A CTFM SYNTHETIC APERTURE SONAR

Michael P. Hayes, B.E. (Hons)

A thesis presented for the degree of Doctor of Philosophy in Electrical and Electronic Engineering, at the University of Canterbury, Christchurch, New Zealand.
September 1989.
ENGINEERING UNIVARY
THESIS

Abstract

This thesis describes the theory and operation of a seafloor imaging Synthetic Aperture Sonar (SAS), based on broadband Continuous Tone Frequency Modulation (CTFM).

Narrowband synthetic aperture techniques are reviewed, and some of the limitations of using such techniques in sonar applications are described. Aperture undersampling is a particular problem when towing single-beam sonars at realistic speeds. However, the mapping rate constraints may be relaxed by using broadband signals, but at the expense of increased self-clutter (or background).

One such broadband signal that is suitable for synthetic aperture operation is CTFM. The signal processing requirements for the CTFM signal are investigated, and are shown to be considerably simplified by decomposing each echo sweep into an ensemble of narrowband components. Images reconstructed from each of these components may be combined in a variety of ways. The relative merits of these differing methods are examined using a computer simulation of a side-scan CTFM sonar.

The temporal phase stability of the acoustic environment is vital to the formation of a synthetic aperture. An experiment was performed which indicates that the phase stability is much better than anticipated, and certainly adequate for the formation of undersea synthetic apertures.

This prediction was confirmed by another experiment in which the prototype CTFM sonar was moved along a fixed cableway under realistic operating conditions. Images of a test target (an air-filled steel buoy) were successfully reconstructed using data measured from this experiment.

Reconstructions obtained from a number of different algorithms are presented, for differing values of the various operating parameters. It is demonstrated that artefacts resulting from aperture undersampling are reduced by using broadband CTFM.
I sincerely thank my supervisor, Dr. Peter Gough ('the Chief'), for his advice, guidance, and enthusiasm during the course of my research. Maybe one day he too will go grey.

I am grateful to my family and friends for all their support and understanding, especially during the duration of my Ph.D. study.

A big thanks is due to all those people who helped in the preparation of this thesis, in particular Bruce McCallum and Catherine Watson for their proof-reading, Peter Gardenier for his typesetting suggestions, my sister Lynley for help in the scissors and sellotape department, and Denise Fastier for her continual encouragement. I would also like to thank Bruce for his colourful contribution to the Department, and for the many useful discussions I had with him.

Thanks are also due to my long suffering flatmates, who have had to put up with my guitar playing, the fire-crackers in the toaster, and the nocturnal nature of my attempts at thesis writing.

The assistance of the technical staff in the development and construction of the sonar is gratefully acknowledged; in particular Art Vernon, Dermot Sallis, and Michael Cusdin. Thanks are also due to Professor J. W. Roy Griffiths and his group at Loughborough Technical University, England, for their assistance with the testing of the sonar.

The project has been supported financially by the New Zealand Defence Scientific Establishment, the New Zealand University Grants Committee, and Marconi Underwater Systems Ltd.

I acknowledge the receipt of a New Zealand University Grants Committee Postgraduate Scholarship and financial assistance from Marconi Underwater Systems Ltd. for the period of my research in the U.K.

Lastly, but not least, I would like to humbly dedicate this thesis to all those whom I owe a beer.
# Contents

Abstract iii

Acknowledgements v

Preface xi

Glossary of Notation xv

Symbols xv

## 1 Preliminary Concepts

1.1 Complex Signal Notation 1

1.1.1 Pre-Envelope 1

1.1.2 Band-Pass Signals 2

1.1.3 Complex Sampling of Real Signals 3

1.1.4 Narrowband Signals 4

1.2 Echo Ranging Systems 4

1.3 Matched Filtering 5

1.4 Effects of Sonar/Target Motion 6

1.4.1 Moving Target/Stationary Sonar 7

1.4.2 Moving Sonar/Stationary Target 8

1.4.3 The Distorted Echo Signal 9

1.4.4 The Doppler Effect 9

1.4.5 Signal Tolerance to Sonar/Target Motion 11

1.4.6 Resolution on the Basis of Sonar/Target Motion 12

1.5 Ambiguity Functions 12

1.5.1 Multiple Target Resolution 13

1.6 The Sonar Equation 14

1.6.1 Pulse Integration 15

1.6.2 Pulse Compression 15

1.7 Reverberation 16

1.7.1 Volume Reverberation 17

1.7.2 Surface Reverberation 17

1.7.3 Signal to Reverberation 18

1.8 Undersea Acoustic Propagation 18

1.8.1 Attenuation 19

1.8.2 Multipath 19

1.8.3 Coherence 21
3.2 Demodulation of CTFM Echo Signals

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.1 Simple CTFM Demodulation</td>
<td>62</td>
</tr>
<tr>
<td>3.2.2 Quadrature Demodulation</td>
<td>64</td>
</tr>
<tr>
<td>3.2.3 Dual-Demodulation</td>
<td>65</td>
</tr>
<tr>
<td>3.2.4 The Effects of Target Motion on CTFM Signal Processing</td>
<td>68</td>
</tr>
</tbody>
</table>

3.3 Spectral Analysis of Demodulated CTFM Signals

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3.1 Range Resolution</td>
<td>71</td>
</tr>
<tr>
<td>3.3.2 Auditory Spectral Analysis</td>
<td>72</td>
</tr>
<tr>
<td>3.3.3 Analogue Spectral Analysis</td>
<td>72</td>
</tr>
<tr>
<td>3.3.4 Digital Spectral Analysis</td>
<td>72</td>
</tr>
</tbody>
</table>

3.4 CTFM Signal Generation

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.1 Analogue Methods</td>
<td>74</td>
</tr>
<tr>
<td>3.4.2 Digital Methods</td>
<td>75</td>
</tr>
</tbody>
</table>

3.5 Properties of CTFM Sonars

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5.1 Maximum Range</td>
<td>75</td>
</tr>
<tr>
<td>3.5.2 Range and Frequency Equalisation</td>
<td>76</td>
</tr>
<tr>
<td>3.5.3 Phase Correction</td>
<td>76</td>
</tr>
<tr>
<td>3.5.4 Reverberation</td>
<td>77</td>
</tr>
<tr>
<td>3.5.5 Crosstalk</td>
<td>77</td>
</tr>
<tr>
<td>3.5.6 Beamwidth Considerations</td>
<td>78</td>
</tr>
</tbody>
</table>

3.6 Comparison of Pulsed and CTFM Sonars | 78 |

4 The Prototype CTFM Synthetic Aperture Sonar

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Electronic Design</td>
<td>81</td>
</tr>
<tr>
<td>4.1.1 The CTFM Sonar Front-End</td>
<td>81</td>
</tr>
<tr>
<td>4.1.2 Data Storage</td>
<td>84</td>
</tr>
<tr>
<td>4.1.3 Computer</td>
<td>85</td>
</tr>
<tr>
<td>4.2 Pre-Processing of the Raw Sonar Data</td>
<td>86</td>
</tr>
<tr>
<td>4.2.1 Chopping up the Sweep</td>
<td>86</td>
</tr>
<tr>
<td>4.2.2 Range Compression</td>
<td>88</td>
</tr>
<tr>
<td>4.2.3 Phase Response</td>
<td>88</td>
</tr>
<tr>
<td>4.2.4 Practical Considerations</td>
<td>89</td>
</tr>
<tr>
<td>4.3 CTFM Phase Monopulse</td>
<td>90</td>
</tr>
<tr>
<td>4.4 Summary</td>
<td>91</td>
</tr>
</tbody>
</table>

5 Measurement of the Acoustic Temporal Phase Stability

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Background</td>
<td>93</td>
</tr>
<tr>
<td>5.2 The Effect of Noise on Phase Accuracy</td>
<td>94</td>
</tr>
<tr>
<td>5.3 Experimental Details</td>
<td>95</td>
</tr>
<tr>
<td>5.4 Experimental Results</td>
<td>97</td>
</tr>
<tr>
<td>5.4.1 Long Term Fluctuations</td>
<td>99</td>
</tr>
<tr>
<td>5.5 Summary</td>
<td>100</td>
</tr>
</tbody>
</table>

6 Simulation and Reconstruction

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Simulation of a CTFM Synthetic Aperture Sonar</td>
<td>101</td>
</tr>
<tr>
<td>6.1.1 Projector Radiation Pattern</td>
<td>101</td>
</tr>
<tr>
<td>6.1.2 Hydrophone Response</td>
<td>103</td>
</tr>
</tbody>
</table>
Preface

The applications of side-scan sonar are diverse. Applications include engineering, geological and hydrographic surveys, mineral exploration, detection of gas leaks, underwater archaeology, cable and pipeline location, oil rig surveys, studies of iceberg scouring, mine hunting, searches for the Loch Ness Monster, and the location of sunken objects such as ships, planes, torpedoes, and mislaid sonars.

Imaging sonars can be subdivided into two main categories; real aperture and synthetic aperture. Real aperture sonars obtain a high azimuth resolution by radiating very narrow beams, and therefore require the use of large apertures and high acoustic frequencies. Moreover, the higher the radiated frequency, the greater the absorption, hence high resolution imaging sonars using real apertures are restricted to short operating ranges.

Synthetic aperture sonars use a slightly different technique to produce side-scan images. A much wider beam is employed and by coherently combining returns from along the aperture, an effective narrow beam is produced. Unlike traditional imaging techniques, synthetic aperture systems achieve a high resolution using small real apertures.

Synthetic aperture methods are widely recognised, particularly in the fields of airborne and spaceborne radar. However, the application of these techniques to sonar is not widespread, and is hampered by slow mapping rates, coherence problems relating to signal propagation through an unhomogeneous medium, and the difficulty of maintaining a straight towfish trajectory.

There have been many different CTFM (Continuous Transmission Frequency Modulation) sonars built at the University of Canterbury. These sonars were originally designed as aids for blind or visually impaired people. The advantage of CTFM sonars in this application, is that they provide a simple auditory display where the range to a target is coded by frequency. These air-sonars were then extended into underwater applications: firstly fish finding sonars, a diver’s sonar, then a narrow beam side-scan sonar designed for target classification.

The work reported in this thesis began in 1985 when I enrolled for the Masters degree course under Dr. Peter Gough (in the Department of Electrical Engineering at the University of Canterbury, Christchurch, New Zealand). Experience with the CTFM classification sonar, showed that even with a continuous target illumination, it was easy to ‘miss’ targets because of the narrow beamwidths required for high angular resolution (de Roos, 1986). Peter was therefore interested in applying widebeam synthetic aperture techniques to overcome this problem. Peter also believed that some of the difficulties associated with the application of synthetic aperture techniques to sonar was due in part, to the use of short acoustic pulses (Gough, 1986).
He therefore initiated the construction of a CTFM synthetic aperture sonar, the operation and results of which are described in this thesis.

The thesis is written in eight chapters. Chapter 1 presents an introduction to undersea imaging with an emphasis on side-scan sonars, and is intended to introduce the terminology, notation, and concepts that are used in subsequent chapters of this thesis. However, much of this material is well known and only dealt with briefly. The chapter starts with an introduction to sonar signals, and covers the concepts of echo ranging (in particular with regards to the effects of target or sonar motion), matched filtering, and ambiguity diagrams. Other topics include the sonar equation, reverberation, undersea acoustic propagation, and target scattering characteristics. The chapter concludes with a review of the properties of undersea imaging systems.

Chapter 2 comprises an introduction to the principles of synthetic aperture processing. This is based largely on the application of Synthetic Aperture Radar (SAR) techniques to imaging sonars, and is written with a sonar flavour. The chapter also looks at the extension of broadband and noncoherent imaging techniques to synthetic aperture processing.

The properties of CTFM sonars are reviewed in Chapter 3. Various methods of signal generation and echo demodulation are presented. Particular attention is paid to the preservation of the echo phase, and the effects of relative sonar/target motion on the CTFM echo signals.

The prototype CTFM synthetic aperture sonar is introduced in Chapter 4, and includes details of the pre-processing performed on the raw data. Particular emphasis is placed on the division of the broadband CTFM sweep into a number of narrowband linear FM components, and the subsequent range compression of these components.

The ocean is often regarded as a ‘random’ inhomogeneous medium, and it has been felt that this will prevent any attempt at synthesising an acoustic aperture in the ocean. Chapter 5 describes an experiment to assess the temporal stability of the sea using the prototype CTFM sonar described in Chapter 4. Note that this chapter is not a study of turbulence, but a measurement to assess the path length stability under actual conditions.

Chapter 6 outlines a simple computer simulation of a CTFM synthetic aperture sonar operating under ideal conditions, and the algorithms used to reconstruct images from both the simulated and measured synthetic aperture data. Possible enhancements to the algorithms are also discussed, as are the effects of phase errors and tow speed errors.

Chapter 7 presents synthetic aperture images reconstructed from data gathered at a sonar range in Loch Linhe, Scotland. These are reconstructed using a number of different algorithms introduced in Chapter 6. The reduction of coherent artefacts using broadband CTFM signals is clearly illustrated.

The thesis concludes in Chapter 8 with a summary of the main results and suggestions for further research and modifications to enhance the operation of the prototype sonar.

While the first three chapters are introductory, the remainder of the thesis is, by and large, original material.

Construction of the sonar commenced in 1985. The transducers for the sonar were designed and tested by Peter Gough and John Knight, and constructed by Art
Vernon. The details of the transducer design can be found in John's Masters thesis (Knight, 1987). Art also built the towfish assembly, the cable and winch system, and numerous acoustic baffles. Dermot Sallis helped in the testing of these baffles; a rather frustrating experience!

The electronics were designed by myself and Dermot Sallis, with contributions from Peter Gough and Michael Cusdin. Dermot built most of the electronics, with contributions from Wiktor Mencel and Richard Cox (test-module). Mention must also be made of the sturdy ‘yellow boxes’ built by Jon Jongens for housing of the sonar sub-systems.

I built the interface between the sonar and the Intel 310 computer system, programmed the test-module, and wrote all the interface software. The test-module software was written in the Motorola 6805 assembly language, and the data collection and pre-processing software was written in Pascal-86 and ASM-86 (the Intel macro assembly language) using the iRMX86 operating system.

I was fortunate enough to accompany Peter to the U.K. for the latter half of 1987 during his sabbatical leave. While in Scotland we obtained most of the experimental data that is presented in this thesis. This was collected over one cold, wet week in early December, in the waters of Loch Linhe at the end of the Fort William pier.

All the image reconstruction algorithms were developed and implemented on the EEE Department’s cluster of VAX computers. I wrote the programs in VAX-11 Pascal using the improc (image processing) utility developed principally by Richard Lane, among others. All the graphical output has been produced using the PLOT79 plotting package using an interface to improc written by Peter Gardenier.

In this thesis I decided to present the images as hidden-line drawings. Since they are not true images they are referred to in the text as image distributions. These hidden-line drawings were found to be more useful than grey-scale images for the display of the reconstructed image data, especially with regards to resolution and dynamic range. Except where explicitly stated, all the image distributions are of the image intensity with either a linear or logarithmic intensity scaling. No other form of image compression is employed.

Papers and presentations prepared during the course of this thesis are listed below in the approximate order of preparation.


Glossary of Notation

The following conventions are adhered to in this thesis:

(i) Temporal signals and spatial distributions are denoted by lower case Roman characters, e.g. \( s(t) \), and their Fourier transforms are denoted by the corresponding upper case characters, e.g. \( S(f) \).

(ii) Complex quantities (having both amplitude and phase) are printed in bold Roman type, e.g. \( A \). This notation also applies to complex fields and signals.

(iii) Bandpass signals are adorned by a tilde, e.g. \( \tilde{s}(t) \).

(iv) Vector quantities are indicated by a superscript arrow, e.g. \( \vec{x} \).

(v) Peak quantities and quadrature signal components are denoted by a superscript caret, e.g. \( \hat{P}_T \), \( \hat{s}(t) \).

(vi) An asterisk denotes the complex conjugate, e.g. \( s^*(t) \).

(vii) Time derivatives are denoted by superscript dots, e.g. \( \dot{R}(t) \), \( \ddot{R}(t) \).

Symbols

\[
\begin{align*}
A_h & \quad (m^2) \quad \text{effective hydrophone area} \\
A_F & \quad (m^2) \quad \text{effective projector area} \\
\dot{A} & \quad (m^2s^{-1}) \quad \text{area mapping rate (one side only)} \\
B & \quad (Hz) \quad \text{signal bandwidth} \\
B_N & \quad (Hz) \quad \text{noise bandwidth} \\
c & \quad (m s^{-1}) \quad \text{speed of propagation} \\
d(t) & \quad \text{demodulated echo signal} \\
D \ell & \quad (dB) \quad \text{receiver directivity index} \\
c(t) & \quad \text{received echo signal} \\
E(f) & \quad \text{received echo spectrum} \\
E & \quad (J) \quad \text{signal energy} \\
E \ell & \quad (dB) \quad \text{echo level} \\
f & \quad (Hz) \quad \text{frequency} \\
f_d(t) & \quad (Hz) \quad \text{instantaneous demodulated frequency} \\
f_d(t) & \quad (Hz) \quad \text{instantaneous Doppler frequency} \\
f_e & \quad (Hz) \quad \text{equivalent frequency} \\
f_f(t) & \quad (Hz) \quad \text{instantaneous frequency} \\
f_R(t) & \quad (Hz) \quad \text{instantaneous received frequency} \\
f_s & \quad (Hz) \quad \text{sampling frequency} \\
f_T(t) & \quad (Hz) \quad \text{instantaneous transmitted frequency} \\
F_h & \quad \text{hydrophone beam pattern} \\
F_P & \quad \text{projector beam pattern}
\end{align*}
\]
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_s$</td>
<td>synthetic beam pattern</td>
</tr>
<tr>
<td>$G_h$</td>
<td>hydrophone gain</td>
</tr>
<tr>
<td>$G_p$</td>
<td>projector gain</td>
</tr>
<tr>
<td>$I_0$</td>
<td>(W m$^{-2}$) reference intensity ($0.667 \times 10^{-18}$ W m$^{-2}$)</td>
</tr>
<tr>
<td>$I_t(n)$</td>
<td>integration improvement factor for $n$ pulses</td>
</tr>
<tr>
<td>$IF_n$</td>
<td>(dB) pulse integration improvement factor</td>
</tr>
<tr>
<td>$IF_p$</td>
<td>(dB) pulse compression improvement factor</td>
</tr>
<tr>
<td>$k$</td>
<td>(m$^{-1}$) wavenumber $= 2\pi/\lambda$</td>
</tr>
<tr>
<td>$L_h$</td>
<td>(m) hydrophone length</td>
</tr>
<tr>
<td>$L_p$</td>
<td>(m) projector length</td>
</tr>
<tr>
<td>$L_r$</td>
<td>(m) length of real aperture</td>
</tr>
<tr>
<td>$L_s$</td>
<td>(m) length of synthetic aperture</td>
</tr>
<tr>
<td>$M$</td>
<td>number of frames per sweep</td>
</tr>
<tr>
<td>$N$</td>
<td>(W) noise power</td>
</tr>
<tr>
<td>$N_i$</td>
<td>number of integrated pulses</td>
</tr>
<tr>
<td>$N_0$</td>
<td>(W m$^{-2}$Hz$^{-1}$) isotropic sea noise intensity</td>
</tr>
<tr>
<td>$NL$</td>
<td>(dB) isotropic noise level</td>
</tr>
<tr>
<td>$p$</td>
<td>(Pa) acoustic pressure</td>
</tr>
<tr>
<td>$prf$</td>
<td>(Hz) pulse repetition frequency</td>
</tr>
<tr>
<td>$P_T$</td>
<td>(W) peak transmitted power</td>
</tr>
<tr>
<td>$q$</td>
<td>(Neper/m) attenuation coefficient</td>
</tr>
<tr>
<td>$r_0$</td>
<td>(m) reference range (1 m)</td>
</tr>
<tr>
<td>$R$</td>
<td>(m) slant range to target</td>
</tr>
<tr>
<td>$R_e$</td>
<td>(m s$^{-1}$) range rate</td>
</tr>
<tr>
<td>$R_a$</td>
<td>(m s$^{-2}$) range acceleration</td>
</tr>
<tr>
<td>$R_u$</td>
<td>(m) maximum unambiguous range</td>
</tr>
<tr>
<td>$R_0$</td>
<td>(m) nearest slant range to target</td>
</tr>
<tr>
<td>$RL_s$</td>
<td>(dB) surface reverberation level</td>
</tr>
<tr>
<td>$RL_v$</td>
<td>(dB) volume reverberation level</td>
</tr>
<tr>
<td>$s_s$</td>
<td>surface reverberation cross-section per unit area</td>
</tr>
<tr>
<td>$s_v$</td>
<td>volume reverberation cross-section per unit volume</td>
</tr>
<tr>
<td>$S$</td>
<td>(W) peak signal power</td>
</tr>
<tr>
<td>$SL$</td>
<td>(dB) projector source level</td>
</tr>
<tr>
<td>$SNR$</td>
<td>(dB) signal to noise ratio</td>
</tr>
<tr>
<td>$SRR$</td>
<td>(dB) signal to reverberation ratio</td>
</tr>
<tr>
<td>$S_s$</td>
<td>(dB/m$^2$) surface scattering strength</td>
</tr>
<tr>
<td>$S_v$</td>
<td>(dB/m$^3$) volume scattering strength</td>
</tr>
<tr>
<td>$S_x$</td>
<td>(m) swath width</td>
</tr>
<tr>
<td>$S_y$</td>
<td>(m) swath length</td>
</tr>
<tr>
<td>$s(t)$</td>
<td>transmitted signal</td>
</tr>
<tr>
<td>$S(f)$</td>
<td>transmitted signal spectrum</td>
</tr>
<tr>
<td>$t$</td>
<td>(s) time (when signal received)</td>
</tr>
<tr>
<td>$t_e$</td>
<td>(s) time since start of received sweep</td>
</tr>
<tr>
<td>$t_r$</td>
<td>(s) instant of signal reflection</td>
</tr>
<tr>
<td>$t_s$</td>
<td>(s) time since start of transmitted sweep</td>
</tr>
<tr>
<td>$t_0$</td>
<td>(s) reference time</td>
</tr>
</tbody>
</table>
\( T \) (s)  sweep or pulse repetition period
\( T_b \) (s)  blanking interval
\( T_c \) (s)  compressed pulse length
\( T_d \) (s)  dwell time (time in beam)
\( T_p \) (s)  uncompressed pulse length
\( T_s \) (s)  sampling interval
\( TL \) (dB)  transmission loss (one-way)
\( TS \) (dB)  target strength
\( v \) (m s\(^{-1}\))  sonar along-track speed
\( W_d \) (Hz)  Doppler bandwidth
\( x \) (m)  target cross-track coordinate
\( y \) (m)  target along-track coordinate
\( y_s \) (m)  sonar track position
\( y_0 \) (m)  position of sonar corresponding to \( R_0 \)
\( z \) (kg m\(^2\) s\(^{-1}\))  specific acoustic impedance = \( \rho c \)
\( \alpha \) (m\(^{-1}\))  attenuation coefficient (\( \alpha = 10^\alpha / \ln 10 \))
\( \Delta f \) (Hz)  frequency bin width
\( \Delta R \) (m)  range bin width
\( \Delta r \) (m)  range resolution
\( \Delta x \) (m)  cross-track pixel size
\( \Delta x \) (m)  cross-track resolution
\( \Delta y \) (m)  along-track pixel size
\( \Delta y \) (m)  along-track resolution
\( \eta \)  Doppler time scale factor
\( \theta \) (radians)  azimuth angle
\( \theta_B \) (radians)  3 dB azimuth beamwidth
\( \theta_r \) (radians)  3 dB array beamwidth
\( \theta_s \) (radians)  3 dB synthesised beamwidth
\( \Theta \) (radians)  effective azimuth beamwidth
\( \lambda \) (m)  wavelength \( \lambda = c/f \)
\( \mu \) (s\(^{-2}\))  sweep rate
\( \rho_a \) (kg m\(^{-3}\))  ambient density
\( \sigma_x \) (m\(^2\))  target backscattering cross-section
\( \tau \) (s)  propagation delay
\( \phi \) (radians)  phase
\( \phi_e \) (radians)  phase error
\( \Phi \) (radians)  phase
\( \Phi_e (t) \) (radians)  instantaneous demodulated phase
\( \psi \) (radians)  elevation angle
\( \psi_B \) (radians)  3 dB elevation beamwidth
\( \omega \) (s\(^{-1}\))  angular frequency = 2\( \pi f \)
\( \Omega \) (radians\(^2\))  effective solid beamwidth

NB. The glossary is not complete, and only includes symbols that apply to more than one chapter.
Chapter 1

Preliminary Concepts

The purpose of this chapter is to introduce the terminology, notation, and concepts that are used in subsequent chapters of this thesis. Much of this material is well known and is dealt with briefly. Readers who are familiar with this material can go directly to Chapter 2 where the concepts more specific to this thesis are introduced.

The complex signal notation is used throughout the thesis, and this is introduced in Section 1.1 as a means of simplifying the analysis of real signals (and fields). The concept of echo ranging (in particular with regards to the effects of target or sonar motion, matched filtering, and ambiguity diagrams) is then described in Sections 1.2 through 1.5.

The sonar equation for noise limited sonars is introduced in Section 1.6, and this is followed by an introduction to pulse integration and pulse compression techniques as a means of improving detection. The sonar equation is then extended in Section 1.7 to include the effects of reverberation.

The remainder of the chapter looks briefly at the vagaries of acoustic propagation (Section 1.8) and backscattering from targets (Section 1.9), and concludes with a discussion of the limitations and properties of undersea imaging systems (Section 1.10).

1.1 Complex Signal Notation

All echo-ranging waveforms can be described in terms of a real function \( s(t) \) having an amplitude function \( a(t) \) and a phase function \( \Phi(t) \), where

\[
s(t) = a(t) \cos \Phi(t)
\]  

(1.1)

When \( a(t) \) is a slowly varying function, it is referred to as the signal envelope. This is usually a rectangular time function, in which case the instantaneous signal frequency \( f_i(t) \) is found from the time differential of the phase function to be

\[
f_i(t) = \frac{1}{2\pi} \frac{d\Phi(t)}{dt}
\]  

(1.2)

This also assumes that the oscillation period changes slowly between cycles, so that it is possible to find a CW signal \( \cos(2\pi f_i t) \) which, at any instant, fits the waveform.

These waveforms are usually band-pass signals, i.e. the spectrum is non-negligible only in a band of frequencies of extent \( \Delta f \), say, centred about some frequency \( \pm f_c \).
To reflect this property, the phase function $\Phi(t)$ is usually split into a carrier term of frequency $f_c$, and a phase modulation term $\phi(t)$ so that

$$\Phi(t) = 2\pi f_c t + \phi(t) \quad (1.3)$$

### 1.1.1 Pre-Envelope

An important property of real signals is that they possess spectra whose real part is even and imaginary part odd, i.e. $S(-f) = S^*(f)$. Hence only one half of the spectrum is required to fully specify a signal waveform. This allows a simplified signal notation, where the real signal $s(t)$ is replaced by a related complex signal $\hat{s}(t)$, with a spectrum $S(f)$ containing only positive frequency components. Of particular importance is the following choice of $s(t)$ as first described by Gabor (1946, p432):

$$s(t) = s(t) + j\hat{s}(t) \quad (1.4)$$

where the imaginary part $\hat{s}(t)$ is chosen to be the Hilbert transformed version of the real signal $s(t)$,

$$\hat{s}(t) = \mathcal{H}\{s(t)\} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\xi)}{t - \xi} d\xi \quad (1.5)$$

Note that this is an improper integral and the Cauchy principal value must be taken to account for the singularity at $t = \xi$ (Kreyszig, 1979, p734). The signal $s(t)$ defined in this above manner is known as an analytic signal and is also referred to as the pre-envelope of $s(t)$ (Haykin, 1983, §2.13). The convenience of this complex signal representation is that the spectrum of $s(t)$ vanishes for negative frequencies.

$$S(f) = \begin{cases} 2S(f), & f > 0 \\ S(f), & f = 0 \\ 0, & f < 0 \end{cases} \quad (1.6)$$

Theoretical analyses involving linear operations are greatly simplified by Gabor’s definition of $s(t)$. For example, the convolution of an analytic signal with the complex conjugate of another is always zero, and similarly, the correlation integral of two analytic signals, or the conjugates of two analytic signals, is also zero (Rihaczek, 1967, p707). For example, the energy of the signal $s(t)$, defined in the real signal notation as

$$E = \int_{-\infty}^{\infty} s(t)^2 dt \quad (1.7)$$

can be evaluated much more conveniently using the complex notation expression

$$E = \frac{1}{2} \int_{-\infty}^{\infty} |s(t)|^2 dt \quad (1.8)$$

where the factor of $\frac{1}{2}$ arises from the fact that although there are no negative frequency components in $S(f)$, the amplitudes of the positive frequency components are double that of $S(f)$.

The biggest problem with using Gabor’s analytic signal is the difficulty of expressing $\hat{s}(t)$, the Hilbert transform of $s(t)$, in a closed form (Cook and Bernfeld, 1967,
p63). However, if $s(t)$ is a narrowband signal, the pre-envelope $s(t)$ can be approximated as

$$s(t) \approx a(t) \exp(j\Phi(t))$$  \hspace{1cm} (1.9)$$

This representation is known as an exponential signal in contrast to the true analytic signal defined by (1.4) (Rihaczek, 1969, p23). The difference being that the analytic signal has only positive frequency components whereas an exponential signal may have some spillover of its spectrum into the negative frequency region. All practical waveforms have some spillover because they are always time limited functions. However, the error in using the approximation of (1.9) to the analytic signal is usually small provided the ratio of the signal bandwidth to the carrier frequency $B/f_c$ is small (Cook and Bernfeld, 1967, p63). This is nearly always the case for radar signals, but sonar signals often have much larger fractional bandwidths and more care must be taken in treating these signals.

Figure 1.1. (a) Amplitude spectrum of bandpass signal $s(t)$, (b) Amplitude spectrum of pre-envelope $s(t)$, (c) Amplitude spectrum of complex envelope $\tilde{s}(t)$.

1.1.2 Band-Pass Signals

Another simplification to the signal notation is to treat the band-pass signal $s(t)$ in terms of a constant carrier signal and a low-pass function $\tilde{s}(t)$. The choice of carrier frequency is purely arbitrary, but is usually chosen so that the first moment of one side of the energy density spectrum, $|S(f)|^2$, with respect to $f_c$ is zero.

$$s(t) = \Re\{\tilde{s}(t) \exp(j2\pi f_c t)\}$$  \hspace{1cm} (1.10)$$

The low-pass function $\tilde{s}(t)$ is called the complex envelope of $s(t)$ and completely represents the information content of the signal $s(t)$, irrespective of the carrier frequency. It is derived simply by translating the pre-envelope to baseband by multiplying $s(t)$ with the complex conjugate carrier signal.

$$\tilde{s}(t) = s(t) \exp(-j2\pi f_c t)$$  \hspace{1cm} (1.11)$$

Alternatively, the complex envelope may be expressed in the complex exponential form as

$$\tilde{s}(t) = a(t) \exp(j\phi(t))$$  \hspace{1cm} (1.12)$$

where the envelope $a(t)$ equals the amplitude of the complex envelope $\tilde{s}(t)$ and that of the pre-envelope $s(t)$, and $\phi(t)$ is the phase modulation function. In general, $\tilde{s}(t)$ is a
complex-valued quantity except when the signal \( s(t) \) is purely amplitude modulated. Also note that because the complex envelope contains negative frequencies, as shown in Figure 1.1, it is not an analytic signal.

1.1.3 Complex Sampling of Real Signals

If \( s(t) \) is a real signal bandlimited to the frequency interval \((-W, W)\), it is well known from the sampling theorem that \( s(t) \) can be completely described by samples at intervals of \( 1/2W \) (Bracewell, 1978, p194). Furthermore, because the spectrum of the complex pre-envelope \( s(t) \) is bandlimited to the interval \((0, W)\), the real signal \( s(t) \) can also be completely described by complex samples at intervals of \( 1/W \) (Woodward, 1964, p41).

1.1.4 Narrowband Signals

Cook and Bernfeld (1967, p61) define a narrowband waveform as one which has the property

\[
S_+(f) = 0, \quad f < 0 \tag{1.13}
\]
\[
S_-(f) = 0, \quad f > 0 \tag{1.14}
\]

where

\[
S_+(f) = \int_{-\infty}^{\infty} s(t) \exp(-j2\pi ft) \, dt \tag{1.15}
\]
\[
S_-(f) = \int_{-\infty}^{\infty} s^*(t) \exp(-j2\pi ft) \, dt \tag{1.16}
\]

In other words, a narrowband signal is one which has no spillover of positive frequency components into the negative frequency region and vice-versa. All sonar signals in practice have some spectral spillover because they are time limited functions, but provided the carrier frequency \( f_c \) is much larger than the modulation bandwidth \( B \), there is minimal spillover (Rihaczek, 1967, p707).

The narrowband definition, (1.14), can also be shown to apply to large time-bandwidth product linear FM chirps, provided the rate of change of frequency is slow. These signals can have extremely large fractional bandwidths \( B/f_c \), but because the instantaneous frequency changes slowly between cycles, there is minimal spectral spillover. Therefore, the complex exponential representation of these signals is a good approximation to the analytic form, although the interpretation of a modulated carrier becomes difficult. However, even though these signals obey the above narrowband definition, they cannot always be treated as such; in particular with regard to the approximation of the Doppler effect as a simple frequency translation (see Section 1.4.4).

1.2 Echo Ranging Systems

Echo ranging systems operate on the principle of transmitting sound energy in the direction of the target and receiving a portion of the transmitted energy which is reflected back from the target. The difference between the time of transmission...
of the sound energy and the time of reception of the reflected energy provides a measure of the range of the target, given the speed of sound propagation in the medium. The fraction of the received backscattered energy provides a measure of the target strength as determined by the sonar range equation (see Section 1.6).

If $s(t)$ is the transmitted signal, the signal $e(t)$ received as an echo from a single isolated target is given by

$$e(t) = As(t - \tau)$$

where $A = A \exp j\phi$ is a complex scale factor, and $\tau$ is the propagation delay. Here it has been assumed that the medium is non-dispersive and that there is no relative motion between the target and the sonar. The effect of such a motion on the received signal is studied in Section 1.4, but for now the propagation delay $\tau = 2R/c$ is considered to be a constant proportional to the target range $R$. Thus the echo signal can be treated simply as a delayed replica of the transmitted signal.

In general, the amplitude scale factor $A$ and phase shift $\phi$ are functions of the transmitted frequency and target aspect. However, this frequency dependence is ignored for the present.

### 1.3 Matched Filtering

All received signals consist of a desired echo component and an undesired noise component resulting from electrical noise introduced in the receiver and ambient acoustical noise incident at the hydrophone. There is also a component consisting of reverberation from unwanted scatterers which can mask the wanted echo signals. Since this reverberation is coherently related to the transmitted signal, it is the principle form of interference in an active sonar system (Winder, 1975, pp291–292). The effect of reverberation is described in Section 1.7 and is ignored for the present.

For reliable target detection, it is necessary to maximise the signal to noise ratio of the received signal by using filtering and/or correlation techniques. When the signal is corrupted by additive white Gaussian noise, the optimum filter is the matched-filter (Deley, 1970, §3.2). Even when the noise is not Gaussian, the matched-filter is still the optimum linear filter.

North (1963, p1021) shows that the transfer function of this filter has the form

$$H(f) = k \exp(-j2\pi f \tau_d) W^*(f)$$

where $W^*(f)$ is the complex conjugate spectrum of the expected echo signal. The corresponding impulse response of this filter is

$$h(t) = kw^*(\tau_d - t)$$

which, aside from an arbitrary gain $k$ and a time delay $\tau_d$, has the same shape as the expected echo signal $w(t)$, but with the time axis reversed. For this reason, the filter is said to be matched to the transmitted signal. The time delay $\tau_d$ has no real significance except that it must be greater than the duration of the input signal for realisability of the filter. In other words, the peak response of a matched-filter must occur after all of the input signal has entered the filter.
The matched-filter output response to an echo signal \( e(t) \) is, using (1.19),

\[
d(t) = \frac{1}{2} \int_{-\infty}^{\infty} h(\xi) e(t - \xi) d\xi \\
= \frac{k}{2} \int_{-\infty}^{\infty} w^*(\xi + \tau_d - t) e(\xi) d\xi
\]

and setting \( k = 1 \), the matched-filter output response is seen to be the crosscorrelation of the actual and expected echo signals. For this reason, the matched-filter is mathematically identical to a correlation receiver, although considerably different in implementation.

The peak output response of the matched-filter occurs at \( t = \tau_d \) and is proportional to the signal energy \( E \). If the two-sided noise spectral density is \( N_0/2 \), the peak signal to noise ratio at the output of the matched-filter is \( 2E/N_0 \). Note that this is independent of the signal waveform and depends only on the signal energy and noise power density (Cook and Bernfeld, 1967, p22).

### 1.4 Effects of Sonar/Target Motion

The round-trip propagation delay is the sum of the outward delay of the transmitted signal to the target and the return delay of the back-scattered signal. When the sonar is stationary, with respect to the medium, the outward and return delays are equal, independent of target motion, provided that the echo signal follows the same acoustic path as the transmitted signal. However if the sonar moves during the signal propagation delay, the outward and return delays are likely to be different. This has a greater implication for sonar than radar because the ratio of boat speeds to the speed of acoustic propagation is usually many times greater than the ratio of aeroplane or satellite speeds compared to the speed of electromagnetic propagation. As a consequence, many of the approximations valid in radar are not appropriate in similar sonar applications.

There is another fundamental difference between sonar and radar echo ranging resulting from the difference between acoustic and electromagnetic propagation. With electromagnetic waves, the speed of propagation is the same for all observers, even when in relative motion (Sears et al., 1982, §23-7). This is one of the postulates of Einstein’s *Special Theory of Relativity*. With acoustic propagation, however, it is possible to distinguish absolute motion with respect to the medium that carries the waves (Krane, 1983, p28). As a consequence, different results are obtained if, say, a sonar is moving toward a target, than if the target was moving toward the sonar, even if the relative velocity is the same in both cases. With electromagnetic propagation, however, the two cases are indistinguishable since only relative motion can be determined. To complicate matters further, the direct line of sight between the sonar and target is not necessarily the quickest acoustic path. Therefore to avoid these additional problems in the following analysis, the medium is assumed to be stationary and the signals are assumed to follow the direct (geometric) path.
1.4.1 Moving Target/Stationary Sonar

To study the effects of target motion on the received echo signal, consider a signal, \( s(t) \), transmitted at \( t = 0 \), reflected from a target with arbitrary motion and then received at \( t = \tau(t) \). The instantaneous propagation delay can be seen from Figure 1.2 to be given by the functional relation

\[
\tau(t) = \frac{c}{2} R \left( t - \frac{\tau(t)}{2} \right)
\]

where \( R(t) \) is the instantaneous target range and \( R(t - \tau(t)/2) \) is the range at the instant of reflection. Provided that both the sonar and the medium are stationary, (1.22) applies to any target motion, with the additional assumptions that the speed of sound, \( c \), is constant over the signal path and that the outward and return delays are equal. Note that if the medium is moving, the outward and return delays are generally no longer equal, and it is then necessary to compensate for the difference.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.2}
\caption{Graphical illustration of the propagation delay measured for a target with arbitrary motion.}
\end{figure}

The propagation delay function (1.22) is usually solved by expanding in a Taylor series about some arbitrary reference point, \( t_0 \),

\[
\tau(t) = \tau(t_0) + \dot{\tau}(t_0)(t - t_0) + \ddot{\tau}(t_0)(t - t_0)^2/2! + \cdots
\]

where \( \dot{\tau} \) and \( \ddot{\tau} \) are the first and second time derivatives of \( \tau(t) \) at \( t_0 \) respectively. The advantage of this series representation is that the higher order terms can be dropped to simplify the analysis, provided that the propagation delay, \( \tau(t) \), is a relatively smooth function of time over the duration of the signal (Kelly and Wishner, 1965, p57). This is often the case except for rapidly accelerating targets and long duration signals (see Section 1.4.5).

The coefficients of (1.23) can be expressed in terms of range and its derivatives, specified at \( t = t_0 - \tau(t_0)/2 \). Whatever the choice of the reference point, \( t_0 \), it defines the instant at which the values of delay and its derivatives are taken. This is important if the target range is changing while the signal is being reflected. Usually the reference point is chosen so that \( \tau(t_0) = t_0 \), i.e. the instant at which the \( t = 0 \) point on the transmitted signal is received (Kramer, 1967, p636). With this particular
choice of \( t_0 \), the target range and its derivatives are specified at \( t_0/2 \), i.e. the instant at which the \( t = 0 \) point on the transmitted waveform is being reflected (Rihaczek, 1969, p59). Successively differentiating (1.22) and (1.23) and equating terms gives

\[
\begin{align*}
\tau(t_0) &= \frac{2 \hat{R}(t_0/2)}{c + \hat{R}(t_0/2)} \approx \frac{2 \hat{R}(t_0/2)}{c} \\
\ddot{\tau}(t_0) &= \frac{2 \ddot{R}(t_0/2)}{c} \left( 1 + \frac{\dot{R}(t_0/2)}{c} \right)^{-3} \approx \frac{2 \ddot{R}(t_0/2)}{c}
\end{align*}
\] (1.24)

where the approximate forms of the above equations are valid when \( \dot{R}(t_0/2) \ll c \).

1.4.2 Moving Sonar/Stationary Target

![Diagram](image)

Figure 1.3. Illustration of the propagation delay measured by a moving sonar, where \( T \) is the target position, \( S' \) is the sonar position when the signal was transmitted, and \( S \) is the sonar position when the echo was received.

