the outputs of stages 10, 8, 7, 6, and 3 must be modulo two added and fed back as the input to stage #1. This 11-tuple may be considered as an 11 digit binary number which may, in turn, be written more compactly in its octal representation as shown below.

10 111 001 001
2 7 1 1
Each octal number is the decimal representation of each three digit binary sequence starting from the right. The "M" sequence generated by a ten stage shift register with the above listed stages fed back may be referred to as sequence no. 2711.

The reciprocal polynomial of a primitive polynomial is also primitive so Peterson's table lists one polynomial for every two maximal length sequence generators. For example, the feedback connections for all 15 bit maximal length sequence generators may be found as follows: The polynomial degree, n, is determined by the equation $2^n - 1 = 15 \implies n = 4$. Peterson's table lists only 23 under fourth degree polynomials so the two primitive 4th degree polynomials are:

$$(23) \implies (10011) \implies x^4 + x^1 + 1$$

and its reciprocal:

$$(11001) \implies x^4 + x^3 + 1$$

which is polynomial number 31. These generators are shown schematically in figure 8 along with their respective "M" sequences, M - 23 and M - 31 written assuming that the initial register states were all 1's.

Up to this point, a binary sequence has been listed as a sequence of 0's and 1's because of convenience. However, a numerical assignment of binary levels of unequal
power (i.e., 0 and 1 as opposed to -1 and +1) is often a disadvantage in signal processing. For example, a correlation of binary signals of non constant power produces a result which places more emphasis on the correlation or similarity existing between the larger magnitude components than the smaller magnitude components of the signals. Since correlation processing techniques are considered in this study, the "M" sequences will be considered as sequences of +1's and -1's rather than 0's and 1's.
CHAPTER III

DETERMINATION OF THE SAMPLING RATE

The receiver which optimally detects a given maximal length time sequence must have prior knowledge of the sequence length \( L \) which is defined to be the number of bits in the coding sequence. If the time sequence is of duration \( T \) seconds, then the bit rate is \( L/T \) bits per second, and this represents the minimum number of time samples per second required in order that the binary time waveform may be uniquely determined from the samples. If it is assumed that the sequence is transmitted throughout one unit of time, \( L \) may be denoted as the minimum sampling rate for maximal length sequences of length \( L \), and it was this sampling rate that was used in the simulation of the sequence detector. Thus, the determination of the sequence length \( L \) is synonymous with the determination of the sampling rate to be used by the sampler shown in figure 5.

The determination of the sampling rate is not only complicated by the existence of more than one maximal length sequence of a given length but by the fact that there are \( L \) shifted versions of each \( L \) bit "M" sequence. Table 1 lists the total number of different binary
waveforms that are possible. Preliminary investigations were aimed at techniques that would reduce the number of different alternatives to be considered. Various linear transformations of the "M" sequences were attempted such as the Fourier and Walsh transforms, but the most successful technique found was the classification of the maximal length time sequences of a given length into groups which had, upon being sampled at the minimum rate as discussed above, the same ratio of the center peak of their autocorrelation functions to the maximum non-center value (Ref. 9). A more precise definition of this ratio may be given by considering an analytical expression for the autocorrelation function of a binary time function represented by its time samples (1 sample per bit). If $s(t)$ represents the binary signal, then the signal may be expressed in terms of its samples, $s_k$, as follows:
\[ s(t) = \sum_{k=1}^{L} s_k f_k(t) \]  
\[ f_k(t) = f(t - k\tau_0) = \begin{cases} 
1 & (k - 1)\tau_0 < t < k\tau_0 \\
0 & \text{otherwise}
\end{cases} \]

such that \( \tau_0 \) represents the width of the binary pulse corresponding to a single bit of the sequence of length \( L \) (see figure 3). Substituting equation (1) into the autocorrelation function

\[ R(t') = \int_{-\infty}^{+\infty} s(t)s(t + t')dt \]

and letting \( t' = k\tau_0 + \tau \) where \( 0 < \tau < \tau_0 \), yields:

\[ R(k\tau_0 + \tau) = \int_{-\infty}^{+\infty} \left[ \sum_{i=1}^{L} s_i f_i(t) \right] \left[ \sum_{i=1}^{L} s_i f_i(t + k\tau_0 + \tau) \right] dt \]

Equation (4) may be reduced to the following equation which represents a function consisting of straight-line segments whose slope may change at the end of every \( \tau_0 \) interval (Ref. 16, p. 178).

\[ R(k\tau_0 + \tau) = (\tau_0 - \tau) \sum_{i=1}^{L-k} s_i s_{i+k} + \tau \sum_{i=1}^{L-k-1} s_i s_{i+k+1} \]

Since \( R(k\tau_0 + \tau) \) consists only of straight line segments, it is uniquely specified by the collection of points
\{R(\tau_0), R(2\tau_0), \ldots, R[(L-1)\tau_0]\}. Thus, the computation of the autocorrelation function need only involve the summation:

$$R(k\tau_0) = \tau_0 \sum_{i=1}^{L-k} s_i s_{i+k}$$  \hspace{1cm} (6)

The autocorrelation ratio discussed above may now be written as:

$$ACR = \frac{\sum_{i=1}^{L} s_i^2}{\max\left\{\sum_{i=1}^{L-k} s_i s_{i+k}\right\}} \hspace{1cm} (7)$$

Figure 9 shows the autocorrelation function (for \(\tau_0 = 1\)) of a 63 bit "M" sequence (sequence no. 147 - see Chapter II) for which the ACR equals 9.00. As stated above, the use of this ratio permits the classification of the sequences into a relatively small number of groups for any given sequence length. Table 2 lists the number of different ratios for some of the sequences, and a comparison of Table 1 and Table 2 shows the reduction in the number of alternatives that need to be considered.

The autocorrelation ratios may now be investigated to determine if their relative magnitudes are different for the case of sequences sampled exactly one sample per bit (herein referred to as a correctly sampled sequence) as
Figure 9. The noise-free autocorrelation of the 63 bit maximal length sequence number M-147.
Table 2. Number of Autocorrelation Ratios for the Various Sampling Rates.

<table>
<thead>
<tr>
<th>L</th>
<th>Number of Autocorrelation Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>63</td>
<td>6</td>
</tr>
<tr>
<td>127</td>
<td>9</td>
</tr>
<tr>
<td>255</td>
<td>10</td>
</tr>
</tbody>
</table>

opposed to sampling either at a rate greater than one sample per bit (herein referred to as oversampling) or at a rate less than one sample per bit (undersampling)\(^1\). That is, can the autocorrelation ratio be considered as a sufficient statistic for testing a set of hypotheses representing the various sequence lengths? If the ACR is a valid sufficient statistic, then the ACR of a sequence of unknown length sampled at \(L_1\) evenly spaced intervals within the pulse duration \(T\) of the signal will either be one of those ratios representing a correctly sampled sequence or be different from these values indicating that the sequence of unknown length was not of length \(L_1\).

If a detector employing the ACR as a sufficient

---

\(^1\) All sampling rates will either be correct (sampling rate equal to sequence length divided by the radar pulse width \(T\)) or incorrect by a factor of approximately a power of 2 since the receiver would only use trial sequence lengths of \(2^n - 1\) to determine the sampling rates in an attempt to correctly sample the sequence.
statistic is to work for sequences embedded in additive noise, the ratios\(^2\) of correctly sampled sequences and those of incorrectly sampled sequences should be significantly different. This was found to be true provided the receiver could eliminate the possibility of undersampling the sequence. The following statement (herein denoted as the basic hypothesis of the sampling rate detector) summarizes the assumption that was tested by the computer simulation of the sampling rate detector designed to operate in an additive Gaussian noise environment for which the noise samples were uncorrelated:

A maximal length time sequence of time duration \(T\) and of length \(L_1\) sampled at \(L_1\) evenly spaced time intervals throughout the time span has an autocorrelation ratio that is significantly larger than those ratios of maximal length sequences of time duration \(T\) but of lengths less than \(L_1\) sampled at the same \(L_1\) intervals (oversampling).

\(^2\)The autocorrelation function is no longer given by equation (5) since the received signal is a stochastic process rather than a square waveform (see figure 6). Throughout the remainder of this chapter, the autocorrelation ratio will be defined by equation (7) where the \(s_k\) terms represent the samples of the received signal. This would be the ACR of the output waveform from a sample and hold device. The autocorrelation of the samples of a signal will be defined to be the discrete set of numbers given by equation (6).
The objective then is to subdivide the set of real numbers representing the range of autocorrelation ratios for all "M" sequences (i.e., all sequences of all lengths) sampled \( L_1 \) times within a given time span \( T \) (where \( L_1 = 2^n - 1 \) is chosen as the initial trial sampling rate) into two decision regions. One region or subset would consist of those ratios such that, if the computed ACR of the received sequence were to be in this subset, the decision would be that the sequence was correctly sampled and its length was indeed \( L_1 \). The other region or subset of real numbers would consist of those ratios such that, if the computed ACR of the received sequence were to be in this subset, the decision would be that the sequence was oversampled and another sampling rate \( L_2 < L_1 \) must be tried. In summary, the receiver samples the received signal at some rate \( L_1/T \), autocorrelates the samples, computes the autocorrelation ratio, determines which decision region the ratio is an element of, and decides either that the received sequence is of length \( L_1 \) or not of length \( L_1 \).

There are two restricting facets of the above basic hypothesis that should be noted. First, the number of subsets or decision regions of autocorrelation ratios is limited to two rather than being set equal to the number of different sequence lengths to be investigated. The latter represents the more ideal situation since there would then exist one decision region corresponding to each
length; therefore, the receiver output would be the sequence length rather than the $L_1$-or-not-$L_1$ decision. The not-$L_1$ output must indicate to the receiver to change the sampling rate and repeat the decision process. However, the number of useable sampling rates is not that large since the use of extremely long sequences is discouraged by their requirement for prohibitively long processing times at the receiver. If $L_L$ represents the longest transmitted sequence and $L_S$ represents the shortest, the greatest number of trials required would be the number of elements in the set \{n:n is an integer and $\log_2 (L_S + 1) \leq n \leq \log_2 (L_L + 1) - 1$\} where the -1 term results from the use of a process of elimination. A parallel type processor could be used to investigate all sampling rates simultaneously.

The second restriction is the receiver's capability of making a decision only when the sampling rate is greater than or equal to the sequence length. This is restrictive only in that it would require the series type processor to investigate the next highest sampling rate upon obtaining an autocorrelation ratio above the threshold where the threshold is defined to be the boundary between the two decision regions. This is due to the possibility that the region above the threshold may contain enough autocorrelation ratios of undersampled sequences such that the
The probability of making an error in the decision would be intolerable.

The basic hypothesis was tested by computer simulation of a receiver designed to sample the "M" sequence, to autocorrelate these samples, and to determine the ACR. Prior to the digital correlation, a random number, obtained from a set of uncorrelated random numbers with a zero-mean normal probability distribution with a variable variance (denoted by $\sigma_n^2$), was added to each sample for the purpose of simulating additive zero-mean Gaussian noise in the transmission channel. The variable $\sigma_n^2$ was used to vary the signal to noise ratio, $S/N$, which was defined as follows:

$$S/N = \frac{E[s^2(t)]}{E[n^2(t)]}$$  \hspace{1cm} (8)

where $E\{\cdot\}$ is the expected value operator of probability theory, $s(t)$ is the signal process, and $n(t)$ is the noise process. For a binary waveform composed of +1 and -1

---

3 A set of normally distributed zero-mean random numbers with unity variance was obtained from the uniform random number generator of the CDC 6600 by adding 12 random numbers uniformly distributed between 0 and 1 and subtracting 6. The variance of $\sigma_n^2$ was obtained by multiplying each normal random number obtained in the above manner by the standard deviation, $\sigma_n$. A discussion of the correlation of these random numbers may be found in Appendix B.
signal levels,

\[ E\{s^2(t)\} = 1 \]  \hspace{1cm} (9)

and for a stationary Gaussian noise process of zero mean,

\[ E\{n^2(t)\} = R_{nn}(0) = \sigma_n^2 \]  \hspace{1cm} (10)

Substituting equations (9) and (10) into equation (8) gives:

\[ \frac{S}{N} = \frac{1}{\sigma_n^2} \]  \hspace{1cm} (11)

Thus, the signal to noise ratio is dependent only on the variance of the noise samples.

