REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE June 1996	3. REPORT TYPE AND DATES COVERED Doctoral Dissertation

4. TITLE AND SUBTITLE

Design of Gradient Index Optical Thin Films

5. FUNDING NUMBERS

6. AUTHOR(S)

Jeffrey J. Druessel, Capt, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology, WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION

AFIT/DS/ENP/96-03

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

WL/MLPJ
Wright-Patterson AFB, OH 45433

10. SPONSORING/MONITORING

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release: distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This dissertation develops an enhancement to existing inverse Fourier transform gradient index design methods, and develops a new optimal design method for gradient index films using a generalized Fourier series approach. Use of an optimal phase function in Fourier-based filter designs reduces the product of index contrast and thickness for desired reflectance spectra. The shape of the reflectance spectrum is recovered with greater fidelity by suppression of Gibbs oscillations and shifting of side-lobes into desired wavelength regions. A new method of gradient index thin film design using generalized Fourier series extends the domain of problems for which gradient index solutions can be found. The method is analogous to existing techniques for layer based coating design, but adds the flexibility of gradient index films. A subset of the coefficients of a generalized Fourier series representation of the gradient index of refraction profile are used as variables in a nonlinear constrained optimization formulation. This method is particularly well suited for the design of coatings for laser applications, where only a few, widely separated wavelength requirements exist. The generalized Fourier series method is extended to determined the minimum film thickness needed, as well as the index of refraction profile for the optimal film.

19960719 076

14. SUBJECT TERMS

Thin Films, Gradient Index, Optical Coatings, Wavelets, Fourier Transforms, Optimal Design, Rugate,

15. NUMBER OF PAGES
181

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

DTIC QUALITY INSPECTED 4

Design of Gradient Index Optical Thin Films

DISSERTATION

Jeffrey J. Druessel, Captain, USAF
AFIT/DS/ENP/96-03

Design of Gradient Index Optical Thin Films

DISSERTATION

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy
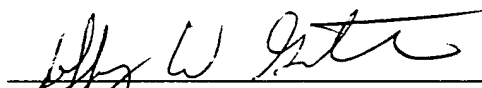
Jeffrey J. Druessel, B.S., M.S., M.S.S.M,
Captain, USAF

June 12, 1996

ii

AFIT/DS/ENP/96-03

Design of Gradient Index Optical Thin Films

Jeffrey J. Druessel, B.S., M.S., M.S.S.M,
Captain, USAF

Approved:
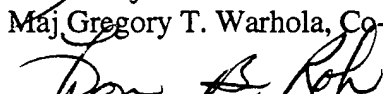
_____   12 Jun 96
Maj Jeffrey W. Grantham, Co-Chairman   Date

_____   14 Jun 96
Maj Gregory T. Warhola, Co-Chairman   Date
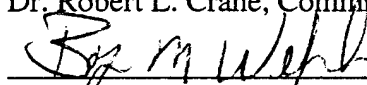
_____   18 Jun 96
Dr. Won B. Roh, Committee Member   Date

RLCrane   18 JUL 96
_____
Dr. Robert L. Crane, Committee Member   Date

_____   18 Jul 96
Dr. Byron M. Welsh, Dean's Representative   Date

Accepted:

_____
Dr. Robert A. Calico, Jr.
Dean, School of Engineering

iii

PREFACE

This research grew out of a desire to create a highly reflective multilayer film with very low resistance for vertical cavity surface emitting semiconductor lasers. The traditional cavity mirrors for these devices are formed by alternating layers of gallium arsenide and aluminum gallium arsenide. The abrupt changes in material which make a good mirror also make a poor conductor, so electrical pumping of these lasers is difficult. One solution is to design a mirror with no abrupt changes in material (lowering the resistance) while maintaining the high reflectance required for the laser cavity mirrors. In the course of the investigation, the original goal of designing, fabricating, and testing a vertical cavity laser device was found to be infeasible in the time available. My research therefore took a more theoretical path, and I decided to address the more general problem of gradient index thin film design.

I would like to express my thanks to my advisor, Maj Jeffrey Grantham, for his aid and support in my ever evolving dissertation. Dr Peter Haaland was also of great help during the development of the SWIFT algorithm. I would also like to thank Maj Gregory Warhola for his willingness to adopt me as my dissertation topic took a more mathematical turn. Finally, I wish to thank my wife Joan for her patience, understanding, and support throughout the years it took to complete my studies.

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Gradient index thin films provide greater flexibility for the design of optical coatings than the more conventional "layer" films. In addition, gradient index films have higher damage thresholds and better adhesion properties. In this dissertation I present an enhancement to the existing inverse Fourier transform gradient index design method, and develop a new optimal design method for gradient index films using a generalized Fourier series approach. The inverse Fourier transform method is modified to include use of the phase of the index profile as a variable in rugate filter design. Use of an optimal phase function in Fourier-based filter designs reduces the product of index contrast and thickness for desired reflectance spectra. The shape of the reflectance spectrum is recovered with greater fidelity by suppression of Gibbs oscillations and shifting of side-lobes into desired wavelength regions. A new method of gradient index thin film design using generalized Fourier series extends the domain of problems for which gradient index solutions can be found. The method is analogous to existing techniques for layer based coating design, but adds the flexibility of gradient index films. A subset of the coefficients of a generalized Fourier series representation of the gradient index of refraction profile are used as variables in a nonlinear constrained optimization formulation. The optimal values of the design coefficients are determined using a sequential quadratic programming algorithm. This method is particularly well suited for the design of coatings for laser applications, where only a few widely separated wavelength requirements exist. The generalized Fourier series method is extended to determined the minimum film thickness needed, as well as the index of refraction profile for the optimal film.

1. INTRODUCTION

1.1. Motivation

The Air Force is interested in optical thin film coatings for a variety of uses, including broad band anti-reflection coatings for low-observable applications, narrow pass filters for sensor protection, and high power laser mirrors. Most current thin film coatings are designed and constructed using a number of slabs of materials with different indices of refraction and thicknesses. The most popular method consists of alternating layers of high and low index material. The properties of the film are determined by the two index values used and the thicknesses of each layer in the film. Recently, there has been a growing interest in inhomogeneous or gradient index thin films, which are characterized by a continuously varying index of refraction throughout the film. The advantages of such gradient index films over the alternating stack type films are based on the elimination of the abrupt interfaces between materials. In a slab based design, very large electric fields can develop at the material interfaces. This can lead to damage of the film in high power laser applications [47:272]. Also, the current deposition methods demonstrate columnar growth patterns in each material. The columnar growth patterns create voids in the structure, which can increase scattering. In addition, the different material properties lead to large inherent stress between layers in the film [47:271-272]. Elimination of the interfaces by using gradient index designs promotes better adhesion to the substrate, reduces internal stress in the film, reduces scattering, and increases the damage threshold for the film [43:2864].

Several basic approaches to thin film design exist, including graphical, analytical, and digital design methods [11:97]. Digital techniques are of particular interest because

they can be used to design films with more complex properties than is possible analytically or graphically. Digital design techniques can be categorized as either refinement or synthesis techniques. Refinement techniques begin with an initial film design and iteratively refine it to achieve a better design. Synthesis techniques on the other hand generate a film design based on the desired film characteristics only. Several methods exist for both refinement of an initial design [11] or synthesis of a film [12]. One of the synthesis techniques relies on the inverse Fourier transform to relate the index of refraction of the film to the transmittance of the film. This technique, originally developed by Sossi and Kard [34,35,36], will be expanded upon in this work. In addition, new design techniques, based on a generalized Fourier series approach, will be developed.

The objective in designing a thin film is to take the performance requirements for the film and generate a design subject to constraints of available indices of refraction and acceptable total thickness. The performance requirements are usually stated as reflectance, transmittance, or absorption versus wavelength, incident polarization or angle. The conventional variables in the design are the number of layers in the film, and the index and/or thickness of each layer. For "simple" filters, such as notch or bandpass filters, anti-reflection coatings, and reflectors, designs by graphical methods or by knowledge of the properties of periodic multilayers are possible [12, 24]. Films with more complex properties require digital design methods.

1.2. Problem Statement

The purpose of this research is to create a design method to map between reflectivity and continuous index of refraction profiles that also allows for additional

2

design constraints, such as limitations on the total film thickness and index of refraction range. The process of designing an optical filter requires a mapping between the space of all possible index profiles and the space of all possible reflectivity characteristics. A mapping from index to reflectivity can be generated from Maxwell's equations. This mapping allows one to evaluate the performance characteristics of a film, in terms of reflectivity as a function of wavelength, angle of incidence, polarization, etc.. Unfortunately, this mapping is not invertable, so it cannot be used to design a film (except by trial and error). Most of the existing design techniques rely on the trial and error approach. The corrections to the index are based on numerical principles of optimal design theory. One exception to this approach is the inverse Fourier transform technique. This design method is derived from Maxwell's equations by making several approximations. However, this method maps one reflectivity profile to many possible index profiles. In addition, there is no mechanism to impose additional constraints on the design, such as available index range or total thickness.

This research is presented in two phases. The first phase is to investigate design of gradient index thin films by the inverse Fourier transform method. The second phase is to create a new method for solving the inverse problem of identifying an index of refraction profile for a given reflectivity, based on a generalized Fourier series approach. Two variations on this theme are explored: a Fourier series method and a wavelet method. The wavelets provide a different framework for analyzing the structure of the index of refraction. A wavelet decomposition of a "signal" is analogous to a Fourier series decomposition, but the information is packaged in terms of "scale" and "shift", rather than the more familiar frequency and phase. Another reason to use wavelets is that each wavelet has a finite extent, as opposed to the sines and cosines of Fourier series. This finite support is helpful because it allows one to focus only on the elements of the design that are important. As an example, the inverse Fourier transform method requires

3

the user to specify the reflectivity over a broader range than the design requirements specify. This is required to achieve sufficient detail in the index of refraction profile. This forces the design to meet more stringent conditions than are really needed. The finite support of the wavelets should allow one to overcome such problems. The theory and concepts of this wavelet formalism are presented in more detail in Appendix C.

1.3. Organization

Chapter 2 presents background information on existing thin film design techniques and the basic theory of thin films. Chapter 3 discusses the inverse Fourier transform method for synthesis of thin films, as well as a modification to the existing theory allowing optimal synthesis of a thin film. Chapter 4 discusses optimal design of gradient index films using Fourier series and wavelet series. The final chapter summarizes the dissertation and presents suggestions for future research in this area. The appendices include an explanation of the numerical considerations involved in implementing the inverse Fourier transform techniques of Chapter 3 (Appendix A), a brief outline of optimization theory (Appendix B), the wavelet theory needed in this research (Appendix C), and the programs used in this work (Appendix D).

2. Background

This chapter presents the historical background of thin film design methods and the current state of the art for gradient index film design and fabrication. In addition, the basic theory of optical properties of thin films necessary for the analysis and synthesis of thin films is presented.

2.1. Historical Perspective

To appreciate the significance of the new approach to gradient index thin film design presented in this dissertation, a brief history of thin film design techniques is needed. Thin films were first discovered in the late 1600's by Robert Hooke and Sir Isaac Newton in the phenomenon known as "Newton's rings" [19:299,24:2]. The first anti-reflection coatings were made by Fraunhaufer in 1817 [24:2]. The modern era of thin film manufacturing began in the 1930's, with the invention of reliable vapor deposition techniques [24:4].

Modern design techniques can be broken into three broad categories; analytical methods, graphical methods, and digital designs. Analytical techniques use a few simple elements as building blocks to design more complex filters. The building blocks used, such as quarter wave stacks, are selected because the relationship between the few variables of the basic element (i.e., index, thickness, number of layers) and the reflectivity (i.e., total reflectivity, pass band width, stop band location) are well known. A specific design is then created by combining a series of these building blocks until the desired profile is obtained [24:164-172]. This technique is effective for fairly simple designs, such as a bandpass filter, but does not work as well for more complex designs. Many of

the existing thin film design techniques rely on the intuition of the designer, based on his experience with basic elements of the design. One example of this is the design of bandpass filters from quarter wave stacks [24:164-172]. Another example is the design of rugate filters, where the basic building blocks are sinusoids [22]. It has been noted that one of the standard rugate filter profiles for a bandpass filter bears a remarkable resemblance to a Morlet wavelet [37].

Graphical techniques are closely related to the analytical methods described above. They consist of various methods to simplify the calculations necessary in the analytical method by graphing the relationships between the variables. Examples of these techniques include the Smith chart, reflection circles, and admittance loci [24:54-66].

The techniques of primary interest here are the digital design methods. Digital design methods can be further subdivided into two categories; refinement and synthesis techniques. Refinement techniques are characterized by the fact that they require an initial starting design. Synthesis techniques, on the other hand, create their design without an initial guess [11: 2876]. There are over a dozen published refinement methods, and about half a dozen synthesis methods. There are undoubtedly many more unpublished design methods that are proprietary to the thin film manufacturing companies. Most of the methods make use of a merit function, which is a measure of how close the existing design is to the desired result. Some of the methods are capable of locating a global minimum, which is the true optimal solution to the problem. Other methods only solve for local minima based on the initial guess. Brief descriptions of a few of the published design methods are given below.

Refinement Methods: [11:2878-2882]

1. *Adaptive Random Search*: This procedure takes a starting design and applies a random change to the parameters of the design. The results at each step are compared to the desired result through a merit function. Changes in the design that reduce the merit

function are kept, while others are discarded. The magnitude of the change in parameters is also changed at each step, making this procedure perform both a local and global search. This means that if the "best" design is very different from the starting design, this method may find it. Local search methods only make small changes to the initial design. This method stops its iteration when the merit function passes a preset threshold or the method stops converging

2. *Damped Least Squares*: This method uses the derivatives of the quantities of interest in the merit function with respect to the construction parameters to determine the changes to be made to the design. Since the design problem is highly non-linear, the size of the change in any one step is limited by a damped least squares algorithm. This method performs a local search from the starting design.

3. *Golden Section Method*: The golden section method takes and initial design and varies each of the construction parameters in sequence. The optimal value for each parameter is found, (within the preset limits for each) before continuing on to change the next parameter. This algorithm converges fairly rapidly, and can find solutions far from the initial design.

4. *Hookes and Jeeves Pattern Search*: This technique is composed of two steps; an exploration and a pattern search. The exploration is to change each parameter up and down by a small amount, and determine which changes improve the merit function. The pattern search extrapolates in the direction of improvement from the exploration step. The step sizes are changed, and then the process is repeated. The process stops when the change in the merit function is smaller than a preset limit.

5. *Simulated Annealing*: The physical process of annealing is heating a sample and allowing the molecules to find a minimum energy state during cooling. This idea is used mathematically in this design approach. This method can deal with arbitrarily large numbers of design parameters, and can find globally optimal solutions.

Synthesis Methods: [23:3790-3791]

1. *Comprehensive Search Method*: This method is limited to films with only a few layers (<6). The global optimal solution is found by testing every possible combination of index and thickness allowed. Clearly, only a limited number of indices and thickness can be used. The advantage of this method is that it finds the truly optimal solution given the constraints on index and thickness. The disadvantage is it takes so much computer time to check every possible combination that only very simple systems can be designed this way.

2. *Gradual-Evolution Method*: This method is a modification of the comprehensive search described above. Instead of designing the whole film in one step, a smaller design, say three layers, is done by comprehensive search. The result is then added as part of the substrate, and an additional few layers are designed on this "new" substrate, again by comprehensive search. Since only a few layers are changed at a time, this method is much faster than the full comprehensive search method. However, there is no longer a guarantee that the global optimal solution will be found.

3. *Minus Filter Method*: A minus filter is a design that transmits all incident radiation except that in a narrow spectral band. Thelen has shown that all dielectric minus filters can be made using quarter wave stacks with several indices of refraction [39:365-369]. This design technique breaks the desired reflectivity profile into several minus filters, and then places the individual designs in series.

4. *Flip-Flop Method*: This unique approach uses many thin layers and only two indices of refraction. The algorithm is to pass through the film in sequences, flipping each layer between high and low index and calculating the resultant merit function. After several passes through the film, this method converges on a solution. The advantage to this technique is it uses only two materials, which is preferred by manufacturers, and it

requires no initial design. The method sometimes results in alternating very thin layers, but this can usually be corrected manually later.

2.2. Gradient Index Film Design and Fabrication

The current gradient index film design techniques focus on the inverse Fourier transform method and the rugate film design method. The inverse Fourier transform method can be classified as a synthesis technique, since no initial design is required. The rugate design method is an analytical design technique.

1. *Inverse Fourier Transform Method*: The use of Fourier transforms in the synthesis of thin films was first proposed by Delano in 1967 [9]. Sossi is generally credited with the development of the inverse Fourier transform method for thin film design [34,35,36]. Sossi's papers introduce a Fourier transform relationship between the logarithmic derivative of index of refraction and the reflectance of the film. This design technique was further developed by Dobrowolski and Lowe in 1978 [10]. Many authors have applied this method to various design problems, and offered modifications to the theory to better suit their needs. These include design of wideband anti-reflectance coatings [44] and high reflectance filters [16,43]. Other work has focused on modifying the theory of the method to improve the quality of the results or ease of computation [4,6,14,42]. This method, along with a modification of my design, will be described in detail in Section 3.2 below.

2. *Rugate Design Method*: A rugate filter is composed of a number of basic elements combined to achieve the desired optical properties of the film. The fundamental rugate design element is a sine wave refractive index profile on a substrate [22]. The basic rugate film has a notch reflectance profile, with the location, width, and height of the notch controlled by five parameters. More complicated rugate films are created by

9

adding several basic rugates in parallel or in series, and including apodization functions or matching layers to suppress sidebands in the reflectance.

The fabrication of films designed using these techniques is done in one of two ways. The first way is to use the concept of the Herpin equivalent index to convert a gradient index design into a two index material system, and make an equivalent film out of discrete layers [3,15,24:191-200]. This method allows the use of gradient index design methods to design conventional alternating layer based films, but does not offer the advantages an actual gradient index film would. The other fabrication possibility is to deposit the gradient index profile as designed. This is currently on the forefront of existing technology, and is an active area of research. There are several material systems that can be used to produce gradient index profiles. Table 2-1 lists some of the materials available, along with the index range and literature reference for additional information.

Table 2-1: Gradient Index Film Materials

Material System	Index Range	Reference
$Ge_{1-x}S_x$	3.5 - 4.0	[22:97]
$Al_{1-x}Ga_xAs$	2.9 - 3.6	[33:588]
$ZnS_{1-x}Se_x$	2.2 - 2.5	[22:97]
SiO_xN_y	1.5 - 2.0	[22:97, 29:179]
$SiO_2 - Ta_2O_5$	1.65 - 2.0	[1:141-145]
$MgF_2 - TiO_2$	1.38 - 2.3	[41:189-196]
$CeF_3 - ZnS$	1.6 - 2.2	[20:61]
$MgF_2 - ZnSe$	1.38 - 2.5	[20:197-204]

2.3. Optical Properties of Thin Films

This section presents the theory and methods for determining the reflectance properties of a film. The first section presents the derivation of the reflectance from Maxwell's equations, and the second section outlines the implementation of this formulation for a single thin film. The third section extends this implementation to gradient index films.

2.3.1. Derivation of Reflectance

This section presents the formal derivation for determining the reflectance of a film starting from Maxwell's equations. Figure 2.1 shows the coordinates used in the derivation. The surface of the film is the x-y plane, and the physical thickness of the film is along z. The origin is at the film-substrate interface, and the thickness of the film is denoted by L.
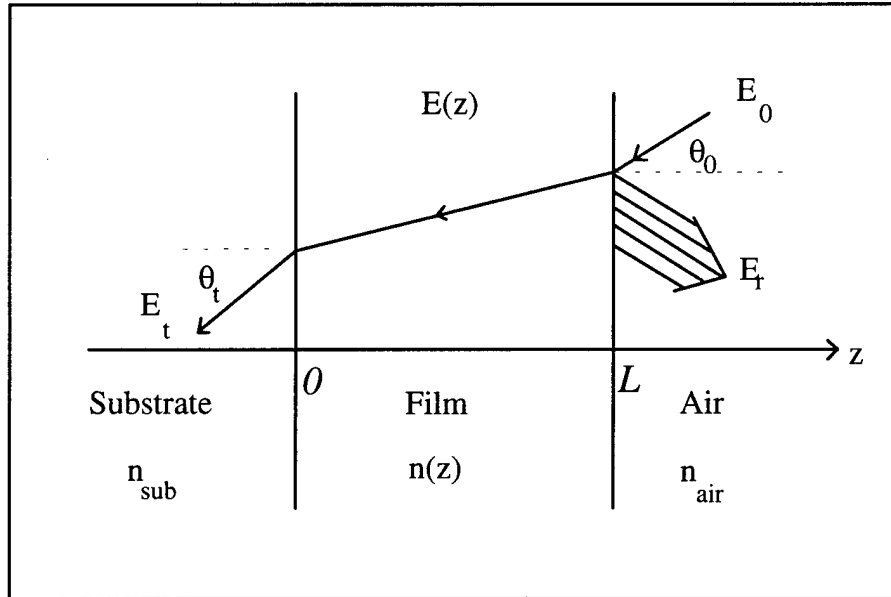
Figure 2.1: Geometry of Thin Film

Assuming a time dependence of exp[$i\omega t$], Maxwell's equations in SI units are

$$curl\ \vec{H} = i\omega\varepsilon E$$
$$curl\ \vec{E} = -i\omega\mu\vec{H}$$
$$div\ \vec{E} = 0 \tag{2.1}$$
$$div\ \vec{H} = 0$$

where E is the electric field, H is the magnetic field, ε is the permittivity, and μ is the permeability.

Two polarizations must be considered to determine the electric field; the transverse electric (TE) case, which has its electric field aligned perpendicular to the plane of incidence, and the transverse magnetic (TM) case, which has its electric field aligned parallel to the plane of incidence. By the symmetry of Maxwell's equations, the TM case can be deduced from the TE case with a substitution of E for H and ε for μ. The plane of incidence is defined by the direction of propagation of the incident light and the normal to the surface. Define the x direction as the direction in which the electric field is aligned, so the fields can be written as

$$\vec{E} = \begin{bmatrix} E_x(z)\exp[i\omega t - ikSy] \\ 0 \\ 0 \end{bmatrix} \qquad \vec{H} = \begin{bmatrix} 0 \\ H_y(z)\exp[i\omega t - ikSy] \\ H_z(z)\exp[i\omega t - ikSy] \end{bmatrix} \tag{2.2}$$

Here S is an invariant, $S = n_0 \sin(\theta_0)$, where n_0 is the index of refraction of the incident medium, and θ_0 is the angle of incidence. Also, ω is the frequency of the incident light, k is the wave number of the light, $k=2\pi/\lambda$, and λ is the wavelength in vacuum. Substituting these expressions for E and H into Maxwell's equations yields

$$Z_0 \frac{\partial H_y(z)}{\partial z} = -ik\left(\varepsilon_r - \frac{S^2}{\mu_r}\right)E_x(z)$$

$$\frac{\partial E_x(z)}{\partial z} = -ik\mu_r Z_0 H_y(z) \tag{2.3}$$

$$H_z(z) = -\left(\frac{S}{Z_0 \mu_r}\right)E_x(z)$$

where Z_0 is the impedance of free space, $Z_0 = \sqrt{\mu_0/\varepsilon_0}$, and

ε_0, μ_0 are the permitivity and permeability of free space,

ε_r, μ_r are the relative permittivity and permeability of the medium.

Let $N^2(z) = \varepsilon_r(z)\mu_r - S^2$, and consider only dielectric materials with $\mu_r = 1$. $N(z)$ is always a real valued function since $S < 1$ for all cases considered here and $\varepsilon_r > 1$ for all dielectric materials. For normal incidence, $S=0$, and N reduces to the usual definition of index of refraction. Also define two new complex valued variables u and v by

$$u(z) = \frac{E_x(z)}{E_t} \qquad v(z) = -\mu_r Z_0 \frac{H_y(z)}{E_t} \tag{2.4}$$

The normalization factor E_t is the complex amplitude of the transmitted wave. Using the new normalized field variables, Maxwell's equations reduce to the following pair of coupled first order ordinary differential equations

$$u'(z) = ikv(z) \quad \text{and} \quad v'(z) = ik N^2(z)u(z) \tag{2.5}$$

with boundary conditions

$$u(0) = 1 \qquad v(0) = -n_{sub}\cos(\theta_t) \tag{2.6}$$

In this formulation, the amplitude reflectance and transmittance of the film, $r(k)$ and $t(k)$ respectively, are found from the Fresnel equations: [40:4274]

$$r(k) = \frac{n_{air} u(L,k)\cos(\theta_0) + v(L,k)\cos(\theta_t)}{n_{air} u(L,k)\cos(\theta_0) - v(L,k)\cos(\theta_t)}$$

(2.7)

$$t(k) = \frac{2n_{air} u(L,k)\cos(\theta_0)}{n_{air} u(L,k)\cos(\theta_0) - v(L,k)\cos(\theta_t)}$$

Here the field variable dependence on the wave number k has been called out explicitly.

2.3.2. Multilayer Thin Films

The optical properties of a multilayer thin film are based upon the Fresnel reflection at interfaces and the interference between multiple reflections. The reflectance of such a film can be determined by using Maxwell's equations and applying the boundary conditions at both interfaces. MacLeod has shown the fields at the first interface are related to the fields at the second interface by [24:35]

$$\begin{bmatrix} E_{01} \\ H_{01} \end{bmatrix} = \begin{bmatrix} \cos(\delta) & i\sin(\delta)/\eta_1 \\ i\eta_1 \sin(\delta) & \cos(\delta) \end{bmatrix} \begin{bmatrix} E_{12} \\ H_{12} \end{bmatrix}$$

(2.8)

where E_{01}, H_{01} represent the field at the interface between the incident media and the film, and E_{12}, H_{12} represent the field at the interface between film and the substrate. The optical admittance of the film material η_1 is given by

14

$$\eta_1 = H_1/E_1 = \frac{n_1 Y_0}{\cos(\theta)} \qquad TM$$

$$= n_1 Y_0 \cos(\theta) \qquad TE$$

(2.9)

with $Y_0 = 1/377$ Siemens. The phase shift δ is given by

$$\delta = 2\pi n_1 d \cos(\theta)/\lambda$$

(2.10)

where n_1 is the index of refraction of the film, d is the thickness of the film, θ is the angle between the direction of propagation and the normal to the surface of the film, and λ is the wavelength in vacuum. Note that the phase shift δ can be complex, and the angle θ_1, found using Snell's Law, may also be complex. The complex amplitude reflectivity, ρ, and the real intensity reflectance, R, can by found from Equation 2.8 by defining an input optical admittance for the assembly as

$$Y \equiv H_{01}/E_{01}$$

(2.11)

Then for an incident medium with an optical admittance of η_0

$$\rho = \frac{\eta_0 - Y}{\eta_0 + Y}$$

$$R = \left(\frac{\eta_0 - Y}{\eta_0 + Y}\right)\left(\frac{\eta_0 - Y}{\eta_0 + Y}\right)^*$$

(2.12)

where the asterisk indicates complex conjugation.

Since the reflectance depends only on the optical admittance of the film, Y, rearrange Equation 2.8 to find Y:

15

$$\begin{bmatrix} B \\ C \end{bmatrix} = \begin{bmatrix} \cos(\delta) & i\sin(\delta)/\eta_1 \\ i\eta_1\sin(\delta) & \cos(\delta) \end{bmatrix}\begin{bmatrix} 1 \\ \eta_2 \end{bmatrix}$$

$$Y = C/B$$

(2.13)

The matrix above is the characteristic matrix for the film, and depends only on the film materials and the angle of incidence of the impinging light. This matrix technique can be directly extended to an assembly of many thin films. So, given q thin film layers, each with specified index and thickness, the characteristic matrix of the assembly is given by

$$\begin{bmatrix} B \\ C \end{bmatrix} = \prod_{r=1}^{q} \begin{bmatrix} \cos(\delta_r) & i\sin(\delta_r)/\eta_r \\ i\eta_r\sin(\delta_r) & \cos(\delta_r) \end{bmatrix}\begin{bmatrix} 1 \\ \eta_s \end{bmatrix}$$

(2.14)

where for each layer r the phase shift δ_r is

$$\delta_r = 2\pi\, n_r d_r \cos(\theta_r)/\lambda$$

(2.15)

and n_r is the index of refraction of the layer, d_r is the thickness of the layer, and the angle θ_r is found using Snell's Law. These equations form the basis for the design of thin film dielectric mirrors.

In addition to determining the reflectivity of a multilayer thin film, the characteristic matrix above can also be used to determine the phase of the reflected field. Writing the optical admittance of the film as $Y=a + bi$, the phase, φ, of the field upon reflection is given by [24:37]

16

$$\tan(\varphi) = \frac{-2b\eta_0}{\left(\eta_0^2 - a^2 - b^2\right)} \tag{2.16}$$

The simplest example of a mirror using these multilayer films is a quarter wave stack. This is a series of layers each with an optical thickness of one quarter of the wavelength of the incident light. If the index and thickness of each layer are chosen so the optical path length $nd = \lambda/4$, and the light is at normal incidence, the phase shift δ for each layer is $\pi/2$. The characteristic matrix then reduces to

$$\begin{bmatrix} B \\ C \end{bmatrix} = \prod_{r=1}^{q} \begin{bmatrix} 0 & i/\eta_r \\ i\eta_r & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \eta_s \end{bmatrix} \tag{2.17}$$

This characteristic matrix is very easy to calculate. An example of such a quarter wave stack is 25 alternating layers of GaAs and AlAs on a GaAs substrate, which have indices of refraction of 3.6 and 3.2 respectively [33:588]. If the design wavelength is 1.0 micron, this mirror is 93.6% reflective at this center wavelength and has a reflectivity versus wavelength profile as shown in Figure 2.2.
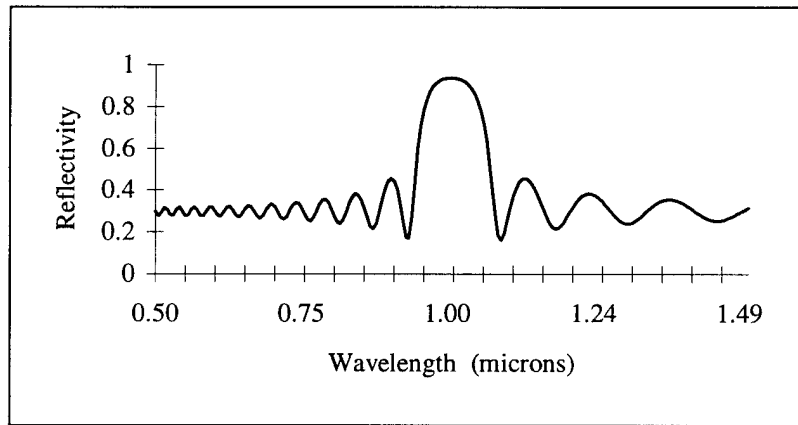
Figure 2.2: Reflectivity vs. wavelength for 25 layer quarter wave stack of GaAs/AlAs on a GaAs substrate.

17

2.3.3. Gradient Index Films

The analysis above assumes the mirror is formed by a discrete number of layers with well-defined thickness and index of refraction. How does one treat a structure where the index varies continuously? One approach, derived by Bovard, is to start from the beginning with Maxwell's equations and develop an expression for the characteristic matrix of the film based on the known index of refraction as a function of position in the film [4:1999-2001]. Unfortunately, the resultant expression contains an infinite series of integrals, which makes evaluation impracticable. Another alternative, and the one implemented here (see **REFLECT.M** in Appendix D), is to approximate the continuous film as a discrete film with many small layers. As a general rule, discrete layers on the order of 5 nm thick give a good approximation for the reflectance of a gradient index film for visible light [6:5429]. Since the formalism presented above includes the complex portion of the fields involved, the interference effects are already incorporated in the calculation. This is of critical importance to the validity of this approach, since high reflectivity depends on destructive interference of the transmitted fields.

The reflectance is calculated by a "C" language version of **REFLECT.M**. The conversion of this program from the MATLAB™ language to "C" increases the speed of each reflectance calculation by at least a factor of ten. The "C" language program, **REFLECT.C**, is also included in Appendix D. The inputs to **REFLECT.C** are the sampled values of the index, the substrate index, the vector of wavenumbers k at which to determine the reflectance, and the array of sample positions in the film x. The output is an array of reflectances, one for each of the input wavenumbers k. The units for x and k must be consistent, i.e. microns and inverse microns or nanometers and inverse nanometers. The programs **REFLECT.M** and **REFLECT.C** are used throughout the dissertation to determine the reflectance characteristics of the gradient index thin film designs.

3.   Thin Film Design Using Inverse Fourier Transforms

Synthesis of a thin film is the process of taking film performance requirements and creating a film that will meet those requirements. Several techniques exist in the literature for performing this task, including the comprehensive search method, the gradual evolution method, the flip-flop method, the minus filter method, and the inverse Fourier transform method [12:101-106]. Of these methods, the inverse Fourier transform method has an advantage in speed of computation. This is due to the use of matrix fast Fourier transform techniques [23:3799]. This chapter includes a description of the current inverse Fourier transform methods, and then presents a derivation of an enhancement, called the Stored Waveform Inverse Fourier Transform, or SWIFT, technique. The SWIFT technique provides a tool for the thin film designer to control the index of refraction range of the film designed with the inverse Fourier transform method. The SWIFT technique determines an optimal phase to use in the inverse Fourier transform, which has the effect of spreading the index contrast more evenly over the total thickness of the film. The SWIFT technique also generates film designs that better achieve the desired film reflectance profile, as compared to films designed without the SWIFT optimal phase. Several examples are given to illustrate the various properties of this new technique.

3.1.   Inverse Fourier Transform Method

One technique for designing continuous gradient index films is based on a Fourier transform relationship proposed by Sossi and Kard [34,35,36] and fully developed by Dobrowolski and Lowe [10] and Bovard [4]. The inverse Fourier transform relation is derived from Maxwell's equations by making several approximations. The film is

19

assumed to be sandwiched between two identical media, and additional assumptions of no dispersion, no absorption, and normal incidence are made. An excellent presentation of this derivation is given (in English) in Bovard [4]. First the basic inverse Fourier transform relation will be stated, and then the end results of the derivation will be presented, in order to illustrate the approximations made to arrive at this inverse Fourier transform relation.

Given a desired reflectance profile, $R(\lambda)$, the index profile, $n(x)$, needed to generate this reflectance is given approximately by a Fourier transform relation. Written in terms of the wavenumber $k=2\pi/\lambda$, and the double optical path length x, the relation is

$$\ln\left[n(x)\right] = \frac{i}{\pi}\int_{-\infty}^{\infty}\frac{Q(k)}{k}\exp\left[i\,\Phi(k)-ikx\right]dk \tag{3.1}$$

where $Q(k)$ is one of several functions of $R(\lambda)$, and $\Phi(k)$ is the complex phase of the index profile. The relation between the double optical path length in the medium, x, and the physical thickness, z, is

$$x = 2\int_{0}^{z}N(u)du \tag{3.2}$$

where $N(z)$ is the index of refraction as a function of physical thickness. Note this is a different functional dependence than $n(x)$, which is the index as a function of double optical path length. The solution of Equation 3.1 is the index of refraction for the film as a function of double optical thickness. The conversion of this index versus double optical path length to an index profile in terms of physical thickness is accomplished by inverting Equation 3.2 to find the physical thickness, z, for a given optical thickness, x:

$$z = \frac{1}{2} \int_0^x \frac{du}{n(u)} \qquad (3.3)$$

The function $N(z)$ is then constructed by noting that $N(z)=n(x)$ when x and z are related by Equation 3.3 above.

Bovard derives a rather complicated series of expressions describing the relationship between the complex amplitude reflection and transmission coefficients ρ and τ as a function of wavelength to the index of refraction profile as a function of position [4]. The amplitude coefficients can be written

$$\rho = \frac{B_1 + B_3 + B_5 + \dots}{1 + B_2 + B_4 + \dots}$$

$$\tau = \frac{1}{1 + B_2 + B_4 + \dots} \qquad (3.4)$$

The functions B_n are a family of functions:

$$B_{2n}(\sigma) = [C_{2n}(\sigma) - iS_{2n}(\sigma)]\exp(-i\pi\sigma_0 x)$$

$$B_{2n+1}(\sigma) = C_{2n+1}(\sigma) - iS_{2n+1}(\sigma) \qquad (3.5)$$

where the wavenumber $\sigma = 1/\lambda$, and x is the double optical phase thickness of the film. The functions C_n and S_n are

$$C_1(\beta) = \int_0^\beta r(\beta_1)\cos(2\beta_1)\,d\beta_1$$

$$S_1(\beta) = \int_0^\beta r(\beta_1)\sin(2\beta_1)\,d\beta_1$$

$$C_2(\beta) = \int_0^\beta \int_0^{\beta_1} r(\beta_1)r(\beta_2)\cos(2(\beta_2 - \beta_1))\,d\beta_2 d\beta_1 \qquad (3.6)$$

$$S_2(\beta) = \int_0^\beta \int_0^{\beta_1} r(\beta_1)r(\beta_2)\sin(2(\beta_2 - \beta_1))\,d\beta_2 d\beta_1$$

or in general

$$C_m(\beta) = \int_0^\beta \int_0^{\beta_1} \dots \int_0^{\beta_{m-1}} r(\beta_1)r(\beta_2)\dots r(\beta_m)\cos(2(\beta_m - \beta_{m-1} + \dots \beta_1))\,d\beta_m \dots d\beta_2 d\beta_1$$

$$\qquad (3.7)$$

$$S_m(\beta) = \int_0^\beta \int_0^{\beta_1} \dots \int_0^{\beta_{m-1}} r(\beta_1)r(\beta_2)\dots r(\beta_m)\sin(2(\beta_m - \beta_{m-1} + \dots \beta_1))\,d\beta_m \dots d\beta_2 d\beta_1$$

The variable β is an optical phase thickness, which is related to the double optical thickness x by

$$\beta = \frac{x\pi}{\lambda} + \frac{\beta_0}{2} \qquad (3.8)$$

where β_0 is the total optical phase thickness. Finally, the function $r(\beta)$ is the logarithmic derivative of the index of refraction:

$$r(\beta) = \frac{n'(\beta)}{2n(\beta)} \qquad (3.9)$$

Equations 3.4 through 3.9 form the relationship between the amplitude reflectance coefficients as a function of wavelength and the index of refraction as a function of position. A Fourier transform relationship between $\rho(k)$ and $n(x)$ is derived by neglecting all terms in Equation 3.4 with $n \geq 2$ and substituting for β using Equation 3.9:

$$\rho(k) \cong \int_{-\infty}^{\infty} \frac{n'(x)}{2n(x)} \exp[ikx]\, dx \tag{3.10}$$

The limits of integration have been extended to infinity without penalty because the index of refraction is assumed to be constant outside the film, so the derivative of the index is zero. The remaining step to achieve Sossi's form for the inverse Fourier transform relation of Equation 3.1 is to replace the complex amplitude reflectance $\rho(k)$ with another function of the reflectance or transmittance, denoted by $Q(k)\exp[i\Phi(k)]$. This complex valued function is usually expressed in polar form in the literature. The Q function is used in place of the amplitude reflectance in an effort to improve the approximation by including some of the effects of the terms neglected in the derivation. The most obvious choice for the Q function is $Q(k) = R(k)^{1/2}$, which is just the magnitude of the complex amplitude reflectance, $\rho(k)$. In this case the "correct" phase function $\Phi(k)$ would be the phase of the complex amplitude reflectance.

Several forms of $Q(k)$ have been used by various authors, including [4:2002, 5:26, 10:3043,42:3673]

$$
\begin{aligned}
Q_1 &= \sqrt{\tfrac{1}{2}\left(T^{-1} - T\right)} \\
Q_2 &= \sqrt{1 - T} \\
Q_3 &= \sqrt{T^{-1} - 1} \\
Q_4 &= w\sqrt{1-T} + (1-w)\sqrt{T^{-1}-1}, \qquad 0 < w < 1 \\
Q_5 &= \ln\left(\gamma + \sqrt{\gamma^2 - 1}\right), \qquad \gamma = 1 + \tfrac{1}{4}\left(T^{-1} - T\right)
\end{aligned}
\tag{3.11}
$$

Here $T(k) = 1 - R(k)$ is the transmission of the non-absorbing film as a function of wave number k. The Q functions in Equation 3.11 are not an exhaustive list of possible Q functions, nor are they exact for any particular problem. The choice of the form for $Q(k)$ depends on the characteristics of the desired reflectance. The first three Q functions above are good for low rejection filters, while the last two are better for high rejection filters. For complicated reflectance profiles, an iterative scheme of linear combinations of these forms yields good results [43:2866].

