

**DESIGN OF DIFFRACTIVE OPTICAL ELEMENTS THROUGH LOW-  
DIMENSIONAL OPTIMIZATION**

A Thesis  
Presented to  
The Academic Faculty

By

David W. Peters

In Partial Fulfillment  
of the Requirement for the Degree of  
Doctor of Philosophy in Electrical and Computer Engineering

Georgia Institute of Technology  
July 2001

Copyright © 2001 by David W. Peters

**DESIGN OF DIFFRACTIVE OPTICAL ELEMENTS THROUGH LOW-  
DIMENSIONAL OPTIMIZATION**

A Doctoral Dissertation

By

David W. Peters

Submitted to the Faculty of the  
School of Electrical and Computer Engineering  
Georgia Institute of Technology

Dr. William Hunt


Ph.D. Dissertation Advisor


June 8, 2001

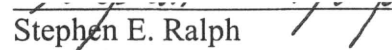


**DESIGN OF DIFFRACTIVE OPTICAL ELEMENTS THROUGH LOW-  
DIMENSIONAL OPTIMIZATION**

Approved:

  
\_\_\_\_\_  
William D. Hunt, Chair

  
\_\_\_\_\_  
John A. Buck

  
\_\_\_\_\_  
Stephen E. Ralph

Date Approved 6/29/01

## Acknowledgements

I would like to thank the members of my committee, Profs. John Buck, William Hunt, Stephen Ralph, and Andrew Zangwill. I owe an extra thanks to Prof. Hunt for taking me on as a graduate student midstream.

I wish to thank my fellow graduate students who graduated before me: Dr. John Callahan and Prof. Albert Titus, with whom I shared a lab for many years, and Drs. Houston Irby, Stan West, and Jim Hoffman. I also wish to thank my fellow graduate students and Student Government members who will soon receive their doctorates, Caroline Clower and Doug Britton. All of these friends helped me reach my goal without first going insane.

Finally I wish to thank my parents who never questioned why it took me so long to graduate, and most importantly to Christine Hung, whose patience and support have been priceless.

# Table of Contents

<b>CHAPTER 1. MOTIVATION, PROBLEM STATEMENT .....</b>	<b>1</b>
<b>CHAPTER 2. BACKGROUND .....</b>	<b>5</b>
2.1 APPLICATIONS.....	5
2.1.1 <i>Historical Background</i> .....	8
2.1.2 <i>Electrical and Optical Interconnects</i> .....	9
2.1.3 <i>Input and Output Optical Couplers</i> .....	11
2.1.4 <i>Surface-Relief Gratings</i> .....	13
2.1.5 <i>Volume Gratings</i> .....	14
2.1.6 <i>Other Diffractive-Optic Structures</i> .....	15
2.2 METHODS OF SOLVING THE ELECTRIC FIELD.....	15
2.3 COMMERCIAL APPLICATIONS.....	19
2.4 DESIGN THROUGH OPTIMIZATION.....	20
<b>CHAPTER 3. THEORY .....</b>	<b>24</b>
3.1 SOLUTION OF THE ELECTRIC FIELD.....	24
3.1.1 <i>Derivation of Equations</i> .....	25
3.2 THE FINITE DIFFERENCE METHOD .....	28
3.2.1 <i>Direct solution method</i> .....	34
3.2.2 <i>The Alternating Direction Implicit (ADI) Method</i> .....	37
3.2.3 <i>Comparison of the ADI and direct-solve methods</i> .....	39
3.3 OPTIMIZATION .....	41
3.3.1 <i>Low-Dimensional Optimization</i> .....	43
3.3.2 <i>Simulated Annealing</i> .....	46
<b>CHAPTER 4. TESTING AND VERIFICATION OF ALGORITHM.....</b>	<b>51</b>
4.1 ELECTROMAGNETIC FIELD SOLVER ALGORITHM TESTING .....	51
4.1.1 <i>Comparison of iterative and direct methods</i> .....	51
4.1.2 <i>Mesh fineness</i> .....	64
4.1.3 <i>Comparison of results to published data</i> .....	67
4.2 SIMULATED ANNEALING (SA) ALGORITHM TESTING .....	69
<b>CHAPTER 5. EXPERIMENTAL RESULTS.....</b>	<b>77</b>

5.1	GRATING-GRATING INTERACTION .....	77
5.2	FORTY FIVE DEGREE MIRROR COUPLERS .....	80
5.3	FOCUSING VOLUME GRATING COUPLERS .....	89
5.4	TRANSMISSIVE DIFFRACTIVE ELEMENT .....	94
<b>CHAPTER 6. SUMMARY OF RESULTS AND CONCLUSION .....</b>		<b>98</b>
<b>CHAPTER 7. REFERENCES .....</b>		<b>102</b>
<b>APPENDIX .....</b>		<b>109</b>

## Table of Figures

<i>FIGURE 1. A generic diffractive optical element .....</i>	<i>3</i>
<i>FIGURE 2. Channel waveguide with grating .....</i>	<i>6</i>
<i>FIGURE 3. Two outcoupling structures.....</i>	<i>7</i>
<i>FIGURE 4. Outcoupled beam resulting from uniform grating .....</i>	<i>8</i>
<i>FIGURE 5. Coordinate system used in all designs .....</i>	<i>27</i>
<i>FIGURE 6. Typical configuration.....</i>	<i>30</i>
<i>FIGURE 7. Graphical demonstration of a. three-point operators and b. a five-point operator .....</i>	<i>31</i>
<i>FIGURE 8. Absorbing boundary conditions .....</i>	<i>34</i>
<i>FIGURE 9. Matrix numbering for direct-solution method.....</i>	<i>35</i>
<i>FIGURE 10. Banded-matrix structure of direct-solution method .....</i>	<i>36</i>
<i>FIGURE 11. Optimization flow chart.....</i>	<i>41</i>
<i>FIGURE 12. Example of search space for optimization algorithm.....</i>	<i>42</i>
<i>FIGURE 13. Example of device optimization.....</i>	<i>43</i>
<i>FIGURE 14. Fit of quadratic curve to selected points.....</i>	<i>46</i>
<i>FIGURE 15. Simulated annealing flow chart.....</i>	<i>47</i>
<i>FIGURE 16. Waveguide used to compare direct solution and implicit solution techniques.....</i>	<i>53</i>
<i>FIGURE 17. Electric field amplitude of unperturbed waveguide .....</i>	<i>55</i>
<i>FIGURE 18. 200 column example.....</i>	<i>57</i>
<i>FIGURE 19. Region of interest in 200 column example .....</i>	<i>58</i>
<i>FIGURE 20. 400 column example.....</i>	<i>59</i>
<i>FIGURE 21. Region of interest in 400 column example .....</i>	<i>60</i>
<i>FIGURE 22. Comparison of one row of data using direct and implicit methods.....</i>	<i>62</i>
<i>FIGURE 23. Percent error of three iterative cases compared to direct-solve case for 600x90 waveguide..</i>	<i>63</i>
<i>FIGURE 24. Four sample waveguides used to measure error.....</i>	<i>65</i>
<i>FIGURE 25. Percent error for four waveguides as a function of number of mesh points per free-space wavelength.....</i>	<i>66</i>
<i>FIGURE 26. Grating described by Ostrowsky and Jacques and used for algorithm verification .....</i>	<i>67</i>
<i>FIGURE 27. Computation region of waveguide showing simulation and absorbing regions (a.), and the imaginary component of the refractive index along the row for the SWR computations.....</i>	<i>71</i>
<i>FIGURE 28. Amplitude profile in direction of propagation resulting from insufficient absorbing region...</i>	<i>72</i>
<i>FIGURE 29. Amplitude profile in direction of propagation resulting from excessive absorbing region.....</i>	<i>73</i>
<i>FIGURE 30. Amplitude profile in direction of propagation resulting from optimal absorbing region .....</i>	<i>73</i>



<i>FIGURE 31. Results of optimization of SWR by varying the maximum absorber value for three different absorber widths .....</i>	<i>74</i>
<i>FIGURE 32. SWR as function of iteration number of the simulated annealing algorithm .....</i>	<i>75</i>
<i>FIGURE 33. Configuration of example of grating-grating interaction .....</i>	<i>78</i>
<i>FIGURE 34. Variation of power coupled into the cover region by two gratings as a function of grating separation distance and grating offset relative to each other .....</i>	<i>79</i>
<i>FIGURE 35. Generic forty-five degree mirror configuration.....</i>	<i>80</i>
<i>FIGURE 36. Resin mirror design variables.....</i>	<i>82</i>
<i>FIGURE 37. Resin droplet limited by surfaces of differing surface tension .....</i>	<i>83</i>
<i>FIGURE 38. Figure of merit at each iteration of the annealing algorithm for the resin mirror design .....</i>	<i>85</i>
<i>FIGURE 39. Figure of merit as a function of the radius of curvature of the mirror.....</i>	<i>86</i>
<i>FIGURE 40. Figure of merit as a function of the distance from the waveguide to the base of the mirror....</i>	<i>87</i>
<i>FIGURE 41. Figure of merit as a function of the depth of the groove below the guiding layer .....</i>	<i>87</i>
<i>FIGURE 42. Figure of merit as a function of the tilt angle of the mirror .....</i>	<i>88</i>
<i>FIGURE 43. Results from seven-dimensional optimization of a volume grating .....</i>	<i>91</i>
<i>FIGURE 44. Polynomial curves for focused Gaussian output beam .....</i>	<i>92</i>
<i>FIGURE 45. Comparison of amplitudes of ideal beam and beam resulting from design process .....</i>	<i>93</i>
<i>FIGURE 46. Diffractive element to be optimized.....</i>	<i>94</i>
<i>FIGURE 47. Figure of merit as a function of duty cycle and grating depth for a transmissive diffractive element.....</i>	<i>96</i>
<i>FIGURE 48. Figure of merit as a function of duty cycle and grating depth for a transmissive diffractive element (two-dimensional view).....</i>	<i>97</i>
<i>FIGURE A1. Duty cycle of less than 0.5 .....</i>	<i>108</i>
<i>FIGURE A2. Duty cycle of greater than 0.5 .....</i>	<i>109</i>

## Table of Tables

<i>TABLE 1. Comparison of direct solution and iterative solution processing times .....</i>	<i>54</i>
<i>TABLE 2. Comparison of results for 90x600 waveguide test with optimized compilation .....</i>	<i>61</i>
<i>TABLE 3. Comparison of theoretical, experimental, and simulated data from surface-relief grating .....</i>	<i>68</i>
<i>TABLE 4. Allowed range for variables in design of angled mirror coupler.....</i>	<i>84</i>
<i>TABLE 5. Optimal design values for resin mirror device.....</i>	<i>88</i>

## Notation

$\epsilon$	permittivity
$\epsilon_r$	relative permittivity
$\epsilon_0$	permittivity of free-space
$\lambda_0$	free-space wavelength of light
$\Lambda$	grating period
$\mu$	permeability
$\rho$	volume charge density
$\omega$	angular frequency of electromagnetic field
$c$	speed of light in a vacuum
$f$	frequency of electromagnetic field
$k_0$	wave number in free-space
$M$	number of columns of grid points in z direction
$N$	number of rows of grid points in x direction
$n$	refractive index
$\Delta n$	refractive index modulation
$N_{\text{eff}}$	effective index of a guided mode
<b>E</b>	vector electric field, without time-dependence
<b>H</b>	vector magnetic field, without time-dependence
<b>D</b>	vector electric displacement, without time-dependence
<b>B</b>	vector magnetic flux density, without time-dependence
$\tilde{E}$	vector electric field, without time-dependence
$\tilde{H}$	vector magnetic field, without time-dependence
$\tilde{D}$	vector electric displacement, without time-dependence
$\tilde{B}$	vector magnetic flux density, without time-dependence

$\hat{\nabla}^2$  Laplacian for coordinate system normalized by  $k_0$

ADI alternating direction implicit

dc duty cycle

FDM finite difference method

fft fast Fourier transform

FoM figure of merit

SA simulated annealing

SWR standing wave ratio

TE transverse electric

TM transverse magnetic

## Summary

The simulation of diffractive optical structures allows for the efficient testing of a large number of structures without having to actually fabricate these devices. Various forms of analysis of these structures have been done through computer programs in the past. However, programs that can actually design a structure to perform a given task are very limited in scope. Optimization of a structure can be a task that is very processor-time intensive, particularly if the optimization space has many dimensions. This thesis describes the creation of a computer program that is able to find an optimal structure while maintaining a low-dimensional search space, thus greatly reducing the processor time required to find the solution. The program can design the optimal structure for a wide variety of planar optical devices that conform to the weakly-guiding approximation with an efficient code that incorporates the low-dimensional search space concept. This work is the first use of an electromagnetic field solver inside of an optimization loop for the design of an optimized diffractive-optic structure.



## Chapter 1. Motivation, Problem Statement

Diffractive optical elements play an important role in data-transfer applications by offering a compact method for redirecting light in a system. The ability to integrate a diffractive grating directly on a waveguide or other device not only saves the space of the bulk-optic equivalent, but can also eliminate the need for alignment of optical elements after the initial fabrication.

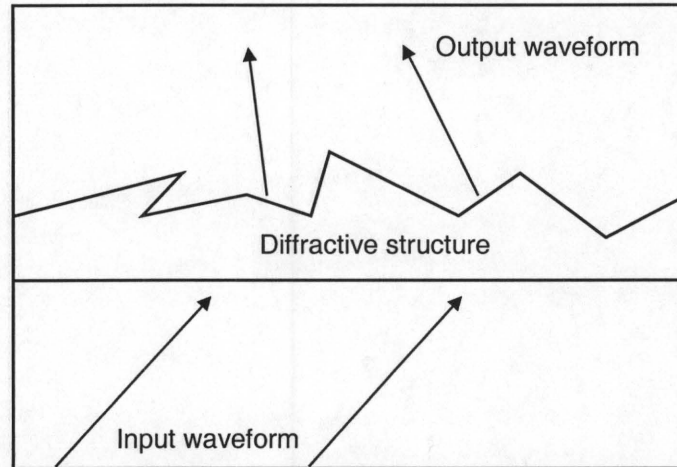
These diffractive elements may be quite complex in their design, with more complex designs allowing for greater freedom in the direction and shape of the redirected beam. These more advanced structures are unable to be designed by analytical methods, and thus numerical methods are required. Analysis of a given diffractive structure is relatively complex, but for planar devices is not overwhelming. However, to design such a structure to perform a desired task when no analytical method is available can be quite challenging. For the class of structure designed by the program described in this thesis there was no previous design program that could determine the optimal structure for a given task.

Some commercially available software offer a method for the design of relatively simple single diffractive gratings, however limitations exist even on the design of these elements as a result of the analysis method used. In the commercial programs the grating is assumed to have infinite extent and a fixed period, duty-cycle, and depth along its length. Most commercial programs only solve for the diffractive orders of free-space to free-space modes; guided modes cannot be analyzed. In addition, these programs cannot design an optimal structure. In the literature are examples of using an analysis method

wrapped inside a simulated annealing loop to find such an optimal structure [1,2].

However, the programs are again limited by the central analysis routine used, thus limiting the type structure that can be analyzed to the very specific structure for which the program was designed.

The goal of my work has been to create a computer program that will design the optimal optical elements in two-dimensions without the limitations of previous programs. A diffractive structure such as the one seen in Figure 1 will be designed by varying parameters of the structure until the particular structure is found that optimizes the output waveform for the given input waveform. The limitations on the parameters of the optical structure will be set by the user at the beginning of the design process. The optical element will not be limited to one type of structure, nor will it be limited to only guided-wave or free-space modes. The structure may be a surface relief or volume grating on a waveguide, a grating for redirecting free-space modes, a reflective mirror, or multiple gratings whose relative distance to each other does not allow for their independent analysis. In this work the devices are limited to planar devices with no variation in one dimension, in the example illustrated in Figure 1 there would be no variation in the dimension perpendicular to the surface of the page.



**FIGURE 1.** A generic diffractive optical element

This thesis describes the development of such a program and its testing and implementation. The program consists of a central analysis routine to analyze the electric fields around the diffractive structure. This routine is embedded inside an optimization algorithm that allows for changes to the structure properties in order to find the best solution.

The central analysis routine consists of solving for the scalar electric field in the region of the diffractive structure. No assumptions are made as to periodicity, uniformity, or length of the grating in the direction of propagation. The scalar Helmholtz equation, discretized using the finite difference method, accomplishes this task. Using this method of analysis, both guided-wave and free-space modes traveling in any direction may be analyzed. In this respect, the work described here is unique to all previous works that have used optimization loops to design diffractive structures. The ability to optimize a structure that is not subject to a long list of limitations on the design results in a program of unique and powerful ability.

As a result of the analysis method used, there are limitations on the devices that may be analyzed. A weakly-guiding approximation is assumed. This assumption precludes the design of structures with high index of refraction variations. Also, structures that involve an anisotropic or non-linear material cannot be designed. Due to the use of the scalar wave equation mode conversion cannot be modeled either.

Optimization loops have been used in optical applications before. The simulated annealing algorithm is the particular optimization routine chosen for this work. A unique method will be shown that allows for the design of optical structures using the annealing algorithm. It finds the optimal solution over a large solution space while maintaining a low-dimensional search space. By keeping the number of dimensions in the search space low, the optimal design may be found with a reasonable amount of computation time.

The chapters of this work will describe the algorithms involved in this program, the verification and testing of sample structures, and the design of several devices. Chapter 2 will describe the background of the diffractive optic structures and the common methods used to analyze them. Chapter 3 will present the theory of the scalar Helmholtz equation, the finite difference method, and the simulated annealing algorithm. Chapter 4 will relate the steps taken in testing and verifying the analysis and optimization portions of the algorithm. In Chapter 5 some sample results will be given that illustrate a sampling of the broad type of devices that can be designed using this computer program. Finally, Chapter 6 will summarize the results and offer conclusions and recommendations for the continuance of this work.



## **Chapter 2. Background**

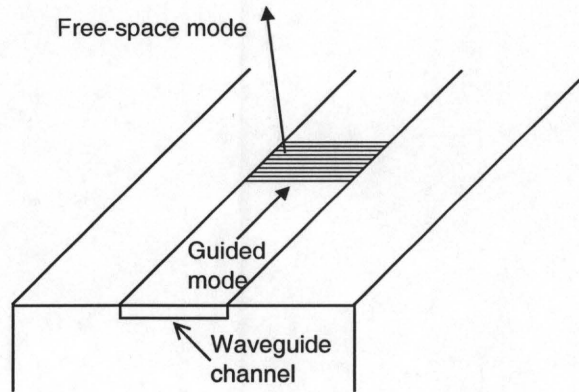
The purpose of this chapter is to introduce some of the background necessary to lay the groundwork for future chapters. This includes the historical development of some of the numerical methods used in my research along with other popular methods that were not chosen. Of particular interest are the development of finite difference Helmholtz equation for solving the electromagnetic field, and the simulated annealing algorithm for finding a global maximum of a function. Other analysis routines are also discussed along with their limitations. Preceding this discussion is the background of some of the applications for the optical devices that necessitate the use of the more complex and time-consuming of these algorithms and why they are becoming increasingly important.

### **2.1 Applications**

The devices which the computer program that I have written can simulate cover a wide range of optical applications. Originally the code was written to simulate waveguide input/output grating couplers as the one shown in Figure 2. Since the waveguiding section of the channel waveguide is often a wide-region it can be reasonably modeled as a planar waveguide. The simplification greatly simplifies the code that may be used and the processing time required, as no variation is assumed in the third dimension. The code developed for this problem was found to be useful for a broader range of diffractive-optic devices with some further adjustments to the boundary



conditions and the code for entering device structure parameters. I will first talk about the waveguide coupler applications in some detail.

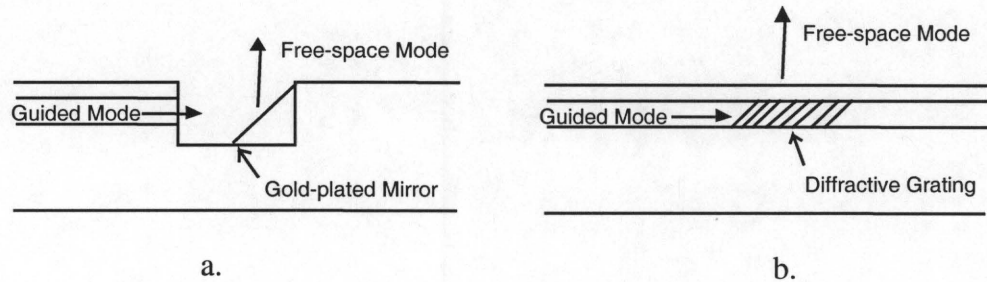


**FIGURE 2.** Channel waveguide with grating

The increase in speed in computer chips has resulted in a strain on the capacity of electrical interconnects that transmit clock signals and data on a chip, between chips, and between boards. One possible solution to the problem is to replace electrical interconnects with optical interconnects. Optical interconnects, although not without their own drawbacks, are free of the problems that limit electrical interconnects, particularly at longer distances of over a few centimeters.

Optical interconnects may be guided-wave or free-space interconnections. Each method has its own advantages and disadvantages. A hybridization of the two modes may be used, but then a method of converting from a guided mode to a free-space mode and from a free-space to guided mode must be devised. There are several ways to accomplish this transition. These methods each have their own weaknesses. The work I propose here investigates two of these methods, mirrors angled at roughly forty-five

degrees as illustrated in Figure 3a, and diffractive grating structures as illustrated in Figure 3b.



**FIGURE 3.** Two outcoupling structures

The parameters of these devices must be chosen carefully and enough degrees of freedom allowed to produce a desired output. A grating coupler may be designed to better accomplish a given task by allowing the parameters of the grating to vary in the direction of propagation of the guided mode of the waveguide. A uniform grating, i.e. a grating with no variation in the index modulation ( $\Delta n$ ), grating period, depth, or tilt angle, will result in an outcoupled beam whose power decreases exponentially with distance along the waveguide since the outcoupled power is proportional to the strength of the mode inside the guide as shown in Figure 4. The characteristics of this beam are not as desirable as a Gaussian beam. Likewise, the mirror couplers have parameters that affect the form that the outcoupled beam takes. A highly-diverging beam results when a simple flat mirror face is used.

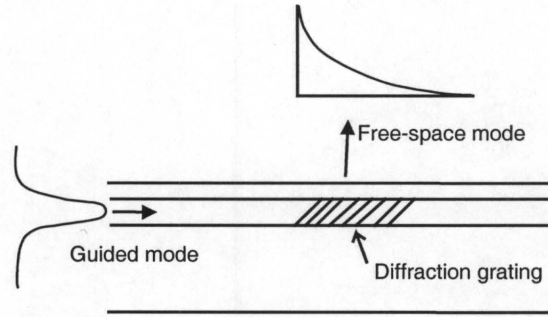


FIGURE 4. Outcoupled beam resulting from uniform grating

No computer program currently exists that allows a user to design a grating or mirror by allowing an algorithm to determine the optimal design of a device within pre-determined limits set by the user. The goal of my research was to develop such a program. The program must be able to allow the user to design a grating or mirror structure to produce an outcoupled beam of a desired shape and propagation direction. The only inputs should be the desired output and the limitations on the parameters of the waveguide grating or mirror. This requirement limits the methods that may be used to evaluate the electromagnetic fields of the guided and free-space modes as some methods make assumptions about the direction of propagation or assume a periodic uniform grating.

### 2.1.1 Historical Background

For the past three decades the optical fiber has been a commonly-used method of transferring digital data by way of guided optics over long distances. The success of fiber optics is encouraging as we look for extensions of this technology into interconnects that cover much smaller distances.

Optical waveguides offer a means of transporting such optical information and have been studied for many years. Although the most familiar form of guided-wave optical information transmission is the optical fiber, planar devices are also important since the fabrication technology overlaps with the processes used in integrated circuit fabrication. As the optical fiber has become a standard method of transmitting data over large distances, integrated optics could become equally important in the future as a means of transmitting large amounts of data over shorter distances. An example would be board-to-board or even chip-to-chip data transfers inside a computer. The number of applications for these optical waveguides are growing as a result of the increasing speed and size of digital computer chips where optical alternatives are now more attractive than the electrical counterpart.

Over the time period that optical fibers and integrated optical waveguides have come to be used with increasing frequency, a similar increase in the interest in numerical modeling and design of these devices has occurred.

Likewise, diffractive optics have been used extensively for many years. The diffractive optic equivalents of lenses and prisms offer compact and economical options in many optical designs. The manipulation of the diffractive orders in these devices offer additional opportunities to utilize these devices in ways that their bulk counterparts cannot.

### ***2.1.2 Electrical and Optical Interconnects***

Metal electrical interconnections have been the standard method for information transfer in VLSI circuit chips. However, speeds have increased to 1GHz on chips for home computers and feature sizes have decreased to below  $0.25\mu\text{m}$  on these chips. As



these trends continue, strains will develop in the use of these electrical interconnections. The bandwidth that may be carried over a given electrical interconnect is limited by the resistance, inductance, and capacitance of the link. Mutual inductance leads to crosstalk between lines, a problem that increases at higher frequencies. As scaling continues to shrink feature sizes the increase in current density in these interconnects will lead to increased problems with electromigration [3]. The RC charging times suffered by electrical interconnects make them impractical for distances over approximately 10cm. As clock frequencies continue to increase this critical distance will grow shorter. Particularly with clock signals where skew is of particular importance the limitations of electrical interconnects are already being observed. The need to distribute a synchronized clock signal across a long distance bars the use of electrical interconnects such as with distributed processors on multiple chips or boards [4].