For a given signal to noise ratio and a given sampling rate, the autocorrelation ratios were plotted as a function of the sequence length for hundreds of different sets of noise samples. The separation of the ratios into the desired groups or decision subsets was noticeable for signal to noise ratios as low as 1/2 for a sampling rate of 31 samples per sequence and, in general, increased not only for an increasing signal to noise ratio but also for an increase in the sampling rate.

These plots are not being presented in this thesis due to their complexity. Instead, the procedure to be described was used to arrive at representative autocorrelation ratios for various signal to noise ratios, and the points representing these ratios were joined by a smooth
curve in order to produce a more readable representation of the results as a function of the signal to noise ratio. A total of 10 trials were run for each sequence (including each shifted version) at each selected signal to noise ratio, and the resulting ratios for the 10 trials were averaged. For example, this produced a set of 378 average ratios for each signal to noise ratio for the 63 bit sequences (see table 1). For a given sampling rate, the autocorrelation ratios were plotted as a function of the signal to noise ratio by using two curves for each sequence length—one curve representing the largest average ratio found for each signal to noise ratio and the other representing the smallest average ratio found. The resulting plots are shown in figures 10, 11, 12, 13, and 14 for sampling rates of 31, 63, 127, 255, and 511 samples per sequence, respectively. The range of autocorrelation ratios between the two curves for each sequence length are typical of the ratios to be expected for the given signal to noise ratio, the given sequence length, and the given sampling rate. For example, from figure 10 it may be noted that, if the sampling rate is 31 samples per sequence (sampling rate equal to 31/T), typical ratios would be in the range of 4 to 6 if the transmitted sequence was of length 31 bits and the signal to noise ratio was 2 but would only be in the range of 2.8 to 3.8 if the transmitted sequence was of length 15 bits for the same signal to noise
Figure 10. Ranges of the average autocorrelation ratios versus the S/N ratio for the sequences of various lengths sampled 31 samples per sequence (solid lines), and an optimum threshold (dashed line).
Figure 11. Ranges of the average autocorrelation ratios versus the S/N ratio for the sequences of various lengths sampled 63 samples per sequence (solid lines), and an optimum threshold (dashed line).
Figure 12. Ranges of the average autocorrelation ratios versus the S/N ratio for the sequences of various lengths sampled 127 samples per sequence.
Figure 13. Ranges of the average autocorrelation ratios versus the S/N ratio for the sequences of various lengths sampled 255 samples per sequence.
Figure 14. Ranges of the average autocorrelation ratios versus the S/N ratio for the sequences of various lengths sampled 511 samples per sequence.
ratio. The discrete set of signal to noise ratios used to plot all the curves in figures 10 through 14 was \( \{0.50, 0.67, 0.85, 1.00, 1.20, 1.43, 1.67, 2.00, 5.00, 10.00, 15.00, 20.00\} \). Note that the signal to noise ratios are plotted on a logarithmic scale.

A comparison of figures 10 through 14 shows that the separation between the curve representing the largest average ratio for oversampled sequences and the curve representing the smallest average ratio for correctly sampled sequences increases as the sampling rate increases. It was deemed impractical to use the large amount of computer time needed to study the cases involving more than 511 samples; thus, the assumption made was that the separation would increase for an increasing sampling rate.

It should be noted that the results shown in figures 10 through 14 were obtained using uncorrelated noise samples (see Appendix B). The existence of any correlation would be expected to result in a decrease in the ACR and a resultant decline in the performance of the detector.

A study of figure 14 (511 samples per sequence) would suggest a threshold for the autocorrelation ratios in the range between 7 and 8. For the cases of lower sampling rates, figures 10 through 13 suggest the possible use of thresholds dependent on the signal to noise ratio. The Monte Carlo technique (see Appendix A) was used in a heuristic investigation into the shape of the threshold.
that would minimize the probability of error; that is, minimize the probability that the receiver will decide either that a given sequence was transmitted when it was not or that a given sequence was not transmitted when it was. This study was done only for the cases of 31 and 63 samples per sequence due to the prohibitive amount of computer time necessary to perform repeated correlations of long sequences. The dashed line on figures 10 and 11 shows the resultant optimum threshold.

A further study of figures 10 through 14 will show that the first restriction implied by the basic hypothesis of the sampling rate detector (restriction to just two decision regions), as discussed earlier in this chapter, is not a mandatory type of restriction. Figures 11, 12, 13, and 14 show some separation existing for all sequence lengths; however, a receiver employing such a multi-threshold scheme would certainly exhibit a much higher probability of error. The second restriction discussed (restriction of not considering undersampling) is not mandatory at high signal to noise ratios since the two curves for each undersampled sequence do separate from the two curves for the correctly sampled sequences for the high S/N ratios. The curves for the undersampled sequences were not shown in the figures for the purpose of clarity. A receiver employing undersampling is not recommended since the slight decrease in both receiver complexity and the
elapsed time to a final decision does not justify this receiver which must operate at a much higher probability of error.

This study was based on an aperiodic autocorrelation processor in that the signal being sampled and correlated was not periodic but was time limited. To the casual observer, a periodic autocorrelation processor (obtained by recycling the received signal through the correlator) appears more advantageous in that the number of different autocorrelation functions reduces by a factor of \( L \). This results from all shifted versions of a given sequence now having identical autocorrelation functions; however, this advantage is more than offset by an increase in the denominator of the ACR caused by an increase in the similarity between a periodic waveform and its shifted version as opposed to the similarity between an aperiodic waveform and its shifted version. This increase in similarity may be demonstrated for binary waveforms by expressing the autocorrelation as a function of the sign or polarity agreements between the samples \( s_i \) of the binary waveform \( s(t) \) and the samples \( s_{i+k} \) of its shifted version \( s(t + \tau) \). If the samples are either +1 or -1, the periodic autocorrelation may be uniquely specified by the set of numbers:
The aperiodic autocorrelation may be uniquely specified by this equation (i.e., equation (12) will be equivalent to equation (6)) if \( s_{i+k} = 0 \) for \( i + k > L \). This decreases the possibility that \( R(k\tau_0) \) will have as large a maximum value for \( k \neq 0 \) as for the periodic autocorrelation. Thus, there exists a greater possibility for similarity (or polarity agreements) for periodic autocorrelation than for aperiodic autocorrelation. This causes a decrease in the ACR for the periodic case which degrades its usefulness as a sufficient statistic.

The above discussion of waveform similarity is important in that it helps to explain the fundamental reason for the truth of the basic hypothesis of the sampling rate detector. If a sequence is oversampled by a factor of about 2, for example, then there is redundancy introduced into the digital correlation and this breeds similarity. The greater the degree of oversampling that exists, the greater the degree of similarity that will exist. Figure 15 illustrates the increase in similarity that exists between the samples of \( x(t) \) and \( x(t + \Delta) \), where \( \Delta \) represents the sampling interval, when one compares a 7 bit "M" sequence sampled at the rate of 14/T samples per second to
Figure 15. (a) Seven bit "M" sequence being sampled. 
(b) Comparison of the polarities of samples used in the correlation of \( x(t) \) and \( x(t + \Delta) \), where \( \Delta \) represents the time interval between samples, for the case of the sequence being sampled correctly. 
(c) Comparison of the polarities for the case of the sequence being oversampled.
the same sequence sampled at the rate of \( 7/T \) samples per second. The measure of similarity expressed by the autocorrelation function of binary samples is the relative number of polarity agreements between the two sets of samples being correlated with respect to the total number of samples (equation (12)). This ratio is 3/7 for the correctly sampled sequence and 5/7 for the oversampled sequence as may be seen by observing the pulse polarity agreements denoted by the "+" signs in figures 15(b) and 15(c).

In summary, this chapter has presented a useful technique for the determination of the length of the maximal length sequence used to encode a biphase modulated waveform given that the time of occurrence and the time duration of the received signal are known. The detection method proposed is certainly not the only one capable of producing satisfactory detectability. A study (Ref. 1) has been done on the frequency domain characteristics of "M" sequences which has produced results that appear very promising. It has shown an interesting null effect in the frequency spectrum that indicates the sequence length independent of the sampling rate used.
CHAPTER IV

STATISTICAL DETECTION THEORY

The basic problem of classical detection theory is illustrated in block diagram form in figure 16. The objective of statistical detection theory is to design a receiver that will make the "best" choice of the samples of signal $s(t)$ given the samples of signal $r(t) = s(t) + n(t)$ where $n(t)$ is a stochastic process denoted as noise. The meaning of "best" will be explained in the following development.

![Block diagram of signal detection](image)

Figure 16. Basic Signal Detection Problem.

Throughout this paper, capital letters will represent a set of time samples of a signal while the small letters will represent the samples themselves. For example, $S = (s_1, s_2, ..., s_k)$ may be considered a vector in a
vector space known as signal space and whose components are the time samples of the signal \( s(t) \). That is, \( s_1 = s(t_1), s_2 = s(t_2), \) etc. where \( t_1 < t_2 \).

Signal detection theory is based on the statistical theory of hypothesis testing. Three important types of problems that are used in the thesis will be developed: binary hypotheses testing, multiple hypotheses testing, and composite hypotheses testing.

For the binary hypotheses problem, there are only two hypotheses, \( H_1 \) and \( H_2 \), corresponding to just two different transmitted signal vectors, \( S_1 \) and \( S_2 \). That is, \( H_1 \) is the hypothesis "\( S_1 \) was sent", and the notation used will be as follows:

\[
\begin{align*}
H_1: & \quad R = S_1 + N \\
H_2: & \quad R = S_2 + N
\end{align*}
\]

where \( R \) denotes a vector in a vector space denoted as "observation space". The basic radar problem dealing with the detection of the presence of a "return" signal is formulated in the above notation by setting \( S_2 = 0 \).

The important step in developing an optimal detector is to determine the decision function. This is the function that maps the observation space into a space known as decision space containing elements which represent the choices available to the receiver. This space is divided into decision regions, and each region corresponds to a
hypothesis in such a manner that, if the received signal is mapped into the decision region corresponding to hypothesis $H_1$, then the detector will decide that $H_1$ is true. The optimal detector works on a decision function which yields the "best possible" decision space; that is, a space whose divisibility into decision regions is the best on the basis of some designated criterion. The mapping of the decision function must be made independently of the transmitted signal that produced a given received signal. Figure 17 illustrates the vector spaces and functional mappings discussed above for the binary hypotheses problem.

![Figure 17. Model of Binary Hypotheses Decision Process.](image)

The optimization criteria used in this thesis will be the minimization of the probability of error in the detector's selection of a hypothesis. The determination of the optimum decision function will be based on the assumptions that the probabilities of the transmitted signals (a priori probabilities) are known and that the noise probability
distribution is known. If \( P_1 \) denotes the a priori probability of \( H_1 \) being true and \( f(R|H_1) \) denotes the probability density function of the received signal vector given that \( H_1 \) is true (derived directly from noise density function), then the probability of error may be written in terms of known expressions. Consider

\[
Pr(\text{error}) = Pr(\text{error}|H_1 \text{ is true}) \cdot Pr(H_1 \text{ is true}) + Pr(\text{error}|H_2 \text{ is true}) \cdot Pr(H_2 \text{ is true}) \quad (14)
\]

Let \( Z_1 \) represent that region of decision space corresponding to the detector's decision that \( H_1 \) is true. Then:

\[
Pr(\text{error}) = \int_{Z_2} f(R|H_1) dR \cdot P_1 + \int_{Z_1} f(R|H_2) dR \cdot P_2 \quad (15)
\]

Since the integral over \( Z_2 \) may be written as unity minus the integral over \( Z_1 \), \( Pr(\text{error}) \) may be written as:

\[
Pr(\text{error}) = P_1 + \int_{Z_1} [f(R|H_2) \cdot P_2 - f(R|H_1) \cdot P_1] dR \quad (16)
\]

\( Pr(\text{error}) \) will be minimum if \( Z_1 \) consists of all those vectors \( R \) for which the integrand contributes negatively to the sum; i.e., \( Z_1 \) must be the set of \( R \)s for which \( g(R) = f(R|H_2) \cdot P_2 - f(R|H_1) \cdot P_1 \) is negative.

\[
\min[Pr(\text{error})] \iff Z_1 = \{R|f(R|H_2) \cdot P_2 - f(R|H_1) \cdot P_1 < 0\} \quad (17)
\]
The inequality expression in implication (17) may be written as:

\[ \frac{f(R|H_2)}{f(R|H_1)} < \frac{P_1}{P_2} \]  

(18)

\( Z_2 \) will consist of all vectors \( R \) for which the inequality sign is reversed.

