2016


Michael J. Nowak
Wright State University

Follow this and additional works at: http://corescholar.libraries.wright.edu/etd_all
Part of the Engineering Commons

Repository Citation
http://corescholar.libraries.wright.edu/etd_all/1476

This Dissertation is brought to you for free and open access by the Theses and Dissertations at CORE Scholar. It has been accepted for inclusion in Browse all Theses and Dissertations by an authorized administrator of CORE Scholar. For more information, please contact corescholar@www.libraries.wright.edu.

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

by

Michael James Nowak
B.S. Astronautical Engineering, U.S. Air Force Academy, 1979
M.S. Aerospace Engineering, University of Texas, Austin, 1992

2016
Wright State University

Kuldip S. Rattan, Ph.D.
Dissertation Director

Frank W. Ciarallo, Ph.D.
Director, Ph.D. in Engineering Program

Robert E. W. Fyffe, Ph.D.
Vice President for Research and Dean of the Graduate School

Committee on Final Examination

Kuldip S. Rattan, Ph.D. Zhiqiang Wu, Ph.D.

Xiaodong Zhang, Ph.D. Michael C. Wicks, Ph.D.

Robert Penno, Ph.D. Michael L. Bryant, Ph.D.
ABSTRACT


As electromagnetic spectrum availability shrinks, there is growing interest in combining multiple functions, such as radar and communications signals, into a single multipurpose waveform. Historically mixed-modulation has used orthogonal separation of different message signals in different dimensions such as time or frequency. This research explores an alternative approach of implementing an in-band, mixed-modulated waveform that combines surveillance radar and navigation functions into a single signal and demonstrates the feasibility of this approach for a limited area navigation solution. The first contribution of this research is the use of reduced phase-angle binary phase shift keying (BPSK) along with overlapped (channelized) spread-spectrum phase discretes based on pseudorandom noise sequences to encode multiple messages in a single pulse. The goal is to determine if these signals can achieve satisfactory message performance in a radar pulse while minimizing the effect on radar performance. For the purpose of this research, radar performance will be evaluated in terms of power spectral density, matched filter auto-correlation for target detection, and the ambiguity function. The second contribution is the development of an algorithm that permits the use of received navigation messages with non-synchronous times of arrival. Numerical simulations of the resulting iterative, non-synchronous geolocation algorithm demonstrate rapid convergence of the estimated and true trajectory despite large temporal differences in message time of arrival.
List of Symbols

Chapter 1

$\phi_{\delta}$  Magnitude of phase angle change
$j$  $\sqrt{-1}$

Chapter 2

$\omega$  Angular frequency (Hz)
$B_w$  Bandwidth (Hz)
$f_0$  Initial LFM frequency (Hz)
$f_s$  Sample rate (Hz)
$\phi_a$  Phase angle - subscript $a$ determines type
$n$  Sample instance (integer)
$t$  Time (sec)
$P$  Power (watts)
$E$  Energy (joules)
$T$  Sample time increment (sec)
$S$  Power Spectral Density
$R_{ab}$  Auto or Cros-correlation of $a$ and $b$
$\chi$  Value of radar ambiguity function
$\nu$  Doppler frequency shift (Hz)
$\tau$  Time delay (sec)
$C$  Covariance matrix
$\sigma^2$  Variance

Chapter 3

$x$  Signal vector
$\kappa$  Boltzmann’s constant
$G_a$  Range Equation: Gain; subscript $a$ determines type
$G_n$  Gold Code Message vector $n$
$T_0$  Antenna temperature
$[SNR]$  Signal to noise ratio
$F$  Receiver noise factor (dB)
$W$  Primitive root of unity

Chapter 6

$m_a$  Message vector - subscript, $n$, determines type

Chapter 7

$T$  Navigation - Master clock reference time
$\tau$  Signal transit time: radar to receiver
$c$  Speed of light
$t$  Time (sec)
Chapter 8

\( \alpha, \beta, \gamma \) Defined variables: IMU position change
\( \delta T \) Time separation between navigation encoded pulses
\( N_{ab} \) IMU measured position change between locations \( a \) and \( b \)
## Contents

1 Introduction ........................................... 1
   1.1 Previous Work ........................................ 3
   1.2 Problem Description / Decomposition ................. 5
   1.3 Outline ................................................. 6
   1.4 Mathematical Notation ................................ 8

2 Background ............................................. 10
   2.1 Digital Linear Frequency Modulated Signal ............ 11
   2.2 Signal Energy and Power Spectral Density Functions .. 12
   2.3 Error / Complementary Error / Q-Functions .......... 14
   2.4 Ambiguity Function ..................................... 16
      2.4.1 Ambiguity Function Properties .................. 17
   2.5 Geometric Dilution of Precision (GDOP) ............... 18

3 Communications Message Performance .................. 22
   3.1 In-Band Mixed Modulation ............................. 23
   3.2 Message Performance .................................... 26
      3.2.1 Reduced Phase Margin Bit Error Rate .......... 27
      3.2.2 Radar Range Equation - Bit Energy ........... 30
   3.3 Encoding Approaches .................................... 32
      3.3.1 Constant Valued Phase Changes ................. 32
      3.3.2 Barker Codes ..................................... 32
      3.3.3 M-Sequences / Gold Sequences .................. 34
      3.3.4 Polyphase Sequences ................................ 38
   3.4 Summary ................................................. 39

4 Radar Performance ........................................ 41
   4.1 Ambiguity Function ..................................... 42
   4.2 Phase Message Effects on Ambiguity Function ........ 43
   4.3 LFM Signal Baseline Performance ...................... 45
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Single Channel Modulation</td>
<td>49</td>
</tr>
<tr>
<td>5.1</td>
<td>Constant Phase Angle Modulation</td>
<td>49</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Constant Small Magnitude Phase Angle</td>
<td>53</td>
</tr>
<tr>
<td>5.2</td>
<td>Barker Encoded Phase Angle Modulation</td>
<td>56</td>
</tr>
<tr>
<td>5.3</td>
<td>Summary</td>
<td>62</td>
</tr>
<tr>
<td>6</td>
<td>Channelized Modulation</td>
<td>63</td>
</tr>
<tr>
<td>6.1</td>
<td>Channelized BPSK Encoding</td>
<td>64</td>
</tr>
<tr>
<td>6.2</td>
<td>Binary M-Sequence Message Recovery</td>
<td>66</td>
</tr>
<tr>
<td>6.3</td>
<td>Channelized Message Radar Performance</td>
<td>69</td>
</tr>
<tr>
<td>6.4</td>
<td>Channelized QPSK Encoding</td>
<td>75</td>
</tr>
<tr>
<td>6.5</td>
<td>M-ary Constellations</td>
<td>78</td>
</tr>
<tr>
<td>7</td>
<td>Radio Based Navigation Problem</td>
<td>80</td>
</tr>
<tr>
<td>7.1</td>
<td>Pseudorange Measurements - Satellite Navigation</td>
<td>81</td>
</tr>
<tr>
<td>7.2</td>
<td>TDOA Measurements - LORAN</td>
<td>84</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Summary</td>
<td>89</td>
</tr>
<tr>
<td>8</td>
<td>Non-Synchronous TDOA Geolocation</td>
<td>90</td>
</tr>
<tr>
<td>8.1</td>
<td>Non-Synchronous Timing Approach</td>
<td>92</td>
</tr>
<tr>
<td>8.2</td>
<td>Non-Synchronous Algorithm Development</td>
<td>94</td>
</tr>
<tr>
<td>8.2.1</td>
<td>Algorithm Singularities</td>
<td>101</td>
</tr>
<tr>
<td>8.2.2</td>
<td>Iterative Geolocation Algorithm</td>
<td>103</td>
</tr>
<tr>
<td>8.3</td>
<td>Non-Synchronous TDOA Algorithm Simulation</td>
<td>105</td>
</tr>
<tr>
<td>8.4</td>
<td>GDOP Considerations for the 3-D Algorithm</td>
<td>109</td>
</tr>
<tr>
<td>9</td>
<td>Summary / Follow-On Research</td>
<td>111</td>
</tr>
<tr>
<td>9.1</td>
<td>Research Summary</td>
<td>111</td>
</tr>
<tr>
<td>9.2</td>
<td>Waveform Development Summary</td>
<td>114</td>
</tr>
<tr>
<td>9.2.1</td>
<td>Constant Phase Angle Encoding</td>
<td>115</td>
</tr>
<tr>
<td>9.2.2</td>
<td>Barker Encoding</td>
<td>115</td>
</tr>
<tr>
<td>9.2.3</td>
<td>M-Sequences - BPSK</td>
<td>115</td>
</tr>
<tr>
<td>9.2.4</td>
<td>M-Sequences- QPSK</td>
<td>116</td>
</tr>
<tr>
<td>9.3</td>
<td>Non-Synchronous Navigation Summary</td>
<td>116</td>
</tr>
<tr>
<td>9.4</td>
<td>Future Work</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td><strong>Bibliography</strong></td>
<td>120</td>
</tr>
</tbody>
</table>
List of Figures

3.1 Intentional Modulation of Radar Pulse with Communications Message . . . 23
3.2 Representative Message Bit . . . . . . . . . . . . . . . . . . . . . . . . . . 25
3.3 Reduced Phase Angle Bit Distance . . . . . . . . . . . . . . . . . . . . . . 28
3.4 Reduced Phase Angle Bit Error Rate . . . . . . . . . . . . . . . . . . . . . 29
3.5 Autocorrelation of $n=13$ Barker Code . . . . . . . . . . . . . . . . . . . 33
3.6 Autocorrelation and Cross-Correlation of Generated M-Sequences . . . . 35
3.7 Three-Sequence Gold Code Correlation . . . . . . . . . . . . . . . . . . . . 37
4.1 PSD/Autocorrelation Unencoded LFM Pulse . . . . . . . . . . . . . . . . . 46
4.2 Ambiguity Function Unencoded LFM Pulse . . . . . . . . . . . . . . . . . . 47
4.3 Ambiguity Function Unencoded LFM Pulse (Top) . . . . . . . . . . . . . . 48
5.1 Time Plots of LFM and Encoded LFM (Real/Imag) . . . . . . . . . . . . . . 50
5.2 Power Spectral Density / Autocorrelation Plots - Constant Phase, 90 Deg . 51
5.3 Ambiguity Plot of Encoded Signal - 90 Degree, Constant Phase . . . . . . 52
5.4 Top View - Ambiguity Plot - 90 Degree, Constant Phase . . . . . . . . . . 53
5.5 Power Spectral Density / Autocorrelation Plots - Constant Phase, 20 Deg . 54
5.6 Ambiguity Plot of Encoded Signal - 20 Degree, Constant Phase . . . . . . 55
5.7 Top View Ambiguity Plot - 20 Degree, Constant Phase . . . . . . . . . . . 56
5.8 Autocorrelation of Barker Sequence (n=13, Rpt=4) . . . . . . . . . . . . . 57
5.9 Power Spectral Density / Autocorrelation Plots - n=13 Barker, 90 Deg . . 58
5.10 Ambiguity Plot of Encoded Signal - n=13 Barker, 90 Deg . . . . . . . . . . 58
5.11 Top View - Encoded Signal - n=13 Barker, 90 Deg . . . . . . . . . . . . . 59
5.12 Power Spectral Density / Autocorrelation Plots - n=13 Barker, 20 Deg . . 60
5.13 Ambiguity Plot of Encoded Signal - n=13 Barker, 20 Deg . . . . . . . . . . 61
5.14 Top View - Ambiguity Plot - n=13 Barker, 20 Deg . . . . . . . . . . . . . 61
6.1 Correlation of Parallel Channel M-Sequences . . . . . . . . . . . . . . . . . 65
6.2 Power Spectral Density / Autocorrelation Plots - BPSK, M-Sequence, 20 Deg 70
6.3 Ambiguity Plot of Encoded Signal - BPSK, M-Sequence, 20 Deg . . . . . . 71
6.4 Top View Ambiguity Plot - BPSK, M-Sequence, 20 Deg . . . . . . . . . . . 71
6.5 Power Spectral Density / Autocorrelation Plots - BPSK, M-Sequence, 90 Deg 73
6.6 Ambiguity Plot of Encoded Signal - BPSK, M-Sequence, 90 Deg . . . . . . 73
6.7  TopView Ambiguity Plot - BPSK, M-Sequence, 90 Deg  
6.8  Traditional QPSK Constellation (a) and Modified QPSK (b)  
6.9  PSD / Autocorrelation Plots - QPSK, M-Sequence, 15/45 Deg  
6.10 Ambiguity Plot of Encoded Signal - QPSK, M-Sequence, ±15/45 Deg  
6.11 Top View Ambiguity Plot - QPSK, M-Sequence, ±15/45 Deg  
7.1  Time Difference of Arrival  
8.1  Radar Pulse Interval  
8.2  Overview of Nonsynchronous Navigation Problem  
8.3  Ambiguous Nonsynchronous Geolocation  
8.4  Estimated (Top) and True (Bottom) Trajectories  
8.5  True and Estimated Trajectory Convergence  
8.6  Plot of Full and Expanded Error Norms After Convergence ≤ 2 Meters
List of Tables

2.1 Dilution of Precision Ratings ........................................ 21
4.1 LFM Pulse and Radar Carrier Parameters ......................... 46
8.1 Dilution of Precision Values at $P_0$ ................................. 109
Acknowledgment

I would like to express my sincere appreciation and thanks to my dissertation committee for their guidance as I pursued a goal I have sought for many years. In particular I would like to thank Dr Michael Wicks who had the initial idea for this research and worked with me painstakingly over great distance and through several medical complications he endured. Further, Dr Zhiqiang (John) Wu provided invaluable support and guidance throughout this research and kept me on track to finish. I would be very remiss if I did not highlight my committee chair, Dr Kuldip Rattan. Even though my Ph.D. course of study began in the controls field, he was accommodating as I moved to the topic of mixed-modulated waveforms with a navigation emphasis. In addition Dr Xiadong (Frank) Zhang, Dr Michael Bryant and Dr Robert Penno were gracious enough to serve on my committee throughout my rather long Ph.D. odyssey.

Finally, I especially want to acknowledge my wife Charlotte. Not many wives would be enthusiastic to have a husband begin an engineering doctoral program as he approached ”middle age”. However, throughout this journey she has been one my most enthusiastic supporters even on those dark days when I thought I would never finish. The completion of this Ph.D. program is due in large part to her love, patience, and sacrifice.
Dedicated to

My parents Donald and Mary Nowak who encouraged and emphasized the importance of education to their eight children and to my wife, Charlotte, whose love and support were instrumental in my achieving this goal.
Introduction

The proliferation of commercial wireless devices and increasing bandwidth needs due to high speed data transfer requirements have created increasing congestion of the electromagnetic spectrum. As demand increases, efforts are being made to find more efficient uses of the available spectrum. Among these approaches are research on dynamic spectrum access that makes use of ‘white space’ between signals and traditional orthogonal approaches such as time (TDMA), code (CDMA), or orthogonal frequency (OFDM) division multiplexing. With the evolution of digital signal processing another approach is mixed modulation through intended modulation on pulse (IMOP).

IMOP differs from other “orthogonal” methods in that the auxiliary message rides atop the primary signal. Previously, real-time dynamic modulation of the primary signal with analog components was difficult at best. In addition, there were obvious cross interference concerns that had less than desirable effects on both the primary and auxiliary signal. [3]. However, advances in both digital electronics and signal processing have opened the door to re-exploring IMOP as a means to increase spectral efficiency.

The contribution of this paper is to explore reduced magnitude phase shift keying IMOP using spread spectrum encoding of signals in order to transmit both radar and communications information in a single radar pulse. A secondary contribution is the devel-
opment and simulation of a non-synchronous navigation algorithm that can implement a
directional radar / navigation architecture using this mixed-modulated waveform in a vi-
able navigation architecture.

Although the auxiliary signal will be used for navigation in this paper, the ability to
introduce an auxiliary signal lends itself to several applications. For the congested spec-
trum environment, a low data rate message can be used to direct communications to other,
less congested areas of the spectrum (e.g. dynamic spectrum access). Further, the use of
small magnitude phase changes and spread spectrum bit encoding provides an inherent low
probability of intercept and detection (LPI/LPD) for an additional degree of security.

The second contribution is demonstrating, through Matlab simulation, the utility of
using the encoded auxiliary signal as a navigation adjunct for areas where navigation cov-
erage is degraded (e.g. urban canyons, mountainous terrains, or high latitude regions).
Figure ?? provides a very basic top level view of a ground-based architecture. In this case a
select number of “participating” or “cooperative” radars (pseudolites) have their waveforms
modified to encode a navigation signal. The minimum navigation message in the waveform
includes the radar’s location to discriminate the source of the pulse (i.e. transmitting radar)
as well as a reference to a common timing signal that is also referenced to some standard.

For the purposes of this research the underlying radar pulse will be assumed to be a
linear frequency modulated (LFM) pulse. The LFM pulse is well known and commonly
used in radar waveforms so its use does not diminish generality of the solution.
1.1 Previous Work

The use of radar waveforms to simultaneously conduct communications functions has been discussed for a number of years. A 1981 report by the Naval Research Laboratory categorized three approaches to multiplexing radar and communications signals into a waveform. [3] These approaches included combining the signal at two different frequencies (e.g. frequency division multiplexing), combining the signals at different times (e.g. time division multiplexing) or using the communications waveform as the radar waveform. In particular, the use of orthogonal frequency division multiplexing (OFDM) as a means to combine radar and communications [10][23][2] and radar and navigation functions [6] have been proposed and studied. In each of these cases sufficient spectral or timing separation between the original waveform and the encoded message is provided to ensure no interference between messages.

The use of integrating communications waveforms as part of the radar signal has not been aggressively pursued due to technological challenges and the impact on performance. In the past, slower electronic processing speeds and signal processing capabilities did not allow for ad hoc mixed modulation. More importantly, radar waveforms are designed to achieve specific range and Doppler sensing characteristics while communications signals vary with the information being transmitted. This can result in range-time sidelobes that degrade the performance of the radar and may not be easily filtered out. [3]

Since 1981, and particularly in the past decade, advances in processing speeds, electronics and digital signal processing have addressed many of the technological challenges previously encountered. Further, it is the hypothesis of this paper that the use of spread spectrum, phase shift keying with reduced phase magnitude changes will minimize the impact on the radar performance.
Recent work has outlined approaches to shape radar waveforms such that they meet performance specifications. [19] [20] [5] [11]. These approaches normally specify constraints on regions of the radar ambiguity function or power spectral density distribution in order to achieve desired range-Doppler or sidelobe performance measures in the waveform.

While work in the use of IMOP has been reported by Kowatsch [12][13] and Song [24], their focus was in the use of spread spectrum pseudonoise and LFM waveforms as a communications means and not as a joint radar/communication waveform. Therefore, the deleterious effect on the ambiguity function (and, correspondingly, the radar performance) of the LFM pulse was not considered. As a result, there were no limitations placed on the magnitude of the associated phase change and conventional binary phase shift keying (BPSK) was employed.

This paper explores the use of reduced magnitude phase changes to encode messages into the LFM waveform while preserving the radar performance characteristics of the radar. From a mixed-modulated signal perspective, this is accomplished by exploring constant phase shift, Barker codes and long duration M-sequences. In addition, by selecting preferred pair M-sequences it will be shown that two message channels can be encoded into a single pulse. Further, application of a mixed-mode radar/navigation signal will be demonstrated. As will be seen, the non-synchronous arrival time of the different radars’ navigation signals at the receiver requires the development of a revised geolocation algorithm. This new algorithm will be demonstrated via Matlab simulation.
1.2 Problem Description / Decomposition

Although the focus of this research is on exploring the possibility of employing in-band, mixed-modulation for a multi-mission capability, the implementation of this signal figures prominently into the research. Developing a mixed-modulated signal that does not have practical application is merely an academic exercise. Therefore, the research will consider both the waveform development and the use-case application.

First, a radar/navigation mixed-modulated signal must be developed and shown to meet performance specifications. In this first problem, the challenge is to find an IMOP set of metrics that provides adequate signal efficiency and a satisfactory bit error rate while minimizing radar performance degradation. In order to reduce the cross-interference effects on the communications signal on the radar performance and improve message recovery, this paper will explore several spread spectrum strategies including constant phase shift, Barker Codes, binary M-sequences and Gold Codes. Further, this research paper will explore the use of reduced phase magnitudes (e.g. $\phi_\delta \ll \frac{\pi}{2}$) in conjunction with longer bit-sample periods in order to reduce radar performance degradation. The hypothesis is that if the phase changes are small enough, subject to message performance at the receiver, the overall perturbation effects on the radar will be mitigated while providing enough signal energy at the receiver for navigation message recovery at an acceptable bit error rate (BER).