For the case of a moving sonar and stationary target (e.g. side-scan sonar) it is necessary to account for the difference in outward and return delays as illustrated in Figure 1.3. Splitting the propagation delay into these two terms gives

\[
\tau(t) = \frac{R(t - \tau(t))}{c} + \frac{R(t)}{c}
\] (1.25)

where \( R(t - \tau(t)) \) is the target range at the instant of transmission and \( R(t) \) is the range at the instant of reception. Again choosing \( \tau(t_0) = t_0 \), the coefficients of (1.23) can be expressed as

\[
\begin{align*}
\tau(t_0) &= \frac{R(t_0) + R(0)}{c} \\
\dot{\tau}(t_0) &= \frac{\dot{R}(t_0) + \dot{R}(0)}{c + \dot{R}(0)} \approx \frac{\dot{R}(t_0) + \dot{R}(0)}{c} \\
\ddot{\tau}(t_0) &= \frac{\ddot{R}(t_0) \left( 1 + \frac{\dot{R}(0)}{c} \right)^2 + \ddot{R}(0) \left( 1 - \frac{\dot{R}(t_0)}{c} \right)^2}{c \left( 1 + \frac{\dot{R}(0)}{c} \right)^3} \approx \frac{\ddot{R}(t_0) + \ddot{R}(0)}{c}
\end{align*}
\] (1.26)
where the approximate forms of the above equations are valid when $\dot{R}(0) \ll c$.

Sometimes it is advantageous to consider that the sonar is stationary and that the targets are moving through the sonar beam (Rihaczek, 1969, pp447-448). The theory of moving target resolution can then be applied. However, one must be careful with this interpretation. Since the sonar moves between transmission and reception, the difference in outward and return paths must be considered, especially when fast tow speeds and long propagation delays are expected. In radar applications however, the displacement between the points of transmission and reception is usually very small and can be neglected.

1.4.3 The Distorted Echo Signal

When there is no relative motion between the sonar and target, the received echo signal $e(t)$ is just a delayed replica of the transmitted signal $s(t)$, scaled in amplitude by the propagation losses and target reflectivity. However, if the target or sonar is moving, the received signal is compressed/stretched by a time scale factor $\eta(t)$, defined as

$$\eta(t) \equiv 1 - \dot{\tau}(t)$$

or in terms of a series expansion about $t_0$ using (1.23) as

$$\eta(t) \equiv \frac{t - \tau(t)}{t - t_0} = 1 - \left(\dot{\tau}(t_0) + \frac{\ddot{\tau}(t - t_0)}{2!} + \cdots\right)$$

Thus in general, the received echo signal can be described as

$$e(t) = A\eta(t)^{\frac{1}{2}}s(t - \tau(t)) = A\eta(t)^{\frac{1}{2}}s(\eta(t)(t - t_0))$$

where $A$ is a complex constant and $\eta(t)^{\frac{1}{2}}$ is an amplitude scale factor to ensure that the signal energy is conserved under the transformation caused by the relative motion.

For a closing range rate ($\dot{R} < 0$) the received signal is compressed by a factor $1/\eta(t)$ ($\eta(t) > 1$), and for an opening range rate ($\dot{R} > 0$) the signal is stretched by a factor $1/\eta(t)$ ($\eta(t) < 1$). The effect of this time scale factor is to distort the spectrum of the received signal. Unless the relative motion is known, there is a mismatch between the received echo and its matched-filter receiver with a consequent degradation of signal to noise at the receiver output.

1.4.4 The Doppler Effect

When the target range rate $\dot{R}$ is constant (or can be considered constant over the signal duration) the higher order range derivatives are zero, and the time scale factor is also constant. Substituting (1.25) into (1.28) for $\dot{R}(t) = \dot{R}$ gives

$$\eta = \frac{1 - \dot{R}/c}{1 + \dot{R}/c} \approx 1 - \frac{2\dot{R}}{c}$$

where the approximation is valid for $\dot{R} \ll c$. The echo spectrum of (1.29) can now simply be solved using the Fourier identity $e(at) \longleftrightarrow E(f/a)/|a|$, giving

$$E(f) = \frac{A}{\sqrt{\eta}} \exp(j2\pi ft_0)S(f/\eta)$$
Apart from a complex scale factor, the spectrum of the echo signal is similar to the spectrum of the transmitted signal but scaled in frequency by the reciprocal of the time scale factor. This phenomenon is known as the Doppler effect, and occurs whenever there is relative motion between the sonar and target. If the range rate $\dot{R}$ is known, the correct matched-filter for optimum detection of the received echo is thus $kS(f/\eta)$, where $k$ is an arbitrary amplitude scale factor.

For narrowband signals, it is common practice to treat the Doppler effect as a simple frequency translation (Cook and Bernfeld, 1967, p64). For example, consider a CW signal with a frequency $f_c$ reflected from a target moving with a radial velocity $v_r$ (i.e. a range rate $\dot{R} = v_r$). From (1.31), the echo signal is observed to have a frequency $\eta f_c = f_c + \nu$, where the frequency shift $\nu$ is

$$\nu = f_c(\eta - 1) \approx -2f_c \frac{v_r}{c}$$

(1.32)

In general, if $\tau(t)$ is the instantaneous propagation delay then the instantaneous demodulated phase measured by a monochromatic CW sonar is

$$\phi_d(t) = -2\pi f_c \tau(t) + \phi$$

(1.33)

where $\phi$ is an arbitrary constant phase shift. Differentiating (1.33) with respect to time, the instantaneous demodulated frequency, or Doppler shift, is thus

$$f_d(t) = -f_c \dot{\tau}(t) = f_c(\eta(t) - 1)$$

(1.34)

The minus sign in (1.33) is necessary since the Doppler shift is defined relative to the transmitted frequency $f_c$. For example, a receding target $\dot{\tau}(t) > 0$ causes a down-Doppler shift, i.e. the received frequency $f_c(1 - \dot{\tau}(t))$ is lower than the transmitted frequency $f_c$.

1.4.4.1 Differences Between Sonar and Radar Doppler

Although qualitatively the same, the effect of Doppler on radar and sonar signals is in general quantitatively different. This is a result of the difference between electromagnetic and acoustic propagation as mentioned in Section 1.4. For example, consider a radar transmitting a CW signal of frequency $f_T$. The Doppler shifted frequency observed at a moving target is

$$f_R = \left(\frac{c - v}{c + v}\right)^\frac{1}{2} f_T$$

(1.35)

where $v$ is the radial velocity of the target relative to the radar (defined to be positive for an increasing range rate). In comparison with an equivalent CW sonar, the acoustic frequency observed at a target is

$$f_R = \left(\frac{c - v_s}{c - v_t}\right) f_T$$

(1.36)

where the sonar speed $v_s$ and the target speed $v_t$ are defined relative to the medium and are positive in the direction away from the sonar in the line of the target.
After reflection, the return signal is also Doppler shifted, and in the radar case the observed echo frequency is

$$f_R = \left(\frac{c - v}{c + v}\right) f_T$$  \hspace{1cm} (1.37)

and in the sonar case is

$$f_R = \left(\frac{c - v_t}{c - v_s}\right) \left(\frac{c + v_s}{c + v_t}\right) f_T$$  \hspace{1cm} (1.38)

Notice that if either the sonar or target is stationary (with respect to the medium) then both the relativistic and non-relativistic approaches, (1.38) and (1.37), give the same value for the Doppler shift.

1.4.5 Signal Tolerance to Sonar/Target Motion

The effect of relative motion between a sonar and a target in its beam is generally a modulation of the phase function of the reflected signal. The longer the signal duration, the greater the distortion leading to an increased decorrelation at the receiver.

To determine whether a signal is tolerant to the effects of a relative target motion, consider a simple pulsed CW signal reflected from a target moving with a constant range rate $\dot{R}$. If the range rate is not so great as to distort the pulse duration, the range change over the pulse duration $T_p$ is $\dot{R}T_p$, producing a carrier phase change of

$$\Delta \phi = \frac{4\pi}{\lambda} \Delta R = \frac{4\pi}{\lambda} \dot{R}T_p$$  \hspace{1cm} (1.39)

If the received echo is processed by a matched-filter, the effect of the relative target motion is to produce a mismatch in the receiver. Assuming that a phase shift of $\Delta \phi \leq \pi/2$ is tolerable, the maximum pulse duration for which the carrier is unaffected by a range rate $\dot{R}$ is thus

$$T_p \leq \frac{c}{8f_c\dot{R}}$$  \hspace{1cm} (1.40)

The complex envelope of a signal, however, is less sensitive to the effects of target motion than the carrier is. Since the highest frequency component in the complex envelope is of the order of the signal bandwidth $B$, it is typically a factor $f_c/B$ more tolerant than the carrier (Rihaczek, 1969, p61). Therefore the maximum time-bandwidth product of a signal, insensitive to a constant range rate $\dot{R}$, is

$$BT_p \leq \frac{c}{8\dot{R}}$$  \hspace{1cm} (1.41)

When this inequality holds, the theory of small time-bandwidth product signals may be applied which is considerably simpler than the theory for large time-bandwidth signals. Equation (1.41) may also be applied when the matched-filter is adjusted for some range rate other than $\dot{R} = 0$, in which case the deviation of $\dot{R}$ from the design value should be used (Rihaczek, 1969, p61).

The tolerance of a signal to a constant range acceleration or other higher order range derivatives can be approached in a similar way. Assuming a constant range acceleration $\ddot{R}$, the change in range over the signal duration is $\ddot{R}T_p^2/2$ and this produces
a phase change $\Delta \phi$ at the receiver of

$$\Delta \phi = \frac{2\pi \tilde{R}T_p^2}{\lambda} \quad (1.42)$$

Again using the assumption that the receiver can tolerate a $\pi/2$ phase shift, the maximum pulse duration where the carrier is unaffected by an acceleration $\ddot{R}$ is thus

$$T_p \leq \frac{1}{2} \sqrt{\frac{c}{f_c \ddot{R}}} \quad (1.43)$$

Similarly, the complex envelope can be regarded to be insensitive to the effects of a constant acceleration when

$$BT_p^2 \leq \frac{c}{4 \ddot{R}} \quad (1.44)$$

Typically the effect of range acceleration on the complex envelope is small, excepting very long duration signals (cf. Stewart and Westerfield, 1959, p880). To illustrate this, consider a simple chirped pulse with a duration $T_p = 40$ ms and a bandwidth $B = 750$ Hz. From (1.41) the maximum tolerable range rate is $\ddot{R} = 6.25$ ms$^{-1}$ and from (1.44) the maximum tolerable range acceleration is $\dddot{R} = 312.5$ ms$^{-2}$ (assuming $c = 1500$ ms$^{-1}$).

### 1.4.6 Resolution on the Basis of Sonar/Target Motion

Rihaczek (1969, p109) shows that a 3 dB drop in a matched-filter output occurs whenever the target motion causes a phase shift of $180^\circ$, independent of whether the motion is one of constant range rate, constant range acceleration, or some other lower order range derivative. This $180^\circ$ phase shift is equivalent to change in path-length of $\lambda/2$, or a range change of $\lambda/4$. Therefore, if the differential range between two targets changes by $\lambda/4$, a filter matched to one target will have a response 3 dB lower for the other target. Furthermore, since bin width is usually defined as the 3 dB width of the filter response, the two targets will fall into different bins provided the differential range, over the measurement interval, is at least $\lambda/2$.

### 1.5 Ambiguity Functions

In Section 1.3, the matched-filter was introduced as a device which correlates the actual echo signal with a function representing the expected noise free echo signal. The response of such a filter, to an echo with parameters different from those expected, is given by the ambiguity function of the filter.

The conventional (Woodward) ambiguity function$^1$ is often expressed as

$$\chi(\nu, \tau) = \int_{-\infty}^{\infty} s(t)s^*(t-\tau)\exp(j2\pi \nu t) \, dt \quad (1.45)$$

$^1$There is a potential confusion in the definition of the ambiguity function (Sinsky, 1978, p129). The definition used here is based on the approach of Rihaczek (1965a, p119), where the ambiguity function is defined as the matched-filter output response. Compare this to Woodward's ambiguity function, which is defined as the cross-correlation between the transmitted and received signals, and gives the time-reversed output from a matched-filter (Woodward, 1964, p120).
where \( s(t) \) is the pre-envelope of the transmitted signal, \( s^*(t - \tau) \) is the complex conjugate of the transmitted signal, but delayed by \( \tau \) (related to target range), and Doppler shifted in frequency by \( \nu \) (related to target range rate).

There are two simplifying assumptions inherent in this definition of the ambiguity function \( \chi(\nu, \tau) \):

(i) Negligible Doppler distortion of the signal complex envelope.

(ii) Negligible effects of range acceleration and higher order range derivatives.

The first assumption requires that the signal has a small time-bandwidth product — i.e. \( T_p B < c/8\bar{R} \), see (1.41). The second assumption requires that the signal has a short duration — i.e. \( T_p < \sqrt{\lambda/4\bar{R}} \), see (1.43). Thus this ambiguity function only applies to signals which have both a small signal duration (with regard to the maximum expected range rate) and a small fractional bandwidth (Rihaczek, 1967, p705). Therefore, \( \chi(\nu, \tau) \) is referred to as the narrowband ambiguity function.

The ambiguity function is a quantitative way to determine the ability of a waveform and its processing filter to resolve two or more targets at arbitrarily different range and range rates (Sinsky, 1978, p127). It is also useful for assessing other signal properties with regards to resolution, measurement accuracy (precision), ambiguity and clutter performance. For a comprehensive treatment of the narrowband ambiguity function see Rihaczek (1969).

For wideband signals, the Doppler effect cannot be approximated as a simple frequency shift (see Section 1.4.4) and so Kelly and Wishner (1965, p59) define the wideband ambiguity function as a function of time delay \( \tau \) and Doppler time scale factor \( \eta \)

\[
\chi_w(\eta, \tau) = \sqrt{\eta} \int_{-\infty}^{\infty} s(t)s^*(\eta(t - \tau))dt
\]  

(1.46)

The assumption made here is that the target is considered to be moving with a constant range rate over the signal duration. When there are rapidly accelerating targets and long duration signals, it is necessary to extend the ambiguity function to include parameters related to target acceleration and other higher order range derivatives (Kelly and Wishner, 1965, p59). For the majority of side-scan sonar applications, however, this added complication is unnecessary.

There have been a number of studies of the properties of the wideband ambiguity function (Harris and Kramer, 1968; Altes, 1973; Sibul and Titlebaum, 1981; Lin, 1988). Analytical solutions of the integral in (1.46) are difficult, and only approximate expressions for the wideband ambiguity function have been derived.

### 1.5.1 Multiple Target Resolution

Each target within the sonar beam can be associated with its own ambiguity function \( \chi(\tau, \nu) \), scaled in magnitude by the target cross-section, and with a phase shift determined by the signal reflection. These ambiguity functions are superimposed by first shifting them appropriately in \( \tau \) and \( \nu \), depending on the target ranges and range rates, and then adding them together. The resultant ambiguity function is the combined receiver response to all the targets present in the beam. However, any slight phase change can cause pronounced fluctuations in the combined response.
because the phase is a very sensitive parameter of the target range and of the target reflection properties. Rather than performing a coherent superposition, it is better to determine the average behaviour of the combined receiver response by adding the squared envelopes $|x(\tau, \nu)|^2$, of the ambiguity functions (Rihaczek, 1969, p114).

The total volume under the normalised ambiguity diagram is unity\(^2\), independent of the signal waveform (Woodward, 1964, p120). Consequently, the total interference contributed by any one target is dependent on its reflection properties and not on the choice of waveform. In other words, there are finite constraints on the achievable resolution in both range and range rate, irrespective of the choice of waveform.

1.6 The Sonar Equation

One form of the general sonar equation applicable to all monostatic active sonars is

$$\text{SNR} = \text{SL} + \text{TS} - 2\text{TL} - (\text{NL} - \text{DI}) + \text{IF}_p + \text{IF}_n$$  \hspace{1cm} (1.47)

where SNR is the peak signal to total rms noise power ratio, SL is the projector source level, TS is the target strength, TL is the one-way transmission loss, NL is the total noise power in the receiver passband, DI is the directivity index of the hydrophone, IF\(_p\) is the improvement factor for chirped signals, and IF\(_n\) is the improvement factor for the integration of several pulses. All these terms are logarithmic ratios measured in dB and are defined as follows (cf. Cutrona, 1975, p340 and de Heering, 1982, pp20–23)

$$\text{SNR} = 10 \log_{10} \frac{\hat{S}}{N}$$  \hspace{1cm} (1.48)

$$\text{SL} = 10 \log_{10} \frac{P_T G_p}{4 \pi r_0^2 I_0}$$  \hspace{1cm} (1.49)

$$\text{TS} = 10 \log_{10} \frac{\sigma_s}{4 \pi r_0^2}$$  \hspace{1cm} (1.50)

$$\text{TL} = 10 \log_{10} \left[ \left( \frac{R}{\tau_0} \right)^2 \exp(qR) \right]$$  \hspace{1cm} (1.51)

$$\text{DI} = 10 \log_{10} G_h$$  \hspace{1cm} (1.52)

$$\text{NL} = 10 \log_{10} \frac{N_0 B_N}{I_0}$$  \hspace{1cm} (1.53)

$$\text{IF}_p = 10 \log_{10} \frac{T_p}{T_c}$$  \hspace{1cm} (1.54)

$$\text{IF}_n = 10 \log_{10} I_i(n)$$  \hspace{1cm} (1.55)

where \(r_0 = 1 \text{ m}\) is the reference range, \(I_0 = 0.667 \times 10^{-18} \text{ Wm}^{-2}\) is the reference intensity of a plane wave of rms pressure 1 \(\mu\text{Pa}\), \(P_T\) is the transmitted power level, \(\hat{S}\) is the peak output signal, \(G_p\) and \(G_h\) are the gains of the projector and hydrophone respectively, \(\sigma_s\) is the effective scattering cross-section of the target, \(q\) is the attenuation coefficient, \(B_N\) is the effective receiver noise bandwidth, \(T_p\) is the transmitted

\(^2\)Provided the signal can be considered narrowband (Sibul and Titlebaum, 1981, p83)
pulse length, and finally $T_c$ is the compressed pulse length. The other parameters are defined in the glossary.

The sonar equation is applicable to all active sonars, and is just the logarithmic equivalent of the radar equation. The main difference is the definition of $N_0$, which is defined here as the background sea noise spectral intensity in $\text{W m}^{-2} \text{Hz}^{-1}$, rather than $\text{W Hz}^{-1}$ as is common in radar. This is because the sonar noise level $NL$ is usually dominated by the background sea noise, unlike most radars which are receiver noise limited (Skolnik, 1980, §2.3). Also note that the background noise is assumed to be isotropic sea noise rather than reverberation. For a reverberant background the terms $NL - DI$ in the sonar equation (1.47) are replaced by an equivalent plane-wave reverberation level $RL$ (see Section 1.7). The parameter $DI$ defined in terms of an isotropic background is now no longer appropriate since reverberation is not isotropic (Urick, 1975, p20).

The sonar equation (1.47) can be expressed in an equivalent non-logarithmic form as

$$\frac{\hat{S}}{N} = \left( \frac{P_T G_p}{4\pi r_0^2 I_0} \right) \left( \frac{\sigma_s}{4\pi r_0^2} \right) \left( \frac{\exp(-2qR) r_0^2}{R^4} \right) \left( \frac{G_h I_0}{N_0 B_N} \right) \left( \frac{T_p}{T_c} \right) (I_i(n)) \quad (1.56)$$

and rearranging terms

$$\frac{\hat{S}}{N} = \left( \frac{P_T G_p \sigma_s A_h \exp(-2qR)}{(4\pi)^2 R^4} \right) \left( \frac{G_h}{N_0 B_N A_h} \right) \left( \frac{T_p}{T_c} \right) (I_i(n)) \quad (1.57)$$

where the first factor is proportional to the received echo signal level, the second factor is proportional to the received noise level, and the last two factors are the pulse compression and integration improvement factors. The advantage of this representation is that it is independent of the reference quantities $I_0$ and $r_0$.

### 1.6.1 Pulse Integration

The reliability of signal detection can be improved by integrating many received pulses, thus improving the signal to noise ratio of the output signal (Brookner, 1978a). The integration improvement factor for the addition of $n$ pulses is described by the function $I_i(n)$ in the sonar equation (1.56). The detection improvement depends on the method of integration (coherent, noncoherent, or a combination of both) under somewhat artificial conditions of constant signal and no fading. These assumptions are known as the Marcum model, after Marcum (1947). Modifications to the Marcum model include fluctuating targets as proposed by Swerling (Swerling, 1960; Bates, 1964).

Integration before the detector is called predetection, or coherent integration, while integration after the detector is called postdetection, or noncoherent integration (Skolnik, 1980, §2.6). Coherent integration requires that the phase of the echo signal be preserved, however noncoherent integration is not concerned with the signal phase and is not as efficient as coherent integration. The loss in efficiency is caused by the nonlinear action of the detector, which converts some of the signal energy into noise in the rectification process.
1.6.2 Pulse Compression

Consider for now a simple pulsed CW sonar with a signal duration $T_p$. The matched-filter for this signal has an effective noise bandwidth $B_N = 1/T_p$ and therefore the signal to noise ratio at the output of this filter for a single pulse is

$$\frac{\hat{S}}{N} = \frac{E A_p A_h \sigma_s \exp(-2qR)}{N_0 \lambda^4 R^4}$$

(1.58)

where $E = P_r T_p$ is the energy in the transmitted signal. It is therefore obvious that to improve the signal to noise ratio it is necessary to increase the peak power level or to lengthen the signal duration.

The maximum acoustic power that a projector can output is dependent on the usual electrical, mechanical and thermal limitations of transducers (Camp, 1970, §8.5) and in addition by the phenomenon of cavitation (Urick, 1975, §4.2). This occurs when the sound pressure amplitude exceeds the hydrostatic pressure of the water causing the water to 'rupture' with a degradation in projector performance. (See Section A.1.2 for an example calculation of the cavitation limit.)

When peak power limited, the alternative is to increase the signal duration with a resultant degradation in range resolution ($\Delta r = c T_p/2$). Thus with a simple pulsed CW sonar, there is a compromise between range resolution and peak power requirements for a desired signal to noise ratio.

The solution to this problem came with the introduction of pulse compression techniques (Klauder et al., 1960). Pulse compression is a means of operating a sonar at long duty cycles to obtain the resolution performance of a short pulse but with the detection capability of a long pulse. This is achieved by modulating the transmitted pulse in such a manner as to increase its bandwidth (usually some form of phase or frequency modulation). The received echo is then processed by a matched-filter that compresses the long pulse to a duration $T_c = 1/B$, where $B$ is the bandwidth of the modulated pulse. The compression factor $T_p/T_c$ (or dispersion factor) is thus equal to the time-bandwidth product of the modulated pulse $T_p B$.

Like pulse integration, there is a wealth of literature on pulse compression (see e.g. Rihaczek, 1969 and Farnett et al., 1970). There have been many different modulation methods proposed, each with different ambiguity functions; the most common method being linear FM since it is relatively easy to generate and to compress (see for example Cook and Bernfeld, 1967, §6.5).

1.7 Reverberation

The sea contains many inhomogeneities that form discontinuities in the physical properties of the medium. The effect of these discontinuities is to intercept and reradiate portions of acoustic energy incident on them, a process known as scattering. The sum of this combined scattered radiation returning to the sonar hydrophone is known as reverberation.

Reverberation is generally split into three categories; volume reverberation from inhomogeneities in the body of the sea, sea surface reverberation produced by scatterers located on or near the sea surface, and bottom reverberation from scatterers
on or near the sea floor. Bottom reverberation is usually the strongest source of reverberation, followed by volume, then sea surface reverberation (Walsh, 1969).

For analysis, bottom and sea surface reverberation are usually grouped together as surface reverberation, since a two dimensional distribution of scatterers is involved. Also note that for sea floor imaging sonars, bottom reverberation (at least the time independent scatterers on the sea floor) is the signal of interest.

1.7.1 Volume Reverberation

In a pinging system the volume of water from which the reverberation can come from at any instant depends on the beamwidth of the transducers and the pulse length of the ping. The beamwidth determines the solid angular boundary of the active reverberation and the length of the ping determines the thickness.

The volume reverberation level from range $R$ is

$$RL_v = SL + S_v - 2TL + 10 \log_{10} V$$

(1.59)

where $S_v$ is the volume scattering strength per unit volume, SL is the source level, $2TL$ is the two-way transmission loss for reverberation at range $R$ (assuming the projector and source are colocated), and $V$ is the active reverberation volume. This volume can be expressed as

$$V = \frac{cT_p \Omega R^2}{4\pi}$$

(1.60)

where $\Omega$ is an equivalent solid angle measure of beamwidth for the two-way power pattern, and $T_p$ is the transmitted pulse length (Burdic, 1984, §12.3).

The volume scattering strength $S_v$ is defined as

$$S_v = 10 \log_{10} \frac{s_v}{4\pi r_0^2}$$

(1.61)

where $s_v$ is the total cross section of the reverberation per unit volume (with units $m^{-1}$). Typically, the magnitude of $S_v$ is in the range $-60$ dB/m$^3$ to $-90$ dB/m$^3$ (Urick, 1975, §8.10), and varies with depth, season, time of day (greater at night), frequency (when below 20 kHz), and concentration of biological organisms.

Notice with this simplified derivation that the reverberation level decreases as $20 \log R$, rather than $40 \log R$, as experienced with a point target. Notice also that the volume reverberation can only be decreased by reducing the transmitted pulse duration or effective beamwidth.

1.7.2 Surface Reverberation

The surface reverberation level from range $R$ is

$$RL_s = SL + S_s - 2TL + 10 \log_{10} A$$

(1.62)

where $S_s$ is the surface scattering strength per unit area, typically in the range $-10$ dB/m$^2$ to $-60$ dB/m$^2$ (Urick, 1975, §8.11), and $A$ is the active reverberation surface area given by

$$A = \frac{cT_p}{2} \Omega R$$

(1.63)
where $\Theta$ is the equivalent two-way azimuth beamwidth.

There are many approximations made in the derivations of these reverberation models, however Urick (1975, §8.4) concludes that the simple reverberation model is valid, at least for short pulses and for relatively short ranges where the multipath effects that complicate a reverberation problem are absent.

### 1.7.3 Signal to Reverberation

Often reverberation is the limitation on active sonar performance (Urick, 1975, p211), but unlike noise, it cannot be overcome by increasing the transmitted source level, since the reverberation level is proportional to the source level.

The echo level $EL$ from a point target is

$$ EL = SL + TS - 2TL $$

and, comparing this with (1.59), the echo to volume reverberation can be written as

$$ EL - RL_v = TS - S_v - 10 \log_{10} \left( \frac{cT_p}{2\Theta} \right) - 20 \log_{10} R $$

and similarly, using (1.62), the echo to surface reverberation level is

$$ EL - RL_s = TS - S_s - 10 \log_{10} \left( \frac{cT_p}{2\Theta} \right) - 10 \log_{10} R $$

Notice that surface reverberation decreases more rapidly with range than volume reverberation since the area of the reverberating region increases as $R$ rather than $R^2$.

From (1.65) and (1.66) it is apparent that the only way to improve the echo to reverberation ratio is to improve the directivity of the sonar beam patterns and to use shorter pulse lengths. Reducing the beamwidths to reduce reverberation, however, also reduces the search (or area mapping) rate of a sonar and provides fewer 'hits' on the target. This problem is usually circumvented by radiating multiple narrow beams (cf. de Moustier, 1986 and Okino and Higashi, 1986).

In shallow water vertical directivity provides little array gain because reverberation tends to be strongly coherent for shallow water in the vertical direction. On the other hand, in the horizontal, reverberation is only weakly coherent and therefore horizontal directivity can help to suppress reverberation (Urick, 1975, §8.15).

Although a reduction of the pulse duration $T_p$ improves the reverberation level, the target strength may also be reduced if the pulse length is shorter than the length of the target (Kinsler et al., 1982, p424). In addition, since the transmitted energy is also reduced in proportion to the signal duration, the echo signal may now become noise limited rather than reverberation limited. The solution here is to use pulse compression techniques as described in Section 1.6.2. This reduces the effective size of the reverberating region by the ratio $T_p/T_c$, where $T_c$ is the compressed pulse length. In addition, Dix and Palmer (1984, p314) have found that the reverberation limited performance of a broadband coherent processor is better than that predicted from consideration of range resolution alone.
1.8 Undersea Acoustic Propagation

As well as the presence of noise and reverberation, an active sonar has to contend with the problems of a low velocity of acoustic propagation, signal attenuation due to viscous absorption, refraction due to inhomogeneities in the medium, echo fluctuations due to multiple propagation paths (multipath), and other distortion effects due to dispersion, medium coherence, and finite amplitude phenomena.

Undersea propagation is notoriously difficult to predict, especially in shallow water environments (Urick, 1982, Ch.9), and is far more complicated than microwave propagation in the atmosphere. The linear acoustic wave equation, subject to time-varying volume and boundary conditions, is often used as a starting point (Knight et al., 1981, p1453). However its scope is limited, since it only applies for propagation of acoustic waves of relatively small amplitude in an isotropic, perfectly elastic, homogeneous, unbounded, and lossless medium (Winder, 1975, p298). In practice, many of the parameters vary in time and space in a fashion which can only be described in a statistical sense. Nevertheless, the simple model has value as a reference or basis from which to predict expected performance.

The following sections look briefly at the effects of absorption, multipath propagation and medium coherence. The more inquisitive reader is referred to the books by Urick (Urick, 1975; Urick, 1982).

1.8.1 Attenuation

The transmission loss is usually described by an inverse square divergence term plus other losses which include absorption, beam refraction and scattering. Under good transmission conditions or mixed-layer thermal conditions these other losses are dominated by the absorption loss which can be described by an absorption coefficient $\alpha$. With this assumption, the total one-way transmission loss can be expressed as

$$ TL = 20 \log_{10} \frac{R}{r_0} + \alpha R \quad (1.67) $$

where $r_0$ is the reference range. The variation in transmission loss due to absorption is usually a slowly changing function of frequency, salinity, depth, and temperature. For complicated empirical formulae see Urick, 1982, Ch.5, and Kinsler et al., 1982, §7.6.

Once absorption loss becomes appreciable it soon dominates the transmission loss equation. For instance, at 30 kHz the two-way absorption loss is approximately $1.5 \times 10^{-2}$ dB/m, or a 6 dB loss for a target at 400 m. However at a range of 4 km, the absorption loss is now 60 dB, an increase of 54 dB, compared with a 20 dB increase in spreading losses over the same interval. Therefore, for long range applications the absorption coefficient must be small, thus requiring the use of low frequencies.

In shallow waters, the transmission loss is often better at short ranges (a result of trapping in the shallow water duct) and poorer at longer ranges (because of boundary losses, see Urick, 1982, p9.1).
One solution to counteract echo fading is to transmit sufficient power to ensure detection under the deepest fading conditions, and to hold the echo levels constant using AGC (Automatic Gain Control) techniques in the receiver (Dix and Palmer, 1984, p309). This is inefficient in terms of transmitter power and there are often better solutions involving diversity reception techniques, as employed in radiocommunications for example.

Spatial diversity techniques for sonar are often restricted by a limited towfish size, so instead it is necessary to use time and/or frequency diversity. Dix and Palmer (1984, p314) show that a broadband coherent processing system acts as such a diversity receiver, reducing the effects of echo fading due to multipath.

\[ R_s = \sqrt{R^2 + 4d_s d_t} \]  

where \( R \) is the direct path slant range, \( d_s \) is the depth of the sonar, and \( d_t \) is the depth of the target. Assuming the transmitted signal can be constrained to follow the direct path (by appropriate beam shaping in elevation, say), the difference between the length of the direct and indirect return paths is \( R_s - R \). Therefore the measured range difference \( \Delta R_s \), is

\[ \Delta R_s = \frac{R_s - R}{2} \approx \frac{d_s d_t}{R} \]  

where the approximation is valid for \( R \gg 4d_s d_t \). Note also that the wave following the indirect path is phase shifted by 180° on reflection at the sea surface.

For example, consider a sonar operating at a depth \( d_s = 15 \) m, and a target at a depth \( d_t = 40 \) m. If the target is at a range of \( R = 100 \) m, then from (1.69), \( \Delta R_s \approx 6 \) m. Therefore, if the range bin width is smaller than \( \Delta R_s \), a ghost image will appear at a range 6 m further than the actual target range. If both the transmitted and received signals follow the indirect path, the ghost target will appear \( 2\Delta R_s \approx 12 \) m.
range difference $\Delta R_s$, is

$$\Delta R_s = \frac{R_s - R}{2} \approx \frac{d_s d_t}{R}$$  \hfill (1.69)$$

where the approximation is valid for $R \gg 4d_s d_t$. Note also that the wave following the indirect path is phase shifted by 180° on reflection at the sea surface.

For example, consider a sonar operating at a depth $d_s = 15$ m, and a target at a depth $d_t = 40$ m. If the target is at a range of $R = 100$ m, then from (1.69), $\Delta R_s \approx 6$ m. Therefore, if the range bin width is smaller than $\Delta R_s$, a ghost image will appear at a range 6 m further than the actual target range. If both the transmitted and received signals follow the indirect path, the ghost target will appear $2\Delta R_s \approx 12$ m further than the actual target range. Imaging the target in shallower waters, with say $d_s = 5$ m, $d_t = 10$ m and $R = 100$ m (typical for Lyttelton Harbour), the multipath range separation reduces to $\Delta R_s \approx 0.5$ m. If the range gating is not fine enough to separate these echoes, interference will occur, destroying the signal coherence.

To reduce the incidence of reflections from the sea surface, it is therefore necessary to reduce the elevation beamwidth of the transducers or to insert acoustic baffles between the transducers and the sea surface.

1.8.3 Coherence

Sonar signals transmitted from a steady source are received at a distant hydrophone as a fluctuating signal. This is a result of the inhomogeneities in the sea, and at its boundaries, which move in an irregular manner with respect to the source and receiver. The effect is to produce scattering and multipath reflections, of constantly varying amplitude and phase, which combine to give a fluctuating echo signal. This fluctuation when measured at a single hydrophone with respect to a known reference is termed temporal coherence (Urick, 1982, Ch.12), and when measured with respect to another spatially separated hydrophone is termed spatial coherence (Urick, 1982, Ch.13). Note that if the movements of the inhomogeneities in the medium were 'frozen', there would be no temporal fluctuation between a fixed source and a fixed receiver. However, the spatial variation would still be present, and thus the performance of any acoustic aperture is limited by the spatial coherence of the medium.

Winder (1975, p301) concludes that the large fluctuations that characterise acoustic propagation are primarily due to the interference of close multipaths generated from the sea surface and sea bottom.

1.9 Target Characteristics

The target strength is the proportion of the transmitted signal energy which is intercepted and reflected by a target. The target strength is thus proportional to the cross-sectional area and reflectivity of the target. In general, target strength varies with both aspect and frequency. Excluded from the definition of targets are inhomogeneities in the sea of indefinite size — the scattering from these objects is considered as reverberation.

Natural objects tend to be rough, unlike most man-made objects which tend to be fairly smooth compared with the wavelengths of sound used in underwater
imaging (Sutton, 1979, p555). Consequently, natural objects tend to produce diffuse reflections and man-made objects often produce specular (mirror-like) reflections.

The Rayleigh parameter, defined as \( R = kH \sin \theta \), is often used as a criterion for the roughness or smoothness of a surface, where \( k \) is the wavenumber, \( H \) is the rms surface roughness and \( \theta \) is the grazing angle of the incident wave (Urick, 1975, p122). When \( R \ll 1 \), the surface is considered primarily as a reflector, and conversely when \( R \gg 1 \), the surface is primarily a scatterer.

At optical wavelengths most surfaces (mirrors excepted) are sufficiently rough so as to scatter light, but at acoustic (and microwave) wavelengths many objects (especially man-made ones) have smooth surfaces that reflect radiation quasi-specularly. This specular nature of man-made objects in particular has made detection and recognition difficult.

These specular returns can produce large image highlights, but at a position dependent upon the geometry of the imaging system. In particular, narrow beamwidth systems can easily 'miss' a target if the specular reflection is not normal to the receiver beam. Consequently, specular reflections are not a stable contribution to the image and therefore not so useful for object recognition (Steinberg, 1976, §10.3). In addition, the weak scattering from corners, edges and other small protuberances can be masked by large sidelobes of these specular reflections. This is unfortunate because these weak scatterers are stable and therefore contribute to object recognition.

Specular effects can be reduced by using some form of diffuse illumination. The simplest way is to use multiple acoustic sources spaced far apart although this may not necessarily be easy to implement in practice due to space restrictions.

Another difference between optical and acoustical imaging is that sound waves usually penetrate objects and often multiple echoes are produced by reflections from inner and back surfaces.

1.9.1 Speckle

Highly coherent systems suffer from the undesirable property of speckle (Goodman, 1976). The resultant images have a speckly appearance, where there are bright and dark areas in the image that are of the same order of size as the details of the object being imaged. These bright and dark areas are not related to any details of the object being imaged and therefore speckle degrades the image quality and resolution (Dainty, 1975).

Speckle results from mutual interference of a coherent signal scattered from elementary areas of a rough surface, or as a result of wavefronts distorted by medium turbulence interfering across an aperture (Abbott and Thurstone, 1979). The size of the image speckle is given roughly by \( \pi c/B \), where \( B \) is the signal bandwidth. Therefore, to reduce the effect of speckle, the speckly image can be smoothed using a low-pass spatial filter with a cutoff frequency significantly lower than \( \pi c/B \) (Ikeda et al., 1979a, p78).

Alternatively, since speckle is not found with incoherent illumination, speckle effects can be reduced by using some form of noncoherent averaging in the production of the image (see for example Dainty, 1975, Ch.4). Both spatial compounding and frequency compounding have been used to reduce the effects of speckle. With spatial compounding, statistically independent echoes from different views are com-
bined (Kirk, Jr., 1975a; van de Lindt, 1977; Burkhardt, 1978), whereas frequency compounding uses different transmitted frequencies (Magnin et al., 1982) or splits a wideband signal into narrower frequency bands (Newhouse et al., 1982).

The result of these techniques is to improve the image contrast (reduce the speckle) by a factor $\sqrt{N}$, where $N$ is the number of independent echoes compounded. However, spatial or frequency compounding results in a loss of resolution due to the requirement that the compounded echoes be uncorrelated. Shankar and Newhouse (1985) show that the resolution can be increased substantially, with only a slight loss in SNR, if partially correlated echoes are compounded. Note, however, that addition of speckle images on an amplitude basis does not change the statistics of intensity, aside from a constant scale factor (Dainty, 1975, §2.3.1).

The speckle reduction techniques described above are only applicable to speckle resulting from coherent illumination of a rough object. If the speckle is generated by medium turbulence, however, then compounding leads to a blurred image. This type of speckle is a result of a convolution of the desired object with a noisy point spread function, and is produced in both coherent and incoherent imaging systems, e.g. astronomical imaging (Bates, 1982). Deconvolution techniques are therefore required to reduce this form of image speckle, and in particular blind deconvolution, since the noisy point spread function is generally unknown (Bates and McDonnell, 1986, Ch.VI).

### 1.10 Undersea Imaging Systems

Undersea imaging systems primarily transmit acoustic signals since electromagnetic waves are readily absorbed in seawater because of the high electrical conductivity. Even though underwater television cameras exist, they are limited in range by the turbidity of water, especially in undersea applications. In very clear waters, the optical visibility range can sometimes extend up to 30–60 m, but usually the visibility in the ocean is limited to much shorter ranges. The visibility range in undisturbed deep ocean waters is typically 6–15 m which reduces to 1–6 m for near-shore waters. However in areas where man is most interested in underwater viewing, such as harbours and estuaries, the visibilities are generally in the range of 0–1 m (Sutton, 1979, p554). Where man is actively working on the sea floor, the disturbed sediment often prohibits any form of optical viewing, even in normally clear waters.

#### 1.10.1 Acoustical Undersea Imaging

Acoustic energy can easily penetrate the mud and silt which is the chief cause of optical turbidity in water. Consequently, acoustical imaging systems can achieve larger ranges than their optical counterparts. However, the resolution of acoustical imaging systems is a lot poorer than that achievable by optical means, due to the much longer wavelengths of sound in water compared to optical wavelengths.

Undersea imaging systems are often divided into two categories; acoustical imaging systems and imaging sonars. Sonars are typically used to indicate where something is located, while acoustical imaging systems (and optical imaging systems) are typically used to indicate what something looks like. The distinction is not so obvious for high resolution imaging sonars, especially since both acoustical imaging systems
and sonars share many properties and implementations. In general, however, sonars have much longer ranges and poorer resolution than acoustical imaging systems.

<table>
<thead>
<tr>
<th>Frequencies</th>
<th>0.1–2 MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelengths</td>
<td>0.75–15 mm</td>
</tr>
<tr>
<td>Apertures</td>
<td>0.1–1 m</td>
</tr>
<tr>
<td>Angular Resolution</td>
<td>0.1–2°</td>
</tr>
<tr>
<td>Ranges</td>
<td>1–100 m</td>
</tr>
<tr>
<td>Depths of Field</td>
<td>0.1–50 m</td>
</tr>
</tbody>
</table>

Table 1.1. Typical characteristics of underwater acoustic imaging systems (from Sutton, 1979, p554).

Typical operating frequencies for acoustical imaging systems are 100 kHz–2 MHz (wavelengths 0.75–15 mm, see Table 1.1) while sonars generally use lower frequencies to obtain longer ranges. Table 1.2 shows the typical operating frequencies and maximum ranges for imaging sonars. Note that long range sonars require the use of long wavelengths since high frequencies signals are readily absorbed. For the relative performances of optical, sonar and acoustical imaging systems see Sutton (1979).

<table>
<thead>
<tr>
<th>Max. Range</th>
<th>Very Short Range</th>
<th>Short Range</th>
<th>Medium Range</th>
<th>Long Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>FQencies</td>
<td>500 kHz–1 MHz</td>
<td>100–500 kHz</td>
<td>20–100 kHz</td>
<td>5–20 kHz</td>
</tr>
<tr>
<td>Wavelengths</td>
<td>1.5–3 mm</td>
<td>3–15 mm</td>
<td>15–75 mm</td>
<td>75–300 mm</td>
</tr>
</tbody>
</table>

Table 1.2. Typical parameters of imaging sonars (compiled from Flemming, 1982b, pp7,8).