The resulting decision function is:

\[ L(R) = \frac{f(R|H_2)}{f(R|H_1)} \]  

(19)

and is denoted as the likelihood ratio. \( L(R) \) represents a nonlinear transformation of the observation space in such a manner as to optimize signal detectability. It has been shown (Ref. 2) that the receiver which bases its decision on the likelihood ratio yields optimum performance for a wide class of criteria including the minimization of the probability of error. The location of the optimum boundary in decision space is given by:

\[ L(R) = \gamma = \frac{P_1}{P_2} \]  

(20)

where \( \gamma \) is denoted as the threshold. The optimum detector will operate according to the following set of inequalities:
If \( L(R) > \gamma \) choose \( H_2 \)
If \( L(R) < \gamma \) choose \( H_1 \) \hspace{1cm} (21)

If \( g(x) \) represents any monotonic function of \( x \), then the operation of the detector may be written as:

\[
\begin{align*}
H_2 & & \\
g[L(R)] & \geq & g(\gamma) \\
H_1 & & 
\end{align*}
\hspace{1cm} (22)

For example, a typical likelihood ratio for the case where \( R \) is only a single element vector (i.e., \( R \) represents only one sample) is given by:

\[
L(R) = \exp \left[-a(2R + a)\right] \hspace{1cm} (23)
\]

where "a" is a constant. The function \( g(x) = (\ln x + a^2)/(2a) \) is a monotonically decreasing function of \( x \) so the detector may operate on the inequalities:

\[
\begin{align*}
H_2 & & \\
R & \geq & \frac{1}{2a} \left[\ln \gamma + a^2\right] \\
H_1 & & 
\end{align*}
\hspace{1cm} (24)

The left hand side of inequality (22) is known as the sufficient statistic and is a result of any attempt to simplify the processing required by the detector at the time of signal reception. The receiver having the sufficient statistic of inequality (24) (i.e., \( R \)) needs only to compare the received signal with a predetermined constant.
The multiple hypotheses problem is a direct extension of the binary problem and will only be discussed briefly. It may be formulated in terms of likelihood ratios also (Ref. 17); however, it will be presented here from the viewpoint of "a posteriori" probabilities. Consider inequality (18) written as:

\[ f(R|H_1) \cdot P_1 > f(R|H_2) \cdot P_2 \implies \text{choose } H_1 \]  

(25)

If both sides of inequality (25) are divided by \( f(R) \), Bayes' formula may be used to write:

\[ f(H_1|R) > f(H_2|R) \implies \text{choose } H_1 \]  

(26)

Since \( f(H_1|R) \) is the "a posteriori" probability of hypothesis \( H_1 \), the binary hypotheses detector is seen to select the hypothesis with the largest "a posteriori" probability. The multiple hypotheses detector merely computes the "a posteriori" probability of all hypotheses upon reception of the signal, and selects the hypothesis having the largest "a posteriori" probability. This type of receiver will be referred to as an a posteriori probability (APP) detector although the detector needs only to compute the product \( f(R|H_i) \cdot P_i \) for each hypothesis since the denominator, \( f(R) \), is common to all terms to be compared.

The composite hypotheses problem is further complicated by the fact that at least one of the hypotheses may represent a signal with one or more random parameters.
For example, a "return" radar signal will most likely have an unknown phase due to phase changes occurring upon reflection of the signal from the target and an unknown frequency due to a doppler shift.

If \( \theta \) represents the parameter and \( f(\theta|H_i) \) is its conditional density, the numerator and denominator terms of the likelihood ratio may be rewritten as follows:

\[
f(R|H_i) = \int_{-\infty}^{\infty} f(R, \theta|H_i) d\theta = \int_{-\infty}^{\infty} f(R|\theta, H_i) f(\theta|H_i) d\theta \quad (27)
\]

Equation (27) shows that the optimal detector merely averages over the parameter \( \theta \) before forming the likelihood ratio. In this thesis, the composite nature of the hypotheses is due to the fact that under any given hypothesis (i.e., given any "M" sequence) there are several different binary sequences (shifted versions). If \( c_m \) \( m = 1, 2, \ldots, L \) represent these component sequences under \( H_i \), then

\[
f(\theta|H_i) = \sum_{m=1}^{L} p(c_m|H_i) \delta(\theta - c_m) \quad (28)
\]

where \( \delta \) is the Dirac delta function. Substituting equation (28) into equation (27) yields:

\[
f(R|H_i) = \sum_{m=1}^{L} f(R|c_m, H_i) p(c_m|H_i) \quad (29)
\]

If the component sequences are equiprobable, the likelihood ratio for a binary double composite hypotheses problem
(both hypotheses are composite) would reduce to:

\[
L(R) = \frac{\frac{1}{K} \sum_{m=1}^{K} f(R|c_m, H_2)}{\frac{1}{L} \sum_{m=1}^{L} f(R|c_m, H_1)}
\] (30)

where hypotheses \( H_1 \) and \( H_2 \) have \( L \) component sequences and \( K \) component sequences respectively.
CHAPTER V

DEVELOPMENT OF THE OPTIMAL SEQUENCE DETECTOR

This chapter presents the derivations used in the design of a receiver for detecting the specific shifted version of the maximal length sequence being transmitted given the number of bits in that sequence. In the initial derivations, only the 30 sequences of length 15 will be considered so that a better understanding of the basic detection procedures might be obtained. The extension to the longer sequences involves very little additional complexity and will be explained toward the end of this chapter.

Let \( H_i \) represent the hypothesis that the \( i \)th "M" sequence was transmitted (\( i = 1, 2 \) for the 15 bit sequences), let \( f(R_k|H_i) \) represent the probability density function of the \( k \) samples of the received signal given that hypothesis \( H_i \) is true, and let \( P_i \) represent the a priori probability of hypothesis \( H_i \). The optimum receiver, in the sense of the minimum probability of error, will determine hypothesis \( H_i \) to be true if the a posteriori probability \( f(R_k|H_i) \cdot P_i / f(R_k) \) is greater for the given received signal \( R_k \) than the a posteriori
probability for \( R_k \) under all other hypotheses (Chapter IV). Since \( f(R_k) \) is invariant under a change in the hypothesis, the detector needs only to compare the numerator terms

\[ f(R_k | H_i) \cdot p_i. \]

The above described a posteriori probability (APP) detector is optimum with respect to the minimization of error probabilities but may not be optimum with respect to the minimization of detection time (elapsed time to a decision being made by the receiver) or with respect to the minimization of receiver memory requirements. The following derivation of a sequential likelihood ratio (SLR) detector is a modification of the optimal detector developed simultaneously by Nolte (Ref. 10) and Fralick (Ref. 5) to adapt to or "learn" an unknown, but fixed, waveform that recurs randomly in time.

The theory of sequential hypothesis testing was developed by Wald (Ref. 18) and first applied to signal detection by Bussgang and Middleton (Ref. 4) and by Blasbalg (Ref. 3). A sequential likelihood ratio detector forms the likelihood ratio following each received sample and decides either that one of the hypotheses may be selected as being true or that another sample should be processed and the decision process repeated. However, prior to 1965, sequential tests of composite hypotheses dealt only with a decision with respect to the hypotheses and not with respect to the random parameters of these
hypotheses. Nolte and Fralick extended the theory of sequential composite hypotheses tests to include a "learning" of the a priori probabilities of the unknown parameters. Both Nolte's and Fralick's detectors were designed for the single composite hypotheses problem:

\[ H_1: R = S(\theta) + N \]
\[ H_2: R = N \]  \hspace{1cm} (31)

where \( S(\theta) \) denotes the composite signal with parameter \( \theta \). The modification to the works of Nolte and Fralick that will be introduced in this thesis will be the design of a detector for the multiple composite hypotheses problem where:

\[ H_1: R = S_1(\theta) + N \] \hspace{1cm} (32)

The maximal length sequence problem for \( L = 15 \) may be treated as a double composite hypotheses problem. The two hypotheses will be:

\[ H_1: R = S_1(\theta) + N \]
\[ H_2: R = S_2(\theta) + N \] \hspace{1cm} (33)

where \( S_1(\theta) \) and \( S_2(\theta) \) are the two maximal length sequences. The hypotheses are composite since there are 15 possible bit arrangements representing each of the two sequences. The signal waveforms comprising hypothesis \( H_1 \) are shown in figure 18 while those comprising hypothesis \( H_2 \) are shown.
in figure 19. The optimum receiver must realize the likelihood ratio:

\[ L(R_k) = \frac{f(R_k | H_2)}{f(R_k | H_1)} \]  

(34)

Since \( H_1 \) and \( H_2 \) are composite, the receiver must average over the signal set under each hypothesis. Letting \( c_m \), \( m = 1, 2, \ldots, L \) represent the set of shifted versions of either maximal length sequence, the numerator and denominator of the ratio become:

\[ f(R_k | H_i) = \sum_{m=1}^{L} f(R_k | c_m, H_i) p_o(c_m | H_i) \quad i = 1, 2 \]  

(35)

where \( p_o(c_m | H_i) \) represents the a priori probability of the \( m \)th component (shifted version) under hypothesis \( H_i \).

The basic operation of the SLR detector is to realize the likelihood ratio \( L(R_k) \) for the received signal after \( k \) samples based on the likelihood ratio \( L(R_{k-1}) \) following \( k-1 \) samples. Since

\[ f(R_k | H_i) = f(R_{k-1}, R_k | H_i) = f(r_k | R_{k-1}, H_i) f(R_{k-1} | H_i) \]  

(36)

where \( r_k \) represents the \( k \)th sample of the received signal, the likelihood ratio may be written in sequential form as:
Figure 18. The binary time waveforms of the 15 shifted versions of sequence M - 23 (see figure 8). These 15 signals comprise composite hypothesis $H_1$ of the 15 bit sequence detector (equation (33)).
Figure 19. The binary time waveforms of the 15 shifted versions of sequence M - 31 (see figure 8). These 15 signals comprise composite hypothesis $H_2$ of the 15 bit sequence detector (equation (33)).
The objective during the remaining development of the SLR detector will be to express \( L(r_k | R_{k-1}) \) only in terms of the statistics of sample \( k-1 \) rather than in terms of all previous samples as expressed by the dependence on \( R_{k-1} \).

Equation (36) may be rewritten as:

\[
L(R_k) = \frac{f(R_{k-1} | H_2)}{f(R_{k-1} | H_1)} \cdot \frac{f(r_k | R_{k-1}, H_2)}{f(r_k | R_{k-1}, H_1)}
\]

where

\[
L(r_k | R_{k-1}) = \frac{f(r_k | R_{k-1}, H_2)}{f(r_k | R_{k-1}, H_1)}
\]

The objective during the remaining development of the SLR detector will be to express \( L(r_k | R_{k-1}) \) only in terms of the statistics of sample \( k-1 \) rather than in terms of all previous samples as expressed by the dependence on \( R_{k-1} \).