To evaluate radar performance autocorrelation and power spectral density (PSD) will be evaluated in terms of the radar ambiguity function. Radar performance depends on how the radar detects targets (autocorrelation and power spectral density) and overall performance in the face of pulse receipt timing differences and Doppler frequency shift (radar ambiguity function). Contrasting change in the ambiguity function provides a point of comparison between an unencoded LFM signal and a mixed-modulated, encoded LFM signal. The objective will be to minimize the difference between the two plots.
The second phase of the dissertation will address the application of this signal to a navigation problem. While GPS is the preferred choice for navigation, there are areas of the world (extreme northern and southern latitudes), man-made and natural obstacles (urban canyons, mountain valleys), and signal interference problems [9] that argue for an adjunct navigation method. Using signals of opportunity such as a hybrid radar/navigation signal may allow geolocation when GPS coverage is not adequate.

The main challenge with mixed modulation is the directional nature of the signal. Dedicated navigation systems are normally omni-directional and rely on simultaneous signal receipt from geographically separated satellites or pseudolites to calculate a position. However, for a directional signal most of the energy is concentrated in a narrow main beam. As a result, instead of simultaneous receipt of signals from several sources, a more realistic case is one where the vehicle receives navigation messages from different pseudolites with a substantial (several seconds or longer) time delay between messages. Current multilateration algorithms are not designed to handle this delay between signals. Implementing a navigation approach will require that the aircraft record message receipt times until it has the minimum number of signals to develop a solution and account for the movement of the vehicle between message receipt points. Such an approach will be developed and simulated in the navigation section of this thesis.

1.3 Outline

The paper is organized as follows. Chapter 2 provides background on foundational concepts used in both the signal development / analysis and navigation aspects of the research. This chapter begins by introducing the linear frequency modulated (LFM) radar signal.
Included in the chapter is a brief discussion on the energy and power signals to lay the foundation for an analysis of the energy contribution of the message phase discretes compared to the energy in the unmodified LFM waveform later in the paper. This is followed by a summary of the error / complementary error functions and radar ambiguity function. The radar ambiguity function, in particular, will be used to ascertain the effect of in-band phase discretes on radar performance. The chapter finishes with a review of geometric dilution of precision (GDOP). GDOP provides a non-dimensional quality metric of radio-based navigation solution quality based on the look angles from the user to the different transmitters.

Chapters 3 through 6 focus on development of the in-band navigation signals and measures of both communications (navigation signal) and radar performance. Chapter three introduces the mechanics of in-band modulation and the various encoding approaches considered. More importantly it uses the radar range equation to help determine required bit energy at the receiver. This equation is crucial as the principal means of minimizing radar performance degradation will be through the use of small magnitude phase changes. The use of small angle phase changes also requires an update of the bit signal-to-noise ratio to bit-error rate relationship. This update is necessary since the bit separation distance varies as a function of the phase angle. Chapter 4 then describes the radar performance metrics to be used in evaluating the impact of the phase discretes. Normally the auto-correlation function and power spectral density are used to help characterize performance. The auto-correlation speaks to how well the return pulse can locate the location of a return while the power spectral density helps highlight how much energy is outside the main beam of the radar pulse. Both of these concepts can be seen in the radar ambiguity function. The message and performance background is necessary to set the stage for chapters 5 and 6 that address single channel constant phase / Barker encoded phase discretes and channelized M-sequence / Gold Code encoding of the communications (navigation) message.
Chapters 7 and 8 concentrate on the practical implementation of a mixed-modulated directional signal to produce a navigation solution. Navigation systems have a single mission and are normally omni-directional in broadcast. This allows for simultaneity of message receipt at a particular epoch of time. This is demonstrated in signals like GPS or LORAN that allow for pseudorange algorithms (GPS) or time difference of arrival algorithms (LORAN) for geolocation.

Use of a combined surveillance radar / navigation signal does not provide for this simultaneity of signal receipt. In radar the strength of the navigation signal is tied to the directional nature of the radar. Chapter 7 develops a modified TDOA approach and a closed-form, iterative solution that in Matlab simulations shows good convergence between the true and estimated aircraft trajectories.

Chapter 9 provides a summary of key findings from this research and discusses areas for follow-on studies.

1.4 Mathematical Notation

In the discussion that follows, the following notation will be used:

Scalars: Scalars are represented by italic characters such as \(a\) or \(b\).

Vectors: Vectors are represented by bold, lower case letters such as \(\mathbf{a}\) or \(\mathbf{b}\). Vectors are assumed to be in column form unless specifically stated otherwise. Vector elements are represented with a non-bold, italicized, lower case letter with a subscript indicating the position of the element in the vector. (e.g. \(r_i\) indicates the \(i^{th}\) element of the \(\mathbf{r}\) vector). The exception to this will be the correlation and power spectral density vectors which will follow the literature’s standard usage of \(R\) and \(S\).
**Matrices:** Matrices are represented by bold upper case letters such as \( A \) or \( B \). The identity matrix is represented by the bold, capital letter \( I \), \( I \). Matrix elements are represented in a similar manner as is done with vectors. Non-bold, italicized, upper case letter with subscripts indicate the position of the element in the matrix. (e.g. \( A_{ij} \) indicates the \( i^{th} \) row and \( j^{th} \) column element of the matrix.

**Unary Matrix and Vector Operations:** Conventional notation will be used to indicate a unary operation on a matrix or a vector. The superscripts \( T \), \( H \), or \( * \) indicate either the real transpose, Hermitian transpose, or the complex conjugate of the matrix or vector, respectively. (e.g. \( A^H \) is the Hermitian transpose of matrix \( A \)).

**Estimated Variables:** Estimates of state vectors or random variables are denoted by a “hat” symbol. (e.g. \( \hat{a} \))

**Calculated and Measured Variables:** Measured variables normally include errors due to measurement noise and other un-modeled perturbations. These variables are identified by the tilde symbol (e.g. \( \tilde{a} \)).
Background

This chapter provides background on the fundamental concepts used in the development of the mixed-modulated waveform. It begins with a review of the digitally sampled, linear frequency modulated (LFM) waveform. The discrete form is required as digital signal processing enables the insertion of varying magnitude phase discretes. Next is a review of signal energy, power and power spectral density (PSD). The PSD of a signal describes how the power is distributed across the spectral bands. For the LFM waveform, most of this power is in the bandwidth of the LFM pulse. As phase changes are added, some of the power is distributed outside the main bandwidth into sides lobes. As will be seen, the objective is to reduce the amount of side lobe power/energy. A review of the error function and complementary error function follows. In particular, the error function will be used in establishing the required bit to noise ratio at the receiver in order to achieve a desired bit error rate (BER).

The chapter continues with an introduction to the radar ambiguity function. This function provides the effect of Doppler frequency shifts and time delays on the returned pulse on the matched filter response. Although the shape of the ambiguity function does not have a clear universal form, it can be tailored to meet desired performance objectives.

The chapter concludes with a review of geometric performance of dilution (GDOP) as it applies to the implementation of such a message waveform. GDOP provides a metric of
the accuracy of the measurement as a function of the geometry of the transmitters relative to the receiver. In essence GDOP is a measure of the covariance matrix of the multilateration algorithm.

### 2.1 Digital Linear Frequency Modulated Signal

The linear frequency modulated (LFM) pulse is one of the more common pulse compression techniques used in radar. Developed during WWII [15], the pulse linearly sweeps a frequency band of width $B$ during the pulse duration $t_p$. As a result the frequency of the pulse changes linearly from $0 \leq t \leq t_p$. Mathematically, this can be represented as

$$\omega(t) = \dot{B}_w t + f_0$$  \hspace{1cm} \text{where} \hspace{1cm} \dot{B}_w = \frac{B}{t_p} \tag{2.1}$$

Since frequency is the time derivative of the phase angle, the instantaneous phase angle of the LFM signal can be expressed as

$$\phi(t) = \int \left( \dot{B}_w t + f_0 \right) dt = \frac{\dot{B}_w t^2}{2} + f_0 t + \phi_I \tag{2.2}$$

where the constant of integration, $\phi_I$, is the initial phase angle which, for a phase coherent radar pulse, can be assumed to be zero without loss of generality.

In a digital radar system, this pulse is transmitted as a series of discrete samples in a radar pulse. For a pulse length of time $t_p$, the sampling frequency, $f_s$, determines the number of samples, $n_p$, in a transmitted pulse:

$$n_p = f_s t_p \tag{2.3}$$
or, in terms of time

\[ t_i = \frac{n_i}{f_s} \text{ where } 0 \leq t_i \leq t_p, \quad i = 1, 2, 3, \ldots, n_p \]  

(2.4)

From this the LFM instantaneous phase angle can be recast in terms of the sample number, \( n \), as opposed to time.

\[ \phi_L(n) = \frac{\dot{B}_w}{2} \left( \frac{n}{f_s} \right)^2 + f_0 \frac{n}{f_s} \]  

(2.5)

The resulting signal of the sampled LFM pulse can be expressed in exponential form:

\[ x_L(n) = e^{2\pi j \phi_L(n)} = e^{2\pi j \left[ \frac{\dot{B}_w}{2} \left( \frac{n}{f_s} \right)^2 + f_0 \frac{n}{f_s} \right]} \]  

(2.6)

or in terms of the signal’s in-phase and quadrature (I-Q) components:

\[ x_L(n) = \cos \left[ \frac{\dot{B}_w}{2} \left( \frac{n}{f_s} \right)^2 + f_0 \frac{n}{f_s} \right] + j \sin \left[ \frac{\dot{B}_w}{2} \left( \frac{n}{f_s} \right)^2 + f_0 \frac{n}{f_s} \right] \]  

(2.7)

### 2.2 Signal Energy and Power Spectral Density Functions

Periodic signals can be described as energy or power signals. If \( x(t) \) is a real valued signal describing the voltage over a resistance \( R \), a current, \( i(t) \), is generated as described by the equation \( i(t) = \frac{x(t)}{R} \). To normalize the power, a resistance of \( R = 1\Omega \) is assumed.[16] From Parseval’s theorem, the energy signal can be written as a function of time or as a function of frequency:

\[ E = \int_{-\infty}^{\infty} |x(t)|^2 \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(\omega)|^2 \, d\omega \]  

(2.8)

An energy signal is defined as one where \( |E| < \infty \). However, there exist classes of signals, including infinite ones, that do not satisfy this constraint. In these cases the power of the signal is computed. The power of a signal is the energy normalized by the signal period. It
can be expressed in the time domain as

\[ P = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 \, dt \]  
(2.9)

while the power can be expressed in terms of the power spectral density (PSD) as:

\[ P = \frac{1}{2\pi} \int_{-\infty}^{\infty} |S_x(\omega)|^2 \, d\omega \quad \text{where} \quad S_x(\omega) = \frac{|X(\omega)|^2}{T} \]  
(2.10)

For digital signals involving discrete time, the functions for energy and power are written as the sum of the individual discrete pulses that make up the signal, or

\[ E_x = \sum_{n=\infty}^{\infty} |x[n]|^2 \delta T \]  
(2.11)

while the power is expressed as

\[ P_x = \frac{1}{N} \sum_{n=0}^{N-1} |x[n]|^2 \delta T \]  
(2.12)

The single variable discrete autocorrelation function can be expressed as

\[ R_{xx}[m] = \sum_{n=0}^{N-1} x[n]x[n - m] \]  
(2.13)

while the power spectral density (PSD) of a vector is the Fourier transform of the autocorrelation vector or

\[ S_{xx}[k] = \sum_{m=0}^{N-1} R_{xx}[m] e^{-2\pi j \frac{mk}{N}} \]  
(2.14)
Combining eqns 2.13 and 2.14 produces

\[ S_{xx}[k] = \sum_{m=0}^{N-1} \left\{ \sum_{n=0}^{N-1} x[n]x^*[n-m] \right\} e^{-2\pi j \frac{mk}{N}} \text{ for } -\pi \leq k \leq \pi \]  

(2.15)

From a matrix perspective the PSD can also be expressed as

\[
R_{xx}(n) = \begin{bmatrix}
S^*_{N-1}(n) & S^*_{N-2}(n) & S^*_{N-3}(n) & \cdots & S_0 & \cdots & 0 & 0 & 0 \\
0 & S^*_{N-1}(n) & S^*_{N-2}(n) & \cdots & S_1 & \cdots & 0 & 0 & 0 \\
0 & 0 & S^*_{N-1}(n) & \cdots & S_2 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & \cdots & S^*_{N-1}(n) & S^*_{N-2}(n) & S^*_{N-3}(n) \\
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & S^*_{N-1}(n) & S^*_{N-2}(n) \\
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & S^*_{N-1}(n) \\
\end{bmatrix}
\]

(2.16)

### 2.3 Error / Complementary Error / Q-Functions

Gaussian distributed noise (AWGN) is, by definition, a random process with a known mean and variance. The stochastic properties of AWGN can be described in terms of a probability density function termed the Gauss error function. It is defined as

\[
erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt
\]

(2.17)

and has properties

\[
erf(-\infty) = -1, \quad erf(+\infty) = 1
\]
\[
erf(-x) = -erf(x), \quad erf(x^*) = [erf(x)]^*
\]
where the asterisk denotes complex conjugation. The **complementary error function** is
defined as

\[
erfc(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-t^2} dt = 1 - \text{erf}(x)
\]  

(2.18)

Note also that

\[
\frac{2}{\sqrt{\pi}} \int_{-\infty}^{x} e^{-t^2} dt = 1 + \text{erf}x
\]

Neither the error nor the complementary error functions have closed form solutions. Instead, the equations are solved through numerical methods and tabulated for lookup. In addition, most engineering software packages such as *Matlab* or *Mathematica* have a built-in function that can compute either function for a specific value of \(x\).

In bit error rate derivations a modified form of the complementary error function, the \(Q\)-function, often appears. The \(Q\)-function is defined as

\[
Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-\frac{z^2}{2}} dz
\]

(2.19)

Like the error and complementary error function, this equation does not have a closed form solution and must be solved numerically. However, through a change of variable the \(Q\)-function can be expressed in terms of the complementary error function:

\[
Q(u) = \frac{1}{2} \text{erfc} \left( \frac{u}{\sqrt{2}} \right)
\]

(2.20)

This form of the equation using the \(\text{erfc}\) function allows for use of existing tabular results or function calls in many mathematical/engineering software packages.


2.4 Ambiguity Function

The radar ambiguity function provides a measure of the response of a matched filter to a finite energy signal in the presence of a time delay, \( \tau \), and Doppler frequency shift, \( \nu \). The periodic ambiguity function can be expressed as \[15\]:

\[
|\chi_T(\tau, \nu)| = \left| \int_0^T u(t)u^*(t + \tau)e^{2\pi j\nu t}dt \right|
\] (2.21)

where the * indicates complex conjugation.

The variables \( \tau \) and \( \nu \) represent the mismatched time or frequency of the function. The magnitude of the function, \( \chi \), is plotted on the z-axis for various values of \( \tau \) and \( \nu \). When \( \tau = 0 \) the function is at matched range meaning that reduction in the magnitude of \( \chi \) for any values along that line are solely attributable to the Doppler frequency shift of the target. In this case the ambiguity function can be viewed as an autocorrelation function with a mismatched frequency. Conversely, when \( \nu = 0 \) the function is frequency matched indicating that the frequency of the matched filter and the Doppler frequency are the same. Values along the \( \nu = 0 \) line are essentially the auto-correlation function between the filter and signal.

Computation of the ambiguity function can be computationally intensive. Fortunately, the function can also be considered as the inverse Fourier transform of the product of Fourier transforms. Going back to the ambiguity function, it can be viewed as the correlation of \( u(\tau)e^{2\pi jf\tau} \) with \( \nu(\tau) \)

\[
\chi_T(\tau, \nu) = \left| \int_0^T u(t)u^*(t + \tau)e^{2\pi j\nu t}dt \right| = |u(\tau)e^{2\pi jf\tau} \otimes \nu(\tau)|
\] (2.22)
2.4.1 Ambiguity Function Properties

The ambiguity function possesses the following properties that are well known in the literature.

1. Maximum value of $\chi(\tau, \nu)$ is $\chi(0, 0)$

$$
\chi(\tau, \nu) \leq \chi(0, 0) = \left| \int_{-\infty}^{\infty} S(t)S^*(t)dt \right| \quad (2.23)
$$

2. The volume under $\chi(\tau, \nu)$ is constant value, $A$,

$$
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \chi(\tau, \nu)^2 d\tau d\nu = A \quad (2.24)
$$

3. $\chi(\tau, \nu)$ is symmetric with respect to the origin

$$
\chi(\tau, \nu) = \chi(-\tau, -\nu) \quad (2.25)
$$

4. The zero Doppler (e.g. $\nu = 0$) is the autocorrelation function

$$
\chi(\tau, 0) = \left| \int_{-\infty}^{\infty} S(t)S^*(t-\tau)dt \right| = |R(\tau)| \quad (2.26)
$$

5. The zero delay (e.g. $\tau = 0$) is the Fourier Transform of the the square of the complex envelope

$$
\chi(0, \nu) = \left| \int_{-\infty}^{\infty} S^2(t)e^{j2\pi\nu t}dt \right| \quad (2.27)
$$

Radar ambiguity functions are normally tailored to meet specific operating environments. For example, the radar algorithm and waveform may be optimized to have precise range resolution at the expense of Doppler resolution or vice-versa. As a result, there is no ideal
ambiguity function shape; instead it depends on the desired performance.

Since the waveform is optimized for performance over a specific area, the objective in introducing a mixed-modulated signal on top of the radar pulse becomes one of minimizing the perturbation to the existing radar ambiguity function while ensuring the communications, or in our case, the navigation message, meets its own performance specifications.

Additional properties of the ambiguity function relevant to the mixed-modulated signal will be addressed later in Chapter 4.

### 2.5 Geometric Dilution of Precision (GDOP)

The accuracy of a multilateration system is a function of the geometry of the transmitters (e.g. pseudolites). In an ideal case, the transmitters will be spatially separated in order to maximize the angular separation between the line-of-sight vector from the receiver to each transmitter. Unfortunately, this ideal case is seldom achievable. As the basis vectors to the transmitters begin to align, positioning errors can develop.

GDOP is a means to determine the quality of the navigation solution based on the location of the transmitters relative to the receiver. For satellite navigation with multiple visible satellites, it is often used in an a priori manner to determine the best subset of the satellites to use to derive a position. For terrestrial based navigation systems, GDOP can be mitigated by judicious placement of transmitters to provide navigation coverage over a particular region. However, in the case of satellite navigation, the problem is more dynamic due to the changing relative motion of the satellites. Yarlagadda [28] and Langley [14] provide an excellent discussion of GDOP. The following derivation is derived from Langley’s
The GDOP equation arises from the least squares solution to the geolocation problem using pseudoranges. The pseudorange problem is one finding the solution to the following equation

$$\Delta P = A\Delta x + e \quad (2.28)$$

where $\Delta P$ is the difference between measured corrected and estimated pseudoranges, $A$ is an $n \times 4$ matrix of partial derivatives of the pseudoranges linearized about an estimated position, and $\Delta x$ is the unknown position. With the exception of the time difference of arrival method, this set of equations is normally solved iteratively. The unknown position can be solved using the following equation

$$\Delta x = (A^TWA)^{-1} A^T W \Delta P \quad (2.29)$$

where $W$ is a weighting matrix that characterizes the weighting of the measurements on the solution. GDOP arises from the covariance of the estimated position, $\Delta x$. Using the expectation function, $E$, we see

$$cov(\Delta x \Delta^T x) = E \left[ (A^TWA)^{-1} A^T WC_{\Delta P} [(A^TWA)^{-1} A^T W]^T \right] \quad (2.30)$$

where $C_{\Delta P} = \Delta P \Delta P^T$ is the covariance matrix of the position estimates. If we assume that all measurements are weighted equally such that $W = I$ and complete the matrix operations, equation 2.30 reduces to

$$C_{\Delta x} = (A^T_{cov(\Delta P)}A)^{-1} \quad (2.31)$$

The covariance function, $cov(\Delta P)$, has the elements of the dilution of precision. It consists
of a matrix whose diagonal terms are the error variances created based on the geometry of
the transmitters (satellites) relative to the receiver as is seen here:

\[
\text{cov}(\Delta P) = \begin{pmatrix}
\sigma_x^2 & \cdot & \cdot & \cdot \\
\cdot & \sigma_y^2 & \cdot & \cdot \\
\cdot & \cdot & \sigma_z^2 & \cdot \\
\cdot & \cdot & \cdot & \sigma_{ctB}^2
\end{pmatrix}
\] (2.32)

From the diagonal elements of the matrix GDOP can be computed as

\[
GDOP = \sqrt{\frac{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_{ctB}^2}{\sigma_{uere}^2}}
\] (2.33)

In the equation, \(\sigma_{uere}\) is the “user equivalent range error” and consists of all errors associated with the characteristics of the navigation system in use. For satellite systems this includes cumulative errors that arise from measurement noise, atmospheric effects, etc.

