Waveform Design for Spaceborne Synthetic Aperture Radar

By

Woo Kyung Lee

A thesis submitted to the University of London for the Degree of Doctor of Philosophy in Electronic Engineering

Department of Electronic and Electrical Engineering
University College London

August 1999
Waveform Design for Spaceborne Synthetic Aperture Radar

By

Woo Kyung Lee

Submitted to the University of London
on 25 August 1999 for the degree of Doctor of Philosophy

Abstract

Synthetic Aperture Radar (SAR), which was first developed in the 1950's as a technique for improving the resolution of military reconnaissance radar, has rapidly matured as a remote sensing tool for a wide range of civilian applications. Although the signal waveform is the very medium that delivers target information, little effort has been made yet on the aspect of waveform design and its influence on the generated SAR image. This thesis aims at addressing the role of waveform design in SAR, and developing new types of waveform that may suit some particular SAR applications.

Several propositions are put forward in regard to the properties of the P code compression output and their corresponding proofs are provided accordingly. Based on the sidelobe formulation at the pulse compression output, a novel pulse compression technique is developed that produces an optimal uniform range sidelobe of the ideal Barker code level. The newly obtained sidelobe pattern provides an optimal compromise between the range resolution and the peak sidelobe level (PSL), regardless of the signal code length, and also retains the merit of strong resistance to the Doppler shift effect. Very low sidelobe regions appear around the mainlobe peak, which is useful in target tracking and recognition.

When distributed targets are concerned, the integrated sidelobe level (ISL) is considered as a useful measure for the image quality. A novel processing technique is proposed to reduce the ISL energy. A sidelobe canceller is generated directly from the incoming signal and combined together with the original pulse compression output in which unwanted sidelobes are significantly eliminated.

With the assumption that many dominant features found in images of urban areas originate from corner reflection phenomena, it is attempted to model the electromagnetic interaction of radar signals with urban structures. The new pulse compression technique successfully accommodates wide dynamic responses of the urban features and provides a high image contrast without sacrificing target resolution.
Acknowledgement

I believe a thesis is a blend of knowledge, a history of self-discipline and a product of creative challenge combined with the passion to make a positive progress, all of which require not to be alone. There are numerous people without whom, I wouldn’t dare to have accomplished what I had put in my thesis.

My sincere gratitude is given to my supervisor Hugh D. Griffiths for having provided such a wonderful opportunity in which his kind supervision and a full support with his scientific expertise has been a source of continuous inspiration.

It is not easy to ask for a trust when I have nothing but a mere confidence out of a shallow experience and a desire to challenge. In this sense, I have been greatly privileged to obtain generous supports from Dr. Soon-Dal Choi, who granted such a warm trust on me.

The profound knowledge and useful advice from professors K. Milne and R. Benjamin have always been great assets, which I am proud of having had around me for years. Especially the positive comment from Prof. R. Benjamin helped me to bring up the courage to put an end to my thesis with a relief. A cordial friendship with David Hall from Matra Marconi Space, has always been and will be continuously appreciated along with his kind support through provisions of various useful materials.

I would like to thank all the members in the antenna and radar group at UCL, without whom my Ph.D. degree would have brought me tears of solitary and painful boredom. They are Richard, Lluis, Joe, Ricardo, Rob, Fabrizio, Glen, Yiev, Sam, Fiona, Jonathan as well as Moon and more, all of whom I would prefer to remember as a flat mate, party members, English teachers, counselors, the great failures in romance, and witnesses of my unforgettable life in England.

I am obliged to thank all people back in SaTReC, Korea, to whom I owe this precious opportunity to open an eye for SAR. I am in debt for their dedication to SaTReC, which has constantly motivated me to take nothing for granted.

And finally there are people in UCL Korean Society, the intimate relationship with whom has certainly relieved me of the stress of living as an alien, although not necessarily in a way to contribute to my thesis. Also the frequent e-mails from my alumni, a source of long-lasting friendship, have constantly amused and consoled the extreme boredom in the course of thesis writing and urged me to accelerate my work. Thanks.
To my family

&

someone who got lost

on the way to me...
## Contents

1 Introduction 1

1.1 Imaging Radar Principle ................................................................. 1
  1.1.1 General Theory ................................................................. 1
  1.1.2 Synthetic Aperture Radar .................................................... 2

1.2 Historical Background ................................................................. 4
  1.2.1 Remote Sensing ................................................................. 4
  1.2.2 Waveform Design ............................................................... 4

1.3 Objective ..................................................................................... 5

1.4 Contribution ................................................................................ 6
  1.4.1 Analysis of Phase Codes ..................................................... 6
  1.4.2 An Optimal Uniform Sidelobe Generation .......................... 6
  1.4.3 New Concept of Sidelobe Reduction .................................... 6
  1.4.4 Complex SAR Environment Analysis .................................. 7

1.5 Structure and Contents of Thesis ................................................ 7

2 SYNTHETIC APERTURE RADAR IMAGES 10

2.1 Radar Remote Sensing ................................................................. 10
  2.1.1 Surface Interaction ............................................................. 10
  2.1.2 Scattering ......................................................................... 13

2.2 SAR Theory ................................................................................ 13
  2.2.1 Principle of Operation ....................................................... 14

2.3 Ambiguity ................................................................................... 19
  2.3.1 Range Ambiguity ............................................................... 20
  2.3.2 Azimuth Ambiguity ............................................................ 22

2.4 Information in SAR Images ......................................................... 24
  2.4.1 Target Distribution ............................................................. 24

3 ELECTROMAGNETIC PHENOMENA in SAR 28

3.1 Geometrical Theory of Diffraction ............................................... 29
  3.1.1 Basic Theory ..................................................................... 29
  3.1.2 Geometrical Optics ............................................................ 30
  3.1.3 Diffraction rays ................................................................. 34
CONTENTS

3.2 SAR Environment Approach ................................................................. 38
  3.2.1 Wedge Structure Scattering ................................................................. 39
  3.2.2 Reflectivity and View Angle ................................................................. 40
3.3 Urban Images ............................................................................................. 42
  3.3.1 Understanding of Urban Radar Response ................................................ 43
  3.3.2 Corner Reflector ........................................................................................ 45
3.4 Urban Area Modelling ............................................................................... 49
  3.4.1 Simulation Strategy .................................................................................. 49
4 WAVEFORM DESIGN ................................................................................... 51
  4.1 Pulse Compression ....................................................................................... 51
    4.1.1 Optimal Matched Filter ........................................................................... 52
    4.1.2 Performance Assessment ........................................................................ 54
    4.1.3 Ambiguity Function ................................................................................ 56
  4.2 Linear FM ...................................................................................................... 57
    4.2.1 Implementation ......................................................................................... 58
    4.2.2 Frequency Spectrum ............................................................................... 59
  4.3 Sidelobe Suppression .................................................................................... 62
    4.3.1 Source of Sidelobes ............................................................................... 62
    4.3.2 Weighting Schemes ............................................................................... 63
    4.3.3 Non-Linear FM ....................................................................................... 67
  4.4 Discrete Coded Waveforms ......................................................................... 67
    4.4.1 Barker Code ............................................................................................ 68
    4.4.2 Maximal Length Sequences Code ............................................................ 69
    4.4.3 Frank Code ............................................................................................... 71
    4.4.4 Huffman Code ......................................................................................... 72
  4.5 Waveform Design for SAR ........................................................................ 72
    4.5.1 Ambiguity Rejection ............................................................................... 73
    4.5.2 Ideal Response Performance ................................................................... 74
5 ANALYSIS OF PHASE CODED WAVEFORM ............................................... 76
  5.1 P-Code Theory ............................................................................................ 76
    5.1.1 P1 and P2 Polyphase Codes ................................................................. 77
    5.1.2 P3 and P4 Polyphase Codes ................................................................. 78
  5.2 Doppler Characteristic ............................................................................... 79
  5.3 Frequency Spectrum and Bandwidth Limitation ........................................ 80
  5.4 ACF of P code .............................................................................................. 83
    5.4.1 Sidelobe Analysis of P code ................................................................... 83
    5.4.2 Sidelobe Equation ................................................................................... 84
  5.5 Propositions and Validations ...................................................................... 86
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.6</td>
<td>Range Sidelobe Reduction</td>
<td>93</td>
</tr>
<tr>
<td>5.6.1</td>
<td>Sidelobe suppression analysis</td>
<td>94</td>
</tr>
<tr>
<td>5.6.2</td>
<td>Finite Quantization Effect</td>
<td>95</td>
</tr>
<tr>
<td>5.6.3</td>
<td>Resolution Loss Compensation</td>
<td>96</td>
</tr>
<tr>
<td>5.6.4</td>
<td>Cross-Correlation Performance</td>
<td>99</td>
</tr>
<tr>
<td>5.7</td>
<td>Comparison with Linear FM</td>
<td>101</td>
</tr>
<tr>
<td>5.8</td>
<td>Digital Pulse Compression Implementation</td>
<td>102</td>
</tr>
<tr>
<td>6</td>
<td>Modified Matched Filter for An Optimal Uniform Sidelobe</td>
<td>105</td>
</tr>
<tr>
<td>6.1</td>
<td>New Matched Filter Scheme without Subtractor</td>
<td>106</td>
</tr>
<tr>
<td>6.2</td>
<td>Woo Filter</td>
<td>107</td>
</tr>
<tr>
<td>6.2.1</td>
<td>New Phase Code Definition</td>
<td>107</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Justifying Formulation and Approximation</td>
<td>112</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Performance Analysis</td>
<td>114</td>
</tr>
<tr>
<td>6.2.4</td>
<td>Doppler Shift Effect</td>
<td>116</td>
</tr>
<tr>
<td>6.3</td>
<td>Exact Time Domain Formulation of Woo filter Response</td>
<td>119</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Derivation of the Doppler Shift Effect</td>
<td>119</td>
</tr>
<tr>
<td>6.4</td>
<td>Frequency Spectrum</td>
<td>120</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Bandwidth Limiting</td>
<td>121</td>
</tr>
<tr>
<td>6.5</td>
<td>Mismatching SNR Loss</td>
<td>121</td>
</tr>
<tr>
<td>6.6</td>
<td>Performance Enhancement</td>
<td>124</td>
</tr>
<tr>
<td>6.6.1</td>
<td>Resolution Compensation Strategy</td>
<td>124</td>
</tr>
<tr>
<td>6.6.2</td>
<td>Weighting Approximation Curve</td>
<td>125</td>
</tr>
<tr>
<td>6.6.3</td>
<td>Peak Ripple Removal</td>
<td>127</td>
</tr>
<tr>
<td>6.7</td>
<td>Non-Perfect Phase Code</td>
<td>127</td>
</tr>
<tr>
<td>6.7.1</td>
<td>Phase Error</td>
<td>128</td>
</tr>
<tr>
<td>6.7.2</td>
<td>Random Noise Effect</td>
<td>129</td>
</tr>
<tr>
<td>6.8</td>
<td>Sampling Effect Consideration</td>
<td>130</td>
</tr>
<tr>
<td>6.8.1</td>
<td>Error Sources</td>
<td>130</td>
</tr>
<tr>
<td>6.8.2</td>
<td>Time Mismatching</td>
<td>131</td>
</tr>
<tr>
<td>6.8.3</td>
<td>Quantization</td>
<td>132</td>
</tr>
<tr>
<td>6.9</td>
<td>Comparison with Linear FM</td>
<td>134</td>
</tr>
<tr>
<td>6.9.1</td>
<td>Weighting function</td>
<td>134</td>
</tr>
<tr>
<td>7</td>
<td>Integrated Sidelobe Energy Reduction</td>
<td>136</td>
</tr>
<tr>
<td>7.1</td>
<td>Integrated Sidelobes</td>
<td>136</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Point Target Discrimination</td>
<td>138</td>
</tr>
<tr>
<td>7.1.2</td>
<td>Integrated Sidelobe Level</td>
<td>139</td>
</tr>
<tr>
<td>7.2</td>
<td>CLEAN Technique Application</td>
<td>141</td>
</tr>
<tr>
<td>7.2.1</td>
<td>CLEAN Process for SAR Image</td>
<td>143</td>
</tr>
</tbody>
</table>
7.3 Sidelobe Cancellation ................................. 146
  7.3.1 Complementary Sequence ...................... 146
  7.3.2 Wide-Swath Application ....................... 148
7.4 Zero Level Sidelobe Energy Reduction Using Woo filter ................................. 150
  7.4.1 Sidelobe Cancellation scheme .................. 150
  7.4.2 Schematic Diagram for Zero Sidelobe Implementation .............................. 156
  7.4.3 Reduction of Sidelobe Energy .................. 157
  7.4.4 Application and Doppler Shift Effect ........ 160
  7.4.5 Woo Amplitude Weighting Cases ............... 161

8 SAR IMAGE SIMULATION ............................. 163
  8.1 Urban Area Image Generation ..................... 163
    8.1.1 Geometry Understanding ...................... 163
    8.1.2 Modeling ......................................... 169
  8.2 Radar Target Response .............................. 173
    8.2.1 Convergence Test on A Simple Structure .... 173
    8.2.2 Look Direction .................................. 173
    8.2.3 Incident Angle .................................. 176
  8.3 Target Response Dynamic Range .................. 177
    8.3.1 Effects on response power level ............. 178
  8.4 SAR Simulation ........................................ 178
    8.4.1 Waveform Performance ....................... 179
    8.4.2 Azimuth Compression ......................... 180
    8.4.3 Sidelobe Energy Reduction ................... 183
    8.4.4 Target Response ................................ 186

9 Conclusion ................................................. 190
  9.1 Summary of Thesis .................................... 190
  9.2 Suggestions for Future Work ...................... 195
List of Figures

1.1 SAR image example - Pusan, South Korea; by SIR-C/X-SAR, 1994  

2.1 Block diagram of typical SAR subsystem illustrating key assemblies  
2.2 Three material properties that affect radar interaction with the surface.  
2.3 Range direction data acquisition procedure corresponding to different radial 
    distances from radar platform. 
2.4 Illustration of Synthetic Aperture Radar formation and the procedure of 
    azimuth resolution improvement  
2.5 Timing diagram of transmitted and received signals  
2.6 Range Ambiguity generation diagram  
2.7 SAR azimuth ambiguity  
2.8 Typical SAR image and its reflectivity amplitude distribution graphs for 
    i) entire domain and ii) dominant distribution  
2.9 SAR image example and its reflectivity amplitude distribution graphs and 
    comparison with known distribution functions 1) Gamma 2) Rayleigh distri-
    bution  
3.1 Typical scattering problem between simple source and target objects  
3.2 Reflection and refraction between two lossy media along planar boundary  
3.3 Propagation modeling along the axial ray and quantitative diagram of energy 
    flow  
3.4 Plane wave propagation fields including diffraction field around half plane  
3.5 Scattered fields around half-plane conductor at $r = 5$ generated by plane 
    wave: (a) Diffraction field $E_d$ magnitude, (b) Scattered field ($E_d + E_r$) and 
    comparison with total field  
3.6 (a) Diffraction field at Wedge structure (b) Diffraction field at 'Vertex 
    point', showing it is analogous to point source  
3.7 ERS-1 SAR raw data after range direction processing is done. (a) A range 
    processed image containing a dominantly bright target. (b) Azimuth direc-
    tion amplitude graph along the dominant bright response  
3.8 SAR image amplitude distribution-Rayleigh and Gamma- analysis and com-
    parison between rural and urban scenes
3.9 Classification of radar backscattering models within urban environment. (a) A simple specular reflection on plane surface with arbitrary smoothness levels; dark image, dark-medium grey, bright grey (b) Combination of multiple reflection and diffractions (c) Strong specular reflection from man-made, cavity structure; Very strong return and exceptionally bright image.


3.11 Typical views of metropolitan area within modern cities (a) Inner area of Sanfrancisco, USA (b) Skyscrapers inside city, USA.

3.12 The effect of diffracted field on wave propagation and its comparison with direct fields. (a) The role of diffraction field in shadow region without presence of any other direct fields (b) Three major backscattering paths via urban structure (c) Quantitative comparison with specular return fields.

3.13 A description of the procedure to analyze the relationship between target structure of corner reflector type and its reflectivity level in SAR environment.

4.1 Pulse compression concept illustration and waveform parameters.

4.2 Graphical description of linear FM ambiguity function and energy distribution on time-frequency shift domain.

4.3 Frequency spectra of linear FM signals having various time-bandwidth products $BT$.

4.4 An example of frequency spectrum in ERS-2 SAR data. The bandwidth of the transmitted Linear FM signal is 15.5 MHz with the time duration of $37.12 \mu s$.

4.5 SAR raw data delivered by a single pulse and its range processed output.

4.6 Sidelobe suppression mechanism. Signal response ripples are cancelled out by leading and lagging echoes.

4.7 Magnitude comparison of Dolph-Chevyshev, Hamming, Han and Blackman weighting functions in frequency domain.

4.8 Pulse compression outputs after various spectrum weighings are applied. The PSL for Dolph-Chevyshev is -40dB. 4 term Blackman window parameters are $a_0=0.35875$, $a_1=0.48829$, $a_2=0.14128$, $a_3=0.01168$.

4.9 (a) Autocorrelation function of Barker code when $N=13$ (b) Ambiguity function plot for Doppler shift range $-0.05B < f_d < +0.05B$.

4.10 M bits shift register and feedback algorithm to generate MLS codes.

4.11 Pulse compression results from MSL. (a) Autocorrelation functions for MLS with length 511 (b) Cross-correlation function for two different MLS codes.
LIST OF FIGURES

4.12 Frank code properties (a) Phase variation diagram of the Frank code for N=5 and N=10 (b) ACF responses of Frank code with and without Doppler shift for N=20 ............................................................... 72

4.13 Ideal sidelobe responses for SAR application in dense area. (a) A superior performance of uniform sidelobe over spurious one having multiple local peaks. (b) Ideal mainlobe isolation .............................................................................. 74

5.1 Phase variations and ACF patterns for P3 and P4 code (N=100) ............... 79

5.2 The ACF outputs for P1 and P3 codes with 5% Doppler frequency shift of signal bandwidth ............................................................ 80

5.3 Power spectral density of P2 and P3 phase coded pulses ........................... 82

5.4 Effect of bandlimiting on the waveform weighting ....................................... 82

5.5 Sampling rate effects on pulse compression output of linear FM ............... 84

5.6 Comparison of the sidelobes between linear FM and P code ....................... 86

5.7 Generation of ACF(q) graphs ........................................................................ 88

5.8 Sidelobe magnitude comparison between two phase coded signals with different code lengths ................................................................. 92

5.9 Peak sidelobe variation with respect to $S_r$ in the interval $[BT, 2BT]$ ........ 92

5.10 Pulse compressor with subtractor ................................................................ 93

5.11 P3 compressed pulse with two types of subtractors .................................... 94

5.12 Comparison of approximately calculated sidelobe level with actual result when P3 code is compressed using sidelobe suppressing processor. The time-bandwidth product $BT=1000.$ .................................................. 95

5.13 Four different cases of ACF of the P3 phase code processed with various quantization levels ............................................................... 97

5.14 Resolution degradation during the pulse compression and sidelobe reduction at postprocessing for the P3 code ........................................... 97

5.15 Modified block diagram for range loss compensation ............................... 98

5.16 A pulse compression output of P3 code with two samples average subtraction technique and its 2D Ambiguity function diagram .................. 99

5.17 Cross correlation function between (a)P3 and P4 codes, (b)p3 and Linear FM codes ........................................................................... 100

5.18 Cross-correlation property of linear FM class waveforms. (a) Peak values of cross-correlation function between opposite sloped linear FM, with respect to pulse compression ratio. (b) CCF between P4 and reversed P4 codes, with(solid) and without adder(dashed). N=BT=706. .......................... 101

5.19 Test of subtractor application to the conventional linear FM. Linear FM ACF response and after subtractor is used in conjunction with the matched filter with different sampling rate $S_r.$ ........................... 102
5.20 Digital pulse generation and compression procedure block diagram of arbitrary waveforms. (a) Linear FM waveform generation in digital format. (b) Digital matched filter, which receives 12-b I/Q samples as input signal and produces 16 bit compressed outputs [4]. ........................................ 103

6.1 Diagrammatic representation of the algorithm of how to incorporate the post-compression part inside the matched filter ........................................ 106
6.2 Incorporated filter characteristic and generated pulse compression output. . 107
6.3 Block diagram on the correlation between the received signal and a new composite filter ................................................................. 110
6.4 Amplitude and Phase angle variation along time axis of Woo filter for (a) s-P3 code (b) s-P4 code. Each of the codes have phase code length of 200 (pulse compression ratio BT=200) ................................................ 111
6.5 The real and imaginary terms comparison between actual correlation output and approximate forms ................................. 113
6.6 Verification of Woo filter output for a s-P4 signal with the code length N=100 (a) $\Omega_1(q)$, $\Omega_2(q)$ and their summation $\Omega_2(q) + \Omega_1(q)$ (b) A performance improvement by the Woo filter and the error level caused by approximate formulation. ............................. 114
6.7 s-P4 code pulse compression outputs by the Woo filters for code lengths N = 200 and 1000. ................................................................. 115
6.8 Illustration of the mainlobe isolation distance with respect to the pulse code length N. ................................................................. 115
6.9 Mainlobe comparison between s-P3 and s-P4 correlation cases. .......... 116
6.10 Woo filter outputs for Doppler shifted signals. Doppler shift $f_d = 1, 2, 3$ and 4 % of the total signal bandwidth $B_d$. Pulse code length N=100. .. 117
6.11 Ambiguity function plot for $N = 100$ and Doppler shift ..................... 118
6.12 The effects of Doppler shift on pulse compression output in range direction shift and peak mainlobe response ............................. 118
6.13 Doppler shift effect on Woo filter outputs in the range direction shift. $f_d = [-0.02B_1, +0.02B_1]$. (a) One sample per code element (b) After oversampling of 2 is imposed. Comparison with a plain linear FM case. ............... 118
6.14 The effect of Doppler shift on the sub-PSL(peak ripple value) ............. 119
6.15 Frequency spectrum diagrams of the signals used in the Woo filter compres­sion on various code length N(100, 500, 1000, 2000) ................................. 121
6.16 Frequency spectrums of Woo filters for different pulse compression ratios . 123
6.17 The graphical relationship between SNR loss and pulse compression ratio . 123
6.18 Phase comparison of asymmetric Woo filter with the original cases of phase code length 100 for (a) s-P3 waveform and (b) s-P4 waveform. (c) Pulse compression output using asymmetric s-P4 code of length N=576. .......................... 126
6.19 The pulse compression outputs generated by asymmetric Woo amplitude weightings and comparison with original Woo filter cases. ................................. 126
6.20 Reduction of sidelobe peak ripples through adjustment of Woo filter elements .............................. 127
6.21 Impact of arbitrary phase errors on uniform sidelobe property. ........................................ 128
6.22 The effects of random phase and amplitude errors on compressed pulse output ................................................................. 129
6.23 Peak response levels and peak to sidelobe ratio variations for Woo filter and matched filter cases corresponding to increased noise levels. Pulse code length is 200. ................................................................. 130
6.24 Misalignment effects on the phase pulse compression output using Woo filter ................................. 131
6.25 Comparison of the filter outputs for quantized phase codes. Number of $Q$ are 7, 9, 11 bits for phase code length $N=573$. ....................................................... 132
6.26 Comparison of magnitude quantized Woo filter compression performances. Number of $Q_{m, mag}$ are 5, 7, 8 bits for phase code length $N=573$. ........................................ 133
6.27 Woo filter magnitude and phase quantization error effects on PSR. Phase code lengths are given as 100, 300, 500 and 700. Comparison of (a) Phase quantized Woo filter and (b) Magnitude quantized Woo filter compression performances ......................................................... 133
6.28 Linear FM and sin weighted FM signal compression outputs ........................................ 134

7.1 Two dimensional sidelobe generation algorithm SAR processing ........................................ 137
7.2 Typical example of SAR image with a very strong reflection occurring from a human-made target ................................................................. 138
7.3 Point target response of SAR image using linear FM signal .................................................. 140
7.4 2D ISL level graphs with respect to the radius relative to 3dB distance ........................................ 141
7.5 CLEAN technique algorithm for one dimensional simple target detection. Profiles of target responses from three different scatterers ........................................ 142
7.6 Magnified point target response of SAR image using linear FM signal ...................................... 143
7.7 ERS-1 raw data after range direction compression with particularly bright tone in the azimuth direction ................................................................. 143
7.8 Flow charts of the CLEAN procedure for SAR processing .................................................... 144
7.9 A amplitude fluctuation along the range direction after each iteration of CLEAN procedure. ................................................................. 144
7.10 Modified range compression output after the CLEAN procedure is completed ........................................ 145
7.11 The comparison of original SAR image with CLEAN processed image ...................................... 145
7.12 Sidelobe cancellation by a Golay complementary pair of length 8. (a) and (b) are ACFs of independent pairs and (c) is their sum ........................................ 147
7.13 Undesired effects in applying complementary codes. (a) Time-delayed ($T_b$) mis-matching effect (b) Cross-correlation energy accumulation ........................................ 148
LIST OF FIGURES

7.14 Wide swath operation using multiple waveforms. (a) Range ambiguity generation from the responses from A and B (b) Cross-correlation responses between linear FMs of opposite slopes for point and distributed targets. BT=100, 2000. .......................................................... 149

7.15 Schematic diagram of sidelobe cancelling scheme in general radar pulse compression system .......................................................... 151

7.16 Comparison between Woo sidelobe and conjugate of the incoming signal in magnitude and phase. Total code length N = 100. Sidelobe domain is bounded as q = [1 N]. .......................................................... 154

7.17 Complex domain comparison between Woo filter compression output and original waveform. .......................................................... 155

7.18 Sidelobe verification for second half sidelobe of s-P4 pulse compression output. N is set as 104=8×13. .......................................................... 155

7.19 Hardware implementation of sidelobe cancelling scheme. Separate I, Q detection for conjugate signal generation and incorporation of Woo filter and SC filter. .......................................................... 158

7.20 The pulse compression output for s-P3, s-P4 codes after sidelobe canceller is introduced. N is given as 1000 and 2000 for s-P3 and s-P4 respectively. .................................................. 158

7.21 ISLR comparison for three different pulse compressions cases of s-P3 code; normal ACF, Woo filter and Woo filter with sidelobe canceller. .................................................. 159

7.22 Sidelobe cancellation on either part of sidelobes only. (N=1000) ........................................ 160

7.23 Doppler shift effect on Woo sidelobe cancelling performance. N=576 and f = 1.15 KHz. .......................................................... 161

7.24 Sidelobe cancelling attempt in conjunction with Woo weighting scheme. N=576 and f = 1.15 KHz. .......................................................... 162

8.1 Bird-eye views of the inner-city area of central London. ........................................ 164

8.2 Range shift and target ambiguity in corner reflector structure between two symmetrically located points P1(W1, H1) and P2(W2, H2). ........................................ 165

8.3 Range shift effect illustrations by the corner reflector within a high or low resolution SAR image. .......................................................... 167

8.4 GTD method approach to the interpretation of the relationship between surface roughness and reflectivity. The surface dimension is randomly distributed with average height h. ........................................ 168

8.5 The returned signal energy (RCS level) from diffracted field. The incident angle is assumed to be 23°. ........................................ 168

8.6 The possible combinations of ray paths during interaction between radar signals and corner reflector formed by two walls and ground plane. A simplified geometry for ground buildings with smooth surface and bare soil ground. ........................................ 171
8.7  Four different cases for tracing the possible ray paths to explain interactions between radar signals and corner reflector formed by two walls and ground plane. Fig.(a),(b),(c) and (d) are mirror geometries devised for the purpose of understanding the complicate reflection mechanism and convenient computer simulation. ................................................................. 172

8.8  Ray tracing method test on corner reflector structure. RCS responses corresponding with respect to view angles. ................................................................. 174

8.9  The energy levels of returning signals with respect to view angles. The distances between two rectangular type walls are altered in the range \([0.1\sim0.5]\) of target width. ................................................................. 175

8.10  The effects of the ratio of building structures to the returning signal levels. For a given separation distance of 0.1W and 0.5W, the heights are altered as 2,3 4 times of target width. In (b) different heights are given for two adjacent building separated by 0.2W. ................................................................. 176

8.11  A comparison of return signal energy levels for various heights buildings.
   (a) D=W,θ = 18° (b) The height of the buildings are fixed as H=3W. ... 177

8.12  The comparison of target resolving performance of various waveforms. Two point targets with identical reflectivity of 60 dB are assumed to be separated by \(\frac{\Delta r}{c} = 2,3/\sqrt{f}\). ................................................................. 180

8.13  The point target response (PTR) of SAR signal processed with Woo filtering in both range and azimuth directions. Range and azimuth axis are expressed in distances \((t_{az},t_{ra}+c)\). ................................................................. 182

8.14  SAR image simulation of four point reflectors. ................................................................. 183

8.15  The point target response of SAR signal processed with Woo filtering and sidelobe reduction technique in both range and azimuth directions. \(T_{az}\) and \(T_{ra}\) are time durations of processed SAR signals in each directions. ................................................................. 184

8.16  2-D SAR simulations for various waveforms and processing methods. A very strong feature is assumed to be present on the dark background. ... 184

8.17  SAR images generated by the Hamming weighted linear FM and s-P code with Woo filtering. The targets have the reflectivities of 0, -20 and -53 dB. The code length of s-P code is given as 576. ................................................................. 187

8.18  ATR procedure for typical SAR images ................................................................. 188

8.19  Simulated SAR images of a tanklike object using different waveform schemes. ... 189

9.1  The Woo filter characteristics and the pulse compression output having an optimal uniform sidelobe. The code length \(N = 1000\) ................................................................. 193

9.2  The integrated sidelobe energy reduction. ................................................................. 194

9.3  SAR image response on the ideal condition that a full phase correction is implemented on the azimuth direction. ................................................................. 197
9.4 The magnitude and phase patterns of the cross-correlation function between linear FM derived phase code signals having opposite slopes. . . . . . . . . . . 198
List of Tables

4.1 Processed output parameters after weighting is implemented .................. 65
4.2 Barker Code Sequences ................................................................. 68

7.1 2-D energy ratio of sidelobe region to mainlobe ................................. 141
7.2 Woo sidelobe pattern with respect to time delay ................................. 151
7.3 Phase signal code values ................................................................. 152
7.4 ISLR comparison for various pulse compression of s-P3 codes. ............... 159
7.5 Parameters for ERS-1 ................................................................. 160
7.6 ISLR comparison when Doppler shift is present .................................. 161

8.1 Roughness of the typical urban features corresponding to various Satellite
    radar systems .................................................................................. 169
8.2 Possible combination of reflection order between rays and urban structure .. 170
8.3 Return signal energy ratio [dB] tables for W-(Width of the walls), D-
    (Distances of the walls from origin), H-(Height). The figures are relative
    ratio to the maximum amount of return energy for the given geometry. (*
    indicates zero response) .................................................................. 178
8.4 Azimuth ambiguity properties comparison .......................................... 181
8.5 ISL energy distributions within the point target response of SAR image .. 186
List of Symbols

- As appeared in the text,

\[ \lambda \] Signal wavelength
\[ \eta \] Radar incident angle
\[ h \] Surface roughness defined as the r.m.s of height distribution
\[ K_s \] Platform height
\[ \theta_i \] Incident angle
\[ \alpha \] Local terrain slope
\[ \langle N_{raw} \rangle \] Average raw data noise power
\[ K(N_0) \] Processor noise gain
\[ K(r) \] Cross-track radiometric correction vector
\[ r \] Range between source and scatterer
\[ T \] Signal time duration
\[ C \] Velocity of light
\[ \delta_{ra} \] Range resolution
\[ \delta_{az} \] Azimuth resolution
\[ B \] Signal bandwidth
\[ S(t) \] Signal in time domain
\[ f_0 \] Carrier frequency
\[ R \] Distance between antenna and target
\[ L_a, D \] Antenna length
\[ \theta_{3dB} \] Antenna 3-dB beam width
\[ V \] Platform velocity
\[ f_p \] Pulse repetition frequency (PRF)
\[ \phi(x) \] Phase as a function of distance x
\[ f_d \] Doppler shift
\[ t_d \] Time delay
\[ \Omega \] Correlation or pulse compression output
\[ L_a \] Synthetic aperture length
\[ R_E \] Earth radius
\[ \tau \] Pulse length
\[ \sigma \] Reflectivity
\[ G \] Cross-track antenna pattern
Signal spectrum

Pm  Mainlobe signal level
Ps  Integrated sidelobe level
Pps Peak sidelobe level
Bp  Processing bandwidth
Vst Relative platform to target velocity
ΔXra Range displacement
ΔZaz Azimuth displacement
fDc Doppler centre frequency
fR Doppler rate
δx, δR Focussed azimuth and range resolutions
PA Amplitude distribution
PI Intensity distribution
I Intensity
L Total number of looks
Γ Gamma function
E Electric field
H Magnetic field
vc Critical angle
vr Reflect angle
vi Incident angle
Ud Diffraction field
D Diffraction coefficient
K_(x) Modified Fresnel integral
Si Signal source
Pi Points that arbitrary rays hit
N0 Gaussian noise spectral density
Wf,t Frequency or time weighting function
χ(t,fd) Ambiguity function
E Energy
H(ω) System transfer function
A(ω),B(ω) Amplitude and phase terms of transfer function
Jn Nth order Bessel function
ψ Phase variation within phased code waveform
N Discrete signal code length
ϕi ith phase term
tc Compressed pulse length
Tb The time duration of individual element of phase code waveform
θn Nth phase element term
Sr  
C1,2  
Woo(t), Woo(n)  
\frac{S}{N}  
PSL  
ISL  
PTR  
(tR,tAz)  
(DR,Daz)  
\sigma_d  
\langle I \rangle  
Wi  
Hi  
Lp  
\theta  
\sigma_i  
\Gamma  
\Psi(rays)  
A_{eff}  

Sampling rate  
Cross-correlation function  
Woo filter in continuous or discrete domain  
Signal to noise ratio  
Peak sidelobe  
Integrated sidelobe  
Point target response  
Time domain location of target object  
Geometrical location of target  
Dynamic range of target response  
Mean intensity  
Distance between \(i\)th wall and the reference axis  
Height of \(i\)th wall  
Prolonged pulse length  
Antenna look angle  
Signal propagation loss  
Reflection coefficient  
Total number of rays involved in actual GTD calculation  
Ratio of effective area to overall wall area
Chapter 1

Introduction

Synthetic Aperture Radar (SAR) technology has continuously advanced during the last 40 years. Especially, with the progress of high speed computing power and large memory storage capacity available these days, it has brought a new era in the field of remote sensing of the Earth. SAR has been widely studied and understood in the past decades. Synthetic Aperture Radar's characteristic comes from its unique way of transmitting and processing complicate phase controlled signals which enable it to achieve a highly improved resolution with a relatively small antenna.

1.1 Imaging Radar Principle

1.1.1 General Theory

Radar is an active system which emits signal waveforms with particular characteristics through the transmit antenna, and receives those signals reflected off a distant surface or objects. The term 'active' implies significant advantages over passive systems which operate with visible or infrared light. While it is less easy to utilize passive systems at night or on a cloudy day, the capability of active radar systems is not affected by such natural factors. Moreover, by flexible control of the parameters of radar signals, such as frequency, polarization, wave shape, its application area can be further diversified. For an imaging radar system, a pulsed signal with finite duration is used to get range information of target objects. These signals cover a small band of frequencies centred on a particular frequency in the microwave region. In most cases, the image resolution in the range direction is dependent on the bandwidth or equivalently the pulse length. Radar images or pictures are composed of several small dots or point elements. What a radar system extracts from the received signal is actually the backscatter returns from the target object or area. They are converted into digital form to be stored on a data recorder and processed later. Each point element in the images corresponds to the unit area of target region and indicates the intensity of scattered signal measured by the radar. Bright points represent area with high reflection characteristics or RCS (Radar Cross Section), while a dark colour means very
1.1 Synthetic Aperture Radar

Weak signals are reflected back from the region. The intensity is not simply determined by the regional RCS characteristics, but is also a function of several variables such as radar view angles, the relative size of the object in comparison with the wavelength, polarization of the pulse, the smoothness of the surface and so on. Mountain areas or tall buildings standing vertically or inclined towards the line of view from the radar will appear with bright intensity, while the backscatter is relatively dark for flat road or sea surface. The backscatter is also sensitive to the electrical property of target region. Wetter regions or objects are apt to appear as bright targets and drier regions can be seen as dark area. Due to its flat surface, an oil layer spread over sea surface tends to reflect away most of the incoming radar signals so that it looks dark on the image.

1.1.2 Synthetic Aperture Radar

The azimuth resolution of radar is generally dependent on the beamwidth of the antenna. The resolution is the minimum distance between scatterers in the target area by which two different objects can be discriminated. In a practical radar system where an antenna is deployed at high altitude, the beamwidth is too wide to gain a desirable resolution due to the finite physical size of antenna. The basic principle of SAR is to overcome this physical limitation by realizing a virtual large antenna during the movement of radar antenna. Due to the similarity between the phase characteristic of the data collected during the sweeping and linear FM, the along track processing of SAR data is performed in a similar way to linear FM pulse compression. A great amount of memory and also time should be consumed, which makes real time processing difficult. In most cases, it is essential to correct the image data through additional processings in a ground station to produce high quality images. Although in theory, the SAR image resolution is entirely dependent on the antenna size, some restrictions on the minimum image resolution are applied to avoid ambiguities and image speckle arising from clutter.

Figure 1.1 shows a typical spaceborne SAR image. The area in the picture is around a big harbour called Pusan, which is positioned near the south-east coast of Korean peninsula. The image was acquired by SIR-C/X-SAR in 1994. The flight was heading in the south and the antenna illumination was to the left side on the image. The sea surface region is shown to be brighter than the river. This can be explained by the waves present on the sea surface while the river stream flows maintaining relatively calm surface. The strong reflectivity from various random scatterers also explains the bright region in the map, which matches to the inner city area composed of multiple complex structures. A number of white objects which are thought to be very large ships are easily recognized on the sea surface.
Figure 1.1: SAR image example - Pusan, South Korea; by SIR-C/X-SAR, 1994
1.2 Historical Background

1.2.1 Remote Sensing

Remote sensing can be described as 'an act of acquisition of information from distant target objects without direct contact from observer' [72]. Since the concept of 'RADAR'(RAdio Detection And Ranging) was first introduced early in this century as a simple method of ship detection in the sea using reflected radio waves, it has matured into highly advanced technology and contributed to opening of new era in remote sensing research field. Now, with the help of recent technology, the purpose of remote sensing has become diversified and is continuing to broaden its operational functions. After a number of SAR studies by ERIM and JPL, the first spaceborne SAR experiment was successfully carried out in mid '70s by NASA, when the earth observing satellite SAR system, SEASAT, was positioned into its space orbit in July 1978. Although its success was damaged by system power failure shortly after its launch, the various data collected during this mission has provided a good motivation for further mission development. Through the Shuttle Imaging Radar (SIR) series and ERS-1,2, JERS and RADARSAT, each of which has been launched with its own purpose, SAR technology has promised to deliver a wide range of geological information over the planet Earth. The microwave frequencies mainly used in SAR systems are highly resistant to weather condition or darkness so that they find their unique application in extreme regions such as polar or tropical zone.

The merit of SAR technology can be attributed to the unique method of processing the received signals. It is because of this complexity that SAR has become the most sophisticated radar technique in the remote sensing domain. Unlike other remote sensing systems using optical or infrared waves which are passive sources, SAR systems can set up specific operation objective by controlling various parameters of microwave source as intended. The primary objective of SEASAT was to image ocean areas, while RADARSAT has focused on arctic observation by mapping ice and northern regions where the use of optical sensor can hardly be favoured due to weather conditions. The ENVISAT-1, which will be launched in the middle of 2000 by ESA is aimed to provide extensive and enhanced opportunity to monitor the global environment. It is expected to deliver highly sophisticated information regarding various environmental parameters by use of improved microwave instruments including the ASAR (Advanced Synthetic Aperture Radar).

1.2.2 Waveform Design

The history of radar waveform design lies in parallel with the pulse compression technique. It has grown out of the need for improved performance during warfare before 1950. Since the peak power available in transmitter tubes was limited, a wider pulse length had to be considered to take advantage of greater average power, which resulted in unacceptable resolution loss. It was not until after the early 1950s that this idea was
brought into fruition with the development of the matched filter and pulse compression concepts. Since then linear FM has been a topic of active researches.

The P code group, first suggested by Lewis [51], belongs to the class of conventional phase codes. Since it is directly derived from linear FM, it shares the merits of linear FM while retaining the unique characteristic of phase code signals. It has been suggested that by employing additional post-compression processing, the sidelobe characteristic of this waveform may be further improved. However, despite its merit and possible potential to contribute to the radar community, very little effort has been made yet to analyze the theoretical aspect of the P code signal and the corresponding pulse compression scheme. Levanon [49] reported an article that compares linear FM and the P codes, but it was rather more concerned about the performance measures than the theoretical understanding.

1.3 Objective

This thesis aims at addressing the role of waveform design technique in SAR and developing new types of waveform that may suit some particular SAR applications. A good waveform normally implies efficient and accurate means to acquire and process the target information. A demand for highly efficient use of limited signal energy resulted in pulse compression techniques while a great deal of effort to reduce unwanted noise at pulse compression output has been motivated by the need of accurate measurement of targets with minimum information loss. However, since SAR is a remote sensing tool mapping wide Earth surface, the required degree of accuracy is more relaxed than other sophisticated target detection radars and the waveform design has not been so much emphasized as the processing technique of received signal. This thesis attempts to relate the waveform performance with the produced SAR image quality.

Much radar literature and researchers in this field classify waveforms largely as linear FM derived signals and phase coding signal group [15][48]. Having motivated the idea of pulse compression, linear FM has been thoroughly studied during several decades. The dominantly popular use of the linear FM scheme is attributed to the easy of implementation, but still better than required performance for many application areas. In comparison, phase coded group has received less attention, which may be explained by the complexity of implementation, high sensitivity to Doppler shift effects and lack of flexibility, while the performance is still not significantly greater compared with the linear FM. However the discrete structure of the phase code signal exhibits a possible potential to manipulate sidelobe patterns as desired, which promises a wide dynamic range response. It also indicates good compatibility with digital signal processing. In this thesis the use of phase code waveform for SAR application is investigated, where it may surpass the performance of conventional linear FM.
1.4 Contribution

1.4.1 Analysis of Phase Codes

As has been done for linear FM signals, this thesis attempts to derive an analytical formulation of the pulse compression procedure of the P code group. Recalling that an in-depth investigation into the sidelobe generation of linear FM compression has contributed to the sidelobe reduction scheme through weighting functions, a similar attempt is made concerning the P codes. Several propositions are constructed in regard to the properties of the P code compression output and their corresponding proofs are provided accordingly through rough mathematical formulations. A closed form expression is obtained to describe the sidelobe pattern at the pulse compression output.

A similar approach is taken to understand the idea of range-sidelobe reduction. Using closed form expressions, the principle of operation of the sidelobe reduction scheme is explained and verified through various simulations. After obtaining a general expression for the reduced sidelobe pattern at the post-compression stage it is shown that, contrary to the claim made by Lewis [50], the sidelobe level thus achieved fails to reach the optimal Barker code level.

1.4.2 An Optimal Uniform Sidelobe Generation

A slightly modified form of the P code is defined and shown to possess a unique sidelobe pattern similar to the P code case. Instead of adopting a post-compression processing which may not be a realistic and feasible scheme to be used in SAR applications, a new composite non-matched filter is developed which is generated by incorporating the post-processing part into the matched filter. Based on the sidelobe pattern formulation for the pulse compression output, a technique is suggested that enables us to produce an optimal uniform sidelobe pattern of the ideal Barker code level. It is claimed that the newly obtained sidelobe pattern provides an optimal compromise between resolution and peak sidelobe level, which has been believed to be obtainable only by the Barker codes. Unlike the Barker code of which the maximum realizable code length is 13, the new technique is not limited by the signal code length and retains the merit of strong resistance to the Doppler shift effect.

1.4.3 New Concept of Sidelobe Reduction

Although a dramatic improvement is achieved in terms of the peak sidelobe level, in some occasions where the distributed targets are mainly concerned instead of point target scatterers, the integrated sidelobe energy level (ISL) is considered as a more useful measure. Especially in SAR environment where target responses of wide dynamic ranges are distributed over broad area, peak sidelobe suppression alone may not be sufficient to experience a
true improvement in the image quality. However, it is verified that conventional integrated sidelobe cancellation schemes fail to work in the complicated SAR environment.

Based on the new filter concept a novel technique is proposed to reduce the amount of the integrated sidelobe energy. The widely stretching sidelobe pattern is considered as one of the disadvantage of the phase codes. By removing unwanted sidelobe originating from strong scatterers a dramatic reduction of integrated sidelobe energy is achieved. Without any loss of range resolution or peak sidelobe level, undesired interference between strong scatterers are fully suppressed.

1.4.4 Complex SAR Environment Analysis

Although it has been believed that an analytical interpretation of the electromagnetic phenomena during SAR operation is necessary for a better understanding of SAR images [30][75][29][10], very little attempt has been made yet in regard to this approach. In this thesis the role of electromagnetic phenomena in SAR environment is analyzed. Taking urban area structure as our main concern, some criteria that may affect the SAR imagery are examined. It is believed that the corner reflector feature plays a dominant role in generating extremely bright responses. Using Geometrical Theory of Diffraction (GTD), a high-frequency electromagnetic analysis technique, the target response is investigated and subsequently the need for wide dynamic response ranges are addressed.

It is claimed that, thanks to the arbitrary control of the sidelobe levels and a strong isolation of the mainlobe response, this new technique outperforms linear FM when a sophisticated target detection and recognition capability is required.

1.5 Structure and Contents of Thesis

This Ph.D thesis is essentially divided into nine Chapters, starting with brief introductory comments on the concept and historical backgrounds of synthetic aperture radar and waveform design at Chapter 1.

Chapter 2 provides a fundamental theory of SAR operation. A detailed analysis on how SAR achieves high resolution radar images is presented from the classical theoretical aspects. Some performance restrictions caused by the pulsed nature of SAR signals are explained. It is verified that the structures of the reflectivity levels of SAR images are described by some designated statistical distributions. It is shown that, on some occasions, the SAR image characteristic deviates from this trend when a heterogeneous area having a wide dynamic range of reflection levels is included within the image.