Sossi proposed to overcome this limitation in the Q function of the Fourier transform relation by using a series of successive approximations [42:3673]. Applying Equation 3.1 to a desired reflectance profile generates an index of refraction profile. The reflectance of this index profile can be determined using the matrix techniques described previously in section 2.3. The resultant reflectance, R_1, will not exactly match the desired reflectance, R_d. This difference can be used as the basis for the successive approximations. By adding the difference between the Q function of the desired reflectance, $Q(R_d)$, and the Q function generated by using R_1, a new, hopefully better Q function can be formed. Mathematically, this algorithm is

$$Q_{i+1} = Q_i + \Delta Q$$

$$\Delta Q = Q(R_d) - Q(R_i) \qquad (3.12)$$

$$Q_1 = Q(R_d)$$

Sossi only illustrated this method on a few simple examples, so the validity of this approach has not been rigorously established. Verly has recently expanded on this approach by incorporating iterations on the phase $\Phi(k)$ as well as the Q function

24

[42:3674]. He claims this approach reduced the thickness of the resultant film significantly.

The phase function $\Phi(k)$ has not received as much study as the Q function has. As was mentioned previously, for the case where the Q function is the square root of the reflectivity, the phase function $\Phi(k)$ is the same as the desired phase of the reflected wave. However, for other forms of $Q(k)$, the interpretation of the phase function is not so straightforward. The phase function was originally set to zero in Sossi's work [35:6]. Later, Dobrowolski and Lowe proposed using various analytic forms for $\Phi(k)$ to improve the spectral fit between the design and film performance [10]. Another expression for the optimal phase is derived in detail in Section 3.2 below.

Since both the $Q(k)$ and $\Phi(k)$ are to some extent arbitrary functions, the Fourier transform relation does not give a single index profile to generate the desired reflectance. This is one of the major drawbacks to this method. The synthesis of a film for given desired reflectance depends on the form chosen for $Q(k)$ and $\Phi(k)$. The difficulty lies in choosing the functional forms that will give the "best" index profile. This is particularly true if the requirements of the film are not rigidly specified over the entire wavelength region of interest. For example, consider the design of a mirror for an optically pumped laser. The requirements are high reflectivity at the lasing wavelength, and high transmission at the pump wavelength. The rest of the reflectance profile does not matter. However, some form for this profile must be chosen in advance to use the inverse Fourier transform design method. The other drawback to this method is it does not allow for the application of additional constraints to the design, such as available index of refraction. The Stored Waveform Inverse Fourier Transform (SWIFT) technique described in the next section addresses the application of constraints on index range to this method.

3.2. The SWIFT Technique

This section describes the Stored Waveform Inverse Fourier Transform technique for calculating a phase function to use in the inverse Fourier transform which will account for constraints on the index of refraction range. First, the original use of this technique in mass spectrometry is discussed and then the theory is derived as it applies to gradient index thin film design.

The reflectivity function $Q(k)$ and the phase function $\Phi(k)$ in Equation 3.1 provide two degrees of freedom in designing an index profile to generate a desired reflectance. Several investigators have explored various functional forms for $Q(k)$, as listed above, and $\Phi(k)$ [10,16,42,43]. In what follows, I present a technique for computing a phase function, $\Phi(k)$, that allows imposition of constraints on filter thickness and maximum index contrast [14]. The method, which derives from an analogous result in pulsed ion cyclotron resonance mass spectrometry [17], has the added benefit of producing the desired reflectance function with higher fidelity than alternative formulations. The new phase function is computed by a method called Stored Waveform Inverse Fourier Transform (SWIFT). The mathematical basis for SWIFT is described by Guan [17,18], and is outlined below.

The connection between pulsed ion cyclotron resonance mass spectrometry work and design of thin films is that both rely on an inverse Fourier transform relationship. In the case of pulsed ion cyclotron resonance mass spectrometry, the inverse Fourier transform relates the voltage profile applied in time to the frequency distribution of the electric fields formed in the device. The objective is to create a field with very sharp turn on and turn off in frequency in order to selectively trap ion species. If the phase of the desired field is arbitrarily set to zero, the voltage profile in time required to generate a notch in frequency can easily exceed the equipment's abilities to switch large voltages. However, if the phase is properly selected for the desired frequency profile, the peak to

26

peak voltage oscillations required can be restricted to within an acceptable range. This method for producing notch field profiles was successfully demonstrated at AFIT by K. Reihl [31].

The goal of the SWIFT technique is to reduce the index range needed by distributing the "power" of the reflectance design region uniformly over a portion of the thickness of the film, referred to as the "power spread region" of the film. The "power" in a function is found by integrating its squared-magnitude. The terminology "power" is borrowed from the signal processing community, where much of the applied Fourier theory was developed. The Fourier transform pair considered here is $\ln(n(x))$ and $\exp[i\Phi(k)]Q(k)/k$. The reflectance design region contains those wavenumbers, k, satisfying $k_0 \leq k \leq k_1$ over which the reflectance design is to be satisfied, not necessarily containing the reflectance for all wavenumbers. Furthermore, the "power spread region" is not necessarily the total thickness of the film. Rather it is the subset of the film occupying positions x satisfying $x_0 \leq x \leq x_1$; the difference $(x_1 - x_0)$ is a design parameter referred to as the "power spread thickness."

The goal of distributing the power uniformly across the thickness of the film is approached by considering the following: imagine the portion of the index, $n(x)$, in the power spread region to be composed of infinitesimals of width dx, located at positions x, each of which is required to have constant power spatial density, denoted by $1/c$. A second requirement, leading directly to the desired expression for $\Phi(k)$, is that the phase contributed to the reflectance by each slab follow a Fourier shift theorem. The results of these two requirements are obtained below and combined to achieve a phase in the reflectance design region to distribute the spatial power density over the power spread region of the index.

Each infinitesimal of power in the reflectance design region is required to be contained in an infinitesimal of the index, all of which (recall) have constant power

27

density, $1/c$. Therefore, the width, dx, of an infinitesimal slab of the index can be related to the power in an infinitesimal slab of reflectance by

$$dx = cG(k)dk \qquad (3.13)$$

where the one-sided power spectral density, $G(k)$, is given by [30:401]

$$G(k) \equiv 2\left|\frac{Q(k)}{k}\right|^2, \qquad k > 0 \qquad (3.14)$$

Now consider the phase shift associated with each infinitesimal index slab. If the phase shift contributed to the reflectance by each index slab of width dx were uniformly constant, the slabs would all be centered at the origin, and their index powers would add up to produce a large power concentrated around the origin. To create the desired uniform power spread in the power spread region, the phase at k is chosen to be that corresponding to shifts of the slabs to positions x according to a Fourier shift theorem, when the slabs are arranged so that if $x < x'$ then the slabs at x and x' contain, respectively, the power in the reflectance design region at wavenumbers k and k' satisfying $k < k'$. If these uniform power slabs in index had finite widths Δx located at position x, the phase of their individual reflectance contributions would be piecewise linearly shifted by kx, so that x is the slope of the phase shift. In the limit of infinitesimal slabs of width dx, this is expressed by imposing the condition

$$x = \frac{d\Phi(k)}{dk} \qquad (3.15)$$

28

To combine the results obtained above from the impositions of constant spatial power density and appropriate phase, the expression in Equation 3.13 can be integrated to sum the index widths, dx, to yield the position, x, of the index slab containing the power at wavenumber, k, according to

$$x = c \int_{k_0}^{k} G(\xi) d\xi + x_0 \qquad (3.16)$$

The constant, c, can be determined by requiring the total power in the power spread region of the index and reflectance design region to be equal. Integrating over the respective regions in index and reflectance yields

$$c = \frac{x_1 - x_0}{\int_{k_0}^{k_1} G(k) dk} \qquad (3.17)$$

Integrating Equation 3.15 and using Equations 3.16 and 3.17 gives the phase function:

$$\Phi(k) = \frac{x_1 - x_0}{\int_{k_0}^{k_1} G(\xi) d\xi} \int_{k_0}^{k} \int_{k_0}^{\xi} G(\eta) d\eta \, d\xi + x_0(k - k_0) + \Phi(k_0), \qquad k_0 \le k \le k_1 \quad (3.18)$$

The last, constant phase, term does not affect the index profile and is set to zero for all that follows.

The SWIFT technique provides a method for indirectly (iteratively) including constraints on available index of refraction in the inverse Fourier transform design technique. As the desired power spread thickness for the phase calculation is increased, the index range decreases (see examples below). However, the SWIFT technique does not address the difficulty of choosing an appropriate form for the Q function. Several

29

examples of applications of this technique to thin film designs are given in Section 3.3 below. The details of the numerical considerations to generate these examples are in Appendix A.

3.3. SWIFT Gradient Index Film Design

This section includes several design examples illustrating the various features of the SWIFT design technique. The first example design is a narrow band reflector, which illustrates the basic trade off between index range and film thickness. The second example is a broadband high reflector. The third and fourth examples are for an optically pumped neodymium : yttrium aluminum garnet (Nd:YAG) laser design, which will be used throughout the dissertation for comparison between design methods. The third example is an output coupling mirror, which has high reflectance at the lasing wavelengths and low reflectance at the pump wavelength. The fourth example is an anti-reflection coating for the gain medium of the laser.

3.3.1. Notch filter design

Consider a narrow-band reflectance filter with a reflectivity of 90% from 580 to 620 nm and 0% outside this band. In this example, the indices of the substrate and the incident medium are the same ($n_{sub} = n_{out} = 1.50$). The desired optical thickness of the film is 30 μm. The optical thickness of the film is a design choice, and is a trade off between the desire to minimize the film thickness and the desire to achieve the best performance possible. The numerical parameters for the sampling are found by requiring the index sample spacing, Δ, to be on the order of five nm, the number of non-zero samples, N_0, of the Q function to be about 25, and the total number of samples, N_s, to be a power of two. A detailed explanation of the relationship between these paramters, and an example of

how to use these values to choose appropriate design parameters is given in Appendix A. The parameters selected for this example, based on these values, are

$$N_s = 2^{14} = 16,384$$
$$d_{Total} = 100 \; \mu \, m$$
$$\Delta = 0.0061 \; \mu \, m$$
$$N_0 \Rightarrow 23 \text{ non-zero samples of } Q$$

$$(3.19)$$

This choice for the total optical thickness of the film should also minimize any aliasing effects, since it is over three times the desired film optical thickness.

The choice of a form for $Q(k)$ is critical to the fidelity with which the computed reflectance spectra match the design goals. For this example a good form of $Q(k)$ was found by trial and error to be an empirical combination of two forms derived by Bovard [5,6] :

$$Q(k) = 0.5[-\ln(T)]^{1/2} + 0.5\left(\frac{1}{\sqrt{T}} - \sqrt{T}\right)^{1/2}$$

$$(3.20)$$

The index profile for this narrow-band filter calculated with Equation 3.1, $\Phi(k)=0$, and the reflectance spectra computed with standard matrix methods are shown in Figure 3.1. Optimization of $\Phi(k)$ for a design "power spread" optical thickness of 10 μm yields the index and the reflectance profiles shown in Figure 3.2. The range of indices is reduced by 20%, while the reflectance spectrum is closer to the design goal with a smaller variation over the notch and smaller $dR/d\lambda$ near its center. Recomputing $\Phi(k)$ for a filter with a design "power spread" optical thickness of 15 μm further reduces the requirements on index contrast, as shown in Figure 3.3. In this case, the reflectance profile is not as good as the design using a 10 micron "power spread" optical thickness. This is due in part to the choice to limit the total film film optical thickness to 30 μm. This design

31

method reduces the index range required by spreading the "active" portion of the index profile over a wider range of optical thickness. When the film design is truncated at 30 μm, some of the "active" portion of the film is lost. If the total film optical thickness is increased, the performance in terms of meeting the desired reflectance profile also improves.
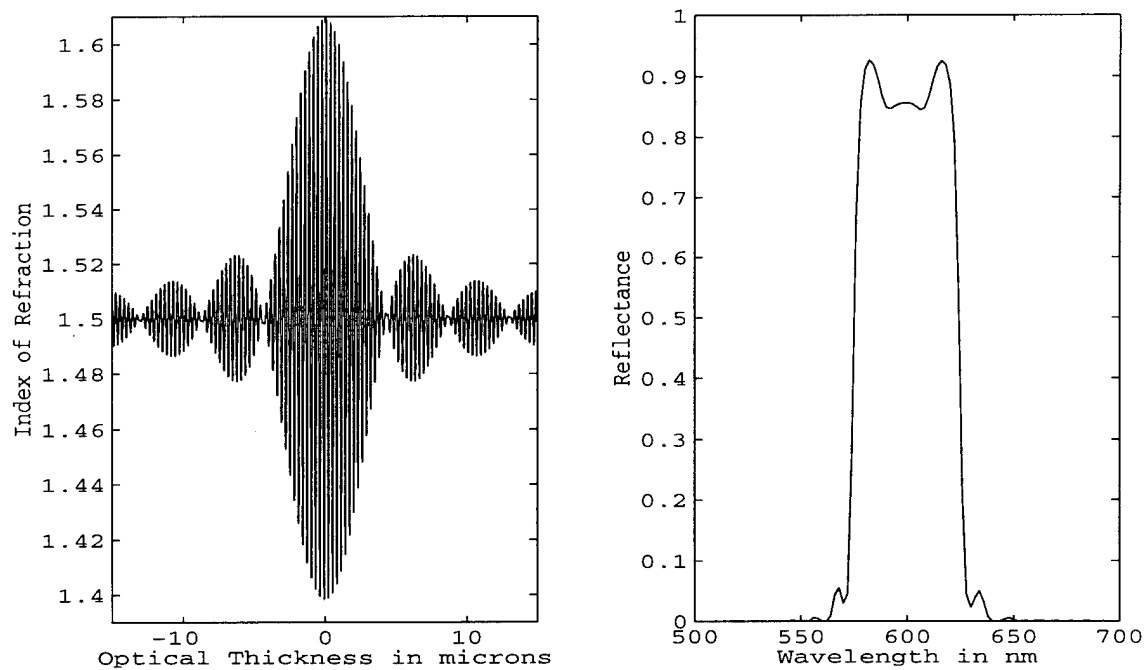
Figure 3.1: Index and reflectance for zero phase narrow notch filter.

32
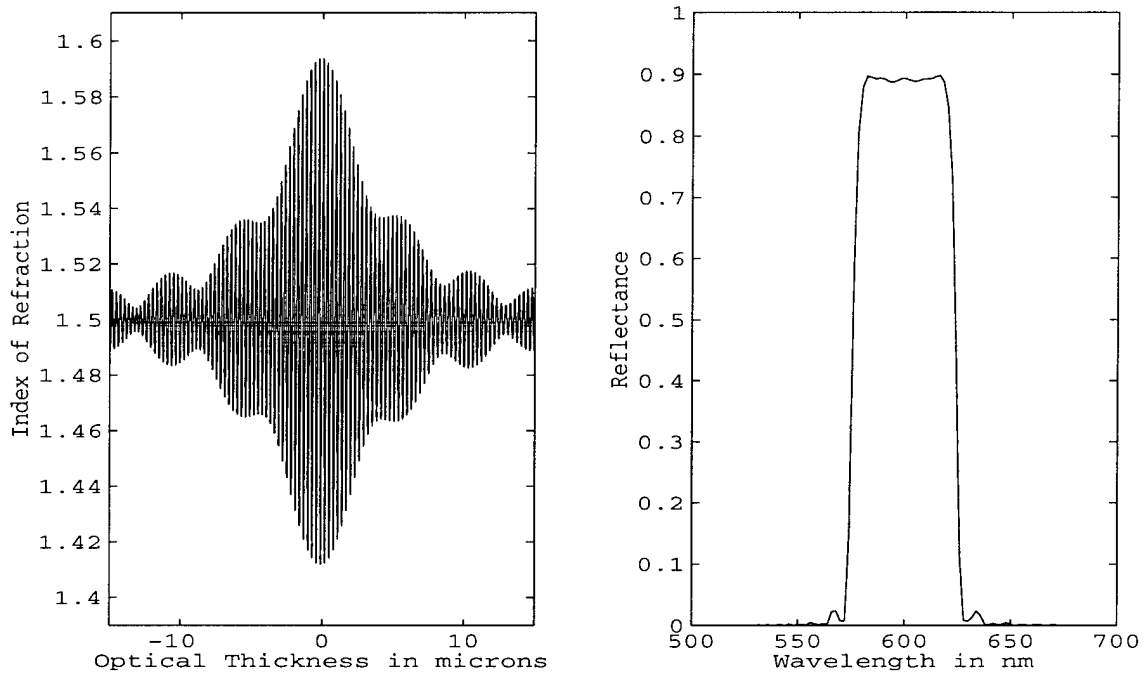
Figure 3.2: Index and reflectance for optimal narrow notch filter with 10 micron phase power spread thickness.
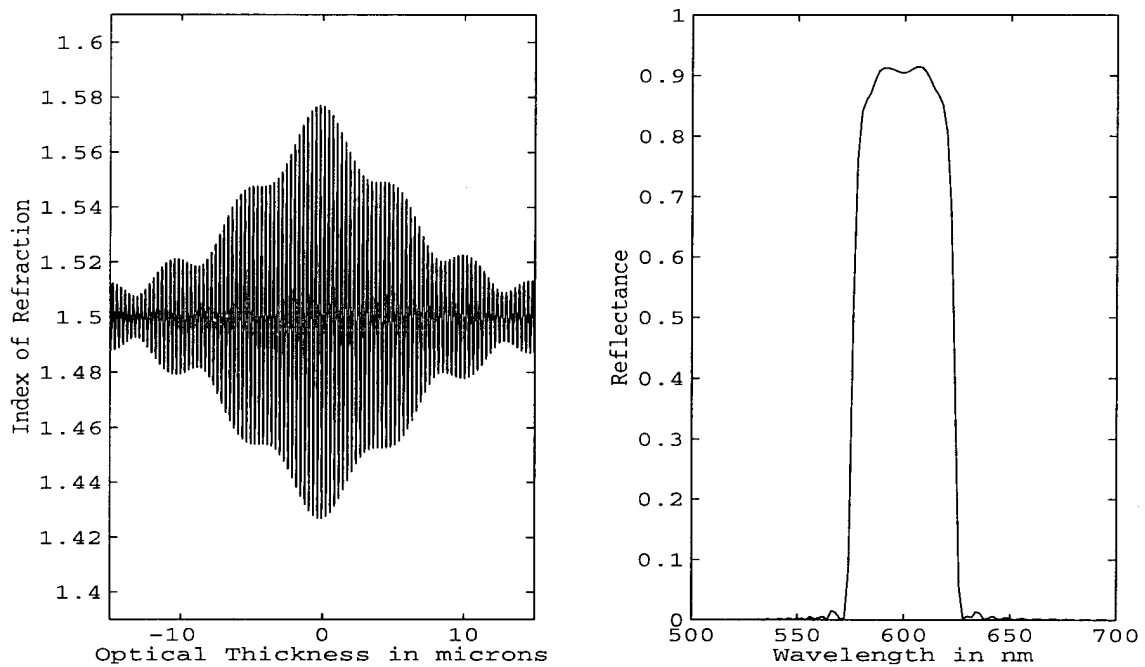
Figure 3.3: Index and reflectance for optimal narrow notch filter with 15 micron phase power spread thickness.

33

A portion of optimal phases used to generate these films are shown in Figure 3.4 and Figure 3.5. The total spatial frequency used in the representation of the Q function is 81.92 μm^{-1}. The phase functions are essentially two linear segments, with a change in slope at the desired notch in reflectivity. This has been emphasized in the two figures by adding a dotted extension to the zero intercept line. The phase functions for these designs are two line segments because the notch in the visible is a very small part of the k-space spanned in the Fourier transform design. The Q function is specified over a range of 81.92 μm^{-1}, but the non-zero portion only spans 0.111 μm^{-1}. As k increases, the integral in the phase function definition (Equation 3.18) is zero before the notch, a continuum of values inside the notch, and a constant after the notch. The two optimal phase functions used are plotted together in Figure 3.6 to illustrate the difference between the two designs.
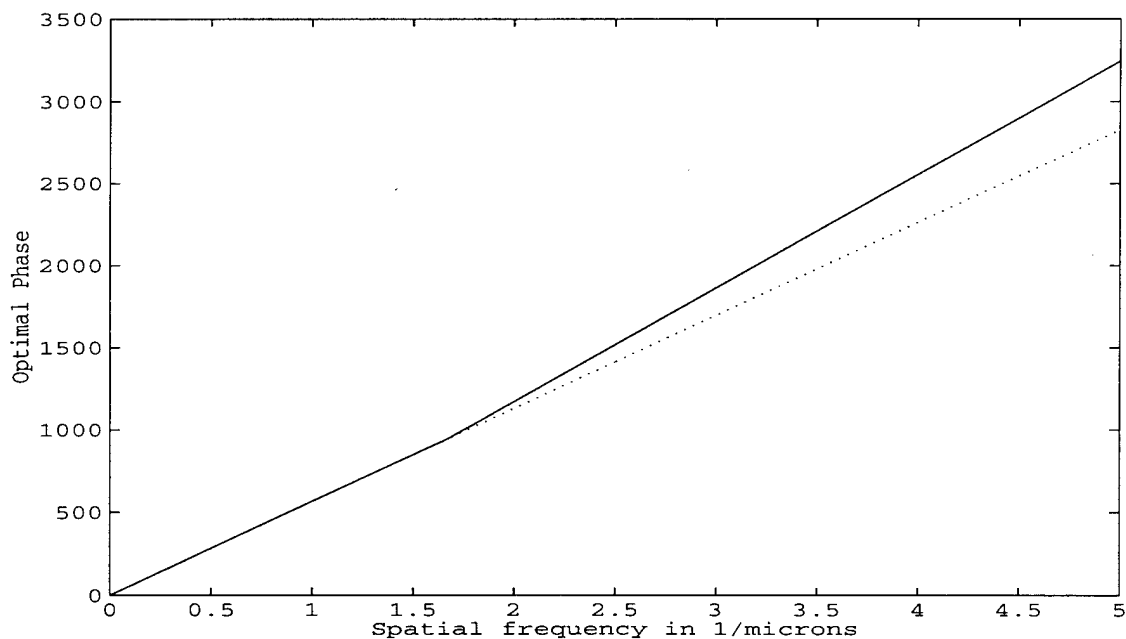
Figure 3.4: Optimal phase for narrow notch filter with 10 micron phase power spread thickness.
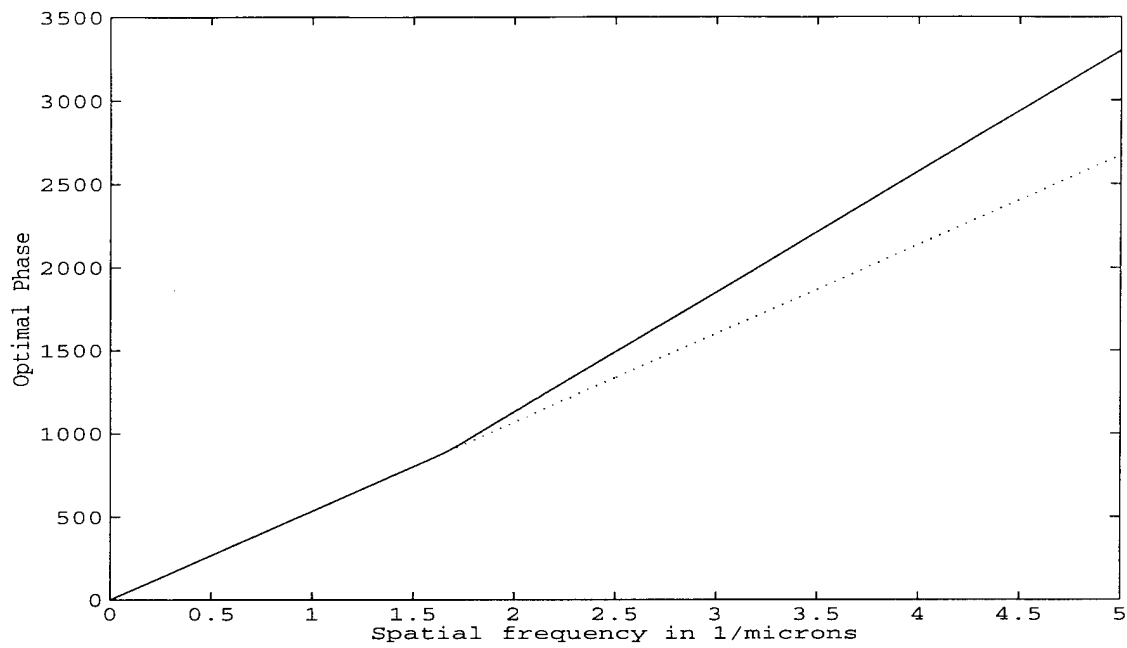
Figure 3.5: Optimal phase for narrow notch filter with 15 micron phase power spread thickness.
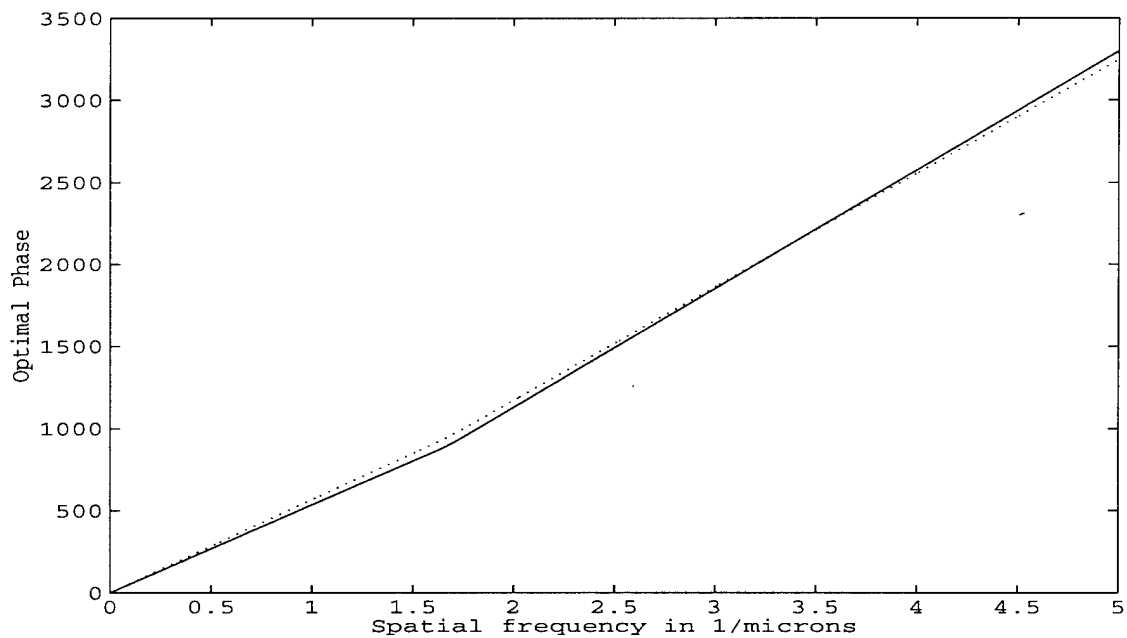
Figure 3.6: Optimal phase for 10 (...) and 15 (__) micron phase optimized notch filter.

35

The effect of the design "power spread" thickness (x_1-x_0) is illustrated in Figure 3.7 below. The relationship between index range and design thickness is approximately cubic over the range of thickness shown. There is no clear relationship between the parameters of the design problem and the design thickness. This is due in part to the fact that all films designed by this method are infinite in extent, and must be truncated by the designer. If the design thickness is chosen to match the desired physical thickness of the film, the output of the inverse Fourier transform will have significant index values far beyond the desired thickness. Truncation of the film will therefore have a significant effect on the film's actual reflectance properties. The SWIFT design method allows the exchange of increased thickness for decreased index range.
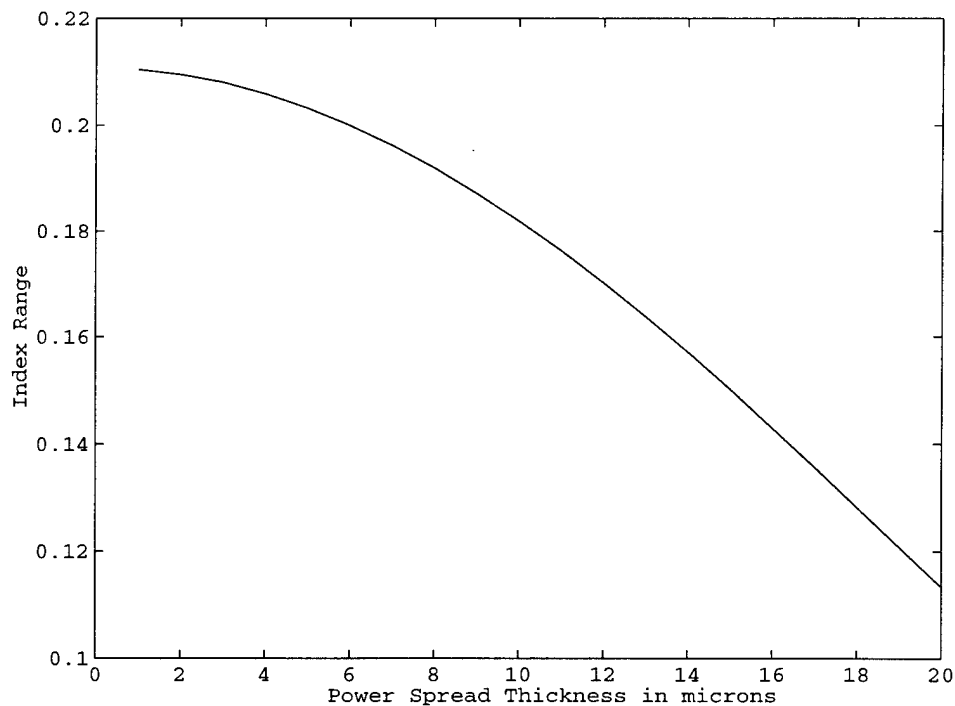
Figure 3.7: Effect of "power spread" thickness on index range.

3.3.2. Broadband Reflector Design

The value of phase optimization is further illustrated in the design of a broadband, high-reflectance rugate mirror over the wavelength range available to Ti:Sapphire lasers. This laser emits at wavelengths between 660 nm and 1.18 μm [33:480]. Presently, it is necessary to change the laser's optics to run it over its full operating range. For this example, specify a reflectivity of 99% between 700 and 1100 nm and 0% outside this band. Simulating a CeF$_3$:ZnS rugate filter, let $n = 1.89$ at each boundary and permit an index variation from 1.6 to 2.2 [21:61]. $Q(k)$ is chosen as in Equation 3.22, and the desired film optical thickness is chosen to be 40 μm.

As before, the numerical parameters for the sampling are found by requiring the index sample spacing, Δ, to be on the order of five nm and the total number of samples, N_s, to be a power of two. Since the reflection band is much larger than in the previous example, the number of non-zero samples of the Q function, N_0, should be about 200. Again, for details refer to Appendix A. The parameters selected for this example, based on these values, are

$$
\begin{aligned}
N_s &= 2^{15} = 32{,}768 \\
d_{Total} &= 200 \; \mu\, m \\
\Delta &= 0.0061 \; \mu\, m \\
N_0 &\Rightarrow 208 \; \text{non-zero samples of } Q
\end{aligned}
\tag{3.21}
$$

Again, this choice for the total optical thickness of the film should also minimize any aliasing effects, since it is much greater than the desired film optical thickness.

The index profiles and the reflectance spectra for this design with $\Phi(k)=0$ are shown in Figure 3.8. The index profile derived with zero phase ranges from 0.9 to 4.0 and is physically unrealizable. The resulting reflectance is a poor approximation to the design goal, with a reflectance as low as 98% over the desired wavelength range, and significant

reflectance outside the design band. Using the SWIFT algorithm for phase modulation with a "power spread" optical thickness of 35 μm, a film can be designed with an index contrast achievable using CeF_3:ZnS and is predicted to have a reflectance of ≥99.8% over the entire design range. The index profiles and the reflectance spectra for the optimized design are shown in Figure 3.9. The optimized design is very close to the design goal. In particular, the edges of the reflectance profile at 700 nm and 1100 nm are very sharp, with small oscillations. This can be very important in some applications. The film is fairly thick (about 25 μm physical thickness), due to the requirement for high reflectance over a 400 nm range, and zero reflectance outside this range. A portion of the optimized phase modulation is shown in Figure 3.10. As before, the rest of the phase is a linear continuation of the phase shown.
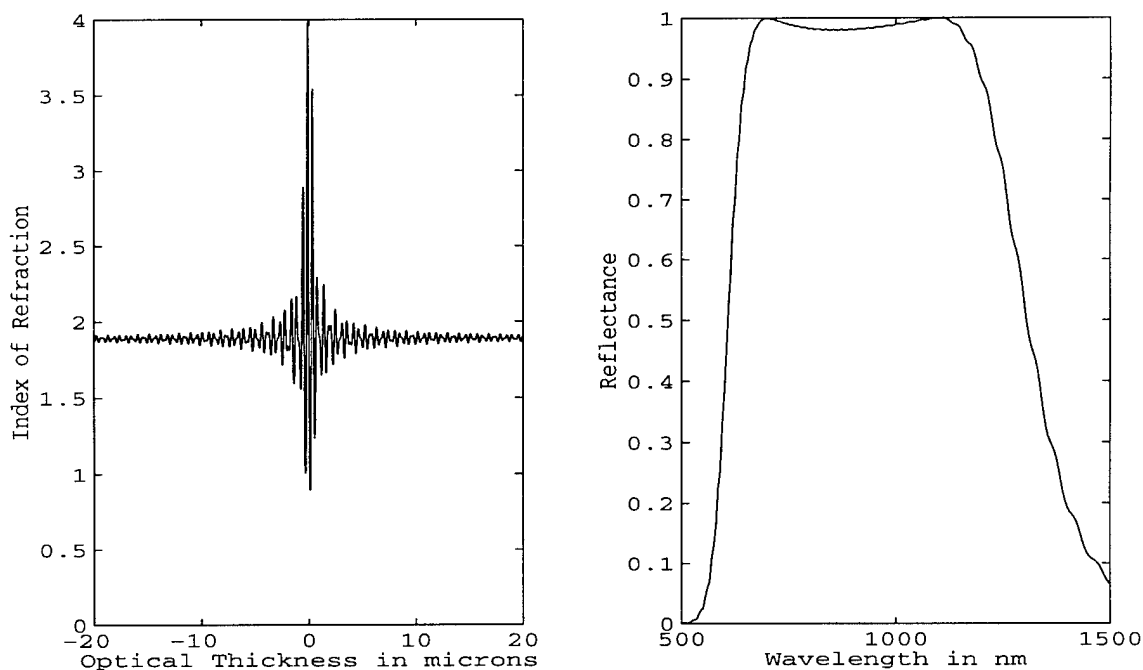
Figure 3.8: Index and reflectance for no phase Ti:Sapphire bandstop filter.
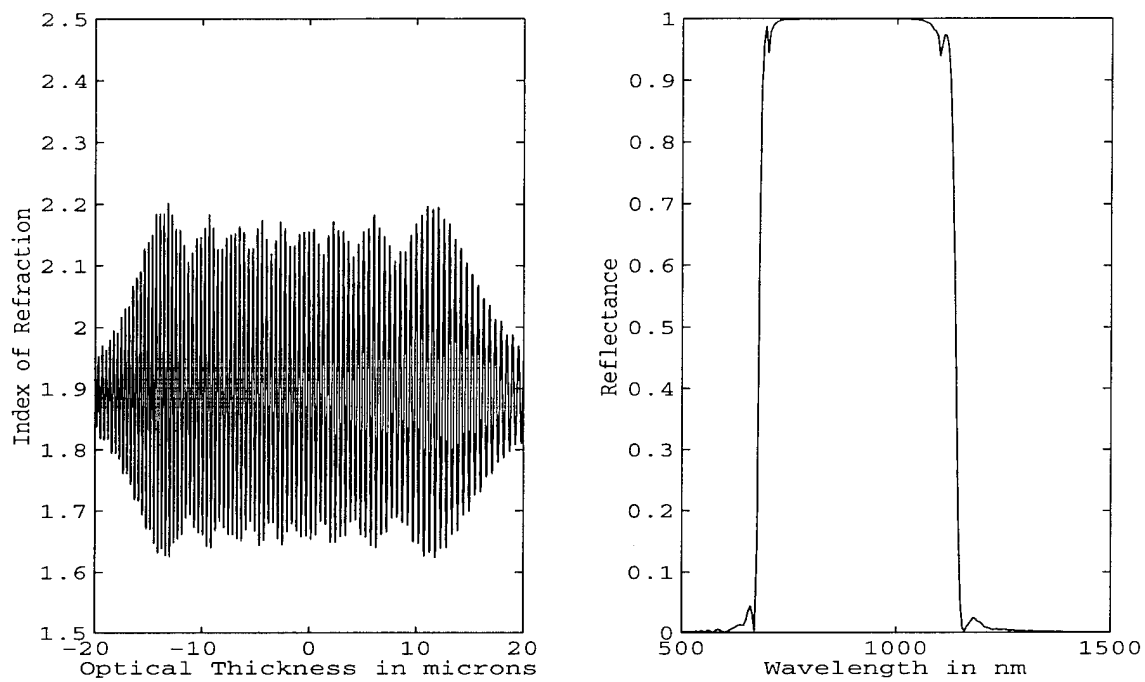
Figure 3.9: Index and reflectance for optimal Ti:Sapphire bandstop filter with 35 micron phase power spread thickness.
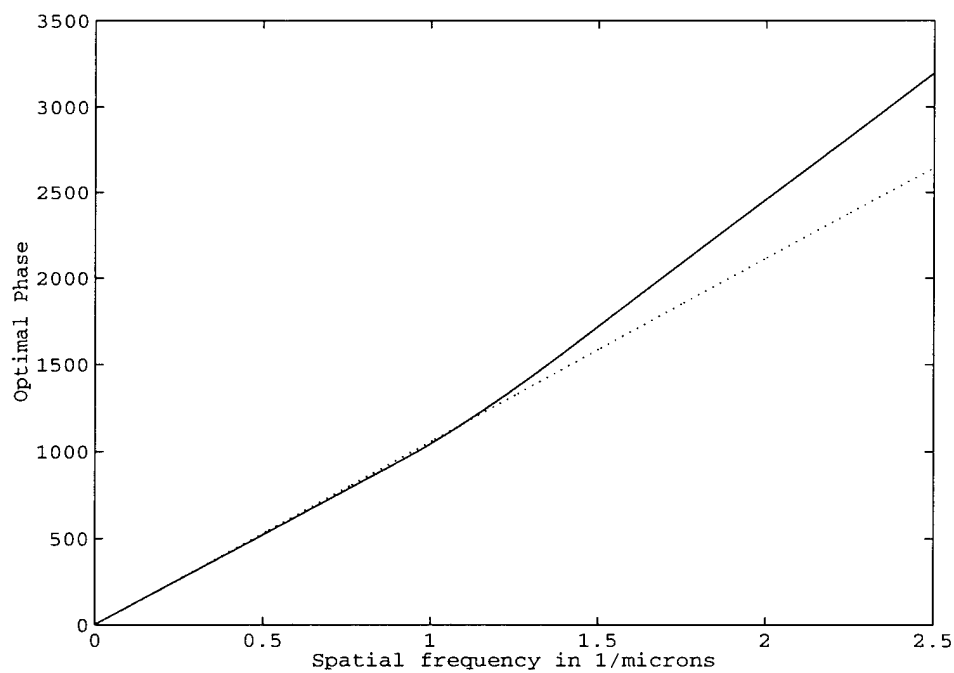
Figure 3.10: Optimal phase for Ti:Sapphire bandstop filter.

3.3.3. Nd:YAG Output Mirror Design

This example is a design for the external cavity mirrors of the Nd:YAG laser. In this case, the mirror must allow the pump laser (at 810 nm) to pass through the mirror, but reflect the other two laser wavelengths of 1.06 μm and 1.33 μm. For this example, the substrate index is n_{sub}=1.816, and the desired reflectance is 99% between 1.00 and 1.40 μm, and zero for all other wavelengths. The design notch is wider than the two design high reflectance points to minimize the "edge" effects and insure good high reflectance values. The previous examples illustrated small oscillations in the reflectance near the turn on and turn off wavelengths. The desired optical thickness for this film is 20 μm, again selected based on the zero phase film calculation.