The GDOP result is a non-dimensional, scalar quantity that provides a measure of ‘goodness’ of the transmitter constellation relative to the user’s position. For GPS navigation, the ideal constellation configuration would have one GPS satellite at zenith and the others distributed uniformly around and just above the horizon (e.g. a tetrahedron). In their paper on user position and GDOP, Dutt, et al [26] provide the following commonly accepted qualitative measures of performance for GDOP.

GDOP is a comprehensive metric of quality but it is only one measurement of dilution of precision. In addition, the covariance signal terms can also provide the following quality metrics on difference aspects of the derived position:

Position Dilution of Precision, PDOP: \[
\sqrt{\frac{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_{ctB}^2}{\sigma_{uere}^2}}
\]
Table 2.1: Dilution of Precision Ratings

<table>
<thead>
<tr>
<th>DOP Value</th>
<th>Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ideal</td>
</tr>
<tr>
<td>2-4</td>
<td>Excellent</td>
</tr>
<tr>
<td>4-6</td>
<td>Good</td>
</tr>
<tr>
<td>6-8</td>
<td>Moderate</td>
</tr>
<tr>
<td>8-20</td>
<td>Fair</td>
</tr>
<tr>
<td>20-50</td>
<td>Poor</td>
</tr>
</tbody>
</table>

Horizontal Dilution of Precision, HDOP: $\sqrt{\sigma_x^2 + \sigma_y^2}/\sigma_{uere}$

Vertical Dilution of Precision, VDOP: $\sigma_z/\sigma_{uere}$

Time Dilution of Precision, TDOP: $\sigma_{ct_B}/\sigma_{uere}$
Communications Message Performance

Fundamental to implementing a successful mixed-modulated signal is identifying a communication encoding scheme that provides adequate bit-throughput at a satisfactory bit error rate while also not adversely affecting the performance of the primary signal, in our case, the radar.

In essence this research will explore two variables, or degrees of freedom, to encode a message. The first is the basis vector of the encoding scheme. Three different approaches will be presented - constant phase change, Barker Code, and binary M-sequences / Gold Codes. As the name implies the constant phase change uses a constant phase angle change to encode a single bit of ‘1’ or ‘0’. Barker Codes incorporate a spread spectrum approach to the communications message and help improve bit recovery at the receiver. The final approaches, binary M-Sequences / Gold Codes are considered together due to their cross-correlation characteristics which allow for “channelization” of multiple messages thereby increasing throughout.

The second variable is the use of reduced phase angle changes along with the encoding approach to reduce the impact on the radar pulse while preserving the communications message performance. The final variable is the type of phase shift keying used. Initially binary phase shift keying (BPSK) will be used in conjunction with encoding approach and the smaller phase angles. In an effort to increase bit-throughput, a modified variation of
quadrature phase shift keying (QPSK) will be implemented. However, as will be seen QPSK does not significantly improve bit throughput due to its negative effect on radar performance.

We begin this chapter with a section on encoding approaches for in-band modulation. This will be followed by sections on factors that drive message performance and then a survey of possible encoding schemes.

### 3.1 In-Band Mixed Modulation

Mixed-modulation is a form of intentional modulation on pulse (IMOP) where the carrier wave is used to transmit another message much like an AM or FM radio broadcast. In the case of mixed-modulated radar and communications, the radar pulse is used as the carrier wave while the communications (navigation) message is the secondary/auxiliary message. This is depicted in Figure 3.1

![Figure 3.1: Intentional Modulation of Radar Pulse with Communications Message](image)

In the graphic an LFM signal is generated and sampled to produce a vector that is a digital representation of the LFM signal. The vector is composed of the instantaneous phase...
angles of the LFM pulse. As separate communications message is also digitally sampled with phase changes, ±φδ. The LFM signal and digit message vectors are element-wise multiplied which imposes a phase shift at each sample instance. The combined digital, mixed-modulated, radar-communications signal is then converted to an analog signal and unconverted to a higher power carrier wave.

For this research, the radar pulse is a linear frequency modulated signal and acts as the carrier wave. The radar signal or pulse, x_r, is produced by a digital waveform generator and sampled producing a vector of length n_p. The navigation signal, x_m, has the same sample rate as the radar signal and employs phase shift keying using some type of encoding (e.g. constant phase shift, Barker Code, etc.) of the phase shift, φδ. For traditional BPSK, the phase shift is φδ = ±π/2 in order to maximize bit separation between the two types of bit values, ‘1’ and ‘0’. In this research, however, reduced magnitude values of φδ in conjunction with various encoding schemes will be explored.

The mixed modulated signal can be expressed as the complex (I-Q) or exponential vector element-wise product of the LFM signal, φ_L(n), and the communications signal, φ_M(n). If the LFM and message vectors are in complex form, the resulting transmitted comm-radar modulated intermediate frequency (IF) signal vector, x_T, is

\[ x_T(m) = e^{jφ_M(m)} ⊙ e^{j2πφ_L(m)} \text{ for } m = 1, 2, \ldots, n \]  \hspace{1cm} (3.1)

where ⊙ is the element-by-element multiplication of the two vectors. The communications message bit-sequence can be done using either sequential or channelized encoding as illustrated in Figure 3.2. In the graphic each block or ‘sample’ represents a sample instance of the digitized baseband navigation message vector, φ_M, while the length of a single bit has a bit-sample length of n_λ. ‘No message’ intervals, n_I, may be used to both improve mes-
sage recovery or to provide an additional degree of freedom in meeting radar performance constraints. The value of a single bit is represented by a series of discrete samples (depicted as $+$ signs) of length $n_\lambda$ while the zero-message interval between one bit and another is $n_I$.

![Diagram of sequential and channelized encoding](image)

**Figure 3.2: Representative Message Bit**

Initially, binary phase shift keying (BPSK) method was selected in order to reduce ambiguity problems during message recovery. Since the phase discretes are of short duration and have a relatively low SNR, BPSK should provide a better navigation message throughput and lower bit error rate. BPSK encoding of the navigation message is done by modifying elements of $\phi_M(m)$ as follows:

$$
\phi_M(m) = \begin{cases} 
-\phi_\delta b, & \text{if } = 0 \\
+\phi_\delta b, & \text{if } = 1 \\
0, & \text{if no message bit}
\end{cases}
$$

where $\phi_\delta$ is the phase shift corresponding to the bit value and $b$ is the basis vector of the encoding scheme.

Sequential bit encoding encodes a single bit at a time by appending the basis vector modified phase change, $\phi_M(m)$, of each subsequent bit-sequence in order. Therefore the maximum bit content in a single pulse is a function of the length of the radar pulse and
the bit sample length. The vector consists of a sequence of small angle phase changes that employ a phase shift bit encoding scheme with each bit separated by an optional interval of zeroes (e.g. no message).

The basis function, \( \phi_M(m) \), for this approach can be expressed using the following step function

\[
\phi_M(m) = \sum_{i=1}^{\beta} \phi_i \left( u[n - n_i] - u[n - (n_i + n_{\lambda})] \right) b(n - n_i + 1) \tag{3.2}
\]

In this equation \( n_i \) is the location of the start of phase bit \( i \) in the vector \( \phi_M(n) \) and \( n_{\lambda} \) is the number of samples over which \( \phi_i \) is held constant to encode a single ‘1’ or ‘0’ bit. The vector \( b \) is the basis function of the encryption scheme used (constant phase change, Barker Coded phase change, Gold Code etc.) and/or any bit pulse shaping (e.g. Hamming, Hanning, Blackman windows etc) that may used to encode a single bit, \( \phi_i \) in order to improve recovery of the baseband message.

Channelized encoding also uses a smaller phase change but is done by selecting two basis function sequences that minimize cross correlation between each other. The resulting navigation message vector, \( \phi_M(n) \) consists of the sum of the phases discretes. Since values of a particular bit sequence will cancel or add to other contributions, it should be possible to keep the overall phase discrete change at a single sample within a tolerable phase shift.

### 3.2 Message Performance

The use of both reduced magnitude phase changes and constant phase change or spread spectrum basis functions requires careful consideration of the bit error rate (BER). Since
each bit is encoded using a bit-sample sequence based on a particular basis function, the individual bit sequence must be long enough to provide adequate bit energy at the receiver through integration.

The effect of using smaller phase changes with BPSK as compared to traditional $\pi/2$ BPSK is that the energy distance between a ‘1’ and ‘0’ bit is reduced. Assuming additive white Gaussian noise, the reduced magnitude phase change requires a higher bit SNR to achieve an $E_b/N_0$ that provides the same bit error rate. The communication signal is a one-way signal from transmitter to receiver so the one-way radar range equation can be used to derive an equation that determines the required M-sequence length. The following subsections will address both the BER for reduced phase margin and required bit-sample length, respectively.

### 3.2.1 Reduced Phase Margin Bit Error Rate

BPSK implementations typically maximize the distance between the ‘1’ and ‘0’ bits to improve detection and reduce bit error rates during message recovery. Haykin shows [8] that the probability of bit error (or BER) for conventional BPSK in terms of a Q-function as:

$$P_{e0} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$  \hspace{1cm} (3.3)

BER can also be expressed as the distance between the different bit values measured in the I-Q plane. The general expression for the BER of a BPSK signal with bits separated by $d_{min}$ is the following Q-function

$$[BER] = Q\left(\frac{d_{min}/2}{\sigma}\right)$$  \hspace{1cm} (3.4)
For smaller phase magnitudes the BER must recomputed to account for the smaller bit-to-
bit distance.

A ‘1’ or ‘0’ bit is defined by a phase discrete value of $+\phi_{\delta}$ or $-\phi_{\delta}$, respectively. This is depicted in Figure 3.3. To ensure adequate bit energy and improved reception at the receiver each bit phase value is multiplied by a predetermined basis vector consisting of either a constant or bipolar (-1, 1) sequence. For the bipolar sequence this will be a pseudorandom generated sequence. Despite the sign (+/-) of the phase changes in a bit sequence, the relative position (and distances) of each phase discrete sample value remains the same.

\[
\phi_1 = \phi \delta \\
\phi_2 = -\phi \delta
\]

This is depicted in Figure 3.3. To ensure adequate bit energy and improved reception at the receiver each bit phase value is multiplied by a predetermined basis vector consisting of either a constant or bipolar (-1, 1) sequence. For the bipolar sequence this will be a pseudorandom generated sequence. Despite the sign (+/-) of the phase changes in a bit sequence, the relative position (and distances) of each phase discrete sample value remains the same.

\[
d_{min} = 2\sqrt{E_b \sin \phi}
\]

Figure 3.3: Reduced Phase Angle Bit Distance

The bit separation is a function of the magnitude of the phase difference, $\phi_\delta$ and the signal bit energy, $E_b$: $d_{min} = 2\sqrt{E_b \sin \phi_\delta}$. In this case the signal energy is a function of the M-sequence bit duration (e.g. bit-sample length, $n_\lambda$). In addition, the signal is assumed to be affected by additive white Gaussian noise (AWGN) with $\sigma^2 = \frac{N_0}{2}$. Substituting into the BER equation

\[
[BER] = Q \left( \frac{2\sqrt{E_b \sin(\phi_\delta)/2}}{\sqrt{\frac{N_0}{2}}} \right) = Q \left( \frac{\sqrt{2E_b \sin(\phi_\delta)}}{\sqrt{N_0}} \right)
\]

Figure 3.4 provides a graphic depiction of BER as a function of $\frac{E_b}{N_0}$ and the magnitude
of the phase discrete angle, $\phi_\delta$. For a traditional BPSK signal with 180 degrees of bit separation (e.g. $\phi_\delta = \pm \frac{\pi}{2}$), the $Q$-function reduces to the standard value of $Q \left( \sqrt{\frac{2E_b}{N_0}} \right)$. The equation can be expressed in terms of the required bit energy to signal ratio as a function of desired BER and phase magnitude change:

$$\frac{E_b}{N_0} = \frac{1}{2} \left[ \frac{Q^{-1}[\text{BER}]}{\sin \phi_\delta} \right]^2$$

(3.6)

Using either equation 3.6 or the plot, the required SNR for a specific BER and phase angle can be determined. This, in turn, drives the next question of bit-sample length in order to achieve the desired SNR value.

Figure 3.4: Reduced Phase Angle Bit Error Rate
3.2.2 Radar Range Equation - Bit Energy

With the required $E_b/N_0$ value established, the required sample length can be determined. The LFM signal is a constant envelope waveform so the energy in one bit can be expressed as

$$E_b = \frac{V^2}{2} T_b = P_R T_b$$  \hspace{1cm} (3.7)

where $T_b$ is the time duration of the bit phase discrete, $V$ is the magnitude of the received LFM waveform, and $P_R$ is the bit power at the receiver relative to a $1\Omega$ resistance. Since the communication message is digitized, the pulse consists of a series of discrete time values, or samples, whose time duration depends on the sampling rate, $f_s$, of the analog-digital converter.

To achieve enough bit energy at the receiver, the duration of the phase change for a single bit must extend over a set period of time (e.g. set number of samples). The time duration, $T_b$, of a bit is merely the product of the number of samples and the sample duration, $\frac{1}{f_s}$, of each sample. Therefore, the energy expression for a single transmitted bit can be rewritten as

$$E_T(\lambda) = P_T \frac{n_\lambda}{f_s}$$  \hspace{1cm} (3.8)

where $n_\lambda$ is the number of samples that make up the bit, $f_s$ is the sampling rate, and $G_T$ is the transmitter gain.

The effect of distance on the received energy must also be considered since radar energy falls off as a square of the distance. Unlike the radar pulse, the communication message is received and processed at the receiver so the one-way radar range equation, or Friis equation, can be used.
\[ E_R = \frac{E_T G_T A_e}{4\pi r^2 L} \]  

(3.9)

In this equation \( E_R \) is the received energy at the aircraft, \( G_T \) is the transmitter aperture gain, \( r \) is the radar to aircraft distance, and \( L \) are energy losses. Further, the effective antenna aperture can be expressed as

\[ A_e = \frac{G_R \lambda^2}{4\pi} \]  

(3.10)

where \( G_R \) is the receiver aperture gain and \( \lambda \) is the signal wavelength. Incorporating both \( A_e \) and the equation 3.8 for the bit energy, \( E_T \), into the one way range equation allows rewriting of the bit energy, \( E_b \), at the receiver as

\[ E_b(n_\lambda, r) = \frac{P_T n_\lambda G_T G_R \lambda^2}{(4\pi r)^2 L f_s} = \left[ \frac{P_T G A_e}{(4\pi)^2 L f_s} \left( \frac{n_\lambda}{r^2} \right) \right] \]  

(3.11)

The first bracketed term in the equation is determined by the performance characteristics of the radar transmitter while the second parenthetical term determines the energy of the communication message pulse at the receiver as a function of sample length and range.

Message bit detection and probability of misidentifying a message bit depends on signal-to-noise ratio. For message recovery the bit energy to noise power spectral density is crucial in ensuring an acceptable BER. For thermal noise, the noise power spectral density, \( N_0 \), can be expressed as

\[ N_0 = \kappa T_0 F \]  

(3.12)

where \( F \) is the receiver noise factor, \( \kappa \) is Boltzmann’s constant \((1.3806488 \times 10^{-23} m^2kgs^{-2}K^{-1})\) and \( T_0 \) is the antenna aperture noise temperature in degrees Kelvin (normally assumed to be 290 degrees).
Rewriting the energy equation in terms of required bit SNR and sample length:

\[
\frac{E_b}{N_0} = \frac{P_T n_\lambda G_T G_R \lambda^2}{(4\pi r)^2 L f_s \kappa T_0 F} \geq [SNR]_{req} \tag{3.13}
\]

\[
n_\lambda \geq \frac{[SNR]_{req} (4\pi r)^2 L f_s \kappa T_0 F}{P_T G_T G_R \lambda^2} \tag{3.14}
\]

This constraint provides a lower limit on the number of samples required to achieve a bit energy level that ensures an adequate BER at the receiver.

### 3.3 Encoding Approaches

As was previously mentioned, short duration phase changes will be used to encode the message into the navigation signal that will be subsequently modulated on the LFM radar pulse. This section details several approaches to encoding the phase discrete.

#### 3.3.1 Constant Valued Phase Changes

As the name implies, this method uses a constant phase change for a period of time, \( \frac{n_\lambda}{f_s} \), to relay a single bit. A constant phase shift to encode a bit is the simplest to implement but does not have desirable characteristics for post-receipt de-correlation. Using a matched filter, the correlation result of a constant phase shift is a triangular function while, ideally, a near-impulse “spike” is preferred. In addition, as will be seen, a constant value phase change does not allow for multiple, channelized signals in a pulse.

#### 3.3.2 Barker Codes

Barker codes are a biphase code used in many radar applications. [21] The term bi-phase refers to the fact that there are only two phase states, +1 and -1. Barker codes are de-
signed so that, when autocorrelated, the peak-sidelobe ratio is \( N : 1 \) where \( N \) is the length of the code. As an example, a Barker code, \( C \), of \( n=13 \) is the following sequence \( C=[1\ 1\ 1\ -1\ -1\ 1\ 1\ -1\ -1\ 1\ -1\ 1] \). A plot of the autocorrelation of \( C \) is shown in Figure 3.5

![Auto-Correlation of Barker Code for n=13](image_url)

**Figure 3.5: Autocorrelation of \( n=13 \) Barker Code**

For instances where \( \tau = 0 \) the value of the function varies between 0 and 1. When there is zero lag the function spikes to a value equal to the length of the Barker code. The combination of a prominent indication of correlation and the high peak-to-sidelobe ratio makes the Barker code a leading candidate to encode communications information. It also lends itself to radar applications due to its excellent correlation properties. To encode the navigation message, the Barker biphase sequence is used as a multiplier to encode a predetermined phase shift, \( \phi_\delta \).

Barker codes offer flexibility in implementation. For example, if the length of the Barker sequence is less than the required number of samples to encode a bit, each element of the Barker sequence can be repeated to expand the bit-sample length so as to meet the
required length. Different length Barker sequences can also be nested to increase length. As an example, a 52-length Barker sequence can be constructed from two Barker sequences of length 4 and 13. This is done by taking one of the sequences, say the n=13 length sequence, and then using each element of this sequence as a multiplier for the n=4 sequence. Combining the 13 n=4 sequences provides a 52-element Barker code with the same high in-phase and low out-of-phase cross correlation properties as a traditional Barker sequence.

In addition, Barker sequences can involve non-integer and complex alphabets. These have been addressed in a paper by Golomb and Schotz [7] who demonstrated that generalized Barker sequences consisting of an alphabet of complex numbers of length \( k \) satisfy the autocorrelation constraint \( |C(\tau)| \leq 1 \) for \( \tau \neq 0 \).

### 3.3.3 M-Sequences / Gold Sequences

Maximum-Length sequences (M-sequences), and their counterpart Gold Codes, are pseudorandom binary sequences that can be produced by the use of binary linear feedback shift register (LFSR) systems using modulo-2 arithmetic and an \( \text{xor} \) function for addition. More conventionally, M-sequences are produced using primitive polynomial functions of degree \( m \) with a binary-valued, \((0, 1)\), seed of the same length. [1] The primitive polynomial (or LFSR implementation) will generate a repeating binary sequence of length \( 2^m - 1 \). [18] Individual M-sequences exhibit high auto-correlation values while minimizing out-of-phase (e.g. when \( \tau \neq 0 \)) auto-correlation values. Further, “preferred pairs” of M-sequences using different starting seeds to generate a pair of m-sequences also exhibit excellent auto-correlation with reduced cross correlation between the bit sequences.

Gold Codes are another variation of M-sequences that are the modulo-2, element-wise product of two M-sequences. Like M-sequences, Gold code sequences can also be
produced from a primitive polynomial that is the result of modulo-2 multiplication of the two m-sequence primitive polynomials. Gold Code sequences also provide the same auto-correlation and slightly better cross-correlation performance.