An analytical approach to relate the target characteristic and returned radar signal is useful for a better understanding of radar images but very often discouraged due to
the extremely complicated nature of the interaction between radar signals and randomly distributed ground targets. However it is known that many man-made structures found in urban areas respond in predictable manners and appear as dominantly bright features within SAR images. Chapter 3 explores a way to interpret this particular case of SAR image. A Geometrical Theory of Diffraction (GTD) theory is introduced as a tool to analyze the electromagnetic behaviours. It is shown that the wide dynamic range of response levels from urban area is the source to modify the SAR image characteristic and the need to separate this phenomenon is addressed.

Chapter 4 presents a comprehensive review of waveform designs starting with classical linear FM theory. At first, the concept of pulse compression is mentioned to introduce the background of the popular use of this waveform in radar applications. Several criteria that may be useful in assessing the waveform performance in relation to SAR are described and a theoretical review on improving the performance of pulse compression is made. As an alternative choice of waveform design, the phase code class waveforms are studied. Based on the performance evaluation of various waveforms, some suggestions are made on how to approach the waveform design that may produce an optimal performance in SAR application.

Chapter 5 provides a detailed analysis of the P-code, which is a relatively new phase code waveform design technique. After theoretical descriptions of these codes are provided, various simulations are performed to investigate the unique property of the P code which distinguishes it from others. The range-sidelobe reduction technique, which is uniquely applicable for this codes, is introduced. At first an in-depth investigation of this unique sidelobe characteristic is presented by taking pure mathematical approaches. Then the principle of sidelobe reduction scheme for the P code is fully explained. Based on thus obtained formulations, a theoretical basis for a further enhanced waveform design is provided.

Chapter 6 devotes to the development of new waveform and pulse compression filter set. It is shown that how a new filter concept, which is denoted as the 'Woo filter', produces a uniform sidelobe pattern that resembles the optimal SAR waveform performance suggested in Chapter 3. Firstly a waveform which is the modified form of the P-code, is defined and its sidelobe characteristic is obtained in the approximate formulations. From this is derived a conjecture in which a uniform sidelobe of Barker code level may be produced. Then a pair of phase code waveforms and the relevant filter is designed, and its performance is measured in various aspects. It is shown that the unique sidelobe characteristic is well preserved in the presence of various noise sources. After discussing that a major sacrifice is the broadened range resolution, a further modification is developed to overcome this.

Chapter 7 extends this new waveform design concept further to take into account its ISL property. It is discussed that when the distributed target is mainly concerned as is the
case of most SAR images, it is rather the integrated sidelobe energy than the peak sidelobe level that may affect final image quality. Some attempts to reduce ISL energy are made using conventional sidelobe cancellation techniques and a conclusion is drawn that they are not applicable to the SAR environment. A technique is suggested which, using the set of waveform and filter developed in Chapter 6, enables to reduce the integrated sidelobe energy while the PSL is still preserved at the Barker code level. After the principle of the ISL reduction process is fully explained, the feasibility of this scheme for practical SAR application is investigated.

Chapter 8 examines the practical application of the waveform set developed in this thesis. At first it presents several factors that may affect the urban SAR image characteristics. Based on the claim that many dominant features found in urban area image originate from corner reflection phenomena, it is attempted to formulate the electromagnetic interaction of radar signals with urban structures. A performance comparison is made between conventional linear FM and the new waveform set when targets having wide dynamic range of reflectivity levels are mixed together. It is argued that conventional linear FM compression fails to deliver a good understanding of SAR image while the new phase code successfully accommodates wide dynamic response and achieves a good target detection capability.

Chapter 9 summarizes the contents of the thesis and conclusions are drawn on the achievement of this research. Discussions are made on how to utilize the new waveform sets for SAR development. Some useful comments are suggested for further work concerning the role of waveform design and application in the SAR environment.

'The social system of science begins with the apprenticeship of the graduate student with a group of his peers and elders in the laboratory of a senior scientist. A formal publication is a formal invitation to criticism' -H. F. Judson (Science writer)
Chapter 2

SYNTHETIC APERTURE RADAR IMAGES

General SAR theory is introduced and explained in detail regarding the principle of operation to provide high resolution SAR images. Also the target distribution characteristic of SAR image is described.

2.1 Radar Remote Sensing

Satellite remote sensing is the use of sensors to collect and interpret signals containing information about the Earth's atmosphere, oceans and land surfaces [27]. Unlike optical and infrared imaging sensors which are passive, SAR uses active electromagnetic waves providing its own illumination on the target area. The transmitted signals travel via various paths through atmosphere and generate dynamic responses by interactions with ground surfaces. SAR is a measuring instrument providing estimates of the complex radar reflectivity or backscattering coefficients of the scene. Figure 2.1 shows a block diagram of the SAR sensor [16]. The enormous amount of raw data generated during SAR missions manifests the important role of signal analysis. In this sense, along with relevant microwave instruments, understanding physical electromagnetic phenomena is essential for accurate interpretation of the received signals.

When the radar signals interact with ground surfaces, they can either be reflected, scattered, absorbed or transmitted (refracted). Except for the case of complete absorption of the electromagnetic wave energy, all the other processes can contribute to the returned signal and affect the final image formation.

2.1.1 Surface Interaction

There are three basic properties that may affect radar interaction with the surface. The dielectric constant is a unique property of each material and determines the reflection
2.1 Surface Interaction

Figure 2.1: Block diagram of typical SAR subsystem illustrating key assembly behaviour. The other contributions come from the roughness and the local slope of the terrain surface.

1. Dielectric Constant: For a typical radar frequency, the complex dielectric constant $\varepsilon_r$ of most dry natural material is found in the range between 3 and 8, which is relatively low [30]. When interacting with this low $\varepsilon_r$ material, most of the EM wave energy is absorbed within deep penetration surface and the reflectivity level becomes very low. On the other hand, the dielectric constant for water is above 80 (where wavelength $\leq 50\mu m$) resulting in strong reflection. The reflectivity of natural material is linearly proportional to the moisture content per unit volume. The observed variation of the reflectivities for the flat vegetation area is mainly due to the moisture content. Hence the analysis of the brightness level of the generated image allows us to classify the type of soil or vegetation being imaged.

2. Roughness: Surface roughness determines the surface scattering and is often the main factor influencing radar return. There is a clear distinction on the nature of reflection between vegetation and non-vegetation area. The reflectivity of natural vegetation area is not easily predictable while non-vegetated smooth surface such as bare soil, water, glass and concrete materials having relatively smooth surface reflect EM wave according to the Snell’s law. A smooth level surface with high dielectric constant will be seen as black in the SAR images. Surface magnitude variation on the order of radar wavelength will scatter radar signals as specified by Rayleigh

---

1. EM spectrum and optical wave behave in different manner depending on the surfaces. Both of them have similar reflectances in non-vegetation surface. But in vegetation area, EM wave loses significant portion of total energy while optical wave still maintains high reflectivity. This difference can encourage the use of composite image derived from both data sets.

2. The detailed Snell’s law will be dealt with in Chap. 3.
2.1 Surface Interaction

Figure 2.2: Three material properties that affect radar interaction with the surface.

criterion, which is given by

\[
h < \frac{1}{k} \cdot \frac{\lambda}{\sin(\eta)}
\]  

(2.1)

where \( \lambda \) is the wavelength and \( \eta \) is the radar incident angle. \( h \) is the surface roughness defined as the root-mean-square of the heights distribution. The Rayleigh criterion sets the constant \( k \) as 8 but can be a finer value as necessary. For the case of ERS-1 operating at the frequency of 5.3 GHz with the incident angle of 23\(^\circ\), the Rayleigh criterion for smooth surface is given as 1.8 cm and any smooth surface with r.m.s. height less than this value will appear as solid black.

3. Terrain Surface: The geometry between incoming signal and terrain slope can determine the amount of reflection level. Even if the surface is not smooth, the perpendicular approach of the signal toward the terrain will produce strong return resulting in bright spot in the image. Thus two surface areas with identical material but different terrain slope appear in different manners. A very common case is found on mountain images. Figure 2.2(c) illustrates a typical example of the mountain area imaging. While one side is brightly illuminated with the help of increased look angle, the other side is hidden in the radar shadow or reflects very weak signals depending on the slope angle. Since radar range measurement relies on the arrival time, large topographic variation can cause confusion in extracting range information.

The geometrical distortion as well as surface property contributes to the received signal power. The relationship between processed pixel power \(< I >\) and actual target
backscatter coefficient $\sigma^0$ is given

$$< I >= K_s \sin(\theta_i) \cdot \frac{\sigma^0}{\sin(\theta_i - \alpha)} + < N_{raw} > \cdot K(N_o) \cdot K_N(r)$$ (2.2)

where $K_s$ is a calibration constant, $\theta_i$ is the nominal incidence angle on the reference ellipsoid corrected by average terrain height and $\alpha$ is the local terrain slope. The second term represents noise effect, where $< N_{raw} >$ is the average raw data noise power, $K(N_o)$ is a processor noise gain and $K_N(r)$ is a cross-track radiometric correction vector given as a function of range $r$. This relationship is commonly utilized for interpreting processed SAR images.

### 2.1.2 Scattering

Any return from a smooth level surface are due to specular scattering where Snell's law strictly rules the movement of the signals and can exist only near vertical incidence [67]. However, arbitrary reflected signal can reflect off other objects and redirected toward receiver increasing the reflection level. It often occurs in vegetation area or man-made structures in urban area. Especially when the vegetation area is wet, EM waves repeat reflections without significant energy loss until they exit. The exit directions will be randomly distributed but even a small percentage of return into receiver will result in bright image level. In general man-made structures consist of smooth materials with high dielectric constants and often forms corner-reflector shapes. Bridge is a good example to have strong reflection. Well arranged urban buildings are typical cases of corner reflector formations and appear as extremely bright targets in the processed image. To understand these electromagnetic interaction mechanism is useful for full understanding and interpreting SAR images.

### 2.2 SAR Theory

Synthetic Aperture Radar (SAR) is an imaging radar technique involving the use of an aircraft or satellite borne antenna to obtain an artificial radar aperture effect by utilizing the forward motion of the vehicle. The aim of SAR processing is to achieve a fine resolution along azimuth direction while range direction resolution is realized through waveform pulse compression. Although the methodologies implemented in range and azimuth direction processings are clearly distinguished from each other, both of them are derived from similar concept-'To spread relevant information widely in time or frequency domains and later rearrange those scattered information so that unique information for relevant area is maximized at corresponding image pixels'.
2.2 Principle of Operation

High resolution is achieved in both range and azimuth directions using pulse compression and azimuth SAR processing respectively.

### Range Resolution

Since the range resolution $\delta_r$ is inversely proportional to the signal time duration $T$ ($\therefore \delta_r = \frac{CT}{2}$), it is desirable to reduce pulse length for high resolution images. On the other hand, the maximum signal intensity is physically limited by the RF power generation technology at the transmitter making long pulse more attractive for a good target detection performance. SAR range processing is aimed at overcoming this trade-off allowing the use of long signal pulse whilst keeping high resolution criterion by means of pulse compression. Figure 2.3 illustrates the procedure of SAR range processing.

For range direction, the information spreading is implemented by sending out long pulse signals on regular intervals, which is restricted by PRF (Pulse Repetition Frequency) to avoid range or azimuth ambiguities. Each target cells within the antenna beam are exposed to the incoming signals during the finite pulse time duration. The location of reflected signals in the time domain are determined by the distances from the antenna. Thus all the target objects residing in the same radial distance from the antenna will be grouped together and treated as a single scatterer during range processing. The separation inside the same range group is done later by azimuth processing. To separate overlapped signals corresponding to each radial distance, the pulse compression technique is introduced. Most of the conventional SAR systems adopt linear FM (Chirp) signal for this purpose due to
its simple but reliable performance. A detailed description on pulse compression theory is treated in Chapter 4. The basic idea is to spread signal of finite bandwidth $B$ intentionally over long time interval $T$ to be compressed later into a short pulse of approximately $1/B$ time-width. Typical expression for linear FM signal is given in complex notation as

$$S(t) = \exp \left\{ j2\pi \left( f_0 t + \frac{B}{2T} t^2 \right) \right\}, \quad \text{for } |t| \leq \frac{T}{2}$$

(2.3)

where $f_0$ is the carrier frequency. The instantaneous frequency versus time changes linearly over the interval $[-T/2 < t < T/2]$. Range processing involves stripping off the carrier frequency and correlating with matched filter. The resulting output has sinc type function. It can be shown that conventional linear FM signal without weighting provides range resolution of approximately $\delta_{RA} = 1/B$ regardless of pulse length. Hence a strong target response is obtained with peak power limited transmitter by expanding waveform pulse duration. A comprehensive analysis on linear FM compression is presented in Chapter 4.

**Azimuth Resolution**

Two targets on the ground separated by $\delta$ along azimuth direction can be resolved only when they are not within the same beam-width at the same time. Thus the minimal azimuth resolution $\delta_{Az}$ is given as

$$\delta_{Az} = R\theta_{beam.width} = R \frac{\lambda}{L_a}$$

(2.4)

where $L_a$ and $\lambda$ represent the antenna length and signal wavelength respectively. In spaceborne radar this will yield resolution of the order of $km$, which is not acceptable. Since the azimuth resolution is inversely proportional to the antenna length, a mere extension of antenna size may improve resolution performance. One of the easiest way to understand the principle operation of SAR is to introduce a virtual large aperture antenna. Instead of building a large antenna on the platform, which is not a practical alternative, SAR sets up a virtual huge antenna along its path. The element spacing $d_A$ of the synthesized antenna is equal to the distance travelled by platform during pulse repetition interval or $d_A = V/f_p(V$ is the platform velocity and $f_p$ is PRF). On the condition that the target area is static, SAR behaves as if it is a part of the large virtual aperture antenna on each moment it transmits and receives pulsed waveform signals. After recording all the signals at each aperture positions, the collected data may be treated as obtained from a conventional aperture radar and processed accordingly.

One the other hand, the azimuth direction processing can be similarly understood analogous to range processing case. The fundamental difference lies in the methodology of signal generation and its collection on return from target scene. A point scatterer on

---

3Stop and Start operation: Although the antenna platform keeps moving during the interval between transmission and receipt of the signal, it can be assumed as static object without significant error [11].
2.2 Principle of Operation

Iso trop ic a l D o p p le r shift lines

(i) Doppler shift signal acquisition

(ii) Doppler shift signal acquisition

Fig. 2.4: Illustration of Synthetic Aperture Radar formation and the procedure of azimuth resolution improvement

The ground can return a signal back only while the target lies within the footprint of the antenna beam and interacts with the transmitted signal. These target reflections are sampled according to the PRF rate. In each sampling procedure the signals travel in different paths of different distances yielding continuous phase variation. After analyzing the phase behaviours, an appropriate correlation filter can be constructed to generate maximum SNR responses at the output. Two targets separated by a certain distance along azimuth direction have different Doppler shift, hence they appear in different positions at the output. The azimuth resolution is established by the 3-dB width of the correlation output produced. There are several ways to derive SAR azimuth resolution and here are described three different approaches.

(i) Virtual Synthetic Aperture Antenna: As mentioned, SAR achieves high resolutions by pretending to be an element of one large array. The virtually synthesized antenna has the length equivalent to the beam width given as

\[ L_A = R \theta_{MB} = R \frac{\lambda}{L_a} \]  \hspace{1cm} (2.5)

Treating this virtual antenna as a conventional aperture antenna, the resolution may be obtained by its 3 dB beamwidth. Since the original antenna pattern is used both on transmit and receive, the net beamwidth is halved. Therefore, the virtual synthetic antenna having equivalent aperture length of \( \lambda / L_a \) provides the resolution of

\[ \delta_{az} = R \frac{1}{2} \left( \frac{\lambda}{R \lambda / L_a} \right) = \frac{L_a}{2} \]  \hspace{1cm} (2.6)
2.2 Principle of Operation

which is the half of the original antenna length.

(ii) Equivalence to linear FM pulse compression: In Figure 2.4(a), the instantaneous distance \( R(x) \) between the radar and targets can be approximated\(^4\) as

\[
R(x) = \sqrt{R(x_0)^2 + x^2} \cong R(x_0) + \frac{x^2}{2R(x_0)} \tag{2.7}
\]

where \( R(x_0) \) is the closest range and hereby the phase variation experienced while traveling the path is given

\[
\phi(x) = \frac{4\pi}{\lambda} \left( R(x_0) + \frac{x^2}{2R(x_0)} \right) \tag{2.8}
\]

The expression in Eqn. (2.8) is shared by the linear FM phase term in Eqn. (2.3) when \( x \) is replaced with the equivalent transformation term \( t \). Instead of time-frequency bandwidth, the parameter relationship between \( x \) and \( \phi(x) \) can be described by introducing spatial-frequency bandwidth concept. Then the frequency variation within aperture length becomes

\[
f_d(x) = \frac{1}{2\pi} \frac{d\phi(x)}{dx} = \frac{2}{\lambda R(x_0)} x \tag{2.9}
\]

The maximum distance \( x \) is equal to the synthetic aperture length or original antenna beamwidth. Similar to linear FM pulse compression case, the minimal spatial resolution can be calculated from the reciprocal of the frequency bandwidth. Hence

\[
\delta_{az} = 1/\max[f_d(x)] = \frac{\lambda R(x_0)}{2} \frac{1}{L_A} = \frac{\lambda R(x_0)}{2} \frac{L_a}{\lambda R(x_0)} = \frac{L_a}{2} \tag{2.10}
\]

is obtained, which is identical to Eqn. (2.6)

(iii) Analytical approach: Although they are easy and straightforward descriptions on the concept of SAR, the analysis in (i) and (ii) do not fully address the principle SAR operation. SAR processing requires to take coherent integration of collected signals where target information is spread over finite time duration. The received signal is a pulse train of time delayed pulse \( S(t-t_d) \), where \( t_d = \frac{2R(x)}{c} \). In ERS-1 case, the slant range is 850km and the corresponding the time delay is \( 2 \times 850km/c = 0.056[s] \).

Since the PRF is 1680 Hz, actually received signal corresponds to 9 cycle preceding pulse signal \( (.1680^*0.056=9.5) \). The received signals undergo range processing through matched filter. The compression is carried out by correlating with a time

\(^{4}\text{In ERS-1 case, the beam footprint is around 5 km while minimum slant range } R(x_0) \text{ is measured as 850 km. Hence, Max}(\frac{x^2}{R(x_0)^2}) = (\frac{x_0}{R(x_0)})^2 = 0.010034 \ll 1.\)
delayed conjugate of the original signal and described as

\[ \Omega_{ra}(t) = \int_{-\infty}^{+\infty} \text{conj}[S(t' - t_d)] \cdot S(t' - t) \, dt' \]  

(2.11)

Substituting \( S(t) \) with linear FM signal in Eqn. (2.3), the output \( \Omega(t) \) is developed as

\[ \Omega_{ra}(t) = \exp[j2\pi f_o(t - t_d)] \int_{\min(t,t_d)+\frac{T}{2}}^{\max(t,t_d)-\frac{T}{2}} \exp[-j \frac{B}{2T} \left\{ (t' - t)^2 - (t' - t_d)^2 \right\}] \, dt' \]  

(2.12)

The first term is a phase term dependent upon target range, while the integral term is for azimuth phase processing analogous to linear FM compression maximizing \( Q(t) \) response at \( t = t_d \). After proceeding the integral calculation, range processed output \( \Omega(x, t) \) is given by

\[ \Omega_{ra}(x, t) = \exp(j2\pi f_o t) \exp[-j \frac{4\pi}{\lambda} R(x)] \sin \left\{ \frac{\pi R}{B} |t - t_d| (T + |t - t_d|) \right\} b \frac{R}{|t - t_d|} \]  

(2.13)

When the time-bandwidth product is sufficiently large, the 3-dB width solution for the amplitude term is approximately obtained as \( |t_{3dB} - t_d| \approx 1/B \), which corresponds to earlier assertion. The first carrier term is stripped away by the linear operation of complex modulation leaving the phase term \( \exp[-j \frac{4\pi}{\lambda} R(x)] \) only. In real situation, the antenna beam pattern imposes a weighting along azimuth direction and the phase term is amplitude modulated. Thus the data collection and processing is done only for the phase terms captured while the amplitude is within -3 dB level of peak response. Using the approximation in Eqns. (2.7) and (2.8), range processed signal is transformed into a function of azimuth distance \( x \) only, which is given as

\[ f_{az}(x, x_o) = \exp \left[ -j \frac{4\pi}{\lambda} \left\{ R(x_o) + \frac{(x - x_o)^2}{2R(x_o)} \right\} \right] \]  

(2.14)

Taking correlation with the conjugate function, the azimuth processed output is obtained as

\[ \Omega_{az}(x) = \exp(-j \frac{4\pi}{\lambda}) \int_{\max(x,x_o)-\frac{T}{A}}^{\min(x,x_o)+\frac{T}{A}} \exp \left[ -j \frac{2\pi}{\lambda R(x_o)} \left\{ (x' - x)^2 - (x' - x_o)^2 \right\} \right] \, dx' \]

\[ = \exp(-j \frac{4\pi}{\lambda}) \sin \left\{ \frac{2\pi}{\lambda R(x_o)} |x - x_d| (L_A + |x - x_o|) \right\} \]  

(2.15)

\(^5\)In most cases, the assumption \( T \gg |t - t_o| \) is valid and the amplitude term becomes a sinc type function \( T \sin(\pi B |t - t_d|) \). The 3-dB width is determined by the first zero or by letting \( \pi B |t - t_d| = 0 \). This also generates 3 dB width of \( \frac{T}{B} \).
2.3 Ambiguity

The azimuth resolution is decided by the 3-dB width of $\Omega_{az}(x)$, which will be denoted as $x_{3dB}$. Because $x_{3dB}$ occurs near $x_o$, it should satisfy the first zero condition for $\Omega_{az}(x)$

$$\frac{2\pi}{\lambda R(x_o)} |x_{3dB} - x_o| L_A = \pi$$  \hspace{1cm} (2.16)

Therefore the azimuth resolution is given

$$\delta_{az} = \frac{\lambda R(x_o)}{2} \frac{1}{L_A} = \frac{\lambda R(x_o)}{2} \frac{D}{\lambda R(x_o)} = \frac{D}{2}$$  \hspace{1cm} (2.17)

All of the separate approaches in (i),(ii) and (iii) reach the same conclusion. Although a number of assumption were made to simplify the procedure, the use of a chirp-like matched filter for azimuth SAR processing is known to produce an excellent result in real situation. For spaceborne SAR, the earth curvature may affect the data generation procedure and should be taken into account [61]. For an orbital SAR moving at the height of $h$, the azimuth resolution can be modified by a constant factor

$$\delta_{h,az} = \frac{R_E}{2(R_E + h)} D$$  \hspace{1cm} (2.18)

where $R_E$ is the Earth’s radius.

2.3 Ambiguity

With the assumption of stationarity, two separate targets can be differentiated by range measurement of the return signals. And when their ranges are known, each target can also be discriminated by Doppler shift information contained in return signal. But without prior knowledge of either range nor velocity, it is not possible to exactly separate two different targets in both domains. The ambiguity factor is the mathematical representation of the relationship between two unknown parameters, the range and velocity of particular target [43].

The range and azimuth ambiguities in SAR are results of non-ideal characteristics of antenna performance and finite pulse bandwidth [49]. The SAR algorithm is designed to deal with distributed target which is to be converted into a large group of unit pixels spread over the target area. To achieve a good resolution image, a great amount of data delivering complicate information should be collected in a short pass time of satellite. During this process, finite length of pulses containing the unique information of arbitrary target area interrupt with each other’s time or frequency interval, which result in severe signal corruption. It is the aim of the SAR processing to resolve this overlapped correlation and to extract original pure information. Unfortunately it is not always possible to achieve
satisfactory results in both directions.

### 2.3.1 Range Ambiguity

In the SAR operating satellite system, the radar transmits a number of finite length pulses in each time interval during the flight over target region. The number of pulses and time interval between each pulses is determined in advance. Because the operation of transmitting and receiving data pulses is performed through the same antenna, transmitted pulses must be separated with each other by a finite time interval long enough not to lose returned signal. When $R_s$ is the shortest slant range between target area and radar, $R_l$ is the longest one and $\tau_p$ is the pulse length, the following relationship should be satisfied to avoid interference among different signals

\[
\begin{align*}
\tau_p &< \frac{2h}{C} \\
\frac{2h}{C} + \tau_p &< \frac{2R_s}{C} \\
\frac{2R_l}{C} + \tau_p &< \frac{1}{PRF}
\end{align*}
\] (2.19)

where $h$ and $C$ represent radar altitude above the surface nadir point and pulse propagation speed respectively. Figure 2.5 illustrates a diagram describing the relationship among radar pulse sequences in the time domain.

The nadir return signals which are reflected from the region vertically below down of flight path begin to arrive just after the trailing edge of transmitted pulses. The target response signals are set to exist between nadir return and the next transmitted signal pulse. On the condition of uniform reflectivity throughout the target area, the magnitude distribution of target response resembles the across-track antenna mainlobe pattern. The strong nadir return can be blocked when the time interval of signal transmission is adjusted to correspond to the nadar return period. The length of the received pulse signal decides
the maximum PRF. In practice there may exist a slight variation in the arrival time of
signal echoes depending on the characteristic of terrain surface. Therefore a sufficient
margin should be allowed between echo intervals.

These descriptions are based on the assumption that the antenna has an ideal perfor­

mance with no sidelobe effect. However, in the practical situation, the antenna sidelobes
appear within returned signals and the returned echoes are no more confined to narrow
finite time intervals but instead they spread over extended time width. In the spaceborne
SAR environment where the beam limited swath width increases in proportion to the in­
creased antenna altitude, the sidelobes from preceding pulses can stretch to the mainlobe
beam area of succeeding echoes. These overlapped signal patterns cause the range am­
biguity. Figure 2.6 illustrates the generation of the range ambiguity caused by antenna
movement during flight.

It shows that the sidelobe patterns from the preceding pulses can arrive back to the
antenna simultaneously with the returned mainlobe pattern of the desired pulse sequence.
A signal corruption occurs when the sidelobes of return signals from strong backscatterers
smear into the signals reflected from nearby weak targets. The integrated RASR (Range
Ambiguity Signal Ratio) is defined as a noise factor to measure the impact of this signal
corruption. Without the prior knowledge of the backscatter distribution of target area
it is difficult to evaluate this sidelobe overlapping effect. When $S_{am}$ and $S_t$ represent the
range ambiguous signal and desired signal powers of the $i$th time interval respectively, the
RASR is given by the ratio of these two as

\[
RASR = \sum_{i=1}^{N} S_{ai} / \sum_{i=1}^{N} S_{i}
\]

\[
= \frac{\sum_{i=1}^{N} \sigma_{ij}^{0} G_{ij}^{2} / R_{ij}^{3} \sin(\eta_{ij})}{\sum_{i=-\infty}^{\infty} \sum_{j \neq 0}^{\infty} \sigma_{ij}^{0} G_{ij}^{2} / R_{ij}^{3} \sin(\eta_{ij})}
\]

(2.20)

where \( N \) is the total number of data streams, \( j \) is the pulse sequence number, \( \eta_{ij} \) is the incident angle, \( \sigma_{ij}^{0} \) is the normalized reflectivity at the point, \( G_{ij} \) is the cross-track antenna pattern and \( R_{ij} \) is given as the distance to target. Considering its close relationship with the antenna sidelobe pattern, the RASR can be efficiently controlled by directly manipulating the antenna beam pattern at the cost of system complexity [24][66].

There is another source of performance degradation in range direction which is resulted from imperfect pulse compression operation. During the pulse compression process, a finite level of sidelobe pattern is generated and spread over the uncompressed pulse length time period. The image quality as well as range direction resolution is affected by this pulse energy leakage. This effect can be measured by two different ways. The ratio of mainlobe level to the integrated total sidelobe level \( \frac{P_{\text{main}}}{P_{\text{sidelobe}}} \) represents the noise effect or the loss of energy during the signal compression procedure. And the ratio between mainbeam and peak sidelobe level \( \frac{P_{\text{main}}}{P_{\text{peak}}} \) indicates the chance of error that weak signals from other targets are obscured by strong target response and eventually lost. Because these are the unique characteristics of the radar waveform, the performance can be altered by the choice of the waveform and the applied signal processing method.

### 2.3.2 Azimuth Ambiguity

In the course of SAR azimuth processing, the Doppler history data corresponding to each point on the ground is collected at a PRF rate. The azimuth ambiguities arise from finite sampling of the Doppler spectrum data at intervals of PRF. Usually the PRF is set high enough to separate two adjacent mainbeams in frequency domain. However the imperfect antenna sidelobe pattern means that the SAR Doppler spectrum is not strictly bandlimited and the desired signal band may be contaminated by aliasing effect [65]. Figure 2.7 describes the frequency spectrum distribution of azimuth ambiguity signals.

The azimuth ambiguity to signal ratio (AASR) is defined as the total ambiguous energy folded into the image processing bandwidth from adjacent ambiguous intervals divided by the unambiguous Doppler energy spectrum falling within the image processing bandwidth.
2.3 Azimuth Ambiguity

[53] Assuming target reflectivity has a distribution of $\sigma$, AASR equation is given as

$$AASR(\tau) \approx \sum_{m,n=-\infty}^{+\infty} \int_{-B_p/2}^{+B_p/2} G^2(f + mf_p, \tau + n/f_p)\sigma(f + mf_p, \tau + n/f_p)df$$

where $G$ is the signal spectrum and $B_p$ is the processing bandwidth. Due to the one-to-one between azimuth time and frequency, $G$ is equivalent to antenna beam equation in the azimuth direction. In reality, without the prior knowledge on the target reflectivity distribution, it is not possible to exactly address this property.

When there is a relatively bright target having extremely high reflectivity, a false response can appear displaced from its original true location. The amount of relative displacement by azimuth ambiguity can be estimated and corrected through SAR image processing. Eqs.(2.22, 2.23) show the displacement in range and azimuth direction,

$$\Delta x_{RA} = (m\lambda f_p/f_R)(f_{DC} + mf_p/2) \quad (2.22)$$
$$\Delta z_{AZ} = mf_pV_{st}/f_R \quad (2.23)$$

where $V_{st}$ is the magnitude of the relative platform to target velocity, $m$ is the ambiguity number. $f_{DC}$ and $f_R$ are the Doppler centred frequency and Doppler rate used in the SAR processing respectively.

Since the azimuth ambiguity is directly derived from both the antenna sidelobe pattern and limited maximum PRF value, it is difficult to expect performance improvement.
2.4 Information in SAR Images

by waveform design only. Instead there are several alternative methods to resolve this problem. The multi-PRF technique makes use of more than one PRF value in collecting SAR Doppler data of multiple images and discriminates the false ambiguous signal by comparing each ones [1][24].

The range of the ambiguous signal is wrongly estimated and is not focussed correctly resulting in image blurring. This dispersion in units of focussed range resolution cell is approximated by [52]

\[ N_{DR} = m \lambda^2 f_S R / (4 V_{st} \delta x \delta R) \]  

(2.24)

where \( R \) is the slant range and \( \delta x, \delta R \) are the focussed azimuth and range resolutions, respectively. The use of low frequency signal is desirable since the ambiguity energy is more focused at shorter wavelength.

2.4 Information in SAR Images

SAR images are generated after considerable amount of information is taken into appropriate processing. The available information is a complicate mixture of electromagnetic signals returned from widely distributed target area. Only a good understanding of these returned signal will lead to a full knowledge on target scene. The interpretation of incoming signal is linked to the Inverse Problem: what properties of an unknown scene can be inferred from the observed backscattered signal? [61]

2.4.1 Target Distribution

Figure 2.8(a) shows a SAR image acquired by X-SAR. The radar was looking on the plain land area where the natural environment and human habitual area are combined together in a single scene. It is seen that most of the natural land is covered with clutter-like surface. The graph shown in Figure 2.8(b) is a statistical description of the amplitude intensity of the pixels of the entire image. This amplitude distribution graph illustrates that overall brightness level is confined in a narrow band relative to the peak response. The peak response is more than 20 times higher than the average value of dominant distribution band.

When the target is constructed from many independently positioned scatterers, the probability density function (PDF) of its radar cross section is usually described by a Rayleigh PDF [48]. This is derived on the two assumptions that the observed in-phase and quadrature components, \( z_1 = A \cos \theta \) and \( z_2 = A \cos \theta \), will be independent Gaussian distributed random variables with zero mean and the observed phase \( \theta \) will be uniformly distributed over \([-\pi, \pi]\). The variance of the Gaussian random variables \( \sigma/2 \) is determined
2.4 Target Distribution

![Complex SAR image and its reflectivity amplitude distribution graphs](image)

(a) Complex SAR image  (b) Brightness level distribution

Figure 2.8: Typical SAR image and its reflectivity amplitude distribution graphs for i) entire domain and ii) dominant distribution

by the scattering amplitude leading to a joint PDF given by

$$P_{z_1,z_2} = \frac{1}{\pi\sigma} \exp\left(-\frac{z_1^2 + z_2^2}{\sigma}\right)$$  \hspace{1cm} (2.25)

Then the amplitude $A$ will have a Rayleigh distribution

$$P_A(A) = \frac{2A}{\sigma} \exp\left(-\frac{A^2}{\sigma}\right), \quad A \geq 0$$  \hspace{1cm} (2.26)

while the intensity or the power $I = A^2$ follows a negative exponential distribution.

$$P_I(I) = \frac{1}{\sigma} \exp\left(-\frac{I}{\sigma}\right), \quad I \geq 0$$  \hspace{1cm} (2.27)

Many SAR image processings adopt multilook processing methods in order to enhance the final image quality. The principle of multilook processing is to average multiple pixels in adjacent area by viewing from different angles. After multiple independent measurements are combined, the mean power $\sigma$ is still preserved while the variance is reduced by the factor of total multilook number denoted as $L$. The resulted measurement has a well-known analytic PDF for both intensity and amplitude distribution cases. Introducing a gamma function $\Gamma(x)$, the PDFs of new multilook intensity $P_I(I)$ and amplitude $P_A(A)$ are calculated respectively as [61]

$$P_I(I) = \frac{1}{\Gamma(L)} \left(\frac{L}{\sigma}\right)^L I^{L-1}e^{-LI/\sigma}, \quad I \geq 0$$  \hspace{1cm} (2.28)

$$P_A(A) = \frac{2}{\Gamma(L)} \left(\frac{L}{\sigma}\right)^L A^{2L-1}e^{-LA^2/\sigma}, \quad A \geq 0$$  \hspace{1cm} (2.29)
2.4 Target Distribution

(a) SAR image

(b) Entire Domain Distribution

(c) Natural Environment Area Distribution

Figure 2.9: SAR image example and its reflectivity amplitude distribution graphs and comparison with known distribution functions 1) Gamma 2) Rayleigh distribution

where $\sigma$ is the mean value of intensity. The processed SAR image is a collection of these reflected signals. Each pixels in the image represents the average intensity values within resolution area.

Figure 2.9(a) shows a SAR image where a relatively wide dynamic range of amplitude level is present. This X-SAR image has been generated through four multilook processing and hence $L=4$ is used to evaluate the gamma distribution. As has been shown earlier in Figure 2.8(b), due to the presence of small number of extreme cases, the total brightness level distribution graph presents no meaningful signature regarding the characteristic of the target scene. Human settlement area appears in very bright tones in contrast with the rest of the image. The existence of these bright pixels shifts the average amplitude level and prevents either gamma or Rayleigh distribution from fitting into the obtained amplitude
distribution (Figure 2.9). It is only after those excessively bright points are taken out that the distribution begins to form a recognizable shape. Figure 2.9(c) confirms that natural land surface area, bounded within dashed line in Figure 2.9(a), exhibits a good consistency with the predicted distribution graph.

It is, hence, deduced that there exist points target responses that may not follow the overall tendency of the image. It is discovered that those highly bright points are generally contributed from a heterogeneous part of the image which is thought to be the human residential area. A good estimation of the target distribution plays an important role in classification and interpretation of SAR images [61]. Therefore there rises a need to deal with SAR images in separate ways depending on the characteristics of target scenes. In general there are two distinguished features possessing separate intensity distributions, which are natural areas such as water, land or trees and urban area comprised of man-made structures respectively. Chapter 3 is devoted to the analysis of the behaviour of signals originated from non-natural environment or man-made structures in urban area.

‘Basic research is what I am doing when I don’t know what I am doing’ - Werner von Braun
Chapter 3

ELECTROMAGNETIC PHENOMENA in SAR

Throughout the long history of SAR development, a great deal of effort has been devoted to proper interpretation and understanding of the produced images. However, very little attention has been paid to the method which employs the fundamental electromagnetic theory in order to explain the interaction between SAR signal and target scene. The lack of this attempt is attributed to the facts that SAR image resolution is far larger than the signal wavelength and each image pixel is generated from a mixture of broad range target responses [81]. Hence even a good analysis on the interaction between individual target points and transmitted signals tends to be useless without the prior knowledge of the complete target scenes. SAR image structure is more influenced by the distribution of target reflectivity, which has motivated many researches to approach in terms of target information theory [34][61].

But when urban structures are concerned, most of the structure in view is man-made artificial objects and they are no longer subjects to any particular distribution type. In many cases the interactions between signals and man-made scatterers show predictable behaviours. It is seen that certain type of structures in urban environment behave in unique manners and appear in the image either as extremely bright objects or dark area depending on the target characteristics and the relative position between radar and scatterers. A proper interpretation of these phenomena and target classification would demand a full understanding of the radar signal response corresponding to scatterers.

To accomplish this, as Byran [10] indicated, it will be necessary to understand both electromagnetic scattering theory and the morphology of the city [81]. There was an attempt to apply the electromagnetic ray theory to simulate SAR responses from simple structures [13] and very recently a brief report is published on how to predict the SAR response from simple reflectors [70]. Both of them are based on the GTD (Geometrical Theory of Diffraction) ray tracing technique. The simulations have been performed on
ideal perfect point conductors, and therefore do not provide sufficient clues to relate SAR image with dynamic urban feature. After a brief introduction regarding this technique, general characteristics of SAR urban scene are discussed along with some suggestions on how to tackle the problems.

### 3.1 Geometrical Theory of Diffraction

The difficulty of analyzing electromagnetic interaction between signals and target objects intensifies dramatically for complicated target structures. Unless the objects are well-defined simple structures, any attempts to derive completely rigorous solutions are discouraged. Recently many near-field problem involving arbitrarily structured objects are approached through numerical analysis. The Finite-Domain-Time-Difference (FDTD) is one of the most prominent methods among these groups [45]. It is useful when the target objects are relatively small and the signal source is located within close range or limited to plane waves. But when the concerned targets are large scaled macro-cells and the signal source is located in the far-field region, Geometrical Theory Diffraction (GTD) rises as the most suitable candidate. Since conceived by Keller [41] in the 1950s, GTD is now established as a leading analytical technique in the prediction of high frequency diffraction phenomena [39].

#### 3.1.1 Basic Theory

Consider a situation in Figure 3.1 in which simple sources and arbitrary shaped targets are positioned in an open field separated by $R(x, x')$. The most straightforward way to calculate the electromagnetic field resulting from the interaction between them will be to refer to Maxwell's equations, which is given in Eqn. (3.1). The ultimate goal of all scattering problem is to obtain field components $E$, the electric field intensity or $H$, the
3.1 Geometrical Optics

Magnetic field intensity at the position of the observer \((x_o,y_o,z_o)\).

\[
\begin{align*}
- \nabla \times E &= j \omega \mu_m H + j \omega (\mu - \mu_m) H + M \\
\nabla \times H &= j \omega \epsilon_m E + j \omega (\epsilon - \epsilon_m) E + J
\end{align*}
\quad (3.1)
\]

Although they explain the scattering phenomena in exact forms, their solution involves three dimensional integrals in the complex variable domains and the difficulty is dramatically heightened with the increased complexity of both the source and target structure. GTD is developed from the known exact solution for simple shapes, which is called canonical problem. GTD is a high-frequency analysis technique. The derivation of its approximate formulation starts from the assumption that when the diffraction body is large by comparison with the wavelength, the scattering response of an object can be well approximated as a sum of responses from individual scatterers or scattering centres. These responses include reflection, refraction and diffracted rays \([44][41]\).

Let the incident field be an infinite plane wave propagating in the \(+z\) direction. Then the field at location \(\vec{r}\) can be described by

\[
\bar{E}(\vec{r}) = E_o e^{-jkz\cdot\vec{r}}
\quad (3.2)
\]

where \(k\) is the wavenumber. When the wavelength of the incident excitation is small relative to the object extent, the backscattered field appears to originate from a set of discrete scattering centres and can be approximated by

\[
\bar{E}^s(k, z) \approx \bar{E}_o e^{jkz} \sum_{m=1}^{M} \frac{\phi(m, k)}{|z - \vec{r}_m|} e^{-jkz\cdot\vec{r}_m}
\quad (3.3)
\]

where \(\vec{r}_m\) is the location of the \(m\)th scattering centre and \(\phi(m, k)\) is frequency dependent coefficient determined by the scattering mechanism. In the far-field, \(|z - \vec{r}_m|\) can be simplified as \(|z|\) only. Each of the components is called scattered ray. The aim of the GTD is to obtain coefficient values \(\phi(m, k)\) for all scattered rays including special occasion of diffraction.

3.1.2 Geometrical Optics

Geometrical Optics (GO) sees the electromagnetic phenomena as a collection of propagation rays. GO interprets electromagnetic fields as the sum of the leading terms in the asymptotic expansions, which are derived from the exact solution. Three types of rays are generally considered to express the whole field: (a) incident, (b) reflected and (c) refracted
3.1 Geometrical Optics

Incident Field

Reflected Field

Reflected Field

Reflected Field

Incident Field

Reflected Field

Reflected Field

Figure 3.2: Reflection and refraction between two lossy media along planar boundary

rays respectively and represented in series form as

\[ E \approx \exp(-jk_n) \sum_{m=0}^{\infty} \frac{E_m}{(j\omega)^m} \]

(3.4)

All the rays are assumed to obey optical laws as well as the energy conservation rule.

Snell's law

The reflected and refracted rays at the boundary between two slightly lossy media are subject to Snell's law. Figure 3.2 is a diagram to depict graphical relationship between the rays concerned. The boundary between the two media is set as infinite planar interface for simplicity. Many structures found in the urban area can be treated similarly. Let this incident ray at the position of \( s \) be expressed as \( E^i(s^i) \), which is the sum of two polarization components namely \( E_{r}^i(s^i) \) and \( E_{n}^i(s^i) \). According to Snell's law, the reflected and refracted field \( E^r(s^r) \) and \( E^n(s^n) \) are determined by the relevant coefficients between the two media and given respectively as

\[ \hat{E}^r(s^r) = \left( \begin{array}{c} \hat{E}_{r}^i(s^i) \\ \hat{E}_{n}^i(s^i) \end{array} \right) = \hat{E}^i(s^i) \left[ \begin{array}{cc} R_e & 0 \\ 0 & R_m \end{array} \right] \exp(-j\tilde{k}_r S^r) \]

(3.5)

\[ \hat{E}^n(s^n) = \left( \begin{array}{c} \hat{E}_{r}^i(s^i) \\ \hat{E}_{n}^i(s^i) \end{array} \right) = \hat{E}^i(s^i) \left[ \begin{array}{cc} T_e & 0 \\ 0 & T_m \end{array} \right] \exp(-j\tilde{k}_n S^n) \]

(3.6)

The \( R_e, R_m, T_e \) and \( T_m \) terms are electrical and magnetic coefficients for reflection and transmission field respectively and determined by the ray propagation angles. For slightly lossy media, which are the case for many real world scatterers, a critical angle \( \nu_c \) is defined

---

\(^1\)This expansion series is known as Luneberg-Kline series
3.1 Geometrical Optics

as

\[ \sin(\nu_c) = \sqrt{\frac{\epsilon_2}{\epsilon_1}} \]  

such that if the angle of the incidence exceeds the critical angle \( \nu_c \), the refracted field cannot penetrate into the second medium and the total incident field is completely reflected. Using the continuity properties of both the electric and magnetic fields at the interface of the two media, Snell’s law establishes the relationship between incident, reflected and refracted angles as

\[ \nu_r = \nu_i \] 
\[ \sin(\nu_i) = \sqrt{\frac{\epsilon_2}{\epsilon_1}} \sin(\nu_i) \]  

The reflected and incident wave make identical angles with the normal to the interface, while the refracted angle is a function of the permittivities. Eqns. (3.8) and (3.9) also specify the wavenumber vector \( \hat{k}_r \) and \( \hat{k}_t \). The continuity property of the field at the interface leads to the derivation of the reflection and transmission coefficients, which are expressed as

i) Incident angle \( \nu_i \leq \nu_c \)

\[
\begin{bmatrix}
R_e & T_e \\
R_m & T_m
\end{bmatrix} = \begin{bmatrix}
\sin(\nu_t - \nu_i) & 2 \cos(\nu_i) \sin(\nu_t) \\
\sin(\nu_t + \nu_i) & \sin(\nu_t + \nu_i)
\end{bmatrix} \begin{bmatrix}
\frac{2 \cos(\nu_i) \sin(\nu_t)}{\sin(\nu_t + \nu_i) \cos(\nu_i - \nu_t)}
\end{bmatrix}
\]  

(3.10)

ii) Incident angle \( \nu_i > \nu_c \)

\[
R_e = \frac{\cos \nu_i - j(\sin^2 \nu_c - \sin^2 \nu_i)^{0.5}}{\cos \nu_i + j(\sin^2 \nu_c - \sin^2 \nu_i)^{0.5}}
\]
\[
R_m = \frac{\sin^2 \nu_c \cos \nu_i - j(\sin^2 \nu_c - \sin^2 \nu_i)^{0.5}}{\sin^2 \nu_c \cos \nu_i + j(\sin^2 \nu_c - \sin^2 \nu_i)^{0.5}}
\]  

(3.11)

Although they have been simplified for planar boundary media, these coefficients can be used to determine the electromagnetic energy flow in the presence of scatterer inside the medium.