As before, the numerical parameters for the sampling are found by requiring the index sample spacing, Δ, to be on the order of five nm and the total number of samples, N_s, to be a power of two. Since the reflection band is about half the size of the previous example, the number of non-zero samples of the Q function, N_0, should be about 100. The parameters selected for this example, based on these values, are

$$
\begin{aligned}
N_s &= 2^{15} = 32{,}768 \\
d_{Total} &= 175 \; \mu\, m \\
\Delta &= 0.00534 \; \mu\, m \\
N_0 &\Rightarrow 100 \text{ non - zero samples of } Q
\end{aligned}
\tag{3.22}
$$

This choice for the total optical thickness of the film should also minimize any aliasing effects, since it is much greater than the desired film optical thickness.

The index profiles and the reflectance spectra for this design with $\Phi(k)$=0 are shown in Figure 3.11. The index profile derived with zero phase ranges from 1.0 to 3.25. The resulting reflectance is a poor approximation to the design goal, with a reflectance as low as 98% over the desired wavelength range. Using the SWIFT algorithm for phase

modulation with a "power spread" optical thickness of 10 μm, a film can be designed with an index contrast achievable using MgF$_2$:ZnSe (index range 1.38 to 2.5) and is predicted to have a reflectance of ≥99.8% over the entire design range. The index profiles and the reflectance spectra for the optimized design are shown in Figure 3.12. A portion of the optimized phase modulation is shown in Figure 3.13 . As before, the rest of the phase is a linear continuation of the phase shown.
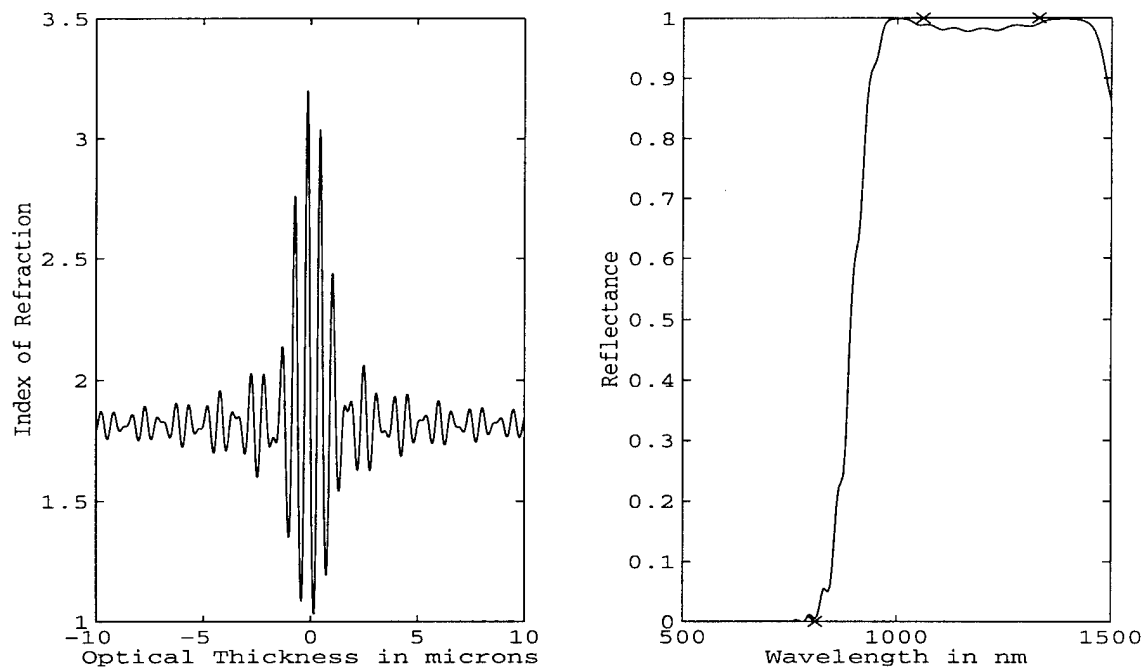
Figure 3.11: Index and reflectance for no phase Nd:YAG output coupler.
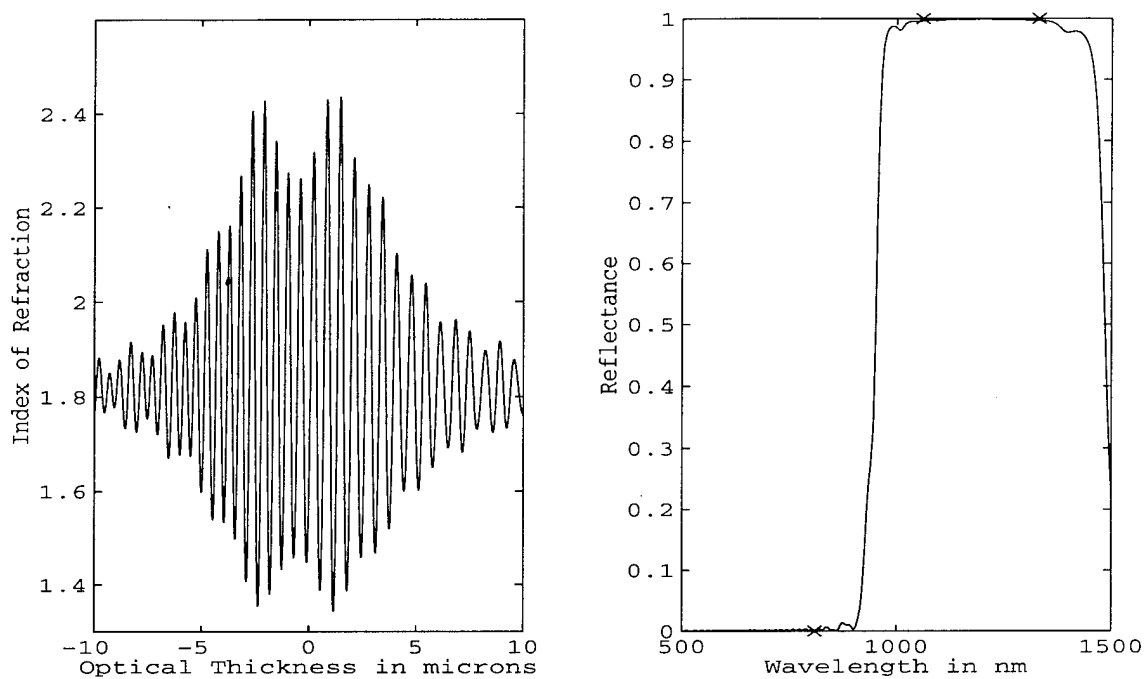
Figure 3.12: Index and reflectance for optimal Nd:YAG output coupler with 10 micron phase power spread thickness.
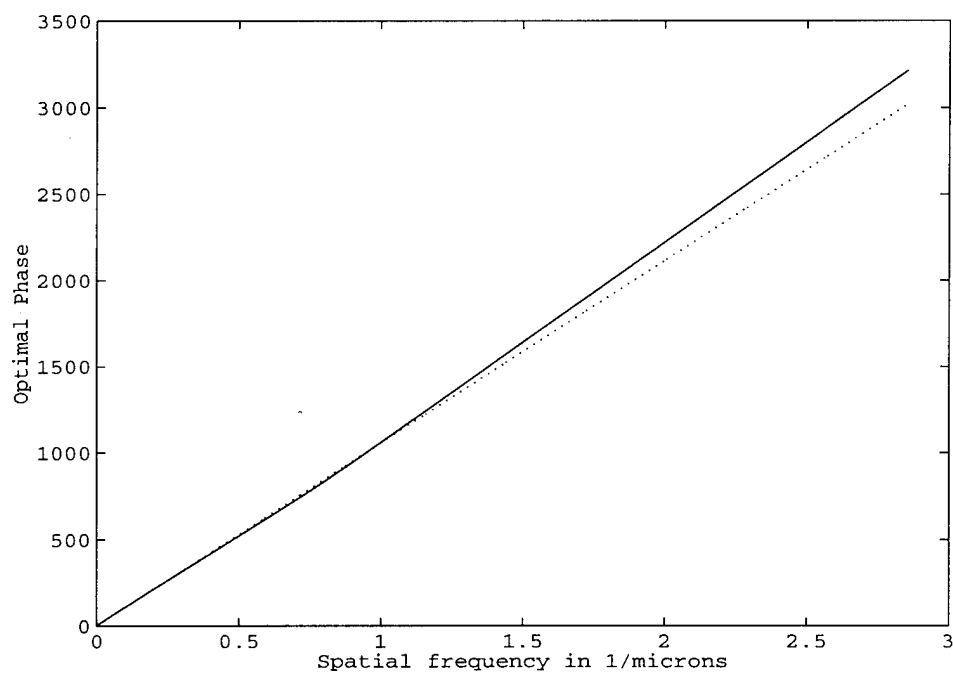
Figure 3.13: Optimal phase for Nd:YAG output coupler with 10 micron phase power spread thickness.

42

One difficulty with using this method to perform realistic designs is the inverse Fourier transform assumes identical entrance and exit media. Using a structure of (Substrate : Film : Air) instead of (Substrate : Film : Substrate) in the reflectance calculation results in significant degradation in the non-reflecting portion of the design, as illustrated in Figure 3.14. The figure shows significant ringing about the nominal 9.5% reflectance from the Substrate : Air interface. The other difficulty with this method lies in the choice of reflectance profile needed to solve the design problem. This problem only specified criteria at three wavelengths, but the SWIFT design requires a reflectance profile be specified over a much larger range. There are an infinite number of possible reflectance profiles that will satisfy the three wavelength design requirements. Unfortunately, there is no way to know *a priori* which reflectance profiles are "good" in terms of design costs (thickness) or performance.
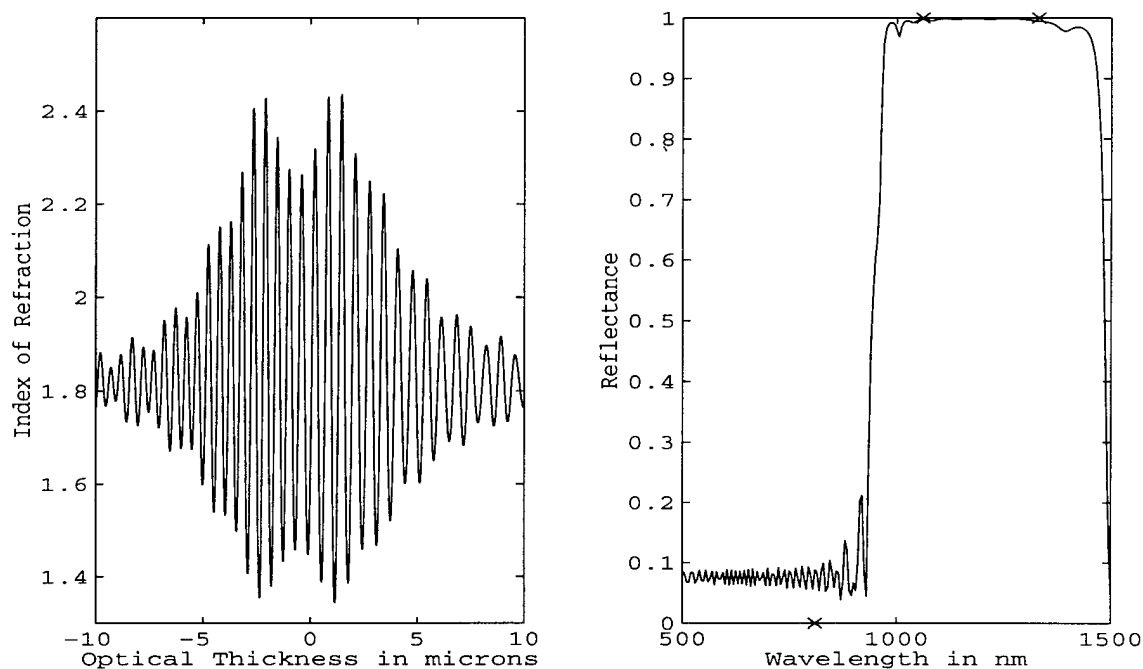
Figure 3.14: Index and reflectance for optimal Nd:YAG output coupler with air boundary and 10 micron phase power spread thickness.

43

3.3.4. Antireflection Coating Design

Another class of design problems to be addressed is that of an anti-reflective coating for an arbitrary substrate. The goal in this type of design problem is to reduce or eliminate the reflectance from the substrate-air interface. To design an anti-reflection coating using the inverse Fourier transform method, some of the assumptions used in the initial derivation must be changed. Specifically, the assumption of identical incident and exit media must be removed. The difference in the incident and exit media can be accommodated by taking advantage of the linearity of the Fourier transform. Consider an index profile consisting of a semi-infinite slab of substrate with an air interface. This single interface has a constant reflectance for all wavelengths (neglecting dispersion), given by the familiar Fresnel relations. The semi-infinite slab -- constant reflectance also satisfies the inverse Fourier transform relation in Equation 3.1. Figure 3.15 illustrates this Fourier transform pair. Figure 3.15 (a) depicts a Heaviside distribution or "step function", and Figure 3.15 (b) illustrates its Fourier transform.
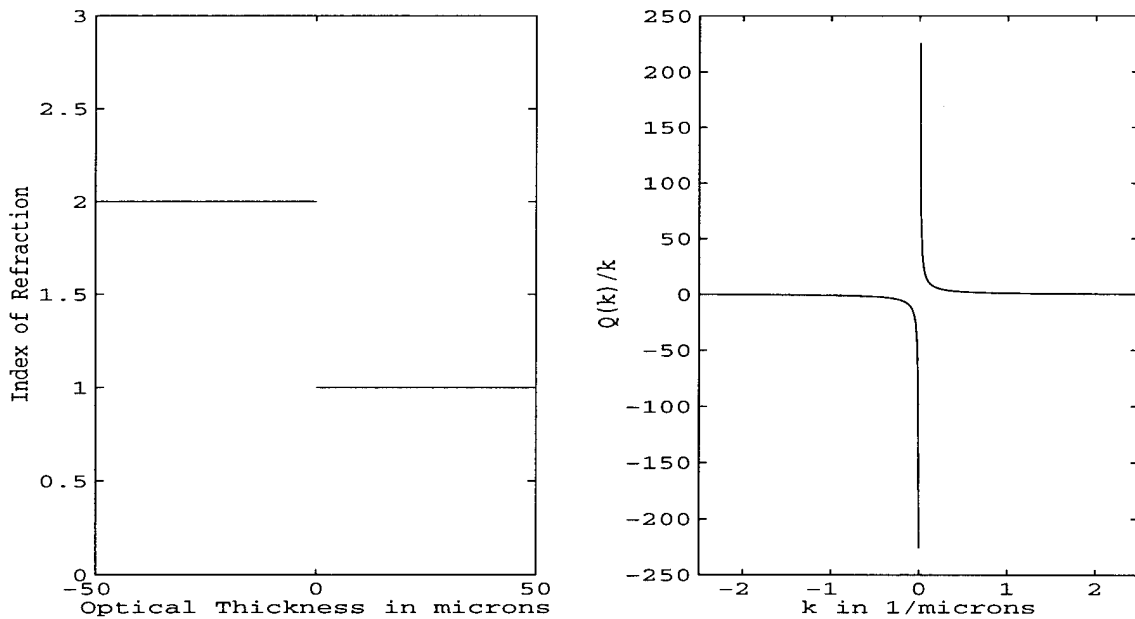
Figure 3.15 : Fourier transform of a semi-infinite slab.

44

Since the Fourier transform is a linear operation, an anti-reflection coating can be designed by superimposing two design elements; the slab-air interface and notch anti-reflector. The Fourier transform relations in Equation 3.1 are

$$\ln\left(\frac{n(x)}{n_0}\right) \Leftrightarrow \frac{Q(k)}{i\pi k} \tag{3.23}$$

The index profile can be separated into a step component, representing the slab-air interface, and a "film" component, which is responsible for the anti-reflective properties. These two elements are multiplied together as the argument of the logarithm, and so their Fourier transforms are summed:

$$\ln\left(\frac{n(x)}{n_0}\right) = \ln\left(\frac{n_{step}(x)\,n_{film}(x)}{n_0}\right)$$

$$= \ln\left(\frac{n_{step}(x)}{n_0}\right) + \ln\left(n_{film}(x)\right) \tag{3.24}$$

$$\Leftrightarrow \frac{1}{i\pi k}\left(Q_{step}(k) + Q_{film}(k)\right)$$

Since the index and Q function for the slab-air interface are known, only the index of the film corresponding to the anti-reflective region needs to be designed. The Q function for the desired anti-reflection property is found by taking the negative of the slab Q function in the wavelength region of interest. Thus, the anti-reflection coating problem reduces to the notch reflector solved previously.

45

As an example, consider the problem of designing an anti-reflection coating for the gain medium of the Nd:YAG laser. The Nd:YAG substrate has an index of refraction of 1.816, which causes a Fresnel reflectance at the substrate-air boundary of 8.4%. In order to eliminate this loss inside the cavity, the gain rod should have a coating which greatly reduces this reflectance at the pump (810 nm) and lasing wavelengths (1.06 μm and 1.33 μm). Thus the anti-reflection design requires the reflectance be minimized between 800 nm and 1.4 μm. The steps in this design process are: 1) design a notch reflector with a reflectance equal to the Fresnel reflectance using air as substrate and exit media; 2) form an index profile for the substrate - air boundary with the interface at the origin; and 3) form the film by dividing the substrate - air film by the notch reflector film. The optical thickness for each step in the final film is the same, and equal to the optical thickness of the notch reflector design. Thus, the division in step 3) above changes both the index and thickness of the film.

The numerical design parameters are found as before. For 200 non-zero sample points of the notch, the appropriate sampling parameters are

$$
\begin{aligned}
N_s &= 2^{15} = 32{,}768 \\
d_{Total} &= 200\ \mu m \\
\Delta &= 0.0061\ \mu m \\
N_0 &\Rightarrow 214\ \text{non-zero samples of } Q
\end{aligned}
\tag{3.25}
$$

Figure 3.16 shows the index profile which results from step one of this design method with no SWIFT phase optimization. Figure 3.17 shows the final index profile and resulting reflectance. The nominal reflectance of 8.4% has been reduced to less than 0.2% across the design wavelength range.
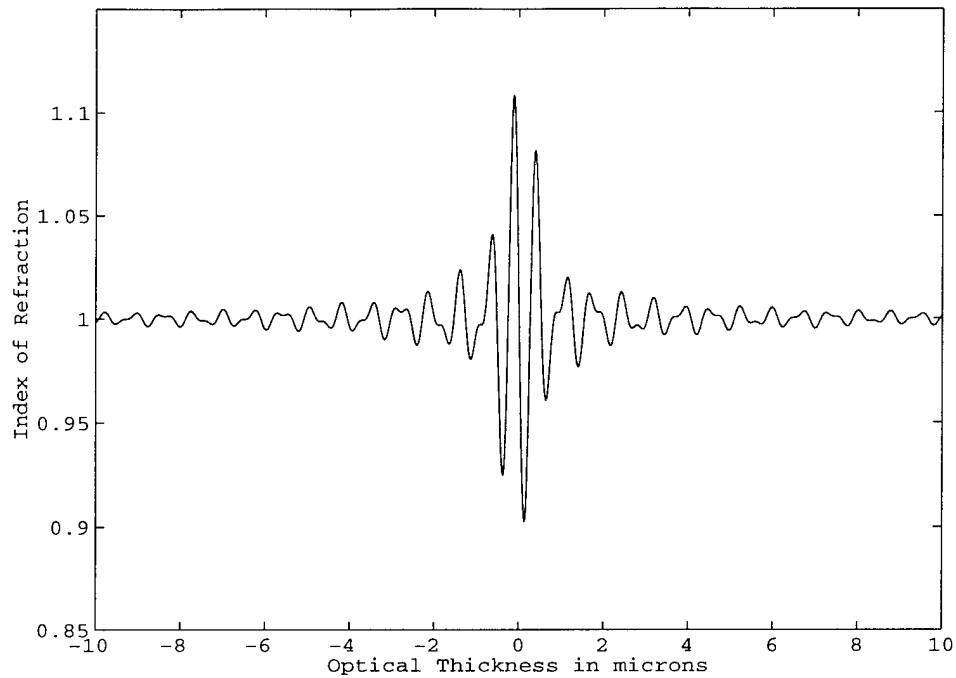
Figure 3.16 : Anti-reflection coating index design without step discontinuity.
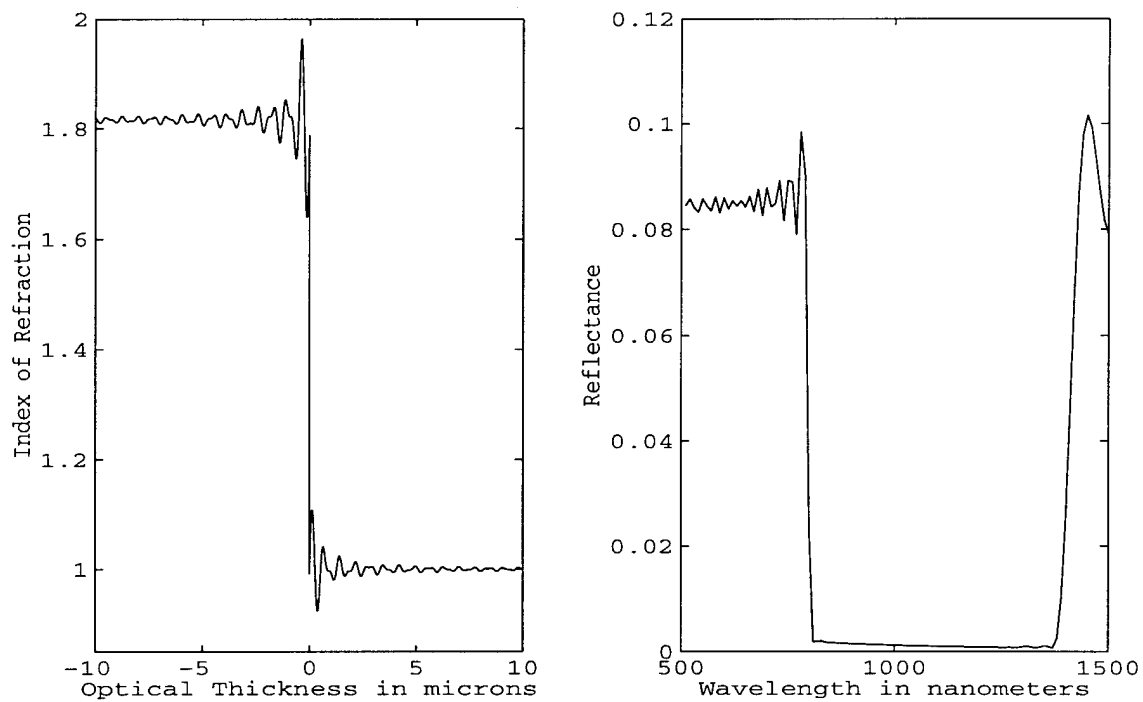
Figure 3.17: Final Fourier transform anti-reflection coating design.

47

There is, however, one problem with this formulation. The slab-air interface is assumed to be at the origin in the above derivation. The film designed by subtracting a notch from the initial Q function profile is also centered on the origin. When this film is combined with the slab-air interface, half of the index profile is modifying the air side (see Figure 3.17). This results in non-physically realizable indices. This problem cannot be solved by shifting the slab-air interface, since the shift would be included as an oscillation in the Q function, and the resulting AR film would be shifted by the same amount. The SWIFT algorithm is also not able to address this problem. While SWIFT does affect the index of refraction range, it does so in a symmetric fashion about the origin. SWIFT does not introduce any linear shifts in the index profile. The inverse Fourier transform method is therefore not well suited to the problem of anti-reflection coating design. Other design methods, which are capable of addressing this Nd:YAG laser anti-reflection design problem, are presented in the next chapter.

3.4. Conclusion

This chapter has introduced the inverse Fourier transform method for the design of gradient index optical coatings, and has presented a modification to this method, called the Stored Waveform Inverse Fourier Transform (SWIFT) technique, that allows the designer to control the index range used in the design. The SWIFT technique determines an optimal phase to use in the inverse Fourier transform, which has the effect of spreading the index contrast more evenly over the total thickness of the film. The SWIFT technique also generates film designs that better achieve the desired film reflectance profile, as compared to films designed without the SWIFT optimal phase. The preceding examples illustrated the class of problems for which the SWIFT method was well suited. Specifically, problems which require a complex reflectance profile over a broad range of

wavelengths. The method has several limitations, however. First, as shown above, the inverse Fourier transform cannot be directly applied to the anti-reflection coating class of problems for substrate -- air interfaces. Second, the entire reflectance profile must be specified by the designer, even if the region of interest is only a small portion of the spectrum. While it is not difficult to arbitrarily specify additional desired reflectance values, it is difficult to choose the profile which will give the best performance in the region of interest. For this reason, the SWIFT method is not well suited for the design of coatings for laser applications in which only a few specific wavelengths are of interest.

4. Design of Gradient Index Films Using Generalized Fourier Series

This chapter presents an alternative method for gradient index thin film design based on an iterative approach using generalized Fourier series. This new method of gradient index thin film design extends the domain of problems for which gradient index solutions can be found. The method is analogous to existing techniques for layer based coating design, but adds the flexibility of gradient index films by varying the index of refraction instead of the thickness of the layers. The coefficients of a generalized Fourier series representation of the gradient index of refraction profile are used as variables in a non-linear constrained optimization formulation. This allows one to design a piece-wise continuous gradient index film with limited number of variables. The optimal values of the design coefficients are determined using a sequential quadratic programming algorithm. The first section outlines the constrained optimization approach as applied to thin film design. The second section illustrates this method using the Fourier series basis to specify the film, and the third section applies this method using a wavelet basis to specify the film. The fourth section describes a method for finding the minimum thickness for an optimal thin film design.

4.1. Optimal Design Problem Formulation

This section describes the constrained optimization problem for the design of gradient index thin films. The process of optimal design requires a statement of the problem to be solved, identification of alternative solutions, and some measure of what constitutes a "best" solution to the problem. For simple problems, it may be possible to exhaustively list all possible solutions to determine which one is best. For more complex problems, however, a more structured approach is needed to insure the "best" solution is

50

found. This structured approach consists of building a mathematical model for the system in question, identifying a set of variables to adjust in the design, and defining a "merit function" to numerically identify the best, or "optimal" solution. The general theory for numerical constrained optimization used in this dissertation is summarized in Appendix B.

To apply this theory, the problem of thin film design must be stated in appropriate form. First the problem should be stated in physical terms, and then translated into a mathematical form. The physical statement of the problem is: given a desired reflectance as a function of wavelength and a range of available index of refraction values, generate an index of refraction as a function of thickness of the film. The film must use index values within the specified range and meet the desired reflectance profile. In the optimal design method, the merit (or objective) function stems from the desired reflectance, and the constraints are due to the requirement of an "acceptable" index profile. In this case "acceptable" means all index values are real (dielectrics only), the index must be the same for all wavelengths (neglect dispersion) and the index values must be within the specified range.

Before this statement of the problem can be translated into a mathematical formulation, the variables of the model must be selected. The majority of the optimal design techniques in the literature focus on finding the thickness of alternating layers of predetermined high and low index materials. To use the theory of optimal design on a gradient index film, a different approach to the definition of variables is needed. Instead of describing the film as a collection of a few discrete layers, the film is specified by a collection of coefficients with respect to a basis. A basis is a linear algebra concept, defined as a set of vectors in a space such that any vector in the space can be represented in one and only one way by a linear combination of these "basis" vectors [38:228]. The two basis systems explored in this work are the Fourier basis of sine and cosine functions, and the basis of Daubechies' wavelets in a multiresolution analysis. The Fourier basis

was selected because of its familiarity. The wavelet basis was selected because it provides a new paradigm for functional analysis. Each will be described further below.

In general, the index of refraction as a function of position in the film can be formed by a weighted sum of the basis elements. The properties of the film must then be determined from this index of refraction. One method for determining the characteristics of a gradient index film is to approximate the film by a series of very thin homogeneous slabs, and then use standard matrix methods on this approximate film. The details of this approach to characterizing the reflectance of a multilayer thin film were discussed in Section 2.3.1.

Now that the problem has been stated in physical terms, and the variables of the design have been selected, the optimal design problem can be stated mathematically:

$$\min_{c_i \in \Re, \; i=1,2...M} \left\| R(k) - R_d(k) \right\|$$

subject to:

$$N_L \leq N(z) \leq N_U$$

where :

$$(4.1)$$

$$N(z) = \sum_{i=1}^{M} c_i \phi_i(z), \qquad z \in [0, L]$$

$$R(k) = f(k; N, L)$$

That is, find the values of c_i, which are restricted to real numbers, such that distance between $R_d(k)$, the design reflectance profile at wavenumbers k, and $R(k)$, the calculated reflectance of the film at wavenumbers k, is minimized. The minimization is subject to the constraint that the index of refraction as a function of position in the film, $N(z)$, must be between the lower and upper limits on the index of refraction, N_L and N_U respectively. The index of refraction as a function of position in the film, $N(z)$, is expressed as a weighted sum of the unknown coefficients, c_i, times the basis functions $\phi_i(z)$. The total thickness of the film is denoted by L. The reflectance of the film is a function of the

wavenumber, k, and the index of refraction, N, over the total film thickness, L. The details of this functional dependence were previously described in Section 2.3.1. In practice, the reflectance is found using the matrix methods previously described in Section 2.3.2. The remaining variable, N_{air}, is the index of refraction of the exit media, assumed in this work to be air. The norm in Equation 4.1 to minimize is the L^2 metric

$$\|R_d - R\| = \left(\int_{k_L}^{k_U} |R_d(k) - R(k)|^2 \, dk \right)^{1/2} \tag{4.2}$$

where k_L and k_U are the lower and upper bounds of the wavenumbers of interest.

The approach to solving the mathematical problem of Equation 4.1 is to approximate it with a discrete version of the problem. The wavenumber range of interest, k_L to k_U is converted to a sequence of wavenumbers of interest. This is not necessarily detrimental, particularly when designing thin films for laser applications. The continuous gradient index film is approximated by dividing it into a number of thin, homogeneous layers for constraint checking and reflectance evaluation. The objective function (or merit function) to minimize for the designs considered here involves the difference between the desired reflectance and the reflectance calculated for the film at a number of different design wavelengths. Numerous merit functions for optimal design of stack-based films have been explored [13:2825-2826]. The merit function, F, used here is the ℓ^2 metric:

$$F = \left(\sum_{i=1}^{K} \left(R_d(k_i) - R(k_i) \right)^2 \right)^{1/2} \tag{4.3}$$

where $R_d(k_i)$ are the design reflectances at specific wavenumbers k_i, $R(k_i)$ is the calculated reflectance of the film at the design wavenumbers k_i, and K is the total number of design

wavelengths used. This merit function was selected because it is continuous with respect to the variable argument $R(k_i)$, and has been shown to provide good convergence for multilayer thin film optimal designs [13:2825-2826].

The main physical constraint on the problem is the index must always remain within a range of acceptable values. This is implemented by building samples of the index profile from the variables of the design, and constraining the samples to be within the range of allowable indices. For N samples, this generates $2N$ constraints (N for the upper bound and N for the lower bound). A statement of the numerical optimization problem, using the same notation as Appendix B with coefficients $(c_1, c_2, ..., c_M)$ as variables is therefore

$$\min_{c_i \in \Re} \qquad F = \left(\sum_{i=1}^{K} \left(R_d(k_i) - R(k_i) \right)^2 \right)^{1/2}$$

subject to:

$$g(j) = N(z_j) - N_U \le 0, \qquad j = 1, 2, ... N$$
$$g(N+j) = N_L - N(z_j) \le 0, \quad j = 1, 2, ... N \qquad (4.4)$$

where:

$$N(z_j) = \sum_{i=1}^{M} c_i \phi_i \left(z_j \right), \qquad j = 1, 2, ... N$$

$$R(k_i) = f\left(k_i, N(z_1), N(z_2)...N(z_N) \right)$$

Here the construction of the index profile from the variables is denoted by a generic basis function "ϕ_i". The specific functional form will depend on the basis used. The calculation of the reflectance is done using the matrix methods described in Chapter 2, and is denoted by the function "f".

4.1.1. Optimal Design Algorithm

This section discusses the implementation of this optimal design method. The numeric optimizations were performed using the MATLAB™ programming environment.

54

The first step in the design process (after the statement of the problem), is to select the generalized Fourier series basis to use. This requires the user to specify a method for generating representations of the basis function and the algorithm for decomposition and reconstruction of an arbitrary function in this basis system. Two example basis systems, Fourier series and wavelets, will be illustrated in Sections 4.2 and 4.3 below.

In order to describe a gradient index film by a set of decomposition coefficients with respect to some basis (the variables in the optimization), a number of parameters of the film are needed. The most important is the total thickness of the film. This is a key parameter because it determines the period of the periodic extension of the interval to the real line. This plays an important part in most representation schemes. Other parameters of the optimization problem include the substrate index, the range of indices to use, and the number of samples to use in the reflectance calculation. In addition, the number of non-zero decomposition coefficients to use as variables in the optimization must be determined.

The sequential quadratic programming optimization method is a local minimization method. It requires an initial value for the variables. The variables considered here are coefficients of basis elements describing the index of the film. The initial estimate of the variables is an initial estimate of the film. In the absence of any knowledge about the final films characteristics, any feasible initial guess will produce a local solution, A feasible guess is defined as one that satisfies the constraints on the index values. It is reasonable to choose initial values corresponding to a flat film with an index in the acceptable range. Note that a "better" initial value for the variables may result in a faster optimization. Also, this technique does not guarantee a global minimum will be found. Different initial values may lead to different solutions.

At this point it is useful to step through the optimization process. The optimization starts with an initial estimate of the film. This initial film is selected based on the designer's experience, and may be the result of other design techniques. In most cases in

this work, the initial design was a constant with the same index of refraction as the substrate. This estimate is then decomposed into coefficients with respect to the basis being used. For most cases, the number of coefficients used as variables in the optimization will be less than the total number of coefficients calculated, which is equal to the number of sample points. The coefficients not used as variables are fixed at their input values. These initial variable estimates are used to start the iterative optimization procedure. The main steps of the iteration are outlined below:

1. Calculate the index represented by the current value of the variables. Method depends directly on choice of basis.

2. Calculate the reflectance of this index profile at the design wavelengths using *REFLECT.M*.

3. Calculate the value of the Merit Function F for this reflectance. (See Equation 4.3 above)

4. Calculate the values of each constraint on the index profile.

5. Compare the Merit Function F and constraint values g_i with the termination criteria.

6. If termination criteria are satisfied, stop. Otherwise continue.

7. The values of the Merit Function F and constraints g_i are used in the sequential quadratic programming algorithm to determine the next guess for the values of the variables.

8. Go back to step 1 and repeat until termination criteria are met or maximum number of iterations is exceeded.

The details of the sequential quadratic programming algorithm used to determine the new estimates for the variables are discussed in Appendix B. The optimization goal is achieved when three termination criteria have been satisfied. There are three numerical tolerances to be met, one on the objective function, one on the variables, and one on the constraints. The objective function tolerance specifies the precision required on the

objective function at the solution. The termination criteria for variables specify the minimum acceptable precision on the values of the variables. The constraint termination criterion is the maximum allowable violation of the constraints at the solution point. All three termination criteria must be satisfied simultaneously to achieve an optimal solution. The default termination tolerances are: objective -- 10^{-4}, variables -- 10^{-4}, and constraints -- 10^{-7}. All three can be adjusted if necessary. In addition, there is a maximum number of iterations allowed, which defaults to $100N$ where N is number of variables but can also be set manually. In the event the maximum number of iterations is exceeded, the best solution found during the optimization is reported, along with the value of the merit function and a warning that the termination criteria were not met.

The programs used to implement this method depend directly on the basis system selected. The specifics of the programs used will be addressed in the context of the examples below. The two basis systems explored in this work are the Fourier basis of sine and cosine functions, and the basis of Daubechies' wavelets in a multi-resolution analysis. Note that these are only two of many possible basis sets that could be used in this technique. Other familiar basis systems include Bessel functions, Legendre polynomials, Hermite polynomials, Laguerre polynomials and Chebyshev polynomials [2:525].

4.2.   Optimal Design using Fourier Series

The first basis system to be considered is the Fourier basis of sines and cosines. This basis was selected because it should be familiar to most readers. The decomposition of an index profile on the interval into Fourier series coefficients is described very briefly below. Three examples of optimal designs using this approach are also given.

4.2.1. Fourier Basis

The most well known basis for decomposing a periodic function is the Fourier basis of sines and cosines. The relation between the function, $f(x)$, and the basis elements is

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx), \qquad 0 < x < 2\pi \qquad (4.5)$$

where the coefficients are given by

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x)\cos(nx)\,dx$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x)\sin(nx)\,dx \qquad n = 0,1,2\ldots \qquad (4.6)$$

Sine and cosine functions form a basis for functions that are square integrable over an interval from 0 to 2π. The first step in building an index profile from these basis elements is to map the film interval, which is 0 to the total thickness of the film, T, to the interval 0 to 2π. This is done by a change of variable from x to z, given by $x = 2\pi z / T$. The next step is to identify a computationally efficient method for constructing samples of the index of refraction profile for a given set of Fourier coefficients. This is done by using the inverse fast Fourier transform algorithm. The inverse fast Fourier transform maps complex Fourier coefficients to samples of a complex valued function. The complex Fourier coefficient c is related to the a and b coefficients above by $c_n = a_n - ib_n$. In addition, the inverse fast Fourier transform algorithm uses both positive and negative frequencies. To insure the index profile is strictly real, the real Fourier coefficients must have even symmetry and the imaginary Fourier coefficients must have odd symmetry.

The optimal design method using Fourier series coefficients for the index of refraction as variables uses the inverse fast Fourier transform for the mapping function ϕ described in Equation 4.4 above. The total thickness of the film must be divided into a number of very thin layers for the reflectivity calculation to be a good approximation of the actual gradient index film reflectivity. The thickness of the layers for this calculation should be less than about 5 nm [6:5429].

The inverse fast Fourier transform yields as many samples of the index of refraction as there are frequencies (both positive and negative). To obtain N samples of the index of refraction profile, the complex coefficients for N frequencies must be specified. Since the index of refraction must be real, the symmetry requirements on the coefficients described above reduces the $2N$ variables down to N variables. This is still much too large. However, by choosing to use only a few low frequency coefficients as variables and setting the coefficients of higher frequencies to zero, the size of the problem can be reduced while still achieving the required number of samples for the index. Thus, the number of frequencies to use as variables becomes one of the design parameters. Since the coefficients in the inverse fast Fourier transform are complex, there are two variables for each frequency used as a design variable.

The MATLAB™ programs used to perform the optimal designs in the Fourier basis system are included in Appendix D. Several functions are combined to make up the program. The script **FFTRUN.M** is the main function, which sets the initial parameters and calls the optimizer. The key element of this implementation is the MATLAB™ optimization toolbox function called **CONSTR.M** [26], which implements the sequential quadratic programming algorithm described in Appendix B. The MATLAB™ help file describing this function is also included in Appendix D. The **CONSTR.M** function requires inputs of: the name of the evaluation function containing the relationship between the variables and the merit function and constraints, the initial variable array, and an array specifying the optimization parameters. The optimization

59

parameters include values for the termination criteria, maximum step size, etc. The function also requires inputs for the lower and upper bounds on the variables, and any parameters to be passed to the evaluation function. The outputs are the optimal values for the variables and an array containing the parameters used in the optimization. The output includes the final values of the termination criteria and the number of iterations used.

The evaluation function **FFTFUN.M** contains the conversion between coefficients of the Fourier series expansion and samples of the index. It also calculates the reflectance of the film at each design wavelength, and the merit function and constraint values used in the optimization. The inputs to this function are: the current variables values, the number of sample points, the desired reflectance at each design wavelength, the substrate index, the film thickness in microns (or nanometers), the array of design wavenumbers k in inverse microns (or inverse nanometers), and the array of sample points x in microns (or nanometers). The units selected for length must be consistent. The outputs of the function are the values of the merit function and the constraints. The reflectance for the design is calculated using **REFLECT.C**, described in Chapter 2.

4.2.2. Nd:YAG Laser Anti-Reflection Coating

The first example of optimal design of a gradient index film is to create an anti-reflection (AR) coating for the gain medium of a neodymium : yttrium aluminum garnet (Nd:YAG) laser. The goal is to eliminate reflections at the two primary laser wavelengths of 1.06 μm and 1.33 μm, as well as the pump laser wavelength of 810 nm. The Nd:YAG material substrate has an index of refraction of 1.816, which means the uncoated surface has a reflectance of 8.4%. The index range for this design example is chosen based on a material system of MgF_2 and ZnSe, which have indices of refraction of 1.38 and 2.5 respectively. These materials have been successfully combined to form intermediate index films with good optical and mechanical properties [20:197-204].