Equation (36) may be rewritten as:

\[
f(r_k | R_{k-1}, H_i) = \frac{f(R_k | H_i)}{f(R_{k-1} | H_i)} \quad i = 1, 2
\]

It should be noted that the left hand side of equation (39) may not be written as \( f(r_k | H_1) \) even if the noise samples are assumed to be independent. This is due to the composite nature of the hypotheses in that the specification of \( H_1 \) does not render the transmitted signal deterministic. The specification of a specific component signal \( c_m \) may be introduced by substituting equation (35) into equation (39).
The term \( f(R_k | c_m, H_i) \) may be rewritten using the assumption of statistically independent noise samples.

\[
f(R_k | c_m, H_i) = f(R_{k-1} | c_m, H_i) f(r_k | c_m, H_i)
\]  

Equation (40) may now be rewritten as:

\[
f(r_k | R_{k-1}, H_i) = \frac{\sum_{m=1}^{L} \frac{f(R_{k-1} | c_m, H_i) p_o(c_m | H_i)}{f(R_{k-1} | H_i)} f(r_k | c_m, H_i)}{f(R_{k-1} | H_i)}
\]  

The term in brackets may be interpreted by starting with the following identity resulting from the application of Bayes' Rule to the density \( f(R_{k-1} | c_m, H_i) \):

\[
f(R_{k-1} | c_m, H_i) p(c_m | H_i) = p(c_m | R_{k-1}, H_i) f(R_{k-1} | H_i)
\]  

where \( p \) is used for the density of \( c_m \) to denote the discrete nature of its distribution as opposed to the continuous distribution of the received signal. \( p(c_m | H_i) \) is simply the a priori probability of component \( c_m \) under hypothesis \( H_i \) and may be written as \( p_o(c_m | H_i) \) to be consistent with equation (35).

\[
p(c_m | R_{k-1}, H_i) = \frac{f(R_{k-1} | c_m, H_i) p_o(c_m | H_i)}{f(R_{k-1} | H_i)}
\]
The right hand side of this equation is the term in brackets in equation (42) and may be interpreted as the a priori probability of component $c_m$ just prior to receiving the $k$th sample. For the purpose of notational shorthand, let $p(c_m | R_{k-1}, H_i) = p_{k-1}(c_m | H_i)$. This notation is consistent with the use of $p_o(c_m | H_i)$ to represent the a priori probability prior to the reception of the first sample.

Equation (42) may now be rewritten as:

$$f(r_k | R_{k-1}, H_i) = \sum_{m=1}^{L} f(r_k | c_m, H_i) p_{k-1}(c_m | H_i)$$

Equation (45) averages over the same signal set prior to receiving the $k$th sample.

It remains only to be shown that the probability $p_{k-1}(c_m | H_i)$ is expressible in terms of the statistics of sample $k-1$ including the a priori probabilities of sample $k-1$, $p_{k-2}(c_m | H_i)$. Equation (44) will not be changed mathematically if the a priori probabilities are indexed ahead by one as follows:

$$p_k(c_m | H_i) = p(c_m | R_k, H_i) = \frac{f(R_k | c_m, H_i) p_o(c_m | H_i)}{f(R_k | H_i)}$$

(46)
Solving equation (44) for $p_0(c_m|H_i)$ and substituting into equation (46) gives:

$$p_k(c_m|H_i) = \frac{f(R_k|c_m,H_1)}{f(R_{k-1}|c_m,H_1)} \frac{f(R_{k-1}|H_i)}{f(R_k|H_i)} p_{k-1}(c_m|H_i) \quad (47)$$

Equations (36) and (39) may be used to rewrite the ratio terms in equation (47) yielding:

$$p_k(c_m|H_1) = f(r_k|c_m,H_i) \frac{1}{f(r_k|R_{k-1},H_i)} p_{k-1}(c_m|H_i) \quad (48)$$

The objective is achieved by substituting equation (45) for the denominator of equation (48) and for both the numerator and denominator of the likelihood ratio, equation (38). The complete set of equations describing the SLR detector are as follows:

$$p_k(c_m|H_i) = \frac{f(r_k|c_m,H_i)}{\sum_{m=1}^{L} f(r_k|c_m,H_i)p_{k-1}(c_m|H_i)} p_{k-1}(c_m|H_i) \quad (49)$$

$$L(R_k) = \frac{\sum_{m=1}^{L} f(r_k|c_m,H_2)p_{k-1}(c_m|H_2)}{\sum_{m=1}^{L} f(r_k|c_m,H_1)p_{k-1}(c_m|H_1)} L(R_{k-1}) \quad (50)$$

These equations describe the sequential realization of the likelihood ratio in that a comparison may be made with
a predetermined constant threshold, given by equation (20), following each received sample. The detector may be said to learn the component signal $c_m$ by means of a comparison of the computed "updated" a priori probabilities, $p_k(c_m|H_i)$, for all components. These probabilities will be referred to as the classification output or the classification probabilities of the SLR detector. Following a given sample, say sample $r_j$, the SLR detector for the 15 bit maximal length sequences would first select one of the two "M" sequences based on whether the likelihood ratio $L(R^j)$ was greater or less than the threshold. The detector would then decide on a specific shifted version of the selected "M" sequence by selecting the component sequence having the greatest classification probability $p_j(c_m|H_i)$.

Each received sample supplies additional information to the receiver to aid in its attempt to discriminate between signal and noise. If the receiver is fortunate enough to receive additional transmissions of the signal (additional radar pulses), the learning process may be continued by processing the entire received signal as though it was modulated by a single binary sequence of length $L' = NL$ instead of length $L$ where $N$ represents the number of times the signal was repeated. The only change to be made in equations (49) and (50) is a change in the upper limit of the summation from $L$ to $L'$. The likelihood ratio will approach zero or infinity, depending on which
hypothesis is true, as the number of received samples
approaches infinity. Also, the classification probability
of the transmitted sequence will approach 1 and those
probabilities of the other \( L - 1 \) sequences will approach
0 as the number of samples approaches infinity.

It is possible for a suboptimum detector to base its
decision on only the classification probabilities since
only one of these probabilities (that of the transmitted
sequence) will converge to 1 as the number of samples
becomes large. However, all of the classification
probabilities under each hypothesis must sum to 1
following any given sample, and this implies that
\[
\lim_{k \to \infty} p_k(c_m | H_i) \text{ is not necessarily zero if } H_i \text{ is false.}
\]
Following any given sample, one of the classification
probabilities under a false hypothesis may be larger than
that classification probability under the true hypothesis
which is converging to 1 even though this latter proba-
bility is larger than all other classification probabil-
ities under the true hypothesis. The result is that an
error will be made in this case if the decision is based
on only the classification probabilities whereas a correct
decision may have been made by the optimum detector pro-
vided the comparison of the likelihood ratio with the
threshold would have identified the true hypothesis.

The computer simulation of the SLR detector was based
on the assumption of Gaussian noise and a unity threshold.
A unity threshold results from ignoring cost considerations and assuming equal a priori statistics on the hypotheses; that is, the probability that either hypothesis is true is $1/2$. The initial a priori probabilities of the components, $p_0(c_m|H_i)$ will also be assumed to be equal under each hypothesis. This latter assumption was made so that any tendency could be noted of the detector's ability to learn a particular shifted version (e.g., the one beginning with $n$ positive pulses where $L = 2^n - 1$) following a fewer number of samples than it would need to learn the other shifted versions. Such a tendency was not noted. The assumption of a unity threshold was made so that any evaluation of the likelihood ratio (for a given set of received samples) that results in a value of $1$ will indicate an inability to make a decision as to which hypothesis is true. This is a result of the implication of a unity likelihood ratio—that a decision can be made only with the help of some a priori knowledge as to which hypothesis is more probable and what penalties or costs are involved in making an incorrect decision versus making a correct decision.

The operational equations of the SLR detector operating in a Gaussian noise environment may easily be derived following a minor change in notation. Let $c_m$ under hypothesis $H_i$ be noted as $c_{mi}$, and let $c_{mi}(k)$ be the $k$th sample of component $c_{mi}$. Then
The SLR detector then operates on the kth sample \( r_k \) in the following manner to produce the likelihood ratio and the classification probabilities:

\[
f(x_k | c_m, H_i) = f(x_k | c_{m_i}(k)) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} [r_k - c_{m_i}(k)]^2} \quad (51)
\]

\[
f(x_k | c_m, H_i) = \frac{\exp\left[-\frac{1}{2\sigma^2}[r_k - c_{m_i}(k)]^2\right]}{\sum_{m=1}^{L} p_{k-1}(c_m|H_i)\exp\left[-\frac{1}{2\sigma^2}[r_k - c_{m_i}(k)]^2\right]} p_{k-1}(c_m|H_i) \quad (52)
\]

\[
L(R_k) = \frac{\sum_{m=1}^{L} p_{k-1}(c_m|H_2)\exp\left[-\frac{1}{2\sigma^2}[r_k - c_{m_2}(k)]^2\right]}{\sum_{m=1}^{L} p_{k-1}(c_m|H_1)\exp\left[-\frac{1}{2\sigma^2}[r_k - c_{m_1}(k)]^2\right]} L(R_{k-1}) \quad (53)
\]

The extension of this binary composite hypotheses detector to the multiple composite hypotheses detector needed for maximal length sequences of length greater than 15 involves only an increase in the computations required. As stated before, the multiple hypotheses detector selects on the basis of the largest a posteriori probability for the given received signal. If \( p_1 \) denotes the a priori probability that hypothesis \( H_1 \) is true, then Bayes likelihood ratio test minimizing the probability of error for the binary hypotheses case is:
\[
\frac{f(r_k|H_2)}{f(r_k|H_1)} \begin{cases} H_2 p_1 \medskip \\
\leq H_1 p_2 \end{cases} \Rightarrow \frac{f(r_k|R_{k-1},H_2)f(R_{k-1}|H_2)}{f(r_k|R_{k-1},H_1)f(R_{k-1}|H_1)} \begin{cases} H_2 p_1 \medskip \\
\leq H_1 p_2 \end{cases} (54)
\]

or

\[
f(r_k|R_{k-1},H_2)f(R_{k-1}|H_2) p_2 \begin{cases} H_2 \medskip \\
\leq H_1 \end{cases} f(r_k|R_{k-1},H_1)f(R_{k-1}|H_1) p_1 (55)
\]

Division of both sides of inequality (55) by the density \(f(r_k)\) will yield the a posteriori probability density function following reception of the \(k\)th sample. Thus the binary detector actually selects the hypothesis corresponding to the largest a posteriori probability just as does the multiple hypotheses detector. The a posteriori density may be written sequentially just as the density \(f(r_k|R_{k-1},H_1)\) was written sequentially before--by the use of equation (45).

\[
f(r_k|R_{k-1},H_1)f(R_{k-1}|H_1) p_i = \left[ \sum_{m=1}^{L} f(r_k|c_m,H_i)p_{k-1}(c_m|H_i) \right] f(R_{k-1}|H_i) p_i (56)
\]

The SLR detector for sequences of length 31 will require the computation of 6 a posteriori probabilities (or 5 likelihood ratios) corresponding to the 6 hypotheses required, and the computation of \(6 \times 31 = 186\) classification
probabilities following the reception of each sample.

The sequential realization of the optimal detector requires more computations than the multiple hypotheses realization of the APP detector where each shifted version is treated as a different hypothesis. However, it was noted following the computer analysis that for a signal to noise ratio of unity, the SLR detector for sequences of length 31 had made the correct decision in better than 85% of the simulation runs before reception of the 15th sample. This means that the computations required by the SLR detector could often be reduced by a factor of 2 and that a decision could be made much sooner than by the APP detector.