Ray Tracing

Most spaceborne SARs use C or X-band radar signals. Their relatively short wavelength compared with ground buildings allows us to understand the movement of electromagnetic signal as a ray propagation. Consider a field movement along a straight axial ray between two points \( r_1 \) and \( r_2 \) in a homogeneous medium. Figure 3.3 illustrates the fields movement along the axial ray \( \hat{r}_2 - \hat{r}_1 \). Suppose that the energy flow is generated...
3.1 Geometrical Optics

Axial Rays

Origin of curvature $\delta_1, \delta_2$

Reference $\bullet O$

Figure 3.3: Propagation modeling along the axial ray and quantitative diagram of energy flow

from a point source target located at $O$ along the axial ray. The task is to derive the electric field at the wavefront of $r_2$ when the incident wave field is given at the position $r_1$. The wavefront area at each $r_1$ and $r_2$ are denoted as $dA_1$ and $dA_2$. In general case, these wavefront will be curved with two radii curvature, which are denoted as $\delta_1$ and $\delta_2$ respectively. The principal rule, ‘Conservation of Energy’, formulates the relationship between the electric fields and the area of the ray tube as

$$|E'(r_1)|^2 dA_1 = |E'(r_2)|^2 dA_2 \quad (3.12)$$

where $E'(r_1)$ and $E'(r_2)$ represent the electric field at each position. Since both the tube area $dA_1$ and $dA_2$ are formed by two different curvatures $\delta_1$ and $\delta_2$, their ratio will be given as

$$\frac{dA_1}{dA_2} = \frac{\delta_1 \delta_2}{(\delta_1 + \Delta r)(\delta_2 + \Delta r)} \quad (3.13)$$

where $\Delta r = |r_2 - r_1|$. Eqns. (3.12) and (3.13) combine together to evaluate the electric field amplitude component at $r_2$. Using the field notation of Eqn. (3.2), the field at $r_2$ is given

$$E'(r_2) = E'(r_1) \frac{\delta_1 \delta_2}{(\delta_1 + \Delta r)(\delta_2 + \Delta r)} \exp(-jk[r_2 - r_1]) \quad (3.14)$$

Eqn. (3.14) implies that the knowledge of the field over one wavefront in the homogeneous space is sufficient to obtain the field at any other point. When both the curvature radii $\delta_1$ and $\delta_2$ approach to infinity, the amplitude proportional term approximates to unity and the field is a simple plane wave flowing toward $\hat{r}_2 - \hat{r}_1$.

The introduction of ray tracing concept allows successful modeling of complicated
Figure 3.4: Plane wave propagation fields including diffraction field around half plane structures residing in the near or far-field. However, it is known that the use of GO technique along generates sharp discontinuities in the certain fields area such as field shadow or reflection boundaries, which are not physically allowed. To provide explanation for those unrealistic phenomena, the diffraction rays are introduced through GTD concept.

3.1.3 Diffraction rays

The main contribution of GTD is found in the concept of diffraction rays that perfect GO method in describing electromagnetic field. The diffracted rays are produced when incident ray hit edges, corners or vertices of boundary surfaces. The most well known solution among diffraction problem is for the conductive half plane of which the exact solution is given by Sommerfeld. [74]. Figure 3.4 shows the electromagnetic field distribution around half plane.

Sommerfeld showed that the incident, reflected and diffracted field components exist in separate form within the exact solution of the total field\(^2\). Figure 3.4 shows that, when a plane wave is normally incident upon the half-plane, the total field area is divided into three different regions depending on the incident angle. The GO field has considered incident and reflected fields only. The different set of components at each three regions cause two discontinuities at the interfaces. This discontinuity are amended by the diffraction field \(U_d\), which itself is a discontinuous function. Each field component can be represented as a function of distance from the diffraction point \(r\) and the angle \(\phi\). It has been well known since Keller devised GTD concept that the diffraction field \(U_d(r, \phi)\) is determined by the incident field component at diffraction point \(\hat{r}_d\) and edge-diffraction coefficient

\(^2\)The exact solution is given by Sommerfeld and for a half plane conductor. A rigorous solution to the wedge problem only exist for perfectly conducting wedge cases.
Generally the diffraction field is expressed as

\[ U^H,V_d(r, \phi) = U_i(\vec{r}_d) \cdot \tilde{D}(\vec{r}_d) \sqrt{\frac{\delta_d}{(\delta_d + r)r}} \exp(-jkr) \]  

(3.15)

where both \((H, V)\) polarizations are considered and \(\delta_d\) is the principal radius of curvature at the observing point. James [39] adopted a compact form to represent this diffraction term \(U_d(r, \phi)\) by dividing it into two separate components associated with incident and reflected field, which are given as

\[ U^H,V_d(r, \phi) = U^i_d(r, \phi) \mp U^r_d(r, \phi) \]  

(3.16)

where the upper and lower signs are for \(E\) and \(H\)-polarization cases respectively. Each diffracted component in Eqn. (3.16) is given as

\[ U^{i,r}_d(r, \phi) = -\text{sgn}(\Gamma^{i,r})K_\pm \left\{ |\Gamma^{i,r}| \sqrt{kr} \right\} \exp(-jkr) \]  

(3.17)

where

\[ \Gamma^i = \sqrt{2} \cos \frac{1}{2}(\phi - \phi_i) \]
\[ \Gamma^r = \sqrt{2} \cos \frac{1}{2}(\phi + \phi_i) \]  

(3.18)

\[ K_\pm(x) = \sqrt{\frac{j}{\pi}} \exp(jx^2) \int_x^\infty \exp(-jt^2)dt \]

\(K_\pm(x)\) is a modified Fresnel integral. The sign terms \(\text{sgn}(\Gamma^{i,r})\) correspond to incident or reflection associated diffractions and are determined by the presence of those diffraction field. As a result discontinuities are generated within diffraction terms at the boundaries between the incident and reflect field regions, which eventually compensate for the GO field discontinuities. For polarization field components, simple diffraction field expressions are obtained as

\[ E^d(r, \phi) = E^d(\vec{r}_d)D^e(\phi, \phi_0) \sqrt{\frac{\delta_d}{(\delta_d + r)r}} \exp(-jkr) \]
\[ H^d(r, \phi) = H^d(\vec{r}_d)D^m(\phi, \phi_0) \sqrt{\frac{\delta_d}{(\delta_d + r)r}} \exp(-jkr) \]  

(3.19)

Eqn. (3.19) implies that diffracted fields behave as if they had emanated from a line source at the edge \(\vec{r}_d\). Away from the optical boundaries where the Fresenel integral argument is sufficiently large, we may use its asymptotic term and then the \(E, H\) diffraction coefficient develops into

\[ D^{e,m} = -\left\{ \sec \left( \frac{\phi - \phi_0}{2} \right) \mp \sec \left( \frac{\phi + \phi_0}{2} \right) \right\} \]  

(3.20)
3.1 Diffraction rays

which forms the basis of the Keller's GTD derivation. The diffraction coefficients for other structures have similar expressions [39]. Although they are valid in most GO field areas, the approximate expression in Eqn. (3.20) indicates that the Keller's GTD coefficients are not applicable in the transition regions close to the optical boundaries. i.e. (∴ \phi - \phi_0, \phi + \phi_0 = 0).

To ensure that the total field is continuous at shadow and reflection boundaries, Uniform Geometrical Theory of Diffraction (UTD) is introduced. UTD is a modified form of GTD and solves discontinuous diffraction field problems by separating field near optical boundaries from other regions [44]. UTD regards the diffraction field around the boundaries as electromagnetic plane waves with half the magnitude of the optical field so as to compensate the discontinuities of the GO field. UTD diffraction coefficients are derived asymptotically through integral approximation of the Fresnel integral in exact solution. The resulting expression for dyadic diffraction coefficient of a perfectly conducting wedge of angle \( \beta = (2 - N)\pi \) is

\[
D^{\epsilon,m}(\phi, \phi') = \frac{-1}{N\sqrt{8j\pi k}} \left[ \begin{array}{c}
\cot \left( \frac{\pi + (\phi - \phi')}{2N} \right) F(kLa^+ (\phi - \phi')) \\
\cot \left( \frac{\pi - (\phi + \phi')}{2N} \right) F(kLa^- (\phi - \phi')) \\
\cot \left( \frac{\pi - (\phi - \phi')}{2N} \right) F(kLa^+ (\phi + \phi')) \\
\cot \left( \frac{\pi + (\phi + \phi')}{2N} \right) F(kLa^- (\phi + \phi'))
\end{array} \right]
\]

(3.21)

where

\[
F(x) = 2j\sqrt{x} \exp(jx) \int_{\sqrt{x}}^{\infty} \exp(-jt^2)dt
\]

(3.22)

\[
a^\pm(\omega) = 2\cos^2 \left( \frac{2N\pi m^\pm - \omega}{2} \right)
\]

(3.23)
in which \( m^\pm \) are the integers that most nearly satisfy the equations

\[
2\pi Nm^\pm - \omega = \pm\pi
\]

(3.24)

\( L \) is a distance parameter and varies depending on the types of illumination. For plane-wave incidence, \( L \) is given as the distance \( r \). The inclusion of complex indefinite integral \( F(x) \) within diffraction coefficients will generate a heavy computational burden during numerical calculation, which otherwise, would have been very straightforward. It will be more convenient to introduce a simplified polynomial approximation form rather than taking approximate integral for individual cases. Depending on the range of \( x \), the Fresnel integral has approximation forms of

• \( x \) is small

\[
F(x) \approx \left\{ \sqrt{\pi x} - 2x \exp(j\frac{\pi}{4}) - \frac{2}{3}x^2 \exp(-j\frac{\pi}{4}) \right\} \exp \left[ j\left(\frac{\pi}{4} + x\right) \right]
\]

(3.25)
3.1 Diffraction rays

Figure 3.5: Scattered fields around half-plane conductor at \( r = 5\lambda \) generated by plane wave: (a) Diffraction field \( E_d \) magnitude, (b) Scattered field \( (E_d + E_r) \) and comparison with total field \( E_t = E_i + E_d + E_r \)

- \( x \) is large

\[
F(x) \approx 1 + j \frac{1}{2x} - \frac{3}{4x^2} - \frac{15}{8x^3} + \frac{75}{16x^4} \quad (3.26)
\]

Although there is no clear boundary for defining small or large \( x \), it is generally accepted that \( F(x) \) can be replaced by Eqn. (3.25) for \( x < 1 \) and by unity for \( x > 10 \) cases which generate an identical result with Eqn. (3.20). Using these UTD expressions for two dimensional scattered field, the electromagnetic field around a half-plane conductor is calculated and the results are shown in Figure 3.5. The incident field is assumed to be a unit amplitude, E-polarized plane wave with incident angle \( 30^\circ \) (\( \phi = 60^\circ \)). The scattered field is measured at all angles with the distance kept at \( 5\lambda \) from the diffraction edge. Figure 3.5(a) shows the field component contributed by the diffraction from the edge. It is seen that most of the diffraction field energy is centred around the transition regions (i.e. incident and reflection field boundaries). The first peak is from \( U_d^\prime(r, \phi) \), while the second one is due to \( U_d^\prime(r, \phi) \). These \( U_d^\prime(r, \phi) \) and \( U_d^\prime(r, \phi) \) components are symmetrical in magnitude but with opposite signs along each sides of the peaks. These abrupt phase changes help to remove discontinuities of the GO field which is non-existent phenomena in real world.

Although a half plane conductor is not a realistic model, it has been chosen to provide a fundamental idea of GTD. The GO field and diffraction ray concept can be extended to various model structures. Both GTD and UTD techniques are not exact solutions but asymptotic approximations for high frequency scattering problems. Once the target geometry is fully understood and electromagnetic source is well defined, the desired field solution is obtained at any observing position. The accuracy of the solution is determined
by the resolution of the model grid and number of rays involved in the calculations. A sophisticated subdivision of target structure may improve the results but with the price of excessive calculation burden, which will grow exponentially. There is no general rule but several factors may be taken into account for GTD modeling. These include

- Relative size of the target structure in comparison to the frequency
- The distance between source and scatterer or observer and scatterer
- Complexity of the scatterer
- Dielectric property of the target
- Surface roughness
- Look angle
- Interaction with other scatterers

The requirement of a well-defined modelling of target structure leads to the conclusion that GTD technique is not suitable to predict the result of SAR observation on natural vegetation area. On the other hand, considering that urban area is mostly comprised of man-made structures which can be described as simple scatterers, the GTD technique may be applicable for analysis purpose.

### 3.2 SAR Environment Approach

The simplicity of the GTD model is based on the assumption that the major contribution towards the scattered field comes from an area in the neighbourhood of some critical points on the scattering surface [13]. Those are generally regarded as 1. Specular, 2. Edge-diffraction, 3. Corner-diffraction points and each of them is termed depending on the path of the ray as well as the shape of scatterers. Traditionally GTD has pursued exact solutions to explain electromagnetic phenomena while the complicate nature of SAR sensing methodology defies such an rigorous approach [18]. However it has been demonstrated that the well established GTD technique still enables us to achieve modest accuracy in predicting reflection level from various urban scene [70][13] mostly for communication signal analysis. GTD is especially useful in urban environment prediction. The reflected signals from this area go through diverse paths combined with multiple reflections and diffractions. GTD technique, provided a good simulation modelling, takes into account these complexity and calculates approximate amount of energy within returned signal. For SAR application, an assumption is made that the propagation field between surfaces within target scene follow the geometrical optics law.
3.2 Wedge Structure Scattering

The diffraction scattering field at a general wedge structure is illustrated in Figure 3.6(a). The diffraction rays propagate in arbitrary directions as long as they obey the Fermat's principle\(^3\). Their behaviours also follow the general diffracted field equation shown in Eqn. (3.15). With the simplified condition that the wedge is perfect conductor with wedge angle \( \beta = (2 - N)\pi \) and the incident field is an arbitrary polarized plane wave, the diffraction field at a receiver point \( \mathbf{R}(r, \theta, \phi) \) becomes

\[
\mathbf{E}^d(\mathbf{R}) = \begin{bmatrix} D^e & 0 \\ 0 & D^m \end{bmatrix} \mathbf{E}^i(\mathbf{R}_d) \left( \frac{\delta_d}{r(r + \delta_d)} \right)^{\frac{1}{2}} \exp(-jkr) \tag{3.27}
\]

\( \mathbf{E}^d(\mathbf{R}) \) is an electric field vector including both \( E \) and \( H \)-polarizations. Basically, the procedure to obtain these new diffraction coefficients is very similar to the previous two-dimensional case but as geometrical characteristic is more complicated, there are more restrictions and variables to consider. Fortunately, SAR environment is treated as a far-field problem and only plane incident wave will be considered. By employing boundary condition for perfect conductor at the edge, it can be shown that, the distance parameter \( L \) within the Fresnel integral of 3.21 should be modified as [39]

\[
L = r \sin^2(\theta_0) \tag{3.28}
\]

while oblique angle term \( \csc \theta_0 \) is additionally included in the Diffraction coefficient.

\(^3\)Fermat’s principle for edge diffraction: ‘The edge diffracted rays from a point \( S \) to a point \( P \) are those rays for which the optical path length between \( S \) and \( P \) with one point on the diffracting edge is stationary with respect to infinitesimal variations in the path’
Since SAR response can be assumed to be a far-field pattern, the Fresnel integral may be replaced with unity. Thus calculated diffraction matrix components $D^e$ and $D^m$ are given respectively as:

$$D^{e,m} = \frac{2}{N} \frac{\sin \frac{\pi}{N} \csc \theta}{\sqrt{8j\pi k}} \left\{ \left( \cos \frac{\pi}{N} - \cos \frac{\phi - \phi_0}{N} \right)^{-1} \pi \left( \cos \frac{\pi}{N} - \cos \frac{\phi + \phi_0}{N} \right)^{-1} \right\} \quad (3.29)$$

This equation is valid for about $kr > 10$, which is the most case in SAR target area.

As is noted from the Eqn. (3.29), the GTD diffraction coefficients asymptotically increase to infinite at $\phi - \phi_0, \phi + \phi_0 = \pi$ and cause singularities. This effect should be compensated during numerical computation. The vertex point diffraction case is somewhat different. Although the derivation procedure is still very similar as shown so far, their propagation characteristics is clearly distinguished from others.

The Figure 3.6(b) shows that the diffracted field from a vertex point appears to be originated from a point source. Using the general notation, the diffraction field from this point may be described as:

$$U^{H,V}_d(r, \phi) = U_i(r_d) \cdot \bar{D}(r_d) \frac{\exp(-jkr)}{r} \quad (3.30)$$

Now it is clear that the field diffracted at vertex point forms a spherical wave pattern and attenuates in proportion to reciprocal of $r$, while the attenuation rate for specular or edge refraction field is proportional to reciprocal of $\sqrt{r}$. Considering the long distance for signals to travel until they reach to antenna platform, the contribution of vertex diffraction waves to the total field may be negligible and are excluded in this thesis.

There may be cases where the diffraction field is again diffracted or reflected from nearby obstacles before they lose substantial energy. This is entirely dependent on the geometrical characteristic of the target scene. Although these phenomena are also taken into account in signal propagation analysis for mobile communication where distance between scatterer and receiver is short, they may be ignored in SAR simulation.

### 3.2.2 Reflectivity and View Angle

It is not sufficient to have a large RCS in one narrow direction since SAR image is formed by a correlation of signals of azimuth length. Therefore, to generate an extremely bright spot, the structure should have an omnidirectional uniform reflectivity while it is in view of SAR. One possibility comes from trihedral structures which are abundant in urban area. However, in a typical spaceborne SAR environment, the antenna beam width is less than $1^\circ$ (: $0.3^\circ$ for ERS-1,2) and the view angle can be considered to be static. Hence the amplitude variation of reflected signal along azimuth direction is dominantly affected by the antenna pattern.
Figure 3.7: ERS-1 SAR raw data after range direction processing is done. (a) A range processed image containing a dominantly bright target. (b) Azimuth direction amplitude graph along the dominant bright response.

The far-field pattern of aperture antenna is given by (3.31)

$$|E(\phi)|^2 = \left[ \frac{\sin \left\{ \pi \left( \frac{L_0}{\lambda} \right) \sin(\phi) \right\}}{\pi \left( \frac{L_0}{\lambda} \right) \sin(\phi)} \right]$$

where $L_0$, $\lambda$ and $\phi$ are antenna length, signal wavelength and antenna angle respectively.

Figure 3.7 verifies the relationship between antenna beam and target reflectivity using the actual SAR raw data acquired by ERS-1. (a) is obtained after range processing alone is implemented. The peak response along the bright line is shown in (b). Using the system parameters for ERS-1, the illumination time during which a point target is within the antenna main pattern is calculated as 0.6467 [sec]. The PRF of 1680Hz results in $1680 \times 0.6467 = 1086$ samples per one point target which is equivalent to the number of pixels contained within a mainlobe of the point target response. The calculated mainlobe width is verified in (b). In practice, the reflectivity of a target tends to be lower than the actual antenna pattern level at higher incident angles, which is evident in the graph. It is attributed to the fact that low grazing angle encourages more reflection to the other side of target and, as a result produces reduced backscattering energy.

4By courtesy of Matra Marconi Space.

5When a single dominant target is present as is the case of Figure 3.7, its point target response can be used as a reference signal for azimuth processing.
3.3 Urban Images

Apart from antenna beam pattern which is a system characteristic, the relative position and orientation of the target structure can affect the reflection mechanism and cause drastic changes in the way for certain type of targets to appear at the radar image.

Figure 3.8: SAR image amplitude distribution-Rayleigh and Gamma- analysis and comparison between rural and urban scenes

Apart from antenna beam pattern which is a system characteristic, the relative position and orientation of the target structure can affect the reflection mechanism and cause drastic changes in the way for certain type of targets to appear at the radar image.

3.3 Urban Images

The close proximity of very bright tones from good radar reflectors is a typical pattern found in residential area. This feature will enable large urban scene to be easily distinguished from agricultural land. Figure 3.8 clearly illustrates the distinction between rural and urban SAR image scenes. In case of rural region, the amplitude distribution is well described by gamma function with multi-look constant L given as 4, while the Rayleigh graph fails to predict the correct statistics. Unlike rural area which comprises large clutter regions, the urban scene contains several strong targets densely spread over the map. The amplitude distribution graph produced by this map shows disappointing discrepancy between the theoretical gamma function and empirical data distribution. Contrary to the gamma function which falls sharply with the increase of amplitude, the actual distribution
3.3 Understanding of Urban Radar Response

When the observed scene is no longer regarded as homogeneous, other type of distribution should be considered. Among others, $K$ distribution has attracted a great attention in the past [38]. The abundance of the bright targets shifts the average reflectivity level further right and changes the peak of otherwise gamma curve. This phenomenon poses a challenge to the correct interpretation of SAR images and encourages to introduce a new type of modelling for urban area. Recently, Frery [2] developed a new type of distribution function that fits some urban or deforested area with widely varying degrees of homogeneity. The new distribution, called $\theta^0$ is produced after several distribution results each of which generated using different parameters are combined together to approximate the statistics of extremely heterogenous regions.

3.3.1 Understanding of Urban Radar Response

Although urban areas represent very small portion of the Earth's surface, over 70 percent of the world's population live in urbanized areas [29]. Naturally, the economic importance and land values are among the highest of all land uses. Despite its importance and high potential for a better understanding of human settlement pattern, the urban area has received much less attention than studies of vegetation or other natural environment. One reason contributing to this lack of attention is the high variability of the urban landscape and the complexities of the interactions between the radar signal and the human build-up environment [34]. This is primarily because those returns from paved roads, garden and bridges are often obscured by the strong reflections from building structures. Within the city, very little information on detailed urban structure can be identified [3]. However the boundaries of larger enclaves within the built-up areas such as parks or playgrounds can usually be accurately mapped since they embrace plain clutter regions. Some important scatterings that are responsible for radar responses in urban environment, are illustrated in Figure 3.9. The classification is executed according to their contribution to the brightness level within radar mapping.

In case of specular reflection over smooth surface, most of the incoming energy is scattered away in opposite direction. The dark appearance in the radar imagery is well observed from pavement, calm water surface and building rooftop surface. While a specular reflection is well understood and thus predictable, the diffusion mechanism is extremely complicated and their behaviour should be dealt with random distribution concept. Radar is very sensitive to the surface roughness. The amount of radar backscattering energy varies depending on the degree of roughness. In general, given the r.m.s height of the land surface, degree of roughness level is determined by the Rayleigh criterion which is defined
3.3 Understanding of Urban Radar Response

Figure 3.9: Classification of radar backscattering models within urban environment. (a) A simple specular reflection on plane surface with arbitrary smoothness levels; dark image, dark-medium grey, bright grey (b) Combination of multiple reflection and diffractions (c) Strong specular reflection from man-made, cavity structure; Very strong return and exceptionally bright image

as

$$\begin{cases} \text{Smooth} < \\ \text{Rough} > \end{cases} \frac{\lambda}{8} \sin(\gamma)$$

where $\lambda$ is the wavelength and $\gamma$ is the radar depression angle [10]. Due to the unpredictable nature and case by case variation, however, it is not practical to attempt to match image brightness level to the specific roughness. Instead many research activities compare the empirical SAR data with optical maps of the scene and attempt to establish the relationship between brightness level and urban land classes [31][38][30][9].

On the other hand, it has been argued that the specular return from man-made objects does not follow random movement but shows a particular pattern [75], which gives rise to a motivation for in-depth analysis [61]. With the condition that the size of target is sufficiently larger than radar wavelength, a ray tracing technique seems to be a good candidate for this purpose.

Plain Building Structure

As a means to verify the use of the GTD technique for urban structure modelling, a plain rectangular pillar is taken as a simplified example of a building structure, which is illustrated in Figure 3.10(a). The rectangular pillar is a simple model for a typical building structure commonly found in urban area. For convenience and easy interpretation of the result, the simulated target is assumed to be coated with perfectly conducting material ($\sigma = \infty$). With the assumption that the building is sufficiently high, the simulation is performed in two-dimensional domain of X-Y plane and the returned signal is calculated
3.3 Corner Reflector

Building roofs may act as a mirror reflector or diffuse scatterer depending on the type of material and relative orientation. Since direct specular returns are not common due to the extremely narrow propagation angles [67], many of the strong returns observed in SAR image are likely to originate from corner reflection phenomena. In fact many empirical results support this claim [12][81]. Despite their important roles in forming SAR images, theoretical models are yet to be developed to account for the behaviour of this phenomena.

Figure 3.11 are some examples of modern cities containing tall buildings of diverse sizes. These type of high building structures do not necessarily represent general city looks, although they are easily found within many big cities in North America and most of the developing countries in Asia.

Total field can be established by taking into account all possible ray paths including...
multiple reflections and diffractions as well as their combinations. At first, all possible ray paths should be traced and subsequently relevant coefficients are calculated for each cases. On the presence of multiple point sources, $S_1, \ldots, S_{i_{\text{max}}}$, the total received field at a point of $\hat{r}$ is given as

\[
\tilde{E}_{\text{total}}(\hat{r}) = \sum_{S_i=1}^{S_{i_{\text{max}}}} \left\{ \sum_{p_1=1}^{p_{1_{\text{max}}}} \Pi^n_{i=1} R_e(i,p_1^i,S_i) \tilde{E}(p_1^i,S_i) \tilde{E}(r,p_1^i) \right\} 
\]

Here $S_i$ designates signal source while $p_1^{D,R}$ represents the point that a ray hits the first and repeats bouncing $n_R$, $n_D$ or $n_{<D,R>}$ times depending on the field characteristics and scatterer geometry. $\tilde{E}(p_1,p_2)$ represents $E$-field component at receiver $p_1$ when a point source is present at $p_2$. $r(p_1,p_2)$ is the distance between $p1$ and $p_2$. Each summation term in Eqn. (3.32) represents reflection, diffraction and combined fields in order. Considering the extreme degree of complexity of inner city structure, it is not practical to attempt to implement an accurate simulation, nor is it necessary. A useful approach will be to find a way to minimize field components in Eqn. (3.32), which may be facilitated by well simplified modelling.
Figure 3.12: The effect of diffracted field on wave propagation and its comparison with direct fields. (a) The role of diffraction field in shadow region without presence of any other direct fields (b) Three major backscattering paths via urban structure (c) Quantitative comparison with specular return fields

**Diffraction Fields**

One of the most distinguished contribution by GTD or its extended form UTD is the inclusion of the diffraction field originated from target discontinuity. The importance of diffraction manifests itself in radar signature when the existence of signal in radar shadow region is to be explained. It has been of particular interest in mobile communication where active researches are still on-going in regard to radar propagation modelling in urban environment [59][60][6]. Due to the high complex buildings surrounding residential and pedestrian area, in many cases, a proper operation of mobile communication relies on indirect diffraction field. Figure 3.12(a) shows the generation of signal field at radar shadow region. But the signal power of diffracted field is so weak that it is meaningful only when direct specular fields are not available.
As a means to describe diffraction field in SAR environment, three different ray paths are considered for a typical urban building in Figure 3.12(b), which comprises a mixture of diffraction and specular returns. A quantitative comparison has been made among three different return paths and shown in Figure 3.12(c). It is clear that the diffraction ray suffers from severe propagation loss compared with the specular return signal. The typical distance between target and spaceborne vehicle is around 800 km and the loss is expected be further below -60 dB. Eqn. (3.29) implies that the attenuation becomes severe at higher frequency. The specular-diffraction represents a diffraction field which is diffracted in the direction of reflection optical boundary and as a result behaves as a plane wave. It is found that its contribution is comparable to the specular return and should not be ignored. The comparison here is made between one single rays of each return path. When the total field are fully taken into account, the typical propagation loss of diffracted field reaches below -100 dB [6]. Especially, when a backscattering from cavity-like reflector coexists, the diffracted field is almost invisible [67]. The various claims developed by theoretical analysis, empirical data and the simulation results provide a good reason to justify the dominance of specular returns over diffracted field in SAR images of urban scenes.

Ground Surface

The ground surface is far from being perfectly conductive in real situation and the signal reflection behaviour on the ground must be adjusted during simulation. The most simple and direct way to handle this problem is to set a new constant for the reflection coefficient on the ground. In the urban area, the majority of the ground surface is covered with concrete or asphalt materials. In the smooth surface, the only amendment required to consider imperfect reflection is simply to replace the coefficient with a new constant, which can be given according to each specific materials. But in case of rough surfaces, the Snell's law of reflection is no more valid and the reflection coefficient becomes a function of incident angle rather than a fixed constant value.

Dynamic Range

The amplitude distributions presented in Figures 3.8 and 2.9 lead to the belief that man-made structures are responsible for the extreme level of brightness. The peak values are found to reach beyond brightness level of 2600 in amplitude while the dominant distribution formed by clutter area is limited below 70. Hence their typical reflectivity gap between the two environments reaches beyond $20 \log_{10}(2600/70) = 31.4$ dB. It should be reminded that strong return signals are power-limited by the system dynamic range before processing and becomes indistinguishable from other returns. Hence this graph does not reflect the true distribution of target reflectivities. Although these empirical data provides a rough insight into the dynamic range necessary to cover both urban and non-
residential area, still a theoretical measure remains unsolved on how to relate a specific target structure to the reflectivity level.

3.4 Urban Area Modelling

Traditionally, the image mapping on the urban area has been in the realm of aerial photography. The results of radar imagery application are often less than satisfactory in identifying objects in urban region. A good target classification requires clear view of the target in terms of resolution and contrast to the surrounding environments. The complex structure of the urban area where multiple strong scatterers are densely positioned makes it difficult to extract the necessary information without being spoiled by interference. A modest approach to understanding of the urban SAR images will be to collect vast amount of empirical database from familiar scene to be used later for comparison with the signal set obtained from the unknown region of interest. A theoretical prediction on the electromagnetic responses that relate the quantitative output to the corresponding target scene would certainly add convenience to this process.

3.4.1 Simulation Strategy

Many structures in urban area act rather like corner reflectors and thus return a high proportion of the radar energy back to the sensor [27]. Several urban SAR radar images lead to the conclusion that internal city structure could be determined because the different orientation and spacing of roads produce different effects on the radar images [10]. In this thesis, the radar signal reflection behaviour is investigated in terms of target response level since the brightness is the very criterion to start with in target classification process. A basic target structure is set up based on the corner reflectors. Several criteria are considered as target parameters, which include height, width, orientation and spacing between buildings. By changing these parameters to take into account various dynamic features found in urban area, the backscattered level from the target scene is measured, and subsequently the required signal dynamic range for proper interpretation is assessed. Figure 3.13 describes the suggested procedure.

Group urban structure is not considered since they involve a statistical interpretation of the target structures within the scene, which is not readily available and beyond the scope of this thesis itself. In addition to that, the lack of empirical data to compare with, may turn the attempt of simulation into obsolete and meaningless output.

*Enough research tends to support your theory. -Murphy's law of research*
Figure 3.13: A description of the procedure to analyze the relationship between target structure of corner reflector type and its reflectivity level in SAR environment.
Chapter 4

WAVEFORM DESIGN

Simple pulsed radar is limited in range sensitivity by the average radiation power, in range resolution by the pulse length. The design of any radar always involves a compromise between the two factors. Waveform design aims to seek an appropriate compromise that best suits the relevant application.

4.1 Pulse Compression

In general, the performance of radar detection is mainly dependent on the energy of the illuminating pulse. A good target detection at long distance range requires a high transmit power. However, to get a good range resolution, a large bandwidth is required, which, in other words, means that pulse time width should be shortened. In addition to that, the increased pulse length of radar signal will accompany highly sensitive response to the frequency Doppler shift which will deteriorate detection performance. More than one matched filter would be required to compensate this effect by tuning each filter into multiple bands respectively [69]. As a result there are two conflicting demands to achieve both good detectability and resolution simultaneously.

A large bandwidth implies more information can be delivered from a target area, which is essential for SAR where the wide target area generates a great amount of information [16]. High range resolution via a short pulse is required for any information to provide practical meaning. Especially in SAR environment, a short pulse is useful in that it can reduce the clutter effect which may prevent to get a clear image of ground area and is also helpful for target classification. The most common form of producing RF radar wave is offered by making use of TWT (Travelling Wave Tubes) of which available peak power is limited [72]. As a result the time duration should be extended to deliver sufficient power to pulsed signals. In general the bandwidth $B$ is defined as an interval over frequency band where the amplitude of the spectrum is kept sufficiently high (usually -3dB) in comparison to the peak one. In simple radar systems the time-bandwidth product $BT$ is always 1. Naturally, it imposes a compromise on radar application between resolution and signal
energy.

The pulse compression theory has been introduced in order to get a high range resolution as well as a good velocity Doppler shift detection. A large time-bandwidth product ($BT > 1$) is achieved by employing frequency or shift modulations. The received pulse is processed through high-gain matched filter by pulse compression algorithm. The matched filter output is in fact an autocorrelation function of the input signal compressed into a short pulse. Accordingly, the pulse generation algorithm or modulation method should be known in advance. There exist several modulation and pulse compression methods each of which has its own unique characteristic and application area. Linear FM modulation is among the most prominent waveform and has been adopted in most SAR systems. The meteorological radar is also another good example where pulse compression technique is found to be very useful by providing additional sensitivity and accuracy, and by reducing data acquisition times [54]. To compensate some disadvantages such as high peak side-lobes, non-linear FM or weighted amplitude modulation have been developed but with the sacrifice of some essential properties such as mainlobe magnitude, range resolution and sidelobe level due to mis-matching in the receiver filter. The time-bandwidth product $BT$ concept which has been introduced to represent the degree of pulse compression is equivalent to the pulse compression ratio. System specification is usually measured with this ratio which is typically more than 100. The increase of $BT$ contributes to improved power capacity inside received pulse without direct increase of peak power from the radar transmitter part. Therefore it is desired to achieve higher $BT$ which is a measure of the waveform performance. Although it is ideal to have pulse shape with rectangular envelopes in both time and frequency domains, due to their closely related relationship, such an implementation is known to be impossible in real situation [15].

4.1.1 Optimal Matched Filter

It is useful to maximize the signal output energy relative to the amount of noise (SNR) in the presence of noise which gives rise to a loss of signal information. When detecting radar targets, the presence or absence of a target is determined by a threshold test on the energy in the received signal. In radar target detection theory matched filter is defined as, given the set of waveform, a receiver filter that provides the maximum obtainable SNR with the presence of stationary additive noise. The matched filter is a function of the transmitted waveform and generally derived on point target condition. When a signal $S(t)$ of arbitrary choice is transmitted and scattered back from a target at some range $R$, then the signal at the receiver input $r(t)$ may be expressed as

$$r(t) = Ae^{-j\phi}S(t - \frac{2R}{c}) + N(t)$$  \hspace{1cm} (4.1)
where $Ae^{-j\phi}$ is a constant dependent on the target characteristic and signal propagation. In general radar environments, $N(t)$ is assumed to be a zero-mean additive Gaussian noise process of a constant power spectral density $N_0 [W/Hz]$ and statistically independent of both the transmitted signal $S(t)$ and target impulse response $h(t)$ at the receiver. The scattered signal output $G(t)$ is a finite-energy random process and given by the convolution integral

$$G(t) = \int_{-\infty}^{+\infty} \left( Ae^{-j\phi}S\left(\frac{2R}{c} - \tau\right) + N(\tau) \right) h(\tau)d\tau \quad (4.2)$$

The output $G(t)$ can be separated as signal and noise related terms respectively. Using the Schwarz's inequality, the SNR can be represented as

$$\left(\frac{S}{N}\right)^2 = \frac{\left[\int_{-\infty}^{+\infty} Ae^{-j\phi}S\left(\frac{2R}{c} - \tau\right) \cdot h(\tau)d\tau\right]^2}{N_0 \int_{-\infty}^{+\infty} h^2(\tau)d\tau} \leq \frac{A^2 \int_{-\infty}^{+\infty} S^2\left(\frac{2R}{c} - \tau\right)d\tau \cdot \int_{-\infty}^{+\infty} h^2(\tau)d\tau}{N_0 \int_{-\infty}^{+\infty} h^2(\tau)d\tau} \equiv 2E \quad (4.3)$$

From the Schwarz's inequality property, the maximum $(S/R)_{\text{max}}$ is achieved when $h(\tau) = kS\left(\frac{2R}{c} - \tau\right)$ is satisfied where $k$ is a constant. Hence the maximum realizable SNR is purely proportional to the energy of original signal. The convolution of Eqn. (4.2) is transformed into correlation operation when a matched filter is used at receiver. Although here is considered only the range variation, the matched filter concept can extend into two dimensional function after taking into account Doppler shift effect.

**General Target Detection**

In addition to simple target detection, the extraction of target information from the received waveform is also another important operation that radar can contribute to. Although the derivation of Eqn. (4.3) is based on the target detection performance only, it is accepted as a good measure for assessing target information as well on the ground that 'a good detection device is usually a good radar for extracting information and vice versa' [72]. It is true for most radar applications and has been widely accepted without questioning. However, when the target in question is modelled as extended radar targets having a certain pattern as is the case in urban area SAR image, the assumption of simple point target used in deriving Eqn. (4.3) may not be valid any more. Extended radar targets exhibit interference and resonance effect instead of simple linear superpositions of point target returns. Hence to achieve a true optimal SNR, the receiver-filter would have to be matched to the group waveforms scattered by the distributed targets, not the transmitted waveform itself. Bell [7] attempted to find a connection between target characteristic and the optimal waveform by referring to information theory.

Let $h_t(t)$ is the impulse response from a given target. Then the optimal matched filter
4.1 Performance Assessment

design will be equivalent to find a waveform \( s(t) \) and an optimal filter \( \xi(t) \) such that the mutual information \( I[r(t) \left| f(t); h(t) \right| s(t)] \) is maximized. It has been shown that, for a given waveform \( s(t) \) with its corresponding frequency spectrum \( \overline{S(f)} \), the optimum receiver filter spectrum \( \xi(f) \) is given as

\[
\xi(t) = K \int_{-\infty}^{+\infty} \frac{\overline{S(f)} \overline{H(f)} e^{-j2\pi f(t_0-t)}}{S_{nn}(f)} df
\]  

(4.4)

where \( K \) is a complex constant and \( s(t) \) is an eigenfunction having energy \( E_x \) and corresponds to the maximum eigenvalue \( \lambda_{max} \) of the integral equation

\[
\lambda_{max} s(t) = \int_{-T/2}^{T/2} s(\tau) \int_{-\infty}^{+\infty} \left| \frac{\overline{H(f)}}{N(f)} e^{j2\pi f(t-\tau)} \right|^2 df
\]  

(4.5)

Then the final signal-to-noise ratio becomes \( \lambda_{max} E_x \) at \( t = t_0 \) [7]. It is shown that when target impulse response is estimated in advance, a significant SNR gain can be achieved using the waveform/optimal-filter design method. Eqn. (4.5) has multiple solutions for various eigenvalues and each eigenfunction produces maximum SNR when its corresponding optimum filter is used. However, since most of the optimum solutions for Eqn. (4.5) generate amplitude variant functions, which is not desirable for transmitter amplifiers operating in saturation, the practical realization may be difficult. Despite of that, it delivers a significant implication that conventional matched filter concept may be altered depending on the target in view so as to bring about further improvement in SNR performance. Since SAR image contains wide dynamic range of target distribution and there are unique solutions for optimum receiver filters corresponding to different target characteristics, a possible suggestion is raised that multiple choice of waveform-receiver filter sets may be utilized to enhance information extraction performance for certain target structure.

4.1.2 Performance Assessment

Figure 4.1 is a typical pulse compression procedure and illustrates a range resolution improvement of target response from an ideal point scatterer. The performance of radar or more precisely, of a pulse compression system are usually measured in four categories.

- **Mainlobe broadening** - Related with the resolution and discrimination power
- **Peak signal to sidelobe ratio (PSLR)** - The maximum level of spurious response of pulse compression relative to the peak
- **Integrated sidelobe ratio (ISLR)** - The energy ratio between integrated sidelobe and mainlobe
- **The losses in signal to noise ratio** - Energy loss that costs non-ideal pulse compression scheme
4.1 Performance Assessment

Figure 4.1: Pulse compression concept illustration and waveform parameters

- **Doppler sensitivity** - Disruption of processed output due to the movement of either target or antenna

Mainlobe width is defined as the -3 dB or half power width relative to the peak response level. It determines the resolving capability of the radar system. -3 dB is set as the minimum fall-off from the peak level to provide discrimination between two separate point targets in range domain. It should be reminded that the minimum resolution width is decided on the basis of intensity level only, excluding the phase information. But as Rihaczek [68] pointed out, if the phase is also considered as well as intensity level, the range resolving capability between two scatterers is no more dictated by the 3-dB width of the mainlobe. Their phases can be formed to hinder a proper discrimination and the resolution can be deteriorated by half. Thus, for coherent detection, 6-dB points may be used as a true resolution criterion. Since our interest lies in SAR images, only intensity information from extended target is mainly concerned. Time resolution improvement will be given as $\frac{T}{\tau_{3dB}}$, where $T$ is the pulse width and $\tau_{3dB}$ is the 3 dB width of pulse compression output.

The PSLR is one of the most concerned and hence actively investigated property when a choice for waveform design or pulse compression scheme is made. The presence of range sidelobes degrades the radar performance by masking weak signals around strong targets. It may not be critical when SAR observation is limited to natural environments only. But when extremely strong scatterers are present in the scene comprised with weak reflectors, high PSLR is particularly hazardous. The lower PSLR is, the wider the dynamic range of detectable target becomes. In fact most of the performance improvement schemes for pulse compression focus on reducing this undesired peak sidelobe level minimizing side effects such as mainlobe broadening.
The integrated sidelobe ratio has attracted relatively less attention compared with the PSLR. It is due to the fact that most of the energy in processed output is already centred around mainlobe area (typically 98~99%), and unlike the peak sidelobe level the impact of integrated energy on overall target detection performance is not significant. But it can be used as a useful measure to assess the efficiency of the waveform compression. This property can be compared to the concept of variance in probability theory in that it attempts to measure the efficiency of energy distribution. The ISLR is not necessarily the ratio between inside and outside precise 3 dB boundary. It is rather related with the fall-off rate of the mainlobe. In this thesis, ISLR will be derived as a mean to calculate the energy concentration around the location of detected target. When the pulse compression output function is given as \( \Omega(x) \) and \( D \) is the designated area around mainlobe, then ISLR is

\[
\text{ISLR}_{(dB)} = 20 \log_{10} \left( \frac{\int_D |\Omega(x)| \, dx}{\int_{Dc} |\Omega(x)| \, dx} \right)
\]

SNR is defined as the ratio of peak signal energy to the average output noise power. For a frequency weighting function given as \( W_f \), the output signal-to-noise ratio is

\[
\frac{S}{N} = \frac{|G(t)_{\text{max}}|^2}{N_0 \int_{-B/2}^{B/2} W_f(f) \, df}
\]

Although the matched filter provides maximum SNR in theory, it is not always desired to adopt perfectly matched filter. In many cases the compression filter is required to undertake adjustment depending on its application purpose. Most of the sidelobe reduction schemes make use of weighting techniques either in time or frequency domains to generate compromised results and the performance deviation from ideal matched filter concerning processing gain is inevitable. SNR loss will be a useful measure in assessing the modified matched filter performance.

The introduction of the frequency Doppler shift into the incoming signal causes mismatching effect and directly results in pulse compression performance degradation. Unlike mis-matching effect caused by arbitrary noise signals, the coherently accumulated Doppler shift along time axis often produces a strong response at the processed output.

### 4.1.3 Ambiguity Function

The ambiguity factor is the mathematical representation of the relationship between two unknown parameters, the range and velocity of particular targets. Since ambiguity function is uniquely decided for each waveform it is often used as a measure to compare performances of various signal used in radar systems. The main properties of ambiguity function are characterized by limited peak power at the origin, limited volume over the
4.2 Linear FM

entire time-frequency domain and the rotational symmetry.

Pulse compression is affected by the movement of a target as well as its range. For example, when a linear FM signal is involved in target detection, the finite velocity of the target in range direction causes target displacement by shifting the frequency sweep inside signal. In phase codes case Doppler shift can cause critical distortion at the matched filter output and deter correct target detection. Ambiguity function is intended to evaluate the Doppler shift effects on pulse compression out of the time-shifted signal. It can be taken as a measure of how accurately different targets can be resolved without causing ambiguity. A general form of ambiguity function is expressed as a function of time delay $\tau$ and Doppler shift $f_d$, and given as

$$\chi(\tau, f_d) = \int_{-\infty}^{+\infty} s_1(t)s_2^*(t+\tau)\exp[-j2\pi f_d t]dt$$  \hspace{1cm} (4.8)

In conventional definition $s_2(t)$ is normally set to be equal to $s_1(t)$ and $\chi(\tau, f_d)$ is a pure evaluation for waveform $s_1$. In this thesis $s_2$ is assumed to have arbitrary forms to take into account mis-matched filter case and $\chi(\tau, f_d)$ will be used to assess the overall pulse compression scheme.

It has been clear since the time of Woodward that the goal of producing radar waveform having a prescribed ambiguity function is not realizable one [80]. Although a great deal of effort has been devoted to this subject, no algorithm has been developed yet for constructing waveform with arbitrary ambiguity surfaces [46]. Due to this fundamental dilemma, ambiguity function is seldom used in radar design stage but regarded as a useful guideline to evaluate waveform characteristics. Instead of attempting to create a new signal possessing an ideal property many works have evolved to synthesize a new set using multiple separate waveform signal [57][19].

4.2 Linear FM

A linear FM waveform, which is also called as chirp waveform due to its similarity to the audio frequency generating chirp sound, is expressed as

$$u(t) = a(t)\cos\left\{2\pi(f_0 t + \frac{k}{2} t^2) + \phi\right\}$$  \hspace{1cm} (4.9)

Here, $a(t)$ is a pure amplitude modulation function and $k$ is a constant which decides the pulse compression ratio. $f_0$ denotes the carrier frequency in the system. For a chirp signal of time duration $T$ and bandwidth $B$, $k$ is given as $\frac{B}{T}$. The instantaneous frequency is
found by differentiating the phase with time which results in

\[ f(t) = \frac{1}{2\pi} \frac{d(2\pi kt)}{dt} = kt \quad (4.10) \]

The pulse generated with this frequency sweep has a dispersion frequency characteristic linearly varying along the time axis. The matched filter aims to correct the frequency dispersion and eventually to recover the original waveform. The time-bandwidth product \( BT \) is also called as a dispersion factor. In many cases it is convenient to adopt complex notation to express signal. If the amplitude envelope \( a(t) \) is simply taken as a non-weighted rectangular signal with time duration \( T \) and discarding the phase term, then \( u(t) \) is expressed in complex form as

\[ u(t) = \text{rect}(t/T) \cdot \exp(j2\pi[f_0 t + \frac{Bt^2}{2T}]) \quad (4.11) \]

The impulse response of matched filter on frequency domain is simply the frequency conjugate of the input pulse. This means there is no phase fluctuation at the peak mainlobe and as a result mainlobe amplitude reduction is prevented [68].