Several values for the thickness of the film are explored, all on the order of hundreds of nanometers. Note that these films are much thinner than any designed in Chapter 3 using the SWIFT technique. For this range of film thicknesses, the 5 nm sampling density desired for the reflectance calculation can be achieved in all cases with 128 samples. This choice for the number of samples, and thus the total number of coefficients, insures the inverse Fourier transform used to calculate the decomposition coefficients is efficient. The thicknesses used in this design are 300 nm, 500 nm, 750 nm, and 1000 nm. In addition, the effect of the number of variables used in the optimization is also explored. The numbers of non-zero coefficients to use as variables are 8, 16, 32, and 64. All other coefficients are fixed at zero. Only even values for the number of coefficients are used because the two coefficients are required to specify one frequency in the Fourier series expansion. The values used here are also powers of two, though they need not be for this method. They were restricted to powers of two for ease of comparison with the wavelet based optimizations which follow. In all cases the initial film design used was a constant index equal to that of the substrate.

The first parameter to vary is total thickness. The number of non-zero coefficients used as variables for these designs is fixed at 16. Figure 4.1 through Figure 4.4 below show the film designs and resultant reflectivity profiles for total thicknesses of 300 nm, 500 nm, 750 nm, and 1000 nm respectively. The reflectances at the design wavelengths for all designs are given in Table 4-1, along with the value of the merit function for the design, which is the root mean square (RMS) error between the desired reflectivity and the design. The optimal designs for both the 300 nm and 500 nm thicknesses reduce the reflectance significantly from the nominal 8.4%, but neither is acceptable for use inside a laser cavity. The designs for 750 nm and 1000 nm both perform much better than the designs for 300nm and 500 nm. This demonstrates the importance of the total thickness as a parameter in the design. This is not a surprising result, but bears emphasizing. Increasing the film thickness enhances the film's performance and decreases the merit

function value for the design. In addition, as the film thickness increases the index range used to achieve the optimal design generally decreases. Figure 4.2 shows the 500 nm design used the entire available index range while the 750 nm and 1000 nm thickness designs (Figure 4.3 and Figure 4.4) did not. This means the index constraints were active at the 500 nm solution point.

Table 4-1: Design results for various thicknesses of Fourier series Nd:YAG AR coatings.

Wavelength (nm)	300 nm film	500 nm film	750 nm film	1000 nm film
810	0.006508	0.001849	3.034×10^{-8}	2.646×10^{-9}
1060	0.005650	0.003242	3.270×10^{-8}	7.896×10^{-9}
1330	0.009282	0.002823	5.603×10^{-8}	2.535×10^{-9}
Total RMS Error	0.012666	0.004679	7.162×10^{-8}	8.705×10^{-9}
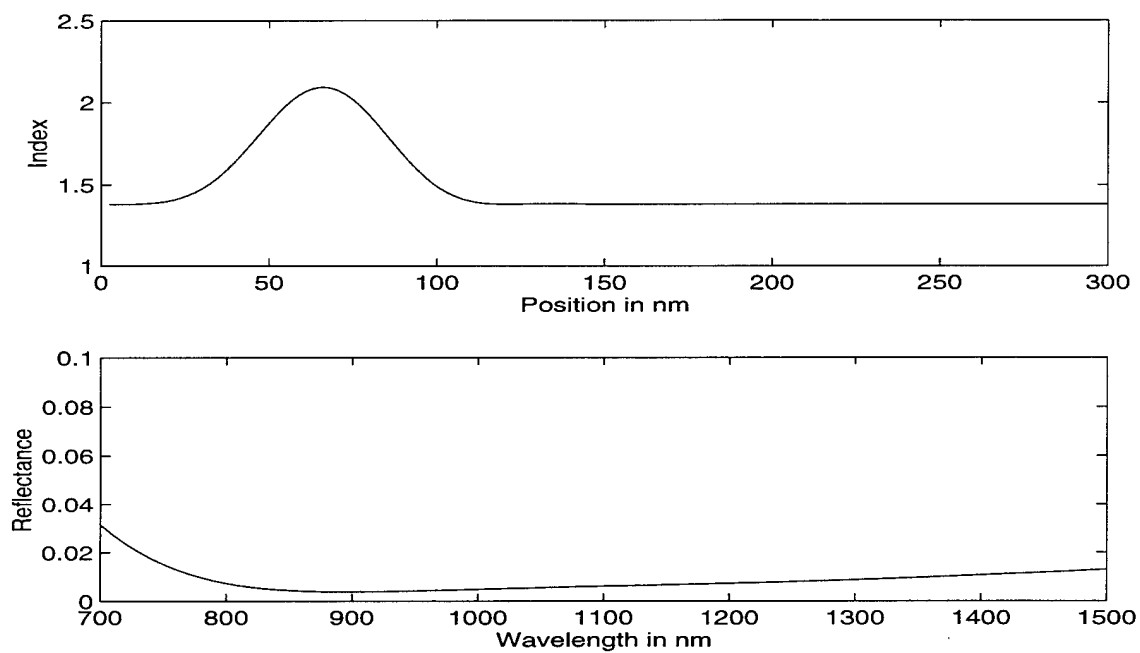
Figure 4.1: Fourier series design for Nd:YAG AR coating for 300 nm thick film.
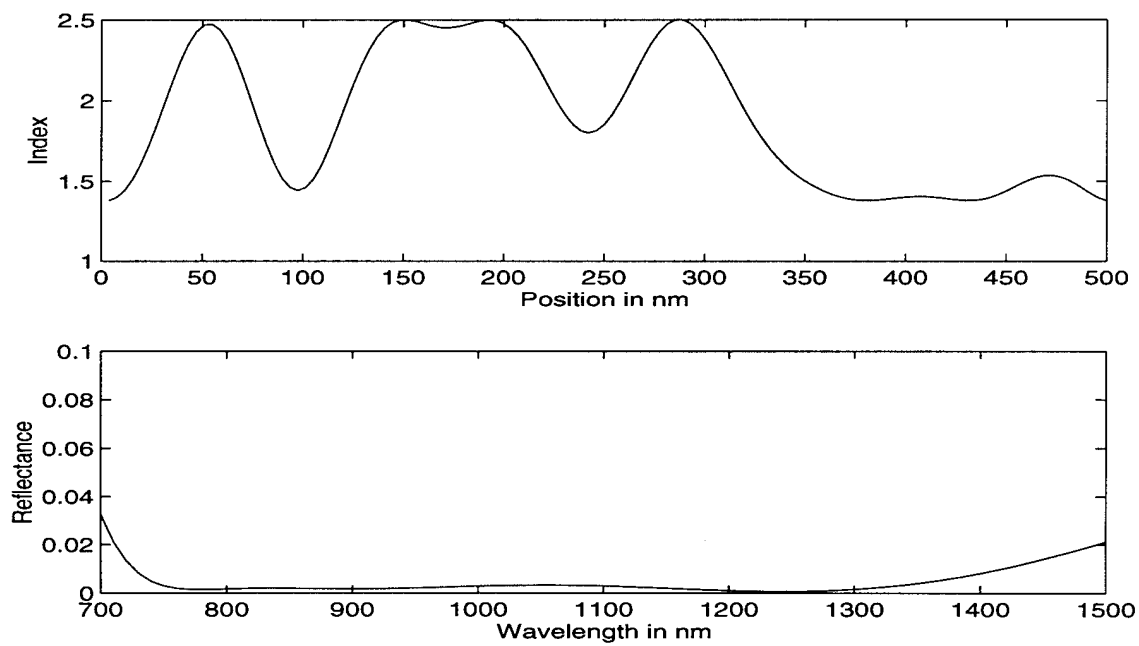
Figure 4.2: Fourier series design for Nd:YAG AR coating for 500 nm thick film.

63
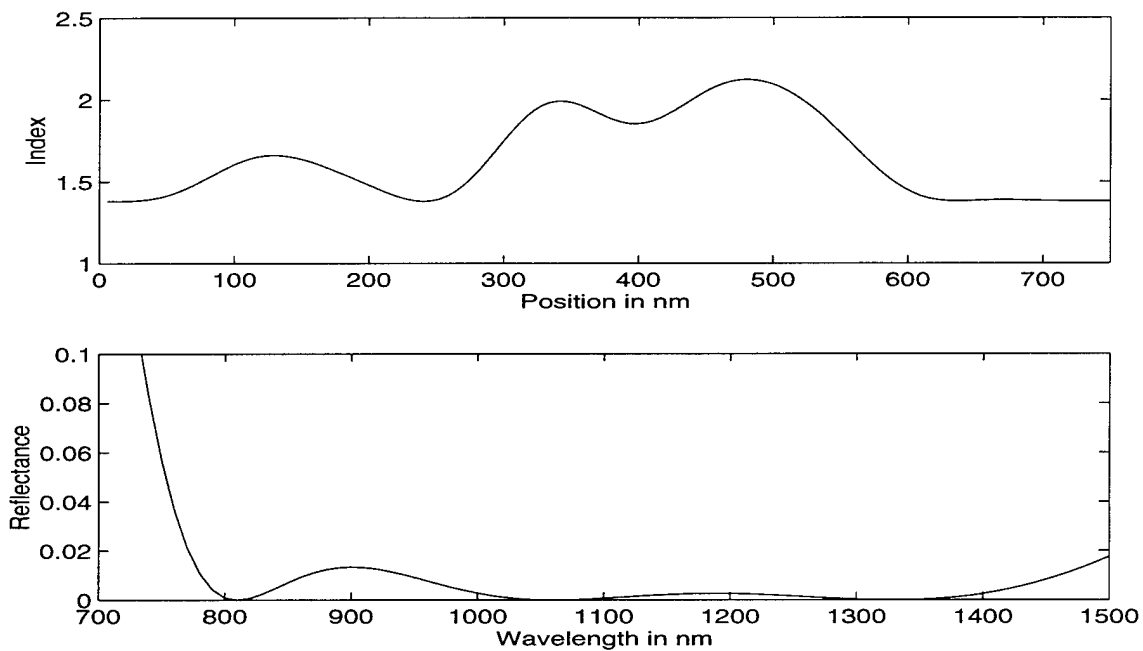
Figure 4.3: Fourier series design for Nd:YAG AR coating for 750 nm thick film.
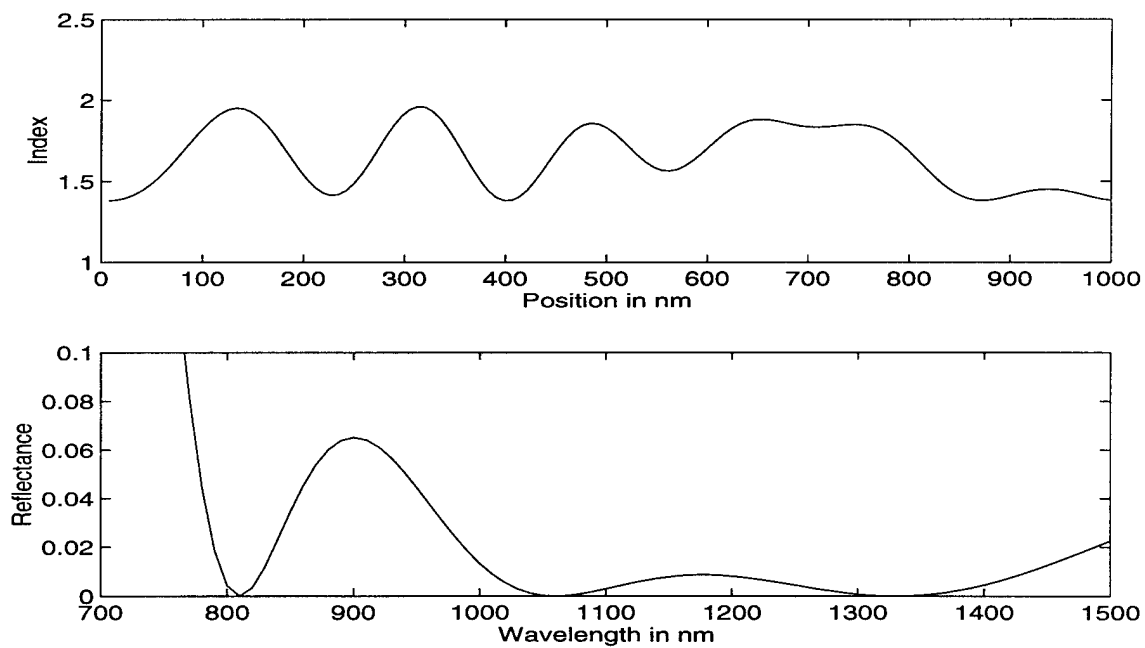
Figure 4.4: Fourier series design for Nd:YAG AR coating for 1000 nm thick film.

64

The next parameter to vary is the number of non-zero frequencies used in the design. For this series of designs, the thickness is fixed at 750 nm, since that is the smaller of the two "adequate" films found above. Figure 4.5 through Figure 4.8 illustrate the effects of varying the number of non-zero coefficients from 8 to 64. In all cases, the film thickness is fixed at 750 nm, and all other parameters are the same as the previous examples. Figure 4.5 shows the optimal design for a 750 nm film using only eight non-zero coefficients as design variables. The performance is quite poor (see Table 4-2), even though the thickness of the film has been shown previously to be sufficient. Figure 4.6 shows the optimal design for 16 coefficients. This is the same film as Figure 4.3 repeated for ease of reference. Figure 4.7 is the optimal film for 32 non-zero coefficient variables. Notice that the general shape of the index profile for both the 16 and 32 coefficient cases is similar, although the 32 coefficient film exhibits more small variations, as expected. The film in Figure 4.8 is the optimal design for 64 coefficient variables. This film has very similar structure to the 32 coefficient film, but many of the small oscillations have been reduced. The errors reported in Table 4-2 for the 16, 32, and 64 coefficient films are all below the numerical conversion criteria, that is, all are equally valid solutions. This example indicates that there is an optimal choice for the number of frequencies to use in the optimization. Too few frequencies yields a poor solution, and too many increases the computation time without significantly affecting the results.

Table 4-2: Design results for Fourier series Nd:YAG AR coatings with various number of design coefficients.

Wavelength (nm)	8 coefficients	16 coefficients	32 coefficients	64 coefficients
810	0.033842	3.034×10^{-8}	3.226×10^{-9}	1.044×10^{-8}
1060	0.041118	3.270×10^{-8}	5.564×10^{-9}	5.121×10^{-8}
1330	0.000179	5.603×10^{-8}	6.432×10^{-9}	4.772×10^{-8}
Total RMS Error	0.053254	7.162×10^{-8}	9.096×10^{-9}	7.077×10^{-8}
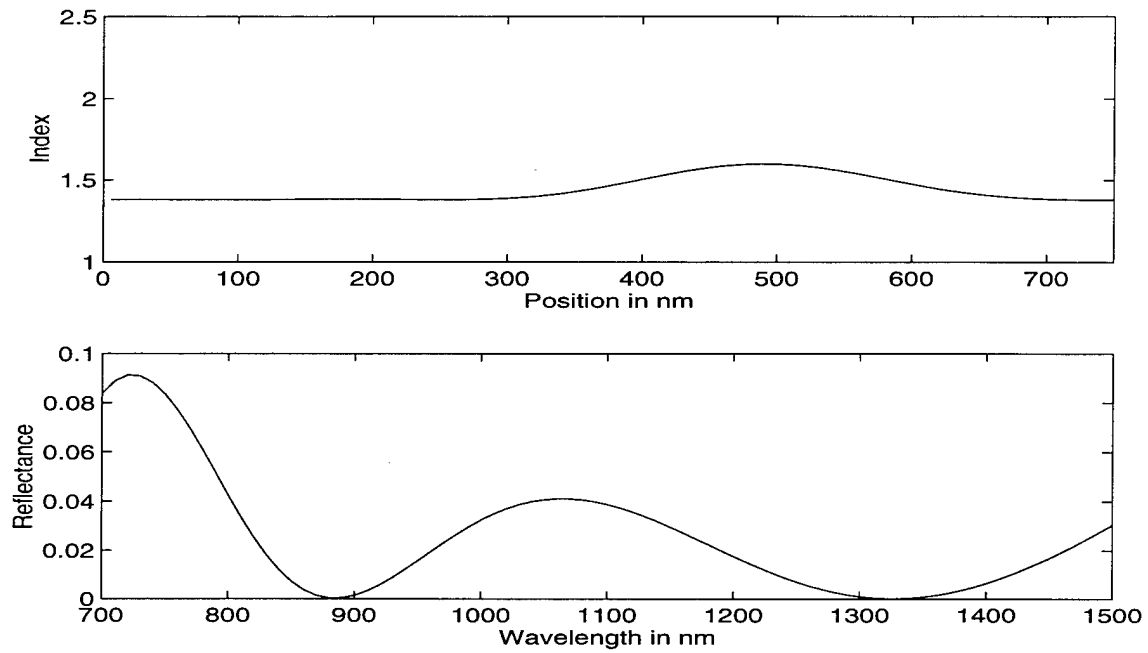
Figure 4.5: Fourier series design for Nd:YAG AR coating with 8 variable coefficients.
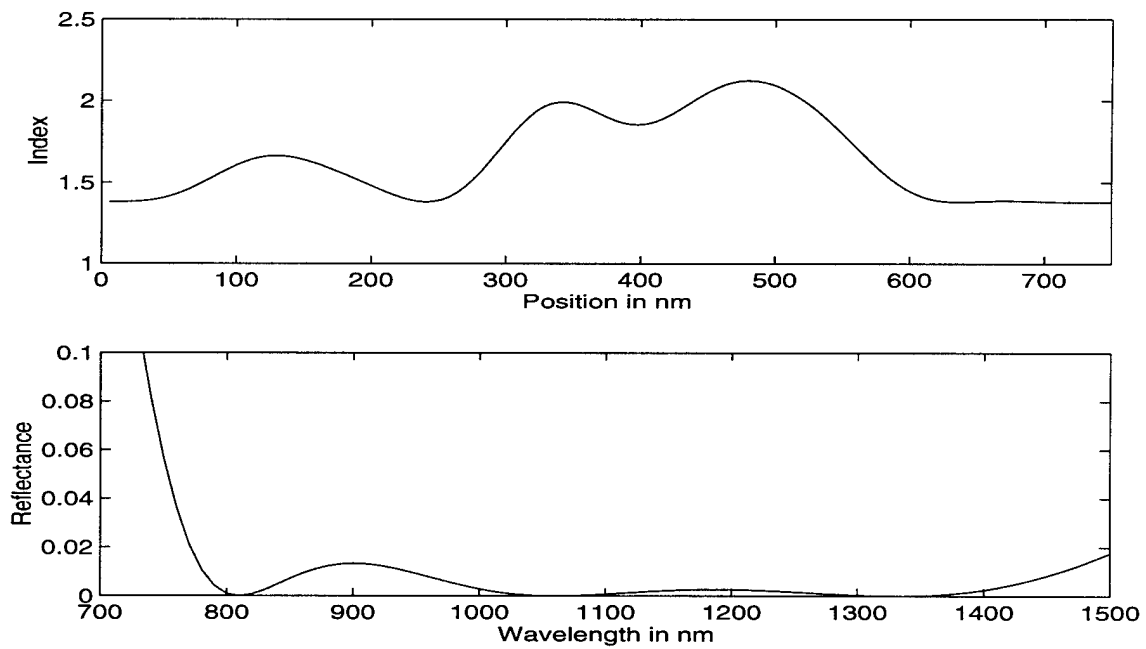
Figure 4.6: Fourier series design for Nd:YAG AR coating with 16 variable coefficients.

66
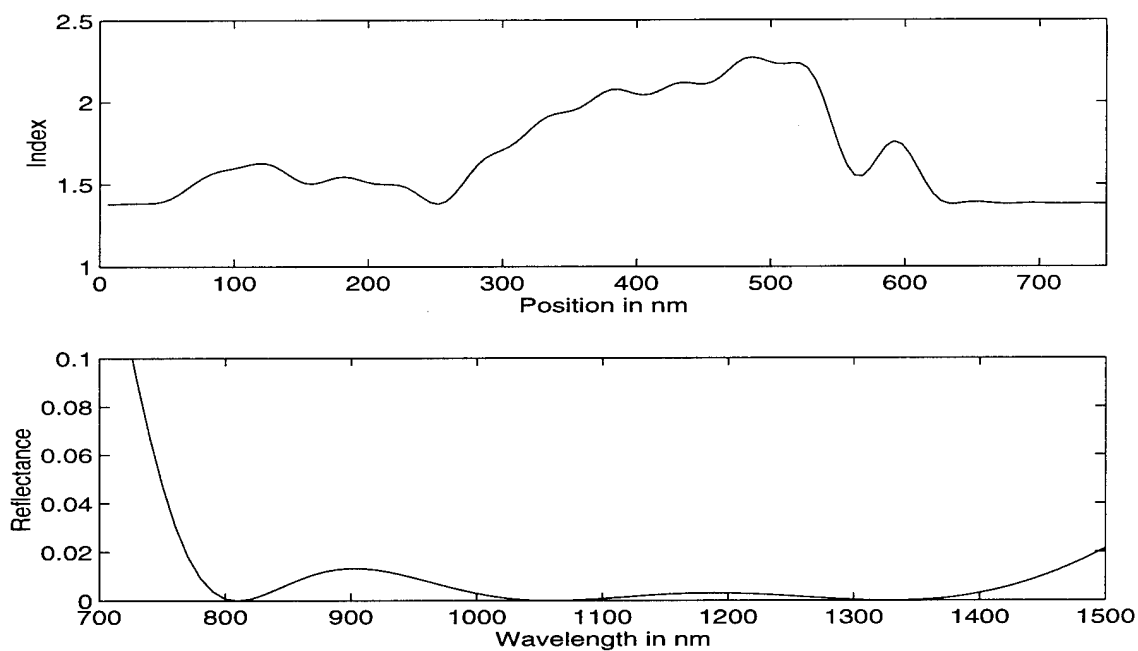
Figure 4.7: Fourier series design for Nd:YAG AR coating with 32 variable coefficients.
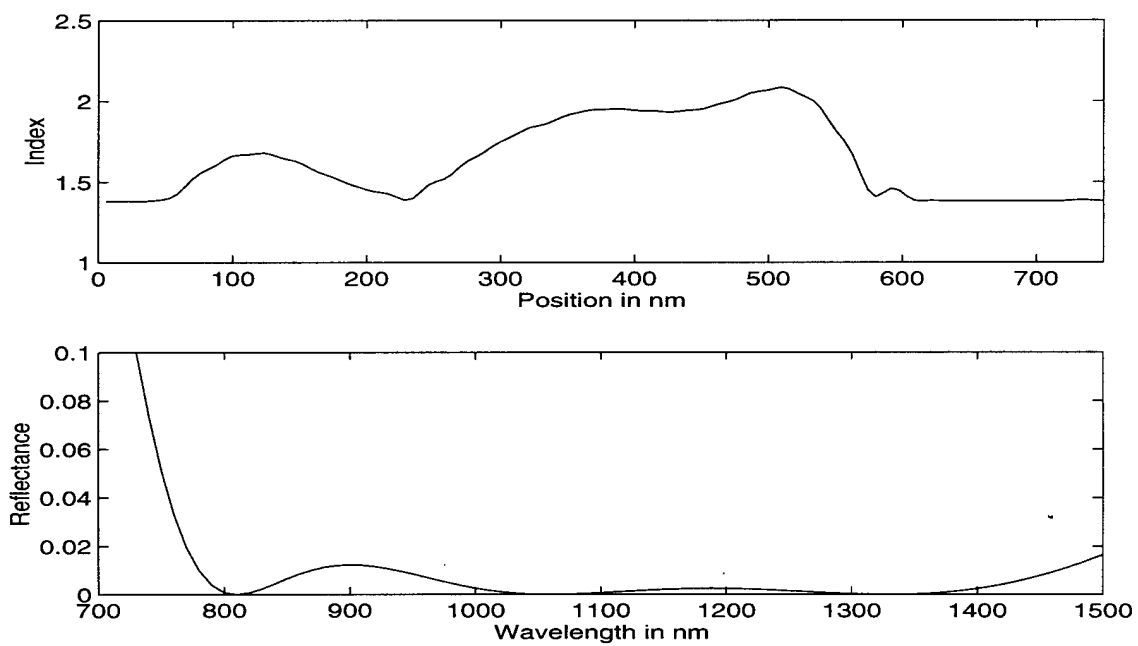
Figure 4.8: Fourier series design for Nd:YAG AR coating with 64 variable coefficients.

67

The 64 variable coefficient film in the example above was also designed using a different initial film to illustrate the effect of the input on the final design. The initial film for this example was a 750 nm film using 64 variables with an index profile given by

$$N(z) = 1.59 + 0.2\sin\left(\frac{40\pi z}{750}\right), \qquad z \in (0,750) \qquad (4.7)$$

This film is a sinusoidal oscillation about an average index of 1.59 and a total of 20 periods over the thickness of the film. The corresponding input to the numerical design is an index profile with two of the 64 coefficients having non-zero initial values. The result of this design is shown in Figure 4.9. Notice that the resulting film is quite different from the previous 64 variable design using a constant initial film shown in Figure 4.8, and the initial oscillation is still present in the final film. The performance of this film is very similar to that of the other 64 coefficient design with a constant initial film (see Table 4-3). The values in the table are all below the numerical termination criteria, indicating that both solutions are equally good. This illustrates both the effect of a poor initial film choice as well as the impact of too many variables in the design space. The difference between the two designs also serves to emphasize that this numerical optimal design method finds a *local* minimum solution, and not necessarily the global minimum. The true optimal design is achieved using the minimum number of variables and the minimum amount of film for the problem. The challenge for the thin film engineer lies in making good estimates of these parameters to use in this optimal design tool.

Table 4-3: Design results for Fourier series Nd:YAG AR coatings with 64 design coefficients for different initial films.

Wavelength (nm)	Sinusoidal Initial film	Constant Initial film
810	4.81×10^{-8}	1.044×10^{-8}
1060	2.13×10^{-9}	5.121×10^{-8}
1330	6.56×10^{-9}	4.772×10^{-8}
Total RMS Error	4.86×10^{-8}	7.077×10^{-8}
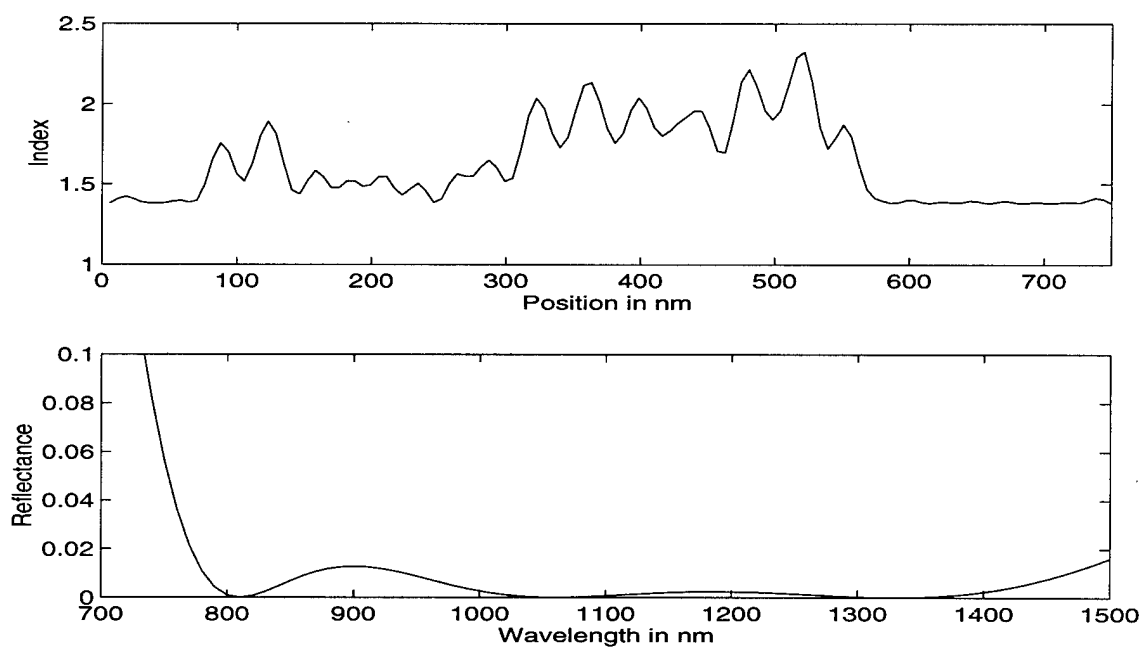
Figure 4.9: Illustration of effect of initial film on Fourier series design of Nd:YAG AR coating using 64 coefficients.

4.2.3. Dichroic Mirror for Nd:YAG Laser

This second example is a design analogous to the external cavity mirrors of the Nd:YAG laser. In this case, the mirror must allow the pump laser at 810 nm to pass through the mirror, but reflect the other two laser wavelengths of 1.06 μm and 1.33 μm. The designs for this coating must be much thicker than the previous anti-reflection designs. This is due to the high reflectivity required at 1.06 μm and 1.33 μm. The design parameters for this problem can be estimated by considering a basic quarter wave stack solution to the design problem.

As was mentioned in Chapter 2, a series of alternating high and low index layers produces a high reflectance band centered on a desired wavelength λ if each layer is one quarter of this wavelength in optical thickness. The reflectance of the film depends on the difference between the high and low index and the number of stack pairs in the film. The quarter wave stack design can be used to determine the design parameters for this Fourier series design by approximating an alternating quarter wave stack pair by a single period of a sinusoid. For this example, the desired high reflectance is centered on a wavelength of 1.2 μm. Using the average index of about 2.0, the quarter wave physical thickness is about 150 nm, which means a sine with a period of 300 nm must be available in the optimization. To decide how many frequencies (and thus coefficients) to use in the optimization, the total thickness of the film must be selected. A film thickness of 3 μm would give about 10 quarter wave pairs, which should provide fairly good reflectance. Dividing this 3 μm film thickness by the 300 nm period of one sinusoid indicates that at least 10 non-zero frequencies are required in the optimization. So choose 32 non-zero coefficients to use as variables, which corresponds to 16 frequencies. The requirement to sample the film every 5 nm and the desire for the number of samples to be a power of two (for Fourier transform efficiency) leads to a choice of 512 samples.

As before, a flat index profile is used as the seed. Figure 4.10 shows the index of refraction and reflectivity for a 3 micron total film thickness. The resultant reflectivity at the three design points is not very good (see Table 4-4). The second design increases the total thickness to 4 μm. This requires a change in the total number of samples to 1024 to maintain the quality of the reflectance calculation. The increase in film thickness changes the estimate of the number of frequencies required calculated above to 14, so the number of coefficients to use as variables remains the same at 32. Figure 4.11 shows the index of refraction profile and resultant reflectivity for a 4 micron design. This design is much better at the design wavelengths (see Table 4-4), and has a root mean squared error of 0.8%. Also note that the index profile shows a fundamental periodicity corresponding to the 300 nm period estimated above.

Table 4-4: Design values for Fourier series Dichroic mirror

Wavelength (nm)	Desired Reflectance	Reflectance for 3 micron film	Reflectance for 4 micron film
810	0.0	0.047	0.003
1060	1.0	0.9551	0.9944
1330	1.0	0.9673	0.9946
Total Error		0.0727	0.00833

.
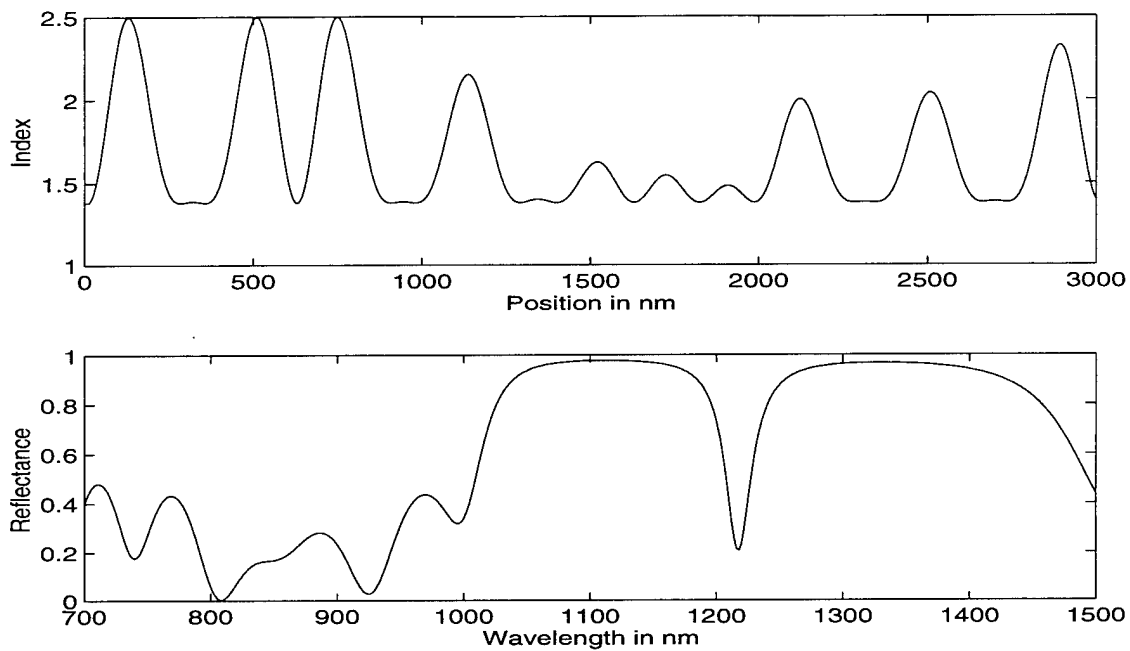
Figure 4.10: Dichroic Mirror Fourier design for 3 micron thick film.
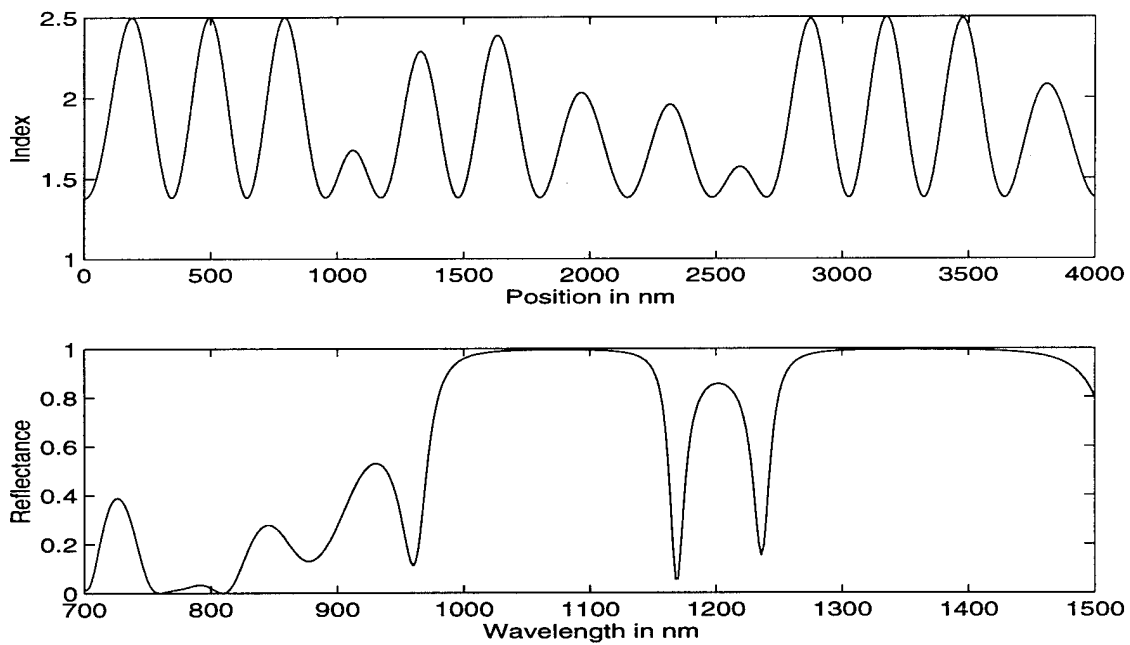
Figure 4.11: Dichroic Mirror Fourier design for 4 micron thick film.

4.2.4. Ti:Sapphire Bandpass film

A third example using the Fourier design method is illustrated in the design of a broadband, high-reflectance rugate mirror over the wavelength range available to Ti:Sapphire lasers. This laser emits at wavelengths between 660 nm and 1.18 μm [33:480]. For this example, specify a reflectivity of >99% between 700 and 1100 nm. Simulating a CeF_3:ZnS rugate filter, let n=1.89 at the substrate and permit an index variation from 1.6 to 2.2 [21:61]. This example is similar to the SWIFT design in Chapter 3, without the additional specification of zero reflectance outside the desired wavelength range. The desired reflectance here was specified to be 1.0 for 41 wavelengths between 700 and 1100 nm (every 10 nm).

The design parameters for this problem can be estimated by considering a quarter wave stack design, as was done above for the dichroic mirror example. This is a very broad range of wavelengths to cover with a singe notch design. However, the notch design considerations can help in making choices on the number of coefficients needed for the design. A conservative estimate can be made using the minimum wavelength in this design, 700 nm, and the maximum index available, 2.2. These values yield a physical thickness for a quarter wave of 80 nm, so a minimum periodicity of twice this number, or 160 nm, is required. A 15 micron film would allow for 93 equivalent quarter-wave periods, and require 93 non-zero frequencies in the optimization. So choose to use 256 non-zero coefficients as variables (corresponding to 128 frequencies) and 2048 samples to model the film, which insures film sample spacing less than 5 nm.

Figure 4.12 illustrates the index of refraction and reflectance of the optimal film. The root sum squared error over the 41 design wavelengths is 1.287. As the figure shows, the reflectance profile is not flat (as desired). To achieve better performance, additional wavelength requirements must be specified, and the total thickness of the film would probably need to be increased as well. Unfortunately, this example as presented requires

almost two days of computation time (on a Sun Microsystems Sparc 20). This example illustrates a weakness in this design method; specifically the extensive computation required for broadband, high reflectivity problems. The high desired reflectance requires both a large total thickness and a large number of variable coefficients. A large number of samples for reflectance calculation are also required. While some speed enhancements of the design program are possible, the basic size of the design problem is too large for efficient application of the generalized Fourier series design method.
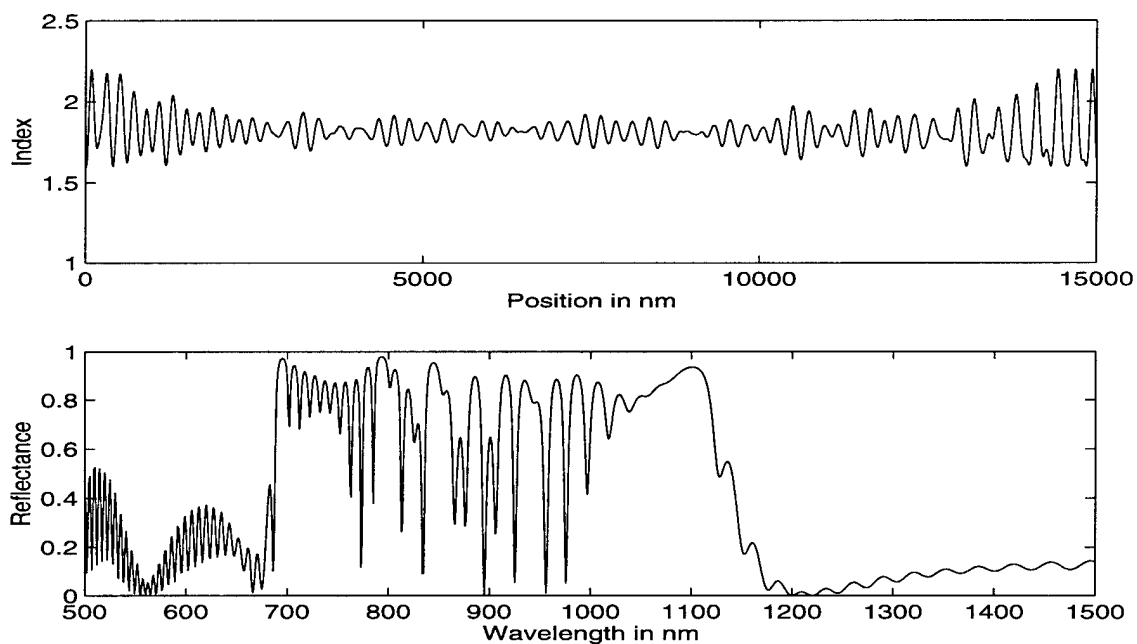
Figure 4.12: Fourier series optimal design for Ti:Sapphire broadband mirror.

4.3. Optimal Design using Wavelets

The second basis system to be considered is a wavelet basis. The wavelet theory used in this analysis is presented in Appendix C. The construction of an index profile on the interval from wavelet coefficients is described briefly below. The same example problems solved by the Fourier series method above are solved using a wavelet basis.

