Theoretical and Practical Limits on the Design and Fabrication of Free-Space Optical Interconnects

by

Vincent V. Wong

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degrees of

Master of Science

and

Bachelor of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1991

© Vincent V. Wong, MCMXCI. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute copies of this thesis document in whole or in part.

Author

Department of Electrical Engineering and Computer Science

May 15, 1991

Certified by

Clifton G. Fonstad

Professor, Electrical Engineering and Computer Science

Thesis Supervisor

Certified by

Gary J. Swanson

Technical Staff, MIT Lincoln Laboratory

Thesis Supervisor

Accepted by

Arthur C. Smith

Chairman, Departmental Committee on Graduate Students

JUL 24 1991
Theoretical and Practical Limits on the Design and Fabrication of
Free-Space Optical Interconnects

by

Vincent V. Wong

Submitted to the Department of Electrical Engineering and Computer Science
on May 15, 1991, in partial fulfillment of the
requirements for the degrees of
Master of Science
and
Bachelor of Science

Abstract

Theoretical and practical issues which arise in the design and fabrication of surface-relief
structures which implement free-space optical interconnects are considered. Two phase
optimization algorithms, simulated annealing and phase retrieval, are utilized for the design
of free-space optical interconnects implementing a Gaussian fanout pattern. The results of
the phase optimization procedures, binary and multi-level phase profiles, are presented and
analyzed. Surface-relief structures which implement the binary-phase profiles are fabricated
and tested. A simple model is presented which accounts for etch depth and linewidth
reproduction errors in the fabrication of binary surface-relief structures. Based on this
model the measured results agreed well with the expected results. This model also shows
that as the degree of optical fanout increases, the tolerances on etch depth control for good
reconstruction error become exceedingly stringent.

The behavior of binary surface-relief phase gratings is analyzed as a function of the
period-to-wavelength ratio and for two different substrate indices of refraction. A rigorous
vector coupled-wave formulation of grating diffraction is used to perform the analysis. The
results of this analysis are compared with those predicted by the use of a Fourier optics
formalism. A general trend towards increasing reconstruction error for decreasing values of
the period-to-wavelength ratio is observed. Furthermore, as the index of refraction of the
grating substrate increases, the grating diffraction behavior agrees more closely with that
predicted by the use of a Fourier optics formalism.

Thesis Supervisor: Clifton G. Fonstad
Title: Professor, Electrical Engineering and Computer Science

Thesis Supervisor: Gary J. Swanson
Title: Technical Staff, MIT Lincoln Laboratory
Acknowledgments

Over the last three years I have been privileged to work in the Binary/Diffractive Optics Group at MIT Lincoln Laboratory. My involvement with and interaction within the group have shaped both my scientific interests and personal character. This thesis represents the culmination of my stay and is comprised of research performed over the past year. By its nature, research cannot be performed successfully by a single individual. Much of this work would not have been possible without the guidance and expertise offered to me by many individuals. To these individuals, I would like to express my most sincere appreciation:

I would like to thank Gary Swanson for advising this thesis at MIT Lincoln Laboratory. He provided direction and insight throughout the course of the project. His unbelievable patience and understanding in the face of seemingly obvious and sometimes not so obvious questions are deeply appreciated. I have truly benefited from our many discussions on, among many other topics, the properties of binary-phase gratings, the scalar and thin-grating approximations and phase optimization algorithms. Many of these discussions have led to both new insight and new paths of investigation. His approachability, relaxed nature and sense of humor have also made working with him that much more enjoyable. Gary has been an advisor in every sense of the word.

I must thank Wilfrid Veldkamp for giving me the privilege and opportunity to work in the group and for his great enthusiasm and support of my endeavors. I also thank Professor Clifton Fonstad for supervising this thesis at MIT. His interest in the project and comments at various stages in the project are gratefully acknowledged.

I am indebted to Jim Leger, who is presently at the University of Minnesota. Jim, who has often taken so much of his valuable time to speak with me, taught me a great deal about Fourier optics, digital signal processing and the Fast Fourier Transform (only to name a few topics). I thank him also for ‘time-sharing’ lab space and equipment with me. Jim has shared many of his experiences with me and described various ‘pitfalls’ to avoid. For this, I am truly grateful.

I thank Bob Knowlden for his help in using the Mann software package and for providing the data for the necessary mask alignment marks. I thank him also for bringing me up to speed on using the DIFRACT code and the MIT Cray account. Finally, Bob’s sincere concern for my progress is acknowledged with my deepest gratitude.

I am grateful to Margaret Stern for her consistent effort to make the fabrication process run smoothly. Her careful characterization of the reactive-ion etching process was essential in meeting the necessary etch depth tolerances. The optical elements described in this thesis could not have been fabricated without her help. I also thank her for sharing with me her vast knowledge in the field of fabrication technology. Finally, her constant encouragement and support for me is deeply appreciated.

I also wish to thank Marsden ‘Denny’ Griswold. He was solely responsible for fabricating all of the optical elements which are described in this thesis. I thank him for the impeccable care with which he performed the fabrication. I would also like to thank Bill Delaney, Shirley Medeiros and Theresa Jay for their guidance and assistance in the fabrication lab.
I thank William ‘Woody’ Goltsos for his detailed explanation of the linewidth measurement technique. Woody’s enthusiasm is only rivaled by his friendly nature. I would also like to thank Michael Holz for explaining to me the simple concepts of the Amacronics processing system.

Michael Farn must also be thanked for expressing a genuine interest in my work. My discussions with him about phase optimization, Fourier optics and rigorous diffraction analysis have been extremely fruitful.

I thank Ed Motamedi of Rockwell International, Inc. for securing funds for the lithographic masks and for his continual support.

My close friends at MIT deserve a special thanks. They have kept me both sane and smiling for the past five years. They, as much as anyone else, have made this thesis possible.

Finally, I would like to thank my family - my Mom, my Dad and my sister, Karen - for their neverending love and support. It is they who have made all that I have ever done possible. To them, I dedicate this thesis.
Contents

Abstract .................................................. 2
Acknowledgements ........................................ 3
Table of Contents ......................................... 5
List of Figures ........................................... 7
List of Tables ............................................ 13

1 Introduction ........................................... 15

2 Diffraction gratings: Theory and Fabrication .......... 23
   2.1 Plane-Wave Decomposition .......................... 25
   2.2 The Transmission Function and Thin-Grating Approximation .......................... 26
   2.3 The Grating Angular Spectrum and Grating Equation .......................... 33
   2.4 Computational Representation of Diffraction Gratings .......................... 40
   2.5 Binary-Phase Gratings: An Example ..................... 43
   2.6 Binary Optics Technology ........................... 55
   2.7 Fabrication Errors .................................. 63

3 Phase Optimization Techniques ......................... 75
   3.1 Simulated Annealing ................................ 77
       3.1.1 Crossed 1D Grating .............................. 84
CONTENTS

3.1.2 2D Binary-Phase Grating ........................................ 97
3.2 Phase Retrieval ..................................................... 104

4 Experimental Results and Discussion .................................. 119
4.1 Measurement Techniques ............................................ 120
  4.1.1 Etch Depth Measurement ....................................... 120
  4.1.2 Linewidth Measurement ......................................... 120
  4.1.3 Power Spectrum Measurement ................................... 121
4.2 Results ............................................................. 123
  4.2.1 1D Binary-Phase Grating ....................................... 124
  4.2.2 2D Crossed Grating ............................................. 131
  4.2.3 2D Binary-Phase Grating ....................................... 136

5 The Scalar and Thin-Grating Approximations ......................... 142
5.1 Regimes of Grating Diffraction .................................... 146
5.2 Theoretical Results ................................................ 153
  5.2.1 1-to-5 fanout grating ......................................... 154
  5.2.2 1-to-6 fanout grating ......................................... 169
  5.2.3 1-to-11 fanout grating ....................................... 183
  5.2.4 Gaussian fanout grating ....................................... 188
5.3 Summary ........................................................... 196

6 Conclusions and Future Research .................................... 197
List of Figures

1.1 Amacronics processing system ........................................... 18

2.1 Wave incident on an arbitrary grating structure. ...................... 27

2.2 Types of phase gratings: (a) planar phase grating; (b) surface-relief phase grating. .................................................. 29

2.3 Wave incident on a phase step of height d. ............................ 31

2.4 Fourier representation of a diffraction grating: (a) grating transmits function; (b) grating angular spectrum. ...................... 34

2.5 k-vector plot of diffracted waves. ........................................ 38

2.6 Phase function of a binary grating. ..................................... 44

2.7 Power spectrum of 50% duty cycle grating with π-phase depth. ...... 47

2.8 Power spectrum of a 50% duty cycle grating for various phase depths

\( \Phi_o: (a) \Phi_o = \frac{3\pi}{4}; (b) \Phi_o = \frac{\pi}{2} \) ........................................ 48

2.8 Power spectrum of a 50% duty cycle grating for various phase depths:

\( \Phi_o: (c) \Phi_o = \frac{\pi}{4} \) ......................................................... 49

2.9 Zero-, first- and third-order diffraction efficiencies for 50% duty cycle grating versus phase depth. ........................................... 50

2.10 Power spectrum of a π-phase grating for various duty cycles \( D \): (a)

\( D = 0.25; (b) D = 0.40 \) ......................................................... 52
LIST OF FIGURES

2.10 Power spectrum of a π-phase grating for various duty cycles D: (c) $D = 0.60$; (d) $D = 0.75$. ........................................ 53

2.11 Zero- and first-order diffraction efficiencies for a π-phase grating versus duty cycle. ........................................ 54

2.12 Arbitrary phase profile $\Phi[n]$. ........................................ 56

2.13 Fabrication process: (a) Optical lithography; (b) Development of photore sist; (c) RIE; and (d) Removal of photore sist. ............... 57

2.14 Phase profile $\Phi[n]$ modulo π. ........................................ 58

2.15 Phase profile of a BOE. ........................................ 59

2.16 Block diagram of algorithm for calculating the phase profile of a BOE. ........................................ 60

2.17 Phase profiles of an arbitrary binary-phase element: (a) Desired profile; (b) Profile due to a linewidth error. ........................................ 69

3.1 Schematic representation of the annealing process ........................................ 78

3.2 Cost versus state space for an arbitrary system. ........................................ 79

3.3 Ideal 1D power spectrum. ........................................ 85

3.4 Surface-relief profile of 1D binary-phase grating optimized with simulated annealing. ........................................ 91

3.5 Portion of the lithographic mask corresponding to the surface-relief shown in Figure 3.4 ........................................ 92

3.6 Power spectrum of optimized 1D binary-phase grating (solid curve) and ideal Gaussian power spectrum (dashed curve). ........................................ 93

3.7 Surface-relief profile of 2D crossed grating. ........................................ 94

3.8 Power spectrum of 2D crossed grating. ........................................ 96

3.9 Ideal 2D power spectrum. ........................................ 98
3.10 Surface-relief profile of 2D binary-phase grating optimized with simulated annealing. ........................................ 100
3.11 Portion of the lithographic mask corresponding to the surface-relief shown in Figure 3.10 ........................................ 101
3.12 Power spectrum of optimized 2D binary-phase grating. ............... 102
3.13 Cross-section of 2D power spectrum of 2D binary-phase grating optimized using simulated annealing (solid curve) and cross-section of ideal 2D power spectrum (dashed curve). ................................. 103
3.14 Block diagram of phase retrieval algorithm. ............................. 105
3.15 Block diagram of modified phase retrieval algorithm ..................... 108
3.16 Squared-error versus number of iterations in phase retrieval algorithm. 111
3.17 Surface-relief profiles of 2D multi-level phase grating optimized with phase retrieval: (a) continuous phase; (b) quantized ($M = 16$) phase. 112
3.18 Power spectrum of 2D multi-level phase grating. ......................... 113
3.19 Cross-section of 2D power spectrum of 2D 16-level phase grating optimized using phase retrieval. ............................ 114
3.20 Portions of the lithographic masks corresponding to the surface-relief shown in Figure 3.17b: (a) Mask 1; (b) Mask 2. ......................... 115
3.20 Portions of the lithographic masks corresponding to the surface-relief shown in Figure 3.17b: (c) Mask 3; (d) Mask 4. ......................... 116
4.1 Experimental setup for power spectrum measurement. .................... 121
4.2 Measured (solid curve) and expected (dashed curve) power spectrums for 1D binary-phase grating. .............................. 126
4.3 Phase function $\Phi[n]$ for 1D binary-phase grating. ...................... 128
4.4 Power spectrums for 2D crossed grating: (a) measured; (b) expected. 133
4.5 Power spectrums for 2D binary-phase grating: (a) measured; (b) expected.

5.1 Phase profile (single period) of 1-to-5 fanout grating. ........................................ 155

5.2 Ideal power spectrum of 1-to-5 fanout grating. ....................................................... 156

5.3 Power spectrum of 1-to-5 fanout grating for various period-to-wavelength ratios: (a) $\Lambda/\lambda = 2$; (b) $\Lambda/\lambda = 5$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve. ....................................................... 158

5.3 Power spectrum of 1-to-5 fanout grating for various period-to-wavelength ratios: (c) $\Lambda/\lambda = 25$; (d) $\Lambda/\lambda = 50$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve. ....................................................... 159

5.4 Power spectrum of 1-to-5 fanout grating - TE (solid curve) and TM (dashed curve) incident polarizations; $\Lambda/\lambda = 50$; $n = 1.5$. ........................................ 160

5.5 Power spectrum of 1-to-5 fanout grating - TE (solid curve) and TM (dashed curve) incident polarizations; $\Lambda/\lambda = 5$; $n = 1.5$. ........................................ 161

5.6 Maximum reconstruction error versus period-to-wavelength ratio for 1-to-5 fanout grating - TE (*) and TM (o) incident polarizations; $n = 1.5$. 162

5.7 Power spectrum of 1-to-5 fanout grating - $\Lambda/\lambda = 9$; $n = 1.5$; TM incident polarization. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve. ....................................................... 164

5.8 Power spectrum of 1-to-5 fanout grating - $\Lambda/\lambda = 50$; $n = 4.0$; TE incident polarization. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve. ....................................................... 165

5.9 Power spectrum of 1-to-5 fanout grating and corresponding reverse parity grating - $\Lambda/\lambda = 2$; $n = 1.5$; TE incident polarization. ........................................ 167

5.10 Power spectrum of 1-to-5 fanout grating and corresponding reverse parity grating - $\Lambda/\lambda = 50$; TE incident polarization: (a) $n = 1.5$; (b) $n = 4.0$. 168
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.11</td>
<td>Phase profile (single period) of 1-to-6 fanout grating.</td>
<td>170</td>
</tr>
<tr>
<td>5.12</td>
<td>Ideal power spectrum of 1-to-6 fanout grating.</td>
<td>171</td>
</tr>
<tr>
<td>5.13</td>
<td>Power spectrum of 1-to-6 fanout grating for various period-to-wavelength</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>ratios: (a) $\Lambda/\lambda = 3$; (b) $\Lambda/\lambda = 6$. Rigorous</td>
<td></td>
</tr>
<tr>
<td></td>
<td>calculation - dashed curve; Fourier optics prediction - solid curve.</td>
<td></td>
</tr>
<tr>
<td>5.13</td>
<td>Power spectrum of 1-to-6 fanout grating for various period-to-wavelength</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>ratios: (c) $\Lambda/\lambda = 25$; (d) $\Lambda/\lambda = 50$. Rigorous</td>
<td></td>
</tr>
<tr>
<td></td>
<td>calculation - dashed curve; Fourier optics prediction - solid curve.</td>
<td></td>
</tr>
<tr>
<td>5.14</td>
<td>Power spectrum of 1-to-6 fanout grating - TE (solid curve) and TM</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>(dashed curve) incident polarizations; $\Lambda/\lambda = 50$; $n = 1.5$.</td>
<td></td>
</tr>
<tr>
<td>5.15</td>
<td>Power spectrum of 1-to-6 fanout grating - TE (solid curve) and TM</td>
<td>176</td>
</tr>
<tr>
<td></td>
<td>(dashed curve) incident polarizations; $\Lambda/\lambda = 6$; $n = 1.5$.</td>
<td></td>
</tr>
<tr>
<td>5.16</td>
<td>Zero-order reconstruction error versus period-to-wavelength ratio for 1-</td>
<td>177</td>
</tr>
<tr>
<td></td>
<td>to-6 fanout grating - TE (*) and TM (o) incident polarizations; $n = 1.5$.</td>
<td></td>
</tr>
<tr>
<td>5.17</td>
<td>Maximum higher-order reconstruction error versus period-to-wavelength</td>
<td>178</td>
</tr>
<tr>
<td></td>
<td>ratio for 1-to-6 fanout grating - TE (*) and TM (o) incident polarizations;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$n = 1.5$.</td>
<td></td>
</tr>
<tr>
<td>5.18</td>
<td>Power spectrum of 1-to-6 fanout grating - $\Lambda/\lambda = 50$; $n = 4.0$;</td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>TE incident polarization. Rigorous calculation - dashed curve; Fourier optics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>prediction - solid curve.</td>
<td></td>
</tr>
<tr>
<td>5.19</td>
<td>Power spectrum of 1-to-6 fanout grating and corresponding reverse parity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>grating - $\Lambda/\lambda = 50$; TE incident polarization: (a) $n = 1.5$;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(b) $n = 4.0182$.</td>
<td></td>
</tr>
<tr>
<td>5.20</td>
<td>Phase profile (single period) of 1-to-11 fanout grating.</td>
<td>184</td>
</tr>
<tr>
<td>5.21</td>
<td>Ideal power spectrum of 1-to-11 fanout grating.</td>
<td>185</td>
</tr>
</tbody>
</table>
5.22 Power spectrum of 1-to-11 fanout grating for various period-to-wavelength ratios: (a) $\Lambda/\lambda = 5$; (b) $\Lambda/\lambda = 10$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve. 186

5.22 Power spectrum of 1-to-11 fanout grating for various period-to-wavelength ratios: (c) $\Lambda/\lambda = 25$; (d) $\Lambda/\lambda = 50$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve. 187

5.23 Power spectrum of 1-to-11 fanout grating - TE (solid curve) and TM (dashed curve) incident polarizations; $n = 1.5$: (a) $\Lambda/\lambda = 10$; (b) $\Lambda/\lambda = 50$ 189

5.24 Maximum reconstruction error versus period-to-wavelength ratio for 1-to-11 fanout grating - TE (*) and TM (o) incident polarizations; $n = 1.5$. 190

5.25 Power spectrum of Gaussian fanout grating for various period-to-wavelength ratios: (a) $\Lambda/\lambda = 5$; (b) $\Lambda/\lambda = 10$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve. 192

5.25 Power spectrum of Gaussian fanout grating for various period-to-wavelength ratios: (c) $\Lambda/\lambda = 25$; (d) $\Lambda/\lambda = 50$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve. 193

5.26 Power spectrum of Gaussian fanout grating - TE (solid curve) and TM (dashed curve) incident polarizations; $\Lambda/\lambda = 50$; $n = 1.5$. 194

5.27 Maximum reconstruction error versus period-to-wavelength ratio for Gaussian fanout grating - TE (*) and TM (o) incident polarizations; $n = 1.5$. 195
List of Tables

2.1 Diffraction efficiencies for phase quantization to $2^M$ levels. ........ 60

3.1 Optimized power spectrum and reconstruction errors for 1D binary-
    phase grating optimized with simulated annealing. ............... 94

3.2 Optimized power spectrum and reconstruction errors for 2D binary-
    phase grating optimized with simulated annealing. ............... 103

3.3 Optimized power spectrum and reconstruction errors for 16-level phase
    grating optimized with phase retrieval. ......................... 117

4.1 Ideal and actual power spectrums for 1D binary-phase grating fabricated
    with a 130Å over-etch. .......................................... 125

4.2 Measured power spectrum and reconstruction errors for 1D binary-
    phase grating. .................................................. 125

4.3 Measured and expected ratios of diffraction efficiencies (normalized to
    the first-order diffraction efficiency) for 1D binary-phase grating. . 129

4.4 Measured and ideal linewidths for 1D binary-phase grating (single period). 130

4.5 Ideal and expected power spectrums for 2D crossed grating fabricated
    with a 50Å overetch for the x-grating and 275Å overetch for the y-grating. 132

4.6 Measured power spectrum and reconstruction errors for 2D crossed
    grating. ......................................................... 134
<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.7</td>
<td>Measured and expected ratios of diffraction efficiencies (normalized to the first-order diffraction efficiency) for 2D crossed grating.</td>
<td>136</td>
</tr>
<tr>
<td>4.8</td>
<td>Ideal and expected power spectrums for 2D binary-phase grating fabricated with a 81Å overetch.</td>
<td>137</td>
</tr>
<tr>
<td>4.9</td>
<td>Measured power spectrum and reconstruction errors for 2D binary-phase grating.</td>
<td>137</td>
</tr>
<tr>
<td>4.10</td>
<td>Measured and expected ratios of diffraction efficiencies (normalized to the first-order diffraction efficiency) for 2D binary-phase grating.</td>
<td>138</td>
</tr>
<tr>
<td>5.1</td>
<td>Reconstruction errors for 1-to-6 fanout grating for various period-to-wavelength ratios.</td>
<td>172</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Even the fastest modern digital computer cannot compare to the brain of an infant in the performance of intelligent information processing such as image processing and pattern recognition. Processing a high-resolution image on a conventional digital computer, which performs approximately $10^9$ operations per second, would take many minutes to many hours. And yet, the human brain, which performs only $10^3$ operations per second, can process this information essentially instantaneously. The key to the power of the brain is its layered structure, its massive parallel processing capability and its ability to process information in a hierarchical manner. In contrast, conventional computers process information in a sequential fashion.

Although integrated circuits have observed exponential growth in processing speed and capability since the birth of the industry, it is improbable that such growth will continue indefinitely. In fact, as circuit dimensions shrink and clock rates increase fundamental and practical limits in the areas of electrical interconnects and power consumption are rapidly being approached [1]. Consequently, intense research efforts are being directed towards developing parallel processing architectures that would far outperform modern conventional computing systems.

A typical parallel computing architecture consists of layers of processing planes,
or decision-planes, where electronic and/or optical devices perform various signal processing functions. Physically, these processing planes correspond to wafers or semiconductor chips containing electronic and optical logic devices. The flow of information, therefore, proceeds from processing layer to processing layer. The power of such an implementation is directly related to the degree of connectivity or “fanout” between processing layers.

Optics has emerged as the technology of choice for implementing these interconnections, due to the inherent parallelism of free-space wave propagation and the potential of achieving massively-interconnected systems. A free-space optical interconnection system consists of three components: a set of light sources such as lasers or LEDs; the free-space optical interconnect (FSOI) which implements the interconnection pattern, and a set of light detectors. The FSOI serves to redirect light from the set of light sources on a given processing plane to the set of light detectors on the next processing plane.

There are two characteristics of an FSOI: the degree of connectivity or fanout and the splitting ratio. The fanout is a measure of the number of detectors connected to each source. A 1-to-1 fanout connects a single source to a single detector. This situation typically occurs when transmitting an image from one plane to another, and can be achieved using refractive optical elements. The other more interesting situation is a 1-to-N fanout, where each source is connected to many detectors. Applications such as neural computing, pattern recognition, optical switching and clock distribution employ this scheme. Clearly, conventional refractive optical elements cannot be used in this situation. Rather, diffractive optical elements, particularly diffraction gratings, which can split an incoming beam into many beams or diffraction orders must be used. By optimizing the grating transmittance over one period, a specified
CHAPTER 1. INTRODUCTION

angular spectrum, also referred to as the far-field diffraction pattern of the grating, can be approximated.

The splitting ratio, which is only relevant in the 1-to-N fanout case, describes the relative powers directed into each of the em N diffraction orders of the grating. In general, one can have a uniform and a non-uniform splitting ratio. Referring to the applications mentioned above, clock distribution and optical switching often require uniform splitting ratios, whereas neural computing and pattern recognition applications call for non-uniform splitting ratios.

The problem at hand, therefore, is to design an FSOI given a specification of its degree of fanout and splitting ratio. Since the efficiency of this element should be high, we restrict ourselves to using pure phase elements (as opposed to elements which contain both amplitude and phase transmittance components). Furthermore, we focus on surface-relief phase gratings, as compared to planar phase gratings, owing to their large potential for integration with electronic and/or optical devices. In general, an exact solution for the grating phase (or surface-relief) profile in this framework does not exist.

A great deal of research has been conducted on the design of free-space optical interconnects. The design algorithms involve the solution of a nonlinear, combinatorial optimization problem. Two techniques, simulated annealing [2] and phase retrieval [3] [4] [5], have found widespread use in this regard.

Much of the research has focused on the design of optical elements which generate arrays of equally intense (i.e. uniform splitting ratio) light spots [6] [7] [8]. Periodic arrays of illumination spots can be used for applications ranging from switching arrays of optical devices [9], optically pumping arrays of vertical-cavity surface-emitting lasers [10] and synchronous clock distribution on VLSI chips [1]. Recently, optical
Figure 1.1: Amacrionics processing system
elements generating arrays of arbitrarily-weighted (i.e. nonuniform splitting ratio) light spots have received much interest. Such elements are essential components in the optical implementation of parallel processing algorithms, neural networks and machine vision systems [11] [12] [13].

In this work, the optimization techniques of simulated annealing and phase retrieval are used to design an FSOI with a 1-to-121 fanout in two dimensions (i.e. an 11-by-11 array of spots) and a Gaussian splitting ratio. This FSOI will ultimately be used for Amacronics, a novel image processing system that mimics the way the human eye works [14]. The Amacronics processing system, shown in Figure 1.1, consists of layered structures of electronic and optical devices and micro-optics. The processing is performed in-plane, and the results are relayed to the next processing plane by the optical interconnect (the top quartz layer in the figure). As shown schematically, the optical interconnect directs light from each light source into several neighboring detectors. The choice of the Gaussian fanout pattern is based on a velocity-extraction technique developed by Allen Waxman and his colleagues [15].

Although designs of FSOIs implementing large fanouts have been realized, experimental demonstrations of FSOIs have only implemented relatively small degrees of fanout [16] [17] [18]. This fact arises from the stability of the optimized solutions with respect to the various sources of error in fabrication, such as over- or under-etching, linewidth reproduction errors, misalignment, and incorrect exposure times. Therefore, it is important to examine the stability of the optimized solution with respect to these sources of error in fabrication. Indeed, the realization of massively-parallel computing systems may ultimately be limited by fabrication capabilities. For certain phase gratings, as the degree of fanout increases, the sensitivity of the angular spectrum of the grating to the etch depth also increases. In this work, the sensitiv-
ity of the angular spectrum of FSOIs to both small perturbations in etch depth and linewidth reproduction errors is investigated.

The design algorithms described above assume that scalar diffraction theory, derived from a simplification of Maxwell's equations, is valid. In short, the scalar theory is valid for large period-to-wavelength ratios [19] [20]. This condition, along with the so-called 'thin-grating approximation', results in a simple Fourier transform relation between the grating transmittance function and the grating angular spectrum. This Fourier relationship forms the basis of the formalism of Fourier optics [20]. Furthermore, such a relationship is ideally suited for numerical computation on standard computers.

However, as optical systems become more compact and device dimensions shrink, the physical period of the grating and the diffracting features within each period rapidly approach the wavelength of the incident radiation. In this regime, scalar diffraction theory breaks down, and the various optimization techniques which utilize this formalism become obsolete. To solve the diffraction problem correctly a full electromagnetic wave theory must be used [21]. In theory one could design and optimize grating profiles using the full electromagnetic wave formalism; however, in practice extremely long computation times preclude this approach. Therefore, to maintain the integrity of solutions obtained using a Fourier optics formalism, it is necessary to determine the regimes where this formalism is valid. In this work, the properties of binary surface-relief phase elements as a function of the period-to-wavelength ratio and for different grating substrate indices of refraction are examined. These properties, of course, are structure-dependent; therefore, the behavior of representative binary-phase gratings is considered.