Underwater acoustical imaging seeks to form two dimensional images using sound waves. Similar techniques are used to those in ultrasonic medical imaging and non-destructive testing (cf. Schueler et al., 1984). Acoustical imaging systems generally produce images, in a vertical plane, of the reflected energy versus vertical and lateral dimensions at some range, similar to that used in optical imaging systems. Sonar images on the other hand have range as one dimension and bearing or track as the other.

1.10.2 Imaging Sonars

The two most common imaging sonar configurations are the sector-scan and side-scan sonars. Both sector-scan and side-scan sonars transmit a beam at some grazing angle to the sea floor. The sound is backscattered from the sea floor and is detected by the sonar. When there is no object present, the bottom backscatters a small fraction of the transmitted energy as a more-or-less uniform background. Any object present intercepts and reflects some of the sound earlier than would the bottom. The reflected energy from an object appears as a bright point on the sonar display at the range of the object. Just behind the bright spot on the display is a dark region due to the shadow cast by the object on the sea floor. In this shadow region, no acoustic energy is backscattered by the sea floor. Shadows are a useful feature of sonar images.
(Gabor, 1969), since they indicate the shapes of targets, whereas the presence of a target is indicated by a bright spot.

Sector-scan sonars display reflected acoustic energy on a polar plot of bearing versus range. This type of display is usually known as a plan position indicator (ppi), and is generated either by mechanically scanning a single beam or by electronically scanning multiple preformed beams.

Side-scan sonars produce images similar to that of sector scan sonars but with a rectangular format of along-track position versus range. A narrow beam is transmitted perpendicularly to the side of the sonar, as the sonar is moved through the water, to obtain spatial coverage in the along-track direction.

1.10.3 Narrow Beam Side-Scan Imaging Sonars

Side-scan (side-looking) sonars image the sea floor by radiating a beam perpendicular to the direction of the sonar motion. A two-dimensional image is generated of the acoustical reflectivity of the sea floor, with one dimension proportional to slant range and the other to along-track coordinate position. Using knowledge of the beam elevation angle, the slant range coordinate can then be converted into cross-track position.

Traditionally, a high along-track resolution is obtained by using a narrow azimuth beamwidth. The narrower the beam, the higher the angular resolution and thus the higher the along-track resolution. To achieve these narrow beamwidths it is necessary to use short acoustical wavelengths compared to the transducer dimension. However, absorption losses increase with frequency, and therefore the choice of operating frequency requires a compromise between angular resolution and operating range. Consequently, high resolution side-scan sonars are usually limited to short range operation.

An advantage of using longer sound wavelengths is that they can penetrate through the mud and sediment on the sea floor (Hamilton, 1972). This is of particular importance for sub-bottom profiling applications or for imaging objects, such as mines or cables, on the sea floor covered in a layer of mud (Dutkiewicz and Dembigh, 1987a). These objects are invisible to optical imaging systems.

1.10.4 Range Resolution

The range accuracy of a sonar is the precision with which the range of a single target can be measured. In principle, given an infinite signal to noise ratio, there is no lower limit to the range accuracy (Rihaczek, 1969, §3.1). Range resolution is often confused with range accuracy, and is defined here as the minimum range separation of two targets that can be simultaneously detected. Usually the 3 dB, or Rayleigh, criterion is used for the resolution measure — i.e. there must be at least a 3 dB dip in the response between targets. The Rayleigh criterion is reasonably satisfactory for two point targets of similar strength. However, if the sidelobes from a strong target swamp the echoes from a weak target, the measure becomes meaningless, even if the targets are theoretically resolved.

(Rihaczek, 1969, p487) shows that range resolution $\Delta r$ is independent of the type
of modulation used and is of the order of the inverse signal bandwidth $B$, where

$$
\Delta r \sim \frac{c}{2B}
$$

(1.70)

Defining the range resolution capability of a signal in a single parameter is difficult since it depends on the shape of the signal spectrum and the measure of the bandwidth used. In this thesis, the 3 dB bandwidth of a signal is used as a measure of its nominal range resolution capability. For example, a pulsed CW sonar with a pulse duration $T_p$ has a 3 dB bandwidth of approximately $1/T_p$, and substituting this into (1.70) gives a nominal range resolution of $cT_p/2$ as expected from simple echo-ranging principles. A pure CW signal on the other hand has no range resolution capability since it has no bandwidth.

1.10.5 Maximum Range

Sonars with a low energy output tend to be noise limited—i.e. the maximum range is achieved when the signal to noise ratio drops below the threshold required for detection. For a given target strength, noise level and directivity index, (1.47) shows that the source level required for target detection, is dependent on the expected transmission losses. The maximum range is limited at the upper frequencies of operation by the attenuation coefficient of seawater and at all frequencies by the spherical spreading loss. Therefore to increase the maximum sonar range, it is necessary to increase the source level or to make the beam patterns more directional.

Increasing the source level increases both the echo level and the reverberation level at a given range. Since reverberation generally decreases less rapidly with range than the echo level, the maximum range of high energy sonars occurs where the echo to reverberation ratio is insufficient for detection—i.e. the maximum range is reverberation limited.

1.10.6 Range Ambiguities

Range ambiguities occur if the autocorrelation of the signal modulation function (complex envelope) has multiple peaks, or equivalently if the spectrum of the modulation function function is not continuous down to zero frequency (Rihaczek, 1969, p80). Therefore, a periodic pulse train with a repetition period $T$ has range ambiguities, and these are spaced by $R_u$, where

$$
R_u = \frac{cT}{2}
$$

(1.71)

i.e. a target displayed at range $R$ may in fact be at a range of $R + nR_u$, where $n$ is a positive integer. $R_u$ is thus the maximum unambiguous range and targets at ranges greater than this are called multiple-time around targets or phantom targets.

The simplest method to distinguish between real and phantom targets is to jitter the signal repetition period $T$. This causes the apparent range of the phantom targets to vary, but only the real targets are displayed at the correct range. The disadvantage of this scheme is that measurements over several repetition periods are required to resolve the ambiguity. For second-time targets only two different repetition periods are required; higher degree multiple-time targets require a correspondingly greater
number of different repetition periods (Skolnik, 1980, p53). Other schemes to resolve range ambiguities include coding of successive pulses, by changing the frequency, phase, pulse amplitude or pulse duration. With radar there is also the option of different polarisations. In practice, the major limitation of these schemes is due to the masking, by nearby strong targets, of the weaker multiple-time around targets which suffer greater transmission losses from being at greater ranges. Also, because several measurements are required to resolve the range ambiguity, there is no improvement in the area mapping rate.

To reduce the likelihood of receiving phantom targets, the transmit or receive elevation beam patterns are designed to prevent illumination of ambiguous swaths (Rihaczek, 1969, p461). Also just as importantly, illuminating only the desired region of the sea floor reduces the level of unwanted reverberation (see Section 1.7). The pulse repetition period is then selected so that the maximum unambiguous range must be greater than the maximum range at which targets are expected — i.e. before the next pulse can be transmitted, the sonar must wait for all the echoes from the previous pulse to be received. Where the width of the swath to be mapped is much less than the minimum target range — e.g. satellite borne SAR — it is possible to have several pulses in transit before the echoes from the earlier pulses arrive. From Figure 1.5 it is apparent that the minimum pulse repetition period is thus

\[ T \geq \frac{2(R_{\text{max}} - R_{\text{min}})}{c} \]  

(1.72)

where \( R_{\text{min}} \) and \( R_{\text{max}} \) are the desired minimum and maximum slant ranges. However, it is necessary to use either separate transducers for transmission and reception, or temporal interlacing to prevent eclipsing of the received echoes by the transmitted pulses (Tomiyasu, 1978, p566). This technique can also be applied when the swath width is illuminated by a number of different beams.

The fact that the speed of acoustic propagation is many orders of magnitude slower than the speed of electromagnetic propagation, necessitates the use of much longer pulse repetition periods for sonars than radars. The major implication of this, is that single-beam high-resolutions side-scan sonars (both real and synthetic aperture) can only operate at very low tow-speeds. This tradeoff between the pulse repetition period and the tow-speed is studied in more detail in Section 2.4.
1.10.7 Azimuth Resolution

The angular resolution of a system can be defined as the minimum angular separation between two sources, targets, or scattering centres on a body at which they can be separately distinguished (Steinberg, 1976, p196). The angular accuracy of a system can be calculated in terms of the beamwidth and SNR, but the angular resolution is also complicated by the effects of phase differences and relative target strengths.

The classical theory of resolution was developed for noncoherent optics, but for acoustics (and coherent microwave radar) coherent resolution criteria need to be considered (Goodman, 1968, §6.5). In any case, angular resolution is proportional to beamwidth (but is not necessarily the same as the beamwidth). In this thesis the nominal angular resolution of a system is defined as its 3 dB beamwidth.

The nominal azimuth resolution of a side-scan sonar is thus \( \theta_B \), where

\[
\theta_B = 2\sin^{-1}\left(0.44\frac{\lambda}{L}\right) \approx 0.88\frac{\lambda}{L}
\]  

(1.73)

and \( L \) is the along-track transducer dimension. The nominal along-track resolution for small beamwidths is thus

\[
\Delta y \approx R\theta_B
\]  

(1.74)

The along-track resolution obtainable is dependent on range and degrades due to beam spreading. Thus targets (and features) of a similar size appear progressively more stretched the further they are away from the sonar. A constant along-track resolution requires all range focusing —i.e. the effective aperture size needs to be increased with focal range to maintain a constant f-number.

1.10.8 Azimuth Ambiguities

If there is continuous radiation over an aperture, angular ambiguities can be avoided. This is analogous to transmitting a continuous signal spectrum to avoid range ambiguities, see Section 1.10.6. However, when an aperture is comprised of a number of discrete sources, as for the case of a line array of point sources, a multi-lobed radiation pattern is produced. The additional lobes are termed grating lobes, and since they are of a comparable magnitude to the main lobe, angular ambiguities exist. As the point sources are extended in length, the grating lobes become spaced further apart, reducing the number of ambiguous lobes. Extending the lengths of each element until the gaps in the aperture are filled produces a single main beam with no ambiguous lobes.

In practice it is often too difficult to implement electrically long antennas with a continuous distribution, or with array elements spaced sufficiently close together to avoid grating lobes. Techniques to avoid grating lobes include the use of aperiodic or random arrays —i.e. arrays with an unequal element spacing (Steinberg, 1976, §7.1). The disadvantage is an increased sidelobe level (cf. Sections 2.2.1 and 2.6).

As mentioned in Section 1.6.2, a non-linear phase function can be used to lengthen the duration of a signal, which can then be compressed in the receiver by removing the non-linear phase function. A similar technique can be applied to disperse and compress an antenna beam (Urkowitz et al., 1962, pp2100–2101). On transmission, a non-linear phase shift is applied to the elements of the transmit array, widening the
beam to greater than $\lambda/L$. This non-linear phase shift is removed upon reception, compressing the beam to the order of $\lambda/L$ as before. The advantage of this technique is that the search area may be swept with a low gain beam, while still retaining the performance of a high gain beam (Rihaczek, 1969, p82). Note that a linear phase shift just steers the beam; in much the same way that a linear phase shift only delays the envelope of a signal.

1.10.9 Towfish Instabilities

Transducers were initially hull mounted, but the images obtained from this configuration were adversely affected by the movement and radiated acoustic noise of the survey vessel (Flemming, 1982b, p6). This led to the development of stable drogues (towfish) towed astern of the survey vessel and under the sea surface. There are a number of technical and economic reasons for deploying a side-scan sonar on a towfish.

The primary advantage is that a properly designed towfish (and tow-system) can decouple the movement of the survey vessel from the sonar. Mechanical and electrical stabilisation techniques are often much more difficult to implement, especially over the periods required for long-range sonars (Somers and Stubbs, 1984, p252). From an operational point of view, the use of a towfish allows a side-scan sonar to operate from a vessel of convenience, which allows easy deployment and servicing. Additionally, deploying the receiver hydrophone some distance away from the survey vessel reduces the received propeller and machinery noise level, which tends to dominate for frequencies below 10 kHz. However, a towed hydrophone is still subject to flow noise caused by turbulent flow across the hydrophone face (Urick, 1975, pp328–334).

Other advantages which accrue from the use of towfish mounted transducers include better propagation conditions — can operate below the thermocline (Cloet et al., 1982, p414) — reduced sea surface multipath effects, and reduced cavitation by operating at greater depths (Somers and Stubbs, 1984, p252).

The three most common modes of towfish instability are roll, pitch and yaw. For examples of images illustrating these three types of towfish instability see Flemming, 1982a, Figs. 7–9.

Rolling of the towfish occurs when the sonar track runs across a swell — i.e. parallel to the wave crests. The beam swings in elevation producing a periodic compression and stretching of the sonargraph in range. The most noticeable effect is that the edge of the water column — the area below the sonar which is not imaged — shifts periodically in phase with the rolling motion. With both port and starboard operation, a large rolling motion can cause an overlap between the two channels directly below the sonar.

Running into a heavy swell causes the survey vessel to abruptly speed up and slow down. In response the towfish rises then sinks in a pitching motion; as the towfish suddenly rises the beam tilts forward, scanning ahead, and as it sinks the beam scans backwards. Target dislocation thus occurs in the along-track direction, with a result that linear features are deformed into wiggles.

Yawing also produces a forward and backward scanning, but this time the beam swings in azimuth. The result is a displacement of targets in the along-track direction, which increase with range. However unlike pitching, yawing causes the port beam
to scan in the opposite direction to that of the starboard beam. Yawing is probably the most adverse form of towfish instability. Compensation requires electrical beam stabilisation to steer the beams in the desired direction, but this is complicated by the influence of pitch and roll (cf. Somers et al., 1978, pp17-18).
Chapter 2

Synthetic Aperture Sonar

Aperture synthesis is a technique that has found widespread use in many applications — optics, radar, sonar, acoustics, and seismology; the Synthetic Aperture Radar (SAR) being the most well known application. Rather than using large apertures (i.e. many wavelengths long), aperture synthesis involves the scanning of an aperture with a smaller array. Image processing techniques are then employed to simulate a larger filled array covering the scanned aperture.

Whatever the scanning method, the basic idea is to reduce the number of receiver (or source) elements at the expense of an increased scanning time. However, the scanning rate must be great enough to prevent motion blurring or a loss in signal to noise ratio. Aside from the saving of hardware, the synthetic aperture technique allows the formation of apertures far larger than is often possible with physical apertures — e.g. side-scan imaging radars (Cutrona, 1970).

An alternative to scanning an aperture with a moving transducer is to move the target instead. This technique generates what is called an inverse synthetic aperture (cf. Chen and Andrews, 1980b; 1980a; Walker, 1980), and has been applied to mapping of the minor planets and the moon (Shapiro, 1968).

Although most synthetic aperture systems use active echo ranging principles, there are applications that synthesise apertures passively; the most common application probably being radio astronomy (Christiansen and Högbom, 1985). This thesis, however, is concerned primarily with the application of active SAR techniques to underwater imaging sonars.

Whereas SAR is a mature technique with sophisticated methods of processing, the unclassified Synthetic Aperture Sonar (SAS) literature is much less abundant and often less advanced (de Heering, 1982). Most of the synthetic aperture sonar literature is concerned with theoretical or model tank work, and has originated from the Nagatsuta Tokyo Institute of Technology (e.g. Ikeda et al., 1979b; Ikeda and Sato, 1980; Ikeda et al., 1985; Sato et al., 1973; Sato and Ikeda, 1977a). The only examples of imaging synthetic aperture sonars found in the literature are an experimental rail based system (Loggins et al., 1982) and a short-range low frequency system designed primarily for sub-bottom imaging (Dutkiewicz and Denbigh, 1987b; 1987a).

Other experimental synthetic aperture sonars have been developed in the fields of ultrasonic medical imaging (Burckhardt et al., 1974) and ultrasonic non-destructive testing (Thomson, 1984). However, these applications are not impeded by the slow
speed of acoustic propagation.

There are a number of differences between a synthetic aperture radar and its sonar equivalent. The major difference is the large disparity in the speed of microwave propagation through the atmosphere (nominally $3 \times 10^8$ ms$^{-1}$) compared to the speed of sound through seawater (nominally 1500 ms$^{-1}$). To achieve a comparable range resolution, radars must therefore transmit a signal bandwidth many orders of magnitude greater than sonar signals. Thus the basic signal processing and storage requirements are many times greater than sonar. On the other hand, a slow speed of acoustic propagation results in a very slow mapping rate for sonars. This is because the propagation delays are much longer for sonar, even when compared with spaceborne radar. Consequently, many of the approximations (and underlying assumptions) made in the analysis of synthetic aperture radars are not applicable to synthetic aperture sonars.

The direct application of SAR techniques to a practical synthetic aperture sonar is, in addition, complicated by the following factors:

(i) Slow mapping rates due to the relatively slow speed of sound in seawater.

(ii) Coherence problems resulting from the difficulty of maintaining a straight trajectory over the length of the synthetic aperture.

(iii) Coherence problems related to an inhomogeneous and turbulent medium.

In general, both these coherence problems are compounded by a slow mapping rate. The slower the speed that the aperture is traversed, the more likely it is that the survey vessel will deviate from a straight track.

It has long been thought that the underwater medium could not support a coherent acoustic synthetic aperture. However, experiments have shown that the medium is not as bad as has been thought and is in fact a lot better (Williams, 1976). Nevertheless, it is the medium coherence that limits the performance of any synthetic aperture sonar. This is examined further in Chapter 5.

Whereas the size constraints of a satellite or aircraft impose a maximum physical aperture length, with sonars it is possible to tow long hydrophone arrays (eels) astern of the survey vessel. Therefore, some of the advantages of synthetic aperture processing are lost on sonar. In addition, a towed array can be configured to provide multiple horizontal beams thus alleviating the mapping rate problems (see Section 2.5.1). However, all-range focusing with towed arrays involves a considerable hardware complexity. A combination of towed array and synthetic aperture techniques may thus be useful for reducing this hardware complexity or for improving system performance. One idea that comes to mind is to tow a ‘thinned’ array and to use synthetic aperture processing to remove the azimuth ambiguities inherent with this array.

This chapter starts out by describing, in Section 2.1, the sonar geometry used in the remainder of this thesis. It then looks at the concepts involved with typical narrowband synthetic aperture processing (the radar case) with regards to resolution (Section 2.3), ambiguities (Section 2.4) and mapping rate (Section 2.5). Section 2.6 looks at the use of broadband signals to improve range resolution and possibly mapping rate, and Section 2.7 looks at the option of incoherent processing, in applications where the signal phase lacks coherence or cannot be measured. Finally, the effects of
phase errors and motion compensation are briefly considered (Section 2.8), and the
chapter concludes in Section 2.9 with a brief discussion on the choice of operating
frequency.

2.1 Synthetic Aperture Side-Scan Sonar Geometry

The geometry of a side-scan sonar, with respect to a single isolated target at \( T \), is
shown in Figure 2.1. The sonar is described by a position vector \( \mathbf{r}_s = (x_s, y_s, z_s) \)
and the target by another position vector \( \mathbf{r}_t = (x_t, y_t, z_t) \). The target position can
also be specified, relative to the sonar position, using a range vector \( \mathbf{r} = \mathbf{r}_t - \mathbf{r}_s \).
It is often convenient to describe this range vector in spherical polar coordinates as
\( \mathbf{r} = (R, \theta, \psi) \), where the slant range \( R \), azimuth angle \( \theta \), and elevation angle \( \psi \) are
calculated from

\[
R = \left| \mathbf{r} \right| = \sqrt{(x_t - x_s)^2 + (y_t - y_s)^2 + (z_t - z_s)^2} \tag{2.1}
\]

\[
\theta = \tan^{-1} \left( \frac{y_t - y_s}{x_t - x_s} \right) \tag{2.2}
\]

\[
\psi = \sin^{-1} \left( \frac{z_t - z_s}{R} \right) \tag{2.3}
\]

Note that the z-axis goes upward and therefore targets below the sonar are at negative
angles of \( \psi \). Similarly with the sonar travelling parallel to the y-axis, targets forward
of the sonar are at positive angles of \( \theta \).

![Figure 2.1. Geometry of a side-scan sonar.](image)

Most side-scan sonars are designed to operate with small grazing angles —
i.e. \( |\psi| < 10^\circ \) — by deploying the sonar close to the sea floor, and so to a good
approximation, the angular target position in azimuth, \( \theta \), is

\[
\theta \approx \sin^{-1} \left( \frac{y_t - y_s}{R} \right)
\]  

(2.4)

To simplify the notation, the sonar is considered to move along a track, parallel to, and above, the y-axis (i.e. \( x_s = 0 \)) and at a constant height \( h = z_s - z_t \) above the sea floor. The sonar beam is also assumed to be perpendicular to the sonar track — i.e. parallel to the x-axis. With these assumptions, the slant range to a target at \((R_0, y_t)\) for a sonar at \((0, y_s)\) is

\[
R(y_s) = \sqrt{R_0^2 + (y_t - y_s)^2}
\]  

(2.5)

where \( R_0 \) is the nearest slant range to the target, which occurs when the sonar is abeam of the target — i.e. when \( y_s = y_t \),

\[
R_0 = \sqrt{x_t^2 + (z_t - z_s)^2}
\]  

(2.6)

The target range locus, as defined by (2.5), is one sheet of a hyperbola with a focus at \((\sqrt{2} R_0, y_t)\). Because of the side-scan geometry, the range histories of the targets are completely defined by only two parameters: the minimum target range \( R_0 \) at broadside, and the sonar position \( y_s = y_t \) at which the target range goes through the minimum. Expanding (2.5) in a binomial series and retaining the first two terms, the target locus can approximated by a parabola, i.e.

\[
R(y_s) \approx R_0 + \frac{(y_t - y_s)^2}{2R_0}
\]  

(2.7)

where the range error, \( R_e \), incurred by neglecting the third term in the series expansion is

\[
R_e = -\frac{(y_t - y_s)^4}{8 R_0^3} \approx -\frac{1}{8} R_0 \tan^4 \theta
\]  

(2.8)

Assuming a range error of \( \lambda/8 \) is tolerable, the parabolic approximation is valid for angles satisfying

\[
|\theta| < \tan^{-1} \left( \frac{\lambda}{R_0} \right)^{\frac{1}{4}} \approx \left( \frac{\lambda}{R_0} \right)^{\frac{1}{4}}
\]  

(2.9)

For example, let \( \lambda = 0.1 \) m and \( R_0 = 100 \) m. From (2.9) the approximation is valid for angles less than \( 10^\circ \) from boresight.

Throughout this thesis, image and target positions are specified either in Cartesian cross-track and along-track \((x, y)\) coordinates, or in polar slant-range and azimuth \((R, \theta)\) coordinates. In many cases the obliquity factor \( \cos \psi \) is ignored, and the target cross-track position is approximated by the minimum slant range \( R_0 \). This introduces a slight geometric distortion which is negligible for small grazing angles.

### 2.1.1 Apparent Target Motion

Although the sonar moves past the targets at a constant along-track velocity, it is useful to assume that the sonar is stationary and that the targets travel through the sonar beam. Thus the theory of moving target resolution can be applied to study the performance of a synthetic aperture sonar.
Assuming a uniform sonar speed $v$ and a stationary medium, the apparent target range can be written as a function of time by substituting $vt$ for $y_t$ in (2.5). This gives

$$R(t) = \sqrt{R_0^2 + (y_t - vt)^2}$$  \hspace{1cm} (2.10)

Differentiating (2.10) with respect to time, the range rate (or relative radial velocity) $\dot{R}(t)$, and range acceleration $\ddot{R}(t)$ are

$$\dot{R}(t) = -v \frac{y_t - vt}{R(t)} \approx -v \sin \theta(t)$$  \hspace{1cm} (2.11)

$$\ddot{R}(t) = \frac{v^2 R_0^2}{R(t)^3} \approx \frac{v^2}{R_0}$$  \hspace{1cm} (2.12)

where $\theta(t)$ is the target position in azimuth and the approximations are valid for small grazing angles and small beamwidths. The higher order range derivatives are found from successive differentiation, but for small angles near broadside they have a negligible effect. Therefore, the apparent target motion can be described in terms of a constant range acceleration (Rihaczek, 1969, p452).

To determine the effect of this range acceleration, (2.12) can be substituted into (1.42) to give

$$T_p < \frac{\sqrt{R_0 \lambda}}{2v}$$  \hspace{1cm} (2.13)

If the above condition is satisfied then the range acceleration has negligible effect on the signal. For example, consider a simple chirped pulse with a centre frequency $f_c = 22.5$ kHz reflected from a target at, say, a near range $R_0 = 50$ m. If the sonar moves at a speed $v = 5$ m s$^{-1}$, then the maximum pulse length from (2.13) is $T_p = 180$ ms.

Therefore, at typical tow speeds it is possible to neglect the effects of range acceleration on most sonar signals, and to treat the apparent target motion as one of constant range rate over the duration of each pulse.

### 2.1.2 Target Phase History

Assuming that a stationary point target is illuminated by a monochromatic CW signal of wavelength $\lambda$, the phase of the demodulated signal as a function of sonar
position $y_s$ is

$$\phi(y_s) = -4\pi \frac{R(y_s)}{\lambda} \approx -4\pi \frac{R_o}{\lambda} - 4\pi \frac{(y_t - y_s)^2}{2R_0\lambda}$$  \hspace{1cm} (2.14)$$

Notice that this phase variation is comprised of a quadratic phase term superimposed on a linear phase term. It is this quadratic variation along the sonar track that is removed by processing to focus the synthetic beam. Also note that as the target range $R_0$ increases, the quadratic phase term has less effect and the resultant phase variation becomes linear. When the target is in the far-field of the synthesised aperture, there is no quadratic phase variation and hence the along-track resolution cannot exceed half the length of the synthetic aperture. This is still a factor of two better than real apertures due to the two way phase shift incurred.

### 2.1.3 Target Doppler History

The result given by (2.14) assumes that the sonar is stationary for each measurement. If, however, the sonar is moving, the echo signals are Doppler shifted, and it is now necessary to calculate the demodulated phase using the instantaneous propagation delay $\tau(t)$. From (1.33) the phase of the instantaneous demodulated signal is

$$\phi_d(t) = -2\pi \frac{c\tau(t)}{\lambda}$$  \hspace{1cm} (2.15)$$

and differentiating this gives the Doppler shift $f_d(t)$, where

$$f_d(t) = \frac{1}{2\pi} \frac{d\phi_d(t)}{dt} = -\frac{c\dot{\tau}(t)}{\lambda}$$  \hspace{1cm} (2.16)$$

In most radar applications, however, the Doppler shift is approximated as

$$f_d(t) \approx -\frac{2\dot{R}(t)}{\lambda} \approx \frac{2v}{\lambda} \sin \theta(t)$$  \hspace{1cm} (2.17)$$

where $\dot{R}(t)$ is found from (2.12). It is thus apparent that the Doppler frequency shift is proportional to the component of the sonar velocity in the direction of the target. When the sonar approaches the target, the angle $\theta$ is positive, and an up-Doppler shift is produced. Conversely, a down-Doppler shift is produced for negative angles of $\theta$ when the sonar is receding from the target. These relationships are summarised in Table 2.1.

The radar approximation effectively assumes that the sonar is stationary during each phase measurement (i.e. $\tau(t) = 2R(t)/c$), and is equivalent to replacing $y_s$ with $vt$ in (2.14). The effect of this approximation is examined in Section 6.1.3 where it is found that the error is simply an additional linear phase shift across the aperture; a similar effect to skewing the beam from broadside. With airborne SAR systems the effect is negligible, but is significant for high speed SAS systems and spaceborne SAR.

If only small angles near broadside are considered (where the $\sin \theta \approx \theta$ approximation is valid), and using the radar approximation, it is apparent from (2.17) and (2.12) that the Doppler target history can be described by a linear FM signal, with a sweep rate $\mu_d$ proportional to the range acceleration $\ddot{R}$, where

$$\mu_d = \frac{2\ddot{R}}{\lambda} \approx \frac{2v^2}{\lambda R_0}$$  \hspace{1cm} (2.18)$$
Approaching Target (up-Doppler) | Receding Target (down-Doppler)
--- | ---
\( y_s < y_t \) | \( y_s > y_t \)
\( \dot{\theta}(t) > 0 \) | \( \dot{\theta}(t) < 0 \)
\( \dot{R}(t) < 0 \) | \( \dot{R}(t) > 0 \)
\( \eta(t) > 1 \) | \( \eta(t) < 1 \)
\( f_d(t) > 0 \) | \( f_d(t) < 0 \)

Table 2.1. Doppler relationships.

In addition to the linear frequency modulation, there is also a slow amplitude modulation due to the target passing through the sonar beam. This acts as a natural aperture shading function.

The time taken for the target to pass through the beam of the sonar is known as the dwell time, or time in beam, and is given by

\[
T_d = \frac{2R_0}{v} \tan \left( \frac{\theta_B}{2} \right) \approx \frac{R_0 \theta_B}{v} \tag{2.19}
\]

where \( \theta_B \approx \lambda/L \) is the 3 dB beamwidth in azimuth.

The Doppler bandwidth of each target history, or Doppler spread, is given approximately by the product of the dwell time \( T_d \) and the Doppler sweep rate \( \mu_d \). Combining (2.20) and (2.18) yields

\[
W_d \approx \frac{2v}{\lambda} \theta_B \approx \frac{2v}{L} \tag{2.21}
\]

Note that the increased Doppler shift at higher frequencies is compensated by a reduction in beamwidth, and therefore the Doppler spread is independent of the signal wavelength and depends solely on the tow speed \( v \) and the along-track transducer dimension \( L \).

2.2 Sampled Apertures

It is well known that a continuous source antenna can be approximated by an array of discrete elements, spaced sufficiently close together (Walter, 1965, p1). Since the continuous aperture has effectively been sampled at discrete regions, sampling theory may be applied to study the properties of the array, e.g. the element spacing requirements.

There are two ways in which an aperture can be sampled. The most common method uses a physical array of antenna elements, typically spaced at regular intervals. The other method uses only a few elements, typically one, which are moved along the aperture in discrete intervals forming what is known as a synthetic array.
2.2.1 Real Arrays

Consider a linear array made up of $N$ transducers, each of length $L$, and uniformly spaced by an amount $\Delta L$. If each transducer element has a far field pattern $F_e(\theta)$ corresponding to an aperture distribution $f_e(y)$, then summing the outputs of the elements together gives a combined aperture distribution of

$$f_r(y) = \sum_{n=1}^{N} f_e(y - n\Delta L) = f_e(y) \otimes \sum_{n=1}^{N} \delta(y - n\Delta L)$$  \hspace{1cm} (2.22)

and a normalised far field pattern $F_r(\theta)$, where

$$F_r(\theta) = F_e(\theta) \frac{\sin[\pi N(\Delta L/\lambda)\sin\theta]}{N \sin[\pi(\Delta L/\lambda)\sin\theta]}$$  \hspace{1cm} (2.23)

The effect of the aperture sampling is to multiply the beam pattern of a transducer element (the element factor) by an array or space factor. The 3 dB beamwidth, or nominal azimuth resolution (see Section 1.10.7), of the real array is thus

$$\theta_r = 2\sin^{-1}\left(\frac{0.44}{L_r}\right) \approx 0.88 \frac{\lambda}{L_r}$$  \hspace{1cm} (2.24)

where $L_r = N \Delta L$ is the effective length of the array. Note that the azimuth resolution is a function of the array length and is independent of the number of array elements. However if $\Delta L > \lambda/2$ (i.e. $N < 2 L_r/\lambda$), the array factor becomes multi-lobed causing additional lobes to appear in the array beam pattern with an amplitude comparable to the main beam. These ambiguous lobes are called grating lobes and occur when the denominator of (2.23) becomes zero — i.e. when $\pi(\Delta L/\lambda)\sin\theta = n\pi$. The angle of the $n$th grating lobe is thus

$$\theta_n = \sin^{-1}\left(\frac{n\lambda}{\Delta L}\right)$$  \hspace{1cm} (2.25)

If the array elements are spaced closer than $\lambda$, the grating lobes occur at invisible angles — i.e. $|\sin\theta| > 1$. However if the array is to be steered from endfire to endfire ($-\pi/2 < \theta < \pi/2$), the array elements need to be spaced closer than $\lambda/2$ to prevent grating lobes appearing in the scanning region (Steinberg, 1976, §5.2). In practice, the elements are usually spaced by less than $\lambda/4$ to reduce the sidelobe level of the array. Note that if the array elements have any directionality — i.e. are not isotropic — ambiguities due to grating lobes are reduced by the element factor.

An alternate method of avoiding grating lobes is to use an aperiodic array, i.e. an array with an unequal element spacing (Steinberg, 1976, §7.1). Since the sampling function is not periodic, neither is its Fourier transform, and hence grating lobes fail to arise.

The array considered so far is termed unfocused, or focused at infinity. The along-track resolution which can be achieved in the far-field is $R\theta_r$, and the limiting near-field along-track resolution is $L_r$, the length of the transducer array. However by focusing the array, it is possible to achieve an along-track resolution of $L$ at the focal range $R_f$. 


2.2.2 Synthetic Arrays

Instead of using $N$ individual transducers, the complex wavefield can be sampled by moving a single transducer along the aperture and storing the measurements. These samples can then be summed together to form an unfocused synthetic beam pattern. Usually the transmit and receive transducers both move and are indeed often the same transducer in pulsed systems. Thus the path length to a target changes twice as much with position than if the transmitter was stationary and the receiver moved. Hence the normalised synthetic beam pattern is

\[ F_s(\theta) = \frac{F_t(\theta)}{N} \frac{\sin[2\pi N(\Delta L/\lambda) \sin \theta]}{\sin[2\pi (\Delta L/\lambda) \sin \theta]} \]  \hspace{1cm} (2.26)

and from analogy with the real beam pattern (2.26), the 3 dB beamwidth, or nominal azimuth resolution, of the synthetic array is

\[ \theta_s = 0.88 \frac{\lambda}{2L_s} \] \hspace{1cm} (2.27)

where $L_s$ is the length of the synthetic aperture. For the remainder of this thesis the 0.88 scale factor is approximated by unity. In practice, the nominal azimuth resolution is slightly poorer than that given by (2.27), but the sidelobe level is better. This is a result of a natural array shading imposed by the real beam pattern.

Notice that the synthetic array beamwidth is half the achievable beamwidth of a real array of the same length, and is due to the two-way phase shifts involved. Alternatively, a synthetic array can achieve the same azimuth resolution as a real array, but by using an array of half the length. However, to avoid grating lobes in the synthetic beam pattern, it is necessary to have the elements spaced twice as closely as those of a real array.

The array analogy is not strictly accurate. For this to be so, the scanning transducer must be stationary during the pulse transmission and reception, and then skip to the next position before transmission of the next pulse. In systems, such as synthetic aperture radar, where the scanning transducer moves during signal transmission and reception, the return signals are Doppler shifted. Provided, however, that the pulse durations are too short to be sensitive to Doppler, the transducer may be considered stationary while transmitting and receiving (Rihaczek, 1969, p447). This effect is considered further in Sections 2.1.3 and 6.1.3.

2.2.2.1 One-Way / Two-Way Beam Patterns

Synthetic beam patterns are often known as two-way beam patterns. This is because the relative phase shift along the aperture depends on both the transmitted and received path lengths. For real arrays, the relative phase shifts depend only on the received path lengths; the transmitted path length being common to all array elements (note that in some applications the converse is true). Hence real beam patterns are known as one-way beam patterns.

2.3 Synthetic Aperture Resolution

There have been many different approaches used to study the performance of synthetic aperture systems, notably in the field of synthetic aperture radar (cf. Ri-
haczek, 1969, Ch.13; Cutrona, 1970; Tomiyasu, 1978). In Section 2.2 the synthetic aperture technique was introduced using the concept of sampled apertures. This section now looks at synthetic aperture processing from the viewpoints of beam compression, Doppler, and close target separability.

### 2.3.1 Beam Compression

The synthetic aperture technique may be viewed as a beam sharpening, or beam compression, process. A real beam of width $\theta_B$ is transmitted and by using beamforming techniques a much narrower beamwidth $\theta_s$ is synthesised. To achieve an along-track resolution, $\Delta y$, at a range $R_0$, requires a synthetic beamwidth $\theta_s = \Delta y/R_0$. This can be achieved by synthesising an aperture of length $L_s$, where

$$L_s \approx \frac{\lambda}{2\theta_s} \approx \frac{R_0 \lambda}{2\Delta y} \quad (2.28)$$

To ensure that a target at a range $R_0$ is illuminated over this aperture, the minimum required real beamwidth is thus

$$\theta_B \approx \frac{L_s}{R_0} = \frac{\lambda}{L} \quad (2.29)$$

and combining (2.29) with (2.28) gives the well known result that

$$\Delta y = \frac{L}{2} \quad (2.30)$$

Equation (2.30) gives the highest along-track resolution obtainable with a transducer of length $L$, assuming that all the reflected echoes from a target are coherently processed. Using only a fraction $\rho$ of the real beamwidth, the along-track resolution degrades to $\Delta y = L/2\rho$, and the beam sharpening factor similarly reduces to

$$\frac{\theta_B}{\theta_s} = \frac{2\rho L_s}{L} = \frac{2\rho R_0 \lambda}{L^2} \quad (2.31)$$

The fact that the along-track resolution is independent of range and wavelength (assuming that the necessary aperture length can be synthesised) is the chief motivation for the use of synthetic aperture processing. The along-track resolution is improved with a shorter transducer, contrary to traditional imaging techniques. However, to achieve this result, it is necessary to synthesise longer apertures at farther ranges and lower frequencies; the asymptotic along-track resolution being $\lambda/4$ (Harger, 1970, p40).

In practice, the along-track resolution is limited by the coherence length of the medium stability and sonar motion. If the coherence length is shorter than the length of the desired synthetic aperture then there is a degradation in along-track resolution with range.

The frequency dependence of a practical synthetic aperture sonar is more difficult to predict. Increasing the frequency reduces the time needed to traverse the synthetic aperture, but also amplifies the phase errors caused by path length variations.

One disadvantage of traditional narrowbeam imaging sonars is that they only manage a few 'hits-on-target'. It is therefore conceivable that certain targets may be
missed altogether, especially if the target highlights are at an aspect not illuminated by the sonar.

With a widebeam sonar, a target is viewed over a greater range of angles, and since it is in the beam for a longer interval, there are more hits-on-target. In theory, a single beam sonar has \( N_t \) hits-on-target, where

\[
N_t \approx \frac{pR_0\lambda}{vTL}
\]

and substituting for the maximum tow speed from (2.38), the minimum number of hits is thus

\[
(N_t)_{\text{min}} \approx \frac{2pR_0\lambda}{L^2}
\]

2.3.2 Doppler Processing

The operation and performance of a synthetic sonar can be evaluated from consideration of the Doppler spread of a target echo. As shown in Section 2.1.3, the Doppler echo signal resembles a linear FM waveform over small angles in azimuth. This suggests that pulse compression techniques may be applied to compress the Doppler echo signal in azimuth.

Figure 2.3 shows the Doppler frequency history for two point targets at the same range as a function of time.

Figure 2.3 shows the Doppler frequency history for two point targets at the same cross-track position, but separated in the along-track direction by \( \Delta y \). Since the targets are at the same range of closest approach, \( R_0 \), the Doppler frequency histories have the same slope \( \mu_d \), but are displaced in time by \( \Delta y/v \) and in frequency by approximately \( \mu_d\Delta y/v \). Substituting for \( \mu_d \) from (2.18), this Doppler frequency difference can also be expressed as

\[
\Delta f_d \approx \frac{2v\Delta y}{\lambda R_0}
\]

Therefore, to resolve the two targets in the along-track direction, the Doppler processor needs to be able to resolve a frequency difference of \( \Delta f_d \). In theory, this requires the targets to be illuminated for a duration of at least \( 1/\Delta f_d \). Equating this to the
dwell time $T_d$ given by (2.20), the along-track resolution $\Delta y = L/2$ is obtained as before.

There are two focused Doppler processing methods. In one method, matched filters similar to pulse compression filters track the Doppler azimuth history of each along-track bin (Wu et al., 1982). Alternatively, a slanted time/frequency reference signal is used to demodulate the Doppler histories to constant frequencies — a process known as *deramping* (Caputi, 1971), similar to heterodyne correlation of CTFM signals as described in Section 3.2. The demodulated frequency is proportional to the delay between the reference and Doppler histories, and therefore a bank of Doppler filters can separate targets at different along-track positions. The minimum number of Doppler filter bins required is

$$N_d = \frac{\rho W_d}{f_d} = \frac{2\rho R_0 \lambda}{L^2} \tag{2.35}$$

where $\rho$ is the fraction of the real beamwidth $\theta_B$ utilised in the reconstruction process. Note that this number is independent of the tow speed and is equal to the beam sharpening factor defined in (2.31).

Although this deramping approach is more computationally efficient than the basic matched filtering (or fast convolution) approach, the processing is complicated by the fact that the Doppler rate is different for targets at different ranges. There are also difficulties of efficiently using global block operations (such as the FFT) because of the staggering of the along-track histories (Sack et al., 1985, p47).