It can be shown that the time impulse response of the matched filter is just the time reverse of input waveform. Some mathematical manipulation leads to the ambiguity function of chirp waveform \(|x(t, \omega_d)|\) given as

\[ |x(t, \omega_d)| = \begin{cases} (1 - \frac{|t|}{T}) \cdot \text{sinc}(\frac{(T-|t|)(\omega_d+2\pi kt)}{2}) & \text{if } |t| < T, \\ 0 & \text{otherwise}. \end{cases} \quad (4.12) \]

Figure 4.2 shows the magnitude plot and energy distribution of linear FM ambiguity function. The shape of diagram indicates that there exists a linear ambiguity in the interpretation of target range and velocity characteristics. This ambiguity between the range and frequency occasionally gives rise to a range error of moving target in SAR analysis. If a scatterer is moving and the velocity components is along the line of sight of the radar, the associated Doppler shift will cause frequency shift which might be recognized as time delay so that the peak time response occurs at the wrong position. Consequently the moving target is located at the wrong place in the range direction in SAR image. Conversely this misplaced distance can be used to determine the velocity of the moving target.

### 4.2.1 Implementation

There are a number of ways to produce a linear FM waveform and to implement associated matched filter. They are largely divided into either active or passive techniques. It can be actively generated by direct use of high power modulation transmitter or by voltage controlled oscillator (VCO) where the frequency is variable with the linearly applied
4.2 Frequency Spectrum

Voltage. To excite a dispersive delay line is one of the examples of passive FM waveform
generation. The surface acoustic wave (SAW) device is the most common and practical
method known for this purpose. This device provides various advantages over any other
methods and can be easily manufactured with extremely low cost. Its simple structure,
small design size and highly reproducible property have been the main reasons for it to
be widely accepted throughout the long history of communication systems. It is also pos­
sible to generate by digital means which has an advantage of flexible bandwidth and time
duration while maintaining a good stability and relatively low error generation rate [23].

4.2.2 Frequency Spectrum

Ideally the signal frequency spectrum should be confined within its designated band­
width $B$. But since the frequency sweep is imposed linearly over finite time and the ampli­
tude envelope is rectangular shaped the spectrum of linear FM reaches out infinitely over
frequency domain. For the sake of good efficiency, which is the goal of pulse compression,
It is desired to restrict frequency spectrum within $B$. It can be shown that linear FM
spectrum equation contains Fresnel integrals, of which the argument is proportional to
$\sqrt{B T}$ [15].

The spectrum of linear FM given in Eqn. (4.11) is expressed using the complex Fresnel
integrals and given as

$$I(f) = \sqrt{\frac{1}{2|k|}} \exp \left[ -j \left\{ \frac{\pi(f-f_0)^2}{k} \right\} \right] \left\{ \begin{array}{ll}
C(X^+) + jS(X^+) + C(X^-) + jS(X^-) & k > 0 \\
C(X^+) - jS(X^+) + C(X^-) - jS(X^-) & k < 0
\end{array} \right. \quad (4.13)$$
4.2 Frequency Spectrum

Figure 4.3: Frequency spectra of linear FM signals having various time-bandwidth products $BT$.

where

$$C(X) = \int_0^X \cos \left( \frac{\pi x^2}{2} \right) dx \quad \text{and} \quad S(X) = \int_0^X \sin \left( \frac{\pi x^2}{2} \right) dx$$

(4.14)

and the parameter $X^\pm$ is given as

$$X^\pm = sgn(k) \left[ k \pm 2(f - f_o) \right] \frac{1}{\sqrt{2k}}$$

(4.15)

Figure 4.3 illustrates some examples of linear FM spectra for three different dispersion factors. The amplitude behaviours correspond to the prediction made from Eqn. (4.13) while the ripples in the spectrum are caused by the nature of the Fresnel integrals in Eqn. (4.14). When $BT$ or dispersion factor $D$ is over 100, approximately 98 to 99 per cent of the energy is confined within the bandwidth $f_o - B/2 < f < f + B/2$. This high efficiency of energy delivery is a great merit of linear FM.

To validate the use of linear FM signal in SAR processing, Figure 4.4 shows a frequency spectrum of received raw SAR signal data. The raw data is acquired by ERS-2 which transmits linear FM signals of 15.5 MHz bandwidth and receives the returned signal at a sampling rate of 18.89 MHz. It is clearly seen that the range of spectrum extension for raw signal data corresponds to the original signal. Figure 4.5 shows an example of the raw data signal produced by one single pulse and compares it with the range processed data. Figure 4.5(b) is not meaningful in this stage since it is before the azimuth direction processing is implemented.
4.2 Frequency Spectrum

Figure 4.4: An example of frequency spectrum in ERS-2 SAR data. The bandwidth of the transmitted Linear FM signal is 15.5 MHz with the time duration of 37.12 $\mu$s

Figure 4.5: SAR raw data delivered by a single pulse and its range processed output
4.3 Sidelobe Suppression

Although the linear FM waveform achieves sufficiently satisfactory result it is not without its disadvantages. Compressed pulse of linear FM waveform has a relatively high sidelobe than usually required, which means it is far from the ideal characteristics of a simple spike pulse. Often sidelobe levels are required to be below -40 dB for surveillance applications on inhomogeneous areas [58] or even below -60 dB for meteorological radars [78]. But the typical PSLR of linear FM waveform is only around -13 dB. Furthermore, in SAR applications, sidelobe corruption often occurs due to the incorrect parameter estimation during processing and as a result the ISLR increases [61].

In an environment where clutter interference problems is persistent or multiple targets are densely scattered, it is important to suppress sidelobes into sufficiently low level. To prevent the target ambiguity due to high sidelobe level weighted amplitude functions have been introduced. Its basic principle is similar to the way a weighted linear array antenna works, which is used to reduce sidelobe or beam width of antenna pattern. The Hamming function, Taylor weighted function are good examples of some algorithms used for sidelobe suppression. They no longer work as matched filters and always involve unavoidable mismatching loss penalty which will diminish the high peak power at the compression output.

4.3.1 Source of Sidelobes

The ambiguity diagram surface has a finite volume proportional to signal energy $E$ and given as [72]

$$
\int \int |\chi(\tau, f_d)|^2 \, d\tau df_d = (2E)^2
$$

(4.16)

Although ideal ambiguity function $|\chi(\tau, f_d)|$ would consists of a single peak with height $2E$ at the origin and zero elsewhere, the conservation of energy does not permit this behaviour. The best approximation to idealized performance will be to construct a very narrow peak at the origin having a sharp fall-off rate and spread the remaining energy widely over $[\tau, f_d]$ domain. In practical terms, the sharpness of the peak response would be decided by specified resolution and the minimum peak sidelobe level could be achieved through uniform spread of the remaining volume.

While the theoretical nature of ambiguity function enforces a compromise between resolution and sidelobe level, imperfect matched filtering deteriorates sidelobe characteristic further. In actual radar systems various error sources are present and as a result, distortional effects perturb the system response. Let the system transfer function $H(\omega)$ be defined as

$$
H(\omega) = A(\omega) \exp[-jB(\omega)]
$$

(4.17)
4.3 Weighting Schemes

where \( A(\omega) \) and \( B(\omega) \) are amplitude and phase terms uniquely specified by the system. When the system deviates from ideal one, both \( A(\omega) \) and \( B(\omega) \) are no longer constant and frequency variant. MacColl\(^1\) considered single frequency terms of the Fourier expansion expressions for both amplitude and phase, which result in

\[
A(\omega) = a_0 + a_1 \cos(\omega)
\]

\[
B(\omega) = b_0 - b_1 \sin(\omega)
\]

For an input signal \( S(t) \), the transfer function of Eqn. (4.17) generates pulse compression output \( \Omega(t) \) of

\[
\Omega(t) = a_0 \left\{ \sum_{n=1}^{\infty} J_0(b_1)[(1 + \frac{na_1}{a_0 b_1})s(t - b_0 + nc) + (-1)^n (1 - \frac{na_1}{a_0 b_1})s(t - b_0 - nc)] \right\}
\]

\( J_n \) is a \( n \)th order Bessel function and originated from the \( \sin \) term in phase components. Since distortion effects can be described with Fourier expansions, Eqn. (4.20) may be extended to include general system distortion cases. Eqn. (4.20) provides not only a physical insight into the physical system response but also some clues leading to the development of sidelobe suppression theory.

4.3.2 Weighting Schemes

When the coefficients of \( a_1 \) and \( b_1 \) are small in Eqn. (4.20) and the Doppler frequency term \( b_0 \) is removed, then using the Bessel function approximation property\(^2\) Eqn. (4.20) can be simplified into

\[
\Omega(t) \approx a_0 \left[ S(t) + \frac{1}{2} (\frac{a_1}{a_0} + b_1)s(t + nc) + \frac{1}{2} (\frac{a_1}{a_0} - b_1)s(t - nc) \right]
\]

A brief interpretation of Eqn. (4.21) is that a small sinusoidal perturbation in the matched filter additionally produce a single pair of echoes, which are amplitude scaled and symmetrically positioned from main response. The linear FM signal produces sinc type output at the matched filter. Hence, with appropriate choices for \( a_0, a_1 \) and \( b_1 \), the scale factors and time shift parameters in Eqn. (4.21) may be adjusted such that the first peak sidelobes are cancelled out. Figure 4.6 illustrates this procedure.

Cosine weighted functions are directly resulted from this approach. A general expression is given as

\[
W_f(f) = \alpha + (1 - \alpha) \cos(2\pi \frac{f}{B})
\]
Figure 4.6: Sidelobe suppression mechanism. Signal response ripples are cancelled out by leading and lagging echoes.

When the same weighting is given to both real and cosine terms, $\alpha$ becomes 0.5 and generates the Han weighting. The Han weighting does not cancel out the first sidelobe peak completely. The Hamming function assigns 0.54 to $\alpha$ to achieve an effective peak sidelobe cancellation. The Dolph-Chebyshev is a closed form solution to the minimum mainlobe width for a given sidelobe level, which is uniformly distributed outside mainlobe. The time domain transformation of Dolph-Chebyshev function is shown as

$$W_{Dolph}(t) = \frac{\cos \pi \sqrt{(Bt)^2 - \beta^2}}{\cosh(\pi \beta)}, \quad -\infty < t < \infty \quad (4.23)$$

where $\beta$ is determined by the specified sidelobe level of $20 \log [1/\cosh(\pi \beta)]$. Since an infinite amount of energy is required (infinite time domain), the Dolph-Chebyshev is not realizable in practice although its performance may be a useful reference for other weighting methods to compare with. Although only one cosine term is used in the Eqn. (4.18), the degree of freedom in waveform design would increase by considering additional terms such as

$$W(f) = a_0 + a_1 \cos(2\pi \frac{f}{B}) + a_2 \cos(2\pi \frac{f}{B}) \ldots \quad (4.24)$$

The choices of the coefficients $a_k$ are at will and a well-designed set would deliver further improved sidelobe suppression capability. As the number of cos terms increases in Eqn. (4.24), the sidelobe level decreases and fall-off rate becomes faster but with an increased mainlobe width.

Both Han and Hamming windows involve only two terms. The Han function gives
4.3 Weighting Schemes

![Figure 4.7: Magnitude comparison of Dolph-Chevyshev, Hamming, Han and Blackman weighting functions in frequency domain](image)

<table>
<thead>
<tr>
<th>Weighting</th>
<th>Parameter</th>
<th>PSL [dB]</th>
<th>Pulse width $\tau_{3dB} \times B$</th>
<th>SNR loss [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dolph-Chevyshev</td>
<td>-40(pre-determined)</td>
<td>1.35</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Hamming</td>
<td>-42.8</td>
<td>1.47</td>
<td>-1.34</td>
<td></td>
</tr>
<tr>
<td>Han</td>
<td>-31.5</td>
<td>1.44</td>
<td>-1.5</td>
<td></td>
</tr>
<tr>
<td>Blackman[3-term]</td>
<td>-67</td>
<td>1.66</td>
<td>-1.71</td>
<td></td>
</tr>
<tr>
<td>Blackman[4-term]</td>
<td>-92</td>
<td>1.90</td>
<td>-2.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Processed output parameters after weighting is implemented

Equal weighting to each terms

$$W_{H_{an}}(f) = 0.5 + 0.5 \cos(2\pi \frac{f}{B})$$  \hspace{2cm} (4.25)

while the coefficients of Hamming windows are modified to achieve lower sidelobe level

$$W_{H_{am}}(f) = 0.54 + 0.46 \cos(2\pi \frac{f}{B})$$  \hspace{2cm} (4.26)

Blackman [8] has developed these further by taking into account more than two terms. Harris [26] devised a four term Blackman-Harris window given as

$$W_{B-H}(f) = a_0 + a_1 \cos(2\pi \frac{f}{B}) + a_2 \cos(2\pi \frac{2f}{B}) + a_3 \cos(2\pi \frac{3f}{B})$$  \hspace{2cm} (4.27)

which further minimizes the peak sidelobe level. With an appropriate choice for the parameters $\{a_1, a_2, a_3\}$, Blackman-Harris window achieves up to -92 dB of sidelobe suppression.

Figure 4.7 compares various amplitudes corresponding each weighting functions in frequency domain. It should be pointed out that the mainlobe area is related to the SNR. Figure 4.7 indicates that the Dolph-Chevyshev weighting is superior to others while the Blackman window shows the worst result in terms of SNR loss. Figure 4.8 shows thus processed outputs generated by various mismatched filters. The unavoidable compromise between the mainlobe width and sidelobe level is clarified throughout the pulse compression outputs.
4.3 Weighting Schemes

(a) Dolph-Chevshev

(b) Hamming output

(c) han

(d) Blackman(4 terms)

Figure 4.8: Pulse compression outputs after various spectrum weighings are applied. The PSL for Dolph-Chevshev is -40dB. 4 term Blackman window parameters are $a_0 = 0.35875$, $a_1 = 0.48829$, $a_2 = 0.14128$, $a_3 = 0.01168$. 
A comprehensive review and comparison on various weighting schemes is provided by Harris \cite{26}. Although the conventional way to compare the mainlobe characteristic is simply to measure -3 dB or -6 dB points from the peak, it fails to reflect the excessive mainlobe expansion such as in Blackman weighting case. The pulse width data shown in Table 4.1 appear not able to deliver sufficient description on the actual graphs in Figure 4.8. It may not be critical in ordinary point target detection. But it becomes necessary to measure the effect of the energy distribution around mainlobe originating from a strong echo as the density of scattered targets grows.

In practice those strong suppression filter mentioned here does not provide as good performance as theoretically predicted due to the in-band ripples, which is caused by the non-ideal behaviour of hardwares involved. For example, Tortoli \cite{62} claims that -48 dB is the maximum sidelobe suppression level attainable by use of Blackman-Harris weighting contrary to the theoretical value of -92 dB. A more precise signal control is necessary to gain pulse compression results that approximate to theoretical values. Digital design approach, instead of SAW filter methods, may be more attractive in this sense. Tortoli has shown that an improved performance is achieved by modifying matched filter to introduce ripples such that spectral ripple in the received signal can be compensated.

The frequency domain function is simpler and thus easier to handle than time domain weighting functions. This explains the dominance of the frequency weighting methods compared to the time domain operation. Since the RF amplifier in transmitter normally operates in saturation mode weighting is implemented at the receiver.

### 4.3.3 Non-Linear FM

Contrary to the weighting approach, the non-linear FM waveform design adopts non-linear variation of frequency component versus time. Since no weighting process is involved low sidelobe levels can be achieved without causing critical mismatching loss at the matched filter output. But because of the non-linear relationship between the time and frequency the system becomes more sensitive to the Doppler shift of moving target. Due to this sensitivity non-linear FM is useful only when there is not serious Doppler shift of fast moving scatterer in the target area. Again it is a matter of compromise.

### 4.4 Discrete Coded Waveforms

In phase coded pulse waveforms, a long pulse of finite duration $T$ is divided into $N$ subpulses and each subpulse contains its unique phase information assigned by the pulse
4.4 Barker Code

<table>
<thead>
<tr>
<th>Code Length</th>
<th>Code elements</th>
<th>Sidelobe level(dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>+-, ++</td>
<td>-6.0</td>
</tr>
<tr>
<td>3</td>
<td>++-</td>
<td>-9.5</td>
</tr>
<tr>
<td>4</td>
<td>+++, ++-</td>
<td>-12.0</td>
</tr>
<tr>
<td>5</td>
<td>+++, ++-</td>
<td>-14.0</td>
</tr>
<tr>
<td>7</td>
<td>+++, ++-</td>
<td>-16.9</td>
</tr>
<tr>
<td>11</td>
<td>+++, ++-</td>
<td>-20.8</td>
</tr>
<tr>
<td>13</td>
<td>+++, +++, +</td>
<td>-22.3</td>
</tr>
</tbody>
</table>

Table 4.2: Barker Code Sequences

The general description for phase code waveform is given as

$$\psi(t) = \sum_{n=1}^{N} a_n P_n(t) \exp[j\{(\omega_0 + \omega_n)t + \theta_n\}]$$

$$= 0 \quad , \text{elsewhere}$$

(4.28)

In most cases the $P_n(t)$ is a unit amplitude pulse of fixed duration $\frac{T}{N}$ between the time interval of $(n-1)\frac{T}{N} \leq t \leq n\frac{T}{N}$. This discrete codes are divided into several groups depending on the variation of parameters $a_n$ and $\theta_n$.

Generally, this type of waveforms require much higher complexity in system implementation than linear FM case. And the fact that its performance is highly vulnerable to the Doppler shift in the returned pulse is another reason for its less popularity compared to linear FM. But in terms of sidelobe level characteristic which is very essential part in many waveform design, the phase codes are known to have attractive property. The range sidelobes of discretely coded pulse exhibits distinguished behaviour from linear FM's. Its direct compatibility with digital system is another great advantage. Here will be given some examples of this type of waveforms. Their performances will be investigated and compared with each other mainly concerning about the sidelobe pattern.

4.4.1 Barker Code

The Barker code is one of the simplest among the discrete waveform group but its unique sidelobe characteristic implicates a very important property.

The Barker code is a binary sequence containing only two different phase components, 0 and $\pi$. The importance of this code lies in its perfect sidelobe level property. The 'perfect' is a term used in the phase code pulse compression area to describe the theoretically minimized sidelobe level. Despite the continuous search for waveforms having this perfect property, only the Barker code sequences are known to have the true 'Perfect' sidelobe property. It is known that only nine Barker sequences exist of which the maximum length
4.4 Maximal Length Sequences Code

The Maximal Length Sequence (MLS) is a binary code class easily generated by a simple algorithm using shift registers. Although its sidelobe level property is not as optimal as that of the Barker code the code length is not limited. The hardware implementation for generating this code is shown in Figure 4.10. The \( m \)-stage shift register has a \( 2^m \) possible states in total. Excluding the zero state which can not contribute to generating any code bit, it can produce up to \( 2^m - 1 \) bits code sequence. When the output sequence is of period \( 2^m - 1 \), then it is called a ‘Maximal Length Sequence’. By adjusting the size of shift register and feedback pattern various length of codes can be obtained. The peak sidelobe level is not precisely related to the length of pulse codes.

To find a MLS code with relatively good performance is not straightforward. The ACF is limited to 13. Table 4.2 shows all the known Barker codes with their peak sidelobe levels. All the ACFs of the Barker codes produce uniform sidelobes at the matched filter outputs while a peak response is produced at zero time lag. The uniform sidelobe level is inversely proportional to the code length. The ACF of Barker code of length \( N = 13 \) is shown in Figure 4.9(a). The peak sidelobe level is derived from the formula \( 20 \cdot \log_{10}(13) = 22.27 \text{ [dB]} \). Due to its extremely limited code length realizable in practice the Barker code is not suitable for a radar system where a large pulse compression ratio is required. Also the high sensitivity to Doppler shift effect as described in Figure 4.9(b) manifests the disadvantage of using discrete phase codes. It is seen that the peak level of ACF diminishes with the increase of the Doppler shift level. However its perfect sidelobe property is often used as a reference in assessing the performance of arbitrary phase code signals.

4.4.2 Maximal Length Sequences Code

Figure 4.9: (a) Autocorrelation function of Barker code when \( N=13 \) (b) Ambiguity function plot for Doppler shift range \(-0.05B < f_d < +0.05B\).
4.4 Maximal Length Sequences Code

Figure 4.10: $M$ bits shift register and feedback algorithm to generate MLS codes

Figure 4.11: Pulse compression results from MSL. (a) Autocorrelation functions for MLS with length 511 (b) Cross-correlation function for two different MLS code
for MLS of length 511 ($m = 9$) are given in Figure 4.11(a). It is shown that the Doppler shift frequency given as a small fraction of pulse bandwidth $0.01B$ results in a severe malfunction. The ACF has relatively flat sidelobe pattern with its peak level around $-30$ dB which is much higher than optimal Barker code level of $20 \log_{10}(1/511) = -54$ dB. The strong sensitivity to Doppler shift means the application area is very limited. The cross-correlation function for different MLS codes with the same length are shown in Figure 4.11(b), which shows a favourable performance for waveform diversity application.

### 4.4.3 Frank Code

Instead of utilizing only two phase components as in binary codes, polyphase codes utilize multiple phase components to construct sequences. Although it is more complicated in terms of hardware implementation the flexible use and combination of multiple phase parameters means that waveform parameters can be adjusted to suit various applications. The Frank code belongs to this type of polyphase code and is also called as a quantized linear phase FM due to its similarity of phase change to linear FM. Its sequence is comprised of $N^2$ phase elements and conveniently expressed as $N$ divided groups each having $N$ different elements. When two indices $i$ and $j$ are assigned for the element of $i$th sequence in $j$th group, the general description for this waveform is given as a double summation

$$\psi(t) = \sum_{j=1}^{N} \sum_{i=1}^{N} p[t - ((i - 1) + (j - 1)N)T] \exp[j(\omega_0 t + \frac{2\pi}{N}(i - 1)(j - 1))]$$ (4.29)

Figure 4.12(a) shows the phase variations of the Frank codes for $N = 5$ and $N = 10$ cases. It is evident that the phase variation pattern resembles that of linear FM. The total time duration of the waveform is $N^2T$ and the effective bandwidth $B$ is given as $\frac{1}{\tau}$, which results in the pulse compression ratio of $N^2$.

Figure 4.12(b) shows the autocorrelation function of the Frank code when the pulse compression ratio is set as 400 or $N = 20$. The peak sidelobes appear at both ends of the sidelobe region at a level of $-37$ dB which is disappointingly high compared to the Barker code level of $20 \log_{10}(400) = -52$ dB. It is also seen that, with the presence of Doppler shift frequency, ACF is disrupted by several sharp spikes appearing along the time axis.
4.4 Huffman Code

Huffman [19][36] defined a class of finite complex sequence \{a_n\} of length \(N+1\) having the nearly ideal aperiodic ACF given by

\[
\Omega(\tau) = \sum_{n=0}^{N-1} a_n a_{n+\tau} = \begin{cases} 
\sum_{n=0}^{N} |a_n|^2, & \tau = 0 \\
0, & 0 < \tau < N \\
\frac{C(0) \zeta^{-N}}{1 - \zeta^{-2N}} \tau N, & \tau N
\end{cases}
\]

where \(\zeta\) is a real number defined for each sequences. All the sidelobe responses are completely eliminated except at \(\tau N\). Although the general theory of Huffman sequence is straightforward, the selection of the sequence constitute a very difficult problem. Most of all it has been shown that precisely uniform amplitude Huffman sequences do not exist. Hence the use of Huffman code is limited in a very limited application area. The fact that Huffman codes are amplitude modulated signals is another reason to discourage their use in radars, in which transmitters with nonlinear characteristics are typically operated at saturation for transmission energy efficiency.

4.5 Waveform Design for SAR

The ‘optimum filter’ is a broadly defined term that usually means minimum loss in resolution and signal detectability for a maximum reduction of the interfering range sidelobes [15]. A bright point scatterer does not appear merely as a single bright point in SAR image because of SAR’s finite bandwidth, finite synthetic aperture length and pulsed nature. Apart from a strong mainlobe a single point scatterer also generates sidelobes and
4.5 Ambiguity Rejection

spurious response called ambiguities. A good SAR system should minimize these unwanted responses.

4.5.1 Ambiguity Rejection

Unfortunately both range and azimuth ambiguities are originated from the pulsed nature of the SAR operation and regarded as unavoidable phenomena regardless of the type of waveform (: refer to Figure 2.5 and 2.7). A number of techniques have been suggested that attempt to overcome this problem and in general they are classified in three different approaches, which are

1. Removal of ambiguous signal during post-processing
2. To adopt a flexible system - multiple antennas
3. Selecting multiple PRFs for either a single or multiple waveforms

To remove a strong ambiguous signal in later stage is the most straightforward method to deal with ambiguity problem. When an ambiguous response is spotted, its source response can be traced in terms of amplitude and phase. Since the relevant ambiguity responses happen in predictable manner they can be eliminated given the information of the ambiguity source [56]. For this technique to work it should be assumed that the point target response is isolated and strong enough to be visible in SAR image, which is not always the case. Especially this simple approach cannot handle general situations where, such as in urban environment, multiple strong scatterers are densely crowded.

To adopt multiple antenna system is another possible option although complex hardware structure has to be involved [47][24]. It attempts to reduce the received ambiguous signal energy by steering antenna or using separate antennas for different waveforms. Thus the waveform design itself does not play a significant role in this technique.

Since the ambiguities are closely related with PRF, many efforts are made to find a suitable choice of PRF that best compromises between opposing characteristics of ambiguities. Using multiple PRF, suggested by Ferrari [1], adopts a scheme that sends out pulses in multiple blocks each of which has their own unique PRF rates. To use multiple waveforms is obsolete unless a good separation is made between different waveforms in use. Callaghan [14] pointed out and proved that the nature of SAR operation does not allow sufficient level of waveform isolation, regardless of the waveform set. This is a predictable conclusion in that the energy conservation rule is still valid in pulse compression. In distributed target environment the ambiguous signals from extended targets affect wide area accumulating unwanted signal energy. Therefore unless a technique to reduce the compressed output energy becomes feasible, the multiple waveforms technique is of little use for SAR application.
4.5 Ideal Response Performance

Since SAR is more than a simple target detection system, there are other factors to consider than general waveform criteria described earlier. Obviously, to hold a good resolution property is an important criterion. Although Eqn. (2.18) implies that azimuth resolution does not receive direct benefit from a waveform design itself, a good waveform design may relax the restriction on the PRF rate which will lead to the reduction of aliasing effect.

A significant amount of work on waveform design have focussed on sidelobe suppression, or more precisely reducing peak sidelobe level. However, a mere suppression of peak sidelobe level would not be enough in some cases when integrated sidelobe energy matters.

Most of the sidelobe suppression techniques involve mainlobe broadening and the performance of a good waveform tends to be judged by the degree of mainlobe width sacrifice. Surely, it is desired to minimize mainlobe width broadening. When strong scatterers are closely located over dark background, the sidelobe energy around mainlobe area will accumulate by a significant level. Hence, it will be highly beneficial to isolate mainlobe area from surrounding spurious sidelobe. Unfortunately, most of the weighting methods accompany mainlobe extension beyond 3 dB points.

From the energy conservation rule and with the assumption that the mainlobe peak level is fixed, the peak sidelobe is optimized to the best possible level only when sidelobe pattern follows uniform distribution. Another great advantage of this type of uniform sidelobe pattern is that the chance of false target detection due to the presence of a strong target response over dark background becomes extremely low. This is graphically described in Figure 4.13(a). In uniform sidelobe case, the local peaks that may be mistaken
as weak point targets are non-existent. In addition to that, since uniform sidelobes behave in a precisely predictable manner as a function of peak signal level, there is a good potential that these unwanted sidelobes may be removed systemically during SAR processing procedure. On the other hand it is also desirable to isolate the mainlobe response from its surrounding spurious signals for the purpose of enhancing image contrast level. With these number of factors considered, an example of optimal pulse compression output that may be suitable for SAR application is illustrated in Figure 4.13(b).

Considering their operational principles explained earlier in this chapter, it is very unlikely that this kind of output can be achieved from the conventional linear FM signals combined with various weighting schemes. However phase coded waveforms allow sophisticated treatment of the pulse compression output thanks to their discrete nature. The rest of thesis is dedicated for this type of waveform and development of pulse compression output resembling the optimal pattern introduced in this chapter.

'Sometimes it can be misleading, for there is no general theorem that maximum output signal-to-noise ratio insures maximum gain of information' - P. M. Woodward -
Chapter 5

ANALYSIS OF PHASE CODED WAVEFORM

This chapter gives a detailed insight into a relatively new phase code technique called P-code. It was first suggested in early '80s by Lewis [51] and further improved later [50]. Felhauer [20] has shown that P3 and P4 codes are special cases of so-called P(n,k) code. The P code has several useful features including ease of implementation, compatibility with digital technique, low cross-correlation between codes and Doppler tolerance of the frequency modulation codes.

Barker codes are known to have perfect sidelobe characteristics by forming ambiguity plots approaching the ideal ‘thumb-tack’ shape. Although their autocorrelation properties are favourable, their actual application is extremely limited due to the short realizable pulse length. The maximum length of this code known is 13, which is too short to deliver enough information. The new polyphase codes introduced in this chapter will be shown to generate peak-limited flat sidelobe level independent of effective pulse compression ratio.

Despite their favourable performance very little attempt has been made to explain the pulse compression performance of the P code. Here is provided a comprehensive insight into the nature of sidelobe generation process. In addition to this, a detailed description for sidelobe reduction is provided, which is to be used for derivation of a new filter concept in Chapter 6.

5.1 P-Code Theory

As is shown in previous chapter, it is the general property of phase codes that increasing pulse compression ratio of phase code gives rise to increased peak sidelobe ratio. This reflects the very clear contrast between linear FM and phase coding and will be taken as an important merit of phase coding over linear FM. Among phase code group, polyphase code such as Frank code is preferred to binary codes for their superior performance in
terms of Doppler tolerance and flexibility in selecting parameters during waveform design stage at the cost of increased system complexity. Like Frank code the P-code is also conceptually derived from a linear frequency modulation waveform (LFMW). The P(n,k) code is another type of conversion from linear FM signal enveloped with optimal weighting function in spectral domain [20]. Here will be introduced four different phase codes each of which will be named as P1, P2, P3 and P4 respectively.

5.1.1 P1 and P2 Polyphase Codes

Similar to Frank code, the P1 code is also derived from the step chirp signal but with modified phase variation which is mainly aimed at enhancing the autocorrelation function (ACF) property. Unlike Frank code, P1 code is generated by placing synchronous oscillator at the centre frequency of the step chirp IF waveform and sampling the baseband waveform at the Nyquist rate. As a result P1 code has \( N \) phase groups each of which also has \( N \) phase elements in it. The \( i \)th element of \( j \)th group is expressed as

\[
\phi_{i,j} = -\frac{(\pi/N)[N - 2j + 1][(j - 1)N + (i - 1)]}{3}
\]

where \( i \) and \( j \) are integers ranging from 1 to \( N \). The following sequence is an example of phase elements array of P1 code for \( N = 3 \),

\[
0 \quad -\frac{2\pi}{3} \quad -\frac{4\pi}{3} \quad 0 \quad 0 \quad 0 \quad 0 \quad \frac{2\pi}{3} \quad \frac{4\pi}{3}
\]

which is a mere rearrangement of Frank code sequence of the same length. P2 code has a similar structure with P1 but is only applicable for even number of \( N \). Each group has the same increment factor as P1 but with different starting point which discriminates the performance between the two codes. Like P1 code, the \( i \)th element of \( j \)th group of P2 code is expressed as

\[
\phi_{i,j} = -\frac{(\pi/N)(N - 1)/N - (\pi/N)(i - 1))[N - 2j + 1]}{3}
\]

where \( i \) and \( j \) are also integers from 1 to \( N \). It is easily noticed that both P1 and P2 code has the same increment of \(-{(\pi/N) * (N + 1 - 2j)}\) in \( j \)th group. When \( N \) is confined only to even numbers, P2 code always forms symmetry code structure.

The ambiguity diagram of P1 code is identical to that of the Frank code for odd \( N \). This is justified by the fact that P1 code is directly derived from the Frank code by rearranging phase elements and P2 code is formed by small phase steering of P1 code. In the ideal situation where no consideration on the bandwidth is necessary, P codes do not provide any superior merit over Frank code. But in real system where only finite spectrum bandwidth is available, P codes show far better performance in terms of sidelobe level. This comes from the difference of phase array between codes. The phase variation of P
code is relatively smooth in the middle of array while the highest phase increment between sample to sample shows up at the ends of the codes. In the Frank code exactly opposite happens. When waveforms are passed through bandwidth limited filter, the contribution from the high frequency parts are reduced or lost. After bandlimited spectrum is recovered, the Frank code is attenuated most heavily in the centre of the array while it works in the other way for P codes. Eventually this bring about a weighting effect which helps to suppress the sidelobe level of the P code signals.

It is clear that a mere manipulation of phase arrangement can bring about a remarkable impact on sidelobe performance in discrete phase codes. Hence a possibility is raised that, within a group of waveforms sharing similar characteristics, a good alternative may be found such that an enhanced pulse compression property is obtained.

### 5.1.2 P3 and P4 Polyphase Codes

One of the most critical disadvantages of phase coding against linear FM is the relatively poor Doppler tolerance. Although the P1, P2 code can be utilized to overcome bandlimiting effect, its ambiguity function shows no difference with that of Frank code in the presence of Doppler shift. The P3 and P4 code are specially aimed at dealing with this problem. As is the case with the P1, P2 code, the generation of these codes are also based on the linear frequency modulation waveform but with slightly different approach. The P3 code is generated by converting frequency modulation signal to baseband using a local oscillator and sampling I and Q phase signals at the Nyquist rate. The difference between P4 and P3 lies in the starting frequency of sampling sweep.

Typical linear FM frequency sweep is

\[ f = f_0 + kt \]  

where \( k \) is the frequency slope. The bandwidth of this signal \( B \) is represented as \( kT \) where \( T \) is the pulse length. The Nyquist sampling interval in the complex domain is decided by compressed pulse length as

\[ t_c = 1/B \]  

When the samples of the waveform are taken by this time interval \( t_c \), the \( i \)th phase \( \phi_i \) of the P3 code sequence is calculated as

\[
\phi_i = 2\pi \int_0^{(i-1)\Delta t_c} [(f_0 + kt) - f_0]dt \\
= \pi k(i - 1)^2 t_c^2 = \pi(i - 1)^2 / BT
\]  

\[(5.6)\]
The P4 code is generated through similar procedure but by making use of different carrier frequency $f_0 + kT/2$ for local oscillator instead of $f_0$. The shift of $kT/2$ in frequency, which is half of total bandwidth, is responsible for the performance difference between the P3 and P4 codes. The phase of this P4 code is given as

$$\phi_i = 2\pi \int_0^{(i-1)t_e} [(f_0 + kt) - (f_0 + kT/2)]dt$$

$$= \pi k(i - 1)^2 t_e^2 - \pi kT(i - 1)t_e$$

$$= \pi(i - 1)^2 / BT - \pi(i - 1) \quad (5.7)$$

The relationship between the P3 and P4 are very similar to that between the P1 and P2 in terms of the frequency increment. As a result the P4 code is more resistant to the finite bandwidth-limiting effect than the P3 code. Because these two codes are directly derived from linear FM there is no restriction on the number of total phase components, while those of the P1 and P2 codes must be set as $N^2$. The phase variations and their associated autocorrelation functions are shown in Figure 5.1.

Apparently these codes look like just no more than an analogue-to-digital conversion of linear FM. But they possess some unique characteristics which are only obtainable in discrete phased code format. An in-depth discussion is provided later.

### 5.2 Doppler Characteristic

The general sidelobe characteristics of P codes discussed in the previous section do not show significant advantages over other conventional phase codes such as Frank code or MLS (Maximal Length Sequence). But when a slight Doppler shift is introduced in the waveform, the P3 and P4 codes show excellent performances. Doppler shift effect should
5.3 Frequency Spectrum and Bandwidth Limitation

The P codes are formed by several contiguous pulses of which the individual pulse length is given as $T_b$. Hence the practical 3 dB bandwidth for this codes will be given as

$$f_{BW} = \frac{1}{T_b}$$  \hspace{1cm} (5.8)

It is worthwhile to refer to frequency spectrum characteristic to measure waveform performance in practical situation. Unlike linear FM waveform which utilizes frequency band
with great efficiency, the discrete phase coded waveform experiences a considerable energy loss during signal processing. The finite length of each pulse elements will generate frequency spectrum which extends to infinite frequency range. As a result the returned phase code signal does not match perfectly to the original waveform after bandlimiting is imposed at receiver filter. This bandlimiting would occur prior to sampling in the A/D conversion process in order to prevent noise foldover and aliasing.

The frequency spectrums of $P$ codes can be calculated by the following steps. When $\theta_n$ is given as a phase constant for the $n$th element of a $P$ code, the time domain expression of general $P$ code $\psi(t)$ is given as

$$\psi(t) = \sum_{n=1}^{N} \exp(j\omega_ct + \theta_n)P_n(t) \quad (5.9)$$

After this signal is shifted to baseband at the receiver, the frequency domain expression corresponding to this signal will be expressed as

$$\psi(f) = \int_{-\infty}^{\infty} \sum_{n=1}^{N} P_n(t)\exp(\theta_n)\exp(-j2\pi ft)dt$$

$$= \sum_{n=1}^{N} \exp(\theta_n) \int_{(n-1)t_b}^{nt_b} \exp(-j2\pi ft)dt$$

$$= \sum_{n=1}^{N} t_b \exp(j[\theta_n - 2\pi f(n - \frac{1}{2}t_b)]sinc(ft_b) \quad (5.10)$$

Figure 5.3 shows the results obtained by performing this calculation on $P_2$ and $P_3$ codes respectively. The results for the $P_1$ and $P_4$ codes do not differ significantly from these diagram. The frequency components above $\frac{f_b}{T_b}$ will be rejected in the receiver, which leads to smoothing effect of the pulses constituting the coded waveform. This bandwidth limiting effect usually causes a performance degradation in discrete phase coded pulse by generating unfavourable mismatch with pulse compressor. This performance degradation becomes severe where the phase increment is high, while relatively small change is detected for low phase increment interval. Eventually, this will appear as weighting effect posed on the returned pulse. In case of the Frank code, the phase variation is high around the middle of the waveform and becomes lower toward each end of the code where the increments approach $180^0$. This is the same as applying the amplitude weighting filter shown in Figure 5.4(a) and will do adverse effect on peak to sidelobe ratio performance. Because the $P_3$ code shares similar phase increment distribution with the Frank code, it will also experience the same negative effect in this aspect. But in the $P_1, P_2$ and $P_4$ codes where the phase increments are relatively high at the each end of the waveforms, the weighting effects happen in opposite manners. In fact this weighting effect helps to
5.3 Frequency Spectrum and Bandwidth Limitation

Figure 5.3: Power spectral density of P2 and P3 phase coded pulses

Figure 5.4: The effective amplitude weighting curve when finite bandlimiting is imposed on incoming signals
increase the peak to sidelobe ratio of the pulse compression output. It is concluded that the loss of partial signal energy residing in the spectrum sidelobe does not necessarily result in performance degradation.

5.4 ACF of P code

Although sidelobe reduction concept during post-compression has been first introduced by Lewis [50], this idea was derived intuitively from the assumption that P code sidelobe characteristic has the similar tendency of Frank Code’s. His claim was based on the fact that pulse compression output can only be changed by about or less than the magnitude of the last pulse into correlator. And in the case of single sideband derived codes like the Frank or P3 code, the successive samples of large sidelobes have nearly the same phase and differ from each other on the order of one code element. Thus it was predicted that sliding window two-sample subtractor existing at the output of matched filter would produce a postcompression output having maximum sidelobe magnitude level of 1. In fact it turned out that the maximum sidelobe magnitude is actually 2 instead of 1, for which an analytical proof will be given later.

To fully understand the algorithm involved in this procedure is very important not only in terms of academic purpose but for the reason that it would also provide a good theoretical motivation to improve or modify the pulse compression performance as wished. Instead of relying on the intuitive guess, here is provided a detailed theoretical description which adopts solid mathematical approaches. It will be shown how the range sidelobe function is generated and why it works only for P code but not linear FM.

5.4.1 Sidelobe Analysis of P code

A convenient criterion for evaluating uniform phase codes of length \( N \) is the peak signal to peak sidelobe ratio (PSR). The PSR of the aperiodic autocorrelation function of a phase code signal is bounded by [20] [72]

\[
PSR[dB] \leq 20 \log(N) = PSR_{\text{max}}[dB]
\]  

(5.11)

This is a very obvious result since the first output of autocorrelation is always equivalent to the uniform amplitude of the signal.

After pulse compression the PSL of P3 or P4 codes appear to be much lower compared to the linear FM case. Considering that the P codes are actually no more than discrete conversions of analog linear FM signals, it may contradict to the well established sidelobe property of linear FM. It is found that this is actually due to the undersampling effect.

\(^1\) On the condition that both incoming signal and waveform of correlated filter have uniform magnitude of 1.
This lower peak sidelobe level is attributed to unique sampling rate used to convert linear FM to P code. The comparison of ACF's among various discrete linear FM signals, each of which is sampled at various sampling rate is demonstrated in Figure 5.5. The sampling rate used for generating P code, which is equivalent to Nyquist rate, leads to the perfect periodic autocorrelation.

### 5.4.2 Sidelobe Equation

Let the P3 polyphase code be represented as continuous time function of \( S_{p3}(t) \) with its expression given as

\[
S_{p3}(t) = \sum_{w=1}^{N} \exp\left[\frac{j\pi}{N}(w - 1)^2\right] \cdot U\left[\frac{t - \left(w - \frac{1}{2}\right)\cdot t_b}{t_b}\right]
\]  

(5.12)

For the purpose of focussing on quick analysis, continuous time domain function \( S_{p3}(t) \) is simplified and converted to a discrete form as \( S_{p3}(w) \). This discrete function comprises \( N \) elements. Then, let the \( q \)th term of the autocorrelation function of \( S_{p3}(w) \) be designated as \( \Psi_{p3}(q) \). This discrete function may be describes as

\[
\overline{S_{p3}}(w) = \sum_{w=1}^{N} \exp\left[\frac{j\pi}{N}(w - 1)^2\right]
\]

(5.13)
and the \( q \)th term of autocorrelation function is

\[
\Psi_{p3}(q) = [S_{p3}(N - q + 1) \cdots S_{p3}(N)] \cdot \text{conj} [S_{p3}(1) \cdots S_{p3}(q)]^T
\]

\[
= \sum_{k=1}^{q} \exp\left[j \frac{\pi}{N}(N - q + k)\right] \cdot \exp[-j \frac{\pi}{N}k^2]
\]

\[
= \sum_{k=1}^{q} \exp\left[j \frac{\pi}{N}(N - q)^2 + 2(N - q)k\right]
\]

\[
= \exp\left(j \frac{\pi}{N}(N - q)^2\right) \cdot \sum_{k=1}^{q} \exp\left[\frac{2\pi}{N}(N - q)k\right]
\]

(5.14)

Eqn. (5.14) has one constant term, which, since both \( N \) and \( q \) are integers, can be further simplified as

\[
\exp\left(j \frac{\pi}{N}(N - q)^2\right) = \exp\left[j\pi(2q + \frac{q^2}{N})\right] \quad (\because \exp(-j\pi \cdot 2q) = 1)
\]

\[
= (-1)^N \exp(j\pi \frac{q^2}{N})
\]

(5.15)

A careful investigation on the second term leads to the discovery that this summation expression is no more than a simple geometric sequence with a geometric ratio of

\[
\exp\left[j \frac{\pi}{N}(N - q)\right] = \exp\left[-j2\pi \frac{q}{N}\right].
\]

Therefore using the formulation for the sum of geometric sequence, the group of \( q \) individual elements in Eqn. (5.14) can be combined together as one simple mathematical expression, which is given as

\[
\sum_{k=1}^{q} \exp\left[j \frac{2\pi}{N}(N - q)k\right] = \exp\left[-j2\pi \frac{q}{N}\right] \cdot \frac{\exp\left[-j2\pi \frac{q^2}{N}\right] - 1}{\exp[-j2\pi \frac{q}{N}]} - 1
\]

(5.16)

Eqn. (5.16) can be further proceeded into a simpler form by manipulating complex exponential term as

\[
\exp\left[-j2\pi \frac{q}{N}\right] \cdot \frac{\exp\left[-j2\pi \frac{q^2}{N}\right] - 1}{\exp[-j2\pi \frac{q}{N}]} = \exp\left[-j\pi \frac{q}{N}\right] \cdot \frac{\exp(-j\pi \frac{q^2}{N}) - \exp[j\pi \frac{q^2}{N}]}{\exp[-j\pi \frac{q}{N}] - \exp[j\pi \frac{q}{N}]}\]

\[
= \exp\left[-j\pi \frac{(q + q^2)}{N}\right] \cdot \frac{\sin(\pi \frac{q^2}{N})}{\sin(\pi \frac{q}{N})}
\]

(5.17)

A part of Eqn. (5.17) may cancel out the complex phase term within Eqn. (5.15). After Eqns. (5.15) and (5.17) are substituted into Eqn. (5.14), \( \Psi_{p3}(q) \) becomes

\[
\Psi_{p3}(q) = (-1)^N \exp(j\pi \frac{q^2}{N}) \cdot \exp\left[-j\pi \frac{(q + q^2)}{N}\right] \cdot \frac{\sin(\pi \frac{q^2}{N})}{\sin(\pi \frac{q}{N})}
\]

(5.18)

It is the sin function term of the variable \( q^2 \) that gives rise to the random noise-like
fluctuation in sidelobe (high frequency term) while the other term with the variable \( q \) decides overall sidelobe shape and magnitude level (low frequency term). However, given this formulation alone, it still looks difficult to reach any intuitive proposition that will eventually lead to a sidelobe reducing scheme.

Although there is such a close relationship between the P code and linear FM, they do not share the same pulse compression sidelobe property, even if the phase code length of P codes increases and sub-pulse length \( t_b \) decreases [49] due to the fundamental difference in the sampling algorithms.

### 5.5 Propositions and Validations

Figure 5.6 is a graphical description on how the P code sidelobe pattern deviates from that of the ordinary linear FM. The sampling rate for the linear FM is chosen as twice higher than that of the P code. Because of the undersampling effect, the -13.2 dB PSL does not appear in the P code case. Instead the sidelobe curves trace the lower pitches of the linear FM sidelobe pattern. This provides a preliminary insight into the unique P code sidelobe behaviour having low sidelobe level.