The goal of this thesis project, therefore, is both to investigate various optimiza-
CHAPTER 1. INTRODUCTION

tion techniques for the design of free-space optical interconnects and also to address theoretical and practical issues in the design and fabrication of free-space optical interconnects. The thesis is arranged in six chapters. In this chapter the concept of free-space optical interconnects was introduced, the motivation of research was given and the project goals were stated.

Chapter 2 is a general discussion of the theory and fabrication of diffraction gratings. It begins with a description of scalar diffraction theory. This formalism is then used to derive the properties of diffraction gratings. The computational representation of a diffraction grating and its angular spectrum is also considered. Next, a simple phase grating is analyzed. The technology of Binary Optics, which is used to fabricate optical elements with arbitrary surface-relief profiles, is then outlined. Finally, sources of error in the fabrication procedure, particularly etch depth errors and linewidth errors, are addressed.

Chapter 3 introduces the techniques of simulated annealing and phase retrieval. These techniques are used to determine the optimum grating phase profiles for implementing the Gaussian fanout described above. The optimized results are presented and analyzed.

Chapter 4 presents the experimental measurements of the surface-relief phase gratings which implement the free-space optical interconnects described in Chapter 3. The measured power spectrums are compared to the expected spectrums determined from etch depth and, in some cases, linewidth measurements.

Chapter 5 considers the issue of the validity of the scalar and thin-grating approximations for surface-relief phase gratings. Several surface-relief phase gratings are analyzed as a function of both the period-to-wavelength ratio and the substrate index of refraction. The chapter begins with a description of the various regimes of
grating diffraction. Next, a rigorous vector coupled-mode formalism, which is used to analyze the above grating structures, is briefly outlined. The results of this analysis are compared with those given in the limit where both the scalar and thin-grating approximations are well-satisfied.

Chapter 6 summarizes the results of the thesis project and also suggests some directions for future research.
Chapter 2

Diffraction gratings: Theory and Fabrication

A diffraction grating is defined as an object which imparts periodic amplitude and/or phase variations on an incident wave. The periodic nature of the grating serves to decompose the incident wave into many exiting waves, or diffraction orders, travelling in different directions, or diffraction angles. In most applications of diffraction gratings, use is made of the grating’s dispersive, or wavelength-dependent, properties and/or the grating’s ability to divide the power contained in an incident wave among various diffracted waves. Diffraction gratings have found widespread use in fields such as spectroscopy, quantum electronics, integrated optics, holography and optical information processing. Specific applications include laser beam combining [22], grating-surface-emitting lasers [23], pulse compression [24] and, as mentioned earlier, free-space optical interconnects [16] [18].

A rigorous analysis of diffraction gratings involves a full vectorial solution of Maxwell’s equations [21]. However, this method of solution, although exact, is quite computationally-intensive and tedious. A simpler approach is to use a scalar theory. This theory is derived from Maxwell’s equations by making an approximation known
as the *scalar approximation*. In this approximation light is treated as a scalar field, as opposed to a vector field. In other words, only one transverse component of the electric or magnetic field is considered at a time, and superposition is used to generate the results for the entire electric or magnetic fields. This treatment neglects the fact that the electric and magnetic fields are coupled through Maxwell's equations [20].

The scalar approximation is valid under two conditions: *i*) the diffracting features must be large compared to the wavelength of the incident light, and *ii*) the diffracted fields must be observed far from the diffracting aperture. In this chapter we assume that these conditions are well-satisfied. Chapter 5 discusses in more detail the situations where the scalar theory breaks down.

In this chapter, the fundamental theory of diffraction gratings in the scalar limit is developed and the fabrication of such grating structures is described. Section 2.1 derives the plane-wave decomposition or angular spectrum of a complex scalar field. Section 2.2 introduces the concept of the grating *transmission function* and the *thin-grating approximation*. Section 2.3 relates the transmission function to the angular spectrum of the thin-grating. The concept of diffraction efficiency is then introduced and the famous "grating equation" is derived. Section 2.4 considers the topic of the computational representation of diffraction gratings. The Discrete Fourier Transform (DFT) and aliasing issues are discussed. Also, the consequences of the DFT representation of grating structures on fabrication are addressed. Section 2.5 utilizes the concepts developed in previous sections to analyze a simple diffraction grating. Section 2.6 describes the grating fabrication technology of Binary Optics. Section 2.7 examines the effect of fabrication errors on the performance of grating structures.
CHAPTER 2. DIFFRACTION GRATINGS: THEORY AND FABRICATION 25

2.1 Plane-Wave Decomposition

Consider a complex scalar field distribution incident on the $xy$ plane at $z = 0$. In
general, this field, denoted $U(x, y, 0)$, varies in the $xy$ plane. Across the $xy$ plane, the
function $U(x, y, 0)$ has a two-dimensional Fourier transform given by

$$A_o(f_x, f_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x, y, 0) \exp[-j2\pi(f_xx + f_yy)] \, dx \, dy \quad (2.1)$$

where $f_x$ and $f_y$ are spatial frequencies. Mathematically, the operation of a
Fourier transform is to decompose an arbitrary function into a superposition of
weighted complex-exponential functions. We show this explicitly by writing $U(x, y, 0)$
as the inverse Fourier transform of its spectrum

$$U(x, y, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_o(f_x, f_y) \exp[j2\pi(f_xx + f_yy)] \, df_x \, df_y. \quad (2.2)$$

The power of such a representation is manifested when such waves propagate
through linear-space-invariant systems, such as free-space.

To gain some physical insight into Eq. (2.2), we consider a unit-amplitude plane
wave, $P(x, y, z)$, travelling with direction cosines $(\alpha, \beta, \gamma)$. The equation for this wave
is given by

$$P(x, y, z) = \exp[j\frac{2\pi}{\lambda}(\alpha x + \beta y + \gamma z)] \quad (2.3)$$

where (from the dispersion relation)

$$\gamma = \sqrt{1 - \alpha^2 - \beta^2}. \quad (2.4)$$

Now, comparing Eq. (2.3) with Eq. (2.2) shows that at $z = 0$ a complex-
exponential function $\exp[j2\pi(f_x x + f_y y)]$ may be regarded as a plane wave prop-
agating with direction cosines

\[ \alpha = \lambda f_x \quad \beta = \lambda f_y \quad \gamma = \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2}. \quad (2.5) \]

The complex amplitude of this wave is then \( A_0(f_x, f_y) \int f_x \, df_Y \), evaluated at \( (f_X = \frac{\alpha}{\lambda}, f_Y = \frac{\beta}{\lambda}) \). For this reason the function

\[ A_0(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x, y, 0) \exp \left[ -j 2\pi \left( \frac{\alpha}{\lambda} x + \frac{\beta}{\lambda} y \right) \right] \, dx \, dy \quad (2.6) \]

is referred to as the angular spectrum of the disturbance \( U(x, y, 0) \).

We conclude that given any disturbance \( U(x, y, z) \) at a plane \( z = z_0 \), we can decompose this disturbance into a superposition of plane waves, each with its own complex amplitude, travelling in various directions. We will use this fact to analyze the effects of diffraction gratings in the following sections.

### 2.2 The Transmission Function and Thin-Grating Approximation

The function of a diffraction grating is to modify an incident wave in amplitude and/or phase to generate multiple exiting waves. Therefore, to characterize a grating, one must specify its effect on an arbitrary wave. This is done by defining a transmission function for the grating [25]. Consider Figure 2.1 which shows a plane wave \( V_0(x, y, z) = A \exp \left[ j 2\pi (\alpha x + \beta y + \gamma z) \right] \) incident on an arbitrary grating at \( z = -d \). The grating region extends from \( z = -d \) to \( z = 0 \) and is assumed to be infinite in the \( x- \) and \( y- \) directions. In the absence of the grating the complex field in the \( xy \) plane at \( z = 0 \) would be that of the incident wave \( V_0(x, y, z) \) evaluated at \( z = 0 \). The presence of the grating, however, results in a complex field at \( z = 0 \) different than
Figure 2.1: Wave incident on an arbitrary grating structure.

$V_o(x, y, z)$. We denote this field by $V(x, y, z)$. The transmission function $T(x, y)$ of the grating is defined as

$$T(x, y) = \frac{V(x, y, 0)}{V_o(x, y, 0)}. \quad (2.7)$$

The most general grating will modify both the amplitude and phase of an incident wavefront. The transmission function corresponding such a grating is given by

$$T(x, y) = |T(x, y)| \exp [j\Phi(x, y)] \quad (2.8)$$

where $|T(x, y)|$ and $\Phi(x, y)$ express the amplitude and phase modulation, respectively. From Eq. (2.7) and Eq. (2.8)
CHAPTER 2. DIFFRACTION GRATINGS: THEORY AND FABRICATION

\[ |T(x, y)| = \frac{|V(x, y, 0)|}{|V_o(x, y, 0)|} \quad (2.9) \]

and

\[ \Phi(x, y) = \angle \{V(x, y, 0)\} - \angle \{V_o(x, y, 0)\} \quad (2.10) \]

where \( \angle \{F\} \) represents the phase of the complex function \( F \). The transmission function will, in general, depend not only on \( x \) and \( y \), but also on the direction of illumination, given by the direction cosines \( \alpha \) and \( \beta \).

As yet, the grating region in Figure 2.1 is unspecified. In general, the grating can impart amplitude and/or phase variations on an incident wave. However, as mentioned earlier we will restrict our discussion to those gratings which only impart phase variations (i.e. \( |T(x, y)| = 1 \)) on an incident wave since amplitude variations necessarily decrease the efficiency of the grating.

There are two types of phase gratings: planar and surface-relief, shown in Figure 2.2. In both cases phase modulation is achieved by modulations in refractive index across the grating. In a planar phase grating (Figure 2.2a) the material properties of the original substrate are modified to induce an index modulation. One example of this approach occurs in acousto-optic modulators, where an acoustic wave is sent travelling though the substrate. The index at a given point (and instant in time) in the substrate is related to the amplitude of the acoustic wave at that point (and time). In a surface-relief grating (Figure 2.2b) regions of the substrate are selectively removed, or etched, to form the index modulation. The fabrication of surface-relief gratings, described in more detail in Section 2.6, is highly compatible with standard semiconductor integrated circuit fabrication techniques. Therefore, unlike planar phase gratings, surface-relief gratings offer a large potential for monolithic integration with
Figure 2.2: Types of phase gratings: (a) planar phase grating; (b) surface-relief phase grating.
electronic and/or optical devices. Consequently, surface-relief phase gratings have
become the method of choice for implementing optical interconnects, and we restrict
our discussion of phase gratings accordingly.

The thin-grating approximation assumes that the grating region in Figure 2.1
is infinitely thin (i.e. \( d \to 0 \)). In this limit the transmittance function \( T(x,y) \) is
independent of the angle of incidence (i.e. \( \alpha \) and \( \beta \)). Stated another way, a grating
is considered thin if a ray incident on the grating at coordinates \((x,y)\) at \( z = -d \)
emerges at approximately the same coordinates \((x,y)\) at \( z = 0 \).

For surface-relief phase gratings typical etch depths are on the order of a wave-
length. Consequently, in the regime where scalar theory is valid (i.e. \( \frac{A}{\lambda} \gg 1 \)), most
gratings can be treated as thin gratings. Of course, any finite etch depth invalidates
the claims of the thin-grating approximation; however, this approximation is a good
one from a convergence-viewpoint for sufficiently shallow etch depths. Chapter 5
considers the thin-grating approximation in more detail.

A consequence of the validity of the thin-grating approximation is that the ac-
cumulated phase through a certain point on the surface-relief grating is (for a given
substrate index and surrounding index) directly proportional to the etch depth at that
point. Therefore, the surface-relief profile or, formally, the etch depth of the phase
grating as a function of position \( d(x,y) \), determines the phase modulation \( \Phi(x,y) \)
or, equivalently, the transmittance function \( \exp[j\Phi(x,y)] \) of the grating. We now
establish the relationship between \( d(x,y) \) and \( \Phi(x,y) \).

Consider Figure 2.3 which shows a unit-amplitude (i.e. \( \Phi(x,y) = 0 \)) plane wave
\( U_{inc}(x,y,z) = \exp[-jkz] \) normally-incident on a phase step of height \( d \), where \( k \) is
the free-space wavenumber \( \frac{2\pi}{\lambda_0} \). Recall that the transmission function is independent
of the angle of incidence in the thin-grating approximation; therefore, the treatment
Figure 2.3: Wave incident on a phase step of height $d$. 
presented here is general to the extent that the approximations are valid. The phase element is assumed to extend infinitely in the \(x\)- and \(y\)-directions, and \(n_1\) and \(n_2\) are the indices of refraction of the two materials on either side of the step.

The goal is to find the phase function of this phase step, \(\Phi(x, y)\), given by eq. (2.10). To calculate \(\Phi(x, y)\), we first choose, without a loss of generality, the plane \(z = -d\) as a reference plane where both the phase of the incident field \(\Phi_{\text{inc}}(x, y)\) and the phase of the transmitted field \(\Phi_t(x, y)\) are equivalently zero. Next, we consider the regions \(x > 0\) and \(x < 0\) separately. It was determined that the phase shift through each of these regions on the step is proportional to the step height. So, for \(x > 0\) the phase of the transmitted wave \(\Phi_t(x > 0, y, 0)\) is given by

\[
\Phi_t(x > 0, y, 0) = -\frac{2\pi}{\lambda_0} n_2 d
\]  
(2.11)

and for \(x < 0\) the phase of the transmitted field \(\Phi_t(x < 0, y, 0)\) is given by

\[
\Phi_t(x < 0, y, 0) = -\frac{2\pi}{\lambda_0} n_1 d.
\]  
(2.12)

Similarly, the phase of the incident wave in the absence of the phase step is found to be

\[
\Phi_{\text{inc}}(x, y, 0) = -\frac{2\pi}{\lambda_0} n_1 d
\]  
(2.13)

for all \(x\). From Eq. (2.10) - Eq. (2.13) the phase function of the grating \(\Phi(x, y)\) is given by

\[
\Phi(x, y) = \begin{cases} 
\frac{2\pi}{\lambda_0} (n_1 - n_2) d & , x > 0 \\
0 & , x < 0.
\end{cases}
\]  
(2.14)
CHAPTER 2. DIFFRACTION GRATINGS: THEORY AND FABRICATION

As mentioned previously, the etch depth $d$ will in general be a complicated function of position; therefore, Eq. (2.14) generalizes to

$$\Phi(x, y) = \frac{2\pi}{\lambda_0} (n_1 - n_2) d(x, y).$$

(2.15)

In formulating Eq. (2.15) the surface-relief profile $d(x, y)$ was assumed to be known. Practically, it is often the phase function $\Phi(x, y)$ that is known (i.e. from computer simulations) and the surface-relief profile that is to be determined. This situation is typified by optimization of the phase function of a diffraction grating to realize a specified angular spectrum, and is described in more detail in Chapter 3.

In this section we have shown that the thin-grating approximation leads to a simple interpretation of the transmission function of a grating. The transmission function describes the effect of the grating structure on the amplitude and/or phase of an incident complex field. Given the incident field distribution at the input plane of the grating, the transmission function relates this field to the field at the output plane of the grating. This result, along with the results of Section 2.1, will be used to analyze the angular spectrum of a diffraction grating in the next section.

2.3 The Grating Angular Spectrum and Grating Equation

In the previous section it was shown that the grating transmission function relates the complex field distribution at the input plane of a grating structure to the complex field distribution at the output plane of the grating structure. Given the output distribution we can decompose the transmitted field into a superposition of weighted complex-exponentials, as described in Section 2.1, which correspond to the plane-wave components of the diffraction grating. In this section we show that the periodic nature
Figure 2.4: Fourier representation of a diffraction grating: (a) grating transmittance function; (b) grating angular spectrum.

of the diffraction grating results in an intuitive description of the properties of the diffraction grating. Specifically, it is shown that the angular spectrum of a diffraction grating consists of discrete plane-wave components whose amplitudes depend on the detailed nature of the grating transmittance function within a single period of the grating. Furthermore, the angles at which these plane-wave components propagate depend on the physical dimension of the grating period and the wavelength of light, a result summarized by the famous "grating equation".

A diffraction grating is by definition a structure that introduces a periodic modulation in amplitude and/or phase to an incident complex field distribution. This fact is expressed in terms of the grating transmission function as $T(x) = T(x + \Lambda)$, where $\Lambda$ is the grating period. We consider a one-dimensional grating for simplicity. Extension of the following discussion to a two-dimensional grating is straightforward.
CHAPTER 2. DIFFRACTION GRATINGS: THEORY AND FABRICATION

The grating transmittance can be expressed as

\[ T(x) = \sum_{m=-\infty}^{+\infty} \delta(x - m\Lambda) * T_{\Lambda}(x) \]  

(2.16)

where \(*\) denotes the convolution operation and \(T_{\Lambda}(x)\) is the transmittance function of the grating over a single period. We recognize the summation term in the convolution as an infinite train of unit amplitude impulses spaced \(\Lambda\) apart. The above relationship is shown schematically in Figure 2.4a, which plots the magnitude component of the transmittance function of an arbitrary grating, \(|T(x)|\). The corresponding phase of the transmission function, \(\Phi(x)\), (not shown) exhibits similar behavior.

Let a unit magnitude uniform plane wave \(U(x, z)\) be normally incident on this grating. The complex field distribution \(U(x, 0)\) in the plane directly after the grating is given simply by the transmittance function \(T(x)\). We next perform a plane-wave decomposition of this field, as described in Section 2.1, by taking the Fourier transform of \(T(x)\) given by Eq. (2.16). Noting that a convolution in the space-domain transforms into a multiplication in the Fourier-domain, the angular plane wave spectrum, \(A(f_x)\), is given by

\[ A(f_x) = \left(\frac{1}{\Lambda} \sum_{k=-\infty}^{+\infty} \delta(f_x - \frac{k}{\Lambda})\right)(T_{\Lambda}(f_x)) \]  

(2.17)

where \(T_{\Lambda}\) denotes the Fourier transform of a single period of the diffraction grating. We recognize the right-hand side of Eq. (2.17) as a sampled version of the Fourier transform of a single period of the diffraction grating. The frequency samples are taken at integer multiples of \(\frac{1}{\Lambda}\). The sampling (i.e. the impulse train in Eq. (2.17)) arises from the periodic nature of the diffraction grating. The sampling, in turn, gives rise to a discrete angular plane wave spectrum for the grating, unlike the continuous
angular plane wave spectrum of non-periodic structures. The above result is shown schematically (again, magnitude only) in Figure 2.4b.

Expressing the complex field $U(x, 0)$ as the inverse Fourier transform of the angular spectrum gives

$$U(x, 0) = \int_{-\infty}^{+\infty} \left( \frac{1}{\Lambda} \sum_{m=-\infty}^{+\infty} \delta(fx - \frac{m}{\Lambda}) \right) \exp [j2\pi fx \Lambda] dfx$$

$$= \frac{1}{\Lambda} \sum_{m=-\infty}^{+\infty} T_{\L}(\frac{m}{\Lambda}) \exp [j2\pi \frac{m}{\Lambda} x], \ m = \text{integer} \quad (2.18)$$

which expresses $U(x, 0)$ as an infinite superposition of plane waves with direction cosines

$$\alpha = m \frac{\lambda}{\Lambda}, \ \gamma = \sqrt{1 - (m \frac{\lambda}{\Lambda})^2}, \ m = \text{integer} \quad (2.19)$$

and amplitudes $a_m$

$$a_m = \frac{1}{\Lambda} T_{\L}(\frac{m}{\Lambda}) \quad (2.20)$$

where the $a_m$'s are in general complex. When $\alpha^2 < 1$ or, equivalently, $m < \frac{\Lambda}{\lambda}$, Eq. (2.18) corresponds to a propagating (i.e. non-zero time-averaged power in the $z$-direction) wave, whereas $\alpha^2 > 1$ or, equivalently, $m > \frac{\Lambda}{\lambda}$, corresponds to an evanescent (i.e. zero time-averaged power in the $z$-direction) wave. Since it is the time-averaged power that is actually detected, evanescent waves are not important for optical interconnect applications. The time-averaged power, or diffraction efficiency $\eta_m$, contained in a propagating order, $m$, is given by the magnitude-squared of its complex amplitude $a_m$: 
\[ \eta_m = |a_m|^2 , \ m < \frac{\Lambda}{\lambda} . \]  

(2.21)

For non-propagating orders physical constraints demand that their \( a_m \) coefficients go to zero. Therefore,

\[ \eta_m = 0 , \ m > \frac{\Lambda}{\lambda} . \]  

(2.22)

This fact shows that Fourier analysis of a grating is not strictly accurate in predicting grating efficiencies, because non-zero values result for non-propagating orders. Fourier analysis gives an accurate description of a grating's efficiency when the ratio of the sum of the efficiencies of the propagating orders to the sum of the efficiencies of all orders is approximately 1 [26]. In the regime where scalar diffraction theory is valid these conditions are nearly always satisfied. For optical interconnect applications these conditions translate into the following: If large fanouts are to be achieved, then one must use grating structures with large period-to-wavelength ratios, \( \frac{\Lambda}{\lambda} \). Chapter 5 discusses the consequences of the period-to-wavelength ratio on grating diffraction in more detail.

In addition to the diffraction efficiency of each order it is also important to determine the diffraction angles for the various orders. These angles must be considered when size constraints are placed on the optical system. For example, in the Amacronics processing system (see Figure 1.1) the detectors (top layer) are spaced 200\( \mu \)m apart, and the source-to-detector spacing is on the order of 1 cm. Given these system constraints the necessary diffraction angles are determined. We show that these angles are simply related to the grating period and the wavelength of light. This relation is the well-known "grating equation".
Figure 2.5: *k*-vector plot of diffracted waves.
To solve for the set of diffraction angles, $\theta_m$, we consider Figure 2.5, which plots the wavevector $\vec{k}$ along with its $x$-component, $k_{m,n} \hat{x}$ for various $m$ values for a transmitting medium of index of refraction $n$. We see that the $\theta_m$'s are given by the equation

$$\sin \theta_m = \frac{k_x}{|k|}$$

$$= \frac{m \lambda}{n \Lambda}, m = \text{integer.} \tag{2.23}$$

This equation is known as the "grating equation", formulated for normal incidence. For non-normal incidence at an angle $\theta_i$ with respect to the $z$-axis, Eq. (2.23) becomes

$$\sin \theta_m = \frac{m \lambda}{n \Lambda} - \sin \theta_i, m = \text{integer.} \tag{2.24}$$

Note that we can also understand the concept of evanescent orders from Eq. (2.23). When $\frac{k_x}{|k|} = 1$, a condition identical to $\frac{m \lambda}{\Lambda} = 1$ ($n = 1$), the angle $\theta_m$ is 90 degrees. This corresponds to a wave propagating along the surface of the grating in the $\hat{x}$-direction. For values of $\frac{m \lambda}{\Lambda} > 1$ the angle is imaginary. Physically, this corresponds to a wave propagating along the grating surface in the $\hat{x}$-direction and attenuating in the $\hat{z}$-direction, the characteristic description of an evanescent wave.

In summary the "grating equation" can be used to determine the angles at which the various diffracted orders propagate. For optical interconnect applications this equation is useful when system constraints are present.

Theoretically, the above derivations are exact (in the scalar regime). However, we show in the next section that practical considerations related to computational methods call for a different representation of grating structures than that assumed in the above derivations.
2.4 Computational Representation of Diffraction Gratings

The transmission functions discussed above were all continuous functions of space. Therefore, to exactly represent the grating transmittance function, an infinite number of values must be used. Furthermore, a Continuous Fourier Transform (CFT), which mathematically decomposes the grating transmittance function into an infinite superposition of complex-exponential, was used to derive the grating angular spectrum. In practice, however, an infinite number of terms cannot be retained. Instead, a finite number of terms $N$ is used to represent both the grating transmittance function and angular spectrum. The operation relating these functions is known as the Discrete Fourier Transform (DFT). In this section we briefly describe the computational representation of a grating structure and its consequences on the design and fabrication of optical interconnects.

The $N$-point DFT $T[m]$ of a function $t[n]$ is given by

$$T[m] = \frac{1}{N} \sum_{n=0}^{N-1} t[n] \exp \left[ j \frac{2\pi}{N} mn \right]$$

(2.25)

where $N$ is the number of samples used to approximate both $t(x)$ and $T(f_X)$. The evolution from the CFT to the DFT involves sampling and truncation in the space domain and sampling in the frequency domain. Note that sampling in the space domain results in a periodic function in the frequency domain; similarly, sampling in the frequency domain results in a periodic function in the space domain. Consequently, the DFT requires that both the space and frequency functions, $t[n]$ and $T[m]$, respectively, be periodic functions with period $N$.

For an arbitrary periodic function in space $t(x) = t(x+\Lambda)$ the sampling rate, $f_s$, is
chosen as \( \frac{N}{\Lambda} \), and the truncation window is chosen to be \( \Lambda \). A truncation window not equal to an integer multiple of a period introduces unwanted frequency components in the spectrum, an effect termed \textit{leakage} [27] [28]. As in the continuous-space case, sampling in the space domain is performed at a rate such that aliasing is minimized. Similarly, frequency samples are taken at intervals of \( \frac{1}{\Lambda} \) to prevent aliasing in the space domain.

In the above framework we can represent a diffraction grating as follows. Given \( t_{\Lambda}(x) \), one period of the ideal continuous-space transmission function \( t(x) = t(x + \Lambda) \), we generate \( t[n] \) by taking \( N \) samples of \( t(x) \) over one period \( \Lambda \) (i.e. \( t[n] = t(n \frac{\Lambda}{N}) \)). In representing the grating period as an \( N \)-element array of transmittance values, we have necessarily introduced space quantization. The grating is represented as an array of pixels whose sizes have been quantized to integer multiples of \( \frac{\Lambda}{N} \). The transmittance of each pixel is constant; however, there are no restrictions on the phase of this transmittance. We add that in most computer optimization problems, one does not begin with an actual function \( t_{\Lambda}(x) \). Rather, one begins the problem with the function \( t[n] \), specified by either the designer or the optimization algorithm. The issues of aliasing described here then deal with the hypothetical function \( t_{\Lambda}(x) \) obtained from a bandlimited-interpolation of the function \( t[n] \) [29].

Since the function \( t_{\Lambda}(x) \) is space-limited, the spectrum \( T_{\Lambda}(f_{X}) \) will not be band-limited; therefore, aliasing must result. Note that this aliasing is only due to the nature of the computation and is not present an experimental sense. However, in designing grating profiles on the computer we must not allow such aliasing to occur simply from a modelling point of view. Practically, we choose \( N \) (i.e. the sampling rate) such that this aliasing is minimized. For a sampling rate of \( \frac{N}{\Lambda} \), the angular spectrum should not have significant contributions above a frequency \( \frac{N}{2\Lambda} \). This is
simply a statement of Nyquist's Sampling Theorem [29].

As an example, we consider the Gaussian fanout element that will be incorporated in the Amacronics system. Since the detectors will be spaced apart by 200µm, let the aperture of the each Gaussian fanout element be 200µm. Let us also choose (arbitrarily) that a minimum of 10 periods of the grating structure should fit over this 200µm aperture. Therefore, the spatial frequency between adjacent angular spectrum components is 50 lines/mm. We are interested in the orders between -/+ 5. The maximum non-zero frequency component in the angular spectrum would be 250 lines/mm. Therefore, the number of samples $N$ needed to prevent significant aliasing would be 10. In general, for a 1-to-$M$ fanout the required number of samples is $N = 2M$. Practically, one typically takes slightly more than the $2M$ samples since the analysis assumed a band-limited angular spectrum. Computationally, such a task is not a problem; one simply has to wait a longer amount of time for the calculation to finish. Practically, however, one must be aware of the fabrication capabilities.