### 2.3.3 Close Target Separability

![Figure 2.4. Differential target range of two targets separated in track.](image)

The theory of moving target resolution (Section 1.4.6) can be applied to the determination of the along-track resolution for a synthetic aperture sonar. Figure 2.4 shows two point targets separated by $\Delta y$, at a nearest range $R_0$ to the sonar track. With the targets at an angle $\theta$, the differential target range, $\Delta R$, can be seen to be approximately $\Delta y \sin \theta$. Therefore, if the targets are observed over the nominal azimuth beamwidth $\theta_B$, the change in differential range is $\Delta R \approx \Delta y \sin \theta_B$. Using
the half-wavelength criterion ($\Delta R = \lambda/2$) the along-track resolution is thus

$$\Delta y = \frac{\lambda}{2 \sin \theta_B} \approx \frac{\lambda}{2 \theta_B}$$  \hspace{1cm} (2.36)

Substituting for the nominal beamwidth, $\theta_B = \lambda/L$, yields the along-track resolution, $\Delta y = L/2$, in agreement with (2.30). Equation (2.36) also applies when a narrow beam is used to track the targets over a larger angle $\theta_B$—e.g. spotlight SAR (cf. Brookner, 1978b).

$$\Delta r \approx \frac{2\lambda}{\sin ^2 \left( \frac{\lambda}{4} \theta_B \right)} \approx \frac{2\lambda}{\theta_B^2}$$  \hspace{1cm} (2.37)

where the approximation is valid for small angles. Comparing (2.36) with (2.37) shows that range resolution in this manner is inferior to the along-track resolution capability of synthetic aperture systems (Rihaczek, 1969, p459). For example, consider a sonar with a wavelength $\lambda = 0.1$ m and a nominal beamwidth $\theta_B = 10^\circ$ (0.175 radians). From (2.36) the highest along-track resolution is $\Delta y = 0.29$ m and from (2.37) the highest range resolution is $\Delta r = 3.28$ m.

Note that this range resolution is achievable without range gating, so in theory a monochromatic CW signal may be used instead of a pulsed signal. Without range gating, however, the image signal to reverberation (clutter) ratio is much poorer. In addition, the resolution of objects in range is limited to selective focus or perspective/parallax, which are only useful with small focal ratios.
2.4 Ambiguities

Up to this point, a monochromatic CW sonar has been assumed. Although an unmodulated CW signal can achieve range resolution on the basis of sonar motion, Section 2.3.3 concludes that it is desirable to transmit a modulated signal. The signal modulation function (or complex envelope) provides the required range resolution capability and the signal carrier provides the phase (or Doppler) measurements necessary for the beam compression.

To achieve the conflicting requirements of a large signal bandwidth for range resolution, a long signal duration for azimuth resolution, and minimal self-clutter, the coherent pulse train is usually the preferred signal (Rihaczek, 1969, Ch.8). This signal has range ambiguities spaced by $cT/2$, where $T$ is the repetition period (see Section 1.10.6, and therefore $T$ must be great enough to permit unambiguous operation over the desired swath width. In addition, since the aperture is now sampled at the pulse repetition frequency (prf), the distance moved per pulse, $vT$, must be small enough to prevent spatial undersampling of the aperture field. There is obviously a compromise between the desired tow speed $v$, swath width and along-track resolution.

Assuming that the along-track dimension $L$ of the transducer and pulse repetition period $T$ have been chosen to give the desired along-track resolution and unambiguous range, the maximum towed speed of the sonar can be estimated from a number of different but equivalent methods (cf. Rihaczek, 1969; Tomiyasu, 1978); two of which are considered here.

2.4.1 Grating Lobe Suppression

One approach is to consider the grating lobes of the synthetic beam pattern caused by (under)sampling along the aperture. With a narrowband sonar the $n$th grating lobe from boresight occurs at an azimuth angle $	heta_n = \sin^{-1} n\lambda/(2vT)$, where $v$ is the sonar speed, $\lambda$ is the acoustic wavelength, and $T$ is the pulse repetition period. Adequate suppression of the grating lobes can be achieved by ensuring that they coincide with the nulls in the real-beam pattern (Rihaczek, 1969). These are similarly spaced in azimuth at angles $\phi_n = \sin^{-1} n\lambda/L$, and equating angles, $\phi_n = \phi_0$ (see Figure 2.6), gives

$$v_{\text{max}} = \frac{L}{2T}$$

(2.38)

i.e. the sonar cannot move more than half the real hydrophone dimension between pulses. Increasing the speed of a narrowband SAS, without introducing grating lobes, either requires a longer hydrophone with a consequential reduction in along-track resolution or a reduced pulse repetition period which causes the range ambiguities to be more closely spaced.

2.4.2 Doppler Sampling

The Doppler signal is sampled at the pulse repetition rate, and therefore to avoid aliasing, the Doppler spread $W_d$ must be less than the prf, assuming that the Doppler signal is bandpass sampled. However, the result remains the same if the Doppler signal is demodulated to baseband before being sampled. The baseband Doppler
spectrum varies from $-W_d/2$ to $W_d/2$ but since this is now a lowpass signal, it must be sampled at twice the maximum frequency $W_d/2$, i.e. $(\text{prf})_{\text{min}} = W_d$ as before. Combining this result with the expression for the Doppler spread (2.21), yields the same maximum tow speed as (2.38).

2.5 Area Mapping Rate

In the previous section it was noted that there was a compromise between tow speed, swath width, and along-track resolution. If the pulse repetition period is too short, range ambiguities occur, and conversely if the pulse repetition period is too long, or if the tow speed is too fast, then angular ambiguities appear in the synthetic beam pattern. There is thus a maximum rate at which a synthetic aperture sonar can image the sea floor.

More specifically, if the desired swath width is $\gamma c T/2$, and $v$ is the tow speed, then the area of the sea floor imaged per second (or area mapping rate) is thus

$$\dot{A} = \frac{1}{2} \gamma v c T$$

(2.39)

where $\gamma$ is the fraction of the maximum unambiguous range utilised.

From (2.39) it is apparent that it is necessary to increase the pulse repetition period $T$, or to increase the sonar speed $v$, to obtain an improvement in mapping rate. However, it must be remembered that increasing $T$ also increases the maximum unambiguous range, but increasing $v$ reduces the unambiguous scanning angle. Combining (2.38) and (2.39) yields the maximum unambiguous area mapping rate,

$$\dot{A}_{\text{max}} = \frac{1}{4} \gamma c L$$

(2.40)

Initially it seems that since both the along-track resolution $\Delta y$ and the mapping rate $\dot{A}$ are functions of the horizontal transducer length $L$, the only way to increase the mapping rate is to sacrifice along-track resolution and the number of integrated pulses by increasing the hydrophone length. However as indicated by Cutrona et al. (1961), the mapping rate can be increased by using multiple beams to provide greater coverage.

There have been a number of different multiple beam configurations proposed to improve the mapping rate. Although the implementations are considerably different, the general result is that the area mapping rate can be improved by a factor $n$, where...
\( n \) is the number of non-overlapping beams. There is, however, a significant increase in the complexity of both the hardware and processing requirements (de Heering, 1982, §3.4.4).

Similar concepts have been proposed to reduce the number of receiver elements in acoustic holographic systems. Keating et al. (1970) and Nitadori (1970) show that the number of elements in the receiver array can be reduced by using a multiplicity of transmitters. To reduce the scanning time the transmitters can either be frequency or modulation coded (Keating et al., 1970; Keating et al., 1978).

### 2.5.1 Multiple Horizontal Beams

The use of multiple adjacent beams in the horizontal plane was proposed by Cutrona et al. (1961) in the context of SAR, and again by Cutrona (1975) in the context of SAS. To determine the number of adjacent beams, Cutrona suggests that transducer length \( L \), pulse repetition period \( T \), and towfish velocity \( v \), are chosen to satisfy the mapping rate requirements without introducing range or azimuth ambiguities. Then the number of beams required is calculated by dividing the resultant along-track resolution by the desired range resolution. The effectively wide beam thus satisfies the along-track resolution requirements, while the \( n \) narrow beams improve the sampling requirements \( n \)-fold.

### 2.5.2 Multiple Vertical Beams

Walsh (1969) and Lee (1979) propose a SAS operating with a number of adjacent vertical beams, possibly at different frequencies. A similar concept is described by Tomiyasu (1981) for a space borne radar. The advantage of illuminating a number of adjacent narrow swaths is that the pulse repetition period may be reduced. This relaxes the range ambiguity requirements, allowing a faster tow speed.

This method is in effect equivalent to a number of synthetic aperture sonars operating in parallel, but each mapping a different narrow swath. The number of beams is therefore calculated by dividing the desired mapping rate by the mapping rate achievable with a single beam.

There are a number of difficulties with this method. Each of the beam patterns needs to be carefully controlled in the vertical direction to avoid illuminating ambiguous swaths (cf. Okino and Higashi, 1986 and de Moustier, 1986). The ambiguity can be reduced by using multiple frequencies, but this may introduce problems relating to illumination of the swath at different frequencies. There is also the additional problem of interlacing the received and transmitted pulses, since there is likely to be several pulses in transit before the echoes arrive (see Section 1.10.6).

### 2.5.3 Multiple Hydrophones

The alternative to radiating a number of narrow beams is to transmit a single wide beam and to have \( n \) adjacent widebeam hydrophone elements (Kock, 1972; Gilmour, 1978; de Heering, 1982, p19). This is a much simpler technique and effectively samples the aperture \( n \) times faster than with a single hydrophone. Although \( n \) independent receivers are required, unlike the other multiple beam techniques the beamforming does not have to be performed in real-time.
There is also an added flexibility with this system. The hydrophones can be grouped together to act as one larger hydrophone, or as a number of smaller hydrophones, thus providing a variable along-track resolution capability without reducing the number of integrated echoes. In addition, the phase difference between pairs of hydrophones may be useful for determining the deviation of the sonar from a straight trajectory (cf. Raven, 1981).

2.5.4 Living with Range Ambiguities

Harger (1970, p27) and Castella (1971, p256) suggest that the mapping rate constraint can be lifted if one accepts range ambiguities. The echo history of a scatterer at an ambiguous range differs from that of an unambiguous scatterer, so a properly designed processor will not respond as well to the ambiguous target — i.e. it suppresses the range ambiguity in the output image. However, the ambiguous returns are smeared out in the correlation process and add to the background noise. This results in a degraded image quality.

Rihaczek (1969, p463) shows that pulse to pulse frequency jumping to suppress range ambiguities, or repetition period staggering to eliminate Doppler ambiguities, can only smear the ambiguous spikes into the background clutter. Often these spikes have an energy equivalent to the desired spike, and hence the suppression of a single spike introduces self-clutter, the total energy of which is of the same order as that of the signal. Suppression of $n$ ambiguous spikes increases the self-clutter $n$-fold, and therefore only those targets whose cross-section is sufficiently higher than the average cross-section per cell will be detectable.

2.6 Broadband Methods

To achieve a resolution in range comparable to the resolution in along-track it is necessary to transmit broadband signals (see Section 1.10.4). These are usually produced using either amplitude modulation (e.g. impulse-like waves), phase modulation (e.g. M-sequence waves), or frequency modulation (e.g. linear FM). In general, a higher signal to noise ratio is obtained by modulating the phase or frequency rather than the amplitude, since more energy may be transmitted. With impulse-like signals, for example, the number of cycles can be so few that detection of the phase and amplitude of the echo signals becomes difficult (Sato and Ikeda, 1977b, p341).

Farhat (1977, p1020) shows that coherent imaging with broadband illumination, whether produced by impulsive, swept frequency, or white noise source, obey similar principles. As well as improving range resolution, broadband signals reduce the effects of coherent artefacts (grating lobes), reduce resonance effects in the image and provide increased object information (assuming that the object is non-dispersive and non-resonant). In addition, a single frequency pulse sonar tends to support the fine lobe structure of a target diffraction pattern, while broadband signals tend to give more of an average value, thereby improving detection at a random aspect (Winder, 1975, p294).

The disadvantages of broadband signals include increased signal processing requirements, more difficult transducer design, and a variable beam pattern. With a
slowly changing linear frequency sweep the signal/beam pattern interaction is predictable, but the behaviour of acoustic arrays is markedly different for pulsed signals. Rather than relying on steady state theory, the transient theory must be considered (Fink, 1980).

The processing complexity of broadband systems is often simplified by filtering the echoes into a number of narrowband frequency components that are processed independently. This technique is termed spectral decomposition because many narrowband signals are derived from a single wideband signal, and has been suggested for both impulsive signals (Nagai, 1984) and linear FM sweeps (Robinson, 1982, §4.3).

Since sonar/target motion has a greater effect on longer duration signals, it is also desirable, if possible, to split long duration signals into shorter components for processing. The processing of these shorter length components is often less complicated than processing the entire signal, and then afterwards they can be recombined to produce the desired resolution.

The converse technique to spectral decomposition has also been proposed (Sato and Ikeda, 1977b; Sato et al., 1981) where a number of CW signals are sequentially transmitted to synthesise the spectrum of an equivalently very short pulse. Sato and Ikeda (1977b) claim that the frequency scanning also acts as the aperture synthesis, without having to move the receiver, and as a method for interpolating or extrapolating the scanning region. This may be useful for reducing the number of sampling points and hence increasing the scanning speed. A similar idea is suggested by Farhat and Chan (1980) where a frequency diverse signal (such as a swept FM signal) is used to provide aperture synthesis. However, this technique only works for off-axis targets, and in addition there is an extra quadratic phase term dependent on the (unknown) target range that must be corrected for.

2.6.1 Smearing of Grating Lobes

The maximum towed speed derived in Section 2.4 assumes narrowband operation where the positions of the grating lobes are fixed relative to the main beam. However for broadband operation, the grating lobes tend to be smeared since their positions are frequency dependent. The broader the bandwidth, the better the overall beam pattern.

When the second grating lobe of the synthetic beam pattern at the highest transmitted frequency is at or beyond the first grating lobe of the lowest transmitted frequency, the grating lobes are completely smeared (defocused) (de Heering, 1984). This condition requires the transmitted signal to have a bandwidth of at least an octave, i.e. a quality factor $Q \leq \sqrt{2}$, assuming that the echo strength is constant over the transmitted bandwidth. Note that the grating lobes are smeared, but not cancelled. Although they are reduced in relationship to the desired main lobe, the smeared grating lobes contribute to the background reverberation and therefore limit the dynamic range of the resultant image. However, the aperture may now be sampled at a faster rate than that required for a narrowband SAS. Thus with a broadband SAS there is a trade-off between mapping rate and image dynamic range. In addition, traversing the aperture more rapidly may improve the overall image quality in conditions where the temporal coherence is affected by turbulence.
Hildebrand (1980) shows that the number of transducer elements \( N \) required to prevent grating lobes in a pulsed array of length \( L \) is

\[
N \geq \frac{2cT_p}{\lambda}
\]  

(2.41)

where \( T_p \) is the pulse length and \( \lambda \) is the wavelength of a gated sinusoid. Thus even a two element interferometer may be unambiguous if the pulse length is short enough to prevent interference effects. However in this case, the amplitude at focus is only twice that in non-overlapping regions of the wave. Therefore, to improve the signal to self-clutter ratio, many more array elements should be added. The region of overlap is now increased and thus grating-lobes can now exist, but these are suppressed in an inverse proportion to the distance from the focus.

### 2.7 Incoherent Synthetic Aperture Processing

Envelope processing (de Heering, 1984, p279) (or incoherent SAS (de Heering, 1982, §6)) is a technique proposed to allow aperture synthesis when the echo signals lack phase coherence. The technique is similar to those proposed by Hanstead (1981) and Wingham (1988), and has the advantage that the towfish position need only be known to within a fraction of the pulse length, rather than the carrier wavelength. Envelope processing is therefore quite robust. Similar concepts related to tomographic imaging are also mentioned by Langenberg et al. (1986).

Instead of using the echo phase, envelope processing is based on the range curvature effect. A point \((x, y)\) is reconstructed by summing the intensities of the raw data along the locus described by (2.5). Like back projection techniques used in computed tomography, envelope processing produces a bright point at the target position sitting on a background 'fog'. Obviously, the finer the range resolution and the wider the beamwidth, the greater the signal to background ratio.

Strictly speaking, noncoherent integration does not improve the signal to noise ratio of the image. However for fluctuating targets and background, it is well known that noncoherent integration of \( n \) pulses produces an improvement of detection equivalent to approximately \( 5 \log_{10} n \), instead of \( 10 \log_{10} n \) achievable with coherent integration. This improvement is due to the fluctuations in the target and background being averaged, rather than a cancellation of the noise.

With these envelope processing schemes 'robustness' is obtained at the cost of reduced resolution and signal to background ratio (de Heering, 1984). However, summation on an intensity basis results in a significant reduction in speckle.

The incoherent along-track resolution of the envelope processing technique has been derived by de Heering (1984) to be

\[
\Delta y_I \approx QL
\]  

(2.42)

Like the resolution obtainable with a coherent synthetic aperture, this is proportional to the length of the real aperture. In addition, it also depends on the \( Q \) of the transmitted signal. Therefore, broadband signals are desirable for aperture synthesis, even when the received echoes lack phase coherence.
2.8 Phase Errors and Motion Compensation

The derivations of the azimuthal response of a synthetic aperture sonar has assumed that the sonar travels along a straight line and that there are no phase instabilities in the medium. In reality the sonar track is not truly straight and path length fluctuations exist. Thus there will be a degradation of performance if phase corrections are not made. However, it may be impossible to make corrections for any unpredictable medium induced phase error. These phase errors become more of a problem when operating at the longer ranges, and limit the size of the aperture that can be coherently synthesised (cf. Elachi and Evans, 1977; Williams, 1976).

There have been a number of theoretical treatments of SAR image degradation due to phase errors introduced by random target motion (e.g. the sea surface, Elachi and Evans, 1977), medium turbulence, and platform motion (Greene and Moller, 1962; Brown and Palermo, 1965; Brown, 1967; Porcello, 1970). Tow path instabilities and fluctuations of the acoustic path length in the medium produce a random phase error along the sonar track. Although these phase errors are random in nature, they are correlated in time, and distance, along the sonar track. For good mapping performance Castella (1971, p269) suggests that these phase errors must be compensated to better than 40° rms.

To reduce the effects of medium turbulence, Ikeda et al. (1979b) and Ikeda and Sato (1980) propose a synthetic aperture sonar system where there is a fixed transmit/receive transducer and a movable hydrophone. The signals received at the fixed transducer are then used as reference signals to reduce the effects of turbulence. This is an application of the technique whereby a reference source is placed close to the object, so that the reference and object signals are subjected to the same turbulence (Goodman et al., 1966).

2.8.1 Motion Compensation

Motional compensation for radars is usually performed using inertial navigation techniques; the position and velocity of the airborne radar platform being derived from integration of the acceleration measurements. For a discussion of processing techniques employed in inertial motion compensation refer to Mims and Farrell, 1972; Kirk, Jr, 1975b; van de Lindt, 1977.

The motion compensation requirements are more severe for sonar than they are for radar (Cutrona, 1975, p348). This is because the time taken to traverse the synthetic aperture is much longer due to the slow tow speeds. Rather than using inertial navigation systems, the position of a sonar towfish (with respect to the medium) may, in principle, be derived from velocity measurements using electrostatic current meters, or alternatively from displacement measurements using propeller driven current meters (de Heering, 1982, §3.6.5).

Porter et al. (1973) suggest that accurate knowledge of the sonar-track can be derived by towing the sonar through a field of implanted transponders, and triangulating the returns (cf. Mitome et al., 1985). However, this is not a very flexible system (Williams, 1976, p300), and to be of any use needs to provide sub-wavelength accuracy.
2.8.2 Passive Autofocusing

In SAR systems, platform navigation can be supplemented by alignment based on the received signal itself. This process is known as autofocusing (Kirk, Jr, 1975a; Kirk, Jr, 1975b). Echoes from a target highlight (natural or artificial — e.g. a beacon) are aligned and then the range rate of other features relative to the highlight is calculated. From this it is possible to deduce the deviation from a straight path (Ellis, 1973, p156; Vant, 1982; Ellis, 1984, p176).

Although primarily suited to spotlight mode SAR (Bladmell and White, 1987), similar techniques may be useful in synthetic aperture sonar imaging. For example, consider a sonar imaging to both port and starboard. Any deviation from a straight track causes the echoes from one side to be shifted up in Doppler while echoes from the other side are shifted down in Doppler. Providing that there are suitably strong reflectors on both sides of the sonar, it may be possible to determine the rate of displacement from the target Doppler histories. Similarly, correlating echoes forward and aft of the sonar may be useful for determining towfish yaw.

2.8.3 Effects of Motional Errors

A qualitative list of the effects of the various sources of phase error is given in Table 2.2. The line of sight range errors are due to target motion (Raney, 1971) and towfish movement in height and cross-track. In high resolution systems these range errors cause the target to translate through a number of resolution cells, a phenomenon known as range walk.

While linear phase errors simply produce an image shift, quadratic phase errors produce a smearing of the image, corresponding to a broadening of the synthetic beam pattern and an increase in the sidelobe level. In particular, it is the along-track component of the velocity error that is primarily responsible for image defocusing (Zeoli, 1986). In addition, there are errors due to towfish roll, yaw, and pitch. These produce lever arm errors and signal modulation as the beam moves past the target (Kirk, Jr, 1975b).

Assuming that the sonar platform is inherently stable, the predominant cause of image defocus is an incorrect assessment of the platform speed (tow speed). Tomiyasu (1978, p575) suggests that defocusing becomes significant when the quadratic phase history deviates by more than $\lambda/4$. Differentiating the instantaneous target range, $R(t)$, with respect to the tow speed $v$, gives

$$\frac{dR(t)}{dv} = \frac{v(t - t_0)^2}{R(t)}$$ (2.43)

The maximum range error $\Delta R_m$ occurs at the extremities of the synthetic aperture, i.e. $t - t_0 = L_s/2v$. Therefore,

$$\Delta R_m = \left(\frac{\lambda L_s}{2vR_m}\right)^2$$ (2.44)

where $R_m$ is the maximum target range at the end of the aperture. If the maximum range error that can be tolerated at each end of the aperture is $\Delta R = \lambda/16$ (equivalent
Table 2.2. The effects of phase errors (compiled from Tomiyasu, 1978, p574). The crosses denote which effects are applicable for a particular source of error. Note that the range errors result from either a towfish movement in cross-track (x) or height (z), or as a result of a moving target. The phase jitter errors result from either a medium instability or a timing error in the sonar electronics.

to the far-field criterion), the maximum allowable tow speed error is thus

\[ \Delta v \approx \frac{v(\Delta y)^2}{R_0 \lambda} \]  

(2.45)

where \( \Delta y = R_0 \lambda/(2L_s) \) is the along-track resolution.

Note that \( \Delta v \) becomes more critical with increasing along-track resolution. For example, let \( \lambda = 0.1 \text{ m}, R_0 = 100 \text{ m}, v = 1.0 \text{ ms}^{-1} \), and \( \Delta y = 1 \text{ m} \). From (2.45) the maximum tolerable tow speed error is \( \Delta v = 0.1 \text{ ms}^{-1} \), a 10% error. A desired along-track resolution of \( \Delta y = 0.2 \text{ m} \), however, requires the tow speed to be known to within 0.4%.

Surprisingly, a tow speed error has less effect at higher operating frequencies. Although the wavelength is shorter, and less tolerant of path length errors, increasing the frequency reduces the beamwidth (assuming a fixed transducer size), and therefore it would seem that the tolerance to a tow speed error would be independent of frequency. However, the phase error is proportional to the square of the integration time, and thus also to the square of the beamwidth. Therefore, the overall result of increasing the frequency is a higher tolerance to tow speed errors.

2.8.4 Range Curvature

During the generation of long apertures, the change in slant range to a point target results in the echo signal lying on a curved locus in measurement space. When the range bin width is smaller than the change in slant range, the echo signal exists in different range bins for different parts of the aperture. This phenomenon is known as range cell migration (Sack et al., 1985, p46), and correction is required to prevent degradation of image quality, especially in widebeam, high resolution systems.

When the sonar beam is skewed in azimuth, so that the beam no longer points perpendicular to the sonar track, the range cell migration consists of a linear component called range walk in addition to the quadratic component known as range...
To show this effect, the target slant range can be expressed, with reference to Figure 2.7, as

$$R(y_s) \approx R_0 + \frac{(R_0 \tan \theta_S)^2}{2R_0} + \frac{(y_t - y_s)^2}{2R_0} + \frac{2(y_t - y_s)R_0 \tan \theta_S}{2R_0}$$ \hspace{1cm} (2.46)

where \( \theta_S \) is the beam skew angle from broadside, \( R_0 \tan \theta_S \) is the offset of the beam centre from the target position \( y_t \), and \( R(y_s) \) is the target slant range at the sonar track position \( y_s \) with respect to the position \( y_c = y_t + R_0 \tan \theta_S \) where the target passes through the beam centre. The first two terms of (2.46) represent the slant range when the target passes through the beam centre, the third term is the range curvature, and the fourth term is the range walk. Note that when there is no beam skew (i.e. \( \theta_S = 0 \)) the range walk component is zero but the range curvature component is still present.

Range walk causes the echo signal to lie along diagonal lines in measurement space, and can be corrected by a simple skewing operation prior to the azimuth compression. To re-align the image, an inverse deskewing operation is performed after the azimuth compression (Sack et al., 1985, p46).

Correction of range curvature is much more involved, and usually requires each image pixel to be individually reconstructed, by tracking the individual locii through measurement space. Alternatively, the echo signal can be correlated with the two-dimensional (range and along-track) impulse response of a point target (Truong et al., 1984). To reduce the level of computation required for range curvature correction (especially with SAR images), methods are employed to use the same correction for groups of targets (van de Lindt, 1977, p423).

Specifically, with \( R_0 \) being the closest slant range to a point target, and \( L \) the aperture length, the range curvature, \( \Delta R_c \), is

$$\Delta R_c = R_0 \left( 1 + \frac{L^2}{4(R_0)^2} \right)^{\frac{1}{2}} - R_0 \approx \frac{L^2}{8R_0}$$ \hspace{1cm} (2.47)

where the binomial approximation is valid for \( R_0 \gg L \). The range curvature can also be expressed in terms of the desired along-track resolution by making the substitution...
\( L = \frac{\lambda R}{2\Delta y} \) in (2.47). This can then be related to the range resolution \( \Delta r \), to gain a measure of the importance of range curvature.

\[
\frac{\Delta R_c}{\Delta r} \approx \frac{\lambda^2 R_0}{32(\Delta y)^2 \Delta r}
\]

(2.48)

This shows that range curvature is more of a problem at longer wavelengths and for high resolution imagery. Tomiyasu (1978, p573) suggests that for ratios of \( \Delta R_c / \Delta r \) in excess of 0.3, compensation of the range curvature is required to prevent image degradation. For example, let \( \lambda = 0.1 \text{ m} \), \( R_0 = 100 \text{ m} \), and \( \Delta r = \Delta y = 1 \text{ m} \), then \( \Delta R_c / \Delta r = 0.03 \) which is a negligible amount, i.e. the target locus can be considered to be a straight line. However if a higher along-track resolution, \( \Delta y = 0.2 \text{ m} \), is desired, then \( \Delta R_c / \Delta r = 0.78 \), i.e. a significant range curvature.

2.9 Summary

The motivation behind a synthetic aperture sonar is that a high along-track resolution can theoretically be achieved by radiating a wide beam and using signal processing techniques to compress the beam. There are a number of advantages over traditional narrow beam imaging sonars; the important ones being:

(i) The along-track resolution is independent of range and frequency (provided that the required beamwidths are realisable and that coherence can be maintained over the length of the desired synthetic aperture).

(ii) An aperture many boat-lengths long may be formed without the hardware complexity of a towed array.

(iii) Widebeam illumination allows many hits on target, and from a number of different views. This may help to improve the detection of certain targets.

The major disadvantages include the large computation requirements for real-time operation, and motion compensation required to correct for sonar deviation from the desired track. For a more complete comparison of synthetic and nonsynthetic aperture sonars refer to Cutrona (1975) and (1977).

There are many conflicting factors that influence the choice of the optimum operating parameters for a synthetic aperture sonar. Oliver (1982) shows that the best performance is obtained if the real aperture is increased as far as possible, consistent with the desired along-track resolution. Reducing the length of the real aperture increases the signal processing and data storage requirements. In addition, the doppler bandwidth is also increased which may require the prf to be increased to ensure adequate sampling, thus reducing the unambiguous range. Finally, the use of a wider real beam spreads the transmitted energy over a wider area, but this is only partially compensated by the increased integration of the received signal (Ellis, 1984).

The choice of operating frequency is more difficult. Table 2.3 summarises the frequency dependence of some of the more important factors. It is not immediately obvious, however, what the optimum operating frequency of a synthetic aperture sonar should be. For instance, increasing the operating frequency makes the sonar more susceptible to path length errors, but by the same token, the synthetic apertures are now shorter thereby easing coherence difficulties. Theory suggests, however, that
a synthetic aperture sonar can give a comparable performance to a nonsynthetic aperture sonar but at a lower frequency (Cutrona, 1975).

<table>
<thead>
<tr>
<th>Factor</th>
<th>Expression</th>
<th>Effect of increased operating frequency</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range resolution</td>
<td>$\Delta r \propto c/(2B)$</td>
<td>no effect</td>
<td>§ 1.10.4</td>
</tr>
<tr>
<td>Along-track resolution</td>
<td>$\Delta y \propto L/2$</td>
<td>no effect</td>
<td>§§ 2.2, 2.3</td>
</tr>
<tr>
<td>Fractional bandwidth</td>
<td>$B/f = B\lambda/c$</td>
<td>reduced</td>
<td>§ 2.6</td>
</tr>
<tr>
<td>Doppler bandwidth</td>
<td>$W_d \approx v/L$</td>
<td>no effect</td>
<td>§ 2.1.3</td>
</tr>
<tr>
<td>Azimuth beamwidth</td>
<td>$\theta_B \approx \lambda/L$</td>
<td>narrower</td>
<td>§§ 2.2, 2.3.1</td>
</tr>
<tr>
<td>Integration time</td>
<td>$T_d \approx R_0\lambda/(vL)$</td>
<td>shorter</td>
<td>§ 2.1.3</td>
</tr>
<tr>
<td>Integrated pulses</td>
<td>$N_i \approx R_0\lambda/(vTL)$</td>
<td>fewer</td>
<td>§ 2.3.1</td>
</tr>
<tr>
<td>Range curvature</td>
<td>$\Delta R_c = \lambda^2 R_0/(8L^2)$</td>
<td>less curvature</td>
<td>§ 2.8.4</td>
</tr>
<tr>
<td>Maximum tow speed</td>
<td>$v \leq L/(2T)$</td>
<td>no effect</td>
<td>§ 2.4</td>
</tr>
<tr>
<td>Area mapping rate</td>
<td>$A \leq cL/4$</td>
<td>no effect</td>
<td>§ 2.5</td>
</tr>
<tr>
<td>Tow speed error</td>
<td>$\Delta v \approx vL^2/(4R_0\lambda)$</td>
<td>more tolerant</td>
<td>§ 2.8.3</td>
</tr>
<tr>
<td>Echo amplitude</td>
<td>$\propto 1/\lambda$</td>
<td>increased</td>
<td>§ 2.8.3</td>
</tr>
<tr>
<td>Absorption loss</td>
<td></td>
<td>increased</td>
<td>§ 1.8.1</td>
</tr>
<tr>
<td>Medium stability</td>
<td></td>
<td>less tolerant</td>
<td>§ 5</td>
</tr>
<tr>
<td>Surface multipath</td>
<td></td>
<td>reduced</td>
<td>§ 1.8.2</td>
</tr>
<tr>
<td>Reverberation</td>
<td></td>
<td>reduced</td>
<td>§§ 1.7, A.2</td>
</tr>
<tr>
<td>Coherence lengths</td>
<td></td>
<td>shorter</td>
<td>§§ 1.8.3, 5</td>
</tr>
<tr>
<td>Bottom penetration</td>
<td></td>
<td>reduced</td>
<td>§ 1.10.3</td>
</tr>
</tbody>
</table>

Table 2.3. The effect of increasing the operating frequency on a synthetic aperture sonar having a single beam, transmitted bandwidth $B$, pulse repetition period $T$, transducer length $L$, tow speed $v$, and range $R_0$, for $\rho = 1$.

Finally, to achieve a range resolution comparable with the along-track resolution it is necessary to transmit very short pulses, or preferably to synthesise short pulses using pulse compression techniques. Furthermore, interference effects (such as grating lobes) may be avoided with the use of such broadband signals.
Chapter 3

Continuous Transmission Frequency Modulation

Traditionally most sonars have been based on pulse-echo technology. This is a consequence of the relative ease with which this type of signal can be generated and analysed. Although the relative merits of FM (frequency modulated) sonars have been recognised for over 50 years, the available technology has made the implementation of FM sonars difficult.

FM sonars became popular during the 1940's when the war effort spurred intensive sonar research and development. Kurie (1946, pp1-3) details many advantages of FM sonar systems over 'conventional' systems, mainly resulting from the continuous insonification of targets. The major difficulty with FM sonars was to generate a linear and stable FM sweep.

With the introduction of pulse compression techniques from radar (see Klauder et al., 1960) the performance of pulse-echo sonars was greatly improved. To distinguish between these pulsed FM sonars (now commonly known as chirped sonars) and traditional continuous FM sonars, the latter became known as CWFM (Continuous Wave FM) sonars and more recently as CTFM (Continuous Transmission FM) sonars. The established radar literature, however, uses the term FMCW for equivalent radar systems (Saunders, 1970, §16.10).

Whereas pulse-echo sonar signals are easily processed in the time domain, FM sonar signals are more readily processed in the frequency domain. This, however, requires the use of accurate and linear frequency modulators, and narrow fractional bandwidth spectrum analysers. These difficulties are now being overcome with the advent of digital frequency synthesisers and digital spectral analysis techniques.

There has been a long history of CTFM sonar development at the University of Canterbury, including such diverse applications as blind aids (Kay, 1974), heart monitors (cardiophone) (Kay et al., 1977; Hodgson and Boys, 1977), fish finding sonars (Smith, 1972; Do, 1977), diver's sonars (Cusdin and de Roos, 1984), robotics (Kay, 1985) and underwater classification sonars (de Roos et al., 1989). All these applications have only been interested in the magnitude, or target strength, of the reflected echoes, and the phase has been ignored. In many cases the only output was an auditory display, and consequently the phase was of no interest anyway.

This chapter looks at the properties of the CTFM signal and how it can be
generated, demodulated and analysed in sonar applications. Particular attention is
directed to the preservation and analysis of the demodulated signal phase, and the
effects of Doppler.

3.1 The CTFM Signal

A CTFM signal can be considered to be a coherent train of FM pulses where the
duration $T_p$ of each pulse is equal to the pulse repetition period $T$. This period $T$
is generally known as the sweep period. Note that unlike chirped sonars, the duty
cycle of CTFM sonars approaches unity.

![Figure 3.1. Repetitive linear FM sweep, where $f_0$ is the initial sweep frequency, $f_1$ is the terminal
sweep frequency, $T$ is the sweep period, and $\mu$ is the sweep rate. Also illustrated is a single echo
sweep delayed by $\tau$.]

The most common CTFM signal is based on a linear FM pulse of the form

$$s_p(t) = \text{rect} \left( \left( \frac{t - \frac{1}{2}T}{T} \right) \right) \exp \left[ j2\pi(f_0t + \frac{1}{2}\mu t^2) \right]$$  \hspace{1cm} (3.1)

where $f_0$ is the initial frequency, $\mu$ is the sweep rate, and $T$ is the sweep period. The
sweep rate is proportional to the slope of the instantaneous signal frequency (see
Figure 3.1) and in terms of the terminal and initial sweep frequencies is

$$\mu = \frac{f_1 - f_0}{T} \hspace{1cm} (3.2)$$

Note that for a decreasing frequency sweep (down-sweep), the sweep rate is defined
to be negative. For example, $\mu = -f_0/2T$ for a one octave, down-sweep.

The linear FM pulse defined by (3.1) is repeated every $T$ seconds to form a
continuous signal. This can be expressed in terms of a convolution of a single pulse
with a replicating function,

$$s(t) = \sum s_p(t - nT) = s_p(t) \odot \sum \delta(t - nT)$$  \hspace{1cm} (3.3)
remembering that the actual transmitted signal is given by the real-part of (3.3).

The spectrum of the transmitted signal has a line structure spaced by $1/T$ and a complex envelope given by the Fourier transform of (3.1), i.e.

$$S_p(f) = \int_0^T \exp \left[j2\pi(f_0t + \frac{1}{2}\mu t^2)\right] \exp(-j2\pi ft) \, dt$$

(3.4)

Equation (3.4) can be solved by completing the square in the exponent and introducing a new variable $\xi^2 = 2|\mu|(t + (f_0 - f)/\mu)^2$, to give

$$S_p(f) = \frac{1}{\sqrt{2|\mu|}} \exp \left(-j\frac{\pi}{|\mu|}(f - f_0)^2\right) \left\{ \int_{\xi_1}^{\xi_2} \cos \left(\pi\xi^2\right) \, d\xi \pm j \int_{\xi_1}^{\xi_2} \sin \left(\pi\xi^2\right) \, d\xi \right\}$$

$$= \frac{1}{\sqrt{2|\mu|}} \exp \left(-j\frac{\pi}{|\mu|}(f - f_0)^2\right) \left\{ (C(\xi_2) - C(\xi_1)) \pm j (S(\xi_2) - S(\xi_1)) \right\}$$

(3.5)

where the positive sign applies for $\mu > 0$ and the negative for $\mu < 0$. The limits of the integrals are $\xi_1 = \sqrt{2|\mu|(f_0 - f)/\mu}$, $\xi_2 = \sqrt{2|\mu|(T + (f_0 - f)/\mu)}$, and $C(\xi)$, $S(\xi)$ are the Fresnel cosine and sine integrals (see Abramowitz and Stegun, 1964, p300). Cook and Bernfeld (1967, p139) show that as the signal time-bandwidth product is increased, the bracketed sum of the Fresnel integrals in (3.5) approximates a rectangle, and thus the transmitted bandwidth $B$ tends to $|\mu|T$, the difference between the minimum and maximum transmitted frequencies. This can be seen from Figure 3.2 which shows the amplitude spectrum described by (3.5). Notice that there is little spectral 'spillover' (cf. Section 1.1.4) and thus the exponential approximation to the analytic signal is valid for large time-bandwidth linear FM signals.

![Figure 3.2](image-url)

**Figure 3.2.** Amplitude spectra of a linear FM pulse, with an initial frequency $f_0 = 30$ kHz, a terminal frequency $f_1 = 15$ kHz and sweep periods (a) $T = 0.8$ ms, (b) $T = 0.8$ s.
3.1.1 The CTFM Echo Signal

To simplify the following analysis of the received echo signal, it is assumed that there is a single stationary point target in a homogeneous medium. The effect of a moving target on a CTFM signal is discussed in Section 3.2.4. Any frequency dependent amplitude and phase fluctuations due to target shape and medium movements are also ignored. The echo signal is thus a scaled replica of the transmitted signal, delayed by $\tau = 2R/c$,

$$e(t, \tau) = A e^{j\phi}$$

(3.6)

where the complex scale factor $A = A e^{j\phi}$ describes the target reflection properties and the propagation losses due to spreading and absorption. Note that the notation $e(t, \tau)$ describes the echo delayed by $\tau$. Compare this to the general echo signal $e(t)$ that constitutes the sum of all the echoes.

From substitution of the transmitted sweep signal (3.3) into (3.6), the CTFM echo signal is

$$e(t, \tau) = e_p(t, \tau) \otimes \sum \delta(t - nT)$$

(3.7)

where each echo sweep is

$$e_p(t, \tau) = A \text{rect} \left( \frac{t - \left( \tau + \frac{1}{2}T \right)}{T} \right) \exp \left[ j2\pi \left( f_0(t - \tau) + \frac{1}{2}\mu(t - \tau)^2 \right) + j\phi \right]$$

(3.8)

Notice that the echo signal has a frequency discontinuity at $t = nT + \tau$, whereas the transmitted signal has a frequency discontinuity at $t = nT$. These frequency discontinuities often generate transients that can mask weaker signals. Usually the largest transient is caused by the transmitted signal flyback, and hence it is desirable to momentarily blank the sonar output just after the start of each transmitted sweep. The flybacks in the echo signal are range dependent, however, and therefore cannot be blanked as easily. Moreover, the transients generated by the echo signal flybacks are usually too small to be of consequence, excluding of course those echoes from large close-range targets.

In applications where this frequency discontinuity cannot be tolerated, linear V-FM (or triangular FM) may be more suitable. With linear V-FM the signal frequency sweeps up for half the sweep period and then sweeps down for the other half. An important property of this modulation is that it is possible to obtain uncorrelated measurements of range and range rate (Cook and Bernfeld, 1967, p98). However, V-FM is difficult to unambiguously demodulate, except for delays much shorter than the sweep period.

3.1.2 Ambiguity Function of the CTFM Waveform

The normalised complex envelope of a single CTFM sweep can be written as

$$\tilde{s}_p(t) = \frac{1}{\sqrt{T}} \text{rect} \left( \frac{t}{T} \right) \exp \left( j\pi \mu t^2 \right)$$

(3.9)

and substituting this into (1.45), the real envelope of the ambiguity function is,

$$|x_p(\tau, \nu)| = \left( 1 - \left| \frac{\tau}{T} \right| \right) |\text{sinc} \left[ BT \left( 1 - \left| \frac{\tau}{T} \right| \right) \left( \frac{\tau}{T} + \frac{\nu}{B} \right) \right]|, \quad |\tau| \leq T$$

(3.10)
where \( \nu = (\eta_0 - 1)f_c \approx -2f_c R/c \) is the Doppler mismatch in the carrier frequency \( f_c \), and \( \tau \) is the mismatch in delay (cf. Glisson et al., 1970, p43).

Equation (3.10) is the narrowband, also known as Woodward’s, ambiguity function and a graph of the contour \( |x_p(\tau, \nu)|^2 = 0.5 \) is shown in Figure 3.3. However as Mitchell and Rihaczek (1968, p417) show, the narrowband ambiguity diagram is only valid for signals with relatively modest sophistication, since it approximates the Doppler effect as a simple frequency translation rather than a time compression (see Sections 1.4.4 and 1.5).

Figure 3.3. Ambiguity diagram of a narrowband linear FM rectangular pulse of bandwidth \( B \), duration \( T \), and carrier frequency \( f_c \), for \( |x_p(\tau, \nu)|^2 = 0.5 \). The equation of the ridge of this contour is, from (3.10), \( \mu \tau + \nu = 0 \). Hence for a down-sweep (i.e. \( \mu < 0 \)), the contour has a positive slope.