It was already mentioned that, despite the fact that P code is a branch of sampled linear FM signal, its unique sampling rate gives rise to a unique characteristics distinguished from ordinary linear FM. To investigate in depth on the property of P code, several simulation results as well as relevant theoretical statements have been produced.

Firstly, let the argument start with the basic linear FM property regarding the peak sidelobe level. For a linear FM signal \( S(t) = exp(-j\pi Bt^2) \) where \(-\frac{T}{2} \leq t \leq \frac{T}{2}\),
Theorem 1: The peak sidelobe level generated by discrete signal, which is sampled from linear FM, is bounded by $1/t$ curve.

- This linear FM sidelobe property is introduced as a provisional knowledge before an in-depth analysis and examination of P code is established. It comes from the ambiguity function of linear FM with zero Doppler shift. When time-bandwidth product $D = BT$ is sufficiently large, the ACF for Linear FM is given

$$ACF_{L-FM}(t) = BT \frac{\sin(\pi Bt)}{\pi Bt}$$ (5.19)

Eqn.(5.19), which is a sinc function, can be seen as a sin function with its magnitude bounded by $T/(\pi t)$ curve and maximum value given as $BT$ at $t = 0$. This equality holds when and only when the sampling frequency is set higher than twice the time bandwidth produce $BT$. The first peak for $t > 0$ or the sidelobe peak appears at $t_{ps}$, which is the minimum solution to satisfy the condition $\frac{dACF}{dt} = d \left( BT \frac{\sin(\pi Bt)}{\pi Bt} \right) /dt = 0$. It can be easily deduced that the PSL is always given as $-13.2$ dB at $t = t_{ps}$, which satisfies $t_{ps} = \tan(t_{ps})$.

Bearing this basic property in mind, the PSL characteristics for undersampling cases, in which $S_r < 2BT$, are examined. For the discrete linear FM signal $S_p(t)$ given in Eqn.(5.12) and (5.13),

Proposition 1: Let $\Psi_p$ be the autocorrelation function of discrete code $\hat{S}_p$ which is a discrete form of linear FM sampled at a rate of $S_r$. When the sampling rate is given as $S_r = BT$, the qth term of sidelobe level $\Psi_{p3}(q)$ is bounded by $(1/\sin(\pi \frac{q}{N}))$ and the PSLR[dB] is given as $20 \log_{10} \left[ \frac{\sqrt{2} \cos((1/x)\sqrt{N})}{\sqrt{\pi^2-4}} \cdot \sqrt{\frac{1}{N}} \right]$ at $q = i$, where $i$ is the nearest integer that satisfies the equality $\alpha = \frac{1}{2} - \frac{1}{x} \cdot \sqrt{N}$ ($-0.5 \leq \alpha < 0.5$).

Proof: After removal of phase terms in Eqn. (5.18), $ACF(q)$ becomes $\sin(\pi \frac{q^2}{N}) / \sin(\pi \frac{q}{N})$. This equation can be viewed as a $\sin(\pi \frac{q^2}{N})$ with its magnitude bounded by the low frequency amplitude term $1/\sin(\pi \frac{q}{N})$. The graph in Figure (5.7) shows the relationship between these curves. For a small $q$, the sidelobe pattern is dominated by the high frequency term $\sin(\pi \frac{q^2}{N})$ and a rough solution for the PSL can be determined by the first peak of $\sin(\pi \frac{q^2}{N})$. This occurs at $q_{peak} = \sqrt{\frac{N}{\pi}}$. After applied into original ACF(q) equation, $q_{peak}$ produces $1/\sin(\frac{\pi q}{\sqrt{N}}) \approx \sqrt{\frac{N}{\pi}}$ as the maximum sidelobe. The phase code length or the pulse compression ratio is assumed to be sufficiently large such that this approximation can be justified. Hence the PSL[dB] is given as $20 \log_{10}(\sqrt{\frac{N}{\pi}})$. But it is found that this approximation does not always correspond to the actual result. Especially the amount of absolute error becomes enlarged as $N$ increases. The graph in Figure 5.7 shows that the first peak of sidelobe occurs just before the crossover with the curve $1/\sin(\pi \frac{q}{N})$. Since the crossover happens at
5.5 Propositions and Validations

$q = \sqrt{\frac{N}{2}}$, the actual maximum peak is higher than $\sqrt{\frac{N}{\pi}}$. Let this maximum peak response take place at $q_{\text{max}} = \sqrt{\frac{N}{2} - \alpha}$, where $\alpha \ll N$. For large $N$ (usually $N > 100$), $\sin(\pi \sqrt{\frac{N}{2} - \alpha})$ can be approximated to $(\pi \sqrt{\frac{N}{2} - \alpha})/N$. Replacing $q$ with $q_{\text{max}}$, the ACF becomes a function of $\alpha$ around maximum peak and is expressed as

$$ACF(\alpha; q_{\text{max}}) = \sin \left[ \frac{\pi}{N} \left( \frac{N}{2} - \alpha \right) \right] \left( \frac{\pi}{N} \sqrt{\frac{N}{2} - \alpha} \right)$$  \hspace{1cm} (5.20)

To find out $\alpha_{\text{max}}$ at $q_{\text{max}}$, the derivative of Eqn. (5.20) in regards to $\alpha$ is calculated such that the following condition is satisfied.

$$\frac{dACF(\alpha; q_{\text{max}})}{d\alpha} = -\sqrt{\frac{N}{2} - \alpha} \sin \left( \frac{\alpha}{\pi} \right) + \frac{N}{2\pi \sqrt{\frac{N}{2} - \alpha}} \cos \left( \frac{\alpha}{\pi} \right)$$ \hspace{1cm} (5.21)

$$= 0 \Big|_{\alpha = \alpha_{\text{max}}}$$

Using $N/2 \gg \alpha$ and trigonometric function approximation rule, it can be shown that Eqn. (5.21) results in the condition

$$\tan \left( \frac{\alpha}{N} \pi \right) \approx \frac{\alpha}{2\pi} = \frac{1}{\pi}$$ \hspace{1cm} (5.22)

$$\therefore \alpha = \frac{N}{\pi^2}$$ \hspace{1cm} (5.23)

Since $q$ has to be an integer, the actual $q_{\text{max}}$ is the nearest integer around non-integer real value

$$\therefore i = \sqrt{\left( \frac{1}{2} - \frac{1}{\pi^2} \right) N}$$ \hspace{1cm} (5.24)
After replacing $q$ with the value $i$ in Eqn. (5.18), the PSL value for P code is obtained as

$$ P_{SL}^p \cong 20 \log_{10} \left[ \frac{\sqrt{2} \cos\left(\frac{1}{\pi} \right)}{\sqrt{\pi^2 - 4}} \cdot \sqrt{\frac{1}{N}} \right] $$

(5.25)

Thus PSL of P code is proportional to $1/\sqrt{N}$. Although Eqn. (5.25) is not an exact solution, the simulation results confirms that it holds a good accuracy within trivial error range. In fact, it is found that -1 dB should be added to compensate the error for most cases.

Proposition 2 : When $1 < BT < S_r < 2BT$, the sidelobe of ACF of $\overline{S}$, which is a discrete form of linear FM $S(t) = \exp(-j\pi \frac{B}{N} t^2)$, $0 \leq t < \frac{T}{2}$ sampled at a rate of $S_r$, is bounded in magnitude by $1/t$ and the inequality $PSL \leq -13.2$ [dB] holds.

Proof : Let the sampling rate $S_r$ be expressed as

$$ S_r = M = \frac{N(p + q)}{p} $$

(5.26)

where $N = BT$ and $p$ and $q$ are arbitrary integers with $1 \leq q < p$. Then the pulse code length $t_b$ is given as $T_b = \frac{p}{N(p+q)}$. Let $\xi$ be a new discrete linear FM signal with sampling rate $S_r \xi = N(p + q)$. And also let the discrete waveform $\overline{S_1}$ be a subset of $\xi$ which is sampled in every $p$ steps starting from $\xi(1)$. Then the discrete signal $\overline{S}$ is equivalent to $\overline{S_1}$. Likewise we can consider $p - 1$ different discrete waveforms $\overline{S_2}, \overline{S_3}, \ldots, \overline{S_p}$ all of which are subsets of $\overline{\xi}$ but with different sampling reference time. Each of these waveforms can be expressed by $\overline{S_1}$ as

$$ \overline{S_u} = \overline{S_1} \cdot [z_1, z_2, z_3, \ldots, z_{N(p+q)}/p] $$

(5.27)

where the $v$th component $z_v$ of complex vector $\overline{Z_u}(u > 1)$ is given as

$$ z_v = \exp \left\{ -j\pi \frac{BT}{N^2} \frac{[(v-1)p + u - 1]^2 - [(v-1)p]^2}{(p + q)^2} \right\} $$

(5.28)

Now suppose that $\overline{\xi}$ is correlated with a matched filter and the correlator outputs are taken in every $p$ steps. This is equivalent to the summation of ACFs generated by $\overline{S_1}, \overline{S_2}, \overline{S_3}, \ldots, \overline{S_p}$. Because each $\overline{S_n}$ are different with each other by pre-determined phase terms only, they behaviours at the correlation outputs will be very similar with each other too. Let $\Psi_\xi(m)$ be the $m$th term of ACF of $\overline{\xi}$. Then the sampled ACF, which is taken in every $t = 1/\{N(p + q)\}$ is the summation of $p$ individual
ACFs,

\[ \Psi_\xi(kp) = \frac{1}{p} \sum_{i=0}^{p-1} \Psi_{\tilde{S}_1}(k) \]  \hspace{1cm} (5.29)

where \( k \) is an integer in the range of \(-M \leq k \leq M\) so that \( \Psi_\xi(m) \) is maximized at \( m = 0 \). A denominator \( p \) is introduced to equalize energy level to the original one. By transforming each ACF expression with the sum of Eqn. (5.27), Eqn. (5.29) can be rearranged using only \( \Psi_{\tilde{S}_1}(k) \) as

\[ \Psi_\xi(kp) = \frac{1}{p} \sum_{i=0}^{p-1} \tilde{Z}_i \cdot \tilde{Z}_i^* \cdot \Psi_{\tilde{S}_1}(k) \]
\[ = \Psi_{\tilde{S}_1}(k) \]  \hspace{1cm} (5.30)

Since \( \Psi_\xi(kp) \) is a part of the ACF of signal \( \xi \) of which the sampling rate is higher than \( 2BT \), its magnitude at \( k \neq 0 \) is be bounded by the ACF sidelobe of the ordinary linear FM. From proposition 1, the sidelobe of linear FM ACF is bounded by \( 1/t \). Therefore \( \Psi_\xi(kp) \) is bounded by a function of \( 1/t \) and the maximum possible PSL is lower than \(-13.2 \text{ dB}\).

**Proposition 3**: Suppose \( \tilde{S} \) is a discrete phased code derived from linear with a sampling rate given as \( S_r \). When the sampling rate is given as \( 1 \ll BT = N < S_r < 2BT \), for the sidelobe of ACF of \( \tilde{S} \), the inequality \( 20 \log_{10} \left[ \frac{\sqrt{2} \cos(\frac{1}{\pi t})}{\sqrt{\pi^2 - 4}} \cdot \frac{\sqrt{1}}{2BT} \right] < \text{PSLR} < -13.2 \text{ [dB]} \) is satisfied.

**Proof**: Let \( \tilde{S}_1 \) be a linear FM phased code with \( N = 2BT \). Because \( BT \) is sufficiently large compared to 1, the normalized ACF of \( \tilde{S}_1 \) belongs to the subset of linear FM ACF. Naturally its ACF output, which will be denoted as \( \Psi_1(k) \), follows Eqn. (5.19). For \( S_r = 2N \) case, \( t \) can be replaced by \( kT/(2N) \). As shown in proposition 1, for \( k = \{2, 4, \ldots, 2N - 2 \} \), the discrete sidelobe pattern approach to that of P code resulting in Eqn. (5.18). Also, from proposition 2, \( \Psi_1(k) \) is bounded as \( \Psi_1(k) < |\sin(\pi \frac{k}{2})/(\pi \frac{k}{2})| \approx 1/(\pi \frac{k}{2}) \), for \( k = \{1, 3, \ldots, 2N - 1 \} \). Now suppose that a new set of discrete linear FM \( \tilde{S} \) has a code length of \( M = N \frac{p+q}{p} \), where \( p \) and \( q \) are arbitrary integers with \( 1 \leq q < p \), and their G.C.M (Greatest Common Measure) is 1. This condition restricts the choice of integers \( M \) and \( N \) such that \( N = pi, (i \text{ is an integer}) \). Similar to \( \Psi_1(k) \), the ACF of \( \tilde{S} \) denoted as \( \Psi(k) \) may have an approximation form of \( \sin(\pi \frac{p}{(p+q)m})/(\pi \frac{p}{(p+q)m}) \), where \( m = 1, 2, \ldots, M \).

From proposition 1, the maximum of lower bound graph occurs at nearest integer of \( \frac{M}{BT} \left( \frac{1}{2} - \frac{1}{\pi^2} \cdot \sqrt{N} \right) \). Then for any integer \( m < \frac{M}{BT} \left( \sqrt{\frac{1}{2} - \frac{1}{\pi^2} \cdot \sqrt{N}} \right) \), there exists
an integer $k$ that satisfies the condition

$$\frac{k}{2} \leq \frac{p}{(p + q)m} < \frac{k + 1}{2} \quad (5.31)$$

i) $k = 2l$: The function $|\Psi_1(x)|$ shows a consistent increase in the interval $\left[\frac{k - 2l}{2} \leq x < \frac{k - 2l + 1}{2}\right]$. Hence

$$|\Psi_1\left(\frac{k}{2}\right)| \leq \left|\Psi\left(\frac{p}{(p + q)m}\right)\right| < \left|\Psi_1\left(\frac{k + 1}{2}\right)\right| \quad (5.32)$$

where $|\Psi_1\left(\frac{k}{2}\right)| = |\Psi_1(l)|$ is a part of P code ACF.

ii) $k = 2l + 1$: The function $|\Psi_1(k)|$ shows a consistent decrease in the interval $\left[\frac{k - 2l + 1}{2} \leq x < \frac{k - 2l}{2}\right]$. Hence $|\Psi_1\left(\frac{k}{2}\right)| \geq \left|\Psi_2\left(\frac{p}{(p + q)m}\right)\right| > \left|\Psi_1\left(\frac{k + 1}{2}\right)\right|$. Here $|\Psi_1\left(\frac{k + 1}{2}\right)| = \left|\Psi_1(l + 1)\right|$ is a part of P code ACF.

From above two cases, a straightforward conclusion is drawn that the magnitude of function $\left|\Psi_2\left(\frac{p}{(p + q)m}\right)\right|$ is lower bounded by the ACF of the P code sidelobe. Thus, for any linear FM phase coded waveform $S_r$ of code length given as $1 \ll BT = N < S_r < 2BT$, there exists at least an integer $m = m_{max}$ such that

$$\therefore |\Psi(m_{max})| > Max (|\Psi_1(2l)|) = \left[\sqrt{2} \cos\left(\frac{1}{\pi}\right) \cdot \sqrt{\frac{1}{BT}}\right]$$

Therefore, from the maximum boundary condition of proposition 1, the inequality

$$20\log_{10}\left[\frac{\sqrt{2} \cos\left(\frac{1}{\pi}\right)}{\sqrt{\pi^2 - 4}} \cdot \sqrt{\frac{1}{BT}}\right] < PSLR \ [dB] < -13.2 \quad (5.34)$$

is satisfied.

Figure 5.8 compares the sidelobes generated from two linear FM derived phased code waveforms with their $S_r$ given as $BT$ and $2BT$ respectively. As expected, the sidelobe of the P code ($\therefore S_r = BT$) exactly overlaps over the other one. On the other hand, the sidelobe for $N = 2BT$ case forms an amplitude envelope of sinc function. Hence the assumption used for proposition 2 and 3 is justified. This graphical description provides an enhanced insight regarding propositions. Between $BT < S_r < 2BT$, as $S_r$ increase, the location of PSL of $\Psi(k)$ will shift toward left. Figure 5.9 shows the PSL along the variation of $BT < S_r < 2BT$. As $S_r$ approach to $BT$, the PSL value converges to the predicted value in Eqn. (5.25) with -1 dB added.

When $S_r$ is set lower than $BT$, the ambiguous secondary peaks are brought into the sidelobe parts (aliasing effect) and comparison of sidelobe level is meaningless.
5.8: Sidelobe magnitude comparison between two phase coded signals with different code lengths: Code length $N=(2BT, BT)$, where $BT$ is the time-bandwidth product.

Figure 5.8: Sidelobe magnitude comparison between two phase coded signals with different code lengths: Code length $N=(2BT, BT)$, where $BT$ is the time-bandwidth product.

Figure 5.9: Peak sidelobe variation with respect to $S_r$ in the interval $[BT, 2BT]$. 
5.6 Range Sidelobe Reduction

**Lemma 1**: For a sufficiently large $BT = N$, the optimum sampling frequency to convert continuous linear FM signal into discrete form such that the peak sidelobe level of ACF is minimized, is given by $BT$ or the pulse compression ratio.

From all the propositions and the relevant simulation results, it is confirmed that the sampling rate used for P code is in fact the optimal value to reduce the peak sidelobe level of the linear FM derived phase code. The optimized peak sidelobe level of the P code is given as $20 \log_{10} \left( \frac{\sqrt{2} \cos \left( \frac{1}{N} \right)}{\sqrt{\pi^2 - 4}} \cdot \sqrt{\frac{1}{N}} \right) - 1$ [dB].

5.6 Range Sidelobe Reduction

In the P3 code, which is a single sideband derived code, successive phase samples of large sidelobes have little phase variation between adjacent code elements. Thus, by simply inserting a subtractor just after pulse compressor, a new code elements which has limited sidelobe characteristic can be achieved. This scheme is shown in Figure 5.10. To implement this in hardware, one element length time-delay as well as a subtractor are additionally needed. Figure 5.11(a) shows the results when this pulse compression scheme is applied to the P3 code for the cases with and without Doppler shift respectively. As predicted, a dramatic sidelobe reduction is achieved. For the pulse compression ratio or code length of 200 the PSL is found to be -33 dB, which is higher than the Barker code level by 6 dB ($\cdot 20 \log_{10}(100) = 40$ dB). When the frequency Doppler shift is introduced, it appears to still maintain a good tolerance. The Doppler shift frequency is given as 5% of the signal bandwidth. In the P4 code case two bits adder is to be used instead of the subtractor. It is due to the difference in the phase generation algorithm between the two codes. The P4 code has a double sideband frequency spectrum. From Eqn. (5.7), unlike the P3 code case, there appears an additional term. It is deduced that an adding operation is required to compensate this term.
5.6 Sidelobe suppression analysis

Figure 5.11: P3 compressed pulse with two types of subtractors

Despite the excellent sidelobe reduction effect, it is not without its disadvantage. Instead of gaining low sidelobe and high Doppler tolerance, its mainlobe width has doubled which causes the pulse compression rate to decrease by half. Also the sidelobe is not fully optimized since they are not uniform pattern. Most of all the complicate hardware implementation poses a practical difficulty. In real situation, when high level noise is produced with increased system complexity, the performance will be deteriorated by low SNR. Figure 5.11(b) shows another simulation result when an additional subtractor or adder is introduced at the output end of Figure 5.10. The mainlobe is more clarified due to deep isolation from sidelobe pattern although its width is now broadened by four times.

5.6.1 Sidelobe suppression analysis

To verify the sidelobe subtraction scheme, the \((q+1)\)th term of pulse compression output is taken along with \(q\)th term. From Eqn. (5.14).

\[
\Psi_{p3}(q+1) = (-1)^N \exp(-j\pi \frac{q+1}{N}) \cdot \frac{\sin\left(\frac{(q+1)^2}{N}\right)}{\sin\left(\frac{q^2}{N}\right)} \tag{5.35}
\]

When the pulse compression ratio is sufficiently large or \(N = BT > 100\) as in the most cases, \(\frac{1}{N}\) approaches to a trivial figure compared to \(\frac{q}{N}\) and this enables to approximate \(q+1\) variable in denominator into \(q\)only term. Bearing this in mind, the procedure to obtain the difference between \(\Psi_{p3}(q+1)\) and \(\Psi_{p3}(q)\) is mathematically straightforward. The \(q\)th term of subtractor output is given as

\[
\Gamma_{p3}(q) = [\Psi_{p3}(q+1) - \Psi_{p3}(q)]
\]

\[
= (-1)^N \exp(-j\pi \frac{q}{N}) \cdot \left\{ \exp(-j\pi \frac{q^2}{N}) \frac{\sin\left(\frac{(q+1)^2}{N}\right)}{\sin\left(\frac{q^2}{N}\right)} - \frac{\sin\left(\frac{q^2}{N}\right)}{\sin\left(\frac{q^2}{N}\right)} \right\} \tag{5.36}
\]
5.6 Finite Quantization Effect

After stripping off sign term \((-1)^N\) and on the approximate condition that satisfies \(N \gg 1, q \gg 1\),

\[
\Gamma_{p3}(q) \cdot (-1)^N \approx \exp(-j\pi \frac{q}{N}) \cdot \frac{1}{\sin(\pi \frac{q}{N})} \cdot \left\{ \sin\left(\pi \frac{(q+1)^2}{N}\right) - \sin\left(\pi \frac{q^2}{N}\right) \right\} \\
= \frac{\exp(-j\pi \frac{q}{N})}{\sin(\pi \frac{q}{N})} \left\{ \sin\left(\pi \frac{q^2}{N}\right) \left[ \cos\left(\pi \frac{2q}{N}\right) - 1 \right] + \cos\left(\pi \frac{q^2}{N}\right) \sin\left(\pi \frac{2q}{N}\right) \right\} \\
= 2 \exp(-j\pi \frac{q}{N}) \left\{ -\sin\left(\pi \frac{q^2}{N}\right) \sin\left(\pi \frac{q}{N}\right) \cos\left(\pi \frac{2q}{N}\right) \right\} \\
= 2 \exp(-j\pi \frac{q}{N}) \cos\left[ \pi \frac{q^2}{N} - q \right] \\
(5.37)
\]

On the condition that \(N\) is sufficiently large compared to 1, \(\exp(-j\pi \frac{q}{N})\) term has been removed and \(\sin(\pi \frac{q^2}{N})\) was approximated to \(\sin(\pi \frac{q}{N})\). Now the postcompression output generated by two-sample subtractor is represented as a simple closed forms. The first term contributes to phase variation only. The output magnitude is only a function of \(\cos(X)\) with maximum magnitude of 2. Figure 5.12 compares the actual amplitude variation at the subtractor output with the approximate function expressed in Eqn. (5.37) for verification purpose. The figure manifests that the approximation curves converges to the exact equation very quickly \((N < 20)\) without significant error margin. A similar form is obtained for the P4 code case by taking summation as

\[
\Gamma_{p4}(q) = \left[ \Psi_{p4}(q + 1) + \Psi_{p4}(q) \right] \\
(5.38)
\]

5.6.2 Finite Quantization Effect

Until now all the simulations in the preceding sections were produced with the assumption that the signal processing was carried out in the ideal condition. However a finite
5.6 Resolution Loss Compensation

interval amplitude quantization may affect the digital waveform performance in actual situation.

In generating waveforms at transmitter, sometimes, it is more efficient to utilize stored data in memory than to produce each phase element value by a direct calculation. Here the sequence of the waveform sample is precomputed, stored and clocked out of memory [23]. Reducing the number of quantization levels so as to minimize the memory storage would contribute favourably to the system complexity and costs.

The chirp signal pulse used in ERS-1 SAR system is extended over a time window of 37.12-μs and the sampling rate is given as 704 points per one pulse. Because the range resolution in SAR image is dependent on the pulse bandwidth, the similar sampling rate should also be applied to the phase coded waveform cases to achieve as good a range resolution property. Eqns. (5.6) and (5.7) show that the number of quantization level for P code is proportional to the pulse compression ratio or time-bandwidth product. The Nyquist sampling rate for discrete pulse code is one per every phase elements and equal to pulse compression ratio. To fully implement 704 different quantization levels, 10 bits of memory space is required. To address this finite quantization effect, Figure 5.13 shows the ACFs of the P3 code for the cases with and without subtractor respectively. It is found that 7 bits quantization level provides a good approximation to the original ACF performance. Although a number of sharp variations are detected in amplitude, the amount of change is so small that they can hardly affect the overall performance. The peak sidelobe level is kept below -50 dB (: −20\log_{10}(704) = −57 dB). Figure 5.13(b) is obtained with the Doppler shift effect considered. Again no significant degradation is observed.

To investigate this effect further, Figure 5.13(c, d) are simulated after the quantization levels are lowered to 6 and 4 bits respectively. It is seen that 6 bits quantized P code still keeps the rough outline of the original ACF but its peak to sidelobe ratio starts to deteriorate leading to the degradation of resolution and SNR property. The 4 bits quantized signal appears to lose all the merits of the P code signal. Summarizing all above results, it is concluded that a 7 bits quantization level for P codes waveform is enough to implement the practical system compatible with that of ERS-1.

5.6.3 Resolution Loss Compensation

The resolution is degraded by half when either subtractor or adder is involved during post-processing procedure. When the two-sample operation is adopted, the computed ACF is possessed with two mainlobes. They are identical in magnitude due to the symmetry property of autocorrelation function. This brings about mainlobe broadening effect or range resolution degradation by half. This process is illustrated in Figure 5.14. Simply enlarging the pulse compression ratio by either reduction of the phase element length or
Figure 5.13: Four different cases of ACF of the P3 phase code processed with various quantization levels

Figure 5.14: Resolution degradation during the pulse compression and sidelobe reduction at postprocessing for the P3 code
increasing bandwidth will be the most straightforward way to overcome this problem. But this will accompany the need of the greater computing power as well as an increased system cost. One possible approach to tackle this problem could be the rearrangement of the signal code such that the resulted autocorrelation function is deprived of the symmetry property.

Here is suggested a simple technique that may serve this purpose. This is accomplished by introducing additional time delay to handle the data stream coming from the matched filter output. The block diagram for this procedure is shown in Figure 5.15. In the P3 code case, the twice broadened mainlobe width is due to the subtraction of identical but one bit shifted ACFs which have symmetry structures. By taking average of two subsequent samples, a non-symmetry autocorrelation property may be gained. The new scheme lets additional mainlobe element be split into two identical ones with their energy level reduced by 6 dB respectively. To achieve the same effect for P4 code, the sign of the first time-delayed data bit should be reversed. After applying this new operation algorithm to handle the matched output stream for P3 code, a modified pulse compression output is obtained and shown in Figure 5.16(a). At first there is a noticeable change in the sidelobe pattern. Although the sidelobe pattern is not uniformly limited, the PSL remains around -50 dB. Now the twice broadened mainlobe found in Figure 5.11 is split into three different parts leaving a mainlobe of width $t_b$, the code element length. The other parts are distinguished from the mainlobe by 6 dB difference.

The 2-D ambiguity function diagram for this waveform has been obtained and shown in Figure 5.16(b). As has been indicated earlier, the P code is a special case of digital conversion from linear FM and naturally they share similar response to the Doppler shift effect. It is shown that the modified P code processing retains the property of strong resistance to the Doppler shift effect.
5.6 Cross-Correlation Performance

The cross-correlation is an output of the correlation between two separate waveforms. For the phased codes $S_1$ and $S_2$ of length $N$, their aperiodic correlation function is

$$
\hat{C}_{1,2}(n) = \left\{ \begin{array}{ll}
\sum_{k=0}^{N-1-n} \hat{S}_1(k)\hat{S}_2(k+n) & 0 \leq n < N-1 \\
\sum_{k=0}^{N-1+n} \hat{S}_1(k-n)\hat{S}_2(k) & -(N-1) \leq n < 0
\end{array} \right. \quad (5.39)
$$

As long as each waveform do not interrupt each other’s signal processing procedure and operate simultaneously by use of the same antenna, it will help to collect increased information on the target area and improve the target classification capability as well [21][55]. The idea of using multiple waveforms for wide-swath SAR application has started from this concept [28].

The cross-correlation function can also be used to measure the similarity between two different waveforms. Figure 5.17(a) is the correlation result of two P3 and P4 phase code. Both of them are derived from the linear FM via single sideband sweeping for the P3 code while the P4 code has a double sideband structure. Because of their different phase alignments, the peak mainlobe splits into two identical ones but dislocated by $T/2$ from each other. Two peak responses appear at -6 dB level of the original ACF peaks. Figure 5.17(b) describe the different features between P codes and linear FM waveform. Because they do not perfectly match with each other, the mainlobe is widened into a broad flat shape. This is easily understood in that the linear FM is assumed to have more than one sample per each phase code element in the P code.

For continuous linear FM signals with pulse compression ratio $D = BT$, the ratio of
5.6 Cross-Correlation Performance

Figure 5.17: Cross correlation function between (a) P3 and P4 codes, (b) p3 and Linear FM codes

cross-correlation peak between two signals of opposite slopes to autocorrelation peak value
is as

\[ s_1(t) = \exp \left[ j2\pi \left( f_0 t + \frac{B}{2T} t^2 \right) \right], \quad s_2(t) = \exp \left[ j2\pi \left( f_0 t - \frac{B}{2T} t^2 \right) \right] \]  \hspace{1cm} (5.40)

\[ \frac{\max \{ CCF(t) \}}{\max \{ ACF(t) \}} = \frac{C_{1,2}(0)}{T} \]

\[ = \frac{2}{\pi D} \sqrt{\frac{\pi D}{2}} \cos^2(x) + j \sin^2(x) dx \]  \hspace{1cm} (5.41)

which is a Fresnel integral and therefore never assumes a zero value. Figure 5.18(a) shows
the peak value graph of thus generated cross-correlation.

Considering their similarity, the same phenomenon is expected in the P code. To
weaken correlation tendency, a reversed P code of opposite slope has been generated by
re-ordering the phase arrangement and correlated with original code. Eqn.(5.42) indicates
the phase generation method for reversed P4 code.

\[ \phi'_i = -\pi(i-1)^2 / BT + \pi(i-1) \] \hspace{1cm} (5.42)

Obviously, the ACF for this code is exactly the same as that of the original P4 code.
The cross correlation function shown in Figure 5.18(b) implies that there exists a strong
discrepancy between the two waveforms. The CCF peak relative of -32 dB corresponds to
the graph in Figure 5.18(a). Hence linear FM formula in Eqn. (5.41) is also applicable to
the phase code.
5.7 Comparison with Linear FM

Recalling that the generation of the P code is based on the concept of the linear FM, it could be argued that the same enhancement on the range sidelobe reduction can be obtained when either two-sample subtractor or adder is attached. In addition to this, there could be a possibility that simply by increasing the pulse compression ratio and reducing $t_b$ value, the performance of the P code approaches to that of the linear FM so that eventually there does not exist any difference between the two cases. But the truth is that P code still maintains its superiority over linear FM at least in terms of low sidelobe level regardless of the pulse compression ratio or the length of the $t_b$. This attributes to the fact that the low sidelobe characteristic of P signals originates from their special property of perfect periodic autocorrelation. This claim is well established mathematically in Propositions 1-3.

A simulation is performed to address the effect of the subtractor on the linear FM and some results are shown in Figure 5.19. A conventional linear FM ACF is compared with the cases where a two-sample operator is connected at the output of matched filter

The sufficiently low peak response level of CCF means that multiple waveforms can be utilized in a single system simultaneously but only for point target detection purposes. The fundamental ambiguity function theory tells that the total output energy is always preserved. Hence the impact on the resulted target image remains the same regardless of the peak CCF level. This will be investigated further in Chapter 7.

Figure 5.18: Cross-correlation property of linear FM class waveforms. (a) Peak values of cross-correlation function between opposite sloped linear FMs, with respect to pulse compression ratio. (b) CCF between P4 and reversed P4 codes, with (solid) and without adder (dashed). $N=BT=706$. 

The sufficiently low peak response level of CCF means that multiple waveforms can be utilized in a single system simultaneously but only for point target detection purposes. The fundamental ambiguity function theory tells that the total output energy is always preserved. Hence the impact on the resulted target image remains the same regardless of the peak CCF level. This will be investigated further in Chapter 7.

5.7 Comparison with Linear FM

Recalling that the generation of the P code is based on the concept of the linear FM, it could be argued that the same enhancement on the range sidelobe reduction can be obtained when either two-sample subtractor or adder is attached. In addition to this, there could be a possibility that simply by increasing the pulse compression ratio and reducing $t_b$ value, the performance of the P code approaches to that of the linear FM so that eventually there does not exist any difference between the two cases. But the truth is that P code still maintains its superiority over linear FM at least in terms of low sidelobe level regardless of the pulse compression ratio or the length of the $t_b$. This attributes to the fact that the low sidelobe characteristic of P signals originates from their special property of perfect periodic autocorrelation. This claim is well established mathematically in Propositions 1-3.

A simulation is performed to address the effect of the subtractor on the linear FM and some results are shown in Figure 5.19. A conventional linear FM ACF is compared with the cases where a two-sample operator is connected at the output of matched filter
The performance of the linear FM-subtractor combination is based on the perfect periodic autocorrelation function of P signal where the phase difference between subsequent elements are kept constant. This can only happen when the signal is sampled at precise time intervals. But this degree of accuracy may not be easily obtainable in practical situation. Therefore the actual result may differ from the ones predicted by simulations. This strict condition is relaxed in the P signal case. Because each code element is a short pulse with finite time length, some moderate sampling time error would be allowed without considerably affecting the matched filter output.

Several drawbacks of the P code signal may appear in real situation. As mentioned in previous section their sinc type frequency spectrum implies that the actual result would not correspond to the theoretical prediction due to finite bandwidth filtering and certain degree of range sidelobe increase may be unavoidable. The broadened mainlobe is another important loss to consider along with the increased system complexity.

5.8 Digital Pulse Compression Implementation

Despite its excellent sidelobe characteristic, the requirement for large time-bandwidth product has been one of the major barrier to the wide use of phased code [78]. However recent advance in digital circuit system has enabled to circumvent this problem. Many
5.8 Digital Pulse Compression Implementation

Figure 5.20: Digital pulse generation and compression procedure block diagram of arbitrary waveforms. (a) Linear FM waveform generation in digital format. (b) Digital matched filter, which receives 12-b I/Q samples as input signal and produces 16 bit compressed outputs [4].

researches have actively investigated on the possible application of digital pulse compression technique [23]. Tortoli’s study [62] showed that high-speed DSP can deliver sidelobe as low as -70 dB using weighted waveform.

It has been described that high sidelobe rejection can be obtained as a trade-off with the peak-widening factor by selecting the proper weighting function. In practice, this level of performance requires stringent manufacturing control. Sidelobe level of -40 dB requires, in both generation and compression stage, amplitude ripples not exceeding 0.2 dB and phase ripples as low as 1 degree. Uncompensated SAW filter distortions ultimately prevent this level of sidelobe rejection, independent of the weighting function used in the compressor. In the digital approach, performance is affected by quantization and truncation of data. The flexibility and precision in data control offers the advantage of full programmability and near-ideal performance for a large class of waveforms.

Figure 5.20 illustrates block diagrams of typical digital signal generator and compressor for linear FM based waveform. Numerical samples of the I/Q baseband signals are stored in erasable programmable memories (EPROMs) with typical accuracy in the 8-12 bits of quantizations. Then these samples are converted to analog format and frequency translated to RF using single-sideband modulator. In the DSP based pulse compression, incoming signal is transformed into FFT format, multiplied by the matched filter samples
and transformed back to the time domain. PDSP 16514 stand-alone FFT processor from GEC Plessey Semiconductor performs forward or inverse FFTs on complex or real data sets up to 1024 points [62]. Data and coefficients are represented by 16 bits format in Figure 5.20(b).

'It was absolutely marvellous working for Pauli. You could ask him anything. There was no worry that he would think a particular question was stupid, since he thought all questions were stupid' - Victor Weisskopf -
Chapter 6

Modified Matched Filter for An Optimal Uniform Sidelobe

It has been shown in Chapter 5 that peak range sidelobe can be reduced through post-processing at the output of phase code pulse compression. Although it is a great advantage to provide low level sidelobe suppression, additionally associated time delay functions and arithmetic operations cause an increase in system complexity as well as a performance degradation. Whether this is really a practical way to apply this scheme to every element of data sets is another challenging question. Particularly, when the amount of data reaches significant level as in the case of SAR, to let every bit of received data go through this stream seems to be hardly a realistic scheme.

In designing or selecting a waveform candidate for specific radar application, some random selections may be better suited than others. One criterion for the selection of a good ‘random’ phase coded waveform is that its autocorrelation function should have equal time-sidelobes [72]. Also one of the common objectives in waveform design is to achieve clear areas in the distribution of ambiguity [15]. A method to achieve this goal is demonstrated in this Chapter. A new technique is presented based on the range sidelobe reduction scheme demonstrated in Chapter 5. This new approach achieves a similar range-sidelobe reduction effect, but without the use of any time delay functions. This is done when the post-processing parts are incorporated into the matched filter itself so as to generate a new correlator. As a result, it no longer works as a matched filter.

A detailed mathematical analysis is undertaken to deliver full understanding why and how it works on this particular waveform. Then, based on the mathematical representation of the pulse correlation output which has been approximated from the exact form, a technique to achieve a perfect sidelobe pattern is developed. The term ‘Perfect Sidelobe’ is used to indicate its resemblance to the Barker code characteristic, which is conceived as a requirement for a good random phase coded waveform [72]. Using this new technique a uniform sidelobe pattern is obtained along the entire range sidelobe. Then, the use of
6.1 New Matched Filter Scheme without Subtractor

The previous time range sidelobe reduction technique can be interpreted as a combination of two separate correlation filters which are matched to the received signal, but displaced in time by $t_i$. Instead of implementing bit-by-bit operations at the output of the matched filter, a pair of time delayed correlators may be put together to form a single set. This new correlator has one bit more element, or is longer by $t_b$ in time duration, than the original pulse replica. The diagram shown in Figure 6.1 illustrates this procedure. This correlation filter is generated by an arithmetic operation between two uniform polyphase codes which are derived from the conjugates of the transmitted signal. The arithmetic operation is dependent upon the given waveform signal. If the signal has a double-sideband derived structure like the P4 code, it should be the $\oplus$ operator while the P3 code, of which phase distribution is single-sideband derived, requires the $\odot$ operator. Therefore the pulse compression correlator is no longer a signal pulse replica and does not have a uniform amplitude pattern.

Figure 6.2 shows the characteristic of thus generated new filter. A P3 code of length 100 has been chosen as a reference signal. The pulse compression performance is shown to be identical to the result obtained from the range sidelobe reduction technique in Chapter 5.

In this stage, for simplicity, only the point target response is considered as the received signal. Although the similar idea can be easily applied to the extended target case, unlike CW radar signals, discrete time polyphase signals bring about a new complication to be considered. This matter will be dealt with in detail later.
Lewis [50] has argued that the adjacent cells within the P3 and P4 autocorrelation outputs differ in magnitude no more than one element cell magnitude. This assertion has been made intuitively on the ground that an additional input into the matched filter cannot cause variation in the output of more than its own magnitude. This intuition has been proved analytically in Chapter 5. It has been shown that, on the condition of sufficiently high pulse compression ratio $BT \gg 100$, the magnitude of autocorrelation function fluctuates according to a specific function which eventually converges to the curve of $2\exp(-j\pi \frac{N}{T}) \cdot \cos \left[ \frac{\pi}{N} (q^2 - q) \right]$. The maximum of this function is 2 instead of 1. Therefore, it is not really true that this method provides an optimum Barker code level (see Figure 6.2(c)). The increased system complexity burden and subsequently added noise terms are additional drawbacks. Instead of resorting to post-processing procedure, here is suggested a rather direct way that fully achieves an ideal unit amplitude sidelobe pattern over the entire sidelobe domain.

6.2.1 New Phase Code Definition

Consider a polyphase sequence $S(t)$ having $N + 1$ elements,

$$S(t) = \sum_{p=-1}^{N} \exp \left( j\pi \frac{N}{T} t^2 \right) \cdot U \left[ \frac{t - (p + \frac{1}{2}) \cdot t_b}{t_b} \right]$$

(6.1)

where $U(t) = 1$ for $|t| < \frac{1}{2}$ and zero elsewhere. $t_b$ is the time duration of one element in phase code sequence. In receiving phased code waveforms, there is an obvious choice for the sampling rate: one sample per bit. When this sampling rate is applied, the continuous time domain function $S(t)$ defined in Eqn. (6.1) can be rearranged as a discrete function.
Let $S(p)_s^{-P3}$ (\(p = 0, \ldots, N\)), which will be denoted as s-P3, be the discrete form of $S(t)$.

\[
\overline{S}_{s^{-P3}} = \sum_{p=0}^{N} \exp\left[\frac{j\pi p^2}{N}\right] \tag{6.2}
\]

$S(p)_s^{-P3}$ is identical to the P3 code apart from the fact that one new term is added at the beginning of the sequence to make the sequence symmetric, which will be useful in reducing calculation complexity for deriving the ACF. The autocorrelation output $\Omega(q)$ for this signal can be shown to be

\[
\Omega(q) = (-1)^N \sin \left[\frac{\pi}{N} q(q-1)\right] \sin \left[\frac{\pi}{N} q-1\right] \tag{6.3}
\]

Then the difference between two adjacent sidelobe is given as

\[
\Omega(q+1) - \Omega(q) = (-1)^N \frac{\sin \left[\frac{\pi}{N} q(q+1)\right]}{\sin \left[\frac{\pi}{N} q\right]} - \frac{\sin \left[\frac{\pi}{N} q(q-1)\right]}{\sin \left[\frac{\pi}{N} q-1\right]} 
\approx (-1)^N 2 \cos \left(\frac{\pi}{N} q^2\right) \tag{6.4}
\]

which is still a function of magnitude 2, but expressed in a simpler form compared with the non-symmetrical P3 case.

**Intuitive Conjecture**

The conjecture presented here is based on the results generated by a P3 code. Eqn. (5.37) is rearranged into an exponential expression as

\[
\Gamma_{p3}(q) = (-1)^N \left\{ \exp \left[\frac{j\pi}{N}(q^2 - q)\right] + \exp(-j\frac{\pi}{N} q^2) \right\} \tag{6.5}
\]

In Eqn. (5.14), the first term is always given as $(-1)^N \exp \left[\frac{j\pi}{N}(q^2 - q)\right]$, which is identical to the first term in Eqn. (6.5). If this term is, somehow, taken out during the correlation procedure, only a single term $(-1)^N \exp(-j\frac{\pi}{N} q^2)$ will be left to represent the range sidelobe which is of a unit magnitude in the entire time domain. This provides a motivation to search for a manipulation by which a uniform sidelobe pattern is achieved. From now s-P3 code is taken as a reference signal, as will be shown later, for the purpose of relaxing the calculation burden.

Let the $q$th term of correlation filter output for s-P3 code be designated as

\[
\Omega_1(q) = \sum_{p=0}^{q-1} \overline{S}_{s^{-P3}}(p) \cdot \overline{S}_{s^{-P3}}(N - q + p + 1) \tag{6.6}
\]

The first term of this summation equation is $\overline{S}_{s^{-P3}}(0) \cdot \overline{S}_{s^{-P3}}(N - q + 1)$, which is
equal to

\[ \mathcal{S}_{s-P3}(N - q + 1) = \mathcal{S}_{s-P3}(q - 1), \quad (\because \mathcal{S}_{s-P3}(0) = 1) \]
\[ = \exp \left[ \frac{\pi}{N} (q - 1)^2 \right] \]
\[ = \cos \left[ \frac{\pi}{N} (q - 1)^2 \right] + j \sin \left[ \frac{\pi}{N} (q - 1)^2 \right] \] (6.7)

Here only even number of \( N \) will be considered for convenience, although exactly the same result but with an opposite sign will be obtained for the case of odd \( N \). Suppose this equation \( \Omega_1(q) \) is used for the subtraction shown in Figure 6.1. By comparing this result with \( \Omega_1(q) - \Omega_1(q - 1) \cong 2 \cos(\frac{\pi}{N} (q - 1)^2) \), which is the sidelobe equation derived in Eqn. (6.4), it can be deduced that the removal of this first term \( \exp \left[ \frac{\pi}{N} (q - 1)^2 \right] \) in \( \Omega_1(q) \) term during the subtraction procedure may produce a sidelobe given as

\[ \left\{ \Omega_1(q) - \exp \left[ \frac{\pi}{N} (q - 1)^2 \right] \right\} - \Omega_1(q - 1) = 2 \cos(\frac{\pi}{N} (q - 1)^2) - \exp \left[ \frac{\pi}{N} (q - 1)^2 \right] \]
\[ = \exp \left[ -\frac{\pi}{N} (q - 1)^2 \right] \] (6.8)

This will produce a dramatic result such that the range sidelobe level is kept constant at unity. To implement this idea in reality, a new correlator set

\[ \Omega_2(q) = \sum_{p=1}^{q} \mathcal{S}^{*}_{s-P3}(p) \cdot \mathcal{S}_{s-P3}(N - q + p) \] (6.9)

is introduced, which is a duplication of \( \Omega_1(q + 1) \) with the first term removed. Under the assumption of a reasonably high pulse code length \( N \) and for \( q \) relatively larger than 1, the following approximate equation is satisfied.

\[ \Omega_2(q) - \Omega_1(q) \cong \cos\left(\frac{\pi q^2}{N}\right) - j \sin\left(\frac{\pi q^2}{N}\right) \]
\[ = \exp(-j\frac{\pi q^2}{N}) \] (6.10)

The correlation filter and its pulse compression output for this scheme may be related as shown in Figure 6.3.