For a period of 20µm and $N = 10$ samples, each pixel of the grating structure is 2.0µm. Present fabrication technology is capable of reproducing linewdths of approximately 0.5µm. Therefore, the specifications in the above example are within present fabrication capabilities and should be feasible. However, as the degree of fanout increases (for a fixed wavelength and assuming that the scalar approximation remains valid) the pixel size decreases. The minimum pixel size that can be fabricated translates into a maximum limit on the degree of fanout (once again for the assumptions made above). Therefore, the realization of optical interconnects with large fanouts will ultimately be limited by fabrication technology.

In this section the computational representation of grating structures and their angular spectrums was presented. It was shown that this representation, known as
the Discrete Fourier Transform, has practical consequences in terms of the fabrication of the grating structures. Large optical fanouts and parallel processing systems implementing such fanouts may ultimately rely on fabrication capabilities. In the next section we analyze a simple example that incorporates the concepts that have been developed in preceding sections.

2.5 Binary-Phase Gratings: An Example

The phases of $t[n]$, which represents the individual pixels of the surface-relief phase grating, will in general take on a continuum of values. The resulting surface-relief profile, however, is exceedingly hard to fabricate. To alleviate this problem, one can restrict the phase of each pixel to take on one value out of a finite number $L$ of possible values. In other words, one can introduce a phase quantization. Quantized-phase gratings are generally classified in two categories: binary-phase gratings (i.e. $L = 2$) and multi-level phase gratings (i.e. $L > 2$). Binary-phase gratings have the advantage of ease of fabrication. Multi-level phase gratings, although more efficient than binary-phase gratings, are much harder to fabricate. Section 2.6 discusses the fabrication of both binary- and multi-level phase gratings, respectively, in more detail. In this section the concepts developed in previous sections will be used to analyze a simple binary phase grating.

Consider the phase profile $\Phi(x)$ shown in Figure 2.6. For simplicity we use a continuous-space phase function, as opposed to a discrete-space phase function (i.e. $\Phi[n] = \Phi(n \frac{A}{N})$). This phase profile is binary with the two phase states being 0 or $\Phi_o$, where $\Phi_o \in [0, \pi]$. The duty cycle $D$ of the grating is defined as the fraction of the period in the binary state $\Phi_o$. As outlined in Section 2.2, the transmittance function $t'(x)$ of this grating is given by $\exp [j\Phi(x)]$. We now investigate the effects of the duty
cycle $D$ and the phase depth $\Phi_0$ on the angular spectrum of the grating.

We first express the transmission function as a convolution of the transmission function over a single period $\Lambda$ of the grating $t_\Lambda(x)$ given by

\[
t_\Lambda(x) = \begin{cases} 
1 & -\frac{\Lambda}{2} < x < -\frac{D\Lambda}{2} \\
\exp[j\Phi_0] & -\frac{D\Lambda}{2} \leq x \leq \frac{D\Lambda}{2} \\
1 & \frac{D\Lambda}{2} < x < \frac{\Lambda}{2} \\
0 & \text{else}
\end{cases}
\] 

(2.26)

and a periodic impulse train $S(x)$ given by

\[
S(x) = \sum_{k=-\infty}^{+\infty} \delta(x - \frac{D\Lambda}{2} - k\Lambda), \quad k = \text{integer}
\] 

(2.27)

The angular spectrum of the grating is then given by the product of the Fourier transforms of $t_\Lambda(x)$ and $S(x)$. 

Figure 2.6: Phase function of a binary grating.
The transform of $t_A(x)$, $T_A(f_x)$, is found to be

$$T_A(f_x) = \int_{-\infty}^{+\infty} t_A(x) \exp[-j2\pi f_x x] \, dx$$

$$= (\exp j\Phi_o - 1) \left[ D\Lambda \text{sinc} \left(2\pi f_x \frac{D\Lambda}{2}\right) \right] + \Lambda \text{sinc} \left(2\pi f_x \frac{\Lambda}{2}\right) \quad (2.28)$$

where $\text{sinc}(x) = \frac{\sin x}{x}$. The transform of $S(x)$, $S(f_x)$, is given by

$$S(f_x) = \frac{1}{\Lambda} \exp \left[j2\pi f_x \frac{D\Lambda}{2}\right] \sum_{k=-\infty}^{+\infty} \delta(f_x - \frac{k}{\Lambda}) \quad , k = \text{integer.} \quad (2.29)$$

Therefore, the angular spectrum $A(f_x)$ of the phase grating is given by the product of Eq. (2.28) and Eq. (2.29):

$$A(f_x) = \sum_{k=-\infty}^{+\infty} \exp \left[j\pi kD \right] \left[ D \left( \exp \left[j\Phi_o\right] - 1 \right) \text{sinc} \left(\pi kD\right) + \text{sinc} \left(k\pi\right) \right] \quad , f_x = \frac{k}{\Lambda} ; k = \text{integer}$$

$$= 0 \quad , \text{else.} \quad (2.30)$$

For a normally-incident plane wave $U(x, z)$ the transmitted field $U_t(x, z)$ is given by

$$U_t(x, 0) = \sum_{k=-\infty}^{+\infty} \{ \exp \left[j\pi kD \right] \left[ D \left( \exp \left[j\Phi_o\right] - 1 \right) \text{sinc} \left(\pi kD\right) + \text{sinc} \left(k\pi\right) \right] \} \exp \left[j2\pi \frac{k}{\Lambda} x \right]$$

$$= \sum_{k=-\infty}^{+\infty} a_k \exp \left[j2\pi \frac{k}{\Lambda} x \right] , k = \text{integer} \quad (2.31)$$

where $a_k$ is the amplitude of the $k^{\text{th}}$ diffraction order. It is useful to express $a_k$ as

$$a_k = [D \left( \exp \left[j\Phi_o\right] - 1 \right) + 1] + \sum_{k \neq 0} \exp \left[j\pi kD \right] D \left( \exp \left[j\Phi_o\right] - 1 \right) \text{sinc} \left(\pi kD\right) \quad (2.32)$$
which separates out the zero-order coefficient. The diffraction efficiency of all propagating orders is, as in Eq. (2.21), given by $|a_k|^2$. Note also that for all propagating orders $\eta_k = \eta_{-k}$. Therefore, the power spectrum of this grating is symmetric with respect to the zero-order component $\eta_0$, and, in fact, the power spectrum of all binary-phase gratings (in the scalar regime) are even-symmetric. We now consider how the diffraction efficiency into the various orders changes with the grating parameters $D$ and $\Phi_o$.

We consider the diffraction efficiency, $\eta_k$, for the two cases described above. First, we examine the case where $D = 0.50$, referred to as a 50\% duty cycle grating. We find $\eta_k$ versus the phase depth $\Phi_o \in [0, \pi]$. The result is used to illustrate the effect of etch depth on the angular spectrum of an arbitrary binary phase grating. Second, we consider the case where $\Phi_o = \pi$, referred to as a $\pi$-phase grating. We find $\eta_k$ versus the duty cycle $D \in [0, 100]$.

The power spectrum for the case $D = 0.50$ and $\Phi_o = \pi$ is plotted in Figure 2.7. Note that the angular spectrum of the grating contains only odd orders (i.e. $k$ odd), which is verified by Eq. (2.31). With the duty cycle fixed at 0.50, Figures 2.8a, b and c show the angular spectrum of the grating for $\Phi_o = \frac{3\pi}{4}$, $\Phi_o = \frac{\pi}{2}$ and $\Phi_o = \frac{\pi}{4}$, respectively. The power spectrum of the grating now not only contains odd orders but also the zero-order. It appears that as $\Phi_o$ decreases, the power in the zero-order increases, while the total power contained into the higher orders decreases. We show in the next section that for an arbitrary binary-phase grating the ratio of the diffraction efficiencies among the higher-orders remains constant as the phase depth is varied and that the sole effect of a change in phase depth is to change the ratios of the zero-order diffraction efficiency to the higher-order diffraction efficiencies.

Figure 2.9 plots the zero-, first-, and third-order diffraction efficiencies, $\eta_0$, $\eta_1$, and
Figure 2.7: Power spectrum of 50% duty cycle grating with $\pi$-phase depth.
Figure 2.8: Power spectrum of a 50% duty cycle grating for various phase depths $\Phi_o$:
(a) $\Phi_o = \frac{3\pi}{4}$; (b) $\Phi_o = \frac{\pi}{2}$.
Figure 2.8: Power spectrum of a 50% duty cycle grating for various phase depths $\Phi_o$: (c) $\Phi_o = \frac{\pi}{4}$.
Figure 2.9: Zero-, first- and third-order diffraction efficiencies for 50% duty cycle grating versus phase depth.
\( \eta_3 \), respectively, versus the phase depth \( \Phi_0 \) for \( \Phi \in [0, \pi] \). For \( \Phi_0 = 0 \) the grating has zero phase depth which effectively says that the grating is not there. So, as expected the grating angular spectrum contains only the zero-order component with a diffraction efficiency \( \eta_0 \) of 1, which corresponds to an undiffracted plane wave. \( \eta_0 \) varies continuously from 1 at \( \Phi_0 = 0 \) to 0 at \( \Phi_0 = \pi \) (see also Figure 2.7). From Eq. (2.31), the functional form of this variation is found to be \( \eta_0 = \cos^2 \frac{\Phi_0}{2} \). Figure 2.9 also shows that the ratio \( \frac{m}{n_3} \) remains constant.

We now consider the case with \( \Phi_0 = \pi \) and \( D \in [0, 100] \). Figures 2.10a - 2.10d plot the grating angular spectrum for \( D = 0.25 \), \( D = 0.40 \), \( D = 0.60 \) and \( D = 0.75 \), respectively. Note that the profiles for the 25% and 40% duty cycle gratings are equivalent to those for the 75% and 60% duty cycle gratings, respectively. In general, binary-phase gratings with duty cycles corresponding to \( D \) and \( (1 - D) \) have equivalent power spectrums. Mathematically, this fact is derived by calculating the diffraction efficiencies \( \eta_k \) from the \( a_k \)'s given in Eq. (2.32) for \( D \) and \( D' = (1 - D) \).

Furthermore, one can show that the power spectrums of arbitrary binary-phase gratings are independent of the parity of the grating phase profile. We mean by 'independent of parity of the grating phase profile' the following: Assume we are given a binary-phase grating, characterized by its 'high' and 'low' phase states, as described earlier. Now, if the 'high' phase states are switched to 'low' phase states and vice versa, the resultant phase grating is also binary. The contention is that the power spectrum of this resultant binary-phase grating is identical to that of the original binary-phase grating. Physically, this symmetry is a direct consequence of the scalar approximation and also the thin-grating approximation. Chapter 5 considers these approximations in more detail. Figure 2.11 plots the zero- and first-order diffraction efficiencies of the \( \pi \)-phase grating as a function of the duty cycle \( D \). As expected, the
Figure 2.10: Power spectrum of a π-phase grating for various duty cycles $D$: (a) $D = 0.25$; (b) $D = 0.40$. 
Figure 2.10: Power spectrum of a π-phase grating for various duty cycles $D$: (c) $D = 0.60$; (d) $D = 0.75$. 
dependencies are even-symmetric about a duty cycle of 50% \((D = 0.50)\).

In this section a simple binary-phase grating was analyzed using the concepts developed in earlier sections. The results were used to illustrate some general properties of the power spectrum of binary-phase gratings. Chapter 5 will revisit these properties in the regime where scalar diffraction theory breaks down. So far, this chapter has been devoted to the theoretical formalism of diffraction gratings. The next section briefly describes how one practically fabricates these grating structures.
2.6 Binary Optics Technology

Recall that the phases $\Phi[n]$ of a transmission function $t[n]$ can in general take on a continuum of values. Consequently, the surface-relief profile corresponding to the grating represented by $t[n]$ is extremely hard to fabricate. Since these grating structures were designed with the intent of incorporating them into parallel processing systems, a practical means of fabricating such surface-relief profiles is of utmost importance. In this section we describe a fabrication technology known as Binary Optics, which has emerged as an efficient and reliable means of fabricating diffractive optical elements, sometimes called binary optic elements (BOEs).

Consider the phase function $\Phi[n]$ shown in Figure 2.12. In general, each element of $\Phi[n]$ will have a unique phase value. In this situation the most straightforward way to fabricate the surface-relief profile is to perform $N$ lithography/etching steps. A lithography/etching step is shown in Figure 2.13. For typical values of $N(\geq 32)$ such a task is unreasonably tedious. Binary optics technology utilizes a stepwise approximation to the original phase function that significantly decreases the number of processing steps. We first present the algorithm for calculating the phase function of a BOE. Then, we describe the techniques for fabricating the corresponding surface-relief profile.

To calculate the phase function of a BOE, two steps are necessary. The first step is to generate a new phase function $\Phi'[n]$ (from the given phase function $\Phi[n]$) defined by the relation

$$\Phi'[n] = |\Phi[n]|_\pi$$  \hspace{1cm} (2.33)

where $|\Phi[n]|_\pi$ represents the function $\Phi[n]$ modulo $\pi$. Figure 2.14 plots the $\Phi'[n]$
Figure 2.12: Arbitrary phase profile $\Phi[n]$. 
Figure 2.13: Fabrication process: (a) Optical lithography; (b) Development of photoresist; (c) RIE; and (d) Removal of photoresist.
Figure 2.14: Phase profile $\Phi[n]$ modulo $\pi$. 
obtained from Eq. (2.33) for the $\Phi[n]$ shown in Figure 2.12. Mathematically, $\Phi[n]$ is equivalent to $\Phi'[n]$ since $\Phi[n] = \Phi[n] + 2l\pi$, $l = \text{integer}$; however, the corresponding surface-relief profiles are quite different.

The second step in the calculation is to approximate the function $\Phi'[n]$ by quantizing it into $2^M$ equally-spaced phase levels, where $M = 1, 2, 3, 4, \ldots$, resulting in a new phase function $\Phi_q'[n]$. Figure 2.15 shows the quantized phase function $\Phi_q'[n]$ corresponding to the $\Phi'[n]$ in Figure 2.14. The phase element corresponding to this quantized phase function is referred to as a binary optic element.

Therefore, the algorithm for calculating the phase profile of a BOE from an ar-
CHAPTER 2. DIFFRACTION GRATINGS: THEORY AND FABRICATION

Figure 2.16: Block diagram of algorithm for calculating the phase profile of a BOE.

<table>
<thead>
<tr>
<th>$M$</th>
<th>Diffraction efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.41</td>
</tr>
<tr>
<td>2</td>
<td>0.81</td>
</tr>
<tr>
<td>3</td>
<td>0.95</td>
</tr>
<tr>
<td>4</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 2.1: Diffraction efficiencies for phase quantization to $2^{M}$ levels.

An arbitrary phase profile is summarized by the block diagram in Figure 2.16. The phase function first passes through a modulo-$\pi$ operation and is then passed through a phase quantization function $Q_M$. An input to the phase quantization function is $M$, which determines the number of quantization levels. One can show that for $M = 4$ (i.e. 16 phase levels) the BOE is 99\% as efficient as the phase element corresponding to the phase function $\Phi'[n]$ [19] [30]. The efficiencies for 2, 4 and 8 quantization levels are provided in Table 2.1.

The minimum feature size $\delta x$ of the grating structure (assuming the representation given in Section 2.4) is given by $\frac{A}{N}$, where $A$ is the period of the grating and $N$ is the number of pixels within one grating period. As outlined in Section 2.4, in the design of diffraction gratings implementing large fanouts the fanout scales linearly with $N$. Consequently, the minimum feature size (for a fixed grating period, perhaps due to a fixed aperture constraint or to insure the validity of scalar diffraction theory) is inversely proportional to the degree of fanout. In addition, we cannot arbitrarily
CHAPTER 2. DIFFRACTION GRATINGS: THEORY AND FABRICATION  61
decrease \( N \) to increase the minimum feature size due to aliasing issues. As an example, consider an incident wavelength of 1\( \mu \)m (for a typical semiconductor laser) and a minimum feature size of 0.5\( \mu \)m. Again, if we assume that \( \Lambda = 10\lambda \), then the maximum number of pixels \( N \) that can be used to represent a diffraction grating is 20. Therefore, an estimate of the maximum degree of fanout for this grating is 10.

Once the desired phase function is calculated, one must design the lithographic masks for patterning and subsequent etching of the surface-relief structure. In general, for \( 2^M \) quantization levels \( M \) lithography/etching steps are necessary to fabricate the BOE. If multiple masks are used (i.e. \( M > 1 \)) an optical alignment is required between mask steps. For \( M = 1 \) the resultant element is a binary phase element, whereas \( M > 1 \) corresponds to a multi-level phase element. Binary phase elements are relatively simple to fabricate since they do not require alignment steps. Multi-level phase elements, which do require careful alignment steps, are much harder to fabricate; however, they also tend to be more efficient than binary phase elements. Therefore, a tradeoff exists between achieving high diffraction efficiency and ease of fabrication.

To design the lithographic masks, one must specify the locations of the phase transitions between steps. These phase transitions correspond to transitions between full opacity and full transparency on the corresponding lithographic mask. These transitions are obtained in a straightforward manner as follows.

The first mask step results in a binary approximation (2 phase-levels, and hence the name "binary optics") to the phase function \( \Phi'[n] \), whose elements take on a continuum of values in the range \([-\pi, \pi]\). Therefore, a binary approximation is made by setting the pixels whose phase is less than 0 to have a phase of 0, and by clipping the phase of pixels with phases greater than or equal to 0 to have a phase of \( \pi \). To
transfer this phase profile to a substrate, one first assigns regions defined to 0 phase to be opaque regions on the amplitude mask, and assigns regions defined to have a \( \pi \)-phase depth to be transparent regions on the amplitude mask. Using this mask the appropriate regions on the substrate can be defined, for subsequent reactive-ion etching, by a lithographic exposure.

The design of mask \#M, where \( M > 1 \), proceeds similarly. However, now the phase clipping occurs at integer multiples of \( \frac{\pi}{2M-1} \). For example, consider the design of mask \#2. Regions of the continuous phase that lie in the ranges \([-\pi, -1(\frac{\pi}{2})]\) and \([0, 1(\frac{\pi}{2})]\) are set to have a phase of 0; regions of the continuous phase that lie in the ranges \([-1(\frac{\pi}{2}), 0]\) and \([1(\frac{\pi}{2}), 2(\frac{\pi}{2})]\) are set to have a phase of \( \frac{\pi}{2} \). To transfer this phase profile to a substrate, one assigns regions defined to have 0 phase to be transparent on the lithographic mask as before, and assigns regions to have a \( \frac{\pi}{2} \)-phase depth to be opaque on the lithographic mask.

The third and fourth masks are designed in a similar fashion, except that now the opaque/transparent transitions on the amplitude masks occur at integer multiples of \( \frac{\pi}{4} \) and \( \frac{\pi}{8} \), respectively.

The etch depths for each mask step can be found from Eq. (2.15), which gives the etch depth \( d[n] \) as a function of the phase depth \( \tilde{\Phi}_q'[n] \). In general, the etch depth for the \( m^{th} \) mask step, \( d(m) \) is given by

\[
d(m) = \frac{\lambda_o}{2^m (n_1 - n_2)}
\]

where \( \lambda_o \) is the free-space wavelength of the incident light and \( n_1 \) and \( n_2 \) are the indices of refraction of the substrate and surrounding medium (typically free-space \( \Rightarrow n_2 = 1 \)), respectively. For a wavelength of 0.6328\( \mu \)m and a fused silica substrate \( (n_1 \approx 1.4570179) \) etch depths range from \( \sim 1.1 \mu \)m for \( m = 1 \) to \( \sim 0.14 \mu \)m = 1400\( \AA \)
for \( m = 4 \). The control of the etch depth is extremely important in the fabrication of optical interconnects. The next section investigates this issue further.

In this section the fabrication technology of Binary Optics was described. This technology can be used to fabricate arbitrary surface-relief phase gratings which implement optical interconnections. The fabrication procedure, borrowed from the semiconductor integrated circuit industry, utilizes optical lithography and reactive-ion etching to construct the surface-relief profile. Clearly, the performance of the fabricated phase grating depends on the reliability of the lithography and etching steps.

In the next section we consider two sources of error in the fabrication procedure described above: an etch depth error and a linewidth error. At Lincoln Laboratory surface-relief structures can be fabricated to an etch depth accuracy of \( \sim 300 \text{Å} \) [31], while linewidth errors are within \( \pm 0.1 \text{µm} \) [31]. It is important to investigate the behavior of grating structures as a function of these fabrication errors since such structures will ultimately be incorporated into real systems. We now consider the effect of these errors on grating power spectrums.

### 2.7 Fabrication Errors

In this section we consider the effects of two fabrication errors, etch depth errors and linewidth errors, on the power spectrum of binary-phase gratings. A similar analysis for arbitrary multi-level phase gratings is beyond the scope of this thesis. The development of improved models for simulating the effects of fabrication errors on the performance of multi-level phase elements is an area of intensive research. We focus on the simpler case of binary-phase gratings.

Etch depth and linewidth errors (among others) are invariably present in any fabrication process; therefore, it is important to investigate the behavior of surface-
relief structures as a function of these errors. Etch depth errors can occur, for example, from fluctuations in the etch rate and/or substrate ‘loading’ effects. Farn et al. [32] studied the effect of etch depth errors on the diffraction efficiency of multi-level diffractive lenses and concluded that the diffraction efficiency of the lens falls off parabolically as a function of the etch depth error. Cox et al. [33] also studied the effect of etch depth errors on the diffraction efficiency of microlenses. They showed that the efficiency of binary-phase microlenses are relatively insensitive to etch depth errors that are typically achieved in the laboratory (\(\sim\) \(\pm\) 5%). Diffractive microlenses, unlike optical interconnects, are conventionally designed to be efficient in a single diffraction order. For optical interconnect applications, however, one is concerned with the diffraction efficiencies of many diffraction orders. Turunen et al. [17] and Fagerholm et al. [7] investigated the effect of etch depth errors on the power spectrum of on-axis and off-axis kinoform interconnects. However, their results are not easily interpreted.

We provide a simple approach to analyze the consequence of etch depth errors on the power spectrum of binary-phase gratings and show that as the optical fanout increases so does the sensitivity of the power spectrum to the etch depth. In addition, we show that for binary-phase gratings an etch depth error only increases/decreases the power contained in the zeroth-order, while leaving the relative amounts of power contained in the higher-orders unchanged. The corresponding changes in the diffraction efficiencies of the higher-orders are determined from power conservation. Therefore, the effect of an etch depth error on the power spectrum of a binary-phase grating can be summarized by the change in the zero-order diffraction efficiency.

Linewidth errors can occur for several reasons: if intimate contact between the mask and substrate is not achieved in the lithographic exposure; if the reactive-ion-
etching process is not anisotropic; over-/under-exposure and uneven photoresist step coverage. Furthermore, it has been observed that the linewidth error is, in general, a function of both the physical dimension of each line and also the surrounding surface-relief patterns on the substrate [31]. Linewidth errors affect the grating power spectrum in a more complicated manner than do etch depth errors. In particular, the diffraction efficiencies of all orders are affected independently. A complete analysis of these effects is well-beyond the scope of this thesis. We do, however, consider the effect of a linewidth error (making suitable simplifications) on the zero-order diffraction efficiency of a binary-phase grating. Unlike the case for etch depth errors, the diffraction efficiencies of the higher-orders cannot be determined from a knowledge of the change in the zero-order efficiency. An accurate model accounting for the above phenomena is an area of ongoing research.

A 1D binary-phase grating is represented as an $N$-element array of phases $\Phi[n] = \phi_n$, where $n \in [0, (N - 1)]$ and the $\phi_n$ take on one of two values, 0 or $\Phi_0$. For reasons that will become clear, we express $\Phi[n]$ as

$$\Phi[n] = (\Phi[n] - \frac{\Phi_0}{2}) + \frac{\Phi_0}{2}$$

$$\equiv \Phi'[n] + \frac{\Phi_0}{2} \quad (2.35)$$

where the function $\Phi'[n]$ takes on one of two values: $-\frac{\Phi_0}{2}$ or $\frac{\Phi_0}{2}$. The corresponding transmittance function $t[n]$ of the grating is given by

$$t[n] = \exp \left( j \frac{\Phi_0}{2} \right) \left[ \cos \left( \Phi'[n] \right) + j \sin \left( \Phi'[n] \right) \right] \quad (2.36)$$

where the Euler relation for exponentials has been used. Since each element of $\Phi'[n]$ is either $\pm \frac{\Phi_0}{2}$, we can express Eq. (2.36) as
\[ t[n] = \exp\left[j\frac{\Phi_0}{2}\right] \left[ \cos\left(\frac{\Phi_0}{2} a[n]\right) + j \sin\left(\frac{\Phi_0}{2} a[n]\right) \right] \]
\[ = \exp\left[j\frac{\Phi_0}{2}\right] \left[ \cos\left(\frac{\Phi_0}{2}\right) + ja[n] \sin\left(\frac{\Phi_0}{2}\right) \right] \] (2.37)

where the elements of \( a[n] \) are \( \pm 1 \) and the evenness and oddness of the cosine and sine functions, respectively, have been used. The grating angular spectrum \( A[m] \) is given by the DFT of Eq. (2.37). After some algebraic manipulation the angular spectrum is found to be

\[ A[m] = \left[ D \exp\left[j\Phi_0\right] + (1 - D) \right] + \frac{1}{2N} \left( \exp\left[j\Phi_0\right] - 1 \right) \sum_{n=0}^{N-1} a[n] \exp\left[j\frac{2\pi m n}{N}\right] \] (2.38)

where \( D \) is the grating duty cycle and the zero-order coefficient \( a_0 \) has been separated out (i.e. summation defined for \( m \neq 0 \)). The above equation defines the \( a_m \) coefficients in the angular plane wave decomposition of a binary-phase grating of arbitrary duty cycle and phase depth. The corresponding diffraction efficiencies \( \eta_m \) are given by the magnitude-squared of the \( a_m \) coefficients. Our goal is to describe the effects of etch depth and linewidth errors on the power spectrum of a binary-phase grating.

From Eq. (2.38) we can see that although the diffraction efficiencies change as a function of phase depth (or, equivalently, the etch depth) the ratios of diffraction efficiencies among the higher-order coefficients (i.e. \( \frac{\eta_k}{\eta_l} \), where \( k \neq 0, l \neq 0 \) and \( k \neq l \)) are constant. Since these ratios are constant we can describe the effect of an etch depth error by the change in the zero-order diffraction efficiency \( |a_0|^2 \). The change in the diffraction efficiencies of the higher-orders can be found from power conservation.

A linewidth error affects the grating power spectrum in a more complicated manner. We distinguish between a positive and negative linewidth error as follows. A
positive linewidth error results in a linewidth which is smaller than the desired value, while a negative linewidth error results in a linewidth which is larger than the target value.

Linewidth errors in general do not redistribute the power in the higher-orders such that their ratios remain constant which was the case for etch depth errors. Consider the 50% duty cycle grating discussed in Section 2.4. For this grating a linewidth error corresponds to a change in the duty cycle \( D \). Referring to Eq. (2.32), we can see that the ratio of diffraction efficiencies, \( \frac{a_k}{a_l} \) where \( k \neq 0, l \neq 0 \) and \( k \neq l \), is a function of the duty cycle \( D \). An arbitrary binary-phase grating also has this property. Therefore, if the relative amount of power contained among the higher-orders is an important consideration, linewidth errors have more deleterious effects than do etch depth errors. We now derive an expression for the effect of both etch depth errors and linewidth errors on the zero-order diffraction efficiency of a binary-phase grating.