4.3.1. Wavelet Basis

The wavelet decomposition and reconstruction theory presented in Appendix C describes a method for analyzing a function using wavelets. This idea can also be used to synthesize a function with certain desired characteristics. The wavelet representation of a function is similar to a Fourier series representation, except the basis elements are shifted and scaled versions of the "mother wavelet" instead of sines and cosines of different frequencies. In either case, the function is completely specified by the coefficients of the series. Given an orthogonal "mother wavelet", ψ, and a "scaling function", ϕ, one can write any function $f \in L^2(\Re)$ as

$$f(x) = \sum_n c_{M,n} \phi_{M,n}(x) + \sum_{m \leq M} \sum_n d_{m,n} \psi_{m,n}(x)$$

where: 								(4.8)

$$\psi_{m,n}(x) \equiv 2^{-m/2} \psi\left(2^{-m} x - n\right)$$

In practice, a continuous function f is represented by a sequence of sampled values. This sequence may be decomposed into a sequence of "approximation" coefficients, $c_{m,n}$, and a sequence of "detail" coefficients, $d_{m,n}$. The approximation coefficients are associated with a scaling function, $\phi(x)$, and the detail coefficients are associated with the wavelet function, $\psi(x)$. In this discrete case, the scaling function $\phi(x)$ can be represented by a sequence of values denoted by $h(n)$ or h_n, and the wavelet

function $\psi(x)$ can also be represented by a sequence of values, denoted by $g(n)$ or g_n (see Appendix C for details). The decomposition algorithm in this multi-resolution analysis requires only the filters h and g to define the wavelets to be used. In fact, the g filter can be derived from the h filter, so only one sequence is needed to completely determine the wavelets. This relationship between h and g is

$$g(n) = (-1)^n h(n-1) \tag{4.9}$$

The coefficients needed to represent the sampled function are obtained using a recursion relation. This relation for the approximation coefficients is

$$c_{m,n} = \sum_{\ell \in Z} c_{m-1,\ell} h(\ell - 2n) \tag{4.10}$$

A similar relation exists for the detail coefficients, $d_{m,n}$:

$$d_{m,n} = \sum_{\ell \in Z} c_{m-1,\ell} g(\ell - 2n) \tag{4.11}$$

Using the two recursion relations for $c_{m,n}$ and $d_{m,n}$, any function can be decomposed into its approximation and detail coefficients from an initial sequence c_0. The recursive relation for reconstruction is

$$c_{m-1,k} = \sum_{n \in Z} c_{m,n} h(k - 2n) + \sum_{n \in Z} d_{m,n} g(k - 2n) \tag{4.12}$$

The design of a film using wavelet coefficients is performed using only a few of these coefficients as variables in the numerical design. The film is created from the span of a few wavelet basis elements, then a large number of samples of the film is generated

using the reconstruction algorithm in Equation 4.12 and zeros for the values of the coefficients not used as variables in the design. This is similar to padding a sequence of numbers prior to performing a numerical Fourier transform. The crux of this point is the relationship between the number of samples of the film needed to adequately model its optical properties, and the number of wavelet coefficients needed to adequately model the film. While a large number of sample points may be needed to determine the reflectivity of the film (using the matrix methods as described in Chapter 2), the film can be specified with relatively few wavelet coefficients.

The MATLAB™ programs used to perform the optimal designs in the wavelet basis system are included in Appendix D. Several functions are combined to make up the program. The script **WAVRUN.M** is the main function, which sets the initial parameters and calls the optimizer. As before, the key element of this implementation is the MATLAB™ optimization toolbox function called **CONSTR.M** [26], which implements the sequential quadratic programming algorithm described in Appendix B. The details of this function were presented in the previous section.

The evaluation function **WAVFUN.M** contains the conversion between coefficients of the wavelet expansion and samples of the index. It also calculates the reflectance of the film at each design wavelength, and the merit function and constraint values used in the optimization. The inputs to this function are: the current variables values, the number of sample points, the desired reflectance at each design wavelength, the substrate index, the film thickness in microns (or nanometers), the array of design wavenumbers k in inverse microns (or inverse nanometers), and the array of sample points x in microns (or nanometers), and the scaling function filter h. The units selected for length must be consistent. The outputs of the function are the values of the merit function and the constraints. The reconstruction of the index of refraction from the wavelet coefficients is accomplished by the function **UP1D.M.** The decomposition of a

function is accomplished using the program **DOWN1D.M**, which is also included. The reflectance is calculated by the "C" language program **REFLECT.C** as before.

4.3.2. Nd:YAG Laser AR Coating

The first example of optimal design of a gradient index film is to create an anti-reflection coating for the gain medium of a neodymium : yttrium aluminum garnet (Nd:YAG) laser. Again the goal is to eliminate reflections at the two primary laser wavelengths of 1.06 μm and 1.33 μm, as well as the pump laser wavelength of 810 nm. The YAG material has an index of refraction of 1.816, which means the uncoated surface has a reflectance of 8.4%. As in the Fourier series design above, the parameters for this design are total film thickness, number of samples, available index range, and number of wavelet coefficients to use. The thicknesses used in this design are the same as in the Fourier example: 300 nm, 500 nm, 750 nm, and 1000 nm. The number of samples of the film to use is determined by the total thickness and the reflectance calculation requirement of approximately 5 nm per sample. The number of samples for this example case is 128. The index range is based on a material system of MgF$_2$ and ZnSe, which have indices of refraction of 1.38 and 2.5 respectively. The wavelet used is Daubechies' 8-tap wavelet [7]. The filter coefficients for this and other Daubechies wavelet are tabulated in Appendix C. The initial variable values for the first examples are for a flat film consisting entirely of the substrate material. This means that no discontinuity is introduced by periodization of the film.

The first parameter to vary is total thickness. The number of non-zero wavelet coefficients for these designs is fixed at 16, for comparison with the Fourier series based designs of the section 4.2.2. Figure 4.13 through Figure 4.16 below show the film designs and resultant reflectivity profiles for total thicknesses of 300 nm, 500 nm, 750 nm, and 1000 nm. The reflectances at the design wavelengths for these films are given in Table 4-5. As with the Fourier series designs for these thicknesses, the optimal designs for both

78

the 300 nm and 500 nm thicknesses reduce the reflectance significantly from the nominal 8.4%, but neither is acceptable for use inside a laser cavity. In the wavelet design case, the 500 nm film has reflectances down to about one hundredth of a percent, as compared to the 500 nm Fourier film design, which had reflectances in the five hundredths of a percent. In at least this example, the wavelet design produced a superior film by a factor of about 3 for the given thickness. The designs for 750 nm and 1000 nm both perform as well as the Fourier designs for these thicknesses did, as indicated by the reflectances in Table 4-5. As before, this demonstrates the importance of the total thickness as a parameter in the design.

Table 4-5: Design results for various thicknesses of Wavelet Nd:YAG AR coatings.

Wavelength (nm)	300 nm film	500 nm film	750 nm film	1000 nm film
810	3.138×10^{-4}	5.200×10^{-4}	3.916×10^{-9}	2.566×10^{-9}
1060	5.560×10^{-4}	9.440×10^{-4}	3.911×10^{-10}	6.612×10^{-9}
1330	6.238×10^{-4}	9.398×10^{-4}	9.863×10^{-10}	7.891×10^{-9}
Total Error	8.926×10^{-4}	1.430×10^{-3}	4.057×10^{-9}	1.061×10^{-8}
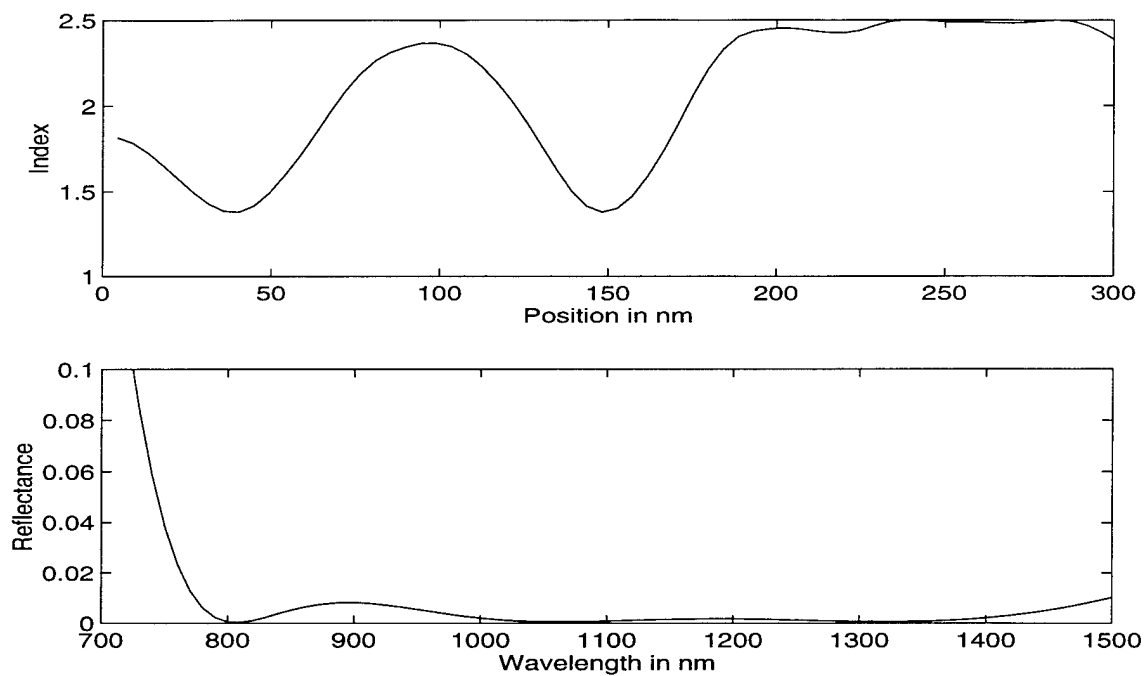
Figure 4.13: Wavelet design of Nd:YAG AR coating for 300 nm thick film.
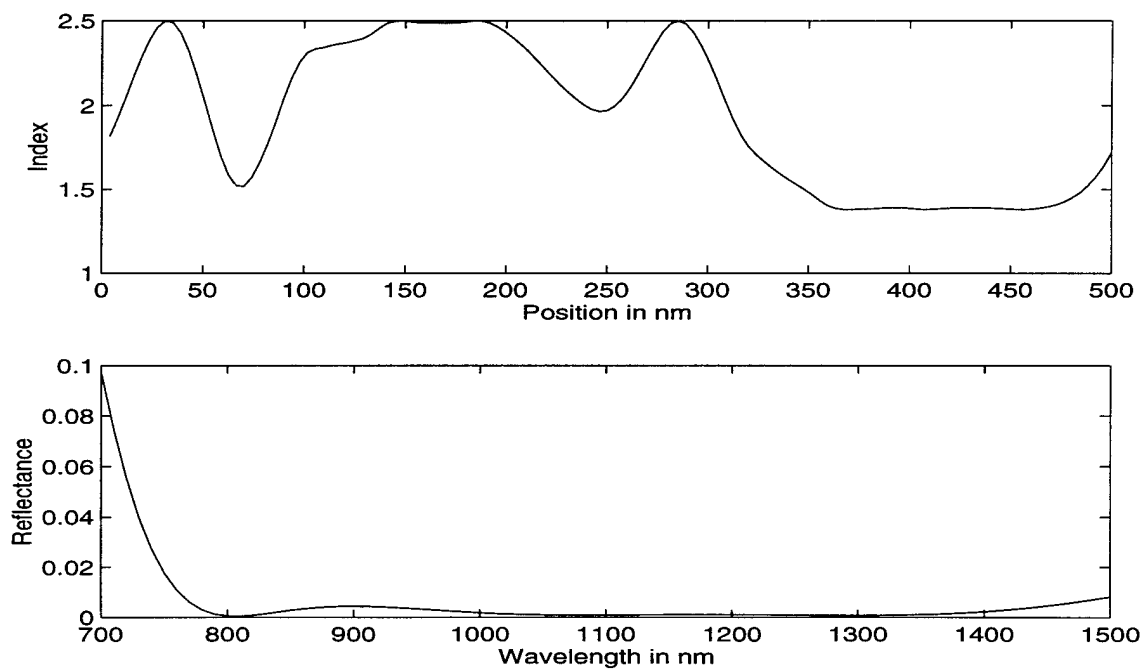
Figure 4.14: Wavelet design of Nd:YAG AR coating for 500 nm thick film.

80
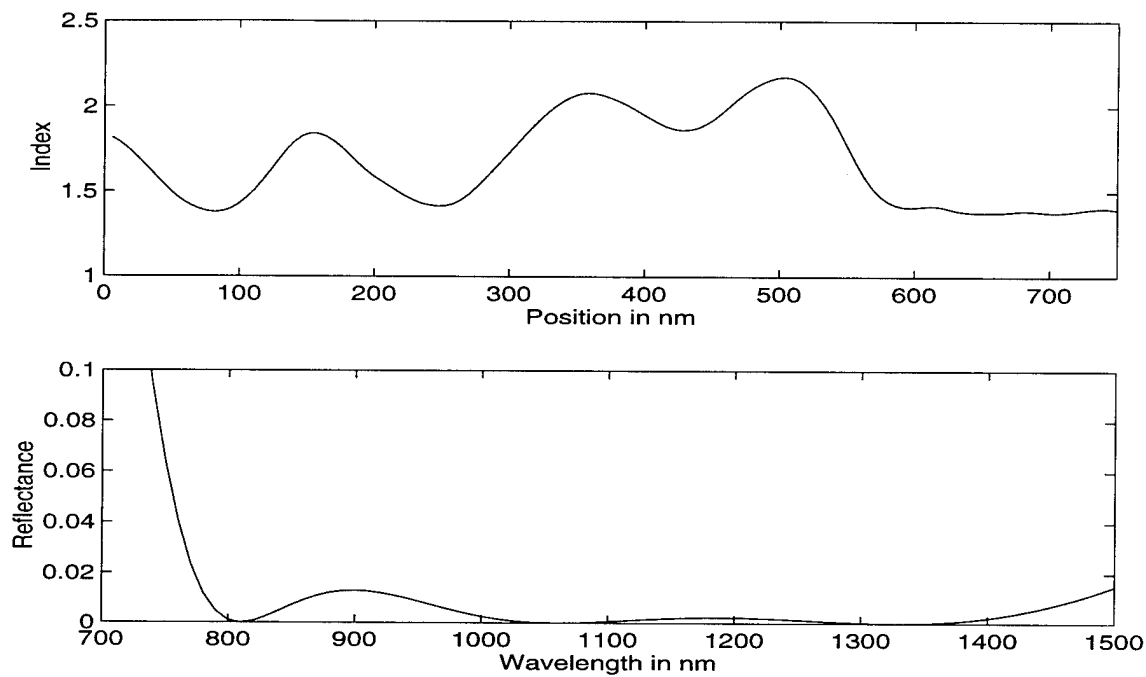
Figure 4.15: Wavelet design of Nd:YAG AR coating for 750 nm thick film.
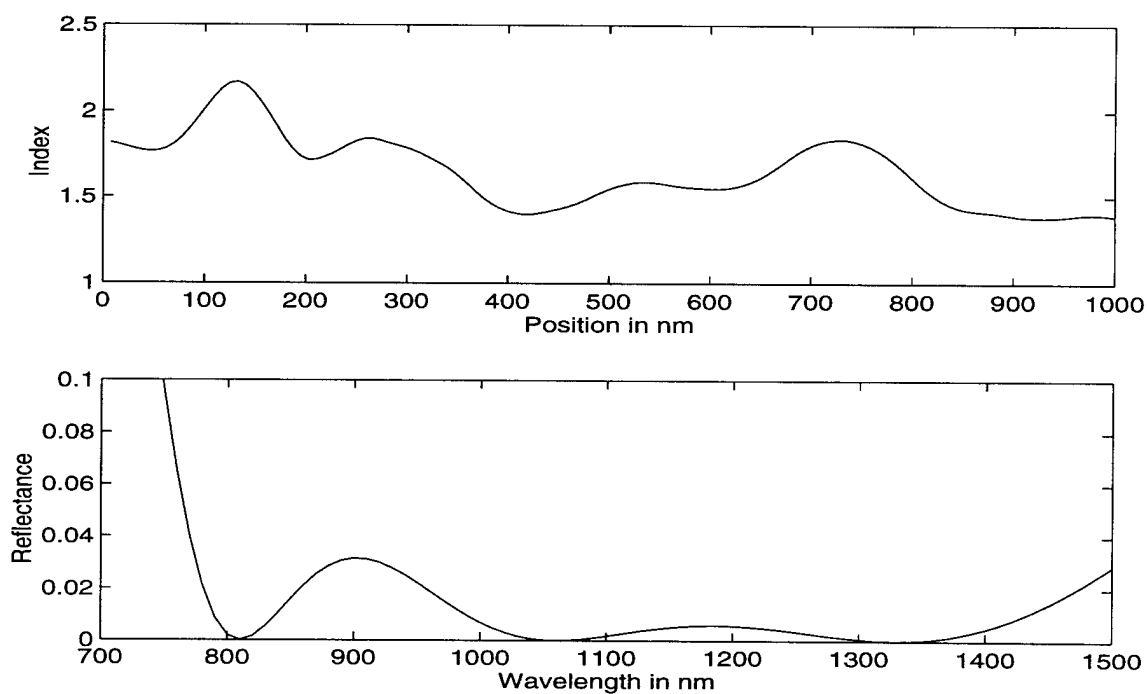
Figure 4.16: Wavelet design of Nd:YAG AR coating for 1000 nm thick film.

The other main design parameter is the number of non-zero wavelet coefficients to use as variables in the optimization. As in the Fourier series examples above, the total thickness is fixed at 750 nm, and the number of samples is 128. The number of wavelet coefficients must be a power of two in order to span the entire film.

Figure 4.17 through Figure 4.20 illustrates the effects of varying the number of non-zero coefficients to use in the design process. Figure 4.17 shows the optimal design for a 750 nm film using only four non-zero wavelet coefficients as design variables. The performance is quite poor (see Table 4-6), even though the thickness of the film has been shown previously to be sufficient. Figure 4.19 is the optimal film for 16 wavelet coefficients. There is a marked improvement in the performance of the film. Figure 4.19 shows the result of the 16 wavelet coefficient design, repeated from Figure 4.15 above for ease of reference. Finally, Figure 4.20 is the 32 wavelet coefficient design. Notice as the number of coefficients used increases, smaller scale feature become evident in the index designs. It is obvious that the additional small variations have little effect on the film performance. However, increasing the number of design variables also increases the computation time for the optimal design. It is therefore important to choose the smallest number of design coefficients that can still achieve the desired performance.

Table 4-6: Design results for Wavelet Nd:YAG AR coatings with various number of design coefficients.

Wavelength (nm)	4 coefficients	8 coefficients	16 coefficients	32 coefficients
810	1.610×10^{-2}	3.462×10^{-4}	3.916×10^{-9}	2.507×10^{-9}
1060	0.531×10^{-2}	5.715×10^{-4}	3.911×10^{-10}	8.336×10^{-9}
1330	1.380×10^{-2}	4.576×10^{-4}	9.863×10^{-10}	5.144×10^{-9}
Total Error	2.186×10^{-2}	8.098×10^{-4}	4.057×10^{-9}	1.011×10^{-8}
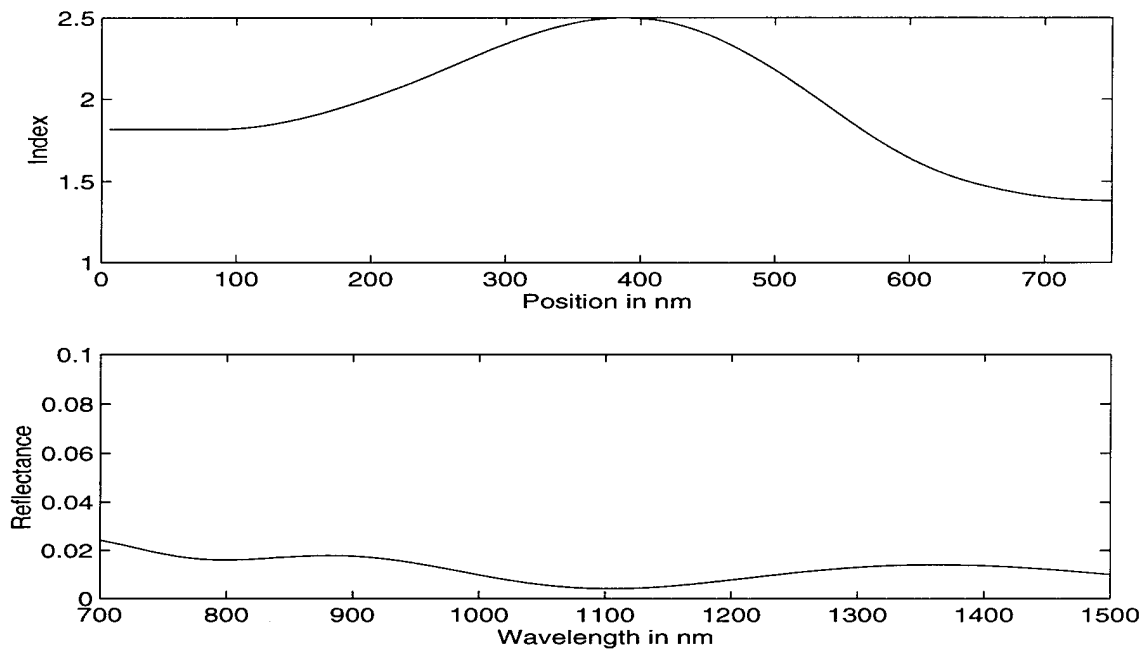
Figure 4.17: Wavelet design of Nd:YAG AR coating using 4 coefficients.
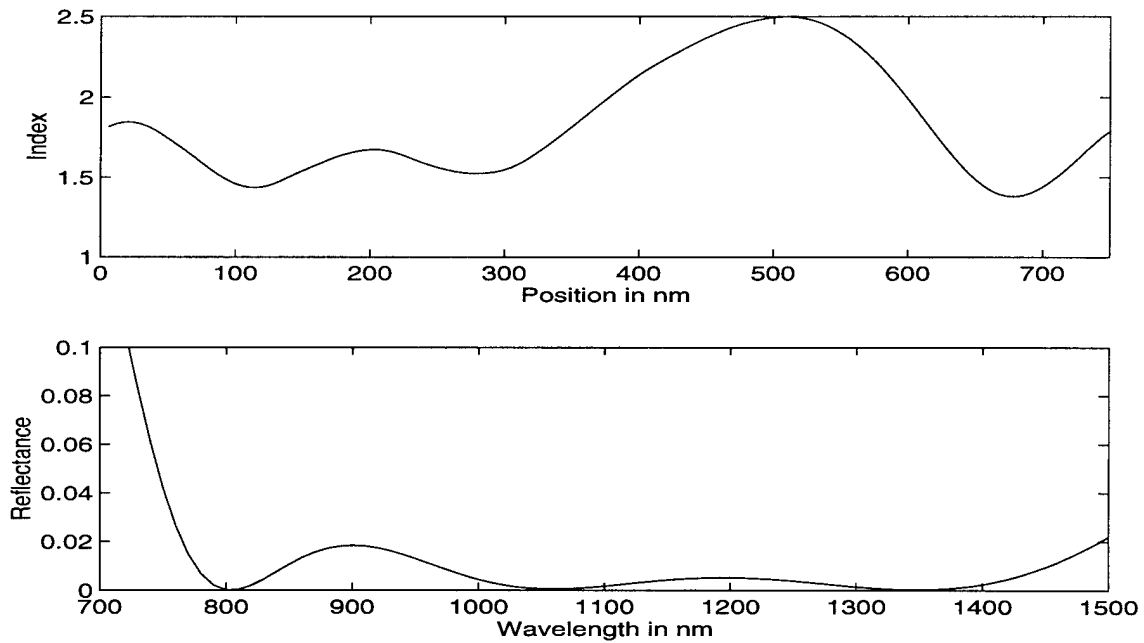
Figure 4.18: Wavelet design of Nd:YAG AR coating using 8 coefficients.

83
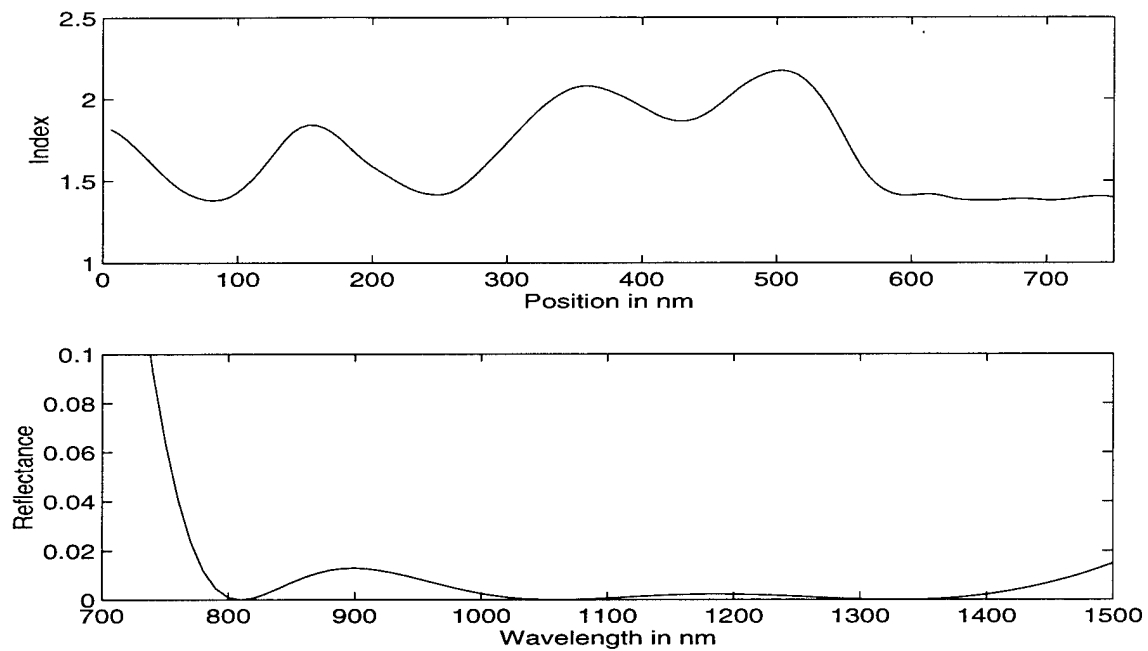
Figure 4.19: Wavelet design of Nd:YAG AR coating using 16 coefficients.
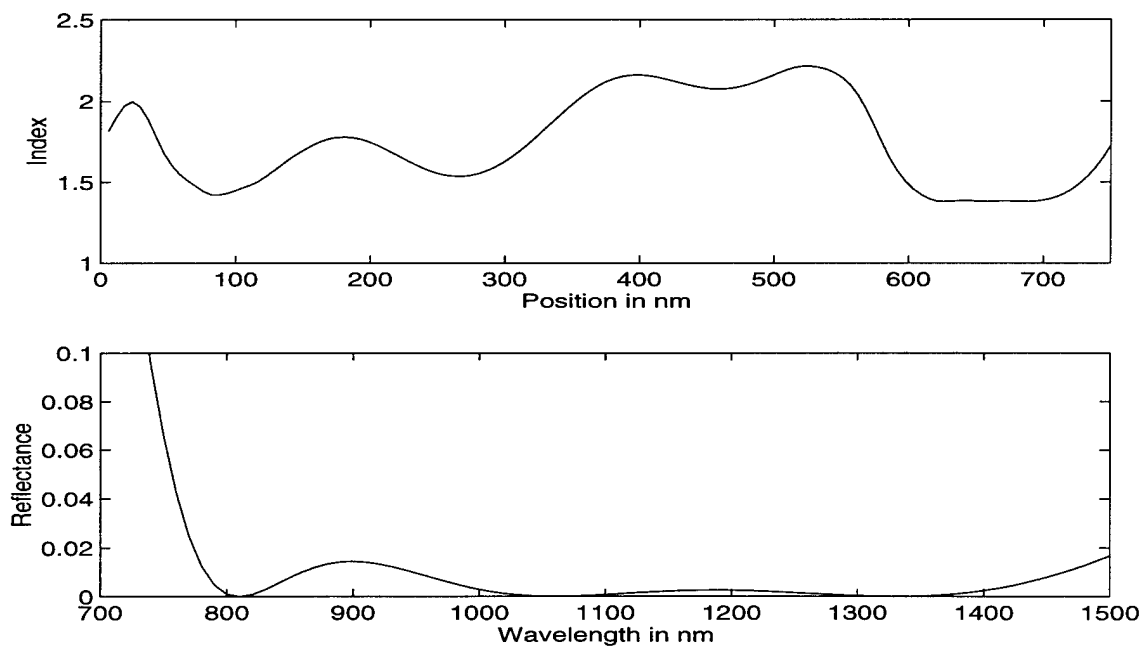
Figure 4.20: Wavelet design of Nd:YAG AR coating using 32 coefficients.

84

4.3.3. Dichroic Mirror

This second example is a design for the external cavity mirrors of the Nd:YAG laser. In this case, the mirror must allow the pump laser (at 810 nm) to pass through the mirror, but reflect the other two laser wavelengths of 1.06 μm and 1.33 μm. The design parameters for this problem can be estimated by considering a basic quarter wave stack solution to the design problem, as was done in the Fourier series example in Section 4.2.3.

As was mentioned in Chapter 2, a series of alternating high and low index layers produces a high reflectance band centered on a desired wavelength λ if each layer is one quarter of this wavelength in optical thickness. The reflectance of the film depends on the difference between the high and low index and the number of stack pairs in the film. As was done in Section 4.2.3, the quarter wave stack design can be used to determine the design parameters for the wavelet based design by insuring the scale corresponding to the minimum quarter wave thickness is included in the design variables. As in the previous Fourier series dichroic mirror example, the desired high reflectance is centered on a wavelength of 1.2 microns. Using the average index of about 2.0, the quarter wave physical thickness is about 150 nm. To decide how many coefficients to use in the optimization, the total thickness of the film must be selected. A film thickness of 3 microns would give about 10 quarter wave pairs, which should provide fairly good reflectance. Dividing this 3 micron film thickness by the 150 nm quarter wave thickness indicates that the optimization must be able to model at least 20 peaks (which is equivalent to the 10 frequencies found in the Fourier series example). The mother wavelets used here include several oscillations over their support. The wavelet detail coefficients must be grouped in powers of two to insure the entire film is spanned by the basis elements used. This means that the detail coefficients which break the film into 16 sections must be included in the design. All detail coefficients of broader scale should also be included. So for this example at least 32 non-zero coefficients are needed as

variables. The requirement to sample the film every 5 nm and the requirement for the number of samples to be a power of two leads to a choice of 512 samples for the film. The wavelet used in this design is Daubechies' 16-tap wavelet. This wavelet has 5 peaks and troughs over its support. The coefficients for this wavelet are tabulated in Appendix C.

Figure 4.21 shows a 3 micron thick wavelet design for this mirror. The design values for the 3 micron film are reported in Table 4-7. The 3 micron film has high reflectance values >98% and a low reflectance value at 810 nm of <0.01%. Figure 4.22 shows a 4 micron thick design for the same mirror. The additional film thickness allows a maximum reflectance value of >99% and keeps the low reflectance < 0.06%. Additional gains in the high reflectance can be achieved by further increases in the total thickness. The specific design values for the 4 micron film are also reported in Table 4-7. The total error is the square root of the sum of the squares of the differences between the desired reflectance and the reflectance of the design at each wavelength (recall Equation 4.3).

Table 4-7: Design values for Wavelet Dichroic mirror.

Wavelength (nm)	Desired Reflectance	Reflectance for 3 micron film	Reflectance for 4 micron film
810	0.0	0.00009	0.0006
1060	1.0	0.9811	0.9939
1330	1.0	0.9801	0.9910
Total Error		0.02744	0.01089
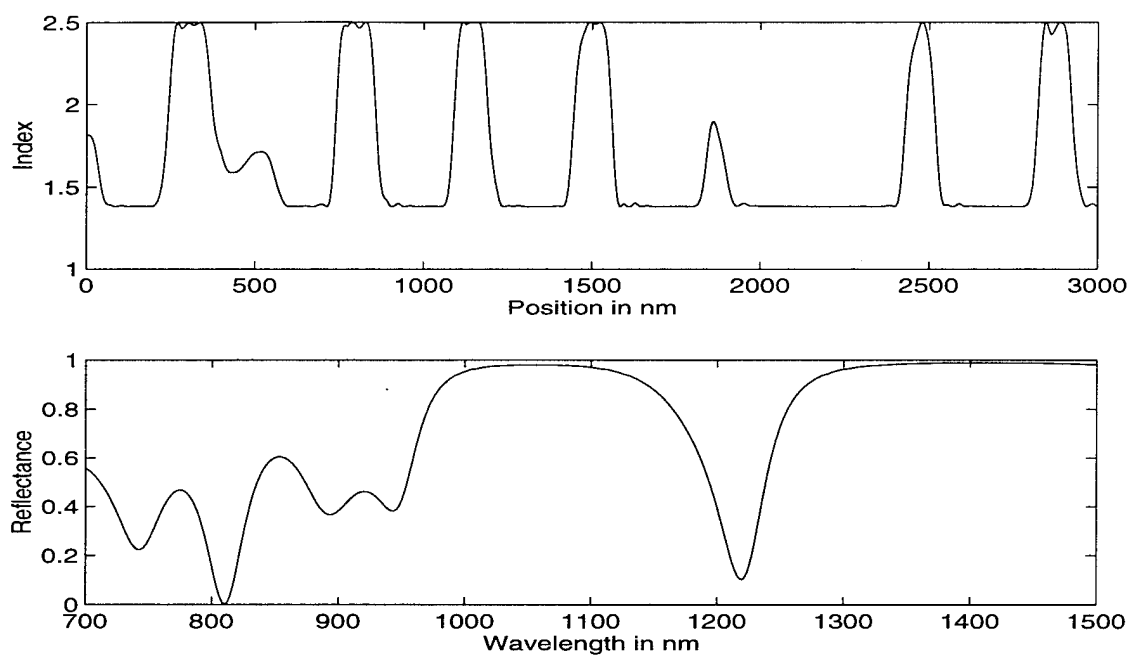
Figure 4.21: Wavelet design of 3 micron thick dichroic mirror.
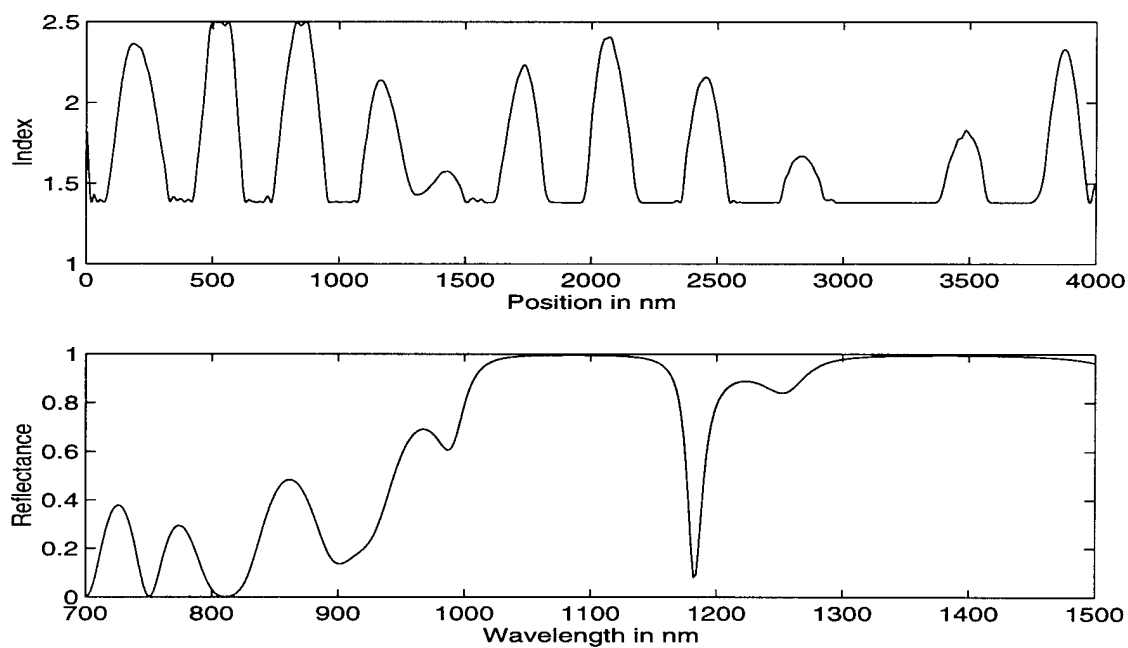
Figure 4.22: Wavelet design of 4 micron thick dichroic mirror.

87

4.3.4. Ti:Sapphire Bandpass film

A third example using the wavelet design method is illustrated in the design of a broadband, high-reflectance rugate mirror over the wavelength range available to Ti:Sapphire lasers. This example is the same as the Fourier series example in Section 4.2.4. This laser emits at wavelengths between 660 nm and 1.18 microns [33:480]. It is presently necessary to change the laser's optics in order to run it over its full operating range. For this example, specify a reflectivity of >99% between 700 and 1100 nm. Emulating a CeF_3:ZnS rugate filter, let n=1.89 at each boundary and permit an index variation from 1.6 to 2.2 [21:61]. This example is similar to the SWIFT design in Chapter 3, without the additional specification of zero reflectance outside the desired wavelength range. The desired reflectance here was specified to be 1.0 for 41 wavelengths between 700 and 1100 nm (every 10 nm). The design parameters used were a 10 micron total thickness, 2048 samples across the film, and 256 variable wavelet design coefficients, and a Daubechies' 16-tap wavelet. Figure 4.23 illustrates the index of refraction and reflectance of the optimal film. The root sum squared error over the 41 design wavelengths is 0.109. All design wavelengths have reflectances >96%, but as the figure shows, the reflectance profile is not flat (as desired). To smooth the reflectance profile, additional wavelength requirements must be specified. In addition, the total thickness of the film would probably need to be increased as well. Unfortunately, this example as presented requires almost two days of computation time (on a Sun Microsystems Sparc 20). As in the Fourier series case, this example illustrates a weakness in this design method, specifically the extensive computation required for broadband, high reflectivity problems. The use of the wavelet basis improves the performance somewhat over the Fourier series basis, but does not solve the problem. The high desired reflectance requires both a large total thickness and a large number of variable coefficients. A large number of samples for reflectance calculation is also required. While some speed enhancements of the design program are possible (such as: lower level

program coding, more efficient wavelet decomposition and reconstruction implementations, more efficient reflectance calculation), the basic size of the design problem is too large for efficient application of the wavelet design method.
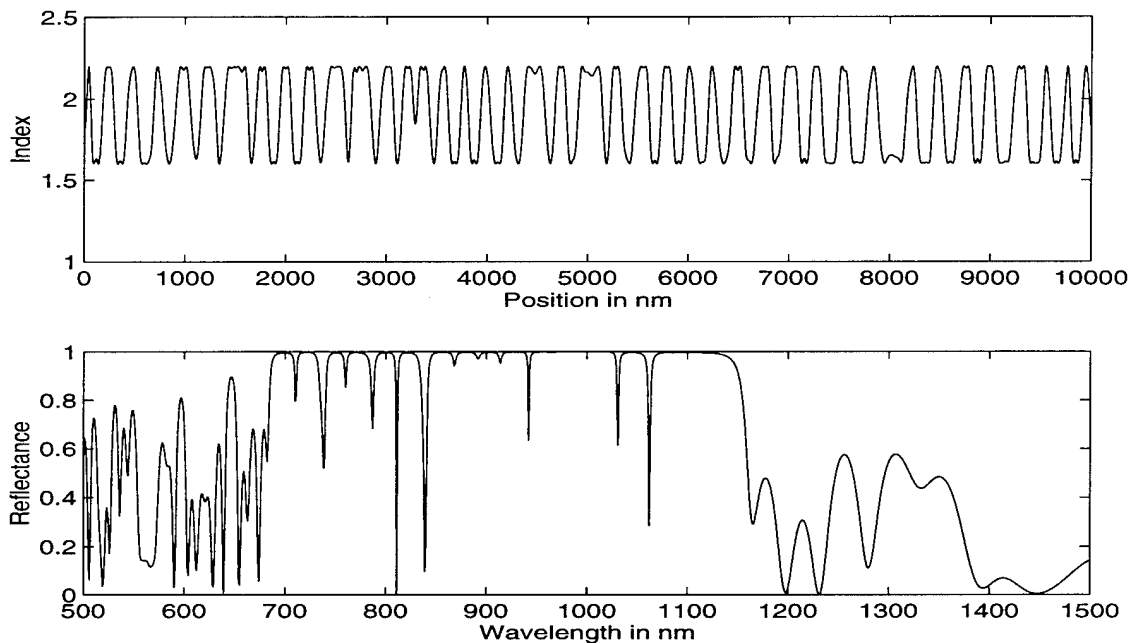
Figure 4.23: Wavelet design for Ti:Sapphire broadband mirror.