To illustrate the correlation properties of M-Sequences fact, two primitive polynomials of degree 5 were used to generate two M-sequences of length 31. The modulo-2 primitive polynomials are \( x^5 + x^2 + 1 \) and \( x^5 + x^3 + 1 \) while the same seed value of \([0 0 0 1]\) is used to start each sequence. The resulting respective m-sequences are:

\[
\begin{align*}
[1000100101100111100011011010] \\
[10001010111011000111100110100]
\end{align*}
\]

![Figure 3.6: Autocorrelation and Cross-Correlation of Generated M-Sequences](image)

After converting the binary (0,1) sequence to a bi-phase (-1,1) sequence, the two sequences are cross-correlated. Figure 3.6 depicts the result of the autocorrelation of the \( x^5 + x^2 + 1 \) generated M-sequence in blue while the cross-correlation of the two generated M-sequences for \( x^5 + x^2 + 1 \) and \( x^5 + x^3 + 1 \) is shown in red. Longer sequence lengths
provide more robust peak-to-sidelobe performance.

M-sequences have been heavily studied for their auto and cross correlation properties. Welch [27] established a lower bound on the maximum cross-correlation of signals in 1974 followed by Massey [17] who provided a variation Welch’s solution. In Welch established a minimum cross-correlation value of

\[(c_{max})^{2^k} \geq \frac{1}{M - 1} \left[ \frac{M}{\binom{L+k-1}{k}} - 1 \right] \quad (3.15)\]

where \(c_{max}\) is the minimum cross-correlation value, \(M\) is the set of vectors, and \(k\) is a positive integer. In addition, Sarwate provided insight on bounds of cross and auto correlation functions as a tradeoff between the two. In his IEEE correspondence [22] he provided his theorem on the relationship between the maximum periodic auto-correlation magnitude, \(\theta_c\), and the maximum out-of-phase periodic auto-correlation, \(\theta_a\):

\[\left( \frac{\theta_c^2}{N} \right) + \frac{N - 1}{N(K - 1)} \left( \frac{\theta_a^2}{N} \right) > 1 \quad (3.16)\]

More recently Tirkel has provided additional insight on cross-correlation of m-sequences [25]. While admitting that analytic techniques to compute the cross-correlation of m-sequences has proven to be intractable, he exhaustively examined cross-correlation patterns of all sequences up to lengths of \(2^{18} - 1\) for some type of predictable pattern. Unfortunately, for each of these cases an analytic solution to the maximum out-of-phase autocorrelation value has not been derived.

A variation of M-sequences are Gold Codes. Gold Codes are generated using pairs of M-sequences with a length of \(2^n - 1\) elements that are combined via an exclusive-or operation. Like M-sequences, the codes have a reduced cross-correlation interaction but
excellent auto-correlation properties.

A principal benefit of Gold Codes and M-sequences in this research is to take advantage of their low cross-correlation and linearity properties in order to encode channelized message streams in a single vector. This can be illustrated in the recovery of a bit sequence from a vector with multiple message sequences.

![Message Correlation with Gold Codes 1, 2, 3](image)

Figure 3.7: Three-Sequence Gold Code Correlation

In the following example, a series of bits were encoded into a single message vector using three distinct Gold Codes with a 50% overlap (time delay) between bit sequences in order to differentiate the bit sequences. The resulting series of phase changes was added together into a single combined baseband message. Figure 3.7 illustrates the result of passing this baseband message through a filter bank of the matched filters for the three Gold
Code basis vectors (sequences). The plots show both a strong autocorrelation and also a low cross correlation with bits encoded using other Gold Code sequences.

The mathematics behind this characteristic is straightforward. Considering the case where $k$ Gold Code bit-encoded sequences are overlapped into a single baseband message, the bit-stream encoded by Gold Code 1 can be recovered as follows,

$$
\phi_{M1}(n) = \sum_{m=0}^{\lambda} \phi_{M}(m)G_1(M+n)
$$ (3.17)

$$
\phi_{M1}(n) = \sum_{m=0}^{\lambda} \{[G_1(n) + G_2(n - \alpha) + \ldots + G_{k}(n - (k-1)\alpha)] G_1(M+n)\} \quad (3.18)
$$

$$
\phi_{M1}(n) = \sum_{m=0}^{\lambda} \{G_1(n) + G_1(m+n)\} + \sum_{m=0}^{\lambda} \{[G_2(n - \alpha) + \ldots + G_{k}(n - (k-1)\alpha)] G_1(M+n)\} \quad (3.19)
$$

The summation term on the right represents the cross-correlation of Gold Code sequence 1 with other Gold sequences. Since these have lower cross-correlation values, the more dominant auto-correlated values will stand out.

The Gold Code example is illustrative only. For his research the focus was on M-Sequence pairs to permit channelization of the message.

### 3.3.4 Polyphase Sequences

Considered, but not pursued, are polyphase sequences. Polyphase sequences are best defined as complex mathematical sequences that use phase shifts to represent bit values. Like M-sequences, polyphase sequences exhibit high autocorrelation values while minimizing
out-of-phase autocorrelation results. To keep the magnitude of the function at unity, a primitive root of unity, $W_n$, normally forms the basis of a polyphase sequence:

$$W_n = e^{-j2\pi r/n}, \quad j = \sqrt{-1}$$  \hspace{1cm} (3.20)

where both $r$ and $n$ are integers. From variations on this basis function, various polyphase sequences can be developed including the matrix-based periodic polyphase code and Frank-Zadoff-Chu sequences.

The drawback of polyphase sequences are two-fold. First, unlike M-sequences, polyphase sequences do not possess the same cross-correlation properties as preferred pair M-sequences. Therefore, channelization is not possible. A second, more concerning problem is the magnitude of the phase changes in a polyphase sequence. The goal of this research is to minimize perturbation to radar performance using the radar ambiguity function as a metric of performance. Therefore, the focus is on reduced angle phase changes while a true polyphase sequence normally employs regulatory spaced bits over $2\pi$ in the I-Q space. This will result in large phase changes for certain bit values that will affect the ambiguity function result.

### 3.4 Summary

This chapter has introduced in-band, mixed-modulation message performance and message encoding approaches. A significant factor in the coding scheme is the required message energy at the receiver to achieve the required BER. For high power radars and limited areas of interest, constant phase and Barker codes are satisfactory. However, their use precludes channelization due to poor cross-correlation properties of the bit steams during
de-correlation of the base band message.

To address the more general application where longer bit-sample lengths are required due to lower power radars and large volumes of interest, M-sequences and Gold Codes will receive the focus of attention in this dissertation.
Radar Performance

The objective of the in-band modulation approach is to ensure each signal’s mission performance is not adversely affected. Performance of the communications message was considered in terms of adequate bit energy at the receiver to achieve the required BER. We now turn to a different set of tools to evaluate radar performance.

Radar performance metrics are based on the ability of the radar to accomplish its mission of detection and/or tracking of targets. Without oversimplifying the process, radar returns are processed through a matched filter that consists of the time-reversed complex conjugate of the transmitted pulse. The ideal LFM autocorrelation response for the target-reflected pulse is a high peak consistent with aircraft position and low sidelobe values. For a mixed-modulated signal, as the magnitude of the in-band phase changes increase, spurious sidelobes appear that can be misidentified as possible returns or can hide returns from actual aircraft. These errors are normally classified as Type I (false positive) or Type II (missed detection).

The power spectral density (PSD) is the Fourier Transform of the auto-correlation result. In essence, it provides a depiction of the distribution of the signal power in terms of frequency. For the unencoded LFM wave, the PSD has a maximum main lobe that is the same as the bandwidth of the LFM pulse. Either side of this bandwidth energy drops off rapidly. As phase discrete changes are introduced and allowed to grow, power moves from
the mainlobe of the LFM pulse to adjacent (side) lobes. Our objective is to try to minimize these power losses from the main lobe.

While both the autocorrelation and PSD functions will be considered, there is another function that provides a more comprehensive measure of radar performance - the radar ambiguity function. When plotted, the radar ambiguity function provides a graphic depiction of the radar’s response in the face of delay from the matched filter range and/or frequency shift due to Doppler effects. In addition, along the zero Doppler plot line, the ambiguity function reduces to become the autocorrelation function of the signal and its matched filter. This chapter continues the background chapter’s discussion on the function and provides an example of the ambiguity function for the basic LFM signal using an unmodulated LFM pulse.

4.1 Ambiguity Function

The ideal (and hypothetical) ambiguity function has a ‘thumbtack’ appearance at the origin (e.g. \( \tau = 0 \) and \( \nu = 0 \)) for a Dirac delta impulse signal. In this situation the radar can provide accurate target location as long as there is no delay or Doppler frequency shift. More realistically, radars need to have some tolerance to time delays and Doppler shifts so the ideal ‘thumbtack’ response is not necessarily a desired matched filter response.

For air-to-air tracking of targets, Doppler shift will always be present unless the target’s radial motion toward the radar is zero. Also, modern radars normally sample the return in the time domain based on various ranges, or range bins. After transmitting the radar pulse, the radar goes into a period of quiescence and ‘listens’ for the return pulse. The radar samples the receiver for any returns based on round-trip times that correspond to
various range bins. If there are two targets at different ranges their respective return should show up during the receiver sample time corresponding to the range of the target.

The challenge with this approach is that the target may not be exactly centered in the range bin of interest. As a very basic example, assume a radar has 20 range bins of 4NM distance each going from 20 NM to 100 NM. If a target is at 22 NM it will fall in the center of the first range bin (20-24 NM). If, however, the target is at 23.5 NM there will be a slight time delay from the center of the range bin, or the matched range.

In a similar manner, Doppler shift due to radial speed between the target and the radar can manifest themselves as the equivalent of delay in time. The ambiguity function provides a three dimensional depiction of the signal energy due to time and Doppler delays from the expected values for a specific range / Doppler combination.

### 4.2 Phase Message Effects on Ambiguity Function

The LFM ambiguity function has the time domain form of

\[
\chi(\tau, \nu) = \left| \int_0^T e^{j2\pi L(t)} e^{-j2\pi L(t-\tau)} e^{j2\pi \nu t} dt \right| \quad (4.1)
\]

where \( \phi_L(t) \) is the instantaneous phase angle at time \( t \) and, for the LFM signal, \( \phi_L(t) = \frac{B_w}{2} t^2 + f_0 t \). Since we are concerned with discrete values this can be expressed as

\[
\chi(p, \nu) = \left| \sum_{n=0}^{N-1} e^{j2\pi \phi_L(n)} e^{-j2\pi \phi_L(n-p)} e^{j2\pi \nu \frac{n}{f_s}} \left( \frac{1}{f_s} \right) \right| \quad (4.2)
\]

The mixed-modulated LFM signal modulates a baseband message vector, \( \phi_m(n) \).
φ_m(n) consists of a series of pseudorandom binary basis vectors, each one multiplied by a small angle phase change corresponding to a ‘1’ or ‘0’ bit. The resulting transmitted intermediate frequency (IF) LFM is the element wise multiplication (Hadamard product) of the discrete LFM vector and the message vector such that

\[ x_T(n) = e^{j2\pi(\phi_L(n)+\phi_m(n))} \quad (4.3) \]

The resulting ambiguity function is

\[ \chi(p,\nu) = \left| \sum_{n=0}^{N-1} e^{j2\pi[\phi_L(n)+\phi_m(n)]} e^{-j2\pi[\phi_L(n-p)+\phi_m(n-p)]} e^{j2\pi\frac{\nu}{f_s}} \left( \frac{1}{f_s} \right) \right| \quad (4.4) \]

where \( p \) has been used for the delay in lieu of \( m \) to avoid confusion with the message subscript. The equation can be rewritten as:

\[ \chi(p,\nu) = \left| \sum_{n=0}^{N-1} e^{j2\pi[\phi_L(n)-\phi_L(n-p)]} e^{j2\pi[\phi_m(n)-\phi_m(n-p)]} e^{j2\pi\frac{\nu}{f_s}} \left( \frac{1}{f_s} \right) \right| \quad (4.5) \]

At this point further reduction is not possible since \( \phi_m(n) \) is a pseudorandom signal of small phase angles of \( \pm \phi_\delta \). However some basic analysis can be made on the equation.

When the correlation delay of the signal and the matched filter is 0 (e.g. \( p=0 \)) the \( \phi \) terms in the exponent cancel resulting in a value of 1 for the leading term multiplied by the Doppler shift exponential term and the bit sample interval.

\[ \chi(0,\nu) = \left| \sum_{n=0}^{N-1} e^{j2\pi\nu\frac{n}{f_s}} \left( \frac{1}{f_s} \right) \right| \quad (4.6) \]

Similarly, when the Doppler frequency shift is zero (\( \nu=0 \)) the Doppler exponential term is one and the ambiguity function reduces to the cross-correlation of the matched filter and
the signal.

\[
\chi(p, 0) = \left| \sum_{n=0}^{N-1} e^{j2\pi[\phi_L(n) - \phi_L(n-p)]} e^{j2\pi[\phi_m(n) - \phi_m(n-p)]} \left( \frac{1}{f_s} \right) \right| \tag{4.7}
\]

Focusing on the second exponential term in the summation, \(\phi_m\) is a function of a binary pseudorandom or constant value basis vector and the magnitude of the phase shift, \(\pm \phi_\delta\). At any instance of correlation lag, \(p\), \(\phi_m\) is one of three values: \(+\phi_\delta\), 0 or \(-\phi_\delta\). It’s apparent that traditional BPSK values of \(\phi_\delta = \pm \frac{\pi}{2}\) that maximize \(d_{\text{min}}\) have a greater impact on the elements of the ambiguity function at zero Doppler.

Based on this, two observations come to light. First, the use of smaller magnitude phase changes will reduce the effect of out-phase correlation factors when the matched filter and signal at \((n - p)\) and \(n\) do not cancel thereby producing a \(\pm 2\phi_\delta\) result. Second, the use of encoding schemes that minimize out-of-phase \((n \neq 0)\) correlation should reduce message correlation effects to the ambiguity function.

The performance of different encoding schemes will be discussed in upcoming chapters. For now, the remainder of this chapter, the focus will be on the PSD, auto-correlation, and ambiguity function of the baseline LFM pulse. This result will be the points of comparison for the other encoding methods and bit phase angle changes.

### 4.3 LFM Signal Baseline Performance

To provide an illustration of the radar performance metrics, the linear frequency modulated (LFM) pulse with the following parameters was considered. The pulsewidth and sample rate provide a digitized signal of 5001 samples that can be used to encode a message sequence. Figure 4.1 provides a depiction of the standard autocorrelation and PSD plots for
Table 4.1: LFM Pulse and Radar Carrier Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse Width (PW)</td>
<td>10 $\mu$sec</td>
</tr>
<tr>
<td>Bandwidth (BW)</td>
<td>5 MHz</td>
</tr>
<tr>
<td>Center Frequency ($f_c$)</td>
<td>2.5 MHz</td>
</tr>
<tr>
<td>Sampling Frequency ($f_s$)</td>
<td>$5 \times 10^8$ Hz</td>
</tr>
<tr>
<td>Signal Amplitude, $A_c$</td>
<td>1</td>
</tr>
<tr>
<td>Radar Power, $P_t$</td>
<td>1 kW</td>
</tr>
<tr>
<td>Antenna Gain, $G_T$</td>
<td>2000</td>
</tr>
<tr>
<td>Antenna Temperature, $T_0$</td>
<td>290 K</td>
</tr>
</tbody>
</table>

the basic unencoded LFM pulse.

![PSD Plot of LFM Signal](image1)

![Auto-Correlation of LFM Signal](image2)

Figure 4.1: PSD/Autocorrelation Unencoded LFM Pulse

The LFM PSD plot shows that much of the power is in the same frequency range as
the bandwidth of the LFM signal, 5 MHz. Outside this region, the power drops off rapidly. The autocorrelation has a defined peak that corresponds to the the target position in time relative to a matched range sample. Offsets of this peak from $\tau=0$ indicate the target range difference (distance) from the expected matched range. More importantly, the autocorrelation shows low sidelobe values ($\approx 13$ dB down from the main lobe to the first side lobe).

Figure 4.2: Ambiguity Function Unencoded LFM Pulse

Those attributes can also be found in the radar ambiguity plot in Figure 4.2. Looking along the $\tau$ axis at $\nu=0$, we see the autocorrelation of the LFM signal. Figure 4.3 provides a top view of the LFM ambiguity function. The ridge extending along the $\nu$ axis is indicative of a Doppler tolerant signal. The angular slant of this ridge in the $x - y$ plane shows that the Doppler offsets from the expected Doppler shift manifest themself as an apparent error.
in target range.

Figure 4.3: Ambiguity Function Unencoded LFM Pulse (Top)
Single Channel Modulation

The most direct implementation of in-band modulation is to apply a constant phase change with a magnitude corresponding to a particular bit value for a specified period time. A secondary approach is to use a message encoding technique such as Barker Codes or pseudorandom sequences to help improve message recovery at the receiver. This chapter will include an example of each of these single channel approaches using traditional binary phase shift keying (BPSK) of the message vector.

5.1 Constant Phase Angle Modulation

Encoding of an individual bit is done by shifting the phase magnitude of \( n_\lambda \) samples of the radar pulse by \( \pm \phi_\delta \). In order to demonstrate the effect of large phase magnitude changes on the radar performance, a phase change of \( \phi_\delta = \pm \frac{\pi}{2} \) was used along with the LFM pulse parameters from Table 4.1. Using the equations for minimum \( \frac{E_b}{N_0} \), equation (3.6), and minimum bit-sample length, equation(3.14), as well as \( 10^{-6} \) BER, the minimum bit-sample length was determined to be as small as one sample. This is due to a combination of the large phase change, the power of the carrier radar signal (1 kW), and the limited employment range of 500 KM. Despite this, and to have a consistent bit-sample length for comparison, a bit sample length of \( n_\lambda = 50 \) was used. The encoded message bit sequence consisted of alternating 1s and 0s was encoded in the message vector as this is more stress-
ing on both the time and, as will be seen, the PSD plots. Figure 5.1 depicts the history of the basic LFM pulse (top), followed by time domain plots of the encoded LFM signal with the message vector. Since the message vector rotates the LFM vector orthogonal to the basic LFM pulse, time plots of both the real and imaginary components of the encoded LFM signal are shown.

![Time Domain Plot of Basic LFM Waveform (Real)](image)

![Time Domain Plot of Encoded LFM Waveform (Real)](image)

![Time Domain Plot of Encoded LFM Waveform (Imag)](image)

Figure 5.1: Time Plots of LFM and Encoded LFM (Real/Imag)

Since the pulse sample-length of the LFM pulse is 5001 samples, this provides a \( \approx 100 \) bits per pulse of information content. The approximation caveat is used since some implementations may include a preamble set of bits such as parity bits in order to define message
start points and to improve error correction.

![PSD Plot of Nav Signal, Constant (Deg=90)](image1)

![Auto-Correlation of Nav Signal, Constant (Deg=90)](image2)

Figure 5.2: Power Spectral Density / Autocorrelation Plots - Constant Phase, 90 Deg

While communications BER performance is easily assured, the effect on radar performance with large phase shifts is more troubling. Figures 5.2, 5.3, and 5.4 depict the significant impact of large phase shifts on the radar performance. In this case the peak-to-sidelobe ratio is only -4.26dB. Since autocorrelation is used for target detection, the existence of significant sidelobes in the autocorrelation plot predict poor detection performance. Compared to the unencoded LFM ambiguity function plot of Figure 4.2, the encoded signal ambiguity plot has poorer Doppler tolerance and sidelobes that can mask targets or cause false positive detections.
Figure 5.3: Ambiguity Plot of Encoded Signal - 90 Degree, Constant Phase
5.1.1 Constant Small Magnitude Phase Angle

A second simulation was run using a smaller phase shift (e.g. ±20 degrees) to explore the effect on radar performance. For this phase shift value, a minimum bit-sample length, \( n_\lambda \), of 37 samples was determined to be required for a 10^{-6} BER using the same radar parameters. Again, to assure communication performance and to have consistency with the previous example, \( n_\lambda = 50 \) was used. Figures 5.5 and 5.6 depict the results of the smaller phase angle shift.
Figure 5.5: Power Spectral Density / Autocorrelation Plots - Constant Phase, 20 Deg

In these plots the radar performance appears to be closer to the unencoded LFM waveform. The autocorrelation plot of Figure 5.5 has a peak-to-sidelobe ratio of -15.6dB does not exhibit the same high autocorrelation sidelobe problems as was the case with a ±90 degree phase shift. Further the radar ambiguity function in Figure 5.6 shows Doppler tolerance approximate to the unencoded LFM signal.

Recovery of the constant phase angle message is done by demodulating the received message from both the carrier wave (radar frequency) and the IF LFM pulse to baseband. While correlation of the baseband message with the constant phase basis vector can be employed, this produces a triangular correlation result. Instead a ‘sample and hold’ approach is used where $n\lambda$ bit samples (corresponding to the length of the constant phase basis vector) are sampled and integrated (summed). A decision is made as to whether the resulting value
is positive or negative. By increasing the bit-sample length, message recovery is more robust in that small errors in determining actual message start and endpoints can be mitigated.