The magnitude of the zero-order coefficient \( a_0 \) is given by

\[
| a_0 | = | D \exp [j \Phi_0] + (1 - D) | \\
= \sqrt{D \cos (\Phi_0) + (1 - D))^2 + (D \sin (\Phi_0))^2}. \tag{2.39}
\]

Clearly, \( \eta_0 \equiv | a_0 |^2 \) is a function of both the phase depth and duty cycles of the grating. The differential of \( | a_0 |^2 \) is given by (to first-order)

\[
d | a_0 |^2 = \frac{\partial}{\partial \Phi_0} (| a_0 |^2) d\Phi_0 + \frac{\partial}{\partial D} (| a_0 |^2) dD \tag{2.40}
\]

where

\[
\frac{\partial}{\partial \Phi_0} (| a_0 |^2) = 2D(D - 1) \sin (\Phi_0) \tag{2.41}
\]
and

$$\frac{\partial}{\partial D}(\mid a_0 \mid^2) = 2(2D - 1)(1 - \cos(\Phi_0)) \quad (2.42)$$

The first term in Eq. (2.40) corresponds to the change in $\mid a_0 \mid^2$ due to an etch depth error for a fixed duty cycle $D$. Similarly, the second term in Eq. (2.40) corresponds to the change in $\mid a_0 \mid^2$ due to a linewidth error, which can be related to the grating duty cycle $D$, for a fixed phase depth $\Phi_0$.

Some qualitative statements can be made from Eq. (2.41) and Eq. (2.42). First, Eq. (2.41) is minimized at $\Phi_0 = \pi$. For binary-phase gratings, $\Phi_0$ is restricted to the range of values $[0, \pi]$. Furthermore, the duty cycle $D$ is always $< 1$. These two facts imply (for a fixed duty cycle $D$) that i) an overetch (i.e. $d\Phi_0 > 0$) decreases the zero-order diffraction efficiency and ii) an underetch (i.e. $d\Phi_0 < 0$) increases the zero-order diffraction efficiency. Second, Eq. (2.42) is minimized for a value of $D = 0.50$. This result is consistent with Figure 2.11 which shows that the zero-order diffraction efficiency is minimized at a duty cycle of 0.50. So, for a fixed phase depth $\Phi_0 \epsilon [0, \pi]$ i) a positive linewidth error (i.e. $dD < 0$) decreases the zero-order diffraction efficiency if the duty cycle is greater than 50% and increases the zero-order diffraction efficiency if the duty cycle is less than 50%. and ii) a negative linewidth error (i.e. $dD > 0$) increases the zero-order diffraction efficiency if the duty cycle is greater than 50% and decreases the zero-order diffraction efficiency if the duty cycle is less than 50%.

These results are useful for a qualitative analysis of binary-phase gratings.

The relation between the linewidth error and duty cycle is now described. Consider the grating surface-relief profile shown in Figure 2.17a. The linewidths of interest are those corresponding to the 'high' phase states (i.e. pixels where $\Phi[n] = \Phi_0$). Suppose the fabrication process results in a positive linewidth error of 10%. The resultant
Figure 2.17: Phase profiles of an arbitrary binary-phase element: (a) Desired profile; (b) Profile due to a linewidth error.
grating for this situation is shown in Figure 2.17b. The new grating duty cycle $D'$ is given by $0.90 ND$. Note that for a negative linewidth error of $-\delta$ (i.e. a percentage change of $\delta \times 100\%$) the new grating duty cycle is given by $D' = (1 + \delta) D$. For spatially-varying linewidth errors the new grating duty cycle can be determined by summing over all 'lines', where for each 'line' the type of error (i.e. positive or negative) is considered.

In the above framework Eqs. (2.40) - (2.42) can be used to determine the change in the zero-order diffraction efficiency $\eta_0$ for known etch depth and linewidth errors. As an example, we consider once again the 50% duty cycle grating with a phase depth $\Phi_0 = \frac{\pi}{2}$ radians. First, we assume that the only fabrication error is a phase depth error of $d\Phi_0 = 0.10$ radians. For a wavelength of $1\mu$m the corresponding etch depth error (overetch) is $\simeq +160\AA$. One can generally etch to an accuracy of $\sim \pm 300\AA$ [31]. From Eq. (2.40) and Eq. (2.41) the change in $|a_0|^2$ is found to be $-0.05$. The ideal zero-order diffraction efficiency can be found from Eq. (2.32): $\eta_0^{\text{ideal}} = 0.50$. Therefore, the percentage change in the the zero-order diffraction efficiency is $-10.0\%$.

Given the change in the zero-order diffraction efficiency and the fact that the ratio of diffraction efficiencies among the higher-orders remains invariant as a function of the grating phase depth, the $\eta_k$ for $k \neq 0$ are determined from power conservation arguments.

We next consider the case where a linewidth error is the sole source of error. For a 50% duty cycle grating the first-order approximation given by Eq. (2.42) predicts that the zero-order diffraction efficiency is not affected by a linewidth error. Therefore, to obtain an estimate of the affect of a linewidth error on the zero-order diffraction efficiency of a 50% duty cycle grating, we assume a linewidth error such that the actual grating duty cycle is 0.48 rather than 0.49. The zero-order diffraction efficiency of a
49% duty cycle grating can be found from Eq. (2.32) and is 0.5002. Once again, we assume a phase depth of $\frac{\pi}{2}$ radians. For a grating period of 20$\mu$m the error in linewidth is -0.1$\mu$m. Practically, one can control the linewidths to within $\pm 0.5\mu$m [31]. From Eq. (2.40) and Eq. (2.42) the change in the zero-order diffraction efficiency is found to be -0.0004, corresponding to a change of approximately $-0.08\%$ change. The two cases above suggest that for present fabrication capabilities (i.e. etch depth control of $\sim \pm 300\AA$ and linewidth control of $\sim \pm 0.5\mu$m) etch depth errors have a more pronounced effect on the zero-order diffraction efficiency than do linewidth errors.

For the case where both of the above errors are present, the change in the zero-order diffraction efficiency is given by the sum of the changes due to each source of error, as expressed by Eq. (2.40). This change is -0.0504, corresponding to a $-10.08\%$ change.

The above formalism can be used to analyze an arbitrary binary-phase grating of duty cycle $D$ and phase depth $\Phi_0$. In designing binary-phase gratings for implementing optical interconnections the duty cycle $D$ (more precisely, the binary states of the grating) and phase depth $\Phi_0$ are optimized such that the grating power spectrum approximates a desired power spectrum. The topic of phase optimization is discussed at length in the next chapter.

For now we assume that an optimized grating phase profile with duty cycle $D$ and phase depth $\Phi_0$ has been determined which implements a 1-to-$N$ fanout ($N$ odd) with a uniform splitting ratio. The zero-order diffraction efficiency of this optimized grating is approximately $\frac{1}{N}$. For simplicity we assume that the only source of error is in etch depth. For a phase depth error of $d\Phi_0$ the fractional change in the zero-order diffraction efficiency $d |a_0|^2$ is given by
\[ d \mid a_0 \mid^2 = 2ND(D - 1) \sin(\Phi_0) d\Phi_0. \] (2.43)

For \( d\Phi_0 = 0.10 \) radians an order of magnitude estimate of \( (d\mid a_0 \mid^2 \) is \( \sim 5N \). For large fanouts (i.e. \( N \) large) the zero-order diffraction efficiency becomes extremely sensitive to etch depth errors. We note, however, that the diffraction efficiencies of the higher-orders are not as sensitive to the etch depth as is the zero-order. This decreased sensitivity is attributed to two reasons: 1) the increase/decrease in the zero-order power must be reflected in a corresponding decrease/increase in the total power contained in the higher-orders (i.e. power conservation) and 2) the changes in the diffraction efficiencies of the higher-orders are such that the ratio of their powers remains constant, as discussed earlier. Given these facts the percentage changes in the diffraction efficiencies of all the higher-orders (i.e. \( m \neq 0 \) and \( m = 1, 2, ..., (N - 1) \) are equal and are given by

\[ d \mid a_m \mid^2 = \frac{-2ND(D - 1)}{(N - 1)} \sin(\Phi_0) d\Phi_0. \] (2.44)

The sensitivity of the higher-orders is less by a factor of \((N - 1)\) than that of the zero-order. Unlike the case for the zero-order diffraction efficiency the sensitivity of the higher-order diffraction efficiencies decreases as the fanout (i.e. \( N \)) increases. Equation (2.44) holds for a binary-phase grating which implements a 1-to-\( N \) fanout with a uniform splitting ratio.

For a binary-phase grating which implements a 1-to-\( N \) fanout with a nonuniform splitting ratio, the percentage change in the zero-order diffraction efficiency is given by
\[ d \left| a_0 \right|^2 = \frac{2D(D - 1) \sin(\Phi_0) d\Phi_0}{\left| a_0 \right|^2}. \quad (2.45) \]

For \( \left| a_0 \right|^2 = \frac{1}{N} \) (i.e. \textit{1-to-N} fanout with a uniform splitting ratio), Eq. (2.45) reduces to Eq. (2.43). The sensitivities of the higher-order diffraction efficiencies can be divided into two cases: if the diffraction efficiency of the \( n^{th} \) is greater/less than \( \frac{1}{N} \). For the former case the sensitivity is less than that estimated by Eq. (2.44), and for the latter case the sensitivity is greater than that estimated by Eq. (2.44). These qualitative statements follow directly from power conservation and the fact that the power ratios between the higher-orders remain constant with a variation in etch depth. These conditions are expressed formally as

\[ \Delta \eta_0 = \sum_{k=-K}^{K} \Delta \eta_k \quad (2.46) \]

where the sum excludes \( k = 0 \) and \( K \) refers to maximum order of interest (i.e. \( K = (N + 1)/2 \)) and

\[ \frac{\eta_l + \Delta \eta_l}{\eta_m + \Delta \eta_m} = \frac{\eta_l}{\eta_m} \quad (2.47) \]

where \( l \neq m, l \neq 0 \) and \( m \neq 0 \). Given the original power spectrum \((\eta_k, k \in [-K, K])\) and the change in the zero-order diffraction efficiency \((\Delta \eta_0 \) given by Eq. (2.40)), the fractional changes in the higher-order diffraction efficiencies \((\Delta \eta_l, l \in [-K, K] \) and \( l \neq 0 \) are expressed, after some algebraic manipulation, by

\[ \Delta \eta_l = \frac{-\Delta \eta_0}{\sum_{k=-K}^{K} \eta_k} \quad (2.48) \]

where, as in Eq. (2.46), the sum excludes \( k = 0 \). Note that for a uniform splitting
ratio,
\[ \sum_{k=-K}^{K} \eta_k = \frac{N - 1}{N} \]
and Eq. (2.48) reduces to Eq. (2.44).

(Note: The gratings which implement the above power spectrums are physically unrealizable since the spectrums are bandlimited. However, real gratings can be designed which closely approximate such power spectrums. The methods used to design such gratings are discussed in Chapter 3.)

Finally, we note that as the fanout increases, the sensitivity of the grating power spectrum to linewidth errors also increases; however, for linewidth errors one also must be concerned with the redistribution of power into the higher-orders of the grating since in general the condition given by Eq. (2.47) is no longer satisfied.

In this section the effects of etch depth errors and linewidth errors on the performance of binary-phase gratings was examined. A first-order approximation was used to describe each of these effects. It was shown that as the degree of fanout increases so does the sensitivity of the phase grating to the above fabrication errors. Furthermore, fabrication process control will be a limiting factor in practically realizing optical interconnects with large fanouts.
Chapter 3

Phase Optimization Techniques

Diffraction gratings offer a convenient means of implementing free-space optical interconnections. In the scalar limit the grating transmittance function and grating angular spectrum are related through a Fourier transform relation. If the amplitude and phase are specified in one domain, then the amplitude and phase in the reciprocal domain are uniquely specified. For free-space optical interconnect applications a design constraint is imposed on the far-field power spectrum of the grating. For this reason the far-field phase distribution can be arbitrary. Given this degree of freedom a transmittance function can be found that exactly results in the specified power spectrum. However, this transmittance function has both amplitude and phase components. For efficiency considerations an additional design constraint that the transmittance function be phase-only (i.e. for a phase grating) is imposed. In this framework an exact solution for a grating transmittance function $T[m, n]$ or, equivalently, a grating phase function $\Phi[m, n]$ does not exist. Consequently, optimization techniques must be used.

The phase optimization problem is formulated by considering the computational representation of phase gratings, as described in Section 2.4. A two-dimensional phase grating is represented as an $M \times N$ array $\Phi[m, n]$ of phases $\phi_{mn}$.
\[ \Phi[m, n] = \phi_{mn}, \quad m = 0, 1, 2, ..., (M - 1) \]
\[ n = 0, 1, 2, ..., (N - 1) \]  

(3.1)

where \( \phi_{mn} \) are the phases of each cell of the array (or pixel of the physical element).

In general, the range of values for these phases form a continuum from 0 on up. However, as discussed in Section 2.6, the resultant phase profiles are exceedingly hard to fabricate. Therefore, we concentrate on the phase optimization of binary optic elements whose phases \( \phi_{mn} \) take on one of \( 2^K \) values: \( \frac{2\pi k}{2^K} \), where \( k = 0, 1, 2, ..., (2^K - 1) \). The optimization problem, therefore, involves determining the phases \( \phi_{mn} \) such that the resultant power spectrum of the grating closely approximates the ideal or desired power spectrum. To quantify the "goodness" of a given set of phases \( \phi_{mn} \) or, equivalently, how well the grating power spectrum approximates the ideal power spectrum, one defines a cost function or merit function \( C \), the specific form of which is described shortly. This cost function depends on the parameters to be optimized which in the phase optimization problem are the phases \( \phi_{mn} \). Associated with each set of phases \( \phi_{mn} \) is a cost. The optimization problem amounts to minimizing the cost \( C \) as a function of the phases \( \phi_{mn} \).

In this chapter the phase optimization techniques of simulated annealing [2] and phase retrieval [3] [4] [5] are introduced. The former technique is well-suited for the design of binary-phase gratings, whereas the latter technique is better adapted for the design of multi-level phase elements. Section 3.1 begins with a brief description of the annealing of solids, the concept from which the simulated annealing algorithm was developed. The simulated annealing algorithm is then considered and the calculated results using this algorithm for the design of two Gaussian fanout elements.
are presented. Section 3.2 begins with a description of the phase retrieval algorithm followed by a presentation of the calculated results using this algorithm.

3.1 Simulated Annealing

In condensed matter physics, annealing denotes a physical process in which an amorphous solid is melted and then slowly cooled such that the atoms of the solid arrange themselves in a crystal lattice. In the amorphous state the energy of the system is high. Conversely, the ground state of the system of particles is associated with the crystalline form. Annealing serves to transform the system of particles from a high-energy state to the ground-state of the system through a process of melting and slow cooling. A schematic of the annealing process is shown in Figure 3.1.

In the liquid state the atoms of the system are in rapid thermal motion. Each configuration of particles corresponds to a different energy of the system. The cooling rate is slow enough such that the system is never far from equilibrium at each temperature (i.e. quasistatic conditions). Consequently, the distribution of energy states as a function of temperature is proportional to the Boltzmann factor \(\exp[-E/k_BT]\), where \(E\) is the energy of the system, \(k_B\) is the Boltzmann constant and \(T\) is the temperature of the system. The Boltzmann factor expresses the probability that a state with energy \(E\) is occupied at a temperature \(T\). As the system is cooled, the thermal motion of the atoms decreases and as \(T \to 0\) the atoms arrange themselves in a lattice, a configuration corresponding to the ground state of the system, as shown in Figure 3.1. From the Boltzmann factor one can qualitatively see that as \(T \to 0\) only those states with low energies will have a significant probability of being occupied.

In 1983 Kirkpatrick et al. [2] realized that a profound analogy existed between the annealing process and the optimization of a function of many parameters. He
ENERGY OF SYSTEM = $E(f_1, f_2, ..., f_n)$

- At a fixed temperature, the distribution of states as a function of energy $= e^{-\frac{E}{kT}}$
- As $T \rightarrow 0$,
  - Thermal motion decreases
  - Atoms assume a lattice arrangement characteristic of the lowest energy state ($E = 0$)

Figure 3.1: Schematic representation of the annealing process
developed this analogy into the concept of simulated annealing. Analogous to the energy of the many-particle system, he defined a cost function $C$. This cost function, which depends on the various system parameters to be optimized, gives a quantitative measure of the "goodness" of a given state of the complex system to be optimized. The optimization problem amounts to minimizing the cost as a function of the parameters of the system.

The traditional approach used to solve the above problem, known as iterative-improvement [2], proceeds as follows: Given an initial state of the system with a corresponding cost $C_0$ a perturbation is applied to the system by varying one of its parameters. This perturbation changes both the state and the cost of the system. If the change in cost $\Delta C = C' - C_0$ is less than zero, then the state change is accepted.
CHAPTER 3. PHASE OPTIMIZATION TECHNIQUES

Otherwise, the state change is rejected. Therefore, only those state changes which lower the cost are accepted. It is clear that this technique is only capable of finding the local minima of the complex system. To illustrate this point, consider Figure 3.2 which plots the cost versus state space for an arbitrary system.

In general, the cost function can contain several maxima and minima. Our objective is to identify the global minimum at point E. Surrounding the global minimum are local minima (i.e. at points B and F). Suppose the initial state of the system is at point C. Given that we accept only those state changes that lower the cost, the algorithm converges to the minimum at point B. Clearly, this is a local minimum of the system. If the system were initialized at point A, the algorithm would also converge to point B. If the system were initialized at point D, then the algorithm would in fact converge to the global minimum at point E. However, since we have no a priori knowledge of how the cost varies as a function of state, we have no means of insuring that the system is always initialized in a state which lies in the valley of the global minimum. Therefore, the iterative-improvement algorithm typically converges to the local minima of a given system. Furthermore, the particular local minima that the algorithm converges to depends solely on the initial configuration of the system.

Kirkpatrick developed the simulated annealing algorithm as a modification of the iterative-improvement algorithm. He introduced the concept of an 'effective' temperature and utilized the so-called Metropolis criterion [34] for state selection. These modifications provide a means of "overcoming" the barriers between the various local minima and the global minimum. The simulated annealing algorithm generates states in the following way:

The goal is to minimize a cost function \( C \) describing a complex system. We assume this cost function depends on \( N \) independent variables. Given an initial
state of the system we then, as in the iterative-improvement algorithm, perturb the system by varying one of its parameters. This perturbation changes both the state of the system and the corresponding cost. If the change in cost $\Delta C \leq 0$, then the state change is accepted unconditionally. If, however, the perturbation results in a cost increase, we do not simply reject the state change, as in the iterative-improvement algorithm. Rather, we accept the state change with a probability given by $\exp[-\Delta C/T]$, where $T$ is an 'effective' temperature with the same units as cost. This exponential factor is analogous to the Boltzmann factor used in the analysis of the annealing of solids. At a non-zero temperature a state change that increases the cost has a finite probability of being accepted. At high 'effective' temperatures all state changes have a high probability of being accepted. This fact is used in determining the 'melting' temperature of the system, which is described below. As the temperature decreases, the probability of accepting state changes which increase the cost decreases. At $T = 0$ the exponential factor is zero, and the acceptance criteria reduces to that for the iterative-improvement algorithm.

The above procedure is repeated by perturbing each of the remaining $N$ variables of the cost function and determining the resultant states. One iteration of the algorithm is completed when all $N$ variables have been perturbed.

In analogy to the annealing process the 'effective' temperature must be chosen such that the complex system is 'melted'. The 'melting' temperature is chosen such that a large percentage ($\geq 90\%$) of the state changes through one iteration are accepted. The 'melting' of the system, therefore, makes the outcome of the optimization problem independent of the initial state of the system, which was a major drawback of the iterative-improvement algorithm.

Once the system is 'melted' the cooling process begins. In cooling the system the
temperature is lowered by a constant ratio from one iteration to the next. Therefore, the cooling schedule specifies an initial temperature $T_0$ (i.e. the melting temperature) and a rate $\frac{T_1}{T_0}$ by which the system is cooled. These two parameters determine the temperature during the $p^{th}$ iteration according to the equation

$$T_p = \left(\frac{T_1}{T_0}\right)^p T_0.$$  \hspace{1cm} (3.2)

The rate of cooling must be slow enough such that the system is not 'frozen' in a metastable state. In annealing a solid one is in a sense minimizing the energy of the solid (i.e. solving an optimization problem). Practically, however, one can determine whether the global minimum (i.e. the crystalline lattice configuration) has been reached or not by simply observing the crystal structure of the annealed solid; consequently, one can experimentally determine an optimum cooling rate. However, in solving combinatorial optimization problems one generally has no knowledge of how the cost varies as a function of the various parameters of the system. Therefore, the cooling rate used in a simulated annealing algorithm is often determined empirically. Typical cooling rates are $\geq 0.9$ \cite{2} \cite{35} \cite{36}.

For a cooling rate arbitrarily close to 1 (assuming a properly 'melted' system), one can theoretically show that the simulated annealing algorithm converges to the global minimum with a probability approaching 1 \cite{37}; however, for such a cooling rate the simulated annealing algorithm does not converge in a finite amount of time. Practically, a non-unity cooling rate must be used; consequently, one can never be sure that the solution of a simulated annealing algorithm is in fact the global minimum of the system. However, the algorithm does provide a systematic means of evaluating the relative merits of a large number of possible states of a complex system and in this regard it has emerged as a powerful technique for dealing with a wide variety of
CHAPTER 3. PHASE OPTIMIZATION TECHNIQUES

combinatorial optimization problems. Practically, one performs several runs of the algorithm for each problem and takes the best result among these runs.

The formalism of simulated annealing can be applied to a wide variety of optimization problems involving the minimization (or maximization) of a function of many independent parameters. Simulated annealing has been used in the optimal design of computers and subsystems [2], digital filter design [38], biology [39] and materials science [40].

In this work we utilize the simulated annealing algorithm to design binary-phase gratings for free-space optical interconnect applications. Our design objective is to obtain a 1-to-121 fanout (i.e. an 11-by-11 array of spots) in two dimensions and a Gaussian splitting ratio. Two approaches were taken to meet this design objective. The first approach involved solving for an optimum 1D binary-phase grating to achieve a 1-to-11 1D fanout. The desired 2D fanout can then be implemented by orthogonally crossing two 1D gratings to form a separable 2D grating. The second approach involved the optimization of a 2D binary-phase grating profile. The grating designed using the former approach requires two lithography/etching steps in addition to an optical alignment step and results in a 3-level phase element, while the grating designed using the latter approach only requires a single lithography/etching step and results in a 2-level phase element. On the other hand, the former approach is less computationally-intensive than the latter. The simulated annealing algorithms used in each of these design approaches and the results of these implementations are now presented.
3.1.1 Crossed 1D Grating

The most general two-dimensional (in the \(x-\) and \(y-\)directions) phase grating is mathematically expressed by Eq. (3.1). A restricted class of 2D phase gratings, known as crossed gratings, satisfies the condition that

\[
\phi_{mn} = \phi_m + \phi_n \tag{3.3}
\]

for all \(m\) and \(n\). The transmittance function of a crossed grating is then given by

\[
T[m, n] = (\exp [j\phi_m]) (\exp [j\phi_n]) \tag{3.4}
\]

which expresses the 2D transmittance function as the product of two 1D transmittance functions. The grating power spectrum \(P[l, q]\) is given by the 2D DFT of Eq. (3.4):

\[
P[l, q] = P[l] P[q] \tag{3.5}
\]

where \(P[l]\) and \(P[q]\) are the 1D DFT's of the respective 1D transmittance functions in Eq. (3.4). If, by design, we impose the conditions \(M = N\) (resulting in a square grating), then \(N\) phases fully specify the grating. By imposing the above conditions, the 2D problem is transformed into a simpler 1D problem, which often reduces the computational requirements of the design. We now describe the design and optimization of a 1D binary-phase grating that implements the desired 1D Gaussian fanout.

We represent the 1D grating by a \(N\)-element array \(\Phi[n]\), where \(n = 0, 1, \ldots, N-1\). Since the grating is to have binary-phase, each element of \(\Phi[n]\) can take on one of two values: 0 or \(\theta_0\), where \(\theta_0 \in [0, \pi]\). Given a set of phases \(\Phi[n]\) the grating angular
Figure 3.3: Ideal 1D power spectrum.
spectrum $A[m]$, where $m = 0, 1, ..., (N - 1)$, is given by the DFT of the grating transmittance function. The grating power spectrum $P[m]$ is the magnitude-squared of the grating angular spectrum. The ideal or desired power spectrum $P_{\text{ideal}}[m]$, as mentioned earlier, falls off in a Gaussian fashion. Figure 3.3 plots the ideal power distribution versus diffraction order number.

The cost function $C$ used in the simulated annealing algorithm is given by

$$C(\phi_n) = \sum_{m=\frac{N}{2}-K}^{\frac{N}{2}+K} (P[m] - P_{\text{ideal}}[m])^2 - (P[N/2] - P_{\text{ideal}}[N/2])^2$$

(3.6)

where $\phi_n$ represents the set of phases $\Phi[n]$, $N$ is the array size and $K$ represents the maximum order of interest on either side of the zero-order beam. For this particular grating $N = 128$ and $K = 5$. The first term represents the squared-error between the grating power spectrum and the ideal power spectrum within the region of interest. The second term represents the squared-error between the power contained in the zero-order of the grating and the ideal amount of power that should be contained in the zero-order. Therefore, the cost function is the squared-error between the grating power spectrum and the ideal power spectrum for all orders except the zero-order. The reason for excluding the zero-order contribution can be understood as follows:

A binary-phase grating is specified by i) the binary phase state of each phase cell (i.e. 0 for a phase of 0, 1 for a phase of $\theta_0$) and ii) the phase depth $\theta_0$ of the grating. We have shown (in Section 2.7) that for binary-phase gratings the sole effect of a change in phase depth is to change the relative power between the zero-order and the higher-orders. Also, the ratio of powers between the higher-orders is invariant with a change in phase depth. Therefore, deviations away from the ideal zero-order power can be corrected for by properly adjusting the phase depth of the binary-phase grating. The procedure above optimizes the binary state of each cell of the array $\Phi[n]$.
or said another way the procedure optimizes the power spectrum of the higher-orders.

After the binary state of the phase grating has been optimized, the zero-order power is optimized by minimizing the cost function $C_{\theta_0}$, which includes the zero-order contribution to the squared-error, given by

$$C_{\theta_0}(\theta_0) = \sum_{m=\frac{N}{2} - M}^{\frac{N}{2} + M} (P[m] - P_{\text{ideal}}[m])^2$$

as a function of the phase depth $\theta_0 \in [0, \pi]$. Given these cost functions the optimization problem can be clearly stated as a minimization of $C_{\theta_0}(\phi_n)$, which optimizes the binary states of the grating, and a minimization of $C_{\theta_0}(\theta_0)$, which optimizes the phase depth of the grating.

At the beginning of the algorithm the array $\Phi[n]$, which represents a single period of the grating, is assigned a random distribution of phases $0$ or $\pi$, representing the binary states of $0$ or $1$, respectively. The cost $C_0$ associated with this initial grating profile is then calculated. Next, the 'melting' temperature of the system is determined. The system is 'melted' when $\geq 90\%$ of the phase changes through one iteration are accepted. Initially, a guess $T_0$ of the melting temperature is made.

An iteration of the algorithm proceeds as follows: One starts with the first phase cell $\Phi[0]$. The phase state of this cell is flipped (i.e. $0$ to $1$, $1$ to $0$). The cost $C$ of this new grating profile is calculated using (3.6). Next, the difference in cost $\Delta C = C' - C_0$ is calculated. If $\Delta C \leq 0$, then the new phase state is accepted unconditionally. If, however, $\Delta C > 0$, then the new phase state is accepted with probability $p = \exp[-\Delta C/T_0]$. A random number $r$ is generated and compared with $p$. If $r \geq p$, then the new phase state is accepted. Otherwise, the new phase state is rejected, and the old state is restored. This procedure is repeated with the second ($\Phi[1]$) and third ($\Phi[2]$) on up to the $N^{th}$ ($\Phi[N-1]$) phase cell. As the phase cells are scanned
a record of the number of accepted phase changes is kept. This record is referred to in determining the 'melting' temperature and also in determining an appropriate termination point for the algorithm. An iteration is completed after all $N$ phase cells have been perturbed.