Russo and Bartberger (1965, p188) show that the inclusion of the Doppler time compression of the frequency sweep rate \( \mu \) results in a drastic shortening of the ambiguity contour. The contours are also rotated slightly and are not symmetric about the origin. When the criterion \( B/f_c < 1/BT \) is fulfilled — i.e. when the fractional bandwidth is less than the inverse time-bandwidth product, or in terms of the sweep rate, when \( \mu < \sqrt{f_c/T^2} \) — Russo and Bartberger (1965, p190) show that the radar assumption yields a valid approximation to the correct ambiguity diagram. When the criterion is not fulfilled, the ambiguity contour is shortened and the effect of time compression on the frequency sweep rate must be considered. Mitchell and Rihaczek (1968, p420) show that the contour is shortened by a factor \((B/f_c)TB\) to where its extent in delay and Doppler equals its width along either axis times the inverse fractional bandwidth \( f_c/B \).

Kramer (1967) and Harris and Kramer (1968) asymptotically evaluated the ambiguity function for the effects of Doppler distortion on large time-bandwidth product linear FM signals. Using exact Doppler compensation (time compression), they show that the Doppler tolerance of a correlation receiver is \( \pm 0.3 \) ms\(^{-1}\) for a sonar with a fractional bandwidth \( B/f_c = 0.2 \), and a time-bandwidth product of \( 5 \times 10^3 \). As a comparison, the narrowband approximation predicts a Doppler tolerance of \( \pm 75 \) ms\(^{-1}\)! This is because the narrowband theory predicts the tolerance loss entirely on the basis of temporal overlap loss, and ignores the mismatch in sweep rate (Kramer, 1967, p630). Furthermore, the acceleration tolerance of wideband linear FM sonars is far better than predicted by narrowband theory (Kramer, 1967, p634).

More recently, Lin (1988) has evaluated the exact wideband ambiguity function.
for linear FM signals, and Hermand and Roderick (1988) have studied the effect of multipath (or target highlight structure) on these signals.

The ambiguity function for a train of linear FM sweeps can be found from superposition of versions of $\chi_{\rho}(\tau, \nu)$ displaced by $T$ along the $\tau$-axis. The effect is to produce subsidiary ridges spaced by $T$, with the inner ridges broken into a number of peaks, spaced in the $\nu$ direction by $1/T$ (Rihaczek, 1969, Fig.6.14). This reduces the ambiguity in Doppler, as expected with a long duration signal.

Most pulse trains are designed to have a duty factor less than 0.5, otherwise the resultant ambiguity diagram has areas of overlap (Bird, 1974, §5.2.3). CTFM signals approach unity duty cycles and therefore a large overlap could be expected. However as Rihaczek (1969, p290) shows, the overlap is insignificant if each pulse has a large time-bandwidth product. In addition, if only a small spread in target velocities is expected, ambiguities from the overlapping areas are further reduced. For example, when imaging the sea-floor with a side-scan sonar, the spread of expected target velocities is limited by the extent of the beamwidth in azimuth.

3.2 Demodulation of CTFM Echo Signals

It is obvious from Figure 3.1 that there is a constant frequency difference between the transmitted and received signals, proportional to the delay between them. It is this property of CTFM signals that is often exploited to determine target range. (Similar concepts can be found in the azimuth compression of synthetic apertures, see Section 2.3.2.)

3.2.1 Simple CTFM Demodulation

The simplest method of CTFM demodulation — sometimes referred to as heterodyne correlation (Smith, 1972, §1.3) or as the spectrum analyser technique (Cook and Bernfeld, 1967, p164) — relies on the fact that there is a substantial overlap between the transmitted and received signals. The received signal is multiplied directly with the complex conjugate of the transmitted signal (a sweep in the opposite direction) to obtain the desired difference frequency signal, i.e. $d(t) = e(t)s^*(t)$. However, it is difficult in practice to generate the complex echo signal $e(t)$ from the real echo $e(t)$, since this requires the use of wideband Hilbert transformers (Manes et al., 1983). So instead, the real echo signal is multiplied with the real transmitted signal, as shown in Figure 3.4. This produces the desired difference frequency component and in addition a sum frequency component around $2f_0$ which is removed by low pass filtering.

This simple demodulation operation may be expressed concisely as

$$d(t) = s(t)e(t) \odot i(t)$$  \hspace{1cm} (3.11)

where $i(t)$ is the impulse response of the low pass filter. Therefore from (3.3) and

\footnote{In radar, there is a similar pulse compression technique called stretch, where the reference signal is usually a delayed and frequency offset version of the transmitted signal to ensure a large overlap (Caputi, 1971).}
the resultant demodulated signal can be expressed as

\[ d_p(t, \tau) = A_1 \text{rect} \left( \frac{t - \frac{1}{2}(T + \tau)}{T - \tau} \right) \cos \left[ -2\pi \left( f_0 + \mu(t - \frac{1}{2}\tau) \right) + \phi_1 \right] \]

\[ + \quad A_1 \text{rect} \left( \frac{t - \frac{1}{2}(2T + \tau)}{\tau} \right) \cos \left[ -2\pi \left( f_0 + \frac{1}{2}\mu(t - (T + \tau)) \right) (T - \tau) + \phi_1 \right] \]

where \( A_1 \) and \( \phi_1 \) are arbitrary amplitude and phase constants. In practice, both \( A_1 \) and \( \phi \) vary with the demodulated frequency and the low-pass filter characteristics. This aspect is considered further in Section 3.5.3.

From inspection of (3.12), the instantaneous demodulated frequency is

\[ f_a(t, \tau) = \begin{cases} 
| -\mu\tau | & \text{for } \tau \leq t \leq T \\
|\mu(T - \tau)| & \text{for } T \leq t \leq T + \tau
\end{cases} \]

and in the interval \( \tau \leq t \leq T \), is directly proportional to the target range as expected. So for a measured frequency difference \( f_a \), the indicated range \( R_a \) is given by

\[ R_a = \frac{cf_a}{2|\mu|} \]

However, for the period \( 0 \leq t_s \leq \tau \) after the start of each transmitted sweep, the indicated range is incorrect. This occurs when the received echo from the previous sweep is demodulated by the current transmitted sweep, and is described by the second term of (3.13). Since this simple demodulation technique cannot differentiate between positive and negative frequency differences, a range ambiguity exists.

To resolve this ambiguity, the CTFM echoes can be demodulated to an IF rather than to baseband, but there still exists an interval where the difference frequency is not proportional to the target range. This unwanted difference frequency is usually ignored by blanking the demodulator output for a period \( T_b \) (called the blind time or lost time) after the start of each transmitted sweep; the longer the blanking the greater the unambiguous range \( cT_b/2 \). However a long blind time is undesirable, because both the signal to noise ratio and range resolution are degraded by a factor

![Figure 3.4. Simple CTFM demodulation system.](image)
\( k_b = T_b/T \). Therefore, as a compromise between blind time and maximum unambiguous range, the blind time is typically limited to 10–30% of the sweep period.

There are a number of other disadvantages to this method of demodulation. With an auditory display, the demodulated signal is heard as an interrupted tone (Do, 1977). Obviously as the blind time is increased, the quality of this tone degrades. This is not so much of a problem for long sweep periods, but at high sweep rates this unwanted modulation may mask the target sounds (Boys et al., 1978, p123). Similarly this discontinuous tone may cause unwanted transients when applied to a spectrum analyser comprised of banks of highly selective band-pass filters (Gough et al., 1984, p271).

The biggest limitation of the simple demodulation technique is that the maximum unambiguous range is only \( cT_b/2 \) (compare this to \( cT/2 \) for a pulsed CW sonar; cf. Section 1.10.6). This can be extended by using a delayed version of the transmitted signal to demodulate the echoes, but at the expense of not being able to unambiguously detect close range targets. Delaying the reference signal shifts the range analysis window but does not improve the range comprehension. For example, delaying the transmitted signal by \( \tau_0 \) shifts the unambiguous range window to

\[
\frac{c}{2} \tau_0 \leq R < \frac{c}{2} (\tau_0 + k_b T)
\]

but the range comprehension (i.e. the range spanned between the minimum and maximum ranges) is still only \( c k_b T/2 \).

![Diagram](image)

**Figure 3.5.** Quadrature CTFM demodulation system.

### 3.2.2 Quadrature Demodulation

Although it is difficult to produce \( e(t) \) from \( e(t) \), it is relatively easy to generate both \( s(t) \) and \( \hat{s}(t) \) using digital synthesis techniques. By multiplying the received echoes
with the in-phase and quadrature phase components of the transmitted signal, as shown in Figure 3.5, it is now possible to differentiate between positive and negative frequency differences — i.e. whether the echoes are higher or lower in frequency than the transmitted signal. Therefore, echoes from the previously transmitted sweep can be distinguished from the echoes of current transmitted sweep. With a down-sweep, for instance, the echoes from the previous sweep are higher in frequency than the current transmitted frequency, and vice-versa for an up-sweep.

\[
f_a(t, \tau) = \begin{cases} 
-\mu \tau, & \tau \leq t - nT < T \\
\mu(T - \tau), & T \leq t - nT < T + \tau 
\end{cases} 
\]  
(3.16)

With the range ambiguity removed, the maximum unambiguous range can now be extended to \( cT/2 \). The demodulated signal still has the undesirable frequency discontinuities at \( t = nT \) and \( t = nT + \tau \) (see Figure 3.6b), but these may be overcome by deliberately aliasing any frequency components outside the range \(-B/2 \leq f_a \leq B/2\) (see Figure 3.6e).

3.2.3 Dual-Demodulation

There have been several methods proposed to overcome the blind time inherent with the simple demodulation technique (Gough et al., 1984; Do, 1986; de Roos et al., 1989). Instead of using a single swept local oscillator, two interlaced local oscillators are required to demodulate the echo signals; one oscillator with an output \( l_1(t) \) for echoes from the current sweep and another oscillator \( l_2(t) \) for echoes from the previous sweep. This is shown in Figure 3.7.

The reference local oscillators have the same sweep rate \( f_t \) as the transmitted signal, but different initial frequencies \( f_{11} \) and \( f_{12} \), where

\[
l_1(t) = \cos \left[ 2\pi \left( f_{11} t + \frac{1}{2} \mu t^2 \right) \right], \quad 0 \leq t < T \\
l_2(t) = \cos \left[ 2\pi \left( f_{12} t + \frac{1}{2} \mu t^2 \right) + \phi_{12} \right], \quad 0 \leq t < T 
\]  
(3.17)

To avoid phase discontinuities in the demodulated signal at the local oscillator flyback, the end of reference sweep \( l_1(t) \) must be frequency and phase continuous with the start of reference sweep \( l_2(t) \). Applying these boundary conditions to (3.17) gives

\[
f_{12} = f_{11} + \mu T \quad (3.18) \\
\phi_{12} = 2\pi \left( f_{11} T + \frac{1}{2} \mu T^2 \right) \quad (3.19)
\]

Note that for a down-sweep, \( l_1(t) \) describes the 'upper' swept local oscillator and \( l_2(t) \) describes the 'lower' swept local oscillator.

To minimise breakthrough it is desirable that \( f_{11} \geq 2(f_0 + |\mu|T) \) for a down-sweep, or \( f_{11} \geq 2f_0 \) for an up-sweep. This can be seen from Figure 3.6 which shows the frequency relationships for a down-sweeping system (cf. de Roos et al., 1989). The desired demodulated echoes are those in the frequency band from \( f_{11} - f_0 \) (zero range) to \( f_{11} - f_0 + \mu T \) (maximum range). These are filtered from the unwanted components with an IF filter and are then brought down to baseband by a third
Figure 3.6. CTFM demodulation time-frequency relationships: (a) baseband echoes, (b) quadrature demodulated echoes (c) simple demodulated echoes, (d) dual demodulated echoes, (e) frequency-wrapped quadrature demodulated echoes. Note that unlike in the simple demodulation case, repeated versions of (b), (d), and (e) corresponding to successive transmitted sweeps butt together with no ambiguity.
A dual CTFM demodulation system consists of a projector, hydrophone, power amp, master swept oscillator, upper swept local osc., pre-amp, fixed local oscillator, upper swept local osc., IF filter, baseband filter, and lower swept local osc.

Figure 3.7. Dual CTFM demodulation system.

demodulator with a fixed local oscillator at $f_{l1} - f_0$. This IF filter is often configured to perform range equalisation (see Section 3.5.2), but with the disadvantage of having the wrong frequency characteristic for a sweep in the other direction.

An alternative method of demodulation is to have the local oscillators sweeping in the opposite direction to the transmitted signal. The sum frequencies, rather than the difference frequencies, are now the desired signal. This approach reduces the maximum frequency of both the swept local oscillators, simplifying the requirements of the frequency synthesisers.

The baseband demodulated signal obtained using the dual-demodulation technique is

$$d_p(t, T) = A_1 \text{rect} \left( \frac{t - \frac{1}{2}(T + 2\tau)}{T} \right) \cos \left[ -2\pi \left( f_0 + \mu(t - \frac{1}{2}\tau) \right) \tau + \phi_1 \right] \quad (3.20)$$

and rewriting the instantaneous phase of this signal,

$$\Phi_a(t, \tau) = -2\pi \mu \tau t - 2\pi(f_0 \tau - \frac{1}{2} \mu \tau^2) + \phi_1, \quad \tau \leq t - nT < T + \tau \quad (3.21)$$

it is apparent that the demodulated signal has a constant frequency $f_a = -\mu \tau$. However, there is still a phase discontinuity at $t = nT + \tau$, even though the transmitted signal and local oscillators may be phase continuous at the start of each sweep. This can only be avoided by ensuring that the demodulated frequency $f_a$ is harmonically related to the sweep period $T$.

To interpret the instantaneous demodulated phase $\Phi_a(t, \tau)$ it is sometimes more intuitive to rewrite (3.21) in terms of the round trip propagation delay $\tau$ and an
equivalent frequency $f_e(t, \tau)$:

$$\Phi_e(t, \tau) = -2\pi f_e(t, \tau) \tau$$  \hspace{1cm} (3.22)

Ignoring the time dependence, (3.22) has a similar appearance to the phase expression derived for a monochromatic CW sonar of frequency $f_e$. Comparing (3.22) with (3.21), this equivalent frequency can be seen to be

$$f_e(t, \tau) = f_0 + \mu(t - nT - \frac{1}{2} \tau), \quad \tau \leq t - nT < T + \tau$$  \hspace{1cm} (3.23)

Note $f_e$ is not the same as the instantaneous received frequency $f_R = f_T(t - \tau)$, but is the mean of the instantaneous received frequency and the extrapolated transmitted frequency (the frequency that would have been transmitted if the sweep was not reset, but allowed to continue), i.e.

$$f_e(t, \tau) = \frac{f_T(t) + f_T(t - \tau)}{2}$$  \hspace{1cm} (3.24)

It is also interesting to compare CTFM demodulation with conventional correlation techniques, mentioned in Section 1.3. With a correlation receiver, the received signal is multiplied by a delayed number of replicas of the transmitted waveform and the output integrated (low pass filtered). The output response is then time gated to separate targets in range. A CTFM demodulator is similar to a correlation receiver, but instead of compressing the signal duration by a factor $T_B$, it compresses the signal bandwidth by the same factor. The output of a CTFM demodulator is a number of narrowband frequency components and these are then range gated in the frequency domain using spectral analysis techniques.

Both techniques give the same signal to noise improvement since they both compress the received signal into the smallest possible resolution cell. Smith (1972, §1.3) shows that hardware requirements for the demodulation technique are, in general, much less than what is required for a correlation receiver. However, the demodulation technique requires substantial modifications to cope with rapidly moving targets. This problem is addressed in Section 3.2.4.

### 3.2.4 The Effects of Target Motion on CTFM Signal Processing

In radar applications, where CTFM signals typically have narrow fractional bandwidths, the Doppler effect is usually considered to be a simple frequency shift applied equally to all the frequencies within the bandwidth (Russo and Bartberger, 1965, p184). The compression and amplitude scaling of the signal are usually ignored because the speed of even the most rapidly moving radar targets is small compared with the speed of electromagnetic propagation. However, large propagation losses, ambient noise and reverberation dictate that sonar signals be broadband and of sufficiently long duration (Kramer, 1967, p627). As a consequence, even modest varying range rates over the signal duration can produce large receiver decorrelations that severely limit system performance.

To study the effect of Doppler on a broadband CTFM signal, consider a single transmitted sweep, described by (3.1), that is reflected by a point target moving
away from the sonar with a constant range rate $\dot{R}$. If $R_0$ is the initial range, the instantaneous propagation delay is thus

$$\tau(t) = \frac{2(R_0 + \dot{R}t)}{c + \dot{R}} = t_0\eta_0 + (1 - \eta_0)t \quad (3.25)$$

where $t_0$ is the time that the echo is first received and $\eta_0$ is the Doppler time scale factor (see Section 1.4.3). Combining (3.25) and (3.8) the received echo sweep for the case of a moving target is now

$$e_p(t) = A\sqrt{\eta_0} \text{rect}\left(\frac{t - (t_0 + \frac{1}{2}T')}{T'}\right) \exp\left[j2\pi \left(f'_0(t - t_0) + \frac{1}{2}\mu'(t - t_0)^2\right)\right] \quad (3.26)$$

where $f'_0$ is the Doppler shifted initial sweep frequency, $\mu'$ is the Doppler modified sweep rate, and $T'$ is the echo sweep period. These modified echo parameters are related to the transmitted sweep parameters as follows:

$$t_0 = \frac{2R(t_0/2)}{c} = \frac{2R_0}{c - \dot{R}} \quad (3.27)$$

$$\eta_0 = \frac{c - \dot{R}}{c + \dot{R}} \approx 1 - 2\frac{\dot{R}}{c} \quad (3.28)$$

$$f'_0 = \eta_0 f_0 \quad (3.29)$$

$$T' = \frac{T}{\eta_0} \quad (3.30)$$

$$\mu' = \eta_0^2 \mu \quad (3.31)$$

From (3.26) it is apparent that the echo signal has the same form as the transmitted signal, but is delayed by an amount $t_0$, and as a result of the time scaling is compressed (or expanded) by a factor $1/\eta_0$. Thus all the frequency components in the transmitted spectrum are scaled in frequency by $\eta_0$ and in amplitude by $\sqrt{\eta_0}$.

As illustrated in Figure 3.8, the Doppler time compression produces a change in slope of the frequency-time history (frequency sweep rate). If this difference is not accounted for, then the demodulated signal contains a residual linear FM sweep due to the mismatch in sweep rate. Therefore dual-demodulation (and other heterodyne demodulation systems) are only appropriate for slowly moving targets. When there is significant Doppler, the echo signal should instead be matched to a set of Doppler shifted references.

To see the effect of this mismatch in sweep rate, let the echo signal described by (3.26) be demodulated using the dual-demodulation or similar technique. From (3.20), the instantaneous frequency of the demodulated signal is

$$f_a(t) = -\mu \tau(t) + (\eta_0 - 1) \left(f_0 + \mu(t - \tau(t))\right) \quad (3.32)$$

where the first term is the desired frequency difference (proportional to the instantaneous propagation delay), and added to this is an extra term due to the Doppler frequency shift. As expected, this additional Doppler term is proportional to the instantaneous transmitted signal frequency at the instant of reflection. From (3.14), this demodulated difference frequency suggests the target is at a range

$$R_a(t) = R(t - \frac{1}{2}\tau(t)) - \frac{c}{2\mu}(\eta_0 - 1) \left(f_0 + \mu(t - \tau(t))\right) \quad (3.33)$$
where \( R(t - r(t)/2) \) is the true target range at the instant the echo was reflected, and the second term is a range error due to the Doppler effect. This range error is a direct result of the coupling between range and range rate measurements when using a linear FM signal.

If a target is approaching the sonar (i.e. \( \dot{R} < 0 \)) the echo spectrum is scaled up in frequency (up-Doppler) by an amount proportional to the radial difference in velocity between the target and sonar. Therefore when a down-sweep is transmitted, the echo frequencies are higher than the projected frequencies. This has the effect of making the indicated target range appear greater than the actual target range (Kurie, 1946, p35). Similarly if the target is receding, the error due to the down-Doppler shift is toward a shorter range indication. The converse is true for a transmitted up-sweep and the four possibilities are summarised in Table 3.1.

<table>
<thead>
<tr>
<th></th>
<th>Up-Doppler ( (\dot{R} &lt; 0) )</th>
<th>Down-Doppler ( (\dot{R} &gt; 0) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up Sweep ( (\mu &gt; 0) )</td>
<td>(-)</td>
<td>(+)</td>
</tr>
<tr>
<td>Down Sweep ( (\mu &lt; 0) )</td>
<td>(+)</td>
<td>(-)</td>
</tr>
</tbody>
</table>

Table 3.1. The effect of Doppler on range indication for a CTFM Sonar, where a '+' means the indicated range is too large and where a '-' means the indicated range is too small.

Provided the target motion is known, the range accuracy can be improved by removing the range error. Furthermore, if the target is moving at a constant, but unknown range rate, it is still possible to calculate the target range with the same accuracy as if the range rate was known (Rihaczek, 1965b).
If the target range at the instant \( t_r = t - \tau(t)/2 \) of signal reflection is \( R(t_r) \), then the indicated range \( R_a(t) \) is the true range at the instant \( t_r + \Delta t(t) \), where

\[
\Delta t(t) = \frac{c}{(c + \dot{R})\mu} \left[ f_0 + \mu(t - \tau(t)) \right]
\]  

(3.34)

is the time shift for a target moving at a constant range rate \( \dot{R} \).

With narrowband signals and small range rates, \( \Delta t(t) \) averaged over a sweep is approximately \( f_c/\mu \), a constant independent of the target motion, where \( f_c \) is the centre sweep frequency. The indicated range is thus independent of the target range rate, hence it can be easily corrected (Klauder et al., 1960, p804; Rihaczek, 1965b). This property has been used to advantage in weapon firing systems (Kurie, 1946, p36), since the indicated range for an up-sweep predicts the actual target range at a time \( \Delta t(t) \) after reflection.

### 3.3 Spectral Analysis of Demodulated CTFM Signals

For single target range measurement (e.g. altimeters), a simple phase locked loop or zero-crossing counting technique can be used to determine the target range. In these single target applications it is also possible to transmit a sinusoidal FM signal rather than a linear FM signal. With sinusoidal FM the instantaneous beat frequency is not constant, but over a modulation cycle the average beat frequency indicates the correct range (Skolnik, 1980, p84).

When there are multiple targets present in the sonar beam, the demodulated signal contains a number of difference frequencies. Provided the system is linear, there is a frequency component corresponding to each target, and thus the combined demodulated signal is

\[
d(t) = \sum_{n=1}^{N} A_n \cos(2\pi f_n t + \phi_n)
\]  

(3.35)

where \( A_n, f_n, \phi_n \) are the amplitude, demodulated frequency, and phase for each of the \( N \) targets. Therefore, spectral analysis techniques are required with CTFM sonars to separate the demodulated echo components. This is in contrast to pulsed sonars where target echoes are separated in the time-domain.

#### 3.3.1 Range Resolution

The theoretical range resolution of a CTFM sonar can be found by taking the Fourier transform of the demodulated signal given by (3.20). This gives

\[
D_p(f, \tau) = A_1 T \exp[j\phi_a(\tau)] \exp \left[-j2\pi(f - \mu\tau)(\tau + \frac{1}{2}T) \right] \text{sinc} \left((f - \mu\tau)T \right)
\]  

(3.36)

where \( \phi_a(\tau) = 2\pi(f_0\tau - \mu\tau^2/2) + \phi_1 \). The compressed CTFM sweep has the shape of a sinc function, with a 3 dB width of 0.88/T, corresponding to a nominal range resolution of

\[
\Delta r = 0.88 \frac{c}{2 |\mu| T}
\]  

(3.37)

Notice that the range resolution of a CTFM depends on the frequency resolution \( \Delta f \) of the spectrum analyser, with the highest resolution obtained when \( \Delta f \sim 1/T \).
However this range resolution is not always needed, in which case the analysis bandwidth can be covered with fewer bins. Moreover since the filters are now less selective, they respond more rapidly. Thus there is a trade-off between range resolution and speed of response, something no pulsed sonar can match (Gough et al., 1984, p271). In addition, it is now possible to make a number of range measurements within the sweep period, whereas a pulsed sonar is limited to one look per pulse.

3.3.2 Auditory Spectral Analysis

The demodulated outputs of CTFM sonars are particularly suited to auditory analysis, especially in applications where vision is limited. Examples include, navigational aids for divers working at night or in murky waters (de Roos et al., 1983; Cusdin and de Roos, 1984), and mobility aids for visually impaired people (Boys et al., 1979). The ear, being a particularly effective spectrum analyser in the range 200 Hz to 12 kHz, is good at discriminating between targets of different surface texture, a task that is hard for a computer to emulate (de Roos, 1986, Ch.5).

3.3.3 Analogue Spectral Analysis

The spectrum from a multi-target CTFM sonar is usually obtained by using banks of bandpass filters, where the number of filters needed depends on the desired range comprehension and range bin width. Obviously the better the filter selectivity, the greater the number of filters that are required to cover a given range interval. To reduce the hardware complexity, a smaller number of bandpass filters may be swept over the analysis bandwidth. However like any serial technique, as well as degrading the response time, it is possible to miss transient events.

If the phase of each range bin is required, it is necessary to use a quadrature demodulation technique with an additional filter bank for the quadrature signals. The two banks, in-phase and quadrature, are then fed to a bank of phase detectors or are time multiplexed between a single phase detector. The multiplexed approach is the preferred method since the peak responses of the filters are staggered throughout the sweep period anyway.

3.3.4 Digital Spectral Analysis

Most digital spectral analysis is based upon the Discrete Fourier Transform (DFT). The two most common implementations being the Chirp Z-Transform (CZT) (cf. Rabiner et al., 1969) and the Fast Fourier Transform (FFT) (cf. Brigham, 1974). The advantage of the FFT is that it is a numerically efficient algorithm for calculating the DFT, whereas the CZT is more suited to hardware implementation.

The FFT can be considered as a bank of parallel, identical constant bandwidth filters with an amplitude response of $|W(f - k\Delta f)|$ for the kth filter, where $\Delta f$ is the line spacing and $W(f)$ is the Fourier transform of the time weighting (or window) function $w(t)$. The phase response is linear and depends on the choice of the FFT time origin (see Section 4.2.3).

In order to achieve the full resolution capability of the CTFM signal it is necessary to use an FFT that spans an entire sweep period $T$ (see Section 3.3.1). Since a block FFT occupies a rectangular region in time-frequency space, while the demodulated
echoes occupy a parallelogram shaped region (see Figure 3.6), it is evident that no single block FFT can produce valid output data for all ranges (Sack et al., 1985, p47). A correlation receiver, in comparison, does not experience this problem, nor does a modified version of the CZT known as the Sliding Chirp Z-Transform (SCZT) (Broderson et al., 1976). Therefore, the analogy of a FFT to a bank of filters is not entirely correct. Each filter in the bank is independent, whereas the FFT is a block operation, thus the outputs are forced to respond at the same time.

To overcome this problem, FFTs may be overlapped. However, this significantly reduces the computational efficiency since most of the data from each FFT must be discarded. Alternatively, the output resolution may be sacrificed by taking shorter length FFTs. Each shorter FFT contains less invalid data points, and therefore, less overlap is required. Moreover, these shorter length FFTs may be reordered and combined to form a number of longer overlapping FFTs. The resolution is now the same as before, but since the FFTs are shorter and less overlap is required, the computation requirements are far smaller. This process is explained further in Section 4.2. Refer also to a similar matched filtering technique called the *step transform* (cf. Perry and Martinson, 1978 and Sack et al., 1985).

Finally, other non-Fourier techniques have been proposed for analysis of CTFM (and chirped) signals (Yamamoto et al., 1984; Liau et al., 1986). Many of these techniques may give a better resolution in certain applications, for example the maximum entropy method. However, most of these non-Fourier techniques only estimate the power spectrum (Kay and Marple, Jr., 1981) and thus the phase information essential for coherent beamforming is lost.

### 3.3.4.1 The Picket Fence Effect

The effect of only measuring a spectrum at discrete frequencies is known as the *picket fence effect* (cf. Bergland, 1969), because it appears as if the continuous distribution has been observed through a picket fence. This occurs whenever spectral analysis with discrete filters is performed, be it the FFT or a bank of parallel or stepped analogue filters (Gade and Herlufsen, 1987, p28).

The picket fence effect introduces an amplitude and frequency error. The amplitude error is limited by the ripple in the filter passband while the frequency error is limited to the line spacing $\Delta f$. The frequency and phase correct are only correct for frequency components that coincide with the centre of a filter. For the FFT this requires that the frequency component is harmonically related to the line spacing $\Delta f$.

The picket fence effect on the FFT may be reduced by interpolating the output samples with sinc functions, or more efficiently by *zero-packing* the input samples (Bates and M'Donnell, 1986, §2.11). Interpolation does not increase the frequency resolution, but it may increase the amount of information that can usefully be extracted from the spectrum.

### 3.3.4.2 Spectral Leakage and Windowing

The discrete and continuous Fourier transforms differ considerably for long duration signals. The difference arises from the truncation of the input signal when the DFT
is performed. The effect of this is that the spectrum is convolved with a sinc($fT_f$) function, where $T_f$ is the analysis period. Although the band-limited input signal may only contain a single frequency component, the effect of the truncation is to introduce additional frequency components. This effect is termed spectral leakage (Brigham, 1974, §6-4,§9-5). It is undesirable because sidelobes of the return from a strong target can sometimes swamp the return from a weaker target, even several range bins away (Carr et al., 1981). In addition, the truncated signal is now no longer bandlimited and therefore aliasing occurs (Bracewell, 1978, p377). This has more of an affect at high frequencies, whereas leakage is uniformly distributed over all frequencies.

To reduce the effect of leakage, it is desirable to window the data with a function that minimises the effect of the truncation (Harris, 1978; Nuttall, 1981). This reduces the response sidelobe level, but at the expense of an increased main lobe width, or smoothing error. An alternative technique is edge extension, where the ends of the truncated segments are smoothly extrapolated to zero (Bates and MCDonnell, 1986, pp70,71).

When a frequency component is harmonically related to the period $T_f$, no leakage is observed with a rectangular window. This is a consequence of the picket fence effect; the sampling points in the frequency domain happen to coincide with the peak of the sinc function and it zeroes. However, the sidelobes of the sinc function can still swamp a smaller nearby signal that is not harmonic. Zero-packing the input data, or sinc-interpolating the output data, reveals this anomaly.

### 3.4 CTFM Signal Generation

#### 3.4.1 Analogue Methods

The CTFM signal is traditionally generated by applying a sawtooth voltage ramp to a Voltage Controlled Oscillator (VCO) (Do, 1984). Typically the voltage ramp is generated by integrating the output of a constant current source, the greater the current, the faster the sweep rate $f$. At the start of each sweep, the integrator is reset to a value corresponding to the initial frequency.

It was noted at an early stage in the development of CTFM sonars that the performance of the sonar was limited by the linearity and stability of the frequency modulated oscillator (Kurie, 1946). Extremely high linearity of the frequency sweep is required, especially for long sweep periods, because any non-linearity causes an error in the demodulated difference frequency. The effect is to make a stationary target appear to wander in range during the sweep period (Do, 1984).

The non-linearity is usually dominated by the VCO; the linearity of the sawtooth generator being relatively easy to achieve. At faster sweep rates the VCO often has difficulty tracking the ramp voltage, resulting in a demodulated frequency drift. To overcome this problem, the ramp voltage was sometimes deliberately distorted to compensate for this behaviour (Cusdin, 1988).

There have been many other corrective techniques applied to improve the linearity of analogue CTFM signal generation (Zehner, 1983). Conventional analogue techniques have rarely reached 0.1% linearity, however, and have now been largely superceded by digital generation techniques.
3.4.2 Digital Methods

The introduction of digital signal generation methods to CTFM sonars allowed high linearity and stability of the frequency modulated oscillator for the first time. Initially quasi-digital techniques were suggested where the VCO was configured to clock through the samples of a sinewave stored in a memory (Zehner, 1983). The waveform was then reconstructed using a Digital to Analogue Converter (DAC) and a low pass filter to reject the unwanted harmonics. The advantage of this method was that quadrature signals could be generated in perfect synchronism; using the same VCO to simultaneously clock through a sampled cosinewave (Zehner and Skinner, 1984).

The advent of larger EPROMs has made the storage of the entire sweep feasible. The signal is reconstructed, simply by using a counter to clock through successive samples which are output to a DAC and filtered as before. Delayed reference sweeps may be stored in a similar fashion and can be generated in perfect synchronism with this method. The number of samples required depends on the time-bandwidth product of the signal and the oversampling factor used. The sampling theorem suggests that at least two samples are needed per cycle, but to simplify filtering of the reconstructed signal, oversampling is required in practice. For a sweep of period $T$, maximum frequency $f_m$, and an oversampling factor $k_s$, $2k_sf_mT$ samples are required to store the entire sweep. This number can be reduced by storing only the baseband of the sweep in which case only $2k_sBT$ samples are needed, where $B$ is the signal bandwidth. Multiplying the baseband signal by a carrier signal, and removing the unwanted sideband by filtering, generates the desired sweep for transmission. However, the filtering operation is difficult and it is often better to generate the single sideband using quadrature techniques (Haykin, 1983, pp141–145).

The disadvantage of storing the entire waveform is that the sweep rate cannot be altered without changing the sweep period or swept bandwidth. For example, if the clock rate is increased by a factor $\alpha$, the entire spectrum is scaled in frequency by $\alpha$, the sweep period $T$ decreases to $T/\alpha$, and hence the sweep rate $\mu$ increases to $\alpha^2\mu$. Alternatively the counter can be short-cycled, but this reduces the transmitted bandwidth as well as the sweep period.

A more flexible method of digitally synthesising a CTFM waveform is to store a single sinewave cycle in memory. By digitally sweeping the clock frequency, a linear FM signal can be produced that has an independently adjustable sweep rate and sweep period.

3.5 Properties of CTFM Sonars

In the following discussion on properties of CTFM systems it is assumed that the echoes are demodulated using dual demodulation or a similar demodulation technique.

3.5.1 Maximum Range

The maximum unambiguous range of a dual-demodulation CTFM sonar with a sweep period $T$ is

$$R_u = \frac{cT}{2}$$

(3.38)
which is identical to the value obtained with a pulse sonar using an equivalent pulse repetition period $T$. However, using the simple demodulation technique, the unambiguous range reduces to $cT_b/2$, where $T_b$ is the blind time.

A CTFM sonar by virtue of its continuous transmission can achieve a comparable source level to that of a pulsed or chirped sonar\(^2\) but using a much lower transmitted power level. Thus the maximum range of a CTFM sonar may be extended further than a comparative pulse sonar by virtue of the greater signal energy available. Gough et al. (1984) claims that for a pulsed sonar operating out to a maximum range of 400 m, a comparable CTFM sonar with the same range resolution can operate out to a maximum range of 700 m. This, however, assumes noise limited operation.

### 3.5.2 Range and Frequency Equalisation

With pulsed sonar systems range equalisation is performed in the time domain using Time Varying Gain (TVG) amplifiers. This is to compensate for the larger spreading losses at longer ranges. With CTFM sonars, the equivalent operation requires a filter with a time varying frequency response, or alternatively, a filter after the demodulator with a fixed frequency response. This second method is the simplest and most common, however it does require that the demodulator can cope with the full dynamic range of the echo signal. Another disadvantage of this method is that the filter response needs to be changed for different sweep rates.

Sometimes it is also necessary to perform frequency compensation of the echo signal prior to demodulation, in particular when using very large bandwidth sweeps. This is to allow for the increased attenuation of the higher frequencies in the medium.

### 3.5.3 Phase Correction

The phase errors in a CTFM sonar can be split into two categories; medium dependent and hardware dependent phase errors. The hardware dependent phase errors result from timing errors, and phase shifts through analogue filters in the demodulation process. The hardware phase error $\phi_h$ can itself be split into two terms

$$\phi_h(t, \tau) = \phi_f(f_R(t, \tau)) + \phi_r(f_d(t, \tau))$$  (3.39)

where the term $\phi_f$ describes the phase shifts prior to demodulation that depend on the received frequency $f_R$, and the term $\phi_r$ describes the range dependent phase shifts that depend on the demodulated frequency $f_d$.

Provided the phase shifts are predictable, and do not vary too greatly with temperature, they can be corrected using a pair of look-up tables, one for the received frequency and another for the demodulated frequency. Where phase accuracy is of great importance, the demodulation process should be implemented using digital signal processing techniques.

The medium dependent phase errors result from acoustic path length variations due to inhomogeneities in the medium. Differentiating (3.21), the demodulated

\(^2\)Provided the transducers can tolerate the increased thermal dissipation when operating at longer duty cycles (Camp, 1970, p191).
phase, with respect to the propagation delay $\tau$ gives

$$\frac{d\Phi_{\text{g}}(t, \tau)}{d\tau} = 2\pi(f_0 + \mu(t - \tau)) = 2\pi f_R(t, \tau)$$  \hspace{1cm} (3.40)

which, as expected, shows that the path length fluctuations produce greater phase errors at the higher frequencies in the received sweep.

There are also phase errors introduced when there is relative motion between the sonar and target. These phase errors are due to decorrelation effects when the Doppler shifted echoes are no longer matched to the reference signal. If the relative target motion is known, however, the phase errors are predictable and can be compensated for. The effect of this mismatch is discussed further in Sections 3.2.4 and 4.2.3.

### 3.5.4 Reverberation

With a continuous transmission system the reverberation arrives from all ranges simultaneously. However, the reverberation intensity from shorter ranges is generally greater than that from longer ranges due to propagation losses. The strong reverberation from very short ranges is a particular problem in continuous transmission systems and appears as acoustical crosstalk. This is discussed in the following section (3.5.5). In a CTFM system, the reverberation from different ranges can be discriminated by spectral analysis of the demodulated reverberation signal. Thus the amount of reverberation appearing in a range bin is proportional to the filter bandwidth of the spectrum analyser.

For instance, with a filter bandwidth $\Delta f$, the effective width of the active reverberation region $\Delta r$ is

$$\Delta r = \frac{c\Delta f}{2|\mu|}$$  \hspace{1cm} (3.41)

If the filter widths for each range are equal, the reverberation volume increases with the square of the range hence the reverberation level is expected to increase as $20 \log R$ relative to the echo level from a constant target. Therefore to obtain a constant echo-to-reverberation level, it is necessary to use a spectrum analyser where the bandwidth of each filter varies inversely in proportion to the square of the centre frequency. This highlights the need to use narrow filter bandwidths at long ranges to overcome the effects of reverberation (cf. Kurie, 1946, p29).

### 3.5.5 Crosstalk

Crosstalk in a CTFM sonar can be either electrical or acoustical in origin. Electrical crosstalk can be minimised by careful circuit design, especially by shielding the small received signals from the large currents generated in the projector power amplifiers. Acoustical crosstalk is due to the direct acoustical coupling between projector and hydrophone and is a problem peculiar to continuous transmission systems.

Pulsed sonar systems can overcome this difficulty by gating the receivers off until the first wanted echo arrives. However as a result of their continuous nature, CTFM sonars must be able to cope with the large crosstalk signals superposed on the weak backscattered echoes.

To reduce the acoustical crosstalk, the hydrophone should be isolated from the projector by as much as possible. This is not always feasible, in particular, when the
sonar is mounted on a towfish. In this case, any direct acoustical coupling through the mounting system should be minimised, and acoustical baffles used to increase the acoustical path length between projector and hydrophone.

The crosstalk can be easily removed in CTFM sonars after demodulation since it appears as a very short range signal. However, it is preferable to remove the crosstalk before demodulation so that the receiver sensitivity can be increased. This requires swept notch filter techniques to track the crosstalk or, possibly, the use of adaptive cancellation techniques similar to echo cancellation techniques used in satellite communication systems (cf. Haykin, 1986, p22).

3.5.6 Beamwidth Considerations

As a result of the wideband nature of CTFM, there is a significant interaction between the signal waveform and beam pattern. For example, a narrowbeam CTFM sonar with an octave bandwidth, say, has a beamwidth that varies by a factor of two over the duration of a sweep.

A number of methods have been proposed to ensure constant beamwidth operation with broadband signals (cf. Smith, 1970; Smith, 1972, Ch.4; Knight, 1987, p64; Gough and Knight, 1989). The simplest and most effective method is to use an array of transducers controlled by a frequency selective shading network. The idea is to control the effective aperture dimensions by shading the transducer array, so that the ratio of effective aperture size to wavelength is kept constant.

Since the transmitted frequency of a CTFM sonar changes relatively slowly, it is a simple matter to shape the transmitted beam pattern, by controlling the gain of the projector array elements throughout the sweep. However, on reception it is necessary to sum the outputs from the hydrophone array using weighted filters. This is because the received signal consists of many different echoes, with different instantaneous frequencies.

3.6 Comparison of Pulsed and CTFM Sonars

There are a great number of waveforms suitable for echolocation. Each has unique characteristics that may be exploited to advantage in particular situations. While there have been a number of comparisons published on the various merits of CTFM versus pulse-echo sonar systems (Kurie, 1946; Kay, 1959; Kay, 1960; de Roos, 1986; de Roos et al., 1989), it must be remembered that all sonars are constrained by the sonar equation (see Section 1.6), and that sufficient energy must be transmitted to allow for signal detection in the presence of noise. In many cases the performance of a particular waveform depends on the nature of the noise or reverberation. As far as signal to white Gaussian noise is concerned, the choice of signal waveform and bandwidth is entirely arbitrary. However, for the case of a signal buried in impulsive noise, a long pulse of low power is superior to a short pulse of high power (Stewart and Westerfield, 1959, p873).

Sonars have been predominantly based on pulsed technology due to the simplicity of generating and processing of pulsed signals. Digital waveform synthesis and spectral analysis techniques now makes CTFM more competitive in these respects. Many of the traditional disadvantages of pulsed signals have been overcome with
pulse compression techniques; the linear FM pulse (or chirp) being the most common example. Many of the properties of the CTFM signal are similar to those of chirped pulses, which is not surprising since CTFM is just a coherent train of chirps, but with a 100% duty cycle.