The grouping of the component signals into composite hypotheses does not hinder the ability of the detector to discriminate between the component signals under a given hypothesis provided the entire sequence has been received. The following derivation shows that the SLR detector and the APP detector are equivalent following the reception of the entire sequence (after the Lth sample). Consider the classification probabilities following the first and second samples.

\[
p_1(c_m|H_i) = \frac{f(r_1|c_m,H_i)}{\sum_{m=1}^{L} f(r_1|c_m,H_i)p_o(c_m|H_i)} p_o(c_m|H_i) \quad (57)
\]
Substituting equation (57) into equation (58) and removing common terms in the numerator and denominator results in:

$$p_2(c_m | H_i) = \sum_{m=1}^{L} \frac{f(r_2 | c_m, H_i) p_1(c_m | H_i)}{\sum_{m=1}^{L} f(r_2 | c_m, H_i) p_1(c_m | H_i)} p_0(c_m | H_i) \quad (59)$$

Substituting equation (59) into the equation for $p_3(c_m | H_i)$ yields:

$$p_3(c_m | H_i) = \sum_{m=1}^{L} \frac{f(r_1 | c_m, H_i) f(r_2 | c_m, H_i) f(r_3 | c_m, H_i)}{\sum_{m=1}^{L} [f(r_1 | c_m, H_i) f(r_2 | c_m, H_i) f(r_3 | c_m, H_i)]} p_0(c_m | H_i) \quad (60)$$

The trend is clear, and the expression for the kth sample may be written as:

$$p_k(c_m | H_i) = \sum_{m=1}^{L} \frac{f(R_k | c_m, H_i)}{\sum_{m=1}^{L} f(R_k | c_m, H_i)} p_0(c_m | H_i) \quad (61)$$

$p_0(c_m | H_i) = 1, 2, \ldots, L$ and $i = 1, 2$ are the a priori probabilities of all component signals, and these are identical to the a priori probabilities in the multiple hypotheses solution which were denoted by $p_i \quad i = 1, 2, \ldots, M$ where $M$ equals the total number of sequences including shifted versions. Also, in the multiple hypotheses
development, \( f(R_k | c_m, H_1) \) may simply be written as \( f(R_k | H_1') \) since there is a separate hypothesis for each component signal. This gives:

\[
p_k(c_m | H_1) = \frac{f(R_k | H_1') \cdot p_i}{\sum_{m=1}^{L} f(R_k | c_m, H_1)}
\]

(62)

The prime denotes that these hypotheses are multiple rather than binary. Upon division of both the numerator and denominator by \( f(R_k) \), equation (62) becomes:

\[
p_k(c_m | H_1) = \frac{a \text{ posteriori probability of } H_1}{\sum_{m=1}^{L} f(R_k | c_m, H_1) / f(R_k)}
\]

(63)

The selection of the component with the largest classification probability by the SLR detector is equivalent to the selection of the hypothesis with the largest a posteriori probability by the APP receiver since the summation term in the denominator of equation (63) is common to all terms under a given binary hypothesis.

When the number of computations required by an optimal detector seems to be impractical, it is well to remember that the major advantage of knowing the performance of the optimal detector may well be the knowledge of what the upper performance limit is. If a detector that is much easier to implement has a performance near that of the optimal detector (or within permissable tolerances), the
more practical receiver may well be considered optimal when all factors are considered. The next chapter develops such a receiver, and it is shown that the most practical receiver is indeed optimal for the case where the received signal is hard-clipped.
CHAPTER VI

CLIPPER CROSS CORRELATOR DETECTOR - SUBOPTIMAL
SEQUENCE DETECTOR

Such considerations as cost, weight, or volume of equipment necessary to implement the optimal detector may make such a detector impractical. This section is concerned with the development of a much more simplified receiver than that described in the last section, and it is shown that this receiver is itself the optimal receiver if the signal has been hard-clipped before arriving at the detector. For example, the signal may have been hard-clipped in the demodulation process.

Given the fact that the received signal is hard-clipped, the optimum detector may be realized by an SLR detector. The only change to be made from the previous chapter is the use of a new density function. If $\alpha$ represents the decision boundary employed by the hard-clipper,

$$f(w_k | c_m, H_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\alpha} \exp\left[-\frac{1}{2\sigma^2}(x - c_m(k))^2\right] dx \delta(w_k + 1)$$

$$+ \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\alpha}^{\infty} \exp\left[-\frac{1}{2\sigma^2}(x - c_m(k))^2\right] dx \delta(w_k - 1)$$

(64)
This is a discrete distribution that expresses the probabilities of the received signal being hard-clipped to either a +1 or a -1. The classification probabilities will not converge as rapidly to 1 or 0 as for the case of the nonclipped signal since the detector's capability of "learning" the noise has been impeded; that is, noise information is discarded in the clipping operation.

The location of the decision boundary \( \alpha \) may be determined for each received sample by the application of Bayes's binary hypotheses decision rule. That is, the a priori probability of each bit of the "M" sequences may be used to determine the optimum clipping threshold. For the computer analysis, it will be assumed that all sequences and their shifted versions occur with equal probability as explained in Chapter V. Since every shifted version of each sequence is equiprobable, the probability that \( c_{mi}(k) \) is +1, for example, does not vary from one sample to the next, and consideration may be limited to a determination of \( \alpha \) for some general sample. Define the hypotheses for the kth sample as:

\[
H_1: \quad c_{mi}(k) = -1 \quad H_2: \quad c_{mi}(k) = +1
\]

The likelihood ratio for the kth sample is:

\[
L(r_k) = \frac{f(r_k | c_{mi}(k) = +1)}{f(r_k | c_{mi}(k) = -1)} \tag{65}
\]
a will then be that value of \( r_k \) for which this ratio will equal the ratio of the probability of occurrence of a -1 to the probability of occurrence of a +1 (threshold).

Since each maximal length sequence of length \( L \) contains \( \frac{L + 1}{2} \) plus 1's and \( \frac{L - 1}{2} \) negative 1's, \( L(\alpha) = \frac{L - 1}{L + 1} \) (see figures 18 and 19). The boundary is then computed as follows:

\[
\frac{\exp\left[-\frac{1}{2\sigma^2}(\alpha - 1)^2\right]}{\exp\left[-\frac{1}{2\sigma^2}(\alpha + 1)^2\right]} = \frac{L - 1}{L + 1} \quad (66)
\]

\[
-\frac{1}{2\sigma^2}(\alpha - 1)^2 + \frac{1}{2\sigma^2}(\alpha + 1)^2 = \ln \frac{L - 1}{L + 1} \quad (67)
\]

\[
\alpha = \frac{\sigma^2}{2} \ln \frac{L - 1}{L + 1} \quad (68)
\]

An understanding of the need for detector simplification may be gained by observing figure 20 which shows the optimal receiver that realizes the signal distributions. The APP detector is shown for comparison purposes since the simplified receiver to be developed is nonsequential in nature. In the last section it was shown that the APP and SLR detectors are equivalent following reception of the final sample of the binary sequence. If each binary sequence (including each shifted version) is assumed to be equilikely, the comparison of the a posteriori
probabilities required in the multiple hypotheses detector reduces to a comparison of the output signal densities conditional on each hypothesis (Chapter IV). A serial type detector is shown for the purpose of simplifying the figure; however, a parallel detector that would perform some or all of the correlations at the same time would certainly be more appropriate if detection time was important.
The simplified detector returns to correlation type processing as was used in the determination of sequence length as discussed in Chapter III. This detector will be denoted as a clipper cross-correlator (CCC), but will be a modified version of the usual CCC configuration shown in figure 21 (Ref. 7).

![Figure 21. Usual CCC Configuration.](image)

Consider the equivalent APP detector and its density expression with $H_i^j$ denoting the multiple hypotheses (one hypothesis for each shifted version)

$$f(W_K|H_i^j)\cdot P_i/f(W_K) = \prod_{k=1}^{K} \left[ \frac{f(w_k|H_i^j)}{f(w_k)} \right] P_i$$

(69)

under the assumption of independent noise samples. Since the APP detector compares the a posteriori probability for each hypothesis, an equivalent comparison may be made with the density

$$\prod_{k=1}^{K} f(w_k|H_i^j)$$

(70)
for each hypothesis since \( f(w_k) \) is common to all terms to be compared and since the a priori probabilities, \( P_i \), are assumed to be equal. This density is given by (from equation (64)):

\[
\prod_{k=1}^{N} \left[ \alpha \exp\left(-\frac{1}{2\sigma^2} [x - cm_i(k)]^2\right) dx \delta(w_k + 1) \right. \\
+ \left. \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2} [x - cm_i(k)]^2\right) dx \delta(w_k - 1) \right] 
\]

(71)

where each \( cm_i \) represents a different hypothesis. The correlator detector that will now be developed must preserve these a posteriori probabilities.

Since \( cm_i(k) = \pm 1 \), the density of \( w_k \) as given by equation (64) may be expressed as two cases. Let \( \Lambda = 1/\sqrt{2\pi\sigma^2} \).

If \( cm_i(k) = 1 \)

\[
f(w_k) = \Lambda \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2} (x-1)^2\right) dx \delta(w_k + 1) \\
+ \Lambda \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2} (x-1)^2\right) dx \delta(w_k - 1)
\]

(72)

If \( cm_i(k) = -1 \)

\[
f(w_k) = \Lambda \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2} (x+1)^2\right) dx \delta(w_k + 1) \\
+ \Lambda \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2} (x+1)^2\right) dx \delta(w_k - 1)
\]

(73)

It can now be shown that the dependence on \( cm_i(k) \) may be eliminated from the analytical expressions for the
a posteriori probabilities by correlating the hard-clipped signal with the stored reference signal \( c_{mi} \) along with adjusting the threshold \( \alpha \).

The following densities will be equivalent to those in equations (72) and (73) provided \( y_k = c_{mi}(k) \cdot w_k \) as performed by the CCC detector.

If \( c_{mi}(k) = +1 \)

\[
\begin{align*}
 f(y_k) &= A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x - 1)^2\right] dx \delta(y_k + 1) \\
 &\quad + A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x - 1)^2\right] dx \delta(y_k - 1) \\
&= A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x + 1)^2\right] dx \delta(y_k + 1) \\
 &\quad + A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x + 1)^2\right] dx \delta(y_k - 1) \\
&= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_k - 1)^2}{2\sigma^2}\right) + \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_k + 1)^2}{2\sigma^2}\right)
\end{align*}
\]

(74)

If \( c_{mi}(k) = -1 \)

\[
\begin{align*}
 f(y_k) &= A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x - 1)^2\right] dx \delta(y_k + 1) \\
 &\quad + A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x - 1)^2\right] dx \delta(y_k - 1) \\
&= A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x + 1)^2\right] dx \delta(y_k + 1) \\
 &\quad + A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x + 1)^2\right] dx \delta(y_k - 1) \\
&= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_k - 1)^2}{2\sigma^2}\right) + \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_k + 1)^2}{2\sigma^2}\right)
\end{align*}
\]

(75)

Due to the symmetry of the normal distribution, these distributions will be equal provided the threshold is negated in one of them. That is:

\[
\begin{align*}
0 &= A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x - 1)^2\right] dx \\
&\quad + A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x + 1)^2\right] dx \\
&= A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x + 1)^2\right] dx \\
&\quad + A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x - 1)^2\right] dx
\end{align*}
\]

(76)

\[
\begin{align*}
0 &= A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x - 1)^2\right] dx \\
&\quad + A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x + 1)^2\right] dx \\
&= A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x + 1)^2\right] dx \\
&\quad + A \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x - 1)^2\right] dx
\end{align*}
\]

(77)

Equations (74) and (75) may be combined into the single density
\[ f(y_k) = A \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2\sigma^2} (x - 1)^2 \right] dx \delta(y_k + 1) \]

\[ + A \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2\sigma^2} (x - 1)^2 \right] dx \delta(y_k - 1) \quad (78) \]

provided the receiver correlates the threshold \( \alpha \) with the stored reference signal. Figure 22 shows this modified CCC detector which is the optimal detector given that the received signal is hard-clipped to equal signal levels.

It is worthwhile to note that the reduction of the density function was based on the Gaussian noise having zero mean and the binary signal having uniform magnitude; otherwise, the relationships of equations (76) and (77) do not hold.

Equation (78) is expressable in terms of the error function as:

Figure 22. Modified CCC Configuration.
\[ f(y_k) = \frac{1}{2}(1 + \text{erf} \gamma) \delta(y_k + 1) \]
\[ + \frac{1}{2}(1 - \text{erf} \gamma) \delta(y_k - 1) \]  
(79)

where

\[ \gamma = \frac{1}{\sqrt{2\sigma}} (\alpha - 1) \]  
(80)

Equation (79) is now equivalent to the term in brackets in expression (71); therefore, the detection procedure is to decide that the true hypothesis is the one corresponding to that reference signal which yields the largest joint probability.

\[ f(Y_k) = \prod_{k=1}^{K} f(y_k) \]  
(81)

For a given received signal, equation (81) will yield a probability for each reference signal \( c_{mi} \) given by:

\[ [(1/2)(1 + \text{erf} \gamma)]^s [(1/2)(1 - \text{erf} \gamma)]^{L-s} \]  
(82)

where \( s \) represents the number of sign agreements between the hard-clipped signal and the stored reference signal; i.e., \( s \) represents the number of samples (out of a total of \( L \) samples) for which \( y_k = 1 \). The optimal detector will now select that reference signal which produces the largest value of expression (82).
The comparison operation may be greatly simplified if a certain monotonic function of expression (82) is used for comparison. Since this expression is a monotonically increasing function of $s$, the parameter $s$ may be compared under each hypothesis, and the hypothesis selected will be the one corresponding to the reference signal that produces the largest value of $s$. Thus, the detector needs only to count the number of positive pulses out of the modified CCC. Figure 23 shows the serial correlator detector for a binary signal with equal signal levels embedded in an additive zero-mean gaussian noise environment with statistically independent noise samples.