Figure 5.6: Ambiguity Plot of Encoded Signal - 20 Degree, Constant Phase
5.2 Barker Encoded Phase Angle Modulation

Barker Codes at small phase angles allow for use of spread spectrum of the message. One of the main benefits in using Barker coding is that recovery of the message can be accomplished using cross-correlation of the baseband message vectors with the basis vector of the Barker code. Due to the low out-of-phase and high in-phase correlation properties of Barker codes bit determination with cross-correlation of the baseband message vector with the basis vector of the Barker sequence is straightforward and the accuracy requirements of where a message starts and stops inherent in the sample-and-hold approach are minimized. In addition, increasing the length of the Barker sequences increases the in-phase to out-of-phase cross-correlation ratio thereby improving bit detection and bit error rate.

Figure 5.7: Top View Ambiguity Plot - 20 Degree, Constant Phase
To extend the Barker sequence, each element of the Barker code was repeated four time, producing a sequence length of 52 elements. The Barker sequence was converted to a binary (-1,1) vector that acts as the basis vector for the signal. A bit value of ‘1’ or ‘0’ is encoded by multiplying this basis sequence by either $+\phi_\delta$ or $-\phi_\delta$. Figure 5.8 shows that the autocorrelation properties of repeating the sequence in this manner still produces similar autocorrelation characteristics with the peak-to-sidelobe ratio remaining constant.

For comparison, a Barker sequence of length 52 was first simulated using a standard $\phi_\delta = \frac{\pi}{2}$. Figures 5.9 and 5.10 depict the PSD / Autocorrelation and ambiguity function plots, respectively.
Figure 5.9: Power Spectral Density / Autocorrelation Plots - n=13 Barker, 90 Deg

Figure 5.10: Ambiguity Plot of Encoded Signal - n=13 Barker, 90 Deg
Figure 5.11: Top View - Encoded Signal - n=13 Barker, 90 Deg

As opposed to the constant phase using $\frac{\pi}{2}$ phase magnitudes, the autocorrelation and ambiguity function show major consequences. This is due due to the continuous large magnitude changes inherent in a Barker code. For a constant phase shift, the phase change was help constant over the bit-sample period. However, in the case of the Barker code, the phase is chaining by $\pm \pi$ radians with each change in Barker element value. This manifests itself in the autocorrelation result which has a -4dB difference between the main lobe and the peak sidelobe of the pulse. As is evident, the possibility of missed targets or false targets is greatly increased which argues against the use of large phase angle Barker codes for message encoding.

The Barker code approach was repeated using a 20 degree phase change and the same 52 bi-sample Barker sequence. The results for the PSD / autocorrelation and ambiguity function are shown in Figures 5.12 and 5.13, respectively. The autocorrelation exhibits low
sidelobes and a high in-phase correlation value while ambiguity function performance is close to the unencoded LFM version.

Figure 5.12: Power Spectral Density / Autocorrelation Plots - n=13 Barker, 20 Deg
Figure 5.13: Ambiguity Plot of Encoded Signal - n=13 Barker, 20 Deg

Figure 5.14: Top View - Ambiguity Plot - n=13 Barker, 20 Deg
5.3 Summary

For small phase angles both the constant phase shift and spread spectrum encoding using Barker sequences provide satisfactory BER performance for the message and neither approach seriously degrades radar performance as seen from the radar ambiguity function. Perhaps the main advantage of the Barker code is that the message recovery through cross-correlation with the message basis vector reduces the probability of error associated with a constant phase angle encoding. This is because the start and ending points of the message must be known with some degree of error. Despite this the constant phase angle approach appears to be an easier method to implement.

There are, however, several drawbacks to Barker codes over a constant phase change. Barker codes require a number of phase change reversals in each encoded bit making it potentially more difficult to implement than a constant phase change. In addition, it is more sensitive to the use of large phase angles for bit encoding.

An additional drawback to both of these approaches is the limitations on bit throughput per pulse. Since constant phase and Barker Codes have poor cross-correlation properties, we are limited to sequentially transmitting one bit at a time. Other spread spectrum methods such as M-sequences and Gold codes have excellent autocorrelation characteristics while also demonstrating superb (e.g. low) cross-correlation properties. The use of encoding with M-sequences and Gold Codes with an eye toward multichannel communications will be taken up in the next chapter.
Channelized Modulation

As was seen in the previous chapter, the use of constant phase angle changes to encode a bit stream while minimizing the effects on the radar waveform is feasible for small phase angle changes. However, the limitation with this approach is that only a single bit stream (channel) can be encoded at a time.

Another approach that offers to improve bit density per pulse is the use of binary M-sequences, and their relative, Gold Code, to encode the bit stream. Both M-sequences and Gold Codes have the added characteristic of offering low cross-correlation properties for preferred pair M-sequences. With channelized pairs of M-sequences this can essentially double the bit throughput.

This chapter begins by describing a BPSK encoded channelized M-sequence bit streams to include the mathematics on the bit recovery for two linearly added message phase vectors. It continues with a discussion of radar performance using channelized sequences. The chapter goes on to consider QPSK encoding of preferred pair sequences in an attempt to increase information content in the bit stream. Finally a brief discussion follows on the feasibility of higher order M-ary PSK bit encoding.
6.1 Channelized BPSK Encoding

Encoding of an individual bit is done by multiplying the M-sequence basis vector of a particular channel by the appropriate phase change corresponding to a ‘1’ or ‘0’. Each message vector, $m_1(n)$ or $m_2(n)$, is constructed by sequentially adjoining the phase modified bit sequences.

If preferred M-sequence pairs are selected so as to minimize cross correlation between overlapping bit sequences, a channelized approach may be used to simultaneously transmit two (or more) message sequences. The resulting communication message vector, $\phi_{MT}(n)$, consists of the sum of the message vectors where some sample values of a particular bit sequence may cancel or add with other message channels. By using a reduced phase magnitude it should be possible to keep the overall phase discrete change for any single sample within a tolerable phase shift.

Normally this vector will be modulated onto an LFM signal and up-converted to a higher power carrier signal. At the receiver this process is reversed and the baseband message is recovered. For this example the baseband message was sent through a filter bank consisting of matched filter basis functions for each M-sequence used to encode the signal.

To illustrate the dual channel approach, two 7-bit M-sequence were generated using primitive polynomials of $x^7 + x + 1$ and $x^7 + x^2 + 1$ with common initial seeds of $[0 0 0 0 0 1]$. The resulting M-sequence basic vectors were of length 127 samples. A single ‘1’ bit was encoded using each M-sequence with a time delay between the channels. The two encoded bit streams were then added to establish the baseband message vector.
The result of the cross-correlation of the two M-sequence basis vectors, as well as the cross correlation of each M-sequence filter with the message vector are shown in Figure 6.1. As can be seen the cross-correlation between the two messages is low (top plot - note vertical scale difference) as compared to the in-phase correlation result of the lower two plots. Also clearly shown is the time delay between the two encoded ‘1’ bits.
6.2 Binary M-Sequence Message Recovery

The pertinent question in channelizing two bit streams with preferred pair M-sequences is whether the individual bit streams can be recovered with a high degree of confidence. Before beginning with this discussion, a review of notation is in order. The bi-phase M-sequence basis vector for an individual channel uses the notation of $m_c$, where, for the two channel case, $c = 1, 2$. A phase message vector for one channel, $\phi_{Mc}$, consisting of $\beta$ bits is composed of $\beta$ adjoined phase modified M-sequences. For Channel 1 the message phase vector is:

$$\phi_{m1}(m) = \sum_{i=1}^{\beta} \left\{ \phi_\delta(i) m_1 (u[m - m_i] - u[m - (m_i + n_\lambda)]) \right\}$$  \hspace{1cm} (6.1)

where $n_\lambda$ is the number of elements of the M-sequence basis vector. For a BPSK encoded message, the complex conjugate of the bit basis vector is also used as the matched filter for bit recovery by the receiver.

The final message vector of phase changes is the complex vector sum of the two phase message vectors (channels) that is modulated on the LFM signal. This message vector, $\phi_{MT}$, can be expressed as

$$\phi_{MT} = \phi_{M1} + \phi_{M2}$$  \hspace{1cm} (6.2)

With this review of notation we turn toward the recovery of the message vector.

After down-converting from the higher power carrier wave, the received digitized signal, $x_r$, is a combination of the transmitted signal, $x_T$, with Doppler and noise. It can be expressed as:

$$x_r(n) = V \cdot e^{j[2\pi\phi_L(n) + \phi_{MT}(n) + \phi_d(n)]} + w(n)$$  \hspace{1cm} (6.3)

where $\phi_L$ is the instantaneous phase contribution of the LFM signal (in Hz) while $\phi_{MT}(n)$ is the M-sequence message vector, $\phi_d$ and is the phase angle changes due to Doppler shift,
and $w(n) = w_i(n) + jw_Q(n)$ is the complex white Gaussian noise, respectively. Practically, the Doppler phase shift can be assumed to be constant over the pulse receipt period since the change in $\nu$ over the time period of a microsecond pulse is negligible.

The received message is conventionally demodulated to baseband using a replica of the LFM signal. Since it is unlikely that a local oscillator or the stored on-board LFM signal will identically match the transmitted version, this demodulation introduces a phase error, $\phi_e$. This results in the following phase angle, $\phi_D(n)$, for the exponent of the demodulated message signal:

$$\phi_D(n) = \phi_{MT}(n) + \phi_d + \phi_w(n) + V_e$$  \hspace{1cm} (6.4)

where $\phi_w$ is the phase angle contribution from the noise $w(n)$.

To illustrate recovery of the multi-channel message, we can consider recovering bits from the first channel of the multi-channel vector. In this case the encoded message vector is demodulated to a baseband vector consisting of the sum of the phase angles of the encoded pulses. To recover the first bit of the message from channel 1, for example, the baseband message vector is cross-correlated with the respective message basis vector, $\phi_{B_1}$. This is shown as:

$$R_{\phi_{MT}\phi_{B_1}}(\tau) = \sum_{n=1}^{n_\lambda} [\phi_{MT}(n) + \phi_d + \phi_N(n) + \phi_e] \phi_{B_1}^*(n + \tau) =$$

$$\sum_{n=1}^{n_\lambda} [(\phi_{M_1}(n) + \phi_{M_2}(n)) + \phi_d + \phi_w(n) + \phi_e] \phi_{B_1}^*(n + \tau)$$  \hspace{1cm} (6.5)

where $n_\lambda$ is the number of samples that make up a single M-sequence encoded bit.

For instances where the filter and message are out-of-phase ($\tau \neq 0$) the equation (6.5)
reduces to the sum of four separate cross-correlation operations:

\[
R(\tau \neq 0) = \sum_{n=1}^{n_\lambda} \phi_{M1}(n)\phi_{B1}^*(n + \tau) + \sum_{n=1}^{n_\lambda} \phi_{M2}(n)\phi_{B1}^*(n + \tau) + \\
\phi_d \sum_{n=1}^{n_\lambda} \phi_{B1}^*(n + \tau) + \phi_e \sum_{n=1}^{n_\lambda} \phi_{B1}^*(n + \tau) + \sum_{n=1}^{n_\lambda} \phi_{w}(n)\phi_{B1}^*(n + \tau)
\]  

(6.6)

The first term is the autocorrelation of the Channel 1 phase angle message vector with its own basis vector. As previously discussed, this autocorrelation term is minimized for out-of-phase, \(\tau \neq 0\), terms. Similarly, the second term is the cross-correlation of the Channel 2 message phase vector with the basis vector for Channel 1. These cross correlation terms are also minimized with M-sequences. As an aside, since the M-sequences will be known \textit{a priori}, a cross correlation can be run to determine the max out-phase correlation value in order to establish a floor for the out-of-phase terms.

The last three terms are the cross-correlation of the M-sequence basis vector with the Doppler shift, phase errors from demodulation, and noise, respectively. Again, both Doppler and the demodulation phase errors will likely be constant over the duration of the received pulse. The cross-correlation function has the effect of multiplying the constant Doppler or phase error by the sum of the M-sequence elements. Since the sum of the M-sequence elements is small, the contribution to the overall correlation value is also small. The last term is the product of two random processes, noise and a pseudorandom binary sequence. The sum of the product of an AWGN distributed random process and a pseudorandom process is also small.
When the basis vector and message 1 are in phase \( \tau = 0 \) the equation (6.5) becomes

\[
R(\tau = 0) = \sum_{n=1}^{n_\lambda} \phi_{M1}(n)\phi_{B1}^*(n) + \sum_{n=1}^{n_\lambda} \phi_{M2}(n)\phi_{B1}^*(n) + \\
\sum_{n=1}^{n_\lambda} \phi_d\phi_{B1}^*(n) + \sum_{n=1}^{n_\lambda} \phi_\epsilon\phi_{B1}^*(n)
\] (6.7)

Proceeding as before, the last three terms are small for the reasons already outlined. However, the first term has two possible outcomes. The first is if the bit is a ‘1’ (e.g. \( \phi_\delta > 0 \)). In this case the element-wise multiplication of the message vector and the message basis vector is

\[
\sum_{a=1}^{n_\lambda} \phi_\delta^2 = n_\lambda \phi_\delta^2
\] (6.8)

The second instance is where a ‘0’ bit has been encoded so that \( \phi_\delta \) of the encoded phase sample is of opposite sign (+/-) to the corresponding element of the message basis vector used for cross-correlation. In this case

\[
\sum_{a=1}^{n_\lambda} \phi_\delta(-\phi_\delta) = -n_\lambda \phi_\delta^2
\] (6.9)

In short, autocorrelation of the message vectors with the appropriate basis vector results in a clear positive/negative result corresponding to a ‘1’ or ‘0’. The magnitude of the result is a function of the length of the bit-sample length.

### 6.3 Channelized Message Radar Performance

A similar approach was taken with the BPSK for binary M-Sequences as was used in the previous chapter that considered constant phase angle and Barker coded phase shifts. The radar and LFM performance parameters outlined in Table 4.1 were again used. The main
difference in this approach is that two independent message streams (channel) were encoded into the pulse. This effectively doubled the pulse-bit throughput.

![PSD Plot of NAV Signal, M-Sequence (Deg=20)](image1)
![Auto-Correlation of NAV Signal, M-Sequence (Deg=20)](image2)

Figure 6.2: Power Spectral Density / Autocorrelation Plots - BPSK, M-Sequence, 20 Deg

Proceeding as before, we begin with a 20 degree phase shift required a minimum \( n_\lambda \) of 37 bit-samples to achieve a \( 10^{-6} \) BER. This required a primitive polynomial of order six for a \( n_\lambda = 63 \) and results in a potential bit density of 78 bits per channel or an overall bit throughput of 156 bits per pulse. From a previous exhaustive simulation of binary M-sequences of order six, the following two preferred pair primitive polynomial and seed combinations were used: \( x^6 + x^5 + x^2 + x + 1 \) \( / [1 0 0 0 0 1] \) and \( x^6 + x^5 + x^3 + x^2 + 1 \) \( / [0 0 0 0 0 1] \). Figures 6.2 and 6.3 provide depictions of the PSD / autocorrelation and ambiguity function plots, respectively.
Figure 6.3: Ambiguity Plot of Encoded Signal - BPSK, M-Sequence, 20 Deg

Figure 6.4: Top View Ambiguity Plot - BPSK, M-Sequence, 20 Deg
Beginning with Figure 6.2, the PSD plot shows improvement in the out-of-band energy levels as compared with using constant phase changes or a Barker approach. Similarly, the autocorrelation plot has low sidelobes and high magnitude when in-phase. The reason for this reduction is due in part to the nature of the message vector. Since the cross-correlation between two preferred pairs is low, summing the two channelized message vector allows for both constructive and destructive changes to the phase shift that is imposed on the primary message waveform.

Referring to the ambiguity plot, Figure 6.3, the impact of the two channel approach causes some reduction in Doppler tolerance, but the shape of the waveform is nearly consistent with the unencoded LFM plot.

For completeness a 90 degree M-sequences was also considered. Although the use of a 90 degree phase shift allowed for a shorter minimum bit-sample length, the same $n_\lambda = 63$ was again used to standardize the comparison with the 20 degree M-sequence plots. Figures 6.5 and 6.6 that follow have the PSD / autocorrelation and ambiguity function plots, respectively.
Figure 6.5: Power Spectral Density / Autocorrelation Plots - BPSK, M-Sequence, 90 Deg

Figure 6.6: Ambiguity Plot of Encoded Signal - BPSK, M-Sequence, 90 Deg
Contrasting with the Barker autocorrelation for 90 degree phase magnitude, the M-sequence has a better peak-to-side lobe ratio of -13dB which is on par with a normal LFM signal. Further, the signal power is spread almost uniformly. The ambiguity function shows excellent range discrimination at low Doppler changes but is less Doppler tolerant than a LFM signal. One conclusion that can be drawn from this is that by changing the magnitude of the phase angle one can modify the Doppler characteristics of the ambiguity function.

A logical extension of channelized BPSK M-sequences is to consider the use of higher order phase shift encoding schemes to improve bit throughput. For this, quadrature phase shift keying (QPSK) will be considered next.
6.4 Channelized QPSK Encoding

In considering QPSK encoding the challenge is applying the same small phase angle perturbation to encode a bit stream on the primary waveform. Figure 6.8a depicts a traditional Gray coded QPSK constellation that maximizes bit distance and, therefore, minimizes bit error rate. The effect of large phase changes inherent in this implementation significantly impacts the shape of the radar ambiguity function plot. Therefore, an alternative QPSK implementation is depicted in the adjacent figure 6.8b.

In this geometry the four phase angles are $\pm \phi_\delta$ and $\pm 3\phi_\delta$. The reason for the difference in phase change magnitudes is to assure a constant $|2\phi_\delta|$ separation between adjacent bits in the constellation. By standardizing the bit-to-bit separation distance, the effective BER for each bit location is the same.

![Normal QPSK Bit Distribution](image1)

![Modified QPSK Bit Distribution](image2)

Figure 6.8: Traditional QPSK Constellation (a) and Modified QPSK (b)

To examine the effects of a QPSK encoding scheme on channelized M-sequences the same set of primitive polynomial and seeds used in the BPSK M-sequence example were used to encode two channels consisting of a sequence of alternating ‘1’ and ‘0’ bits. To
keep the overall phase shift small, a phase shift vector of $[-30 - 15 + 15 + 30]$ was used. This required recomputing the minimum bit-sample length for the smaller phase angle of 15 degrees. Using the same radar constraints, the minimum $n_\lambda$ is 112 requiring a primitive polynomial of order seven. Two 7-bit M-sequence were generated using the same primitive polynomials of $x^7 + x + 1$ and $x^7 + x^2 + 1$ with common initial seeds of $[0 0 0 0 0 1]$ discussed earlier in this chapter. Since the M-sequences are of length 127, and since QPSK encodes two bits per sequence, the channelized QPSK can encode 156 bits in our implementation. This is identical to the two channel BPSK method discussed earlier.

![PSD Plot of QPSK Encoded Nav Signal, (Phase= +/- 45, 15)](image1.png)

![Absolute Value of Auto-Correlation of QPSK Encoded Navigation Signal](image2.png)

Figure 6.9: PSD / Autocorrelation Plots - QPSK, M-Sequence, 15/45 Deg

Figures 6.9 and 6.10 depict radar performance items of the PSD, autocorrelation, and ambiguity function plots. In examining the plots the autocorrelation and PSD plots are fairly consistent with the two-channel BPSK case. However, the ambiguity plots shows
decreased doppler tolerance as compared to two channel BPSK.

Increasing the phase angles to help reduce the minimum bit sample length in order to preserve a satisfactory BER is not necessarily an option. Since the higher magnitude phase angle for QPSK is three times the size of the smaller one, moving to a 20 degree phase angle would also required a 60 degree phase change. The larger phase change creates the same performance pressure on the ambiguity function seen in earlier examples.

Figure 6.10: Ambiguity Plot of Encoded Signal - QPSK, M-Sequence, ±15/45 Deg
Aside from accepting reductions in communications performance (e.g. BER) or radar performance, the other option is to adjust the overall architecture. For example, increasing radar power from 1 kW to 5 kW reduces the minimum sample length for a $\pm 15$, $\pm 45$ degree QPSK phase shift from 112 to 23. Similarly, reducing communications coverage from 500 KM to 300 KM for a 1 kW radar reduces minimum bit-sample length from 112 to 41 at the same $\pm 15$, $\pm 45$ degree phase angle settings. This in turn would will increase bit throughput.