Separate transmit and receive transducers are required for CTFM operation (and any other continuous transmission system), unlike monostatic pulsed sonars which only require one transducer time multiplexed between transmission and reception. But, unlike chirped sonars, CTFM sonars are not 'blind' for the duration of the transmitted pulse length, and are therefore useful for high resolution at close range.

A major practical difficulty of CTFM sonars is the problem of coping with the very large dynamic range of received echo signals. The biggest signal is the direct acoustic crosstalk between the projector and hydrophone transducers. The smallest signal of interest is from a weak target at maximum range. With a pulsed sonar, acoustic crosstalk is avoided by gating off the receiver until the first desired echo is received. The dynamic range in echo levels between near and far range targets can be reduced further using TVG techniques. The equivalent processing for CTFM signals is much more difficult to implement, and requires the use of swept notch filters or adaptive cancellation techniques to track the unwanted crosstalk.

The prime advantage of the CTFM signal is that it can provide continuous observation of targets while at the same time having a large bandwidth suitable for rejecting reverberation and environmental noise from unwanted ranges. Furthermore, very low peak power levels, well below the cavitation limit, are required for reverberation limited operation. This suggests that CTFM is particularly suited to long-range high-frequency applications (i.e. where there are large absorption losses), where conventional sonars are noise limited because of peak power constraints. The advantage is lost in short-range low-frequency applications, however, where sonars tend to be reverberation limited.

Finally, CTFM is a simple technique for generating a large bandwidth signal, but with a relatively slow change in instantaneous frequency (sweep rate). This slow linear frequency sweep simplifies the processing requirements considerably, since the broadband CTFM signal can be easily divided into narrowband components.
Chapter 4

The Prototype CTFM Synthetic Aperture Sonar

This chapter briefly outlines the prototype CTFM sonar constructed for the phase stability experiments described in Chapter 5, and the synthetic aperture experiments described in Chapter 7. Also included is a description of the pre-processing of the raw data — the splitting up of the broadband CTFM echo signals into narrow bands, and the range compression and range gating of these echoes (Section 4.2).

The sonar was originally designed with two overlapping hydrophones to provide differential phase measurements, similar to a radar phase monopulse system. This is briefly described in Section 4.3; however, neither the phase stability nor the synthetic aperture experiments made use of this facility.

4.1 Electronic Design

A general system block diagram of the prototype synthetic aperture sonar is illustrated in Figure 4.1. The sonar is split into three major sub-systems: the CTFM sonar front-end, the data recorder for storage of the demodulated echo signals, and an Intel 310 computer system for pre-processing of the raw data and for crude real-time displays at sea. These sub-systems are briefly described over the next few sections.

Refer to Table B.1 for a summary of the various sonar parameters, and also to Appendix A for an assessment of the theoretical performance of the sonar.

4.1.1 The CTFM Sonar Front-End

The CTFM sonar front-end comprises the electronics required to generate both the transmitted signal and reference local oscillators, to demodulate the received echoes, and to interface with the data recorder and computer sub-systems. Also included are the power supplies and a ‘test-module’ for diagnostic and calibration purposes.

The front-end is split into two physically separate units; the CTFM transmitter/receiver unit, and the underwater electronics. These units connect via a multicore umbilical cable, that is fastened to a steel towing cable.
Figure 4.1. System block diagram of the prototype sonar.

Sonar and Data Recording Block Diagram
4.1.1.1 Frequency Synthesisers

The transmitted and reference signals are generated using standard digital frequency synthesiser techniques. The transmitter is configured to sweep linearly from 30 to 15 kHz with a 0.8192 s sweep period (hereafter approximated to 0.8 s), and the two reference local oscillators required for dual-demodulation sweep from 90–75 kHz and 75–60 kHz (see Section 3.2.3).

The operating frequency was chosen as a compromise between being low enough so that extreme towfish stability was not required, and yet high enough so that the design of the transducers did not result in large physical dimensions. In addition, it was hoped that the choice of operating frequency would provide some sub-bottom penetration.

The desired maximum operating range was 400 m, and this required a sweep period of at least 0.53 s to avoid range ambiguities (see Section 3.5.1). A longer value was chosen to provide a guard band and to reduce the demodulated bandwidth requirements from 15 kHz to around 10 kHz. This simplified the design of the IF filter required in the receivers.

4.1.1.2 Receivers

There are two identical receivers in the sonar front-end, matched to close tolerances with regard to their phase characteristics. This is important for phase 'monopulse' applications as described in Section 4.3.

The received echoes are amplified, high-pass filtered (to remove low frequency sea noise), and then fed to a pair of balanced demodulators. To prevent overloading of the demodulators, a gain control and peak level detector are provided. Apart from the transmitted power level, this is the only other operator control.

The dynamic range of CTFM sonars is usually limited by the first demodulator, since this must cope with both the strong crosstalk signal and the weakest echoes. Therefore, to improve the dynamic range of the sonar, a swept notch filter was implemented to track and attenuate the crosstalk signal prior to this first demodulator. However, this circuit was not operational for any of the experiments described in this thesis prior to the Marlborough Sounds trial in March 1989.

After the first demodulator the unwanted sum frequencies are removed by an elliptical IF filter. This filter has notches at 45 kHz and 60 kHz to reject the crosstalk components and a frequency characteristic tailored to correct for the spreading and absorption losses. This gives approximately 40 dB of range equalisation. The echoes are finally brought down to baseband using a local oscillator fixed at 60 kHz.

4.1.1.3 Underwater Electronics

The underwater electronics consists of the hydrophone preamplifiers, the projector power amplifiers and projector beam shaping electronics. These are housed in the towfish (streamlined body) along with the hydrophone and projector arrays.

There are four identical power amplifiers, each connected to four transducer elements in the projector array. The gain of each of these power amplifiers being controlled by a single-chip microprocessor to provide some control of the projector beam
pattern. This microprocessor monitors the instantaneous frequency of the transmitted signal and changes the effective length of the projector array by controlling the gain of each of the power amplifiers.

Initial specifications for the sonar required the projector to have a 20° beamwidth independent of the transmitted frequency. Subsequently it was found that a constant beamwidth was unnecessary, and in fact undesirable (see Section 6.2.2), making the beam shaping circuitry redundant. Nevertheless, the beam shaping circuitry is still useful for maintaining a constant source level with frequency, if desired.

The underwater electronics connects to the rest of the sonar via an undersea multicore cable. This has four twisted-wire pairs; one pair for the transmitted signal, a pair each for the two hydrophone channels, and the extra pair for the projector power supply. The DC supply for the preamplifiers is sent separately using centre-tapped transformers on each of the two hydrophone balanced pairs. Apart from the advantages of power supply decoupling, this also allows the projector and hydrophones to be separately switched on or off.

### 4.1.1.4 Test-Module

Also included in the sonar front-end is a 'test-module', designed to simulate a delayed CTFM echo from a stationary target. Two output channels are provided to simulate the differential delay experienced by a pair of displaced hydrophones. This unit was found to be invaluable, since it allowed the sonar to be tested in the laboratory without recourse to a large body of water! In particular, it was useful for phase calibration measurements of the receivers.

The test-module was also employed to generate a delayed transmitted signal so that the acoustic crosstalk could be observed. Usually the crosstalk is removed by a notch filter in the receiver, but by transmitting a delayed sweep the crosstalk echo component demodulates to a different frequency outside the range of the notch filter.

The test-module is designed to simulate a 30–15 kHz linear frequency sweep with a sweep period of 0.8192 s, i.e. mimicking the transmitted signal. The entire sweep is stored in a number of EPROMs, and therefore unlike the transmitted signal, a change of sweep parameters requires a new waveform to be programmed.

The test-module is configured to produce delays corresponding to targets at ranges of 0 m up to 400 m, in steps of 50 m. The test-module is microprocessor controlled, and may be easily reprogrammed to synthesise many other different delays, with a minimum step size of 6.25 μs (≈ 5 mm in range, see Section B.9). In addition, it is possible to generate an additional fine delay in increments of 52.1 ns. This extra delay facility was designed to allow one channel to be delayed with respect to the other channel, simulating an echo from an off-axis target. This incremental delay can be changed on the fly (in real-time) to, say, simulate a moving target.

### 4.1.2 Data Storage

The demodulated echoes from the two receiver channels are stored on video tape using a commercial Pulse Code Modulation (PCM) unit together with a VHS Video Cassette Recorder (VCR). In conjunction, the PCM unit and VCR act as a very high quality stereo cassette recorder, with a 20 kHz bandwidth, a very large dynamic range
(96 dB for 16-bit operation or 84 dB for 14-bit operation), and most importantly, negligible wow and flutter. Although 16-bit operation provides a greater dynamic range, 14-bit operation was found to be preferable, since the PCM has better error correction in this mode.

Ironically, after the analogue signals have been reconstructed by the PCM unit on playback, they are resampled for transmission to the computer for processing. This was a deliberate decision so that the sonar was not tied to any particular data storage device. However, direct digital interfacing of the tape recorder to the computer may be preferable, particularly for reasons of synchronisation.

An important requirement of a coherent CTFM sonar is to record the start of sweep information along with the demodulated echoes, for reasons of synchronisation. On the current system this is achieved by storing a continuous 18.75 kHz tone on one channel, and on the other channel an 18.75 kHz tone that is phase shifted by 180° at the start and middle of each sweep. On playback these synchronising signals are filtered from the data, summed, rectified and then thresholded to regenerate the start of sweep signal. This start of sweep signal is necessary to reset the sampling system, to reduce the effects of any timing drift over the sweep period, and to provide a regular timing reference.

With the current system, a free running oscillator generates the 25 kHz sampling signal, although it would be desirable to generate it directly from the 18.75 kHz synchronising signals, since these are locked to the data on tape. This modification would reduce some of the phase error introduced by variations in the VCR tape speed, and drift in the frequency of the 25 kHz sampling signal.

A further modification would be to lock the sonar master oscillator to the PCM master oscillator. This helps to prevent ‘beating’ between the sonar and PCM master oscillators. Note that on recording, the PCM master oscillator ‘free-runs’, but on playback, it is locked to the recorded video data to minimise the effects of a variable tape speed.

4.1.3 Computer

Pre-processing of the raw demodulated data is performed using an Intel 310 computer system. This is a Multibus-I product with an 80286 processor board running the iRMX86 operating system. Added to the system is a Multibus array processor (DSP Systems AP-4). This device uses 16-bit two’s complement, fixed point arithmetic and can compute a FFT of up to 1024 complex samples in approximately 12 ms (including the loading and unloading of the input and output arrays). Although most of the ‘number-crunching’, for the images presented elsewhere in this thesis, was performed using VAX computers, the array processor allows the 310 system to be used as a real-time spectrum analyser. This is useful for display purposes when gathering data at sea.

A custom built Multibus board provides the interface with the CTFM sonar. The raw demodulated data from the sonar is sent serially in packets of 12-bit samples and these are buffered by the sonar interface board and converted into 16-bit integers. The conversion simply scales the 12-bit samples toward the MSB to help reduce truncation error during the FFT (cf. DSP Systems, 1983, p13). The raw samples are then blocked into frames of 1024 samples (see Section 4.2), before being transferred
to the array processor for real-time spectral analysis, or directly to hard disk for subsequent processing.

With each sample stored as 16-bit integer, sampling the two channels at a rate $f_s = 25$ kHz produces 100,000 bytes of data per second, or approximately 6 Mb of data per minute. Coincidentally, the largest contiguous file that could be stored was also 6 Mb for a 20 Mb (15 Mb formatted) disk. Therefore the longest measurement period is limited to one minute with two channel operation, or to two minutes with single channel operation. Any longer measurement interval either requires a larger disk capacity or real-time data extraction of the features of interest.

4.2 Pre-Processing of the Raw Sonar Data

The pre-processing of the raw demodulated data involves dividing each sweep into a number of contiguous sections (or frames), and computing the complex spectrum of each frame using a FFT algorithm. Essentially this operation divides the broadband transmitted signal into a number of narrowband linear FM pulses. These are then pulse compressed using the FFT to produce a number of range measurements per sweep, albeit at a poorer range resolution than if the entire sweep was compressed.

There are a number of benefits from processing the demodulated CTFM echoes in this manner. Often the full range resolution is not required anyway, in which case the frames can be averaged to reduce the effects of noise, coherent speckle, and target fluctuation. Secondly, no single block FFT can provide full range resolution over all ranges, because of the manner in which the echo sweeps are staggered in time (see Section 3.3.4). However, by taking shorter length FFTs, rearranging the outputs, and combining them with another FFT, it is possible to obtain nearly the full range resolution capability of the transmitted bandwidth (Sack et al., 1985, p47). Thirdly, shorter FFTs are faster and therefore the raw data may be coarsely range gated in real-time. The unwanted range bins can then be rejected to reduce the data rate, leaving the desired range bins for subsequent processing. Finally, splitting a CTFM sweep up into frames can considerably simplify the signal processing requirements, especially when broadband sweeps are transmitted (see Section 6.2.1).

4.2.1 Chopping up the Sweep

At the start of each transmitted sweep, the sampled demodulates echoes are resynchronised with the start of sweep timing mark to reduce the effects of any timing error. The first few samples for a period $T_b$ (the blanking interval) after the start of each sweep are ignored to avoid transients generated by the transmitted signal flyback and start of sweep synchronising signals. The following samples are then divided into $M$ contiguous frames, each of $N_m$ samples, where

$$T = T_b + MN_m T_s \tag{4.1}$$

and where $T_s = 1/f_s$ is the sampling interval. With the current prototype sonar, $T_b = 0$, and the sweep period $T$ is divided into $M = 20$ frames each of $N_m = 1024$ samples per channel. A frame size of 1024 samples was chosen simply because this was the largest FFT size that could be directly evaluated with the array processor. The other parameters were then chosen so that $MN_m = f_s T$. 

Figure 4.2. The parallelogram in the upper figure is the region of the unambiguous echoes from a transmitted sweep. The lower figure shows the region of the unambiguous demodulated echoes.

Each frame may be considered to be the demodulated echoes from a number of narrowband chirps, each with a duration $T_m = N_m T_s$, and a bandwidth $B_m$ of approximately $B / M$, where $B$ is the total sweep bandwidth.

From inspection of Figure 4.2 it is obvious that the demodulated echoes in each frame are derived from different parts of the transmitted sweep. For instance, consider an echo delayed by $\tau$. The mean received frequency of this echo (ignoring Doppler) in the $m$th frame is

$$f_m(\tau) = \begin{cases} 
 f_0 + \mu(t_m + T - \tau), & t_m + \frac{1}{2}T_m < \tau \\
 f_0 + \mu(t_m - \tau), & t_m - \frac{1}{2}T_m > \tau \\
 \text{Flyback}, & \text{otherwise}
\end{cases}$$

(4.2)

where $t_m$ is the time centre of the $m$th frame relative to the start of the sweep being transmitted, i.e.

$$t_m = T_b + (m - 1)T_m, \quad m = 1, 2, \cdots M$$

(4.3)

Note that when $\tau < t_m$, the echo is from the sweep currently being transmitted, otherwise the echo is from the previously transmitted sweep. Also note that the frame may encompass the flyback period of some of the echo signals, and for these echoes the frame is referred to as the glitch frame.

Because of the changing received frequency, the beam pattern of the hydrophone also changes. In addition, there are many overlapping sweeps simultaneously received, and so it is useful to define the nominal received beamwidth as

$$\theta_{B_m}(\tau) = \frac{\lambda_m(\tau)}{L_h}$$

(4.4)
where $\lambda_m(\tau) = c/f_m(\tau)$ is the mean received wavelength of the echo delayed by $\tau$. In side-scan applications off-axis echoes are also Doppler shifted in frequency. This, however, has a negligible effect on the beam pattern, except at very low frequencies.

### 4.2.2 Range Compression

The range compression operation is performed by taking the Fourier Transform of the demodulated echoes in each frame. Although a discrete Fourier Transform is involved, the description of the process is simpler within the continuous framework. The differences between the continuous approximation and the discrete computation are discussed further in Section 4.2.4.

Ignoring the glitch frame, the demodulated echo signal (3.20) written in complex notation for the $m$th frame is

$$d_m(t, r) = A \text{rect} \left( \frac{t - t_m}{T_m} \right) \exp j\phi_a(t, r),$$

where the demodulated phase $\phi_a(t, r)$ is given by (3.21). Taking the Fourier Transform of (4.5) and shifting the time origin to the frame centre $t_m$ gives

$$D_m(f, r) = A T_m \text{sinc} ((f - f_a)T_m) \exp j\phi_m(r)$$

where $\phi_m(r) = \phi_a(t_m, r)$ with this particular choice of origin.

As expected, the demodulated echo spectrum has a characteristic sinc($r$) response resulting from the rectangular time gating function, and this is centred at $fa = \mu r$, the demodulated frequency for the echo delayed by $r$. The nominal range resolution of each frame is thus

$$\Delta r_m \approx \frac{c}{2 |\mu| T_m} \approx \frac{c}{2B_m}$$

which can be seen to improve with a longer frame duration $T_m$; the highest range resolution being obtained when $T_m = T$. However, it is not possible to obtain the highest resolution at all ranges using a single FFT. If the FFT straddles the echo flyback, a glitch is produced and in the worst case complete cancellation can occur. To avoid this possibility, the glitch frame is usually rejected, and therefore it is desirable to have many short length frames to minimise wastage.

### 4.2.3 Phase Response

In a coherent imaging system the range compression operation must preserve the phase relationship of the echo signals. This is true of the FFT as can be seen from (4.6).

To determine the phase relationship between frames, (4.6) can be combined with (3.21) to yield

$$\Phi_m(\tau) = \begin{cases} 
2\pi \left( f_0 + \mu(t_m + T - \frac{1}{2}T_m) \right) \tau, & t_m + \frac{3}{2}T_m < \tau \\
2\pi \left( f_0 + \mu(t_m - \frac{1}{2}T_m) \right) \tau, & t_m - \frac{1}{2}T_m > \tau \\
\text{Phase glitch}, & \text{otherwise}
\end{cases}$$

(4.8)
where it is assumed that the time origin of the FFT coincides with the frame centre \( t_m \). Notice that consecutive frames from the same sweep differ in phase by \( 2\pi \mu T_m \), and therefore they can be coherently added with this as a phase correction to increase the range resolution.

The result given by (4.8) assumes that the target is stationary, i.e. \( f_a = \mu r \) is a constant frequency. If, however, the target is moving, \( f_a \) is no longer constant and decorrelation occurs. The result of the decorrelation is a broadening of the response function with a consequential reduction in the peak amplitude and an additional phase shift.

For the case of a target moving with a constant radial velocity, the demodulated signal contains a residual FM component. If \( \Delta f_m \) is the change in frequency over the frame duration \( T_m \), the residual FM produces a phase shift of

\[
\phi_r = \pm \tan^{-1} \left( \frac{S(\xi)}{C(\xi)} \right)
\]

where the plus sign applies when \( \Delta f_m > 0, \xi = \sqrt{|\Delta f_m| T_m / 2} \), and \( S(\xi), C(\xi) \) are the Fresnel sine and cosine integrals. Usually the phase shift is small for typical frame durations and can be ignored.

### 4.2.4 Practical Considerations

The discrete spectrum of each frame calculated using a DFT (or FFT) approximates the continuous spectrum given by (4.6) (cf. Bracewell, 1978, p377), but with a number of differences due to truncation effects. Since the spectrum of a rectangular function is not bandlimited, truncating a waveform violates the sampling theorem. The replicated parts of the sampled spectrum thus overlap producing a form of aliasing known as leakage. These leakage effects are more pronounced at higher frequencies and therefore mainly affect longer ranges.

To minimise spectral leakage, the input samples may be windowed (usually with a Hamming window since this can be performed by the array processor on the fly). However, windowing the input data broadens the width of the response (increased smoothing error) thereby degrading resolution. Moreover, the frames should not be windowed individually if they are going to be subsequently combined to improve the range resolution.

With the time function (4.5) sampled at \( nT_s \), the line spacing of the discrete spectrum is thus \( f_s / N_m = 1 / T_m \), and this corresponds to a time-gate width of

\[
\Delta \tau_m = \frac{f_s}{|\mu| N_m}
\]

or an equivalent range-gate width of

\[
\Delta R_m = \frac{c f_s}{2 |\mu| N_m} = \frac{c}{2 |\mu| T_m}
\]

Notice that this expression for the range gate width \( \Delta R_m \) is the same as (4.7), the expression for the range resolution \( \Delta r_m \) in the continuous case. However, targets in adjacent bins are not resolved since the Rayleigh criterion requires at least a 3 dB dip in the response. Therefore, to achieve the theoretical resolution \( \Delta r_m \) the
output of the FFT must be oversampled. In general, this is accomplished in practice by zero packing the data prior to the FFT and using a longer length FFT (see Section 3.3.4.1).

4.3 CTFM Phase Monopulse

The prototype sonar was built with two overlapping hydrophones (displaced in track) each with its own receiver. It was thought that the medium stability would prevent the formation of a coherent synthetic aperture, and that differential phase measurements obtained from a pair of displaced hydrophones could be integrated to form coherent sub-apertures. Obviously integration over the entire aperture is impossible since the signal magnitude often falls below the noise floor. Therefore, integration is limited to sub-apertures where the signal magnitude is sufficient. Perturbation techniques then could be applied to align the phases at the edges of the sub-apertures to form a full-sized aperture (Barron and Gough, 1986).

A typical phase monopulse configuration is shown in Figure 4.3, where the two displaced hydrophones are designated A and B. These are separated in track by a distance $\Delta L$, and are at an angle $\theta$ to the target at $T$. With this hydrophone geometry there is obviously a slight difference in target range for the two hydrophones. This differential range $\Delta R$ is related to the angle of arrival $\theta$, as

$$\theta = \sin^{-1} \left( \frac{\Delta R}{\Delta L} \right)$$  \hspace{1cm} (4.12)

assuming that $R \gg \Delta L$.

![Figure 4.3. A typical phase monopulse configuration where two hydrophones are separated by $\Delta L$.](image)

With a phase monopulse one attempts to infer $\Delta R$ from the difference in phase measured at the two hydrophones, and then calculates the angle of arrival using (4.12). This technique is well known in both radar and sonar applications (see e.g. Rhodes, 1980; Sherman, 1985; Henderson, 1987). There are difficulties, however, with this method. Taking the phase difference is a non-linear operation which causes
problems when there is more than one dominant target in any range annulus (Kliger and Olenberger, 1975; Gough, 1983).

Phase monopulse applications for sonar have therefore concentrated mainly on side-scan bathymetry (Cloet et al., 1982), where the two hydrophones are displaced vertically, and narrow beams are used in azimuth to reduce the possibility of multiple interfering targets. This monopulse configuration estimates the echo angle of arrival in elevation, which is then combined with the target range to estimate the target depth. Similar techniques have also been suggested as an extension to both synthetic aperture radar (Graham, 1974) and synthetic aperture sonar (Spiess and Anderson, 1983).

The phase difference obtained when using a CTFM sonar in a phase monopulse configuration is not quite so obvious because of the changing sweep frequency. To determine the differential phase let \( \tau_A \) be the delay to hydrophone A, and \( \tau_B \) be the delay to hydrophone B. Using (4.8), the differential phase measured for the \( m \)th frame is thus

\[
\Delta \Phi_m(\tau) = \begin{cases} 
2\pi \left( f_0 + \mu(t_m + T - \frac{1}{2}T) \right) \Delta \tau, & t_m + \frac{1}{2}T_m < \tau_A \\
2\pi \left( f_0 + \mu(t_m - \frac{1}{2}T) \right) \Delta \tau, & t_m - \frac{1}{2}T_m > \tau_B \\
\text{Phase glitch,} & \text{otherwise}
\end{cases}
\]  

(4.13)

where it is assumed that \( \tau_A < \tau_B \), and where \( \Delta \tau = \tau_B - \tau_A \) is the differential delay, and \( \tau = (\tau_A + \tau_B)/2 \) is the mean delay. Note that the term in brackets is simply the instantaneous received frequency given by (4.2) and therefore the phase difference may be expressed concisely as

\[
\Delta \Phi_m(\tau) = 2\pi f_m(\tau) \Delta \tau
\]  

(4.14)

These phase difference expressions assume, of course, that there is at the most only a single reflecting point target in each range bin. Nevertheless, they show that a CTFM sonar may be used in a phase monopulse, or phase interferometer configuration, provided corrections are made for the changing received frequency, i.e.

\[
\theta = \sin^{-1} \left( \frac{c \Delta \Phi_m(\tau)}{2\pi f_m(\tau) \Delta L} \right)
\]  

(4.15)

4.4 Summary

A CTFM synthetic aperture sonar with an octave bandwidth has been developed. This sonar has a wide beamwidth and unlike most broadband sonars the beamwidth is allowed to vary with the changing transmitted frequency.

Apart from the demodulation of the CTFM echoes, all the signal processing and signal generation is performed digitally to ensure linearity of the sweep for accurate timing and phase measurements. Ideally the echo demodulation should also be performed digitally. This would reduce the effect of frequency dependent phase shifts introduced by the analogue receiver electronics, and in addition, should help to improve the dynamic range of the sonar.
The sonar has been configured to provide differential phase measurements using a pair of displaced hydrophones. For a single isolated target at a given range, the phase difference depends on both the angle of arrival of the incident plane wave and the instantaneous received frequency. Therefore, provided there is no more than a single dominant target in any range annulus, the differential CTFM phase information may be used to estimate an image.
Chapter 5

Measurement of the Acoustic Temporal Phase Stability

The ultimate performance of any synthetic aperture sonar is limited by the acoustic path stability, or coherence, of the ocean medium. The longer the correlation time of the temporal coherence, the larger the acoustic aperture that can be synthesised. Therefore, to assess the potential performance of a synthetic aperture sonar, an experiment was performed to measure the temporal phase stability of the medium using the sonar described in Chapter 4. This chapter discusses the experimental set up and the results obtained.

5.1 Background

Although there is a vast literature concerning the theory of wave propagation in a random medium, experiments to determine the coherence of the ocean have concentrated mainly on the temporal amplitude fluctuations of the received echoes (Urick, 1979, Ch.11). Experiments to measure the temporal phase stability of the underwater medium are somewhat rare. Some groups have made phase stability experiments at frequencies lower than is usual for active sonar systems (e.g. Steinberg and Birdsell, 1966; Stanford, 1974; Williams and Battestin, 1976; Williams, 1976). Most of these experiments were performed over one-way acoustic paths using a bistatic configuration of projector source and hydrophone. The conclusion reached from these experiments was that the acoustic phase stability of the ocean was remarkably good, and much better than anticipated. Williams (1976, p70) added that the phase stability was good enough to form synthetic apertures in the ocean.

Lee (1979, p60) cites an experiment conducted by Stone et al. (1974) in which the rms fluctuations of acoustic path length in the ocean was measured to be of the order of one-hundredth of a wavelength (3.6°) at 10 kHz, over intervals of one minute for a 2500 m total acoustic path length. This experiment also employed a bistatic configuration, with a source at a depth of 70 m and a receiver array on the sea floor (at an unspecified depth).

Similar results were obtained by Christoff et al. (1982), but at a higher frequency of 100 kHz. A pulsed CW sonar was used, with a 140 μs pulse length, and a 1 s pulse repetition period, operating over a 48 m acoustic path. Here again a bistatic
configuration was employed, but this time with the source on the sea floor at a nominal depth of 12 m, and with the receiver hydrophone at different heights above the sea floor (from 3 m up to 9 m). For small heights above the sea floor (3 m), the phase standard deviation was as small as 2.3° (0.1 mm in 48 m) over a 20 min. interval. However, the phase stability was noted to get poorer as the hydrophone was moved closer to the surface. At a height of 9 m above the sea floor (i.e. 3 m below the surface), the phase standard deviation was as large as 18° (0.7 mm in 48 m) when measured over one 2 min. interval. This degradation in phase stability when operating close to the surface is probably caused by short delay (sea surface) multipaths.

5.2 The Effect of Noise on Phase Accuracy

Apart from propagation effects, phase errors are introduced by movement of the source or receiver, timing errors and phase instabilities in the measuring apparatus, and finally through the effects of a limited signal to noise ratio.

The range measurement precision of a CW signal can be determined from narrowband noise theory. Although the noise received at the hydrophone is broadband, a matched-filter receiver filter suppresses the out-of-band noise. Thus the matched-filter output signal consists of the wanted CW echo signal, superimposed with narrowband noise fluctuating at the same mean frequency as the signal. This can be expressed as

\[
x(t) = A \cos(2\pi f_c t) + n(t) \\
= (A + n_c(t)) \cos(2\pi f_c t) - n_s(t) \sin(2\pi f_c t)
\]

where \( n_c(t) \) and \( n_s(t) \) are the in-phase and quadrature narrowband noise components (Haykin, 1983, §5.11).

![Figure 5.1. Phase shift induced by narrowband noise.](image)

If \( N_c \) and \( N_s \) are the mean amplitudes of the in-phase and quadrature noise components, the mean phase shift \( \varphi \) due to noise can be seen from Figure 5.1 to be approximately

\[
\varphi \approx \tan \varphi \approx \frac{N_s}{A}
\]

where it is assumed that the signal is large compared to the noise — i.e. \( N_c \ll A \) (Rihaczek, 1969, p43). The variance of the signal phase \( \sigma^2_\varphi \) is similarly

\[
\sigma^2_\varphi = \frac{1}{A^2} \sigma^2_N = \frac{1}{A^2} \sigma^2_N
\]
where \( \sigma_N^2 \), the variance or average power of the noise, has been substituted for \( \sigma_N^2 \). From the theory of matched filters, the peak signal to noise ratio is \( 2E/N_0 \) (provided the noise is Gaussian and white, see Section 1.3), where \( E \) is the signal energy and \( N_0 \) is the noise power density. If \( A \) is the peak signal level, the average noise power is thus \( \sigma_N^2 = A^2N_0/2E \), giving a phase variance of

\[
\sigma_\phi^2 = \frac{N_0}{2E}
\]

(5.5)

The effect of this noise-induced phase shift is to simulate a range error, with a standard deviation \( \sigma_R = (\lambda/4\pi)\sigma_\phi \) (assuming a two-way path), or

\[
\sigma_R = \frac{\lambda}{4\pi} \sqrt{\frac{N_0}{2E}}
\]

(5.6)

where \( \lambda \) is the wavelength of the CW signal. Thus the range measurement precision, or accuracy, is dependent on the wavelength and the signal to noise ratio.

For phase accuracy measurements, a bistatic configuration with a self-radiating source, rather than a reflecting target, is preferable. This configuration has only one-way spreading losses (proportional to \( R^2 \)) in contrast to two-way spreading losses (proportional to \( R^4 \)) suffered by a monostatic configuration, and therefore can achieve a greater signal to noise ratio.

Rather than using a CW signal, a pulsed (or chirped) signal is often preferable to reduce the effects of reverberation, multipaths, and returns from other strong scatterers. With a continuous system, reverberation is received from all ranges simultaneously; however, with a modulated waveform it is possible to gate out reverberation from all the undesired ranges, and some of the multipath signals. Similarly, it is also desirable to have highly directive beam patterns. In addition to reducing the size of the active reverberation region (see Section 1.7), this also allows separation of multipaths in angle (in particular image interference from the sea surface, see Section 1.8.2).

The disadvantage of using a pulsed signal is that the ability to continuously monitor the medium fluctuations is lost. Instead the medium is sampled at the pulse repetition frequency. In addition, the integration period is greatly reduced, especially if short pulses are required for fine range gates.

Using a CTFM sonar, the acoustic path length variation can be continuously monitored. The echo sweeps can either be divided into many short duration chirps to observe the frequency dependence of the medium stability, or into fewer long duration chirps to obtain a better SNR and hence a better discrimination against multiple arrivals in the multipath propagation pattern.

### 5.3 Experimental Details

Before attempting to synthesise an acoustic synthetic aperture with the CTFM sonar described in Chapter 4, an experiment was performed to measure the temporal phase stability of the sonar range set up for the synthetic aperture experiment. This sonar range was located in Loch Linnhe, off the end of the Fort William pier, where the experiment was conducted over several days at the beginning of December 1987.
Loch Linnhe is a sea loch with considerable tidal flows (up to 6 knots at the pier end) with a significant amount of fresh water influx after rainfall. Thermal and salinity layering was therefore expected, but measurement with a commercial salinity and temperature meter found no noticeable layering.

At mean high tide the water depth varied from 5 m (next to the pier) to 20 m (some 70 m out from the pier), with a tidal variation of about 3 m. Because of the high tidal flows, measurement of the phase stability at times of slack water around the turn of the tide was desirable. In addition, operating at high tide with the sonar in mid-water, helped to minimise the effects of sea surface and sea floor multipath. However, biological scatter (from either seaweed or schools of fish) interfered with almost every measurement made at slack high water, so most measurements were obtained on the incoming and outgoing tides.

The sonar was configured to run on a cableway arrangement designed to minimise any tow path variations during the synthetic aperture experiments. This consisted of two vertically displaced steel cables strung between the end of the pier and an underwater gantry situated some 50 m out from the pier. Attached to this cableway was a pulley mounted cradle that carried the sonar and a three-axis navigation unit for monitoring sonar displacement. Unfortunately, gyroscopes in the navigation unit radiated acoustic noise, preventing any displacement measurements to be made during an experiment. Measurements between experiments, however, indicated that the effect of the sonar moving with the tide was a phase error always less than 45° (at 30 kHz). In slack or slowly moving water, the lateral displacement produced a phase error of less than 10°—i.e. a maximum towfish deviation of 1.4 mm.

Photo of buoys

Figure 5.2. The test target used is the steel-shell, air-filled buoy on the left.

The target, as shown in Figure 5.2, was an air-filled steel buoy with a nominal
diameter of 4 foot, and a theoretical target strength of \(-10\) dB (see Section A.1.5). The target was tethered to the sea floor by large blocks of concrete and shackled to the blocks so that it remained stationary under all conditions of tidal flow. During the experiments the reflecting surface of the buoy was about 7 m to 9 m below the surface (as indicated in Figure 5.3) and located approximately 66 m from the cableway. The optimum phase response of the prototype sonar is for targets somewhere in the range of 150 m to 200 m, but the steep profile of Loch Linnhe made it necessary to deploy the target at a much closer range than optimal.

![Figure 5.3. The geometry of the test target deployed in 10 m of water some 66 m out from the sonar track (N.B., not to scale).](image)

### 5.4 Experimental Results

The CTFM echoes were demodulated using a dual-demodulation technique (see Section 3.2.3), then sampled and stored on a video cassette recorder as described in Section 4.1.2. Sections of the demodulated data where the signal to reverberation ratio was sufficient, were later replayed and transferred to the Intel 310 computer system for analysis as explained in Section 4.2.

Because of a limited hard-disk capacity, only 1 min. periods of contiguous data (75 transmitted sweeps) could be stored before analysis. Although a longer measurement period was desirable, 1 min. was still sufficient to establish 'proof of concept', since even at a very modest tow speed of 1 ms\(^{-1}\), a 60 m synthetic aperture could be traversed.

To determine the range to the leading edge of the target, a spectral analysis of the demodulated echoes was performed over the full sweep, giving a nominal range resolution of 5 cm. An example of the range resolution obtained, when the entire 15 kHz bandwidth was analysed, is shown in Figure 5.4. This established that the leading edge of the sphere was 66.6 m from the sonar, and 25 dB above the
surrounding sea floor clutter. The demodulated echo sweeps were then divided into twenty equal length segments (or frames), each of 40.96 ms duration (say 40 ms for simplicity), and a 1024 point FFT was computed for each frame to gate the echoes into range bins.

Splitting the demodulated echoes into twenty frames provided independent measurements of the phase stability at twenty different frequency bands (within the transmitted bandwidth), albeit at the expense of the degradation of the range resolution to 1 m — compare this to the 5 cm range resolution available when the full sweep is used. However, a resolution of 1 m was still adequate to discriminate between the desired echo from the front edge of the sphere and the unwanted echoes from the concrete blocks, shackles and surrounding sea floor clutter.

The backscattering cross-section of a true sphere should be independent of both aspect and frequency — provided the sphere is large in terms of wavelength and has a rigid surface (Urick, 1975, §9.1). Practical test spheres have a backscattering cross-section that can vary with aspect and frequency (Hickling and Means, 1968). These differences result from welding seams, dents, suspension fittings (e.g. eyebolts, shackles and ropes) and other deviations from a true sphere. Moreover, it is not uncommon for the echo from the leading edge to be smaller than subsequent echoes Freedman, 1964, and that the first arrival echo from the leading edge of a sphere is less frequency dependent than subsequent echoes, from say the shackles etc. (de Roos et al., 1989, p1467).

Both these effects were observed with this experiment; the desired echo from the front edge of the sphere (bin 66 m) was not the strongest echo signal, but it varied the least with the changing sweep frequency. The adjacent bin (67 m) had the largest peak magnitude, but also the largest fluctuation over the sweep. This result is probably caused by coherent interference between multiple scatterers in the 67 m
range bin (for example, the shackles and concrete blocks) or is possibly due to short delay multipaths (from the sea floor or sea surface).

Therefore, the temporal phase stability was estimated from phase measurements of the 66 m range bin, for each of the twenty frequency bands. Over a 1 min. interval a sample population of 75 phase measurements per frequency band was obtained, and from all this data, histograms of the phase fluctuation around the mean frequency were calculated. Representative histograms are shown in Figure 5.5 for four of these frequency bands.

![Histograms of phase fluctuations (in degrees) around the mean phase at frequencies of (a) 20 kHz, (b) 24 kHz and (c) 28 kHz where the phase is estimated over a 40 ms period and the mean phase calculated from 75 phase estimates covering 1 min.](image)

The standard deviation of the phase fluctuations was expected to vary with the mean frequency of each band, since any acoustic path length fluctuations would have a greater effect on the phase measured at higher frequencies. However, the phase fluctuation was measured to be approximately 10° irrespective of frequency.

The signal to reverberation ratio was estimated to be better than 15 dB for all phase stability measurements, given the nature of the sea floor and the beamwidths of the sonar. Coincidentally, the phase accuracy of this technique calculated using (5.5) is also 10° for a 15 dB signal to reverberation ratio. In conclusion, the standard deviation of 10° is not a measurement of the acoustic stability, but a limit imposed by the physical and electronic stability of the test equipment or a limit imposed by the SNR.

### 5.4.1 Long Term Fluctuations

Over time scales much greater than 1 min., it is easy to perceive that the tidal movement changes the entire body of water within the acoustic path. As a consequence, the phase of any particular frequency band would show some long-term random drift. Interestingly, this mechanism is not particularly apparent in the long term measurements made by Christoff et al. (1982). As a comparison, the phase of frame 8 (corresponding to a mean frequency of 24 kHz) is plotted over a 15 min. period in Figure 5.6.
5.5 Summary

The acoustic medium (as measured in Loch Linnhe) is remarkably stable at frequencies around 25kHz. The phase standard deviation was found to be better than 10° (when averaged over a 40 ms interval). This corresponds to an approximate path length fluctuation of one part in 25,000. Therefore, the medium should be sufficiently stable to allow the formation of synthetic apertures, certainly for periods of around one minute.
Chapter 6

Simulation and Reconstruction

This chapter outlines a simple computer simulation of a CTFM synthetic aperture sonar operating under ideal conditions (Section 6.1), and the algorithms used to reconstruct images from both the simulated and measured synthetic aperture data (Section 6.2). Images reconstructed from simulated test targets are presented in Section 6.3, while Chapter 7 is devoted entirely to images reconstructed from measured synthetic aperture data.

6.1 Simulation of a CTFM Synthetic Aperture Sonar

To ascertain the validity of the reconstruction algorithms it was necessary to simulate the operation of a CTFM side-scan sonar. The simulation made many simplifying assumptions and approximations, but nevertheless, it was invaluable for the development of the reconstruction algorithms. More sophisticated reconstruction algorithms would of course require better models of the signal propagation conditions and target backscattering characteristics. However, the simple model was adequate to show 'proof of concept' of a CTFM synthetic aperture sonar under ideal operating conditions.

A number of ideal point targets (i.e. targets where the reflected energy appears to radiate from a point) could be represented. More complicated extended or distributed targets can be formed from the addition of these simulated single point target measurements. In addition, the output data format was designed to be compatible with the real sonar output, so that the same reconstruction algorithms could operate on both the real and simulated data.

6.1.1 Projector Radiation Pattern

To determine the received echo amplitude it is necessary to estimate the beam patterns for the transmitted and received signals. Because of the changing swept frequency, there is a considerable interaction of the CTFM signal on both the hydrophone and projector beam patterns. Fortunately the rate of frequency change is slow, which allows the beam patterns to be calculated using the instantaneous signal frequency.

The far-field radiation pattern of a projector can be expressed in terms of the
aperture pressure field distribution \( p(\eta, \zeta) \) (cf. Silver, 1949, §6.2) as

\[
p(R, y, z, \lambda) = \frac{j}{\lambda R} e^{-jkR} \int p(\eta, \zeta) \exp\left[ j k \left( \frac{\eta y + \zeta z}{R} \right) \right] d\eta d\zeta
\]

(6.1)

where \( R = (x^2 + y^2 + z^2)^{1/2} \) (see Figure 6.1), and \( k = 2\pi/\lambda \).