4.4. Dual Goal Optimization

One feature to note from the previous examples is the importance of the thickness of the film on the design. In practice, one of the objectives of a thin film design is to minimize the total thickness of the material, and another is to minimize the complexity of the film. A thinner, less complex film costs less to manufacture and can be produced more rapidly than a thicker, more complicated film. The traditional multilayer slab films are created by assuming a number of layers and then optimizing the thicknesses. The gradient index design process is almost the exact opposite: the thickness is fixed and the

"layer" structure is optimized. One of the aims of this research therefore was to develop a method for finding an optimal film that not only satisfied the requirements on reflectivity, but also minimized the total thickness of the film.

The idea that there exists a minimal thickness for some level of film performance is based on the hypothesis that the merit function for the optimal solution depends monotonically on the total thickness of the film. This hypothesis can be justified by the following argument. For a given substrate and film index range (containing the substrate index), and any arbitrary desired reflectance, $R_{desired}(k_i)$, the merit function for the optimal film design is given by

$$F = \sum_{i=1}^{K} \left(R_{desired}(k_i)^2 - R_{calculated}(k_i)^2 \right)^{\frac{1}{2}}$$ (4.13)

where $R_{calculated}(k_i)$ is determined by matrix methods as before. The film design with minimum thickness consists of the substrate-air interface only. This "film" of zero thickness has the same reflectance for all wavelengths (neglecting dispersion), given by the Fresnel relation:

$$R_{calculated} = \left(\frac{n_{sub} - 1}{n_{sub} + 1} \right)^2$$ (4.14)

and value of the corresponding the merit function is found from Equation 4.13 above. Now if some small amount of film is included, the optimal design must have a merit function that is lower than the zero film case, because the optimal solution space includes the previous solution of substrate-air only. (Note this argument is only valid if the substrate index is in the allowed index range.) So the minimum merit function corresponding to the optimal solution is a non-increasing function of the total film

thickness. As the total thickness of the film increases, the minimum merit function will decrease until the termination tolerance of the optimization is reached. Once there is "sufficient" film, increasing the film thickness may produce different films, but the merit function should remain at the termination tolerance level. Figure 4.24 illustrates this theoretical relationship between minimum merit function value and total film thickness.

Having established a monotonic relationship between the minimum merit function and the total film thickness, all that remains is to determine the minimum thickness for the desired performance level. This can be done by subtracting the acceptable performance level (in terms of the merit function) from the merit function curve. The zero crossing is then the minimum acceptable film thickness.

The argument above provides a sound heuristic basis for this approach, but there are some additional issues associated with the implementation of this dual goal optimization algorithm. The first is to establish to what extent the curve of Figure 4.24 is actually produced. In practice, the optimization algorithm requires an initial film design, and the output depends on the initial design. To ensure a non-increasing merit function, the initial design used to seed the optimizer should be a previous (thinner) design padded with an appropriate amount of substrate. However, there are a number of difficulties with
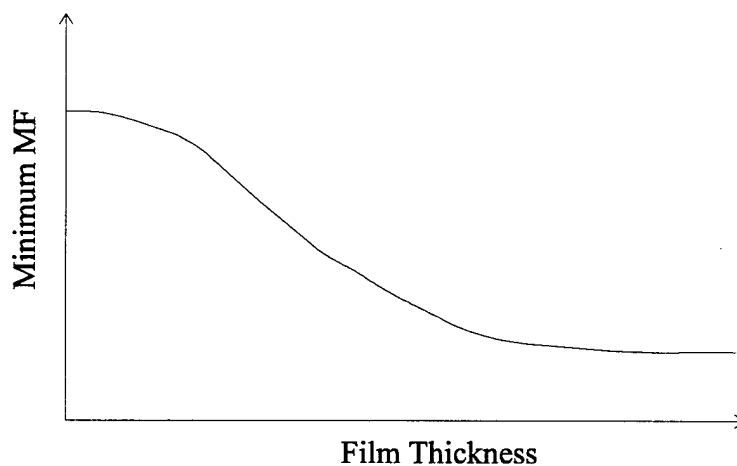
Figure 4.24: Theoretical dependence of merit function on film thickness.

this approach. First, the number of coefficients used must be a power of two, so a seed design formed by padding a previous solution must be resampled and decomposed, keeping only the few coefficients to be used as variables in the optimization. This makes the input seed much different from the padded version of the previous film, since the few coefficients used cannot accurately model the flat substrate and the gradient index film. Unfortunately, this resampled version of the index profile does not have the same reflectance as the original, since the resampled representation is not very good. While the differences between an index profile and the padded and resampled index profile are small, the coefficients that are used in the optimization are quite different, and can destroy the non-increasing nature of the merit function versus thickness curve. This can create shallow local minimums in this curve, which could lead to multiple roots for a minimum thickness. This problem can be reduced by increasing the number of coefficients used in the optimization, but that also increases the computation time.

The second consideration in implementing this idea is that the optimizer only identifies local minima. If the seed value is a padded version of a previous solution, this initial guess may be in a local minimum for the new thickness. (Recall that the same effective film was an optimal solution before.) This approach may yield higher merit function values than other seed values that reach different optimal solutions. Similarly, the size of the change in thickness from the previous solution to the new input can affect the outcome. For example, to get from a solution at thickness "A" to a new solution at thickness "B", one could take one step or many steps. The film (and merit function) generated can be very different for the different initial guesses. This problem of different answers for the same thickness precludes the direct application of most root finding algorithms. This difficulty can be worked around by estimating a curve similar to that of Figure 4.24 using the optimizer, and then finding the root from a numerical fit to the curve.

The question of which seed film to choose remains unanswered. Choosing the same film for all thicknesses often yields lower merit function values than the "pad the previous solution" approach. However, the constant seed value does not guarantee a non-increasing merit function versus thickness. A solution is to use a constant seed value until the merit function increases. When the merit function increases, the optimal solution for that thickness is recalculated using a padded version of the previous solution as the initial seed film. If this film also has a larger merit function, the previous merit function can be assigned to the current thickness. This is justified by the fact that a thinner film padded with substrate is certainly a solution for a larger thickness. This approach compromises between minimizing the value of the merit function and ensuring a non-increasing merit function versus thickness.

The minimization algorithm outlined above was implemented using the MATLAB™ programming language. Both the Fourier series optimization and the wavelet based optimization were implemented. The programs used are included in Appendix D. The Fourier series version starts with the function **FFTMIN.M**. This is a script function that requires an initial design in the MATLAB™ workspace that defines the variables. The function is similar to the **FFTRUN.M** function described previously to perform a single Fourier series optimal design. The **FFTMIN.M** function includes a loop which iterates several times to calculate the value of the merit function for different thicknesses. The program uses a constant film seed to start the optimizations (function **FLATEVAL.M**) unless the merit function increases. In this case the seed is formed by using the previous output padded with substrate and resampled using a cubic spline interpolation between points. This is implemented in the function **FPADEVAL.M**. The evaluation function for the optimization is the same as before, **FFTFUN.M**. The result of **FFTMIN.M** is an array of film thicknesses and merit function values. The minimum thickness for a specific desired merit function is found by using this data to define a functional dependence of the merit function on the film thickness. The functional

93

dependence is a cubic spline interpolation between data points. The minimum thickness is found by using this function in the MATLAB™ function **FZERO.M**. This function uses Brent's algorithm to determine the zero of a function. Brent's algorithm for root finding is a combination of a simple bisection technique and inverse quadratic interpolation [30:267-269]. The bisection method locates the zero of a function by finding two values for the function argument which have opposite sign, and then halving the interval between successive estimates until a solution point is found. This method works reliably, but the convergence may be slow. The inverse quadratic interpolation method has the advantage of very rapid convergence in most cases, but not all. This method fits a quadratic between three points of the function and uses the root of the quadratic as the estimate for the next guess. The combination of the two methods guarantees at least linear convergence, but in most cases performs much better.

Consider the example of the wavelet design of a Nd:YAG anti-reflection coating, as described in Section 4.3.2 above. The reflectance at the design wavelengths for use inside a laser cavity should be less than 0.01%. This will be the definition of an "acceptable" film. The examples of the various input thicknesses in that section demonstrated that a thickness of 500 nm was insufficient to generate a good anti-reflection coating, while the thickness of 750 nm or greater was acceptable. The number of non-zero coefficients is set at 32, to try to more accurately model a film padded with substrate. Figure 4.25 shows the result of the algorithm above to determine the merit function dependence on thickness. The open circles represent the data points, and the solid line is the cubic spline interpolation between points. Note that in this case the merit function is monotonically decreasing with increasing thickness. The minimum acceptable thickness is found by using the **FZERO.M** function described above, which determines the thickness with a merit function with a log of -4. The thinnest acceptable film was found to be 563 nm thick. This index profile and reflectance are illustrated in Figure 4.26, and the reflectance values at the design wavelengths are tabulated in Table 4-8.

94
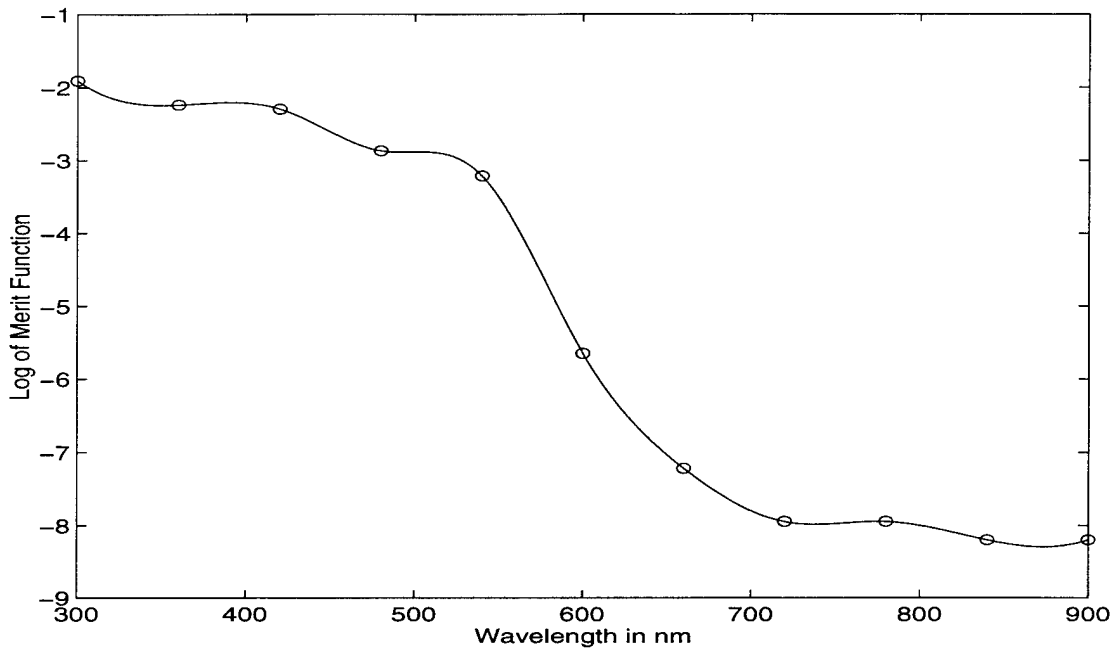
Figure 4.25: Log of Merit Function vs. Thickness for Nd:YAG AR coating design using wavelet method.
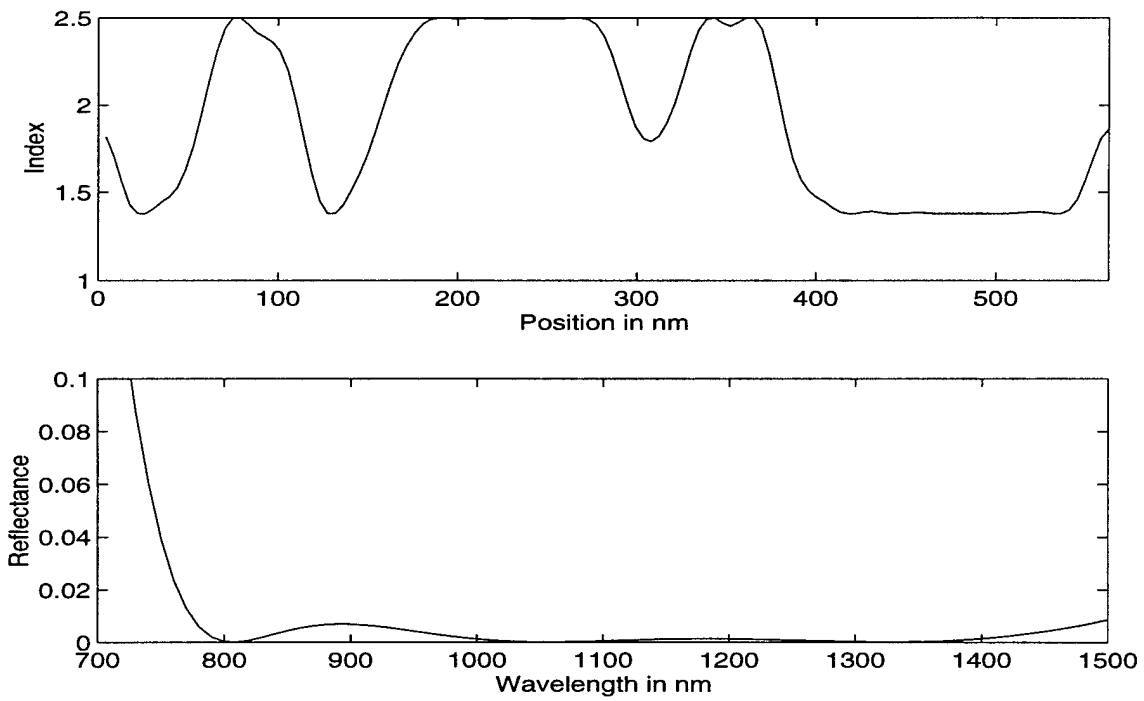
Figure 4.26: Index of Refraction and Reflectance of thinnest optimal film found using wavelet method.

95

This technique for finding a thinnest acceptable film can also be used with the Fourier series optimal design method. Figure 4.27 shows the estimate of the merit function dependence on thickness using the Fourier series method. Note that in this case the merit function has several points where it is constant with increasing thickness. These points correspond to thicknesses where the optimized solution was worse than the thinner solution. The minimum acceptable thickness is again found by using the FZERO.M function on a cubic spline interpolation of the data. The result is a film 568 nm thick which agrees well with the one found above. The minimum film found using the Fourier series method is shown in Figure 4.28, and the reflectance values at the design wavelengths are tabulated in Table 4-8.

Table 4-8: Design Values for Minimum Thickness Antireflection Ccoating

Wavelength (nm)	Reflectance for Wavelet film	Reflectance for Fourier Series film
810	9.53×10^{-5}	1.48×10^{-4}
1060	1.68×10^{-4}	2.69×10^{-4}
1330	1.85×10^{-4}	3.06×10^{-4}
Total Error	2.68×10^{-4}	4.33×10^{-4}
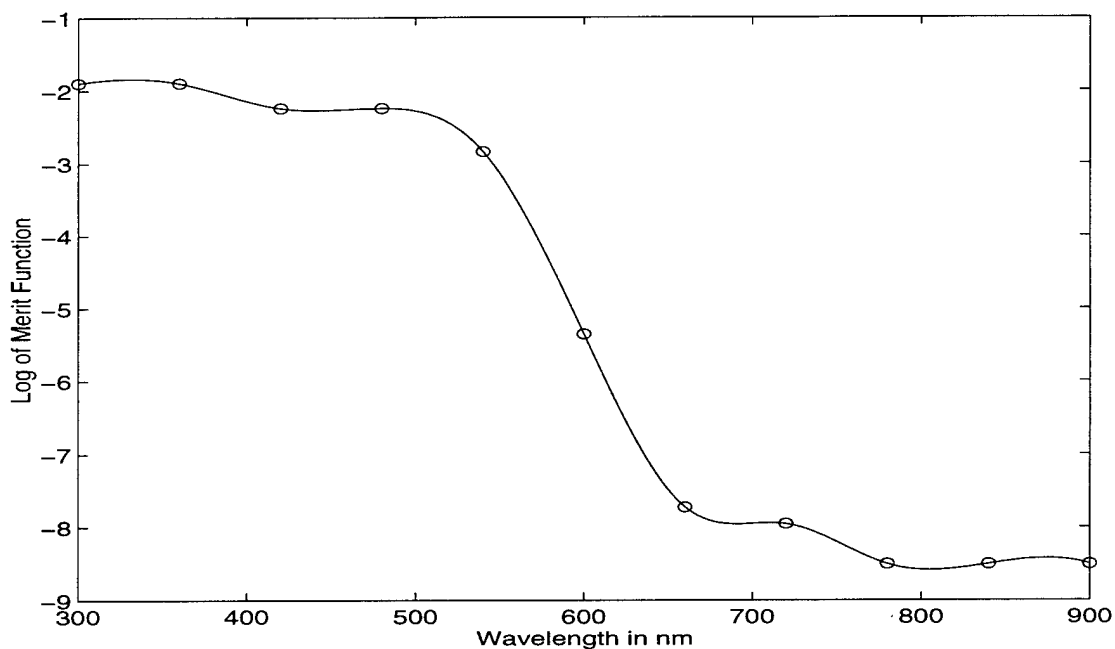
Figure 4.27: Log of Merit Function vs. Thickness for Nd:YAG AR coating design using Fourier series method.
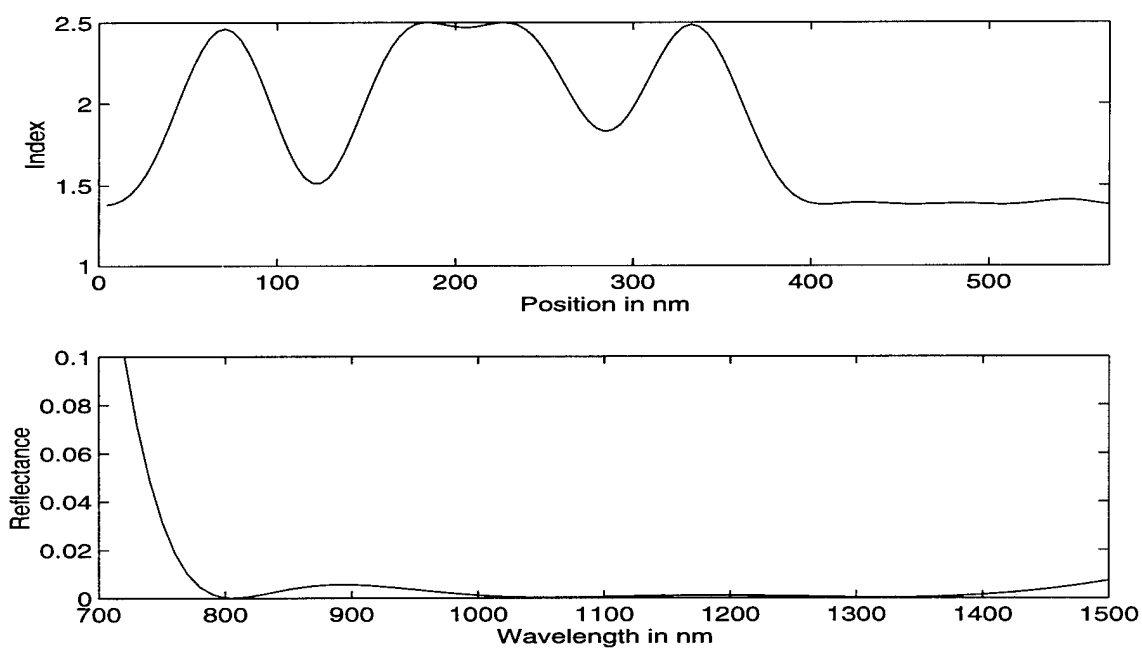
Figure 4.28: Index of Refraction and Reflectance of thinnest optimal film found using Fourier series method.

4.5. Comparison of Fourier Series and Wavelet Methods

The preceding sections described two similar methods for optimal design of gradient index thin films. In each case, similar examples were presented to illustrate the effect of the various parameters of the design. While the general optimal design method is the same, the choice of basis function does make a difference in the final design. Each of the two bases has advantages and disadvantages. In general, the Fourier series basis has the advantage of allowing the designer to choose any number of frequencies to use as design variables, so the number of design variables used can vary, for example, by factors of two. In contrast, the wavelet basis formulation requires the number of coefficients used as variables to be a power of two. For example, if 64 variable coefficients are not sufficient to solve the problem, the next design would need to use 128 variable coefficients. The Fourier series method also completed the optimization in a shorter time than the wavelet design for comparable examples, but this may be misleading. Comparison of computation time is not a "fair" test of the methods. The Fourier series decomposition and reconstruction was performed using the fast Fourier transform algorithm which is built in to the MATLAB™ kernel. The wavelet based decomposition and reconstruction was per-formed by higher level MATLAB™ programs, so the computation of coefficients was not as efficient. Theoretically, the discrete wavelet decomposition and reconstruction algorithms are $O(N)$ computations, where N is the number of samples of the function [8:152]. The fast Fourier transform, on the other hand, requires $O(N \log_2 N)$ computations [30: 408], where N is a power of 2. So if both algorithms are implemented at the same level (e.g., both compiled in C), the wavelet based computation of coefficients should take less time to compute.

The wavelet based approach does offer a number of advantages over the Fourier approach. First, for a given design problem and parameter set, the optimal solution using the wavelet method is usually found in fewer iterations of the main optimization loop

than the Fourier method (see Table 4-9 through Table 4-12 below). This would lead to faster computation times if both methods are implemented at the same level. In addition, the optimal wavelet based designs also tend to achieve a lower merit function value then the optimal Fourier series design for the same problem.

Table 4-9: Fourier series Nd:YAG designs with various thicknesses.

Thickness (nm)	Iterations
300	1996
500	5863
750	7996
1000	1971

Table 4-10: Wavelet Nd:YAG designs with various thicknesses

Thickness (nm)	Iterations
300	1160
500	6015
750	1794
1000	631

Table 4-11: Fourier series Nd:YAG designs with constant thickness.

Variables	Iterations
8	310
16	2465
32	7996
64	18411

Table 4-12: Wavelet Nd:YAG designs with constant thickness

Variables	Iterations
4	76
8	502
16	1507
32	1794

4.6. Conclusions

This chapter has introduced a new gradient index thin film design method based on a generalized Fourier series approach. This new method of gradient index thin film design extends the domain of problems for which gradient index solutions can be found. The method is analogous to existing techniques for layer based coating design, but adds the flexibility of gradient index films by varying the index of refraction instead of the thickness of the layers. The index of refraction for a film is specified by a few coefficients with respect to a basis. These coefficients are used as variables in a sequential quadratic programming optimization routine to design an index of refraction profile. Two basis systems were used to illustrate this method; Fourier series and a Daubechies wavelet multi-resolution analysis. The method works quite well for problems which require specific reflectances at a few specific wavelengths, both for high-reflection and anti-reflection applications. This method is not well suited for problems which require a large number off reflectance points be met. This technique is a good complement to the SWIFT algorithm of Chapter 3, since the strengths of the generalized Fourier series method are the weaknesses of SWIFT, and visa versa.

5.   Conclusion

Gradient index thin films provide greater flexibility for the design of optical coat-ings than the more conventional "layer" films. In addition, gradient index films have higher damage thresholds and better adhesion properties. The design of such gradient index thin films, however, is a difficult problem. There are many different design techniques, based on vastly different approaches to the problem. Two aspects of gradient index thin film design have been presented. The first was the Stored Waveform Inverse Fourier Transform (SWIFT) enhancement of the inverse Fourier transform design method, and the second was optimal design of gradient index films using generalized Fourier series.

The inverse Fourier transform method was modified using the SWIFT technique to include use of the phase of the index profile as a variable in rugate filter design. The SWIFT technique has two primary effects on the inverse Fourier transform design of gradient index thin films. First, it reduces the index of refraction range used in the design. The index range is reduced because the method shifts the high index contrast at the center of the film toward the edges, more evenly distributing the "work" of the film over the entire thickness. Second, the reflectance profile of the film designed using SWIFT is closer to the desired reflectance profile than films with no phase function. The edges of the reflectance profile, where the reflectance changes rapidly from low to high or high to low, are much sharper, with fewer, smaller oscillations than the non-SWIFT designs. Use of an optimal phase function in Fourier-based filter designs reduces the product of index contrast and thickness for desired reflectance spectra. The SWIFT technique enables the designer to generate an optimal design, including constraints on the index range available for the film.   A parameter of the SWIFT formulation, the "power spread

101

thickness", controls the index range. Increasing this parameter causes a decrease in the index range of the resultant film. The quantitative relationship between this "power spread thickness" parameter and the index range of the film is not clear, so the method for constraining the index range is one of designer trial and error or a stepwise search through the possible parameter values until an acceptable film is found.

Several examples of this technique were presented. A basic narrow bandstop filter design was used to illustrate the trade off between index contrast and film thickness using the SWIFT technique. The SWIFT design method was used to reduce the index contrast needed for the film by increasing the "power spread thickness" parameter, thus distributing the index contrast more uniformly across the film thickness. A second example of a broadband reflector, suitable for a Ti:Sapphire laser mirror, illustrated the SWIFT technique's affect on the reflectance profile. The shape of the reflectance spectrum is recovered with greater fidelity by suppression of Gibbs oscillations and shifting of side-lobes into desired wavelength regions. A third example, that of a dichroic mirror for an Nd:YAG laser, illustrated some of the difficulties with applying the SWIFT technique. This design method is a good choice for problems with complicated desired reflectance profiles which span a broad range of wavelengths. It has several drawbacks, however. One is the assumption that the entrance and exit media are the same. This is generally not the case in practical applications, and the anti-reflectance portion of the design is not well satisfied. The inverse Fourier transform approach is also not well suited to problems with a few reflectance requirements sparsely spread in wavelength, since there is no way to select an optimal desired Q function based on only a few reflectance requirements. A fourth example, that of an anti-reflection coating for the Nd:YAG, further illustrated the limitations on the application of the SWIFT design technique. The attempt to modify the method to solve an anti-reflection problem met with little success.

The second aspect of gradient index design presented was optimal design of gradient index films using generalized Fourier series (GFS). This new method of gradient index thin film design extends the domain of problems for which gradient index solutions can be found. The method is analogous to existing techniques for layer based coating design, but adds the flexibility of gradient index films by varying the index of refraction instead of the thickness of the layers. The coefficients of a GFS representation of the gradient index of refraction profile are used as variables in a non-linear constrained optimization formulation. This allows one to design a piecewise continuous gradient index film with limited number of variables. The optimal values of the design coefficients are determined using a sequential quadratic programming algorithm. Two basis systems, the familiar Fourier series and the newer wavelet representations, were successfully used in a number of examples to illustrate the features of this design method.

The first examples in each basis were to design anti-reflection coatings for the Nd:YAG laser. These illustrated the effect of the film thickness and number of design variables used on the index design. In each case, there is some minimum thickness of film needed to achieve the design goals. Similarly, one must use a sufficient number of generalized Fourier series coefficients as index design variables to design a film that meets the reflectance goals. However, using too many variables complicates the index design without improving the reflectance of the design, and should be avoided. The second example was the dichroic mirror for the Nd:YAG laser. This example illustrated the ability to successfully specify both high and low reflectances in this design method, and allows comparison with the similar, unsuccessful SWIFT design. The third example was the broadband reflector for a Ti:Sapphire laser. Both the Fourier series and wavelet versions of the generalized Fourier series design method had difficulty with this design, illustrating a limitation of this design method. Each of the two basis systems used has advantages and disadvantages. In general, the Fourier series basis has the advantage of

103

choosing any number of frequencies to use as design variables, so the number of design variables used can be chosen to be any product of small prime numbers. In contrast, the wavelet basis formulation requires the number of coefficients used as variables to be a power of two. The wavelet based approach does offer a number of advantages over the Fourier series approach. First, for a given design problem and parameter set, the optimal solution using the wavelet method is usually found in fewer iterations of the main optimization loop than the Fourier method. Since the wavelet decomposition algorithm has lower computational complexity than the fast Fourier transform algorithm, the wavelet basis approach should also take the least computation time. In addition, the optimal wavelet based designs also tend to achieve a lower merit function value than the optimal Fourier series design for the same problem. The GFS approach is ideal for the design of coatings for laser applications, were only a few reflectance values are specified. In contrast, this optimal design method is not very effective for problems with required reflectances specified for a broad, dense set of wavelengths. In current form, some of the larger problems, such as the Ti:Sapphire example, required days of computation time to arrive at a solution.

The generalized Fourier series method was also extended to determine the minimum film thickness needed, as well as the index of refraction profile for the optimal film. The minimum acceptable thickness for a film is determined by calculating the optimal film design for a number of thicknesses. The merit function for the film, which is a measure of performance for the film, was shown to be constant or monotonically decreasing as the film thickness increased. For a pre-defined level of acceptable performance, the minimum film thickness to achieve that performance can be determined. This thickness minimization was illustrated for both the wavelet and Fourier series bases using the Nd:YAG anti-reflector example described previously. The two methods produced similar films and minimum thicknesses.

Comparison between the SWIFT method and the generalized Fourier series method is complicated by the fact that the two techniques approach the problem of gradient index film design differently. The SWIFT approach states the problem as an attempt to map a reflectance specified over all wavelengths to a continuous, infinite extent index of refraction profile. In contrast, the GFS approach is formulated to map discrete reflectance requirements to a continuous index profile of finite extent. Further, in the SWIFT design the index profile is truncated after the design is complete, so the film thickness is chosen *a posteriori*, while in the GFS case, the film thickness must be chosen *a priori*. The similarity of the examples for the SWIFT method and the GFS method allows some comparison between the two techniques. Table 5-1 summarizes the three examples used and the applicability of each of the design techniques to those examples. The value in the table is physical thickness of the film for the best case example, and the NA indicates the method does not apply to the example. Note that the index profiles plotted in Chapter 3 are all plotted as a function of double optical path length, so the values in the table do not appear to match the illustrations. The physical thickness for the films were calculated from the double optical path lengths using Equation 3.3. In general, the GFS solution is achieved with much thinner films than the SWIFT solution. However, as the table indicates the SWIFT technique was not effective in the design of an anti-reflective coating, and the GFS method was not effective for the broadband reflector Ti:Sapphire mirror design. The generalized Fourier series method is a nice complement to the SWIFT method, since the strength of each method is the weakness of the other.

105

Table 5-1: Comparison of SWIFT and GFS design physical thickness.

Design Example	SWIFT	GFS-Wavelet	GFS-Fourier
Anti-Reflection Coating	NA	563 nm	567 nm
Dichroic Mirror	11.1 μm	4 μm	4 μm
Ti:Sapphire Mirror	15.9 μm	NA	NA

This work revealed several avenues for future research. In the SWIFT research area, these include characterizing the quantitative relationship between the "power spread thickness" parameter and the index range needed, investigating and quantifying the effects of truncating the film to a finite thickness on the resultant reflectance profile, and exploring methods to successfully incorporate a different entrance and exit media into the formulation. In the case of the generalized Fourier series optimal design technique, other areas for study include: developing faster computational algorithms and codes to allow the timely solution of larger scale problems, investigating the effects of other basis systems on optimization performance, and exploring the potential to use the multi-resolution wavelet decomposition strategy to enhance the optimization algorithm.

Appendix A : SWIFT Numerical Design Considerations

The discrete implementation of the inverse Fourier transform relation in Equation 3.1 is accomplished using the fast Fourier transform (FFT) algorithm in the MATLAB™ programming environment. MATLAB™ is a "technical computing environment for high-performance numeric computation and visualization" [27:*i*], which includes a high level programming capability. The MATLAB™ programs developed for this work are included in Appendix D. The implementation of the SWIFT technique is done using two programs: **QUE.M** and **SWIFT.M.** The program **QUE.M** uses a valid expression for *Q(k)* such as one of the forms of Equation 3.11 to generate a sampled version of *Q(k)/k* for notch reflectance profiles. **SWIFT.M** calculates the phase function for given *Q(k)/k* using the SWIFT technique (Equations 3.13 and 3.18), and then calculates the index profile using the built in inverse Fourier transform (Equation 3.1). This program numerically estimates the integral in Equation 3.18 from the sampled values of *Q(k)* using a simple trapezoid summation algorithm. A third program, **REFLECT.M**, is used to estimate the reflectance for a gradient index film by approximating the film by thin, homogeneous layers at each sample index value and implementing Equations 2.12 through 2.15.

The inputs to **QUE.M** are the lower and upper wavelength limits for the notch reflector in microns, the reflectivity of the notch (between 0 and 1), the number of samples to use, and the total thickness of the film to model in microns. Note that the total thickness of the film here is not the same as the desired film thickness, nor the "power spread" thickness of the SWIFT theory. The relationship among these three design parameters will be further discussed below. The functional dependence of *Q(k)* on the reflectivity or transmission of the film (such as in Equation 3.11) is hard coded in the

program. The output of the **QUE.M** program is a vector containing the samples values of $Q(k)/k$.

The inputs to the **SWIFT.M** program are the vector of $Q(k)/k$ values output by **QUE.M**, the number of samples, the lower and upper wavelength values in microns used to generate the Q function, the total film thickness, and the "power spread" thickness for the SWIFT algorithm. Again, this need not be the same as the desired film thickness. The output of the program is a matrix; the first column contains the sample values of the index, the second column contains the optical thickness of that layer, and the third column is the position of the sample in optical thickness.

The inputs to the **REFLECT.M** program are the index matrix of index values and layer thicknesses from **SWIFT.M**, and the number of plot points. The index of the substrate and exit medium are fixed in the code, as is the wavelength range to calculate. The output is a matrix with the reflectance in column one and the wavelength in microns in column two.

There are a few numerical computation issues associated with using the discrete implementation of the Fourier transform. The functions **QUE.M** and **SWIFT.M** above require the user to define the sampling scheme to use in the Fourier transform calculation. The goal is to accurately specify an index of refraction profile over some finite thickness of film. There are several related parameters that must be specified to calculate a discrete Fourier transform. Note that the Fourier transform variables in the theory are the double optical thickness, x, of the index of refraction and the wavenumber, k, of the Q function. However, the reflectance calculation requires the index profile be specified in terms of optical thickness. For consistency, all thicknesses specified as inputs to and outputs from the programs are in optical thickness, and the conversion to double optical thickness in done inside the programs as necessary.

The parameters for the index profile are: number of sample values, denoted by N_s, the spacing of samples in optical thickness, denoted by Δ, the total optical thickness of the film sampled, d_{Total}, and the design optical thickness of the film, d_{film}. The parameters for the Q function in k-space are: the number of samples, N_s, the spacing of samples, $\delta = 1/(d_{Total})$, and the total range of k-space sampled, $1/(\Delta)$. Clearly, the number of samples can be found from the simple relation: $N_s = d_{Total}/\Delta$. The critical frequency, or Nyquist frequency, is related to the sample spacing of the index by $f_c=1/(2\Delta)$. The various parameters for specifying the numerical sampling are illustrated in Figure A.1 below. The sampling densities actually used for calculation are much higher than depicted in this figure. The method for determining the sampling density needed is illustrated in the example below. All the examples in Section 3.3 use the same approach.
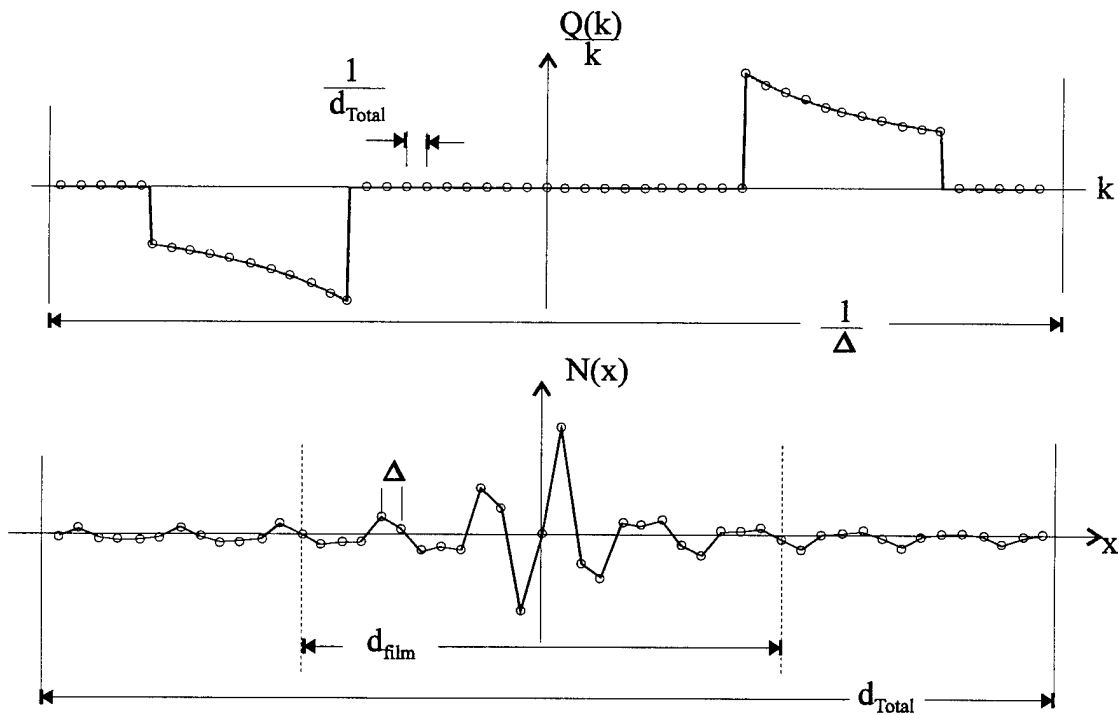
Figure A.1: Illustration of numerical parameters for discrete Fourier transform.

Acceptable values for the design parameters are found by considering the physical constraints on the design problem. First, the number of samples, N_s, should be a power of two to take advantage of the efficiency of the fast Fourier transform algorithm. The sample spacing of the index of refraction profile, Δ, should be on the order of one sample every five nm, to ensure the reflectivity calculation is accurate. Also, the Q function, which is the starting point for this design, must be adequately sampled. The number of non-zero samples of the Q function is denoted by N_0. Finally, the total spatial thickness of the film, d_{Total}, must be sufficiently large to minimize spatial aliasing in the index. Aliasing arises from the finite sampling of a function. Any power that lies outside the Nyquist range is folded back inside the range, generating in this case an inverse Fourier transform that is incorrect [30:387]. Viewed another way, this means that the sampling of the Q function must be sufficiently high to insure the index profile calculated is near the substrate value at the edges of the film.

To illustrate the process to determine acceptable parameter values for a problem, the parameters for the first example in Section 3.3.1 will be found below. Recall this example is to design a narrow-band reflectance filter with a reflectivity of 90% from the lower (initial) design wavelength, λ_i=580, to the upper (final) design wavelength, λ_f=620 nm, and 0% outside this band. The indices of the substrate and the incident medium are the same ($n_{sub} = n_{out} = 1.50$). The desired optical thickness of the film is d_{film}=30 μm. The numerical parameters for the sampling are found by requiring the index sample spacing, Δ, to be on the order of five nm in optical thickness, the number of non-zero samples, N_0, of the Q function to be about 25, and the total number of samples, N_s, to be a power of two. The first step in determining appropriate parameter values is to make a rough estimate, based on the values above. The second step will be to evaluate the numerical values produced, and then select a set of new numerical values that satisfy all the constraints.