Some of the primary drawbacks that effect electrical interconnects are frequency dependent. As frequencies increase in future systems these limitations will cause difficulties in data transfer in computer systems. With increasing frequency the skin effect causes the current to flow in an increasing thin layer at the surface of the conductor thus increasing the current density near the surface. Electrical lines also act as antennas that may both transmit and receive leading to crosstalk between lines; this effect also increases with frequency. It has been shown [5] that the limitations on electrical interconnects are a function of their aspect ratio, the ratio of their length to their cross-sectional area. Further miniaturization is therefore not a panacea, since decreasing chip sizes to decrease the length between chip locations will require a similar decrease in the area of the interconnects themselves.



### ***2.1.3 Input and Output Optical Couplers***

As in the case of electrical interconnects there are drawbacks to optical interconnects that must also be considered. Optical sources cannot be fabricated in silicon, the material of choice for VLSI circuits. Thus circuits must be a hybridization of different materials. However, multiple quantum well GaAs devices have been flip-chip bonded to VLSI circuits [6]. Flip-chip bonding is becoming increasingly common; Lucent operates a foundry for making these devices. In a circuit with optical interconnects it will become necessary to convert the signal from electrical-to-optical and optical-to-electrical. This drawback is less cumbersome if the amount of optical power required can be minimized. Each connection in the optical path should therefore be as lossless as possible. Each input or output coupler can be designed to minimize the amount of loss in the conversion between the free-space and guided modes. Another drawback to optical features on a largely electrical circuit is the packaging and fabrication differences. To mitigate this problem the optical interconnects should be fabricated to cause the least interference with the electrical feature fabrication steps.

For the long-distance interconnections for clock signals mentioned in the previous section guided-wave optical interconnects have already been used in a 1GHz clock-signal distribution system between Cray supercomputer boards [4]. Volume gratings acted as couplers in a UV-sensitive polyimide waveguide 45cm in length. The need to develop new interconnect technologies, including optical interconnects, is recognized as necessary to overcome the performance limitations of current technology [7].

One of the problems with optical interconnects is the loss of energy at each conversion point between guided and free-space modes. A look at the different forms

that input and output couplers may take is in order with a focus on the limitations of each method.

A common method of coupling light into a waveguide in the lab is through prism coupling, which is achieved by placing a prism over a waveguide [8]. The prism couples a beam in the prism into the guided mode of the underlying waveguide through the coupling of energy from the evanescent tail of the beam that undergoes total internal reflection into a guided mode in the waveguide. Using prism coupling outside the lab is impractical as a result of the mechanical devices needed to press the waveguide and the prism into close proximity. The prism is also bulky and non-planar. Furthermore, the mode that the light is coupled into as it exits the waveguide and prism may not be a desirable shape. The direction of the outcoupled beam will also be determined by the angles of the prism; for the user to pick a particular outcoupled beam angle would be difficult.

End-fire coupling is also possible for coupling into or out of a waveguide. Outcoupling with this method has the drawback that all of the light must enter or exit at an end of the waveguide; this will likely mean that the outcoupling must also occur at a cleaved edge of the sample, which may be inconvenient. It also precludes the possibility of outcoupling a percentage of the light from one source at several locations along the waveguide as one would like to do for clock distribution on a chip.

In a similar manner angled mirrors have been used to couple light into and out of waveguides. Similar to the end-fire coupler the angled mirrors out-couple all of the energy in one location. However, they do allow the outcoupled light to exit normal to the waveguide surface and allow out-coupling at locations other than a cleaved edge.

Gratings that perform guided-mode to free-space radiation mode conversion can take several forms. The two basic categories are surface-relief gratings and volume gratings. In either type of coupler the grating may exist either in the waveguide layer itself or in an adjacent layer. If the grating is in a layer of the waveguide structure other than the actual guiding layer then the coupling is achieved through the interaction of the grating with the evanescent fields that lie outside the guiding layer. These evanescent fields decay quickly, and the coupling of power from the guiding mode to a radiation mode is not as strong.

#### ***2.1.4 Surface-Relief Gratings***

A surface-relief grating involves the cutting away of one layer of the waveguide or substrate layer to form a grooved pattern. In a periodic surface-relief grating these cut-outs form regularly spaced teeth. These teeth may be rectangular in shape, blazed, where the sides of the teeth are not orthogonal to the direction of propagation in the waveguide, or any other shape. The cutting away to form the teeth is achieved by using lithographic methods common in the fabrication of semiconductor devices with dry and wet etching and electron beam lithography all used for this purpose.

At times a double surface-relief grating is used to achieve a desired output. One surface-relief grating may be used to outcouple light from a waveguide to a free-space mode. A short distance away, on top of another glass or polymer layer, a second surface-relief grating then focuses the light. Past work [9] has tended to ensure that the layer between the two gratings is sufficiently thick that the two gratings operate independently from one another. The second grating only interacts with the far-field result from the first

grating. However, it could be beneficial for the two gratings to be brought into closer contact as constructive interaction could lead to an improved output.

### **2.1.5 Volume Gratings**

Volume gratings in polymers are common devices. The volume grating formation involves a writing procedure in the photosensitive polymer. Interference fringes form a sinusoidally varying exposure of the polymer that in turn result in an index modulation,  $\Delta n$ , profile that also varies sinusoidally. The wavelength and angle of the interference fringes may be controlled by the angle of the incident interfering beams. The  $\Delta n$  produced by this process is dependent on the time of exposure in the writing process, the material used, and the wavelength of operation of the device. For some materials at some wavelengths a  $\Delta n$  of as large as 0.06 can be achieved. Non-idealities also exist in this type of grating. Upon curing the polymer the polymer will shrink a small percentage. This shrinkage, however, is usually predictable and can be taken into account in the design of the grating period. Multiple holographic gratings can be exposed in a single polymer. Care must be taken not to overexpose the photopolymer in this process. Each individual hologram is thus exposed less than a single hologram would be such that the total exposure does not result in overexposure. Examples of this type of photopolymer are the polymers of Du Pont's HRF series [10, 11]. The index modulation that can be attained in these two polymers is a function of the wavelength. At 633nm HRF-600 has an independently measured  $\Delta n$  of 0.036, although Du Pont reports higher values. At longer operating wavelengths the measured  $\Delta n$  falls off in both materials, for HRF-600 at 1300nm the  $\Delta n$  is 0.015 [12].



Volume gratings have the advantage that they are planar devices unlike the surface relief structures. The primary drawback is that they must be constructed out of special materials that may or may not be convenient materials to use in an application that may have several other processing steps. The materials themselves are often lossy, an inconvenient quality for a waveguiding material. For this reason the grating is sometimes placed in a layer other than the guiding layer. Such a grating has a low coupling coefficient as a result of only interacting with the exponential tails of the guided mode.

### ***2.1.6 Other Diffractive-Optic Structures***

The program also allows the modeling of diffractive-optic structures besides the waveguide grating structures that the program was originally intended to design. This ability is the result of the versatility of the scalar Helmholtz equation. As both free-space and guided modes may be modeled there are a wide class of structures that may be designed. Simple diffractive lenses and prism-like structures are easily designed with the grating depth and period used as variables. An example of such a structure will be shown in the results chapter.

## **2.2 Methods of Solving the Electric Field**

Various methods have been used to model the electromagnetic fields in optical waveguides. Some of these methods have the advantage of speed of computation, while others are able to handle more general structures or are considered more accurate [13].

The beam propagation method (BPM) is computationally efficient and is often used in waveguide structures where the direction of the beam remains largely in one direction. Error increases as the angle of the beam diverges from the center. Wide-angle BPM techniques attempt to minimize this error [14]. However, even at angles as small as  $20^\circ$  from the forward direction of propagation there is considerable error in most versions of the beam propagation method [15]. This reference also describes the inability of the BPM technique to accurately handle structures with “significant occurrences” of local evanescent fields. For gratings that direct a beam a large angle away from the direction of propagation of the incoming mode this method is infeasible. In the case of a waveguide grating coupler this angle may be  $90^\circ$ , and in the region of a grating evanescent modes are to be anticipated.

Rigorous coupled wave analysis is another method for analyzing waveguide modes. It is the most-efficient method for determining the coupling between guided modes in a waveguide, such as between forward and backward propagating modes. However, it is not as useful when dealing with more general diffractive optic structures that may involve guided and free-space modes and gratings that are aperiodic.

The finite difference time domain allows the user to see the advancement of a light pulse with time, which is of importance if interested in the transients of a system. However, it is a computationally-intensive code.

The static finite element method (FEM) and finite difference method (FDM) based on either the scalar or vector Helmholtz equation are similar methods for solving for electromagnetic fields. The finite difference method with a rectangular mesh is a special case of the finite element method with rectangular shape functions. The FDM

yields a very simple program with efficient code, while the FEM can achieve more flexibility in the shape of the structures modeled and placement of grid points.

By the late 1970's diffractive optical structures were being analyzed using the finite element method. The methods were limited in scope, largely as a result of the computer resources available at the time. Moaveni, et al., [16] used the FEM to model diffraction gratings. The goal was to determine the distribution of diffractive orders from illumination of the grating with a plane wave. Since a FEM was employed the grating groove shape could be arbitrary, however the grating was assumed to be infinite and periodic. By making this assumption only one grating period needed to be considered, using the appropriate boundary conditions to model its identical neighbors. The assumption made the problem solvable given the computer resources available at the time.

During this same period Yeh, et al., [17,18] also used a FEM in optical applications. The FEM was used to determine the field components in a single mode unperturbed waveguide such as an optical fiber. The mode profile could be obtained for a waveguide with a non-homogeneous refractive index profile. It is this that differentiated the work from previous methods. Marcatili and Goell had methods that would approximate the propagation characteristics of modes in unperturbed waveguides. However, their methods made assumptions that made them unusable for a more general case. For the FEM elements far from the guiding structure a decaying parameter was used for the boundary conditions. The cross-sections of the waveguides were divided up into elements numbering in the hundreds for their calculations. It is noted that for two decimal place accuracy the entire 4MB memory of their mainframe had to be used.



In parallel to the FEM approaches several researches used the similar finite difference method (FDM). Beaubien and Wexler [19] used the scalar Helmholtz equation with a five-point finite-difference operator to solve for the modes, including the higher order modes of a dielectric waveguide of arbitrary cross-section. The method of solution of the system of equations was iterative using successive overrelaxation, a method later shown to be considerably slower than other methods [20]. This paper was soon followed by an invited paper by Wexler on numerical solutions for electric fields [21] in which he describes much of the mathematics behind the emerging field of computational electromagnetics.

The important aspect of the early FEM and FDM formulations is their ability to handle arbitrary shapes and refractive index profiles. Although the computers at the time severely limited the use of these methods to very simple structures, their versatility makes them important.

To reduce the computer processor time and memory requirements Mabaya et al. introduced a scalar formulation of the FEM.[22] The modes of the guiding structures were found to be quite accurate using the scalar form. The scalar formulation also avoided problems at the time of spurious solutions, particularly when solving for the higher-order modes.

The scalar form of the Helmholtz equation offers one primary advantage in either the FEM or FDM formulations. Only one field component needs to be stored in memory and used in computations. The savings of memory and, more importantly, in processor time required when running the program makes the scalar Helmholtz equation very attractive in many two-dimensional applications where a full vector analysis is not

required. The scalar Helmholtz equation has been used with the FDM on the type of two-dimensional grating structures proposed here [23]. Particularly in a program such as an optimization program where the central analysis routine is run many times, the time savings of the scalar method is necessary for a reasonable run time for the program for problems of a reasonably large size.

### **2.3 Commercial Applications**

Several applications have been commercially developed that will analyze, and in some cases design, optical systems. A brief overview of this currently available software with the strengths and limitations of each is in order.

Of the applications for optical systems the most popular programs use a ray-tracing approach to design systems of lenses in macro-optics. In programs such as Oslo by Sinclair Optics, ray-tracing is primarily used, with physical optics used to calculate modulation transfer functions and point spread functions. No guided wave structures may be analyzed. Another popular program, Zemax Optical Design Program, by Focus Software Inc. is similar to Oslo. Zemax allows gratings with the grating spacing a variable. It allows for optimization using 20 default merit functions. Again the ability to design guided-wave systems is not present. Code V is a third optics program that is primarily designed for lens design applications.

Optics programs that are somewhat closer to solving the type problems described in this work include GLAD and Gsolver. These programs are designed to better handle physical optics. The GLAD program can solve for the near-field diffraction patterns from gratings using a Fresnel diffraction method. Although gratings may be analyzed,

GLAD cannot analyze deep sub-wavelength period gratings. Gsolver by the Grating Solver Development Company comes closest to analyzing the type diffractive elements described in the introduction. It allows for varying one or more parameters of a grating, using rigorous coupled wave analysis to determine efficiencies of grating orders. It allows for multi-step gratings, with either a triangular or blazed orientation. However, Gsolver assumes a uniform grating of infinite extent with constant grating period and depth across its length and will only solve for one grating with plane wave input, thus cannot solve for grating-grating interaction. Guided-waves are also not supported by the Gsolver program.

Thus there is not a commercially available program that can design a grating with a non-uniform period or grating depth. Likewise there is no program that will allow for the design of such a structure using an optimization or other method. Although a program such as Gsolver may be used as a good approximation of a grating with finite extent, its inability to model waveguides and guided modes disallows its use in this application. Most of these commercial programs are limited primarily by the analysis routines they use. None use an a method such as the Helmholtz equation that allows a large variation in structure design. The commercial programs sacrifice flexibility of design for speed of computation.

## **2.4 Design through optimization**

As there are no commercial programs that can design the diffractive optical structures of interest here, the next question is whether any researchers have published works in this area. There have been attempts to design diffractive optic structures using

an optimization routine and a central analysis method [1,2]. The limitation here is the same as in the commercial applications. The method is suitable for a narrow range of structures, but is too narrow in scope to optimize a generic diffractive optic structure. A more general program is needed.

Optimization algorithms may be used to design structures by iteratively trying many structures until the optimal structure for a given situation is found. Common optimization algorithms include local algorithms such as the gradient descent (or ascent) method that require the user to input an initial guess that is in the neighborhood of the solution for each dimension of the search space, as the program will only find the nearest local extrema. In contrast, global search algorithms allow the user to overcome local maxima in the search for the global maximum. To be efficient there must be some form of memory from one iteration to the next. Without a memory of the past best solution and an algorithm to use that information in the following iterations, then the search for the maximum becomes an exhaustive search of all possible combinations. Two routines that encompass these features are the simulated annealing algorithm and the genetic algorithm. As previously stated, these algorithms have been used in applications related to optics. A brief description of these research efforts is described here.

The genetic algorithm was developed in the early 1970's [24]. In the genetic algorithm "parents" are used to create the next iteration (generation) of trial points. In a simple one-dimensional example where the maximum of a function defined along the x-axis is desired, the location of a trial point on the x axis is defined by a binary string of numbers. Several trial points are chosen and the value of the function at those points is found. Two of these points are then chosen, those with higher function values have a



greater probability of being chosen; these two points are now the parents. The first half of the binary string of one parent is concatenated with the second half of the binary string of the second parent; the resulting “child” consists of the genetic code of the two parents. After multiple generations of selecting parents with higher function values, a natural selection takes place resulting in finding the maximum. As in a natural genetic situation, mutations, or random changes in the binary code, are also allowed to facilitate fuller coverage of the domain range.

The simulated annealing (SA) algorithm was developed in the mid 1970's [25]. It is similar to the genetic algorithm in that it mimics a naturally occurring process. In this case the annealing of metals through slow cooling to ensure the strongest molecular configurations. After its development, the SA algorithm was quickly incorporated into optics applications. The most common use over the years for the simulated annealing algorithm in diffractive optics has been the formation of binary optic arrays that will transform a plane wave into a pattern of points of light in the far-field diffractive pattern. The mathematical basis for the SA algorithm will be given in Chapter 3.

We have seen that although the methods to analyze the electromagnetic fields of these structures have been around for many years, the use of these methods to implement a design of a structure through optimization has not been attempted. The commercially available programs are not designed to handle such problems and recent research that has implemented optimization algorithms has not resulted in programs with the degree of flexibility necessary to model the type of devices previously mentioned.

The design of the type of diffractive-optic structures described above requires a computer program comprised of two basic parts. The first is to employ one of the

methods mentioned in this section for determining the electric field of both the guided wave and the radiation modes; the algorithm that I have chosen, the scalar Helmholtz equation discretized using the finite difference method, is described in Section 3.1. The second required algorithm is an optimization routine that forms a loop around the electric field solver; this part of the computer program is discussed in Section 3.2.

## **Chapter 3. Theory**

In this chapter is a synopsis of the applicable theory for the major parts of the design algorithm. The scalar Helmholtz equation is derived for continuous space; then it is altered for use with the finite difference method (FDM) in a discretized space. The alternating direction implicit (ADI) method is explained along with the rationale for its employment in this algorithm in the place of a direct solution of the system of equations. Finally the optimization method developed that uses a small-dimensional search space is described, along with the simulated annealing (SA) algorithm, which is the optimization algorithm used.

### **3.1 Solution of the Electric Field**

To determine the direction and magnitude of the out-coupled light from a non-uniform grating or from structures in which multiple gratings have evanescent fields that interact with one another the electric field must be solved in the region of the grating. Since analytical solutions are not possible for such irregular gratings a numerical solution must be found that approximates the true solution closely enough to be considered acceptable. Some methods for finding this solution were given in the previous chapter. The method to be chosen must meet two requirements to be useful for this application. First the method must be able to handle light in both guided-wave and free-space modes traveling in any direction. Second, the method needs to allow for the efficient



computation of the fields since the computational time involved in such an optimization algorithm can be large.

The following sections describe the method by which the electric field in the discretized region of the structure of interest is to be found. If it is then necessary to determine the electric field at a point outside the solution region, the electric field at the edge of the solution space can be used along with the Fourier transform to calculate the field in the far field region.

### 3.1.1 Derivation of Equations

Starting with Faraday's law and Gauss' law for the electric field:

$$\nabla \times \tilde{\mathbf{E}} = -\frac{\partial \tilde{\mathbf{B}}}{\partial t} \quad (3.1)$$

$$\nabla \cdot \tilde{\mathbf{D}} = \rho \quad (3.2)$$

Where above, and in the derivations that follow,  $\nabla$  is the del operator defined as  $(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$ ,  $\rho$  is the volume charge density, and  $\tilde{\mathbf{E}}$ ,  $\tilde{\mathbf{H}}$ ,  $\tilde{\mathbf{D}}$ , and  $\tilde{\mathbf{B}}$  are the time-dependent vector electric field, magnetic field, electric displacement, and magnetic flux density, respectively, and  $t$  is time.

Then taking the curl of both sides of Eqn. 3.1

$$\nabla \times (\nabla \times \tilde{\mathbf{E}}) = \nabla \times \left( -\mu \frac{\partial \tilde{\mathbf{H}}}{\partial t} \right) \quad (3.3)$$

where  $\mu$  is the permeability, and replacing the left-hand side with the vector identity

$$\nabla \times \nabla \times \tilde{\mathbf{E}} \equiv \nabla \nabla \cdot \tilde{\mathbf{E}} - \nabla^2 \tilde{\mathbf{E}} \quad (3.4)$$

results in

$$\begin{aligned}
\nabla\nabla \cdot \tilde{\mathbf{E}} - \nabla^2 \tilde{\mathbf{E}} &= -\mu \frac{\partial}{\partial t} (\nabla \times \tilde{\mathbf{H}}) \\
&= -\mu \frac{\partial}{\partial t} \left( \sigma \tilde{\mathbf{E}} + \epsilon \frac{\partial \tilde{\mathbf{E}}}{\partial t} \right) \\
&= -\mu \sigma \frac{\partial \tilde{\mathbf{E}}}{\partial t} - \mu \epsilon \frac{\partial^2 \tilde{\mathbf{E}}}{\partial t^2}
\end{aligned} \tag{3.5}$$

where  $\sigma$  is the conductivity and  $\epsilon$  is the permittivity.

Assuming that  $\tilde{\mathbf{E}}$  is time harmonic

$$\tilde{\mathbf{E}} = \mathbf{E} e^{j\omega t} \tag{3.6}$$

where  $\omega$  is equal to  $2\pi f$ , where  $f$  is the frequency of the oscillation of the electric field, and  $\mathbf{E}$  represents the vector electric field, without the time-dependence;  $\mathbf{E}$  may have x, y, and z components. By substituting in to Eqn. 3.5 we obtain the wave equation:

$$\begin{aligned}
(\nabla\nabla \cdot \mathbf{E} - \nabla^2 \mathbf{E}) e^{j\omega t} &= (-j\omega\mu\sigma + \omega^2\mu\epsilon) \mathbf{E} e^{j\omega t} \\
\nabla\nabla \cdot \mathbf{E} - \nabla^2 \mathbf{E} &= (-j\omega\mu\sigma + \omega^2\mu\epsilon) \mathbf{E}
\end{aligned} \tag{3.7}$$

Since  $\mathbf{E} = \mathbf{D}/\epsilon$ , where  $\epsilon$  may change with position, the  $\nabla \cdot \mathbf{E}$  in the first term may be rewritten as

$$\nabla \cdot \mathbf{E} = \nabla \cdot (\epsilon^{-1} \cdot \mathbf{D}) = \mathbf{D} \cdot \nabla (\epsilon^{-1}) + \epsilon^{-1} (\nabla \cdot \mathbf{D}) \tag{3.8}$$

For a source-free medium  $\nabla \cdot \mathbf{D} = 0$ , thus eliminating the second term of Eqn. 3.8. The first term may be put back in terms of  $\mathbf{E}$ , rather than  $\mathbf{D}$ .

$$\mathbf{D} \cdot \nabla (\epsilon^{-1}) = -\epsilon \mathbf{E} (\epsilon^{-2} \nabla \epsilon) = -\mathbf{E} (\epsilon^{-1} \nabla \epsilon) \tag{3.9}$$

Substituting Eqn. 3.9 into Eqn. 3.7, and assuming a non-conducting medium where  $\sigma = 0$ , results in

$$\nabla (\mathbf{E} \epsilon^{-1} \nabla \epsilon) + \nabla^2 \mathbf{E} + \omega^2 \mu \epsilon \mathbf{E} = 0 \tag{3.10}$$

The first term of Eqn. 3.10 contains the term  $\nabla \epsilon$ . If we assume that structures to be modeled are weakly-guiding, that is there are not large steps in the refractive index profile, then we may drop this term from the equation.

Making the further substitutions of:

$$\frac{1}{\sqrt{\mu\epsilon}} = \frac{c}{n} \text{ and } k_o = \frac{\omega}{c} \quad (3.11)$$

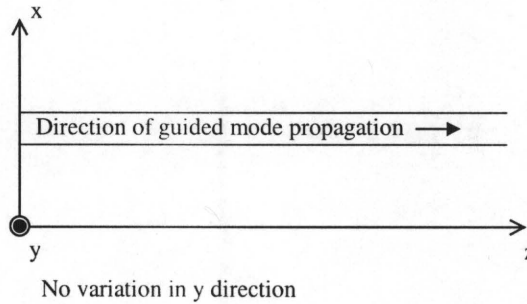
where  $c$  is the vacuum speed of light and  $n$  is the refractive index of the material. The third term of Eqn. 3.10 may be simplified

$$\omega^2 \mu \epsilon \mathbf{E} = \omega^2 \frac{n^2}{c^2} \mathbf{E} = n^2 k_o^2 \mathbf{E} \quad (3.12)$$

for the final wave equation in a homogeneous, source-free region

$$\nabla^2 \mathbf{E} + n^2 k_o^2 \mathbf{E} = 0 \quad (3.13)$$

Throughout this work the coordinates shown in Figure 5 will be used, where the planar devices show no variation in the  $y$  direction.



**FIGURE 5.** Coordinate system used in all designs

For a transverse electric, TE, mode of a waveguide it is known that the components  $E_x = E_z = 0$ . Therefore the  $\mathbf{E}$  of Eqn 3.13 may be replaced with the scalar  $E_y$ .

If the coordinates in the x and z directions are normalized by  $k_0$  then Eqn. 3.13 may be reduced to:

$$\hat{\nabla}^2 E_y + \varepsilon E_y = 0, \quad (3.14)$$

where  $\varepsilon$  is the relative permittivity in the space to be discretized and  $\hat{\nabla}^2$  is defined for the new normalized coordinate system to be

$$\hat{\nabla}^2 \equiv \frac{\partial^2}{\partial \hat{x}^2} + \frac{\partial^2}{\partial \hat{y}^2} + \frac{\partial^2}{\partial \hat{z}^2} \quad (3.15)$$

where  $\hat{x} \equiv k_0 x$ ,  $\hat{y} \equiv k_0 y$  and,  $\hat{z} \equiv k_0 z$ .

Eqn. 3.14 is referred to as the scalar Helmholtz equation; the normalized version shown here is that used in [23]. The scalar Helmholtz equation is unable to model waveguides in anisotropic media or interactions between TE and transverse magnetic, TM, modes. However, for a large class of waveguide structures the scalar Helmholtz equation offers the attractive benefit of computational simplicity over its vector counterparts. Only one component of the electric field must be used, rather than the x, y, and z components of the electric and magnetic fields.

Unlike beam propagation methods that are intended for use where the propagation is largely unidirectional, the scalar Helmholtz equation can accurately describe the effect of diffraction off of gratings.