![Figure 6.1. Transducer beam pattern geometry.](image)

If the projector is assumed to be rectangular (length \( L_p \) and depth \( D_p \)) with a uniform phase and amplitude pressure distribution \( p_0 \), the integral of (6.1) is separable and can be solved to give

\[
p(x, y, z, \lambda) = \frac{j p_0}{\lambda R} e^{-j2\pi \frac{x}{\lambda}} L_p \text{sinc} \left( \frac{L_y y}{\lambda R} \right) D_p \text{sinc} \left( \frac{D_y z}{\lambda R} \right)
\]

(6.2)

Taking the squared modulus of (6.2) and replacing \( L_p D_p \) with the projector aperture area \( A_p \), the intensity of the field can be written as

\[
I(R, y, z, \lambda) = \frac{I_0 A_p^2}{\lambda^2 R^2} \text{sinc}^2 \left( \frac{L_y y}{\lambda R} \right) \text{sinc}^2 \left( \frac{D_y z}{\lambda R} \right)
\]

(6.3)

where \( I_0 = |p_0|^2/(2\rho c) \) and \( \rho c \) is the specific acoustic impedance of a plane wave (cf. Kinsler et al., 1982, §5.10). Equation (6.3) can be rewritten in spherical polar coordinates (see Section 2.1) as

\[
I(R, \theta, \psi, \lambda) = \frac{P_T G_p}{4\pi R^2} |F_p(\theta, \psi, \lambda)|^2
\]

(6.4)

where \( P_T = I_0 A_p \) is the total transmitted acoustic power (assuming that the aperture field is constant), \( G_p \) is the projector gain on axis,

\[
G_p = 4\pi \frac{A_p}{\lambda^2}
\]

(6.5)

and \( |F_p(\theta, \psi, \lambda)|^2 \) is the normalised projector beam pattern (intensity pattern),

\[
|F_p(\theta, \psi, \lambda)|^2 = \text{sinc}^2 \left( \frac{L_y}{\lambda} \cos \psi \sin \theta \right) \text{sinc}^2 \left( \frac{D_y}{\lambda} \sin \psi \right)
\]

(6.6)
6.1.2 Hydrophone Response

The frequency dependence of the hydrophone response can be approached in a similar manner to that of the projector response. Consider now a target with a backscattering cross-sectional area \( \sigma \), at a position \((R, \theta, \psi)\) in the far-field of a hydrophone. If the incident field intensity at the target is denoted by \( I \), the intensity of the field received at the hydrophone is \( I_\text{R}/(4\pi R^2) \), assuming that the target scatters the incident wave as an isotropic spherical wave. The received acoustic power is thus

\[
P_R = \frac{I_\sigma A_h}{4\pi R^2} |F_h(\theta, \psi, \lambda)|^2
\]

(6.7)

where \( A_h \) is the effective hydrophone area, and \( |F_h(\theta, \psi, \lambda)|^2 \) is the normalised hydrophone beam pattern. \( |F_h(\theta, \psi, \lambda)|^2 \) can be derived in a similar way to that of the projector using the reciprocity theorem for electroacoustic devices, (cf. Camp, 1970, §8.3).

Combining (6.7), (6.4), (6.5), and assuming that the hydrophone is situated near the projector, the ratio of the received acoustic power to the total transmitted power is

\[
\frac{P_R}{P_T} = \frac{A_p A_h \sigma}{(4\pi)^2 \lambda^2 R^4} |F_p(\theta, \psi, \lambda)|^2 |F_h(\theta, \psi, \lambda)|^2
\]

(6.8)

Therefore, provided the target reflectivity characteristic is frequency independent, the amplitude of the on-axis echo signal is expected to be inversely proportional to the transmitted wavelength.

In practice, the echo signals from off-axis targets are Doppler shifted in frequency due to the movement of the sonar. In general, this frequency shift is small compared with the transmitted frequency, hence the effect on the beam patterns is negligible.

6.1.3 The Effect of Sonar Displacement During Signal Propagation

If the sonar moves between transmission and reception, the difference in outward and return paths must be considered, especially when fast tow speeds and long propagation delays are expected.

![Figure 6.2. The effect of sonar displacement during signal propagation.](image)

To determine the effect of this displacement on the measured phase, consider Figure 6.2. Let \((0, y'_S)\) be the sonar position when the signal was transmitted and
(0, y_s) be the position when the reflected echo was received. The total acoustic path length is thus \( R' + R \), and the propagation delay \( \tau = (R' + R)/c \). During this time \( \tau \), the sonar moves a distance \( y_t - y_s = v \tau \), or

\[
y_s - y_t = \frac{v}{c} (R' + R) = a(R' + R)
\]

(6.9)

where \( a = v/c \). With reference to the triangle STS' in Figure 6.2 and application of the cosine rule, it follows that

\[
R'^2 = R^2 + a^2 (R' + R)^2 + 2a(R' + R)(y_t - y_s)
\]

(6.10)

which can be rearranged to give

\[
R' = \left( \frac{1 + a^2}{1 - a^2} \right) R + \left( \frac{2a}{1 - a^2} \right) (y_t - y_s)
\]

(6.11)

and

\[
R' + R = \left( \frac{2}{1 - a^2} \right) R + \left( \frac{2a}{1 - a^2} \right) (y_t - y_s)
\]

(6.12)

Thus the propagation delay \( \tau(y_s) \), written as a function of the instantaneous sonar position \( y_s \), is

\[
\tau(y_s) = \frac{R + R'}{c} = \frac{2R(y_s)}{c(1 - a^2)} + \frac{2a(y_t - y_s)}{c(1 - a^2)}
\]

(6.13)

where \( R = R(y_s) \) is the target range at the instant of reception.

If necessary, the first term in (6.13) can be compensated by changing the wavelength used in the reconstruction to \( (1 - a^2) \). However, this correction is usually extremely small and can be neglected. The second term represents an additional Doppler shift of \( 2av/(1 - a^2) \) (van de Lindt, 1977, p431). This is a constant term and produces a slight skew of the image in azimuth. The skew can be compensated by inserting a linear phase taper across the receiving synthetic aperture, so that the synthetic beam is steered forward by an angle of \( 2a \) (cf. Castella, 1971, p270).

With CTFM sonars it is necessary to have a separate hydrophone and projector, and these are usually separated to reduce acoustic crosstalk. If \( \Delta P \) is the along-track displacement, the round-trip path length can be described by a quadratic equation,

\[
R' + R = \alpha + \sqrt{\alpha^2 + \beta}
\]

(6.14)

where

\[
\alpha = \left( \frac{1}{1 - a^2} \right) R + \left( \frac{a}{1 - a^2} \right)(y_t - y_s) + \left( \frac{a}{1 - a^2} \right) \Delta P
\]

\[
\beta = \frac{\Delta P^2 + 2\Delta P(y_t - y_s)}{1 - a^2}
\]

(6.15)

(6.16)

Note that when \( \Delta P = 0 \), then \( \beta = 0 \) and \( R' + R = 2\alpha \), giving the same result as (6.12). For small displacements \( \Delta P \), the effect is to simply shift the along-track image position by \( \Delta P/2 \).
6.1.4 The Simulation Algorithm

This simulation algorithm attempts to model the demodulated signal produced by a CTFM side-scan sonar as it passes a number of ideal point targets. For a single isolated point target at \((x_t, y_t)\), the demodulated waveform can be described in terms of an amplitude factor proportional to the target strength and a phase term proportional to the acoustic path length. In addition, there is a slow amplitude modulation caused by the sonar beam pattern moving past the target. The demodulated signal can thus be described as

\[
d(t) = \text{Re} \left\{ \frac{A}{\lambda(t)} \text{sinc} \left( \frac{L_p}{\lambda(t)} \sin \theta(t) \right) \text{sinc} \left( \frac{L_h}{\lambda(t)} \sin \theta(t) \right) \exp j \Phi(t) \right\}
\]  

(6.17)

where \(A\) is the reflection coefficient of the target, \(L_p\) and \(L_h\) are the along-track dimensions of the projector and hydrophone transducers, and the other parameters are calculated as follows.

Step 1: The sonar along-track position \(y_s(t) = vt\).

Step 2: The angle of arrival (azimuth):

\[
\theta(t) = \tan^{-1} \left( \frac{y_t - vt}{x_t} \right)
\]

(6.18)

Step 3: The target range:

\[
R(t) = \sqrt{x_t^2 + (y_t - vt)^2}
\]

(6.19)

Step 4: The propagation delay \(\tau(t)\) using (6.13) to allow for the movement of the sonar during signal propagation.

Step 5: The transmitted wavelength \(\lambda(t) = c/f_T(t)\), where

\[
f_T(t) = \begin{cases} 
  f_0 + \mu(t_s - \tau(t)) & t_s \geq \tau(t) \\
  f_0 + \mu(t_s + T - \tau(t)) & t_s < \tau(t)
\end{cases}
\]

(6.20)

and \(t_s = t \mod T\) is the time since the start of the transmitted sweep.

Step 6: If necessary, the Doppler shifted received wavelength \(\lambda(t)/\eta(t)\) can be calculated, where the Doppler coefficient \(\eta(t)\) is

\[
\eta(t) = \frac{c - v \sin \theta(t)}{c + v \sin \theta(t)}
\]

(6.21)

but the slight shift in received frequency is negligible in all but the most extreme cases.

Step 7: The CTFM demodulated phase (assuming dual-demodulation or some other similar technique, see Section 3.2.3):

\[
\Phi(t) = \begin{cases} 
  2\pi \left( f_0 + \mu \left( t_s - \frac{1}{2} \tau(t) \right) \right) \tau(t) & t_s \geq \tau(t) \\
  2\pi \left( f_0 + \mu \left( t_s + T - \frac{1}{2} \tau(t) \right) \right) \tau(t) & t_s < \tau(t)
\end{cases}
\]

(6.22)
The demodulated CTFM signal (6.17) is sampled at a rate \( f_s \geq |\mu| (2\tau_{\text{max}} + W_d) \) to prevent aliasing, where \( W_d \) is the Doppler spread (see Section 2.1.3). The sampled signal is then chopped into frames and range compressed, as described in Sections 4.2.1 and 4.2.2.

### 6.1.5 Approximations and Assumptions

There are many simplifying approximations and assumptions made in the simulation algorithm. The important ones are listed below.

(i) The medium is assumed to be homogeneous — i.e. the refractive index is constant — and therefore ray paths are straight and the acoustic path length is simply the slant range \( R \).

(ii) The medium is also assumed to be linear, isotropic, and infinite, so that waves can be considered to propagate spherically, and without distortion.

(iii) The sonar is considered to travel at a constant speed \( v \), and without any deviation from a straight track. The medium is also assumed to be stationary so that the sonar water and ground speeds are the same.

(iv) Slant range distortion is ignored, i.e. \( \psi = 0 \).

(v) The target reflectivity is assumed to be aspect and frequency independent, i.e. there is no frequency signature.

(vi) Phase change on reflection is also ignored.

(vii) The projector and hydrophone beam patterns are sinc functions of the instantaneous transmitted and received frequencies (see Sections 6.1.1 and 6.1.2), and are perpendicular to the sonar track.

(viii) Losses due to absorption and spreading are ignored. These are assumed to be compensated for in the receiver.

(ix) Multipath and image interference effects are ignored.

(x) Finally, the signal to noise and signal to reverberation ratios are both assumed infinite.

### 6.2 CTFM Synthetic Aperture Reconstruction

The coherent reconstruction algorithm for CTFM synthetic aperture sonars can be sectioned into five steps:

Step 1: Filtering of broadband CTFM echoes into a number of narrowband chirps (Section 4.2).

Step 2: Range compression of the narrowband chirps (see Section 4.2.2).

Step 3: Phase correction of range compressed chirps. This is necessary to compensate for the frequency dependent phase shifts introduced by the analogue receiver electronics.

Step 4: Azimuth compression of the phase corrected and range compressed chirps into a number of narrowband images.
Step 5: Addition of the narrowband images to produce the resultant image.

The first two steps are explained in Chapter 4 (see Sections 4.2.1 and 4.2.2), and the other operations are explained in the sections which follow.

6.2.1 Phase Correction

As noted in Section 3.5.3, the analogue demodulation electronics introduce frequency dependent phase shifts. These include phase shifts that depend on the received frequency (i.e. phase shifts introduced before the first demodulator), and those that are range dependent (i.e. due to phase shifts of the demodulated, or range coded, signal).

If the ensemble of narrowband images are to be coherently added together, then the frequency dependent phase shifts must be compensated to give optimum performance. The range dependent phase shifts have little effect, however, provided that there is only a slight phase shift across any range locus.

In general, it is easier to compensate for the phase shifts after the range compression operation, but before the azimuth compression operation. The simplest method is to employ a two dimensional lookup table, with coordinates frame number and range bin. This lookup table may also be used to compensate for frequency dependent amplitude effects, say for example, if the projector-hydrophone response is not flat with frequency.

6.2.2 Azimuth Compression

The azimuth compression operation is essentially a matched filtering process, where the measured data is correlated with the expected amplitude and phase along the locus defined by (2.5) for each image point.

With respect to the $m^{th}$ complex narrowband image $C_m(x, y)$, the azimuth compression operation is the discrete analogue of

$$C_m(x, y) = \int_{-\frac{L_s}{4}}^{\frac{L_s}{4}} D_m(R, y + y') W \left( \frac{y'}{L_s} \right) \exp \left[ -j \Phi_m(x, y') \right] dy'$$

(6.23)

where $R = \sqrt{x^2 + y'^2}$ is the range from the along-track measurement point $(0, y + y')$ to the image point $(x, y)$, $L_s$ is the length of the synthetic aperture, $W(y'/L_s)$ is the along-track window function, and $\Phi_m(x, y')$ is the phase correction function needed to focus the synthetic aperture.

The synthetic aperture is considered to extend over some fraction $\rho$ of the main lobe of the hydrophone beam pattern, i.e.

$$L_s \approx 2x \tan \left( \frac{\rho \theta_{Bm}}{2} \right)$$

(6.24)

where $\theta_{Bm} \approx \lambda_m/L_h$ is the nominal 3 dB beamwidth corresponding to a wavelength $\lambda_m$ (see Section 4.2.1). Notice that the beamwidth is a function of frequency, and therefore longer apertures need to be synthesised at the lower transmitted frequencies to achieve a constant along-track resolution. For this reason, the transducer beamwidths should be allowed to change in azimuth with frequency, contrary to most real aperture side-scan sonars where a fixed beamwidth is desirable.
6.2.3 Addition of Narrowband Images and Speckle Reduction

There are a number of ways that the ensemble of narrowband images can be combined to form the resultant image. The simplest method is to simply add the intensities of each of the narrowband images, i.e.

\[ I_{nc}(x, y) = \sum_{m} |C_m(x, y)|^2 \]  \hspace{1cm} (6.25)

where \( I_{nc}(x, y) \) is the image resulting from the noncoherent addition. Alternatively, the ensemble of narrowband complex images may be coherently added together to produce a broadband complex image. This has an intensity \( I_{co}(x, y) \), where

\[ I_{co}(x, y) = \left| \sum_{m} C_m(x, y) \right|^2 \]  \hspace{1cm} (6.26)

Note that it is necessary to interpolate the narrowband images onto a finer pixel grid in cross-track, commensurate with the increased range resolution obtained using the full transmitted bandwidth. This additional step is not necessary with noncoherent addition.

Aside from its simplicity, noncoherent addition is robust with regard to the effects of frequency dependent phase errors introduced by the analogue demodulation electronics. Moreover, noncoherent addition has the important advantage that it may reduce the coherent speckle found in each of the narrowband images.

Speckle reduction techniques require statistical independence between members of the ensemble of speckle images — i.e. a lack of correlation. Robinson (1982, p101) suggests that the images obtained by different parts of a linear FM sweep are suitably independent for this sort of processing — even when the images are generated by scanning an aperture. This is because scattering patterns measured at different frequencies are likely to exhibit distortions that vary with frequency in a manner that is at least quasi-random (cf. Abbott and Thurstone, 1979, §VI).

If the speckle is primarily caused by the medium turbulence rather than the target scattering, then Shift-And-Add (SAA) based methods — both incoherent (cf. Bates and Cady, 1980; Bates and Robinson, 1981) and coherent (cf. Bates and Robinson, 1982; Robinson, 1982; Minard et al., 1985) — may be more appropriate than simple noncoherent image addition. These methods require a single dominant target highlight, so that the images may be aligned by simple shifts before addition. With multiple target highlights, however, SAA processing produces 'ghosts' in the resultant image (Bates and Cady, 1980).

An alternative method of obtaining independent images for speckle reduction, is to divide the full synthetic aperture into a number of shorter sub-apertures. Images are reconstructed from each of these sub-apertures and then noncoherently summed together. The resultant image has a poorer along-track resolution than an image produced using the full synthetic aperture, but the coherent speckle is reduced, producing a higher quality image. This is a common technique in SAR, where it is known as multi-look processing (cf. Kirk, Jr, 1975a, p331; Porcello et al., 1976; van de Lindt, 1977, p421; Ellis, 1984, p176).
6.2.4 Narrowband Approximations

The image reconstruction is simplified if each frame can be considered narrowband (in the Doppler sense, see Section 1.4.4). This condition requires that there is negligible range rate distortion of the signal envelope, and negligible effects of range acceleration and other higher order range derivatives (see Section 1.5).

Starting with the requirement of negligible range rate distortion, (1.41) states that the signal time-bandwidth product must be less than \( c/(8\dot{R}) \). Assuming that the maximum range rate of interest occurs at the edge of the nominal real beamwidth \((\theta_{Bm}/2)\), combining (2.12) and (4.4) gives

\[ \dot{R}_{\text{max}} \approx \frac{v\lambda_m}{2L} \]

and therefore the maximum time-bandwidth product of each frame is

\[ B_m T_m \leq \frac{cL}{4v\lambda_m} \]  

(6.28)

To achieve this condition it is necessary to divide the signal into \( M \) components, each of a duration \( T_m = T/M \) and bandwidth \( B_m \approx B/M \), where

\[ M \geq \left( \frac{4v\lambda_m BT}{cL} \right)^{\frac{1}{2}} \]

(6.29)

The condition for negligible effects of range acceleration can be similarly derived, but using (1.43) and (2.12) to give

\[ M \geq \frac{vT}{\sqrt{\lambda_m R_0/2}} \]

(6.30)

Except for very fast tow speeds \((v > 10 \text{ m s}^{-1})\) and close range targets, the range acceleration in general has little effect on individual sweeps. For example, using the prototype sonar parameters (listed in Table B.1) and choosing the worst case wavelength \( \lambda_m = 0.05 \text{ m} \), the closest operating range \( R = 50 \text{ m} \), and \( v = 5 \text{ m s}^{-1} \), then (6.29) requires that \( M > 6 \) and (6.30) requires that \( M > 4 \).

6.2.5 Algorithm Enhancements and Optimisations

Although not explicitly mentioned in the algorithm, it is advantageous to reconstruct synthetic aperture images with the cross-track index changing the most rapidly. This helps to reduce the regeneration of the along-track window function.

Computation throughput may be increased by exploiting parallel computer architectures to reconstruct each of the narrowband images simultaneously. The algorithm could also be optimised by applying efficient global operations (such as the FFT) to groups of pixels (cf. Sack et al., 1985). In addition, two dimensional correlation techniques may provide a simpler solution, where the range and azimuth compression are performed in a single operation.

Where the refractive index of the medium is not constant, but is known — for instance when there is thermal layering — then ray tracing techniques (cf. Tarng and Yang, 1987) could provide a better estimate of the propagation delays and arrival angles.
approximate function. The graphs are plotted as a function of along-track position for a target at 100 m, and cover a 40 m aperture length. The lefthand graphs correspond to the lowest sweep frequencies, while those in the righthand column correspond to the highest sweep frequencies. The dashed lines in the top two graphs show the shape of the hydrophone beam pattern, at each of these two different frequencies.

Three different approximations are illustrated. The graphs in the first row show the phase errors that result, if the effect of the sonar motion during the signal propagation is not considered. As predicted in Section 6.1.3, the error is simply a linear phase taper across the aperture, equivalent to steering the hydrophone beam pattern slightly away from boresight.

The second row (Figure 6.3c,d) shows the effect of approximating the phase correction function as a simple parabola. Although the range error is amplified at the higher frequency, the beamwidth is narrower, hence the approximation is better.

The third row (Figure 6.3e,f) shows the effect of ignoring the change in received frequency along the locus. Apart from a constant phase error, the effect over the most of the main beam lobe is minimal, especially at the higher frequencies where the beamwidth is narrower.

None of the images presented in this thesis were reconstructed using any one of these three approximations. However, the use of any one of these approximations would considerably reduce the computation required to generate an image.

### 6.3 Image Reconstructions from Simulated Data

For the purposes of the simulation, parameters close to those used with the prototype sonar were chosen. These are listed in Table 6.1. Note that the parameters have been selected so that the number of samples per sweep is a integer power of two. The real aperture size and the tow speed have been chosen to undersample the synthetic aperture by a factor of four.

In the development of the reconstruction algorithms many different target configurations were tried, but only a few of these are presented here. While only single point targets are shown, the response to more complicated target arrangements may be found from superposition.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial sweep frequency</td>
<td>$f_0 = 30.0 \text{ kHz}$</td>
</tr>
<tr>
<td>Sweep period</td>
<td>$T = 0.64 \text{ s}$</td>
</tr>
<tr>
<td>Frames per sweep</td>
<td>$M = 16$</td>
</tr>
<tr>
<td>Range bin width</td>
<td>$\Delta R_m = 0.8 \text{ m}$</td>
</tr>
<tr>
<td>Hydrophone length</td>
<td>$L_h = 250 \text{ mm}$</td>
</tr>
<tr>
<td>Projector length</td>
<td>$L_p = 250 \text{ mm}$</td>
</tr>
<tr>
<td>Tow speed</td>
<td>$v = 0.781 \text{ m s}^{-1}$</td>
</tr>
<tr>
<td>Terminal sweep frequency</td>
<td>$f_1 = 15.0 \text{ kHz}$</td>
</tr>
<tr>
<td>Sampling frequency</td>
<td>$f_s = 25.6 \text{ kHz}$</td>
</tr>
<tr>
<td>Samples per frame</td>
<td>$N_m = 1024$</td>
</tr>
<tr>
<td>Max aperture length</td>
<td>$L_a = 64.0 \text{ m}$</td>
</tr>
<tr>
<td>Hydrophone separation</td>
<td>$\Delta L_h = 35 \text{ mm}$</td>
</tr>
<tr>
<td>Projector displacement</td>
<td>$\Delta L = 0 \text{ mm}$</td>
</tr>
<tr>
<td>Speed of sound</td>
<td>$c = 1500 \text{ m s}^{-1}$</td>
</tr>
</tbody>
</table>

Table 6.1. Simulated CTFM synthetic aperture sonar parameters.
Figure 6.3. Phase errors introduced by approximations in the reconstruction process. (a)(b) Ignoring the effect of sonar displacement during signal propagation, (c)(d) Approximating the target range using the binomial approximation, (e)(f) Assuming received wavelength constant over the locus. Plots a,c,e are for Frame 3 (15.8 kHz) and plots b,d,f are for Frame 5 (28.9 kHz).
6.3.1 Range Compression of Simulated Data

The amplitude of the range compressed simulated data is plotted in Figure 6.4, in terms of track and range. This figure illustrates range curvature, and leakage in range whenever the response is not centred on a range bin. Notice how the dips in Figure 6.4, resulting from the picket fence effect, have been ameliorated by sinc-interpolation in range (see Figure 6.5). However, the interpolation is not so effective at the edges of the image, due to the truncation of the range data. Also notice how the response differs with frequency. For high frequencies in the sweep (Figure 6.5c), the response is narrower and the peak greater than for frequencies at the lower end of the sweep (Figure 6.5a).

![Image](image.png)

Figure 6.4. The amplitude distribution of frame 3 (15.8 kHz) of the simulated data, for a target at (100.4,32.0) plotted on a linear scale. The along-track direction is across the page to the left, and the cross-track direction slopes back to the right. The image distribution is centred at the target position, and only covers a small portion of swath.

6.3.2 Azimuth Compression of Simulated Data

The combined effect of undersampling the aperture and using narrowband signals is clearly shown in Figure 6.6. Here image distributions from three of the sixteen narrowband frames are presented. Note that the target has been deliberately positioned half way between cross-track pixels (the 'worst-case') and the aperture has been undersampled by a factor of four. These and other details of each image reconstruction are summarised in Table C.1.

Note that each of the figures show only a small portion of the available swath, with the image distributions being centred on the desired target. The actual sonar track is off the lower left-hand edge of the picture, and therefore the cross-track direction slopes away to the right. This is true of all the image distributions displayed in this thesis.

The result of this aperture undersampling is that grating lobes appear in the synthetic beam pattern, as evidenced by the along-track image ambiguities. Note how the ambiguities are spaced closer together at the high frequency end of the sweep.
Figure 6.5. The sinc-interpolated amplitude distribution of the simulated data for a target at (100.4, 32.0) plotted on a linear scale. (a) Frame 3 (15.8 kHz), (b) Frame 12 (22.4 kHz), Frame 5 (28.9 kHz).
(Figure 6.6c), and because of the octave sweep, the second ambiguity of Figure 6.6c is in a similar position to the first ambiguity of Figure 6.6a.

In addition to being attenuated by the real beam pattern, the ambiguities away from the main lobe are seen to be smeared slightly in the along-track direction, as well as being broader in the cross-track direction. This is because the differential delays across the synthetic aperture, to the ambiguous image points, become longer than the effective pulse width. The region of overlap is thus smaller, reducing the effects of interference (see Section 2.6.1).

Also note that because all three images share the same bandwidth, they each have a similar cross-track response. Whereas the images have side-lobes in cross-track, side-lobes are not apparent in the along-track direction. This is because the real beam pattern acts as a natural windowing function in the along-track direction (provided the synthetic aperture extends over most of the main lobe), while no attempt has been made to window the data before range compression.

Figure 6.7 shows the reconstructed image distribution for a target at a position (100.4, 32.0) calculated by both coherent and noncoherent addition of the sixteen narrowband images. As expected from the coherent addition, the cross-track resolution of Figure 6.7a has increased markedly from that shown in Figure 6.6 (approximately fifteen-fold if the glitch frame is ignored), and there is no sign of the coherent artefacts. The cross-track resolution of Figure 6.7b resulting from noncoherent addition is, by comparison, unchanged. Figure 6.7b also shows signs of the coherent artefacts in the background. These can be seen more clearly when plotted on a logarithmic scale (see Figure 6.8). The displayed dynamic range of these logarithmic image distributions is 20 dB.

The highest point of the background is 14 dB below the height of the main peak in the noncoherent case and 24 dB in the coherent case. The discrepancy is due to some coherent cancellation of the artefacts, and a much reduced region of overlap where interference effects can occur.

Figure 6.9 shows the reconstructed image distribution for a target lying halfway between two image pixels in cross-track (the 'worst-case' when using the full range resolution). Consequently, this image appears broader than Figure 6.7 in the cross-track direction. Sinc-interpolation of both Figure 6.8 and Figure 6.9 in the cross-track direction would reveal the same characteristic response.

Figure 6.10 shows the effect of splitting the received sweeps into thirty-two, rather than sixteen, frames. The image distribution (Figure 6.10a) formed from coherent addition of the thirty-two narrowband images is, as expected, much the same as Figure 6.8a. If anything, this image distribution should be slightly more faithful, since the narrowband approximations are more accurate. In addition, the glitch-frame is shorter, and therefore has less effect.

The non-coherent image distribution (Figure 6.10b) has a poorer cross-track resolution, as expected from reducing the bandwidth of each frame. The artefact level is still comparable with that in Figure 6.8b, even though twice as many narrowband images have been summed. This is because the artefacts in the narrowband images have increased in proportion to the reduction in range resolution.

Doubling the tow speed should halve the spacing between the grating lobes in each narrowband synthetic beam pattern. This is confirmed by the image distributions
Figure 6.6. The intensity distribution of three narrowband images reconstructed from simulated data for a target at (100.4,32.0). (a) Frame 3 (15.8 kHz), (b) Frame 12 (22.4 kHz), (c) Frame 5 (28.9 kHz). The image distributions are calculated on a pixel grid of 0.05 m cross-track by 0.1 m along-track, and each image covers a total area of 6.4 m by 25.6 m.
Figure 6.7. The intensity distribution of images reconstructed from simulated data for a target at (100,4,32.0). (a) Coherent addition, (b) Noncoherent addition. The image distributions are calculated on a pixel grid of 0.05 m cross-track by 0.1 m along-track, and each image covers a total area of 6.4 m by 25.6 m. Tow speed $v = 0.78$ m s$^{-1}$, 4 times undersampled.
Figure 6.8. These image distributions are the same as those of Figure 6.7, except presented on a logarithmic scale with a dynamic range of 20 dB.
Figure 6.9. The intensity distribution of images reconstructed from simulated data for a target at $(100.025,32.0)$, plotted on a logarithmic scale with $20$ dB dynamic range. (a) Coherent addition, (b) Noncoherent addition. The image distributions are calculated on a pixel grid of $0.05$ m cross-track by $0.1$ m along-track, and each image covers a total area of $6.4$ m by $25.6$ m.
shown in Figure 6.11. In addition, the artefact level is greater than that of Figure 6.8 because of the extra ambiguous energy.

Finally, halving the transmitted bandwidth from 15 kHz to 7.5 kHz degrades the cross-track resolution of both the coherent and noncoherent images (see Figure 6.12). The artefact level thus increases by 3 dB. Inspection of Figure 6.12b reveals that there are gaps in the artefacts, because the transmitted sweep no longer covers an octave bandwidth.

6.3.3 The Effects of Speed of Sound and Tow Speed Errors

Although not illustrated, an error in the estimate for the speed of sound defocuses the reconstructed image, and scales the image in the cross-track direction. There is also a slight image skew in the along-track direction, resulting from the sonar movement during the propagation delay being improperly accounted for. However, all these effects were found to have little impact on the final image for estimates within $10 \text{ m s}^{-1}$ of their correct value.

The effect of an error in the estimate of the tow speed is much more significant. This is illustrated in Figure 6.13 for two point targets at the same cross-track but displaced by 0.5 m in the along-track direction. The top two plots (Figure 6.13a,b) show an along-track slice through the images when reconstructed using the correct tow speed. The plot on the left is for frequencies at the low end of the transmitted sweep, and the plot on the right is for frequencies at the high end of the sweep. Note that although the along-track scaling is different for the two plots, the resolution is still the same.

The effect of overestimating the tow speed is illustrated in Figure 6.13c,d (+0.6% error) and Figure 6.13e,f (+1.2% error). Predictably, the synthetic beam pattern becomes more defocused in azimuth with greater tow speed errors. The effect is more pronounced for lower frequencies because of the longer apertures that need to be synthesised (see Section 2.8.3). In addition to the loss of main lobe height and increased side-lobe level, notice that spurious image detail has been generated in Figure 6.13e. This is obviously unacceptable, highlighting the need for accurate monitoring of the tow speed in high resolution applications.

6.4 Summary

This simulation was intended to test the CTFM synthetic aperture reconstruction algorithms under ideal conditions. Although the simulation did not model such effects as signal propagation and target backscattering, it proved invaluable for testing a variety of image reconstruction algorithms over different target configurations. More sophisticated reconstruction algorithms (e.g. for speckle reduction) would, of course, require better simulation models.

The use of broadband CTFM may relax the tow speed restrictions imposed by narrowband synthetic aperture processing. The greater the transmitted bandwidth, the smaller the compressed pulse length, and hence the region of interference is reduced. The coherent artefacts are not cancelled, but smeared, and thus add to the background clutter. With an octave bandwidth and undersampling the aperture by
Figure 6.10. The same as Figure 6.8, but with the sweep divided into 32, rather than 16, frames.
(a) Coherent addition, (b) Noncoherent addition.
Figure 6.11. The same as Figure 6.8, but with $v = 1.56 \text{ ms}^{-1}$, i.e. 8 times undersampled. (a) Coherent addition, (b) Noncoherent addition.
Figure 6.12. The same as Figure 6.8, but with the sweep bandwidth reduced from 15 kHz to 7.5 kHz. (a) Coherent addition, (b) Noncoherent addition.
Figure 6.13. The effect of overestimating the tow speed. (a) Frame 3 (15.8 kHz) \( v = 0.781 \text{ ms}^{-1} \), (b) Frame 5 (28.9 kHz) \( v = 0.781 \text{ ms}^{-1} \), (c) Frame 3 (15.8 kHz) \( v = 0.786 \text{ ms}^{-1} \), (d) Frame 5 (28.9 kHz) \( v = 0.786 \text{ ms}^{-1} \), (e) Frame 3 (15.8 kHz) \( v = 0.791 \text{ ms}^{-1} \), (f) Frame 5 (28.9 kHz) \( v = 0.791 \text{ ms}^{-1} \).
a factor of four, the minimum signal to self-clutter ratio is 24 dB. This is probably acceptable in most applications.

While addition of the narrowband frames on a noncoherent basis may be robust with regard to the effects of phase errors and speckle, the cross-track resolution is poorer than images generated using coherent addition. Therefore, the signal to self-clutter ratio is also poorer. Although more frames may be combined, splitting the sweep into a greater number of frames further degrades the cross-track resolution, with a corresponding degradation in self-clutter.

Finally, the along-track tow speed must be accurately determined for high resolution synthetic aperture imagery. However, the speed of sound is not so critical, and can be estimated reasonably accurately using empirical formulae.
Chapter 7

Image Reconstructions from Sea Trials

Although the short term phase stability of the undersea medium was shown in Chapter 5 to be sufficiently encouraging for the formation of underwater synthetic apertures, it remained to be seen whether or not this could be achieved in practice.

There is the additional problem of unwanted deviation of the sonar from the desired track. The slow mapping rate of a single beam synthetic aperture sonar, together with the influence of tidal and wave motion, can result in significant deviations from the desired sonar track. Conceptually, the influence of these motion errors can be compensated by accurately monitoring the position of the sonar and correcting the received signals accordingly. Ultimately, however, it is the medium induced phase errors that limit the performance of any synthetic aperture sonar. Therefore, in an attempt to minimise the effect of unwanted sonar motion, an experiment was performed in which the sonar was constrained to follow a fixed path (cf. Loggins et al., 1982).

The details of this experiment are described in Section 7.1, and synthetic aperture images reconstructed from this experiment are presented in Sections 7.2 and 7.3. The effect of imprecisely estimating the tow speed or the speed of sound is shown in Section 7.3.

7.1 Imaging the Test Target

The experimental details for the towed synthetic aperture experiment are similar to those described in Section 5.3 for the phase measurement experiments. In addition, the towfish was configured to slide along the taut wire cableway, so that the experiments could be repeated at different towfish velocities. However, all the results shown here were obtained with the sonar's towfish travelling at approximately 1 knot: a limit fixed by the towing mechanism (a pneumatic winch).

The sonar at this stage of development had the projector beam pattern controlled by a microcomputer, to maintain the beamwidth at a nominal 20°, independent of frequency. The hydrophones, however, had a frequency dependent beam pattern, and at the high end of the sweep their beam patterns were narrower than that of the projector. Consequently, the signal to noise ratio was poorer at the higher
transmitted frequencies. After this experiment the projector was modified to have a frequency dependent beam pattern similar to that of the hydrophones.

The images presented in this chapter were reconstructed using the algorithm described in Section 6.1.4. To reduce the number of parameters listed in each figure caption, the more pertinent parameters are tabulated in Table C.2. Of the many different runs performed, to date images have only been reconstructed from data collected on two of the runs. The first of the runs was performed at a tow speed of \( v = 0.51 \text{ m s}^{-1} \), and the images reconstructed from this are presented in Section 7.2. The second of the runs was performed on the following day at a slightly faster tow speed of \( v = 0.70 \text{ m s}^{-1} \). The images reconstructed from the second run are presented in Section 7.3. In both cases the maximum tow speed predicted by narrowband theory is approximately 0.16 m s\(^{-1}\) (see Section 2.4), hence the aperture was significantly undersampled on both these runs.

### 7.2 Results of Image Reconstructions

The intensity distribution of the measured data is shown in Figure 7.1. Although the along-track resolution is predictably poor, at least two targets at different ranges are apparent. The closer target (66.5 m) is the desired test-target, but the identity of the other target is unknown. This could be a reflection from the floating marker buoy (and coil of unused rope) used to mark the presence of the test target. This mysterious target appears to lie almost normal to the end of the cableway, and is therefore only detected at the low end of the frequency sweep when the beamwidths are wider, and when the sonar is near the end of the cableway.

The combined effect of undersampling the aperture and using narrowband reconstruction algorithms is clearly shown in Figure 7.2. This confusing image is the result of moving the towfish along the aperture about three times faster than the maximum speed required to obey the sampling criterion. The raw data for this image distribution is taken from the demodulated echoes of a sequence of chirped pulses, 40 ms long, separated in time by 0.8 s, with the frequency of the chirped pulse centred at 29 kHz and sweeping down over 750 Hz during the transmitted pulse. At the towfish speed of 0.51 m s\(^{-1}\), the samples are separated along the aperture by approximately 0.4 m. The image reconstruction algorithm is based on a 0.5 m cross-track by 0.1 m along-track pixel grid, and the resultant intensity distribution covers a sea floor area of 16 m cross-track by 32 m along-track, which includes the test target. In all the figures (which show only a small portion of the available swath), the sonar track is off the lower left-hand edge of the picture, and the bottom line of the along-track pixels are at a distance of 65 m from the track.

Although an image of the test target is evident in the displayed intensity distribution (Figure 7.2), the undersampled aperture has produced three artefacts, and what is worse, one of the artefacts is stronger than the desired image. In addition, the whole image shows the effects of coherent speckle. This image distribution is obviously unacceptable.

Figure 7.3 shows a selection of four (out of the twenty) narrowband image distributions from four different frequency bands. Each distribution, covering the same 16 m by 32 m area of the sea floor, is computed from the demodulated echoes of
Figure 7.1. The intensity distribution of the range-compressed measured data, with the along-track direction going across the page to the left. Note that this image distribution only covers a small portion of the swath, with the near range at 65 m and the far range at 71 m. (a) All twenty frames, side by side (b) Frame 8 (≈ 29 kHz). Note that the presence of the far range target is not so obvious in this image distribution.
Figure 7.2. The intensity distribution of a narrowband image of the test target calculated at 29.0 kHz. The image distribution is calculated on a pixel grid of 0.5 m cross-track by 0.1 m along-track, and covers a total area of 16 m by 32 m.

Chirped pulses centred at frequencies of approximately 29 kHz, 26 kHz, 23 kHz and 21 kHz. Note how the position of the desired image (that of the test target) remains fixed, while the displacement of the artefacts from the desired image change with frequency. Also note that, in Figure 7.3(d), the sea floor clutter appears to be stronger. Actually, since the echoes from the target become weaker at frequencies of 20 kHz and below, the clutter has simply been accentuated by the normalisation of the displays.

The image distribution shown in Figure 7.4 was generated by adding all twenty narrowband image distributions non-coherently. That is, the intensity of each narrowband image distribution was added. The effect of this addition is to smear out the artefacts while adding the intensities of the desired image. Although the combined desired image is in the correct position, its resolution is poor and some artefacts are still apparent. Simulations indicate that the artefacts could be further reduced if the echo strength of the target is held constant over the entire transmitted bandwidth. Despite these artefacts, addition on an intensity basis is robust and may prove to be the algorithm-of-choice for fast or real-time imaging.

Figure 7.5 shows the effect of coherently combining the twenty narrowband images. Only the centre portion of the reconstructed image distribution is displayed (since the higher resolution requires a magnified display to see the details). This image shows the buoy and the concrete blocks where the along-track display now covers only 6 m and the cross-track display from 65 m to 68 m.

The image details of the main target (the buoy) are demonstrably better than
Figure 7.3. The intensity distributions of the images generated from undersampled apertures at four different frequency bands: (a) centred at 29.0 kHz, (b) at 26.0 kHz, (c) at 23.0 kHz and (d) at 20.7 kHz. All four image distributions are calculated on a pixel grid of 0.5 m cross-track by 0.1 m along-track.
Figure 7.4. The resultant intensity distribution when all 20 narrowband images are added non-coherently (0.5 m by 0.1 m pixel grid).

Figure 7.5. The resulting intensity distribution when all 20 narrowband images are added coherently to create an image on a 0.05 m cross-track by 0.1 m along-track pixel grid covering an area of 6 m (cross-track) by 12 m (along-track), where the target is now in the centre of the displayed area.
Figure 7.6 shows the central portion of the image created by the non-coherent addition algorithm displayed to the same magnification as that used in Figure 7.5. Note that the cross-track resolution is noticeably worse, but the along-track resolution is maintained. Comparing Figure 7.5 and Figure 7.6, it is apparent that the fully coherent image has higher resolution, smaller artefacts, but a more speckled (noisy) appearance.

7.3 The Effects of Speed of Sound and Tow Speed Errors

The images presented in this section were generated from data obtained during the second run, in which the towfish was pulled along the aperture at a slightly faster speed of \( v = 0.70 \text{ m s}^{-1} \). All the displayed image distributions were generated by coherently adding all twenty narrowband images, and are similar to Figure 7.5. Furthermore, each image distribution covers an area of 3.2 m (cross-track) by 6.4 m (along-track), and they are all centred about the expected target position, for comparison.
Figures 7.7 and 7.8 illustrate the effect of incorrectly assessing the along-track tow speed. In both these figures, the top image distribution is a reference that was reconstructed using the tow speed estimated from stopwatch measurements. Consequently, this image distribution has the sharpest response. However, there are artefacts in cross-track, because of incomplete cancellation of the narrowband image sidelobes. This is a direct result of a mistake that occurred while on location in Scotland, in which each of the demodulated echo frames were irrevocably Hamming windowed before they were range compressed (see Section 4.2.4).

Incorrectly estimating the along-track tow speed has two effects. Firstly, the image distributions are incorrectly scaled in the along-track direction, as revealed by the shift in position of the peak response. Secondly, and more importantly, the along-track response becomes defocused. Note how an error, as small as 0.1 m s\(^{-1}\), in the along-track speed can produce a significant defocusing effect (see Figure 7.8b); the along-track sidelobe level has increased, the width of the response is broader, and the peak response has decreased. Overestimating the tow speed even further has caused the image distribution in Figure 7.8c to break up. This is obviously undesirable, emphasising the point made in Section 6.3.3, that the along-track tow speed must be accurately determined for high resolution synthetic aperture imagery.

Incorrectly estimating the speed of sound is not so severe as incorrectly estimating the along-track tow speed, since in practice the speed of sound can be estimated sufficiently accurately from measurements of the water temperature, depth, and salinity (cf. Urick, 1975, §5.4).