The above procedure becomes very computationally-intensive as the array dimensions increase. Since the array dimension limits the degree of fanout (see Section 2.4), the above fact essentially limits the simulated annealing algorithm to the design of optical interconnects with relatively small fanouts ($\leq 100 - by - 100$ array of spots).

If, through this first iteration at a temperature $T_0$, less than 90% of the phase changes were accepted, then the new guess of the 'melting' temperature is doubled. Another iteration of the algorithm is performed, and the percentage of accepted phase changes is determined. If this percentage is $\geq 90\%$, then the cooling of the system begins. If this percentage is $< 90\%$, the algorithm keeps iterating until this condition is satisfied.

Once the system is 'melted', the cooling begins. The temperature, $T_p$, of the system during the $p^{th}$ iteration is given by Eq. (3.2). Typical cooling rates are $\geq 0.9$; therefore, after 'melting' the system, the program performs successive iterations, as described above, at decreasing temperatures. After 10 consecutive iterations are executed in which a small percentage ($\leq 1\%$) of the phase changes are accepted after each iteration, the phase depth $\theta_0$ is optimized. Given the phase states (i.e. 0 or 1) of the grating, the cost $C_\theta$ in (3.7) is calculated as a function of the phase depth $\theta_0 \in [0, \pi]$. The optimum phase depth is determined by minimizing Eq. (3.7). After the phase depth has been optimized, the set of phases $\phi_n$ then constitutes the solution of the optimization problem.

The above algorithm was used to calculate an optimum 1D binary-phase profile for
implementing the ideal power spectrum shown in Figure 3.3. The various parameters used in the optimization routine were as follows: i) a 128-element (i.e. $N = 128$) array to represent a single period of the grating; ii) a desired fanout of $I \to I I$ (i.e. $K = 5$); iii) a melting temperature $T_0 = 1.0$; and iv) a cooling rate of 0.95. The algorithm took 226 iterations and approximately 15 minutes of CPU time on a VAX 8600 system.

In determining the optimum phase profile described below several runs of the simulated annealing algorithm were made and the 'best' result among these runs was taken. Each result was evaluated based on two figures of merit. The first figure of merit is the total diffraction efficiency of the grating, $\eta_t$, defined as

$$\eta_t = \frac{\sum_{m=\frac{N}{2}}^{\frac{N}{2}+K} P[m]}{\sum_{m=0}^{N-1} P[m]}$$

(3.8)

where $P[\frac{N}{2} + m] \equiv \eta_m$ is the power contained in the $m^{th}$ diffraction order or, equivalently, the diffraction efficiency of the $m^{th}$ diffraction order. $\eta_t$ is the ratio of the total power contained within the orders from $-K$ to $K$ and the total power contained within all diffraction orders. In general, one seeks to maximize $\eta_t$ since it represents the total amount of useful power directed in the desired directions.

Note that the quantity $\eta_t$ contains no information about the amount of power contained in each of the orders from $-K$ to $K$. For optical interconnect applications the relative ratio of power among these diffraction orders is of prime importance. For our purposes the power spectrum should be Gaussian, as shown in Figure 3.3. The second figure of merit, similar in form to the cost function $C$ in Eq. (3.6), quantifies the degree (i.e. excellent, good, poor) to which the grating power spectrum approximates the ideal power spectrum. We define a reconstruction error $\Delta R[m]$ for each order $m$, where $m$ is an integer in the range $[-K, K]$, as
\[ \Delta R[m] = \frac{\left| \frac{\eta_m / \eta_t - \eta_m^{\text{ideal}} / \eta_t^{\text{ideal}}}{} \right|}{\eta_m^{\text{ideal}} / \eta_t^{\text{ideal}}} \quad (3.9) \]

where \( \eta_m \) is the diffraction efficiency of the \( m^{th} \) diffraction order of the grating, \( \eta_m^{\text{ideal}} \) is the ideal diffraction efficiency of the \( m^{th} \) diffraction order of the grating, \( \eta_t \) is the total diffraction efficiency of the grating and \( \eta_t^{\text{ideal}} \) is the ideal total diffraction efficiency. As in Eq. (3.8), \( \eta_t^{\text{ideal}} \) is defined as

\[ \eta_t^{\text{ideal}} = \frac{\sum_{m=-N+K}^{N+K} P_{\text{ideal}}[m]}{\sum_{m=0}^{N-1} P_{\text{ideal}}[m]} . \quad (3.10) \]

For the Gaussian power spectrum shown in Figure 3.3, \( \eta_t^{\text{ideal}} = 0.93451 \). In Eq. (3.9) the term \( \frac{\eta_m}{\eta_t} \) represents the fraction of the total power in the orders of interest contained in the \( m^{th} \) diffraction order for the grating. Similarly, the term \( \frac{\eta_m^{\text{ideal}}}{\eta_t^{\text{ideal}}} \) represents the target value for the term described above. We can then take the maximum reconstruction error

\[ \Delta R_{\text{max}} \equiv \max_m \Delta R[m] \]

among all the orders as our figure of merit.

The surface-relief profile of the optimized 1D grating is shown in Figure 3.4. The optimized phase depth is \( \theta_0 = 0.78\pi \) radians = 2.405 radians. The etch depth required to obtain this phase depth can be found using Eq. (2.15) with a specification of the indices of refraction \( n_1 \) and \( n_2 \) and the wavelength of light \( \lambda_o \). The grating will be tested using a Helium-Neon laser; therefore, \( \lambda_o = 0.6328 \mu \text{m} \). The surrounding medium of the grating is free-space, and consequently \( n_2 = 1.0 \). The index \( n_1 \) is determined by the substrate material of the grating and also the wavelength of incident light. Fused silica, which is transparent at a wavelength of 0.6328 \( \mu \text{m} \), is the chosen substrate material. The index of refraction \( n \) of fused silica at a wavelength of 0.6328
Figure 3.4: Surface-relief profile of 1D binary-phase grating optimized with simulated annealing.
\[ n^2 - 1 = \frac{0.6961663}{\lambda_o^2 - (0.0684043)^2} \frac{\lambda_o^2}{\lambda_o^2 - (0.1162414)^2} + 0.4079426 \frac{\lambda_o^2}{\lambda_o^2 - (0.896161)^2} \]

\[ + 0.8974794 \frac{\lambda_o^2}{\lambda_o^2 - (9.896161)^2} \]

where \( \lambda_o \) is expressed in \( \mu \)m. From (3.11) \( n_2 = 1.4570179 \). Therefore, the necessary etch depth is found to be 6348 Å.

As described in Section 2.6, the surface-relief profile shown in Figure 3.4 can be fabricated by utilizing lithography and etching techniques. The necessary lithographic mask was designed using the format of a Mann 3600 Pattern Generator. The grating
Figure 3.6: *Power spectrum of optimized 1D binary-phase grating (solid curve) and ideal Gaussian power spectrum (dashed curve).*

period was chosen to be 128\(\mu\)m. Figure 3.5 shows a portion of the lithographic mask.

The power spectrum of this grating is plotted in Figure 3.6 along with the ideal Gaussian power spectrum (both normalized to the zero-order power). As shown, the grating power spectrum is a good approximation to the desired power spectrum. The maximum reconstruction error, occurring for the \(\pm\) orders, is 12.1%, while the reconstruction errors for the orders from -4 to 4 are all below 5%. The total diffraction efficiency of the grating is 0.795. These grating characteristics compare favorably with prior published results [16] [18] [42] [43]. Table 3.1 summarizes these results.

By crossing two 1D gratings with surface-relief profiles as shown in Figure 3.4
Table 3.1: Optimized power spectrum and reconstruction errors for 1D binary-phase grating optimized with simulated annealing.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\eta_m / \eta_0$</th>
<th>$\Delta R[m]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.149</td>
<td>0.0453</td>
</tr>
<tr>
<td>±1</td>
<td>0.135</td>
<td>0.00112</td>
</tr>
<tr>
<td>±2</td>
<td>0.117</td>
<td>0.0275</td>
</tr>
<tr>
<td>±3</td>
<td>0.0866</td>
<td>0.00295</td>
</tr>
<tr>
<td>±4</td>
<td>0.0563</td>
<td>0.0379</td>
</tr>
<tr>
<td>±5</td>
<td>0.0312</td>
<td>0.121</td>
</tr>
</tbody>
</table>

Figure 3.7: Surface-relief profile of 2D crossed grating.
(single period), one forms a 2D crossed grating with the surface-relief profile in Figure 3.7 (single period). As shown, the resulting phase element has 3 phase levels. From Eq. (3.3) these phase levels are $0, \theta_0$ and $2\theta_0$, where $\theta_0 = 2.4504$ radians. To fabricate this surface-relief profile, two lithography/etching steps are necessary with an $x-y$ alignment in between.

If we had constrained $\theta_0$ to be $\pi$ in the design of the 1D binary-phase grating, then the resulting 2D crossed grating would also have binary-phase. Although this phase element would be easier to fabricate, such a design constraint reduces the number of possible grating configurations (i.e. to those constrained to have a $\pi$ phase depth); consequently, the total diffraction efficiency and reconstruction errors of an optimized 'constrained' grating are (at best) lower and higher, respectively, than those achievable with optimized 'unconstrained' gratings.

The power spectrum of the 2D crossed grating is given by Eq. (3.5). The total diffraction efficiency of the 2D crossed grating, $\eta_{2D}^t$, is simply the square of the corresponding 1D total diffraction efficiency, $\eta_{1D}^t$; therefore, $\eta_{2D}^t = 0.63207$, approximately 20% lower than the 1D total diffraction efficiency. Figure 3.8 plots the power spectrum of the 2D crossed grating shown (single period) in Figure 3.7.

As described above, one way of obtaining a 2D power spectrum is to cross two 1D binary-phase gratings. In general, the resulting 2D surface-relief profile has 3 phase levels (see Figure 3.7). Practically, a binary-phase element is much easier to fabricate than a multi-level (i.e. number of levels greater than 2) element, and for this reason it is important to investigate 2D binary-phase gratings. In the special case where the phase depth of the 1D binary-phase grating referred to above is $\pi$, the 2D surface-relief profile formed by crossing two of these 1D gratings is also binary with a phase depth of $\pi$. Note, however, that constraining a 2D binary-phase grating to have a $\pi$
Figure 3.8: Power spectrum of 2D crossed grating.
phase depth does not guarantee a separable phase profile (i.e. satisfying Eq. (3.3)).
In general the set of possible configurations for a 2D phase grating formed by crossing
two 1D \(\pi\)-phase gratings is a very small subset of the set of possible configurations for
an arbitrary 2D binary-phase grating (let alone the set of possible configurations for
an arbitrary 2D binary-phase grating with a phase depth of \(\pi\)). In the next section
we describe the design of a 2D binary-phase grating to implement a 2D Gaussian
fanout.

### 3.1.2 2D Binary-Phase Grating

A general 2D binary-phase grating is given by Eq. (3.1) where \(\phi_{mn}\) is either 0 or
\(\theta_0 \in [0, \pi]\). For simplicity we assume \(M = N\), corresponding to a square grating.
The power spectrum \(P[l, p]\) of this grating is given by the 2D DFT of the grating
transmittance function

\[
P[l, p] = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \exp[j\phi_{mn}] \exp[j\frac{2\pi}{N}(lm + pn)].
\]  

Equation 3.12

The desired power spectrum \(P_{\text{ideal}}[l, p]\) is plotted in Figure 3.9. This spectrum was
obtained by sampling a circularly-symmetric 2D Gaussian. The sampling, however, is
done in a rectangular coordinate system due to the computational representation of
the grating (see Section 2.4). Consequently, \(P_{\text{ideal}}[l, p]\) has 8-fold symmetry about its
maximum value and an additional design constraint in the phase optimization problem
that the grating power spectrum have this 8-fold symmetry must be imposed.
Clearly, to satisfy this design constraint the phase profile must also have 8-fold sym-
metry. With this condition on the phase profile of the grating, \(N^2/8\) phases must be
optimized, as opposed to \(N^2\) phases for a fully arbitrary 2D phase grating.

A cost function \(C\) similar to that for the 1D grating described earlier is defined as
Figure 3.9: Ideal 2D power spectrum.
\[ C(\phi_{mn}) = \sum_{n=\frac{N}{2}-K}^{\frac{N}{2}+K} \sum_{m=\frac{N}{2}-K}^{\frac{N}{2}+K} (P[m, n] - P_{\text{ideal}}[m, n])^2 - (P[\frac{N}{2}, \frac{N}{2}] - P_{\text{ideal}}[\frac{N}{2}, \frac{N}{2}])^2. \]

(3.13)

For this particular grating \( N = 32 \) and \( K = 5 \) As in the 1L optimization the squared-error of the zero-order does not contribute to the cost. The cost function \( C_{\theta_0} \) for phase depth optimization is simply Eq. (3.7) extended to two dimensions:

\[ C_{\theta_0}(\theta_0) = \sum_{n=\frac{N}{2}-K}^{\frac{N}{2}+K} \sum_{m=\frac{N}{2}-K}^{\frac{N}{2}+K} (P[m, n] - P_{\text{ideal}}[m, n])^2 \]

(3.14)

where \( \theta_0 \in [0, \pi] \).

The procedures for determining the ‘melting’ temperature \( T_0 \), scanning of phase cells and state selection are identical to those for the 1D optimization problem described in the previous section. For the 2D optimization, however, the phase depth \( \theta_0 \) is optimized along with the binary states of the grating through each iteration. An iteration consists of an optimization of the binary states of each phase cell followed by an optimization of the phase depth for this set of optimized binary states. The algorithm stops after 10 consecutive iterations are executed in which less than 1% of the phase changes are accepted through each iteration. In the 1D optimization only the binary states of the grating are optimized through each iteration and the phase depth is optimized after the ‘10 consecutive iteration’ condition above is satisfied. By optimizing the phase depth after each iteration, as opposed to optimizing the phase depth after the ‘10 consecutive iteration’ condition is satisfied, the algorithm traverses a larger region of the solution space (i.e. the state space corresponding to all the possible grating phase configurations).

The various parameters used in the 2D optimization were as follows: i) a 32 X 32
array (i.e. $N = 32$) to represent a single period of the grating; ii) a desired fanout of 11-by-11 (i.e. $K = 5$); iii) a melting temperature $T_0 = 5.12$; and iv) a cooling rate of 0.90. The algorithm stopped after 84 iterations and required 204 minutes of CPU time on a VAX 8600 system.

The surface relief profile of the optimized 2D binary-phase grating is shown in Figure 3.10. As shown, this phase profile possesses 8-fold symmetry. The phase depth of the grating is $\theta_0 = 0.9170\pi$ radians = 2.8808 radians. The corresponding etch depth is 5054 Å (assuming $\lambda_o = 0.6328 \mu m$ and $n_{quartz} = 1.4570179$). The corresponding lithographic mask was designed with a grating period of 32\mu m. Figure 3.11 shows a portion of the lithographic mask. Fabricating this 2D binary-phase grating, as
Figure 3.11: Portion of the lithographic mask corresponding to the surface-relief shown in Figure 3.10
opposed to the 2D crossed grating, requires one less alignment step (i.e. requires no alignment) and one less lithography/etching step.

The power spectrum of this grating is plotted in Figure 3.12. The total diffraction efficiency of this grating is 0.75858, ~ 20% higher than for the 2D crossed grating. The reconstruction errors, however, are somewhat higher. The increased reconstruction error is due to the fact that a smaller-sized array (i.e. \( N = 32 \) versus \( N = 128 \)) was used to represent this grating, thereby increasing the effects of space quantization. Barnard et al. have investigated the dependence of reconstruction errors on array size for binary-phase elements and concluded that reconstruction errors varied as \( \frac{1}{N} \) [44]. Therefore, in a qualitative sense the increased reconstruction errors for the 2D binary-phase grating compared to those for the 2D crossed grating are reasonable.
Figure 3.13: Cross-section of 2D power spectrum of 2D binary-phase grating optimized using simulated annealing (solid curve) and cross-section of ideal 2D power spectrum (dashed curve).

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\eta_m/\eta_t$</th>
<th>$\Delta R[m]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0233</td>
<td>0.151</td>
</tr>
<tr>
<td>±1</td>
<td>0.0220</td>
<td>0.146</td>
</tr>
<tr>
<td>±2</td>
<td>0.0170</td>
<td>0.0466</td>
</tr>
<tr>
<td>±3</td>
<td>0.0107</td>
<td>0.128</td>
</tr>
<tr>
<td>±4</td>
<td>0.00731</td>
<td>0.122</td>
</tr>
<tr>
<td>±5</td>
<td>0.00384</td>
<td>0.239</td>
</tr>
</tbody>
</table>

Table 3.2: Optimized power spectrum and reconstruction errors for 2D binary-phase grating optimized with simulated annealing.
CHAPTER 3. PHASE OPTIMIZATION TECHNIQUES

Figure 3.13 plots a slice through the peak of the power spectrum shown in Figure 3.12 for the central 11 orders (solid curve) along with the desired power distribution (dashed curve) (both normalized to the zero-order power). Table 3.2 summarizes these results.

In the previous two sections we have shown that the technique of simulated annealing is well-suited to the design of binary-phase gratings. This technique was used to design one- and two-dimensional binary-phase gratings which implement a 2D Gaussian power spectrum. The technique of simulated annealing can be extended to the design of multi-level phase gratings; however, the increased number of possible phase configurations (i.e. a larger state space) requires that one perturb a larger number (directly proportional to the increase in possible phase configurations) of phase cells at a given temperature. Consequently, as the number of phase levels increases the necessary computation times become impractical (i.e. several days or weeks) [7].

In the next section we describe a phase optimization technique, known as \textit{phase retrieval}, which is ideally-suited for the design of multi-level phase elements.

\section{3.2 Phase Retrieval}

In the scalar limit the grating transmittance function and the grating angular spectrum are related through a Fourier transform. In designing free-space optical interconnects constraints on both the transmittance function and angular spectrum are imposed. The technique of phase retrieval [3] [4] accounts for these constraints in an iterative Fourier transform algorithm for the optimization of the grating phase. In this section we describe the phase retrieval technique and apply it to the design of a multi-level phase grating that implements the 2D Gaussian fanout shown in Figure 3.9.
Figure 3.14: Block diagram of phase retrieval algorithm.
CHAPTER 3. PHASE OPTIMIZATION TECHNIQUES

The phase retrieval algorithm, first developed by Gerchberg and Saxton [3] who applied it to electron microscopy, is shown in Figure 3.14. The object space corresponds to the grating and the image space corresponds to the grating power spectrum. The basic idea of the algorithm is to transform back and forth between these domains, while satisfying the constraints in one domain before returning to the other domain.

The algorithm starts in the object domain with an initial guess of the grating transmittance function. Since the grating is to be phase-only, the guess is typically implemented [3] [45] as a random configuration of phases, which we denote $\Phi_0[m,n]$. However, in general the solution that is obtained depends on this initial guess [45]. In some cases a crude estimate of the solution offers a superior starting point [46], while in other cases a random configuration performs better [47]. In any case having an initial input close to the true solution reduces the number of necessary iterations. Therefore, any knowledge of the form of the optimum grating phase should be utilized in specifying the initial input to the algorithm.

Given an initial phase profile the corresponding transmittance function is given by $T_0[m,n] = \exp[j\Phi_0[m,n]]$. The $r^{th}$ iteration of the algorithm proceeds as follows. A trial solution (which at the start of the algorithm is $\exp[j\Phi_0[m,n]]$) for the grating transmittance $\exp[j\Phi_r[m,n]]$ is Fourier transformed, yielding an angular spectrum $A_r[l,p] = |A_r[l,p]| \exp[j\angle A_r[l,p]]$ which in general contains both non-uniform magnitude and non-uniform phase components. Next, the image domain constraint is imposed. In our problem the image domain constraint is that the grating power spectrum be Gaussian. Therefore, we replace the magnitude $|A_r[l,p]|$ with a Gaussian distribution $|G[l,p]|$. This Gaussian distribution is given mathematically by the square root of the ideal power spectrum $P_{\text{ideal}}[l,p]$. The phase $A_r[l,p]$ is not changed since it is a free parameter in the problem.
The resulting function \( |G[l, p]| \exp[j \Lambda_r[l, p]] \) is then inverse Fourier transformed to yield a new transmittance function \( \overline{T}_r[m, n] = |\overline{T}_r[m, n]| \exp[j \overline{T}_r[m, n]] \), which in general also contains both non-uniform magnitude and non-uniform phase components. Next, the object domain constraint, which in our problem is that the grating be phase-only, is imposed. This constraint is formally expressed as \( |\overline{T}_r[m, n]| = 1 \). The resulting transmittance function \( \overline{T}_r[m, n] = \exp[j \overline{T}_r[m, n]] \) then serves as the input for the \((r + 1)\)th iteration of the algorithm.

As a measure of the convergence of the algorithm, we use the squared error \( \gamma \) defined as

\[
\gamma = \sum_{l = \frac{N}{2} - K}^{\frac{N}{2} + K} \sum_{p = \frac{N}{2} - K}^{\frac{N}{2} + K} (P[l, p] - P_{\text{ideal}}[l, p])^2.
\]  

(3.15)

where \( N \) is given by the array size \( N \times N \) and \( K \) is given by the desired 2D fanout of \((2K + 1) - by - (2K + 1)\). In our case \( N = 128 \) and \( K = 5 \). Typically, \( \gamma \) decreases rapidly for the first few iterations and then decreases more slowly for later iterations.

When \( \gamma = 0 \), an exact Fourier transform pair has been found that satisfies all the constraints. The phase \( \Theta[m, n] \equiv \angle \overline{T}_r[m, n] \) then represents the phase profile of the optimized grating. In most cases, however, an exact Fourier transform pair does not exist; consequently, a stopping criteria for the iterative procedure must be determined. The most straightforward approach, which is the approach we take, is to stop the algorithm after a certain number \( r_{\text{max}} \) of iterations and observe the grating power spectrum. If this spectrum adequately approximates (i.e. based on certain tolerances) the desired grating power spectrum, then the optimization problem has been ‘solved’. Otherwise, the algorithm can be executed again with a larger value of \( r_{\text{max}} \).

In the phase retrieval algorithms described above all elements of the array rep-
Figure 3.15: Block diagram of modified phase retrieval algorithm.
representing the grating phase profile can be modified simultaneously. This fact is in contrast to the situation in the simulated annealing algorithm, where each phase cell is considered individually. In this regard the phase retrieval algorithm tends to converge much more rapidly than the simulated annealing algorithm.

The above algorithm results in a grating phase profile which varies continuously. As discussed in Section 2.6 the corresponding surface-relief profile is very hard to fabricate. In Section 2.6 it was shown that we can approximate the continuous phase profile quite well by quantizing a modulo-$\pi$ version of the continuous phase profile $\Theta[m,n]$. The resulting phase profile $\Theta_q[m,n]$ defines the surface-relief profile of a binary optic element, as described in Section 2.6. The modulo-$\pi$/quantization operation $MOD_\pi/Q_M$, where $M$ denotes the number of phase levels used in the stepwise-approximation, can be incorporated into the phase retrieval algorithm, as depicted in Figure 3.15. This modified phase retrieval algorithm allows one to optimize phase profiles which can practically be fabricated with existing technologies.

The quantization scheme above does not work well for binary quantization ($M = 2$). Because of the coarse quantization the algorithm can encounter situations (on its path to the solution of the problem) where the object phase does not change significantly from iteration to iteration, causing the algorithm to stagnate (i.e. $\gamma$ remaining essentially constant for many iterations). Solutions to overcome this stagnation problem have been proposed [48] [49] [50]. These solutions involve a gradual introduction of the binary quantization, as opposed to the 'hard-clipping' operation corresponding to $Q_2$. However, the necessary modifications to the phase retrieval algorithm shown in Figure 3.15 are not easily implemented. This technique has been used to design quantized, blazed phase structures and amplitude holograms; however, binary-phase free-space optical interconnects have not been designed using this technique. Future
CHAPTER 3. PHASE OPTIMIZATION TECHNIQUES

research efforts will be directed towards this end. Our immediate goal was to design a multi-level free-space optical interconnect implementing a Gaussian fanout using the algorithm shown in Figure 3.15.

The parameters used in this optimization problem were as follows: i) a 128 X 128 array (i.e. $N = 128$) to represent a single period of the grating; ii) a desired fanout of $11 - by - 11$ (i.e. $K = 5$); iii) $M = 16$ (i.e. a 16-level surface-relief profile) and iv) $r_{max} = 100$.

The initial phase configuration that was used is given by

$$\Phi_0[m, n] = 2\pi \beta \exp \left[-\alpha \left((m - \frac{N}{2})^2 + (n - \frac{N}{2})^2\right)\right] \quad (3.16)$$

where $\alpha$ and $\beta$ are variables. Several runs of the algorithm were made for various $\alpha$'s and $\beta$'s. The optimum values were found to be $\alpha_{opt} = 10.0$ and $\beta_{opt} = 3.0$.

The optimization required $\sim 20$ minutes of CPU time on a VAX 8600 system. The squared-error as a function of the number of iterations is shown in Figure 3.16. The squared-error decreases very rapidly for the first 6 or 7 iterations and then changes relatively slowly thereafter. The algorithm was stopped after 100 iterations since, as shown, the squared-error appeared to have converged to a minimum value.

Figures 3.17(a) and (b) show the final continuous modulo-$\pi$ phase profile and the corresponding quantized modulo-$\pi$ phase profile. Recall from Section 2.6 that an $M$-level phase element, where $M = 2^L$ and $L = 1, 2, ..., \ldots$ requires $L$ lithography/etching processing steps. Therefore, 4 lithographic masks are needed to fabricate the above 16-level phase element. Figures 3.20a through 3.20d show portions of the appropriate lithographic masks. The grating period used in the mask designs was $128\mu m$.

The power spectrum of the above multi-level grating is shown in Figure 3.18. Figure 3.19 plots a slice through the peak of the grating power spectrum shown in
Figure 3.16: Squared-error versus number of iterations in phase retrieval algorithm.
Figure 3.17: Surface-relief profiles of 2D multi-level phase grating optimized with phase retrieval: (a) continuous phase; (b) quantized ($M = 16$) phase.
Figure 3.18: Power spectrum of 2D multi-level phase grating.
Figure 3.19: Cross-section of 2D power spectrum of 2D 16-level phase grating optimized using phase retrieval.
Figure 3.20: Portions of the lithographic masks corresponding to the surface-relief shown in Figure 3.17b: (a) Mask 1; (b) Mask 2.
Figure 3.20: Portions of the lithographic masks corresponding to the surface-relief shown in Figure 3.17b: (c) Mask 3, (d) Mask 4.
Table 3.3: Optimized power spectrum and reconstruction errors for 16-level phase grating optimized with phase retrieval.

Figure 3.18 (solid curve) and the desired grating power spectrum (dashed curve) for the central 11 diffraction orders (both normalized to the zero-order power). The agreement between the actual and ideal power spectrums is nearly exact.

The two power spectrums and reconstruction errors are compared in Table 3.3. The maximum reconstruction error is only 3.8%. An improved reconstruction compared to the gratings described earlier is expected since we are dealing with a multi-level phase element and not one with binary-phase. The multi-level element possesses a larger state space than that for binary-phase gratings; consequently, the global minimum of the state space corresponding to the multi-level phase element will be less than or equal to the corresponding global minimum of the state space for the binary-phase element. This multi-level phase element also has an improved total diffraction efficiency, 85%, over those gratings described earlier. This improvement is reasonable since, as noted above, a multi-level phase element provides more flexibility in spectrum shaping than does a binary-phase element.

However, the improved characteristics of the multi-level phase element come at a great expense. A multi-level element is far more complex to fabricate than a simple binary-phase element. This complexity arises from the required alignment steps between each of the lithography/etching steps, as described in Section 2.6. Therefore,
a tradeoff exists between high efficiency/good reconstruction error and fabrication complexity.