6.5 M\text{-}ary Constellations

The use of higher order PSK encoding schemes appears somewhat limited as the phase angle requirements eventually grow and affect the performance of the radar pulse. In an
effort to constrain phase magnitude values, consideration was given to quadrature ampli-
tude modulation (QAM). QAM involves both phase angle and magnitude (signal energy) changes to define eight unique constellation points and allows for transmitting three bits at a time. Implementing this involves discrete phase angle values and using different the bit-sample lengths to modulate the magnitude of the bit energy (e.g. amplitude).

One drawback to QAM is that the amplitude portion needs to be done by varying the bit-sample length. That is, a different order primitive polynomial is required. Low cross-correlation values arise from preferred pair sequences of similar order primitive polynomials or linear feedback shift registers (LFSR). The same cannot be said for M-sequences of different lengths generated by different order primitive polynomial or LFSRs. As a result, the low cross correlation characteristic of M-sequences cannot be leveraged in a QAM-type architecture.
Radio Based Navigation Problem

Previous chapters in this dissertation focused on the development of a mixed-modulated waveform that satisfied both navigation (communication) and radar performance requirements. In that development, the power of the main transmitter played a significant role in the minimum bit-sample length to encode a single bit. For a typical radar this transmitted power is confined to a narrow main beam (main lobe). Outside the main lobe, radiated power falls off rapidly with the exception of the back lobe and, depending on radar design, some sidelobes. Still, in these areas the radiated power is reduced from the main beam.

In the navigation scenario under consideration a number of surveillance-type radars are positioned in an area of interest. Each radar provides coverage of a circular area with the radar at the center by scanning in azimuth. The directional nature of the radar energy beam means that signal reception is a function of the orientation of the radar’s main beam relative to the receiver. Unfortunately, the directional nature of the radar presents a problem to traditional geolocation algorithms.

Before discussing application of the directional mixed-modulated navigation signal in a navigation architecture, this chapter reviews current radio-based navigation methods. Radio based navigation involves the use of omnidirectional transmitters using a common time reference. This allows for near simultaneous receipt of signals from several transmitters at any particular epoch. From this, pseudoranges or time difference of arrival algorithms can be used to determine current position at the message receipt point. This chapter will
explore the two most used wide area radio based navigation methods, GPS and LORAN, in order to better understand the geolocation algorithm each uses.

7.1 Pseudorange Measurements - Satellite Navigation

The GPS constellation is a satellite-based system that uses pseudorandom range measurements from several GPS satellites broadcasting an omnidirectional signal to a receiver in order to provide a geolocation solution. The omnidirectional characteristics of GPS are due to the semi synchronous orbit the satellites occupy approximately 11,000 NM above the earth. At this orbital altitude, the area of coverage of a particular GPS satellite is only limited by the horizon and any line of sight obstructions between the receiver and the satellite.

At a basic level, GPS satellites use highly accurate, low drift rate clock broadcasting a time varying pseudorandom signal that is time synchronized to a world reference master clock. The receiver’s on-board clock is also synchronized with the same world reference master clock and has the ability to replicate the time varying pseudorandom sequence. By comparing the receiver’s on-board generated code with the received pseudorandom code a time difference is calculated that corresponds to the distance from the transmitting satellite to the receiver. Since the satellite orbit (ephemeris data) is known with exceptional accuracy, the relative position between a satellite and the receiver at a particular instance of time can be determined with a high degree of accuracy.

The receiver’s position is computed by solving a set of simultaneous equations based on pseudorange measurements from the receiver to a set of satellites. The pseudorange
from the satellite to the transmitter can be expressed as

\[ \rho_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2 + ct_b} \]  \hspace{1cm} (7.1)

where \([x_i, y_i, z_i]\) is the position of the satellite in the earth centered frame of reference, \([x, y, z]\) is the position of the receiver, and \(ct_b\) is any clock bias between the receiver and world clock multiplied by the speed of light. In the complete analysis, additional terms for relativistic effects due to satellite motion and atmospheric effects on the signal path are also involved. However, these terms are ignored as the discussion is meant to compare different approaches.

The geolocation equation using pseudorandom distances is a set of non-linear equations given by equation (7.2). These equations are a function of the satellite location (known) and the receiver position (unknown).

\[
\begin{bmatrix}
\rho_1 \\
\rho_2 \\
\rho_3 \\
\rho_4 \\
\end{bmatrix} = 
\begin{bmatrix}
\sqrt{(x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2 + ct_b} \\
\sqrt{(x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2 + ct_b} \\
\sqrt{(x_3 - x)^2 + (y_3 - y)^2 + (z_3 - z)^2 + ct_b} \\
\sqrt{(x_4 - x)^2 + (y_4 - y)^2 + (z_4 - z)^2 + ct_b} \\
\end{bmatrix} \hspace{1cm} (7.2)
\]

The solution to this set of equations is well documented in books and articles. For this paper, the solution put forth by Yarlagadda [28] is implemented.

To solve the set of equations, the pseudorange equations from each satellite, \(i\), are linearized via a Taylor series expansion about the user’s approximate position, \(\hat{x}, \hat{y}, \hat{z}\). The linearized equations are subtracted from the original pseudorange equation to get the effect of changes in position on the pseudorange. Ignoring non-linear higher order terms,
the linearized equations are

\[ d\rho_i = \left[ \frac{x_i - \hat{x}}{\hat{\rho}_i} dx + \frac{y_i - \hat{y}}{\hat{\rho}_i} dy + \frac{z_i - \hat{z}}{\hat{\rho}_i} dz - ct_b \right] \]

(7.3)

where \( \hat{\rho}_i = \sqrt{(x_i - \hat{x})^2 + (y_i - \hat{y})^2 + (z_i - \hat{z})^2} \). Equation (7.3) can be written in the matrix form as

\[
\begin{bmatrix}
  d\rho_1 \\
  d\rho_2 \\
  d\rho_3 \\
  d\rho_4
\end{bmatrix} \approx \begin{bmatrix}
  \frac{x_1 - \hat{x}}{\hat{\rho}_1} & \frac{y_1 - \hat{y}}{\hat{\rho}_1} & \frac{z_1 - \hat{z}}{\hat{\rho}_1} & 1 \\
  \frac{x_2 - \hat{x}}{\hat{\rho}_2} & \frac{y_2 - \hat{y}}{\hat{\rho}_2} & \frac{z_2 - \hat{z}}{\hat{\rho}_2} & 1 \\
  \frac{x_3 - \hat{x}}{\hat{\rho}_3} & \frac{y_3 - \hat{y}}{\hat{\rho}_3} & \frac{z_3 - \hat{z}}{\hat{\rho}_3} & 1 \\
  \frac{x_4 - \hat{x}}{\hat{\rho}_4} & \frac{y_4 - \hat{y}}{\hat{\rho}_4} & \frac{z_4 - \hat{z}}{\hat{\rho}_4} & 1
\end{bmatrix} \begin{bmatrix}
  dx \\
  dy \\
  dz \\
  -ct_b
\end{bmatrix}
\]

(7.4)

The approximate symbol in the equation (7.3) is due to ignoring the higher order Taylor series terms. Adding an error term, \( \epsilon \), the previous equation can be re-expressed in compact matrix notation

\[ d\bar{\rho} = A d\bar{x} + \bar{\epsilon} \]

(7.5)

where \( A \) is a square \( 4 \times 4 \) matrix. The solution to equation (7.5) is done through a least squares approach that minimizes \( \bar{\epsilon} \). In order for a solution to exist, the matrix, \( A \) must be nonsingular. The singular case arises if two or more rows or columns of the matrix are linear combinations of each other. However the satellites are in fixed orbits and not aligned with each other. Further only satellites within the field of view of the receiver (e.g. above the horizon) can be used which restricts the allowable vector space. Therefore, since the pseudo range vectors are linearly independent, the inverse to \( A \) exists.

In the general case, if more than four satellites are available, more pseudorandom measurements may be taken resulting in an overdetermined system of equations. The benefit of additional measurements is that each measurement provides additional information that helps reduces the error term.
Solving for the overdetermined case requires use of the pseudoinverse, \((A^T A)^{-1}\). This produces

\[ d\bar{x} = (A^T A)^{-1} A^T d\bar{\rho} \]  

(7.6)

It is easily seen that for the determined \(n \times n\) case, equation 7.6 reduces to \(d\bar{x} = A^{-1} d\bar{\rho}\) since

\[ (A^T A)^{-1} A^T = A^{-1} A^{-T} A^T = A^{-1} \]  

(7.7)

The set of equations is still not solvable in closed form as \(A\) is a function of present position \((\hat{x}, \hat{y}, \hat{z})\). Therefore an iterative approach is employed to solve the system using an initial estimate of the current location. The iterations continue until the difference between the \(d\bar{x}_k - d\bar{x}_{k+1}\) is down to an acceptable error. With an IMU on board, the position can be updated using inertial measurements of acceleration thereby providing a more accurate initial ‘guess’ for subsequent updates.

So far, the effects of processor, receiver and measurement noise have been ignored. However, these effects can be mitigated by incorporating a Kalman filter into the navigation scheme.

### 7.2 TDOA Measurements - LORAN

While the GPS algorithm uses pseudorange measurements to achieve a geolocation solution, pseudorange measurements require very accurate timing criteria at both the transmitter and receiver. Until the advent of atomic clocks with extremely low drift rates this was not feasible. Another approach has been used that reduces the receiver reliance on a common time reference and instead relies on differences in the signal time of arrival at the receiver, or TDOA.
TDOA (time difference of arrival) is illustrated in Figure 7.1 for two-dimensional geolocation using three transmitters, A, B, and C. In this arrangement a receiver records the time of arrival of a pulse from three different transmitters. If the transmitters broadcast simultaneously, the difference in arrival time is equivalent to the difference in distances between the receiver and the transmitters. These time differences uniquely identify a point in a plane for the two-dimensional case. The principal benefit of TDOA is that while time synchronization is required between ground stations, there is no requirement for time synchronization between the aircraft/vehicle and the ground stations.

The LORAN chain, in operation until the late 1990s, used a variation of this approach through a series of master and slave stations with a slight time delay in transmission between the master and slave stations. In the LORAN system the master station, \( M \), sends out a pulse at time \( T \) so it arrives at the receiver at \( T + \tau_M \). Slave station 1, \( S_1 \), then sends a pulse a very short time, \( \delta T_1 \), after the master so it arrives at \( T + \delta T_1 + \tau_{S_1} \). Similarly slave station 2 follows with a receiver arrival time at \( T + \delta T_2 + \tau_{S_2} \). \( \delta T \) is selected to be

![Figure 7.1: Time Difference of Arrival](image-url)
very small so any movement of the vehicle between the master signal and slave signal is negligible.

The benefit of TDOA is that the receiver only needs to record the arrival time of the pulse using its on-board clock which does not have to be synchronized with any external source. For the two-dimensional geolocation problem, \((x, y)\), the difference in arrival times of the pulses produces three pseudorange difference equations:

\[
\begin{bmatrix}
c [T + \tau_M - (T + \delta T_2 + \tau_{S2})] \\
c [T + \tau_M - (T + \delta T_3 + \tau_{S3})] \\
c [T + \delta T_2 + \tau_{S2} - (T + \delta T_3 + \tau_{S3})]
\end{bmatrix}
= \begin{bmatrix}
c [\tau_{M-S2} - \delta T_2] \\
c [\tau_{M-S3} - \delta T_3] \\
c [\tau_{S2-S3} + (\delta T_3 - \delta T_2)]
\end{bmatrix}
\tag{7.8}
\]

where \(\tau_{a-b} = \tau_a - \tau_b\) and \(c\) is the speed of light.

Since the slave station delay from the master station, \(\delta T_i\), is established for each slave station, the time difference can be determined as well as the pseudorange difference by multiplying the result by \(c\). Again, atmospheric error, Doppler effects on the frequency, and clock errors between transmitters are not addressed in this background discussion. While our focus is on the two-dimensional LORAN example, the TDOA approach is extensible to the three-dimensional realm.

Solving the TDOA system of equations follows a somewhat similar path as was used for pseudorange measurements. The general two-dimensional pseudorange difference equation generated by the TDOA measurement can be expressed as:

\[
c \tau_{a-b} = \sqrt{(x_a - \bar{x})^2 + (y_a - \bar{y})^2} - \sqrt{(x_b - \bar{x})^2 + (y_b - \bar{y})^2}
\tag{7.9}
\]

Recasting equation 7.8 the change in time difference equations in terms of present position
of the receiver can be written as

\[
\begin{bmatrix}
  c [\tau_{M-S2} - \delta T_2] \\
  c [\tau_{M-S3} - \delta T_3] \\
  c [\tau_{S2-S3} + (\delta T_3 - \delta T_2)]
\end{bmatrix}
= \begin{bmatrix}
  \sqrt{(x_1 - \tilde{x})^2 + (y_1 - \tilde{y})^2} - \sqrt{(x_2 - \tilde{x})^2 + (y_2 - \tilde{y})^2} \\
  \sqrt{(x_1 - \tilde{x})^2 + (y_1 - \tilde{y})^2} - \sqrt{(x_3 - \tilde{x})^2 + (y_3 - \tilde{y})^2} \\
  \sqrt{(x_2 - \tilde{x})^2 + (y_2 - \tilde{y})^2} - \sqrt{(x_3 - \tilde{x})^2 + (y_3 - \tilde{y})^2}
\end{bmatrix}
\]

(7.10)

where \(\tilde{x}, \tilde{y}\) represent present position of the vehicle at message receipt.

At first glance, a closed form solution appears to be as intractable as the pseudorange problem. However, Bucher and Misra [4] have derived a closed form three-dimensional solution as part of a 2002 paper that implements the algorithm in a synthesizable VHDL model. For completeness, their closed form solution follows.

For clarity purposes the following notation is used. Transmitters are identified by \(i, j, k, l\) while the receiver position is the standard \(x, y, z\). Pseudorange differences between transmitter \(a\) and \(b\) is shown as \(R_{ab}\) while distances between, for example, the \(x\) position component of two transmitters is \(x_{ab} = x_a - x_b\). Receiver position is \(x, y, z\) without subscripts while while capital letters (e.g. A, B, C, … etc) are used as intermediate variables. Note that distances are sign (e.g. \(\pm\)) sensitive in these calculations; they are not 2-norm, absolute distances.

The fundamental equations that compute the \(x, y\) and \(z\) coordinates are

\[
x = Gz + H \quad y = A(Gz + H) + Bz + C \quad \text{where } z = \frac{N}{2M} \pm \sqrt{\left(\frac{N}{2M}\right)^2 - \frac{P}{M}} \quad (7.11)
\]

In the equations above, the \(z\) component is computed first and used to determine \(x, y\). The
capital letter terms are intermediate values computed as follows:

\[
A = \begin{bmatrix} R_{ik}x_{ji} - R_{ij}x_{ki} \\ R_{ij}y_{ki} - R_{ik}y_{ji} \end{bmatrix} \quad B = \begin{bmatrix} R_{ik}z_{ji} - R_{ij}z_{ki} \\ R_{ij}y_{ki} - R_{ik}y_{ji} \end{bmatrix}
\tag{7.12}
\]

\[
C = \frac{R_{ik} \left[ R_{ij}^2 + x_i^2 - x_j^2 + y_i^2 - y_j^2 + z_i^2 - z_j^2 \right] - R_{ij} \left[ R_{ik}^2 + x_i^2 - x_k^2 + y_i^2 - y_k^2 + z_i^2 - z_k^2 \right]}{2 \left[ R_{ij}y_{ki} - R_{ik}y_{ji} \right]}
\tag{7.13}
\]

\[
D = \begin{bmatrix} R_{kl}x_{jk} - R_{kj}x_{lk} \\ R_{kj}y_{lk} - R_{kl}y_{jk} \end{bmatrix} \quad E = \begin{bmatrix} R_{kl}z_{jk} - R_{kj}z_{lk} \\ R_{kj}y_{lk} - R_{kl}y_{jk} \end{bmatrix}
\tag{7.14}
\]

\[
F = \frac{R_{kl} \left[ R_{kj}^2 + x_k^2 - x_j^2 + y_k^2 - y_j^2 + z_k^2 - z_j^2 \right] - R_{kj} \left[ R_{kl}^2 + x_k^2 - x_l^2 + y_k^2 - y_l^2 + z_k^2 - z_l^2 \right]}{2 \left[ R_{kj}y_{lk} - R_{kl}y_{jk} \right]}
\tag{7.15}
\]

\[
G = \frac{E - B}{A - D} \quad H = \frac{F - C}{A - D} \quad I = AG + B
\tag{7.16}
\]

\[
J = AH + C \quad K = R_{ik}^2 + x_i^2 - x_k^2 + y_i^2 - y_k^2 + z_i^2 - z_k^2 + 2x_{ki}H = 2y_{ki}J
\tag{7.17}
\]

\[
L = 2 \left[ x_{ki}G + y_{ki}I + 2z_{ki} \right]
\tag{7.18}
\]

\[
M = 4R_{ik}^2 \left[ G^2 + I^2 + 1 \right] - L^2
\tag{7.19}
\]

\[
N = 8R_{ik}^2 \left[ G(x_i - H) + I(y_i - J) + z_i \right] + 2LK
\tag{7.20}
\]

\[
P = R_{ik}^2 \left[ (x_i - H)^2 + (y_i - J)^2 + z_i^2 \right] - K^2
\tag{7.21}
\]

It should be noted that the \( z \) term in equation (7.11) yields two result with the positive result being the correct one. However, as will be seen, the use of terrestrial based pseudo-lites produces unacceptable errors in the \( z \)-axis that propagate into \( x, y \) errors due to GDOP.
7.2.1 Summary

By definition, radio-based navigation systems are designed to be omnidirectional so that simultaneous reception of all signals is possible at a particular epoch. However, this study considers the case where the navigation message is integrated into a directional transmitter, specifically radar.

For a combined radar-communications message the directional nature of the radar will cause variations in the received power of the signal over time especially when the receiver is outside the main beam of the radar. While there is some power leakage in the radar sidelobes, the design of the radar may limit this. As a result, simultaneity of signal receipt at a specific epoch cannot be assured and the TDOA approach is not feasible without some modification. An algorithm that accounts for the non-simultaneity of message receipt is the focus of the next chapter.
Non-Synchronous TDOA Geolocation

The key challenge with a directional navigation signal is the non-simultaneous nature of the signals at the receiver. Traditional navigation systems usually employ a signal that provides wide area coverage in order to ensure temporal simultaneity in the receipt of multiple navigation signals. This is done by either an omni-directional transmitted waveform (e.g. LORAN, VOR) or a directionally radiated signal at an altitude that enables large area coverage (e.g. GPS).

Further, the ubiquitous nature of traditional navigation signals allows the receiver to “lock on” and maintain continuous reception from different transmitters or satellites. Multilateration techniques rely on this characteristic for simultaneous receipt of information so that time or angle difference of arrival (TDOA, AODA) or pseudoranges are all measured relative to the transmitters at a particular instance in time (epoch).

Conversely, in the case of directional navigation signals, signal receipt may be “episodic” in that there is no guarantee the minimum required number of radar beams will focus on the receiver at the same epoch. Depending on radar power and signal strength outside the radar main beam, a more likely scenario is one where the receiver picks up navigation information from different radars at different times and slightly different locations. As a result, the multilateration algorithm needs to account for non-synchronous receipt of navigation information from the radar sites.
In developing the non-synchronous geolocation algorithm the general case will be considered where the aircraft clock is not synchronized to a common true time (world time) reference. By not relying on a receiver that must be time synchronized to the pseudolite clock, the challenges associated with precision time transfer at accuracies on the order of $10^{-8}$ to achieve meter-level accuracy are averted. However, it is crucial that the ground radar clocks share a common time reference. Although each radar does not have to be precisely aligned to the reference, the offset of each radar clock from the world clock must be known and either be shared or the same for all radars. This information can then be provided to the receiver as part of the navigation message.

Besides a compatible signal receiver, the aircraft/vehicle must have some type of inertial measurement unit (IMU) that allows for measurement of direction and velocity as well as an on-board clock to measure time of arrival of signals. The IMU is required for two reasons. The first is to smooth the navigation system’s estimated vehicle trajectory between updates and, second, for estimating position propagation and measured spatial distances between updates to allow for a navigation solution/update. The on-board clock of the vehicle is assumed to have drift rates consistent with that of an atomic clock, but it is not synchronized with the master clock of the ground radars. The proliferation of relatively low cost, low drift, atomic clock chips and chip-based IMUs mean these are not onerous requirements.

The lack of a ubiquitous navigation signal means timing is especially critical in order to associate pulse receipt with a particular transmission time. Therefore, this chapter begins with a discussion of approaches to ensure that pulse receipt can be unambiguously correlated to a specific transmission event. Following this is the development of the non-synchronous geolocation algorithm from first principles. The development section
also includes a discussion of conditions that induce ambiguity into the solution for certain radar/vehicle geometries and corrective measures. Finally, simulations are presented to show convergence and accuracy.