![Diagram of simplified realization of optimal detector for hard-clipped signal]

Figure 23. Simplified Realization of Optimal Detector for Hard-Clipped Signal.
As stated previously, the correlation processor is suboptimal for reception of a binary signal not hard-clipped. A determination of the increase in the probability of error by using the suboptimal detector in place of the optimal receiver must be done by Monte Carlo techniques (see Appendix A) since such a determination by the analytical approach would require an unwarranted amount of computations. The analytical approach to the determination of the probability of error for a multiple hypotheses problem of \( N \) hypotheses would require determination of an \( N - 1 \) dimensional decision boundary in a likelihood ratio decision space, followed by \( (N - 2) \cdot (N - 1) \) probability of error expressions that must be summed to give the total probability of error.
CHAPTER VII

PERFORMANCE OF THE OPTIMAL AND SUBOPTIMAL DETECTORS

A study of the performance of the optimal detector should first include an illustration of the "learning" characteristics of the sequential likelihood ratio detector. The detector will be said to have "learned" the correct sequence by the kth sample provided first, that the likelihood ratios have identified the correct "M" sequence for the kth and all following samples, and second, that the classification probability of the correct shifted version is larger than for all other shifted versions of the same "M" sequence for the kth and all following samples. The 31 bit detector was simulated for the purpose of producing computer generated plots that illustrate the variation of the likelihood ratios and the classification probabilities as a function of the number of samples of the received waveform that have been processed.

For the 31 bit detector, there are 6 a posteriori probabilities, \( f(R_k | H_i) \cdot P_i / f(R_k) \) \( i = 1, \ldots, 6 \), corresponding to the 6 "M" sequences of length 31. However, it is more convenient to form 5 likelihood ratios by using one of the a posteriori probabilities as a reference, say...
\[ f(R_k | H_1) \cdot P_1 / f(R_k) \]. This is due to the fact that the a posteriori probabilities typically reduce by more than a factor of 10 from one sample to the next necessitating the use of the power-of-ten multipliers and producing a very impractical scale for the plots. The five likelihood ratios are:

\[ L_i(R_k) = \frac{f(R_k | H_i)}{f(R_k | H_1)} \quad i = 2, 3, 4, 5, 6 \quad (83) \]

The decision process associated with the likelihood ratios is twofold. First, if all five ratios are less than the threshold \( P_1 / P_1 \), the proper decision is that hypothesis \( H_1 \) is true. Second, if at least one ratio is greater than the threshold, the proper decision corresponds to that ratio which is the largest. If \( L_j(R_k) > L_i(R_k) \) for the set \( \{ i : i \neq j \} \), then \( H_j \) is taken to be the true hypothesis.

Figure 24 shows the plot of the 5 ratios up through the 23rd sample for a trial run using a signal to noise ratio of unity. The plots may be identified by the symbol (or lack of symbol) drawn at each sample point. The line plot corresponding to the ratio representing the "M" sequence which was actually present in the simulation contains no symbols while the other four ratios are denoted by the set of symbols \( \{ +, x, \circ, \tau \} \) representing the plotted values. Consecutive points were then joined by straight lines for the purpose of clarity.
Fig. 24. The five likelihood ratios of the 31 bit detector versus the number of samples processed by the receiver for an average trial run under a signal to noise ratio of unity. The inset shows a magnification of the plot for the first 17 samples along with a horizontal line denoting the threshold corresponding to equal a priori probabilities. The sample points of the ratio corresponding to the sequence received are represented by dots joined by a smooth line while the symbols +, x, ◊, and † denote samples of the other four ratios.
This plot was chosen for illustrative purposes since it represents a trial simulation run that resulted in the sequence being "learned" by the 14th sample - a very typical trial for a signal to noise ratio of unity. The inset shows a magnified version of the plot up to the 17th sample that allows a visual determination of the exact sample on which the sequence became correctly learned. The horizontal line shown in the inset represents a threshold of unity which results from the assumption used throughout this study—that the a priori probabilities of the "M" sequences were assumed to be equal (i.e., $P_i = P_1$ for $i = 2, 3, \ldots, n$ where $n$ represents the number of "M" sequences that exist for the given sequence length being considered). Note that the correct decision was not made until the 14th sample even though the ratio corresponding to the actual "M" sequence received was larger than the other ratios beginning with the 12th sample. The maximum ratio was less than unity for the 12th and 13th samples indicating that the hypothesis to be taken as true was the one appearing in the denominator of each of the five ratios.

Figure 24 illustrates a case for the true hypothesis appearing in the numerator of one of the likelihood ratios. Due to the tremendous value that this ratio attains for samples following the 23rd, it was deemed impractical to plot the ratio for the 124 samples (four repetitions of
the sequence) that were actually processed. The magnitudes of the five ratios following the 31st, 62nd, 93rd, and the 124th samples for the trial run of figure 24 are shown in table 3.

Figure 25 is an illustration of the likelihood ratios for the "worst" trial run out of 1116 trials (1116 trials = 6 trials per shifted version times 31 shifted versions per "M" sequence times 6 different "M" sequences) for a signal to noise ratio of unity. The correct decision was not made until the 37th sample--well into the 2nd repetition of the sequence. Table 4 lists the ratio magnitudes for this trial run following each of the first four repetitions of the sequence.

The decision of the correct "M" sequence as determined by the likelihood ratios does not constitute a correct decision by the detector since the problem of which shifted version of the "M" sequence was transmitted is still unresolved. This decision is made by investigating the classification probabilities as discussed at the beginning of this chapter.

There are 186 classification probabilities of the 31 bit detector but only 31 of these are of interest at any given time. Once the likelihood ratios have identified the most probable "M" sequence, the detector needs only to investigate that set of 31 classification probabilities associated with this "M" sequence. Figure 26 shows three
Table 3. The five likelihood ratios for the 6-hypotheses 31 bit SLR detector following each of the first four sequence repetitions in an average trial run for S/N = 1.
<table>
<thead>
<tr>
<th>Sequence Repetitions ( (N) )</th>
<th>Likelihood Ratios ( L_i(R_{31}, \cdot N) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( i = 1 )</td>
</tr>
<tr>
<td>1</td>
<td>( 1.6 \times 10^{-3} )</td>
</tr>
<tr>
<td>2</td>
<td>( 2.0 \times 10^{-3} )</td>
</tr>
<tr>
<td>3</td>
<td>( 2.1 \times 10^{-8} )</td>
</tr>
<tr>
<td>4</td>
<td>( 2.8 \times 10^{-14} )</td>
</tr>
</tbody>
</table>

Table 4. The five likelihood ratios for the 6-hypotheses 31 bit SLR detector following each of the first four sequence repetitions in the "worst" trial run out of 1116 trials for \( S/N = 1 \).
Figure 25. The five likelihood ratios of the 31 bit detector versus the number of samples processed by the receiver for the "worst" trial run out of 1116 trials under a signal to noise ratio of unity. The sample points of the ratio corresponding to the sequence received are represented by dots joined by a smooth line while the symbols +, x, ◊, and † denote samples of the other four ratios.
of the classification probabilities (plotted versus the number of samples processed) associated with an "M" sequence that was the one being received in a signal to noise ratio environment of unity. These plots are from an average trial run and show the classification probability of the correct shifted version (smooth line plot) being greater than those of the other two shifted versions for all samples greater than the 8th. The correct decision was made by the 9th sample as may be seen from the figure given the fact that the other two line plots shown were selected because they were the ones converging to zero the least rapidly of the other 30 classification probabilities. In fact, the other 28 probabilities were less than 0.005 for the 9th and all succeeding samples.

Figure 26 is only representative of the set of classification probabilities corresponding to the "M" sequence actually present in the simulated trial. These probabilities must sum to unity for each of the six hypotheses ("M" sequences); thus, there must always be at least one classification probability in each of the other 5 sets of 31 probabilities that does not converge to zero (see Chapter V). A study of these other sets showed that at least 2 and usually not more than 3 of these probabilities were significantly different from zero following the processing of any given sample except the first few samples. The only implication to be drawn from the
Figure 26. The classification probability (smooth line plot) of the shifted version actually received and the two out of the other 30 classification probabilities that converged to zero the least rapidly (sample points denoted by the symbols + and x) for an average trial run in a signal to noise ratio environment of unity.
observation of these other sets is that the two or three shifted versions having the largest probabilities were the most similar sequences of the 31 to the actual sequence present.

An observation of the starting point (sample one) of the line plots of these probabilities might lead one to the false belief that all 31 of the plots begin at the same point as did the line plots of the 5 likelihood ratios. Two values of probability may exist following the processing of the first sample under a given hypothesis. A given classification probability is a function of the difference between the received sample and the first bit of the reference sequence (equation (52)). Since only binary sequences are being considered, only two possible values may exist for this difference corresponding to the fact that the receiver, in reality, can only make one of two possible decisions following observation of only the first sample--either a sequence beginning with the one binary bit or the other is present. By the same reasoning, there may exist four different values following the second sample, eight following the third, and sixteen following the fourth sample. However, there may exist only 31 \(2^5 - 1\) different values following the fifth sample since it is impossible for a 31 bit "M" sequence to contain five consecutive zeros (Ref. Chapter II). The phrase "may exist" was used above rather than "will exist" since the
additive noise may be such that the difference on which the classification probabilities are dependent may be the same for different reference signal combinations. For example, if the first bit of all reference signals consists of either a "0" or a "1" as for binary sequences, a received signal of "1/2" would render all the classification probabilities equal following the first sample.

Figure 27 contains three line plots of classification probabilities from the "worst" trial out of 1116 simulation runs for a signal to noise ratio of unity. The other 28 probabilities were less than 0.005 for the 17th and all succeeding samples for this trial. Only 35 samples were plotted for the purpose of clarity; however, the line plot corresponding to the correct shifted version remained above the other plots through the 124th sample. In fact, this plot remained within a distance of $5 \times 10^{-5}$ of unity from the 48th through the 124th sample.

The trial run producing figure 27 was "worst" with respect to deciding the identity of the shifted version present and was not the "worst" trial with respect to deciding the "M" sequence present. For example, the latter trial run, for which the ratios were shown in figure 25, did have the correct shifted version identified by the 11th sample but the detector could not identify the correct "M" sequence until the 37th sample. In more than 90% of the trials, the correct shifted version was
Figure 27. The classification probability (smooth line plot) of the shifted version actually received and the two out of the other 30 classification probabilities that converged to zero the least rapidly (sample points denoted by the symbols + and x) for the "worst" trial out of 1116 trials using a signal to noise ratio of unity.
identified earlier than the "M" sequence. However, for the trial of figure 27, the "M" sequence was identified following the processing of the 29th sample but the detector could not make the final correct decision until the 30th sample as seen from the figure.

The above discussion has centered on only the "worst" trial and one average trial for just one signal to noise ratio and one sequence length. Receiver performance is better illustrated by a study of the probability of error as a function of the number of samples processed for various signal to noise ratios. The probability of error for the SLR detector is the probability that the transmitted sequence will not be the sequence selected by the receiver. The Monte Carlo technique (see Appendix A) was used to determine the probability of error following each sample beginning with the first and ending with the last sample of the fourth iteration of the sequence. Figure 28 shows the "optimum" probability of error curves for 15 bit "M" sequence detection while figure 29 shows the curves for 31 bit detection. The smoothness of these curves attests to the accuracy of the Monte Carlo technique which employed 1080 trials for the 15 bit detector and 1116 trials for the 31 bit detector for each signal to noise ratio.