**Development and Implementation**

Following the intuition in the previous section that involves rough approximations, an exact mathematical formulation is pursued to confirm the assertion.
The $q$th term of correlator output $\Omega_1(q)$ is calculated as

$$
\Omega_1(q) = \sum_{p=0}^{q-1} \exp[-\frac{j\pi}{N} p^2] \cdot \exp[\frac{j\pi}{N} (N - q + 1 + p)^2] 
$$

$$
= \sum_{p=0}^{q-1} \exp[-\frac{j\pi}{N} p^2] \cdot \exp[\frac{j\pi}{N} (\zeta_1^2 + 2\zeta_1 p + p^2)], \quad (\because \zeta_1 = N - q + 1) 
$$

$$
= \sum_{p=0}^{q-1} \exp[-\frac{j\pi}{N} p^2 + \frac{j\pi}{N} p^2] \cdot \exp[\frac{j\pi}{N} (\zeta_1^2 + 2\zeta_1 p)] 
$$

$$
= \exp(\frac{j\pi}{N} \zeta_1^2) \cdot \sum_{p=0}^{q-1} \exp(\frac{j\pi}{N} \cdot 2\zeta_1 p) 
$$

(6.11)

The similar procedure is used for obtaining $\Omega_2(q)$ as

$$
\Omega_2(q) = \sum_{p=1}^{q} \exp[-\frac{j\pi}{N} p^2] \cdot \exp[\frac{j\pi}{N} (N - q + p)^2] 
$$

$$
= \sum_{p=1}^{q} \exp[-\frac{j\pi}{N} p^2] \cdot \exp[\frac{j\pi}{N} (\zeta_2^2 + 2\zeta_2 p + p^2)], \quad (\because \zeta_2 = N - q) 
$$

$$
= \exp\left\{\frac{j\pi}{N} (N-q)^2\right\} \cdot \sum_{p=1}^{q} \exp\left[\frac{j\pi}{N} \cdot 2(N-q)p\right] 
$$

(6.12)

After computing the geometric sequence summation, it is shown that each $\Omega_2(q)$ and $\Omega_1(q)$ are given as

$$
\Omega_1(q) = (-1)^N \frac{\sin\left[\frac{\pi}{N} q (q-1)\right]}{\sin\left[\frac{\pi}{N} (q-1)\right]} 
$$

$$
\Omega_2(q) = (-1)^N \exp(-j\pi q \frac{\sin\left(\frac{\pi}{N}\right)}{\sin\left(\frac{\pi}{N}\right)} q) 
$$

(6.13)

(6.14)

Since the correlators that generate $\Omega_2(q)$ and $\Omega_1(q)$ are of the same length, they can be
Figure 6.4: Amplitude and Phase angle variation along time axis of Woo filter for (a) s-P3 code (b) s-P4 code. Each of the codes have phase code length of 200 (pulse compression ratio $BT=200$)

combined together to become a new single filter $\overline{W}(p)$. Let this new filter $\overline{W}(p)$ be denoted as the Woo filter for convenience. Then the Woo filter for the s-P3 code $Woo_{s-P3}(t)$ is described in the time domain as

$$Woo_{s-P3}(t) = \sum_{p=1}^{N} \left\{ \exp\left(\frac{-j\pi}{N}p^2\right) - \exp\left[\frac{-j\pi}{N}(p-1)^2\right] \right\} \cdot U\left[ \frac{t-(p-\frac{1}{2}) \cdot t_b}{t_b} \right]$$

$$= \sum_{p=1}^{N} 2j \sin\left[\frac{\pi}{N}(p-\frac{1}{2})\right] \cdot \exp\left[-j\frac{\pi}{N}(p^2 + p - \frac{1}{2})\right] \cdot U\left[ \frac{t-(p-\frac{1}{2}) \cdot t_b}{t_b} \right]$$

Let $\overline{Woo}(n)$ be a discrete form converted from $Woo(t)$. After substituting Eqns. (6.13) and (6.14) into Eqn. (6.10), it can be shown that the correlation between $\overline{Woo}_{s-P3}(n)$ and the s-P3 codes results in

$$\overline{Woo}_{s-P3}(n) \otimes \tilde{N}_{p3}(n') = \Omega_2(q) - \Omega_1(q)$$

$$\approx (-1)^N \exp(-\frac{\pi q^2}{N})$$

A similar procedure may be applied for a s-P4 code. Figure 6.4 shows the magnitude and phase angle distribution of the Woo filters $Woo(t)$ along the time axis. The appearance of the amplitude envelope implies that the Woo filter is generated through a similar procedure of weighting equivalent to reducing low frequency components in s-P3 code or high frequency components for s-P4 code. They can also be compared to conventional weighting techniques such as Ham, Han or Dolph-Chebyshev. But it will be shown that their performances in pulse compression are remarkably distinct.
6.2 Justifying Formulation and Approximation

There have been a few assumptions and approximations in the previous section in the course of deriving the Woohler concept for a s-P3 code. Here is given a detailed explanation on how to proceed and justify the approximate derivation. Instead of the s-P3 code used in previous section, the s-P4 code is taken for this purpose.

For an s-P4 code of code length $N$, the $q$th term of the ACF in the first sidelobe domain $(1 \leq q < N - 1)$ is given as

$$
\Omega_1(q) = \sum_{p=0}^{q-1} \exp[-\frac{j\pi}{N} p^2 + j\pi p] \cdot \exp[\frac{j\pi}{N}(N - q + 1 + p)^2 - j\pi(N - q + 1 + p)]
$$

( Let $N - q + 1$ be replaced by \( \zeta_1 \) )

$$
= \exp[\frac{j\pi}{N}(\zeta_1^2 - N\zeta_1)] \cdot \sum_{p=0}^{q-1} \exp[\frac{j\pi}{N} \cdot 2\zeta_1 p]
$$

(6.17)

The summation term is a geometric sequence and is subsequently computed as

$$
\sum_{p=0}^{q-1} \exp \left[ \frac{2\pi}{N} \zeta_1 p \right] = \exp[j\pi(q - 1)\frac{\zeta_1}{N}] \cdot \frac{\sin(\frac{\pi(\zeta_1^2)}{N})}{\sin(\frac{\pi\zeta_1}{N})}
$$

(6.18)

Substituting Eqn. (6.18) into Eqn.(6.17), $\Omega_1(q)$ is given as

$$
\Omega_1(q) = \frac{\sin(\frac{\pi q\zeta_1}{N})}{\sin(\frac{\pi\zeta_1}{N})}
$$

(6.19)

A similar process can be used for obtaining $\Omega_2(q)$, which leads to the expression

$$
\Omega_2(q) = \sum_{p=1}^{q} \exp[-\frac{j\pi}{N} p^2 + j\pi p] \cdot \exp[\frac{j\pi}{N}(N - q + p)^2 - j\pi(N - q + p)]
$$

( Let $N - q$ be replaced with $\zeta_2$ )

$$
= \exp[\frac{j\pi}{N}\zeta_2] \frac{\sin(\frac{\pi q\zeta_2}{N})}{\sin(\frac{\pi\zeta_2}{N})}
$$

(6.20)

Under the summation $\Omega_2(q) + \Omega_1(q)^1$ becomes

$$
\Omega_2(q) + \Omega_1(q) = -\exp[-\frac{j\pi}{N} q] \frac{\sin(\frac{\pi q^2}{N}) \cdot (-1)^{q+1}}{\sin(\frac{\pi q}{N})} + \frac{\sin(\frac{\pi q(q - 1)}{N}) \cdot (-1)^{q+1}}{\sin(\frac{\pi(q - 1)}{N})}
$$

(6.21)

---

1 It has been explained in Chapter 5 that the double sideband structure of P4 code requires (+) operation instead of (-) to experience the range sidelobe reduction effect.
6.2 Justifying Formulation and Approximation

Figure 6.5: The real and imaginary terms comparison between actual correlation output and approximate forms

With the assumption that \( N \) is sufficiently large and that \( q \) is relatively greater than 1 but not very close to \( N \) (i.e. \( 1 \ll q \ll N \)), it can be shown that Eqn. (6.21) develops into an approximate form of

\[
\Omega_2(q) + \Omega_1(q) \cong \frac{(-1)^q}{\sin\left(\frac{\pi q}{N}\right)} \left\{ \exp\left[ -j\frac{\pi}{N}q \sin\left(\frac{\pi q^2}{N}\right) - \sin\left(\frac{\pi}{N}q(q-1)\right) \right] \right\} \\
= (-1)^q \exp\left(-j\frac{\pi}{N}q^2\right) = \exp(j\pi q) \exp\left(-j\frac{\pi}{N}q^2\right) \tag{6.22}
\]

A similar procedure for \( N+1 < q < 2*N \) produces an identical result\(^2\). Unlike the s-P3 code case a linear phase term appears in the s-P4 modulation.

From the relationship between original signals and Woo filter correlation, the Woo filter for a s-P4 code is represented as

\[
W_{\text{woo}_{s-P4}}(t) = \sum_{p=1}^{N} (-1)^p 2j \sin\left(\frac{\pi}{N}(p - \frac{1}{2})\right) \cdot \exp\left(-j\frac{\pi}{N}(p^2 - p + \frac{1}{2})\right) \cdot U\left[\frac{t - (p - \frac{1}{2}) \cdot t_b}{t_b}\right] \tag{6.23}
\]

Figure 6.5 illustrates how accurately the approximate Eqns. (6.22) and (6.10) follow the actual outputs. The discrepancy error at the beginning of the ACF output has been predicted from the condition which assumes that \( q \) is very large (i.e. \( q \gg 1 \)). The simulation tests confirm that very good agreement is found between the curves. Even for a moderate pulse compression ratio of \( N = 50 \), the discrepancy disappears rapidly beyond the range bin of 5.

The curves representing \( \Omega_1, \Omega_2 \) and their sum \( \Omega_1 + \Omega_2 \) are shown respectively in Figure

\(^2q = \{N, N+1\} \) designates the mainlobe elements.
6.2 Performance Analysis

Figure 6.6: Verification of Woo filter output for a $s^4$ signal with the code length $N=100$
(a) $\Omega_1(q), \Omega_2(q)$ and their summation $\Omega_2(q) + \Omega_1(q)$ (b) A performance improvement by
the Woo filter and the error level caused by approximate formulation.

6.6(a) to validate the mathematical approximation. Figure 6.6(b) presents a comparison
between the sidelobe structures each generated by the Woo filter and the post-compression
range sidelobe reduction technique. A very obvious improvement is made by the Woo filter.
Their integrated sidelobe energies are the same, but through the Woo filter the PSL is
optimally minimized for a given total sidelobe energy. The error graph shows that a
marginal discrepancy occurs between Eqn. (6.22) and the actual result when $q$ is close to
sidelobe boundaries.

6.2.3 Performance Analysis

Figure 6.7 shows the pulse compression outputs when the Woo filter is used as a
correlator. The signal consists of 200 and 1000 sub-pulses each having length of $t_b$ while
their corresponding Woo filters have 199 and 999 elements respectively. As predicted, a
uniform sidelobe pattern has been achieved throughout the entire time range bins except
at the beginnings and ends of the sidelobe sequences. Now the uniform sidelobe, which
is decided solely by the signal code length, is kept at a constant level consistently over
whole sidelobe ranges. Apart from four ripple peaks that appear at both ends of the
sidelobes, the uniform levels, which now can be considered as the PSR, are 46 dB and 60
dB respectively. These exactly correspond to $20\log_{10}(200)$ and $20\log_{10}(1000)$, the Barker
code level for these signal code lengths. The existence of those ripples has been predicted
from the mathematical expressions in Eqns. (6.16) and (6.22) which are assumed to work
only for large $q$ values. The highest peak ripple reaches 2 dB above the uniform level,
which is not significant.

Another important property to notice is the complete isolation of the mainlobe by the
Figure 6.7: s-P4 code pulse compression outputs by the Woo filters for code lengths $N = 200$ and 1000.

Figure 6.8: Illustration of the mainlobe isolation distance with respect to the pulse code length $N$. 
6.2 Doppler Shift Effect

The pulse compression responses of Woo filters corresponding to various Doppler frequency shifted signals are shown in Figure 6.10. In Figure 6.11 are shown the ambiguity functions for different ranges of the Doppler frequency shift $\Delta f_d/B$. The Woo filter is a linear combination of the matched filters for linear FM derived phase codes. Therefore it would be logical to anticipate that its response would be similar to that of linear FM, which

Figure 6.9: Mainlobe comparison between s-P3 and s-P4 correlation cases.

extremely low sidelobe level around it, which gives rise to a mainlobe sharpening effect. This is due to the fact that, for a small $q$, the two separate summations $\Omega_1$ and $\Omega_2$ in Eqns. (6.16, 6.22) have nearly identical values. This phenomenon is clarified in Figure 6.6(a). The benefit of this kind of strong mainlobe protection will be received mostly by the target detection capability, when strong scatterers are densely mixed with each other [5]. The range distances between the mainlobe and sidelobe, which is created by the isolations, are $4f_{3dB}$ for $N = 200$ and $10f_{3dB}$ for $N = 1000$ cases. Eqns. (6.19, 6.20) indicate that, as $N$ increases, $\Omega_1$ and $\Omega_2$ approach each other more closely and the mainlobe isolation distance is widened. Figure 6.8 illustrates their relationship and confirms that the isolation distance increases linearly in proportion to the pulse code length $N$.

Unlike the s-P4 code, the s-P3 code does not have a symmetric frequency sweep. Hence the ACF of s-P3 and s-P4 differ in phase terms. The subtle difference between the two cases is illustrated in Figure 6.9. The s-P4 modulation is clearly preferable, since both mainlobe peaks are in phase with each other.

Because the Woo filter is one bit shorter than the processed signal and its structure is symmetrical in the time domain, the mainlobe width is broadened by a factor of two. This can be understood as the signal power in the sidelobe part has been pushed towards the mainlobe area by the Woo filter. Although this still remains as the major drawback of this technique, it may be easily compensated by increasing the signal bandwidth.

6.2.4 Doppler Shift Effect

The pulse compression responses of Woo filters corresponding to various Doppler frequency shifted signals are shown in Figure 6.10. In Figure 6.11 are shown the ambiguity functions for different ranges of the Doppler frequency shift $\Delta f_d/B$. The Woo filter is a linear combination of the matched filters for linear FM derived phase codes. Therefore it would be logical to anticipate that its response would be similar to that of linear FM, which

$^3f_{3dB}$ represents 3dB mainlobe width in the time domain.
6.2 Doppler Shift Effect

![Graphs showing Doppler shift effects](image)

Figure 6.10: Woo filter outputs for Doppler shifted signals. Doppler shift $f_d = 1, 2, 3$ and $4\%$ of the total signal bandwidth $B_d$. Pulse code length $N=100$.

is very robust to the Doppler frequency shift effect. However, the discrete phase structure makes its performance deviate from the original linear FM characteristic. Some distinctive characteristics are found regarding the mainlobe, ripple peaks and range direction shift.

Setting the zero Doppler shift response as a reference level, the mainlobe peak variation is investigated in Figure 6.12. The mainlobe peak affected by the Doppler shift shows a periodic behaviour. The peaks are not necessarily lower than the zero Doppler case.

Although, like the linear FM signal, mainlobe peaks are shifted away from the zero Doppler peak point with increasing Doppler shift, it follows a stairway curve rather than a smooth linear graph seen in the linear FM case. This is illustrated in Figure 6.13(a). The Doppler shift range is set as -2 to 2 [$\%$] of the signal bandwidth. The fact that the sampling rate is 1 for each phase element may have provoked a loss of the frequency shift information. However, even if oversampling is imposed on the received signal, the discrete nature of the phase distribution in the signal would lead to the same range shift effect. A comparison with the linear FM case is made in Figure 6.13(b) after two samples per code element are taken and, confirms this claim.

The peak ripples (PR) appearing at the edges of sidelobe also change according to the frequency shift. The sidelobes at both sides behave in symmetric ways. It is seen in Figure 6.14 that the PR follows a linear graph versus Doppler shift with a slope of 1.02 (which can be approximated to 1) in the positive frequency domain, while it stays at a constant level in negative frequency region. This constant level is about 1 dB higher than the uniform sidelobe level.
6.2 Doppler Shift Effect

Figure 6.11: Ambiguity function plot for $N = 100$ and Doppler shift (a) $-2<\frac{\Delta f}{B} < 2[\%]$ and (b) $-5<\frac{\Delta f}{B} < 5[\%]$

Figure 6.12: The effects of Doppler shift on pulse compression output in range direction shift and peak mainlobe response

Figure 6.13: Doppler shift effect on Woo filter outputs in the range direction shift. $f_d = [-0.02B, +0.02B]$. (a) One sample per code element (b) After oversampling of 2 is imposed. Comparison with a plain linear FM case.
6.3 Exact Time Domain Formulation of Woo filter Response

A frequency shift in the signal creates a phase ramp rather than a phase staircase. The ramp means that there is also a linear phase shift within each bit of the incoming signal. When only one sample is taken per each bits, the pulse compression output may not fully reflect a rapid phase shift change within a relatively long pulse duration. A tapped delay line correlation may be no longer applicable when distributed targets are observed. To take distributed target environments into account, it will be necessary to derive a continuous time domain representation instead.

6.3.1 Derivation of the Doppler Shift Effect

For a given incoming signal $S(t)$ and its corresponding Woo filter $Woo(t)$, the ambiguity function $\chi_0(\tau, f_d)$ is given as

$$\chi_0(\tau, f_d) = \int_{-\infty}^{\infty} S(t - \tau) Woo(t) \exp(j2\pi f_d t) dt$$

where

$$S(t) = \sum_{n=0}^{N} \exp(j\pi \frac{n^2}{N}) U \left[ \frac{t - (n + \frac{1}{2}) t_b}{t_b} \right]$$

$$Woo(t) = \sum_{n=0}^{N} \frac{WOO_n \cdot U}{t_b} \left[ \frac{t - (n + \frac{1}{2}) t_b}{t_b} \right]$$

To proceed with the discrete summation, it is convenient to divide the time delay such that $\tau$ represent a subpulse within the range $0 \leq \tau < Nt_b$. Let $\tau$ be designated as

$$Mt_b \leq \tau \leq Mt_b + \varepsilon,$$

where
Then the ambiguity function in Eqn. (6.24) can be rearranged as

\[
\chi_0(\tau, f_d) = \sum_{n=0}^{M-1} \int_{nt_b}^{(n+1)t_b} WOO_n \cdot \overline{S}(N - M + n) \exp(j2\pi f_d t) dt \\
+ \sum_{n=0}^{M-1} \int_{nt_b + \varepsilon}^{(n+1)t_b + \varepsilon} WOO_n \cdot \overline{S}(N - M + n + 1) \exp(j2\pi f_d t) dt \\
= \sum_{n=0}^{M-1} WOO_n \cdot \overline{S}(N - M + n + 1) \exp\left[j2\pi f_d (nt_b + \varepsilon) \right] \frac{\sin(\pi f_d \varepsilon)}{\pi f_d} \\
+ \sum_{n=0}^{M-1} WOO_n \cdot \overline{S}(N - M + n) \exp\left[j2\pi f_d (nt_b + \frac{\varepsilon}{2}) \right] \frac{\sin[\pi f_d (t_b + \varepsilon)]}{\pi f_d} 
\]

(6.29)

Here \(WOO_n\) and \(\overline{S}(n)\) represent the nth term of the discretely transformed signal of \(S(t)\) and \(Woo(t)\) respectively. It is the parameter \(\varepsilon\) that distinguishes Eqn. (6.29) from the plain discrete ambiguity equations.

### 6.4 Frequency Spectrum

The low efficiency in utilizing the signal bandwidth is another disadvantage of using phase coded waveforms. Unlike the continuous linear FM signal of which the power spectrum is well defined within the designated bandwidth area, the power spectrum of the phase coded signal spreads out indefinitely over the whole frequency domain. To implement a filter utilizing all the energy in the received signal is not a feasible scheme.

There exist two different factors that may cause SNR loss during the Woo filter correlation procedure. Since the Woo filter is not perfectly matched to the received signal, a Signal-to-Noise ratio (SNR) loss is expected at the correlation output. On the other hand the non-realizable filter characteristic of infinite bandwidth is another source to add further loss to the SNR.

The output signal to noise ratio (SNR) in the pulse compression process is defined as the ratio of the peak instantaneous output signal power to the output noise power, and is represented as

\[
\frac{S}{N} = \left| \frac{Max \left( \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) * H(\omega) \exp(j\omega t) d\omega \right)}{N_0 \nabla f} \right|^2
\]

(6.30)

\(N_0\) is the white Gaussian noise amplitude in the frequency domain and \(\nabla f\) is the bandwidth of the given compression filter. The SNR loss is defined as the reduction of
6.4 Bandwidth Limiting

A phased code waveform can be considered as a group of rectangular subpulses each having their own phase components. And these types of subpulses have a baseband spectrum envelope of \( \sin(\pi f T_c)/\pi f T_c \). The selected phase coding determines the fine structure of the overall frequency spectrum.

In Figure 6.15, the frequency spectra have been plotted for s-P4 codes having different code lengths. Their appearances look more symmetrical than those of the P codes, which is attributed to the symmetric nature of the phase distribution within the signal. The increased code length brings about a smoothing effect on the spectrum envelopes.

6.5 Mismatching SNR Loss

Typically there exists a 1 or 2 dB SNR loss when the receiver filter for phase coded signal is amplitude weighted [72]. Although the Woo filter does not belong to the amplitude weighting class, considering the amplitude envelope along the time axis shown in Figure 6.4, a similar outcome is expected. When only a time weighting function \( W(t) \) is imposed on the receiver part, and no restriction is given on the bandwidth, the SNR loss is simply
given by

\[
\text{SNR loss} = \frac{\left[ \int_0^{Nt_b} W(t)dt \right]^2}{Nt_b \int_0^{Nt_b} W^2(t)dt} \quad (\because \text{Parserval's theorem})
\] 

(6.31)

For discrete phase code cases, Eqn. (6.31) is modified into a discrete summation expression.

\[
\text{SNR loss}_{\text{Discrete}} = \frac{\left( \sum_{i=1}^{N} W(i) \right)^2 \cdot t_b^2}{Nt_b \sum_{i=1}^{N} W(i)^2 t_b} = \frac{\left( \sum_{i=1}^{N} W(i) \right)^2}{N \sum_{i=1}^{N} W(i)^2}
\]

(6.32)

Without any amplitude weighting, the highest attainable SNR for uniform discrete phase codes is

\[
\text{SNR}_{\text{max}} = \frac{Nt_b}{N_0}
\]

(6.33)

where \( N \) is the number of code elements, \( N_0 \) is the input noise power per unit bandwidth and \( t_b \) is the length of one bit subpulse.

Although the Woo filter is originated from a simple time-domain superimposition of correlators, it is defined in a totally new manner in both time and frequency domains. As shown in Figure 6.4 its phase as well as the magnitude characteristics are distinguished from the original signals.

The receiver input noise is assumed to be white Gaussian, having a power density in the entire frequency region of \( \frac{N_0}{2} \text{W/Hz/s} \). When frequency weighting function is denoted as \( W(f) \) with bandwidth \( \nabla f \), then the average output noise power is given as

\[
\sigma^2 = \frac{N_0}{2} \int_{-\frac{\nabla f}{2}}^{\frac{\nabla f}{2}} W^2(f) df
\]

(6.34)

Like a typical phase code, the Woo filter can also be considered as a combination of multiple rectangular subpulses each of which has its own unique amplitude and phase property. The procedure to obtain a frequency weighting function corresponding to the
Woo filter is described as

\[ W(f) = \int_{-\infty}^{\infty} \bar{W}_o(t) \cdot \exp(-j2\pi tf) dt \]

\[ = \sum_{p=1}^{N} \bar{W}_o(p) \cdot \int_{(p-1)t_b}^{pt_b} \exp(-j2\pi tf) dt \]

\[ = \sum_{p=1}^{N} 2j\sin\left[\frac{\pi}{N}(p - \frac{1}{2})\right] \cdot \exp\left[-j\frac{\pi}{N}(p^2 + p - \frac{1}{2})\right] \cdot \left\{ \exp[-j2\pi ft_b(p + 0.5)] \cdot \frac{\sin(\pi ft_b)}{\pi f} \right\} \]  

(6.35)

Figure 6.16 illustrates the frequency spectrum characteristic of the Woo filter calculated through Eqn. (6.35). It has a sinc function based shape, but the main beam is split into two separate ones. This phenomenon is understood from the nature of the Woo filter generation. It has been created by a subtraction or addition of two similar polyphase sequences and during the process, their DC components are removed. The lack of DC component causes a deep hole to emerge around 0 Hz and justifies the split of main beam into two symmetrical parts.

The SNR loss by use of the Woo filter is calculated using Eqn. (6.32). The best
attainable SNR remains the same, but the actual SNR for Woo filter should be adjusted.

\[
\text{SNR loss}_{\text{Woo}} = \frac{1}{N_t} \left( \max_{t} (S(t) \otimes \text{Woo}(t))^2 \right) dt
\]

\[= \frac{\sum_{i=1}^{N} \text{conj}(\text{Woo}(i)) \cdot S(i)^2 \cdot t_b^2}{N_t \sum_{i=1}^{N} |\text{Woo}(i)|^2 t_b} = \frac{\sum_{i=1}^{N} \text{Woo}(i) \cdot S(i)^2}{N \sum_{i=1}^{N} |\text{Woo}(i)|^2} \quad (6.36)
\]

Eqn. (6.36) is a discrete form expression of the SNR loss when Woo filter is used at the receiver. Figure 6.17 is a graphical representation of the relationship between the SNR loss and the pulse compression ratio. Although the SNR loss appears to grow with the increase of the pulse compression ratio, it eventually stabilizes and converges to a constant value -3.02 dB beyond \(N = 400\). The margin of the change is so small that it will be justified to say that the Woo filter has a penalty of about -3 dB SNR loss regardless of the pulse compression ratio or signal code length \(N\). This is a fairly predictable result from the amplitude curve of the Woo filter shown in Figure 6.4. When relatively strong target objects are in the view of the radar, a small loss in SNR does not pose severe degradation on the overall image quality.

### 6.6 Performance Enhancement

The Woo filter has a symmetric structure and the code length is one bit shorter than the original signal. This property generates two identical peaks at the correlation output resulting in a resolution degradation by half. As long as the Woo filter is obtained through the summation (s-P4 case) or subtraction (s-P3 case) between one bit shifted phased codes, the resolution sacrifice cannot be avoided.

#### 6.6.1 Resolution Compensation Strategy

However, if only the symmetric nature of the Woo filter is abolished, then the undesired mainlobe broadening may be prevented. The problem is, then, to find an asymmetric filter that best imitates the Woo filter characteristics.

From Eqn. (6.15) it could be argued that, for large \(N\) the phase variation of Woo filter is similar to the original waveform, although its code length is one bit shorter than the original waveform. Consequently an intuition may be drawn that, if the magnitude curve of the Woo filter is taken as an amplitude weighting function to be imposed on the original waveform in the signal code domain of either \(\{1, ..., N - 1\}\) or \(\{2, ..., N\}\), a good approximation to the Woo filter could be made, but in a non-symmetrical way.
6.6 Weighting Approximation Curve

6.6.2 Weighting Approximation Curve

From the Woo filter equation the pure amplitude weighting function may be directly derived. The \( p \)th element of the Woo filter array is

\[
W_{ooN}(p) = \begin{cases} 
\exp(j\pi \frac{p^2}{N}) - \exp(j\pi \frac{(p-1)^2}{N}) & ; \text{s-P3 code} \\
\exp[-\frac{i\pi}{N} p^2 + j\pi p] - \exp[-\frac{i\pi}{N} (p-1)^2 + j\pi(p-1)] & ; \text{s-P4 code}
\end{cases}
\]

(6.37)

where \( p = 1, 2, ..., N - 1 \). In both cases, the magnitude components \( |W_{ooN}(p)| \) are given as

\[
|W_{ooN}(p)| = \left| \exp\left(j\pi \frac{p^2 + p + 1/2}{N}\right) \cdot \exp\left(-j\pi \frac{p + 1/2}{N}\right) - \exp\left(j\pi \frac{p + 1/2}{N}\right) \right|
\]

\[
= 2 \sin \left( \frac{\pi}{N} \right) \left( p + \frac{1}{2} \right),
\]

(6.38)

Based on the intuitive prediction, a correlation simulation is performed for a non-symmetric Woo filter. Firstly, it is verified that the phase characteristic of the asymmetric Woo filter is very closely related to that of the original signal. Figures 6.18(a,b) show the phase graphs of the s-P3 and s-P4 codes having code length of 100 and compare with the phases of the asymmetric Woo filters of code length 99. Very good agreements are found between the phase curves, especially in the interval where the phase variation rate is low.

Figure 6.18(c) shows the pulse compression output generated by the asymmetric Woo filter for a s-P4 signal of code length \( N = 576 \). It is very obvious that the new pulse compression scheme makes a considerable improvement in the range resolution. The overall sidelobe level has been further suppressed to -57.3 dB, which is even below the optimum Barker code level (\( 20 \log_{10}(576) = 55 \text{ dB} \)). The low uniform sidelobe property is well preserved, although the most prominent achievement is found in the reduced mainlobe width.

Figure 6.19 illustrates the pulse compression outputs generated by the asymmetric Woo filters for different phase code lengths and compared with the symmetric Woo filter cases. It is seen that even the highest ripple peak is below the uniform sidelobe level achieved by the Woo filter. The new uniform sidelobe is found to be lower by 2 dB than the corresponding optimal Barker code level. However it is found that 2 dB of further mismatching loss is incurred regardless of the code length. As a result, the actual sidelobe level remains the same as the original Woo filter response. The mainlobe width is reduced to the one element phase code length \( t_b \). Now the 3 dB pulse width of the pulse compression output is equivalent to that of the plain linear FM case, while the sidelobe level is optimally suppressed. On the other hand, the mainlobe isolation effect is no longer available. It can be argued that the reduced mainlobe energy is forced to spill into nearby region as a result
Figure 6.18: Phase comparison of asymmetric Woo filter with the original cases of phase code length 100 for (a) s-P3 waveform and (b) s-P4 waveform. (c) Pulse compression output using asymmetric s-P4 code of length N=576.

Figure 6.19: The pulse compression outputs generated by asymmetric Woo amplitude weightings and comparison with original Woo filter cases.
Figure 6.20: Reduction of sidelobe peak ripples through adjustment of Woo filter elements

6.6 Peak Ripple Removal

Although the impact of peak ripples on pulse compression performance is trivial, it would serve an academic interest to make an attempt to remove it. The graphical illustration of peak ripple reveals that its behaviour resembles a unit pulse response seen in typical control systems, where one of the common techniques to reduce the peak response is to slow down the rising time through a feedback control. Adopting a similar idea, some elements in the Woo filter are modified to emulate this control technique. Figure 6.20 illustrates two examples produced by adjusting the first 8 elements in the Woo filter. The adjustment has been done by assigning unique gain weightings to each element. The maximum improvement is no more than 1 dB reduction in the peak ripples while unwanted ripples are found to spread over the whole sidelobe interval. This phenomenon also corresponds to the result experienced in typical control system responses when similar attempts are made. Considering that the mainlobe isolation is attributed to the existence of the peak ripples, it would be more useful to accept them as an unique property of Woo filter.

6.7 Non-Perfect Phase Code

In a pulse compression radar system, signal distortions or imperfect signal processing can cause performance degradation [71]. The radar transmitter, receiver components or
6.7 Phase Error

Figure 6.21: Impact of arbitrary phase errors on uniform sidelobe property.

the transmission media are the typical sources of the signal distortion. Considering that the unique property of Woo filter relies on the precise manipulation of signal phase, it will be worthwhile to investigate what may influence and alter the signal phase distribution.

### 6.7.1 Phase Error

For signals with rectangular envelopes of duration $T$ the autocorrelation function for the main peak is given as [71]

$$\chi(0, f_d) = \frac{\sin(\pi f_d T)}{\pi f_d T} = \frac{\sin(\pi \Delta \theta)}{\pi \Delta \theta}$$  \hspace{1cm} (6.39)

Here $\Delta \theta$ is equivalent to the total phase error in wavelength. Therefore the concept of Eqn. (6.39) can be extended to include the time domain phase error effect, which may be written as

$$\chi(\tau, f_d) = \frac{\sin(\pi N \Delta \theta_n)}{\pi N \Delta \theta_n}$$  \hspace{1cm} (6.40)

where

$$\Delta \theta_n = \frac{\Delta f_n}{f}, \quad \Delta T_n = \frac{\Delta T}{T}$$  \hspace{1cm} (6.41)

$N$ is the total code length and $\Delta \theta_n$ is the fractional phase error within one code bit in the time or frequency domain.

However the Woo filter correlation is not a matched filter procedure and Eqn. (6.39) cannot be used directly to assess the performance deterioration. The ambiguity function in Eqn. (6.29) has been adopted instead for this purpose. Simulations have been performed using Monte Carlo techniques to describe the random phase error effects and Figure 6.21 shows the results. The average random phase error is given as the percentage ratio to the minimum phase variation within the waveform or $\Delta \theta_n / \frac{2\pi}{N}$. It is seen that as long as
6.7 Random Noise Effect

Figure 6.22: The effects of random phase and amplitude errors on compressed pulse output

the phase error is limited to the minimum signal phase variation, the Woo filter responses maintain the unique performance. Although the uniform sidelobe property weakens for large phase errors, the mainlobe peak remains without a significant loss. Naturally an increase in signal code length increases the sensitivity to the random phase error.

6.7.2 Random Noise Effect

The transmitted signals are vulnerable to various noise sources. To determine the sensitivity of the waveform signal to arbitrary variations in amplitude and phase, uniformly distributed error signals are generated and added to the signal. Then simulations are performed to estimate the changes at the compression outputs. Assuming the noise signals are represented by arbitrary random noise vectors, the Monte Carlo simulation technique is adopted. The random noises are assumed to affect both I and Q signal components such that the noise signals are represented as vector forms. The maximum amplitude of random noise is set as a percentage relative to the signal amplitude. Figure 6.22 shows the simulation results. For 2% of noise level relative to original signal, despite the corruption of the uniform sidelobes, the overall performance is not critically affected. The random noise-like spikes are kept below the highest ripple level. As the noise level grows up above 5% of the signal level, the uniform sidelobe characteristic starts to show a considerable disruption due to increased noise levels. These effects naturally result in a reduced peak to sidelobe ratio (PSR).
Figure 6.23: Peak response levels and peak to sidelobe ratio variations for Woo filter and matched filter cases corresponding to increased noise levels. Pulse code length is 200.

The peak sidelobe level increases consistently with the increase of noise level. Figure 6.23 illustrates the relationship between the noise level and the PSR. In both the matched and the Woo filter cases, the peak response maintain its original peak value in the presence of strong noise signals. As for the PSR property, however, the performance of Woo filter deteriorates along increased noise level suggesting a relatively high sensitivity to the noise effect. The PSR for the matched filter remains quite stable with about 3 dB reduction in PSR for 50% of noise level intrusion. However the PSR of Woo filter output degrades linearly by 20 dB from -43.9 to -24.6 dB. The crossover point of two PSR curves is measured as -28.6 dB at 32.6% noise level. This level of noise error is far higher than actually experienced in reality.

6.8 Sampling Effect Consideration

6.8.1 Error Sources

The act of expanding a function into a sampling series inevitably involves a possibility of incurring various errors. Several different kinds of errors have been identified and studied intensively over past decades [35]. The main ones are

- **Truncation error**: occurs when all the samples outside a finite interval are ignored and always happens in practice. If some of the data inside the interval is missing, it leads to information loss error.

- **Amplitude error**: occurs when the retrieved sample value is different from the correct value. In practice, the sample values are usually calculated as an average over a small interval.

- **Time jitter error**: arises from a possible difference between available sample points and correct sets.
When a digital pulse compression technique is employed in association with data stored in memory allowing limited number of quantization levels, a precise control of sidelobe pattern has to be compromised.

### 6.8.2 Time Mismatching

When the incoming signal is sampled and arranged to be correlated with the compression filter, it is not always guaranteed that the two sets are perfectly lined up with each other in the time domain to produce theoretical results. There may exist some cases when the two sets are misaligned due to the random delay in designating the sampling points and the compression output deviates from the expected behaviour. This effect becomes obvious when distributed targets are concerned, where multiple signals from scattered target objects overlap.

In Figure 6.24 are plotted several graphs, each corresponding to different time shifts. The time shift is represented as the ratio of time misalignment $t_a$ to $t_b$, which is the length of one subpulse constituting the phase code. Up to 5% of time-jitter error permits no significant corruption either in the uniform sidelobe pattern or 3 dB width. A time shift error above 10% brings about a noticeable distortion in the sidelobe part. Also the uniform sidelobe characteristic is lost. When $t_a/t_b$ is 0.5, which is an extreme case, the 3 dB width is widened by a factor of two. However no significant corruption is found below 10% of time shift in terms of the 3 dB width.

- **Aliasing error**: arises when the number of sample per band is insufficient, i.e. *undersampling error*.
6.8 Quantization

Figure 6.25 illustrates various pulse compression outputs when the quantization schemes are introduced to incoming phase code signals, as well as the Woo filters. The pulse bandwidth used for ERS-1,2 is 15.5 MHz while the pulse duration is 37.1 μs. These specifications generate a phase code length or the pulse compression ratio $BT$ of 573. From $9 < \log_2(573) = 9.16 < 10$ at least 10 bits of quantization is required theoretically to fully implement this phased code waveform. Figure 6.25 demonstrates that an 11 bits quantized signal produces a uniform sidelobe pattern. Actually, even this degree of quantization does not fully regenerate the original outputs, but still there exists a small amount of fluctuation of 1 dB throughout the sidelobe areas. In fact, the number of required phase codes in constructing the s-P codes are not exactly decided by their code length since they are derived from linear FM conversion of which the phase distribution follows a square function. Hence to assign a precisely quantized phase value to each code elements is not practically feasible and some errors are always induced.

Since the Woo filter is also an amplitude variant function, it is necessary to consider the effect of the amplitude quantization effect as well. Figure 6.26 illustrates the compression outputs when the Woo filters are quantized by different levels. Since the amplitude envelope of the Woo filter can be approximated to a simple smooth sine function, less complicated and less quantization levels are required than in the phase quantization case. It is confirmed that for a code length of 500, 6 bits of amplitude quantization produces a sufficiently good uniform sidelobe and the 7 bit quantization virtually eliminates any quantization effect.

To investigate the specific relationship between quantization bits and PSR levels of compression outputs, Figure 6.27 is presented. The PSR versus quantization bits are shown for the various code lengths. All the graphs follow knee shaped curves which begin to flatten from around 5-bits. Similar type of curves are illustrated in Figure 6.27(b) to graphically relate the magnitude quantization with performance degradations. As in
Figure 6.26: Comparison of magnitude quantized Woo filter compression performances. Number of $Q_{n,mag}$ are 5, 7, 8 bits for phase code length $N=573$.

Figure 6.27: Woo filter magnitude and phase quantization error effects on PSR. Phase code lengths are given as 100, 300, 500 and 700. Comparison of (a) Phase quantized Woo filter and (b) Magnitude quantized Woo filter compression performances.
6.9 Comparison with Linear FM

6.9.1 Weighting function

It has been explained that, although there is such a close relationship between linear FM and P-codes, the unique sidelobe property of discrete phase code is not shared by linear FM. However since the amplitude weighting technique does not make use of the phase parameters, it may be directly applied to the continuous linear FM signal. For a linear FM signal $S(t)$ of which the time duration is $T$, the time domain amplitude weighting function equivalent to Woo filter is given as

$$S(t) : \exp(j\pi \frac{B}{T} t^2) \quad \Leftrightarrow \quad W(t) : \sin(\pi \frac{t}{T})$$

It is difficult to expect a continuous signal to have the uniformly distributed sidelobe pattern seen in the phased code case. But their close connection raises a possibility that the compression output of Woo-weighted linear FM may have a sidelobe that resembles the pattern shown in Figure 6.18. In fact it is found that a ‘sin’ weighting in the time domain generates a mixture of the two cases. The pulse compression outputs are compared in Figure 6.28.

The sidelobe fluctuation is neatly suppressed in a simple curve, thus optimizing the PSR. The rapid decay of sidelobe power should also be noticed. The inevitable sacrifice of 3 dB mainlobe width broadening is detected and measured as 1.2 wider than the case without
any weighting. The unique property of the Woo filtering that pushes away the undesired energies around the mainlobe and dissipates them throughout the whole sidelobe domain, is also found. However the dramatic PSL reduction proportional to the code length, as seen in the Woo filter processing for the s-P codes, is not likely to occur in linear FM case.

'Thers is something fascinating about science. One gets such wholesale returns of conjecture out of such a trifling investment of facts' -Mark Twain-
Chapter 7

Integrated Sidelobe Energy Reduction

A major technical barrier to the use of pulse compression in the surveillance radar is the presence of range sidelobes upon pulse compression and the data corruption introduced by these sidelobes. This range sidelobe corruption can hamper the observation of weak phenomena occurring near strong or extended phenomena [58]. In many cases, waveform designs for pulse compression radar regard peak sidelobe level as the most critical parameter in measuring target detection performance. This may be true when simple point target scatterers are observed but when a widely distributed target area is present within the view of the imaging radar, the peak sidelobe level (PSL) criterion alone is not a sufficient performance measure. What becomes more important in the distributed target scene such as seen in SAR image is the integrated sidelobe level (ISL) rather than the PSL. Here we discuss how to evaluate the integrated sidelobe energy and how to reduce it by means of the Woo filter.

A conclusion is drawn that the sidelobe energy can be further reduced at the pulse compression output through a cancellation scheme. This is made possible when the received signal is driven into two different paths each to generate their own sampled data sets, which are then correlated with appropriate filters respectively. The two outcomes are combined together to produce a new composite pulse compression result in which unwanted sidelobes are significantly reduced.

7.1 Integrated Sidelobes

The so-called ‘Radar uncertainty principle’ is a well known property of radar ambiguity function $\chi(r, \phi)$, which is formulated as

$$\iint_{-\infty}^{+\infty} |\chi(r, \phi)|^2 dr d\phi = |\chi(0, 0)|^2$$

(7.1)
Figure 7.1: Two dimensional sidelobe generation algorithm SAR processing

Eqn. (7.1) is sometimes called the 'law of the conservation of ambiguity' and implies that all signals are equally good or bad as long as they are not compared against a specific environment [15]. In a spaceborne SAR environment where the Doppler frequency shift imposed on signal is not significantly high, this integrated sidelobe energy is regarded as a less important criterion than other waveform properties. SAR processing requires two mutually independent correlation processes in range and azimuth directions. Both correlations are based on the conventional linear FM pulse compression and create unavoidable sidelobes in each dimension. Figure 7.1 is a two dimensional illustration of how sidelobes are generated by the SAR processing. The signal compressions in range and azimuth directions produce sidelobe patterns similar to the linear FM compression output although each in direction the signals are amplitude-weighted by the antenna sidelobe patterns. This effect should also be considered in deriving a exact point target response.

Although each direction processing is independent from the other, it is the range sidelobe that provides a sources for azimuth processing such that total sidelobe field stretches out widely over two dimensional domain. Hence, though the azimuth compression is entirely dependent on the azimuth antenna pattern and PRF, the unwanted sidelobes may be reduced by suppressing the contribution from range sidelobes, which can be done by the range sidelobe suppression. Therefore it can be claimed that a good choice of waveform and range sidelobe suppression methodology may affect the azimuth processing result and final image quality.

In the ordinary point target detection the importance of ISL is not so much emphasized as PSL, since with a good compression scheme, the finite amount of the sidelobe energy rapidly attenuates and spreads away over widely extended area. However in SAR imagery all the point target responses mutually affect each other through spread sidelobes. As the frequency of sidelobe overlap increases, the sidelobes begin to accumulate up to the level that cannot be completely ignored. Unlike natural vegetation areas covered with low reflection level clutter, most urban areas are comprised of wide dynamic range scatterers.
7.1 Point Target Discrimination

The images shown in Figure 7.2 are good examples of a SAR image containing a strong target object. Figure 7.2(a) is taken by the X-SAR system and resulted from four-look processing. The overall appearance is divided by two distinguished features of small but bright residential area and dark vegetation land. The existence of a strong scatterer, which is thought to be a artificial man-made structure, in the clutter region signifies a typical case that is rarely seen in natural environment but commonplace in urban scenes. Figure 7.2(b) is a magnified look around the point target within the dashed line of (a). There are seen a few closely located objects but it is very difficult to examine these fully since resolution cells in both range and azimuth directions are corrupted by strong sidelobes. Due to this strong reflector, this image fails to deliver a detailed information on the target scene. When such objects appear frequently forming a dense population of strong reflections, the image interpretation and target recognition will be severely confused. In ordinary pulse compression a typical threshold level that masks the surrounding area is around 45 dB\(^1\). The point scatterer shown in Figure 7.2 has a peak response level 50 dB higher than the surrounding area and becomes a dominant feature.

---

\(^1\)This has been drawn from the sidelobe pattern of the ordinary linear FM waveform pulse compression as well as empirical data sets of SAR images [30].
7.1 Integrated Sidelobe Level

The PSL may work as a sufficient measure for the point target detection case. However when a widely distributed target area is concerned for radar imaging purpose, other criteria may have to be considered.

The integrated sidelobe energy is defined as the energy contained in the sidelobe region or the total unwanted signal energy outside main peak area. If the sidelobe region is assumed to begin beyond the 3-dB point of the compressed output, the ISL for range pulse compression may be given as

$$\text{ISL (dB)} = 10 \log_{10} \left( \frac{\int_{-\infty}^{t_{+3dB}} |\chi(t)|^2 dt + \int_{-\infty}^{t_{-3dB}} |\chi(t)|^2 dt}{\int_{t_{-3dB}}^{t_{+3dB}} |\chi(t)|^2 dt} \right)$$

(7.2)

This can be further developed to include Doppler shift effect by simply replacing the range compression output $\chi(t)$ with $\chi(t, f_d)$ and taking double integrals over time and frequency.

Eqn. (7.2) alone may be useful when target detection is performed within the one dimensional spatial domain. But SAR processing requires two separate compressions and the concept of integrated sidelobe energy around or outside the mainlobe area should be modified to provide a proper measure.