8.1 Non-Synchronous Timing Approach

Ensuring a unique relationship between the transmitted and received pulse can be handled one of two ways. With all radars operating from a common time reference, the most obvious approach is to incorporate the transmission time as part of the navigation message in the radar pulse. The challenge with this is whether a pulse has enough bits to encode the time. Realistically, the timing requirement to achieve non-WAAS GPS accuracies of ±10m is roughly on the order of $3.3333 \times 10^{-8}$ seconds. If the negative exponent is assumed, this requires $\log_2(10^9) \approx 30$ bits of timing information coded into the pulse.

One way to reduce the bit requirements is to transmit the difference relative to a secondary periodic timing source. If, for example, the secondary source was operating at 100 Hz, the period would be $10^{-2}$ and the transmitted pulse would encode the time difference from the start of the current cycle. This provides an unambiguous range of 3000 KM between cycles. It also reduces the bit requirement to $\log_2(10^7) \approx 24$ bits of timing information.

The second, and preferred, approach is to establish a periodic rate of when the radar starts sending a sequence of navigation encoded pulses. In this option, the radar includes navigation information in the radar pulse precisely every $\delta T$ seconds to avoid ambiguity problems. The time delay associated with different pulse retransmission times can be factored out by either a code within the navigation pulse or through using the modulus of the
interval, $\delta T$, between pulses. This is illustrated in Figure 8.1.

Figure 8.1: Radar Pulse Interval

In this figure, pulses transmitted at instances $T_1, T_2,$ and $T_4$ are received at various aircraft clock times indicated by the triangles. Each pulse receipt is associated with a different transmission time. Taking the difference of arrival clock time at the receiver yields the wrong TDOA. In looking at pulses 1 and 2, the blue pulse transmitted at $T_1$ is received at $t_1$ while the orange pulse transmitted at $T_2$ is received at $t_2$. In aircraft clock time, the first signal is received at $t_1 = T_1 + t_c + \tau_1$ and the second at $t_2 = T_2 + t_c + \tau_2$, where $t_c$ is the unknown difference (bias) between the aircraft clock and world clock. Since $T_2 = T_1 + \delta T$ we can write the time difference between $t_1$ and $t_2$ as

$$t_2 - t_1 = T_1 + \delta T + t_c + \tau_2 - (T_1 + t_c + \tau_1) = \tau_2 - \tau_1 + \delta T$$  \hspace{1cm} (8.1)

Since $\delta T$ is constant and known, there are two ways to proceed. One way is to take the modulus of $t_2 - t_1$ with respect to $\delta T$ which, as long as $t_2 - t_1 << \delta T$, factors out any $n\delta T$ elements of the interval. The second way is to sequentially number the transmission pulses and encode that number as part of the navigation message in the pulse. This way the problem becomes one of taking the difference of the sequence numbers times $\delta T$ and removing that from the aircraft clock time delay.
It is important to re-emphasize that $\delta T$ must be sufficiently large to ensure there is no ambiguity in the association. Practically, the frequency and period of the secondary timing source, $T$, should be at least twice the unambiguous range of the area of navigation. For example, if the area of navigation is of size $m \times m$, then

$$T_{\text{min}} = \frac{m\sqrt{2}}{c}$$

For a $1,000 \times 1,000$ km area, this would required a $T_{\text{min}}$ of 4.7 msec between pulses. As can be seen in the example, the ‘twice the usable distance’ rule still allows for a relatively high data throughput at the receiver depending on the radar scan rate and half-power beamwidth.

Of the two approaches (secondary timing source or periodic transmission), the second using periodic transmission time is preferred since it minimizes the data requirements per pulse. However, it also places a constraint on the radar as the radar now has to schedule the pulse transmission of the navigation encoded pulse to a highly time sensitive transmission interval. Errors in transmission times of even $1\mu \text{sec}$ can induce errors on the order of 300 meters which emphasizes the need for atomic clock accuracy.

### 8.2 Non-Synchronous Algorithm Development

Figure 8.2 provides a general overview of the non synchronous TDOA problem. In the diagram, an aircraft flight path is depicted by the black line moving from left to right. Four radars, $R_1, R_2, R_3, R_4$, send a pulse modulated by a navigation message. The navigation message includes, at a minimum, a code providing a reference to the transmitting radar. As previously discussed, it is also possible to include a counter that identifies the transmission
A necessary requirement for this approach is that the aircraft have an inertial measurement unit (IMU) on board to measure acceleration, velocity and direction of the aircraft. The IMU position is used to establish an accurate spatial difference relationship between message receipt points. Initially, the IMU does not need to know its true position; this will be determined by the nonsynchronous TDOA algorithm with periodic position updates.

At a message receipt time, the aircraft clock time and estimated position \( (t_n, \hat{P}_n) \) are recorded. The time of arrival is based on the aircraft clock while the estimated position is from the IMU. The transit time of the pulse from each radar to the receiver, \( \tau_n \), is shown in figure (8.2). This time is not known, however, it factors into the calculations as will be seen later. Between receipt times \( t_1, t_2, t_3, \) and \( t_4 \), the change in position of the aircraft is recorded as \( N_{12}, N_{23} \) and \( N_{34} \) where \( N_{ab} \) is defined as the difference in IMU position be-
between message receipt point $a$ point $b$. The position change along with the time of receipt connect the four message receipt points in a unique solution.

Based on these calculations, the clock times, $t_n$, at each message receipt point are given by the following systems of equations.

$$
t_1 = T_1 + t_c + \tau_1
$$
$$
t_2 = T_1 + n_2\delta T + t_c + \tau_2
$$
$$
t_3 = T_1 + n_3\delta T + t_c + \tau_3
$$
$$
t_4 = T_1 + n_4\delta T + t_c + \tau_4
$$

A time difference relationship can be established between these times that removes unknowns such as the clock bias, $t_c$, and the actual (world clock) transmission time at $T_1$. The differences between the time equations reduces to:

$$
t_{12} = \tau_1 - \tau_2 - n_2\delta T
$$
$$
t_{13} = \tau_1 - \tau_3 - n_3\delta T
$$
$$
t_{23} = \tau_2 - \tau_3 + \delta T(n_2 - n_3)
$$
$$
t_{34} = \tau_3 - \tau_4 + \delta T(n_3 - n_4)
$$

where $t_{ab} = t_a - t_b$.

As mentioned earlier, the term $\delta T$ is a standard separation between pulse transmissions in the network radar array and is intentionally set much larger than any allowable $\tau_n$. Therefore, it can be removed by either taking the modulus of the time differences with $\delta T$. If the pulse sequences are sequentially numbered and encoded in the message, another approach is to use the multiply the difference between the pulse number by $\delta T$. As a result
we have a series of equations that relate the difference in distances from the each radar at each point based on the aircraft clock time.

\[ t_{12} = \tau_1 - \tau_2 \quad t_{13} = \tau_1 - \tau_3 \quad t_{23} = \tau_2 - \tau_3 \quad t_{34} = \tau_3 - \tau_4 \quad (8.5) \]

It is important to note that these time differences are **NOT** true time difference of arrivals since the arrival clock times, \( t_n \), do not occur at precisely the same location. Instead each arrival time is at a displaced location from other message receipt points.

The pseudorange distances of the vehicle at each message receipt point from the transmitting radar in terms of the message transit time, \( \tau_n \), are given by

\[
\bigg| \hat{P}_1 - R_1 \bigg| = c\tau_1 \quad \bigg| \hat{P}_2 - R_2 \bigg| = c\tau_2 \quad \bigg| \hat{P}_3 - R_3 \bigg| = c\tau_3 \quad \bigg| \hat{P}_4 - R_4 \bigg| = c\tau_4 \quad (8.6)
\]

where the \( || \) term is the 2-norm of the vector results. For the three-dimensional case, there are a total of sixteen unknowns - twelve from the \( x, y, z \) components of \( \hat{P}_1, \hat{P}_2, \hat{P}_3, \hat{P}_4 \) and four from \( \tau_1, \tau_2, \tau_3, \tau_4 \). However, message receipt positions \( P_a \) and \( P_b \) are related by the known IMU position change, \( N_{ab} \). Therefore the position of the vehicle at any of the three message receipt points in terms of a single position (in this case \( P_4 \)) can be obtained using the following procedure. \( \hat{P}_1 \) is related to \( \hat{P}_4 \) through the cumulative IMU changes as \( \hat{P}_1 = \hat{P}_4 - (+N_{12} + N_{23} + N_{34}) \). Similarly, \( \hat{P}_2 \) and \( \hat{P}_3 \) are substituted in equation (8.6) as shown in equation (8.7).

\[
\bigg| \hat{P}_4 - N_{12} - N_{23} - N_{34} - R_1 \bigg| = c\tau_1 \\
\bigg| \hat{P}_4 - N_{23} - N_{34} - R_2 \bigg| = c\tau_2 \\
\bigg| \hat{P}_4 - N_{34} - R_3 \bigg| = c\tau_3 \\
\bigg| \hat{P}_4 - R_4 \bigg| = c\tau_4 \quad (8.7)
\]
At this point it appears there are four equations in seven unknowns. However, the signal transit times, \( \tau_n \) for \( n = 1, 2, 3, 4 \), are related through the time difference relationships of equation (8.5). As a result the system of equations can be reduced to one where the unknown variables are the \( X, Y, Z \) components of \( P_4 \) and \( \tau_4 \). The remaining terms are measured IMU or time of arrival values and known radar locations.

The solution of this system begins with squaring both sides of the sides of each equation and expanding the quadratic terms. Next the relationships between \( P_4 \) and the other message receipt positions are used to reduce the number of unknown positions to a function of \( P_4 \). This produces the following system of equations:

\[
\begin{align*}
(X - x_{N14} - x_{R1})^2 + (Y - y_{N14} - y_{R1})^2 + (Z - z_{N14} - z_{R1})^2 &= (c\tau_1)^2 \\
(X - x_{N24} - x_{R2})^2 + (Y - y_{N24} - y_{R2})^2 + (Z - z_{N24} - z_{R2})^2 &= (c\tau_2)^2 \\
(X - x_{N34} - x_{R3})^2 + (Y - y_{N34} - y_{R3})^2 + (Z - z_{N34} - z_{R3})^2 &= (c\tau_3)^2 \\
(X - x_{R4})^2 + (Y - y_{R4})^2 + (Z - z_{R4})^2 &= (c\tau_4)^2
\end{align*}
\]

where \( X, Y, \) and \( Z \) refer to the components of \( P_4 \) and the notation \( N_{14} = N_{12} + N_{23} + N_{34} \), and \( N_{24} = N_{23} + N_{34} \). In these equations, the IMU changes, \( N_{12}, N_{23}, \) and \( N_{34} \), as well as the radar locations are known vector values. For convenience, we define

\[
\begin{align*}
\alpha &= N_{14} + R_1 \\
\beta &= N_{24} + R_2 \\
\gamma &= N_{34} + R_3
\end{align*}
\]
Substituting these terms and expanding the left side results in

\[ X^2 - 2X\alpha_x + \alpha_x^2 + Y^2 - 2Y\alpha_y + \alpha_y^2 + Z^2 - 2Z\alpha_z + \alpha_z^2 = (c\tau_1)^2 \]
\[ X^2 - 2X\beta_x + \beta_x^2 + Y^2 - 2Y\beta_y + \beta_y^2 + Z^2 - 2Z\beta_z + \beta_z^2 = (c\tau_2)^2 \]
\[ X^2 - 2X\gamma_x + \gamma_x^2 + Y^2 - 2Y\gamma_y + \gamma_y^2 + Z^2 - 2Z\gamma_z + \gamma_z^2 = (c\tau_3)^2 \]
\[ X^2 - 2X x_{R4} + x_{R4}^2 + Y^2 - 2Y y_{R4} + y_{R4}^2 + Z^2 - 2Z z_{R4} + z_{R4}^2 = (c\tau_4)^2 \]

(8.10)

Initially, the system of equations appears to be non-linear due to the \(X^2, Y^2,\) and \(Z^2\) terms. However, these can be removed through elementary row operations. After subtracting the bottom equation for \((c\tau_4)^2\) from the first three, moving constant terms from the left to the right side of the equations, and placing the result in matrix form gives:

\[ Ax = b \]

(8.11)

where

\[ A = \begin{bmatrix} (x_{R4} - \alpha_x) & (y_{R4} - \alpha_y) & (z_{R4} - \alpha_z) \\ (x_{R4} - \beta_x) & (y_{R4} - \beta_y) & (z_{R4} - \beta_z) \\ (x_{R4} - \gamma_x) & (y_{R4} - \gamma_y) & (z_{R4} - \gamma_z) \end{bmatrix} \]

(8.12)

\[ x = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \]

(8.13)

and

\[ b = \frac{1}{2} \begin{bmatrix} c^2 (\tau_1^2 - \tau_4^2) - \alpha_x^2 + x_{R4}^2 - \alpha_y^2 + y_{R4}^2 - \alpha_z^2 + z_{R4}^2 \\ c^2 (\tau_2^2 - \tau_4^2) - \beta_x^2 + x_{R4}^2 - \beta_y^2 + y_{R4}^2 - \beta_z^2 + z_{R4}^2 \\ c^2 (\tau_3^2 - \tau_4^2) - \gamma_x^2 + x_{R4}^2 - \gamma_y^2 + y_{R4}^2 - \gamma_z^2 + z_{R4}^2 \end{bmatrix} \]

(8.14)
with the values of $\tau_n$ subject to the previously discussed relationship

$$
t_{12} = \tau_1 - \tau_2 \quad t_{13} = \tau_1 - \tau_3 \quad t_{24} = \tau_2 - \tau_4 \quad t_{34} = \tau_3 - \tau_4
$$

(8.15)

Further reduction of the equations (8.14) and (8.15) is not possible. Although equation (8.15) appears to be a system of four equations in four unknowns, there are an infinite number of solutions to the linear system. This can be proven by observing the matrix form of equation (8.15) as shown in equation (8.16).

$$
\begin{bmatrix}
1 & -1 & 0 & 0 \\
1 & 0 & -1 & 0 \\
0 & 1 & 0 & -1 \\
0 & 0 & 1 & -1 \\
\end{bmatrix}
\begin{bmatrix}
\tau_1 \\
\tau_2 \\
\tau_3 \\
\tau_4 \\
\end{bmatrix}
= 
\begin{bmatrix}
t_{12} \\
t_{13} \\
t_{24} \\
t_{34} \\
\end{bmatrix}
$$

(8.16)

in this system the $4 \times 4$ matrix is singular precluding a unique solution. Although, it would appear that adding more information through another message point, $P_5$, would help. However, adding another message receipt point is offset by the addition of another unknown term, $\tau_5$, to the system.

Despite the lack of a firm value for $\tau_n$, solving the system of equation can be approached iteratively for a linear system of equations. Fundamental to this hypothesis is that the relationship between the time difference of arrival values and IMU measured distances between message receipt points, are tied to specific radar transmitters. Also there exists a unique set of points from $P_1$ to $P_4$ that satisfy these relationships. Specifically, what is required is that the inverse of equation (8.12) exists.
8.2.1 Algorithm Singularities

In order to have a unique solution to the linear system shown in equations (8.11) through (8.14) matrix $A$ must be non-singular. Specifically we are interested in any set of values that result in a zero-valued determinant.

\[
\begin{vmatrix}
  x_{R4} - \alpha_x & y_{R4} - \alpha_y & z_{R4} - \alpha_z \\
  x_{R4} - \beta_x & y_{R4} - \beta_y & z_{R4} - \beta_z \\
  x_{R4} - \gamma_x & y_{R4} - \gamma_y & z_{R4} - \gamma_z
\end{vmatrix} = 0 \quad (8.17)
\]

The algebra for the determinant can get fairly complicated, but there is an alternative approach. A zero determinant can result if there is a zero column or row or if one or more rows or columns are linearly dependent. Reviewing equation (8.17), each row is a function of the respective $x, y, z$ components of radar number four and $\alpha, \beta$ and $\gamma$. Further, $\alpha, \beta$ and $\gamma$ are functions of the IMU position changes between $R_4$ and $R_1, R_2,$ and $R_3$, respectively.

Examining the $x$-component of the first row and substituting for $\alpha_x$

\[
x_{R4} - \alpha_x = x_{R4} - (x_{N12} + x_{N23} + x_{N34} + x_{R1}) = 0
\]

or

\[
x_{R4} - x_{R1} = x_{N12} + x_{N23} + x_{N34} = x_{N14} \quad (8.18)
\]

In this equation the sum of the $N$-terms represent the $x$-component change between message receipt from radar 1 and radar 4. From this equation it can be seen that if the change in the $x$-component receipt positions between radars 1 and 4 is the same as the $x$-component distance traveled between the same message receipt points, the $x$-term is zero. Comparing this to the $y$ and $z$ terms in row one produces the same conclusion. A similar conclusion can be reached in looking at other rows.

These equations provide the conditions for a zero determinant; namely, that if the
IMU position change between the receipt of signal from two radars is equal to the distance between the two radars, a singular matrix results. This event occurs when the line-of-sight from a radar to the vehicle is orthogonal to the vehicle’s flight path and the radars are all located to one side of the flight path. The extreme example is when all four radars meet this constraint and is illustrated in Figure 8.3:

![Figure 8.3: Ambiguous Nonsynchronous Geolocation](image)

In figure (8.3), flight path $F_1$ is the true trajectory of the vehicle where the change in IMU position and message receipt points between adjacent radars are identical with the spatial separation of the radars. As can be seen, an arbitrary $F_2$ has the same characteristics as $F_1$ in that the IMU changes and differences in times of arrival are equal. For that matter there are an infinite number of acceptable parallel solutions for $F_1$ and $F_2$ with the same characteristics. Fortunately this situation is highly unlikely due to the requirement that the flight path of the vehicle must exactly coincide with signal receipt orthogonal to the vehicle path at all measurement points. If even one of the radar receptions is not perpendicular, the distance between neighboring radar, $N_{ab}$, will change relative to other parallel paths.

To preclude an inadvertent stumble into this singular matrix scenario the solution is to first evaluate the determinant of the matrix for condition. If the matrix is poorly condi-
tioned, the current set of measurements should be rejected for a new set of measurements or additional measurements from other radar sites with better geometry should be added in lieu of earlier measurements. In essence, this is similar to a GPS receiver selecting an optimal set of satellites by computing a pre-measurement GDOP of the visible satellites.

8.2.2 Iterative Geolocation Algorithm

Assuming a non-singular $A$ matrix, the closed-form solution of the non-synchronous geolocation algorithm is

$$
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} = \frac{1}{2}
\begin{bmatrix}
(x_{R4} - \alpha_x) & (y_{R4} - \alpha_y) & (z_{R4} - \alpha_z) \\
(x_{R4} - \beta_x) & (y_{R4} - \beta_y) & (z_{R4} - \beta_z) \\
(x_{R4} - \gamma_x) & (y_{R4} - \gamma_y) & (z_{R4} - \gamma_z)
\end{bmatrix}^{-1}
\begin{bmatrix}
c^2 (\tau_1^2 - \tau_4^2) - \alpha_x^2 + x_{R4}^2 - \alpha_y^2 + y_{R4}^2 - \alpha_z^2 + z_{R4}^2 \\
c^2 (\tau_2^2 - \tau_4^2) - \beta_x^2 + x_{R4}^2 - \beta_y^2 + y_{R4}^2 - \beta_z^2 + z_{R4}^2 \\
c^2 (\tau_3^2 - \tau_4^2) - \gamma_x^2 + x_{R4}^2 - \gamma_y^2 + y_{R4}^2 - \gamma_z^2 + z_{R4}^2
\end{bmatrix}
$$

(8.19)

To start the iterations, we begin by finding an estimated value of $\tau_4$ based on the current estimated position. This value of $\tau_4$ determines the values of $\tau_3$, $\tau_2$ and $\tau_1$ which permits a direct solution to $P_4$ (e.g. $x_4$, $y_4$ and $z_4$) using equation(8.19). With $P_4$, $N_{12}$, $N_{23}$, and $N_{34}$ positions $P_1$, $P_2$ and $P_3$ can be computed. The revised position estimates are used to update the values of $\tau_n$ using the known radar positions. This iterative approach continues until the change in position in $P_4$ is within an acceptable tolerance.

For an initial position computation with no $a$ priori position information, the estimated position update can be the average $x$, $y$, $z$ location of the three radars used for navigation.
Subsequent updates should start with the IMU computed position based on prior navigation signal updates.