The effect of incorrectly assessing the speed of sound is much the same as incorrectly assessing the tow speed, except that the image is now incorrectly scaled in the cross-track direction. Figure 7.9 illustrates this effect for a speed of sound error of ±10 m s\(^{-1}\) (≈ 0.7%). Note that while the peak response has shifted several pixels in cross-track, the defocusing has produced only a slight degradation in the along-track resolution. Furthermore, it appears that the true speed of sound is closer to 1490 m s\(^{-1}\) than 1500 m s\(^{-1}\) as estimated.

### 7.4 Summary

On comparing Figures 7.1 and 7.5, it is apparent that the synthetic aperture processing has significantly improved the along-track image resolution. The theoretical along-track resolution of the prototype sonar is \(\Delta y = 0.125\) m, which compares favourably to the value of 0.2 m estimated from Figure 7.5.

Although the aperture was significantly undersampled, the suppression of image artefacts using broadband CTIFM has been demonstrated. While artefacts resulting from the aperture undersampling are apparent in each of the narrowband images, they occur at different positions and hence are smeared in the resultant image.

Noncoherent addition of the narrowband images produces a smoother, less speckly image, but with a reduced cross-track resolution and an increased background due to self-clutter. However, noncoherent addition is robust with regard to frequency dependent phase errors introduced by the sonar.

Finally, for accurate high resolution synthetic aperture imagery, the along-track tow speed must be accurately estimated, however, the speed of sound is not so critical.
Figure 7.7. The effect of varying the tow speed in the reconstruction of the test sphere below the true speed. $c = 1500 \text{ ms}^{-1}$, $\Delta x = 0.05 \text{ m}$, $\Delta y = 0.1 \text{ m}$. (a) $v = 0.70 \text{ ms}^{-1}$, (b) $v = 0.69 \text{ ms}^{-1}$, (c) $v = 0.66 \text{ ms}^{-1}$. 
Figure 7.8. The effect of varying the tow speed in the reconstruction of the test sphere above the true speed. \( c = 1500 \text{ ms}^{-1}, \Delta x = 0.05 \text{ m}, \Delta y = 0.1 \text{ m}. \) (a) \( v = 0.70 \text{ ms}^{-1}, \) (b) \( v = 0.71 \text{ ms}^{-1}, \) (c) \( v = 0.73 \text{ ms}^{-1}. \)
Figure 7.9. The effect of varying the speed of sound in the reconstruction of the test sphere. $v = 0.70 \text{ ms}^{-1}$, $\Delta x = 0.05 \text{ m}$, $\Delta y = 0.1 \text{ m}$. (a) $c = 1500 \text{ ms}^{-1}$, (b) $c = 1490 \text{ ms}^{-1}$, (c) $c = 1510 \text{ ms}^{-1}$.
Chapter 8

Conclusions and Recommendations for Future Research

This thesis reviews synthetic aperture processing techniques for undersea seafloor imaging, and extends them to broadband signals, in particular CTFM. A prototype synthetic aperture sonar based on CTFM has been developed, with which experiments were performed to determine the acoustic stability of the undersea medium and to synthesise an undersea synthetic aperture.

In this chapter, conclusions from the work presented in this thesis are drawn in Section 8.1, and recommendations for future research and development are presented in Section 8.2.

8.1 Conclusions

There are a number of substantive conclusions that can be made from this experimental and theoretical investigation. They include:

(i) Broadband operation is essential for high resolution synthetic aperture sonar imagery. Apart from providing a high cross-track resolution, broadband operation allows the tow speed constraints to be relaxed, albeit at the expense of increased self-clutter (or background). With broadband signals the effective pulse length is short, hence interference effects (e.g. grating lobes) caused by undersampling the aperture are reduced. The ambiguous signal energy, however, appears as background self-clutter. Thus, there is a trade-off between mapping rate and image quality.

(ii) There are a number of advantages in using CTFM instead of one of the many other possible broadband signals, one of which is that it can be straightforwardly decomposed into a number of narrowband components. Images may be generated from each narrowband component, and then coherently added to obtain the full across-track resolution commensurate with the transmitted bandwidth. Alternatively, the narrowband images may be noncoherently averaged to reduce the effects of coherent speckle, but at the expense of a degraded across-track resolution.
The acoustic phase stability ultimately limits the performance of any synthetic aperture sonar. However, under the conditions of intended operation, such as harbour, coastal, and other shallow water surveys, the acoustic phase stability does not appear to be as severe a problem as initially thought. This thesis demonstrates that it is possible to synthesise an undersea acoustic aperture under realistic operating conditions.

Although the aperture was significantly undersampled, the suppression of image artefacts using broadband CTFM has been demonstrated. While artefacts resulting from the aperture undersampling are apparent in each of the narrow-band images, they occur at different positions and hence are smeared in the resultant image.

Finally, for accurate high resolution synthetic aperture imagery, the along-track tow speed must be accurately estimated, otherwise significant image degradation occurs. However, the speed of sound is not so critical.

8.2 Recommendations

In the light of experience gained by working with the prototype sonar, the following modifications and enhancements are suggested:

(i) The vertical (elevation) beamwidth should be decreased, and the vertical beam pattern preferably tailored to illuminate only the desired swath. This would reduce reverberation (both surface and volume), reduce the likelihood of range ambiguities, and help to reduce sea surface multipath and image interference effects.

(ii) The horizontal (azimuth) beam width should also be decreased, thereby reducing the length of the aperture to be synthesised. This is important if the sonar is to be operated without towfish motion compensation. There are two alternatives for reducing the beamwidth; by either increasing the hydrophone length, or the transmitted frequency. Increasing the hydrophone length appears to be the better option, since it would allow an increase in the area mapping rate, albeit at a reduced along-track resolution.

(iii) Extensive attempts must be made to minimise the acoustic crosstalk between the projector and hydrophone, thereby improving the dynamic range of the sonar. Possible solutions include the development of more effective acoustic baffles, increased separation of the projector and hydrophone, the application of adaptive echo cancellation techniques, or a combination of all three. Alternatively, it would be interesting to mount a hydrophone directly on the face of the projector. Although the magnitude of the crosstalk signal would increase, cancellation of the crosstalk should be more straightforward than when there are multiple, unknown, crosstalk paths.

(iv) The CTFM demodulation electronics should be upgraded with an increased digital signal processing content. This would eliminate the frequency dependent phase shifts introduced by the analogue filters in the current system. Preferably, the whole receiver should be digitised, with the echo signals being
sampled directly at the hydrophone preamplifiers in the towfish. The digital signals could then be time multiplexed up a single coaxial cable, or optical fibre, to the data recorders on board the survey vessel. With the advances in digital signal processing technology, full digital demodulation of the CTFM echoes is not too unrealistic.

(v) The recorded data should also be stored digitally, and in a suitable format for direct interfacing with a computer. This would alleviate most of the timing errors encountered with the prototype sonar, since the start-of-sweep information may be readily encoded along with the echo data.

(vi) Upgrading the sonar for both port and starboard operation would not only double the area mapping rate, but may also prove useful for determining the towfish trajectory using the measured data. To reduce interference effects from the other channel, the transmitted CTFM signals can be arranged to sweep in opposite directions. Alternatively, triangular FM (or VFM) may be better for reasons of symmetry.

(vii) The performance of the sonar could be markedly improved by employing an array of independent hydrophones, each with its own preamplifier and receiver. This would allow the area mapping rate to be increased in proportion to the number of hydrophones. This may be necessary for aperture synthesis in regions of limited temporal coherence. Practical disadvantages include an increase in hardware complexity, data storage, and signal processing requirements. However, the signal processing complexity could be reduced by employing a single high speed digital CTFM demodulator that is time multiplexed between the hydrophone channels. In addition, measurements of the phase difference between array elements may be useful for determining towfish yaw and displacement from a straight trajectory.

(viii) Last, but by no means least, some form of failsafe recovery device is mandatory for use with towed sonars, so that minor accidents do not become major catastrophes!

In order to make synthetic aperture sonar a viable undersea mapping tool, further research is require in the following areas:

(i) Auto-focusing techniques for the determination of the towfish trajectory from the measured data.

(ii) Precise navigation techniques for motion compensation and the determination of the towfish along-track velocity.

(iii) Speckle reduction techniques, particularly multi-look processing and shift-and-add type methods.

(iv) The effect of multipath arrivals, the sound velocity profile, and other signal path distortions in the formation of an undersea synthetic aperture.
Appendix A

Specifications and Theoretical Performance

In this appendix, the specifications and preliminary performance of the CTFM synthetic aperture sonar in shallow water are presented. Apart from tests in a water tank (Knight, 1987), the operating performance of the sonar has not been measured. The figures presented here are based on theoretical predictions with some inspired guesswork. Since most acoustic phenomena of the sea are extremely variable, these figures give only a rough qualitative estimate of the operating performance. Measurements to confirm these predictions are complicated by the need for a large unrestricted body of water.

Most of the experimentation with the sonar has been conducted in Lyttelton Harbour (Christchurch, New Zealand), with the exception of the experiments performed in Loch Linnhe, Scotland (see Chapter 5). Lyttelton Harbour is unfortunately a commercial waterway, and after dredging, the bottom layer is stirred up. The water depth is typically 10 m, and the harbour bed is composed predominantly of sand and mud sediments.

A.1 The Sonar Equation

This section is primarily concerned with the signal to noise ratio of the sonar. The signal to reverberation ratio is estimated separately in Section A.2. Although the sonar was built with two identical partially overlapping hydrophones (cf. Knight, 1987, Ch.6), only the signal from one of these is considered in the following analysis.

A.1.1 Projector Source Level

The maximum transmitted power level was estimated at $P_T = 2 \text{ W}$ (when averaged over a sweep), and from (1.50) this corresponds to a projector source level of

$$\text{SL} = 10 \log_{10} \frac{P_T G_p}{I_0 4\pi r_0^2} = 189 \text{ dB re } 1 \mu \text{Pa at } 1 \text{ m}$$

(A.1)

where the on-axis gain of the projector $G_p$ was derived from the effective projector area $A_p$ using

$$G_p = \frac{4\pi A_p}{A^2} = 35 \text{ at } 22.5 \text{ kHz}$$

(A.2)
This expression assumes that the hydrophone dimensions are large compared to the wavelength $\lambda$ and that one side of the hydrophone is shielded from, or is insensitive to, acoustic signals (Burdic, 1984, p358).

### A.1.2 Cavitation Limit

Cavitation occurs when the acoustic pressure fluctuation developed by the projector exceeds the hydrostatic pressure. The hydrostatic pressure with depth $h$ is $p_a = p_{\text{atm}} + \rho gh$, where $p_{\text{atm}}$ is atmospheric pressure, $g$ is the gravitational constant, and $\rho$ is the density of water. Thus the pressure intensity required for the onset of cavitation is

$$I_c = \frac{p_{\text{atm}}}{2pc} \left( 1 + \frac{\rho gh}{p_{\text{atm}}} \right)^2 \text{ (W m}^{-2}\text{)}$$

(A.3)

With $c = 1500 \text{ m s}^{-1}$, $\rho = 1 \times 10^3 \text{ kg m}^{-3}$, $g = 9.81 \text{ m s}^{-2}$, $p_{\text{atm}} = 101.3 \text{ kPa}$, and $P_c = I_c A_p$ the maximum transmitted power is thus

$$P_c = 3.421(1 + \frac{h}{10.3})^2 A_p \text{ (kW)}$$

(A.4)

For the projector of area $A_p = 0.0125 \text{ m}^2$ operating at a depth of $h = 5 \text{ m}$, cavitation occurs for a transmitted power level of $P_c = 94 \text{ W}$, assuming that the projector develops a uniform pressure field across its aperture without any 'hot spots'. In practice, cavitation effects have been found to decrease with increasing frequency and reduced pulse length (Urick, 1975, Fig.4.6). The frequency dependence of cavitation below 10 kHz is slight, but rapidly increases for frequencies thereafter. At 30 kHz a wide variation in the cavitation threshold has been found (Urick, 1975, Fig.4.5), but on average, is a factor of three better than given by (A.4). Thus the peak power limit imposed by cavitation on the projector is in the order of 300 W. From (A.1) this corresponds to a source level of 213 dB.

### A.1.3 Pulse Compression Improvement

If the entire sweep of duration $T_s$ is compressed to the order of the transmitted bandwidth (i.e. $T_c = 1/B$), the pulse compression improvement factor becomes

$$\text{IF}_p = 10 \log_{10} \frac{T}{T_c} = 10 \log_{10} TB \text{ (dB)}$$

(A.5)

However, if only a fraction $1/M$ of the sweep is coherently compressed, the pulse compression improvement factor is reduced by a factor $1/M$, i.e.

$$\text{IF}_p = 10 \log_{10} \frac{T M}{T_c / M} = 10 \log_{10} \frac{TB}{M} \text{ (dB)}$$

(A.6)

Note that the sonar noise bandwidth is unchanged. Reducing the length of the analysis period to $T/M$ reduces the signal energy by a factor $M$, but this also reduces the noise energy by a similar amount. Therefore the overall signal to noise ratio only drops by a factor $1/M$.

For comparison with a pulse echo sonar, it is useful to include the pulse compression improvement factor with the projector source level, since it is the total signal
energy rather than the peak power level that determines the signal to noise ratio. Therefore with $T = 0.8 \text{ s}$ and $B = 15 \text{ kHz}$, the pulse compression gives a 41 dB improvement, and the equivalent source level is $SL' = SL + 10 \log_{10} P_p = 230 \text{ dB}$. To achieve a comparable range resolution and signal to noise ratio with an unchirped pulse echo sonar, it is necessary to transmit a very short pulse, $T_p = 1/B = 66.7 \mu\text{s}$, with a peak power level of $P_T = 25 \text{ kW}!$ This peak power level is certainly well over the cavitation limit, see Section A.1.2.

A.1.4 Noise

The total noise spectrum level is

$$N_{total} = N_{ship} + N_{sea} + N_{ot} + N_{th} \quad \text{(dB/Hz)} \quad \text{(A.7)}$$

where $N_{ship}$ is the shipping noise, $N_{sea}$ is the sea state noise, $N_{ot}$ is the noise due to the ocean turbulence, and $N_{th}$ is the ocean thermal noise. Note these are specified for a 1 Hz band, and thus the noise level $NL$ for a receiver equivalent noise bandwidth $BN$ is

$$NL = N_{total} + 10 \log_{10} BN$$

$$= 10 \log_{10} \frac{N_0 BN}{I_0} \quad \text{(dB)} \quad \text{(A.8)}$$

For the frequency range 15–30 kHz, the shipping and ocean turbulence noise can be neglected, and the combined thermal and sea state noise is approximately 42 dB/Hz (see Table A.1.4). It is interesting to note that for this frequency band and a sea state of 1, the combined thermal and sea state noise is frequency independent. For a noise bandwidth, $BN = 15 \text{ kHz}$, (A.9) predicts a total noise level $NL = 84 \text{ dB}$.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Thermal Noise</th>
<th>Sea State Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Sea State=1)</td>
<td>(Sea State=1)</td>
</tr>
<tr>
<td>15.0 kHz</td>
<td>8.5 dB/Hz</td>
<td>33.0 dB/Hz</td>
</tr>
<tr>
<td>22.5 kHz</td>
<td>12.0 dB/Hz</td>
<td>30.0 dB/Hz</td>
</tr>
<tr>
<td>30.0 kHz</td>
<td>14.5 dB/Hz</td>
<td>28.0 dB/Hz</td>
</tr>
</tbody>
</table>

Table A.1. Frequency dependence of sea noise and attenuation coefficient (compiled from EDO Corporation (1980)).

The directivity index of the hydrophone is approximately

$$DI = 10 \log_{10} G_h = 10 \log_{10} \frac{4 \pi A_h}{\lambda^2} \quad \text{(dB)} \quad \text{(A.10)}$$

where $G_h$ is the hydrophone gain. At the mean transmitted frequency, 22.5 kHz, the theoretical hydrophone gain is $G_h = 47$, or $DI = 17 \text{ dB}$. This is only an approximate parameter (cf. Urick, 1975, §3.3), because it assumes that the incident wave is perfectly coherent and the incident sea noise is isotropic.
A.1.5 Target Strength

A.1.5.1 Air-Filled Spheres

The target strength of thin-walled air-filled spheres is difficult to predict. Diercks and Hickling (1967) and Hickling and Means (1968), show that excitation of flexural waves around the surface of these spheres can produce variations in target strength with frequency of 40 dB or more. With solid spheres, compressional waves penetrate into the material of the sphere to produce a wide variation of target strength with frequency (Urick, 1975, §9.4). Further complications in target strength prediction arise from welding seams, dents, suspension fittings (e.g. eyebolts, shackles and ropes), and other deviations from a true sphere (see Section 5.3).

It is assumed for the present purposes that the frequency averaged target strength \( TS \) of a sphere is that of an ideal rigid sphere of radius \( a \), where

\[
TS = \begin{cases} 
10 \log_{10} \left( \frac{a^2}{4\pi^2} \right) & 2\pi a \gg \lambda \\
10 \log_{10} \left( \frac{100\pi^4a^6}{94\pi^2} \right) & 2\pi a \ll \lambda 
\end{cases} \quad (A.11)
\]

and \( r_0 \) is the reference distance, usually 1/m (Urick, 1975, Table 9.1).

For example, the nominal 4 foot diameter sphere \( (a = 0.61 \text{ m}) \) used for the phase stability experiment in Loch Linnhe has a calculated target strength of -10 dB. Compare this with the 18 inch diameter sphere \( (a = 0.23 \text{ m}) \) used for the towed synthetic aperture experiments in Lyttelton Harbour that has a calculated target strength of -18 dB (cf. Smith, 1972, p227).

A.1.5.2 Mines

Most modern mines are quasi-cylindrical objects of about 1 to 2 metres in length and half a metre in diameter. The measured target strengths range from +10 dB within a few degrees of beam aspect, to much smaller values at intermediate aspects (Urick, 1975, §9.10).

Approximating a mine as a cylinder of length \( L \) and radius \( a \), the target strength at long ranges is

\[
TS = 10 \log_{10} \left( \frac{aL^2}{4\lambda r_0^2} \right) \quad (\text{dB}) \quad (A.12)
\]

where the direction of the incident wave is assumed to be normal to the axis of the cylinder.

For example, if \( a \) is taken as 0.2 m, \( L = 1.5 \text{ m} \), \( \lambda = 66.7 \text{ mm} \) \( (f = 22.5 \text{ kHz}) \), and \( r_0 = 1 \text{ m}, (A.12) \) gives a target strength 5 dB. At short ranges (in the near field of the cylinder) the target strength is likely to be lower, since the intensity falls off like \( 1/R \) (cylindrical spreading), whereas at longer ranges the intensity falls off like \( 1/R^2 \) due to spherical spreading (Urick, 1975, pp284,285).

A.1.6 Signal to Noise Ratio
Table A.2. Transmission loss as a function of range, where $2TL_s$ is the two-way spreading loss, and $2TL_a$ is the two-way absorption loss calculated at $5^\circ$ C (compiled from EDO Corporation, 1980).

<table>
<thead>
<tr>
<th>Range</th>
<th>$2TL_s$</th>
<th>$2TL_a$</th>
<th>$2TL_a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 m</td>
<td>68 dB</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>100 m</td>
<td>80 dB</td>
<td>0.4</td>
<td>1.5</td>
</tr>
<tr>
<td>150 m</td>
<td>87 dB</td>
<td>0.6</td>
<td>2.3</td>
</tr>
<tr>
<td>200 m</td>
<td>92 dB</td>
<td>0.8</td>
<td>3.0</td>
</tr>
<tr>
<td>250 m</td>
<td>96 dB</td>
<td>1.1</td>
<td>3.8</td>
</tr>
<tr>
<td>300 m</td>
<td>99 dB</td>
<td>1.3</td>
<td>4.5</td>
</tr>
<tr>
<td>350 m</td>
<td>102 dB</td>
<td>1.5</td>
<td>5.3</td>
</tr>
<tr>
<td>400 m</td>
<td>104 dB</td>
<td>1.7</td>
<td>6.0</td>
</tr>
</tbody>
</table>

Consider a $-18$ dB target at maximum range ($R = 400$ m). From Table A.1.6 the combined two-way transmission loss for a target at 400 m is approximately 108 dB at midband frequencies. If the entire echo is coherently compressed, the signal to noise ratio per sweep is

$$\text{SNR} = SL + TS - 2TL - (NL - DI) + IF_p + IF_n$$

$$= 189 - 18 - 108 - (84 - 17) + 41 + 0 = +37 \text{ dB}$$  \hspace{1cm} (A.13)

Note that if the total sweep is split into $M=20$ frames, and if each frame is compressed separately, the signal to noise ratio per frame drops by $10 \log_{10} 20 = 13$ dB. However, the signal to noise ratios quoted above are not necessarily achievable, since it has been assumed that the sonar is sea noise limited (see Sections A.3 and 1.10.5).

### A.2 Signal to Reverberation Ratio

#### A.2.1 Volume Reverberation

The hydrophone having a slightly more directive response than the projector determines the size of the reverberating region. Approximating $\Omega$ by $\lambda^2/A_h$, the effective active reverberation volume (1.60) becomes

$$V = \frac{cT_c \lambda^2}{2 A_h R^2} \text{ (m}^3\text{)}$$  \hspace{1cm} (A.14)

assuming that the transmitted signal is compressed to a pulse of duration $T_c$. Now combining (A.14) with (1.60), the volume reverberation level is given by

$$RL_v = 10 \log_{10} \left( \frac{P_T A_T}{I_0} \frac{1}{R^2 \exp(2qR)} \frac{cT_c}{2 A_h} s_v \frac{1}{4\pi} \right) \text{ (dB)}$$  \hspace{1cm} (A.15)

Of more importance is the echo to volume reverberation level. This can be determined by combining (1.65) with (A.14) to give

$$EL - RL_v = 10 \log_{10} \left( \frac{A_h}{(\lambda R)^2} \frac{2 \sigma_v}{cT_c s_v} \right) \text{ (dB)}$$  \hspace{1cm} (A.16)
Notice that although the reverberation level from volume scatterers is independent of frequency, the echo to volume reverberation level increases with frequency. This is because both the projector and hydrophone directivity improve with frequency, thus increasing the source level, and reducing the size of the reverberation volume $V$. Therefore the echo level goes up, but the reverberation level stays the same.

Assuming the sweep is split into twenty frames, the effective compressed pulse duration $T_c$ is approximately $1/750$ Hz, or 1.33 ms. Let the range of interest be 100 m, then the effective reverberation volume calculated using (A.14) is $2.7 \times 10^3$ m$^3$ at the centre sweep frequency 22.5 kHz. Choosing a ball-park figure of $-70$ dB/m$^3$ for $S_v$ (i.e. $s_v = 1.25 \times 10^{-6}$ m$^{-1}$), and a target strength of $-18$ dB as before (i.e. $\sigma_s = 0.2$ m$^2$), (A.16) gives an echo to reverberation level of 18 dB. At a range of 400 m the reverberation volume increases to $4.3 \times 10^4$ m$^3$ and $\text{EL} - \text{RL}_v$ decreases to 6 dB. This value is probably too small for reliable detection. However, if the full transmitted sweep is compressed so that $T_c = 1/B$, or 66.6 $\mu$s, the effective reverberation volume becomes $2.1 \times 10^3$ m$^3$, giving $\text{EL} - \text{RL}_v = 19$ dB. Thus emphasising that a good signal to reverberation level requires the use of fine range gates and narrow beam patterns.

### A.2.2 Surface Reverberation

The equivalent azimuth beamwidth $\Theta$ of a rectangular hydrophone of length $L_h$ is $\lambda/L_h$, hence the active reverberation surface area from (1.63) is

$$ A = \frac{cT_c}{2} \frac{\lambda}{L_h} R \quad (\text{m}^2) \quad (A.17) $$

Note that if the projector is more directional than the hydrophone, then $L_p$ should replace $L_h$ in (A.17).

The surface reverberation level, obtained by combining (A.17) with (1.63), is

$$ \text{RL}_s = 10 \log_{10} \left( \frac{P_T A_p}{4 \Theta \lambda} \frac{1}{R^3 \exp (2qR)} \frac{cT_c}{2} \frac{1}{L_h} \frac{s_s}{4\pi} \right) \quad (\text{dB}) \quad (A.18) $$

Note that unlike volume reverberation, surface reverberation increases with frequency. This is because the source level increases with the square of the frequency, whereas the reverberation area reduces in proportion to the frequency.

Finally, the echo to surface reverberation level from (1.66) and (A.17) is

$$ \text{EL} - \text{RL}_s = 10 \log_{10} \left( \frac{L_h}{\lambda R} \frac{2}{cT_c} \frac{\sigma_s}{s_s} \right) \quad (\text{dB}) \quad (A.19) $$

Notice that the echo to surface reverberation level improves with frequency, but not as rapidly as the echo to volume reverberation level does.

Again choosing the mid-sweep frequency (22.5 kHz), and an effective pulse width $T_c$ of 1.33 ms, (A.17) gives the effective reverberation surface area $A = 26$ m$^2$ at a range of 100 m. Assuming a grazing angle of less than 5 degrees, and a relatively smooth sea surface, it would seem that $S_s$ is in the order of $-60$ dB/m$^2$ (Urick, 1975, §8.11). This gives an echo to sea surface reverberation level of 28 dB for the $-18$ dB sphere. At 400 m this decreases by 6 dB to 12 dB, a figure which is barely adequate for reliable detection.
The backscattering strength for a sand/mud sea bottom is typically $-35$ dB/m$^2$ at shallow grazing angles (McKinney and Anderson, 1964). Assuming the sonar insonifies a similar area of sea floor as given by (A.17), the echo to sea floor backscatter is $13$ dB for the $-18$ dB sphere again at $100$ m using (A.19) with $T_c = 1.33$ ms and $\lambda = 66.7$ mm.

The expected signal to reverberation ratios are summarised in Table A.3 for a $-18$ dB target at ranges of $100$ m and $400$ m. These values assume a compressed pulse length $T_c = 1.33$ ms commensurate with only compressing $1/20$th of a sweep. If pulse compression of the entire sweep is performed, another $13$ dB should be added to the listed values. Note that at near ranges, the reverberation is dominated by backscatter from the sea floor, but at longer ranges volume reverberation becomes more significant.

<table>
<thead>
<tr>
<th>$R$</th>
<th>EL $- RL_v$ (volume)</th>
<th>EL $- RL_s$ (sea surface)</th>
<th>EL $- RL_a$ (sea bottom)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 m</td>
<td>18 dB</td>
<td>28 dB</td>
<td>13 dB</td>
</tr>
<tr>
<td>400 m</td>
<td>6 dB</td>
<td>12 dB</td>
<td>7 dB</td>
</tr>
</tbody>
</table>

Table A.3. Range dependence of signal to reverberation ratio for $TS = -18$ dB, $S_v = -70$ dB/m$^3$, $S_s = -60$ dB/m$^2$ (surface), $S_a = -35$ dB/m$^2$ (bottom), $T_c = 1.33$ ms, $\lambda = 66.7$ m.

### A.3 Noise Limited Range

Active sonars are normally reverberation limited, i.e. sufficient energy is transmitted so that the maximum range is limited by the signal to reverberation ratio rather than the signal to background noise ratio (see Section 1.10.5). The transition between a noise limited sonar to a reverberation limited sonar occurs, when the reverberation level becomes greater than the directive noise level, $NL - DI$. Assuming that volume reverberation is the dominant source of unwanted reverberation, this can be expressed in non-logarithmic terms as

\[
P_T T_p > \frac{2 N_0 B_N R^2 \exp 2q R \lambda^2}{A_s c} \quad (J)
\]

where $T_p$ is the transmitted pulse length, and $A$ is the smaller of $A_h$ or $A_p$. This expression gives the minimum transmitted energy required for reverberation limited operation out to a range $R$. Note that this expression is unaffected by pulse compression, since both the noise and reverberation are reduced by the same factor.

Let $N_0 = 11 \times 10^{-15}$ Wm$^{-2}$Hz$^{-1}$, $B_N = 15$ kHz, $R = 400$ m, $q = 0.5$ m$^{-1}$ (i.e. $\alpha = 2.1 \times 10^{-3}$ dB/m), $A = 0.01$ m$^2$, $s_v = 1.25 \times 10^{-6}$ m$^{-1}$, $c = 1500$ ms$^{-1}$, and $T_p = 0.8$ s, then from (A.20) the maximum transmitted power required is in the order of $50$ mW. In a quieter environment with, say $S_v = -60$ dB/m$^3$ (i.e. $s_v = 12.5 \times 10^{-6}$ m$^{-1}$), the transmitted power could be increased to $0.5W$. 

A.4 CTFM Sonar Dynamic Range

The required dynamic range for a pulse echo sonar receiver is determined by the largest expected target at the nearest range of interest, and the smallest detectable target at the fartherest range of interest. The largest signal in a CTFM sonar system is the acoustic crosstalk between the projector and hydrophone. This crosstalk level can be expressed as

\[
CT = SL - CA \quad (\text{dB}) \tag{A.21}
\]

where \( CA \) is the crosstalk attenuation factor, conservatively estimated to be better than \(-20 \, \text{dB}\) for the current CTFM sonar. This can be improved by:

(i) the use of acoustic baffles
(ii) increasing the separation between the projector and hydrophone
(iii) making the projector and hydrophone more directional.

The received echo level \( EL \) for a \(-30 \, \text{dB}\) target at a range of 400 m is

\[
EL = SL + TS - 2TL + DI
\]

\[
= 189 - 30 - 108 + 17 = 68 \, \text{dB} \tag{A.22}
\]

where the values are calculated for the mean transmitted frequency. Therefore the receiver must have a dynamic range of at least 100 dB to detect this signal among the crosstalk. Once the crosstalk has been removed, the dynamic range requirements are drastically reduced. For example, if the largest expected target is \(+10 \, \text{dB}\) at a range of 50 m, the required dynamic range reduces to 78 dB. After range correction, the dynamic range only needs to be as large as the spread of the target strengths — in this case \(10 - (-30) = 40 \, \text{dB}\). Note that this value is much larger than what most display devices can cope with, highlighting the need for some form of dynamic range compression.
## Appendix B

### Sonar Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrophone length</td>
<td>$L_h = 255$ mm</td>
</tr>
<tr>
<td>Hydrophone depth</td>
<td>$D_h = 65$ mm</td>
</tr>
<tr>
<td>Hydrophone sensitivity</td>
<td>$\approx -180$ dB re 1 V/$\mu$Pa</td>
</tr>
<tr>
<td>Projector length</td>
<td>$L_p = 223$ mm</td>
</tr>
<tr>
<td>Projector depth</td>
<td>$D_p = 55$ mm</td>
</tr>
<tr>
<td>Initial transmitted frequency</td>
<td>$f_0 = 30.0$ kHz</td>
</tr>
<tr>
<td>Terminal transmitted frequency</td>
<td>$f_1 = 15.0$ kHz</td>
</tr>
<tr>
<td>Swept bandwidth</td>
<td>$B = 15.0$ kHz</td>
</tr>
<tr>
<td>Sweep period</td>
<td>$T = 0.8192$ s</td>
</tr>
<tr>
<td>Sweep rate</td>
<td>$\mu = -18.3105$ kHz/s</td>
</tr>
<tr>
<td>Sweep time-bandwidth product</td>
<td>$BT = 12288$</td>
</tr>
<tr>
<td>Maximum demodulated frequency</td>
<td>$(f_{a})_{\text{max}} = 9.75$ kHz</td>
</tr>
<tr>
<td>Fraction of bandwidth demodulated</td>
<td>$k_b = 0.65$</td>
</tr>
<tr>
<td>Sampling frequency</td>
<td>$f_s = 25.0$ kHz</td>
</tr>
<tr>
<td>Oversampling factor</td>
<td>$k_s = 1.282$</td>
</tr>
<tr>
<td>Maximum unambiguous range</td>
<td>$R_u = 614.4$ m</td>
</tr>
<tr>
<td>Maximum operating range</td>
<td>$R_{\text{max}} = 400.0$ m</td>
</tr>
<tr>
<td>Number of samples per sweep</td>
<td>$N_s = 20480$</td>
</tr>
<tr>
<td>Number of samples per FFT</td>
<td>$N_m = 1024$</td>
</tr>
<tr>
<td>Number of FFTs per sweep</td>
<td>$M = 20$</td>
</tr>
<tr>
<td>Number of samples ignored per sweep</td>
<td>$N_b = 0$</td>
</tr>
<tr>
<td>FTT analysis period</td>
<td>$T_m = 40.96$ ms</td>
</tr>
<tr>
<td>FFT frequency bin width</td>
<td>$\Delta f_m = 24.414$ Hz</td>
</tr>
<tr>
<td>Delay bin width</td>
<td>$\Delta r_m = 1.333$ ms</td>
</tr>
<tr>
<td>Range bin width</td>
<td>$\Delta R_m = 1.0$ m</td>
</tr>
</tbody>
</table>

Table B.1. Summary of default parameters for the prototype CTFM synthetic aperture sonar (assuming $c = 1500$ m s$^{-1}$).
## Appendix C

### Image Reconstruction Parameters

<table>
<thead>
<tr>
<th>Fig.</th>
<th>$M$</th>
<th>$v$</th>
<th>$\Delta x$</th>
<th>$\Delta y$</th>
<th>$I_m$</th>
<th>$(x_m, y_m)$</th>
<th>$B$</th>
<th>File</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.6a</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td></td>
<td></td>
<td>15</td>
<td>recon_yi.3</td>
</tr>
<tr>
<td>6.6b</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td></td>
<td></td>
<td>15</td>
<td>recon_yi.12</td>
</tr>
<tr>
<td>6.6c</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td></td>
<td></td>
<td>15</td>
<td>recon_yi.5</td>
</tr>
<tr>
<td>6.7a</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td>0.95</td>
<td>73,129</td>
<td>15</td>
<td>recon_yi.addc</td>
</tr>
<tr>
<td>6.7b</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td>0.06</td>
<td>73,129</td>
<td>15</td>
<td>recon_yi.addi</td>
</tr>
<tr>
<td>6.8a</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td>0.95</td>
<td>73,129</td>
<td>15</td>
<td>recon_yi.addc</td>
</tr>
<tr>
<td>6.8b</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td>0.06</td>
<td>73,129</td>
<td>15</td>
<td>recon_yi.addi</td>
</tr>
<tr>
<td>6.9a</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td>0.41</td>
<td>66,129</td>
<td>15</td>
<td>recon_ni.addc</td>
</tr>
<tr>
<td>6.9b</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td>0.06</td>
<td>65,129</td>
<td>15</td>
<td>recon_ni.addi</td>
</tr>
<tr>
<td>6.10a</td>
<td>32</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td>1.00</td>
<td>73,129</td>
<td>15</td>
<td>recon_li.addc</td>
</tr>
<tr>
<td>6.10b</td>
<td>32</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td>0.03</td>
<td>73,129</td>
<td>15</td>
<td>recon_li.addi</td>
</tr>
<tr>
<td>6.11a</td>
<td>16</td>
<td>1.56</td>
<td>0.05</td>
<td>0.1</td>
<td>0.23</td>
<td>73,129</td>
<td>15</td>
<td>recon_ji.addc</td>
</tr>
<tr>
<td>6.11b</td>
<td>16</td>
<td>1.56</td>
<td>0.05</td>
<td>0.1</td>
<td>0.01</td>
<td>73,129</td>
<td>15</td>
<td>recon_ji.addi</td>
</tr>
<tr>
<td>6.12a</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td>0.98</td>
<td>73,129</td>
<td>7.5</td>
<td>recon_hi.addc</td>
</tr>
<tr>
<td>6.12b</td>
<td>16</td>
<td>0.78</td>
<td>0.05</td>
<td>0.1</td>
<td>0.06</td>
<td>73,129</td>
<td>7.5</td>
<td>recon_hi.addi</td>
</tr>
</tbody>
</table>

Table C.1. This table lists some of the parameters used in the reconstruction of the figures displayed in Chapter 6. All units are in m s$^{-1}$, m, and kHz. The column headed $I_m$ is the relative intensity of the brightest pixel in the image, and $(x_m, y_m)$ is the image coordinate of this pixel.
<table>
<thead>
<tr>
<th>Fig.</th>
<th>( v )</th>
<th>( c )</th>
<th>( \Delta x )</th>
<th>( \Delta y )</th>
<th>( I_m )</th>
<th>( (x_m, y_m) )</th>
<th>( (x_m, y_m) )</th>
<th>File</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2</td>
<td>0.51</td>
<td>1500</td>
<td>0.5</td>
<td>0.1</td>
<td>0.26</td>
<td>67.0,22.6</td>
<td>5,227</td>
<td>move3.xva.8</td>
</tr>
<tr>
<td>7.3a</td>
<td>0.51</td>
<td>1500</td>
<td>0.5</td>
<td>0.1</td>
<td>0.16</td>
<td>66.5,27.1</td>
<td>4,272</td>
<td>move3.xva_A</td>
</tr>
<tr>
<td>7.3b</td>
<td>0.51</td>
<td>1500</td>
<td>0.5</td>
<td>0.1</td>
<td>0.26</td>
<td>67.0,22.6</td>
<td>5,227</td>
<td>move3.xva.8</td>
</tr>
<tr>
<td>7.3c</td>
<td>0.51</td>
<td>1500</td>
<td>0.5</td>
<td>0.1</td>
<td>0.92</td>
<td>67.0,21.9</td>
<td>5,220</td>
<td>move3.xva_12</td>
</tr>
<tr>
<td>7.3d</td>
<td>0.51</td>
<td>1500</td>
<td>0.5</td>
<td>0.1</td>
<td>0.56</td>
<td>67.0,21.4</td>
<td>5,215</td>
<td>move3.xva.15</td>
</tr>
<tr>
<td>7.4</td>
<td>0.51</td>
<td>1500</td>
<td>0.5</td>
<td>0.1</td>
<td>3.9</td>
<td>66.5,27.2</td>
<td>4,273</td>
<td>move3.xva.addi</td>
</tr>
<tr>
<td>7.5</td>
<td>0.51</td>
<td>1500</td>
<td>0.05</td>
<td>0.1</td>
<td>42</td>
<td>66.45,27.2</td>
<td>65,65</td>
<td>move3.xxa.addc</td>
</tr>
<tr>
<td>7.6</td>
<td>0.51</td>
<td>1500</td>
<td>0.05</td>
<td>0.1</td>
<td>4.0</td>
<td>66.60,27.2</td>
<td>68,65</td>
<td>move3.xxa.addi</td>
</tr>
<tr>
<td>7.7a</td>
<td>0.70</td>
<td>1500</td>
<td>0.05</td>
<td>0.1</td>
<td>44</td>
<td>66.50,16.9</td>
<td>33,33</td>
<td>move4.xua.addc</td>
</tr>
<tr>
<td>7.7b</td>
<td>0.69</td>
<td>1500</td>
<td>0.05</td>
<td>0.1</td>
<td>39</td>
<td>66.50,16.8</td>
<td>33,32</td>
<td>move4.xug.adddc</td>
</tr>
<tr>
<td>7.7c</td>
<td>0.66</td>
<td>1500</td>
<td>0.05</td>
<td>0.1</td>
<td>21</td>
<td>66.50,16.3</td>
<td>33,27</td>
<td>move4.xue.addc</td>
</tr>
<tr>
<td>7.8a</td>
<td>0.70</td>
<td>1500</td>
<td>0.05</td>
<td>0.1</td>
<td>44</td>
<td>66.50,16.9</td>
<td>33,33</td>
<td>move4.xua.addc</td>
</tr>
<tr>
<td>7.8b</td>
<td>0.71</td>
<td>1500</td>
<td>0.05</td>
<td>0.1</td>
<td>33</td>
<td>66.60,17.2</td>
<td>33,36</td>
<td>move4.xuh.adddc</td>
</tr>
<tr>
<td>7.8c</td>
<td>0.73</td>
<td>1500</td>
<td>0.05</td>
<td>0.1</td>
<td>38</td>
<td>66.50,17.3</td>
<td>33,37</td>
<td>move4.xuf.adddc</td>
</tr>
<tr>
<td>7.9a</td>
<td>0.70</td>
<td>1500</td>
<td>0.05</td>
<td>0.1</td>
<td>44</td>
<td>66.50,16.9</td>
<td>33,33</td>
<td>move4.xua.addc</td>
</tr>
<tr>
<td>7.9b</td>
<td>0.70</td>
<td>1490</td>
<td>0.05</td>
<td>0.1</td>
<td>56</td>
<td>66.05,16.9</td>
<td>24,33</td>
<td>move4.xud.adddc</td>
</tr>
<tr>
<td>7.9c</td>
<td>0.70</td>
<td>1510</td>
<td>0.05</td>
<td>0.1</td>
<td>36</td>
<td>66.90,17.0</td>
<td>41,34</td>
<td>move4.xub.adddc</td>
</tr>
<tr>
<td>—</td>
<td>0.70</td>
<td>1495</td>
<td>0.05</td>
<td>0.1</td>
<td>57</td>
<td>66.25,16.9</td>
<td>28,33</td>
<td>move4.xuc.adddc</td>
</tr>
</tbody>
</table>

Table C.2. This table lists some of the parameters used in the reconstruction of the figures displayed in Chapter 7. All units are in \( \text{m s}^{-1} \), m, and kHz. The column headed \( I_m \) is the relative intensity of the brightest pixel in the image, and \((x_m, y_m)\) is the image coordinate of this pixel. This corresponds to a target position \((x_m, y_m)\). Note, the high resolution images reconstructed from the run \texttt{move3} are centred around the point (66.6, 27.2), whereas the images reconstructed from the run \texttt{move4} are centred around the point (66.5, 16.9).
References


BATES, R.H.T. (1964), Graphical Presentation of the Statistic of Fluctuating Target Detection by Pulse Radar, Special Report SR-120, MITRE Corporation, Bedford, Massachusetts, November.


EDO Corporation (1980), ‘Sonar and transducer computer’.


HAYES, M.P. and GOUGH, P.T. (1989), 'Results from an experimental synthetic aperture sonar', September. To be presented at the 18th International Symposium on Acoustical Imaging, Santa Barbara, California.