In this chapter the phase optimization techniques of simulated annealing and phase retrieval were described. Simulated annealing, developed from simple concepts in condensed matter physics, is well-suited for the design of binary-phase gratings. On the other hand, phase retrieval, based on an iterative Fourier transform algorithm, is more applicable to the design of multi-level phase gratings. These techniques were used to design several phase gratings with Gaussian power spectrums. Parameters such as diffraction efficiency and reconstruction error were described for each of these gratings. In the next chapter the experimental setup is considered and the power measurements for each of these gratings are presented and analyzed.
Chapter 4

Experimental Results and Discussion

In the previous chapter phase optimization techniques were described and applied to the design of phase gratings for optical interconnection applications. In particular, the techniques were applied to the design of binary-phase and multi-level phase gratings which implement a Gaussian fanout. Given these optimized gratings, corresponding lithographic masks were designed and the gratings were fabricated on fused silica substrates (1" X 0.020"). For a perfect fabrication process the actual grating power spectrums are given by the ideal spectrums in Chapter 3. (We neglect finite aperture effects, which broaden the ideal discrete grating power spectrum by a factor of the form $\frac{\sin S\theta}{\sin \theta}$ where $S$ is the number of periods within the finite aperture.) However, the fabrication process is not perfect. Two common errors are etch depth errors and linewidth errors, as described in Section 2.7. If these errors are known, their effects on the grating power spectrums can be determined.

In this chapter the experimental measurements are presented, particularly of the grating etch depths, and in some cases of the grating linewidths and the grating power spectrums. First, the measurement techniques are briefly discussed. Second,
the results for the 1D binary-phase grating and the corresponding 2D crossed grating, the 2D binary-phase grating and the multi-level phase grating are presented. For each grating the measured etch depths are compared with the ideal values. The formalism outlined in Section 2.7 is then used to determine the expected grating power spectrums. The measured grating power spectrums are then presented and compared with the expected power spectrums.

4.1 Measurement Techniques

In this section the techniques used to measure the grating etch depths, linewidths and power spectrums are described.

4.1.1 Etch Depth Measurement

The grating etch depths were measured with a Tencor Instruments\(^1\) - Alpha-step 200 stylus profilometer which has an accuracy of ±100Å. The etch depths at various locations on the grating were measured, since the etch rate varies as a function of space [31]. The reported values were obtained by averaging the etch depths of successive scans at various locations across the grating. The averaging process, therefore, effectively assumes that the etch depth is uniform across the corresponding grating. Given this ‘averaged’ etch depth the formalism described in Section 2.7 can be used to predict the diffraction efficiencies for the grating.

4.1.2 Linewidth Measurement

The grating linewidths were measured with a Leitz\(^2\) microscope (Model 307). A balanced illumination technique [51] provided measurements to an accuracy of ±0.10μm.

\(^1\)©Tencor Instruments
\(^2\)©Leitz-Wetzlar, Inc.
Figure 4.1: *Experimental setup for power spectrum measurement.*

The measured linewidths are used to predict the zero-order diffraction efficiency of the grating using the theory developed in Section 2.7.

### 4.1.3 Power Spectrum Measurement

The experimental setup for the power spectrum measurement is shown in Figure 4.1. A collimated HeNe ($\lambda = 0.6328\mu m$) is used to illuminate the grating. A HeNe laser was chosen for convenience because (i) a HeNe source was readily available and (ii) the HeNe wavelength is in the visible region of the spectrum, making it easier to work with (as opposed to a laser emitting in the infrared region of the spectrum) in the lab. An aperture was used to ensure that the beam width was smaller than the finite size of the grating ($\sim 2.5\text{mm}$). Illumination of the grating with a beam width which is larger than the grating extent effectively increases the DC component of the grating power spectrum. This DC component is an error that results in an overestimation of the zero-order diffraction efficiency of the grating.

The truncated HeNe beam is incident on the grating. Since the positions of the
HeNe laser and the limiting aperture were fixed, the grating was mounted on an 
x-y-z stage. These three degrees of freedom were used to align the grating to the 
optical axis defined by the illuminating beam. A 30cm lens is used to focus the 
diffracted waves into an evenly-spaced array of spots, where each spot corresponds to 
a specific diffraction order. The power contained in each spot is then measured with a 
photodetector (United Detector Technologies\textsuperscript{3} - Model 350 Linear/Log Optometer). 
Finally, to eliminate any stray light effects, the measurements were performed in 
complete darkness.

Two sets of power measurements were made for each grating: i) the power con-
tained in the orders of interest and ii) the total incident power on the grating. The 
first set of measurements provides information on the relative amounts of power con-
tained in each order, whereas the second measurement provides information on the 
total diffraction efficiency of the grating.

The total diffraction efficiency, as defined in Eq. (3.8), is normalized to the total 
transmitted power. The total transmitted power can be obtained from the measured 
incident power by accounting for the reflection losses suffered upon passing through 
the grating structure. At each air-substrate/substrate-air interface, a reflection occurs 
which reduces the amount of transmitted power into the substrate/air medium. The 
associated reflection coefficient $R$ for both cases is given by

$$R = \frac{(n - 1)^2}{(n + 1)^2} \quad (4.1)$$

where $n$ is the index of refraction of the substrate. In reality there are an infinite 
number of reflections and back-reflections (i.e. Fabry-Perot effects) occurring within 
the grating structure; however, a good approximation (for a substrate index of $\sim 1.5$)

\textsuperscript{3}\textsuperscript{3} \textcopyright United Detector Technologies, Inc.
can be made by assuming only 2 reflections (i.e. air/substrate interface upon incidence and substrate/air interface upon transmission). The normalized (to the incident power) total transmitted power $P_{tn}$, therefore, is given by

$$P_{tn} = (1 - R)^2. \quad (4.2)$$

For a wavelength of 0.6328$\mu$m and a fused silica substrate ($n = 1.4570179$ at $\lambda = 0.6328\mu$m), $P_{tn} = 0.932$. The total transmitted power $P_t$ is related to the total incident power $P_{inc}$ by

$$P_t = P_{tn} P_{inc} \quad (4.3)$$

where $P_{tn}$ is given by (4.2) and $P_{inc}$ is experimentally measured.

In this section the etch depth, linewidth and power spectrum measurement techniques were described. These techniques are used to characterize the properties of each of the fabricated gratings. In the next section the experimental results are presented and analyzed.

### 4.2 Results

This section presents the measured results, including etch depths, and in some cases linewidths and power spectrums, of the fabricated gratings. The order of presentation is as follows: i) the 1D binary-phase grating shown in Figure 3.4; ii) the 2D crossed grating shown in Figure 3.7 and iii) the 2D binary-phase grating shown in Figure 3.10.

Before proceeding, we mention a few words regarding nomenclature. In the following discussions the 'ideal' grating power spectrum refers to the appropriate grating power spectrum described in Chapter 3 (i.e. optimized with simulated annealing or
phase retrieval). The 'actual' or 'expected' grating power spectrum refers to the power spectrum of an over-/under-etched 'ideal' grating. The etch depth of the 'actual' grating is determined experimentally, as described in Section 4.1.1. The 'measured' grating power spectrum refers to the experimentally measured power spectrum of the fabricated grating. Therefore, the 'actual' grating differs from the 'measured' grating in that it only accounts for etch depth errors, which are also assumed to be uniform across the grating, in the fabrication process. The 'measured' grating may also suffer from non-uniform etch depth errors, linewidth errors, non-vertical sidewalls, etc... Therefore, to the extent that etch depth errors predominate over all other fabrication errors, the 'measured' values should agree well with the 'expected' values.

4.2.1 1D Binary-Phase Grating

As described earlier, the fabrication of a crossed grating requires two lithography/etching steps with an optical alignment in between. Both lithography/etching steps serve to fabricate a 1D grating. The optical alignment insures that these 1D gratings are fabricated orthogonally to one another. Etch depth errors or linewidth errors are generally present in both lithography/etching steps. Misalignment is another source of error. Consequently, the measured power spectrum of the resultant grating is not easily analyzed. A natural alternative is to analyze the behavior of the 1D grating before forming the 2D crossed grating. In this section we present the measured results for a 1D binary-phase grating whose ideal phase profile is shown in Figure 3.4. In the next section we present the measured results for the 2D crossed grating formed from this 1D grating.

In Section 3.1.1 the ideal etch depth was calculated to be 5400Å. The measured etch depth of the grating was 5530Å, corresponding to an over-etch of 130Å (dΦ₀ =
Table 4.1: Ideal and actual power spectrums for 1D binary-phase grating fabricated with a 130Å over-etch.

<table>
<thead>
<tr>
<th>m</th>
<th>$\eta_m/\eta_t^{ideal}$</th>
<th>$\eta_m/\eta_t^{actual}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.142</td>
<td>0.119</td>
</tr>
<tr>
<td>± 1</td>
<td>0.128</td>
<td>0.138</td>
</tr>
<tr>
<td>± 2</td>
<td>0.112</td>
<td>0.120</td>
</tr>
<tr>
<td>± 3</td>
<td>0.0825</td>
<td>0.0888</td>
</tr>
<tr>
<td>± 4</td>
<td>0.0537</td>
<td>0.0578</td>
</tr>
<tr>
<td>± 5</td>
<td>0.0297</td>
<td>0.0320</td>
</tr>
</tbody>
</table>

Table 4.2: Measured power spectrum and reconstruction errors for 1D binary-phase grating.

<table>
<thead>
<tr>
<th>m</th>
<th>$\eta_m/\eta_t^{meas}$</th>
<th>$\Delta R[m]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5</td>
<td>0.0345</td>
<td>0.0770</td>
</tr>
<tr>
<td>-4</td>
<td>0.0536</td>
<td>0.0725</td>
</tr>
<tr>
<td>-3</td>
<td>0.0956</td>
<td>0.0761</td>
</tr>
<tr>
<td>-2</td>
<td>0.114</td>
<td>0.0521</td>
</tr>
<tr>
<td>-1</td>
<td>0.143</td>
<td>0.0368</td>
</tr>
<tr>
<td>0</td>
<td>0.122</td>
<td>0.0258</td>
</tr>
<tr>
<td>1</td>
<td>0.145</td>
<td>0.0503</td>
</tr>
<tr>
<td>2</td>
<td>0.114</td>
<td>0.0491</td>
</tr>
<tr>
<td>3</td>
<td>0.0921</td>
<td>0.0371</td>
</tr>
<tr>
<td>4</td>
<td>0.0522</td>
<td>0.0971</td>
</tr>
<tr>
<td>5</td>
<td>0.0328</td>
<td>0.0243</td>
</tr>
</tbody>
</table>

$\pm 0.01877\pi)$. Given the ideal power spectrum (see Table 3.1), Eq. (2.41) can be used to determine the effect of the overetch on the grating power spectrum. The results are shown in Table 4.1. The total diffraction efficiencies of the ideal and actual gratings, $\eta_t^{ideal}$ and $\eta_t^{actual}$, are 0.79503 and 0.79068, respectively. As expected, the zero-order diffraction efficiency of the overetched grating is 15.84% less than that of the ideal grating and the higher-order diffraction efficiencies of the actual grating are 2.12% greater than those of the ideal grating. Clearly, the zero-order diffraction efficiency is more sensitive to etch depth errors than are the higher-order diffraction efficiencies.
Figure 4.2: Measured (solid curve) and expected (dashed curve) power spectrums for 1D binary-phase grating.

The measured grating power spectrum and the reconstruction errors (see Eq. (3.9)) with respect to the expected grating power spectrum are summarized in Table 4.2. The measured power spectrum is plotted along with the expected power spectrum in Figure 4.2 (both normalized to the zero-order diffraction efficiency). The total diffraction efficiency of the grating, $\eta_t^{\text{meas}}$, was measured to be 0.7692 which compares well with the expected value of 0.7906. The reconstruction errors are all below 10%. These reconstruction errors are with respect to the expected power spectrum, which was determined based on an experimentally measured etch depth. The accuracy of the etch depth measurement, however, is ±100Å. Given this accuracy the values for the reconstruction errors are reasonable.
CHAPTER 4. EXPERIMENTAL RESULTS AND DISCUSSION

Note that the measured power spectrum is not even-symmetric, as would be expected for a binary-phase grating. Two factors can account for this observed asymmetry. First, the assumption that the etch depth is uniform across the grating, although simple to model, is in general not valid. Etch depths measured at various locations across the grating varied by as much as 75 Å from the reported 'averaged' value. In general, the etch depth varies continuously across the substrate. Consequently, the grating phase profile also varies across the substrate. The resultant power spectrum of this varying-depth grating is in general not even-symmetric. A case in point is the well-known 'blazed grating', whose phase varies linearly with position. The ideal blazed grating concentrates all of the incident power into the first diffraction order (i.e. \( \eta_1 = 1.0 \) and \( \eta_k = 0 \) for all \( k \neq 1 \)). Clearly, this power spectrum is not even-symmetric.

Second, the contention that the power spectrum of a binary-phase grating is even-symmetric assumes the validity of the scalar approximation (and also the thin-grating approximation - see Chapter 5). Consider Figure 4.3 which plots the phase function \( \Phi[n] \), which represents a single period of the ideal 1D grating, as a function of the array index \( n \). When fabricating the grating represented by this array, a grating period \( \Lambda \) must be chosen to insure the validity of the scalar approximation. A sufficient condition on the grating period is \( \Lambda >> N\lambda \), where \( N \) is the dimension of \( \Phi[n] \) and \( \lambda \) is the wavelength of light (i.e the physical dimension of each pixel must be much larger than the wavelength of light). In our case \( N = 128 \), \( \Lambda = 128 \mu m \) and \( \lambda = 0.6328 \mu m \).

From Figure 4.3 we see that there are a few 'lines' which are only a single pixel wide. Therefore, the scalar approximation is not strictly valid, and we should not expect the measured power spectrum to be symmetric. However, since there are only a few small features in the profile, scalar theory should still give an accurate prediction.
Figure 4.3: Phase function $\Phi[n]$ for 1D binary-phase grating.
of the grating power spectrum. Therefore, the measured diffraction efficiencies should be close to those predicted by scalar diffraction theory; however, the spectrum itself will not in general be even-symmetric.

Recall that for a binary-phase grating the ratios of the higher-order diffraction efficiencies remains constant as a function of the grating etch depth. Table 4.3 summarizes the measured and expected ratios of the higher-order diffraction efficiencies. The ratios are all with respect to the diffraction efficiency of the +1 order, $\eta_1$. The measured ratios (for $|m| \neq 0, 1$) are approximately 10% lower than the expected ratios, while the ratio for $m = 0$ is approximately 7% lower. Note also that the diffraction efficiency of the $\pm 1$ orders are approximately equal. This information suggests that the fabricated grating concentrates a greater percentage of the incident power into the $\pm 1$ diffraction orders than desired. If the surface-relief profile had non-vertical sidewalls and/or a non-uniform etch depth, such an enhancement of the first-order diffraction efficiency could occur. However, since both of these effects are difficult to model (and to experimentally measure), we can simply state that they are plausible.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\eta_m/\eta_1$ - measured</th>
<th>$\eta_m/\eta_1$ - expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5</td>
<td>0.238</td>
<td>0.261</td>
</tr>
<tr>
<td>-4</td>
<td>0.369</td>
<td>0.406</td>
</tr>
<tr>
<td>-3</td>
<td>0.658</td>
<td>0.724</td>
</tr>
<tr>
<td>-2</td>
<td>0.784</td>
<td>0.869</td>
</tr>
<tr>
<td>-1</td>
<td>0.987</td>
<td>1.00</td>
</tr>
<tr>
<td>0</td>
<td>0.843</td>
<td>0.904</td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>0.787</td>
<td>0.869</td>
</tr>
<tr>
<td>3</td>
<td>0.634</td>
<td>0.724</td>
</tr>
<tr>
<td>4</td>
<td>0.359</td>
<td>0.406</td>
</tr>
<tr>
<td>5</td>
<td>0.226</td>
<td>0.261</td>
</tr>
</tbody>
</table>

Table 4.3: Measured and expected ratios of diffraction efficiencies (normalized to the first-order diffraction efficiency) for 1D binary-phase grating.


Table 4.4: Measured and ideal linewidths for 1D binary-phase grating (single period).

<table>
<thead>
<tr>
<th>Line#</th>
<th>Measured linewidth (μm)</th>
<th>Ideal linewidth (μm)</th>
<th>Deviation (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5 ±0.1</td>
<td>1.0</td>
<td>-0.5</td>
</tr>
<tr>
<td>2</td>
<td>0.6 ±0.1</td>
<td>1.0</td>
<td>-0.4</td>
</tr>
<tr>
<td>3</td>
<td>12.7 ±0.1</td>
<td>13.0</td>
<td>-0.3</td>
</tr>
<tr>
<td>4</td>
<td>0.6 ±0.1</td>
<td>1.0</td>
<td>-0.4</td>
</tr>
<tr>
<td>5</td>
<td>0.7 ±0.1</td>
<td>1.0</td>
<td>-0.3</td>
</tr>
<tr>
<td>6</td>
<td>16.7 ±0.1</td>
<td>17.0</td>
<td>-0.3</td>
</tr>
<tr>
<td>7</td>
<td>35.2 ±0.2</td>
<td>35.0</td>
<td>+0.2</td>
</tr>
</tbody>
</table>

explanations.

The measured grating linewidths (i.e. the 'high' phase states in Figure 4.3) are summarized in Table 4.4 for a single period of the grating. Given this data, the formalism described in Section 2.7, in particular Eq. (2.40) and Eq. (2.42), can be used to predict the change in the zero-order diffraction efficiency. The phase depth of the grating is (from Section 3.1.1) 0.78π. The duty cycle $D$ of this grating, determined from Figure 4.3, is 0.5313. The change in the duty cycle, $ΔD$, due to the linewidth errors is -0.02. From this information the change in the zero-order diffraction efficiency due to linewidth errors is found to be -0.00443. Accounting for this additional source of error, the expected quantity $η_0/η_1$ in Table 4.1 is decreased to 0.115. However, the corresponding measured value was 0.122, which results in an even larger reconstruction error of 0.0609.

A possible explanation for this discrepancy could be as follows. The analysis above assumes that all periods of the grating are equivalent (i.e. have the same linewidths), which is not true in general. Therefore, although the values in Table 4.4 indicate a decrease in the duty cycle, the duty cycle across the entire grating may in fact be larger than the target value, resulting in a larger zero-order diffraction efficiency. The change in the linewidths might also explain the lower ratios in Table 4.3. Recall that
linewidth errors change the diffraction efficiencies of all orders independently, unlike etch depth errors. Therefore, a linewidth error may have also increased the diffraction efficiencies of the +1 and -1 orders at the expense of the diffraction efficiencies of the other orders.

To the extent that the assumptions made, i.e. uniform etch depth, equivalent linewidths from period to period, the scalar approximation, are valid, the measured values are within experimental error. In theory, the fabrication of a binary-phase grating is simple; however, we see from the above discussion that many practical problems exist which make such a task non-trivial. In the next section we present the experimental results for a 2D crossed grating.

4.2.2 2D Crossed Grating

The ideal surface-relief profile of the 2D crossed grating is shown in Figure 3.7. We briefly review the fabrication process for this grating. First, the 1D binary-phase grating described in the previous section is fabricated. We refer to this grating as the 'x-grating'. Next, an identical 1D binary-phase grating whose 'lines' are orthogonal to the 'lines' of the 'x-grating' is fabricated. We refer to this grating as the 'y-grating'. For each fabrication step above the same errors as described in the previous section can occur. The optical alignment step associated with the 'crossing' further complicates the fabrication process.

The ideal etch depth for both the x— and y— gratings is 5400Å. The measured etch depths were as follows: 5450Å for the x-grating and 5675Å for the y-grating. Both gratings have been overetched (50Å for the x-grating and 275Å for the y-grating). From the measured etch depths the expected grating power spectrum can be calculated using the formalism in Section 2.7. The results are shown in Table 4.5. In the
Table 4.5: Ideal and expected power spectrums for 2D crossed grating fabricated with a 50Å overetch for the x-grating and 275Å overetch for the y-grating.

<table>
<thead>
<tr>
<th>$l, p$</th>
<th>$\eta_{lp}/\eta_{l}^{\text{ideal}}$</th>
<th>$\eta_{lp}/\eta_{l}^{\text{actual}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0</td>
<td>0.0201</td>
<td>0.0146</td>
</tr>
<tr>
<td>±1, 0</td>
<td>0.0165</td>
<td>0.0142</td>
</tr>
<tr>
<td>±2, 0</td>
<td>0.0125</td>
<td>0.0123</td>
</tr>
<tr>
<td>±3, 0</td>
<td>0.00681</td>
<td>0.00911</td>
</tr>
<tr>
<td>±4, 0</td>
<td>0.00288</td>
<td>0.00592</td>
</tr>
<tr>
<td>±5, 0</td>
<td>0.000884</td>
<td>0.00328</td>
</tr>
<tr>
<td>0, ±1</td>
<td>0.0165</td>
<td>0.0199</td>
</tr>
<tr>
<td>0, ±2</td>
<td>0.0125</td>
<td>0.0173</td>
</tr>
<tr>
<td>0, ±3</td>
<td>0.00681</td>
<td>0.0128</td>
</tr>
<tr>
<td>0, ±4</td>
<td>0.00288</td>
<td>0.00832</td>
</tr>
<tr>
<td>0, ±5</td>
<td>0.000884</td>
<td>0.00461</td>
</tr>
</tbody>
</table>

The measured grating power spectrum and the reconstruction errors with respect to the expected grating power spectrum are summarized in Table 4.6. Figures 4.4a and 4.4b plot the experimentally measured and expected 2D power spectrums, respectively. The total diffraction efficiency of the grating, $\eta_{l}^{\text{meas}}$, was measured to be 0.6172. Most of the reconstruction errors are within experimental error (i.e. the accuracy of the etch depth measurement). However, the reconstruction errors for $(l, p)$ of (-5,0),(4,0),(5,0),(0,3) and (0,0) are slightly higher. As mentioned in the previous section, several factors may account for this fact.

Non-vertical sidewalls may have increased the power directed into these diffraction orders. Similarly, linewidth errors may have redistributed the power in such a way
Figure 4.4: Power spectrums for 2D crossed grating: (a) measured; (b) expected.
Table 4.6: Measured power spectrum and reconstruction errors for 2D crossed grating.

that these orders are enhanced. For this grating the linewidths were not measured; however, it is feasible that such errors were present. In addition, non-uniform etch depths could have introduced significant asymmetries in the grating power spectrum, as discussed previously.

Since the gratings were overetched, the zero-order diffraction efficiency is expected to be lower than the ideal value, as expressed in Table 4.5. However, the measured value was much higher than this expected value and even higher than the ideal value. A possible explanation for this discrepancy could have been the following.

Lithographic exposure involves achieving intimate contact between the lithographic mask and the photoresist-coated substrate. Intimate contact insures good replication of the lithographic mask pattern onto the photoresist layer, which in turn results in a more anisotropic etch. If intimate contact is not achieved, air gaps are present in between the mask and substrate. Upon subsequent exposure the light that passes through the lithographic mask is diffracted within this air gap. The diffraction effectively decreases the amount of incident energy on the photoresist and results in an exposure with a non-uniform intensity profile. For a given exposure time, which
is calculated assuming intimate contact has been achieved, the non-uniform intensity profile incident on the photoresist layer results in a replicated linewidth which is smaller than the ideal linewidth.

When performing the second lithographic exposure, the lithographic mask was contacted with a non-planar substrate. In particular, the substrate had the surface-relief profile of the first binary-phase grating, which contained 'low' and 'high' levels. Therefore, intimate contact could not have been achieved between the mask and the substrate (i.e. in regions where a clear area on the lithographic mask overlapped a 'low' level of the substrate surface-relief profile). Therefore, the replicated linewidths (on the photoresist) for the second exposure could have been smaller than the expected values. These replicated linewidths, in turn, result (after subsequent reactive-ion-etching) in a grating with a duty cycle $D$ which is higher than desired. The duty cycle of the 1D binary-phase grating which is used to construct the 2D crossed grating is, as mentioned in the previous section, 0.5313. Recall that for a grating whose duty cycle is greater than 0.50 a positive change in the duty cycle (i.e. $\Delta D > 0$) increases the zero-order diffraction efficiency. The above arguments offer a qualitative explanation for the high value of the measured zero-order diffraction efficiency. The effect of non-intimate mask-to-substrate contact on the higher-order diffraction efficiencies is not as easily determined. The analysis is beyond the scope of this thesis.

The ratios of the higher-order diffraction efficiencies are shown in Table 4.7. Once again the ratios are all with respect to the diffraction efficiency of the +1 order, $\eta_1$. With the exception of the ratio for the zero-order, all ratios are within 10% of the expected values. The deviations could be a result of the fabrication errors described earlier, such as non-uniform etch depths, linewidth errors and non-intimate contact. The fact that the ratios are within 10% of the expected ratios suggests that the


\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
$l, p$ & $\eta_{lp}/\eta_1$ - measured & $\eta_{lp}/\eta_1$ - expected & $l, p$ & $\eta_{lp}/\eta_1$ - measured & $\eta_{lp}/\eta_1$ - expected \\
\hline
-5,0 & 0.266 & 0.232 & 0,5 & 0.210 & 0.231 \\
-4,0 & 0.386 & 0.418 & 0,4 & 0.379 & 0.418 \\
-3,0 & 0.671 & 0.643 & 0,3 & 0.586 & 0.642 \\
-2,0 & 0.855 & 0.869 & 0,2 & 0.852 & 0.869 \\
-1,0 & 0.956 & 1.00 & 0,1 & 1.05 & 1.00 \\
0,0 & 1.44 & 0.733 & 0,0 & 1.03 & 0.733 \\
1,0 & 1.00 & 1.00 & 0,1 & 1.00 & 1.00 \\
2,0 & 0.821 & 0.869 & 0,2 & 0.841 & 0.869 \\
3,0 & 0.613 & 0.643 & 0,3 & 0.703 & 0.642 \\
4,0 & 0.348 & 0.418 & 0,4 & 0.414 & 0.418 \\
5,0 & 0.179 & 0.232 & 0,5 & 0.241 & 0.231 \\
\hline
\end{tabular}
\caption{Measured and expected ratios of diffraction efficiencies (normalized to the first-order diffraction efficiency) for 2D crossed grating.}
\end{table}

Cumulative effects of the fabrication errors on the higher-orders is not significant; however, as shown these errors have a pronounced effect on the zero-order diffraction efficiency.

From the above discussion we see that the extra alignment and lithography/etching steps associated with the fabrication of a crossed grating, as opposed to the fabrication of a binary-phase grating, introduce another level of difficulty in the fabrication process. Binary-phase gratings, owing to their relative ease of fabrication (though far from trivial), are desirable from this standpoint. In the next section we present the experimental results for a 2D binary-phase grating.

\subsection{2D Binary-Phase Grating}

The ideal surface-relief profile of the 2D binary-phase grating is shown in Figure 3.10. The ideal etch depth of this grating is 6349Å. The measured etch depth was 6430Å, corresponding to an overetch of 81Å. Given this fact the expected grating power spectrum can be determined. The results of this analysis are shown in Table 4.8.
Table 4.8: Ideal and expected power spectrums for 2D binary-phase grating fabricated with a 81Å overetch.

<table>
<thead>
<tr>
<th>$l, p$</th>
<th>$\eta_l / \eta_{ideal}$</th>
<th>$\eta_l / \eta_{actual}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0</td>
<td>0.0226</td>
<td>0.0168</td>
</tr>
<tr>
<td>$(\pm 1, 0); (0, \pm 1)$</td>
<td>0.0213</td>
<td>0.0214</td>
</tr>
<tr>
<td>$(\pm 2, 0); (0, \pm 2)$</td>
<td>0.0165</td>
<td>0.0166</td>
</tr>
<tr>
<td>$(\pm 3, 0); (0, \pm 3)$</td>
<td>0.0104</td>
<td>0.0104</td>
</tr>
<tr>
<td>$(\pm 4, 0); (0, \pm 4)$</td>
<td>0.00708</td>
<td>0.00713</td>
</tr>
<tr>
<td>$(\pm 5, 0); (0, \pm 5)$</td>
<td>0.00373</td>
<td>0.00375</td>
</tr>
</tbody>
</table>

Table 4.9: Measured power spectrum and reconstruction errors for 2D binary-phase grating.