First, since the sample spacing in k-space is the inverse of the total thickness (see Figure A.1), the requirement for 25 non-zero samples of the Q function over notch design wavelengths can be used to determine the total double optical thickness:

$$d_{Total} = \frac{N_0}{\dfrac{1}{\lambda_i} - \dfrac{1}{\lambda_f}} = \frac{25}{\dfrac{1}{0.58} - \dfrac{1}{0.62}} = 225 \ \mu m \qquad (A.1)$$

Note that for this calculation the spatial frequencies are defined as $1/\lambda$, rather than $2\pi/\lambda$. This is to be consistent with the QUE.M and SWIFT.M programs in Appendix D. This value for d_{Total} can be used along with the sample spacing, Δ, of the index profile to estimate the number of samples needed, N_s. The sample spacing, Δ, must be multiplied by a factor of 2 to convert the requirement on the sample spacing in the index from optical thickness (required by the reflectance calculation program) to double optical thickness, which is the variable of the Fourier transform relation. The sample spacing value to use in these calculations is therefore $\Delta=2(0.005)=0.010 \ \mu m$.

$$N_s = \frac{d_{Total}}{\Delta} = \frac{225}{0.010} = 22,400 \qquad (A.2)$$

At this point, constrain the number of samples, N_s , to be a power of two. Unfortunately, the N_s found above is not a power of two, and the closest powers of 2 are $2^{14}=16,384$, and $2^{15}=32,786$. In the interest of minimizing computation time, choose the lower number of samples, and then determine the impact on the other parameters. So, for $N_s=2^{14}=16,384$, and keeping the original desired sample spacing in double optical path length, $\Delta=0.010 \ \mu m$, the new total film thickness, d_{Total}, would be

111

$$d_{total} = N_s \Delta$$
$$= (16{,}384)(0.010) \tag{A.3}$$
$$= 163.84 \text{ nm}$$

Using this new value for d_{Total}, the number of non-zero samples of the Q function for this parameter choice can be determined by

$$N_0 = d_{total}\left(\frac{1}{\lambda_i} - \frac{1}{\lambda_f}\right)$$
$$= (163.84)\left(\frac{1}{0.58} - \frac{1}{0.62}\right) \tag{A.4}$$
$$\Rightarrow 18 \text{ non - zero samples of } Q(k)$$

At this point the designer can decide to accept this set of parameters, or decide to try a different combination of values. For example, one might decide that 18 samples of the Q function is not enough, and decide to adjust the total thickness again to increase the number of samples. This would also affect the sample spacing of the index profile. To illustrate this, choose to keep the total number of samples $N_s = 2^{14} = 16{,}384$, but now choose to increase the total film thickness to be $d_{Total} = 200$ μm. Using this new value for d_{Total}, the number of non-zero samples of the Q function is

$$N_0 = d_{total}\left(\frac{1}{\lambda_i} - \frac{1}{\lambda_f}\right)$$
$$= (200)\left(\frac{1}{0.58} - \frac{1}{0.62}\right) \tag{A.5}$$
$$\Rightarrow 23 \text{ non - zero samples of } Q(k),$$

and the sample spacing in index, Δ, would be

$$\Delta = \frac{d_{total}}{N_s}$$

$$= \frac{200}{16,384} \qquad\qquad (A.6)$$

$$= 0.0122 \quad \mu m$$

This corresponds to an optical thickness sample spacing of 0.0061 μm, which is sufficient.

This collection of design parameters is closer to the original objectives, so they are selected to use in the example. The final parameters required as input to the programs in Appendix D are expressed in optical path length (not double optical path length). The conversion from optical path length to double optical path length is done in the programs as needed. The reason for this is to keep the input consistent with that required by the reflectance calculation. Therefore the final parameter values for total film thickness, d_{Total}, and sample spacing in index, Δ, are the values found above divided by 2. For the example of Section 3.3.1 the parameters are:

$$N_s = 2^{14} = 16,384$$

$$d_{Total} = 100 \ \mu m$$

$$\Delta = 0.0061 \ \mu m \qquad\qquad (A.7)$$

$$N_0 \Rightarrow 23 \ \text{non-zero samples of } Q$$

This choice for the total optical thickness of the film should also minimize any aliasing effects, since it is over three times the desired film optical thickness. The same process is used to determine the parameters used in all of the examples in Chapter 3.

Appendix B : Optimization Theory

Introduction

The process of design requires a statement of the problem to be solved, identification of alternative solutions, and some measure of what constitutes a "best" solution to the problem. For simple problems, it may be possible to exhaustively list all possible solutions and "obviously" determine which one is best. For more complex problems, however, a more structured approach is needed to ensure the "best" solution is found. This structured approach consists of building a mathematical model for the system in question, identifying a set of variables to adjust in the design, and defining a "merit function" to numerically identify the best, or "optimal" solution. This section will summarize the theory of optimal design, which will be applied to the problem of gradient index thin film design in Chapter 4.

Statement of Optimal Design Problem

To express a design problem in mathematical form, one first must define the physical system to be modeled and identify the inputs to and outputs from this system. The next step is to construct a mathematical model of the system which approximates its performance by correctly mapping inputs to outputs. The model consists of a set of design variables and a collection of functions of these variables that define the objective and the constraints on the problem. The objective function (sometimes referred to as the merit function) is usually constructed such that its minimum value represents the optimal

design. The formal definition of the constrained optimization problem with n real-valued variables is therefore

$$
\begin{array}{c}
\underset{x \in \Re^n}{\text{minimize}} \quad f(x) \\[4pt]
\text{Subject to:} \\[4pt]
g_i(x) = 0 \qquad i = 1, 2, \ldots, m_e \\[4pt]
g_j(x) \le 0 \qquad i = m_e + 1, \ldots, m \\[4pt]
x_{\min} \le x \le x_{\max}
\end{array}
\tag{B.1}
$$

where x is a vector of the design variables, $f(x)$ is the objective function, the $g_i(x)$'s are the constraint functions, and the $x_{\min}$ and $x_{\max}$ are vector bounds on the design variables. It should be noted here that there are two general classes of constrained optimization problems. The first class is characterized by the fact that all the functions in Equation B.1 are linear functions of the design variables. This class of problems can be readily solved by the techniques of linear programming. The second class of problems, and the one with which the rest of this paper is concerned, have non-linear functions of the design variables as objective functions or constraints. This format of the optimal design problem with the constraints all expressed as less than or equal to zero is called the Null Negative Form [28:17]. The solution of the problem expressed by Equation B.1 is denoted by x^*.

A few additional definitions/concepts are needed before describing the methods used to solve the optimal design problem. First, a point x_o in the vector space $\Re^n$ is a "feasible" point if all constraints $g_i(x_o)$ are satisfied by equality or inequality, as required by the constraint. Conversely, a point in $\Re^n$ where this is not true is called an "infeasible" point. The second concept is that of "active" constraints. A constraint is said to be active if it is equal to zero when evaluated at the solution point ($g_i(x^*){=}0$). Clearly, all equality constraints are active by definition.

Now that the mathematical problem has been defined, how is a solution determined? Conceptually, the most straightforward method would be to select a feasible point x_k in the vector space $\Re^n$ as a start, then explore the neighborhood by adjusting each element of the vector by a fixed amount. Then evaluate the objective function and constraints at each test point, and identify the feasible point for which the objective function is smallest. This test point would become the new starting point, x_{k+1}, and the procedure would be repeated until a minimum value for the objective function is found. While this method is easily visualized, it is not practical. There are much more efficient techniques for determining the solution to the constrained optimal design problem. The current state of the art in non-linear optimization is a method known as Sequential Quadratic Programming (SQP) [28:365-372]. The basic theory for this method is outlined below.

Kuhn-Tucker Conditions and SQP

The general approach to solving non-linear constrained optimization problems is to transform the problem into an easier sub-problem, which can be iterated to find the true solution. The foundation of this approach is a set of necessary conditions on the solution of a general constrained optimization problem known as the Kuhn-Tucker (KT) equations: [26:2-22]

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0$$

$$\lambda_i^* g_i(x^*) = 0 \qquad i = 1,\ldots,m \qquad \text{(B.2)}$$

$$\lambda_i^* \geq 0 \qquad i = m_e + 1,\ldots,m$$

The first equation is the gradient of the Lagrangian, and describes a canceling of the gradients between the objective function, f, and the constraints, g_i, weighted by Lagrange multipliers, λ_i. Only active constraints are included in this expression, so the Lagrange multiplier for any inactive constraint is set to zero. The solution point x^* is assumed to be a "regular" point, which means that gradients of the constraints are linearly independent at that point. Note that the KT equations are not sufficient conditions for a solution, so each point which satisfies Equation B.2 must be checked individually to determine if it is the optimal solution. An additional sufficiency condition can be formed using information on the second derivative of the Lagrangian. For any regular solution point of the KT equations, x^*, if the Hessian of the Lagrangian at x^* is positive-definite, then the solution point is a local constrained minimum [28:184]. A positive-definite matrix is symmetric and has only positive, real elements. The mathematical expressions for the Lagrangian (denoted by L) and the Hessian of the Lagrangian (denoted by H), are given below:

$$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x), \qquad x = (x_1, x_2, \ldots, x_n)$$

$$H(x,\lambda) = \begin{bmatrix} \dfrac{\partial^2 L}{\partial x_1 \partial x_1} & \dfrac{\partial^2 L}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 L}{\partial x_1 \partial x_n} \\[2mm] \dfrac{\partial^2 L}{\partial x_2 \partial x_1} & \dfrac{\partial^2 L}{\partial x_2 \partial x_2} & & \\[2mm] \vdots & & \ddots & \vdots \\[2mm] \dfrac{\partial^2 L}{\partial x_n \partial x_1} & \cdots & \cdots & \dfrac{\partial^2 L}{\partial x_n \partial x_n} \end{bmatrix} \qquad (B.3)$$

These necessary and sufficient conditions for the solution of the original constrained optimization problem can be used in building a sequence of easier sub-problems to solve.

117

The easier sub-problem to be solved is to determine the best "direction" in the vector space of variables to move in for the next iteration, as well as the length of step to take along this direction. Again using the null negative form of the problem in Equation B.1, form the Lagrangian as

$$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \tag{B.4}$$

The simplified problem, called the Quadratic Programming (QP) sub-problem, is to solve the quadratic approximation to the Lagrangian above subject to linearized constraints [28:367]. The quadratic approximation to the Lagrangian is found by expanding in a Taylor's series and keeping terms up to second order in x. For any fixed point in the vector space of variables, x_k, the QP sub-problem can be written as

$$\begin{array}{l}
\underset{d \in \mathfrak{R}^n}{\text{minimize}} \quad \frac{1}{2} d^T H d + \nabla f(x_k)^T d \\[2mm]
\text{subject to:} \\[1mm]
\nabla g_i(x_k)^T d + g_i(x_k) = 0, \qquad i = 1,2,\ldots,m_e \\[1mm]
\nabla g_i(x_k)^T d + g_i(x_k) \leq 0, \qquad i = m_e + 1,\ldots,m
\end{array} \tag{B.5}$$

Here the functions f and g_i are the original objective function and constraints, H is a positive-definite approximation to the Hessian of the Lagrangian, and d is the new step direction for the next iteration. Bold typeface indicates a vector or matrix. The superscript T indicates matrix transpose. The solution to this QP sub-problem is used to determine the next iteration point $x_{k+1} = x_k + \alpha d$, where the step length α is to be determined. The step length α is needed to stabilize the algorithm, because it is possible

that the solution to the QP sub-problem is infeasible. It is found by performing a line search along the direction d and minimizing a merit function, which includes penalties for infeasible solutions.

In summary, the SQP algorithm is:

1. Select a starting point, x_o, and an initial estimate for λ_o. (k=0)

2. Solve the QP sub-problem above to determine the step direction d and the new estimate of λ_{k+1}. ($k=k+1$)

3. Set $x_{k+1} = x_k + \alpha d$.

4. Test for termination. If criteria not satisfied, goto step 2.

The SQP algorithm for solution of non-linear optimization problems is available in the MATLAB™ Optimization Toolbox [26, 27]. MATLAB™ is an interactive software package for scientific and engineering numeric computation [27:4]. The next section outlines the numerical methods used by MATLAB™ to implement this algorithm.

MATLAB™ Implementation of SQP

The numerical optimizations performed in this research were done using the MATLAB™ software package and in particular the MATLAB™ optimization toolbox [26,27]. The MATLAB™ software uses the SQP algorithm described above in its constrained optimization routines. To implement the theory above, the program must estimate the gradients and the Hessian at each point. The implementation of the SQP is done in three main stages [26:2-26]:

1. Update the estimate for the Hessian matrix of the Lagrangian.

2. Solve the QP sub-problem.

3. Perform line search and merit function minimization.

Equality vs. Inequality Constraints

The MATLAB™ algorithm differentiates between equality and inequality constraints. Equality constraints are active by definition, and have non-zero Lagrange multipliers. Inequality constraints are included in null negative form, and may or may not be active. The program uses an active set strategy approach to handle the inequality constraints. This means that an estimate of which constraints are active is calculated at each iteration. This estimate of activity figures in the merit function used in the determination of step size. Also, the zero tolerances for the equality and inequality constraints are controlled separately. This can impact the results, depending on how the problem is originally stated.

Gradient Estimation

If the objective functions and constraints are not known analytically, the partial derivatives needed to estimate the gradients must be calculated numerically. MATLAB™ uses an adaptive finite difference technique to find the values of the gradients as needed. The maximum and minimum perturbations for this calculation can be controlled by the user.

BFGS Method for Hessian Estimation

The key to the SQP algorithm is knowledge of the Hessian of the Lagrangian for the problem. Since this matrix is computationally expensive to calculate directly, it is estimated and refined through iteration using the formula devised by Broyden, Fletcher, Goldfarb, and Shanno (BFGS) [26:2-26]. In addition, the estimate for the Hessian is forced to remain positive-definite, so the solution to the original problem is a minimum.

The optimization goal is achieved when three termination criteria have been satisfied. There are three numerical tolerances to be met, one on the objective function, one on the variables, and one on the constraints. The objective function tolerance specifies the precision required on the objective function at the solution. The termination criteria for variables specify the minimum acceptable precision on the values of the variables. The constraint termination criterion is the maximum allowable violation of the constraints at the solution point. All three termination criteria must be satisfied simultaneously to achieve an optimal solution. The default termination tolerances are: objective -- 10^{-4} , variables -- 10^{-4} , and constraints -- 10^{-7}. All three can be adjusted if necessary.

Appendix C : Wavelet Theory

Wavelets are a relatively new mathematical technique. They are a form of generalized Fourier series, using wavelets as basis elements instead of sines and cosines. These new basis elements will be used to solve Maxwell's equations for the fields inside the film, and establish a new form for the mapping between index of refraction and reflectivity. The sections below will first introduce some of the basic concepts of wavelets, and then describe the technique of multi-resolution analysis.

Basic Wavelet Concepts

The term wavelets was first coined in 1982 by Morlet, who used a mathematical technique similar to Fourier series but using "little waves" as the basis elements, instead of sines and cosines [8:2]. The wavelets must be oscillatory (the wave part) and must also decay rapidly (the little part). An example of such a wavelet, called the Morlet wavelet, is shown in Figure C.1.
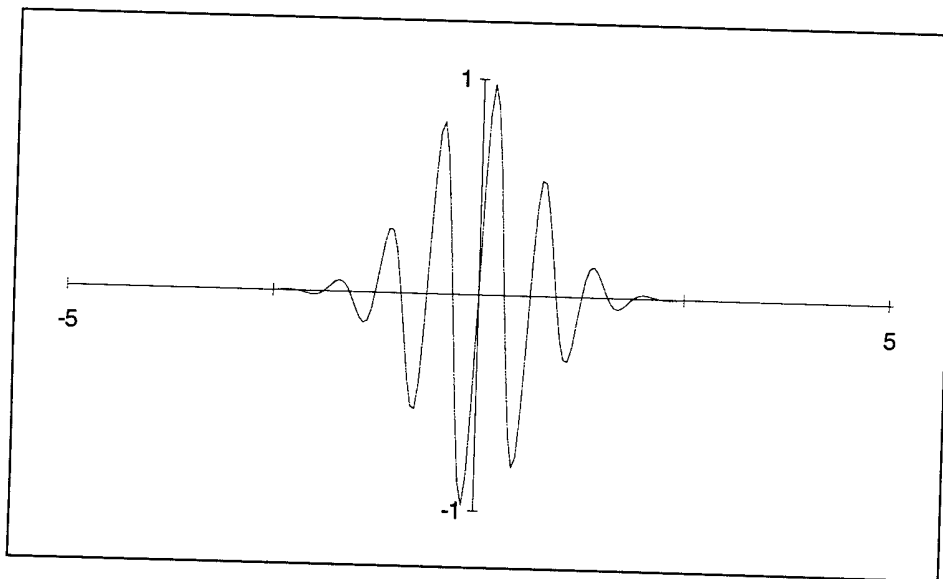
Figure C.1: Morlet Wavelet

In Fourier analysis the pieces used to decompose and reconstruct a signal are sines and cosines of different frequencies. In wavelet analysis the building blocks are translated and dilated (shifted and scaled) versions of a single wavelet, called the "mother wavelet" [46:2]. Denoting the mother wavelet by $h(t)$, the rest of the building blocks are given by

$$h_{a,b}(t) = \frac{1}{\sqrt{|a|}} h\left(\frac{t-b}{a}\right) \qquad (C.1)$$

where a is the scale and b is the shift [7:909]. The mother wavelet h must satisfy the condition

$$\int_{-\infty}^{\infty} h(x)dx = 0 \qquad (C.2)$$

which implies that $\hat{h}(0) = 0$, where the ^ denotes Fourier transform. If the shift b and scale a vary continuously, the continuous wavelet transform based on this mother wavelet, h, maps a one dimensional function $f(x)$ to a two dimensional surface $W_h f(a,b)$. The mathematical form of this transform is

$$W_h f(a,b) = \int_{-\infty}^{\infty} f(x) h_{a,b}(x)dx \qquad (C.3)$$

If the shift and scale are restricted to a discrete sublattice, the continuous wavelet transform above becomes a discrete wavelet transform. The wavelets become [7:910]

$$h_{m,n}(t) = a_0^{-m/2} h(a_0^{-m}t - nb_0) \qquad m,n \in \mathbb{Z} \qquad (C.4)$$

where

$$a = a_0^m$$

$$b = nb_0 a_0^m \qquad\qquad\qquad\text{(C.5)}$$

Z is the set of integers

For this work it is convenient to choose a_0=2, and b_0=1, although other choices are possible. With this choice of a_0 and b_0 it is possible to construct an $h(t)$ such that $h_{m,n}$ form an orthonormal basis for $L^2(\Re)$ [7:911], where $L^2(\Re)$ is the set of all measurable functions of finite energy, defined by

$$L^2(\Re) = \left\{ f : f \text{ is measurable and } \int_{-\infty}^{\infty} |f(x)|^2 dx < \infty \right\} \qquad\text{(C.6)}$$

Multiresolution Analysis

A multiresolution analysis is a technique of representing a function $f \in L^2(\Re)$ as a limit of successive approximations, each of which is a smoothed version of f [7:913, 25]. A multiresolution analysis consists of: [7:915, 45:152-158]

(i.) A group of embedded subspaces $V_m \subset L^2(\Re)$, such as

$$... \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} ... \qquad\text{(C.7)}$$

(ii.) These subspaces have only the zero element in common, and the sum of the spaces span all of $L^2(\Re)$:

$$\bigcap_{m \in Z} V_m = \{0\} \qquad \overline{\bigcup_{m \in Z} V_m} = L^2(\Re) \qquad\text{(C.8)}$$

(iii.) If a function is an element of one space, then the same function scaled by a factor of 2 is an element of the next higher space:

$$f(x) \in V_m \qquad f(2x) \in V_{m-1} \qquad \text{(C.9)}$$

(iv.) There exists a function $\phi \in V_0$ such that shifted versions of this function form a basis for the space V_0:

$$V_0 = \overline{span\{\phi_{0,n}, n \in Z\}}$$
$$\phi_{m,n}(t) = 2^{-m/2} \phi\left(2^{-m} t - n\right) \qquad \text{(C.10)}$$

The first property above means that any feature that can be seen at coarse resolution (like V_1) can also be seen at finer resolutions (like V_0). Naturally, the finer resolutions also contain more detailed features as well. The second property describes the limits of this ladder of subspaces. At the coarsest resolution there is nothing left but the zero element of the space, while in the limit of the finest resolution the V_m converge to $L^2(\Re)$. The third property is the same scaling feature already identified as a property of wavelets. The last property is the building block for the analysis. For a fixed dilation (scale), translations (shifts) of the function ϕ are basis elements for that subspace. These functions ϕ are called scaling functions, or father wavelets. Only real scaling functions will be considered here, although complex ϕ are possible.

Since a space V_m has a basis given by $\phi_{m,n}$, then any element of the space can be written as a linear combination of these basis elements. That is if $x(t) \in V_m$, then there exist coefficients $c_{m,n}$ such that

$$x(t) = \sum_{n \in Z} c_{m,n} 2^{-m/2} \phi\left(2^{-m} t - n\right)$$
$$= \sum_{n \in Z} c_{m,n} \phi_{m,n}(t) \qquad\qquad \text{(C.11)}$$

The $c_{m,n}$ are called approximation coefficients. They are found by taking the inner product of the function $x(t)$ with the scaling function $\phi_{m,n}$

$$c_{m,n} = \int_{-\infty}^{\infty} x(t) \phi_{m,n}(t) dt$$
$$= \left\langle x, \phi_{m,n} \right\rangle. \qquad\qquad \text{(C.12)}$$

The Dirac bracket notation is used to denote the inner product of two functions. For $f \in L^2(\Re)$, a projection operator P_m can be defined by

$$P_m f = \sum_{n \in Z} \left\langle f, \phi_{m,n} \right\rangle \phi_{m,n} \qquad\qquad \text{(C.13)}$$

which gives the expression for that part of $f(t)$ that is an element of V_m. Since all the elements in V_m are in V_{m-1} (by the construction of the ladder of subspaces), what elements are in V_{m-1} that are not in V_m? Construct a projection operator Q_m that identifies such elements by

$$Q_m f = P_{m-1} f - P_m f \qquad\qquad \text{(C.14)}$$

By the Classical Projection Theorem, the $Q_m f$ is orthogonal to V_m, and so defines a space of all such elements, called W_m, whose inner product with elements of V_m is zero. In set notation, that is [45:162]

$$W_m = \left\{ f \in V_{m-1} : \quad \langle f, v \rangle = 0 \; \forall \; v \in V_m \right\} \tag{C.15}$$

So by definition of W_m and Q_m

$$P_{m-1} f = P_m f + Q_m f$$
$$\text{c} \tag{C.16}$$
$$V_{m-1} = V_m \oplus W_m$$

where the symbol "$\oplus$" is the orthogonal direct sum of the two spaces. The direct sum is the set of vectors formed by adding one element from each of the orthogonal subspaces. Equation C.16 is a key point in the multiresolution analysis theory. Each subspace V_{m-1} consists of the direct sum of two other subspace, V_m and W_m. This idea is illustrated graphically in Figure C.2.
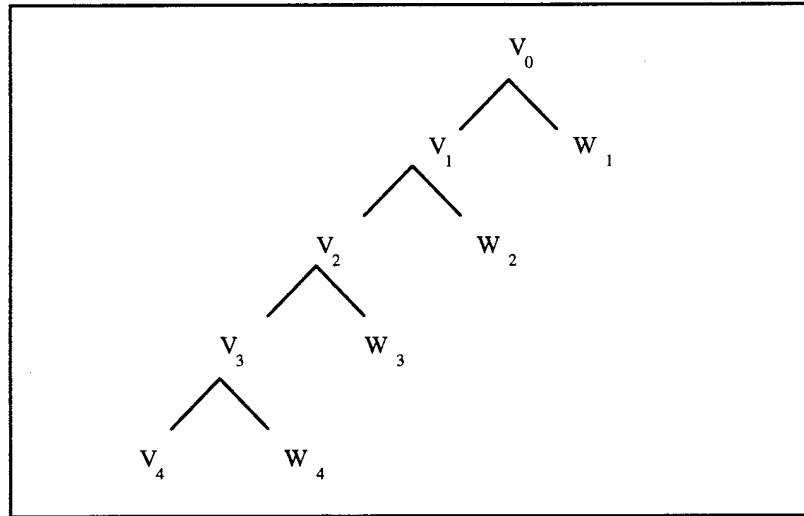
Figure C.2: Structure of Nested Subspaces

By the nature of the ladder of subspaces and the fact that W_m is orthogonal to V_m, the W_m's are all mutually orthogonal. Furthermore, the combination of all the W_m's spans the entire space of $L^2(\Re)$ functions:

$$\bigoplus_{m \in Z} W_m = L^2(\Re) \qquad\qquad (C.17)$$

Now identify the basis for each W_m as $\{ \Psi_{m,n}: n \in Z \}$. This set of functions forms a basis for the entire space $L^2(\Re)$, and they are the wavelets we seek. The details of the techniques used to construct the scaling function $\phi_{m,n}$ and the wavelets $\Psi_{m,n}$ are not necessary to understand the use of this theory. Several different types of scaling functions and wavelets have been developed in the literature. These include the Haar basis above, which is the simplest, as well as Daubechies' compactly supported wavelets [7:980]. Illustrations of some of Daubechies' wavelets and their associated scaling functions are shown in Figure C.3. The different wavelet-scaling function pairs shown are labeled by the number of filter coefficients, or "taps", needed to completely specify each wavelet. The functional form of the scaling functions and wavelets are not required to implement the method of multiresolution analysis, as will be seen in the next section.
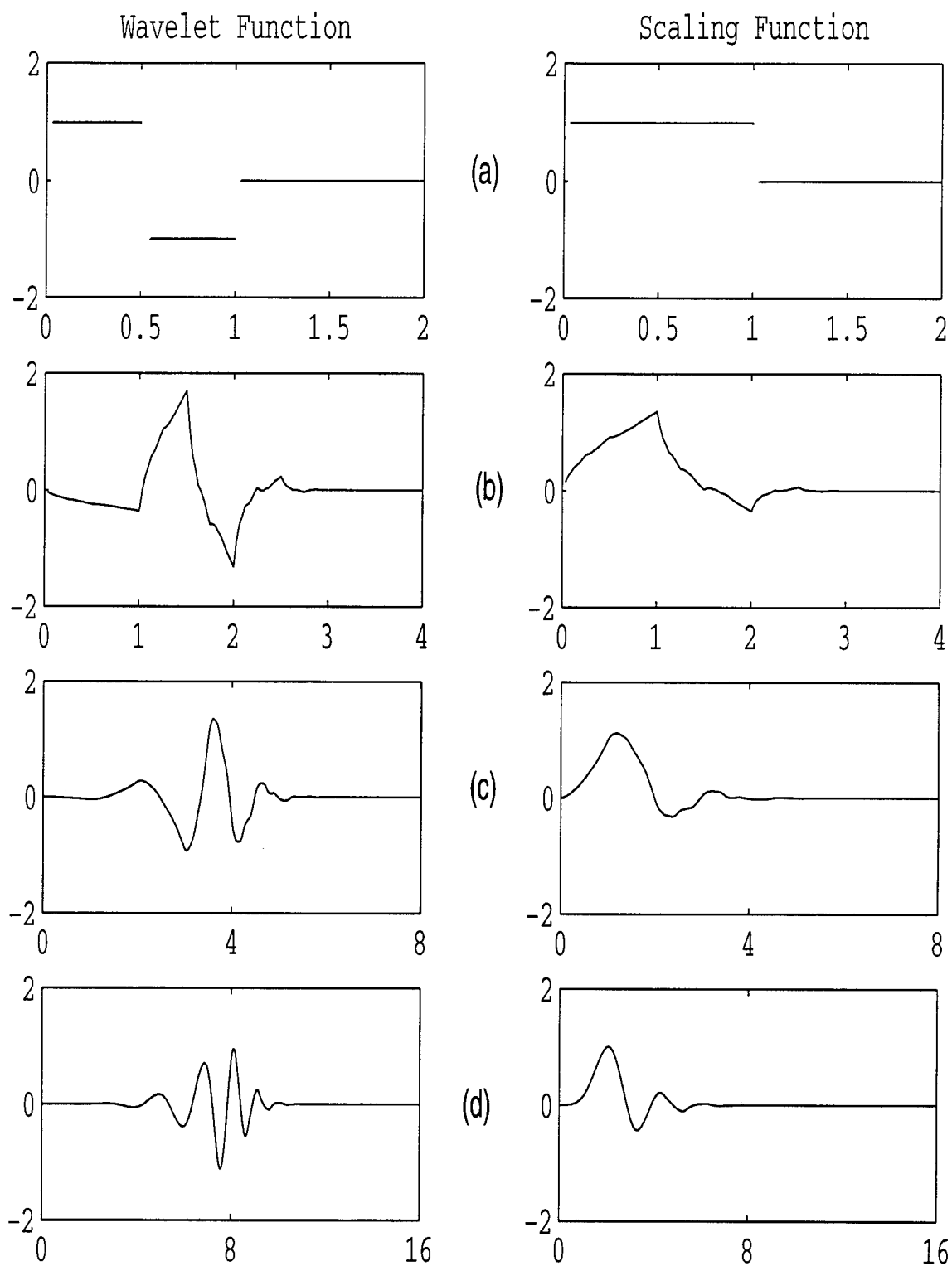
Figure C.3: Illustration of Daubechies' compactly supported wavelets and scaling functions: (a) 2 tap (Haar), (b) 4 tap, (c) 8 tap, (d) 16 tap.

Function Decomposition and Reconstruction

To make use of the multiresolution analysis theory, a method is needed to represent the function to be analyzed in each of the subspaces, preferably without performing numerous integrations. To develop such a method, define the coefficients of expansion for each level "m" by

$$P_m f(t) = \sum_{n \in Z} c_{m,n} \phi_{m,n}(t)$$

$$Q_m f(t) = \sum_{n \in Z} d_{m,n} \psi_{m,n}(t)$$

(C.18)

The coefficients $c_{m,n}$ are found by taking the inner product of the projection of $f(t)$ onto V_m with the scaling function $\phi_{m,n}$. Applying the relationship between the projection operators in Equation C.17 to this expression for $c_{m,n}$ yields

$$c_{m,n} = \left\langle P_m f, \phi_{m,n} \right\rangle$$
$$= \left\langle P_{m-1} f, \phi_{m,n} \right\rangle - \left\langle Q_m f, \phi_{m,n} \right\rangle$$

(C.19)

However, the second inner product is zero because $Q_m f(t) \perp \phi_{m,n}$ by construction. This leaves a recursive relationship between the coefficients:

$$c_{m,n} = \sum_{\ell \in Z} c_{m-1,\ell} \left\langle \phi_{m-1,\ell}, \phi_{m,n} \right\rangle$$

(C.20)

Similarly for $d_{m,n}$

$$d_{m,n} = \langle Q_m f , \psi_{m,n} \rangle$$
$$= \langle P_{m-1} f , \psi_{m,n} \rangle - \langle P_m f , \psi_{m,n} \rangle \qquad \text{(C.21)}$$
$$= \sum_{\lambda \in Z} c_{m-1,\lambda} \langle \phi_{m-1,\lambda} , \psi_{m,n} \rangle .$$

The recursive relations above show all that is needed to decompose a function is to find the inner products between a scaling function and the scaling function or wavelet of the next level down. Consider the explicit form of one of these inner products

$$\langle \phi_{m-1,\ell} , \phi_{m,n} \rangle = 2^{-(m-1)/2} 2^{-m/2} \int_{-\infty}^{\infty} \phi\left(2^{-(m-1)} t - \ell\right) \phi\left(2^{-m} t - n\right) dt \qquad \text{(C.22)}$$

Let $\dfrac{u}{2} = 2^{-m} t - n$. Then

$$\langle \phi_{m-1,1} , \phi_{m,n} \rangle = 2^{-1/2} \int_{-\infty}^{\infty} \phi\left(\frac{u}{2}\right) \phi\left(u - [1 - 2n]\right) du \qquad \text{(C.23)}$$

Define a sequence $h(n)$ by

$$h(n) \doteq 2^{-1/2} \int_{-\infty}^{\infty} \phi\left(\frac{t}{2}\right) \phi\left(t - n\right) dt \qquad \text{(C.24)}$$

then

$$\langle \phi_{m-1,\lambda} , \phi_{m,n} \rangle = h(\lambda - 2n) \qquad \text{(C.25)}$$

Notice this relation is **independent of the level m!** This means that the coefficients for any level can be found from the coefficients for the level above. (Also note the sequence $h(n)$ defined in Equation C.25 above is not related to the wavelet $h(t)$ defined in the previous section. This conflicting notation is standard within the wavelet

131

literature.) The coefficients required to specify the Daubechies' wavelets shown in Figure

C.3 are listed in Table C-1.

Table C-1: Daubechies' wavelet coefficients [7:980]

h=2:	0.7071067811865	h=16:	0.054415842243
	0.7071067811865		0.312871590914
			0.675630736297
h=4:	0.482962913145		0.585354683654
	0.836516303738		-0.015829105256
	0.224143868042		-0.284015542962
	-0.129409522551		0.000472484574
			0.128747426620
h=8:	0.230377813309		-0.017369301002
	0.714846570553		-0.044088253931
	0.630880767930		0.013981027917
	-0.027983769417		0.008746094047
	-0.187034811719		-0.004870352993
	0.030841381836		-0.000391740373
	0.032883011667		0.000675449406
	-0.010597401785		-0.000117476784

Thus, there is a recursion relation between levels that can be exploited to calculate the coefficients:

$$c_{m,n} = \sum_{\ell \in Z} c_{m-1,\ell} h(\ell - 2n) \tag{C.26}$$

A similar relation exists for the detail coefficients, $d_{m,n}$:

$$\langle \phi_{m-1,\ell}, \psi_{m,n} \rangle = g(\ell - 2n)$$

$$g(n) = 2^{-1/2} \int_{-\infty}^{\infty} \psi\left(\tfrac{t}{2}\right) \phi(t - n)\, dt \tag{C.27}$$

$$d_{m,n} = \sum_{\ell \in Z} c_{m-1,\ell}\, g(\ell - 2n)$$

The decomposition algorithm in this multiresolution analysis requires only the filters h and g to define the wavelets to be used. In fact, the g filter can be derived from the h filter, so only one sequence is needed to completely determine the wavelets. This relationship between h and g is

$$g(n) = (-1)^n h(n - 1) \tag{C.28}$$

Using the two recursion relations for $c_{m,n}$ and $d_{m,n}$, any function can be decomposed into its approximation and detail coefficients from an initial sequence c_0. Figure C.4 illustrates the relationship between these coefficients. Note the structure of the coefficients is the same as the structure of the underlying spaces, V_m and W_m.

Now that the technique used to decompose a function has been identified, how is the original function reconstructed from these coefficients? As can be seen from Figure C.4, only the coarsest (lowest) approximation coefficient and the detail coefficients are needed to reconstruct the original c_0.
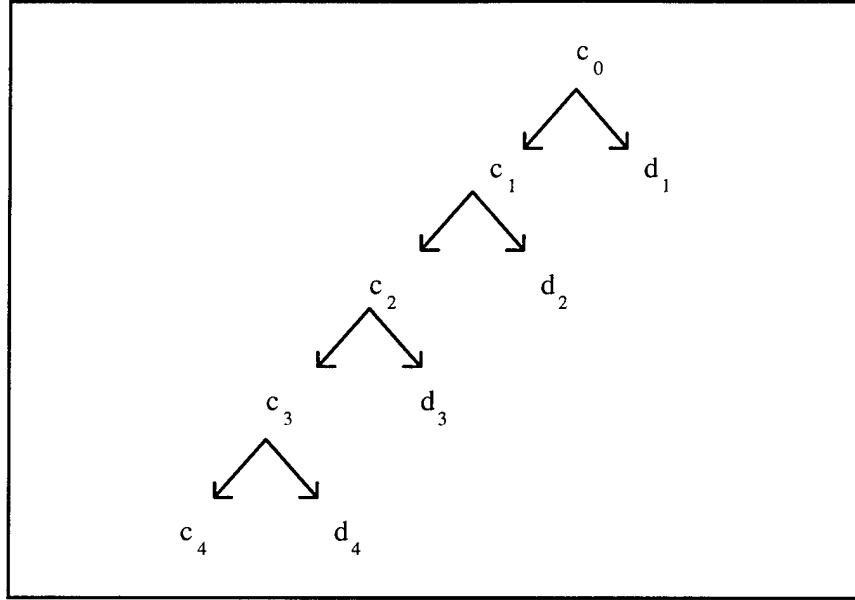
Figure C.4: Wavelet decomposition tree.

The recursive relation for reconstruction is

$$c_{m-1,k} = \sum_{n \in Z} c_{m,n} h(k - 2n) + \sum_{n \in Z} d_{m,n} g(k - 2n) \qquad (C.29)$$

A further simplification to the decomposition-reconstruction algorithm can be made by treating the initial sequence of coefficients as part of a periodic sequence. Wavelet transforms have the property that the transform of a periodic function is also periodic. In the case of the multiresolution decomposition, if the top sequence c_0 has N values, and N is a power of 2 (i.e., $N=2^M$ for some integer M), then the next level coarse coefficients c_1 will be 2^{M-1} periodic, as will the details for that level, d_1. So the lowest level of decomposition consists of a 1-periodic coarse coefficient sequence (a constant), and a 1-periodic detail coefficient sequence.

134

This appendix includes the MATLAB™ programs used to generate the thin film designs presented in this dissertation. The MATLAB™ language is designed to efficiently handle matrix and vector manipulations. The reader should bear in mind that each variable represents a matrix. MATLAB™ also does not require pre-definition of variables or variable typing. The colon operator (:) is used to access a series of elements of an array or matrix. For example *A(1:5)* refers to the first five elements of an array named *A*. Comments in the programs are denoted by the percent sign (%). For other details of the MATLAB™ programming language, see the user's manuals [26,27].

The first three programs are used in conjunction with the SWIFT algorithm described in Chapter 3. The function **QUE.M** calculates the *Q* function for a notch reflector. The **SWIFT.M** function converts the *Q* function into and index profile. The **REFLECT.M** function determines the reflectance of an input index profile. Each of these programs is described in some detail in Chapter 3.

The next four files are used to perform the Fourier series based optimal designs of Chapter 4. The top level function is **FFTRUN.M**. The optimization is performed using the MATLAB optimization toolbox function **CONSTR.M**. The help file for this program is included to explain the input and output of this program. The evaluation function is **FFTFUN.M**, which converts the optimization variables to merit function and constraint values. The fourth program is a "C" language program, **REFLECT.C**, to calculate the reflectance for an input index at specific wavelengths.

The next group of programs are used in the wavelet based optimal film design. They are a top level program, **WAVRUN.M**, the evaluation function **WAVFUN.M**, and the wavelet coefficient to function conversion function called **UP1D.M**. The

CONSTR.M and **REFLECT.C** programs are also used along with these programs. The program **DOWN1D.M**, which is the complement to **UP1D.M** is also included.