The received signal $S'$ is a summation of the reflections $S_i$ from the distributed target scenes of $M$ each of which is collected at regular interval along the azimuth path.

$$S(t_R, t_{Az}) = S_1(t_R - t^i_{R} + t^i_{Az}) + \ldots + S_M(t_R - t^M_{R} + t^M_{Az}, f_D)$$

$$= \sum_{i=1}^{M} S_i(t_R - t^i_{R} + t^i_{Az}, t^i_{Az})$$

(7.3)

Each $t_R$ and $t_{Az}$ represents the time reference for a point target location specified as $(D_R, D_{Az})$. The parameter $f_D^i$ is inserted to take into account the Doppler shift due to the antenna movement. When the target scene is assumed to be stationary and with other effects caused by payload instability ignored, the Doppler shift is purely a function of the antenna position $(D_R, D_{Az})$ relative to the target point. Thus the variable $f_D^i$ may be removed from the expression. The focusing of this signal is done by a convolution with the SAR impulse response and generates the 2-D function $\chi(D_R, D_{Az})$, which is denoted as the point target response (PTR).

$$\chi(D_R, D_{Az}) = S(t_R, t_{Az}) \otimes h(t_R, t_{Az})$$

$$= \sum_{i=1}^{M} S_i(t_R - t^i_{R} + t^i_{Az} - t^i_{Az}) \otimes h((t_R, t_{Az}))$$

(7.4)

A two dimensional point target response after SAR processing is shown in Figure 7.3. The parameters used for the simulation are based on the ERS-1 SAR specification and the
7.1 Integrated Sidelobe Level

As in the conventional one-dimensional pulse compression case, the performance of the SAR processing can be measured by the efficiency of accumulating as much signal power around the precise target point. For this purpose Eqn. (7.2) is extended to two dimensions to give a new definition of SAR integrated sidelobe energy ratio. The ISL is concerned about the distance $R$ from the target point regardless of the energy spreading direction. Therefore although the PTR consists of two independent variables, it will be convenient to transform them into new coordinate domain to include the parameter $R$ as an independent variable. Any point location denoted as $(D_R, D_{Az})$ is displaced from the origin by $R = \sqrt{D_R^2 + D_{Az}^2}$ and uniquely specified as $(R \cos(\theta), R \sin(\theta))$. Taking appropriate conversion of variables in Eqn. (7.4) and using the concept of Eqn. (7.2), the 2-D ISL expression is given by

$$ ISL_{2-D'}|_{R=R_2d} [dB] = 10 \log_{10} \left( \frac{\int_{0}^{2\pi} \int_{-\infty}^{+\infty} \chi(D_R, D_{Az})^2 \, dR \, d\theta}{\int_{0}^{R_{2d}} \chi(D_R, D_{Az})^2 \, dR} \right) $$

while the efficiency of the energy concentration at point target area is given by

$$ \Psi(R_{2d})[dB] = 10 \log_{10} \left( \frac{\int_{0}^{R_{2d}} \chi(R)^2 \, dR}{\int_{0}^{+\infty} \chi(R)^2 \, dR} \right) $$

A smaller ISL is desired to prevent energy dispersion onto surrounding zones maximizing the target detection capability. This parameter is particularly useful in mapping an urban area where very strong target objects are closely located with each other and a target response isolation becomes essential. Since the signal pattern along the azimuth direction is shaped by the antenna beam pattern and pre-determined pulse repetition rate.
7.2 CLEAN Technique Application

Another way of suppressing sidelobes further is to pass the received data from pulse compression through a sidelobe suppression filter [77][58][5]. Unlike weighting schemes which impose amplitude weightings during or before pulse compression on incoming signals, this procedure is performed after compression and can be classified as post-processing. The P code sidelobe reduction technique described in Chapter 5 is such a case. A similar example is found in meteorological radar where a special filter designed to minimize the ISL of the processed pulse is adopted [77].

Coherent destructive interference in the sidelobe responses of strong scatterers within a target can cause the image of the target to break into small pieces [64]. But when the point target response is a pre-defined function and the incoming signals are simple superpositions

<table>
<thead>
<tr>
<th>Distance/R3dB</th>
<th>1</th>
<th>3</th>
<th>5</th>
<th>9</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mainlobe[dB]</td>
<td>-1.239</td>
<td>-0.389</td>
<td>-0.196</td>
<td>-0.105</td>
<td>-0.056</td>
</tr>
</tbody>
</table>

Table 7.1: 2-D energy ratio of sidelobe region to mainlobe

Figure 7.4: 2D ISL level graphs with respect to the radius relative to 3dB distance (PRF) only, the ISL_{2D} is no longer within the entire control of the waveform design. Although the sidelobe pattern along the azimuth direction is also implemented from the similar linear FM autocorrelation processing, since the returned signal is weighted by the antenna beam pattern, its shape does not conform to the range direction sidelobe. The ambiguity effect in both range and azimuth direction is not considered here for convenience because they are not directly relevant factors in assessing the pure waveform performance. The ISL_{2-D}[dB] for the point target response of ERS-1 is calculated as -4.810[dB]. Table 7.1 exhibits 2-D ISL variation with respect to the radius as an indication of the energy concentration rate around target point.

7.2 CLEAN Technique Application

Although the sidelobe pattern along the azimuth direction is also implemented from the similar linear FM autocorrelation processing, since the returned signal is weighted by the antenna beam pattern, its shape does not conform to the range direction sidelobe. The ambiguity effect in both range and azimuth direction is not considered here for convenience because they are not directly relevant factors in assessing the pure waveform performance. The ISL_{2-D}[dB] for the point target response of ERS-1 is calculated as -4.810[dB]. Table 7.1 exhibits 2-D ISL variation with respect to the radius as an indication of the energy concentration rate around target point.
Figure 7.5: CLEAN technique algorithm for one dimensional simple target detection. Profiles of target responses from three different scatterers

of independently radiated point source responses, the original target distribution may be traced back through reverse processing. This idea, the so-called ‘CLEAN Algorithm’, was first introduced in the radio astronomy to reduce sidelobe-induced artifacts. Originally designed for non-coherent radiation, it was further extended to include coherent radiation as well [76].

Figure 7.5 is a graphical illustration of the CLEAN technique. When multiple point scatterers exist over random Gaussian noise background, there can be a case that a few dominant targets exist with individual sidelobes strong enough to shadow adjacent weak target responses. The CLEAN process spots the brightest point image and measures the complex amplitude $|s| e^{j\phi}$ corresponding to position $(u, v)$. From this, the original radiation field can be calculated and subtracted from the measured data. After subtraction the new data set is without the contribution from the strong target signal. The repeated processing with the remaining data set and removal of the brightest point response in each step, will theoretically extract the point scatterer responses only without sidelobes. The weak target responses obscured by strong sidelobes may also be retrieved as long as their SNRs are higher than pre-determined threshold value. Tsao [76] applied this concept to remove the sidelobe artifacts generated by N large random thinned array antenna and showed that absolute threshold value is given as

$$\text{Peak}_{\text{Thres.}} = \frac{1}{N} \sum_{m=1}^{M} |S_m|^2 \left(1 - e^{-\sigma^2_{\phi m}}\right) + \frac{1}{2SNR}$$  \hspace{1cm} (7.7)

where $M$ is total point target number and $\sigma^2_{\phi m}$ is the phase-error variance. It has been shown that this method works successfully when the target is considered to consists of finite number of simple scatterers. However this approach has not yet been verified for air-to-ground microwave imaging such as SAR, where the image content is largely contiguous.
Figure 7.7: ERS-1 raw data after range direction compression with particularly bright tone in the azimuth direction

and differentiation from image cell to image cell or from region to region is by texture and intensity.

7.2.1 C L E A N Process for S A R I m a g e

To investigate the feasibility of the ‘CLEAN Process’ on SAR, a SAR raw data has been taken containing a relatively strong target object on dark clutter background. The SAR raw data set was obtained by ERS-1 during its flight over southern England region 1997. It is revealed that the CLEAN technique approach, which is originally designed for reducing high speckle artifacts in target detection radar or antenna, is not readily applicable to the complicated SAR processed image.

After a rectangular range-azimuth processing is implemented, a bright point object is detected within a rural vegetation area. Figure 7.6 illustrates the resulting images. Figure 7.7 is the result after range compression only has been accomplished. The bright tone across the azimuth direction line is a typical outcome caused by a relatively high reflection cell in the region, although the detailed information on target is still not available yet.

To recover the image content without being contaminated by the strong reflector cell,

---

²By courtesy of Martra Marconi Space at Portsmouth
the 'CLEAN' procedure is enforced. Firstly the peak values are detected along the azimuth line from the processed image and their locations are traced back in the original SAR raw data. Then a simulated target response corresponding to the detected peak value is subtracted from the original raw data. Finally the remaining new raw data set is sent to the range compression step again. Subsequently the processing is reversed as shown in the flow-diagram in Figure 7.8 to repeat the process until the threshold limit is met with the maximum detectable brightness level in the image. Figure 7.9 illustrates the procedures. It describes the amplitude fluctuation along one dimensional range direction specified by the dashed line in Figure 7.7. It is found that the threshold condition is met after the third iteration is finished. In each step, the peak values and the accompanied sidelobes are removed from the raw data. It should also be noted that the overall clutter environment is not affected by the removal of the strong target signals. Unlike the simple point target detection case, SAR is a mixture of distributed clutter, and the mutual interaction among the residing scatterers is accumulated over the entire image domain. The randomly behaving low amplitude part in Figure 7.7 contains the accumulated energy from a wide area and the removal of one particular dominant target would not alter its
7.2 CLEAN Process for SAR Image

Figure 7.10: Modified range compression output after the CLEAN procedure is completed

Although this ‘CLEANed’ image naturally leads to the belief that all the signals responsible for the bright point target are completely taken out, the finally obtained SAR image proves it is not the case. Figure 7.11 is a comparison between the original and ‘CLEAN’ed image sets. Contrary to anticipation, still the bright point target is generated from the CLEANed data set, although its peak magnitude is lowered by 6.8 dB. That is 2 dB below the peak amplitude suppression shown in Figure 7.9. This is attributed to the fact that the appearance of the bright point target is due to focussing of spread energy along the azimuth line and the mere removal of the peak amplitude spots fails to detect and completely erase the relevant raw signals. The focussing algorithm has more to do with phases than amplitudes and the ‘CLEAN’ process based on the amplitude information only is not suitable for the SAR application. The CLEAN scheme may be effective when there exists an extremely strong target.

Because SAR is concerned over a distributed clutter, it is not desirable to introduce a maximum threshold criterion. To be able to see the effect of cleaning nearby area spoiled
by strong sidelobes, the dominant target may have a dynamic range of

$$\sigma_d \geq \alpha \frac{[<I> + ISL]}{PSL}$$

(7.8)

where $I$ is the mean intensity of the target reflection level. $\alpha$ is a constant that can be variable depending on the scene nature but it is found that $\alpha$ of 0.5 or higher may induce a noticeable change. In the previous example three independent strong scatterers form a bright cluster. Thus the mixed energy level is intensified by their strong sidelobes. This situation, where strong reflection points are residing closely, hinders proper CLEAN operation since it is not possible to extract a pure response without contamination by other spurious responses. However this will be one of the most frequently confronted case in urban area and the application of CLEAN technique is discouraged.

### 7.3 Sidelobe Cancellation

Although most pulse compression schemes are mainly concerned with the ACF of a single pulse, there are special cases where it will be useful to utilize two different codes at the same time. If the nature of the conventional pulse compression method sets a theoretical limit on generating near zero energy level sidelobes, a new tactic should be considered. There have been a number of attempts to achieve this goal especially among phase codes groups thanks to the relatively easy manipulation of finite unknown variables.

A simple straightforward way would be to seek multiple waveforms of the same length which generate compression outputs such that their combined sidelobes are cancelled with each other while the peak responses add up. The Welti and Golay codes [79] are among these classes. They are defined as a set of pairs with the property that produce auto-correlation functions of the same magnitudes but opposite signs in the sidelobes without affecting the mainlobes. As a result the combined correlation output contains only one single peak response which is twice as high as that of the original autocorrelation peak voltage.

### 7.3.1 Complementary Sequence

The notion of complementary sequence was first introduced by Golay [22] in connection with his study of infrared spectrometry. The property of a complementary sequence is characterized by having a perfectly zero level ACF everywhere except at zero time shift, which is known to be ideal but non-existent in conventional pulse compression. Since the discovery of complementary sequences, they have found wide applications notably in communication areas such as spread-spectrum systems or code division multiple access (CDMA).
The condition for a pair of sequences \( \alpha = \{\alpha_1, ..., \alpha_P\} \) and \( \beta = \{\beta_1, ..., \beta_Q\} \) to be a complementary sequence is

- Identical sequence length: \( P = Q = N \)
- Sidelobe \( \Rightarrow ACF_\alpha(\tau) + ACF_\beta(\tau) = \sum_{n=0}^{N-\tau-1} (\alpha_n\alpha_{n+\tau} + \beta_n\beta_{n+\tau}) = 0, \tau \neq 0 \)
- Peak Response: \( ACF(0) = 2N \)

These conditions can be called ‘Strict conditions for complementary sequence’, in that zero value is required for all sidelobe bins. This perfect sidelobe suppression may be an ideal output but not necessarily essential in radar imaging. The binary sequences \( (\{\alpha\}, \{\beta\}) = (\{-1,-1,1,1+1,-1+1,-1\}, \{+1,+1,-1,+1,-1,-1\}) \) are one of typical examples called Golay pairs. It is known and proved that infinite number of Golay binary sequence pairs can be constructed [63]. These binary sequences can be easily extended to polyphase sequences without losing the unique complementary property by multiplying identical phase sequences of the pair. The new polyphase sequence pair \( (\alpha^p, \beta^p) \) will be defined as

\[
\alpha_n^p = \alpha_n \exp\left(\frac{2\pi}{m} j - n\right), \beta_n^p = \beta_n \exp\left(\frac{2\pi}{m} j - n\right)
\]  

(7.9)

Figure 7.12 shows the resulted ACFs corresponding to \( \alpha, \beta \) and the combined polyphase sequence \( \{\alpha^p_n, \beta^p_n\} \). As mentioned, it is shown that the individual ACF for each \( \alpha, \beta \) behaves as an ordinary binary sequence while their sidelobe terms sum to zero for all range bins except zero lag. Although it may sound a very attractive solution to the sidelobe suppression problems, there are several practical difficulties in adopting this technique for SAR. These codes belong to the conventional phase code classes and are bound to suffer the usual Doppler sensitivity problems. Only a slight amount of Doppler shift will introduce enough sidelobe corruption at the sidelobes that they will no longer perfectly
cancel. Another problem, which makes this approach pessimistic, is the practical feasibility of this scheme. There is an assumption that the two autocorrelation outputs are perfectly lined up with each other, which means the two separate incoming signals should be processed in a perfectly synchronized manner. This may be feasible in some digital communication system where coded signals are precisely synchronized in time. However there is no guarantee for this degree of precision in radar environments. The difficulty becomes severe especially when multiple distributed targets are involved. Figure 7.13(a) illustrates for the cases where ratio of time misalignment $T_d$ to code bit length $T_b$ is given as 0.1 and 0.5. A mere 0.1$T_b$ of time lag appears to trigger considerable disruption in the sidelobe cancelling.

Most of all when multiple waveforms are concerned, unwanted noise is induced in the course of cross-correlation of which the energy level is comparable to that of ACF. Figure 7.13(b) compares an ACF graph with the CCF generated between the two waveforms. Although the peaks of CCF seems to be sufficiently low, their integrated energy level relative to the ACF energy reaches

$$\frac{\int |CCF(t)|^2 \, dt}{\int |ACF_{main}(t)|^2 \, dt} = -2.7 \, [dB] \quad (7.10)$$

Considering that actual sidelobe contribution comes in the form of integrated energy this level is not acceptable.

### 7.3.2 Wide-Swath Application

Range ambiguity occurs in pulsed radar system when returns from different ranges arrive simultaneously and cause difficulty in locating target object in the correct position.
The swath width of SAR is restricted in order to prevent this occurrence, although a broad swath width provides several advantages [17]. Using more than one single waveform, which is referred to as waveform-diversity, has been suggested and studied to resolve this ambiguity problem [14][17]. For a successful application these multiple transmit waveforms should be distinguished and separated from each other at the receiver, which subsequently demands to reduce the cross-correlation function (CCF) levels among the waveforms. As in the Woo filter sidelobe example, given a finite constant energy, an optimum CCF distribution would be a uniform pattern and as shown in Chapter 4, this type of output is obtained using two linear FM waveforms having opposite frequency slopes.

Figure 7.14(a) describes a waveform diversity SAR operation transmitting and receiving two orthogonal linear FM signals. Here one important principle should be pointed out, that to suppress merely the peak level of CCF is not beneficial to SAR performance. In the point target detection case, it is the peak response of CCF that determines the visibility of the main target responses. It is verified that the ratio of \( \max(CCF) \) to \( \max(ACF) \) can be lowered by increasing the pulse compression ratio or BT product. However SAR is looking over a widely distributed region and the sidelobe energy from one part tends to spill over the other’s accumulating the total energy. Therefore even if the peak CCF is successfully reduced to a minimum level, there is not much advantage in terms of ambiguity resolution. Figure 7.14(b) compares CCF levels for BT=100 and 2000 cases. Their peak CCF responses are calculated as -20 dB and -32.8 dB respectively. However, when a uniformly distributed target is considered, the total CCF level reaches up to the peak of
ACF that any meaningful target response is concealed.

As with autocorrelation cases, the construction of cross-correlation complementary pair may be found useful for this purpose. It has been proved that for any given Golay pair \( \{\alpha, \beta\} \), there exist only two correlation pairs which are given as

\[
C_{p_1} = (\beta^*, -\alpha^*); \quad C_{p_2} = (-\beta^*, \alpha^*)
\]  

(7.11)

\( \beta^* \) is the reverse conjugate of \( \beta : e^{i\phi} = e^{-i\phi} \). But again, the energy accumulation inside ACF and CCF responses from a distributed target area diminishes the merit of this approach. It is not possible to build a Golay sequence with arbitrary code length. Even if its existence is known still it is not straightforward to produce the sequence, which should be searched for with the help of computer.

### 7.4 Zero Level Sidelobe Energy Reduction Using Woo filter

Until now most of the range sidelobe reduction techniques have focused on the strategy of achieving lower peak sidelobe level. The motivation is very clear that they are aimed at detecting point targets or objects without considerable information loss. Therefore a high priority is often given to preventing unwanted strong noise in the sidelobe region that may obscure or deteriorate weak signals from other target objects. But when distributed target is concerned for remote sensing purpose, integrated sidelobe level (ISL) becomes a critical measure to assess the overall image quality.

#### 7.4.1 Sidelobe Cancellation scheme

The quantitative range sidelobe pattern is unknown until the received signal enters the pulse compression procedure and its peak response is evaluated. Although the produced image is a group of peak responses of individual signals, their peak values are already affected and corrupted by numerous other adjacent signals making it difficult to recover the correct original signals. The necessary information to estimate the original signal waveform so as to predict resulted sidelobe pattern is not known until the signal reaches the filter. Especially when the received signal is coming from a distributed target scene instead of a single point target, the precise information is only available in the original raw data and should be extracted directly before the compression process is finished.

The best way to implement a sidelobe canceller corresponding to the incoming signal from target is to extract the relevant information from the received signal itself. As illustrated in Figure 7.15, a sidelobe canceller processed in separate path may combine with the pulse compression output and produce a new output of discrete format. The Woo filter scheme used for the s-P code signal is an ideal subject for this purpose. Whilst most of the waveforms used for pulse compression generate random noise-like sidelobe
Figure 7.15: Schematic diagram of sidelobe cancelling scheme in general radar pulse compression system

\[
\begin{array}{c|c}
\text{Time delay range} & \text{Woo Sidelobe} \\
\hline
\text{s-P3 code} & 1 < q < N - 1 \\
& (-1)^N \exp \left( -\frac{\pi q^2}{N} \right) \\
& N + 2 < q < 2 + N \\
& (-1)^N \exp \left( -\frac{\pi (2N + 1 - q)^2}{N} \right) \\
\text{s-P4 code} & 1 < q < N - 1 \\
& (-1)^{N+q} \exp \left( -\frac{\pi q^2}{N} \right) \\
& N + 2 < q < 2 + N \\
& (-1)^{N+q} \exp \left( -\frac{\pi (2N + 1 - q)^2}{N} \right)
\end{array}
\]

Table 7.2: Woo sidelobe patten with respect to time delay

patterns making it hardly practical strategy to rebuild sidelobe cancelling signals, the uniform sidelobe patterns generated by Woo filter casts a promising hope of realizing this scheme.

Woo filter Sidelobe investigation

In Chapter 6 it has been shown that the sidelobes generated by a Woo filter, can be approximated by a unit amplitude curve. The Woo sidelobe curve is described in Table 7.2 in a discrete format. For a signal of code length given as N+1, the corresponding Woo filter has N components and the mainlobe position is designated as q = \{N, N + 1\}. It should be noted that < is used instead of ≤ indicating these formulation are approximate form on the condition of sufficiently large N and may not be valid for some q values which are near at each ends of the sidelobes.

Although it is very obvious that Woo sidelobe is not identical with the original signal, a close investigation of the sidelobe pattern reveals a consistent trend between original signals and the resulting sidelobe outputs. In fact, from the phase variation of the Woo sidelobe part, it can be conjectured that the sidelobe is linked to the conjugate of the original signal. From Eqn. (6.1) original signal equations for the s-P3 and s-P4 codes of length N + 1 are derived and arranged in Table 7.3.
Table 7.3: Phase signal code values

<table>
<thead>
<tr>
<th>S-P3 code</th>
<th>Range</th>
<th>Phase Code Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ≤ q ≤ N + 1</td>
<td>[ \exp \left( \frac{\pi}{N}(q - 1)^2 \right) ]</td>
<td></td>
</tr>
<tr>
<td>S-P4 code</td>
<td>1 ≤ q ≤ N + 1</td>
<td>[ \exp \left( \frac{\pi}{N}(N - q + 1)^2 \right) ]</td>
</tr>
</tbody>
</table>

The phase code value for the s-P4 code can be rearranged as

\[
s - P4(q) = \exp \left( j \frac{\pi}{N} \frac{N}{2} - q + 1 \right)^2 \]
\[
= \exp \left( j \frac{\pi}{4} \right) \exp \left( \frac{\pi}{N}(q - 1)^2 \right) \exp \left( -j\pi(q - 1) \right) \]
\[
= (-1)^{q-1} \exp \left( j \frac{\pi}{N}(q - 1)^2 \right) \exp \left( j \frac{N}{4}\pi \right) \quad (7.12)
\]

Now it can be easily deduced that the Woo sidelobe in the range of 1 ≤ q ≤ N - 1 is related with the complex conjugate of the original signal shifted by one bit. Let the qth term of the Woo sidelobe equations for s-P3 and s-P4 be designated as Woo\_S(q)\_sp3 and Woo\_S(q)\_sp4 respectively. From Tables 7.2 and 7.3 the following relationship is derived connecting the received signal and the sidelobes of pulse compression outputs.

\[
Woo\_S(q)\_sp3 \equiv (-1)^N \exp \left[ -j \frac{\pi}{N} q^2 \right] = (-1)^N \text{conj} \{s-P3(q+1)\} \quad (7.13)
\]

\[
Woo\_S(q)\_sp4 \equiv (-1)^{N+q} \exp(-j \frac{\pi}{N} q^2) \]
\[
= (-1)^N \text{conj} \{s-P4(q+1)\} \cdot \exp \left( j \frac{N}{4}\pi \right) \quad (7.14)
\]

Eqn. (7.13) and (7.14) are valid for the first sidelobe part which satisfies q < N - 1.

While each sidelobes are perfectly symmetric in the s-P4 cases, the s-P3 code generates a sign reversed pattern at the other sidelobe region. Keeping in mind that their mainlobes are occupying two time step periods and both Woo filters and waveform signals to be compressed are of symmetrical shapes, Eqn. (7.13) and (7.14) can be modified to resolve the relationships for the range N + 2 < q < 2N, which are given as

\[
Woo\_S(q)\_sp3 = -(-1)^N \exp \left[ -j \frac{\pi}{N}(2 \cdot N + 1 - q)^2 \right] \]
\[
= -(-1)^N \exp \left[ -j \frac{\pi}{N}(N - q')^2 \right], \quad (\because 1 < q' = q - N - 1 < N - 1) \]
\[
= -(-1)^N \text{conj} \{s-P3(N + 1 - q')\} \quad (7.15)
\]

for the s-P3 code. Similarly, the Woo filter output for the s-P4 code is expressed by the original code as
7.4 Sidelobe Cancellation scheme

\[ \text{Woo}_{-}S(q)_{sp4} = (-1)^{N+q} \exp \left[ -j \frac{\pi}{N} (2N + 1 - q)^2 \right] \]
\[ = (-1)^{N+q} \exp \left[ -j \frac{\pi}{N} (N - q')^2 \right], \quad (\because 1 < q' = q - N - 1 < N - 1) \]
\[ = (-1)^N \text{conj} \{s-P4(N + 1 - q')\} \cdot \exp \left( j \frac{N}{4} \pi \right) \]  
(7.16)

Unlike the s-P3 code cases where the arrays of complex conjugate of original codes constitute the sidelobe parts, there appears an extra constant phase term for s-P4 case. This phase term is determined solely by the code length and may be ignored or replaced by a pure real constant when \( N \) is given as a multiple of integer 4. Then the phase term will be replaced as a real value of

\[ N : 4 \times (2n - 1), (n = 1, 2, \ldots) \implies \exp \left( j \frac{N}{4} \pi \right) = -1 \]  
(7.17)
\[ N : 4 \times 2n, (n = 1, 2, \ldots) \implies \exp \left( j \frac{N}{4} \pi \right) = 1 \]  
(7.18)

The variable phase terms in Eqn. (7.17) and (7.18) should be taken into account in designing the sidelobe cancellation system. The main task to be tackled in realizing the sidelobe reduction scheme will be how to generate the complex conjugate of the incoming signals and how to mix it with the generated sidelobe.

Now it is clear that Woo filter produces range sidelobes that are defined by incoming signals. The fact that sidelobes are represented by a simple manipulation of the received signal raises a possibility that sufficient information to remove unwanted sidelobe energy may be acquired during the compression procedure. The close link between signals and sidelobes shown in Eqn. (7.15) and (7.16) implies that enough information is already provided. A sufficient condition that enables to implement this scheme will be how to directly and automatically reproduce sidelobe pattern out of the complicate mixture of signals without interfering mainlobe parts.

**Sidelobe Phase Variation Verification**

The information needed to construct the sidelobe cancelling waveform are deterministic parameters contained inside incoming signals. It has been conjectured that the phase variation of the range sidelobe can be approximated by the sign reversed form of the original signal phase function. But this speculation is based on a rough mathematical approximation and should be verified. Hence prior to designing a method to construct a signal conjugate, it would be necessary to measure how accurately they resemble each other.

Figure 7.16 clarifies the similarity between an original signal and its relevant range sidelobes generated by the Woo filter. An s-P3 code with a relatively small number of
7.4 Sidelobe Cancellation scheme

Figure 7.16: Comparison between Woo sidelobe and conjugate of the incoming signal in magnitude and phase. Total code length $N = 100$. Sidelobe domain is bounded as $q = [1 \ N]$.

code length $N=100$ is chosen for clear visualization. Only first sidelobe region or $q = [1 \ N]$ is considered here. Some parts at both ends of the sidelobe deviates from the original signal while most of the region remains fitted to the uniform magnitude. In case of the phase terms the two curves are overlapped with very little variation over the whole time span. The shift effect in the right direction of Woo sidelobe relative to the original waveform is evidently noticed. Figure 7.17 visualizes their discrepancy more effectively through separate comparisons of real and imaginary parts. Now the previous claim is clearly confirmed, that the Woo filter produces a sidelobe pattern which has a excellent agreement with a one bit shifted form of original waveform. The two graphs show that the discrepancy between the two waveforms are mostly caused by the real terms. The imaginary components are completely fitted to each other while the graph of real part components confirms there exists a discrepancy at the beginning but the sidelobe converges to the original signal as the time delay increases.

The s-P4 waveform should be treated with caution. Unless $N$ is set as a multiple of 4, the sidelobe of pulse compression output cannot be linked to the original waveform by a simple operation in the real number domain. After $104=8\times13$ is given as $N$, the pulse compression has been performed. In this case the second half part of sidelobe is compared with the original waveform according to Eqn. (7.16). The code length is given as $N+1 = 105$ and the Woo filter length is as $N = 104$. Let the discrete pulse compression output be designated as $\tilde{x}(q), \{1 \leq q \leq 2N\}$. From Eqn. (7.16) $\tilde{x}(q)$ is related to the
Figure 7.17: Complex domain comparison between Woo filter compression output and original waveform.

Figure 7.18: Sidelobe verification for second half sidelobe of s-P4 pulse compression output. 
N is set as 104=8×13.
original waveform $\overline{S}(q)$ as

$$\{\overline{S}(q), q = \{N + 2, ..., 2N\}\} = (-1)^q \exp \left[-j \frac{\pi}{N} (N - q')^2 \right], \quad (\because q' = q - N - 1)$$

$$= \text{conj} \{s-P4(N), s-P4(N - 1), ..., s-P4(1)\} \quad (7.19)$$

Figure 7.18 illustrates the comparison of real and imaginary parts respectively. The second sidelobe region is compared with the reversed s-P4 code components \{s-P3(N), ..., s-P4(1)\}. Both real and imaginary part graphs show excellent agreement with each other except at the ends of real terms.

### 7.4.2 Schematic Diagram for Zero Sidelobe Implementation

Now that it is clear that the Woo filter generates sidelobes which are duplicates of complex conjugates of the original signals, a practical scheme to eliminate this predetermined sidelobe is discussed.

**i) s-P3 code**: In the first half of the sidelobe domain, the $q$th component in Eqn. (7.13) is expressed as

$$\overline{X}(q) = \exp \left[-j \frac{\pi}{N} (N - q)^2 \right] = s-P3^*(N-q+1) \quad (7.20)$$

Here $\overline{X}(q)$ is equivalent to one of the original code elements which faces the end of the signal part being correlated. Now the remaining task is how to extract the conjugate form of the signal element in a particular location. To pick up an element in a particular position can be carried out by introducing a new correlator which will be denoted as a sidelobe canceller or $SC_{sp}(n)$. From Table 7.2 and Eqn. (7.20), keeping in mind that the sidelobe is also variable depending on $N$, $SC_{sp}$ may be formed as \{-1, 0, 0, 0 ...\} for s-P3 code. The later part is decided according to the other half of sidelobe domain. From Eqn. (7.15), $q$th ($\cdot 1 \leq q' \leq N-1$) element of sidelobe is

$$\overline{X}(q') = \begin{cases} 
-\exp \left[-j \frac{\pi}{N} (N-q')^2 \right] = s-P3^*(N-q'+1) , & N = \text{even} \\
+\exp \left[-j \frac{\pi}{N} (N-q')^2 \right] = -s-P3^*(N-q'+1) , & N = \text{odd}
\end{cases} \quad (7.21)$$

which is similar to Eqn. (7.20) but with $N$ dependent sign.

The required element $\overline{X}(q)$ is found at the end of the signal part being correlated. As a result, $SC_{sp}$ should have a form of \{..., 0, 0, (-1)^N\}. The integration of two cases yields a sidelobe canceller for s-P3 code as

$$SC_{sp3} = \{-1, 0, 0, 0, ..., 0, 0, (-1)^N\} \quad (7.22)$$
ii) s-P4 code: Unlike the s-P3 case the s-P4 pulse compression generates identical results in both sidelobe domain. From Eqn. (7.14) and (7.16) the qth term of each sidelobe parts is given as

$$\bar{x}(q) \cong (-1)^N s-p4^*(N - q + 1) \exp(j\pi\frac{N}{4})$$  \hspace{1cm} (7.23)

Taking into account the phase term, the sidelobe cancelling filter for the s-P4 code is expressed as

$$SC_{sp4} = (-1)^N \exp(j\pi\frac{N}{4})\{-1, 0, 0, 0, ..., 0, 0, -1\}$$ \hspace{1cm} (7.24)

Two steps are required to extract the sidelobe cancelling waveform from the incoming signals. Firstly the conjugate form of the processing signal should be constructed out of received signal. Afterward it has to be passed through pre-determined sidelobe cancelling filter $SC_{sp3}$ or $SC_{sp4}$ in a manner precisely synchronized with the incoming signal. The generation of signal conjugate form can be done by manipulating both I and Q signals in the pre-matched filter stage. Instead of a summation between two quadrature signals, the use of subtraction will enable to achieve the desired result, which can be described as $I - Q$.

The diagram in Figure 7.19 illustrates the hardware implementation of this scheme. When a sufficient degree of accuracy is enforced on the synchronized sampling and signal shift, they will combine to produce a clean output of which the sidelobe energy is cancelled. Those area where approximate equations are not applied, will not benefit from this effect. The peak mainlobe is not affected by this operation. The procedure of adding two separate matched filter outputs has been similarly employed to realize Welti codes at coherent QPSK systems and shown to be practicable [79].

7.4.3 Reduction of Sidelobe Energy

The dramatic sidelobe energy reduction effect is well illustrated in Figure 7.20. Unlike the s-P3 cancelling, the performance of s-P4 varies depending on the length of the codes and Woo cancelling filter setup. Therefore the specific filter type should be assigned according to the code length. Assuming the algorithm is designed to work only for $N = 8n$ (\(n = 1, 2, \ldots\)), a simulation is performed for $N = 8 \times 250 = 2000$.

The sidelobe energy is reduced in a considerable amount without altering mainlobe peak, which implies that this sidelobe cancelling technique does not involve any resolution.

\[\text{By taking received signals as basis for generating sidelobe canceller, the information on received signal is kept and a precise synchronization is promised.}\]
7.4 Reduction of Sidelobe Energy

Figure 7.19: Hardware implementation of sidelobe cancelling scheme. Separate I, Q detection for conjugate signal generation and incorporation of Woo filter and SC filter.

Figure 7.20: The pulse compression output for s-P3, s-P4 codes after sidelobe canceller is introduced. N is given as 1000 and 2000 for s-P3 and s-P4 respectively.
loss experienced in most conventional sidelobe reduction techniques. The peak sidelobes appear at both ends of the sidelobe and remain in the same level as Woo filter case. This has been predicted in Figure 7.17 where a unit level discrepancy is found in the beginning of sidelobe. This difference results in $-20\log_{10}(N) [\text{dB}]$ peak sidelobe level, which means still the peak sidelobe level is decided by the code length. As the time delays increase the Woo sidelobe equation approaches to the conjugate of the incoming signal and hence the sidelobe energy is reduced further. The minimum level is found in the middle of the sidelobe where the accuracy of the approximation is maximized. Apart from increased system complexity and the loss of mainlobe isolation property seen in the Woo sidelobe, there has been no sacrifice in terms of the range resolution and peak sidelobe level.

As a quantitative measure of sidelobe energy reduction, ISLR is calculated and tabulated. Around 15 dB of total ISLR gain is achieved through sidelobe canceller compared with the Woo filter case while it is 30 dB improvement compared with the original P code case. Considering that the approximation accuracy used in sidelobe canceller derivation improves with an increased $N$, the ISLR gain is expected to arise as well. Figure 7.21 highlights a significant improvement of ISLR brought by the sidelobe canceller.

As an interesting variation, this process can be controlled with a slight adjustment on the Woo sidelobe cancelling filter such that the reduction effect happens at either side only. After removing components at either side of the sidelobe cancelling filter, simulations are performed and presented in Figure 7.22. The performance improvement from a normal P code is evident in regard to the sidelobes as well as the mainlobe isolation.
7.4 Application and Doppler Shift Effect

The various simulated results so far have validated the use of sidelobe energy reduction technique in both s-P3 and s-P4 cases. However their successful operation relies on precise alignment of the multiple phase code arrays and thus there is a fear that this technique becomes highly sensitive to the outside noise. Doppler shift effect is among the major source that might bring about a breakdown of the precise phase alignment. Here will be investigated regarding this impact. Since this technique is devised for a use particularly in SAR operation the parameters adopted for simulation work will be based on the conventional SAR system.

Using the parameters in the Table 7.5 referring ERS-1 the maximum Doppler shift

---

**Table 7.5: Parameters for ERS-1**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ERS-1 parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
<td>785 Km</td>
</tr>
<tr>
<td>Pulse Width</td>
<td>37.1 µs</td>
</tr>
<tr>
<td>Sampling Rate</td>
<td>18.97 Mbits/s</td>
</tr>
<tr>
<td>Pulse Bandwidth</td>
<td>15.5 MHz</td>
</tr>
<tr>
<td>Samples/Pulse</td>
<td>704</td>
</tr>
<tr>
<td>Antenna Length</td>
<td>11.5 m</td>
</tr>
<tr>
<td>Frequency</td>
<td>5.3 GHz</td>
</tr>
<tr>
<td>PRF</td>
<td>1680 Hz</td>
</tr>
<tr>
<td>Velocity($V_o$)</td>
<td>6628 m/s</td>
</tr>
<tr>
<td>Look Angle</td>
<td>23°</td>
</tr>
</tbody>
</table>
Table 7.6: ISLR comparison when Doppler shift is present.

<table>
<thead>
<tr>
<th>$P_{code}$</th>
<th>$Woo$ filter</th>
<th>With Sidelobe Canceller</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.5f_{d_{max}}$</td>
<td>-15.8</td>
<td>-27.4</td>
</tr>
<tr>
<td>$f_{d_{max}}$</td>
<td>-15.2</td>
<td>-26.8</td>
</tr>
<tr>
<td>$2f_{d_{max}}$</td>
<td>-13.9</td>
<td>-25.0</td>
</tr>
</tbody>
</table>

Figure 7.23: Doppler shift effect on Woo sidelobe cancelling performance. $N=576$ and $f_{d_{max}} = 1.15$ KHz.

The frequency $f_{d_{max}}$ is calculated as

$$\text{Azimuth beam width } \theta = \frac{\lambda}{D}$$

$$f_{d_{max}} = \frac{2V_s \sin(\theta)}{\lambda}$$

which results in $f_{d_{max}} = 1.1527$ KHz or 42.7° phase shift for a single pulse duration. This represents less than 0.01% of the pulse bandwidth. This prior knowledge seems to lessen the worry of the degradation by Doppler shift impact. Various simulations are performed to explore this effect as illustrated in Figure 7.23. $N$ is chosen to be equal to the system pulse compression ratio $BT = 15.5 \times 37.1 = 576$. There exists an obvious performance degradation with the increase of the Doppler shifts. However even for the case where the Doppler shift is twice higher than the maximum worst case, still a significant improvement is found in terms of overall sidelobe energy reduction. The ISLR gains are calculated in Table 7.6. Although these are performed for the case of the s-P3 case the similar results are expected for the s-P4 case.

7.4.5 Woo Amplitude Weighting Cases

It has been shown that a pure amplitude weighted Woo filter still achieves a sufficiently peak sidelobe level suppression. Furthermore its application has been further encouraged in that there is not any resolution loss during the procedure. However since this method is not based on the exact complex sidelobe cancellation, it is doubtful whether it would
Figure 7.24: Sidelobe cancelling attempt in conjunction with Woo weighting scheme. N=576 and $f_{d_{\text{max}}}$ = 1.15 KHz.

still accomplish similar sidelobe energy reduction effect. In fact the simulated result draw a disappointing conclusion.

For the cases with and without Doppler frequency shift effect the pure amplitude weighting in conjunction with sidelobe canceller produces Figure 7.24. The peak sidelobe appears in the central area of sidelobe at a level of -50 dB and there is not found any sidelobe cancelling effect throughout the entire range bins. Actually these results are worse than the case without sidelobe cancelling filter and the use of this technique should be abandoned.

‘Equations are more important to me, because politics is for the present, but an equation is something for eternity’—Albert Einstein
Chapter 8

SAR IMAGE SIMULATION

An extremely bright spot frequently found in SAR images of urban scenes is generally considered to originate from a corner reflector phenomenon, which is generated by the combination of building walls and ground surface. But at what critical height a particular surface object will act as a corner reflector to a radar signal of a certain wavelength, and to what degree building walls at certain orientation will cause multiple reflection has been left unsolved [81].

An attempt is made to solve this problem with the help of GTD analysis. A geometrical analysis of typical urban structure is described, and a suitable ray tracing methodology is suggested. Subsequently, SAR processing is simulated using the reflected signals from arbitrary target structures, which are generated by GTD ray analysis. Different types of waveform are employed for various cases and the superior performance of the phase coded waveform developed in this thesis is verified through comparison with conventional waveform results.

8.1 Urban Area Image Generation

With the main emphasis laid on the understanding of urban structure and its radar responses, a dominant reflector that appears in this environment is investigated through simplified modelling.

8.1.1 Geometry Understanding

As a typical example of a modern urbanized area, Figure 8.1 shows a part of central London taken from a bird-eye view. As discussed in Chapter 3, many of the structures appear to have the potential to function as corner reflectors through multiple reflection processes. Here is provided a detailed insight into multiple corner reflection phenomena.
Corner Reflector Effect

The structure of urban buildings does not necessarily constitute a conventional corner reflector since the ground surface is distributed widely over an infinite ground plane and the height of walls is arbitrary. It is assumed that the signal source and receiver are located in the far-field and the wavelength is very small compared to the size of the structures. A ray signal $\Upsilon_n$ ($n = 1, 2, \ldots$) within the received signal is given as

$$
\Upsilon_n = \kappa E_i e^{-jkR_i} \Gamma_{rc1} e^{-jkr_1} \Gamma_{rc2} e^{-jkr_2} \cdots \Gamma_{rcN} e^{-jkR_r} \\
= \kappa E_i e^{-jk(R_i + R_r)} \Gamma_{rcN} \prod_{p=1}^{N-1} \Gamma_{rep} e^{-jkr_p}
$$

(8.1)

where $\kappa$ is a proportional constant including the propagation loss and $E_i$ is the signal strength of incident ray in a position where the signal can be considered as a far-field plane wave. $R_i$ is the distance between the incident or return rays and antenna, $\Gamma_x$ is a reflection coefficient of the wall $x$ and $r_p$ is the travel distance of the ray paths in each reflection interval. $N$ is the number of reflections and the minimum $N$ for a complete signal return would be 3.
8.1 Geometry Understanding

Figure 8.2: Range shift and target ambiguity in corner reflector structure between two symmetrically located points \( P_1(W_1, H_1) \) and \( P_2(W_2, H_2) \).

**Range Shift and Target Confusion Effect**

SAR is a range detecting device and the cross-track reflectivity map of the image is based on the sensor to target distance. When the local terrain deviates from a smooth surface, a geometric distortion occurs due to the high structure of the target. Depending on the slope of the target relative to the view angle, these effects are referred as ‘layover’ or ‘foreshortening’ errors (refer to Chapter 3).

In addition to these typical phenomena, a dihedral reflector structure comprised of urban buildings brings about another cause of range displacement error and confusion for target location measurement. In the course of the multiple reflections between buildings and ground surface, a ray propagates a long distance from when it first enters the structure until it hits a final reflector and bounces back to the receiver. As a result a time delay occurs associated with this extended travel path. This effect will be considerable in macro structures found in many modern urban cities. From the coordinates shown in Figure 8.2, when the locations of the two specular reflection points are designated as \( P_1(W_1, H_1) \) and \( P_2(W_2, H_2) \), the prolonged path length \( L_p \) due to multiple reflection is calculated as

\[
L_p = \sqrt{W_1^2 + W_2^2 + (H_1 + H_2)^2}
\]  
(8.2)

The extra propagation path increases in proportion to the heights of the building walls as well as their displacement. On the other hand, layover distortion makes the roof of buildings appear at a nearer slant range than the base, and the actual range position...
displacement will be

\[ \tilde{L}_n = \sqrt{W_1^2 + W_2^2 + (H_1 + H_2)^2 - H_{1,2} \cos(\theta)} \]  

(8.3)

where \( \theta \) is the antenna look angle. In addition to this, two return signals from \( P_1 \) and \( P_2 \) follow identical ray paths due to their symmetry, and are indistinguishable from each other. They share the same amount of time delay and attenuation factor during reflections. Hence although they are separated in range and azimuth directions, their return signals are ambiguous with each other. When their sizes are not large compared to the resolution cell, they will appear as a single point target. Since their signals will coherently add at the receiver, the combined signal will result in further increase of the target reflectivity. On the other hand, when their displacement is larger than the resolution cell, their new point image will share the same mirror image which is a joint reflection phenomena of \( P_1 \) and \( P_2 \) and given as

\[ P_3 = \sigma_1 \sigma_{p1,p2} \sigma_2 \Gamma_{p1} \Gamma_{p2} \Gamma_G \]  

(8.4)

where \( \sigma_1 \) is the signal propagation loss between reflections and \( \Gamma_x \) is reflection coefficients of walls and ground.

Figure 8.3 clarified the range shift effect through SAR point target simulations. When the image resolution is high enough (\( \delta < \tilde{L}_p \)) to separate the two walls, each point target appears as mirrored and range shifted images having identical responses and displaced in range by the distance between the walls given in Eqn. (8.3). On the other hand, a low resolution image (\( \delta \geq \tilde{L}_p \)), which will be a more frequent case, does not resolve the two structures and treats them as a single point target.

Surface Reflection

A strong response from man-made targets in radar imagery is attributed to two features; their geometrical structure and high dielectric constants. Strong returns from metal bridges and railway tracks are good examples that combines these two factors. The surface roughness is another major factor to contribute to the signal strength. Unless the target surface is perfectly conductive and smooth, the reflected signal is no more localized to one particular direction but instead randomly diffuses away in arbitrary direction. The degree of signal dispersion is dependent on the surface roughness. Even if the dielectric constant is not sufficiently high, a smooth surface tends to obey Snell's law with good efficiency by focussing most of the energy in one particular direction. Contrary to this, the interaction with a rough surface is less predictable and the distribution of the reflected signal follows the Rayleigh criterion [81].

The depression angle of the incoming signal affects the reflection process as well. At
very low depression angle or high incident angle, most of the energy reaching a surface is reflected away, while a strong energy return is seen at high depression angle. This can be compared to the diffraction phenomena of the GTD theory. At a rough surface the reflection field may be treated as a diffracted field originating from the edges of randomly distributed scatterers. As explained in Chapter 3, the diffracted field is most strong at the angle of boundary between field shadows and diminishes as the angle moves far away. This boundary angle is the same as the reflection angle and the difference between incident and reflection angle is merely twice the incident angle. Figure 8.4 is an illustration of this phenomenon, where an arbitrary surface is modeled as a group of random pitches. In the smooth surface, the average height of the random pitch model is very low and accordingly the standard deviation of the height distribution is low as well. Since the height differences among sharp edges are small, the so-called RBS (Randomly Bounced Signal) rarely exists and the reflection is dominated only by the diffraction at the vertex points of each barriers. Therefore, at low depression angle, majority of the incoming signal is reflected away from the direction of the incident field and very little energy is left to contribute to RCS. This theory justifies the quantitative relationship between low depression angle on the smooth surface and dark appearance in the image. The higher the average height dimension and also the bigger the incident angle, the more chance there is that the incoming signal is deterred from direct reflection and bounced back or randomly diffuse out. Since the behaviour is not precisely predictable, the term ‘RBS’ is used within the Figure 8.4.