A comparison of these two figures shows a significant improvement in detector performance of the 31 bit detector.
Figure 28. Probability of error versus the number of samples processed by the receiver for the 15 bit sequential likelihood receiver detecting an "M" sequence and its shifted version in additive gaussian noise. These curves represent the lower limit on the probability of error for the given signal to noise ratio for any 15 bit "M" sequence detector.
15 BIT SEQUENTIAL DETECTOR FOR THE NON CLIPPED SIGNAL
Figure 29. Probability of error versus the number of samples processed by the receiver for the 31 bit sequential likelihood receiver detecting an "M" sequence and its shifted version in additive gaussian noise. These curves represent the lower limit on the probability of error for the given signal to noise ratio for any 31 bit "M" sequence detector.
31 BIT SEQUENTIAL DETECTOR FOR THE NON CLIPPED SIGNAL
over the 15 bit detector if one divides the sample numbers by the sequence length for which each detector is designed. One may then compare the probability of error of the two detectors for a given percentage of the sequence length that has been processed. An example of this improvement would be the reduction in the probability of error from 0.55 following 7 bits of a 15 bit sequence to 0.32 following 15 bits of a 31 bit sequence for a signal to noise ratio of unity. Note that the abscissa scales of the two plots differ by a factor of 2 which is approximately the ratio 31/15; thus, a feeling for the degree of improvement may be gained just by a visual comparison of the two plots.

Consider once again the biphase modulated radar receiver discussed in Chapter I and the fact that additive noise in the transmission channel will create random phase shifts in the waveform. It may be inconvenient or impractical to use a linear analog demodulator that would preserve the gaussian noise information. The result would be the use of a demodulator that performs a hard-clipping operation in conjunction with demodulation yielding, once again, a binary sequence. As discussed in Chapter VI, the clipping operation discards much of the noise information which, in turn, increases the probability of error.

Figures 30 and 31 show the results of the error analysis on the 15 bit and 31 bit sequential detectors
designed to optimally detect a hard-clipped "M" sequence in gaussian noise. The equations describing these detectors, corresponding to equations (52) and (53) describing the optimal detector for the non-clipped signal, may be derived from equation (64) giving the conditional density of the hard-clipped received signal \( w_k \). Upon evaluation of the error function of probability theory \( \text{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz \), the classification probabilities and the likelihood ratio for the binary composite hypotheses case are given by the following set of equations.

Letting

\[
d_{\text{mi}}(k) = \frac{\alpha - c_{\text{mi}}(k)}{\sqrt{2\sigma}}
\]

(84)

gives:

\[
p_k(c_m|H_i) = \frac{1 \pm \text{erf} \ d_{\text{mi}}(k)}{\sum_{m=1}^{L} p_{k-1}(c_m|H_i)[1 \pm \text{erf} \ d_{\text{mi}}(k)]} p_{k-1}(c_m|H_i)
\]

(85)

for \( w_k = \mp 1 \) and, using the same notation:

\[
L(R_k) = \frac{\sum_{m=1}^{L} p_{k-1}(c_m|H_2)[1 \pm \text{erf} \ d_{\text{mi}}(k)]}{\sum_{m=1}^{L} p_{k-1}(c_m|H_1)[1 \pm \text{erf} \ d_{\text{mi}}(k)]} L(R_{k-1})
\]

(86)
Figure 30. Probability of error versus the number of samples processed by the 15 bit sequential likelihood receiver designed to detect an "M" sequence and its shifted version in additive gaussian noise that has been hard-clipped to the binary levels of the "M" sequence. These curves represent the lower limit on the probability of error for the given signal to noise ratio for any 15 bit "M" sequence detector following a hard-clipper.
15 BIT SEQUENTIAL DETECTOR FOR THE HARD CLIPPED SIGNAL

PROBABILITY OF ERROR

SAMPLE NUMBER

S/N = 1/4
S/N = 1/2
S/N = 2
S/N = 4
Figure 31. Probability of error versus the number of samples processed by the 31 bit sequential likelihood receiver designed to detect an "M" sequence and its shifted version in additive gaussian noise that has been hard-clipped to the binary levels of the "M" sequence. These curves represent the lower limit on the probability of error for the given signal to noise ratio for any 31 bit "M" sequence detector following a hard-clipper.
31 BIT SEQUENTIAL DETECTOR FOR THE HARD CLIPPED SIGNAL
A comparison of figures 30 and 31 with figures 28 and 29, respectively, demonstrates the degradation in system performance that has been caused by the clipping operation. Table 5 shows the increase in the probability of error following the processing of just \(L\) samples by the \(L\) - bit detectors for \(L\) equal to 15 and 31. Table 5 also gives a further numerical indication of the lower probability of error for the 31 bit detector than for the 15 bit detector as discussed earlier.

<table>
<thead>
<tr>
<th>Probability of Error Following the 15th Sample for the 15 Bit Detector</th>
<th>Probability of Error Following the 31st Sample for the 31 Bit Detector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-clipped</td>
<td>Hard-clipped</td>
</tr>
<tr>
<td>.71</td>
<td>.88</td>
</tr>
<tr>
<td>.43</td>
<td>.69</td>
</tr>
<tr>
<td>.22</td>
<td>.52</td>
</tr>
<tr>
<td>.10</td>
<td>.35</td>
</tr>
<tr>
<td>.01</td>
<td>.08</td>
</tr>
</tbody>
</table>

Table 5. Comparison of the Probabilities of Error for the Optimal Detection of a Hard-clipped Signal to That for the Optimal Detection of a Non-clipped Signal for the 15 Bit and 31 Bit Optimal "M" Sequence Detectors.

Chapter VI contained the derivation of a correlation type detector that was shown to be equivalent (with respect to the probability of error) to the nonsequential realization of the optimal detector for hard-clipped binary signals provided two restrictions are satisfied. These restrictions are that the noise must be zero-mean
gaussian and that the received signal must be clipped to binary levels of equal power (Ref. Chapter VI). Since these restrictions were met throughout this study, table 5 presents the complete error analysis for the 15 bit and 31 bit correlation detectors since these detectors are constrained to making a decision following the reception of the final sample of the received sequence.
CHAPTER VIII

SUMMARY AND DISCUSSION OF FUTURE RESEARCH

Given that the receiver has knowledge of the time interval in which the signal is being received, the number of pulses comprising the maximal length sequence that is being transmitted (sequence length) may be determined by the following procedure. A sampling rate is selected that is equal to the number of bits comprising the longest maximal length sequence expected to be received divided by the time duration of the received signal. The time samples of the received signal are correlated with themselves (time autocorrelation) and the ratio of the center peak of the autocorrelation (waveform power) to the maximum noncenter value is compared with a predetermined threshold. If this ratio exceeds the threshold, the receiver decides that the number of samples equals the number of bits of the transmitted maximal length sequence; otherwise, a new sampling rate is chosen which is related to the previous rate by the expression: new rate = (1/2)(old rate - 1). The above procedure is repeated for this new rate.

Knowledge of which maximal length sequence was received given the sequence length may optimally be
determined by computing expression (87) given below for each hypothesis $H_i$ where $H_i$ is the hypothesis that the $i$th maximal length sequence was received where $i = 1, 2, \ldots, n$ such that $n$ is the number of maximal length sequences of the given length under consideration.

\[
\sum_{m=1}^{L} f(r_k | c_m, H_i) p_{k-1}(c_m | H_i) f(r_{k-1} | R_{k-2}, H_i) \cdot p_i \quad (87)
\]

where

\[
f(r_{k-1} | R_{k-2}, H_i) = \sum_{m=1}^{L} f(r_{k-1} | c_m, H_i) p_{k-2}(c_m | H_i) \quad (88)
\]

and

\[
p_k(c_m | H_i) = \frac{f(r_k | c_m, H_i)}{\sum_{m=1}^{L} f(r_k | c_m, H_i) p_{k-1}(c_m | H_i)} p_k(c_m | H_i) \quad (89)
\]

$p_i$ = a priori probability of hypothesis $H_i$

$c_m$ = $m$th shifted version of the $i$th sequence

$r_k$ = $k$th sample of the received signal

$R_k$ = the set of the first $k$ samples of the received signal

$L$ = sequence length

Expression (87) is the numerator of the a posteriori probability of the set of samples $R_k$ given the hypothesis
The jth maximal length sequence will be said to be present following the kth sample if expression (87) is greater for $i = j$ than for all $i \neq j$.

If knowledge is also desired as to which shifted version of the maximal length sequence is present in the received signal, the probabilities of equation (89), denoted as the classification probabilities, are investigated for each of the L shifted versions, and the shifted version $c_m$ with the largest classification probability will be said to be present. This sequential realizability allows decisions to be made by the receiver following any given sample rather than only after the Lth sample as is the case with a nonsequential detector designed to receive an L bit sequence.

The correlation type processor shown in figure 23 is equivalent to the nonsequential optimal detector for the situation where the received signal is hard-clipped before reaching the detector provided the noise is zero-mean gaussian and the magnitudes of the binary levels of the clipped signal are equal. This receiver is far more simplified with respect to implementation than even the nonsequential detector that is optimal for the reception of a nonclipped binary signal, but its performance is considerably worse as can be seen from a study of table 5 for the 15 bit and the 31 bit maximal length sequence detectors.
At the time of this writing, further investigation is planned into three aspects of this problem of the detectability of maximal length sequences. First, a sequence length or sampling rate detector needs to be designed for the case of a hard-clipped signal. The sampling rate detector proposed in Chapter III operates at such a higher probability of error on the hard-clipped signal than on the nonclipped signal that it was deemed practical to search for a different type of processor. Whereas the analytical development of the optimal sampling rate detector seemed to be beyond the present development of mathematics even with the recent advances of dynamic programming techniques, preliminary investigation has indicated a possibility of deriving an approximate expression for the optimal detector for the case of hard-clipped signals. As may be seen from equation (64), the hard-clipping operation reduces the density functions from continuous functions to impulse distributions which, in turn, simplify required integrals of these functions. This near optimal detector for the hard-clipped signal may prove to be a better detector of the nonclipped signal than the one proposed in this study.

The second area of further investigation would include additional investigation into the frequency domain sampling rate detector (Ref. 1). Further study would include an error analysis of both this detector and the one proposed
in this thesis so that the relative efficiencies of the frequency domain detectors and time domain detectors could be compared.

Finally, investigation is planned into the practicality of using a sampling rate for the sequence detector that is some multiple of the sequence length. Although this additional sampling magnifies the required number of computations, it will reduce the number of samples that need to be processed in order to accomplish a given probability of error. However, an increase in the sampling rate may invalidate the assumption that the noise samples are independent (see Appendix B). It will then be necessary to rederive the detector equations for consideration of an mth order Markov noise process.
The Monte Carlo method is a technique of solving various problems in computational mathematics by constructing for each problem a random process with parameters equal to the required quantities of that problem. The unknowns are determined approximately by carrying out observations on the random process, and by computing its statistical characteristics which are approximately equal to the required parameters.

The Monte Carlo method was used in this thesis to determine the optimum threshold for the proposed sampling rate detector and to determine the probability of error for the proposed sequence detectors. Both applications employed the assumption that the a priori probabilities of all $M$ sequences of a given length (including each shifted version) were equal.

The determination of the optimum threshold for the sampling rate detector involved a dual application of the Monte Carlo method. Consider, first, the sampling rate detector to be sampling at a given rate $l_1/T$ and to be
receiving only those sequences of length $L_1$ (i.e., the detector is always correctly sampling the sequence). The Monte Carlo method then employs the following procedure in determining the probability that a sequence of length $L_1$ will be identified as not being of length $L_1$ (error of omission):

1) Select a threshold that is well below those values that might be thought to represent expected values of the optimum threshold (for example, an initial threshold value of 3 would be appropriate for $L_1$ equal to 31 and S/N equal to 1/2--see figure 10).

2) Select one "M" sequence of length $L_1$ and add each bit to a random number (noise sample).

3) Autocorrelate the resulting sequence and compute the autocorrelation ratio (ACR).

4) Note whether or not the ACR is below the threshold. This may be done by letting $e_1(k)$ be a quantity that is 1 if the ACR was below the threshold on the ith trial (this being the first trial) and 0 otherwise. The $k$ may be used to denote the various thresholds used (e.g., $e_1(1)$ would represent the result from comparing the ACR with the initial threshold value).

5) Increase the threshold by an increment equal to the accuracy desired in the value of the optimum threshold. A value of 0.1 was used in obtaining the results described in this thesis.
6) Repeat step 4 (i.e., determine $e_1(2)$).