The next six programs are used to find the minimum thickness optimal film. The first three, **FFTMIN.M**, **FLATEVAL.M**, and **FPADEVAL.M**, are for the Fourier series version of the algorithm. The other three, **WAVMIN.M**, **NFLTEVAL.M**, and **NPADEVAL.M**, are used in the wavelet implementation of the algorithm.

```
function Q = que(li,lf,nchan,scale,dtot)
```

```
%     This program creates a Q function for a notch reflective filter over an input range
from
%     li to lf in 'nchan' channels.  The output is Q(f)/f, as required for the inverse Fourier
%     transform relation with n(x).  Inputs are:
%          li = the lower wavelength (microns)
%          lf = the upper wavelength (microns)
%          nchan = number of channels          (must be even number; mod(2) preferred)
%          scale = the desired reflectivity between 0 & 1.
%          dtot  = the total thickness of film to model
%                        (note: should be about 4 times the actual desired film thickness.)
%
% References:
%     Bovard, Appl. Opt. 29:24, 1990.
%     Dobrowolski and Lowe, Appl. Opt. 17:3039, 1978.
%     Druessel, Grantham, Haaland, Opt. Let. 18:1583, 1993.
```

```
% ('freq' and 'nchan' are the total freq range and number of channels)
% The Nyquist frequency is found by taking 1/ 2 * total thickness.  The total frequency
% range is twice the Nyquist frequency.  Note the total thickness input is an optical
% thickness in microns.  Since the theory relates the double optical thickness to the
% wavenumber, the total thickness must be doubled before it is used in calculations.
% freq is the total frequency range modeled. Delta is the size of a singe step in k.
freq=nchan/2/dtot;
delta=freq/nchan;
% The max spatial frequency is 1/ (min wavelength)
% convert endpoints to spatial frequencies.
```

```
fi=1/lf;                    % fi is initial frequency
ff=1/li;                    % ff is final frequency
f=zeros(nchan,1);
```

```
% Treat the left half plane as positive frequencies and the right half plane as negative
% frequencies so that the Nyquist frequency is 'freq/2' and it falls in
%   channel '(nchan/2)+1'. The DC component is in channel 1.
```

```
il=round(fi/delta)+1;
ih=round(ff/delta)+1;
f(il:ih)=ones(ih-il+1,1);
% must divide by frequency of the bin for Q(x)/x
```

```
for i=il:ih
   f(i)=f(i)/((i-1)*delta);
end
%  Scale Q using one of various techniques outlined in Bovard's paper.
%

Q(:)=f*sqrt(scale);

% the next two lines are one alternative scaling function
%Q(:)=0.5*f*sqrt(-log(1-scale));
%Q(:)=Q(:)+0.5*f*sqrt((1/sqrt(1-scale)-sqrt(1-scale)));
return;
```

```
function [nx,p] = swift(Q,nchan,li,lf,dtot,dtarget)
%
%  swift(Q,nchan,li,lf,dtot,dtarget)
%
%  This program computes the index of refraction n(x) and
%  phase function p for an input symmetric modified
%  transmission function Q.
%  Arguements:
%   Q    = modified transmission vector (input from que.m)
%   nchan = # of elements in vector 'Q' (i.e., # of channels)
%   li   = lowest non-zero wavelength in Q
%   lf   = highest non-zero wavelength in Q
%   dtot = total thickness of film modeled
%   dtarget=swift optimization parameter
%
%  Output :
%   nx = n(x) with average n = nsub  (coded in program)
%    p = optimal phase function used
%
%  From Guan and McIver, J.Chem.Phys., 92:5841, 1990.
%  Bertrand, Appl. Opt., 27:1998, 1988.
%  Druessel, Grantham, and Haaland, Opt. Letters,18:1583, 1993.


%  the value nsub is the index of the substrate.
nsub=1.50;

%  'freq' is the twice the Nyquist frequency.
%  It's determined by nchan/(2*thickness)
freq=nchan/2/dtot;

% define spatial freq endpoint.  note low=1/high
fl=1/lf;
fh=1/li;

%  'nchan' is the no. of elements in vector 'Q'
%  For the organization (frequency tagging) of vector 'Q',
%   see notes in function QUE.M.
%
%  'delta' is the frequency (1/microns) per bin
```

```
delta=freq/nchan;

%   Calculate the x-axis in units of 1/microns

xf=(1:nchan/2+1)*delta -delta;

%  'ichanl' and 'ichanh' track the lowest and highest non-zero
%    bins in the frequency domain.

ichanl=round(fl/delta)+1;
ichanh=round(fh/delta)+1;
if (ichanh>nchan),ichanh=nchan-1;,end;

t=dtarget;
%  Must double the desired value for calculations, since the FT relation is in terms of
%  twice the optical thickness.  This thickness contains most of the index change
%, but the tails are not included.


%  t0 is the starting point for desired actual thickness.
%  note that the desired thickness is a FWHM, not 0 to 0.


t=2*t;
t0=(1/delta-t)/2


%---perform filtering on data------
%  The Fourier transform of a function with sharp turn on has significant high frequency
%  components.  Since these would be truncated in practice anyway, the input function
%  is filtered to smooth the sharp turn on features a little.
%   specify the number of filters
m=1;
%   specify the filtering bandwidth (Hz)
deltaw=0.5*delta;
if deltaw > freq
  m=0;
end
ff=zeros(nchan,1);
k=round(0.5*deltaw/delta)
if k==0
  m=0;
end
if m > 0
  for j=1:m
   il1=ichanl-j*k-1;
```

```
    if il1 <= 1
    il1=2;        % insures dc component still 0
    end
  ih1=ichanh+j*k+1;
    if ih1 > nchan/2
    ih1=nchan/2;
    end
%   Test to ensure filtering doesn't exceed array dimensions
  for i=il1:ih1
  ik1=i-k;
  ik2=i+k;
    if ik1 <=1
      ik1=2;
    end
    if ik2 > nchan/2
      ik2=nchan/2;
    end
  ff(i)=(sum(Q(ik1:ik2))-0.5*(Q(ik1)+Q(ik2)))/(2*k);
  end
  Q=ff;
 end
end
Q=Q';




%---begin phase function calculation----

j=sqrt(-1);
g=(abs(Q)).^2;
gx=zeros(nchan,1);
gsum=0;
gxsum=0;
ichanl
ichanh
denom=delta*(sum(g(1:ichanh+m*k+1))-0.5*(g(ichanl-m*k-1)+g(ichanh+m*k+1)));
for i=1:nchan
  gsum=gsum+g(i);
  gx(i)=delta*(gsum-0.5*(g(1)+g(i)));
  gxsum=gxsum+gx(i);
  gy=delta*(gxsum-0.5*(gx(1)+gx(i)));
  shift=2*pi*i*delta*t0;

%   Assume the initial phase (P0) is such that p(0)=0.
```

141

```
%
   p(i)=shift+2*pi*gy*t/denom;
end

%  Perform inverse fourier transform using matlab built in function.
omeg=Q .*exp(p*j);
ftemp=ifft(omeg);

%  calculate the optical thickness step.
%  note in the theory the index is a function of twice
%  the optical thickness, so must divide by 2.
%
del=1/(2*freq);
%  Perform the normalization.
ftemp=ftemp/del/pi;
ftemp=imag(ftemp);

xt=(1:nchan)*del-(nchan*0.5)*del;
ft(:)=ftemp';

%  nx is a nchan x 3 matrix.  the first column is the index
%  the second column is the thickness of that layer.
%  the third column is the position of the layer (for plotting).

nx(:,1)=nsub*exp(real(ft));
nx(:,2)=del ./nx(:,1);
nx(:,3)=xt';

return;
```

```
function R = reflect(n,M)
% reflect(n,M) calculates the Reflectivity of an  index profile 'n' in nchan sections.
%     n   = input index matrix, with index in column 1 and thickness of layer in column
2.
%            (Assumes first row is on substrate)
%     M   = number of points for output plot
%    cutoff  = min percentage of peak index to include
%
nchan=size(n)
nchan=nchan(1);
xt=[1:nchan]*n(2,1)*n(2,2);
%
incident=1;
NGaAs=3.5986;
NAlAs=2.9742;
Nsub=1.52;
Nair=1.;
% assume incidence from substrate
N0=Nsub;
Nout=Nair;
nx=n(:,1);
d =n(:,2);
%  if incident from air, reverse order of indicies.
%
if incident == 1
  N0=Nair;
  Nout=Nair;
  nx=fliplr(nx')';
  d=fliplr(d')';
end
eta0=N0;
%  calculate phase delay delta and admittance for each layer.
%  note this phase delay does not yet include the division by
%  lambda, since lambda must be a variable to plot Reflectivity.
%  delta is a constant here.  this assumes the input was generated
%  by the IFFT method using NOFAZ or SWIFTQ.
delta=2*pi*nx(1)*d(1);
eta = nx;
%
etaout=Nout;
%
```

```
%  now form the vector of [B C]  used to find Reflectivity.
%  Also set the values of lambda to be ploted.  The constant
%  M is the number of points to calculate.
lambda=zeros(1,M);
film=zeros(2,2);
J=sqrt(-1);
%  initialize the BC vector at the output end.
for i=1:M
  lambda(i)=0.5+i*1/M;
  BC=[1,etaout]';
  film(1,1) = cos(delta/lambda(i));
  film(2,2) = cos(delta/lambda(i));
  sini=J*(sin(delta/lambda(i)));
  for j=1:nchan
    film(1,2) = sini/eta(nchan-j+1);
    film(2,1) = sini*eta(nchan-j+1);
    BC = film*BC;
  end
  Y=BC(2)/BC(1);
  R(i) = (abs((eta0 - Y)/(eta0+Y)))^2;
end
%
%  now plot the output, Reflectivity vs wavelength.
plot(lambda,R)
title('Reflectivity vs Wavelength')
xlabel('Wavelength in microns')
ylabel('Reflectivity')
grid
R=R';
R(:,1)=R;
R(:,2)=lambda';
return;
```

FFTRUN.M

```
%  Run for AR using fft.  This is a script file.  It is assumed the variables below
%  are defined in the MATLAB workspace.
tic
%
[out,options]=constr('fftfun',input,optin,[],[],[],N,Rin,Nsub,din,k,x);
I=sqrt(-1);
fftN=zeros(N,1);
fftN(1:Nd)=out(1:Nd)+I*out(Nd+1:2*Nd);
fftN(N-Nd+2:N)=conj(flipud(fftN(2:Nd)));
indexout=real(ifft(fftN));  % should be real by construction, just make sure.

rplot=reflectn(indexout,Nsub,kplot,x');
time=toc
eval(['save ',file])
```

```
function [f,g]=fftfun(d,N,R,Nsub,din,k,x)

% [f,g]=fftfun(d,R,Nsub,din,k,h) is a function to minimize in thin film optimization
% This version uses sins as the building blocks instead of wavelets.
% The first arguement is the vector of variables.  R contains
% the desired index profile, din is the thickness of the film, k is the vector
% of wave numbers specified, and h is the wavelet filter to use.
%
%        d = vector of variable detail coefficients
%        R = desired Reflectance vs k
%      Nsub= index of substrate
%      din = total thickness of film
%        k = wavenumber vector
%        x = position vector

I=sqrt(-1);
Nd=length(d)/2;
fftN=zeros(N,1);
fftN(1:Nd)=d(1:Nd)+I*d(Nd+1:2*Nd);
fftN(N-Nd+2:N)=conj(flipud(fftN(2:Nd)));
Nx=real(ifft(fftN));  % should be real by construction, just make sure.

Rcalc=reflectn(Nx,Nsub,k,x');

% f is the function to minimize, sum of squared error in reflectance

f=sqrt(sum((R-Rcalc).^2));

% the g's are the constraints

Nmin=1.38;
Nmax=2.5;
g(1:N)=Nmin-Nx;
g(N+1:2*N)=Nx-Nmax;
return
```

```
function[x, OPTIONS,lambda, HESS]= constr(FUN,x,OPTIONS,VLB,VUB
          GRADFUN,P1,P2,P3,P4,P5,P6,P7,P8,P9,P10,P11,P12,P13,P14,P15)
```

```
%CONSTR    Finds the constrained minimum of a function of several variables.
%
%      X=CONSTR('FUN',X0) starts at X0 and finds a constrained minimum to
%      the function which is described in FUN (usually an M-file: FUN.M).
%      The function 'FUN' should return two arguments: a scalar value of the
%      function to be minimized, F, and a matrix of constraints, G:
%      [F,G]=FUN(X). F is minimized such that G < zeros(G).
%
%      X=CONSTR('FUN',X,OPTIONS) allows a vector of optional parameters to
%      be defined. For more information type HELP FOPTIONS.
%
%      X=CONSTR('FUN',X,OPTIONS,VLB,VUB) defines a set of lower and upper
%      bounds on the design variables, X, so that the solution is always in
%      the range VLB < X < VUB.
%
%      X=CONSTR('FUN',X,OPTIONS,VLB,VUB,'GRADFUN') allows a function
%      'GRADFUN' to be entered which returns the partial derivatives of the
%      function and the  constraints at X:  [gf,GC] = GRADFUN(X).
%
%      X=CONSTR('FUN',X,OPTIONS,VLB,VUB,GRADFUN,P1,P2,..) allows
%      coefficients, P1, P2, ... to be passed directly to FUN:
%      [F,G]=FUN(X,P1,P2,...). Empty arguments ([]) are ignored.


%      Copyright (c) 1990 by the MathWorks, Inc.
%      Andy Grace 7-9-90.
```

OPTIONS for CONSTR.M

```
function OPTIONS=foptions(parain);
%FOPTIONS Default parameters used by the optimization routines.
%       In MATLAB itself:
%           FMIN and FMINS.
%       In the Optimization Toolbox:
%           FMINU, CONSTR, ATTGOAL, MINIMAX, LEASTSQ, FSOLVE.
%       The parameters are:
%       OPTIONS(1)-Display parameter (Default:0). 1 displays some results
%       OPTIONS(2)-Termination tolerance for X.(Default: 1e-4).
%       OPTIONS(3)-Termination tolerance on F.(Default: 1e-4).
%       OPTIONS(4)-Termination criterion on constraint violation.(Default: 1e-6)
%       OPTIONS(5)-Algorithm: Strategy:  Not always used.
%       OPTIONS(6)-Algorithm: Optimizer: Not always used.
%       OPTIONS(7)-Algorithm: Line Search Algorithm. (Default 0)
%       OPTIONS(8)-Function value. (Lambda in goal attainment. )
%       OPTIONS(9)-Set to 1 if you want to check user-supplied gradients
%       OPTIONS(10)-Number of Function and Constraint Evaluations.
%       OPTIONS(11)-Number of Function Gradient Evaluations.
%       OPTIONS(12)-Number of Constraint Evaluations
%       OPTIONS(13)-Number of equality constraints.
%       OPTIONS(14)-Maximum number of iterations. (Default 100*no. of variables)
%       OPTIONS(15)-Used in goal attainment for special objectives.
%       OPTIONS(16)-Minimum change in variables for finite difference gradients.
%       OPTIONS(17)-Maximum change in variables for finite difference gradients.
%       OPTIONS(18)- Step length. (Default 1 or less).

%       Andy Grace 7-9-90.
%       Copyright (c) 1984-94 by The MathWorks, Inc.

if nargin<1; parain = []; end
sizep=length(parain);
OPTIONS=zeros(1,18);
OPTIONS(1:sizep)=parain(1:sizep);
default_options=[0,1e-4,1e-4,1e-6,0,0,0,0,0,0,0,0,0,0,0,0,1e-8,0.1,0];
OPTIONS=OPTIONS+(OPTIONS==0).*default_options;
```

```
/*

    reflectx.C    .MEX file corresponding to reflectn.M
                  Calculates reflectance for input index profile

    The calling syntax is:

                [R] = reflectx(index,Nsub,k,x)


    Jeffrey J. Druessel  Aug 19, 1995

*/                                                          */
/*-------------------------------------------------------------*/
/*  function R = reflect(index,Nsub,k,x)                  */
/*  reflect(index,Nsub,k,x) calculates the Reflectivity of a   */
/*  index profile  in nchan sections.                     */
/*      index = input index matrix, with index in column 1    */
/*              and thickness of layer in column 2              */
/*      Nsub  = index of substrate                        */
/*        k   = wavenumber of points to plot              */
/*        x   = position of index values (units must match k)  */
/*                                                        */
/*  ref:  MacLeod's book: Thin Film Optical Filters, Ch 2,   */
/*    New York:MacMillian , 1986.                         */
/*-------------------------------------------------------------*/


#include <stdio.h>
#include <math.h>
#include "mex.h"          /*  Matlab provided header file for running with matlab */
#include "nrutil.h"       /*  Matlab provided header file for running with matlab */
#include "nrutil.c"       /*  Matlab provided functions for running with matlab */

/* Input Arguments */

#define     INDEX_IN    prhs[0]
#define     NSUB_IN     prhs[1]
#define     K_IN        prhs[2]
#define     X_IN        prhs[3]
```

```c
/* Output Arguments */

#define       R_OUT        plhs[0]

#define max(A,B)     ((A) > (B) ? (A) : (B))
#define min(A,B)     ((A) < (B) ? (A) : (B))

#define pi 3.14159265


static
#ifdef __STDC__
int reflect(double n[],double x[],double R[],double k[],
            int nchan,double Nsub,int M)
#else
reflect(n,x,R,k,nchan,Nsub,M)
double n[],x[],R[],k[];
int nchan;
double Nsub;
int M;
#endif

{
static double Nair=1.00;
double N0,Nout,eta0,etaout;
double a,b,c,d,RB,IB,RC,IC,RB0,IB0,RC0,IC0,RY,IY;
double *nx,*d1;
double *delta,*eta;
double *lambda;
int i,j;

/* allocate vectors using dvector above */

nx=dvector(1,nchan);
d1=dvector(1,nchan);
delta=dvector(1,nchan);
eta=dvector(1,nchan);
lambda=dvector(1,M);

/* assume light is incident from air, film vector starts at substrate */
  N0=Nair;
  Nout=Nsub;
  nx[nchan]=n[0];
```

```
d1[nchan]=x[0];
for (i=1; i<nchan; i++)
{
  nx[i]=n[nchan-i];
  d1[i]=x[nchan-i] - x[nchan-i-1];
}


/*  calculate phase delay delta and admittance eta for each layer.   */
/*  note this phase delay does not yet include the division by            */
/*  lambda, since lambda must be a variable to plot Reflectivity.   */
/*  The parameter eta0 is the admittance of the incident media.         */
/*                                                                                         */
/*                                                                                         */

eta0=N0;

for (i=1; i<=nchan; i++)
{
  delta[i]=2.0 * pi * nx[i] * d1[i];
  eta[i] = nx[i];
}
etaout=Nout;


/*  now form the vector of [B C]  used to find Reflectivity.  */
/*  [B C] are both complex, so carry RB for real and IB for imag*/
/*  Also set the values of lambda to be ploted. The values     */
/*  below are for wavelengths between 500 and 1500 nm. The     */
/*  constant M is the number of points to calculate.           */

for (i=1;i<=M;i++) {
  lambda[i]=2.0 * pi / k[i-1];
  RB0=1.0;
  IB0=0.0;
  RC0=etaout;
  IC0=0.0;
  for (j=1;j<=nchan;j++) {
        a=cos(delta[nchan-j+1]/lambda[i]);
        b=(sin(delta[nchan-j+1]/lambda[i]))/eta[nchan-j+1];
        c=(sin(delta[nchan-j+1]/lambda[i]))*eta[nchan-j+1];
        d=a;

    RB = a*RB0 - b*IC0;
```

```c
        IB = a*IB0 + b*RC0;
        RC = d*RC0 - c*IB0;
        IC = d*IC0 + c*RB0;
                RB0=RB;
                IB0=IB;
                RC0=RC;
                IC0=IC;
        }
        RY=(RC*RB + IC*IB)/(RB*RB+IB*IB);
        IY=(RB*IC - RC*IB)/(RB*RB+IB*IB);

        R[i-1] =((eta0-RY)*(eta0-RY)+IY*IY)/((eta0+RY)*(eta0+RY) + IY*IY);
}
free_dvector(nx);
free_dvector(d1);
free_dvector(delta);
free_dvector(eta);
free_dvector(lambda);
return;
}


#ifdef __STDC__
void mexFunction(
        int             nlhs,
        Matrix *plhs[],
        int             nrhs,
        Matrix *prhs[]
        )
#else
mexFunction(nlhs,plhs,nrhs,prhs)
int nlhs,nrhs;
Matrix *plhs[], *prhs[];
#endif
{
        double *index,*R;
        double *x,*k,*Nsub;
        unsigned int    m,n;
        int nchan,M;
        /* Check for proper number of arguments */

        if (nrhs != 4) {
                mexErrMsgTxt("REFLECTX requires four input arguments.");
        } else if (nlhs > 1) {
                mexErrMsgTxt("REFLECTX requires one output argument.");
```

```
        }
        /* Check the dimensions of index, x.  Determine number of k points */

        m = mxGetM(INDEX_IN);
        n = mxGetN(INDEX_IN);
        nchan=max(m,n);
        if (!mxIsNumeric(INDEX_IN) || mxIsComplex(INDEX_IN) ||
                !mxIsFull(INDEX_IN) || !mxIsDouble(INDEX_IN) ||
                (min(m,n) != 1)) {
                mexErrMsgTxt("REFLECTX requires that INDEX be a vector.");
        }
        m = mxGetM(X_IN);
        n = mxGetN(X_IN);
        if (!mxIsNumeric(X_IN) || mxIsComplex(X_IN) ||
                !mxIsFull(X_IN) || !mxIsDouble(X_IN) ||
                (max(m,n)!=nchan) || (min(m,n) != 1)) {
                mexErrMsgTxt("REFLECTX requires that X be a vector with same lenght
as INDEX.");
        }

        m = mxGetM(K_IN);
        n = mxGetN(K_IN);
        M = max(m,n);
        if (!mxIsNumeric(K_IN) || mxIsComplex(K_IN) ||
                !mxIsFull(K_IN) || !mxIsDouble(K_IN) ||
                (min(m,n) != 1)) {
                mexErrMsgTxt("REFLECTX requires that K be a vector.");
        }
        /* Create a matrix for the return argument */
        R_OUT = mxCreateFull(M, 1, REAL);

        /* Assign pointers to the various parameters */
        R = mxGetPr(R_OUT);
        index = mxGetPr(INDEX_IN);
        Nsub = mxGetPr(NSUB_IN);
        k = mxGetPr(K_IN);
        x = mxGetPr(X_IN);

        /* Do the actual computations in a subroutine. Nsub is a double in the */
        /* subroutine, so dereference before passing the value.              */

        reflect(index,x,R,k,nchan,*Nsub,M);
        return;
}
```

```
%  Run for AR using  wavelet with filter h
tic

input=nguess(Nbound,Nsub,din,N,k,h);
VLB=-10*ones(Nvar,1);
VUB=10*ones(Nvar,1);
VLB(Nvar)=0;
VUB(Nvar)=100;

[out,options]=constr('wavfun',input(N-Nvar+1:N),optin,VLB,VUB,[],
                     input(1:N-Nvar),Rin,Nsub,din,k,x,h);

indexout=d_to_n(input(1:N-Nvar),out,h);
rplot=reflectn(indexout,Nsub,kplot,x');
time=toc
eval(['save ',file])
```

WAVFUN.M

```
function [f,g]=wavfun(d,P1,R,Nsub,din,k,x,h)
% [f,g]=wavfun(d,P1,R,Nsub,din,k,h) is a function to minimize in thin film optimization
% The first arguement is the vector of variables.  P1 on are parameter
% vectors.  P1 contains the first level detail coefficeints, and R contains
% the desired index profile, din is the thickness of the film, k is the vector
% of wave numbers specified, and h is the wavelet filter to use.
%
%       d = vector of variable detail coefficients
%       P1 = vector of fixed detail coefficients
%       R = desired Reflectance vs k
%     Nsub= index of substrate
%     din = total thickness of film
%       k = wavenumber vector
%       x = position vector
%       h = wavelet filter


Nd=length(d);
NP1=length(P1);
N=Nd+NP1;
decomp(1:NP1)=P1;
decomp(NP1+1:N)=d;
Nx=up1d(decomp',h);


Rcalc=reflectn(Nx,Nsub,k,x');


% f is the function to minimize, sum of squared error in reflectance

f=sqrt(sum((R-Rcalc).^2));


% the g's are the constraints

Nmin=1.38;
Nmax=2.5;


g(1)=Nx(1)-Nsub;
g(2:N+1)=Nmin-Nx;
g(N+2:2*N+1)=Nx-Nmax;


return
```

```
function out=up1d(A,h);
%  UP1D builds the rebuilds the sample value matrix from the matrix
%  of wavelet decomposition coefficients  A  with filter h.
%  This is a reconstruction for a 1D decomposition
%  of coulumns of a matrix.
%
%     A  =   matrix of wavelet coefficients. Should be power of 2.
%     h  =   column vector,wavelet filter to use.
%
%  First check that h is a column vector.
%
temp=size(h);
if temp(2)>temp(1)
  h=h';
end
L=length(h);
[N,P]=size(A);
M=log2(N);


% The matlab function spdiags(B,d,N,N) creates an N x N
% sparse matrix with the elements of vector B on the diagonal d.
% For more than one non-zero diagonal, B is a matrix and d is a vector
% of which diagonals the columns of B go on.


% for the wavelet filters, there are 2*L non-zero diagonals,
% where L is the length of the filter h for the wavelet.


% build filter matrix for h.
B=ones(N,1)*[h',h'];
ddd=[0:1:L-1,-N:L-N-1];
H=spdiags(B,ddd,N,N);



% build filter matrix for g.
g=flipud(h);
g(2:2:L)=-g(2:2:L);


B=ones(N,1)*[g',g'];
ddd=[0:1:L-1,-N:L-N-1];
G=spdiags(B,ddd,N,N);
```

```
% use transpose of G and H for reconstruction
% Input is assumed to be arranged  | D1     D1  |
%                                   | D2     D2  |
%                                   | D3     D3  |
%                                   | C      C   |

C=A;
N2=1;
end
pointN=N-2*N2+1;
for i=M:-1:1
   Din=A(pointN:pointN+N2-1,:);
   Cin=C(pointN+N2:pointN+2*N2-1,:);
   N2=N2*2;

   D=zeros(N2,P);
   C=zeros(N2,P);
   C(1:2:N2,:)=Cin;
   C(2:2:N2,:)=zeros(N2/2,P);
   D(1:2:N2,:)=Din;

   for j=1:log2(N/N2)
     C=[C;C];
     D=[D;D];
   end
   C=G'*D+H'*C;
   pointN=pointN-N2;

end
out=C;
return
```

DOWN1D.M

```matlab
function out=down1d(A,h);
%  DOWN1D builds the matrix of wavelet decomposition coefficients
%  for an input matrix A  with filter h.  This is a 1D decomposition
%  of each colunm in the input matrix.
%
%     A  =   matrix of sampled data. Should be power of 2.
%     h  =   column vector,wavelet filter to use.
%


%  First check that h is a column vector.
%
temp=size(h);
if temp(2)>temp(1)
   h=h';
end
L=length(h);
[N,P]=size(A);
M=log2(N);


% The matlab function spdiags(B,d,N,N) creates an N x N
% sparse matrix with the elements of vector B on the diagonal d.
% For more than one non-zero diagonal, B is a matrix and d is a vector
% of which diagonals the columns of B go on.


% for the wavelet filters, there are 2*L non-zero diagonals,
% where L is the length of the filter h for the wavelet.


% build filter matrix for h.
B=ones(N,1)*[h',h'];
ddd=[0:1:L-1,-N:L-N-1];
H=spdiags(B,ddd,N,N);


% build filter matrix for g.
g=flipud(h);
g(2:2:L)=-g(2:2:L);


B=ones(N,1)*[g',g'];
ddd=[0:1:L-1,-N:L-N-1];
G=spdiags(B,ddd,N,N);
```

```
% This decomposition operates on  columns.
% the D matrices are the interactions at that level,
%  and the C matrix is the input for the next level of decomposition.
% Output is arranged  | D1     D1  |
%                     | D2     D2  |
%                     | D3     D3  |
%                     | C      C   |

out=zeros(N,P);
Ain=A;
pointN=1;
N2=N;
for i=1:M
   N2=N2/2;
   D=G*Ain; D=D(1:2:N,:);
   C=H*Ain;C=C(1:2:N,:);
   Ain=[C;C];

   out(pointN:pointN+N2-1,:)=D(1:N2,:);
   pointN=pointN+N2;
end
   out(N,:)=C(1,:);
return
```

FFTMIN.M

```
% Optimization script for finding the minimum film thickness
% for a given reflection tolorance using fourier series method.
% The matlab workspace must have a valid design loaded.
% The design should be too thin to produce the desired performance.
% The first part of the script calculates the value of the error as
% a function of film thickness. The starting thickness, stop thickness,
% and number of point is hard coded in below. A curve is then fit to
% these points, and the min thickness is determined from the curve.
%
% This script uses a simple bisection method to find the zero intercept
% of error-Rtol.
% the variables used are:
%      data is an array of each point and its function value
tic
load yagfft3
file='fminpad';
Nvar=32;
optin(14)=10000;
start=300;
stop=900;
points=10;
data=zeros(points+1,2);
dx=(stop-start)/points;


% Find index for starting thickness

din=start;
x=linspace(0,din,N+1);,x(1)=[];
input=zeros(2*Nd,1);
input(1)=Nsub*N/2;
[out,options]=constr('fftfun',input,optin,[],[],[],N,Rin,Nsub,din,k,x);
I=sqrt(-1);
fftN=zeros(N,1);
fftN(1:Nd)=out(1:Nd)+I*out(Nd+1:2*Nd);
fftN(N-Nd+2:N)=conj(flipud(fftN(2:Nd)));
indexout=real(ifft(fftN));  % should be real by construction, just make sure.
data(1,:)=[din,options(8)];

oldindex(:,1)=indexout(:,1); olddin=din;
```

```
for i=1:points
  newdin = start + i*dx;
  flateval;
  data(i+1,:)=[newdin,options(8)];
  if ( data(i,2)<data(i+1,2))
     indexin=oldindex(:,i);
     fpadeval;
     data(i+1,:)=[newdin,options(8)];
  end
  oldindex(:,i+1)=indexout(:,1);   olddin=newdin;
  eval(['save ',file])
end

time=toc;
```

FLATEVAL.M

`% Script to evaluate a point in frunmin with constant seed`

```
x=linspace(0,newdin,N+1);x(1)=[];
[out,options]=constr('fftfun',input,optin,[],[],[],N,Rin,Nsub,newdin,k,x);
I=sqrt(-1);
fftN=zeros(N,1);
fftN(1:Nd)=out(1:Nd)+I*out(Nd+1:2*Nd);
fftN(N-Nd+2:N)=conj(flipud(fftN(2:Nd)));
indexout=real(ifft(fftN));  % should be real by construction, just make sure.
```

FPADEVAL.M

% Script to evaluate a point in frunmin with substrate padding

```
x=linspace(0,newdin,N+1);x(1)=[];
pad=((newdin/olddin-1)*N);
padN=ceil(pad);
newx=linspace((pad-padN)*olddin/N,newdin,N+padN+1);
newindex=[Nsub*ones(padN+1,1);indexin];
index=interp1(newx,newindex,x,'spline');

fftin=fft(index);
input(1:Nd)=real(fftin(1:Nd));
input(Nd+1:2*Nd)=imag(fftin(1:Nd));

[out,options]=constr('fftfun',input,optin,[],[],[],N,Rin,Nsub,newdin,k,x);
I=sqrt(-1);
fftN=zeros(N,1);
fftN(1:Nd)=out(1:Nd)+I*out(Nd+1:2*Nd);
fftN(N-Nd+2:N)=conj(flipud(fftN(2:Nd)));
indexout=real(ifft(fftN));  % should be real by construction, just make sure.
```

```
% Optimization script for finding the minimum film thickness
% for a given reflection tolorance.
% The matlab workspace must have a valid design loaded.
% The design should be too thin to produce the desired performance.
% The first part of the script calculates the value of the error as
% a function of film thickness. The starting thickness, stop thickness,
% and number of point is hard coded in below. A curve is then fit to
% these points, and the min thickness is determined from the curve.
%
% This script uses a simple bisection method to find the zero intercept
% of error-Rtol.
% the variables used are:
%       data is an array of each point and its function value
tic
load nmin350
file='nminflta';
Nvar=32;
optin(14)=10000;
start=540;
stop=900;
points=6;
data=zeros(points+1,2);
dx=(stop-start)/points;
logrtol=log10(Rtol);


% Find index for starting thickness

din=start;
x=linspace(0,din,N+1);,x(1)=[];
indexin=Nsub*ones(N,1);
input=down1d(indexin,h);
[out,options]=constr('wavfun',input(N-Nvar+1:N),optin,VLB,VUB,[],
                        input(1:N-Nvar),Rin,Nsub,din,k,x,h);
indexout=d_to_n(input(1:N-Nvar),out,h);
data(1,:)=[din,options(8)];


oldindex(:,1)=indexout(:,1); olddin=din;

for i=1:points
  newdin = start + i*dx;
  nfltevala;
  data(i+1,:)=[newdin,options(8)];
```

```
  if (data(i,2)<data(i+1,2))
    indexin=oldindex;
    npadeval;
    data(i+1,:)=[din,options(8)];
  end
  oldindex(:,i+1)=indexout(:,1);   olddin=newdin;
  eval(['save ',file])
end

time=toc;
```

NFLTEVAL.M

```
%  Script to evaluate a point in wavmin with constant seed

x=linspace(0,newdin,N+1);x(1)=[];
[out,options]=constr('minfun',input(N-Nvar+1:N),optin,VLB,VUB,[],
                     input(1:N-Nvar),Rin,Nsub,newdin,k,x,h);
indexout=d_to_n(input(1:N-Nvar),out,h);
```

NPADEVAL.M

```
%  Script to evaluate a point in nrunmin with substrate padding

x=linspace(0,newdin,N+1);x(1)=[];                    \
pad=((newdin/olddin-1)*N);
padN=ceil(pad);
newx=linspace((pad-padN)*olddin/N,newdin,N+padN+1);
newindex=[Nsub*ones(padN+1,1);oldindex];
index=interp1(newx,newindex,x,'spline');
input=down1d(index,h);
[out,options]=constr('minfun',input(-Nvar+1:N),optin,VLB,VUB,[],
                     input(1:NNvar),Rin,Nsub,newdin,k,x,h);
indexout=d_to_n(input(1:N-Nvar),out,h);
```

Bibliography

1   Adolph, J. B. *et al.* "Properties of multilayer and graded index Si-based coatings deposited in dual-frequency plasma," in *Inhomogeneous and Quasi-Inhomogeneous Optical Coatings*, J. A. Dobrowolski, P. G. Verly, Editors, *Proc. SPIE,* **2046**:179-188 (1993).

2   Arfken, G. *Mathematical Methods for Physicists* (Third Edition). Orlando: Academic Press, 1985.

3   Berning, P. H. "Use of equivalent films in the design of infrared multilayer antireflection coatings," *Journal of the Optical Society of America,* **52** (4): 431-436 (April 1962).

4   Bovard, B. "Derivation of a matrix describing a rugate dielectric thin film," *Applied Optics,* **27** (10):1998-2005 (May 1988).

5   Bovard, B. "Rugate filter design: the Modified Fourier transform technique," *Applied Optics,* **29** (1):24-30 (January 1990).

6   Bovard, B. "Rugate filter theory: an overview," *Applied Optics,* **32** (28):5427-5442 (October 1993).

7   Daubechies, I. "Orthonormal Basis of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics,* **XLI**:909-996 (1988).

8   Daubechies, I. *Ten Lectures on Wavelets.* Philadelphia: Society for Industrial and Applied Mathematics, 1992.

9   Delano, E. "Fourier synthesis of multilayer filters," *Journal of the Optical Society of America,* **57** (12):1529-1533 (December 1967).

10   Dobrowolski, J. A. and D. Lowe. "Optical thin film synthesis program based on the use of the Fourier transforms," *Applied Optics,* **17** (19):3039-3050 (October 1978).

11   Dobrowolski, J. A. and R. A. Kemp. "Refinement of optical multilayer systems with different optimization procedures," *Applied Optics,* **29** (19):2876-2893 (July 1990).

12   Dobrowolski, J. A. "Computer design of optical coatings," *Thin Solid Films,* **163**:97-110 (1988).

13   Dobrowolski, J. A., F. C. Ho, A. Belkind and V. A. Koss. "Merit functions for more effective thin film calculations," *Applied Optics,* **28** (14):2824-2831 (July 1989).

14   Druessel, J. J., J. Grantham and P. Haaland. "Optimal phase modulation for gradient index optical filters," Optics Letters, **18** (19):1583-5 (October 1993).

15   Epstein, L. I. "The design of optical filters," *Journal of the Optical Society of America,* **41** (11): 806-810 (November 1952).

16   Fabricius, H. "Gradient-index filters: designing filters with steep skirts, high reflection, and quintic matching layers," *Applied Optics,* **31** (25):5191-6 (September 1992).

17   Guan, S. "General phase modulation method for stored waveform inverse Fourier transform ion cyclotron resonance mass spectrometry," *Journal of Chemical Physics,* **91**(2):775-7 (July 1989).

18   Guan, S. and R. McIver, Jr. "Optimal phase modulation in stored waveform inverse Fourier transform excitation for Fourier transform mass spectrometry. I., Basic algorithm," *Journal of Chemical Physics,* **92** (10):5841-6 (May 1990).

19   Hecht, E. and A. Zajac. *Optics.* Reading, MA: Addison-Wesley Publishing Co., 1979.

20   Ho, F. C., C. Lee and R Tsai. "Characterization of inhomogeneous films formed by co-evaporation of ZnSe and MgF_2," in *Inhomogeneous and Quasi-Inhomogeneous Optical Coatings*, J. A. Dobrowolski, P. G. Verly, Editors, *Proc. SPIE,* **2046**:197-204 (1993).

21   Jacobsson, R. in *Physics of Thin Films*, G. Haas, M. Francombe, and R. Hoffman, eds., New York : Academic Press, pp. 51-97, 1975.

22   Johnson, W. E. and R. L. Crane. "Introduction to rugate filter technology," in *Inhomogeneous and Quasi-Inhomogeneous Optical Coatings*, J. A. Dobrowolski, P. G. Verly, Editors, *Proc. SPIE,* **2046**:88-108 (1993).

23   Li, L., and J. A. Dobrowolski. "Computation speeds of different optical thin-film synthesis methods," *Applied Optics,* **31**(19):3790-3799 (July 1992).

24   MacLeod, H. A. *Thin Film Optical Filters* (Second Edition). New York: MacMillian Publishing Company, 1986.

25   Mallat, S. G. "A theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **11** (7):674-93 (July 1989).

26   Matlab Optimization Toolbox Users Guide, The Math Works, Inc., Natick, MA, 1992.

27   Matlab version 4.0, The Math Works, Inc., Natick, MA, 1992.

28   Papalambros, P. Y. and D. J. Wilde. *Principles of Optimal Design: Modeling and Computation.* Cambridge: Cambridge University Press, 1988.

29   Poitras, D. *et al.* "Characterization of inhomogeneous films formed by co-evaporation of ZnSe and MgF_2," in *Inhomogeneous and Quasi-Inhomogeneous Optical Coatings*, J. A. Dobrowolski, P. G. Verly, Editors, *Proc. SPIE,* **2046**:197-204 (1993).

30   Press, W. H., B. P. Flannery, S. A. Teukolsky and W. T. Vetterling. *Numerical Recipes.* Cambridge: Cambridge University Press, 1986.

31   Reihl, K. *Collisional detachment of negative ions using Fourier transform mass spectrometry.* Ph. D. dissertation. School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB Ohio, 1992.

32   Rigler, A. K. and R. J. Pegis. "Optimization methods in Optics," *The Computer in Optical Research: Methods and Applications*, ed. by B. R. Frieden, Springer-Verlag: Berlin, 1980.

33   Saleh, B. E. A. and M. C. Teich. *Fundamentals of Photonics.* New York: John Wiley and Sons, 1991.

34   Sossi, L. "On the synthesis of interference coatings," *Eesti NSV Tead. Akad. Toim. Fuus. Mat.* **26**:28-36 (1977), translation by J. A. Dobrowolski, Translation Services of the Canada Institute for Scientific and Technical Information, National Research Council of Canada, Ottawa, Ontario, Canada K1A 0R6.

35   Sossi, L. "On the theory of the reflection and transmission of light by a thin inhomogeneous dielectric film," *Eesti NSV Tead. Akad. Toim. Fuus. Mat.* **23**:229-237 (1974). (See Ref 34)

36   Sossi, L. and P. Kard. "On the theory of the reflection and transmission of light by a thin inhomogeneous dielectric film," *Eesti NSV Tead. Akad. Toim. Fuus. Mat.* **17**:41-48 (1968). (See Ref 34)

37   Southwell, W. H., R. L. Hall and W. J. Gunning. "Using wavelets to design gradient-index interference coatings", in *Inhomogeneous and Quasi-Inhomogeneous Optical Coatings*, J. A. Dobrowolski, P. G. Verly, Editors, *Proc. SPIE,* **2046**:46-59 (1993).

38   Stakgold, I. *Green's Functions and Boundary Value Problems.* New York: John Wiley and Sons, Inc., 1979.

39   Thelen, A. "Design of optical minus filters," *Journal of the Optical Society of America,* **61** (2):365-369 (February 1971).

40   Tikhonravov, A. V. and J. A. Dobrowolski. "Quasi-optimal synthesis for antireflection coatings: a new method," *Applied Optics* **32** (22):4265-75 (August 1993).

41   Tsai, R., C. Wei and F. C. Ho. "Optical and structural properties of composite films prepared by co-evaporation of TiO_2 and MgF_2," in *Inhomogeneous and Quasi-Inhomogeneous Optical Coatings*, J. A. Dobrowolski, P. G. Verly, Editors, *Proc. SPIE,* **2046**:189-196 (1993).

42   Verly, P. G. and J. A. Dobrowolski. "Iterative correction process for optical thin film synthesis with the Fourier transform method," *Applied Optics,* **29** (25):3672-84 (September 1990).

43   Verly, P. G. *et al.* "Synthesis of high refection filters with the Fourier transform method," *Applied Optics,* **28** (14):2864-75 (July 1989).

44   Verly, P. G., J. A. Dobrowolski and R. R. Wiley. "Fourier transform method for the design of wideband antireflection coatings," *Applied Optics,* **31** (19): 3836-3846 (July 1992).

45   Warhola, G. T *et al.* Application of Wavelets to Signal Processing Short Course Notes, Air Force Institute of Technology, Wright-Patterson AFB OH, 20-22 March, 1991.

46   Young, R. K. *Wavelet Theory and Applications.* Boston: Kluwer Academic Publishers, 1993.

47   Zukic, M. and K. H. Guenther. "Optical coatings with graded index layers for high power laser applications: design," in *Laser Optics for Intracavity and Extra-cavity Applications*, P. M. Fauchet, Editor, *Proc. SPIE,* **895**:271-277 (1988).