### 3.2 The Finite Difference Method

The finite difference method is based upon the idea of replacing a differential equation with a corresponding difference equation. This requires replacing the continuous function  $f(x)$  with a discretized approximation  $f_d(i)$ , where  $i$  is the index of the

discretized points. A relatively large number of points are used per free-space wavelength (at least 15 points per  $\lambda_0$  typically). For this reason the discretization of the continuous function does not result in concerns over losing information as a result of the Nyquist criterion as in some other discretization applications.

For a two-dimensional function this discretization is accomplished by laying a mesh over  $f(x, z)$ ;  $f_d(i, j)$  is then defined at every node in the mesh, but not for the continuum of points between the mesh points. The method then involves the solving of the discretized equation at the mesh points rather than the original equation over the continuum of the space. In the simplest form of the method, a rectangular mesh is laid over the area of interest as shown in Figure 6.

The boundary conditions shown in Figure 6 are typical for the type problem encountered. The boundary condition on the left represents the source of optical power into the region, either a plane wave or a mode from a waveguide. If it is a waveguide mode, then the mode is determined by a matrix method for multiplayer guides as described by Kogelnik in [38]. If any waveguide parameters are changed between iterations of the optimization loop then the program recalculates the mode for this new guide structure. If no modes are present for a given waveguide then the program does not proceed. The method in which the three zero boundary conditions are handled will be discussed later. Also note the numbering convention of the grid points, with the columns numbered 0 to  $M-1$ , and the rows numbered 0 to  $N-1$ . The numbering of these points will be referenced in following sections in this chapter.



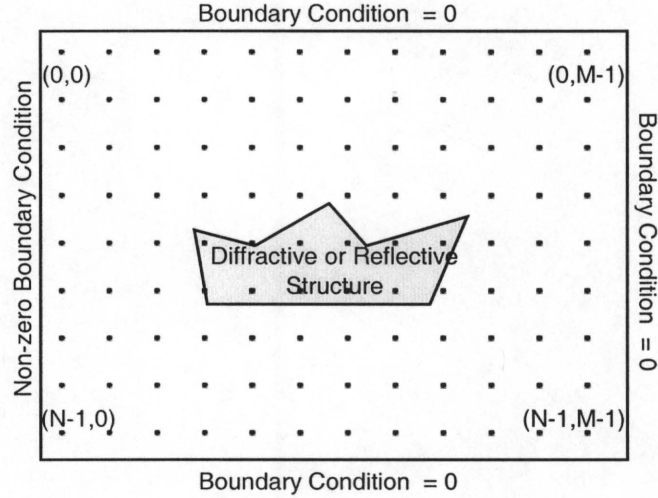


FIGURE 6. Typical configuration

The finite difference version of the scalar Helmholtz equation of Eqn. 3.14 requires an approximation of the second derivative in the discretized space. If we take the Taylor series expansion of  $E(x + \Delta x)$  one obtains:

$$E(x_o + \Delta x) = E(x_o) + \Delta x E'(x_o) + \frac{(\Delta x)^2}{2!} E''(x_o) + \frac{(\Delta x)^3}{3!} E'''(x_o) + \dots \quad (3.16)$$

and likewise for  $E(x - \Delta x)$

$$E(x_o - \Delta x) = E(x_o) - \Delta x E'(x_o) + \frac{(\Delta x)^2}{2!} E''(x_o) - \frac{(\Delta x)^3}{3!} E'''(x_o) + \dots \quad (3.17)$$

By adding Eqns. 3.16 and 3.17 an expression for the second derivative is obtained

$$E''(x_o) = \frac{E(x_o + \Delta x) + E(x_o - \Delta x) - 2E(x_o)}{(\Delta x)^2} + O(\Delta x^2) \quad (3.18)$$

where  $O(\Delta x^2)$  represents the error in the formula introduced when the terms of order higher than  $\Delta x^2$  are discarded. For the  $\nabla^2$  operator of Eqn. 3.14 the second derivative representation of Eqn. 3.18 is used in both the x and z dimensions. Using the subscripted

notation used for the mesh points  $E(x_0 + \Delta x)$  is represented as  $E_{i+1,j}$ , and  $E(z_0 + \Delta z)$  is represented as  $E_{i,j+1}$ . Eqn. 3.14 then becomes in discretized form:

$$\frac{E_{i+1,j} + E_{i-1,j} - 2E_{i,j}}{(\Delta x)^2} + \frac{E_{i,j+1} + E_{i,j-1} - 2E_{i,j}}{(\Delta z)^2} + \varepsilon_{i,j}E_{i,j} + O(\Delta x^2) + O(\Delta z^2) = 0 \quad (3.19)$$

where  $i$  and  $j$  index the mesh points in the  $x$  and  $z$  directions, respectively,  $E$  is the scalar  $y$ -component of the electric field,  $E_y$ ,  $\Delta x$  and  $\Delta z$  represent the distance between the mesh points in each dimension, and  $\varepsilon_{i,j}$  is the relative permittivity at mesh point  $i,j$ .

Oftentimes the  $\Delta x$  and  $\Delta z$  terms are set to be equal as this results in the lowest error. For this case, Eqn. 3.19 can be further simplified to

$$\frac{1}{h^2} [E_{i+1,j} + E_{i-1,j} + E_{i,j+1} + E_{i,j-1} - 4E_{i,j}] + \varepsilon_{i,j}E_{i,j} + O(h^2) = 0 \quad (3.20)$$

where  $h$  represents the mesh spacing in both dimensions. Eqns. 3.19 and 3.20 represent five-point equations, whereas Eqn. 3.18 was a three-point equation. This terminology refers to the number of mesh points contained in the equation. A graphical representation of the five-point operator for a sample point in the discretized mesh is shown in Figure 7.

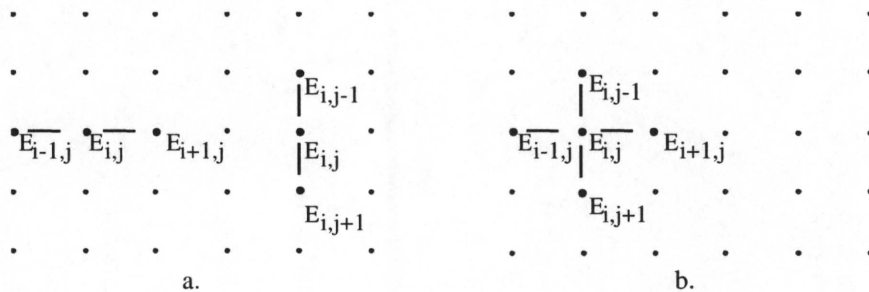


FIGURE 7. Graphical demonstration of a. three-point operators and b. a five-point operator

The uses of finite difference equations presents two problems. The first is the error involved in Eqn. 3.20. The second is the practical necessity of having to impose finite boundaries on the simulation space, where no boundaries exist in the actual device. These non-physical boundaries must be handled in a manner that does not significantly affect the numerical results inside the simulation region.

In Eqn. 3.20 the term  $O(h^2)$  represents the error involved in representing the second derivatives with the three-point operator first derived in Eqn. 3.18. This implies that the error will decrease at a rate of the square of the rate of decrease in the mesh spacing. This is true for all mesh points in uniform regions of the simulation space. However, the simple three-point operator does not take into account a discontinuity in the second derivative that is present at a boundary of two regions of differing refractive index.[26] The result is an error at these individual points on the order of  $k^2\Delta n^2$ . Where  $\Delta n$  is the index difference at the boundary. A more complex operator may be used to account for this discontinuity, however for the current work it was decided against using this operator. The error term is only present at boundary elements, which comprise only a fraction of the total number of elements. To further minimize the possibility of errors, large index steps are avoided using the current formulation of the program. This limitation currently precludes the modeling of semiconductor waveguides with high refractive indices compared to air.

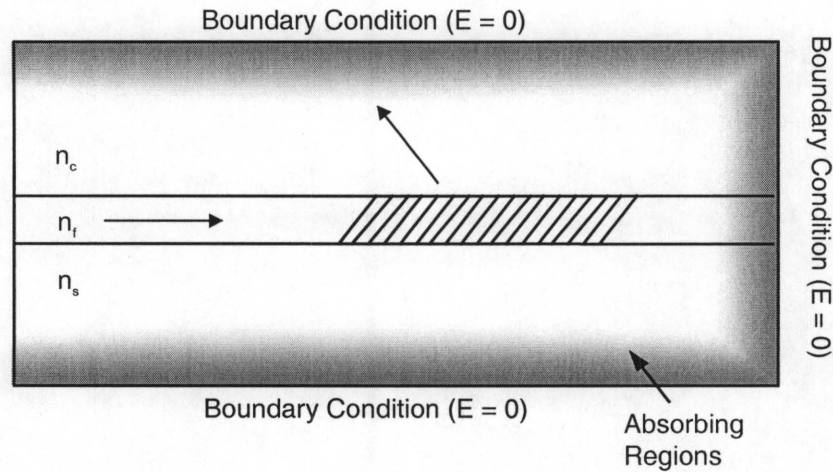
In any finite computation area the boundaries of the computational space are of concern. The physical device size is larger than the area to be modeled numerically, therefore there is an artificial termination of the electric field at some point in the waveguide and above and below the waveguide. Merely setting the electric field to zero

at these boundaries is the equivalent of a perfect metal mirror and results in a reflection that will interfere with the modeling of the electric field in the solution space. A numerical method of dealing with the boundaries so as to render them incapable of interfering with the solution in the modeled area is to introduce absorbers at the boundaries.

For the absorbing boundaries the complex refractive index,  $\hat{n}$ , must be considered.

$$\hat{n} = n(1 + i\kappa) \quad (3.21)$$

where  $n$  is the real refractive index and  $\kappa$  is the extinction coefficient.[27]. The value of  $\kappa$  is ramped up quadratically from 0 at the edge of the modeled area to a maximum after several rows [28]. The real part of the refractive index is unchanged from the surrounding media. Such an arrangement of absorbing regions around the waveguide is illustrated in Figure 8 as the shaded areas. As the electromagnetic field enters these regions it is gradually attenuated as it is absorbed, while the reflections from the boundaries into the simulation region are minimized. With a sufficient number of absorber rows, the electric field may be reduced to a value near zero. The electric field may then be set to zero behind these absorbing regions since after the absorbers the reflections are minimal from the near-zero field. The value of the maximum and the number of rows required is dependent on the waveguide to be modeled.



**FIGURE 8.** Absorbing boundary conditions

If the mesh contains  $N$  rows and  $M$  columns as in the example shown in Figure 6 then there will be  $MN$  nodes. Thus, the problem is transformed into a system of  $MN$  equations and  $MN$  unknowns. The resulting matrix will thus be quite large for a large mesh. However, the matrix can be sparse, and this can save considerable computation time when a solution is sought. The method is not necessarily limited to rectangular grids; triangular and curved grids have also been used. The geometry of the problem dictates the type of grid that is employed. For the planar waveguides to be studied the rectangular grid conveniently discretizes the problem domain.

### 3.2.1 *Direct solution method*

The direct solution method involves the use of the same five-point representation of the discretized scalar Helmholtz equation of Eqn. 3.20, directly solving for the solution at each grid point by forming a large matrix and using Gaussian elimination to obtain the solution. As each grid point in the  $M \times N$  two-dimensional space is an unknown there are  $MN$  equations and  $MN$  unknowns. A matrix could be set up to solve for these unknowns,



however a square  $MN$  matrix would result in an enormous matrix that would quickly exhaust the computer memory presently available for a problem of any reasonable size. In addition to the memory requirement is the question of processor time. Solving a full-matrix would require on the order of  $M^3N^3$  divisions and multiplies, a considerable amount of processor time for large values of  $M$  and  $N$ . All operator counts in this chapter are taken from Strang [29]. Luckily, with the five-point scalar Helmholtz equation we can form a banded matrix with only five non-zero diagonals, which may be solved with many fewer operations.

The numbering of the matrix elements becomes important when solving such a banded matrix. If we assume that in our  $M \times N$  space that  $N$  is smaller than  $M$  then we could number the grid points as shown in Figure 9.

1 ■	N+1 ■	2N+1 ■		(M-1)N+1 ■
2 ■	N+2 ■	2N+2 ■	...	(M-1)N+2 ■
3 ■	N+3 ■	2N+3 ■		(M-1)N+3 ■
4 ■	N+4 ■	2N+4 ■		(M-1)N+4 ■
⋮	⋮	⋮		⋮
N-1 ■	2N-1 ■	3N-1 ■		MN-1 ■
N ■	2N ■	3N ■	...	MN ■

**FIGURE 9.** Matrix numbering for direct-solution method

If we consider the point labeled  $N+3$  in the above figure, then we see that the five unknowns involved in the linear equation for this point will be 3,  $N+2$ ,  $N+3$ ,  $N+4$ , and  $2N+3$ . The  $N+3$  value will lie on the main diagonal since this is the equation for the  $N+3^{\text{th}}$  point. Note that two of the terms, the  $N+2$  and  $N+4$  terms, are immediately adjacent numerically, while the terms  $2N+3$  and 3 are a factor of  $N$  from the central

value. As all of the equations have the same form the resulting matrix formed from these equations will have a band structure as shown in Figure 10.

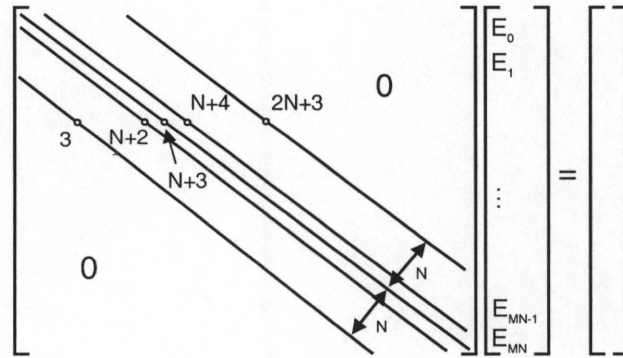


FIGURE 10. Banded-matrix structure of direct-solution method

Three of the non-zero diagonals are the main diagonal and the immediate sub- and super-diagonals. The other two non-zero diagonals are a distance  $N$  from the main diagonal; the bandwidth,  $w$ , is defined as  $N+1$ . [29] If the grid points were numbered consecutively horizontally as opposed to vertically, then the bandwidth would be  $M+1$ . If  $M$  is larger than  $N$ , then the bandwidth would be larger.

A banded matrix may be solved much faster than a full matrix, and the memory requirements are greatly reduced. The number of divisions and multiplies,  $P_1$ , required to find the upper and lower triangular matrices is given by

$$P_1 = \frac{1}{3} w(w-1)(3MN - 2w + 1) \quad (3.22)$$

For forward elimination and back-substitution an additional  $4MN$  steps are required.

Assuming a bandwidth of  $N+1$ , then the total number of multiply and divide operations required is given by

$$P = \frac{1}{3} (N+1)(N)(3MN - 2N - 1) + 4MN \quad (3.23)$$

For large  $M$  and  $N$  this is on the order of  $MN^3$ . From this relation we see the much stronger dependence of one dimension over the other on processing time and the importance of numbering the grid points properly.

Only the matrix elements inside the outermost non-zero diagonals must be stored in memory, therefore in the case shown in Figure 10,  $2N+1$  diagonals of numbers must be stored. Although all but five of these  $2N+1$  diagonals are zero, in solving the banded matrix the linear algebra routine uses these spaces in memory, thus necessitating the allocation of memory for all  $2N+1$  diagonals. Using the LINPACK routines for banded matrices, the full square matrix of order  $MN$  (with its  $M^2N^2$  elements) is replaced by storing only the central diagonals in a form that requires a matrix of size  $3MN^2$  elements. Again we note here in the case of the memory required, as in the case of processor time there is a stronger dependence on one dimension than on the other. It is for these dual reasons of memory and processing-time requirements that when numbering the grid points for the direct-solution method the numbering should be consecutive along the shorter dimension to minimize the bandwidth of the banded matrix.

### ***3.2.2 The Alternating Direction Implicit (ADI) Method***

A modification of the two-dimensional FDM considers the computational problems associated with the  $\nabla^2$  operator. In the FDM earlier described, the Laplace operator is a five-point operator as seen in Eqn. 3.20. When numbering mesh points it would be impossible for all five points to have successive numbers resulting in the two diagonals that for large problems are far from the central diagonal as described in the previous section. The ADI method splits the two-dimensional Laplacian operator into

two one-dimensional operators [20, 30]. Hadley [23] used this method for solving the type waveguide problems that are described here.

The basic ADI equations are as follows:

$$\begin{aligned} (\omega_n + \nabla_z^2)E^{n+\frac{1}{2}} &= (\omega_n + \nabla_x^2)E^n \\ (\omega_n + \nabla_x^2)E^{n+1} &= (\omega_n + \nabla_z^2)E^{n+\frac{1}{2}} \end{aligned} \quad (3.24)$$

The  $\omega_n$  term is a complex acceleration parameter whose value may be allowed to vary at each step in the iteration; it is unrelated to  $\omega$ , the frequency of oscillation of the light,  $E$  is the scalar electric field, in the case of a TE mode,  $E_y$ . Each equation of Eqn. 3.24 is half of an iteration. The right-hand side in each of the two parts is a known value from the previous half-iteration. The  $\nabla_x^2$  and  $\nabla_z^2$  terms are defined as

$$\begin{aligned} \nabla_x^2 E_{i,j} &= \frac{E_{i+1,j} + E_{i-1,j} - 2E_{i,j}}{(\Delta x)^2} + \frac{\epsilon_{i,j}}{2} E_{i,j} \\ \nabla_z^2 E_{i,j} &= \frac{E_{i,j+1} + E_{i,j-1} - 2E_{i,j}}{(\Delta z)^2} + \frac{\epsilon_{i,j}}{2} E_{i,j} \end{aligned} \quad (3.25)$$

where the subscripted  $E$ 's are the scalar electric field values at each grid point,  $\epsilon_{i,j}$  is the relative permittivity at grid point  $i,j$ , and  $\Delta x$  and  $\Delta z$  are the mesh spacings in the  $x$  and  $z$  directions, respectively.

Through limited experimentation I have been unable to better the rate of convergence realized by the acceleration parameters recommended by Hadley in [23]. Two values for  $\omega_0$  are alternated between after each full iteration;  $\omega_1 = -200$  and  $\omega_2 = 0.05b + j1.1b$ , where  $b = Re(\epsilon/2)$  for a representative value of  $\epsilon$  in the structure. I have used the permittivity of the substrate layer for this value.

By dividing the Laplacian into the two separate equations of Eqn. 3.25, one that contains the  $\nabla_x^2$  term and one that contains the  $\nabla_z^2$  terms, there will only be three terms in each equation. Since only three adjacent elements at a time are considered in any single equation it is possible to form a tridiagonal matrix. All unknowns belong to one row or column, and each row may be solved for in succession. Then the other direction is considered, again resulting in a tridiagonal matrix, this time using the columns rather than the rows. Tridiagonal matrix solving routines are computationally fast with the number of multiplies and divides directly proportional to the order of the matrix. For each iteration there will be  $8MN$  operations.

The two directions are alternated between until the solution converges. It has been shown that the ADI method converges considerably faster than the other common iterative method for solving such problems, the successive over-relaxation method [23]. As for the memory required to store the matrix, two matrices are required, one with the electric field values at each grid point prior to the latest iteration, and a second one with the new values. The total memory requirement for matrix-solving is thus  $2MN$ .

### ***3.2.3 Comparison of the ADI and direct-solve methods***

The memory requirement is still much larger than the memory needed for the ADI method and thus the direct solution approach, though faster, can only be used on relatively small problems currently with a desktop computer system. The processor time required to solve the direct solution method is proportional to  $(N+1)^2NM$ , assuming  $N$  is the smaller dimension. In other words, the processor time is highly dependent on  $N$ . For a simulation involving a long, slender waveguide grating structure, where  $N$  is much smaller than  $M$  this method would be very attractive. In contrast, one iteration in the



iterative method takes processor time directly proportional to  $MN$ . However, the iterative method by definition must be repeated until it converges. The number of iterations required is itself dependent on the size of the matrix, specifically the number of grid points between the non-zero boundary condition and the farthest grid point away, for my simulations this generally was the  $M$  dimension. The time per iteration, coupled with the number of iterations required, means that the processor time for the iterative method is directly proportional to  $N$ , but is superlinearly proportional to  $M$ . This idea will be further explored in Chapter 4, where a comparison of the computational time to run the direct-solution method and the ADI method will be given.

The primary limitation to the direct-solution method is the large amount of memory required compared to the ADI method. The largest matrices present in the ADI method are the electric field strength values, the dielectric permittivity for each grid point, and a third matrix used in the solution process. Each of these matrices would have a size  $MN$ , where each value is a complex number, for a total of  $3MN$  complex values that must be stored. All other variables in the program are on the order of only one of the dimensions or are single values. The largest matrix that must be stored in the direct-solution method is the matrix containing the central diagonals of the banded matrix, comprised of  $MN(3N+1)$  complex numbers, a much larger requirement than the ADI method. A further complication is the necessity to use double precision variables in the matrix as otherwise errors can occur during pivoting of the matrix.

Consider a structure that requires a region of  $40\lambda_0 \times 8\lambda_0$  to simulate. At 20 points per free-space wavelength this corresponds to a grid size of  $800 \times 160$  grid points. In the direct-solution program, for each complex number two double precision numbers of 8

bytes each are allocated. Thus the working matrix of the solution algorithm would itself take  $(61568000 \text{ matrix entries}) * (16 \text{ bytes}) = 939 \text{ MB}$ . Even for this relatively small structure a relatively large amount of memory is required by today's computing standards.

### 3.3 Optimization

A method for analyzing a given waveguide structure has been described in the previous section, however it is only able to analyze a single out-coupling structure with the grating or mirror parameters already specified by the user. To be able to design an optimal waveguide for a given purpose, the above algorithm must be nested inside an optimization algorithm. The basic concept of such an optimization algorithm is shown in Figure 11.

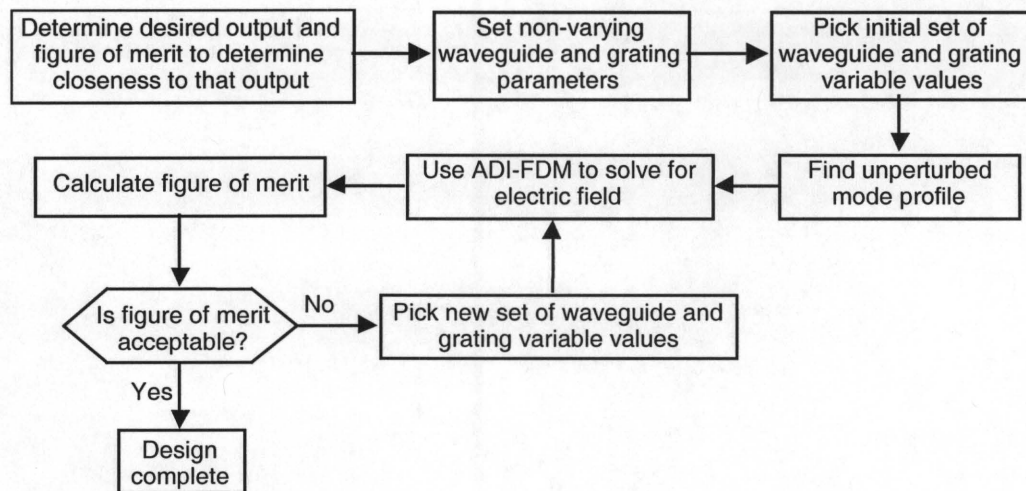
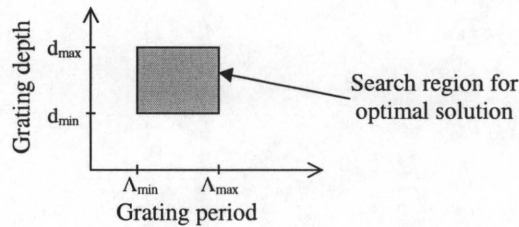


FIGURE 11. Optimization flow chart

Note that the electric field solver is inside the loop formed by the optimization routine. This loop is dependent on the calculation of a figure of merit (FoM) that must be defined by the user prior to running the program. The FoM should indicate how close the output

waveform of the current output coupler configuration is to the desired output. If the FoM is not acceptable then a new structure is created and the electric field again solved for. This loop continues until a structure with an acceptable FoM is found.

The programmer must decide the boundaries for the selection of the set of waveguide and grating variable values. For a simulation in which the grating depth and period of a surface-relief grating on a waveguide are to be varied to find an optimal structure the program user must specify a maximum and minimum value for each of these parameters. In general, the extrema are the result of processing limitations. Once the extrema are determined the optimization algorithm knows over what range to search for an optimal structure; the shaded region in Figure 12.



**FIGURE 12.** Example of search space for optimization algorithm

In Figure 12, we see that the variation of two parameters in the structure leads to a two-dimensional search space for the optimization algorithm. In the general case, allowing  $N$  parameters to vary leads to an  $N$ -dimensional search space for the optimization algorithm to cover. An increase in the number of dimensions greatly increases the amount of time required for the algorithm to find the solution.

### 3.3.1 Low-Dimensional Optimization

By using the optimization loop we have a powerful tool to find a solution, but a solution to what? In an ideal case the waveguide could be broken into small squares each one of a width and height of  $\lambda/20$  as shown in Figure 13, each square corresponding to one mesh point in the discretized simulation space. The optimization algorithm would be allowed to change the refractive index of each of these squares individually until the optimal out-coupled beam is found. This approach would place no restrictions on the finding of a solution; the algorithm would have complete freedom to find a solution, and could result in the finding of an interesting and non-intuitive device.