<table>
<thead>
<tr>
<th>$l, p$</th>
<th>$\eta_l / \eta_{meas}$</th>
<th>$\Delta R(l, p)$</th>
<th>$l, p$</th>
<th>$\eta_l / \eta_{meas}$</th>
<th>$\Delta R(l, p)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5, 0</td>
<td>0.00359</td>
<td>0.0414</td>
<td>0, -5</td>
<td>0.00342</td>
<td>0.0867</td>
</tr>
<tr>
<td>-4, 0</td>
<td>0.00772</td>
<td>0.0837</td>
<td>0, -4</td>
<td>0.00693</td>
<td>0.0275</td>
</tr>
<tr>
<td>-3, 0</td>
<td>0.0109</td>
<td>0.0479</td>
<td>0, -3</td>
<td>0.0109</td>
<td>0.0460</td>
</tr>
<tr>
<td>-2, 0</td>
<td>0.01691</td>
<td>0.0218</td>
<td>0, -2</td>
<td>0.0159</td>
<td>0.0411</td>
</tr>
<tr>
<td>-1, 0</td>
<td>0.0223</td>
<td>0.0434</td>
<td>0, -1</td>
<td>0.0221</td>
<td>0.0332</td>
</tr>
<tr>
<td>0, 0</td>
<td>0.0328</td>
<td>0.957</td>
<td>0, 0</td>
<td>0.0328</td>
<td>0.957</td>
</tr>
<tr>
<td>1, 0</td>
<td>0.0224</td>
<td>0.0448</td>
<td>0, 1</td>
<td>0.0222</td>
<td>0.0369</td>
</tr>
<tr>
<td>2, 0</td>
<td>0.0168</td>
<td>0.0151</td>
<td>0, 2</td>
<td>0.0157</td>
<td>0.0502</td>
</tr>
<tr>
<td>3, 0</td>
<td>0.0106</td>
<td>0.0182</td>
<td>0, 3</td>
<td>0.0104</td>
<td>0.00862</td>
</tr>
<tr>
<td>4, 0</td>
<td>0.00764</td>
<td>0.0717</td>
<td>0, 4</td>
<td>0.00679</td>
<td>0.0474</td>
</tr>
<tr>
<td>5, 0</td>
<td>0.00345</td>
<td>0.0793</td>
<td>0, 5</td>
<td>0.00317</td>
<td>0.155</td>
</tr>
</tbody>
</table>

Recall that the power spectrum of this grating contained 8-fold symmetry; therefore, the condition $\eta_{0p} = \eta_{l0}$ for $l = p$ holds, as expressed in the table. The total diffraction efficiencies of the ideal and actual grating, $\eta_l^{ideal}$ and $\eta_l^{actual}$, are 0.75858 and 0.75749, respectively.

The measured power spectrum and reconstruction errors with respect to the expected power spectrum are summarized in Table 4.9. The measured and expected power spectrums are also plotted in Figures 4.5a and 4.5b, respectively. Figure 4.5a shows that the measured power spectrum is fairly symmetric, as would be expected
Table 4.10: Measured and expected ratios of diffraction efficiencies (normalized to the first-order diffraction efficiency) for 2D binary-phase grating.

from a binary-phase grating. The total diffraction efficiency of the grating, $\eta_{t, \text{meas}}$, was measured to be 0.70770. With the exception of the zero- and (0,5)-orders the reconstruction errors are all below 10%, which is within the experimental error margin (i.e. accuracy of the etch depth measurement).

The slightly larger reconstruction error for the (0,5)-order may be due to a non-uniform etch depth, which would introduce an asymmetry into the power spectrum. Furthermore, as with the 1D binary-phase grating this grating contains diffracting features whose dimensions are on the order of wavelength of light; consequently, a power spectrum analysis based on scalar diffraction theory may not be strictly valid.

Note that an overetch would lead us to expect a reduced (from the ideal value) zero-order diffraction efficiency, as expressed in Table 4.8. However, the measured value was higher than both the expected and ideal values. To address this discrepancy, we refer to Table 4.10 which gives the ratio of the power spectrum with respect to the first-order diffraction efficiencies $\eta_{10}$ and $\eta_{01}$.

With the exception of the (5,0)- and (0,5)-orders the ratios of all the higher-orders
Figure 4.5: Power spectra for 2D binary-phase grating: (a) measured; (b) expected.
agree well with the expected values. Therefore, we would expect any etch depth errors to be fairly uniform across the grating since a highly non-uniform etch depth would result in large asymmetries in the grating power spectrum (e.g. the 'blazed grating' example). As shown, the zero-order ratio is much higher than expected. This fact suggests that the grating was under-etched rather than over-etched. Given the accuracy of the etch depth measurement (±100 Å), the stylus profilometer may in fact have overestimated the etch depth of the grating. Such an overestimation of the etch depth would result in an underestimation of the (expected) zero-order diffraction efficiency.

Apart from an apparent underetch the 2D binary-phase grating behaves nearly as expected. Recall that the 1D binary-phase grating also exhibited good characteristics. However, the 2D crossed grating, which required extra processing steps, was less accurate. The main point here is that the simpler the grating fabrication process the better since such surface-relief gratings are ultimately to be incorporated into practical systems.

In this chapter the experimental results and measurements for the fabricated gratings (corresponding to the optimized surface-relief profiles described in Chapter 3) were presented. The measured results agreed well with theory. The discrepancies between the experimental and theoretical results arise from both practical and theoretical considerations. First, measurement accuracy limits the precision by which we can evaluate the fabrication process. Second, the theoretical models used to predict the grating power spectrum do not account for fabrication errors such as non-uniform etch depths, non-vertical sidewalls or non-uniform linewidth errors. The development of more complete models which do account for these effects is presently an area of intense research. Third, the theoretical models themselves may not, in fact, be
appropriate for analyzing the above grating structures. The issue here is the validity of the approximations on which such models are based, particularly the scalar and thin-grating approximations. In the next chapter we focus on the issue of the validity of the scalar and thin-grating approximations in analyzing surface-relief grating structures.
Chapter 5

The Scalar and Thin-Grating Approximations

In theory, Maxwell's equations can be solved rigorously to provide an exact solution to any grating diffraction problem. In practice, the methods for obtaining rigorous solutions are very computationally-intensive. This fact limits the usefulness of such methods in the design of surface-relief phase gratings. Most design algorithms, such as simulated annealing and phase retrieval (see Chapter 3), utilize a Fourier-transform relation, which is ideally-suited for numerical computation, between the grating's surface-relief profile and its power spectrum. This relationship is the cornerstone of the well-known formalism of Fourier optics [20].

Because of the various properties and symmetries of the Fourier transform, Fourier optics provides a simple, intuitive description of grating diffraction, unlike a rigorous analysis of grating diffraction using Maxwell's equations. This fact, along with the fact that Fourier-transform-based algorithms are easily implemented on a computer, makes the use of Fourier optics for describing grating diffraction very attractive. However, one must remember that Fourier optics is an approximate theory. The Fourier transform relationship described above is derived from Maxwell's equations by making
two approximations: the scalar approximation and the thin-grating approximation. These approximations were briefly discussed in Chapter 2. In this chapter we reconsider these concepts in more detail.

Recall that the scalar approximation is a good one when the dimensions of the diffracting features of the grating are much larger than the wavelength of light. This condition is formally expressed as the limit $\Lambda/\lambda \to \infty$, where $\Lambda$ is the grating period and $\lambda$ is the wavelength of light. Actual gratings, however, are not infinite in extent. Furthermore, it is often advantageous and desirable to design and fabricate small and compact optical elements. The degree to which such optical elements can be 'shrunk' or miniaturized (given that we would still like to design such elements using a Fourier optics formalism, as opposed to a rigorous electromagnetic wave formalism) is directly related to the issue of the validity of the scalar approximation. Consequently, it is important to investigate the conditions under which the scalar approximation is well-satisfied for actual gratings.

The thin-grating approximation, which is only meaningful in the scalar limit (i.e. it is a further restriction on the grating given the validity of the scalar approximation), is a good one if the effect of the surface-relief phase grating on an incident wavefront at a given point is to delay (in phase) the wavefront by an amount proportional to the thickness of the grating at that point. This condition is true in the limit $d/\lambda \to 0$, where $d$ is the grating etch depth (once again, assuming that the value of $\Lambda/\lambda$ is sufficiently large).

In general, a grating is designed to have an optimum phase depth (e.g. determined from a phase optimization algorithm such as simulated annealing). This phase depth is practically implemented by etching away selected regions of the substrate surface, as described in Chapter 2. For a given phase depth the physical etch depth decreases
as the substrate index of refraction \( n \) increases (see Eq. (2.15)). Therefore, the limit 
\( d/\lambda \to 0 \) corresponds to the limit \( n \to \infty \). Clearly, for actual substrate materials this 
condition is not satisfied. Typical optical materials such as fused silica and germanium 
have refractive indices of approximately 1.5 and 4.0, respectively.

Experimental data suggest that surface-relief phase gratings (with sufficiently 
large periods) fabricated on such substrates can be well-approximated as 'thin' gratings. However, as the grating period decreases, the thin-grating approximation becomes a poorer one. We note that as the grating period decreases to a point where 
the scalar approximation breaks down, the entire notion of a transmission function and/or a thin-grating is irrelevant. However, for intermediate values of the period-to-wavelength ratio, the condition on the physical etch depth such that the grating can be considered 'thin' is in fact coupled to the period-to-wavelength ratio. This fact complicates the task of determining the validity of the thin-grating approximation for arbitrary grating structures. Therefore, it is important to investigate grating behavior for various period-to-wavelength ratios and also various substrate indices of refraction to determine the regime where both the scalar and thin-grating approximations are well-satisfied. In this regime a Fourier optics formalism can be used to accurately describe the grating diffraction characteristics.

The above discussion leads to a simple conclusion: Fourier optics can be reliably used to describe grating diffraction of actual surface-relief phase gratings when the dimensions of the grating are consistent with both the scalar and thin-grating approximations.

In this chapter we perform computer simulations to analyze the behavior of actual surface-relief gratings which implement ideal binary-phase gratings for various period-to-wavelength ratios and substrate indices of refraction to illustrate the regimes where
the scalar and thin-grating approximations are well-satisfied. The ideal binary-phase gratings considered implement free-space optical interconnections. Four gratings are analyzed: i) a grating implementing a 1-to-5 fanout with a uniform splitting ratio; ii) a grating implementing a 1-to-6 fanout with a uniform splitting ratio; iii) a grating implementing a 1-to-11 fanout with a uniform splitting ratio and iv) a grating implementing a 1-to-11 fanout with a Gaussian splitting ratio. The last grating is, in fact, the 1D binary-phase grating described in Section 3.1.1.

We note that our analysis is specific to the particular gratings considered. A similar analysis for arbitrary grating structures will in general give different results. It is not our intent to make general statements on grating behavior as a function of the period-to-wavelength ratio and/or the substrate index of refraction. In fact, it is not clear whether such a generalization is possible. The main purpose of this chapter is to provide some insight on the issue of the validity of the scalar and thin-grating approximations.

This chapter is divided into two sections. Section 5.1 begins with a description of the different regimes of grating diffraction. Next, a rigorous vector coupled-wave theory of grating diffraction, developed by Moharam and Gaylord [21], is outlined. This formalism is used to analyze the diffraction characteristics of the gratings described earlier. This approach produces exact formulations of grating diffraction problems without any approximations. An extensive computer program, DIFFRACT, was written to implement this theory. DIFFRACT was used to determine the power spectrums of the above gratings for various period-to-wavelength ratios and substrate indices of refraction. The results of this analysis are compared to grating diffraction results calculated using a Fourier optics formalism in Section 5.2.
5.1 Regimes of Grating Diffraction

Various theoretical formulations, all derived from Maxwell's equations, can be employed to analyze grating diffraction. These formulations differ in the approximations made in their development. In theory, Maxwell's equations can be solved rigorously to yield exact solutions to all grating diffraction problems; however, under certain conditions approximate theories can provide extremely good solutions to the problem at hand.

The choice of using an approximate theory over a rigorous one is purely a practical one (e.g. ease of implementation, necessary computation time, etc...). The use of an approximate theory provides an accurate description of experimentally observed phenomena only if the physical parameters of the system are consistent with the approximations made in the theory. Therefore, the particular choice of utilizing one grating diffraction formalism over another one is dependent on the physical parameters of the grating diffraction problem. The physical dimensions of the grating, particularly the period and the etch depth, compared to the wavelength of light provide a basis for selection among the various formalisms. The issue involved concerns the validity of the scalar and thin-grating approximations, as described earlier.

The analysis of grating diffraction can be categorized in three general regimes: i) a regime where the scalar approximation is not well-satisfied; ii) a regime where the scalar approximation is well-satisfied and the thin-grating approximation is not well-satisfied and iii) a regime where both the scalar and thin-grating approximations are well-satisfied. A complete description of the various diffraction theories which can be employed in the above regimes is beyond the scope of this thesis. An excellent review on this topic is given in [21]. We do, however, address the qualitative aspects
CHAPTER 5. THE SCALAR AND THIN-GRAting APPROXIMATIONS

of the required method of solution for grating problems in each of the above regimes.

Scalar diffraction theory can be used in the second and third regimes above, while a vectorial formulation of grating diffraction, particularly to account for polarization effects, is necessary in the first regime. The onset of polarization effects, in fact, establishes the range of validity of scalar diffraction theory. Scalar diffraction theory can be reliably used to predict grating diffraction characteristics when polarization effects are very small. We begin our discussion with the simplest case, corresponding to the third regime above.

The third regime is the familiar case where a Fourier optics formalism [20] can be used to describe grating diffraction. In this regime scalar diffraction theory is combined with a simplified model of the surface-relief profile (i.e. the thin-grating approximation) to yield simple, intuitive results for grating diffraction. This formalism was discussed in Chapter 2.

In the second regime a scalar theory can still be used; however, the transmission function of the grating is not straightforwardly related to the grating surface-relief profile, as was the case in the third regime (see Eq. (2.15)). We give a qualitative explanation for the differences between these two regimes. The fact that the grating is not 'thin' suggests that the ridges of the surface-relief profile can be thought of as waveguides. Therefore, the surface-relief grating can be thought of as a set of homogeneous multi-mode (since λ/λ is large) waveguides. The field distribution at the output plane of the grating is determined by solving the scalar coupled-mode problem defined by the grating structure, which in this discussion is assumed to have a binary surface-relief profile.

The field distribution at the output plane has a uniform amplitude. However, in general the phase profile is not binary, as a consequence of the different phase velocities
associated with each 'waveguide'. Given the the phase profile at the output plane of the grating, the corresponding grating transmission function can be determined from Eq. (2.8) and the grating angular spectrum can be calculated using Eq. (2.6). The grating power spectrum is then given by the squared-magnitude of the angular spectrum.

We note that if the surface-relief profile has a center of symmetry, then the phase profile at the output plane of the grating will also have a center of symmetry. Similarly, if the surface-relief profile does not have a center of symmetry, then the phase profile at the output plane of the grating will not have any symmetry associated with it. These symmetry considerations are related to the grating power spectrum in that any symmetry in the grating phase profile is preserved in the grating power spectrum. In other words, the symmetric grating phase profile described above has a corresponding grating power spectrum which is also symmetric. On the other hand, the power spectrum corresponding to a grating phase profile which contains no symmetry has in general no symmetry associated with it. Recall that the power spectrums for 'thin' binary-phase gratings are necessarily symmetric. Clearly, we see that if a phase grating whose surface-relief profile is binary cannot be considered 'thin', the grating diffraction behavior is more complicated.

(Note: A 'thin' binary-phase grating, regardless of the symmetry of its surface-relief profile or lack thereof, has a symmetric power spectrum. This symmetry arises from the fact that the grating phase profile is binary. A 'thick' grating whose surface-relief profile is binary will not in general have a phase profile at the output plane of the grating which is binary. However, the corresponding power spectrum of this 'non-binary' phase profile will be symmetric if the grating surface-relief profile is symmetric since the phase profile at the output plane of the grating will be symmetric. Consider
a 'thick' grating whose surface-relief profile does not contain any symmetry. Since
the phase profile at the output plane of this 'thick' grating is in general not binary,
the corresponding power spectrum will not be symmetric since the phase profile at
the output plane of the grating will not contain any symmetry. We see that there are
two properties of a phase grating (in the limit that scalar diffraction theory is valid)
that result in a symmetric power spectrum: i) a phase profile which is binary and
ii) a surface-relief profile which is symmetric. Because of this fact, a comparison of
the power spectrums of 'thin' and 'thick' phase gratings whose surface-relief profile
is binary may be slightly misleading.)

Therefore, the essential difference between the second and third regimes is in the
calculation of the grating transmission function. In the latter the transmission func-
tion is related to the grating surface-relief profile in a simple fashion, whereas in the
former a non-trivial scalar coupled-mode problem must be solved to determine the
transmission function. In both regimes, however, once the grating transmission func-
tion is determined, the grating angular spectrum (and, therefore, the grating power
spectrum) is determined by a plane wave decomposition, as discussed in Chapter 2.

In the first regime above scalar diffraction theory completely breaks down, and
a rigorous solution to Maxwell's equations is required. The inaccuracy of the scalar
formalism is due to the fact that it does not account for polarization effects, which
become increasingly important as the dimensions of the diffracting features approach
the wavelength of light. The necessary formalism must treat light as a vector field,
as opposed to a scalar field. This fact immediately suggests an increased complexity
in the required formalism compared to a scalar treatment.

The most common rigorous solution techniques are the vector coupled-wave ap-
proach [21] [52] [53] and the modal approach [54] [55] [56]. These techniques in
CHAPTER 5. THE SCALAR AND THIN-GRATING APPROXIMATIONS

general involve an expansion of the field both inside and outside the grating medium in terms of a suitable set of basis functions (e.g. plane waves, space harmonics, etc...). The diffracted fields are determined by matching boundary conditions at the interfaces. Although exact, such methods do not lend themselves to much intuitive insight. Furthermore, the implementation of such algorithms is extremely computationally-intensive. Because of these facts, such methods are not easily utilized in grating design. However, as the grating dimensions approach the wavelength of light, such methods are indispensable for an accurate description of the grating diffraction characteristics.

Note that from symmetry arguments alone (i.e. not 'intuition' about the grating diffraction behavior) one can conclude that if the grating surface-relief profile has a center of symmetry, the corresponding grating power spectrum will also have a center of symmetry. Similarly, if the grating surface-relief profile does not contain any symmetry, we would not expect the grating power spectrum to contain any symmetry either. Recall that when scalar diffraction theory is not valid, the concept of a grating transmission function is irrelevant; the notion of a grating phase profile (e.g. a binary-phase grating) loses its meaning. It is just this fact that makes a rigorous analysis of grating diffraction 'non-intuitive'.

From the above discussion one can conclude that the appropriate method of solution for a particular grating diffraction problem depends on the regime in which the problem is specified. Once the regime has been determined one can proceed to a solution straightforwardly. Often, however, the determination of the operating regime is difficult. This difficulty arises from the fact that the criteria for determining the operating regime depend on the physical dimensions of a grating structure. Clearly, there are an infinite number of possible grating surface-relief profiles. Consequently,
CHAPTER 5. THE SCALAR AND THIN-GRATING APPROXIMATIONS

every grating has its own unique specification of operating regimes.

In this work we investigate the behavior of surface-relief phase gratings for various period-to-wavelength ratios and substrate indices of refraction. These surface-relief gratings practically implement ideal binary-phase profiles. The particular binary-phase profiles we consider are optimized (e.g. by simulated annealing) to implement free-space optical interconnections. Once again, we do not attempt to establish general conditions for the validity of the scalar and thin-grating approximations for arbitrary surface-relief grating structures. The intent of our investigation is primarily to shed some light on the issues of the validity of the scalar and thin-grating approximations.

To correctly analyze our gratings for small period-to-wavelength ratios, we must use a rigorous formulation of grating diffraction, as discussed earlier. Among the many formalisms we have chosen a rigorous vector coupled-wave formalism first employed by Moharam and Gaylord [21].

An extensive computer code, DIFFRACT [57], has been developed to implement this formalism. The necessary computation time for a given grating diffraction problem is a strong function of the number of propagating orders from the grating. The number of propagating orders is dependent on the period-to-wavelength ratio and can be determined from the grating equation (see Eq. (2.23)). As the period-to-wavelength ratio increases, so does the number of propagating diffraction orders. In practice, gratings with period-to-wavelength ratios larger than 50 become unreasonable to try to solve using the DIFFRACT code. For this reason we only perform calculations for period-to-wavelength ratios up to 50.

A grating is specified in the DIFFRACT program by inputting the transition points specifying its binary surface-relief profile. The wavelength of light which is
used in all of our simulations was 1 µm. The medium surrounding the grating has an index of refraction of 1.0, corresponding to free space. We consider two grating substrate indices of refraction: \( n = 1.5 \), corresponding to a fused silica substrate, and \( n = 4.0 \), corresponding to a germanium substrate. The last parameter in our simulations was the grating period. As mentioned earlier, period-to-wavelength ratios up to 50 are considered, corresponding to a maximum period of 50 µm. We note that the period-to-wavelength ratios are specified for the wavelength of light in free space.

Given the above input parameters the \textit{DIFFRACT} code calculates the power contained in the reflected and transmitted diffraction orders of the grating. The reflected orders propagate in free space (with an index of refraction of 1.0) and the transmitted orders propagate within the grating medium (with an index of refraction of either 1.5 or 4.0, as described above), which is assumed to be infinite in extent along the direction of the optical axis. This transmitted power spectrum is compared with the corresponding power spectrum calculated using a Fourier optics formalism (i.e. for a similar grating with a period of infinite extent and a substrate index of refraction approaching infinity) as a function of both the period-to-wavelength ratio and also the substrate index of refraction. In the next section we present the results of this comparative study.

We note that it is not strictly correct to compare the rigorously-calculated transmitted power spectrum above with experimental data. Actual surface-relief phase gratings are fabricated on substrates with finite thickness. The \textit{DIFFRACT} simulations above do not account for the finite thickness of the grating substrate. Rather, the substrate is assumed to have an infinite thickness. The effect of the finite substrate thickness is significant only for small values of the period-to-wavelength ratio. Therefore, when evaluating experimental data taken from grating structures with
small period-to-wavelength ratios, the finite thickness of the grating substrate must be accounted for.

5.2 Theoretical Results

In this section the results of a rigorous coupled-wave analysis of surface-relief gratings are presented. These surface-relief gratings are practical implementations of idealized binary-phase gratings. The binary-phase gratings considered implement free-space optical interconnects. Specifically, the binary-phase gratings we considered are as follows: i) a grating implementing a 1-to-5 fanout with a uniform splitting ratio; ii) a grating implementing a 1-to-6 fanout with a uniform splitting ratio; iii) a grating implementing a 1-to-11 fanout with a uniform splitting ratio and iv) a grating implementing a 1-to-11 fanout with a Gaussian splitting ratio.

The power spectrums of these gratings are examined as a function of the period-to-wavelength ratio and also for two different substrate indices of refraction. The goals of such an analysis are as follows: i) to confirm the fact that scalar diffraction theory is only valid for large period-to-wavelength ratios; ii) to qualitatively illustrate the effects of the period-to-wavelength ratio on the performance of binary-phase gratings and iii) to investigate the effects of the substrate index of refraction on the validity of the thin-grating approximation. The results of the rigorous formulation are compared to those predicted by a Fourier optics treatment of grating diffraction. The deviation between the two sets of results may have practical implications in terms of grating design and fabrication.
5.2.1 1-to-5 fanout grating

Figure 5.1 shows the phase profile of a single period of this grating. The specifications of this grating (i.e. the transition points and phase depth) were taken from [22]. The grating phase depth is 2.986 radians. Clearly, the phase profile of this grating is symmetric; therefore, the corresponding power spectrum should also be symmetric. Figure 5.2 plots the power spectrum of this grating calculated using a Fourier optics formalism. This power spectrum can be thought of as ideal in the sense that the binary-phase grating corresponding to this power spectrum has a period of infinite extent (as a consequence of the scalar approximation) and has an etch depth of zero thickness (as a consequence of the thin-grating approximation). Consequently, the power spectrums of actual surface-relief gratings which implement this ideal binary-phase grating will differ from this ideal power spectrum. The degree to which the two spectrums deviate from one another provides a measure of the validity of scalar and thin-grating approximations.

The deviation can be quantified by the maximum reconstruction error $\Delta R_{\text{max}}$ among all diffraction orders, as defined in Chapter 3:

$$\Delta R_{\text{max}} \equiv \max_m \Delta R[m]$$  \hspace{1cm} (5.1)

where $\Delta R[m]$, the reconstruction error for each diffraction order, is given by

$$\Delta R[m] = \frac{\left| \frac{\eta_m}{\eta_t} - \frac{\eta_{m,\text{ideal}}}{\eta_{t,\text{ideal}}} \right|}{\eta_{m,\text{ideal}}/\eta_{t,\text{ideal}}}$$  \hspace{1cm} (5.2)

where $\eta_m$ is the diffraction efficiency of the $m^{th}$ diffraction order of the actual surface-relief grating, $\eta_{m,\text{ideal}}$ is the ideal diffraction efficiency of the $m^{th}$ diffraction order of the ideal binary-phase grating, $\eta_t$ is the total diffraction efficiency of the
Figure 5.1: Phase profile (single period) of 1-to-5 fanout grating.
Figure 5.2: Ideal power spectrum of 1-to-5 fanout grating.
CHAPTER 5. THE SCALAR AND THIN-GRATING APPROXIMATIONS

actual surface-relief grating and $\eta_{\text{ideal}}$ is the total diffraction efficiency of the ideal binary-phase grating.

Physically, as the period-to-wavelength ratio increases the maximum reconstruction error should approach zero. For small values of the period-to-wavelength ratio, it is unclear what type of behavior to expect. This uncertainty arises from the fact that a rigorous analysis of grating diffraction offers very little insight into the expected results.

Figures 5.3a through 5.3d plot the grating power spectrums for the orders from -2 to +2 of the 1-to-5 fanout grating for period-to-wavelength ratios of 2, 5, 25 and 50, respectively. The substrate index of refraction in each case is 1.5 (fused silica), corresponding to a physical etch depth of 0.9505$\mu$m, and the incident polarization is TE. In all of the plots the solid curve corresponds to the ideal power spectrum calculated using a Fourier optics formalism.

The most apparent feature of these plots is that the power spectrums for increasing period-to-wavelength ratios agree more closely with the ideal power spectrum. The maximum reconstruction errors are 0.793, 0.326, 0.130 and 0.0498 for period-to-wavelength ratio values of 2, 5, 25 and 50, respectively. Therefore, at a period-to-wavelength ratio of 50, the grating power spectrum is within $\pm$5% of the ideal spectrum. This result suggests that at a period-to-wavelength ratio of 50, both the scalar approximation and the thin-grating approximation are well-satisfied.

These results have a practical significance. Consider the application of optical switching. Assume that the system tolerances are specified such that the power contained in each beam be above a certain threshold value. For simplicity we choose this threshold value at 50% of the mean value of the ideal higher-order diffraction efficiencies. Given this threshold value the maximum tolerable reconstruction error
Figure 5.3: Power spectrum of 1-to-5 fanout grating for various period-to-wavelength ratios: (a) \( \Lambda/\lambda = 2 \); (b) \( \Lambda/\lambda = 5 \). Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.
Figure 5.3: Power spectrum of 1-to-5 fanout grating for various period-to-wavelength ratios: (c) $\Lambda/\lambda = 25$; (d) $\Lambda/\lambda = 50$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.
Figure 5.4: Power spectrum of 1-to-5 fanout grating - TE (solid curve) and TM (dashed curve) incident polarizations; $\Lambda/\lambda = 50; n = 1.5$.

is approximately 0.500. Therefore, a grating could be fabricated with a minimum period-to-wavelength ratio of approximately 2.5 (assuming an incident polarization of TE) that would still meet the above design specification. We have considered a particular application where the deviation in absolute power, rather than the percentage deviation (which is what the reconstruction error represents) in power, is important. In applications where the percentage deviation in power is important, there is less flexibility in decreasing the period-to-wavelength ratio of the surface-relief grating.