The ray signals may bounce back to the transmitter contributing to increased RCS or repeat bouncing on other pitches until they escape in arbitrary directions. From the illustration, it can be conjectured that the second RBS is more likely to head toward original source eventually than the first RBS. The diffraction field from first RBS spreads
out in all direction and contributes to the total RCS increase. However, it is doubtful that the tendency of bouncing backward is proportional to the increased average height. Once the height increases above a certain critical level compared to the wavelength, the surface no longer operates as a random surface and the reflection response is governed by a number of dominant pitches which may be seen as independent scatterers.

The impact of the depression angle parameter on radar imagery may be roughly explained in this way. Figure 8.5 shows the relationship between the diffraction power and reflection angle and it becomes obvious that a smaller incident angle results in the higher backscatter energy and makes target appear with bright RCS level. This may even enable us to draw a quantitative measure of the target reflectivity corresponding to specific target scene once their average height is known in advance.
Table 8.1: Roughness of the typical urban features corresponding to various Satellite radar systems

Table 8.1 [37][12] lists the average vertical dimension of some typical target scenes and their appearance in the radar imagery. The roughness is divided as smooth, intermediate and rough surfaces depending on the height and the wavelengths. The pavement and grass areas behave in different manners because of the different interactions between signal and surface. The dark image of the oil surface on the sea can be understood in this way.

In fact, even if the mechanism of the signal interaction with arbitrary surface is verified, the quantitative prediction of energy return is still a very complicate problem since there are other factors such as wetness, conductivity and homogeneity of the surface that may affect the final output. Despite these unpredictable circumstances, this approach may provide a useful measure for relative comparison between different target scenes.

8.1.2 Modeling

Based on the urban geometry in Figure 8.1 where perpendicularly aligned buildings and ground plane form a corner reflector structure, a simplified model is developed for GTD computer simulation. The GTD technique is aimed to accomplish an electromagnetic analysis by tracing all possible interactions of radar signals with the scatterers. In fact it is not the complexity of the GTD mathematical formulation that limits its application but rather the complexity of target structures, which can pose a difficult challenge to the appropriate modeling procedure. Once an appropriate modeling is devised so that a sufficient number of ray paths, which contribute to delivering significant energy, are fully taken into account, the remaining task is merely a matter of calculation burden assigned to the computer.

Here the target object in question no more acts like a conventional corner reflector. Two arbitrary sized walls are standing over infinite plane and for simplicity, both of them as well as ground plane are assumed to have smooth surfaces. The large gap between corner axis and actual building locations merely adds more complexity in finding valid ray paths. The energy delivered by diffraction signals from the building edges decays at the rate inversely proportional to their journey distance. Hence their contribution to overall response is negligible and ignored in this simulation (as validated in Chapter 3). At the
8.1 Modeling

Table 8.2: Possible combination of reflection order between rays and urban structure

moment this scatterer is assumed to stand alone without being interfered by surrounding structures.

At first it looks like that there exist extremely diverse manners in which radar signal can behave within the regions. But as long as the radar signals follow Snell's reflection law, and keeping in mind that a ray has to experience 3 reflections from each walls to return backward, all possible combinations can be fully traced. Initially an incoming ray can hit either one among three surfaces first and bounce backward through the other two surfaces. Therefore, the number of possible combination is given by $3 \times 2 \times 1 = 6$ cases. Table 8.2 lists the 6 different cases.

For each incoming signal ray $\vec{E}_i^n$, the returned signal $\vec{E}_p^n$ will have the form

$$\vec{E}_p^n(x) = -|E_i^n|^2 |\Gamma_{wp}| \Gamma_{wp}(\phi_{wp}, \theta_{wp}) \cdot \exp\{-jk(R_i + d_1 + d_2 + R_r)\}$$

(8.5)

where $\Gamma_{wp}(\phi_{wp}, \theta_{wp})$ is the reflectivity of wall p and a function of the ray incident direction. $d_1$ and $d_2$ are the distances between walls that the ray travels and $R_i$ and $R_r$ are the distances from the antenna to the walls that the incident ray $E_i^n$ hits the first and the last respectively.

Figure 8.6 is the graphical illustration describing the 6 different ray paths. Even after the attempt to determine the individual paths is successfully carried out, there is another difficulty still to overcome. Unlike a conventional corner reflector, not all rays that hit wall 1 firstly bounce backward through wall 2 and ground planes due to asymmetrical geometry orientation. A similar assertion is also applicable to cases 3 - 6. Given the incident angle and the exact location of the first hitting spot, the next reflection behaviour can be calculated and whether they will return back or scatter away can be decided afterwards. However depending on the size of the targets relative to the wavelength, the required number of rays can reach millions, and it will be extremely inefficient to repeat the ray path tests and evaluate reflection directions for all individual ray paths. To tackle this problem, a mirror image modeling is devised. When the reflection is subject to Fermat's principle, a mirror image can be set up in virtual space for convenience and the visibility of the target is determined by the line of sight drawn in the virtual space.

Here is set up a three-dimensional mirror space and a line of sight test is imposed
Figure 8.6: The possible combinations of ray paths during interaction between radar signals and corner reflector formed by two walls and ground plane. A simplified geometry for ground buildings with smooth surface and bare soil ground.

for mirrored wall image. The mirror images of case 1 and case 2 are adjacent to each other and can be considered as one single case. Similarly, the cases 3 and 4 combine to become one single case. Eventually there are four different types of ray paths to consider. Figure 8.7 draws hence generated mirror image spaces for each cases. The shadow region of incident wall is bounded by its line of sight extensions, which is created by a light source from radar. Only those rays that pass through a specific zone which is common to both shadow region and the mirror image of the other wall, experience corner reflections and return back toward radar. This region is indicated by the shaded area in Figure 8.7.

How many rays should be employed to gain a sufficiently reliable result is a matter of the relative size and complexity of the scatterers. Grid lines of size \((N_1, M_1)\) and \((N_2, M_2)\) are drawn within two walls, which represent the density of the rays to be tested. The parameters \(N_i\) and \(M_i\) are determined by the relative sizes of the walls in comparison to wavelength. The larger number of grid points does not necessarily guarantee the better accuracy since the result will asymptotically follow convergence pattern and should be compromised against the calculation time. A choice for the size of grid is made by increasing the employed rays until a converging tendency is detected at the simulation output. To have reflections at both walls 1 and 2 is a necessary and sufficient condition for a ray to bounce back to antenna. In the ray tracing model shown in Figure 8.6, the total number of rays involved in actual GLD calculation is given by

\[
\Psi(\text{rays}) = A_{eff}(\theta, X_{w1}, Y_{w2}, N_1, M_1) \cdot N_1M_1N_2M_2
\] (8.6)
Figure 8.7: Four different cases for tracing the possible ray paths to explain interactions between radar signals and corner reflector formed by two walls and ground plane. Fig. (a), (b), (c) and (d) are mirror geometries devised for the purpose of understanding the complicate reflection mechanism and convenient computer simulation.
where $A_{\text{eff}}$ is the ratio of effective area to overall wall area. It is a function of incident angle, the size and location of walls and always less than 1. When the grid is built with a $100 \times 100$ size matrix, then the number of possible ray paths to be considered can approach the order of 100 million! Since SAR processing requires a collection of data for different view angles, this figure can further grow up to more than billions. Hence an efficient algorithm to reduce overall grid numbers and search for valid rays with good efficiency should be sought. It is found that the mirrored image modelling in Figure 8.6 provides a good efficiency in this sense.

8.2 Radar Target Response

8.2.1 Convergence Test on A Simple Structure

In strict sense, GTD simulation is a numerical approximation technique and hence there is a preliminary condition that the convergence should be guaranteed. Using a dihedral reflector standing on an infinite ground surface, the validity of ray tracing method is examined and the results are shown in Figure 8.8. For convenience, all the walls and ground are given perfect conductivities ($\sigma = \infty$). Hence it is assumed that the reflection coefficients $\Gamma_r$ are -1 for all areas and there is no attenuation during reflections. Compared to the wavelength (3\(^\text{rd}\)m) of typical spaceborne SAR, most of urban building walls are very large and considered as infinite planes. Then the resolution of the ray tracing grid is decided by the relative size of target objects and their distances. For a given width $W_{1,2}$ and height $H_{1,2}$, the grid resolutions for ray tracing are set as $W_{1,2}/u_{1,2}$ and $H_{1,2}/v_{1,2}$ for each dimensions, where $u$ and $v$ indicate grid sizes. A simulation is performed for $W_{1,2} = H_{1,2} = 10m$. The result shown in Figure 8.8(a) confirms that a factor of 20 for both $u_{1,2}$ and $v_{1,2}$ or a resolution 0.5$m$ delivers a good convergence performance. The asymmetrical distribution of the number of valid rays for case I, II is related to the incoming signal intensity that changes with respect to orientation and corrected to produce symmetrical E-fields.

8.2.2 Look Direction

The appearance of urban targets in SAR image is sensitive to the geometrical relationship between a look direction and feature orientation. It was observed that the detected signal level from urban area is highly dependent on the look angle and the same target can appear with different brightness levels. A corner reflector shows a broad RCS pattern over the entire look angle domain. It is a typical outlook of the modern city that the building structures are aligned over planned grids lines. In this environment a more realistic model will be a group of corner reflector structures displaced by finite distances from each other. Both the heights of the buildings and spaces between them are no longer uniform.

\(^1\)The influence of look direction on radar backscatter has been referred to as the ‘Cardinal Effect’ [12].
Figure 8.8: Ray tracing method test on corner reflector structure. RCS responses corresponding with respect to view angles.
Many researchers have emphasized the need for a quantitative relationship between these parameters and brightness levels in the radar image [30].

Figure 8.9 and 8.10 show simulated results on two arbitrary building structures performed by changing their distances. Signal levels are measured by the ratio of return signal energy to that of incoming signal. The larger the separation between them, the less chance of rays to bounce back and hence the return signal level weakens. As the separation approaches half of the width of the buildings (0.5W), there is found a dramatic reduction within the return signal level. Figure 8.9(b) shows that asymmetric structures generate beam patterns derived from the results of symmetric cases shown in Figure 8.9(a).

In spaceborne radar application where the incident angle is very steep, the heights of buildings appears relatively low compared to their widths, which results in reduction of number of possible ray paths. The effect of the height of buildings on RCS level is investigated in Figure 8.10. When the separation between the walls is large, a longer propagation path is required for a ray to complete the corner reflection, of which the chance is increased by higher building. In case the separation is small (here is given as 0.1W), the signal energy gain due to increased height is relatively moderate while a considerable improvement is achieved for large separation case. There is found a 3.5 dB improvement for small separation case which is comparable to $20 \log(4/3) = 2.5$ dB. On the contrary, a mere increase by 1.5 times in height brings about 10 dB $>10 \log(1.5^2) = 3.5$ dB improvement for distance $x_1=y_1=0.5W$ case. When building structures are displaced by considerable distances, their reflectivity becomes more sensitive to their height, which may reflect their average height distribution.
8.2 Incident Angle

The incident angle is defined as the angle between the radar beam and the perpendicular lines to the surface ground of incidence. A larger incident angle is desired when the target is covered with relatively flat surface to enhance range resolution. But on the other hand, when high structures are present in an urban area, the radar shadows become longer with the increase of incident angle and conceal other land cover region. Several empirical data on urban scenes have led to the conclusion that there is no absolute ideal incident angle that is universally suitable to any type of region. Henderson [32] claimed that the minimum required incident angle for proper observation is at least 20-23°. Other studies [37] also found that the best performances were seen around 40-45° incidences. However, there has been no theoretical approach to quantify and understand their mutual relationship.

Here simulations are performed to measure quantitative relationship between incident angles and various building structures. Figure 8.11(a) shows that there is a linear increase of return energy levels with respect to building heights. The distance D is set to be equal to W of the buildings, which is a more realistic description of typical urban feature. A particularly important property to notice is the cut off angle, which is defined as the minimum incident angle required for the corner reflection to happen. Unlike the case where buildings are closely located, a considerable gap between the structures sets a critical angle for reflection. Although it is very difficult to consider a generalized model for urban structure, the model used for this simulation generates critical look angle at [25-30°] range. It is found that this range interval remains the same regardless the incident angle once the heights of the buildings are relatively larger than the width. The similar condition
8.3 Target Response Dynamic Range

The contrast and dynamic range are used as a means to measure the image quality. Contrast is defined as the ratio of the intensity of the strongest target to the average background [76]. If a wide dynamic range of input is employed to enhance the contrast in urban scenes, their surrounding natural features, which tend to be more homogeneous, are suppressed and incur an loss of information. Therefore it is not always preferable to have a wide dynamic range response in terms of target detection and recognition. However if the detection is mainly concerned about an urban area which is densely occupied by strong scatterers having wide response spectrums, there is a clear advantage of using waveform with a good dynamic response. It is reported that 20 dB dynamic range of brightness level provides enough contrast to distinguish surfaces in rural area but causes a confusion in metropolitan area composed of a complex mix of build environment and natural surfaces [31]. It is observed that the image contrast level above 60 dB range is required to avoid confusion in target recognition and classification. The observed brightness level have been related mostly with the average heights and density of the buildings.
8.3 Effects on response power level

The complex nature of electromagnetic interaction within urban area discourages to draw an absolute quantitative measure of dynamic ranges corresponding to arbitrary target models. However there is little doubt that the target reflectivity varies in proportion to the size and orientation of the target structure.

With the assumption that the dominant source for bright return from urban buildings are specular returns, a rough measure to determine necessary dynamic range to distinguish different features is approximately attained. Table 8.3 lists SAR response dynamic ranges originating from various target specifications. The energy level is calculated as a relative ratio of the responses. The target size is assumed to be sufficiently larger than the wavelength of the radar signal which is assumed to be in L-C band. The maximum response is obtained as -10.3 dB at (D, H, \( \phi \)) = (0, 5W, 45°) while the minimum is -62.8 dB at (0.5W, 2W, 20°). The group of height H below 2W corresponds to residential building, while H=5W represents tall buildings in a modern city. From these results it can be claimed that for a mixture of wide range of urban buildings to be distinguished with each other, at least 50 dB of input dynamic range is required for a sufficient information extraction. Considering that this figure is based on only the corner reflector scatterers with strong specular returns, the actual image contrast in the image involving diffuse areas will demand much strict condition of wider range. As Xia [81] has reported, if 20 dB range is accepted as the range of reflectivities in the natural area, up to 70 dB of input range becomes the safe threshold value to guarantee an effective coverage of the complicated urban structure.

8.4 SAR Simulation

The point spread function integrates the response from the observed pixel and its surrounding pixels. It is known that the response from the observed pixel accounts for approximately 50% of the total response, with the balance coming from its neighbours [30]. Unlike homogeneous area, if the surrounding cover is dissimilar, which can occur quite frequently in urban regions, the point-spread function can significantly affect the response from a single cover class. As a means to assess the performance of waveforms,
several SAR images are generated mainly focussing on the point target response (PTR) of SAR processing.

8.4 Waveform Performance

Based on the figures regarding wide dynamic ranges within urban SAR images, the performance of various waveforms are compared. The 70 dB dynamic range is a figure applicable for a extremely heterogeneous inner-city region. In an ordinary residential area, the average heights \( H \) of the buildings are expected not to exceed 3W. Table 8.3 shows that the dynamic range for this case is around 47 dB. This figure has come from when only specular returns among strong corner reflectors are considered. If the general reflectivity range of 20 dB in rural area is added, then the required level will reach up to 67 dB. Hence, instead of using conventional definition of range resolution based on 3 dB sidelobe level, a highly strict condition would have to be introduced according to the dynamic range. The Hamming weighting, which is generally acceptable in most situations for its moderate loss of resolution and SNR, provides an insufficient level of -42 dB. This may be enough for most SAR applications concerning natural environments but when a detailed analysis and classification is required on complicated urban scenes, the insufficient target isolation level makes Hamming weighting incompetent. The Blackman weighting scheme shows the most prominent performance in terms of peak sidelobe suppression. But the excessively broadened mainlobe width makes its use impractical. For a given bandwidth of \( \nabla f \), -40 dB sidelobe level is obtained after \( 5\nabla f \) of mainlobe broadening and \( 5.7\nabla f \) mainlobe width should be sacrificed to achieve -60 dB isolation. Apart from that, this level of sidelobe level is very difficult to obtain in practice unless an extreme degree of precision is guaranteed for phase control [4]. Although the Dolph-Chebyshev filter is not realizable in practice, its optimal performance serves as a useful reference for comparison. It has been shown that -60 dB sidelobe level accompanies \( 4.8\nabla f \) of mainlobe width and -70 dB requires \( 5.7\nabla f \) width, which is identical to the 3-term Blackman weighting case. Contrary to the Dolph-Chebyshev case where the mainlobe width is a function of both bandwidth and given sidelobe level, the P code combined with the Woo filter generates a mainlobe width as a function of bandwidth only. Its performance is not affected by the given sidelobe level, which is decided solely by the time-bandwidth product. For a given range resolution, the new pulse compression scheme enables us to achieve arbitrary sidelobe level without losing range resolution as long as no constraint is applied to signal pulse width. Figure 8.12 compares the target resolving capabilities of Hamming, Blackman, Dolph-Chebyshev and Woo weighting filters. The two point targets are assumed to be separated in time domain by \( \sqrt{2} \) and \( \sqrt{3} \) respectively. Except in the Woo filter case, all the other weighting methods fail to distinguish two different targets by 60 dB isolation. Figure 8.12(d) shows

\[ ^{2}\text{Here the optimal pattern is defined by the criterion of minimum mainlobe width for a given sidelobe level.} \]
8.4 Azimuth Compression

Figure 8.12: The comparison of target resolving performance of various waveforms. Two point targets with identical reflectivity of 60 dB are assumed to be separated by $\frac{4x}{c} = 2.3/\sqrt{f}$.

that 60 dB isolation is easily achieved by employing 25 MHz frequency bandwidth for 40μs pulse duration signal ($\cdot BT=1000$)\(^1\).

8.4.2 Azimuth Compression

One of the most critical dilemmas in developing a waveform for SAR application is that only range direction processing can receive, if any, the benefit of a good waveform design while azimuth direction property is left unaffected. The belief that image property in azimuth direction is beyond the control of waveform being used explains the lack of work related with the waveform design for SAR.

The new phase coding scheme introduced in this thesis is a result of the unique ma-

\(^1\)Here the bandlimiting effect is not considered. In practice, a slight increase of mainlobe beamwidth is expected but the sidelobe response is actually improved due to the weighting effects [20][42].
8.4 Azimuth Compression

<table>
<thead>
<tr>
<th>Property</th>
<th>1680 (ERS-1)</th>
<th>1240 (Nyquist)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak Response (Ambiguity Main Peak)</td>
<td>-26.4 dB</td>
<td>-19.8 dB</td>
</tr>
<tr>
<td>Ambiguity Energy (AASR)</td>
<td>-20.1 dB</td>
<td>-12 dB</td>
</tr>
</tbody>
</table>

Table 8.4: Azimuth ambiguity properties comparison

Manipulation of linear FM utilizing a specific sampling scheme and an appropriate receiver filter. Since the SAR signal in azimuth direction can be approximated as a weighted linear FM waveform with the sampling rate decided by PRF, which is in the control of user, the similar improvement seen in range processing may be achieved in azimuth direction as well. The unambiguous condition in azimuth resolution sets the minimum PRF $f_p$ as

$$f_p \geq \frac{2V_s \theta}{\lambda}$$

where $V_s$ is the velocity of the satellite and $\theta$ is antenna beamwidth. In practice, a higher PRF is preferable in order to suppress the unwanted integrated azimuth ambiguity energy. In ERS-1, the minimum boundary $f_p|_{\text{min}}$ is 1240 Hz while the actual PRF is selected as 1680 Hz = 1.35$f_p|_{\text{min}}$. The use of high PRF is rather intended to cut off the intrusion of ambiguous energy than preventing aliasing effect. Hence if it is guaranteed that Nyquist PRF rate can be used without causing severe disruption by ambiguity energy, the sampled signal can be regarded as a P code signal such that Woo filter suppression can be applied in the azimuth direction as well.

To investigate the feasibility, the amount of ambiguity energy corresponding to various PRF values are calculated using Eqn. (2.21). Naturally, compared with the actual case where the sampling rate is high enough to reject alias signals, the Nyquist PRF rate produces increased ambiguity energy. Table 8.4 is a quantitative comparison between the two cases. The Nyquist PRF rate deteriorates the ambiguity energy rejection performance by 7 dB in peak response and 8 dB in ambiguity energy ratio respectively. However, it should be reminded that these values are calculated with the assumption of uniform reflectivity. When the urban area is the main concern, its relatively strong reflectivity level is protected from contamination by azimuth ambiguity energy coming from adjacent area. Eqn. (2.23) implies that the azimuth ambiguity response originates from a point displaced by approximately 10 km. Considering the extremely widened dynamic range, 7-8 dB of ambiguity energy loss seems to be acceptable. Since the PRF rate is reduced, the ambiguity level in range direction becomes less severe.

Figure 8.13 illustrates PTR when Woo filtering is implemented in both directions. Not only in range direction but also in azimuth direction is brought a sidelobe elimination effect. The point target is isolated from the surrounding by a few hundred metres in both range and azimuth directions, which is sufficiently larger than most man-made structures.
The deep hollow near -100 dB energy level is well below the dynamic range of very strong specular return signals and promises an excellent insulation of the sidelobe energy.

It may contradict conventional information theory that the better performance is achieved with lower PRF rate and hence less available information. In fact there is clearly a loss of information, which appears as a resolution degradation by half. The Woo amplitude weighting may be applied to avoid this sacrifice but still the increased sampling interval results in a broadened mainlobe. To highlight the excellent performance obtainable with this PTR, SAR simulations are performed using conventional linear FM with Hamming weighting and the Woo filter respectively and compared in Figure 8.14. Four strong reflectors are assumed to align along the azimuth direction. Although their 3 dB resolutions are set as the same, the broadened mainlobe width in the Hamming weighting image produces enlarged target responses. The most prominent improvement by the Woo filter is found in the reduction of unwanted sidelobes in both range and azimuth directions. On the assumption that the dynamic range of the image is wide enough to provide a good discrimination among various reflectors, the scatterers reflection level is set as 60 dB above the background level. In this degree of contrast the existence of sidelobe cannot be ignored. As shown in Figure 8.14(a), the Hamming weighting fails to suppress the strong interference of sidelobe signatures. Contrary to this the application of the Woo filter scheme in both directions fully removes the sidelobe effect. As the distance between the scatterers becomes narrow the Hamming weighting image loses resolving capability rapidly while the Woo weighting image still maintains the distinction.
8.4 Sidelobe Energy Reduction

Urban SAR images consist of a group of target responses having broad dynamic ranges. Even when the PTR itself satisfies the criterion for good point target detection, the accumulation of energy coming from various other sources may elevate the noise level and hinder target detection at certain points. Especially for the case when a very bright area borders with dark areas having relatively low reflectivity, the integrated sidelobe energy from bright sources may deteriorate those low level signals resulting in poor image quality. A subtle border line that distinguish two separate areas may be lost after losing weak contrast and the amount of information contained in image texture can be diminished. Since the texture information plays a significant role in target classification, the loss of contrast and target information will certainly hamper the proper understanding of processed SAR image. Regardless of the peak sidelobe level, the conservation of energy rule will not tolerate reduction of accumulated sidelobe energy in any specific part of the image.

The sidelobe reduction scheme introduced here does not attempt to overturn the energy conservation rule, but it seeks to emulate the CLEAN concept which performs sidelobe energy reduction task at post-processing stage. To illustrate the dramatic improvement brought by this technique, 2-D SAR simulations are performed. At first a simple point target SAR response is simulated in Figure 8.15 and it is seen that most of the energy residing in sidelobes directions are eliminated except in the sidelobe tails.

To investigate signals from distributed targets, a small target region is set up having a random distribution of high reflectivities and the sidelobe energy leakage into nearby area are investigated for various cases. Figure 8.16(a) is generated by a plain linear FM signal with the Hamming weighting imposed in both range and azimuth direction processings. Figures 8.16(b) and (c) are obtained using s-P codes and the Woo filter in the range.

Figure 8.14: SAR image simulation of four point reflectors.

8.4.3 Sidelobe Energy Reduction

Figure 8.15: SAR image simulation of a point target.
Figure 8.15: The point target response of SAR signal processed with Woo filtering and sidelobe reduction technique in both range and azimuth directions. $T_{az}$ and $T_{ra}$ are time durations of processed SAR signals in each directions.

Figure 8.16: 2-D SAR simulations for various waveforms and processing methods. A very strong feature is assumed to be present on the dark background.
8.4 Sidelobe Energy Reduction

direction only. A distributed target occupying $10\delta_{ra} \times 10\delta_{az}$ area is assumed to have a very high reflectivity over dark background. In all cases the background noise is assumed to be sufficiently low so as to highlight the dramatic effect of sidelobe reduction. In Figure 8.16(b) the code length is set to fit the ERS specification and the uniform sidelobe level is given as $20\log_{10}(573) = 55$ dB. Due to the presence of bright targets, the accumulated sidelobe level increases by 5 dB. Figure 8.16(c) is for the similar case but the code length is increased to 2000 and, as a result, the range sidelobe level is reduced further below -60 dB. To clarify the mainlobe protection the range dimension is reduced to $0.1T$. An isolation of mainlobe by a sharp gap is eminent.

One of the disadvantages of using conventional phased code signals is that their sidelobes are not attenuated away from the peak response, which presents the risk of having high accumulated sidelobe energy. Although the sidelobe level is already suppressed below conventional linear FM case level, the new energy reduction method further attempts to remove a significant portion of the sidelobe energy. Figure 8.16(d) is obtained when energy reduction is applied only to range direction processing. If the distribution of target reflectivities is confined to a small dynamic range, this degree of energy reduction would not produce considerable improvement. But when several extremely bright scenes are mixed within dark natural environment, this technique can generate a great enhancement. The buildings, as demonstrated earlier, tend to be the brightest spots due to the strong specular returns while their surrounding features such as pavements, playground and grass fields show very low reflectivities. This kind of mixture of extremely heterogeneous textures can be properly interpreted without the image corruption only when a good contrast between targets and low sidelobe low energy accumulation is provided. Here the accumulated sidelobe energy is reduced as low as -100 dB level.

As a means to quantify the efficiency of energy condensation, ISL distribution around mainlobes are calculated for different waveforms and shown in Table 8.5. The energy condensation is measured by the relative amount of the energy within given ranges from the peak point compared to the total energy. When the Hamming weighting is applied to both range and azimuth direction processings, the actual energy residing within 3 dB radius $\delta_r$ is only 60% of total energy, which is attributed to the broadened mainlobe pattern. It is only after the range is extended to $3\delta_r$ that most of the pulse compressed output energy is recovered. Contrary to this, when Woo filtering is applied, the peak response holds most of the energy within the designated 3 dB distance. The amount of partial ISL energy does not change with the increase of the range distance since the peak response is surrounded by a strong isolating area. The energy condensation rate is proportional the signal code length, which again demonstrates the flexible use of phase code. Even when only range processing receives the Woo filter processing, due to the practical difficulty of PRF control, still a superior performance is obtained over the Hamming filter only case.
8.4 Target Response

The energy condensation rate is meaningful for indicating the target contrast level in the image or maximum utilization of the waveform.

8.4.4 Target Response

The highly isolated pattern of the pulse compression output may find application in target detection and recognition. A target scene may be regarded as consisting of background clutter and a small number of significant objects, which are generally man-made, such as buildings, bridges or vehicles [61]. In practice, there are three types of scatterers that play a dominant role in signal reflection, which are,

1. smooth surfaces
2. Discontinuities
3. Cavity-like or corner reflectors

A smooth surface return is a plain specular return, which is very strong but rarely observed due to its narrow focus. Discontinuities generate diffraction fields introduced in GTD theory and it was shown earlier that their backscattering energy is so weak that it can be ignored at the presence of other specular returns. Rihaczek [67] also observed that when one examines the intensity image of a ground vehicle, every response discernible by eye will have been generated by a cavity-like reflector.

Target Contrast

With the specification for ERS-1, the s-P code can produce a dynamic energy level in the range up to 55 dB (\( \cdot 20 \log_{10}[576] = 55 \)). However the linear FM with the Hamming weighting provides a fixed PSL of 46 dB. Hence when a wide range of contrast level is required due to the dynamic nature of scatterers in the target area, the flexible phase code is favoured. Figures 8.17 compares the SAR images generated by the s-P code and linear FM waveform when three reflectors having wide response levels are seen in the target area.

Although their resolutions are set to be the same, due to the broad mainlobe response of linear FM, the target response appears larger in the linear FM case than the actual
8.4 Target Response

Figure 8.17: SAR images generated by the Hamming weighted linear FM and s-P code with Woo filtering. The targets have the reflectivities of 0, -20 and -53 dB. The code length of s-P code is given as 576.

target size. This indicates that when the target dynamic range is wide, the practical resolution size is larger than defined by the 3 dB rate. This phenomenon does not happen in s-P code case where the mainlobe pattern is sharp and protected from the sidelobe by a strong isolation. Although the three different scatterers are assumed to have equal sizes, they appear in different dimensions in the linear FM image due to the nature of the mainlobe response while the original size is maintained in the s-P code. When -53 dB of weak signal is present linear FM loses the detectability since it is out of dynamic range while s-P code successfully recovers. Even -20 dB of linear FM causes a confusion within the SAR image when it is situated nearby the sidelobe generated by a strong reflector. It may be misunderstood as a part of the sidelobe pattern depending on its position. On the other hand the strong isolation of the mainlobe response and low uniform sidelobe in the s-P code prevents this error. When the goal of SAR imaging is not only confined to the target detection purpose but also required to consider recognition and statistical analysis of targets within urban area, a wide dynamic range of signal response should be favoured.

Target Detection and Recognition

The scenario here assumes a situation where very strong reflectors that generate sidelobe levels comparable to the underlying noise level, are distributed densely at close distances to each other. The reduction of sidelobe energy is desired in order to avoid a confusion between background noise and sidelobe signatures. It also helps to enhance ATR (Automatic Target Recognition) performance [61] and increases the chance of correct
target detection. All of these will eventually result in an accurate target discrimination within the dense target scene. The flow chart for a typical ATR procedure of SAR image is shown Figure 8.18. The contribution of sidelobe reduction can be found in each stages depending on the degree of background noise but the procedure between target discrimination and classification is most likely to receive the benefit. Target classification is based on pattern-matching or recognition. The pattern-recognition is performed at the edges of potential target objects and therefore the sidelobe removal at adjacent area of strong scatterers provides a great advantage for precise discrimination and classification by suppressing the pattern distortion.

Figure 8.19 shows simulated SAR images of a tank-like vehicle. The signature of the man-made structure is dominated by a number of strong point responses. Figure (a) specifies those dominant point signatures of the target structure\(^4\). Figures (b),(c) and (d) compares SAR images generated using different waveforms. The advantage of phase code accompanied by a sidelobe cancelling scheme is clearly demonstrated. The enhancement is not dramatic when the Woo filtering is applied only to range direction. It is after a full azimuth direction processing is accomplished using sidelobe canceller that the original point signatures become visible.

'I know I've got a degree. Why does that mean I have to spend my life with intellectuals? I've got a life-saving certificate but I don't spend my evening diving for a rubber brick with my pyjamas on' - Victoria Wood -

\(^4\)The point signatures are obtained from [61]
Figure 8.19: Simulated SAR images of a tanklike object using different waveform schemes.
Chapter 9

Conclusion

9.1 Summary of Thesis

Waveform design for SAR  Synthetic Aperture Radar (SAR), which was first developed in the 1950's as a technique for improving the resolution of military reconnaissance radar, has rapidly matured as a remote sensing tool for wide range of civilian applications. While accurate and reliable processing techniques or motion compensation problems were once popular research areas, many current works now deal with SAR applications, which involve acquisitions of diverse data format and post-processing for better understandings of the SAR image. Although the signal waveform is the very medium that delivers target information, very little effort has been made yet on the aspect of waveform design and the subsequent effect on the generated image. This thesis has attempted to relate the waveform design with SAR operation and to introduce a new type of waveform that may contribute to enhanced SAR image quality for complex urban application. An ideal pulse compression performance for SAR application has been chosen as having uniform level of sidelobe while the mainlobe is well separated from surrounding areas.

Pulse compression  Pulse compression is inevitable in many radar systems where the peak transmit power is limited. The peak sidelobe level is a common measure for target detection while the integrated sidelobe energy plays an important role in the image quality over a distributed target area. It is conceived that to construct waveforms having a arbitrarily prescribed ambiguity function is not a realizable one. The energy conservation law implies that a compromise should be made between the range resolution and sidelobe suppression level. This thesis shows that sidelobe energy can be reduced without altering the mainlobe width by introducing sidelobe canceller. The sidelobe canceller is constructed directly from the incoming signal, and hence the performance is reliable and also the difficulty of time-synchronization is relaxed.
Why have phase codes been chosen as starting point of waveform design?
While linear FM has established itself as one of the most popular waveforms, the phase code class also has been a topic of active research as an alternative choice for particular radar applications. The research performed in this thesis starts from the phase code signals in the belief that the discrete nature of finite code length signal would make it easier to manipulate the sidelobe of a finite length and may open a way to reduce the unwanted sidelobe energy instead of merely suppressing the peak sidelobe level.

What criteria have been considered? It is a common technique to employ weighting schemes to suppress the peak range sidelobe at the linear FM pulse compression output. In theory the ideal sidelobe suppression is described by Dolph-Chevyshev weighting where a minimum mainlobe width is achieved for a given sidelobe level. The optimum sidelobe is given as a uniform pattern over the entire range delay domain except the mainlobe and used as a reference for arbitrary pulse compression schemes. The Barker Codes produce autocorrelation outputs equivalent to the Dolph-Chevyshev weighting cases. Although the maximum realizable code length is limited to 13, the uniform sidelobe of the Barker code is regarded as an ideal property and used as a reference for the phase code compression designs. A novel polyphase code waveform design technique has been developed that achieves a uniform sidelobe of Barker code level, but is unlimited by the code length and robust to the frequency Doppler shift effect.

P code analysis The P code group is not a new phase code class. Unlike the Frank or Huffman codes which are generated through particular algorithms, the phase distribution of the P code merely follows that of the linear FM signal and can be considered as a discrete form of linear FM sampled at a particular rate. It was found that by setting the sampling rate $S_r$ equivalent to the pulse compression ratio $B_T$, a unique phase distribution property is obtained at the correlation output. Normally, linear FM signal has a fixed PSL (Peak Sidelobe Level) of -13.2 dB when a sufficient sampling rate is imposed on the received signal. But for the P3 or P4 code of length $N$, the PSL is a function of $N\, \text{and given as}$

$$20 \log_{10} \left[ \frac{\sqrt{2} \cos\left(\frac{1}{\pi} \right)}{\sqrt{\pi^2 - 4}} \cdot \sqrt{\frac{1}{N}} \right]$$

at time delays $\pm i$ which are the nearest integers that satisfy the equality $i + \alpha = \pm \sqrt{N \left(\frac{1}{2} - \frac{1}{\pi^2}\right)}$, $-0.5 \leq \alpha < 0.5$. A typical PSL is below -30 dB for $N = 200$ and can be further lowered by increasing the pulse compression ratio $B_T$. This property is attributed to the unique sampling procedure during which sequences generating higher sidelobe parts are ignored and skipped.
The relationship between the sampling rate and the sidelobe pattern generated by the discrete linear FM signal, has been further investigated. When $1 \ll BT < S_r < 2BT$, the sidelobe of ACF of $S$, which is a discrete linear FM sampled from $S(t) = \exp(-j \pi \frac{B}{T} t^2)$, $0 \leq t < \frac{T}{2}$ at a rate of $S_r$, is bounded in magnitude by $1/t$ and the inequality PSLR $<-13.2$ dB holds. Subsequently it is proved that in the interval $1 \ll BT < S_r < 2BT$, the sampling rate that generates minimal sidelobe pattern is given as $BT$. Thus the low sidelobe profile of the P code group has been analytically explained.

An analytical approach has been taken to explain the range-sidelobe reduction scheme for P codes implemented at the output of the matched filter. It does not actually achieve the Barker code sidelobe level nor is its performance optimized, since the sidelobe pattern is not uniform and the peak mainlobe is broadened by a factor of two. The postprocessing scheme, required to achieve this sidelobe reduction effect, is not easily realizable in complicated radar systems and may not suit the SAR data sets.

**New type of pulse compression filter construction** Based on the calculated sidelobe pattern of the s-P code which is a symmetrical form of the P code, a novel pulse compression filter has been developed and denoted as the ‘Woo filter’. The Woo filter rearranges incoming signals of any time delay such that the sidelobe of compression output is kept at a unit constant level and the mainlobe peak is preserved with a slight SNR loss of 3 dB regardless of the code length. The critical drawback of range resolution loss, which may be overcome by increasing bandwidth, is compensated by introducing an asymmetric form of Woo filter weighting.

**Optimum uniform sidelobe generation** By spreading the sidelobe signal power uniformly over the entire sidelobe domains, the peak sidelobe power is ideally minimized. The uniform sidelobe response means that the unwanted clutter peak has been optimized to reach the theoretical minimum for a given signal which is specified by the bandwidth and pulse length. It may be compared with the Dolph-Chebyshev weighting, which is ideal but unrealizable. The Dolph-Chebyshev weighting enables us to achieve a pre-determined uniform sidelobe but with a severe sacrifice on the range resolution proportional to the PSL. The Woo filter achieves the similar effect but without mainlobe broadening, and it is realizable. The property of having a known level of sensitivity can be utilized for calibration purpose of surveillance system. High sensitivity to the environment helps to improve the capability to detect weak signals from small scatterers.

Although the asymmetric Woo filter successfully compensates the resolution loss preserving the uniform sidelobe level, the wider base of the mainlobe and the absence of mainlobe isolation makes it less desirable than the standard Woo filter case. Although it seems to be very difficult to remove the ripple peaks appearing at both ends of the
sidelobes, their contribution to the overall performance may be ignored as the code length increases. In fact it is the existence of these ripples that enables the Woo filter to bring particular protection against very close nearby sidelobe interferers by isolating the main-lobe. It is particularly desirable to have low sidelobes or sidelobe free regions close to the main peak in tracking applications or in target recognition system [5]. The PSL is solely a function of pulse code length and hence can be set as desired by adjusting the signal bandwidth or time duration. Speckles originated from coherent and non-uniform sidelobes tend to break up the images of large size targets into smaller target images. The harmonic pattern of conventional linear FM sidelobe has more chance to trigger this phenomenon. The uniform sidelobe may help to suppress this occurrence.

**Advantages of s-P code + Woo filter** Some of the advantages of the new waveform design and pulse compression technique are

- Unlimited pulse compression ratio or code length provides the maximum flexibility to the system designer.
- Close relationship with linear FM \(\rightarrow\) Good resistance to the Doppler shift effect.
- Optimized sidelobe pattern without losing range resolution.
- Sidelobe free regions \(\rightarrow\) Tracking applications and target recognition.
- Simplicity \(\rightarrow\) No need for weighting or postprocessing.
9.1 Summary of Thesis

The fact that an improved pulse compression process is realized using a smaller number of available data sets, may contradict to the conventional information theory. However it should be reminded that no practical gain in terms of information is achieved since the range resolution is still limited by the reciprocal of the signal bandwidth. What the Woo filter performs is an efficient rearrangement of the pulse compression output with a slight loss of the peak response level such that unwanted clutter noise is evenly distributed.

Integrated sidelobe energy. The term ‘ISL energy reduction’ is used instead of ‘sidelobe suppression’ to indicate that the energy residing within the sidelobe is actually reduced. When it is desired to coherently cancel the sidelobe, what becomes important is the precise control of time-synchronization. A phase code signal compatible with the sophisticated digital signal processing may satisfy this demand. The uniform sidelobe property makes it convenient to generate the sidelobe canceller directly from the incoming signal. The constructed Woo sidelobe canceller eliminates a significant portion of the sidelobe and subsequently reduces the ISL. For a code length \( N = 1000 \), more than 30 dB of ISL energy reduction is achieved. The efficiency of the sidelobe reduction is proportional to the code length.

Urban SAR environment analysis. In an urban area which is densely occupied by strong scatterers having wide target response levels, there is a clear advantage of using waveform with a good dynamic response. It is observed that the image contrast level above 60 dB range is required to avoid confusion in target recognition and classification.
9.2 Suggestions for Future Work

The observed brightness level have been related mostly with the average heights and positions of the buildings, which combine together to form corner reflectors. The fact that radar return forming very bright images are very sensitive to the polarization type also supports the important role of multiple reflection involving walls and grounds. The GTD technique provides a tool to estimate the return signal level. Contrary to some previous assertions concerning SAR electromagnetic phenomenon, it is found that the contribution of diffraction field is insignificant and may be ignored in the presence of strong specular returns. Up to 70 dB of input range becomes the safe threshold value to guarantee an effective coverage of the complicate urban structure. This level of uniform sidelobe level is realized with the s-P code and Woo filtering of code length 2000.

**Waveform design and azimuth processing** The Woo filter design concept is based on the linear FM signal. As conventional linear FM weightings are implemented in the azimuth processing to reduce the sidelobe level, on the condition of precise PRF control and compensation of antenna beam pattern, the Woo filter concept can be applied to the azimuth processing as well so as to achieve a similar effect seen in the range processing.

9.2 Suggestions for Future Work

**Sensitivity of the uniform pattern to outside noise** The property of a uniform sidelobe pattern forms the basis for the ISL reduction. This property is well preserved in the presence of the Doppler shift effect when the spaceborne SAR is concerned. However, since the success of the sidelobe cancellation relies on the precise control of the time-synchronization between two independent filters, some factors that may affect against this operation should be further investigated. The imperfect signal sampling procedure including finite quantization effect is the most likely candidate to trigger an adverse impact. Although this technique is expected to provide a valuable contribution to the target detection and data communication fields, it is still not clear to what degree of precision is required for a complicated SAR image generation. The system noise sensitivity may provoke a severe disruption of this technique when the incoming signal level is very weak. In this thesis the application of this technique is aimed for the target areas containing strong scatterers and the extreme system sensitivity to outer noise has not been fully considered. A realistic model to take into account these undesired noise sources may be necessary to guarantee a reliable ISL energy reduction.

**Azimuth direction processing** In theory the azimuth direction SAR signal is approximated to the linear FM but with the amplitude weighted by the antenna azimuth
beam pattern. Provided that the weighting effect is corrected and the sampling rate is chosen such that the P code waveform pattern is emulated, the same sidelobe reduction technique may be applicable to azimuth direction processing as well. Various simulations in which some imperfect sampling effects are taken into account show that this technique is robust to modest deviation. However it is not easy to tell how accurately the simulation predicts the actual situation. Unlike the spaceborne SAR where the orbit path is relatively stable, the airborne SAR involves some critical motion compensation and the complexity is further added.

It is not likely to obtain a signal pattern sampled in azimuth direction that exactly matches the P code. Then it becomes necessary to develop a phase compensation scheme which modifies the azimuth signal pattern so that the signal phase deviation lies within the acceptable range. As is done for the airborne SAR motion compensation, a similar strategy may be considered which includes an antenna weighting compensation and signal re-sampling. If the PRF is higher than desired for Woo filtering, we should be able to interpolate linearly between the two nearest actual pulses (or possibly parabolically between the three nearest ones) to infer what the signal would have been from a pulse at the desired spacing.

Also a modification may be directed to the Woo filter design according to the arbitrary PRF rate. Although individual pulse phases are ambiguous, there should normally be only one set of interpretation that will fit the total parabolic phase excursion for SAR returns from a stationary target. Hence, following such ambiguity resolution, it should be possible to interpolate the data that would have been obtained with perfect Woo sampling. Of course the interpolation, and also the curve fitting, will introduce some noise-like errors, and so will degrade the process to some extent.

Ideally, after the azimuth signal correction is fully implemented, the SAR image will consist of dominant mainlobe responses only. Figure 9.3 illustrates the expected image signal level after the Woo filter processings are executed in both range and azimuth directions. The target area is identical to Figure 8.16 case. Only the mainlobe responses of strong target area are visible in the image.

**Crosscorrelation canceller for wide swath implementation** The sidelobe cancellation in this thesis has been performed and verified only for the autocorrelation cases. In fact a similar attempt has been made for the cancellation of crosscorrelation sidelobe pattern but without success. Unlike ACF case it is not straightforward to relate the sidelobe patterns of CCF with the original signal codes, and the derivation of sidelobe canceller is discouraged. However there is a good motivation that may promise a way to successfully implement the similar sidelobe cancellation effect. Figures 9.4 shows the
amplitude and phase pattern of the CCF between two linear FM derived phase codes. The fact that the amplitude pattern is stable at a constant level and the phase characteristic shares a parabolic pattern with the original signal, suggests a possibility that a sidelobe cancellation filter may be implemented for the CCF case as well. Further research would be required on how to relate the received signals to the generated CCF output.

Group urban environment statistics When several buildings are assumed to align along particular street patterns, their behaviour can not be explained by a group of independent scatterers. Several observation results in the past have confirmed that the combination of buildings, walls and ground surface can act as dihedral reflectors when the orientation of the reflector feature are within 10-20° of the perpendicular to the radar look direction [33][12][9]. Where either residential or commercial area comprising low story (2-3) buildings are dominant features, the close distances between buildings and the low average height diminish the corner reflection effect. If the walls of the buildings confronting toward radar direction are smooth mirror reflectors, specular flashes are so concentrated in particular angle directions that they are rarely observed and the critical angle around 20° can not be explained. It is conceived that there is no general rule to produce a satisfactory and fruitful conclusion when such a diverse environment as the urban scene is involved.

Instead of separate analysis for each point scatterer, in which case corner reflector is a dominant factor, a realistic modelling of closely located buildings would have to be set up and examined. In this case statistical modellings such as the height distribution and the target density should be considered.

‘Science becomes dangerous only when it imagines that it has reached its goal’ -George Bernard Shaw-
Figure 9.4: The magnitude and phase patterns of the cross-correlation function between linear FM derived phase code signals having opposite slopes.
Bibliography


[75] J. M. Berenyi Tajbakhsh, M-J. Kim and R. E. Burge. Images of urban areas by a synthetic aperture radar simulator. Conf. on SAR data processing for remote sensing, ROME, ITALY 290-300 Sep 1994