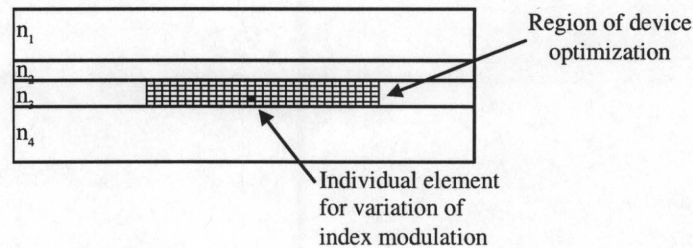


FIGURE 13. Example of device optimization

Unfortunately, there are two major drawbacks to this approach. First, a physical problem: even if a wonderfully efficient device were designed using such an approach it would be impossible to build a device where a volume in a waveguide could have a refractive index that is independent of the refractive index of all of the adjacent points. The second is a computational problem. By allowing each volume element in the waveguide to vary independently in refractive index, the number of dimensions over which the annealing algorithm would have to search for a solution would be in the



thousands, if not millions. Although computer speeds continue to increase, such a manner of problem solution is not currently feasible.

A lower number of dimensions in the search space can be used if the degrees of freedom are limited to the variables in the grating manufacture that may be controlled by the researcher. This approach thus reduces the computation time required and also guarantees a result that is physically realizable by limiting the possible solutions to those consisting of regular fringes in the waveguide, or a pattern of relief structures on the surface of the waveguide or substrate, rather than a random assemblage of blocks. The individual parameters in a waveguide may each be a single dimension over which the solution space covers or may be allowed to vary in a more complex form. If one allows the grating period to vary along the length of the grating in the direction of propagation then we could model this variation using only three parameters by allowing the period to vary quadratically with the three coefficients of the equations comprising three independent dimensions in the search space.

If finding the perfect grating for coupling into and out of a waveguide were a simple function whose optimal solution were easy to determine then a simple search routine such as a gradient descent method could quickly find the solution. However, a problem in which several degrees of freedom are allowed and which may have many local minima is not suited to such a solution method since the method would assume a solution is found when a local minimum is encountered. A method that is able to extricate itself from local minima in order to find the global minimum is required in such a case. Since this is the case in the problems to be studied this type of routine must be used. Two types of routines that fall into this pattern are genetic algorithms and



simulated annealing. Each method is able to find the best solution while avoiding local minima. The simulated annealing routine similarly finds solutions without falling in local minima and has been used previously in other optical waveguide applications [18, 19]; it will be discussed further in the next section.

The method by which the simulated annealing algorithm works is by sorting through possible cases within the boundaries set by the designer. These boundaries are limitations on the values that a variable may take. As an example let us consider a grating where the period of the grating is allowed to change with distance along the waveguide. We want to allow the grating period to change quadratically; this requires the setting of three parameters in the equation

$$ax^2 + bx + c = 0 \tag{3.26}$$

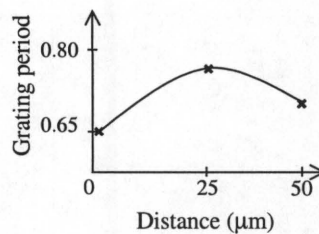
For an optimization with three variables the limitations on each variable creates a three-dimensional cube from which the simulated annealing algorithm must search for the solution. In a formula of the form above there may be a large number of curves that over the range of  $x$  of interest would maintain the grating period within the boundaries set by the user. Consider limiting the grating period in the range  $0.65\mu\text{m}$  and  $0.8\mu\text{m}$  for a length of  $0$  to  $50\mu\text{m}$ . The formula in Eqn. 3.27 will produce values that will fall inside the range restriction, but changing  $b$  slightly as in Eqn. 3.28 results in values that fall outside the acceptable range.

$$1x10^{-5}x^2 + 0.002x + 0.65 \tag{3.27}$$

$$1x10^{-5}x^2 + 0.003x + 0.65 \tag{3.28}$$

Therefore the range of acceptable values for  $a$ , is dependent on the values chosen for  $b$ , and a simple region such as the example in Figure 12 cannot be specified by the user *a priori* for the optimization algorithm to search for an optimal set of parameter values.

This problem may be avoided by specifying not the coefficients of the quadratic equations, but rather the values for the grating period at three points along the waveguide. The three points are each inside the acceptable range of the period. As three points define a parabola, the quadratic curve may be derived from the values for these three points, with  $a$ ,  $b$ , and  $c$  solved for. An example of this is shown in Figure 14.



**FIGURE 14.** Fit of quadratic curve to selected points

The equation of the curve can then be used to determine the period along the entire length of the grating. The specification of values for the points along the curve is also much more intuitive than the specifications of coefficients of a quadratic equation.

### 3.3.2 *Simulated Annealing*

The physical process of annealing is a metallurgical process of allowing a metal to gradually cool so that it may find a state with minimum energy and thus the strongest configuration. In the process of cooling a configuration with a higher energy state may temporarily exist as the atoms search for the lowest-energy configuration. Simulated

annealing algorithms mimic this physical process by allowing the temporary existence of higher energy states in the search for the global minimum; in this way local minima are avoided. A temperature parameter is employed that gradually “cools” with successive iterations, making it less likely as time goes by to accept higher energy states.

In the simulated annealing algorithm, summarized in the flow chart in Figure 15, rather than molecules trying to find the lowest energy configuration a set of parameters are varied to determine the minimum value for a cost function. The user first specifies a range over which each parameter in the problem may be varied. A cost function that determines the “goodness” of the solution must also be specified such that by minimizing the cost function the optimal solution to the problem can be obtained.

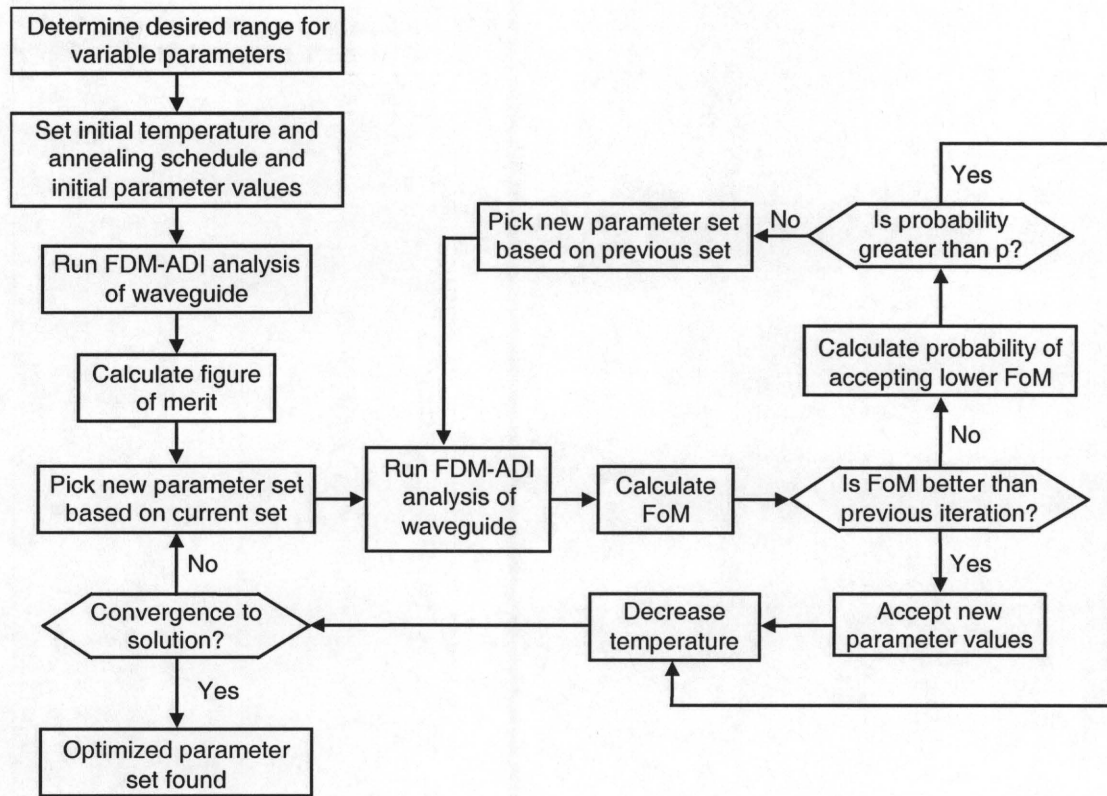


FIGURE 15. Simulated annealing flow chart



The user initially picks a starting point for each variable (dimension) in the specified sample space. The cost function is then calculated. The algorithm then selects a second set of variable values. Assume that the a parameter in the  $i^{\text{th}}$  dimension,  $\phi^i$ , is allowed to vary between  $[\phi_{\min}^i, \phi_{\max}^i]$ , then the next value for  $\phi$  will be chosen by

$$\phi_{k+1}^i = \phi_k^i + y^i (\phi_{\max}^i - \phi_{\min}^i) \quad (3.29)$$

where  $y^i$  is a value in the range  $[-1,1]$  determined by the function:

$$y^i = \text{sgn}\left(u^i - \frac{1}{2}\right) T^i \left[ \left(1 + \frac{1}{T^i}\right)^{|2u^i-1|} - 1 \right]. \quad (3.30)$$

In Eqn. 3.30,  $u^i$  is a random number chosen with a random number generator with a uniform distribution in the range  $[0,1]$ . Note that each dimension has its own temperature schedule and randomly generated value for  $u$ . The cost function is then calculated for the  $k+1^{\text{th}}$  iteration. There are now two possibilities:

In the first possibility, if the resulting  $k+1^{\text{th}}$  cost is found to have a lower value (better solution to the problem) than for the  $k^{\text{th}}$  iteration then the new set of variable values are immediately accepted. The next trial set of variable values are then chosen with the distribution function centered on the  $k+1^{\text{th}}$  set.

The second possibility is that the  $k+1^{\text{th}}$  FoM is found to have a worse value than the  $k^{\text{th}}$  cost. The new set of variable values is not immediately discarded, however. The new set of values may still be accepted with a probability,  $p$ , dependent on the difference in the cost between the current and previous iterations and the temperature at that iteration as shown in Eqn. 3.31, where  $c$  is the cost function to be minimized.

$$p = e^{\frac{-(c_{k+1}-c_k)}{T_k}} \quad (3.31)$$

If the  $k+1^{\text{th}}$  set of values is accepted then the next iteration of the algorithm will start with this set of parameter values as the center of the distribution function for picking the next set of variable values. In the other case, with probability  $1-p$ , the  $k+1^{\text{th}}$  set is not accepted. In this case the algorithm returns to the  $k^{\text{th}}$  set of variable values. The next set of variable values is again chosen from a distribution centered about the set of parameters  $\phi_k^i$ .

The probability,  $p$ , in Eqn. 3.31 is therefore dependent on a temperature parameter,  $T$ , and how much higher the cost function is compared to the current cost function value. The value of  $T$  is gradually lowered with succeeding iterations according to Eqn 3.32,

$$T(k_i) = T_o e^{-c_i k_i^{\frac{1}{D}}} \quad (3.32)$$

where  $D$  indicates the total number of dimensions,  $T_o$  is the initial temperature, and  $c_i$  is a constant. As the temperature  $T$  is lowered the exponent in Eqn. 3.31 is increased, thus making it less likely that a value with a worse FoM is picked as the number of iterations is increased. The routine is therefore allowed to jump wildly over the sample space at the beginning of the algorithm in search of a minimum. As time passes the algorithm jumps less as the temperature decreases and the algorithm eventually settles on a solution.

Variations on the basic simulated annealing algorithm try to improve the efficiency of multivariable optimizations. Somewhat more advanced algorithms use a separate temperature parameter for each variable, which is varied independently depending on the speed at which a solution is converging. This advanced annealing



algorithm is referred to as an adaptive simulated annealing algorithm [31,32]. It has been used in several engineering applications to date [33-5], including electromagnetic applications [36], and phase kinoforms [37].

In summary, an algorithm has been developed that will allow for the optimal design of a diffractive optic structure. This design is accomplished by solving for the electric field present in the region of the structure using the scalar Helmholtz equation discretized using the finite difference method. This analysis routine is then placed inside a simulated annealing optimization loop that is able to change the parameters of the structure that the inner analysis routine analyzes until the optimal structure is found.

## Chapter 4. Testing and Verification of Algorithm

Prior to the use of the code in any design, the code must be verified to ensure that the correct results are obtained. The two parts of the code, the electromagnetic field solver and the optimization routine may be viewed separately. Of less concern is the adaptive simulated annealing code since this has been extensively used and tested by others [320-4]. It is the purpose of this chapter to investigate potential problems in the algorithms and determine if acceptable results can be obtained from the code. Comparisons are made to both theoretical expectations and published experimental results.

### 4.1 Electromagnetic field solver algorithm testing

In solving for the electric field there is the concern that the mesh fineness is sufficiently compact to solve the problem to desired accuracy, but as coarse as possible to minimize computational requirements. Also in question is the number of iterations required in the ADI method for convergence to a solution, and whether the solution that it has converged to is the correct one. The algorithm is then compared to a published result to corroborate the method.

#### 4.1.1 *Comparison of iterative and direct methods*

The iterative method of solving the scalar Helmholtz equation by the finite difference method (FDM) for a two-dimensional problem takes many iterations to converge to a solution. Three issues arise from using the iterative method. The first is

that the amount of computational time required for its execution when compared to a direct solution of the equations. The second is ensuring that the iterative solution is the correct solution. Finally, the number of iterations required to reach this solution must be determined.

Comparison of the direct method solution and the iterative solution gives us an idea of the number of iterations required for the iterative solution to converge and whether the iterative solution is in fact the same solution as that obtained by the direct solution. It also offers an opportunity to observe the computational time required for each solution method.

A sample waveguide was simulated using two computer programs to determine the difference between the amount of processor time required and any differences in the solutions arrived at by an implicit solution and a direct solution. The implicit solution is found by the alternating direction implicit (ADI) method, while the direct solution is found through solving the banded matrix that is generated by the FDM.

The waveguide used for the simulations is shown in Figure 16. A free-space wavelength  $1.55\mu\text{m}$  is used. In both the x direction and in the z direction 20 grid points are used per free-space wavelength. At the left-hand boundary the fundamental mode of the waveguide is introduced. The other three boundaries use an absorbing boundary condition. The absorber consists of 15 rows with the imaginary component increasing parabolically to a maximum of 1.4.

n = 1.0	
n = 1.386	t = 0.5 $\mu$ m
n = 1.485	t = 1.0 $\mu$ m
n = 1.386	

**FIGURE 16.** Waveguide used to compare direct solution and implicit solution techniques

The above waveguide was simulated twice, each time the width of the simulated region was 80 grid points. The first simulation was for a waveguide with a length of 200 grid points. For the second simulation the waveguide length was doubled to 400 grid points.

For the 80 x 200 region the times for the direct solution of the matrix and the implicit solution were recorded. For the implicit solution different numbers of iterations were tried. All times were measured on a IBM RS6000 computer with the code compiled in C with the exception of the LINPACK banded-matrix solver used in the direct solution, which is written in FORTRAN, and is called from the C program. For these times no optimization codes were used when compiling the program. The times recorded for the algorithm to run are presented in the first column of Table 1.

Optimization codes are optional flags that may be used when compiling a program. They allow the compiler to unroll loops and make other changes to the program for more efficient computation. Different flags allow the program to make more or less aggressive changes to improve speed. More aggressive optimization routines may lead to slightly different numerical results as a result of changes in rounding methods. The `-O2` flag is the most commonly used optimization flag and in my programs leads to a factor of two improvement in required processor time, with no changes in the numerical



results. The more aggressive  $-O3$  flag can theoretically lead to slight numerical differences in program execution, although in my programs I never found a difference in numerical results in the  $-O3$  flag to the same program compiled with the  $-O2$  flag. However, the  $-O3$  flag led to only marginal speed improvements over the  $-O2$  flag, so I chose the  $-O2$  flag for most compilations to avoid the possibility of numerical rounding errors.

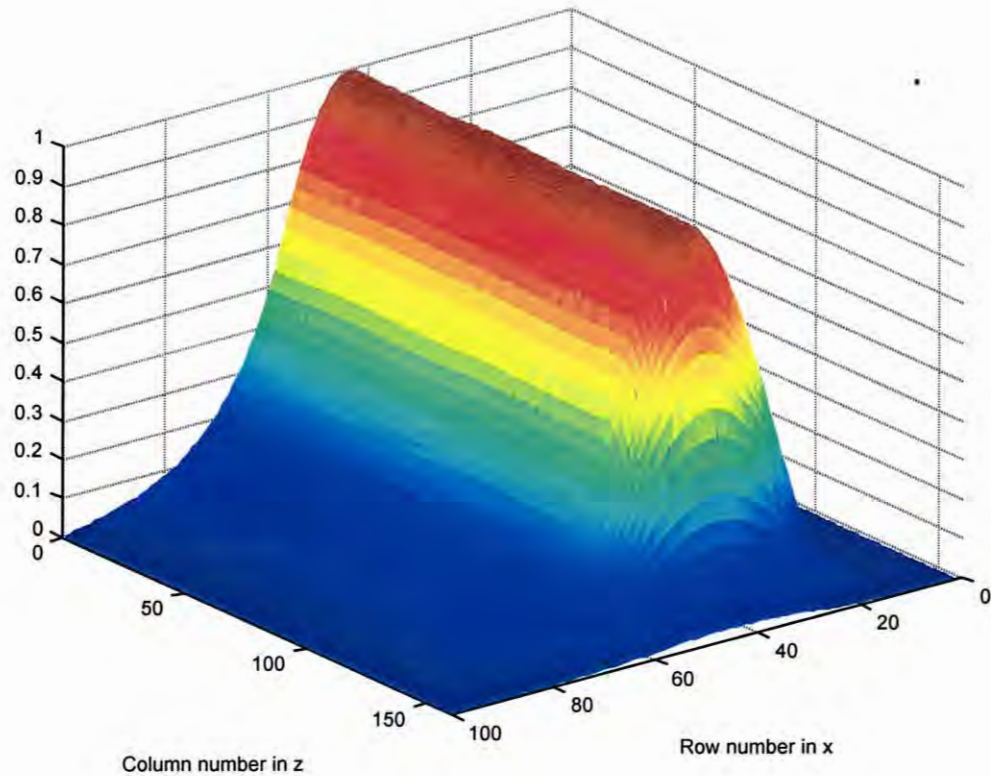
Simulation space size	80x200	80x400
Direct solve	58 sec.	117 sec.
100 adi iterations	32 sec.	
150 adi iterations	47 sec.	93 sec.
200 adi iterations	62 sec.	124 sec.
250 adi iterations		155 sec.

**TABLE 1.** Comparison of direct solution and iterative solution processing times

Shown in Figure 17 is the result of a typical simulation; in this case a  $160 \times 100$  region. The plot is shown for illustrative purposes, as all of the simulations in this section are similar. The boundary condition on the back-left side of the plot, where  $z = 0$ , is the  $TE_0$  mode of the waveguide. The other boundary conditions are set to zero behind the absorbing layers. The mode can be seen to travel down the waveguide in the  $z$  direction until it is attenuated by the absorbing region at the right-hand side of the figure. Note that the absorbing region attenuates the field quickly, with the reflected energy so low as to be imperceptible in the plot. The row of data that corresponds to the maximum amplitude of



the mode is used for some of the following comparisons. In the case of the sample waveguide below, this would correspond to the row of data where  $x = 45$ .



**FIGURE 17.** Electric field amplitude of unperturbed waveguide

Plotted in Figure 18 is the amplitude of the field in the row of grid points corresponding to the maximum amplitude of the fundamental mode of the waveguide. Some reflection can be seen from the absorbing boundaries. Figure 19 shows the region of interest more clearly. It is evident that the case of 100 ADI iterations is insufficient. The case of 150 iterations is considerably closer to the direct solve solution, while the 200 iteration case is so close to the direct solution case that the two curves are

indifferentiable. The difference in the computer processing time between the direct solution and the 200 iteration implicit solution is marginal.

The same waveguide with the length in the z direction doubled to 400 grid points was then simulated. The width was left at 80 rows. The time required for the two methods was again measured; the results are shown in the second column of Table 1. Note that the time required for the direct-solve method and for the ADI method where the same number of iterations are performed scales by a factor of two. However, whereas in the 80x200 case 200 iterations of the implicit solution closely matched the direct solution, in this longer waveguide more iterations are required for the field at the far end of the waveguide (farthest from the side with the non-zero boundary condition) to converge. The time required for the iterative method therefore scales super-linearly with the number of grid points. The implicit method now has a significant time penalty. The field amplitudes are plotted in a similar manner to the previous example in Figure 20 and Figure 21. Here the 250 iteration case is very close to the direct solution.