7) Repeat steps 5 and 6 until the trial threshold is well above those values that might be thought to represent expected values of the optimum threshold (for example, a final value of 6 would be appropriate for $L_1$ equal to 31 and $S/N = 1/2$--see figure 10). The set of binary numbers \{e_1(1), e_1(2), \ldots, e_1(K)\} is now determined.

8) Repeat steps 2 through 7 for another "M" sequence. That is, perform a second trial and determine another set of values denoted by \{e_2(1), e_2(2), \ldots, e_2(K)\}.

9) Continue the above procedure until the number of trials, $N$, equals that number required to produce the desired accuracy (to be discussed below). Since equal a priori probabilities were assumed for the maximal length sequences, each sequence (including each shifted version) should be used in a total of $N/M$ trials where $M$ equals the total number of sequences of length $L_1$.

10) Following the final trial, the number of errors of omission, denoted by $E(k)$, must be determined for each value of the threshold where:

$$E(k) = \sum_{i=1}^{N} e_i(k) \quad (90)$$
11) The relative frequency of this error of omission, 
\( E(k)/N \), is used as the approximation to the proba­
bility of a sequence of length \( L_1 \) being identified
as not being of length \( L_1 \) given that the threshold
has the value identified as the kth threshold.

The optimum threshold was defined to be the threshold
which minimizes the probability of error which must
include, in addition to the probability of the error of omission, the probability of the error of inclusion.
This latter error probability is the probability that the
receiver will decide that a sequence not of length \( L_1 \) was
of length \( L_1 \), and this probability may also be approximated
by the Monte Carlo method. This involves using as trial
sequences (step 2) all "M" sequences which are oversampled
by the receiver which is using a sampling rate of \( L_1/T \).
However, it was found that the change in the optimum
threshold was insignificant upon limiting the Monte Carlo
analysis only to those sequences of length \( L_2 = \frac{L_1 - 1}{2} \); that is, by limiting the analysis to those sequences that
are oversampled by a factor of approximately 2.

As the threshold is increased, the probability of the
error of omission, denoted as \( P(\text{omission}) \), increases while
the probability of the error of inclusion, denoted as
\( P(\text{inclusion}) \), decreases which, in turn, yields a proba­
bility of error which is maximum for some value of the
threshold. Both of the above denoted probabilities are conditional and should be written as follows:

\[ P(\text{omission}|\text{bit sequence sampled } L_1 \text{ times}) \]
\[ P(\text{inclusion}|L_2 = \frac{1}{2} (L_1-1) \text{ bit sequence sampled } L_1 \text{ times}) \]

The probability of error is then:

\[ P(\text{error}) = P(\text{omission}|\text{bit seq. sampled } L_1 \text{ times}) \cdot P(\text{receiving } L_1 \text{ bit sequence}) \]
\[ + P(\text{inclusion}|L_2 = \frac{1}{2} (L_1-1) \text{ bit seq. sampled } L_1 \text{ times}) \cdot P(\text{receiving } L_2 \text{ bit sequence}) \]  

(91)

Since this thesis considered the probability of receiving a sequence of length \( L_1 \) to be equal to the probability of receiving a sequence of the next shorter length \( (L_2) \), the probability of error becomes the average of the probability of omission and the probability of inclusion. The optimum threshold is then the threshold for which this average, \( \frac{1}{2} [P(\text{omission}) + P(\text{inclusion})] \), is minimum.

Equal a priori probabilities (the marginal probabilities in equation (91) were assumed in this study for the convenience of illustrating the appearance of a typical optimum threshold. In an application where this assumption is not good and where the actual a priori probabilities are unknown, a mini-max procedure (Ref. 17) may be used to determine the optimum threshold. This will be the threshold for which the probability of omission equals the
probability of inclusion.

The accuracy of the Monte Carlo method depends both on the probability that is being approximated by the relative frequency being computed and on the number of trials. If \( p \) represents the desired probability (e.g., the probability of the error of omission), then the variance of the relative frequency term may be found as follows:

\[
\text{var}\left\{ \frac{E(k)}{N} \right\} = \frac{\text{var}\{E(k)\}}{N^2} = \frac{1}{N^2} \sum_{i=1}^{N} \text{var}\{e_i(k)\}
\]

\[
= \frac{1}{N^2} \sum_{i=1}^{N} [E\{e_i^2(k)\} - E^2\{e_i(k)\}]
\]

\[
= \frac{1}{N^2} \sum_{i=1}^{N} [0^2(1-p) + 1^2p - (0(1-p) + 1\cdot p)^2]
\]

\[
= \frac{1}{N^2} \sum_{i=1}^{N} p(1-p) = \frac{p}{N} (1-p)
\]

(92)

The variance along with the Chebyshev inequality may be used to provide an estimate on the number of trials required to achieve a relative frequency within given bounds provided \( pN \gg 1 \). If \( \delta \) is the bound on the error of the Monte Carlo method, then

\[
\frac{E(k)}{N} = p \pm \delta
\]

(93)

provided (see Ref. 14 for derivation)
An example of this estimate may be given by considering the sampling rate detector using 63 samples per sequence in an environment with a signal to noise ratio of 1/2. Following about 1000 trials, the optimum threshold was seen to be convergent to 4.5 for which the probability of omission was about 1/3 and the probability of inclusion was about 1/2. Equation (94) indicates that, if the desired error for each probability is to be less than 0.07, 1837 trials must be used to compute the probability of the error of inclusion and 3673 trials for the probability of the error of omission. The actual numbers of trials used were 1860 (186 sequences of length 31 times 10 trials per sequence) for P(inclusion) and 3780 (378 sequences of length 63 times 10 trials per sequence) for P(omission).

The optimum threshold curves denoted by the dashed lines in figures 10 and 11 do not extend to the right edge of either of the figures because both P(omission) and P(inclusion) become so small that the use of the Monte Carlo method is impractical. For example, for the sampling rate of 63/T, P(omission) is less than 0.001 for S/N > 5. According to equation (94), an error of less than 0.07 would require approximately 2 million trials, and an error this large (δ = 0.07) is not even consistent with the

\[
N = \frac{9(1 - p)}{p \delta^2}
\]
magnitude of the probability of error that is being sought.

The Monte Carlo method was used with respect to the sequence detector to determine an estimate of the probability of error that existed following the processing of each sample of the received signal. The probability of error following reception of the kth sample is the probability that the sequence that was actually transmitted has not been identified following the processing of the kth sample. The Monte Carlo method applied to this problem involves the following steps:

1) Select one "M" sequence of length $L_1$ and add each bit to a random number (noise sample).

2) Compute the likelihood ratios and the classification probabilities, and determine the sequence to be selected by the receiver from the information of the likelihood ratios and classification probabilities.

3) Let $s_i(k)$ be a quantity that is 1 if the sequence was correctly determined in the ith trial by the kth sample (this includes being correctly determined following all samples between the kth and the final sample processed—sample number $4 \cdot L_1$ in this study) and is 0 otherwise.

4) Repeat steps 1 through 4 until the number of trials, $N$, equals that number required to produce the desired accuracy (see equation (94)). Since equal a priori probabilities were assumed for the maximal length
sequences, each sequence (including each shifted version) should be used in a total of \(N/M\) trials where \(M\) equals the total number of sequences of length \(L_1\).

5) Following the final trial, the number of successful identifications, denoted by \(S(k)\), must be determined for each sample where:

\[
S(k) = \sum_{i=1}^{N} s_i(k) \quad (95)
\]

6) The relative frequency of successful identifications, \(S(k)/N\), is used as an approximation to the probability that an error is not made in identifying the transmitted sequence by the \(k\)th sample. The probability of error is approximated by \(1 - S(k)/N\).

Application of the Monte Carlo method in the determination of small probabilities of error is so time consuming as to be impractical in most cases. This is due to the fact that the number of necessary trials varies inversely as the cube of the desired probability as may be seen from equation (94). For small \(p\),

\[
N = \frac{9}{p\delta^2} \quad (96)
\]

But the desired error bounds should be at least as small as \(\delta\) (i.e., \(\delta < p\)); hence,

\[
N > \frac{9}{p^3} \quad (97)
\]
This implies that at least 9 million trials are needed for a probability of error as small as only 0.01.
APPENDIX B

CORRELATION OF NOISE SAMPLES

The aperiodic time autocorrelation of the time samples of a stochastic noise process is given by (equation (6), Chapter III):

\[ R(k\tau_o) = \tau_o \sum_{i=1}^{L-k} n_i n_{i+k} \]  \hspace{1cm} (98)

where \( k \) is an integer, \( \tau_o \) is the time interval between samples, \( L \) is the number of samples, and \( n_i \) is the \( i \)th time sample of the noise process. The samples may be said to be uncorrelated if \( R(k\tau_o) \) for all \( k \neq 0 \) is several orders of magnitude smaller than \( R(k\tau_o) \) for \( k = 0 \). That is,

\[ R(k\tau_o) \bigg|_{k \neq 0} << R(0) \]  \hspace{1cm} (99)

If the values of \( R(k\tau_o) \) are plotted for all values of \( k \) and adjacent points are joined by a straight line, the plot should be a rough approximation to the function \( f(x) = \delta(x) \). This is the discrete time autocorrelation analogy to the statistical autocorrelation of a stochastic process known
as "white noise". A stationary stochastic noise process, \( n(t) \), is said to be white noise provided its statistical autocorrelation, \( R_{nn}(\tau) = E\{n(t)n(t + \tau)\} \), is given by:

\[
R_{nn}(\tau) = \delta(\tau) \quad (100)
\]

The random numbers used to simulate the noise samples may be considered to be uncorrelated since they meet condition (99) as may be seen from figure 32. The points \( R(k\tau_o) \) for \( k = 0, 1, \ldots, 550 \) have been joined by straight lines to add some continuity to the figure, and \( R(k\tau_o) \) has been plotted only for positive \( k \)'s since \( R(k\tau_o) \) equals \( R(-k\tau_o) \). The noise autocorrelation is of interest only up to the 511th point since a sequence of 511 noise samples was the longest sequence used in the simulations.

In a practical sense, the assumption of uncorrelated noise samples is valid for many common noise interferences such as thermal noise unless the time interval between samples is very small and/or the bandwidth of the system is very wide. This statement may be clarified by noting that the autocorrelation of the noise samples, \( R(k\tau_o) \), which may be envisioned as a set of impulses, approaches the autocorrelation, \( R_{nn}(\tau) \), of the stochastic process \( n(t) \) as the sampling interval, \( \tau_o \), becomes very small.

\[
\lim_{\tau_o \to 0} R(k\tau_o) = R_{nn}(\tau) \quad (101)
\]
Figure 32. The autocorrelation of 550 of the random numbers used to simulate noise samples in the simulations of the sampling rate detector and the sequence detector. The "x" on the vertical axis marks the center peak.
Pure white noise is defined to have a power spectrum, $S(\omega)$, (Fourier transform of autocorrelation function) that is uniform for all $\omega$ ($\mathcal{F}\{k\delta(\tau)\} = k$). However, such noise cannot exist since this implies infinite average power for the noise process. That is, if

$$S_{nn}(\omega) = \int_{-\infty}^{\infty} R_{nn}(\tau)e^{-j\omega\tau}d\tau = k$$  \hspace{1cm} (102)$$

then,

$$E\{n^2(t)\} = R_{nn}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{nn}(\omega)d\omega = \frac{k}{2\pi} \int_{-\infty}^{\infty} d\omega \hspace{1cm} (103)$$

Equation (103) implies that the power spectrum must experience cutoff at some frequency. However, the spectrum may be relatively constant below the cutoff frequency, and, in this case, white noise is a good assumption provided the frequency range over which $S_{nn}(\omega)$ is relatively constant is much broader than the bandwidth of the communication system.

A power spectrum that is relatively constant over the frequency range of interest but is not perfectly constant will correspond to an autocorrelation function with an appearance similar to the autocorrelation of figure 32. That is, the center peak or lobe will have a finite width. If the sampling interval $\tau_o$ becomes small enough, condition (99) will not hold for small $k$ since the values of $R(k\tau_o)$ for small $k$ will correspond to values of the autocorrelation $R_{nn}(\tau)$ that are well up on the center lobe. This
implies a degree of correlatedness between adjacent samples which violates the assumption of uncorrelated noise samples.
BIBLIOGRAPHY