Figure 5.4 plots the grating power spectrum for TE and TM incident polarizations and a period-to-wavelength ratio of 50 and a substrate index of refraction of 1.5. As
Figure 5.5: Power spectrum of 1-to-5 fanout grating - TE (solid curve) and TM (dashed curve) incident polarizations; $\lambda/\lambda = 5; n = 1.5.$
Figure 5.6: Maximum reconstruction error versus period-to-wavelength ratio for 1-to-5 fanout grating - TE (*) and TM (o) incident polarizations; $n = 1.5$.

shown, the profiles are nearly identical, again suggesting that the scalar approximation is well-satisfied. To illustrate the effect of the incident polarization on the grating power spectrum for small values of the period-to-wavelength ratio, consider Figure 5.5 which plots the grating power spectrum for TE and TM incident polarizations and a period-to-wavelength ratio of 5. The substrate index of refraction is 1.5. Clearly, at a period-to-wavelength ratio of 5 the grating power spectrum is sensitive to the incident polarization, suggesting that the scalar approximation is not well-satisfied at a period-to-wavelength ratio of 5.

Figure 5.6 plots the maximum reconstruction error of the grating as a function
of the period-to-wavelength ratio for both TE and TM incident polarizations and a substrate index of refraction of 1.5. The minimum period-to-wavelength ratio considered was 2. This value was chosen such that at least 5 beams were propagating in both the incident medium and grating medium. The maximum period-to-wavelength ratio considered was 50 due to practical considerations, as mentioned earlier.

There are three important features of this plot. First, as the period-to-wavelength ratio increases the TE and TM curves approach each other. This result is reasonable since we know that in the limit of an infinite period-to-wavelength ratio, the two cases are degenerate.

Second, at small and moderate values of the period-to-wavelength ratio the maximum reconstruction errors for the TE- and TM-polarized incident waves deviate significantly. For period-to-wavelength ratios greater than 5, the power spectrum for an incident TE-polarized wave deviates further from the ideal spectrum than does the power spectrum for an incident TM-polarized wave. For period-to-wavelength ratios down to approximately 8, the maximum reconstruction error for the TM-polarized case is below approximately 0.12. Therefore, if an optical system is designed such that a grating reconstruction error of 0.12 does not affect its performance significantly and also such that it can utilize TM-polarized light, then the results shown in Figure 5.6 establish a lower bound for a grating period. The particularly low value of the maximum reconstruction error for the TM-polarized case at a period-to-wavelength ratio of 9 is an unexpected result. The reconstruction error at this point is 0.0287, which is even lower than the value at a period-to-wavelength ratio of 50, 0.0348. The power spectrum for an incident TM-polarized wave and a period-to-wavelength ratio of 9 is shown in Figure 5.7 along with the ideal power spectrum.

The last point to note is that the plot contains no information on the actual
Figure 5.7: Power spectrum of 1-to-5 fanout grating - $\Lambda/\lambda = 9; n = 1.5$; TM incident polarization. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.
CHAPTER 5. THE SCALAR AND THIN-GRATING APPROXIMATIONS

Figure 5.8: Power spectrum of 1-to-5 fanout grating - $\Lambda/\lambda = 50; n = 4.0$; TE incident polarization. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.

amount of power contained in the transmitted orders. For example, the plot shows that at a period-to-wavelength ratio of 5, both the TE- and TM-polarized cases have a maximum reconstruction error of approximately 0.32; however, as shown in Figure 5.5 the power spectrums for these two polarizations differ dramatically.

To illustrate the effects of the substrate index of refraction on the grating power spectrum in the scalar limit, we considered the grating power spectrum for a TE-polarized incident wave and a substrate with an index of refraction of 4.0 at a period-to-wavelength ratio of 50. The physical etch depth of this grating is 0.1584$\mu$m, which is 6 times smaller than that for a grating fabricated on a substrate with an index of
refraction of 1.5. Therefore, this grating more closely approximates a 'thin' grating and should agree more closely to the ideal spectrum than a similar grating fabricated on a substrate with an index of refraction of 1.5. The power spectrum of this grating is shown in Figure 5.8 along with the ideal spectrum. The spectrums are nearly identical. Furthermore, a comparison with Figure 5.3d shows that the spectrum of the grating fabricated on a substrate with an index of refraction of 4.0 does in fact show better agreement with the ideal spectrum then does a similar grating fabricated on a substrate with an index of refraction of 1.5. The maximum reconstruction errors for each case are 0.0465 and 0.00400 for \( n = 1.5 \) and \( n = 4.0 \), respectively. The etch-depth-to-wavelength ratios are 0.9505 and 0.1584 for the \( n = 1.5 \) and \( n = 4.0 \) cases, respectively.

To further evaluate the validity of the scalar and thin-grating approximations, we utilized the following fact. If the above approximations are well-satisfied for a given binary surface-relief phase grating, then the power spectrum of this grating should be fairly independent of the parity of the phase states of the grating. Figure 5.9 illustrates a case where this property is not true. The plot shows the power spectrums of both the surface-relief grating implementing the ideal binary-phase grating shown in Figure 5.1 and its corresponding reverse parity surface-relief grating for TE incident polarization, a period-to-wavelength ratio of 2 and a substrate index of refraction of 1.5. Clearly, the grating power spectrum is highly-dependent on the parity of the grating.

Figures 5.10a and 5.10b show plots similar to Figure 5.9 except for a period-to-wavelength ratio of 50 and two different substrate indices of refraction (\( n = 1.5 \) in 5.10a and \( n = 4.0 \) in 5.10b). In both cases the power spectrum is fairly independent of the parity of the grating (i.e. within \( \pm 10\% \) of the mean value between the two
Figure 5.9: Power spectrum of 1-to-5 fanout grating and corresponding reverse parity grating - \( \Lambda/\lambda = 2; n = 1.5 \); TE incident polarization.
Figure 5.10: Power spectrum of 1-to-5 fanout grating and corresponding reverse parity grating - $\Lambda/\lambda = 50$; TE incident polarization: (a) $n = 1.5$; (b) $n = 4.0$
grating power spectrums), suggesting that the dimensions of the surface-relief grating are consistent with both the scalar and thin-grating approximations. As expected, the power spectrums of the original surface-relief grating and its corresponding reverse parity grating agree more closely for the higher index substrate.

5.2.2 1-to-6 fanout grating

Figure 5.11 shows the phase profile of a single period of this grating. As for the 1-to-5 fanout grating, this grating is also symmetric. The specifications for this grating are given in [22]. The ideal power spectrum of this grating (i.e. calculated using a Fourier optics formalism) is shown in Figure 5.12. Unlike the power spectrum for the 1-to-5 fanout grating, the zero-order diffraction efficiency of this grating is suppressed. Consequently, the phase depth of this grating is $\pi$ (see Eq. (2.39)).

Note that in the ideal case the power spectrum contains 6 beams each of which contain essentially the same amount of power, as shown in Figure 5.12. The zero-order beam is not optimized to contain any useful power. Rather, it is minimized. For a 1-to-$N$ fanout grating where $N$ is an even number, this minimization is necessary if the grating is to be efficient and if the desired grating phase profile is to be binary (which is often the case since binary-phase elements are relatively simple to fabricate). Therefore, in the discussion below we consider the reconstruction errors for the zero-order and higher-order beams separately. As in the 1-to-5 fanout grating earlier, we consider the question of "How much can the grating period be reduced until period-to-wavelength effects become significant (i.e. for a particular optical system with a given set of performance tolerances)?"

If in a particular application the power contained in the zero-order beam is unimportant, then the reconstruction error associated with this order can be disregarded.
Figure 5.11: Phase profile (single period) of 1-to-6 fanout grating.
Figure 5.12: *Ideal power spectrum of 1-to-6 fanout grating.*
Table 5.1: *Reconstruction errors for 1-to-6 fanout grating for various period-to-wavelength ratios.*

<table>
<thead>
<tr>
<th>$\lambda/\Lambda$</th>
<th>$\Delta R[0]$</th>
<th>$\Delta R_{m\neq 0}^{m\neq 0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>19.7</td>
<td>0.701</td>
</tr>
<tr>
<td>6</td>
<td>4.46</td>
<td>0.313</td>
</tr>
<tr>
<td>25</td>
<td>0.743</td>
<td>0.0840</td>
</tr>
<tr>
<td>50</td>
<td>0.259</td>
<td>0.0316</td>
</tr>
</tbody>
</table>

In some applications, however, such as optical switching the power contained in the zero-order may be constrained to be below a certain threshold value. In such a situation the reconstruction error of the zero-order beam becomes extremely important.

Figures 5.13a through 5.13d plot the grating power spectrums, for the orders from -3 to +3, of the 1-to-6 fanout grating for period-to-wavelength ratios of 3, 6, 25 and 50, respectively. Also plotted (solid curves) are the ideal power spectrums. The index of refraction of the grating substrate is 1.5 in each of these cases and the incident polarization is TE. The physical etch depth for this index of refraction, for a phase depth of $\pi$ and a wavelength of 1.0 $\mu$m, is 1.0 $\mu$m. The reconstruction errors for each of these cases are summarized in Table 5.1. As expected, as the period-to-wavelength ratio increases the agreement between the actual spectrum with the ideal spectrum improves. At a period-to-wavelength ratio of 50 the agreement is fairly good, suggesting that the the grating characteristics can be accurately modelled using a Fourier optics formalism.

We next considered the polarization effects at a period-to-wavelength ratio of 50 and for a substrate index of refraction of 1.5. Figure 5.14 shows that these effects are very small. For a smaller period-to-wavelength ratio such as 6, these effects become appreciable. This point is conveyed in Figure 5.15.

Figure 5.16 plots the zero-order reconstruction error as a function of the period-to-
Figure 5.13: Power spectrum of 1-to-6 fanout grating for various period-to-wavelength ratios: (a) $\Lambda/\lambda = 3$; (b) $\Lambda/\lambda = 6$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.
Figure 5.13: Power spectrum of 1-to-6 fanout grating for various period-to-wavelength ratios: (c) $\Lambda/\lambda = 25$; (d) $\Lambda/\lambda = 50$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.
Figure 5.14: Power spectrum of 1-to-6 fanout grating - $TE$ (solid curve) and $TM$ (dashed curve) incident polarizations; $\Lambda/\lambda = 50; n = 1.5$. 
Figure 5.15: Power spectrum of 1-to-6 fanout grating - TE (solid curve) and TM (dashed curve) incident polarizations; \( \Lambda/\lambda = 6; n = 1.5 \).
Figure 5.16: Zero-order reconstruction error versus period-to-wavelength ratio for 1-to-6 fanout grating - TE (*) and TM (o) incident polarizations; $n = 1.5$. 
Figure 5.17: Maximum higher-order reconstruction error versus period-to-wavelength ratio for 1-to-6 fanout grating - TE (•) and TM (○) incident polarizations; n = 1.5.
wavelength ratio for both TE and TM incident polarizations and for a substrate index of refraction of 1.5. Figure 5.17 represents a similar plot for the maximum higher-order reconstruction error. Both the zero-order and higher-order reconstruction errors decrease as the period-to-wavelength ratio increases, as expected. Also, as the period-to-wavelength ratio increases the grating power spectrum becomes less sensitive to the incident polarization.

Figure 5.16 also shows that the zero-order reconstruction errors for TE and TM polarizations as a function of the period-to-wavelength ratio are nearly identical. Note, however, that we cannot determine whether the actual zero-order diffraction efficiencies for the TE- and TM-polarized cases are equal since, as noted before, this plot contains no information on the absolute power contained in each order. The higher-order reconstruction errors, on the other hand, are quite different for the TE and TM polarizations. For all values of the period-to-wavelength ratio the maximum higher-order reconstruction error for the TM-polarized case is lower than that for the TE-polarized case. Note that for this grating we do not observe an unexpectedly low reconstruction error at a moderate value of the period-to-wavelength ratio, as we did for the 1-to-5 fanout grating. The essential feature of Figure 5.17, however, is that the grating power spectrum is clearly dependent on the period-to-wavelength ratio.

At a period-to-wavelength ratio of 50, the power contained in the higher-orders are within approximately ±3%, corresponding to a maximum higher-order reconstruction error of 0.03, of the ideal values. However, the zero-order reconstruction error is approximately 26%. Although this value is relatively high, the deviation in absolute power is relatively small. Consider an optical switching application where the tolerance on the zero-order diffraction efficiency is that it be no more than half of the mean value of the ideal higher-order diffraction efficiencies. For such a system
tolerance a zero-order reconstruction error of 26% is clearly tolerable. In fact, a reconstruction error up to approximately 4.00, which occurs for a period-to-wavelength ratio of approximately 6.5 (assuming TE polarization), would adequately meet the above design specification. However, at a period-to-wavelength ratio of 6.5 the maximum higher-order reconstruction error is 0.267. Depending on the application, this value of reconstruction error may or may not be tolerable. For the optical switching application described above, a 27% error in the higher-orders still keeps all of the higher-order beams above a 50% threshold value. Consequently, for this hypothetical application the grating period can be decreased to approximately 6.5\(\mu\)m and still stay within the system design tolerances.

We next considered the effect of the substrate index of refraction on the grating power spectrum for a period-to-wavelength ratio of 50. Figure 5.18 plots the grating power spectrum for a TE-polarized incident wave and for a substrate index of refraction of 4.0. The physical etch depth of this grating is 0.167\(\mu\)m. This etch depth is smaller by a factor of 6 than the corresponding etch depth for a substrate with an index of refraction of 1.5; therefore, the grating fabricated in the higher index substrate should more closely approximate a 'thin' grating than a similar grating fabricated in the lower index substrate. Comparing Figure 5.18 with Figure 5.13, we see that the power spectrum of the grating fabricated on the higher index material does agree more closely with the ideal spectrum than does the grating fabricated on the lower index material. The maximum reconstruction errors are 0.259 and 0.0570 for the \(n = 1.5\) and \(n = 4.0\) cases, respectively.

Finally, we examined the similarity between the 1-to-6 fanout grating in Figure 5.11 and the corresponding reverse parity grating at a period-to-wavelength ratio of 50 and for substrate indices of refraction of \(n = 1.5\) and \(n = 4.0\). As shown in
Figure 5.18: Power spectrum of 1-to-6 fanout grating - $\Lambda/\lambda = 50; n = 4.0$; TE incident polarization. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.
Figure 5.19: Power spectrum of 1-to-6 fanout grating and corresponding reverse parity grating - $\Lambda/\lambda = 50$; TE incidence polarization: (a) $n = 1.5$; (b) $n = 4.0$
Figures 5.19a \((n = 1.5)\) and 5.19b \((n = 4.0)\), the grating power spectrum is relatively independent of the parity of the surface-relief grating, suggesting that the grating dimensions are consistent with the scalar and thin-grating approximations.

### 5.2.3 1-to-11 fanout grating

The phase profile of a single period of this grating is shown in Figure 5.20. This grating differs from the two previous gratings in that the ideal phase profile is not symmetric. Therefore, although the surface-relief profile of an actual grating implementing this ideal phase profile may be binary, the power spectrum of this grating will not be strictly symmetric. The design specifications for this grating were taken from [43]. The phase depth of the grating is 2.589 radians. The ideal power spectrum of this binary-phase grating is shown in Figure 5.21.

Figures 5.22a through 5.22d plot the grating power spectrums, for the orders from -5 to +5, of the 1-to-11 fanout grating for period-to-wavelength ratios of 5, 10, 25 and 50, respectively. Also plotted in each figure is the ideal power spectrum (solid curve), calculated using a Fourier optics formalism, of the binary-phase grating represented in Figure 5.20. The substrate index of refraction in each case is 1.5, corresponding to a physical etch depth of 0.8241\( \mu \)m, and the incident polarization is TE.

The lack of symmetry in the grating power spectrum is very apparent for small values of the period-to-wavelength ratio. As the period-to-wavelength ratio increases, the agreement between the actual grating power spectrum and the ideal symmetric grating power spectrum improves. The maximum reconstruction errors are 0.786, 0.446, 0.115 and 0.111 for period-to-wavelength ratio values of 5, 10, 25 and 50, respectively. The maximum reconstruction error at a period-to-wavelength ratio of 50, 0.111, is higher than those for the 1-to-5 and 1-to-6 (higher-orders) fanout gratings at
Figure 5.20: Phase profile (single period) of 1-to-11 fanout grating.
Figure 5.21: Ideal power spectrum of 1-to-11 fanout grating.
Figure 5.22: Power spectrum of 1-to-1 far out grating for various period-to-wavelength ratios: (a) $\Lambda/\lambda = 5$; (b) $\Lambda/\lambda = 10$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.
Figure 5.22: Power spectrum of 1-to-11 fanout grating for various period-to-wavelength ratios: (c) \( \Lambda/\lambda = 25 \); (d) \( \Lambda/\lambda = 50 \). Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.
the same period-to-wavelength ratio. This increased reconstruction error is probably due to the fact that the feature sizes for this grating are finer than for the two previous gratings. In general, as the degree of fanout increases, the feature sizes of the grating tend to decrease. By reducing the feature sizes, the high spatial frequency content of the grating is enhanced. Since the dimensions of the diffracting features of this grating are smaller than those of the two previous gratings, a larger value of the period-to-wavelength ratio is required in order for scalar diffraction theory to be reliably used to predict the grating diffraction behavior.

Polarization effects are addressed in Figures 5.23a and 5.23b. Plotted are the grating power spectrums for TE- and TM-polarized incident waves for period-to-wavelength values of 10 and 50 in 5.23a and 5.23b, respectively. The substrate index of refraction in each case is 1.5. Evidently, at a period-to-wavelength ratio of 10 the diffraction characteristics are quite sensitive to the incident polarization. At a period-to-wavelength ratio of 50, this sensitivity is lower.

Figure 5.24 plots the maximum reconstruction error of the grating as a function of the period-to-wavelength ratio for both TE- and TM-polarized incident waves and for a substrate index of refraction of 1.5. At small values of the period-to-wavelength ratio, the TM-polarized wave results in a significantly higher reconstruction error than the TE-polarized wave. As the period-to-wavelength ratio increases, the reconstruction error becomes less sensitive to the incident polarization, as shown also in Figure 5.23b.

5.2.4 Gaussian fanout grating

In the three previous sections several surface-relief phase gratings which implement ideal binary-phase gratings were analyzed. The binary-phase gratings were optimized
Figure 5.23: Power spectrum of 1-to-11 fanout grating - TE (solid curve) and TM (dashed curve) incident polarizations; $n = 1.5$: (a) $\Lambda/\lambda = 10$; (b) $\Lambda/\lambda = 50$
Figure 5.24: Maximum reconstruction error versus period-to-wavelength ratio for 1-to-11 fanout grating - TE (*) and TM (o) incident polarizations; $n = 1.5$. 
(once again, using a Fourier optics formalism) for a uniform splitting ratio. The first two gratings had surface-relief profiles which were symmetric. Consequently, their power spectrums were also symmetric. The surface-relief profile of the third grating did not contain any symmetry; consequently, its power spectrum did not contain any symmetry. In this section we consider another binary-phase grating whose surface-relief profile does not contain any symmetry. This grating, however, is optimized for a non-uniform splitting ratio, a Gaussian splitting ratio in particular. The design specifications for this grating are given in Section 3.1.1. The phase profile of a single period of this grating is shown in Figure 4.3. The grating phase depth is 2.450 radians.

Figures 5.25a through 5.25d plot the grating power spectrums, for the orders from -5 to +5, of the Gaussian fanout grating for period-to-wavelength ratios of 5, 10, 25 and 50, respectively. The substrate index of refraction in each case is 1.5, corresponding to a physical etch depth of 0.780μm. The incident polarization is TE. In all of the plots the solid curve corresponds to the ideal power spectrum calculated using a Fourier optics formalism, as described in Section 3.1.1. The most apparent feature of these plots is the asymmetry of the power spectrum, particularly at the lower values of the period-to-wavelength ratio. The asymmetry is less pronounced for larger values of the period-to-wavelength ratio, as expected.

We next considered the effect of the incident polarization on the grating power spectrum. Figure 5.26 plots the grating power spectrum for TE and TM incident polarizations, a period-to-wavelength ratio of 50 and for a substrate index of refraction of 1.5. The profiles are nearly identical, suggesting that the scalar approximation is well-satisfied in this regime. However, the fact that the spectrum is not symmetric (although strictly speaking it is only symmetric in the limit of an infinite substrate index of refraction), suggests that the deviation between reality and the ideal case is
Figure 5.25: Power spectrum of Gaussian fanout grating for various period-to-
wavelength ratios: (a) $\Lambda/\lambda = 5$; (b) $\Lambda/\lambda = 10$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.
Figure 5.25: Power spectrum of Gaussian fanout grating for various period-to-wavelength ratios: (c) $\Lambda / \lambda = 25$; (d) $\Lambda / \lambda = 50$. Rigorous calculation - dashed curve; Fourier optics prediction - solid curve.
Figure 5.26: Power spectrum of Gaussian fanout grating - TE (solid curve) and TM (dashed curve) incident polarizations; $\Lambda/\lambda = 50$; $n = 1.5$. 
Figure 5.27: *Maximum reconstruction error versus period-to-wavelength ratio for Gaussian fanout grating* - TE (•) and TM (○) incident polarizations; \( n = 1.5 \).

due to the finite thickness of the grating etch depth.

Figure 5.27 plots the maximum reconstruction error of the grating as a function of the period-to-wavelength ratio for both TE and TM incident polarizations and for a substrate index of refraction of 1.5. As shown, the deviation between the maximum reconstruction errors for the TE- and the TM-polarized waves at period-to-wavelength ratios above 7 is nearly constant. This result is in contrast to the case for the 1-to-5, the 1-to-6 and the 1-to-11 fanout gratings, all of which showed substantial deviations between the two polarizations. As the period-to-wavelength ratio increases, the maximum reconstruction error decreases. The rate at which this error decreases,
however, is lower than the corresponding rate for the three previous gratings we considered. Furthermore, the maximum reconstruction errors of 0.269 and 0.321 at a period-to-wavelength ratio of 50 for TE and TM polarizations, respectively, are higher than those for the three gratings considered earlier.

5.3 Summary

In this chapter the diffraction behavior of surface-relief structures which implement ideal binary-phase gratings as a function of the period-to-wavelength ratio and the grating substrate index of refraction was considered. It was shown that the grating behavior deviates significantly from that predicted from a Fourier optics formalism when the period-to-wavelength ratio is small. The degree of deviation and the specific form of the deviation depend on the specific surface-relief grating being considered. Practically, a sufficiently large grating period must be chosen such that the deviation is within a specified system tolerance. One must also consider the physical etch depth of the surface-relief grating. The issue here is the validity of the thin-grating approximation. Several examples were given to qualitatively illustrate the consequences of both the scalar and the thin-grating approximations. The essential point of this chapter is that the physical limits (i.e. grating dimensions) which are implicitly imposed by the grating phase optimization procedure (i.e. Fourier-based phase optimization) must be considered when one practically fabricates surface-relief structures which implement the optimized phase profiles.
Chapter 6

Conclusions and Future Research

The goal of this thesis project has been both to investigate various phase optimization techniques for the design of free-space optical interconnects and to address the theoretical and practical issues which arise in the design and fabrication of the surface-relief structures which implement the free-space optical interconnects. To date, most of the research effort in the field of free-space optical interconnects has been directed towards the development of more efficient and powerful design algorithms. These algorithms involve the optimization of the phase profiles which are to be implemented by the surface-relief structures.

Optimum phase profiles have been determined which achieve 1001 X 1001 fanouts. However, the corresponding grating power spectrums are extremely sensitive to the phase depth of the grating. A simple model was presented to determine this sensitivity. It was shown that as the degree of fanout increases the sensitivity of the zero-order reconstruction error increases, while the corresponding sensitivity of the higher-order reconstruction errors decreases. Therefore, when possible, optical systems utilizing free-space optical interconnects should be designed around the zero-order beam.

Two phase optimization techniques, simulated annealing and phase retrieval, were used to design binary and multi-level phase profiles which implement a Gaussian
fanout. The method of simulated annealing is well-suited for the optimization of binary-phase profiles, while the method of phase retrieval is more applicable to the optimization of multi-level phase profiles. Binary surface-relief phase gratings are much easier to fabricate than multi-level surface-relief phase gratings, although the latter are more efficient and have lower reconstruction errors than the former.

Surface-relief structures were fabricated on fused silica substrates to implement the ideal binary-phase profiles optimized by simulated annealing. The effects of etch depth errors and linewidth errors in the fabrication process were analyzed with a simple model. Experimental measurements of etch depth and, in some cases, linewidth errors were made for the fabricated surface-relief structures. Based on these error measurements and the simplified model mentioned above, the experimentally measured power spectrums agreed well with the expected power spectrums.

Essentially all phase optimization algorithms, such as simulated annealing and phase retrieval, are based on the formalism of Fourier optics. This formalism is derived from Maxwell’s equations by making two approximations, the scalar approximation and the thin-grating approximation. These approximations result in an idealized grating surface-relief profile, particularly a grating whose period is infinite in extent and whose etch depth approaches zero. Clearly, actual surface-relief profiles will differ from this idealized profile. Consequently, Fourier optics can be reliably used to describe the grating diffraction of actual surface-relief phase gratings when the dimensions of the grating are consistent with both the scalar and thin-grating approximations (i.e. within a specified error criterion, such as a maximum reconstruction error of 10%).

The grating diffraction behavior of binary surface-relief phase gratings was analyzed as a function of the period-to-wavelength ratio and for two different substrate
indices of refraction. The intent of this analysis was to qualitatively illustrate the regimes of grating diffraction where the scalar and thin-grating approximations both are and are not well-satisfied. A rigorous vector coupled-wave formulation of grating diffraction was used to perform the analysis. The results of this analysis were compared with those predicted by the use of a Fourier optics formalism. A general trend of increasing reconstruction error for decreasing values of the period-to-wavelength ratio was observed. Furthermore, as the index of refraction of the grating substrate increases, the grating diffraction behavior agrees more closely with that predicted by the use of a Fourier optics formalism.

Much work needs to be done before free-space optical interconnects can be reliably incorporated into large-scale optical systems. First, fabrication technology, particularly etch depth control and linewidth reproduction, must be improved. Presently, fabrication technology cannot meet the stringent requirements on process control posed by free-space optical interconnects implementing large fanouts. Consequently, the realization of surface-relief structures which implement large fanouts may ultimately be limited by fabrication capability. Second, a comprehensive theoretical model of grating diffraction accounting for more general types of fabrication errors, such as non-uniform etch depths, non-uniform linewidth errors and non-vertical sidewalls, should be developed. Such a model could quantify the sensitivity of a grating's performance to the various errors in fabrication (i.e. tell us which source of error has the most impact on the grating diffraction behavior). Furthermore, such a model would not only help in explaining experimentally observed phenomena, but also help in evaluating the fabrication process.

A third area of future research would be to further investigate the behavior of binary surface-relief phase gratings as a function of the period-to-wavelength ratio
and the substrate index of refraction. Such an investigation might provide some "intuition" for the grating diffraction process in the regime where a rigorous analysis is required. It is questionable whether phase optimization techniques will eventually be based on rigorous diffraction theories; however, an improved understanding of grating diffraction in the 'rigorous regime' will undoubtedly enhance our understanding of grating diffraction in the 'Fourier regime' since the 'Fourier regime' is simply a limiting case of the 'rigorous regime'. Finally, a natural progression from the analysis of binary surface-relief phase gratings is to the analysis of multi-level surface-relief phase gratings.
Bibliography


[51] P. N. Everett and W. C. Goltsos, to be published.