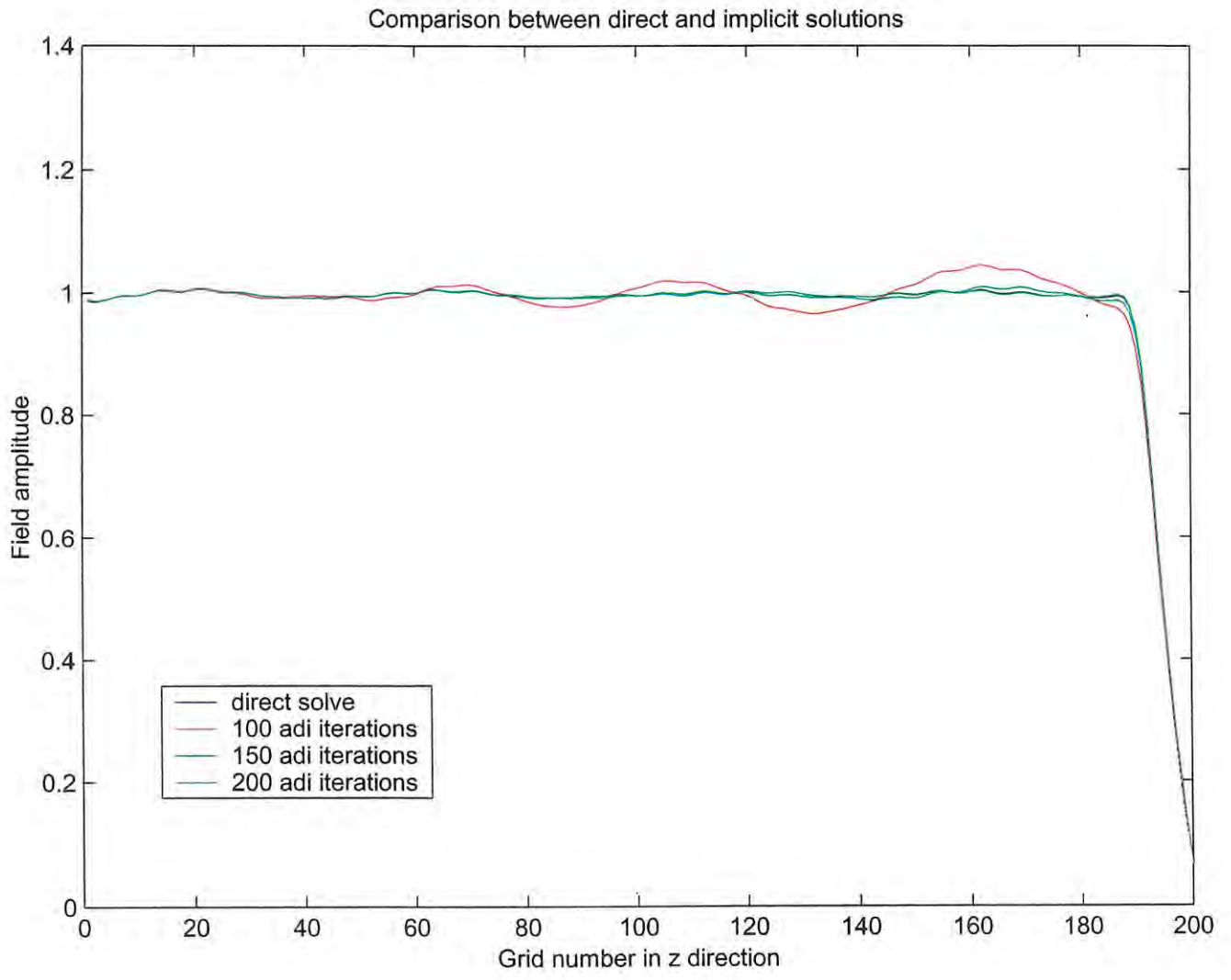


FIGURE 18. 200 column example

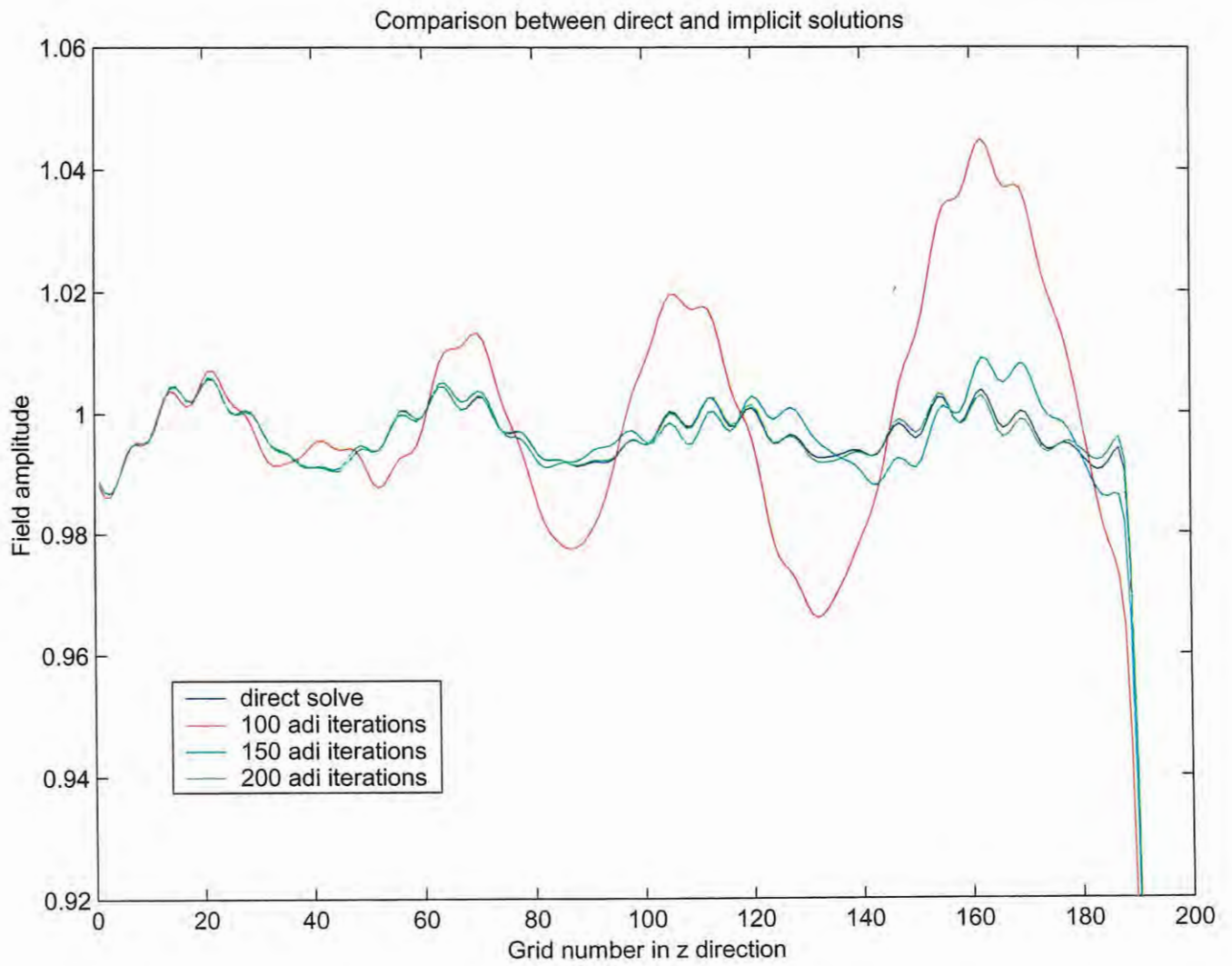


FIGURE 19. Region of interest in 200 column example



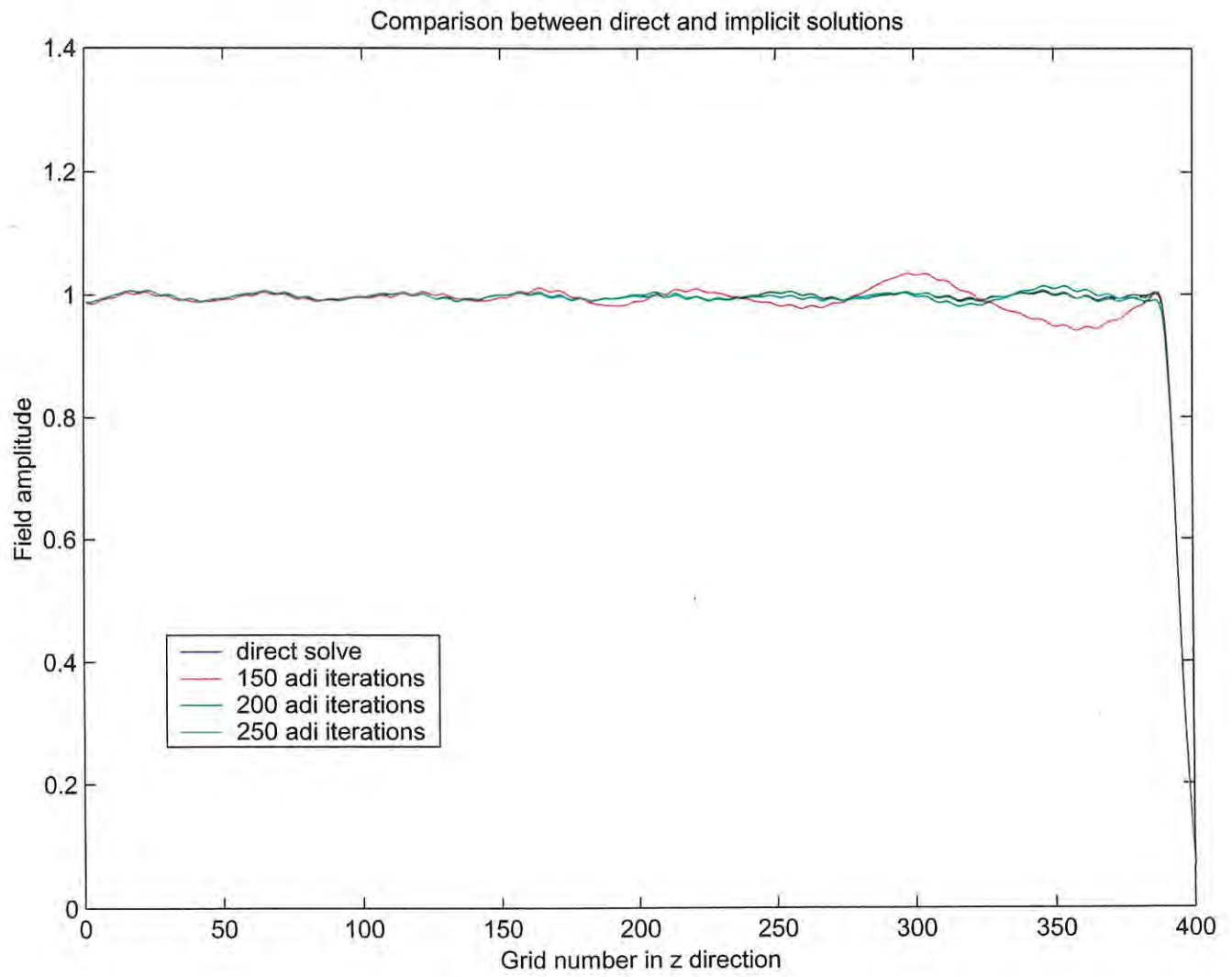


FIGURE 20. 400 column example

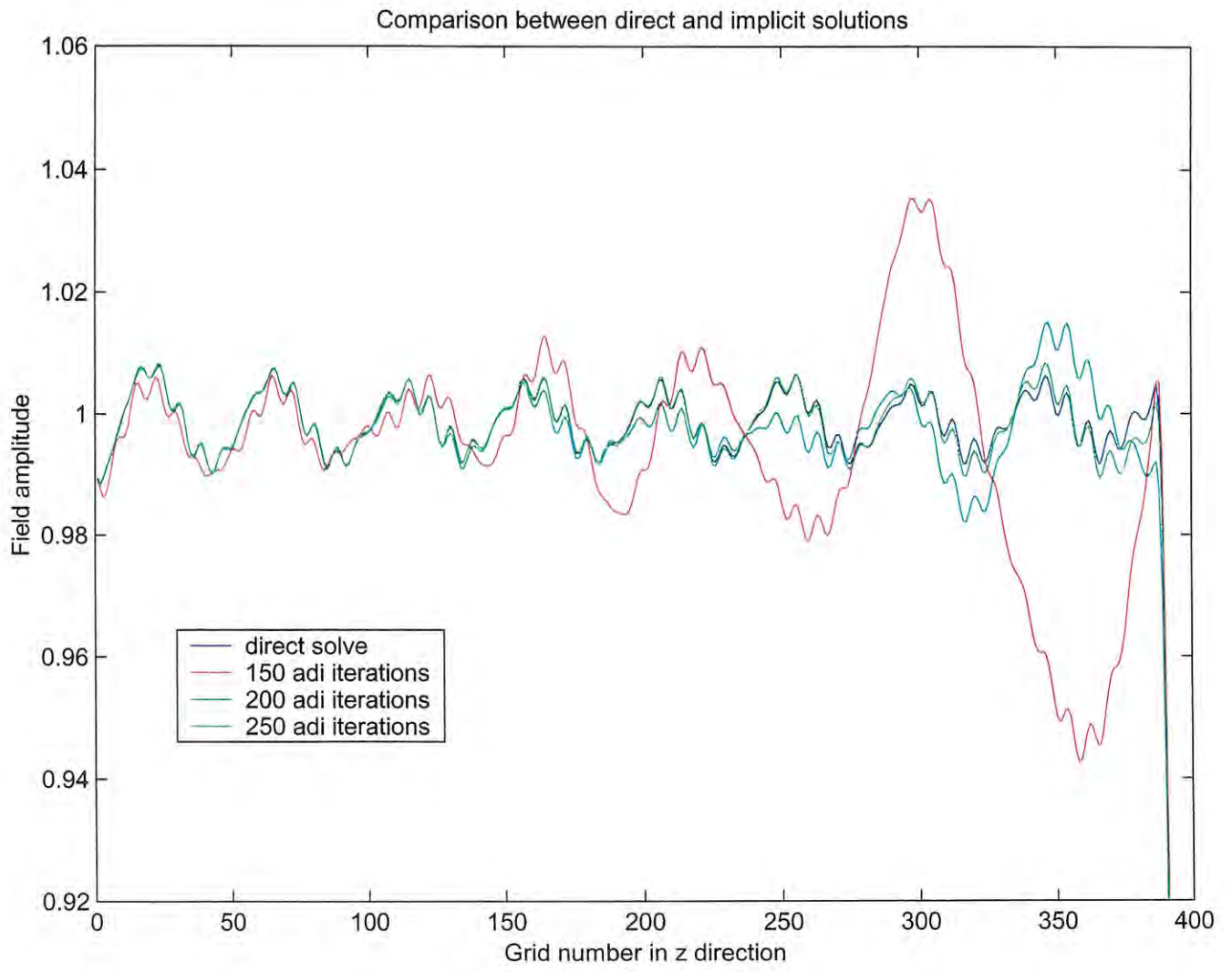


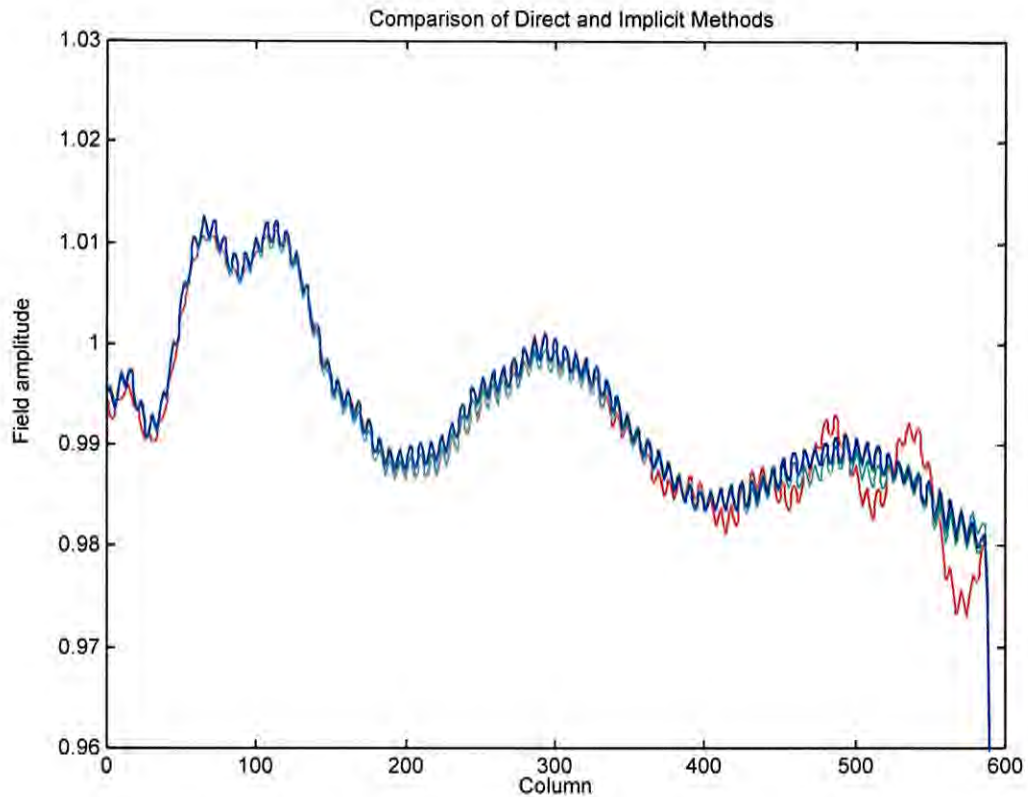
FIGURE 21. Region of interest in 400 column example

All of the code was then optimized at compilation with the exception of the graphics routine, which is a small portion of the code and is identical to both methods. The waveguide of Figure 16 is again used, with the solution space dimensions increased to 90X600 grid points. The results are shown in Table 2.

Direct solve	67 sec.
250 adi iterations	133 sec.
300 adi iterations	158 sec.
400 adi iterations	214 sec.

**TABLE 2.** Comparison of results for 90x600 waveguide test with optimized compilation

These results are plotted in Figure 22. The 300 iteration case is very close to the direct-solve solution. The 400 iteration case lies almost exactly on top of the direct solve solution with the exception of the far right of the plot. It should be noted that the implicit case is solved using single precision variables, while the direct solve case is solved using double precision. The double precision is necessitated by the mathematics involved in pivoting during the direct solve matrix inversion. For the direct solution method the use of double-precision accuracy is necessitated by the sensitivity of the method to round-off error in the case of an ill-conditioned matrix.



**FIGURE 22.** Comparison of one row of data using direct and implicit methods

The percent error between the three iterative cases and the direct-solve case are plotted in Figure 23. Note that in all three cases the relative error is under one percent for all grid points. It is evident that the highest error is at the highest-numbered grid points, farthest from the non-zero boundary condition. It is at these grid points that the greatest difference is seen between the three iterative cases. For the 400 iteration case the percent error remains under 0.1% throughout the length of the waveguide. The number of iterations required is partially at the discretion of the program user, as even the lowest number of iterations in this example resulted in less than 1% deviation from the direct-solution case. It should be remembered that error is also introduced from the coarseness



of the discretization, and may be much larger than the error introduced from the number of iterations used.

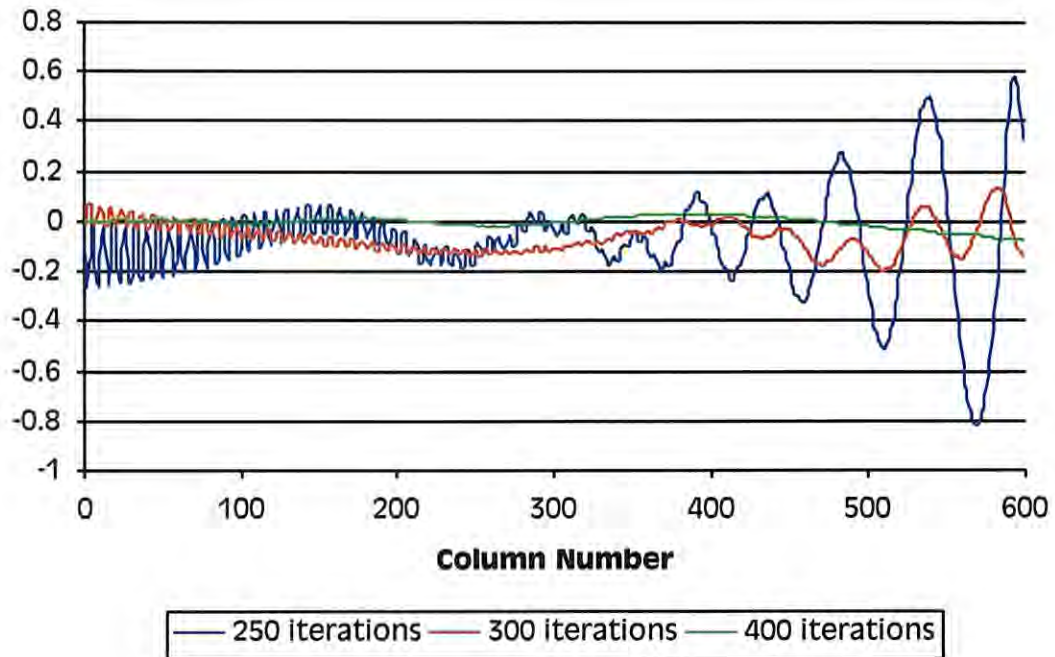


FIGURE 23. Percent error of three iterative cases compared to direct-solve case for 600x90 waveguide

The iterative cases for the sample waveguides illustrate that the number of iterations needed is a function of the number of grid points in the length of the waveguide. The processor time for one iteration is proportional to the number of grid points in the simulation. Therefore, the direct solution method would appear to have the advantage for such problems, as the time required for computation of the implicit method increases more than linearly with the length of the simulation space. However, the direct solution method requires considerably more memory than the implicit method, thus

making its employment impossible for large structures with the computer resources presently available to me.

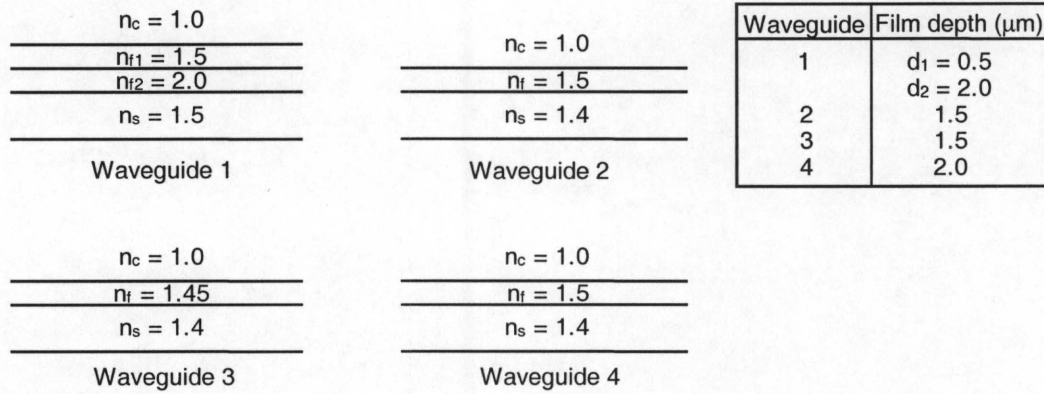
The tests do prove that when the iterative method converges it does reach the same solution as the direct solution of the banded matrix. The iterative method may then be used on larger structures with the confidence that the correct solution is reached although in those cases a comparison to the direct solution cannot be made.

#### **4.1.2 Mesh fineness**

The larger the number of points used, the larger the matrices become for a structure of a given size. Since the region to be discretized is two-dimensional, the matrix size increases with the square of the number of points per wavelength. As we have seen, the matrix size determines the amount of processor time required, thus it is important to determine the minimum number of grid points required for a reasonable solution to be obtained.

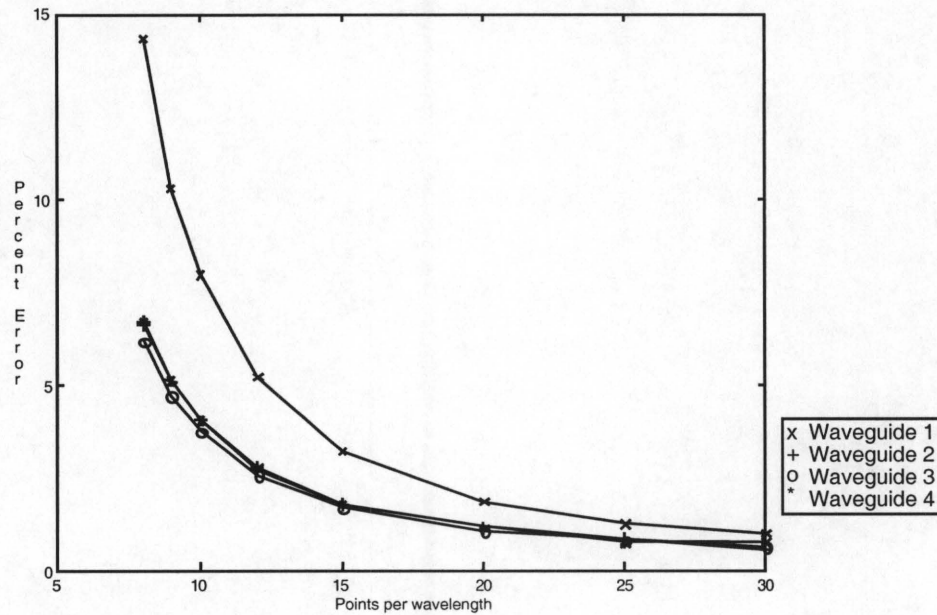
To determine the minimum number of points per wavelength that would result in a solution with an acceptable amount of error, I tested the set of waveguides shown in Figure 24. Each waveguide was tested with the number of points per free-space wavelength varied from 8 to 30. The waveguides were unperturbed, that is no grating or mirror structures were introduced. Plot increment in microns in the direction transverse to the direction of propagation was left constant for all simulations at 0.1 microns, corresponding to 15.5 points per wavelength. Variations in the plot increment in the transverse direction did not have a pronounced effect on these calculations since the guided mode is traveling unperturbed through the waveguide. For all simulations there

were 30 free-space wavelengths in the direction of propagation plus 30 rows of absorbers at the end.



**FIGURE 24.** Four sample waveguides used to measure error

The effective index of the lowest order waveguide mode was then calculated from the resulting numerical solution by looking at the phase of the data and calculating the wavelength of the mode. This value was compared to the value expected by using the multilayer method for slab waveguides presented by Kogelnik in Tamir [38]. The result is shown in Figure 25.



**FIGURE 25.** Percent error for four waveguides as a function of number of mesh points per free-space wavelength

Further testing was done to determine if these numbers would change significantly if double-precision rather than single-precision variables were used. No difference was detected when the above tests were performed again using double-precision numbers. Similarly, the width of the modeled region was increased by over 50% with no change in the above results. The error is therefore from the approximations involved in the scalar Helmholtz equation itself and not from round-off error in the computations or windowing effects.

From the data plotted in Figure 25 we see that the greatest error is found in the waveguide with the most confined mode. This is expected since the scalar Helmholtz equation assumes the weakly-guiding approximation. There are methods that improve the three-point difference operator to accommodate the error encountered at steps in the refractive index in a slab waveguide [39, 40]. However, for this work I decided a



sufficient number of grid points could be used to successfully model a large class of structures. The use of more complex operators to lessen the number of mesh points required would partially offset the gains realized by using fewer mesh points.

#### 4.1.3 Comparison of results to published data

A comparison of the out-coupled angle of a slab waveguide was performed to ensure that the simulated results agree with both theoretical and experimental results of previously published data. The comparison was made to an early paper by Ostrowsky and Jacques where the outcoupling angle of a surface-relief grating on a planar waveguide was measured and compared to the theoretical results [41]. The grating used is illustrated in Figure 26.

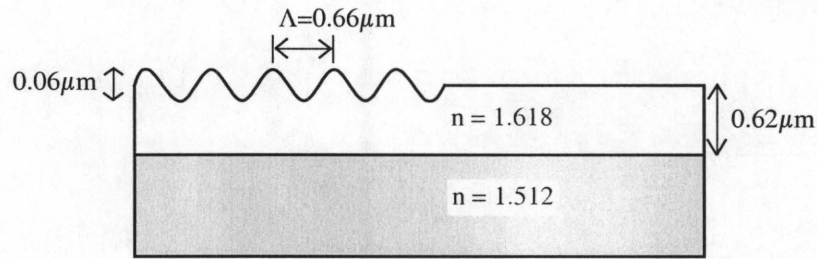


FIGURE 26. Grating described by Ostrowsky and Jacques and used for algorithm verification

The theoretical outcoupling angle is given by the equation 4.1.

$$\theta = \sin^{-1} \left( N_{eff} - \frac{\lambda_o}{\Lambda} \right) \quad (4.1)$$

where  $\theta$  is the angle of the out-coupled beam measured from the normal to the surface of the waveguide,  $N_{eff}$  is the effective refractive index seen by the guided mode, and  $\Lambda$  is the grating pitch. The authors state that a theoretical angle of  $38.67^\circ$  is expected. The

experimental result that they obtained was  $37.95^\circ$  for the  $TE_0$  mode at a free-space wavelength of 632.8nm.

The same structure was then simulated using the electric field solver of my program. Different levels of mesh fineness were used. In each case the result was within  $1^\circ$  of the theoretical result. For the finest mesh 46 points per free-space wavelength ( $\lambda_0$ ) were used in both the x and z directions. A second trial used 46 points/ $\lambda_0$  in z and 20 points/ $\lambda_0$  in x. A third trial used 23 points/ $\lambda_0$  in z and 20 points/ $\lambda_0$  in x. The results are shown in Table 3.

	Angle (deg.)
Theoretical	38.67
Experimental	37.95
46pts./ $\lambda_0$ in x, 46pts./ $\lambda_0$ in z	38.64
20pts./ $\lambda_0$ in x, 46pts./ $\lambda_0$ in z	38.79
20pts./ $\lambda_0$ in x, 23pts./ $\lambda_0$ in z	39.30

**TABLE 3.** Comparison of theoretical, experimental, and simulated data from surface-relief grating

This test against published data offered an opportunity to test a two-dimensional structure in which the beam of final interest is not a guided wave, but rather a free-space wave traveling at an angle very different from the direction of the original guided mode of the waveguide. This type of out-coupling structure is similar to those to be simulated so offered the best test of the electric field algorithm. The close agreement between the simulated, theoretical, and experimental validates the results of this algorithm.