**Overdetermined Case**

It is also possible to use additional measurements in an effort to improve accuracy of the solution. A general form of the system of equations for \( n \) radars is

\[
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} = \frac{1}{2}
\begin{bmatrix}
(x_n - x_{N1n} - R_{x1}) & (y_n - y_{N1n} - R_{y1}) & (z_n - z_{N1n} - R_{z1}) \\
(x_n - x_{N2n} - R_{x2}) & (y_n - y_{N2n} - R_{y2}) & (z_n - z_{N2n} - R_{z2}) \\
\vdots & \vdots & \vdots \\
(x_n - x_{N(n-1),n} - R_{x(n-1)}) & (y_n - y_{N(n-1),n} - R_{y(n-1)}) & (z_n - z_{N(n-1),n} - R_{z(n-1),n})
\end{bmatrix}
\begin{bmatrix}
c^2 (\tau_1^2 - \tau_n^2) + r_n^2 - \alpha_{x1}^2 - \alpha_{y1}^2 - \alpha_{z1}^2 \\
c^2 (\tau_1^2 - \tau_n^2) + r_n^2 - \alpha_{x2}^2 - \alpha_{y2}^2 - \alpha_{z2}^2 \\
\vdots \\
c^2 (\tau_{n-1}^2 - \tau_n^2) + r_n^2 - \alpha_{x(n-1)}^2 - \alpha_{y(n-1)}^2 - \alpha_{z(n-1)}^2
\end{bmatrix}^{-1}
\]

where \( r_n^2 = x_n^2 + y_n^2 + z_n^2 \) for the location of radar \( n \) and \( \alpha_{xm}, \alpha_{ym}, \alpha_{zm} \) are the \( x, y, z \) components of \( \alpha_m \) where \( \alpha_m = N_{m,m=1} + N_{m+1,m+2} + \ldots + N_{n-1,n} + R_m \) for \( m + 1, 2, \ldots, n - 1 \). Assuming the inverse exists, the well known solution to this overdetermined system of equations is

\[
x = (A^T A)^{-1} A^T b
\]
8.3 Non-Synchronous TDOA Algorithm Simulation

To demonstrate the effectiveness of the non-synchronous least squares TDOA algorithm presented in the last section, a Matlab simulation was encoded. For the simulation the navigation area of interest was deemed to be a $1,000 \times 1,000$ km square and had five radars distributed randomly, but uniformly around the area. The area of navigation is assumed to use a Cartesian coordinate system with the $[0 \ 0]$ point at the southwest corner and the $[1000 \ 1000]$ coordinate at the northeast corner. The vector components are the also in kilometers.

The simulation uses two vectors in the calculations. $\mathbf{P}$ is the actual (true) position of the vehicle and is used to determine the exact time of arrival of a pulse at the receiver and to provide a standard for comparison. The aircraft’s computed position both from its IMU and the geolocation algorithm is stored in $\hat{\mathbf{P}}$. The simulation generates a random interval time between receipt of navigation messages from consecutive radars. The maximum interval can be set in the program. For the simulation that follows, the interval time varied between 1-4 seconds between consecutive receipts of signal from different radars. The change in position at each point is stored for computation.

For this simulation the two-dimensional geolocation algorithm is run after radar/navigation signals from three radars are received. Information on the measured IMU position change at each message receipt point is stored. This information is used to spatially relate all of the message receipt points to each other. The algorithm computes the position of the vehicle at the last message receipt point. With this a revised set of transit times, $\tau_n$, for the updated locations to the radars is computed. This allows for another iteration of the algorithm. The iterations continue until the location change from one iteration to the next is within a difference of 0.1 meters. In practice, since the algorithm is a closed form solution, only two iterations are normally required.
Figure 8.4: Estimated (Top) and True (Bottom) Trajectories

In the simulation the actual (true) initial position, \( \mathbf{P} = [150 \ 300 \ 1] \) as referenced to the \( x, y, z \) grid location in kilometers. At \( t = 0 \), the vehicle’s navigation system has no knowledge of its position so the initial estimate is set at \([50 \ 150 \ 1]\). As will be discussed
in the next chapter, only a two-dimensional, $x, y$ position was computed. The reason for this was due to dilution of precision problems in the $z$-axis that introduced additional errors in the interdependent computations for $x$ and $y$. Practically, the aircraft altimeter can provide an accurate measurement of altitude relative to sea level so this does not seriously degrade this adjunct navigation method.

Using these parameters and the algorithm, the following trajectory and error plots were produced. In figure 8.4 the true flight path of the vehicle is shown in blue while the plot showing the estimated position is in red. The square area on the estimated position plot $\hat{P}$ highlights the initial correction from the estimated position to the true position.

The convergence of $P$ and $\hat{P}$ can be more clearly seen in the figure 8.5. The asterisked points represent the location of a three point measurement collect and subsequent update. The plot picks up after the fourth point in the simulation and demonstrates how the plot converges on the true position.

![Overlaid 2-Dimensional Plot of P_Hat, P_True](image)

**Figure 8.5: True and Estimated Trajectory Convergence**
The initial error in position was over 180km between true and estimated position. Figure 8.6 depicts the 2-norm of the error between the estimate and true trajectory after the two converge to within two meters. In the simulation, the interval between radar navigation message receipt times was randomly varied from one to five seconds. Convergence occurred after receiving 60 radar (navigation) updates and approximately three minutes (173 seconds). The subfigure at right in figure 8.6 further expands a subsection of the error plots to more clearly see the difference. In these plots, the norm of the $x, y$ error is less than 0.5 meters. Individual plots of the $x$ and $y$ errors show, as expected, approximately the same error.

The 0.5 meter error can be explained, in part, by the contribution of altitude on the the two-dimensional solution. The time of arrival measurements are based on the three-dimensional signal travel time to the vehicle. The algorithm computes the two-dimensional solution. Although the contribution of altitude to the transit time $\tau$ is almost negligible as compared with the $x$ and $y$ position contribution, it does introduce a slight discrepancy. This error can be reduced by incorporating the aircraft’s altimeter reading and the known radar elevations into the computation for the $z$-axis.

Figure 8.6: Plot of Full and Expanded Error Norms After Convergence $\leq 2$ Meters
The simulation and algorithm demonstrate that the use of embedded navigation messages into a radar use can result in a viable alternative navigation adjunct capability in the presence of degraded GNSS navigation.

8.4 GDOP Considerations for the 3-D Algorithm

Initially a three-dimensional solution was sought for the geolocation problem using navigation information from four radars. The major issue with the three-dimensional solution is the effect of dilution of precision (DOP). As previously discussed, DOP is related to the geometry of the transmitters (radars) relative to the receiver. The preferred geometry is one where the angular separation between the transmitters and the receiver is maximized. As the angular separation decreases, the accuracy of the measurement decreases. DOP provides a qualitative assessment of this spatial geometry.

In initial simulations, the computed $x, y, z$ position had an error between the estimated and true trajectories that seemed to diverge with time. The simulation was run again and DOP measurements were computed using the indicated radar locations and the initial true position of the vehicle as shown in Table 8.1.

<table>
<thead>
<tr>
<th>DOP Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal DOP</td>
<td>1.51</td>
</tr>
<tr>
<td>Vertical DOP</td>
<td>7.58e+05</td>
</tr>
<tr>
<td>Geometric DOP</td>
<td>871</td>
</tr>
</tbody>
</table>

These are acceptable values for the horizontal $(x, y)$ component of DOP but the vertical component is extremely high. The high value of VDOP drives overall GDOP as can be seen in the Table 8.1.
The reason for the high VDOP is the poor vertical angular separation between the radars and the vehicle. In the simulation, the radars are ground based with a terrain elevation between 0.1 and 0.5 kilometers in elevation. The vehicle is assumed to be at a constant 1.0 kilometers in altitude so the difference in altitude is much smaller than the horizontal distances of the vehicle to the radars. At these ranges the difference in angular site angle from one radar to another is almost negligible. Going back to three-dimensional solutions, Eqn 8.19 we also see that these minimal angular changes contribute little to the z-axis solution since the z-component values of \( \alpha, \beta \) and \( \gamma \) are negligible as compared with \( z_{R4} \).

Since the z-component is interdependent with the \( x, y \) component, errors in the z-axis manifest themselves in related errors in \( x \) and \( y \). A separate implementation of the non-synchronous geolocation algorithm was run where the vertical position component was computed. Beginning with a true initial location of \([150 \ 300 \ 1]\) and an initial position estimate of \([50 \ 200 \ 1]\), the algorithm computed its updated estimated position at \([15.468 \ 264.087 \ -176e+03]\) after the first iteration. The poor vertical DOP is apparent as well as its effect on the \( x \) and \( y \) components.

Increasing vehicle altitude does not take care of the problem since the issue is the lack of angular difference between the line-of-sight elevation angle for each radar. An increase in altitude merely changes each radar’s elevation look-angle by the same amount.

In practice this result is consistent with terrestrial-based navigation such as LORAN. LORAN and other pseudorange systems do not compute the altitude component, instead the vehicle relies on its much more accurate altimeter. In doing this, the minimum number of required transmitters also goes from four to three and the computational complexity of the simulation is reduced.
Summary / Follow-On Research

9.1 Research Summary

This thesis has addressed the use of in-band intentional modulation on pulse (IMOP) of a combined radar/communications signal. For this study, the communications signal was a navigation message that could enable a regional, limited area navigation solution as an adjunct capability to existing global navigation capabilities such as GPS. The thesis began by decomposing the problem into two distinct areas: waveform development and a nonsynchronous navigation algorithm.

For the purposes of this research the linear frequency modulated (LFM) signal was used as the intermediate frequency (IF) carrier of a baseband navigation message. The baseband message modulated the LFM radar pulse with short duration phase changes (discretes). This mixed-modulated signal is then up-converted to a higher frequency carrier (typically C or S band) for transmission in order to take advantage of the higher power carrier wave to improve message transmission characteristics. Message performance is measured in terms of bit error rate (BER) at the receiver. Since a reduced magnitude phase angle \( \phi_\delta \ll \frac{\pi}{2} \) is used to designate a ‘1’ or ‘0’ bit, a revised plot of BER to bit energy to noise ratio was developed. Using this plot, the minimum bit energy to noise ratio is determined. The bit energy to noise ratio along with transmission distance and radar characteristics (power, gain, etc.) are used in the one-way radar range equation to determine
minimum bit duration (bit-chip length) for each message bit. The result of these calculations along with the LFM pulse sample rate and LFM pulse duration control maximum number of bits per pulse.

Radar performance is normally addressed through achieving a required target detection, $P_d$, while reducing Type I (false positive) and Type II (false negative) detection errors. Further, modern radars are generally evaluated in terms of how they handle time delays in range and Doppler frequency shift effects. For these reasons the radar ambiguity function was selected as the performance metric. The radar ambiguity function provides a three-dimensional depiction of the performance of a signal and its matched filter in the face of time delays and Doppler frequency changes. Further, at zero Doppler shift ($\nu = 0$) the ambiguity function is the autocorrelation of the signal and matched filter. Autocorrelation is the primary means of target detection over a range of time delays between the signal and the matched filter. Its result can illuminate sidelobe responses that may obscure actual targets (Type II error) or produce false detections (Type I error).

The main challenge with using the ambiguity function is that desired performance of the radar drives the desired shape of the plot; therefore, there is no ideal plot. Normally there is a tradeoff in radar range resolution and Doppler resolution. As a result, the objective metric is to minimize perturbation of the combined radar/communication ambiguity function from the baseline LFM radar ambiguity function plot.

The second area of interest was the implementation of a mixed radar / navigation signal in a functional navigation architecture. As has been pointed out, radar is normally a directional signal with most of its power in the radar main beam. As the receiver’s antenna moves out of the main beam, the signal energy and the baseband message bit energy decreases. This decrease in bit energy to expected Gaussian distributed noise can result in
increased BER.

The challenge with using traditional navigation algorithms is that they assume an omni-directional signal and near simultaneous message arrival at the receiver from several different transmitters. At a particular location and epoch, the receiver can use either pseudoranges (GPS) or time difference of arrival (TDOA) to calculate present position. Coupled with filtering techniques such as Kalman filtering, the effects of measurement and other source of noise can be minimized especially if there is some a priori understanding of the noise distribution, mean and variance characteristics.

For the directional radar there is no guarantee that the minimum number of radars required for a geolocation solution will illuminate the receiver at a particular epoch. Depending on radar density, a more likely scenario is one where radars illuminate the receiver with adequate energy for message receipt at times separated by several seconds or more. The literature is silent on this non-synchronous case, most likely because navigation signals are normally dedicated mission signals, designed to be somewhat omnidirectional. Therefore, large temporal arrival differences are not a factor.

To demonstrate a feasible implementation of this non-synchronous navigation architecture, a modified algorithm as developed that accounted for time delay of message arrival by incorporating vehicle movement between messages. Vehicle movement can be determined using an on-board inertial measurement unit (IMU) that records velocity and direction and translates that into position differences between message receipt points. This provides a relationship between all of the message receipt points so that the message receipt points can be expressed in terms of a single message receipt position, normally at the last message receipt location.
9.2 Waveform Development Summary

In this study the use of constant phase shift, Barker Codes and M-sequences were evaluated. In particular the use of small angle phase discrete in conjunction with the different encoding schemes were explored.

Based on the revised BER plot for small angle (minimum bit separation distance) phase angles and the one-way radar wave equation, it is feasible to obtain adequate BER at range for a navigation message. The main reasons for this are the higher transmit power of the carrier wave used to transmit the combined radar/navigation message and the one-way travel of the signal from the transmitter. One-way navigation signal travel means the received signal energy/power at the receiver is a $\frac{1}{r^2}$ multiple instead of a $\frac{1}{r^4}$ multiple.

The minimum bit chip length required to achieve a specific BER is a function of both the BER equation and the one-way radar link budget (Friis transmission) equation were both modified to include the reduction in phase magnitude from a traditional BPSK of $\pm \frac{\pi}{2}$. In addition the radar range equation was modified to determine required minimum bit chip length as a function of radar parametrics and the phase discrete magnitude. These results follow:

\[
[SNR]_{req} = Q\left(\sqrt{\frac{2E_b}{N_0}} \sin(\phi_{\delta})\right) \quad (9.1)
\]

\[
n_{\lambda} \geq \frac{[SNR]_{req}(4\pi r)^2 Lf_s \kappa T_0 F}{P_T G_T G_R \lambda^2} \quad (9.2)
\]

Equations (9.1) and (9.2) demonstrate that the minimum sample-bit length (bit period) can be explicitly determined in closed form as long as desired BER and certain radar parameters are known ahead of time. The main challenge is the effect on the radar performance.
9.2.1 Constant Phase Angle Encoding

In evaluating the constant phase angle encoding of the waveform, it was demonstrated that low magnitude phase angle changes allow both a satisfactory message receipt at the receiver while minimizing effects on the radar ambiguity function. An additional benefit of constant phase angles is that implementation is less complex than one where the phase magnitude is changing frequently from chip to chip.

9.2.2 Barker Encoding

Barker codes were considered in light of the correlation properties (high in-phase and low out-of-phase values) for better message detection and recovery. However, Matlab simulations showed higher sidelobes in the PSD and auto-correlation than was observed in the constant phase change approach. This is due to the fact that the phase is changing frequently but, because it is held constant for a short period of time for each Barker element, additional frequency contributions are created outside the main lobe of the LFM signal even with reduced phase angle magnitudes. In addition, Barker codes also suffer from the same lack of cross-correlation between Barker sequences that would allow for multiple signals in one radar pulse.

9.2.3 M-Sequences - BPSK

The research then moved on to M-sequences. Like Barker codes, M-sequences offer both excellent auto-correlation and out-of-phase cross-correlation attributes. In addition, M-sequences also have excellent cross-correlation properties with other ‘preferred pair’ M-sequences. This additional property opens the door to channelized messages in a single pulse using small angle phase changes.
Based on the Barker Code performance, it was initially thought that the constant binary phase shifts of an M-sequence would induce the same problems as the Barker Code. However, the radar ambiguity function, PSD and auto-correlation functions were all well behaved. Further, channelized signals using preferred pair M-sequences also performed well and permitted twice the signal content to be passed per pulse.

### 9.2.4 M-Sequences- QPSK

In an effort to increase signal density, an attempt was made to explore the use of QPSK encoding. In QPSK a two-bit symbol is encoded requiring four distinct points in the I-Q plane. Unfortunately, to maintain a constant bit separation distance, the phase angles quickly grew to a point where the radar ambiguity function and PSD were adversely affected. In order to reduce the phase angle requirements, the length of the minimum bit-chip length was extended but this had the effect of reducing the number of bits that could be encoded. In short, moving to QPSK increased complexity without increasing bit throughput.

### 9.3 Non-Synchronous Navigation Summary

The primary difference between a directional radar with an embedded, in-band, mixed-modulated navigation signal and traditional radio based navigation is the lack of an omnidirectional signal. Radar waveforms normally cover a narrow finite region with an energy/power profile that decreases rapidly outside the radar main beam. Navigation waveforms and transmitters are designed to ensure a receiver has continuous access to the navigation signal for geolocation purposes. Radio-based navigation algorithms assume a near simul-
aneous message receipt at the receiver from a specific number of geographically separated transmitters. This is not the case with a directional radar that may have a revisit rate of several seconds.

To overcome the delay between message receipt times, a revised algorithm was developed that incorporated the position change between message receipt points based on an on-board inertial measurement unit (IMU). The algorithm is based on the concept that there is a unique set of time difference of arrival and IMU separation between message receipt points. In addition, a restriction was placed on the periodicity of navigation message transmit times in order to provide an unambiguous relationship between message receipt and message transmission times.

The algorithm uses the time difference of arrival of navigation signal from different radars employing an iterative approach. An estimated message transit time, \( \tau_n \), is computed based on the distance between the estimated position and the location of the transmitting radar. Using the pseudo-time difference of arrival relationships, an estimate of the remaining estimated message transit time for the other radar and message receipt points can be determined. These values, along with the IMU recorded spatial separation between message receipt points, are used to compute the position of the vehicle at the last message receipt point. From there, the vehicle position at previous message receipt points can be determined using the IMU position changes. The revised final position is then used to update the current position of the vehicle as a starting point for the next geolocation iteration. Matlab simulations of the algorithm demonstrated rapid convergence of the estimated and true position even with large errors in the initial estimated position.

The main challenge in implementation is that the vehicle position change for any two message receipt points cannot be the same as the spatial separation of the transmits-
ters. When this occurs a singular matrix inversion results. The physical realization of this problem is when the radar to vehicle line-of-sight between any two message receipt points are both orthogonal to the vehicle flight path/trajectory. In this scenario there are an infinite number of allowable parallel solutions. Fortunately, this problem can be resolved by checking the condition of the matrix and, if the matrix is poorly conditioned for inversion, dropping some or all the measurements and obtaining new ones from different radars.

9.4 Future Work

As this dissertation is in two parts, the recommendations for future research is also in two parts.

From a waveform perspective, this research focused on the feasibility of using an LFM signal to act as an IF carrier for a baseband phase modulated signal. Follow-on research on other radar forms such as non-linear frequency modulation, Barker coded radar pulses and noise radar employing an embedded pseudonoise auxiliary signal is a logical extension of this work. Further, the effect of in-band modulation on other growing communications such as WiFi or other wide area communications is warranted.

Other areas for research include application of the low data rate message for dynamic spectrum access approaches to spectrum use as well as low probability of intercept/low probability of detection (LPI/LPD) communications. The main challenge to LPI/LPD is the use of cyclostationary processes that can be sued to detect the embedded message. As was seen in the PSD charts, for constant phase angle encoding, the message imparts a modulation on the power distribution as a function of frequency. Research on variable length communications that circumvents cyclostationary exploitation is an area of interest for future research.
From a navigation perspective, the algorithm was developed in order demonstrate the feasibility of using a directional, non-synchronous signal in a navigation architecture. This dissertation did not consider the effects of measurement nose and the use of Kalman filter type methods to improve the geolocation solution. Additional sensitivity analysis of the effect of spatial distribution of the radar and effect of extended time delays between message receipt points ware but two areas of future work.

A final area is the most challenging, implementation of this architecture in moving/flying pseudolites. Specifically, embedding navigation signals in aircraft radars in order to provide an auxiliary communications pathway would be of benefit. The problems associated with this application are diverse. Besides combined Doppler effects, transmitter position/velocity vector changes, and potential multipath errors, the optimal geometry for such a system is an sea of study. It may also be possible in this case to use the same transmitter for multiple signal updates within a geolocation update cycle for the receiver.
Bibliography


[26] V.B.S. Srilatha Indira Dutt, G. Sasi Bhushana Rao, S. Swapna Rani, Swarna Ravindra Babul, Rajkumar Goswami and Ch. Usha Kumari. Investigation of GDOP for Pre-