Additional testing of various aspects of the post-processing of data was also performed. In many simple waveguides the output angle of the out-coupled light could best be found by taking the fast Fourier transform (fft) of a row of data. The results of the fft algorithm were verified, and it was later used inside the annealing loop as part of a method for determining the figure of merit for a structure. It was noted that the farther the row of data is from the waveguide surface the less the evanescent fields interfere with the determining of the outcoupled angle of the far-field radiation via the fft. Also, since a large percentage of the data points in the fft correspond to non-physical frequencies, a large number of fft points works best to allow for determining the out-coupled angle with sufficient accuracy. For example, consider a simulation using 20 points to represent one free-space wavelength: the fft point corresponding to a sinusoid with a period of five grid points cannot represent a radiating mode. Usually 8192 points were used in any simulation requiring a fft.

## **4.2 Simulated Annealing (SA) Algorithm Testing**

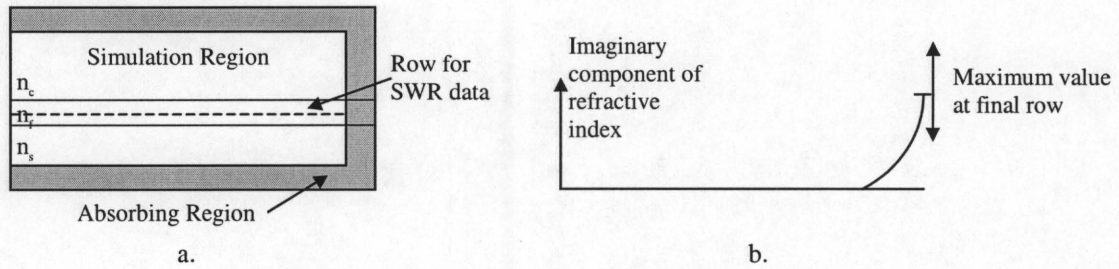
A simple test was needed to determine if the simulated annealing algorithm could be used with the electric field solver and if it in fact would converge to the correct solution. A simple test would be to allow the SA algorithm find the optimal strength for the absorbing layers at the end of the simulation region. I had determined these values largely through trial-and-error on first writing the electric field analysis routine; so the approximate solution was already known.

A typical waveguide structure is shown in Figure 27a; the computation region is shown. The region over which computations are carried out covers the simulation and



absorbing regions. The simulation region is where the solution of the electric fields is desired; the absorbing region is a non-physical region used to insulate the simulation region from the effects of the boundaries as described in the previous chapter. The strength of the absorbing regions should increase parabolically as one moves away from the simulation region with the vertex of the parabola at the first row [28]. The number of variables left to the user is low. The first is the number of rows to be used. The second is the maximum absorbance in the absorbing region, the row farthest from the simulation region, as illustrated in Figure 27b. The number of rows does not lend itself well to an optimization algorithm: the more rows of absorbers the smaller the reflectance will be back into the simulation region. However, a greater number of rows means increased computation time. The user must decide when a sufficiently small reflectance from the boundaries has been achieved. In general, around one to two free-space wavelengths are required for the absorbing regions. For the case of 20 grid points per  $\lambda_0$ , 30 absorber rows may be required. However, the strength of the absorbance is a variable that has a clear minimum, as will be demonstrated. It is therefore simple to minimize with the SA algorithm. Although for such a simple search a more trivial algorithm could be used, the problem is well-suited for testing the SA algorithm.



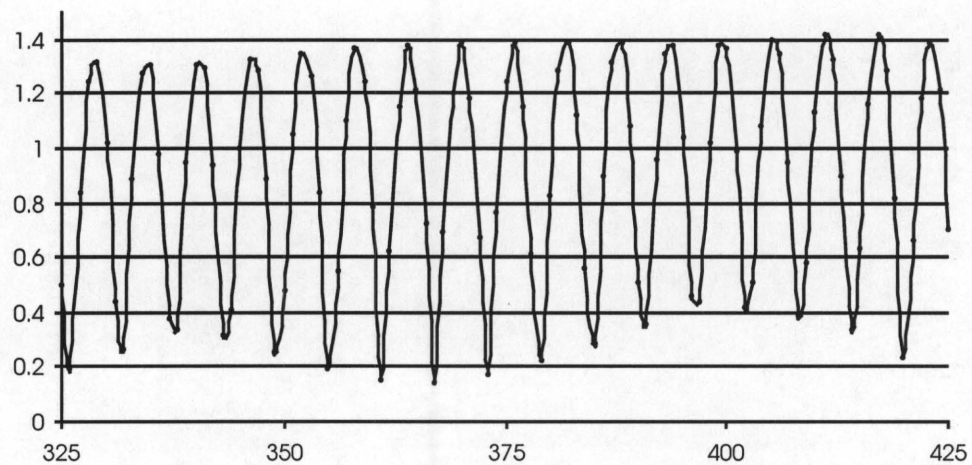


**FIGURE 27.** Computation region of waveguide showing simulation and absorbing regions (a.), and the imaginary component of the refractive index along the row for the SWR computations

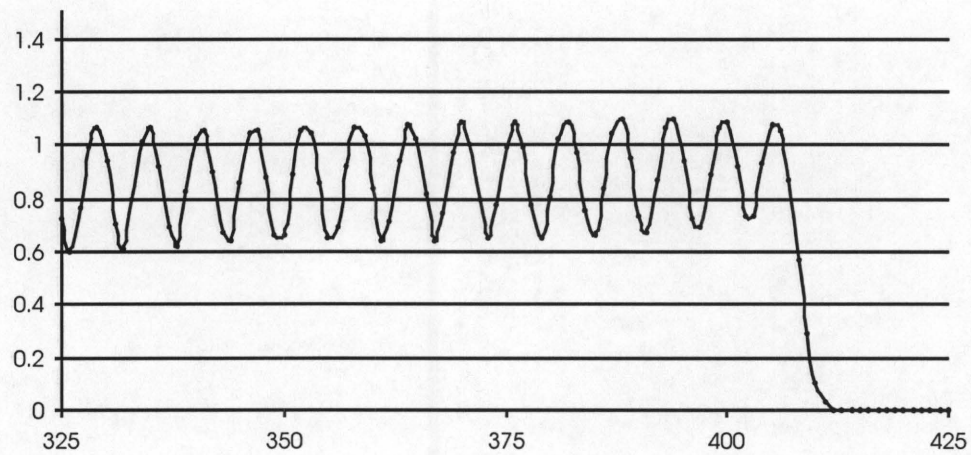
The absorbing regions around the simulation space are modeled by increasing the imaginary portion of the refractive index of the material used in the waveguide, while leaving the real portion unchanged. The strength of the absorbing region may be modeled with one variable, the value of the imaginary part of the refractive index at the last row of the absorbing region. The test for the simulated annealing algorithm is thus a simple one-dimensional search.

Ideally one would expect to see all of the energy traveling in one direction in the waveguide with no reflected energy for the end of the waveguide. Any reflection from the end of the computational region will result in a standing wave as it counter-propagates through the waveguide. If the imaginary portion of the refractive index of the absorbing layers is too small then the electric field is not fully attenuated before it strikes the final edge of the computational region, where the electric field is set to zero. The edge of the computational region thus acts as a metal mirror, reflecting the remaining energy back into the simulation region. However, if the absorbing region is made too strong then the energy will reflect off of the absorbing region itself, again resulting in a standing wave seen in the waveguide. The three cases of absorbers that are too weak, too

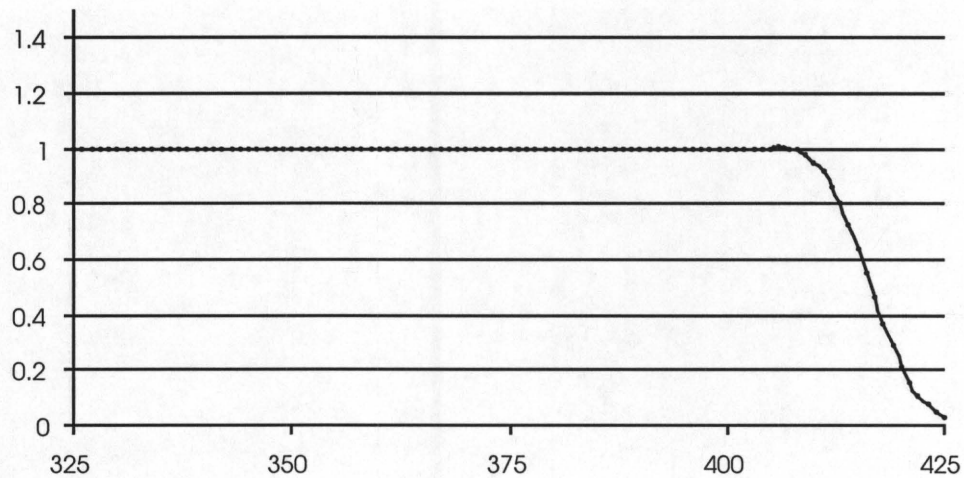
strong, and optimal can be seen in Figure 28, Figure 29, and Figure 30, respectively. All three graphs use the same range for the amplitude of the electric field, and all show the last 100 columns of an unperturbed waveguide. The data is taken from the row of data corresponding to the maximum value of the electric field in the waveguide. In the first, an insufficient absorbing region leads to a large reflection. In the second graph an overly strong absorbing region leads to excessive damping and a similar large reflection. The third graph shows an absorbing region that results in almost no backward reflection.



**FIGURE 28.** Amplitude profile in direction of propagation resulting from insufficient absorbing region



**FIGURE 29.** Amplitude profile in direction of propagation resulting from excessive absorbing region



c.

**FIGURE 30.** Amplitude profile in direction of propagation resulting from optimal absorbing region

The figure of merit for the annealing algorithm was to minimize the standing wave ratio (SWR) in the waveguide. The row of data points in the film region with the highest amplitude was chosen to measure the SWR. Over the length of the guide the



highest and lowest amplitudes were divided. The value of the imaginary portion of the refractive index was allowed to vary over two orders of magnitude, from 0.1 to 10. Three different simulations were run, using 10, 20, and 30 rows of absorbers with one free-space wavelength discretized into 20 points in each direction.

The results are predictable. The lowest SWR is found in the case with the highest number of absorbers as seen in Figure 31. The SWR in the case of 10 rows is never found to be that low. However, 20 rows results in a quite low SWR. This example was run with 20 points/ $\lambda_0$ . In general, one free-space wavelength is the minimum width of the absorber regions found to achieve good results.

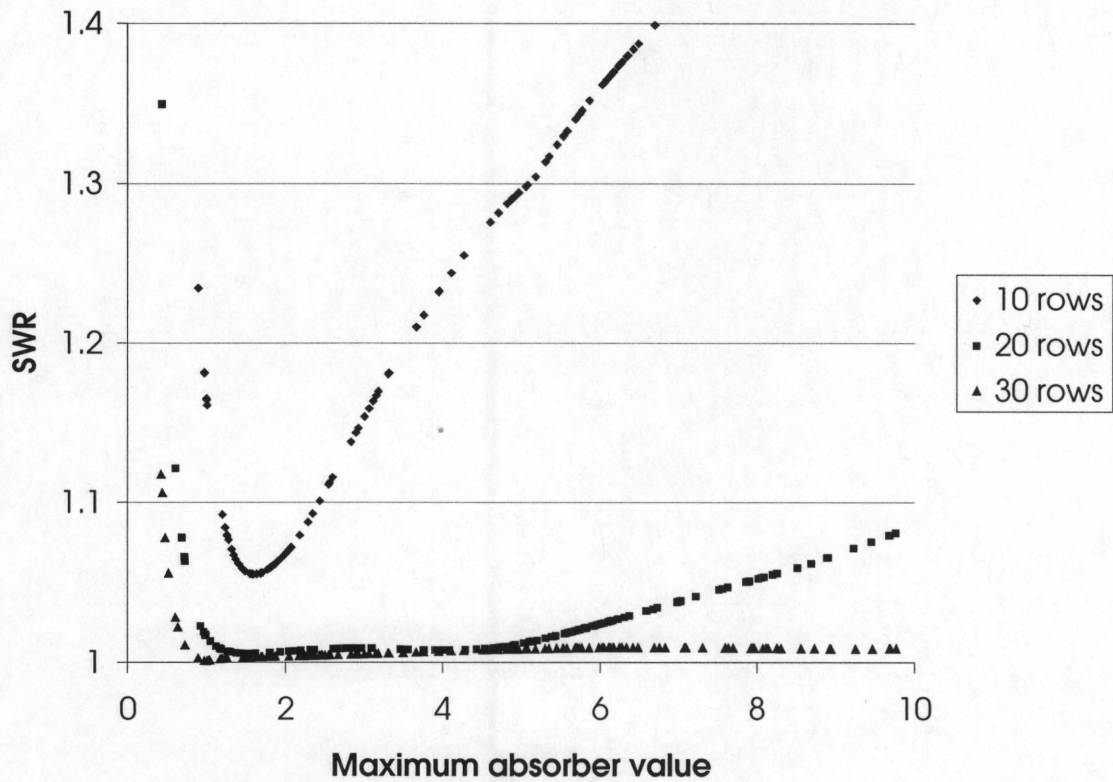
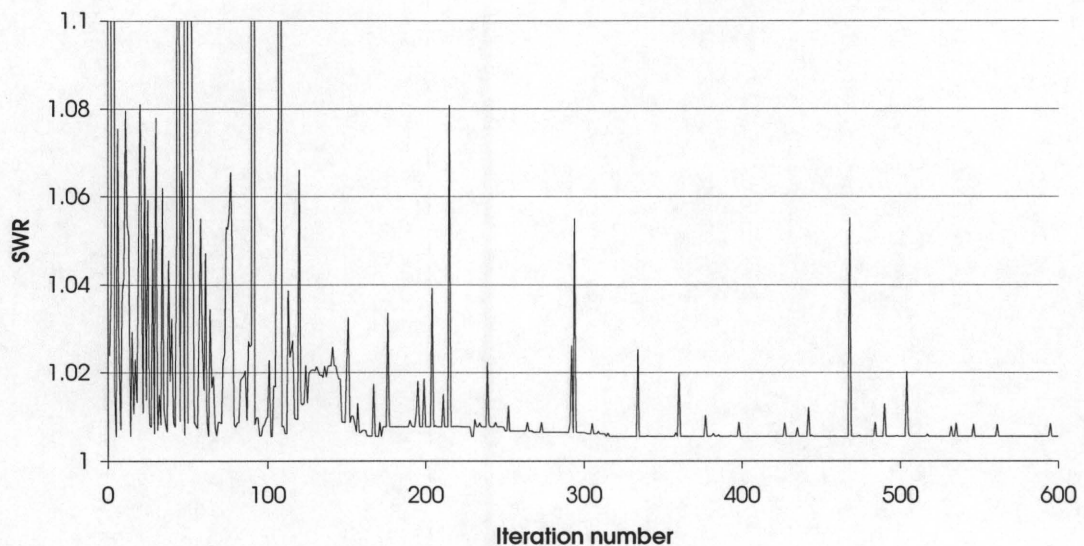


FIGURE 31. Results of optimization of SWR by varying the maximum absorber value for three different absorber widths



The acceptance of absorber values that result in higher SWR values can be seen in the early iterations of Figure 32. Although too large an initial temperature led to an excessive number of iterations to be required to find the solution of such a simple problem, it does allow us to better see the process that the SA algorithm uses to find a solution. One can see the large swings in the SWR in the early iterations of the algorithm, as would be expected with the high temperature value assigned at the time. These swings to large SWR values become less frequent as the temperature cools and the values chosen for the next iteration are closer to the current value. Also note that as the temperature decreases the probability of accepting a point with a significantly worse FoM is much lower. In the higher iterations, absorber values that result in a large SWR are rejected as can be seen by the sharp spikes, immediately followed by a return to the previous value.

The SA algorithm found a minimum in a systematic fashion for a given waveguide structure faster than would be possible using trial and error.



**FIGURE 32.** SWR as function of iteration number of the simulated annealing algorithm

In this chapter I have presented the testing of the algorithm to be used against experimental results and against other numerical approaches. The individual parts of the algorithm have been separately tested to ensure their accuracy compared to known theoretical and experimental results. The low relative error encountered implies that designs made using the program may be viewed with confidence. The iterative method was found to result in the same solution as directly inverting the matrix, however the amount of processing time involved is greater for simulations with the non-zero boundary condition far from the most distant grid point. Unfortunately, as a result of the large memory requirements for the direct-solution method, I was forced to use the iterative method for the great majority of the simulations that I performed.

## Chapter 5. Experimental Results

This chapter seeks to demonstrate how the algorithm may be used as a design tool for the design of optical diffractive elements. The examples given are by no means meant to represent an exhaustive set of the type structures that could be designed. With relatively minor changes to the code quite different structures may be designed. The examples given here represent four very different optical structures. The first is a dual structure comprised of two surface-relief gratings on two different waveguide layers. Their relative separation is too small to consider the gratings independently. The second is a metallic mirror used to couple a waveguide mode into a free-space mode traveling normal to the waveguide. The third is a grating coupler that again takes a guided mode and results in a near-normal free-space mode. The final example is a free-space to free-space diffractive element that serves to redirect a plane wave from a  $45^\circ$  angle in fused silica to a normally exiting plane wave in air.

### 5.1 Grating-Grating Interaction

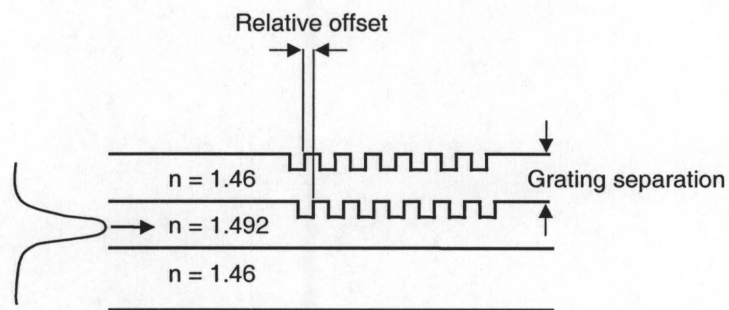
Even relatively simplistic gratings become impossible to correctly solve analytically when two gratings are brought into close proximity to each other. At this point the two gratings do not interact only with the far-field patterns produced by the other gratings, but rather the near-field patterns overlap. In the literature are examples of structures made with gratings separated for the sole purpose of keeping their near fields from interacting [9]. This near-field interaction could be used to the benefit of the



application however. For example, the design could allow for preferential coupling into the superstrate, rather than downward into the substrate. Designing a structure to take advantage of these interactions is quite simplistic with the program.

Let us assume a multilayer waveguide where there are two gratings. The first lies on the guiding layer while the second lies on a superstrate layer covering the guiding layer. If this superstrate layer is of sufficient thickness then the two gratings may be considered independently, the first as an out-coupling grating, and the second as a means to re-direct the light into the air. The two may be placed closer together and used to complement each other. A simple means of illustrating this is to take two gratings and allow the program to vary the layer thickness between the two gratings and the relative offset of one grating to the other to maximize the out-coupled light into the superstrate.

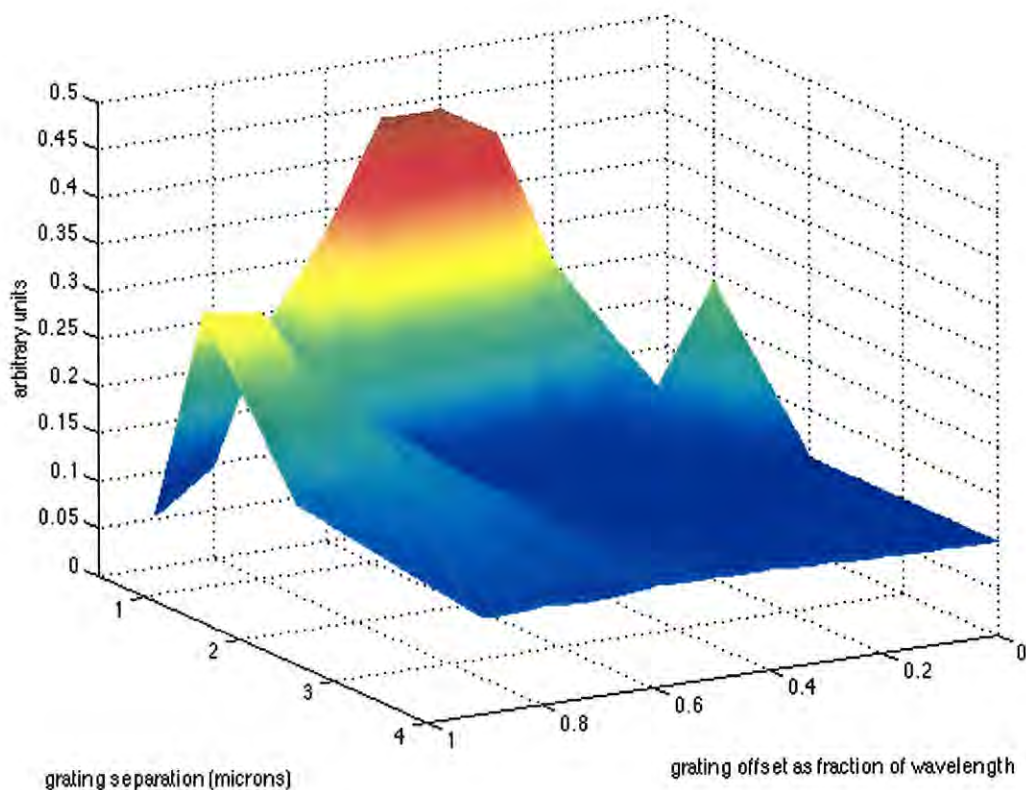
Figure 33 shows the configuration used for this demonstration. The material parameters and wavelength are those used in [9]. The grating period,  $\Lambda$ , is  $0.8\mu\text{m}$ , while the free-space wavelength of the incident light,  $\lambda_0$ , is  $1.3\mu\text{m}$ .



**FIGURE 33.** Configuration of example of grating-grating interaction



In the simulation, the thickness of the layer above the guiding layer is allowed to vary from  $0.5\mu\text{m}$  to  $4.0\mu\text{m}$ . This range was chosen since a value less than  $0.5\mu\text{m}$  may be difficult to fabricate, while by  $4.0\mu\text{m}$  we should begin to see the two gratings acting independently as the evanescent fields separate. The figure of merit is to maximize the output power into the superstrate. To measure this, the electric field amplitude in the last row of grid points before the onset of the absorbing region above the cover of the waveguide is used. The field amplitude at each grid point is squared and then these values summed along the length of the row. The resulting output is shown in Figure 34.



**FIGURE 34.** Variation of power coupled into the cover region by two gratings as a function of grating separation distance and grating offset relative to each other

The variation as a function of grating separation is predictable. As the gratings are separated farther from each other the offset of one grating from the other in the direction of propagation makes less of a difference. Finally at a distance of  $4.0\mu\text{m}$  there is almost no variation as the only the far-field pattern from the first grating interacts with the second one as can be seen by the near flat line at the front edge of Figure 34. However, at a distance of  $0.5\mu\text{m}$  we see a considerable effect of one grating's relative location in the direction of propagation to the other one. If the two gratings could be aligned with sufficient accuracy to one another during fabrication, then this interaction could be used to greatly enhance the outcoupled energy from the top surface of the waveguide. At a shift of  $0.4\Lambda$ , or  $0.32\mu\text{m}$ , the energy coupled through the top of the waveguide is a factor of 4.0 larger than the energy coupled out when no interaction is possible due to a wide separation between gratings.

## 5.2 Forty Five Degree Mirror Couplers

Another example of a device seen in the literature that could be designed using my program is a resin-based  $45^\circ$  output coupling mirror [42]. An output coupler from a waveguide mode to a free-space mode can be fabricated with a mirror placed at around  $45^\circ$  to the direction of propagation in the waveguide as shown in Figure 35.

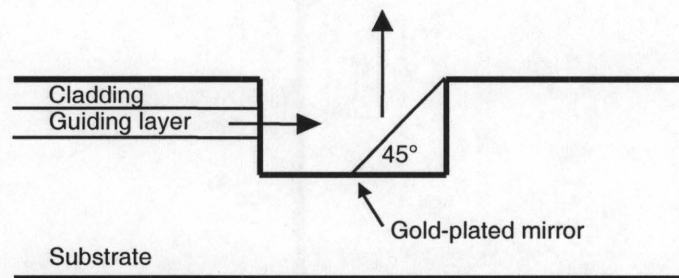


FIGURE 35. Generic forty-five degree mirror configuration

The waveguide ends with a vertical facet with the light coupling out of the waveguide into a free-space mode in a manner similar to light exiting from an optical fiber. This light is still co-linear with the waveguide. The angled mirror is used to redirect this energy to propagate perpendicular to the surface of the waveguide. Unlike in the case of a grating coupler the mirror coupler must be placed at the end of a waveguide and cannot be used for the partial out-coupling of energy from a waveguide. The formation of this mirror has been achieved by various methods. The mirror may be cut directly into the surface of the waveguide using a diamond blade with a  $45^\circ$  angle on the saw blade. The mirror may also be made by making a rectangular groove across the waveguide as illustrated by the dark lines in Figure 35. A liquid resin is then used to form a sloped-sidewall on which a gold plating may be added. This approach will be further investigated below as it offers a greater degree of design freedom than does the diamond blade method.

There are several variables that the designer may vary when building such a mirror. The distance from the end of the waveguide to the mirror; the depth into the substrate that the groove extends; the angle of the mirror surface; and the radius of curvature of the mirror surface. The curvature in the mirror surface is the result of using a liquid resin. This curvature may be controlled and thus adds another variable that may be optimized in the design of the mirror. The curvature also allows the mirror to act as a focusing element. These four variables that may be varied in the design process are labeled in Figure 36.



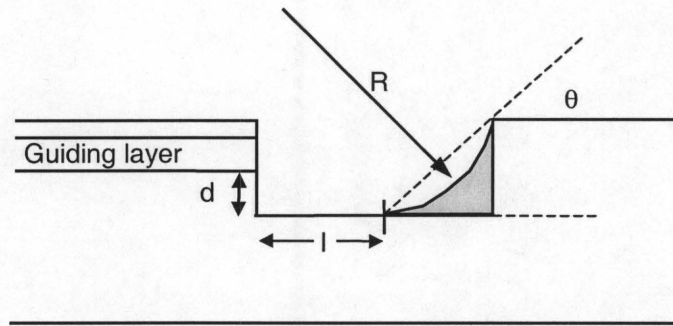


FIGURE 36. Resin mirror design variables

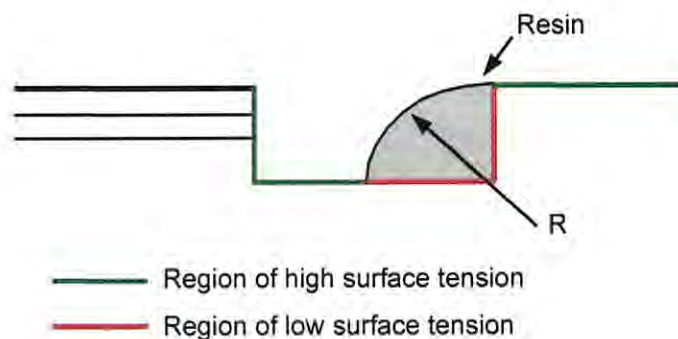
The first two design variables, the distance from the waveguide to the mirror and the depth of the groove into the substrate, are straight-forward in their implementation. The width of the groove is simply made wider or narrower, and likewise the depth of the groove deeper or shallower. The distance to the mirror is an important factor since this variable determines to what extent the beam is allowed to diverge before striking the mirror. In a single mode waveguide the divergence will be rapid. Since the waveguide mode has an evanescent tail in the substrate it is logical that the groove will have to be deeper than the guiding layer or some of the power may be lost into the substrate.

The final two design variables are related; they are the result of the material properties of the resin that is introduced as a liquid in the groove corner opposite the waveguide. The resin is then solidified and a gold plating added. The mirror's surface is therefore the shape of the hardened drop of liquid. The physics of the liquid interacting with the surface of the groove therefore determines the mirror's properties, both its angle and its curvature. Since the dimensions are so small for these mirrors the effect of gravity is minimal; it is the surface tension that primarily controls the droplet's shape.

The angle of the mirror is determined by where in the groove the base of the mirror starts. If the base of the mirror is close to the back surface of the rectangular



groove then a steeper mirror is formed than if the base of the mirror is close to the waveguide-side of the groove. If a liquid resin is used for the wedge of material behind the mirror then the point at which the mirror's base is fixed is determined by the surface tension and surface wetability of the materials used. The surface wetability can be changed by treating the surface at the base of the groove prior to addition of the resin material [42]. Such treatment is analogous to waxing the paint on an automobile--after waxing the water will form droplets with steep angles to the surface of the paint whereas an unwaxed surface results in the sheeting of the water. Treating the part of the base of the groove such that the resin will not sheet allows the point of the base of the resin to be fixed. Notice that the radius of curvature of the drop of resin is inverted in Figure 37 compared to the mirror surface in Figure 36; this inversion occurs as a result of shrinkage during the curing process. The degree of shrinkage and its effect on the radius of curvature are known and may therefore be accounted for in the design to produce the desired radius of curvature in the final device.



**FIGURE 37.** Resin droplet limited by surfaces of differing surface tension

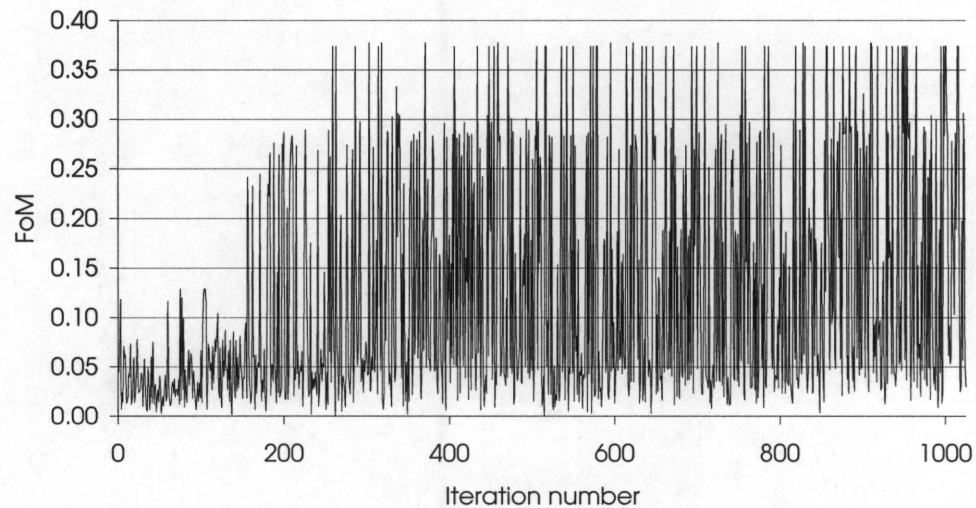
A simulation of this device was run to determine the optimal values for each of these four variables so as to produce a Gaussian beam with a given beam since a Gaussian form is usually preferred in free-space propagation. The simulation space was chosen to have a length of 400 grid points in the direction of propagation, with 360 grid points in the direction perpendicular to propagation. The mirrored device was centered in this region. Twenty points per free-space wavelength were used in the simulation. A wavelength of  $1.55\mu\text{m}$  was used with a refractive index for the guiding layer of 1.485, and for the substrate and cladding a value of 1.386. The gold was placed as a uniform layer if viewed vertically, as if the gold was evaporated onto the surface. The refractive index for gold at  $1.55\mu\text{m}$  was used. The ranges of the four variables to be optimized as defined in Figure 36 are given in Table 4.

variable	Allowed range
d	0-3.0 $\mu\text{m}$
l	0-3.0 $\mu\text{m}$
R	8-30 $\mu\text{m}$
$\theta$	0.4-0.9 radians

**TABLE 4.** Allowed range for variables in design of angled mirror coupler

The results of the four-dimensional search are shown in the following figures. The figure of merit for the design was an overlap integral of a Gaussian beam of beam waist  $5.1\mu\text{m}$  a distance of  $3.64\mu\text{m}$  above the surface of the waveguide. These values

were largely chosen for simplicity of calculations in the configuration used in the design. Other values could have been used. The output beam was not expected to be centered in the simulation region, since as the distance to the mirror was varied the output beam would likewise shift in position, and we do not care if the beam is shifted by a small amount. This shifting was accounted for in the calculation of the FoM by centering the maximum of the ideal beam with the maximum of the out-coupled beam. Figure 38 shows the progression of the figure of merit as the simulated annealing algorithm progressed.

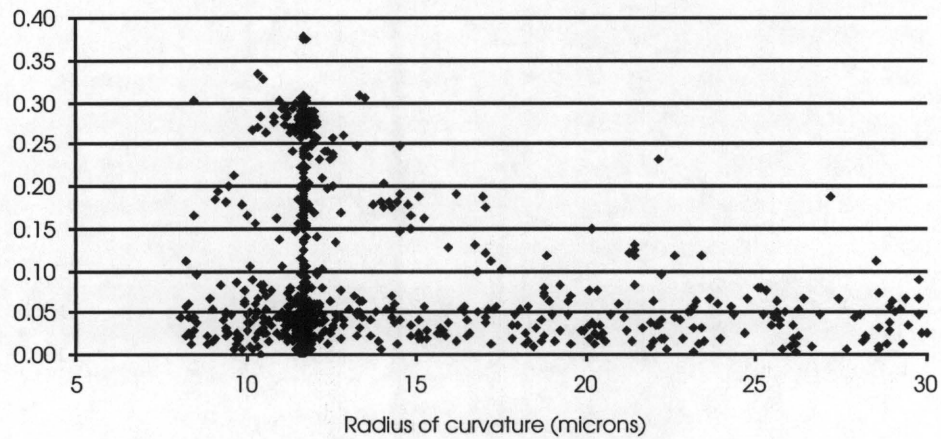


**FIGURE 38.** Figure of merit at each iteration of the annealing algorithm for the resin mirror design

Although the graph appears to be erratic, we do notice that the figure of merit does not increase beyond a certain value. After looking at the data being produced, the algorithm was terminated since a clear maximum had been found. . If the simulation had been run with a lower initial temperature then the erratic jumps in FoM could have been lessened in later iterations. The maxima can be seen in the data for the FoM as projected

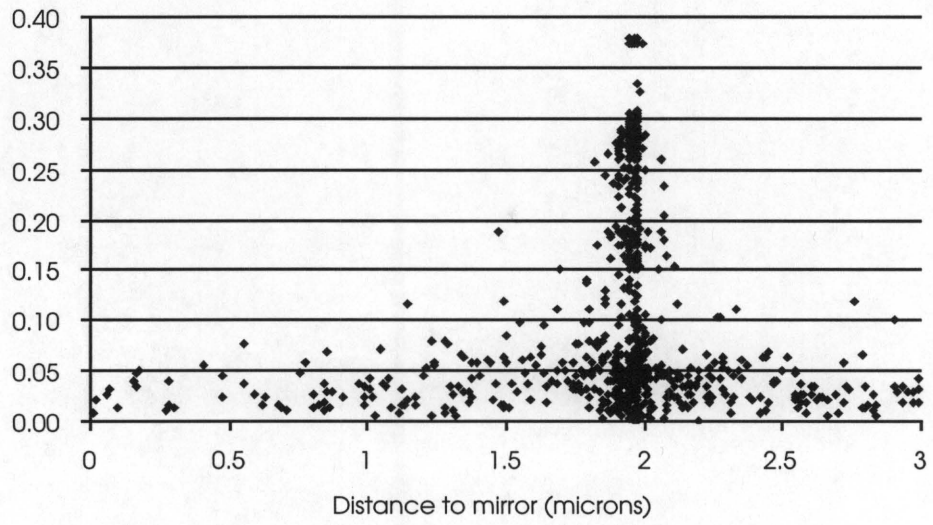


onto each of the four axes in the following four figures. From these figures we can see a clear maximum in each case. It is interesting to note the sharpness of the peaks in the four cases, as this sharpness gives us an idea as to the sensitivity of the FoM to each parameter. Fabrication tolerances must be tight enough to allow the distances found to be able to actually obtain the designed structure. For example, the figure of merit is highly sensitive to the tilt angle of the mirror as seen in Figure 42.

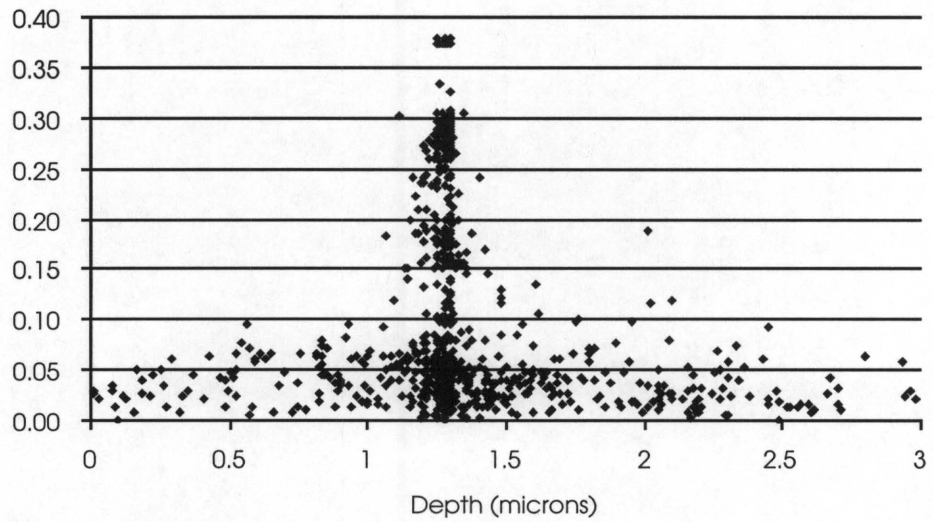


**FIGURE 39.** Figure of merit as a function of the radius of curvature of the mirror

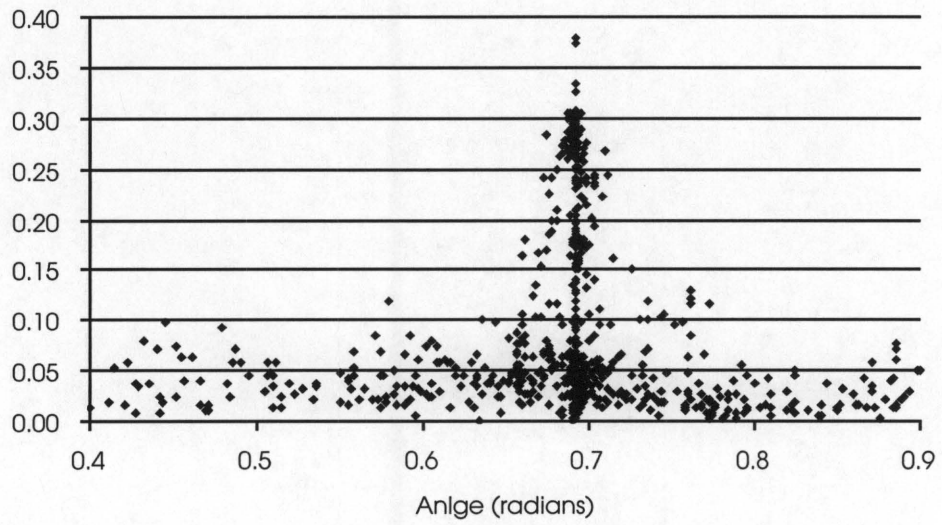




**FIGURE 40.** Figure of merit as a function of the distance from the waveguide to the base of the mirror



**FIGURE 41.** Figure of merit as a function of the depth of the groove below the guiding layer



**FIGURE 42.** Figure of merit as a function of the tilt angle of the mirror

The final values that resulted in the highest figure of merit for the resin mirror device are as follows:

variable	Optimal value
d	1.30 $\mu\text{m}$
l	1.97 $\mu\text{m}$
R	11.72 $\mu\text{m}$
$\theta$	0.692 radians

**TABLE 5.** Optimal design values for resin mirror device

### 5.3 Focusing Volume Grating Couplers

I have designed a volume grating coupler using a five-dimensional search space. The criterion that was to be optimized was that the amplitude of the out-coupled beam should have a Gaussian form with a beam waist of a given width. This was accomplished, however, the out-coupled mode had phase-fronts that indicated a highly-diverging beam. I then varied the search parameters. Now a seven-dimensional search space was employed, the index modulation was allowed to vary following a cubic spline, while the grating period varied along a quadratic spline. This method differs from those described in the previous examples in that each parameter of the structure does not form one dimension of the search space. Instead two fundamental grating parameters form seven dimensions in the search space by each being parameterized by multiple variables. By using this method the grating period and refractive index modulation are allowed to vary along the length of the waveguide, as opposed to constants. The FoM was again to achieve a Gaussian beam, but this time more than just the amplitude was considered. I forced the out-coupled beam to have a flat phase-front a given distance above the surface of the waveguide.

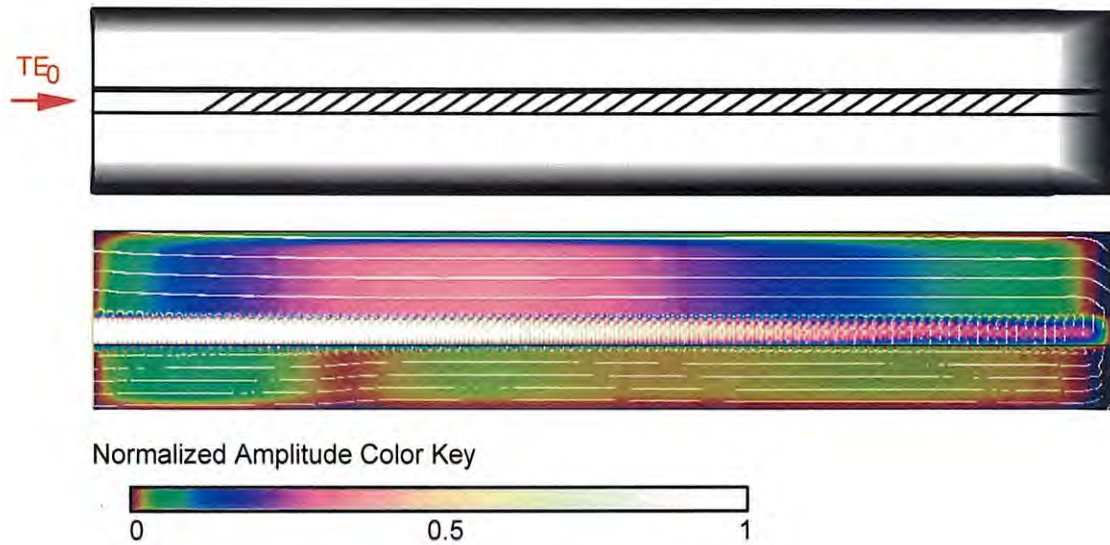
The goal of the design process was to obtain a volume grating that would out-couple a Gaussian beam with a beam waist occurring at a distance of 2 microns above the surface of the waveguide. The FoM was defined to be the overlap integral of a Gaussian beam with the output of the structure after each iteration. The Gaussian beam had a fixed position over the waveguide in this case, forcing the output to have a maximum at a predetermined location to maximize the FoM. The row of data points immediately below the absorber region at the top of the simulation region was used for the grating output (a

distance of  $2\mu\text{m}$  above the waveguide surface). The chosen FoM encourages not only the outcoupling of power in the correct amplitude shape, but also with the same phase as the ideal output waveform. If the phase of one was skewed from the other, then the successive differences in phase would lead to the canceling of terms when the values were added together for all data points.

The period and index modulation of a volume grating were allowed to vary along the length of the guide. The index modulation was allowed to vary cubically (four degrees of freedom), while the period was allowed to vary quadratically (three degrees of freedom). The range over which each of the three grating period variables could vary was limited to values around the range of the grating period that would produce a normally-outcoupled beam. The index modulation was allowed to vary from 0 to a maximum of 0.2. This value is higher than is currently possible to obtain with given photorefractive materials that can only achieve a  $\Delta n$  of 0.06 or less. However, since the simulation space must be relatively small as a result of computing limitations the increase to 0.2 allows a large amount of the incident power to be outcoupled over a relatively small distance in the direction of propagation.

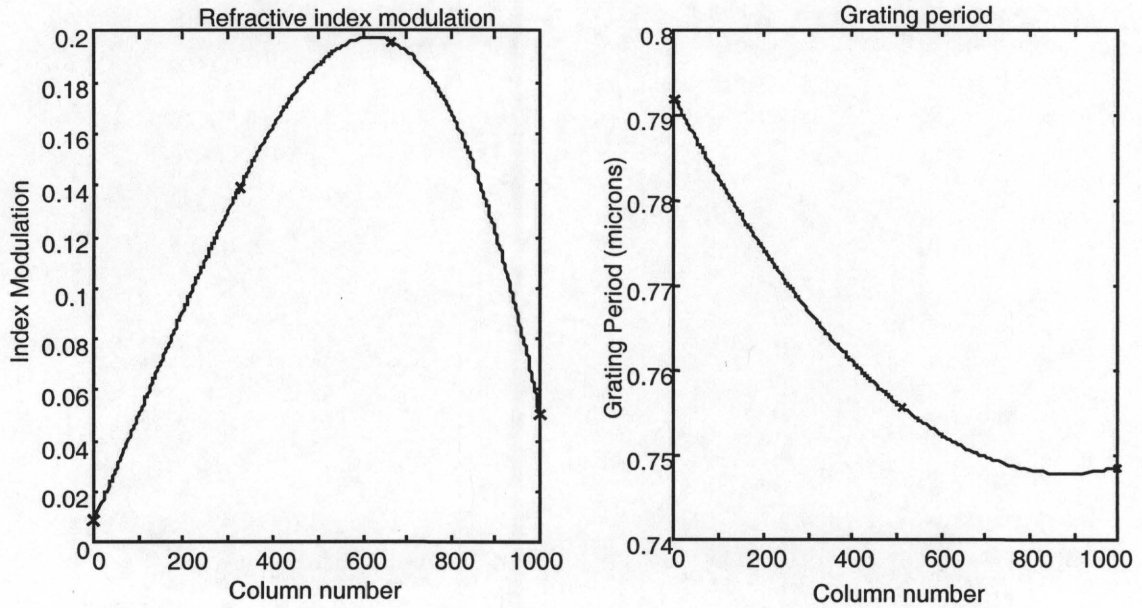
The waveguide structure used is shown in the top of Figure 43, the resulting amplitude profile of the electric field is shown below it. The waveguide has the indices of refraction and layer thicknesses of waveguide 1 of Figure 24. The white lines in the substrate and air regions represent equiphase lines. Note the phasefronts seen at the top of the simulation region are near horizontal, indicating light coupling vertically out of the waveguide, and with a beam waist at that point.





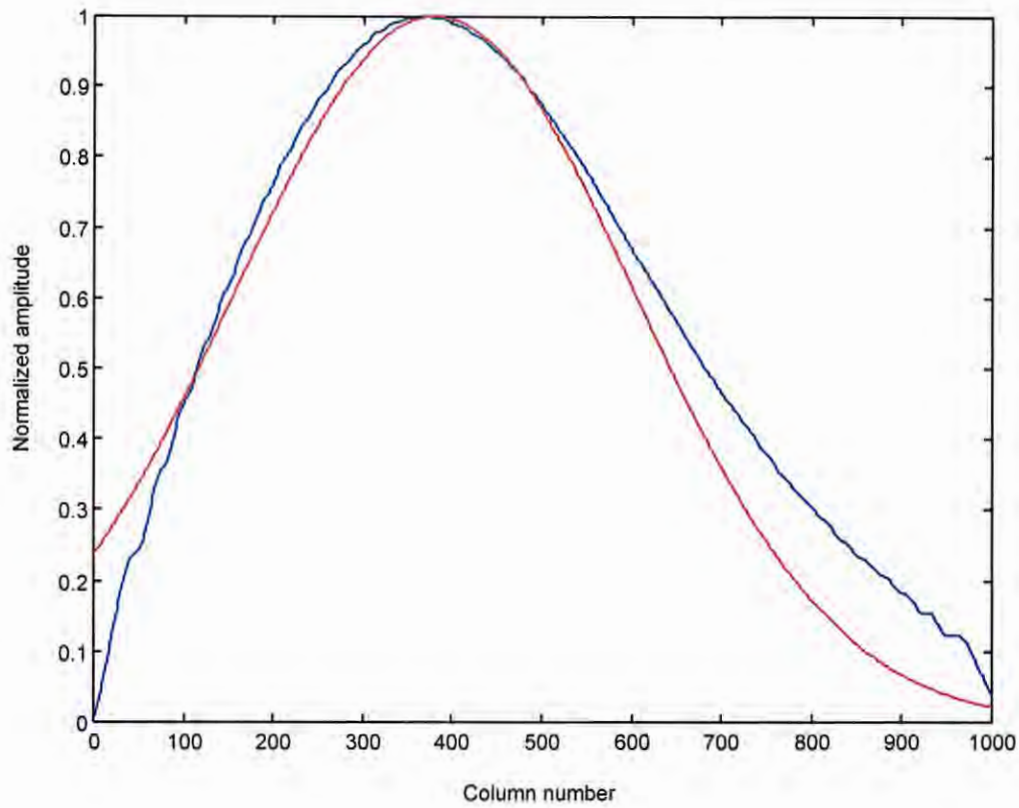
**FIGURE 43.** Results from seven-dimensional optimization of a volume grating

The modulation of the index of refraction and the variation of the period as a function of length that resulted in the optimal result are shown in Figure 44. Marks on the graphs denote points where the splines were defined by the program. From these four points in the index modulation plot, and three points in the grating period plot, the curves that are shown were defined.



**FIGURE 44.** Polynomial curves for focused Gaussian output beam

In Figure 45, the amplitude of the outcoupled beam (in blue) as measured above the waveguide is shown compared to the ideal beam (in red). Since the powers associated with these beams would correspond to the amplitudes squared, the regions of greatest deviation between the two curves would appear to be smaller as they occur at lower amplitudes.



**FIGURE 45.** Comparison of amplitudes of ideal beam and beam resulting from design process

From the phasefronts shown in Figure 43 we can see that the additional degrees of freedom allowed by the seven-dimensional search space resulted in the ability to design a structure where not only the magnitude, but also the phase of the out-coupled beam, could be controlled. The use of the polynomial splines to define the index modulation and grating period along the length of the waveguide allowed the optimization to occur as it reduced the number of dimensions to seven. To date, this is the largest number of dimensions over which a design has been attempted using this program. As this design had not only the largest number of dimensions, but was also large in the number of grid points in the analysis of the electric field, a brief note should also be made here about the

computer times required. A Sun SparcStation 2 Creator was used for the simulation. The program was run in the background, with little other computational activity performed on the machine. A typical simulation such as the one described above would require around 1000 iterations of the annealing loop to converge to a solution. Each electric field analysis would take around 20 minutes using optimized code, thus a full simulation would take approximately two weeks to run.

#### 5.4 Transmissive Diffractive Element

A transmissive grating was designed to redirect a plane wave. The grating was to be etched into fused silica (glass) using multiple mask steps. A beam propagating at  $45^\circ$  inside the fused silica plate was to be redirected to a normal beam in the air as shown in Figure 46.

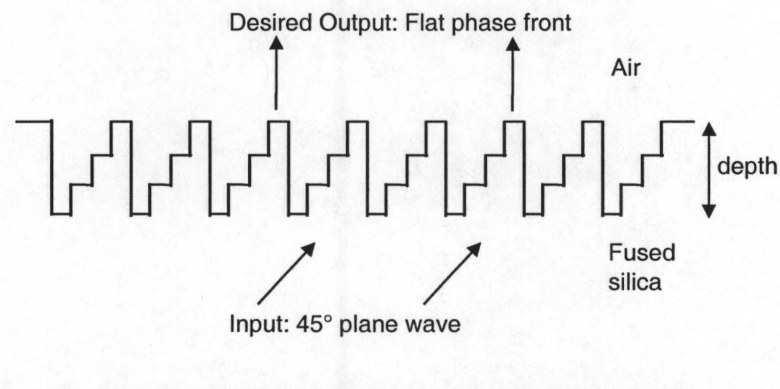


FIGURE 46. Diffractive element to be optimized

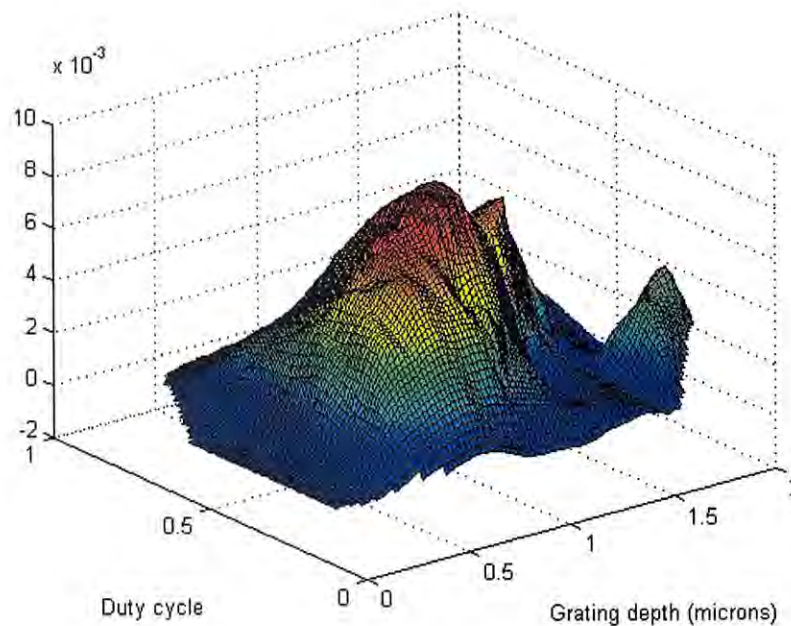
The Gsolver program was used to determine the optimal grating depth. The grating period is not a variable since the angle of incidence and transmission are fixed.



For the desired angles a grating period of  $0.829\mu\text{m}$  is required. The Gsolver program determined that the maximum deflection into the correct order would occur at  $1.12\mu\text{m}$ . This value was found assuming a 50% duty cycle on a four-layer (two mask) grating. A second simulation using eight layers (three masks) did not lead to a significantly improved result. It would, however, cause increased difficulty in fabrication and was therefore not considered. Conversations with the fabrication staff at Sandia National Labs led to determining what values were reasonable for such gratings in fused silica. As a result of learning the fabrication difficulties of etching fused silica with high aspect ratios, the total depth of the grating was limited in all simulations to be less than  $2.0\mu\text{m}$ . Although varying the duty cycle is listed as an option in Gsolver, questionable results led to conversations with the writer of the program that confirmed that the option is undefined for multiple layer gratings. Therefore, the only Gsolver result is for a grating with a duty cycle of 50%.

The same element was then designed using my program. Both the grating depth and duty cycle were allowed to vary. Duty cycle is not commonly defined for multilayer gratings. However, in the design of the masks used to create these structures there is no reason that the levels must all be of equal width. To take advantage of this extra degree of freedom that could be exploited, I derived a definition for duty cycles for these structures for use in the program. The derivation is shown in the Appendix. The simulated annealing algorithm was allowed to pick a duty cycle between 0 and 1. The formulas derived in the Appendix were then used to determine the width of each of the levels in the grating. As with the Gsolver program, the depth was allowed to vary up to  $2.0\mu\text{m}$ . The FoM in this simulation was taken from the top row of data in the simulation

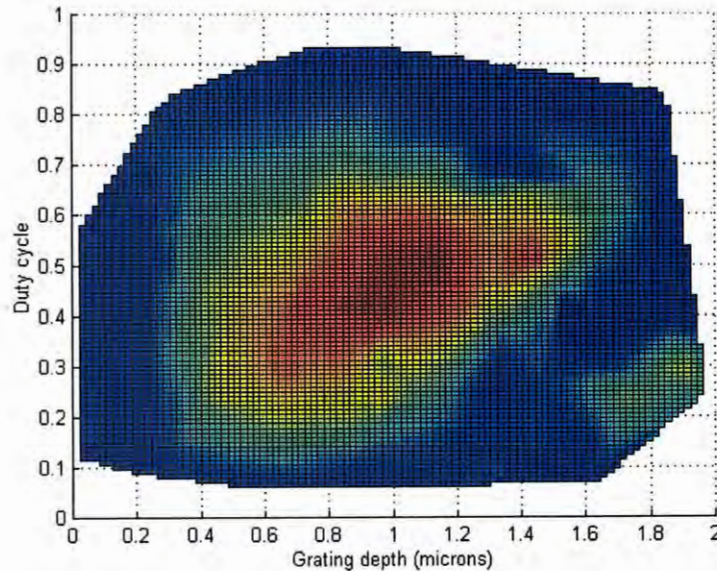
region, immediately before the onset of the absorbing region. The dc component of the fft was to be maximized. This FoM thus tries to maximize power into a plane wave exiting normally from the grating. The plots of the FoM for these simulations are shown in Figure 47 and Figure 48. Note that the maximum is a broad region, allowing for some fabrication error in both the duty cycle and grating depth.



**FIGURE 47.** Figure of merit as a function of duty cycle and grating depth for a transmissive diffractive element

A similar maximum to the one found by the Gsolver program was found when the duty cycle was limited to 50%, the value found by my program being  $1.10\mu\text{m}$ , as opposed to the  $1.12\mu\text{m}$  found by Gsolver. However, if both the duty cycle and the grating depth were allowed to vary, then a higher maximum can be seen to occur at a duty cycle of 42%, as defined by the definition in the Appendix, and a grating depth of

0.92 $\mu\text{m}$ . Efforts to fabricate this device were hampered by the large depth of the grating for such a small period. The deepest ridge in the grating proved especially troublesome, as fused silica does not etch as easily as many other materials.



**FIGURE 48.** Figure of merit as a function of duty cycle and grating depth for a transmissive diffractive element (two-dimensional view)

In this chapter I have shown a sampling of diffractive optic structures that can be designed with the program I have written. These examples are not meant to represent an exhaustive set of the type structures that may be designed, however it is hoped that the variety of structures described illustrates the versatility of the program. Both free-space and guided-wave structures have been demonstrated.



## Chapter 6. Summary of Results and Conclusion

I have developed a new method for the optimal design of diffractive optical structures such as out-coupling devices for planar waveguides. The method employs the scalar Helmholtz equation, allowing modeling of light traveling in any direction inside the simulation region. The scalar formulation is necessitated by the computational power limitations that currently exist. The fields to be simulated may be guided modes, free-space modes, or a combination of the two.

In previous programs for the design and analysis of diffractive structures the method used methods for the analysis of the fields that force the user to make assumptions on the parameters of the structure, such as infinite grating extent. The program described in this thesis makes no such assumptions. The user of the program does not need to be an expert in the field, as long as he is able to set the limits on the parameters to be varied. These limitations on parameter values could be the result of the limits of fabrication, thus the optimal structure that can actually be constructed can be found.

These parameters are varied by the simulated annealing algorithm to find the optimal design of a grating or of a mirror. Through the use of polynomial splines the parameters of the out-coupling gratings may be modeled using a low number of parameters; this allows the simulated annealing algorithm to search over a relatively small number of dimensions while allowing freedom for a large number of possible structures. The method is flexible in that it can simulate a variety of out-coupling structures and free-space diffractive elements.



The use of a search space with low dimensionality is of prime importance. With an analysis routine that takes several minutes to run, a method had to be devised to limit the number of iterations required by the simulated annealing algorithm. This can be achieved by limiting the number of dimensions in the search space. In some structures the parameters to be varied easily conform to a low-dimensional search space. In other, more complex structures such as the volume grating, the use of polynomial splines to characterize the device to be designed was required to maintain a low number of dimensions while allowing sufficient freedom of design.

The method was first tested against theoretical and experimental results, with the outcome being deemed successful. Sample structures were then designed by the program; these structures span a large range of device type. Both guided wave and free-space applications were considered. Multiple diffractive elements in one design were considered, allowing the evanescent fields surrounding the two elements to interact. Complex volume gratings whose properties varied in the direction of guided-mode propagation were parameterized and designed. An out-coupling reflective structure was also designed. To illustrate the flexibility of the program, a free-space to free-space diffractive element that is unlike the guided-mode applications was also designed. Many of these designs are examples taken from the literature or current research activities, thus showing the usefulness of this program.

Future work that would lead to improvements to this code could take two tracks. One option would be to replace the standard five-point representation of the discretized scalar Helmholtz equation with a more complicated expression. Although this would seemingly make the code more complex and thus run slower the reverse might actually

be the case as a less fine mesh would be required [43]. A nine-point representation would lead to a relaxation of the requirements on mesh fineness while only slightly increasing the bandwidth of the banded-matrix in the direct solution method. It would, however, preclude the use of the tridiagonal matrices produced by the current ADI method.

The second track that could lead to significant advances in the usefulness of the program is simply to wait. Throughout the design and testing of this computer code, many of the decisions made as to the algorithms to be used and the type problems that could be solved were the result of trying to minimize either the computer processor or memory requirements. It is unfortunate that many decisions as to the direction of the development of the code had to be made for such reasons unrelated to the purpose of the code. However, the rapid increase of computer capabilities with time offers a method for the code to become increasingly useful for more complex structure design without any changes to the code required. Larger structures could be designed in shorter amounts of time. As computer memory sizes available to the average computer user increase the use of the direct-solution method by using Gaussian elimination on the large banded matrix becomes a much more attractive option for solving even larger problems than is currently possible.

Major variations in the code itself could lead to the possible simulation of more complex structures. By changing the internal analysis routine to a full three-dimensional vector-based analysis of the fields then the type of structures that could be simulated would be greatly expanded. A more realistic first step would be the coding of a two-dimensional vector code. Again, this change would require many times the computer

processor and memory requirements of the current code. However when the time comes, any of these changes would be relatively simplistic to do given the current modular framework of the algorithm. As the analysis routine operates independently of the annealing algorithm, it may be replaced in the future when the computer technology allows for it. In summary, the framework of the algorithm described in this thesis allows for considerable variation and future expansion and may be updated by future users to remain current.

## Chapter 7. References

- [1] J. Wu, T. Kondo, and R. Ito, "Design of waveguide grating based on simulated annealing and narrow-band modulation," *Jpn. J. Appl. Phys.*, vol 32, pp. 4973-4979, 1993.
- [2] J. Wu, T. Kondo, and R. Ito, "Optimal Design for Broadband Quasi-Phase-Matched Second-Harmonic Generation Using Simulated Annealing," *J. of Lightwave Tech.*, vol 13, pp. 456-460, 1995.
- [3] B. Lunitz and J. Jahns, "Tolerant design of a planar-optical clock distribution system," *Optics Comm.*, vol. 134, pp. 281-8, 15 Jan. 1997.
- [4] T. Li, S. Tang, F. Li, M. Dubinovsky, L. Wu, R. Wickman, and R. T. Chen, "Polymer waveguide based 1-GHz clock signal distribution system" in *Optical characterization techniques for high-performance microelectronic device manufacturing III*, D. DeBusk, R. T. Chen, eds. Austin, Texas : SPIE vol 2877, pp. 111-117, Oct. 1996.
- [5] D. A. B. Miller and H. M. Ozaktas, "Limit to the bit-rate capacity of electrical interconnects from the aspect ratio of the system architecture," *J. of Parallel and Distributed Computing*, vol. 41, pp. 42-52, 25 Feb. 1997.



- [6] A. V. Krishnamoorthy and D. A. B. Miller, "Firehose architectures for free-space optically interconnected VLSI circuits," *J. of Parallel and Distributed Computing*, vol. 41, pp. 109-114, Feb. 1997.
- [7] Semiconductor Industry Association. *International Technology Roadmap for Semiconductors: 1999 edition*. Austin, TX:International SEMATECH, 1999.
- [8] P. K. Tien and R. Ulrich, "Theory of prism-film coupler and thin-film light guides," *J. of the Opt. Soc. Of Am.*, vol. 60, pp. 1325-1350, Oct. 1970.
- [9] S. Sheard, T. Liao, G. Yang, P. Prewett, and J. Zhu, "Focusing waveguide grating coupler using diffractive doublet," *Appl. Opt.*, vol. 36, pp. 4349-4353, 1 July 1997.
- [10] W. K. Smothers, B. M. Monroe, A. M. Weber, and D. E. keys, "Photopolymers for holography," in *Practical holography IV*, Stephen A. Benton ed. Los Angeles, CA: SPIE, vol. 1212, pp. 20-28, 1990.
- [11] A. M. Weber, W. K. Smothers, T. J. Trout, and D. J. Mickish, "Hologram recording in Du Pont's new photopolymer materials," in *Practical holography IV*, Stephen A. Benton ed. Los Angeles, CA: SPIE, vol. 1212, pp. 20-28, 1990.

[12] Q. Huang, and P. R. Ashley, "Holographic Bragg grating input-output couplers for polymer waveguides at an 850nm wavelength," *Appl. Opt.*, vol 36, pp. 1198-1203, 20 Feb. 1997.

[13] V. P. Tzolov, D. Feng, S. Tanev, and Z. J. Jakubczyk, "Modeling tools for integrated and fiber optical devices," in *Integrated Optics III*, G. C. Righini and S. I. Najafi, eds. San Jose, CA : SPIE vol. 3620, pp. 162-173, Jan. 1999.

[14] G. R. Hadley, "Multistep method for wide-angle beam propagation," *Opt. Letters*, vol. 17, pp. 1743-1745, 15 Dec. 1992.

[15] C. Vassallo, "The limitations of wide-angle BPM in non-uniform systems," Conference Edition. Summaries of the Papers Presented at the Topical Meeting, Integrated Photonics Research. 1996 Technical Digest Series. vol. 6, pp. 108-110, 1996.

[16] M. K. Moaveni, H. A. Kahlor, and S. Shamma, "Application of finite-elements to the analysis of diffraction gratings," *Int. J. Electronics*, vol. 40, no. 3, pp 225-236, 1976.

[17] C. Yeh, S. B. Dong, and W. Oliver, "Arbitrarily shaped inhomogeneous optical fiber or integrated optical waveguides," *J. Appl. Phys.*, vol. 46, no. 5, pp 2125-2129, May 1975.

- [18] C. Yeh, K. Ha, S. B. Dong, and W. P. Brown, "Single-mode optical waveguides," *Appl. Opt.*, vol. 18, no 10, pp 1490-1504, 15 May 1979.
- [19] M. J. Beaubien and A. Wexler, "An Accurate finite-difference method for higher-order waveguide modes," *IEEE Trans. on Microwave Theory Tech.*, vol MTT-16, no. 12, pp 1007-1017, Dec. 1968.
- [20] E. L. Wachspress and G. J. Habetler, "An alternating-direction-implicit iteration technique," *SIAM J.*, vol. 8, pp. 403-424, June 1960.
- [21] A. Wexler, "Computation of electromagnetic fields," *IEEE Trans. on Microwave Theory Tech.*, vol MTT-17, no. 8, pp 416-439, Aug. 1969.
- [22] N. Mabaya, P. E. Lagasse, and P. Vandenbulcke, "Finite Element Analysis of Optical Waveguides," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-29, no. 6, pp 600-605, June 1981.
- [23] G. R. Hadley, "Numerical simulation of reflecting structures by solution of the two-dimensional Helmholtz equation," *Opt. Lett.*, vol. 19, pp. 84-86, Jan. 15, 1994.
- [24] J. H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, 1975.

- [25] S. Kirkpatrick, C. D. Gelatt Jr, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, vol. 220, pp. 671-680, 13 May 1983.
- [26] C. Vassallo, "Improvement of finite difference methods for step-index optical waveguides," *IEE Proc. J*, vol 139, pp. 137-142, Apr. 1992.
- [27] M. Born and E. Wolf, *Principles of optics, Sixth edition*, Cambridge University Press, Cambridge, 1980.
- [28] J. P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comp. Phys.*, vol. 114, pp. 185-200, 1994.
- [29] Strang, G. *Linear algebra and its applications*, Academic Press, New York, 1976.
- [30] D. W. Peaceman and H. H. Rachford Jr., "The numerical solution of parabolic and elliptic differential equations," *J. Soc. Indust. Appl. Math.*, vol. 3, pp. 28-41, Mar. 1955.
- [31] L. Ingber, "Simulated annealing: practice and theory," *Mathl Comput. Modelling*, vol 18, pp. 29-57, Nov. 1993.
- [32] L. Ingber, "Adaptive Simulated annealing (ASA): lessons learned," *Control and Cybernetics*, vol 25, pp. 33-54, Jan. 1996.



- [33] B. Chandler, C. Rekeczky, Y. Nishio, A. Ushida, "Adaptive simulated annealing in CNN template learning," *IEICE Transactions on Fundamentals of Electronics*, vol. E82-A, pp. 398-402, Feb. 1999.
- [34] S. Chen, B.L. Luk, "Adaptive simulated annealing for optimization in signal processing applications," *Signal Processing*, vol. 79, pp. 117-28, Nov. 1999.
- [35] A.K. Shaw, T. Manickam, "A parameter adaptive simulated annealing algorithm for frequency estimation," in *IEEE International Conference on Systems Engineering*. Fairborn, OH: IEEE, 1989, pp. 213-216.
- [36] Y. Wang, W. Yan, G. Zhang, "Adaptive simulated annealing for the optimal design of electromagnetic devices," *IEEE Transactions on Magnetics*, vol.32, pp.1214-17, May 1996.
- [37] S. Nozaki, Y. Chen, Z. Nakao, "An adaptive simulated annealing applied to optimization of phase distribution of kinoform," in *Proc. of the 5th Intl. Conf. on Soft Computing and Information/Intelligent Systems Methodologies for the Conception, Design and Application of Soft Computing* (T Yamakawa and G. Matsumoto, eds.). Fukuoka, Japan, 1998, pp.979-82.
- [38] Tamir, T., Guided-wave optoelectronics, Springer-Verlag, New York, 1990.

- [39] C. Vassallo, "Improvement of finite difference methods for step-index optical waveguides," *IEE Proc. J*, vol 139, pp. 137-142, Apr. 1992.
- [40] C. Vassallo, "Interest of improved three-point formulas for finite-difference modeling of optical devices," *J. Opt. Soc. Am. A*, vol. 14, pp. 3273-3284, Dec. 1997.
- [41] D. B. Ostrowsky and A. Jacques, "Formation of optical waveguides in photoresist films," *Appl. Phys. Lett.*, vol. 18, pp. 556-557, 15 June 1971.
- [42] H. Terui, K. Shutoh, "Novel micromirror for vertical optical path conversion formed in silica-based PLC using wettability control of resin," *J. of Lightwave Tech.*, vol.16, pp.1631-9, Sept. 1998.
- [43] C. Vassallo, "Interest of improved three-point formulas for finite-difference modeling of optical devices," *J. Opt. Soc. Am. A*, vol. 14, pp. 3273-3284, Dec. 1997.

## Appendix

### Derivation of Duty Cycle for 4-Layer Grating

The assumption is made that the duty cycle should reference the percent of a unit area of the grating that is filled. Two cases are considered: the first where the duty cycle is less than 0.5, the other when it is greater.

#### *Case 1: Duty Cycle < 0.5*

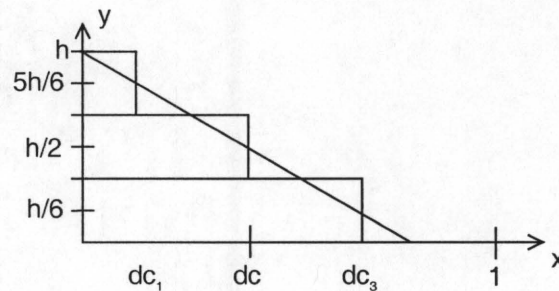


FIGURE A1. Duty cycle of less than 0.5

The above figure shows a duty cycle,  $dc$ , of less than 0.5. The area underneath the diagonal line is less than half the area of the square. The units in the  $x$  direction are normalized to the period of the grating. The length of the middle block is clearly  $dc$ . However,  $dc_1$  and  $dc_3$  must be computed. The equation of the line may be put in the form

$$y = mx + b \quad (\text{A.1})$$

The slope  $m$  may be easily calculated from looking at the figure:

$$m = \frac{-h}{dc} \quad (\text{A.2})$$

Substituting in the known point  $(dc, h/2)$  gives us the value for  $b$ :

$$b = h \quad (\text{A.3})$$

The full equation for the sloped line is thus:

$$y = \frac{-h}{2} x + h \quad (\text{A.4})$$

The value for  $dc_3$  occurs at  $y = h/6$ . Substituting in for  $y$  in A.4 yields:

$$dc_3 = \frac{5}{3} dc \quad (\text{A.5})$$

The value for  $dc_1$  occurs at  $y = 5h/6$ . Substituting in for  $y$  in A.4 yields:

$$dc_1 = \frac{1}{12} dc \quad (\text{A.6})$$

**Case 2: Duty Cycle > 0.5**

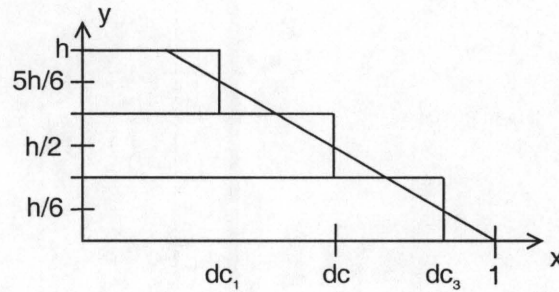


FIGURE A2. Duty cycle of greater than 0.5

The above figure shows a duty cycle,  $dc$ , of greater than 0.5. The area underneath the diagonal line is greater than half the area of the square. The units in the  $x$  direction are normalized to the period of the grating. The length of the middle block is clearly  $dc$ .

However,  $dc_1$  and  $dc_3$  must be computed. The equation of the line may be put in the form

$$y = mx + b \quad (\text{A.7})$$

The slope  $m$  may be easily calculated from looking at the figure:



$$m = \frac{-\frac{h}{2}}{1-dc} \quad (\text{A.8})$$

Substituting in the known point  $(dc, h/2)$  gives us the value for  $b$ :

$$b = \frac{-\frac{h}{2}}{1-dc} \quad (\text{A.9})$$

The full equation for the sloped line is thus:

$$y = \frac{-\frac{h}{2}}{1-dc} x + \frac{\frac{h}{2}}{1-dc} \quad (\text{A.10})$$

The value for  $dc_3$  occurs at  $y = h/6$ . Substituting in for  $y$  in A.10 yields:

$$dc_3 = \frac{5}{3} dc - \frac{2}{3} \quad (\text{A.11})$$

The value for  $dc_1$  occurs at  $y = 5h/6$ . Substituting in for  $y$  in A.10 yields:

$$dc_1 = \frac{1}{3} dc + \frac{2}{3} \quad (\text{A.12})$$

It should be noted that for  $dc = 0.5$  the two sets of equations yield identical results.

## Vita

David W. Peters was born on August 8, 1970 in Huntsville, Alabama. He graduated from Grissom High School in Huntsville in 1988. After receiving a President's Scholarship he graduated from the Georgia Institute of Technology with a Bachelor's of Electrical Engineering with Highest Honor in March 1992. He attended the Georgia Institute of Technology for graduate school after receiving a President's Fellowship and a National Science Foundation fellowship. He received the Master of Science in Electrical Engineering in June 1993. In the summer of 1994 he worked as a visiting researcher at the U. S. Army Research Labs. In the summer of 2000 he worked as a summer intern at Sandia National Labs where he worked on the design of diffractive-optic structures. With the exception of those summers he worked as a graduate teaching assistant in the School of Electrical and Computer Engineering, with the primary responsibility of lecturing the course Circuits and Electronics. He received the institute-wide CETL teaching excellence award in 1999, and the School of Electrical and Computer Engineering Outstanding Teaching Assistant Award four times.